Elucidating the Structure and Bioactivity of a *Rhodococcus fascians* Non-Ribosomal Peptide

Jonathan James Ford

John Innes Centre

Department of Molecular Microbiology

A thesis submitted December 2022 to the University of East Anglia for the degree of Doctor of Philosophy.

Abstract

The rise of antimicrobial resistance has necessitated the discovery of novel antimicrobial compounds as a matter of urgency. Non-Ribosomal Peptides (NRPs) are a class of structurally diverse natural products used clinically as antibiotic and anticancer agents. Actinonin and the matlystatins are actinobacterial NRPs which inhibit metalloproteinases by the chelation of their catalytic metal ion. This bioactivity is conferred by the N-hydroxyl-2-pentyl succinamic acid group, referred to as the 'warhead' group. Actinonin is the most potent natural inhibitor of peptide deformylase (PDF, a key enzyme involved in bacterial protein synthesis) and has been the focus of much research into the development of therapeutic PDF inhibitors as antimicrobials. Therefore, there is interest in the identification of other warhead-containing natural products which may have potent bioactivities.

In this thesis, a genetic probe for warhead-containing NRP biosynthetic gene clusters (BGCs) is identified and a number of putative clusters are presented. One such BGC was present in the pathogenicity-associated megaplasmid of the plant pathogen *Rhodococcus fascians*. The product of the cluster was isolated from *R. fascians* and structural analysis indicated that it was the antimycobacterial agent lydiamycin A, originally identified from *Streptomyces lydicus.* Consideration of the BGC and extensive NMR analysis informed a structural revision of lydiamycin A. Lydiamycin A was confirmed to inhibit PDF *in vitro.* The LydA PDF is encoded within the lydiamycin BGC and has been experimentally determined to be resistant against lydiamycin A. This result implicates LydA as an important self-immunity determinant for *R. fascians*. Finally, *in planta* competition assays frame lydiamycin as an important ecological factor that may enhance the fitness of *R. fascians* during niche colonisation.

This study highlights the power of targeted genome mining for the identification of BGCs that may produce structurally related natural products. This enabled structural revision and bioactivity characterisation of lydiamycin A.

Acknowledgements

It has been a privilege to carry out my PhD at the John Innes Centre, thanks to all of the wonderful people who go out of their way to ensure that it is an inviting and productive place to work. I would like to thank every member of the Mol Micro department for making it such a friendly workplace and always being there to help. My sincere gratitude to Dr. Andy Truman for all the guidance, help and advice throughout the project and for the extensive support whilst I have been writing this thesis. I have really appreciated the freedom to explore different aspects of my project and gain varied experience with biology and chemistry. Thank you to everyone in the Truman lab (Natalia, Louis, Liam) for making every day a joy to come to work. In particular, I am hugely grateful to Javier for his training and guidance throughout my project and him always going above and beyond to help me progress. Thank you to everyone in office 219 (Cat, Alaster, Nathan) for those necessary moments of minimal productivity, with the occasional gem of scientific advice.

I would like to thank Dr. Martin Rejzek and Dr. Edward Hems for their unending help with all things chemistry and their steadfast resolve that there is no such thing as a 'stupid question'. Thank you also to my secondary supervisor Prof. Barrie Wilkinson for all the support and helpful conversations. I couldn't have finished this work without the support from Dr. Joe Sallmen for his Mycobacterial wizardry and unflappable optimism, and Dr. Guy Polturak for his crucial plant advice. Thanks also to collaborators of the project: Dr. Isolde Francis and Dr. Lena Bögeholz.

Thank you to those who have made Norwich such a fun city over the last few years (Tom, Alan, Lauren). Thank you to my family for always supporting me and keeping me sane: Dad, Mum, Ashley, Nicola and Nanna and Poppa. Finally, none of this would have been possible without (soon to be Dr.) Ocean – it has been quite the four years. Thank you.

This work was funded by the BBSRC Norwich Research Park Biosciences Doctoral Training Partnership.

Author's Declaration

The research described in this thesis was conducted at the John Innes Centre between October 2018 and December 2022. All data described here are original and were obtained by the author unless otherwise attributed in the text. No part of this thesis has previously been submitted for a degree at this or any other academic institution.

The results described in **Chapter 5** contributed to the following publication:

Alba Pacheco-Moreno, Francesca L Stefanato, Jonathan J Ford, Christine Trippel, Simon Uszkoreit, Laura Ferrafiat, Lucia Grenga, Ruth Dickens, Nathan Kelly, Alexander DH Kingdon, Liana Ambrosetti, Sergey A Nepogodiev, Kim C Findlay, Jitender Cheema, Martin Trick, Govind Chandra, Graham Tomalin, Jacob G Malone, Andrew W Truman (2021) 'Pan-genome analysis identifies intersecting roles for Pseudomonas specialized metabolites in potato pathogen inhibition' *eLife* 10:e71900.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Natural Products

Natural products (NPs) are specialised metabolites produced outside of primary metabolism. As such, NPs are not essential for growth under laboratory conditions but instead confer fitness advantages in the native environment of an organism [1, 2]. The most prolific microbial NP producers include fungi and bacteria, in particular the actinomycetes [3]. NPs are incredibly chemically diverse and serve a wide range of biological roles, ranging from siderophores for the scavenging of trace iron, to biosurfactants for enhancement of motility and quorum sensing molecules for the orchestration of growth phases [4-6].

Over 22,000 bioactive natural products have been described to date, with 45% of microbial NPs produced by the actinomycetes [3]. Due to the massive chemical and bioactivity diversity, NPs and their derivatives have been co-opted as pharmaceutical compounds since the discovery of penicillin over 80 years ago [7, 8]. Over the last forty years, one-third of small molecule drugs approved by the US Food and Drug Administration (FDA) were NPs or derived/synthesised from NPs [9].

The clinical use of NPs and their derivatives has contributed to the extension of the human lifespan by 30 years during the 20th century [10]. In particular, NP antimicrobials, such as oxytetracycline and erythromycin, have curbed infectious disease mortality. Antimicrobials alongside NP immunosuppressants, such as cyclosporine and tacrolimus, have heralded huge developments in surgical techniques and organ transplantations [11]. Over 30% of approved anticancer agents are NP-derived [9], including doxorubicin and dactinomycin. This small selection of molecules exemplifies the chemical diversity of NP-derived drugs, as illustrated in **Figure 1**.

Natural products continue to play an important role in the pharmaceutical industry. The market for antibiotics from *Streptomyces* species alone had a value of $25 billion in 2001 [12]. However, the emergence of antibiotic resistance is inevitable and is predicted to cause 10 million deaths per year by 2050 [13]. As such, it is more important than ever to isolate novel NPs and characterise their bioactivities in search for antimicrobial compounds.

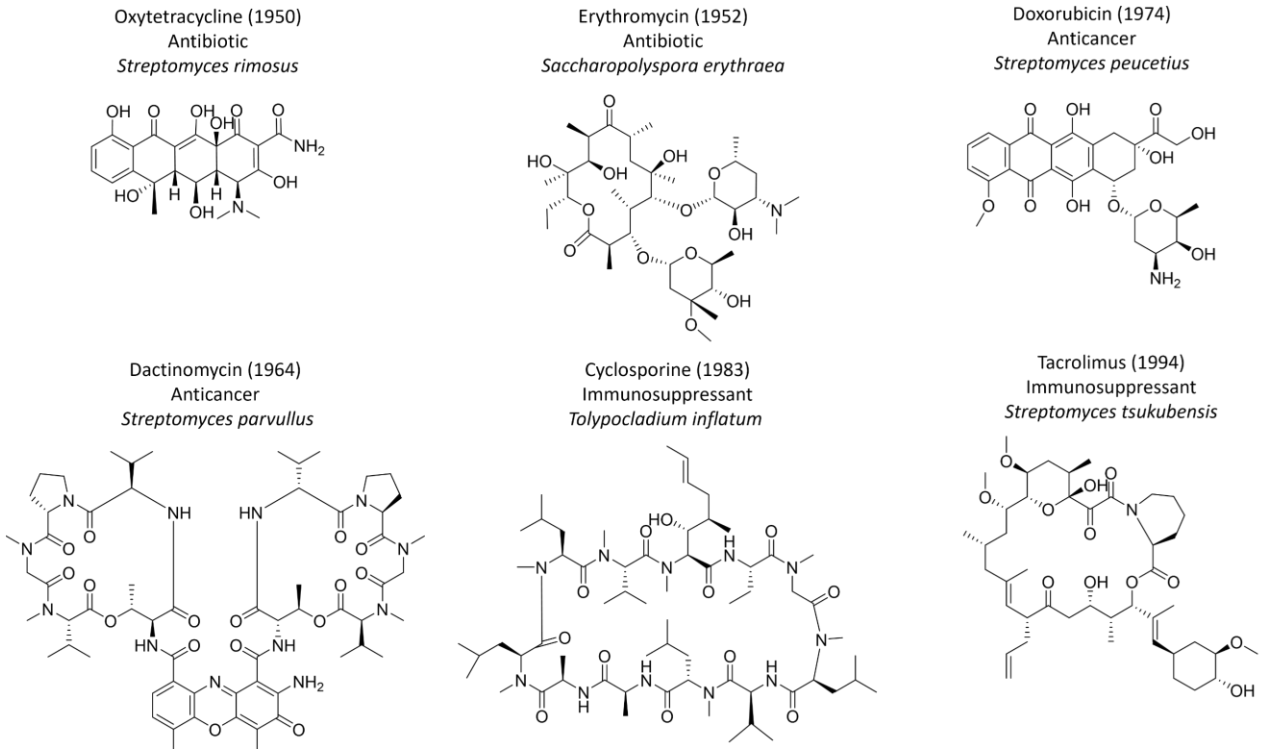


Figure 1. Structures of clinically important natural product drugs. Producer strain, drug type and year of approval by regulatory authorities are denoted.

1.2 A Brief History of Antibiotic Discovery

Infectious diseases have been a major cause of mortality throughout human history. As recently as the late 19th century, infectious diseases were responsible for 72% of deaths [14]. NPs have been indirectly leveraged by humankind for millennia for their therapeutic properties. Antibiotic-producing microbes were present in 'medicinal soil' and poultices of mouldy bread used to treat open wounds, described in the Eber's papyrus from 1550 B.C. [15]. As the field of bacteriology was developing throughout the 18th and 19th centuries, so was our understanding of bacterial antagonism. The concept of one microorganism antagonising the growth of another microorganism was described in 1877 by Louis Pasteur [16].

The very first clinical antibiotics were discovered during the synthesis and derivatisation of dyes. Paul Ehrlich investigated a group of arsenic-containing dyes for the ability to selectively kill the causative agent of syphilis, *Treponema pallidum*. Ehrlich identified arsphenamine, which was commercialised in 1910 as 'Salvarsan'. Salvarsan was the first effective syphilis treatment and was lauded by the media as the 'magic bullet'. The next

major antibiotic was discovered by Gerhard Domagsk in 1932 by the derivatisation and screening of over one thousand azobenzene dyes [17]. The sulphonamido-containing lead compound was found to be effective against *Streptococcus pyogenes* infection and was marketed in 1935 as Prontosil.

The first marketed microbial NP antimicrobial was penicillin. In 1928, Alexander Fleming demonstrated that *Penicillium* medium filtrate selectively killed *Staphylococcus aureus* [18]. In 1939 the bioactive NP, penicillin, was extracted from the growth medium by Howard Florey, Ernst Chain and Norman Heatley and shown to protect mice from *Streptococcus pyogenes* and *Staphylococcus aureus* infection [19]. Optimisation of the growth of the *Penicillium* mould and extraction of penicillin was prioritised throughout and after World War II. Dorothy Hodgkin determined the structure of penicillin in 1949 [20] and ascertained that it contained a beta-lactam group, which has since allowed the development of later generations of beta-lactam based antibiotics which can overcome resistance mechanisms. The use of penicillin has saved tens of millions of lives globally since its discovery [21].

The development of penicillin paved the way for the isolation of effective antibiotics from microorganisms and brought about a surge in research. Selman Waksman developed the platform for screening soil-derived *Streptomyces* strains for antimicrobial activity, by detecting the zones of inhibition against a reporter strain. This 'Waksman platform' led to a twenty-year period, which is now referred to as the 'golden age' of antibiotic discovery, where almost all major antibiotics used today were discovered. This included the discovery of streptomycin in 1944, which was purified from the soil bacterium *Streptomyces griseus* and proved to be the first antibiotic effective at treating tuberculosis [22, 23]. In the late 1940s, *Streptomyces venezuelae* was found to produce an antimicrobial, which was chemically characterised and termed chloramphenicol [24-26]. In 1953, vancomycin was isolated from the actinobacterium *Amycolatopsis orientalis* and was asserted to be an effective treatment against penicillin-resistant *Staphylococcus aureus* infection [27, 28] (**Figure 2**).

This method of screening soil bacteria led to the discovery of about half of the drugs used today [29] and also demonstrated the great potential that actinomycetes possess for the production of secondary metabolites. It is estimated that actinomycetes produce over two third of the antibiotics in use today and members of the *Streptomyces* genus account for ~75% of this amount [30].

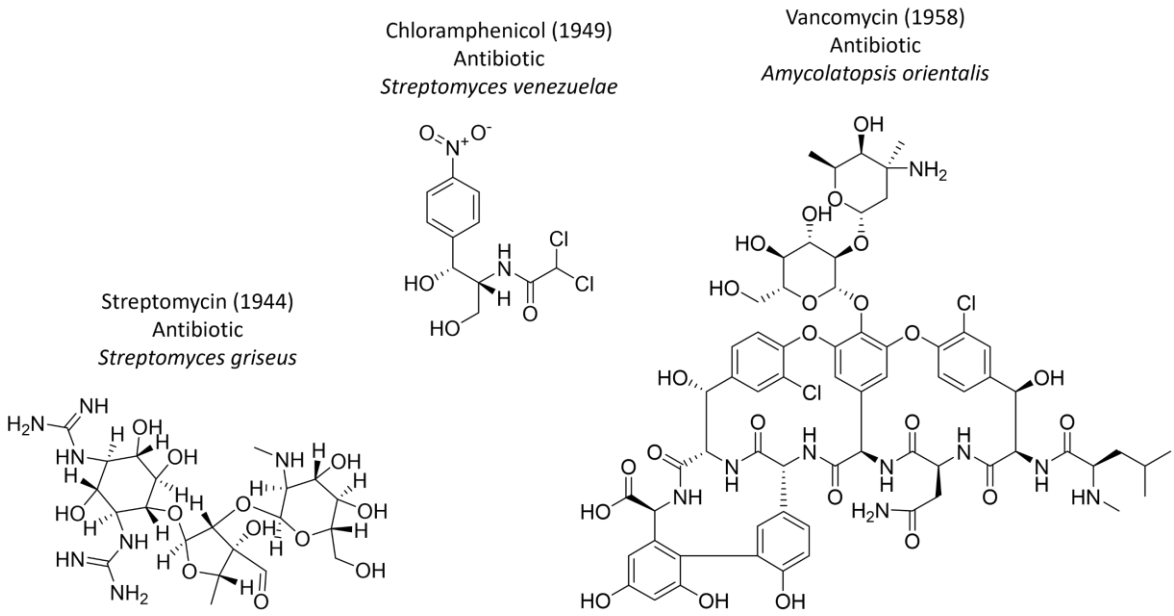


Figure 2. Structures of antibiotics discovered during the 'Golden age of discovery'. Producer strain, drug type and year of approval by regulatory authorities are denoted.

However, the pace of antibiotic discovery slowed from the 1960s onwards. The re-discovery of known antibiotics plagued the Waksman platform and made finding novel NPs near impossible. This is exemplified by the report that approximately 1% of soil actinomycetes produce streptomycin whereas the discovery of daptomycin required the screening of 10^7 actinomycetes [31]. At this point, the pharmaceutical companies minimally invested in antimicrobial development as it was considered that the existing drugs were effective. Instead, synthetic compound libraries were screened for bioactive molecules which could be derivatised and optimised. However, the rate of antimicrobial discovery was very low (GSK reported the screening of 500,000 synthetic compounds that did not lead to further development [32]) and this approach had the disadvantage of being unable to identify antimicrobials from novel chemical classes [33]. This lack of drug discovery caused a lag in approved new drugs which encroached into the early 2000s. The development of antibiotic resistance and emergence of new infectious diseases (e.g. HIV and COVID-19) has highlighted the importance of discovering bioactive NPs with novel chemistry.

1.3 Antibiotic Resistance

Interest in antimicrobial discovery has steadily increased as the threat of antimicrobial resistance (AMR) has become apparent. AMR is recognised as one of the most serious global threats to human health, with over 35,000 deaths caused by AMR infections per year in the US [34] and the global burden estimated at over 1 million deaths annually [35]. With an increasing rate of clinically-relevant bacterial strains developing antimicrobial resistance and thus the reduction of efficacy of available antimicrobials, it is predicted that AMR may cause 10 million deaths annually by 2050 [13], although this magnitude is disputed [36].

The development of AMR is inevitable as within several years of an antibiotic's discovery, resistant bacteria strains are reported [37]. For example, *S. aureus* infections were initially effectively treated by penicillin but within three years, penicillin-resistance was observed to readily occur in the laboratory and also in clinical isolates [38, 39]. In 1960, methicillin was synthesised and found to be effective at treating penicillin-resistant *S. aureus* infections [40, 41]. However, just a year later, methicillin-resistant *S. aureus* (MRSA) was detected [42] and to this day remains a serious global infectious agent [43]. Vancomycin was one of the most effective treatments of MRSA infection [44], however vancomycin-resistant *S. aureus* (VRSA) has recently been reported [45]. One of the last effective treatments of VRSA infection is daptomycin, although reports of daptomycin-resistant *S. aureus* are emerging [46, 47]. This timeline of continual antimicrobial discovery and the emergence of resistant strains exemplifies the importance of finding new antimicrobials, particularly those with novel structures, modes of action and with high barriers to resistance.

1.4 Antibiotic Discovery in the Postgenomic Era

Whilst overcoming the problem of AMR requires a concerted reform of healthcare practices and policy [48], one undeniably important aspect is the identification of novel antimicrobial compounds. In order to overcome the problem of re-discovering bioactive natural products by the classical Waksman approach, several new methods have been employed in the literature.

Many researchers have reasoned that the screening of microbes that are unlikely to have been subjected to the classical Waksman approach may produce novel bioactive NPs. Therefore, many studies have isolated microbes from diverse and extreme environments.

The fasamycins and formicamycins were isolated from ant-associated *Streptomyces* and show activity against MRSA and vancomycin-resistant *Enterococcus faecium* (VRE) [49]. Novel antimicrobial NPs have also been isolated from deep sea sponge-associated actinomycetes [50], hyper-arid desert environments [51, 52], polar regions [53] and pristine caves [54].

Another method of de-replicating isolate screening is by culturing the 'unculturable'. It is predicted that less than 1% of microbes have been cultured in the laboratory as the limited growth medium conditions do not support the growth of the 'unculturable' majority. Methods have developed to support 'unculturable' bacterial growth by more closely mimicking the natural environment. This includes the growth of monocultures with exchange of native nutrients via semi-permeable membranes, the use of co-culture, the addition of native growth factors and the use of anoxic or nutrient limited conditions [55-58]. The growth of the previously unculturable proteobacteria *Eleftheria terrae* resulted in the identification of the antimicrobial teixobactin, which is an effective treatment of MRSA *in vivo* [59].

Aside from the screening of novel and exotic microbial isolates, a great proportion of modern drug discovery is genomic-led. Natural product biosynthesis genes are often clustered together and transcriptionally linked in biosynthetic gene clusters (BGCs). The first two *Streptomyces* genomes were published in the early 2000s and genomic analysis indicated that they both featured 30+ BGCs, far exceeding the number of NPs detected [60-63]. This notion of talented producer strains has been expanded to actinomycetes in general, with many genera including *Saccharopolyspora* and *Rhodococcus* featuring large quantities of cryptic BGCs [64]. It is now predicted that only 10% of BGCs are expressed under routine laboratory conditions [65]. Therefore, many researchers are finding novel NPs by bioinformatically identifying cryptic BGCs, in a process called 'genome mining', and working to activate BGC expression for NP isolation and characterisation [66, 67].

The development of sequencing technologies has drastically reduced the price of bacterial genome sequencing and the number of bacterial genomes deposited into publicly accessible databases has been exponentially increasing. The NCBI database currently contains over 460,000 prokaryotic genomes, of which 38,000 are actinomycetes. One of the most widely used genome mining tools is the 'antibiotics and secondary metabolite analysis shell' (antiSMASH), which can identify 71 different types of BGCs from genomic data [68]. AntiSMASH predicts the function of query sequence gene products by sequence alignment and the identification of conserved domains, and also predicts accessory genes and cluster boundaries [69].

Once a cryptic BGC has been bioinformatically identified, there are many different reported methods for the induction of BGC expression. Perhaps the simplest is the 'One Strain – Many Compounds' (OSMAC) approach, whereby a single strain is grown under varied conditions (including media composition, pH, temperature, oxygenation levels, vessel types and co-culture) [70]. This is considered to result in varied transcription, proteome and therefore metabolome states in a single strain and has been widely reported to activate the expression of cryptic fungal and bacterial BGCs [71].

Further, precise genetic manipulation can induce cryptic BGC expression. For instance, negative regulators may be knocked out or positive regulators overexpressed, either at a global or pathway-specific level [72]. Engineering of ribosome genes has also resulted in enhanced secondary metabolite production [73]. In the case of a native producer with limited genetic tractability, heterologous expression using a 'superhost' strain is often favourable. Several cloning methods have been developed for the capture of BGCs up to 60-100 Kb in size and their conjugation into heterologous hosts [74]. Superhost *Streptomyces* strains have been engineered for enhanced secondary metabolite production by the removal of known BGCs and non-essential genes [75-77]. Heterologous expression also has the benefit of being able to introduce a strong constitutive promotor to drive NP biosynthesis and enhance compound titre [78, 79]. This heterologous expression approach has led to the expression of cryptic BGCs and the characterisation of novel NPs [72, 80].

1.5 Non-Ribosomal Peptides

Non-Ribosomal Peptides (NRPs) are a major class of NPs that are biosynthesised by large multimodular enzyme complexes called Non-Ribosomal Peptide Synthetases (NRPSs). NRPSs are organised into modules, which are each made up of three core domains: an adenylation (A) domain, a condensation (C) domain and a peptidyl carrier protein (PCP). In linear NRPS systems each module incorporates a single amino acid residue into the NRP. NRP biosynthesis proceeds with an A domain specifically adenylating and activating an amino acid residue. This is then loaded onto the 4-phosphopantetheine arm of the PCP domain and shuttled to the C domain. The C domain catalyses the condensation of the tethered amino acid residue to the nascent peptide chain of the preceding module via amide bond formation [81]. Each module may also contain tailoring domains which

catalyse specific modifications, such as epimerisation or oxidation. The final module features a terminal thioesterase (TE) domain which catalyses the offloading and/or macrocyclization of the NRP. Final modification and maturation of the NRP may be facilitated by other tailoring enzymes encoded in the BGC.

Canonical protein synthesis is limited to the twenty regular proteinogenic amino acids of which corresponding transfer RNA molecules (tRNAs) exist. However, NRP biosynthesis is not limited in this way and the adenylation domains may be specific to non-proteinogenic amino acids, of which there are over 500 examples. It is common for an NRPS BGC to encode enzymes for the biosynthesis of non-proteinogenic building blocks from proteinogenic amino acids or other primary metabolites [81]. Further, there is variability in the NRPS-mediated and post-NRPS modifications of these residues and the cyclisation of off-loaded peptide backbones. These factors result in the huge structural diversity of NRPs.

The actinomycetes are some of the most prolific producers of NRP NPs, where the large structural diversity of NRPs supports a wide range of bioactivities which enhance the competitiveness of the bacterium. For example, *S. coelicolor* produces coelichelin, a NRP which is produced under iron deplete conditions and functions as a siderophore [5, 82]. Further, *Streptomyces scabies* and *Streptomyces acidiscabies* produce thaxtomin A as a virulence factor for the phytopathogenesis of common potato scab [83, 84]. Actinomycete NRPs are not only important specialised metabolites for the enhancement of ecological fitness of the producing strain, but many have also been approved as drugs. For example, several *Streptomyces* strains produce the NRP antibiotics vancomycin and dactinomycin, and *Actinomyces antibioticus* produces the NRP anticancer agent dactinomycin [85, 86] (**Figure 3**). Of the currently marketed drugs, nearly 30 are NRPs, which translates to a value of hundreds of millions of dollars to the pharmaceutical industry [81].

NRP BGCs are also attractive targets for genome mining as the well characterised and repetitive nature of the multi-modular NRPS complex is effectively identified bioinformatically. Further, the identity of incorporated amino acid residues can be predicted by consideration of the adenylation domain sequences [87, 88]. The functional and chemical diversity, along with the relative ease of bioinformatic identification, make NRPs an attractive target for genomic-led drug discovery and the identification of novel classes of bioactive NPs. In this project, the biosynthesis of two actinomycete NRPs (actinonin and the matlystatins) are investigated.

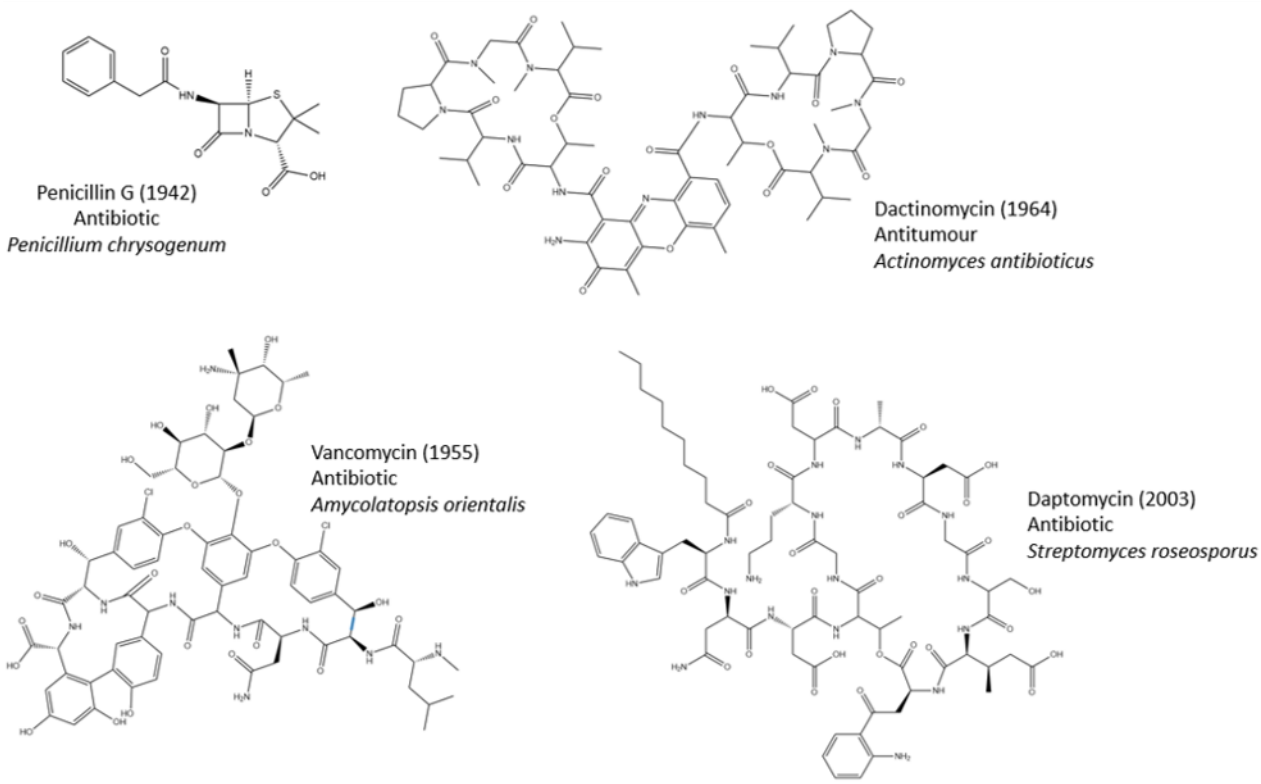


Figure 3. Structures of clinically relevant non-ribosomal peptides. Producer strain, drug type and year of approval by regulatory authorities are denoted.

1.6 Actinonin

Actinonin was isolated from *Streptomyces* sp. NCIMB 8845 in 1962, by Gordon and colleagues [89]. Actinonin was characterised as an antibiotic, active against both Gram-positive bacteria and Gram-negative bacteria, at a minimum inhibitory concentration (MIC) ranging from 20-100 µg/ml. Moreover, the *in vivo* toxicity of actinonin proved to be very low, with no toxic effects occurring in mice up to 400 mg per kg of body mass [89]. In 1968, Attwood determined that actinonin exerts its antimicrobial effects in a bacteriostatic manner [90]. The structure of actinonin was determined by IR and UV spectroscopy and mass spectrometry (MS) (**Figure 4**) [91].

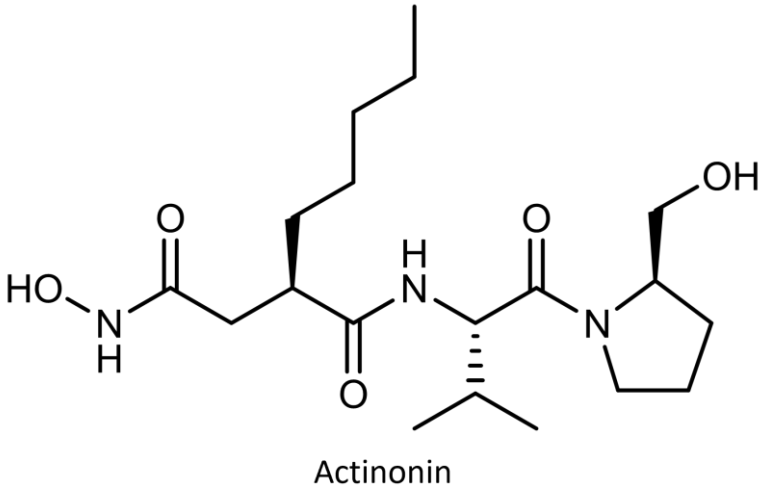


Figure 4. Structure of Actinonin.

The same group went on to develop a series of actinonin derivatives of various structural permutations, including various stereoisomers, the removal of the pentyl side chain and the addition of branched and unbranched hydrocarbon side chains at various R sites [92-95]. Structure activity relationship (SAR) analysis of these actinonin derivatives suggested that the antimicrobial activity was conferred by the hydroxamic acid moiety [96, 97].

A specific molecular target of actinonin was identified in 1985. Takeuchi and colleagues isolated actinonin from bacterial culture fractions during the screening for aminopeptidase M (APM) inhibitors. This interaction was characterised and actinonin was determined to be a potent APM inhibitor (K_i = 170 nM) [98]. APM is a membrane bound zinc-dependent metalloproteinase expressed on the surface of human kidney and brain cells which preferentially cleaves N-terminal amino acids from oligopeptides [99].

Having determined that actinonin was able to inhibit the metalloproteinase APM, the activity of actinonin against other metalloproteinases were also characterised. APM is a member of the enkephalinase family of metalloproteinases which are important for regulating the pain response in humans. The enkephalins are pentapeptide ligands localised to the central nervous system which bind to opioid receptors and activate pain inhibition pathways [100-106]. Enkephalins have been considered as analgesic (pain killing) drugs, however the analgesia rapidly dissipates in 10 minutes in rat models [107], which was attributed to the degradation of enkephalins by enkephalinase proteins. Actinonin was found to inhibit three enkephalinase metalloproteinases at <10 µM IC_{50} values [108-111] and the co-administration of enkephalin and actinonin to mice extended the analgesic effect by approximately 50% [111]. Therefore, these are further reports of actinonin inhibiting metalloproteinases and having an effect *in vivo* at non-toxic concentrations.

Actinonin has also been reported to have an anti-proliferative effect upon tumour cells. *In vitro* human breast cancer cell lines, grown within collagen gels, were challenged with actinonin (20 μg/mL) and some had completely abolished tumour cell proliferation [112]. This observation was further rationalised as the actinonin-sensitive cancer cell lines were overexpressing gelatinases whereas actinonin-resistant cancer cell lines lacked gelatinase activity [112]. The anti-proliferative effect of actinonin was also tested *in vivo* where the treatment of actinonin doubled the survival of mice transplanted with leukaemia cells by reducing the tumour growth rate [113]. Therefore, these data were the first to suggest that actinonin has an anti-proliferative effect on some types of human tumour cells.

Also, the observation that cells overexpressing gelatinases were more sensitive suggests that actinonin may inhibit gelatinases. This is the first indication that actinonin may inhibit matrix metalloproteinases (MMPs), which are a group of proteins that work in concert to remodel the extracellular matrix (ECM) [114]. There is a huge body of work researching the role of MMPs in tissue remodelling during human development [115, 116] and the development of diseases, including cancer development and metastasis [117-121], inflammatory diseases [122, 123] and vascular diseases such as stroke and atherosclerosis [124-127]. Therefore, this frames actinonin as a potential anti-cancer drug and alludes to actinonin having more diverse therapeutic value, however further work was required to determine whether actinonin was able to inhibit any other MMPs.

It was not until 1999, nearly four decades after the initial discovery of actinonin as an antibiotic, that the antimicrobial mechanism was elucidated. During the screening of compounds which may be able to inhibit the peptide deformylase (PDF) enzyme, actinonin was highlighted [128].

PDF is involved in bacterial protein synthesis (**Figure 5**), where only the initiator tRNA ($tRNA_i$), harbouring a formylmethionine residue (fMet), can initiate translation. The process of formylation/de-formylation is well understood. Firstly, the formyltransferase (FMT) enzyme uses the 10-formyltetrahydrofolate (10-fth) formyl donor to formylate the methionine bound to the initiator tRNA (Met-$tRNA_i^{Met}$) to generate fMet-$tRNA_i^{Met}$ [129, 130]. The fMet-$tRNA_i^{Met}$ initiates translation, forming a growing polypeptide with an N-terminal formylmethionine residue. The PDF enzyme localises to the exit of the ribosome and catalyses the post-translational deformylation of the polypeptide by cleaving the formylmethionine amide bond [131]. The structures of various PDF proteins have been determined and each feature a catalytic metal ion which is essential for hydrolysis of the

formylmethionine amide bond, with examples of iron, zinc and nickel-dependent PDFs [132-134].

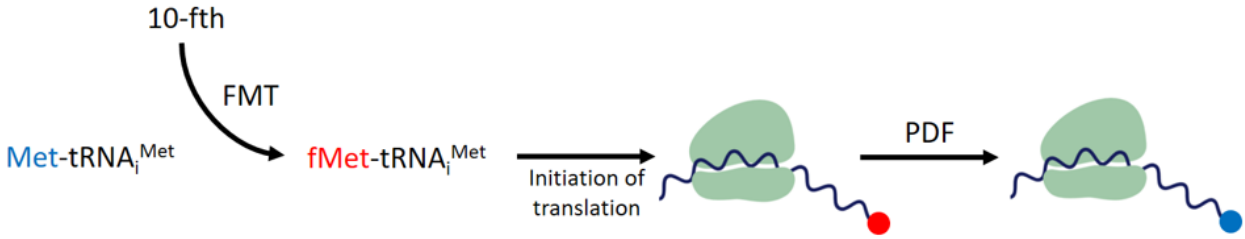


Figure 5. Schematic of formylation/de-formylation during bacterial protein synthesis. Blue denotes methionine, whereas red denotes formylated methionine. $tRNA_i^{Met}$ = initiator tRNA (anti-codon specificity to methionine); 10-fth = 10-formyltetrahydrofolate; FMT = formyltransferase; PDF = peptide deformylase. Adapted from [135].

The screening of a chemical library using a PDF assay led to the identification of actinonin, which transpired to be the most potent natural PDF inhibitor ever described with a sub-nanomolar IC_{50} against the *E. coli* Fe-PDF enzyme *in vitro* [128]. To further characterise the antimicrobial effect of actinonin, the *E. coli* PDF gene was put under the control of the arabinose inducible promotor P_{BAD}. There was a positive correlation between arabinose concentration in the media and the actinonin MIC [128], demonstrating that actinonin was exerting its antimicrobial effects by inhibiting PDF. The crystal structure of actinonin bound to *E. coli* PDF was published shortly after and clearly indicates bidentate coordination of the catalytic zinc ion by the two oxygen atoms of the actinonin hydroxamate group [136] (PDB:1G2A) (**Figure 6**). Therefore, actinonin inhibits PDF activity by occupying the active site and preventing the entry of formylmethionine. The development of actinonin and actinonin derivatives as antimicrobial drugs is explored in the following section **1.6.1**.

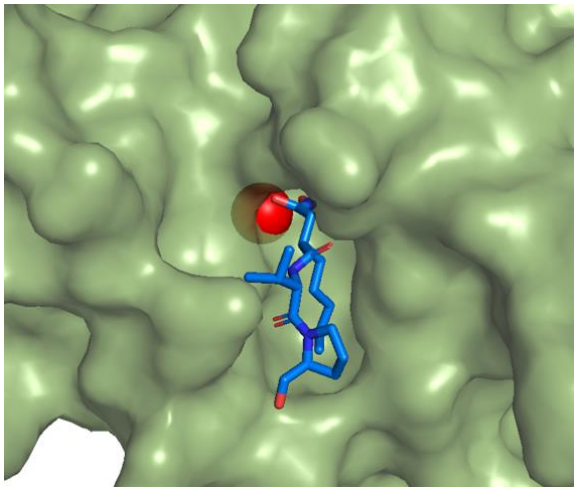


Figure 6. Crystal structure of actinonin bound to *E. coli* PDF. Surface of PDF shown in green. The active site catalytic nickel ion is show in red. PDB:1G2A.

Other varied bioactivities have also been described for actinonin. Actinonin has anti-malarial function, as it can inhibit a membrane metalloproteinase present on the apicoplast (organelle involved in lipid synthesis) of the causative agent of malaria, *Plasmodium falciparum* [137, 138]. As such, actinonin can immediately halt apicoplast biogenesis and cause the death of the parasite [139].

Actinonin has also been implicated in reducing acute renal failure (ARF), which is the loss of kidney failure which occurs in 20% of critically ill patients and has a mortality rate of 70% [140]. ARF can be caused by the ischemia/reperfusion (the loss and then recovery of blood supply to kidney tissue) or sepsis [141, 142]. Meprins are metalloproteinases present on surface of kidney tubules which have been implicated in the pathology of ARF [143]. Actinonin is able to tightly inhibit the meprins (K_i = 20nM [144]) - thus decreasing kidney damage resulting from ischemia/reperfusion [145, 146] and sepsis [147, 148] in mice models. Therefore, actinonin offers a novel method of treating both ischemia/reperfusion-induced and sepsis-induced ARF.

This same meprin protein has also been implicated in the pathology of atherosclerosis, as it can metabolise vasoactive peptides. In a mouse model of atherosclerosis, treatment with actinonin decreased the plaque volume and supressed lipid deposition in carotid arteries, thus exerting a protective role [149]. Further, the meprin inhibitory activity of actinonin has been correlated to a decrease of pulmonary edema in a murine model of acute lung injury but a mechanism has not been investigated [150].

Therefore, actinonin is of scientific interest due to its hydroxamic acid group, which mediates the inhibition of a wide range of metalloproteinases and thus confers a number of bioactivities which may serve a therapeutic function.

1.6.1 Development of Therapeutic Peptide Deformylase Inhibitors

Peptide deformylase inhibitors (PDFIs) were considered to be very interesting leads for antimicrobial drug development throughout the 2000s. This was supported by reports that PDF activity was essential for bacterial survival, present in all bacteria and that non-toxicity was guaranteed as no mammalian PDF counterparts had been described [151-153]. Actinonin may have been the most potent PDFI *in vitro*, however it had limited activity *in vivo* due to poor bioavailability [154] and being rapidly exported by efflux pumps [155].

There was therefore considerable interest in derivatising actinonin to enhance bioavailability for use as a therapeutic antimicrobial compound. Further structural reports of PDF proteins in complex with actinonin [156] along with existing understanding of metalloprotease inhibition allowed the development of a generic PDFI skeleton which featured a terminal pharmacophore for the metal ion chelation and a peptide backbone mimicking the formylmethionine PDF substrate (**Figure 7**) [151].

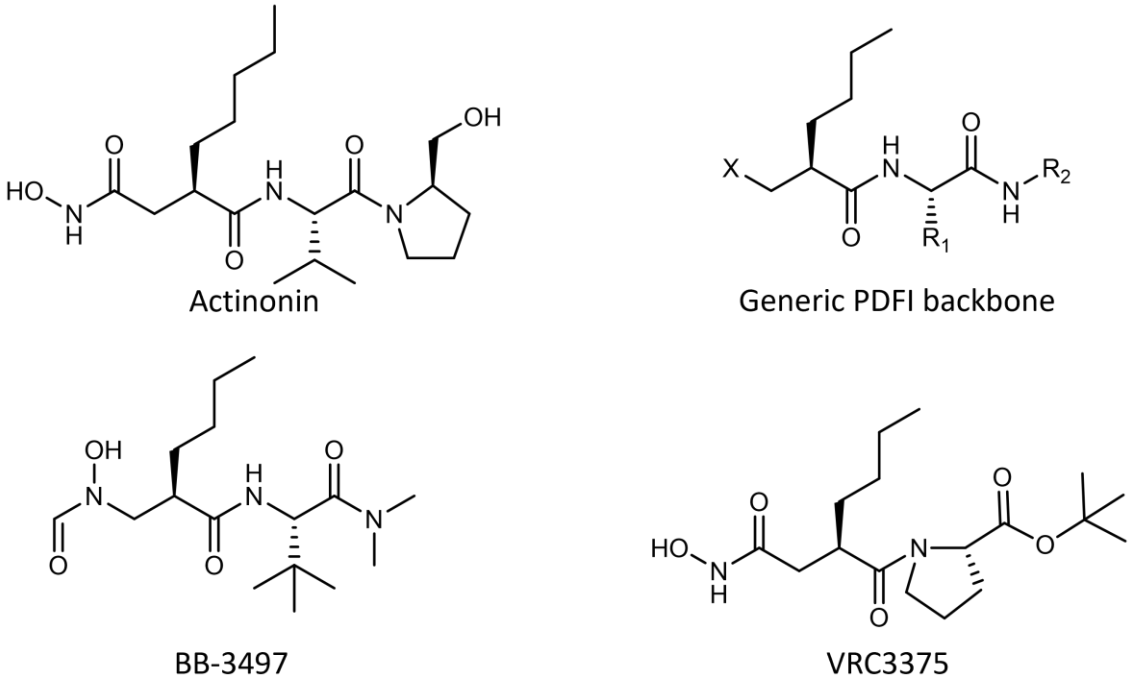


Figure 7. Structures of actinonin and first-generation PDF inhibitors

Two PDFIs with *in vivo* activity were developed by applying this logic to the synthetic development of chelator-based chemical libraries. BB-3497 features a so-called 'reverse hydroxamate' group (**Figure 7**) and strongly inhibited *E. coli* Ni-PDF with an IC_{50} of 7 nm and had sufficient bioactivity for the complete treatment of mice infected with *Staphylococcus aureus* [136]. Extensive derivatisation of BB-3497 and SAR analysis indicated that a terminal hydroxamic acid (alike actinonin) increased the potency to an IC_{50} of 1 nM whereas modification of the peptide backbone portion of the compound did not enhance potency [157-159].

Using similar methodology, VRC3375 was identified with strong *in vitro* Ni-PDF inhibition with an IC_{50} of 4 nM [160] (**Figure 7**). VRC3375 offered sufficient bioavailability for the curing of *S. aureus* infection in mice and proved to have low toxicity with toleration up to 400 mg/kg with no adverse effects [160]. These studies all indicate that the pentyl chain and the hydroxamate metal-chelating group are important moieties for a PDFI [161].

There was interest in the pharmaceutical industry in developing these types of PDFIs for clinical use. A considerable quantity of derivatised variants of the generic PDFI backbone were synthesised and tested. Novartis were developing VIC-104959 (**Figure 8**), which was effective against multidrug resistant clinical isolates [162] and entered phase I clinical trials as an oral antibacterial agent [163]. However, some adverse effects have since been reported [164]. Oscient Pharmaceuticals were developing BB-83698 (**Figure 8**) which passed phase I clinical trials with no significant adverse effects upon intravenous infusion [163, 165]. However, both of these compounds were discontinued after phase I clinical trials.

Glaxo Smith Kline (GSK) developed the GSK1322322 PDFI (**Figure 8**) for intravenous and oral treatment of acute bacterial skin infections, which proved effective against multidrug resistant bacterial isolates (including MRSA) *in vitro* [166, 167]. GSK1322322 proved to have good pharmacokinetics and tolerability in the phase I clinical trial [168-170]. GSK1322322 was also subjected to a phase II clinical trial, however had worse side effects, withdrawal symptoms and clinical success rates as compared to the standard antibiotic treatment linezolid [171]. There were also problems with potentially reactive metabolites [172] which led to the discontinuation of GSK1322322 development.

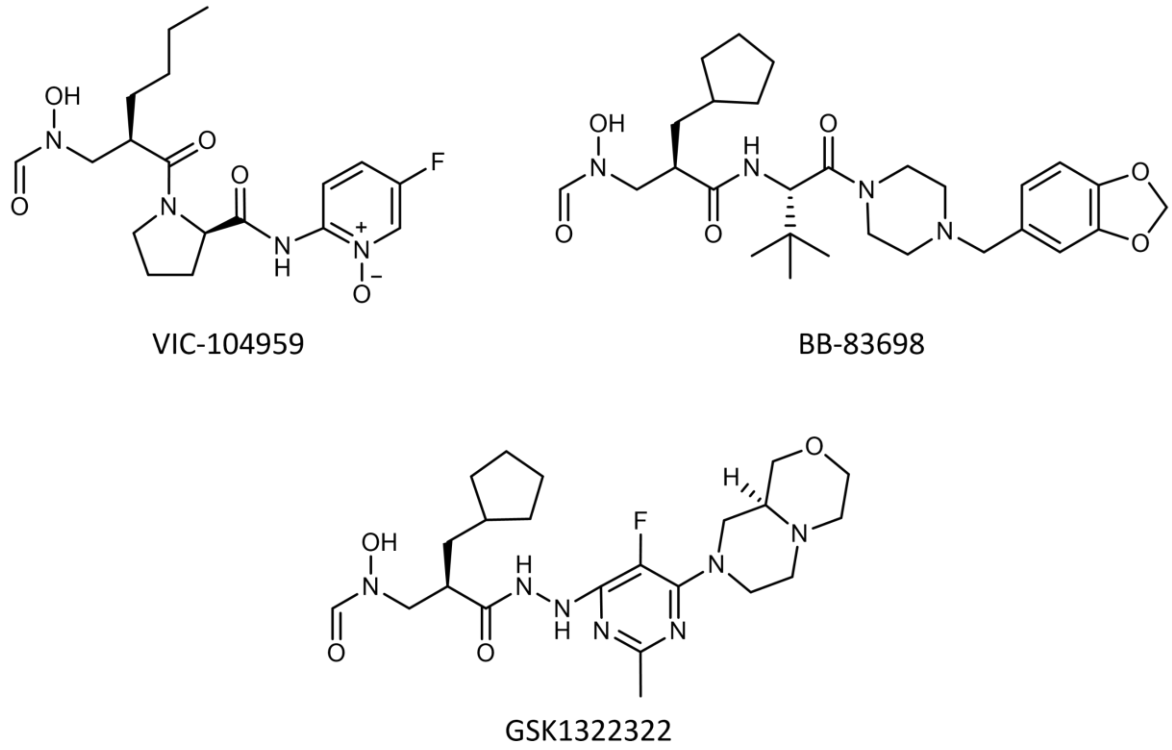
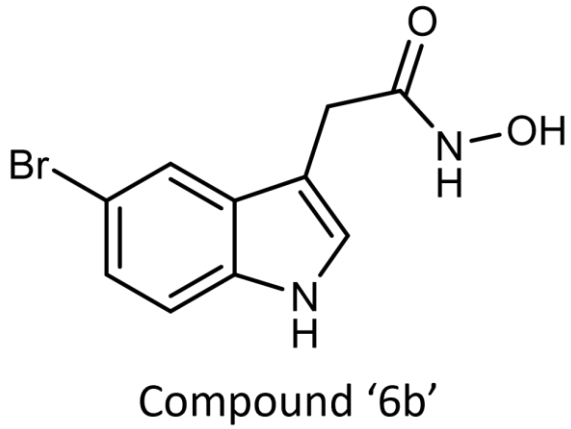


Figure 8. PDF inhibitors that reached clinical trials.

Compounds featuring the hydroxamate group (e.g. VRC3375) and those featuring the reverse hydroxamate group (e.g. BB-3497, VIC-104959, BB-83698 and GSK1322322) all show good efficacy *in vitro* and have good bioavailability *in vivo*. Therefore, the development and characterisation of chemically related compounds is of interest. Although no PDFI has been approved for clinical use, they represent a novel class of antibiotics with a completely unique mode of action. Therefore, they would not suffer from cross-resistance against existing classes of antibiotics and thus represent important lead compounds for tackling multidrug resistant bacteria.

However, the prevailing notion that PDFI drugs represented a perfect antimicrobial drug solution were soon curtailed. Eukaryotic PDF proteins were identified [173, 174], firstly in plants where they were characterised as being essential for chloroplast protein processing [175-177]. Then, the human PDF homolog (HsPDF) was identified [178, 179] and characterised as being essential in mitochondrial protein processing [180-182]. This weakened the notion that PDFIs would be entirely specific to bacteria and therefore would offer little-to-no toxic side effects. However, there are reports of structural differences between bacterial and eukaryotic PDF proteins. Bacterial PDF proteins have a wider active site pocket which allows the entry of bulky inhibitors whereas the HsPDF protein has a

narrower active site pocket [180, 183]. Therefore, bulky PDFIs specific against bacterial PDFs have been rationally designed, including a hydroxamate-containing bromoindole compound (termed '6b' - **Figure 9**) which was as potent as actinonin against bacterial PDF but was 13,000x more potent against bacterial PDF compared to HsPDF *in vitro* [184].



Compound '6b'

Figure 9. Compound '6b' developed as PDFI selective against bacterial PDF. Adapted from [184]

Another hurdle with PDFI development is the development of antimicrobial resistance. The development of PDFI resistance was originally considered impossible as the PDF-mediated deformylation of translated polypeptides was considered essential for protein synthesis. However, PDFI resistance can be gained by developing inactivating mutations in the FMT gene, which formylates the Met-$tRNA_i^{Met}$ prior to translation. Strains lacking FMT activity are resistant to PDF inhibition as PDF is no longer essential for proteins synthesis. This loss of FMT function leading to PDFI resistance has been observed in *Bacillus subtilis, Pseudomonas aeruginosa, S. aureus* and *Salmonella enterica* [185-189]. Mutations in the IF-2 initiation factor can also allow formylation-independent translation, as observed in *P. aeruginosa* and *Salmonella typhimurium* [190, 191]. FMT bypass mutants in *P. aeruginosa* have also been demonstrated to increase expression of efflux pumps for the export of PDFIs [192]. Resistance gain via modification of the target has been observed in *Streptococcus pneumoniae* whereby mutations in residues involved with catalytic metal ion coordination in the PDF protein resulted in reduced growth rate but enhanced resistance against actinonin [193]. Therefore, there are several reported PDFI resistance mechanisms – either formylation-independent translation by bypassing FMT activity, modification of the PDF target or the expression of efflux pumps. Nevertheless, PDFIs represent a novel mode of antimicrobial activity, which suggests that there would be no cross-resistance with current antibiotics.

Another of the fundamental assumptions of PDFIs has also been undermined by the discovery of eukaryotic PDF proteins. Originally, PDFIs were assumed to be absolutely selective for bacterial PDFs. However, the HsPDF protein was identified and determined to be essential for mitochondrial protein processing [182]. Therefore, PDFIs have some cytotoxicity [113, 179] and this may contribute to the moderate side-effects observed in the clinical trials of PDFIs as antimicrobial agents.

However, this PDFI cytotoxicity may also contribute to the anti-proliferative activity of actinonin against human cancer cells, which had previously been attributed to MMP inhibition [112]. It has been reported that the inhibition of HsPDF results in the degradation of respiratory function and ATP levels [194] and leads to cancer cell death [195]. Indeed, a group of actinonin derivatives were screened against varied human cancer cell lines and only the derivatives that showed *in vitro* activity against HsPDF were potent antiproliferative agents [182]. This positions HsPDF as an interesting novel target for anticancer drugs. *hsPDF* has been identified to be overexpressed in breast, colon and lung human cancer lines and treatment with actinonin reduced cancer cell proliferation [196].

In a recent study, actinonin derivatives were developed for the optimisation of HsPDF inhibition in HsPDF overexpressing cancers. This led to the identification of compound '15m' (**Figure 10**) which suppressed several human cancer cell lines *in vitro* whilst not inhibiting the growth of a normal cell line [197]. This compound was also tolerated well *in vivo*, demonstrating minimal side effects and supressing tumour growth by 72% [197]. Therefore, there is renewed interest in the development of actinonin-like hydroxamate-containing PDFIs for use as anticancer agents.

There is also a recent report of using actinonin directly as an antiproliferative agent by overcoming its poor bioavailability characteristics. Actinonin was bound to a polymer via the hydroxamate group to form an inactive pro-drug conjugate which should slowly release free actinonin in the body [198]. When injected into mice infected with metastatic cancer cells, the actinonin-polymer conjugate inhibited the metastatic spread of cancer whereas free actinonin was unable to [198]. This study attributed the inhibition of metastasis to the inhibition of MMP and did not investigate a direct cytotoxic effect of HsPDF inhibition. However, this represents an interesting method of enhancing the bioactivity of actinonin and deserves further investigation in the future.

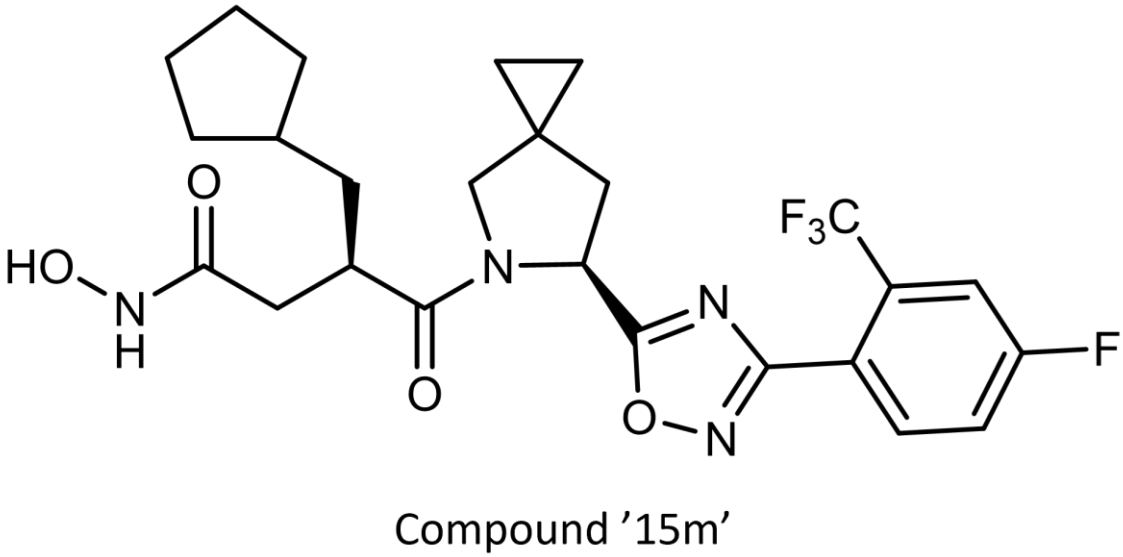


Figure 10. Compound '15m' developed as anti-cancer PDFI selective against HsPDF. Adapted from [197]

Therefore, there is lots of interest in the literature for the development of actinonin derivatives and hydroxamate-containing PDFIs. Developments in the rational drug design of compounds active against bacterial PDF and inactive against HsPDF has allowed the development of effective PDFI antimicrobial agents. On the other hand, PDFIs designed for the specific inhibition of HsPDF have proven to be potent and selective antiproliferative agents when used against cancers overexpressing HsPDF. This thesis will investigate the biosynthesis of actinonin and the characterisation of a natural product featuring a similar warhead moiety.

1.7 The Matlystatins

The matlystatins are a group of natural products which are of interest as they also feature the same N-hydroxyl-2-pentyl succinamic acid group as actinonin. The matlystatins are rather underexplored as natural products with considerably less research compared to actinonin. Therefore, it is possible that the entire range of bioactivities that the matlystatins possess has yet to be uncovered.

The matlystatins were discovered in 1992 during a screening of bacteria fermentation extracts, as type IV collagenase inhibitors [199]. The producer strain was identified as

Actinomadura atramentaria (SANK 61488) by considering its colony morphology, cell wall composition and physiological properties [199, 200]. It was found to produce five structural congeners of matlystatin (designated A, B, D, E and F) (**Figure 11**) [199].

The structure and stereochemistry of matlystatin A and matlystatin B were both unambiguously determined by their synthesis. The synthesised compounds had identical mass spectra, NMR spectra and specific rotation to the natural matlystatins [201, 202]. The remaining structural congeners; matlystatins D, E and F were structurally characterised by tandem MS and 2D NMR (**Figure 11**) [203].

Matlystatin A was determined to be able to inhibit collagenase (MMP-2) *in vitro* with an IC_{50} of 0.56 μM [204] and was also able to inhibit APN with approximately equal potency to actinonin [205]. Alike actinonin, the metalloproteinase inhibitory activity of matlystatin A was attributed to the hydroxamic acid group chelating the catalytic metal ion. There is some limited evidence that matlystatin A may have anticancer activity, as it was able to inhibit the invasion of human fibrosarcoma cancer cells into a basement membrane analogue *in vitro* [204].

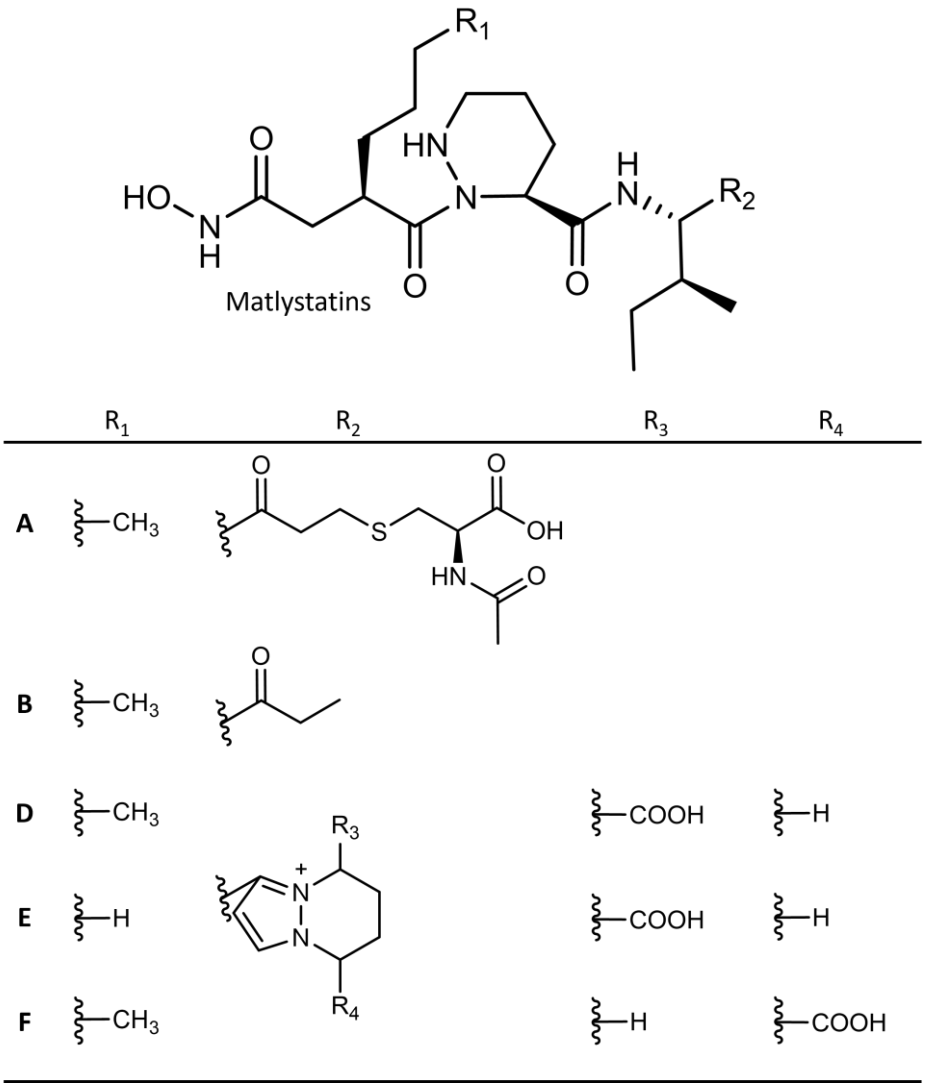


Figure 11. Structure of matlystatin congeners.

There have been several SAR studies conducted upon the matlystatins and their derivatives in order to improve their efficacy. For example, Sugimura and colleagues demonstrated that by altering the pentyl side chain there could be more favourable steric interactions with the target metalloproteinase, leading to greater inhibition [206]. One such chemically synthesised derivative featured a nonyl side chain and boasted an IC_{50} of 1.2 nM against MMP-9 – a potency 475x greater than matlystatin B [206]. Another group generated ten different stereoisomers of matlystatin B and tested their inhibition against collagenase, which determined that the most effective configuration was that of matlystatin B [202].

1.8 The Bioactive Warhead Group

It is known that the bioactivities of actinonin and the matlystatins are conferred by their hydroxamic acid. In particular, the potent PDFI activity of actinonin is being actively researched for the development of antimicrobial and anticancer agents.

However, up until recently very little was known about the biosynthesis of actinonin and the matlystatins. Considering the structures of actinonin and the matlystatins (**Figure 12**), they both feature the N-hydroxy-2-pentyl succinamic acid group (which will hereinafter be referred to as the 'warhead' group due to its role in binding the active site metal of metalloproteases). The warhead group contains the terminal hydroxamic acid moiety which is known to confer the bioactivity of these two molecules.

Taking a closer look at the pseudotripeptidic structures of actinonin and the matlystatins (**Figure 12**) gives an indication as to their biosynthesis. The structure of actinonin consists of the warhead group condensed to a valine and proline residue (which undergoes further modification). The structure of the matlystatins consists of the warhead group condensed to a piperazic acid, an isoleucine, and a variable portion (R_2, **Figure 12**).

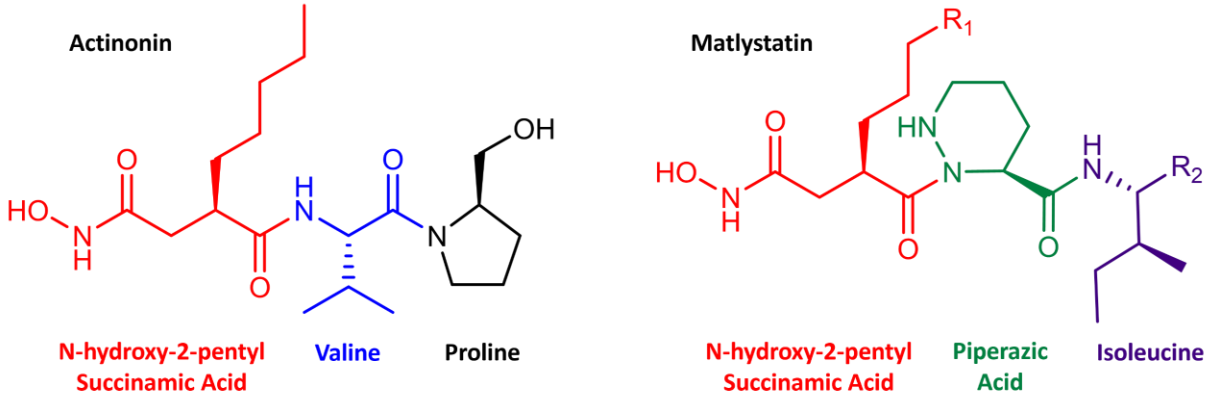


Figure 12. Comparison of pseudotripeptidic structures of actinonin and matlystatin. The monomeric units incorporated during the biosynthesis of these natural products are denoted.

Non-proteinogenic features such as the warhead and the piperazic acid residue are indicative that the compounds are produced by NRPS systems, however the biosynthesis of these molecules had not been reported. Further, it was unknown how the warhead group was biosynthesised. Therefore, a collaborative project between the Truman and

Kaysser (University of Tübingen) groups aimed to understand the biosynthesis of both natural products in addition to the biosynthesis of the warhead functional group [207].

The warhead group has also been reported in some other bacterial compounds, including propioxatin and BE16627B which also both have metalloproteinase inhibitory activity [208-213]. However, the actinonin and matlystatins were selected for characterisation of their biosynthesis as the producer strains were readily available.

1.9 Actinonin and Matlystatin Biosynthetic Gene Clusters

A candidate matlystatin BGC was identified in *Actinomadura atramentaria* DSM 43919, a native matlystatin producer, by searching for a suitable NRPS BGC within its genome [207]. This BGC features a NRPS/polyketide synthase (PKS) system and ethylmalonyl-Coenzyme A (EMC) pathway-like genes (**Figure 13**, **Table 1**). The native producer of actinonin, *Streptomyces* sp. NCIMB 8845 was sequenced, and the genome analysed using antiSMASH. This identified a putative actinonin BGC which again featured a NRPS system and an EMC-like pathway (**Figure 13**, **Table 2**).

With the putative BGCs of two warhead-containing natural products having been identified, the genes could be compared in order to suggest which genes are involved in warhead biosynthesis (**Figure 13**). Note that homologs of *matQ* and *matR* are not present in the putative actinonin BGC but highly similar homologs (>70% similarity) are present elsewhere in the *Streptomyces* sp. NCIMB 8845 genome (*actM* and *actN*, respectively).

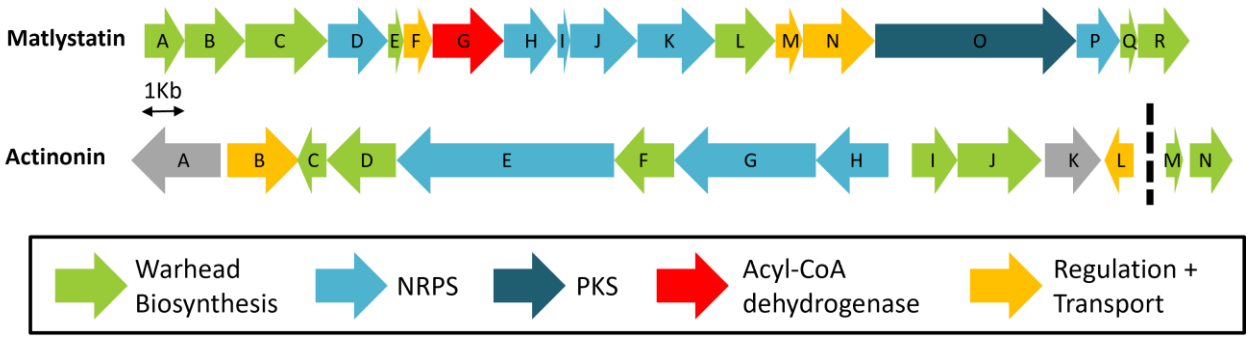


Figure 13. Schematic overview of the matlystatin and actinonin biosynthetic gene clusters (BGC). Genes are classified into functional groups, based on their predicted function. *actM* and *actN* are not present in the actinonin BGC and are instead found elsewhere in the genome.

Table 1. Details of proposed matlystatin biosynthesis genes. The size of each gene product is denoted. For each gene, the most homologous characterised protein was identified by BLAST analysis. The accession number, percentage identity and function of the homolog is annotated.

Gene	Size (aa)	Protein homolog	Accession number	Identity (%)	Predicted function
matA	392	TUE45_03041, *Streptomyces reticuli*	CUW28310.1	45	AurF-like N-oxygenase
matB	482	GA0115244_12547, *Streptomyces sp. DvalAA-19*	SCE40150.1	77	Methylmalonyl-CoA mutase, N terminal domain
matC	618	BE04_36800, *Sorangium cellulosum*	KYF49978.1	52	Asparagine synthetase B
matD	429	SAMN05444320_112113, *Streptoalloteichus hindustanus*	SHG70271.1	61	L-ornithine N5-oxygenase
matE	139	DFR75_10426, *Nocardia ignorata*	TDP37676.1	72	Methylmalonyl-CoA epimerase
matF	233	DVA86_13415, *Streptomyces armeniacus*	AXK33500.1	51	FMN-binding negative transcriptional regulator
matG	547	EOO71_02185, *Myxococcaceae bacterium*	RYZ43788.1	44	Acyl-CoA dehydrogenase
matH	421	OQI_27585, *Streptomyces pharetrae* CZA14	OSZ57397.1	35	Non ribosomal peptide synthetase
matI	97	A3L23_04981, *Rhodococcus fascians* D188	AMY56279.1	44	Non-ribomsomal peptide synthetase. Ppant attachment site
matJ	522	ABE83_07175, *Streptomyces* sp. CFMR 7	ALC31766.1	65	Amino acid adenylation protein
matK	605	GA0070563_12811, *Micromonospora carbonacea*	SCF50255.1	32	Non ribosomal peptide synthetase adenylation
matL	448	SAMN06272789_5495, *Streptomyces* sp. 1331.2	SOB85221.1	71	Crotonyl-CoA carboxylase/reductase
matM	204	BCF44_114157, *Kutzneria buriramensis*	REH38132.1	53	ArsR family transcriptional regulator
matN	521	ABS94_17420, *Variovorax* sp. SCN 67-85	ODU15778.1	34	MFS transporter
matO	1549	mscl, *Sorangium cellulosum*	AHB82059.1	45	Non ribosomal peptide synthetae/polyketide synthase
matP	313	Snas_4057, *Stackebrandtia nassauensis* DSM 44728	ADD43709.1	46	Thioesterase
matQ	136	CO540_15495, *Micromonospora* sp. WMMA2032	ATO15060.1	87	Methylmalonyl-CoA mutase
matR	417	DVA86_20870, *Streptomyces armeniacus*	AXK34738.1	61	Methylmalonyl-CoA mutase-associated GTPase MeaB

Table 2. Details of the proposed actinonin biosynthesis genes. The size of each gene product is denoted. For each gene, the most homologous characterised protein was identified by BLAST analysis. The accession number, percentage identity and function of the homolog is annotated.

Gene	Size (aa)	Protein homolog	Accession number	Identity (%)	Predicted function
actA	678	B1H18_00560, *Streptomyces tsukubensis*	OON82982.1	63	CocE/NonD family hydrolase
actB	554	ADK88_20190, *Streptomyces* sp. NRRL F-2295	KOU05014.1	69	C4-dicarboxylate ABC transporter
actC	152	ACZ90_58010, *Streptomyces albus* subsp. albus	KUJ64477.1	72	Methylmalonyl-CoA epimerase
actD	481	GA0115244_12547, *Streptomyces* sp. DvalAA-19	SCE40150.1	70	Methylmalonyl-CoA mutase, N terminal domain
actE	1589	DMF65_11600, *Acidobacteria* bacterium	PYS97535.1	39	Non-ribosomal peptide synthetase
actF	442	SNL152K_6626, *Streptomyces* sp. NL15-2K	GCB49292.1	68	Crotonyl-CoA reductase
actG	1027	C4J99_2447, *Pseudomonas synxantha*	AZE78232.1	32	Non-ribosomal peptide synthetase
actH	505	ASL14_11480, *Paenibacillus* sp. IHB B 3084	ALP36685.1	42	Non-ribosomal peptide synthetase
actI	340	TUE45_03041, *Streptomyces reticuli*	CUW28310.1	49	AurF-like N-oxygenase
actJ	619	BE18_19210, *Sorangium cellulosum*	KYF89188.1	53	Asparagine synthetase B
actK	463	WSS_A32995, *Rhodococcus opacus* M213	EKT78339.1	41	Amidase
actL	203	H181DRAFT_05163, *Streptomyces* sp. WMMB 714	SCK55771.1	37	TetR family transcriptional regulator
actM (pred.)	137	DVA86_10075, *Streptomyces armeniacus*	AXK32946.1	90	Methylmalonyl-CoA mutase, small subunit
actN (pred.)	327	DVA86_20870, *Streptomyces armeniacus*	AXK34738.1	88	Methylmalonyl-CoA mutase-associated GTPase MeaB

Both the putative actinonin and matlystatin BGCs contained genes which were homologous to genes previously shown to constitute the ethylmalonyl-CoA (EMC) pathway. The EMC pathway is utilised by proteobacteria and actinobacteria for acetate assimilation when they do not have the capability to perform the glyoxylate pathway [214, 215]. Preceding the EMC pathway, two CoA-activated acetate molecules are enzymatically transformed to the CoA-activated C4 acid, crotonyl-CoA.

The EMC pathway proceeds by the carboxylation of crotonyl-CoA by the crotonyl-CoA carboxylase/reductase (CCR) enzyme, to produce (2S)-ethylmalonyl-CoA. The stereochemistry is inverted by the ethylmalonyl-CoA epimerase to produce (2R)-ethylmalonyl-CoA. The ethylmalonyl-CoA mutase alters the carbon backbone by shifting the carboxylic acid group from the alpha to beta carbon position, to produce methylsuccinyl-CoA (**Figure 14**) [216]. Following the EMC pathway, the methylsuccinyl-CoA is ultimately converted to glyoxylate and enters central metabolism.

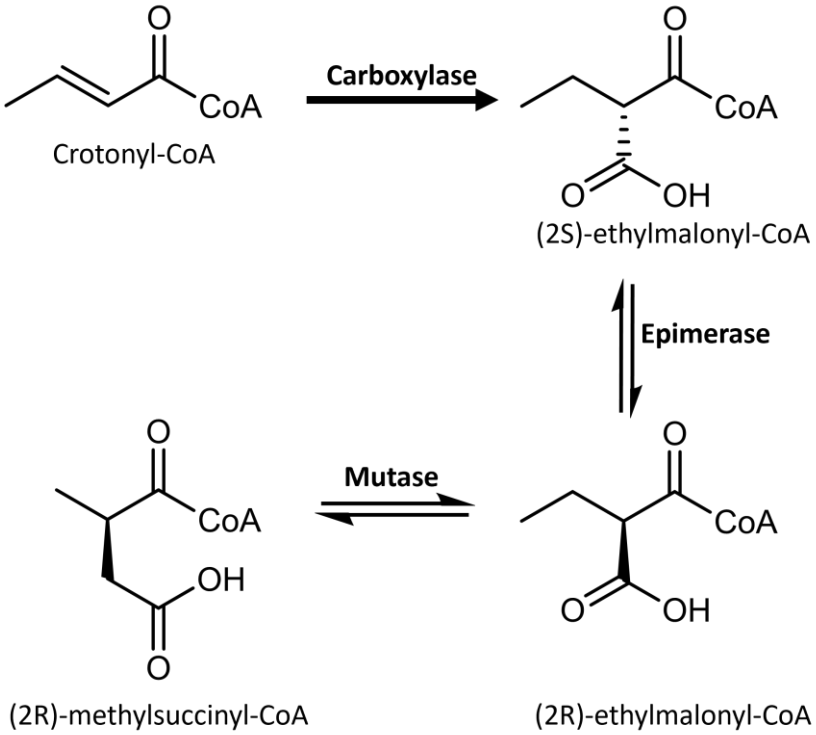


Figure 14. The ethylmalonyl-CoA (EMC) pathway. Abbreviations; crotonyl-CoA reductase/carboxylase (CCR), ethylmalonyl-CoA epimerase (epimerase), ethylmalonyl-CoA mutase (mutase). Adapted from [216]

The EMC-like genes were candidate warhead biosynthesis genes as they are present in both the actinonin and matlystatin BGCs, where they may catalyse the carboxylation, epimerisation and carbon backbone rearrangement of CoA-activated short chain unsaturated carboxylic acids. The asparagine synthetase and N-oxygenase genes are also present in both BGCs and were therefore also candidate warhead biosynthesis genes.

Leipoldt and Santos-Aberturas *et al*. proposed that the warhead biosynthesis pathway (**Figure 15**) starts with the CoA-activated trans-octenoyl acid, which is carboxylated at the alpha position by the CCR to produce 2S-hexylmalonyl-CoA. The acyl-CoA epimerase alters the stereochemistry to yield 2R-hexylmalonyl-CoA. This enables the acyl-CoA mutase to shift the CoA-bound carboxyl group from the alpha to beta position, producing 2R-2-pentyl succinyl-CoA. This is then amidated by the asparagine synthetase and then oxygenated by the N-oxygenase to yield a N-hydroxy-2-pentyl-succinamic acyl-CoA (warhead) moiety that could be incorporated into the molecule via the NRPS [207].

In order to go about ratifying this putative warhead biosynthesis pathway, a stable isotope labelling experiment was conducted by Leipoldt and Santos-Aberturas *et al*. [1,2,3,4-

$^{13}C_4$]octanoic acid was fed to the actinonin native producer strain, *Streptomyces* sp. NCIMB 8845. This labelled substrate was predicted to be converted *in vivo* to 2E-octenoyl-CoA. By considering the chemical transformations and carbon backbone rearrangements proposed in the putative pathway, the eventual position of the labelled carbons in the actinonin product was predicted. This feeding experiment was conducted and the labelled actinonin was isolated and structural characterised by ^{13}C NMR. The labelled carbons were incorporated as predicted [207], thus supporting the proposed warhead biosynthesis pathway (**Figure 15**). It is possible that some of these transformations occur at a later stage, although deletion of the N-oxygenase gene in *Streptomyces* sp. NCIMB 8845 completely abolished actinonin production [217], which is consistent with an early-stage biosynthetic step.

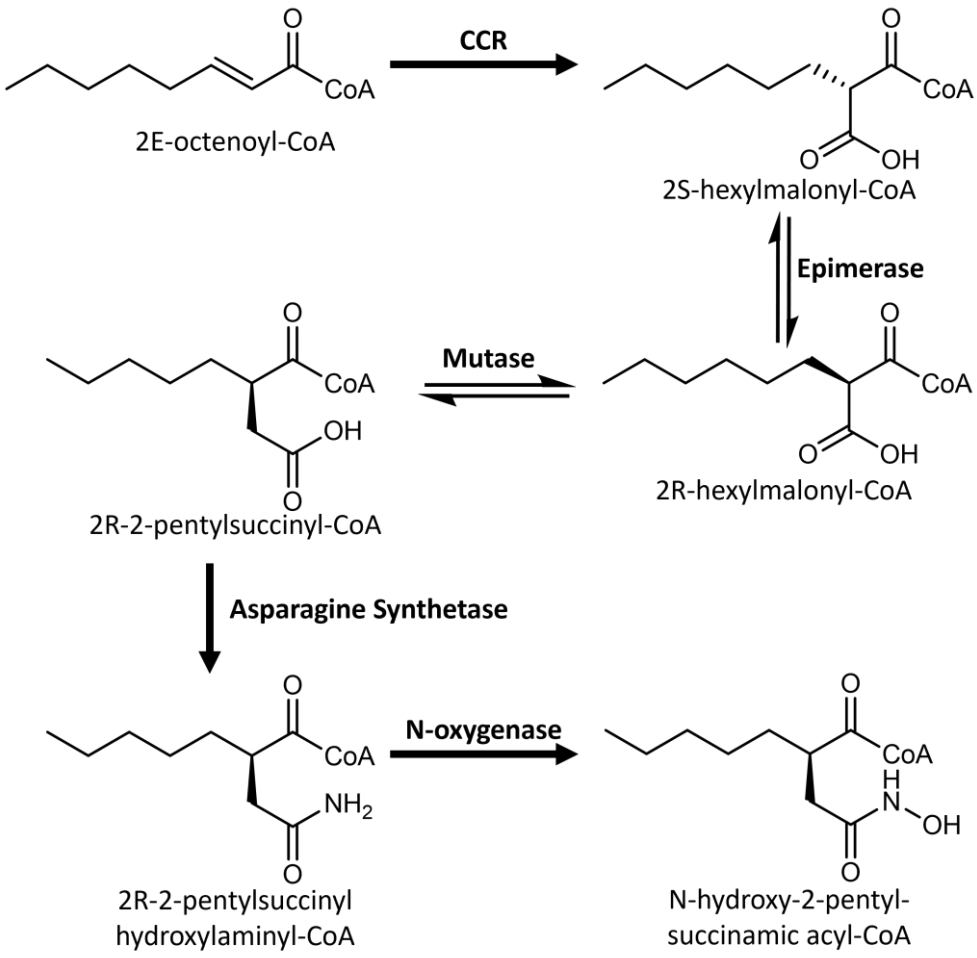


Figure 15. Proposed biosynthesis pathway of N-hydroxy-2-pentyl-succinamic acid 'warhead' group. CCR = crotonyl-CoA carboxylase. Adapted from [207].

Leipoldt and Santos-Aberturas *et al*. also investigated matlystatin biosynthesis in general (**Figure 16**). They proposed that the first module of the NRPS/PKS system adenylates piperazic acid and condenses it to the warhead. At module 2, isoleucine is specifically adenylated, loaded and condensed to the pseudodipeptide. The final module is a PKS module and is responsible for the condensation of methylmalonyl-CoA to the upstream polypeptide. The thioesterase domain releases the pseudotripeptide intermediate compound.

To yield matlystatin B, the intermediate compound is spontaneously decarboxylated. In order to yield all other matlystatin congeners, the intermediate compound is decarboxylated by the MatG enzyme to produce a reactive vinyl ketone moiety. A nucleophile attacks the electron rich vinyl ketone moiety to produce the final matlystatin compounds (**Figure 16**). This hypothesis would explain the diversity of matlystatin congeners (**Figure 11**) and was supported by the supplementation of the *A. atramentaria* growth media with unnatural nucleophiles, which resulted in the identification of novel matlystatin derivatives with altered C termini [207].

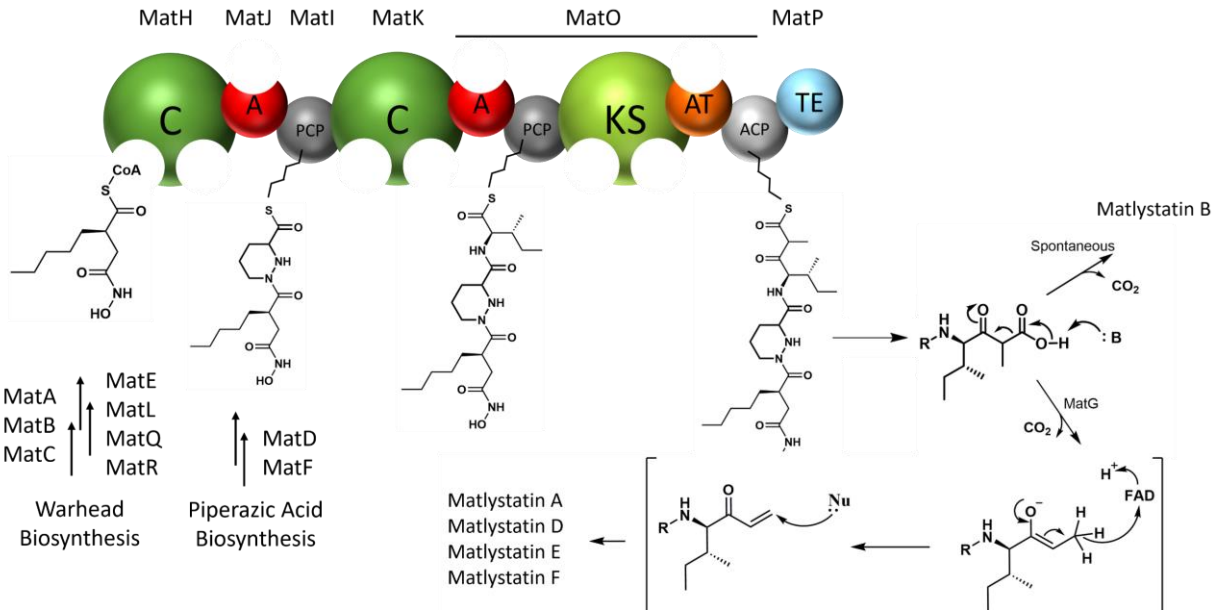


Figure 16. Proposed matlystatin biosynthesis. The NRPS/PKS system produces a pseudopeptide consisting of the warhead group, piperazic acid, isoleucine and methylmalonic acid. Abbreviations; C= condensation domain, A = adenylation domain, PCP = peptidyl carrier protein, KS = ketosynthase domain, AT = acyltransferase domain, ACP = acyl carrier protein, TE = thioesterase domain. Adapted from [207]

The work of Leipoldt and Santos-Aberturas *et al*. on actinonin and matlystatin makes up the background work for this thesis. The warhead biosynthesis pathway has been proposed (**Figure 15**) and was supported by a stable isotope labelling experiment [207]. However, a number of important questions remain unanswered.

