A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

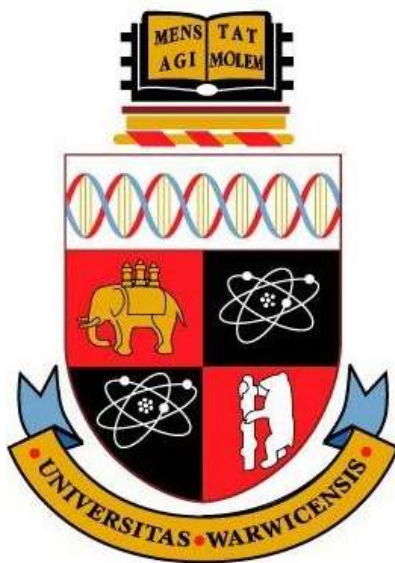http://wrap.warwick.ac.uk/147032

warwick.ac.uk/lib-publications

Chemical probing of polyketide bio-assemblies

Robert D. Jenkins

Supervisor: Dr Manuela Tosin

Thesis submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy in Chemistry

Department of Chemistry

University of

Warwick

September 2019

<u>Contents</u>

Declaration

The experimental work reported in this thesis is original research carried out by the author, unless otherwise stated, in the Department of Chemistry, University of Warwick, between October 2015 and December 2018. No material has been submitted for any other degree, or at any other institution.

Results from other authors are referenced in the usual manner throughout the text.
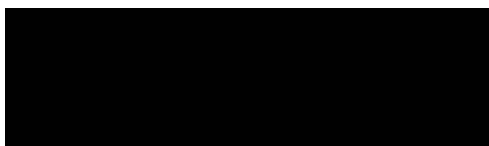
Date: 23/09/2019

Robert D. Jenkins

Acknowledgements

This thesis is written in memory of my grandmother, who sadly passed away shortly before I began my PhD. She was an extremely stubborn, witty and determined lady, with a strong work ethic, who I greatly admired and was inspired by. I can only imagine how delighted she would have been to have glanced at this work, although, I am certain she would not have understood a word of it.

Marian Alice "Bunny" Irwin-Singer (1917-2015)

List of Figures

List of Schemes

List of Tables

Abbreviations and Definitions

Abbreviations	Definitions
5'-ClDA	5'-chlorodeoxyadenosine
5'-FDA	5'-fluorodeoxyadenosine
6-DEB	6-deoxyerythronolide B
6-DEBS	6-deoxyerythronolide B synthase
6-MSA	6-methylsalicylic acid
6-MSAS	6-methylsalicylic acid synthase
Ac	Acetyl
Ac_2O	Acetic anhydride
AcOH	Acetic acid
ACP	Acyl carrier protein
Act	Actinorhodin
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
Ar	Argon gas
ARO	Aromatase
AT	Acyl transferase
ATP	Adenosine triphosphate
AU	Absorbance units
BGC	Biosynthetic gene cluster
BLAST	Basic local alignment search tool
Bor	Borrelidin
Br_2	Elemental bromine
C	Condensation domain
CAD or CID	Collisionally Activated/Induced Dissociation
$CDCl_3$	Deuterated chloroform
CDI	Carbon diimidazole
cDNA	Complementary DNA
CE	Collision energy
CH	Cyclohexane
$CHCl_3$	Chloroform
CLF	Chain length factor
CoA	Coenzyme A
COSY	Correlation spectroscopy

Abbreviations	Definitions
CpTiCl$_3$	Cyclopentenyl titanium trichloride
CuI	Copper iodide
Cy	Cyclisation domain
CYC	Cyclase
CYP450	Cytochrome P450
Cys	Cysteine
D	Deuterium
D$_2$O	Deuterated water
D6-DMSO	Deuterated dimethyl sulphoxide
Da	Daltons
DCC	Dicyclohexyl carbodiimide
DCM	Dichloromethane
DEB	Deoxyeryronolide B
DEBS	Deoxyeryronolide B synthase
DH	Dehydratase
DI	Deionised
DIBAL	Diisobutyl aluminium hydride
DIPEA	N,N-Diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	Dimethyl formamide
DMNB	4,5-dimethoxy-2-nitrobenzyl
DMP	Dess-Martin periodinane
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DPCK	Dephosphocoenzyme A kinase
DPPA	Diphenyl phosphoryl azide
E	Epimerisation domain
E. coli	*Escherichia coli*
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDC.HCl	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
EDTA	Ethylenediaminetetraacetic acid
EIC	Extracted ion chromatogram
ER	Enoyl reductase
Ery	Erythromycin
ESI	Electrospray ionisation

Abbreviations	Definitions
Et$_2$O	Diethyl ether
Et$_3$N	Triethylamine
EtOAc	Ethyl acetate
EtOH	Ethanol
FA	Formic acid
Fab	Fatty acid biosynthesis
FAD	Flavin adenine dinucleoside
FAS	Fatty acid synthase
Fig	Figure
FT	Fourier transform
FTICR	Fourier transform ion cyclotron resonance
FT-MS	Fourier transform-mass spectrometry
GABA	γ-Aminobutyric acid
gDNA	Genomic DNA
Glu	Glutamine
H$_2$O	Water
H$_2$SO$_4$	Sulphuric acid
HATU	*N,N,N′,N′*-Tetramethyl-O-(7-azabenzotriazol-1-yl)uronium hexafluorophosphate
HCl	Hydrochloric acid/hydrogen chloride
His	Histidine
HMBC	Heteronuclear multiple-bond correlation spectroscopy
HOBt	Hydroxybenzotriazole
HPLC	High performance liquid chromatography
HPLC	High pressure liquid chromatography
HR	High resolution
HSQC	Heteronuclear single-quantum correlation spectroscopy
IBX	Iodoxybenzoic acid
IPA	Isopropyl alcohol
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IR	Infrared
K$_2$CO$_2$	Potassium carbonate
Kan	Kanamycin
KAS	Keto-acyl synthase
KMNO$_4$	Potassium permanganate
KOH	Lithium hydroxide

Abbreviations	Definitions
KR	Ketoreductase
KS	Ketosynthase
KSQ	Ketosynthase decarboxylase
Las	Lasalocid
LB	Lysogeny broth
LC	Liquid chromatography
LED	Light emitting diode
LiOH	Lithium hydroxide
$[MH]^+$	Proton adduct mass ion
$[MNa]^+$	Sodium adduct mass ion
m/z	Mass to charge ratio
MALDI	Matrix assisted laser desorption ionisation
MB	Maltose Bennett
MCAT	Malonyl-CoA: *holo* acyl carrier protein transacylase
*m*CPBA	*meta*-chloroperoxybenzoic acid
Me	Methyl
MeCN	Acetonitrile
MeOD	Deuterated methanol
MeOH	Methanol
$MgBr_2.OEt_2$	Magnesium bromide diethyl etherate
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulphate
MS	Mass spectrometry
MS^2	Tandem mass spectrometry
MT	Methyltransferase
MW	Molecular weight
N_2	Nitrogen gas
NaCl	Sodium Chloride
NAD^+	Nicotinamide dinucleotide oxidised form
NADH	Nicotinamide dinucleotide reduced form
$NADP^+$	Nicotinamide dinucleotide phosphate oxidised form
NADPH	Nicotinamide dinucleotide phosphate reduced form
$NaHCO_3$	Sodium hydrogen carbonate/Sodium bicarbonate
NaI	Sodium iodide
NaOH	Sodium hydroxide

Abbreviations	Definitions
NMR	Nuclear magnetic resonance
NOESY	Nuclear overhauser effect spectroscopy
NRP	Non-ribosomal peptide
NRPS	Non-ribosomal peptide synthase
OD_{600nm}	Optical Density at 600 nm
ORF	Open reading frame
PanK	Pantothenate kinase
PCC	Pyridinium chlorochromate
PCP	Peptidyl carrier protein
PCR	Polymerase chain reaction
Pd	Palladium
Pd/C	Palladium on activated carbon
PhMe	Toluene
PhOH	Phenol
PhSH	Thiophenol
P_i	Inorganic phosphate
PKS	Polyketide synthase
PLE	Pig liver esterase
PPant	4'-phosphopantetheine
PPAT	Phosphopantetheine adenyltransferase
ppm	Parts per million
QIT	Quadrupole ion trap
RF	Retention factor
RhFRED	Ferredoxin reductase
RNA	Ribonucleic acid
RP	Reverse phase
Rpm	Revolutions per minute
RT	Room temperature
S. coelicolor	*Streptomyces coelicolor*
S. lasaliensis	*Streptomyces lasaliensis*
SAM	*S*-Adenosyl methionine
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Se	Selenium
Ser	Serine
Sfp	4'-phosphopantetheinyl transferase
SNAc	*N*-acetylcysteamine

Abbreviations	Definitions
TBE	Tris borate-EDTA
TBS	*tert*-butyldimethylsilyl
TE	Thioesterase
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
THF	Tetrahydrofuran
TIC	Total ion current
$TiCl_4$	Titanium tetrachloride
TLC	Thin layer chromatography
TLM	Thiolactomycin
Tmn	Tetronomycin
TOF	Time of flight
TSB	Tryptic soy broth
Tsn	Tetronasin
UPLC	Ultra-pressure liquid chromatography
UV	Ultra violet
Vis	Visible
W	Watts
WT	Wildtype

__Abstract__

Polyketides constitute an invaluable source of products for the agrochemical, pharmaceutical and biotechnology industries. Due to the enzyme-bound covalent nature of their intermediates, polyketide biosynthesis is challenging to study. The use of chemical probes offers a complementary approach to molecular biology and enzymology to characterise the nature of biosynthetic intermediates.

My PhD work involved the synthesis and the use of novel chemical probes to investigate biosynthetic pathways leading to several polyketide compounds of biomedical interest, including thiolactomycin and tetronasin. The importance of these compounds is introduced in Chapter 1, together with an overview of the current knowledge of enzyme mechanisms involved in polyketide assembly.

Chapter 2 focuses on the use of chemical probes for the investigation of thiolactomycin bio-assembly. Thiolactomycin is a thiotetronate natural product produced by a soil bacterium discovered in the early 1980's. Its antibiotic, antimalarial and antituberculotic activities have been extensively studied and associated to type II fatty acid synthase inhibition, however its polyketide biosynthetic origins have only been recently unveiled. The chemical probes utilised to investigate thiolactomycin assembly were designed to mimic malonate building blocks utilised in polyketide biosynthesis, and were found to intercept putative early stage intermediates, which were characterised by LC-HRMS. The species in question were isolated from *in vivo* (live microorganisms) and *vitro* (recombinant enzymes) experiments; parallel investigations with other thiotetronate producing organisms corroborated our early results. In order to shed light on the late stages of intermediate processing, synthetic substrates mimicking established biosynthetic intermediates were prepared as *N*-acetyl cysteamine thioesters (SNAcs). Their processing and conversion to thiolactomycin by the enzyme machinery was demonstrated *in vivo* by supplementation of mutant bacterial cultures. Unnatural substrate mimics were also prepared and were shown to produce unnatural analogues of thiolactomycin by the same machinery to a small extent. Further investigations *in vitro* involved the expression of recombinant proteins and their incubation with synthetic

substrates in order to monitor specific individual transformations. Overall some insights into the final biosynthetic steps were gathered, but future work will be required to underpin the exact mechanisms of sulphur insertion and cyclisation leading to thiotetronate formation.

Chapter 3 focuses on biosynthetic investigations on the complex polyether tetronasin, a natural product of polyketide origin used as an additive in ruminal feedstock exerting antibiotic activity as an ionophore. The structure of tetronasin contains numerous chiral centres and several rings including tetrahydrofuran, pyran and cyclohexane moieties of unclear biosynthetic origin. Chemical probes mimicking polyketide malonate building blocks were synthesised and supplemented to cultures of the tetronasin producer *S. longisporoflavus* and of two mutant strains with inactivated genes shown to be essential for tetronasin production in the hope of capturing biosynthetic intermediates for detection by LC-HRMS. A few highly abundant and advanced species were detected and partially characterised; these provided key snapshots on the timing of the THF ring formation and other transformations in tetronasin bio-assembly, including *O*-methylation. Additional novel substrates were also prepared to probe the mechanism of tetronate ring formation; in preliminary *in vivo* experiments they did not prove fruitful, however they are currently being utilised for *in vitro* experiments concerning enzyme activity and crystallisation.

Chapter 4 entails the synthesis of new chemical probes bearing structural variations on the original 'chain termination' probes devised by the Tosin group; in particular, variations involving the replacement of an amide moiety with isosteric moieties were devised and implemented. The ultimate aim of this task was to investigate the tolerance of PKSs towards unnatural substrates throughout the polyketide assembly process. The new probes were tested on a variety of *in vitro* and *in vivo* PKS systems and their efficiency in capturing polyketide biosynthetic intermediates evaluated. Limited success was observed across most systems which appeared to be strongly linked to either probe toxicity or poor *in vivo* hydrolysis of the probes methyl ester precursors.

To overcome the latter issue, probes bearing photo-cleavable esters for a more controlled and efficient *in situ* generation of active carboxylate probes were synthesised. Unfortunately, this had little effect on the abundance and array of polyketide intermediates detected *in vitro* an *in vivo*, leading to the conclusion that changing of the probe amide functionality has a clear negative effect on its use for PKS species capture.

Chapter 5 consists of conclusions and suggestions for future work for the different areas of investigation discussed in the three previous chapters. Following this, Chapter 6 reports experimental details of chemical synthesis, biological experiments methods and analytical methods utilised in this work. Chapter 7 and 8 are the appendix and bibliography respectively.

"Progress is made by trial and failure; the failures are generally a hundred times more numerous than the successes; yet they are usually left unchronicled."

-Sir William Ramsay

1 Introduction

1.1 Natural products and their importance

1.1.1 Overview of natural products

Natural products, also known as secondary metabolites, refer to non-essential organic compounds which are produced by a myriad of microorganisms, such as bacteria, fungi and plants. Such compounds commonly ensue some form of advantage to the producer in their natural environment, such as a defensive mechanism.[1] Secondary metabolite natural products are by and large evolved out of primary metabolic pathways but are produced under certain environmental stimulus. Alkaloid, terpene, non-ribosomal peptide enzymes and polyketide synthases are prominent examples of biosynthetic enzymes responsible for their production; examples of these are quinine (**1**), lovastatin (**2**) and tyrothricin (**3**).[2] Most secondary metabolites present diverse bioactivities targeting metabolism within other organisms and are estimated to constitute 80% of marketed medicines, as themselves (**Fig. 1.1**) or as templates for bioactive scaffolds in synthetic drugs.[3]

Quinine   Lovastatin   Tyrothricin
1   **2**   **3**
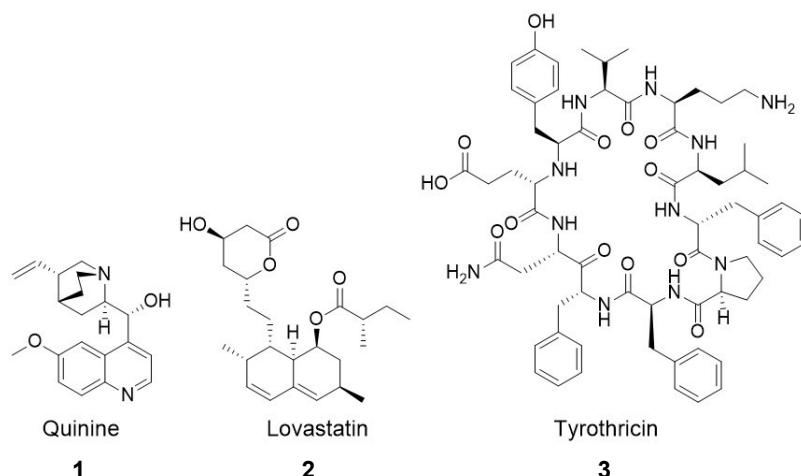
Figure 1.1 – **Examples of diverse natural products. Quinine (1) is an antimalarial alkaloid derived from *Cinchona* trees, lovastatin (2) is a fungal polyketide product used as a cholesterol lowering drug, and tyrothricin (3) is a bacterial non-ribosomal peptide used as a topical antibiotic.**

Not only do these compound find roles in the pharmaceutical industry, many natural products are crucial in the agrochemical and farming industrial sector for use as pesticides and also cattle feedstocks to prevent infections.[4] As previously mentioned, primary metabolism is largely responsible for generating flux for secondary metabolism and often the production of these compounds diverges from intermediates on primary metabolic pathways.[1] For example, the structurally complex anticancer agent taxol is produced in nineteen steps from the universal diterpene biosynthetic intermediate geranylgeranyl diphosphate.[5] As well as utilising primary metabolism, some secondary metabolite natural product pathways appear to have had an evolutionary link with some primary metabolism pathways and enzymes, such as in the case of polyketides, which share many building blocks and mechanisms of construction as seen in the biosynthesis of fatty acids.[6]

As well as offering a plethora of indispensable pharmaceutical and agrochemical compounds, natural products appear in our daily lives as seemingly benign compounds such as flavones like cyanidin (**4**), whilst others, like aflatoxin B$_1$ (**6**), are some of the most toxic compounds known to us (**Fig. 1.2**).[1,7]
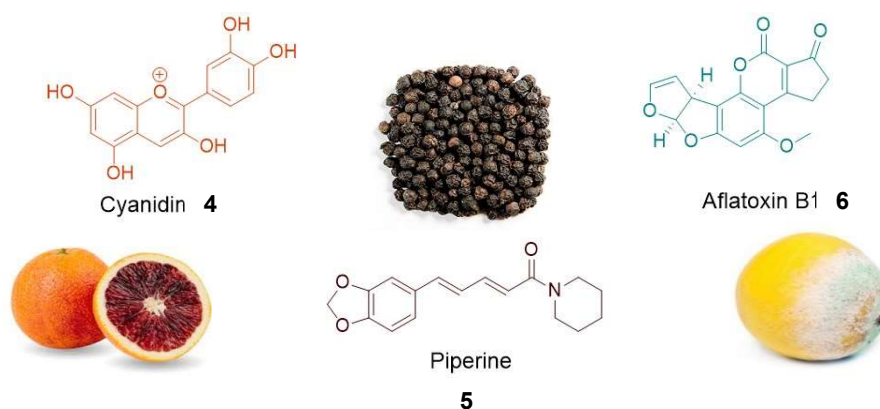
Figure 1.2 – **Some natural products encountered outside industrial application. Cyanidin (4) is an anthocyanin responsible for the blood orange colour, piperine (5) is the main flavour compound of black pepper and aflatoxin B$_1$ (6) is a potent carcinogen produced by *Aspergillus* fungi which commonly grows on fruit and nuts.**

Whilst a large number of natural products may result from a single biosynthetic pathway, many natural products are also the result of more than one primary/secondary metabolic pathway (**Fig. 1.3**). For instance the psychoactive natural product tetrahydrocannabinol (**7**) contains moieties deriving from the fatty acid, polyketide and terpene biosynthetic pathways.[8] In contrast, the antibiotic monensin A (**8**) is produced entirely from polyketide biosynthesis and additional tailoring steps, in spite of its overwhelming structural complexity, teixobactin (**9**) is also biosynthesised by a single type of biosynthetic machinery in the form of a non-ribosomal peptide synthetase (NRPS).[9]
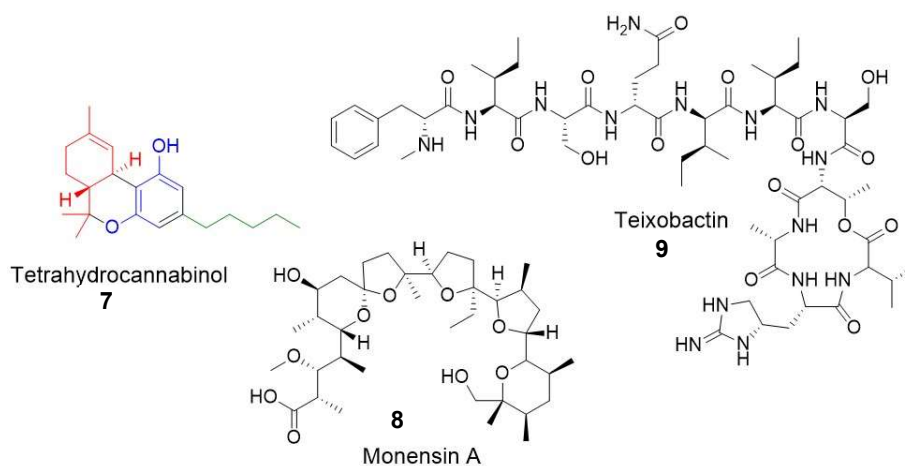
Figure 1.3 – **Tetrahydrocannabinol (7) derives from fatty acid, polyketide and terpene biosynthetic pathways, whereas monensin A (8) and teixobactin (9) are produced by a polyketide synthase and non-ribosomal peptide synthetase respectively.**

1.1.2 Natural products in the pharmaceutical and agrochemical industry

Natural products have been the starting point for the discovery and development for the majority of marketed drugs and agrochemicals. As of 1993, over 50% of the most prescribed drugs in the US were either themselves natural products or possessed some form of natural product "inspired" scaffold (**Fig. 1.4**).[10,11] Natural product-derived compounds also represent key players in a multitude of drug groups such as anticancer/cardiovascular treatments, and analgesic compounds to name a few.[10,12] Natural products have been used in traditional medicine worldwide for over a millennia.[13]
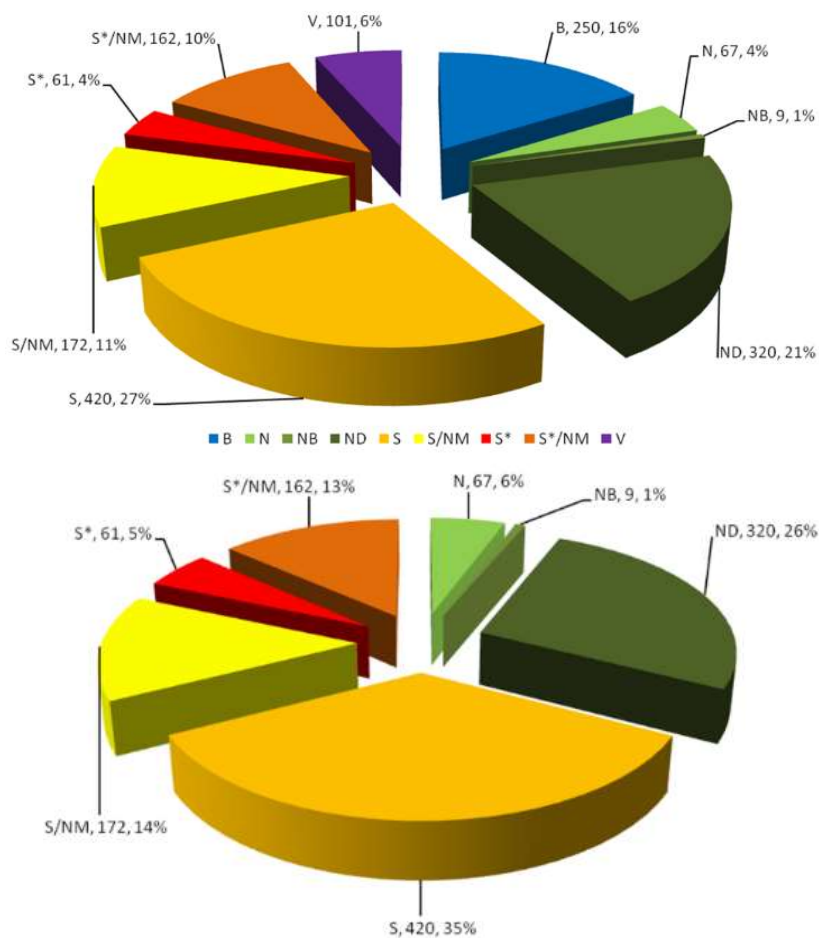
B	Biological macromolecule	V	Vaccine
S	Synthetic drug	S*	Synthetic drug, NP pharmacophore
NB	Botanical drug (defined mixture)	ND	Natural product derivative
NM	Natural product mimic	N	Unaltered natural product

Figure 1.4 – **Pie charts illustrating approved drugs from 1981-2014 categorised by origin. Top pie chart illustrates natural product representation among all drugs (1562) including biological and vaccines. Bottom pie chart shows natural product representatives across small molecule drugs (1211).**[11]

A similar history emerges for natural product-derived agrochemicals. Pesticides are dominated by natural product-derived/inspired derivatives and represent more than a third of new active ingredient registries between 1997-2010 (**Fig. 1.5**).[4]
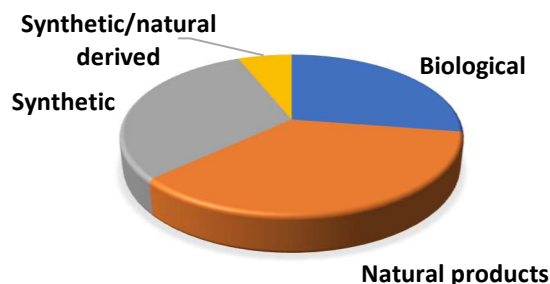
Figure 1.5 – **Pie chart showing all new active ingredient registries for conventional pesticides and biopesticide: natural products represent more than a third.**[4]

1.1.3 Antimicrobial natural products and antibiotic resistance

In the area of antimicrobials, natural products are easily the most important class of compounds. Since the "golden age" of antibiotic discovery between 1940-1950, most antibiotic compounds developed in the past 60 years have been the products of semi-synthesis from well documented bioactive natural product scaffolds,[14] very few compounds like teixobactin (**9**) novel mechanisms of action are being discovered.[15] The lack of new antibiotic discovery (**Fig. 1.6**) coupled with the irresponsible use of antibiotics has made concern of global antibiotic resistance very real.[16]
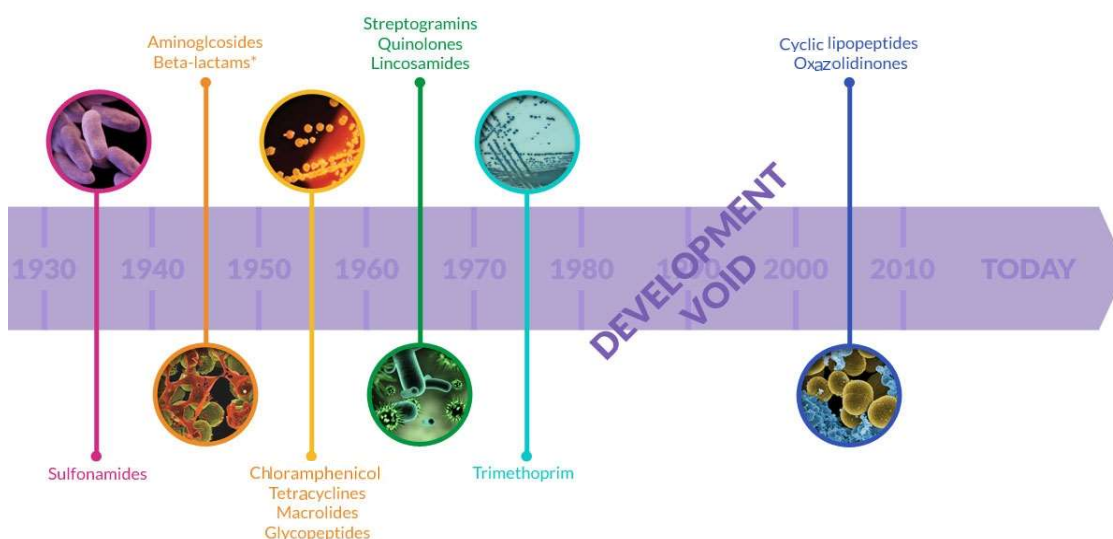
Figure 1.6 **(from www.antibioticresearch.org.uk/about-antibiotic-resistance) – Developmental timeline of new marketed antibiotics. Since the "golden age" of discovery, very few antibiotics with novel mechanisms of action have been discovered with a thirty-year void after around 1975.**

The quest for new antibiotic molecules is a continuous challenge, with scientists looking at natural products generated in diverse environments and from different biosynthetic pathways. For instance, **9** was discovered through specialised culture techniques used on previously "unculturable" bacteria and shows potent cell wall inhibitory effects through binding to lipid II.[15,17]

1.2   Polyketide natural products

1.2.1   Overview of Polyketides

Polyketides are a family of secondary metabolites produced by bacteria, fungi and plants. The biosynthesis of these compounds is tied closely to that of fatty acid biosynthesis from primary metabolism where both share the same simple building blocks. However, polyketide biosynthesis presents more flexibility with respect to enzymatic modification which allows a greater range of diversity and structure complexity.[1,18] Polyketide assembly is carried out by a minimal set of chemical reactions. These will be briefly presented in the following paragraphs. An estimated £10 billion annual value has been placed on polyketide derived pharmaceutical compounds.[19] Polyketides are produced by a range of multifunctional enzymes known as polyketide synthases (PKSs). There are several types of PKSs according to their structural organisation and workings, however the general mechanisms of polyketide assembly are the same across the different enzyme families.[18] Polyketide biosynthesis has been largely investigated only in the last 70 years. Isotopic labelling and NMR spectroscopy have been key tools in deciphering the originally cryptic mechanism of their construction, and has led to new strategies to exploit their biosynthesis and to upregulate production or produce novel unnatural analogues.[20]

1.2.2   Biosynthesis of polyketides

1.2.2.1   Overview of polyketide biosynthesis

Polyketide biosynthesis begins by condensing a "starter" unit, which is commonly an acetate, onto an "extender" unit such as malonate *via* a decarboxylative Claisen condensation (**Scheme 1.1, top right**). An extended polyketide chain is then formed from these units and this can be further elongated using more extender units. At each stage of elongation, other chemical alterations can be made on the growing chain to diversify the carbon skeleton depending on the nature of the PKS enzymes involved. PKSs are broadly classified into three types, these are defined as type I, II and III. Type I PKSs are large multifunctional mega-enzymes also defined as assembly lines for which intermediates are shuttled along and modified within multiple modules; these are compartmental enzyme units responsible for at least one process of carbon chain extension. Type II and III PKSs are smaller, monofunctional dimeric enzymes but perform the same transformations as type I enzymes. Each module contains different domains which are briefly described in the next sections.

1.2.2.2   Domains within polyketide biosynthesis

In polyketide biosynthesis, there are two essential domain types: ketosynthases (KSs) and acyl carrier proteins (ACPs), which are responsible for malonyl decarboxylative Claisen condensation leading to polyketide chain formation/elongation. The ketosynthase domain has a conserved cysteine thiol residue which binds the starter unit or a polyketide intermediate to allow the nucleophilic attack of a decarboxylated malonate extender unit bound to the ACP. The selection of starter and extender units is mediated by AT (acyl transferase) domains.
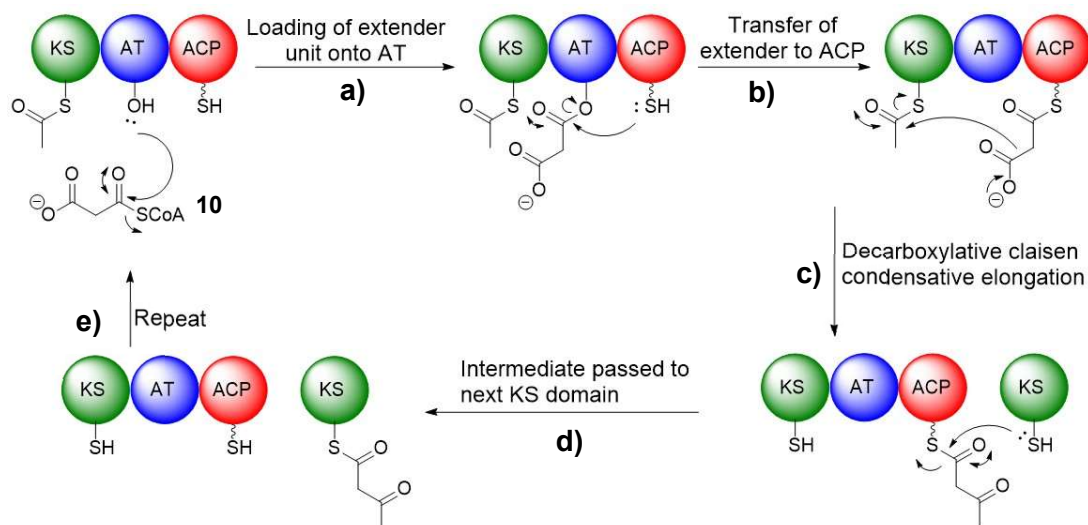
Scheme 1.1 – **Stepwise scheme of polyketide elongation. A): A malonyl CoA unit (10) is loaded onto the AT domain which has substrate specificity for it. B): This extender is then shuttled to the ACP. C): malonyl-ACP decarboxylation results in a reactive nucleophilic enolate which attacks a KS-bound acyl chain (an acetyl starter unit in this case). D): the resulting elongated carbon chain is passed to the downstream KS domain (in modular systems) or back to the previous KS domain (iterative systems). E): the elongation process can now be repeated or other reactions can occur on enzyme-bound intermediates such as release from the PKS via hydrolysis.**

These are used to recruit specific coenzyme-A (CoA) starter/extender units. AT domains have a conserved serine residue which is used to bind and shuttle the extender units to the ACP domain. These domains have been studied to the extent that their specific substrate can be bioinformatically predicted in unelucidated polyketide genes.[21,22] Acyl carrier proteins are small dynamic proteins of approx. 10-12 kDa in size and have conserved serine residues, which are non-functional as nascent translated "*apo*" form. The serine residues are post-translationally modified by a phosphopantetheinyl transferase enzyme (PPTase) with a phosphopantetheine cofactor which acts as a swinging "arm" to access other domains within the same or adjacent protein modules carrying polyketide intermediates: this is known as the "*holo*" form of the ACP (**Scheme 1.2**). The ACP's key role is to accept extender units from the AT domain and provide these to KS's for decarboxylative Claisen condensation. Although acetate and malonate are the most common respective starter and extender units, there are many examples of more functionalised substrates utilised in PKS biosynthesis to afford more complex products.
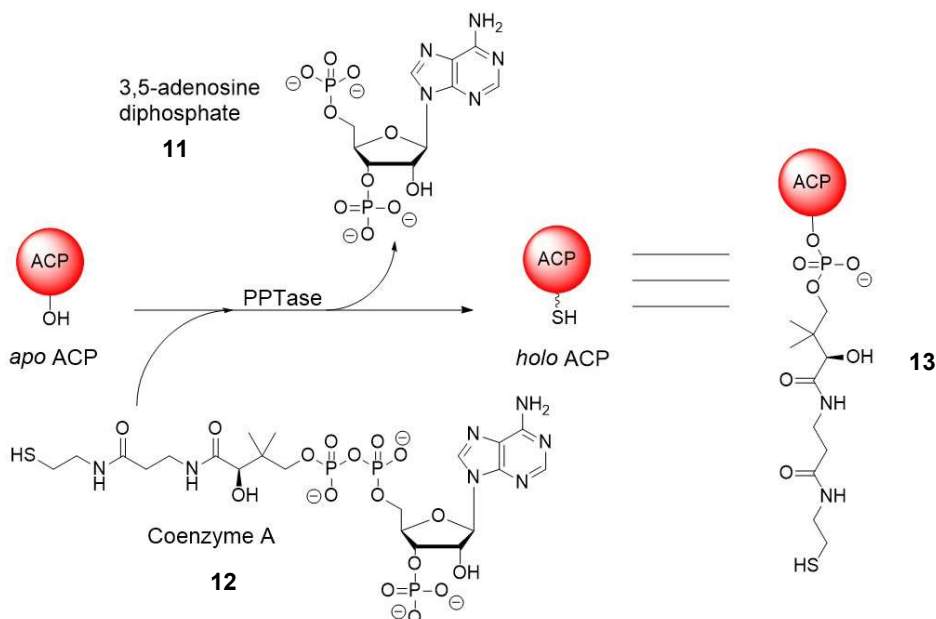
Scheme 1.2 **– A pantetheine cofactor is transferred from coenzyme A (12) onto the inactive *apo*-ACP enzyme/domain using a PPTase enzyme to release 3,5-adenosine diphosphate (11) and generate the active *holo*-ACP (13); the arm is often abbreviated to a wavy line showing only the thiol group.**

Additional domains can also be present within PKSs to modify the polyketone intermediates ahead of further chain extension. These include the ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) domains (**Scheme 1.3**).[18] KR domains utilise the reduced cofactor nicotinamide adenine dinucleotide (NADH) as a hydride source to convert ketones to secondary alcohols in a stereospecific fashion to release oxidised nicotinamide adenine dinucleotide (NAD$^+$). DH domains can then dehydrate secondary alcohols formed by the action of KR domains to give α, β-unsaturated thioesters. ER domains can finally act on the latter substrates using NADH to convert them to fully saturated thioesters.
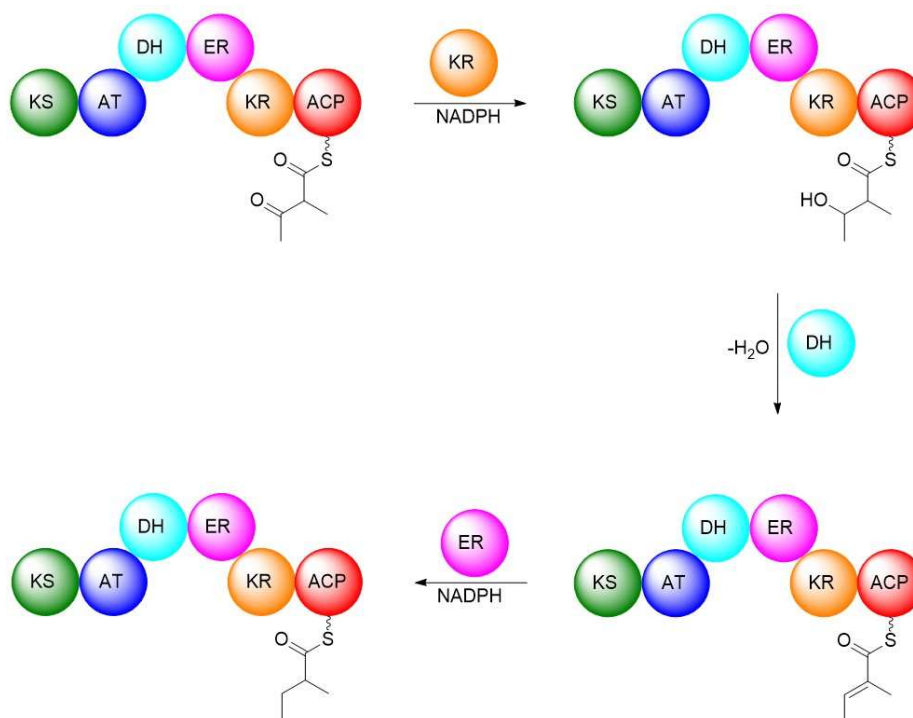
Scheme 1.3 **– Stepwise processing of a nascent elongated polyketide by the KR, DH and ER domains. First step: KR domain reduces ketone of ACP-bound polyketide using NADPH as a hydride source; second step: DH domain dehydrates the alcohol to eliminate water and leave an unsaturated thioester; third step: ER domain uses NADPH to reduce the unsaturated thioester to a fully saturated thioester.**

Many PKSs possess some but not all of these domains within individual modules; this leads to polyketide structure of much higher diversity in comparison to fatty acids Many PKSs also have a terminal thioesterase domain (TE) with a catalytic serine hydroxyl group: this is responsible for holding the completed polyketide chain and allowing its hydrolysis from the PKS, or cyclisation. Many PKSs have alternative chain release mechanisms: these include reductive release whereby NADPH is used to free an aldehyde product, release by lactone formation or even passing to other attached machineries such as an NRPS which will be examined in detail (**Section 1.2.3**).

1.2.2.3   Domain architecture of type I polyketide synthases

Type I PKSs can be categorised as iterative or modular systems. Modular PKSs present grouped sets of domains which are each responsible for extending a polyketide chain once with a single extender malonate substrate. This type of enzyme can consist of several modules each comprising the essential AT, KS and ACP domains along with any combination of the KR, DH and ER "reductive" domains.

The key difference with respect to iterative systems is that each module within the modular PKS is utilised once for a chain extension before the intermediate is passed onto the KS of the module downstream. Consequently, the modular PKSs can be large (the erythromycin A producing PKS protein deoxyerythronolide B synthase (DEBS) is approximately two megadaltons)[21] and the products may be very diverse. They are however more predictable from a bioinformatics perspective and the polyketide backbone structure can be directly traced back to the domain architecture of the modules in most cases. This can be clearly exemplified by the polyketide deoxyerythronolide B (DEB) which is produced by *S. erytheae* (**Fig. 1.7**).[21,23]
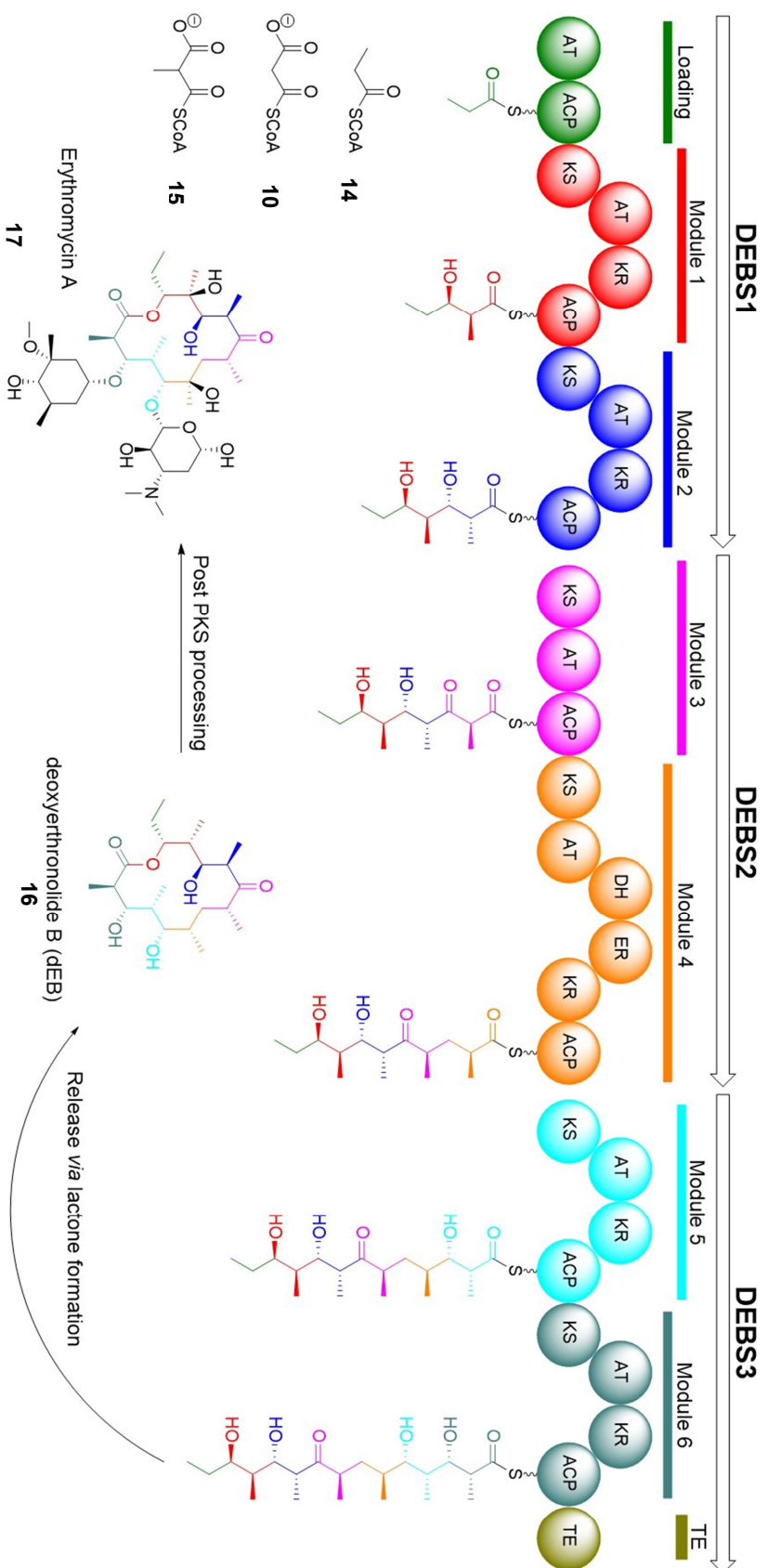
Figure 1.7 – **Domain architecture of the DEBS type I modular PKS responsible for making the macrolide deoxyerythronolide (16) from a propionate starter unit (14) as well as the malonyl (10) and methylmalonyl (15) extender units. dEB (16) is then processed post-PKS into the final natural product erythromycin A (17). The putative intermediates after elongation and processing by present KR, DH and ER domains are shown bound on each ACP. The bioinformatically proposed domain architecture of DEBS matches the macrolide structure.**

Whilst having a simpler domain setup, iterative PKSs can be highly cryptic due to the unpredictable polyketide chain length and usage of KR, DH and ER domains. These systems use a single module repetitively to extend a polyketide chain; intermediates are shuttled back to the same KS for further elongation. Post-PKS tailoring can heavily modify the original polyketide products such as in dynemicin A (**21**) which is believed to be constructed from two heptaketide subunits produced by the same iterative PKS machinery (**19** and **20**, **Scheme 1.4**).[24,25] The majority of known type I iterative PKS natural products are produced in fungi whereas bacteria tend to harbour modular PKSs.
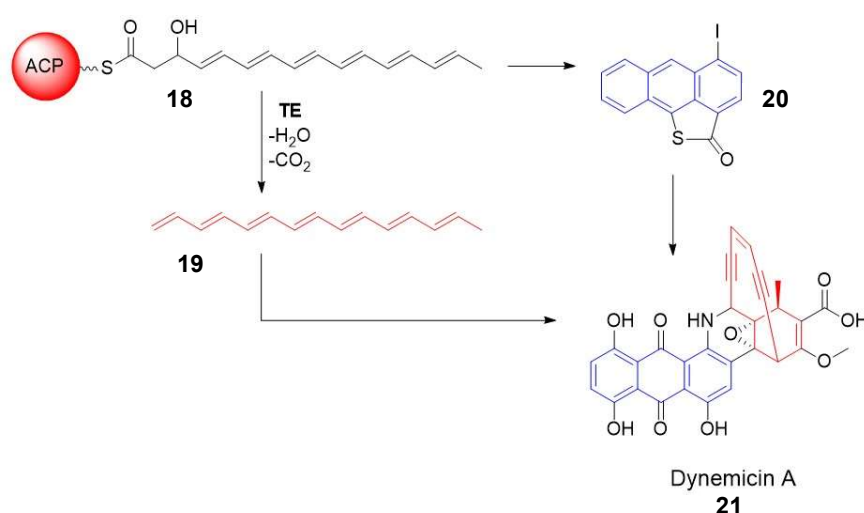
Scheme 1.4 – **Dynemicin A (21) is biosynthesised by an iterative PKS which builds a conjugated octaketide backbone (18); two chains from the same PKS proposedly undergo separate transformations through two very different intermediates (19 and 20) and are tethered in the final compound, the exact biosynthetic steps remain to be fully elucidated.**

The repeating nature of the type I iterative PKS also raises the question of how the correct length polyketide is reached and reacted before further elongation can occur. Post-translational enzymes are proposed to only fold correct size polyketide chains to further process them; this presumably prevents incomplete intermediates being processed. Type II iterative systems tend to include a special chain length factor (CLF) domain which mediates and allows the next transformations to begin only when a particular length of polyketide chain is reached.[26]

1.2.2.4 Domain architecture of type II and III polyketide synthases

Type II PKSs perform identical mechanistic transformations as type I systems: the domains involved, however, are discreet monofunctional proteins; many of them are homodimers (**Scheme 1.5**, **top**).[27] Actinorhodin, (which will be discussed more later), is an important and well-studied type II iterative PKS product. Type III PKSs consist of individual multifunctional proteins for which the entire biosynthesis occurs on coenzyme-A (CoA) bound intermediates instead of ACPs (**Scheme 1.5**, **bottom**).[28] They are exemplified by chalcone and stilbene synthases and are more extensively studied in plants where they are most commonly found.[29]
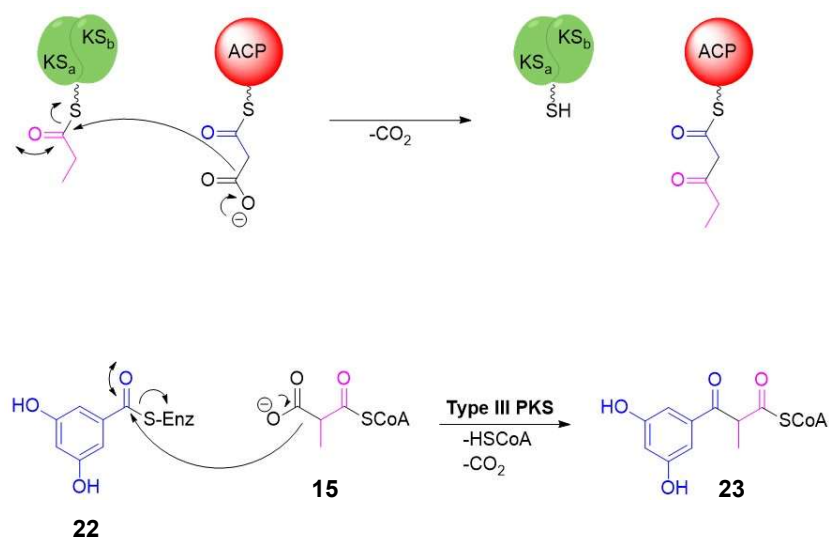
Scheme 1.5 – **Scheme showing polyketide biosynthesis in type II and II systems. Top: independent monomeric ACP and homo-dimeric KS proteins perform the decarboxylative Claisen condensation in type II systems. Bottom: An enzyme bound starter unit (22) undergoes decarboxylative Claisen condensation from an CoA bound extender unit such as methylmalonyl CoA (15) (this can also be enzyme bound) which leads to the elongated polyketide 23.**

1.2.3 Non-ribosomal peptide synthetases (NRPSs) and hybrid PKS-NRPSs

Beside purely polyketide derived products, there are also molecules made by 'hybrid' biosynthetic machineries. Perhaps the most well-known compounds of this kind are polyketide-peptide compounds which are produced by hybrid polyketide/non-ribosomal peptide synthetases (PKS-NRPSs). These enzymes comprise a PKS region fused directly to a non-ribosomal peptide synthetase (NRPS) region. The *B. cereus* antibiotic metabolite zwittermicin A is a PKS-NRPS product (**Scheme 1.6**);[30] thiotetronates which will be extensively discussed later (**Section 1.5.1**) are also PKS-NRPS products.
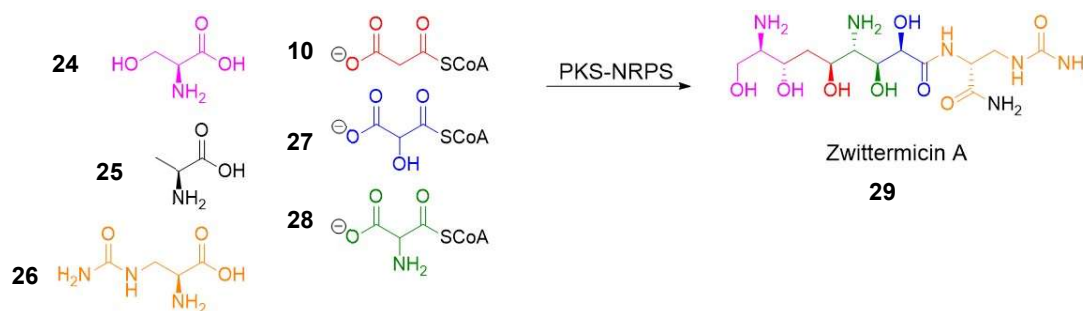
Scheme 1.6 *– **Zwittermicin A (29) is synthesised from amino acids and several acyl coenzyme A derivatives, such as citrulline (26), alanine (25), serine (24), aminomalonyl (28), hydroxymalonyl (27) and malonyl CoA (10). A polyketide synthase non-ribosomal peptide synthetase (PKS-NRPS) is responsible for its production.***

NRPSs are multifunctional giant proteins that assemble peptide chains together independently of the ribosome. NRPSs can utilise non-proteinogenic amino acids and fatty acids as building blocks activated by adenylation (**Scheme 1.7**). Peptide biosynthetic intermediates can also undergo several chemical transformations: common examples include amide *N*-methylation, epimerisation of side chains and cyclisation of serine/cysteine amino acid side chains to form oxazole and thiazole moieties. Like PKS products, NRPS products are made by bacteria, fungi and plants, and display significant bioactive properties, including antibiotic, antifungal, anticancer, immunosuppressant and others. NRPSs have a similar domain architecture to type I PKSs where specific domains are essential and the non-essential domains are used for tailoring intermediates. They also undergo similar chain release *via* hydrolysis, reduction or even self-cyclisation to form

lactams/lactones. Similar to PKS biosynthesis, building blocks are loaded onto the enzyme and elongated in an assembly line fashion.
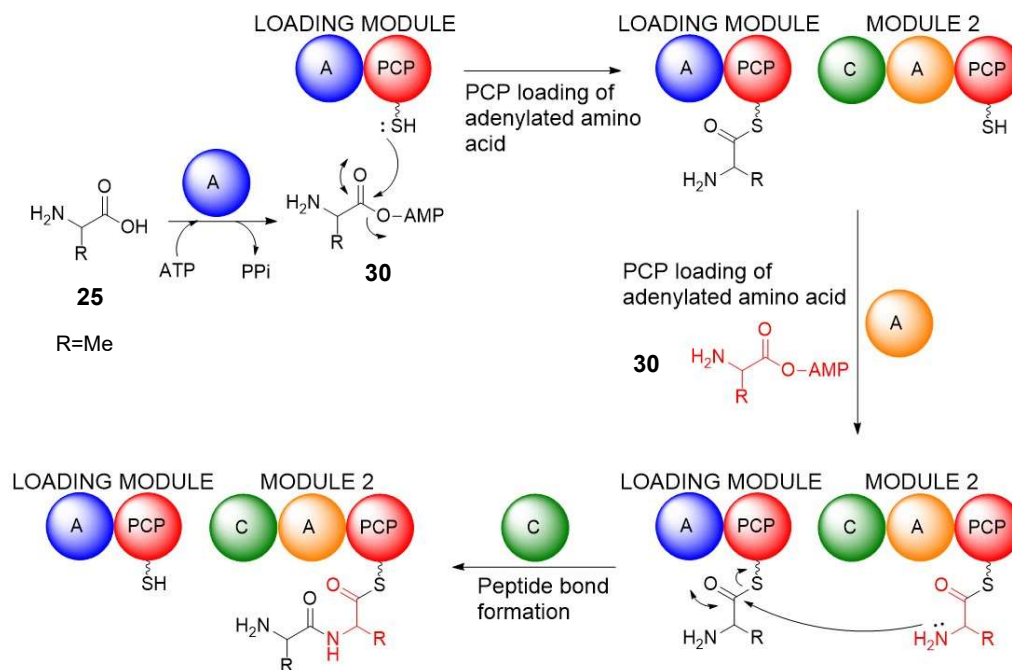
Scheme 1.7 – **Peptide formation on an NRPS. An amino acid such as alanine (25) is activated by the A domain using ATP to give an amino acid adenylate (30) (top left): this is then loaded onto the PCP domain (top right). A second alanine amino acid is adenylated (30) and loaded onto the downstream PCP domain (bottom right); condensation of the two amino acids leads to peptide elongation on the downstream PCP which can then be further processed (bottom left).**

The essential domains in NRPSs are the condensation (C), adenylation (A) domains and the peptidyl carrier protein (PCP), which is occasionally referred to as the thiolation (T) domain. The PCP domain is a small dynamic domain with a conserved serine residue post-translationally modified with a pantetheine cofactor that shuttles amino acids around and holds them during peptide bond formation. These domains are essentially the same in modified structure and function as the ACP domain from PKSs. The A domain has specificity for certain amino acids and will recruit these and attach them to the NRPS PCP; this process requires ATP and proceeds by forming an activated adenosine monophosphate (AMP) amino acid ester and tethering them to release AMP and

pyrophosphate rather than forming a coenzyme-A adduct as PKS extender units are loaded. The C domain is responsible for forming the peptide bond between two loaded PCP-amino acids on two adjacent modules. This occurs by the free amine from the downstream amino acid attacking the thioester bond of the upstream amino acid to form an elongated peptide on the downstream PCP; further amino acids can then be loaded on downstream modules leading to increasing peptide chain length and further processing.

As with PKSs, NRPS biosynthesis can utilise non-essential domains to further decorate and change the nascent peptide such as the methylation (M) domain which will methylate the peptide bond nitrogen, the epimerisation domain (E) will reverse the stereochemistry of a given amino acid stereocentre, and the cyclisation (Cy) domain which will condense a serine or cysteine dipeptide motif to form the respective oxazole or thiazole. There are many examples where PKS and NRPS enzymes are both involved intimately in the formation of natural products which contain both peptide and polyketide elements. The combinatorial use of non-essential PKS tailoring domains KR, DH or ER coupled with a large substrate range[31] from both PKS and NRPS biosynthetic routes essentially leads to limitless diversity of PKS-NRPS products.[32]

1.2.4   Biosynthesis of fatty acids

Fatty acids are primary metabolites that are essential for all organisms. They have important functions associated to energy storage, cell structure and many more. They are produced by fatty acid synthases (FASs), which utilise acetate and malonate units to construct carbon chains in a similar fashion to PKSs (**Scheme 1.8**). PKSs evolved from FASs and contain the same domains, FASs however are iterative and use all domains in each extension. When the correct chain length is reached, the TE domain mediates the hydrolysis of a linear carboxylic acid or "fatty acid" product.
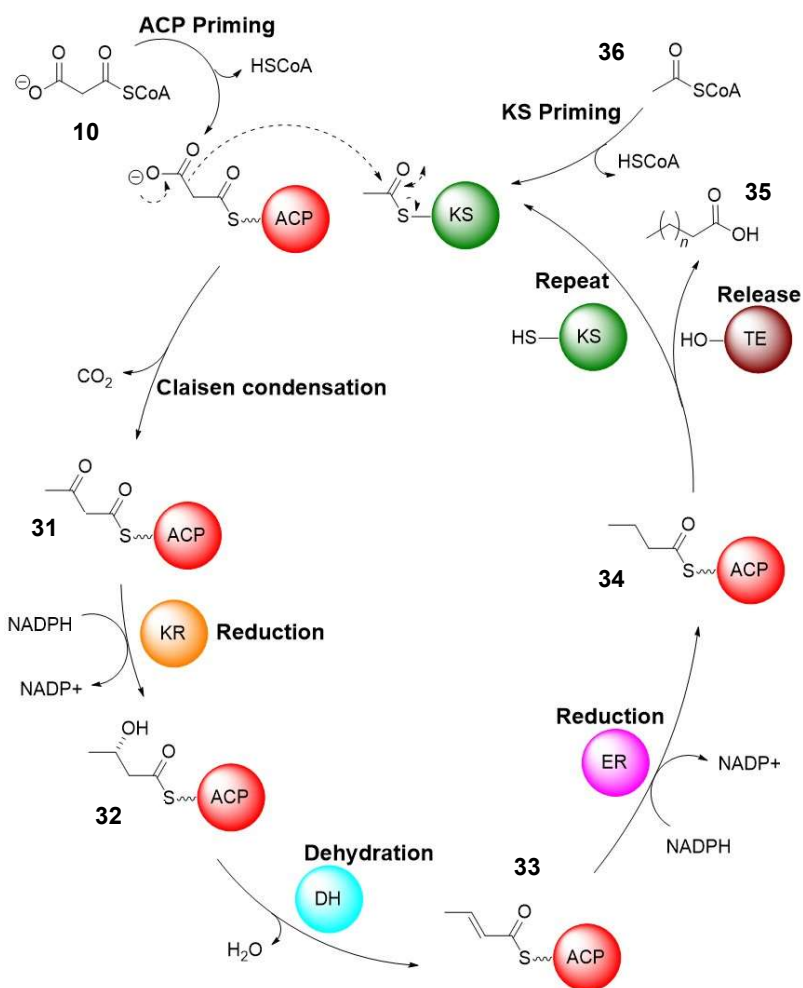
Scheme 1.8 – Overview of fatty acid biosynthesis. First the KS domain is primed with an acetyl coenzyme-A starter unit (36) and the ACP is primed with malonyl coenzyme-A (10), Claisen condensation occurs to elongate the carbon chain (31). Then KR uses NADPH to reduce the ketone to a hydroxyl group (32), water is then eliminated by the DH domain to give an unsaturated thioester (33), which is again reduced with ER using NADPH to give a saturated thioester (34). This can then either be hydrolysed by TE to give a fatty acid product (35) or can be further elongated in the cycle.

FAS are classified into type I and type II enzyme groups. Both types are iterative systems unlike in PKSs. Type II are discrete proteins only found in prokaryotes, whereas type I are single large multifunctional enzymes similar to type I iterative PKSs and are found in eukaroytes.[33] Much of the early stage investigations into PKS biosynthesis relied on the more extensively studied and less complex fatty acid pathway; this will be further presented in the next section.[20,34]

1.3   Early investigation of polyketides biosynthesis.

1.3.1   Isotope labelling of polyketide products

The earliest ground-breaking polyketide research was carried out by Arthur Birch in the 1950s whilst studying the biosynthesis of 6-methyl salicylic acid (6-MSA, **38**): this is a small metabolite with mild antibiotic activity and is a precursor to many other more complex metabolites such as the fungal toxin patulin (**30**). Birch theorised that four units of acetate could be joined together *via* Claisen condensation to form a linear polyketone chain, which could ultimately dehydrate to form a phenolic product: this was shown by feeding radiolabelled acetate to the fungal 6-MSA producer *Penicilium patulum* and analysis of the degradation products. Birch's predicted incorporation pattern was observed in the final compound as shown in degradation studies and selective radioactivity measurement of the final product (**Fig. 1.8, top**).[35] Similar isotopic experiments followed using deuterium and oxygen-18 established the biosynthetic origins of the *Aspergillus terreus* polyketide product terrein (**41**, **Fig. 1.8**): these showed that a portion of the product backbone is derived from three acetates.[36,37] The presence of an intact CD_3 labelled carbon also indicated that acetate was the starter unit in the biosynthetic pathway leading to **41**.
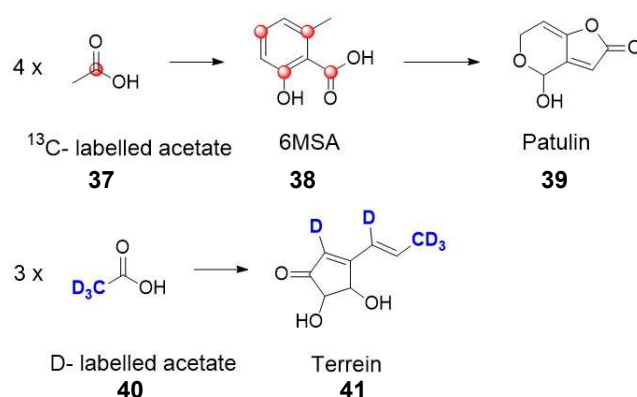
***Figure 1.8 –* Top: [13]C labelled acetate (37) fed to the *P. patulum* leading to labelled 6MSA (38) which is a precursor to patulin (39). Bottom: deuterium labelled acetate (40) feeding in *A. terreus* showed that terrein (41) is produced partly by incorporation of three acetates and another C_2 fragment.**

Further use of double ^{13}C labelled acetate allowed even greater insight into the specific chemical transformations occurring to produce complex polyketides such as with the use of NMR coupling constants. If a doubly labelled acetate (**42**) used in a biosynthetic pathway becomes fragmented (C-C bond breakage) then the characteristic coupling constant will be lost; this can indicate which C-C bonds come from intact substrates and which have been integrated by other means. Spiciferone phytotoxins (**44**) are produced by the fungus *C. spicifer* and are responsible for leaf spot disease. Feeding double labelled acetate gave low yields of labelled product but combined with ^{13}C labelled methionine (**43**), moderate level production of labelled spiciferone A (**45**) was achieved and the difference in C-C coupling constants indicated that specific carbon atoms were incorporated by several *S*-adenosyl methionine (SAM) methylation events and not from the acetate units used in PKS carbon chain assembly (**Scheme 1.9**).[37,38]
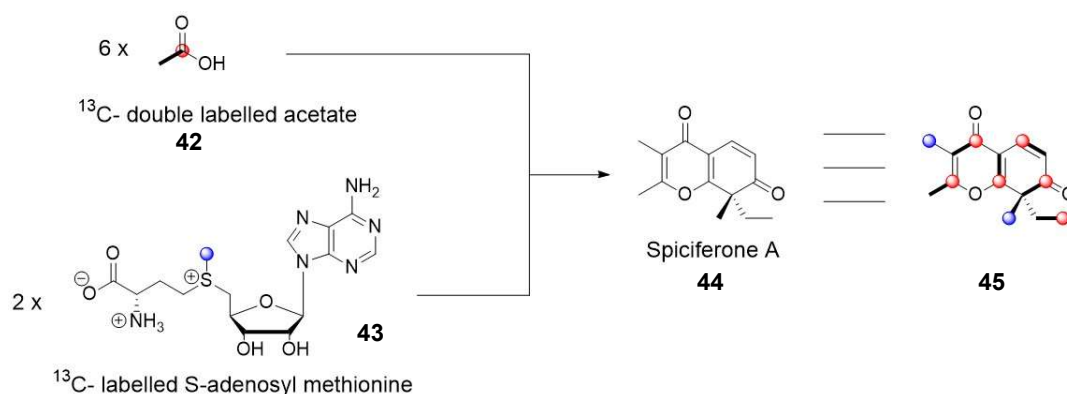
Scheme 1.9 – Feeding of doubly ^{13}C labelled acetate (42) and singly ^{13}C labelled SAM (43) to *S. spicifer* resulted in labelled spiciferone A (44). NMR analysis of adjacent carbon coupling in spiciferone A indicates the source of each carbon atom (45)

Throughout the isotopic studies of polyketide synthesis, it became clear that more than just acetate moieties could be a starter unit for a polyketide product due to the presence of singly or doubly deuterated carbons originally from a triple labelled acetate, suggesting the processing of acetate into other starter units.[20]

1.3.2 Linking polyketide biosynthesis to fatty acid biosynthesis.

Possibly the most influential work in bridging the gap between the knowledge of both polyketides and fatty acids was the study of the previously mentioned macrolide antibiotic erythromycin from *S. erythrea* and also the type II PKS product actinorhodin which is a strongly blue coloured antibiotic produced by *S. coelicolor*. Two publications in the 1984 and 1990 investigating these two systems laid the groundwork on the discovery of all following polyketide products and the genes responsible for making them.[39,40]

1.3.3 Erythromycin, actinorhodin and discovering biosynthetic gene clusters

Actinorhodin (**46**) is a benzoisochromanequinone antibiotic compound produced by *S. coelicolor*, with a vivid blue colour due to its conjugated ring structure (**Fig. 1.9, top**).[41] Through early pioneering work on gene deletion by Malpartida *et al.*, a library of non-producing mutants lacking the characteristic colour were generated; it was found that inserting a particular DNA segment from the wildtype strain into any of the mutants restored actinorhodin production as observed by the blue phenotype.[39] The DNA segment could even be inserted into other species (*S. parvulus*) which allowed actinorhodin production; this indicated that all biosynthetic genes responsible for making actinorhodin were on this single DNA segment and all of these genes may be clustered in close proximity. This is nowadays known colloquially as a biosynthetic gene cluster (BGC).[20,39] The sequence of the clustered genes showed homology to some of those in fatty acid biosynthesis such as the ACP, KS and KR domains.[20]
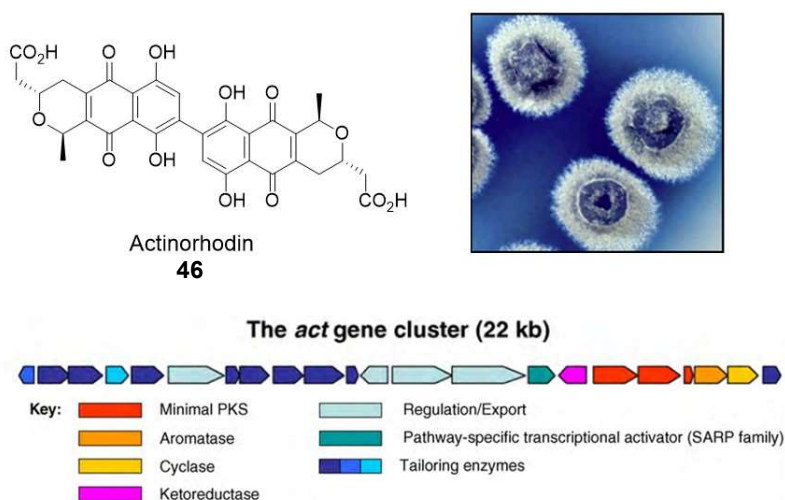
Figure 1.9 – **Top: Actinorhodin (46) is a blue antibiotic polyketide produced by *S. coelicolor*. Bottom: All of the biosynthetic genes are clustered together in the genome which includes the PKS and many other regulatory and tailoring genes.**

Further investigation into the product has led to categorisation of the PKS responsible for making actinorhodin as a type II PKS, which, similarly to type II FASs,[42] consists of discrete monofunctional enzymes. Unlike FASs, the *act* gene cluster was shown by sequencing to not contain any DH or ER domains.[41] This knowledge tied the biosynthetic route of fatty acids to those of complex polyketide antibiotics for the first time, and all biosynthetic genes in the *act* (actinorhodin) cluster have since been characterised and most of the biosynthetic pathway has been elucidated (**Fig. 1.9, bottom**).[43]

Studies on the biosynthesis of the macrolide antibiotic erythromycin A (**17**) constitute the other landmark example of linking the PKS and FAS pathways, as well as becoming the most studied example of 'modular' PKS. Several genes within the producer, *S. erythraea,* were found to have homology to FAS genes by "walking" along the genome nearby to genes likely associated with resistance to erythromycin.[40] The multifunctional mega-synthase 6-deoxyerythronolide B synthase (DEBS) was first discovered in this way by Cortes *et al.* and found to have high homology to type I FASs.[40] The DEBS PKS has since

been found to be made of three proteins, DEBS 1, 2 and 3, which are encoded by three separate genes (*eryA I, II* and *III*) (**Fig. 1.10**), with each protein comprising two modules.
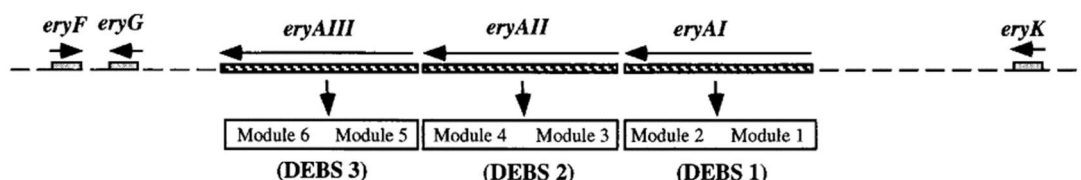
Figure 1.10– **The erythromycin (17) macrolide core 6-dEB (16) is constructed by the type I modular 'DEBS' PKS consisting of DEB1, DEB2 and DEB3 proteins, encoded by the** *eryAI*, *eryAII* **and** *eryAIII* **genes respectively.**

1.3.4 DEBS domain architecture, function and prediction of modular PKS products

Following the discovery of the erythromycin biosynthetic gene cluster, the composition and the workings of DEBS proteins have been investigated for many years.[22] A key study to understanding the enzymology of the DEBS system will be further explained in **Section 1.4.1**, this involved the swapping of domains and modules and scrutinising the aglycone products which were consequentially produced. As well as DEBS gene manipulation, *in vitro* reconstitution of the DEBS proteins to produce **16** has provided valuable insight into modular polyketide biosynthesis. In particular was the work of Lowry *et al.* which was an early example.[44]

The remarkable correspondence between the structure of **16** and the function of domains present in each PKS module as predicted by DNA sequence analyses (**Fig. 1.7**) was noted earlier on, over the years it has since become known as the 'co-linearity' rule and is respected for the majority of known modular polyketide biosynthetic pathway.[45] Nowadays several bioinformatic tools are utilised in conjunction with genome sequencing analyses to identify PKS biosynthetic gene clusters and predict the structure of the corresponding metabolites.[46,47] For example, gene mining of an *Aspergillus* fungus for genes associated with terpene and polyketide substrate biosynthesis revealed a putative

BGC harbouring a polyketide-diterpenoid product. Reconstitution of the cluster in a heterologous host lead to the production of novel hybrid terpene polyketide products.[48] An important aspect of natural product biosynthesis from a bioinformatics point of view is that polyketide biosynthetic genes in bacteria and fungi are clustered together. Nonetheless, several challenges related to PKS biosynthetic investigation remain; domains in modules can occasionally be inactive, even entire modules can be skipped leading to unobvious products. Iterative PKSs are also cryptic owing to the unpredictability of products as discussed with dynemicin (**Scheme 1.4**).

1.3.5   Use of SNAc thioesters in polyketide biosynthetic investigation

Over the years, chemical probes have proved valuable tools to investigate the substrate specificity and processing of PKSs in their own right and in complementation to molecular biology and protein chemistry studies. The early elucidation of the coenzyme A and pantetheine cofactor structures has led to the use of a strategy for studying polyketide biosynthesis with synthetic substrates that mimic cofactor-bound substrates and intermediates. A heavily truncated form of the pantetheine arm has been used to deliver putative polyketide substrates as an artificial mimic for the corresponding natural coenzyme-A form (**Fig. 1.11**): this is an *N*-acetylcysteamine (SNAc) moiety to which substrates and intermediates can be anchored to as thioesters.
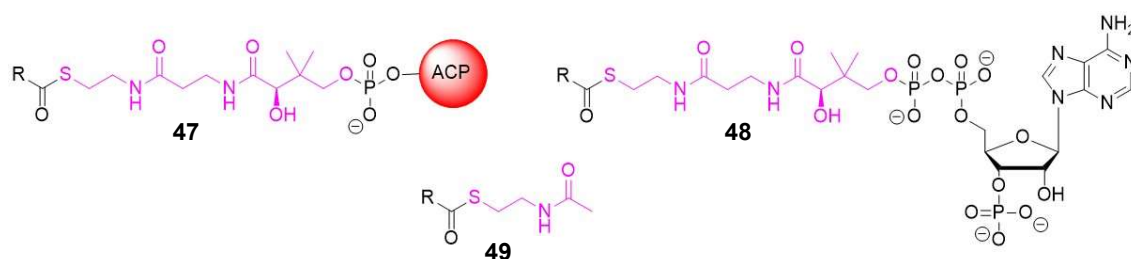
Figure 1.11 – **Structure of an intermediate bound to coenzyme-A (48) and *holo*-ACP (47), both include the pantetheine arm (pink). The "SNAc" thioester substrate (49) is a heavily truncated synthetic form which can be used to provide synthetic substrates into PKSs as a coenzyme-A mimic.**

This approach was first demonstrated by Cane *et al.*[49,50] for the study of 6-dEB assembly in *S. erythraea*. Synthetic erythromycin starter units with ^{13}C or deuterium labels were prepared as *N*-acetyl cysteamine (SNAc) thioesters: incorporation of the labels was found in the final erythromycin scaffold when these substrates were fed to cultures of *S. erythrae* (**Scheme. 1.10**).[49,50]
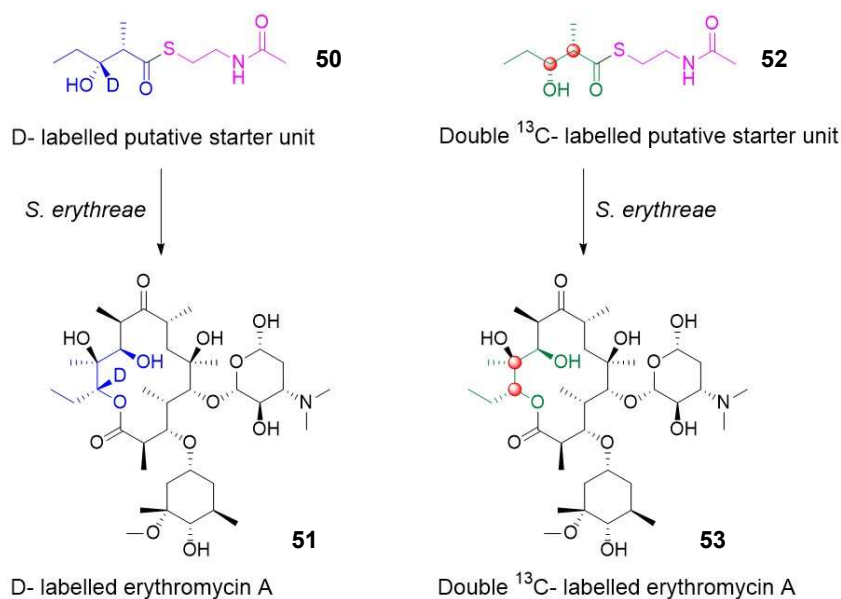
Scheme 1.10 – **Feeding of a deuterium-labelled putative starter unit as an SNAc thioester (50) to *S. erythraea* leads to incorporation in the final product (51). Similarly, the doubly ^{13}C labelled analogue (52) is processed to give double ^{13}C labelled erythromycin (53).**

This practice has become common as a simplified way of incorporating synthetic intermediates into polyketide products. SNAc substrates have also been to confirm the stereochemistry of the starter units in complex polyketide products: for instance, enantiomerically pure isotopically labelled SNAc starter units can be fed into a PKS system and only the starter unit with the correct stereochemistry will be processed into the final polyketide product. This allows stereochemical confirmation of discreet chiral centres within a polyketide; an example of this is the study of tetronasin which will be revisited later (**Section 1.6.1**).[51] The use of SNAc thioesters has been instrumental in understanding many aspects of polyketide biosynthesis. In the next section the use of pantetheine analogues for PKS biosynthesis investigation will be briefly presented.

1.3.6 The role and structure of the ACP in polyketide biosynthesis

The critical realisation of the close link between PKSs and FASs has allowed further in-depth mechanistic studies. The *act* PKS ACP of actinorhodin biosynthesis was found to have a high sequence homology to that of the homologous FAS carrier protein from *E. coli*: this suggested a similar structure and function for the two ACPs. Solution state NMR studies confirmed little difference in chemical shifts between the *apo* and *holo* forms in both proteins.[52,53] In addition, the actinorhodin ACP was shown to be phosphopantetheinylated by the machinery in *E. coli* which normally converts its own native FAS ACP from *apo* to *holo* which suggested very close recognition between the FAS and PKS ACPs.[54] This reinforced that PKSs construct intermediates in the same fashion as FASs; fatty acid biosynthesis has been instrumental in understanding how polyketide biosynthesis occurs.

On a similar note, the prementioned strategy of using SNAc thioesters has since been built on and adapted following the discovery of ACP-bound intermediates. The entire pantetheine arm can be chemically synthesised with a polyketide intermediate attached and loaded onto the inactive *apo*-ACP *via* a well-established procedure using enzymes knows as PanK, DPCK, PPAT and Sfp (these will be further discussed in **Section 2.3.4**. This synthetic version of a *holo*-ACP is known as the "*crypto*" ACP and can be purified and manipulated *in vitro* for several purposes: for example, to investigate the involvement of biosynthetic enzymes which may act on an ACP-bound polyketide chain, or the interactions and structural features of ACP-bound intermediates with other domains. FASs have also been investigated with the use of *crypto*-ACPs given that both incorporate the ACP in their biosynthetic machinery toolbox.

An example of FAS investigation with this approach is the use of crosslinking *crypto*-ACPs to fully characterise the structure and movement of the ACP domain, which is very small and dynamic and thus hard to resolve. Worthington *et al.* showed the irreversible linking of the keto-acyl synthase (KAS) and ACP proteins in the *E. coli* fatty acid biosynthesis by post-translationally inserting a synthetic cross-linkable pantetheine arm onto the *apo*-ACP (**Fig. 1.12**).[55]
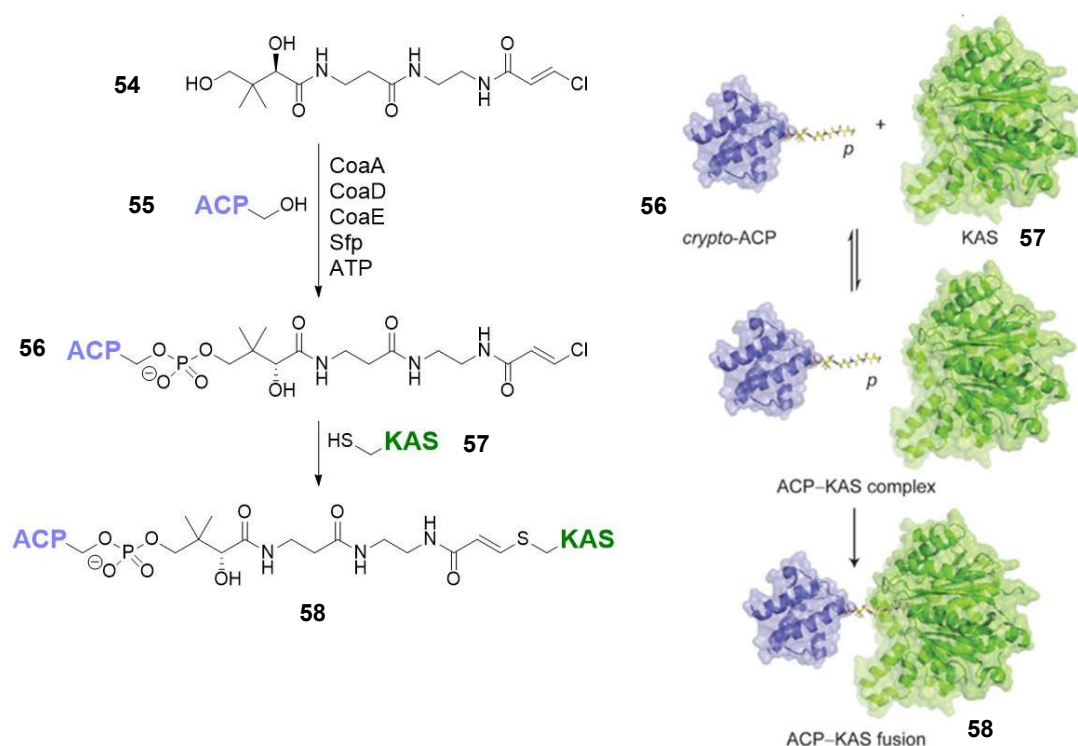
Figure 1.12 – **Synthetic pantetheine crosslinker (54) is loaded onto an *apo*-ACP (55) in a multienzyme process to give a *crypto*-ACP (56). This irreversibly binds to the active site cysteine residue of the KAS enzyme (57) forming a dimeric fusion protein (58).**

The technique for FAS was further developed resulting in a crystal structure of the ACP-FabA complex in several conformations offering a biosynthetic "snapshot" of fatty acid biosynthesis.[56] Similarly these probe crosslinkers were able to stitch together the KS and ACP domains within the DEBS system to delve into the domain-domain interactions on the type I modular PKS. Moreover NRPS systems are also being investigated with the goal of crosslinking their analogue of the ACP known as the peptidyl carrier protein PCP which is responsible for the shuttling of activated amino acids.[57]

1.4 Manipulation and 'redirection' of natural product biosynthesis

In parallel to the research investigating how natural products are produced, there has also been great interest and effort in trying to manipulate and direct biosynthetic pathways

into biosynthesising "unnatural" products: these are effectively biosynthetic analogues of natural products and may hold equal or superior bioactivity value. Many different strategies have been employed for this, including purely genetic manipulation of biosynthetic machineries to produce new compounds, and the use of synthetic chemistry to provide exotic substrates for incorporation into biosynthetic machineries, or a mixture of both.[58] In the following paragraphs, some examples illustrating these strategies are briefly presented.

1.4.1   Genetic manipulation to modify scaffolds of natural products

DEBS has been the focus of much molecular biological research. McDaniel *et al.* was able to inactivate many of the module domains and express these modified DEBSs in *S. coelicolor* and *S. lividans*.[59] HPLC purification and analysis of culture extracts lead to the detection of numerous 6-dEB unnatural macrolides (**59-61, Fig. 1.13**).[59] These ring structures had modifications specific to what particular domain had been inactivated: for example, an enone rather than an a diol moiety within the macrolide scaffold was generated where a KR domain was inactivated and a DH added. Even some AT domains were swapped from different modules allowing the uptake of malonate units where normally methylmalonates would be incorporated. Libraries of differentiated macrolides have been prepared, which would be very difficult to access rapidly by total synthesis.
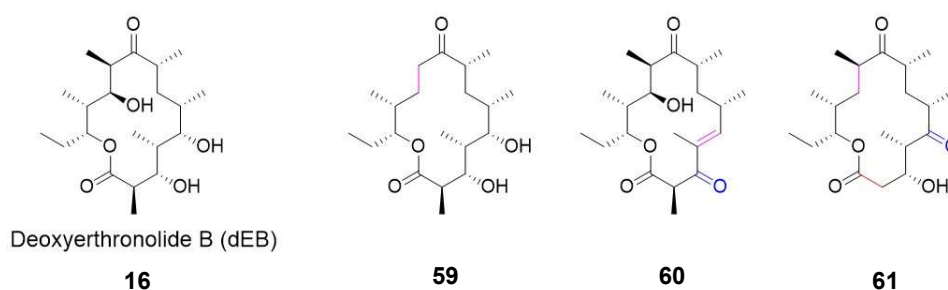
Figure 1.13 – **Over 50 unnatural analogues of the DEB macrolide (16) were described by McDaniel et al. with one, two and three points of modification. AT domain swap from module two removed a methylene group (59), deletion of a KR and addition of a DH lead to an enone group (60) and modifying several modules can lead to structures with three points of differentiation (61).**

Although AT domains were swapped around within the same PKS system, mutations in these domains can lead to highly diminished titre of products or even complete abolishment. Clearly, in this type of "plug and play" synthetic biology approach, the interactions between the different PKS modules are significantly affected.[60] In another example of manipulating DEBS for novel compound production, the DEBS1 protein was fused to the TE domain from the final DEBS3 PKS: this was shown to release a triketide product in the form of a 6-membered lactone (**62, Scheme 1.11**).[61] These experiments have paved the way to more sophisticated approaches to unnatural products in a multitude of natural product assembly lines.
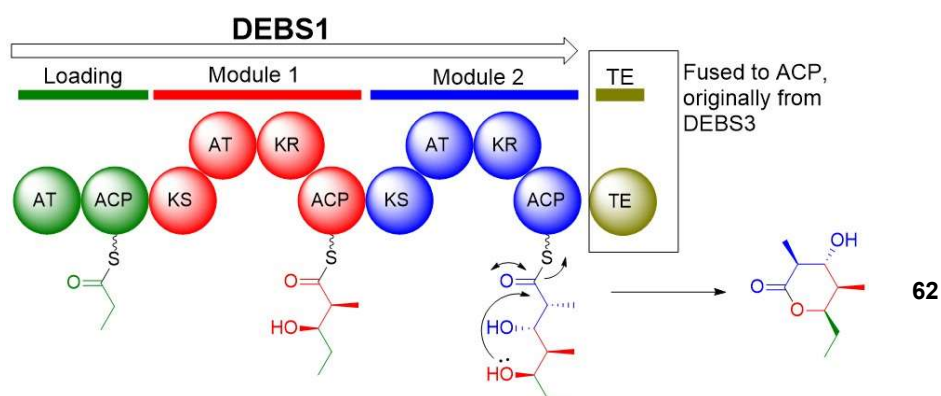
Scheme 1.11 – **Fusing modules one to three of DEBS1 with the terminal TE from DEBS3 leads to early lactone formation and release of an unnatural product (62).**

1.4.2   Combining biosynthetic pathways for novel natural products

Another strategy arising from genetic manipulation involves the merging of two biosynthetic pathways. A landmark example has been provided by the fluorination of the anticancer agent salinisporamide A. The naturally occurring chlorine atom in salinisporamide A derives from chlorodeoxyadenosine (5'-ClDA), which is synthesised from SAM and chloride by the chlorinase salL. Fluorine is widely used in medicinal chemistry and drug design for the remarkable properties conferred to bioactive compounds.[62] In nature however, there are rare examples of fluorinated natural products;

the majority of halogenated natural products contain chlorine, bromine or less often iodine.[63] One of the few known fluorinase enzymes is flA from *S. cattleya*: this synthesises fluoroacetate (from SAM and fluoride), which can then be incorporated into other metabolites.[64] When the chlorinase involved in salinisporamide A synthesis in producer *S. tropica* was chromosomally replaced with the *S. cattleya* fluorinase, the novel product fluorosalinisporamide was produced (**Scheme 1.12**).[65] Moore *et al.* had previously shown that supplementing synthetic fluorodeoxyadenosine (5'-FDA) to the producer organism with the inactivated chlorinase also produces fluorosalinisporamide.
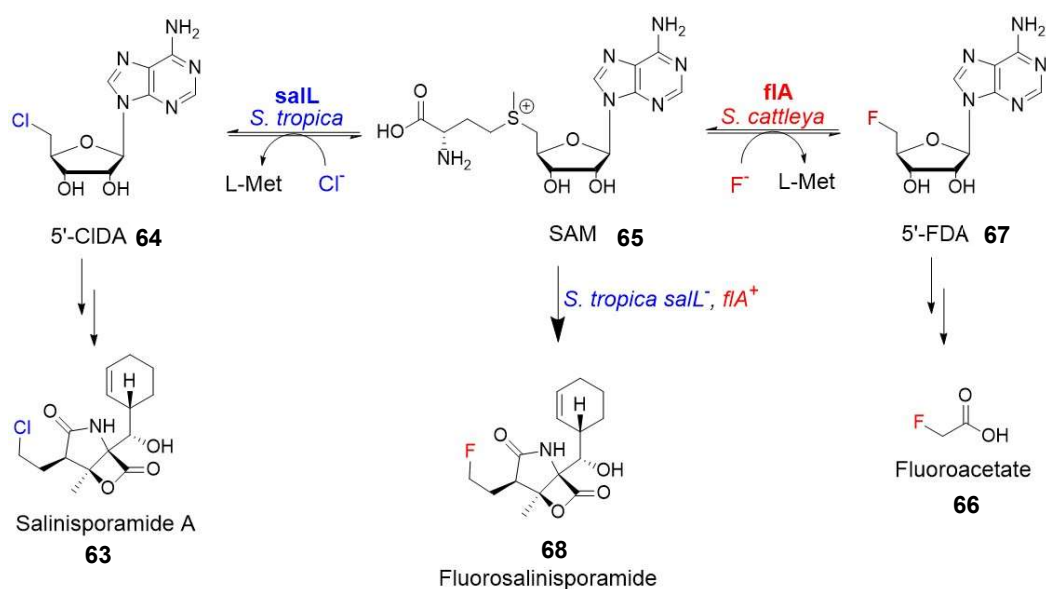
Scheme 1.12 – **Salinisporamide A (63) is produced by *S. tropica* using 5-ClDA (64), which is generated by salL using SAM (65) and chloride ions. *S. cattleya* makes the fluorine analogue 5-FDA (67) from SAM (65) and fluoride ions with flA, 5-FDA is used to produce fluoroacetate (66) which is further processed into more complex molecules. Inactivation of salL and addition of flA into *S. tropica* leads to production of fluorosalinisporamide (68).**

It is considerably simpler to direct natural product biosynthesis from more reactive to less reactive halogens as the both bromide and iodide undergo the same oxidative processes to form hypohalides as chloride ions do *via* chlorinases.[63,66] Many chlorinase enzymes also utilise (albeit less efficiently) bromide and iodide ions, so supplementation of culture

growth media with these salts can lead to production of natural product containing the lower reactivity halogen analogues.[58]

1.4.3 Synthetic alteration of natural products (Bio-chem)

Historically, many unnatural products have also been produced by semi-synthesis, whereby synthetic chemistry has been used to modify natural product scaffolds. This has been particularly important in the development of new penicillin and cephalosporin analogues to overcome bacterial resistance or improve drug properties. The semi-synthetic cephalosporin compound cephazolin is made *via* fermentation of the fungal product cephalosporin C (which contains the pharmacophore β-lactam scaffold) followed by chemical alterations (tetrazole and thiadiazole heterocycle formation) (**Scheme 1.13**).[67]
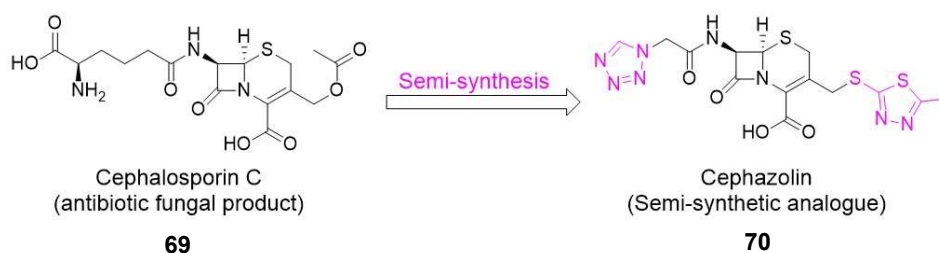
Scheme 1.13 – **Cephalosporin C (69) is a fungal antibiotic which is fermented on a large scale and then chemically converted to cephazolin (70).**

Artemisinin is an antimalarial plant natural product derived from terpene biosynthesis. Many semi-synthetic analogues have been made directly from the completed product which also exhibit higher potency and other bioactivities (**Scheme 1.14, top**).[68,69] Artemisinin is now manufactured on large scale from an analogue of a late stage artemisinin biosynthetic intermediate: artemisinic acid is produced by bioengineered yeast fermentation before undergoing several chemical steps to a semi-synthetic version of artemisinin (**Scheme 1.14, bottom**).[70] This is of particular importance given the previously fragile production of this compound from harvesting the producing plant.
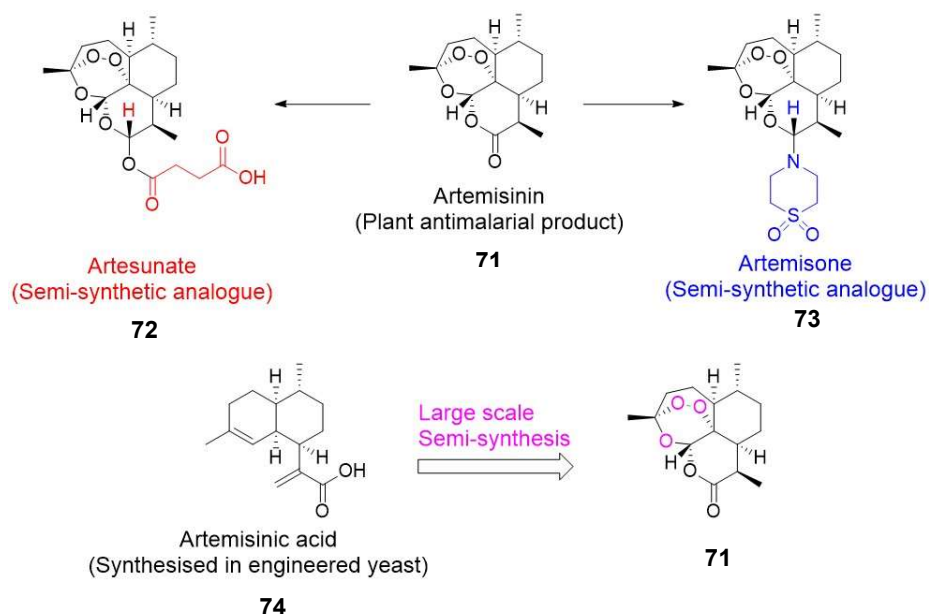
Scheme 1.14 – **(Top) Several semi-synthetic analogues of artemisinin (71) have been produced by chemical alteration of the complete natural product. These include artesunate (72) and artemisone (73). (bottom) Semi-synthetic artemisinin (71) is now chemically produced on large scale from artemisinic acid (74) which is fermented from bioengineered yeast.**

1.4.4 Biotransformation of synthetic compounds to novel products (Chem-bio)

New compounds can be made by flexible biosynthetic machineries to utilise synthetic substrates rather than their native ones. Indeed, some biosynthetic enzymes maybe display relaxed substrate specificity and may be used to generate new compounds when provided with a variety of substrates. Genetic manipulation may be required to allow or help this. Some polyketide synthases have been mutated to increase substrate tolerance. This is the case for the type III PKS enzyme HsPKS1 from the plant *Huperzia serrata,* which ordinarily produces the chalcone product naringenin. As well as directing the biosynthesis of purified HsPKS1 with synthetic coenzyme-A starter units to produce novel alkaloid/PKS products, a mutant with a structural modification in the KsPKS1 enzyme (S348G) vastly increased substrate tolerance to synthesise several diverse products in larger titres (**Scheme 1.15**).[71]
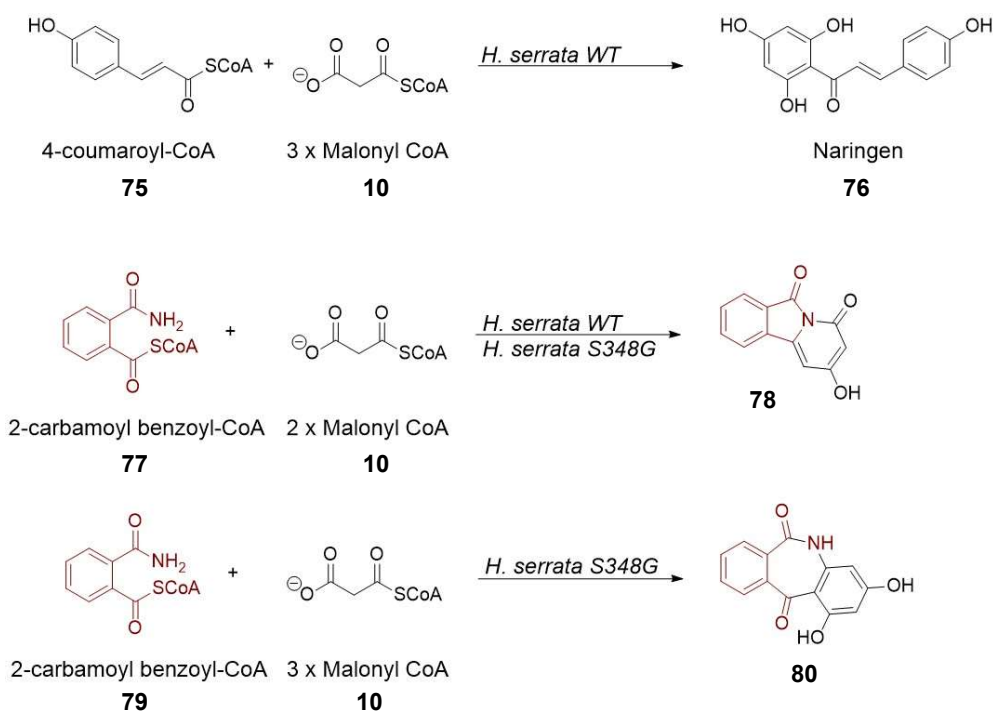
Scheme 1.15 – H. serrata normally produces the chalcone PKS product naringenin (76) using malonyl-CoA (10) and coumaryl-CoA (75), a novel indolone product (78) is produced when supplemented with 2-carbamoyl benzoyl-CoA (77). Incubation with the S348G mutant PKS gave higher titres of the indolone product along with another novel alkaloid-like product (80).

1.4.5 Chemical probes for natural product biosynthesis investigations

The final area of discussion regarding understanding and manipulating biosynthesis concerns the use of chemical probes. The FAS ACP and KS crosslinking agents mentioned previously (**Fig. 1.12**) are examples of how synthetic pantetheine derivatives were used to interfere with the ACP function to answer questions about protein interactions with other domains.[56] This has been adapted to investigate NRPS systems as shown by Bloudoff *et al.* by linking the C domain with synthetic probes which are non-hydrolysable: this was pursued in order to understand the exact mechanism of how this

domain catalyses amide bond formation **(Scheme 1.16)**,[72] and co-crystallisation of the substrate with the C domain revealed information on substrate-protein interactions.
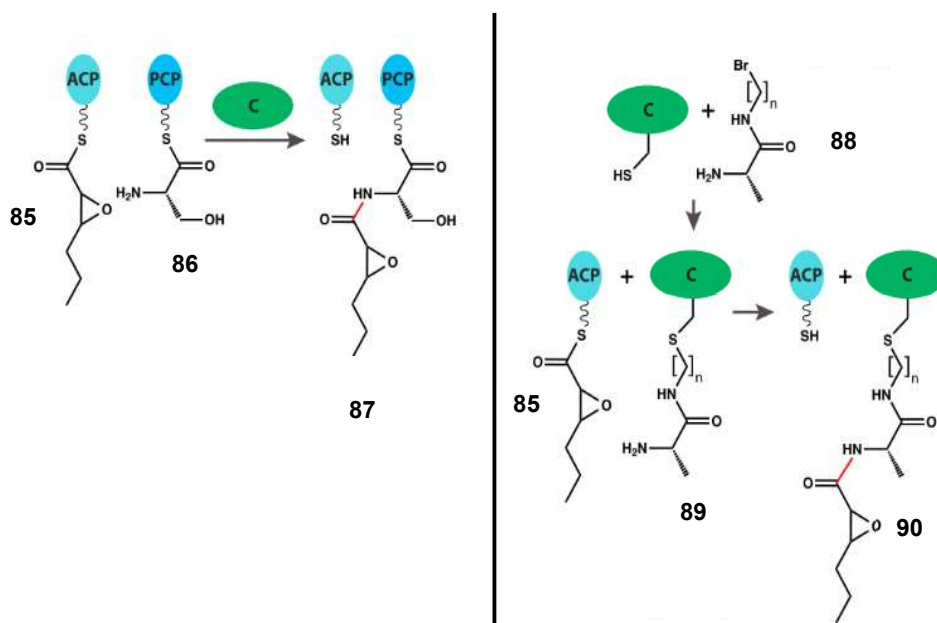
Scheme 1.16 – (*Taken directly from reference*)[72] Condensation of a polyketide (85) and an amino acid (86) to product (87) is mediated by the C domain (left). Synthetic amino acid substrates (88) are C-bound (89) and undergo the same condensation to deliver the non-hydrolysable coupled product (90) for crystallisation studies (right).

SNAc thioesters are effectively used to probe substrate specificity of PKSs in terms of whether unnatural SNAc substrates can be turned over. A particularly interesting example was provided by the investigation of the biosynthesis of borrelidin from *S. rochei*. Curran *et al.* were specifically interested in the PKS fifth module, which is iterative and fully reducing and recruits methylmalonyl as the extender unit **(Scheme 1.17)**.[73] This single module was expressed and supplemented with the relevant coenzyme-A substrates: under these conditions a variety of SNAc compounds were elongated to give unnatural products.
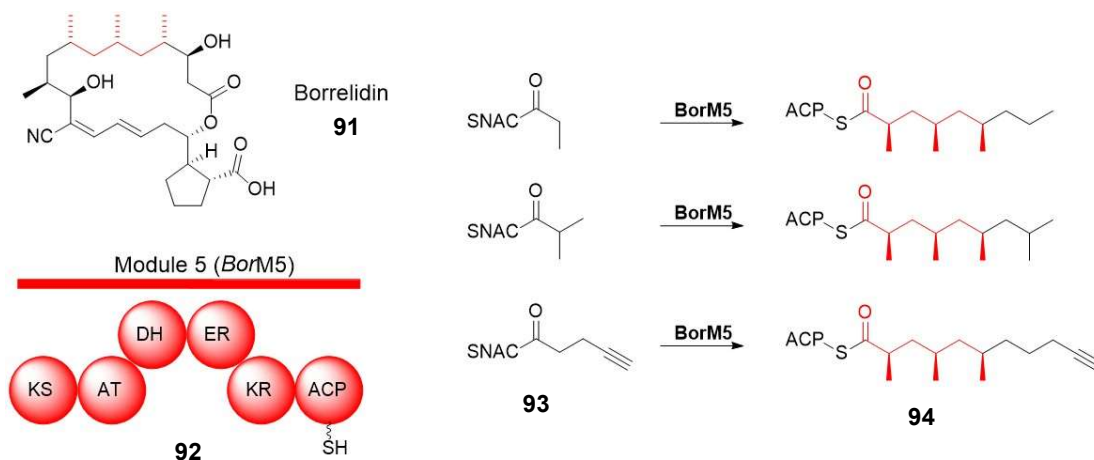
Scheme 1.17 – **Borrelidin (91) is a type I modular PKS product but part of its ring structure (red) is made by the iterating module BorM5 (92). Incubation of unnatural SNAc substrates (93) with the standalone module leads to three chain iterations (94) as observed by intact protein MS.**

1.4.6   Chemical probes for off-loading polyketide intermediates

Several examples of SNAc substrates or other chemical probes similar to those described in the previous paragraph exist. The following section will focus on the description of novel chemical probes developed by Tosin *et al.* which intercept polyketide biosynthetic intermediates. This concept has been developed from early work on malonyl coenzyme A mimics used *in vitro* for the study of type III PKSs, and refined over the years to generate small molecule lipophilic probes for exploration of type I and II polyketide systems *in vivo*. The idea behind this approach is that captured biosynthetic intermediates during polyketide assembly constitute "snapshots" of PKSs in action. In general, only the final products of polyketide biosynthesis can be isolated, and the nature of the biosynthetic intermediates is mostly predicted.

1.4.6.1   Malonyl-CoA mimic probes for offloading type III PKS intermediates

The use of 'intermediate capturing' probes was first demonstrated when a carba(dethia) analogue of malonyl-CoA was incubated with a type III PKS and shown to off-load

enzyme-bound intermediates. The synthetic malonyl-CoA surrogate undergoes a similar decarboxylative Claisen condensation reaction as the native ACP-malonate substrate normally involved in chain elongation (**Scheme 1.1**). The intercepted intermediates, which have the canonical thioesters replaced with methylene ketone moieties, cannot be reloaded onto the PKS enzymes, hence they can be detected by liquid chromatography mass spectrometry (LC-MS) analyses of enzymatic extracts, and further characterised by high resolution (HR) tandem MS (MS^2) (**Scheme 1.18**).[74]
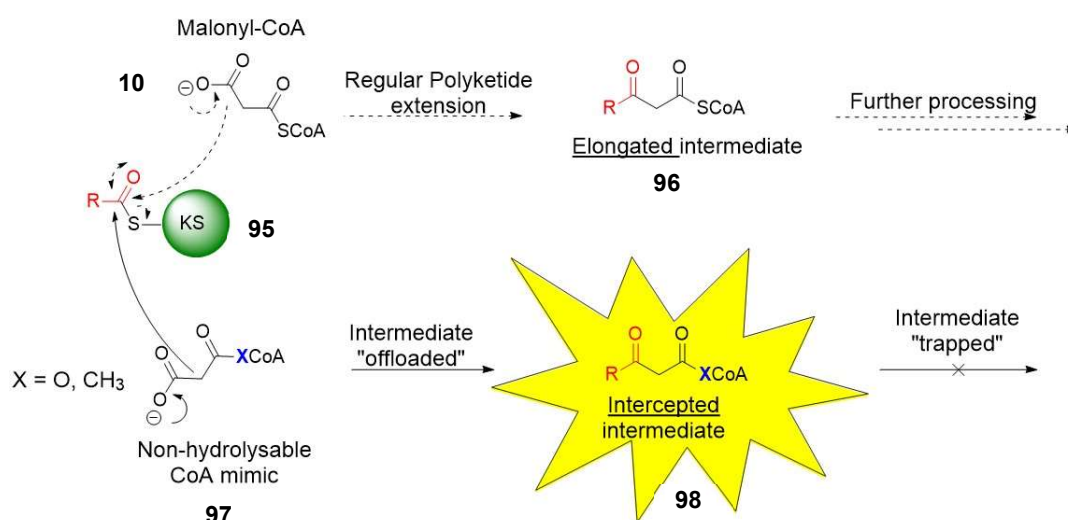
Scheme 1.18 – **(Top) Normal polyketide extension from the Claisen condensation of CoA/ACP tethered malonate (10) building blocks abstracts the KS bound intermediate (95) to elongate the polyketide chain (96) which can then be further processed by virtue of the labile thioester. (bottom) A synthetic non-hydrolysable CoA mimic (97) prepared from pantetheine has a ketone/ester rather than the thioester, the decarboxylative Claisen condensation occurs but leads to an "intercepted" intermediate (98) which is not enzyme bound, this intermediate is stable and cannot be processed.**

These species offered a "snapshot" into polyketide biosynthesis in action which would be otherwise difficult to gather by other means as polyketides intermediates remain PKS-bound throughout natural product assembly and generally only end products can be isolated and characterised.[75] This proof of concept study on type III PKS enzymes *in vitro* has since been applied to the study of much more complex biosynthetic pathways, including type I modular and iterative polyketide assemblies.

