UNIVERSITY OF CAMBRIDGE

Investigating Intermediates in 6-Methylsalicylic Acid Biosynthesis

Helen K. Potter

Girton College

This dissertation is submitted for the degree of Doctor of Philosophy

Preface

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

Helen Potter
29th September 2010

Abstract

6-Methylsalicylic acid (6-MSA) is one of the oldest known polyketides. It is synthesised *in vivo* by the polyketide synthase 6-methylsalicylic acid synthase (6-MSAS), a multifunctional enzyme which uses its active sites iteratively. The stereochemistry of the hydroxyl produced from the single ketoreduction, as well as the order of dehydration, cyclisation and aromatisation steps, remain cryptic, despite extensive study.

Holo 6-MSAS was heterologously expressed in *E. coli* and purified in two steps. A non-hydrolysable carba(dethia)malonyl-N-acetylcysteamine analogue was synthesised and used to off-load enzyme-bound intermediates from 6-MSAS. In assays with acetyl-CoA and acetoacetyl-CoA alone, diketide and triketide intermediates were off-loaded and detected by HPLC-HR-ESI-MS. In the presence of NADPH, the off-loaded triketide was reduced by the ketoreductase domain of 6-MSAS. A potential dehydrated intermediate was also observed.

The dehydratase domain of 6-MSAS has recently been reassigned as a thioester hydrolase. To test this theory, the catalytic histidine residue in 6-MSAS was mutated to an alanine and the abolition of production of 6-MSA *in vivo* was observed. Mutated 6-MSAS was still able to produce the shunt product triacetic acid lactone. Incubation of mutated 6-MSAS with acetyl-CoA, malonyl-CoA, NADPH and carba(dethia)malonyl-N-acetylcysteamine saw only the off-loading of diketide and triketide analogues.

To investigate the stereochemistry of ketoreduction in 6-MSA biosynthesis, steps were made to synthesise the resolved diastereomeric reduced-triketide CoAs which would be the substrates for the ketoreductase domain. Attempts to phosphopantetheinylate *apo* 6-MSAS *in vitro* with three different phosphopantetheinyltransferases were unsuccessful. Limited proteolysis of both *holo* and *apo* 6-MSAS found that the *apo* synthase rapidly lost a C-terminal fragment while *holo* 6-MSAS was much more stable under the same conditions. Attempts were made to express the acyl carrier protein domain from 6-MSAS to overcome these problems.

These experiments represent the first use of the non-hydrolysable analogue methodology in a Type I iterative polyketide synthase and provide a framework for future experiments investigating intermediates in the biosynthesis of 6-MSA.

Acknowledgements

Thanks must first go to Dr Jonathan Spencer for giving me this fascinating project and for his supervision and guidance in the early stages of this work. After Joe's untimely death, I am deeply grateful to Dr Finian Leeper and Professor Peter Leadlay for stepping in and not only providing supervision, but keeping the research group together and allowing it to continue to thrive. Their continued interest in this project and guidance has been invaluable.

I would like to thank Dr Fanglu Huang for her biochemical expertise and patience in teaching a chemist. Thanks must also go to Dr Manuela Tosin for her synthetic advice and guidance in the field of non-hydrolysable analogues. I am extremely grateful that she was able to run all the samples on the Orbitrap for me in addition to analysing the data, without which this would be a very sparse thesis indeed. I am very thankful for the assistance of Dr Andrew Truman, whether in his commitment to the continued running of the LC-MS, dedication to tea-time, proof-reading of this thesis or as a general sounding board for ideas. I would like to acknowledge the contribution of Geoff Taster, my part III student who took on a difficult project with enthusiasm. I am also indebted to Mike Kenning for assistance with the analysis of the 6-MSAS gene sequence.

All of the members of the Spencer and Leadlay groups have been an excellent source of ideas and encouragement. I would especially like to acknowledge the contributions of Andy H, Nick, Frank, Jenny, Habib, Shu and Chris to the tea-time ritual of commiseration.

Outside the lab, the constant support of my friends has been invaluable. Thanks to Jenny, Adam, James, Anna, Mehdi, Carina and Rebecca to name but a few. Without you it would have been a much poorer experience.

Finally I must extend my heartfelt thanks to my family. Mum, Dad, Sarah, Deborah, Chris and Ellie – there are no words.

Table of Contents

Chapter 3: The dehydratase/thioester hydrolase domain of 6-MSAS

Chapter 4: The stereochemistry of ketoreduction by 6-MSAS

Chapter 5: Materials and methods

Addendum: Additional experiments in 6-methylsalicylic acid biosynthesis and synthetic approaches towards 5-amidinyl teicoplanin

Abbreviations

6-dEB	6-Deoxyerythronolide B
6-MSA	6-Methylsalicylic acid
6-MSAS	6-Methylsalicylic acid synthase
6-MSAS-DHm	6-Methylsalicylic acid synthase with an inactivated DH domain
ACP	Acyl carrier protein
AT	Acyl transferase domain
ATP	Adenosine-5'-triphosphate
ATX	A 6-MSAS from *Aspergillus terreus*
ATX-DHm	ATX with an inactivated DH domain
BSA	Bovin serum albumin
B. subtilis	*Bacillus subtilis*
Cb	Carbenicillin
CoA	Coenzyme A
CoASH	Coenzyme A (free thiol)
CLF	Chain length factor
DCC	N,N'-Dicyclohexylcarbodiimide
DCE	1,2-Dichloroethane
DCM	Dichloromethane
DEBS	6-Deoxyerythronolide B synthase
DH	dehydratase domain
DIPEA	N,N-Diisopropylethylamine
DK	Dephospho-coenzyme A kinase
DMAP	4-Dimethylaminopyridine
DMSO	Dimethylsulphoxide
DNA	5’-deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
E. coli	*Escherichia coli*

EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ER	Enoyl reductase domain
FAS	Fatty acid synthase
FPLC	Fast protein liquid chromatography
FTMS	Fourier-transform mass spectrometry
GABA	γ-aminobutyric acid
HATU	2-(1H-7-Azabenzotriazol-1-yl)--1,1,3,3-tetramethyl uronium hexafluorophosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His_6-	Hexahistidine affinity tag
HPLC	High performance liquid chromatography
HPLC-HR-ESI-MS	High performance liquid chromatography, high resolution electrospray ionisation mass spectrometry
HR-PKS	Highly reducing PKS
ID	Interdomain region
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KR	Ketoreductase domain
KS	Ketosynthase domain
LB	Luria-Bertani
LC-MS	Liquid chromatography – mass spectrometry
LDD	Leucine-aspartate-aspartate
MAT	Malonyl-CoA:ACP transacylase domain
mFAS	Mammalian fatty acid synthase
MPT	Malonyl/palmitoyl transferase domain
MS	Mass spectrometry
NAC	N-acetyl cysteamine
NADP+	Nicotinamide adenine dinucleotide phosphate (oxidised form)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
NR-PKS	Non-reducing PKS

NRPS	Non-ribosomal peptide synthetase
ORF	Open reading frame
PCR	Polymerase chain reaction
PK	Pantothenate kinase
PKS	Polyketide synthase
PLE	Pig liver esterase
PMSF	Phenylmethylsulfonyl fluoride
P. patulum	*Penicillium patulum*
PPTase	4'-Phosphopantetheinyltransferase
PR-PKS	Partially reducing PKS
PT	Phosphopantetheineadenyltransferase
PT	Product template domain
PVDF	Polyvinylidene fluoride
SAT	Starter unit acyl transferase domain
S. aurantiaca	*Stigmatella aurantiaca*
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodium dodecylsulphate
SDS-PAGE	Sodium dodecylsulphate polyacrylamide gel electrophoresis
S. erythraea	*Saccharopolyspora erythraea*
TAL	Triacetic acid lactone
TE	Thioesterase domain
TH	Thioester hydrolase domain
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TLCK	Tosyl lysyl chloromethyl ketone
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultraviolet
UV/Vis	UV/Visible spectroscopy

CHAPTER 1

Introduction

1.1 Natural products

Natural products can be defined as compounds which are derived from a natural source such as animals, plants and micro-organisms and which often have biological activities.[1] They have been used in crude form by human beings for medicinal purposes for centuries.

The interest in purified natural products in a pharmaceutical context began after the mass production of penicillin during World War II.[2] The antibiotics streptomycin[3], gentamicin[4], tetracycline[5] and erythromycin[6] were discovered within a few years and shifted the focus of the pharmaceutical industry to compounds which could be obtained by fermentation.