1.10 Thesis objectives

Objective 1: Ratification of Proposed Warhead Biosynthesis Pathway

Work is presented towards the ratification of the proposed warhead biosynthesis pathway via its reconstitution *in vitro*. This is presented in chapter 2 and the specific objectives were to:

i. Synthesise and purify the proposed starting substrate, octenoyl-CoA
ii. Purify the proposed warhead biosynthesis proteins
iii. Develop *in vitro* assays and LC-MS analysis methodology

Objective 2: Identification and Characterisation of Warhead-Containing Natural Products

The work of Truman *et al*. represents the first report of putative warhead biosynthesis genes. An analysis of the phylogenetic abundance of these putative warhead biosynthesis genes led to the identification of putative novel warhead-containing NP BGCs at the end of chapter 2. A single putative novel warhead-containing NP BGC was selected and the strain, *Rhodococcus fascians D188*, retrieved. In chapter 3, the product of the cluster was investigated, and the specific objectives were to:

i. Identify the NP by comparative metabolomic analysis
ii. Purify the compound
iii. Elucidate the structure by LC-MS and NMR analysis

In chapter 4, the bioactivity of this natural product is examined. The specific objectives were to:

i. Determine whether the compound exerts antimicrobial activity
ii. Determine the mechanism of resistance of the producing strain
iii. Determine whether the compound exerts *in planta* bioactivity

Objective 3: Structural Characterisation of a Pseudomonas Cyclic Lipopeptide

In chapter 5, I contributed to a related project by isolating and structurally characterising a cyclic lipopeptide natural product. In practical terms, this work gave me experience of NP isolation and structural elucidation prior to the work performed for chapters 3 and 4.

Chapter 2: *In Vitro* Reconstitution of the Proposed Warhead Biosynthesis Pathway

2.1 Introduction

The actinomycete NPs actinonin and matlystatin both feature a N-hydroxy-2-pentyl-succinamic acyl-CoA group. The group contains a hydroxamic acid moiety which confers the metalloproteinase inhibitory activity of the compounds by chelating the catalytic metal ion of target proteins. The optimisation of actinonin derivatives has received much attention for the development of peptide deformylase inhibitors with antimicrobial or antiproliferative activity. Due to this bioactivity, the N-hydroxy-2-pentyl-succinamic acyl-CoA group is referred to as the 'warhead' group.

However, the biosynthesis of the warhead group, and actinonin and matlystatin in general, remained unknown. In 2017, the work of Leipoldt and Santos-Aberturas *et al.* identified candidate actinonin and matlystatin BGCs and proposed a putative warhead biosynthesis pathway (**Figure 15**) based on comparative genomics [207]. This proposed pathway had been supported by an isotopic labelling experiment [207], however the specific biochemical steps had not been experimentally ratified. In order to prove, or otherwise determine, the proposed warhead biosynthesis pathway, work is presented in this chapter towards the *in vitro* reconstitution of the pathway. The first three proposed reactions, that of the crotonyl-CoA carboxylase/reductase (CCR), hexylmalonyl-CoA epimerase and hexylmalonyl-CoA mutase proteins, are considered. These three putative warhead biosynthesis proteins are homologous to the known proteins of the ethylmalonyl-CoA mutase (EMC) pathway (**Figure 14**). However, the putative warhead biosynthesis proteins are proposed to have unique substrate specificity and process a longer fatty acyl-CoA thioester with an eight-carbon acyl chain rather than the known four-carbon acyl chain of the EMC pathway substrate.

In order to reconstitute the proposed warhead biosynthesis pathway, the following aims were defined:

i. Synthesis of the predicted starting substrate, 2-octenoyl-CoA
ii. Production and purification of the proposed warhead biosynthesis proteins
iii. *In vitro* reactions between the substrate and protein

In order to develop methodology for these aims, the literature was reviewed for protein purification and assay conditions, as detailed below.

2.1.1 Crotonyl-CoA Carboxylase/Reductase (CCR) Activity in Secondary Metabolism

The CCR enzyme was first characterised in 1995, at which time only the reductase activity was known, thus it was referred to as 'Crotonyl-CoA Reductase'. A CCR protein was purified from *Streptomyces collinus* and was shown to reduce crotonyl-CoA to butyryl-CoA (**Figure 17**), in the presence of NADPH and absence of bicarbonate [218]. This specific enzyme had tight substrate specificity as it was unable to reduce any other acyl-CoA thioesters [218].

The reductive carboxylation activity of the CCR enzyme was first identified in 2007, during a study of the metabolism of *Rhodobacter sphaeroides*. Most bacteria are able to assimilate acetate using the glyoxylate cycle. However, *R. sphaeroides* lacks vital enzymes of the glyoxylate cycle and is still able to grow on media with acetate as the sole carbon source. This led to the discovery of an independent metabolism pathway for the assimilation of acetate, named the ethylmalonyl-CoA (EMC) pathway [219] (**Figure 14**). The *R. sphaeroides* CCR enzyme involved in the EMC pathway was heterologously expressed in *E. coli* and was found to carboxylate crotonyl-CoA to ethylmalonyl-CoA, in the presence of bicarbonate (**Figure 17**). The CCR was shown to retain the ability to reduce crotonyl-CoA to butyryl-CoA in the absence of bicarbonate, albeit at a tenth the rate of the carboxylation reaction, suggesting that the carboxylation reaction is the physiologically relevant function [219]. Further, in the case of the carboxylation reaction the reactive carbon species was shown to be CO_2, rather than the bicarbonate itself [219, 220].

The proposed CCR carboxylation mechanism (**Figure 17**) proceeds with the transfer of a hydride ion from NADPH to the beta carbon of the α/β unsaturated enoyl-CoA substrate, in this case crotonyl-CoA [221]. Carboxylation proceeds via the electrophilic attack of CO_2 at the alpha carbon to produce ethylmalonyl-CoA. Whereas, the CCR reduction reaction proceeds by the incorporation of a solvent-derived proton at the alpha carbon, to produce butyryl-CoA [220].

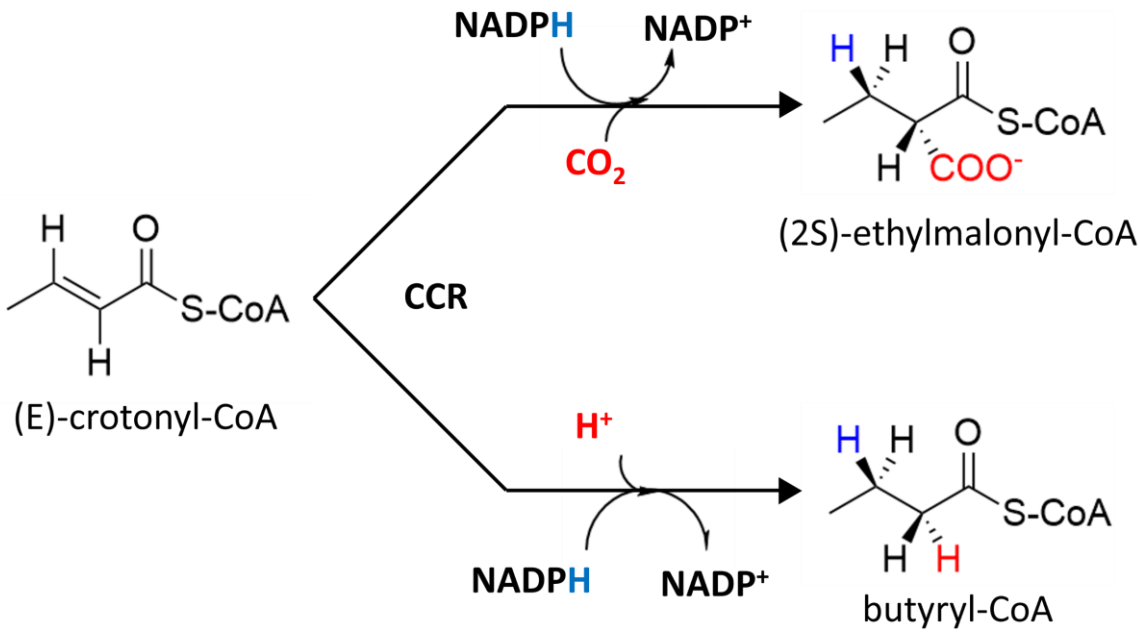


Figure 17. Mechanism of Crotonyl-CoA carboxylase/reductase (CCR). Carboxylation reaction shown on top and reduction reaction shown on bottom. Adapted from [221]

Although the CCR enzyme was originally characterised as functioning in primary metabolism, there are also many reports of CCR proteins involved in secondary metabolism. CCR genes are often co-localised with polyketide synthases, as they have been shown to produce derivatised extender units for the production of polyketides. There is a wide range of reported CCR substrates and several examples of CCR proteins with substrate promiscuity.

For example, SalG is a CCR protein with substrate promiscuity which is involved in the biosynthesis of the antitumour compounds, the salinosporamides [222]. SalG catalyses the carboxylation of 2-pentenoyl-CoA and 4-chloro-crotonyl-CoA for incorporation by the hybrid NRPS/PKS system into salinosporamide E and A respectively [223, 224]. This study also identified that the SalG CCR could carboxylate unnatural halocrotonic acids, including 4-bromo-crotonate and 4-fluoro-crotonate, for the production of halogenated salinosporamide derivatives [224]. There is also a CCR-mediated carboxylation reaction involved in the biosynthesis of the immunosuppressant tacrolimus by *Streptomyces tsukubaensis* [225]. The TcsC CCR protein carboxylates 2-pentenoyl-CoA to produce propylmalonyl-CoA, which is further modified to give rise to the allylmalonyl-CoA extender unit for incorporation by a PKS module of the hybrid NRPS/PKS system [226] (**Figure 18**).

The CCR proteins encoded in the proposed warhead biosynthesis pathway for the biosynthesis of actinonin and matlystatin are proposed to carboxylate 2-octenoyl-CoA. There is precedence in the literature for CCR proteins with this substrate specificity. This includes the PteB CCR of the filipin III PKS cluster in *Streptomyces avermilitis* [227], the TgaD CCR of the soce-thuggacin A hybrid NRPS/PKS system in *Sorangium cellulosum* [228] and the CinF CCR of the cinnabaramide A PKS system in *Streptomyces* sp. JS360 [229, 230]. Each of these natural products (**Figure 18**) has interesting bioactivities, as filipin III is a reported antifungal [231], soce-thuggacin A is a reported antibiotic [232] and cinnabaramide A has weak antitumour activity [233]. This reinforces the importance of identifying and characterising novel natural products which are biosynthesised with non-canonical polyketide extender units.

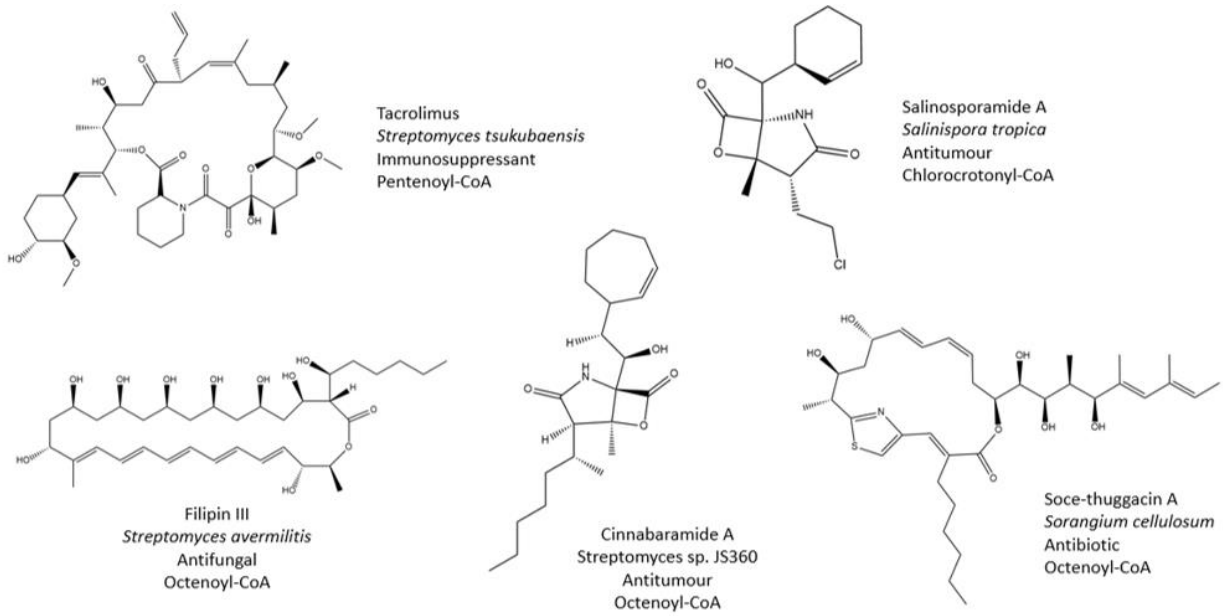


Figure 18. Structures of NPs that are biosynthesised using crotonyl-CoA carboxylase/reductase (CCR) proteins. Producer strain, bioactivity and substrate specificity of CCR is denoted.

Therefore, there are several examples of octenoyl-CoA specific CCR enzymes being utilised for the synthesis of non-canonical extender units in PKSs [234]. However, the proposed function of the actinonin and matlystatin octenoyl-CoA specific CCRs is novel as they represent the direct involvement of CCR, and by extension an EMC-like pathway, in natural product biosynthesis.

2.1.2 CCR Protein Purification and Functional Characterisation in the Literature

For this project, the actinonin and matlystatin CCR proteins needed to be expressed, purified and their activities assessed. Therefore, the reported methodology of octenoyl-CoA specific CCR protein purification and characterisation was considered to guide experimentation.

Reports of the purification of the octenoyl-CoA specific PteB and CinF CCR proteins exist in the literature, and both appear to be readily expressed and soluble using standard conditions: *E. coli* BL21(DE3), N-terminal histidine affinity tags and nickel affinity chromatography [227, 230]. Therefore, these reports indicate that the actinonin and matlystatin CCR genes should be cloned with an N-terminal hexa-histidine tag, which is unlikely to impede protein function.

As for literature reports of *in vitro* CCR activity assays, both of the above papers use LC-MS to detect the production of the carboxylated product. Consideration of the reported reaction conditions indicate that 1-2 mM of the acyl-CoA substrate, 2-4 mM NADPH, 33-80 mM bicarbonate and 50-100 mM Tris-HCl at pH 7.5 [227, 230] should be a good starting point for the optimisation of reaction conditions. To expand, the NADPH and bicarbonate are cofactors required for CCR function (**Figure 17**). Reported CCR reactions proceeded at 30 ^{o}C overnight. Separation was achieved using a reverse phase analytical column with a shallow gradient of acetonitrile or methanol against water and detection was achieved in both negative and positive mode [227, 230]. Therefore, these reports indicate a good starting point for the optimisation of CCR protein purification, *in vitro* assay conditions and product detection by LC-MS.

2.1.3 Acyl-CoA Epimerase and Acyl-CoA Mutase Activity in Secondary Metabolism

Acyl-CoA epimerase and acyl-CoA mutase proteins are categorised based on their substrate specificity. For instance, the EMC pathway is made up of an ethylmalonyl-CoA epimerase (ECE) and an ethylmalonyl-CoA mutase (ECM) (**Figure 14**) and is used by some bacteria for the assimilation of acetate in primary metabolism. The ECM mutase was determined to be homodimeric [235].

The methylmalonyl-CoA specific epimerase and mutase proteins have been more extensively studied as they are widely distributed in all organisms, except plants. In humans, these proteins are involved in the breakdown of odd chain fatty acids, cholesterol and the amino acids valine, isoleucine, threonine and methionine [236, 237]. These metabolites are broken down to propionyl-CoA, which enters the methylmalonyl-CoA pathway (**Figure 19**). Propionyl-CoA is carboxylated by propionyl-CoA carboxylase to produce 2S-methylmalonyl-CoA [238]. The stereocentre is inverted by methylmalonyl-CoA epimerase (MCE) to generate 2R-methylmalonyl-CoA. Then the methylmalonyl-CoA mutase (MCM) protein rearranges the carbon backbone to generate succinyl-CoA [239] which feeds into the tricarboxylic acid cycle of central metabolism. Deficiency of MCE or MCM in humans can lead to methylmalonic aciduria, where the body cannot break down these amino acids and as such, methylmalonic acid builds up in the blood and leads to kidney or liver damage [237, 240].

The methylmalonyl-CoA pathway (**Figure 19**) is used by some bacteria for the conversion of propionyl-CoA to succinyl-CoA, including the following genera: *Rhodospirillum*, *Propionibacterium*, *Micrococcus*, *Rhizobium* and *Mycobacterium* [241]. However, there are also sporadic occurrences of MCM being utilised for the generation of methylmalonyl-CoA extender units for polyketide biosynthesis. Depending on the strain and fermentation conditions the methylmalonyl-CoA can be derived via carboxylation of propionyl-CoA or by backbone rearrangement of succinyl-CoA (i.e. running the MCM reaction 'in reverse') [242]. For instance, *Streptomyces hygroscopicus* drives methylmalonyl-CoA biosynthesis via carboxylation of propionyl-CoA for incorporation into the immunosuppressant rapamycin [243]. However, *Saccharopolyspora erythraea* drives methylmalonyl-CoA biosynthesis via the 'reverse' methylmalonyl-CoA mutase conversion of succinyl-CoA for incorporation into the antibiotic erythromycin [244, 245].

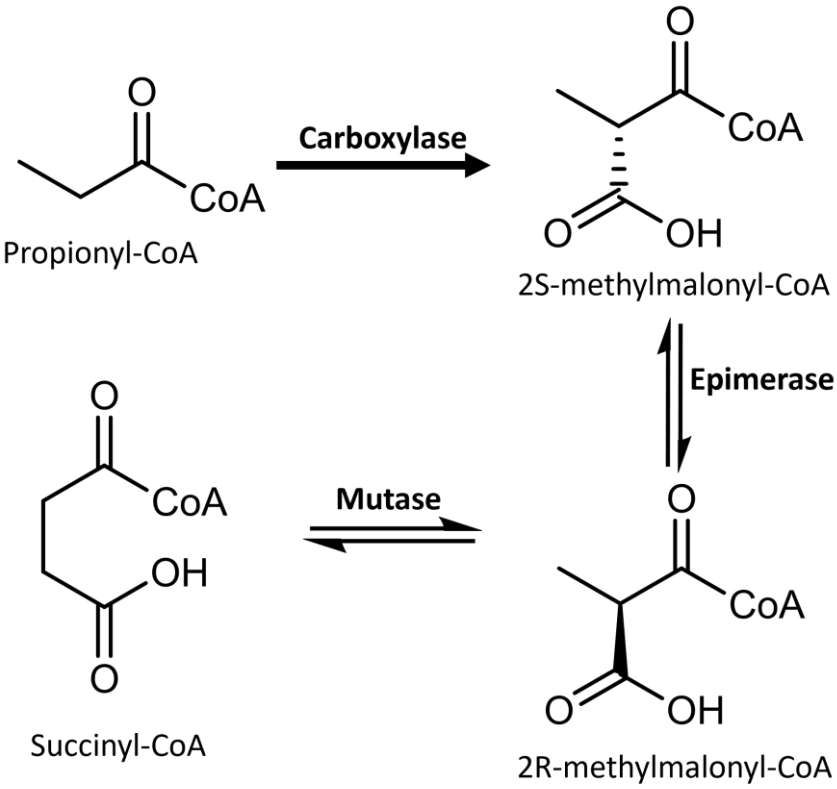


Figure 19. Methylmalonyl-CoA pathway. Occurs in humans and some bacteria.

The *Propionibacterium shermanii* MCM protein is one of the best structurally and functionally characterised mutases [246]. Alike many bacterial MCM proteins, the *P. shermanii* mutase is an αβ heterodimer which requires the coenzyme B12 cofactor for function [242]. Coenzyme B12 is a structurally complex molecule with a coordinated cobalt ion. Coenzyme B12 binds the MCM α subunit and is cleaved to generate a radical in the active site [247]. This radical facilitates the exchange of the CoA thioester group and a proton between two adjacent substrate carbon atoms in order to achieve carbon backbone rearrangement [242, 248] (**Figure 20**). In this way, MCM catalyses the backbone rearrangement between methylmalonyl-CoA and succinyl-CoA (**Figure 20**).

There are also reports of a chaperone-like protein which is involved in MCM αβ heterodimer activity. The first characterised protein was MeaB from *Methylobacterium extorquens*, which was essential for MCM activity [249]. MeaB was determined to form a complex with the MCM heterodimer [250, 251] and is proposed to prevent the cleaved coenzyme B12 remaining in the active site which would otherwise inactivate the mutase [252].

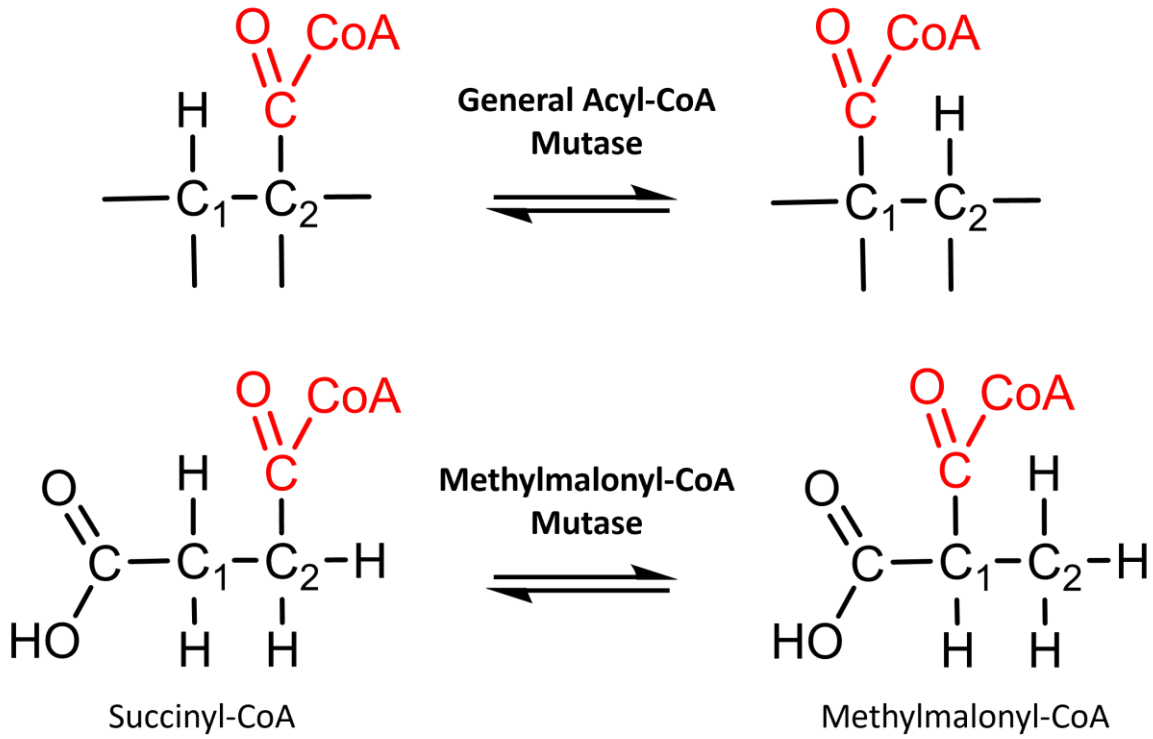


Figure 20. Schematic of general acyl-CoA mutase and methylmalonyl-CoA mutase mechanism.

2.1.4 Acyl-CoA Epimerase and Mutase Protein Purification and Functional Characterisation in the Literature

Consideration of the matlystatin BGC (**Figure 13**,**Table 1**) suggests that there are three acyl-CoA mutase genes; the mutase α subunit (*matB*), the mutase β subunit (*matQ*) and the MeaB-like protein (*matR*). As for the actinonin BGC (**Figure 13**, **Table 2**), there is a mutase α subunit (*actD*) present, however the β subunit and the MeaB-like protein are not present in the BGC. By querying the *Streptomyces* sp. NCIMB 8845 genome with the matlystatin counterparts, highly similar homologs of the mutase β subunit and the MeaB-like genes were identified outside of the actinonin BGC and their genes were named *actM* and *actN*, respectively. Both the matlystatin and actinonin BGCs contain putative acyl-CoA epimerase genes.

The actinonin and matlystatin acyl-CoA epimerase and acyl-CoA mutase genes both share ~30% and ~70% sequence similarity to their ethylmalonyl-CoA and methylmalonyl-CoA specific counterparts in bacteria, respectively. Therefore, the sequence similarity and likely quaternary structure is more supportive of the candidate actinonin and matlystatin

proteins being more similar to their methylmalonyl-CoA counterparts than the ethylmalonyl-CoA proteins. This may guide the protein purification and *in vitro* assay methodological development. However, the actinonin and matlystatin acyl-CoA epimerase and acyl-CoA mutase genes are predicted to be specific to hexylmalonyl-CoA (**Figure 15**), of which no proteins have yet been reported.

As such, a literature review was conducted to understand how the ethylmalonyl-CoA and methylmalonyl-CoA epimerase and mutase proteins have been previously expressed, purified and characterised. This should serve as a basis for the purification and characterisation of the actinonin and matlystatin hexylmalonyl-CoA epimerase and mutase proteins.

The ethylmalonyl-CoA specific epimerase and mutase proteins have been studied during investigation of the EMC pathway in *R. sphaeroides*. Protein purification proceeded using very standard methodology (produced in *E. coli* BL21(DE3) with N-terminal histidine affinity tags and purified using nickel affinity chromatography) [235]. Reported purification of the MCM protein is more complex, as the MCM is a heterodimer whereas the ECM is a homodimer. MCM protein solubility has been enhanced by co-expressing both the MCM α and β subunits in the same expression construct, each with a N-terminal histidine tag [249]. Purification of the MeaB MCM chaperone protein has also been reported with C-terminal histidine tagging [249].

Therefore, these reports indicate that N-terminal histidine tagging of the actinonin and matlystatin proposed hexylmalonyl-CoA epimerase and mutase genes and C-terminal tagging of the MeaB-like protein should not impede protein function. Further, the co-expression of the α and β subunits of the proposed hexylmalonyl-CoA mutase proteins may enhance solubility.

LC-MS is one of the major *de facto* analytical techniques for *in vitro* enzymatic characterisation. However, LC-MS analysis of acyl-CoA epimerase and mutase *in vitro* activity is limited as all proposed substrates and products (2S and 2R hexylmalonyl-CoA and pentylsuccinyl-CoA) have identical *m/z* values. Further, it is unlikely that the two stereoisomers of hexylmalonyl-CoA will be separated using standard reverse phase chromatography conditions. One method of overcoming the analytical limitations are to use radiolabelled substrates to distinguish the substrate and product [235].

However, the use of radioactive material can be avoided by performing a coupled acyl-CoA epimerase/mutase assay. Methylmalonyl-CoA epimerase activity has been determined by

using excess methylmalonyl-CoA mutase in order to drive the conversion of 2R-methylmalonyl-CoA to succinyl-CoA. Succinyl-CoA and methylmalonyl-CoA were separated using reverse phase chromatography and detected using UV absorption [253].

These reports also provide a basis for the conditions required for a coupled acyl-CoA epimerase/mutase assay. In brief, the substrate was biosynthesised by an *in situ* CCR reaction and the coupled assay was performed in the presence of coenzyme B12 in 80 mM Tris-HCl (pH 7.8). The reaction proceeded at 30^{o}C and aliquots were removed at time points and quenched with formic acid [235, 253].

To summarise, the first three steps of the proposed warhead biosynthesis pathway are planned on being ratified. This represents some novel biochemistry as specificity towards hexylmalonyl-CoA has not been reported for the acyl-CoA epimerase and mutase proteins. Further, this is the first proposal of the CCR, acyl-CoA epimerase and acyl-CoA mutase (EMC-like) proteins being directly involved in natural product biosynthesis rather than just the biosynthesis of polyketide extender units.

2.1.5 Chapter Aims

Objective 1: *In vitro* Reconstitution of Proposed Warhead Biosynthesis Pathway

I aimed to reconstitute the proposed warhead biosynthesis pathway *in vitro*. To achieve this, the specific objectives were:

i. Clone each warhead biosynthesis gene into a suitable expression vector
ii. Heterologously express the genes and optimise protein purification
iii. Synthesise and purify the proposed initial substrate of the pathway, octenoyl-CoA
iv. Perform and analyse *in vitro* reactions between purified proteins and substrate(s)

Objective 2: Phylogenetic Analysis of the Acyl-CoA Mutase Protein

The acyl-CoA mutase genes of the proposed warhead biosynthesis pathway are interesting as they are theorised to encode proteins with specificity towards hexylmalonyl-CoA, which has not been described previously. Therefore, these genes may represent effective 'probes' for the identification of the warhead biosynthesis pathway in other strains. A bioinformatic analysis of actinomycete acyl-CoA mutase genes was carried out with the aim of identifying novel warhead-NP BGCs. To achieve this, the specific objectives were:

i. Build a phylogenetic tree of all actinomycete acyl-CoA mutase gene homologs
ii. Perform a co-association analysis to determine whether the proposed warhead biosynthesis genes and other biosynthetic machinery genes are present in close proximity to the acyl-CoA mutase gene
iii. Assess whether this analysis highlights any candidate novel warhead-NP BGCs

2.2 Results

2.2.1 Purification of Putative Warhead Biosynthesis Proteins

The putative warhead biosynthesis genes of both the actinonin and matlystatin BGCs were cloned into pET28a. The pET28a expression plasmid was selected as it supports histidine tagging at either the N- or C- termini of proteins. Previous reports of the purification of proteins related to the warhead biosynthesis proteins were considered to determine whether N-terminal or C-terminal was more likely to result in soluble protein. The matlystatin and actinonin warhead biosynthesis genes were amplified from *A. atramentaria* DSM 43919 genomic DNA (gDNA) and *Streptomyces* sp. ATCC 14903 gDNA, respectively, and purified. These inserts were digested and ligated into digested pET28a backbone to generate the expression constructs outlined in **Table 3** (pJF01-14). Note that the *matQ* and *matR* genes were unable to be amplified from gDNA and therefore these inserts were codon optimised for *E. coli* expression and synthesised for construct generation. The sequences of all constructs were confirmed by colony PCR and by sequencing the insert.

Table 3. Plasmid details of actinonin and matlystatin putative warhead biosynthesis genes. Each were cloned into pET28a. His tag terminality and plasmid names are reported. An asterisk denotes genes which are not present in the biosynthetic gene cluster, however are present elsewhere in the genome. A 'CO' gene prefix denotes that the gene insert was codon optimised and synthesised. Abbreviations: CCR (crotonyl-CoA carboxylase/reductase), Asp Syn (asparagine synthetase), N-oxy (N-oxygenase).

		Gene	Accession Number	Size (Kb)	Size (kDa)	His Tag	Name
Matlystatin	CCR	*matL*	WP_019634553.1	1.34	47.7	N-	pJF01
	Epimerase	*matE*	WP_026341874.1	0.42	15.2	N-	pJF02
	Mutase (α)	*matB*	WP_019634562.1	1.44	52.1	N-	pJF03
	Mutase (β)	*COmatQ*	WP_019634548.1	0.41	14.2	N-	pJF04
	MeaB	*COmatR*	WP_084693477.1	1.21	42.6	C-	pJF05
	Asp Syn	*matC*	WP_019634561.1	1.85	68.7	C-	pJF06
	N-oxy	*matD*	WP_157408083.1	1.27	47.6	C-	pJF07
Actinonin	CCR	*actF*	WP_086815437.1	1.32	47.8	N-	pJF08
	Epimerase	*actC*	WP_030890453.1	0.45	16.6	N-	pJF09
	Mutase (α)	*actD*	WP_086815435.1	1.44	52.1	N-	pJF10
	Mutase (β)	*actM**	WP_030874632.1	0.41	14.3	N-	pJF11
	MeaB	*actN**	WP_030885530.1	0.98	33.4	C-	pJF12
	Asp Syn	*actJ*	WP_030878155.1	1.85	68.8	C-	pJF13
	N-oxy	*actI*	WP_078872911.1	1.03	38.1	C-	pJF14

Plasmids pJF01-14 were transformed into *E. coli* expression strains and subject to a standard protein purification protocol to gauge the titre, solubility and purity of resultant protein preparations. Four *E. coli* expression strains were tested, including BL21(DE3) and its derivatives; SoluBL21, Rosetta and NiCo21. These strains are reported to offer enhanced protein solubility and purity in a number of ways; SoluBL21 has uncharacterised mutations which aid protein folding, Rosetta encodes tRNAs for rare codons and NiCo21 features four modified proteins which bind to a chitin pre-treatment column, which would otherwise bind the nickel column. In brief, *E. coli* expression strain cultures were induced with IPTG and grown at 30 ^{o}C for four hours. The cells were pelleted, lysed and centrifuged. The clarified supernatant was subject to nickel affinity chromatography and fractions analysed by SDS-PAGE. See materials and methods for details.

Table 4. Purification of warhead biosynthesis proteins using different *E. coli* expression strains. A cross mark indicates that there was no expression of a correctly sized protein. A tick mark indicates that there was expression of a correctly sized protein. A tick mark in the mass spectrometry (MS) column indicates that the protein was confirmed to be expressed by tryptic digest and MS.

		Gene	Plasmid	BL21	SoluBL21	Rosetta	NiCo21	MS?
Matlystatin	CCR	*matL*	pJF01		X		X	
	Epimerase	*matE*	pJF02				✓	
	Mutase (α)	*matB*	pJF03		X		X	
	Mutase (β)	*COmatQ*	pJF04				X	
	MeaB	*COmatR*	pJF05				X	
	Asp Syn	*matC*	pJF06		X			
	N-oxy	*matD*	pJF07		X			
Actinonin	CCR	*actF*	pJF08	✓			✓	✓
	Epimerase	*actC*	pJF09	✓				✓
	Mutase (α)	*actD*	pJF10	✓	X	X	✓	✓
	Mutase (β)	*actM**	pJF11	✓			✓	✓
	MeaB	*actN**	pJF12	✓	X	X	X	✓
	Asp Syn	*actJ*	pJF13					
	N-oxy	*actI*	pJF14					

Table 4 summaries the results of over twenty preliminary protein purification attempts of the warhead biosynthesis proteins using different *E. coli* expression strains. A tick indicates that a protein of the expected size was observed by SDS-PAGE analysis using a specific expression strain. The actinonin warhead biosynthesis proteins proved to be more readily soluble and more effectively purified using standard methodology than their matlystatin counterparts (**Table 4**). As the warhead biosynthesis genes of the actinonin and matlystatin clusters are highly similar, only one 'set' of proteins are likely to be required for the *in vitro*

reconstitution of the pathway. Therefore, purification of the actinonin proteins was prioritised.

There were considerable differences in the resultant solubility, titre and purity of proteins between different *E. coli* expression strains, as exemplified by the purification of ActD (**Appendix Figure 95**). Overall, use of the *E. coli* NiCo21 expression strain and a chitin pre-treatment of the clarified lysate resulted in the most effective and consistent nickel affinity chromatography purification results. Therefore, this methodology served as the baseline for further optimisation.

Figure 21 demonstrates the purity of actinonin warhead biosynthesis proteins afforded by this methodology. Each of the bands indicated by a red arrow were excised, analysed by MS and their identity confirmed as the target protein. The ActF (CCR), ActC (acyl-CoA epimerase) and ActM (acyl-CoA mutase β subunit) proteins were expressed in large quantities and were readily soluble. Although the protein preparations were not pure, they contained soluble protein and thus should be suitable for *in vitro* assays. However, the ActD (acyl-CoA mutase α subunit) protein was very poorly expressed and/or soluble. Further, the ActN (MeaB-like) protein was not expressed using the *E. coli* NiCo21 strain. Expression of ActN using *E. coli* BL21(DE3) resulted in the purification of a small amount of ActN (as confirmed by MS) but there was a huge amount of contaminating protein (**Appendix Figure 96**) that rendered the sample unsuitable for *in vitro* assays. Therefore, further optimisation of ActD and ActN purification was required in order to generate the αβ mutase heterodimer for *in vitro* characterisation.

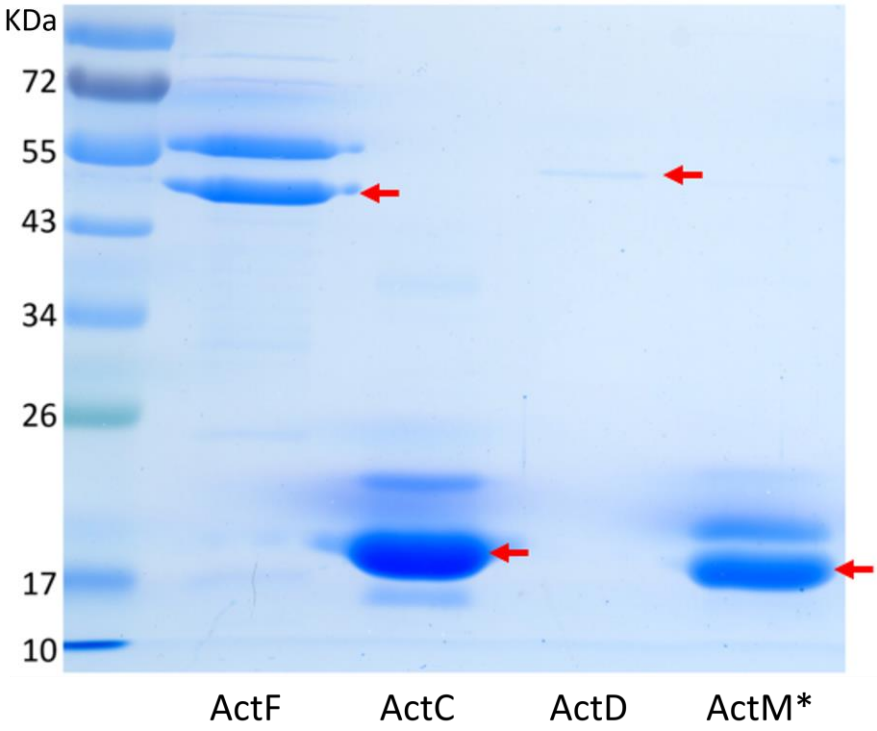


Figure 21. SDS-PAGE analysis of purification of actinonin warhead biosynthesis proteins. *E. coli* NiCo21 strains, containing pJF08, pJF09, pJF10, pJF11 plasmids were purified by chitin and nickel affinity chromatography to purify ActF (CCR), ActC (epimerase), ActD (mutase α) and ActM* (mutase β), respectively. Masses of ladder proteins are shown in kDa. Expected protein masses (kDa) are as follows; ActF (47.8), ActC (16.6), ActD (52.1), ActM* (14.3). Arrows denote that the band was analysed by mass spectrometry. All were confirmed to be the expected protein.

2.2.2 Optimisation of ActD Purification

There are reports in the literature that co-expression of the acyl-CoA mutase α and β subunits enhanced protein solubility and that the MeaB chaperone is essential for activity (see section **2.1.4**). Therefore, constructs were produced for the dual expression of *actD* and *actM*, along with compatibility for the tri-expression of the MeaB-like mutase chaperone *actN*. pACYCDuet-1 was selected as a suitable co-expression plasmid as it features a different origin of replication and antibiotic resistance cassette, as compared to pET28a, and therefore both of these plasmids should be stable in a single expression strain. The pJF15 plasmid was generated by cloning the *actD* and *actM* genes into pACYCDuet-1, to produce both proteins with N-terminal histidine tags.

The co-expression *E. coli* NiCo21 strains were prepared as follows. The di-expression (acyl-CoA mutase α/β) strain was transformed with the pJF15 plasmid, and the tri-expression (acyl-CoA mutase α/β and MeaB-like) was transformed with the pJF15 and pJF12 plasmids.

Figure 22. SDS-PAGE analysis of protein purification of acyl-CoA mutase heterodimer, by di- and tri-expression systems. *E. coli* NiCo21 strains were purified by chitin and nickel affinity chromatography. The di-expression system consists of pJF15, which encodes *actD* (mutase α) and *actM** (mutase β). The tri-expression system consists of pJF15 and pJF12, which encodes *actN* (MeaB). An empty vector (EV) control purification is shown. Masses of ladder proteins are shown, in kDa. Protein masses (kDa) are as follows; ActD (52.1), ActM (14.3), ActN (33.4). Arrows denote that the band was analysed by mass spectrometry. The middle band was confirmed to be ActN.

The chitin and nickel affinity chromatography purification proceeded as above, and the resultant protein preparations were analysed by SDS-PAGE (**Figure 22**). The co-expression of *actD* and *actM* did not result in any enhanced solubility of the acyl-CoA mutase subunits, as compared to the protein preparation of an empty vector (EV) negative control strain. Whereas the tri-expression of *actD*, *actM* and *actN* resulted in three novel protein bands in the protein preparation, of approximately the expected sizes. This was suggestive that the co-expression of the MeaB-like chaperone protein (ActN) enhanced the solubility and/or stability of the ActD/ActM acyl-CoA mutase heterodimer. The indicated bands were

excised and analysed by MS and the middle band was confirmed as ActN, although the identity of the remaining bands was inconclusive.

In summary, the optimisation of putative warhead biosynthesis protein purification resulted in soluble CCR and acyl-CoA epimerase proteins from the actinonin pathway (ActF and ActC, respectively), which were confirmed by MS (**Figure 21**), along with some promising acyl-CoA mutase protein preparations. Although these protein preparations were impure, the proteins were soluble and therefore likely to be folded correctly and functionally active. Therefore, the synthesis of the proposed starting substrate of the warhead biosynthesis pathway, octenoyl-CoA, was focussed upon.

2.2.3 Octenoyl-CoA Synthesis

The proposed starting substrate of the warhead biosynthesis pathway is 2E-octenoyl-CoA **Figure 15**, [207]. 2E-octenoyl-CoA is an α,β-unsaturated acyl-CoA thioester (**Figure 23**), effectively consisting of the C8 unsaturated fatty acid 2-octenoic acid linked to CoA via a thioester bond. 2E-octenoyl-CoA was not commercially available and therefore required synthesising.

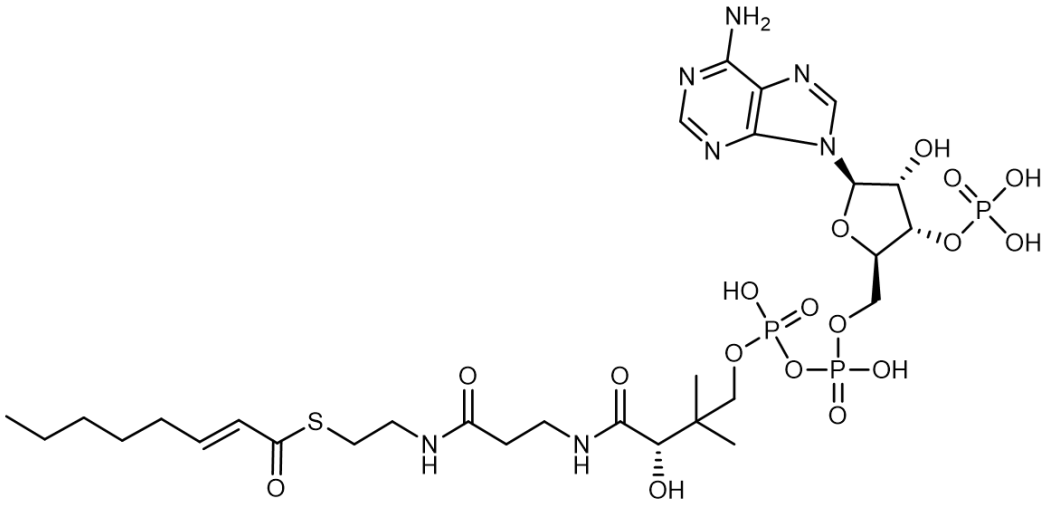


Figure 23. Structure of 2E-octenoyl-CoA.

There are two major synthetic routes to enoyl-CoA compounds reported in the literature. The first of which is the mixed anhydride method which was developed in the 1960s [254, 255] and has remained an effective method for the synthesis of enoyl-CoA thioesters [256]. In this method, the 2-octenoic acid is converted into a carboxylic anhydride (**1**, **Figure 24**) by ethyl chloroformate. In the second reaction step, free coenzyme A is reacted with the

carboxylic anhydride. This second reaction limits the overall yield as it may yield the desired 2E-octenoyl-CoA or the propoxycarbonyl-CoA side product (**2**, **Figure 24**). One report indicated that the yield of octenoyl-CoA via the mixed anhydride method is 57% [257]. Purification of a related enoyl-CoA from the crude reaction mixture has been reported using reverse phase C18 chromatography [256].

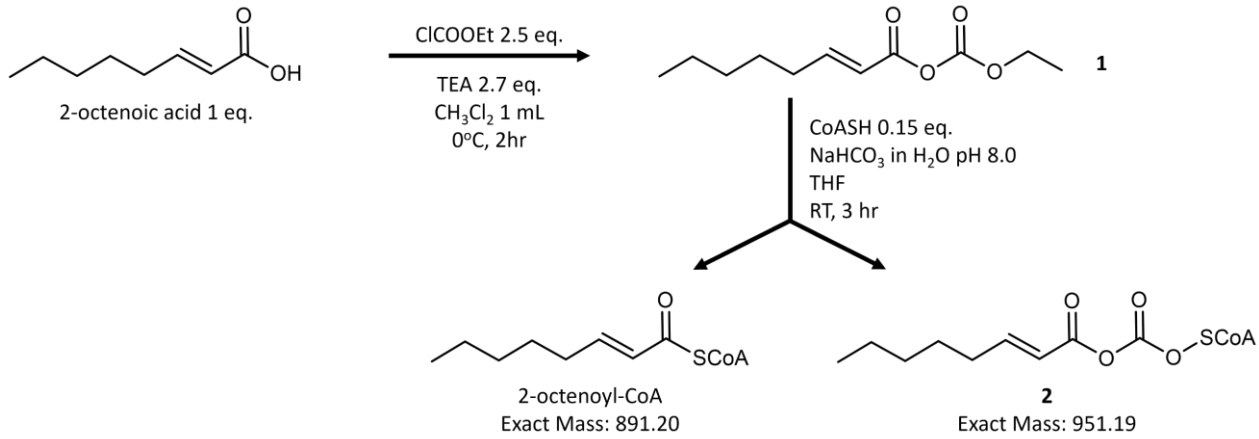


Figure 24. Schematic of 2E-octenoyl-CoA synthesis by mixed anhydride method. Abbreviations: TEA = triethylamine, THF = tetrahydrofuran, CoASH = Coenzyme A, RT = room temperature.

Octenoyl-CoA was initially synthesised using the mixed anhydride method (**Figure 24**). 2-octenoic acid was dissolved in dichloromethane and reacted with ethyl chloroformate in the presence of triethylamine for 2 hours at 0 °C. The solvent was removed by rotary evaporation and the mixed anhydride dissolved in tetrahydrofuran. This was then reacted with free coenzyme A in aqueous solution for 3 hours at room temperature. The volume was reduced by rotary evaporation and the crude reaction mixture was separated by reverse phase flash chromatography. Fractions were analysed by LC-MS and those containing octenoyl-CoA were combined and dried. This was further purified by semi-preparative scale reverse phase HPLC.

This yielded a 13.7 mg sample of an orange oil, which was analysed by LC-MS (**Figure 25**). The major purified product proved to be octenoyl-CoA, as shown by the purple extracted ion chromatogram in **Figure 25**. However, the sample also contained the propoxycarbonyl-CoA side product (observed $[M-H]^-$ = 950.15, **2**, **Figure 24**) as denoted by the asterisk in **Figure 25**. The estimated purity of the 13.7 mg octenoyl-CoA sample is 70%, which represents a yield of 9%. This is considerably lower than the 57% yield previously reports for this method [257].

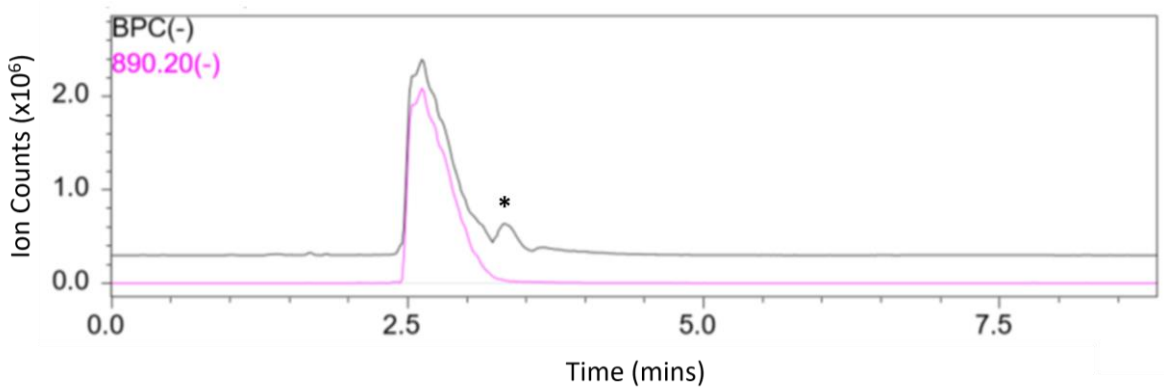


Figure 25. LC-MS analysis of octenoyl-CoA sample synthesised using mixed anhydride method. Asterisk denotes the major contaminant. Purple spectrum is the extracted ion chromatogram of octenoyl-CoA $[M-H]^-$ = 890.20).

Octenoyl-CoA proved to be relatively unstable as it was prone to degradation over several months. This necessitated several separate syntheses and purifications of the starting substrate. The PyBOP method was used for the second octenoyl-CoA synthesis as it provides a reduced reaction duration and less harsh solvents and reagents, as compared to the mixed anhydride method, and therefore may provide a greater yield due to a reduced rate of octenoyl-CoA degradation.

PyBOP (benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate) is a coupling reagent developed in the 1990s for peptide synthesis [258, 259]. The PyBOP coupling reaction is proposed to proceed via the mechanism illustrated in **Figure 26** [260, 261]. The PyBOP method has been used for the synthesis of enoyl-CoA compounds, including 2E-hexenoyl-CoA (C6) [262] and decadienoyl-CoA (C10) [263], which have been purified form the crude reaction mixture by reverse phase chromatography [262, 263].

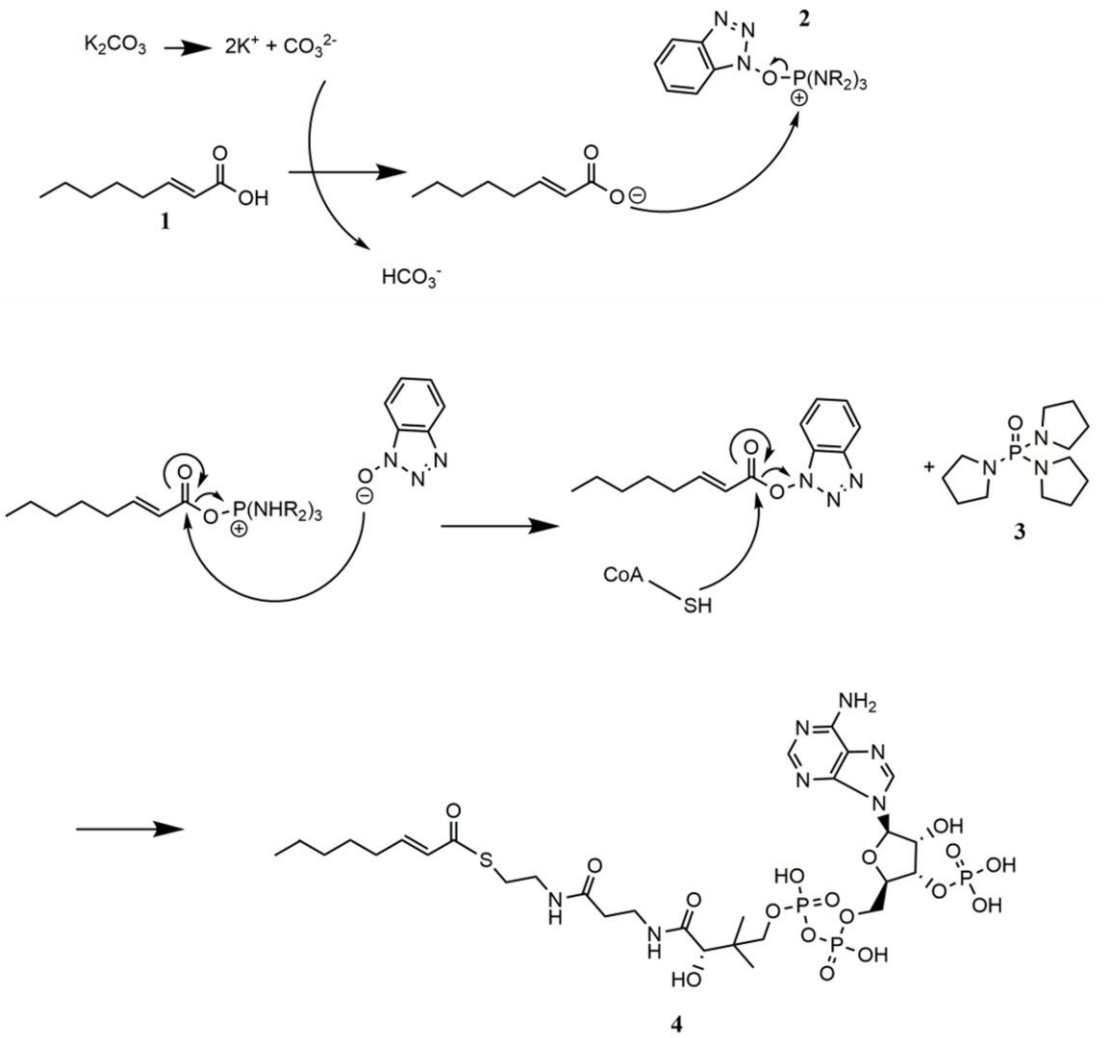


Figure 26. Schematic of 2E-octenoyl-CoA synthesis by PyBOP method. 1 = 2-octenoic acid; 2 = PyBOP (benzotriazol-1-yl-oxytripyrrolidinophosphonium); 3 = tris(pyrrolidino)phosphine oxide; 4 = octenoyl-CoA.

In brief, 2-octenoic acid, free CoA, PyBOP and K_2CO_3 were reacted in THF/H_2O at room temperature for 2.5 hours (see materials and methods for details). The crude reaction mixture was concentrated and analysed by LC-MS (**Appendix Figure 97**), which indicated that all CoA had been converted to octenoyl-CoA, however there was a large quantity of the PyBOP breakdown product contaminant tris(pyrrolidino)phosphine oxide (**3**, **Figure 26**).

Strong anion exchange (SAX) chromatography was tested for the separation of the PyBOP octenoyl-CoA crude reaction mixture but provided no benefit over reverse phase chromatography (**Appendix Figure 98**).

This methodological optimisation led to the final octenoyl-CoA synthesis, whereby the PyBOP reaction was performed and yielded 100 mg of crude reaction mixture. This was separated by reverse phase flash chromatography, to yield 39 mg of material (**Appendix Figure 99**). This material was further separated by preparative-scale reverse phase HPLC (**Figure 27**), which yielded an 8.9 mg sample of fluffy white material. HR-MS analysis (**Appendix Figure 100**) indicated that this sample was predominantly octenoyl CoA (observed $[M-H]^-$ = 890.1968, Δ0.1 ppm). Dr. Martin Rejzek (John Innes Centre) performed ^1H NMR analysis on this sample (**Appendix Figure 101**), which indicated the presence of octenoyl-CoA at a purity of approximately 82%. This represents a yield of only 6.3%, however is a sufficient amount of octenoyl-CoA to perform several thousand *in vitro* assay reactions.

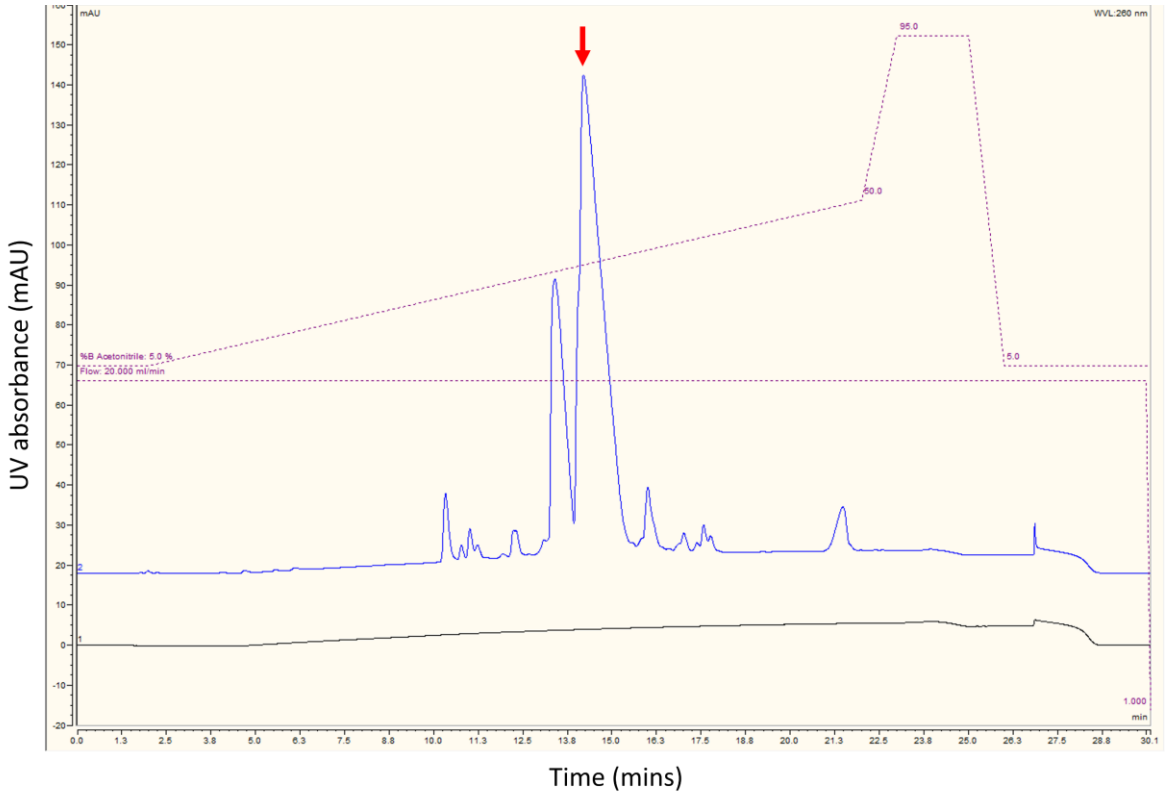


Figure 27. Preparative scale reverse phase HPLC purification of flash chromatography octenoyl-CoA sample. Detection at 260 nm. Black line is blank injection and blue line is sample injection. The red arrow denotes which peak is octenoyl-CoA, as determined by HR-MS analysis.

2.2.4 CCR *in vitro* assay

The first reaction of the proposed warhead biosynthesis pathway is the carboxylation of octenoyl-CoA by the CCR protein (ActF) to produce 2S-hexylmalonyl-CoA (**Figure 28**).

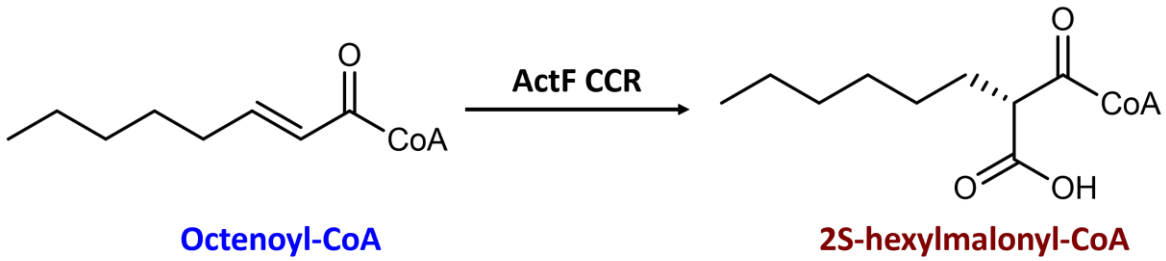


Figure 28. First reaction of the proposed warhead biosynthesis pathway.

ActF was reacted with 2 mM octenoyl-CoA in the presence of 4 mM NADPH and 33 mM $NaHCO_3$ at pH 8.0 at 30°C for 30 minutes. Negative control samples were prepared using denatured protein. The reactions were quenched and analysed by LC-MS (**Figure 29**). The detection of masses equal to hexylmalonyl-CoA were only present in the sample with native ActF. Therefore, this was the first indication that ActF was able to carboxylate octenoyl-CoA to hexylmalonyl-CoA.

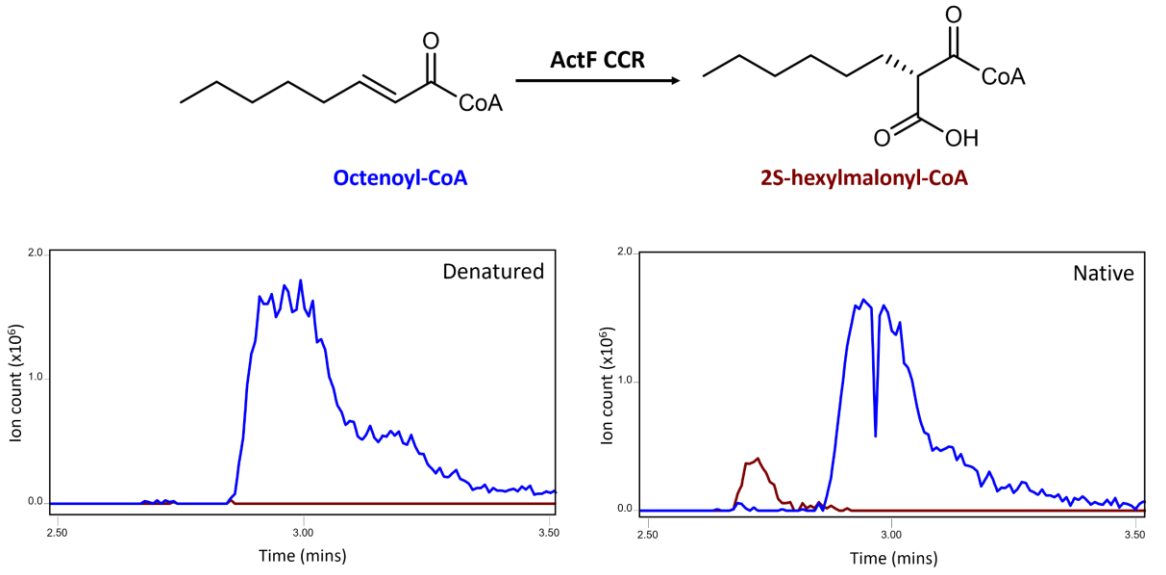


Figure 29. LC-MS analysis of Octenoyl-CoA / ActF CCR *in vitro* assay. *In vitro* reactions between ActF and Octenoyl-CoA. Left panel is negative control, using denatured ActF, and right panel is using native ActF. The extracted ion chromatogram of octenoyl-CoA (*m/z* 892.21; $[M+H]^+$) is plotted in blue. The extracted ion chromatogram of hexylmalonyl-CoA (*m/z* 938.22; $[M+H]^+$), the product of the reaction, is plotted in brown.

The substrate specificity of ActF was also tested. In the literature, there are reports of octenoyl-CoA specific CCR proteins which are also able to carboxylate the C4 crotonyl-CoA to ethylmalonyl-CoA [230] (the first step of the EMC pathway, **Figure 14**). Crotonyl-CoA is commercially available and was used as the substrate in another *in vitro* ActF assay using the same reaction conditions and analysis methodology as above. The detection of masses equal to ethylmalonyl-CoA were only present in the sample with native ActF protein (**Figure 30**). Therefore, this experiment indicated that ActF also has the ability to carboxylate crotonyl-CoA.

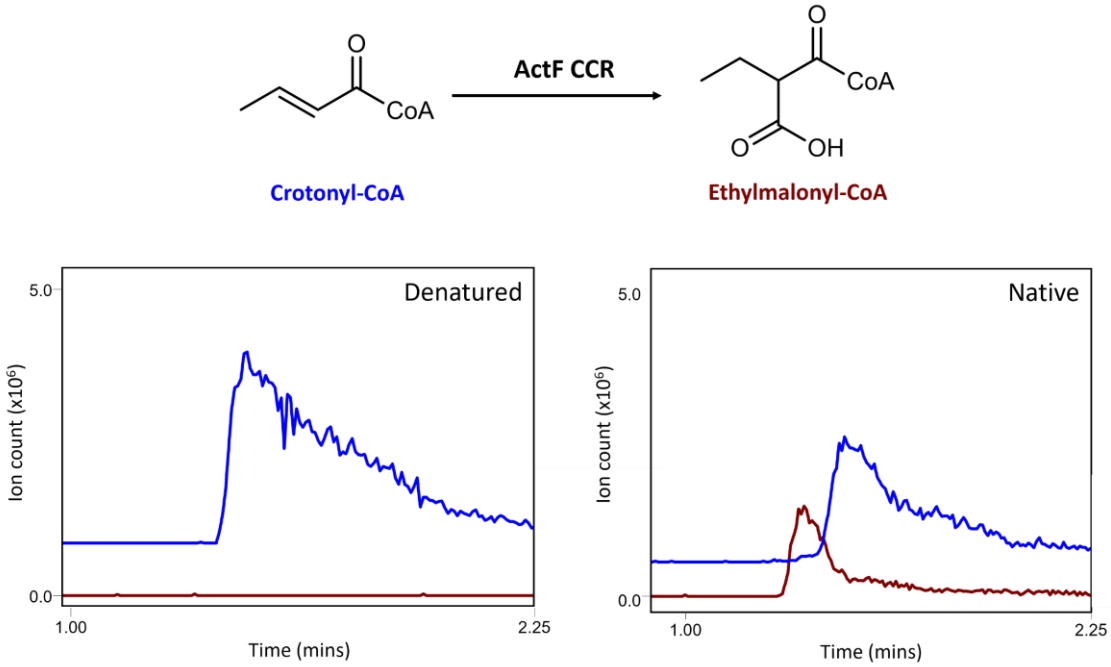


Figure 30. LC-MS analysis of Crotonyl-CoA / ActF CCR in vitro assay. *In vitro* reaction between ActF and crotonyl-CoA. Left panel is negative control, using denatured ActF, and right panel is using native ActF. The extracted ion chromatogram of crotonyl-CoA (*m/z* 836.14; $[M+H]^+$) is plotted in blue. The extracted ion chromatogram of ethylmalonyl-CoA (*m/z* 882.15; $[M+H]^+$), the product of the reaction, is plotted in brown.

Throughout the purification of octenoyl-CoA and the analysis of *in vitro* assay reactions, attempts were made to optimise LC conditions to reduce the broadness of the enoyl-CoA peaks. Different concentrations of formic acid and trifluoracetic acid were added to mobile phases in an attempt to ensure that all analytes were equally protonated. Further, different LC gradients, columns (including C18 polar and HILIC), sample concentrations and MS instruments were tested. However, there was little improvement in enoyl-CoA peak shape and the long tail was always observed.

I attempted to optimise the CCR turnover efficiency using carbonic anhydrase, which is a commercially available enzyme that catalyses the interconversion of bicarbonate and free carbon dioxide [264]. As the carboxylation species of CCR has been determined to be carbon dioxide rather than bicarbonate [220], the addition of 10 µg/mL carbonic anhydrase is reported to enhance the CCR turnover rate [220]. In the case of ActF, the addition of carbonic anhydrase to the crotonyl-CoA / CCR *in vitro* assay enhanced turnover by 260%, as determined by comparison of extracted ion chromatogram peak areas (**Figure 31**). Therefore, carbonic anhydrase was added to all further CCR *in vitro* assays.

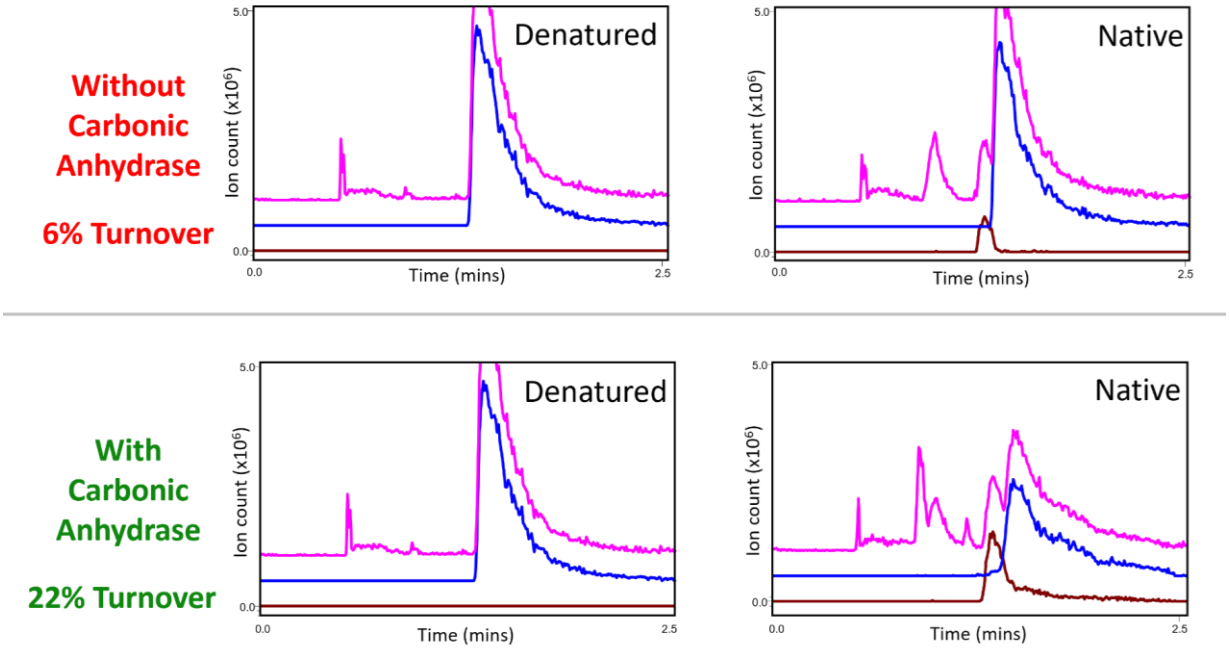


Figure 31. LC-MS analysis of effect of carbonic anhydrase (CA) to ActF CCR *in vitro* assay substrate turnover. *In vitro* reaction between ActF and crotonyl-CoA, with and without CA. Left panels are negative control, using denatured ActF, and right panels are using native ActF. The base peak chromatogram is plotted in pink. The extracted ion chromatogram of crotonyl-CoA (*m/z* 836.14; $[M+H]^+$) is plotted in blue. The extracted ion chromatogram of ethylmalonyl-CoA (*m/z* 882.15; $[M+H]^+$), the product of the reaction, is plotted in brown. The turn-over of substrate (in negative control) to product (with native enzyme) was calculated from the peak areas.

2.2.5 Development of a More Robust CCR *in vitro* assay

The above ActF assays indicate that a compound of equal mass to the expected carboxylated product was produced. A more advanced CCR assay was developed in order to more accurately characterise the CCR activity. Further assays were performed in the presence of stable isotopically labelled sodium bicarbonate ($NaH^{13}CO_3$) which dissociated into labelled carbon dioxide ($^{13}CO_2$) in the aqueous *in vitro* assay conditions. LC-MS analysis of the CCR carboxylation of crotonyl-CoA in the presence of $^{13}CO_2$ indicated that a

compound of the expected mass of ^{13}C-labelled ethylmalonyl-CoA was produced (**Figure 32**). The incorporation of ^{13}C was confirmed by LC-MS analysis of the isotopic patterning of the carboxylated product – enrichment of the [M+1] isotope was only present when the reaction was performed in the presence of $NaH^{13}CO_3$ (**Figure 33**). These data confirmed that ActF was carboxylating crotonyl-CoA. This method may serve as a more robust *in vitro* assay for characterising ActF activity.

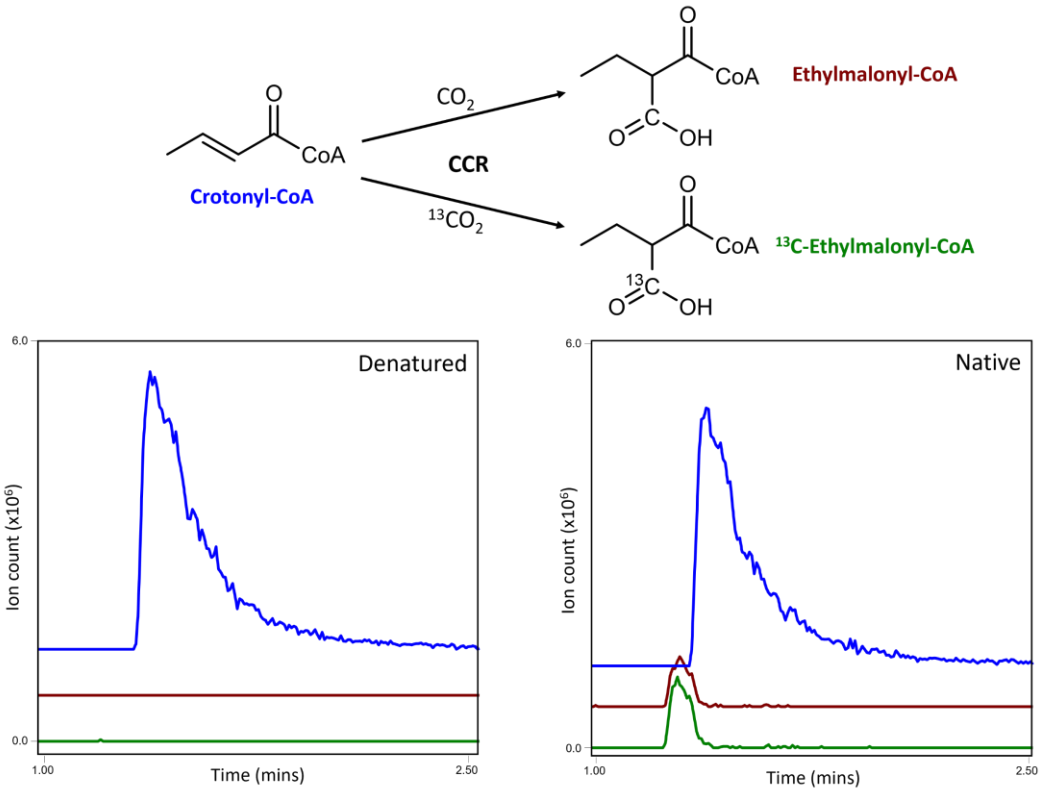


Figure 32. LC-MS analysis of ActF assay using stable isotope labelled carbon. *In vitro* reaction between ActF and crotonyl-CoA, in presence of labelled $^{13}CO_2$. Overall schematic (top) demonstrates predicted incorporation of ^{13}C by the CCR enzyme. Bottom panels are LC-MS results of *in vitro* assays; left panel is negative control, using denatured ActF, and right panel is using native ActF. The extracted ion chromatogram (EIC) of crotonyl-CoA (*m/z* 836.14; $[M+H]^+$) is plotted in blue. The EIC of ethylmalonyl-CoA (*m/z* 882.15; $[M+H]^+$) is plotted in brown. The EIC of ^{13}C-ethylmalonyl-CoA (*m/z* 883.15; $[M+H]^+$) is plotted in green.

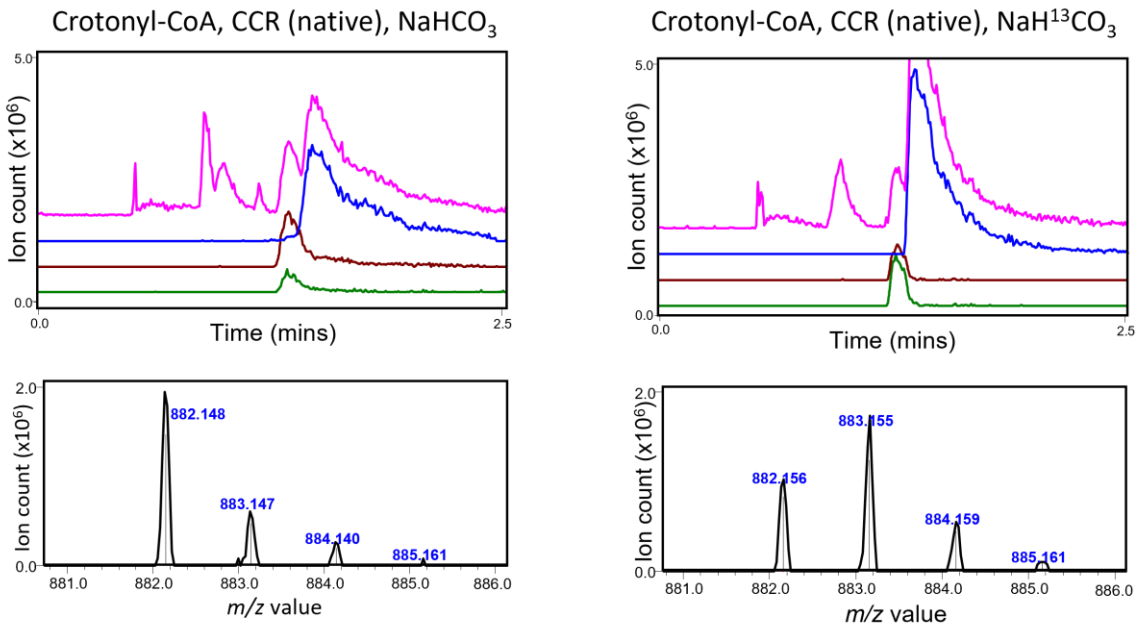


Figure 33. LC-MS analysis of ActF assay - isotopic patterning of carboxylated products. *In vitro* reactions between ActF and crotonyl-CoA. Reactions took place in $NaH^{12}CO_3$ aqueous buffer (left panel) or $NaH^{13}CO_3$ aqueous buffer (right panel). LC-MS analysis of *in vitro* assays (top) show base peak chromatogram plotted in purple. The extracted ion chromatogram (EIC) of crotonyl-CoA (*m/z* 836.14; $[M+H]^+$) is plotted in blue. The EIC of ethylmalonyl-CoA (*m/z* 882.15; $[M+H]^+$) is plotted in brown. The EIC of ^{13}C-ethylmalonyl-CoA (*m/z* 883.15; $[M+H]^+$) is plotted in green. For each *in vitro* assay condition, the range of detected *m/z* ratios associated with the carboxylated product is shown (bottom), demonstrating the isotopic patterning.

2.2.6 Scale-up of *in vitro* assays

The above *in vitro* assays had demonstrated that ActF was able to carboxylate both crotonyl-CoA and the proposed starting substrate of the warhead biosynthesis pathway, octenoyl-CoA. However, in order to progress with the *in vitro* reconstitution of the following proposed warhead biosynthesis reactions, the availability of substrate had to be addressed. As each subsequent warhead biosynthesis reaction uses the prior product as the substrate, there are two options for substrate generation.

Firstly, a 'one-pot' reaction could be performed whereby the initial substrate, octenoyl-CoA, and a number of preceding enzymes are provided for *in situ* substrate generation. *In vitro* reactions were performed with octenoyl-CoA in the presence of ActF, ActC and the purified preparation from the ActD, actM and ActN tri-expression strain. However, LC-MS analysis indicated that the 2-pentylsuccinyl-CoA product was not produced, as there were no peaks with new retention times.

Alternatively, each substrate could be biosynthesised or synthesised and purified. This latter approach has the benefit of more precisely controlling the amount of substrate provided and using more controlled *in vitro* reaction conditions. However, the losses involved with multiple rounds of chromatographic purification would likely necessitate the scaling-up of reactions.

The utility of using the proposed natural starting substrate, octenoyl-CoA, is limited. Octenoyl-CoA readily degraded in aqueous solution and was not stable when lyophilised and stored at -20 ^{o}C. Further, free CoA is expensive and therefore limited the synthesis scale. Therefore, experiments were conducted to test the feasibility of enhancing *in situ* substrate generation and also whether enoyl-CoA analogues could be synthesised and used.

2.2.6.1 Optimisation of enzymatic substrate generation

2.2.6.1.1 In situ octenoyl-CoA generation by fatty acyl-CoA ligase protein

The fatty acyl-CoA ligase (FACL) proteins are a class of enzymes which catalyse the formation of acyl-CoA thioesters from the fatty acid and CoA. The FACL mechanism occurs in two steps: firstly, the fatty acid is adenylated by ATP and then the acyl-AMP intermediate is reacted with CoA to generate the acyl-CoA thioester [265, 266]. The FACL proteins are categorised based on their substrate specificity. The medium chain FACLs (MC-FACLs) catalyse the formation of C4-C12 acyl-CoA thioesters. Therefore, MC-FACLs presented as a potential solution for the *in situ* biosynthesis of octenoyl-CoA.

The reveromycin BGC of *Streptomyces* sp. SN-593 has been reported to encode a MC-FACL protein, RevS [267]. RevS has been purified and determined to be able to effectively ligate 2-octenoic acid to generate octenoyl-CoA [267]. The RevS protein was purified in line with this existing methodology. In brief, the RevS gene was codon optimised, synthesised and cloned into pET28a. This expression construct was transformed into the *E. coli* NiCo21 strain, and the protein purified by chitin column pre-treatment followed by nickel affinity chromatography. However, the RevS protein was inconsistently soluble between different protein purifications. Further, LC-MS analysis of *in vitro* octenoic acid / CoA ligase reactions with purified soluble RevS protein never indicated that octenoyl-CoA was produced.

2.2.6.1.2 Optimisation of CCR turnover efficiency

Optimisation of the CCR turnover efficiency was explored as a method of enhancing substrate availability for later warhead biosynthesis reactions. The addition of carbonic anhydrase had produced the greatest turnover, of 22%, in a crotonyl-CoA / ActF reaction. However, the literature was searched for octenoyl-CoA specific CCR proteins that may offer enhanced turnover efficiency. The CinF CCR protein was identified in the *Streptomyces* sp. JS360 cinnabaramide BGC [229] and characterised as having specificity towards the carboxylation of 2-octenoyl-CoA [230]. Although CinF appears to have a similar turnover efficiency to the actinonin CCR protein, the CinF *in vitro* assays were not supplemented with carbonic anhydrase [230], which had resulted in the tripling of turnover efficiency for the ActF *in vitro* assay. Therefore, the *cinF* gene was codon optimised, synthesised and cloned into pET28a. This expression construct was transformed into the *E. coli* NiCo21 strain, and the protein purified by chitin column pre-treatment and nickel affinity chromatography. SDS-PAGE analysis indicated that the CinF protein was highly soluble and purified in considerable amounts (**Figure 34**).

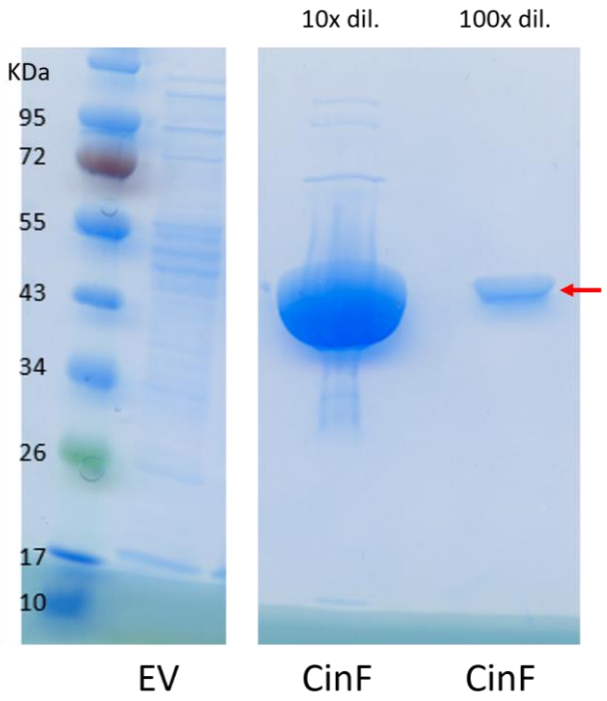


Figure 34. SDS-PAGE analysis of purification of CinF protein. *E. coli* NiCo21 (pET28a::*cinF*) lysate supernatant was purified by chitin and nickel affinity chromatography. The purified protein was analysed at 10x and 100x dilution. An empty vector (EV) control purification is shown. Masses of ladder proteins are shown, in kDa. The mass of CinF is 48.1 kDa. The arrow denotes that the band was analysed by mass spectrometry. The protein was confirmed to be CinF.

However, no CinF activity was detected in *in vitro* assays using either octenoyl-CoA or crotonyl-CoA substrates using identical conditions to the literature. This is contradictory to the literature reports of CinF activity and further troubleshooting is required to attempt to get the CinF protein working in our hands and to determine whether it is more efficient than the actinonin CCR protein.

2.2.6.2 Synthesis of Enoyl-CoA Analogue

The expense of octenoyl-CoA synthesis and its inherent instability limit the scale-up of warhead biosynthesis *in vitro* reactions. The N-acetylcysteamine (SNAC) thioesters present as suitable compounds as they feature the terminal section of CoA (**Figure 35**), can be synthesised from inexpensive reagents and have been widely used in the literature as enoyl-CoA analogues [268].