1.4.6.2   Intermediate offloading in type I PKS using simplified probes

In an attempt to utilise probes that are simpler in structure compared to CoA derivatives, truncated *N*-acetyl β-keto esters (**97**) bearing some of the original functionality and pattern of carbon chain substitution were synthesised. Recombinant DEBS3 was expressed, purified and incubated with the chemical probes (*N*-acetyl β-keto acids (**98**) generated from the corresponding esters by the use of pig liver esterase) in the presence of the necessary cofactors and substrates. Putative intercepted intermediates (**99-102**) from the DEBS3 PKS were identified and characterised by LCMS. This was also confirmed using deuterium labelled probes leading to labelled intermediates.[76] The next step involved the use of the ester probes *in vivo* with several strains of the erythromycin producer *Saccharopolyspora erythraea* (e.g. wildtype and with inactivated DEBS2 and/or DEBS3 PKS genes). By relying on the action of esterases *in vivo* to generate the active β-keto carboxylates *in situ*, putative erythromycin intermediates were offloaded from the DEBS PKS and retrieved by organic extraction of the fermentation broth for LC-HR-MS[n] characterisation (**Scheme 1.19**).[77]
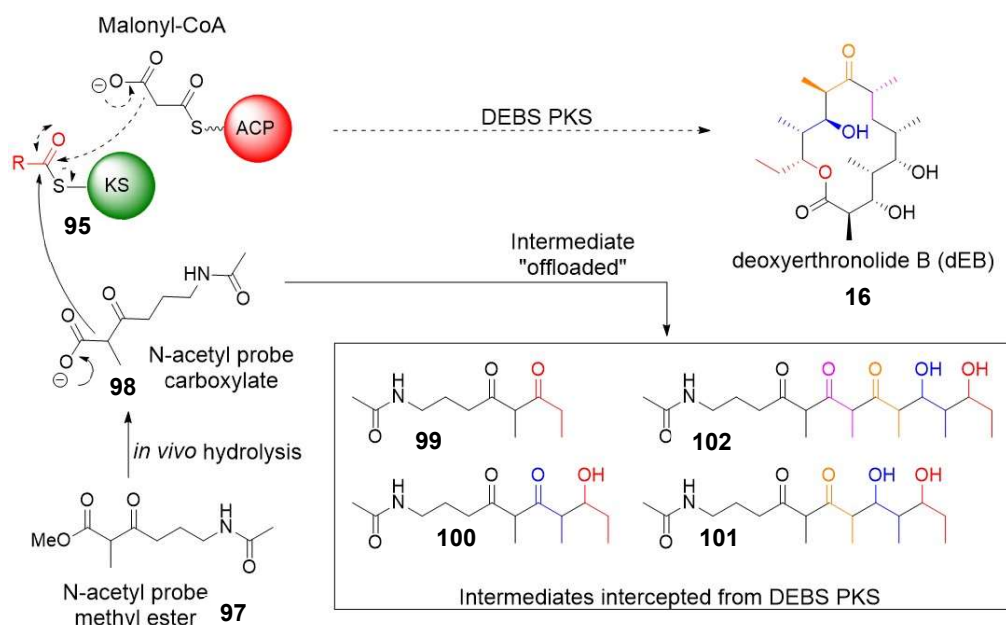
***Scheme 1.19* – Simple *N*-acetyl methyl ester probes (97) when fed to *S. erythraea in vivo* are hydrolysed by esterase to the active carboxylate (98) which intercept intermediates from KS (95) during the biosynthesis of dEB (16) by DEBS. Mono (99), tri (100) tetra (101) and pentaketides (102) were detected by MS.**

Following this proof of concept study *in vivo*, further challenging biosynthetic pathways have been investigated with these probes, such as the modular type I PKS product lasalocid A from *S. lasaliensis.* Lasalocid A belongs to a large class of polyketide products known also as polyether ionophores. These compounds show activity against protozoa, bacteria and viruses,[78] and are particularly important in ruminal animal feedstocks by improving feedstock conversion to improve their growth and health.[79] The canonical mechanism of action for ionophores is metal cation binding: metal mono/divalent cations are "wrapped" up by the ionophore in an oxygen rich cavity by chelating to furan/pyran type rings to form a lipophilic "blanket"; this allows them to be transported through the hydrophobic cell membrane of the organism which then disrupts the highly controlled membrane potential to cripple biochemical processes within the cell.[78,80] Some have further pharmaceutical interest such as salinomycin (an anticancer stem cell inhibitor lead).[81] Cation selectivity for polyether ionophore can also vary: for example, lasalocid A binds preferentially to potassium over sodium ions,[82] whereas monensin shows the reversed selectivity. Tosin *et al.* showed that the complex polyketide intermediates leading to lasalocid A formation can be offloaded *in vivo* from the *lasA* PKS by using chemical probes, providing information about the timing and mechanism of the polyether ring formation otherwise not accessible **(Scheme 1.20)**.[83, 84]

A mutant strain of *S. lasaliensis* bearing the last ACP domain of the PKS in an inactivated form has proved particularly useful for the accumulation of advanced polyether intermediates closely resembling the natural product but bearing structural modifications. Further 'second generation' probe design has indeed included long *N*-acyl chain instead of *N*-acetyl to improve cell permeability, and the featuring of chemical "handles" in the *N*-acyl chain to allow derivatisation of captured intermediates,[85] such as a "clickable" alkyne for 1,3-dipolar cycloaddition. Azido moieties have also been included for the reverse complementary cycloaddition or Staudinger reactions. The characterisation of intercepted intermediates has also provided preliminary information about the relative kinetics of each step in lasalocid A assembly based on intermediate accumulation abundance.
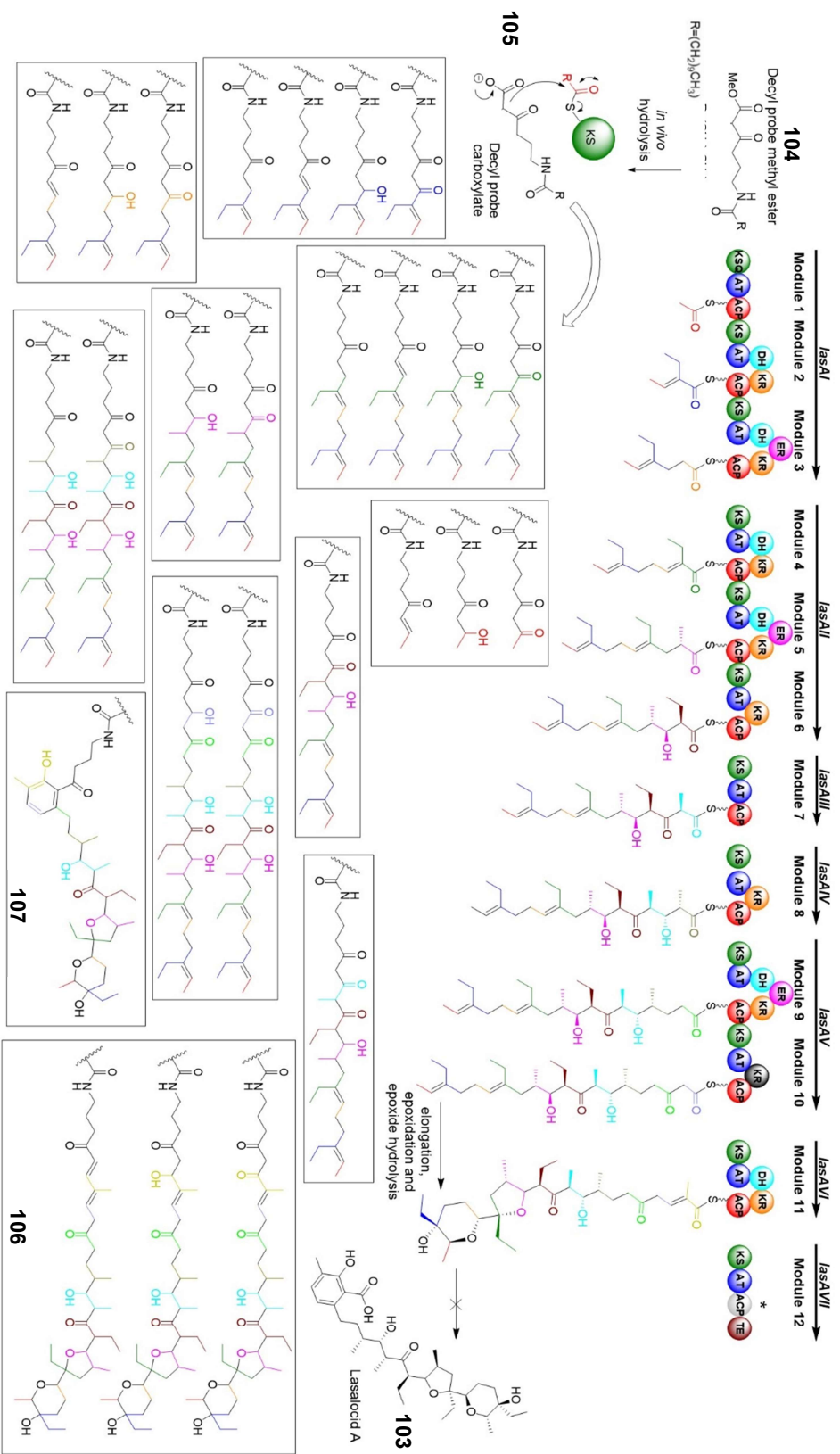
Scheme 1.20 - Lasalocid A (103) is produced by a twelve-module type I PKS in *S. lasaliensis*. A mutant strain with an inactivated twelfth module ACP was generated (ΔACP12) and cultures were fed in vivo with the long chain probe (104) which undergoes hydrolysis to the active carboxylate (105). A wide range of offloaded intermediates were detected by LC-HRMS and confirmed by MS². These include late stage polyether (106) and phenol (107) containing intermediates.

1.4.6.3 Iterative PKS investigation with "2nd generation" chemical probes

As mentioned earlier (**Section 1.2.2**), iterative polyketide systems are more difficult to decipher owing to unpredictability of chain length and timing of domain usage. The use of second generation probes have also shed light on the biosynthesis of 6-methylsalicylic acid (**38**) (by isolating biosynthetic intermediates including reduced/dehydrated mono-tetraketides from which a linear biosynthetic pathway was concluded),[86] and other 6-MSA analogues from fungal polyketide pathways.[87] Recently NRPS chain termination probes have also been designed and proved by our group to intercept peptide intermediates in the making of peptide natural product scaffolds, such as that of antitumor antibiotic echinomycin.[88] The use of both polyketide and non-ribosomal peptide probes can now allow the study of challenging hybrid PKS-NRPS enzymes.

1.4.7 Aims of this thesis

The main aim of my PhD work was to investigate challenging PKS pathways which are yet to be fully understood utilising chemical probes of different nature. Two particular PKS assemblies were the subject of this thesis work: firstly, those leading to thiotetronate compounds, which are interesting small molecules from a pharmaceutical perspective; the understanding of how exactly these products are biosynthesised would be hugely beneficial for development of new drugs, including antibiotics and antimalarials. The second pathway in exam was that leading to tetronasin, a tetronate polyether antibiotic whose assembly details were still unclear (in terms of ring formation timing and mechanism) at the beginning of my work. Background information on these two families of PKS natural products will be discussed in the following paragraphs and provide the starting point of my investigations, which will be detailed in chapters 2 and 3. In parallel to these pathway investigations, I also undertook the synthesis of a range of novel chemical probes to be used to explore PKS substrate specificity and for innovative 'unnatural product' generation. These efforts will be reported and discussed in chapter 4, before summarising conclusions and discussing future work plans in chapter 5.

1.4.8 Redirecting biosynthesis with minimal biosynthetic machinery

Another approach to unnatural products involves removing essential genes from biosynthetic gene clusters: this leads to incomplete biosynthetic intermediates which may undergo spontaneous transformations to afford 'shunt' products, often featuring new scaffolds. This has been achieved for the actinorhodin biosynthetic pathway by utilising its 'minimal' PKS machinery i*n vitro* and *in vivo* to produce the octaketides SEK4 and SEK4b. The "minimal" *act* PKS system consist of a bifunctional KS/AT, CLF and ACP (**Scheme 1.21**).[89] In the absence of any enzymes other than the bifunctional KS/AT, CLF and ACP, an octaketide polyketone intermediate is produced, which spontaneously cyclises to the aromatic shunt products SEK4 and SEK4b.

Naturally, the KR, Cyclase (CYC) and aromatase (ARO) are required to convert the octaketide to the final PKS intermediate (S)-DPNA which is then further processed and dimerised to actinorhodin. Incubation of recombinant minimal PKS in the presence of acetyl-CoA and malonyl-CoA (or malonyl CoA alone) leads to SEK product formation *in vitro.* This system has been utilised in this work for testing of chemical probes and will be further explored in **Section 4.2.3**.
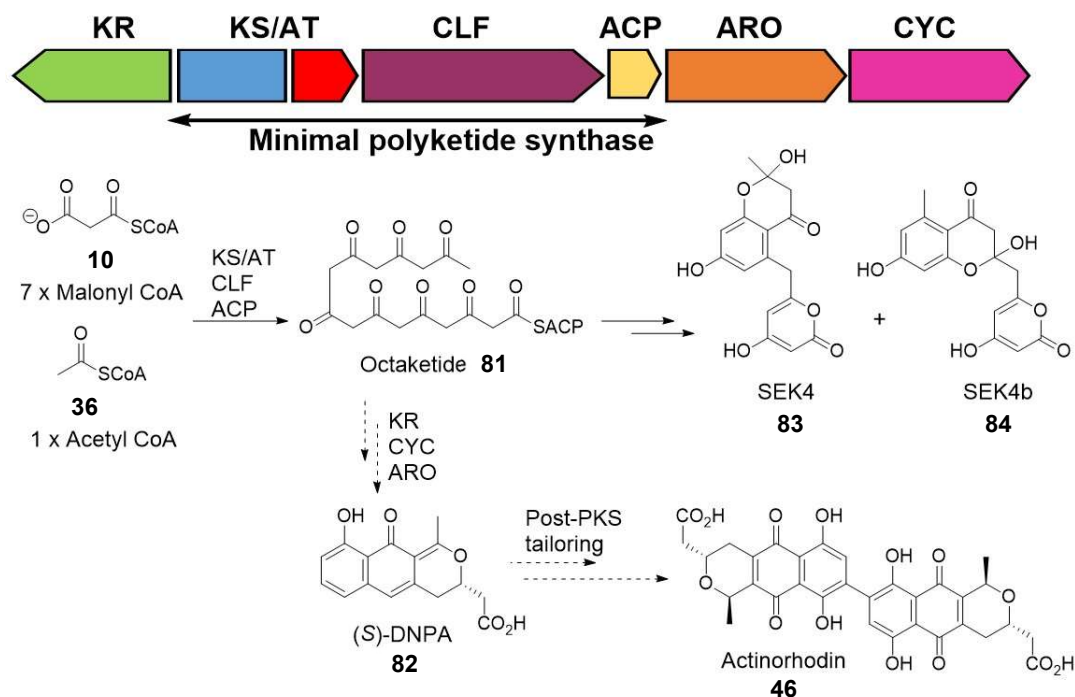
Scheme 1.21 – **The *act* gene cluster PKS includes KR, KS/AT, CLF, ACP, ARO and CYC. The KS/AT, CLF and ACP enzymes convert malonyl-CoA (10) and acetyl-CoA (36) to a linear octaketide (81) which KR, CYC and ARO convert into (S)-DNPA (82). Non-PKS enzymes convert this cyclised intermediate into actinorhodin (46). In the minimal PKS with just octaketide constructing enzymes, the SEK4 (83) and SEK4b (84) shunt products are formed spontaneously.**

1.5   Thiotetronate natural products

1.5.1   Thiotetronate discovery and activity

Thiotetronates are a family of natural products which all possess a thiotetronate ring. The most well-known, being the earliest discovered thiotetronate, is thiolactomycin (**108, Fig. 1.14**). Isolated from a Japanese soil sample in 1981, this molecule is produced by a *Nocardia* sp. strain, which has since been renamed *Lentzea* sp.*,* and it exhibits broad spectrum antibiotic activity [90] *in vitro* and *in vivo* against a variety of microbes.[91]

Its structure was elucidated with the use of X-ray crystallography and NMR spectroscopy.[92] Since its discovery, other analogues have been isolated from other actinobacteria: for example, thiotetromycin (**109**) was discovered soon after from *Streptomyces* sp. OM-674 and Tü3010 (**110**) is produced by *Streptomyces olivaceus*.[93] Interestingly thiolactomycin was recently found to be produced also from marine bacteria of the *Salinispora* genus such as in *S. pacifica*.[94]
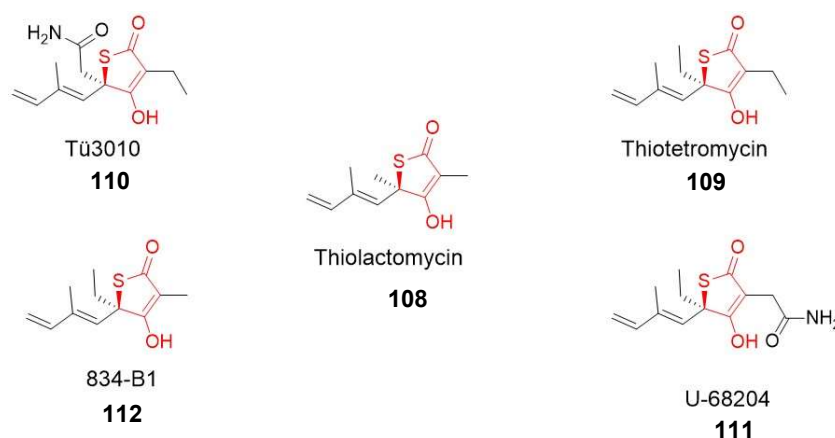
Figure 1.14 – **The family of thiotetronate natural products all contain a thiotetronate ring (red) with various substituents around the central scaffold, thiolactomycin is the most well studied (108).**

The activities of thiotetronate compounds have been extensively studied since their discovery, with many different gram positive and negative bacteria destroyed or inhibited by their effects. Examples of a few include tuberculosis causing *Mycobacterium tuberculosis*, malaria causing *Plasmodium falciparum* and also *Trypanosoma brucei* which causes the African sleeping sickness disease.[95–97] There is moderate pharmaceutical interest in the thiotetronate class as they have a number of medicinally valuable properties such as a good bioavailability and low toxicity as shown in mice.[98] The proven mechanism of action of these thiotetronate compounds is by interfering with the biosynthesis of fatty acids by reversible inhibition of type II FAS which exists only in archaea and bacteria; no activity of these compounds is observed against the mammalian type I FAS.[93] Synthetic analogues of thiotetronates have also been produced

which target type I FAS and lead to weight loss in mouse model assays as well as exhibiting cytotoxicity to human breast cancer cells.[99] Thiolactomycin is observed to inhibit the β-keto acyl-acyl carrier protein synthases in type II fatty acid biosynthesis,[100] these are colloquially referred to as the FabB/F enzymes and are responsible for mediating the decarboxylative Claisen condensation to elongate the fatty acid (**Scheme 1.22**).
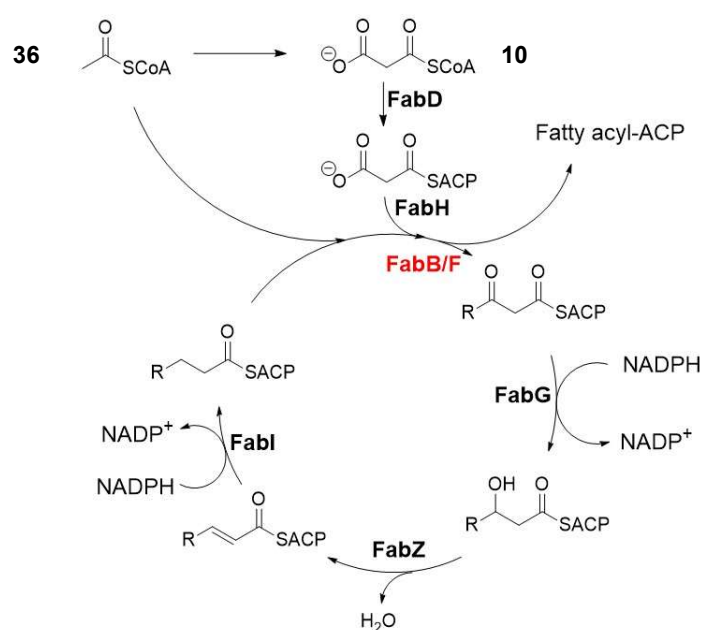
Scheme 1.22 – **Type II fatty acid biosynthesis proceeds from acetyl-CoA (36) and malonyl-CoA (10) to continuously elongate a fatty acid chain mediated by standalone Fab enzymes. Thiolactomycin reversibly binds the FabB/F enzymes (red), inhibiting the decarboxylative Claisen condensation step.**

Mycobacterium species possess homologues of FabB and FabF known as the Kas A/B enzymes which mediate the same elongation step; these similarly are inhibited by thiotetronates and prevent formation of crucial cell wall component mycolic acid (**Fig. 1.17, top**).[101,102] Thiolactomycin (**108**) is believed to act as a mimic for the ACP-bound malonate extender unit and competitively binds into the enzyme pocket to prevent both substrates binding and therefore elongation of the fatty acid chain (**Fig. 1.17, bottom**).[100]
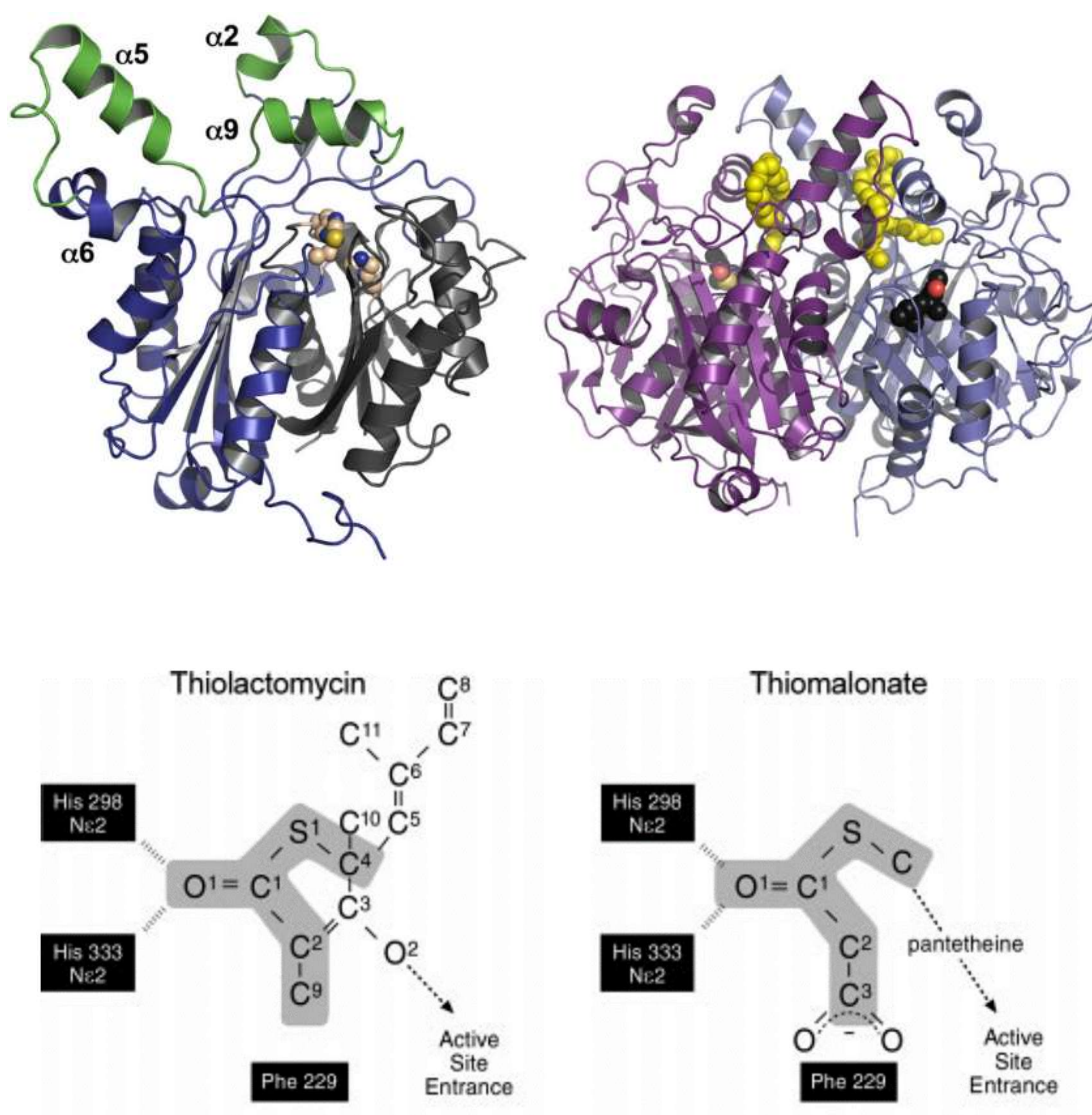
Figure 1.15 – **(Top left) Crystal structure of KasA monomer, a type II FAS from *M. tuberculosis*. The cysteine, histidine and histidine catalytic triad residues are shown in a brown space filling model. (Top right) Crystal structure of the KasA dimer bound to thiolactomycin, PEG fills the fatty acid pocket shown in a yellow space filling model and thiolactomycin is shown in a black space filling model. (Bottom) Diagram illustrating the binding of thiolactomycin to a FabB enzyme active site as a mimic for the native ACP malonate substrate.**

Other members of the thiotetronate family also unsurprisingly display an identical range of activities with variable efficacy; for instance, **110** has been shown to have a fifteen fold increased antibacterial activity compared to that of thiolactomycin and thiotetromycin.[103]

1.5.2 Biosynthetic origins of thiotetronate natural products.

Despite the properties and activities of thiotetronate compounds being extensively studied since the mid-1980s, there had been little definitive proposal of their biosynthetic origins. One of the earliest conclusions was made in 2003 by Brown *et al.* who deduced that the chemical backbone of thiolactomycin might be made by a polyketide synthase utilising an acetate starter unit and three extender units of methylmalonate on the basis of feeding experiments with ^{13}C labelled acetate and propionate; other thiotetronates also fitted this hypothesis by utilising an ethylmalonate building block instead.[104] They also speculated that the tetraketide backbone (**115, Scheme 1.23**) could be built using these building blocks, at which stage a number of pathways of S insertion could lead to the ring formed product such as a thiirane formation or through a persulphide thioester intermediate. *Lentzea* sp. fermentation broths supplemented with ^{35}S-labelled cysteine produced labelled thiolactomycin, this supports the idea that desulphuration of cysteine is likely source of the sulphur atom in the final natural product, possibly *via* a thiirane or a persulphide thioester intermediate.
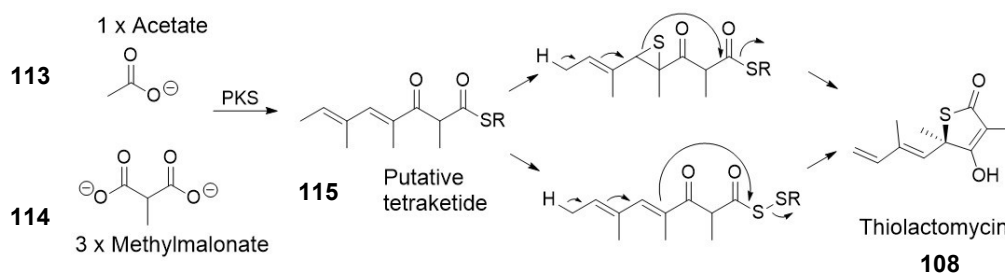
Scheme 1.23 – **Hypothesis of thiolactomycin (108) biosynthesis: thiolactomycin backbone (115) could be constructed by a PKS using acetate (113) and methylmalonate (114) before an unknown mechanism of sulphur transfer from cysteine and ring closure.**

In 2015, two groups used different genomic mining approaches to discover that thiotetronates are produced by hybrid PKS-NRPS enzymes. Tao *et al.* sequenced the genomes of several thiotetronate producing bacteria and found all of the genomes to contain a conserved BGC for a putative PKS-NRPS. When selected genes or indeed the

whole cluster was inactivated, production of their respective products ceased as confirmed by LC-HRMS.[93] Several further deletion experiments identified the limits of each cluster and the essential genes within each cluster. As well as the essential PKS-NRPS, several other genes including a putative cytochrome P450 monooxygenase, cysteine desulphurase and a sulphur transferase were proven to be crucial for production. A self-resistance gene was also found in each of the clusters translating unsurprisingly to an additional FabB/F protein involved in type II fatty acid production (**Fig. 1.16**).[93] Simultaneously, Tang *et al.* found almost identical clusters in the genomes of *S. afganiensis* and *S. pacifica* which were also confirmed by gene inactivation; similar essential genes were present in these clusters such as the putative cytochrome P450.[94]

The PKS-NRPS loading module TlmA in *Lentzea* sp. is proposed to have an AT domain specific for malonate which is then decarboxylated by the unusual ketosynthase decarboxylase domain (KSQ) to effectively give an acetate starter unit. Unlike previous belief that the PKS was modular, TlmB was proposed to be an iterative modular PKS with a methylmalonate specific AT domain: these extender units would decarboxylate to generate the polyketide chain, and partial reduction and dehydration by the KR and DH domains could give rise to a postulated tetraketide diene backbone. The NRPS region is predicted to have a cyclisation (Cy) domain normally important for oxazole/thiazole formation, an A domain specific for cysteine and terminal PCP and TE domains.
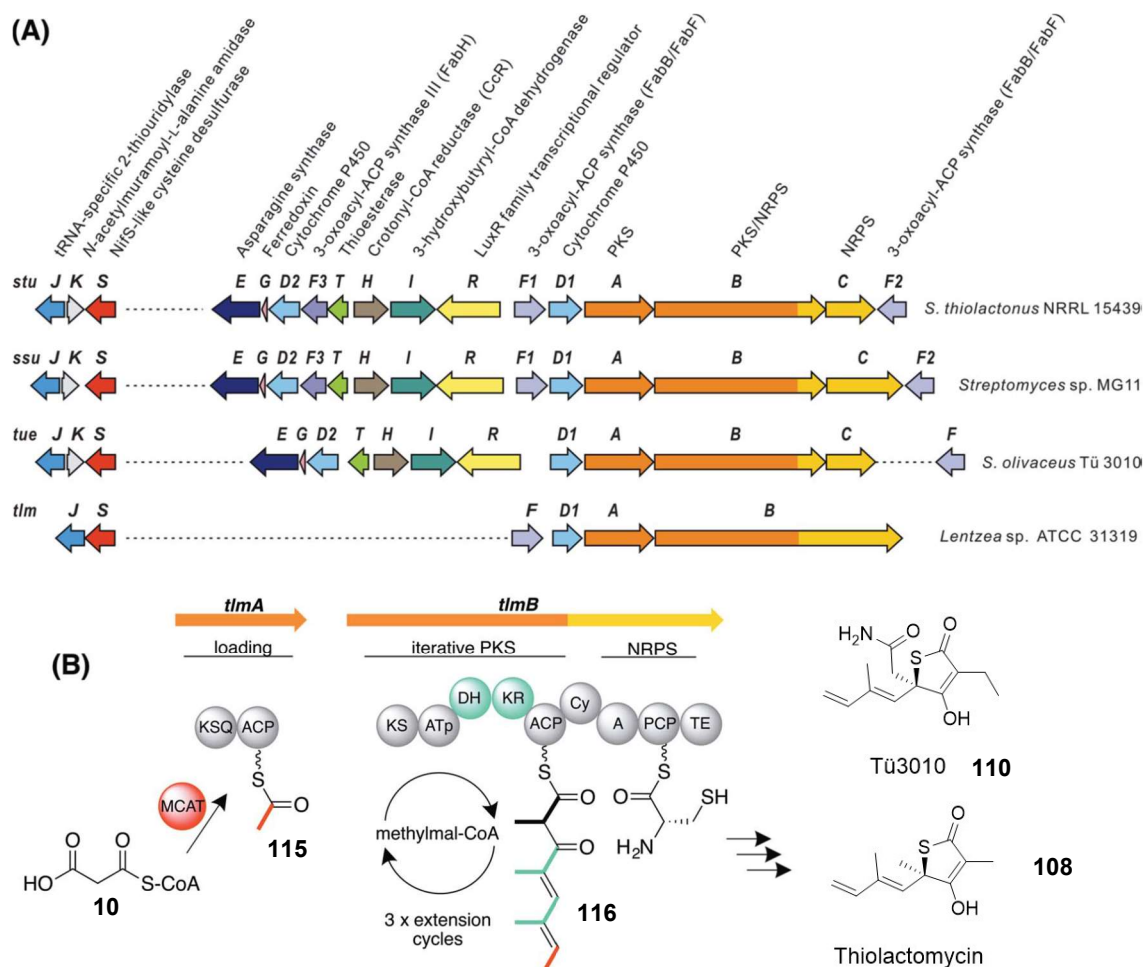
Figure 1.16 – (A) Sequence alignment of four thiotetronate producing actinobacteria. Thiolactomycin (108) is produced by Lentzea sp. whilst the other three strains produce Tü3010 (110). Lentzea sp. presents the simplest cluster which contains the hydrid PKS-NRPS, a putative cytochrome P450 (TlmD1), a putative cysteine desulphurase (TlmS), a putative sulphur transferase (TlmJ) and the self-resistance FAS gene (TlmF). (B) Proposed thiolactomycin PKS-NRPS where TlmA is a loading module specific for malonyl-CoA (10): the latter is decarboxylated to acetate (115), which the iterative PKS-NRPS TlmB can extend with methylmalonyl-CoA (15) to a tetraketide backbone (116). The latter is then converted to thiolactomycin by a yet unknown mechanism of sulphur insertion and cyclisation.

Based on the essential putative cysteine desulphurase, sulphur transferase and CYP450 monooxygenase enzymes found, a mechanism of sulphur insertion and ring closure has been proposed, whereby a tetraketide diene intermediate (**116**) would be epoxidised (**117**) and converted to a thiirane (**118**) to allow for thiolactone formation as speculated by

Brown *et al.* previously. The sulphur would be introduced as persulphide in a catalytic cycle loading cysteine and releasing alanine mediated by TlmJ and TlmS (**Scheme 1.24**).[93]
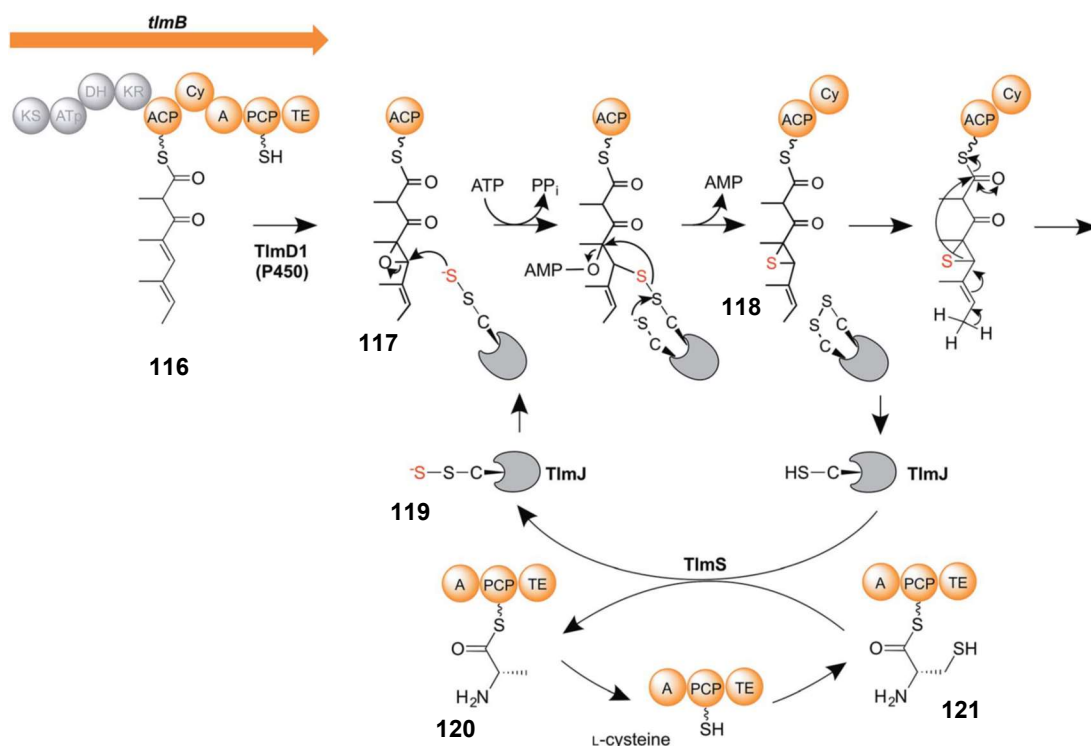
Scheme 1.24 – **Proposed mechanism of sulphur insertion and thiolactone formation from a tetraketide intermediate (116). Epoxidation (117) and thiirane formation (118) of the backbone would allow the intermediate to simultaneously be released from the enzyme and form the thiolactone ring. Cysteine is proposed to be loaded onto the PCP (121), desulphurised to alanine (122) and released using the TE domain. TlmJ acts as the persulphide carrier (119) being replenished by TlmS.**

The Cy domain in the NRPS region of TlmB is postulated to mediate the collapse of the thiirane ring and formation of the thiotetronate ring, which would be in contrast to the typical function of these domains. An extensive review of carbon-sulphur bond formation in sulphur containing natural products appears to suggest this is a novel mechanism of insertion which needs further investigation.[105,106]

1.5.3 Chemical synthesis of thiotetronate derivatives

Despite the relatively low molecular weight (MW) and minimal chirality of thiotetronate compounds, their chemical synthesis has proved challenging. The vast majority of syntheses has been focused on thiolactomycin being the simplest compound of this class. The first racemic synthesis was reported in 1984 where the thiotetronate core was assembled by a self-condensation between a thiolate and an ester (**128**) followed by alkylation with its side chain giving the final compound in a 10% overall yield (**Scheme 1.25**).[107]
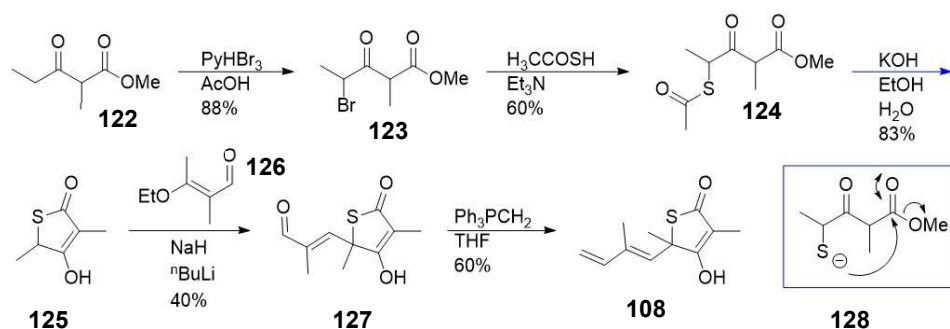
Scheme 1.25**– **First total synthesis of racemic thiolactomycin (108) is a five-step procedure from a simple keto-ester (122) starting material with an overall yield of 10%. The synthesis is centred around a base- catalysed ring closure via the hydrolysis of a thioester (124) which reveals a thiolate (128) and in turn self-cyclises onto the methyl ester to close the ring (125).

This procedure (centred around base-catalysed ring closure) has been improved and utilised many times to generate analogues with more elaborated ring decorations.[108] The main limitation of this procedure is the lack of chiral control and the poor yield concerning the side chain addition. Chambers *et al.* published the first enantioselective synthesis of thiolactomycin a few years after reporting a different procedure which used (*S*)-lactate to fix the chirality and built the sidechain before forming the thiolactone core later.[109]

A much more flexible, high yielding and asymmetric procedure was developed by McFadden *et al.* in 2002.[110] This route ensured total stereochemical control by the use of an amino acid derived starting material. This procedure installed the sidechain group before closing the thiotetronate ring in the final step with a Dieckmann cyclisation (**Scheme 1.26**). Even with more steps, the route contained higher yielding reactions using cheaper materials and allowed easier functionalisation for synthesising analogues.
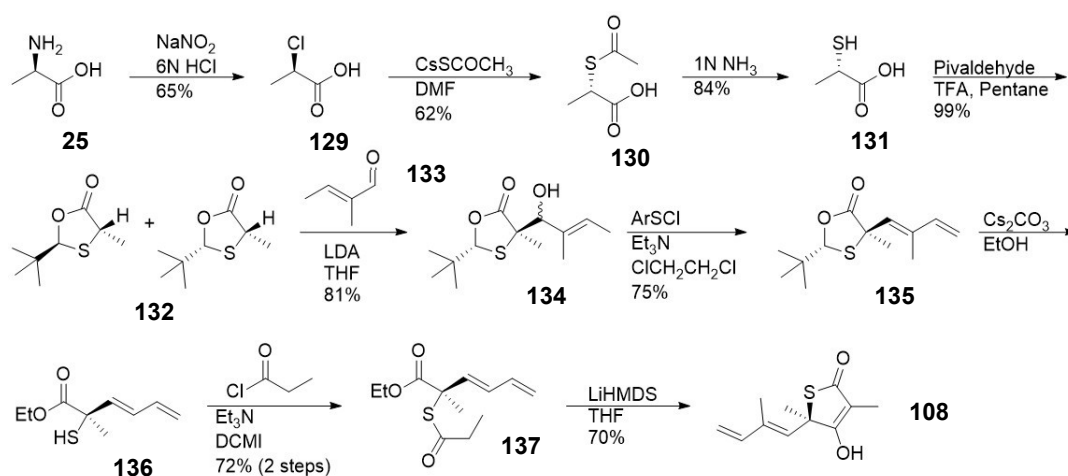
Scheme 1.26 – **Asymmetric total synthesis of thiolactomycin (108). The chiral centre is introduced in the starting material thiolactic acid (131) produced from *D*-alanine (25). All functionality is installed before the final step Dieckmann cyclisation.**

The two routes herein discussed are the two most typical synthetic routes to producing thiotetronate scaffolds. Since these were established, there has been little methodology development and much of the synthetic work has been focused on developing analogues in an attempt to improve properties or potency. Because of this, there is clearly a need and interest in further understanding of how such a small and complex natural product is biosynthesised, this may pave the way for chemoenzymatic methods to produce and diversify this pharmaceutically relevant molecule.

1.6 Tetronate natural products

1.6.1 Tetronate structure and activity

While thiotetronate compounds are a small and select group of compounds biosynthesised by very similar machinery, tetronate compounds are a much larger class of natural products and show greater variety in compound size, activity and biological origins.[111] Many tetronates exhibit antibacterial properties but also display great potential for treating specific cancers although none are currently marketed.[112] Many tetronate natural products are produced by both bacteria and fungi by PKSs whereas all the currently known thiotetronate family are solely bacterial products. Many tetronate products incorporate the ring in extremely complicated motifs such as that of abyssomicin C (**140, Fig. 1.17**) which contains a spirocyclic tetronate ring.
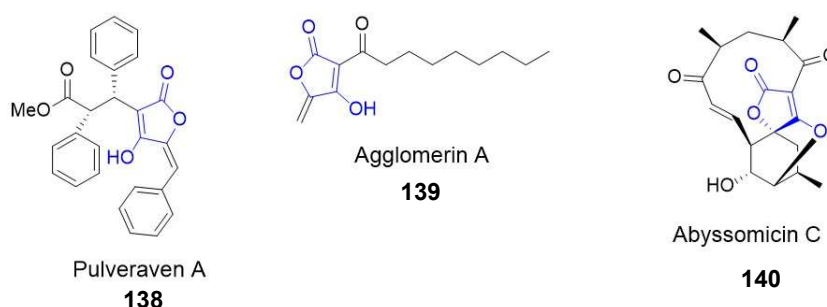
Figure 1.17 – **Pulveraven A (138), agglomerin A (139) and abyssomicin C (140) are diverse natural products all containing a tetronate ring (blue).**

Tetronates employ different biosynthetic machinery for the incorporation of other functionality surrounding the ring. For instance, the bacterial product agglomerin A (**139**) appears to contain a fatty acid chain, whereas the fungal metabolite pulveraven A (**138**) and is postulated to contain phenylalanine or tyrosine derived moieties. The assembly of the tetronate ring has been well studied and results from a thioester ACP-tethered intermediate condensing with a glyceryl-ACP (**Scheme 1.27**).[112]
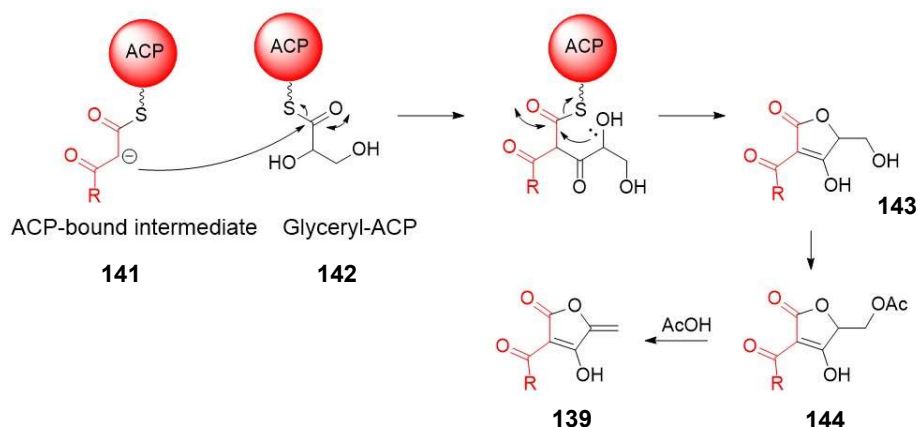
Scheme 1.27 – **Condensation of a keto-thioester containing an ACP-bound intermediate (141) with a glycerate-ACP (142) and ring cyclisation releases a tetronate ring (143). This can be dehydrated to the alkene (144) as is the case in agglomerin A (139) biosynthesis by acetylation (144).**

1.6.2 Tetronasin discovery and bioactivity.

Tetronasin is a polyketide natural product produced by the actinobacterium *Streptomyces longisporoflavus* and has both tetronate and polyether motifs in its structure (**145, Fig. 1.18**). Like other polyether compounds, it has antibiotic activity as an ionophore. It was first reported in 1988 having potent activity against ruminal bacteria, protozoa and anaerobic fungi.[113] It was used commercially as a cattle growth promoter by improving feedstock efficiency, in 2006 it was banned in many European countries.[114] A related compound called tetronomycin (**146**) has minor structural differences to tetronasin but intriguingly has the inverted absolute stereochemistry in all ten chiral centres. Tetronomycin is produced by *Streptomyces* sp. and both compounds exhibit similar activity as ionophores.[115]
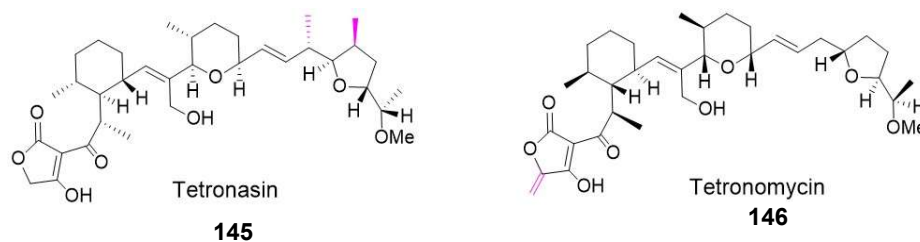
Figure 1.18 – **Structure of tetronasin (145) and related ionophore tetronomycin (146), aside from a few structural differences (pink), both compounds are enantiomers with ten inverted stereocentres.**

The first total synthesis of tetronasin was completed in 1993 by Hori *et al.* who had previously completed the total synthesis of tetronomycin.[116,117] Others have also reported improved synthesis for several of the fragments of tetronasin[118] including the pyran and cyclohexane ring closure in a single step.[119] Despite the success from a synthetic perspective, the biosynthesis of these compounds remains unclear and there has been great interest in understanding how key transformations occur in terms of chemoenzymatically producing bioactive analogues.

1.6.3   Tetronasin and tetronomycin biosynthesis

Investigating the biosynthesis of tetronasin started in the 1980s with several publications reporting incorporation of ^{13}C, ^{18}O and deuterium labelled acetate and propionate units.[120,121] Bulsing *et al.* in 1994 postulated that the tetronasin carbon skeleton was consistent with a polyketide product made from seven acetate units and six propionate units, this was confirmed using ^{13}C labelled acetate and propionate. Feeding of these two labelled substrates to cultures of the producer organism *S. longisporoflavus* resulted in labelled tetronasin (**150-152, Scheme 1.28**).[122–125] as detected by NMR spectroscopy.[126] A series of publications involving isotopically labelled C_4 to C_9 SNAc starter units confirmed the stereocentres for several chiral centres in the starting portion of the tetonasin polyketide backbone.

54
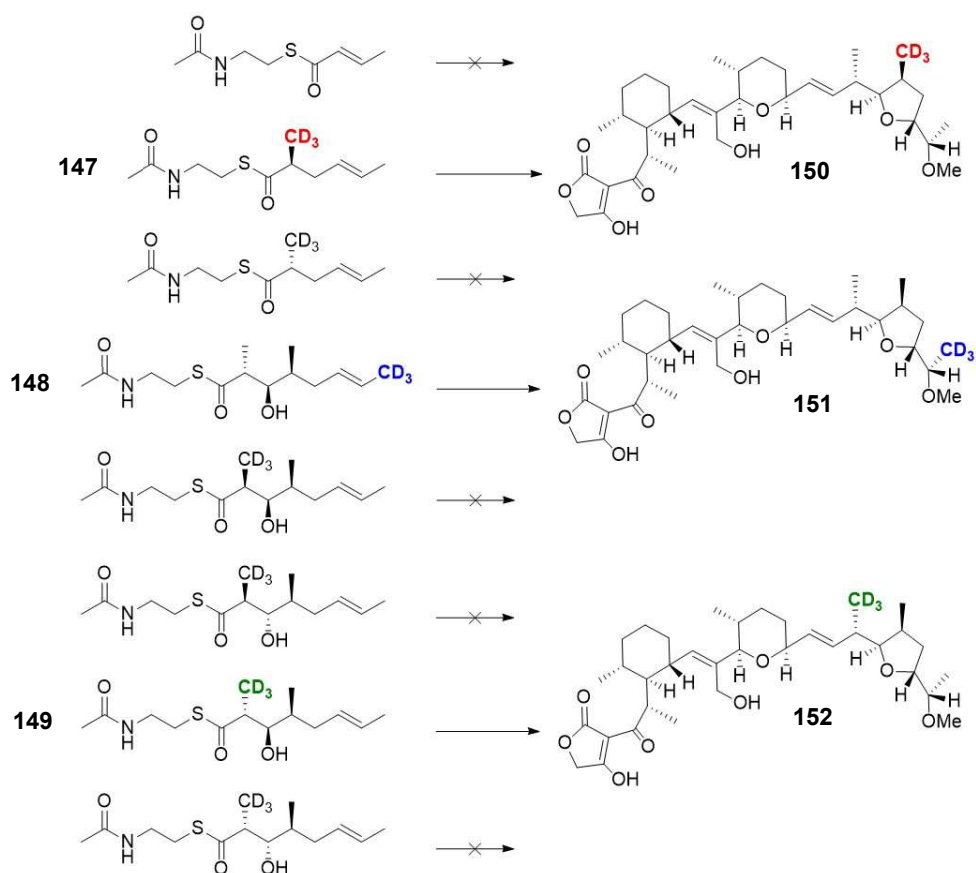
Scheme 1.28 – Deuterium labelled SNAc substrates of varying stereochemistry were fed to *S. longisporoflavus* and only the native isomers (147-149) were processed and observed in the final product (150-152) by NMR spectroscopy. This confirmed the stereochemistry of several chiral centres within tetronasin.

This was followed by analogous experiments with fluorinated SNAc substrates where only correct stereoisomers were processed and the fluorinated ionophores observed by NMR spectroscopy.[127–129] Radioactive glycol was also incorporated into tetronasin and suggested to be a precursor to the tetronate ring (**Scheme 1.29**).[51]
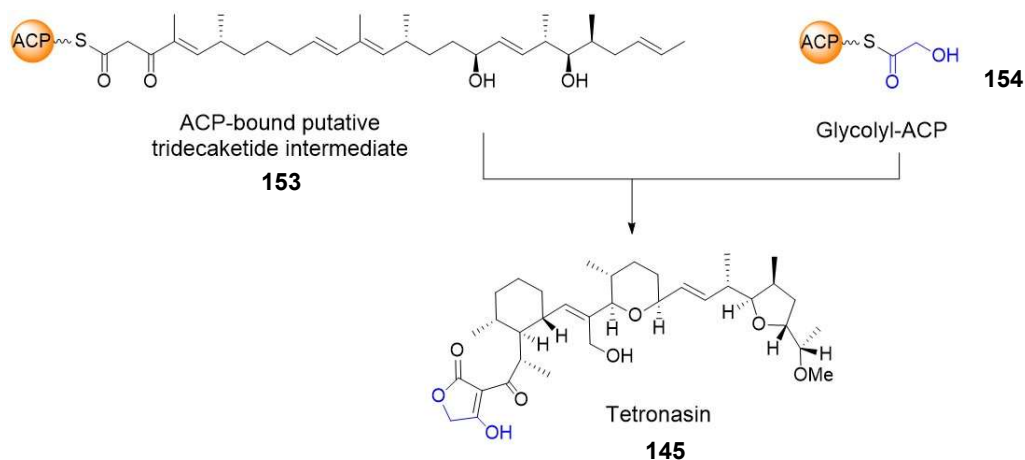
Scheme 1.29 – Tetronasin (145) appears to be made from a putative tridecaketide (153) with incorporation of a glycolyl-ACP (154) which provides a fragment (blue) in the tetronate ring.

More recently, a bioinformatic study of the tetronomycin producer genome revealed a BGC containing a six genes (*tmnAI*, *tmnAII*, *tmnAIII*, *tmnAIV*, *tmnAV* and *tmnAVI*) corresponding to a putative type 1 modular PKS responsible for constructing the carbon skeleton; gene deletion experiments confirmed this was the case.[130] The gene cluster also contains a number of other essential non-PKS genes which share high homology to those found in the related tetronate ionophore chlorothricin (**Fig. 1.21, left**). These include epoxidases and epoxide hydrolases which likely introduce the THF ring, however it is unclear whether that occurs post-polyketide elongation. The alkene side chain tetronate moiety in tetronomycin was demonstrated to originate from glycerate which is then dehydrated unlike tetronasin which appears to use glycolate instead (**Fig. 1.21, right**).[131]
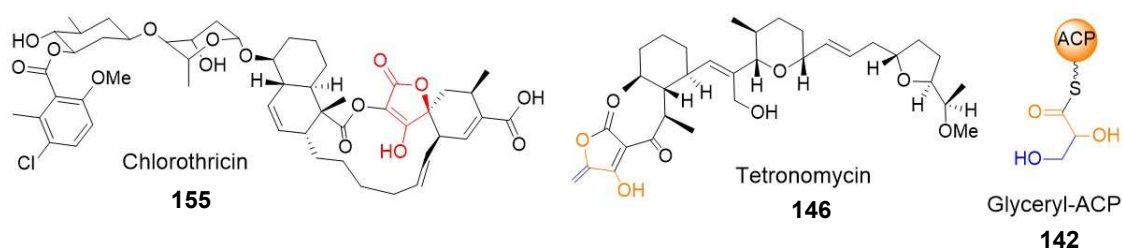
Figure 1.19 – (Left) Chlorothricin (155) is a tetronate natural product which has genes within its BGC highly homologous to tetronomycin (146). (Right) The tetronate moiety is composed partly of a dehydrated glycerate from a glyceryl-ACP (142).

In the case of tetronomycin and tetronasin, the backbone also undergoes hydroxylation and *O*-methylation at an unknown stage and genes responsible for this were previously unassigned.[130] The exact mechanisms of pyran and cyclohexane formation have not been deduced, although a metal ion mediated cyclisation has been proposed. (**157, Scheme 1.30**).[51,132,133]
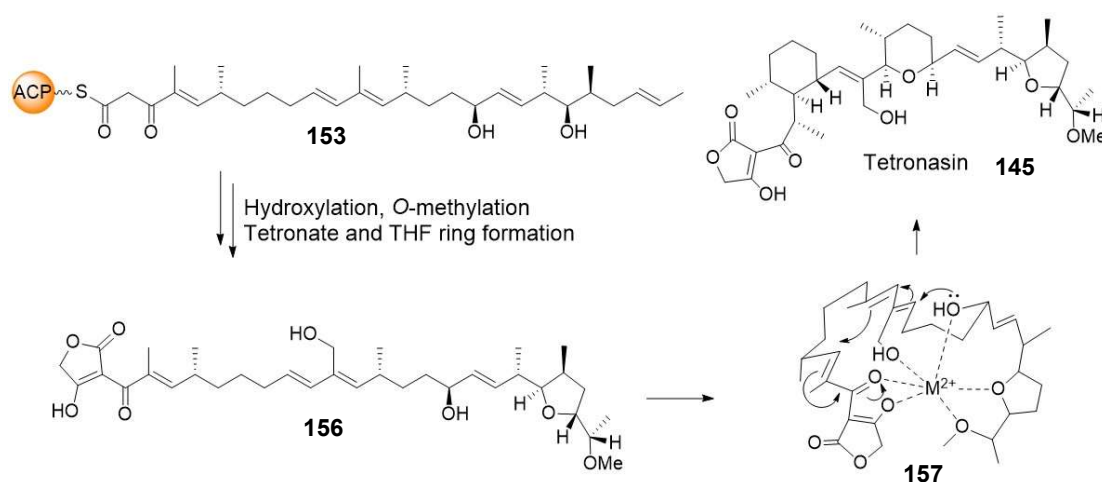
Scheme 1.30 – **Proposed tetronasin biosynthesis: the putative PKS backbone (153) undergoes tetronate and THF ring formation along with hydroxylation (156). A metal catalysed double ring formation (157) leads to the final product (145).**

The domain architecture of the tetronomycin gene cluster consists of twelve modules utilising an acetate starter unit and both malonyl and methylmalonyl extender units at specific extensions. An inactive KR domain is also present in the final module. Tetronasin derives likely from a similar PKS domain architecture to tetronomycin, differing only by the use of methylmalonate rather than malonate for two elongations (**Fig. 1.20**).
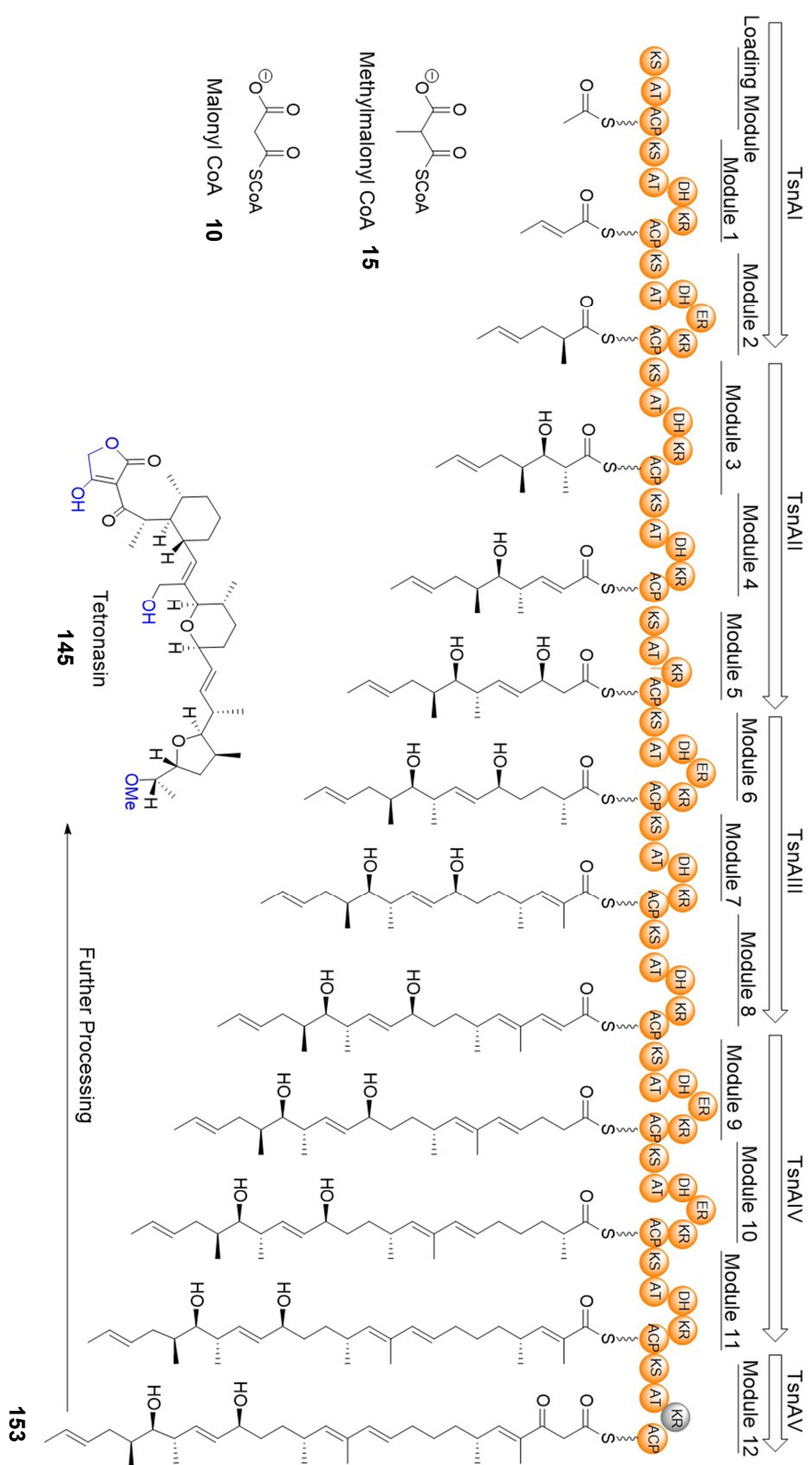
Figure 1.20 – **Putative twelve module type I PKS responsible for producing the tetronasin tridecaketide carbon backbone (153) which is then released and converted to tetronasin (145). Both malonyl (10) and methylmalonyl (15) extender units are utilised and each module is comprised of a different combination of KR, DH and ER domains. Module twelve also has an inactive KR domain.**

58

As discussed, despite extensive investigation, there remains many points of the biosynthesis which remain unclear. Current gene cluster analysis shows a putative *O*-methyl transferase, epoxide hydrolase and other various oxidative genes such as an epoxidase and a monooxygenase but at what point these may act on the polyketide chain is unknown and requires further investigation (**Fig. 1.21**).
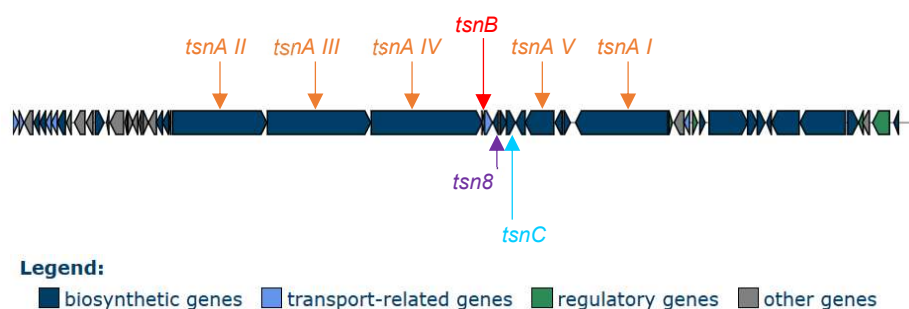
Figure 1.21 – **Biosynthetic gene cluster for tetronasin, *tsnA* (orange) are the modular PKS, *tsn8* (purple) is a methyl transferase, *tsnB* (red) is an epoxide hydrolase and *tsnC* (cyan) is an epoxidase. The cluster also contains other putative biosynthetic genes such as monooxygenases and oxidoreductases which have not been functionally assigned.**

2 Chemical probing of thiotetronate biosynthesis

2.1 Chemical capture of thiolactomycin biosynthetic species *in vivo*

Polyketide biosynthetic intermediates leading to thiolactomycin were proposed at the time of the TLM gene cluster identification (**Chapter 1.5.1**) as represented in **Scheme 2.1**.[93] In order to investigate effectively the nature of the biosynthetic intermediates, methyl ester probes prepared by myself were supplemented to solid cultures of the native **108** producer *Lentzea* sp. as well as with mutant strains prepared by Dr Marie Yurkovich (Leadlay group) at the University of Cambridge.
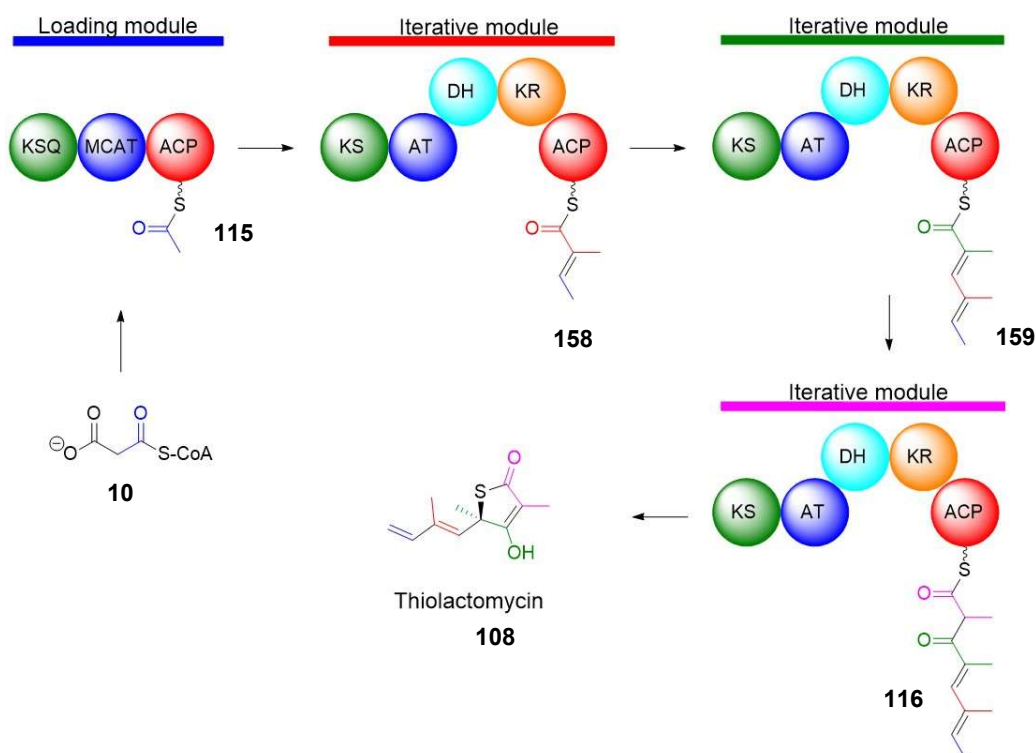
Scheme 2.1 **– Proposed biosynthetic pathway of thiolactomycin tetraketide backbone (116): malonate CoA (10) uptake and decarboxylation provides an acetyl starter unit on the loading module (115), then three rounds of carbon chain extension via decarboxylative Claisen condensation with methylmalonyl building blocks lead to a putative tetraketide ultimately leading to final product (108).**

The structure of probes used are shown below and results of the studies undertaken with them are summarised in the next sections.

2.1.1 *In vivo* capture of polyketide intermediates from the TLM wild type producer *Lentzea* sp. and in mutant/heterologous expression strains

The decanamido (**104**) and fluoromalonyl (**160**) probes had previously shown to be very effective in offloading intermediates from iterative type I PKS systems such as 6MSAS, and were therefore tested first. A new methylmalonyl probe (**161**) prepared by myself for the first time was also tested in order to more closely resemble the structure of the natural methylmalonyl extender unit used in thiolactomycin biosynthesis.
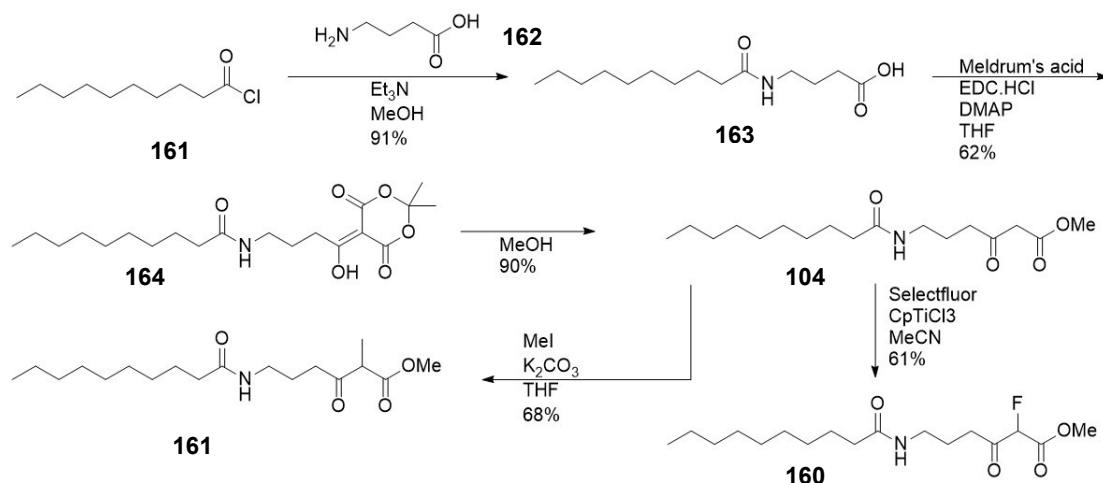
Scheme 2.2 – **Synthetic route to methylmalonyl probe (161) and fluoromalonyl probe (160).**

The methylmalonyl probe (**161**) and fluoromalonyl probe (**160**) were synthesised in a single step from the previously reported decanamido probe (**104**).[85] The synthesis of the decanamido probe (**104**) began with amide formation between decanoyl chloride (**161**) and γ-amino butyric acid (GABA) (**162**) to form the amide carboxylic acid (**163**) in high yield without purification. 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide

61

hydrochloride (EDC.HCl) was then used to couple Meldrum's acid to the carboxylate leading to a conjugated adduct (**164**) in moderate yield. When refluxed in MeOH, this undergoes thermolysis to release acetone and CO_2, affording the decanamido probe (**104**). Fluorination of **104** was carried out using Selectfluor to give the fluoromalonyl probe (**160**) in reasonable yield and the preparation of the methylmalonyl probe (**161**) was using methyl iodide and potassium carbonate with a moderately high yield. All three probes (**104, 160, 161, Fig. 2.1, top**) were used in 2 mM concentration for the wildtype (WT) strain growing on solid media, showing mild to no toxicity (**Fig. 2.1, bottom**). Higher concentrations had previously been tested by fellow group member Dr Judith Havemann and had detrimental effects towards bacteria growth, whereas lower concentrations (1 mM) had proved ineffective for intermediate detection.
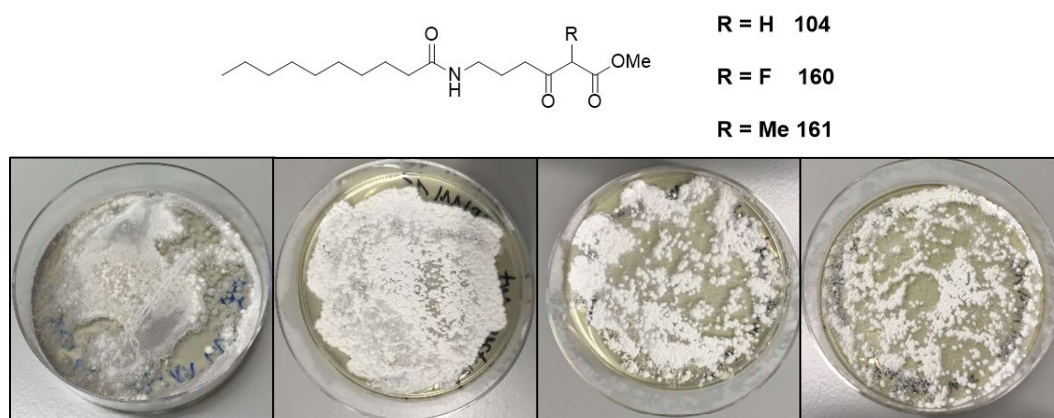
Figure 2.1 – **(Top): Structure of chemical probes for *in vivo* experimentation. These include no malonyl substitution as in 104, malonyl methylation as in 161, and malonyl fluorination as in 160. (Bottom): plates of WT *Lentzea* sp. supplemented with (from left to right) no probe, 2mM decanamido probe (104), 2mM methylmalonate probe (161), and 2mM fluoromalonate probe (160).**

The fermentation of solid cultures was carried out for 5 days with probe containing agar, after which time, plates were sliced into small pieces and extracted by soaking in EtOAc. Concentrated extracts were then re-dissolved in MeOH for analysis by LC-HRMS on a ThermoFisher LTQ Orbitrap instrument. Firstly, the peak area size for the intact probe mass peak was compared with that of the hydrolysed and decarboxylated equivalent to estimate the ratio and therefore approximate efficiency of probe hydrolysis. The methylmalonyl probe (**161**) showed the poorest level of *in vivo* hydrolysis and

decarboxylation at around 15%, the decanamido probe (**104**) was slightly more efficient (30-35%). The fluoromalonate probe (**160**) displayed considerably higher extent of hydrolysis, with around 90% detected. Putative thiolactomycin diketide and triketide species with varying degrees of processing were captured and observed with all three probes (**Table 2.1**).[134]
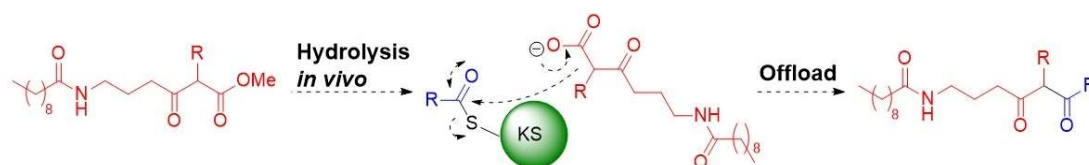
Intermediate	104 (Decanamido)	161 (Methylmalonyl)	160 (Fluoromalonyl)
	Y	Y	Y
	N	N	Y
	N	Y	Y
	Y	Y	Y
	N	Y/T.A.	Y/T.A.
	N	N	N
	N	Y/T.A.	Y/T.A.

Table 2.1 – **Summary of putative thiolactomycin intermediates captured with 104, 160 and 162: Y indicates detected, N indicates not detected and T/A detected in trace amounts. The fluoromalonyl probe (160) was able to capture a wide variety of intermediates in comparison to the other two probes.**

The fluoromalonate and methylmalonate probes also captured a tetraketide (**166**) in trace amounts and all intermediates were characterised by MS^2 (**167-168, Fig. 2.2**). Captured intermediates were consistent with those which would be KS-bound and about to undergo decarboxylative Claisen condensation from an ACP-bound methylmalonate. Control extracts which did not contain any probe did not show any of the same parent ions as confirmation of putative species.
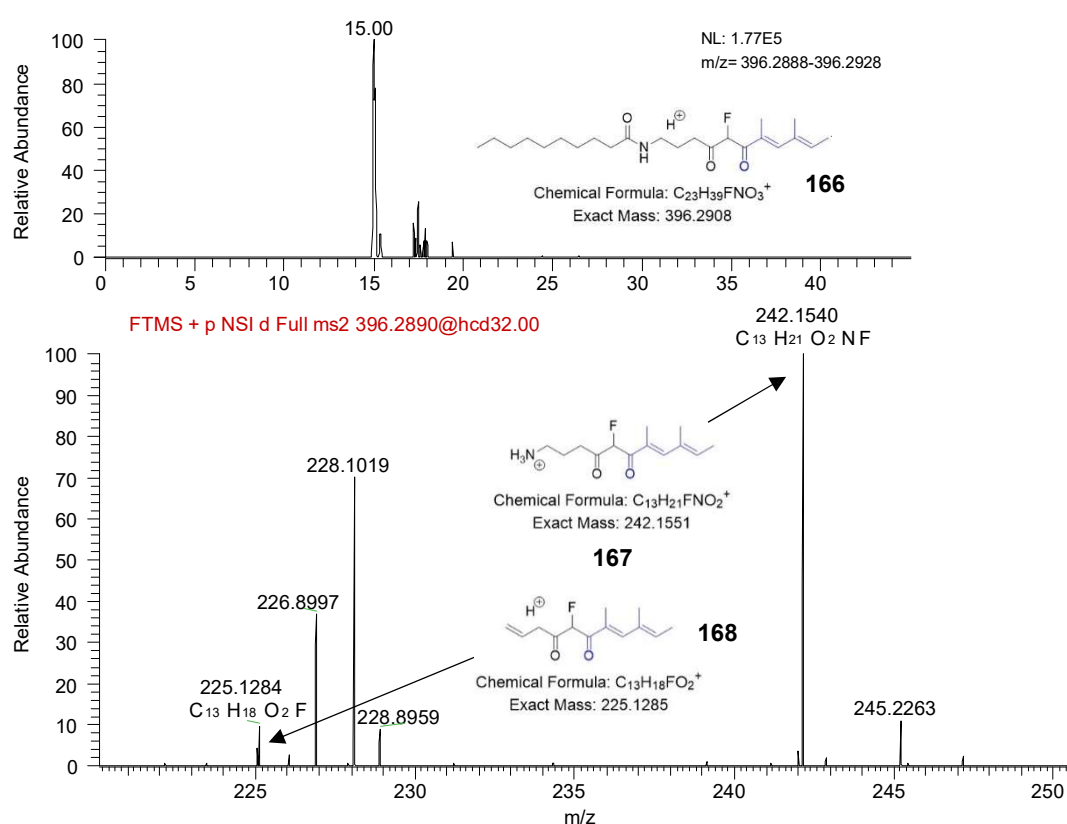
Figure 2.2 – **LC-HRMS extracted ion chromatogram for captured tetraketide (166) with fluoromalonyl probe (160) (top) and two diagnostic MS^2 fragments (167 and 168) of the parent ion (bottom).**

Some captured species (**169**) had further processing consistent with capture of an intermediate and subsequent reduction or dehydration by the KR and DH domains (**Fig. 2.3**).
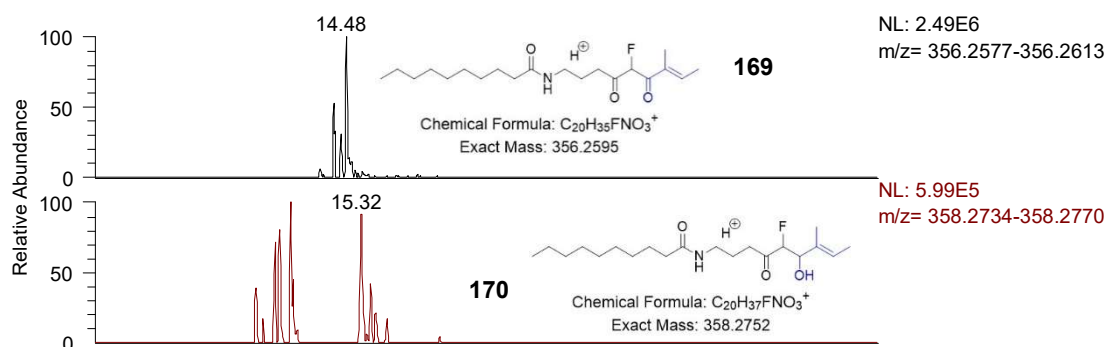
Figure 2.3 – **LC-HRMS extracted ion chromatogram for captured triketide (169) showing post capture reduction by the KR domain to give the detectable reduced analogue (170).**

This points towards similar structural properties of the putative captured species to ACP-bound intermediates. Many of the same intermediates from the wildtype extracts were also detected using other mutant strains of *Lentzea* sp. such as the *tlmD1* and *Cy* knockout strains at the same retention times (see ChemComm publication in **Appendix 7.5**, or summarised data in table here) This was fairly unsurprising as these domains are proposed to be involved in post-PKS biosynthesis and should not influence the polyketide backbone formation.

An Erasmus student working in our group at the same time of my investigations, Sophia Harringer, also showed that similar intermediates can be captured in the assembly of Tü3010 (**110**) in *S. thiolactonus,* where an extra methyl group is incorporated when ethylmalonate in used a single elongation (**171, Fig. 2.4**).[134]
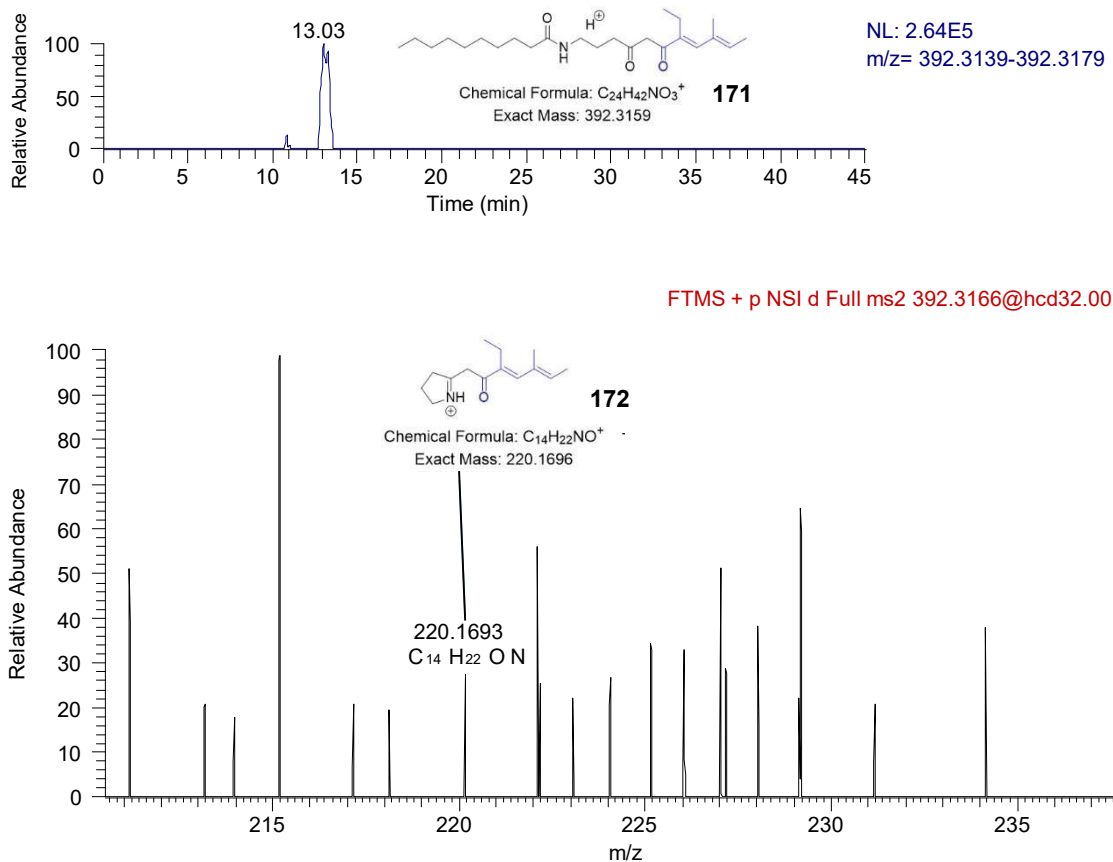
Figure 2.4 – **LC-HRMS extracted ion chromatogram for captured tetraketide (171) from *S. thiolactonu*s with the decanamido (104) probe (top); a diagnostic MS² fragment (172) for this was also detected (bottom).**

As well as whole gene deletion strains such as the Δ*tlmA* and Δ*tlmD1*, our collaborators in Cambridge also generated a number of *tlm* clusters (harboured by native or heterologous hosts) with single point mutations in key catalytic residues of the TE, Cy, and PCP domains within *tlmB*. This was pursued in the effort to shed light on the mechanisms of S insertion and cyclisation.[135] *S. coelicolor* sp. M1154 or *S. lividans* TK24 have been used as heterologous hosts for the *tlm* cluster (**Table 2.2**).

Strain	Domain mutated
Lentzea sp. ATCC31319	Cy (S1832A)
S. coelicolor M1154:Tlm	Cy (D1824A, D1829A)
S. coelicolor M1154:Tlm	Cy (D2083A)
S. coelicolor M1154:Tlm	PCP (S2678A)
S. lividans TK24:Tlm	PCP (S2678A)
S. lividans TK24:Tlm	Cy (D1824A, D1829A)
S. lividans TK24:Tlm	TE (S2800A)

Table 2.2 – **Mutation strains/clusters used in conjunction with the fluoromalonyl probe (160)**

No thiolactomycin was produced in each of these strains as expected, however also no polyketide intermediates were intercepted by the fluoromalonyl probe either. For heterologous hosts, natural product formation often occurs at a much lower titre than that of the wildtype host. No thiolactomycin was produced in these strains and it is possible that enzyme-bound intermediates were present at too lower level to be efficiently intercepted. For the *Lentzea* ΔCy mutant, the apparent lack of intercepted intermediates has been unexpected and is worthy of further future investigation.

2.1.2   *In vivo* capture of putative sulphur containing intermediates from the TLM wildtype producer *Lentzea* sp. and in mutant/heterologous expression strains

Based on the proposed mechanism of thiolactomycin formation from the tetraketide through an epoxide and thiirane formation (**Scheme 2.3**),[93,134] intermediates to corroborate these steps were searched for. This included tetraketides containing intact epoxide and thiirane rings and also ring opening products of these with either water or methanol which the extracts were subjected to. The Δ*tlmD1* strain became an important control strain from which such intermediates should not be detected.
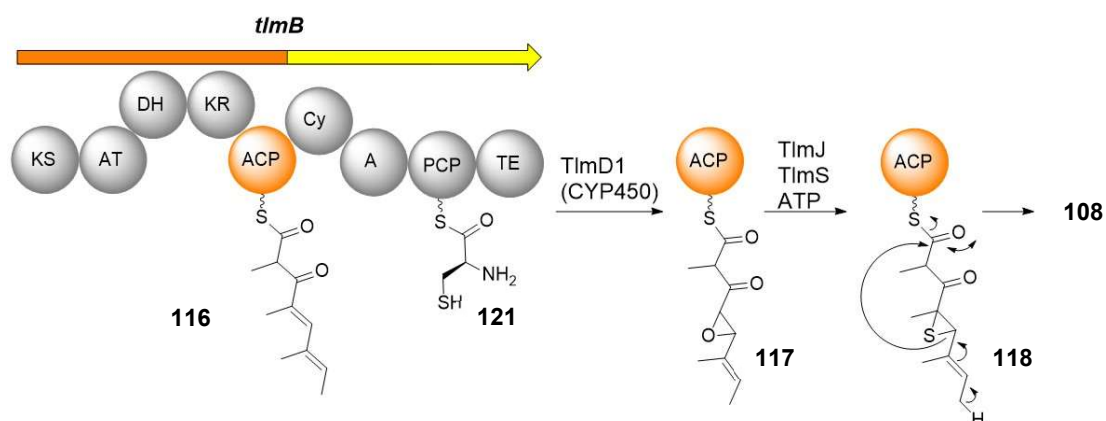
Scheme 2.3 **– Proposed late stage thiolactomycin biosynthesis: TlmD1 catalyses epoxide formation (117) on a tetraketide intermediate (116) followed by thiirane formation (117) mediated by TlmJ and TlmS using sulphur from PCP-bound cysteine (121), ring formation and thiolactomycin (108) release.**

Analysis of wild type *Lentzea* sp. extracts grown in the presence of **160** revealed a putative fluorinated and ring-opened thiirane intermediate with plausible MS^2 fragments (**173, Fig. 2.5**). An analogous putative species was also found in *S. thiolactonus* extracts, with an extra methyl group derived from the use of ethylmalonate for a single elongation step. Control extracts lacking in **160** did not contain putative intermediates of this kind, as well as in the Δ*tlmD1* mutant extracts.

The Cy domain was proposed to mediate the final heterocycle formation step whereby the final polyketide intermediate is released forming thiolactomycin. However, analysis of extracts derived from this mutant did not reveal any putative sulphur containing species.
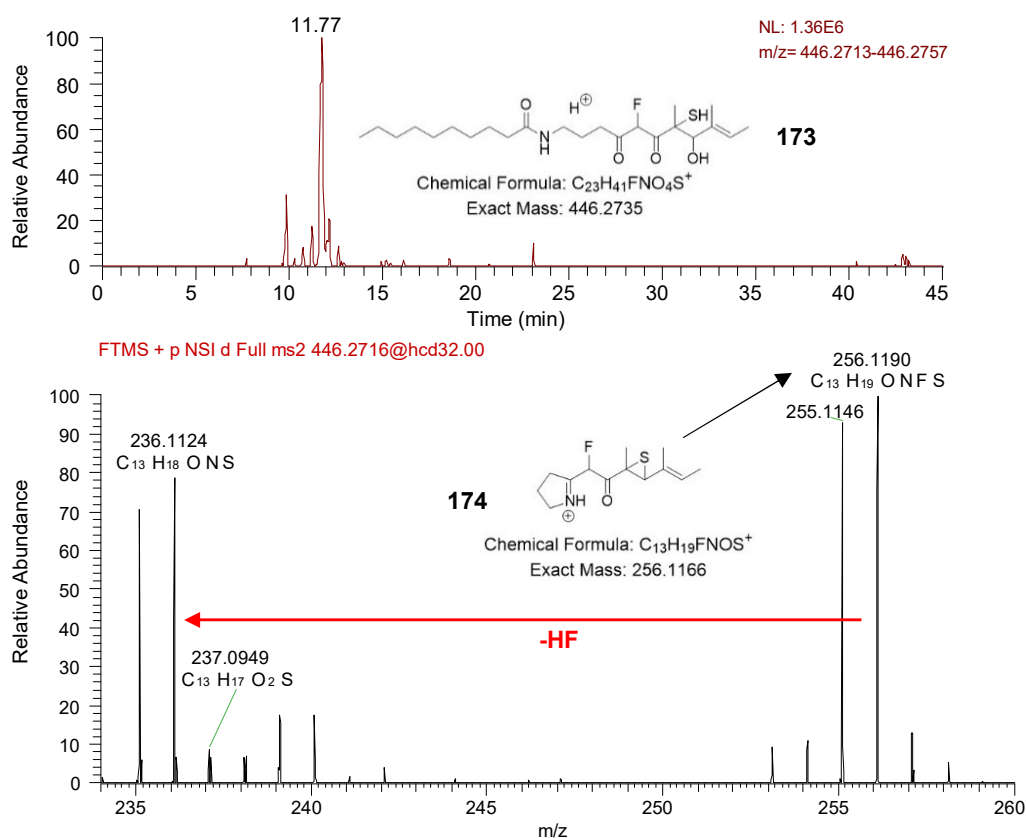
Figure 2.5 – **LC-HRMS ion extracted chromatogram of thiirane derived species (173) and a derived MS² fragment (174).**

Two putative pentaketides (**175-176**) containing the possible remnants of a thiirane ring were detected, these correspond to ring opening by MeOH or H_2O (**Fig. 2.6, A&B**). Both were detected with clear MS² fragment species consistent with both structures.

A)

NL: 7.74E4
m/z= 516.3109-516.3161

Chemical Formula: $C_{27}H_{47}FNO_5S^+$
Exact Mass: 516.3153

175

FTMS + p NSI d Full ms2 516.3135@hcd32.00 [110.0000-527.0000]

344.1693
$C_{17}H_{27}O_3NFS$

Chemical Formula: $C_{17}H_{27}FNO_3S^+$
Exact Mass: 344.1690

176
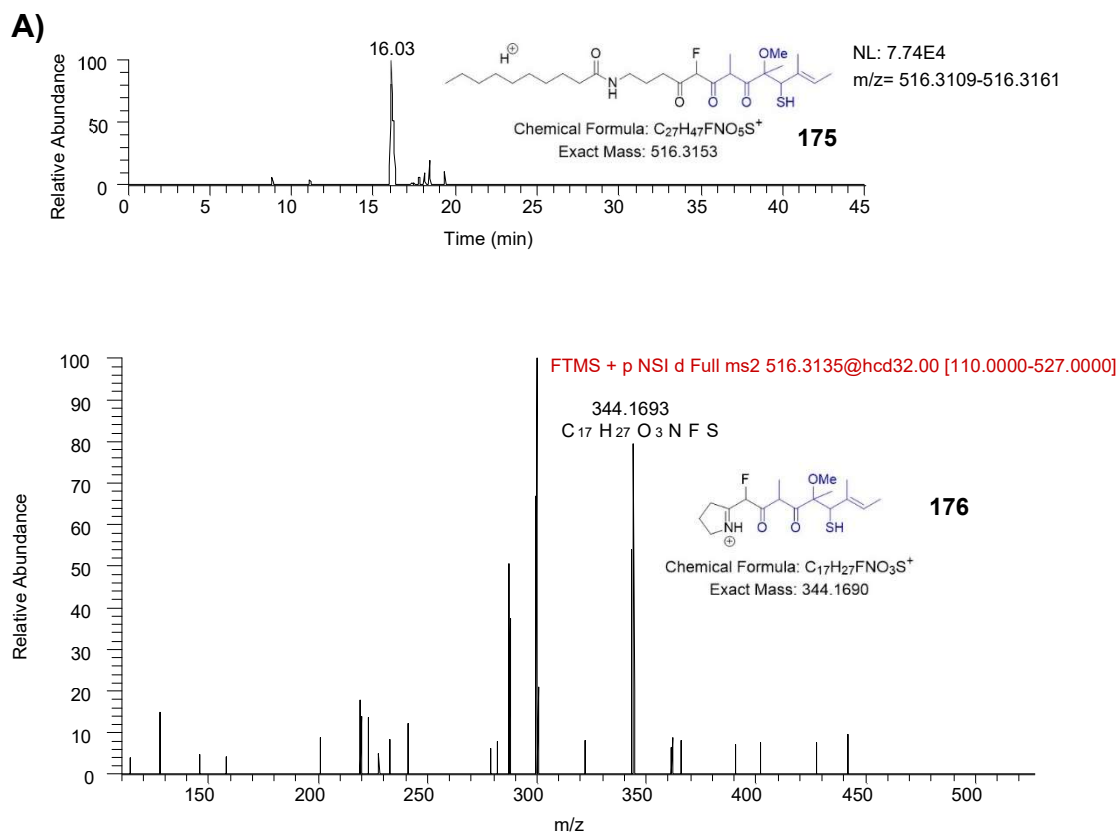
Figure 2.6 – **(A&B) Putative thiirane ring opened species from MeOH (175) and H$_2$O (177) with their respective MS^2 fragments (176) and (178).**

These are structurally similar to the putative sulphur containing tetraketide intermediates previously shown (**Fig. 2.5**). The detection of both tetraketide and pentaketide species bearing similar functionalities suggests that S insertion could happen at enzyme-bound triketide or tetraketide stage.

B)

NL: 8.76E4
m/z= 502.2972-502.3022

Chemical Formula: $C_{26}H_{45}FNO_5S^+$
Exact Mass: 502.2997

177
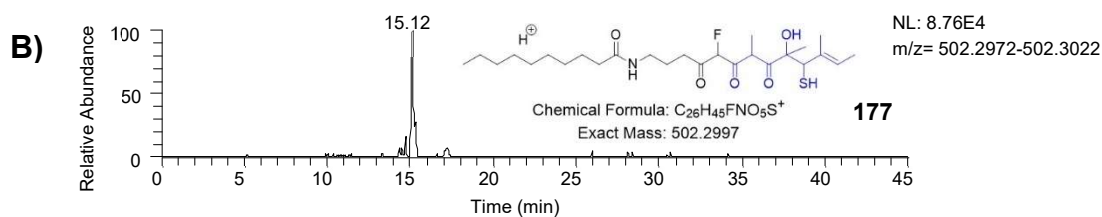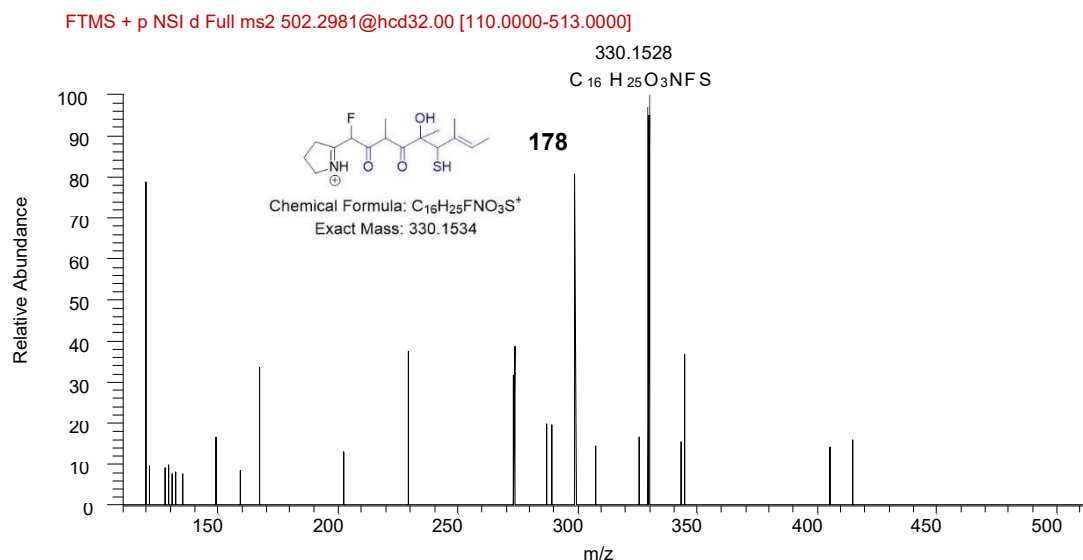
The lack of detection of putative S species from *Lentzea* sp. point mutation strains (*S. coelicolor* and *S. lividans*) is consistent with the lack of early stage intercepted intermediates, this is not surprising as the lack of early intermediates suggests no later biosynthetic steps are occurring.

It was noted during early experiments with the *Lentzea* sp. WT strain and ΔCy, Δ*tlmD1* and Δ*tlmA* mutants, that the phenotype of the Δ*tlmD1* knockout strain is characterised by a very strong green colour readily extractable into organic solvent (**Fig. 2.7**). An attempt to isolate a compound related to the green colour by HPLC was made, but with no success. The possible origins of this colour are discussed in chapter 5.

Figure 2.7 – **Comparison of *Lentzea* sp. WT (left) and the CYP450 inactivated Δ*tlmD1* mutant (right) phenotype- the agar became green coloured during Δ*tlmD1* growth.**

2.2 Chemical capture of TlmB-derived biosynthetic intermediates *in vitro*

2.2.1 *In vitro* capture of polyketide intermediates from TlmB

In parallel to polyketide intermediate capture *in vivo*, an investigation of the workings of recombinant TlmB in the presence of the chemical probes was carried out *in vitro*. Marie Yurkovich (Cambridge) was able to amplify *tlmB* from *Lentzea* sp. genomic DNA and insert it into a pIB139 vector for protein expression in *E. coli* (BL21) and purification for use.[136] When incubated with the necessary polyketide building blocks and cofactors acetyl CoA (**36**), methylmalonyl CoA (**15**) and NADPH (**Scheme 2.4**), TlmB was shown by our chemical probes to produce di (**181**), tri (**182**) and tetraketide (**183**) species similar to those previously reported from *in vivo* studies.[134]
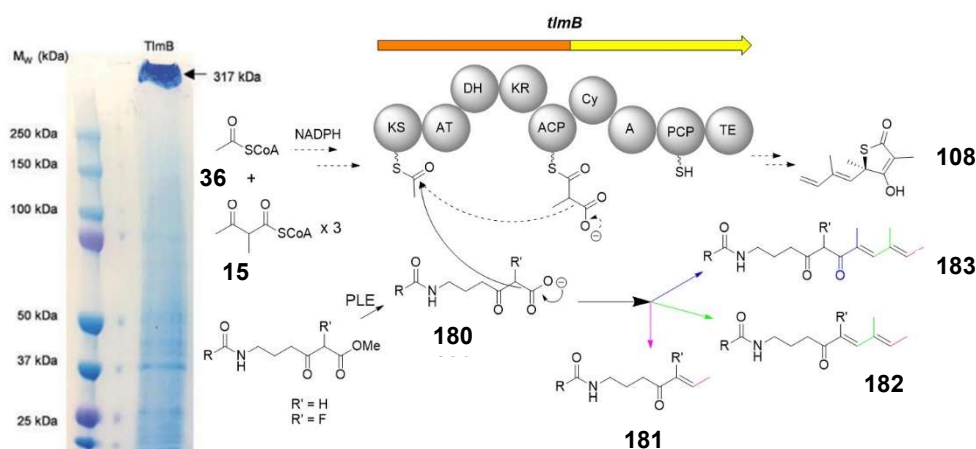
Scheme 2.4 – (Left) TlmB expression and purification. (Right) incubation of TlmB with NADPH, acetyl (36) and methylmalonyl (15) coenzyme-A substrates can produce polyketide intermediates but not thiolactomycin (108). Fluoromalonyl (160) and decanamido probe (104) treated with PLE to give the carboxylate (180) which captured intermediates of di (181), tri (182) and tetraketide (183) length.

Note: The experiments discussed on this page were completed solely by M. Yurkovich.

The fluoromalonyl (**160**) and decanamido (**104**) probes were incubated with pig liver esterase (PLE) to generate the corresponding carboxylate which captured said intermediates as detected by LC-HRMS with the expected MS^2 fragments (**Fig. 2.8**).
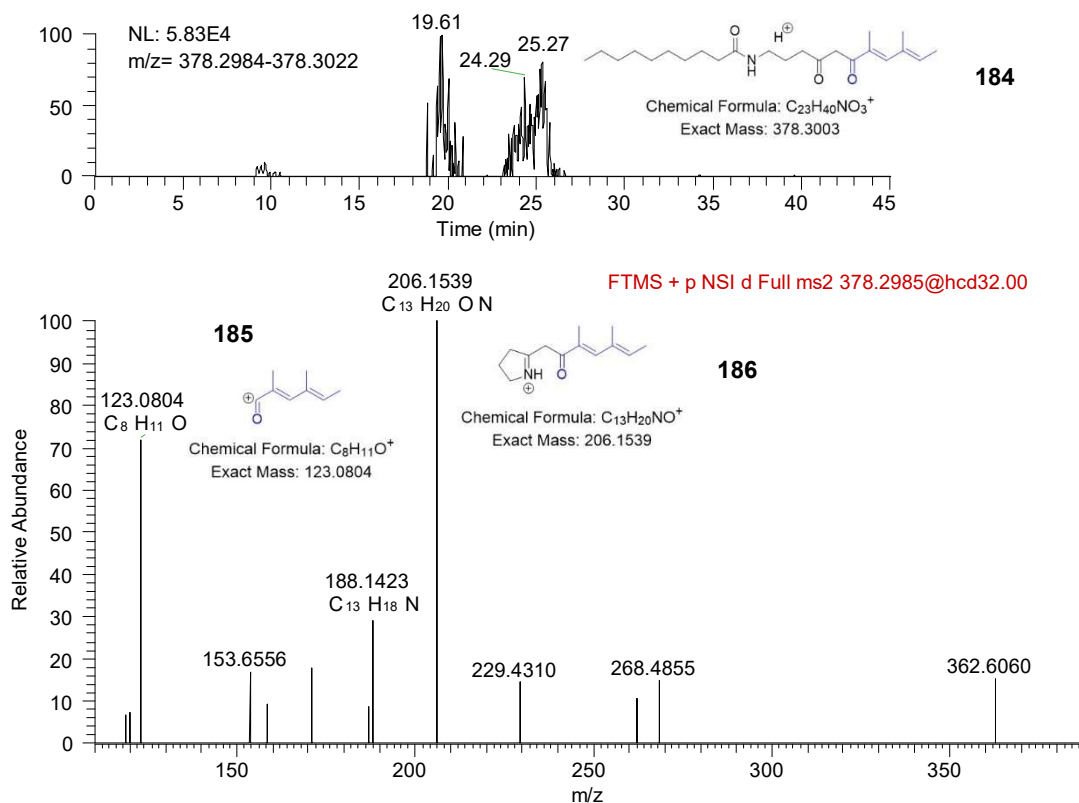
Figure 2.8 – **Top: captured tetraketide (184) using decanamido (104) probe detected; (bottom: two expected MS^2 fragments (185 and 186) also detected.**

The most advanced polyketide species detected was the tetraketide illustrated in **Fig. 2.5** No thiolactomycin or other related shunt products were observed in these experiments as many of the other essential proteins such as TlmD1, TlmS and TlmJ were not present in the first instance. A synthetic SNAc substrate was therefore synthesised (**Scheme 2.5**) in order to further probe the workings of the *in vitro* machinery in the presence of additional domains.

2.2.1.1 Synthesis of SNAc triketide mimic (192)
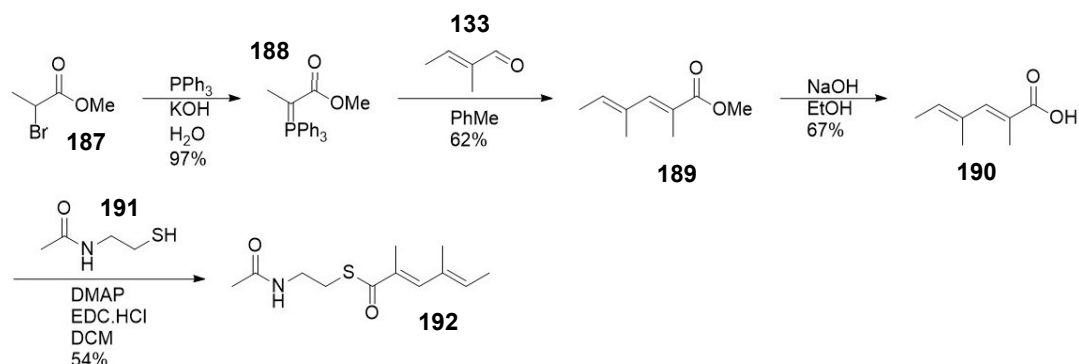
Scheme 2.5 – Synthetic route to triketide SNAc thioester (192)

Alkylation of methyl 2-bromopropionate (**187**) with triphenylphosphine and treatment with base gave the ylide (**188**) in excellent yield. A Wittig reaction with tiglic aldehyde (**133**) gave the diene methyl ester (**189**) in modest yield. Hydrolysis of the methyl ester gave the corresponding carboxylic acid (**190**) in 67% yield, which was then coupled to *N*-acetyl cysteamine (**191**) to afford the triketide mimic (**192**) in 54% yield (**Scheme 2.5**).

The synthetic SNAc triketide mimic was incubated with TlmB in the hope that the mimic could provide an ACP or KS-bound substrate (using boiled enzyme controls in duplicate). In the presence of both the decanamido (**104**) and fluoromalonyl (**160**) probes, a tetraketide (**166**) species similar to those previously characterised *in vivo* was identified. This proved that the triketide is an effective intermediate in thiolactomycin biosynthesis (**Fig. 2.9**. This is a rare example showing the ability of an iterative PKS-NRPS machinery to load and process SNAc substrates. *Note: this assay performed by M. Yurkovich*.
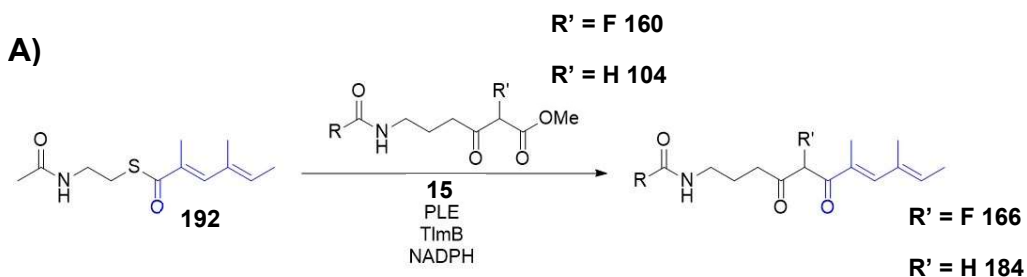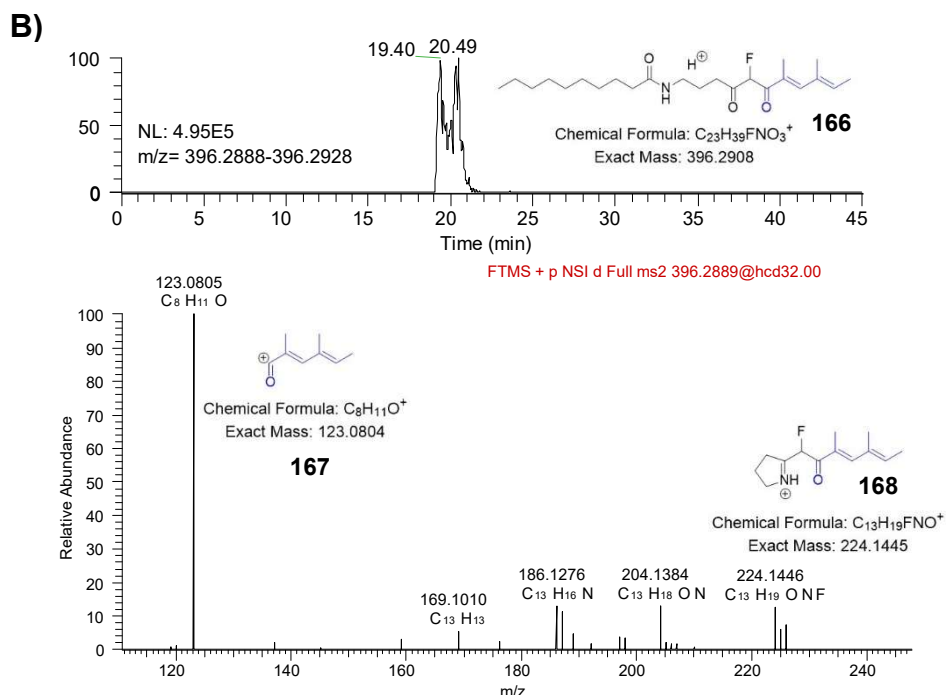
Figure 2.9 – **(A) SNAc triketide (192) can be loaded onto TlmB and then recaptured using the fluoromalonyl (160) and decanamido (104) probes in their carboxylate form after PLE hydrolysis. The tetraketide formed from probe capture (166) offloaded is the same as was detected from the *in vitro* assays containing TlmB and CoA substrates in TLM synthesis, (B) LC-HRMS extracted ion chromatogram for the tetraketide (166) using the fluoromalonyl probe (160) and two diagnostic MS[2] fragments (167 and 168).**

2.2.2   Capture of putative sulphur containing intermediates from the TlmB *in vitro*

Several intermediates consistent with thiirane formation were found in extracts derived from TlmB *in vitro* assays.   These include putative tetraketides which may have undergone epoxide/thiirane formation and ring opening (**193 and 173, Fig. 2.10, A**); elongated pentaketides with the same MS[2] fragmentation (**194-195**) as shown in **Fig. 2.6** were also detected. For these experiments (set up by Marie Yurkovich) TlmB was given all the required cofactors, substrates and supplemented with cysteine for sulphur donation. A previously undetected putative intact thiirane was also detected, albeit in relatively low abundance; this further complements the observation of putative ring-opened tetraketides (**Fig. 2.10, B**) which are presumably derivatives.