The emergence of antibiotic resistance in a clinical environment has led to a biological arms race. The development of new antibiotics to treat infections that have become resistant to older antibiotics is in direct competition with the ability of micro-organisms to mutate and gain resistance to these new drugs. The emergence of so-called 'superbugs' such as methicillin-resistant *Staphylococcus aureus* (MRSA)[7],

vancomycin-resistant *enterococci* (VRE)[8] and extensively drug-resistant tuberculosis (XDR-TB)[9] has revealed how finely balanced the contest is.

Advances in natural product chemistry did not just yield antibiotics. Anti-fungals such as amphotericin[10], immunosuppressants like rapamycin[11] and the cholesterol-lowering statins[12], such as compactin and mevinolin have all entered clinical use.

Natural products are biosynthesised as part of secondary metabolism in organisms. Primary metabolism encompasses all processes in a living organism which are necessary for the maintenance of life. Secondary metabolism is concerned with the production of molecules which are not absolutely necessary for life such as antibiotics, pigments, pheromones and signalling molecules in quorum sensing.

Natural product structures make excellent lead compounds in the search for improved pharmaceuticals as they have a wide variety of targets within cells.[13] Both gene expression and DNA transcription can be altered using compounds which alkylate DNA, inhibit DNA polymerases or inhibit DNA topisomerases.[14] Protein biosynthesis can be disrupted by compounds such as the aminoglycosides which bind to one of the ribosomal subunits. This has been exploited to yield novel aminoglycoside analogues which can cause readthrough of unnatural stop codons.[15] Mammalian fatty acid synthase (mFAS) is a useful target as cancer cells express a much higher level of this enzyme than normal cells. An analogue of the natural product methylenolactocin which inhibits mFAS was shown to be active against human breast cancer cells.[16] Plant-derived natural products, such as the benzophenanthridine alkaloid chelerythine and other structurally related compounds, have been shown to trigger cell death by the intrinsic apoptotic pathway. For the majority of significant targets within a cell, it has been postulated that there exists at least one cognate natural product ligand.[14]

1.2 Polyketides

Polyketides are an extremely successful class of pharmaceutically useful natural products. The most successful cardiovascular drugs of all time, the statins, are of polyketide origin. Of natural product and natural product-derived compounds which were approved for clinical trials in 2005-2007, more than a third were based on polyketides.[17]

Clinically important polyketides include the antibiotic erythromycin (**4**) and its semi-synthetic derivative azithromycin, the enediyne class of anticancer agents[18], the immunosuppressant rapamycin and the anti-hypercholesterolemic lovastatin. The success of polyketides against a wide range of cellular targets is due to their great structural diversity. The class ranges from 6-methylsalicylic acid (6-MSA, **1**) and orsellinic acid (**2**), which contain only eight carbons, up to maitotoxin (**3**), a marine polyketide which is the largest and most toxic non-biopolymer natural product known with 168 carbons in its backbone.[19]

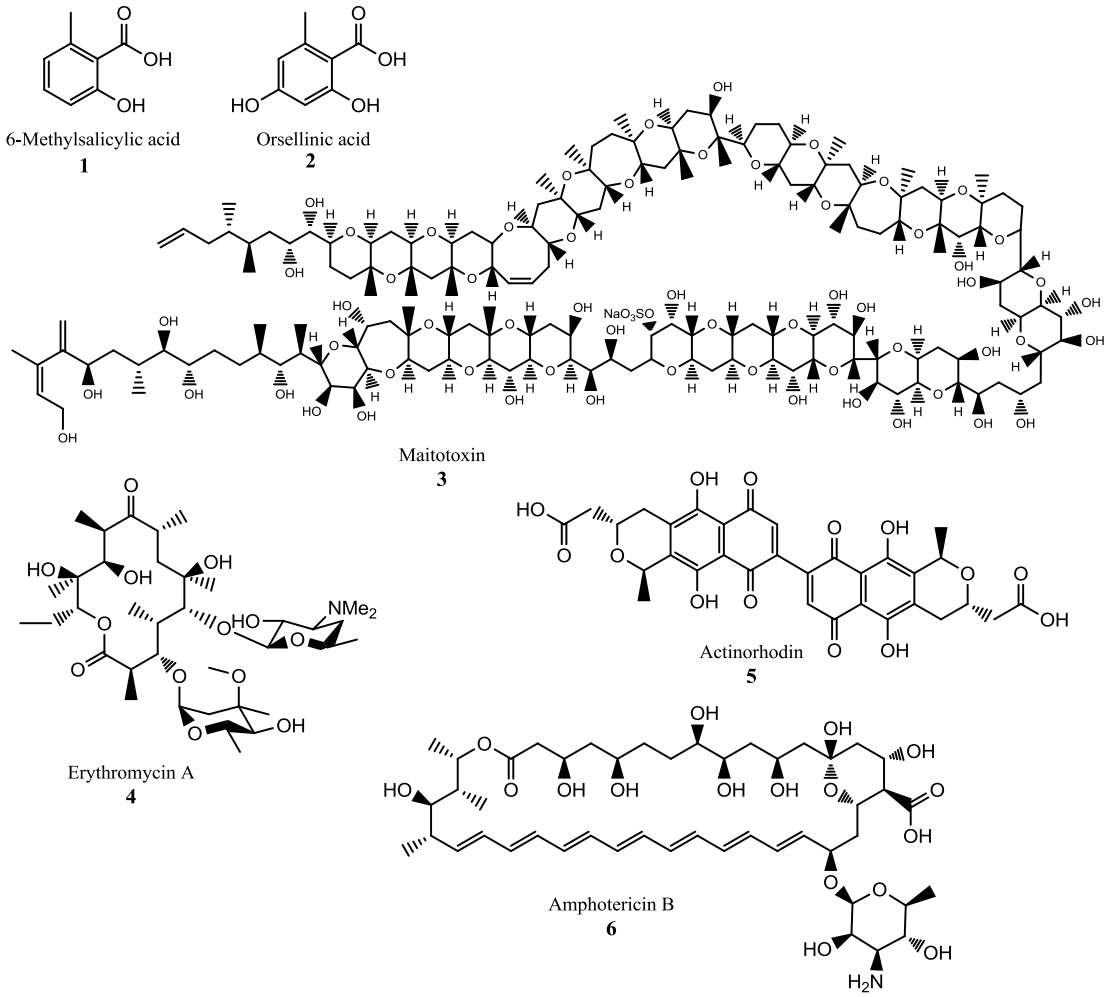


Figure 1.1: A great diversity in polyketide-derived skeletons exists, from small aromatic polyketides (6-MSA, **1** and orsellinic acid, **2**) , macrolides such as erythromycin A (**4**) and amphotericin B (**6**) and polyaromatic compounds like actinorhodin (**5**) to very large polyethers (maitotoxin, **3**).

The variation in size is not the only reason for the efficacy of polyketides. The polyketide class encompasses a vast spectrum of functionality, from the macrocycles such as erythromycin (**4**), aromatic compounds such as actinorhodin (**5**), polyenes such as amphotericin B (**6**) and polyethers such as monensin and maitotoxin (**3**) (Figure 1.1).

Despite the great array of structures, all polyketides are biosynthesised by a common mechanism from a common pool of building blocks.

1.2.1 The acetate hypothesis

The earliest forays into investigating polyketide biosynthesis were undertaken by James Collie in 1893. After boiling dehydroacetic acid (**7**) with barium hydroxide he discovered that one of the products was the aromatic compound orcinol (**8**) (Figure 1.2).[20] He proposed that this could be formed via a polyketone intermediate. Further investigations such as the formation of a naphthalene derivative from diacetylacetone and the formation of orcinol from ethyl acetoacetate[21] led Collie to the hypothesis that all of these natural products could be formed by repeated condensation of acetate units. He also extended his theory to include the biosynthesis of sugars by hydrolysis of pyrones and from there to starches and celluloses.[22] With no available methods to test the theory the acetate hypothesis failed to gain ground in the natural product field for the next 45 years.[23]

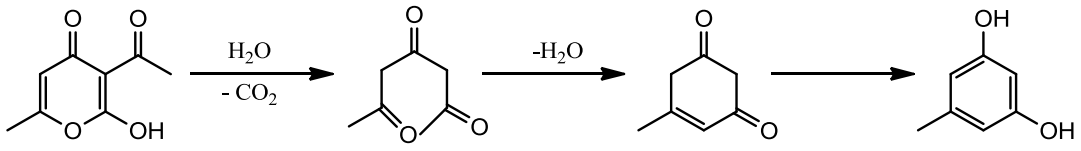
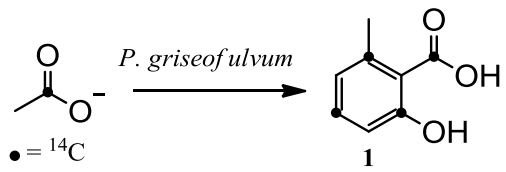


Figure 1.2: Collie's synthesis of orcinol (**8**) from dehydroacetic acid (**7**).