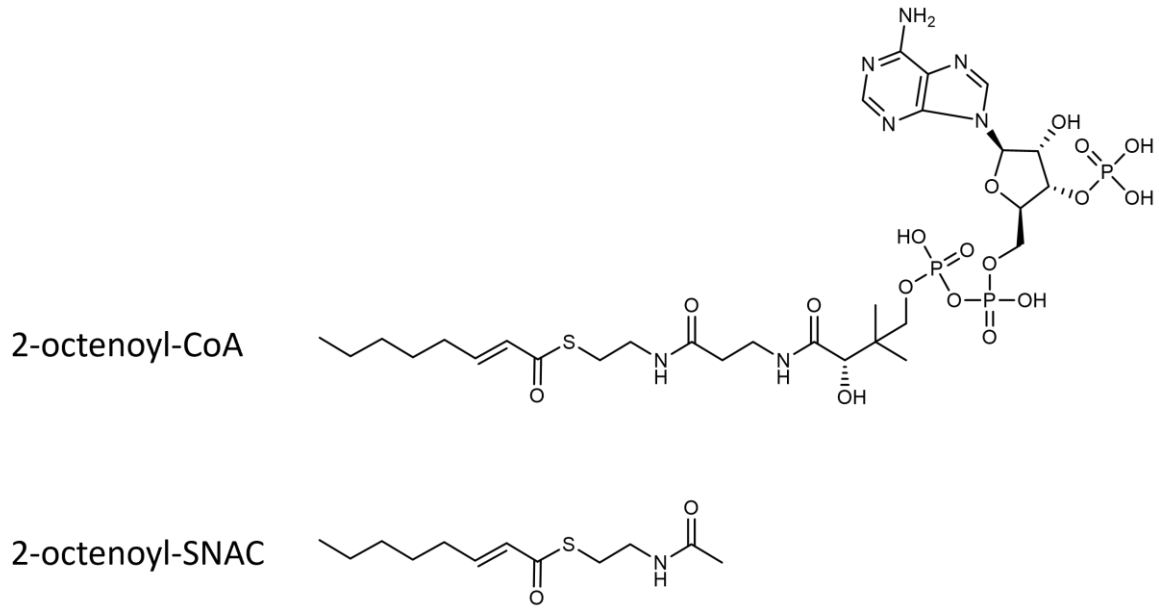


Figure 35. Structural comparison of 2-octenoyl-CoA and 2-octenoyl-SNAC.

The peptide coupling reagent diphenyl phosphoryl azide (DPPA) was used to generate the octenoyl-SNAC thioester according to existing methodology [269] (see materials and methods for details). As the retention factor of octenoyl-sNAC was reported (0.5 in ethyl acetate) [270], purification proceeded by multiple rounds of normal phase flash chromatography and the fractions were analysed by LC-MS and thin layer chromatography. This yielded 17 mg of octenoyl-SNAC, which was pure by ^{1}H NMR and also proved stable over a period of several months, as tested by LC-MS. This purification represents a yield of

only 2%, which could be improved in the future by minimising the number of flash chromatography steps by the incorporation of a HPLC separation step.

An *in vitro* assay was performed using the octenoyl-SNAC compound and ActF. However, LC-MS analysis of the reaction indicated that the carboxylated product (hexylmalonyl-SNAC) was not produced. Further, CinF was unable to carboxylate octenoyl-SNAC in our hands, in contrast to literature reports [230]. Attempts to troubleshoot this reaction include the testing of a range of temperatures, pH and reaction durations, however the carboxylated product was never detected by LC-MS.

The octenoyl-SNAC compound presents as a highly-scalable octenoyl-CoA analogue and may allow the complete validation of the proposed warhead biosynthesis pathway if the biosynthetic proteins exhibit sufficient substrate promiscuity for the SNAC thioester compounds. However, further troubleshooting is required to promote octenoyl-SNAC carboxylation as observed in the literature.

Overall, although the first warhead biosynthesis reaction had been reconstituted *in vitro*, there were clear bottlenecks with the scale-up of reactions for subsequent proposed warhead biosynthesis reactions. In particular, the one-pot *in vitro* reaction represents an elegant method for the generation of intermediate substrates. However, further methodological troubleshooting was required to optimise the in vitro assay conditions and protein activity. Combined with reduced access to the laboratory due to COVID lockdowns, focus was switched to a bioinformatic analysis.

2.2.7 Phylogenetic Analysis of Actinomycete Acyl-CoA Mutase Genes

The actinonin and matlystatin BGCs both feature NRPS genes and the putative warhead biosynthesis genes (**Figure 13**). Therefore, it was proposed that these newly identified genes may be conserved between warhead-NP BGCs and may serve as a genetic probe for the identification of novel warhead-NP BGCs, especially as there are a limited number of characterised examples of NPs with similar warheads, and no other characterised BGCs.

MatB (WP_019634562.1) is the acyl-CoA mutase (α subunit) from the matlystatin biosynthesis pathway. The α subunit of acyl-CoA mutase heterodimers has been reported to confer substrate specificity [242, 246]. The proposed warhead biosynthesis pathway (**Figure 15**) necessitates that the substrate specificity of the α/β acyl-CoA mutase is for hexylmalonyl-CoA. This is in contrast to the previously reported ethylmalonyl-CoA and

methylmalonyl-CoA specific acyl-CoA mutases (see **Section 2.1.3**). Therefore, this implicates MatB as an effective 'probe' for hexylmalonyl-CoA specific mutases. By extension, the presence of the remaining proposed warhead biosynthesis genes and the NRPS biosynthetic machinery in proximity to a 'hit' hexylmalonyl-CoA mutase gene may indicate the presence of a warhead-NP BGC. This approach may lead to the identification of novel warhead-NPs.

Firstly, a phylogenetic analysis of acyl-CoA mutase genes was performed. This analysis was limited to actinobacterial genomes as they are the most prolific specialised metabolite producers, and this approach reduced the necessary computing power. In brief, all actinobacterial protein sequences which contain the methylmalonyl-CoA mutase pfam domain (PF01642) were downloaded. These 13,000 sequences were clustered based on sequence similarity, using the EFI-EST tool [271], to generate a list of 1,615 representative sequences (similarity $<e^{-50}$, ID>40%). These sequences were aligned and used to generate a phylogenetic tree (see materials and methods for details).

Secondly, a co-association analysis was performed to determine whether genes of interest were present nearby the identified actinomycete acyl-CoA mutase genes. The genes of interest include the proposed warhead biosynthesis genes (CCR, acyl-CoA epimerase, acyl-CoA mutase β subunit, MeaB-like, asparagine synthetase and N-oxygenase) in addition to known biosynthetic machinery (including NRPS A, PCP, C, TE domains and PKS ketosynthase genes). In brief, the EFI-GNT [272] tool was used to retrieve the fifteen upstream and fifteen downstream pfam domains for each 'hit' acyl-CoA mutase gene. The presence of the genes of interest was determined for each acyl-CoA mutase gene. This data was converted to a binary matrix for visualisation on the phylogenetic tree using iToL [273] (see materials and methods for details).

The actinomycete methylmalonyl-CoA mutase phylogenetic tree, decorated with the warhead biosynthesis gene co-association analysis, is presented in **Figure 36**. The presence of a filled blue circle represents the presence of a co-associated warhead biosynthesis gene and the presence of a filled red circle represents the presence of a co-associated biosynthetic machinery gene. Overall, the actinomycete acyl-CoA (α) genes are often co-associated with the methylmalonyl-CoA mutase (β) subunit gene and MeaB-like gene (especially in the 'lower left' quadrant of the tree). These may represent copies of primary metabolism MCM genes which may be co-transcribed.

In contrast, the co-association of proposed warhead biosynthesis genes (e.g. asparagine synthetase and N-oxygenase) is rare in the tree. The actinonin and matlystatin acyl-CoA

mutase (α) genes are located in close proximity to one another, in a single clade which is coloured red in **Figure 36**. The members of this clade are of interest as they share high sequence similarity to the proposed encoded hexylmalonyl-CoA specific acyl-CoA mutase proteins of the actinonin and matlystatin BGCs. Further, this clade is enriched in the presence of both proposed warhead biosynthesis genes and NRPS/PKS biosynthetic machinery genes (**Figure 36**). Therefore, it is possible that the strains are represented in this clade of interest are producers of novel warhead-NPs. This 'warhead clade' will be further analysed in the following chapter.

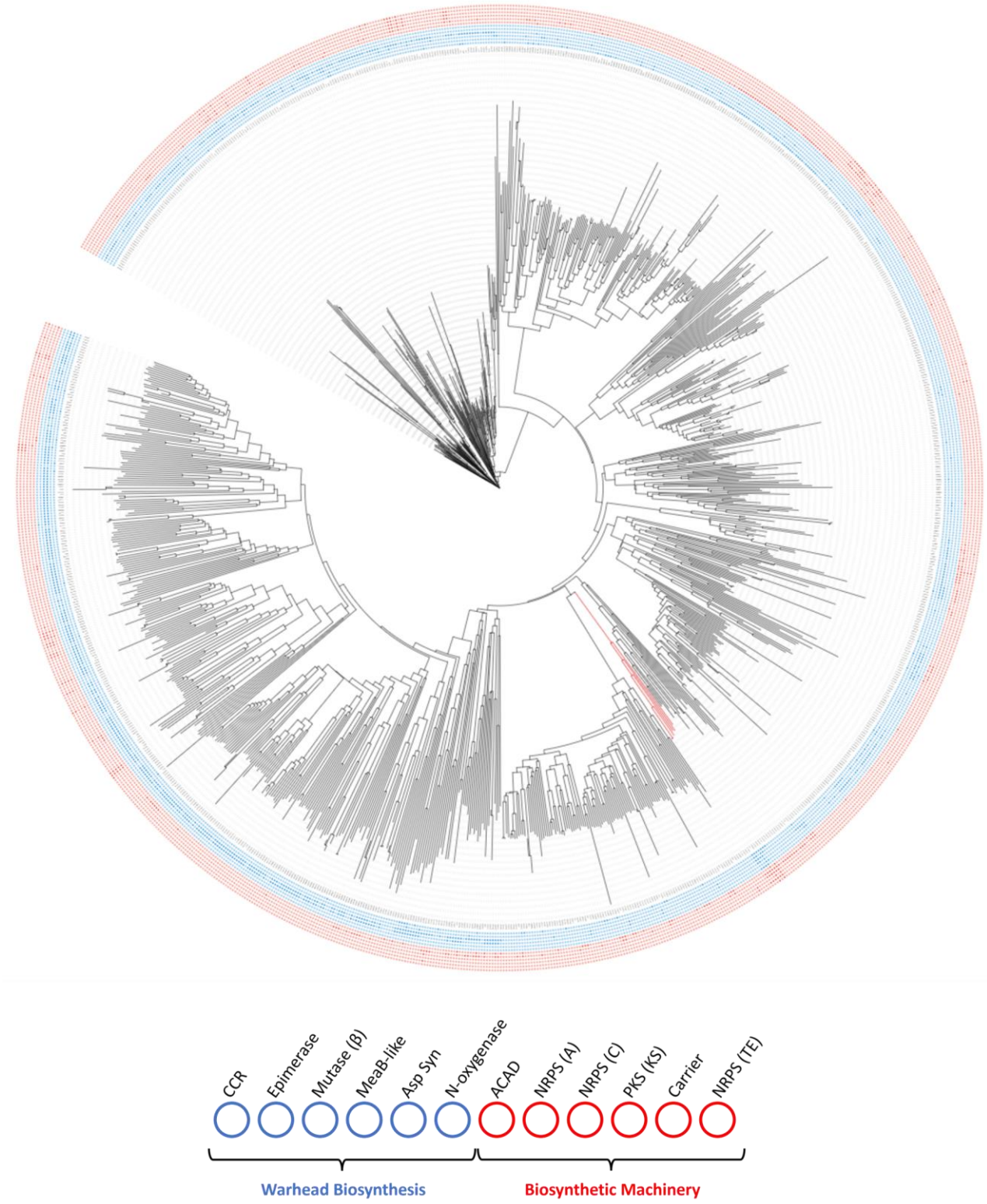


Figure 36. Phylogenetic tree of actinomycete methylmalonyl-CoA mutase (α) homologs. Decorated with co-association analysis of proposed warhead-NP biosynthesis genes. A filled blue circle indicates the presence of a proposed warhead biosynthesis gene. A filled red circle indicates the presence of an NRPS or PKS gene. The 'warhead clade' is coloured in red. Abbreviations; CCR = Crotonyl-CoA carboxylase/reductase, Asparagine Synthetase, NRPS = Non-Ribosomal Peptide Synthetase (Adenylation, Condensation), PKS = Polyketide Synthase (Ketosynthase), ACAD = Acyl-CoA dehydrogenase, TE= Thioesterase.

2.3 Discussion and conclusions

2.3.1 Summary / Discussion of Results

In summary, the *in vitro* reconstitution of the warhead biosynthesis pathway required the purification of the proposed warhead biosynthesis genes and the synthesis of the proposed starting substrate, octenoyl-CoA.

Following the optimisation of protein purification methodology, the proposed warhead biosynthesis proteins encoded by the actinonin pathway proved to be of greater solubility than their counterparts encoded by the matlystatin BGC. In particular, the ActF (CCR), ActC (acyl-CoA epimerase) and ActM (acyl-CoA mutase β subunit) proteins were sufficiently soluble and were purified to an extent required for *in vitro* assays. A preliminary purification experiment indicated that the tri-expression of ActM, ActD (acyl-CoA mutase α subunit) and ActN (MeaB-like) greatly enhanced the solubility of ActD.

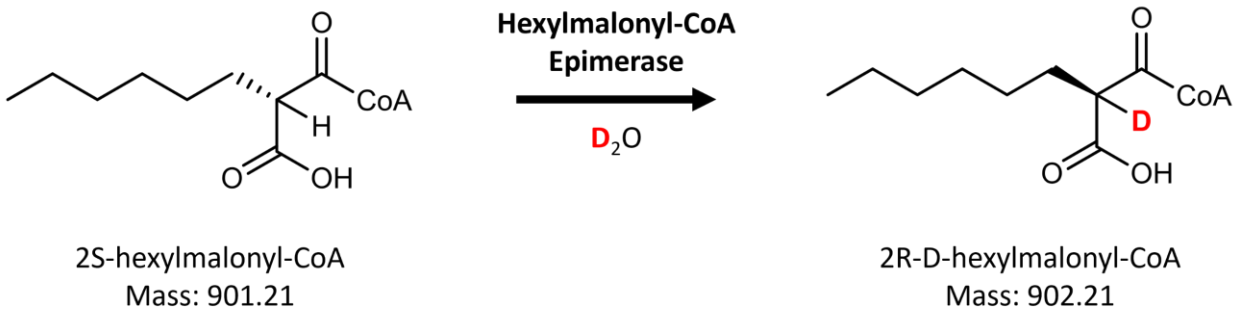
Two synthetic routes and several purification options were tested for the preparation of octenoyl-CoA. Overall, the PyBOP synthesis method, followed by reverse phase chromatography purification, proved to offer the greatest yield and purity. The octenoyl-SNAC thioester analogue was also synthesised and purified as an alternative substrate option.

The first proposed warhead biosynthesis reaction (the carboxylation of octenoyl-CoA by CCR) was tested *in vitro* using the ActF protein. LC-MS analysis indicated that ActF was able to carboxylate both octenoyl-CoA and crotonyl-CoA and the addition of carbonic anhydrase greatly enhanced turnover efficiency. The use of stable isotope labelled bicarbonate confirmed that ActF was carboxylating the substrate.

However, the progression of warhead biosynthesis *in vitro* reconstitution slowed as it became clear that subsequent reactions would require either *in situ* substrate generation or up-scaled *in vitro* reactions followed by the purification of intermediate substrates. The previously characterised CinF CCR protein was effectively purified, however CinF displayed no activity towards octenoyl-CoA in our hands. Attempts to use the inexpensive and stable octenoyl-SNAC thioester analogue for higher scale reactions were hampered by the lack of CinF or ActF activity. Attempts to generate octenoyl-CoA *in situ* using the previously described RevS MC-FACL protein were also curtailed by a lack of solubility and activity.

Although the full *in vitro* reconstitution of the proposed warhead biosynthesis pathway was unable to be completed due to the reasons outlined above, there are several possible options for carrying out this work in the future. The ActC (acyl-CoA epimerase) protein proved to be extremely soluble and effectively purified. The preliminary tri-expression experiment indicated that ActD and ActM (acyl-CoA mutase α and β subunits) could be effectively purified in the presence of ActN (MeaB-like) protein. Therefore, if the problems with the CCR *in vitro* reaction scale-up and/or *in situ* generation of hexylmalonyl-CoA could be rectified, then the proposed hexylmalonyl epimerase and mutase *in vitro* assays could be performed.

A novel acyl-CoA epimerase *in vitro* assay has been theorised, although it has not been able to be empirically tested nor optimised. The substrate and product of the acyl-CoA epimerase have equal masses and only differ by the chirality of a single stereocentre – therefore they cannot be distinguished by LC-MS analysis. It is reported that the methylmalonyl-CoA epimerase flips the stereochemistry by rearrangement of the hydrogen atom about the alpha carbon [239]. Further, the incorporated hydrogen atom is known to be solvent derived [274]. Thus, an advanced epimerase assay has been proposed, whereby conducting the reaction in D_2O should result in incorporation of a deuterium atom at the alpha carbon by the epimerase enzyme (**Figure 37**). This has practical considerations, as the reaction must be quenched, lyophilised and the re-dissolved in H_2O to replace all exchangeable deuterium atoms with hydrogen. In a reaction where the epimerase enzyme has flipped the stereochemistry, the product will be detectable by LC-MS as it would feature a *m/z* ratio of [M+1], where M is the *m/z* ratio of the substrate.



2S-hexylmalonyl-CoA
Mass: 901.21

2R-D-hexylmalonyl-CoA
Mass: 902.21

Figure 37. Schematic of proposed advanced acyl-CoA epimerase assay. Completion of the reaction in D_2O will result in the incorporation of a deuterium atom at the alpha carbon, which is detectable by LC-MS.

Alongside this work, a bioinformatic analysis was performed using actinobacterial methylmalonyl-CoA mutase (α) genes as probes to identify putative warhead-NP BGCs. This highlighted that a single mutase clade, the 'warhead clade', was enriched with co-associated genes of interest. Therefore, these sequences may represent novel warhead-NP BGCs. This analysis will be expanded upon in the next chapter and the product of one of these novel warhead-NP BGCs isolated and characterised.

There are a number of notable limitations of this co-association analysis. Firstly, the pfam domains of interest are selected manually. Therefore, if a particular reaction can be catalysed by different proteins featuring different pfam domains, these would have to be researched and added to the analysis. Further, poor quality genome sequences hamper this approach as the presence of short contigs increases the risk of any potential BGCs to be split over multiple sequences. In these cases, the true full complement of co-associated genes of interest would not be detected. However, these downsides are rather generalised and could affect any large scale bioinformatic analysis. This phylogenetic and co-association analysis approach has the benefit of pointing the user towards a particular clade or number of sequences which may represent novel BGCs. Therefore, it greatly reduces the number of genomic regions that need to be manually analysed.

Chapter 3: Discovery and Structural Elucidation of Lydiamycin A

3.1 Introduction

3.1.1 The 'Warhead Clade'

Following the phylogenetic and co-association analysis of actinomycete methylmalonyl-CoA mutase genes (**Figure 36**), there was a single clade which was enriched in co-associated genes of interest. As such, this 'warhead clade' contains mutase homologs belonging to putative warhead-containing natural product (warhead-NP) BGCs. The sequences contained within the representative nodes were collated and manually curated to account for instances where the exact same strain had been sequenced multiple times. The dereplicated sequences were used to generate a new phylogeny and co-association analysis of the 'warhead clade' using a sequence from a nearby clade (Frankia sp. EI5c) as an outgroup (**Figure 38**).

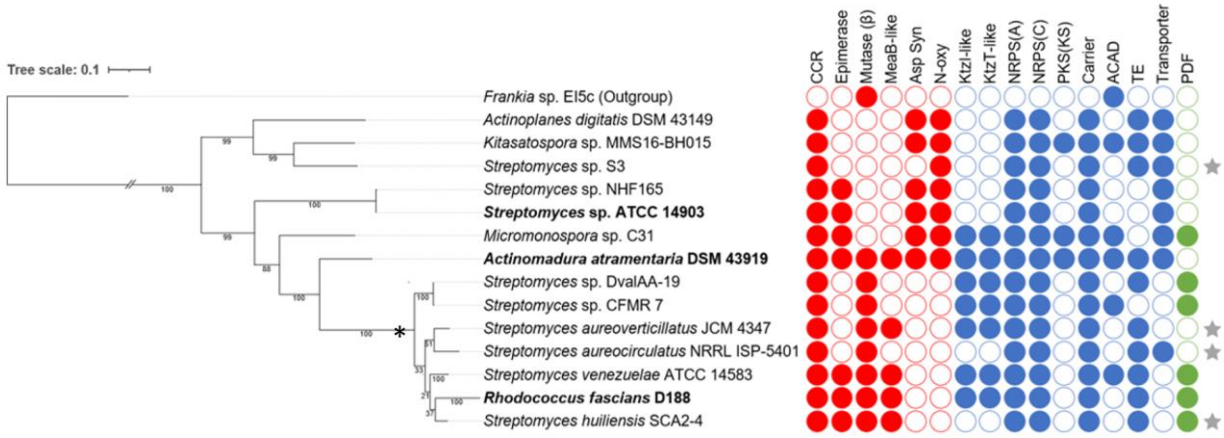


Figure 38. Phylogenetic and co-association analysis of the actinomycete 'warhead' clade. Clade of actinobacterial hexylmalonyl-CoA mutase sequences that are highly co-associated with the proposed warhead biosynthesis genes (red), biosynthetic machinery (blue) and potential resistance gene (green). A filled circle indicates the presence of a homologous gene in close proximity to the mutase gene. The producers of characterised warhead-containing natural products are indicated in bold. Sequences present at the end of a contig are indicated with a star. Phylogenetic tree built using RAxML (bootstrap n=300) and tree manually dereplicated and trimmed. Abbreviations; CCR = Crotonyl-CoA carboxylase/reductase, Asparagine Synthetase, N-oxygenase, NRPS = Non-Ribosomal Peptide Synthetase (Adenylation, Condensation), PKS = Polyketide Synthase (Ketosynthase), ACAD = Acyl-CoA dehydrogenase, TE= Thioesterase, PDF = Peptide Deformylase.

The ‘warhead clade’ contains the producers of the known warhead-NPs actinonin and the matlystatins (*Streptomyces* sp. ATCC 14903 and *Actinomadura atramentaria* DSM 43919, respectively). Also present are representatives of *Actinoplanes*, *Micromonospora* and *Rhodococcus*, which would represent the first of their genera to produce warhead-NPs (**Figure 38**).

Interestingly, there seems to be some discrepancies in the putative warhead biosynthesis genes present in the analysis. For instance, the lower sub-clade of particularly homologous mutase genes (denoted by an asterisk in **Figure 38**) lack the presence of asparagine synthetase and N-oxygenase genes. This suggests that the putative warhead-NP clusters of this sub-clade may biosynthesise products featuring a carboxylated warhead rather than the fully mature hydroxamate group. Whereas, the genomic regions outside of this sub-clade feature the asparagine synthetase and N-oxygenase but most lack some key putative warhead biosynthesis genes (e.g. the small mutase subunit and MeaB-like chaperone). It is possible that in these cases the small mutase subunit is located elsewhere in the genome, as is the case for actinonin biosynthesis by *Streptomyces sp.* ATCC 14903 [207].

Considering the presence of co-associated biosynthetic machinery (blue circles in **Figure 38**), there is a great abundance of NRPS genes – including adenylation domains, condensation domains and thioesterase domains. The presence of the piperazic acid biosynthesis pair KtzI-like and KtzT-like domains appears to be phylogenetically conserved, as they are observed within the lower sub-clade (denoted with an asterisk). The presence of these two domains is indicative of the potential biosynthetic pathway incorporating piperazic acid residue(s) into the peptide backbone. Finally, the peptide deformylase (PDF) domain (PF01327) is also present in some of the putative clusters and may present a self-immunity determinant as the warhead has been shown to confer the PDF inhibitory activity of actinonin [128]. This is investigated in the following chapter.

I selected the *Rhodococcus fascians* D188 potential warhead-NP (hereafter referred to as ‘D188nrp’) BGC for further characterisation, as the mutase homolog is phylogenetically distinct from the known actinonin and matlystatin mutases. Further, the size and specificity of the NRPS is very different to the actinonin and matlystatin pathways. In addition, the D188nrp cluster also features a PDF gene would also allow the determination of whether this represents a self-immunity determinant against the product of the D188nrp cluster (**Figure 39**). Finally, *R. fascians* D188 is an economically relevant plant pathogen and the D188nrp cluster is located on a pathogenicity-associated megaplasmid. As such, it is

possible that the D188nrp product may play a role in plant pathogenicity, which is explored further in **Section 4.2.3**.

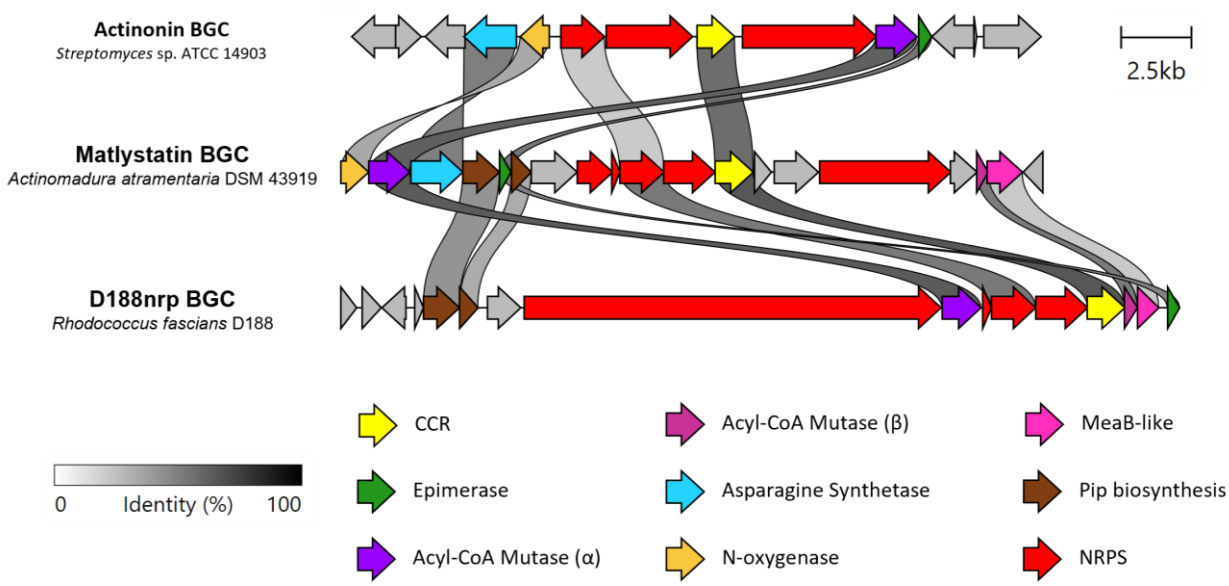


Figure 39. Comparison of the actinonin and matlystatin BGCs to the putative warhead BGC from *Rhodococcus fascians* D188. Gene similarity percentage is represented by the grayscale link between genes.

3.1.2 *Rhodococcus fascians* D188

Rhodococci are ubiquitous in the environment, having been isolated from the soil, seawater, plants, the arctic and desert environments [275-278]. *Rhodococcus* sp. are best known for their bioremediation activities and have been used to decontaminate soils and waters of complex hydrocarbon mixtures, including crude oil and fuels [279]. *Rhodococcus* have recently come to light as under-explored talented specialised metabolite producers as their genomes tend to encode a large number of orphan BGCs with uncharacterised products [280, 281]. There is considerable chemical diversity in the characterised *Rhodococcus* natural products (**Figure 40**), including the antimicrobials aurachin and rhodostreptomycin [282, 283] and the siderophores rhodobactin and rhodochelin [284, 285]. There are two pathogenic *Rhodococcus* species – *R. equi* infects foals and causes pneumonia [286] and *R. fascians* is the causative agent of leafy gall disease in plants [287].

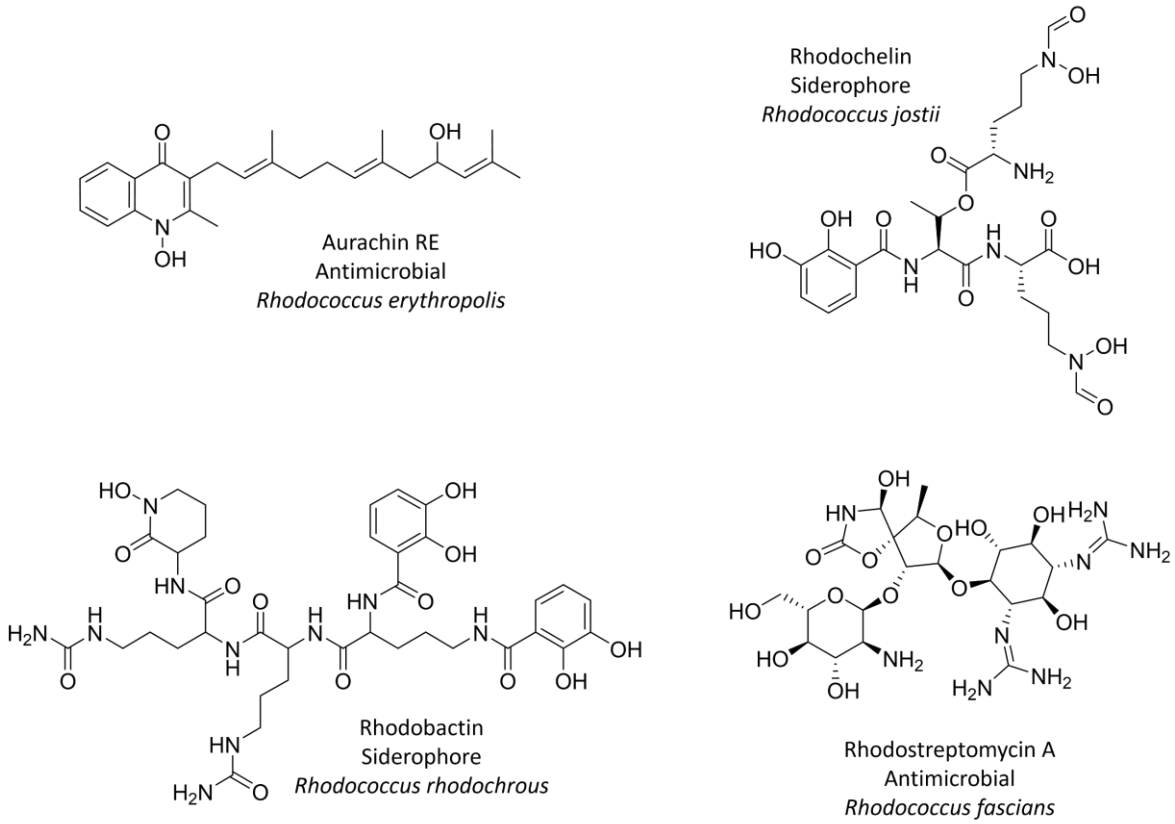


Figure 40. Structures of natural products produced by *Rhodococci*. Producing organism is denoted.

R. fascians causes leafy gall disease by infiltrating the plant and biosynthesising cytokinins – plant hormones that disrupt normal development. The symptoms of leafy gall disease range from deformed leaves and excessive flowering to the excessive growth of deformed shoots that form a densely packed leafy gall [288, 289]. These symptoms culminate in a severely stunted and bushy plant, often with deformed or aborted flowers, which causes considerable economic losses to the ornamental plants industry [290]. *R. fascians* has a very wide host range and leafy gall disease has been reported in over 40 plant families and 160 species ranging across monocotyledons and dicotyledons [291].

Leafy gall disease was first described on diseased sweet pea plants in 1927 [292, 293] and the causative agent was later isolated as an orange gram positive, non-motile, rod-shaped bacterium termed *Phytomonas fascians* [287], to reflect its fasciation-inducing phytopathogenicity. Advances in taxonomy led to the eventual reclassification of this bacterium to *Rhodococcus fascians* [294].

R. fascians is a well-adapted epiphyte that can survive on the surface of a plant for months prior to initiating infection [291]. The production of orange carotenoid pigments [295] protects *R. fascians* from UV irradiation and the production of a slime layer [276] may

enhance biofilm formation and provide desiccation resistance. The switch to endophytic colonisation is coordinated by the *attenuation* (*att*) operon which biosynthesis a quorum sensing-like att molecule, of unknown structure [293, 296]. High att concentrations are proposed to trigger the infiltration of the bacteria, establishment of an endophytic population and the expression of the *fasciation* (*fas*) operon [291, 296].

The *fas* operon biosynthesises a mix of cytokinins [297] which function in concert to modulate plant gene expression. This plant hormone imbalance results in excessive differentiated tissue growth that originates from existing shoot meristems and also *de novo* formed shoot meristems [298] and blocks tissue maturation [299]. Further, cytokinin imbalance promotes plant biosynthesis of other plant hormones, auxin and putrescine [300, 301], which function to exacerbate infection symptoms. Taken together, this abrogated plant hormone balance results in uncontrolled shoot growth which develops into a gall. Symptomatic tissues become sink regions, enriched in carbon sources and amino acids [302], which the metabolism of *R. fascians* shifts to take advantage of and to detoxify plant-derived primary metabolites [303] in order to support niche establishment.

R. fascians D188 is the wild type strain and features a conjugative linear 199 Kb megaplasmid, termed pFiD188, that has been determined to be vital for pathogenicity [304]. The *R. fascians* D188-5 strain lacks pFiD188 and is avirulent *in planta* [305, 306]. Conversely, the conjugation of pFiD188 into *Rhodococcus* isolates lacking virulence genes conferred the ability to cause leafy gall disease [307]. pFiD188 features regions required for plasmid maintenance which are shared with other Rhodococcus linear plasmids (**Figure 41**). There are also regions unique to *R. fascians* (shown in blue in **Figure 41**) which are involved in host interactions and disease progression [289]. This includes both known genetic determinants of leafy gall disease, the *att* and *fas* operons [296, 308]. Interestingly, the *D188nrp* cluster is also present in the *R. fascians* unique region of pFiD188 (**Figure 41**), which implicates the product as a potential pathogenicity factor.

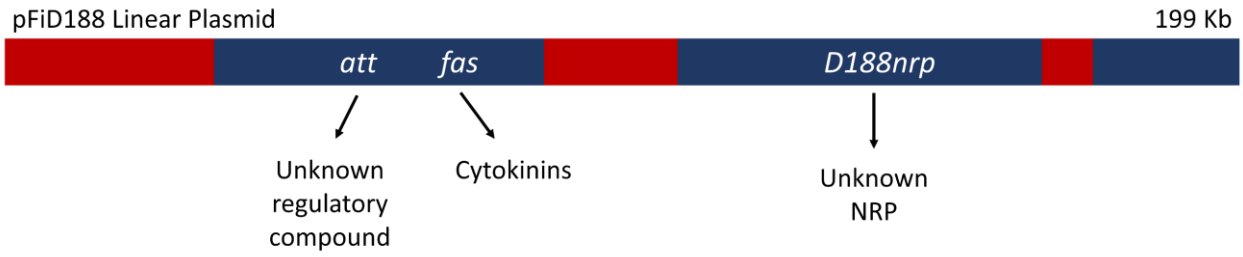


Figure 41. Schematic of pFiD188 megaplasmid. Features regions involved in plasmid maintenance (red) and regions unique to *R. fascians* (blue). Adapted from [289].

The *D188nrp* cluster has been investigated in the literature previously. Francis and colleagues were interested in elucidating whether this cluster is involved in plant pathogenicity [304]. The D188 Δ*nrp* deletion mutant was generated by excision of the adenylation domain of the third module (**Figure 43**) in order to prevent biosynthesis of the *D188nrp* product. However, the *D188nrp* product was not determined to be involved in virulence as the D188 Δ*nrp* and D188 WT strains were indistinguishable in an *in planta* leafy gall disease assay [304]. Therefore, the function of the *D188nrp* product remained elusive and there had not been an attempt to isolate the compound nor elucidate the structure.

We therefore established a collaboration with the Francis laboratory (California State University Bakersfield, USA), who provided the *R. fascians* WT, D188-5 (plasmid free) and D188 Δ*nrp* strains in order for us to determine the product of the *D188nrp* cluster and isolate the putative warhead-NP.

3.1.3 Chapter Aims

Bioinformatic analysis has highlighted a number of potentially novel warhead-NP producers. The *R. fascians D188nrp* cluster was selected for further investigation as it is phylogenetically and enzymatically distinct from the characterised warhead-NPs actinonin and matlystatin. As such, the *D188nrp* product may feature interesting chemistry and may be involved in the plant pathogenicity of *R. fascians*. The overall aim of this chapter is to identify the metabolite(s) produced by the *D188nrp* cluster and determine their structure. To achieve this, the specific aims of this chapter are:

i. Use comparative metabolomics to identify the product(s) of the *R. fascians D188nrp* cluster.
ii. Isolate the product(s) of the *D188nrp* cluster.
iii. Elucidate the structure of the *D188nrp* cluster product(s).

3.2 Results

3.2.1 Analysis of the *R. fascians D188nrp* Biosynthetic Gene Cluster

The *R. fascians* pFiD188 megaplasmid sequence (GenBank: JN093097.1) was analysed by antiSMASH [68], which identified a candidate NRPS BGC spanning 55 Kb. However, there were many genes with predicted functions unrelated to secondary metabolite biosynthesis. Comparison with the candidate actinonin and matlystatin BGCs, this region was trimmed to the 29.7 Kb *D188nrp* cluster shown in **Figure 42**. Each of the gene sequences were analysed by BLAST [309] and CDART [310] to predict gene function based on sequence homology and domain content respectively, which is summarised in **Table 5**.

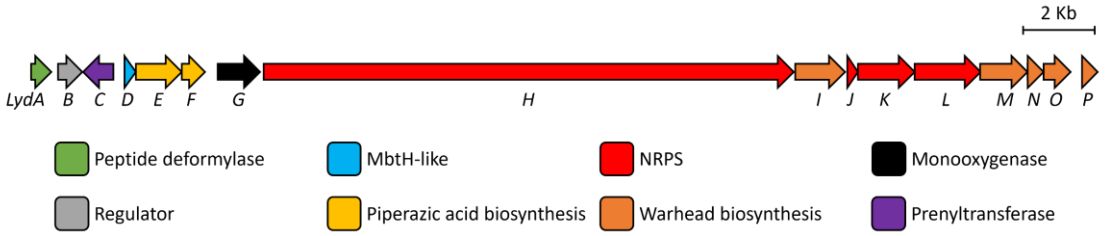


Figure 42. Schematic of D188nrp cluster from *Rhodococcus fascians* D188.

Table 5. Details of predicted protein functions of the genes present in the D188nrp cluster. Abbreviations: CCR (crotonyl-CoA carboxylase/reductase)

Gene name	Protein accession number	Pfam domain	Predicted function	Size (AA)	Size (KDa)
lydA	WP_015586162.1	PF01327	Peptide deformylase	188	20.71
lydB	AET25245.1	PF00196	LuxR-type regulator	226	24.86
lydC	WP_015586164.1	PF01040	UbiA family prenyltransferase	282	29.44
lydD	WP_015586165.1	PF03621	MbtH-like	103	11.50
lydE	WP_015586166.1	PF13434	KtzI-like	424	46.95
lydF	WP_202902192.1	PF04299	KtzT-like	218	24.51
lydG	WP_015586168.1	PF01494	Monooxygenase	394	42.68
lydH	WP_015586169.1	Multiple	NRPS	4925	537.17
lydI	AET25252.1	PF01642	Hexylmalonyl-CoA mutase (α)	461	49.76
lydJ	AET25253.1	PF00550	NRPS	99	11.06
lydK	WP_015586172.1	PF00501	NRPS	527	56.99
lydL	WP_015586173.1	PF00668	NRPS	614	67.49
lydM	WP_015586174.1	PF08240	CCR	440	47.68
lydN	WP_015586175.1	PF02310	Hexylmalonyl-CoA mutase (β)	151	16.37
lydO	WP_015586176.1	PF03308	MeaB-like	251	26.66
lydP	WP_015586177.1	PF13669	Hexylmalonyl-CoA epimerase	144	15.89

The *D188nrp* cluster features genes predicted to encode a crotonyl-CoA carboxylase/reductase (CCR), a hexylmalonyl-CoA epimerase, the hexylmalonyl-CoA mutase (small and large subunits) and the MeaB-like mutase chaperone, which have all been previously proposed to collectively catalyse the first three steps of the putative warhead biosynthesis pathway (**Figure 15**) [207, 217]. The lack of late-stage warhead maturation genes encoding an asparagine synthetase and an N-oxygenase suggest that the D188nrp product may feature an intermediate, carboxylated, warhead moiety.

The *D188nrp* cluster features four predicted non-ribosomal peptide synthetase (NRPS) genes and a MbtH-like gene. MbtH-like proteins have been shown previously to function as chaperones in NRPS systems [311]. Therefore, these proteins are predicted to biosynthesis the peptide backbone of the cluster product. Consideration of the predicted domain architecture of the NRPS (**Figure 43**) suggest that the largest NRPS protein, LydH, is made up of four modules, whilst the *lydJ*, *lydK* and *lydL* genes are predicted to encode a standalone carrier protein, adenylation domain and condensation domain, respectively. As such, as there are five total NRPS modules it is possible to predict that the D188nrp product will feature a pentapeptide backbone with a terminal warhead group.

In order to predict the identity of the peptide backbone of the D188nrp product, the sequences of the NRPS domains were analysed. In particular, key residues in the substrate-binding pocket of adenylation domain sequences (known as the Stachelhaus sequence) can be analysed to predict the substrate specificity and therefore which amino acid residue is incorporated at each module [87]. Several tools were used to predict adenylation substrate specificity; NRPSpredictor2 [88], NRPS/PKS Analysis [312] and NRPSsp [313] (**Figure 43**). However, there was no consensus result reached by all tools.

As the genetically related actinonin and matlystatin NRPS genes have been described and their adenylation domain specificity proposed, their Stachelhaus sequences were manually compared to those of the *D188nrp* adenylation domains (**Figure 44**). The Stachelhaus sequences of the standalone LydK adenylation domain and the MatJ adenylation domains were found to be identical (denoted by asterisk in **Figure 44**). As MatJ has been proposed to specifically activate piperazic acid [207], this analysis suggests that LydK may also activate piperazic acid for incorporation into the peptide backbone. Further supporting this prediction is the presence of *ktzI*-like and *ktzT*-like genes in the *D188nrp* cluster (*lydE* and *lydF* – 48% and 47% sequence similarity, respectively). KtzI and KtzT have been described in the literature to work in concert to convert ornithine to piperazic acid [314].

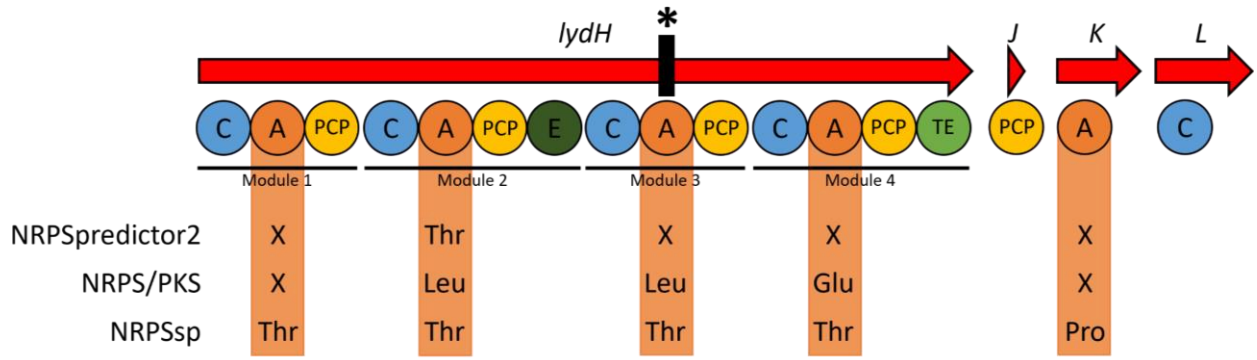


Figure 43. Schematic of NRPS genes of Rhodococcus fascians D188nrp cluster and predicted domain architecture. The predicted adenylation specificity using three different tools is denoted. The position that is disrupted in the Δ*nrp* mutant is denoted with an asterisk. Abbreviations: adenylation (A), condensation (C), peptide carrier protein (PCP), thioesterase (TE).

Known Substrate	Adenylation Domain	Stachelhaus sequence								
	LydH M1	D	V	W	S	V	A	M	V	
	LydH M2	D	F	W	T	L	G	A	V	
	LydH M3	D	V	W	T	F	G	S	V	
	LydH M4	D	V	W	N	F	G	G	F	
	LydK	D	V	Q	F	N	A	H	M	*
Leu	MatO	D	A	W	F	L	G	N	I	
Pip	MatJ	D	V	Q	F	N	A	H	M	*
Val	ActH	D	F	W	N	I	G	M	V	
Pro	ActE	D	V	Q	Q	V	A	H	V	

Figure 44. Stachelhaus sequence comparison of D188nrp adenylation modules. Adenylation domains from the related warhead-containing natural product clusters, actinonin and the matlystatin, are included as references. Sequences were aligned to the GrsA reference sequence and the residue numbers are 235, 236, 239, 278, 299, 301, 322, 330 from left to right. Identical sequences are denoted by an asterisk. Abbreviations: Piperazic acid (Pip), Module x (Mx).

In the proposed matlystatin biosynthetic pathway, the piperazic acid is proposed to be activated by the standalone adenylation domain MatJ, loaded onto the standalone MatI carrier protein, and then condensed to the warhead via the standalone condensation domain MatH [207]. In the genetic context, these three genes are located in the same direction in the order *matH*, *matI*, *matJ*. There is a direct correlation in the *D188nrp* NRPS system as the *lydJ*, *lydK* and *lydL* genes are located next to one another and in the same

direction. LydJ is a predicted standalone carrier protein, LydK is a predicted standalone adenylation domain with specificity to piperazic acid and LydL is a predicted standalone condensation domain. As such, it is possible that LydK activates the piperazic acid, then LydJ shuttles piperazic acid to LydL for condensation between the piperazic acid and the warhead molecule. Biosynthesis is then likely to proceed with four additional monomers being condensed to the piperazic acid-warhead dipeptide by the four modules of the LydH NRPS assembly line. Module four of LydH is predicted to feature a terminal thioesterase domain which may catalyse product release as either a linear or macrocylised liberated product [81]. Therefore, taken together these data indicate that the D188nrp product may feature a pentapeptide backbone (with at least one piperazic acid residue) and a terminal carboxylated warhead group.

3.2.2 Identification of candidate D188nrp products by comparative metabolomics

In order to identify candidate products of the *D188nrp* cluster, the *R. fascians* D188 WT and *R. fascians* D188 Δ*nrp* strains were analysed by comparative metabolomics. Both strains were cultured in a range of production media and their metabolites extracted at both five and twelve days of incubation. Extracts of sterile production media were also prepared. These samples were then analysed by untargeted LC-MS in order to determine the metabolomic profiles. The WT and Δ*nrp* base peak chromatograms were manually compared and any compounds present in the WT and absent in the Δnrp and media only chromatograms were candidate *D188nrp* cluster products.

A range of production media (PM) were used in an attempt to induce the production of secondary metabolites. This included media previously used to produce *Rhodoccocus* natural products, such as aurachin PM, rhodostreptomycin PM and MinA minimal medium [283, 315, 316]. A range of media which have been used to induce actinomycete natural product production were also tested, including; yeast extract broth (YEB), actinonin PM, matlystatin PM, bottromycin PM, SM7 and SM12 [207]. See **Section 6.1.4** for details.

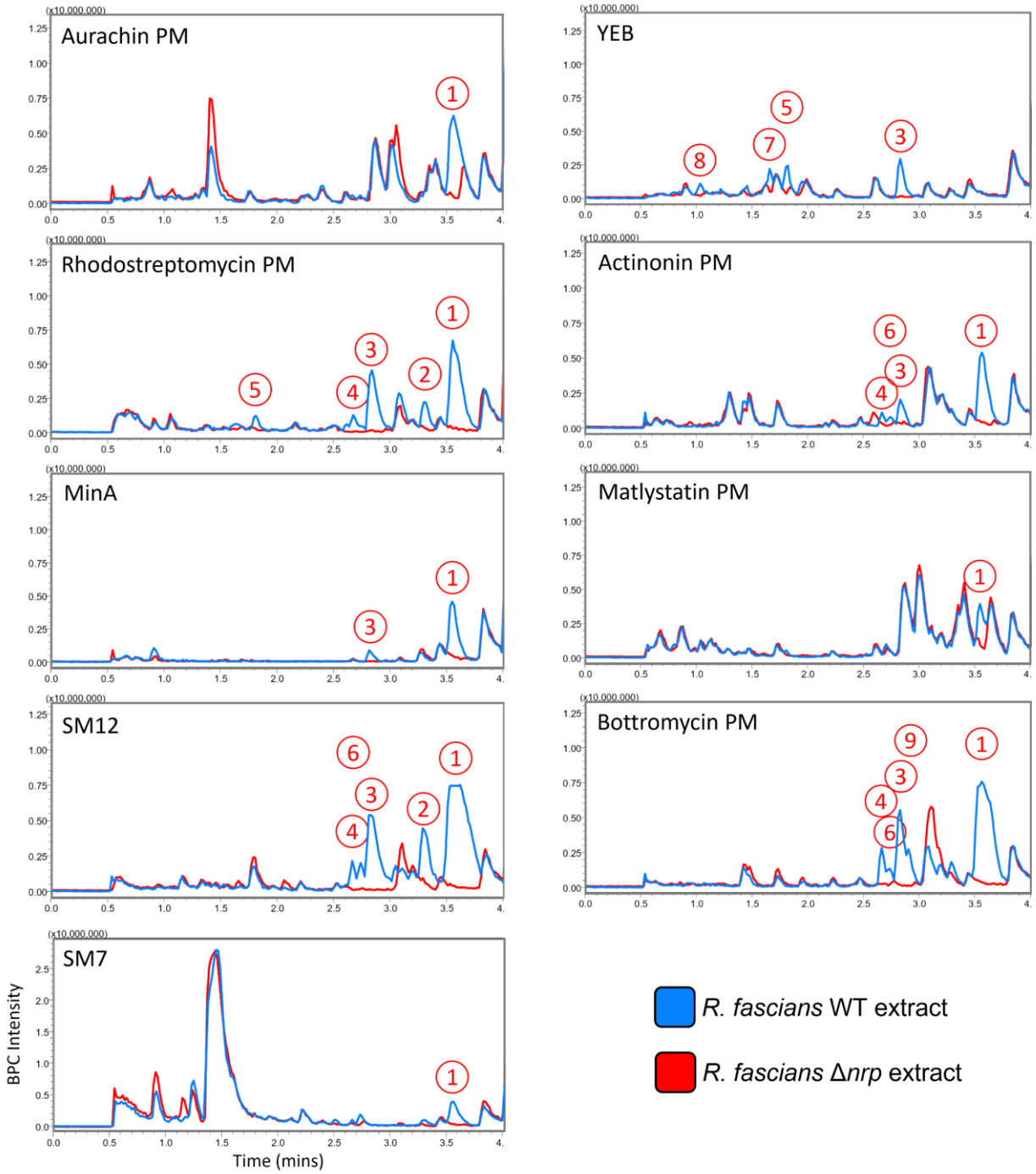


Figure 45. Comparative metabolomics experiment of R. fascians WT and *Δnrp* strains at 5 days growth. Abbreviations: PM = Production Medium, YEB = Yeast Extract Broth, BPC = Base Peak Chromatogram.

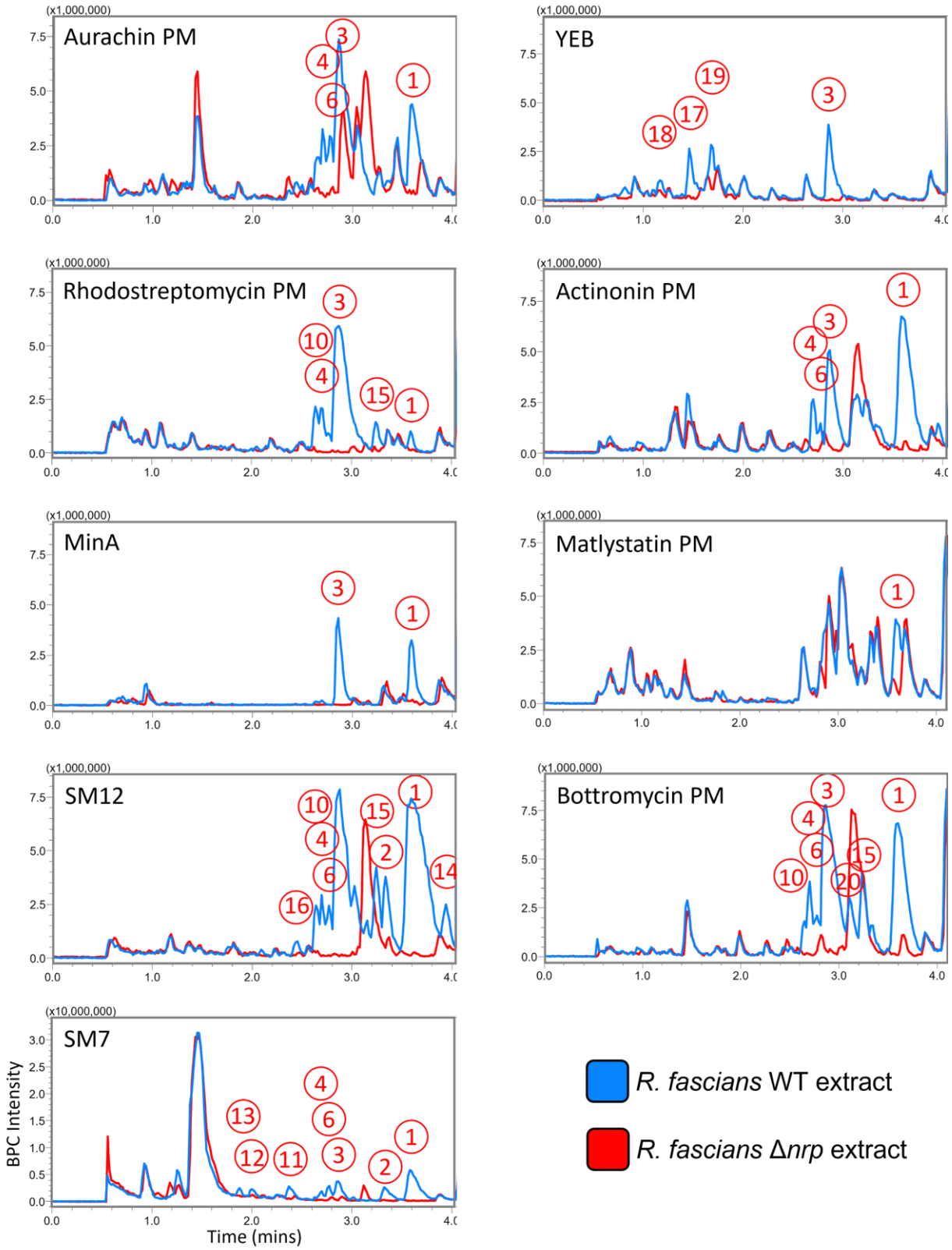


Figure 46. Comparative metabolomics experiment of R. fascians WT and *Δnrp* strains at 12 days growth. Abbreviations: PM = Production Medium, YEB = Yeast Extract Broth, BPC = Base Peak Chromatogram.

This comparative metabolomics analysis indicated that there were a large number of candidate *D188nrp* cluster products produced at both 5 days (**Figure 45**) and 12 days (**Figure 46**) of culture. These different compounds of interest can be primarily differentiated by their retention times and precursor parent ion *m/z* values. As the same compounds were identified in different production media and sampling time points, they were each denoted with an arbitrary number. In total, twenty candidate D188nrp products were identified, and their retention times and precursor parent ion *m/z* values are indicated in **Table 6**. Compounds were only considered candidates when they were present in both of the duplicate *R. fascians* WT extracts and absent in the duplicate *R. fascians* Δ*nrp* extracts and media only extracts. The chromatograms displayed above have been cropped at four minutes retention time as there were no compounds of interest later in the runs.

Table 6. Overview of candidate D188nrp products identified by comparative metabolomics.

Compound #	Retention Time (mins)	Major Ion *m/z*
1	3.55	664.37
2	3.31	650.35
3	2.84	682.38
4	2.67	570.30
5	1.81	530.30
6	2.74	689.35
7	1.66	659.34
8	1.04	603.31
9	2.90	703.36
10	2.64	668.36

Compound #	Retention Time (mins)	Major Ion *m/z*
11	2.37	580.31
12	2.00	577.27
13	1.88	605.26
14	3.94	700.36
15	3.25	664.37
16	2.44	654.35
17	1.48	553.34
18	1.16	594.33
19	1.70	659.34
20	3.10	655.28

There was a large variance in retention times and predicted major parent ion *m/z* ratio in the above twenty candidate *D188nrp* products. Further analysis was required to determine whether these compounds are structurally related. However, those compounds that shared similar retention times and similar masses (including similarities in post decimal place masses) may be structurally related. Common structural modifications can be predicted by comparison of the *m/z* differences between some of the compounds. For example, the *m/z* of compound #1 is 18.0 Daltons less than compound #3 which may be indicative that compound #1 is a dehydrated form of compound #3 via the loss of water. Further, compound #2 may be a demethylated version of compound #1 due to the comparative loss of 14.0 Daltons. A complication of the above data is that it was unknown

whether all of observed major ion *m/z* values were the protonated parent ion or whether other adducts had been detected, e.g. sodiated.

The overall relationship between observed candidate D188nrp product production in the various culture media and sampling timepoints is presented in **Figure 47**. From this analysis, it is clearer that the growth duration of *R. fascians* WT plays a large role in observed metabolite production. For instance, compounds #10 through #20 were only observed at the 12-day timepoint. It was also clear that some production media induce a large amount of varied metabolite production, for instance *R. fascians* WT is producing eight different compounds after 12 days of growth in SM7. Whereas matlystatin production media only supports the production of compound #1 after either 5 or 12 days of *R. fascians* WT growth. There is also variation in the abundance of the compounds and the amount of background complexity (from the bacterial metabolome and the media itself) in the different production media and sampling timepoints.

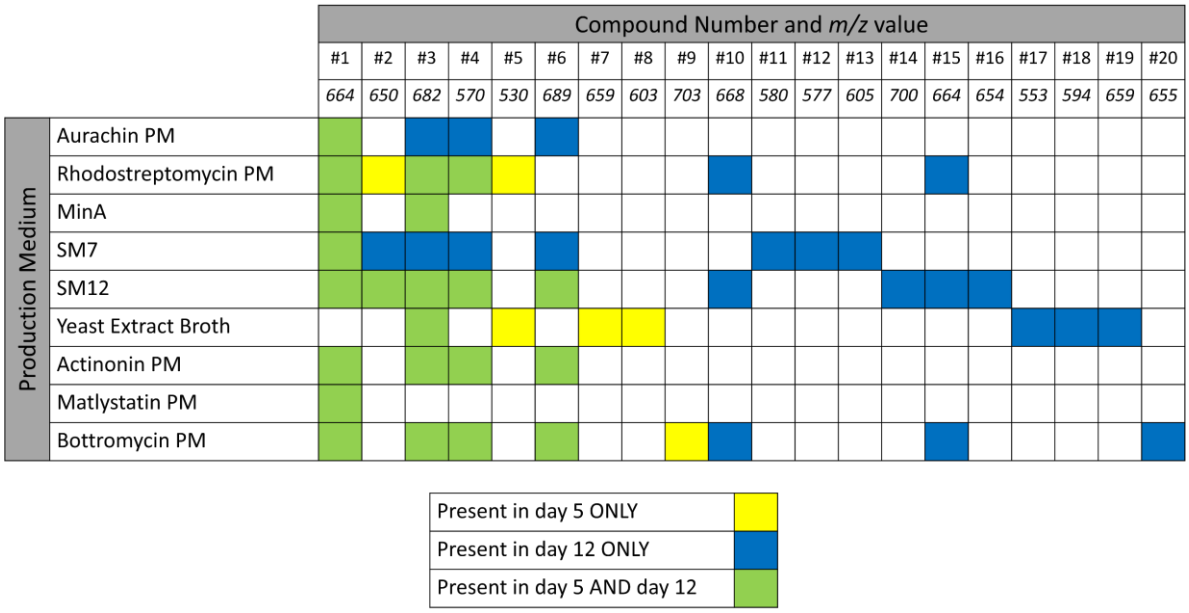


Figure 47. Matrix of candidate D188nrp product production across production media and sampling timepoints. The colour of the filled box denotes which sampling timepoint(s) the compound was observed at. Each candidate D188nrp product is denoted with the assigned reference number (#) and the observed major *m/z* value.

The compound #1 (*m/z* = 664.37) was consistently produced in the most abundant quantities and so was selected as the candidate D188nrp product for isolation and characterisation. The varied production media and sampling time points were assessed to identify conditions that provided a large amount of compound #1 whilst maintaining good separation from background compounds. Based on these criteria, *R. fascians* WT grown in

SM12 medium for 5 days was selected as an appropriate starting point for testing compound #1 isolation.

3.2.3 Purification of compound #1 metabolite

3.2.3.1 Preliminary scale up experiments for the purification of compound #1.

Several small-scale experiments were conducted to understand how best to scale *R. fascians* growth and compound #1 purification. Firstly, the *R. fascians* WT strain was grown in increasing volumes of SM12 medium to determine whether production was maintained at larger growth volumes (25 mL, 50 mL, 200 mL, 400 mL). After five days of growth, an equal amount of culture from each condition was extracted in methanol and analysed by LC-MS. The extracted ion chromatogram peak area for compound #1 was directly comparable between the conditions. The production of compound #1 was maintained throughout the scale-up process (**Figure 48**). This experiment indicated that the *R. fascians* WT strain could be grown in at least 400 mL of SM12 medium (in 2000 mL conical flasks) and scale-up was unlikely to pose a technical challenge in further experiments.

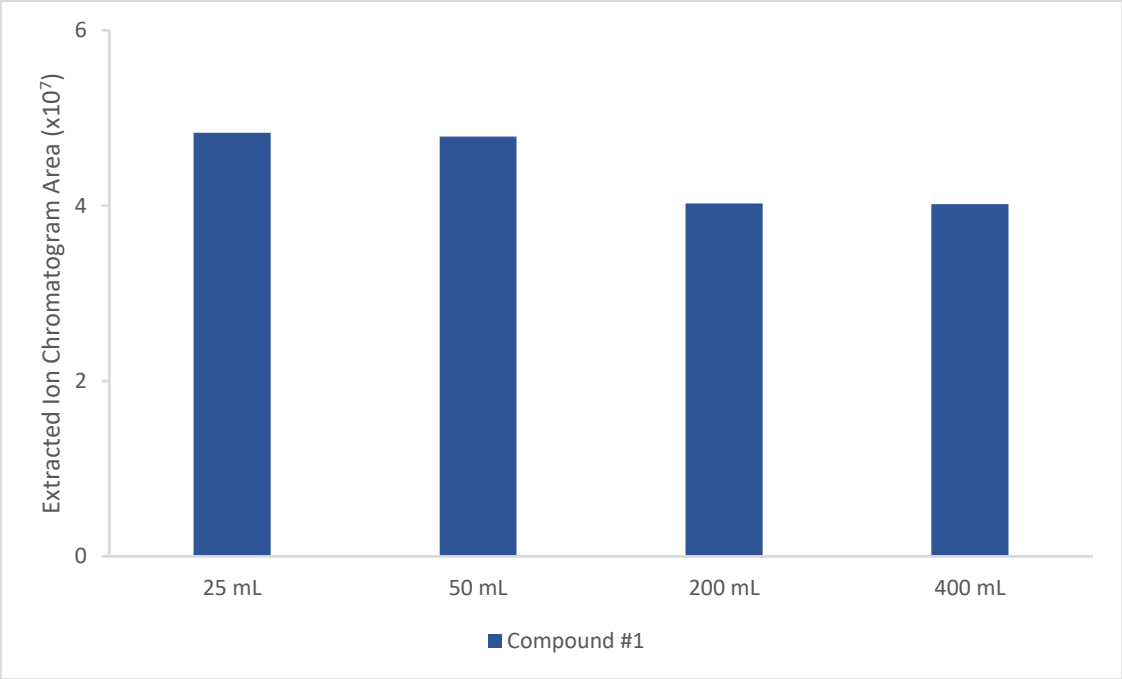


Figure 48. Comparison of *R. fascians* putative D188nrp compound #1 production in different volumes of SM12 growth medium. *R. fascians* D188 was grown in SM12 production media in a range of volumes; 25, 50, 200 and 400 mL. An equal volume of culture was extracted and analysed by LC-MS. The area of the extracted ion chromatograms of compound #1 ($[M+H]^+$ = 664.37) was calculated to assess amount of compound produced. N=1.

Another preliminary experiment was conducted to understand whether the candidate D188nrp products were more readily isolated from the cell pellet or the culture supernatant. Fermentations of *R. fascians* D188 WT grown in SM12 medium were centrifuged. The pellet was extracted in 70% acetonitrile, before being filtered and analysed by LC-MS. The growth medium supernatant was extracted with ethyl acetate, leaving a bulk aqueous fraction. Both the supernatant-derived organic and aqueous fractions were concentrated and analysed by LC-MS. Comparison of the compound #1 extraction ion chromatogram peak areas (**Figure 49**) indicates that compound #1 is present in all three extracts but the most effective extraction technique overall was the extraction of the culture supernatant with ethyl acetate. Therefore, for future larger-scale purifications, the organic extraction of the supernatant was used as it extracted a large quantity of compound #1 and represented a simple method of extraction as the ethyl acetate was readily removed by rotary evaporation to concentrate the fraction and allow further processing.

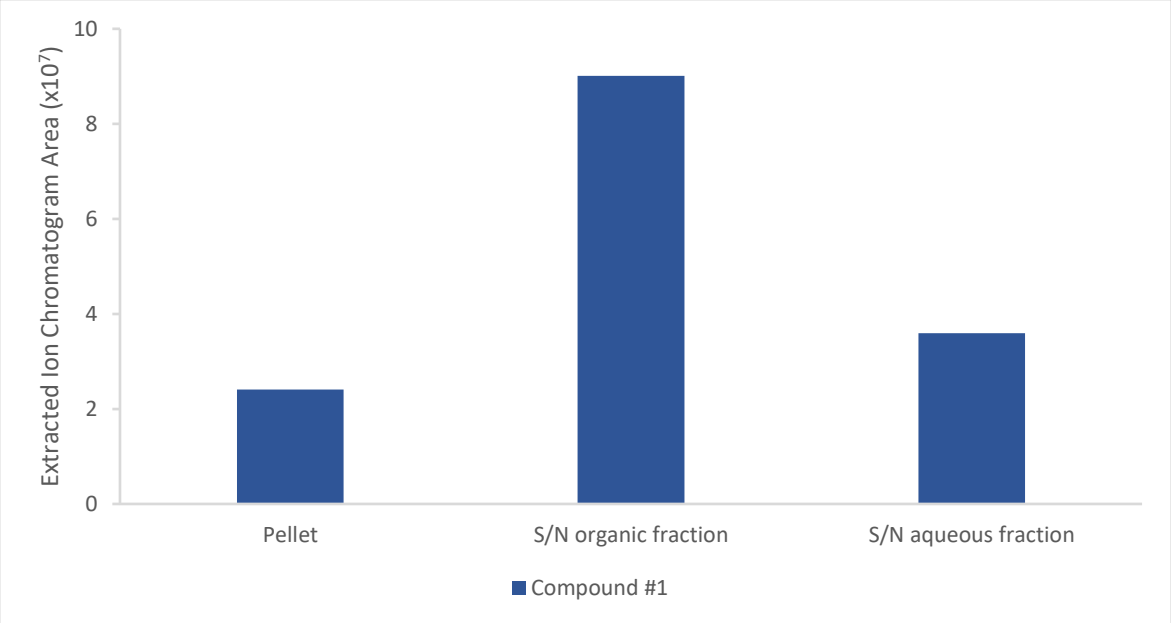


Figure 49. Comparison of putative D188nrp compound extraction techniques. *R. fascians* D188 WT was grown in SM12 production media and extracted using three methods: (1) from the pellet, and from the liquid/liquid extraction of the growth medium supernatant into (2) ethyl acetate organic fraction and (3) aqueous fraction. The extracts were analysed by LC-MS. The area of the extracted ion chromatograms of compound #1 ($[M+H]^+$ = 664.37) was calculated to assess amount of compound produced. N=1.

3.2.3.2 First large-scale purification of compound #1

A larger scale extraction was performed to isolate compound #1 for structural characterisation. In brief, *R. fascians* D188 WT was cultured in 0.6 L SM12 for five days and the culture centrifuged. The culture supernatant was extracted with ethyl acetate and this organic fraction was washed with water. The organic fraction was then concentrated by rotary evaporation.

In order to separate the organic fraction it was fractionated by flash chromatography using a C18 reverse phase cartridge (Biotage). Flash chromatography was chosen as an initial separation technique as it offers separation of a large amount of material, albeit with relatively poor separation characteristics. A gradient of 5-100% acetonitrile across a reverse phase C18 cartridge was used whilst measuring UV absorption at 210 nm (as a standard wavelength) and at 237 nm (which was the absorbance maxima of compound #1 determined during the above comparative metabolomics experiment).

Separation of the organic fraction using flash chromatography (**Appendix Figure 102**) resulted in the separation of two peaks, termed fractions A and B. Both fractions were concentrated to dryness to yield 9 mg of fraction A and 64 mg of fraction B. LC-MS analysis indicated that both fractions A and B contained compound #1 (**Appendix Figure 103**). Fractions A and B were more thoroughly analysed by LC-MS to determine which of the twenty putative D188nrp products were present in each fraction. This analysis indicated that fraction A contained large amounts of compound #1 with minimal amounts of other compounds of interest (**Appendix Figure 104**). LC-MS analysis of fraction B indicated that the major compounds of interest present are #1 and #14 (*m/z* 664.37 and 700.36, respectively) which were well separated (**Appendix Figure 105**).

Fraction A was further purified by HPLC using a semi-preparative reverse phase C18 column (**Appendix Figure 106**) whilst detecting absorbance at 237 nm. This process yielded 0.30 mg of material, which proved pure by LC-MS (**Appendix Figure 107**).

In order to test another method of separation, fraction B (64 mg) was fractionated by mass-guided HPLC using a preparative-scale reverse phase column. Fraction B contained relatively large amounts of compounds #1 and #14 with small amounts of compound #2 (**Appendix Figure 105**). As such, the HPLC separation was performed with mass-guided fractionation for the isolation of the above compounds ($[M+H]^+$ = 664.37, 700.36 and 650.35, respectively). This process afforded a 3.2 mg fraction of enriched compound #1 and two fractions of less than 0.5 mg, one enriched with compounds #2 and one enriched with

compound #14. However, all fractions still featured major contaminants that were difficult to remove (**Appendix Figure 108**).

At this stage, the most effective purification was achieved by the manual HPLC separation of Fraction A, which afforded 0.30 mg of pure compound #1. However, this proved to be an insufficient quantity for NMR analysis (data not shown).

3.2.3.3 Second large-scale purification of compound #1

The most effective compound #1 isolation methodology of **Section 3.2.3.2** proved to be the ethyl acetate extraction of *R. fascians* growth medium, followed by separation by reverse phase flash chromatography and HPLC. This methodology was therefore employed again at a larger scale in order to isolate enough compound #1 for structural characterisation.

In brief, the *R. fascians* WT strain was grown in 2 L of SM12 medium for five days. The culture was centrifuged, and the supernatant extracted with ethyl acetate. This organic fraction was washed with water before being dried *in vacuo* to afford 140 mg of material. This organic fraction was separated by reverse phase flash chromatography, again whilst measuring UV absorption at 210 and 237 nm (**Appendix Figure 109**). Fractions with a high abundance and purity of compound #1 (**Appendix Figure 110**) were collected and dried to yield 27 mg of an off-white amorphous powder. This material was then further purified using a preparative-scale reverse phase C18 column featuring a 45 to 70% gradient in order to enhance separation of any closely eluting contaminants (**Appendix Figure 111**). The resultant single peak (eluting at 13.25 minutes) from multiple injections was collected and dried down to afford 20.1 mg of a white amorphous solid. LC-MS and NMR analysis indicated that this sample was pure compound #1, as explored below.

3.2.4 Structural determination of compound #1

With a large quantity of compound #1 to hand, it was possible to begin determining the structure of compound #1. A combination of LC-MS-MS and nuclear magnetic resonance (NMR) analysis was performed to achieve this.

3.2.4.1 Accurate Mass and MS2 Fragmentation of Compound #1

A sample of compound #1 was analysed by High Resolution-Mass Spectrometry (HR-MS) in order to determine its accurate mass and to obtain a detailed MS-MS fragmentation pattern. The accurate $[M+H]^+$ *m/z* value of compound #1 was determined to be 664.3679 (**Appendix Figure 112**). The *m/z* value of compound #1 was submitted to the Natural Products Atlas [317] in order to indicate whether compound #1 may be a previously reported natural product. The natural product with the most similar mass was found to be lydiamycin A, with a protonated mass of 664.3665 (Δ2.1 ppm to the isolated compound #1). Lydiamycin A was identified from *Streptomyces lydicus* as a antimycobacterial agent [318] and features a pentapeptide cyclised depsipeptide backbone with a terminal carboxylated acyl chain.

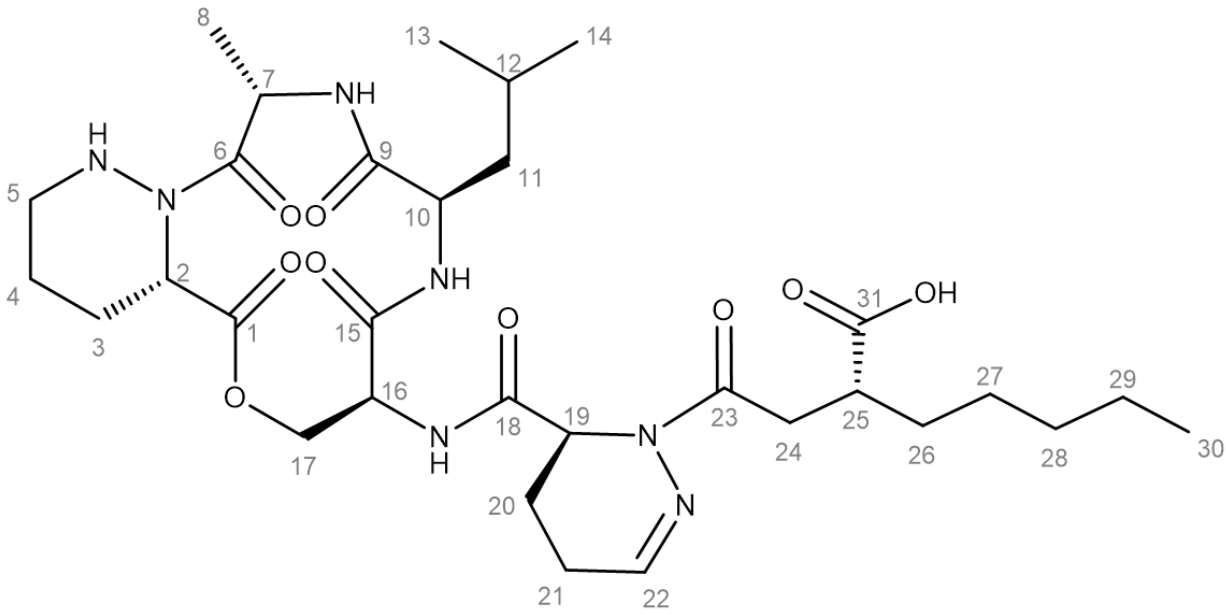


Figure 50. Published structure of Lydiamycin A. Carbon atoms are numbered according to the most recent publication reporting lydiamycin structure [319].

The reported lydiamycin A structure (**Figure 50**) was of interest as it aligns with the biosynthetic potential of the *D188nrp* cluster as described in **Section 3.2.1**. In brief, the *D188nrp* product was predicted to feature a pentapeptide backbone (including at least one piperazic acid residue) and a terminal carboxylated warhead group. Therefore, it is possible that the isolated *D188nrp* product, compound #1, is the previously described lydiamycin A natural product.

In order to test this hypothesis, the MS-MS fragmentation pattern of compound #1 was assessed. MS-MS of a linear peptide results in fragmentation at peptide bonds and the ionised fragments are detected. Fragmentation gives rise to two major series of peptide fragments as there is a resultant N-terminal and C-terminal fragment at each broken peptide bond (known as the *b* and *y* ions respectively). This fragmentation pattern can be analysed to inform the identity of peptide residues and the connectivity of adjacent residues.

Macrocyclic peptides are likely to linearise under MS conditions prior to fragmentation. Consideration of the lydiamycin A structure, the ester linkage of the macrocycle (denoted with an asterisk in **Figure 51A**) is a likely linearisation point. This may occur through an elimination reaction (**Figure 51B**) to give rise to a linearised lydiamycin A species as shown in **Figure 51C**. This features a modified serine residue with a carbon-carbon double bond.

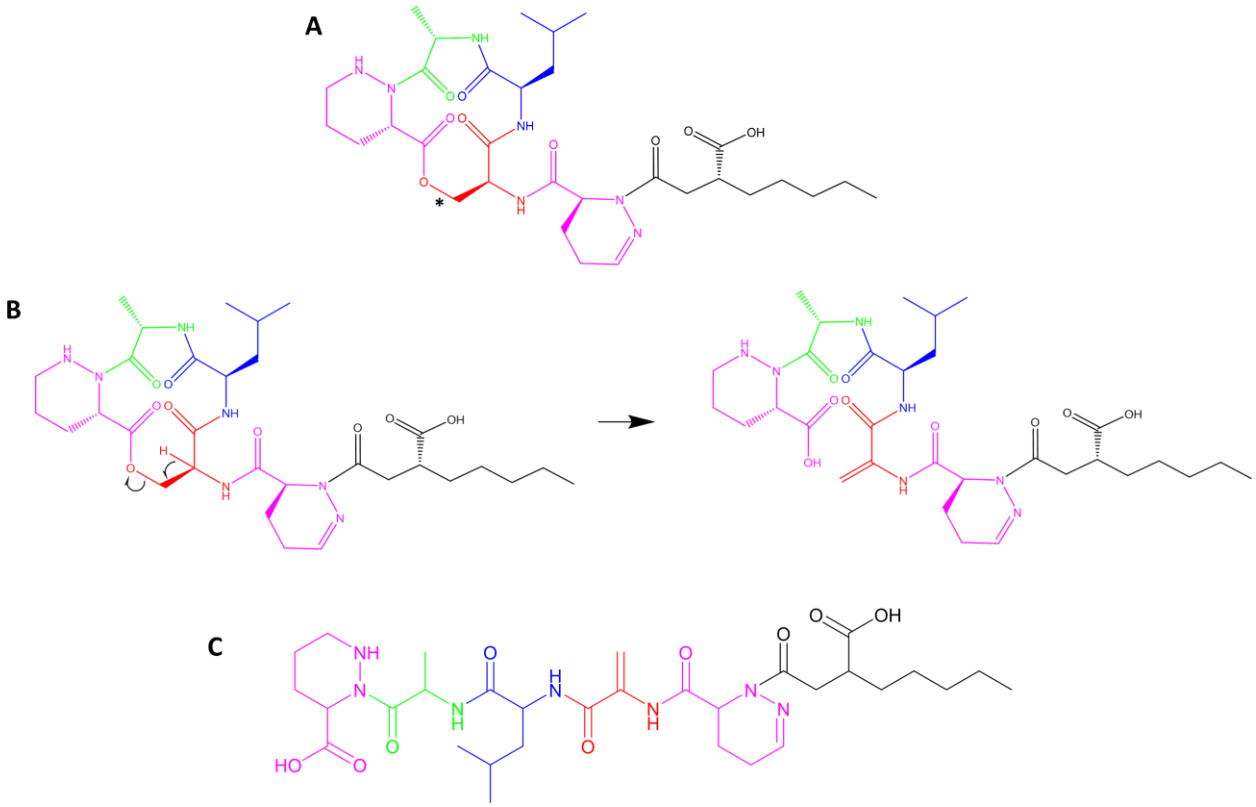


Figure 51. Scheme of predicted lydiamycin A linearisation. The backbone residues are colour coded: Piperazic acid in purple, serine in red, leucine in blue and alanine in green.

The MS-MS fragmentation pattern of compound #1 (**Figure 52**) was analysed to indicate whether the peptide backbone structure is likely to be the same as lydiamycin A. The *b* ion series fragmentation pattern equal to the loss of piperazic acid, alanine, isoleucine/leucine,

a modified serine and a piperazic acid/pentylsuccinamic acid dipeptide from the parent mass ion was observed (**Figure 52**). Further, all but one fragment from the lydiamycin A *y* ion pattern was present – these are indicated by the green masses in the **Figure 52**. Where predicted peptide fragment ion masses are absent, there are 'preceding' larger fragments which account for this discrepancy. For example, the predicted fragment of the pentylsuccinamic acid alone was not detected (*m/z* = 171.10), but the pseudodipeptide fragment of pentylsuccinamic acid/piperazic acid was detected (*m/z* = 281.15). It is common for some fragmentation ions to be absent from the MS-MS pattern.

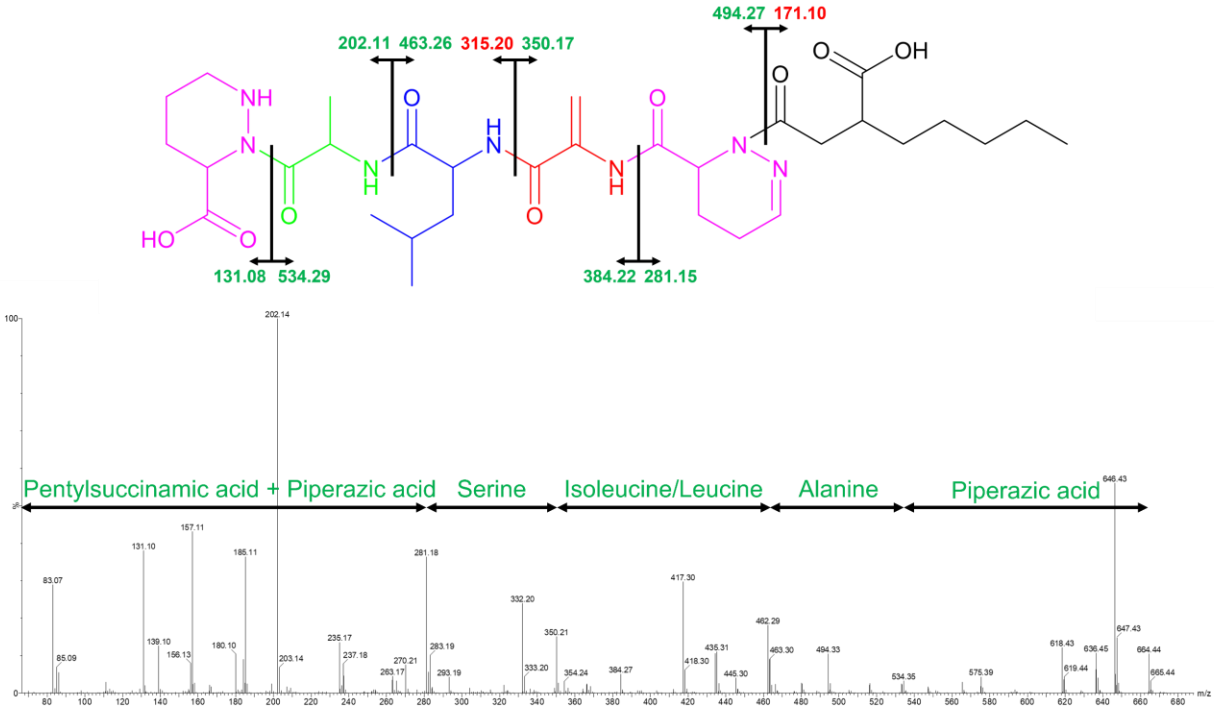


Figure 52. MS-MS fragmentation pattern of compound #1. The backbone residues are colour coded: Piperazic acid in purple, serine in red, leucine in blue and alanine in green. Mass losses of known amino acids are annotated. Expected *y*-ion and *b*-ion series masses are annotated at each amide bond. Masses in green are observed in MS2 data. Asterisk denotes that this residue is predicted to be site of molecule linearisation under MS conditions.

These HR-MS and MS-MS data indicated that the isolated compound #1 has an identical peptide backbone connectivity and a near identical mass (Δ2.1 ppm) to lydiamycin A. However, MS experiments cannot distinguish between isoleucine and leucine residues, nor do they provide stereochemical information. MS-MS experiments also likely necessitate the linearisation of the molecule at the ester bond and give rise to a modified serine residue, thus obfuscating native macrocycle structure. Therefore, these MS data suggested that the isolated compound #1 is lydiamycin A. However, further NMR experiments were

required to confirm this annotation and confirm the identity of the intramolecular cyclisation.

3.2.4.2 Compound #1 Structural Elucidation by NMR

In order to determine whether the isolated compound #1 is lydiamycin A, NMR experiments were conducted in the same NMR solvent as published spectra and compared to these data [319]. To ensure that low intensity NMR peaks were detected, compound #1 was prepared at a relatively high concentration of 35 mg/mL in $CDCl_3$ and analysed in a 3 mm tube in a 600 MHz NMR spectrometer.