75

A)

NL: 1.98E5
m/z= 460.2868-460.2914

Chemical Formula: C₂₄H₄₃FNO₄S⁺
Exact Mass: 460.2891 **193**

NL: 8.05E5
m/z= 446.2713-446.2757

Chemical Formula: C₂₃H₄₁FNO₄S⁺
Exact Mass: 446.2735 **173**
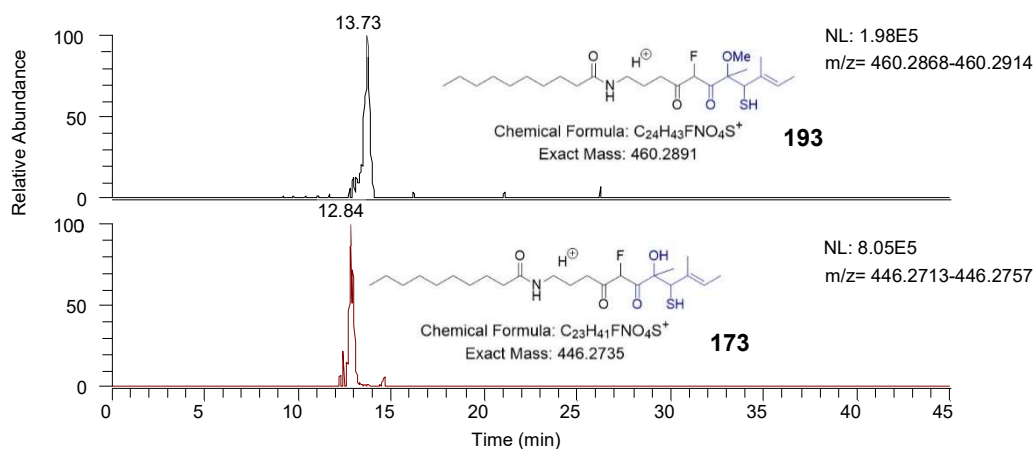
FTMS + p NSI d Full ms2 446.2725@hcd32.00 [110.0000-457.0000]

274.1281
C₁₃H₂₁O₂NFS

Chemical Formula: C₁₃H₂₁FNO₂S⁺
Exact Mass: 274.1272 **194**
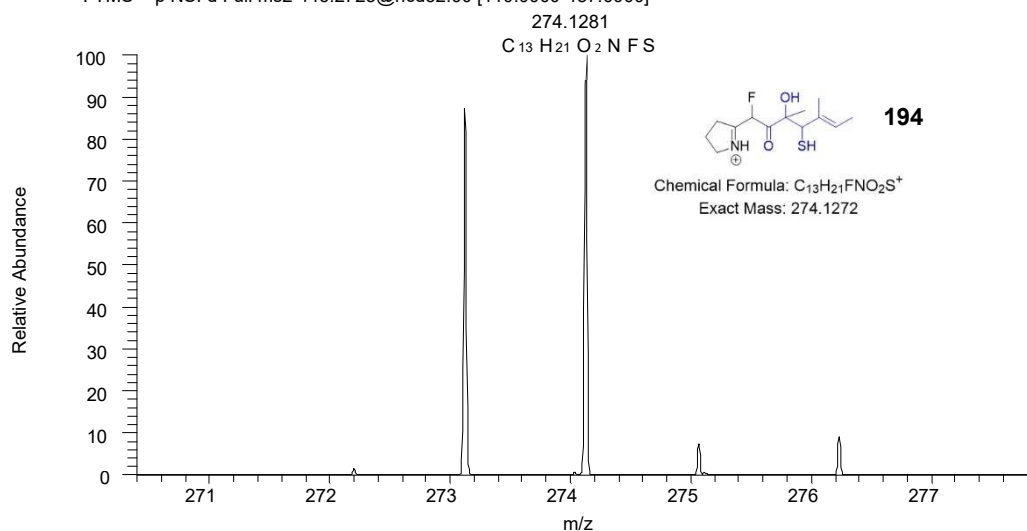
FTMS + p NSI d Full ms2 460.2869@hcd32.00 [110.0000-471.0000]

288.1435
C₁₄H₂₃O₂NFS

Chemical Formula: C₁₄H₂₃FNO₂S⁺
Exact Mass: 288.1428 **195**

76
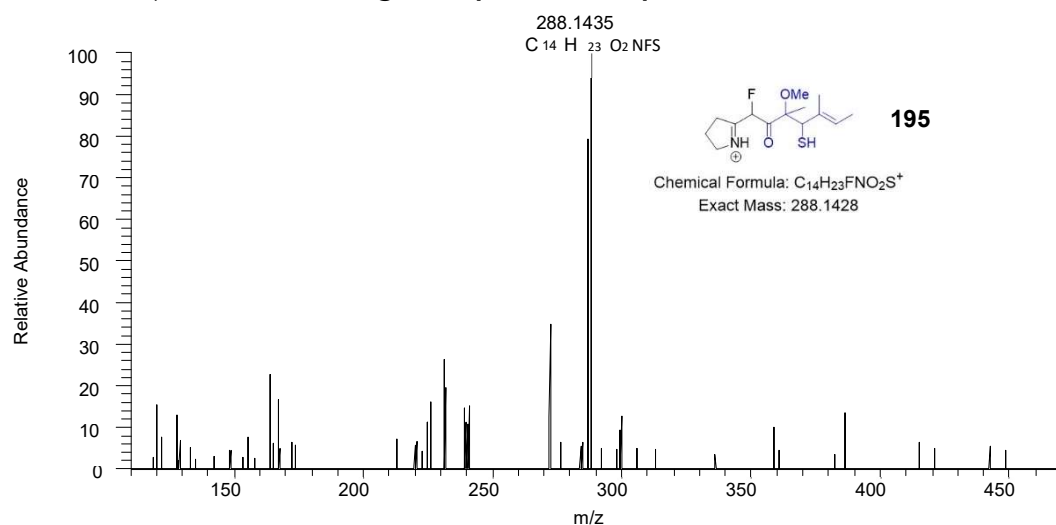
B)
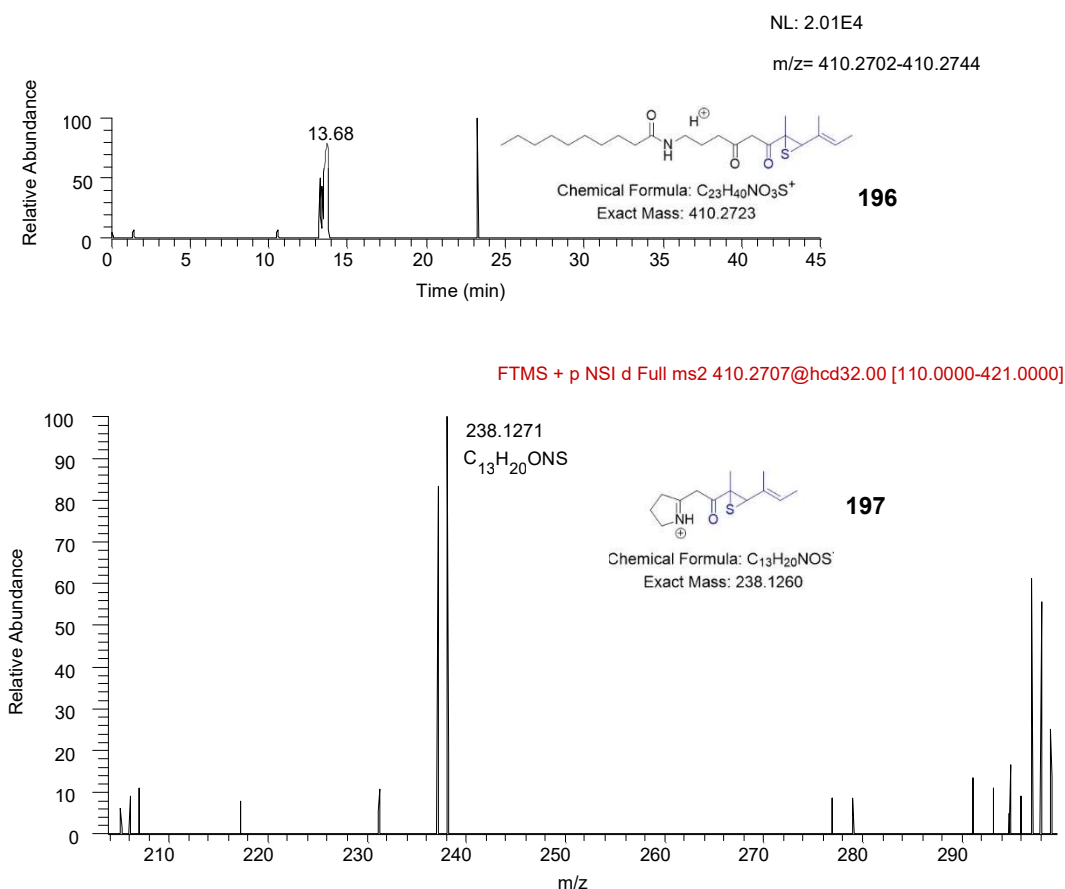
Figure 2.10 – **(A) Putative ring-opened thiirane tetraketides (193 and 173) captured by fluoromalonyl probe (160) detected in in vitro assays along with characteristic fragments (194 and 195). (B) Intact thiirane tetraketide (196) along with MS² fragmentation (197) detected in capture experiment with decanamido probe (104).**

Instead of TlmJ and TlmS (which were cloned but proved to be insoluble proteins), the soluble homologues TueJ and TueS from *S. thiolactonus* were used; these should have similar functionality. What proved particularly puzzling in these *in vitro* experiments is that the capture of putative sulphur-containing tetraketide and pentaketides was achieved regardless of whether TlmD1, TueJ or TueS were present in the TlmB assays. This suggests the two sulphur related genes TlmJ and TlmS are not involved in addition of the sulphur into the polyketide backbone.

It is possible that the presence of sulphur in the previously discussed thiirane related species may not be directly related to thiolactomycin formation itself. In fact, to date no *in vitro* reconstitution of the whole TLM biosynthetic machinery has been achieved, and the true nature of species generated *in vivo* can be complex to elucidate. Nonetheless, on the basis of our data at this stage, we started to consider whether S insertion and thiotetronate ring formation would occur from other proteins independent of TlmB itself.

Given that the *S. pacifica* TLM cluster has no essential sulphur transfer related genes, it is not surprising that TueJ and TueS are not essential. However, TlmD1 was thought to generate an epoxide to aid thiirane formation and we could not experimentally validate this hypothesis with the aid of chain termination probes.

2.3   Investigation of late stage thiotetronate biosynthesis with synthetic substrates

Given the successful priming of the SNAc triketide (**192**) onto TlmB, other putative intermediates of thiolactomycin were synthetically prepared as SNAc thioesters in order to see whether the bacteria could process these into thiotetronate products. As the Δ*tlmA* strain possesses all of the required machinery for making thiolactomycin except for the loading module and therefore cannot produce the compound, it was thought that this strain could process synthetic substrates comprising of at least an acetate unit.

An array of compounds was devised in order to act as mimics of ACP-bound biosynthetic species as earlier envisaged (**Scheme 2.1**) and based also on the probe captured intermediates.

2.3.1 Synthesis of putative intermediate SNAc substrates

The structure and synthesis of the SNAc substrates is reported in the following paragraphs except the triketide (**192**) which was previously discussed (**Scheme 2.5**). All were synthesised as SNAc thioesters to allow priming of the ACP by acting as CoA mimics.

2.3.1.1 Synthesis of acetyl 'monoketide' mimic (199)
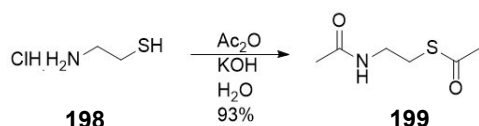
Scheme 2.6 – Synthetic route to acetyl SNAc thioester (199)

Acetyl SNAc (**199**) was synthesised as a monoketide mimic in a single step by acetylation of cysteamine hydrochloride (**198**) with acetic anhydride in 93% yield (**Scheme 2.6**).

2.3.1.2 Synthesis of diketide mimic (201)
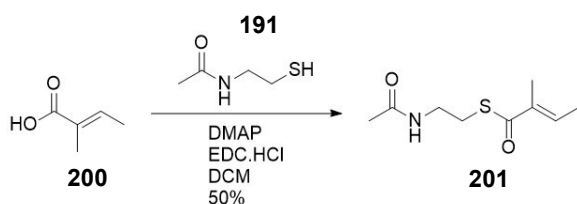
Scheme 2.7 – Synthetic route to diketide SNAc thioester (201)

A single step thioester coupling of tiglic acid (**200**) with *N*-acetyl cysteamine (**191**) (SNAc) gave the product (**201**) in 50% yield (**Scheme 2.7**); this was completed by Aiste Andriulyte (URSS project student, summer 2017).

2.3.1.3 Synthesis of tetraketide mimic (207)

The postulated tetraketide (**207**) contains one chiral centre; an asymmetric synthesis was considered unnecessary as the stereocentre is acidic and if a single enantiomer was made, it would likely racemise. Two synthetic routes were used to access this compound.
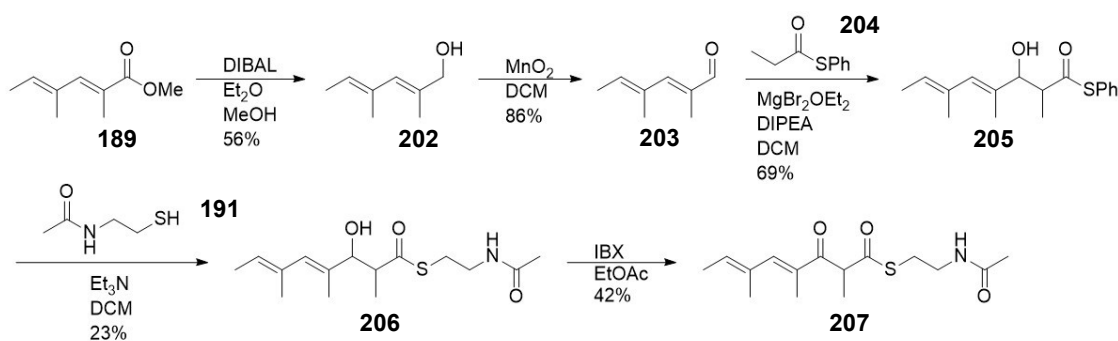
Scheme 2.8 – First synthetic route to tetraketide SNAc thioester (207)

The first synthesis (**Scheme 2.8**) began with the previously prepared diene ester (**189**) which was reduced with DIBAL to the corresponding alcohol (**202**) and selectively oxidised back to the aldehyde (**203**) with reasonable yields in both steps. A magnesium mediated aldol with the thiophenol thioester (**204**) gave the secondary alcohol product (**205**) in good yield; transesterification with SNAc (**191**) gave the functionalised thioester (**206**) in poor yield, which was oxidised to the final product (**207**) in 42% yield.

Multiple issues emerged with this route particularly in the penultimate step which involved swapping of a thiophenol thioester for the NAc thioester. Excess SNAc had to be used to push the transesterification towards completion, and purification of the product from the SNAc was extremely difficult as thiophenol Michael addition side-products were difficult to separate. The final oxidation step also led to isomerisation of the terminal alkene, likely due to trace thiol from the previous step undergoing addition-elimination onto the final product, allowing the alkene to convert to *cis* configuration. The final product was obtained in enough quantity for biological use but an improved Reformatsky

aldol route was used for subsequent preparation: this avoided using thiophenol, preventing alkene isomerism and improved the final product purification (**Scheme 2.9**).
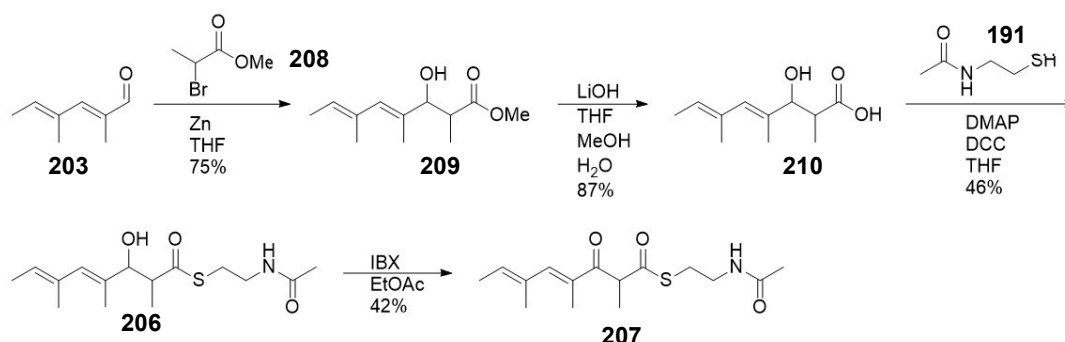
*Scheme 2.9 – **Improved synthetic route to tetraketide SNAc thioester (207)***

A Reformatsky aldol reaction of aldehyde **203** with a bromoester (**208**) gave the methyl ester aldol product (**209**) in 75% yield. Ester hydrolysis gave the corresponding carboxylic acid (**210**) in high yield followed by coupling to SNAc (**191**) to give the previously synthesised secondary aldol product (**206**) in modest yield. IBX oxidation gave the tetraketide product (**207**) in 42% yield.

2.3.2   Restoring thiolactomycin production in a non-producing Δ*tlmA* strain

All four synthetic SNAc compounds were used to supplement a non-producing *Lentzea* sp. Δ*tlmA* strain to check whether production of thiolactomycin could be restored. A control wildtype extract containing thiolactomycin was used to compare retention times and characteristic MS^2 fragments. *Lentzea* sp. Δ*tlmA* was cultured on agar plates and after significant mycelium growth, the SNAc compounds (**192**, **199**, **201** and **207**) were added to them as methanol solutions. The plates were allowed to dry and the bacteria further grow for three days before organic solvent extraction. No production of thiolactomycin was observed when plates were supplemented with acetyl SNAc (**199**). However, the

other more advanced polyketide mimic substrates (**192, 201, 207**) were successfully processed by the cells and the TLM enzymatic machinery: in the extracts of their experiments, thiolactomycin production was clearly detected at the same retention time and with comparable MS^2 fragments to those of wildtype cultures (**Fig. 2.11**). The diketide (**201**) substrate was processed most efficiently with a modest abundance of thiolactomycin production restored; in comparison, the triketide substrate (**192**) was processed in much lower abundance, and for the tetraketide (**207**) an additional peak matching the mass of thiolactomycin was observed (**108**) at a shorter retention time.
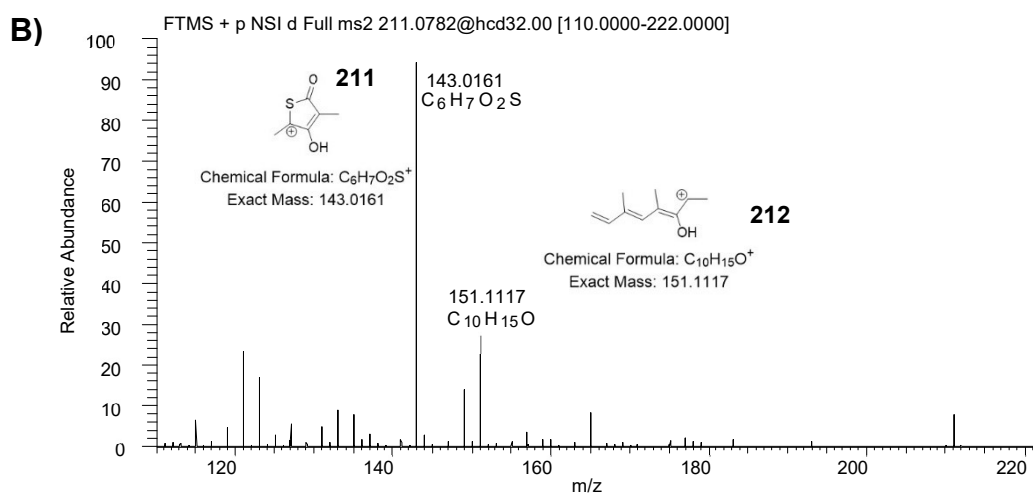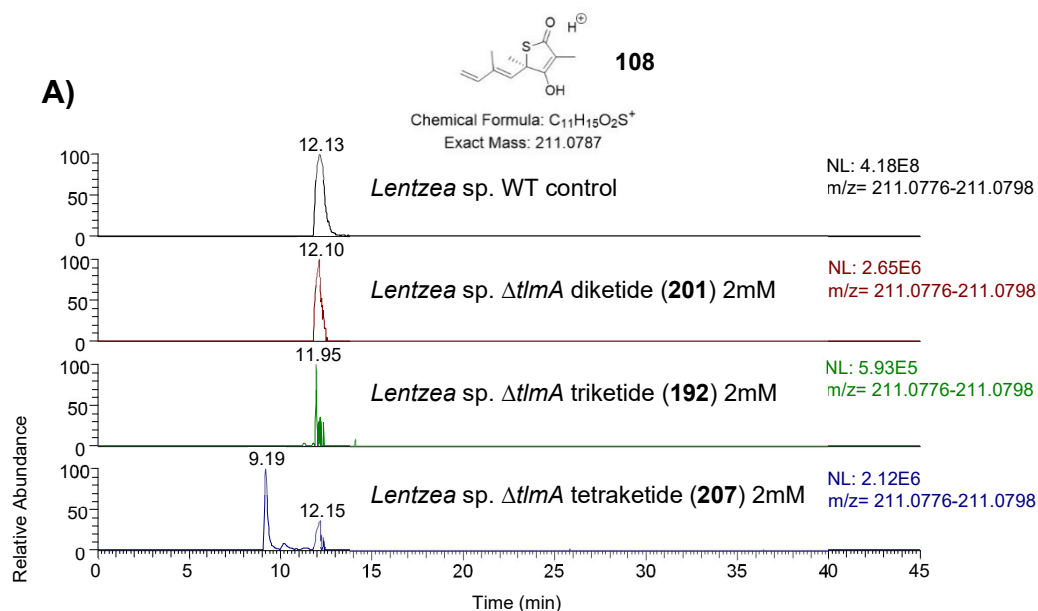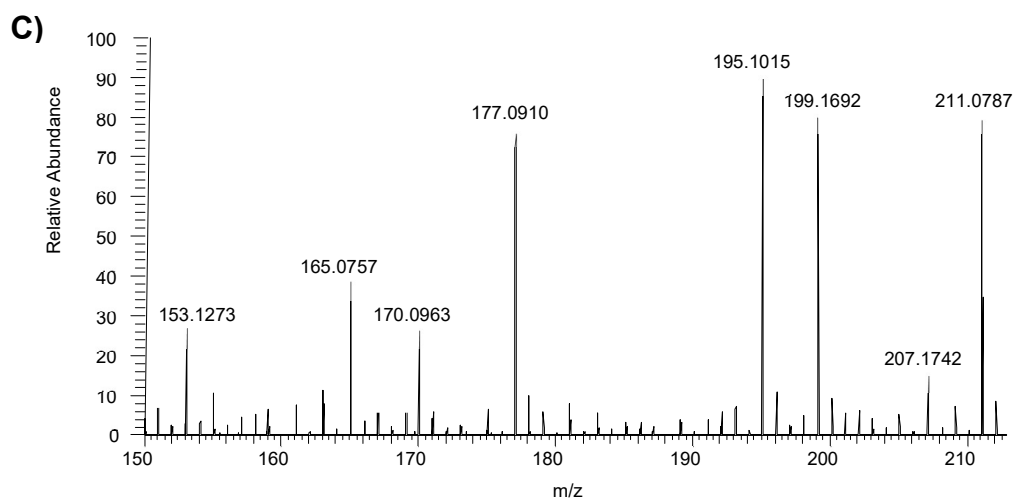
Figure 2.11 **– (A) Detection of restored thiolactomycin (108) product in plates supplemented with the three putative intermediate mimics (201, 192 and 207) at the same retention time as the natural compound from a wildtype extract. (B) Two MS^2 fragments (211 and 212) characteristic for thiolactomycin were detected across all of the restored product peaks. (C) MS^2 trace for unknown identical mass peak at retention time 9.19 from tetraketide (207) supplemented extracts; the structure of this compound remains unclear.**

The identity of the same mass peak with the retention time of 9.19 remains unclear as the MS^2 fragmentation pattern offered little insight. It could possibly be a thiolactomycin isomer as the tetraketide SNAc is a racemate and perhaps one or more of the biosynthetic enzymes is unable to process the unnatural stereochemistry leading to a shunt product, this is still currently under investigation. Nonetheless, these experiments show for the first time that synthetic intermediate mimics can be incorporated into thiotetronate biosynthesis to generate 'unnatural' thiolactomycin.

2.3.3   *In vitro* capture of putative late stage intermediates from TlmB using an SNAc substrate and a chain termination probe

Following on from the *in vitro* assays (**Section 2.2**) showing the activity of TlmB primed by CoA substrates or an SNAc triketide substrate (**192, Scheme 2.5**), Marie Yurkovich further performed a range of other *in vitro* experiments with TlmB and other enzymes in

83

the presence of NADPH, cysteine, the SNAc triketide **192** (to prime TlmB) and the fluoromalonyl 'chain termination' probe (**160**). The results of these assays were analysed my myself and Dr Tosin. Previously, the only captured species from TlmB *in vitro* whilst using the SNAc triketide (**192**) was the triketide portion of the substrate, which only indicated the priming of TlmB but not any form of elongation or processing. Several assays containing TlmB, TlmD1, TueJ, TueS, cysteine, NADPH and **192** were supplemented with the fluoromalonyl probe (carboxylate of **160** after PLE treatment) and subjected to LC-HRMS^n. Some interesting putative intermediates were intercepted in these experiments. They included both species possibly arising from water/methanol catalysed thiirane ring opening (tetraketides **173** and **193**) with MS^2 fragments previously observed (**Fig. 2.10**).

More intriguing was the detection of putative pentaketide intermediates possibly resulting from the same thiirane ring opening (**175** and **177**), featuring MS^2 fragments previously found (**Fig. 2.12**). This suggests that sulphur insertion might occur at both triketide and tetraketide (enzyme-bound) stages. Curiously, the exact same results were detected in parallel assays in the absence of TlmD1, TueJ and TueS, implying that the putative sulphur containing intermediates can be formed independently from these enzymes.

B)

288.1434

C$_{14}$H$_{23}$O$_2$NFS

Chemical Formula: C$_{14}$H$_{23}$FNO$_2$S$^+$
Exact Mass: 288.1428

195

Relative Abundance (y-axis, 0–100)

m/z (x-axis, 150–450)

344.1705

C$_{16}$H$_{25}$O$_3$NFS

Chemical Formula: C$_{16}$H$_{25}$FNO$_3$S$^+$
Exact Mass: 330.1534

176

Relative Abundance (y-axis, 0–100)

m/z (x-axis, 150–500)
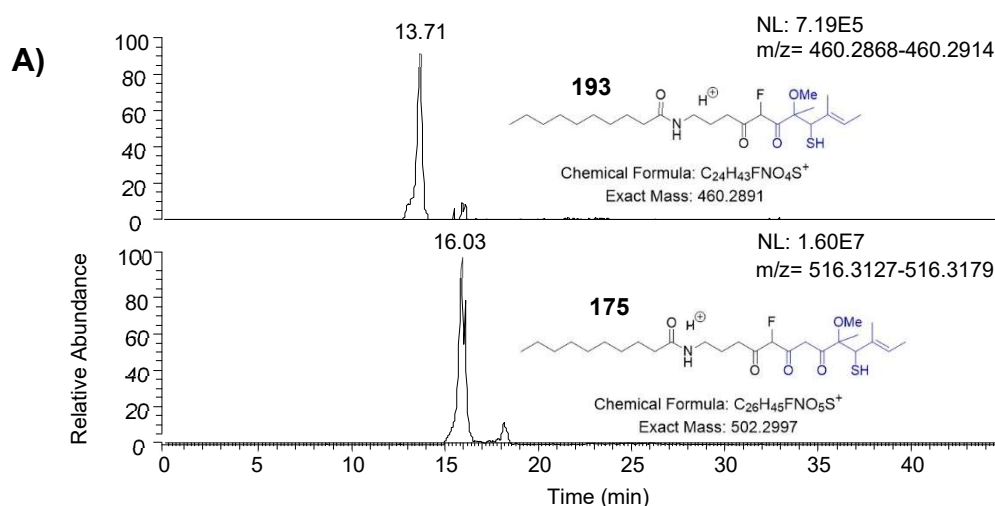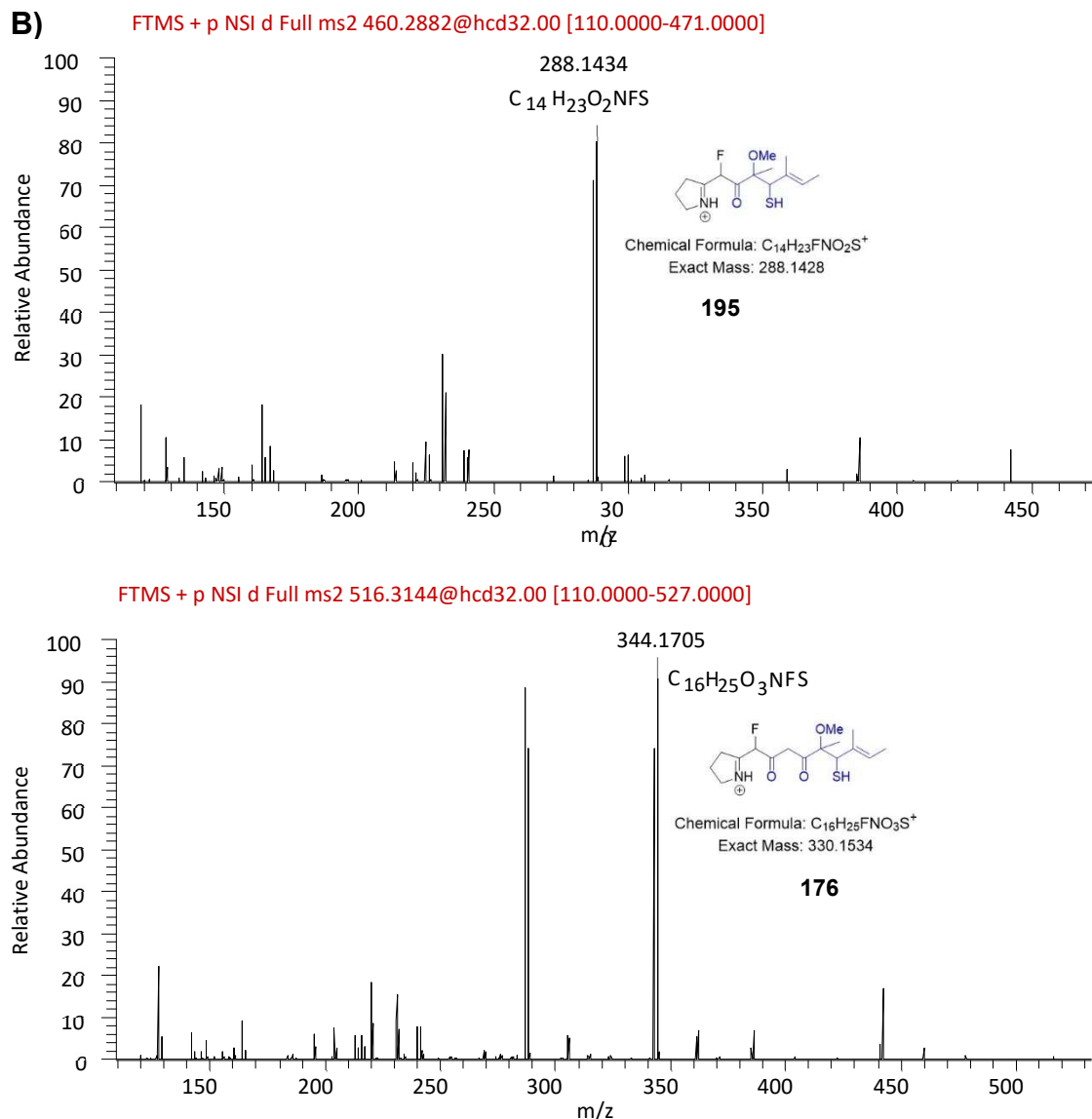
Figure 2.12 – **(A) Detection of intercepted thiirane ring opened tetraketide (193) as well as the pentaketide analogue (175), the H$_2$O opened analogous species were also detected. (B) MS^2 fragments detected for the former parent ions (195 detected for 193 and 176 detected for 175). All intermediates were detected in TlmB assays regardless of TlmD1, TueJ and TueS inclusion.**

As further probed and discussed in **Section 2.3.5**, TlmD1 may not be involved in oxidation preceding sulphur insertion.

2.3.4 Expression and purification of the putative P450 enzyme TlmD1 and substrate incubation

Following the verification that the SNAc tetraketide (**207**) can be processed and converted to thiolactomycin *in vivo* (**Fig. 2.11**), I prepared several substrates for testing *in vitro* as substrates for TlmD1; the predicted P450 enzyme to further investigate the likelihood of the tetraketide being epoxidised whilst enzyme bound. These results, as well as novel compound preparation, are briefly summarised in the following paragraphs.

2.3.4.1 Synthesis of a malonyl tetraketide SNAc mimic (216)

Compound **216** was prepared for testing *in vitro* and *in vivo* as a possible substrate for TlmD1 and/or enzymes involved in TLM biosynthesis, possibly leading to an unnatural thiotetronate, desmethyl-TLM. As earlier described for compound **207**, two routes were undertaken. Issues with the thiophenol aldol route (**Scheme 2.10**) were rectified by using a Reformatsky reaction scheme (**Scheme 2.11**).
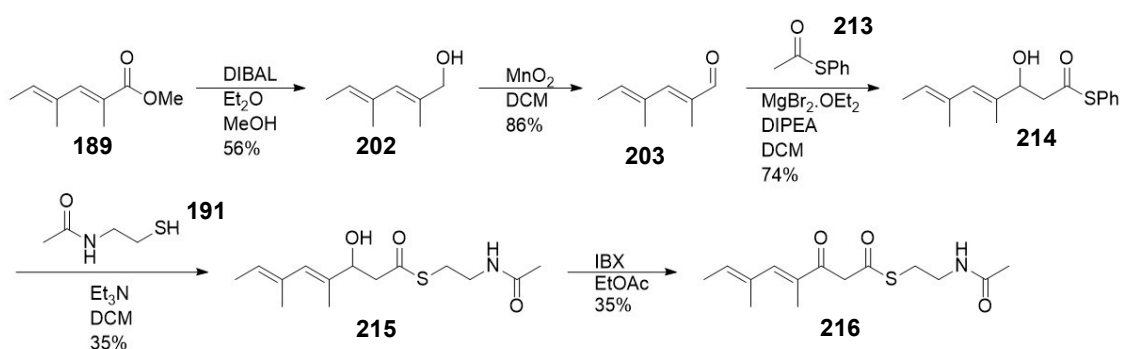
*Scheme 2.10 – **First synthetic route to malonyl tetraketide SNAc (216)***

The original synthesis began with ester (**189**) reduction, using DIBAL, to the corresponding alcohol (**202**) and selective oxidation back to the aldehyde (**203**) in

reasonable yields. A magnesium-mediated aldol reaction with the thiophenol thioester (**213**) gave the secondary alcohol product (**214**) in good yield, transesterification with SNAc (**191**) gave the functionalised thioester (**215**) in modest yield which was followed by oxidation of the alcohol to give the final product (**216**) in 35% yield. Similar issues with purification of the product in the final two steps led to the use of the second route.
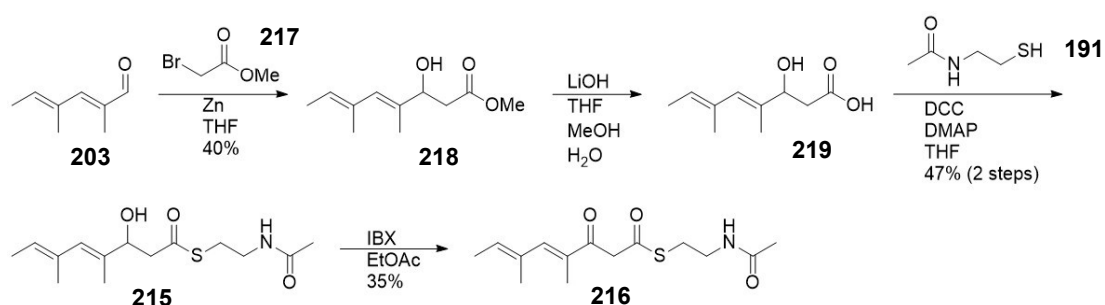
Scheme 2.11 – Second synthetic route to malonyl tetraketide SNAc (216)

The synthesis (**Scheme 2.11**) begins with a Reformatsky reaction with the aldehyde (**203**) using the bromoester (**217**) to give the aldol product (**218**) in 40% yield. Ester hydrolysis to the crude carboxylic acid (**219**) and coupling with SNAc (**191**) gave the alcohol thioester (**215**) in 47% of two steps. Finally, IBX oxidation gave the homologated tetraketide (**216**) in 35% yield.

2.3.4.2 Testing of SNAc triketides and tetraketides as substrates for recombinant TlmD1

The previously reported triketide and tetraketide mimics (**192** and **207**) along with the unnatural homologated mimic (**216**) were used in enzymatic assays by Marie Yurkovich and Karen Chan in Prof. Leadlay's group; we analysed the results of these assays. Each substrate was incubated overnight with NADPH and TlmD1 fused to a ferredoxin reductase partner in order to regenerate the active oxidised iron species (TlmD1-

RhFRED, created by Marie Yurkovich).[137] In each substrate assay, a new species was initially observed with an increased MW exactly matching an atom of oxygen could be consistent with epoxidation (**Fig. 2.13**).
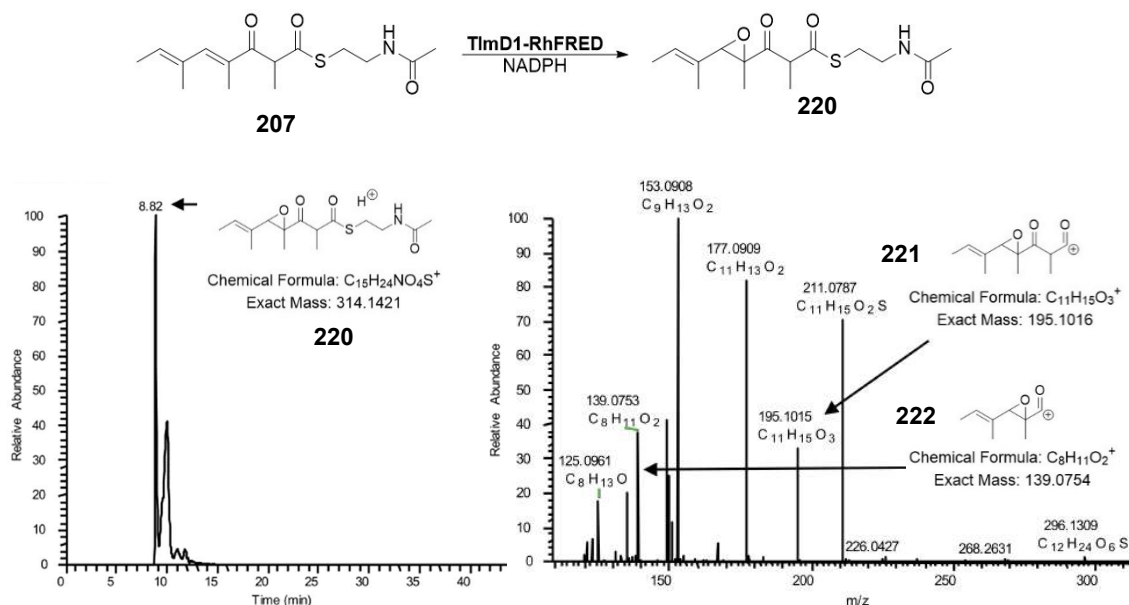
Figure 2.13 – **(Top) In vitro assay performed using SNAc tetraketide mimic (207) to check for possible epoxidation (220) (3 μM TlmD1-RhFRED, 2mM NADPH, 1mM 207 diluted to 100μL TLM buffer (10 mM MgCl₂, 50 mM Tris-HCl, pH 7.2) incubated at RT for 24 h. (Bottom) Extracted ion chromatogram for assay extraction shows putative epoxide product (220) as the protonated ion and two MS² fragments (221 and 222) supporting the structure also detected.**

MS² fragments consistent with this oxygen addition were also detected, however, control extracts with boiled enzymes also showed detection of this species. This indicated that TlmD1 was not performing the reaction and that perhaps the diene moiety is susceptible to oxidation; by itself, in the assay buffer or, more likely, in the MS ionisation process. These assays were performed several times varying temperature and time, I also analysed extracts in parallel by analytical HPLC: only UV peaks corresponding to the starter SNAc molecules were found in every case.

Note: The assay shown above was performed by M. Yurokovich and I analysed the data.

2.3.5 Testing alternative biosynthetic proposal with synthetic substrates

During experiments designed to establish the validity of the epoxide/thiirane formation mechanism previously proposed by Tao *et. al*, a development was published by Tang *et al.* which led to a new biosynthetic proposal for thiolactone ring formation. *S. pacifica* CNS863 is a marine bacterium which also produces thiolactomycin but with a marginally different gene cluster (**Fig. 2.14**).[94] The PKS-NRPS region in this microbe has an identical domain constitution but its modular architecture differs slightly whereby the NRPS TE domain is found further from the other PKS genes. Interestingly, the cluster appears to lack the sulphur metabolism related genes found in the *Lentzea* sp. cluster (TlmJ and TlmS);[94] presumably there are primary metabolic enzymes with similar function that may act in this pathway. As well as a similar resistance gene transcribing a FAS enzyme (*tlmE*), there is a putative cytochrome P450 monooxygenase (encoded by the gene *tlmF* in this cluster) with a function likely similar to TlmD1, although sequence similarity is less than 50%; a single TE domain is also present in the cluster with an unclear function.
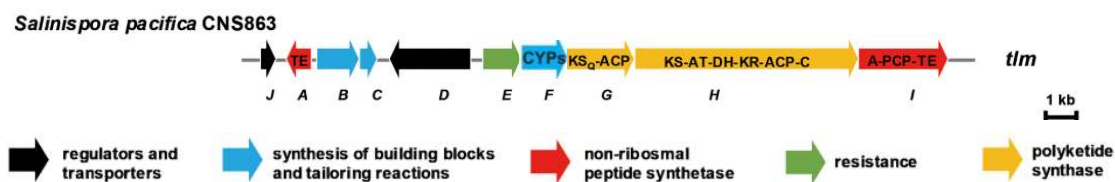
Figure 2.14 – **Gene cluster from *S. pacifica* responsible for producing thiolactomycin. The cluster contains the same PKS-NRPS domains as *Lentzea* sp. with a few minor differences; there is also an essential putative CYP450 monooxygenase (named *tlmF* in contrast to *tlmD1* for *Lentzea* sp.)**

Heterologous expression of the minimal cluster including just the PKS-NRPS region and CYP450 (*tlmF-I*) in *S. coelicolor* M1552 resulted in the production of thiolactomycin. The cluster from *Lentzea* sp. has also been heterologously expressed in *S. albus* by Marie Yurkovich, also producing thiolactomycin. From heterologous expression of the *S. pacifica* cluster without the putative CYP450 TlmF, a shunt product was produced instead, which was purified and characterised by NMR spectroscopy (**223, Scheme**

2.12).[138] The product was shown to be a 6-membered thiolactone, 2 Da heavier than TLM. The formation of TLM has then been rationalised by the proposal of a thiocarboxylic acid (**224**) formation, *O* addition and 5-member ring formation, with the latter two events catalysed by TlmF (**route a, Scheme 2.12**). In the absence of TlmF, Michael addition of thiocarboxylate could occur leading to the formation of the 6-membered thiolactone shunt product (**route b, Scheme 2.12**).
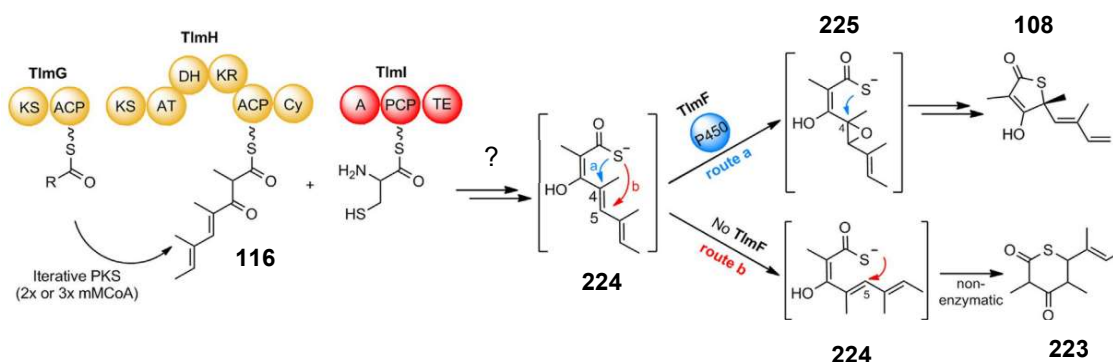
Scheme 2.12 – **Proposed late stage biosynthetic pathway of thiolactomycin from the *S. pacifica* cluster (by Moore *et al.*).[138] By an unknown mechanism, the thiocarboxylic acid intermediate (224) is formed/ released, which is epoxidised (225) to allow ring closure and dehydration to thiolactomycin (108), absence of the CYP450 leads to the 6-membered shunt product (223).**

The isolation of the shunt product (**223**) seems to indicate that the presence of TlmF/TlmD1 is crucial in leading to the correct ring size formation, however it still does not provide direct evidence for the formation of an epoxidised intermediate and still leaves unanswered questions regarding the mechanism of sulphur insertion.

In order to gather insights into the new possible scenario hypothesised by Moore *et al.*, a trifluorinated SNAc diketide substrate (**226**) was by prepared in order to generate truncated late stage biosynthetic intermediates from which terminal proton abstraction (**Scheme 2.13**) could not occur. Hopefully, trifluorinated late stage intermediates may be detectable by LC-HRMS and provide clues on the late mechanistic steps leading to TLM formation.
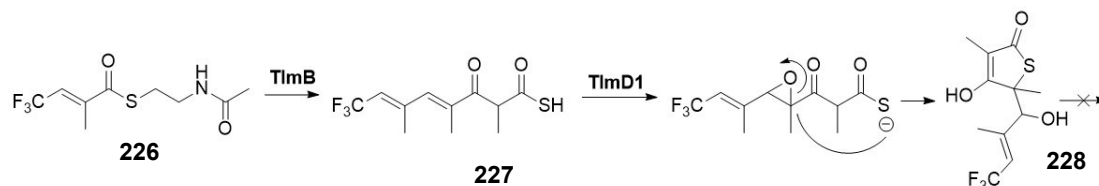
Scheme 2.13 – Processing of the CF₃ containing diketide mimic (226) by Lentzea sp. ΔtlmA would lead to the corresponding tetraketide (227) followed by the subsequent thiolactone (228) which cannot undergo final dehydration and may be detectable by LC-HRMS.

2.3.5.1 Synthesis of trifluorinated diketide SNAc mimic (226)
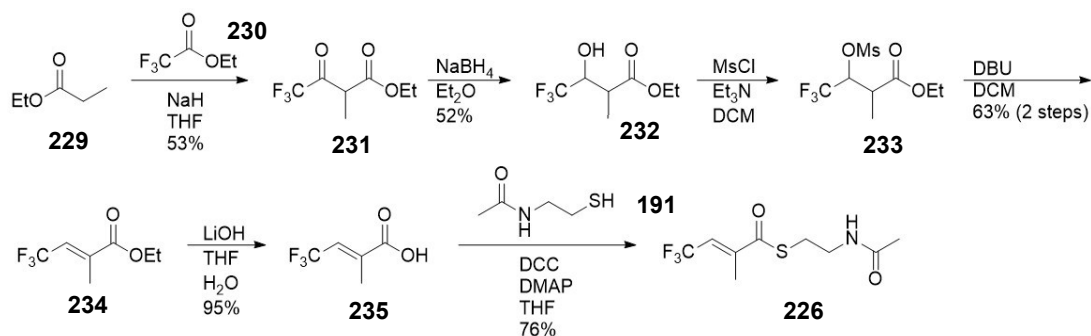
Scheme 2.14 – Synthetic route to trifluorinated diketide mimic (226)

The synthesis (**Scheme 2.14**) begins with a Claisen condensation of ethyl propionate (**229**) and ethyl trifluoroacetate (**231**) to give the keto-ester product (**231**) in 53% yield. Reduction of the ketone with sodium borohydride lead to the alcohol (**232**) in a modest 52% yield. Formation of the crude mesylate (**233**) and elimination with DBU gave the unsaturated ester (**234**) in 63% yield after purification over two steps. Hydrolysis of the ester gave the carboxylic acid (**235**) in 93% crude yield followed by coupling to SNAc (**191**) to give the final diketide mimic (**226**) in 76% yield.

2.3.5.2   Fermentation of *Lentzea* sp. Δ*tlmA* and wildtype with fluorinated SNAc

Feeding of the terminal CF$_3$ substrate **226** to both the *Lentzea* sp. Δ*tlmA* mutant and the wildtype on plate lead to the detection of two possible stalled intermediates of thiolactomycin biosynthesis; neither of these masses were detected in non-supplemented control extracts for both strains (**228/236, 227, Fig. 2.15**). Higher peak abundance was observed in the wildtype strain likely because of the faster growth and higher tolerance to the potentially toxic substrate. The putative parent ions would be consistent with a proposed thiocarboxylic acid intermediate undergoing oxidative 1,5 ring closure possibly mediated by a P450 enzyme. Unfortunately, the poor abundance of MS^2 fragmentation meant no full characterisation could be made to support this proposal. The same experiments carried out on a *Lentzea* sp. mutant strain lacking TlmD1 did not show these putative species.
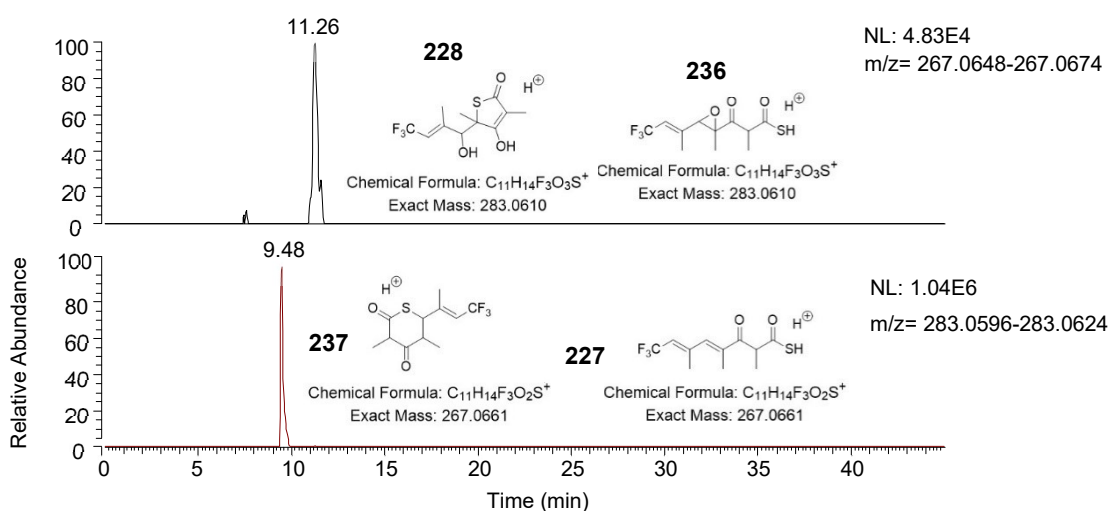
Figure 2.15 – **Extracted ion chromatogram of the two putative intermediates of the fluorinated substrate supplementation in *Lentzea* sp. strains. (Bottom) A putative linear thioacid (227) or ring closed analogue (237) was detected in the wildtype extract. (Top) The putative epoxidised linear or cyclised derivatives (236 or 228) were also detected in the wildtype extract.**

In order to further confirm these intermediates in a different manner, the same feeding experiments were also performed on the Tü3010 thiotetronate producer *S. olivaceus*. It

was predicted that similar intermediates might be generated as in the case of thiolactomycin or perhaps with further ramification as Tü3010 also utilises ethylmalonyl building blocks and has a carboxamide sidechain. Unfortunately, no intermediates were detected to corroborate the putative ions detected in the *Lentzea* sp. feeding study; this could due to the different phenotype displayed by *S. olivaceus* where the plate mycelium prevented even distribution of the substrate over the plate. Instead of low concentration over the entire plate (as with the *Lentzea* sp. feeding experiments), high concentration droplets of substrate dissolved in MeOH evaporated leaving inhibition zones and could have caused poor enzyme processing (**Fig. 2.16**). Small zones of inhibition are seen on the *S. olivaceus* plates indicating toxicity of the substrate in higher concentrations but the toxicity in *Lentzea* sp. appears to be lower.
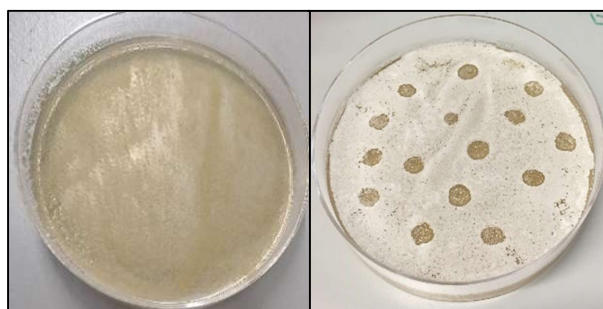
Figure 2.16 – **Comparison of *Lentzea* sp. wildtype (left) and *S. olivaceus* wildtype (right), substrate 226 does not distribute on the *S. olivaceus* plate possibly leading to poor uptake, the substrate appears to be toxic given the zones of inhibition.**

Overall, the *in vivo* experiments carried out with the trifluorinated substrate **226** were not able to shed further light into TLM formation. The preliminary detection of the two previously mentioned species supports the idea that TlmD1 is involved in epoxidation of the unsaturated carbon backbone to allow closure of the 5-membered thiotetronate ring in thiolactomycin biosynthesis.

2.3.5.3 Attempted synthesis of a thiocarboxylic acid intermediate to probe TlmD1 function

In addition to the use of the fluorinated substrate just mentioned, a synthetic route to the proposed thiocarboxylic acid intermediate (**239**) was planned in order to investigate the post-PKS epoxidation mechanism suggested by Tang *et al.*[138] The synthesis (**Scheme 2.15**) was however not completed due to a number of compound stability and purification issues.
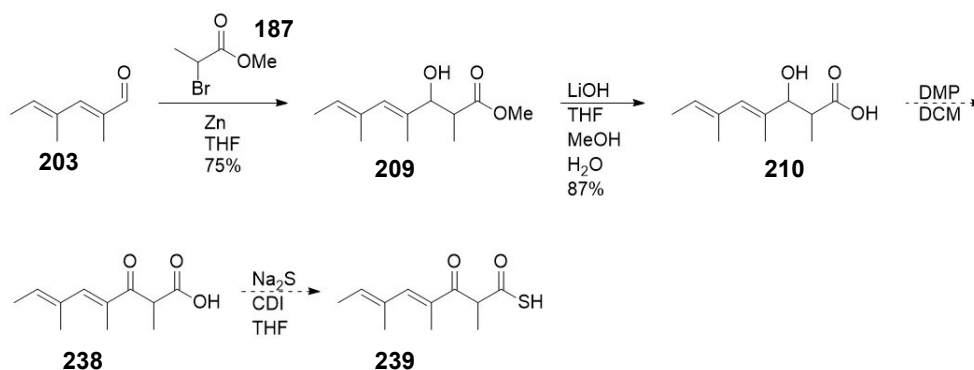
Scheme 2.15 – Planned synthetic route to tetraketide thiocarboxylic acid substrate (239)

The Reformatsky reaction product (**209**) (previously discussed in **Scheme 2.11**) underwent ester hydrolysis to give the hydroxy carboxylic acid (**210**) in 87% yield. Oxidation of this to the corresponding keto-acid (**238**) proved extremely problematic. Both IBX and DMP were utilised, however, degradation was observed with IBX which required heating to initiate any form of reaction. The use of DMP lead to significant decomposition and purification resulted in mixtures of DMP related contaminants along with peaks consistent with the product. It is possible that some Michael "like" addition onto the 1,4 electrophilic starting material is the cause of this; significant isomerism of the substrate terminal alkene was also observed again likely due to addition/elimination of any nucleophile in the reaction, the synthesis was not further progressed.

2.3.5.4 Use of a synthetic cysteine mimic to probe sulphur insertion

According to the original proposal by Tao *et al.*, the PCP domain is presumed to load a cysteine amino acid residue in order to provide sulphur for thiolactone ring formation;[93] desulphuration of cysteine presumably would generate alanine (**Scheme 2.16, top**) or perhaps even dehydroalanine. A non-hydrolysable mimic for a PCP-tethered cysteine was synthesised by PhD student Daniel Leng (Tosin group) to probe intermediates in NRPS biosynthesis (unpublished results). We envisaged this compound might help in providing evidence of cysteine processing within TLM biosynthesis (**Scheme 2.16, bottom**).
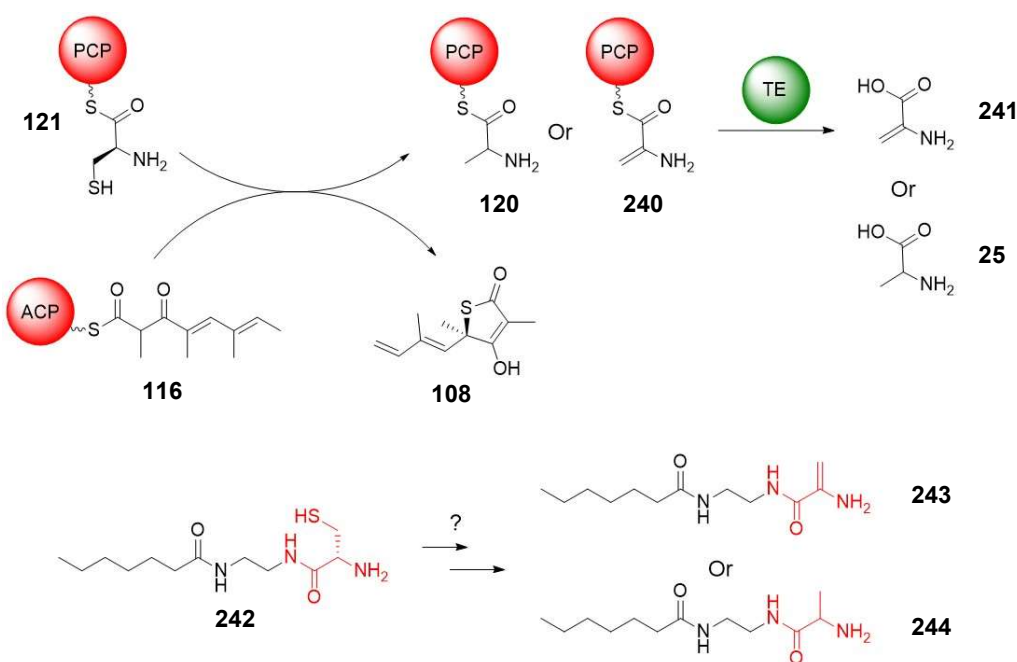
Scheme 2.16 **– (Top) PCP-bound cysteine (121) is used as a sulphur donor for thiolactomycin biosynthesis via an unclear mechanism which presumably leaves either PCP-bound alanine (120) or the dehydro equivalent (250). Hydrolysis of these via TE then leads to alanine (25) or dehydroalanine (241) as by-products. (Bottom) Cysteine based probe (242) made by Daniel Leng is designed to mimic the PCP-bound cysteine but has no thioester linkage so abstraction of sulphur should leave a detectable by-product either as the alanine (244) or dehydroalanine (243) analogue and be detectable by LC-HRMS.**

Fellow group member Daniel Leng and I supplemented **242** to two heterologous expression strains harbouring the TLM cluster with point mutations in the PCP active residue (**Table 2.2**). These strains do not produce thiolactomycin (Vinzent Schulz, Master thesis, University of Cambridge): as the mutations prevents the PCP from having any cysteine bound to it, the lack of thiolactomycin production in these mutants supports the idea that the cysteine moiety is very important, either by providing the S atom of the thiolactone moiety, or by acting as a possible intermediate carrier during cyclisation. The latter scenario will be further discussed in **Section 2.4.4.**

Upon feeding of *S. coelicolor* M1154:Tlm (PCP S2678A) and *S. lividans* TK24:Tlm (PCP S2678A) with probe **242** (2 mM), no restoration of thiolactomycin production was observed. H_2S elimination from probe **242** was observed in high abundance, however the exact cause remains unclear; this elimination could be happening during the mass spectrometry ionisation process since **242** and the H_2S eliminated species coelute. Formation of the alanine or dehydroalanine by-products was not observed. Also, masses corresponding to probe processing were searched for, however nothing meaningful was found. Several reasons could explain this lack of insights, from poor probe permeability and reactivity to inefficient processing of the mutated TLM machinery. Nonetheless a simple explanation could come from the necessity of substrates to be covalently attached to their carrier proteins to be efficiently processed. Indeed ^{35}S-labelled cysteine supplemented to *Nocardia* (now *Lentzea*) sp. was shown to lead to ^{35}S-labelled thiolactomycin. The covalent attachment of substrates to enzymes and their processing will be further explored and discussed in the next sections.

2.3.6   Synthesis and enzymatic loading of pantetheine substrate

To further investigate the possible mechanisms behind sulphur insertion and cyclisation, an ACP-bound tetraketide intermediate (**116**) was chemoenzymatically prepared. Several PKSs, NRPs and accessory enzymes have shown recalcitrance in processing free small molecules and show a distinct preference for enzyme (carrier protein)-bound substrates.

A molecule mimicking the natural tetraketide intermediate was chemically joined to pantetheine *via* thioester linkage formation. This advanced thioester species (**245**) was chemoenzymatically converted to the corresponding CoA derivative and used to convert the standalone TlmB *apo*-ACP to the corresponding *crypto*-ACP **249**. A well-established method using four enzymes (PanK, PPAT, DPCK and Sfp) was used to perform this conversion.[139] The *crypto*-ACP **249** was then incubated with other recombinant enzymes and cofactors in order to investigate the possible processing of its tetraketide moiety in an enzyme-bound form (**Scheme 2.17**).
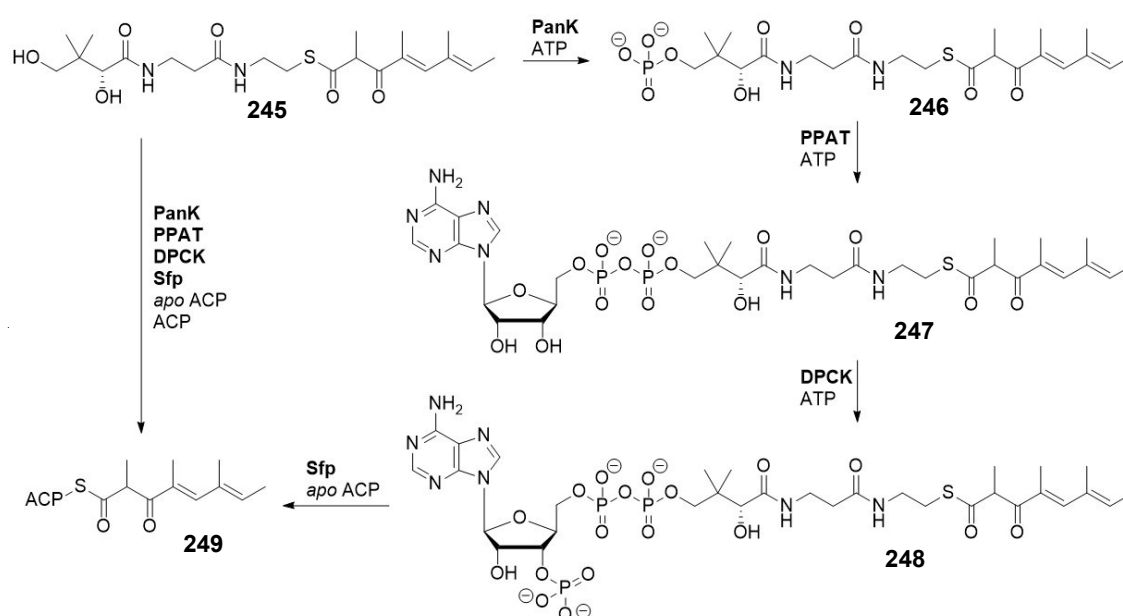
Scheme 2.17 – **Chemoenzymatic conversion of pantetheine substrate and** *apo*-**ACP to functionalised** *crypto*-**ACP (20 mM Tris buffer, pH 7.5, RT, 1 h***). The synthetic pantetheine (245) is phosphorylated* **(246) with PanK followed by adenylation (247) using ATP and PPAT, the pentose is then phosphorylated by DPCK to give the CoA linked pantetheine (248) which is then loaded onto the** *apo*-**ACP by Sfp giving the** *crypto*-**ACP (249).**

The pantetheine substrate (**245**) was synthesised as a diastereomeric mixture due to the inherent instability of the malonyl chiral centre, as previously mentioned in the synthesis of the SNAc analogue (**207**).

2.3.6.1 Synthesis of pantetheine tetraketide mimic (245)

Compound **245** was prepared according to **Scheme 2.18**.
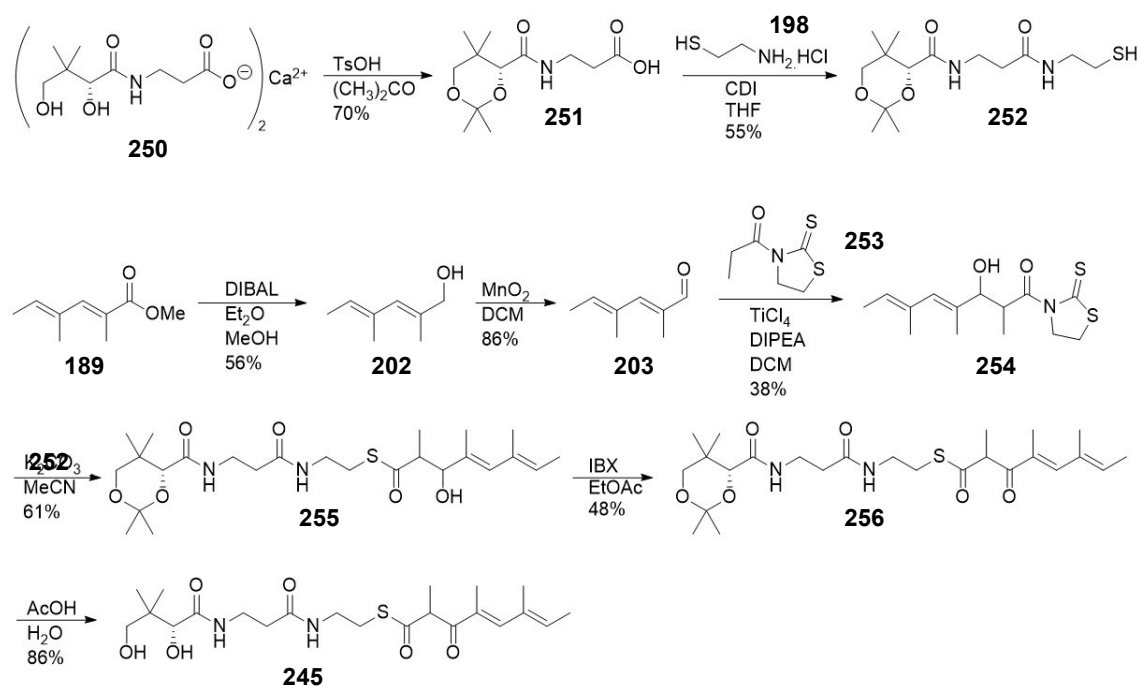
Scheme 2.18 – Synthetic route to pantetheine tetraketide substrate (245).

The diene ester (**189**) was reduced to the alcohol (**202**) and oxidised to the aldehyde (**203**) as previously mentioned (**Scheme 2.10**). A titanium aldol reaction with the thiazolidine thione (**253**) gave the aldol product (**254**) in 38%. This was reacted with the acetonide protected pantetheine (**252**), prepared in two steps by acetonide protection of the pantoic acid hemicalcium salt (**250**) followed by coupling to cysteamine HCl (**198**) to give the thioester product (**255**) in 61% yield. Oxidation of the alcohol using IBX gave the keto-thioester product (**256**) in modest yield and finally removal of the acetonide protecting group gave **245** in high yield.

2.3.6.2 Preliminary ACP pantetheine loading assay

Before utilisation of the *crypto*-ACP with other proteins, the correct formation of *crypto*-ACP from *apo*-ACP was verified. PanK, PPAT, DPCK and Sfp were expressed and purified (by group member Panward Prasongpholchai) along with the standalone ACP region of *tlmB*. The enzymes and cofactors were incubated together for thirty minutes before protein analysis by mass spectrometry. Analysis of the extracts appeared to show complete conversion of the *apo*-ACP to the *crypto*-ACP. However, another species was also present; the *holo*-ACP, possibly from *crypto*-ACP thioester hydrolysis (**Fig. 2.17**).
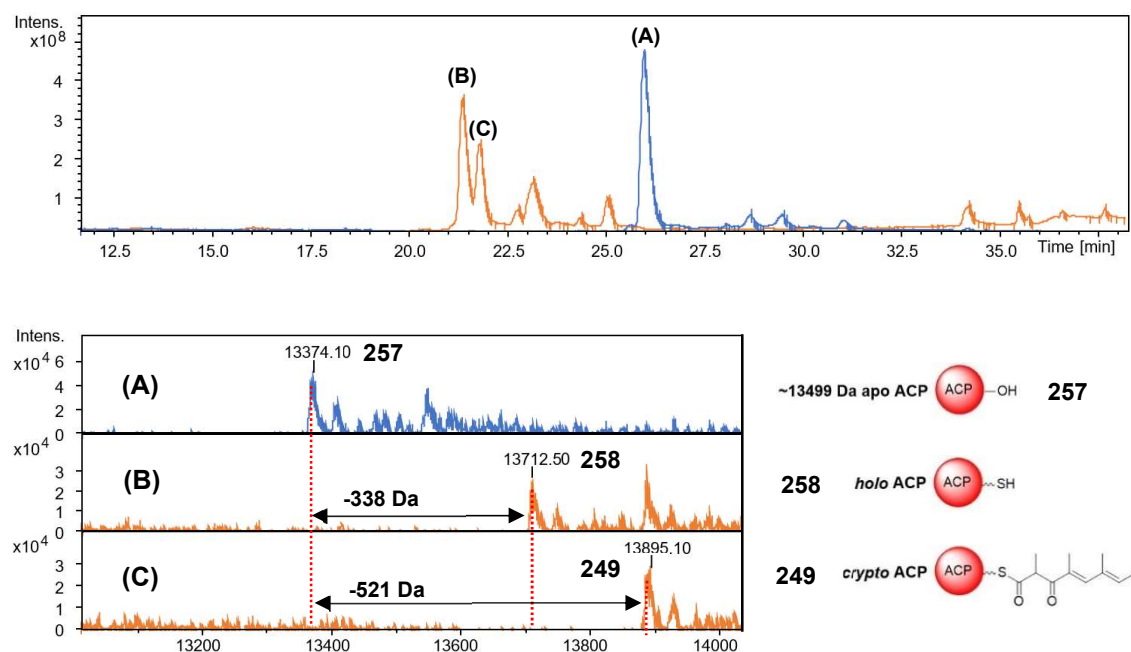
Figure 2.17 – **(Top) LC chromatogram of purified *apo*-ACP (blue) and then after incubation with PanK, PPAT, DPCK, Sfp, *apo*-ACP and 245 in 20 mM tris buffer pH 7.5 at RT for 0.5h (orange). The a*po*-ACP peak (A) disappeared and two new peaks appeared (B and C). (Bottom) HRMS spectra for each of the peaks: (A) is the *apo*-ACP (257), (C) is the *crypto*-ACP (249) (approx. 521Da increase), (B) is the *holo*-ACP (258) which is formed from thioester hydrolysis of the *crypto*-ACP. This indicates the assay was successful and the pantetheine arm was loaded but that thioester hydrolysis also readily occurs.**

This loading experiment was repeated and similar results were obtained. Despite thioester hydrolysis of the *crypto*-ACP to the *holo*-ACP, this assay mixture was immediately

treated with other enzymes and cofactors to minimise ACP-bound tetraketide degradation and hopefully maximise the chance of observing some transformations.

2.3.6.3 Enzyme assays with standalone "*crypto*" ACP

The first enzymes to be incubated with the *crypto*-ACP were both available forms of recombinant TlmD1; these were the standalone TlmD1 protein and the reductase-fused protein (TlmD1-RhFRED, both constructs prepared by Marie Yurkovich). Both of the putative P450 purified enzymes were expressed alongside the standalone cyclisation (Cy) domain and *apo*-ACP (**Fig. 2.18**). However due to lack of available MS instrument time, only the TlmD1-RhFRED was used for these assays in the first instance. It was envisaged that intact protein MS would show whether an epoxide is being formed on the tetraketide whilst enzyme bound. The sulphur insertion mechanism was also investigated by using the previously discussed cysteine mimic probe (**242**) to act as an artificial PCP-bound cysteine. The use of this mimic as a sulphur donor for making thiolactomycin was also expected to show whether an elimination or reduction mechanism occurs to leave PCP bound alanine/dehydroalanine and maybe help understand the role of the Cy domain.
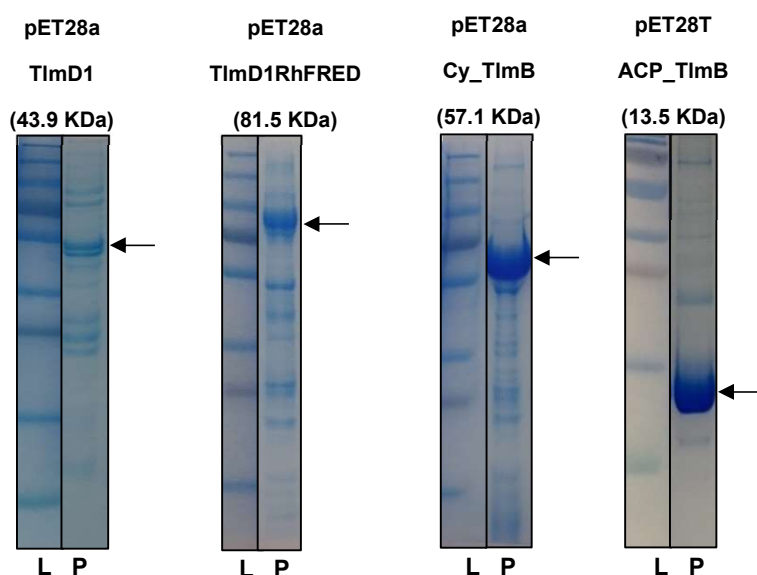
Figure 2.18 – **SDS-PAGE gels for purified proteins and their sizes (L = ladder, P = purified protein fraction).**

The loading assay was incubated as previously mentioned (**2.3.6.2**) to generate the *crypto*-ACP. Aliquots of this were then incubated for thirty minutes with a range of different enzyme/cofactor concentrations (**Scheme 2.19** and **Table 2.3**) before running MS analysis of the ACP species. Protein and cofactor concentrations were estimated according to previous *in vitro* work by Marie Yurkovich.[137]
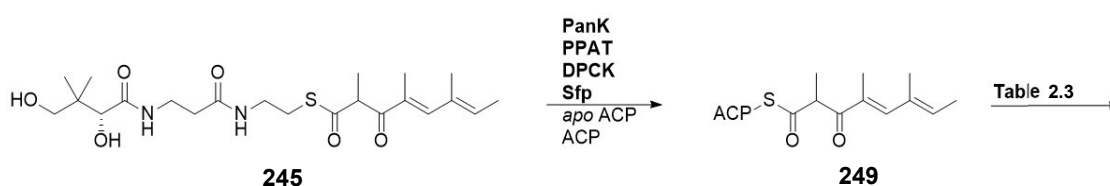
Scheme 2.19 – Enzymatic assays with *crypto*-ACP. The synthetic pantetheine derivative (245) was loaded onto the *apo*-ACP to generate the corresponding *crypto*-ACP (249), which was then incubated with a variety of enzymes and cofactors.

Assay #	*Crypto*-ACP	TlmD1-Rhfred	NADPH	Cy (TlmB)	Cysteine probe (242)
1	50 μM	3 μM	2 μM		
2	50 μM	20 μM	2 μM		
3	50 μM			5 μM	250 μM
4	50 μM			20 μM	250 μM
5	50 μM	3 μM	2μM	5 μM	
6	50 μM	15 μM	2μM		
7	50 μM	15 μM	2μM	15μM	

Table 2.3 – Table of enzyme/cofactors used for assays performed with *crypto*-ACP; various concentrations and combinations of enzymes and cofactors were tested (20 mM tris buffer, pH 7.5, RT, 2 h).

Due to time and instrument constraints, control experiments for these assays were not run, if anything interesting was observed then specific experiments would be rerun with the necessary boiled enzyme controls. This however was not necessary as analysis of all reaction traces showed the presence of *crypto*-ACP and *holo*-ACP alone as observed previously. No novel peaks corresponding to new ACP species were detected under any of the conditions described in the previous table.

It is not clear whether the hydrolysis of the *crypto*-ACP might lead to highly abundant reactive thiolate species which may hinder the assay in some way or if perhaps the other enzymes utilised were not active. The resolution of the LCMS instrument was not particularly sensitive towards these types of intact protein experiments and deconvoluting the spectra proved difficult, so very small intensity peaks may have been overlooked. In either case, few conclusions could be drawn from these experiments which may benefit from repeating them in the near future.

2.4   Investigating substrate specificity in thiotetronate assembly

Given the successful processing of SNAc substrates mimicking natural intermediates into thiolactomycin formation using the Δ*tlmA* strain, other "unnatural" SNAc compounds were synthesised in the hope of chemoenzymatically producing thiolactomycin analogues. Recently, Li *et al.* demonstrated that manipulation of the thiotetromycin cluster in *Streptomyces afghaniensis* allowed production of more ramified thiotetronate analogues with increased antibiotic activity; this was achieved by adapting a pair of oxidation and amidation enzymes into the cluster to activate a C-H centre and allow conversion to a carboxamide.[140]

2.4.1 Synthesis of "unnatural" SNAc substrates

A range of novel or mono, di, tri and tetraketide SNAc compounds were synthesised for feeding experiments with selected thiotetronate producers (**Fig. 2.19**). These would incorporate unnatural moieties to alter the properties of the final compound in terms of activity and give information about the promiscuity of the enzyme machinery.
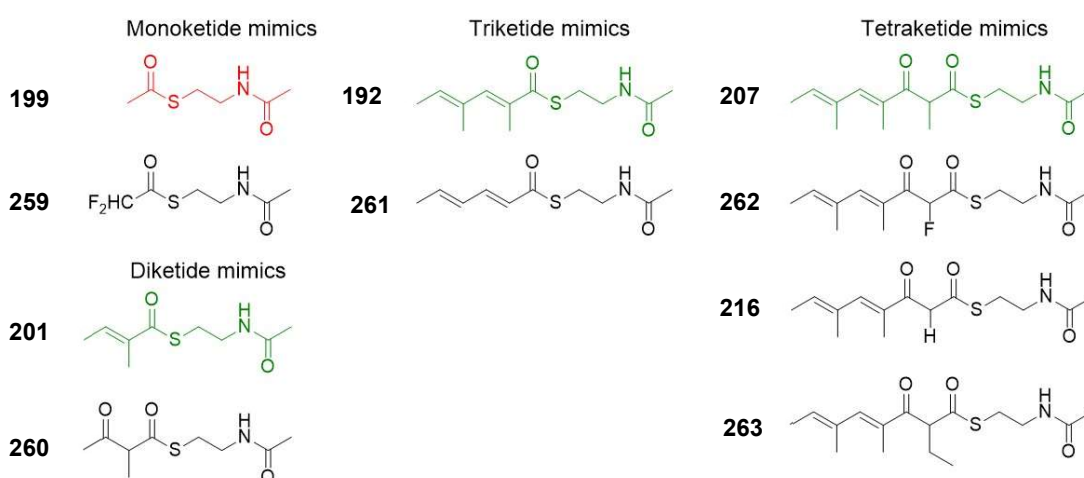
Figure 2.19 – **Structures of all SNAc substrates synthesised. Some of these (in green) were shown to be processed to generate thiolactomycin (Section 2.3.2), except for the monoketide mimic (in red); the remaining (black) were expected to possibly lead to unnatural thiotetronate formation.**

Compounds **259** and **262** were synthesised in the hope that fluorine could be incorporated into a thiotetronate compound: if successful, this would be the first example of its kind. Compounds **261** and **263** were designed in order to probe the specificity of the iterative PKS as many thiotetronates are different only in the extender unit they use: these may show the limit of what alkyl group size can be positioned around the ring and on the side chain. Finally, **260** would show whether the non-reduced or dehydrated substrates could be primed and then processed before elongation.

2.4.1.1 Synthesis of difluorinated monoketide mimic (259)
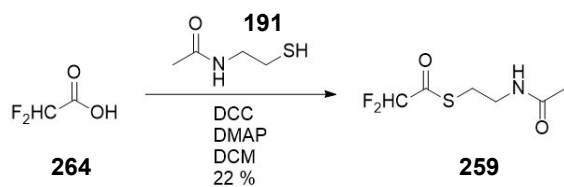
Scheme 2.20 – Synthetic route to difluoroacetyl monoketide SNAc (259)

A single step coupling of difluoroacetic acid (**264**) with SNAc (**191**) afforded the product (**259**) in 22% yield (**Scheme 2.20**).

2.4.1.2 Synthesis of unreduced diketide mimic (260)
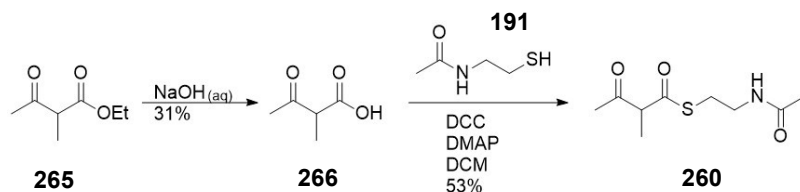
Scheme 2.21 – Synthetic route to unreduced diketide SNAc (260)

Hydrolysis of the commercially available keto ester (**265**) gave the corresponding carboxylic acid (**266**) in 31% yield, coupling with SNAc (**191**) gave the final compound (**260**) in 53% yield (**Scheme 2.21**).

2.4.1.3 Synthesis of non-branched triketide mimic (261)
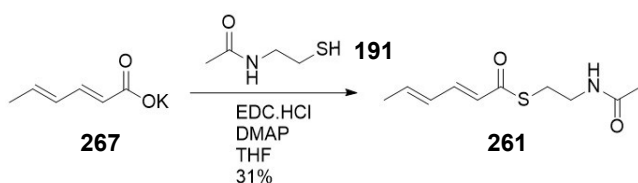
Scheme 2.22 – Synthetic route to the non-branched triketide SNAc (261)

Coupling of the potassium sorbate (**267**) with SNAc (**191**) gave the non-branched triketide (**261**) in 31% yield; some solubility issues were observed, which may explain the poor yield (**Scheme 2.22**) (synthesis completed by Aiste Andriulyte, URSS project student).

2.4.1.4 Synthesis of fluoromalonyl tetraketide mimic (262)

Similar to the previous tetraketide syntheses, two synthetic routes were undertaken to make compound **262**. The first route used a thiophenol aldol reaction, which presented the issues previously encountered (**Scheme 2.23**). A Reformatsky reaction was used again offering easier purification, purer intermediates and pure final product (**Scheme 2.24**).
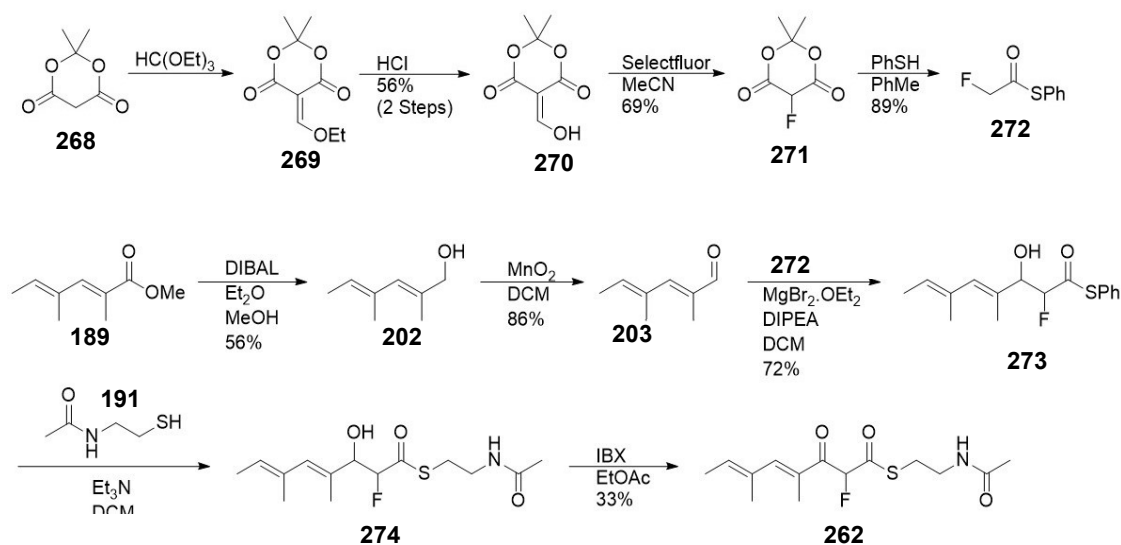
Scheme 2.23 – **First synthetic route to achiral fluoromalonyl tetraketide SNAc (262)**

Meldrum's acid (**268**) was converted to the enol ether (**269**) followed by acid treatment to give the enol (**270**) in 56% yield over two steps. This was then converted to fluoro Meldrum's acid (**271**) using Selectfluor in 69% yield followed by reflux with thiophenol to give the fluoroacetyl thioester (**272**) in 89% yield. This was used in an aldol reaction with aldehyde (**203**) to give the aldol product (**273**) in 72% yield. Thioester exchange

with SNAc (**191**) gave the SNAc ester (**274**) in 31% yield, followed by IBX oxidation to the final product (**262**) in low yield.
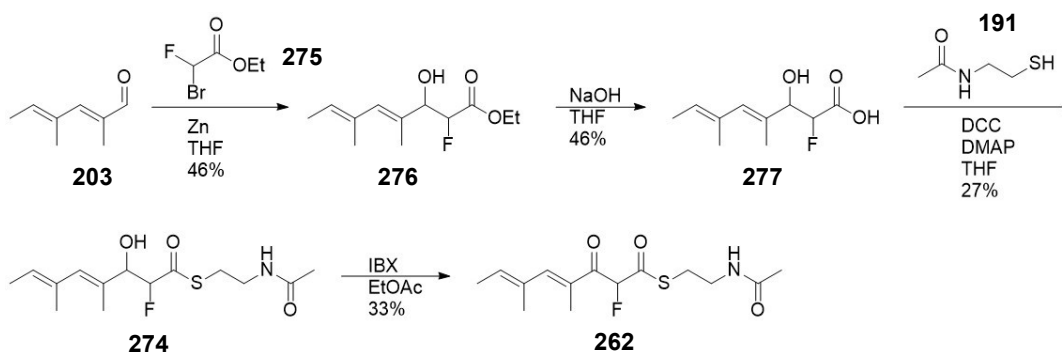
Scheme 2.24 – Second synthetic route to achiral fluoromalonyl tetraketide SNAc (262)

A Reformatsky reaction of the aldehyde (**203**) with the fluorobromo ester (**275**) gave the aldol product (**276**) in a moderate yield; ester hydrolysis to the acid (**277**) gave a poor yield. Coupling with SNAc (**191**) followed by IBX oxidation of the functionalised thioester (**274**) gave the final product (**262**) in acceptable yield (**Scheme 2.24**).

2.4.1.5   Synthesis of ethylmalonyl tetraketide mimic (263)
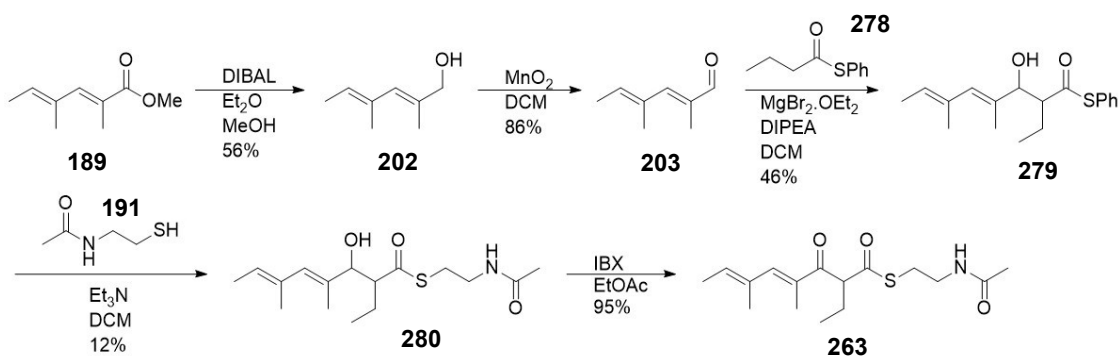
Scheme 2.25 – Synthetic route to ethylmalonyl tetraketide SNAc (263)

The diene aldehyde (**203**) was subjected to a magnesium bromide mediates aldol reaction with a butyl thioester (**278**) to give the aldol product (**279**) in average yield. Thioester

106

exchange with SNAc (**191**) was extremely poor yielding but the resulting alcohol (**280**) was oxidised to the final compound (**263**) in extremely good yield. Unlike the previous tetraketide examples, enough of this substrate was synthesised for all necessary feeding experiments with the poorer thiophenol aldol route (**Scheme 2.25**). The final compound was purer than had been seen with other examples, but NMR still showed a mix of terminal alkene structural isomers. It was decided in the first instance to use this "crude" mixture for experiments, the other improved Reformatsky procedure was not used to remake this substrate; full NMR characterisation was also not achieved given the complicated mix of inseparable structural isomers.

2.4.1.6 Attempted synthesis of trifluoromethylmalonyl tetraketide mimic (285)

Synthesis of another halogenated substrate was attempted (**Scheme 2.26**). A boron and peroxide-mediated aldol reaction of the diene aldehyde (**203**) with a trifluoro methyl ester (**281**) gave the alcohol product (**282**) in 55% yield. Complete degradation was observed in the ester hydrolysis step to form the β-hydroxy carboxylic acid (**283**). A literature search revealed little precedent for the synthesis of β-hydroxy carboxylic acids with a trifluoromethyl substituent in the α position. Proton NMR analysis of the crude product revealed a reverse aldol reaction had occurred back to the aldehyde and so the route was no longer pursued. The electron withdrawing nature of the CF$_3$ group may lead to increased acidity of the alcohol proton allowing the reverse aldol to occur.
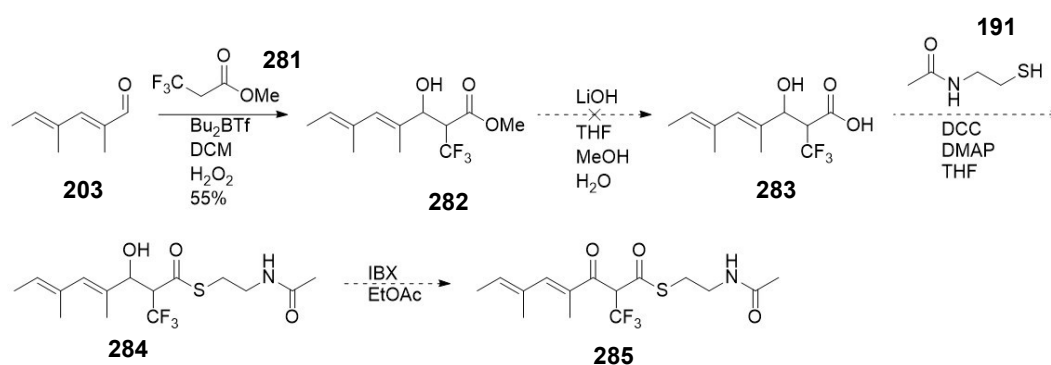
Scheme 2.26 – Attempted synthetic route to trifluoromethylmalonyl tetraketide SNAc (285)

2.4.2 Testing of novel SNAc substrates in wildtype *Lentzea* sp. and Δ*tlmA* mutant

Firstly, all "unnatural" substrates were tested with the *Lentzea* sp. Δ*tlmA* strain for production of unnatural analogues. After repeated feeding experiments with these substrates, all except **216** showed no significant sign of any novel thiotetronate production. Supplementation of **216** led to the detection of a parent ion matching des-methyl thiolactomycin (**286**) as observed by LC-HRMS. Analysis of the wildtype *Lentzea* sp. control extract showed an identical mass peak but at a marginally later retention time, indicating perhaps a different regioisomer (**Fig. 2.20**). MS^2 analysis of the wildtype extract shows an identical ring fragment to that generated from thiolactomycin fragmentation, suggesting that the thiotetronate ring is intact and that the side chain is missing a methyl group. This could be due to malonyl coenzyme-A (**10**) usage instead of the usual methylmalonyl coenzyme-A (**15**) for a single round of chain extension. No assignable/diagnostic MS^2 fragments were collected for the putative structure **286** due to the low abundance of the species.
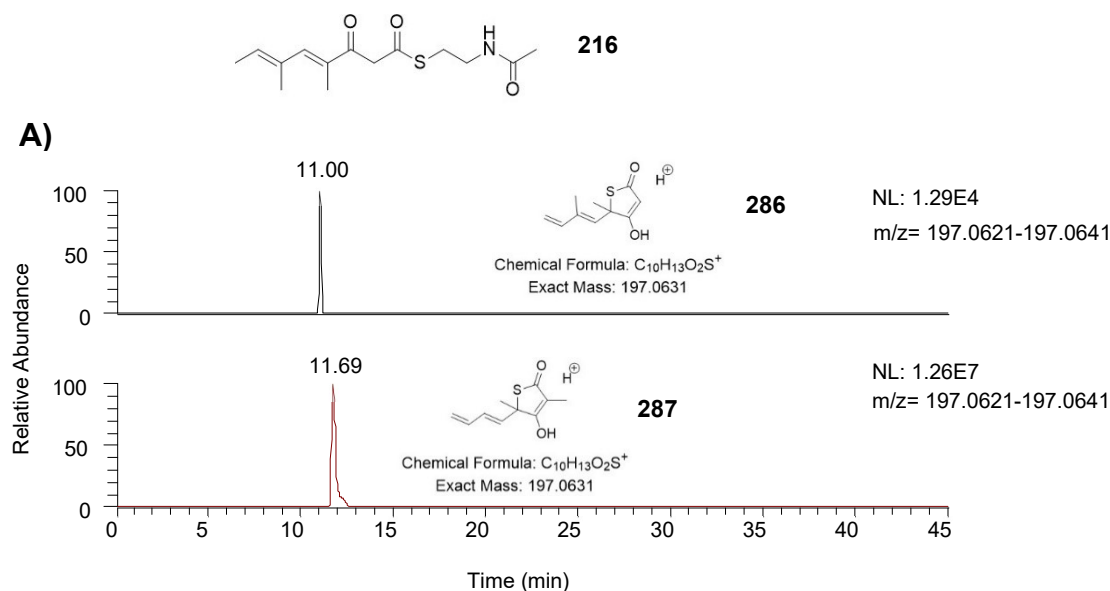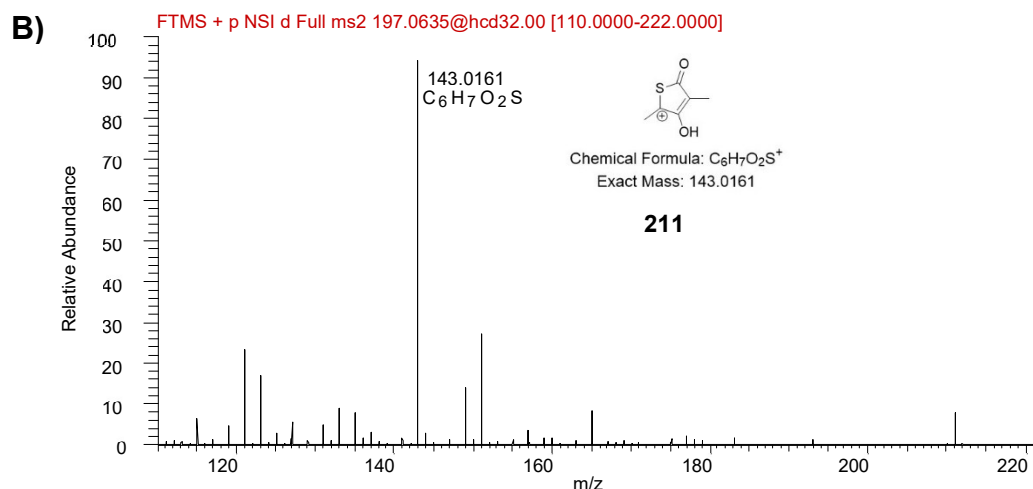
Figure 2.20 – **(A) Extracted ion chromatograms of (top) 216 supplemented *Lentzea* sp. Δ*tlmA* extracts and (bottom) control wildtype *Lentzea* sp. Both detect a mass corresponding to des-methyl thiolactomycin (286 or 287) but with different retention times. (B) MS² for the wildtype strain detected ion supports the hypothesis that the side chain is missing a methyl group as the whole ring fragment (211) was detected.**

As the processing of substrates **259** and **261-263** would lead to unnatural thiotetronates, they were also tested in the wildtype *Lentzea* strain, for which better growth and higher tolerance to exogenous substrates has been observed in comparison to the Δ*tlmA* mutant. Unfortunately, no novel thiotetronate production was observed in these experiments, revealing a tight substrate specificity for TlmB.

2.4.3 Testing of novel SNAc substrates on TueB in wildtype *S. olivaceus* and the heterologous strain *S. avermilitis* harbouring the *stu* cluster

Substrates were also tested in two other strains, the Tü3010 (**110**) wildtype producer *S. olivaceus* (*tue* cluster) and also in the *S. thiolactonus* (*stu*) cluster which was inserted and heterologously expressed in *S. avermilitis* (work of Prof. Yuhui Sun of Wuhan University). The *stu* cluster produces Tü3010 (**110**) as well as thiotetromycin (**109**) (which appears to be an intermediate of Tü3010) (**Fig. 2.21**),[93] this cluster inserted into

S. avermilitis lacks the *stuA* gene (the PKS loading module) and therefore cannot produce any thiotetronate products in a similar manner to the *Lentzea* sp. Δ*tlmA* strain.
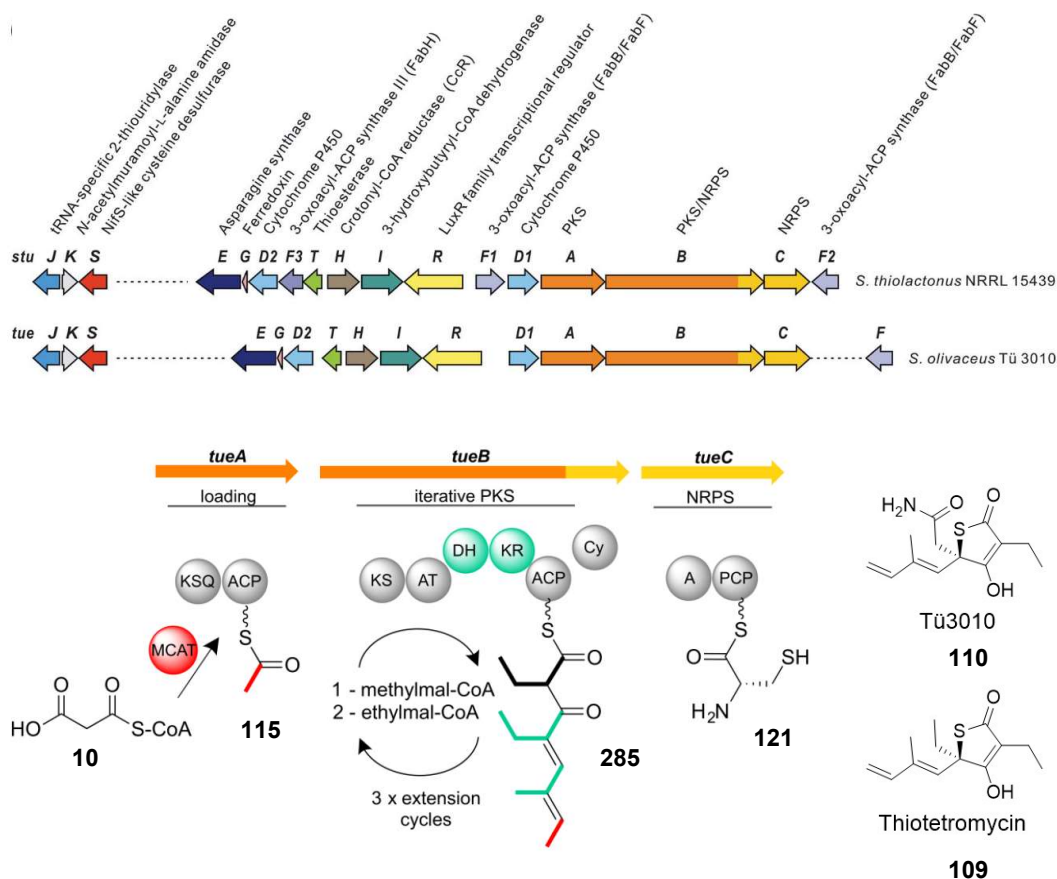
Figure 2.21 – **(Top) Gene clusters for *S. olivaceus* (*tue*) and *S. thiolactonus* (*stu*). (Bottom) Proposed domain architecture for the *tue* cluster (which appears to be the same as *stu*), an acetate starter unit (115) derived from malonyl coenzyme-A (10) is elongated once with malonyl coenzyme-A and twice with ethylmalonyl coenzyme-A to give a tetraketide backbone (285) which utilises cysteine-PCP (121) to generate thiotetromycin (109) which is further processed into Tü3010 (110).**

Experiments aiming at restoring Tü3010 production using natural substrates as previously done with *Lentzea* sp. Δ*tlmA* were planned for the *S. avermilitis*:Stu Δ*stuA* strain, however this strain still shows a low level of Tü3010 production, so careful data analysis of extracts was carried out looking at comparative levels of Tü3010 between unfed and fed strains. An overview of the results obtained from supplementation of *S. avermilitis:*Stu Δ*stuA* and *S. olivaceus* WT with substrates **201**, **192**, **216**, **207** and **262** is given below (**Table 2.4**).
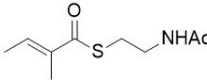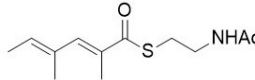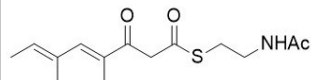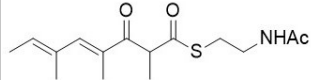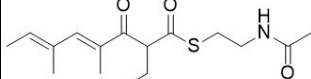
	Wildtype *S. olivaceus*	*S. avermitilis*:Stu Δ*stuA*
201	Tü3010 production unaffected.	Tü3010 production unaffected.
192	Severely diminished Tü3010 production.	Increased production of Tü3010 (compared to control unfed strain). Novel peak at m/z 254.1 (possibly des-methyl Tü3010).
216	Tü3010 production unaffected.	Tü3010 production unaffected, no novel thiotetronates detected.
207	Tü3010 production unaffected. Novel peak at m/z 211.1 (possibly thiolactomycin).	Tü3010 production unaffected. Novel peak at m/z 211.1 (possibly thiolactomycin).
262	Tü3010 production unaffected. Novel peak at m/z 225.1 (possibly des-methyl thiotetromycin).	Tü3010 production unaffected. Novel peak at m/z 225.1 (possibly des-methyl thiotetromycin).