The revival of the acetate hypothesis and interest in the field of polyketide biosynthesis was due to experiments carried out by Birch in the 1950s. To confirm his acetate hypothesis Birch administered ^{14}C-labelled acetate to a culture of *Penicillium griseofulvum* which produced the polyketide 6-MSA (**1**).[24] The biosynthesis was proposed to feature the head-to-tail condensation of four acetate units, a reduction at one of the keto-groups to form a hydroxyl, an aldol condensation to form the 6-membered ring followed by dehydration and enolisation to yield the aromatic product. This putative pathway would lead to the incorporation of radioactivity at four sites in the product 6-MSA. The pattern of distribution of radioactivity was determined by a series of degradation reactions to produce

identifiable fragments and confirmed that 6-MSA was assembled from an acetate precursor.



1

Figure 1.3: The incorporation of labelled acetate into 6-MSA (**1**) by feeding the labelled compound to *Penicillium griseovfulvum*.

1.3 Fatty acid biosynthesis

Much work of the early understanding in the field of polyketide biosynthesis came from insights gained in the related field of fatty acid biosynthesis.

Fatty acids are synthesised *de novo* in all living organisms for use in membrane biosynthesis, although other uses such as post-translational modification of proteins and bacterial quorum sensing are also known.[25] Their biosynthesis proceeds by repeated condensation of C_2 units until a chain of the correct length is formed. The two-carbon unit is from malonyl-CoA (**9**), which reacts by decarboxylative Claisen condensation with an initial acyl unit. Throughout this process both the malonate and acyl compounds are bound to the enzyme catalysing the reaction, fatty acid synthase (FAS), as thioesters.

The fatty acid synthase is divided up into distinct functionalities, known as domains, each of which catalyses a separate reaction in the biosynthetic pathway. In animals and fungi the FAS is a single large protein, with the active sites distributed along the polypeptide, also known as a type I synthase. In bacteria the FAS active sites are all on separate proteins which associate in solution to perform reactions This is termed a type II synthase.

A malonate unit is attached to the acyl carrier protein (ACP) domain as a thioester. Rather than directly attaching the malonate to a cysteine residue, a serine residue on the ACP is modified post-translationally to attach a 4'-phosphopantetheine arm.[26] This is added by the enzyme *holo*-ACP synthase which transfers the 4'-phosphopantetheine moiety from coenzyme A (CoA). The phosphopantetheine arm has a terminal thiol to which the malonate group becomes attached as a thioester (Figure 1.4). The addition of a flexible 18 Å arm to the ACP allows the growing fatty acid chain to be delivered to the active sites of other domains while still remaining bound to the enzyme. The ACP in type II systems has been shown to be able to load malonyl units onto itself without need for a transferase domain.[27] In Type I systems a malonyl-CoA:ACP transacylase domain (MAT) is used to transfer the malonyl group from malonyl-CoA to ACP.

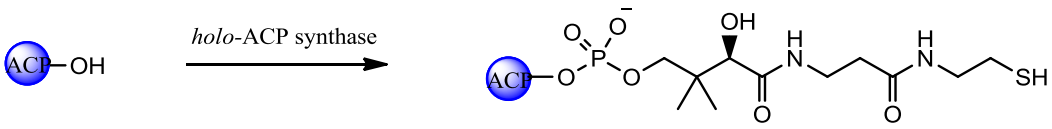


Figure 1.4: The addition of a 4'-phosphopanthetheine arm to the active site serine on the ACP is required for activity.

The ketosynthase (KS) domain catalyses the decarboxylative condensation of a malonyl unit, which is bound to the ACP, with an initial acetyl unit (transferred by MAT), which is bound to a cysteine residue in its active site as a thioester (Figure 1.5). This leads to a chain which has been extended by two carbons and is attached to the ACP. The β-ketothioester formed is then further reacted to form the saturated acyl thioester. Firstly, the ketone is reduced to a hydroxyl group by the ketoreductase (KR) domain, followed by elimination to form a double bond which is catalysed by the dehydratase (DH) domain. The double bond is further reduced to the final saturated product by the enoyl reductase (ER) domain. Following this round of reductive processing, the extended chain is transferred back to the active cysteine in the KS domain. A malonate unit is transferred on to the ACP by the MAT and the process repeats. This cycle repeats until a chain of the correct length is formed

(usually C14, C16 or C18), when the chain is hydrolysed from the ACP by the action of the thioesterase (TE) domain to form the free acid (Figure 1.5). In fungal FAS the chain is released as a CoA-ester after transfer from the ACP by a bi-functional malonyl/palmitoyl transferase (MPT).[25]

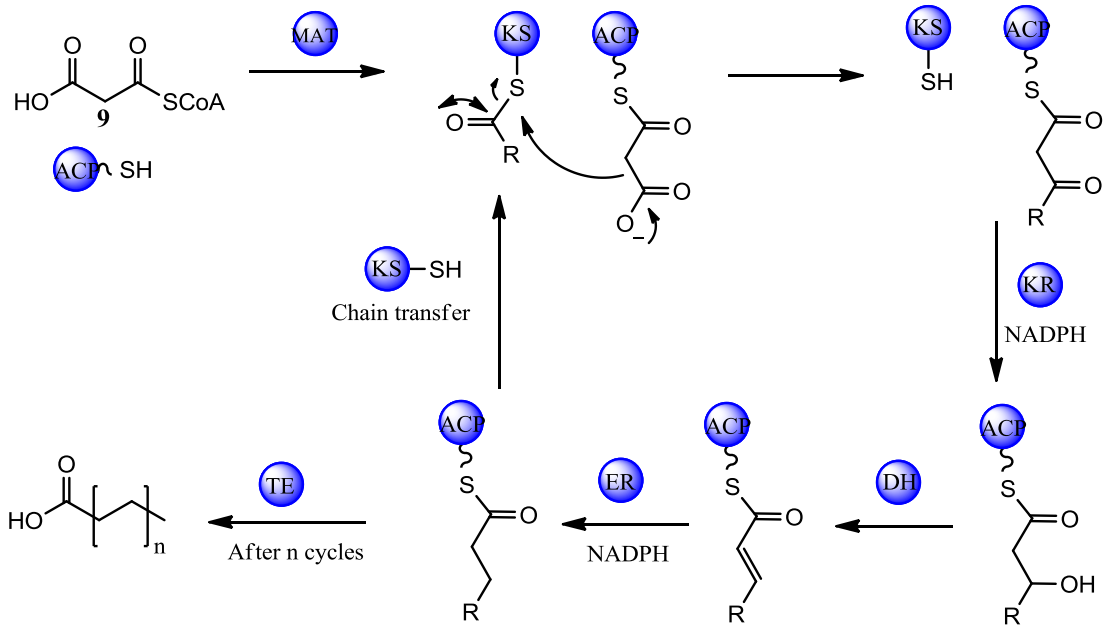


Figure 1.5: Fatty acid biosynthesis proceeds *via* the decarboxylative Claisen condensation of malonyl-ACP with the acyl chaing on the KS. The β-ketoester is then reduced by the KR, the hydroxyl is dehydrated by the DH and the double bond is reduced by the ER. The chain can then be transferred to the KS for another round of elongation.

While the mechanism of fatty acid biosynthesis has long been known, the architecture of the synthase has remained elusive until recently. High-resolution crystal structures of all the components in type II fatty acid biosynthesis have been obtained, including a number of enzyme/inhibitor and enzyme/substrate complexes.[28] These provide the basis for much of the work done on type I systems.