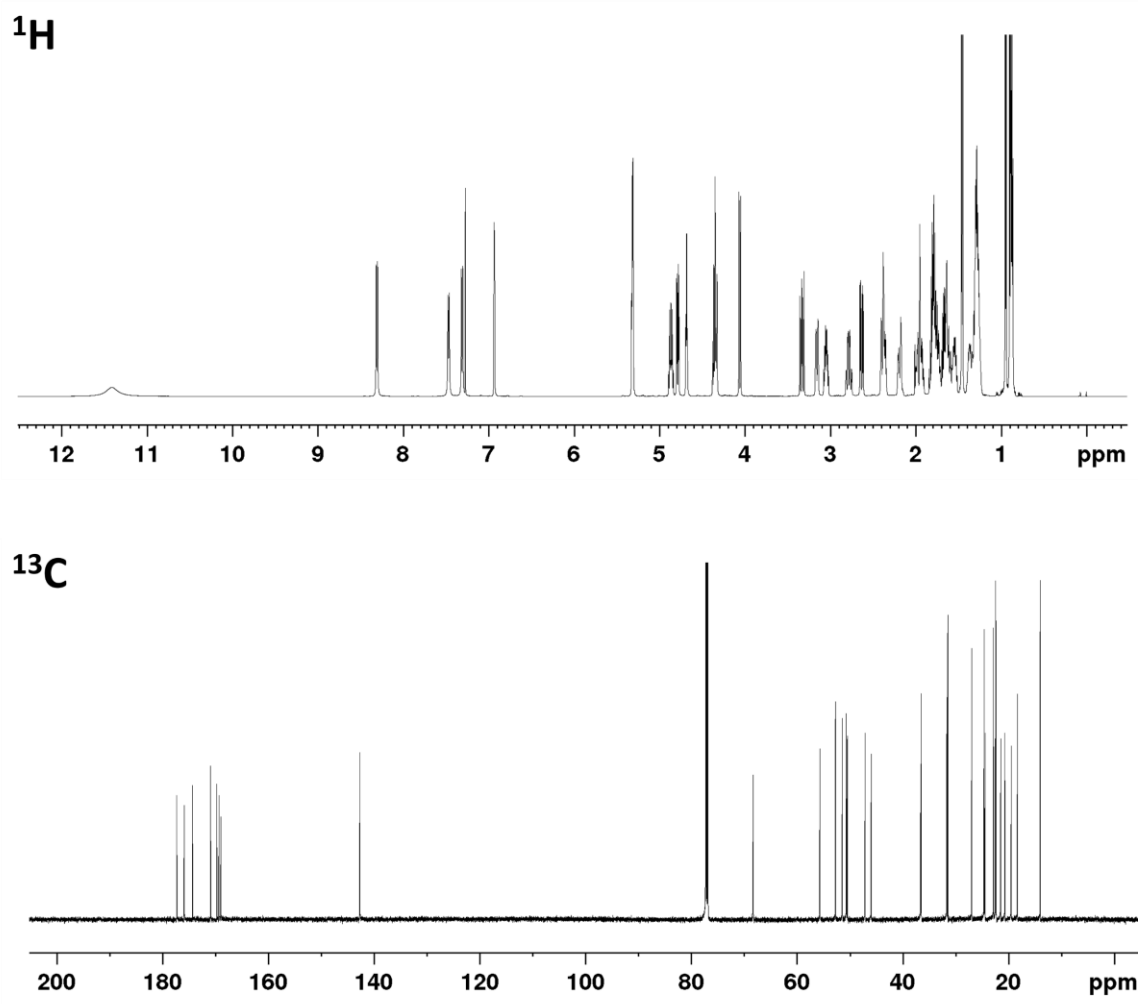


Figure 53. 1D NMR analysis of compound #1. Recorded in $CDCl_3$.

Initially, the proton and carbon-13 1D NMR experiments were conducted in $CDCl_3$ (**Figure 53**). Qualitative consideration of these spectra indicated that the compound #1 sample was very pure as there were no contaminating minor peaks. In the literature, there are two reported tabulated datasets for 1D lydiamycin A NMR spectral data, both of which were recorded in $CDCl_3$ [318, 319]. The observed carbon and proton spectral data of compound #1 were compared to these datasets (**Appendix Table 25**, **Table 26**). The observed ^{13}C, ^{1}H peak shifts and the ^{1}H multiplicity and coupling constants were nearly identical to the published lydiamycin A data.

Further, HMBC and COSY 2D experiments were conducted (**Appendix Figure 113**) which also proved to be identical to the published lydiamycin A data [319]. Taken together, the MS and NMR data demonstrate that the isolated compound #1 is the same compound previously isolated and characterised as lydiamycin A. As such, we can assign lydiamycin A as a product of the *R. fascians D188nrp* cluster. Therefore, a new lydiamycin A BGC has been identified and *R. fascians* D188 is the first non-streptomycete producer of lydiamycin. The full NMR assignment of lydiamycin A, including HMBC, COSY and 1,1-ADEQUATE cross-peaking, is described in **Table 7**.

Table 7. Lydiamycin A NMR assignment (carbon and proton) and observed crosspeaks (HMBC, COSY, 1,1-ADEQUATE). All NMR recorded in $CDCl_3$. Atom numbering of the revised structure is used (see Figure 61).

Position	Δ_C ppm	Δ_H ppm (multiplicity, *J* in Hz)	HMBC (C to H)	COSY	1,1-ADEQUATE
1	169.8		16, 17		2
2	52.7	5.31, m	3, 4	3	3
3	24.4	2.38, 1.79, m	4, 5, 2		
4	21.5	1.63, 1.59, m	2, 5, 5-NH	3, 5	5
5	47.1	3.16, br d (13.4) 2.78, qd (13.1, 3.5)	3, 4	4, 5-NH	
5-NH		4.33, br d			
6	174.4		2, 5-NH, 7, 7-NH, 8		7
7	50.5	4.86, qd (10.4, 7.3)	7-NH, 8	7-NH, 8	8
7-NH		7.31, d (10.4)		7	
8	18.3	1.46, d (7.3)	7	7, 7-NH	7
9	169.0		7, 7-NH, 10, 11		10
10	51.4	4.35, m	7-NH, 10-NH, 11, 12	10-NH, 11	11
10-NH		7.47, d (8.9)		10	
11	36.5	1.77, m	10, 10-NH, 12	10	10
12	24.6	1.63, m	10		11, 13, 14
13	22.8	0.94, d (6.6)	11, 12	12	
14	22.3	0.89,(6.6)	12		
15	169.4		10, 10-NH, 16, 16-NH, 17		16
16	50.7	5.31, m	16-NH, 17	16-NH, 17	17
16-NH		8.31, d (9.9)		16	
17	68.3	4.79, dd (11.4, 4.3) 4.06, d (11.4)	16	16	16
18	171.0		16, 16-NH, 19, 20		19
19	55.6	4.68, t (5.0)	20, 21	20, 21	20
20	19.4	2.38, 1.95, m	19, 21, 22	19, 21, 22	19
21	20.7	2.17, 1.95, m	19, 20, 22	19, 20, 22	22
22	142.7	6.93, m	20, 21	20, 21	21
23	177.4		19, 24, 25, 30		24
24	46.0	3.05, m	25, 26, 30	25, 30	25, 30
25	31.6	1.75, 1.54, m	24, 30		24
26	26.9	1.37, 1.26, m	24, 25, 27	25	25
27	31.5	1.29, 1.26, m	25, 26	25, 26, 29	28
28	22.4	1.28, m	26	26, 29	29
29	14.0	0.87, t (6.9)	27, 28		28
30	36.6	3.33, dd (17.0, 12.3) 2.63, dd (17.0, 5.5)	24	24	24
31	175.9		30		30
31-OH		11.41, br s			

3.2.4.3 Bioinformatics and Novel NMR data informs Lydiamycin A Structural Revision

A structural inconsistency of lydiamycin A is apparent when the BGCs of the known warhead-containing natural products, actinonin and matlystatin, are considered. Matlystatin B biosynthesis is proposed to proceed with the condensation of the N-hydroxy-2-pentyl succinamic acid warhead moiety to piperazic acid (**Figure 16**). This warhead group is proposed to be derived from a dicarboxylic acid (pentyl succinic acid) which has undergone further modification to generate the terminal hydroxamate group (**Figure 15**). The carboxylic acids have been distinguished as the alpha-proximal and beta-proximal acids based on their proximity to the pentyl side chain. The proposed matlystatin biosynthetic pathway suggests that it is the alpha-proximal carboxylic acid that becomes CoA-activated and therefore biosynthesis proceeds via the condensation of the alpha-proximal carboxylic acid of the warhead to piperazic acid (**Figure 54**). Actinonin is also biosynthesised via condensation of the alpha-proximal carboxylic acid of the warhead at the first condensation module [207, 217].

However, with the identification of the lydiamycin BGC it has become apparent that the existing published lydiamycin A structure is inconsistent with the biosynthetic theory. The published lydiamycin A structure necessitates the CoA-activation of the beta-proximal carboxylic acid of the dicarboxylic acid (pentyl succinic acid) (**Figure 54**). However, this is not supported by the putative biosynthetic capability of the lydiamycin biosynthetic machinery. It is likely that the first three steps of the warhead biosynthetic pathway (**Figure 15**) are identical in matlystatin and lydiamycin biosynthesis and produce a warhead molecular that is CoA-activated at the alpha-proximal carboxylic acid. Further, there are no genes present in the lydiamycin BGC which are predicted to have functionality that could account for CoA-activation of the dicarboxylic acid via the beta-proximal carboxylic acid.

As such, it was proposed that lydiamycin A is instead biosynthesised via CoA-activation of the alpha-proximal warhead carboxylic acid and thus adopts a structure more alike the known warhead-containing natural products, actinonin and the matlystatins. This proposed lydiamycin A structure is illustrated in **Figure 55** along with a new atom numbering system which will be used hereinafter.

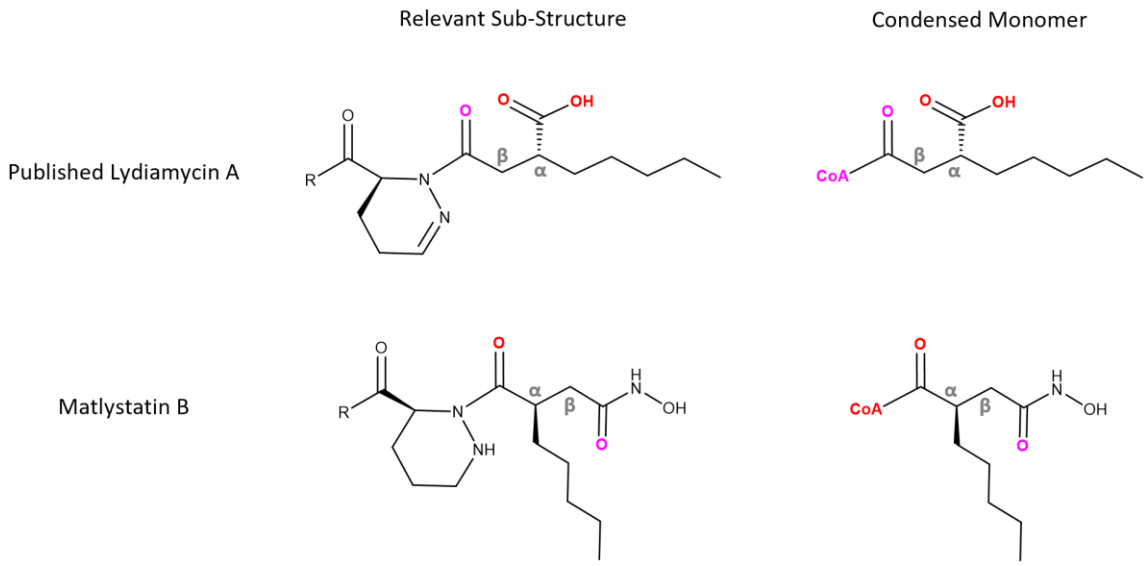


Figure 54. Schematic of structural inconsistency of published Lydiamycin A structure. The alpha-proximal carboxylic acid is highlighted in red and the beta-proximal carboxylic acid is highlighted in purple.

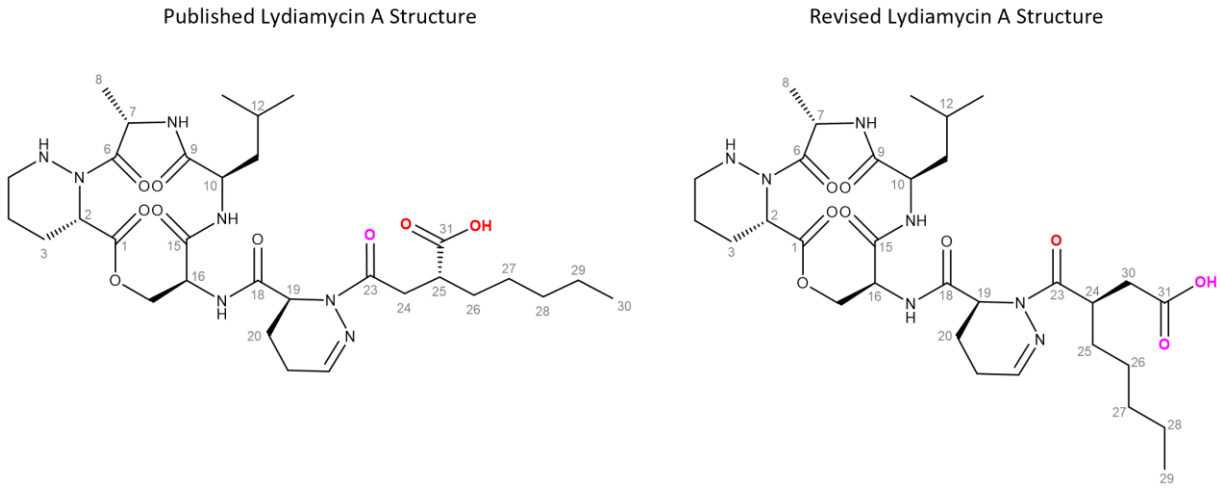


Figure 55. Comparison of published and proposed, revised, lydiamycin A structures. The alpha-proximal carboxylic acid is highlighted in red and the beta-proximal carboxylic acid is highlighted in purple.

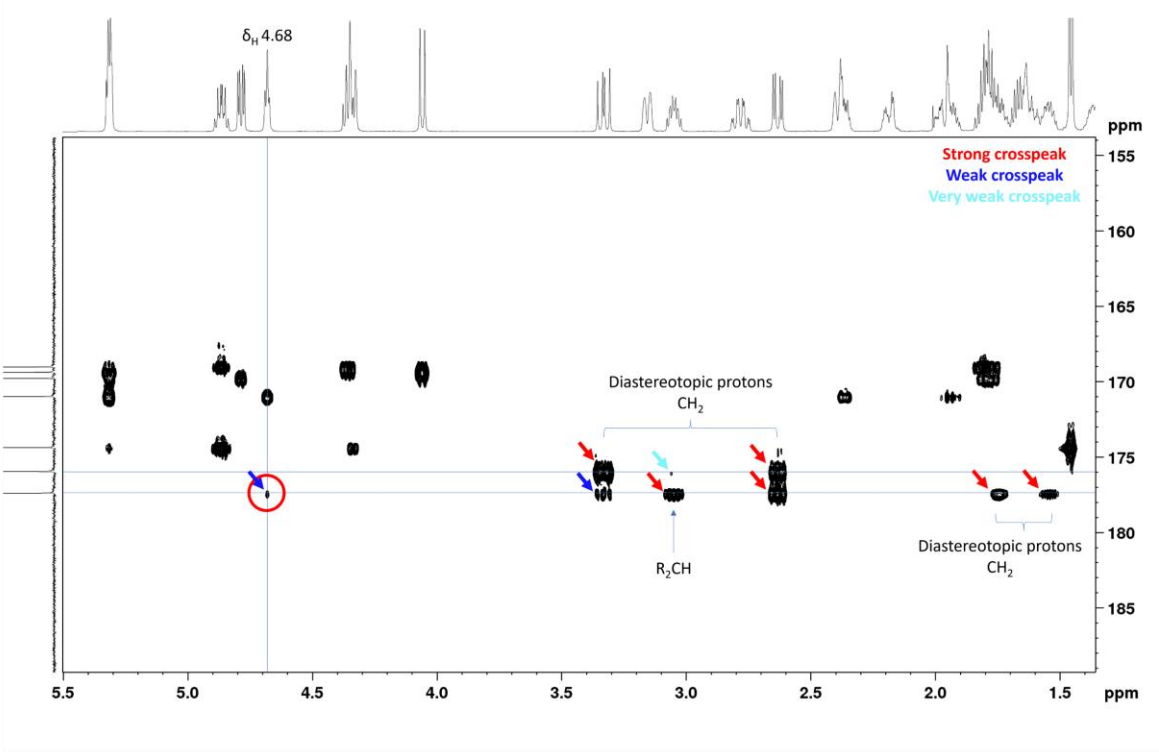
In order to ratify this proposed lydiamycin A structural revision, attempts were made to crystallise the compound and perform X-Ray Diffraction (XRD) analysis, which would provide a full structural and stereochemical determination. However, many iterations of solvents, solute concentrations, temperatures and practical set-ups failed to provide crystals of sufficient quality for structural elucidation by XRD.

Therefore, the proposed lydiamycin A structural revision was ratified by further NMR experimentation. Principally, this problem required the distinction of the closely shifted carbon peaks of the warhead carbonyl groups and determination of their structural environments. The ^{13}C NMR spectrum features seven strongly deshielded carbon peaks (δ_C 169.0 – 177.4 ppm) which are carbonyl carbons. Consideration of the HMBC 2D spectrum (**Table 7**) allowed the assignment of five of these based on proton-carbon connectivity – the δ_C 169.8, 174.4, 169.0 and 169.4 ppm peaks are the carbonyl carbons making up the macrocycle (carbons 1, 6, 9 and 15 respectively). Similarly, the δ_C 171.0 ppm peak is assigned as carbon 18. This leaves two remaining carbonyl carbons (δ_C 175.9 and 177.4 ppm) as the two warhead carbonyl carbons.

A novel HMBC crosspeak was detected (illustrated by the red circle in **Figure 56**) which correlated the δ_C 177.4 ppm carbonyl carbon to proton 19 (δ_H 4.68 ppm), which is a proton within the piperazic acid residue adjacent to the warhead. Therefore, this crosspeak allowed the distinction of the two warhead carbonyl carbon atoms; the δ_C 177.4 ppm carbon as the piperazic-proximal carbonyl and the δ_C 175.9 ppm carbon as the piperazic-distal carbonyl.

Consideration of the relative intensities of the proton-to-carbon crosspeaks in the HMBC data (**Figure 56**) gave some indication as to whether the existing published lydiamycin A structure or the revised structure was correct. The relative intensity of crosspeaks are described in the table within **Figure 56** and may be considered an estimation of the distance between atoms. The piperazic-distal carbonyl carbon (δ_C 175.9 ppm) had crosspeaks with a pair of diastereotopic protons (both strong signals) and an R_2CH proton (very weak signal). The piperazic-proximal carbonyl carbon (δ_C 177.4 ppm) had crosspeaks with proton 19 (weak signal), the same pair of diastereotopic protons (one weak and one strong signal), the same R_2CH proton (strong signal) and a different pair of diastereotopic protons (both strong signals).

The characterisation of these crosspeaks was supported by the identification of diastereotopic protons (which have identical δ_C crosspeaking and coupling constant values). Further, a HSQC experiment was performed to determine the multiplicity of the protons on each carbon (**Figure 57**). The known CH_3 carbons 13 and 14 were used as references and as such, positive values are illustrated in black and relate to a multiplicity of either CH or CH_3 whereas negative values are illustrated in blue and relate to a multiplicity of CH_2 (**Figure 57**). Therefore, the carbon at δ_C 46.0 ppm was confirmed to only have a single proton (R_2CH).



	Observed Crosspeaks			
Piperazic-Distal Carbonyl C (δ_C 175.9)		Diastereo (H, H)	R_2CH (H)	
Piperazic-Proximal Carbonyl C (δ_C 177.4)	Proton 19 (H)	Diastereo (H, H)	R_2CH (H)	Diastereo (H, H)

Strong crosspeak
Weak crosspeak
Very weak crosspeak

Figure 56. Overview of HMBC NMR analysis of lydiamycin A.

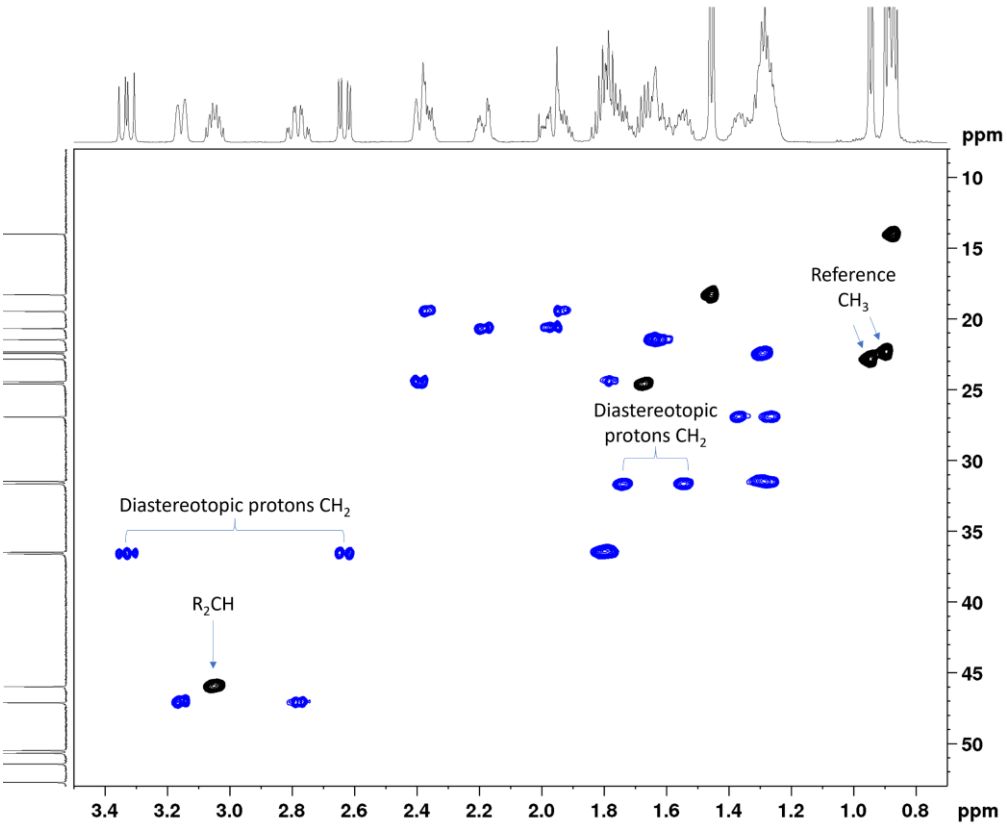


Figure 57. Overview of HSQC NMR analysis of lydiamycin A.

The observed HMBC crosspeaks between the relative protons and the warhead carbonyl carbons (and their relative intensities) was mapped onto both the published and proposed revised lydiamycin A structures (**Figure 58**) to determine which structure best fitted the observed NMR data. When considering the published lydiamycin A structure, the very weak crosspeak between the R_2CH proton and the adjacent piperazic-distal carbonyl carbon was weaker than might be expected. Further, there were inconsistencies with the relative strengths of crosspeaks of diastereotopic protons to the piperazic-proximal carbonyl carbon as crosspeaks through less bonds were much weaker than crosspeaks through more bonds. Also, if the published structure was correct, it would be expected that the pair of diastereotopic protons on the pentyl chain would interact with the piperazic-distal carbonyl carbon, but these crosspeaks were not observed.

In contrast, the revised lydiamycin A structure fits the observed HMBC correlations (**Figure 58**) better as the relative crosspeak strengths are better correlated with the number of bonds between atoms. Therefore, this analysis suggests that the revised lydiamycin A structure is correct, and that the warhead is condensed to piperazic acid via the alpha-proximal carbonyl. However, this analysis is qualitative, and the result is indicative rather than conclusive. Further NMR analysis was required to corroborate with this finding.

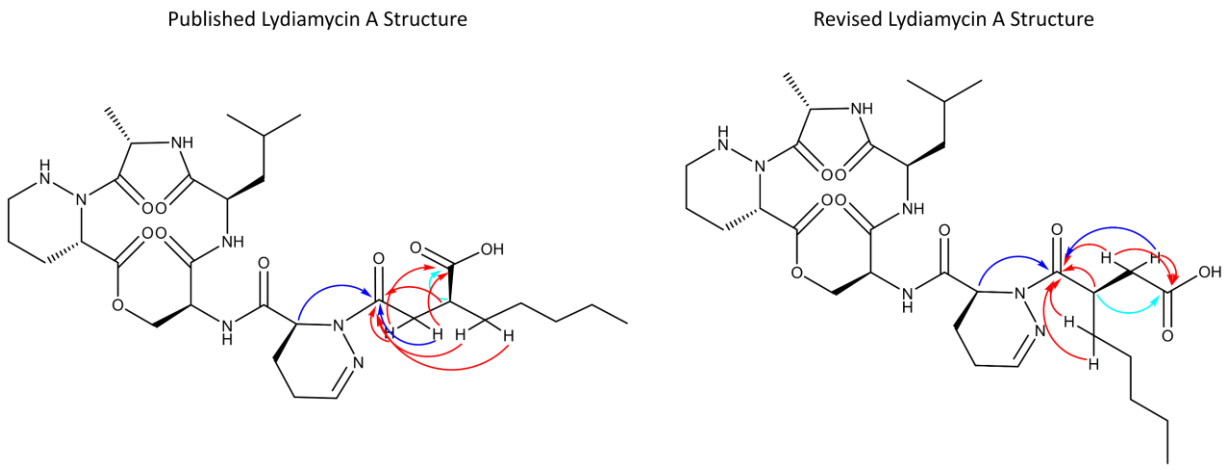
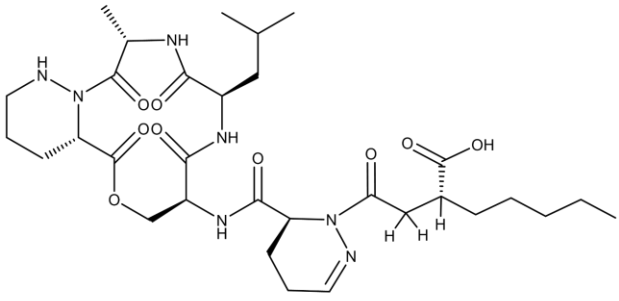


Figure 58. Schematic of correlation of observed HMBC data to the published and proposed lydiamycin A structures.

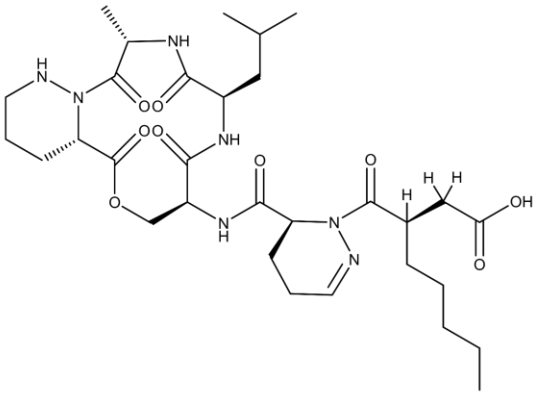
A 1,1-ADEQUATE NMR experiment was performed on lydiamycin A, which is a 2D experiment which correlates between a carbon atom and a proton on a directly adjacent carbon. In this way, data is only acquired between carbons and protons that are two bonds away, which provides more precise detail than HMBC which can correlate over three or four bonds. This experiment should therefore distinguish between the two proposed lydiamycin A structures, as the proton multiplicity of the carbon atoms adjacent to the warhead carbonyl carbons differs between the two structures. In the published lydiamycin A structure, the carbon adjacent to the piperazic-acid proximal carbonyl carbon has two protons and the carbon adjacent to the piperazic-acid distal carbonyl carbon has a single proton. By contrast, in the revised lydiamycin A structure this proton multiplicity is reversed (**Figure 59**).

Analysis of the 1,1-ADEQUATE NMR data (**Figure 60**) indicated that the piperazic-distal carbonyl carbon (δ_C 175.9 ppm) had crosspeaks with a pair of diastereotopic protons (δ_H 3.33, 2.63 ppm) on the adjacent carbon. The piperazic-proximal carbonyl carbon (δ_C 177.4 ppm) was found to have a crosspeak with a single proton at δ_H 3.05 ppm. Cross-referencing this with the HSQC proton multiplicity data (**Figure 60**) is consistent with the pair of diastereotopic protons having CH_2 multiplicity and the proton at δ_H 3.05 ppm having R_2CH multiplicity. Therefore, the observed 1,1-ADEQUATE data is contrary to the published lydiamycin A structure and supports the revised lydiamycin A structure.

Published Lydiamycin A Structure

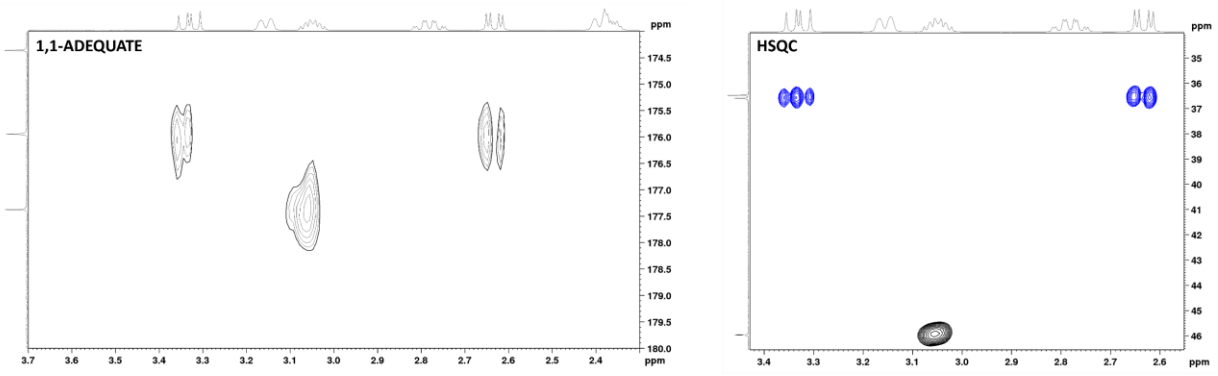


Revised Lydiamycin A Structure



	Proton Multiplicity of Adjacent Carbon	
	Piperazic-Distal Carbonyl C (δ_C 175.9)	Piperazic-Proximal Carbonyl C (δ_C 177.4)
THEORETICAL: **Published** Lydiamycin A	Single (CH)	Double (CH_2)
THEORETICAL: **Revised** Lydiamycin A	Double (CH_2)	Single (CH)

Figure 59. Comparison of expected 1,1-ADEQUATE NMR distinction of published and revised lydiamycin A structures.



	Proton Multiplicity of Adjacent Carbon	
	Piperazic-Distal Carbonyl C (δ_C 175.9)	Piperazic-Proximal Carbonyl C (δ_C 177.4)
THEORETICAL: **Published** Lydiamycin A	Single (CH)	Double (CH_2)
THEORETICAL: **Revised** Lydiamycin A	Double (CH_2)	Single (CH)
OBSERVED	Double (CH_2)	Single (CH)

Figure 60. Overview of 1,1-ADEQUATE NMR analysis of lydiamycin A.

In conclusion, lydiamycin A has been identified as the product of the *R. fascians D188nrp* cluster. However, consideration of the proposed biosynthetic pathway of the newly discovered lydiamycin A BGC identified a structural inconsistency, and a revised structure was proposed. A novel HMBC crosspeak allowed for the distinction of key warhead carbonyl carbons, and consideration of the relative HMBC crosspeak intensities was more supportive of the revised lydiamycin A structure. The 1,1-ADEQUATE and HSQC NMR experiments were crucial for the characterisation of the immediately adjacent carbon atoms and their proton multiplicity, which is supportive of the revised lydiamycin A structure (**Figure 61**).

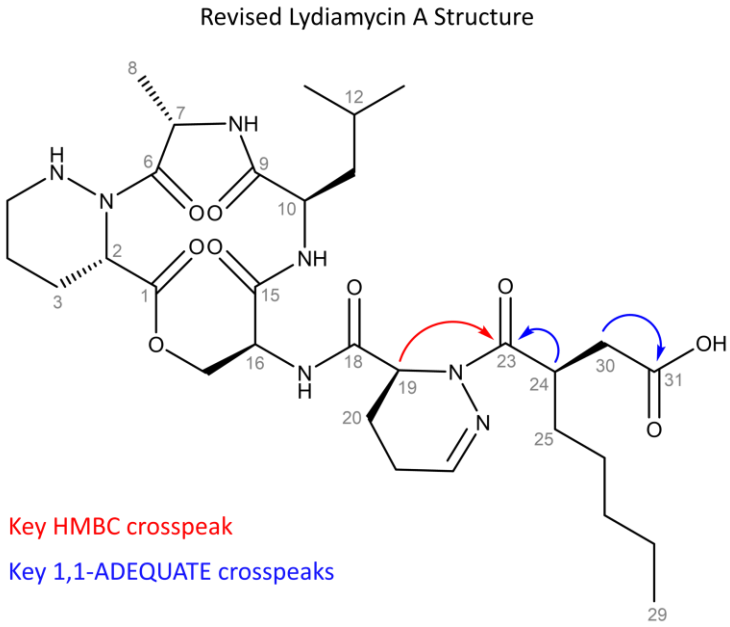


Figure 61. Revised lydiamycin A structure with key NMR crosspeaks highlighted.

3.2.4.4 Assigning stereochemistry using published PGME data

Having revised the structure of the warhead portion of lydiamycin A, the stereochemistry at C24 was evaluated by a colleague, Dr. Edward Hems (John Innes Centre). In the work of Hwang et al., they had assigned this stereocentre based on phenylglycine methyl ester (PGME) derivatisation of lydiamycin A [319]. This methodology requires the comparison of ^{1}H NMR spectra of the (*S*)-PGME and (*R*)-PGME amides of lydiamycin A. Edward Hems reanalysed these published spectra in relation to the revised lydiamycin A structure and determined that the stereochemistry at C24 is (*R*) (**Appendix Figure 114,** [320]), as shown in **Figure 61**. This is consistent with the proposed biosynthetic logic as the proposed warhead biosynthesis pathway (**Figure 15**) would result in an *R* stereocentre.

3.2.4.5 Potential Lydiamycin A congeners from *Rhodoccocus fascians* D188

Earlier manual LC-MS comparison of the *R. fascians* D188 WT and ΔNRP strains led to the identification of twenty potential *D188nrp* cluster products (**Section 3.2.2**), the major product of which was identified as lydiamycin A. However, this metabolomic analysis had limited scope as any compounds that ionised weakly or otherwise produced small base peak chromatograms would have been overlooked during the manual comparison. Further, this analysis only provided precursor *m/z* value and retention time information rather than any more robust structural information.

In order to enhance this metabolomic analysis, new *R. fascians* cultures were subject to mass spectral networking analysis. *R. fascians* D188 WT and Δ*nrp* strains were grown in both SM12 and YEB media for five days and extracted. The extracts were analysed by LC-MS-MS, which was submitted for mass spectral networking analysis. This technique networks structurally related metabolites together based on shared MS-MS fragmentation patterns and has the advantage of visualising potential lydiamycin BGC products as the colour of each node indicates whether the compound was detected in the WT or Δ*nrp* extracts. Media only samples were used as negative controls to identify metabolites not associated with *R. fascians*.

The full results of the mass spectral networking analysis are shown in **Figure 62**. Each node represents a compound precursor mass and the width of each edge is proportional to the structural relatedness of each compound. The colour of each node indicates which strain is producing the compound (red indicates production by the WT strain and blue indicates production by the Δ*nrp* strain). The size of each node is representative of the relative

amount of each compound in the extracts. Overall, the *R. fascians* metabolites network into four major networks with a great number of smaller networks. Networks one and two contain compounds which are produced in roughly equal amounts by both the WT and Δ*nrp* strains and are therefore unrelated to the lydiamycin cluster. Whereas networks three and four contain some metabolites that are only produced by the WT strain and never by the Δ*nrp* strain (**Figure 62**). This is suggestive that the metabolites within networks three and four may be products of the lydiamycin BGC.

Networks three and four were investigated further and the precursor *m/z* value of each node labelled (**Figure 63**). Lydiamycin A is the most abundant compound in network three (denoted with an asterisk) and the wide edge thickness indicates that lydiamycin A shares a considerable degree of structural similarity to the other compounds in the network. The presence of highly structurally similar compound which are detected only in the Δ*nrp* extracts (blue nodes) may be indicative of shunt metabolites of lydiamycin biosynthesis given that the NRPS only has a deletion in a single A domain.

Network four appeared to consist mainly of sodiated adducts of metabolites found in network three. For instance, the major node of network four has a *m/z* value of 686.346 which corresponds to a sodiated adduct of lydiamycin A $[M+Na]^+$ (**Figure 63**). As such, network three is likely to consist of the lydiamycin cluster products, detected as protonated ions.

It is interesting that lydiamycin A is highly structurally related to a number of compounds that are absent when the BGC is disrupted. This is another indication that *R. fascians* may produce a number of structurally related lydiamycin congeners. Consideration of the mass differences between compounds can suggest structural variances. For instance, the 678.381 metabolite may be a methylated version of lydiamycin A (M+14). However, isolation and structural characterisation of these metabolites is required to confirm and determine where the modification is introduced.

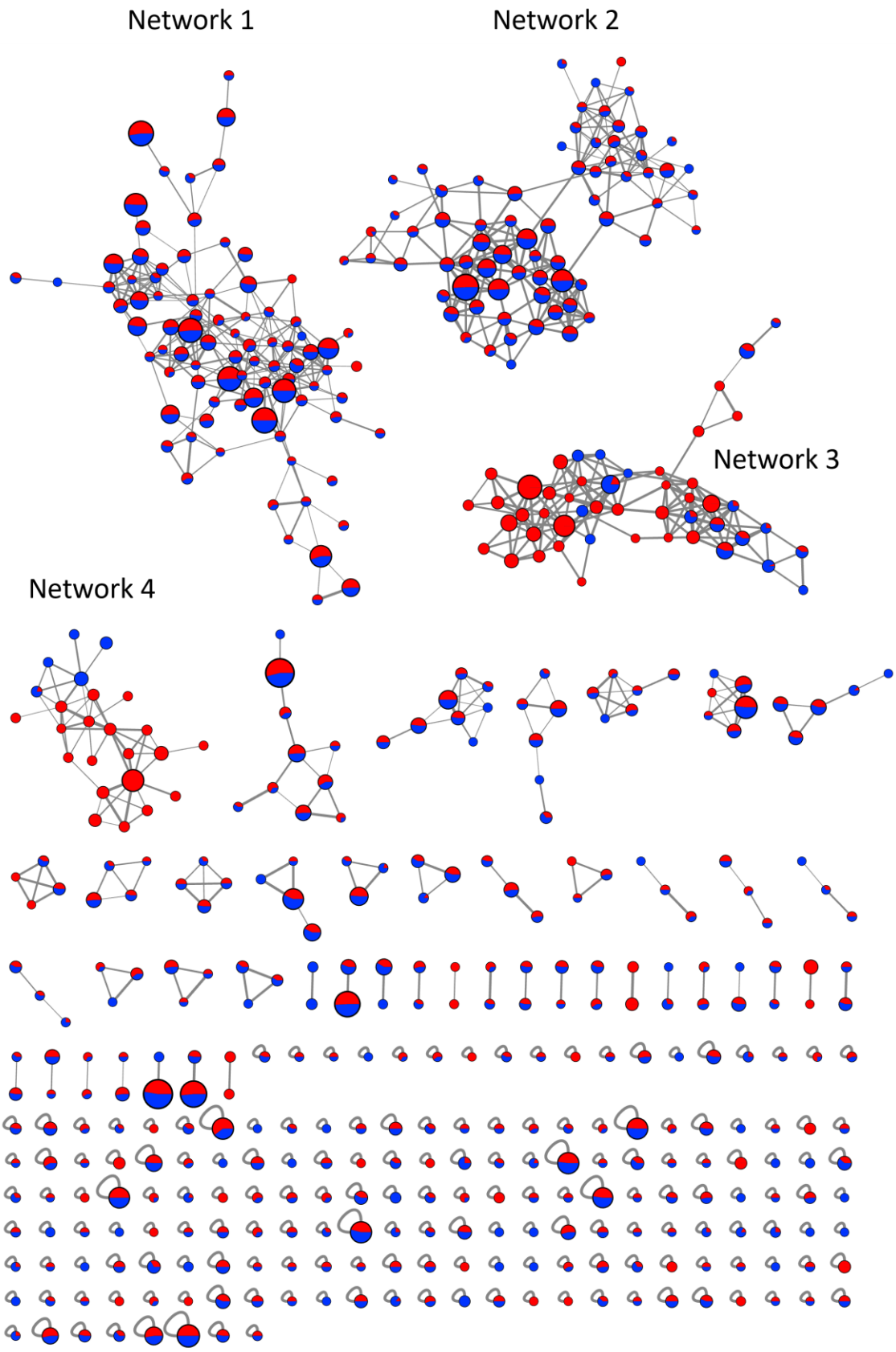


Figure 62. Overview of *R. fascians* metabolomic spectral networks.

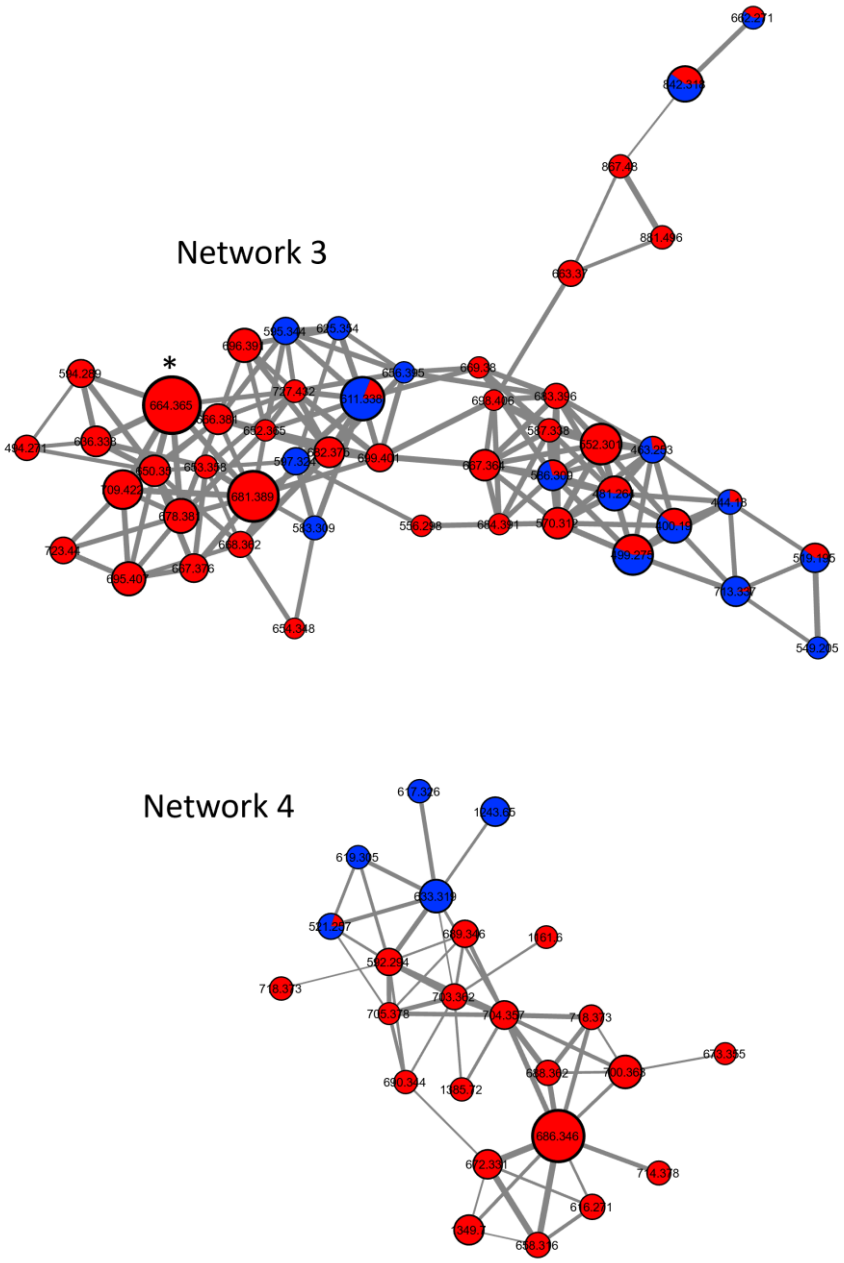


Figure 63. *R. fascians* metabolomic spectral networks that are enriched with potential D188nrp products. Lydiamycin A is denoted with an asterisk.

This spectral networking analysis cannot be directly compared to the previous manual comparative metabolomic analysis as these analyses have different scopes. The manual analysis was performed with nine production media across two timepoints, whereas the spectral networking is a snapshot of production in two media at a single time point. Further work would be required to fully characterise the metabolomic profile and induction conditions of *R. fascians*. Nevertheless, these analyses indicate that *R. fascians* produces lydiamycin A and a number of structurally related metabolites which may represent novel lydiamycin congeners.

3.3 Discussion and conclusions

3.3.1 Summary / Discussion of Results

This investigation began with the bioinformatic analysis of actinomycete genomes in order to identify putative warhead-NP BGCs. This resulted in the identification of a single small clade of sequences which had a co-association with putative warhead biosynthesis genes and the necessary biosynthetic machinery for peptide backbone biosynthesis. There were several potentially new genera of warhead-NP producers and this bioinformatic approach was very effective for the identification of new compound leads. As this approach only requires knowledge of one pathway-specific gene and some of the likely accessory biosynthetic genes, it may be used for other projects and the scope could be increased to include non-actinomycetes. However, it is not possible to rely entirely on the results of the co-association analysis as several sequences are present at the end of contigs (denoted by a grey star in **Figure 38**) and there have been instances of warhead-containing natural products that are biosynthesised using genes outside of the BGC (for instance the mutase small subunit in actinonin biosynthesis [207]). Therefore, the true complement of co-associated genes of interest may be larger.

It is also interesting that the co-association of these putative warhead-NP biosynthesis genes were constrained to one clade of mutases, which could indicate that these mutases are specific for warhead-NP biosynthesis. This phylogenetic association with actinonin, matlystatin and lydiamycin suggests that they are hexylmalonyl-CoA specific whereas most of the other sequences in the phylogenetic tree are likely ethylmalonyl-CoA or methylmalonyl-CoA mutases.

The *R. fascians D188nrp* cluster was selected for further investigation as the bacterium is an important plant pathogen and it was proposed that the putative warhead-NP may be involved in pathogenesis. Comparative metabolomic analysis led to the isolation of the major *D188nrp* product, lydiamycin A. Lydiamycin A was first isolated in 2006 as an antimycobacterial agent produced by *Streptomyces lydicus* [318]. In this initial report, the macrocyclic pentapeptide structure with a pentylsuccinic acid tail was deduced by MS and NMR analysis. Further, three congeners (lydiamycin B, C, D) were identified which had varied hydroxylation modifications and state of desaturation in the backbone piperazic acid residue. However, questions remained about the chemistry of the molecule as the stereochemistry of the pentyl carboxylic acid and the adjacent piperazic acid was not initially determined.

There were subsequently multiple attempts to clarify the stereochemistry of lydiamycin A and B by the synthesis of their diastereomers [321, 322]. However, none of the diastereomers had identical NMR data to their isolated natural counterparts, which was indicative of incorrect stereochemical assignment in the initial report. It was not until 2020 that the work of Hwang *et al*. rectified the incorrect stereocentre assignment and determined the remaining two stereocentres [319]. An advanced Marfey's analysis revised the absolute configuration of C-2 from *R* to *S*, which explained why the above diastereomer synthetic attempts failed. The identity of the unassigned stereocentres was determined – the terminal piperazic acid by advanced Marfey's method and the acyl chain stereochemistry was determined by derivatisation of the free carboxylic acid with known isomers of PGME and comparison of the proton NMR shifts to the natural compound [319, 320].

As such, there have been a number of studies into the structure and stereochemistry of the lydiamycins, most of which have been from a synthetic chemistry viewpoint. However, the identification of the lydiamycin BGC and prediction of a likely biosynthetic pathway has suggested that the published lydiamycin A structure is unsupported by biosynthetic logic. A revised lydiamycin A structure was proposed where the pentylsuccinic acid group is installed via alpha-proximal condensation. Extensive NMR experimentation supports the revised structure. As such, consideration of the biosynthetic potential of the BGC is presented as an important tool which may inform structural nuances that are invisible by routine chemical derivatisation or NMR analysis.

The identification of the lydiamycin BGC and the structural revision has allowed the proposal of a lydiamycin A biosynthetic pathway (**Figure 64**). The warhead biosynthesis proteins are proposed to biosynthesise the carboxylated warhead. It is proposed that the warhead is CoA-activated at the alpha-proximal carboxylic acid, which is condensed to piperazic acid by LydL. It is proposed that subsequent NRPS modules of LydH condense a serine residue, followed by a leucine, alanine and piperazic acid. The LydE and LydF proteins biosynthesise the nonproteinogenic piperazic acid residues. Finally, it is predicted that there is thioesterase-catalysed cyclisation of the peptide to generate mature lydiamycin A.

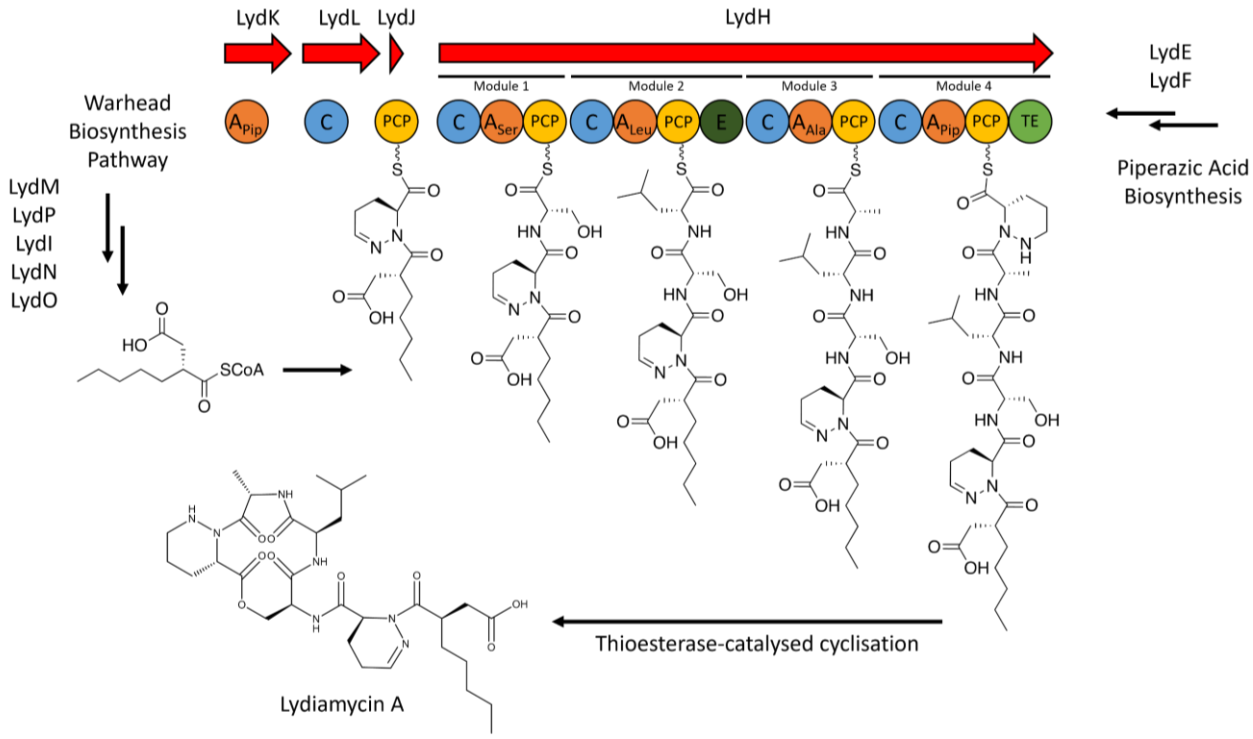


Figure 64. Proposed lydiamycin A biosynthesis pathway. Abbreviations: Adenylation (A), Condensation (C), Peptide carrier protein (PCP), Thioesterase (TE).

This works represented the first identification of the lydiamycin BGC. However, during the course of this work, a very similar BGC was reported by Libis *et al*. in 2022 [323] whilst investigating the large-scale expression of BGCs for natural product identification. However, Libis *et al*. did not report on the structural revision of lydiamycin A. Interestingly, their lydiamycin cluster was from *Streptomyces aureoverticillatus* strain JCM 4347 – a strain which was identified in the actinomycete bioinformatic analysis in this work (**Figure 38**). Comparison of the confirmed lydiamycin producing clusters, *S. aureoverticillatus* and *R. fascians* (**Figure 65**), indicate that they are highly similar, however the *S. aureoverticillatus* cluster is lacking the predicted regulator, prenyltransferase, MbtH-like and epimerase genes. The lack of an epimerase is possibly an artifact of sequencing or cluster boundary determination, as it would not be possible to generate the warhead without it. This comparison also suggests that either the regulator and prenyltransferase genes are non-essential or possibly that they contribute to product variation, as it is not reported that the *S. aureoverticillatus* cluster produces any potential structural congeners.

Two of the strains present in the same warhead sub-clade (denoted by an asterisk in **Figure 38**) have now been confirmed to produce lydiamycin A. As such, it is possible that this subclade represents producers of lydiamycin as they also share similar co-association profiles – including the lack of hydroxamate-maturation genes (asparagine synthetase and N-oxygenase). The cluster of another member of this sub-clade, *Streptomyces venezuelae* ATCC 14583, was found to be very similar to the known lydiamycin BGCs (**Figure 65**) and found to feature identical NRPS architecture to the *R. fascians* and *S. aureoverticillatus* clusters (**Appendix Figure 115**). This supports the hypothesis that this sub-clade may consist of lydiamycin producers, however this would have to be determined experimentally.

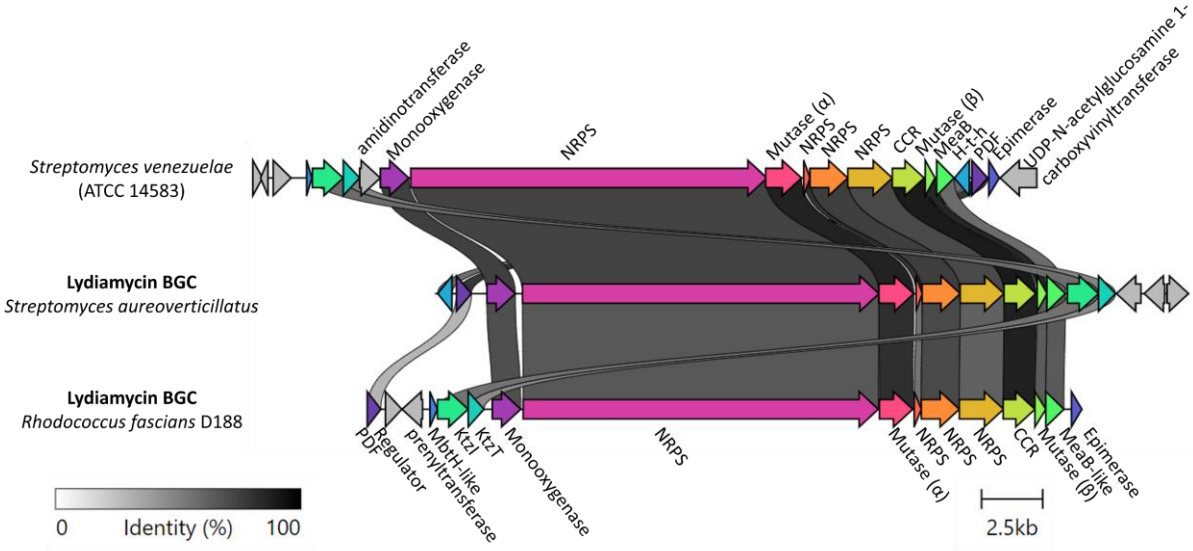


Figure 65. Comparison of known lydiamycin BGCs and proposed lydiamycin cluster from *S. venezuelae*. Sequence similarity is denoted by greyscale linker. *S. aureoverticillatus* cluster is described in [323].

Comparative metabolomics and spectral networking analyses have identified a number of metabolites that are structurally related to lydiamycin A and their production is completely absent in the Δ*nrp* strain. This is suggestive that *R. fascians* may produce a number of lydiamycin congeners, although these metabolites need to be isolated and characterised to confirm this. Interestingly, the known lydiamycin B, C and D congeners are not present in the culture extracts (as protonated or sodiated species). This is suggestive that different lydiamycin BGCs may have different tailoring enzymes, NRPS specificity promiscuity or regulatory circuits that may determine the possible lydiamycin congener chemical diversity. This is supported by the lydiamycin BGC identified in *S. aureoverticillatus* having a different range of genes in the predicted cluster boundary.

This work highlights the importance of investigating known BGCs in different producers, as they may feature different congener potential. The identification of homologous core biosynthetic machinery with varied tailoring capacity has been exploited in the literature. Morgan and colleagues identified novel congeners of the lipodepsipeptide ramoplanin by identification of related BGCs and led to the identification of chersinamycin from the actinomycete *Micromonospora chersina* [324]. The examination of a conserved BGC found in three fungal species resulted in the identification of structurally diverse cyclic depsipeptides from the beauveriolide family [325]. This relatively underexplored approach may allow the systematic screening of novel BGCs homologous to known clinical natural products, perhaps enhancing the chemodiversity of congeners available for SAR studies and pharmacological optimisation.

3.3.2 Future Perspectives

It would be interesting to gain more information as to the chemical diversity potential of the *R. fascians* lydiamycin cluster. This may involve performing additional fermentations using varied production media and sampling timepoints and analyse the metabolite content of extracts using the spectral networking methodology. This would enhance our understanding of putative lydiamycin congeners. This could be expanded to include other secondary metabolite induction techniques, for example supplementation with chemical induction agents, engineering the cluster regulatory system or supplementation with plant exudates.

Work is ongoing to isolate some of the potential lydiamycin congeners. Once purified, these can be structurally characterised (which should be aided by the above NMR assignment work), and their bioactivity assessed.

It would be beneficial to perform some more genetic experimentation – primarily to attempt to discern the cluster boundaries either by mutation in *R. fascians* or heterologous expression of different sections of the cluster. Further, work is ongoing to express the late-stage warhead maturation genes (asparagine synthetase and N-oxygenase) in *R. fascians* D188. This may lead to the isolation of a hydroxamated lydiamycin compound which has the potential to exhibit greater bioactivity than lydiamycin A.

Chapter 4: Lydiamycin A Bioactivity

4.1 Introduction

In the previous chapter the plant pathogen *R. fascians* was found to produce the warhead-NP lydiamycin A and the BGC was identified. Based on biosynthetic logic and novel NMR data, the structure of lydiamycin A was revised. The only published bioactivity of lydiamycin A is as an antimycobacterial agent. However, the question still remains as to the physiological role of lydiamycin in the arsenal of *R. fascians.*

It is possible that lydiamycin A may have a role in plant colonisation or leafy gall disease development. This notion is supported by the fact that the lydiamycin BGC is present on the pathogenicity-associated pFiD188 megaplasmid along with the known genetic determinants of leafy gall disease (see **Section 3.1.2**). Further, the presence of a copy of peptide deformylase (PDF) in the lydiamycin BGC may represent a self-immunity determinant. The fact that the structurally related warhead-NP, actinonin, is the most potent natural PDF inhibitor (PDFI) (see **Section 1.6**) supports this. These potential lydiamycin bioactivities will be tested in this chapter.

4.1.1 The Reported Bioactivity of Lydiamycin A

Lydiamycins A-D were isolated in 2006 from *Streptomyces lydicus* strain HKI0343 and found to be effective antimycobacterial agents [318]. Lydiamycin A was reported to selectively inhibit a number of mycobacteria (*M. smegmatis* SG 987, *M. aurum* SB66 and *M. vaccae* IMET 10670), with minimum inhibitory concentration (MIC) values of 25, 6.3 and 3.1 µg/mL respectively [318]. Further, lydiamycin A was tested against strains of *M. tuberculosis*, including the wild type H37Rv strain and a multi-drug resistant clinical isolate, and produced MICs of 12.5 and 25 µg/mL respectively [318]. The minimum bactericidal concentration (MBC) of lydiamycin A was also determined as the concentration required for complete inhibition of growth against the above *M. tuberculosis* strains and was double the MIC values. These assays were performed using solid agar techniques. Further, the lydiamycins were found to be non-toxic to human cell lines [318].

However, a later study reported that the antimycobacterial potency of lydiamycin A was much weaker than first indicated. The work of Hwang and colleagues used a microplate-based liquid medium assay to determine the antimycobacterial activity against *M. tuberculosis* mc^2 6230. Mycobacterial growth was plotted against lydiamycin A concentration in order to calculate MIC_{50} and MIC_{90} values (drug concentrations at which 50% and 90% of *M. tuberculosis* had completely arrested growth). MIC_{50} and MIC_{90} values were determined to be 31 μg/mL and 286 μg/mL, respectively, and complete killing of the mycobacteria was never observed (MBC >300 μg/mL) [319].

Lydiamycin A has also been tested for other bioactivities. No anti-cancer properties were identified against a screen of cancer cell lines, including gastric, liver, lung, colon and breast cancer [319]. Further, lydiamycin A did not exhibit antimicrobial activity against a screen of pathogenic bacteria (*Staphylococcus aureus*, *Enterococcus faecalis*, *Klebsiella pneumonia*, *Salmonella enterica* and *E. coli*) nor pathogenic fungi (*Candida albicans*, *Aspergillus fumigatus*, *Trichophyton rubrum* and *Trichophyton mentagrophytes*) [319].

Therefore, the major reported bioactivity of lydiamycin A is as an antimycobacterial agent, however, there remain inconsistencies in the field as to the potency of this activity. Further, the exact molecular target of lydiamycin was unknown. Overall, previous reports have presented the lydiamycins as interesting molecules that may represent a novel platform for antimycobacterial drug development as they are relatively potent and human cytotoxicity has not been observed.

The lydiamycins are cyclic lipodepsipeptides (CLPs), which indicates that they are macrocylised (cyclic), they feature fatty acid lipid tails (lipo-) and that at least one of the amide bonds in the peptide backbone are replaced by an ester linkage (depsipeptide). The two major drives of CLP investigation is their uses as antimicrobials effective against multi-drug resistant pathogens and their production by biocontrol strains for the control of plant diseases [326]. For instance, the CLP daptomycin is produced by *Streptomyces roseosporus* and has been approved for the treatment of complicated skin and soft tissue infections [327]. Daptomycin is effective against a number of multi-drug resistant organisms, including methicillin-resistant *S. aureus* (MRSA), vancomycin-resistant enterococci (VRE) and penicillin-resistant *Streptococcus pneumoniae* [328]. In terms of the agricultural importance of CLPs, biocontrol strains are known to produce CLPs that confer antimicrobial and antifungal bioactivity [329]. For instance, the pseudophomins are CLPs produced by *Pseudomonas fluorescens* that are active against phytopathogenic fungi [330] and the antimicrobial activity of a viscosin-like CLP produced by *Pseudomonas* sp. is explored in the next chapter.

The range of reported CLP bioactivities is also expanding, with examples of biosurfactants, herbicides, insecticides, anticancer agents and immunosuppression agents [331-334]. Therefore, this wide range of CLP bioactivities indicates that it is important to thoroughly test the bioactivities of CLPs.

4.1.2 Chapter Aims

Lydiamycin A was first reported as an antimycobacterial agent. However, this work presents the first identification of the lydiamycin BGC, *Rhodococcus fascians* as a producer and the revised structure of lydiamycin A. In this chapter, the antimycobacterial activity of lydiamycin A will be tested. Further, several more exploratory and broader bioactivities of the lydiamycins will be tested. To achieve this, the specific aims of this chapter are:

i. Determine the antimycobacterial activity of lydiamycin A
ii. Assess whether lydiamycin A inhibits PDF
iii. Assess whether lydiamycin production plays a role in *R. fascians* plant colonisation or leafy gall disease progression

4.2 Results

4.2.1 Quantifying the Antimicrobial Activity of Lydiamycin A.

With purified lydiamycin A to hand, the first bioactivity investigated was the reported antimycobacterial activity. The reporter strain selected was *Mycobacterium smegmatis* mc^2 155 as this species had been used in the initial report, where the MIC of lydiamycin A was determined as 25.0 µg/mL [318]. Further, *M. smegmatis* is fast-growing and non-pathogenic and is widely used as a model system to study *M. tuberculosis* [335] so may allow wider parallels to be drawn from the results.

In order to estimate the MIC of lydiamycin A against reporter strains, the spot-on-lawn assay was used (see materials and methods for details). In brief, different concentrations of lydiamycin A were spotted onto solid agar inoculated with the reporter strain. The plate was incubated until growth was observed and the MIC was determined as the concentration at which growth is absent around the applied compound. A spot-on-lawn assay was performed with *M. smegmatis* challenged with lydiamycin A. However, the MIC of lydiamycin A against *M. smegmatis* proved greater than the highest concentration used of 600 µg/mL (**Figure 66**).

In order to determine whether *E. coli* was susceptible to lydiamycin A, another spot-on-lawn assay was performed. The *E. coli* NR698 strain was selected as it features a modified outer membrane that is more permeable to extracellular compounds and therefore represented the *E. coli* strain that was most likely to be susceptible to lydiamycin A [336]. *E. coli* NR698 was challenged with lydiamycin A (**Figure 67**), however proved resistant against the highest concentration tested (MIC > 600 µg/mL).

Therefore, in our lab, the spot-on-lawn assay was not suitable for quantifying the antimicrobial activity of lydiamycin A as the reported antimycobacterial activity was not being replicated.

Figure 66. Spot-on-lawn assay of lydiamycin A against *M. smegmatis* mc^2155. Lydiamycin A concentration is denoted in µg/mL. Negative control spot of methanol (diluent solvent) and positive control spot of 50 µg/mL kanamycin are indicated. *M. smegmatis* is grown on MOADC.

Figure 67. Spot-on-lawn assay of lydiamycin A against *E. coli* NR698. Lydiamycin A concentration is denoted in µg/mL. Negative control spot of methanol (diluent solvent) and positive control spot of 50 µg/mL kanamycin are indicated. *E. coli* NR698 is grown on SNA.

Due to the problems of detecting the antibacterial activity of lydiamycin A on solid medium bioassays, an assay was developed for the comparison of growth curves to assess the antimicrobial activity of lydiamycin A. In brief, the *E. coli* NR698 or *M. smegmatis* mc^2155 strains were grown in triplicate in 96-well plates, with and without challenge with lydiamycin A (see materials and methods for details). Several preliminary experiments were performed in order to determine the minimum inoculum required which supported normal growth, as this would offer maximal sensitivity to the addition of an antimicrobial agent. This was also tested in parallel with the addition of solvents required for the solubilisation and addition of lydiamycin A. Previous experimentation has indicated that lydiamycin A is soluble in MeOH and DMSO, therefore reporter strain sensitivity towards 5% v/v of these two solvents were also tested. These preliminary experiments indicated that MeOH was tolerated far better than DMSO and that the optimal starting inoculum of both the reporter strains were O.D. 0.05 (**Appendix Figure 116, Figure 117**). These experiments also suggested that *M. smegmatis* may be the favoured reporter strain as it has greater resistance to 5% MeOH, as compared to the *E. coli* strain, and therefore it may be easier to delineate the effects of lydiamycin A from the solvent using *M. smegmatis*.

Having demonstrated that methanol was a suitable low-toxicity lydiamycin diluent for growth curves, the two reporter strains were challenged with lydiamycin A. In brief, both reporter strains were grown in triplicate using a starting inoculum concentration of O.D. 0.05. Strains were grown in the presence of 5% v/v methanol or 50 µM lydiamycin A (at 5% v/v methanol). This concentration of lydiamycin A was chosen as it is approximately equal to the MIC and MIC_{50} against *Mycobacteria* reported in the literature (38 and 47 µM respectively [318, 319]). Growth curves were visualised by a solid line representing the median O.D. reading and a shaded envelope representing the minimum and maximum O.D. reading bounds at each time point. This representation of raw data is used hereinafter.

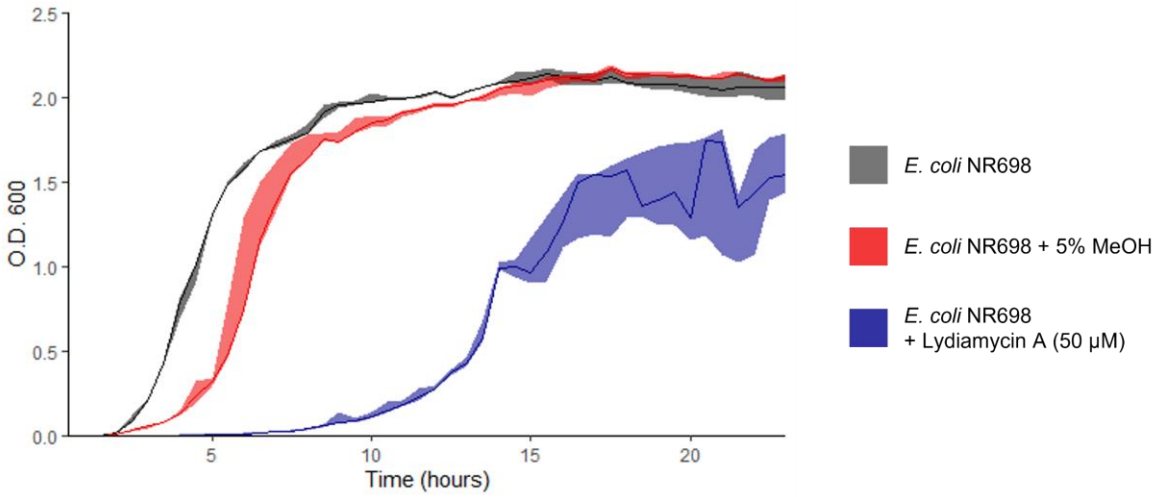


Figure 68. Growth curve of *E. coli* NR698 grown in the presence and absence of lydiamycin A. Figure generated using a R script written by Dr. Alaster Moffat (John Innes Centre). All subsequent growth curve figures were also generated using this script.

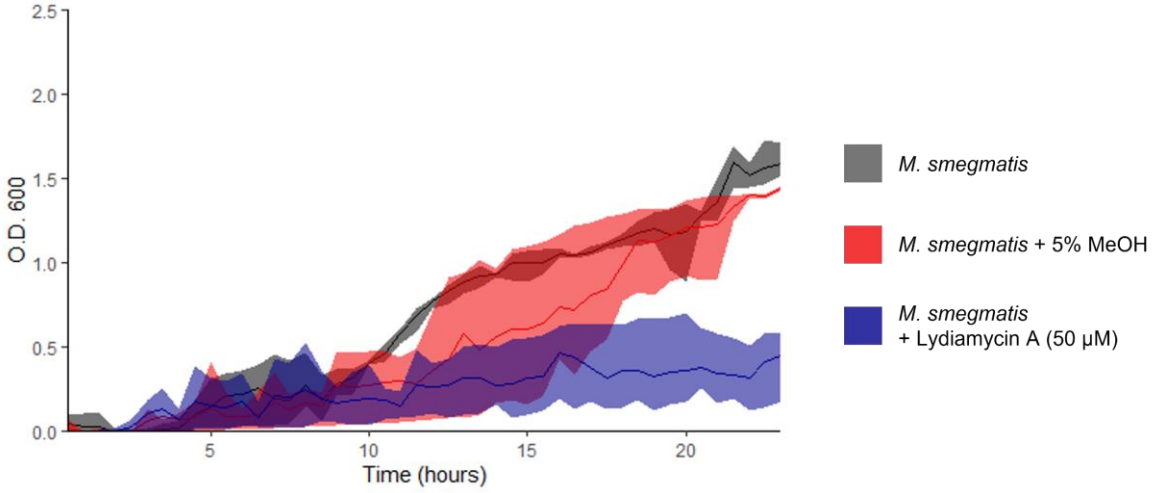


Figure 69. Growth curve of *M. smegmatis* grown in the presence and absence of lydiamycin A.

When *E. coli* NR698 was challenged with 50 µM lydiamycin A (**Figure 68**) the growth was delayed by approximately ten hours and the treated cultures did not reach the same O.D. at stationary phase as the untreated or methanol control cultures. Therefore, lydiamycin A proves to have a moderate antimicrobial effect upon the growth of *E. coli* NR698 at this concentration, however the strain is able to almost entirely recover.

Lydiamycin A also caused a moderate antibacterial effect upon *M. smegmatis* (**Figure 69**). As compared to the methanol treated cultures, those treated with 50 µM lydiamycin A had greatly reduced growth. However, it was clear from this experiment that the *M. smegmatis* growth curves should be run for a longer period as they did not appear to reach stationary phase by 24 hours. Further, there was a large amount of variance in the *M. smegmatis* O.D. readings at each time point. Consideration of the microplates after the experiment

indicated that was caused by aggregation of the bacteria in the wells. Efforts to reduce these drawbacks are addressed in the following section.

Overall, these growth curve experiments indicated that lydiamycin A has a moderate antimicrobial effect upon the *E. coli* NR698 and *M. smegmatis* reporter strains. Further optimisation of the *M. smegmatis* assay was required.

4.2.2 Is LydA an immunity determinant?

The *lydA* PDF gene is located at the edge of the predicted lydiamycin cluster boundary (**Figure 42**) and it is unclear whether it is part of the BGC. However, analysis of the *R. fascians* genome indicated that there are a further two proposed PDF genes present on the chromosome (45% and 40% sequence similarity to *lydA*), which are likely housekeeping copies. Further, structurally related warhead-NPs have been reported to be potent PDF inhibitors [128]. Therefore, it was proposed that *lydA* may represent a self-immunity determinant and that the LydA PDF protein may be resistant to the proposed PDFI activity of lydiamycin A, thus conferring resistance to the producer strain. Unfortunately, we did not have experimental expertise in *Rhodococcus* genetic manipulation in order to knock out *lydA* function in the producer strain. Instead, this hypothesis was tested by cloning the *lydA* gene into *M. smegmatis*, which had been proven to be lydiamycin A sensitive, and determining whether *lydA* expression conferred resistance.

4.2.2.1 Cloning *lydA* into pJAM2

The pJAM2 plasmid [337] was selected for these experiments as it is routinely used for mycobacterium protein expression and has been proven to be stable in *M. smegmatis* [338]. Protein expression is controlled by the pAMI acetamide inducible promotor.

The pJAM2:*lydA* plasmid was generated in *E. coli*. In brief, the *lydA* gene was amplified from *R. fascians* D188 gDNA. This was inserted into the pJAM2 backbone by ligation. The resultant pJAM2:*lydA* vector was verified by sequencing and then electrotransformed into *M. smegmatis* mc^2155. The resultant pJAM2:*lydA* plasmid confers kanamycin resistance and the acetamide-inducible expression of LydA (**Appendix Figure 118**). An empty vector (EV) control was also generated by transforming pJAM2 into *M. smegmatis*.

4.2.2.2 Optimisation of *M. smegmatis* growth curves

The variation observed in previous *M. smegmatis* growth curves had been proposed to be caused by aggregation of the cells in the microtitre wells. In an attempt to reduce this effect, the *M. smeg*:pJAM2(EV) and *M. smeg*:pJAM2(*lydA*) strains were grown in the presence and absence of 0.25% TWEEN-80 surfactant, which has been used in mycobacterial culture in the literature [337, 338]. In addition to this, 5% v/v methanol and 0.2% v/v acetamide were added to the cultures to mimic the addition of solubilised lydiamycin and inducer.

The addition of 0.25% TWEEN-80 reduced the variability of *M. smegmatis* O.D. readings at each time point and thus resulted in a reduced envelope of raw data variance on the plotted growth curve (**Figure 70**). This is likely due to reduced aggregation of *M. smegmatis* in the microtitre plate well, as this was observed to a far reduced degree at the end of the platereader run. Further, the cultures supplemented with TWEEN-80 had a faster growth rate and achieved a greater O.D. value as they reached stationary phase (**Figure 70**). Perhaps this was due to increased planktonic growth and therefore an enhanced bacterial surface area to exchange nutrients and waste products in the culture to support more efficient growth. Following this experiment, all *M. smegmatis* pre-cultures and growth curve microtitre cultures were supplemented with 0.25% TWEEN-80. Further, this experiment confirms that the production of LydA protein did not cause any growth defects as compared to the EV control (**Figure 70**).

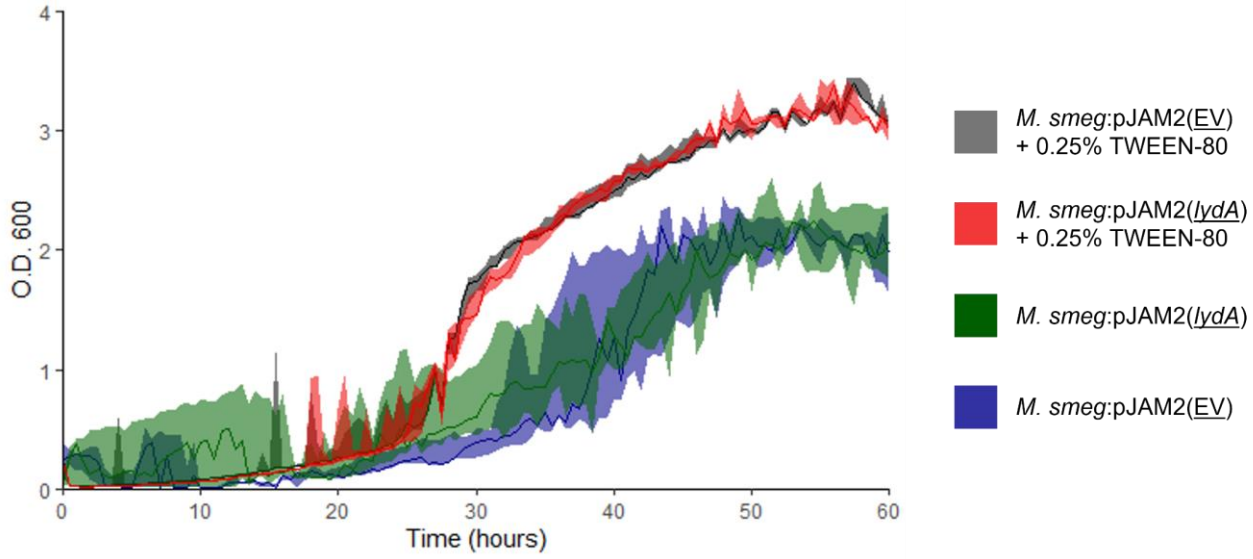


Figure 70. Comparison of absence and presence of TWEEN-80 surfactant on the growth of *M. smegmatis*.

4.2.2.3 Does LydA Confer Lydiamycin A Resistance?

In order to determine whether the expression of *lydA* confers resistance against lydiamycin A, the *M. smeg*:pJAM2(EV) and *M. smeg*:pJAM2(*lydA*) strains were grown in the absence and presence of 50 µM lydiamycin A (**Figure 71**). All cultures were supplemented with 0.2% acetamide and untreated cultures were supplemented with 5% v/v methanol to match the content of the treated cultures.

Consistent with previous experimentation, the *M. smeg*:pJAM2(EV) control strain was sensitive to lydiamycin A as indicated by the blue line in **Figure 71**. However, when challenged with 50 µM lydiamycin A, the *M. smeg*:pJAM2(*lydA*) strain demonstrated resistance against lydiamycin A (indicated by the green line in **Figure 71**). This *M. smegmatis* culture expressing the LydA PDF protein had slightly slower growth compared to untreated cultures but ultimately reached the same culture density. Taken together, this experiment indicated that the LydA PDF protein is able to confer resistance against lydiamycin A. Taken a step further, this result suggested that PDF is a molecular target of lydiamycin A and may represent the mechanism of action of the antimycobacterial activity reported in the literature.

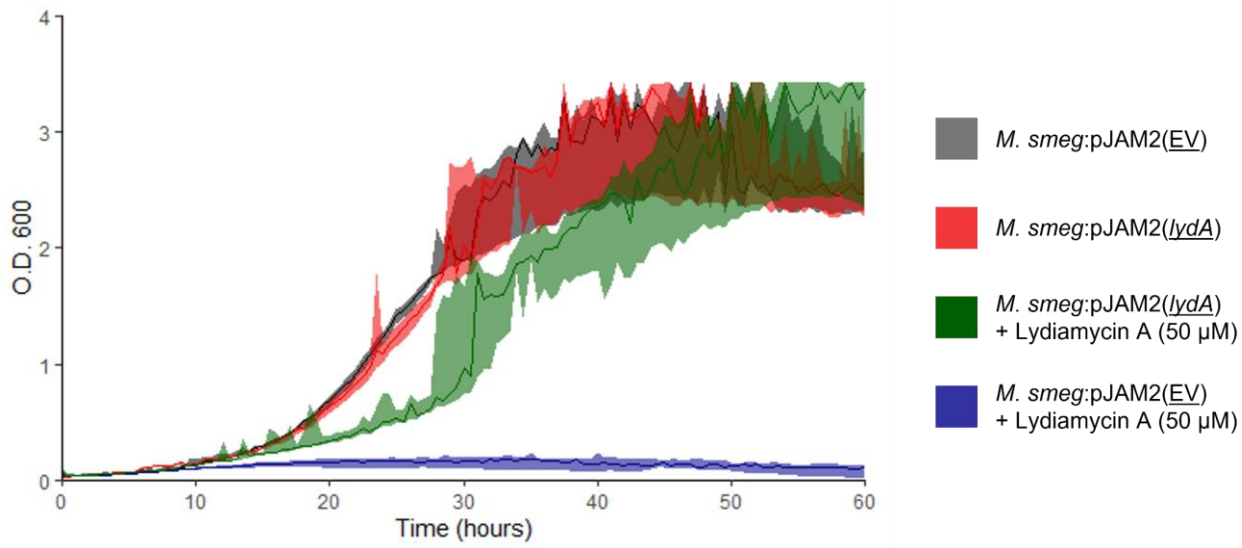


Figure 71. Effect of *M. smegmatis* expression of *lydA* on the sensitivity to lydiamycin A.

4.2.2.4 Is lydiamycin resistance conferred by PDF titration?

To further characterise the observed LydA-mediated resistance against lydiamycin A, an experiment was conducted to determine whether any overexpressed PDF protein could have caused this effect by effectively titrating lydiamycin A from the culture. To this end, the housekeeping *M. smegmatis* PDF gene (*MsPDF*) was cloned into pJAM2 and transformed into *M. smegmatis* to generate the *M. smeg*:pJAM2(*MsPDF*) strain. This specific PDF gene was selected as it was likely to fold correctly during overexpression due to being a native *M. smegmatis* protein.

The *M. smeg*:pJAM2(EV) and *M. smeg*:pJAM2(*MsPDF*) strains were grown in the presence and absence of 50 µM lydiamycin A, all under induction conditions. Again, the EV control strain grew normally when untreated and was killed when treated with lydiamycin A (shown by the grey and blue lines respectively in **Figure 72**). The MsPDF expression strain grew a little slower than the EV control strain but ultimately reached the same culture density. However, the treated MsPDF expression strain (green line in **Figure 72**) was completely killed by the presence of 50 µM lydiamycin A. These data indicated that the housekeeping MsPDF protein was unable to confer resistance against lydiamycin A, which is indicative that the observed LydA-mediated resistance is not mediated by titration of lydiamycin. Instead, it is possible that the LydA PDF is specifically resistant against inactivation by lydiamycin A. This may be mediated by residue changes that affect the shape of the active site and thus block the binding of lydiamycin A to LydA. This hypothesis will be considered in **Section 4.3.1**.

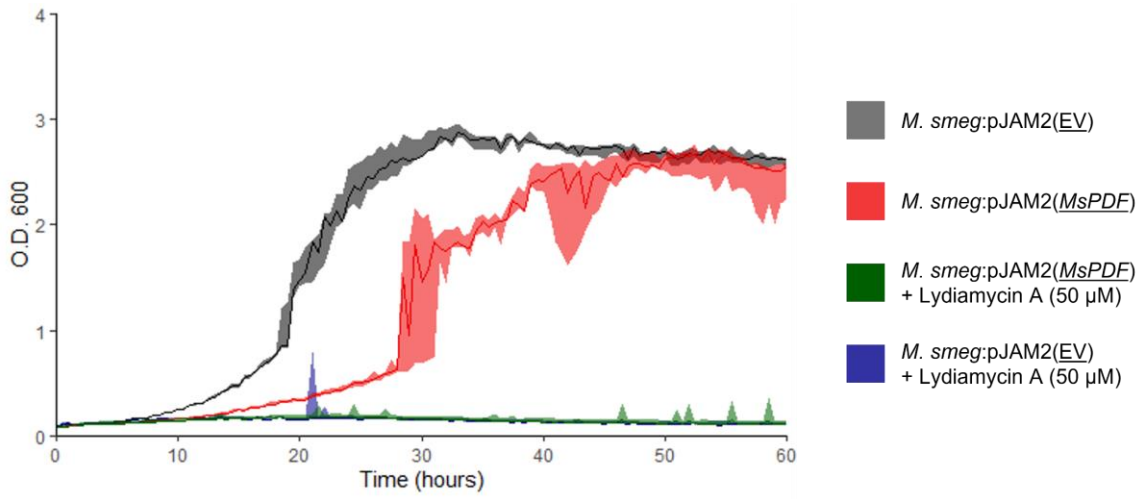


Figure 72. Effect of *M. smegmatis* expression of *MsPDF* on the sensitivity to lydiamycin A.