Table 2.4 – **Summary of results from feeding experiments of unnatural SNAc substrates to *S. olivaceus* and *S. avermitilis* (heterologous host for *stu* cluster without *stuA*) strains. Substrate 192 had the dramatic effect of decreasing Tü3010 production in *S. olivaceus* and possibly generating a novel thiotetronate in *S. avermilitis*. The two tetraketide substrates 207 and 262 are also apparently processed into thiolactomycin (108) and des-methyl thiotetromycin (287) respectively.**

The results indicate two particularly interesting observations. Firstly, the triketide substrate (**192**) supplemented to *S. avermilitis* appeared to show production of a novel thiotetronate having undergone oxidation and amidation to a des-methylated Tü3010 derivative (**286**, **Fig. 2.22**). *S. olivaceus* appears to produce a same MW thiotetronate likely from incorrect malonate loading (similar to previously seen in *Lentzea* sp. wildtype extracts) but the retention times are different suggesting both are different regioisomers. Assigning the regioisomer produced in *S. olivaceus* is difficult owing to the lack of diagnostic fragments to pin point where the methyl group is absent, however the triketide substrate (**192**) feeding should lead to a methyl group missing on a branch of the

thiotetronate ring (**286**). This position normally undergoes oxidative amidation which shows the post PKS-NRPS machinery can still function with this different ring structure.
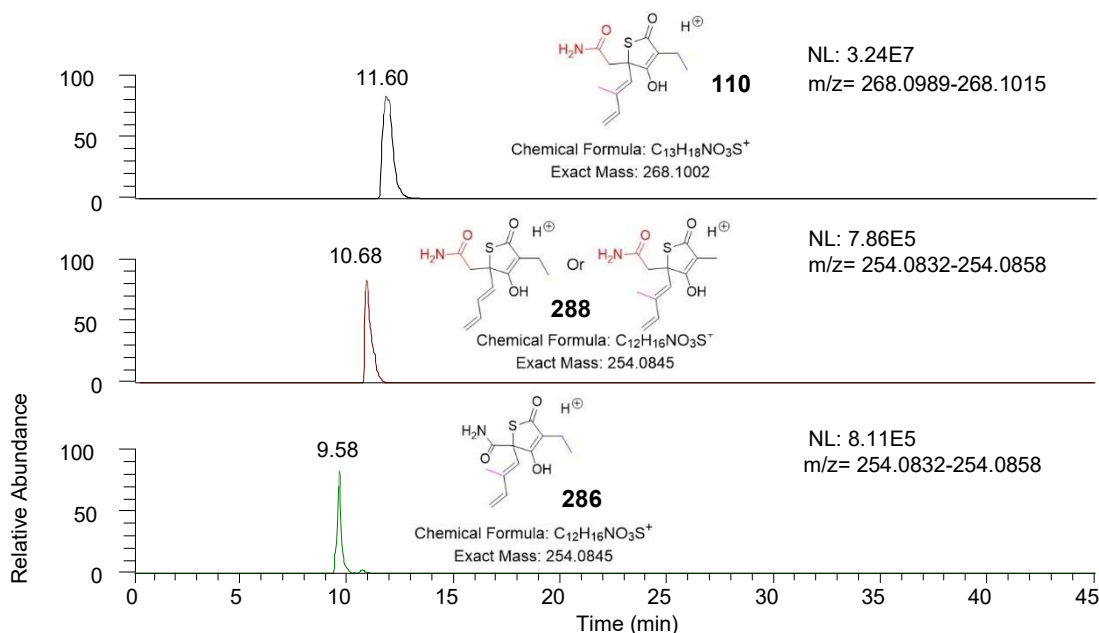
Figure 2.22 – **Extracted ion chromatograms of Tü3010 (110) (top), in the same wildtype extract was detected a des-methyl Tü3010 analogue (288) (middle) with unclear regiochemistry. In the triketide SNAc (192) supplemented *S. avermilitis*:Stu Δ*stuA* extracts was detected another des-methyl Tü3010 analogue (286) (bottom) which has a different retention time indicating perhaps a regiomeric isomer.**

Despite the promising detection of the new species in the extracts of *S. avermilitis*:Stu Δ*stuA* supplemented with the triketide SNAc (**192**), MS^2 fragment analysis offered little confirmation. Interestingly, supplementing the same substrate to *S. olivaceus* did not give the same Tü3010 analogue (**286**) and instead the production of Tü3010 (**110**) itself was severely hindered. All other substrate extracts had high levels of Tü3010 production and therefore little effect on the biosynthesis, this appears to confirm that there is some kind of processing of the triketide with these Tü3010 producing clusters in that both processing and inhibition occurs in each of the two strain. The second observation is the possible processing of the malonyl substituted tetraketides **207** and **262** into thiolactomycin (**108**) and des-methyl thiotetromycin (**287**) respectively in both strains (**Scheme 2.27**), these

peaks are absent in the control extracts but are not of high enough intensity to gather sufficient MS^2 data for characterisation from. Their presence would suggest that the PKS-NRPS machinery involved in thiotetronate ring formation are not as substrate specific but the enzymes responsible for oxidation and amidation are highly specific.
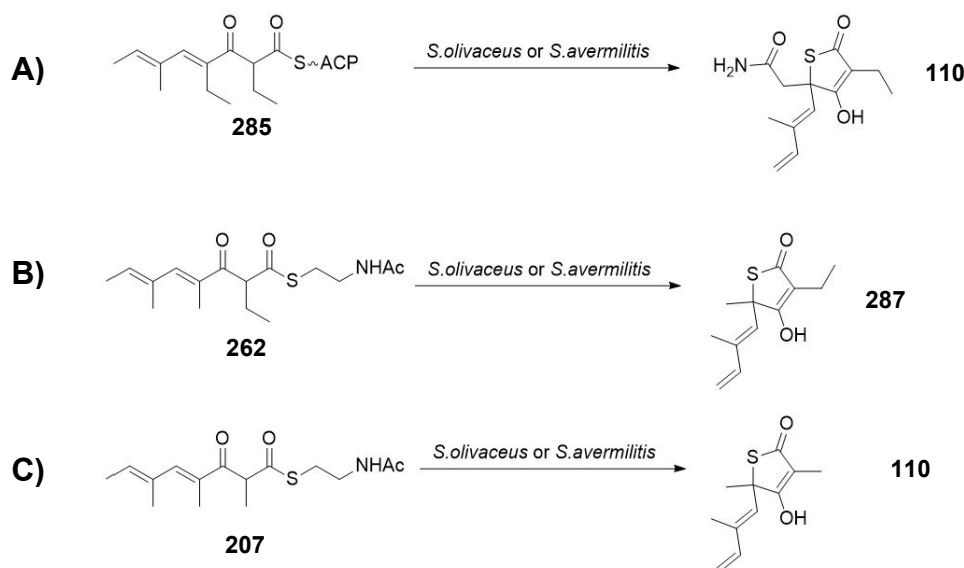
Scheme 2.27 – **A) The native tetraketide backbone (285) built by both the *tue* and *stu* PKS-NRPS is converted into Tü3010 (110) and is observed in control extracts for both strains used. B) Extracts of both strains supplemented with ethylmalonyl tetraketide (262) appear to process the substrate into a novel des-methyl thiotetromycin product (287). C) Analogous experiments with the methylmalonyl tetraketide (207) lead to detection of thiolactomycin (110). Clear MS^2 fragments could not be detected to confirm the structures but assuming they are correct, the oxidation/amidation enzyme machinery does not appear to process these products further.**

2.4.4 Synthesis of additional substrates for mechanistic investigations of thiotetronate ring formation

A synthetic route was also planned to access two thiolactomycin analogues (**295** and **296, Scheme 2.28**) with a dual purpose. Due to the lack of characteristic MS^2 fragments for the putative thiotetronates found in microbial extracts, it was hoped that synthetic compounds could be used as standards for them. Also, modification of the thiotetronate ring methyl substituent of thiolactomycin has received limited attention from synthetic

chemists. Groups that have been placed in this position instead of the methyl moiety have included various other aliphatic alkyl groups such as ethyl, butyl and even benzyl. These constitute minor changes in size and electronic properties. Exploration of modifications at this ring position may lead to compounds of improved potency or scope. The preparation of **295** and **296** was planned around an already validated route involving a tandem thioester hydrolysis and Dieckmann-like cyclisation of **290** to form a thiotetronate ring before introducing the carbon side chain by alkylation (**Scheme 2.28**).[107]
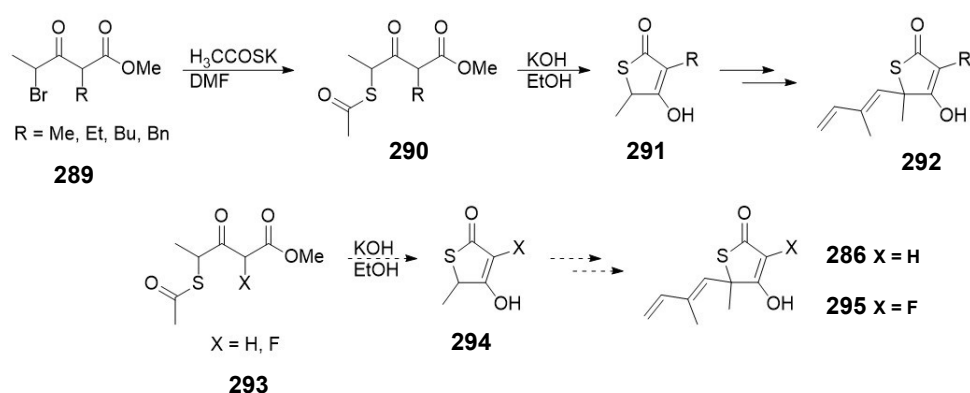
Scheme 2.28 – **(Top) The most widely reported routes for thiotetronates start from bromo keto ester substrates with alkyl groups (R) already installed as malonyl substituents (289). These compounds are typically alkylated at the α-carbonyl position bearing the halogen with thioacetate to give thioesters (290) which are then cyclised in basic conditions to give the ring formed products (291) before further side-chain addition (292). (Bottom) Similar thioester substrates with either no substitution or a halogen (293) may also be able to cyclise to give novel thiotetronate rings (294), which can be similarly processed towards new thiotetronates (295 and 286).**

Synthesis (**Scheme 2.28**) of the non-methylated thiotetronate ring (**286**) began with bromination of a keto methyl ester (**296**). A trace amount of bis brominated side-product was observed by NMR which corresponded to the brominated product (**297**) undergoing a second bromine addition to the more acidic malonyl position. Separation by chromatography was unsuccessful and the purity of the desired product was estimated by proton NMR as approximately ~81% w/w (giving an approximate yield of 91%). Alkylation of the α- carbonyl position bearing the bromide with potassium thioacetate gave full conversion to the thioester product (**298**) by TLC. Given the number of similar

impuriies and stereoisomers, purification of deemed unfeasible and so the crude thioester was used in the cyclisation reaction.
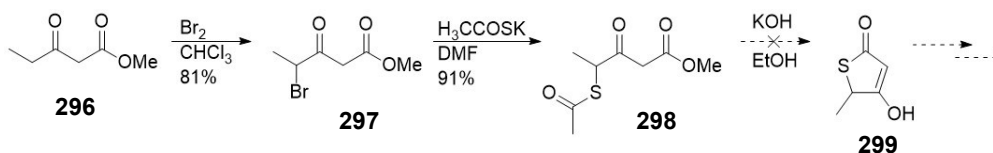
Scheme 2.29 – **Attempted synthetic route to des-methyl thiotetronate (286)**

The synthesis of a thiotetronate molecule bearing a fluorine substituent in place of the methyl malonate group was also devised (**Scheme 2.30**).
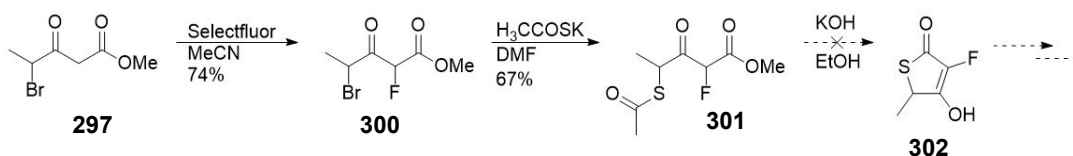
Scheme 2.30 – **Attempted synthetic route to a fluoro-thiotetronate substrate (295)**

The previously prepared bromide **300** was treated with Selectfluor to afford the desired compound (**301**) in 74% yield. Alkylation with potassium thioacetate gave the fluorinated thioester molecule (**301**), the NMR spectra was complex even after chromatography but the yield was estimated at 67%. Issues in compound purification from trace amounts of side products were again experienced, therefore product **301** was tested for cyclisation in basic conditions without further purification. Several trial cyclisation reactions were attempted to isolate **302** but were unsuccessful as they showed little evidence of product formation. The use of previously reported KOH catalysed cyclisation conditions[108] were not successful for either substrate (**298** and **301**). Proton NMR analysis suggested possibly a trace of product from the cyclisation of **298** (**Table 2.5**), but a clean sample for

further characterisation could not be obtained. The lack of any similar reported compounds for NMR shift comparison also prevented any effort of characterisation.
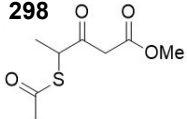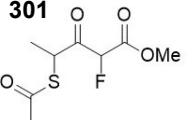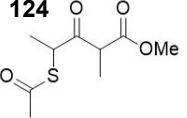
	KOH/EtOH (Lit. conditions)	K₂CO₃/MeOH (Our conditions)
298 (structure)	Possibly trace product based on COSY NMR data, not confirmed or isolated pure	No evidence of any cyclised product
301 (structure)	No evidence of any cyclised product	No evidence of any cyclised product
124 (structure)	Clear product formation (previously reported)	Not tested

Table 2.5 – **Summary of results for attempted cyclisation of substrates 298, 301 and 124.**

A literature search of these types of ring closures shows the need for a carbon-based substituent in the malonate position which may explain why the outcome of our reactions was poor. Use of milder conditions such as K₂CO₃ and MeOH also resulted in no cyclised product. The previously reported reaction with the methyl malonate substrate (**124**) were repeated by us to validate the reported conditions; for this substrate, the corresponding cyclised product **125** was clearly observed by proton NMR and also isolated, indicating that the base-catalysed reaction is very substrate dependant. A possible explanation for our results may be due to the decreased electrophilicity of the methyl ester carbon towards attack by the thiolate. Both the unsubstituted (**298**) and fluorine substituted (**301**) analogues likely have increased enol character due to the more acidic malonate position, even this was not clearly observed by proton NMR. Increased tautomerisation would indeed result in decreased electrophilicity of the ester carbonyl which in turn would impair the Dieckmann cyclisation and allow ester hydrolysis to occur instead (**Scheme 2.31**).
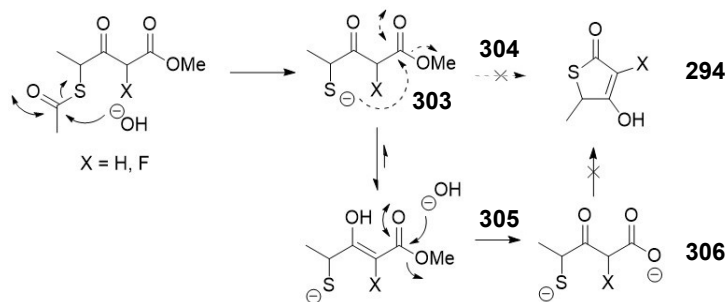
Scheme 2.31 – **Possible side reaction in cyclisation with non-carbon substituted analogues, lower electrophilicity of methyl ester carbonyl in the substrate (303) could slow Dieckmann cyclisation (304) allowing hydrolysis of methyl ester to happen instead (305), this prevents product (294) formation and leads to a side product (306) which probably further reacts un unknown ways.**

There may be also a lesson to learn from these cyclisation results when compared to the *in vivo* feeding experiments with tetraketide SNAc substrates (**Section 2.4.2**). The success of the substrates cyclising (in order of most successful being **124 > 298 > 301**) is consistent with the ability of the corresponding tetraketide SNAc analogue (**207 > 216 > 262**) to be processed by *Lentzea* sp. Δ*tlmA* to generate thiotetronates. It could be speculated that this provides evidence for the way in which the thiolactomycin ring is closed (**Scheme 2.32**).
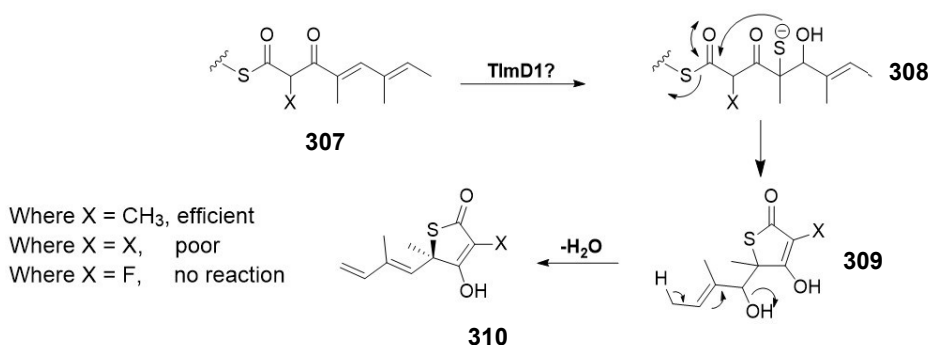
Scheme 2.32 – **Possible alternative mechanism of thiolactone formation which would rationalise the *in vivo* data gathered for SNAc processing in *Lentzea* sp. Δ*tlmA* and synthetic substrate cyclisation. Similar to the base-catalysed cyclisation previously discussed, a thiolate (308), deriving from PCP-bound cysteine through a yet unknown mechanism from the tetraketide (307), would perform a nucleophilic attack on the thioester of ACP-bound tetraketide. This ring closure would give a hydrated thiotetronate (309) species which can be subjected to H_2O elimination to give thiolactomycin (310). In this proposed mechanism, the substituent in the malonate position is important in influencing the fate of intramolecular cyclisation.**

The scenario of **Scheme 2.33** would also explain the nature of some of the intermediates containing S by us intercepted and also the latest findings by Tang *et al.*[138]. In fact, the 6-membered ring shunt product (**223**) they observed may result from conjugate addition of sulphur to the ACP-bound tetraketide in the absence of TlmD1.
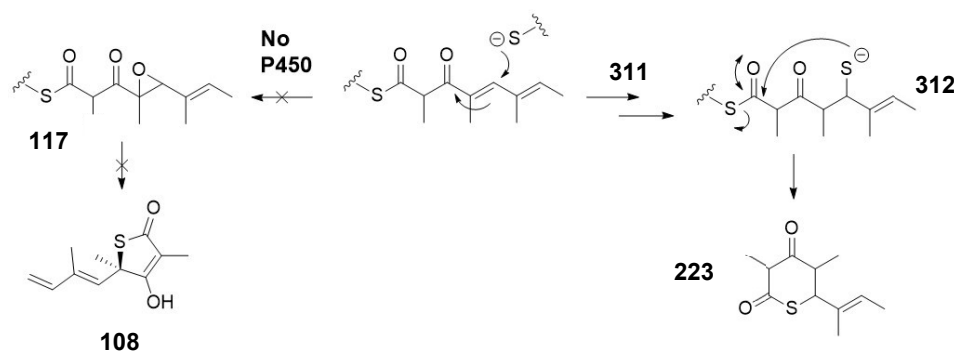
Scheme 2.33 – **Possible alternative mechanism for a 6-membered ring shunt product formation. In the absence of TlmD1, thiol/thiolate conjugate addition (311) may occur and deliver the sulphur atom of intermediate (312) primed to form a 6-membered ring (223).**

This hypothesis remains speculative at this time, in chapter 5, some ideas of future work will be suggested and discussed in view of probing different mechanisms and timing of events associated to sulphur insertion and cyclisation.

2.5   Synthesis of co-crystallisation substrate for TlmB

As part of a collaboration with Prof. Marco Diaz (Sao Paulo, Brazil), we also undertook the preparation of a substrate (**315**) for crystallisation studies of TlmB and other domains deriving from this intriguing PKS/NRPS. The substrate is an amide analogue of the natural tetraketide intermediate ultimately converted to TLM; it is non-hydrolysable therefore it should be stable in a variety of conditions utilized for crystallisation trials.
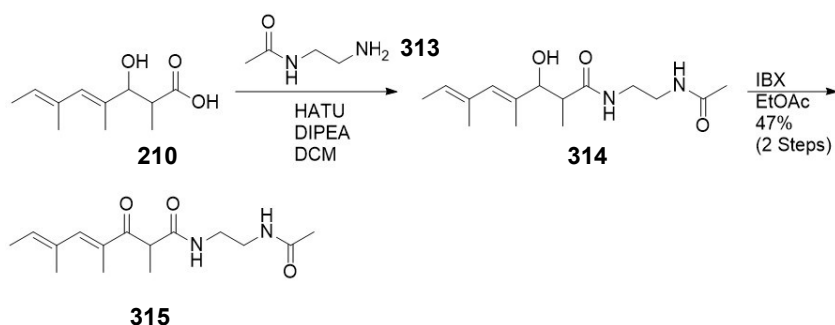
Scheme 2.34 – **Synthetic route to amide tetraketide SNAc thioester (315)**

The synthesis of **315** began with the previously prepared hydroxy acid (**210**), which was amide coupled to *N*-acetyl ethylene diamine (**313**) giving the β-hydroxy amide (**314**). This was partially purified by silica chromatography and used in the IBX oxidation to afford the substrate (**315**) with a 47% yield over two steps.

Substrate **315** is currently undergoing crystallisation trials set up by the group of Prof. Marcio Dias.

2.6 Summary of results

As detailed earlier in the chapter (**Section 2.1**) and already published,[134] two previously synthesised probes (**104** and **160**) and a novel methylmalonyl substituted probe **161** were successfully used *in vivo* on *Lentzea* sp. strains to capture putative intermediates of the thiolactomycin pathway. This experimentally corroborated the proposed iterative type I PKS responsible for producing thiolactomycin. These captured species not only varied from diketide to tetraketide length, but also in the degree of processing where intermediates were apparently offloaded but also further reduction/dehydration had taken place; no definitive evidence for capture of the proposed late stage thiolactomycin intermediate tetraketide (**116**) was gathered though.

In general, the probes were well tolerated by all strains tested up to a concentration of 2 mM. Based on the proportion of intact probe compared to its deprotected and decarboxylated by-product observed by LC-HRMS, the fluoromalonyl (**160**) probe has so far proved the most efficiently processed by cells. This is also consistent with the findings delivered by this substrate compared to others utilised in similar experiments. Investigation of other thiotetronate producing organisms such as *S. thiolactonus* lead to capture of analogous PKS species expected from the biosynthesis of those products (thiotetromycin and Tü3010). Some putative sulphur containing species of both tetraketide and pentaketide length were also detected *in vivo*. The presence of this species do not necessarily support the previously envisaged mechanism of thiolactone formation *via* a P450-mediated epoxide formation.[93] Also, the supplementation of chemical probes to mutant strains of *Lentzea* sp. lacking genes/domains such as *tlmD1* or the *tlmB_Cy* shed no further light on their function/ mechanism.

In terms of *in vitro* assays, similar probe-captured thiolactomycin intermediates were detected from incubation of the iterative PKS-NRPS TlmB with its required cofactors/substrates as was seen *in vivo*. The previously detected sulphur containing species were present in some *in vitro* assays, however the detection occurred in the absence of TlmD1, TlmS and TlmJ which were initially thought to be essential for sulphur insertion in the polyketide carbon backbone.

A synthetic SNAC thioester substrate (**192**) was prepared as a mimic for the native ACP-bound triketide and was successfully used to prime TlmB, indicating that non-native hydrolysable substrates can be used to investigate TlmB. Other SNAc mimics of ACP-bound intermediates were then prepared for mechanistic investigations (**201**, **199** and **207**). A non-thiolactomycin producing mutant strain of *Lentzea* sp. lacking the loading PKS module TlmA (*Lentzea* sp. Δ*tlmA*) was supplemented *in vivo* with each of the SNAc substrates (**199**, **192**, **201** and **207**) and for the latter three substrates, the production of thiolactomycin was detected, clearly demonstrating the involvement of these polyketide species as natural intermediates in thiolactomycin biosynthesis. Surprisingly, no restored production of TLM was observed for **199** supplementation extracts. *In vitro* incubation of the SNAc triketide and tetraketide (**192** and **207**) with the putative monooxygenase

TlmD1 initially appeared to mediate an epoxidation reaction given the mass increase of the substrate of 16 Da along with expected MS^2 fragments, however, detection of the same species in boiled enzyme extracts discounted the possibility that this species is effectively formed. Analytical HPLC analysis of the same experiments revealed no new peaks formed from the oxidation of the SNAc polyketide intermediate mimics.

'Unnatural' PKS intermediates of thiolactomycin biosynthesis were also synthesised as SNAc esters and tested with the *Lentzea* sp. Δ*tlmA* mutant in order to probe whether 'unnatural' thiotetronate analogues could be produced, however, only one substrate (**216**) appeared to give rise to a novel product (**286**) for which no diagnostic MS^2 fragments could be acquired. The unnatural SNAc substrates were also tested on a heterologous host expressing the Tü3010 (**110**) producing cluster with an inactivated loading module. Low levels of Tü3010 were still observed in extracts from the mutant strain itself, therefore comparative analysis looking at restored production of Tü3010 and other metabolites proved challenging. Supplementation of **192** did give rise to a new peak with m/z 254.1, indicating that an unnatural thiotetronate (**286**) may have been produced but could not be confirmed by MS^2. Interestingly though, feeding of the same substrate to the wildtype Tü3010 producer *S. olivaceus* did lead to diminished production of indicating some inhibition of the PKS-NRPS by this substrate. Supplementing of the two substituted tetraketide substrates **207** and **263** also appeared to give rise to thiolactomycin (**108**) and des-methyl thiotetromycin (**287**) respectively but no diagnostic MS^2 fragments were found.

Based on a recent report that the P450 in thiotetronate producing gene clusters is responsible for the 5-membered ring formation as opposed to sulphur insertion,[138] a terminal trifluorinated SNAc (**226**) was prepared to shed light on the later steps of thiotetronate biosynthesis. Supplementation of the substrate in *Lentzea* sp. led to detection in the organic extract of two species which could support the newer hypothesis of the P450 forming an epoxide to allow a five membered ring to form, however no further characterisation of these species could be gathered to confirm their structure.

I also synthesised the proven late stage tetraketide intermediate as a pantetheine thioester (**245**) in order to load it onto the ACP domain of TlmB *in vitro*. This was successful and

the loaded *crypto*-ACP was detected by LCMS; treatment of this with various purified recombinant proteins was performed in order to investigate the later biosynthetic steps, but unfortunately no novel peaks corresponding to processing of the ACP-bound tetraketide were detected.

Finally, synthetic routes to generate unnatural thiotetronate compounds synthetically were also devised and attempted. A trend in the success of the thiolactone ring forming step was observed in relation to the malonyl group substituent, which may possibly be linked to the production of unnatural thiotetronate analogues using SNAc substrates *in vivo*. Future work related to the investigation of sulphur insertion and to intermediate cyclisation leading to TLM will be discussed in **Chapter 5**.

3 Chemical probing of ring formation in tetronate antibiotic biosynthesis

As outlined in the introduction, tetronate antibiotics are of great interest due to their ranging activities; however, several biosynthetic steps leading to their formation remain unclear. An example of interesting and intriguing pathways includes those leading to tetronasin and tetronomycin (**145** and **146**, **Fig. 3.1**). Our starting point of investigation focused on the formation of the tetrahydrofuran (THF) ring. Previous studies on the biosynthetic construction of other polyethers such as lasalocid A (**103**) revealed that THF ring formation was likely occurring while the polyketide chain was still enzyme bound [83] (**Fig. 1.15**). Therefore, we wanted to establish whether this step was occurring in a similar manner on the tetronasin modular PKS assembly line; we were also interested in understanding the timing of other ring formation, including the tetronate rings.
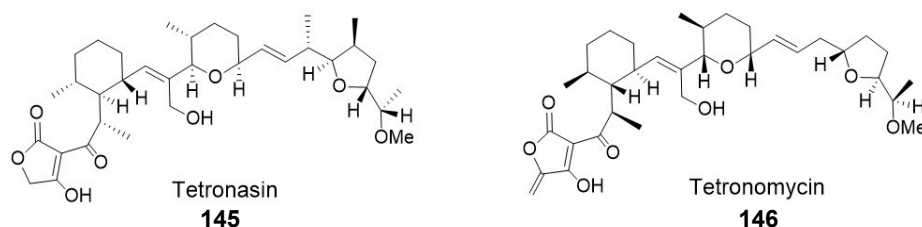
Figure 3.1 – **Structures of tetronasin (145) and tetronomycin (146).**

The timing and mechanism of formation for the cyclohexane and pyran ring has been the subject of a collaboration with Rory Little, a PhD student in the Leadlay group at the University of Cambridge, and Prof Marcio Dias at the University of Sao Paulo (now at Warwick). Some of the work carried out by the collaborators will be mentioned throughout this chapter in relation to fully elucidating the biosynthesis of tetronasin and tetronomycin, our collective work has been recently published in Nature Catalysis.[141]

In the following section, I will present my work that focused on understanding the timing and the mechanisms of formation of the THF and tetronate ring moieties in tetronasin assembly.

3.1.1 THF ring formation mechanism and timing in tetronasin biosynthesis

Based on the presence of a putative epoxidase and an epoxide hydrolase gene within the tetronasin BGC (*tsnB*, *tsnC*, **Fig. 1.23**), it is thought that the THF ring was likely to be the result of epoxidation of an alkene in the polyketide backbone followed by enzyme-catalysed ring opening and cyclisation, this type of transformation has been well reported. Below, polyether ring formation for the antibiotic monensin is illustrated; this pathways has been extensively studied and has constituted the starting point for the investigation of other polyether assemblies (**Scheme 3.1**),[142,143] they are also interesting in that many of these transformations feature Baldwin disfavoured ring closures.[144]
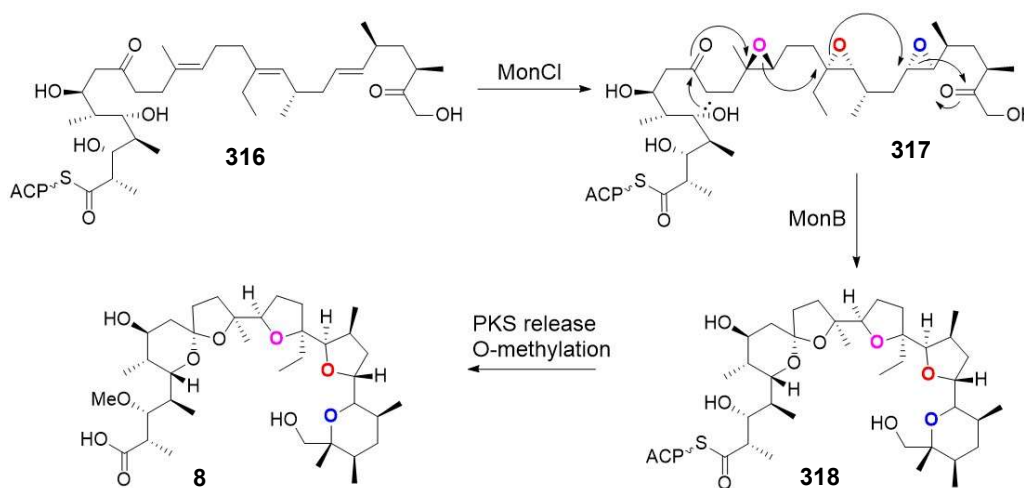
Scheme 3.1 – **Proposed mechanism of THF and pyran ring formation in monensin A (8). The PKS-bound backbone (316) undergoes conversion to a triple epoxide (317) by the epoxidase MonCl; the triple epoxide leads to the formation of two THF and two pyran rings in a cascade mechanism mediated by the epoxide hydrolase MonB. The last intermediate (318) is then released from the PKS and undergoes *O*-methylation.**

In tetronasin biosynthesis, the timing of the THF ring formation based on this type of mechanism could theoretically occur at any point after tetraketide formation on module four: indeed, at this stage both the terminal alkene and the hydroxyl group moieties required for epoxidation and cyclisation should be present (**Scheme 3.2**). The alcohol *O*-methylation could also occur at any point after cyclisation.
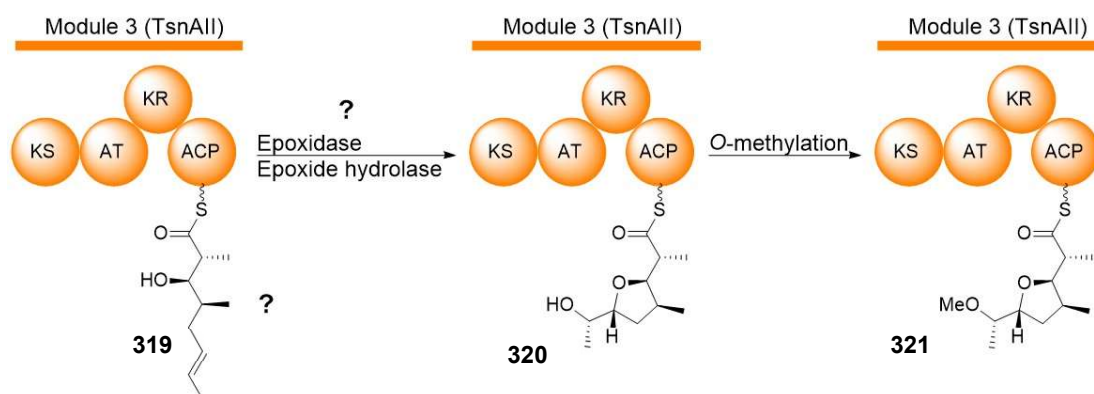
Scheme 3.2 – **Earliest possible THF formation stage based on epoxide/ring open mechanism: the ACP-bound tetraketide (319) can be epoxidised allowing ring opening with the secondary alcohol to form product (320); *O*-methylation can occur at any time after this stage (321).**

In order to shed light into the timing of THF ring formation, we planned to use the wildtype *S. longisporoflavus* strain and two mutants Δ*tsn11* and Δ*tsn15* (provided by Rory Little, Leadlay group, Cambridge) to perform feeding experiments with our chain termination probes. The mutant strains each have a single gene knockout for the *tsn11* and *tsn15* genes within the tetronasin BGC which Rory Little had shown to be essential for tetronasin biosynthesis;[141] *tsn11* shares homology with several proposed Diels-Alderase enzymes from various natural products including *ChlE3* from the chlorothricin (**155**) BGC. Moreover, *tsn15* has close homology for known Diels-Alder like cyclase enzymes present in numerous spirotetronate natural products. Both of these genes have homologues in the tetronomycin BGC which strongly suggests both products use an identical mechanism of ring formation. It was also hoped that experiments with both the aforementioned strains could also help to shed light on the timing and mechanisms of cyclohexane and pyran ring closure.

3.1.1.1 Use of chain terminator probes on *S. longisporoflavus* in solid culture

The wildtype and two mutant strains of *S. longisporoflavus* were grown on tsn-medium-B plates as previously reported (recipe given in **Experimental 6.2.5**). Good tetronasin production was detected by LC-HRMS in the wildtype plate EtOAc extracts. These conditions were used for all feeding experiments with the strains on plate; experiments with liquid cultures were run utilising the same media. In the first instance, 1-2 mM concentrations of chemical probe were used, with only the fluoromalonyl (**160**) and decanamido (**104**) probes utilised. The phenotype of the wildtype, Δ*tsn11* and Δ*tsn15* strains on plate appeared fairly consistent in terms of plate colour and mycelium growth. However, the specific probe used and its concentration had a significant impact: moderate toxicity was seen above 1 mM concentration with less mycelium across each strain; interestingly the fluoromalonyl probe (**160**) causes a more-rounded spore like mycelium, whereas control plates and decanamido probe (**104**) supplemented plates had a more filament-like growth. All plates fed with probes were noticeable lighter in colour than the respective control plate and the wildtype plates tended to be darker than their Δ*tsn11* and Δ*tsn15* counterparts (**Fig. 3.2**).
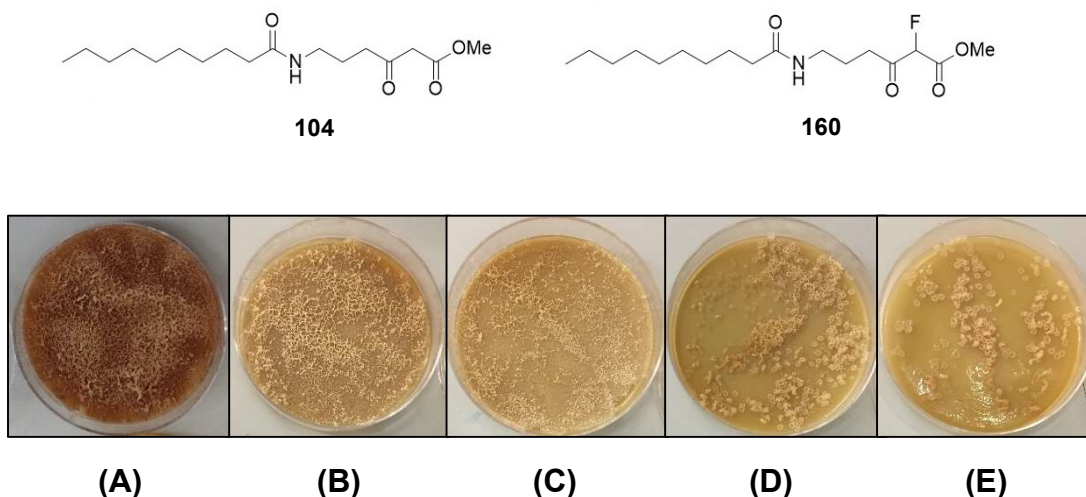
Figure 3.2 – **Appearance of feeding *S. longisporoflavus* wildtype with probes and control plate: (A) control plate (no probe present), (B) 1 mM decanamido probe, (C) 2 mM decanamido, (D) 1 mM fluoromalonyl probe and (E) 2 mM fluoromalonyl probe.**

LC-HRMS analysis of the wildtype control extracts showed a large abundance of ions corresponding to the tetronasin sodium adduct, which was present in all plates with probes at different concentrations. However, tetronasin production appeared reduced in plates supplemented with the decanamido probe (**104**) at 1 mM concentration with further reduction at 2 mM (**Fig. 3.3**). As for treatment with the fluoromalonyl probe (**160**), the effect was even more pronounced; in both cases this is likely due to bacterial growth inhibition due to the presence of the probes. As previously observed, the abundance of the hydrolysed and decarboxylated derivatives of the fluoromalonyl probe was significantly higher than that of the decanamido probe.
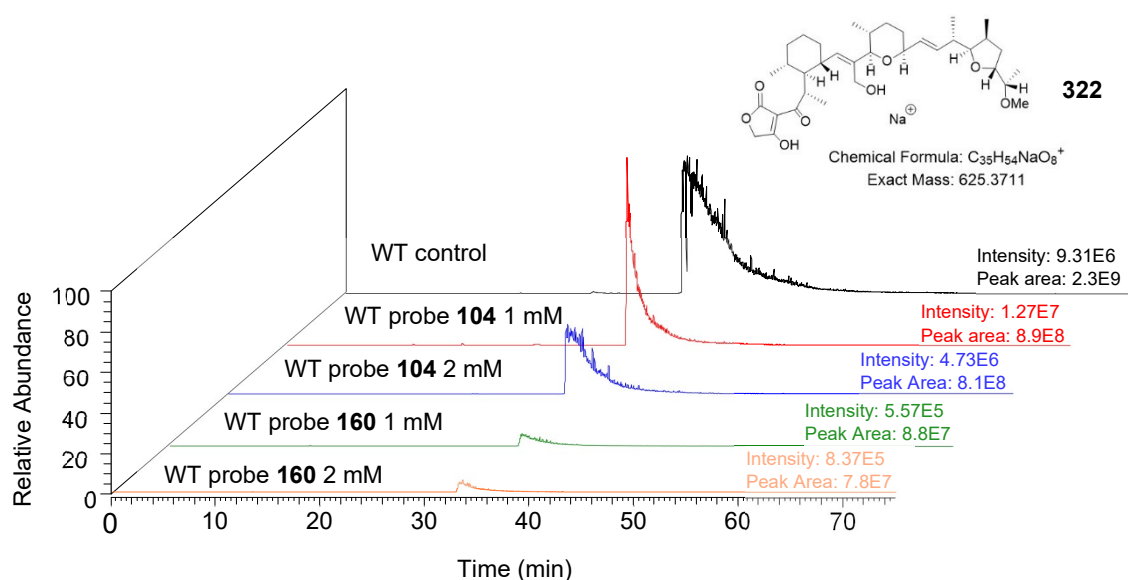
Figure 3.3 – **Extracted ion chromatogram for tetronasin [M+Na]+ (322) adduct detected in** *S. longisporoflavus* **WT plate extracts varying probes and concentration: a large decrease is seen using probe 104 and even greater with probe 160.**

Interestingly, a peak with the exact same mass of tetronasin was detected in both the Δ*tsn11* and Δ*tsn15* mutants; the retention time was fractionally different to tetronasin, the MS^2 profile had minor differences to that of tetronasin (**Fig. 3.4**). As shown here and in previously reported literature,[145] tetronasin does not fragment easily and the backbone remains intact with just dehydration masses observed; the MS^2 profile of this same mass peak observed in the mutant extracts have far greater abundance of the dehydrative fragments than observed in the tetronasin MS^2 profile from the wildtype extracts. From

this data, it is difficult to assess whether a small quantity of tetronasin is still produced by the mutants or whether the peak of m/z 625.3711 corresponds to a novel/ different species. Since both *tsn11* and *tsn15* are required for production of tetronasin, these peaks could both be intermediates or shunt products of tetronasin which have the same mass: this would suggest that no atoms are added or removed in the reactions that are mediated by Tsn11 and Tsn15, possibly supporting the idea of a cycloaddition/Diels-Alder reaction.
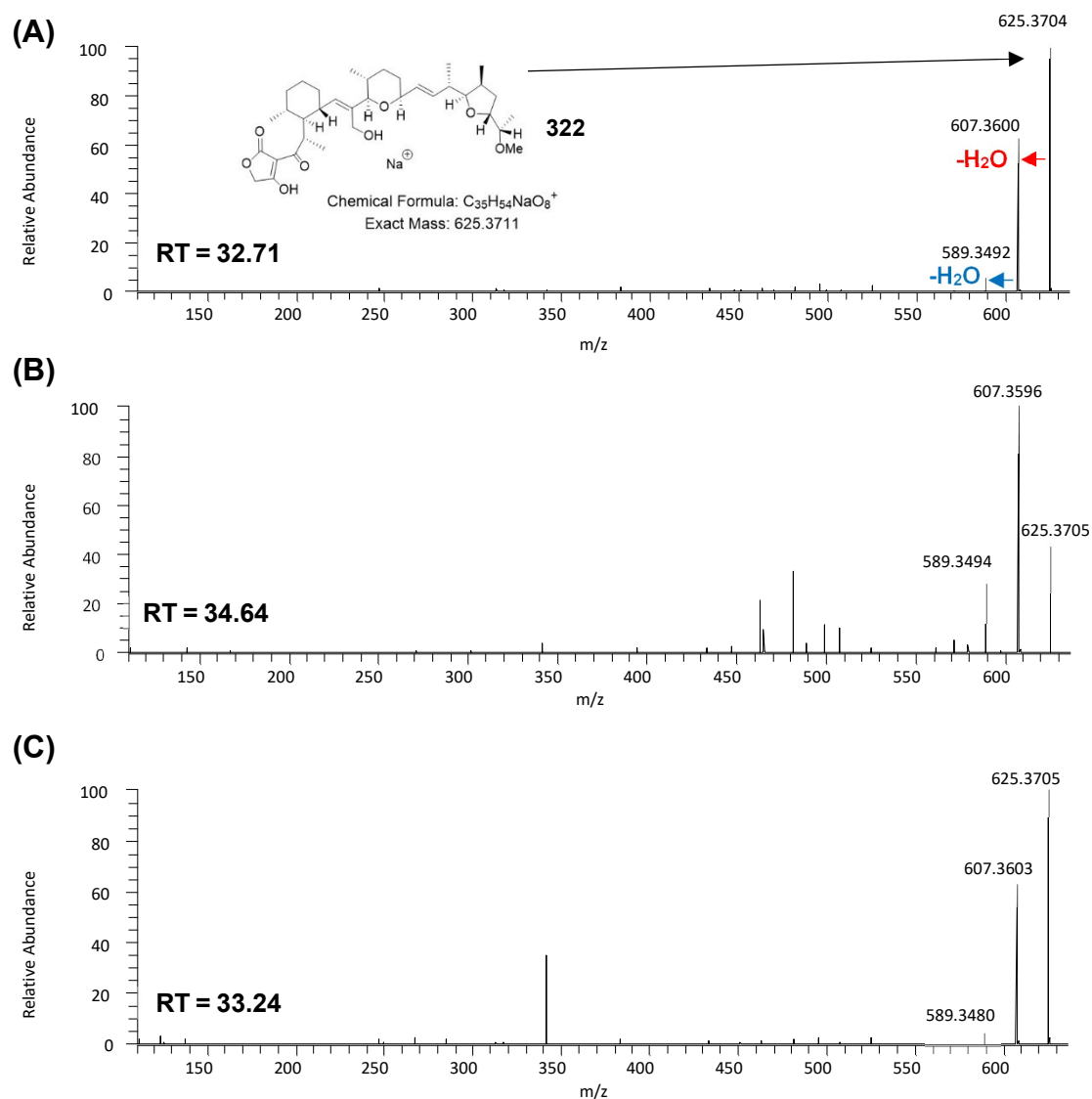
Figure 3.4 – **MS^2 spectrum for mass matching the tetronasin sodium adduct (322) in *S. longisporoflavus* wildtype (A), Δ*tsn11* (B) and Δ*tsn15* (C) control extracts along with the parent ion retention times; the profiles of each are different and the abundances of dehydration events are varied, indicating three products with the same mass.**

No further investigation of these extracts was carried out to determine the nature of these putative shunt products by myself as further studies were carried on by Rory Little (Cambridge).

Instead, we focused on the detection and analysis of putative intercepted polyketide intermediates in culture extracts. The deletion of the genes *tsn11* and *tsn15* in *S. longisporoflavus* should not greatly disrupt the polyketide backbone construction. However, despite the large number of intermediates that potentially could be intercepted from the twelve-module PKS and the abundance of tetronasin production, few intermediate parent ions were detected with even fewer which could be confirmed by MS^2 fragmentation. The fluoromalonyl probe (**160**) intercepted every intermediate that was seen in the decanamido probe (**104**) extracts in an equivalent or higher abundance along with several different ones. These were still restricted to simple di and triketide length intermediates (not shown here) which could be confirmed by MS^2 but could just as easily be products of other metabolic pathways; this lack of evidence for intercepted polyketides was consistent in all three strains tested with both probe concentrations under the same conditions. It became apparent early on that the tetronasin PKS must be a very efficient assembly line and difficult to interfere with; nonetheless we pursued further investigation changing the fermentation conditions (**Section 3.1.1.2**).

3.1.1.2   Use of chain terminator probes on *S. longisporoflavus* in liquid culture

Given the little success in offloading intermediates through fermentation on solid media (partially due to probe toxicity issues), liquid cultures of *S. longisporoflavus* administering the probe in smaller quantities over a longer period of time were set up. Feeding experiments previously conducted by the group on the lasalocid A producer *S. lasaliensis* in this way proved successful in by-passing probe toxicity issues. Due to the better solubility of the fluoromalonyl probe (**160**) in MeOH (essential for supplementation) and its superior offloading capability, this probe was initially added all at once in 1 mM overall concentration. Eventually, the probe was added as a MeOH solution in aliquots over four days after the first day of fermentation to achieve the final

concentration of 1 mM. The production of tetronasin in the wildtype liquid feedings was comparable to that seen on plates, comparable probe methyl ester hydrolysis and decarboxylation levels were also observed. Like in the solid culture experiments, several putative intercepted diketide and triketide species were found but rewardingly, more advanced species were also observed. Two particular putative intercepted species were detected in all three strains in good abundance. The first correlated to a possible intercepted heptaketide species (**323, Fig. 3.5**); the other was consistent with a heptaketide being subjected to epoxidation and ring opening to form a THF ring and backbone dehydration (**324, Fig. 3.5**). The retention times of both these intermediates correlated well with structures of other polyketide intermediate species previously characterised by the group from other assembly lines.
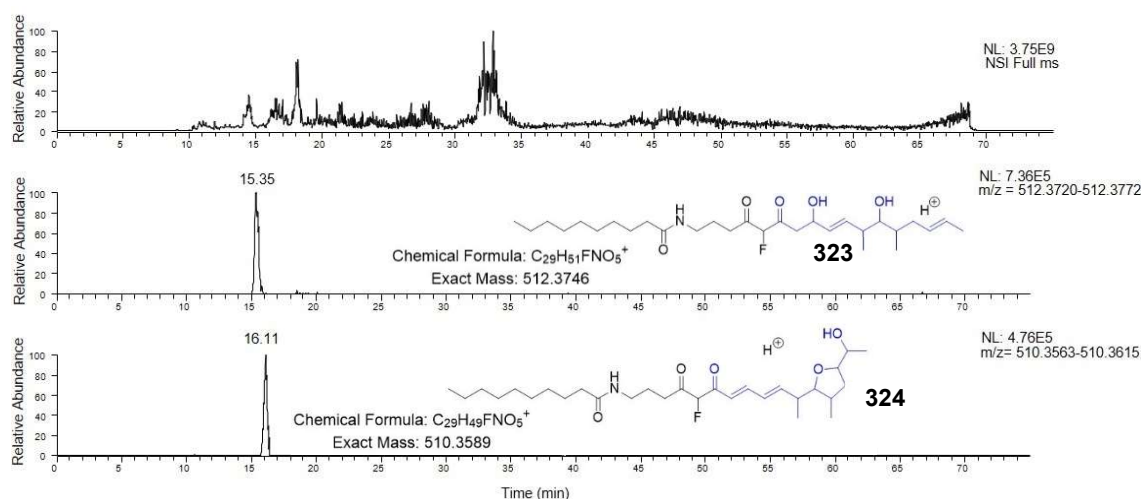
Figure 3.5 – **LC-HRMS chromatogram showing total ion current (top) and the detection of two putative ions for an offloaded linear heptaketide (323) (middle) and a ring closed intermediate after dehydration (324) (bottom).**

Unfortunately, MS^2 fragmentation of these species was not very informative. Similar to the fragmentation of tetronasin itself, several water losses from these molecules were observed and, more interestingly, loss of HF which clearly indicated that these ions are related to the probe in some form. This pattern was observed for both the putative linear heptaketide (**Fig. 3.6 upper**) and also the oxidised and cyclised analogue (**Fig. 3.6,**

lower). Whilst no diagnostic fragments resulting from carbon-carbon bond cleavage were characterised, many low abundance masses related to the loss of a decyl carbon chain were observed: these are important because a diagnostic fragment of the probe is obtained by amide bond cleavage (leaving an ammonium species diagnostic of the captured species and a decanoyl fragment with a mass/charge ratio of 155 Da).
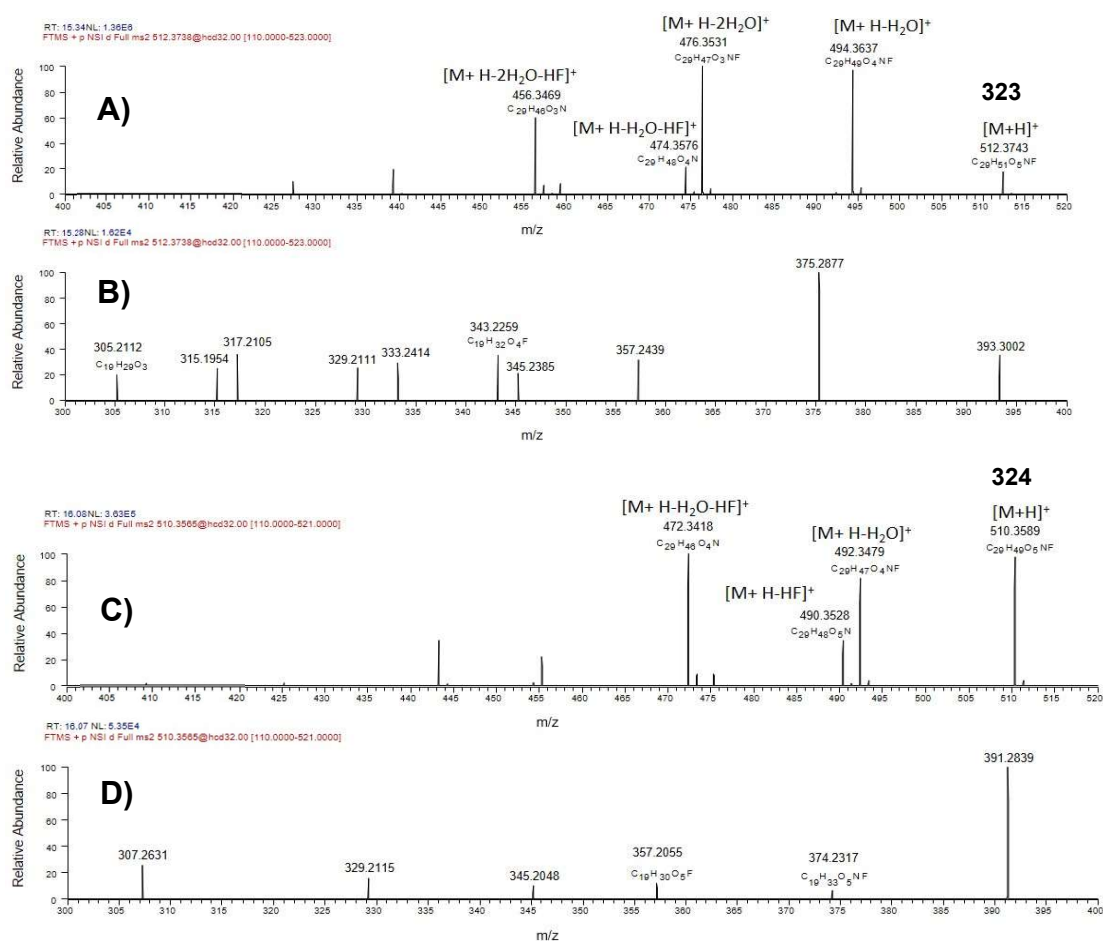
Figure 3.6 – **LC-HRMS^2 analysis of putative linear offloaded heptaketide (323) (upper) and cyclised analogue (324) (lower). Several combinations of HF and H$_2$O loss can be seen for both putative intermediates (A and C). Few intermediates are detected indicating fragmentation through the molecule centre (B and D) but several masses match the loss of a decanoyl chain fragment through amide breakage as is usually seen in amide containing chemical probes.**

Several attempts were made to further characterise the putative intermediate structures. Extracts were re-run with a higher collision energy value to increase the abundance of the

300-400 mass MS^2 fragments but this was unsuccessful. Many of the MS^2 values detected from loss of H_2O and HF were also placed in a list for targeted MS^3 analysis. Unfortunately, this too gave no new fragments for characterisation. It had been observed previously that lasalocid A intercepted intermediates were quite resistant towards fragmentation so these similar structures might expectedly behave the same but at this stage, the structure of these ions cannot be definitively confirmed. If these were confirmed, it indicates that the THF ring is being formed by an epoxidation and cyclisation cascade (**Scheme 3.3**) occurring on the heptaketide (**325**) whilst still bound to the polyketide synthase. It is unclear whether the module five KR domain would have acted on the intermediate before or after the epoxidation and ring cyclisation.
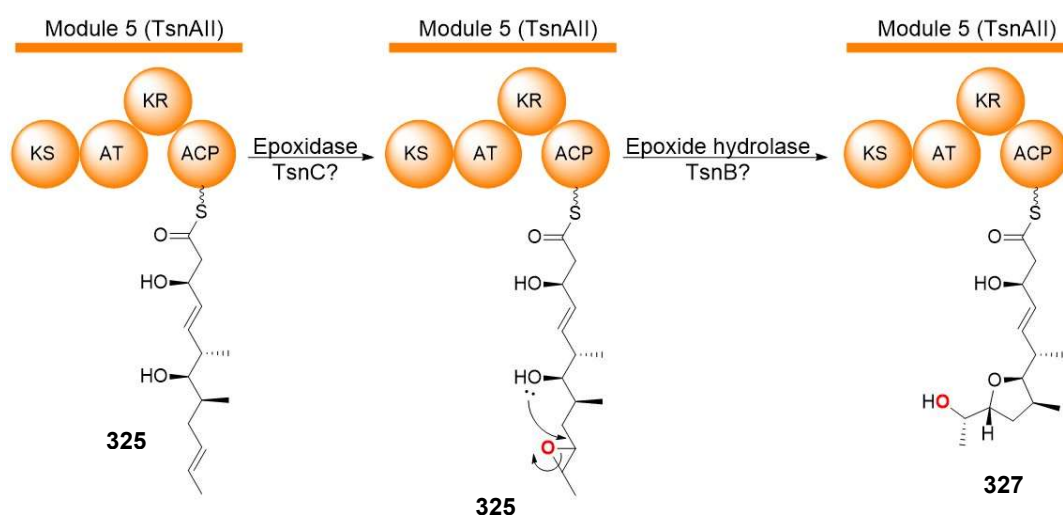
Scheme 3.3 – **Proposed THF formation timing and mechanism based on LC-HRMS analysis: the linear heptaketide (325) is converted to an epoxide (326) via an epoxidase (TsnC) which is then ring opened by an epoxide hydrolase (TsnB) to give the THF intermediate (327).**

As well as these species, other putative intermediates from downstream modules were detected but once more the MS^2 characterisation proved difficult. The most interesting amongst them was the putative dehydrated undecaketide (**328**), featuring a hydroxyl group at the branch position of C-15 and *O*-methylation at position C-2 beside the THF ring (**Fig. 3.7, A**). This species was detected in extracts from all *S. longisporoflavus* strains (WT, Δ*tsn11,* Δ*tsn15*). Initial MS^2 fragmentation of it did not clearly support the

132

proposed structure. Additional feeding experiments in liquid culture were then setup with a 2 mM final probe concentration in an attempt to increase the abundance of putative intermediates and MS² fragments. This time, putative undecaketide **328** was more abundantly detected and clearer MS² fragments were gathered that were consistent with proposed structure (**Fig. 3.7, B&C**).
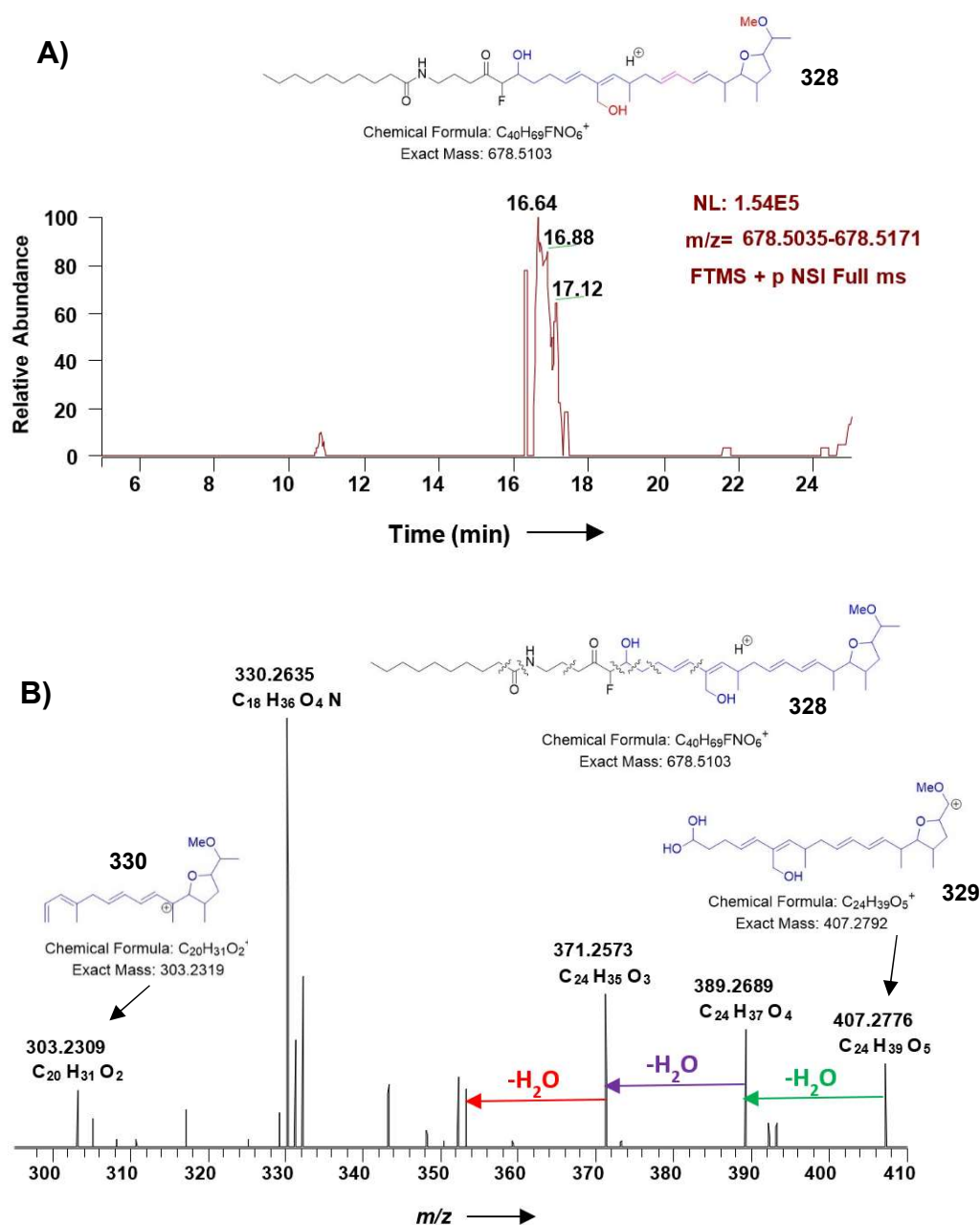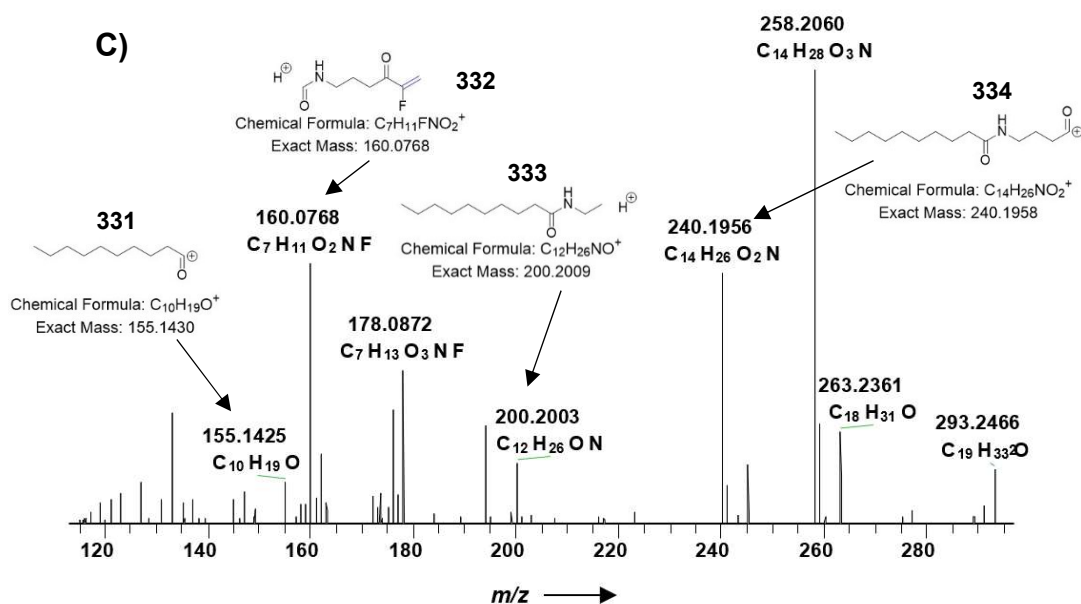
Figure 3.7 – **(A) Putative captured undecaketide (328) having undergone hydroxylation, and *O*-methylation (red) and dehydration (pink) detected in experiment extract, (B) two fragments (329-330) consistent with amide bond cleavage and dehydration were detected, (C) numerous probe related fragments (331-334) were also detected.**

In particular, fragment **329** (m/z 407) would result from a McLafferty rearrangement (**Scheme 3.4**) of species **328** leading to an aldehyde which could then undergo hydration and demethylation.
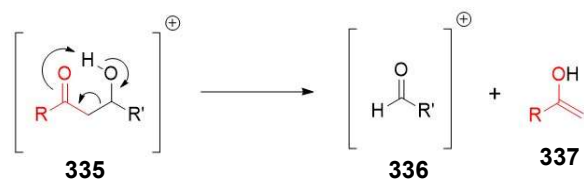
Scheme 3.4 – **Example of a McLafferty rearrangement: A ionized hydroxy-ketone (335) undergoes the rearrangement to afford a neutral enol fragment (337) and an ionized aldehyde fragment (336) which then undergo hydration.**

Based on these data, we speculate that P450-mediated hydroxylation followed by *O*-methylation could occur during polyketide chain formation between the octaketide and

undecaketide stage, however we cannot exclude that the transformations are occurring after intermediate capture. Combining all the information obtained by our experiments, a biosynthetic sequence including ring formation timing was proposed for tetronasin formation (**Scheme 3.5**).
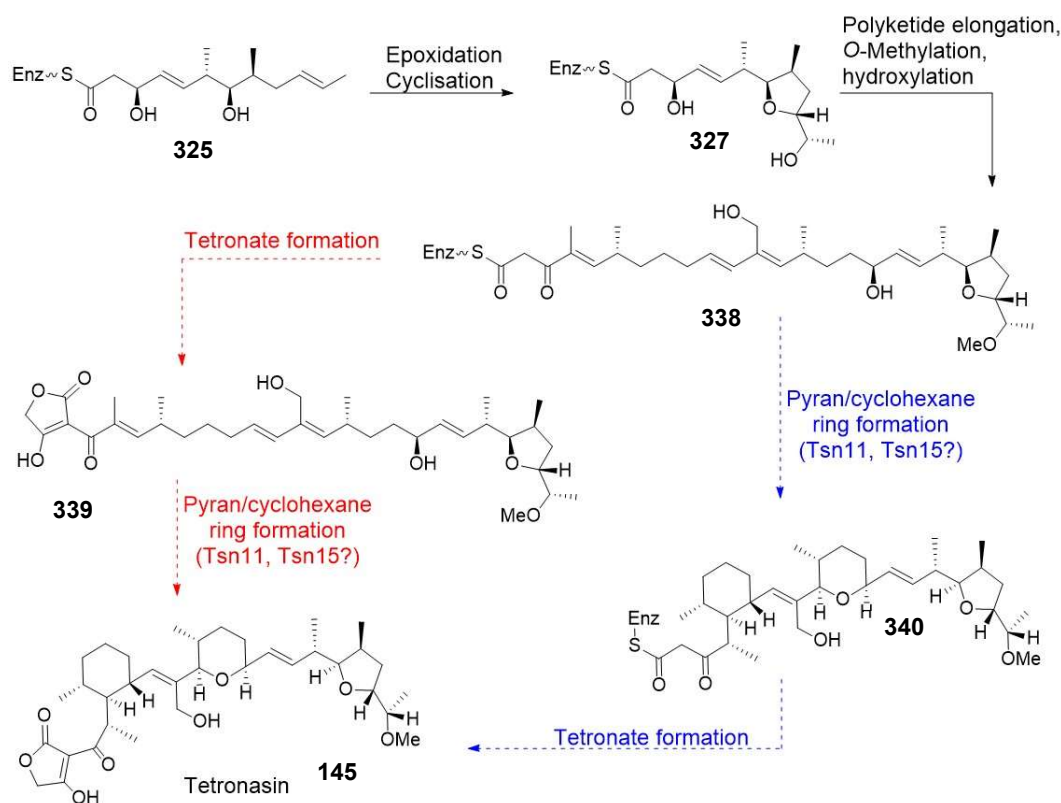
Scheme 3.5 – **Proposed mechanism of tetronasin ring formation (145) biosynthesis: a polyketide bound heptaketide intermediate (325) undergoes epoxidation and ring closure to give the THF ring containing heptaketide (327), which is then elongated to an undecaketide (338) and undergoes *O*-methylation and hydroxylation during these stages but the order of events is unclear. Either tetronate formation could lead to PKS release (339) followed by pyran/cyclohexane ring formation (red), or the two rings could be formed whilst PKS-bound (340) followed by final tetronate forming PKS release (blue). The undecaketide (338) may even be released from the PKS before any ring formation.**

It still remains unclear whether the tetronate ring is formed with the pyran/cyclohexane ring formation occurring during or post-PKS construction. To gather further information about the overall timing of ring formation, it was envisaged that a chemical probe could be designed to capture late stage PKS intermediates preceding tetronate formation. The design and utilisation of this type of probe will be discussed in the following section.

3.1.2 Determining the relative timing of ring formations in tetronasin with chemical probes

As mentioned in the introduction (**Section 1.6**), the tetronate portion of tetronasin and tetronomycin is postulated to derive from glycolyl (**341**) and glyceryl-bound ACP respectively. A chemical probe was therefore designed to effectively mimic the function of glycolyl-ACP involved in tetronate formation and compete with it to capture PKS-bound intermediates, interfering in regular tetronate formation (**Fig. 3.8**). Isolating these species and gathering MS^2 fragmentation may reveal whether the pyran or cyclohexane rings are already in place and therefore indicate whether tetronate formation and polyketide release is effectively the final step of the biosynthesis.
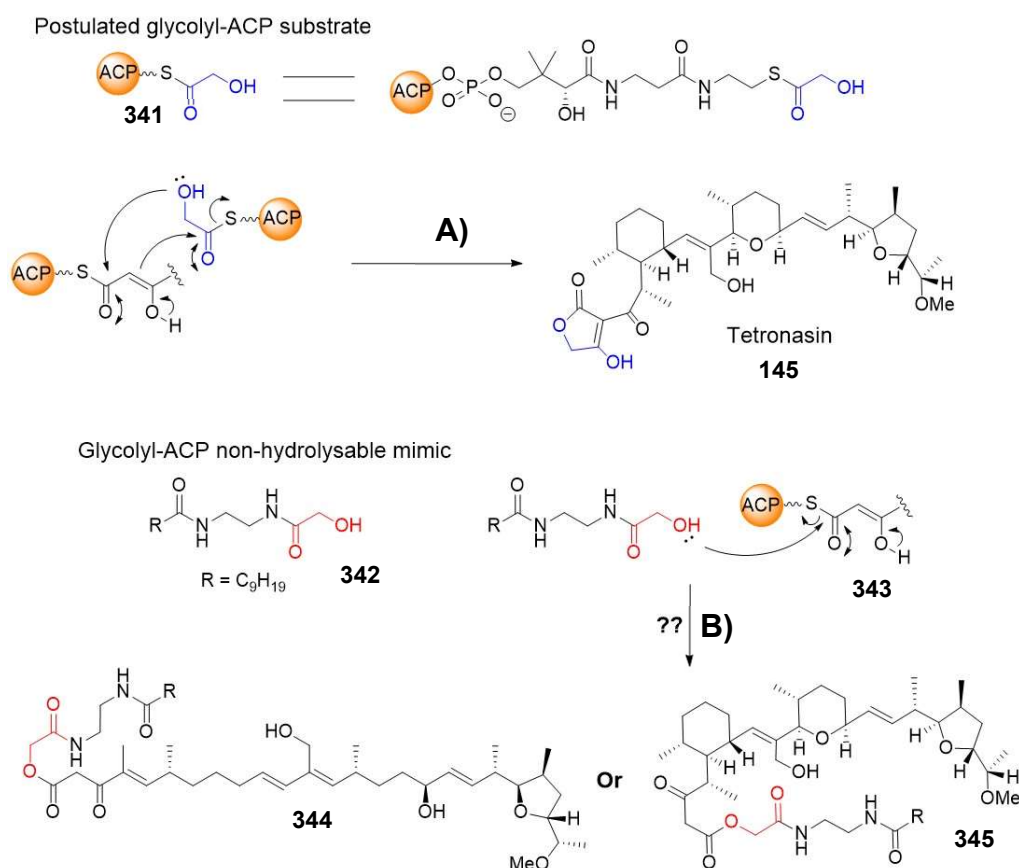
Figure 3.8 – **A glycolic acid-based probe (342) was designed to mimic glycolyl-ACP (341) allegedly involved in tetronate ring formation (A) in tetronasin (145). 342 may capture the PKS intermediate (343) by competing with glycolyl-ACP (B), hopefully allowing detection of a linear chain (344) or the pyran/cyclohexane ring closed species (345) to elucidate timing of events in tetronasin biosynthesis.**

As well as this substrate, a glycine-derived probe synthesised by group member Daniel Leng for NRPS intermediate capture[88] was readily available and had been used for feeding experiments as well (**Fig. 3.9, upper**). This probe has a slightly shorter *N*-acyl chain band and has an amine functional group (**346**) instead of an alcohol moiety as in the glycolate based probe (**345**).

The rational for its use was twofold: the amine is an isostere for an alcohol group and any intermediate captured would be tethered by an amide linkage rather than an ester, leading to better hydrolytic stability of the captured intermediate; amide formation would also lead to clearer MS^2 fragments, allowing better characterisation of any putative intercepted species. A glyceryl probe (**347**) was also synthesised for the purpose of investigating tetronomycin biosynthesis and tetronate ring formation in that pathway (**Fig. 3.9, lower**) at a later stage.
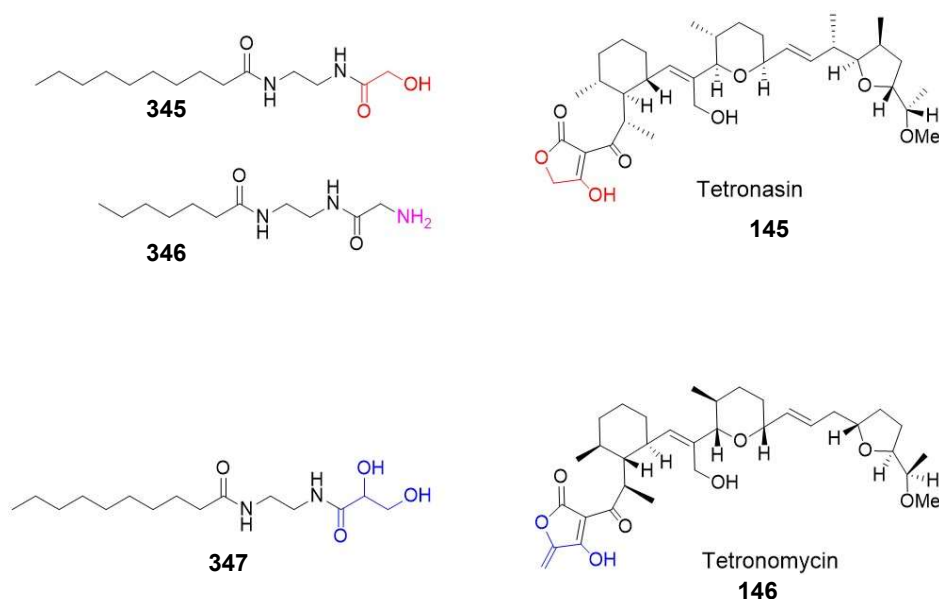
Figure 3.9 – **(Upper) As well as using the glycolate based probe (345) for hopefully intercepting tetronasin (145) biosynthetic intermediates, the glycine probe (346) originally used for NRPS species capture was tested in the hope that intercepted intermediates would be more stable owing to the amide linkage. (Lower) A glycerate probe (347) was also synthesised for intercepting intermediates of the tetronomycin (146) biosynthetic pathway; this is because glyceryl-ACP is the believed source of the three-carbon fragment of the tetronate ring (blue) experiments with this pathway were conducted in this work.**

3.1.3 Synthesis of non-hydrolysable glycolyl/glyceryl ACP mimics

Both the glycolyl- (**342**) and glyceryl- (**347**) based probes were prepared according to the synthetic scheme below.
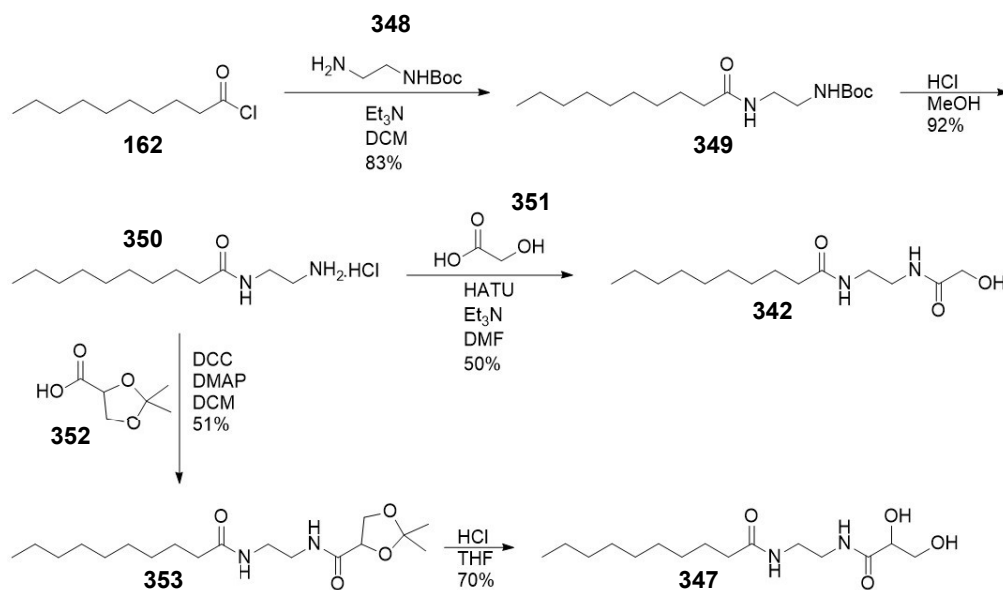
Scheme 3.6 – Synthetic route to glycolyl probe (342) and glyceryl probe (347)

Reaction between decanoyl chloride (**162**) and mono-Boc protected ethylene diamine (**348**) gave the amide product (**349**) in 83% yield after recrystallisation. This was then deprotected in acidic conditions to give the amine HCl salt (**350**) in high yield. This intermediate was used to prepare the two probes. Coupling of the salt with glycolic acid (**351**) gave the glycolyl substrate (**342**) in a 50% yield but purification by chromatography proved difficult; recrystallisation from hot EtOAc eventually gave pure product but the yield suffered as a result. For the second probe, the amine salt (**350**) was coupled with the commercially available racemic diol protected glyceryl acid (**352**) to give the amide (**353**) in a modest 51% yield, followed by acidic deprotection of the acetonide to afford the final glyceryl substrate (**347**) in 70% yield (**Scheme 3.6**).

3.1.4 Feeding of *S. longisporoflavus* strains with the glycolyl/glycine ACP mimic probes

The wildtype and mutant strains of *S. longisporoflavus* ($\Delta tsn11$ and $\Delta tsn15$) were grown on plates supplemented with the glycolyl probe (**342**) and the glycine probe (**346**, provided by Daniel Leng) in both 1 mM and 2 mM concentration. All possible intercepted intermediates resulting from different cyclisation patterns and timings were searched for. No apparent toxicity was displayed by the two probes towards any of the strains at the different concentrations: the same uniform dark brown colour and extent of mycelium growth was observed in all plates and tetronasin was detected in the wildtype extracts in reasonably consistent abundance. Interestingly, whilst the extracts containing the glycine probe displayed an extremely strong parent ion and MS^2 fragments, the glycolyl probe displayed poor proton and sodium adduct intensity, suggesting that this substrate is likely more degraded.

Unfortunately, for both probes, no reliable intercepted species were detected; time constraints meant that these experiments could not be repeated in either solid or liquid culture. It is possible that the increased polarity of these two probes affects their bioavailability and distribution in bacterial cells, preventing their access to the target proteins in sufficient amount to observe any detectable effect. Hydrolysis of the ester in any captured species with the glycolyl probe also cannot be ruled out. It is also possible that the predicted species are further processed *in vivo* by additional enzymes and undergo further transformations. Carrying out the experiments *in vitro* with the recombinant proteins may overcome these obstacles and allow detection of the species previously envisaged to be intermediates of tetronasin biosynthesis.

3.2 Understanding substrate-enzyme interactions in tetronasin ring

In order to gather mechanistic information on the THF ring formation in tetronasin, we began a collaboration with Prof. Marco Dias (Sao Paulo, Brazil) who crystallised the putative epoxidase (TsnC) and epoxide hydrolase (TsnB) enzymes involved in the THF ring formation (personal communication) in the absence of any substrates. There is considerable interest in understanding the molecular basis of polyether ring formation, however literature information in the area is sparse. Some epoxidases have been previously characterized in terms of their activity and specificity, such as Lsd18 from the lasalocid A BGC which is responsible for a double epoxidation of the polyketide backbone.[146] Given the reported use of recombinant Lsd18 to turn over simple alkenes (mimic for six carbon motif of the natural substrate) into epoxides[147] and the acquisition of a co-crystal structure for the subsequent epoxide hydrolase Lsd19 with lasalocid A (**103**),[148] it was hoped that generating basic mimic substrates for the TsnC and TsnB may allow a co-crystal structure to be obtained.

3.2.1 Design of substrates for co-crystallisation with TsnC and TsnB

At the point of designing these substrates, our feeding experiment data suggested that a PKS-bound heptaketide (**325**) would be a substrate of epoxidation and THF ring formation. However, the synthesis of such a substrate would be complex due to the presence of multiple stereocentres. We previously considered that a tetraketide substrate would be the shortest possible molecule whereby the THF ring could be formed (**Section 3.1.1**). Therefore, we reasoned that a mimic of this molecule could be a starting point for *in vitro* mechanistic investigations and structural studies. The natural polyketide intermediates are tethered to the PKS *via* a thioester linkage during the two transformations so ideally an SNAc derivative would have been synthesised; due to possible hydrolytic instability of a thioester group (especially in protein crystallisation conditions), a more robust and simple methyl ester functionality was chosen. Chirality

and branching groups were removed in order to simplify the synthesis and make compounds rapidly accessible in good amounts for testing (**Fig. 3.10**).
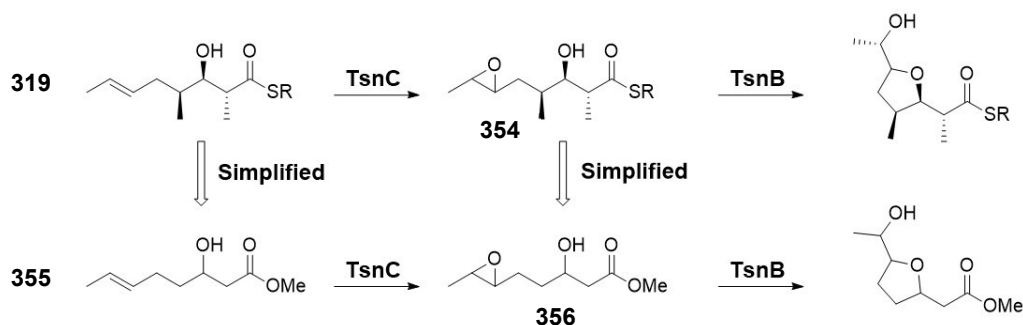
Figure 3.10 – **Design of simple mimics for the putative substrates for co-crystallisation with TsnC and TsnB. The proposed TsnC and TsnB enzyme-bound substrates 319 and 354 were simplified respectively to the partial mimics 355 and 356 allowing fast synthesis but maintaining structural similarity.**

3.2.2 Synthesis of substrates for co-crystallisation with TsnB and TsnC

Substrates **355** and **356** were prepared using the synthetic route below (**Scheme 3.7**).
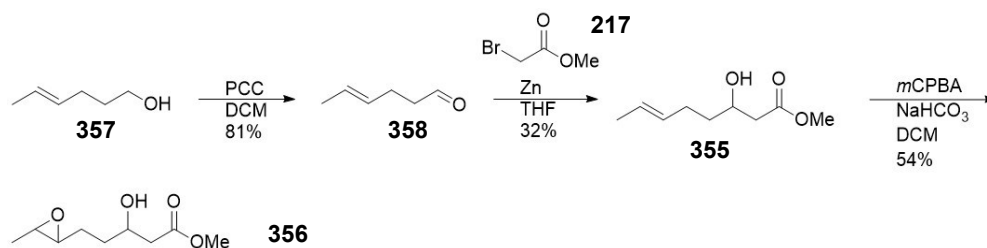
Scheme 3.7 – **Synthetic route to epoxidase (TsnC) substrate (355) and the epoxide hydrolase (TsnB) substrate (356).**

PCC oxidation of commercially available (E)-hex-4-enol (**357**) gave the aldehyde (**358**) in 81% yield; a Reformatsky reaction with methyl bromoacetate (**217**) gave the racemic hydroxy ester product (**355**) in 32% yield after purification: this was sent to Marcio Dias for use without chiral separation. Treatment with *m*CPBA gave the epoxide (**356**) as a mix of stereoisomers in an unknown ratio. The NMR data was complex and indicated a mix of products but LCMS showed a single mass peak for the epoxide. This was also sent without any form of chiral separation. *In vitro* testing of these two compounds and co-crystallisation trials with TsnC and TsnB in the Dias group are still ongoing.

3.1   Summary of results

As outlined in the introduction, the ring formation mechanisms and timing in tetronasin biosynthesis were not well understood at the beginning of our investigations. Initial growth of wild type and mutant *S. longisporoflavus* strains on plate all appeared to show production of species matching the mass ion of tetronasin, but the MS^2 profile and retention time for species detected for the mutant strains was slightly different. Initial toxicity screening of probes **104** and **160** showed a marked decrease in metabolite abundance across the three strains for 1 mM to 2 mM probe concentration, with probe **160** displaying a larger impact on production then **104** did. The phenotype of the plates also varied in terms of colour and the appearance of bacterial mycelium. Intermediate detection was particularly poor on solid cultures and with the exception of di/triketide type captured species, no detectable mass ions pertinent to tetronasin biosynthesis were found despite the clear production of tetronasin in the presence of the probes.

Repeat experiments run in liquid culture showed comparable metabolite production, and more putative intermediates as result of improved ester hydrolysis and decarboxylation. Two significantly abundant species detected were two putative heptaketides (**323-324**) corresponding to a nascent linear PKS chain as expected from the predicted domain architecture, but also oxidised heptaketides as the result of epoxidation followed by dehydration. MS^2 characterisation of these two putative intermediates did not offer highly

convincing fragments though, and attempts to use higher energy fragmentation techniques did not improve their structural elucidation either.

An undecaketide species (**328**) was also detected earlier on, indicating that hydroxylation and *O*-methylation of the PKS intermediate may occur whilst enzyme bound. Several MS^2 fragments were detected following experiments at higher probe concentration, supporting the envisaged structure; however, it remained unclear whether the fully formed pyran and cyclohexane rings could be present in this species given that the formation of these rings would not lead to a mass change and therefore detection.

A chemical probe of different nature (**342**) was prepared in order to investigate the later stages of tetronasin biosynthesis by acting as a glycolyl- ACP mimic. This and a previously synthesised NRPS probe based in glycine were tested in various concentrations in solid culture with the three *S. longisporoflavus* strains available to us. No toxicity was observed but no pertinent species were detected either. An analogous probe (**347**) based on glycerate was prepared for the investigation of tetronomycin biosynthesis but no studies were conducted on *Streptomyces* sp. (the producer of tetronomycin) due to time constraints. In addition, two compounds (**355-356**) were synthetically prepared to act as putative substrate mimics and for co-crystallisation studies with TsnC and TsnB, the enzymes allegedly involved in THF ring formation in tetronasin. Our collaborator Prof. Marcio Dias is currently working with the putative epoxidase and epoxide hydrolase enzymes and results of this work will be reported in due course. Following my investigations, the mechanism of formation of cyclohexane and pyran ring formation has been published.[141] Further work towards the complete understanding of tetronasin biosynthesis will be discussed in Chapter 5.

4   Exploring PKS substrate specificity with novel chemical probes

4.1   Expanding the complexity of methyl ester chain terminator probes

Section 1.4.7 in the introductory chapter outlined the evolution and utilisation of chain terminator chemical probes within our group to explore polyketide biosynthesis from the perspective of generating new "unnatural" products and also deducing the mechanism and timing of biosynthetic steps.[84,85] The use of probes based on synthetic malonyl coenzyme-A (**358**)[74] evolved into the use of *N*-acetyl cysteamine mimics (**359**) which are relatively simple substrates to prepare.[76] The lipophilicity of these molecules has been increased by replacing the *N*-acetyl moiety with an *N*-decyl moiety (**360**) in order to aid cellular uptake through the hydrophobic cell membrane of the polyketide producing organism.[85,86] (**Scheme. 4.1**).

Scheme 4.1 – Chain termination probe development. From a malonyl-CoA mimic (38) to a pantetheine analogue (359), and N-acyl cysteamine mimics (360), which can be utilised for in vivo studies.

The chain termination probes have also been diversified in other ways including the addition of clickable tags, radioactive labels, malonyl substitutions and modifying the ester group (**Fig. 4.1**). Most of these changes have positively affected the properties of the probes for offloading intermediates. Adding a fluorine substituent in the malonate position, for instance, vastly increased the rate of methyl ester hydrolysis leading to increased amounts of the "active" malonate probe *in vivo*, allowing a more efficient capture of polyketide intermediates in terms of structural variety and abundance of species.
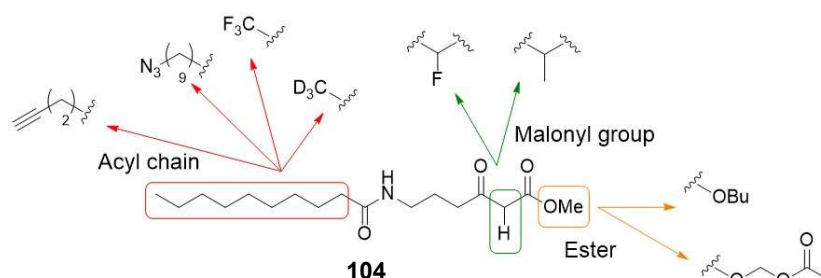
Figure 4.1 – **Examples of some key chemical changes previously made to the decanamido probe 104. The majority of previously reported changes have been made to the acyl chain (red), the ester group (orange) and substituting the malonyl group (green).**

The amide moiety in **104** remained a key point of structural diversification that had not been previously explored in our group. This was interesting from the standpoint of exploring polyketide substrate specificity and ultimately generating unnatural compounds carrying different functionality and potentially displaying novel properties. As mentioned in chapter 2, the methyl malonyl substituent for the decanamido probe had also not been synthesised which may be a more suitable for PKS systems utilising the methylmalonyl extender unit. Chapter 4 focuses on amide group changes which had not previously been explored.

With this in mind, several functional group changes were planned to the amide group of **104** as illustrated below (**Fig. 4.2**).
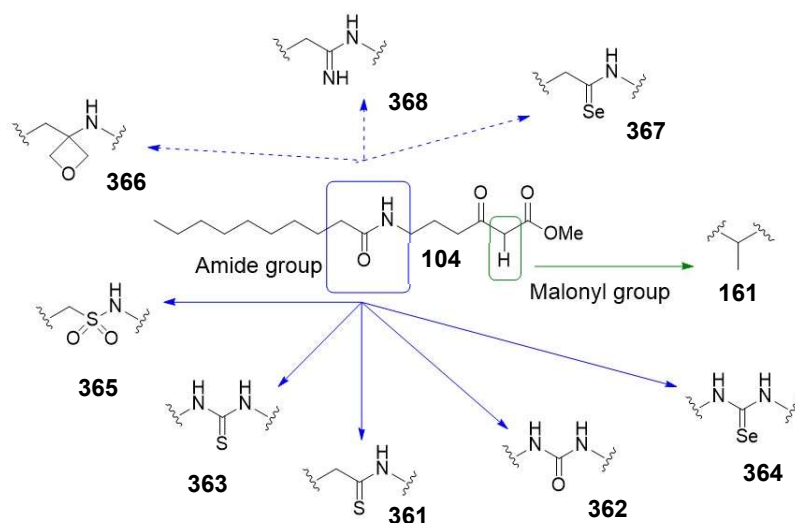
Figure 4.2 – **Overview of functional group changes made to the decanamido probe (104). Aside from the malonyl substitutions (161) previously explored, other analogues were based on amide variation (361-368). Solid line analogues were successfully synthesised, whereas dotted lines represent unsuccessful synthetic routes.**

Most modifications were envisaged as isosteric replacements (commonly used in medicinal chemistry)[149] that may affect properties such as hydrogen bonding (*e.g.* replacement of an amide with an amidine moiety), or replacements that would maintain the same hydrogen bonding properties but change the nature of the amide moiety (*e.g.* hydrolysable, versus a non-hydrolysable amino oxetane, **366**). Also, it was envisaged that chain termination probes containing different atoms such as selenium would offer more easily detectable offloaded intermediates by LC-HRMS given its highly characteristic isotope pattern.

All the diversified probes illustrated in **Fig. 4.2** were envisaged as methyl esters in the first instance in order to devise a straightforward chemical synthesis for the substrates and preliminarily evaluate their suitability and usefulness for *in vivo* studies.[85,86] Our attempts to prepare substrates **361-368** and utilise them as novel probes for polyketide biosynthesis will be described in the next sections.

4.1.1 Synthesis of chain termination probes bearing novel functional groups

4.1.1.1 Synthesis of thioamide probe (361)
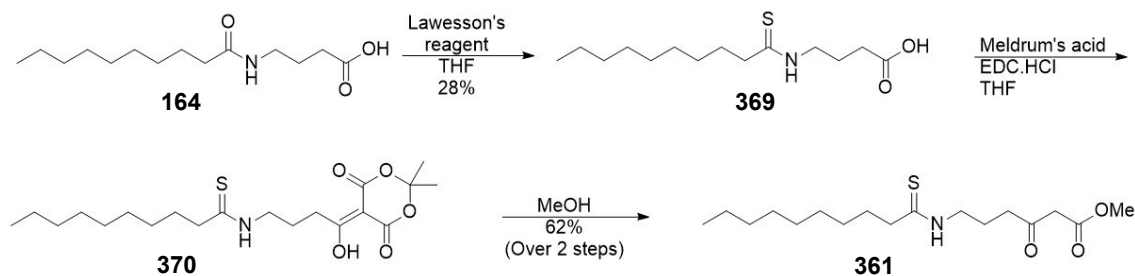
Scheme 4.2 – **Synthetic route to thioamide probe (361).**

The synthesis of **361** was carried out from the amide carboxylic acid intermediate (**164**). Lawesson's reagent was used to convert the amide into the thioamide (**369**); the low yield of this product was due to a competing reaction of thiocarboxylic acid formation as well as thioamide formation. The thioamide (**369**) was then coupled to Meldrum's acid giving the adduct (**370**). This was not purified and taken immediately into the thermolysis step to give the final probe (**361**) in 62% yield over the last two steps (**Scheme 4.2**).

4.1.1.2 Synthesis of sulphonamide probe (365)
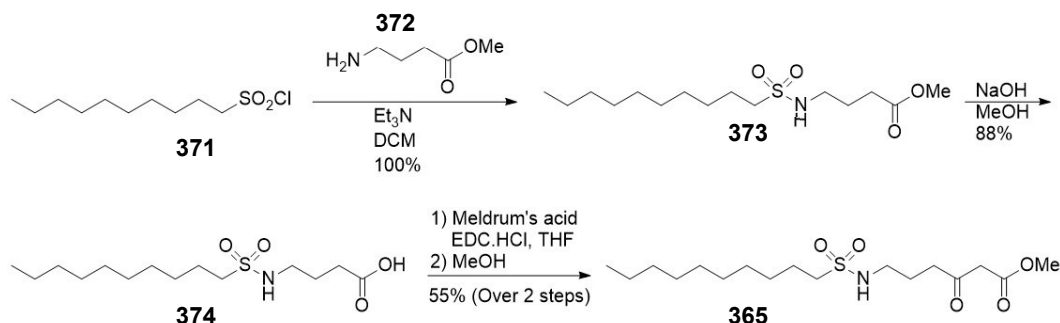
Scheme 4.3 – **Synthetic route to sulphonamide probe (365).**

Reaction between the long chain sulphonyl chloride (**371**) and the methyl ester of GABA (**372**) afforded the sulphonamide product (**373**) in quantitative yield. The ester was hydrolysed in basic conditions to give the corresponding carboxylic acid (**374**), which was coupled to Meldrum's acid. The adduct was isolated and utilised without further purification in the final reaction with methanol to afford the sulphonamide probe (**365**) in 55% yield (over the last two steps and final purification) (**Scheme 4.3**).

4.1.1.3 Synthesis of urea probe (362)
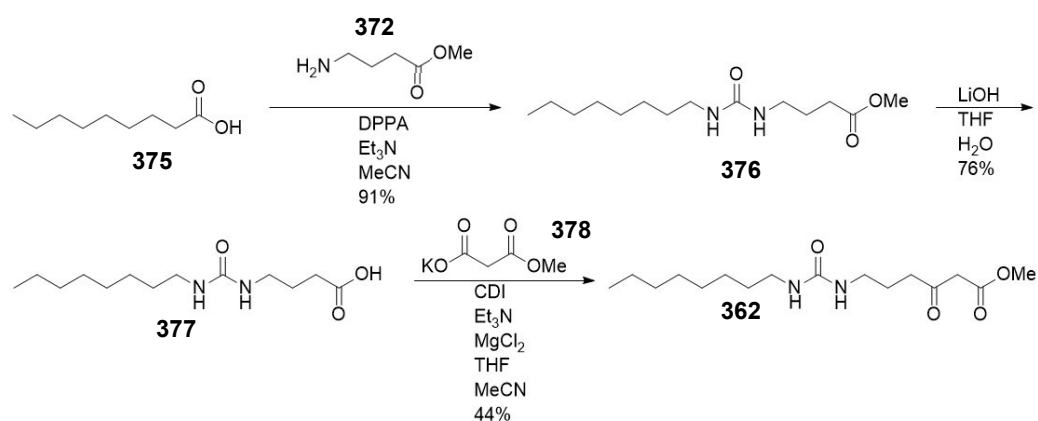
*Scheme 4.4 – **Synthetic route to urea probe (362).***

The synthesis of **362** (**Scheme 4.4**) began with a Curtius rearrangement of nonanoic acid (**375**) using diphenylphosphoryl azide (DPPA) to give an isocyanate intermediate upon heating and loss of N_2. Opening of the isocyanate with the GABA methyl ester (**372**) delivered the urea product (**376**) in 91% yield. Basic hydrolysis of the methyl ester gave the carboxylic acid (**377**) which was insoluble in THF and was eventually filtered rather than extracted, leading to some loss in yield. The previous chemistry with Meldrum's acid did not work on this substrate, therefore an alternative procedure was attempted, using carbonyl di-imidazole (CDI) to activate the carboxylic acid and displace it using a magnesium enolate formed *in situ* from the monopotassium salt of dimethyl malonate (**378**). The urea probe (**362**) was obtained in a 44% yield after purification (**Scheme 4.4**).

148

4.1.1.4 Synthesis of thiourea probe (363)
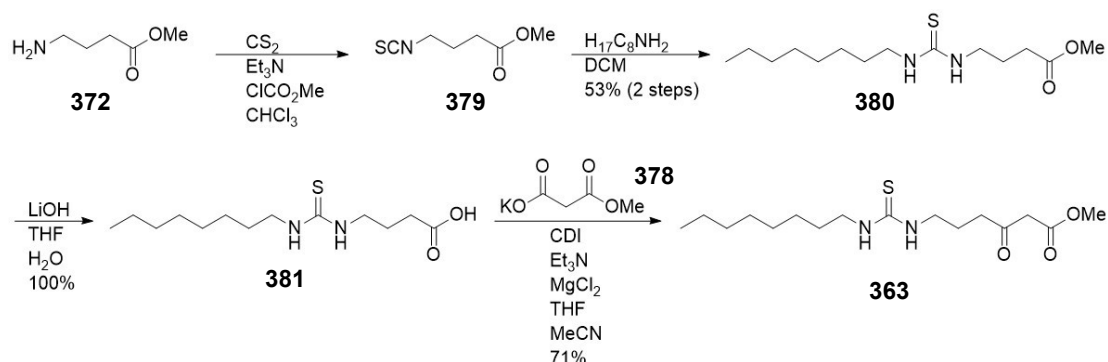
Scheme 4.5 – Synthetic route to thiourea probe (363).

GABA methyl ester (**372**) was first converted to the isothiocyanate intermediate (**379**) using carbon disulphide and methyl chloroformate. Without purification, the isothiocyanate was further reacted with *n*-octylamine to give the thiourea (**380**) in a 53% yield over two steps. Quantitative yield was achieved in the ester hydrolysis giving the carboxylic acid intermediate (**381**) which was further reacted without purification. The final thiourea probe (**363**) was obtained in a 71% yield using the magnesium enolate reaction previously employed (**Scheme 4.5**).

4.1.1.5 Attempted synthesis of selenoamide probe (367)

The synthesis of **367** was planned around the use of Woollins' reagent to convert the previously made amide intermediate (**164**) to the selenoamide analogue (**382**) directly (**Scheme 4.6**). However, no product was isolated and instead an unexpected side product (**383**) was obtained in 50% yield. This was characterised as a lactone resulting from dehydrative cyclisation of the starting material; acetic anhydride has also been reported as a reagent for this type of transformation.[150] The higher temperature reaction conditions required for with Woolins' reagent and instability of the selenoamide may be the cause of this. No other routes towards **367** were pursued and an alternative selenium functional group for the amide replacement was chosen instead.
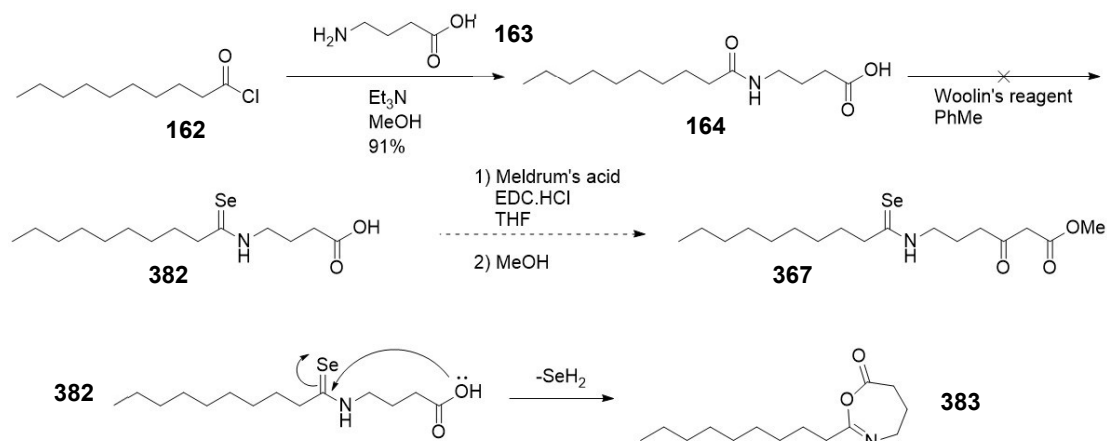
Scheme 4.6 – **Attempted synthesis of selenoamide probe (367). No product was obtained in the step to make the selenoamide (382) from 164; instead, an azo-lactone side product (383) was isolated.**

4.1.1.6 Synthesis of selenourea probe (364)

Synthesising a selenourea derivative instead of a selenoamide offered a safe, more desirable and well characterised route. GABA methyl ester (**372**) was first converted in quantitative yield to the formamide (**384**), which was then dehydrated using phosphorus oxychloride to the isocyanide (**385**) in good yield. Refluxing in selenium afforded the isoselenocyanate (**386**) which was converted to the selenourea (**387**) using octylamine. A mild ester hydrolysis accessed the carboxylic acid (**388**), which was used crude in the final reaction to give selenourea probe (**364**) in 50% yield over two steps (**Scheme 4.7**).
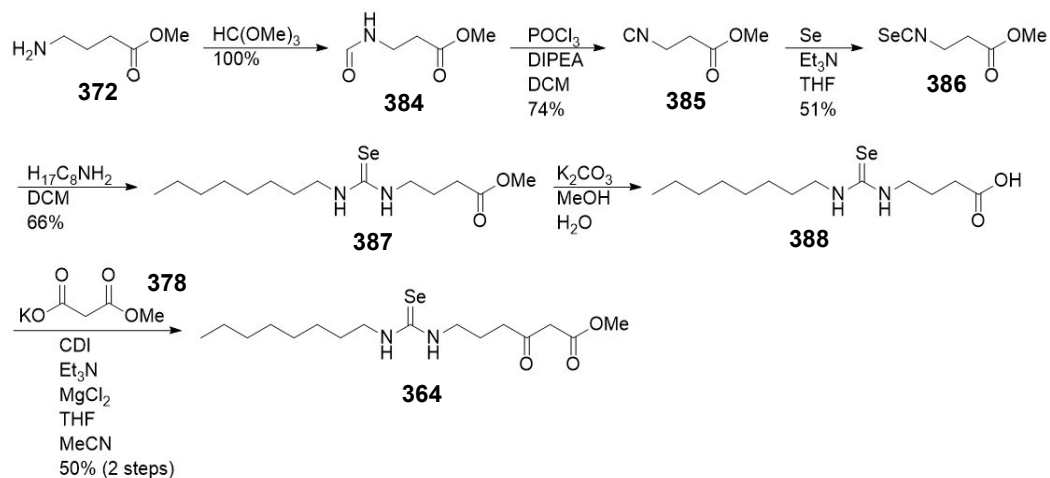
Scheme 4.7 – **Synthetic route to selenourea probe (364).**

4.1.1.7 Attempted synthesis of amino-oxetane probe (366)
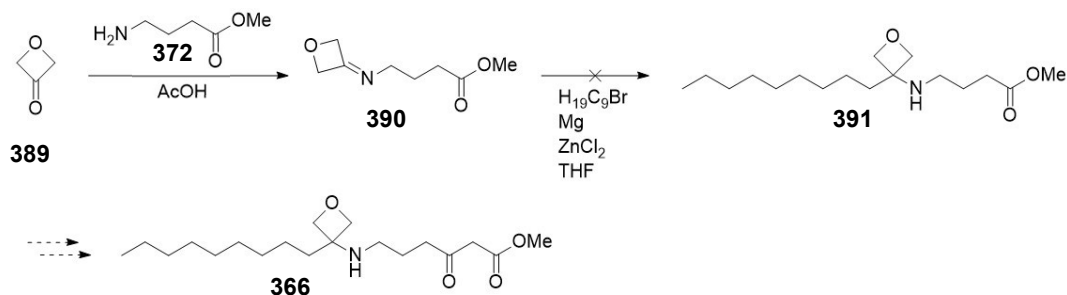
Scheme 4.8 – Attempted synthesis of oxetane probe (366).

One of many bioisosteric functional groups for amides are α-amino oxetanes.[151] It was envisaged that an oxetane probe (**366**) could have similar substrate recognition as the *N*-decanamido probe (**104**), in particular these oxetanes are of interest as amide isosteres due to their H-bonding properties but resistance to proteases.[152] Imine formation between oxetane-2-one (**389**) and GABA methyl ester (**372**) appeared to be successful and NMR of the crude product (**390**) looked promising. However multiple attempts of addition of an organozinc reagent to the imine gave no observable product (**391**). There is scarce literature for addition onto 2-imino-oxetane substrates, possibly due to steric hindrance or oxetane instability. The oxetane probe synthesis (**Scheme 4.8**) was not pursued further.

4.1.1.8 Attempted synthesis of amidine probe (368)
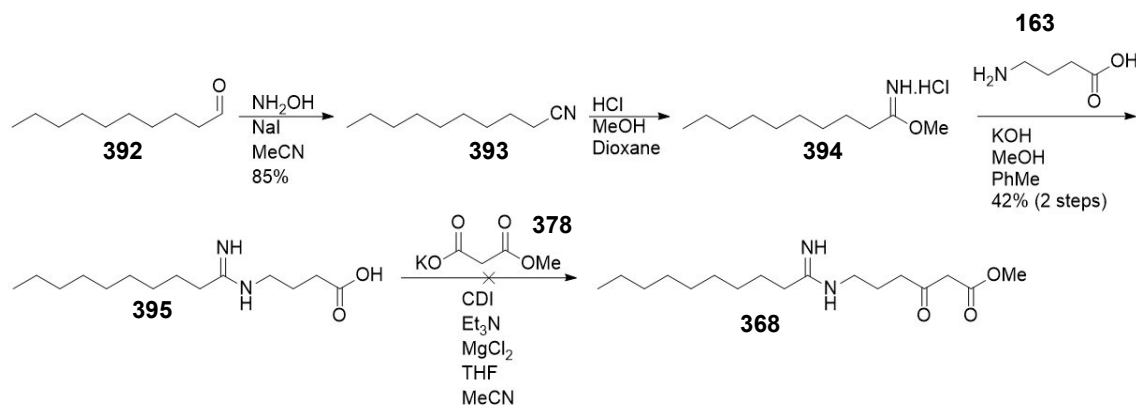
Scheme 4.9 – Attempted synthesis of amidine probe (368).

Decanal (**392**) was first converted to an oxime intermediate, which was catalytically dehydrated to decanitrile (**393**) using NaI in high yield. A Pinner reaction was then used to generate the amidine (**395**) through reaction of the imidate salt (**394**) with GABA (**163**) in 42% yield over two steps (**Scheme 4.9**). Attempts to react this intermediate further to the final amidine probe (**395**) were hampered by the reactivity of the amidine functional group towards intramolecular cyclisation. Protection of amidine carboxylic acid N-H with Boc anhydride was also unsuccessful due to the insolubility of **395** in any solvent other than methanol, which lead to esterification. Synthesis of the amidine probe (**368**) from the thioamide probe (**361**) was attempted by a method of silver salt catalysed amination;[153] product formation was observed by LCMS but appeared to degrade during work-up or purification.

4.1.2 Evaluation of novel probes *in vivo* on 6MSAS heterologously expressed in *E. coli*

The novel probes were initially checked for their efficacy in offloading intermediates from the iterative type 1 PKS 6MSAS, for which the group had previously reported successful intermediate capture (**Fig. 4.3**).[86] A pET-28a vector containing 6MSAS from *P. patulum* was used. Transformation of this vector into BAP1 *E. coli* and induction with IPTG allowed heterologous expression of 6MSAS and therefore production of 6MSA in *E. coli*. Probes **104** and **361-365** were supplemented to these expression cultures for a period of 24 h, and cultures were then extracted in EtOAc as per the previously reported method before concentration and resuspension of extracts in MeOH for LC-HRMS[n] analysis.
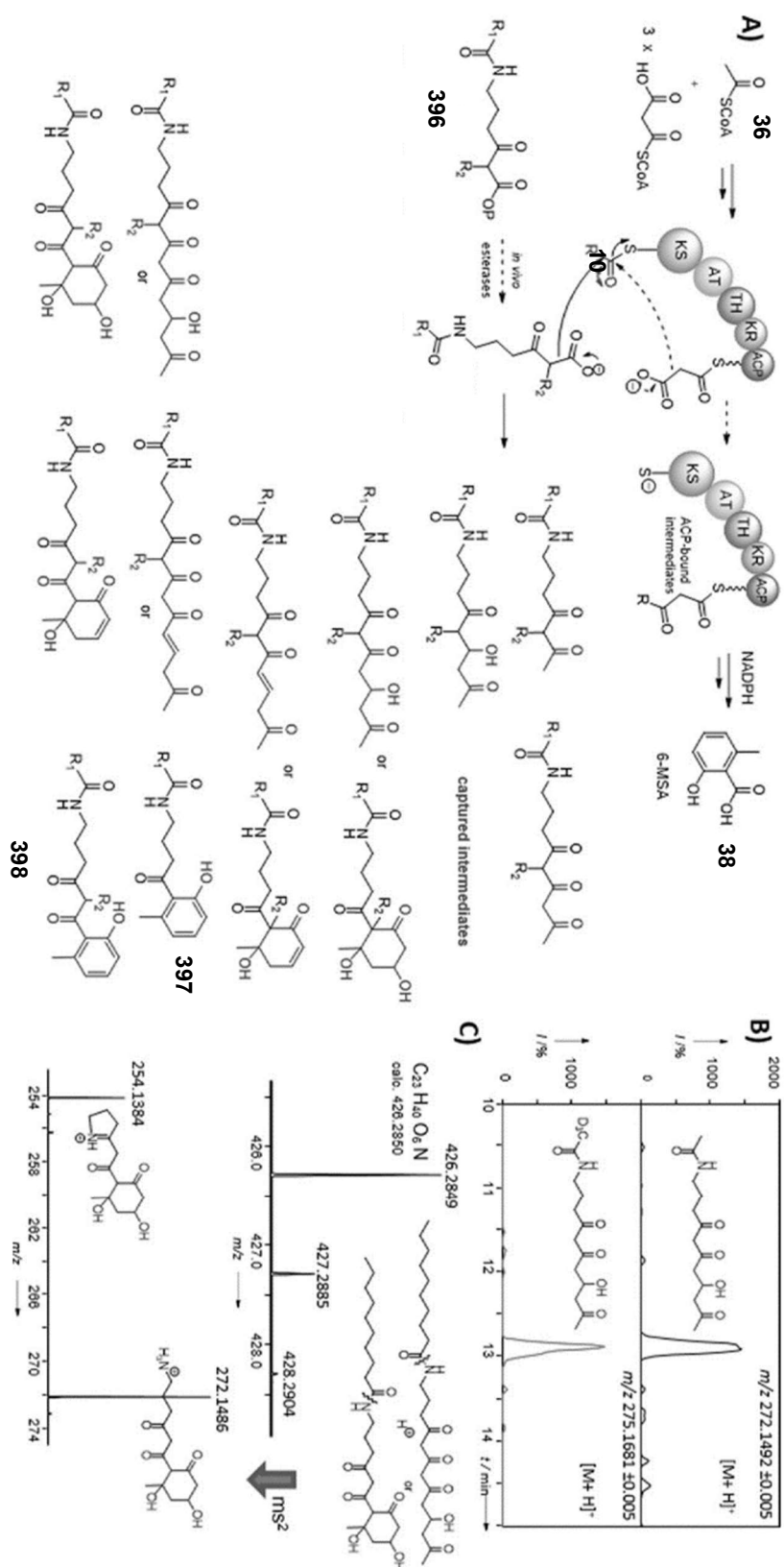
Figure 4.3 – **A) 6MSA (38) is produced by the iterative five domain enzyme 6MSAS from acetyl coenzyme-A (36) and malonyl coenzyme-A (10). Chain terminators (396) fed to the culture are hydrolysed *in vivo* the active carboxylate and capture intermediates from the PKS of ranging length including phenolic tetraketides (397) and pentaketides (398). B) Putative captured intermediate parent ions are detected by LC-HRMS. C) MS² fragmentation of respective parent ion leads to diagnostic intermediate fragments.**

153

4.1.2.1 Analysis of novel probe methyl ester hydrolysis to their active carboxylate form

In the first instance the extent of *in vivo* methyl ester hydrolysis for all probes was looked at in each feeding experiment by comparing the approximate intensity of probes parent ions to that of monoketide species resulting from decarboxylation of β-ketoacids (**Fig. 4.4**).
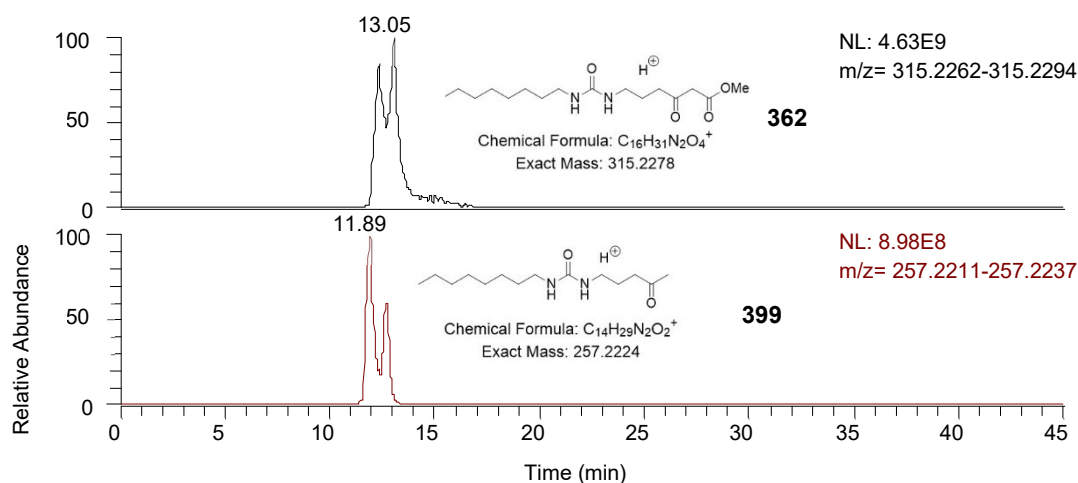
Figure 4.4 – **LC-HRMS chromatogram comparing the intensity of the urea probe (362) to that of the hydrolysed and decarboxylated derivative (399) in order to determine the approximate efficiency of methyl ester hydrolysis *in vivo*.**

Poor ester hydrolysis generally correlates to poor polyketide intermediate capture because of low amounts of active probes generated *in situ* (**Table 4.1**). All probes were detected mostly as proton adducts, with the exception of the sulphonamide derivative (**365**) appearing in the form of sodium adduct. The thiourea derivative (**363**) displayed the highest hydrolysis ratio (approximately 39%), with most of the other compounds displaying 15-20% hydrolysis. Interestingly the selenourea derivative (**364**) appeared mostly intact and was mainly detected as a dehydro-compound. This was confirmed by the extremely diagnostic isotopic pattern for selenium containing species. Besides,

cultures used with this probe visually appeared less optically dense, suggesting toxicity towards the bacteria.