The fungal FAS was crystallised and revealed to be an $\alpha_6\beta_6$ heterododecamer with a mass of 2.6 MDa. The two chains have the domains required for fatty acid biosynthesis shared between them. The 210 kDa α chain encodes for the ACP, KR,

KS, a phosphopantetheinyltransferase (PPT)and part of the MPT domain, while the 230 kDa β chain carries the AT, ER, DH and the majority of the MPT. The structure is barrel-like, with a central wheel of six α chains capped by two domes of three β chains. The fungal FAS has a highly rigid structure with reaction chambers in each dome (Figure 1.6).[29, 30]

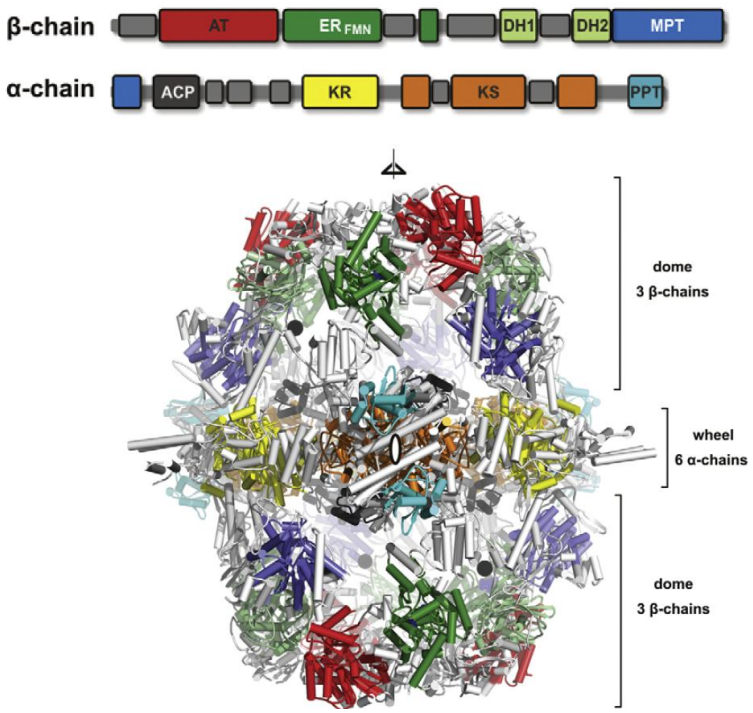


Figure 1.6: The domain organisation and structure of fungal FAS. The central wheel is made up of α-chains while the domes are formed from β-chains. The MAT domain is represented as AT. (Reprinted from *Current Opinion in Structural Biology,* Leibundgut, M.; Maier, T.; Jenni, S.; Ban, N., The multienzyme architecture of eukaryotic fatty acid synthases, 714-725, Copyright (2008), with permission from Elsevier.)[31]

The mammalian FAS (mFAS), despite carrying out an identical series of reactions to the fungal FAS, has a very different structure. mFAS exists as an X-shaped dimer split into two sections. The lower section consists of a 'condensing block' and contains the

KS and MAT domains, while the upper section is a 'modifying block' which houses the KR, DH and ER domains. There are also pseudo-KR (ΨKR) and pseudo-methyltransferase (ΨMT) domains which have no known function but appear to play an important structural role. The ACP and TE domains were not visualised in the structure, probably due to a high degree of conformational flexibility, but the anchor point for the ACP is in the centre of the upper portion of the lateral clefts of mFAS (Figure 1.7).[29, 32]

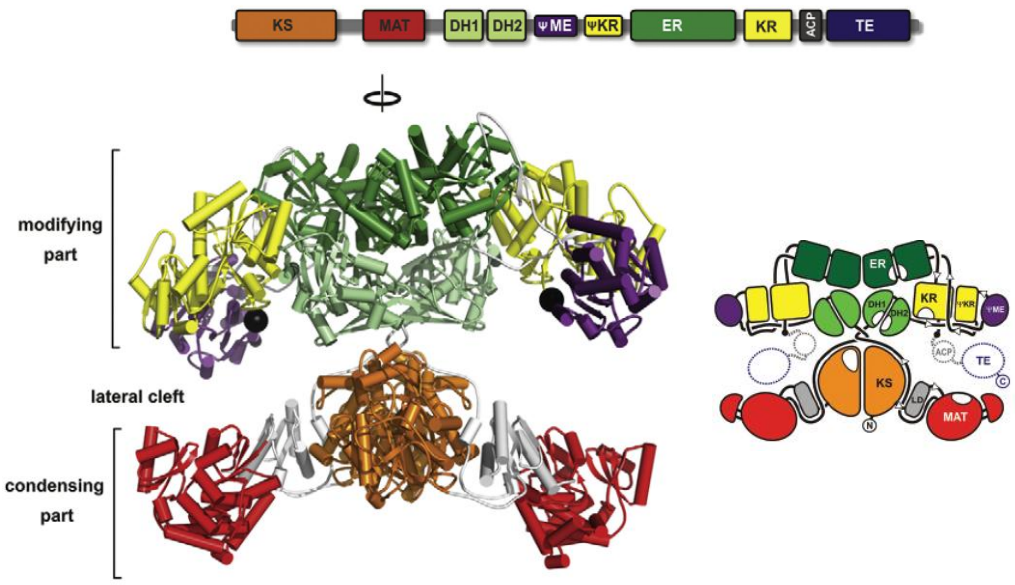


Figure 1.7: The domain organisation and structure of mammalian FAS. mFAS is a head to head dimer of two identical chains and is divided into condensing and modifiying blocks. The ACP and TE domains were not visible in the crystal structure (Reprinted from *Current Opinion in Structural Biology,* Leibundgut, M.; Maier, T.; Jenni, S.; Ban, N., The multienzyme architecture of eukaryotic fatty acid synthases, 714-725, Copyright (2008), with permission from Elsevier.)[31]

1.4 Polyketide synthases

The enzymes which assemble polyketides *in vivo*, polyketide synthases (PKSs) are analogous to fatty acid synthases. The greater structural diversity of polyketide

products comes from the fact that PKSs can vary the degree of reduction after each step. This can lead to formation of a ketone, hydroxyl, alkene or methylene functionality at C-3 in the chain after each condensation. Additional diversity comes from the fact that PKSs do not only use malonyl-CoA as an extender unit. Systems which employ methylmalonyl-CoA and methoxymalonyl-CoA are also known.[33] PKSs can employ a wide variety of starter units and also feature C-methylation domains for the introduction of branching.

1.4.1 Type I modular PKSs

Type I modular PKSs are analogous to Type I FASs in that all the domains are present on a single polypeptide. Unlike FAS, however, each domain is only used once. The domains are formed into modules which collectively perform one condensation step and associated modification of the polyketide chain before transfer to the following module.

The first known modular PKS was 6-deoxyerythronolide B synthase (DEBS) from *Saccharopolyspora erythraea*. Sequence analysis of the *S. erythraea* genome found three large open reading frames (ORFs) which coded for three giant polypeptides (approximately 350 kDa each). By sequence comparison to FAS domains, regions of the polypeptides were assigned biosynthetic functions.[34] The DEBS megasynthases function as a 'molecular assembly line', passing the growing polyketide chain from one module to the next (Figure 1.8).