4.2.2.5 Can LydA confer actinonin resistance?

An experiment was conducted to determine whether LydA could confer resistance against actinonin. This was supported by the logic that actinonin and lydiamycin A are structurally similar (**Figure 73**). However, there are some important structural differences. The actinonin warhead features a hydroxamic acid whereas the lydiamycin A warhead features a carboxylic acid. Further, the peptide backbones of actinonin and lydiamycin A are dissimilar and this region is known to be important for the affinity of a PDFI to the PDF active site [339]. Therefore, this experiment was performed to understand the range of PDFI compounds that LydA may confer resistance against.

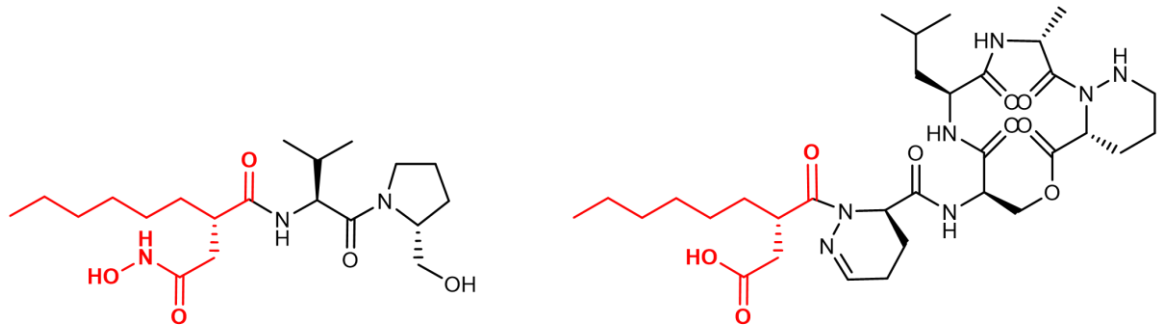


Figure 73. Comparison of lydiamycin A and actinonin structures. Warhead groups coloured in red.

The *M. smeg*:pJAM2(EV) and *M. smeg*:pJAM2(*lydA*) strains were grown in the presence and absence of 50 µM actinonin. The LydA expression strain grew in the absence of actinonin (indicated by the red line in **Figure 74**), albeit with a slower growth rate as compared to the EV control strain. The *M. smegmatis* EV control strain was effectively killed when treated with 50 µM actinonin (**Figure 74**). Interestingly, the expression of LydA conferred resistance when treated with 50 µM actinonin (indicated by the green line in **Figure 74**). The growth rate of the treated LydA expression strain was severally reduced over the first forty hours of the experiment, however the final culture density of this strain was consistent with the untreated strains by the sixty-hour mark. These preliminary data indicate that the LydA PDF is also able to confer resistance against actinonin (**Figure 71**).

Overall, these *M. smegmatis* growth curve assays have indicated that the LydA PDF is able to confer resistance against lydiamycin A and this interaction is likely caused by a specific resistance as resistance was not observed by the overexpression of the housekeeping PDF. There is also some preliminary data to suggest that LydA also confers resistance against actinonin.

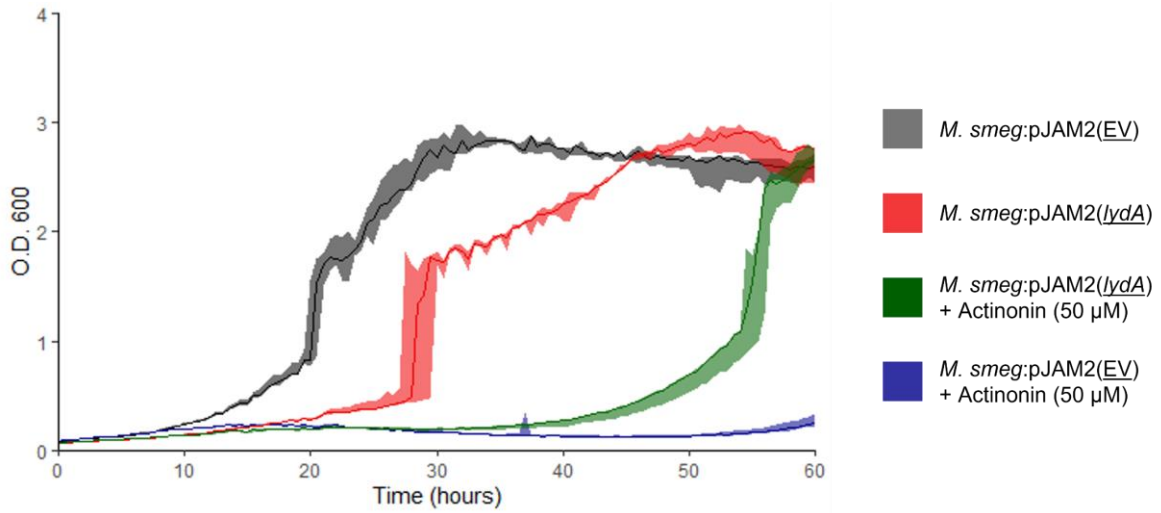


Figure 74. Effect of *M. smegmatis* expression of lydA on the sensitivity to actinonin.

4.2.2.6 *In vitro* characterisation of lydiamycin A PDF inhibition

The above *Mycobacterium* growth curve assays have indicated that the LydA PDF protein is able to confer resistance against lydiamycin A. This was suggestive that the molecular target of lydiamycin A was PDF. In order to experimentally prove this hypothesis, a collaboration was established with the Rodnina laboratory (Max Planck Institute for Multidisciplinary Sciences, Germany), with whom a sample of lydiamycin A was shared.

These collaborators (Lena Bögeholz, Evan Mercier) used established methodology for the *in vitro* assessment of PDF activity [340-343]. In brief, the activity of PDF was measured using a colorimetric coupled enzyme assay. A formylated dipeptide, formyl-Met-Leu-*p*-nitroaniline (fML-pNA) can be deformylated by PDF. The dipeptide product can then be hydrolysed by recombinant aminopeptidase protein to release *p*-nitroaniline, which is monitored spectrophotometrically at 405 nm (see materials and methods for details). They determined that lydiamycin A inhibited *E. coli* PDF function in a dose-dependent manner, with an IC_{50} of 4.6 µM (**Figure 75**).

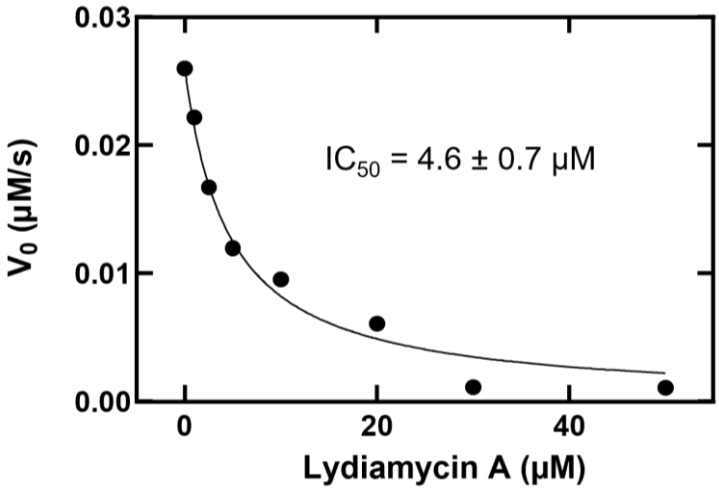


Figure 75. *In vitro* inhibition of *E. coli* PDF by lydiamycin A. Experiment performed and figure prepared by Lena Bögeholz.

Therefore, this data indicated that lydiamycin A is able to inhibit *E. coli* PDF protein. In the future, this experiment may be repeated with the LydA protein to determine whether the observed *in vivo* resistance corresponds to an increased *in vitro* IC_{50} value.

4.2.3 *In planta* assays

As *R. fascians* D188 is a well-documented epiphyte and the causative agent of leafy gall disease, experiments were performed to determine whether lydiamycin A production played a role *in planta*. This notion was supported by the fact that the lydiamycin BGC is present on the pathogenicity-associated megaplasmid along with the known genetic determinants of leafy gall disease. Therefore, assays were performed to assess the bioactivity of lydiamycin A in relation to plants. Three possible hypotheses were tested:

1. Is lydiamycin A directly toxic to plants?
2. Is lydiamycin A involved in leafy gall development?
3. Does the production of lydiamycin enhance the ecological fitness of *R. fascians in planta*?

Nicotiana benthamiana and *Nicotiana tabacum* (tobacco) species were selected as suitable plant models for *R. fascians* infection and leafy gall formation as they have been widely used in previous literature. Existing *Nicotiana* methodologies include seedling assays [344], whole plant assays [344] and excised leaf assays [304], which are explained below.

A preliminary experiment was conducted to determine whether *R. fascians* biosynthesised lydiamycin A when grown on plants. *N. benthamiana* seedlings were inoculated with either *R. fascians* WT or *R. fascians* Δ*nrp* strains and allowed to grow for another week. The seedlings were then extracted in methanol and extracts were analysed by LC-MS. The extracted ion chromatograms for lydiamycin A ($[M+H]^+$ = 664.37) are presented in **Figure 76** and indicate that lydiamycin A is present in all *R. fascians* WT replicates and none of the *R. fascians* Δ*nrp* replicates. This indicates that lydiamycin A is produced by *R. fascians* WT when grown on plants and therefore supports the further characterisation of the *in planta* role of lydiamycin A.

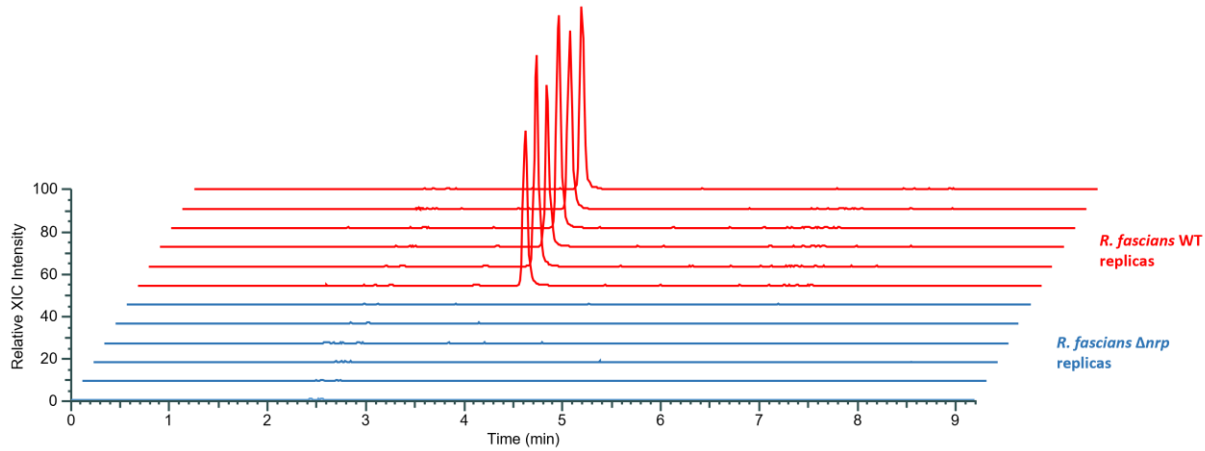


Figure 76. Production of lydiamycin A by *R. fascians in planta*.

4.2.3.1 Is lydiamycin A directly toxic to plants?

A root length assay, based on the methodology of Savory and colleagues [344], was developed to determine the phytotoxicity of lydiamycin A. In brief, *N. benthamiana* seeds were surface sterilised and grown on an agar plate that is positioned vertically in a plant growth chamber. After a week of growth, the seedlings are inoculated with either *R. fascians* WT or Δ*nrp* strains and sterile solution was used for mock treatment. The seedlings were grown for another week and the amount of root growth over this week period was recorded. The production of a phytotoxic compound would cause a reduction in root growth as compared to the mock treatment seedlings.

The results of the root length assay are displayed in **Figure 77**. Each dot represents a single seedling data point and the median and interquartile range are represented by the error bars. There was a significant (t-test, $p<0.00005$) reduction in root length growth in *R. fascians* WT and Δ*nrp* infected seedlings, as compared to the mock treatment. This

indicates that the growth of *R. fascians* on the seedlings is stunting root growth, which has been observed previously in the literature [344]. However, there was no significant difference (t-test, p=0.985) between the root length growth of *R. fascians* WT and Δ*nrp* infected seedlings. As the only difference between these strains is the production of lydiamycin, this result indicates that lydiamycin does not act as a general phytotoxin. This experiment was repeated once (data not shown) and also once by a collaborator (Dr. Isolde Francis, California State University Bakersfield, USA) and similar results were observed.

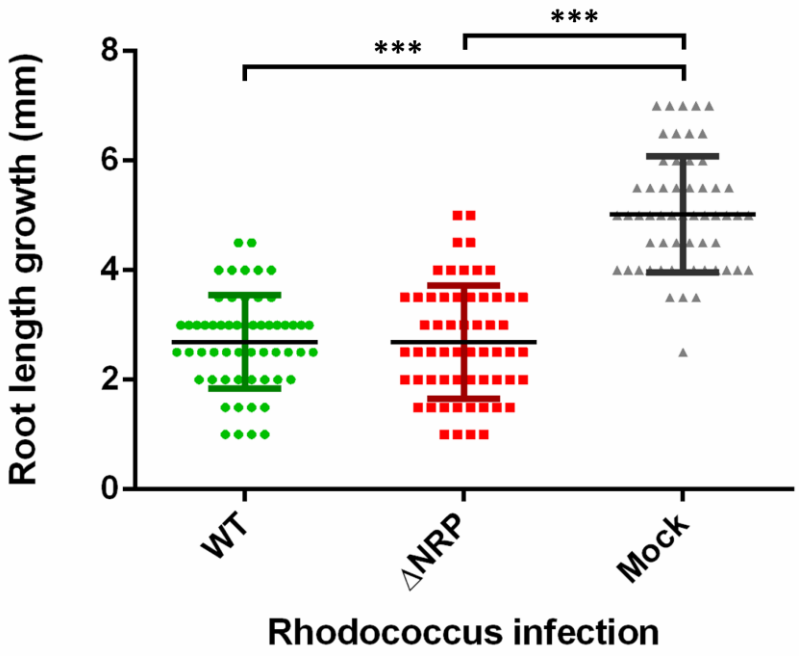


Figure 77. *N. benthamiana* root length assay to determine phytotoxic effect of lydiamycin production. *** = t-test significance where p<0.00005

4.2.3.2 Is lydiamycin A involved in leafy gall disease pathogenesis?

In order to assess whether the production of lydiamycin affected the occurrence rate or severity of leafy gall disease, two assays were used. In the first, whole tobacco plants were infected, according to the methodology of Savory and colleagues [344], and leafy gall disease monitored. In brief, sterile *N. benthamiana* plants were grown aseptically for six weeks. The meristem of the plants were pinched with forceps and then infected with either *R. fascians* WT or Δ*nrp* or with sterile mock infection. Plants were allowed to grow for another four weeks and were then imaged. Nine *N. benthamiana* plants were grown under each of the three conditions (uninfected, *R. fascians* WT infection, *R. fascians* Δ*nrp* infection). 100% of *R. fascians* WT infected and 100% of *R. fascians* Δ*nrp* infected plants

developed leafy galls, whereas none of the uninfected plants developed leafy gall disease symptoms. Representative images of plant disease phenotypes are illustrated in **Figure 78**.

Figure 78. Representative images of infected *N. benthamiana* plants.

There were no striking visual differences between the leaky galls developed in *R. fascians* WT and Δ*nrp* infected tobacco plants. The masses of the plants and masses of the gall relative to the plant (expressed as a percentage) were recorded. The masses of *R. fascians* WT infected and *R. fascians* Δ*nrp* infected plants were statistically lower than the masses of uninfected plants (**Figure 79**. T-test: p=0.048, 0.024 respectively). This suggests that the growth of *R. fascians* was stunting plant growth, perhaps due to the utilisation of plant-derived carbon and nitrogen sources by the bacteria during the later stages of infection [302, 303]. However, there was no significant difference in plant mass between *R. fascians* WT infected and *R. fascians* Δ*nrp* infected plants (**Figure 79**). Further, there was no significant difference in the relative gall mass between *R. fascians* WT infected and *R. fascians* Δ*nrp* infected plants (**Figure 80**). These data indicate that lydiamycin production does not play a role in leafy gall disease development.

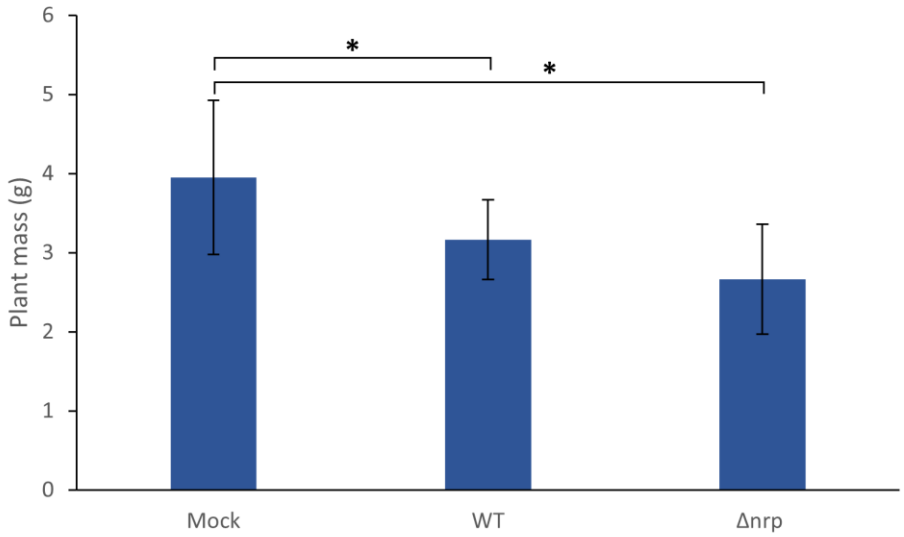


Figure 79. Mass of *N. benthamiana* plants infected with mock treatment or *R. fascians* WT or Δ*nrp* strains. Error bars represent standard deviation. * = t-test significance where p<0.05

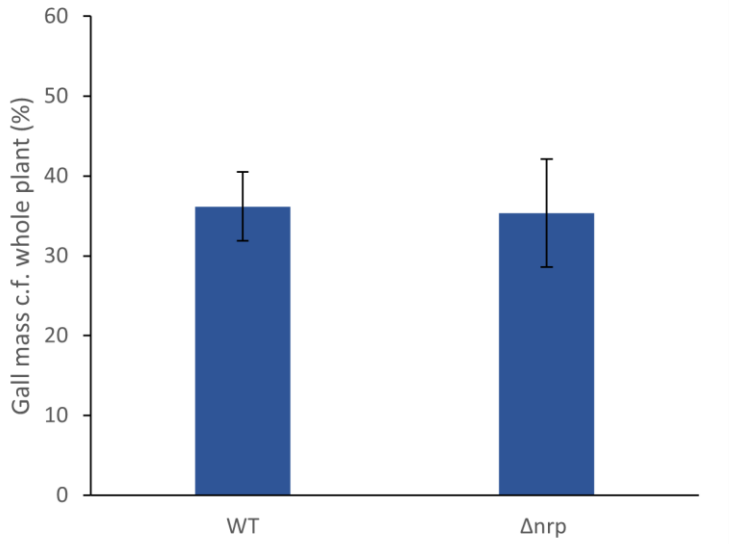


Figure 80. Relative mass of leafy gall (%) of *N. benthamiana* plants infected with *R. fascians* WT or Δ*nrp* strains. Error bars represent standard deviation.

An excised leaf assay, as described by Francis and colleagues [304], was also used to characterise leafy gall disease pathogenesis caused by the *R. fascians* WT and Δ*nrp* strains. In brief, sterile *N. tabacum* leaves from four-week-old plants were placed on agar plates and a drop of culture was applied. The plates were incubated for three weeks and the phenotypes recorded. Again, three infection conditions were tested: *R. fascians* WT, *R. fascians* Δ*nrp* and a mock infection using sterile solution.

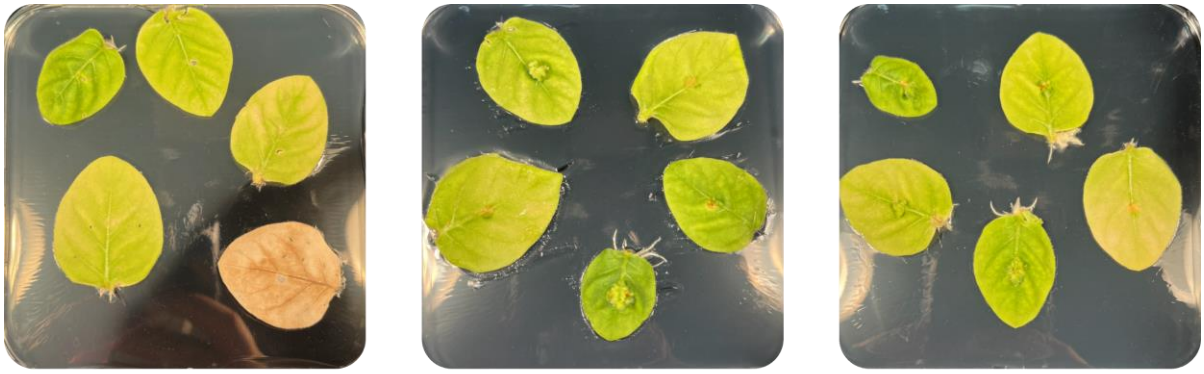


Figure 81. Representative images of infected *N. tabacum* excised leaves.

Excised leaves infected with either *R. fascians* WT or Δ*nrp* strains resulted in the growth of gall-like structures that appear as bundles of tissue developing vertically from the surface of the plate (**Figure 81**). These gall-like structures never developed on uninfected (mock) excised leaves. The severity and frequency of these gall-like structures were approximately equal between the *R. fascians* WT and Δ*nrp* infected leaves, which is suggestive that lydiamycin production is not involved with leafy gall pathogenesis.

These whole plant and excised leaf assays both indicate that lydiamycin is not involved in the progression of leafy gall disease as the *R. fascians* Δ*nrp* strain caused very similar plant disease phenotypes as the *R. fascians* WT strain. This is consistent with the work of Francis and colleagues, who previously determined that the *D188nrp* cluster was not involved in leafy gall disease development [304].

4.2.3.3 Does the Production of Lydiamycin enhance the Ecological Fitness of *R. fascians in planta*?

The final hypothesis as to why *R. fascians* biosynthesises lydiamycin A during plant infection is concerned with niche colonisation. Perhaps lydiamycin A is produced in order to suppress competing bacteria and enhance the survivability and colonisation abilities of *R. fascians*. This hypothesis is supported in theory as only the producer, *R. fascians*, is likely to produce the lydiamycin resistant PDF, LydA, and therefore most other epiphytes or soil bacteria are likely to be sensitive to the antimicrobial activity of lydiamycin A.

Testing this hypothesis required comparing a lydiamycin-producing strain to a strain that lacks LydA. A preliminary spot-on-lawn experiment was conducted to determine the MIC of lydiamycin A against *R. fascians* WT (which has the LydA gene) and *R. fascians* D188-5 (which lacks the megaplasmid and therefore does not have the LydA gene). The MIC of lydiamycin A against *R. fascians* WT was determined to be >600 µg/mL, whereas the MIC of lydiamycin A against *R. fascians* D188-5 was determined to be approximately 10 µg/mL (**Figure 82**). Therefore, the D188-5 strain is at least sixty times more sensitive to lydiamycin A than the wild type strain. Therefore, this data supports the hypothesis that lydiamycin A production may enhance fitness by suppressing bacteria lacking resistance against lydiamycin A.

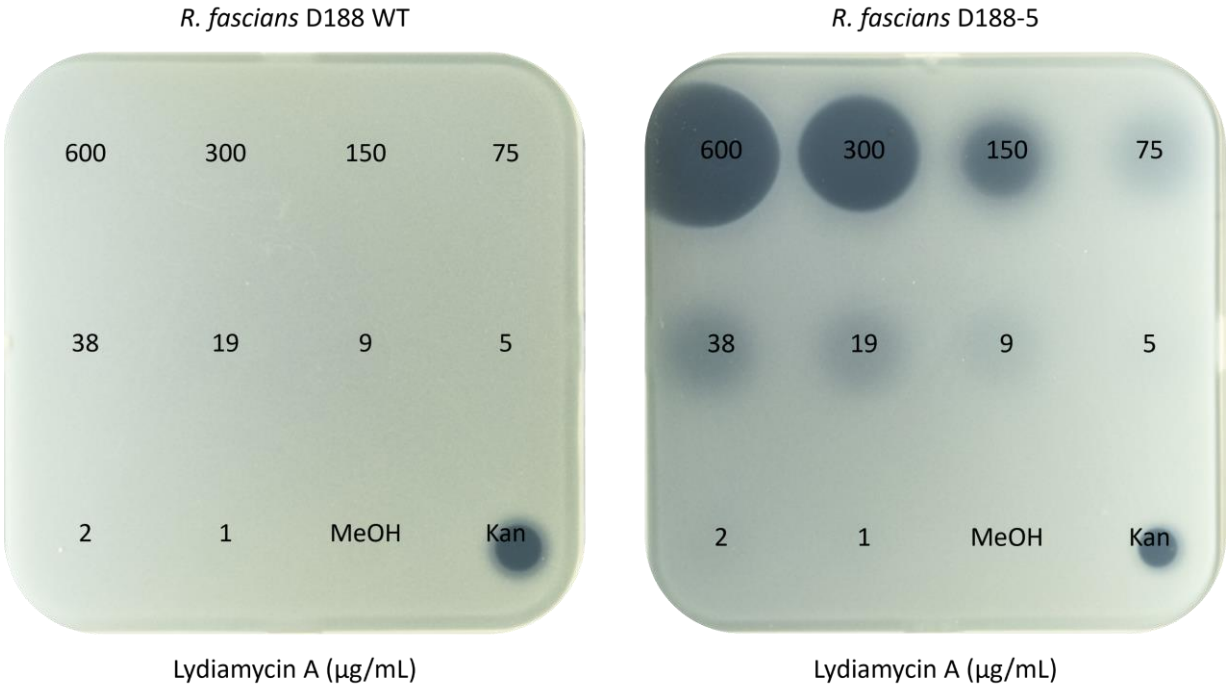


Figure 82. Spot-on-lawn assays of lydiamycin A against *R. fascians* WT and D188-5 strains.

In order to test this hypothesis further *in planta*, a competition assay was performed to test the relative fitness of *R. fascians* strains growing on *N. benthamiana* seedlings. Two pairs of *R. fascians* strains were tested – WT compared to plasmid free, and Δ*nrp* compared to plasmid free. These pairings were chosen as the plasmid-free strain represents a competing soil bacterium, as it lacks the ability to produce lydiamycin A and is lydiamycin sensitive, whilst sharing the same chromosome as megaplasmid-containing *R. fascians* strains – thus being directly comparable. Further, these pairings facilitated effective strain selection as the plasmid-free D188-5 strain features a chromosomal streptomycin

resistance cassette [305]. The competition assay proceeded with an equal (1:1) mix of *R. fascians* strains applied to a seedling and the seedling grown for 7 days. The bacteria were recovered, serially diluted and plated on selective media to determine the resultant ratio of bacteria. Relative fitness was calculated by comparing the test strain CFU/mL to the plasmid free strain CFU/mL, as detailed in materials and methods.

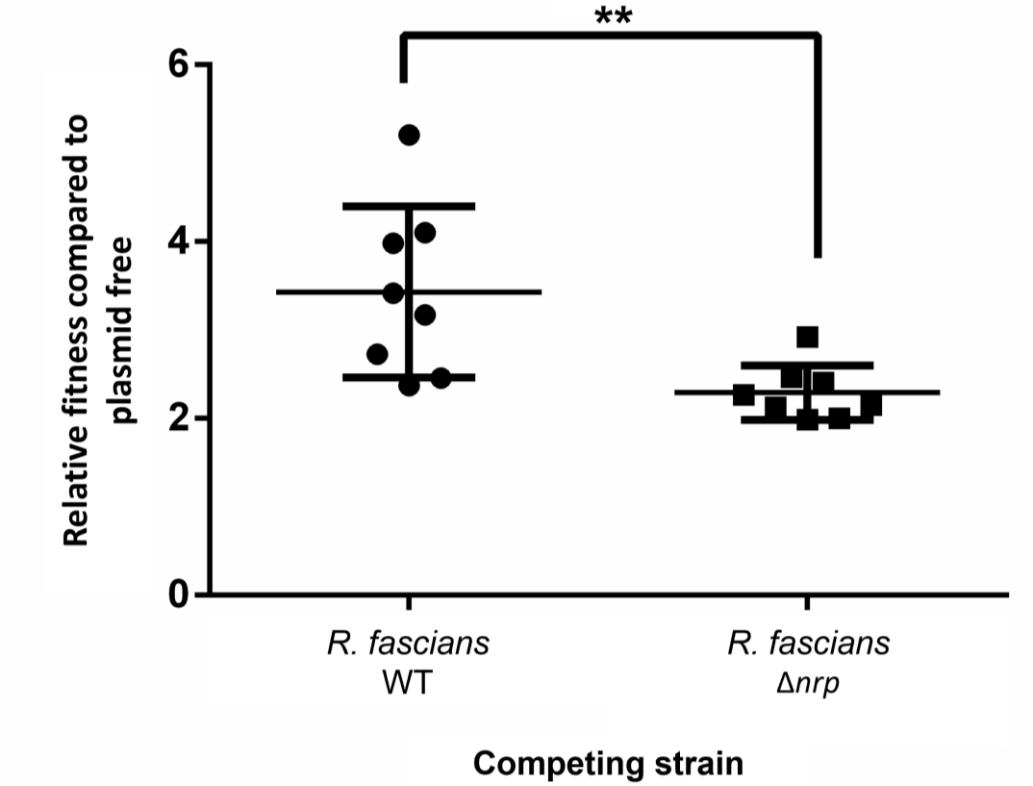


Figure 83. Competition assay of plasmid free *Rhodococcus fascians* against WT and Δ*nrp* strains. Relative fitness of each test strain was calculate as compared to the plasmid free strain. Individual replicates are indicated as dots. The mean of these replicates is shown as a black cross-bar and error bars indicate the standard error of the mean. A two-tailed t-test showed a significant different between WT and Δ*nrp* relative fitness values ($p=0.0068$, $p<0.01$, **).

Both the *R. fascians* WT and Δ*nrp* strains out-competed the plasmid-free strain, as they both had relative fitness values greater than 1 (**Figure 83**). This was consistent with the notion that the pFiD188 megaplasmid confers adaptions for plant colonisation and disease progression. There was also a significant (t-test, $p=0.0068$) difference in the relative fitness between the WT and Δ*nrp* strains, indicating that the production of lydiamycin enhanced fitness.

To ensure that this significant result was reproducible, the plant competition assay was repeated using biological replicates of the *R. fascians* WT and Δ*nrp* strains, with bacteria

sampled at three timepoints; 3, 7 and 14 days post infection. Over these time points, there was an increase in survivability of the WT strain compared to the Δ*nrp* strain and by day 14 there was a significantly greater amount of WT cells compared to the Δ*nrp* cells ($p<0.01$) (**Figure 84**).

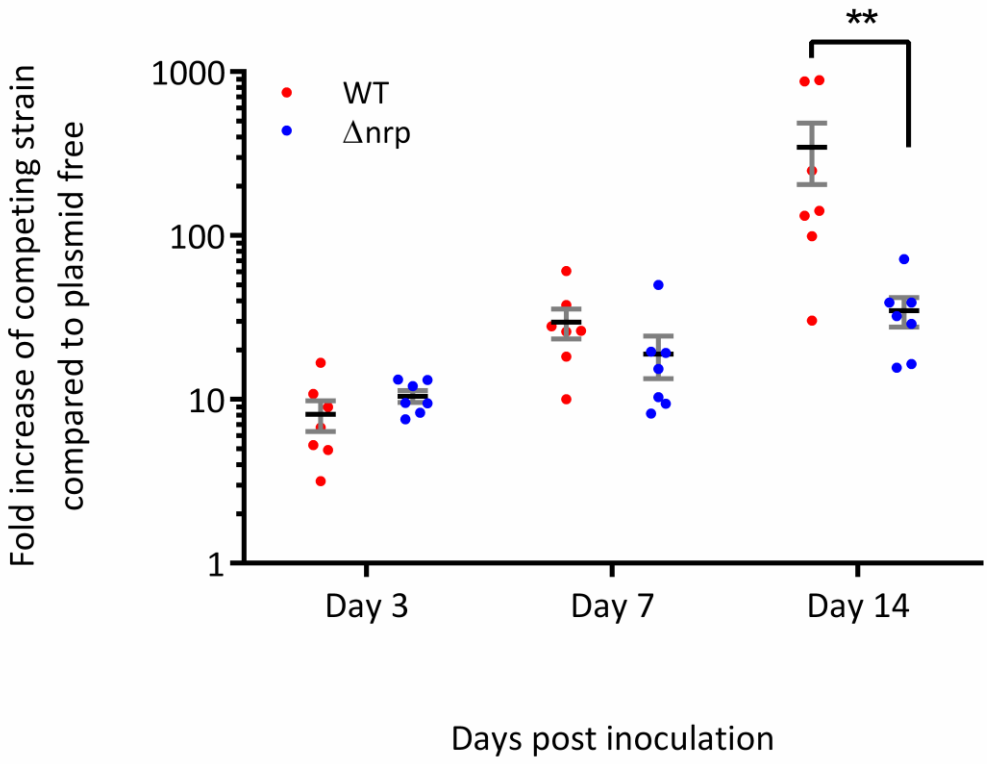


Figure 84. Competition assay of plasmid free *Rhodococcus fascians* against WT and Δ*nrp* strains. Percentage of WT (red) and Δ*nrp* (blue) are plotted as a fold increase compared to plasmid free. Individual biological replicates are indicated as dots for both WT (red) and Δ*nrp* (blue) at 3 time points. The mean of these biological replicates is shown as a black cross-bar and error bars (grey) indicate the standard error of the mean. A Two-way ANOVA showed a significant interaction between strain and time point ($p<0.05$) with a significant effect of both days post inoculation ($p<0.01$) and competing strain ($p<0.05$). Sidak's multiple comparisons showed a significant difference between WT and Δ*nrp* at day 14 post inoculation ($p<0.01$, **).

Therefore, these data indicate that the WT strain was able to outcompete the plasmid-free strain to a significantly greater degree than the Δ*nrp* strain was able to outcompete the plasmid-free strain. As the only difference between the WT and Δ*nrp* strains is the ability to produce lydiamycin, this indicates that the production of lydiamycin is enhancing the fitness of *R. fascians in planta.* This is the first indication that the production of lydiamycin enhances the fitness of *R. fascians* and implicates lydiamycin as an important NP for ecological competitiveness.

Further experimentation is required to characterise the mechanisms promoting this observed fitness advantage. However, it is possible that lydiamycin may suppress the

growth of competing bacteria and therefore enhance the ability of *R. fascians* WT to establish a niche and survive on the surface of the plant. More advanced experiments with more complex microbial communities are required to understand this hypothesis further.

4.3 Discussion and conclusions

4.3.1 Summary / Discussion of results

Chapter 3 described the isolation of lydiamycin A and the structural revision based on biosynthetic logic. Existing literature had reported that the major bioactivity of lydiamycin A was as an antimycobacterial agent. However, this chapter explored whether there was a more physiologically relevant role of lydiamycin, considering the identification of the plant pathogen *R. fascians* as a novel producer.

There have been large variances in the reported antimycobacterial potency of lydiamycin A in the literature [318, 319]. This variance was also observed in this work as the MIC of lydiamycin A against *M. smegmatis* varied greatly depending on the methodology used. The MIC of lydiamycin A against *M. smegmatis* was determined to exceed 600 µg/mL using a spot-on-lawn assay, however growth curve assays indicated that the MIC was approximately 50 µM. This growth curve assay was further optimised for the use of *M. smegmatis* as a reporter strain and was used to test whether *lydA* was an immunity determinant.

The expression of the LydA PDF protein conferred complete resistance against 50 µM lydiamycin A in the *M. smegmatis* growth curve assay. This mechanism was further characterised as the housekeeping *M. smegmatis* PDF protein did not confer lydiamycin resistance. These data suggest that the LydA protein is specifically resistant to lydiamycin and thus confers resistance to the producing organism. There is precedence in the literature for gain of PDFI resistance by modification of the target. Actinonin-resistant *Streptococcus* strains were generated, and harboured mutations in the PDF gene thought to disrupt metal ion coordination which decreased susceptibility to actinonin but slowed the growth rate [193].

Having experimentally confirmed that the LydA PDF was resistant to lydiamycin A (**Figure 71**) and the *E. coli* and *M. smegmatis* housekeeping PDF proteins were sensitive to lydiamycin A (**Figure 72, Figure 75**), these sequences were aligned in order to understand

whether a particular residue may be the resistance determinant. To broaden this analysis, the *R. fascians* housekeeping PDF sequences were added and proposed to be sensitive (as the D188-5 strain is sensitive – **Figure 82**). The PDF sequences encoded in the warhead clade potential clusters (**Figure 38**) were added, which I hypothesise are also resistant to PDFIs. One hypothesis for how some PDF proteins may be resistant to lydiamycin is that the surface geometry in and around the active site may occlude the binding of lydiamycin to the catalytic metal ion due to occlusion of the bulky peptide backbone. To test this hypothesis, the reported crystal structure of actinonin bound to *E. coli* PDF (PDB: 1G2A) was analysed and the residues making up the active site surface were determined to be 41-50 and 86-97. These surface regions were annotated on the alignment of the resistant and sensitive PDF sequences (represented by purple boxes on **Appendix Figure 119**). A single residue (*E. coli* PDF #97) was identified which was tyrosine in all six putative resistant PDFs and always a different identity (R, N, Q, L) in the sensitive PDFs.

In order to indicate whether the bulky tyrosine aromatic group may confer resistance via occlusion of the active site, the *lydA* sequence was submitted to AlphaFold 2 [345, 346] to generate a protein structure model. This LydA model was superimposed against the *E. coli* PDF structure. Consideration of the active site geometry (**Figure 85**) indicates that the LydA Y113 residue (red surface) occupies more space than the corresponding R97 residue (yellow surface) in the *E. coli* PDF structure. It is the bulky aromatic phenol group of Y113 which is predicted to be occupying the active site.

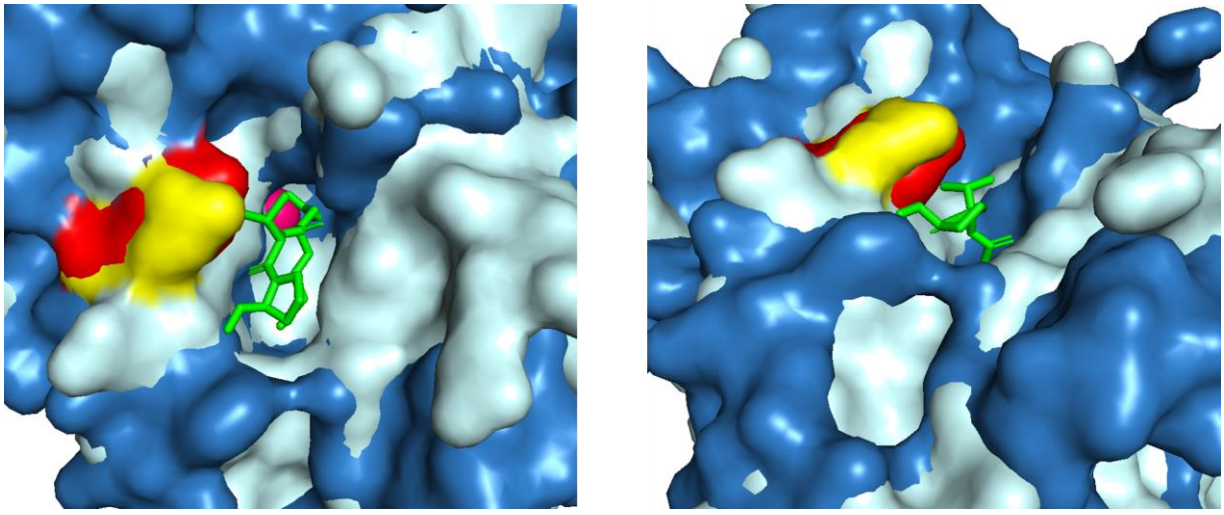


Figure 85. Comparison of sensitive and resistant PDF active site surface geometry. The *E. coli* PDF crystal structure (PDB: 1G2A) is sensitive to lydiamycin A and the surface is coloured in light blue. The LydA AlphaFold generated model is resistant to lydiamycin A and is coloured in blue. The *E. coli* PDF model features actinonin (green) docked to the catalytic zinc ion (pink). The right image is rotated 90^{o} forward in respect to the left image. The potential resistance determinant residue is coloured is *E. coli* R97 (yellow surface) and LydA is Y113 (red surface). This tyrosine residue may occlude the active site in the resistant LydA protein.

Therefore, this alignment and structural analysis suggests that the Y113 residue may function as a resistance determinant by blocking the entry and docking of lydiamycin A to the LydA catalytic zinc ion. This would represent the first reported example of PDFI resistance gained by the change in shape of the active site, as the only reported PDFI resistance gained by target mutation is via modification of the catalytic metal ion coordination residues [193].

However, the LydA structure has been generated using AlphaFold 2 and it may be necessary to determine the structure by crystallography and XRD to guarantee the model accuracy. Further, a more extensive alignment of likely resistant and sensitive PDF sequences would more thoroughly test the importance of this tyrosine residue. Finally, this potential resistance determinant would need to be experimentally confirmed. For example, generating a mutant LydA protein lacking the Y113 residue and generating a mutant *E. coli* PDF protein featuring the tyrosine residue to determine whether the former loses resistance and the latter gains resistance.

The potency of lydiamycin A PDF inhibition was determined *in vitro* by the Rodnina laboratory. Lydiamycin A was determined to inhibit *E. coli* PDF with an IC_{50} of 4.6 µM. Comparatively, actinonin is the most potent natural PDFI and exhibits an IC_{50} of 3 nM against *E. coli* Ni-PDF [128], over 1,000x more potent than lydiamycin A. Perhaps this is to be expected as SAR analyses of PDFIs have determined that the presence of a hydroxamate group greatly enhances potency [184]. This could be further explored by producing a lydiamycin A derivative with the terminal hydroxamate group and determination of whether it is a more potent PDF inhibitor.

Further, comparison the peptide backbones of actinonin and lydiamycin A may answer why actinonin has antiproliferative activity [197], whereas lydiamycin is reported to not be cytotoxic [319]. The balance between antimicrobial and antiproliferative potency is thought to be caused by the relative affinity of a PDFI towards the human PDF protein (HsPDF) compared to the affinity towards bacterial PDF proteins. HsPDF features a narrow active site pocket, whereas bacterial PDF proteins feature a wider active site pocket that allows the entry of bulky inhibitors [180, 183]. Therefore, it is possible that the bulky peptide backbone of lydiamycin A confers greater selectivity towards bacterial PDF than HsPDF and therefore moderate antimicrobial activity is observed and no cytotoxicity activity. This may position lydiamycin A as an interesting candidate compound for a human antimicrobial agent and reinforces the notion that the compound may be an effective lead for antimycobacterial agents to combat multidrug resistant *M. tuberculosis* infection [318]. This also highlights the importance of isolating and characterising lydiamycin derivatives

(such as the potential lydiamycin congeners described in **Section 3.2.4.5**) which may offer enhanced potency.

In planta experiments were conducted with the objective of exploring whether the biosynthesis of lydiamycin by the plant pathogen *R. fascians* confers a physiologically or ecologically relevant fitness advantage. Three hypotheses were tested. Firstly, a root length growth assay determined that the production of lydiamycin A was not directly toxic to the plant. In theory, continued plant growth is advantageous for the survival of the bacterium as bacterial growth could be supported by exudate production. Secondly, excised leaf and whole plant assays indicated that the production of lydiamycin was not required for leafy gall disease progression, as the frequency and severity of gall formation was approximately equal between *R. fascians* WT and Δ*nrp* infected tobacco plants. Again, this is consistent with reported leafy gall disease pathogenesis research as the production of cytokinins is considered to be the major virulence factor and the operon responsible for their production is already described [291, 306, 347].

Finally, the hypothesis that lydiamycins are produced to enhance *R. fascians* survival and niche colonisation on the plant was tested. Two plant competition assays were performed, and both indicated that the WT *R. fascians* strain outcompeted the plasmid-free strain to a significantly greater extent than the Δ*nrp* strain outcompeted the plasmid-free strain. These data suggest that the presence of the pFiD188 megaplasmid is important for plant colonisation, which is consistent with the literature. However, these data also indicate for the first time that the production of lydiamycin enhances the fitness of *R. fascians* compared to *R. fascians* unable to produce lydiamycin. This frames lydiamycin as an important ecological NP which may enhance the fitness of *R. fascians* throughout its epiphytic colonisation and during leafy gall disease development. It is possible that this effect is mediated by the antimicrobial activity of lydiamycin which may supress competing bacteria on the surface of the plant or in the soil, whilst the *R. fascians* producer is resistant to lydiamycin due to the LydA protein. This hypothesis will need further investigation in order to determine the exact mechanism of how lydiamycin confers an ecological advantage to *R. fascians*. For instance, competition assays may be performed in more complex microbial environments that may more accurately represent the soil diversity.

Overall, this chapter has identified that lydiamycin A is a PDF inhibitor and that the lydiamycin BGC encodes a resistant PDF (LydA) which may prevent binding of lydiamycin to the active site metal ion by a novel mechanism. An ecological role of lydiamycin production by *R. fascians* has been proposed, as *in planta* competition assays indicated that lydiamycin A production confers a fitness advantage against competing bacteria.

4.3.2 Future perspectives

In the future, it would be beneficial to more thoroughly characterise the LydA PDF mechanism of lydiamycin resistance. There is some preliminary evidence to suggest that a single tyrosine residue at the entrance of the active site is able to confer resistance by occluding the entry of lydiamycin. This hypothesis could be tested by the mutation of sensitive and resistance PDF genes, computational docking experiments or by crystallisation and structural characterisation. The affinity of lydiamycin A towards these PDF proteins could also be quantified using the *in vitro* PDF assay.

The other major question for the future is the extent to which lydiamycin production plays a role in *R. fascians* niche colonisation and survival *in planta*. It would be beneficial to generate a strain of *R. fascians* which is lacking the entire lydiamycin BGC as this would allow characterisation of a strain lacking the ability to produce lydiamycin which was also sensitive to lydiamycin (as it lacks *lydA*). This would be a useful analogous strain to the competitive bacteria in the soil as the lydiamycin cluster is not widespread in nature. More broad competition assays could be conducted using synthetic or natural microbial communities to gain an understanding of the ecological role of lydiamycin production in more natural environments. Further, the *in planta* experiments described in this chapter were focussed only on large scale effects (e.g. plant mass, gall development). In the future, a more thorough assessment of the bacterial and plant conditions (e.g. microscopy, transcriptional analysis) may shed more light on this interaction.

Chapter 5: Pseudomonas Cyclic Lipopeptides

This work was performed in order to gain practical experience of cyclic lipopeptide isolation and characterisation, prior to undertaking the work described in Chapters 3 and 4. The work described in this chapter was published in the paper of Pacheco-Moreno and Stefanato *et al*. [348].

5.1 Introduction

Previous work by the Truman and Malone laboratories (John Innes Centre) had identified a number of pseudomonads which may be active against the causative agent of potato scab – *Streptomyces scabies*. In this work, 240 pseudomonas strains were isolated from potato fields and were phenotypically characterised for their motility, protease production, siderophore production and on-plate suppression of *S. scabies* [348]. Of these strains, 69 isolates were sent for whole-genome sequencing, of which approximately half were *S. scabies* suppressive. Each of the genomes were analysed by antiSMASH to determine their specialised metabolite potential. An association matrix was produced to compare the predicted specialised metabolite potential to the observed phenotype of each strain. In doing so, a strong association between *S. scabies* suppression and the presence of putative cyclic lipopeptide (CLP) BGCs was observed (**Figure 86**) [348]. Therefore, it was proposed that the *Pseudomonas* CLPs may be facilitating *S. scabies* antagonism.

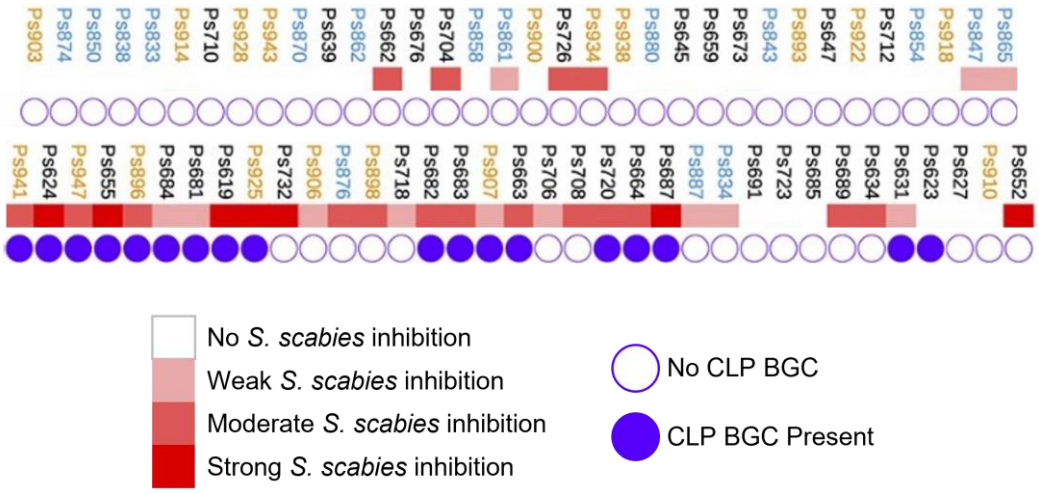


Figure 86. Association of *S. scabies* antagonism and cyclic lipopeptide (CLP) clusters in *Pseudomonas* strains. *S. scabies* suppression (red colour scales) and presence of putative cyclic lipopeptide biosynthetic gene clusters (purple circles) are indicated for each *Pseudomonas* strain. Adapted from [348].

Pacheco-Moreno and colleagues further investigated the Ps682 strain in order to characterise the CLP responsible for *S. scabies* antagonism. The Ps682 CLP BGC shared homology with the viscosin BGC. The authors produced a mutant Ps682 strain with a disrupted NRPS gene in the CLP BGC (Ps682 Δ*visc*) which was unable to antagonise *S. scabies* (**Figure 87A**). The Ps682 WT and Δ*visc* strains were cultured, extracted and analysed by LC-MS. Comparison of the exact mass and fragmentation patterns indicated that Ps682 produced a viscosin-like CLP (**Figure 87B**) [348], hereafter referred to as '682CLP'.

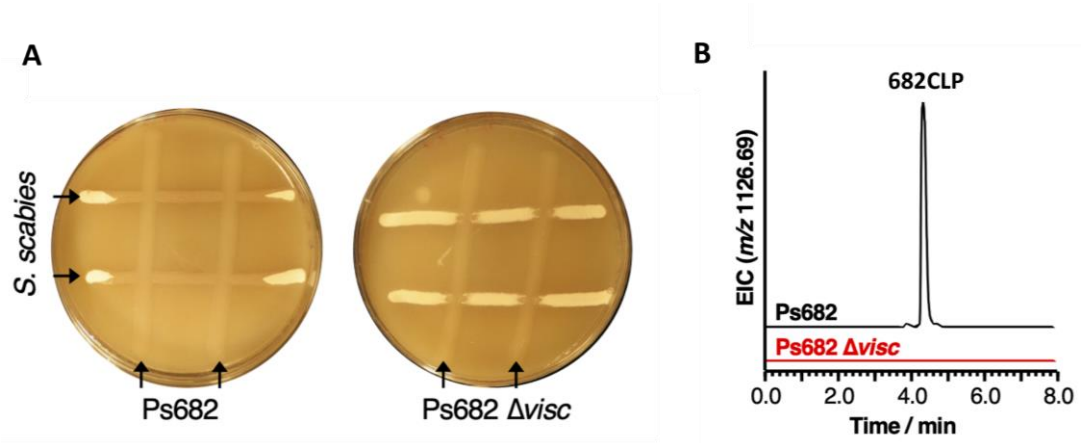


Figure 87. Identification of the Ps682 682CLP compound. **A.** Cross-streak competition assays of *S. scabies* (horizontal) against the Ps682 WT and Δ*visc* strains (vertical). **B.** Metabolomic profiles of the Ps682 WT and Ps682 Δ*visc* strains with extracted ion chromatograms of m/z 1126.69 – the mass of 682CLP. Adapted from [348].

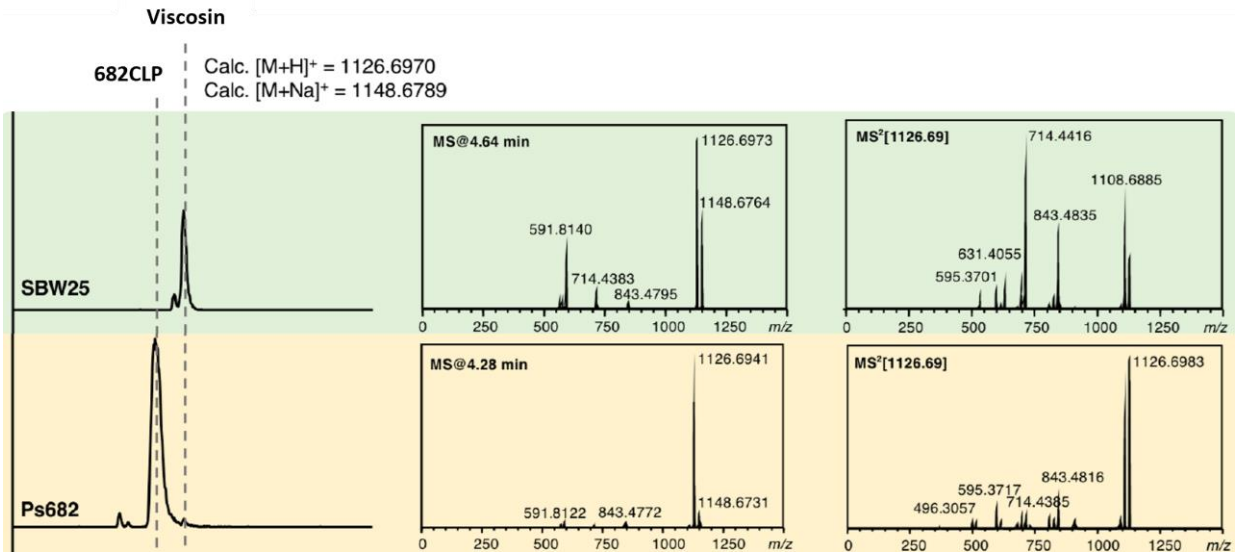


Figure 88. LC-MS/MS analysis of *Pseudomonas* SBW25 and Ps682. Comparison of the retention time and MS2 fragmentation pattern of $[M+H]^+$ = 1126.697 ion from SBW25 (viscosin producer) extract and Ps682 (682CLP producer) extract. Adapted from [348].

Pacheco-Moreno and colleagues used LC-MS analysis to compare 682CLP and viscosin. Although both compounds share the same mass, the retention times and MS-MS spectra are distinct [348] (**Figure 88**). This data indicated that 682CLP shared some structural similarity with viscosin but was structurally distinct.

Due to the genetic and structural similarity of 682CLP to viscosin, the published viscosin-like CLPs were considered in order to direct the structural elucidation of 682CLP. Only three viscosin-like CLPs have masses equal to the observed 682CLP mass: viscosin, white-line-inducing principle (WLIP) [349] and massetolide F [350]. All three of these natural products are comprised of a cyclic nonapeptidic backbone with a 3-hydroxy decanoic acid (3HDA) tail, however, are distinguished by the identity and/or configuration of the amino acid residues (**Table 8**). Therefore, the 682CLP compound may be one of these three known viscosin-like compounds or may represent a novel viscosin-like CLP.

Table 8. Structures of viscosin, WLIP and massetolide F nonapeptidic backbones. All of these viscosin-like CLPs have a mass of 1125.69 and feature a terminal 3-hydroxy decanoic acid tail. Highlighted columns indicate differentiating amino acid residues/configurations.

	1	2	3	4	5	6	7	8	9
Viscosin	L-Leu	D-Glu	D-Thr	D-Val	L-Leu	D-Ser	L-Leu	D-Ser	L-Ile
WLIP	L-Leu	D-Glu	D-Thr	D-Val	D-Leu	D-Ser	L-Leu	D-Ser	L-Ile
Massetolide F	L-Leu	D-Glu	D-Thr	D-Val	L-Leu	D-Ser	L-Leu	D-Ser	L-Leu

5.1.1 Chapter Aims

Previous work by Pacheco-Moreno and colleagues have identified that *Pseudomonas* Ps682 produces a viscosin-like CLP that is proposed to mediate the on-plate *S. scabies* antagonism [348]. However, the exact structure of 682CLP has not been determined. Therefore, the specific aims of this chapter are:

i. Develop methodology for the purification of the 682CLP compound
ii. Structurally characterise 682CLP by LC-MS/MS and NMR analysis
iii. Assess the activity of pure 682CLP against *S. scabies*

5.2 Results

5.2.1 682CLP Purification and Structural Elucidation

5.2.1.1 682CLP Purification

Attempts to purify 682CLP from *Pseudomonas* Ps682 grown in liquid media failed. Therefore, purification proceeded from solid media. *Pseudomonas* Ps682 was grown on fifteen 140 mm diameter King's Broth agar plates supplemented with 5 mM Mg (KB+Mg), which were extracted into ethyl acetate. This organic fraction was washed with water and dried to yield 43 mg of crude material. This crude organic fraction was separated firstly by reverse phase flash chromatography and then by reverse phase semi-preparative scale HPLC (**Appendix Figure 120**), whilst monitoring the elution of 682CLP by LC-MS analysis (see materials and methods for details). This yielded 0.5 mg of pure 682CLP.

Due to scalability limitations with the purification of 682CLP from solid medium, this process was repeated twice more to yield a total of 1.7 mg of pure 682CLP.

5.2.1.2 682CLP Structural Elucidation

Pure 682CLP was analysed by HR-MS/MS to produce an accurate mass and to gain MS/MS information. The exact mass of 682CLP was found to be 1126.6964 ($[M+H]^+$) (**Appendix Figure 121**), which is consistent with the masses of viscosin, WLIP and massetolide F (calculated masses = 1126.6970, Δ0.5 ppm).

The 682CLP MS/MS spectrum was also analysed to confirm the connectivity of the compound. It was predicted that the viscosin-like CLP would undergo elimination of the ester bond under MS conditions to give rise to a linearised peptide that is susceptible to fragmentation (**Figure 89**).

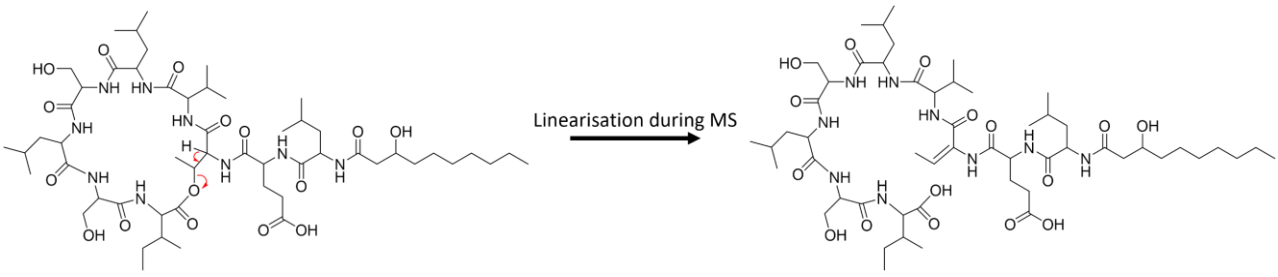


Figure 89. Predicted linearisation of viscosin-like CLP under MS conditions.

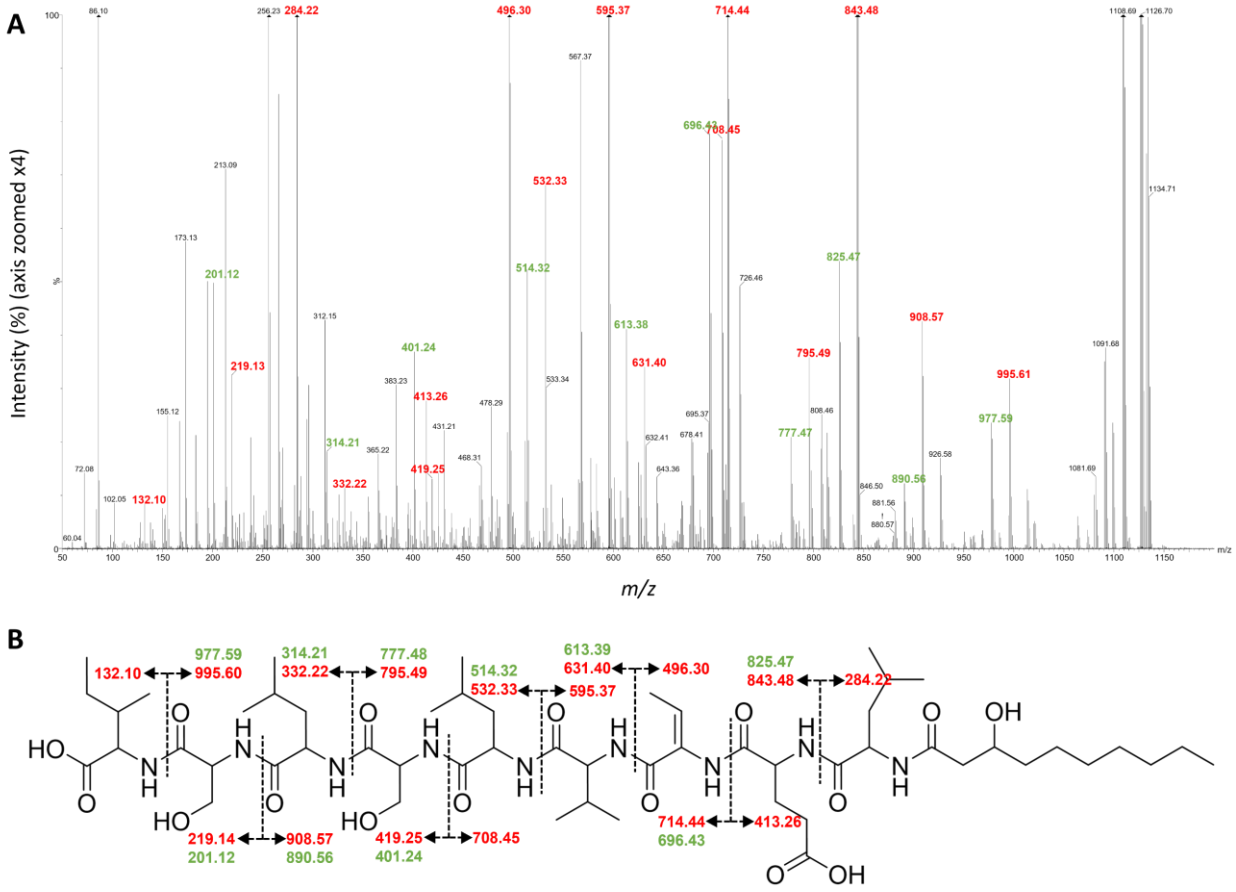


Figure 90. LC-MS/MS analysis of 682CLP. **A.** Full range of MS2 fragments from 682CLP parent ion. Intensity zoomed x4 for clarity. **B.** Corresponding 682CLP structure. Predicted masses of fragments and mono-dehydrated fragments are indicated in red and green, respectively.

The 682CLP MS/MS spectrum was consistent with the connectivity of the candidate viscosin-like CLPs (**Figure 90**). All of the predicted *y*-series and *b*-series fragment ions were observed (labelled in red), in addition to many fragments corresponding to a mono-dehydrated species (labelled in green). Therefore, this MS/MS data confirms that the connectivity of 682CLP is hydroxy decanoic acid (HDA)-Leu-Glu-Thr-Val-Leu-Ser-Leu/Ile-Ser-Leu/Ile. A limitation of this analysis is that leucine and isoleucine residues cannot be differentiated by MS as they have the same mass. Therefore, the identity of 682CLP could not be distinguished by MS analysis alone as the stereochemistry of the critical 5th leucine residue nor the identity of the 9th residue could be resolved (**Table 8**).

To further characterise the 682CLP structure, NMR experiments were carried out by Dr. Sergey Nepogodiev (John Innes Centre). In brief, pure 682CLP (1.0 mg) was dissolved in DMF-d7 and a series of 1D and 2D NMR spectra were acquired (see materials and methods for details). Analysis carried out by Dr. Sergey Nepogodiev confirmed that the connectivity of 682CLP was 3HDA-Leu1-Glu2-Thr3-Val4-Leu5-Ser6-Leu7-Ser8-Ile9 (Supplementary Figure 5 in [348]).

Therefore, the identify of 682CLP was confirmed to not be massetolide F as this features a leucine residue at position nine (**Table 8**), whereas NMR analysis confirms that the ninth residue of 682CLP is isoleucine. This suggests that the identity of 682CLP is either viscosin or WLIP, which are distinguishable by the configuration of the leucine residue at position 5 (**Figure 91**). Unfortunately, the configuration of residue 5 could not be determined by NMR. Further, comparison of the 682CLP NMR data to the published WLIP spectra [348, 349] by Dr. Sergey Nepogodiev indicated some minor shift differences (such as the γ-CH2 group of Glu2 (WLIP = δ_H 2.54 ppm, δ_C 30.3 ppm; 682CLP = δ_H 2.24 ppm, δ_C 34.8 ppm), suggesting that the identity of 682CLP is not WLIP.

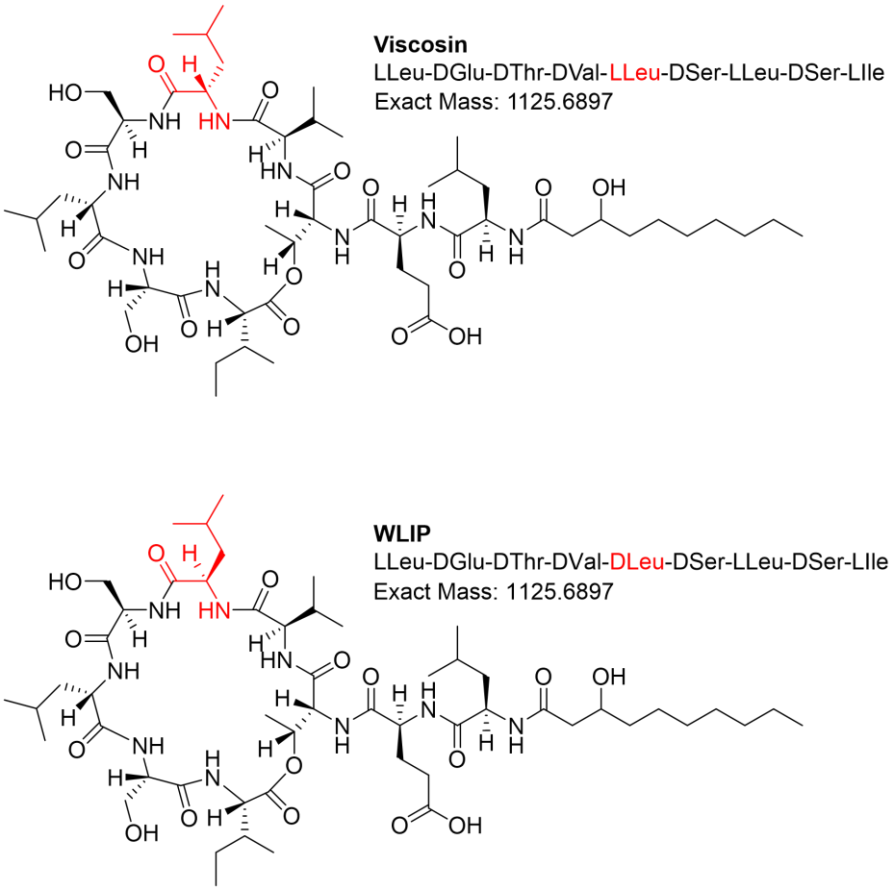


Figure 91. Comparison of viscosin and WLIP structures.

Comparison of the retention times was also performed in an attempt to determine the identity of 682CLP. In brief, cultures of Ps682, *Pseudomonas* SBW25 (viscosin producer) and *Pseudomonas* sp. LMG 2338 (WLIP producer) were extracted and analysed by LC-MS (**Figure 92**). Viscosin demonstrated a retention time of 9.34 minutes and WLIP demonstrated a retention time of 8.86 minutes. 682CLP eluted at a retention time of 8.86

minutes, which is suggestive that it is WLIP. However, this is in contention with the inconsistency of Ps682 NMR data as compared to published WLIP NMR data.

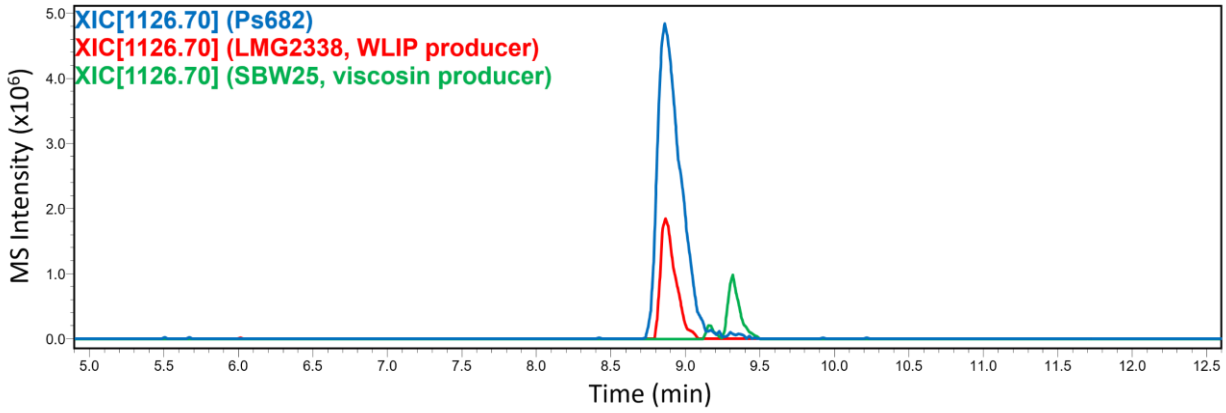


Figure 92. LC-MS comparison of Ps682 compared to *Pseudomonas* producers of WLIP and viscosin. LC-MS Extracted Ion Chromatograms (XICs) of ion [M+H$^+$] = 1126.70 from three *Pseudomonas* strains; Ps682, *Pseudomonas* sp. LMG 2338 (WLIP producer) and *Pseudomonas* SBW25 (Viscosin producer). Demonstrates the distinct retention time of viscosin (9.34 minutes), compared to WLIP (8.86 minutes).

Overall, the MS/MS data and NMR data were supportive that 682CLP shared the same backbone structure as viscosin and WLIP. However, the exact absolute configuration of 682CLP could not be determined and therefore the compound was reported as viscosin Isomer (viscosin I) [348].

5.2.2 Viscosin I Bioactivity

In order to test whether viscosin I was the compound active against *S. scabies*, disk diffusion assays were conducted. In brief, viscosin I was dissolved to a range of concentrations (10-100 µg/mL) and applied to filter disks. These disks were applied to agar plates of *S. scabies*, which were incubated and imaged daily. There was a clear concentration-dependent inhibition of *S. scabies* caused by viscosin I (**Figure 93**) and the minimum inhibitory concentration was estimated to be 20 µg/mL. Long-term growth of *S. scabies* in the presence of viscosin I indicates that the antimicrobial effect is only temporary and *S. scabies* growth resumes after 2-3 days (**Figure 94**). These data indicate that viscosin I is able to inhibit the growth of *S. scabies* and is the chemical determinant for the inhibition of *S. scabies* by Ps682, as observed in **Figure 87** [348].

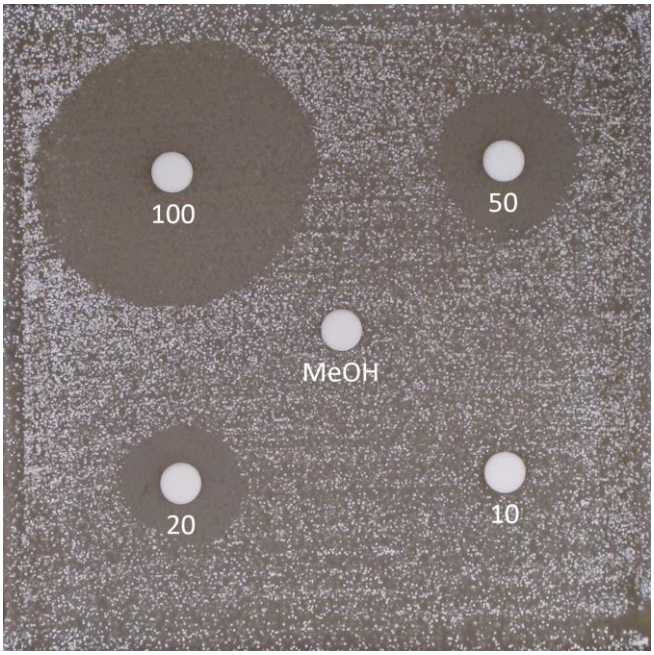


Figure 93. Disk diffusion assay of viscosin I against *S. scabies*. Concentrations are indicated (µg/mL), alongside a methanol control. The medium is Instant Potato Medium. Imaged at two days growth.

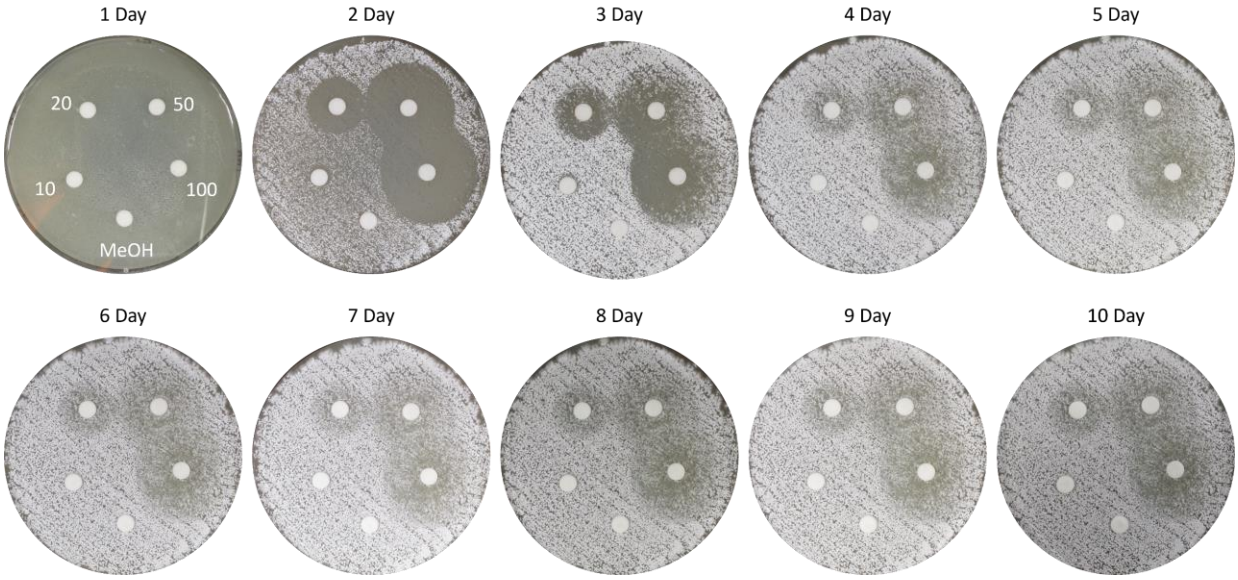


Figure 94. Timecourse of disk diffusion assays of viscosin I against *S. scabies.* Concentrations are indicated (µg/mL), alongside a methanol control. The medium is Instant Potato Medium.

5.3 Discussion and Conclusions

Prior to this work, Pacheco-Moreno and colleagues had determined that *Pseudomonas* Ps682 produced a viscosin-like CLP (682CLP) which was predicted to mediate *S. scabies* antagonism [348]. As described in this chapter, I developed methodology for the purification of 682CLP and isolated 1.7 mg of pure 682CLP across several purifications. LC-MS/MS analysis indicated that 682CLP features the same peptide backbone connectivity as viscosin, WLIP and massetolide F. Extensive NMR experimentation and analysis performed by Dr. Sergey Nepogodiev confirmed that the ninth residue was isoleucine and as such the structure of 682CLP was consistent with WLIP and viscosin. Comparison of the 682CLP retention time indicated that it was more structurally alike to WLIP than viscosin. However, inconsistencies with published WLIP NMR data and the inability to determine the absolute configuration of 682CLP limited further characterisation. Therefore, the identity of 682CLP was reported as viscosin isomer (viscosin I) [348].

Having purified viscosin I, it was possible to determine that viscosin I inhibits *S. scabies* growth with an MIC of approximately 20 µg/mL. This is consistent with the reported MIC values of viscosin-like CLPs against other Gram-positive strains [351]. The proposed antimicrobial mode of action of viscosin-like CLPs is via destabilisation of the cell membrane, possibly via pore formation [351, 352]. This work indicated that viscosin I was the chemical determinant of *S. scabies* inhibition and helped to frame CLPs as important antimicrobial NPs produced by *Pseudomonas* biocontrol strains [348].

Chapter 6: Materials and Methods

6.1 Materials

6.1.1 Strains

Table 9. Strains used during this study.

Strain	Genotype/Description	Application
E. coli DH5α (NEB)	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	Plasmid host for molecular cloning
E. coli BL21(DE3) (NEB)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ* *DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	Protein expression
E. coli NiCo21(DE3) (NEB)	can::CBD fhuA2 [lon] ompT gal (λ DE3) [dcm] arnA::CBD slyD::CBD glmS6Ala ΔhsdS λ	Protein expression with chitin pre-treatment
E. coli Rosetta(DE3) (Novagen)	F-ompT $hsdS_B(r_B^- m_B^-)$ gal dcm (DE3) pRARE (Cam^R)	Protein expression
E. coli SoluBL21 (Amsbio)	Not Reported	Protein expression
M. smegmatis mc^2155 (ATCC 700084)	Wild Type	Mycobacterial reporter strain
R. fascians D188	Wild Type	Lydiamycin producer
R. fascians D188-5 [305]	pFiD188-free. $Strep^R$	Avirulent strain
R. fascians D188 Δ*nrp* [304]	Disruption in *lydH*.	Disrupted lydiamycin production
Pseudomonas Ps682 [348]	Wild Type. Environmental biocontrol isolate	Viscosin I producer
Pseudomonas fluorescens SBW25	Wild Type	Viscosin producer
Pseudomonas LMG 2338	Wild Type	WLIP producer
S. scabies 87-22	Wild Type	Reporter strain, causative agent of potato scab
E. coli NiCo21::pET28a: *cinF*	Heterologous expression of *cinF* CCR gene	CinF purification

E. coli NiCo21::pET28a: *revS*	Heterologous expression of *revS* acyl-CoA ligase gene	RevS purification
M. smegmatis mc^{2}155::pJAM2:*lydA*	Heterologous expression of *lydA* PDF gene	Testing LydA-mediated lydiamycin resistance
M. smegmatis mc^{2}155::pJAM2:*msPDF*	Heterologous expression of *msPDF* PDF gene	Testing MsPDF-mediated lydiamycin resistance

Expression strains of *E. coli* BL21, SoluBL21, Rosetta and NiCo21 were also transformed with the pJF01-pJF15 plasmids, for the heterologous expression of warhead biosynthesis proteins. However, these are too numerous to list.

6.1.2 Plasmids

Table 10. Plasmids used during this study.

Plasmid	Features	Application	Resistance Marker
pET28a (Merck)	T7 RNA polymerase under *lac* control. N- or C- His tagging.	Heterologous expression in *E. coli*	Kanamycin
pACYCDuet-1 (Merck)	T7 RNA polymerase under *lac* control. N- or C- His tagging.	Heterologous co-expression of two genes in *E. coli*	Chloramphenicol
pJAM2 [337]	Acetamide-inducible promotor. C-terminal His tagging.	Heterologous expression in *M. smegmatis*	Kanamycin

6.1.3 Vectors

Table 11. Vectors constructed during this study.

Construct	Features	Application	Resistance Marker
pJF01	pET28a with *matL* insert. N-terminal His tag	Expression of matlystatin CCR	Kanamycin
pJF02	pET28a with *matE* insert. N-terminal His tag.	Expression of matlystatin epimerase	Kanamycin

pJF03	pET28a with *matB* insert. N-terminal His tag.	Expression of matlystatin mutase α subunit	Kanamycin
pJF04	pET28a with codon optimised *matQ* insert. N-terminal His tag.	Expression of matlystatin mutase β subunit	Kanamycin
pJF05	pET28a with codon optimised *matR* insert. C-terminal His tag.	Expression of matlystatin MeaB-like protein	Kanamycin
pJF06	pET28a with *matC* insert. C-terminal His tag.	Expression of matlystatin asparagine synthetase protein	Kanamycin
pJF07	pET28a with *matD* insert. C-terminal His tag.	Expression of matlystatin N-oxygenase	Kanamycin
pJF08	pET28a with *actF* insert. N-terminal His tag.	Expression of actinonin CCR	Kanamycin
pJF09	pET28a with *actC* insert. N-terminal His tag.	Expression of actinonin epimerase	Kanamycin
pJF10	pET28a with *actD* insert. N-terminal His tag.	Expression of actinonin mutase α subunit	Kanamycin
pJF11	pET28a with *actM* insert. N-terminal His tag.	Expression of actinonin mutase β subunit	Kanamycin
pJF12	pET28a with *actN* insert. C-terminal His tag.	Expression of actinonin MeaB-like protein	Kanamycin
pJF13	pET28a with *actJ* insert. C-terminal His tag.	Expression of actinonin asparagine synthetase	Kanamycin
pJF14	pET28a with *actI* insert. C-terminal His tag.	Expression of actinonin N-oxygenase	Kanamycin
pJF15	pACYCDuet-1 with *actD* insert (N-terminal His tag) and *actM* insert (N-terminal His tag)	Dual expression of actinonin mutase α and β subunits.	Chloramphenicol
pJAM2: *lydA*	pJAM2 with *lydA* insert. C-terminal His tag.	Expression of lydiamycin PDF in *M. smegmatis*	Kanamycin
pJAM2: *msPDF*	pJAM2 with *msPDF* (WP_011727226.1) insert. C-terminal His tag.	Expression of *M. smegmatis* housekeeping PDF in *M. smegmatis*	Kanamycin
pET28a: *cinF*	pET28a with codon optimised *cinF* (CBW54676.1) insert. N-terminal His tag. Synthesised by TWIST Bioscience.	Expression of *cinF* CCR in *E. coli*	Kanamycin

pET28a: *revS*	pET28a with codon optimised *revS* (BAK64635.1) insert. N-terminal His tag. Synthesised by TWIST Bioscience.	Expression of *revS* Acyl-CoA ligase in *E. coli*	Kanamycin

6.1.4 Chemicals, Reagents and Media

Unless otherwise stated, all chemicals and media components were purchased from Sigma Aldrich, except for the following: agar (Melford), NaCl and glucose (Fisher Scientific), yeast extract (Merck), soya flour (Holland and Barrett) and peptone (BD Biosciences). All enzymes were supplied from New England Biolabs (NEB) unless otherwise specified. All solvents for extractions and chromatography were supplied by Fisher Scientific. Ultrapure water was obtained using a Milli-Q purification system (Merck). All media was autoclaved prior to use and chemical solutions were filter sterilised using a 0.22 µM syringe filter.

All media used are defined below and were made up in 1L of Milli-Q water. Solid media were made with 2% agar unless stated otherwise. Antibiotics were added where necessary to the following concentrations: Kanamycin (50 µg/mL for *E. coli* and 20 µg/mL for *M. smegmatis*); Chloramphenicol (25 µg/mL), Streptomycin (100 µg/mL).

Table 12. Bacterial growth media used during this study.