Probe	Ion	Intact probe peak area	Monoketide probe peak area	Approx. % hydrolysis
Decanamido (104)	MH+	1.91 E11	2.69 E10	14.1
Urea (362)	MH+	3.01 E11	3.54 E10	11.7
Thiourea (363)	MH+	2.08 E11	8.16 E10	39.3
Thioamide (361)	MH+	2.04 E11	3.35 E08	16.1
Sulphonamide (365)	MNa+	1.81 E09	3.02 E08	16.7
Selenourea (364)	[M-H2]H+	9.56 E10	5.05 E08	0.5

Table 4.1 – **Peak areas of probes and respective decarboxylated intensities indicating hydrolysis efficiency, the thiourea appears to hydrolyse most abundantly with most other probes at around 15%. Note- selenourea appears as the [M-H$_2$]H$^+$ and the sulphonamide ion appears as the sodium adduct.**

As the extender unit for 6MSA is a malonyl group, the methylmalonate (**161**) probe was not used as the PKS would likely be more preferential for unsubstituted malonyl probes.

4.1.2.2   Analysis of putative intermediates captured by novel probes

Some putative 6MSA polyketide intermediates were initially detected along with diagnostic MS^2 fragments; other putative later stage biosynthetic intermediates also were detected but could not be confirmed due to the lack of recognisable MS^2 fragments. Most probes were able to offload di and triketide putative intermediates as unreduced ketones, hydroxy ketone and "enone" species, some tetraketide species were also detected and confirmed with MS^2 of low abundance (**Fig. 4.5**). Whether these intermediates were cyclised or linear could not be confirmed.

There is also the possibility that some intermediates are being offloaded and further processed independently of the enzyme machinery, especially if the steps are thermodynamically favourable such as dehydrative aromatisation. Control extracts of *E. coli* expressing 6MSAS but lacking the chain terminator probe were checked for confirming the lack of identical mass peaks to exclude non-related artefacts, none of the mentioned species were detected in these extracts.
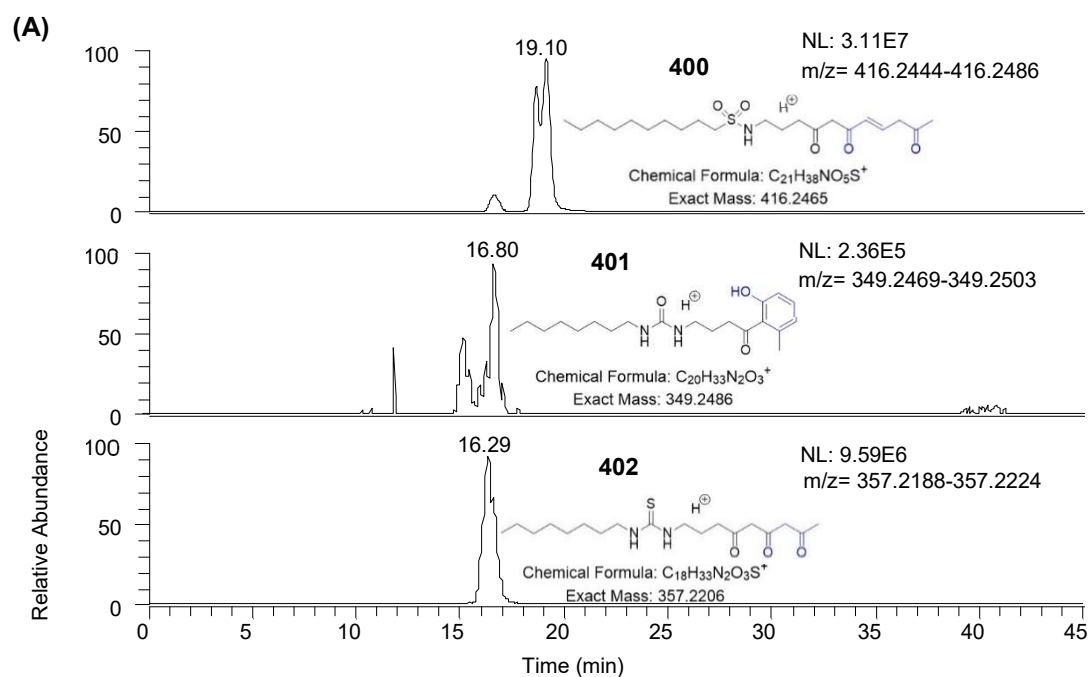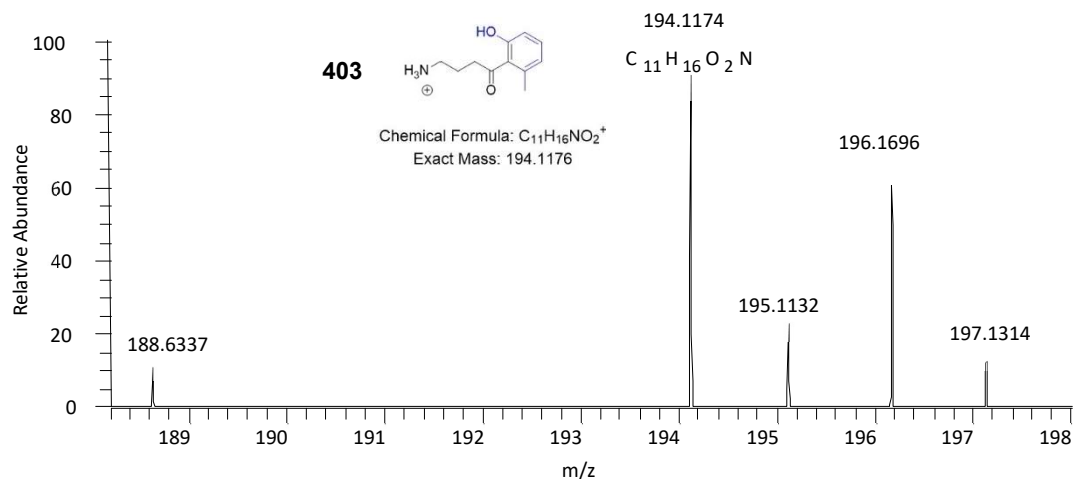
(C)

FTMS + p NSI d Full ms2 416.2448@hcd32.00 [110.0000-427.0000]
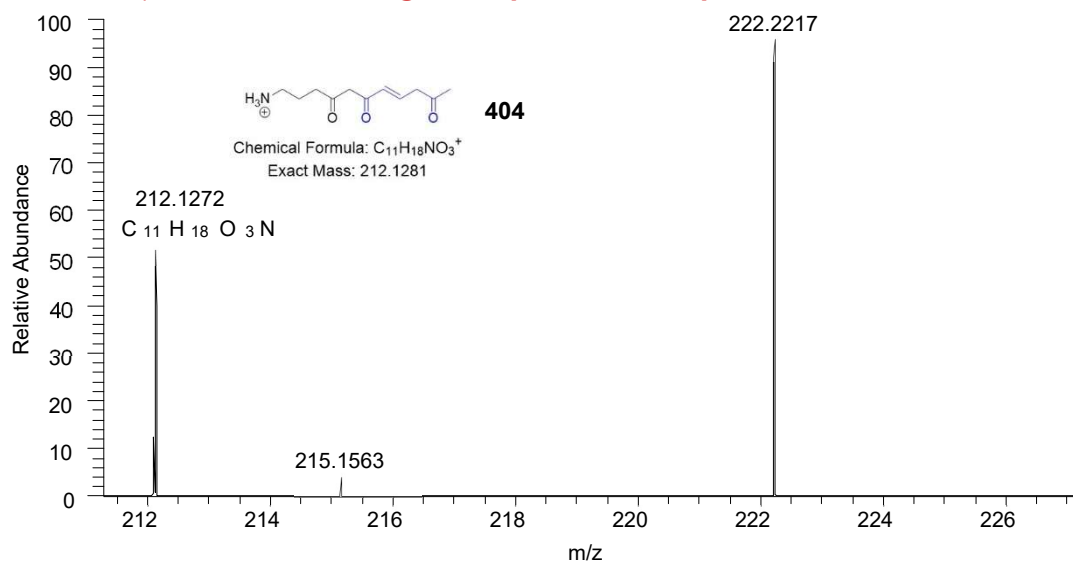
(D)

FTMS + p NSI d Full ms2 357.2199@hcd32.00 [110.0000-368.0000]
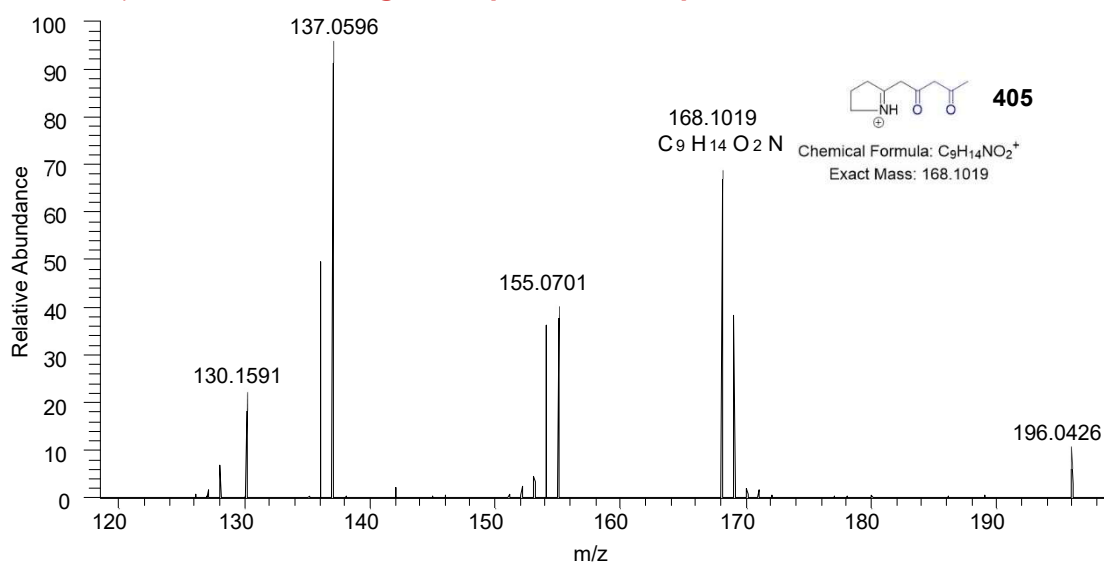
Figure 4.5 – **(A) LC-HRMS traces showing detection of putative captured polyketide intermediates from extracts using the thiourea (403), urea (402) and sulphonamide (401) probe. (B) MS² fragment (404) for the urea captured phenol (402). (C) MS² fragment (404) for the sulphonamide captured tetraketide (401). (D) MS² fragment (405) for the thiourea captured triketide (403).**

No particular probe appeared to conclusively capture a wider range of intermediates or in greater abundance than others, and there was no particular correlation between the abundance of active carboxylate probe and its offloading capability (**Table 4.2**). Unlike previously reported, no pentaketides were detected and the decanamido (**104**) probe also showed a poorer range of detectable intermediates.
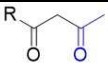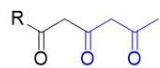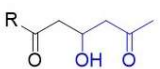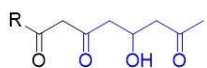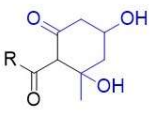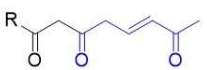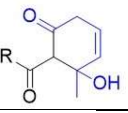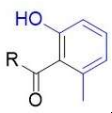
Intermediate	Decanamido (104)	Urea (362)	Thiourea (363)	Thioamide (361)	Sulphonamide (365)
	Y	Y	Y	Y	Y
	Y	Y	Y	Y	Y
	Y	Y	Y	Y	Y
Or	Y	Y	Y	Y	Y
Or	Y	Y	Y	N	Y
	N	Y	Y	N	Y

Table 4.2 – Summary of intermediates captured with novel probes where Y = detected, N = not detected.

Control experiments were also run feeding each probe to cultures for the untransformed *E. coli* to confirm that any captured intermediates were from offloading of 6MSAS and not from other possible FAS or PKS pathways. Analysis of these extracts showed many of the putative di and triketide and even some of the non-dehydrated tetraketides were in these extracts. This may be the result of the 6MSA intermediates being common substrates for another pathway or even the reaction of the probe with an unknown substrate or pathway. Overall the true offloading ability of these novel probes remains inconclusive in this system.

4.1.3   Evaluation of novel probes *in vivo* with lasalocid A producer S. *lasaliensis*

A modular type 1 PKS was also used to test the novel methyl ester probes: this was the las A synthase in *S. lasaliensis*. The twelve module PKS utilises malonyl, methylmalonyl and ethylmalonyl extension units along with diverse domain architecture and an inactive KR in the tenth module (**Fig. 1.15**). Much of the previous results of intermediate capture from this assembly line have been obtained from a mutant strain with an inactivated ACP in the twelfth module, preventing the production of lasalocid A and leaving polyketide intermediates bound to the enzyme.[85] The urea (**362**) and thiourea (**363**) were the most available probes in hand at the time of the preliminary feeding experiments, and were therefore used for testing.

The decanamido (**104**) probe was also tested as intermediates had previously been captured from the same assembly line and would therefore act as a standard. Feeding experiments were carried out using the previously reported method, adding the probes into the agar before streaking precultures of *S. lasaliensis* ACP12 (S970A) and allowing growth for five days. Issues emerged with these preliminary feedings: toxicity for the thiourea (**363**) probe was observed immediately when compared to other probes such as the urea (**362**) (**Fig. 4.6**).
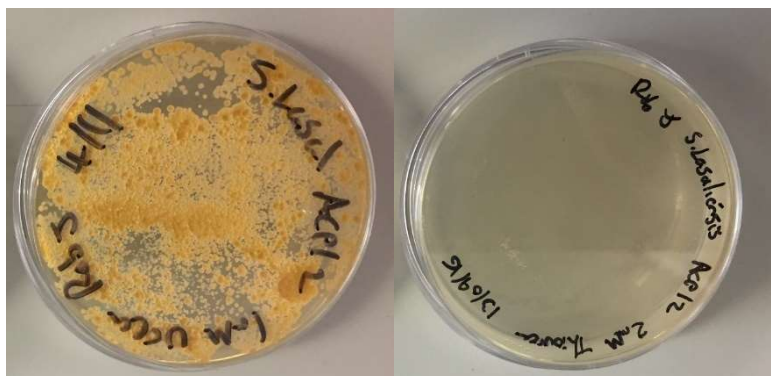
Figure 4.6 **– Thiourea (363) probe is toxic on *S. lasaliensis* plate (right) unlike urea (362) (left).**

Both the urea (**362**) and decanamido (**104**) probe extracts were analysed by LC-HRMS. Comparable ratio of intact probe and decarboxylated probe was observed and some early intermediate parent ions were found including diketide and triketide species with a few tetraketides (**Fig. 4.7**) confirmed by MS^2. Previously reported late stage intermediates were detected in low quantities but no diagnostic fragments for these were detected.

(B)
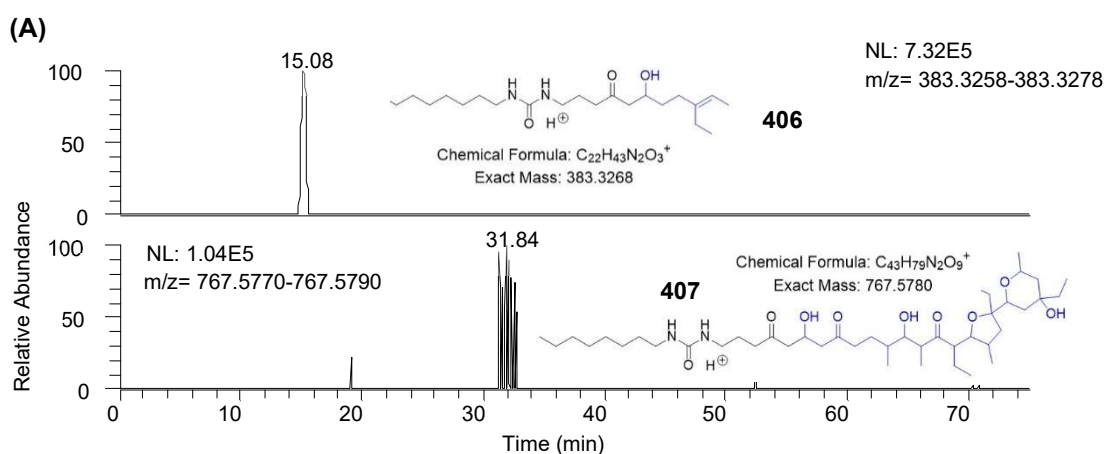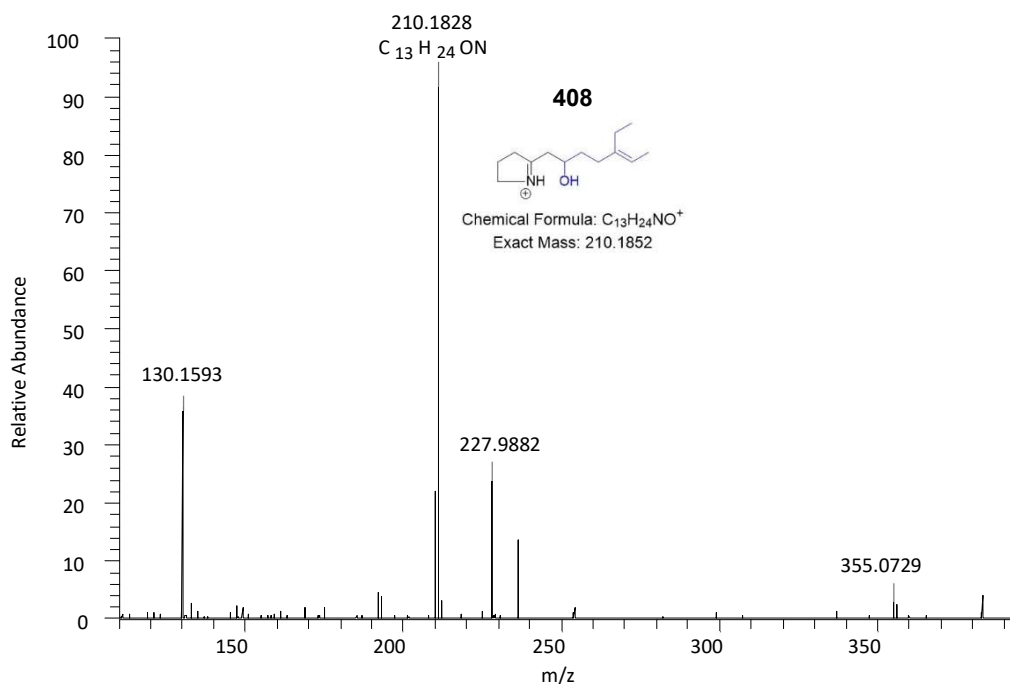
Figure 4.7 – **(A) LC-HRMS of extract from *S. lasaliensis* ACP12 (S970A) after supplementation with urea probe (362), a putative tetraketide (406) and polyether undecaketide (407) were detected. (B) MS^2 fragment (408) was found supporting the captured tetraketide (406). No fragments were detected for the putative polyether species.**

The decanamido (**104**) probe also showed fewer captured intermediates than had previously been reported;[83] liquid cultures of the feeding experiments were also conducted in parallel and the results obtained were identical. Given the lower efficacy of this probe than previously seen, this perhaps explains why intermediates from the novel probes were not detected well, if at all. It is difficult to pin-point the cause, but perhaps variation in the strain stock used or changes in the media composition may have had an effect on metabolic pathways within the organism altering the expression of the lasalocid cluster. Many of the probes were however tested in other systems *in vivo* and as different types of esters, which will be discussed in the forthcoming section.

161

4.1.4 Evaluation of novel probes *in vivo* with thiolactomycin producer *Lentzea* sp.

The novel amide functionalised probes were tested on the thiolactomycin (**108**) producer *Lentzea* sp. as intermediates had already been successfully offloaded from this system with malonyl substituted probes as previously discussed (**Section 2.1.2**). The thiourea (**363**) and selenourea (**364**) probes showed severe toxicity towards the strain as had been seen in *S. lasaliensis* feeding plates; very little or no mycelium could be seen, and unsurprisingly, no intermediates were detected along with no thiolactomycin production (**Fig. 4.8**).

*Figure 4.8 – **Toxicity in *Lentzea* sp. WT (left to right) of sulphonamide (365), thioamide (361), urea (362), thiourea (363) and selenourea (364) probes in 2 mM concentration.***

The sulphonamide probe showed no toxicity, whereas the urea and thioamide exhibited toxicity above 2 mM concentration but the growth was still limited in comparison to the classic amide probes. Whilst nowhere near as successful as the previously mentioned malonate substituted probes, some early intermediates were detected using the urea, thioamide and sulphonamide probes. These were restricted to diketide and hydroxy diketide species in all cases along with the dehydrated analogue using the thioamide only.

These probes clearly have a poor or non-existent ability in this strain but in retrospect, the iterative PKS responsible for thiolactomycin biosynthesis has been difficult to study and showed high specificity in substrate selection. The overriding factor in capturing thiolactomycin intermediates has been the abundance of the probe "active" carboxylate form, which has so far appeared significantly poorer in amide functionalised probes. This likely is due to the *in vivo* esterases having a poor recognition and activity towards them.

4.2 Designing and testing of photolabile probes

As well as testing probes *in vivo* and taking advantage of esterase activity to generate the carboxylate from the methyl ester, the group has also utilised other types of ester groups for the *in vitro* testing of chemical probes without the need for additional enzymes such as pig liver esterase (PLE) to hydrolyse methyl esters. The main example is with ester groups which can be cleaved by photolysis with specific wavelengths of light: this has the advantage of efficiently cleaving the ester without the need for enzymes or specific chemical reagents and conditions to remove it. The probes can also be added to the culture medium and the carboxylate be generated *in situ* without significant disturbance to the bacteria.[154,155] The 4,5-dimethoxy-2-nitrobenzyl (DMNB) group was used due to its relatively high wavelength absorption to avoid damage to the bacteria which would occur using higher energy near UV photo-cleavable groups; this group can be used to protect groups such as amines, alcohols and in this case a carboxylic acid.[156] Upon excitation at around 350 nm wavelength of light, the photolabile moiety is proposed to undergo several radical bond cleavages releasing an aldehyde by-product (**414**) and the unprotected functional group (**415**) (**Scheme 4.10**).[157]
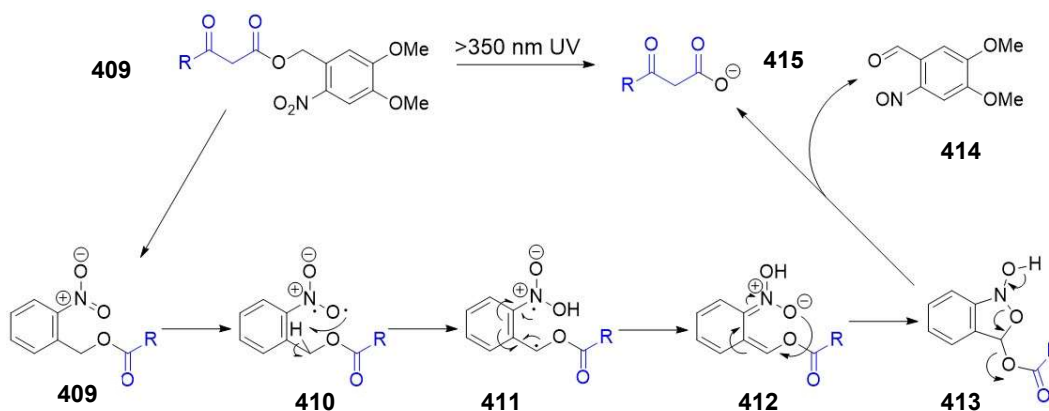
Scheme 4.10 – Proposed mechanism of photo-cleavage of the DMNB group (409) involving several radical and ion intermediates (410-413) leading to an aldehyde by-product (414) and the unprotected malonate (415).

163

Previously in the group, Samantha Kilgour showed that photocleavable ester groups could be added the chain terminator probe and photolysed *in vivo* in *S. lasaliensis* ACP12 (S970A) cultures unmasking the carboxylate to offload lasalocid A intermediates (**Scheme 4.11**).[158]
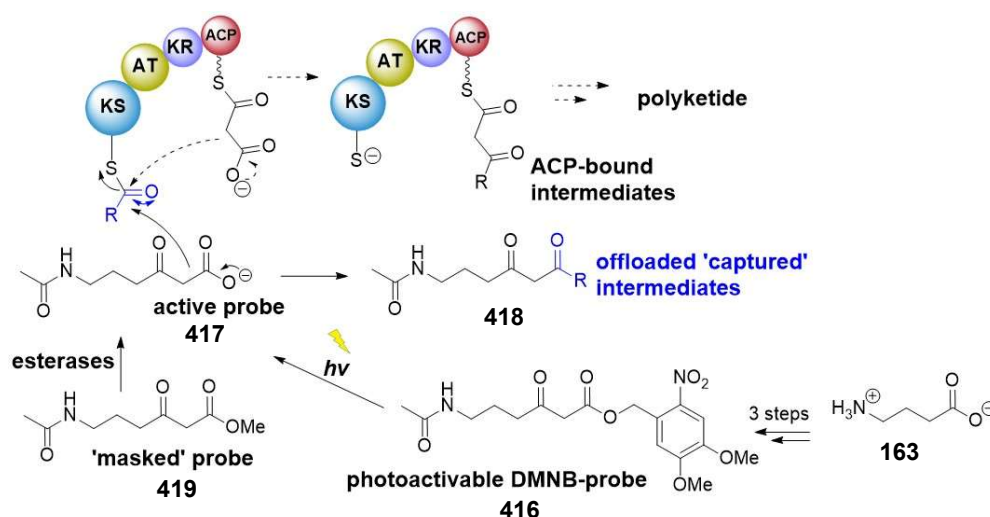
Scheme 4.11 – **Acetyl chain terminator probe with photo-cleavable ester (416) undergoes conversion to the carboxylate (417) upon irradiation and can then capture polyketide intermediates (418), this is unlike the methyl ester probes (419) which relies on esterase enzymes.**

This was achieved using a custom-built lightbox containing a circular 22W UVA lamp to surround the feeding cultures with the necessary wavelength of light to cleave the photolabile ester (**Fig. 4.9**). Building upon this work, each of the prior discussed methyl ester chemical probes were resynthesised as DMNB esters in order to test their efficacy.
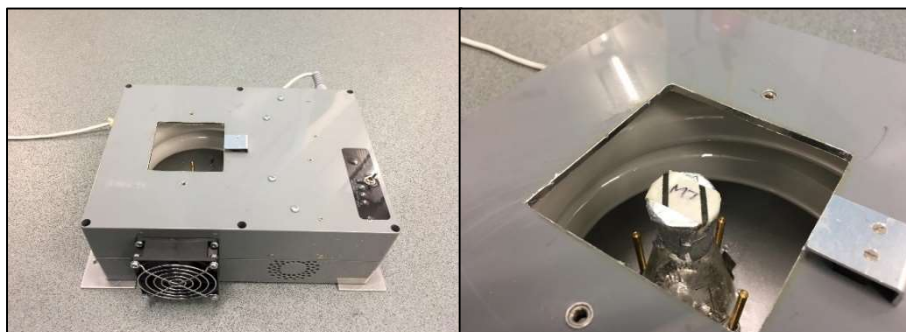
Figure 4.9 – **Image of custom-built lightbox (left), flasks fit inside box surrounded by 22W UVA lamp delivering light to cleave photolabile group (right).**

4.2.1 Synthesis of photolabile chemical probes with novel functional groups

Chemical probes with stable functional groups were synthesised *via* the Meldrum's acid adduct; the remaining probes were synthesised by a malonate potassium salt (**Scheme 4.12**) in a Claisen-like condensation fashion as previously shown (**Scheme 4.4**).
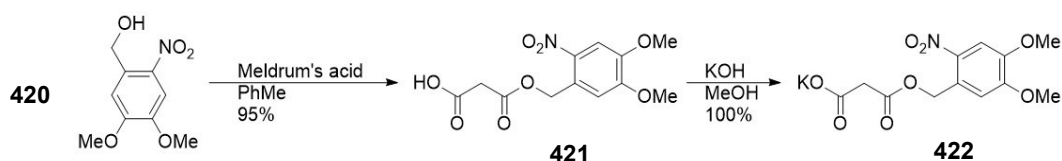
Scheme 4.12 – Synthetic route to photo-cleavable potassium salt (422)

As before with the methyl ester probes, it was envisaged that a potassium half malonate ester (**422**) could be used allowing access to the probes in a single step; this would be suitable for probes with a functional group which would not tolerate the other route through a Meldrum's acid adduct. The synthesis of this began with refluxing Meldrum's acid with the photolabile alcohol (**420**) where the carboxylic acid (**421**) was obtained in high purity and yield after filtration. This was neutralised with KOH and concentrated giving the potassium salt which was used without purification or characterisation (**Scheme 4.12**).

4.2.1.1 Synthesis of decanamido (423) and fluoro (424) photolabile probe
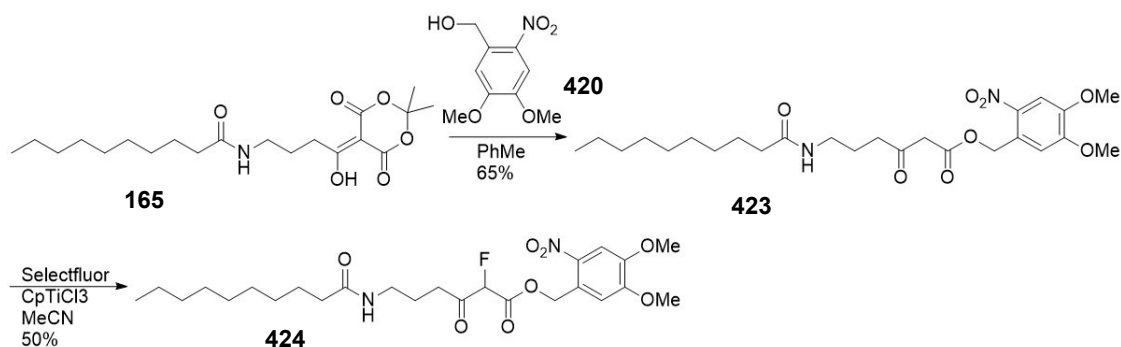
Scheme 4.13 – Synthetic route to decanamido (423) and fluoro (424) photolabile probe.

Following the route reported by Kilgour *et al.*,[158] this compound was synthesised and characterised by MChem student Ben Westwood. The synthesis began from the Meldrum's acid adduct (**165**) which was reacted in anhydrous conditions with the alcohol (**420**) giving the photolabile probe (**423**) in 65% yield after purification. This was then converted to the fluoro analogue (**424**) using Selectfluor in a 50% yield (**Scheme 4.13**).

4.2.1.2 Synthesis of azide photolabile probe (426)
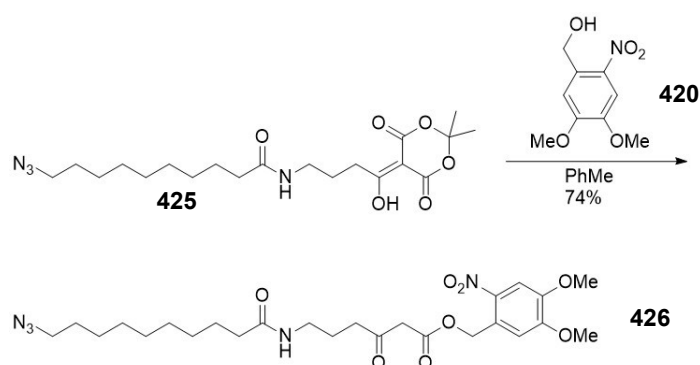
Scheme 4.14 – Synthetic route to azido photolabile probe (426).

The previously reported Meldrum's acid adduct with an azido group on the chain[83] (**425**) was refluxed with the benzyl alcohol (**420**) to deliver the azide containing photo probe (**426**) in 74% yield after purification (**Scheme 4.14**).

4.2.1.3 Synthesis of sulphonamide photolabile probe (427)
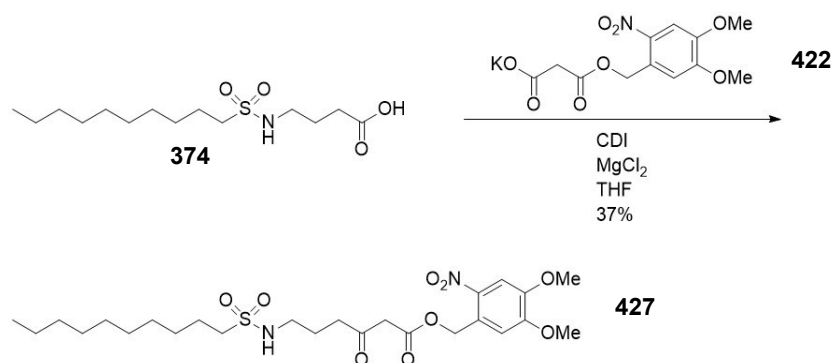
Scheme 4.15 – Synthetic route to sulphonamide photolabile probe (427).

166

The previously synthesised sulphonamide acid (**374**) was converted directly into the photoester product (**427**) with a one pot procedure using the malonate ester potassium salt (**422**) in 37% yield after purification. Despite the low yield, the crude product was fairly pure even before purification (**Scheme 4.15**).

4.2.1.4   Synthesis of azido fluoro photolabile probe (428)
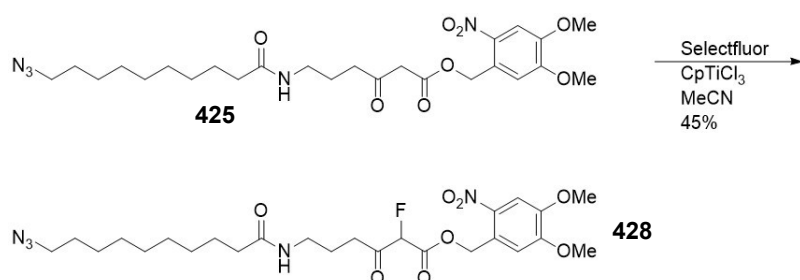
Scheme 4.16 – Synthetic route to azido fluoro photolabile probe (428).

The azide and fluorine containing probe (**428**) was synthesised in a single step from the azide photolabile probe (**426**). Fluorination using a titanium catalyst gave product in 45% yield after purification. This probe was not used in any feeding experiments as the compound showed instability and degradation during characterisation; this was noted in any probe containing both an azide and a fluoromalonate functionality (**Scheme 4.16**).

4.2.1.5   Synthesis of urea photolabile probe (429)
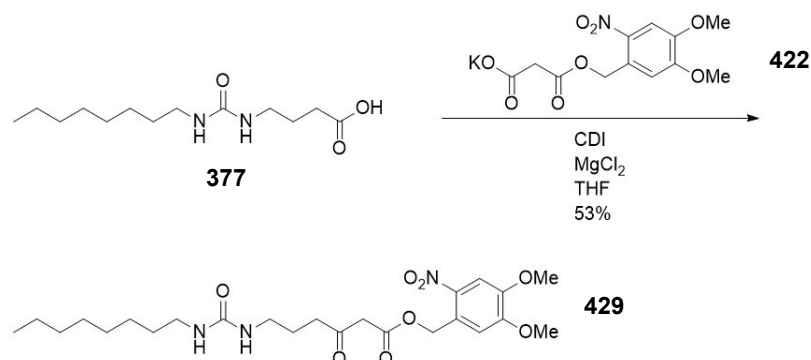
Scheme 4.17 – Synthetic route to urea photolabile probe (429).

167

The urea photolabile probe (**429**) was synthesised from the urea carboxylic acid (**377**) with the malonate ester potassium salt (**422**) in 53% yield (**Scheme 4.17**).

4.2.1.6   Synthesis of thiourea photolabile probe (430)
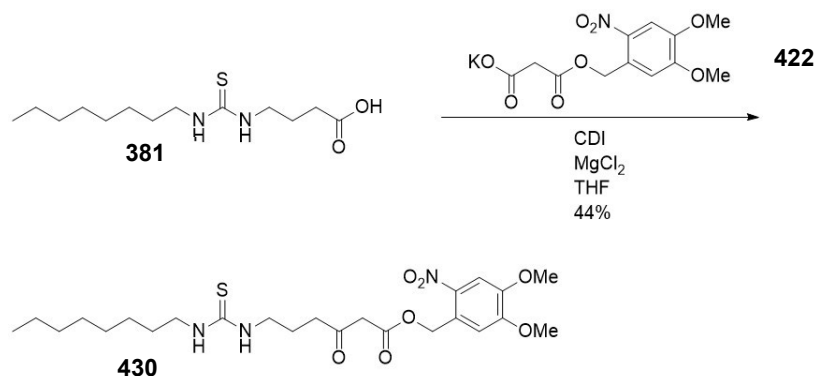
Scheme 4.18 – Synthetic route to thiourea photolabile probe (430).

The thiourea photolabile probe (**430**) was synthesised using the single step procedure from the thiourea carboxylic acid (**381**) with the malonate ester potassium salt (**422**) in 44% yield (**Scheme 4.18**).

4.2.1.7   Synthesis of thioamide photolabile probe (431)
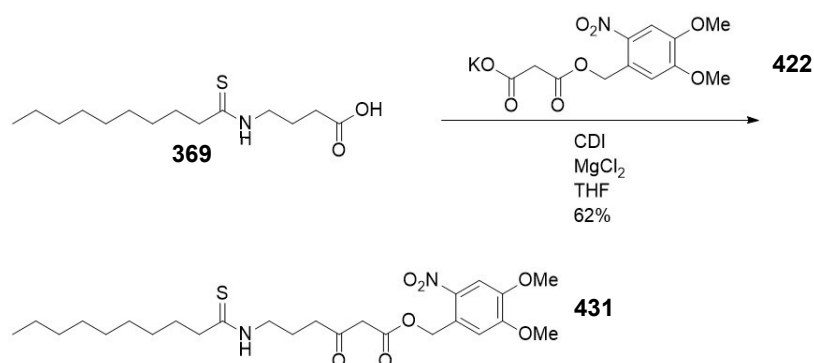
Scheme 4.19 – Synthetic route to thioamide photolabile probe (431).

The thioamide photolabile probe (**431**) was synthesised from the thioamide carboxylic acid (**369**) with the malonate ester potassium salt (**422**) (**Scheme 4.19**).

4.2.1.8   Synthesis of fluorophore photolabile probe (433)
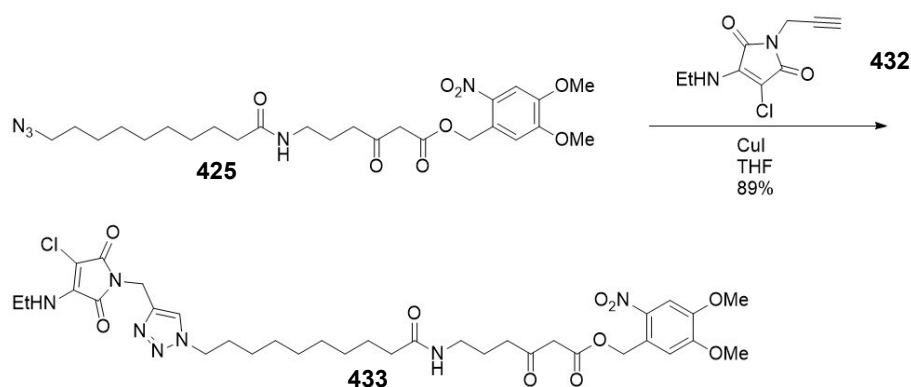
Scheme 4.20 – Synthetic route to fluorophore photolabile probe (433).

The fluorophore probe (**433**) was synthesised by a [2+3] cycloaddition of the azide photolabile probe (**425**) and a fluorescent maleamide alkyne (**432**), which was donated by the group of Professor Rachel O'Reilly,[159] and obtained in 89% yield after purification (**Scheme 4.20**). It was envisioned that using a fluorescent probe might allow isolation of captured intermediates by HPLC using UV-VIS detection in the visible light region.

4.2.2   Evaluation of photolabile probes *in vivo* with *S. lasaliensis* and *S. longisporoflavus*

Feeding experiments in *S. lasaliensis* ACP12 (S970A), *S. lasaliensis* WT and *S. longisporoflavus* WT were jointly performed by myself and by MChem student Ben Westwood in liquid culture form. Only the decanamido (**104**) photolabile probe was used in both these systems as the novel amide photolabile probes were synthesised at a later date. The decanamido probe was used to corroborate previously the detected

intermediates using the acetyl chain photolabile probe (**416**) which had shown success in the *S. lasaliensis* system.[158] All experiments were run with no probe and no photolysis controls and in duplicates.

4.2.2.1   Evaluation *in vivo* with *S. lasaliensis* ACP12 (S970A)

Preliminary feeding experiments using this strain involved supplementation of the photolabile probe (2 mM) and irradiation of it on the same day that the cultures were inoculated. It was observed that these cultures did not grow, likely due to induced toxicity whilst the seed cultures were still in early levels of growth. For this reason, all later experiments were carried out with probe addition and irradiation after one day of growth to allow the bacteria to reach a higher cell density. These repeat experiments were more successful and reduced toxicity was observed. Comparing the abundance of photolysed probe, probe carboxylate and decarboxylated probe in extracts showed that ester hydrolysis is still occurs moderately even in non-irradiated cultures. Photolysed extracts contained the deprotected and decarboxylated probe in higher amounts when photolysis was carried out but the ester still remained in moderate amount (**Fig. 4.10**).
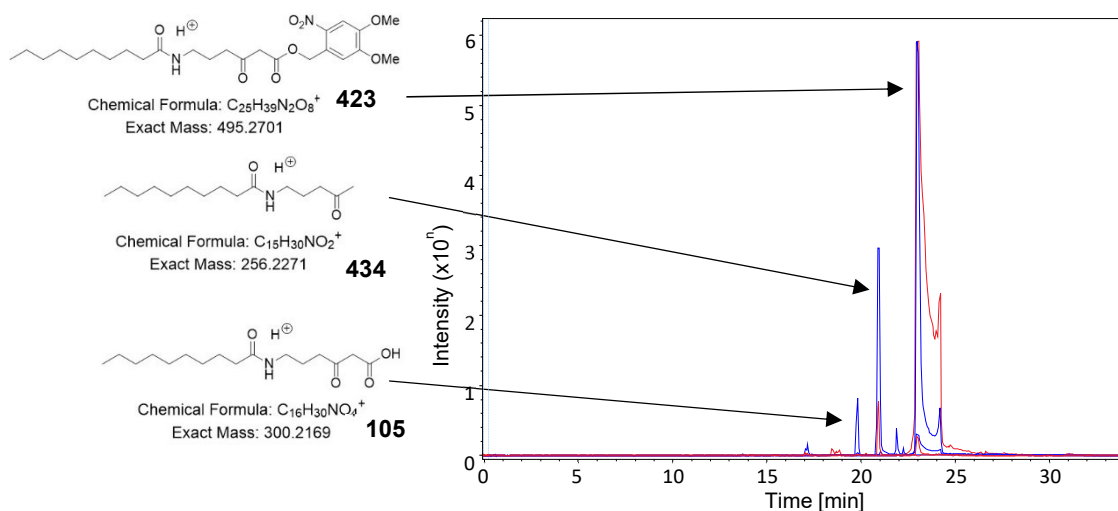
Figure 4.10 – **LC-HRMS traces showing extracted ion chromatogram for photolysed (blue) and non-photolysed (red) extracts. More decarboxylated probe (434) is detected in irradiated extract along with less intact probe (423) which is being consumed, lower abundance of hydrolysed probe (105) was seen in non-irradiated control.**

Unfortunately, no clear intermediates were captured from these experiments; this again is likely due to the *S. lasaliensis* ACP12 (S970A) strain behaving different to previously reported. The *S. lasaliensis* wildtype strain was then used, for which the production of lasalocid A was observed and confirmed that the *lasA* PKS was active.

4.2.2.2   Evaluation *in vivo* with *S. lasaliensis* WT

A similar level and proportion of intact, hydrolysed and decarboxylated probe was observed in the wildtype *S. lasaliensis* extracts. There was also moderate lasalocid A (**103**) production which was unaffected by the use of irradiation or addition of the probe, indicating no toxicity or disruption to the PKS (**Fig 4.11**).
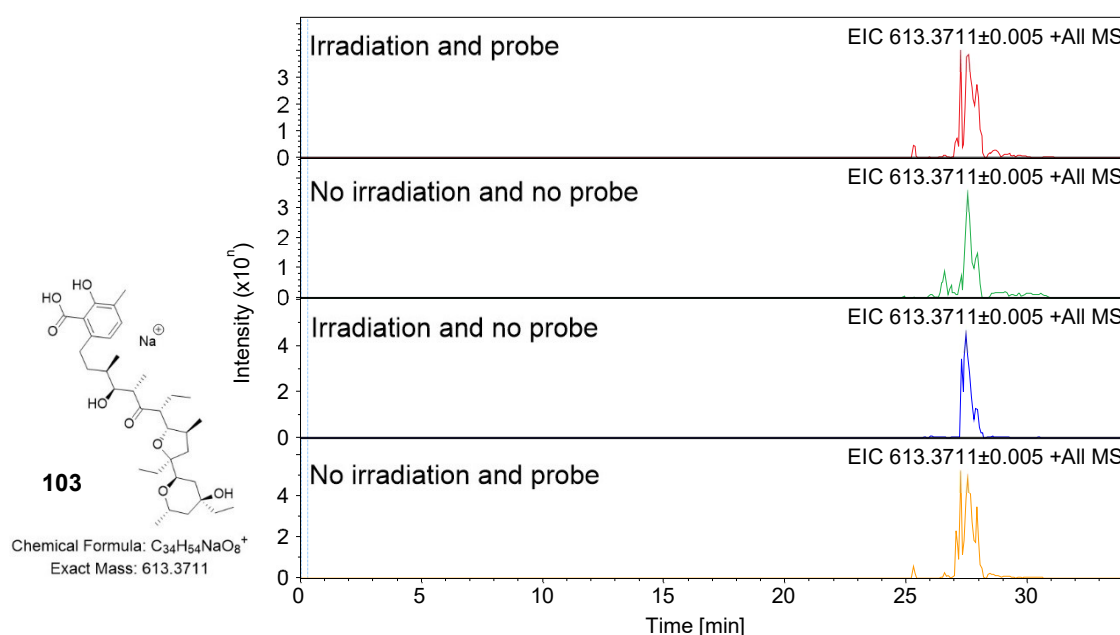
Figure 4.11 – **LC-HRMS traces showing extracted ion chromatogram for each combination of probe and irradiation extract, lasalocid A (103) production is moderate and fairly constant across all traces.**

Early polyketide intermediates such as di- and triketides were detected, although their pattern of abundance was different to that observed with methyl ester probes and the ACP12 (S970A) mutant strain. There were few late stage intermediates detected which were consistent to those detected with the *N*-acetyl photolabile probe (**417**) by Samantha Kilgour.[160] Also, the intensity of those intermediates was not high enough to gather MS^2 data so these could not be confirmed (**Fig. 4.12**). Despite this, none of the putative species were present in probe absent controls (not shown).
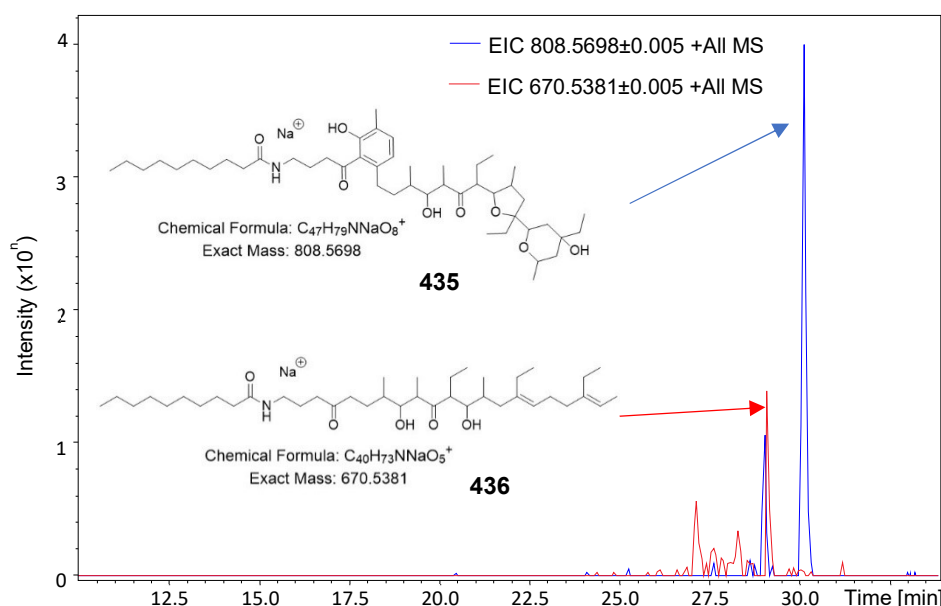
Figure 4.12 – **LC-HRMS trace (EIC) showing detection of two putative captured lasalocid A intermediates (435 and 436) consistent with those found using the acetyl photolabile probe (417).**

4.2.2.3 Evaluation of 423 and 424 *in vivo* with *S. longisporoflavus* WT

Probes **423** and **424** were also tested for intermediate capture from the *tsn* PKS in *S. longisporoflavus*. An array of putative polyketide intermediates was searched for, no confirmed intermediates from this pathway were detected with the photolabile probe. Comparable levels of intact and decarboxylated probe were observed to those observed in the *S. lasaliensis* experiments but no tetronasin production was observed: this likely

explains the lack of intercepted intermediates. The reasons for lack of tetronasin production remains unclear. Similar to experiments performed on *S. lasaliensis* ACP12 (S970A) mutant, cultures were given one day to grow to a higher density but perhaps *S. longisporoflavus* is a less robust strain. Overall, we observed that the strain grows more slowly than *S. lasaliensis*, so perhaps a longer period of growth is required prior to the use of the photolabile probes. Further experiments using photolabile substrates and further changing the conditions of their use were not performed due to lack of time.

4.2.3   Evaluation of photocleavable probes *in vitro* with minimal actinorhodin PKS from *S. coelicolor*

As mentioned in the introduction (**Section 1.4.5**), *S. coelicolor* produces the blue polyketide antibiotic actinorhodin using a type II iterative PKS: the final product is a dimer made from two identical biosynthetic intermediates (octaketides). The KS/CLF and ACP enzymes within the cluster (the minimal PKS) can together generate the actinorhodin shunt product SEK4 or SEK4b (**Scheme 1.16**).

The "minimal PKS" consisting of the KS/CLF and ACP enzymes has been previously reported and widely employed for studying polyketide biosynthesis *in vitro*. We used these enzymes as a model *in vitro* PKS system to test the offloading of polyketide intermediates by the novel photolabile probes. Quartz cuvettes containing each probe dissolved in DMSO were irradiated at 365 nm for 45 mins by a circular 22W UVA lamp to generate the active carboxylate probes *in situ* before their addition to *in vitro* PKS assays. These experiments were all carried out with the help of Panward Prasongpholchai (PhD student in the Tosin group), who provided purified recombinant KS/CLF and ACP.

TLC analysis was used to check for complete probe photolysis, which appeared to be around 45 mins; the cuvettes also began to turn yellow upon irradiation due to generation of the benzaldehyde by-product (**414**). Surprisingly, the already yellow fluorophore probe (**433**) turned red upon irradiation; this is likely caused by changes in the environment affecting the maleimide chromophore group (**Fig. 4.13**).

173
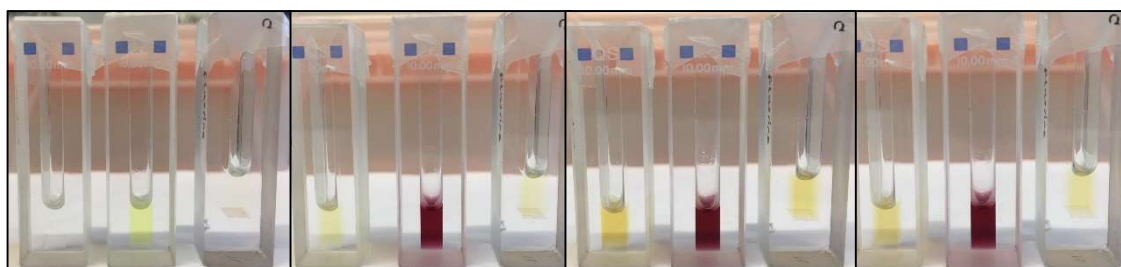
Figure 4.13 – **Photolysis of three probes in cuvettes at time = 0, 15, 30 and 45 mins (left to right). Probes in each vial are (left to right) thioamide (431), fluorophore (433) and fluoromalonyl (424).**

Various ratios of probe to malonyl-ACP were tested to maximise intermediate capture and monitor SEK4 and SEK4b production accordingly. Malonyl-ACP to probe ratios of 4:1, 1:1, 1:5 and 1:10 were tested; inhibition of SEK production was noticeable but was not significant even with 1:5 and 1:10 so it was decided that further assays were to be performed with the high concentration. Compared to *in vivo* testing, hydrolysed and decarboxylated probe species were present in abundance despite the TLC suggesting high conversion to the carboxylate (**Fig. 4.14**).
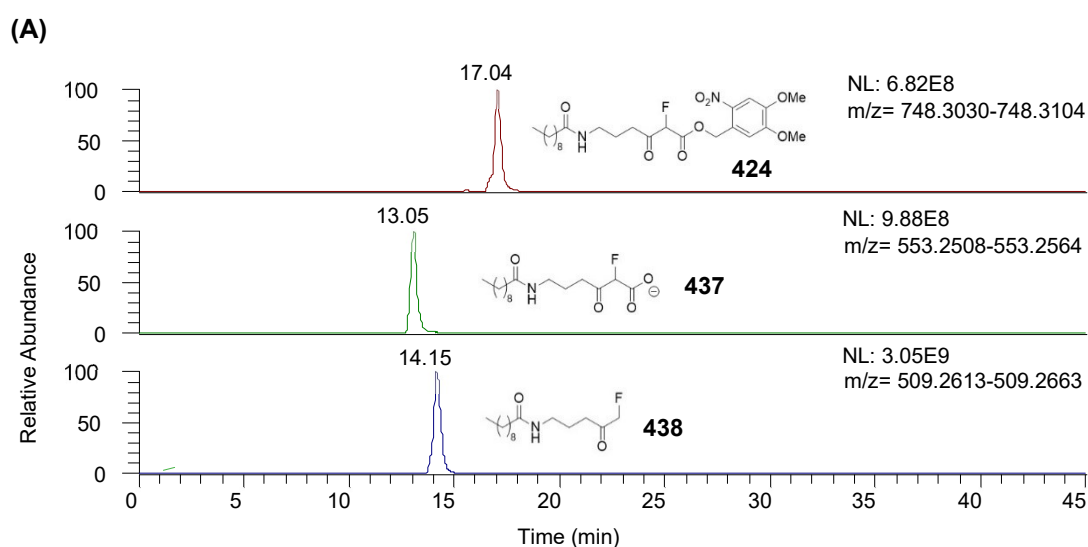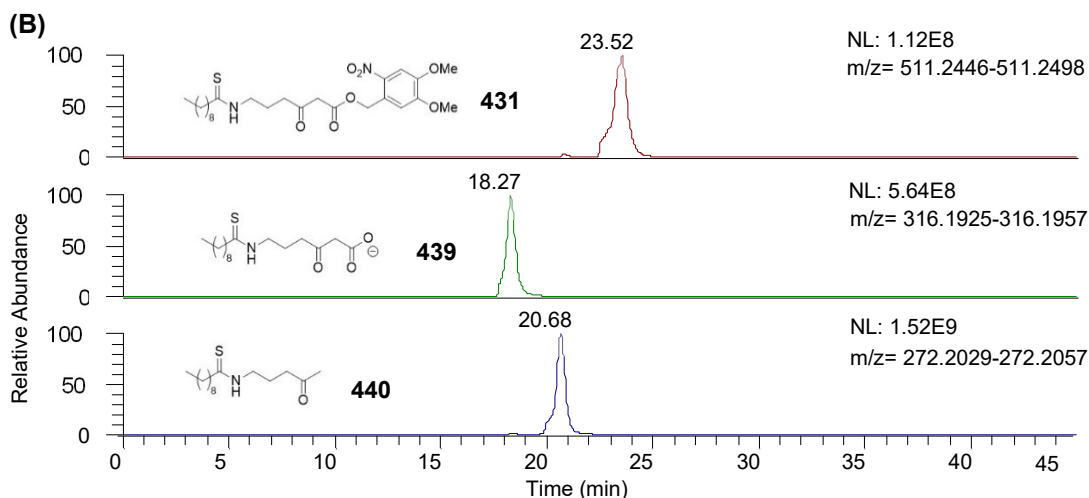
Figure 4.14 – **LC-HRMS extracted ion chromatograms showing intact probe (red), carboxylate (green) and decarboxylated probe (blue) for both (A): fluoro probe (424) and (B): thioamide probe (431). Much higher intensities of the latter two compounds compared to the intact probe indicate efficient photocleavage.**

All enzymatic assays for SEK4/4b were set up as detailed in **Experimental 6.3.4** and in duplicates. Briefly, photolysed solutions of probe in DMSO were incubated in various concentrations with the preformed *act* malonyl-ACP, *act* KS-CLF and malonyl-CoA in tris buffer at 30 °C for 16 h before extraction with MeOH and analysis by LC-HRMS. ACP SEK4/4b production was confirmed before captured intermediate search.

Initially for many of the probes, putative di, tri and tetraketides were observed; some but not all of which had supporting MS^2 fragments. They were also detected in boiled enzyme control experiments for which SEK4/4b compounds are not produced, suggesting that these species were not true biosynthetic intermediates. Experiments with probes **423, 431, 429** and **433** in SEK4/4b generating assays were repeated several times; one putative advanced intermediate not present in control assays with boiled enzyme was a putative pentaketide generated by the use of the urea photolabile probe (**429**). This was detected and characterized in both 1:5 an 1:10 *act* malonyl-ACP: probe ratio extracts with MS^2 diagnostic fragments (**Fig. 4.15**). Though the linear chain structures are shown, the intermediates are likely cyclised from undergoing aldol type reactions.
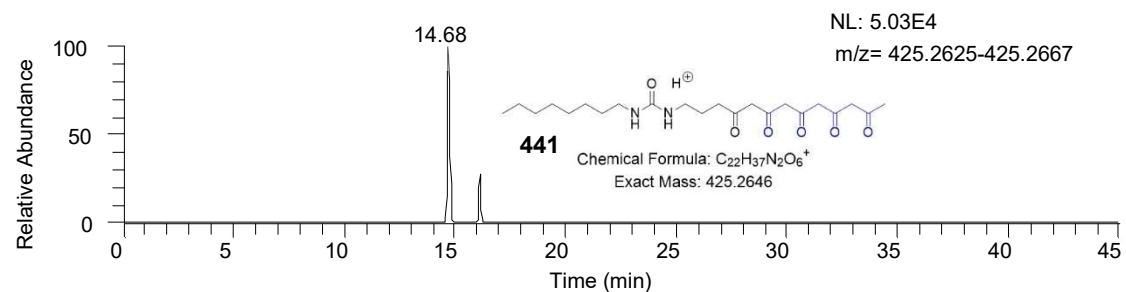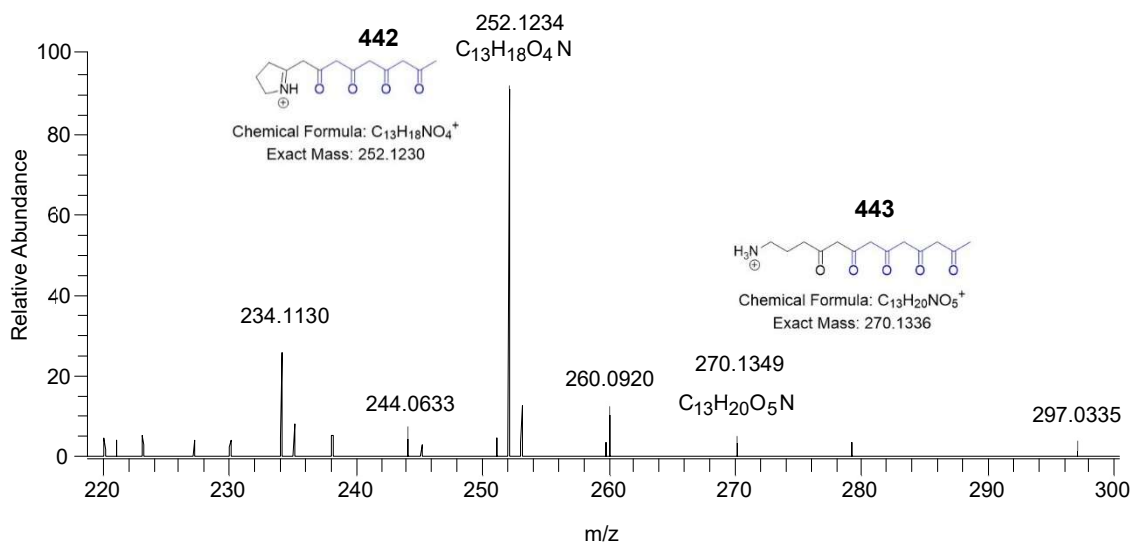
Figure 4.15 – LC-HRMS extracted ion chromatograms of putative pentaketide (441) captured by urea probe (Top) along with two diagnostic MS² fragments (442 and 443) (Bottom).

Similar pentaketide molecular weight intermediates were not observed in samples arising from other probes. In general type II iterative biosynthesis is very challenging to study *in vitro* due to the very rapid processing of intermediates. The detection of a pentaketide intermediate in these experiments is encouraging and consistent with other similar putative intermediate species observed by Samantha Kilgour in an ACP-bound form.[158,160] In addition to the tight working of discrete proteins to generate SEK4/4b, a bottleneck in the use of unnatural substrates is likely the tight specificity of the PKS in substrate recognition. This will be further discussed and addressed in Chapter 5 (future work).

176

4.3 Summary of results

A series of chemical probes with modifications centred around the amide group of **104** were prepared for offloading intermediates in several modular and iterative PKS systems. The idea was to probe PKS substrate specificity and the extent to which unnatural substrates can be used to generate unnatural intermediates or products.

Initially all probes were tested in *E. coli* liquid cultures heterologously expressing the iterative type I PKS 6-MSAS, for which species capture has been previously reported by our group. Preliminary results showed poor probe hydrolysis to the active carboxylate and unconvincing evidence of intermediates intercepted from the pathway of interest.

The lasalocid A PKS from *S. lasaliensis* was also utilised as a model system for intermediate capture in type I modular PKS assembly. Solid culture experiments showed strong toxicity of certain probes (such as the selenourea **364** and thiourea **363**) towards this strain. Very few putative unnatural intermediate species were detected and most of the captured species corresponded to early stages of the biosynthesis. Parallel experiments with *Lentzea* sp. (the producer of thiolactomycin *via* a hybrid iterative PKS-NRPS) were not more successful. Toxicity was observed for all sulphur and selenium containing probes. The extent of probe hydrolysis and decarboxylation on solid culture with *S. lasaliensis* and *Lentzea* sp. were equally poor, indicating that these new substrates present a different bioavailability for cellular uptake and/or enzymatic processing.

Almost all of the novel probe bearing amide substitutions were re-synthesised as photo-cleavable esters in the effort to improve the hydrolysis efficiency and therefore the abundance of captured species. From *in vivo* liquid cultures of *S. lasaliensis* and *S. longisporoflavus,* a few putative intermediates species were detected but lacked clear MS^2 diagnostic fragmentation. *In vitro* experimentation of the same substrates subjected to quantitative photolysis prior to their use with type II minimal *act* PKS enzymes (generating the octaketides SEK4/4b) also proved challenging, with a possible putative pentaketide intermediate intercepted by the urea probe (**429**).

5 <u>Conclusions and further work</u>

5.1 Investigation of thiotetronate biosynthesis

In this work, the study of thiotetronate biosynthesis has focused on two aspects: the use of non-hydrolysable chemical probes for intermediate capture and biosynthetic mechanism/timing elucidations, and the use of hydrolysable SNAc substrates to investigate PKS substrate specificity and flexibility towards the possible making of unnatural products. Both lines of inquiry have offered some clear conclusions as well as further opportunities for future investigations.

5.1.1 Chemical probing of early stage thiotetronate biosynthesis

As outlined in the introduction, the use of chain terminator probes has been previously useful for investigating a variety of PKSs of bacterial and fungal origin. In the investigation of thiotetronate assembly with these tools, *in vivo* (*Lentzea* sp. strains) intercepted intermediates[134] ranging from diketides to pentaketides consistent with proposed ACP-bound intermediates after each round of chain extension, reduction and dehydration have been detected and characterised by LC-HR-MS[n].[104]Also demonstrated was the importance of hydrolysis rate of the probe methyl ester: more abundant intermediates were offloaded with the fluoromalonyl probe (**160**) than the methyl malonate probe (**161**), despite the latter being a closer chemical mimic to the natural methyl malonate substrate for chain extension.

 Putative tetraketide species allegedly containing sulphur were also detected, which may indicate the timing of S insertion (see following section). Crucial to the corroboration of all the *in vivo* results was the complementary data gathered from the use of recombinant TlmB protein *in vitro*: virtually identical intercepted polyketide species were identified and characterised. So far, these species have been characterised only by LC-MS: future

investigations could include scaling up of experiments to attempt species purification and characterisation by NMR. Also, the preparation of a chiral chain terminator probe which would produce diastereomers when capturing PKS intermediates containing stereocentres would be helpful: this may allow the stereoselectivity of PKS processing to be better understood.

The most intriguing part of the investigation concerned the attempted elucidation of the late stages of thiotetronate biosynthesis, focusing on how a small number of enzymes can be involved in the formation of the complex thiolactone moiety. The presence of two essential sulphur transfer related proteins, TlmJ and TlmS, in *Lentzea* sp. and also the putative monooxygenase TlmD1 makes the epoxide/thiirane proposal (**Section 2.3.5**) extremely attractive,[93] however for a similar cluster in *S. pacifica* there are no essential sulphur transfer genes[94] which raises questions on the true function and relevance of TlmJ and TlmS. Many extracts deriving from supplementation of chain termination probes to *Lentzea* sp. WT strain appeared to show a variety of intermediates which may derive from epoxide and/or thiirane intermediate ring opening. Tetraketide and pentaketide stage intermediates deriving from epoxide/thiirane moieties were found, with the latter resulting from direct offloading of nascent tetraketide species. Regarding S and O containing tetraketide species, we can speculate that either epoxidation and/or thiirane formation occurs at triketide level whilst PKS-bound, r that O/S insertion occur post-intermediate offloading. At this point no further *in vivo* experiments were able to provide more insights. Essentially all *in vivo* experiments carried out with point mutants of the *tlm* cluster heterologously expressed gave no detectable intermediates which may have offered clues on the exact role of domains in the NRPS region such as the Cy and TE.

The data gathered through the *in vitro* experiments by Marie Yurkovich was also corroborative of what has been observed *in vivo;* this was crucial given the ambiguity of MS^2 fragmentation acquired for certain species detected *in vivo*. The same putative tetraketide and pentaketide were detected but of clear importance was their presence in TlmB assays where TlmD1, TlmS and TlmJ were not utilised. This strongly suggests that sulphur insertion from cysteine into the polyketide chain may be independent from genes/proteins that were initially thought to be essential. One could envisage a Michael-

like addition of the cysteine thiol onto the conjugated polyketide intermediate followed by elimination to release de-hydroalanine; supplementation of *Lentzea* sp. ΔCy strain with the cysteine mimic probe (**242**) was hoped to shed light on this being the case, however no clear evidence was gathered.

The proposal for the formation of an epoxide intermediate remains to date experimentally unverified. Indeed, incubation of the SNAc triketide (**192**) and tetraketide (**207**) substrates with recombinant TlmD1 did not lead to the formation of a new species. Later on, an ACP-bound tetraketide was shown not to be processed by recombinant TlmD1 either. It is possible that for some reason the recombinant TlmD1 is not active, however the collective evidence from our studies and others[94,138] point towards a different role for TlmD1.

5.1.2   Use of SNAc substrates for biosynthetic elucidation and manipulation

As highlighted in **Section 2.4.1**, many SNAc substrates were prepared for the purpose of biosynthetic elucidation and also for attempted thiotetronate analogue production. With exception of the acetyl-CoA mimic (**199**), all other putative "natural" intermediate substrates of thiolactomycin (**192**, **201** and **207**) were unequivocally shown to be processed *in vivo* by the Δ*tlmA Lentzea* sp. strain to restore thiolactomycin production, albeit to differing extents. This is the first evidence of 'unnatural' thiolactomycin generated by a chemoenzymatic approach employing synthetic substrates and engineered cells as biosynthetic factories. The ineffectiveness of **199** to provide starter unit is not surprising for three reasons: 1) the true starter unit for thiolactomycin biosynthesis is malonate which TlmA converts to an acetyl; 2) the lack of a functional TlmA protein may result in lost protein interactions with TlmB which are essential for loading of the acetyl; 3) the abundance of acetyl-CoA in many metabolic pathways likely means that unnecessary priming of the PKS would occur without strong substrate specificity.

As well as showing that TlmB is capable of processing synthetic polyketide mimic SNAc substrates, this also confirms that the tetraketide **207** is a clear mid/late stage intermediate

in thiolactomycin biosynthesis. TlmB showed a strong substrate specificity: even substrates without any huge electronic or structural differences (such as **263**) were not processed by it *in vivo*. Indeed, for only a single non-natural substrate there was evidence of processing: this was for the non-methylated tetraketide SNAc analogue (**216**). The lack of supporting MS^2 data gathered means that this unnatural processing cannot be confirmed with absolute confidence, despite the retention time of the putative non-methylated thiotetronate generated being consistent with naturally occurring desmethyl TLM. It was noticed by NMR analysis that this particular substrate presents considerable enol tautomerisation, therefore it may electronically not be similar to the native methylmalonate substrate despite being very similar structurally. This same point also applies to the fluorinated tetraketide substrate (**262**). In any case, there seems to be few avenues left to pursue in terms of chemoenzymatic synthesis of novel TLM analogues based on the inability of TlmB to process unnatural substrates shown by my experiments.

No putative epoxide or thiirane related species were found which would indicate the steps involved in conversion of these SNAc substrates into thiolactomycin. Looking further at the alternative mechanistic proposal of the P450's involvement in forming the 5-membered ring from a thiocarboxylic acid tetraketide,[138] preparation and testing of the trifluorinated diketide (**226**) gave two detectable species that would support such a mechanism. As mentioned in **Section 2.3.5.2**, there was reasonable MS^2 detection at the retention times in question but no sense could be made of the fragment masses. This could be due to the thiolactone moiety fragmenting in an unpredictable manner. Nonetheless, the main caveat of these experiments is that they still don't confirm the envisaged epoxidation, S insertion and ring closure. A possible method for further investigating this would be to prepare the entire tetraketide SNAc and pantetheine substrate with a terminal CF_3 group in order to test them with TlmD1 *in vitro*.

Experimentation of the SNAc TLM intermediate mimics with other thiotetronate strains followed the paradigm of high substrate specificity and seemingly few supplementation extracts showed detection of putative novel compounds. A big issue of working with the Tü3010 cluster containing *S. olivaceus* and the *S. avermilitis* heterologous expression host was the poor distribution of substrates on plate. Rather than droplets of substrate in MeOH

spreading evenly across the plate upon application, as seen in *Lentzea* sp., the drops remained in single points on the plate leading to pockets of high concentration. Many of these plates had zones of inhibition growth around the evaporated droplet areas which may explain the lack of any detectable novel thiotetronates. Despite this, an intriguing pattern emerged for the supplementation of triketide (**192**) in the *S. olivaceus* and *S. avermilitis* extracts. Levels of Tü3010 were severely diminished in *S. olivaceus* by over a thousand-fold.

The lack of any novel thiotetronate being produced seems to suggest that the PKS machinery has being hijacked but is then stalling, with either chain extension or some of the tailoring steps involved in the carboxamide moiety installation unable to act on the unnatural substrate (likely enzyme-bound); this would explain the dramatic decrease in Tü3010 titre. The parallel supplementation of the same triketide SNAc substrate to the heterologous host *S. avermilitis* containing the *stu* cluster without *stuA* showed a species with *m/z* 254 consistent with the formation of the expected unnatural thiotetronate desmethyl Tü3010 (**286**) in very low amount. This result seems to support the idea that, in the absence of natural triketide substrate, StuB can process unnatural triketides, albeit to a much lower extent. Feeding experiments performed with tetraketide SNAc substrates **207** and **263** in *S. olivaceus* WT showed slightly diminished production of Tü3010 and possible *m/z* species attributable to thiolactomycin and des-methyl thiotetromycin; *S. avermitilis:*Stu Δ*stuA* was fed with the same substrates and similar species were identified but with both strains there were no clear supporting MS^2 fragments. If these detections are genuine, it suggests that the unnatural SNAc substrates can be processed by StuB to generate new species, however to a much lower extent and more likely incorrectly in terms of ring size and/or stereochemistry.

In the future it would be valuable to scale up these experiments and attempt the isolation and thorough characterisation of these novel species. Also, crystallisation of TlmB and StuB proteins/domains with surrogate substrates would prove useful to shed light on the factors that dictate substrate specificity and tolerance at different stages of polyketide chain extension and processing.

5.1.3 Synthesis of thiotetronate analogues and *in vitro* investigation of thiolactomycin machinery

As outlined in **Section 2.4.4**, an attempt was made to prepare synthetic thiolactomycin analogues for use as internal standards for LC-HRMS extract comparison, as well as for antimicrobial testing; recent work reported in the literature has indeed shown that thiotetronate ring modifications have a dramatic effect on potency.[140] Despite the preparation of many synthetic thiolactomycin analogues in the literature (mainly with modifications on the alkyl sidechain),[108] there are very few examples of modifications directly on the thiolactone moiety, especially of non-alkyl groups. Attempted synthesis of the des-methyl and fluoro thiolactone ring structures (**286** and **295)** using well documented conditions for cyclisation gave no clear products. There may be some inherent instability issues for these rings, or perhaps the electronics of the substrate are not suited for this type of cyclisation. This may explain the lack of processing observed with the unnatural tetraketide SNAc substrates **216** and **262**.

In vitro manipulation of the TlmB ACP also proved difficult. Pleasingly, the conversion of the recombinant *apo*-ACP to a *crypto*-ACP loaded with a tetraketide substrate was successful. Unfortunately, none of the assays performed with this species and recombinant TlmD1, Cy domain and a chemical surrogate of PCP-bound cysteine (**242**) yielded further insight into the ACP-bound species processing including sulphur insertion and cyclisation. It is possible that the recombinant enzyme utilised are incorrectly folded and therefore are not functional, or simply that we need further comprehensive assays including TlmD1, TlmJ, TlmS and intact TlmB, ideally with ACP-bound and PCP-bound substrates, to observe clear substrate processing. (e. g. by intact protein MS) Preparing the earlier mentioned trifluoro tetraketide analogue **227** and loading onto the *apo*-ACP could be a desirable avenue to pursue. Any chemical transformations would likely lead to a stalled intermediate to allow easier MS detection in a protein-bound form, and fluorine NMR could also prove useful for reaction monitoring

Overall, thiotetronates appear to be constructed in a way unlike any other known sulphur containing natural product. Much has been learned about substrate processing in

183

thiolactomycin biosynthesis through this work. Success has been achieved in determining unequivocally the nature of the linear polyketide intermediates involved in its carbon skeleton formation, however there are still several steps which remain cryptic. TlmD1 appears to be involved in S insertion in a way different from what originally hypothesised. So far, no direct evidence of polyketide chain oxidation by TlmD1 has been found. According to the recent study by Moore and co-workers, TlmD1 appears important in dictating ring formation size, although the mechanistic basis of its workings remain obscure. TlmD1 is predicted to be a cytochrome P450 enzyme by BLAST searches (see **Appendix 7.4**). A Phyre2 search has predicted structural similarities with several P450 enzymes involved in the hydroxylation of diverse natural products, but also with a putative epoxidase (**Appendix 7.2**) and a couple of other enzymes involved in C-S bond or S containing product formation (*e.g.* Sgvp, a thioether synthase in griseoviridin biosynthesis, and LnmZ, a hydrolyase involved in leinamycin biosynthesis). An interesting possibility to consider is that, as a putative P450 enzyme, TlmD1 might be controlling the oxidation state of the sulphur atom ultimately inserted into thiolactomycin, or it might tailor the nature of the S containing species responsible for C-S bond formation. Ultimately, in all possible scenarios envisaged for thiotetronate formation so far, the formation of a new C-S bond must occur.

Going back to our attempts to characterise shunt products of the biosynthetic pathway, the *Lentzea* sp. Δ*tlmD1* mutant grown on plate appeared green in colour (**Fig. 2.7**), which may be indicative of excess sulphide species present. Also, from *in vitro* experiments with recombinant enzymes (**Section 2.3.3**) it has been observed that: 1) the titre of putative intercepted tetraketide and pentaketide species allegedly containing S as thiols or thiiranes (**Fig. 2.12**) and detected in recombinant TlmB assays alone significantly decreases (to almost depletion) in the presence of TlmD1; 2) the titre of these species is restored and actually increased when TlmJ and TlmS are also present with both TlmB and TlmD1 (observations by Dr Tosin, personal communication). Assuming that the characterised species are effectively intercepted biosynthetic intermediates, these unexpected finding suggests that: 1) TlmB is mainly responsible for both carbon chain formation and S insertion; 2) the recombinant TlmD1 is somehow active and works with

other enzymes to process of S species for C-S bond formation. So far, the reconstitution of thiolactomycin production *in vitro* and the function of TlmD1 have remained elusive. It's possible that perhaps not all key proteins interacting with TlmD1 have been considered/investigated. Recent analysis by Dr Tosin of a biosynthetic gene cluster in *S. cattleya* encoding for thiotetronate biosynthesis (U-68204) has revealed the presence of genes/proteins associated to sulphur processing; in particular, SCATT_57700 encodes for a putative pyridine nucleotide disulphide oxidoreductase (**Fig 5.1**).
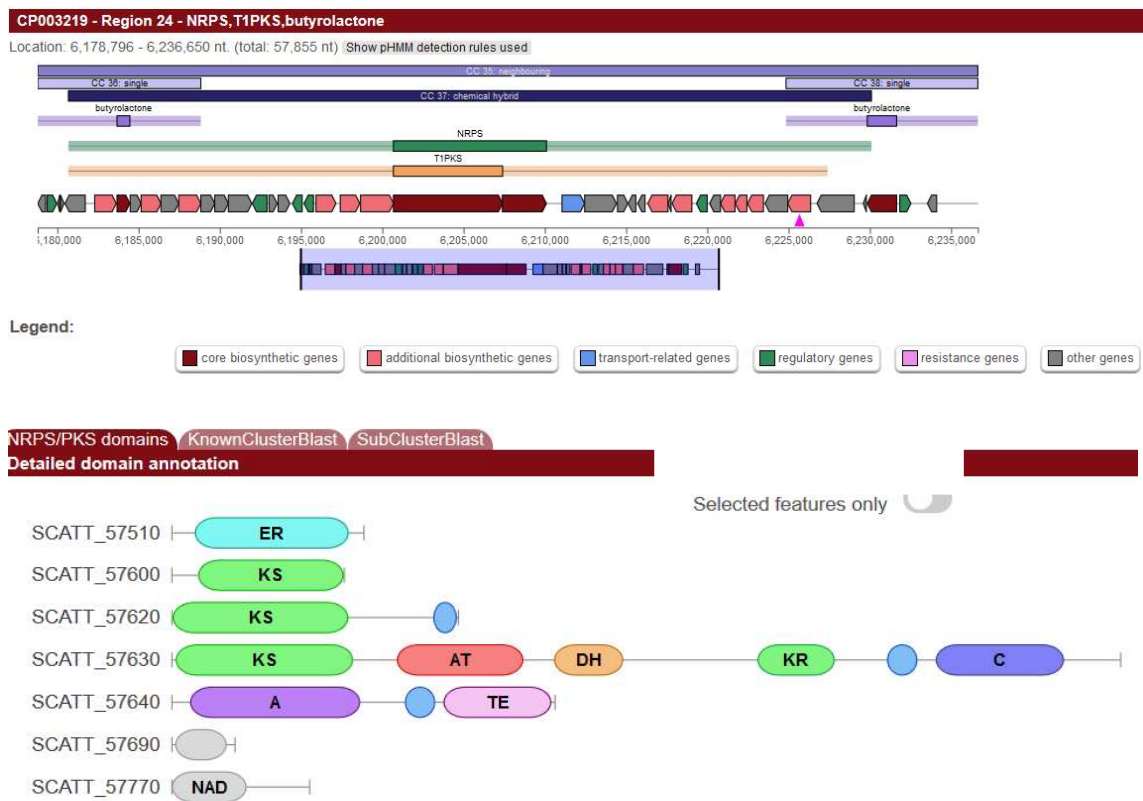
Figure 5.1 - BGC within *S. cattleya* for a putative PKS-NRPS for biosynthesis of a thiotetronate. SCATT 57770 encodes a putative sulphur processing enzyme.

Examples of proteins of this class include disulphide forming enzymes found in the biosynthesis of depsipeptide anticancer agents such as FK228, and of dithiopyrrolone

antibiotics such as holomycin. Closer inspection of other genes close to SCATT_57700 revealed SCATT_57690, encoding for a putative cyclase, and SCATT_57670, a DbsA oxidoreductase involved in disulphide formation. The *S. cattleya* cluster is also peculiar in that it harbours genes responsible for sugar diphosphate nucleotide biosynthesis and processing, including glycosyl transfer; hence we cannot exclude that the aforementioned proteins associated to sulphur processing are related to sugar rather than polyketide/nonribosomal peptide processing. At the light of this information, new hypothesis concerning the late stage events of thiotetronate formation and tailoring remain to be explored and are object for current investigation in the Tosin group.

5.2 Investigation of tetronasin biosynthesis

The comprehensive study of tetronasin assembly has been a collaborative effort with Rory Little (Cambridge) and Prof Marcio Dias (Sao Paulo), with each group focusing on different aspects; the collection of our efforts has been recently been published (**Appendix 7.5**).[141] My work has focused on the use of chemical probes to gather insights on the mechanism and timing of ring formation in tetronasin assembly. It became clear immediately that the production of tetronasin was unaffected by the culture medium. This is in contrast to that on *Lentzea* sp. where growth in liquid has a highly detrimental effect of thiolactomycin production. The addition of the decanamido probe (**104, Fig. 3.2**) or fluoromalonyl probe (**160, Fig. 3.2**) in any concentration drastically reduced the production of tetronasin; the extent of growth was also diminished indicating a moderate toxicity level. This may explain the lack of putative intermediates detected from the tetronasin assembly line. Aside from a few high abundance species, few to none were observed for early stage intermediates. Previously reported experimentation[85] with the similar type I modular PKS from *S. lasaliensis* showed a vast range of captures ketide species from all modules. The *S. longisporoflavus* assembly line has proven much less susceptible to polyketide capture which is either a result of the toxicity of maybe the efficiency of tetronasin production preventing probe interception.

The few intermediates detected were all from extracts of liquid cultures where the probe was added in small quantities periodically rather than being readily available on plate. The extent of probe hydrolysis did not change vastly from solid to liquid culture, this highlights the more efficient method of probe introduction is in liquid where toxicity can be minimised. The two putative heptaketide species (**323** and **324**) detected support epoxidation and THF ring formation; the fact that no other intermediates of similar elongation extent were found suggests that heptaketide formation may be rate limiting and interception can occur more readily. Without clear diagnostic MS^2 fragments, this remains a speculative proposal.

The only other putative intermediate detected was an undecaketide (**328**) which seemingly has undergone hydroxylation and *O*-methylation as well as the previously mentioned epoxidative THF ring formation. Several of the MS^2 fragments matched structures which could feasibly result from the proposed undecaketide fragmenting in several ways. Based on this species' proposed structure, THF ring formation is thought to precede other ring formation and likely occurs when the nascent polyketide chain is still bound to the PKS; hydroxylation and *O*-methylation may be also occurring on a PKS-bound intermediate and can again precede the remaining ring formation. In contrast to the investigation of THF ring formation, little was further understood about the timing of tetronate ring formation. There is already a reasonable understanding of tetronate biosynthesis learned from other natural products pathways, which involved the condensation of a glycerate group with a malonate. Unfortunately, the glycolate probe (**342**) prepared and tested in *S. longisporoflavus* offered no novel insight into the timing of the ring formation. The probe may not have the right bioavailability to reach the biosynthetic enzymes, or simply it may not fit into the enzyme active site. Also, even if intermediates are captured, the ester linkages generated may hydrolyse *in vivo* (although no evidence of this was found) (**Scheme 5.1**). A probe designed to allow a tetronate ring to form may offer more success, this probe could include modifications to allow more distinctive fragmentation and therefore better characterisation such as **414** as a simple example. Perhaps Cl or Br containing probes could be incorporated to give a distinct isotope pattern or even a "clickable" handle for capturing all probe related products for cleaner LC-HRMS profiles.
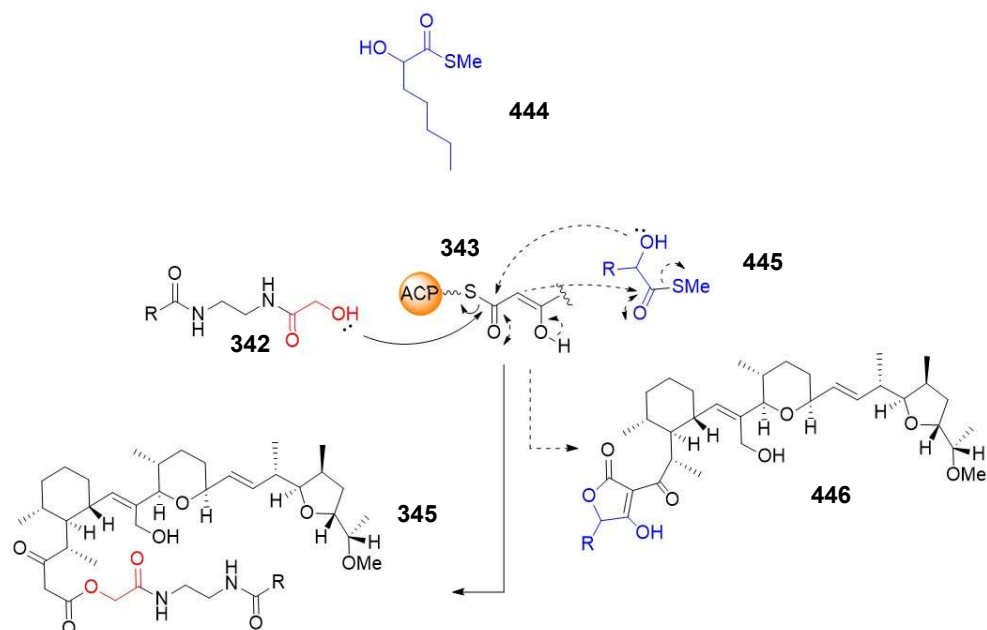
Scheme 5.1 – **(Solid arrows) Interception of a tetronasin late stage intermediate (343) with the glycolate probe (342) would lead to an ester linked product (345) which could be hydrolysed. Formation of the tetronate ring may be crucial for the biosynthetic enzyme machinery to function. (Dotted arrows) A probe designed to allow tetronate ring formation such as 444 would likely be more suitable and lead to the more stable tetronate adduct (446).**

Further experiments with glycine probe **356** were hoped to lead to more stable and thus detectable intermediates owing to the amide linkage which would be less susceptible to hydrolysis. Unfortunately, this also showed no evidence of intermediate capture for reasons likely similar to that of the glycolate probe (bioavailability and substrate recognition). Also, a probe based on glycerate has been prepare for future investigation of tetronomycin biosynthesis. Probes **342**, **346** and **347** should be valuable tools for further *in vitro* studies with recombinant enzymes since there would be no issue in their bioavailability. Attempted co-crystallisation of the putative epoxidase and epoxide hydrolase by Marcio Dias with the truncated substrates I prepared (**355** and **356**) is still in progress.

5.3 Design and testing of novel chemical probes *in vitro* and *in vivo*

5.3.1 Preparation and testing of novel methyl ester probes

As described in **Section 4.1**, a variety of chemical probes bearing amide functionality changes were prepared and tested in the *in vitro* PKS assays (*act* minimal system) and *in vivo* with the PKSs responsible for producing 6MSA, thiolactomycin and lasalocid A respectively. Throughout all systems utilised, the efficiency of hydrolysis was considerably poorer than that seen with probes bearing a simple amide moiety; either the unnatural functionality is not recognised well by the esterase enzymes within the microorganisms, or the change of amide moiety influences the bioavailability and distribution of the probe within cells. Some of new compounds also displayed toxicity towards microorganisms, in particular the selenium and sulphur containing probes. This could be due to sulphur groups being inherently toxic or perhaps from degradation or interference with S/Se metabolism; extracts of the thiourea probe (**363**) for example indeed showed the presence of a urea by-product and possible release of H_2S.

The poor hydrolysis of these probes along with the toxicity issues seems likely the reason for poor capture of intermediates *in vivo*. A direct comparison could be made in the case of *Lentzea* sp. and *S. lasaliensis* where the decanamido probe (**104**) or fluoromalonyl probe (**160**) supplemented extracts showed a clear range of PKS intermediates.[85,134] The diversity and abundance of species was unparalleled with any of the novel probes in question. Some odd results were obtained in the 6MSA feeding experiments where preliminary data suggested that possible 6MSA putative intermediates were detected, however the same species were detected in *E. coli* extracts containing no 6MSA plasmid, suggesting that these species may derive from other PKS/FAS pathways. Overall, modification of the original amide probe did not lead to improvement in the ability to capture PKS species and/or generate unnatural products.

5.3.2 Preparation and testing of probes with photocleavable esters

The synthesis of novel chemical probes as photocleavable esters was hoped to resolve issues of poor substrate hydrolysis and aid intermediate capture. These compounds were prepared in a novel way from a photolabile malonyl potassium half ester (**422**), this was adapted from a well reported procedure to make β-keto methyl esters from the respective methyl malonyl potassium half ester. The method proved to be extremely mild and tolerated groups such as the thiourea and thioamide well.

Despite lengthy irradiation of plates and liquid cultures, comparison with the non-irradiated samples showed mildly improved ester cleavage. A higher abundance of hydrolysed and decarboxylated probe was detected but not at significant levels as was expected. This is most definitely due to the poor penetration of the light through the glass culture flasks and also the plastic agar plates; the MYM and dextrin media used for production of lasalocid A and tetronasin respectively is also dark brown and may also hinder the ability of light to reach the probe and cleave the ester group. The obvious consequence of this is that few intermediates were detected in *S. lasaliensis* and none were detected in *S. longisporoflavus*. Given the difficulty in capturing polyketides in *S. longisporoflavus* with the fluoromalonyl probe (**160**), not finding any captured species was not surprising but a series of putative intermediates of lasalocid A that had been previously detected and characterised from experimentation with *S. lasaliensis* were found. It is also possible that the photolabile probes can undergo deprotection outside of the bacterial cells to give highly polar carboxylates which will then be unlikely to permeate and may also decarboxylate prematurely.

The effectiveness of intermediate capture could possibly be improved if the probes were incubated with the cells for a longer period of time before irradiation to maximise cellular uptake before conversion to the carboxylate. The use of specialised flasks to allow maximum permeation of the specific wavelength of light may also improve cleavage of the ester. Despite the clear advantage of using irradiation to generate the carboxylate compared to relying on esterase's, the practical aspects of the running of the experiments need improvement.

In vitro assays with the *act* minimal PKS enzymes were set up to shed light on whether any probe with changes on the highlighted amide moiety would be capable of offloading polyketide intermediates. For these experiments, quantitative photo-cleavage of the ester was achieved separately prior to incubation within the assays, as shown by TLC analysis. Despite this, no clear SEK4/4b intermediates were captured. This suggests that these probes are poorer ACP-malonate mimics. In several cases the presence of the probes did not significantly interfere in SEK4/4b production.

The type II *act* minimal system is notoriously a hard system to investigate due to the rapid processing of enzyme-bound substrates; the use of SNAc mimics for probing this system has not proved fruitful to date.[160] Interestingly, a putative pentaketide intermediate (**441**) was intercepted by the urea photoprobe (**429**) (**Section 4.2.3**); the nature of this intermediate is in agreement with putative enzyme-bound pentaketide species characterised by a previous student in the group, Samantha Kilgour, using a chemoenzymatically modified ACP probe.[158]

Future work with the new substrates reported in this chapter will involve the testing of the more promising and interesting compounds (*e.g.* the urea photoprobe (**429**) on other *in vitro* systems and cell lysates). In particular, the fluorescent probe (**433**) offers the possibility to 'track' its journey within cells and enzymes; efforts in this direction are currently in progress.

6   Experimental

6.1   General chemistry methods

6.1.1   Chemicals and solvents

Unless otherwise specified, all reactions were performed in oven-dried round bottom flasks under argon. Chemicals were purchased from Sigma Aldrich, Fisher Scientific and VWR International (AR grade) and were used without any further purification. Dry dichloromethane (DCM), methanol (MeOH), tetrahydrofuran (THF), *N,N*-diisopropylethylamine (DIPEA), pyridine, triethylamine (Et$_3$N), dimethylformamide (DMF), toluene, diethyl ether (Et$_2$O) and acetonitrile (MeCN) were purchased from Sigma Aldrich or dried using solvent towers. Reagent grade dichloromethane (DCM), ethyl acetate (EtOAc), methanol (MeOH), acetonitrile (MeCN), acetone, dimethylsulphoxide (DMSO), cyclohexane, toluene, acetic acid (AcOH), triethylamine (Et$_3$N), *N,N*-diisopropylamine (DIPEA), petroleum ether, diethyether (Et$_2$O) and tetrahydrofuran (THF) were purchased from Fisher Scientific. Analytical thin-layer chromatography was performed on aluminium sheets pre-coated with silica gel 60 (F$_{254}$, Merck) using the indicated solvents, the spots were visualized under ultra-violet light (short wavelength) and using a stain of potassium permanganate, anisaldehyde, ninhydrin or vanillin, followed by heating using a heat gun. Silica column chromatography was carried out for purification with silica gel which was purchased from Sigma Aldrich (Tech Grade, pore size 60 Å, 230-400 mesh).

6.1.2 Compound analysis and characterisation

All IR spectra were recorded as neat samples or thin films on a Bruker Alpha-T FTIR spectrometer with 24 scans. ^{1}H-, ^{13}C- and ^{19}F- NMR spectra were recorded on Bruker Avance instruments, including DPX-300 MHz, DPX-400 MHz and DPX-500 MHz; ^{77}Se NMR spectra were recorded on a Bruker AVIII-600 MHz instrument. NMR samples were submitted in CDCl$_3$, MeOD, d$_6$-DMSO or D$_2$O unless otherwise stated. All coupling constant (J) values are quoted in Hertz (Hz) and are rounded to the nearest 0.1 for carbon couplings and 0.01 for proton couplings; all chemical shifts in parts per million (ppm). For ^{1}H- and ^{13}C- NMR data, the chemical shifts were reported relative to the solvent signal, such as CDCl$_3$ (δ_H 7.26 ppm, δ_C 77.2 ppm), D$_2$O (δ_H 4.79 ppm), d^6-DMSO (δ_H 3.31, 4.78 ppm, δ_C 49.0 ppm) and CD$_3$OD (δ_H 2.50, δ_C 39.5 ppm). ^{77}Se- and ^{19}F- NMR data were reported relative to a standard as calibrated but not with a specific solvent signal in the sample. High-resolution mass spectra (HR-MS) of novel compounds were obtained using electrospray ionization (ESI) on a MaXis UHR-TOF (Bruker Daltonics) or on Bruker MaXis (ESI-HR-MS). Melting points were recorded on a Stuart SMP20 digital melting point apparatus.

6.1.3 Synthetic procedures

6.1.3.1 Synthesis of 4-decanamidobutanoic acid (163)
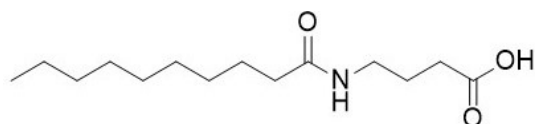
Decanoyl chloride (4.10 mL, 19.8 mmol) was added dropwise to a solution of γ-aminobutyric acid (2.00 g, 19.4 mmol) and Et$_3$N (8.10 mL, 58.2 mmol) in anhydrous

MeOH (120 mL) under argon at 0 °C. The reaction was stirred at RT for 16 h and then concentrated. The crude material was diluted with H_2O (200 mL) and acidified to pH 1 using 1M HCl. The aqueous was then extracted using $CHCl_3$ (3 x 40 mL) and the combined organics were dried over $MgSO_{4(s)}$, filtered and concentrated to give the title compound (**163**) (4.56 g, 91%): **^{1}H NMR** (400 MHz, $CDCl_3$): δ 5.78 (1H, br s, N*H*), 3.33 (2H, q, *J* = 6.6, NC*H*$_2$), 2.40 (2H, t, *J* = 7.0, C*H*$_2$CO$_2$), 2.19-2.16 (2H, m, C*H*$_2$CONH), 1.85 (2H, quin, *J* = 6.9, NHCH$_2$C*H*$_2$), 1.61-1.59 (2H, m, NHCOCH$_2$C*H*$_2$), 1.35-1.18 (12H, m, C*H*$_2$), 0.87 (3H, t, *J* = 6.8, C*H*$_3$). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[85]

6.1.3.2 Synthesis of N-(4-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-ylidene)-4-hydroxybutyl)decanamide (164)
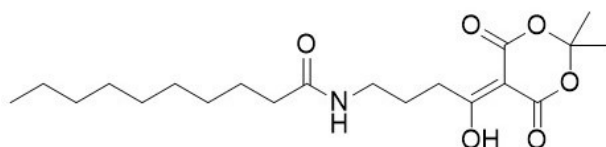
To a solution of 4-decanamidobutanoic (**163**) (3.00 g, 11.6 mmol), Meldrum's acid (1.68 g, 11.7 mmol) and DMAP (1.77 g, 14.5 mmol) in anhydrous THF (230 mL) under argon at 0 °C was added EDC.HCl (2.46 g, 12.8 mmol). The reaction was then stirred at RT for 16 h and concentrated. The crude solid was dissolved in DCM (75 mL) and washed with 1M HCl (75 mL) then H_2O (75 mL). The organic solution was then dried over $MgSO_{4(s)}$, filtered and concentrated to give a crude solid. Purification by silica chromatography with a gradient of 60-100% EtOAc in cyclohexane gave the title compound (**164**) as a yellow oil (2.79 g, 62%): **TLC** (Rf = 0.11, EtOAc); **^{1}H NMR** (400 MHz, $CDCl_3$): δ 15.31 (1H, br s, O*H*), 5.91 (1H, br s, N*H*), 3.34 (2H, q, *J* = 6.3, NHC*H*$_2$), 3.09-3.06 (2H, m, C-*H*$_2$COH), 2.18-2.16 (2H, m, C*H*$_2$CONH), 1.92 (2H, quin, *J* = 6.2, NHCH$_2$C*H*$_2$), 1.74 (6H, s, C(C*H*$_3$)$_2$), 1.63-1.61 (2H, m, C*H*$_2$CH$_2$CONH), 1.34–1.18 (12H, m, C*H*$_2$), 0.87 (3H, t, *J* = 6.8, C*H*$_3$). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[85]

6.1.3.3 Synthesis of methyl 6-decanamido-3-oxohexanoate (104)
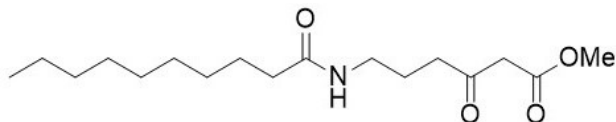
A stirring solution of N-(4-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-ylidene)-4-hydroxybutyl)decanamide (**165**) (2.79 g, 7.27 mmol) in anhydrous MeOH (120 mL) under argon was refluxed for 20 h. The reaction was then concentrated and purified by silica column chromatography with a gradient of 0-100% EtOAc in cyclohexane to give the title compound (**104**) as a white solid (2.28 g, 90%): **TLC** (Rf = 0.21, 1:1 EtOAc : cyclohexane); **MP:** 67 °C; **^{1}H NMR** (400 MHz, CDCl$_3$): δ 5.65 (1H, br s, N*H*), 3.74 (3H, s, OC*H$_3$*), 3.46 (2H, s, COC*H$_2$*CO), 3.27-3.25 (2H, m, NHC*H$_2$*), 2.61 (2H, t, *J* = 6.8, C*H$_2$*COCH$_2$CO), 2.15-2.13 (2H, m, CH$_2$CONH), 1.81 (2H, quin, *J* = 6.8, NHCH$_2$C*H$_2$*), 1.62-1.60 (2H, m, C*H$_2$*CH$_2$CONH), 1.33–1.20 (12H, m, C*H$_2$*), 0.87 (3H, t, *J* = 6.8, C*H$_3$*) Proton NMR data gathered for this compound were in accordance with those reported in the literature.[85]

6.1.3.4 Synthesis of methyl 6-decanamido-2-methyl-3-oxohexanoate (161)
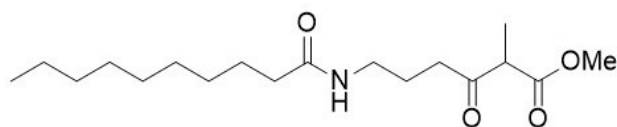
To a solution of methyl 6-decanamido-3-oxohexanoate (**104**) (0.20 g, 0.64 mmol) and potassium carbonate (0.23 g, 1.69 mmol) in anhydrous THF (6 mL) at 0 °C was added dropwise methyl iodide (0.05 mL, 0.77 mmol). After stirring at RT for 24 h, the reaction was quenched with H$_2$O (10 mL) and concentrated. The remaining aqueous phase was extracted with DCM (3 x 10 mL) and the combined organic layers were dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica

column chromatography with a gradient of 0-55% EtOAc in cyclohexane afforded the title compound (**161**) as a white solid (0.14 g, 68%): **TLC** (Rf = 0.24, 1:1 EtOAc : cyclohexane); **MP:** 62 °C; **^{1}H NMR** (400 MHz, CDCl$_3$): δ 5.64 (1H, br s, N*H*), 3.73 (3H, s, OC*H*$_3$), 3.54 (1H, q, *J* = 7.3, CH$_3$C*H*), 3.25 (2H, m, NHC*H*$_2$), 2.66 (1H, dt, *J* = 18.1, 6.9, C*H*$_2$COCH), 2.54 (1H, dt, *J* = 18.1, 6.9, C*H*$_2$COCH), 2.14 (2H, m, C*H*$_2$CONH), 1.80 (2H, quin, *J* = 6.8, NHCH$_2$C*H*$_2$), 1.60 (2H, m, C*H*$_2$CH$_2$CONH), 1.34 (3H, d, *J* = 7.2, C*H*$_3$CH), 1.32-1.21 (12H, m, C*H*$_2$), 0.87 (3H, t, *J* = 6.9, C*H*$_3$); **^{13}C NMR** (125 MHz, CDCl$_3$): 205.8 (CH$_2$*C*OCH), 173.4 (*C*ONH), 171.1 (*C*O$_2$CH$_3$), 52.6 (CO*C*HCO), 52.4 (O*C*H$_3$), 38.9 (NH*C*H$_2$), 38.7 (*C*H$_2$COCHCO), 36.9 (*C*H$_2$CONH), 31.9 (*C*H$_2$), 29.5 (*C*H$_2$), 29.4 (*C*H$_2$), 29.4 (*C*H$_2$), 29.3 (*C*H$_2$), 25.8 (*C*H$_2$CH$_2$CONH), 23.3 (NHCH$_2$*C*H$_2$), 22.7 (*C*H$_2$), 14.1 (*C*H$_3$), 12.8 (*C*H$_3$CH); **IR** (neat) v_{max} = 3308, 2920, 2851, 1737, 1699, 1638, 1551, 1327, 1262 cm^{-1}; **HRMS (ESI)**: calculated for C$_{18}$H$_{34}$NO$_4$ [M+H]$^+$: 328.2482, found: 328.2479.