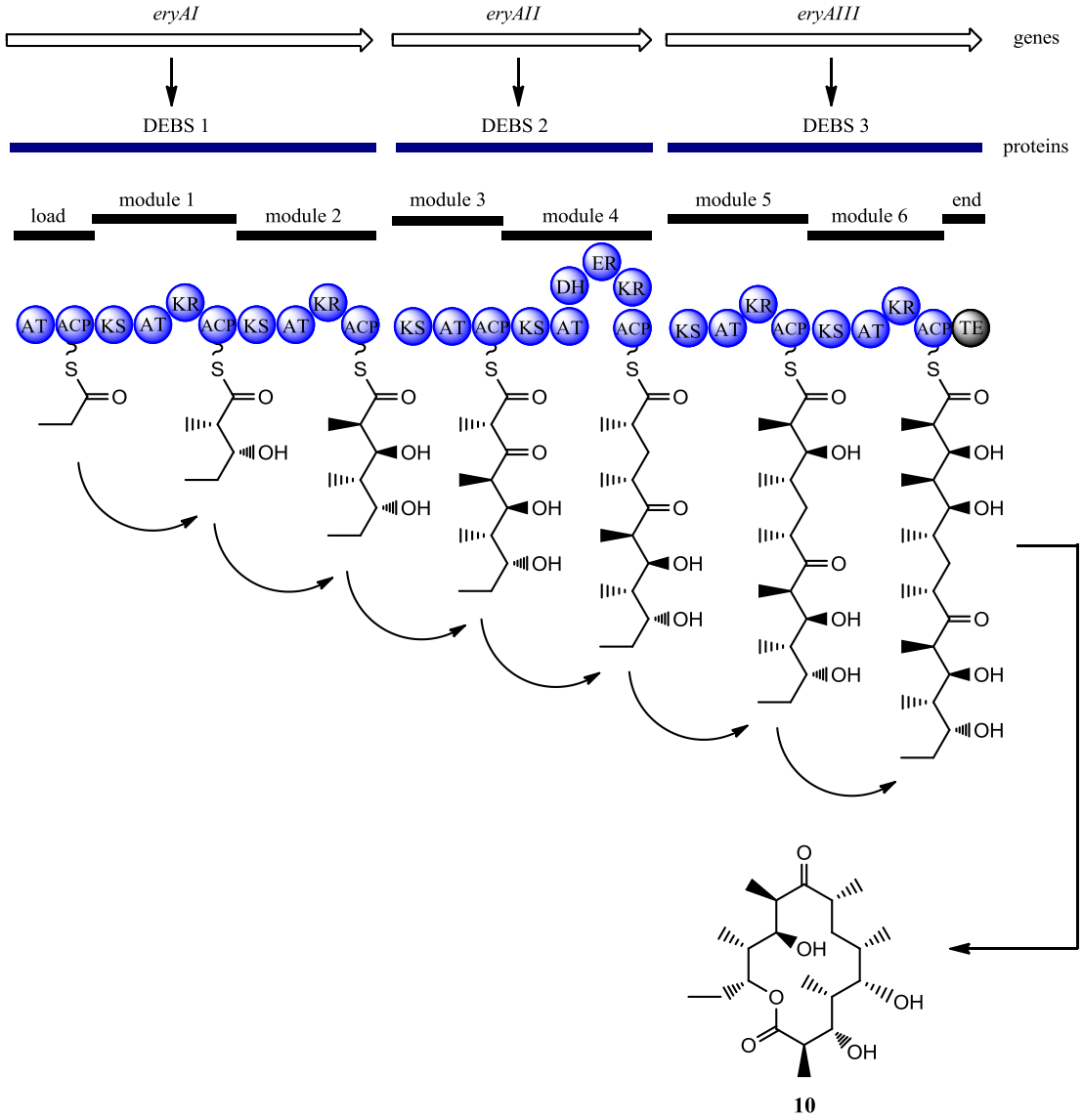


Figure 1.8: Biosynthesis of 6-deoxyerythronolide B (**10**) by the DEBS type I modular polyketide synthases.

The sequence of domains corresponds exactly to the functionality observed in the product 6-deoxerythronolide B (6-dEB, **10**). This is the principle of co-linearity, where the domains present on the synthase dictate the final structure of the product. Not all Type I modular PKSs conform to this rule. The rapamycin PKS contains modules which have KR, DH and ER domains that are not required to act to form the final product. These could be inactive and only remain to be edited out of the sequence, or they could be required for the structural integrity of the synthase. Alternatively, it

is possible that the domains are active and further oxidation happens once product is off-loaded from the PKS.

A recently discovered sub-class of the modular PKSs is the '*trans*-AT' group. These synthases have an AT domain which is supplied as a discrete protein rather than one per extension module.

Modular Type I PKSs are dimeric and are proposed to adopt the same structure as mFAS, a head-to-head, tail-to-tail dimer. This structure is considerably more complicated than the iterative mFAS since a modular PKS can contain more than one covalently linked set of modules and must also be able to interact with modules on other polypeptide chains.

1.4.2 Type I iterative PKSs

Type I iterative PKSs are mostly found in fungi and consist of a single large polypeptide with multiple domains distributed along it. Fungal PKSs use a single set of active sites iteratively, which leads to questions of how the molecular programming of these enzymes occurs which enables them to produce a single product.

Fungal PKSs can be subdivided into three classes based on their product: highly-reducing, partially reducing and non-reducing.[17]

Highly-reducing fungals PKSs, such as the lovastatin synthases LovB and LovF, yield products with a high degree of saturation. The production of lovastatin also requires a discrete ER domain, LovC. Interestingly LovB does possess a putative ER domain, but the need for LovC suggests that it is inactive.[35] It is also proposed that LovB catalyses the intramolecular Diels-Alder reaction needed to form lovastatin (Figure 1.9). Evidence suggests that this cyclisation occurs on an enzyme-bound hexaketide which is subsequently extended to the nonaketide.[36]

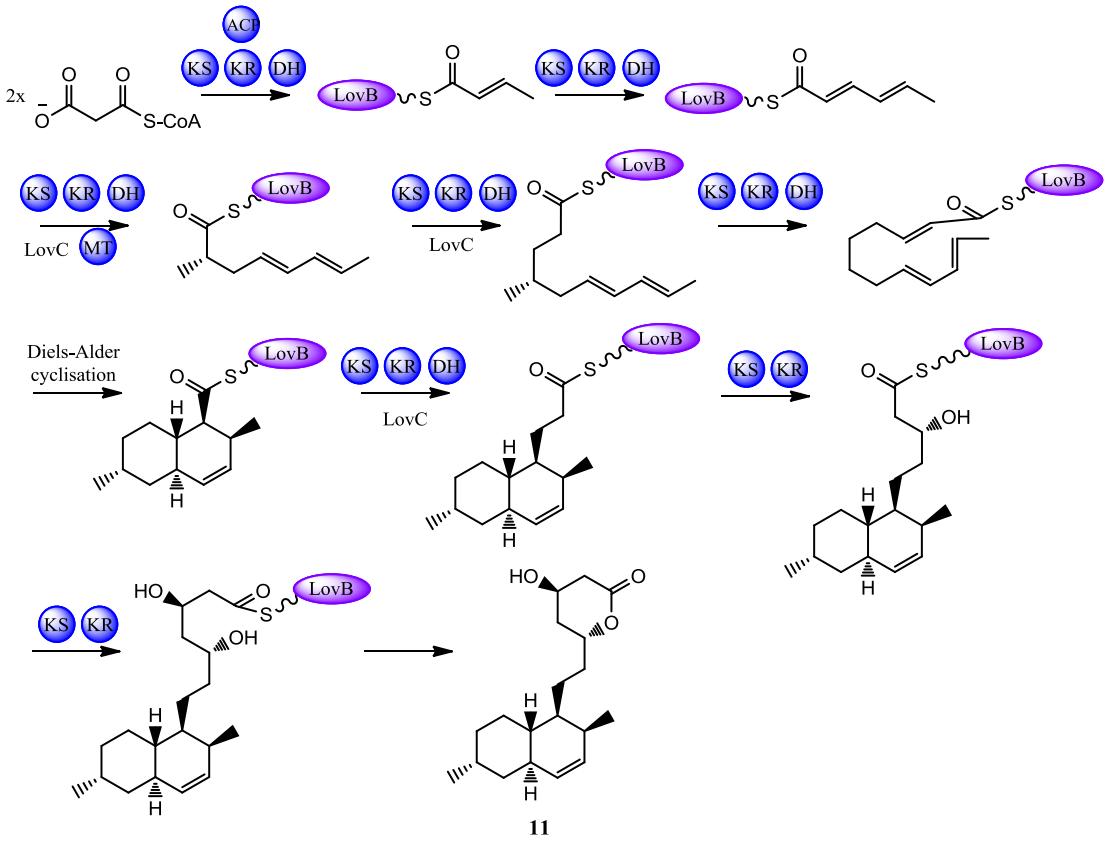


Figure 1.9: Biosynthesis of dihydromonacolin L (**11**), a precursor to lovastatin, by a highly reducing type I iterative PKS.[37]

Partially-reducing PKSs are typified by 6-methylsalcylic acid synthase (6-MSAS). This performs only one ketoreduction in three condensation cycles to form the aromatic compound 6-MSA. 6-MSAS will be further discussed in section 1.6.1.