Medium	Application	Ingredients
Lysogeny Broth (LB)	*E. coli* culture, *R. fascians* culture	1% Bacto-tryptone, 0.5% yeast extract, 1% NaCl, pH 7.0
SOC	*E. coli* transformation recovery medium	2% tryptone, 0.5% yeast extract, 0.058% NaCl, 0.2% $MgCl_2$, 0.25% $MgSO_4$, 0.36% glucose
Autoinduction Media (AIM)	*E. coli* protein expression	1.2% tryptone, 2.4% yeast extract, 0.33% $(NH_4)_2SO_4$, 0.68% KH_2PO_4, 0.71% Na_2HPO_4, 0.05% glucose, 0.2% lactose, 0.015% $MgSO_4$
Lennox Broth (L)	*Pseudomonas* culture	1% Bacto-tryptone, 0.5% yeast extract, 1% NaCl, 0.1% glucose
King's Broth with Mg (KB+Mg)	*Pseudomonas* Ps682 culture	2% Peptone, 1.5% Glycerol, 0.16% K2HPO4, 0.5% 1M MgSO4 (after autoclaving)
Middlebrook 7H9 Broth with OADC Enrichment (MOADC liquid)	*M. smegmatis* liquid culture	0.47% Middlebrook 7H9 broth base, 0.2% glycerol, 10% MOADC growth supplement (after autoclaving)

Middlebrook 7H10 Agar with OADC Enrichment (MOADC solid)	*M. smegmatis* solid culture	1.9% Middlebrook 7H10 agar base, 0.5% glycerol, 10% MOADC growth supplement (after autoclaving)
Aurachin PM	*R. fascians*, comparative metabolomics	1% starch, 1% glucose, 1% glycerol. 0.25% corn steep powder, 0.5% bacto peptone, 0.2% yeast extract, 0.1% NaCl, pH 7.3
Rhodostreptomycin PM	*R. fascians*, comparative metabolomics	1% starch, 2% glucose, 2.5% soytone. 0.4% dry yeast, 0.1% beef extract, 0.005% K2HPO4, 0.2% NaCl, pH 7.0
Minimal media A (MinA)	*R. fascians*, comparative metabolomics	1.05% K2HPO4, 0.45% KH2PO4, 0.1% (NH4)2SO4, 0.05% sodium citrate, 0.025% MgSO4-7H2O, 0.001% thiamine with 20mM pyruvate and 5mM histidine
Yeast Extract Broth (YEB)	*R. fascians*, comparative metabolomics	0.5% peptone, 0.1% yeast extract, 0.5% beef extract, 0.5% sucrose, 0.5% MgSO4-7H2O, pH 7.2
Actinonin PM	*R. fascians*, comparative metabolomics	1% glucose, 1% DIFCO soluble starch, 2% corn liquor step, 2% soy flour, 0.25% NH4Cl, 0.3% NaCl, 0.6% CaCO3, pH 6.2
Matlystatin PM	*R. fascians*, comparative metabolomics	3% glucose, 7% glycerol, 1% bacto peptone, 1% soy flour, 1% corn steep liquor, 0.5% MgSO4, 0.5% NH4NO3, 0.5% NaCl
Bottromycin PM	*R. fascians*, comparative metabolomics	1% glucose, 1.5% starch, 0.5% yeast extract, 1% soy flour, 0.5% NaCl, 0.3% CaCO3, pH 7.0
Screening medium 7 (SM7)	*R. fascians*, comparative metabolomics	2.09% MOPS, 1.5% L-proline, 2% glycerol, 0.25% sucrose, 0.15% L-glutamate, 0.05% NaCl, 0.2% K2HPO4 with 2mM MgSO4, 0.2mM CaCl2, 0.5% trace salts No. 1, pH 6.5
Screening medium 12 (SM12)	*R. fascians*, comparative metabolomics	1% soy flour, 5% glucose, 0.4% peptone, 0.4% beef extract, 0.1% yeast extract, 0.25% NaCl, 0.5% $CaCO_3$, pH 7.6
Soya Flour Mannitol (SFM)	*S. scabies* bioassay	2% mannitol, 2% soy flour, 100 mM CaCl2
Instant Potato Medium (IPM)	*S. scabies* bioassay	2% Instant mashed potato

Soft Nutrient Agar (SNA)	Spot-on-lawn bioassays	0.8% Difco Nutrient Broth, 0.7% agar

Table 13. Plant growth medium used in this study.

Medium	Application	Ingredients
½ MS + 3% Sucrose	Tobacco plant growth	0.22% Murashige & Skoog Medium Including Vitamins, 3% sucrose, 0.8% agar

6.1.5 Primers

All primers were synthesised by Eurofins Genomics or Integrated DNA Technologies (IDT) at a 25 nanomole synthetic scale with standard desalting. Primers were diluted to 100 µM in Milli-Q water and stored at -20 ^{o}C. Working stocks were prepared at 10 µM and stored at 4 ^{o}C.

Table 14. Primers used during this study.

Primer Name	Sequence (5' to 3')	Application	Restriction Site
M CCR For	GATACACATATGGACGCTCTCGCCGAGGC	Amplification of *matL* for pJF01 construction	NdeI
M CCR Rev	GATACAAAGCTTGCAAAGCAGGGGATCAAAGCGG		HindIII
M Epi For	GATACACATATGTTCACGCGCGTCGACCACGT	Amplification of *matE* for pJF02 construction	NdeI
M Epi Rev	GATACAAAGCTTGGGGAACGAACACTGGTCCTCCGAAC		HindIII
M Mut For	GATACACATATGGCGGGCCCAGGCGAGTTC	Amplification of *matB* for pJF03 construction	NdeI
M Mut Rev	GATACCAAGCTTCTCCTCGTTTCGCTTGGTTGCCG		HindIII
CO M B12 For	GATACACATATGGTTCGCCGTGTCCGTGTTGT	Amplification of codon optimised *matQ* for pJF04 construction	NdeI
CO M B12 Rev	TGTATCAAGCTTTTATTCCGCGACTTCCAGAG		HindIII
CO M MeaB For	GATACACCATGGTTAGGACCTGGCGGAGCAGC	Amplification of codon optimised	NcoI

CO M MeaB Rev	TCTATCAAGCTTGTGGTGGTGGTGGTGGTGAAATGCATCCGGCAACCAAC	*matR* for pJF05 construction	HindIII
M AspSyn For	AAAAAACCATGGATGTGTGGCATCAC	Amplification of *matC* for pJF06 construction	NcoI
M AspSyn Rev	TTTTTTAAGCTTTCACAGGGCCAGCC		HindIII
M Noxy For	GATACACCATGGATGAGGCGGCG	Amplification of *matD* for pJF07 construction	NcoI
M Noxy Rev	TGTATCAAGCTTTCAGATCTCCCCGAT		HindIII
A CCR For	CATAGACATATGCACAGCATGATCGAAGCCGT	Amplification of *actF* for pJF08 construction	NdeI
A CCR Rev	GATACAAAGCTTCACGGCAGCAGCGGCCCG		HindIII
A Epi For	GATACACATATGATCACTCGTATCGACCACATCGGC	Amplification of *actC* for pJF09 construction	NdeI
A Epi Rev	GATACAAAGCTTGGTACCGTCATCGGCTGACCGC		HindIII
A Mut For	CAGCCAGACCATGGGTAAACCCG	Amplification of *actD* for pJF10 construction	NdeI
A Mut Rev	GATACCAAGCTTTCCTGCCTGCGGGAGCGCGTCCTCGATC		HindIII
A B12 For	GATACACATATGGGTGTGTCGGGTCCGATCCG	Amplification of *actM* for pJF11 construction	NdeI
A B12 Rev	GATACAAAGCTTCACGGTGAGCTTCCGGCGGC		HindIII
A MeaB For	GATACACCATGGGTGGATGTGGCAGAGCTGGTGGC	Amplification of *actN* for pJF12 construction	NcoI
A MeaB Rev	GATACACCATGGTGGATGTGGCAGAGCTGGTGG		HindIII
A AspSyn For	AAAAAACCATGGATGTGTGGCATCAC	Amplification of *actJ* for pJF13 construction	NcoI
A AspSyn Rev	TTTTTTAAGCTTTCACAGGGCCAGCC		HindIII
A Noxy For	GATACACCATGGATGTACGCAAGTGCT	Amplification of *actI* for pJF14 construction	NcoI
A Noxy Rev	TGTATCAAGCTTTCACGGGGCGGCGTGCA		HindIII

pACAB12F1	GATACAGAATTCCATCACCATCATCACCACAGCATGGGTGTGTCGGGTCCGATCC	Amplification of *actM* and *actD* for pJF15 construction	EcoRI
pACAB12R1	TGTATCAAGCTTTCAGGCGGCGACGGGGC		HindIII
pACAMutF1	GATACAGATATCATGGGTAAACCCGGAGAGTACCCCT		EcoRV
pACAMutR1	GGTATCCCTAGGGTGGTGATGATGGTGATGGCTGCCGGCCTGCGGGAGC		AvrII
pJAM LydA For	GATACAGGATCCATGCCTGTCTCTGAACTTCTGC	Amplification of *lydA* for pJAM2:*lydA* construction	BamHI
pJAM LydA Rev	GATACATCTAGACTATCGTGTTGGCCAATCGTTG		XbaI
pJAM MsPDF For	CATGCCCGAGGTAGTTTTCGGATCCATGGCCGTCGTCCCGATCC	Amplification of *M. smegmatis* housekeeping PDF for pJAM2:*MsPDF* construction	Gibson
pJAM MsPDF Rev	AGTGGTGGTGGTGGTGGTGTCTAGATCAGTGCCCGAACGGATCG		Gibson
Chk T7 For	TAATACGACTCACTATAGGG	For sequencing and checking of pET28a MCS	
Chk T7 Rev	GCTAGTTATTGCTCAGCGG		

6.2 General Methods

6.2.1 *Escherichia coli*

6.2.1.1 *E. coli* growth and maintenance

Unless otherwise stated, *E. coli* strains were grown in liquid LB medium at 37 ^{o}C with shaking at 250 rpm for 16-18 hours. *E. coli* strains were also grown on solid LB medium at 37 ^{o}C until colonies were visible (18-24 hours). Plates were stored at 4 ^{o}C. Long-term *E. coli* stocks were stored at -70 ^{o}C in 25% glycerol.

6.2.1.2 *E. coli* transformation

6.2.1.2.1 Making *E. coli* chemically competent cells

10 mL of LB was inoculated with a single colony of the *E. coli* strain and grown overnight at 37 ^{o}C. 500 mL of LB (in 1 L flask) was inoculated with the pre-culture (1% v/v) and grown at 37 ^{o}C until an OD_{600} of 0.4 is reached. The following steps were performed at 4 ^{o}C or on ice.

The culture was transferred to ten 50 mL tubes and centrifuged at 1000 x *g* for 8 minutes. The supernatant was discarded by decanting. Each pellet was resuspended gently in 20 mL sterile competency buffer (15% glycerol, 7.5% 1M $CaCl_2$ in Milli-Q H_2O). Two suspensions were combined and centrifuge as above. The supernatant was discarded by decanting. Each pellet was resuspended gently in 20 mL sterile competency buffer and centrifuged as above. The supernatant was discarded by decanting. Each pellet was resuspended gently in 4 mL sterile competency buffer and 100 µL aliquots were dispensed into microcentrifuge tubes. These were flash frozen in liquid N_2 and stored at -70 ^oC.

6.2.1.2.2 *E. coli* heat-shock transformation

Chemically competent cells were thawed on ice. Plasmid DNA was mixed with the cells (2 µL purified vector per 100 µL cells, 10 µL ligation mixture per 100 µL cells) and left on ice for 20 minutes. The tubes were placed in a 42 ^oC water bath for 45 seconds and then placed on ice for 2 minutes. The cells were mixed with 1 mL SOC medium and incubated at 37 ^oC for 1 hour with shaking at 250 rpm. 100-300 uL of the culture was plated onto LB agar plates (with appropriate antibiotic) and grown overnight.

Commercial 5-alpha high efficiency *E. coli* competent cells (NEB C2987H) were also used for chemical transformation, according to the manufacturer's protocol, when required.

6.2.1.2.3 Making *E. coli* electrocompetent cells

10 mL of LB was inoculated with a single colony of the *E. coli* strain and grown overnight at 37 ^oC. 500 mL of LB (in 1 L flask) was inoculated with the pre-culture (1% v/v) and grown at 37 ^oC until an OD_{600} of 0.4 was reached. The following steps were performed at 4 ^oC or on ice. The culture was transferred to ten 50 mL tubes and centrifuged at 800 x *g* for 20 minutes. The supernatant was discarded by decanting. Each pellet was resuspended gently in 50 mL sterile 10% glycerol. The suspensions were centrifuged as above, and the supernatant was discarded by decanting. Each pellet was resuspended gently in 25 mL sterile 10% glycerol and centrifuged as above. The supernatant was discarded by decanting. Each pellet was resuspended in 5 mL sterile 10% glycerol and combined into a single 50 mL tube. This was centrifuged as above. The supernatant was discarded by decanting. The pellet was resuspended in 1 mL of 10% glycerol and 50 µL aliquots were dispensed into microcentrifuge tubes. These were flash frozen in liquid N_2 and stored at -70 ^oC.

6.2.1.2.4 *E. coli* electrotransformation

Electrocompetent cells were thawed on ice and mixed with 0.5 – 1.0 µL of plasmid DNA. This was transferred to a 2 mm electroporation cuvette. The outside of the cuvette was dried, and it was ensured that no bubbles had formed. The cells were electroporated at 25 µFD, 2.5 KV, 200 Ω. The cells were mixed with 1 mL SOC medium and incubated at 37 ^{o}C for 1 hour with shaking at 250 rpm. 0.1 – 1.0 mL of the culture was plated onto LB agar plates (with appropriate antibiotic) and grown overnight.

6.2.1.3 *E. coli* plasmid isolation

Single colonies were picked and inoculated in 10 mL LB with appropriate antibiotic selection and grown overnight. Cultures were centrifuged for 10 minutes at 1,538 x *g* in an Eppendorf Centrifuge 5810. Plasmid DNA was purified using the Wizard Plus SV Miniprep kit (Promega) according to the manufacturer's protocol. DNA was eluted using 50 µL Milli-Q H_2O.

6.2.2 *Rhodococcus fascians*

6.2.2.1 *R. fascians* growth and maintenance

Unless otherwise stated, *R. fascians* strains were grown in liquid LB medium at 28 ^{o}C with shaking at 250 rpm for 24-48 hours. *R. fascians* was also grown on solid LB medium at 28 ^{o}C until colonies were visible (2-3 days). Plates were stored at 4^{o}C. Long term *R. fascians* stocks were stored at -70 ^{o}C in 25% glycerol.

6.2.2.2 *R. fascians* gDNA extraction

gDNA was extracted from *R. fascians* using the Fast DNA SPIN Kit for Soil (MP Biomedicals), according to the manufacturer's protocol.

6.2.3 *Mycobacterium smegmatis*

6.2.3.1 *M. smegmatis* growth and maintenance

Unless otherwise stated, *M. smegmatis* was grown in liquid MOADC medium at 37^{o}C with shaking at 250 rpm for 18-24 hours. Aerobic conditions were maintained by taping lids slightly open. Sterile glass beads and 0.25% TWEEN-80 were added to reduce cellular aggregation. *M. smegmatis* strains were also grown on solid MOADC medium at 37 ^{o}C until colonies were visible to the naked eye (2-3 days). Plates were stored at 4 ^{o}C. Long-term *M. smegmatis* stocks were stored at -70 ^{o}C in 25% glycerol.

6.2.3.2 *M. smegmatis* transformation

6.2.3.2.1 Making *M. smegmatis* electrocompetent cells

10 mL of MOADC was inoculated with a single colony of *M. smegmatis* and grown overnight at 37 °C. 100 mL of MOADC (in 250 mL flask) was inoculated with the pre-culture (1% v/v) and grown at 37 °C until an OD_{600} of 0.2-0.8 was reached. The culture was incubated on ice for 1.5 hours. The following steps were performed at 4 °C or on ice. The culture was transferred to two 50 mL tubes and centrifuged at 5,000 x *g* for 10 minutes. The supernatant was discarded by decanting. Each pellet was resuspended in 50 mL sterile 10% glycerol and centrifuged as above. The supernatant was discarded by decanting. Each pellet was resuspended in 25 mL sterile 10% glycerol and centrifuged at 2,000 x *g* for 10 minutes. The supernatant was discarded by decanting. This wash step was repeated once more. Each pellet was resuspended in 500 µL sterile 10% glycerol and 50 µL aliquots were dispensed into microcentrifuge tubes. These were flash frozen in liquid N_2 and stored at -70 °C.

6.2.3.2.2 *M. smegmatis* electrotransformation

Electrocompetent cells were thawed on ice and mixed with 1.0 – 5.0 µL of plasmid DNA. This was transferred to a 2 mm electroporation cuvette and placed on ice for 1 minute. The outside of the cuvette was dried, and it was ensured that no bubbles had formed. The cells were electroporated at 25 µFD, 2.5 KV, 1000 Ω. The cells were mixed with 1 mL MOADC medium and transferred to a 15 mL tube. The cells were incubated at 37 °C for 2 hours with shaking at 250 rpm. 0.1 – 1.0 mL of the culture was plated onto MOADC agar plates (with appropriate antibiotic) and grown for 1-3 days.

6.2.3.3 *M. smegmatis* gDNA extraction

gDNA was extracted from *M. smegmatis* using the Fast DNA SPIN Kit for Soil (MP Biomedicals), according to the manufacturer's protocol.

6.2.4 *Pseudomonas*

6.2.4.1 *Pseudomonas* growth and maintenance

Unless otherwise stated, *Pseudomonas* strains were grown in liquid LB medium at 30 °C with shaking at 250 rpm for 16-20 hours. *Pseudomonas* strains were also grown on solid LB medium at 30 °C until colonies were visible (1-2 days). Plates were stored at 4 °C. Long-term *Pseudomonas* stocks were stored at -70 °C in 25% glycerol.

6.3 Cloning and Sequencing

6.3.1 *E. coli* colony PCR

Single colonies were picked with sterile toothpicks and inoculated into 20 µL sterile Milli-Q water for use as template DNA. GoTaq polymerase (Promega) was used for colony PCR reactions. The PCR reaction mixture and program shown in **Table 15** and **Table 16** was performed using a T100 Thermal Cycler (Bio-Rad).

Table 15. Reaction conditions for colony PCR of *E. coli* template

Component	Volume (µL)
GoTaq Polymerase	0.2
Forward Primer (10 µM)	1.0
Reverse Primer (10 µM)	1.0
DMSO	1.0
Green GoTaq Buffer (5X)	4.0
$MgCl_2$ (25 mM)	2.8
dNTPs (10 mM)	0.4
Template DNA	1.0
H_2O	7.6
TOTAL	20.0

Table 16. PCR cycling conditions for *E. coli* colony PCR

Temperature (°C)	Time	Number of cycles
98	4 mins	x1
98	30 secs	x30
Annealing Temp.	30 secs	
72	45 secs (per kb)	
72	5 mins	x1

6.3.2 High-Fidelity PCR for cloning

Q5 polymerase (NEB) was used for high-fidelity PCR reactions for cloning purposes. The PCR reaction mixture and program shown in **Table 17 and Table 18** was performed using a T100 Thermal Cycler (Bio-Rad). Template DNA was usually gDNA.

Table 17. Reaction conditions for high fidelity PCR using Q5 polymerase

Component	Volume (µL)
Q5 Polymerase	0.5
Forward Primer (10 µM)	2.5
Reverse Primer (10 µM)	2.5
DMSO	2.5
Q5 Buffer (5X)	10.0
dNTPs (10 mM)	1.0
Template DNA	150 ng
H_2O	Up to 50
TOTAL	50

Table 18. PCR cycling conditions for high fidelity PCR using Q5 polymerase

Temperature (°C)	Time	Number of cycles
95	2 mins	x1
95	45 secs	x35
Annealing Temp.	45 secs	
72	20 secs (per kb)	
72	5 mins	x1

6.3.3 Agarose Gel Electrophoresis

0.8% Agarose gels with 3 uL ethidium bromide per 100 mL were used for electrophoresis of DNA. DNA samples were mixed with 6x Gel Loading Dye (NEB) and loaded into the agarose gel wells. 5 µL of 1Kb Plus DNA Ladder (NEB) was also loaded. Gel electrophoresis proceeded at 90V for 50-75 minutes in 1x TBE (Tris/Borate/EDTA) running buffer. Gels were visualised and imaged using UV light.

6.3.4 Purification of DNA from Agarose Gels

50 µL Q5 PCR reactions were separated by gel electrophoresis and viewed under low intensity UV light. A scalpel was used to excise the DNA band into a 2 mL microcentrifuge tube. The DNA was purified using the GFX PCR DNA and Gel Band Purification Kit (Cytiva) according to the manufacturer's protocol. The DNA was eluted using 2x 10 µL aliquots of 70 ^{o}C H_2O.

6.3.5 DNA Digestions

All restriction enzymes were sourced from NEB and used with the supplied CutSmart Buffer. Reactions were performed in microcentrifuge tubes and proceeded at 37 ^{o}C for 1 hour (<20 µL scale) or overnight (100 µL scale). Reaction mixtures were prepared as described in **Table 19 and Table 20**, and immediately gel purified or stored at -80 ^{o}C for 30 minutes to inactivate the enzyme.

Table 19. Reaction mixture for digestion of purified PCR product

Component	Volume (µL)
Linear DNA	10.0
CutSmart Buffer (10x)	1.2
Enzyme A	0.4
Enzyme B	0.4
TOTAL	12.0

Table 20. Reaction mixture for digestion of plasmid

Component	Volume (µL)
Plasmid DNA	10 µg
CutSmart Buffer (10x)	10.0
Enzyme A	2.0
Enzyme B	2.0
Milli-Q H_2O	Up to 100
TOTAL	100

6.3.6 DNA concentration quantification

The concentration and purity of gDNA, purified amplified DNA and digested vector backbones was determined using a NanoDrop 2000 spectrophotometer (Thermo Scientific) according to the manufacturer's protocol.

6.3.7 DNA Ligations

Ligations were performed with T4 DNA Ligase (Thermo) using a 3-fold molar excess of insert to vector. Reaction mixtures were prepared as described in **Table 21** and proceeded at 18 ^{o}C overnight. A negative control ligation reaction lacking insert was always performed to check that the vector was sufficiently digested. The entire 10 µL reaction mixture was used to transform chemically competent *E. coli* DH5α cells (see **Section 6.2.1.2.2**).

Table 21. Reaction mixture for ligation reaction

Component	Volume (µL)
T4 Ligase (5U/µL)	0.5
T4 Ligase Buffer (10x)	1.0
PEG (50% w/v)	1.0
Digested Vector	Variable
Digested Insert	4.0
Milli-Q H_2O	Up to 10
TOTAL	10.0

6.3.8 Gibson Assembly

Gibson assembly was performed using NEBuilder HiFi DNA Assembly Master Mix (NEB) using the manufacturer's protocol for the assembly of 2-3 fragments. The reaction was incubated at 50 ^{o}C for 60 minutes and 2 µL of the reaction mixture used to transform chemically competent *E. coli* DH5α cells.

6.3.9 Sequencing

All vectors were sequenced by Eurofins Genomics using a Mix2Seq Kit according to the manufacturer's protocol. Each T7-based vector was sequenced once with the 'check T7' forward and reverse primers along with the primers used to amplify the insert. DNA sequences were analysed using BLAST (https://blast.ncbi.nlm.nih.gov) to ensure that the PCR amplification had not introduced any mutations.

6.4 Protein purification

E. coli expression strains were grown at a 1 L scale and proteins were purified by affinity chromatography using gravity columns.

6.4.1 *E. coli* Induction

6.4.1.1 *E. coli* induction using isopropylthio-β-galactoside (IPTG)

10 mL of LB with appropriate antibiotics was inoculated with either a single colony or 20 µL of thawed glycerol stock of the *E. coli* expression strain and grown overnight at 37 °C. 1 L of LB (in 2 L flask) with appropriate antibiotics was inoculated with 10 mL of the pre-culture and grown at 37 °C until an OD_{600} of 0.6 was reached. 100 µL of a 1M IPTG solution was added to the culture (0.1 mM final concentration). The culture was incubated for 4 hours at 30 °C.

6.4.1.2 *E. coli* induction using autoinduction medium (AIM)

10 mL of LB with appropriate antibiotics was inoculated with either a single colony or 20 µL of thawed glycerol stock of the *E. coli* expression strain and grown overnight at 37 °C. 1 L of AIM (in 2 L flask) with appropriate antibiotics was inoculated with 10 mL of the pre-culture and grown at 18°C for 1-3 days with shaking at 250 rpm.

6.4.2 Protein Purification Buffers

Buffers were prepared as described in **Table 22** in Milli-Q water. Buffers were stored at 4 °C.

Table 22. Composition of protein purification buffers.

Buffer	Composition
Lysis and Binding	20 mM TRIS-HCl, 500 mM NaCl, 10 mM Imidazole, 10% glycerol, pH 7.9
Wash #1	20 mM TRIS-HCl, 500 mM NaCl, 20 mM Imidazole, 10% glycerol, pH 7.9
Wash #2	20 mM TRIS-HCl, 500 mM NaCl, 40 mM Imidazole, 10% glycerol, pH 7.9
Elution	20 mM TRIS-HCl, 500 mM NaCl, 250 mM Imidazole, 10% glycerol, pH 7.9
Storage	50 mM TRIS-HCl, 100 mM NaCl, 10% glycerol, pH 7.9

6.4.3 *E. coli* cell lysis and clarification

The *E. coli* culture was split equally between two 1000 mL centrifuge bottles and centrifuged at 12,000 x *g*, 4 °C for 30 minutes using a Fiberlite F9-6x1000 rotor (Thermo) in a Sorvall LYNX 6000 centrifuge (Thermo). The following steps were performed at 4 °C or on ice using ice-cold buffers. The supernatant was discarded by decanting. Each pellet was resuspended in 6 mL Lysis and Binding Buffer and these suspensions were combined. The cell suspension was vortexed until homogenised. The cell suspension was lysed using an EmulsiFlex B-15 homogeniser (Avestin) according to the manufacturer's protocol. The cells were homogenised 3-4 times for full lysis.

The lysate was split equally between two 30 mL Nalgene Oak Ridge tubes (Thermo) and centrifuged at 45,000 x *g*, 4 °C for 30 minutes using a Fiberlite F20-12x50 rotor (Thermo) in a Sorvall LYNX 6000 centrifuge (Thermo). The supernatant was filtered using a 0.45 µm syringe filter. A 50 µL aliquot of the crude clarified lysate was stored for later SDS-PAGE analysis.

6.4.4 Protein Purification by Gravity Affinity Chromatography

The *E. coli* BL21, SoluBL21 and Rosetta expression strains were purified by a single round of nickel affinity chromatography. The *E. coli* NiCo21 expression strain was purified by chitin affinity chromatography followed by a single round of nickel affinity chromatography. Ice-cold buffers were used, and these steps were performed at 4 °C.

6.4.4.1 Chitin Affinity Chromatography

An Econo-Pac chromatography column (Bio-Rad) was loaded with 4 mL of chitin resin (NEB) and the storage buffer was drained. 15 mL of Milli-Q H_2O was added, rotated for 5 minutes and drained. 15 mL of Lysis and Binding Buffer was added, rotated for 5 minutes and drained. The clarified *E. coli* NiCo21 lysate was added and mixed by rotation for 1 hour. The resin was allowed to settle, and the flowthrough was collected for further purification.

6.4.4.2 Nickel Affinity Chromatography

An Econo-Pac chromatography column (Bio-Rad) was loaded with 2 mL of Ni-NTA Agarose (Qiagen) and the storage buffer was drained. 15 mL of Milli-Q H_2O was added, rotated for 5 minutes and drained. 15 mL of Lysis and Binding Buffer was added, rotated for 5 minutes and drained. The chitin-treated NiCo21 flowthrough or clarified *E. coli* BL21, SoluBL21 or Rosetta lysate was added and mixed by rotation for 1 hour. The buffer was drained, and an aliquot labelled as 'non-binding' stored for later SDS-PAGE analysis. 15 mL of Wash Buffer #1 was added, rotated for 10 minutes, drained and a flowthrough sample retained. The

above step was repeated once. 15 mL of Wash Buffer #2 was added, rotated for 10 minutes, drained and a flowthrough sample retained. The above step was repeated once. 2.5 mL of Elution Buffer was added, rotated for 10 minutes, drained and the flowthrough retained in a 15 mL tube. The above step was repeated once, and the elution flowthroughs combined.

6.4.5 Protein Buffer Exchange and Storage

The protein buffer was exchanged for Storage Buffer. The selected centrifugal filter molecular weight cut-off must be less than the purified protein weight to ensure that the protein is retained. The centrifugal filter is placed within a 50 mL tube. Ice-cold buffers were used, and these steps were performed at 4 ^{o}C.

The affinity chromatography eluent was applied to the Amicon Ultra-15 Centrifugal Filter (Merck) and centrifuged at 5,000 x *g* at 4 ^{o}C in a 6x50mL 7588 rotor (Heraeus) using a Biofuge Primo centrifuge (Sorvall) until 250 µL of prefiltered sample remained (~30 minutes). 5 mL of Storage Buffer was added, and the tube was centrifuged as above until 250 µL of prefiltered sample remained. This step was repeated twice more. The final 250 µL protein sample was transferred to a microcentrifuge tube and mixed with 25 µL of 100% glycerol. 20 µL aliquots were transferred to 0.2 mL PCR tubes and stored at -70 ^{o}C.

6.4.6 Protein Analysis

6.4.6.1 Protein Analysis by SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

The above protein purification samples (crude lysate, non-binding, washes 1-4 and purified) were analysed by SDS-PAGE to determine protein mass and purity. Purified protein samples were often diluted 10x or 100x in storage buffer to avoid overloading the SDS-PAGE gel. The following buffer was prepared as described in **Table 23** in Milli-Q water.

Table 23. Composition of SDS-PAGE buffers.

Buffer	Composition
10X SDS-PAGE Tris-Gly Running Buffer	3% TRIS, 14.4% Glycine, 1% SDS

10 µL protein samples were mixed 1:1 with 2X SDS-PAGE Loading Buffer (Invitrogen), heated at 105 ^{o}C for 5 minutes and centrifuged at 15,871 x *g* for 1 minute. The 20 µL protein samples and 5 µL of Color Prestained Protein Standard, Broad Range ladder (NEB) were loaded into a Novex 12% Tris-Glycine SDS-PAGE gel (Invitrogen). Electrophoresis proceeded in 1X SDS-PAGE Tris-Gly Running Buffer at 200 V for 30 minutes. Gels were

visualised by shaking in InstantBlue Protein Stain (Expedeon) for 1-18 hours. Gels were washed in water with shaking for 1 hour, three times, and imaged.

6.4.6.2 Protein Analysis by LC-MS

Once separated by SDS-PAGE, protein bands were sequenced to determine the identity of the purified protein. This was performed by Carlo de Oliveira Martins and Gerhard Saalbach (John Innes Centre). In brief, SDS-PAGE protein bands were excised, de-stained, washed and dehydrated. The gel slices were incubated with trypsin. The resulting peptides were analysed by LC-MS/MS using a Synapt G2-Si mass spectrometer equipped with an Acquity UPLC (Waters). Peptide sequences were analysed using Scaffold (Proteome Software) and compared against the target peptide sequence and the *E. coli* protein database to determine the identity of the protein band.

6.5 Octenoyl-CoA and Octenoyl-SNAC Synthesis

6.5.1 Octenoyl-CoA Synthesis

6.5.1.1 Mixed Anhydride Method

2-octenoic acid (154 µL, 1.0 eq) was added to dry dichloromethane (7.0 mL) at 0 ^{o}C. Triethylamine (334 µL, 2.7 eq) and ethyl chloroformate (213 µL, 2.5 eq) were added and the reaction mixture was stirred for 2 hours at 0 ^{o}C. The solvent was removed using rotary evaporation and the mixed anhydride residue suspended in tetrahydrofuran (5.0 mL). A solution of 100mg (0.15 eq) coenzyme-A (in 5 mL $NaHCO_3$ in H_2O, pH 8.0) was added and the mixture stirred at room temperature for three hours. The reaction was dried by rotary evaporation and then lyophilisation.

6.5.1.2 PyBOP method

Coenzyme A (100 mg, 1 eq), 2-octenoic acid (43 µL, 2 eq), PyBOP (104 mg, 1.5 eq) and K_2CO_3 (72 mg, 4 eq) were dissolved in 10 mL THF/H_2O (1:1). The reaction was stirred and proceeded at room temperature for 2.5 hours. Samples were taken at 1.25 and 2.5 hours for LC-MS analysis. The reaction was dried by rotary evaporation and then lyophilisation.

6.5.2 Octenoyl-CoA purification

6.5.2.1 Flash Chromatography

Crude reaction mixtures were separated by flash chromatography using a Biotage Isolera system equipped with a 12 g SNAP Ultra C18 cartridge (Biotage). The solvent system used

was ACN (B) against H_2O (A) and the flowrate was 12 mL/min. The cartridge was equilibrated with 3 column volumes (CVs) of 5% ACN. Crude reaction mixtures were solubilised in 3 mL 50% MeOH and loaded onto the cartridge. Separation proceeded with a linear 5-100% gradient for 10 CVs and a 2 CV isocratic hold at 100% B. Elution was monitored at 260 nm and 210 nm UV absorption and 9 mL fractions were collected. Each fraction was sampled and analysed by LC-MS (see **Section 6.5.3**). Fractions containing octenoyl-CoA were combined and dried by rotary evaporation.

6.5.2.2 Semi-Preparative scale HPLC

Flash chromatography fractions were further separated by reverse phase HPLC using an UltiMate 3000 HPLC instrument (Thermo) equipped with a semi-preparative Luna C18 100 mm x 10 mm column (Phenomenex). The solvent system used was ACN (+0.1% TFA) (B) against H_2O (A) and the flowrate was 5 mL/min. The column was equilibrated with 5% B for 3 CVs. The sample was solubilised in 50% MeOH to a concentration of 7 mg/mL and filtered using a 0.45 µm PTFE filter (Whatman). The sample was injected and separated with an isocratic hold of 5% B for 2.5 CV and a linear gradient of 5-70% B over 16 CVs. The UV absorption at 265 nm was recorded. Peaks were manually collected and stored on ice. The samples were sampled for LC-MS analysis (see **Section 6.5.3**) and then lyophilised to dryness.

6.5.2.3 Strong Anion Exchange (SAX) HPLC

Following separation of the crude reaction mixture by flash chromatography, the efficacy of SAX chromatography was tested. This was performed on an UltiMate 3000 HPLC instrument (Thermo) equipped with a 50mm x 10mm POROS 50 HQ strong anion exchange column (Thermo). The solvent system used was 1 M NH_4HCO_3 in Milli-Q H_2O (B) against 5 mM NH_4HCO_3 in Milli-Q H_2O (A) and the flowrate was 7 mL/min. The column was equilibrated with 100% A for 5 CVs. The sample was solubilised in Milli-Q H_2O to a concentration of 10 mg/mL and filtered using a 0.45 µm PTFE filter (Whatman). 0.2 mL of the sample was injected and a linear gradient of 0-40 % B over 17 CVs, linear gradient of 40-100% B over 3.5 CVs and an isocratic hold at 100% B for 2.5 CVs was performed. The UV absorption at 265 nm was recorded. Peaks were manually collected and stored on ice. The samples were sampled for LC-MS analysis (see **Section 6.5.3**) and then lyophilised to dryness.

6.5.2.4 Preparative-scale HPLC

The most effective purification was the preparative-scale reverse phase HPLC separation of flash chromatography fractions. This used an UltiMate 3000 HPLC instrument (Thermo)

equipped with a preparative-scale Gemini-NX C18 (150mm x 21.1mm) column (Phenomenex). The solvent system used was ACN (+0.1% FA) (B) against 10 mM aqueous ammonium formate (pH 4.5) (A) and the flowrate was 20 mL/min. The column was equilibrated with 5% B for 1 CV. The sample was solubilised in 1 mL Milli-Q H_2O to a concentration of 32 mg/mL and filtered using a 0.45 µm PTFE filter (Whatman). The sample was injected (6 mg per injection) and separated with a linear gradient of 5-50% B over 8 CVs and an isocratic hold at 95% B for 1 CV.The UV absorption at 260 nm was recorded. Peaks were manually collected and stored on ice. The fractions were sampled for HR-MS analysis (see **Section 6.5.3**) and then lyophilised to dryness. Octenoyl-CoA HR-MS analysis is presented in **Appendix** Figure 100. HR-MS analysis of preparative-scale HPLC octenoyl-CoA sample.**Figure 100**. Octenoyl-CoA NMR analysis is presented in **Appendix** Figure 101**.**

6.5.3 Octenoyl-CoA Identification by LC-MS

6.5.3.1 LC-MS analysis

Samples were analysed using the Shimadzu ion-trap time-of-flight (IT-ToF) MS system, with separation achieved with an Agilent UHPLC with a Phenomenex Kinetex C18 column (2.6 µm, 50 x 2.1 mm, 100 Å). The sample was eluted with a linear gradient of 5-70% ACN (+0.1%FA) / H_2O (+0.1%FA) over 6 minutes with a flow rate of 0.6 mL/min and detected by the IT-ToF MS. Detection was performed in positive and negative mode and a mass range of *m/z* 200 – 2000. MS^2 events were also detected, with a mass range of *m/z* 200-1000. Data was analysed using LabSolutions software (Shimadzu).

6.5.3.2 HR-MS analysis

HR-MS/MS analysis was performed by Gerhard Saalbach and Carlo de Oliveira Martins (John Innes Centre). High-resolution mass spectra were acquired on a Synapt G2-Si mass spectrometer equipped with an Acquity UPLC (Waters). Aliquots of the samples were injected onto an Acquity UPLC® BEH C18 column, 1.7 µm, 1x100 mm (Waters) and eluted with a gradient of acetonitrile/0.1% formic acid (B) in water/0.1% formic acid (A) with a flow rate of 0.08 mL/min at 45 °C. The concentration of B was kept at 1% for 1 min followed by a gradient up to 40% B in 9 min, ramping to 99%B in 1 min, kept at 99%B for 2 min and re-equilibrated at 1%B for 4 min. MS data were collected in positive mode with the following parameters: resolution mode, positive ion mode, scan time 0.5 s, mass range m/z 50-1200 calibrated with sodium formate, capillary voltage = 2.5 kV; cone voltage = 40 V; source temperature = 125 °C; desolvation temperature = 300 °C. Leu-enkephalin peptide was used to generate a lock-mass calibration with 556.2766, measured every 30 s during the run. For MS2 fragmentation, a data directed analysis (DDA) method was used with the following parameters: precursor selected from the 4 most intense ions; MS2 threshold:

5,000; scan time 0.5 s; no dynamic exclusion. In positive mode, collision energy (CE) was ramped between 10-30 at low mass (m/z 50) and 15-60 at high mass (m/z 1200).

6.5.4 Octenoyl-SNAC Synthesis

This reaction was performed under nitrogen atmosphere. 2-octenoic acid (0.56 mL, 4 mmol) was dissolved in dimethylformamide (10 mL) and cooled to 0 ^{o}C . To this solution, DPPA (1.4 mL, 4 mmol) and triethylamine (1.8 mL, 13 mmol) was added. The reaction was stirred for 2 hours at 0^{o}C. N-acetyl cysteamine (3.6 mL, 34 mmol) was added and the reaction stirred at room temperature for 24 hours. The reaction mixture was transferred to a 0.5 L separatory funnel and water (25 mL) was added. Liquid/liquid extraction was performed with 4 x 50 mL ethyl acetate and the organic extract combined. The organic extract was dried over $MgSO_4$ and concentrated by rotary evaporation.

6.5.5 Octenoyl-SNAC Purification

Octenoyl-sNAC purification proceeded by normal phase flash chromatography using a Biotage Isolera equipped with a 25 g Sfär HP-Sphere normal phase cartridge (Biotage). The solvent system used was ethyl acetate (B) against hexane (A). The cartridge was equilibrated with 0% B for 3 CVs. The crude reaction mixture was solubilised in toluene and loaded onto the cartridge. Separation proceeded with a linear 0-70% B gradient over 2 CVs, a linear gradient of 70-100% B over 5 CVs and an isocratic hold at 100% B for 5 CVs. The UV absorbance at 265 nm was recorded. Each fraction was sampled and analysed by TLC (using ethyl acetate). TLC plates were stained using potassium permanganate stain. Fractions of interest were also analysed by LC-MS using the direct injection Advion Expression Compact MS instrument in positive mode to identify fractions containing octenoyl-SNAC (*m/z* = 244.13; $[M+H]^+$).

6.6 *In vitro* enzyme assays

6.6.1 CCR Reaction Conditions

In vitro reactions between CCR and acyl-CoA thioester substrates were prepared according to **Table 24**.

Table 24. Composition of CCR *in vitro* assay reactions.

Component	Volume (µL)	Final Concentration
1M TRIS-HCl (pH 8.0)	5.0	100 mM
140 mM NADPH	1.4	4 mM
20 mM enoyl-CoA or octenoyl-SNAC	2.5	1 mM
CCR protein preparation	2.5	unknown
120 mM $NaHCO_3$	13.8	33 mM
200 µg/mL carbonic anhydrase	2.5	10 µg/mL
Milli-Q H_2O	22.3	
TOTAL	50.0	

The negative control was substituted with denatured CCR enzyme (heated at 95 ^oC for 5 minutes and centrifuged at 15,871 x *g* for 1 minute). The *in vitro* assay reaction proceeded at 30 ^oC, 500 rpm for 30 minutes, unless otherwise stated. Reactions were quenched with 50 µl of -20^oC MeOH and centrifuged at 15,871 x *g* for 5 minutes. A 80 µl aliquot of the supernatant was transferred to a 2 mL LC-MS vial with an insert.

Where stated, the $NaHCO_3$ was replaced with $NaH^{13}CO_3$, at the same concentration of 33 mM.

6.6.2 LC-MS analysis of *in vitro* assays

Samples were analysed using the Shimadzu IT-ToF MS system, with separation achieved with an Agilent UHPLC with a Phenomenex Kinetex C18 column (2.6 µm, 50 x 2.1 mm, 100 Å). The sample was eluted with a linear gradient of 5-70% ACN (+0.1%FA) / H_2O (+0.1%FA) over 6 minutes with a flow rate of 0.6 mL/min and detected by the IT-ToF MS. Detection was performed in positive mode and a mass range of *m/z* 200 – 2000. MS^2 events were also detected, with a mass range of *m/z* 200-1000. Data was analysed using LabSolutions software (Shimadzu).

6.7 Bioinformatic Analysis of Warhead NP BGCs

6.7.1 Acyl-CoA Mutase Phylogeny and Co-association Analysis

6.7.1.1 Retrieving Representative Actinobacterial Acyl-CoA Protein Sequences

The acyl-CoA mutase phylogenetic tree was constructed as follows. The MatB protein sequence (WP_019634562.1) was submitted to the InterPro search (www.ebi.ac.uk/interpro/search/sequence/) in order to identify the methylmalonyl-CoA mutase InterPro entry (IPR006099). All actinobacterial protein sequences belonging to the IPR006099 InterPro entry (~13,000 sequences at time of analysis – February 2022) were downloaded in .FASTA format. Mutase protein sequences identified in a MatB BLAST analysis that were highly similar (>60%) but absent in the InterPro database were added to the sequence list (WP_019634562.1, WP_176725504.1, WP_210837750.1, WP_063760750.1, WP_189742219.1, WP_223771548.1, WP_104814753.1, WP_184992507.1, QES31823.1, QNF54062.1, QHF95772.1). Sequences were submitted to the Enzyme Similarity Tool (Enzyme Function Initiative; EFI-EST [271]: https://efi.igb.illinois.edu/efi-est/) to produce a protein sequence similarity network (SSN). Edges were drawn between nodes where the similarity e-value was less than e^{-50}. Protein sequences of >40% sequence similarity were collated into 1,615 representative nodes to generate the SSN. The SSN was exported and visualised using Cytoscape [353] (https://cytoscape.org/). The SSN node table was exported as a .csv file and opened using Excel. The UniProt retrieval (https://www.uniprot.org/id-mapping/) tool was used to download the .FASTA protein sequences of the first accession number of each representative node (**File 1**).

6.7.1.2 Building Phylogenetic Tree of Representative Actinobacterial Acyl-CoA Protein Sequences

The .FASTA protein sequences (**File 1**) were uploaded to the CIPRES server [354](http://www.phylo.org/) and an alignment was performed using the ClustalW method [355] with default parameters. The .FASTA alignment file was uploaded to the CIPRES server and submitted for phylogenetic tree building using the 'RAxML-HPC Blackbox' method [356], using default parameters. The program determined that a bootstrap value of n=350 was sufficient. The best scoring tree was uploaded to iTOL [273] (https://itol.embl.de/) for visualisation.

6.7.1.3 Co-association analysis

The co-association analysis pipeline starts with a SSN file. In order to perform a co-association analysis of just the first protein sequence of each representative node, **File 1** was submitted to EFI-EST and the network lacking clustering (ID=100%) was downloaded. This SSN was submitted to the Genome Neighbourhood Tool (Enzyme Function Initiative; EFI-GNT [272]: https://efi.igb.illinois.edu/efi-gnt/). Parameters of neighbourhood size = 15 and minimal co-occurrence percentage lower limit = 20 were set. The resultant coloured SSN was opened in Cytoscape and the node table exported as a .csv file. This was opened in Excel and the presence of Pfam domains of interest for each accession number was determined. The Pfam domains of interest were as follows: PF00107 (CCR), PF13669 (Acyl-CoA epimerase), PF02310 (Acyl-CoA mutase β subunit), PF03308 (MeaB-like), PF00733 (Asparagine synthetase), PF11583 (N-oxygenase), PF02770 (Acyl-CoA dehydrogenase), PF00501 (NRPS, adenylation), PF00668 (NRPS, condensation), PF00109 (PKS, ketosynthase), PF00550 (carrier) and PF00975 (Thioesterase). Excel was used to translate the EFI-GNT output into a binary matrix to denote whether each Pfam domain of interest was present in proximity to each Acyl-CoA mutase accession. This was converted into an iTOL binary annotation file, as detailed here https://itol.embl.de/help.cgi#binary. This annotation file was used to display the co-association analysis on the phylogenetic tree using iTOL.

6.7.1.4 Phylogenetic and co-association analysis of warhead clade

The warhead clade was identified in the above analysis. The sequences making up the representative nodes of the warhead clade were subject to another phylogenetic and co-association analysis. The EFI-EST SSN informed which protein sequences were present in each representative node. These .FASTA protein sequences, in addition to the A0A166QLU4 sequence as a phylogenetic outgroup, were retrieved as detailed above. The sequences absent in the UniProt database were also added, as detailed above. In total, 45 sequences were submitted to CIPRES for ClustalW alignment and RAxML-HPC Blackbox tree building, using default parameters. The program determined that a bootstrap value of n=300 was sufficient and de-replicated 20 sequences that were 100% identical. The tree was uploaded to iTOL.

The co-association analysis was performed using the same parameters as detailed above. Additional Pfam domains of interest were added: PF13434 (KtzI-like), PF04299 (KtzT-like), PF07690 (transporter) and PF01327 (peptide deformylase). This co-association analysis was converted into a binary annotation file and displayed on the tree using iTOL. The tree was manually trimmed to remove phylogenetically identical Acyl-CoA mutase leaves.

6.7.2 Bioinformatic Analysis of the Lydiamycin BGC

AntiSMASH [68] (https://antismash.secondarymetabolites.org/) was used for general analysis of the lydiamycin BGC. NRPSsp [313] (http://www.nrpssp.com/), PKS/NRPS Analysis [312] (http://nrps.igs.umaryland.edu/) and NRPSpredictor2 [88] were used for predicting the adenylation specificity of the *lyd* NRPS genes. Clinker [357] (https://github.com/gamcil/clinker) was used to compare BGCs. Sequence data for the *S. venezuelae* (NZ_CP029193.1) and *S. aureoverticillatus* (OL452061.1) clusters was used.

6.7.3 Structural Comparison of LydA and E. coli PDF proteins

The LydA protein sequence was submitted to AlphaFold2 [345, 346] and the best structural model was downloaded. The crystal structure of actinonin bound to *E. coli* PDF (PDB: 1G2A) was superimposed against the LydA structural model using PyMOL (https://pymol.org). PDF protein sequences were aligned by ClustalW using MEGA-X [358] and the alignment visualised using ESPript 3.0 (https://espript.ibcp.fr/).

6.8 Comparative Metabolomics

6.8.1 Production Media Screening

The *R. fascians* WT and Δ*nrp* strains were analysed by comparative metabolomics. The following liquid media were used (as detailed in **Table 12**): Aurachin production medium (PM) [315], Rhodostreptomycin PM [283], Actinonin PM, Matlystatin PM, Bottromycin PM [207], MinA [316], SM7, SM12 and YEB.

10 mL of LB was inoculated with a single colony of the *R. fascians* strain and grown for 2 days at 30 ^{o}C. 10 mL of each of the above media (in duplicate) were aliquoted into sterile bunged 50 mL tubes and inoculated with the pre-culture (1% v/v). The inoculum for MinA samples was centrifuged and washed twice with MinA. Media only samples were also prepared in duplicate. All cultures were incubated at 30 ^{o}C with shaking at 250 rpm. 0.75 mL aliquots of each culture were sampled after 5 and 12 days of fermentation. These were extracted in 0.75 mL MeOH by shaking vigorously for 1 hour. These samples were centrifuged at 15,871 x *g* for 30 minutes and 0.6 mL of clarified supernatant aliquoted into 2 mL LC-MS vials.

6.8.2 LC-MS analysis of Extracts

Samples were subjected to LC-MS analysis using a Shimadzu Nexera X2 UHPLC coupled to a Shimadzu IT-ToF mass spectrometer. Samples (5 µL) were injected onto a Phenomenex Kinetex 2.6 µm C18 column (50 x 2.1 mm, 100 Å), eluting with a linear gradient of 5 to 100% acetonitrile in water + 0.1% formic acid over 6 minutes with a flow-rate of 0.6 mL/min at 40 °C. Positive mode mass spectrometry data was collected between m/z 300 and 2000 with an ion accumulation time of 10 ms featuring an automatic sensitivity control of 70% of the base peak. The curved desolvation line temperature was 300 °C and the heat block temperature was 250 °C. MS/MS data was collected in a data-dependent manner using collision-induced dissociation energy of 50% and a precursor ion width of 3 Da. The instrument was calibrated using sodium trifluoroacetate cluster ions prior to every run.

Data was analysed using LabSolutions software (Shimadzu). The base peak chromatograms of the *R. fascians* WT and Δ*nrp* extracts from each media were overlayed and manually compared. Compounds present in the WT extract and absent in the Δ*nrp* and media only extracts were candidate *nrp* BGC products.

6.8.3 Mass Spectral Networking

A more advanced comparative metabolomics experiment was performed to compare the metabolites produced by *R. fascians* WT and Δ*nrp* strains in both SM12 and YEB media.

10 mL of LB was inoculated with a single colony of the *R. fascians* strain and grown for 2 days at 30 ^{o}C. 10 mL of each of the above media (in triplicate) were aliquoted into sterile bunged 50 mL tubes and inoculated with the pre-culture (1% v/v). Media only samples were also prepared in triplicate. All cultures were incubated at 30 ^{o}C with shaking at 250 rpm. 0.75 mL aliquots of each culture were sampled after 5 days of fermentation. These were extracted in 0.75 mL MeOH by shaking vigorously for 1 hour. These samples were centrifuged at 15,871 x *g* for 30 minutes and 0.6 mL of clarified supernatant aliquoted into 2 mL LC-MS vials.

Samples were subjected to LC-MS-MS analysis using a Waters Acquity UHPLC coupled to a Q Exactive Orbitrap Mass Spectrometer (Thermo). Samples (5 µL) were injected onto a Phenomenex Kinetex 2.6 µm C18 column (50 x 2.1 mm, 100 Å), eluting with a linear gradient of 5 to 95% acetonitrile in water + 0.1% formic acid over 6 minutes with a flow-rate of 0.6 mL/min at 40 °C. Positive mode mass spectrometry data was collected between *m/z* 200 and 2000. MS/MS data was collected in a data-dependent manner.

MS-MS data was converted to mzML format. A molecular network was created using the online workflow (https://ccms-ucsd.github.io/GNPSDocumentation/) on the GNPS website (http://gnps.ucsd.edu). The data was filtered by removing all MS/MS fragment ions within +/- 17 Da of the precursor m/z. MS/MS spectra were window filtered by choosing only the top 6 fragment ions in the +/- 50Da window throughout the spectrum. The precursor ion mass tolerance was set to 0.1 Da and a MS/MS fragment ion tolerance of 0.1 Da. A network was then created where edges were filtered to have a cosine score above 0.7 and more than 6 matched peaks. Further, edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. Finally, the maximum size of a molecular family was set to 100, and the lowest scoring edges were removed from molecular families until the molecular family size was below this threshold. The spectra in the network were then searched against GNPS' spectral libraries. The library spectra were filtered in the same manner as the input data. All matches kept between network spectra and library spectra were required to have a score above 0.7 and at least 6 matched peaks. The network was visualised using Cytoscape (version 3.8.2) [353].

6.9 Lydiamycin A Purification

6.9.1 Identification of Lydiamycin A by LC-MS

Samples were subjected to LC-MS analysis using a Shimadzu Nexera X2 UHPLC coupled to a Shimadzu IT-ToF mass spectrometer. Samples (5 µL) were injected onto a Phenomenex Kinetex 2.6 µm C18 column (50 x 2.1 mm, 100 Å), eluting with a linear gradient of 5 to 100% acetonitrile in water + 0.1% formic acid over 6 minutes with a flow-rate of 0.6 mL/min at 40 °C. Positive mode mass spectrometry data was collected between m/z 300 and 2000 with an ion accumulation time of 10 ms featuring an automatic sensitivity control of 70% of the base peak. The curved desolvation line temperature was 300 °C and the heat block temperature was 250 °C. MS/MS data was collected in a data-dependent manner using collision-induced dissociation energy of 50% and a precursor ion width of 3 Da. The instrument was calibrated using sodium trifluoroacetate cluster ions prior to every run.

6.9.2 Preliminary Purification Tests

6.9.2.1 Testing Lydiamycin A extraction techniques

10 mL of LB was inoculated with a single colony of the *R. fascians* WT strain and grown for 2 days at 30 °C. 10 mL of SM12 was aliquoted into sterile bunged 50 mL tubes and inoculated with the pre-culture (1% v/v). The culture was fermented at 30 °C for 12 days. The culture was centrifuged at 6,500 x *g* for 20 minutes using a 6x50mL 7588 rotor (Heraeus) in a Biofuge Primo centrifuge (Sorvall).

The supernatant was removed and extracted with 5 mL ethyl acetate with vigorous shaking for 20 minutes. The organic extract was removed and washed with 5 mL water using a separatory funnel. A sample of the supernatant-derived organic and aqueous fractions was removed for LC-MS analysis, as described in **Section 6.9.1**. The cell pellets were extracted in 2.5 mL 70% ACN with shaking for 20 minutes and centrifuged (as above). The organic extract was sampled and analysed by LC-MS, as described in **Section 6.9.1**.

6.9.2.2 Testing Lydiamycin A production in larger culture volumes

10 mL of LB was inoculated with a single colony of the *R. fascians* WT strain and grown for 2 days at 30 ^{o}C. This pre-culture was used to inoculate SM12 (at 1% v/v) at varied media and flask volumes (25 mL in 250 mL flask, 50 mL in 250 mL flask, 200 mL in 2 L flask, 400 mL in 2 L flask). All cultures were grown at 30 ^{o}C for 5 days, with shaking at 250 rpm. A 0.75 mL aliquot of the culture was extracted in an equal volume of MeOH for 30 minutes, with vigorous shaking. The extract was centrifuged at 15,871 x *g* for 30 minutes and 0.6 mL of the supernatant was transferred to a 2 mL LC-MS tube. These were analysed by LC-MS, as described in **Section 6.9.1**. Peak area was calculated using default integration parameters, with a peak width of 3 seconds.

6.9.3 Large-Scale Lydiamycin A Purification

6.9.3.1 *R. fascians* fermentation and extraction

10 mL of LB was inoculated with a single colony of the *R. fascians* WT strain and grown for 2 days at 30 ^{o}C. 0.5 mL of this culture was used to inoculate 50 mL of LB (in 250 mL flask), which was cultured at 30 ^{o}C for 1 day with 250 rpm shaking. This pre-culture was used to inoculate 2x 1.0 L of SM12 medium (in 2000 mL flasks) at 1% v/v. These cultures were fermented at 30 ^{o}C for five days with shaking at 250 rpm. The cultures were split equally between two 1000 mL centrifuge bottles and centrifuged at 12,000 x *g* for 30 minutes using a Fiberlite F9-6x1000 rotor (Thermo) in a Sorvall LYNX 6000 centrifuge (Thermo).

The supernatant was equally split between two 2 L separatory funnels and a liquid/liquid extraction was performed with two volumes of ethyl acetate. The mixture was shaken often over the course of an hour. The organic fraction was combined, and the volume was reduced to ~0.5 L by rotary evaporation. The organic fraction was washed with equal volumes of Milli-Q water, three times. The organic fraction was dried over $MgSO_4$ and dried by rotary evaporation.

6.9.3.2 Flash Chromatography

The crude organic fraction was separated by flash chromatography using a Biotage Isolera system equipped with a 30 g Sfär C18 cartridge (Biotage). The solvent system used was ACN (B) against H_2O (A) and the flowrate was 25 mL/min. The cartridge was equilibrated with 3 CVs of 5% B. The crude organic fraction was solubilised in 3 mL MeOH and loaded onto the cartridge. Separation proceeded with a 5% B hold for 2CVs, a linear gradient from 5-100% B over 12 CVs and a 100% B isocratic hold for 1.5 CVs. Elution was monitored at 210 nm and 237 nm and 22 mL fractions were collected. Each fraction was sampled and analysed by LC-MS (as described in **Section 6.9.1**). Fractions containing lydiamycin A were combined and dried by rotary evaporation and lyophilisation.

6.9.3.3 Preparative-Scale HPLC

Flash chromatography fractions were further separated by reverse phase HPLC using an UltiMate 3000 HPLC instrument (Thermo) equipped with a preparative-scale Gemini-NX C18 (150 mm x 21.1mm) C18 column (Phenomenex). The solvent system used was ACN (+0.1% FA) (B) against H_2O (A) and the flowrate was 20 mL/min. The column was equilibrated with 5% B. The sample was solubilised in 3 mL MeOH to a concentration of 9 mg/mL and filtered using a 0.45 µm PTFE filter (Whatman). The sample was injected (9 mg per injection) and separated with an isocratic hold of 5% B for 1 CV, a linear gradient of 5-45% B for 2 CVs, a linear gradient of 45-70% B for 6.5 CVsThe UV absorption at 237 nm was recorded. Peaks were manually collected and stored on ice. The fractions were sampled for LC-MS analysis (as described in **Section 6.9.1**) and then lyophilised to dryness.

6.10 Lydiamycin A Structural Elucidation

6.10.1 HR-MS/MS

HR-MS/MS analysis was performed by Gerhard Saalbach and Carlo de Oliveira Martins (John Innes Centre). High-resolution mass spectra were acquired on a Synapt G2-Si mass spectrometer equipped with an Acquity UPLC (Waters). Aliquots of the samples were injected onto an Acquity UPLC® BEH C18 column, 1.7 µm, 1x100 mm (Waters) and eluted with a gradient of acetonitrile/0.1% formic acid (B) in water/0.1% formic acid (A) with a flow rate of 0.08 mL/min at 45 °C. The concentration of B was kept at 1% for 1 min followed by a gradient up to 40% B in 9 min, ramping to 99%B in 1 min, kept at 99%B for 2 min and re-equilibrated at 1%B for 4 min. MS data were collected in positive mode with the following parameters: resolution mode, positive ion mode, scan time 0.5 s, mass range m/z 50-1200 calibrated with sodium formate, capillary voltage = 2.5 kV; cone voltage = 40 V; source temperature = 125 °C; desolvation temperature = 300 °C. Leu-enkephalin peptide

was used to generate a lock-mass calibration with 556.2766, measured every 30 s during the run. For MS2 fragmentation, a data directed analysis (DDA) method was used with the following parameters: precursor selected from the 4 most intense ions; MS2 threshold: 5,000; scan time 0.5 s; no dynamic exclusion. In positive mode, collision energy (CE) was ramped between 10-30 at low mass (m/z 50) and 15-60 at high mass (m/z 1200). Lydiamycin A HR-MS analysis is presented in **Appendix Figure 112**.

6.10.2 NMR analysis

Pure lydiamycin A (7.0 mg) was dissolved in 200 µL $CDCl_3$ and transferred into a 3 mm NMR tube. This sample was subjected to a series of 1D and 2D NMR experiments on a Bruker Avance Neo 600 MHz spectrometer equipped with a TCI cryoprobe at 298 K. The NMR experiments carried out were Proton (16 scans), Carbon (900 scans), HMBC (8 scans), COSY (1 scan), HSQC (6 scans), 1,1-ADEQUATE (197 scans). Spectra were analysed using Bruker TopSpin 4.0. The lydiamycin A NMR assignment is presented in **Table 7**.

6.11 Lydiamycin A Bioactivity Assays

6.11.1 Spot-on-Lawn assays

10 mL of LB or MOADC was inoculated with a single colony of *E. coli* NR698 or *M. smegmatis* mc^2155, respectively, and grown overnight at 37 ^oC with shaking at 250 rpm. This pre-culture was used to inoculate 50 mL of the respective media (in 250 mL flasks), which were grown at 37 ^oC with 250 rpm shaking until an OD_{600} of 0.40 was reached. SNA and solid MOADC media were melted and cooled to 42 ^oC in a water bath. The *E. coli* culture was added to SNA and *M. smegmatis* added to MOADC at a concentration of O.D. 0.02 and were poured into 120 x 120 mm plates (50 mL each). A concentration series was prepared by serial dilution of pure lydiamycin in methanol from 600 µg/mL to 1 µg/mL. A positive control of 50 µg/mL kanamycin in water was prepared. 5 µL spots were applied to the plates and allowed to dry. The plates were grown at 37 ^oC for 1-3 days and imaged.

6.11.2 *M. smegmatis* Growth Curves

10 mL of MOADC with appropriate antibiotics was inoculated with a single colony of *M. smegmatis* mc^2155 and grown overnight at 37 ^oC with shaking at 250 rpm. The *M. smegmatis* cultures were inoculated into 200 µL of MOADC medium in 96-well microplates to a final concentration of $O.D_{600}$ 0.05, in triplicate. All cultures were supplemented with 0.25% TWEEN-80, 0.2% acetamide and the appropriate antibiotics. Each strain was grown either in the presence of methanol (5% v/v) or 50 µM lydiamycin A (in 5% v/v methanol), unless otherwise stated. Media only samples were also grown in triplicate. Fermentation proceeded at 37 ^oC with 500 rpm shaking. Growth was measured by $O.D_{600}$ detection every

30 minutes for 72 hours using a SPECTROstar Nano UV/Vis microplate reader (BMG Labtech). The data was processed in Excel. The figure was generated using a R script written by Dr. Alaster Moffat (Truman lab, John Innes Centre).

6.11.3 *In vitro* Deformylation Assay

This assay was performed by Lena Bögeholz (Rodnina laboratory, Max Planck Institute for Multidisciplinary Sciences, Germany). PDF from *E. coli* was purified as described [340, 341]. In short, His_6-tagged PDF was expressed in BL21(DE3)pLysS cells (NEB) and purified by Cobalt-Talon affinity chromatography (Clontech) followed by ion exchange chromatography using a HiTrapQ column (Cytiva). PDF was stored in buffer A (25 mM HEPES, pH 7.5, 70 mM NH_4Cl, 30 mM KCl, 7 mM $MgCl_2$, 0.2 mM $CoCl_2$, 1 mM TCEP, and 10% (v/v) glycerol) at -80 °C.

To determine the inhibitory effect of lydiamycin A on PDF, a colorimetric coupled enzyme assay utilizing a model peptide formyl-methionyl-leucyl-p-nitroaniline (fML-pNA, Bachem) as the substrate was performed. PDF (10 nM) was incubated with fML-pNA (20 µM) and *Aeromonas* aminopeptidase (0.8 U/mL) in the presence of lydiamycin A (0-50 µM) in buffer A with 1% DMSO [342, 343]. Formation of *p*-nitroaniline upon deformylation was monitored at 405 nm (ε_{405} = 10,600 $M^{-1}cm^{-1}$) every 5 s at room temperature. The initial velocity was determined by linear regression and the half maximal inhibitory concentration (IC50) was estimated by hyperbolic fitting. The activity of PDF in the presence of 1% DMSO was verified by Michaelis Menten kinetics.

6.11.4 Plant Assays

6.11.4.1 Root growth assay

120 x 120 mm plates were prepared, each with 50 mL of ½ MS + 3% sucrose + 0.8% agar media. *N. benthamiana* seeds were surface sterilised in 1.5% bleach solution for 30 minutes with strong vortexing. Under aseptic conditions, the seeds were washed in sterile Milli-Q water four times and aliquoted onto sterile filter paper. Twenty seeds were placed on each plate (in a line 2 cm from one edge). Plates were closed with micropore tape and stored in the dark at 4 ^oC for 3 days to align germination. The plates were placed vertically and grown under 16/8 hour light at 28 ^oC for 5 days.

10 mL of LB was inoculated with a single colony of *R. fascians* and grown overnight at 30 ^oC with shaking at 250 rpm. 0.5 mL of culture was centrifuged (4,000 x *g* for 10 minutes), the supernatant discarded by decanting and the pellet resuspended in 1 mL sterile Milli-Q water. This step was repeated once more, and the pellet resuspended in sterile 100 mM

$MgCl_2$ solution to an $O.D_{600}$ value of 0.5. The length of the plant root was marked on the plate. Each seedling was inoculated with 4 µL of culture. Each *R. fascians* strain and a mock infection using 4 µL of 100 mM $MgCl_2$ was tested using 60 seedlings. The plate was closed with micropore tape and returned to the same growth conditions for 7 days. The length of the plant root was marked and the distance between the 0 and 7 d.p.i marks is the root length growth. Data was processed using GraphPad.

6.11.4.1.1 LC-MS analysis of seedlings

Following the root length assay described above, six groups of five seedlings from each condition were transferred separately to 2 mL centrifuge tubes containing 1 mL MeOH and shaken vigorously for 30 minutes. The tubes were centrifuged at 15,871 x *g* for 30 minutes and 0.6 mL of the supernatant transferred to a 2 mL LC-MS tube. These extracts were analysed by LC-MS for the presence of lydiamycin A (as described in **Section 6.9.1**).

6.11.4.2 Leafy gall assay

GA-7 Magenta Vessels (Sigma) were prepared, each with 150 mL of ½ MS + 3% sucrose + 0.8% agar media. *N. benthamiana* seeds were surface sterilised in 1.5% bleach solution for 30 minutes with strong vortexing. Under aseptic conditions, the seeds were washed in sterile Milli-Q water four times and aliquoted onto sterile filter paper. One seed was placed into each Magenta vessel, and these were closed and grown with 16/8 hour light at 28 ^oC for 6 weeks.

10 mL of LB was inoculated with a single colony of *R. fascians* and grown overnight at 30 ^oC with shaking at 250 rpm. 0.5 mL of culture was centrifuged (4,000 x *g* for 10 minutes), the supernatant discarded by decanting and the pellet resuspended in 1 mL sterile Milli-Q water. This step was repeated once more, and the pellet resuspended in sterile 100 mM $MgCl_2$ solution to an $O.D_{600}$ value of 0.5. The meristem of each plant was wounded with sterile forceps and infected with 5 µL of *R. fascians* WT culture, *R. fascians* Δ*nrp* culture or sterile 100 mM $MgCl_2$ solution (n = 9 plants each). The plants were returned to the same growth conditions as above for four weeks and imaged. The galls were also removed, and the masses of gall and plant measured.

6.11.4.3 Excised leaf assay

GA-7 Magenta Vessels (Sigma) were prepared, each with 150 mL of ½ MS + 3% sucrose + 0.8% agar media. *N. tabacum* seeds were surface sterilised in 1.5% bleach solution for 30 minutes with strong vortexing. Under aseptic conditions, the seeds were washed in sterile Milli-Q water four times and aliquoted onto sterile filter paper. One seed was placed into

each Magenta vessel, and these were closed and grown with 16/8 hour light at 28 °C for 4 weeks. Leaves were aseptically excised and placed onto 120 x 120 mm ½ MS + 3% sucrose + 0.8% agar plates.

10 mL of LB was inoculated with a single colony of *R. fascians* and grown overnight at 30 °C with shaking at 250 rpm. Leaves were infected with 5 µL of *R. fascians* WT culture, *R. fascians* Δ*nrp* culture or sterile 100 mM $MgCl_2$ solution (n = 15 leaves each). The plates were returned to the same growth conditions as above for three weeks and imaged.

6.11.4.4 Competition assay

120 x 120 mm plates were prepared, each with 50 mL of ½ MS + 3% sucrose + 0.8% agar media. *N. benthamiana* seeds were surface sterilised in 1.5% bleach solution for 30 minutes with strong vortexing. Under aseptic conditions, the seeds were washed in sterile Milli-Q water four times and aliquoted onto sterile filter paper. 36 seeds were placed onto each plate. Plates were closed with micropore tape and stored in the dark at 4 °C for 3 days to align germination. The plates were grown under 16/8 hour light at 28 °C for 7 days.

10 mL of LB was inoculated with a single colony of *R. fascians* and grown overnight at 30 °C with shaking at 250 rpm. 0.5 mL of culture was centrifuged (4,000 x *g* for 10 minutes), the supernatant discarded by decanting and the pellet resuspended in 1 mL sterile Milli-Q water. This step was repeated once more, and the pellet resuspended in sterile 100 mM $MgCl_2$ solution to an $O.D_{600}$ value of 0.5. 1:1 mixes of *R. fascians* D188-5/WT and *R. fascians* D188-5/Δ*nrp* were prepared. The second competition assay used biological replicates of *R. fascians* WT and Δ*nrp* strains (n = 7 each) from separate colonies, against the same D188-5 biological isolate.

A serial dilution of these inocula was prepared in LB (from 10^{-1} to 10^{-8} dilution) and 5 µL spots were plated onto both LB and LB Streptomycin plates. These plates were incubated at 30 °C for 1-3 days until individual colonies could be counted. The CFU/mL was calculated. The CFU/mL of the test strain was equal to CFU/mL (LB – LB streptomycin).

The same inoculum was used to infect the seedlings. 4 µL of the inoculum (approximately $2x10^5$ cells) was applied to each seedling. In the first experiment, 8 groups of 7 seedlings were infected per condition. In the second experiment, 7 groups of 8 seedlings per timepoint were infected per condition. The plates were grown under 16/8 hour light at 28 °C. In the first experiment, bacteria were isolated at 7 d.p.i., whereas in the second experiment, bacteria were isolated at 3, 7 and 14 d.p.i. For both experiments, the seedlings were aseptically transferred to a 2 mL microcentrifuge tube containing 1 mL sterile 100 mM

$MgCl_2$ and vigorously mixed for 1 hour. These were serially diluted in LB (from 10^{-1} to 10^{-8} dilution) and 5 µL spots were plated onto both LB and LB Streptomycin plates. These plates were incubated at 30 ^oC for 1-3 days until individual colonies could be counted. The CFU/mL was calculated.

The data and statistics were processed in GraphPad by Dr. Catriona Thompson (John Innes Centre). In the first experiment, relative fitness was calculated: ln(test CFU 7dpi/ test CFU 0dpi)/ln(D188-5 CFU 7dpi/D188-5 CFU 0dpi).

6.12 Viscosin I Purification and Structural Elucidation

6.12.1Viscosin I Purification

6.12.1.1 *Pseudomonas* Ps682 Fermentation and Extraction

10 mL of L was inoculated with a single colony of the *Pseudomonas* Ps682 and grown overnight at 30 ^oC. Fifteen 140 mm diameter KB+Mg agar plates were each inoculated with 0.6 mL of the Ps682 pre-culture and grown for 24 hours at 30 ^oC. The agar plates were cut, transferred to a 2 L Duran bottle and extracted in 500 mL ethyl acetate for 2 hours with frequent agitation. The organic extract was filtered through muslin cloth and transferred to a 2 L separatory funnel. The organic extract was washed with 3x 200 mL of Milli-Q H_2O and dried over $MgSO_4$ followed by rotary evaporation.

6.12.1.2 Flash chromatography

The crude organic fraction was separated by flash chromatography using a Biotage Isolera system equipped with a 12 g SNAP Ultra C18 cartridge (Biotage). The solvent system used was MeOH (B) against H_2O (A) and the flowrate was 12 mL/min. The cartridge was equilibrated with 3 CVs of 70% B. The crude organic fraction was solubilised in 1 mL MeOH and loaded onto the cartridge. Separation proceeded with a linear gradient of 70-100% B over 7 CVs followed by an isocratic hold of 100% B for 3 CVs. Elution was monitored at 210 nm and 260 nm and 22 mL fractions were collected. Each fraction was sampled and analysed by LC-MS (as described in **Section 6.9.1**). Fractions containing viscosin I were combined and dried using a Genevac EZ-2 evaporator on the 'HPLC' setting.

6.12.1.3 Semi-preparative HPLC

Flash chromatography fractions were further separated by reverse phase HPLC using an UltiMate 3000 HPLC instrument (Thermo) equipped with a semi-preparative scale Luna C18 100mm x 10 mm column (Phenomenex). The solvent system used was MeOH (B) against H_2O (A) and the flowrate was 5 mL/min. The column was equilibrated with 50% B. The sample was solubilised in MeOH to a concentration of 4 mg/mL and filtered using a 0.45

µm PTFE filter (Whatman). The sample was injected and separated with a linear gradient of 50-100% B over 10 CVs and an isocratic hold at 100% B for 3 CVs. The UV absorption at 210 nm was recorded. Peaks were manually collected and stored on ice. The fractions were sampled for LC-MS analysis (as described in **Section 6.9.1**) and dried using a Genevac EZ-2 evaporator on the 'HPLC' setting.

6.12.1.4 Known viscosin-like CLP extraction and MS comparison

CLP production of *Pseudomonas* SBW25, *Pseudomonas* Ps682 and *Pseudomonas* sp. LMG 2338 was compared by LC-MS. 10 mL of L was inoculated with a single colony of each of the above strains and grown overnight at 30 ^{o}C. 0.1 mL of each pre-culture was used to inoculate 40 mm KB+Mg agar plates, which were incubated at 30 ^{o}C for 24 hours. Half of each plate was decanted into a sterile 50 mL tube and extracted with 10 mL 50% ethanol for 3 hours with vigorous shaking. A 2 mL aliquot was centrifuged (15,871 x *g*, 5 minutes) and 0.5 mL of the supernatant was removed for LC-MS analysis. The sample was analysed by LC-MS as described in **Section 6.9.1**, but with a linear gradient of 5-100% over 15 minutes.

6.12.2 Viscosin I Structural Elucidation

6.12.2.1 HR-MS/MS Analysis

Pure viscosin I was analysed by HR-MS/MS by Gerhard Saalbach and Carlo de Oliveira Martins (John Innes Centre). High-resolution mass spectra were acquired on a Synapt G2-Si mass spectrometer equipped with an Acquity UPLC (Waters). Aliquots of the samples were injected onto an Acquity UPLC® BEH C18 column, 1.7 µm, 1x100 mm (Waters) and eluted with a gradient of acetonitrile/0.1% formic acid (B) in water/0.1% formic acid (A) with a flow rate of 0.08 mL/min at 45 °C. The concentration of B was kept at 1% for 1 min followed by a gradient up to 40% B in 9 min, ramping to 99% B in 1 min, kept at 99% B for 2 min and re-equilibrated at 1% B for 4 min. MS data were collected in positive mode with the following parameters: resolution mode, positive ion mode, scan time 0.5 s, mass range m/z 50-1200 calibrated with sodium formate, capillary voltage = 2.5 kV; cone voltage = 40 V; source temperature = 125 °C; desolvation temperature = 300 °C. Leu-enkephalin peptide was used to generate a lock-mass calibration with 556.2766, measured every 30 s during the run. For MS2 fragmentation, a data directed analysis (DDA) method was used with the following parameters: precursor selected from the 4 most intense ions; MS2 threshold: 5,000; scan time 0.5 s; no dynamic exclusion. In positive mode, collision energy (CE) was ramped between 10-30 at low mass (m/z 50) and 15-60 at high mass (m/z 1200). The HR-MS analysis of Viscosin I is presented in **Appendix Figure 121**.

6.12.2.2 NMR Analysis

Pure viscosin I (1.0 mg) was dissolved in 1 mL MeOH-d4 in order to exchange exchangeable protons with deuterium. This was dried under a stream of nitrogen at room temperature. The sample was dissolved in N,N-dimethylformamide-d7 (DMF-d7) and NMR spectra were acquired on a Bruker Avance Neo 600 MHz spectrometer equipped with a TCI cryoprobe. The experiments were carried out at 298 K with the residual DMF solvent used as an internal standard (δ_H/δ_C 2.75/29.76). The residual solvent signal from H_2O was suppressed through a presaturation sequence in 1D ^1H. Resonances were assigned through 1D ^1H (32 scans), ^{13}C (30,000 scans) and DEPT135 (25,600 scans) experiments, and 2D COSY (10 scans), HSQCed (36 scans), HMBC (40 scans), TOCSY (50 scans), and HSQC-TOCSY (25 scans) experiments. Spectra were analyzed using Bruker TopSpin 3.5 and Mestrelab Research Mnova 14.0 software. Dr. Sergey Nepogodiev (John Innes Centre) carried out the NMR experiments and performed all NMR data analysis.

6.13 Viscosin I Bioactivity Assays

6.13.1 Disk Diffusion Assays

S. scabies 87-22 spore solution was diluted 1:100 in sterile Milli-Q water. 60 uL aliquots were spread onto 100 mm square SFM or IPM agar plates using sterile cotton buds and allowed to dry for 30 mins. Pure viscosin I was diluted in MeOH to produce a range of concentrations from 10-100 µg/mL. 100 µL of each of these solutions was applied to a sterile 6 mm filter disk, in 5x 20 µL applications, and allowed to dry between each application. A methanol only control disk was also produced. The disks were applied to the agar plate, which was incubated at 30^{o}C for 48 hours.

Chapter 7: References

1. Vining, L.C., *FUNCTIONS OF SECONDARY METABOLITES.* Annual Review of Microbiology, 1990. **44**(1): p. 395-427.
2. Maplestone, R.A., M.J. Stone, and D.H. Williams, *The evolutionary role of secondary metabolites—a review.* Gene, 1992. **115**(1-2): p. 151-157.
3. Bérdy, J., *Bioactive Microbial Metabolites.* The Journal of Antibiotics, 2005. **58**(1): p. 1-26.
4. Kaspar, F., P. Neubauer, and M. Gimpel, *Bioactive Secondary Metabolites from Bacillus subtilis: A Comprehensive Review.* Journal of Natural Products, 2019. **82**(7): p. 2038-2053.
5. Challis, G.L. and J. Ravel, *Coelichelin, a new peptide siderophore encoded by the Streptomyces coelicolor genome: structure prediction from the sequence of its non-ribosomal peptide synthetase.* FEMS Microbiol Lett, 2000. **187**(2): p. 111-4.
6. Ćirić, A., J. Petrović, J. Glamočlija, M. Smiljković, M. Nikolić, D. Stojković, and M. Soković, *Natural products as biofilm formation antagonists and regulators of quorum sensing functions: A comprehensive review update and future trends.* South African Journal of Botany, 2019. **120**: p. 65-80.
7. Newman, D.J., G.M. Cragg, and K.M. Snader, *The influence of natural products upon drug discovery.* Nat Prod Rep, 2000. **17**(3): p. 215-34.
8. Demain, A.L., *Importance of microbial natural products and the need to revitalize their discovery.* J Ind Microbiol Biotechnol, 2014. **41**(2): p. 185-201.
9. Newman, D.J. and G.M. Cragg, *Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019.* Journal of natural products, 2020. **83**(3): p. 770-803.
10. Lederberg, J., *Infectious History.* Science, 2000. **288**(5464): p. 287-293.
11. Starzl, T.E., *History of clinical transplantation.* World J Surg, 2000. **24**(7): p. 759-82.
12. Hranueli, D., J. Cullum, B. Basrak, P. Goldstein, and P.F. Long, *Plasticity of the streptomyces genome-evolution and engineering of new antibiotics.* Curr Med Chem, 2005. **12**(14): p. 1697-704.
13. Resistance, R.o.A., R.o.A.R. ., J. O'Neill, and Grande-Bretagne, *Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations : December 2014*. 2014: Review on Antimicrobial Resistance.
14. Christensen, S.B., *Drugs That Changed Society: History and Current Status of the Early Antibiotics: Salvarsan, Sulfonamides, and β-Lactams.* Molecules, 2021. **26**(19).
15. Haas, L.F., *Papyrus of Ebers and Smith.* J Neurol Neurosurg Psychiatry, 1999. **67**(5): p. 578.
16. BRUNEL, J., *Antibiosis from Pasteur to Fleming*.* Journal of the History of Medicine and Allied Sciences, 1951. **VI**(Summer): p. 287-301.
17. Hörlein, H., *The Chemotherapy of Infectious Diseases caused by Protozoa and Bacteria: (Section of Tropical Diseases and Parasitology).* Proc R Soc Med, 1936. **29**(4): p. 313-24.
18. Fleming, A., *On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of B. influenzæ.* British journal of experimental pathology, 1929. **10**(3): p. 226-236.

19. Chain, E., H.W. Florey, A.D. Gardner, N.G. Heatley, M.A. Jennings, J. Orr-Ewing, and A.G. Sanders, *PENICILLIN AS A CHEMOTHERAPEUTIC AGENT.* The Lancet, 1940. **236**(6104): p. 226-228.
20. Hodgkin, D.C., *The X-ray analysis of the structure of penicillin.* Adv Sci, 1949. **6**(22): p. 85-9.
21. Kardos, N. and A.L. Demain, *Penicillin: the medicine with the greatest impact on therapeutic outcomes.* Applied Microbiology and Biotechnology, 2011. **92**(4): p. 677-687.
22. Schatz, A., E. Bugle, and S.A. Waksman, *Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria.∗†.* Proceedings of the Society for Experimental Biology and Medicine, 1944. **55**(1): p. 66-69.
23. Dickinson, L., *Effect of streptomycin on experimental tuberculosis in guinea-pigs.* British journal of pharmacology and chemotherapy, 1947. **2**(1): p. 23-26.
24. Ehrlich, J., Q.R. Bartz, R.M. Smith, D.A. Joslyn, and P.R. Burkholder, *Chloromycetin, a New Antibiotic From a Soil Actinomycete.* Science, 1947. **106**(2757): p. 417.
25. Smadel, J.E. and E.B. Jackson, *Chloromycetin, an Antibiotic With Chemotherapeutic Activity in Experimental Rickettsial and Viral Infections.* Science, 1947. **106**(2757): p. 418-9.
26. Bartz, Q.R., *Isolation and characterization of chloromycetin.* J Biol Chem, 1948. **172**(2): p. 445-50.
27. Geraci, J.E., F.R. Heilman, D.R. Nichols, W.E. Wellman, G.T. Ross, E.W. With The Technical Assistance Of, and R. Dorothy, *Some Laboratory and Clinical Experiences with a New Antibiotic, Vancomycin.* Proc. Staff, Meetings Mayo Clinic, 1956. **31**(21): p. 564-82.
28. Geeaci, J.E., F.R. Heilman, D.R. Nichols, and W.E. Wellman, *Antibiotic Therapy of Bacterial Endocarditis. VII. Vancomycin for Acute Micrococcal Endocarditis. Preliminary Report.* Proceedings of Staff Meetings of the Mayo Clinic, 1958. **33**(7): p. 172-81.
29. Davies, J., *Where have All the Antibiotics Gone?* The Canadian journal of infectious diseases & medical microbiology = Journal canadien des maladies infectieuses et de la microbiologie medicale, 2006. **17**(5): p. 287-290.
30. Challis, G.L. and D.A. Hopwood, *Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by Streptomyces species.* Proceedings of the National Academy of Sciences, 2003. **100**(suppl 2): p. 14555-14561.
31. Baltz, R.H., *Renaissance in antibacterial discovery from actinomycetes.* Current Opinion in Pharmacology, 2008. **8**(5): p. 557-563.
32. Payne, D.J., M.N. Gwynn, D.J. Holmes, and D.L. Pompliano, *Drugs for bad bugs: confronting the challenges of antibacterial discovery.* Nature Reviews Drug Discovery, 2007. **6**(1): p. 29-40.
33. Brown, D.G. and H.J. Wobst, *A Decade of FDA-Approved Drugs (2010–2019): Trends and Future Directions.* Journal of Medicinal Chemistry, 2021. **64**(5): p. 2312-2338.
34. CDC, *Antibiotic resistance threats in the United States, 2019 https://stacks.cdc.gov/view/cdc/82532.* 2019.
35. Murray, C.J.L., K.S. Ikuta, F. Sharara, L. Swetschinski, G. Robles Aguilar, A. Gray, C. Han, C. Bisignano, P. Rao, E. Wool, S.C. Johnson, A.J. Browne, M.G. Chipeta, F. Fell,

S. Hackett, G. Haines-Woodhouse, B.H. Kashef Hamadani, E.A.P. Kumaran, B. McManigal, R. Agarwal, S. Akech, S. Albertson, J. Amuasi, J. Andrews, A. Aravkin, E. Ashley, F. Bailey, S. Baker, B. Basnyat, A. Bekker, R. Bender, A. Bethou, J. Bielicki, S. Boonkasidecha, J. Bukosia, C. Carvalheiro, C. Castañeda-Orjuela, V. Chansamouth, S. Chaurasia, S. Chiurchiù, F. Chowdhury, A.J. Cook, B. Cooper, T.R. Cressey, E. Criollo-Mora, M. Cunningham, S. Darboe, N.P.J. Day, M. De Luca, K. Dokova, A. Dramowski, S.J. Dunachie, T. Eckmanns, D. Eibach, A. Emami, N. Feasey, N. Fisher-Pearson, K. Forrest, D. Garrett, P. Gastmeier, A.Z. Giref, R.C. Greer, V. Gupta, S. Haller, A. Haselbeck, S.I. Hay, M. Holm, S. Hopkins, K.C. Iregbu, J. Jacobs, D. Jarovsky, F. Javanmardi, M. Khorana, N. Kissoon, E. Kobeissi, T. Kostyanev, F. Krapp, R. Krumkamp, A. Kumar, H.H. Kyu, C. Lim, D. Limmathurotsakul, M.J. Loftus, M. Lunn, J. Ma, N. Mturi, T. Munera-Huertas, P. Musicha, M.M. Mussi-Pinhata, T. Nakamura, R. Nanavati, S. Nangia, P. Newton, C. Ngoun, A. Novotney, D. Nwakanma, C.W. Obiero, A. Olivas-Martinez, P. Olliaro, E. Ooko, E. Ortiz-Brizuela, A.Y. Peleg, C. Perrone, N. Plakkal, A. Ponce-de-Leon, M. Raad, T. Ramdin, A. Riddell, T. Roberts, J.V. Robotham, A. Roca, K.E. Rudd, N. Russell, J. Schnall, J.A.G. Scott, M. Shivamallappa, J. Sifuentes-Osornio, N. Steenkeste, A.J. Stewardson, T. Stoeva, N. Tasak, A. Thaiprakong, G. Thwaites, C. Turner, P. Turner, H.R. van Doorn, S. Velaphi, A. Vongpradith, H. Vu, T. Walsh, S. Waner, T. Wangrangsimakul, T. Wozniak, P. Zheng, B. Sartorius, A.D. Lopez, A. Stergachis, C. Moore, C. Dolecek and M. Naghavi, *Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis.* The Lancet, 2022. **399**(10325): p. 629-655.