6.1.3.5 Synthesis of methyl 6-decanamido-2-fluoro-3-oxohexanoate (160)
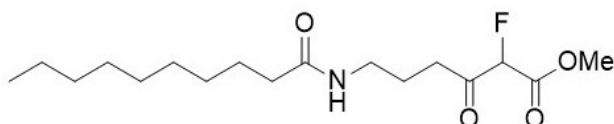
To a solution of methyl 6-decanamido-3-oxohexanoate (**104**) (1.00 g, 3.19 mmol) and cyclopentadienyltitanium trichloride (0.04 g, 0.16 mmol) in anhydrous MeCN (40 mL) under argon at RT was added Selectfluor (1.24 g, 3.51 mmol) followed by stirring for 16 h. The reaction was concentrated, re-dissolved in H$_2$O (40 mL) and the resulting aqueous solution was then extracted with DCM (4 x 10 mL). The combined organics were washed with brine (30 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 50-100% EtOAc in petroleum ether afforded the title compound (**160**) as a white solid (0.65 g, 61%): **TLC** (Rf = 0.48, 3:1 EtOAc : cyclohexane), **MP:** 48 °C; **^{1}H NMR** (300 MHz, CDCl$_3$): 5.51 (1H, br s, N*H*), 5.26 (1H, d, *J* = 48.9 (C*H*F), 3.86 (3H, s, OC*H*$_3$), 3.27 (2H,

q, J = 6.5, NHCH_2), 2.81-2.64 (2H, m, CH_2COCHF), 2.19-2.11 (2H, m, CH_2CONH), 1.84 (2H, quin, J = 6.9, NCH$_2$CH_2), 1.64-1.58 (2H, m, CH_2CH$_2$CONH), 1.34-1.18 (12H, m, CH_2), 0.88 (3H, t, J = 6.7, CH_3); ^{19}**F NMR** (282 MHz, CDCl$_3$): -113.5 (S, CFH). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[85]

6.1.3.6 Synthesis of methyl 2-(triphenyl-λ^5-phosphanylidene)propanoate (188)
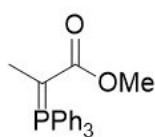
A solution of methyl 2-bromopropionate (4.00 mL, 35.9 mmol) and triphenylphosphine (9.40 g, 35.9 mmol) in H$_2$O (100 mL) was refluxed for 16 h. The resulting solution was cooled to RT and diluted with 2M KOH (250 mL). The precipitated solution was diluted with DCM (50 mL) and stirred for 20 min. The organic layer was separated and the aqueous was extracted with DCM (3 x 40 mL). The combined organic extracts were dried over MgSO$_{4(s)}$, filtered and concentrated to give the title compound (**188**) as a green/yellow solid (12.2 g, 97%). ^{1}**H NMR** (500 MHz, CDCl$_3$): δ 7.72-7.42 (15H, m, Ar-H), 3.16, 3.14 (3H, br s, OCH_3), 1.68-1.53 (3H, m, CH_3CCO$_2$). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[161]

6.1.3.7 Synthesis of methyl (2E,4E)-2,4-dimethylhexa-2,4-dienoate (189)
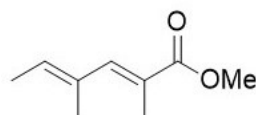
A solution of methyl 2-(triphenyl-λ^5-phosphanylidene)propanoate (**188**) (12.2 g, 34.9 mmol) and (E)-2-methylbut-2-enal (**133**) (3.10 mL, 31.7 mmol) in anhydrous toluene (90 mL) was refluxed under argon for 72 h. The reaction was then concentrated, triturated with Et$_2$O, filtered and the filtrate concentrated to give the crude product. Purification by

silica column chromatography with a gradient of 0-40% EtOAc in cyclohexane gave the title compound (**189**) as a pale-yellow oil (3.48 g, 62%): **TLC** (Rf = 0.42, 3:1 petroleum ether : Et$_2$O); **^{1}H NMR** (400 MHz, CDCl$_3$): δ 7.13 (1H, s, C*H*CCO$_2$), 5.72 (1H, q, *J* = 6.3, CH$_3$C*H*C), 3.75 (3H, s, OC*H*$_3$), 2.00 (3H, s, C*H*$_3$CCO$_2$), 1.84 (3H, s, C*H*$_3$CCHCCO$_2$), 1.75 (3H, s, C*H*$_3$CH). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[162]

6.1.3.8   Synthesis of (2E,4E)-2,4-dimethylhexa-2,4-dienoic acid (190)
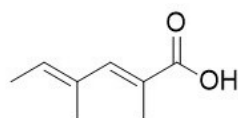
To a solution of methyl (2E,4E)-2,4-dimethylhexa-2,4-dienoate (**189**) (1.00 g, 6.17 mmol) in EtOH (10 mL) was added 2M NaOH (10 mL, 20.00 mmol) and the reaction was stirred at 50 °C for 16 h. The reaction was then cooled, concentrated and the remaining aqueous was washed with Et$_2$O (5 mL). The aqueous was acidified to pH 2 using 1M HCl and extracted using EtOAc (4 x 10 mL). The combined EtOAc extracts were dried over MgSO$_{4(s)}$, filtered and concentrated to give the product (**190**) as a brown solid (0.58 g, 67%): **TLC** (Rf = 0.71, 3:2 EtOAc : cyclohexane); **^{1}H NMR** (400 MHz, CDCl$_3$): δ 7.13 (1H, s, C*H*CCO$_2$), 5.72 (1H, q, *J* = 6.9, CH$_3$C*H*C), 3.75 (3H, s, OC*H*$_3$), 2.00 (3H, s, C*H*$_3$CCO$_2$), 1.84 (3H, s, C*H*$_3$CCHCCO$_2$), 1.75 (3H, d, *J* = 6.9, C*H*$_3$CH). Proton NMR data gathered for this compound were in accordance with those reported in the literature .[163]

6.1.3.9   Synthesis   of   S-(2-acetamidoethyl)   (2E,4E)-2,4-dimethylhexa-2,4-dienethioate (192)
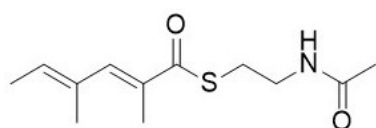
To a solution of (2E,4E)-2,4-dimethylhexa-2,4-dienoic acid (**190**) (40 mg, 0.29 mmol) in anhydrous DCM (2.5 mL) under argon at 0 °C was added EDC.HCl (60 mg, 0.31 mmol) and DMAP (2 mg, 0.02 mmol), the solution was stirred at 0 °C for 15 min. *N*-acetyl cysteamine (**191**) (0.03 mL, 0.29 mmol) was then added to the reaction which was stirred at RT for 4 h. The reaction was diluted with DCM (10 mL) and washed with 0.01M HCl (2 x 5 mL) then brine (2 x 5 mL). The organic solution was then dried over $MgSO_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-90% EtOAc in cyclohexane gave the title compound (**192**) as a white solid (37 mg, 54%): **TLC** (Rf = 0.16, 1:1 EtOAc : cyclohexane); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.12 (1H, s, C*H*CCSO), 5.91 (1H, br s, N*H*), 5.82 (1H, q, *J* = 6.7, C*H*CH$_3$), 3.46 (2H, q, *J* = 6.1, C*H$_2$*NH), 3.08 (2H, t, *J* = 6.3, SC*H$_2$*), 1.96 (3H, s, C*H$_3$*CONH), 2.06 (3H, s, C*H$_3$*CCSO), 1.87 (3H, s, C*H$_3$*CCHCCSO), 1.77 (3H, d, *J* = 6.9, C*H$_3$*CH); **^{13}C NMR** (125 MHz, CDCl$_3$): 194.9 (OC*S*), 170.3 (NH*C*O), 142.8 (*C*HCCSO), 133.2 (*C*HCH$_3$), 132.9 (*C*CHCCSO), 132.5 (*C*CSO), 39.9 (*C*H$_2$NH), 28.7 (S*C*H$_2$), 23.3 (*C*H$_3$CONH), 16.0 (*C*H$_3$), 14.3 (*C*H$_3$CH), 14.1 (*C*H$_3$CCHCCSO); **IR** (thin film) v_{max} = 3277, 3077, 2925, 2855, 1649, 1547, 1185, 1019 cm^{-1}; **HRMS (ESI)**: calculated for C$_{12}$H$_{19}$NNaO$_2$S [M+Na]$^+$: 264.1029, found: 264.1039.

6.1.3.10 Synthesis of S-(2-acetamidoethyl) ethanethioate (199)
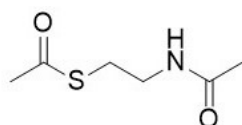
Aqueous KOH (6M) was added at room temperature (RT) to a stirring solution of cysteamine hydrochloride (**198**) (2.00 g, 17.6 mmol) in H$_2$O (23 mL) until the pH reached 8. Acetic anhydride (5.16 mL, 54.6 mmol) was added dropwise to the solution along with 6M KOH to maintain a pH of 9. After this, the reaction was stirred at RT for 1.5 h. The reaction was then saturated with NaCl$_{(s)}$ and extracted with DCM (5 x 20 mL), the combined organics were dried over $MgSO_{4(s)}$, filtered and concentrated to afford the title compound (**199**) as a colourless oil (2.65 g, 93%): **^{1}H NMR** (300 MHz, CDCl$_3$): δ 5.83

(1H, br s, N*H*), 3.44 (2H, q, *J* = 6.1, NC*H₂*), 3.02 (2H, t, *J* = 6.4, SC*H₂*), 2.36 (3H, s, NCOC*H₃*), 1.97 (3H, s, SCOC*H₃*). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[164]

6.1.3.11 Synthesis of S-(2-acetamidoethyl) (E)-2-methylbut-2-enethioate (201)
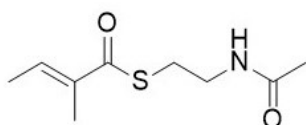
To a solution of (E)-2-methylbut-2-enoic acid (**200**) (0.13 g, 1.30 mmol), *N*-acetyl cysteamine (**191**) (0.17 g, 1.40 mmol) and DMAP (0.03 g, 0.30 mmol) in anhydrous DCM (2 mL) at 0 °C under argon was added 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC.HCl) (0.28 g, 1.40 mmol) followed by stirring at RT for 16 h. The reaction was then diluted with 1M HCl (10 mL) and extracted with DCM (3 x 5 mL). The combined organics were washed with saturated NaHCO$_{3(aq.)}$ (10 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 20-100% EtOAc in petroleum ether afforded the title compound (**201**) as a white solid (0.13 g, 50%): **TLC** (Rf = 0.27, 3:1 EtOAc : petroleum ether); **¹H NMR** (300 MHz, CDCl₃): δ 6.87 (1H, q, *J* = 6.8, CH₃C*H*), 5.87 (1H, br s, N*H*), 3.45 (2H, q, *J* = 5.9, C*H₂*NH), 3.06 (2H, t, *J* = 6.1, SC*H₂*), 1.96 (3H, S, CONHC*H₃*), 1.87 (3H, s, C*H₃*CCOS), 1.84 (3H, d, *J* = 7.0, C*H₃*CH). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[165]

Synthesis of (2E,4E)-2,4-dimethylhexa-2,4-dien-1-ol (202)
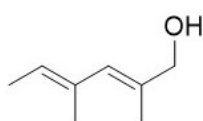
To a solution of methyl (2E,4E)-2,4-dimethylhexa-2,4-dienoate (**189**) (2.08 g, 12.8 mmol) in anhydrous Et$_2$O (30 mL) under argon at 0 °C was added diisobutylaluminium hydride (DIBAL, 1M in hexane, 26.8 mL, 26.8 mmol) over 5 mins. The solution was then stirred at RT for 1 h and quenched with slow dropwise addition of MeOH (1.5 mL). The resulting solution was diluted with Et$_2$O (25 mL) and a saturated aqueous potassium tartrate tetrahydrate solution (40 mL), and stirred at RT for further 3 h. The organic layer was then separated, washed with brine (20 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-50% Et$_2$O in petroleum ether gave the title compound (**202**) as a cloudy oil (0.95 g, 56%): TLC (Rf = 0.27, 1:1 Et$_2$O : petroleum ether); **¹H NMR** (400 MHz, CDCl$_3$): δ 5.89 (1H, s, C*H*CCH$_2$OH), 5.42 (1H, q, *J* = 6.8), 4.04 (2H, d, *J* = 5.4, C*H$_2$*OH), 1.81 (3H, s, C*H$_3$*CCH$_2$OH), 1.75 (3H, s, C*H$_3$*CCHCCH$_2$OH), 1.69 (2H, d, *J* = 6.8, C*H$_3$*CHC), 1.33 (3H, t, *J* = 6.0, O*H*); Proton NMR data gathered for this compound were in accordance with those reported in the literature.[166]

6.1.3.12 Synthesis of (2E,4E)-2,4-dimethylhexa-2,4-dienal (203)
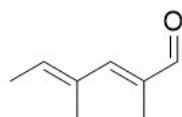
To a solution of methyl (2E,4E)-2,4-dimethylhexa-2,4-dienoate (**202**) (0.95 g, 7.54 mmol) in anhydrous DCM (80 mL) under Argon was added manganese dioxide (3.31 g, 37.70 mmol) followed by stirring at RT. TLC confirmed reaction completion after 16 h and was then filtered through Celite. The filtrate was concentrated to give the title compound (**203**) (0.80 g, 86%). **¹H NMR** (400 MHz, CDCl$_3$): δ 9.38 (1H, s, C*H*O), 6.73 (1H, s, C*H*CCHO), 5.99 (1H, q, *J* = 7.0, C*H*CH$_3$), 1.98 (3H, s, C*H$_3$*CCHCCHO), 1.96 (3H, s, C*H$_3$*CCHO), 1.82 (3H, d, *J* = 7.0, C*H$_3$*CH). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[167]

6.1.3.13 S-phenyl propanethioate (204)
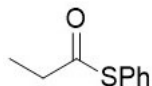
To a solution of thiophenol (1.02 mL, 10.0 mmol) and pyridine (0.81 mL, 10.0 mmol) in DCM (10 mL) under argon at 0 °C was added dropwise propionyl chloride (0.87 mL, 10.0 mmol). The reaction was then allowed to warm to RT over 1 h. The reaction was diluted with H_2O (20 mL) and extracted using DCM (3 x 5 mL), the combined organic layers were dried over $MgSO_{4(s)}$, filtered and concentrated to give the title compound (204) as a colourless oil which was used without further purification (1.78 g, quant): ^1H **NMR** (500 MHz, CDCl$_3$): δ 7.41 (5H, s, Ar-*H*), 2.69 (2H, q, *J* = 7.5, C*H*$_2$CH$_3$), 1.23 (3H, t, *J* = 7.5, CH$_2$C*H*$_3$). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[168]

6.1.3.14 Synthesis of S-phenyl (4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-dienethioate (205)
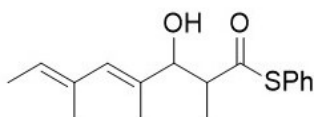
To a solution of S-phenyl propanethioate (204) (0.16 g, 0.99 mmol) and (2E,4E)-2,4-dimethylhexa-2,4-dienal (203) (0.13 g, 1.08 mmol) in DCM (5 mL) under argon at RT was added in one portion MgBr$_2$.OEt$_2$ (0.36 g, 1.38 mmol). After stirring for 5 mins, DIPEA (0.34 mL, 1.97 mmol) was added dropwise followed by stirring at RT for 1 h. The reaction was then diluted with EtOAc (8 mL) and H_2O (8 mL), the aqueous phase was adjusted the pH 5 using 1M HCl and the biphasic solution was stirred vigorously for 5 mins. The organic layer was separated and aqueous extracted using EtOAc (4 x 5 mL), the combined organic extracts were washed with brine (10 mL), dried over $MgSO_{4(s)}$,

filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-20% EtOAc in petroleum ether afforded the title compound (**205**) as a yellow oil (0.19 g, 69%) with a relative stereochemistry ratio of 1:1 *syn* and *anti* as shown by NMR: **TLC** (Rf = 0.16, 9:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.45-7.35 (5H, m, Ph), 5.97 (1H, s, C*H*CCOH), 5.90 (1H, s, C*H*CCOH), 5.44 (1H, q, *J* = 7.0, CH$_3$C*H*), 4.38-4.35 (1H, m, C*H*OH), 4.23-4.18 (1H, m, C*H*OH), 3.04-2.97 (1H, m, C*H*COS), 2.27-2.24 (1H, m, O*H*), 2.24-2.21 (1H, m, O*H*), 1.80 (3H, s, C*H*$_3$CCOH), 1.75 (3H, s, C*H*$_3$CCHCCOH), 1.69 (3H, d, *J* = 6.8, C*H*$_3$CH), 1.28 (3H, d, *J* = 7.0, C*H*$_3$CHCOS), 1.16 (3H, d, *J* = 7.1, C*H*$_3$CHCOS); **^{13}C NMR** (125 MHz, CDCl$_3$): 201.6 (O*C*S), 201.3 (O*C*S), 134.6 (Ph), 134.5 (Ph), 133.1 (*C*COH), 133.0 (*C*HCCOH), 132.9 (*C*COH), 132.5 (*C*CHCCHCOH), 132.2 (*C*CHCCHCOH), 131.2 (*C*HCCOH), 129.5 (Ph), 129.4 (Ph), 129.3 (Ph), 129.2 (Ph), 127.6 (OCS*C*), 127.4 (OCS*C*), 125.4 (CH$_3$*C*H), 124.9 (CH$_3$*C*H), 80.7 (*C*OH), 77.4 (*C*OH), 51.6 (*C*HCOS), 51.3 (*C*HCOS), 16.6 (*C*H$_3$CCHCCOH), 16.5 (*C*H$_3$CCHCCOH), 15.4 (*C*H$_3$CHCOS), 14.4 (*C*H$_3$CH), 13.7 (*C*H$_3$CH), 12.7 (*C*H$_3$CH), 12.1 (*C*H$_3$CHCOS); **IR** (thin film) v_{max} = 3430, 2972, 2917, 1687, 1440, 1326, 947 cm^{-1}; **HRMS (ESI)**: calculated for C$_{17}$H$_{22}$NaO$_2$S [M+Na]$^+$: 313.1234, found: 313.1231

6.1.3.15 Synthesis of S-(2-acetamidoethyl) (4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-dienethioate (206)
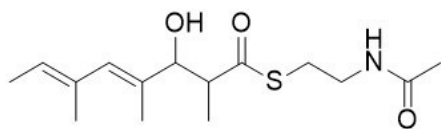
To a solution of S-phenyl (4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-dienethioate (**205**) (0.18 g, 0.62 mmol) and *N*-acetyl cysteamine (**191**) (0.08 g, 0.67 mmol) in anhydrous DCM (6 mL) under argon at RT was added anhydrous Et$_3$N (0.13 mL, 0.65 mmol) followed by stirring for 6 h. The reaction was then concentrated and purified by silica column chromatography with a gradient of 0-70% EtOAc in petroleum ether to afford the

title compound (**206**) as a colourless oil (0.04 g, 23%) with a relative stereochemistry ratio of 1:1 *syn* and *anti* as shown by NMR: **TLC** (Rf = 0.36, 95:5 DCM : MeOH); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 5.92 (1H, s, C*H*CCOH), 5.89 (1H, s, C*H*CCOH), 5.87 (1H, br s, N*H*), 5.80 (1H, br s, N*H*), 5.45-5.35 (1H, m, CH$_3$C*H*CCH$_3$), 4.31 (1H, d, *J* = 5.1, C*H*OH), 4.17 (1H, d, *J* = 8.8, C*H*OH), 3.54-3.48 (2H, m, C*H*$_2$N), 3.14-2.95 (2H, m, SC*H*$_2$), 2.95-2.85 (1H, m, C*H*CSO), 2.33 (1H, d, *J* = 2.3, O*H*), 2.18 (1H, d, *J* = 2.3, O*H*), 1.95 (3H, s, NHCOC*H*$_3$), 1.77 (3H, s, C*H*$_3$CCHOH), 1.74 (3H, s, CH$_3$CHCC*H*$_3$), 1.73 (3H, s, CH$_3$CHCC*H*$_3$), 1.70-1.66 (3H, C*H*$_3$CHCCH$_3$), 1.20 (3H, d, *J* = 7.0, C*H*$_3$CHCSO), 1.06 (3H, d, *J* = 7.1, C*H*$_3$CHCSO); **^{13}C NMR** (125 MHz, CDCl$_3$): 203.7 (*C*SO), 203.4 (*C*SO), 170.6 (NH*C*O), 170.5 (NH*C*O), 133.4 (*C*HCCOH), 133.1 (*C*CHCCOH), 133.0 (*C*CHCCOH), 132.9 (*C*COH), 132.8 (*C*COH), 131.1 (*C*HCCOH), 125.8 (CH$_3$*C*HCCH$_3$), 125.1 (CH$_3$*C*HCCH$_3$), 81.1 (*C*OH), 77.9 (*C*OH), 52.1 (*C*CSO), 51.9 (*C*CSO), 39.7 (*C*H$_2$N), 39.6 (*C*H$_2$N), 28.9 (S*C*H$_2$), 28.8 (S*C*H$_2$), 23.3 (3H, s, NHCO*C*H$_3$), 16.8 (3H, s, CH$_3$CHC*C*H$_3$), 16.6 (3H, s, 15.4 (*C*H$_3$CHCSO), 14.5 (*C*H$_3$CHCCH$_3$), 13.8 (*C*H$_3$CHCH$_3$), 12.6 (*C*H$_3$CCHOH), 12.0 (*C*H$_3$CHCSO); **IR** (thin film) v_{max} = 3293, 2976, 2932, 2859, 1655, 1544, 1471, 983, 681 cm^{-1}; **HRMS (ESI)**: calculated for C$_{15}$H$_{25}$NNaO$_3$S [M+Na]$^+$: 322.1447, found: 322.1450

6.1.3.16 Synthesis of S-(2-acetamidoethyl) (4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-dienethioate (207)
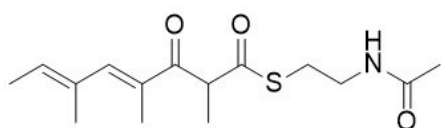
To a solution of S-(2-acetamidoethyl) (4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-dienethioate (**206**) (0.08 g, 0.28 mmol) in EtOAc (14 mL) was added iodoxybenzoic acid (IBX) (45% wt., 0.52 g, 0.84 mmol) followed by stirring at reflux for 1 h. The reaction was then filtered through Celite and the filtrate was concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-70% EtOAc

in cyclohexane afforded the title compound (**207**) as a pale grey oil (0.04 g, 42%) (as a 4:1 keto: enol tautomeric ratio as shown by NMR): **TLC** (Rf = 0.39, EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): (**Keto form**) δ 7.05 (1H, s, C*H*CCO), 5.85 (1H, q, *J* = 7.0, C*H*CCHCCO), 5.84 (1H, br s, N*H*), 4.47 (1H, q, *J* = 7.0, C*H*CSO), 3.45-3.43 (2H, m, C*H*$_2$NH), 3.09-3.06 (2H, m, SC*H*$_2$), 1.98 (3H, s, C*H*$_3$CCO), 1.95 (3H, s, C*H*$_3$CONH), 1.90 (3H, s, C*H*$_3$CCHCCO), 1.79 (3H, d, *J* = 7.0, C*H*$_3$CH), 1.42 (3H, d, *J* = 7.0, C*H*$_3$CHCSO): (**Enol form**) δ 7.08 (1H, s, C*H*CCOH), 5.57 (1H, q, *J* = 7.0, C*H*CCHCCOH). **^{13}C NMR** (125 MHz, CDCl$_3$): (**Keto form**) 197.6 (*C*COCH), 197.3 (O*C*S), 170.4 (*C*ONH), 145.5 (*C*HCCO), 133.7 (*C*HCCHCCO), 133.4 (*C*CHCCO), 133.2 (*C*COCHCSO), 55.0 (CO*C*HCSO), 39.5 (*C*H$_2$NH), 28.7 (S*C*H$_2$), 23.2 (*C*H$_3$CONH), 16.1 (*C*H$_3$CCHCCO), 15.1 (*C*H$_3$CHCSO), 14.3 (*C*H$_3$CHC), 13.4 (*C*H$_3$CCO); (**Enol form**) 140.7 (*C*HCCOH), 127.0 (*C*HCCHCCO); **IR** (thin film) v_{max} = 3285, 3069, 2933, 2871, 1653, 1541, 1476, 1432, 1276, 983 cm^{-1}; **HRMS (ESI)**: calculated for C$_{15}$H$_{23}$NNaO$_3$S [M+Na]$^+$: 320.1291, found: 320.1292.

6.1.3.17 Synthesis of methyl (4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-dienoate (209)
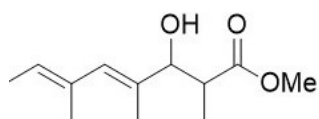
To stirring anhydrous THF (8 mL) at 70 °C under argon was added in quick succession activated zinc dust (0.40 g, 6.10 mmol), (2E, 4E)-2,4-dimethylhexa-2,4-dienal (**203**) (0.42 g, 3.39 mmol) and then methyl bromopropionate (**208**) (0.68 mL, 6.10 mmol) followed by stirring at 70 °C for 5 min. The reaction colour became green along with some gas evolution. After this, the reaction stirred at RT for 15 mins followed by concentration. The residue was diluted in H$_2$O (15 mL) and extracted with EtOAc (4 x 10 mL). The combined organics were washed with 1M HCl (10 mL), H$_2$O (10 mL) and brine (15 mL), then dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-15% EtOAc in petroleum ether afforded the title compound (**209**) as a colourless oil (0.54 g, 75%) with a relative

stereochemistry ratio of 1:1 *syn* and *anti* as shown by NMR: **TLC** (Rf = 0.48/0.58, 4:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 5.93 (1H, s, C*H*CCOH), 5.88 (1H, s, C*H*CCOH), 5.42 (1H, q, *J* = 6.8, CH$_3$C*H*CCH), 5.38 (1H, q, *J* = 6.8, CH$_3$C*H*CCH), 4.29 (1H, d, *J* = 5.0, CO*H*), 4.11 (1H, d, *J* = 8.6, CO*H*), 3.73 (3H, s, OC*H$_3$*), 3.68 (3H, s, OC*H$_3$*), 2.77-2.64 (1H, m, C*H*CO$_2$), 2.40 (1H, br s, O*H*), 2.32 (1H, br s, O*H*), 1.76-1.72 (3H, m, C*H$_3$*CCOH), 1.76-1.72 (3H, m, C*H$_3$*CCHCCOH), 1.70-1.66 (3H, m, H$_3$CCC*H$_3$*), 1.16 (3H, d, *J* = 7.1, C*H$_3$*CHCOH), 1.06 (3H, d, *J* = 7.1, C*H$_3$*CH); **^{13}C NMR** (125 MHz, CDCl$_3$): 176.6 (*C*O$_2$), 176.2 (*C*O$_2$), 133.2 (*C*HCCOH), 133.2 (*C*COH), 133.0 (*C*CHCCOH), 132.9 (*C*CHCCOH), 130.6 (*C*CHCCOH), 125.4 (CH$_3$*C*H), 124.7 (CH$_3$*C*H), 80.7 (*C*OH), 52.0 (O*C*H$_3$), 51.9 (O*C*H$_3$), 43.5 (*C*HCO$_2$), 43.1 (*C*HCO$_2$), 16.7 (*C*H$_3$CHCCH), 16.6 (*C*H$_3$CHCCH), 14.5 (*C*H$_3$CCHCCOH), 14.4 (*C*H$_3$CHCO$_2$), 13.8 (*C*H$_3$CCHCCOH), 12.5 (*C*CHCOH), 11.3 (*C*H$_3$CHCO$_2$); **IR** (thin film) υ$_{max}$ = 3456, 2949, 1720, 1456, 1197, 1014 cm^{-1}; **HRMS (ESI)**: calculated for C$_{12}$H$_{20}$NaO$_3$ [M+Na]$^+$: 235.1305, found: 235.1307.

6.1.3.18 Synthesis of (4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-dienoic acid (210)
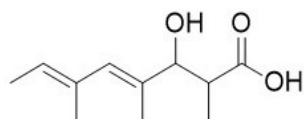
To a solution of methyl(4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-dienoate (**209**) (0.54 g, 2.53 mmol) in 3:3:1 THF : MeOH : H$_2$O (21 mL) at RT was added LiOH (0.30 g, 12.6 mmol) followed by stirring for 6 h. The reaction was then concentrated to leave an aqueous solution which was diluted further with H$_2$O (10 mL). The solution was then washed with Et$_2$O (20 mL), acidified to pH 2 using 1M HCl and extracted with EtOAc (4 x 10 mL). The combined organic layers were dried over MgSO$_{4(s)}$, filtered and concentrated to give the title compound (**210**) as a colourless oil (0.43 g, 87%) with a relative stereochemistry ratio of 1:1 *syn* and *anti* as shown by NMR: **^{1}H NMR** (500 MHz, CDCl$_3$): δ 5.97 (1H, s, C*H*CCOH), 5.90 (1H, s, C*H*CCOH), 5.44 (1H, q, *J* = 6.9, (CH$_3$C*H*CCH), 5.40 (1H, q, *J* = 7.0, (CH$_3$C*H*CCH), 4.38 (1H, d, *J* = 4.8, C*H*OH), 4.14

(1H, d, J = 9.2, C*H*OH), 2.81-2.74 (1H, m, C*H*CO$_2$), 2.74-2.67 (1H, m, C*H*CO$_2$), 1.77 (3H, s, C*H$_3$*CCOH), 1.75 (3H, s, C*H$_3$*CCOH), 1.75 (3H, s, C*H$_3$*CCHCCOH), 1.73 (3H, s, C*H$_3$*CCHCCOH), 1.71-1.66 (3H, d, J = 7.0, CH$_3$CH), 1.17 (3H, d, J = 7.1, C*H$_3$*CHCOH), 1.09 (3H, d, J = 7.1, CH$_3$C*H*COH); **^{13}C NMR** (125 MHz, CDCl$_3$): 179.6 (*C*O$_2$), 179.4 (*C*O$_2$), 133.8 (*C*HCCOH), 132.9 (*C*COH), 132.8 (*C*COH), 132.7 (*C*CHCCOH), 130.9 (*C*HCCOH), 125.8 (CH$_3$*C*H), 125.0 (CH$_3$*C*H), 80.7 (*C*OH), 77.1 (*C*OH), 43.1 (*C*HCO$_2$), 42.7 (*C*HCO$_2$), 16.7 (*C*H$_3$CHCCH), 16.6 (*C*H$_3$CHCCH), 14.6 (*C*H$_3$CCHCCOH), 13.8 (*C*H$_3$CCHCCOH), 12.3 (*C*H$_3$CHCO$_2$), 14.3 (*C*CHCOH), 10.9 (*C*H$_3$CHCO$_2$); **IR** (thin film) v_{max} = 2976, 2917, 2860, 1708, 1558, 1456, 1186, 1005, 732 cm^{-1}; **HRMS (ESI)**: calculated for C$_{11}$H$_{18}$NaO$_3$ [M+Na]$^+$: 221.1148, found: 221.1147.

6.1.3.19 Alternative synthesis of S-(2-acetamidoethyl) (4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-dienethioate (206)
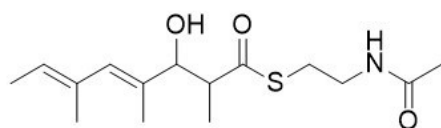
To a solution of (4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-dienoic acid (**210**) (0.10 g, 0.50 mmol), *N*-acetyl cysteamine (**191**) (0.06 g, 0.52 mmol) and DMAP (0.01 g, 0.05 mmol) in anhydrous THF (3 mL) under argon at 0 °C was added DCC (0.11 g, 0.55 mmol) followed by stirring at RT for 16 h. The reaction was then concentrated, resuspended in EtOAc (2 mL) and filtered through Celite. The filtrate was concentrated and purified by silica column chromatography with a gradient of 50-100% EtOAc in petroleum ether to afford the title compound (**206**) as a pale-yellow oil (0.07 g, 46%) with a relative stereochemistry ratio of 1:1 *syn* and *anti* as shown by NMR: **TLC** (Rf = 0.36, 95:5 DCM : MeOH); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 5.92 (1H, s, C*H*CCOH), 5.77 (1H, br s, N*H*), 5.40 (1H, q, J = 6.8, C*H*CCHCCOH), 4.31 (1H, br s, *H*COH), 3.44-3.42 (2H, m, C*H$_2$*NH), 3.05-3.03 (2H, m, SC*H$_2$*), 2.93-2.91 (1H, m, C*H*COS), 2.21 (1H, d, J = 2.6, O*H*), 1.95 (3H, s, C*H$_3$*CONH), 1.77 (3H, s, C*H$_3$*CCOH), 1.73 (3H, s, CH$_3$CCHCCOH), 1.68 (3H, d,

J = 6.6, C*H*$_3$CHC), 1.20 (3H, d, J = 6.9, C*H*$_3$CHCO); **^{13}C NMR** (125 MHz, CDCl$_3$): 203.4 (S*C*O), 170.3 (NH*C*O), 132.9 (*C*CHCCOH), 132.9 (*C*COH), 131.0 (*C*HCCOH), 125.0 (*C*HCCHCCOH), 77.8 (51.7 (*C*HCO), 39.5 (*C*H$_2$NH), 28.6 (S*C*H$_2$), 23.2 (*C*H$_3$CONH), 16.6 (*C*H$_3$CCHCCOH), 14.4 (*C*H$_3$CCOH), 13.7 (*C*H$_3$CHC), 11.8 (*C*H$_3$CCO); **IR** (thin film) v_{max} = 3293, 2976, 2932, 2859, 1655, 1544, 1471, 983, 681 cm^{-1}; **HRMS (ESI)**: calculated for C$_{15}$H$_{25}$NNaO$_3$S [M+Na]$^+$: 322.1447, found: 322.1450

6.1.3.20 Synthesis of S-phenyl ethanethioate (213)
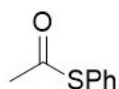
To a solution of thiophenol (1.30 mL, 12.7 mmol) and K$_2$CO$_3$ (3.37 g, 25.4 mmol) in EtOAc (100 mL) under argon at RT was added acetic anhydride (1.56 mL, 16.5 mmol). After stirring for 48 h, the reaction was filtered through Celite. The filtrate was concentrated to give the title compound (**213**) as a colourless oil which was used without further purification (2.63 g, 100%): **^{1}H NMR** (300 MHz, CDCl$_3$): δ 7.16-6.96 (5H, m, Ar-*H*), 2.27 (3H, COC*H*$_3$). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[169]

6.1.3.21 Synthesis of S-phenyl (4E,6E)-3-hydroxy-4,6-dimethylocta-4,6-dienethioate (214)
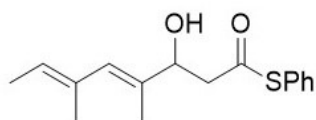
To a solution of S-phenyl ethanethioate (**213**) (0.15 g, 0.99 mmol) and (2E,4E)-2,4-dimethylhexa-2,4-dienal (**203**) (0.13 g, 1.08 mmol) in DCM (5 mL) under argon at RT was added in one portion MgBr$_2$.OEt$_2$ (0.36 g, 1.38 mmol). After stirring for 5 mins,

DIPEA (0.34 mL, 1.97 mmol) was added dropwise followed by stirring at RT for 1 h. The reaction was then diluted with EtOAc (8 mL) and H$_2$O (8 mL), the aqueous phase was adjusted the pH 5 using 1M HCl and the biphasic solution was stirred vigorously for 5 mins. The organic layer was separated and aqueous extracted using EtOAc (4 x 5 mL), the combined organic extracts were washed with brine (10 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-20% EtOAc in petroleum ether afforded the title compound (**214**) as a yellow oil (0.20 g, 74%): **TLC** (Rf = 0.48, 3:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.35 (5H, br s, Ph), 5.96 (1H, s, C*H*CCOH), 5.43 (1H, q, *J* = 6.8, CH$_3$C*H*), 4.55-4.53 (2H, m, C*H*OH), 2.98-2.86 (2H, m, C*H*$_2$COS), 2.52 (1H, d, *J* = 2.3, O*H*), 1.81 (3H, s, C*H*$_3$CCOH), 1.75 (3H, s, C*H*$_3$CCHCCOH), 1.69 (3H, d, *J* = 6.8, C*H*$_3$CH); **^{13}C NMR** (125 MHz, CDCl$_3$): 197.3 (S*C*O), 134.49 (Ph), 133.95 (CH$_3$CH*C*), 132.8 (*C*COH), 130.6 (*C*HCCOH), 129.6 (Ph), 129.3 (Ph), 127.3 (OCS*C*), 125.23 (CH$_3$*C*H), 74.5 (*C*OH), 49.0 (COH*C*H$_2$), 16.5 (*C*H$_3$CCHCCOH), 13.7 (*C*H$_3$CH), 13.6 (*C*H$_3$CCOH); **IR** (thin film) v_{max} = 3409, 2915, 2857, 1699, 1440, 1021, 745 cm^{-1}; **HRMS (ESI)**: calculated for C$_{16}$H$_{20}$NaO$_2$S [M+Na]$^+$: 299.1076, found: 299.1078.

6.1.3.22 Synthesis of S-(2-acetamidoethyl) (4E,6E)-3-hydroxy-4,6-dimethylocta-4,6-dienethioate (215)
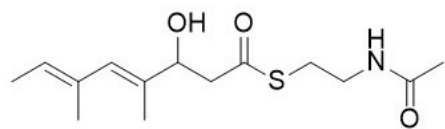
To a solution of S-phenyl (4E,6E)-3-hydroxy-4,6-dimethylocta-4,6-dienethioate (**214**) (0.25 g, 0.91 mmol) and *N*-acetyl cysteamine (**191**) (0.12 g, 1.01 mmol) in anhydrous DCM (7 mL) under argon at RT was added anhydrous Et$_3$N (0.13 mL, 0.91 mmol) followed by stirring for 6 h. The reaction was then concentrated and purified by silica column chromatography with a gradient of 0-70% EtOAc in petroleum ether to afford the

title compound (**215**) as a colourless oil (0.89 g, 35%): **TLC** (Rf = 0.30, 95:5 DCM : MeOH); ^{1}**H NMR** (500 MHz, CDCl$_3$): δ 5.94 (1H, s, C*H*CCOH), 5.81 (1H, br s, N*H*), 5.41 (1H, q, *J* = 6.8, C*H*CH$_3$), 4.50 (1H, m, *H*COH), 3.45 (2H, m, C*H*$_2$NH), 3.05 (2H, m, SC*H*$_2$), 2.86 (1H, dd, *J* = 15.0, 8.9, C*H*CO), 2.78 (1H, dd, *J* = 15.0, 3.8, C*H*CO), 2.45 (1H, d, 2.88, O*H*), 1.96 (3H, s, NHCOC*H*$_3$), 1.78 (3H, s, C*H*$_3$CCOH), 1.73 (3H, s, CH$_3$CC*H*CCOH), 1.68 (3H, d, *J* = 6.9, C*H*$_3$CH); ^{13}**C NMR** (125 MHz, CDCl$_3$): 199.0 (S*C*O), 170.4 (NH*C*O), 134.1 (*C*COH), 132.7 (*C*CHCCOH), 130.6 (*C*HCCOH), 125.4 (CH$_3$*C*HCCH$_3$), 74.7 (*C*OH), 49.6 (*C*H$_2$CO), 39.4 (*C*H$_2$NH), 28.9 (S*C*H$_2$), 23.2 (*C*H$_3$CONH), 16.5 (*C*H$_3$CCHCH$_3$), 13.7 (*C*H$_3$CH), 13.5 (*C*H$_3$CCOH); **IR** (thin film) v_{max} = 3294, 2921, 2858, 1686, 1656, 1548, 1372, 1290, 1073 cm^{-1}; **HRMS (ESI)**: calculated for C$_{14}$H$_{23}$NNaO$_3$S [M+Na]$^+$: 308.1291, found: 308.1293.

6.1.3.23 S-(2-acetamidoethyl)　　　(4E,6E)-4,6-dimethyl-3-oxoocta-4,6-dienethioate (216)
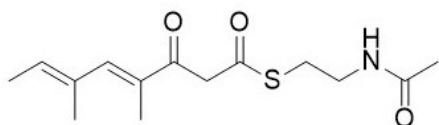
To a solution of S-(2-acetamidoethyl) (4E,6E)-3-hydroxy-4,6-dimethylocta-4,6-dienethioate (**215**) (0.05 g, 0.19 mmol) in EtOAc (11 mL) was added IBX (45% wt.) (0.35 g, 0.56 mmol) followed by stirring at reflux for 1 h. The reaction was then filtered through Celite and the filtrate was concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-100% EtOAc in cyclohexane afforded the title compound (**216**) as a light yellow oil (0.02 g, 35%) (as a 2:1 keto:enol tautomeric ratio): **TLC** (Rf = 0.40, EtOAc); ^{1}**H NMR** (500 MHz, CDCl$_3$): (**Keto form**) δ 6.94 (1H, s, C*H*CCO), 6.00 (1H, br s, N*H*), 5.85 (1H, q, *J* = 6.9, C*H*CH$_3$), 4.00 (2H, s, COC*H*$_2$CSO), 3.48-3.46 (2H, m, C*H*$_2$NH), 3.11-3.09 (2H, m, SC*H*$_2$), 1.97 (3H, s, C*H*$_3$CONH), 1.97 (3H, s, C*H*$_3$CCO), 1.90 (3H, s, C*H*$_3$CCHCCO), 1.79 (3H, d, *J* = 7.2, C*H*$_3$CH): (**Enol form**) δ 12.70 (1H, s, O*H*), 7.01 (1H, s, C*H*CCOH), 5.69-5.68 (1H, m, C*H*CSO), 5.68-5.67 (1H,

m, C*H*CCHCCOH), 1.92 (3H, s, C*H₃*CCO), 1.83 (3H, s, C*H₃*CCHCCOH), 1.76 (3H, d, J = 6.2, C*H₃*CH); **^{13}C NMR** (125 MHz, CDCl₃): (**Keto form**) 194.4 (O*C*CH₂*C*SO), 193.5 (*C*SO), 170.5 (*C*ONH), 146.5 (*C*HCCO), 133.6 (*C*COCH₂CSO), 133.6 (*C*HCCHCCO), 130.5 (*C*HCH₃), 52.9 (*C*H₂CSO), 39.3 (*C*H₂NH), 29.2 (S*C*H₂), 23.2 (*C*H₃CONH), 16.0 (*C*H₃CCHCCO), 14.2 (*C*H₃CH), 12.9 (*C*H₃CCO): (**Enol form**) 171.7 (*C*OH), 140.0 (*C*HCCOH), 134.0 (*C*HCH₃), 133.3 (*C*CHCCOH), 125.7 (*C*COH), 97.3 (*C*CSO), 16.3 (*C*H₃CCHCCOH), 14.3 (*C*H₃CH), 13.7 (*C*H₃CCOH); **IR** (thin film) v_{max} = 3278, 2927, 2856, 1652, 1616, 1557, 1372, 1289, 1074 cm^{-1}; **HRMS (ESI)**: calculated for C₁₄H₂₁NNaO₃S [M+Na]$^{+}$: 306.1134, found: 306.1129.

6.1.3.24 Synthesis of S-(2-acetamidoethyl) (2E,4E)-hexa-2,4-dienethioate (261)
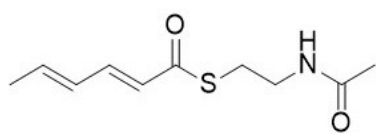
To a solution of potassium (2E,4E)-hexa-2,4-dienoate (0.10 g, 0.67 mmol), *N*-acetyl cysteamine (**191**) (0.09 g, 0.72 mmol) and DMAP (0.02 g, 0.15 mmol) in anhydrous THF (2 mL) at 0 °C under argon was added EDC.HCl (0.14 g, 0.72 mmol) followed by stirring at RT for 16 h. The reaction was then diluted with 1M HCl (5 mL) and extracted with DCM (4 x 5 mL). The combined organics were washed with saturated NaHCO₃₍aq.₎ (5 mL), dried over MgSO₄₍s₎, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 50-80% EtOAc in petroleum ether afforded the title compound (**261**) as a pale yellow solid (0.04 g, 31%): **TLC** (Rf = 0.14, 1:1 EtOAc : cyclohexane); **^{1}H NMR** (500 MHz, CDCl₃): δ 7.12 (1H, dd, J = 15.2, 10.7, C*H*CHCHCOS), 6.26 (1H, dq, J = 13.2, 6.6, C*H*CH₃), 6.17 (1H, dd, J = 15.1, 10.7, CH₃CHC*H*), 6.09 (1H, d, J = 15.0, C*H*COS), 5.89 (1H, br s, N*H*), 3.47 (2H, q, J = 6.1, NC*H₂*), 3.11 (2H, t, J = 6.3, SC*H₂*), 1.96 (3H, s, NHCOC*H₃*), 1.88 (2H, d, J = 6.6, C*H₃*CH). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[170]

6.1.3.25 Synthesis of methyl (4E,6E)-3-hydroxy-4,6-dimethylocta-4,6-dienoate (218)
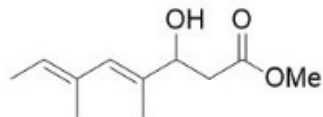
To stirring anhydrous THF (5 mL) at 70 °C under argon was added in quick succession activated zinc dust (0.20 g, 3.05 mmol), (2E,4E)-2,4-dimethylhexa-2,4-dienal (**203**) (0.21 g, 1.69 mmol) and then methyl bromoacetate (**217**) (0.29 mL, 3.05 mmol) followed by stirring at 70 °C for 5 min. After the exotherm and colour change, the reaction stirred at RT for 15 min followed by concentration. The residue was diluted in H_2O (10 mL) and extracted with EtOAc (4 x 10 mL). The combined organics were washed with 1M HCl (10 mL), H_2O (10 mL) and brine (15 mL) then dried over $MgSO_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-15% EtOAc in petroleum ether afforded the title compound (**218**) as a colourless oil (0.13 g, 40%): **TLC** (Rf = 0.26, 4:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 5.94 (1H, s, C*H*CCOH), 5.40 (1H, q, *J* = 5.4, CH$_3$C*H*), 4.48-4.42 (2H, m, C*H*OH), 3.72 (3H, s, OC*H*$_3$), 2.68 (1H, d, *J* = 2.6, O*H*), 2.65-2.52 (2H, m, C*H*$_2$CO$_2$), 1.78 (3H, s, C*H*$_3$CCOH), 1.73 (3H, s, C*H*$_3$CCHCCOH), 1.68 (3H, d, *J* = 6.8, C*H*$_3$CH); **^{13}C NMR** (125 MHz, CDCl$_3$): 173.2 (*C*O$_2$), 134.4 (*C*COH), 132.9 (*C*CHCCOH), 130.4 (*C*HCCOH), 125.2 (*C*HCH$_3$), 74.1 (*C*OH), 52.0 (*C*H$_3$O), 40.2 (*C*H$_2$CO$_2$), 16.7 (*C*H$_3$CCHCH$_3$), 13.8 (*C*H$_3$CH), 13.7 (*C*H$_3$CCOH); **IR** (thin film) v_{max} = 3437, 2951, 2917, 1732, 1437, 1161, 1016, 883 cm^{-1}; **HRMS (ESI)**: calculated for C$_{11}$H$_{18}$NaO$_3$ [M+Na]$^+$: 221.1149, found: 221.1148.

6.1.3.26 Alternative synthesis of S-(2-acetamidoethyl) (4E,6E)-3-hydroxy-4,6-dimethylocta-4,6-dienethioate (215)
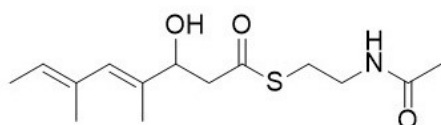
LiOH (80 mg, 3.33 mmol) was added to a solution of methyl (4E,6E)-3-hydroxy-4,6-dimethylocta-4,6-dienoate (**218**) (132 mg, 0.67 mmol) in 3:3:1 MeOH : THF : H$_2$O (5 mL) and stirred for 3 h. The reaction was then concentrated, diluted in H$_2$O (2 mL) and washed with diethyl ether (2 x 10 mL). The aqueous layer was then acidified using 1M HCl to pH 2 and extracted in EtOAc (5 x 8 mL). The combined organics were dried over MgSO$_{4(s)}$, filtered and concentrated to the crude carboxylic acid (**219**) (94 mg). To a solution of the carboxylic acid (**219**), *N*-acetyl cysteamine (**191**) (65 mg, 0.54 mmol) and DMAP (7 mg, 0.05 mmol) in anhydrous THF (3 mL) under argon at 0 °C was added DCC (116 mg, 0.56 mmol) followed by stirring at RT for 16 h. The reaction was then concentrated, resuspended in EtOAc (2 mL) and filtered through Celite. The filtrate was concentrated and purified by silica column chromatography with a gradient of 50-100% EtOAc in petroleum ether to afford the title compound (**215**) as a colourless oil (89 mg, 47%): **TLC** (Rf = 0.30, 95:5 DCM : MeOH); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 5.94 (1H, s, C*H*CCOH), 5.81 (1H, br s, N*H*), 5.41 (1H, q, *J* = 6.8, C*H*CH$_3$), 4.51-4.49 (1H, m, *H*COH), 3.46-3.44 (2H, m, C*H*$_2$NH), 3.06-3.04 (2H, m, SC*H*$_2$), 2.86 (1H, dd, *J* = 15.0, 8.9, C*H*CO), 2.78 (1H, dd, *J* = 15.0, 3.8, C*H*CO), 2.45 (1H, d, *J* = 2.9, O*H*), 1.96 (3H, s, NHCOC*H*$_3$), 1.78 (3H, s, C*H*$_3$CCOH), 1.73 (3H, s, CH$_3$CCHCCOH), 1.68 (3H, d, *J* = 6.9, C*H*$_3$CH); **^{13}C NMR** (125 MHz, CDCl$_3$): 199.0 (S*C*O), 170.4 (NH*C*O), 134.1 (*C*COH), 132.7 (*C*CHCCOH), 130.6 (*C*HCCOH), 125.4 (CH$_3$*C*HCCH$_3$), 74.7 (*C*OH), 49.6 (*C*H$_2$CO), 39.4 (*C*H$_2$NH), 28.9 (S*C*H$_2$), 23.2 (*C*H$_3$CONH), 16.5 (*C*H$_3$CCHCH$_3$), 13.7 (*C*H$_3$CH), 13.5 (*C*H$_3$CCOH); **IR** (thin film) v_{max} = 3294, 2921, 2858, 1686, 1656, 1548, 1372, 1290, 1073 cm^{-1}; **HRMS (ESI)**: calculated for C$_{14}$H$_{23}$NNaO$_3$S [M+Na]$^+$: 308.1291, found: 308.1293.

6.1.3.27 Synthesis of ethyl 4,4,4-trifluoro-2-methyl-3-oxobutanoate (231)
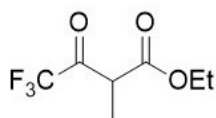
To a solution of ethyl propionate (2.00 mL, 17.4 mmol in anhydrous THF (20 mL) at RT under argon was added sodium hydride (60% in oil) (0.46 g, 19.2 mmol), the reaction was then warmed to 50 °C, ethyl trifluoroacetate (3.11 mL, 26.2 mmol) was added dropwise and the reaction was stirred at 50 °C for 16 h. The reaction was cooled to RT and added dropwise to ice cold 10% H_2SO_4 (25 mL), the mixture was extracted using EtOAc (5 x 10 mL). The combined organics were washed with brine (30 mL), dried over $MgSO_2$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-30% Et_2O in petroleum ether afforded the title compound (**231**) as an orange oil (1.82 g, 53%): **TLC** (Rf = 0.55, 3:1 petroleum ether : EtOAc); **^{1}H NMR** (300 MHz, $CDCl_3$): δ 4.22 (2H, q, J = 7.1, OCH_2CH$_3$), 3.09 (1H, q, J = 7.2, CHCH$_3$), 1.48 (3H, d, J = 7.2, CHCH_3), 1.27 (3H, t, J = 7.2, OCH$_2$CH_3); **^{19}F NMR** (282 MHz, $CDCl_3$): -77.2 (s, CF_3). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[171]

6.1.3.28 Synthesis of ethyl 4,4,4-trifluoro-3-hydroxy-2-methylbutanoate (232)
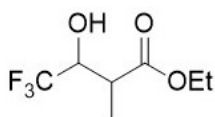
To a solution of ethyl 4,4,4-trifluoro-2-methyl-3-oxobutanoate (**231**) (1.22 g, 6.2 mmol) in anhydrous Et_2O (20 mL) at 0 °C under argon was added portion wise sodium borohydride (0.25 g, 6.49 mmol) over 10 mins followed by stirring at RT for 16 h. The reaction was cooled to 0 °C and quenched by dropwise addition of 1M HCl (10 mL). The reaction was then extracted with Et_2O (4 x 10 mL), the combined organics were dried over $MgSO_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-30% Et_2O in petroleum ether afforded the title compound (**232**) as a colourless oil (0.64 g, 52%), NMR appeared to show a single relative stereochemistry which was not further characterised: **TLC** (Rf = 0.42, 2:1 petroleum ether : Et_2O); **^{1}H NMR** (300 MHz, $CDCl_3$): δ 4.47-4.39 (1H, m, F$_3$CCH), 4.20 (2H, q, J = 7.1, OCH_2), 2.90-2.82 (1H, m, CHCO$_2$), 1.33 (3H, d, J = 7.3, CHCH_3), 1.29

(3H, d, J = 7.2, OCH$_2$CH_3); ^{19}**F NMR** (282 MHz, CDCl$_3$): -76.4 (d, J = 7.2, CF_3). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[172]

6.1.3.29 Synthesis of ethyl (E)-4,4,4-trifluoro-2-methylbut-2-enoate (234)
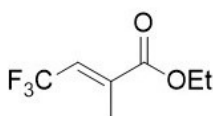
To a solution of ethyl 4,4,4-trifluoro-3-hydroxy-2-methylbutanoate (**232**) (0.64 g, 3.20 mmol) and anhydrous Et$_3$N (1.11 mL, 7.99 mmol) in anhydrous DCM (8 mL) under argon at 0 °C was added dropwise mesyl chloride (0.31 mL, 3.99 mmol) followed by stirring at RT for 16 h. The reaction was then diluted with H$_2$O (10 mL), acidified to pH 2 using 1M HCl and then extracted with DCM (3 x 10 mL). The combined organics were washed with saturated NaHCO$_{3(aq.)}$ (15 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to afford the mesylated intermediate (**233**) (0.71 g). To a solution of the mesylate (**233**) (0.71 g) in anhydrous DCM (5 mL) at RT under argon was added dropwise 1,8-Diazabicyclo[5.4.0]undec-7-ene (0.48 mL, 3.20 mmol) followed by stirring for 2 h. The reaction was then concentrated and purified by silica column chromatography with a gradient of 0-30% Et$_2$O in petroleum ether to afford the title compound (**234**) as a colourless oil (0.37 g, 63%): **TLC** (Rf = 0.31, 9:1 petroleum ether : Et$_2$O); ^{1}**H NMR** (300 MHz, CDCl$_3$): δ 6.67 (1H, q, J = 8.9, F$_3$CCH), 4.27 (2H, q, J = 7.1, OCH_2CH$_3$), 2.09 (3H, s, CH_3), 1.33 (3H, t, J = 7.1 OCH$_2$CH_3); ^{19}**F NMR** (282 MHz, CDCl$_3$): 58.8–58.9 (m, CF_3). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[173]

6.1.3.30 Synthesis of (E)-4,4,4-trifluoro-2-methylbut-2-enoic acid (235)
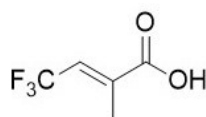
To a solution of ethyl (E)-4,4,4-trifluoro-2-methylbut-2-enoate (**234**) (0.61 g, 3.37 mmol) in THF (15 mL) and H_2O (4 mL) at RT was added LiOH (0.81 g, 33.4 mmol) followed by stirring for 2 h at 80 °C. The reaction was then cooled to 0 °C, acidified to pH 2 using 6M HCl and extracted with EtOAc (4 x 10 mL). The combined organics were dried over $MgSO_{4(s)}$, filtered and concentrated to give the title compound (**235**) as a colourless oil which solidified on standing (0.49 g, 95%): **^{1}H NMR** (500 MHz, $CDCl_3$): δ 6.80 (1H, q, J = 7.7, C*H*), 2.11 (3H, s, C*H*$_3$); **^{13}C NMR** (125 MHz, $CDCl_3$): 170.8 (CO_2), 138.6 (CH_3C), 128.2 (q, J = 35.1, *C*H), 122.5 (q, J = 271.9, *C*F_3), 13.3 (*C*H$_3$); **^{19}F NMR** (282 MHz, $CDCl_3$): -59.7–59.6 (m, C*F*$_3$); **IR** (thin film) v_{max} = 3466, 2917, 2618, 1708, 1665, 1425, 1126, 1094 cm^{-1}; **HRMS (ESI)**: calculated for $C_5H_4F_3O_2$ [M-H]$^-$: 153.0169, found: 153.0162.

6.1.3.31 S-(2-acetamidoethyl) (E)-4,4,4-trifluoro-2-methylbut-2-enethioate (226)
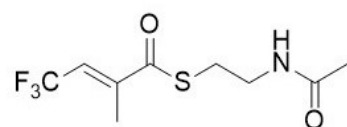
To a solution of (E)-4,4,4-trifluoro-2-methylbut-2-enoic acid (**235**) (0.49 g, 3.19 mmol) and DMAP (0.04 g, 0.32 mmol) in anhydrous THF (15 mL) under argon at 0 °C was added *N*-acetyl cysteamine (**191**) (0.38 g, 3.19 mmol) then DCC (0.72 g, 3.51 mmol) followed by stirring for 16 h at RT. The reaction was concentrated, resuspended in cold EtOAc (15 mL) and filtered through Celite. The filtrate was concentrated and purified by silica column chromatography with a gradient of 50-100% EtOAc in petroleum ether to afford the title compound (**226**) as a cloudy oil (0.62 g, 76%): **TLC** (Rf = 0.33, 3:1 EtOAc : petroleum ether); **^{1}H NMR** (500 MHz, $CDCl_3$): δ 6.55 (1H, q, J = 8.0, C*H*), 5.77 (1H, br s, N*H*), 3.48 (2H, q, J = 6.3, C*H*$_2$NH), 3.14 (2H, t, J = 6.5, SC*H*$_2$), 2.16-2.13 (3H, m, CC*H*$_3$), 1.98 (3H, s, NHCOC*H*$_3$); **^{13}C NMR** (125 MHz, $CDCl_3$): 193.3 (O*C*S), 170.5 (NH*C*O), 146.0 (*C*CSO), 123.9 (q, J = 35.4, CF$_3$*C*H), 122.1 (*C*F$_3$), 39.4 (*C*H$_2$NH), 29.4 (S*C*H$_2$), 23.5 (NHCO*C*H$_3$), 13.9 (*C*H$_3$CCSO); **^{19}F NMR** (282 MHz, $CDCl_3$): -58.8–58.9

(m, C*F*₃); **IR** (thin film) υ_{max} = 3286, 3076, 2933, 1653, 1554, 1216, 1116, 1032 cm^{-1}; **HRMS (ESI)**: calculated for $C_9H_{12}F_3NaO_2S$ [M+Na]$^+$: 278.0433, found: 278.0432.

6.1.3.32 Synthesis of (R)-3-(2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamido)propanoic acid (251)
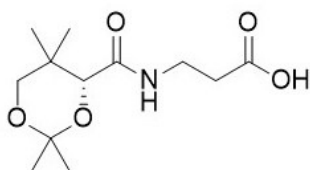
A solution of D-Pantothenic acid hemicalcium salt (**250**) (2.00 g, 8.39 mmol), *p*-toluene sulphonic acid monohydrate (1.92 g, 10.1 mmol) and 3Å molecular sieves (20 g) in acetone (100 mL) under argon at RT was stirred for 16 h. The reaction was filtered through Celite and the filtrate was concentrated then re-dissolved in EtOAc (50 mL). The organic solution was washed with brine (2 x 20 mL), dried over MgSO₄₍ₛ₎, filtered and concentrated to give the crude product. The crude product was dissolved in minimal hot EtOAc and diluted with cyclohexane. The solvent was decanted and the remaining solid dried to give the title compound (**251**) as a white solid (1.51 g, 70%). **¹H NMR** (300 MHz, CDCl₃): δ 6.99 (1H, br s, N*H*), 4.10 (1H, s, C*H*CONH), 3.68 (1H, d, *J* = 11.6, OC*H*₂), 3.65-3.42 (2H, m, C*H*₂NH), 3.29 (1H, d, *J* = 11.6, OC*H*₂), 2.65-2.63 (2H, m, C*H*₂CO₂H), 1.46 (3H, s, OCC*H*₃), 1.43 (3H, s, OCC*H*₃), 1.05 (3H, s, OCH₂CC*H*₃), 0.98 (3H, s, OCH₂CC*H*₃). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[174]

6.1.3.33 Synthesis of (R)-N-(3-((2-mercaptoethyl)amino)-3-oxopropyl)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamide (252)
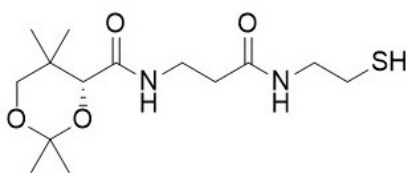
To a solution of (R)-3-(2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamido)propanoic acid (**251**) (1.51 g, 5.81 mmol) in THF (40 mL) under argon at RT was added CDI (1.41 g, 8.72 mmol). After stirring for 1 h, cysteamine hydrochloride (**198**) (1.15 g, 8.72 mmol) was added, after 16 h the reaction was concentrated. The crude product was dissolved in DCM (40 mL), washed with brine (10 mL, dried over MgSO$_{4(s)}$, filtered and concentrated. Purification by silica column chromatography isocratically with 100% EtOAc afforded the title compound (**252**) as a white solid (1.02 g, 55%): **TLC** (Rf = 0.18, EtOAc); **^{1}H NMR** (300 MHz, CDCl$_3$): δ 6.98 (1H, br s, OCHCON*H*), 6.22 (1H, br s N*H*CH$_2$CH$_2$SH), 4.01 (O*H*CONH), 3.67 (1H, d, *J* = 11.7, OC*H*), 3.61-3.32 (4H, m, NC*H$_2$*), 2.66 (2H, q, *J* = 8.2, C*H$_2$*SH), 2.46 (2H, t, *J* = 6.3, C*H$_2$*CONH), 3.27 (1H, d, *J* = 11.7, OC*H*), 1.46 (3H, s, OCC*H$_3$*), 1.42 (3H, s, OCC*H$_3$*), 1.36 (1H, t, *J* = 8.4, S*H*), 1.04 (3H, s, OCH$_2$CC*H$_3$*), 0.97 (3H, s, OCH$_2$CC*H$_3$*). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[175]

6.1.3.34 Synthesis of S-(2-(3-((R)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamido)propanamido)ethyl) (4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-dienethioate (255)
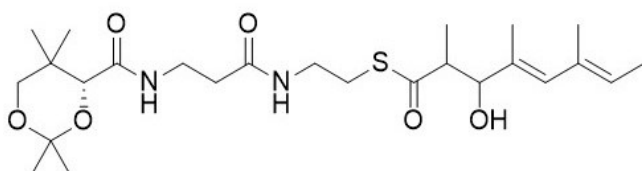
A solution of (4E,6E)-3-hydroxy-2,4,6-trimethyl-1-(2-thioxothiazolidin-3-yl)octa-4,6-dien-1-one (**254**) (0.23 g, 0.76 mmol), (R)-N-(3-((2-mercaptoethyl)amino)-3-oxopropyl)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamide (**252**) (0.29 g, 0.91 mmol), K$_2$CO$_3$ (0.31 g, 2.28 mmol) in dry MeCN (5 mL) under argon at RT was stirred for 2 h followed by concentration. Purification by silica column chromatography with a gradient of 50-100% EtOAc in petroleum ether afforded the title compound (**255**) as a colourless sticky oil (0.23 g, 61 %) NMR appeared to show a single relative stereochemistry which was not further characterised: **TLC** (Rf = 0.16, 1:3 cyclohexane : EtOAc); **^{1}H NMR** (500 MHz,

CDCl$_3$): δ 6.98 (1H, br s, OHCCON*H*), 6.25-6.16 (1H, m, N*H*CH$_2$CH$_2$S), 5.98 (1H, s, C*H*CHOH), 5.39 (1H, q, *J* = 6.5, CH$_3$C*H*CCH$_3$), 4.42 (1H, q, *J* = 4.3, CO*H*), 4.07 (1H, d, *J* = 4.5, OC*H*CONH), 3.59-3.31 (2H, m, NHC*H$_2$*CH$_2$CONH), 3.59-3.31 (2H, m, NHC*H$_2$*CH$_2$S), 3.70 (1H, d, J= 11.7, OC*H*), 3.30 (1H, d, *J* = 11.7, OC*H*), 3.14-2.96 (2H, m, NHCH$_2$C*H$_2$*S), 2.95-2.89 (1H, m, SCOC*H*), 2.40 (2H, t, *J* = 6.6, C*H$_2$*CONH), 1.76 (3H, s, COHCC*H$_3$*), 1.73 (s, 3H, COHHCCHCC*H$_3$*), 1.68 (3H, d, *J* = 6.8, C*H$_3$*CHCCH$_3$), 1.46 (3H, s, (C*H$_3$*)$_2$CO$_2$), 1.42 (3H, s, (C*H$_3$*)$_2$CO$_2$), 1.02 (OCH$_2$C(C*H$_3$*)$_2$, 0.96 (OCH$_2$C(C*H$_3$*)$_2$; **^{13}C NMR** (125 MHz, CDCl$_3$): 203.1 (*C*SO), 203.0 (*C*SO), 171.2 (*C*ONH), 171.1 (*C*ONH), 133.1 (COHCCH*C*), 133.1 (COH*C*CH), 130.5 COHCC*H*), 130.4 COHCC*H*), 124.9 (COHCCHCC*H*), 124.8 (COHCCHC*C*H), 99.1 (*C*O$_2$), 77.7 (*C*OH), 77.6 (*C*OH), 77.3 (O*C*CONH), 77.2 (O*C*CONH), 71.6 (O*C*H$_2$C(CH$_3$)$_2$), 51.7 (SCO*C*H), 39.6 (NH*C*H$_2$CH$_2$CONH), 39.5 (NH*C*H$_2$CH$_2$CONH), 36.4 (NHCH$_2$*C*H$_2$CONH), 36.3 (NHCH$_2$*C*H$_2$CONH), 35.3 (NHCH$_2$*C*H$_2$SCO), 35.2 (NHCH$_2$*C*H$_2$SCO), 33.1 (OCH$_2$*C*(CH$_3$)$_2$), 29.6 (O$_2$C(*C*H$_3$)$_2$), 28.6 (S*C*H$_2$), 28.5 (S*C*H$_2$), 22.3 (OCH$_2$C(*C*H$_3$)$_2$), 19.0 (OCH$_2$C(*C*H$_3$)$_2$), 18.8 (O$_2$C(*C*H$_3$)$_2$), 16.8 (*C*H$_3$CCHCH$_3$), 15.0 (COHC*C*H$_3$), 14.9 (COHC*C*H$_3$), 13.8 (*C*H$_3$CHCCH$_3$), 10.8 (SCOCH*C*H$_3$), 10.7 (SCOCH*C*H$_3$); **IR** (thin film) v_{max} = 3308, 2990, 2935, 2872, 1651, 1525, 1377, 1096, 900 cm^{-1}; **HRMS (ESI)**: calculated for C$_{25}$H$_{42}$N$_2$NaO$_6$S [M+Na]$^+$: 521.2656, found: 521.2659.

6.1.3.35 S-(2-(3-((R)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamido)propanamido)ethyl) (4E,6E)-2,4,6-trimethyl-3-oxoocta-4,6-dienethioate (256)
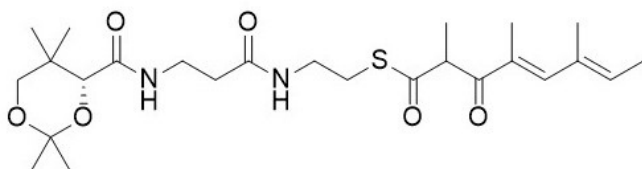
A solution of S-(2-(3-((R)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamido)propanamido)ethyl) (4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-

dienethioate (**255**) (0.23 g, 0.46 mmol) and IBX (0.51 g, 1.81 mmol) in EtOAc (20 mL) under argon was stirred at 80 °C for 3 h. The reaction was cooled, filtered through Celite, concentrated and purified by silica column chromatography with a gradient of 50-100% EtOAc in petroleum ether to afford the title compound (**256**) as a pale yellow oil (0.11 g, 48%), NMR appeared to show a single relative stereochemistry which was not further characterised: **TLC** (Rf = 0.39, EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.06-7.04 (1H, m, OHCCON*H*), 7.03 (1H, s, COCC*H*) 6.11 (1H, br s, CON*H*), 5.84 (1H, q, *J* = 7.0, CH$_3$C*H*C), 4.46 (1H, q, *J* = 6.9, SCOC*H*), 4.07 (1H, s, OC*H*CONH), 3.61 (1H, d, *J* = 11.7, OC*H*), 3.60-3.34 (4H, m, NC*H$_2$*), 3.21 (1H, d, *J* = 11.7, OC*H*), 3.13-2.95 (2H, m, C*H$_2$*S), 2.41 (2H, t, *J* = 5.9, C*H$_2$*CONH), 1.98 (3H, s, C*H$_3$*CHCCHCCH$_3$), 1.90 (3H, s, CH$_3$CHCC*H$_3$*), 1.79 (3H, d, *J* = 7.0, C*H$_3$*CHCCH$_3$), 1.46 (6H, s, O$_2$C(C*H$_3$*)$_2$), 1.43-1.40 (3H, m, SCOCHC*H$_3$*), 1.41 (O$_2$C(C*H$_3$*)$_2$), 1.04 (OCH$_2$C(C*H$_3$*)$_2$), 0.97 (OCH$_2$C(C*H$_3$*)$_2$); **^{13}C NMR** (125 MHz, CDCl$_3$): 197.6 (SCOCH*C*O), 197.5 (SCOCH*C*O), 197.3 (S*C*O), 197.2 (S*C*O), 171.5 (NHCH$_2$CH$_2$*C*ONH), 170.2 (OCH*C*ONH), 145.6 (CH$_3$CHC*C*H), 133.8 (CH$_3$*C*HCCH$_3$), 133.5 (CH$_3$CHCCH*C*), 133.3 (CH$_3$CHC*C*H$_3$), 99.2 (*C*O$_2$), 77.3 (O*C*HCONH), 71.6 (*C*H$_2$OCO), 55.1 (SCO*C*H), 39.5 (NH*C*H$_2$CH$_2$CONH), 36.0 (*C*H$_2$CONH), 34.9 (NH*C*H$_2$CH$_2$S), 33.1 (*C*(CH$_3$)$_2$), 29.6 (O$_2$C(*C*H$_3$)$_2$), 28.8 (*C*H$_2$S), 22.3 (OCH$_2$C(*C*H$_3$)$_2$), 18.8 (O$_2$C(*C*H$_3$)$_2$), 16.2 (CH$_3$CHC*C*H$_3$), 15.2 (SCOCH*C*H$_3$), 14.5 (*C*H$_3$CHCCH$_3$), 13.5 (CH$_3$CHCCHC*C*H$_3$); **IR** (thin film) v_{max} = 3305, 2988, 2936, 2871, 1654, 1525, 1197, 1097 cm^{-1}; **HRMS (ESI):** calculated for C$_{25}$H$_{40}$N$_2$NaO$_6$S [M+Na]$^+$: 519.2499, found: 519.2499.

6.1.3.36 Synthesis of S-(2-(3-((R)-2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl) (4E,6E)-2,4,6-trimethyl-3-oxoocta-4,6-dienethioate (245)
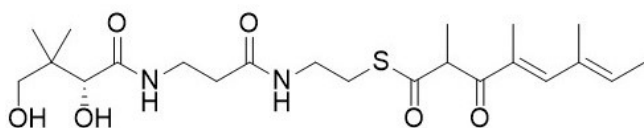
A solution of S-(2-(3-((R)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamido)propanamido)ethyl) (4E,6E)-2,4,6-trimethyl-3-oxoocta-4,6-dienethioate (**256**) (0.11 g, 0.22 mmol) in acetic acid (4 mL) and H_2O (2 mL) was stirred at RT for 16 h. The reaction was then concentrated and purified by silica column chromatography with a gradient of 0-10% MeOH in DCM to afford the title compound (**245**) (0.60 g, 86%) as a colourless oil, NMR shows approximately 1:1 ratio of both diastereoisomers with respect to the malonate chiral centre: **TLC** (Rf = 0.34, 95:5 DCM : MeOH); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.35 (1H, br s, HOCCON*H*), 7.03 (1H, s, COCC*H*), 6.25 (1H, br s, CON*H*), 5.87 (1H, q, *J* = 6.8, CH$_3$C*H*CCHCCO), 4.48 (1H, q, *J* = 6.9, SCOC*H*), 4.01-3.99 (1H, m, HOC*H*CONH), 3.86 (1H, br s, *H*OCHCONH), 3.60-3.41 (1H, m, *H*OCH$_2$), 3.60-3.41 (2H, m, NHC*H*$_2$CH$_2$CONH), 3.60-3.41 (2H, m, NHC*H*$_2$CH$_2$S), 3.50-3.48 (2H, m, HOC*H*$_2$), 3.20-3.17 (2H, m, C*H*$_2$S), 2.41 (2H, t, *J* = 5.6, C*H*$_2$CONH), 1.97 (3H, s, C*H*$_3$CCOCH), 1.91 (COCCHCC*H*$_3$), 1.80 (3H, d, *J* = 6.9, C*H*$_3$CHCCH$_3$), 1.41 (1H, d, *J* = 3.3, SCOCHC*H*$_3$), 1.40 (1H, d, *J* = 3.3, SCOCHC*H*$_3$), 1.04 (3H, s, C(C*H*$_3$)$_2$), 1.03 (3H, s, C(C*H*$_3$)$_2$), 0.93 (3H, s, C(C*H*$_3$)$_2$), 0.92 (3H, s, C(C*H*$_3$)$_2$); **^{13}C NMR** (125 MHz, CDCl$_3$): 198.4 (SCOCH*C*O), 198.3 (SCOCH*C*O), 197.9 (S*C*O), 173.6 (HOCH*C*ONH), 172.0 (NHCH$_2$CH$_2$*C*ONH), 171.9 (NHCH$_2$CH$_2$*C*ONH), 146.1 (COC*C*H), 134.3 (COCCHC*C*H), 134.2 (COCCHC*C*H), 133.6 (OC*C*), 133.5 (OC*C*), 133.1 (COCCH*C*), 133.0 (COCCH*C*), 77.9 (HO*C*H), 77.8 (HO*C*H), 71.1 (HO*C*H$_2$), 55.0 (SCOC*H*), 54.9 (SCOC*H*),39.6 (NH*C*H$_2$CH$_2$CONH), 35.7 (*C*H$_2$CONH), 35.2 (*C*H$_2$CH$_2$S), 35.1 (*C*H$_2$CH$_2$S), 33.3 (HOCH$_2$*C*COH), 28.9 (*C*H$_2$S), 28.8 (*C*H$_2$S), 22.0 (C(*C*H$_3$)$_2$), 21.9 (C(*C*H$_3$)$_2$), 20.5 (C(*C*H$_3$)$_2$), 20.4 (C(*C*H$_3$)$_2$), 16.2 (*C*H$_3$CCHCCO), 15.3 (SCOCH*C*H$_3$), 15.2 (SCOCH*C*H$_3$), 14.5 (*C*H$_3$CHCCHCCO), 13.5 (COC*C*H$_3$); **IR** (thin film) v_{max} = 3327, 2935, 2874, 1649, 1529, 1443, 1041, 956 cm^{-1}; **HRMS (ESI)**: calculated for C$_{22}$H$_{36}$N$_2$NaO$_6$S [M+Na]$^+$: 479.2186, found: 479.2188.

6.1.3.37 Synthesis of 1-(2-thioxothiazolidin-3-yl)propan-1-one (253)
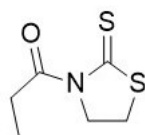
A solution of thiazolidine-2-thione (4.76 g, 40.0 mmol), acetic anhydride (18.9 mL, 200 mmol) and pyridine (19.4 mL, 240.1 mmol) was stirred at 115 °C for 1 h. The solution was then cooled, diluted with H_2O (100 mL) and extracted with DCM (3 x 30 mL). The combined organic extracts were washed with H_2O (2 x 30 mL) and then saturated $NaHCO_{3(aq.)}$ (30 mL), dried over $MgSO_{4(s)}$, filtered and vigorously concentrated. The crude was purified by silica column chromatography to give the title compound (**253**) as a yellow oil (6.28 g, 78%): **TLC** (Rf = 0.38, 1:3 EtOAc : cyclohexane); **^{1}H NMR** (300 MHz, CDCl$_3$): δ 4.59 (2H, t, *J* = 7.5, NC*H*$_2$), 3.31-3.22 (2H, m, SC*H*$_2$), 3.31-3.22 (2H, m, COC*H*$_2$), 1.18 (3H, t, *J* = 7.2, C*H*$_3$CH$_2$). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[176]

6.1.3.38 Synthesis of (4E,6E)-3-hydroxy-2,4,6-trimethyl-1-(2-thioxothiazolidin-3-yl)octa-4,6-dien-1-one (254)
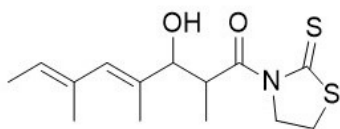
To a solution of 1-(2-thioxothiazolidin-3-yl)propan-1-one (**253**) (0.48 g, 2.74 mmol) in anhydrous DCM (8 mL) under argon at -40 °C was added dropwise 1M TiCl$_4$ in DCM (2.90 mL, 2.90 mmol). DIPEA (0.52 mL, 2.90 mmol) was added after 5 mins followed by stirring at -40 °C for 2 h. The reaction was cooled to -78 °C and a solution of and (2E,4E)-2,4-dimethylhexa-2,4-dienal (**203**) (0.20 g, 1.61 mmol) in anhydrous DCM (3 mL) was added dropwise. TLC confirmed completion after 40 mins and the reaction was quenched with dropwise addition of saturated $NH_4Cl_{(aq)}$ (10 mL). The organic layer was then separated and the aqueous extracted with DCM (3 x 10 mL). The combined organic extracts were dried over $MgSO_{4(s)}$, filtered and concentrated to give the crude product Purification by silica column chromatography with a gradient of 0-30% EtOAc in cyclohexane afforded the title compound (**254**) as a sticky yellow oil (0.18 g, 38%), NMR appeared to show a single relative stereochemistry which was not further characterised:

TLC (Rf = 0.62, 1:1 EtOAc : cyclohexane); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 5.94 (1H, s, C*H*CCOH), 5.40 (1H, q, J = 6.8, CC*H*CH$_3$), 4.80-4.78 (1H, m, C*H*CO), 4.51-4.49 (2H, m, NC*H$_2$*), 4.37-4.35 (1H, m, *H*COH), 3.30-3.27 (2H, m, SC*H$_2$*, 2.47 (1H, d, J = 2.5, O*H*), 1.75 (3H, s, C*H$_3$*CCOH), 1.74 (3H, s, C*H$_3$*CCHCCOH), 1.68 (3H, d, J = 6.8, C*H$_3$*CHC), 1.20 (3H, d, J = 6.9, C*H$_3$*CHCO); **^{13}C NMR** (125 MHz, CDCl$_3$): 201.7 (*C*S), 178.7 (N*C*O), 133.0 (*C*HCOH), 133.0 (*C*CHCCOH), 130.3 (*C*HCCOH), 124.8 (*C*HCCHCCOH), 77.2 (*C*OH), 56.5 (N*C*H$_2$), 42.6 (*C*HCO), 28.5 (S*C*H$_2$), 16.7 (*C*H$_3$CCHCCOH), 14.7 (*C*H$_3$CCOH), 13.7 (*C*H$_3$CHC), 11.1 (*C*H$_3$CHCO); **IR** (thin film) v_{max} = 3439, 2973, 2929, 2854, 1696, 1496, 1412, 1276, 1212, cm^{-1}; **HRMS (ESI)**: calculated for C$_{14}$H$_{21}$NNaO$_2$S$_2$ [M+Na]$^+$: 322.0906, found: 322.0907.

6.1.3.39 Synthesis of S-(2-acetamidoethyl) 2,2-difluoroethanethioate (259)
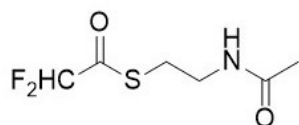
To a solution of *N*-acetyl cysteamine (**191**) (0.17 mL, 1.59 mL), dimethylamino pyridine (DMAP) (0.08 g, 0.64 mmol) and difluoroacetic acid (**264**) (0.10 mL, 1.60 mmol) in DCM (12 mL) was added dicyclohexyl carbodiimide (DCC) (0.40 g, 1.94 mmol) under argon at RT. After stirring for 16 h, the reaction was concentrated, re-dissolved in cold EtOAc (10 mL) and filtered under vacuum. The filtrate was concentrated to dryness and purified by silica column chromatography with a gradient of 25-100% EtOAc in petroleum ether to afford the title compound (**259**) as a cloudy oil (0.07 g, 22%): **TLC** (Rf = 0.35, EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 5.89 (1H, s, C*H*CCH$_2$OH), 5.88 (1H, t, J = 54.0, CF$_2$*H*), 5.76 (1H, br s, J = 5.4, N*H*), 3.49 (2H, q, J = 6.3, C*H$_2$*NH) , 3.19 (2H, t, J = 6.4, SC*H$_2$*), 1.99 (3H, s, C*H$_3$*); **^{13}C NMR** (125 MHz, CDCl$_3$): 191.7 (t, J = 29.6, O*C*S), 170.4 (NH*C*O), 108.9 (t, J = 254.3, *C*F$_2$H), 38.8 (*C*H$_2$NH), 28.2 (S*C*H$_2$), 23.2 (*C*H$_3$); **^{19}F NMR** (282 MHz, CDCl$_3$): -123.7 (d, J = 53.9); **IR** (thin film) v_{max} = 3277,

3083, 2931, 2855, 1651, 1549, 1092 cm^{-1}; **HRMS (ESI)**: calculated for $C_6F_2H_9NNaO_2S$ [M+Na]$^+$: 220.0214, found: 220.0214.

6.1.3.40 Synthesis of 2-methyl-3-oxobutanoic acid (266)
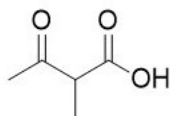
To ethyl 2-methylacetoacetate (**265**) (1.00 mL, 7.07 mmol) was added 0.5M NaOH (15 mL) *via* dropping funnel, followed by stirring of the mixture for 16 h at RT. The reaction was then washed with Et$_2$O (2 x 10 mL), acidified to pH 2 using 1M HCl, saturated with NaCl$_{(aq)}$ and extracted using EtOAc (4 x 10 mL). The combined organics were dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude compound. Purification by silica column chromatography with a gradient of 0-20% EtOAc in petroleum ether afforded the title compound (**266**) as a colourless oil (0.25 g, 31%): **TLC** (Rf = 0.10, 1:1 petroleum ether : EtOAc); **^{1}H NMR** (300 MHz, CDCl$_3$): δ 3.57 (1H, t, *J* = 7.2, CH$_3$C*H*), 2.31 (3H, s, C*H$_3$*CO), 1.40 (3H, d, *J* = 7.2, CHC*H$_3$*). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[177]

6.1.3.41 Synthesis of S-(2-acetamidoethyl) 2-methyl-3-oxobutanethioate (260)
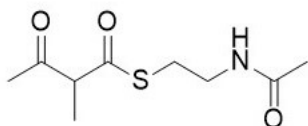
To a solution of 2-methyl-3-oxobutanoic acid (**266**) (0.25 g, 2.17 mmol), *N*-acetyl cysteamine (**191**) (0.23 g, 2.39 mmol) and DMAP (0.03 g, 0.22 mmol) in anhydrous DCM (10 mL) at RT was added DCC (0.49 g, 2.39 mmol) in one portion followed by stirring for 16 h. The reaction was then concentrated, dissolved in EtOAc (10 mL), filtered through Celite and the filtrate concentrated to give the crude product. Purification by silica column chromatography with a gradient of 50-100% EtOAc in petroleum ether

afforded the title compound (**260**) as a colourless oil (0.25 g, 53%): **TLC** (Rf = 0.30, EtOAc); **^{1}H NMR** (300 MHz, CDCl$_3$): δ 5.89 (1H, br s, N*H*), 3.77 (1H, q, *J* = 7.1, CH$_3$C*H*), 3.55-3.35 (2H, m, C*H$_2$*N), 3.16-2.97 (2H, m, SC*H$_2$*), 2.24 (3H, s, C*H$_3$*CO), 1.98 (3H, s, NHCOC*H$_3$*), 1.40 (2H, d, *J* = 7.1, CHC*H$_3$*). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[178]

6.1.3.42 Synthesis of 5-(hydroxymethylene)-2,2-dimethyl-1,3-dioxane-4,6-dione (270)
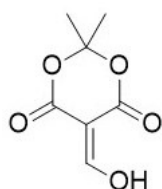
A solution of Meldrum's acid (**268**) (10.0 g, 69.4 mmol) in triethylorthoformate (38.0 mL, 228 mmol) was stirred under argon at 85 °C for 4 h. The reaction was then concentrated to give the crude enol ether product (**269**) which was stirred vigorously in 2M HCl (120 mL) at RT for 1 h. The solution was then extracted with Et$_2$O (4 x 30 mL), the combined organic extracts were washed with brine (100 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to give the title compound (**270**) as an orange solid which was used without further purification (6.71 g, 56%). **^{1}H NMR** (300 MHz, CDCl$_3$): δ 8.55 (1H, s, C*H*OH), 1.77 (6H, s, 2 x C*H$_3$*). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[179]

6.1.3.43 Synthesis of 5-fluoro-2,2-dimethyl-1,3-dioxane-4,6-dione (271)
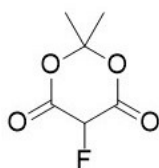
To a solution of 5-(hydroxymethylene)-2,2-dimethyl-1,3-dioxane-4,6-dione (**270**) (1.00 g, 5.81 mmol) in anhydrous MeCN (40 mL) under argon at -40 °C was added in one portion Selectfluor (2.06 g, 5.81 mmol), the reaction was then stirred at -40 °C for 2 h then allowed to warm to RT over a further 3 h. The reaction was then diluted with DCM (24 mL), 1M HCl (18 mL) and brine (50 mL), the resulting solution was extracted using DCM (4 x 10 mL). The combined organic extracts were dried over $MgSO_{4(s)}$, filtered and concentrated. The resulting crude product was triturated with Et_2O (20 mL) and filtered under suction. The filter cake was dried to give the title compound (**271**) as a white solid (0.65 g, 69%). **^{1}H NMR** (300 MHz, $CDCl_3$): δ 5.65 (1H, d, J = 44.9, CHF), 1.86 (3H, s, H_3CCO_2), 1.84 (3H, s, H_3CCO_2); **^{19}F NMR** (282 MHz, $CDCl_3$): -205.5 (d, J = 44.9, CFH). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[180]

6.1.3.44 Synthesis of S-phenyl 2-fluoroethanethioate (272)
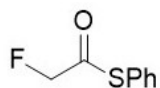
A solution of 5-fluoro-2,2-dimethyl-1,3-dioxane-4,6-dione (**271**) (1.00 g, 6.17 mmol) and thiophenol (1.10 mL, 6.79 mmol) in toluene (30 mL) was stirred under argon at 110 °C for 16 h. The reaction was then concentrated and purified by silica column chromatography with a gradient of 0-15% EtOAc in petroleum ether to give the title compound (**272**) as a colourless oil (0.95 g, 89%): **TLC** (Rf = 0.79, 1:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, $CDCl_3$): δ 7.45 (5H, m, Ph), 5.00 (2H, d, J = 47.1, CH_2F); **^{13}C NMR** (125 MHz, $CDCl_3$): 195.6 (d, J = 25.2, OCS), 134.9 (Ph), 129.9 (Ph), 129.5 (Ph), 125.1 (d, J = 5.1, CS), 84.7 (d, J = 84.7, CH_2F); **^{19}F NMR** (282 MHz, $CDCl_3$): -205.5 (S, CH_2F); **IR** (thin film) v_{max} = 3060, 2926, 1701, 1477, 1340 cm^{-1}; **HRMS (ESI)**: calculated for C_8FH_7NaOS [M+Na]$^+$: 193.0094, found: 193.0096.

6.1.3.45 Synthesis of S-phenyl (4E,6E)-2-fluoro-3-hydroxy-4,6-dimethylocta-4,6-dienethioate (273)
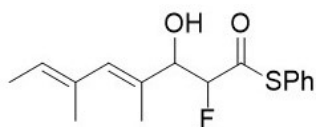
To a solution of S-phenyl 2-fluoroethanethioate (**272**) (0.26 g, 1.54 mmol) and (2E,4E)-2,4-dimethylhexa-2,4-dienal (**203**) (0.21 g, 1.69 mmol) in DCM (8 mL) under argon at RT was added in one portion $MgBr_2.OEt_2$ (0.58 g, 2.16 mmol). After stirring for 5 mins, DIPEA (0.54 mL, 3.08 mmol) was added dropwise followed by stirring at RT for 1 h. The reaction was then diluted with EtOAc (10 mL) and H_2O (10 mL), the aqueous phase was adjusted to pH 5 using 1M HCl and the biphasic solution was stirred vigorously for 5 mins. The organic layer was separated and aqueous extracted using EtOAc (4 x 5 mL), the combined organic extracts were washed with brine (10 mL), dried over $MgSO_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-20% EtOAc in petroleum ether afforded the title compound (**273**) as a yellow oil (0.33 g, 72%) with a relative stereochemistry ratio of 1:1 *syn* and *anti* as shown by NMR: **TLC** (Rf = 0.40, 4:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, $CDCl_3$): δ 7.50-7.35 (5H, m, Ph), 6.02 (1H, s, C*H*CCOH), 5.97 (1H, s, C*H*CCOH), 5.50-5.48 (1H, m, CH_3C*H*), 5.10 (1H, dd, *J* = 10.5, 4.6, C*H*F), 5.00 (1H, dd, *J* = 10.5, 4.6, C*H*F), 4.55-4.47 (1H, m, C*H*OH), 4.46-4.39 (1H, m, C*H*OH), 2.56 (1H, br s, O*H*), 2.23 (1H, br s, O*H*), 1.87 (3H, d, *J* = 11.0, C*H_3*CCOH), 1.77-1.74 (3H, m, C*H_3*CH), 1.72-1.68 (3H, m, C*H_3*CCHCCOH); **^{13}C NMR** (125 MHz, $CDCl_3$): 197.9 (d, *J* = 35.8, OC*S*), 196.7 (d, *J* = 30.4, OC*S*), 135.0 (Ph), 134.9 (Ph), 134.6 (*C*HCCOH), 132.7 (*C*CHCCOH), 132.6 (*C*HCCOH), 132.5 (*C*CHCCOH), 131.0 (Ph), 130.1 (Ph), 129.6 (Ph), 129.5 (Ph), 126.1 (d, *J* = 24.4, CH_3*C*H), 125.9 (*C*COH), 125.8 (*C*COH), 125.6 (OCS*C*), 125.5 (OCS*C*), 123.2 (d, *J* = 14.9, CH_3*C*H), 96.7 (d, *J* = 104.6, *C*HF), 95.1 (d, *J* = 104.3, *C*HF), 77.7 (d, *J* = 21.9, *C*OH), 77.1 (d, *J* = 21.9, *C*OH), 23.6 (d, *J* = 3.9, *C*H_3CH), 16.6 (d, *J* = 5.4, *C*H_3CH), 15.1 (d, *J* = 10.1, *C*H_3CCOH), 14.7 (d, *J* = 10.0, *C*H_3CCOH), 13.8 (d, *J* = 13.8, *C*H_3CCHCCOH); **^{19}F NMR** (282 MHz, $CDCl_3$): -201.0 (dd, *J* = 48.2, 23.8, CH*F*), -194.6 (dd, *J* = 48.1, 14.7, CH*F*); **IR** (thin film) υ_{max} = 3429,

2917, 2857, 1698, 1440, 1108 cm^{-1}; **HRMS (ESI)**: calculated for $C_{16}FH_{19}NaO_2S$ [M+Na]$^+$: 317.0982, found: 317.0980.

6.1.3.46 Synthesis of S-(2-acetamidoethyl) (4E,6E)-2-fluoro-3-hydroxy-4,6-dimethylocta-4,6-dienethioatedienethioate (274)
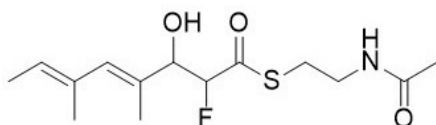
To a solution of S-phenyl (4E,6E)-2-fluoro-3-hydroxy-4,6-dimethylocta-4,6-dienethioate (**273**) (0.15 g, 0.51 mmol) and *N*-acetyl cysteamine (**191**) (0.24 g, 2.04 mmol) in anhydrous DCM (6 mL) under argon at RT was added anhydrous Et$_3$N (0.07 mL, 0.51 mmol) followed by stirring for 6 h. The reaction was then concentrated and purified by silica column chromatography with a gradient of 0-70% EtOAc in petroleum ether to afford the title compound (**274**) as a colourless oil (0.05 g, 31%): **TLC** (Rf = 0.38, EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 6.00 (1H, s, C*H*CCOH), 5.93 (1H, s, C*H*CCOH), 5.84 (1H, br s, N*H*), 5.80 (1H, br s, N*H*), 5.49-5.41 (1H, m, CH$_3$C*H*), 5.40-5.35 (1H, m, CH$_3$C*H*), 5.06-4.85 (1H, m, C*H*F), 4.62-4.32 (1H, C*H*OH), 3.65-3.36 (2H, m, C*H$_2$*NH), 3.27-2.97 (2H, m, SC*H$_2$*), 1.98 (3H, s, NHCOC*H$_3$*), 1.97 (3H, s, NHCOC*H$_3$*), 1.85 (3H, s, C*H*CCHCCOH), 1.74 (3H, s, C*H$_3$*CCOH), 1.69 (1H, d, J = 6.8, C*H$_3$*CHCCH$_3$), 1.67 (3H, s, C*H*CCHCCOH), 1.51 (3H, d, J = 7.3, C*H$_3$*CHCCH$_3$); **^{13}C NMR** (125 MHz, CDCl$_3$): 198.9 (*C*SO), 198.7 (*C*SO), 170.9 (NH*C*O), 170.7 (NH*C*O), 134.0 (CH$_3$*C*HCCH$_3$), 132.7 (CH$_3$CH*C*CH$_3$), 132.5 (CH$_3$CH*C*CH$_3$), 132.2 (CH$_3$*C*HCCH$_3$), 130.7 (*C*COH), 130.2 (*C*COH), 126.1 (*C*HCCOH), 125.9 (*C*HCCOH), 96.2 (d, J = 194.3, *C*FH), 96.1 (d, J = 194.7, *C*FH), 77.0 (*C*OH), 76.8 (*C*OH), 39.1 (*C*H$_2$N), 39.0 (*C*H$_2$N), 28.3 (S*C*H$_2$), 28.2 (S*C*H$_2$), 23.6 (NHCO*C*H$_3$), 23.3 (NHCO*C*H$_3$), 16.6 (*C*H$_3$CCOH), 16.5 (*C*H$_3$CCOH), 14.9 (*C*H$_3$CCHCCOH), 14.8 (*C*H$_3$CCHCCOH), 13.8 (*C*H$_3$CH); **^{19}F NMR** (282 MHz, CDCl$_3$): -203.3 (CH*F*), -202.7 (CH*F*); **IR** (thin film) v_{max} = 3304, 2925, 2856,

1651, 1547, 1436, 1115, 1045 cm^{-1}; **HRMS (ESI)**: calculated for C$_{14}$H$_{22}$FNNaO$_3$S [M+Na]$^+$: 326.1197, found: 201.0930.

6.1.3.47 Synthesis of S-(2-acetamidoethyl) (4E,6E)-2-fluoro-4,6-dimethyl-3-oxoocta-4,6-dienethioate (262)

To a solution of S-(2-acetamidoethyl) (4E,6E)-2-fluoro-3-hydroxy-4,6-dimethylocta-4,6-dienethioatedienethioate (**274**) (0.06 g, 0.19 mmol) in EtOAc (3 mL) at RT under argon was added IBX (0.16 g, 0.57 mmol) followed by stirring at 78 °C for 3 h. The reaction was then cooled, filtered through Celite and the filtrate concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-70% EtOAc in petroleum ether afforded the title compound (**262**) as a yellow oil (0.02 g, 33%): **TLC** (Rf = 0.25, 95:5 DCM : MeOH); **¹H NMR** (500 MHz, CDCl$_3$): δ 7.01 (1H, s, C*H*CCOCHF), 5.95 (1H, q, *J* = 7.0, CH$_3$C*H*), 5.90 (1H, d, *J* = 49.1, C*H*F), 5.88 (1H, br s, N*H*), 3.49-3.42 (2H, m, C*H$_2$*NH), 3.19-3.06 (2H, m, SC*H$_2$*), 2.02 (3H, s, C*H$_3$*CCOCHF), 1.96 (3H, s, NHCOC*H$_3$*), 1.93 (3H, s, CH$_3$CHCC*H$_3$*), 1.82 (3H, d, *J* = 7.0, C*H$_3$*CH); **¹³C NMR** (125 MHz, CDCl$_3$): 194.9 (d, *J* = 29.1, OC*S*), 190.5 (d, *J* = 18.4, OC*C*HF), 170.5 (*C*ONH), 149.7 (*C*HCCOCHF), 136.2 (CH$_3$*C*H), 133.7 (CH$_3$CHC*C*H$_3$), 131.2 (*C*COCHF), 93.9 (d, *J* = 199.6, *C*HF), 39.1 (*C*H$_2$NH), 28.4 (S*C*H$_2$), 23.3 (NHCO*C*H$_3$), 16.1 (CH$_3$CHC*C*H$_3$), 14.6 (*C*H$_3$CH), 13.3 (*C*H$_3$CCOCHF); **¹⁹F NMR** (282 MHz, CDCl$_3$): -186.7 (d, *J* = 6.9, CH*F*); **IR** (thin film) v_{max} = 3305, 2926, 2855, 1658, 1548, 1438, 1035 cm^{-1}; **HRMS (ESI)**: calculated for C$_{14}$H$_{20}$FNNaO$_3$S [M+Na]$^+$: 324.1040, found: 324.1042.

Synthesis of ethyl (4E,6E)-2-fluoro-3-hydroxy-4,6-dimethylocta-4,6-dienoate (276)
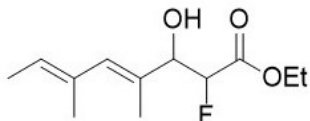
To stirring anhydrous THF (10 mL) at 70 °C under argon was added in quick succession activated zinc dust (0.40 g, 6.09 mmol), (2E,4E)-2,4-dimethylhexa-2,4-dienal (**203**) (0.42 g, 3.39 mmol) and then ethyl bromofluoroacetate (**275**) (0.72 mL, 6.09 mmol) followed by stirring at 70 °C for 5 mins. After the exotherm and colour change, the reaction stirred at RT for 15 mins followed by concentration. The residue was diluted in 1M HCl (15 mL) and extracted with EtOAc (4 x 10 mL). The combined organics were washed with brine (15 mL) then dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-20% EtOAc in petroleum ether afforded the title compound (**276**) as a colourless oil (0.36 g, 46%) with a relative stereochemistry ratio of 1:1 *syn* and *anti* as shown by NMR: **TLC** (Rf = 0.45, 4:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 5.96 (1H, s, C*H*CCHOH), 5.43 (1H, q, *J* = 6.25, C*H*CH$_3$), 4.96 (1H, d, *J* = 48.7, FC*H*), 5.0 (1H, dd, *J* = 48.7, 1.6, FC*H*), 4.45-4.30 (1H, m, C*H*OH), 4.26 (2H, q, *J* = 7.1, CO$_2$C*H$_2$*), 2.41 (1H, d, *J* = 4.8, O*H*), 2.23 (1H, d, *J* = 6.3, O*H*), 1.84 (3H, s, CH$_3$CHCC*H$_3$*), 1.83 (3H, s, CH$_3$CHCC*H$_3$*), 1.74 (3H, s, C*H$_3$*CCHOH), 1.69 (3H, s, C*H$_3$*CH), 1.31 (3H, t, *J* = 7.1, CH$_2$C*H$_3$*), 1.30 (3H, t, *J* = 7.1, CH$_2$C*H$_3$*); **^{13}C NMR** (125 MHz, CDCl$_3$): 168.2 (*C*O$_2$), 168.0 (*C*O$_2$), 133.5 (*C*HCCHOH), 132.6 (d, *J* = 5.2, *C*CHOH), 132.5 (*C*HCCHOH), 130.7 (d, *J* = 2.8, CH$_3$CH*C*CH$_3$), 130.3 (d, *J* = 2.2, CH$_3$CH*C*CH$_3$), 125.9 (CH$_3$*C*H), 125.8 (CH$_3$*C*H), 91.0 (d, *J* = 96.9, *C*HF), 89.5 (d, *J* = 96.5, *C*HF), 77.4 (d, *J* = 20.0, *C*HOH), 77.1 (d, *J* = 22.1, *C*HOH), 61.9 (O*C*H$_2$), 61.8 (O*C*H$_2$), 16.6 (*C*H$_3$CCHOH), 16.5 (*C*H$_3$CCHOH), 14.4 (*C*H$_3$CCHCH$_3$), 14.3 (*C*H$_2$CH$_3$), 13.9 (*C*H$_3$CCHCH$_3$), 13.8 (*C*H$_3$CH); **IR** (thin film) v_{max} = 3475, 2980, 2917, 1741, 1465, 1280, 1206, 1023 cm^{-1}; **HRMS (ESI)**: calculated for C$_{12}$H$_{19}$FNaO$_3$ [M+Na]$^+$: 253.1210, found: 253.1212.

6.1.3.48 Synthesis of (4E,6E)-2-fluoro-3-hydroxy-4,6-dimethylocta-4,6-dienoic acid (277)
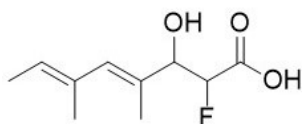
A solution of ethyl (4E,6E)-2-fluoro-3-hydroxy-4,6-dimethylocta-4,6-dienoate (**276**) (0.36 g, 1.56 mmol) was stirred at reflux in THF (4 mL) and 2M NaOH (3 mL) for 2 h. The reaction was then concentrated, diluted in H_2O (10 mL) and washed with Et_2O (2 x 5 mL). The remaining aqueous solution was acidified to pH 2 using 1M HCl, extracted in EtOAc (5 x 5 mL) and the combined organics dried over $MgSO_{4(s)}$, filtered and concentrated to a brown/orange oil. This solidified on standing and proved to be the title compound (**277**) (0.15 g, 46 %), NMR appeared to show a single relative stereochemistry which was not further characterised: **^{1}H-NMR** (500 MHz, $CDCl_3$): δ 6.01 (1H, s, 5.45 (1H, q, J = 6.7, CH_3CH), 5.1 (1H, dd, J = 47.9, 3.1, CHF), 4.50 (1H, dd, J = 24.6, 2.50, $CHOH$), 1.86 (3H, s, CH_3CCHCH_3), 1.75 (3H, s, CH_3CCHOH), 1.69 (3H, d, J = 6.9, CH_3CH); **^{13}C NMR** (125 MHz, $CDCl_3$): 170.7 (d, J = 24.8, CO_2), 132.3 ($CHCCHOH$), 132.3 (CH_3CCHCH_3), 130.4 ($CCHOH$), 125.8 (CH_3CH), 89.9 (d, J = 191.7, CHF), 76.6 ($CHOH$), 16.4 (CH_3CCHOH), 14.5 (CH_3CCHCH_3), 13.7 (CH_3CHCCH_3); **IR** (thin film) v_{max} = 3333, 2992, 2951, 1720, 1437, 1211, 1103, 1026 cm^{-1}; **HRMS (ESI)**: calculated for $C_{10}H_{14}FO_3$ [M-H]$^-$: 201.0932, found: 201.0930.

6.1.3.49 Alternative synthesis of S-(2-acetamidoethyl) (4E,6E)-2-fluoro-3-hydroxy-4,6-dimethylocta-4,6-dienethioatedienethioate (274)
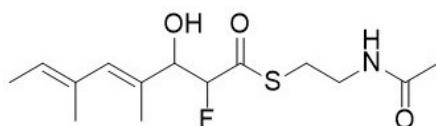
To a solution of (4E,6E)-2-fluoro-3-hydroxy-4,6-dimethylocta-4,6-dienoic acid (**277**) (0.12 g, 0.72 mmol), DMAP (0.01 g, 0.07 mmol) and *N*-acetyl cysteamine (**191**) (0.09 g,

0.75 mmol) in anhydrous THF (5 mL) under argon at 0 °C was added DCC (0.16 g, 0.79 mmol) followed by stirring at RT for 16 h. The reaction was then concentrated, triturated in minimal cold EtOAc and filtered through Celite. The filtrate was concentrated and purified by silica column chromatography with a gradient of 50-100% EtOAc in petroleum ether to afford the title compound (**274**) as a pale yellow oil (0.06 g, 27%), NMR appeared to show a single relative stereochemistry which was not further characterised: **TLC** (Rf = 0.38, EtOAc); ^{1}**H NMR** (500 MHz, CDCl$_3$): δ 6.00 (1H, s, C*H*CCHOH), 5.88 (1H, br s, N*H*), 5.44 (1H, q, J = 6.9, CH$_3$C*H*), 4.96 (1H, dd, J = 48.0, 2.7, C*H*F), 4.51-4.42 (1H, m, C*H*OH), 3.62-3.35 (2H, m, SC*H*$_2$), 3.23-2.99 (2H, m, C*H*$_2$NH), 1.97 (3H, s, NHCOC*H*$_3$), 16.6 (3H, s, C*H*$_3$CCHOH), 14.8 (3H, s, C*H*$_3$CCHCH$_3$), 13.9 (3H, d, J = 6.9, C*H*$_3$CHCCH$_3$); ^{13}**C NMR** (125 MHz, CDCl$_3$): 198.8 (d, J = 29.3, S*C*O), 170.7 (NH*C*O), 132.6 (CH$_3$CH*C*CH$_3$), 132.2 (*C*HCCHOH), 130.8 (*C*CHOH), 125.9 (CH$_3$*C*HCCH$_3$), 96.2 (d, J = 195.6, *C*HF), 77.0 (*C*HOH), 30.9 (*C*H$_2$NH), 28.3 (S*C*H$_2$), 23.4 (NHCO*C*H$_3$), 16.6 (*C*H$_3$CCHOH), 14.8 (*C*H$_3$CCHCH$_3$), 13.9 (*C*H$_3$CHC*C*H$_3$); **IR** (thin film) v_{max} = 3304, 2925, 2856, 1651, 1547, 1436, 1115, 1045 cm^{-1}; **HRMS (ESI)**: calculated for C$_{14}$H$_{22}$FNNaO$_3$S [M+Na]$^+$: 326.1197, found: 201.0930.