The non-reducing PKSs form aromatic compounds such as orsellinic acid (**2**), 1,3,6,8-tetrahydroxynaphthalene and norsolorinic acid (**12**), the first intermediate on the pathway to aflatoxin B_1.[38] Norsolorinic acid synthase has been extensively studied by Townsend and co-workers (Figure 1.10). Analysis of the amino acid sequence of the synthase using the Udwary-Merski algorithm, which predicts inter-domain linker regions, found two new domains within the PKS. [39] The first was called a product template (PT) domain and is proposed to control the chain-length and folding of the growing polyketide. The PT domain has been crystallised and shown to be able to

bind both linear and bicyclic substrate mimics.[40] The second new domain was discovered at the N-terminus of the PKS and was found to be a starter unit acyl transferase (SAT). This transfers a hexanoate starter unit synthesised by a dedicated FAS onto the PKS.[41] Dissection of the norsolorinic acid synthase into individual domains and their reconstitution *in vitro* has enabled the function of each domain to be assigned.[42]

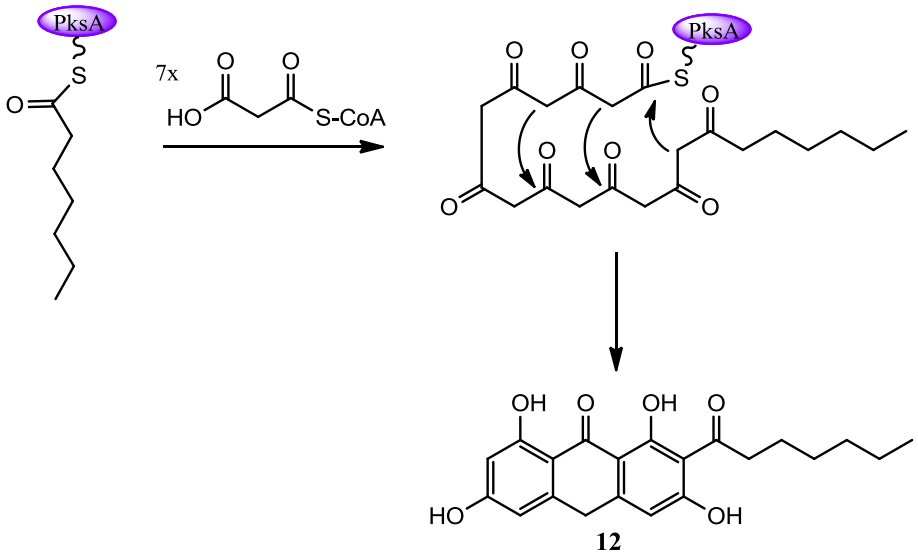


Figure 1.10: Biosynthesis of norsolorinic acid (**12**) by the non-reducing type I iterative PKS, PksA.[42]

Although all three classes of type I iterative PKSs carry out similar reactions, the makeup of their synthases are very different. HR-PKSs feature KS, AT, KR, DH, ER and ACP domains, along with a C-methyltransferase domain. NR-PKSs lack any domains from the reductive loop, but instead contain SAT and PT domains, alongside Claisen cyclase domains or thioesterase domains for off-loading. PR-PKSs have the most simple domain structure of all, containing only KS, AT, DH, KR and ACP domains along with a core domain of unknown function (Figure 1.11).[43]

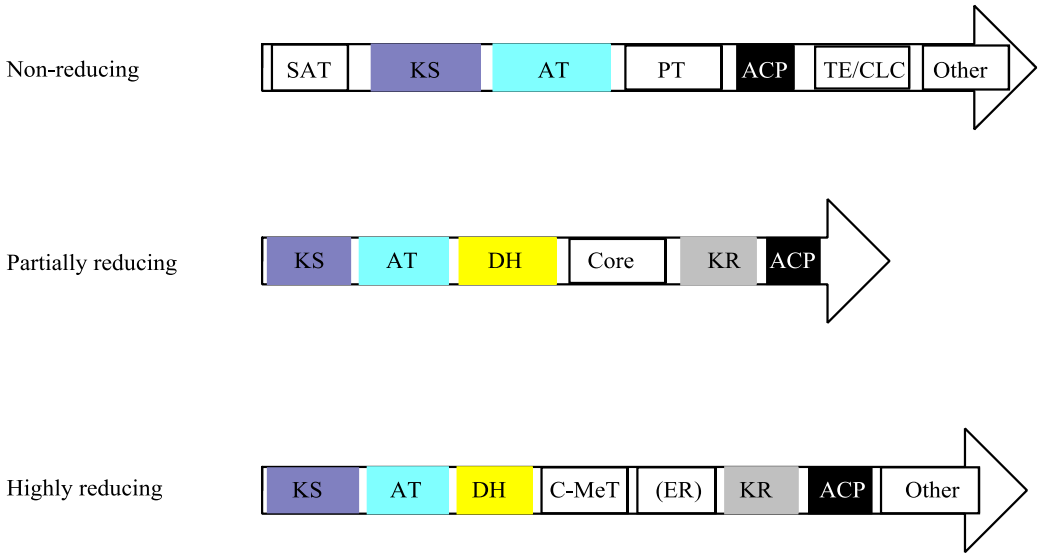


Figure 1.11: The domain architecture of the three classes of type I iterative PKS.[43]

A type I iterative PKS has also been found to synthesise polyunsaturated fatty acids (PUFAs) in marine microbes which inhabit cold environments.[44]

1.4.3 Type II PKSs

Type II PKSs, like bacterial type II FASs, are associated complexes of discrete proteins. The "minimal PKS" consists of two KS-like enzymes (KS_α and KS_β) and an ACP. KS_β has been shown to be important in controlling chain length of products and is also known as the 'chain length factor' (CLF). It is unclear whether an MCAT domain is recruited from fatty acid biosynthesis since the ACP has been demonstrated to be capable of self-malonylation *in vitro*.[45] Other proteins encoding ketoreductases, aromatases and cyclases can also act on the polyketide chain.[46]

1.4.4 Type III PKSs

Type III PKSs, like type II PKSs act in an iterative manner. Instead of the multi-enzyme complex, a single KS-like domain is used to carry out all decarboxylation, condensation, cyclisation and aromatisation reactions. Rather than utilising substrates bound to an ACP, type III PKSs act on CoA thioesters directly. Type III PKSs

such as chalcone synthase and stilbene synthase use a wide variety of acyl-CoA starter units to generate diversity and typically give mono- and bi-cyclic aromatic products.[17]

1.5 Engineering of polyketide synthases

Traditionally, the discovery of new polyketide drugs has been from one of two routes: the isolation of a new compound from its natural source or synthetic modification of a known polyketide. The improved understanding of the genetics behind polyketide biosynthesis has opened up a new avenue of investigation: the engineering of polyketide synthases to produce novel polyketides with new or improved activities.[47] There have been two complementary approaches: combinatorial biosynthesis and mutasynthesis.

1.5.1 Combinatorial biosynthesis

The modular organisation of type I PKSs has made them an attractive target for engineering. Initially it was hoped that simple module-swapping and domain-swapping experiments would provide a way to form libraries of novel polyketides. While experiments in this line were successful to some extent, the engineered PKSs had very low efficiency which made them unsuitable for high-throughput screening.[47]

Subsequent investigations have found that although domain-swapping is theoretically possible, all domains exhibit a degree of substrate specificity and have preferred partner domains with which they interact. Nevertheless, it has been possible to reduce and increase chain length, alter the degree of ketoreduction and alter the starter unit and extender unit specificity through genetic engineering.[48]

An emerging area of study is the linker regions between catalytic domains. It is proposed that these mediate the interaction of the domains and thus careful selection of the linker would be necessary to increase the efficiency of the engineered synthase. To date over 200 novel polyketides have been synthesised using the combinatorial approach. [47]

1.5.2 Mutasynthesis

Mutasynthesis is an improvement on precursor-directed biosynthesis, which added alternative starter units to cultures that produced polyketides to generate novel analogues. The mutasynthetic approach to forming novel polyketides relies on the inactivation of the PKS pathway at an early stage which removes production of the natural polyketide, usually by disrupting the biosynthesis of the starter unit. A synthetic precursor is then fed to the culture which can be taken up by the cells and attached to the PKS after the inactivated module. This precursor will then be extended as the normal substrate would have been to form an altered polyketide scaffold.[49]

This approach has been used successfully to generate novel rapamycin analogues which were formed by a strain unable to synthesise the natural 4,5-dihydoxycyclohex-1-enecarboxylic acid starter unit.[50] Mutasynthesis is also used commercially to synthesise the veterinary antiparasitic drug doramectin by addition of cyclohexanecarboxylic acid to a culture of *Streptomyces avermitilis*. The strain is engineered to be unable to synthesise the branched fatty acid starter unit normally used in the biosynthesis of avermectin.[51]

1.6 Biosynthesis of 6-methylsalicylic acid (6-MSA)

6-MSA (**1**) is a polyketide produced by a wide variety of fungi and bacteria. It is the first identifiable intermediate in the biosynthesis of patulin (**13**), which is used as an

antibiotic in veterinary medicine and is also a cause of apple spoilage in cider manufacture (Figure 1.12).