36. de Kraker, M.E.A., A.J. Stewardson, and S. Harbarth, *Will 10 Million People Die a Year due to Antimicrobial Resistance by 2050?* PLoS medicine, 2016. **13**(11): p. e1002184-e1002184.
37. Saga, T. and K. Yamaguchi, *History of antimicrobial agents and resistant bacteria.* 2009.
38. Eriksen, K.R., *Studies on induced resistance to penicillin in a pneumococcus type I.* Acta Pathol Microbiol Scand, 1945. **22**(4): p. 398-405.
39. Plough, H.H., *Penicillin resistance of Staphylococcus aureus and its clinical implications.* Am J Clin Pathol, 1945. **15**: p. 446-51.
40. Knox, R., *A new penicillin (BRL 1241) active against penicillin-resistant staphylococci.* British medical journal, 1960. **2**(5200): p. 690-693.
41. White, A. and D.T. Varga, *Antistaphylococcal activity of sodium methicillin: 2,6-dimethoxyphenyl penicillin; penicillin X-1497.* Arch Intern Med, 1961. **108**: p. 671-8.
42. Barber, M., *METHICILLIN-RESISTANT STAPHYLOCOCCI.* Journal of Clinical Pathology, 1961. **14**(4): p. 385-&.
43. Lee, A.S., H. de Lencastre, J. Garau, J. Kluytmans, S. Malhotra-Kumar, A. Peschel, and S. Harbarth, *Methicillin-resistant Staphylococcus aureus.* Nature Reviews Disease Primers, 2018. **4**(1): p. 18033.
44. Sorrell, T.C., D.R. Packham, S. Shanker, M. Foldes, and R. Munro, *Vancomycin Therapy for Methicillin-Resistant Staphylococcus aureus.* Annals of Internal Medicine, 1982. **97**(3): p. 344-350.
45. Chang, S., D.M. Sievert, J.C. Hageman, M.L. Boulton, F.C. Tenover, F.P. Downes, S. Shah, J.T. Rudrik, G.R. Pupp, W.J. Brown, D. Cardo, and S.K. Fridkin, *Infection with vancomycin-resistant Staphylococcus aureus containing the vanA resistance gene.* N Engl J Med, 2003. **348**(14): p. 1342-7.

46. Mangili, A., I. Bica, D.R. Snydman, and D.H. Hamer, *Daptomycin-Resistant, Methicillin-Resistant Staphylococcus aureus Bacteremia.* Clinical Infectious Diseases, 2005. **40**(7): p. 1058-1060.
47. Marty, F.M., W.W. Yeh, C.B. Wennersten, L. Venkataraman, E. Albano, E.P. Alyea, H.S. Gold, L.R. Baden, and S.K. Pillai, *Emergence of a Clinical Daptomycin-Resistant Staphylococcus aureus Isolate during Treatment of Methicillin-Resistant Staphylococcus aureus Bacteremia and Osteomyelitis.* Journal of Clinical Microbiology, 2006. **44**(2): p. 595.
48. Shallcross, L.J., S.J. Howard, T. Fowler, and S.C. Davies, *Tackling the threat of antimicrobial resistance: from policy to sustainable action.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2015. **370**(1670): p. 20140082.
49. Qin, Z., J.T. Munnoch, R. Devine, N.A. Holmes, R.F. Seipke, K.A. Wilkinson, B. Wilkinson, and M.I. Hutchings, *Formicamycins, antibacterial polyketides produced by Streptomyces formicae isolated from African Tetraponera plant-ants.* Chemical Science, 2017. **8**(4): p. 3218-3227.
50. Xu, D., L. Han, C. Li, Q. Cao, D. Zhu, N.H. Barrett, D. Harmody, J. Chen, H. Zhu, P.J. McCarthy, X. Sun, and G. Wang, *Bioprospecting Deep-Sea Actinobacteria for Novel Anti-infective Natural Products.* Front Microbiol, 2018. **9**: p. 787.
51. Goodfellow, M., I. Nouioui, R. Sanderson, F. Xie, and A.T. Bull, *Rare taxa and dark microbial matter: novel bioactive actinobacteria abound in Atacama Desert soils.* Antonie van Leeuwenhoek, 2018. **111**(8): p. 1315-1332.
52. Abdelkader, M.S.A., T. Philippon, J.A. Asenjo, A.T. Bull, M. Goodfellow, R. Ebel, M. Jaspars, and M.E. Rateb, *Asenjonamides A–C, antibacterial metabolites isolated from Streptomyces asenjonii strain KNN 42.f from an extreme-hyper arid Atacama Desert soil.* The Journal of Antibiotics, 2018. **71**(4): p. 425-431.
53. Liu, J.-T., X.-L. Lu, X.-Y. Liu, Y. Gao, B. Hu, B.-H. Jiao, and H. Zheng, *Bioactive natural products from the antarctic and arctic organisms.* Mini Reviews in Medicinal Chemistry, 2013. **13**(4): p. 617-626.
54. Adam, D., M. Maciejewska, A. Naômé, L. Martinet, W. Coppieters, L. Karim, D. Baurain, and S. Rigali, *Isolation, Characterization, and Antibacterial Activity of Hard-to-Culture Actinobacteria from Cave Moonmilk Deposits.* Antibiotics, 2018. **7**(2): p. 28.
55. Kaeberlein, T., K. Lewis, and S.S. Epstein, *Isolating "Uncultivable" Microorganisms in Pure Culture in a Simulated Natural Environment.* Science, 2002. **296**(5570): p. 1127-1129.
56. Stewart, E.J., *Growing unculturable bacteria.* J Bacteriol, 2012. **194**(16): p. 4151-60.
57. Vartoukian, S.R., R.M. Palmer, and W.G. Wade, *Strategies for culture of 'unculturable'bacteria.* FEMS microbiology letters, 2010. **309**(1): p. 1-7.
58. Stevenson, B.S., S.A. Eichorst, J.T. Wertz, T.M. Schmidt, and J.A. Breznak, *New Strategies for Cultivation and Detection of Previously Uncultured Microbes.* Applied and Environmental Microbiology, 2004. **70**(8): p. 4748-4755.
59. Ling, L.L., T. Schneider, A.J. Peoples, A.L. Spoering, I. Engels, B.P. Conlon, A. Mueller, T.F. Schäberle, D.E. Hughes, S. Epstein, M. Jones, L. Lazarides, V.A. Steadman, D.R. Cohen, C.R. Felix, K.A. Fetterman, W.P. Millett, A.G. Nitti, A.M. Zullo, C. Chen, and K. Lewis, *A new antibiotic kills pathogens without detectable resistance.* Nature, 2015. **517**(7535): p. 455-459.

60. Bentley, S.D., K.F. Chater, A.M. Cerdeño-Tárraga, G.L. Challis, N.R. Thomson, K.D. James, D.E. Harris, M.A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C.W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C.H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M.A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B.G. Barrell, J. Parkhill, and D.A. Hopwood, *Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2).* Nature, 2002. **417**(6885): p. 141-147.
61. Ikeda, H., J. Ishikawa, A. Hanamoto, M. Shinose, H. Kikuchi, T. Shiba, Y. Sakaki, M. Hattori, and S. Ōmura, *Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis.* Nature Biotechnology, 2003. **21**(5): p. 526-531.
62. Ikeda, H., S.Y. Kazuo, and S. Omura, *Genome mining of the Streptomyces avermitilis genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters.* J Ind Microbiol Biotechnol, 2014. **41**(2): p. 233-50.
63. Challis, G.L., *Exploitation of the Streptomyces coelicolor A3(2) genome sequence for discovery of new natural products and biosynthetic pathways.* J Ind Microbiol Biotechnol, 2014. **41**(2): p. 219-32.
64. Nett, M., H. Ikeda, and B.S. Moore, *Genomic basis for natural product biosynthetic diversity in the actinomycetes.* Nat Prod Rep, 2009. **26**(11): p. 1362-84.
65. Katz, L. and R.H. Baltz, *Natural product discovery: past, present, and future.* J Ind Microbiol Biotechnol, 2016. **43**(2-3): p. 155-76.
66. Ziemert, N., M. Alanjary, and T. Weber, *The evolution of genome mining in microbes–a review.* Natural product reports, 2016. **33**(8): p. 988-1005.
67. Zarins-Tutt, J.S., T.T. Barberi, H. Gao, A. Mearns-Spragg, L. Zhang, D.J. Newman, and R.J.M. Goss, *Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products.* Natural product reports, 2016. **33**(1): p. 54-72.
68. Blin, K., S. Shaw, A.M. Kloosterman, Z. Charlop-Powers, G.P. Van Wezel, M.H. Medema, and T. Weber, *antiSMASH 6.0: improving cluster detection and comparison capabilities.* Nucleic Acids Research, 2021. **49**(W1): p. W29-W35.
69. Blin, K., M.H. Medema, M.A. Fischbach, P. Cimermancic, P. Zakrzewski, R. Breitling, T. Weber, V. de Jager, and E. Takano, *antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences.* Nucleic Acids Research, 2011. **39**(suppl_2): p. W339-W346.
70. Bode, H.B., B. Bethe, R. Höfs, and A. Zeeck, *Big Effects from Small Changes: Possible Ways to Explore Nature's Chemical Diversity.* ChemBioChem, 2002. **3**(7): p. 619-627.
71. Pan, R., X. Bai, J. Chen, H. Zhang, and H. Wang, *Exploring structural diversity of microbe secondary metabolites using OSMAC strategy: A literature review.* Frontiers in Microbiology, 2019. **10**: p. 294.
72. Baltz, R.H., *Genetic manipulation of secondary metabolite biosynthesis for improved production in Streptomyces and other actinomycetes.* J Ind Microbiol Biotechnol, 2016. **43**(2-3): p. 343-70.
73. Ochi, K., Y. Tanaka, and S. Tojo, *Activating the expression of bacterial cryptic genes by rpoB mutations in RNA polymerase or by rare earth elements.* J Ind Microbiol Biotechnol, 2014. **41**(2): p. 403-14.

74. Nah, H.-J., H.-R. Pyeon, S.-H. Kang, S.-S. Choi, and E.-S. Kim, *Cloning and heterologous expression of a large-sized natural product biosynthetic gene cluster in Streptomyces species.* Frontiers in microbiology, 2017. **8**: p. 394.
75. Gomez-Escribano, J.P. and M.J. Bibb, *Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters.* Microb Biotechnol, 2011. **4**(2): p. 207-15.
76. Komatsu, M., T. Uchiyama, S. Ōmura, D.E. Cane, and H. Ikeda, *Genome-minimized Streptomyces host for the heterologous expression of secondary metabolism.* Proceedings of the National Academy of Sciences, 2010. **107**(6): p. 2646-2651.
77. Komatsu, M., K. Komatsu, H. Koiwai, Y. Yamada, I. Kozone, M. Izumikawa, J. Hashimoto, M. Takagi, S. Omura, K. Shin-ya, D.E. Cane, and H. Ikeda, *Engineered Streptomyces avermitilis Host for Heterologous Expression of Biosynthetic Gene Cluster for Secondary Metabolites.* ACS Synthetic Biology, 2013. **2**(7): p. 384-396.
78. Li, S., J. Wang, X. Li, S. Yin, W. Wang, and K. Yang, *Genome-wide identification and evaluation of constitutive promoters in streptomycetes.* Microbial Cell Factories, 2015. **14**(1): p. 172.
79. Luo, Y., L. Zhang, K.W. Barton, and H. Zhao, *Systematic Identification of a Panel of Strong Constitutive Promoters from Streptomyces albus.* ACS Synthetic Biology, 2015. **4**(9): p. 1001-1010.
80. Baltz, R.H., *Streptomyces and Saccharopolyspora hosts for heterologous expression of secondary metabolite gene clusters.* J Ind Microbiol Biotechnol, 2010. **37**(8): p. 759-72.
81. Sussmuth, R.D. and A. Mainz, *Nonribosomal Peptide Synthesis-Principles and Prospects.* Angewandte Chemie-International Edition, 2017. **56**(14): p. 3770-3821.
82. Lautru, S., R.J. Deeth, L.M. Bailey, and G.L. Challis, *Discovery of a new peptide natural product by Streptomyces coelicolor genome mining.* Nature Chemical Biology, 2005. **1**(5): p. 265-269.
83. Healy, F.G., M. Wach, S.B. Krasnoff, D.M. Gibson, and R. Loria, *The txtAB genes of the plant pathogen Streptomyces acidiscabies encode a peptide synthetase required for phytotoxin thaxtomin A production and pathogenicity.* Mol Microbiol, 2000. **38**(4): p. 794-804.
84. King, R.R., C.H. Lawrence, M.C. Clark, and L.A. Calhoun, *Isolation and characterization of phytotoxins associated with Streptomyces scabies.* Journal of the Chemical Society, Chemical Communications, 1989(13): p. 849-850.
85. Waksman, S.A. and H.B. Woodruff, *Bacteriostatic and Bactericidal Substances Produced by a Soil Actinomyces.* Proceedings of the Society for Experimental Biology and Medicine, 1940. **45**(2): p. 609-614.
86. Hollstein, U., *Actinomycin. Chemistry and mechanism of action.* Chemical Reviews, 1974. **74**(6): p. 625-652.
87. Stachelhaus, T., H.D. Mootz, and M.A. Marahiel, *The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases.* Chemistry & biology, 1999. **6**(8): p. 493-505.
88. Rottig, M., M.H. Medema, K. Blin, T. Weber, C. Rausch, and O. Kohlbacher, *NRPSpredictor2--a web server for predicting NRPS adenylation domain specificity.* Nucleic Acids Res, 2011. **39**(Web Server issue): p. W362-7.
89. Gordon, J.J., G.A. Miller, and B.K. Kelly, *ACTINONIN - AN ANTIBIOTIC SUBSTANCE PRODUCED BY AN ACTINOMYCETE.* Nature, 1962. **195**(4842): p. 701-&.

90. Attwood, M.M., *AN INVESTIGATION INTO MODE OF ACTION OF ACTINONIN.* Journal of General Microbiology, 1969. **55**: p. 209-&.
91. Gordon, J.J., J.P. Devlin, A.J. East, W.D. Ollis, I.O. Sutherland, D.E. Wright, and L. Ninet, *STUDIES CONCERNING ANTIBIOTIC ACTINONIN .1. CONSTITUTION OF ACTINONIN - NATURAL HYDROXAMIC ACID WITH ANTIBIOTIC ACTIVITY.* Journal of the Chemical Society-Perkin Transactions 1, 1975(9): p. 819-825.
92. Anderson, N.H., W.D. Ollis, J.E. Thorpe, and A.D. Ward, *STUDIES CONCERNING ANTIBIOTIC ACTINONIN .2. TOTAL SYNTHESIS OF ACTINONIN AND SOME STRUCTURAL ANALOGUES BY ISOMALEIMIDE METHOD.* Journal of the Chemical Society-Perkin Transactions 1, 1975(9): p. 825-830.
93. Broughton, B.J., P.J. Warren, K.R.H. Wooldridge, D.E. Wright, W.D. Ollis, and R.J. Wood, *STUDIES CONCERNING ANTIBIOTIC ACTINONIN .4. SYNTHESIS OF STRUCTURAL ANALOGUES OF ACTINONIN BY MIXED ANHYDRIDE METHOD.* Journal of the Chemical Society-Perkin Transactions 1, 1975(9): p. 842-846.
94. Devlin, J.P., W.D. Ollis, and J.E. Thorpe, *STUDIES CONCERNING ANTIBIOTIC ACTINONIN .5. SYNTHESIS OF STRUCTURAL ANALOGUES OF ACTINONIN BY ANHYDRIDE-ESTER METHOD.* Journal of the Chemical Society-Perkin Transactions 1, 1975(9): p. 846-848.
95. Devlin, J.P., W.D. Ollis, J.E. Thorpe, R.J. Wood, B.J. Broughton, P.J. Warren, K.R.H. Wooldridge, and D.E. Wright, *STUDIES CONCERNING ANTIBIOTIC ACTINONIN .3. SYNTHESIS OF STRUCTURAL ANALOGUES OF ACTINONIN BY ANHYDRIDE-IMIDE METHOD.* Journal of the Chemical Society-Perkin Transactions 1, 1975(9): p. 830-841.
96. Vanderwerf, C.A., *Steric effects in organic chemistry (Newman, Melvin S., ed.).* Journal of Chemical Education, 1957. **34**(4): p. A180.
97. Broughton, B.J., P. Chaplen, W.A. Freeman, P.J. Warren, K.R.H. Wooldridge, and D.E. Wright, *STUDIES CONCERNING ANTIBIOTIC ACTINONIN .8. STRUCTURE-ACTIVITY-RELATIONSHIPS IN ACTINONIN SERIES.* Journal of the Chemical Society-Perkin Transactions 1, 1975(9): p. 857-860.
98. Umezawa, H., T. Aoyagi, T. Tanaka, H. Suda, A. Okuyama, H. Naganawa, M. Hamada, and T. Takeuchi, *PRODUCTION OF ACTINONIN, AN INHIBITOR OF AMINOPEPTIDASE-M, BY ACTINOMYCETES.* Journal of Antibiotics, 1985. **38**(11): p. 1629-1630.
99. Sjöström, H., O. Norén, and J. Olsen, *Structure and function of aminopeptidase N.* Cellular peptidases in immune functions and diseases 2, 2002: p. 25-34.
100. Hughes, J., T.W. Smith, H.W. Kosterlitz, L.A. Fothergill, B.A. Morgan, and H.R. Morris, *Identification of two related pentapeptides from the brain with potent opiate agonist activity.* Nature, 1975. **258**(5536): p. 577-80.
101. Hughes, J., T. Smith, B. Morgan, and L. Fothergill, *Purification and properties of enkephalin — The possible endogenous ligand for the morphine receptor.* Life Sciences, 1975. **16**(12): p. 1753-1758.
102. Sar, M., W.E. Stumpf, R.J. Miller, K.J. Chang, and P. Cuatrecasas, *IMMUNOHISTOCHEMICAL LOCALIZATION OF ENKEPHALIN IN RAT-BRAIN AND SPINAL-CORD.* Journal of Comparative Neurology, 1978. **182**(1): p. 17-37.
103. Arvidsson, U., R.J. Dado, M. Riedl, J.-H. Lee, P.Y. Law, H.H. Loh, R. Elde, and M.W. Wessendorf, *delta-Opioid receptor immunoreactivity: distribution in brainstem and spinal cord, and relationship to biogenic amines and enkephalin.* Journal of Neuroscience, 1995. **15**(2): p. 1215-1235.

104. Waldhoer, M., S.E. Bartlett, and J.L. Whistler, *Opioid receptors.* Annu Rev Biochem, 2004. **73**: p. 953-90.
105. Harasawa, I., H.L. Fields, and I.D. Meng, *Delta opioid receptor mediated actions in the rostral ventromedial medulla on tail flick latency and nociceptive modulatory neurons.* Pain, 2000. **85**(1-2): p. 255-62.
106. Vanderah, T.W., *Delta and kappa opioid receptors as suitable drug targets for pain.* Clin J Pain, 2010. **26 Suppl 10**: p. S10-5.
107. Belluzzi, J.D., N. Grant, V. Garsky, D. Sarantakis, C.D. Wise, and L. Stein, *Analgesia induced in vivo by central administration of enkephalin in rat.* Nature, 1976. **260**(5552): p. 625-626.
108. de la Baume, S., C.C. Yi, J.C. Schwartz, P. Chaillet, H. Marcais-Collado, and J. Costentin, *Participation of both 'enkephalinase' and aminopeptidase activities in the metabolism of endogenous enkephalins.* Neuroscience, 1983. **8**(1): p. 143-51.
109. Gros, C., B. Giros, and J.C. Schwartz, *IDENTIFICATION OF AMINOPEPTIDASE-M AS AN ENKEPHALIN-INACTIVATING ENZYME IN RAT CEREBRAL MEMBRANES.* Biochemistry, 1985. **24**(9): p. 2179-2185.
110. Malfroy, B., J.P. Swerts, A. Guyon, B.P. Roques, and J.C. Schwartz, *High-affinity enkephalin-degrading peptidase in brain is increased after morphine.* Nature, 1978. **276**(5687): p. 523-6.
111. Hachisu, M., T. Hiranuma, Y. Shibazaki, K. Uotani, S. Murata, T. Aoyagi, and H. Umezawa, *COMPOSITE EFFECTS OF ACTINONIN WHEN INHIBITING ENKEPHALIN-DEGRADING ENZYMES.* European Journal of Pharmacology, 1987. **137**(1): p. 59-65.
112. Sayama, K., Y. Goto, T. Iguchi, Y. Takeda, and A. Matsuzawa, *Effects of an antibiotic protease inhibitor, actinonin on the growth within collagen gels of non-metastatic and metastatic mouse mammary tumors of the same origin.* Cancer Lett, 1995. **94**(2): p. 171-7.
113. Xu, Y., L.T. Lai, J.L. Gabrilove, and D.A. Scheinberg, *Antitumor activity of actinonin in vitro and in vivo.* Clin Cancer Res, 1998. **4**(1): p. 171-6.
114. Birkedal-Hansen, H., W. Moore, M. Bodden, L. Windsor, B. Birkedal-Hansen, A. DeCarlo, and J. Engler, *Matrix metalloproteinases: a review.* Critical Reviews in Oral Biology & Medicine, 1993. **4**(2): p. 197-250.
115. Ghajar, C.M., S.C. George, and A. Putnam, *Matrix metalloproteinase control of capillary morphogenesis.* Critical Reviews™ in Eukaryotic Gene Expression, 2008. **18**(3).
116. Werb, Z. and J.R. Chin, *Extracellular Matrix Remodeling during Morphogenesis a.* Annals of the New York Academy of Sciences, 1998. **857**(1): p. 110-118.
117. Liabakk, N.B., I. Talbot, R.A. Smith, K. Wilkinson, and F. Balkwill, *Matrix metalloprotease 2 (MMP-2) and matrix metalloprotease 9 (MMP-9) type IV collagenases in colorectal cancer.* Cancer Res, 1996. **56**(1): p. 190-6.
118. Azzam, H.S., G. Arand, M.E. Lippman, and E.W. Thompson, *Association of MMP-2 activation potential with metastatic progression in human breast cancer cell lines independent of MMP-2 production.* J Natl Cancer Inst, 1993. **85**(21): p. 1758-64.
119. Gilles, C., M. Polette, J. Piette, C. Munaut, E.W. Thompson, P. Birembaut, and J.M. Foidart, *High level of MT-MMP expression is associated with invasiveness of cervical cancer cells.* Int J Cancer, 1996. **65**(2): p. 209-13.

120. Hotary, K.B., E.D. Allen, P.C. Brooks, N.S. Datta, M.W. Long, and S.J. Weiss, *Membrane type I matrix metalloproteinase usurps tumor growth control imposed by the three-dimensional extracellular matrix.* Cell, 2003. **114**(1): p. 33-45.
121. Sakamoto, T. and M. Seiki, *Integrated functions of membrane-type 1 matrix metalloproteinase in regulating cancer malignancy: Beyond a proteinase.* Cancer Sci, 2017. **108**(6): p. 1095-1100.
122. Di Sebastiano, P., F.F. di Mola, L. Artese, C. Rossi, G. Mascetta, H. Pernthaler, and P. Innocenti, *Beneficial effects of Batimastat (BB-94), a matrix metalloproteinase inhibitor, in rat experimental colitis.* Digestion, 2001. **63**(4): p. 234-9.
123. Muri, L., D. Grandgirard, M. Buri, M. Perny, and S.L. Leib, *Combined effect of non-bacteriolytic antibiotic and inhibition of matrix metalloproteinases prevents brain injury and preserves learning, memory and hearing function in experimental paediatric pneumococcal meningitis.* J Neuroinflammation, 2018. **15**(1): p. 233.
124. Asahi, M., K. Asahi, J.C. Jung, G.J. del Zoppo, M.E. Fini, and E.H. Lo, *Role for matrix metalloproteinase 9 after focal cerebral ischemia: effects of gene knockout and enzyme inhibition with BB-94.* J Cereb Blood Flow Metab, 2000. **20**(12): p. 1681-9.
125. Koistinaho, M., T.M. Malm, M.I. Kettunen, G. Goldsteins, S. Starckx, R.A. Kauppinen, G. Opdenakker, and J. Koistinaho, *Minocycline protects against permanent cerebral ischemia in wild type but not in matrix metalloprotease-9-deficient mice.* J Cereb Blood Flow Metab, 2005. **25**(4): p. 460-7.
126. Kuzuya, M., K. Nakamura, T. Sasaki, X.W. Cheng, S. Itohara, and A. Iguchi, *Effect of MMP-2 deficiency on atherosclerotic lesion formation in apoE-deficient mice.* Arterioscler Thromb Vasc Biol, 2006. **26**(5): p. 1120-5.
127. Luttun, A., E. Lutgens, A. Manderveld, K. Maris, D. Collen, P. Carmeliet, and L. Moons, *Loss of matrix metalloproteinase-9 or matrix metalloproteinase-12 protects apolipoprotein E-deficient mice against atherosclerotic media destruction but differentially affects plaque growth.* Circulation, 2004. **109**(11): p. 1408-14.
128. Chen, D.Z., D.V. Patel, C.J. Hackbarth, W. Wang, G. Dreyer, D.C. Young, P.S. Margolis, C. Wu, Z.-J. Ni, J. Trias, R.J. White, and Z. Yuan, *Actinonin, a Naturally Occurring Antibacterial Agent, Is a Potent Deformylase Inhibitor.* Biochemistry, 2000. **39**(6): p. 1256-1262.
129. Guillon, J.M., Y. Mechulam, J.M. Schmitter, S. Blanquet, and G. Fayat, *Disruption of the gene for Met-tRNA(fMet) formyltransferase severely impairs growth of Escherichia coli.* Journal of bacteriology, 1992. **174**(13): p. 4294-4301.
130. RajBhandary, U.L., *Initiator transfer RNAs.* J Bacteriol, 1994. **176**(3): p. 547-52.
131. Rajagopalan, P.T., A. Datta, and D. Pei, *Purification, characterization, and inhibition of peptide deformylase from Escherichia coli.* Biochemistry, 1997. **36**(45): p. 13910-8.
132. Chan, M.K., W. Gong, P.T.R. Rajagopalan, B. Hao, C.M. Tsai, and D. Pei, *Crystal Structure of the Escherichia coli Peptide Deformylase.* Biochemistry, 1997. **36**(45): p. 13904-13909.
133. Dardel, F., S. Ragusa, C. Lazennec, S. Blanquet, and T. Meinnel, *Solution structure of nickel-peptide deformylase 11Edited by A. R. Fersht.* Journal of Molecular Biology, 1998. **280**(3): p. 501-513.
134. Rajagopalan, P.T.R., X.C. Yu, and D. Pei, *Peptide Deformylase: A New Type of Mononuclear Iron Protein.* Journal of the American Chemical Society, 1997. **119**(50): p. 12418-12419.

135. Piatkov, K.I., T.T.M. Vu, C.-S. Hwang, and A. Varshavsky, *Formyl-methionine as a degradation signal at the N-termini of bacterial proteins.* Microbial cell (Graz, Austria), 2015. **2**(10): p. 376-393.
136. Clements, J.M., R.P. Beckett, A. Brown, G. Catlin, M. Lobell, S. Palan, W. Thomas, M. Whittaker, S. Wood, S. Salama, P.J. Baker, H.F. Rodgers, V. Barynin, D.W. Rice, and M.G. Hunter, *Antibiotic activity and characterization of BB-3497, a novel peptide deformylase inhibitor.* Antimicrob Agents Chemother, 2001. **45**(2): p. 563-70.
137. Amberg-Johnson, K., S.B. Hari, S.M. Ganesan, H.A. Lorenzi, R.T. Sauer, J.C. Niles, and E. Yeh, *Small molecule inhibition of apicomplexan FtsH1 disrupts plastid biogenesis in human pathogens.* Elife, 2017. **6**.
138. Goodman, C.D., T. Uddin, N.J. Spillman, and G.I. McFadden, *A single point mutation in the Plasmodium falciparum ftsh1 metalloprotease confers actinonin resistance.* bioRxiv, 2020: p. 2020.05.13.092882.
139. Uddin, T., G.I. McFadden, and C.D. Goodman, *Validation of Putative Apicoplast-Targeting Drugs Using a Chemical Supplementation Assay in Cultured Human Malaria Parasites.* Antimicrobial Agents and Chemotherapy, 2018. **62**(1): p. 17.
140. Bellomo, R., C. Ronco, J.A. Kellum, R.L. Mehta, P. Palevsky, and t.A. workgroup, *Acute renal failure – definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group.* Critical Care, 2004. **8**(4): p. R204.
141. Versteilen, A.M., F. Di Maggio, J.R. Leemreis, A.B. Groeneveld, R.J. Musters, and P. Sipkema, *Molecular mechanisms of acute renal failure following ischemia/reperfusion.* Int J Artif Organs, 2004. **27**(12): p. 1019-29.
142. Majumdar, A., *Sepsis-induced acute kidney injury.* Indian journal of critical care medicine : peer-reviewed, official publication of Indian Society of Critical Care Medicine, 2010. **14**(1): p. 14-21.
143. Bylander, J., Q. Li, G. Ramesh, B. Zhang, W.B. Reeves, and J.S. Bond, *Targeted disruption of the meprin metalloproteinase beta gene protects against renal ischemia-reperfusion injury in mice.* Am J Physiol Renal Physiol, 2008. **294**(3): p. F480-90.
144. Kruse, M.N., C. Becker, D. Lottaz, D. Kohler, I. Yiallouros, H.W. Krell, E.E. Sterchi, and W. Stocker, *Human meprin alpha and beta homo-oligomers: cleavage of basement membrane proteins and sensitivity to metalloprotease inhibitors.* Biochemical Journal, 2004. **378**: p. 383-389.
145. Ongeri, E.M., O. Anyanwu, W.B. Reeves, and J.S. Bond, *Villin and actin in the mouse kidney brush-border membrane bind to and are degraded by meprins, an interaction that contributes to injury in ischemia-reperfusion.* American Journal of Physiology-Renal Physiology, 2011. **301**(4): p. F871-F882.
146. Carmago, S., S.V. Shah, and P.D. Walker, *Meprin, a brush-border enzyme, plays an important role in hypoxic/ischemic acute renal tubular injury in rats.* Kidney Int, 2002. **61**(3): p. 959-66.
147. Wang, Z., C. Herzog, G.P. Kaushal, N. Gokden, and P.R. Mayeux, *ACTINONIN, A MEPRIN A INHIBITOR, PROTECTS THE RENAL MICROCIRCULATION DURING SEPSIS.* Shock, 2011. **35**(2): p. 141-147.
148. Holly, M.K., J.W. Dear, X. Hu, A.N. Schechter, M.T. Gladwin, S.M. Hewitt, P.S.T. Yuen, and R.A. Star, *Biomarker and drug-target discovery using proteomics in a new rat*

model of sepsis-induced acute renal failure. Kidney International, 2006. **70**(3): p. 496-506.
149. Gao, P., R.W. Guo, J.F. Chen, Y. Chen, H. Wang, Y. Yu, and L. Huang, *A meprin inhibitor suppresses atherosclerotic plaque formation in ApoE(-/-) mice.* Atherosclerosis, 2009. **207**(1): p. 84-92.
150. Chepurnova, D., V. Samoilova E, D. Verin A, A.G. Fesenko, A. Anisimov, and A.A. Korotaeva, *Inhibition of Meprins Reduces Pulmonary Edema in LPS-Induced Acute Lung Damage.* Bull Exp Biol Med, 2019. **166**(6): p. 719-721.
151. Yuan, Z., J. Trias, and R.J. White, *Deformylase as a novel antibacterial target.* Drug Discov Today, 2001. **6**(18): p. 954-961.
152. Giglione, C., M. Pierre, and T. Meinnel, *Peptide deformylase as a target for new generation, broad spectrum antimicrobial agents.* Molecular Microbiology, 2000. **36**(6): p. 1197-1205.
153. Mazel, D., S. Pochet, and P. Marlière, *Genetic characterization of polypeptide deformylase, a distinctive enzyme of eubacterial translation.* Embo j, 1994. **13**(4): p. 914-23.
154. Broughton, B.J., P. Chaplen, W.A. Freeman, P.J. Warren, K.R.H. Wooldridge, and D.E. Wright, *Studies concerning the antibiotic actinonin. Part VIII. Structure–activity relationships in the actinonin series.* Journal of the Chemical Society, Perkin Transactions 1, 1975(9): p. 857-860.
155. Goemaere, E., A. Melet, V. Larue, A. Lieutaud, R. Alves de Sousa, J. Chevalier, L. Yimga-Djapa, C. Giglione, F. Huguet, M. Alimi, T. Meinnel, F. Dardel, I. Artaud, and J.M. Pagès, *New peptide deformylase inhibitors and cooperative interaction: a combination to improve antibacterial activity.* J Antimicrob Chemother, 2012. **67**(6): p. 1392-400.
156. Guilloteau, J.-P., M. Mathieu, C. Giglione, V. Blanc, A. Dupuy, M. Chevrier, P. Gil, A. Famechon, T. Meinnel, and V. Mikol, *The crystal structures of four peptide deformylases bound to the antibiotic actinonin reveal two distinct types: a platform for the structure-based design of antibacterial agents.* Journal of molecular biology, 2002. **320**(5): p. 951-962.
157. Smith, H.K., R.P. Beckett, J.M. Clements, S. Doel, S.P. East, S.B. Launchbury, L.M. Pratt, Z.M. Spavold, W. Thomas, and R.S. Todd, *Structure–activity relationships of the peptide deformylase inhibitor BB-3497: modification of the metal binding group.* Bioorganic & medicinal chemistry letters, 2002. **12**(24): p. 3595-3599.
158. Davies, S.J., A.P. Ayscough, R.P. Beckett, R.A. Bragg, J.M. Clements, S. Doel, C. Grew, S.B. Launchbury, G.M. Perkins, and L.M. Pratt, *Structure–activity relationships of the peptide deformylase inhibitor BB-3497: modification of the methylene spacer and the P1' side chain.* Bioorganic & medicinal chemistry letters, 2003. **13**(16): p. 2709-2713.
159. Davies, S.J., A.P. Ayscough, R.P. Beckett, J.M. Clements, S. Doel, L.M. Pratt, Z.M. Spavold, S.W. Thomas, and M. Whittaker, *Structure–activity relationships of the peptide deformylase inhibitor BB-3497: modification of the P2' and P3' side chains.* Bioorganic & medicinal chemistry letters, 2003. **13**(16): p. 2715-2718.
160. Chen, D., C. Hackbarth, Z. Ni, C. Wu, W. Wang, R. Jain, Y. He, K. Bracken, B. Weidmann, and D. Patel, *Peptide deformylase inhibitors as antibacterial agents: identification of VRC3375, a proline-3-alkylsuccinyl hydroxamate derivative, by*

using an integrated combinatorial and medicinal chemistry approach. Antimicrobial agents and chemotherapy, 2004. **48**(1): p. 250-261.
161. Boularot, A., C. Giglione, I. Artaud, and T. Meinnel, *Structure-activity relationship analysis and therapeutic potential of peptide deformylase inhibitors.* Curr Opin Investig Drugs, 2004. **5**(8): p. 809-22.
162. Credito, K., G. Lin, L.M. Ednie, and P.C. Appelbaum, *Antistaphylococcal activity of LBM415, a new peptide deformylase inhibitor, compared with those of other agents.* Antimicrob Agents Chemother, 2004. **48**(10): p. 4033-6.
163. Chen, D. and Z. Yuan, *Therapeutic potential of peptide deformylase inhibitors.* Expert Opinion on Investigational Drugs, 2005. **14**(9): p. 1107-1116.
164. Rolan, P., H. Sun, C. Macleod, K. Bracken, and T.G. Evans, *Pharmacokinetics and unexpected safety issues of LBM415, a novel oral peptide deformylase inhibitor.* Clin Pharmacol Ther, 2011. **90**(2): p. 256-62.
165. Ramanathan-Girish, S., J. McColm, J.M. Clements, P. Taupin, S. Barrowcliffe, J. Hevizi, S. Safrin, C. Moore, G. Patou, H. Moser, A. Gadd, U. Hoch, V. Jiang, D. Lofland, and K.W. Johnson, *Pharmacokinetics in animals and humans of a first-in-class peptide deformylase inhibitor.* Antimicrob Agents Chemother, 2004. **48**(12): p. 4835-42.
166. O'Dwyer, K., M. Hackel, S. Hightower, D. Hoban, S. Bouchillon, D. Qin, K. Aubart, M. Zalacain, and D. Butler, *Comparative analysis of the antibacterial activity of a novel peptide deformylase inhibitor, GSK1322322.* Antimicrob Agents Chemother, 2013. **57**(5): p. 2333-42.
167. Butler, D., D. Chen, K. O'Dwyer, T. Lewandowski, K. Aubart, and M. Zalacain, *Potent sub-MIC effect of GSK1322322 and other peptide deformylase inhibitors on in vitro growth of Staphylococcus aureus.* Antimicrob Agents Chemother, 2014. **58**(1): p. 290-6.
168. Naderer, O.J., L.S. Jones, J. Zhu, M. Kurtinecz, and E. Dumont, *Safety, tolerability, and pharmacokinetics of oral and intravenous administration of GSK1322322, a peptide deformylase inhibitor.* J Clin Pharmacol, 2013. **53**(11): p. 1168-76.
169. Naderer, O.J., E. Dumont, J. Zhu, M. Kurtinecz, and L.S. Jones, *Safety, tolerability and pharmacokinetics of repeat dosing of the antibiotic GSK1322322, a peptide deformylase inhibitor: a randomized placebo-controlled study.* Journal of Antimicrobial Chemotherapy, 2013. **68**(8): p. 1901-1909.
170. Naderer, O., L.S. Jones, J. Zhu, M.D. Coffin, M. Kurtinecz, and E. Dumont, *The Effect of Food and Formulation on the Pharmacokinetics, Safety, and Tolerability of GSK1322322 in Healthy Volunteers.* Clin Pharmacol Drug Dev, 2015. **4**(1): p. 49-55.
171. Corey, R., O.J. Naderer, W.D. O'Riordan, E. Dumont, L.S. Jones, M. Kurtinecz, and J.Z. Zhu, *Safety, tolerability, and efficacy of GSK1322322 in the treatment of acute bacterial skin and skin structure infections.* Antimicrob Agents Chemother, 2014. **58**(11): p. 6518-27.
172. Hoover, J.L., C.M. Singley, P. Elefante, P. DeMarsh, M. Zalacain, and S. Rittenhouse, *Reducing Antibacterial Development Risk for GSK1322322 by Exploring Potential Human Dose Regimens in Nonclinical Efficacy Studies Using Immunocompetent Rats.* Antimicrob Agents Chemother, 2017. **61**(11).
173. Giglione, C., A. Serero, M. Pierre, B. Boisson, and T. Meinnel, *Identification of eukaryotic peptide deformylases reveals universality of N-terminal protein processing mechanisms.* Embo j, 2000. **19**(21): p. 5916-29.

174. Giglione, C., S. Fieulaine, and T. Meinnel, *N-terminal protein modifications: Bringing back into play the ribosome.* Biochimie, 2015. **114**: p. 134-146.
175. Serero, A., C. Giglione, and T. Meinnel, *Distinctive features of the two classes of eukaryotic peptide deformylases.* J Mol Biol, 2001. **314**(4): p. 695-708.
176. Dirk, L.M., M.A. Williams, and R.L. Houtz, *Eukaryotic peptide deformylases. Nuclear-encoded and chloroplast-targeted enzymes in Arabidopsis.* Plant Physiol, 2001. **127**(1): p. 97-107.
177. Giglione, C., O. Vallon, and T. Meinnel, *Control of protein life-span by N-terminal methionine excision.* Embo j, 2003. **22**(1): p. 13-23.
178. Nguyen, K.T., X.B. Hu, C. Colton, R. Chakrabarti, M.X. Zhu, and D.H. Pei, *Characterization of a human peptide deformylase: Implications for antibacterial drug design.* Biochemistry, 2003. **42**(33): p. 9952-9958.
179. Lee, M.D., C. Antczak, Y. Li, F.M. Sirotnak, W.G. Bornmann, and D.A. Scheinberg, *A new human peptide deformylase inhibitable by actinonin.* Biochem Biophys Res Commun, 2003. **312**(2): p. 309-15.
180. Escobar-Alvarez, S., Y. Goldgur, G. Yang, O. Ouerfelli, Y. Li, and D.A. Scheinberg, *Structure and Activity of Human Mitochondrial Peptide Deformylase, a Novel Cancer Target.* Journal of Molecular Biology, 2009. **387**(5): p. 1211-1228.
181. Serero, A., C. Giglione, A. Sardini, J. Martinez-Sanz, and T. Meinnel, *An unusual peptide deformylase features in the human mitochondrial N-terminal methionine excision pathway.* J Biol Chem, 2003. **278**(52): p. 52953-63.
182. Lee, M.D., Y.H. She, M.J. Soskis, C.P. Borella, J.R. Gardner, P.A. Hayes, B.M. Dy, M.L. Heaney, M.R. Philips, W.G. Bornmann, F.M. Sirotnak, and D.A. Scheinberg, *Human mitochondrial peptide deformylase, a new anticancer target of actinonin-based antibiotics.* Journal of Clinical Investigation, 2004. **114**(8): p. 1107-1116.
183. Fieulaine, S., C. Juillan-Binard, A. Serero, F. Dardel, C. Giglione, T. Meinnel, and J.L. Ferrer, *The crystal structure of mitochondrial (Type 1A) peptide deformylase provides clear guidelines for the design of inhibitors specific for the bacterial forms.* J Biol Chem, 2005. **280**(51): p. 42315-24.
184. Boularot, A., C. Giglione, S. Petit, Y. Duroc, R. Alves de Sousa, V. Larue, T. Cresteil, F. Dardel, I. Artaud, and T. Meinnel, *Discovery and Refinement of a New Structural Class of Potent Peptide Deformylase Inhibitors.* Journal of Medicinal Chemistry, 2007. **50**(1): p. 10-20.
185. Cai, Y., P. Chandrangsu, A. Gaballa, and J.D. Helmann, *Lack of formylated methionyl-tRNA has pleiotropic effects on Bacillus subtilis.* Microbiology, 2017. **163**(2): p. 185-196.
186. Newton, D.T., C. Creuzenet, and D. Mangroo, *Formylation is not essential for initiation of protein synthesis in all eubacteria.* J Biol Chem, 1999. **274**(32): p. 22143-6.
187. Margolis, P.S., C.J. Hackbarth, D.C. Young, W. Wang, D. Chen, Z. Yuan, R. White, and J. Trias, *Peptide deformylase in Staphylococcus aureus: resistance to inhibition is mediated by mutations in the formyltransferase gene.* Antimicrob Agents Chemother, 2000. **44**(7): p. 1825-31.
188. Duroc, Y., C. Giglione, and T. Meinnel, *Mutations in three distinct loci cause resistance to peptide deformylase inhibitors in Bacillus subtilis.* Antimicrob Agents Chemother, 2009. **53**(4): p. 1673-8.

189. Nilsson, A.I., A. Zorzet, A. Kanth, S. Dahlström, O.G. Berg, and D.I. Andersson, *Reducing the fitness cost of antibiotic resistance by amplification of initiator tRNA genes.* Proc Natl Acad Sci U S A, 2006. **103**(18): p. 6976-81.
190. Steiner-Mosonyi, M., C. Creuzenet, R.A. Keates, B.R. Strub, and D. Mangroo, *The Pseudomonas aeruginosa initiation factor IF-2 is responsible for formylation-independent protein initiation in P. aeruginosa.* J Biol Chem, 2004. **279**(50): p. 52262-9.
191. Zorzet, A., M.Y. Pavlov, A.I. Nilsson, M. Ehrenberg, and D.I. Andersson, *Error-prone initiation factor 2 mutations reduce the fitness cost of antibiotic resistance.* Mol Microbiol, 2010. **75**(5): p. 1299-313.
192. Caughlan, R.E., S. Sriram, D.M. Daigle, A.L. Woods, J. Buco, R.L. Peterson, J. Dzink-Fox, S. Walker, and C.R. Dean, *Fmt bypass in Pseudomonas aeruginosa causes induction of MexXY efflux pump expression.* Antimicrob Agents Chemother, 2009. **53**(12): p. 5015-21.
193. Margolis, P., C. Hackbarth, S. Lopez, M. Maniar, W. Wang, Z. Yuan, R. White, and J. Trias, *Resistance of Streptococcus pneumoniae to deformylase inhibitors is due to mutations in defB.* Antimicrob Agents Chemother, 2001. **45**(9): p. 2432-5.
194. Escobar-Alvarez, S., J. Gardner, A. Sheth, G. Manfredi, G. Yang, O. Ouerfelli, M.L. Heaney, and D.A. Scheinberg, *Inhibition of Human Peptide Deformylase Disrupts Mitochondrial Function.* Molecular and Cellular Biology, 2010. **30**(21): p. 5099-5109.
195. Sheth, A., S. Escobar-Alvarez, J. Gardner, L. Ran, M.L. Heaney, and D.A. Scheinberg, *Inhibition of human mitochondrial peptide deformylase causes apoptosis in c-myc-overexpressing hematopoietic cancers.* Cell Death & Disease, 2014. **5**: p. 9.
196. Randhawa, H., S. Chikara, D. Gehring, T. Yildirim, J. Menon, and K.M. Reindl, *Overexpression of peptide deformylase in breast, colon, and lung cancers.* Bmc Cancer, 2013. **13**: p. 7.
197. Hu, L., X. Cai, S. Dong, Y. Zhen, J. Hu, S. Wang, J. Jiang, J. Huang, Y. Han, Y. Qian, Y. Yuan, and W. Hu, *Synthesis and Anticancer Activity of Novel Actinonin Derivatives as HsPDF Inhibitors.* Journal of Medicinal Chemistry, 2020. **63**(13): p. 6959-6978.
198. Kousalová, J., M. Šírová, L. Kostka, V. Šubr, J. Kovářová, K. Běhalová, M. Studenovský, M. Kovář, and T. Etrych, *Metastatic spread inhibition of cancer cells through stimuli-sensitive HPMA copolymer-bound actinonin nanomedicines.* Nanomedicine: Nanotechnology, Biology and Medicine, 2022. **44**: p. 102578.
199. Ogita, T., A. Sato, R. Enokita, K. Suzuki, M. Ishii, T. Negishi, T. Okazaki, K. Tamaki, and K. Tanzawa, *MATLYSTATINS, NEW INHIBITORS OF TYPE-IV COLLAGENASES FROM ACTINOMADURA-ATRAMENTARIA .1. TAXONOMY, FERMENTATION, ISOLATION, AND PHYSICOCHEMICAL PROPERTIES OF MATLYSTATIN-GROUP COMPOUNDS.* Journal of Antibiotics, 1992. **45**(11): p. 1723-1732.
200. Miyadoh, S., S. Amano, H. Tohyama, and T. Shomura, *Actinomadura atramentaria, a New Species of the Actinomycetales.* International Journal of Systematic and Evolutionary Microbiology, 1987. **37**(4): p. 342-346.
201. Tamaki, K., S. Kurihara, and Y. Sugimura, *TOTAL SYNTHESIS OF MATLYSTATIN-A.* Tetrahedron Letters, 1993. **34**(52): p. 8477-8480.
202. Tamaki, K., T. Ogita, K. Tanzawa, and Y. Sugimura, *SYNTHESIS AND DETERMINATION OF THE ABSOLUTE-CONFIGURATION OF MATLYSTATIN-B.* Tetrahedron Letters, 1993. **34**(4): p. 683-686.

203. Haruyama, H., Y. Ohkuma, H. Nagaki, T. Ogita, K. Tamaki, and T. Kinoshita, *MATLYSTATINS, NEW INHIBITORS OF TYPE-IV COLLAGENASES FROM ACTINOMADURA-ATRAMENTARIA .3. STRUCTURE ELUCIDATION OF MATLYSTATIN-A TO MATLYSTATIN-F.* Journal of Antibiotics, 1994. **47**(12): p. 1473-1480.
204. Tanzawa, K., M. Ishii, T. Ogita, and K. Shimada, *MATLYSTATINS, NEW INHIBITORS OF TYPEIV COLLAGENASES FROM ACTINOMADURA-ATRAMENTARIA .2. BIOLOGICAL-ACTIVITIES.* Journal of Antibiotics, 1992. **45**(11): p. 1733-1737.
205. Fujii, H., M. Nakajima, T. Aoyagi, and T. Tsuruo, *Inhibition of tumor cell invasion and matrix degradation by aminopeptidase inhibitors.* Biological & Pharmaceutical Bulletin, 1996. **19**(1): p. 6-10.
206. Tamaki, K., K. Tanzawa, S. Kurihara, T. Oikawa, S. Monma, K. Shimada, and Y. Sugimura, *SYNTHESIS AND STRUCTURE-ACTIVITY-RELATIONSHIPS OF GELATINASE INHIBITORS DERIVED FROM MATLYSTATINS.* Chemical & Pharmaceutical Bulletin, 1995. **43**(11): p. 1883-1893.
207. Leipoldt, F., J. Santos-Aberturas, D.P. Stegmann, F. Wolf, A. Kulik, R. Lacret, D. Popadic, D. Keinhorster, N. Kirchner, P. Bekiesch, H. Gross, A.W. Truman, and L. Kaysser, *Warhead biosynthesis and the origin of structural diversity in hydroxamate metalloproteinase inhibitors.* Nature Communications, 2017. **8**.
208. INAOKA, Y., H. TAMAOKI, S. TAKAHASHI, R. ENOKITA, and T. OKAZAKI, *PROPIOXATINS A AND B, NEW ENKEPHALINASE B INHIBITORS I. TAXONOMY, FERMENTATION, ISOLATION AND BIOLOGICAL PROPERTIES.* The Journal of Antibiotics, 1986. **39**(10): p. 1368-1377.
209. INAOKA, Y., S. TAKAHASHI, and T. KINOSHITA, *PROPIOXATINS A AND B, NEW ENKEPHALINASE B INHIBITORS II. STRUCTURAL ELUCIDATION.* The Journal of Antibiotics, 1986. **39**(10): p. 1378-1381.
210. INAOKA, Y., S. TAKAHASHI, and S. SATO, *Propioxatins A and B, new enkephalinase B inhibitors III. Total synthesis of propioxatin A.* The Journal of Antibiotics, 1986. **39**(10): p. 1382-1385.
211. Inaoka, Y. and S. Naruto, *Propioxatins A and B, new enkephalinase B inhibitors. IV. Characterization of the active site of the enzyme using synthetic propioxatin analogues.* The Journal of Biochemistry, 1988. **104**(5): p. 706-711.
212. Okuyama, A., K. Naito, H. Morishima, H. Suda, S. Nishimura, and N. Tanaka, *Inhibition of growth of human tumor cells in nude mice by a metalloproteinase inhibitor.* Annals of the New York Academy of Sciences, 1994. **732**(1): p. 408-410.
213. Naito, K., S. Nakajima, N. Kanbayashi, A. Okuyama, and M. Goto, *Inhibition of metalloproteinase activity of rheumatoid arthritis synovial cells by a new inhibitor [BE16627B; LN-(N-hydroxy-2-isobutylsuccinamoyl)-seryl-L-valine].* Agents and Actions, 1993. **39**(3): p. 182-186.
214. Schneider, K., R. Peyraud, P. Kiefer, P. Christen, N. Delmotte, S. Massou, J.C. Portais, and J.A. Vorholt, *The ethylmalonyl-CoA pathway is used in place of the glyoxylate cycle by Methylobacterium extorquens AM1 during growth on acetate.* J Biol Chem, 2012. **287**(1): p. 757-66.
215. Kornberg, H.L. and N.B. Madsen, *The metabolism of C2 compounds in microorganisms. 3. Synthesis of malate from acetate via the glyoxylate cycle.* Biochem J, 1958. **68**(3): p. 549-57.
216. Peyraud, R., P. Kiefer, P. Christen, S. Massou, J.-C. Portais, and J.A. Vorholt, *Demonstration of the ethylmalonyl-CoA pathway by using ¹³C*

metabolomics. Proceedings of the National Academy of Sciences, 2009. **106**(12): p. 4846-4851.
217. Wolf, F., F. Leipoldt, A. Kulik, D. Wibberg, J. Kalinowski, and L. Kaysser, *Characterization of the Actinonin Biosynthetic Gene Cluster.* Chembiochem, 2018. **19**(11): p. 1189-1195.
218. Wallace, K.K., Z.-Y. Bao, H. Dai, R. Digate, G. Schuler, M.K. Speedie, and K.A. Reynolds, *Purification of Crotonyl-CoA Reductase from Streptomyces collinus and Cloning, Sequencing and Expression of the Corresponding Gene in Escherichia coli.* European Journal of Biochemistry, 1995. **233**(3): p. 954-962.
219. Erb, T.J., I.A. Berg, V. Brecht, M. Muller, G. Fuchs, and B.E. Alber, *Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway.* Proc Natl Acad Sci U S A, 2007. **104**(25): p. 10631-6.
220. Erb, T.J., V. Brecht, G. Fuchs, M. Müller, and B.E. Alber, *Carboxylation mechanism and stereochemistry of crotonyl-CoA carboxylase/reductase, a carboxylating enoyl-thioester reductase.* Proceedings of the National Academy of Sciences, 2009. **106**(22): p. 8871-8876.
221. Wilson, M.C. and B.S. Moore, *Beyond ethylmalonyl-CoA: The functional role of crotonyl-CoA carboxylase/reductase homologs in expanding polyketide diversity.* Natural Product Reports, 2012. **29**(1): p. 72-86.
222. Williams, P.G., G.O. Buchanan, R.H. Feling, C.A. Kauffman, P.R. Jensen, and W. Fenical, *New Cytotoxic Salinosporamides from the Marine Actinomycete Salinispora tropica.* The Journal of Organic Chemistry, 2005. **70**(16): p. 6196-6203.
223. Eustáquio, A.S., R.P. McGlinchey, Y. Liu, C. Hazzard, L.L. Beer, G. Florova, M.M. Alhamadsheh, A. Lechner, A.J. Kale, Y. Kobayashi, K.A. Reynolds, and B.S. Moore, *Biosynthesis of the salinosporamide A polyketide synthase substrate chloroethylmalonyl-coenzyme A from S-adenosyl-l-methionine.* Proceedings of the National Academy of Sciences, 2009. **106**(30): p. 12295-12300.
224. Liu, Y., C. Hazzard, A.S. Eustáquio, K.A. Reynolds, and B.S. Moore, *Biosynthesis of salinosporamides from alpha,beta-unsaturated fatty acids: implications for extending polyketide synthase diversity.* Journal of the American Chemical Society, 2009. **131**(30): p. 10376-10377.
225. Schreiber, S.L. and G.R. Crabtree, *The mechanism of action of cyclosporin A and FK506.* Immunology today, 1992. **13**(4): p. 136-142.
226. Mo, S., D.H. Kim, J.H. Lee, J.W. Park, D.B. Basnet, Y.H. Ban, Y.J. Yoo, S.-w. Chen, S.R. Park, E.A. Choi, E. Kim, Y.-Y. Jin, S.-K. Lee, J.Y. Park, Y. Liu, M.O. Lee, K.S. Lee, S.J. Kim, D. Kim, B.C. Park, S.-g. Lee, H.J. Kwon, J.-W. Suh, B.S. Moore, S.-K. Lim, and Y.J. Yoon, *Biosynthesis of the Allylmalonyl-CoA Extender Unit for the FK506 Polyketide Synthase Proceeds through a Dedicated Polyketide Synthase and Facilitates the Mutasynthesis of Analogues.* Journal of the American Chemical Society, 2011. **133**(4): p. 976-985.
227. Yoo, H.G., S.Y. Kwon, S. Kim, S. Karki, Z.Y. Park, and H.J. Kwon, *Characterization of 2-Octenoyl-CoA Carboxylase/Reductase Utilizing pteB from Streptomyce avermitilis.* Bioscience Biotechnology and Biochemistry, 2011. **75**(6): p. 1191-1193.
228. Buntin, K., H. Irschik, K.J. Weissman, E. Luxenburger, H. Blöcker, and R. Müller, *Biosynthesis of thuggacins in myxobacteria: comparative cluster analysis reveals basis for natural product structural diversity.* Chem Biol, 2010. **17**(4): p. 342-56.

229. Rachid, S., L. Huo, J. Herrmann, M. Stadler, B. Köpcke, J. Bitzer, and R. Müller, *Mining the cinnabaramide biosynthetic pathway to generate novel proteasome inhibitors.* Chembiochem, 2011. **12**(6): p. 922-31.
230. Quade, N., L. Huo, S. Rachid, D.W. Heinz, and R. Muller, *Unusual carbon fixation gives rise to diverse polyketide extender units.* Nat Chem Biol, 2011. **8**(1): p. 117-24.
231. Whitfield, G.B., T.D. Brock, A. Ammann, D. Gottlieb, and H.E. Carter, *Filipin, an antifungal antibiotic: isolation and properties.* Journal of the American Chemical Society, 1955. **77**(18): p. 4799-4801.
232. Irschik, H., H. Reichenbach, G. Höfle, and R. Jansen, *The thuggacins, novel antibacterial macrolides from Sorangium cellulosum acting against selected Gram-positive bacteria.* The Journal of antibiotics, 2007. **60**(12): p. 733-738.
233. Stadler, M., J. Bitzer, A. Mayer-Bartschmid, H. Müller, J. Benet-Buchholz, F. Gantner, H.-V. Tichy, P. Reinemer, and K.B. Bacon, *Cinnabaramides A–G: Analogues of Lactacystin and Salinosporamide from a Terrestrial Streptomycete.* Journal of Natural Products, 2007. **70**(2): p. 246-252.
234. Chan, Y.A., A.M. Podevels, B.M. Kevany, and M.G. Thomas, *Biosynthesis of polyketide synthase extender units.* Nat Prod Rep, 2009. **26**(1): p. 90-114.
235. Erb, T.J., J. Rétey, G. Fuchs, and B.E. Alber, *Ethylmalonyl-CoA Mutase from Rhodobacter sphaeroides Defines a New Subclade of Coenzyme B12-dependent Acyl-CoA Mutases.* Journal of Biological Chemistry, 2008. **283**(47): p. 32283-32293.
236. Martínez, M.A., A. Rincón, L.R. Desviat, B. Merinero, M. Ugarte, and B. Pérez, *Genetic analysis of three genes causing isolated methylmalonic acidemia: identification of 21 novel allelic variants.* Mol Genet Metab, 2005. **84**(4): p. 317-25.
237. Forny, P., D.S. Froese, T. Suormala, W.W. Yue, and M.R. Baumgartner, *Functional characterization and categorization of missense mutations that cause methylmalonyl-CoA mutase (MUT) deficiency.* Hum Mutat, 2014. **35**(12): p. 1449-58.
238. Kaziro, Y., S. Ochoa, R.C. Warner, and J.-Y. Chen, *Metabolism of Propionic Acid in Animal Tissues: VIII. CRYSTALLINE PROPIONYL CARBOXYLASE.* Journal of Biological Chemistry, 1961. **236**(7): p. 1917-1923.
239. Mazumder, R., T. Sasakawa, Y. Kaziro, and S. Ochoa, *Metabolism of propionic acid in animal tissues. IX. Methylmalonyl coenzyme A racemase.* J Biol Chem, 1962. **237**: p. 3065-8.
240. Abily-Donval, L., S. Torre, A. Samson, B. Sudrié-Arnaud, C. Acquaviva, A.M. Guerrot, J.F. Benoist, S. Marret, S. Bekri, and A. Tebani, *Methylmalonyl-CoA Epimerase Deficiency Mimicking Propionic Aciduria.* Int J Mol Sci, 2017. **18**(11).
241. Halarnkar, P.P. and G.J. Blomquist, *Comparative aspects of propionate metabolism.* Comparative Biochemistry and Physiology Part B: Comparative Biochemistry, 1989. **92**(2): p. 227-231.
242. Cracan, V. and R. Banerjee, *Novel B12-Dependent Acyl-CoA Mutases and Their Biotechnological Potential.* Biochemistry, 2012. **51**(31): p. 6039-6046.
243. Jung, W.S., Y.J. Yoo, J.W. Park, S.R. Park, A.R. Han, Y.H. Ban, E.J. Kim, E. Kim, and Y.J. Yoon, *A combined approach of classical mutagenesis and rational metabolic engineering improves rapamycin biosynthesis and provides insights into methylmalonyl-CoA precursor supply pathway in Streptomyces hygroscopicus ATCC 29253.* Appl Microbiol Biotechnol, 2011. **91**(5): p. 1389-97.

244. Reeves, A.R., I.A. Brikun, W.H. Cernota, B.I. Leach, M.C. Gonzalez, and J. Mark Weber, *Engineering of the methylmalonyl-CoA metabolite node of Saccharopolyspora erythraea for increased erythromycin production.* Metabolic Engineering, 2007. **9**(3): p. 293-303.
245. Reeves, A.R., I.A. Brikun, W.H. Cernota, B.I. Leach, M.C. Gonzalez, and J.M. Weber, *Effects of methylmalonyl-CoA mutase gene knockouts on erythromycin production in carbohydrate-based and oil-based fermentations of Saccharopolyspora erythraea.* Journal of Industrial Microbiology and Biotechnology, 2006. **33**(7): p. 600-609.
246. Mancia, F., N.H. Keep, A. Nakagawa, P.F. Leadlay, S. McSweeney, B. Rasmussen, P. Bosecke, O. Diat, and P.R. Evans, *How coenzyme B12 radicals are generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 A resolution.* Structure, 1996. **4**(3): p. 339-50.
247. Brooks, A.J., M. Vlasie, R. Banerjee, and T.C. Brunold, *Co−C Bond Activation in Methylmalonyl-CoA Mutase by Stabilization of the Post-homolysis Product Co2+Cobalamin.* Journal of the American Chemical Society, 2005. **127**(47): p. 16522-16528.
248. Banerjee, R., *Radical Carbon Skeleton Rearrangements: Catalysis by Coenzyme B12-Dependent Mutases.* Chemical Reviews, 2003. **103**(6): p. 2083-2094.
249. Korotkova, N. and M.E. Lidstrom, *MeaB is a component of the methylmalonyl-CoA mutase complex required for protection of the enzyme from inactivation.* J Biol Chem, 2004. **279**(14): p. 13652-8.
250. Padovani, D., T. Labunska, and R. Banerjee, *Energetics of interaction between the G-protein chaperone, MeaB, and B12-dependent methylmalonyl-CoA mutase.* J Biol Chem, 2006. **281**(26): p. 17838-44.
251. Hubbard, P.A., D. Padovani, T. Labunska, S.A. Mahlstedt, R. Banerjee, and C.L. Drennan, *Crystal structure and mutagenesis of the metallochaperone MeaB: insight into the causes of methylmalonic aciduria.* J Biol Chem, 2007. **282**(43): p. 31308-16.
252. Padovani, D. and R. Banerjee, *Assembly and Protection of the Radical Enzyme, Methylmalonyl-CoA Mutase, by Its Chaperone.* Biochemistry, 2006. **45**(30): p. 9300-9306.
253. Bobik, T.A. and M.E. Rasche, *HPLC assay for methylmalonyl-CoA epimerase.* Anal Bioanal Chem, 2003. **375**(3): p. 344-9.
254. Goldman, P. and P.R. Vagelos, *The Specificity of Triglyceride Synthesis from Diglycerides in Chicken Adipose Tissue.* Journal of Biological Chemistry, 1961. **236**(10): p. 2620-2623.
255. Shah, P.P. and E. Staple, *Synthesis of coenzyme A esters of some bile acids.* Steroids, 1968. **12**(5): p. 571-6.
256. Valenzano, C.R., Y.-O. You, A. Garg, A. Keatinge-Clay, C. Khosla, and D.E. Cane, *Stereospecificity of the Dehydratase Domain of the Erythromycin Polyketide Synthase.* Journal of the American Chemical Society, 2010. **132**(42): p. 14697-14699.
257. Peter, D.M., B. Vögeli, N.S. Cortina, and T.J. Erb, *A Chemo-Enzymatic Road Map to the Synthesis of CoA Esters.* Molecules, 2016. **21**(4): p. 517.
258. Coste, J., D. Le-Nguyen, and B. Castro, *PyBOP®: A new peptide coupling reagent devoid of toxic by-product.* Tetrahedron Letters, 1990. **31**(2): p. 205-208.

259. Нфеg-Jensen, T., M.H. Jakobsen, and A. Holm, *A new method for rapid solution synthesis of shorter peptides by use of PyBOP®.* Tetrahedron Letters, 1991. **32**(44): p. 6387-6390.
260. Al-Warhi, T.I., H.M.A. Al-Hazimi, and A. El-Faham, *Recent development in peptide coupling reagents.* Journal of Saudi Chemical Society, 2012. **16**(2): p. 97-116.
261. El-Faham, A. and F. Albericio, *Peptide Coupling Reagents, More than a Letter Soup.* Chemical Reviews, 2011. **111**(11): p. 6557-6602.
262. Kopp, F., U. Linne, M. Oberthür, and M.A. Marahiel, *Harnessing the Chemical Activation Inherent to Carrier Protein-Bound Thioesters for the Characterization of Lipopeptide Fatty Acid Tailoring Enzymes.* Journal of the American Chemical Society, 2008. **130**(8): p. 2656-2666.
263. Imker, H.J., D. Krahn, J. Clerc, M. Kaiser, and C.T. Walsh, *N-acylation during glidobactin biosynthesis by the tridomain nonribosomal peptide synthetase module GlbF.* Chem Biol, 2010. **17**(10): p. 1077-83.
264. Keilin, D. and T. Mann, *Carbonic anhydrase. Purification and nature of the enzyme.* Biochem J, 1940. **34**(8-9): p. 1163-76.
265. Hisanaga, Y., H. Ago, N. Nakagawa, K. Hamada, K. Ida, M. Yamamoto, T. Hori, Y. Arii, M. Sugahara, S. Kuramitsu, S. Yokoyama, and M. Miyano, *Structural Basis of the Substrate-specific Two-step Catalysis of Long Chain Fatty Acyl-CoA Synthetase Dimer*.* Journal of Biological Chemistry, 2004. **279**(30): p. 31717-31726.
266. Jogl, G. and L. Tong, *Crystal Structure of Yeast Acetyl-Coenzyme A Synthetase in Complex with AMP.* Biochemistry, 2004. **43**(6): p. 1425-1431.
267. Miyazawa, T., S. Takahashi, A. Kawata, S. Panthee, T. Hayashi, T. Shimizu, T. Nogawa, and H. Osada, *Identification of Middle Chain Fatty Acyl-CoA Ligase Responsible for the Biosynthesis of 2-Alkylmalonyl-CoAs for Polyketide Extender Unit *.* Journal of Biological Chemistry, 2015. **290**(45): p. 26994-27011.
268. Franke, J. and C. Hertweck, *Biomimetic Thioesters as Probes for Enzymatic Assembly Lines: Synthesis, Applications, and Challenges.* Cell Chemical Biology, 2016. **23**(10): p. 1179-1192.
269. Yamada, S., Y. Yokoyama, and T. Shioiri, *New synthesis of thiol esters.* The Journal of Organic Chemistry, 1974. **39**(22): p. 3302-3303.
270. Tse, M.L., R.E. Watts, and C. Khosla, *Substrate tolerance of module 6 of the epothilone synthetase.* Biochemistry, 2007. **46**(11): p. 3385-93.
271. Gerlt, J.A., J.T. Bouvier, D.B. Davidson, H.J. Imker, B. Sadkhin, D.R. Slater, and K.L. Whalen, *Enzyme function initiative-enzyme similarity tool (EFI-EST): a web tool for generating protein sequence similarity networks.* Biochimica Et Biophysica Acta (BBA)-Proteins and Proteomics, 2015. **1854**(8): p. 1019-1037.
272. Zallot, R., N. Oberg, and J.A. Gerlt, *Discovery of new enzymatic functions and metabolic pathways using genomic enzymology web tools.* Current Opinion in Biotechnology, 2021. **69**: p. 77-90.
273. Letunic, I. and P. Bork, *Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation.* Nucleic acids research, 2021. **49**(W1): p. W293-W296.
274. Fuller, J.Q. and P.F. Leadlay, *Proton transfer in methylmalonyl-CoA epimerase from Propionibacterium shermanii. The reaction of (2R)-methylmalonyl-CoA in tritiated water.* Biochem J, 1983. **213**(3): p. 643-50.

275. Larkin, M.J., L.A. Kulakov, and C.C.R. Allen, *Biodegradation by Members of the Genus Rhodococcus: Biochemistry, Physiology, and Genetic Adaptation*, in *Advances in Applied Microbiology*. 2006, Academic Press. p. 1-29.
276. Cornelis, K., T. Ritsema, J. Nijsse, M. Holsters, K. Goethals, and M. Jaziri, *The Plant Pathogen Rhodococcus fascians Colonizes the Exterior and Interior of the Aerial Parts of Plants.* Molecular Plant-Microbe Interactions®, 2001. **14**(5): p. 599-608.
277. Eriksson, M., J.O. Ka, and W.W. Mohn, *Effects of low temperature and freeze-thaw cycles on hydrocarbon biodegradation in Arctic tundra soil.* Appl Environ Microbiol, 2001. **67**(11): p. 5107-12.
278. Röttig, A., P. Hauschild, M.H. Madkour, A.M. Al-Ansari, N.H. Almakishah, and A. Steinbüchel, *Analysis and optimization of triacylglycerol synthesis in novel oleaginous Rhodococcus and Streptomyces strains isolated from desert soil.* Journal of Biotechnology, 2016. **225**: p. 48-56.
279. Kuyukina, M.S. and I.B. Ivshina, *Bioremediation of contaminated environments using Rhodococcus*, in *Biology of Rhodococcus*. 2019, Springer. p. 231-270.
280. Ceniceros, A., L. Dijkhuizen, M. Petrusma, and M.H. Medema, *Genome-based exploration of the specialized metabolic capacities of the genus Rhodococcus.* BMC Genomics, 2017. **18**(1): p. 593.
281. Doroghazi, J.R. and W.W. Metcalf, *Comparative genomics of actinomycetes with a focus on natural product biosynthetic genes.* BMC Genomics, 2013. **14**: p. 611.
282. Kitagawa, W. and T. Tamura, *A Quinoline Antibiotic from Rhodococcus erythropolis JCM 6824.* The Journal of Antibiotics, 2008. **61**(11): p. 680-682.
283. Kurosawa, K., I. Ghiviriga, T.G. Sambandan, P.A. Lessard, J.E. Barbara, C. Rha, and A.J. Sinskey, *Rhodostreptomycins, Antibiotics Biosynthesized Following Horizontal Gene Transfer from Streptomyces padanus to Rhodococcus fascians.* Journal of the American Chemical Society, 2008. **130**(4): p. 1126-1127.
284. Dhungana, S., R. Michalczyk, H. Boukhalfa, J.G. Lack, A.T. Koppisch, J.M. Fairlee, M.T. Johnson, C.E. Ruggiero, S.G. John, M.M. Cox, C.C. Browder, J.H. Forsythe, L.A. Vanderberg, M.P. Neu, and L.E. Hersman, *Purification and characterization of rhodobactin: a mixed ligand siderophore from Rhodococcus rhodochrous strain OFS.* BioMetals, 2007. **20**(6): p. 853-867.
285. Bosello, M., L. Robbel, U. Linne, X. Xie, and M.A. Marahiel, *Biosynthesis of the Siderophore Rhodochelin Requires the Coordinated Expression of Three Independent Gene Clusters in Rhodococcus jostii RHA1.* Journal of the American Chemical Society, 2011. **133**(12): p. 4587-4595.
286. Von Bargen, K. and A. Haas, *Molecular and infection biology of the horse pathogen Rhodococcus equi.* FEMS Microbiology Reviews, 2009. **33**(5): p. 870-891.
287. Tilford, P.E., *Fasciation of Sweet Peas Caused by Phytomonas fascians N. Sp.* 1936: p. p. 383-394.-USDA.
288. Goethals, K., D. Vereecke, M. Jaziri, M. Van Montagu, and M. Holsters, *LEAFY GALL FORMATION BY RHODOCOCCUS FASCIANS.* Annual Review of Phytopathology, 2001. **39**(1): p. 27-52.
289. Stes, E., I. Francis, I. Pertry, A. Dolzblasz, S. Depuydt, and D. Vereecke, *The leafy gall syndrome induced by Rhodococcus fascians.* FEMS Microbiology Letters, 2013. **342**(2): p. 187-194.
290. Putnam, M.L. and M.L. Miller, *Rhodococcus fascians in Herbaceous Perennials.* Plant Disease, 2007. **91**(9): p. 1064-1076.

291. Stes, E., O.M. Vandeputte, M. El Jaziri, M. Holsters, and D. Vereecke, *A successful bacterial coup d'état: how Rhodococcus fascians redirects plant development.* Annu Rev Phytopathol, 2011. **49**: p. 69-86.
292. Brown, N., *Sweet pea fasciation, a form of crowngall.* Phytopathology, 1927. **17:29–30**.
293. Goethals, K., D. Vereecke, M. Jaziri, M. Van Montagu, and M. Holsters, *Leafy gall formation by Rhodococcus fascians.* Annu Rev Phytopathol, 2001. **39**: p. 27-52.
294. Goodfellow, M., *Reclassification of Corynebacterium fascians (Tilford) Dowson in the Genus Rhodococcus, as Rhodococcus fascians comb. nov.* Systematic and Applied Microbiology, 1984. **5**(2): p. 225-229.
295. Prebble, J., *The Carotenoids of Corynebacterium fascians Strain 2 Y.* Microbiology, 1968. **52**(1): p. 15-24.
296. Maes, T., D. Vereecke, T. Ritsema, K. Cornelis, H.N.T. Thu, M. Van Montagu, M. Holsters, and K. Goethals, *The att locus of Rhodococcus fascians strain D188 is essential for full virulence on tobacco through the production of an autoregulatory compound.* Molecular Microbiology, 2001. **42**(1): p. 13-28.
297. Simón-Mateo, C., S. Depuydt, D.E.O.M. CL, F. Cnudde, M. Holsters, K. Goethals, and D. Vereecke, *The phytopathogen Rhodococcus fascians breaks apical dominance and activates axillary meristems by inducing plant genes involved in hormone metabolism.* Mol Plant Pathol, 2006. **7**(2): p. 103-12.
298. de, O.M.C.L., M. Van Montagu, E. Prinsen, K. Goethals, and M. Holsters, *De novo cortical cell division triggered by the phytopathogen Rhodococcus fascians in tobacco.* Mol Plant Microbe Interact, 2001. **14**(2): p. 189-95.
299. Depuydt, S., L. De Veylder, M. Holsters, and D. Vereecke, *Eternal youth, the fate of developing Arabidopsis leaves upon Rhodococcus fascians infection.* Plant Physiol, 2009. **149**(3): p. 1387-98.
300. Stes, E., E. Prinsen, M. Holsters, and D. Vereecke, *Plant-derived auxin plays an accessory role in symptom development upon Rhodococcus fascians infection.* The Plant Journal, 2012. **70**(3): p. 513-527.
301. Stes, E., S. Biondi, M. Holsters, and D. Vereecke, *Bacterial and plant signal integration via D3-type cyclins enhances symptom development in the Arabidopsis-Rhodococcus fascians interaction.* Plant Physiol, 2011. **156**(2): p. 712-25.
302. Depuydt, S., S. Trenkamp, A.R. Fernie, S. Elftieh, J.-P. Renou, M. Vuylsteke, M. Holsters, and D. Vereecke, *An Integrated Genomics Approach to Define Niche Establishment by Rhodococcus fascians* Plant Physiology, 2008. **149**(3): p. 1366-1386.
303. Vereecke, D., K. Cornelis, W. Temmerman, M. Jaziri, M. Van Montagu, M. Holsters, and K. Goethals, *Chromosomal locus that affects pathogenicity of Rhodococcus fascians.* J Bacteriol, 2002. **184**(4): p. 1112-20.
304. Francis, I., A. De Keyser, P. De Backer, C. Simón-Mateo, J. Kalkus, I. Pertry, W. Ardiles-Diaz, R. De Rycke, O.M. Vandeputte, and M. El Jaziri, *pFiD188, the linear virulence plasmid of Rhodococcus fascians D188.* Molecular plant-microbe interactions, 2012. **25**(5): p. 637-647.
305. Desomer, J., P. Dhaese, and M.V. Montagu, *Conjugative transfer of cadmium resistance plasmids in Rhodococcus fascians strains.* Journal of Bacteriology, 1988. **170**(5): p. 2401-2405.

306. Crespi, M., E. Messens, A.B. Caplan, M. van Montagu, and J. Desomer, *Fasciation induction by the phytopathogen Rhodococcus fascians depends upon a linear plasmid encoding a cytokinin synthase gene.* The EMBO Journal, 1992. **11**(3): p. 795-804.
307. Savory, E.A., S.L. Fuller, A.J. Weisberg, W.J. Thomas, M.I. Gordon, D.M. Stevens, A.L. Creason, M.S. Belcher, M. Serdani, M.S. Wiseman, N.J. Grünwald, M.L. Putnam, and J.H. Chang, *Evolutionary transitions between beneficial and phytopathogenic Rhodococcus challenge disease management.* eLife, 2017. **6**: p. e30925.
308. Crespi, M., D. Vereecke, W. Temmerman, M. Van Montagu, and J. Desomer, *The fas operon of Rhodococcus fascians encodes new genes required for efficient fasciation of host plants.* J Bacteriol, 1994. **176**(9): p. 2492-501.
309. Johnson, M., I. Zaretskaya, Y. Raytselis, Y. Merezhuk, S. McGinnis, and T.L. Madden, *NCBI BLAST: a better web interface.* Nucleic acids research, 2008. **36**(suppl_2): p. W5-W9.
310. Geer, L.Y., M. Domrachev, D.J. Lipman, and S.H. Bryant, *CDART: protein homology by domain architecture.* Genome research, 2002. **12**(10): p. 1619-1623.
311. Baltz, R.H., *Function of MbtH homologs in nonribosomal peptide biosynthesis and applications in secondary metabolite discovery.* Journal of Industrial Microbiology and Biotechnology, 2011. **38**(11): p. 1747-1747.
312. Bachmann, B.O. and J. Ravel, *Chapter 8. Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data.* Methods Enzymol, 2009. **458**: p. 181-217.
313. Prieto, C., C. García-Estrada, D. Lorenzana, and J.F. Martín, *NRPSsp: non-ribosomal peptide synthase substrate predictor.* Bioinformatics, 2011. **28**(3): p. 426-427.
314. Du, Y.-L., H.-Y. He, M.A. Higgins, and K.S. Ryan, *A heme-dependent enzyme forms the nitrogen–nitrogen bond in piperazate.* Nature Chemical Biology, 2017. **13**(8): p. 836-838.
315. Nachtigall, J., K. Schneider, G. Nicholson, M. Goodfellow, H. Zinecker, J.F. Imhoff, R.D. Süssmuth, and H.-P. Fiedler, *Two new aurachins from Rhodococcus sp. Acta 2259.* The Journal of Antibiotics, 2010. **63**(9): p. 567-569.
316. Temmerman, W., D. Vereecke, R. Dreesen, M. Van Montagu, M. Holsters, and K. Goethals, *Leafy gall formation is controlled by fasR, an AraC-type regulatory gene in Rhodococcus fascians.* J Bacteriol, 2000. **182**(20): p. 5832-40.
317. van Santen, J.A., E.F. Poynton, D. Iskakova, E. McMann, Tyler A. Alsup, T.N. Clark, C.H. Fergusson, D.P. Fewer, A.H. Hughes, C.A. McCadden, J. Parra, S. Soldatou, J.D. Rudolf, E.M.-L. Janssen, K.R. Duncan, and R.G. Linington, *The Natural Products Atlas 2.0: a database of microbially-derived natural products.* Nucleic Acids Research, 2021. **50**(D1): p. D1317-D1323.
318. Huang, X., E. Roemer, I. Sattler, U. Moellmann, A. Christner, and S. Grabley, *Lydiamycins A–D: Cyclodepsipetides with Antimycobacterial Properties.* Angewandte Chemie International Edition, 2006. **45**(19): p. 3067-3072.
319. Hwang, S., D. Shin, T.H. Kim, J.S. An, S.-I. Jo, J. Jang, S. Hong, J. Shin, and D.-C. Oh, *Structural Revision of Lydiamycin A by Reinvestigation of the Stereochemistry.* Organic Letters, 2020. **22**(10): p. 3855-3859.
320. Yabuuchi, T. and T. Kusumi, *Phenylglycine methyl ester, a useful tool for absolute configuration determination of various chiral carboxylic acids.* J Org Chem, 2000. **65**(2): p. 397-404.

321. Chen, B., L. Dai, H. Zhang, W.F. Tan, Z.S. Xu, and T. Ye, *Towards the stereochemical assignment of natural lydiamycin A.* Chemical Communications, 2010. **46**(4): p. 574-576.
322. Li, W., J. Gan, and D. Ma, *A Concise Route to the Proposed Structure of Lydiamycin B, an Antimycobacterial Depsipeptide.* Organic Letters, 2009. **11**(24): p. 5694-5697.
323. Libis, V., L.W. MacIntyre, R. Mehmood, L. Guerrero, M.A. Ternei, N. Antonovsky, J. Burian, Z. Wang, and S.F. Brady, *Multiplexed mobilization and expression of biosynthetic gene clusters.* Nature Communications, 2022. **13**(1): p. 5256.
324. Morgan, K.T., J. Zheng, and D.G. McCafferty, *Discovery of Six Ramoplanin Family Gene Clusters and the Lipoglycodepsipeptide Chersinamycin*.* Chembiochem, 2021. **22**(1): p. 176-185.
325. Yin, Y., B. Chen, S. Song, B. Li, X. Yang, and C. Wang, *Production of Diverse Beauveriolide Analogs in Closely Related Fungi: a Rare Case of Fungal Chemodiversity.* mSphere, 2020. **5**(5): p. e00667-20.
326. Geudens, N. and J.C. Martins, *Cyclic Lipodepsipeptides From Pseudomonas spp. – Biological Swiss-Army Knives.* Frontiers in Microbiology, 2018. **9**.
327. Micklefield, J., *Daptomycin structure and mechanism of action revealed.* Chemistry & biology, 2004. **11**(7): p. 887-888.
328. Tran, T.T., J.M. Munita, and C.A. Arias, *Mechanisms of drug resistance: daptomycin resistance.* Ann N Y Acad Sci, 2015. **1354**: p. 32-53.
329. Castaldi, S., A. Cimmino, M. Masi, and A. Evidente, *Bacterial Lipodepsipeptides and Some of Their Derivatives and Cyclic Dipeptides as Potential Agents for Biocontrol of Pathogenic Bacteria and Fungi of Agrarian Plants.* Journal of Agricultural and Food Chemistry, 2022. **70**(15): p. 4591-4598.
330. Pedras, M.S.C., N. Ismail, J.W. Quail, and S.M. Boyetchko, *Structure, chemistry, and biological activity of pseudophomins A and B, new cyclic lipodepsipeptides isolated from the biocontrol bacterium Pseudomonas fluorescens.* Phytochemistry, 2003. **62**(7): p. 1105-1114.
331. Guo, D.-L., B. Wan, S.-J. Xiao, S. Allen, Y.-C. Gu, L.-S. Ding, and Y. Zhou, *Cyclic Lipopeptides with Herbicidal and Insecticidal Activities Produced by Bacillus clausii DTM1.* Natural Product Communications, 2015. **10**(12): p. 1934578X1501001235.
332. Inès, M. and G. Dhouha, *Lipopeptide surfactants: Production, recovery and pore forming capacity.* Peptides, 2015. **71**: p. 100-112.
333. Lee, Y., C. Phat, and S.-C. Hong, *Structural diversity of marine cyclic peptides and their molecular mechanisms for anticancer, antibacterial, antifungal, and other clinical applications.* Peptides, 2017. **95**: p. 94-105.
334. Sajid, M., M.S. Ahmad Khan, S. Singh Cameotra, and A. Safar Al-Thubiani, *Biosurfactants: Potential applications as immunomodulator drugs.* Immunology Letters, 2020. **223**: p. 71-77.
335. Mohan, A., J. Padiadpu, P. Baloni, and N. Chandra, *Complete Genome Sequences of a Mycobacterium smegmatis Laboratory Strain (MC2 155) and Isoniazid-Resistant (4XR1/R2) Mutant Strains.* Genome Announc, 2015. **3**(1).
336. Ruiz, N., B. Falcone, D. Kahne, and T.J. Silhavy, *Chemical Conditionality: A GeneticStrategy to Probe Organelle Assembly.* Cell, 2005. **121**(2): p. 307-317.
337. Triccas, J.A., T. Parish, W.J. Britton, and B. Gicquel, *An inducible expression system permitting the efficient purification of a recombinant antigen from Mycobacterium smegmatis.* FEMS Microbiology Letters, 1998. **167**(2): p. 151-156.