6.1.3.50 Synthesis of S-phenyl butanethioate (278)
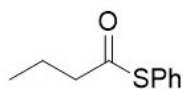
To a solution of thiophenol (1.02 mL, 10.0 mmol) in anhydrous DCM (20 mL) under argon at 0 °C was added anhydrous pyridine (0.97 mL, 12.0 mmol) followed by dropwise addition of butyryl chloride (1.13 mL, 11.0 mmol), then stirring at RT for 16 h. H$_2$O (20 mL) was added to the reaction and the solution was extracted in DCM (3 x 10 mL). The combined organics were washed with sat. NaHCO$_{3(aq.)}$ (20 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-10% EtOAc in petroleum ether afforded the title

compound (**278**) as a colourless oil (1.97 g, 100%): **TLC** (Rf = 0.82, 9:1 petroleum ether : EtOAc); **^{1}H NMR** (300 MHz, CDCl$_3$): δ 7.41 (5H, s, Ar-H), 2.64 (2H, t, *J* = 7.4, SCOC*H$_2$*), 1.75 (2H, sext, *J* = 7.5, CH$_3$C*H$_2$*), 1.00 (3H, t, *J* = 7.4, C*H$_3$*CH$_2$). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[181]

### 6.1.3.51 Synthesis of S-phenyl (4E,6E)-2-ethyl-3-hydroxy-4,6-dimethylocta-4,6-dienethioate (279)
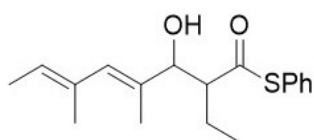
To a solution of S-phenyl butanethioate (**278**) (0.46 g, 2.53 mmol) and (2E,4E)-2,4-dimethylhexa-2,4-dienal (**203**) (0.40 g, 2.78 mmol) in DCM (13 mL) under argon at RT was added in one portion MgBr$_2$.OEt$_2$ (0.91 g, 3.54 mmol). After stirring for 5 mins, DIPEA (0.88 mL, 5.03 mmol) was added dropwise followed by stirring at RT for 1 h. The reaction was then diluted with EtOAc (20 mL) and H$_2$O (15 mL), the aqueous phase was adjusted the pH 5 using 1M HCl and the biphasic solution was stirred vigorously for 5 mins. The organic layer was separated and aqueous extracted using EtOAc (4 x 10 mL), the combined organic extracts were washed with brine (30 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-20% EtOAc in petroleum ether afforded the title compound (**279**) as a yellow oil (0.36 g, 46%) with a relative stereochemistry ratio of 1:1 *syn* and *anti* as shown by NMR: **TLC** (Rf = 0.21, 9:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.43-7.33 (5H, m, Ph), 5.94 (1H, s, C*H*CCOH), 5.89 (1H, s, C*H*CCOH), 5.44 (1H, q, *J* = 6.6, CH$_3$C*H*), 4.27 (1H, d, *J* = 7.6, C*H*OH), 4.22 (1H, d, *J* = 6.9, C*H*OH), 2.92-2.83 (1H, m, C*H*CSO), 2.21 (1H, br s, O*H*), 2.05 (1H, br s, O*H*), 1.84 (3H, s, CH$_3$CHCC*H$_3$*), 1.81 (3H, s, CH$_3$CHCC*H$_3$*), 1.76-1.73 (3H, m, C*H$_3$*CCOH), 1.70 (3H, d, *J* = 6.9, C*H$_3$*CH), 1.59-1.50 (2H, m, CH$_3$C*H$_2$*), 1.07-0.99 (3H, m, C*H$_3$*CH$_2$); **^{13}C NMR** (125 MHz, CDCl$_3$): 201.3 (O*C*S), 200.3 (O*C*S), 134.4 (Ph), 134.3 (Ph), 133.4 (*C*COH), 132.9 (*C*COH), 132.8 (*C*CHCCHCOH), 132.7 (*C*CHCCHCOH), 132.2

(*C*HCCOH), 131.9 (*C*HCCOH), 129.5 (Ph), 129.2 (Ph), 127.7 (Ph), 127.6 (Ph), 125.3 (CH$_3$*C*H), 125.1 (CH$_3$*C*H), 79.4 (*C*OH), 78.2 (*C*OH), 59.3 (*C*HCOS), 58.6 (*C*HCOS), 23.5 (*C*H$_2$CH$_3$), 21.3 (*C*H$_2$CH$_3$), 16.6 (*C*H$_3$CCOH), 13.8 (*C*H$_3$CH), 13.7 (*C*H$_3$CCHCCOH), 13.3 (*C*H$_3$CCHCCOH), 12.1 (CH$_2$*C*H$_3$), 11.7 (CH$_2$*C*H$_3$); **IR** (thin film) υ_{max} = 3428, 2964, 2930, 1697, 1440, 974, 745 cm^{-1}; **HRMS (ESI)**: calculated for C$_{18}$H$_{24}$NaO$_2$S [M+Na]$^+$: 327.1389, found: 327.1390.
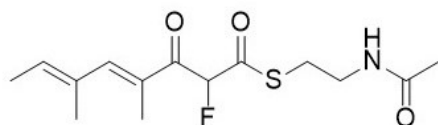
6.1.3.52 Synthesis of S-(2-acetamidoethyl) (4E,6E)-2-ethyl-3-hydroxy-4,6-dimethylocta-4,6-dienethioate (280)
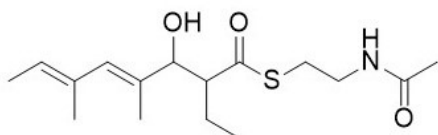
To a solution of S-phenyl (4E,6E)-2-ethyl-3-hydroxy-4,6-dimethylocta-4,6-dienethioate (**279**) (0.42 g, 1.16 mmol) and *N*-acetyl cysteamine (**191**) (0.41 g, 3.47 mmol) in anhydrous DCM (5 mL) at RT under argon was added anhydrous Et$_3$N (0.04 mL, 0.23 mmol) followed by stirring for 7 h. The reaction was then concentrated and purified by silica column chromatography with a gradient of 25-100% EtOAc in petroleum ether to afford the title compound (**280**) as a colourless oil (0.04 g, 12%) with a relative stereochemistry ratio of 1:1 *syn* and *anti* as shown by NMR: **TLC** (Rf = 0.35, EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 5.90 (1H, s, C*H*CCOH), 5.87 (1H, s, C*H*CCOH), 5.77 (1H, br s, N*H*), 5.40 (1H, q, *J* = 6.7, CH$_3$C*H*), 5.34 (1H, q, *J* = 6.6, CH$_3$C*H*), 4.30 (1H, d, *J* = 6.6, C*H*OH), 4.20 (1H, d, *J* = 7.2, C*H*OH), 3.50-3.31 (2H, m, C*H*$_2$NH), 3.01-2.94 (2H, m, SC*H*$_2$), 2.83-2.76 (2H, m, C*H*CSO), 1.96 (3H, s, NHCOC*H*$_3$), 1.94 (3H, s, NHCOC*H*$_3$), 1.81-1.78 (2H, CH$_3$C*H*$_2$), 1.72 (2H, s, C*H*$_3$CCHCCCOH), 1.80 (3H, s, C*H*$_3$CCOH), 1.61 (3H, s, C*H*$_3$CCOH), 1.67 (3H, d, *J* = 6.9, C*H*$_3$CH), 1.48 (3H, d, *J* = 6.8, C*H*$_3$CH), 0.98-0.91 (3H, m, C*H*$_3$CH$_2$); **^{13}C NMR** (125 MHz, CDCl$_3$): 202.9 (O*C*S), 202.8 (O*C*S), 170.4 (*C*ONH), 170.3 (*C*ONH), 135.4 (*C*CHCCOH), 133.3 (*C*CHCCOH), 133.0 (*C*COH), 132.9 (*C*COH), 131.8 (*C*HCCOH), 127.3 (*C*HCCOH), 125.5 (CH$_3$*C*H), 122.7 (CH$_3$*C*H), 78.6 (*C*OH), 77.6 (*C*OH), 59.9 (*C*HCSO), 59.7 (*C*HCSO), 39.8 (*C*H$_2$NH), 28.7 (S*C*H$_2$), 28.6 (S*C*H$_2$), 23.3 (NHCO*C*H$_3$), 23.2 (NHCO*C*H$_3$), 21.4 (CH$_3$*C*H$_2$), 21.1

(CH$_3$*CH$_2$*), 16.7 (CH$_3$*CH*), 15.1 (CH$_3$*CH*), 13.9 (CH$_3$*C*CHCCCOH), 13.8 (CH$_3$*C*COH), 13.7 (CH$_3$*C*COH); **IR** (thin film) υ_{max} = 3291, 2965, 2930, 1655, 1549, 1439, 986 cm^{-1}; **HRMS (ESI)**: calculated for C$_{16}$H$_{27}$NNaO$_3$S [M+Na]$^+$: 336.1604, found: 336.1601.

6.1.3.53 Synthesis of S-(2-acetamidoethyl) (4E,6E)-2-ethyl-4,6-dimethyl-3-oxoocta-4,6-dienethioate (263)
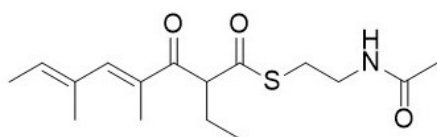
To a solution of S-(2-acetamidoethyl) (4E,6E)-2-ethyl-3-hydroxy-4,6-dimethylocta-4,6-dienethioate (**280**) (0.04 g, 0.13 mmol) in EtOAc (5 mL) at RT under argon was added IBX (0.11 g, 0.40 mmol) followed by stirring at 78 °C for 4 h. The reaction was then cooled, filtered through Celite and the filtrate concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-90% EtOAc in petroleum ether afforded the title compound (**263**) as a yellow oil (0.04 g, 95%), NMR analysis showed a 1:1 ratio *cis* and *trans* isomers for the CH$_3$-CH alkene of the compound: **TLC** (Rf = 0.27, 3:1 EtOAc : petroleum ether); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.11 (1H, s, C*H*CCOCHCSO), 7.01 (1H, s, C*H*CCOCHCSO), 5.86 (1H, q, *J* = 6.9, CH$_3$*CH*), 5.28 (1H, br s, N*H*), 5.58 (1H, q, *J* = 6.9, CH$_3$*CH*), 4.36-4.30 (1H, m, C*H*CSO), 3.50-3.35 (2H, m, C*H$_2$*NH), 3.12-2.98 (2H, m, SC*H$_2$*), 2.00-1.96 (2H, m, CH$_3$*CH$_2$*), 1.98 (1H, s, CH$_3$CHCC*H$_3$*), 1.95 (3H, s, NHCOC*H$_3$*), 1.94 (3H, s, NHCOC*H$_3$*), 1.90 (3H, s, C*H$_3$*CCO), 1.87 (3H, s, C*H$_3$*CCO), 1.81 (1H, s, CH$_3$CHCC*H$_3$*), 1.80 (3H, d, *J* = 6.9, C*H$_3$*CH), 1.57 (3H, d, *J* = 7.0, C*H$_3$*CH), 0.98-0.90 (2H, m, CH$_3$C*H$_2$*); **^{13}C NMR** (125 MHz, CDCl$_3$): 196.7 (*C*O), 196.6 (*C*O), 196.5 (*C*SO), 196.3 (*C*SO), 170.5 (*C*ONH), 145.7 (*C*HCCO), 140.9 (*C*HCCO), 137.1 (*C*CHCCO), 134.0 (*C*CHCCO), 133.8 (CH$_3$*C*H), 133.5 (*C*CO), 132.2 (*C*CO), 127.2 (CH$_3$*C*H), 62.9 (*C*CSO), 62.7 (*C*CSO), 40.0 (*C*H$_2$N), 28.8 (*C*H$_2$S), 24.1 (CH*C*H$_2$), 23.9 (CH*C*H$_2$), 23.3 (NHCO*C*H$_3$), 22.9 (NHCO*C*H$_3$), 16.2 (*C*H$_3$CH), 15.3 (*C*H$_3$CH), 14.5 (CH$_3$CHC*C*H$_3$), 13.7 (*C*H$_3$CCO), 13.5 (CH$_3$CHC*C*H$_3$);

IR (thin film) v_{max} = 3288, 2967, 2930, 2875, 1652, 1543, 1239, 983 cm^{-1}; **HRMS (ESI)**: calculated for $C_{16}H_{25}NNaO_3S$ [M+Na]$^+$: 334.1447, found: 334.1449.

6.1.3.54 Synthesis of methyl (4E,6E)-3-hydroxy-4,6-dimethyl-2-(trifluoromethyl)octa-4,6-dienoate (282)
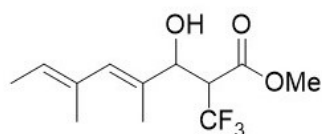
To stirring solution of methyl-3,3,3-trifluoropropanoate (**281**) (1.34 g, 8.00 mmol) in anhydrous DCM (32 mL) at -78 °C under argon was added dropwise 1M dibutyl boron triflate in DCM (8.00 mL, 8.00 mmol) followed by stirring for 1 h. The reaction was then warmed to 0 °C for 1 h. The reaction was re-cooled to -78 °C and (2E,4E)-2,4-dimethylhexa-2,4-dienal (**203**) (0.75 g, 6 .00 mmol) in anhydrous DCM (8 mL) was added dropwise). This was followed by stirring for 1 h and then the reaction was warmed to 0 °C for 1 h. The reaction was then quenched with 0.5M pH 7 phosphate buffer (15 mL), MeOH (15 mL) and then 30% hydrogen peroxide solution (15 mL) followed by stirring at RT for 3 h. The reaction was then extracted with DCM (4 x 15 mL), the combined organic layers were washed with brine (30 mL) dried over MgSO$_{4(s)}$, filtered and concentrated to the crude product. Purification by silica column chromatography with a gradient of 0-20% EtOAc in petroleum ether afforded the title compound (**282**) (0.88 g, 55%) as a brown oil with a relative stereochemistry ratio of 2:1 as shown by NMR but the major isomer was not characterised as *syn* and *anti*: **TLC** (Rf = 0.46, 4:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 5.98 (1H, s, *H*CCCHOH), 5.89 (1H, s, *H*CCCHOH), 5.45-5.35 (1H, m, H$_3$CC*H*), 4.65-4.57 (1H, m, C*H*OH), 3.81 (3H, s, OC*H$_3$*), 3.69 (3H, s, OC*H$_3$*), 3.51-3.44 (1H, m, C*H*CF$_3$), 2.26 (1H, d, *J* = 5.1, O*H*), 1.81 (3H, s, *H$_3$*CCCHCCHOH), 1.78 (3H, s, *H$_3$*CCCHCCHOH), 1.73 (3H, s, *H$_3$*CCCHOH), 1.71 (3H, s, *H$_3$*CCCHOH), 1.69 (3H, d, *J* = 7.3, *H$_3$*CCH); **^{13}C NMR** (125 MHz, CDCl$_3$): 167.2 (*C*O$_2$), 167.1 (*C*O$_2$), 134.2 (H*C*CCHOH), 133.5 (H*C*CCHOH), 132.6 (*C*CHOH), 131.5

(*C*CHOH), 126.2 (CH$_3$*C*H), 126.1 (CH$_3$*C*H), 125.3 (*C*CHCHOH), 125.0 (*C*CHCHOH), 123.2 (q, *J* = 280.3, *C*F$_3$), 76.0 (d, *J* = 2.7, *C*OH), 75.2 (d, *J* = 1.9, *C*OH), 55.2 (q, *J* = 26.1, H*C*CF$_3$), 53.8 (q, *J* = 25.9, H*C*CF$_3$), 53.0 (3H, s, O*C*H$_3$), 52.8 (3H, s, O*C*H$_3$), 16.4 (H$_3$*C*CCHOH), 13.8 (H$_3$*C*CH), 12.9 (3H, s, H$_3$*C*CCHCCHOH), 12.1 (3H, s, H$_3$*C*CCHCCHOH); **IR** (thin film) v_{max} = 3454, 2957, 2873, 1752, 1439, 1321, 1265 cm^{-1}; **HRMS (ESI)**: calculated for C$_{12}$H$_{17}$F$_3$NaO$_3$ [M+Na]$^-$: 289.1022, found: 289.1024.

6.1.3.55 Synthesis of methyl 4-bromo-3-oxopentanoate (297)
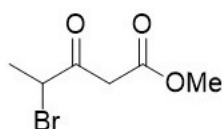
To a solution of methyl 3-oxopentanoate (**296**) (0.48 mL, 3.84 mmol) in chloroform (1 mL) at RT under argon was added dropwise a solution of 1M Br$_2$ in CHCl$_3$ (3.84 mL, 3.84 mmol). After stirring for 16 h, the reaction was diluted with H$_2$O (10 mL) and extracted with DCM (3 x 5 mL) the combined organic layers were dried over MgSO$_{4(s)}$, filtered and concentrated to the crude product. Purification by column chromatography with a gradient of 0-10% EtOAc in petroleum ether afforded the title compound (**297**) as a pale orange oil, the purity was estimated as 80% by H NMR (0.81 g, 81%): **TLC** (Rf = 0.44, 9:1 petroleum ether : EtOAc); **^{1}H NMR** (300 MHz, CDCl$_3$): δ 4.61 (1H, q, *J* = 6.8, C*H*Br), 3.90-3.64 (2H, m, C*H$_2$*), 3.75 (3H, s, OC*H$_3$*), 1.78 (3H, d, *J* = 6.8, C*H$_3$*CHBr). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[182]

6.1.3.56 Synthesis of methyl 4-bromo-2-fluoro-3-oxopentanoate (300)
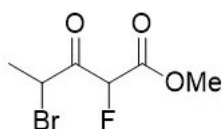
To a solution of methyl 4-bromo-3-oxopentanoate (**297**) (0.32 g, 1.53 mmol) in anhydrous MeCN (8 mL) under argon at RT was added cyclopentadienyltitanium trichloride (0.02 g, 0.08 mmol) then by Selectfluor (0.71 g, 1.99 mmol) followed by stirring for 2 h. The reaction was then concentrated, diluted with H_2O (10 mL) and extracted with DCM (4 x 5 mL). The combined organics were dried over $MgSO_{4(s)}$, filtered and concentrated to the crude product. Purification by silica column chromatography with a gradient of 0-15% EtOAc in petroleum ether afforded the title compound (**300**) as a colourless oil (0.32 g, 91%). NMR appeared to show a single relative stereochemistry which was not further characterised: **TLC** (Rf = 0.23, 9:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, CDCl₃): δ 5.63 (1H, dd, *J* = 48.1, 14,4, C*H*F), 4.95-4.85 (1H, m, C*H*Br), 3.88 (3H, s, OC*H₃*), 1.80 (3H, d, *J* = 6.7, *H₃*CCHBr), 1.79 (3H, d, *J* = 6.7, H₃CCHBr); **^{13}C NMR** (125 MHz, CDCl₃): 193.7 (d, *J* = 22.8, *C*O), 193.4 (d, *J* = 22.6, *C*O), 164.4 (d, *J* = 23.6, *C*O₂), 164.3 (d, *J* = 23.8, *C*O₂), 88.9 (d, *J* = 99.1, *C*HF), 88.3 (d, *J* = 100.1, *C*HF), 53.7 (O*C*H₃), 53.5 (O*C*H₃), 41.6 (*C*HBr), 41.0 (*C*HBr), 19.1 (*C*H₃CHBr), 18.9 (*C*H₃CHBr); **^{19}F NMR** (282 MHz, CDCl₃): -197.5 (C*F*), 195.9 (C*F*); **IR** (thin film) v_{max} = 2958, 1764, 1734, 1439, 1262, 1134, 1104 cm⁻¹; **HRMS (ESI)**: calculated for $C_6H_8BrFNaO_3$ [M+Na]⁺: 248.9533, found: 248.9534.

6.1.3.57 Synthesis of methyl 4-(acetylthio)-3-oxopentanoate (298)
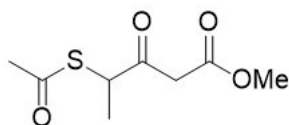
To a stirring solution of methyl 4-bromo-3-oxopentanoate (**297**) (0.25 g, 0.90 mmol) in anhydrous DMF (3 mL) at RT under argon was added potassium thioacetate (0.10 g, 0.90 mmol), the reaction was stirred at RT for 16 h. Water (10 mL) was added and the reaction was extracted with EtOAc (3 x 10 mL). The combined organics were washed with brine (4 x 10 mL), dried over MgSO₄, filtered and concentrated to the crude product. Purification by silica column chromatography with a gradient of 0-35% EtOAc in

petroleum ether afforded the title compound (**298**) as a colourless oil (0.24 g). Proton NMR showed a complex mixture which could not be characterised, the material was used without further analysis: **TLC** (Rf = 0.21, 4:1 petroleum ether : EtOAc).

6.1.3.58 Synthesis of methyl 4-(acetylthio)-2-fluoro-3-oxopentanoate (301)
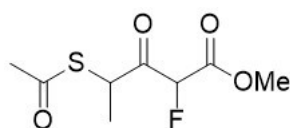
To a stirring solution of methyl 4-bromo-2-fluoro-3-oxopentanoate (**300**) (0.38 g, 1.57 mmol) in anhydrous DMF (3 mL) at RT under argon was added potassium thioacetate (0.18 g, 1.57 mmol), the reaction was stirred at RT for 16 h. Water (10 mL) was added and the reaction was extracted with EtOAc (3 x 10 mL). The combined organics were washed with brine (4 x 10 mL), dried over MgSO$_4$, filtered and concentrated to the crude product. Purification by silica column chromatography with a gradient of 0-30% EtOAc in petroleum ether afforded the title compound (**301**) as a colourless oil (0.23 g). Proton NMR showed a complex mixture which could not be characterised, the material was assumed pure and used without further analysis: **TLC** (Rf = 0.24, 4:1 petroleum ether : EtOAc).

6.1.3.59 Synthesis of methyl 4-bromo-2-methyl-3-oxopentanoate (123)
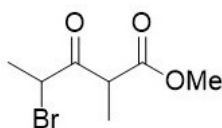
To a solution of methyl 2-methyl-3-oxopentanoate (**122**) (3.19 g, 22.1 mmol) in CHCl$_3$ (6 mL) at RT under argon was added dropwise Br$_2$ (1M in CHCl$_3$) (22.2 mL, 22.2 mmol) followed by stirring for 16 h. The reaction was washed with H$_2$O (50 mL), dried over MgSO$_{4(s)}$, filtered, concentrated and purified by silica chromatography with a gradient of 0-10% EtOAc in petrol to afford the title product (**123**) as a pale-yellow oil with a relative

stereochemistry ratio of 1:1 *syn* and *anti* as shown by NMR: **^{1}H NMR** (300 MHz, CDCl$_3$): δ 4.79-4.57 (1H, m, C*H*Br), 4.15-4.05 (1H, m, C*H*CH$_3$), 3.72 (3H, s, OC*H*$_3$), 1.76 (3H, d, *J* = 6.5, C*H*$_3$CHBr), 1.42 (3H, d, *J* = 7.1, C*H*$_3$CH). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[183]

6.1.3.60 Synthesis of (4E,6E)-N-(2-acetamidoethyl)-2,4,6-trimethyl-3-oxoocta-4,6-dienamide (315)
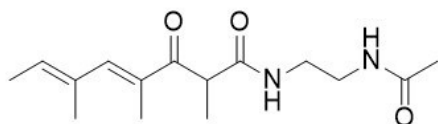
To a stirring solution of (4E,6E)-3-hydroxy-2,4,6-trimethylocta-4,6-dienoic acid (**210**) (0.05 g, 0.25 mmol), *N*-(2-aminoethyl)acetamide (**133**) (0.03 mL, 0.28 mmol), and DIPEA (0.09 mL, 0.53 mmol) in anhydrous DCM (2 mL) at RT under argon was added HATU (0.12 g, 0.30 mmol). After 16 h, the reaction was diluted with 0.1M HCl (10 mL) and extracted in DCM (3 x 5 mL). The combined organics were dried over MgSO$_{4(s)}$, filtered, concentrated and purified by silica column chromatography with a gradient of 0-10% MeOH in EtOAc to afford the semi-pure hydroxy amide intermediate (**314**) (0.05 g). To a stirring solution of the intermediate (**314**) (0.05 g) in EtOAc (2.5 mL) was added IBX (0.10 g, 0.34 mmol) and stirred at 80 °C for 2 h. The reaction was cooled, filtered through Celite and the filtrate concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-10% MeOH in EtOAc afforded the title compound (**315**) as a white solid (0.03 g, 47%): **TLC** (Rf = 0.21, 9:1 EtOAc : MeOH); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.11 (1H, s, C*H*CCO), 6.88 (1H, br s, COCHCON*H*), 6.18 (1H, br s, CON*H*CH$_3$), 5.86 (1H, q, *J* = 6.8, CH$_3$C*H*), 4.19 (1H, q, *J* = 7.2, C*H*CONH), 3.43-3.29 (4H, m, NHC*H*$_2$C*H*$_2$NH), 1.97 (3H, s, C*H*$_3$CCOCHCONH), 1.96 (3H, s, NHCOC*H*$_3$), 1.90 (3H, C*H*$_3$CCHCCO), 1.79 (3H, d, *J* = 7.0, C*H*$_3$CH), 1.41 (3H, d, *J* = 7.2, C*H*$_3$CHCONH); **^{13}C NMR** (125 MHz, CDCl$_3$): 202.0 (*C*O), 172.4 (CO*C*HCONH), 171.0 (NH*C*OCH$_3$), 146.0 (*C*HCCOCHCONH),

240

134.1 (CH₃C*H*CCH), 133.5 (*C*COCHCONH), 133.2 (CH₃CH*C*CH₃), 48.0 (CO*C*HCONH), 40.5 (COCHCONH*C*H₂), 39.7 (*C*H₂NHCOCH₃), 23.3 (NHCO*C*H₃), 17.8 (*C*H₃CHCONH), 16.2 (*C*H₃CCHCCO), 14.5 (*C*H₃CHCCH₃); **IR** (thin film) v_{max} = 3281, 3086, 2980, 2933, 1666, 1635, 1553, 1445, 1244, 1204 cm^{-1}; **HRMS (ESI)**: calculated for $C_{15}H_{24}N_2NaO_3$ [M+Na]$^+$: 303.1679, found: 303.1674.

6.1.3.61 Synthesis of *tert*-butyl (2-decanamidoethyl)carbamate (349)
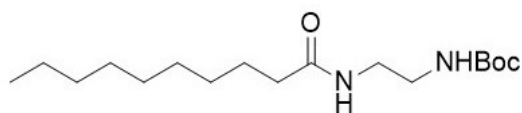
To a solution of *N*-Boc-1,2-diaminoethane (**348**) (0.98 g, 6.13 mmol) and Et₃N (1.71 mL), 12.3 mmol) in anhydrous DCM (30 mL) under argon at 0 °C was added decanoyl chloride (**162**) (1.40 mL, 6.74 mmol) followed by stirring at RT for 1 h. The reaction was then washed with H_2O (3 x 20 mL), dried over MgSO_{4(s)}, filtered and concentrated to give the crude product. This was recrystallised by dissolving in minimal hot EtOAc, diluting with petroleum ether, cooling to 0 °C and filtration. The filtrate was collected and dried to give the title compound (**349**) as a white powder (1.50 g, 83%): **¹H NMR** (500 MHz, CDCl₃): δ 6.11 (1H, br s , CON*H*), 4.89 (1H, br s, N*H*CO₂), 3.38-3.32 (2H, m, CONHC*H*₂), 3.30-3.24 (2H, m, C*H*₂NHCO₂), 2.16 (2H, t, *J* = 7.7, C*H*₂CONH), 1.64-1.59 (2H, m, C*H*₂CH₂CONH), 1.44 (9H, s, 3 x C*H*₃), 1.33-1.21 (12H, m, C*H*₂), 0.88 (3H, t, *J* = 6.8, C*H*₃CH₂); **¹³C NMR** (125 MHz, CDCl₃): 174.0 (*C*ONH), 157.0 (NH*C*O₂), 79.9 (NHCO₂*C*), 40.9 (CONH*C*H₂), 40.4 (*C*H₂NHCO₂), 37.0 (*C*ONH), 32.0 (*C*H₂), 29.6 (*C*H₂), 29.5 (*C*H₂), 29.4 (*C*H₂), 29.4 (*C*H₂), 28.5 (2 x *C*H₃), 25.9 (*C*H₂CH₂CONH), 22.8 (*C*H₂), 14.2 (*C*H₃CH₂); **IR** (thin film) v_{max} = 3352, 3318, 2919, 2850, 1685, 1643, 1535, 1282, 1176, 978 cm^{-1}; **HRMS (ESI)**: calculated for $C_{17}H_{34}N_2NaO_3$ [M+Na]$^+$: 337.2462, found: 337.2452.

6.1.3.62 Synthesis of N-(2-aminoethyl)decanamide hydrochloride (350)
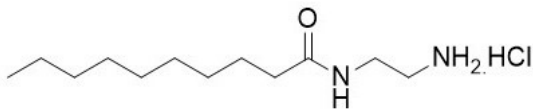
A solution of tert-butyl (2-decanamidoethyl)carbamate (**349**) (1.50 g, 4.78 mmol) in MeOH (15 mL) and conc. HCl (3 mL) was stirred for 16 h. The reaction was then vigorously concentrated, dissolved in minimal hot MeOH and precipitated by addition of ice-cold Et$_2$O (50 mL). The suspension was filtered and the cake collected and dried to give the title compound (**350**) as a white solid (1.10 g, 92%). **^{1}H NMR** (500 MHz, MeOD): δ 3.44 (2H, t, J = 6.0, CONHCH_2), 3.04 (2H, t, J = 6.0, CH_2NH$_2$), 2.24 (2H, t, J = 7.6, CH_2CONH), 1.65-1.57 (2H, m, CH_2CH$_2$CONH), 1.37-1.24 (12H, m, CH_2), 0.90 (3H, t, J = 6.8, CH_3CH$_2$); **^{13}C NMR** (125 MHz, MeOD): 177.7 (*C*ONH), 40.9 (*C*H$_2$NH$_2$), 38.2 (CONH*C*H$_2$), 36.9 (*C*H$_2$CONH), 33.0 (*C*H$_2$), 30.6 (*C*H$_2$), 30.5 (*C*H$_2$), 30.4 (*C*H$_2$), 30.4 (*C*H$_2$), 26.7 (*C*H$_2$CH$_2$CONH), 23.7 (*C*H$_2$), 14.4 (*C*H$_3$); **IR** (neat) v_{max} = 3456, 2976, 2948, 1736, 1456, 1375, 1252, 1197 cm^{-1}; **HRMS (ESI)**: calculated for C$_{12}$H$_{27}$N$_2$O [M+H]$^+$: 215.2118, found: 215.2115.

6.1.3.63 N-(2-(2-hydroxyacetamido)ethyl)decanamide (342)
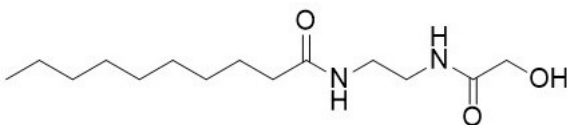
To a solution of tert-butyl (2-decanamidoethyl)carbamate (**350**) (0.20 g, 0.80 mmol), glycolic acid (**351**) (0.09 g, 1.20 mmol) and Et$_3$N (0.34 mL, 2.39 mmol) in anhydrous DMF (8 mL) under argon at RT was added HATU (0.46 g, 1.20 mmol) followed by stirring for 16 h. The reaction was then concentrated, diluted in H$_2$O (20 mL) and extracted with EtOAc (5 x 5 mL). The combined organics were washed with brine (20 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to the crude product, purification by silica column chromatography with a gradient of 0-15% MeOH in DCM afforded the

semi-pure product. This was then dissolved in minimal hot MeOH and then diluted in excess Et_2O followed by cooling to 0°C. The suspension was filtered and the filter cake collected and dried to give the title compound (**342**) as a white solid (0.11 g, 50%): **TLC** (Rf = 0.19, 85:15 DCM : MeOH); **^{1}H NMR** (500 MHz, d_6-DMSO): δ 7.85-7.81 (1H, m, *H*NCOCH$_2$OH), 7.81-7.76 (1H, m, CON*H*), 5.48 (1H, t, *J* = 5.7, O*H*), 3.78 (2H, d, *J* = 5.7, C*H*$_2$OH), 3.17-3.06 (4H, m, NHC*H*$_2$C*H*$_2$NH), 2.03 (2H, t, *J* = 7.5, C*H*$_2$CONH), 1.50-1.42 (2H, m, C*H*$_2$CH$_2$CONH), 1.27-1.15 (12H, m, C*H*$_2$), 0.85 (3H, t, *J* = 6.8, C*H*$_3$); **^{13}C NMR** (125 MHz, d_6-DMSO): 172.5 (*C*ONH), 172.1 (NH*C*OCH$_2$OH), 61.5 (*C*H$_2$OH), 38.4 (NH*C*H$_2$), 38.2 (NH*C*H$_2$), 35.5 (*C*H$_2$CONH), 31.3 (*C*H$_2$), 29.0 (*C*H$_2$), 28.9 (*C*H$_2$), 28.7 (*C*H$_2$), 28.7 (*C*H$_2$), 25.3 (*C*H$_2$CH$_2$CONH), 22.2 (*C*H$_2$), 14.0 (*C*H$_3$); **IR** (neat) v_{max} = 3256, 3083, 2915, 2848, 1631, 1545, 1245 1075 cm^{-1}; **HRMS (ESI)**: calculated for $C_{14}H_{28}N_2NaO_3$ [M+Na]$^+$: 295.1992, found: 295.1994.

6.1.3.64 Synthesis of N-(2-decanamidoethyl)-2,2-dimethyl-1,3-dioxolane-4-carboxamide (353)
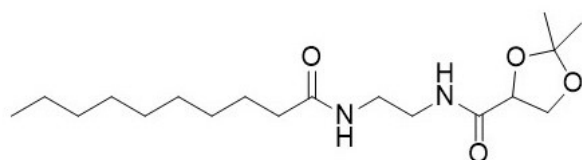
To a solution of N-(2-aminoethyl)decanamide hydrochloride (**350**) (0.15 g, 0.60 mmol), 2,2-dimethyl-1,3-dioxolane-4-carboxylic acid (**352**) (0.11 g, 0.72 mmol), and DMAP (0.22 g, 1.79 mmol) in anhydrous DCM (10 mL) under argon at 0 °C was added DCC (0.19 g, 0.90 mmol) followed by stirring at RT for 16 h. The reaction was concentrated, diluted with EtOAc (15 mL) and filtered through Celite. The filtrate was washed with 1M HCl (10 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to the crude product. Purification by silica column chromatography with a gradient of 70-100% EtOAc in petroleum ether afforded the title compound (**353**) as a white solid (0.10 g, 51%): **TLC** (Rf = 0.25, EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.01 (1H, br s, N*H*COCHO), 6.04 (1H, br s CON*H*), 4.47 (1H, dd, *J* = 7.4, 5.4, NHCOC*H*), 4.28 (1H, t, *J* = 8.2, OCHC*H*$_2$O),

4.06 (1H, dd, J = 8.7, 5.3, OCHCH_2O), 3.47-3.36 (4H, m, NHCH_2CH_2NH), 2.16 (2H, t, J = 7.7, CHCONH), 1.60-1.59 (2H, m, CH_2CH$_2$CONH), 1.50 (3H, s, O$_2$CCH_3), 1.39 (3H, s, O$_2$CCH_3), 1.34-1.20 (12H, m, CH_2), 0.87 (3H, t, J = 6.9); ^{13}**C NMR** (125 MHz, CDCl$_3$): 174.2 (CONH), 172.8 (NHCOCHO), 111.2 (MeCO$_2$), 75.1 (NHCOCHO), 67.9 (NHCOCHCH$_2$O), 40.1 (CH$_2$NHCOCHO), 39.4 (CONHCH$_2$), 36.9 (CH$_2$CONH), 32.0 (CH$_2$), 29.6 (CH$_2$), 29.5 (CH$_2$), 29.4 (CH$_2$), 29.4 (CH$_2$), 26.3 (O$_2$CCH$_3$), 25.9 (CH$_2$CONH), 25.1 (O$_2$CCH$_3$), 14.1 (CH$_3$CH$_2$); **IR** (thin film) ν_{max} = 3309, 2920, 2851, 1644, 1544, 1262, 856 cm^{-1}; **HRMS (ESI)**: calculated for C$_{18}$H$_{34}$N$_2$NaO$_4$ [M+Na]$^+$: 365.2411, found: 365.2406.

6.1.3.65 Synthesis of N-(2-(2,3-dihydroxypropanamido)ethyl)decanamide (347)
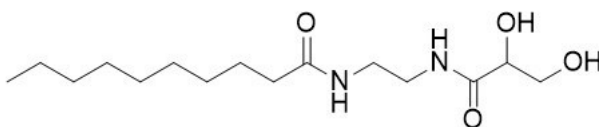
A solution of N-(2-decanamidoethyl)-2,2-dimethyl-1,3-dioxolane-4-carboxamide (**353**) (0.10 g, 0.30 mmol) in THF (1 mL) and 2M HCl (1 mL) was stirred for 16 h. The reaction was then concentrated and purified by silica column chromatography with a gradient of 0-20% MeOH in DCM to afford the title compound (**347**) as a pale yellow solid (0.06 g, 70%): **TLC** (Rf = 0.35, 85:15 DCM : MeOH); ^{1}**H NMR** (500 MHz, MeOD): δ 4.07-4.03 (1H, m, OHCHCH$_2$OH), 3.76 (1H, dd, J = 11.2, 2.8, OHCHCH_2OH), 3.70 (1H, dd, J = 11.2, 5.1, OHCHCH_2OH), 3.36-3.34 (4H, m, NHCH_2CH_2NH), 2.19 (2H, t, J = 7.6, CH_2CONH), 1.64-1.55 (2H, m, CH_2CH$_2$CONH), 1.35-1.24 (12H, m, CH_2), 0.90 (3H, t, J = 6.9, CH_3); ^{13}**C NMR** (125 MHz, MeOD): 177.0 (NHCOCHOH), 175.5 (CONH), 74.1 (NHCOCHOH), 65.3 (OHCHCH_2OH), 40.0 (NHCH$_2$), 39.7 (NHCH$_2$), 37.0 (CH$_2$CONH), 33.0 (CH$_2$), 30.6 (CH$_2$), 30.4 (CH$_2$), 30.4 (CH$_2$), 30.3 (CH$_2$), 26.9 (CH$_2$CH$_2$CONH), 323.7 (CH$_2$), 14.4 (CH$_3$); **IR** (neat) ν_{max} = 3383, 3300, 2964, 2917, 2849, 1640, 1620, 1545, 1069, 683 cm^{-1}; **HRMS (ESI)**: calculated for C$_{15}$H$_{30}$N$_2$NaO$_4$ [M+Na]$^+$: 325.2098, found: 325.2100.

6.1.3.66 Synthesis of (E)-hex-4-enal (358)
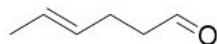
To a solution of (E)-hex-4-en-1-ol (**357**) (0.50 mL, 4.25 mmol) in anhydrous DCM (8 mL) under argon a RT was added PCC (1.37 g, 6.37 mmol) followed by stirring for 16 h. The reaction was triturated with Et$_2$O (20 mL) and filtered through Celite. The filtrate was gently concentrated to the crude product (**358**) as a colourless oil (0.34 g, 81%) and used without further purification: **TLC** (Rf = 0.81, 2:3 Et$_2$O : petroleum ether); **^{1}H NMR** (300 MHz, CDCl$_3$): δ 9.76 (1H, t, *J* = 1.7, C*H*O), 5.55-5.35 (2H, m, CH$_3$C*H*C*H*CH$_2$), 2.54-2.44 (2H, m, C*H*$_2$CHO), 2.36-2.26 (2H, m, C*H*$_2$CH$_2$CHO), 1.67-1.62 (3H, m, C*H*$_3$CH). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[184]

6.1.3.67 Synthesis of methyl (E)-3-hydroxyoct-6-enoate (355)
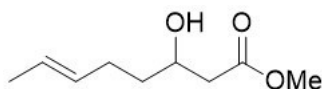
To a solution of anhydrous THF (8 mL) under argon at reflux was added in quick succession activated zinc dust (0.34 g, 5.14 mmol), (E)-hex-4-enal (**358**) (0.34 g, 3.42 mmol) and then methyl bromopropionate (**217**) (0.49 mL, 5.14 mmol) followed by stirring at 70 °C for 10 mins. After the exotherm and colour change, the reaction stirred at RT for 15 mins followed by concentration. The sticky solid was dissolved in 1M HCl (10 mL) and extracted in EtOAc (4 x 10 mL), the combined organics were washed with brine (15 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to the crude product. Purification by silica column chromatography with a gradient of 0-20% EtOAc in petroleum ether afforded the title compound (**355**) as a colourless oil (0.19 g, 32%): **TLC** (Rf = 0.29, 1:4 EtOAc : petroleum ether); **^{1}H NMR** (300 MHz, CDCl$_3$): δ 5.52-5.35 (2H, m, CH$_3$C*H*C*H*CH$_2$), 4.07-3.97 (1H, m, C*H*OH), 3.71 (3H, s, CO$_2$C*H*$_3$), 2.85 (1H, d, *J* = 3.8, O*H*), 2.56-2.41 (2H, m, C*H*$_2$CO$_2$), 2.20-2.02 (2H, m, C*H*$_2$CH$_2$CHOH), 1.65 (3H, d,

J = 4.9, CH_3CH), 1.60-1.45 (2H, m, CH_2CHOH). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[185]

6.1.3.68 Synthesis of methyl 3-hydroxy-5-(3-methyloxiran-2-yl)pentanoate (356)
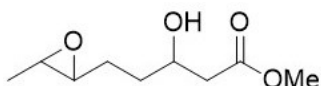
To a solution of methyl (E)-3-hydroxyoct-6-enoate (355) (0.05 g, 0.29 mmol) in anhydrous DCM (2 mL) under argon at 0 °C was added *m*-CPBA (70% wt.) (0.08 g, 0.32 mmol) followed by stirring at RT for 6 h. The reaction was quenched with saturated NaHCO$_{3(aq.)}$ (2 mL) and extracted with Et$_2$O (4 x 2 mL). The combined organics were dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-60% EtOAc in petroleum ether afforded the title compound (356) as a colourless oil (0.03 g, 54%). The purified product comprised of several isomers in unknown quantities, the product was used as a mixture without subsequent separation of isomers: **TLC** (Rf = 0.20, 1:4 EtOAc : petroleum ether); **^{1}H NMR** (300 MHz, CDCl$_3$): δ 4.45-4.30 (1H, m, O*H*), 4.14-3.80 (2H, m, CH$_3$C*H*OC*H*), 3.72-3.68 (3H, m, OC*H$_3$*), 2.81-2.42 (3H, m, *H*OCHC*H$_2$*CO$_2$), 2.18-1.58 (4H, m, C*H$_2$*C*H$_2$*), 1.32-1.08 (3H, m, C*H$_3$*CHOCH).

6.1.3.69 4-decanethioamidobutanoic acid (369)
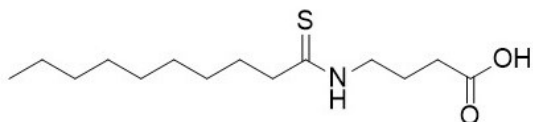
To an argon purged stirring solution of 4-decanamidobutanoic acid (164) (0.50 g, 1.94 mmol) in anhydrous THF (20 mL) at RT, Lawesson's reagent (2,4-Bis(4-methoxyphenyl)-1,3,2,4-dithiadiphosphetane-2,4-dithione) (0.39 g, 0.97 mmol) was added. The reaction was refluxed for 20 mins then cooled and filtered through Celite

using EtOAc (10 mL) to wash the filter cake. The filtrate was concentrated and purified by silica column chromatography with a gradient of 0-50% EtOAc in cyclohexane to give the title compound (**369**) as a white solid (0.15 g, 28%): **TLC** (Rf = 0.65, 1:1 EtOAc : cyclohexane); ^{1}**H NMR** (400 MHz, CDCl$_3$): δ 7.57 (1H, br s, N*H*), 3.76-3.70 (2H, m, NC*H$_2$*), 2.68-2.61 (2H, m, C*H$_2$*CS), 2.50 (2H, t, *J* = 6.8, C*H$_2$*CO$_2$H), 2.01 (2H, quin, *J* = 6.7, NHCH$_2$C*H$_2$*), 1.80-1.74 (2H, m, NHCOCH$_2$C*H$_2$*), 1.35-1.20 (12H, m, C*H$_2$*), 0.88 (3H, t, *J* = 6.8, C*H$_3$*); ^{13}**C NMR** (125 MHz, CDCl$_3$): 206.3 (NH*C*S), 177.9 (*C*O$_2$H), 47.3 (NHCO*C*H$_2$), 45.3 (NH*C*H$_2$), 31.9 (*C*H$_2$), 31.4 (*C*H$_2$CO$_2$H), 29.7 (*C*H$_2$), 29.5 (*C*H$_2$), 29.5 (*C*H$_2$), 29.4 (*C*H$_2$), 29.4 (*C*H$_2$), 29.4 (*C*H$_2$CH$_2$CS), 22.7 (NHCH$_2$*C*H$_2$), 14.1 (*C*H$_3$); **IR** (neat) v_{max} = 3205, 3038, 2916, 2847, 1691, 1536, 1295,1168 cm^{-1}; **HRMS (ESI)**: calculated for C$_{14}$H$_{27}$NNaO$_2$S [M+Na]$^+$: 296.1660, found: 296.1655.

6.1.3.70 Synthesis of Methyl 6-decanethioamido-3-oxohexanoate (361)
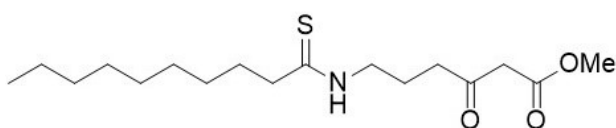
To an argon purged stirring solution of 4-decanethioamidobutanoic acid (**369**) (0.35 g, 1.28 mmol), Meldrum's acid (0.19 g, 1.28 mmol) and DMAP (0.20 g, 1.60 mmol) in anhydrous THF (20 mL) at 0 °C was added EDC.HCl (0.27 g, 1.41 mmol) and the solution was stirred at RT for 16 h. The reaction was then concentrated and re-dissolved in DCM (20 mL). The organic solution was washed with 1M HCl (2 x 10 mL) and then H$_2$O (2 x 10 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude Meldrum's acid adduct (**370**). This was then dissolved in anhydrous MeOH (8 mL) and stirred at 110 °C for 10 mins with microwave irradiation. The reaction was then concentrated and purified by silica column chromatography with a gradient of 0-35% EtOAc in cyclohexane to give the title compound (**361**) as a white solid (0.26 g, 62%): **TLC** (Rf = 0.59, 1:1 EtOAc : cyclohexane); ^{1}**H NMR** (400 MHz, CDCl$_3$): δ 7.70 (1H, br s, N*H*), 3.74 (3H, s, OC*H$_3$*), 3.70-3.68 (2H, m, NHC*H$_2$*), 3.48 (2H, s, COC*H$_2$*CO), 2.71 (2H, t, *J* = 6.4, C*H$_2$*COCH$_2$CO), 2.64-2.61 (2H, m, C*H$_2$*CS), 1.98 (2H, quin, *J* = 6.6, NHCH$_2$C*H$_2$*), 1.76-

1.71 (2H, m, CH$_2$CH$_2$CS), 1.35-1.20 (12H, m, CH$_2$), 0.87 (2H, t, J = 6.8); ^{13}C NMR (125 MHz, CDCl$_3$): 206.1 (CS), 203.3 (CH$_2$COCH$_2$), 167.7 (CO$_2$CH$_3$), 52.5 (OCH$_3$), 48.9 (COCH$_2$CO), 47.3 (CH$_2$COCH$_2$CO), 45.4 (NHCH$_2$), 40.8 (CH$_2$CS), 31.9 (CH$_2$), 29.8 (CH$_2$), 29.5 (CH$_2$), 29.4 (CH$_2$), 29.3 (CH$_2$CH$_2$CS), 29.3 (CH$_2$), 28.9 (CH$_2$), 22.7 (CH$_2$), 21.4 (NHCH$_2$CH$_2$), 14.1 (CH$_3$); IR (neat) v_{max} = 3210, 2916, 2847, 1738, 1708, 1537, 1325, 1250, 1003 cm^{-1}; HRMS (ESI): calculated for C$_{17}$H$_{32}$NO$_3$S [M+H]$^+$: 330.2097, found: 330.2094.

6.1.3.71 Synthesis of methyl 4-(decylsulphonamido)butanoate (373)
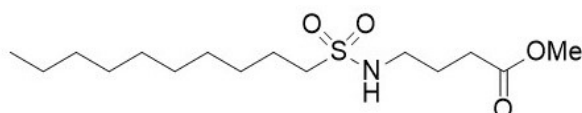
To a solution of methyl 4-aminobutyrate hydrochloride (372) (0.46 g, 2.77 mmol) and Et$_3$N (1.20 mL, 8.31 mmol) in anhydrous DCM (10 mL) at RT under argon was added n-decylsulphonyl chloride (371) (1.00 g, 4.15 mmol). The solution was then stirred at RT for 16 h. The reaction was then diluted with 1M HCl (30 mL) and extracted with DCM (3 x 10 mL). The combined organic extracts were dried over MgSO$_{4(s)}$, filtered, concentrated and purified by silica column chromatography with a gradient of 0-40% EtOAc in cyclohexane to afford the title compound (373) as a white solid (0.91 g, quant.): TLC (Rf = 0.24, 3:1 cyclohexane : EtOAc); ^{1}H NMR (500 MHz, CDCl$_3$): δ 4.32 (1H, t, J = 6.2, NH), 3.69 (3H, s, OCH$_3$), 3.18 (2H, q, J = 6.7, NHCH$_2$), 3.03-2.99 (2H, m, CH$_2$SO$_2$), 2.43 (2H, t, J = 7.0, CH$_2$CO$_2$), 1.89 (2H, quin, J = 6.9, NHCH$_2$CH$_2$), 1.82-1.76 (2H, m, CH$_2$CH$_2$SO$_2$), 1.43-1.39 (2H, m, CH$_2$CH$_2$CH$_2$SO$_2$), 1.33-1.22 (12H, m, CH$_2$), 0.88 (3H, t, J = 6.9, CH$_3$); ^{13}C NMR (125 MHz, CDCl$_3$): 173.6 (CO), 52.8 (CH$_2$SO$_2$), 51.8 (OCH$_3$), 42.6 (NHCH$_2$), 31.9 (CH$_2$), 30.9 (CH$_2$CO), 29.5 (CH$_2$), 29.3 (CH$_2$), 29.2 (CH$_2$), 29.1 (CH$_2$), 28.3 (CH$_2$CH$_2$CH$_2$SO$_2$), 25.4 (NHCH$_2$CH$_2$), 23.7 (CH$_2$CH$_2$SO$_2$), 22.7 (CH$_2$), 14.1 (CH$_3$); IR (neat) v_{max} = 3282, 2918, 2848, 1728, 1423, 1370, 1171, 1133 cm^{-1}; HRMS (ESI): calculated for C$_{15}$H$_{31}$NNaO$_4$S [M+Na]$^+$: 344.1866, found: 344.1865.

6.1.3.72 Synthesis of 4-(decylsulphonamido)butanoic acid (374)
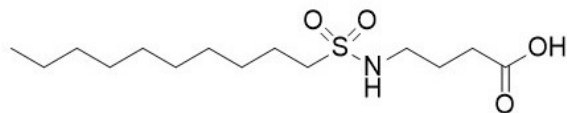
To a solution of methyl 4-(decylsulphonamido)butanoate (**373**) (0.91 g, 2.77 mmol) in MeOH (14 mL) at RT was added 1M NaOH (5.60 mL, 5.60 mmol), the reaction was then stirred at reflux for 16 h. The solvent was then concentrated and the aqueous solution acidified to pH 1 using 1M HCl and extracted using EtOAc (4 x 15 mL). The combined organic extracts were dried over MgSO$_{4(s)}$, filtered and concentrated to afford the title compound (**374**) as a white solid (0.76 g, 88%): **^{1}H NMR** (500 MHz, d_6-DMSO): δ 12.08 (1H, br s, O*H*), 7.01 (1H, t, *J* = 5.8, N*H*), 2.95-2.92 (2H, m, NHC*H$_2$*), 2.94-2.91 (2H, m, C*H$_2$*SO$_2$), 2.25 (2H, t, *J* = 7.4, C*H$_2$*CO$_2$), 1.65 (2H, quin, *J* = 7.1, NHCH$_2$C*H$_2$*), 1.64-1.61 (2H, m, C*H$_2$*CH$_2$SO$_2$), 1.37-1.34 (2H, m, C*H$_2$*CH$_2$CH$_2$SO$_2$), 1.31-1.19 (12H, m, C*H$_2$*), 0.56 (3H, t, *J* = 6.9, C*H$_3$*); **^{13}C NMR** (125 MHz, d_6-DMSO): 174.6 (*C*O$_2$), 51.2 (*C*H$_2$SO$_2$), 42.1 (NH*C*H$_2$), 31.8 (*C*H$_2$), 3.12 (*C*H$_2$CO$_2$), 29.4 (*C*H$_2$), 29.3 (*C*H$_2$), 29.1 (*C*H$_2$), 29.0 (*C*H$_2$), 28.0 (*C*H$_2$CH$_2$CH$_2$SO$_2$), 25.5 (NHCH$_2$*C*H$_2$), 23.7 (*C*H$_2$CH$_2$SO$_2$), 22.6 (*C*H$_2$), 114.4 (*C*H$_3$); **IR** (neat) v_{max} = 3283, 2957, 2917, 2848, 1692, 1423, 1316, 1286, 1133 cm^{-1}; **HRMS (ESI)**: calculated for C$_{14}$H$_{29}$NNaO$_4$S [M+Na]$^{+}$: 330.1710, found: 310.1711.

6.1.3.73 Synthesis of methyl 6-(decylsulphonamido)-3-oxohexanoate (365)
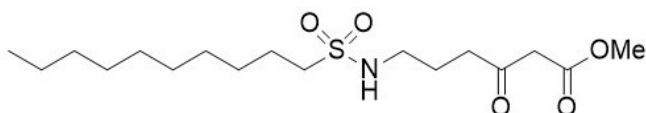
A suspension of MgCl$_2$ (0.06 g, 0.68 mmol) and methyl potassium malonate (**378**) (0.15 g, 0.98 mmol) in anhydrous THF (4 mL) under argon was stirred at 50 °C for 4 h. The suspension was then cooled to RT and to it was added dropwise a solution of 4-(decylsulphonamido)butanoic acid (**374**) (0.02 g, 0.65 mmol) in anhydrous THF (6 mL) under argon which was treated with CDI (0.13 g, 0.78 mmol) at 0 °C and stirred at RT

for 1 h. The reaction was then stirred at RT for 16 h and concentrated and re-dissolved in EtOAc (30 mL). The organic solution was washed sequentially with saturated NaHCO$_{3(aq.)}$ (30 mL), 1M HCl (30 mL) then brine (10 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Silica chromatography purification with a gradient of 0-45% EtOAc in cyclohexane afforded the title compound (365) as a pale pink solid (0.13 g, 55%): **TLC** (R$_f$ = 0.49, 1:1 cyclohexane : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 4.36 (1H, t, *J* = 6.3, N*H*), 3.74 (3H, s, OC*H*$_3$), 3.48 (COC*H*$_2$CO$_2$), 3.14 (2H, q, *J* = 6.6, NHC*H*$_2$), 2.99-2.96 (2H, m, C*H*$_2$SO$_2$), 2.70 (2H, t, *J* = 6.7, C*H*$_2$COCHCO$_2$), 1.86 (2H, quin, *J* = 6.6, NHCH$_2$C*H*$_2$), 1.79-1.76 (2H, m, C*H*$_2$CH$_2$SO$_2$), 1.42-1.39 (2H, m, C*H*$_2$CH$_2$CH$_2$SO$_2$), 1.34-1.21 (12H, m, C*H*$_2$), 0.88 (3H, t, *J* = 6.9, C*H*$_3$); **^{13}C NMR** (125 MHz, CDCl$_3$): 202.3 (CH$_2$*C*OCH$_2$), 167.8 (*C*O$_2$), 52.7 (*C*H$_2$SO$_2$), 52.5 (O*C*H$_3$), 48.9 (*C*H$_2$CO$_2$), 42.2 (NH*C*H$_2$), 39.4 (*C*H$_2$COCH$_2$CO$_2$), 31.8 (*C*H$_2$), 29.5 (*C*H$_2$), 29.3 (*C*H$_2$), 29.2 (*C*H$_2$), 29.1 (*C*H$_2$), 28.3 (*C*H$_2$CH$_2$CH$_2$SO$_2$), 23.9 (NHCH$_2$*C*H$_2$), 23.7 (*C*H$_2$CH$_2$SO$_2$), 22.6 (*C*H$_2$), 14.1 (*C*H$_3$); **IR** (neat) v_{max} = 3248, 2957, 2918, 2848, 1748, 1711, 1425, 1316, 1133, 1067 cm^{-1}; **HRMS (ESI)**: calculated for C$_{17}$H$_{33}$NNaO$_5$S [M+Na]$^+$: 386.1972, found: 386.1974.

6.1.3.74 Synthesis of methyl 4-(3-octylureido)butanoate (376)
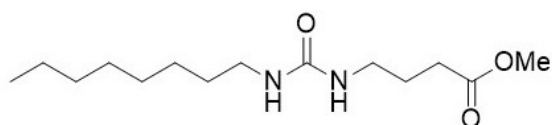
To a solution of nonanoic acid (375) (1.30 mL, 7.39 mmol), anhydrous Et$_3$N (2.40 mL, 17.1 mmol) and anhydrous MeCN (30 mL) under argon at RT, DPPA (1.40 mL, 6.36 mmol) was added dropwise. The solution was gently heated to 50 °C for 3 h. Methyl 4-aminobutyrate hydrochloride (1.02 g, 6.64 mmol) was then added and the reaction was stirred at 50 °C for 2 h. The reaction was then cooled, concentrated and the crude solid diluted with EtOAc (60 mL). The organic phase was washed sequentially with saturated NaHCO$_{3(aq.)}$ (30 mL), 1M HCl (30 mL) and brine (30 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography

with a gradient of 40-100% EtOAc in cyclohexane afforded the title compound (**376**) as a white solid (1.56 g, 91%: **TLC** (Rf = 0.41, 3:1 EtOAc – cyclohexane); **^{1}H NMR** (400 MHz, CDCl$_3$): δ 4.55 (2H, br s, N*H*CON*H*), 3.67 (3H, s, OC*H*$_3$), 3.21 (2H, t, *J* = 6.9, C*H*$_2$CH$_2$CH$_2$CO), 3.13 (2H, t, *J* = 7.2, C*H*$_2$NH), 2.38 (2H, t, *J* = 7.1, C*H*$_2$CO), 1.82 (2H, quin, *J* = 7.0, C*H*$_2$CH$_2$CO), 1.49-1.46 (2H, m, C*H*$_2$CH$_2$NH), 1.34-1.19 (10H, m, C*H*$_2$), 0.87 (3H, t, *J* = 6.9, C*H*$_3$) ; **^{13}C NMR** (125 MHz, CDCl$_3$): 174.2 (*C*O$_2$CH$_3$), 158.3 (NH*C*ONH), 51.7 (O*C*H$_3$), 40.7 (*C*H$_2$NH), 39.9 (*C*H$_2$CH$_2$CH$_2$CO), 31.8 (*C*H$_2$), 31.3 (*C*H$_2$CO), 30.2 (*C*H$_2$CH$_2$NH), 29.3 (*C*H$_2$), 29.2 (*C*H$_2$), 26.9 (*C*H$_2$), 25.3 (*C*H$_2$CH$_2$CO), 22.6 (*C*H$_2$), 14.1 (*C*H$_3$); **IR** (neat) v_{max} = 3331, 2923, 2851, 1725, 1614, 1570, 1201, 1176 cm^{-1}; **HRMS (ESI)**: calculated for C$_{14}$H$_{28}$N$_2$NaO$_3$ [M+Na]$^+$: 295.1992, found: 295.1998.

6.1.3.75 Synthesis of 4-(3-octylureido)butanoic acid (377)
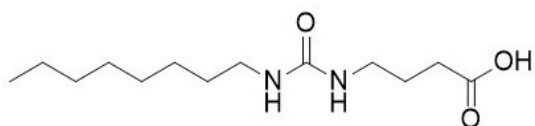
To a solution of methyl 4-(3-octylureido)butanoate (**376**) (1.74 g, 6.40 mmol) in THF (40 mL) and H$_2$O (20 mL) at RT was added LiOH (0.31 g, 12.8 mmol), followed by stirring for 16 h. The reaction was then concentrated and the remaining aqueous was diluted with H$_2$O (100 mL) and acidified to pH 2 using 1M HCl. The reaction was filtered under suction and the filter cake was washed with minimal H$_2$O, collected and dried thoroughly to give the title compound (**377**) as a white solid (1.26 g, 76%): **^{1}H NMR** (500 MHz, d_6-DMSO): δ 12.04 (1H, br s, CO$_2$*H*), 5.81 (1H, t, *J* = 5.2, N*H*CH$_2$CH$_2$CH$_2$CO), 5.75 (1H, t, *J* = 5.1, N*H*CO), 2.98-2.94 (2H, m, C*H*$_2$CH$_2$CH$_2$CO), 2.96-2.92 (2H, m, C*H*$_2$NHCONH), 2.17-2.15 (2H, m, C*H*$_2$CO$_2$H), 1.57 (2H, quin, *J* = 7.2, C*H*$_2$CH$_2$CO), 1.34-1.32 (2H, m, C*H*$_2$CH$_2$NHCONH), 1.28-1.19 (10H, m, C*H*$_2$), 0.85 (3H, t, *J* = 6.9, C*H*$_3$) ; **^{13}C NMR** (125 MHz, d_6-DMSO): 175.1 (*C*O$_2$H), 158.6 (NH*C*ONH), 39.7 (*C*H$_2$CH$_2$CH$_2$CO), 39.1 (*C*H$_2$NH), 31.7 (*C*H$_2$CO), 30.5 (*C*H$_2$CH$_2$NH), 29.4 (*C*H$_2$), 29.3 (*C*H$_2$), 29.2 (*C*H$_2$), 26.9 (*C*H$_2$), 26.2 (*C*H$_2$CH$_2$CO), 22.6 (*C*H$_2$), 14.4 (*C*H$_3$); **IR** (neat) v_{max}

= 3330, 2954, 2921, 2850, 1692, 1612, 1517, 1209, 632 cm^{-1}; **HRMS (ESI)**: calculated for $C_{13}H_{26}N_2NaO_3$ [M+Na]$^+$: 281.1836, found: 281.1834.

6.1.3.76 Synthesis of methyl 6-(3-octylureido)-3-oxohexanoate (362)
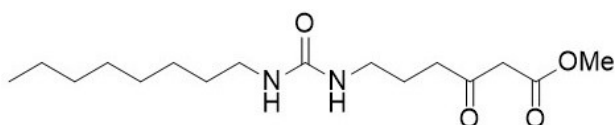
MgCl$_2$ (0.11 g, 1.10 mmol) was added to a solution of methyl potassium malonate (**378**) (0.15 g, 0.93 mmol) and Et$_3$N (0.19 mL, 1.28 mmol) in anhydrous MeCN (2 mL) under argon at 0 °C, the suspension was then stirred at RT for 4 h. To the reaction was added a solution of 4-(3-octylureido)butanoic acid (**377**) (0.15 g, 0.58 mmol) in anhydrous THF (6 mL) which was treated with CDI (0.11 g, 0.70 mmol) under argon at 0 °C and stirred at RT for 4 h. After 16 h stirring at RT, the reaction was filtered using EtOAc (10 mL) and H$_2$O (10 mL) to wash the filter cake, the filtrate was concentrated and the remaining aqueous was diluted with H$_2$O (10 mL) and extracted using EtOAc (3 x 10 mL). The combined organic extracts were washed with saturated NaHCO$_{3(aq.)}$ (10 mL) followed by 1M HCl (10 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 20-100% EtOAc in cyclohexane afforded the title compound (**362**) as a white solid (0.08 g, 44%): **TLC** (Rf = 0.23, 1:3 EtOAc : cyclohexane); **^{1}H NMR** (400 MHz, CDCl$_3$): δ 4.45 (1H, br s, N*H*CH$_2$CH$_2$CH$_2$CO), 4.39-4.36 (1H, m, N*H*CONH), 3.74 (3H, s, OC*H*$_3$), 3.47 (2H, s, COC*H*$_2$CO), 3.19 (2H, q, *J* = 6.2, C*H*$_2$CH$_2$CH$_2$CO), 3.13 (2H, q, *J* = 7.2, C*H*$_2$NH), 2.62 (2H, t, *J* = 6.7, C*H*$_2$CO), 1.81 (2H, quin, *J* = 6.8, C*H*$_2$CH$_2$CO), 1.49-1.45 (2H, m, C*H*$_2$CH$_2$NH), 1.37-1.17 (10H, m, C*H*$_2$), 0.87 (3H, t, *J* = 6.85, CH$_3$); **^{13}C NMR** (125 MHz, CDCl$_3$): 203.0 (CH$_2$*C*OCH$_2$), 167.8 (*C*O$_2$Me), 158.2 (NH*C*ONH), 52.4 (O*C*H$_3$), 49.0 (CO*C*H$_2$CO), 40.7 (*C*H$_2$NH), 40.3 (*C*H$_2$CO), 39.6 (*C*H$_2$CH$_2$CH$_2$CO), 31.8 (*C*H$_2$), 30.2 (*C*H$_2$CH$_2$NH), 29.3 (*C*H$_2$), 29.2 (*C*H$_2$), 26.9 (*C*H$_2$), 23.9 (*C*H$_2$CH$_2$CO), 22.6 (*C*H$_2$), 14.1

(C*H₃*); **IR** (thin film) ν_{max} = 3324, 2922, 2851, 1747, 1713, 1611, 1566, 1323, 1254, 606 cm⁻¹; **HRMS (ESI)**: calculated for $C_{16}H_{29}N_2O_4$ [M-H]⁻ : 313.2133, found: 313.2139.

6.1.3.77 Synthesis of methyl 4-(3-octylthioureido)butanoate (380)
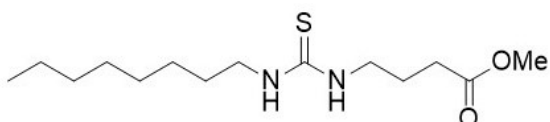
To a solution of methyl 4-aminobutyrate hydrochloride (**372**) (0.20 g, 1.30 mmol) and carbon disulphide (0.09 mL, 1.43 mmol) in anhydrous CH*Cl₃* (3 mL) under argon at RT was added anhydrous Et₃N (0.38 mL, 2.73 mmol) was added dropwise. The solution was stirred for 30 min at RT. Methyl chloroformate (0.11 mL, 1.43 mmol) was then added dropwise and the reaction was stirred at reflux for 4 h. The reaction was then cooled, washed with 1M HCl (5 mL) followed by brine (5 mL), dried over MgSO₄₍ₛ₎, filtered and concentrated to give the isothiocyanate intermediate (**379**) (0.18 g). To a solution of the isothiocyanate (**379**) in anhydrous DCM (3 mL) under argon at RT was added dropwise *n*-octylamine (0.19 mL, 1.15 mmol). After 16 h, the reaction was washed with sequentially with saturated NaHCO₃₍ₐq.₎ (5 mL), 1M HCl (5 mL) and brine (5 mL). The organic layer was then dried over MgSO₄₍ₛ₎, filtered and concentrate to give the crude product. Purification by silica column chromatography with a gradient of 0-60% EtOAc in cyclohexane gave the title compound (**380**) as a colourless oil (0.20 g, 53%): **TLC** (Rf = 0.17 3:1 cyclohexane : EtOAc); **¹H NMR** (500 MHz, CDCl₃): δ 6.16 (1H, br s, N*H*CH₂CH₂CH₂CO), 6.09 (1H, br s, N*H*CS), 3.70 (3H, s, OC*H₃*), 3.49 (2H, br s, C*H₂*CH₂CH₂CO), 3.37 (2H, br s, C*H₂*NH), 2.44 (2H, t, *J* = 6.5, C*H₂*CO), 1.92 (2H, quin, *J* = 6.6, C*H₂*CH₂CO), 1.63-1.60 (2H, m, C*H₂*CH₂NH), 1.40-1.20 (10H, m, C*H₂*), 14.1 (3H, t, *J* = 6.9, C*H₃*); **¹³C NMR** (125 MHz, CDCl₃): 181.5 (*C*S), 174.7 (*C*O₂), 52.0 (O*C*H₃), 44.0 (NH*C*H₂), 43.8 (NH*C*H₂), 31.8 (*C*H₂CO), 29.3 (*C*H₂), 29.2 (*C*H₂), 28.8 (*C*H₂CH₂NH), 26.9 (*C*H₂), 26.9 (*C*H₂), 23.7 (*C*H₂CH₂CO), 22.6 (*C*H₂), 14.1 (*C*H₃); **IR** (neat) ν_{max} = 3251, 2923, 2853, 1734, 1548, 1405, 1220 cm⁻¹; **HRMS (ESI)**: calculated for $C_{14}H_{28}N_2NaO_2S$ [M+Na]⁺: 311.1764, found: 311.1766.

6.1.3.78 Synthesis of 4-(3-octylthioureido)butanoic acid (381)
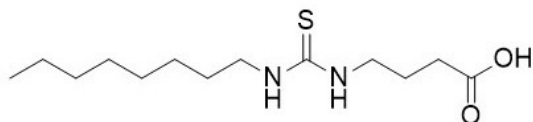
To a solution of methyl 4-(3-octylthioureido)butanoate (**380**) (0.20 g, 0.69 mmol) in THF (24 mL) and H_2O (6 mL) at 0 °C was added dropwise 1M LiOH (2.70 mL, 2.75 mmol) followed by stirring at RT for 16 h. The reaction was concentrated and the remaining aqueous was acidified to pH 2 using 1M HCl. The aqueous layer was then extracted using EtOAc (5 x 5 mL) and the combined organic layers were washed with brine (10 mL), dried over $MgSO_{4(s)}$, filtered and concentrated to give the title compound (**381**) as a white solid (0.21 g, 100%): **^{1}H NMR** (500 MHz, CDCl$_3$): δ 6.11 (1H, br s, N*H*CH$_2$CH$_2$CH$_2$CO), 6.04 (1H, br s, N*H*), 3.53 (2H, br s, C*H*$_2$CH$_2$CH$_2$CO), 3.35 (2H, br s, C*H*$_2$NH), 2.49 (2H, t, J = 6.6, C*H*$_2$CO), 1.95 (2H, quin, J = 6.8, C*H*$_2$CH$_2$CO), 1.62-1.59 (2H, m, C*H*$_2$CH$_2$NH), 1.39-1.21 (10H, m, C*H*$_2$), 0.88 (3H, t, J = 6.9, CH$_3$); **^{13}C NMR** (125 MHz, CDCl$_3$): 181.5 (*C*S), 177 (*C*O$_2$), 44.2 (*C*H$_2$NH), 43.8 (*C*H$_2$CH$_2$CH$_2$CO), 31.8 (*C*H$_2$), 30.7 (*C*H$_2$), 29.7 (*C*H$_2$CO), 29.2 (*C*H$_2$), 29.1 (*C*H$_2$), 28.8 (*C*H$_2$CH$_2$NH), 26.9 (*C*H$_2$), 23.7 (*C*H$_2$CH$_2$CO), 22.6 (*C*H$_2$), 14.1 (*C*H$_3$); **IR** (neat) v_{max} = 3232, 3070, 2919, 2849, 1692, 1557, 1316, 1133, 1063 cm^{-1}; **HRMS (ESI)**: calculated for C$_{13}$H$_{25}$N$_2$O$_2$S [M-H]$^-$: 273.1642, found: 273.1644.

6.1.3.79 Synthesis of methyl 6-(3-octylthioureido)-3-oxohexanoate (363)
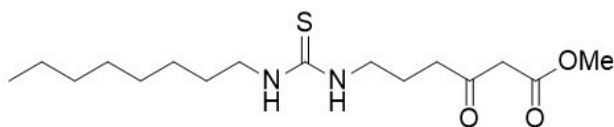
A suspension of MgCl$_2$ (0.04 g, 0.36 mmol) and methyl potassium malonate (**378**) (0.08 g, 0.52 mmol) in anhydrous THF (1 mL) under argon was stirred at 50 °C for 4 h. The suspension was then cooled to RT and to it was added dropwise a solution of 4-(3-octylthioureido)butanoic acid (**381**) (0.09 g, 0.34 mmol) in anhydrous THF (2 mL) under

argon which was treated with CDI (0.07 g, 0.41 mmol) at 0 °C and stirred at RT for 1 h. The reaction was then stirred at RT for 16 h and concentrated and re-dissolved in EtOAc (20 mL). The organic solution was washed sequentially with saturated NaHCO$_{3(aq.)}$ (10 mL), 1M HCl (10 mL) then brine (10 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Silica column chromatography purification with a gradient of 0-60% EtOAc in cyclohexane afforded the title compound (**363**) as a white solid (0.08 g, 71%): **TLC** (Rf = 0.35, 1:1 cyclohexane : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 6.07 (2H, br s, N*H*CSN*H*), 3.74 (3H, s, OC*H*$_3$), 3.50 (2H, s, COC*H*$_2$CO), 3.47 (2H, br s, C*H*$_2$CH$_2$CH$_2$CO), 3.36-3.33 (2H, m, C*H*$_2$NH), 2.70 (2H, t, *J* = 6.2, C*H*$_2$CO), 1.92 (2H, quin, *J* = 6.4, C*H*$_2$CH$_2$CO), 1.62-1.60 (2H, m, C*H*$_2$CH$_2$NH), 1.40-1.21 (10H, m, C*H*$_2$), 0.88 (3H, t, *J* = 6.9, C*H*$_3$); **^{13}C NMR** (125 MHz, CDCl$_3$): 203.6 (CH$_2$*C*OCH$_2$), 181.5 (*C*S), 167.6 (*C*O$_2$), 52.5 (O*C*H$_3$), 48.9 (CO*C*H$_2$CO), 44.0 (*C*H$_2$NH), 43.7 (*C*H$_2$CH$_2$CH$_2$CO), 40.2 (*C*H$_2$COCH$_2$CO), 31.8 (*C*H$_2$), 29.2 (*C*H$_2$), 29.1 (*C*H$_2$), 28.8 (*C*H$_2$CH$_2$NH), 26.9 (*C*H$_2$), 22.6 (*C*H$_2$), 22.4 (*C*H$_2$CH$_2$CO), 14.1 (*C*H$_3$); **IR** (neat) v_{max} = 3244, 3075, 2919, 2850, 1737, 1708, 1560, 1396, 1245, 1008 cm^{-1}; **HRMS (ESI)**: calculated for C$_{16}$H$_{31}$N$_2$O$_3$S [M+H]$^+$: 331.2050, found: 331.2053.

6.1.3.80 2-nonyl-5,6-dihydro-1,3-oxazepin-7(4H)-one (383)
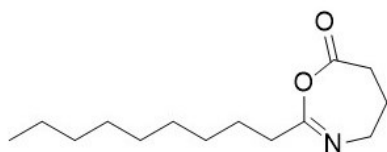
To a solution of 4-decanamidobutanoic acid (**164**) (0.30 g, 1.17 mmol) in anhydrous toluene (8 mL) under argon at RT, Woolins' reagent (2,4-Diphenyl-1,3,2,4-diselenadiphosphetan-2,4-diselenide, 0.31 g, 0.58 mmol) was added. The reaction was stirred under reflux for 1.5 h then cooled and concentrated. Purification by column chromatography with a gradient of 0-40% EtOAc in cyclohexane afforded **383** (instead of the desired **382**) as a red/orange oil (0.19 g, 50%): **TLC** (Rf = 0.69, 1:1 EtOAc : cyclohexane); **^{1}H NMR** (400 MHz, CDCl$_3$): δ 3.80 (2H, t, *J* = 7.5, NC*H*$_2$), 2.88 (2H, t, *J* = 7.5, C*H*$_2$CNO), 2.59 (2H, t, *J* = 8.1, C*H*$_2$CO$_2$), 2.04-2.00 (2H, m, NCH$_2$C*H*$_2$), 1.62 (2H,

quin, J = 7.5, CH_2CH_2CNO), 1.38-1.19 (12H, m, CH_2), 0.87 (3H, t, J = 6.8, CH_3); ^{13}C

NMR (125 MHz, CDCl$_3$): 175.3 (CO$_2$), 174.6 (CNO), 45.5 (NCH$_2$), 36.8 (CH$_2$CNO),

33.8 (CH$_2$CO$_2$), 31.9 (CH$_2$), 29.5 (CH$_2$), 29.4 (CH$_2$), 29.3 (CH$_2$), 29.2 (CH$_2$), 24.2

(CH$_2$CH$_2$CNO), 22.7 (CH$_2$), 17.2 (NCH$_2$CH$_2$), 14.1 (CH$_3$); **IR** (thin film) v_{max} = 2922,

2852, 1737, 1692, 1318, 1021 cm^{-1}; **HRMS (ESI)**: calculated for C$_{14}$H$_{25}$NNaO$_2$

[M+Na]$^+$: 262.1778, found: 262.1777.

6.1.3.81 Synthesis of methyl 4-formamidobutanoate (384)
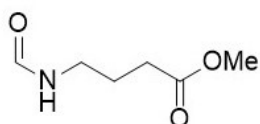
A solution of methyl 4-aminobutyrate hydrochloride (**372**) (0.50 g, 3.25 mmol) and

trimethyl orthoformate (1.20 mL, 10.9 mL) were stirred at reflux for 6 h. The reaction

was then concentrated to give the title compound (**384**) as a colourless oil (0.49 g, 100%).

^{1}**H NMR** (500 MHz, CDCl$_3$): δ 8.16 (1H, s, CHO), 5.96 (1H, br s, NH), 3.67 (3H, s,

OCH_3), 3.34 (2H, q, J = 6.6, NHCH_2), 2.38 (2H, t, J = 7.2, CH_2CO$_2$), 1.86 (2H, quin, J

= 6.9, NHCH$_2$CH_2). Proton NMR data gathered for this compound were in accordance

with those reported in the literature.[186]

6.1.3.82 Synthesis of methyl 4-isocyanobutanoate (385)
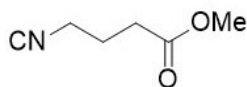
To a solution of methyl 4-formamidobutanoate (**384**) (0.49 g, 3.25 mmol) and anhydrous

DIPEA (1.70 mL, 9.76 mmol) in anhydrous DCM (9 mL) under argon at 0 °C was added

slowly dropwise phosphorus oxychloride (0.36 mL, 3.90 mmol), the solution was stirred

at RT for 2 h. The reaction was then carefully quenched with 20% K$_2$CO$_3$ solution (6 mL)

and stirred for another 30 mins. H$_2$O (20 mL) was then added and the reaction was

extracted with DCM (4 x 10 mL) The combined organic layers were dried over MgSO$_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-50% EtOAc in cyclohexane afforded the title compound (385) as a yellow oil (0.30 g, 74%): **TLC** (Rf = 0.28 3:1 cyclohexane : EtOAc); **^{1}H-NMR** (500 MHz, CDCl$_3$): δ 8.16 (1H, s, *H*CONH), 5.96 (1H, br s, N*H*), 3.67 (3H, s, OC*H*$_3$), 3.34 (2H, q, *J* = 6.6, NHC*H*$_2$), 2.38 (2H, t, *J* = 7.2, C*H*$_2$CO$_2$), 1.86 (2H, quin, *J* = 6.9, C*H*$_2$CH$_2$CO$_2$). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[186]

6.1.3.83 Synthesis of methyl 4-isoselenocyanatobutanoate (386)
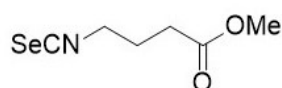
To a solution of methyl 4-isocyanobutanoate (385) (0.20, 1.61 mmol) and anhydrous Et$_3$N (0.16 mL, 1.12 mmol) in anhydrous THF (15 mL) under argon was added selenium (0.38 g, 4.82 mmol). The reaction was stirred at reflux for 16 h in darkness and then filtered through Celite. The filtrate was concentrated to give the title compound (386) as a brown oil (0.17 g, 51%): **TLC** (Rf = 0.43 3:1 cyclohexane : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 3.71 (3H, s, OC*H*$_3$), 3.74-3.69 (2H, m, NC*H*$_2$), 2.49 (2H, t, *J* = 7.1, C*H*$_2$CO), 2.05 (2H, quin, *J* = 6.9, NCH$_2$C*H*$_2$); **^{13}C NMR** (125 MHz, CDCl$_3$): 172.4 (*C*O$_2$), 123.4 (Se*C*N), 52.0 (*C*H$_3$), 44.6 (N*C*H$_2$), 30.4 (*C*H$_2$CO), 24.7 (NCH$_2$*C*H$_2$); **^{77}Se NMR** (115 MHz, CDCl$_3$): -355.8 (*Se*CN); **IR** (thin film) v_{max} = 2949, 2133, 1731, 1435, 1169 cm^{-1}; **HRMS (ESI)**: calculated for C$_6$H$_9$NNaO$_2$Se [M+Na]$^+$: 229.9691, found: 229.9687.

6.1.3.84 Synthesis of methyl 4-(3-octylselenoureido)butanoate (387)
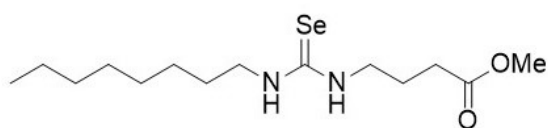
To a solution of methyl 4-isoselenocyanatobutanoate (**386**) (0.17 g, 0.81 mmol) in anhydrous DCM (5 mL) under argon was added *n*-octylamine (0.16 mL, 0.97 mmol) at RT. The reaction was stirred in darkness for 2 h and then concentrated. Purification of the crude by silica chromatography with a gradient of 0-50% EtOAc in cyclohexane gave the title compound (**387**) as a pink solid (0.18 g, 66%): TLC (Rf = 0.56, 1:1, cyclohexane : EtOAc; **¹H NMR** (500 MHz, CDCl₃): δ 6.64 (1H, br s, N*H*), 6.51 (1H, br s, N*H*), 3.70 (3H, s, OC*H₃*), 3.60 (2H, br s, C*H₂*CH₂CH₂CO), 3.21 (2H, br s, C*H₂*NH), 2.45 (2H, s, C*H₂*CO), 1.93 (2H, s, C*H₂*CH₂CO), 1.66-1.62 (2H, m, C*H₂*CH₂NH), 1.39-1.23 (10H, m, C*H₂*), 0.88 (3H, t, *J* = 6.9, C*H₃*); **¹³C NMR** (125 MHz, CDCl₃): 175.1 (*C*O₂), 52.1 (O*C*H₃), 31.8 (*C*H₂), 31.8 (*C*H₂), 29.2 (*C*H₂CO₂), 29.1 (*C*H₂CH₂NH), 26.9 (*C*H₂), 23.3 (*C*H₂CH₂CO₂), 22.6 (*C*H₂), 22.6 (*C*H₂), 14.1 (*C*H₃); **⁷⁷Se NMR** (115 MHz, CDCl₃): 178 (NHC*Se*); **IR** (thin film) v_{max} = 3225, 2924, 2854, 1736, 1558, 1507, 1437, 1437, 1174 cm⁻¹; **HRMS (ESI)**: calculated for $C_{14}H_{28}N_2NaO_2Se$ [M+Na]⁺: 359.1209, found: 359.1224.

6.1.3.85 Synthesis of Methyl 6-(3-octylselenoureido)-3-oxohexanoate (364)
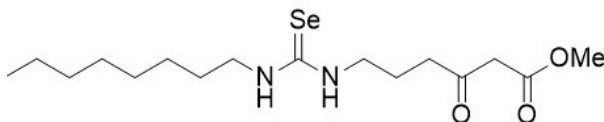
To a solution of methyl 4-(3-octylselenoureido)butanoate (**387**) (0.09 g, 0.27 mmol) in MeOH (7 mL) and H₂O (2 mL) at RT was added K₂CO₃ (0.07 g, 0.54 mmol) followed by stirring in darkness for 16 h. The reaction was then acidified to pH 2 using 1M HCl and concentrated. The remaining aqueous layer was extracted with DCM (5 x 2 mL). The combined organic extracts were dried over MgSO₄₍ₛ₎, filtered and concentrated to give the crude carboxylic acid (**388**) (0.10 g). A suspension of MgCl₂ (0.03 g, 0.28 mmol) and methyl potassium malonate (0.06 g, 0.40 mmol) in anhydrous THF (2 mL) under argon was stirred at 50 °C for 4 h. The solution was then cooled to RT and to it was added dropwise a solution of the carboxylic acid (**388**) (0.10 g) in anhydrous THF (2 mL) which was treated with CDI (0.05 g, 0.32 mmol) and stirred at RT for 1 h. The resulting solution

was stirred in darkness for 16 h at RT. The reaction was concentrated, diluted with H_2O (10 mL) and extracted using EtOAc (4 x 5 mL). The combined EtOAc extracts were washed sequentially with 0.1 M HCl (10 mL), saturated $NaHCO_{3(aq.)}$ (10 mL) and brine (10 mL), dried over $MgSO_{4(s)}$, filtered and concentrated to give the crude product. Purification by silica column chromatography with a gradient of 0-50% EtOAc in cyclohexane afforded the title compound (**364**) as a light brown solid (0.05 g, 50%): **TLC** (Rf = 0.36, 1:1 cyclohexane : EtOAc); **^{1}H NMR** (500 MHz, $CDCl_3$): δ 6.51 (2H, br s, N*H*), 3.74 (3H, s, OC*H*$_3$), 3.70 (2H, br s, C*H*$_2$CH$_2$CH$_2$CO$_2$), 3.49 (2H, s, COC*H*$_2$CO), 3.2 (2H, br s, C*H*$_2$NH), 2.71 (2H, br s, C*H*$_2$COCHCO$_2$), 1.93 (2H, br s, C*H*$_2$CH$_2$CO), 1.63 (2H, br s, C*H*$_2$CH$_2$NH), 1.40-1.19 (10H, m, C*H*$_2$), 0.87 (3H, t, *J* = 6.8, C*H*$_3$); **^{13}C NMR** (125 MHz, $CDCl_3$): 204.0 (CH$_2$*C*OCH$_2$), 167.6 (*C*O$_2$), 52.5 (O*C*H$_3$), 48.9 (CO*C*H$_2$CO), 41.0 (*C*H$_2$COCH$_2$CO$_2$), 31.8 (*C*H$_2$), 29.7 (*C*H$_2$), 29.2 (*C*H$_2$), 29.1 (*C*H$_2$CH$_2$NH), 26.9 (*C*H$_2$), 22.2 (*C*H$_2$), 22.1 (*C*H$_2$), 14.1 (*C*H$_3$); **^{77}Se NMR** (115 MHz, $CDCl_3$): 177.5 (NHC*Se*); **IR** (thin film) v_{max} = 3230, 2923, 2853, 1741, 1711, 1556, 1437, 1319, 1263 cm^{-1}; **HRMS (ESI)**: calculated for $C_{16}H_{30}N_2NaO_3Se$ [M+Na]$^+$: 401.1314, found: 401.1319.

6.1.3.86 Synthesis of *n*-decanitrile (393)

n-Decanal (**392**) (0.71 mL, 3.75 mmol) was added to a solution of hydroxylamine hydrochloride (0.34 g, 4.88 mmol) and NaI (0.28 g, 1.88 mmol) in MeCN (15 mL) at RT. The reaction was then stirred at 85 °C for 1.5 h. The reaction was cooled to RT and stirred whilst adding 5% $NaS_2O_{3(aq)}$ (27 mL). The reaction was further stirred for an additional 15 mins and concentrated. The remaining aqueous solution was extracted using DCM (3 x 10 mL) and the combined organics were dried over $MgSO_{4(s)}$, filtered, concentrated and purified by silica column chromatography with a gradient 0-2% EtOAc in cyclohexane to give the title compound (**393**) as a yellow oil (0.49 g, 85%): **TLC** (Rf = 0.26, 1:50

EtOAc : cyclohexane); **^{1}H NMR** (400 MHz, CDCl$_3$): δ 2.33 (2H, t, J = 7.1, CH_2CN), 1.68-1.64 (2H, m, CH_2CH$_2$CN), 1.45-1.42 (2H, m, CH_2CH$_2$CH$_2$CN), 1.36-1.20 (10H, m, CH_2), 0.88 (3H, t, J = 6.8, CH_3). Proton NMR data gathered for this compound were in accordance with those reported in the literature. [187]

6.1.3.87 Synthesis of decanimidamidobutanoic acid (395)
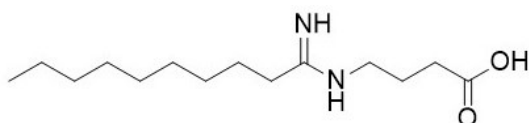
To an argon purged stirring solution of decanitrile (**393**) (0.80 g, 5.23 mmol) in anhydrous MeOH (0.21 mL, 5.75 mmol) at RT was added 4M HCl in dioxane (2.62 mL, 10.5 mmol). The solution was stirred at RT for 7 h then vigorously concentrated. The residue was triturated with Et$_2$O (20 mL) and filtered. The filtered solid was collected and dried to give the crude imidate salt (**394**) (0.66 g). To a solution of the imidate salt (**394**) (0.66 g, 2.98 mmol) and GABA (**163**) (0.31 g, 2.98 mmol) in MeOH (3 mL) at 0 °C 6M KOH was added until the pH reached 9.6. Toluene (3 mL) was then added and the reaction was stirred at RT for 16 h. The reaction was neutralised to pH 7 using 1M HCl and vigorously concentrated to give a white viscous oil. This was triturated with cold EtOH (3 mL) and filtered, the filtrate was concentrated to the crude product as a grey oil. Purification by silica column chromatography with a gradient of 0-50% MeOH in DCM afforded the title compound (**395**) as a very hygroscopic white solid (0.56 g, 42%): **TLC** (Rf = 0.71, 3:2 DCM : MeOH); **^{1}H NMR** (500 MHz, d_6-DMSO): δ 9.65 (1H, br s, NH), 9.08 (1H, br s, NH), 8.81 (1H, br s, NH), 3.20-3.18 (2H, m, NHCH_2), 2.38-2.36 (2H, m, CH_2CN$_2$), 2.27 (2H, t, J = 7.2, CH_2CO), 1.73 (2H, quin, J = 7.2, CH_2CH$_2$CO), 1.60-1.57 (2H, m, CH_2CH$_2$CN$_2$), 1.32-1.20 (10H, m, CH_2), 0.86 (3H, t, J = 6.9 CH_3); **^{13}C NMR** (125 MHz, d_6-DMSO): 175.4 (CO$_2$), 167.8 (CN$_2$), 41.6 (NHCH$_2$), 32.7 (CH$_2$CN$_2$), 31.7 (CH$_2$), 31.5 (CH$_2$CO$_2$), 29.3 (CH$_2$), 29.1 (CH$_2$), 28.9 (CH$_2$), 28.5 (CH$_2$), 27.1 (CH$_2$CH$_2$CN$_2$), 23.3 (CH$_2$CH$_2$CO$_2$), 22.6 (CH$_2$), 14.4 (CH$_3$); **IR** (thin film) v_{max} = 3218, 3048, 2925, 2855, 1680, 1636, 1476 cm^{-1}; **HRMS (ESI)**: calculated for C$_{14}$H$_{29}$N$_2$O$_2$ [M+H]$^+$: 257.2224, found: 257.2227.

6.1.3.88 Synthesis of 3-((4,5-dimethoxy-2-nitrobenzyl)oxy)-3-oxopropanoic acid (421)
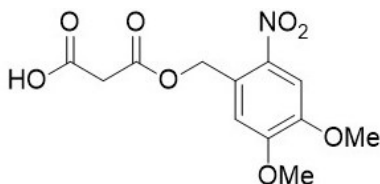
A stirring solution of (4,5-dimethoxy-2-nitrophenyl)methanol (**420**) (0.30 g, 1.41 mmol) and Meldrum's acid (0.20 g, 1.41 mmol) under argon in anhydrous toluene (7 mL) was refluxed for 3 h, then cooled and filtered under suction. The filter cake was collected and dried to give the title compound (**421**) as a yellow solid (0.40 g, 95%): **^{1}H NMR** (300 MHz, D_2O): δ 7.87 (1H, s, C*H*CNO$_2$), 7.23 (1H, s, C*H*CH$_2$O), 5.59 (2H, s, C*H$_2$*CO$_2$), 4.00 (3H, s, CH$_2$CCHCOC*H$_3$*), 3.97 (3H, s, C*H$_3$*OCCHCNO$_2$). Proton NMR data gathered for this compound were in accordance with those reported in the literature.[188]

6.1.3.89 Synthesis of potassium 3-((4,5-dimethoxy-2-nitrobenzyl)oxy)-3-oxopropanoate (422)
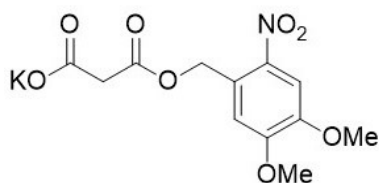
To a solution of 3-((4,5-dimethoxy-2-nitrobenzyl)oxy)-3-oxopropanoic acid (**421**) (0.40 g, 1.34 mmol) in MeOH (20 mL) at 0 °C was added dropwise a solution of KOH (0.08 g, 1.34 mmol) in MeOH (2 mL). After stirring for 10 mins, the reaction was vigorously concentrated to give the title compound (**422**) as a yellow solid (0.46 g, 100%) which was used without further purification. **^{1}H NMR** (300 MHz, D_2O): δ 7.79 (1H, s, C*H*CNO$_2$), 7.19 (1H, s, C*H*CH$_2$O), 5.23 (2H, s, C*H$_2$*CO$_2$), 3.98 (3H, s, CH$_2$CCHCOC*H$_3$*), 3.93 (3H, s, C*H$_3$*OCCHCNO$_2$).

6.1.3.90 Synthesis of 4,5-dimethoxy-2-nitrobenzyl 6-decanamido-3-oxohexanoate (423)
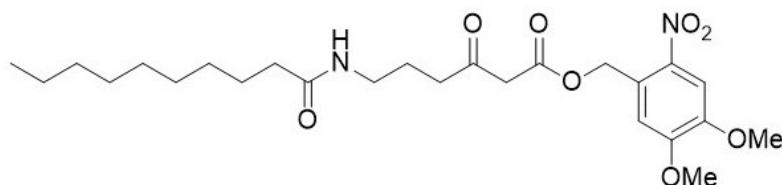
A stirring solution of N-(4-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-ylidene)-4-hydroxybutyl)decanamide (**165**) (0.10 g, 0.26 mmol) and 4,5-dimethoxy-2-nitrobenzyl alcohol (**420**) (0.07 g, 0.31 mmol) under argon in anhydrous toluene (5 mL) was refluxed for 4 h then concentrated. Purification of the crude product by silica chromatography with a gradient of 20-90% EtOAc in petroleum ether afforded the title compound (**423**) as a white solid (0.08 g, 65%): **TLC** (Rf = 0.40, 3:1 EtOAc : petroleum ether); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.76 (1H, s, C*H*CNO$_2$), 7.18 (1H, s, C*H*CCNO$_2$), 5.62 (2H, s, CO$_2$C*H$_2$*C), 5.59-5.55 (1H, m, N*H*), 4.08 (3H, s, NO$_2$CCHCOC*H$_3$*), 3.99 (3H, s, NO$_2$CCCHCOC*H$_3$*), 3.63 (2H, s, COC*H$_2$*CO), 3.29 (2H, q, *J* = 6.6, CONHC*H$_2$*), 2.65 (2H, t, *J* = 6.9, COCH$_2$COC*H$_2$*), 2.17 (2H, t, *J* = 7.6, C*H$_2$*CONH), 1.84 (2H, quin, *J* = 7.2, C*H$_2$*CH$_2$NH), 1.63 (2H, quin, *J* = 7.0, C*H$_2$*CH$_2$CONH), 1.60-1.55 (10H, m, C*H$_2$*), 1.30-1.28 (3H, m, C*H$_2$*C*H$_3$*); **^{13}C NMR** (125 MHz, CDCl$_3$): 202.5 (CH$_2$*C*OCH$_2$), 173.5 (R*C*ONH), 166.6 (R*C*O$_2$R), 153.9 (NO$_2$CC*C*HCOCH$_3$), 148.3 (NO$_2$C*C*HCOCH$_3$), 139.6 (NO$_2$*C*), 126.8 (NO$_2$C*C*), 110.5 (NO$_2$*C*H), 108.2 (NO$_2$C*C*H), 64.1 (CO$_2$*C*H$_2$C), 56.8 (H$_3$*C*OCCHC), 56.4 (H$_3$*C*OCCHCNO$_2$), 49.2 (CO$_2$*C*H$_2$CO), 40.5 (CONH*C*H$_2$), 38.5 (CO$_2$CH$_2$CO*C*H$_2$), 36.9 (*C*H$_2$CONH), 31.9 (*C*H$_2$CH$_2$COCH$_2$), 29.5 (CH$_2$), 29.4 (CH$_2$), 29.3 (CH$_2$), 29.3 (CH$_2$), 25.8 (CH$_2$), 23.6 (CH$_2$), 22.7 (*C*H$_2$CH$_3$), 14.1 (CH$_3$); **IR** (thin film) v_{max} =3326, 2921, 2846, 1709, 1750, 1214, 1630, 1271, 1517, 1325 cm^{-1}; **HRMS (ESI)**: calculated for C$_{25}$H$_{37}$N$_2$O$_8$ [M-H]$^-$: 493.2551, found: 493.2555.

6.1.3.91 Synthesis of 4,5-dimethoxy-2-nitrobenzyl 6-(10-azidodecanamido)-3-oxohexanoate (426)
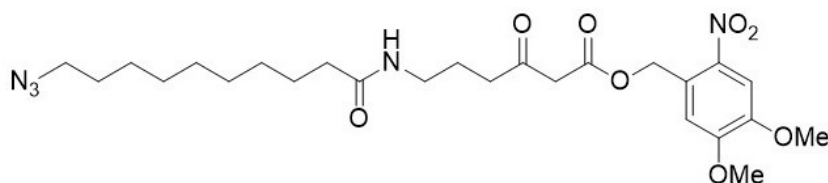
A solution of 10-azido-N-(4-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-ylidene)-4-hydroxybutyl)[85] (**425**) (0.30 g, 0.71 mmol) and 4,5-dimethoxy-2-nitrobenzyl alcohol (**420**) (0.15 g, 0.71 mmol) in anhydrous toluene (6 mL) under argon was stirred at 110 °C for 16 h followed by concentration to the crude product. Purification by silica column chromatography with a gradient of 60-100% EtOAc in petroleum ether afforded the title compound (**426**) as an off-white solid (0.28 g, 74%): **TLC** (Rf = 0.30, EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.70 (1H, s, NO$_2$CC*H*), 7.10 (1H, s, CO$_2$CH$_2$CC*H*), 5.60 (2H, s, CO$_2$C*H$_2$*), 5.58 (1H, br s, N*H*), 4.05 (3H, s, *H$_3$*COCCHCNO$_2$), 3.96 (3H, s, *H$_3$*COCCHC), 3.60 (2H, s, C*H$_2$*CO$_2$), 3.29-3.22 (2H, m, NHC*H$_2$*), 3.29-3.22 (2H, m, N$_3$C*H$_2$*), 2.62 (2H, t, J = 6.9, C*H$_2$*COCH$_2$CO$_2$), 2.14 (2H, t, J = 7.6, C*H$_2$*CONH), 1.81 (2H, quin, J = 6.9, NHCH$_2$C*H$_2$*), 1.64-1.57 (2H, m, N$_3$CH$_2$C*H$_2$*), 1.64-1.57 (2H, m, C*H$_2$*CH$_2$CONH), 1.38-1.23 (10H, m, C*H$_2$*); **^{13}C NMR** (125 MHz, CDCl$_3$): 202.6 (CH$_2$*C*OCH$_2$), 173.5 (*C*ONH), 166.7 (*C*O$_2$), 154.0 (NO$_2$CCH*C*), 148.4 (OCH$_2$C*C*COCH$_3$), 139.8 (OCH$_2$*C*CCOCH$_3$), 126.9 (NO$_2$*C*), 110.6 (OCH$_2$CC*C*OCH$_3$), 108.3 (NO$_2$CC*H*), 64.3 (*C*O$_2$CH$_2$), 56.9 (H$_3$*C*OCCHC), 56.6 (H$_3$*C*OCCHCNO$_2$), 51.6 (N$_3$*C*H$_2$), 49.3 (*C*H$_2$CO$_2$), 40.6 (*C*H$_2$COCH$_2$CO$_2$), 38.4 (NH*C*H$_2$), 36.9 (*C*H$_2$CONH), 29.4 (*C*H$_2$), 29.3 (*C*H$_2$), 29.3 (*C*H$_2$), 29.2 (N$_3$CH$_2$*C*H$_2$), 28.9 (*C*H$_2$), 26.8 (*C*H$_2$), 25.8 (*C*H$_2$CH$_2$CONH), 23.7 (NHCH$_2$*C*H$_2$); **IR** (thin film) υ_{max} = 3396, 3299, 2927, 2854, 2094, 1748, 1715, 1644, 1523, 1277, 1068 cm^{-1}; **HRMS (ESI)**: calculated for C$_{25}$H$_{37}$N$_5$NaO$_8$ [M+Na]$^+$: 558.2534, found: 558.2533.

6.1.3.92 Synthesis of 4,5-dimethoxy-2-nitrobenzyl 6-(decylsulphonamido)-3-oxohexanoate (427)
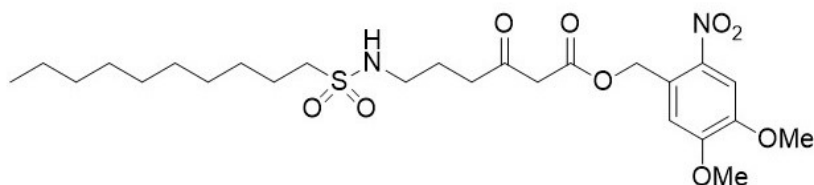
To a solution of 4-(decylsulphonamido)butanoic acid (**374**) (0.07 g, 0.24 mmol) in anhydrous THF (3 mL) under argon at RT was added CDI (0.05 g, 0.29 mmol), followed by stirring for 2 h. Potassium 3-((4,5-dimethoxy-2-nitrobenzyl)oxy)-3-oxopropanoate (**422**) (0.12 g, 0.36 mmol) and $MgCl_2$ (0.02 g, 0.24 mmol) were added to the reaction followed by stirring for 16 h. The reaction was concentrated, partitioned between H_2O (10 mL) and EtOAc (10 mL) and the aqueous layer was acidified to pH 2 using 1M HCl. The aqueous layer was extracted using EtOAc (3 x 5 mL) and the combined organic layers were washed with both saturated $NaHCO_{3(aq.)}$ (10 mL) and brine (10 mL), dried over $MgSO_{4(s)}$, filtered and concentrated to the crude product. Purification by silica chromatography with a gradient of 0-50% EtOAc in petroleum ether afforded the title compound (**427**) as a white solid (0.05 g, 37%): **TLC** (Rf = 0.31, 3:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.73 (1H, s, O$_2$NCC*H*), 7.13 (1H, s, O$_2$NCCC*H*), 5.60 (2H, s, CO$_2$C*H*$_2$), 4.23 (1H, t, *J* = 6.3, N*H*), 4.05 (3H, s, O$_2$NCCHCOC*H*$_3$), 3.96 (3H, s, CH$_2$CCHCOC*H*$_3$), 3.62 (2H, s, COC*H*$_2$CO$_2$), 3.14 (2H, q, *J* = 6.5, NHC*H*$_2$), 3.01-2.96 (2H, m, C*H*$_2$SO$_2$), 2.72 (2H, t, *J* = 6.7, C*H*$_2$COCH$_2$CO$_2$), 1.86 (2H, quin, *J* = 6.6, NHCH$_2$C*H*$_2$), 1.82-1.73 (2H, m, C*H*$_2$CH$_2$SO$_2$), 1.45-1.36 (2H, m, C*H*$_2$CH$_2$CH$_2$SO$_2$), 1.35-1.21 (12H, m, C*H*$_2$), 0.88 (3H, t, *J* = 6.9, C*H*$_3$CH$_2$); **^{13}C NMR** (125 MHz, CDCl$_3$): 202.3 (*C*O), 166.7 (*C*O$_2$), 154.0 (O$_2$NCCHCOCH$_3$), 148.5 (CH$_2$CC*H*COCH$_3$), 140.0 (CO$_2$CH$_2$*C*), 126.8 (O$_2$N*C*), 110.8 (CO$_2$CH$_2$*C*H), 108.4 (O$_2$NC*C*H), 64.4 (CO$_2$*C*H$_2$), 56.9 (O$_2$NCCHCO*C*H$_3$), 56.6 (CH$_2$CCHCO*C*H$_3$), 52.9 (*C*SO$_2$), 49.3 (CO*C*H$_2$CO$_2$), 42.3 (NH*C*H$_2$), 39.7 (*C*H$_2$COCH$_2$CO$_2$), 32.0 (*C*H$_2$), 29.6 (*C*H$_2$), 29.5 (*C*H$_2$), 29.4 (*C*H$_2$), 29.2 (*C*H$_2$), 28.4 (*C*H$_2$CH$_2$CH$_2$SO$_2$), 24.2 (NHCH$_2$*C*H$_2$), 23.9 (*C*H$_2$CH$_2$SO$_2$), 22.8 (*C*H$_2$), 14.3 (*C*H$_3$CH$_2$); **IR** (thin film) v_{max} = 3295, 2923, 2852, 1749, 1714, 1522, 1276, 1169, 1140

cm^{-1}; **HRMS (ESI)**: calculated for C$_{25}$H$_{40}$N$_2$NaO$_9$S [M+Na]$^+$: 567.2347, found: 567.2351.

6.1.3.93 Synthesis of 4,5-dimethoxy-2-nitrobenzyl 6-(10-azidodecanamido)-2-fluoro-3-oxohexanoate (428)

To a solution of 4,5-dimethoxy-2-nitrobenzyl 6-(10-azidodecanamido)-3-oxohexanoate (**426**) (0.18 g, 0.34 mmol) and cyclopentadienyltitanium trichloride (0.01 g, 0.02 mmol) in anhydrous MeCN (10 mL) under argon at RT was added Selectfluor (0.13 g, 0.37 mmol). After 16 h stirring, the reaction was concentrated, dilute with H$_2$O (15 mL) and extracted with DCM (3 x 10 mL). The combined organic extracts were dried over MgSO$_{4(s)}$, filtered and concentrated to the crude product. Purification by silica chromatography with a gradient of 10-80% EtOAc in petroleum ether afforded the title compound (**428**) as a yellow solid (0.09 g, 45%): **TLC** (Rf = 0.46, 3:1 EtOAc : petroleum ether); ^{1}**H NMR** (500 MHz, CDCl$_3$): δ 7.75 (1H, s, NO$_2$CC*H*), 7.14 (1H, s, CO$_2$CH$_2$CC*H*), 5.85 (1H, d, *J* = 14.8, C*H*$_2$O), 5.58 (1H, d, *J* = 14.8, C*H*$_2$O), 5.55 (1H, br s, N*H*), 5.37 (1H, d, *J* = 49.0, C*H*F), 4.05 (3H, s, *H*$_3$COCCHCNO$_2$), 3.97 (3H, s, *H*$_3$COCCHC), 3.31-3.22 (2H, m, NHC*H*$_2$), 3.31-3.22 (2H, m, N$_3$C*H*$_2$), 2.86-2.68 (2H, m, C*H*$_2$COCH$_2$CO$_2$), 2.14 (2H, t, *J* = 7.5, C*H*$_2$CONH), 1.88-1.79 (2H, m, NHCH$_2$C*H*$_2$), 1.63-1.57 (2H, m, N$_3$CH$_2$C*H*$_2$), 1.63-1.57 (2H, m, C*H*$_2$CH$_2$CONH), 1.39-1.24 (10H, m, C*H*$_2$); ^{13}**C NMR** (125 MHz, CDCl$_3$): 201.3 (d, *J* = 23.0, *C*OCHF), 173.6 (*C*ONH), 163.7 (d, *J* = 24.0, *C*O$_2$), 154.0 (NO$_2$CCH*C*), 148.7 (OCH$_2$C*C*COCH$_3$), 139.7 (OCH$_2$*C*CCOCH$_3$), 126.0 (NO$_2$*C*), 110.6 (OCH$_2$CC*C*OCH$_3$), 108.4 (NO$_2$CC*H*), 91.4 (d, *J* = 24.0, *C*HF), 65.2 (CO$_2$*C*H$_2$), 56.9 (H$_3$*C*OCCHC), 56.6 (H$_3$*C*OCCHCNO$_2$), 51.6 (N$_3$*C*H$_2$), 38.4 (NH*C*H$_2$), 36.9 (*C*H$_2$CONH), 35.9 (*C*H$_2$COCH$_2$CO$_2$), 29.4 (*C*H$_2$), 29.3 (*C*H$_2$), 29.3 (*C*H$_2$), 29.2 (N$_3$CH$_2$*C*H$_2$), 28.9 (*C*H$_2$), 26.8 (*C*H$_2$), 25.8 (*C*H$_2$CH$_2$CONH),

23.0 (NHCH$_2$CH$_2$); **^{19}F NMR** (282 MHz, CDCl$_3$): -195.0 (s, CHF); **IR** (thin film) v_{max} = 3397, 3295, 2927, 2854, 2094, 1765, 1732, 1644, 1523, 1278, 1067 cm^{-1}; **HRMS (ESI)**: calculated for C$_{25}$H$_{36}$FN$_5$NaO$_8$ [M+Na]$^+$: 576.2440, found: 576.2444.