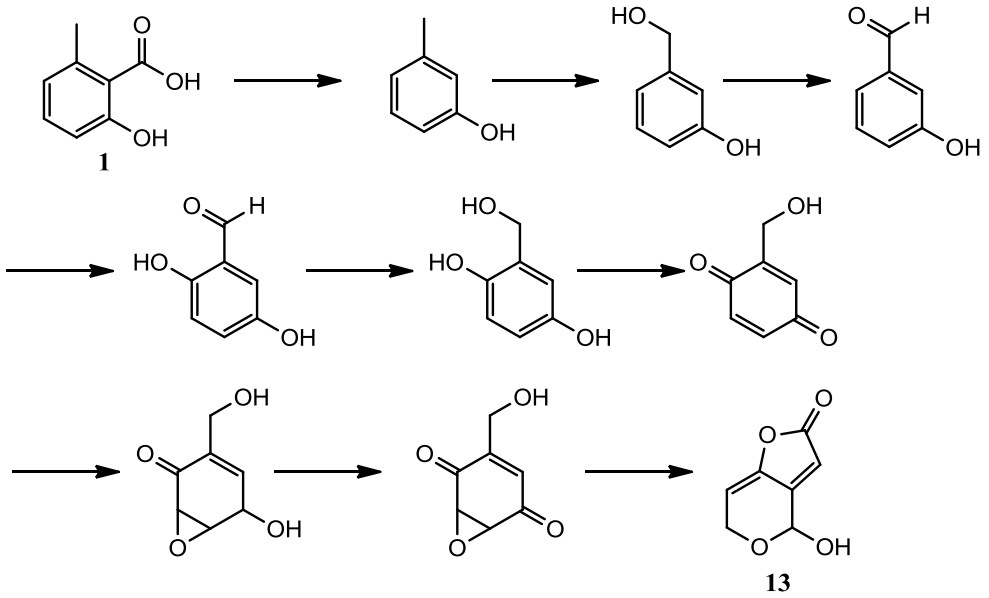


Figure 1.12: The biosynthesis of patulin (**13**) from 6-MSA (**1**).[52]

6-MSA was found to be made up of four acetate units through the feeding studies of Birch.[24] Later studies found that the rate of incorporation of ^{14}C-acetate into 6-MSA in *P. patulum* was decreased when 6-MSA or structurally related aromatic acids were added to the culture medium, suggesting a degree of product inhibition in 6-MSA biosynthesis.[53]

1.6.1 6-Methylsalicylic acid synthase

6-MSA is assembled from one unit of acetyl-CoA (**14**) and three units of malonyl-CoA (**9**) by 6-methylsalicylic acid synthase (6-MSAS).

Initial purification of 6-MSAS from *P. patulum* was attempted using ammonium sulphate fractionation, followed by sedimentation in an ultracentrifuge and sucrose density-gradient centrifugation. This gave enzyme extract of 100 times the purity of

cell free extract and an approximate molecular weight of 1.1 – 1.5 MDa.[54] Later purification followed much the same model, but found that the inclusion of glycerol improved activity.[55] 6-MSAS was finally purified to homogeneity by ammonium sulphate fractionation, followed by poly(ethylene glycol) precipitation, DEAE-Sepharose ion-exchange chromatography, hydroxyapatite chromatography before a final ion-exchange chromatography step. 6-MSAS was found to consist of a single type of subunit of molecular weight 180 kDa. Gel filtration analysis confirmed that the enzyme existed as a tetramer with a total molecular weight of approximately 750 kDa, which was significantly lower than had previously been determined.[56]

The 6-MSAS gene from *P. patulum* was identified by immunological screening of a genomic DNA expression library. The gene was found to be contained in a single reading frame which was 5322 base pairs long with a single 69 base pair intron in the N-terminal part of the synthase. 6-MSAS was calculated to be a protein made up of 1774 amino acid with a predicted molecular mass of 190 731 Da (Figure 1.13). Comparison of the amino acid sequence with known FASs and PKSs found little similarity between 6-MSAS and the yeast and *Penicillium* FASs but 6-MSAS did show much higher similarity to the rat FAS. Homology to previously sequenced PKSs and FASs allowed the identification of KS, AT, DH, KR and ACP domains within the synthase. Interestingly no thioesterase domain for product release was detected within the synthase.[57]

Figure 1.13: The domain organisation of 6-MSAS. The single polypeptide contains KS, AT, DH, KR and ACP domains.

Investigations into the substrate specificity of 6-MSAS found that while acetyl-CoA was the preferred starter unit, propionyl-CoA was also accepted at 1/13th of the rate of acetyl-CoA to form 6-ethylsalicylic acid.[58] 6-MSAS also accepts acetoacetyl-CoA as

a starter unit at 80% of the rate of acetyl-CoA, to form 6-MSA.[56] A variety of starter units were subsequently found to be incorporated into the corresponding alkylsalicylic acid, including butyryl-CoA and crotonyl-CoA, but branched alkyl chains were not accepted by the synthase.[52]

The 6-MSAS has been successfully heterologously expressed in *Eschericia coli*, *Saccharomyces cerevisiae*[59] and *Streptomyces coelicolor*.[60] In each of these systems production of 6-MSA was observed, revealing that no other genes were required for the biosynthesis. In particular, no discrete thioesterase from *P. patulum* was required to release 6-MSA from the synthase. Other type I iterative PKSs with no thioesterase domain release their products using cyclisation or reductive domains.

Heterologously expressed 6-MSAS has been used in metabolic engineering studies in yeast in which the aim was to maximise polyketide production by modification of the phosphopantetheinyltransferase[61] and acetyl-CoA carboxylase genes.[62] 6-MSAS has also been transformed into the tobacco plant where it provided some resistance against tobacco mosaic virus, although not as effectively as salicylic acid.[63]

Before crystal structures of any PKS or FAS were available for comparison, the structure of 6-MSAS was investigated by cross-linking and limited proteolysis studies. The thiol-modifying cross-linking agent 1,3-dibromopropanone was found to inactivate 6-MSAS by selectively modifying the reactive cysteine in the KS domain and the thiol on the end of the phosphopantetheine arm attached to the ACP. Cross-linked 6-MSAS dimers were observed. Incubation of 6-MSAS with acetyl-CoA prevented both inactivation and cross-linking, by protecting the thiol in the KS domain, while incubation with malonyl-CoA prevented cross-linking but not inactivation. These results suggest that the KS of one subunit is linked to the ACP of another.[56, 64] Studies with the thiol-binding reagent cerulenin found that this bound specifically to the reactive cysteine in the KS.[65] These investigations gave results which were analogous with those achieved in investigations of mammalian FAS, and so it is implied that there are great structural similarities between 6-MSAS and mFAS.

1.6.2 Mechanism of 6-MSA biosynthesis

Experiments with partially purified 6-MSAS found that triacetic acid lactone (TAL) was produced in the absence of nicotinamide adenine dinucleotide phosphate (NADPH).[54] When the purified synthase was incubated with acetyl-CoA and malonyl-CoA with no NADPH it was found that TAL was formed as the exclusive product at 10% of the rate of 6-MSA biosynthesis.[56] TAL is also seen as a shunt product in fatty acid biosynthesis in the absence of NADPH.[66]

Incubation of 6-MSAS with 3-pentynoyl-N-acetylcysteamine, an inhibitor which reacts with any dehydratase activity, completely stopped 6-MSA production. Oxidation of NADPH only fell to 60% of the rate seen without inhibitor, confirming that the reduction step was unaffected.[55] These results suggest that the ketoreductase domain acts on a six-carbon intermediate and the resulting alcohol is then dehydrated to from a double bond (Figure 1.14). For the formation of 6-MSA, a *cis* double bond is required to allow cyclisation of the tetraketide intermediate, but polyketide dehydratases generally from *trans* double bonds. It is possible that a *trans* bond is formed which is then isomerised to the *cis* or dehydration could happen on a cyclised reduced tetraketide intermediate. The mechanistic considerations underlying the dehydration step are discussed in section 3.1.2.