338. Brown, A.C. and T. Parish, *Instability of the acetamide-inducible expression vector pJAM2 in Mycobacterium tuberculosis.* Plasmid, 2006. **55**(1): p. 81-86.
339. Jain, R., D. Chen, R.J. White, D.V. Patel, and Z. Yuan, *Bacterial Peptide deformylase inhibitors: a new class of antibacterial agents.* Curr Med Chem, 2005. **12**(14): p. 1607-21.
340. Ranjan, A., E. Mercier, A. Bhatt, and W. Wintermeyer, *Signal recognition particle prevents N-terminal processing of bacterial membrane proteins.* Nat. Commun., 2017. **8**: p. 15562.
341. Bogeholz, L.A.K., E. Mercier, W. Wintermeyer, and M.V. Rodnina, *Kinetic control of nascent protein biogenesis by peptide deformylase.* Sci Rep, 2021. **11**(1): p. 24457.
342. Wei, Y. and D. Pei, *Continuous spectrophotometric assay of peptide deformylase.* Anal. Biochem., 1997. **250**(1): p. 29-34.
343. Nguyen, K.T. and D. Pei, *High-throughput screening of peptide deformylase inhibitors.* Methods Mol. Med., 2008. **142**: p. 117-30.
344. Savory, E.A., A.J. Weisberg, D.M. Stevens, A.L. Creason, S.L. Fuller, E.M. Pearce, and J.H. Chang, *Phytopathogenic Rhodococcus Have Diverse Plasmids With Few Conserved Virulence Functions.* Front Microbiol, 2020. **11**: p. 1022.
345. Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.
346. Mirdita, M., K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, and M. Steinegger, *ColabFold: making protein folding accessible to all.* Nature Methods, 2022. **19**(6): p. 679-682.
347. Pertry, I., K. Václavíková, M. Gemrotová, L. Spíchal, P. Galuszka, S. Depuydt, W. Temmerman, E. Stes, A. De Keyser, M. Riefler, S. Biondi, O. Novák, T. Schmülling, M. Strnad, P. Tarkowski, M. Holsters, and D. Vereecke, *Rhodococcus fascians impacts plant development through the dynamic fas-mediated production of a cytokinin mix.* Mol Plant Microbe Interact, 2010. **23**(9): p. 1164-74.
348. Pacheco-Moreno, A., F.L. Stefanato, J.J. Ford, C. Trippel, S. Uszkoreit, L. Ferrafiat, L. Grenga, R. Dickens, N. Kelly, A.D.H. Kingdon, L. Ambrosetti, S.A. Nepogodiev, K.C. Findlay, J. Cheema, M. Trick, G. Chandra, G. Tomalin, J.G. Malone, and A.W. Truman, *Pan-genome analysis identifies intersecting roles for Pseudomonas specialized metabolites in potato pathogen inhibition.* eLife, 2021. **10**: p. e71900.
349. Rokni-Zadeh, H., W. Li, A. Sanchez-Rodriguez, D. Sinnaeve, J. Rozenski, C. Martins José, and R. De Mot, *Genetic and Functional Characterization of Cyclic Lipopeptide White-Line-Inducing Principle (WLIP) Production by Rice Rhizosphere Isolate Pseudomonas putida RW10S2.* Applied and Environmental Microbiology, 2012. **78**(14): p. 4826-4834.
350. Gerard, J., R. Lloyd, T. Barsby, P. Haden, M.T. Kelly, and R.J. Andersen, *Massetolides A−H, Antimycobacterial Cyclic Depsipeptides Produced by Two Pseudomonads Isolated from Marine Habitats.* Journal of Natural Products, 1997. **60**(3): p. 223-229.
351. Geudens, N., M.N. Nasir, J.-M. Crowet, J.M. Raaijmakers, K. Fehér, T. Coenye, J.C. Martins, L. Lins, D. Sinnaeve, and M. Deleu, *Membrane Interactions of Natural Cyclic*

Lipodepsipeptides of the Viscosin Group. Biochimica et Biophysica Acta (BBA) - Biomembranes, 2017. **1859**(3): p. 331-339.
352. De Vleeschouwer, M., T. Van Kersavond, Y. Verleysen, D. Sinnaeve, T. Coenye, J.C. Martins, and A. Madder, *Identification of the molecular determinants involved in antimicrobial activity of pseudodesmin a, a cyclic lipopeptide from the viscosin group.* Frontiers in microbiology, 2020. **11**: p. 646.
353. Shannon, P., A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker, *Cytoscape: a software environment for integrated models of biomolecular interaction networks.* Genome research, 2003. **13**(11): p. 2498-2504.
354. Miller, M.A., W. Pfeiffer, and T. Schwartz. *The CIPRES science gateway: a community resource for phylogenetic analyses*. in *Proceedings of the 2011 TeraGrid Conference: extreme digital discovery*. 2011.
355. Thompson, J.D., T.J. Gibson, and D.G. Higgins, *Multiple sequence alignment using ClustalW and ClustalX.* Current protocols in bioinformatics, 2003(1): p. 2.3. 1-2.3. 22.
356. Stamatakis, A., *RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.* Bioinformatics, 2014. **30**(9): p. 1312-1313.
357. Gilchrist, C.L. and Y.-H. Chooi, *Clinker & clustermap. js: Automatic generation of gene cluster comparison figures.* Bioinformatics, 2021. **37**(16): p. 2473-2475.
358. Kumar, S., G. Stecher, M. Li, C. Knyaz, and K. Tamura, *MEGA X: molecular evolutionary genetics analysis across computing platforms.* Molecular biology and evolution, 2018. **35**(6): p. 1547.

Chapter 8: Appendix

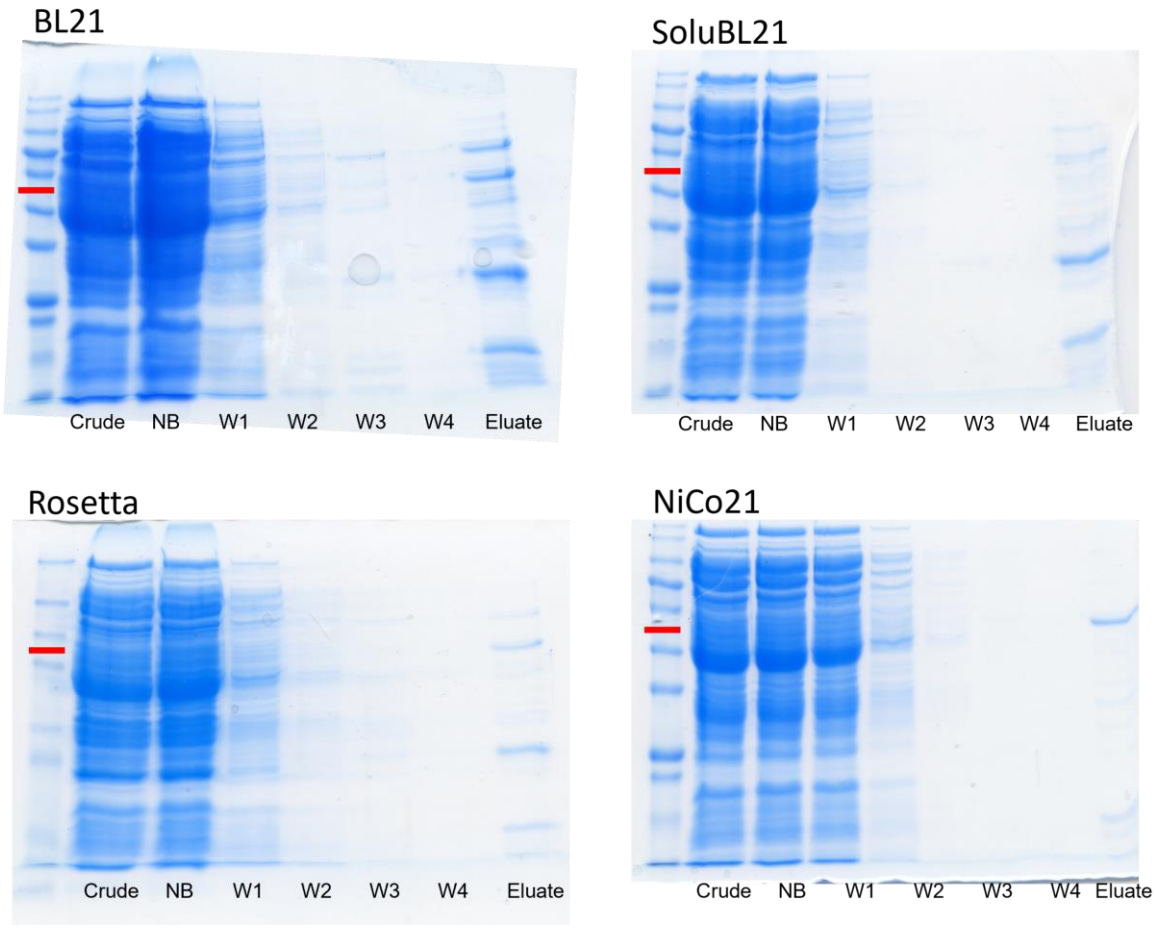


Figure 95. SDS-PAGE comparison of ActD protein purification using different *E. coli* expression strains. Expected size of protein indicated by red line. Lanes are the clarified lysate (crude), non-binding column fraction (NB), washes with increasing imidazole concentration (W1-W4) and the final eluted purified protein (eluate).

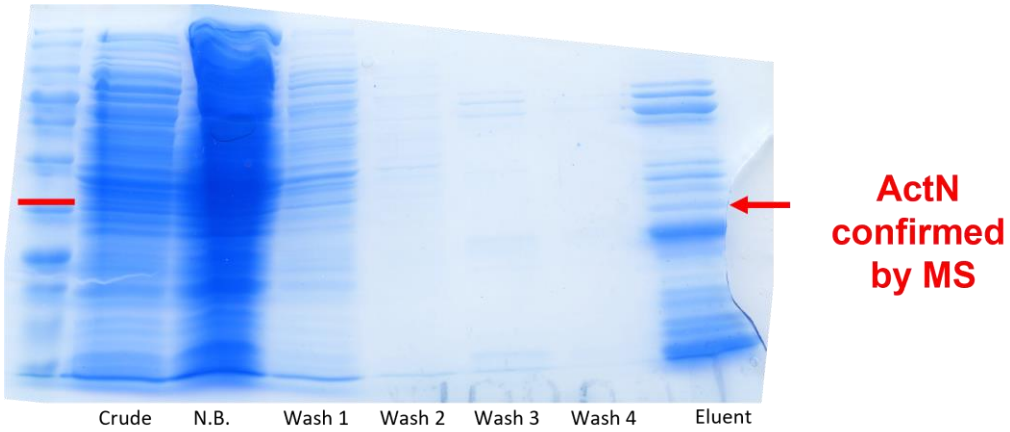


Figure 96. ActN (actinonin MeaB) protein purification attempt. Used *E. coli* BL21(DE3) expression strain and nickel affinity chromatography. Abbreviations: N.B. = non-binding.

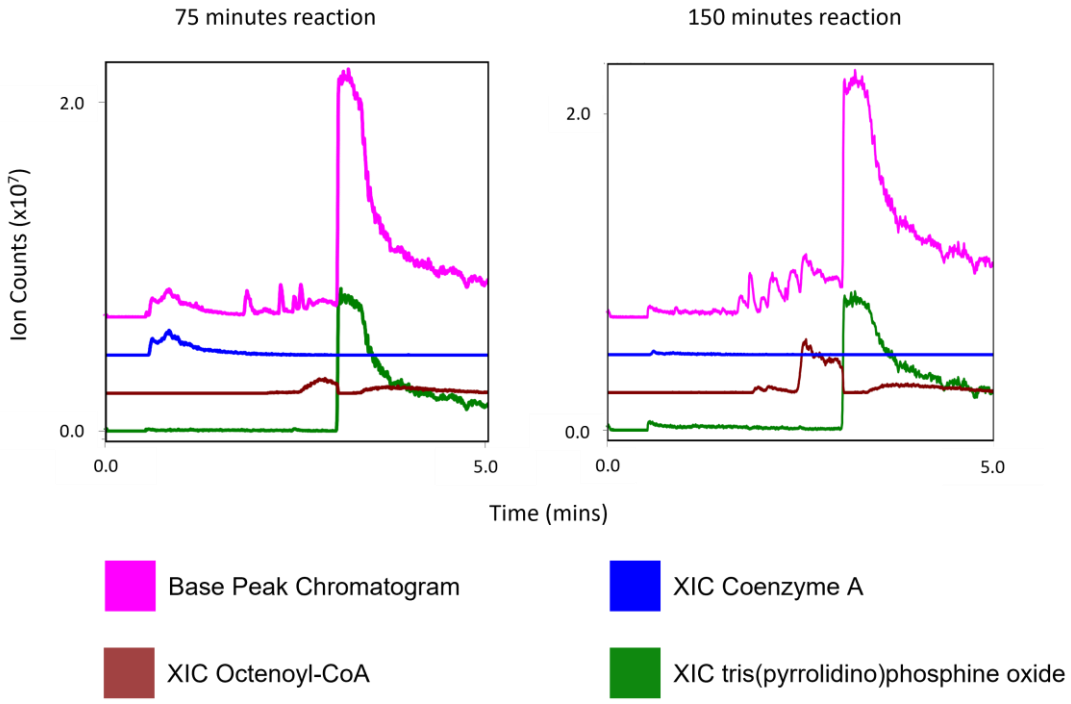


Figure 97. LC-MS analysis of Octenoyl-CoA synthesis using PyBOP method. Sampled at 75 minutes (left panel) and upon completion (right panel) of 150 minute reaction. The base peak chromatogram is plotted in purple. The extracted ion chromatogram of Coenzyme A (*m/z* 768.123; $[M+H]^+$) is plotted in blue, Octenoyl-CoA (*m/z* 892.211; $[M+H]^+$) is plotted in brown and the PyBOP breakdown product, tris(pyrrolidino)phosphine oxide, (*m/z* 258.173; $[M+H]^+$) is plotted in green.

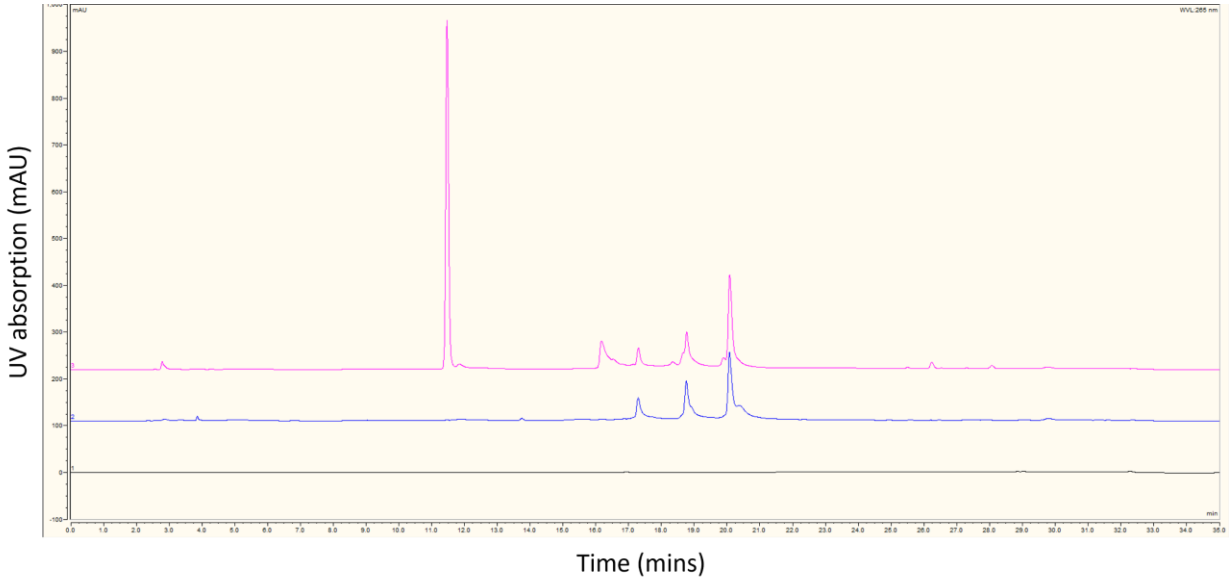


Figure 98. Comparison of octenoyl-CoA purification methods. C18 reverse phase separation of PyBOP octenoyl-CoA crude reaction mixture. Absorption measured at 265 nm. Separation of crude reaction mixture by C18 alone is shown in pink. Separation of crude reaction mixture by SAX followed by C18 is shown in blue. Blank is shown in black. Subsequent LC-MS analysis indicated that the 18.5 minute peak contains octenoyl-CoA.

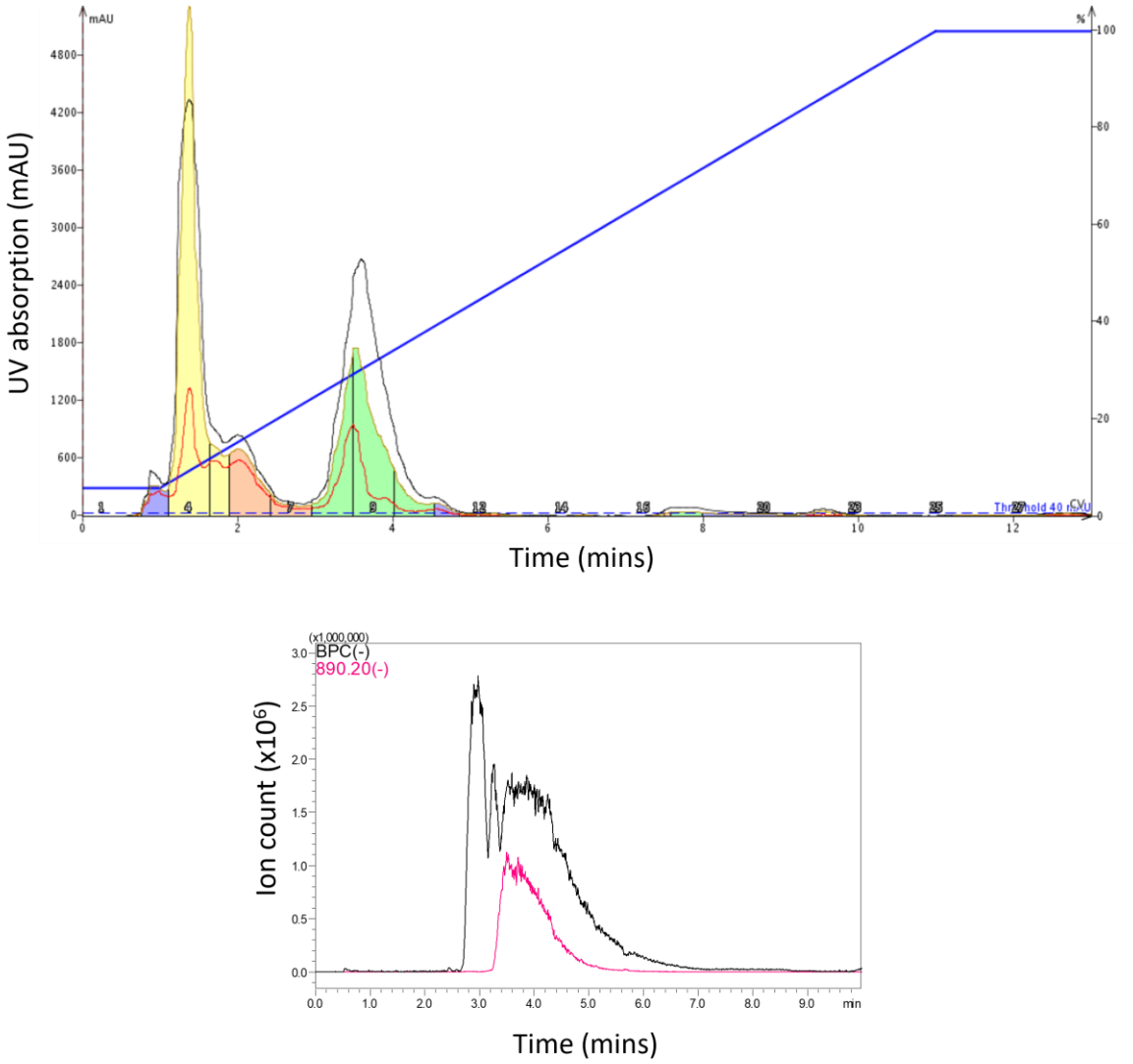


Figure 99. Separation of PyBOP octenoyl-CoA crude reaction mixture by reverse phase flash chromatography. Bottom is the LC-MS analysis of fractions 8/9 with extracted ion chromatogram of octenoyl-CoA in pink and base peak chromatogram in black.

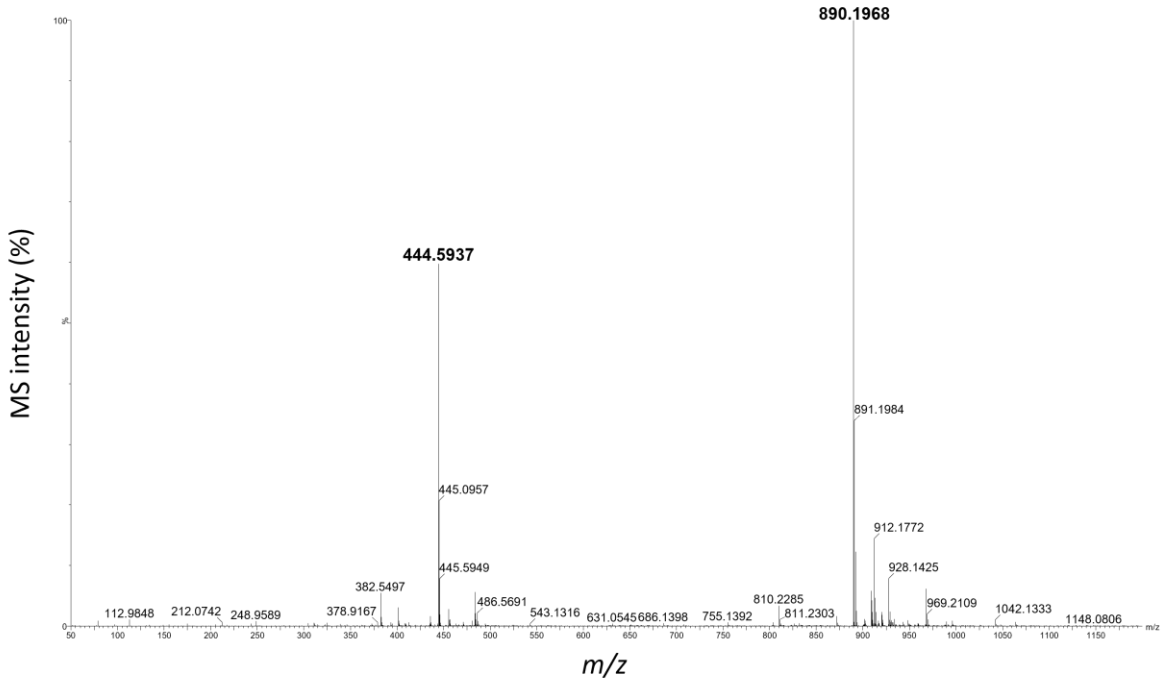


Figure 100. HR-MS analysis of preparative-scale HPLC octenoyl-CoA sample.

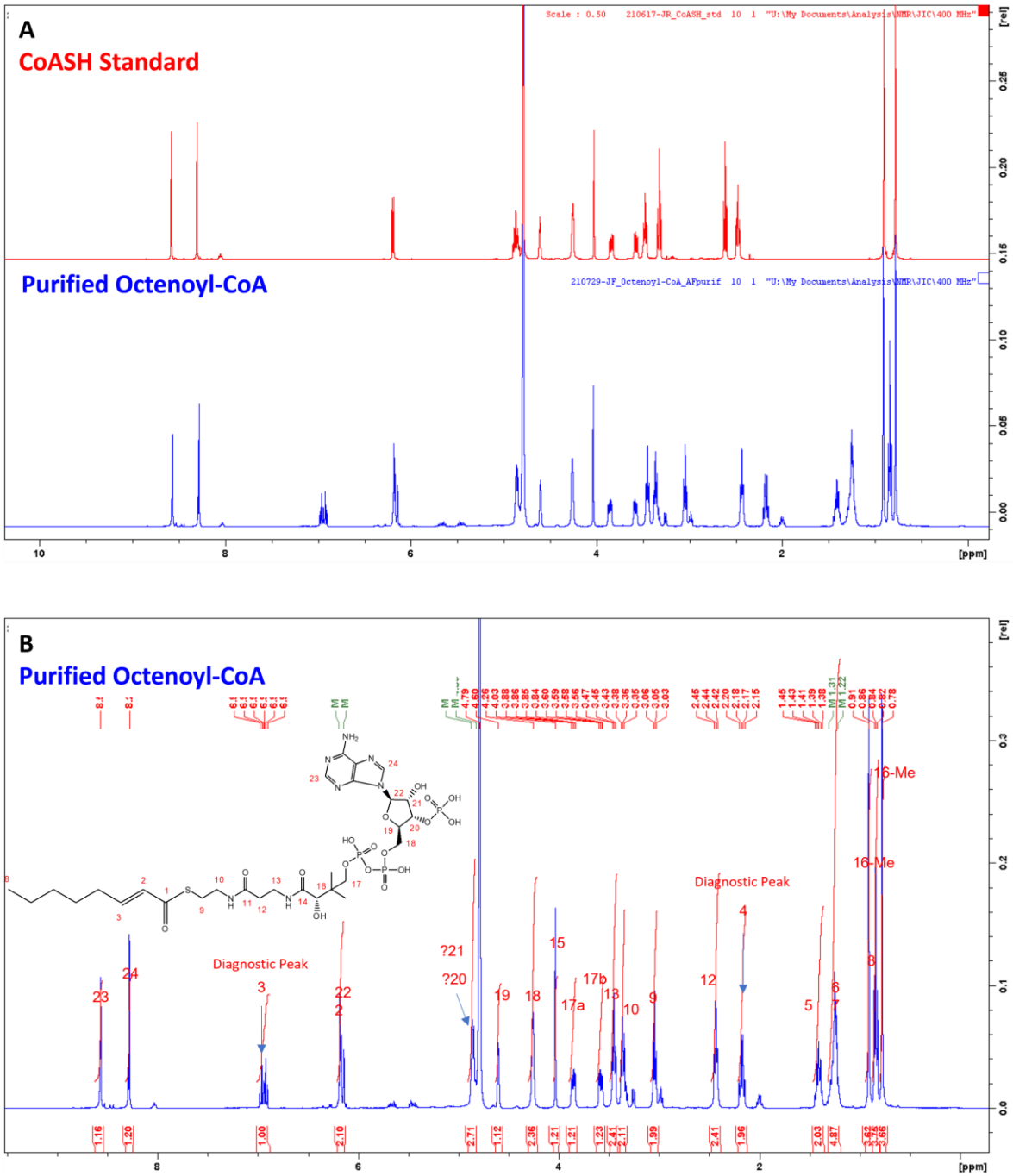


Figure 101. ^{1}H NMR analysis of purified octenoyl-CoA. Performed by Dr. Martin Rejzek. **A.** Comparison of purified octenoyl-CoA to a standard of free coenzyme A. **B.** Assignment of ^{1}H spectrum to the structure of octenoyl-CoA.

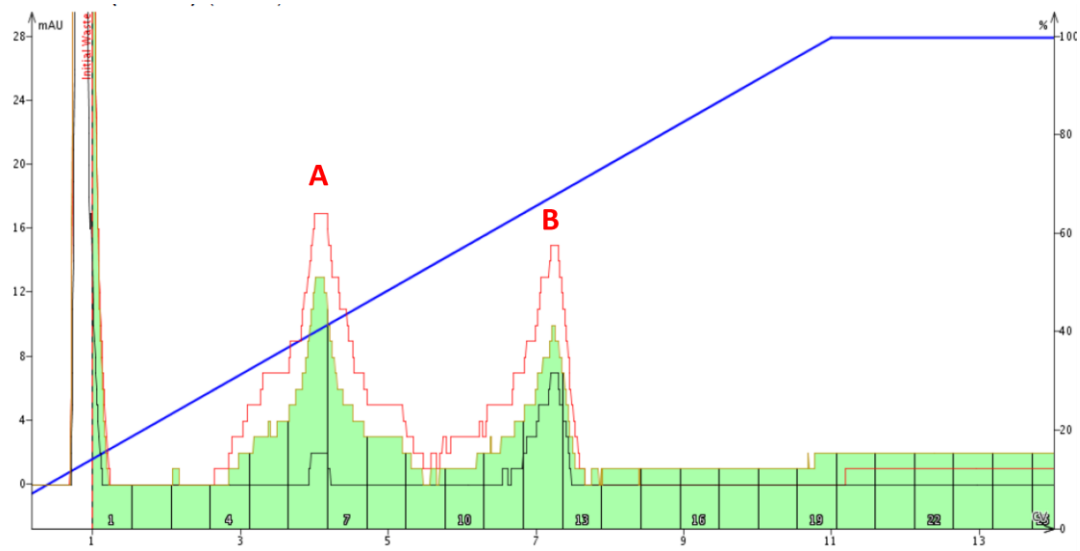


Figure 102. Overview of flash chromatography purification of compound #1 from R. fascians SM12 culture organic extract.

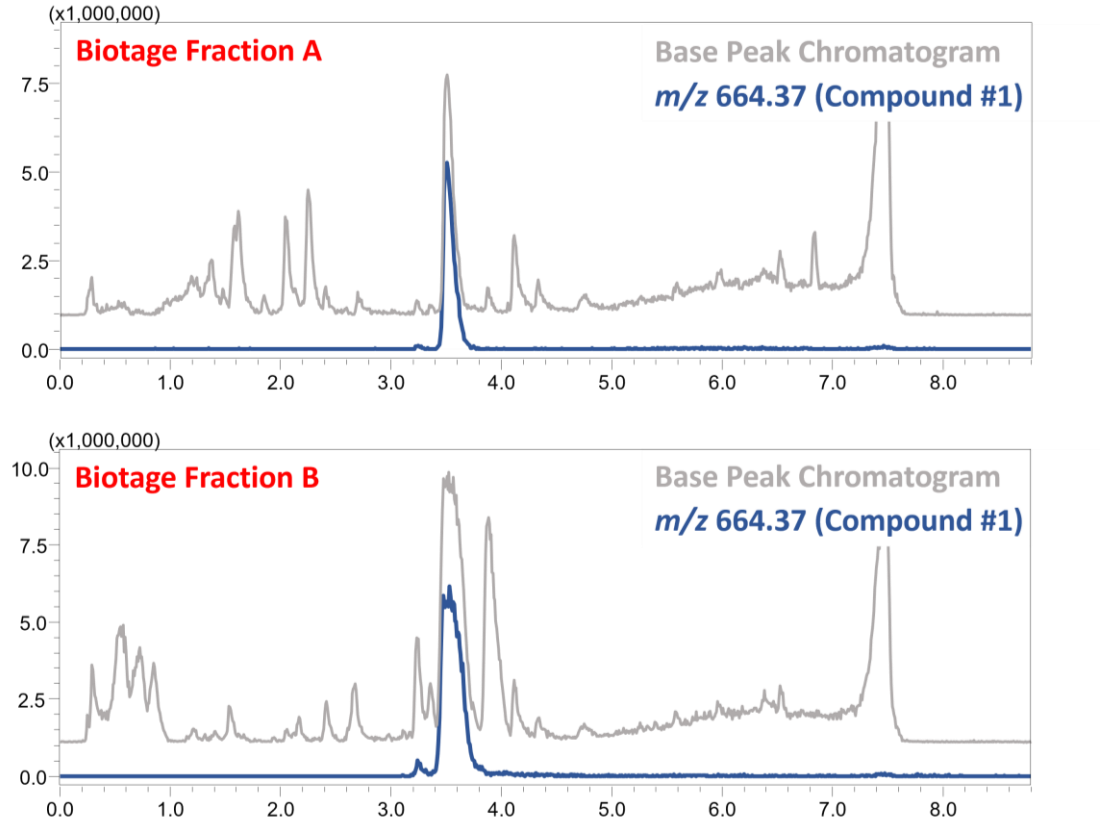


Figure 103. LC-MS analysis of biotage fractions A and B.

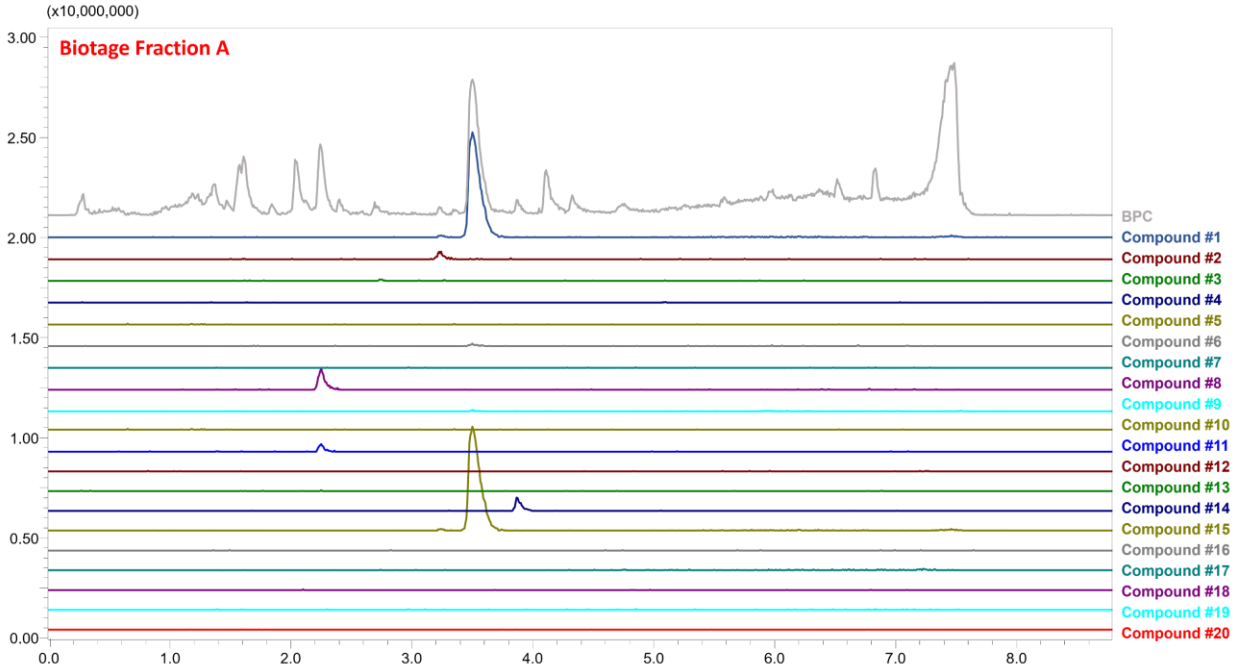


Figure 104. LC-MS analysis of the presence of the twenty candidate D188nrp products in biotage fraction A.

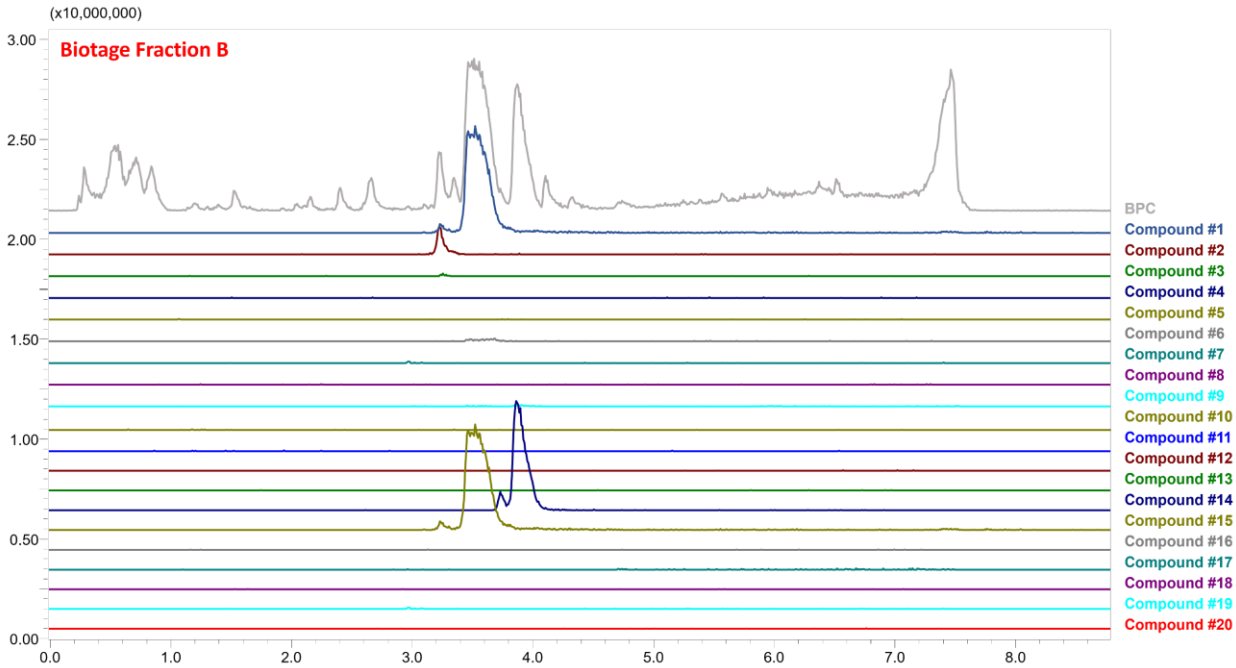


Figure 105. LC-MS analysis of the presence of the twenty candidate D188nrp products in biotage fraction B.

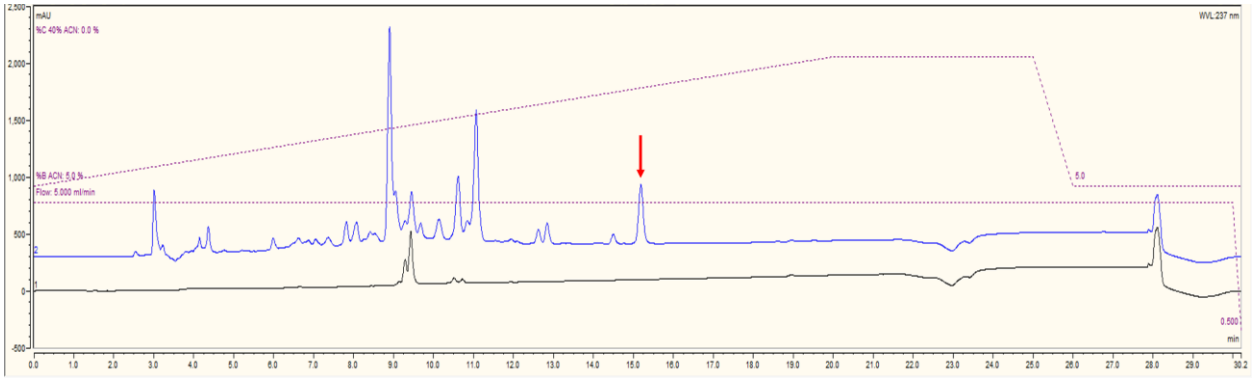


Figure 106. HPLC separation of biotage fraction A. UV absorbance measured at 237 nm. Blank run in black, sample injection in blue. Red arrow indicates peak of compound #1, as determined by LC-MS.

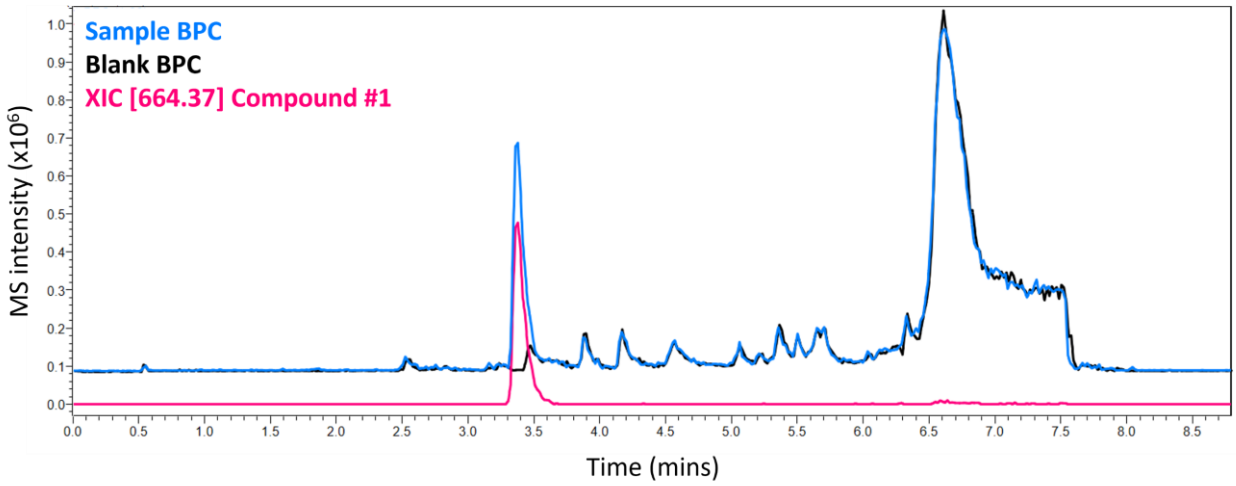


Figure 107. LC-MS analysis of semi-preparative scale HPLC purification of compound #1 from biotage fraction A. The blank BPC is indicated as the column was dirty in this run.

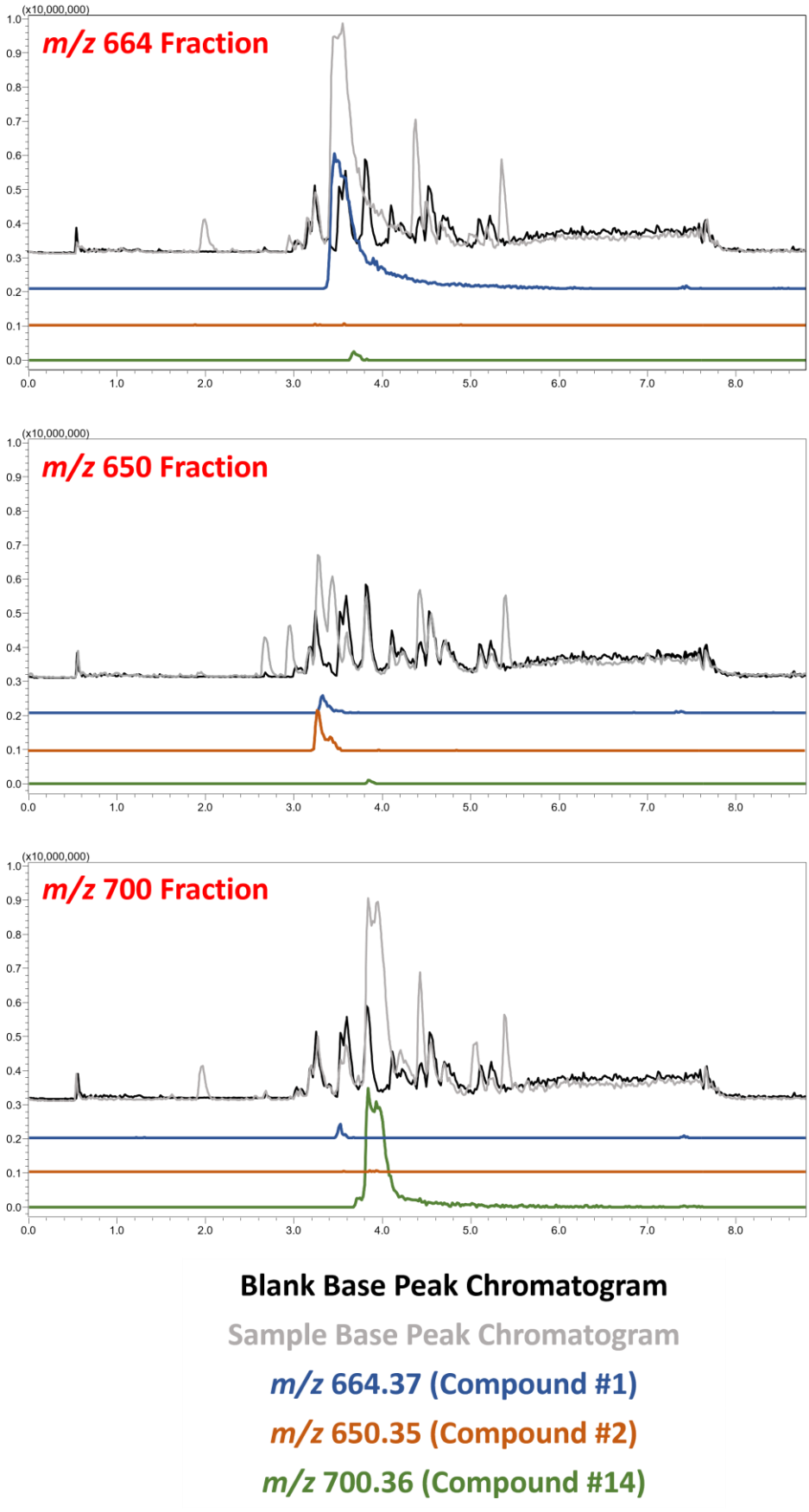


Figure 108. LC-MS analysis of the mass-guided separation of biotage fraction B.

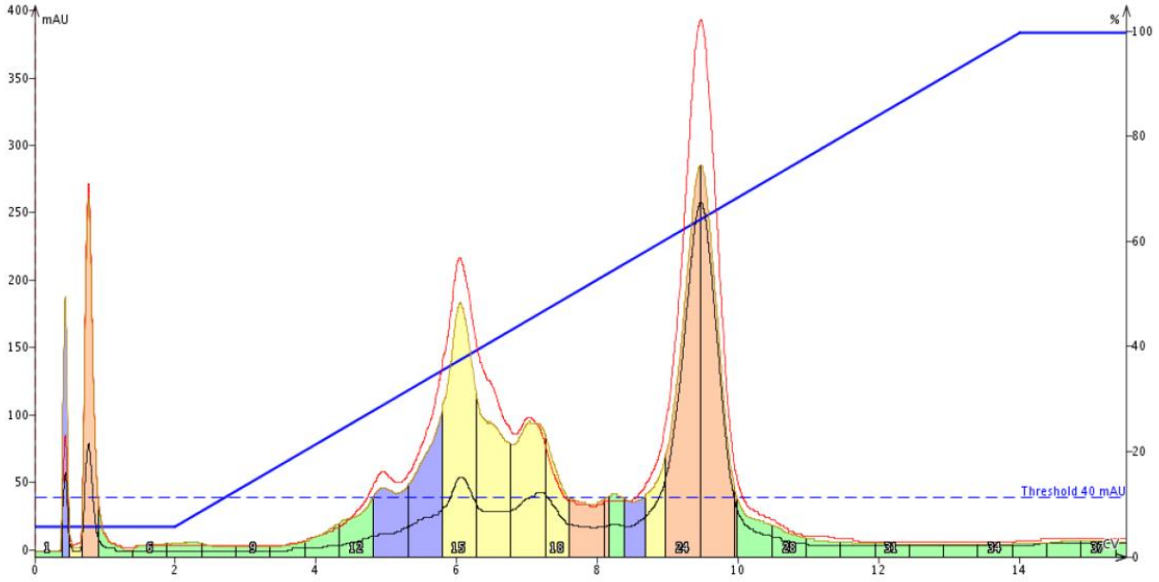


Figure 109. Reverse phase flash chromatography separation of the 2L organic extract for isolation of compound #1.

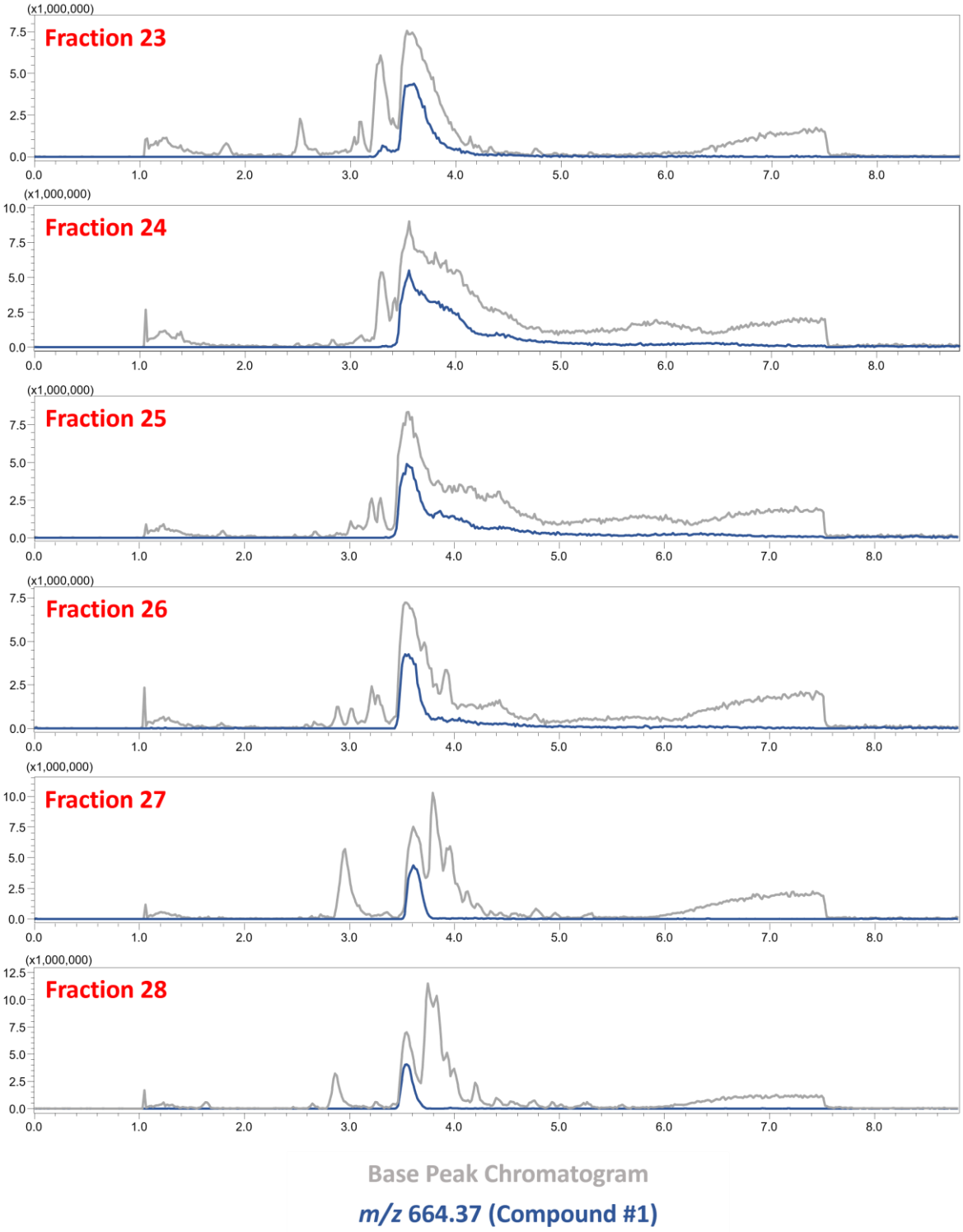


Figure 110. LC-MS analysis of large scale compound #1 flash chromatography fractions.

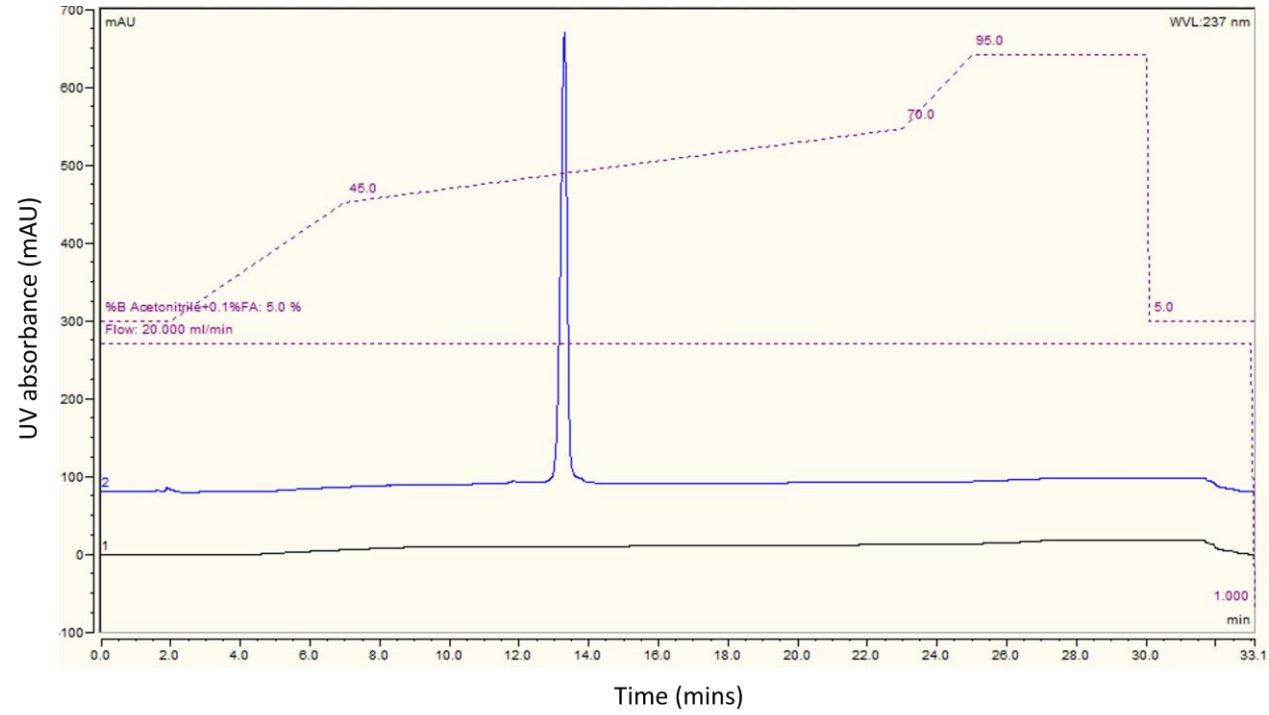


Figure 111. Preparative-scale HPLC purification of compound #1 flash chromatoraphy fractions 24-26.

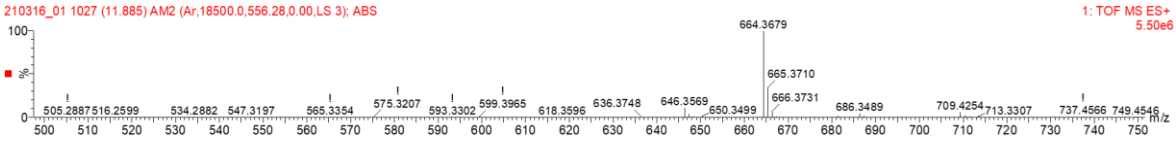


Figure 112. HR-MS analysis of compound #1. Predicted Lydiamycin A $[M+H]^+$ = 664.3665.

Table 25. Proton NMR spectra comparison between published lydiamycin A and observed compound #1 data. All NMR recorded in $CDCl_3$.

	REPORTED [319] δH (J IN HZ)	REPORTED [318] δH (J IN HZ)	OBSERVED δH (J IN HZ)
1			
2	5.30, m	5.28, m	5.31, m
3	2.38, 1.77, m	2.35, 1.77, m	2.38, 1.78 m
4	1.63, 1.59, m	1.60, m	1.63, 1.54, m
5	3.14, m, 2.77, qd (12.0, 3.0)	3.12, br d (13.5), 2.75, br dq (13.5, 3.8)	3.15, br d (14.2), 2.78, qd (13.0, 3.5)
5-NH	4.32, d (12.0)	4.31, br d	4.33, br d
6			
7	4.85, qd (10.5, 7.0)	4.83, qd (10.5, 7.25)	4.86, qd (10.5 ,7.3)
7-NH	7.30, d (10.5)	7.27. d (10.5)	7.30, d (10.3)
8	1.44, d (7.0)	1.43, d (7.25)	1.46, d (7.3)
9			
10	4.35, td (8.5, 8.5)	4.33, m	4.35, m
10-NH	7.46, d (8.5)	7.43, d (7.8)	7.48, d (9.0)
11	1.78, m	1.77, m	1.78, m
12	1.65, m	1.63, m	1.63, m
13	0.93, d (6.5)	0.93, d (6.5)	0.94, d (6.6)
14	0.88, d (6.5)	0.87, d (7.8)	0.89, d (6.7)
15			
16	5.31, m	5.28, m	5.31, m
16-NH	8.30, d (10.0)	8.25, d (9.85)	8.31, d (9.5)
17	4.77, dd (11.5, 4.5), 4.04, d (11.5)	4.75, dd (11.4, 4.4), 4.03, br d (11.4)	4.78, dd (11.4, 4.4), 4.05, d (11.4)
18			
19	4.67, t (5.0)	4.65, br t (5.7)	4.70, br t (5.0)
20	2.35, 1.92, m	2.35, 1.93, m	2.38, 1.92 m
21	2.17, 1.95 m	2.15, 1.92, m	2.18, 1.95 m
22	6.92, m	6.90, br s	6.93, m
23			
24	3.32, dd (17.0, 12.0), 2.62, dd (17.0, 5.5)	3.30, dd (16.8, 12.2), 2.60, dd (16.8, 5.3)	3.33, dd (18.0, 13.2) 2.63, dd (16.9, 5.2)
25	3.04, m	3.02, m	3.05, m
26	1.72, 1.53, m	1.71, 1.52, m	1.78, 1.54, m
27	1.35, 1.25, m	1.35, 1.24, m	1.36, 1.26, m
28	1.29, 1.25, m	1.27-1.26, m	1.28, 1.26, m
29	1.27, m	1.27-1.26, m	1.26, m
30	0.86, t (7.0)	0.85, t (6.5)	0.87, t (6.9)
31			
31-OH	11.37, br s		11.40, br s

Table 26. Carbon NMR spectra comparison between published lydiamycin A and observed compound #1 data. All NMR recorded in $CDCl_3$.

	REPORTED [319] δC, MULT	REPORTED [318] δC, MULT	OBSERVED δC
1	169.8, qC	169.1, qC	169.1
2	52.7, CH	52.7, CH	52.7
3	24.4, CH2	24.4, CH2	24.4
4	21.4, CH2	21.4, CH2	21.5
5	47.1 CH2	47.1, CH2	47.1
5-NH			
6	174.3, qC	174.3, qC	174.4
7	50.5, CH	50.5, CH	50.5
7-NH			
8	18.3	18.2	18.3
9	169.0, qC	169.3, qC	169.4
10	51.4, CH	51.4, CH	51.5
10-NH			
11	36.4, CH2	36.6, CH2	36.6
12	24.6, CH	24.6, CH	24.6
13	22.8, CH3	22.7, CH3	22.8
14	22.3, CH3	22.2, CH3	22.3
15	169.3, qC	169.7, qC	169.8
16	50.6, CH	50.7, CH	50.7
16-NH			
17	68.3, CH2	68.2, CH2	68.2
18	171.0, qC	170.9, qC	171.0
19	55.6, CH	55.6, CH	55.6
20	19.4, CH2	19.3, CH2	19.4
21	20.6, CH2	20.6, CH2	20.7
22	142.7, CH	142.7, CH	142.7
23	175.9, qC	175.9, qC	176.0
24	36.6, CH2	36.5, CH2	36.5
25	46.0, CH	46.0, CH	46.0
26	31.6, CH2	31.6, CH2	31.7
27	26.9, CH2	26.8, CH2	26.9
28	31.5, CH2	31.5, CH2	31.5
29	22.4, CH2	22.4, CH2	22.4
30	14.0, CH3	13.9, CH3	14.0
31	177.3, qC	177.3 qC	177.4
31-OH			

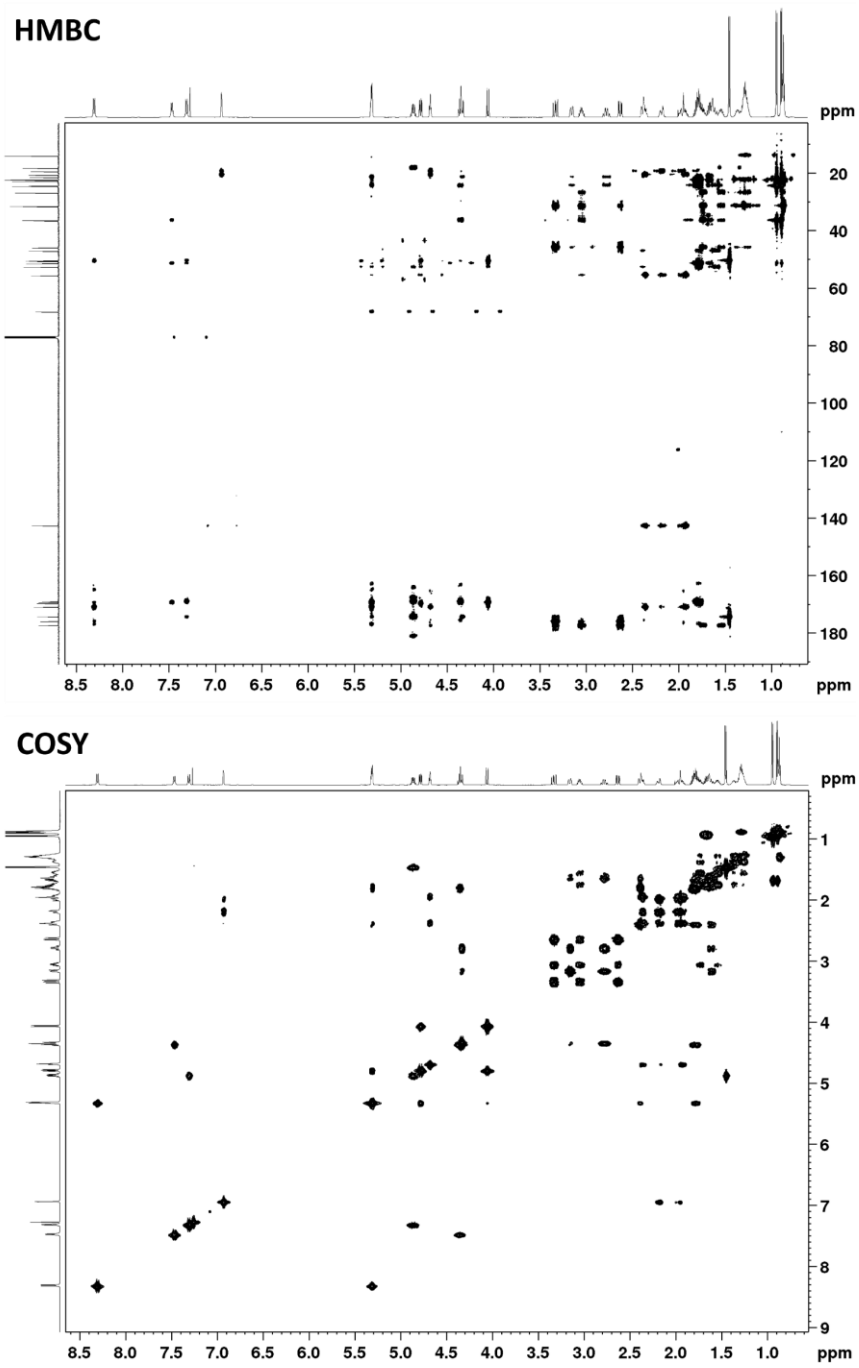


Figure 113. HMBC and COSY NMR spectra of compound #1. Measured in $CDCl_3$.

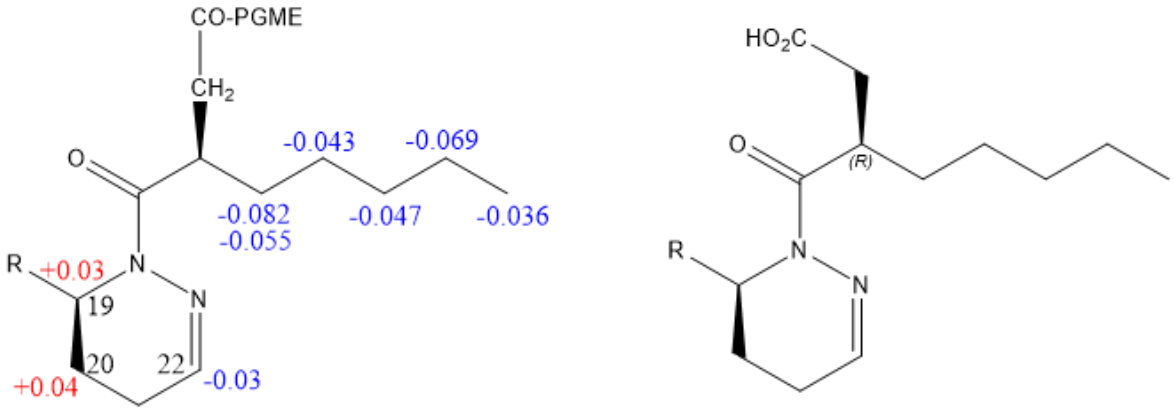


Figure 114. Stereochemistry analysis of C24 using existing PGME derivatisation data. Analysis and figure prepared by Edward Hems.

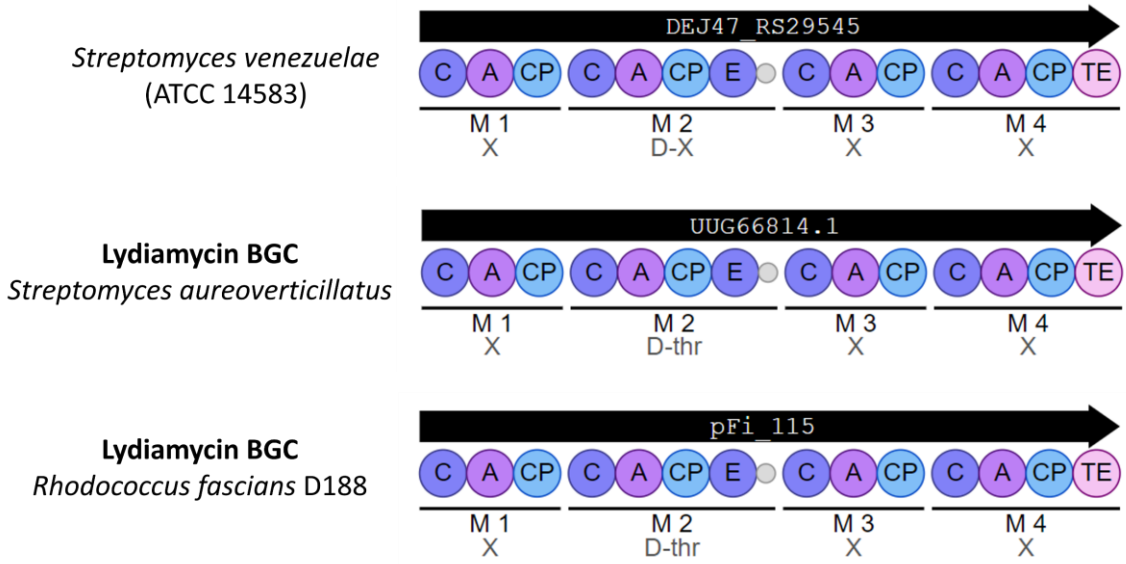


Figure 115. Comparison of NRPS domain architecture of known lydiamycin producers and *S. venezuelae* cluster. Produced using antiSMASH and NRPSpredictor2.

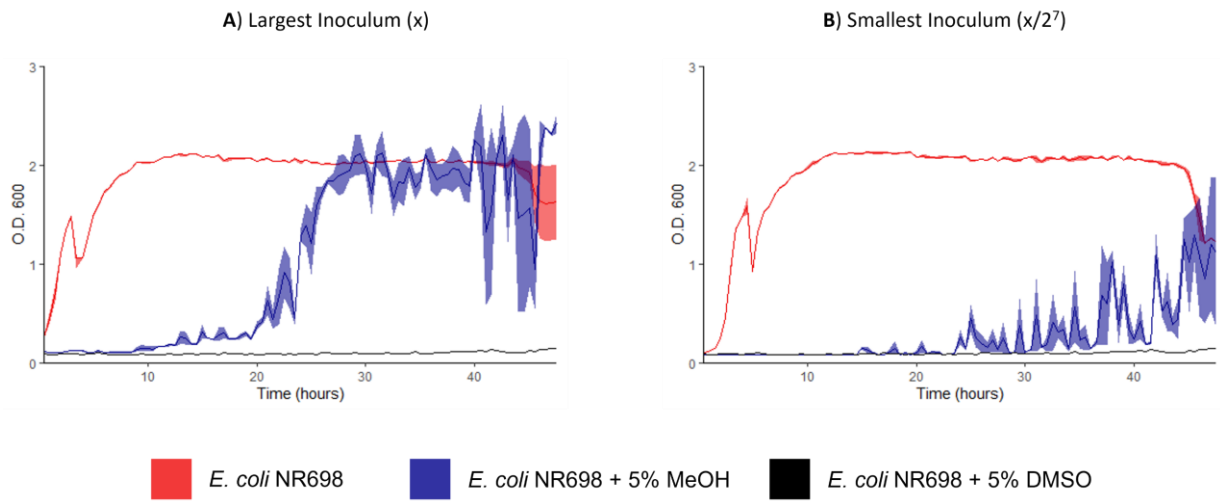


Figure 116. Comparison of inoculum concentration and presence of solvent on *E. coli* NR698 growth. Where x = O.D. 0.05. Reporter strain and reporter strain plus MeOH is duplicate data (thick line represents average value and shaded envelope represents minimum and maximum values). DMSO value is single (not replicated).

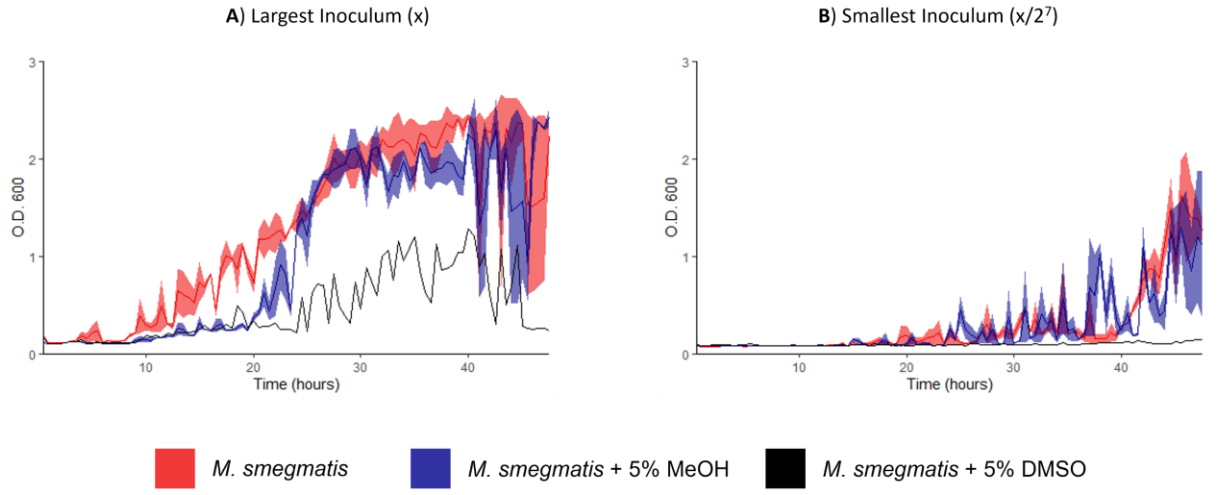


Figure 117. Comparison of inoculum concentration and presence of solvent on *M. smegmatis* growth. Where x = O.D. 0.05. Reporter strain and reporter strain plus MeOH is duplicate data (thick line represents average value and shaded envelope represents minimum and maximum values). DMSO value is single (not replicated).

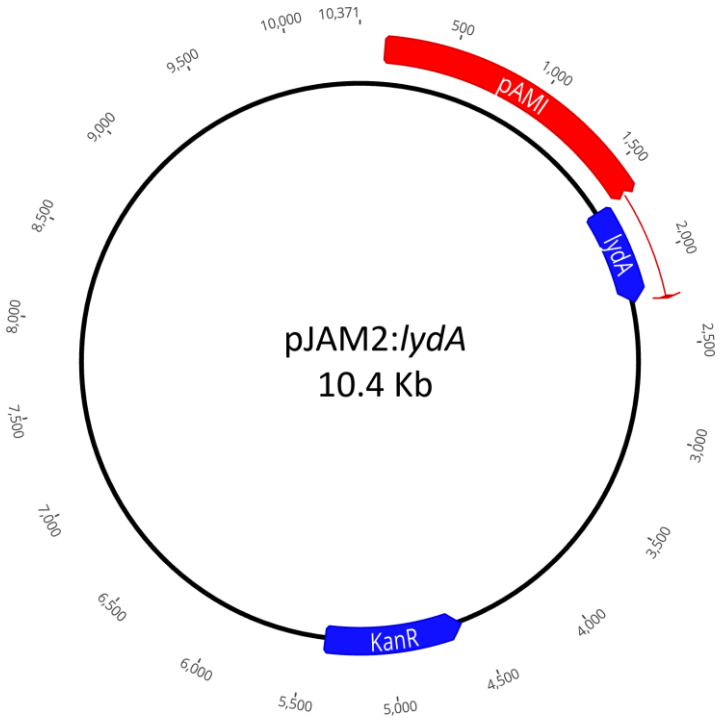


Figure 118. Plasmid map of pJAM2:lydA.

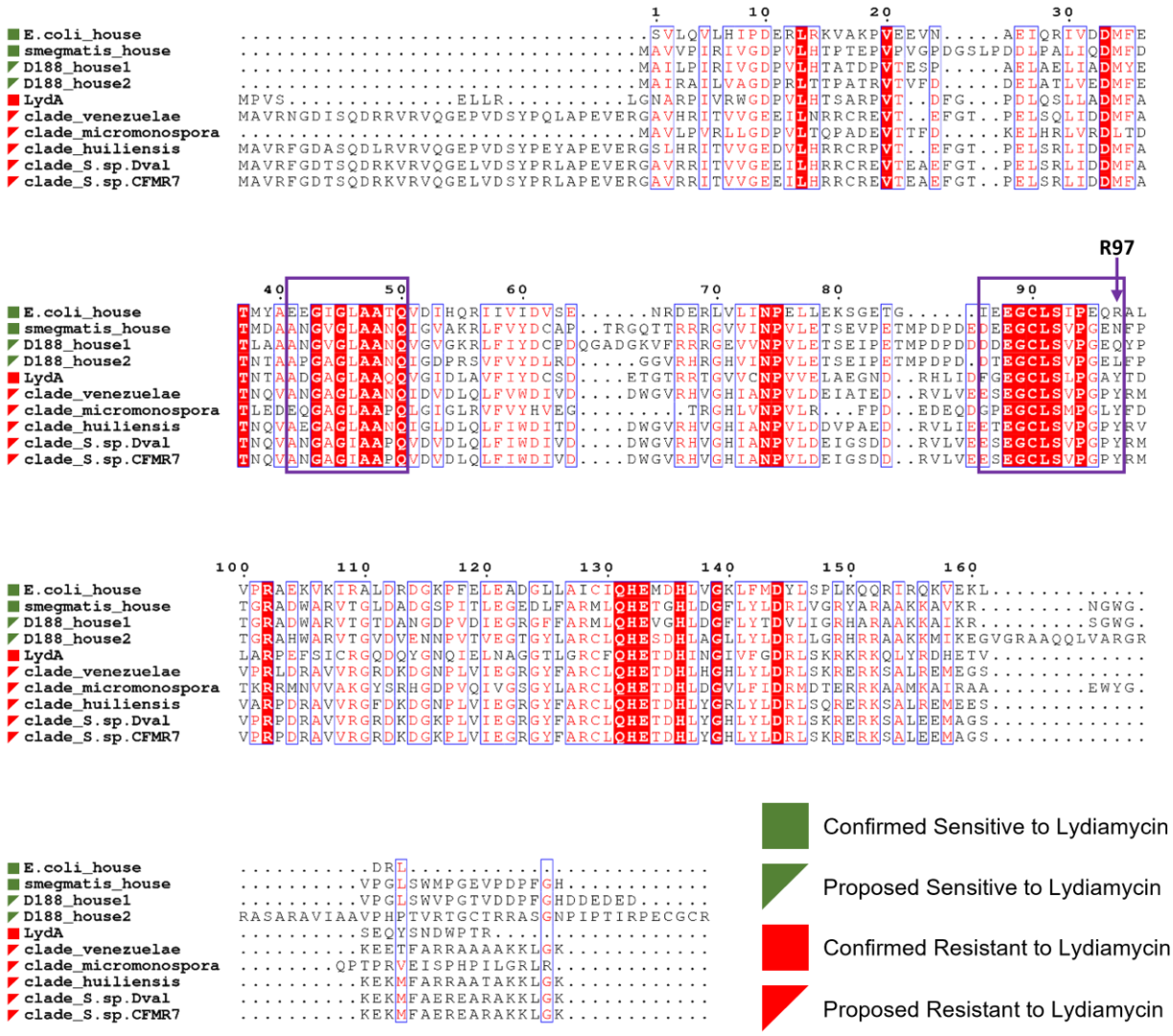


Figure 119. Alignment of PDF genes for prediction of lydiamycin immunity determinant. The sensitivity or resistance against lydiamycin is indicated. The purple boxes indicate the residues which are located on the surface of the protein in or around the active site. The residues are numbered according to the *E. coli* PDF gene. A possible resistance determinant residue (*E. coli* R97) is indicated.

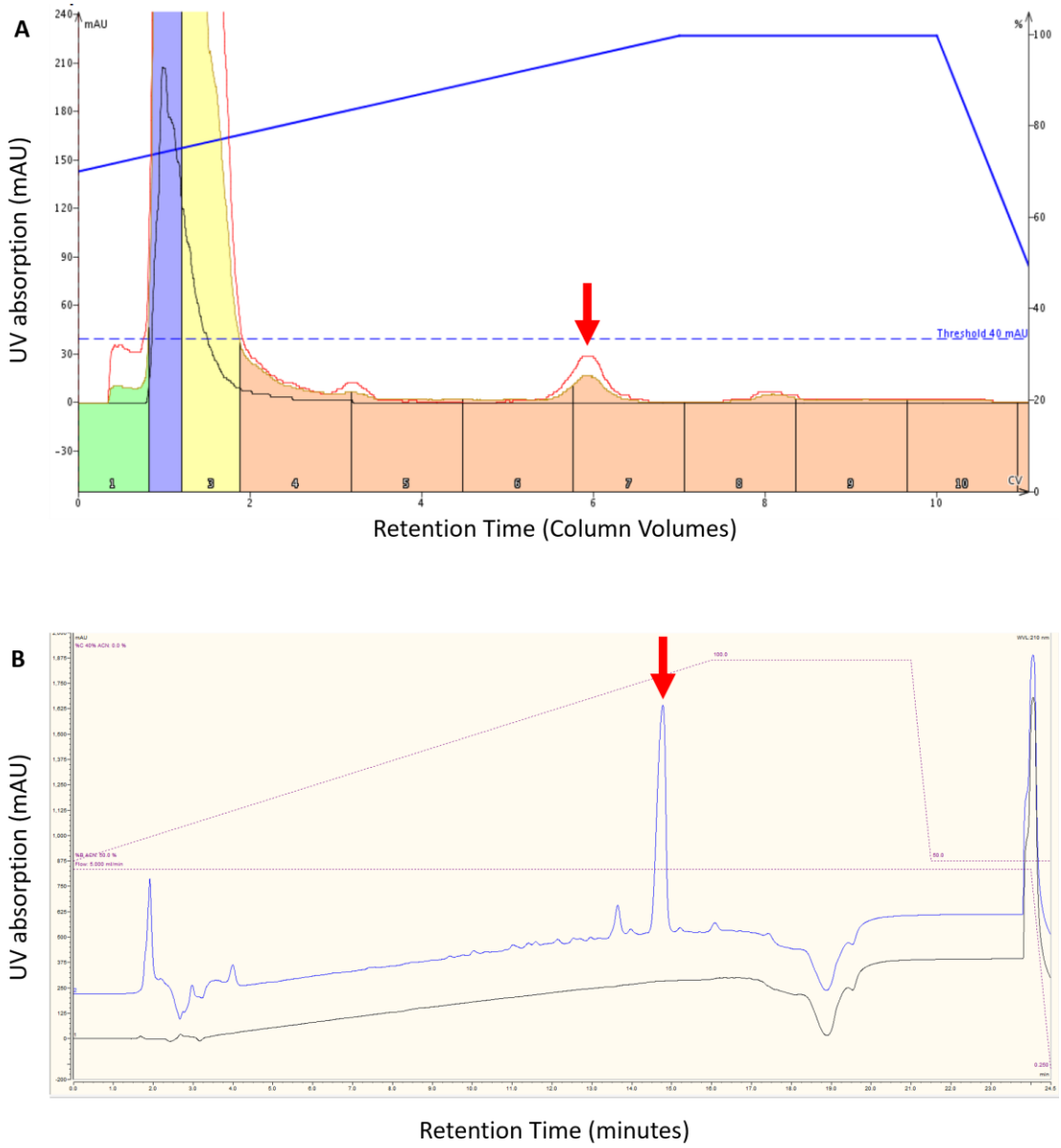


Figure 120. Overview of 682CLP purification. **A.** Fractionation of *Pseudomonas* Ps682 organic extract by flash chromatography. UV absorbance measured at 210 nm. Blue line indicates gradient of MeOH/H_2O. Each peak was analysed using LC/MS and 682CLP was identified in peak labelled with red arrow. **B.** Fractionation by semi-preparative HPLC. Blue trace is sample injection, black trace is blank injection. Purple line indicates gradient of MeOH/H_2O. UV absorbance measured at 210 nm. Each peak was analysed using LC/MS and 682CLP was identified in peak labelled with red arrow.

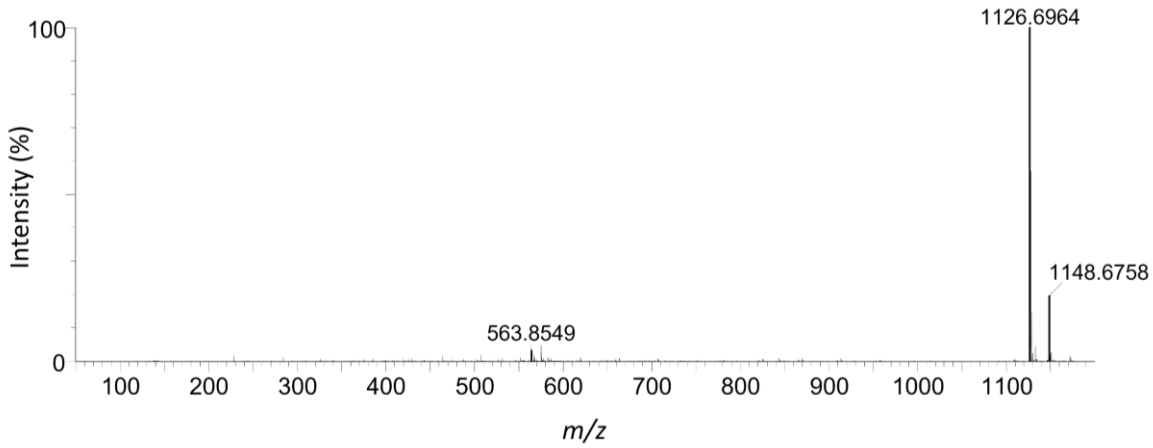


Figure 121. HR-MS/MS analysis of 682CLP. Predicted $[M+H]^+$ = 1126.6970