6.1.3.94 Synthesis of 4,5-dimethoxy-2-nitrobenzyl 6-decanamido-2-fluoro-3-oxohexanoate (424)
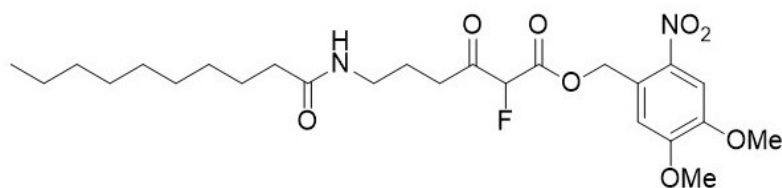
To a solution of 4,5-dimethoxy-2-nitrobenzyl 6-decanamido-3-oxohexanoate (**423**) (50 mg, 0.10 mmol) and cyclopentadienyltitanium trichloride (1 mg, 0.01 mmol) in anhydrous MeCN (1 mL) under argon at RT was added Selectfluor (38 mg, 0.11 mmol). After 16 h stirring, the reaction was concentrated, dilute with H$_2$O (4 mL) and extracted with EtOAc (4 x 4 mL). The combined organic extracts were washed with brine (10 mL) dried over MgSO$_{4(s)}$, filtered and concentrated to the crude product. Purification by silica chromatography with a gradient of 0-40% EtOAc in petroleum ether afforded the title compound (**424**) as a white solid (26 mg, 50%): **TLC** (Rf = 0.38, 1:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.75 (1H, s, NO$_2$CCH), 7.14 (1H, s, CO$_2$CH$_2$CCH), 5.83 (1H, d, J = 14.8, CO$_2$CH_2), 5.56 (1H, d, J = 14.8, CO$_2$CH_2), 5.53 (1H, br s, NH), 5.37 (1H, d, J = 48.9, CHF), 4.05 (3H, s, H_3COCCHCNO$_2$), 3.97 (3H, s, H_3COCCHC), 3.27 (2H, qq, J = 13.6, 6.62, NHCH_2), 2.86-2.68 (2H, m, CH_2COCH$_2$CO$_2$), 2.14 (2H, t, J = 8.0, CH_2CONH), 1.83 (2H, tq, J = 14.0, 6.9, NHCH$_2$CH_2), 1.64-1.57 (2H, m, CH_2CONH), 1.33-1.21 (12H, m, CH_2), 0.87 (3H, t, J = 6.9, CH$_2$CH_3); **^{13}C NMR** (125 MHz, CDCl$_3$): 201.3 (d, J = 23.0, CH$_2$$COCH_2$), 173.7 ($C$ONH), 163.6 (d, J = 23.7, CO$_2$), 154.1 (NO$_2$CCHC), 148.7 (OCH$_2$CCCOCH$_3$), 139.7 (OCH$_2$$C$CCOCH$_3$), 125.9 (NO$_2$$C$), 110.6 (OCH$_2CCCOCH_3$), 108.4 (NO$_2CCH$), 91.4 (d, J = 198.8, CHF), 65.2 (CO$_2$$CH_2$), 56.9 ($H_3$$C$OCCHC), 56.6 ($H_3$$C$OCCHCNO$_2$), 38.4 (NH$CH_2$), 36.9 ($CH_2$CONH), 35.9 ($CH_2$COCHF), 32.0 ($CH_2$), 29.6 ($CH_2$), 29.5 ($CH_2$),

29.5 (CH_2), 29.4 (CH_2), 25.9 (CH_2CH_2CONH), 23.0 ($NHCH_2CH_2$), 22.8 (CH_2), 14. (CH_3CH_2); 19F NMR: -113.2 (CFH); **IR** (thin film) v_{max} = 3301, 2924, 2853, 1765, 1733, 1644, 1522, 1276, 1221, 1066 cm^{-1}; **HRMS (ESI)**: calculated for $C_{25}H_{37}FN_2NaO_8$ [M+Na]$^+$: 535.2426, found: 535.2424.
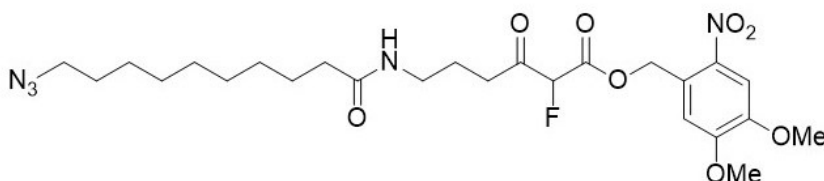
6.1.3.95 Synthesis of 4,5-dimethoxy-2-nitrobenzyl 6-(3-octylureido)-3-oxohexanoate (429)
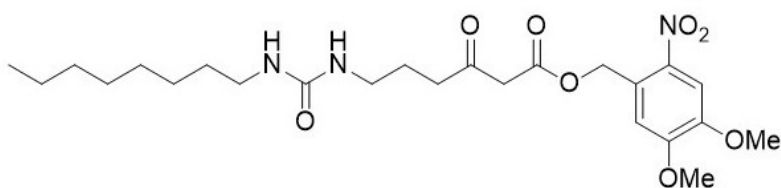
To a solution of 4-(3-octylureido)butanoic acid (**377**) (0.08 g, 0.30 mmol) in anhydrous THF (5 mL) under argon at RT was added CDI (0.06 g, 0.36 mmol), followed by stirring for 2 h. Potassium 3-((4,5-dimethoxy-2-nitrobenzyl)oxy)-3-oxopropanoate (**422**) (0.15 g, 0.45 mmol) and MgCl$_2$ (0.03 g, 0.30 mmol) were added to the reaction followed by stirring for 16 h. The reaction was concentrated, partitioned between H$_2$O (10 mL) and EtOAc (10 mL) and the aqueous layer was acidified to pH 2 using 1M HCl. The aqueous layer was extracted using EtOAc (3 x 5 mL) and the combined organic layers were washed with both saturated NaHCO$_{3(aq.)}$ (10 mL) and brine (10 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to the crude product. Purification by silica chromatography with a gradient of 0-50% EtOAc in petroleum ether afforded the title compound (**429**) as a white solid (0.08 g, 53%): **TLC** (Rf = 0.49, EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.73 (1H, s, O$_2$NCCH), 7.14 (1H, s, O$_2$NCCCH), 5.59 (2H, s, CO$_2$CH_2), 4.35-4.23 (2H, m, NHCONH), 4.05 (3H, s, O$_2$NCCHCOCH_3), 3.96 (3H, s, CH$_2$CCHCOCH_3), 3.60 (2H, s, COCH_2CO$_2$), 3.21 (2H, q, J = 6.5, NHCH_2CH$_2$CH$_2$CO), 3.13 (2H, q, J = 6.6, CH$_2$NHCONH), 2.63 (2H, t, J = 6.8, CH_2CH$_2$CO), 1.81 (2H, quin, J = 6.8, CH_2CH$_2$CO), 1.51-1.44 (2H, m, CH_2CH$_2$NHCONH), 1.33-1.22 (10H, m, CH_2), 0.87 (3H, t, J = 6.8, CH_3CH$_2$); **^{13}C NMR** (125 MHz, CDCl$_3$): 203.0 (CO), 166.8 (CO$_2$), 158.3 (NHCONH), 154.0 (O$_2$NCCHCOCH$_3$), 148.5 (CH$_2$CCHCOCH$_3$), 139.9 (CO$_2$CH$_2$$C$), 126.9 (O$_2NC$),

110.8 (CO$_2$CH$_2$CH), 108.3 (O$_2$NCCH), 64.3 (CO$_2$CH$_2$), 57.0 (O$_2$NCCHCOCH$_3$), 56.6 (CH$_2$CCHCOCH$_3$), 49.3 (COCH$_2$CO$_2$), 40.9 (CH$_2$NHCONH), 40.6 (CH$_2$COCH$_2$CO$_2$), 39.6 (NHCH$_2$CH$_2$CH$_2$CO), 32.0 (CH$_2$), 30.3 (CH$_2$), 29.4 (CH$_2$CH$_2$NHCONH), 29.3 (CH$_2$), 27.0 (CH$_2$), 24.3 (CH$_2$CH$_2$CO), (CH$_2$), 22.8 (CH$_2$), 14.3 (CH$_3$CH$_2$); **IR** (thin film) υ_{max} = 3341, 2925, 2854, 1745, 1710, 1622, 1580, 1278, 1223, 1069 cm^{-1}; **HRMS (ESI)**: calculated for C$_{24}$H$_{37}$N$_3$NaO$_8$ [M+Na]$^+$: 518.2473, found: 518.2475.

6.1.3.96 Synthesis of 4,5-dimethoxy-2-nitrobenzyl 6-(3-octylthioureido)-3-oxohexanoate (430)

To a solution of 4-(3-octylthioureido)butanoic acid (**381**) (0.08 g, 0.30 mmol) in anhydrous THF (5 mL) under argon at RT was added CDI (0.06 g, 0.36 mmol), followed by stirring for 2 h. Potassium 3-((4,5-dimethoxy-2-nitrobenzyl)oxy)-3-oxopropanoate (**422**) (0.15 g, 0.45 mmol) and MgCl$_2$ (0.03 g, 0.30 mmol) were added to the reaction followed by stirring for 16 h. The reaction was concentrated, partitioned between H$_2$O (10 mL) and EtOAc (10 mL) and the aqueous layer was acidified to pH 2 using 1M HCl. The aqueous layer was extracted using EtOAc (3 x 5 mL) and the combined organic layers were washed with both saturated NaHCO$_{3(aq.)}$ (10 mL) and brine (10 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to the crude product. Purification by silica chromatography with a gradient of 20-80% EtOAc in petroleum ether afforded the title compound (**430**) as a white solid (0.07 g, 44%): **TLC** (Rf = 0.27, 1:1 petroleum ether : EtOAc); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.73 (1H, s, O$_2$NCCH), 7.09 (1H, s, O$_2$NCCCH), 5.94 (1H, br s, CH$_2$NHCS), 5.86 (1H, NHCH$_2$CH$_2$CH$_2$CO), 5.57 (2H, s, CO$_2$CH$_2$), 4.04 (3H, s, O$_2$NCCHCOCH$_3$), 3.97 (3H, s, CH$_2$CCHCOCH$_3$), 3.61 (2H, s, COCH$_2$CO$_2$), 3.58-3.45 (2H, m, NHCH$_2$), 3.40-3.19 (2H, m, CH$_2$NHCSNH), 2.71 (2H, t, J = 6.4, CH$_2$COCH$_2$CO$_2$), 1.93 (2H, q, J = 6.5, CH$_2$CH$_2$CO), 1.63-1.56 (2H, m, CH$_2$CH$_2$NHCS),

1.38-1.19 (10H, m, C*H*$_2$); ^{13}**C NMR** (125 MHz, CDCl$_3$): 203.3 (*C*OCH$_2$CO$_2$), 193.1 (*C*S), 166.7 (*C*O$_2$), 153.9 (O$_2$NCC*C*HCOCH$_3$), 148.7 (CH$_2$C*C*HCOCH$_3$), 140.1 (CO$_2$CH$_2$*C*), 126.4 (O$_2$N*C*), 111.1 (CO$_2$CH$_2$*C*H), 108.4 (O$_2$NC*C*H), 64.5 (CO$_2$*C*H$_2$), 56.9 (O$_2$NCCHCO*C*H$_3$), 56.6 (CH$_2$CCHCO*C*H$_3$), 49.3 (CO*C*H$_2$CO$_2$), 44.1 (*C*H$_2$NHCS), 42.6 (*C*H$_2$CH$_2$CH$_2$CO$_2$), 40.5 (*C*H$_2$COCH$_2$CO$_2$), 31.9 (*C*H$_2$), 29.4 (*C*H$_2$), 29.3 (*C*H$_2$), 29.0 (*C*H$_2$CH$_2$NHCS), 27.0 (*C*H$_2$), 22.8 (*C*H$_2$), 22.7 (*C*H$_2$CH$_2$CO), 14.2 (*C*H$_3$CH$_2$); **IR** (thin film) v_{max} = 3367, 2925, 2854, 1747, 1714, 1522, 1277, 1067 cm^{-1}; **HRMS (ESI)**: calculated for C$_{24}$H$_{36}$N$_3$O$_7$S [M-H]$^-$: 510.2279, found: 510.2273.

6.1.3.97 Synthesis of 4,5-dimethoxy-2-nitrobenzyl 6-decanethioamido-3-oxohexanoate (431)

To a solution of 4-decanethioamidobutanoic acid (**391**) (0.08 g, 0.30 mmol) in anhydrous THF (5 mL) under argon at RT was added CDI (0.06 g, 0.36 mmol), followed by stirring for 2 h. Potassium 3-((4,5-dimethoxy-2-nitrobenzyl)oxy)-3-oxopropanoate (**422**) (0.15 g, 0.45 mmol) and MgCl$_2$ (0.03 g, 0.30 mmol) were added to the reaction followed by stirring for 16 h. The reaction was concentrated, partitioned between H$_2$O (10 mL) and EtOAc (10 mL) and the aqueous layer was acidified to pH 2 using 1M HCl. The aqueous layer was extracted using EtOAc (3 x 5 mL) and the combined organic layers were washed with both saturated NaHCO$_{3(aq.)}$ (10 mL) and brine (10 mL), dried over MgSO$_{4(s)}$, filtered and concentrated to the crude product. Purification by silica chromatography with a gradient of 0-50% EtOAc in petroleum ether afforded the title compound (**431**) as a white solid (0.09 g, 62%): **TLC** (Rf = 0.54, 1:1 petroleum ether : EtOAc); ^{1}**H NMR** (500 MHz, CDCl$_3$): δ 7.73 (1H, s, O$_2$NCC*H*), 7.53 (br s, N*H*), 7.10 (1H, s, O$_2$NCCC*H*), 5.58 (2H, s, CO$_2$C*H*$_2$), 4.05 (3H, s, O$_2$NCCHCOC*H*$_3$), 3.97 (3H, s, CH$_2$CCHCOC*H*$_3$), 3.70 (2H, q, *J* = 6.5, NHC*H*$_2$), 3.60 (2H, s, COC*H*$_2$CO$_2$), 2.71 (2H, t, *J* = 6.6, C*H*$_2$CSNH),

2.65-2.58 (2H, m, CH_2CO$_2$), 1.98 (2H, q, J = 6.7, NHCH$_2$CH_2), 1.79-1.69 (2H, m, CH_2CH$_2$CSNH), 1.34-1.19 (12H, m, CH_2), 0.87 (3H, t, J = 6.9, CH_3CH$_2$); **^{13}C NMR** (125 MHz, CDCl$_3$): 206.5 (CSNH), 202.9 (CO), 166.7 (CO$_2$), 153.9 (O$_2$NCCHCOCH$_3$), 148.6 (CH$_2$CCHCOCH$_3$), 140.0 (CO$_2$CH$_2$$C$), 126.5 (O$_2NC$), 111.0 (CO$_2CH_2$$C$H), 108.4 (O$_2NCC$H), 64.5 (CO$_2$$CH_2$), 56.9 (O$_2$NCCHCO$CH_3$), 56.6 (CH$_2$CCHCO$CH_3$), 49.2 (CO$CH_2CO_2$), 47.5 (CH$_2COCH_2CO_2$), 45.3 (NH$CH_2$), 40.9 ($CH_2$CSNH), 32.0 ($CH_2$), 29.6 ($CH_2$), 29.5 ($CH_2$), 29.5 ($CH_2$), 29.4 ($CH_2CH_2$CSNH), 29.1 ($CH_2$), 22.8 ($CH_2$), 21.8 (NHCH$_2$$CH_2$), 14.2 ($CH_3CH_2$); **IR** (thin film) υ_{max} = 3337, 2923, 2852, 1747, 1715, 1522, 1276, 1102 cm^{-1}; **HRMS (ESI)**: calculated for C$_{25}$H$_{38}$N$_2$NaO$_7$S [M+Na]$^+$: 533.2292, found: 533.2289.

6.1.3.98 Synthesis of 4,5-dimethoxy-2-nitrobenzyl 6-(10-(4-((3-chloro-4-(ethylamino)-2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)methyl)-1H-1,2,3-triazol-1-yl)decanamido)-3-oxohexanoate (433)

A solution of 3-chloro-4-(ethylamino)-1-(prop-2-yn-1-yl)-1H-pyrrole-2,5-dione **(432)** (kindly provided by Jon Husband, Prof Rachel O'Reilly's group) (5 mg, 0.20 mmol), 4,5-dimethoxy-2-nitrobenzyl 6-(10-azidodecanamido)-3-oxohexanoate **(426)** (15 mg, 0.03 mmol) and copper (I) iodide (2 mg, 0.01 mmol) in anhydrous THF (0.5 mL) was stirred at RT under argon for 72 h. The reaction was then concentrated and purified by silica chromatography with a gradient of 0-10% MeOH in DCM to afford the title compound **(433)** as a yellow solid (16 mg, 89%): **TLC** (Rf = 0.27, 95:5 DCM : MeOH); **^{1}H NMR** (500 MHz, CDCl$_3$): δ 7.73 (1H, s, NO$_2$CCH), 7.52 (1H, s, N$_3$CH), 7.16 (1H, s, CO$_2$CH$_2$CCH), 5.78 (1H, br s, CONH), 5.60 (CO$_2$CH_2), 5.41 (1H, br s, CH$_3$CH$_2$NH), 4.78 (2H, s, NCH_2CN$_3$), 4.29 (2H, t, J = 7.1, N$_3$CH_2), 4.05 (3H, s, H_3COCCHCNO$_2$),

3.96 (3H, s, H_3COCCHC), 3.65 (2H, quin, J = 7.0, CH$_3$CH_2NH), 3.61 (2H, s, CH_2CO$_2$), 3.27 (2H, q, J = 7.0, CONHCH_2), 2.63 (2H, t, J = 7.0, CH_2COCH$_2$CO$_2$), 2.14 (2H, t, J = 7.6, CH_2CONH), 1.90-1.82 (2H, m, N$_3$CH$_2$CH_2), 1.82 (2H, quin, J = 6.9, CONHCH$_2$CH_2), 1.61-1.57 (2H, m, CH_2CH$_2$CONH), 1.30-1.23 (3H, m, CH_3CH$_2$NH), 1.30-1.23 (10H, m, CH_2); ^{13}**C NMR** (125 MHz, CDCl$_3$): 202.7 (COCH$_2$CO$_2$), 173.6 (CONH), 166.8 (CO$_2$), 165.0 (NCO), 164.0 (NCO), 154.0 (NO$_2$CCHC), 148.4 (OCH$_2$CCCOCH$_3$), 142.8 (CN$_3$), 140.7 (NHCCON), 139.8 (OCH$_2$$C$CCOCH$_3$), 127.0 (NO$_2$$C$), 122.7 (N$_3$$C$H), 110.6 (OCH$_2CCCOCH_3$), 108.3 (NO$_2CCH$), 64.2 (2H, s, CO$_2$$CH_2$), 57.0 (H$_3$$C$OCCHC), 56.6 (H$_3$$C$OCCHCNO$_2$), 50.5 (N$_3$$CH_2$), 49.3 (CO$CH_2CO_2$), 40.7 (CONH$CH_2CH_2CH_2$), 38.6 (CONH$CH_2$), 38.2 (CH$_3$$CH_2$NH), 36.8 ($CH_2$CONH), 33.5 (N$CH_2CN_3$), 29.1 (N$_3CH_2$$CH_2$), 28.9 ($CH_2$), 28.9 ($CH_2$), 28.7 ($CH_2$), 28.7 ($CH_2$), 26.3 ($CH_2$), 25.7 ($CH_2CH_2$CONH), 23.7 (CONHCH$_2$$CH_2$), 16.3 ($CH_3CH_2$NH); **IR** (thin film) v_{max} = 3323, 2929, 2854, 1714, 1657, 1522, 1276, 1067, 733 cm^{-1}; **HRMS (ESI)**: calculated for C$_{34}$H$_{46}$N$_7$NaO$_{10}$ [M+Na]$^+$: 770.2887, found: 770.2893.

6.2   General biology methods

6.2.1   Reagents and kits used

Unless specified, all chemicals were purchased from Sigma-Aldrich. GeneJET Gel Extraction Kit and GeneJET plasmid miniprep kit was purchased from Thermo Fisher Scientific. Q5 high-fidelity DNA polymerase was purchased from New England Biolabs (NEB). Restriction enzymes and ligases were purchased from Thermo Fisher Scientific and NEB. All were stored at -20 ºC and utilised in the supplied buffers and at suggested temperatures. All primers were purchased from Sigma-Aldrich. Electrophoresis-grade agarose was purchased from Invitrogen. GelRed for DNA detection on agarose gel under UV light was purchased from Thermo Fisher Scientific. GeneRuler 1Kb from Thermo Fisher was used to estimate DNA fragment sizes. Antibiotics were purchased from Sigma-Aldrich (UK) and stock solutions were sterilised through a 0.22 µM filter and

stored at -20 ºC. All bacterial strains and plasmids utilised in the reported studies are either property of the Tosin Group or gifts from the Leadlay group at the University of Cambridge and the Sun group at Wuhan University, unless stated otherwise. PCR was performed using an Eppendorf Mastercycler Personal. Agarose gel electrophoresis was performed a Mini-Sub Cell GT system and visualised at 365 nm by a UVP Ultraviolet (UV) transilluminator. NanoDrop Lite spectrophotometer (Thermo Scientific) was utilised to measure plasmid DNA concentration. Vector p28TEV (kan^R, P_{T7}, ori_{pBR322}, $lacI$, NTerm His$_6$-tag), a derivative of pET28T with TEV cleavage site instead of thrombin cleavage site, was used for insertion of ACP, TlmJ and TlmS from the thiolactomycin gene cluster.

6.2.2   Bacterial strains

Strain	Type
Escherichia coli	BL21(DE3) (New England Biolabs)
Escherichia coli	TOP10 (Thermo Fisher)
Escherichia coli	BAP1 (kind gift from the Leadlay group)[189]
Streptomyces olivaceus	Wildtype (kind gift from Prof Yuhui Sun)[93]
Streptomyces albus/tlm	Strain engineered to contain the *tlm* cluster (kind gift from Prof Yuhui Sun)[93]
*Streptomyces avermilitis:*Stu Δ*stuA*	Strain engineered to contain the *stu* cluster but with inactivation of PKS gene s*tuA* (kind gift from Prof Yuhui Sun)[93]
Lentzea sp.	Wildtype thiolactomycin producer (NRRL 15439), purchased from NBRC, Tosin group)
Lentzea sp. Δ*tlmA*	Thiolactomycin producer but with inactivation of gene *tlmA* (kind gift from Marie Yurkovich)[134]

Lentzea sp. ΔCy	Thiolactomycin producer but with inactivation of the *tlmB* Cy domain (kind gift from Marie Yurkovich)[134]
Lentzea sp. Δ*tlmD1*	Thiolactomycin producer but with inactivation of gene *tlmD1* (kind gift from Marie Yurkovich)[134]
Streptomyces lasaliensis	Wildtype lasalocid A producer (NRRL 3382), reported by our group)[85]
Streptomyces lasaliensis ACP12 (S970A)	Lasalocid A producer but with inactivation of module 12 ACP domain within the PKS (reported by our group)[85]
Streptomyces longisporoflavus	Wildtype tetronasin producer (kind gift from Rory Little)
Streptomyces longisporoflavus Δ*tsn11*	Tetronasin producer but with inactivation of gene *tsn11* (kind gift from Rory Little)
Streptomyces longisporoflavus Δ*tsn15*	Tetronasin producer but with inactivation of gene *tsn15* (kind gift from Rory Little)
Lentzea sp. Cy (S1832A)	Thiolactomycin producer but with point mutation in the *tlmB* Cy domain (kind gift from Leadlay group)
S. coelicolor Cy (D1824A, D1829A)	Strain engineered to contain the *tlm* cluster but with point mutations in the *tlmB* Cy domain (kind gift from Leadlay group)
S. coelicolor Cy (D2083A)	Strain engineered to contain the *tlm* cluster but with a point mutation in the *tlmB* Cy domain (kind gift from Leadlay group)
S. coelicolor PCP (S2678A)	Strain engineered to contain the *tlm* cluster but with a point mutation in the *tlmB* PCP domain (kind gift from Leadlay group)

S. lividans PCP (S2678A)	Strain engineered to contain the *tlm* cluster but with a point mutation in the *tlmB* PCP domain (kind gift from Leadlay group)
S. lividans Cy (D1824A, D1829A)	Strain engineered to contain the *tlm* cluster but with point mutations in the *tlmB* Cy domain (kind gift from Leadlay group)
S. lividans TE (52800A)	Strain engineered to contain the *tlm* cluster but with point mutations in the *tlmB* TE domain (kind gift from Leadlay group)

Table 6.1 – List of bacteria strains used in this work.

6.2.3 Primers

Construct	Forward primer	Reverse primer
pET-28TEV-*tlmB*_ACP	ATACATATGACCGCCCC GGTCG	ATTGCTCGAGTTATTCCGG CTCCGGTTCGC
pET-28TEV-*tlmJ*	GCTATCATATGAGCGGT GGTGTCGATTC	AATTCTCGAGTTAGCTGGT CGCGTCGATGG
pET-28TEV-*tlmS*	ATTTATCATATGCACCC GCAGGC	TTCTCGAGTCACTGCTTGAC CACCTTCC

Table 6.2 – List of primers used for amplifying tlmB_ACP, tlmJ and tlmS.

6.2.4 Plasmids

Plasmids were constructed by myself unless otherwise stated.

Construct	Forward primer
pET-28T-*tlmB*_ACP	pET-28 vector with the thiolactomycin cluster *tlmB* ACP domain insert
pET-28T-*tlmJ*	pET-28 vector with the thiolactomycin cluster *tlmJ* gene insert
pET-28T-*tlmS*	pET-28 vector with the thiolactomycin cluster *tlmS* gene insert
pET-28T-*tlmB*_Cy	pET-28 vector with the thiolactomycin cluster *tlmB* Cy domain insert (kind gift from Marie Yurkovich)[137]
pET-28T-*tlmD1*-RhFRED	pET-28 vector with the thiolactomycin cluster *tlmD1* fused reductase insert (kind gift from Marie Yurkovich)[137]
pET-28T-*tlmD1*	pET-28 vector with the thiolactomycin cluster *tlmD1* gene insert (kind gift from Marie Yurkovich)[137]
pET-28T-6MSA	pET-28 vector with the *P. patulum* 6MSAS gene insert (Previously reported by our group)[86]

Table 6.3 – List of plasmids used in this work.

6.2.5   Buffer/solution recipes

Buffer/Solution	Recipe
Tris-borate EDTA (TBE) buffer (10X)	108 g tris-base, 55 g boric acid, 40 mL 0.5 M EDTA, diluted to 1 L and adjusted to pH 8.3.
1.5M Tris-HCl (pH 8.8)	18.1 g tris-base, diluted to 0.1 L and adjusted to pH 8.8 with HCl.

1M Tris-HCl (pH 6.8)	12.1 g tris-base, diluted to 0.1 L and adjusted to pH 6.8 with HCl.
SDS-PAGE loading dye (4X)	2 mL 1M Tris-HCl pH 6.8, 0.8 g SDS, 4 mL glycerol, 4 mL 2-mercaptoethanol, 10 µL 10% bromophenol blue.
SDS-PAGE running buffer (10X)	30 g tris-base, 144 g glycine, 10 g SDS in 1 L and adjusted to pH 8.3.
SDS-PAGE staining solution	10 mL AcOH, 45 mL H$_2$O, 45 mL MeOH, 0.25 g Coomassie brilliant blue R250.
SDS-PAGE destaining solution	10 mL AcOH, 45 mL H$_2$O, 45 mL MeOH.
Protein storage buffer	(10 mM imidazole 20 mM tris-base, 100mM NaCl, 10% (v/v) glycerol)
Protein lysis buffer (20 mM Tris-base, 100 mM NaCl, 20 mM imidazole, 10% (v/v) glycerol)	2.42 g tris-base, 5.84 g NaCl, 1.36 g imidazole, 100 mL glycerol diluted to 1 L with distilled H$_2$O and adjusted to pH 7.5 with HCl.
Protein elution buffer 1 (50 mM imidazole 20 mM tris-base, 100 mM NaCl, 10% (v/v) glycerol)	2.42 g tris-base, 5.84 g NaCl, 3.40 g imidazole, 100 mL glycerol diluted to 1 L with distilled H$_2$O and adjusted to pH 7.5 with HCl.
Protein elution buffer 2 (300 mM imidazole 20 mM tris-base, 100mM NaCl)	2.42 g tris-base, 5.84 g NaCl, 20.4 g imidazole, 100 mL glycerol diluted to 1 L with distilled H$_2$O and adjusted to pH 7.5 with HCl.

Table 6.4 – List of buffers and solutions used in this work.

6.2.6 Isolation of *Lentzea* sp. genomic DNA (gDNA)

A 4 mL culture of *Lentzea* sp. in *Lentzea* inoculum media was cultured for 3 days, then centrifuged and the supernatant discarded. The cells were resuspended in 450 µL of Sol.1 (10% sucrose, 50 mM Tris-HCl (pH 8.0), 10 mM EDTA). 50 µL of 30 mg/mL lysozyme was added, vigorously mixed and incubated at 37 °C for 1 h with shaking. 10 µL of Proteinase K (5 mg/mL) and 150 µL of 3.3% SDS were added; the solution was gently inverted a few times before incubation at 37 °C for 1 h without shaking until fully lysed. 500 µL of PCI (Phenol:Chloroform:Isoamyl alcohol 25:24:1) was added followed by vigorous shaking and vortexing. The emulsion was centrifuged at 11,000 rpm for 30 s, the upper layer was decanted and 700 µL IPA added to it. After gentle inversion several times and centrifuged at 11,000 rpm for 30 s, the supernatant was discarded to give gDNA pellet 1 mL of 70% EtOH was added. 500 µL of 1X TE buffer containing RNase A (final 0.1 mg/mL) was added to the DNA pellet followed by incubation at 37 °C for 2 h to allow complete dissolution. The DNA was warmed to 50 °C and gently mixed, 1 mL of 70% EtOH was added, mixed well and the supernatant discarded. 500 µL of 0.2X TE buffer was added to the DNA pellet followed by incubation at 37 °C for 2 h before storage at -20 °C.

6.2.7 Polymerase chain reaction and amplified DNA purification

For all PCR reactions, the master mix was composed of 20 µL of 5X Q5 reaction buffer, 20 µL 5X Q5 GC enhancer, 10 µL of 2 mM dNTPs, 5 µL forward primer (10 µM), 5 µL reverse primer (10 µM), 1 µL template gDNA (from *Lentzea* sp.), 39 µL distilled H_2O and 1 µL Q5 polymerase. 2 x 50 µL aliquot reactions were prepared from each master mix. PCR reactions were carried out using the below parameters (**Table 6.5**) and then held at 4 °C after completion. PCR products were analysed by agarose gel electrophoresis in a tank filled with TBE buffer on 0.5% agarose in TBE buffer gels using GelRed for DNA detection under UV light.

Gene	Initial temp/time	Denature temp/time	Anneal temp/time	Extension temp/time	Extension cycles	Final extension temp/time
tlmJ	98 °C/30 s	98 °C/10 s	69 °C/15 s	72 °C/33 s	30	72 °C/ 120 s
tlmS	98 °C/30 s	98 °C/10 s	69 °C/15 s	72 °C/33 s	30	72 °C/ 120s
*tlmB_* ACP	98 °C/30 s	98 °C/10 s	71 °C/15 s	72 °C/11 s	30	72 °C/ 120s

Table 6.5 – List of PCR conditions used in gene amplification

Amplified DNA fragments were purified after gel electrophoresis using GeneJET gel extraction kit according to the manufacturer's instructions. The concentration of the purified DNA solution was determined using a Nanodrop Lite spectrophotometer.

6.2.8 Double digestion and ligation of DNA into p28-TEV vector

Double-digest reactions were set up with restriction endonuclease enzymes (Xho1, Nde1) from Thermo Fisher Scientific (0.5 μL of each enzyme), 5 μL of 10X digest buffer and 500 ng DNA insert diluted with distilled H_2O to a 50 μL total volume and incubated at 37 °C for 1 h. After completion, reactions were purified using a GeneJET PCR purification kit according the to manufacturer's instructions the give digested DNA, for which an approximate concentration was measure using a Nanodrop Lite spectrophotometer (Thermo Scientific). Ligation of digested inserts into digested vector was achieved by reaction of, 2.5 μL 4X T4 Anza ligase and 50 ng digested vector with a 3X stoichiometric excess of insert DNA diluted to a total volume of 10 μL with distilled H_2O, reactions were incubated at RT for 30 mins. After completion, the reactions were

transformed into *E. coli* TOP10. Correct constructs were confirmed by formation of single colonies which were subsequently purified using the Thermo Fisher Scientific miniprep kit and then sent for sequencing.

6.2.9   Expression and purification of TlmD1, TlmD1-RhFRED, Cy and *tlmB*-ACP domains

A single colony of *E. coli* BL21(DE3) transformed with the desired plasmid (**Table 6.5**) was selected and incubated with 5 mL LB media containing 100 μg/mL of ampicillin or 50 μg/mL of kanamycin overnight at 37 ℃ with agitation. The resulting mixture was added to 500 mL LB media containing 100 μg/mL of ampicillin or 50 μg/mL of kanamycin and shaken at 37 ℃ until its optical density at 600 nm reached 0.6. The gene expression was induced by addition of 0.5 mM isopropyl-β-D-33 thiogalactoside (IPTG) and cells were incubated overnight at 15 °C. Following protein production, cells were harvested by centrifugation (5000 rpm, 20 mins, 4 ℃) and resuspended in 5 mL lysis buffer. Cells containing overproduced proteins were lysed by sonication and the resulting lysate was centrifuged (18000 rpm, 30 mins, 4 ℃). The resulting supernatant was loaded onto HisTrap HP 1 mL column pre-loaded with 100 mM $NiSO_4$. A gradient elution with buffer containing 20 mM imidazole was applied over a volume of 15 mL to wash off unbound proteins. Subsequently, a buffer containing 50 mM imidazole was applied over a volume of 5 mL to wash off loosely bound proteins and finally to elute His6-tagged recombinant protein, a buffer containing 300 mM imidazole was applied in a total volume of 15 mL. Presence of the protein was confirmed by SDS-PAGE and protein fractions were pooled and concentrated using a Vivaspin Centrifugal Concentrator. The concentrated protein was flash frozen in liquid nitrogen and stored at -80 ℃. SDS-PAGE analysis of the TlmD1, 1-RhFRED, Cy and *tlmB*-ACP purified proteins showed discrete bands at approximately 44 kDa, 82 kDa, 57 kDa and 14 kDa respectively.

6.2.10 Expression and purification of Sfp, PanK, DPCK, PPAT, *act* ACP and *act* KS/CLF

Sfp, PanK, DPCK, PPAT expression constructs for these enzymes were available in the group from previous related work.[74] The plasmid encoding for *act apo*-ACP was a kind gift of Dr Hui Hong (Cambridge); the plasmid encoding for *act* KS/CLF was obtained from *S. coelicolor* CH999 (kind gift of Dr John Crosby, Bristol). The enzymes were expressed and purified by fellow PhD student Panward Prasongpholchai for use in assays *via* a previously described method reported in the PhD thesis of Dr Samantha Kilgour (previous student in the Tosin's group).[160]

6.2.11 Protein analysis by SDS-PAGE

During purification, a sample was kept for Sodium Dodecyl Sulphate-Protein Agarose Gel Electrophoresis (SDS-PAGE) analysis (**Table 6.6**).

	10% (5 mL)	12% (5 mL)	15% (5 mL)	18% (5 mL)
H₂O	1.9 mL	1.6 mL	1.1 mL	0.6 mL
30% Acrylamide mix	1.7 mL	2 mL	2.5 mL	3 mL
1.5 M Tris (pH 8.8)	1.3 mL	1.3 mL	1.3 mL	1.3 mL
10% SDS	50 μL	50 μL	50 μL	50 μL
10% APS	50 μL	50 μL	50 μL	50 μL
Tetramethylethylenediamine (TEMED)	2 μL	2 μL	2 μL	2 μL

Table 6.6 – Recipe for SDS page gels used for protein analysis.

A PageRuler^{TM} Plus Prestained Protein Ladder from Thermo Fisher was run alongside all samples. SDS-PAGE loading dye was added to each sample. A 1X SDS PAGE running buffer was used to run the gel. The gels were loaded with the samples and run at 200V. The gels were stained with Coomassie blue stain and destained with destaining solution before viewing with white light.

6.3   Bacteria feeding experiments with chemical probes

6.3.1   Media recipes

All media and glassware were sterilized prior to use by autoclave (Astell).

Media	Recipe
TSB	30 g tryptic soy broth powder in 1 L distilled H_2O.
TSB solid media (TSA)	30 g tryptic soy broth powder, 15 g agar in 1 L distilled H_2O.
Lentzea inoculum	30 g glucose, 5 g peptone, 5 g meat/beef extract, 3 g NaCl, 2 g yeast extract, 2 g $CaCO_3$ in 1 L tap H_2O, adjusted pH to 7.0.
Maltose Bennett's solid media	1 g yeast extract, 1 g beef extract, 2 g NZ Amine, 10 g maltose, 20 g agar in 1 L distilled H_2O, adjusted pH to 7.3.
Oatmeal solid media	20 g oatflour (blended oats), 20 g agar in 1 L distilled H_2O.
ISP2 solid media	4.0 g yeast extract, 10.0 g malt extract, 4.0 g dextrose, 20 g agar in 1 L distilled H_2O, adjusted pH to 7.2
LB media	10 g tryptone, 5 g yeast extract, 10 g NaCl in 1 L distilled H_2O.
LB solid media (LBA)	10 g tryptone, 5 g yeast extract, 15 g agar, 10 g NaCl in 1 L distilled H_2O.
M79 media	10 g glucose, 10 g peptone, 2 g yeast extract, 6 g NaCl, 5 g casein hydrolysate in 1 L tap H_2O, adjusted pH to 7.1.

MYM media	4 g maltose, 4 g yeast extract, 10 g malt extract in 1 L distilled H_2O, adjusted to pH 7.1.
MYM solid media	4 g maltose, 4 g yeast extract, 10 g malt extract, 15 g agar in 1 L distilled H_2O, adjusted to pH 7.1.
Tsn medium-A	30 g tryptic soy broth powder, 5 g yeast extract, 3 g $CaCO_3$ in 1 L distilled H_2O.
Tsn medium-B	30 g tryptic soy broth powder, 5 g yeast extract, 3 g $CaCO_3$, 100 g dextrin in 1 L distilled H_2O.
Tsn medium-B agar	30 g tryptic soy broth powder, 5 g yeast extract, 3 g $CaCO_3$, 100 g dextrin, 20 g agar in 1 L distilled H_2O.

Table 6.7 – List of media used for bacterial growth in this work.

6.3.2   *In vivo* feeding experiment procedures with methyl ester probes

Unless stated otherwise, liquid cultures were incubated at 30ºC at 180 rpm (Innova 44 incubator/shaker, New Brunswick Scientific) and solid cultures were incubated on plate at 30ºC an oven.

6.3.2.1   Feeding of chemical probes (361-365) to *Lentzea* sp. strains

Stocks of *Lentzea* sp. wildtype and mutant strains (50 µL) were grown in 10 mL *Lentzea* inoculum media for 3 days. 100 µL of these precultures were used to inoculate Maltose Bennetts agar plates (5 mL) containing the respective chemical probe (2mM final concentration, added as MeOH solution of up to 100 µL to melted agar). All feedings and control experiments were set up in duplicates/triplicate, control experiments were either with no probe or without bacteria. After 5 days of incubation at 30 °C the agar was cut

into small pieces and extracted with 10 mL EtOAc, which was removed *in vacuo* and the residue was dissolved in 1 mL HPLC grade MeOH for mass spectrometry analysis.

6.3.2.2   Feeding of chemical probes (104, 160 and 161) to *S. lividans* and *S. coelicolor*

Stocks of *S. lividans* TK24 and *S. coelicolor* M1154 strains heterologously expressing the *tlm* cluster with point mutations were grown in 10 mL TSB media for 3 days. 100 µL of this preculture was used to inoculate TSA plates (5 mL) containing the respective chemical probe (2 mM final concentration, added as MeOH solution of up to 100 µL to melted agar). All feedings and control experiments were set up in duplicates/triplicate; control experiments were either with no probe or without bacteria. After 5 days of incubation at 30 °C the agar was reduced to small pieces and extracted with 10 mL EtOAc, which was removed *in vacuo* and the residue was dissolved in 1 mL HPLC grade MeOH for mass spectrometry analysis.

6.3.2.3   Feeding of chemical probes (104, 362 and 363) to *S. lasaliensis* strains (solid)

Stocks of *S. lasaliensis* strains were grown in 10 mL M79 media for 3 days. 100 µL of this preculture was used to inoculate MYM plates (5 mL) containing the respective chemical probe (2 mM final concentration, added as MeOH solution of up to 100 µL to melted agar). All feedings and control experiments were set up in duplicates/triplicate, control experiments were either with no probe or without bacteria. After 5 days of incubation at 30 °C the agar was reduced to small pieces and extracted with 10 mL EtOAc, which was removed *in vacuo* and the residue was dissolved in 1 mL HPLC grade MeOH for mass spectrometry analysis.

6.3.2.4 Feeding of chemical probes (104, 362 and 363) to *S. lasaliensis* strains (liquid)

Stocks of *S. lasaliensis* strains were grown in 10 mL M79 media for 3 days, 100 µL of this preculture was used to inoculate MYM liquid cultures in 50 mL Erlenmeyer flasks with springs (10 mL) which were incubated for 1 day. Respective chemical probes dissolved in 80 µL MeOH (2 mM final concentration) were added daily in 20 µL aliquots. All feedings and control experiments were set up in duplicates/triplicate, control experiments were given HPLC grade MeOH lacking any dissolved probe. After 4 days of incubation, at 30 °C the liquid cultures were extracted with 20 mL EtOAc (vortexed periodically over a 2h period), the organic extracts were concentrated *in vacuo* and the residue was dissolved in 1 mL HPLC grade MeOH for mass spectrometry analysis.

6.3.2.5 Feeding of chemical probes (104 and 160) to *S. longisporoflavus* strains (plate)

Stocks of *S. longisporoflavus* strains were grown in 10 mL Tsn medium-A media for 3 days, 100 µL of this preculture was used to inoculate Tsn medium-A agar plates (5 mL) containing the respective chemical probe (2mM final concentration, added as MeOH solution of up to 100 µL to melted agar). All feedings and control experiments were set up in duplicates/triplicate, control experiments were either with no probe or without bacteria. After 5 days of incubation at 30 °C the agar was reduced to small pieces and extracted with 10 mL EtOAc, which was removed *in vacuo* and the residue was dissolved in 1 mL HPLC grade MeOH for mass spectrometry analysis.

6.3.2.6 Feeding of chemical probes (104 and 160) to *S. longisporoflavus* strains (liquid)

Stocks of *S. longisporoflavus* strains were grown in 10 mL Tsn medium-A for 3 days, 100 µL of this preculture was used to inoculate Tsn medium-B cultures in 50 mL

Erlenmeyer flasks with springs (10 mL) which were incubated for 1 day. Respective chemical probes dissolved in 80 µL MeOH (2 mM final concentration) were added daily in 20 µL aliquots. All feedings and control experiments were set up in duplicates/triplicate, control experiments were given HPLC grade MeOH lacking any dissolved probe. After 4 days of incubation, at 30 °C the liquid cultures were extracted with 20 mL EtOAc (vortexed periodically over a 2h period), the organic extracts were concentrated *in vacuo* and the residue was dissolved in 1 mL HPLC grade MeOH for mass spectrometry analysis.

6.3.2.7 Feeding of SNAc substrates (192, 199, 207, 201, 216 and 259-263) to *Lentzea* sp. strains

Stocks of *Lentzea* sp. wildtype of mutant strains were grown in 10 mL *Lentzea* inoculum media for 3 days; 100 µL of this preculture was used to inoculate Maltose Bennetts agar plates (5 mL). All feedings and control experiments were set up in duplicates/triplicates. After 2 days of fermentation at 30 °C. Respective SNAc substrates in 50 µL MeOH were added in 2 µL droplets evenly around each plate (2mM final concentration) then allowed to dry for 10 mins. Control plates were distributed with HPLC grade MeOH lacking any dissolved substrate. After a further 3 days of incubation at 30 °C the agar was reduced to small pieces and extracted with 10 mL EtOAc, which was removed *in vacuo* and the residue was dissolved in 1 mL HPLC grade methanol for mass spectrometry analysis.

6.3.2.8 Feeding of SNAc substrates (201,192, 207, 262 and 263) to *S. olivaceus*

Stocks of *S. olivaceus* were grown in 10 mL *Lentzea* inoculum media for 3 days, 100 µL of this preculture was used to inoculate oatmeal agar plates (5 mL). All feedings and control experiments were set up in duplicates/triplicates. After 2 days of fermentation at 30 °C. Respective SNAc substrates in 50 µL MeOH were added in 2 µL droplets evenly around each plate (2 mM final concentration) then allowed to dry for 10 mins. Control

plates were distributed with HPLC grade MeOH lacking any dissolved substrate. After a further 3 days of incubation at 30 °C the agar was reduced to small pieces and extracted with 10 mL EtOAc, which was removed *in vacuo* and the residue was dissolved in 1 mL HPLC grade methanol for mass spectrometry analysis.

6.3.2.9 Feeding of SNAc substrates (201,192, 207, 262 and 263) to *S. avermilitis:*Stu stuA strain

Stocks of *S. avermillitis* were grown in 10 mL TSB media for 3 days, 100 µL of this preculture was used to inoculate ISP2 agar plates (5 mL). All feedings and control experiments were set up in duplicates/triplicates. After 2 days of fermentation at 30 °C. Respective SNAc substrates in 50 µL MeOH were added in 2 µL droplets evenly around each plate (2mM final concentration) then allowed to dry for 10 mins. Control plates were distributed with HPLC grade MeOH lacking any dissolved substrate. After a further 3 days of incubation at 30 °C the agar was reduced to small pieces and extracted with 10 mL EtOAc, which was removed *in vacuo* and the residue was dissolved in 1 mL HPLC grade methanol for mass spectrometry analysis.

6.3.2.10 Feeding of chemical probes (104, 361-365) to *E. coli* expressing 6-MSAS

E. coli BAP1 cells were transformed with the 6-MSAS encoding plasmid pKOS007-109[86,189] and grown in LB media (10 mL) containing 100 µg/mL carbenicillin to an A600= 0.6 at 37 ºC. IPTG (0.5 mM final concentration) and the respective chemical probe (2 mM final concentration, added as MeOH solution of up to 100 µL) were added, the cultures were incubated at 37 °C for further 24 h. All feedings and control experiments were set up in duplicates/triplicate; control experiments either lacked chemical probe or were carried out on *E. coli* strains not harbouring the 6-MSAS encoding plasmid. Each bacterial culture was extracted with 20 mL EtOAc (vortexed periodically over a 1 h

period), the organic extracts were concentrated *in vacuo* and the residue was dissolved in 1 mL HPLC grade MeOH for mass spectrometry analysis.

6.3.3 *In vivo* feeding experiment procedures with photolabile probes

Bacterial cultures were irradiated as stated below (**6.3.3.1**) unless stated otherwise, liquid cultures were incubated at 30 °C, at 180 rpm (Innova 44 incubator/shaker (New Brunswick scientific), solid cultures were incubated on plate at 30 °C.

6.3.3.1 Irradiation procedures

For *in vitro* experiments, samples were placed in a quartz cuvette for irradiation with a home built light box, equipped with a circular 22 W UVA lamp (365 nm) for 45 mins. The light box was purposely built by Rod Wesson (Electronics workshop, University of Warwick, UK). Cultures in liquid media were placed within a home built light box, equipped with a circular 22 W UVA lamp, able to be fixed to an Innova™ 4300 Incubator Shaker, New Brunswick Scientific platform for temperature and shaking.

6.3.3.2 Use of photolabile probes (417 and 423) on *S. lasaliensis* strains

Stocks of *S. lasaliensis* strains were grown in 10 mL M79 media for 3 days. 100 μL of this preculture was used to inoculate MYM liquid cultures in 50 mL Erlenmeyer flasks with springs (10 mL) which were incubated for 1 day and then respective chemical probes (2mM final concentration, added as MeOH solution of up to 100 μL) were added. Over 4 days, each flask was irradiated for 1 h per day with the light box inside of the shaker. All feedings and control experiments were set up in duplicates/triplicates; control experiments were either non-irradiated cultures and/or cultures lacking the probe. After 4 days of incubation, at 30 °C the liquid cultures were extracted with 20 mL EtOAc

(vortexed periodically over a 2 h period), the organic extracts were concentrated *in vacuo* and the residue was dissolved in 1 mL HPLC grade MeOH for mass spectrometry analysis.

6.3.3.3   Use of photolabile probes (423-424) on *S. longisporoflavus* strains

Stocks of *S. longisporoflavus* strains were grown in 10 mL Tsn medium-A for 3 days. 100 µL of this preculture was used to inoculate Tsn medium-A cultures in 50 mL Erlenmeyer flasks with springs (10 mL) which were incubated for 1 day and then respective chemical probes (2 mM final concentration, added as MeOH solution of up to 100 µL) were added. Over 4 days, each flask was irradiated for 1 h per day with the light box inside of the shaker. All feedings and control experiments were set up in duplicates/triplicates; control experiments were either non-irradiated cultures and/or cultures lacking the probe. After 4 days of incubation, at 30 °C the liquid cultures were extracted with 20 mL EtOAc (vortexed periodically over a 2 h period), the organic extracts were concentrated *in vacuo* and the residue was dissolved in 1 mL HPLC grade MeOH for mass spectrometry analysis.

6.3.4   Chemical probing of SEK4/4b formation

6.3.4.1   *In vitro* reconstitution of SEK4/4b formation by a minimal type II PKS

This procedure and all related assays were carried out by Panward Prasongpholchai (PhD student in the Tosin group); data analysis was performed by myself.

For the reconstitution of SEK4/4b in vitro, *apo-act*ACP was first converted to the malonyl *act*ACP and then combined with KS-CLF. Formation of malonyl *act*ACP was performed in 20 mM Tris pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$ with 100 µM *apo*-ActACP, 400 µM malonyl CoA and 3 µM Sfp in a total volume of 50 µL. The mixture was incubated

at 30 °C for 1 hr then immediately used for the *in vitro* reconstitution of SEK4/4b formation. This was performed in 20 mM Tris pH 7.5, 100 mM NaCl with 40 µM malonyl *act* ACP and 3 µM KS-CLF in a total volume of 100 µL. The mixture was incubated at 30 °C overnight. The reaction was quenched and extracted by addition of 150 µL HPLC grade MeOH then centrifuged at 14800 rpm for 10 mins. Supernatant was collected and stored in -20 °C for mass spectrometry analysis.

6.3.4.2   Use of photolabile probes (423, 433, 426-431) in the formation of SEK4/4b

Chemical probes (2 mM in DMSO) were photolysed as previously described for 45 mins prior to their addition to the SEK4/4b assays. The photolysed probes were added to the assays in different equivalents with respect to the concentration of malonylated-ActACP (40 µM), ranging from 10 µM (0.1 eq.), 40 µM (1 eq.), 160 µM (4 eq.), 280 µM (7 eq.) and 400 µM (10 eq.), prior to the incubation at 30 °C overnight, assay tubes were then centrifuged and sampled for orbitrap fusion LC-HRMS analysis.

6.3.5   Chemical probing of TLM formation *in vitro*

1.3.5.1 Conversion of *apo*-tlmB ACPs to *crypto*-tlmB ACP

Conversion of *apo*-TlmB ACP to *crypto*-TlmB-ACP was carried out using the synthetic pantetheine analogue (**245**). The conversion was performed in 20 mM Tris pH 7.5 with the subsequent addition of 100 mM NaCl, 10 mM MgCl$_2$, 90 µM *apo*-TlmB ACP, 300 µM synthetic pantetheine (**245**), 8 mM ATP, 5 µM PanK, 5 µM PPAT, 5 µM DPCK and 2 µM Sfp in a total volume of 60 µL. The mixture was incubated at RT for 1 h then immediately utilised for *in vitro* assays with Tlm proteins as described in **Section 1.3.5.2.** 2 µL aliquots were taken and diluted 10-fold for LCMS analysis (2 µL injection).

1.3.5.2 Enzymatic assays of recombinant Tlm proteins

Several assays were set up as summarised below (**Table 6.8**) in order to gather information on the mechanism of sulphur insertion in the formation of thiolactomycin. After formation of the *crypto*-ACP described above in **Section 1.3.5.1** which was typically done in a 60 µL total volume PCR tube, two 27.8 µL (giving a 50 µM final concentration) aliquots (yielding two separate assays) were taken and added to two PCR tubes. The respective variable enzymes/cofactors were then added to each tube followed by addition of 20 mM Tris pH 7.5, 100 mM NaCl to make up a final volume of 50 µL and incubated at RT for 2 hr then analysed by mass spectrometry as described in **6.4.4.2**. 2 µL aliquots were taken and diluted 5-fold for LCMS analysis (2 µL injection).

crypto-TlmB_ACP	TlmD1-RhFred	NADPH	Cy	DL probe (242)
50 µM	3 µM	2 mM	-	-
50 µM	20 µM	2 mM	-	-
50 µM	-	-	5 µM	250 µM
50 µM	-	-	20 µM	250 µM
50 µM	3 µM	2 mM	5 µM	-
50 µM	15 µM	2 mM	5 µM	-
50 µM	15 µM	2 mM	15 µM	-

Table 6.8 – Assay conditions used for *crypto*-ACP assays; values are final volume concentrations.

6.4 LC-HRMS^n analysis of extracts and enzymatic assays

All extracts from thiotetronate and 6MSA related experiments were analysed by LC-HRMS^n analysis on a Thermo Fisher Scientific Orbitrap Fusion (Q-OT-qIT, Thermo) instrument as described in section (**6.4.4.1** and **6.4.1.3**), unless otherwise stated. Organic extracts from *in vivo* experiments using photolabile probes on *S. lasaliensis* and *S. longisporoflavus* strains were analysed using a Bruker MaXis Impact UHR-TOF instrument **(6.4.2.1).** Analyses of ACP species from *in vitro* assays were carried out using a Bruker Amazon X instrument coupled with an Agilent 1260 HPLC **(6.4.4.2).**

6.4.1 Orbitrap Fusion analyses of small molecules

Organic extracts from *in vivo* experiments with chemical probes were analysed in a similar manner to that previously reported,[83,85] however the LC-HRMS parameters used for each type of experiment were subject to the below specifications. Three methods were used: *S. lasaliensis* and *S. longisporoflavus* extracts were analysed with a 75 min method (**6.4.1.2**); thiotetronate and SEK4/4b related samples were run using a 45 min method (**6.4.1.1);** and finally, 6-MSAS experiments used a steeper 45 min gradient (**6.4.1.3).**

6.4.1.1 Analyses of thiotetronate and SEK4/4b related experiments

Reverse phase chromatography was used to separate the mixtures prior to MS analysis. Two columns were utilised: an Acclaim PepMap μ-precolumn cartridge 300 μm i.d. x 5 mm 5 μm 100 Å and an Acclaim PepMap RSLC 75 μm x 15 cm 2 μm 100 Å (Thermo Scientific). The columns were installed on an Ultimate 3000 RSLCnano system (Dionex). Mobile phase buffer A was composed of 0.1% aqueous formic acid and mobile phase B was composed of 100% MeCN containing 0.1% formic acid. Samples were loaded onto

the μ-precolumn equilibrated in 50% aqueous MeCN containing 0.1% TFA for 8 min at 10 μL min^{-1} after which compounds were eluted onto the analytical column following a 45 min gradient for which the mobile phase B concentration was increased from 50% B to 80% over 15 min, then maintained at 80% B for 23 minutes, then decreased to 50% over 1 min, followed by a 5 min wash at 50% B. Eluting cations were converted to gas phase ions by electrospray ionization and analysed. Survey scans of precursors from 150 to 1500 m/z were performed at 60K resolution (at 200 m/z) with a 4×10^5 ion count target. Tandem MS was performed by isolation at 1.6 Th with the quadrupole, HCD fragmentation with normalized collision energy of 32, and rapid scan MS analysis in the ion trap. The MS2 ion count target was set to 2 x 10^5 and the maximum injection time was 50 ms. A filter targeted inclusion mass list was used to select the precursor ions.

6.4.1.2 Analyses of *S. lasaliensis* and *S. longisporoflavus* experiments

Reverse phase chromatography was used to separate the mixtures prior to MS analysis. Two columns were utilised: an Acclaim PepMap μ-precolumn cartridge 300 μm i.d. x 5 mm 5 μm 100 Å and an Acclaim PepMap RSLC 75 μm x 15 cm 2 μm 100 Å (Thermo Scientific). The columns were installed on an Ultimate 3000 RSLCnano system (Dionex). Mobile phase buffer A was composed of 0.1% aqueous formic acid and mobile phase B was composed of 100% MeCN containing 0.1% formic acid. Samples were loaded onto the μ-precolumn equilibrated in 50% aqueous MeCN containing 0.1% TFA for 8 min at 10 μL min^{-1} after which compounds were eluted onto the analytical column following a 75 min gradient for which the mobile phase B concentration was increased from 50% B to 80% over 15 min, then maintained at 80% B for 50 minutes, then decreased to 50% over 1 min, followed by a 9 min wash at 50% B. Eluting cations were converted to gas phase ions by electrospray ionization and analysed. Survey scans of precursors from 150 to 1500 m/z were performed at 60K resolution (at 200 m/z) with a 4×10^5 ion count target. Tandem MS was performed by isolation at 1.6 Th with the quadrupole, HCD fragmentation with normalized collision energy of 32, and rapid scan MS analysis in the

ion trap. The MS2 ion count target was set to 2 x 10^5 and the maximum injection time was 50 ms. A filter targeted inclusion mass list was used to select the precursor ions.

6.4.1.3   Analyses of 6-MSAS related experiments

Reverse phase chromatography was used to separate the mixtures prior to MS analysis. Two columns were utilised: an Acclaim PepMap μ-precolumn cartridge 300 μm i.d. x 5 mm 5 μm 100 Å and an Acclaim PepMap RSLC 75 μm x 15 cm 2 μm 100 Å (Thermo Scientific). The columns were installed on an Ultimate 3000 RSLCnano system (Dionex). Mobile phase buffer A was composed of 0.1% aqueous formic acid and mobile phase B was composed of 100% MeCN containing 0.1% formic acid. Samples were loaded onto the μ-precolumn equilibrated in 50% aqueous MeCN containing 0.1% TFA for 8 min at 10 μL min^{-1} after which compounds were eluted onto the analytical column following a 45 min gradient for which the mobile phase B concentration was increased from 50% B to 95% over 15 min, then maintained at 95% B for 23 minutes, then decreased to 50% over 1 min, followed by a 5 min wash at 50% B. Eluting cations were converted to gas phase ions by electrospray ionization and analysed. Survey scans of precursors from 150 to 1500 m/z were performed at 60K resolution (at 200 m/z) with a 4×10^5 ion count target. Tandem MS was performed by isolation at 1.6 Th with the quadrupole, HCD fragmentation with normalized collision energy of 32, and rapid scan MS analysis in the ion trap. The MS2 ion count target was set to 2 x 10^5 and the maximum injection time was 50 ms. A filter targeted inclusion mass list was used to select the precursor ions.

6.4.2   Bruker instrument setup and analysis

6.4.2.1   MaXis Impact UHR-TOF analyses of small molecules

UPLC-HR-ESI-MS analyses of extracts on a MaXis Impact UHR-TOF (Bruker Daltonics) were carried out as follows: samples (5 μL) were injected onto an Acquity

UPLC HSS T3 (150 mm x 1.0 mm, 1.8 µm) or Agilent Eclipse C18 (1.8 µm, 100 mm x 2.1 mm). The mobile phase consisted of a gradient of H_2O and MeCN (HPLC grade, each with 0.1% TFA). The following solvent (A =1% TFA in H2O, B =1% TFA in MeCN) gradient was applied: 10% B 0-2.7 min; 10-100% B 2.7-42.7 min; 100% B 42.7-52.7 min; 100-10% B 52.7-55.7 min; 10% B 55.7-67.7 min, using an Acquity UPLC HSS T3 column at a flow rate of 0.05 mL/min. Spectra were recorded in positive ionisation mode, scanning from *m/z* 0 to 2000 with the resolution set at 45K. Gathered data was internally calibrated once it has been acquired from the instrument using the Bruker Daltronics Data Analysis software. A series of peaks in the spectra were chosen, in as wide a range as possible. The software was instructed to fit a calibration curve to the peaks chosen, and this was then applied to the spectrum. Each spectrum was internally calibrated separately, with its own calibration curve, and a range spanning the whole *m/z* range required, selected ion search within 10 ppm was performed.

6.4.2.2   Small protein analyses (Bruker Amazon X)

All ACP-related species deriving from *in vitro* experiments were analysed on a Bruker Amazon X ion-trap mass spectrometer coupled with an Agilent 1260 HPLC, the column used was an Agilent Zorbax Eclipse C18, 150 x 4.6 mM, 5 µM. The parameters were set as follows: capillary 4500V, end plate off set 500V, Nebuliser pressure 40 psi, dry gas 8 L/min, dry temperature 210 ºC, mass range m/z 100-3000, ICC 200000, flowrate 1 mL/min, with 20% flow to MS. Mobile phase buffer A was composed of 0.1% aqueous formic acid and mobile phase B was composed of 100% MeCN containing 0.1% formic acid using a flowrate of 0.2 mL/min. The gradient utilised is shown below (**Table 6.9**). Gathered data were internally calibrated once it had been acquired from the instrument using the Bruker Daltronics Data Analysis software before analysis.

Time (mins)	H_2O + 0.1% FA (%)	MeCN + 0.1% FA (%)	Flow rate (mL/min)
0.0	95	5	0.2
5.0	95	5	0.2
35.0	0	100	0.2
40.0	0	100	0.2
45.0	95	5	0.2

Table 6.9 – **Solvent gradient used for ACP analyses on Amazon X instrument.**

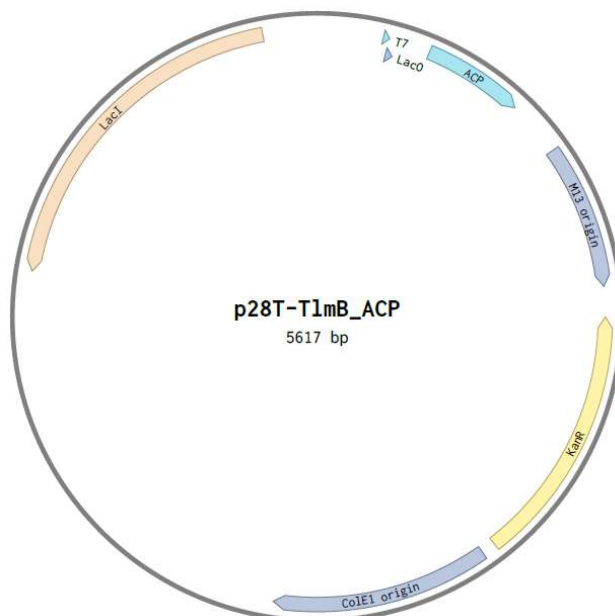# 7   Appendix

7.1   Plasmid maps

Figure 7.1 – **Plasmid map of standalone ACP from TlmB in pET28T vector.**
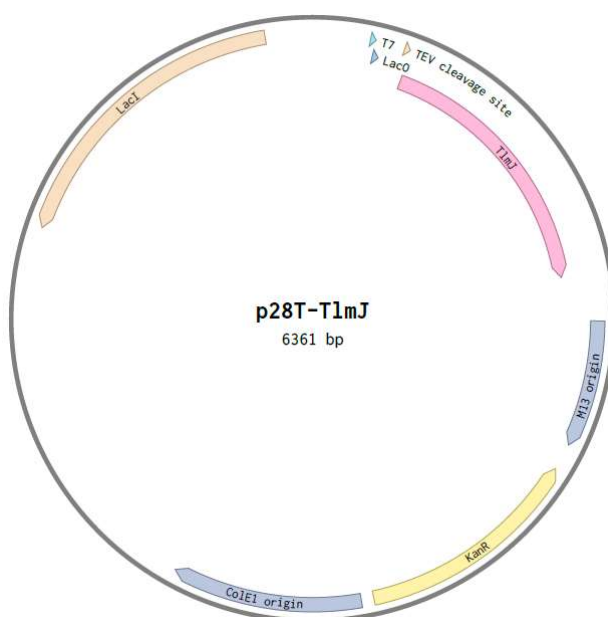

Figure 7.2 - **Plasmid map of TlmJ in pET28T vector.**
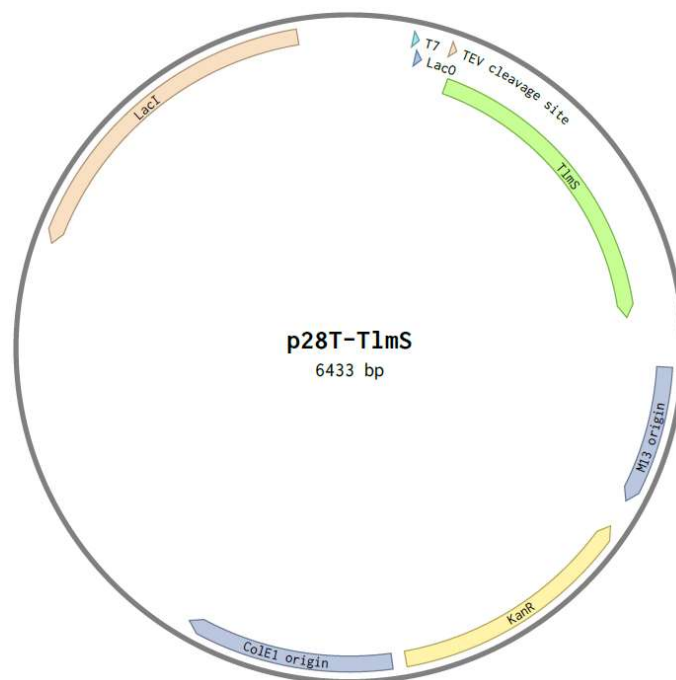
Figure 7.3 - **Plasmid map of TlmS in pET28T vector.**
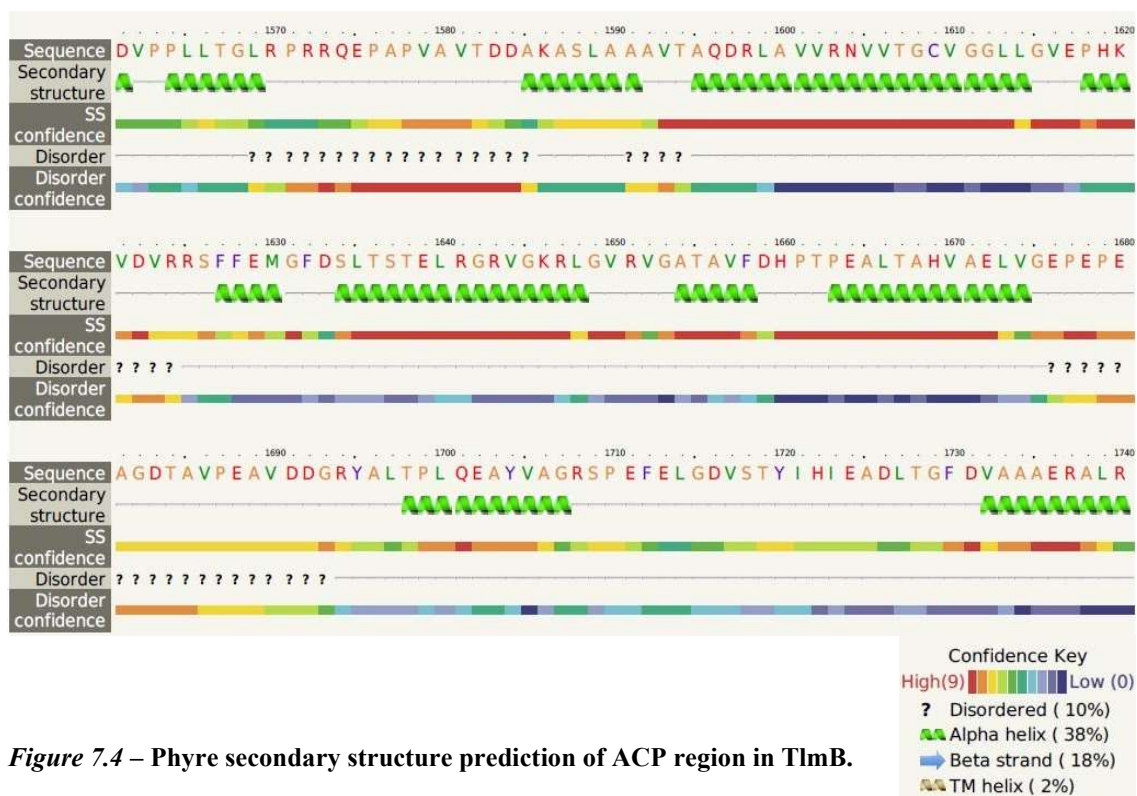
7.2 Phyre predictions

Figure 7.4 – **Phyre secondary structure prediction of ACP region in TlmB.**
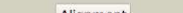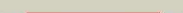
#	Template	Alignment Coverage	3D Model	Confidence	% i.d.	Template Information
1	c4l0eA_	Alignment		100.0	28	**PDB header:**oxidoreductase **Chain:** A: **PDB Molecule:**p450 monooxygenase; **PDBTitle:** structure of p450sky (cyp163b3), a cytochrome p450 from skyllamycin2 biosynthesis (heme-coordinated expression tag)
2	c2zbxA_	Alignment		100.0	34	**PDB header:**oxidoreductase **Chain:** A: **PDB Molecule:**cytochrome p450-su1; **PDBTitle:** crystal structure of vitamin d hydroxylase cytochrome p4502 105a1 (wild type) with imidazole bound
3	c2z36A_	Alignment		100.0	31	**PDB header:**oxidoreductase **Chain:** A: **PDB Molecule:**cytochrome p450 type compactin 3'',4''- **PDBTitle:** crystal structure of cytochrome p450 moxa from nonomuraea2 recticatena (cyp105)
4	c3a4hA_	Alignment		100.0	36	**PDB header:**oxidoreductase **Chain:** A: **PDB Molecule:**vitamin d hydroxylase; **PDBTitle:** structure of cytochrome p450 vdh from pseudonocardia autotrophica2 (orthorhombic crystal form)
5	d1cpta_	Alignment		100.0	24	**Fold:**Cytochrome P450 **Superfamily:**Cytochrome P450 **Family:**Cytochrome P450
6	d1lfka_	Alignment		100.0	27	**Fold:**Cytochrome P450 **Superfamily:**Cytochrome P450 **Family:**Cytochrome P450

Figure 7.5 – **Phyre top predicted structurally similar enzymes for TlmD1.**

7.3 Nucelotide sequences

```
GTGTACGCCGCCATCAGGAAGGCGGGCAACAGGCCCGTCGAGGTGGTGCTGGAGACCGGCATGGCCTACC
TGCCGCCCGGTCTGCGGGCGTGGCTGCTGACCTCGTACGACGACATCGAGTTCGCGCTGGGCGATCCCCG
GTTCCGCAAGGAGCTGGACAAGGCGATGGCGGTGTTCGGCCACGCCGGCCAGGACTCCGTGCTCTACGAC
AACATGGCGAACAACGACCCGCCTGAGCACACCGGCTGCGCAAGCCGCTCAACGGCGCGTTCACCGCGC
GGGCGGCGACGGCACGGCGTGAAGACGTGCGCGCGGTTGCACACCGAGTGCTGGACCGGGTGGCCGGCAA
GCCGGAGTTCGATCTGGTCGAGGACTTCGCGCTGCCGTTCTCGATCGAGGTGATCTGCGGGCTCCTCGGT
GTGCCCGTGGCCGATCGGCACACGTTCTCGTCGTGGGCCCAGACGATCACGAGTGCGGCGTCGCGGGACG
CAATCCAGCGCGACGCCGGGCTGATGGCCAGCTACCTGCGCGATCTCATGGCGGCCAAGCGGAACTCCGG
CGACGAGGACGTGCTGGCCCTGCTGGCCGGGCATCCCGGCGTCACCGAGCGCGAAGCCGTCTCCCAGGCC
TACGCGTTGCTGGTCGCCGGGTACGAGACGACGGCGAACCTGCTCGCGGCGGGTCTGCTCGCGCTCGACG
CCGACCCCGCTGTGCACGCACGTCTCTGGGCCGATCCGGCGCTCGTTCCCGGTGCGGTGCAGGAGATGTT
GCGCCACCAGTCGCCGTTCAACCTGTCGCTCTACCGCTACGTGAGCGAGGACGTCGAGGTCGGCGGCACG
CGGATCCCGGCGGGCTCGATCGTGTTCCTCGCGTTCGCGGCGGCGAACCGGGACGACGAGAGGTTCGCCG
```

```
CACCGGACGCGTTCGACCTCGACCGGCCCGGTGGCGACCACATCGCCTTCGGCGGCGGTGTGCACAACTG
CATCGGCAAGCACCTCGCGAGGGTCGAGGCCGAGGTCGCGTTCGAGGCGTTGATCGAACGCTGTCCCGCC
CTGGAGATCCGGACCTCCCGCGCGGACCTGGAGTGGAAACCGAGCCCCACGTTCCGCGGCGTGCGGCACC
TGCTCGTCGGACCGGGCCCGGCCGAGGTGACCCGATGA
```

*Figure 7.6 – Nucleotide sequence for *tlmD1*.*

```
ATGAGCGGTGGTGTCGATTCCGCGGTGGCGGCGGCGCGCGCCGTCGAAGCGGGACACGACGTGACGGGCG
TGCACCTGGCGTTGTCGGCGAAGCCCGGCACCCTGCGCACCGGCGCACGTGGCTGCTGCACGATCGAAGA
CGCCCACGACGCCCGCCGAGCCGCGGACCTGATGGGCATCCCGTTCTACGTCTGGGACTTCGCCGAGCGC
TTCACCGAAGAGGTCGTGGAGGACTTCGTCGCGGAGTACGCGGCCGGTCGCACGCCGAACCCGTGCCTGC
GCTGCAACGAGCGGATCAAGTTCGAGGCGTTGCTGGACAAGGCGATCGCCCTCGGTTTCGACGCCGTCTG
CACCGGCCACTACGCGCGCCTGGAGATGGTCGACGGCCGTCCCGAGCTGCGCCGTTCCGCCGACATGGGC
AAGGACCAGTCGTACGTGCTGGCCTCGCTGACGGCGGAGCAGCTCTCGCACGCGATGTTCCCGTTGGGGG
ACACGACGAAGGACGACGTGCGCGTCGAGGCGGCGTCACGCGGGCTGTCGGTGGCGAAGAAGCCGGACAG
CCACGACATCTGCTTCATCCCGGACGGCGACACGCAGAAGTTCCTGGCGGGCCGCATGGAGACCAAGCCG
GGCCTGCTGATCGACGACGCGACCGGCGCGGTGCTGGGCCGGCACGCGGGCGTGCACGAGTTCACCGTCG
GGCAGCGCAAGGGCCTCGGCATCGACGGCCCCGCCGCCGACGGCCGTCCCCGGTACGTGCTGGCGCTGGA
GCCGGTGTCAGGGAACGTGCGGGTCGGCGCGGTCGAAAAGCTCGCGGTGACCGGGATTTCCGCGTCGTCG
CCGGTCTCGCACGTCGAGTTGGACGGACCGGTCGAGTGCGTCGCGCAGGTGCGTGCGCACGGTGGCACGG
CTCCGGCCGTCGCCGAACTGGTCGACGGCGTGCTGCGGGTGCAGTTGCGGGAGCCGCTGAAGGGCGTGGC
GCCCGGTCAGGCCGTCGCGATCTACCGCGAGGACGCGGCGGGCGACATCGTGCTGGCCAGCGCGACCATC
GACGCGACCAGCTAA
```

*Figure 7.7 – Nucleotide sequence for *tlmJ*.*

```
ATGCACCCGCAGGCGGTCGCGGCCATGACCGAGGCGTTGTCCACCACGGGCAACGCCTCGTCGCTGCACA
CCTCCGGCAGGCGGGCGCGGCGGGCGATCGAAGAGGCCCGCGAGAACATCGCCGACGCCCTGGGCGCCCG
GCCGTCCGAGGTGCTGTTCACCGGCGGTGGCACGGAGAGCGACAACCTCGCGGTCAAGGGCATCTACTGG
GCGAGCGGTCGTCGCCGTGTGCTGGCATCGGCCGTCGAGCACCACGCCGTGCTGGACGCCGTCGAGTGGC
TCGAACAGCAGGGCGCCGAGGTCACCTGGCTGCCCACGGACCGCTTCGGCCGGGTGTCGCCGGAGTCGGT
CGCCGAGGCGCTGGACGACGACGTCGCGCTCGTCACCACGATGTGGGCGAACAACGAGGTCGGCACCATC
AACCCCGTGCACGACATCGCTCGGCTGTGCGCTGAGCACGGAGTGCCGTTCCACACGGACGCGGTGCAGG
CCGTGGGCGCGGTCCCGGTCGATTTCGCCGCCTCCGGCGCCTCCGCGCTCACGATGACCGGCCACAAGGT
```

```
CGGCGGTCCGTACGGCATCGGCGTGCTGCTGCTCGCACGGGACGTGAAGTGCACGCCGTTGATGCACGGC
GGCGGCCAGGAGCGTGACGTGCGGTCCGGCACGCTGGACGTGCCGTCGATCGTCGGCCTGGCGCGCGCGG
TGCGGATCTCGGTGGAGGAACAGGCTTCCCGGGCTGTGGAGATCGCGGCGCTGCGGGACGAGCTGATCGC
GTCGGTGCGCGCGGTGGTGCCGGACATGATCGTGAACGGCGACCCGGCGGACCGGCTGCCCGGCAACGCG
CACCTGACGTTCCCCGGCTGCGAGGGCGACAGCCTGCTGATGCTGTTGGACGCCAAGGGCATCGAGTGCT
CGACGGGCTCCGCGTGCACGGCCGGTGTCGCGCAGCCCTCGCACGTGCTGCTCGCGATGGACGTCGAACC
GGCGCTGGCCCGTGGATCTTTGAGGTTCTCGCTCGGGCATACGTCGACGCGGGCGGACGTTGAACAACTC
GCATCGGTCATCGGCCCGGTGGTGGAGAGGGCCCGCACGGCGGGCCTGGCAGGTATGAGGCGGTCCCGGA
AGGTGGTCAAGCAGTGA
```

Figure 7.8 – Nucleotide sequence for tlmS.

```
ATGACCGCCCCGGTCGCGGTGACCGACGACGCGAAGGCCTCGCTCGCCGCAGCGGTGACGGCGCAGGACC
GGCTCGCGGTGGTGCGGAACGTGGTGACCGGGTGTGTCGGCGGGCTGCTCGGCGTCGAACCGCACAAGGT
GGACGTGCGCCGTTCCTTCTTCGAGATGGGCTTCGACTCGCTCACGTCGACCGAACTGCGCGGCCGGGTC
GGCAAGCGGCTCGGTGTCCGGGTGGGCGCGACCGCCGTGTTCGACCACCCGACCCCGGAGGCGCTCACCG
CGCACGTCGCCGAACTCGTCGGCGAACCGGAGCCGGAATAA
```

Figure 7.9 – Nucleotide sequence for ACP region of tlmB.

7.4 Miscellaneous

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
Lentzea sp. ATTC31319, thiolactomycin biosynthetic gene cluster sequence, strain ATCC31319	2139	2139	100%	0.0	100.00%	LN879412.1
Streptomyces afghaniensis strain NRRL5621 thiotetroamide biosynthesis gene cluster, complete sequence	230	230	73%	1e-55	72.16%	KT282101.1
Streptomyces incarnatus strain NRRL 8089 sequence	224	224	71%	7e-54	72.14%	CP011497.1
Streptomyces olivaceus thiotetronate Tu3010 biosynthetic gene cluster sequence, strain Tu3010	202	202	73%	3e-47	71.54%	LN879414.1

Figure 7.10 – BLAST search results for *tlmD1* nucleotide sequence.

7.5 Publication list

1) Havemann, J.; Yurkovich, M. E.; <u>Jenkins, R</u>.; Harringer, S.; Tao, W.; Wen, S.; Sun, Y.; Leadlay, P. F.**;** Tosin, M.;* 'Chemical Probing of Thiotetronate Bio-Assembly', *Chem. Commun.* 2017, **53**, 1912-1915.

2) Yurkovich, M. E.; <u>Jenkins, R</u>.; Sun, Y.; Tosin, M**.;*** Leadlay, P. F.;* 'The Polyketide Backbone of Thiolactomycin is Assembled by an Unusual Type I Iterative Synthase', *Chem. Commun.* 2017**, 53**, 2182-2185.

3) Little, R.; Paiva, F.; <u>Jenkins, R</u>.; Hong, H.; Sun, Y.; Demydchuk, Y.; Samborskyy, M.; Tosin, M.; Leeper, F.; Dias, M.; Leadlay, P. F.;* 'Unexpected enzyme-catalysed [4+2] cycloaddition and rearrangement in polyether antibiotic biosynthesis', *Nat. Cat.*, 2019, **2**, 1045–1054.

4) Kilgour, S. L.; <u>Jenkins, R</u>.; Tosin, M.;* 'A photoactivatable small molecule probe for the *in vivo* capture of polyketide intermediates, *Chem. Eur. J.*, 2019, **25**, 16511.

8 <u>Bibliography</u>

1   C. T. Walsh and Y. Tang, in *Natural Product Biosynthesis: Chemical Logic and Enzymatic Machinery, Major Classes of Natural Product Scaffolds and Enzymatic Biosynthetic Machinery*, The Royal Society of Chemistry, 2017, pp. 7–54.

2   P. M. Dewick, in *Medicinal Natural Products, Secondary Metabolism: The Building Blocks and Construction Mechanisms*, 2009, pp. 7–38.

3   A. L. Harvey, *Drug Discov. Today*, 2008, **13**, 894–901.

4   C. L. Cantrell, F. E. Dayan and S. O. Duke, *J. Nat. Prod.*, 2012, **75**, 1231–1242.

5   R. Croteau, R. E. B. Ketchum, R. M. Long, R. Kaspera and R. Mark, *Phytochem. Rev.*, 2010, **5**, 75–97.

6   C. P. Ridley, H. Y. Lee and C. Khosla, *PNAS*, 2008, **105**, 4595–4600.

7   P. Casati, *Front. Plant Sci.*, 2012, **3**, 1–15.

8   Â. Carvalho, E. H. Hansen, O. Kayser, S. Carlsen and F. Stehle, *FEMS Yeast Res.*, 2017, **17**, 1–11.

9   W. Hüttel, J. B. Spencer and P. F. Leadlay, *Beilstein J. Org. Chem.*, 2014, **10**, 361–368.

10   D. J. Newman, G. M. Cragg and K. M. Snader, *Nat. Prod. Rep.*, 2000, **17**, 215–234.

11   D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2016, **79**, 629–661.

12   D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2007, **70**, 461–477.

13   D. A. Dias, S. Urban and U. Roessner, *Metabolites*, 2012, **2**, 303–336.

14   J. Clardy, M. A. Fischbach and C. T. Walsh, *Nat. Biotechnol.*, 2006, **24**, 1541–1550.

15   L. L. Ling, T. Schneider, A. J. Peoples, A. L. Spoering, I. Engels, B. P. Conlon, A.

Mueller, 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, *Nature*, 2015, **517**, 455–459.

16   M. G. Moloney, *Trends Pharmacol. Sci.*, 2016, **37**, 689–701.

17   T. Homma, A. Nuxoll, A. B. Gandt, P. Ebner, I. Engels, T. Schneider, F. Götz, K. Lewis and B. P. Conlon, *Antimicrob. Agents Chemother.*, 2016, **60**, 6510–6517.

18   P. M. Dewick, in *Medicinal Natural Products, The Acetate Pathway: Fatty Acids and Polyketides*, 2009, pp. 39–131.

19   K. J. Weissman and P. F. Leadlay, *Nat. Rev. Microbiol.*, 2005, **3**, 925–936.

20   J. Staunton and K. J. Weissman, *Nat. Prod. Rep.*, 2001, **18**, 380–416.

21   C. Khosla, D. Herschlag, D. E. Cane and C. T. Walsh, *Biochemistry*, 2014, 2875–2883.

22   L. Katz, in *Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates*, Academic Press, 2009, vol. 459, pp. 113–142.

23   K. J. Weissman, *Nat. Prod. Rep.*, 2016, **33**, 203–230.

24   D. R. Cohen and C. A. Townsend, *Nat. Chem.*, 2018, **10**, 231–236.

25   D. R. Cohen and C. A. Townsend, *Angew. Chem. Int. Ed.*, 2018, **57**, 5650–5654.

26   Y. Tang, S. C. Tsai and C. Khosla, *J. Am. Chem. Soc.*, 2003, **125**, 12708–12709.

27   B. Shen, *Curr. Opin. Chem. Biol.*, 2003, **7**, 285–295.

28   D. Yu, F. Xu, J. Zeng and J. Zhan, *IUBMB Life*, 2012, **64**, 285–295.

29   I. J. Flores-Sanchez and R. Verpoorte, *Plant Physiol. Biochem.*, 2009, **47**, 167–174.

30   B. M. Kevany, D. A. Rasko and M. G. Thomas, *Appl. Environ. Microbiol.*, 2009, **75**, 1144–1155.

31   Y. A. Chan, A. M. Podevels, B. M. Kevany and M. G. Thomas, *Nat. Prod. Rep.*, 2009, **26**, 90–114.

32   L. Du, C. Sánchez and B. Shen, *Metab. Eng.*, 2001, **3**, 78–95.

33   J. Beld, D. J. Lee and M. D. Burkart, *Mol. Biosyst.*, 2015, **11**, 38–59.

34   A. Chen, R. N. Re and M. D. Burkart, *Nat. Prod. Rep.*, 2018, **35**, 1029–1045.

35   A. J. Birch, R. A. Massy-Westropp and C. J. Moye, *Aust. J. Chem.*, 1955, **8**, 539–544.

36   R. A. Hill, R. H. Carter and J. Staunton, *J. Chem. Soc., Perkin Trans. 1*, 1981, 2570–2576.

37   T. J. Simpson, in *Application of Isotopic Methods to Secondary Metabolic Pathways*, 1998.

38   H. Nakajima and T. Hamasaki, *J. Chem. Soc., Perkin Trans. 1*, 1994, 1865–1869.

39   F. Malpartida and D. A. Hopwood, *Nature*, 1984, **309**, 462–464.

40   J. Cortes, S. F. Haydock, G. A. Roberts, D. J. Bevitt and P. F. Leadlay, *Nature*, 1990, **348**, 176–178.

41   S. Okamoto, T. Taguchi, K. Ochi and K. Ichinose, *Chem. Biol.*, 2009, **16**, 226–236.

42   S.-C. (Sheryl) Tsai, *Annu. Rev. Biochem.*, 2018, **87**, 503–531.

43   A. Osbourn, *Trends Genet.*, 2010, **26**, 449–457.

44   B. Lowry, T. Robbins, C.-H. Weng, R. V O'Brien, D. E. Cane and C. Khosla, *J. Am. Chem. Soc.*, 2013, **135**, 16809–16812.

45   B. Callahan, M. Thattai and B. I. Shraiman, *PNAS*, 2009, **106**, 19410–19415.

46   M. H. Medema and M. A. Fischbach, *Nat. Chem. Biol.*, 2016, **11**, 639–648.

47   N. Ziemert, M. Alanjary and T. Weber, *Nat. Prod. Rep.*, 2016, **33**, 988–1005.

48 W.-G. Wang, L.-Q. Du, S.-L. Sheng, A. Li, Y.-P. Li, G.-G. Cheng, G.-P. Li, G. Sun, Q.-F. Hu and Y. Matsuda, *Org. Chem. Front.*, 2019, **6**, 571–578.

49 D. E. Cane and C. Yang, *J. Am. Chem. Soc.*, 1987, **109**, 1255–1257.

50 D. E. Cane, P. C. Prabhakaran, W. Tan and W. R. Ott, *Tetrahedron Lett.*, 1991, **32**, 5457–5460.

51 B. Rawlings, *Nat. Prod. Rep*, 1997, **14**, 523–556.

52 M. P. Crump, J. Crosby, C. E. Dempsey, J. A. Parkinson, M. Murray, D. A. Hopwood and T. J. Simpson, *Biochemistry*, 1997, **36**, 6000–6008.

53 M. Andrec, R. B. Hill and J. H. Prestegard, *Protein Sci.*, 1995, **4**, 983–993.

54 J. Crosby, D. H. Sherman, M. J. Bibb, W. P. Revill, D. A. Hopwood and T. J. Simpson, *Biochim. Biophys. Acta*, 1995, **1251**, 32–42.

55 A. S. Worthington, H. Rivera, J. W. Torpey, M. D. Alexander and M. D. Burkart, *ACS Chem. Biol.*, 2006, **1**, 687–691.

56 C. Nguyen, R. W. Haushalter, D. J. Lee, P. R. L. Markwick, J. Bruegger, G. Caldara-festin, K. Finzel, D. R. Jackson, F. Ishikawa, B. O. Dowd, J. A. Mccammon, S. J. Opella, S. Tsai and M. D. Burkart, *Nature*, 2013, **505**, 427–431.

57 J. L. Meier and M. D. Burkart, *Synthetic Probes for Polyketide and Nonribosomal Peptide Biosynthetic Enzymes*, Elsevier Inc., 1st edn., 2009, vol. 458.

58 R. J. M. Goss, S. Shankar and A. A. Fayad, *Nat. Prod. Rep.*, 2012, **29**, 870–889.

59 R. McDaniel, A. Thamchaipenet, C. Gustafsson, H. Fu, M. Betlach, M. Betlach and G. Ashley, *PNAS*, 1999, **96**, 1846–1851.

60 B. J. Dunn and C. Khosla, *J. R. Soc. Interface*, 2013, **10**, 1–10.

61 K. E. H. Wiesmann, J. Cortkl, M. J. B. Brown, A. Cutter, P. F. Leadlay and J. Staunton, *Chem. Biol.*, 1995, **2**, 583–589.

62 S. Purser, P. R. Moore, S. Swallow and V. Gouverneur, *Chem. Soc. Rev.*, 2008, **37**, 320–330.

63   C. S. Neumann, D. G. Fujimori and C. T. Walsh, *Chem. Biol.*, 2008, **15**, 99–109.

64   H. Deng, S. L. Cobb, A. R. McEwan, R. P. McGlinchey, J. H. Naismith, D. O'Hagan, D. A. Robinson and J. B. Spencer, *Angew. Chem. Int. Ed.*, 2006, **45**, 759–762.

65   A. S. Eustáquio, D. O'Hagan and B. S. Moore, *J. Nat. Prod.*, 2010, **73**, 1212–1217.

66   L. C. Blasiak and C. L. Drennan, *Acc. Chem. Res.*, 2009, **42**, 147–155.

67   G. Volpato, R. C. Rodrigues and R. Fernandez-Lafuente, *Curr. Med. Chem.*, 2010, 3855–3873.

68   T. Efferth, M. R. Romero, D. G. Wolf, T. Stamminger, J. J. G. Marin and M. Marschall, *Clin. Infect. Dis.*, 2008, **47**, 804–811.

69   R. K. Haynes, B. Fugmann, J. Stetter, K. Rieckmann, H. Heilmann, H. Chan, M. Cheung, W. Lam, H. Wong, S. L. Croft, L. Vivas, L. Rattray, L. Stewart, W. Peters, B. L. Robinson, M. D. Edstein, B. Kotecka, D. E. Kyle, B. Beckermann, M. Gerisch, M. Radtke, G. Schmuck, W. Steinke, U. Wollborn, K. Schmeer and A. Römer, *Angew. Chem. Int. Ed.*, 2006, **45**, 2082–2088.

70   J. Turconi, R. Guevel, G. Oddon, R. Villa, A. Geatti, M. Hvala, K. Rossen, R. Go and A. Burgard, *Org. Process Res. Dev.*, 2014, 417−422.

71   H. Morita, M. Yamashita, S.-P. Shi, T. Wakimoto, S. Kondo, R. Kato, S. Sugio, T. Kohno and I. Abe, *PNAS*, 2011, **108**, 13504–13509.

72   K. Bloudoff, D. A. Alonzo and T. M. Schmeing, *Cell Chem. Biol.*, 2016, **23**, 331–339.

73   S. C. Curran, A. Hagen, S. Poust, L. J. G. Chan, B. M. Garabedian, T. De Rond, M. Baluyot, J. T. Vu and J. D. Keasling, *ACS Chem. Biol.*, 2018, **13**, 2261–2268.

74   M. Tosin, D. Spiteller and J. B. Spencer, *ChemBioChem*, 2009, **10**, 1714–1723.

75   S. Smith and S. C. Tsai, *Nat. Prod. Rep.*, 2007, **24**, 1041–1072.

76   M. Tosin, L. Betancor, E. Stephens and W. M. A. Li, *ChemBioChem*, 2010, **11**,

539–546.

77   M. Tosin, Y. Demydchuk, J. S. Parascandolo, B. Per, F. J. Leeper and P. F. Leadlay, *Chem. Commun.*, 2011, **47**, 3460–3462.

78   D. A. K. Ii, D. A. F. Meujo and M. T. Hamann, *Expert Opin. Drug Discov.*, 2009, **4**, 109–146.

79   C. J. Dutton, B. J. Banks and C. B. Cooper, *Nat. Prod. Rep.*, 1995, **12**, 165–181.

80   J. Rutkowski and B. Brzezinski, *Biomed Res. Int.*, 2013, **2013**, 31.

81   S. Zhou, F. Wang, E. T. Wong, E. Fonkem, T. Hsieh, J. M. Wu and E. Wu, *Curr. Med. Chem.*, 2013, **20**, 4095–4101.

82   Y. N. Antonenko and L. S. Yaguzhinsky, *Biochim. Biophys. Acta*, 1988, **938**, 125–130.

83   I. Wilkening, S. Gazzola, E. Riva, J. S. Parascandolo and M. Tosin, *Chem. Commun.*, 2016, **52**, 10392–10395.

84   M. Tosin, L. Smith and P. F. Leadlay, *Angew. Chem. Int. Ed.*, 2011, **50**, 11930–11933.

85   E. Riva, I. Wilkening, S. Gazzola, W. M. A. Li, L. Smith, P. F. Leadlay and M. Tosin, *Angew. Chem. Int. Ed.*, 2014, **53**, 11944–11949.

86   J. S. Parascandolo, J. Havemann, H. K. Potter, F. Huang, E. Riva, J. Connolly, I. Wilkening, L. Song, P. F. Leadlay and M. Tosin, *Angew. Chem. Int. Ed.*, 2016, **55**, 3463–3467.

87   H. Kage, E. Riva, J. S. Parascandolo, M. F. Kreutzer, M. Tosin and M. Nett, *Org. Biomol. Chem.*, 2015, **13**, 11414–11417.

88   Y. T. C. Ho, D. J. Leng, F. Ghiringhelli, I. Wilkening, D. P. Bushell, O. Ko, E. Riva, J. Havemann, D. Passarella and M. Tosin, *Chem. Commun.*, 2017, **53**, 7088–7091.

89   C. W. Carreras and R. Pieper, *J. Am. Chem. Soc.*, 1996, **7863**, 5158–5159.

90   T. Hayashi, T. Noto, Y. Nawata, H. Okazaki, M. Sawada and K. Ando, *J. Antibiot.*, 1982, **35**, 771–777.

91   T. Noto, S. Miyakawa, H. Oishi, H. Endo and H. Okazaki, *J. Antibiot.*, 1981, **16**, 399–407.

92   H. Sasaki, H. Oishi, T. Hayashi, I. Matsuura, K. Ando and M. Sawada, *J. Antibiot.*, 1982, **35**, 396-400.

93   W. Tao, M. E. Yurkovich, S. Wen, K. E. Lebe, M. Samborskyy, Y. Liu, A. Yang, Y. Liu, Y. Ju, Z. Deng, M. Tosin, Y. Sun and P. F. Leadlay, *Chem. Sci*, 2016, **7**, 376–385.

94   X. Tang, J. Li, N. Millán-Aguiñaga, J. J. Zhang, E. C. O'Neill, J. A. Ugalde, P. R. Jensen, S. M. Mantovani and B. S. Moore, *ACS Chem. Biol.*, 2015, **10**, 2841–2849.

95   S. M. Jones, J. E. Urch, M. Kaiser, R. Brun, J. L. Harwood, C. Berry and I. H. Gilbert, *Bioorg. Med. Chem.*, 2004, **12**, 683–692.

96   S. M. Jones, J. E. Urch, M. Kaiser, R. Brun, J. L. Harwood, C. Berry and I. H. Gilbert, *J. Med. Chem.*, 2005, **48**, 5932–5941.

97   A. Mahajan, R. Hans, K. Chibale and V. Kumar, *RSC Adv.*, 2014, **4**, 15180–15215.

98   K. Kapilashrami, G. R. Bommineni, C. A. MacHutta, P. Kim, C. T. Lai, C. Simmerling, F. Picart and P. J. Tonge, *J. Biol. Chem.*, 2013, **288**, 6045–6052.

99   J. M. McFadden, S. M. Medghalchi, J. N. Thupari, M. L. Pinn, A. Vadlamudi, K. I. Miller, F. P. Kuhajda and C. A. Townsend, *J. Med. Chem.*, 2005, **48**, 946–961.

100   A. C. Price, K. H. Choi, R. J. Heath, Z. Li, S. W. White and C. O. Rock, *J. Biol. Chem.*, 2001, **276**, 6551–6559.

101   L. Kremer, J. D. Douglas, A. R. Baulard, C. Morehouse, M. R. Guy, D. Alland, L. G. Dover, J. H. Lakey, W. R. Jacobs, P. J. Brennan, D. E. Minnikin and G. S. Besra, *J. Biol. Chem.*, 2000, **275**, 16857–16864.

102   S. R. Luckner, C. A. Machutta, P. J. Tonge and C. Kisker, *Structure*, 2009, **17**,

1004–1013.

103 K. Young, H. Jayasuriya, J. G. Ondeyka, K. Herath, C. Zhang, S. Kodali, A. Galgoci, R. Painter, V. Brown-driver, R. Yamamoto, L. L. Silver, Y. Zheng, J. I. Ventura, J. Sigmund, S. Ha, A. Basilio, F. Vicente, F. Pelaez, P. Youngman, D. Cully, J. F. Barrett, D. Schmatz, S. B. Singh and J. Wang, *Antimicrob. Agents Chemother.*, 2006, **50**, 519–526.

104 M. S. Brown, K. Akopiants, D. M. Resceck, H. A. I. McArthur, E. McCormick, K. A. Reynolds, V. Commonwealth and V. Uni, *J. Am. Chem. Soc.*, 2003, **125**, 10166–10167.

105 K. L. Dunbar, D. H. Scharf, A. Litomska and C. Hertweck, *Chem. Rev.*, 2017, **117**, 5521–5577.

106 A. J. Waldman, T. L. Ng, P. Wang and E. P. Balskus, *Chem. Rev.*, 2016, **117**, 5784–5863.

107 C. J. Wang and J. M. Salvino, *Tetrahedron Lett.*, 1984, **25**, 5243–5246.

108 A. Kamal, A. A. Shaik, R. Sinha, J. S. Yadav and S. K. Arora, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 1927–1929.

109 M. S. Chambers, E. J. Thomas and D. J. Williams, *J. Chem. Soc., Chem. Commun.*, 1987, **0**, 1228–1230.

110 J. M. McFadden, G. L. Frehywot and C. A. Townsend, *Org. Lett.*, 2002, **4**, 3859–3862.

111 T. Weixin, Z. Manghong, D. Zixin and S. Yuhui, *Sci. China Chem.*, 2013, **56**, 1364–1371.

112 L. Vieweg, S. Reichau, R. Schobert, P. F. Leadlay and R. D. Süssmuth, *Nat. Prod. Rep*, 2014, **31**, 1554–1584.

113 C. J. Newbold, R. J. Wallace, N. D. Watt and A. J. Richardson, *Appl. Environ. Microbiol.*, 1988, **54**, 544–547.

114　C. J. Newbold, R. J. Wallace and N. D. Walker-bax, *FEMS Microbiol Lett*, 2018, **338**, 161–167.

115　C. Keller-Juslén, H. D. King, M. Kuhn, H. R. Loosli, W. Pache, T. J. Petcher, H. P. Weber and A. von Wartburg, *J. Antibiot.*, 1982, **35**, 142–150.

116　K. Hori, H. Kazuno, K. Nomura and E. Yoshii, *Tetrahedron Lett.*, 1993, **34**, 2183–2186.

117　M. M. Faul and B. E. Huff, *Chem. Rev.*, 2000, **100**, 2407–2473.

118　S. V Ley, J. A. Clase, D. J. Mansfield and H. M. I. Osborn, *J. Heterocycl. Chem.*, 1996, **33**, 1533–1544.

119　S. V. Ley, D. S. Brown, J. A. Clase, A. J. Fairbanks, I. C. Lennon, H. M. I. Osborn, E. S. E. Stokes and D. J. Wadsworth, *J. Chem. Soc., Perkin Trans. 1*, 1998, **1**, 2259–2276.

120　D. Doddrell, E. Laue, F. Leeper, J. Staunton, A. Davies and A. B. Davies, *J. Chem. Soc., Chem. Commun.*, 1984, **0**, 1302–1304.

121　A. K. Demetriadou, E. Laue, J. Staunton, G. A. F. Ritchie, A. Davies and alan, *J. Chem. Soc., Chem. Commun.*, 1985, **0**, 408–410.

122　H. C. Hailes, C. M. Jackson, P. F. Leadlay, S. V Ley and J. Staunton, *Tetrahedron Lett.*, 1994, **35**, 307–310.

123　H. C. Hailes, S. Handa, P. F. Leadlay, I. C. Lennon, S. V Ley and J. Staunton, *Tetrahedron Lett.*, 1994, **35**, 311–314.

124　H. C. Hailes, S. Handa, P. F. Leadlay, I. C. Lennon, S. V Ley and J. Staunton, *Tetrahedron Lett.*, 1994, **35**, 315–318.

125　G.-J. Boons, J. A. Claise, I. C. Lennon, S. V Ley and J. Staunton, *Tetrahedron Lett.*, 1995, **51**, 5417–5446.

126　M. J. Bulsing, E. D. Laue, F. J. Leeper, J. Staunton, D. H. Davies, G. A. F. Ritchie, A. Davies, A. B. Davies and R. P. Mabelisb, *J. Chem. Soc., Chem. Commun.*, 1984,

0, 1301–1302.

127   C. J. Dutton and J. Staunton, *Tetrahedron Lett.*, 1996, **37**, 3511–3514.

128   S. L. Less, S. Handa, K. Millburn, P. F. Leadlay, C. J. Dutton and J. Staunton, *Tetrahedron Lett.*, 1996, **37**, 3515–3518.

129   S. L. Less, P. F. Leadlay, C. J. Dutton and J. Staunton, *Tetrahedron Lett.*, 1996, **37**, 3519–3520.

130   Y. Demydchuk, Y. Sun, H. Hong, J. Staunton, J. B. Spencer and P. F. Leadlay, *ChemBioChem*, 2008, **9**, 1136–1145.

131   Y. Sun, H. Hong, F. Gillies, J. B. Spencer and P. F. Leadlay, *ChemBioChem*, 2008, **9**, 150–156.

132   G.-J. Boons, D. S. Brown, J. A. Clase, I. C. Lennon and S. V Ley, *Tetrahedron Lett.*, 1994, **35**, 319–322.

133   G.-J. Boons, I. C. Lennon, S. V Ley, E. S. E. Owen, J. Staunton and D. J. Wadsworth, *Tetrahedron Lett.*, 1994, **35**, 323–326.

134   J. Havemann, M. E. Yurkovich, R. Jenkins, S. Harringer, W. Tao, S. Wen, Y. Sun, P. F. Leadlay and M. Tosin, *Chem. Commun.*, 2017, **53**, 1912–1915.

135   V. Schulz, Masters Thesis, University of Cambridge, 2017.

136   M. E. Yurkovich, R. Jenkins, Y. Sun, M. Tosin and P. F. Leadlay, *Chem. Commun.*, 2017, **53**, 2182–2185.

137   M. E. Yurkovich, PhD Thesis, University of Cambridge, 2018.

138   X. Tang, J. Li and B. S. Moore, *ChemBioChem*, 2017, **18**, 1072–1076.

139   R. Sanichar and J. C. Vederas, *Org. Lett.*, 2017, **19**, 1950–1953.

140   J. Li, X. Tang, T. Awakawa and B. S. Moore, *Angew. Chem. Int. Ed.*, 2017, **56**, 12234–12239.

141   R. Little, F. Paiva, R. Jenkins, H. Hong, Y. Sun, Y. Demydchuk, M. Samborskyy,

M. Tosin, F. Leeper, M. Dias and P. Leadlay, *Nat. Cat.*, 2019, **2**, 1045–1054.

142 A. Migita, M. Watanabe, Y. Hirose, K. Watanabe, T. Tokiwano, H. Kinashi and H. Oikawa, *Biosci. Biotechnol., Biochem.*, 2009, **73**, 169–176.

143 A. R. Gallimore, C. B. W. Stark, A. Bhatt, B. M. Harvey, Y. Demydchuk, V. Bolanos-Garcia, D. J. Fowler, J. Staunton, P. F. Leadlay and J. B. Spencer, *Chem. Biol.*, 2006, **13**, 453–460.

144 K. Hotta, X. Chen, R. S. Paton, A. Minami, H. Li, K. Swaminathan, I. I. Mathews, K. Watanabe, H. Oikawa, K. N. Houk and C.-Y. Kim, *Nature*, 2012, **483**, 355.

145 T. Fonseca, N. P. Lopes, P. J. Gates and J. Staunton, *J. Am. Soc. Mass. Spectrom.*, 2004, **15**, 325–335.

146 A. Minami, M. Shimaya, G. Suzuki, A. Migita, S. S. Shinde, K. Sato, K. Watanabe, T. Tamura, H. Oguri and H. Oikawa, *J. Am. Chem. Soc.*, 2012, **134**, 7246–7249.

147 G. Suzuki, A. Minami, M. Shimaya, T. Kodama, Y. Morimoto, H. Oguri and H. Oikawa, *Chem. Lett.*, 2014, **43**, 1779–1781.

148 F. T. Wong, K. Hotta, X. Chen, M. Fang, K. Watanabe and C.-Y. Kim, *J. Am. Chem. Soc.*, 2015, **137**, 86–89.

149 G. A. Patani and E. J. LaVoie, *Chem. Rev.*, 1996, **96**, 3147–3176.

150 A. A. Abdel-hafez and B. A. Abdel-wahab, *Bioorg. Med. Chem.*, 2008, **16**, 7983–7991.

151 J. A. Bull, R. A. Croft, O. A. Davis, R. Doran and K. F. Morgan, *Chem. Rev.*, 2016, **116**, 12150–12233.

152 M. McLaughlin, R. Yazaki, T. C. Fessard and E. M. Carreira, *Org. Lett.*, 2014, **16**, 4070–4073.

153 A. Okano, R. C. James, J. G. Pierce, J. Xie and D. L. Boger, *J. Am. Chem. Soc.*, 2012, **134**, 8790–8793.

154 P. Klán, T. Šolomek, C. G. Bochet, A. Blanc, R. Givens, M. Rubina, V. Popik, A.

Kostikov and J. Wirz, *Chem. Rev.*, 2013, **113**, 119–191.

155 M. J. Hansen, W. A. Velema, M. M. Lerch, W. Szymanski and B. L. Feringa, *Chem. Soc. Rev.*, 2015, **44**, 3358–3377.

156 I. Aujard, C. Benbrahim, M. Gouget, O. Ruel, J.-B. Baudin, P. Neveu and L. Jullien, *Chem. Eur. J.*, 2006, **12**, 6865–6879.

157 M. Gaplovsky, Y. V Il'ichev, Y. Kamdzhilov, S. V Kombarova, M. Mac, M. A. Schwörer and J. Wirz, *Photochem. Photobiol. Sci.*, 2005, **4**, 33–42.

158 S. Kilgour, R. D. Jenkins and M. Tosin, *Chem. Eur. J.*, 2019, **25**, 16511.

159 A. B. Mabire, M. P. Robin, W.-D. Quan, H. Willcock, V. G. Stavros and R. K. O'Reilly, *Chem. Commun.*, 2015, **51**, 9733–9736.

160 S. Kilgour, PhD Thesis, University of Warwick, 2018.

161 Y. Elemes and C. S. Foote, *J. Am. Chem. Soc.*, 1992, **114**, 6044–6050.

162 P. Patel and G. Pattenden, *Tetrahedron Lett.*, 1985, **26**, 4789–4792.

163 C. Kuroda, C. Y. Tang, M. Tanabe and M. Funakoshi, *Bull. Chem. Soc. Jpn.*, 1999, **72**, 1583–1587.

164 M. S. Lee, G. Qin, K. Nakanishi and M. G. Zagorski, *J. Am. Chem. Soc.*, 1989, **111**, 6234–6241.

165 D. M. Roberts, C. Bartel, A. Scott, D. Ivison, T. J. Simpson and R. J. Cox, *Chem. Sci*, 2017, **8**, 1116–1126.

166 J. E. Moses, J. E. Baldwin, S. Brückner, S. J. Eade and R. M. Adlington, *Org. Biomol. Chem.*, 2003, **1**, 3670–3684.

167 P. Patel and G. Pattenden, *J. Chem. Soc., Perkin Trans. 1*, 1991, 1941–1945.

168 R. L. Danheiser and J. S. Nowick, *J. Org. Chem.*, 1991, **56**, 1176–1185.

169 M. Barbero, S. Cadamuro, S. Dughera and P. Venturello, *Synthesis (Stuttg).*, 2008, **2008**, 3625–3632.

170   M. Mir Mohseni, T. Höver, L. Barra, M. Kaiser, P. C. Dorrestein, J. S. Dickschat and T. F. Schäberle, *Angew. Chem. Int. Ed.*, 2016, **55**, 13611–13614.

171   C. Aubert, J.-P. Bégué, M. Charpentier-Morize, G. Nee and B. Langlois, *J. Fluor. Chem.*, 1989, **44**, 361–376.

172   S. Watanabe, Y. Sakai, T. Kitazume and T. Yamazaki, *J. Fluor. Chem.*, 1994, **68**, 59–61.

173   T. Kawasaki, T. Ichige and T. Kitazume, *J. Org. Chem.*, 1998, **63**, 7525–7528.

174   A. L. Sewell, M. V. J. Villa, M. Matheson, W. G. Whittingham and R. Marquez, *Org. Lett.*, 2011, **13**, 800–803.

175   N. M. Gaudelli and C. A. Townsend, *J. Org. Chem.*, 2013, **78**, 6412–6426.

176   N. Chen, H. Du, W. Liu, S. Wang, X. Li and J. Xu, *Phosphorus, Sulfur Silicon Relat. Elem.*, 2015, **190**, 112–122.

177   M. Tanaka, O. Oota, H. Hiramatsu and K. Fujiwara, *Bull. Chem. Soc. Jpn.*, 1988, **61**, 2473–2479.

178   M. Häckh, M. Müller and S. Lüdeke, *Chem. Eur. J.*, 2013, **19**, 8922–8928.

179   F. Sagui, C. De Micheli, G. Roda, P. Magrone, R. Pizzoli and S. Riva, *J. Mol. Catal. B Enzym.*, 2012, **75**, 27–34.

180   H. Kamaya, M. Sato and C. Kaneko, *Tetrahedron Lett.*, 1997, **38**, 587–590.

181   C. Simion, I. Hashimoto, Y. Mitoma, A. M. Simion and N. Egashira, *Phosphorus, Sulfur Silicon Relat. Elem.*, 2010, **185**, 2480–2488.

182   U.S. Patent 21639, 2003.

183   A. Kamal, A. A. Shaik, S. Azeeza, M. S. Malik and M. Sandbhor, *Tetrahedron: Asymmetry*, 2006, **17**, 2890–2895.

184   J. D. Aebi, D. T. Deyo, C. Q. Sun, D. Guillaume, B. Dunlap and D. H. Rich, *J. Med. Chem.*, 1990, **33**, 999–1009.

185   J. Iqbal, A. Pandey and B. P. S. Chauhan, *Tetrahedron*, 1991, **47**, 4143–4154.

186   O. Kreye, O. Türünç, A. Sehlinger, J. Rackwitz and M. A. R. Meier, *Chem. Eur. J.*, 2012, **18**, 5767–5776.

187   R. Ballini, D. Fiorini and A. Palmieri, *Synlett*, 2003, **2003**, 1841–1843.

188   Y. Ryu and A. I. Scott, *Tetrahedron Lett.*, 2003, **44**, 7499–7502.

189   H. K. Potter, PhD Thesis, University of Cambridge, 2011.