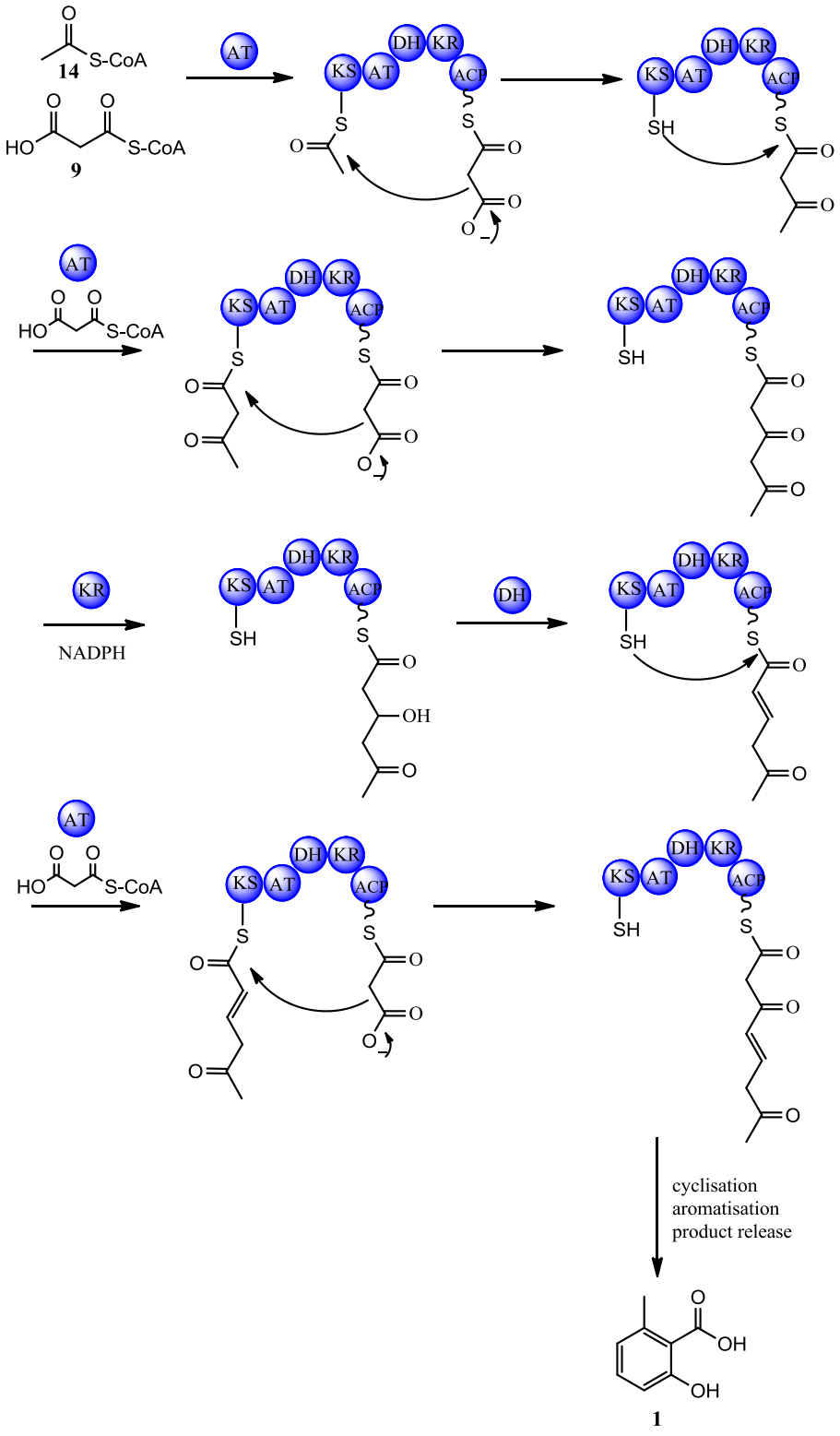


Figure 1.14: The proposed biosynthetic pathway to 6-MSA from acetyl-CoA (**14**) and malonyl-CoA (**9**), as catalysed by 6-MSAS.

Initial experiments using tri-deuteriated acetate found that there was high degree of retention of deuterium at chain building sites, but that the incorporation was not uniform for all sites in 6-MSA.[67] Further experiments with both tri-deuteriated acetate and mono-deuteriated acetate found that deuterium retention at C-3 and C-5 was consistent with the removal of hydrogen at these postions being stereospecific. As the hydrogens at C-3 and C-5 will be removed during the aromatisation sequence, this suggests that the formation of the aromatic ring happens under enzymatic control.[68]

A series of elegant experiments involving chiral malonates was performed by Spencer and Jordan to determine which hydrogen atoms were removed in the formation of the aromatic ring. (R)- and (S)-[1-^{13}C, 2-^{2}H] malonate were converted to malonyl-CoA using succinyl-CoA transferase. As the transferase was unable to distinguish between the two acid groups in malonate, each chiral malonate reaction gave rise to two chiral malonyl-CoAs. Incorporation of these malonyl-CoA mixtures would give eight possible labelling patterns in the final 6-MSA. Four possible mechanisms were suggested and the appropriate labelling patterns for 6-MSA produced from each malonyl-CoA mixture were calculated. The mass spectra of 6-MSA obtained after incubation with both the (R)- and (S)-[1-^{13}C, 2-^{2}H] malonate-derived malonyl-CoAs with acetyl-CoA as a starter unit was compared with the predicted spectra. It was determined that the hydrogen atoms at C-3 and C-5 in 6-MSA arose from hydrogen atoms with opposite stereochemistries in the polyketide chain.[69] To determine the absolute stereochemistry of the hydrogen retained at each position, unlabelled acetoacetyl-CoA was used as a starter unit. This gave an unlabelled C-3 in 6-MSA and allowed the labelling at C-5 to be probed directly. Comparison of the predicted and obtained spectra of 6-MSA after incubations with (R)- and (S)-[1-^{13}C, 2-^{2}H] malonate-derived malonyl-CoAs showed that the hydrogen retained at C-5 came from the *pro*-(R) hydrogen of malonyl-CoA. This meant that the hydrogen retained at C-3 came from the *pro*-(S) hydrogen of malonyl-CoA (Figure 1.15). [70]

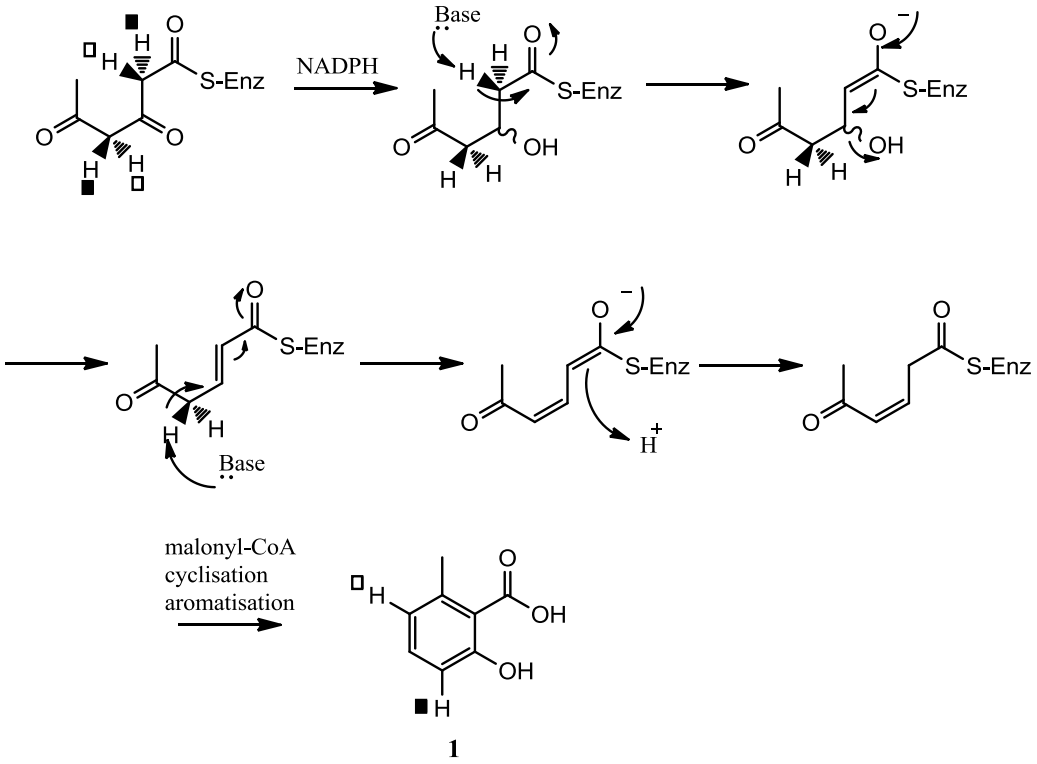


Figure 1.15: One possible mechanism to explain the stereochemistry of hydrogen atom removal: Removal of hydrogens from the same face of the C_6 intermediate to form a *cis* double bond.[52]

As the 6-MSAS gene contains no thioesterase domain, the mechanism of product release is unclear. It is possible that 6-MSA is hydrolysed by an activated water molecule which can be deprotonated by a base situated on the enzyme. Alternatively, a mechanism involving the formation of a ketene has been proposed. The ketene could then be hydrated to form 6-MSA (Figure 1.16).[56]

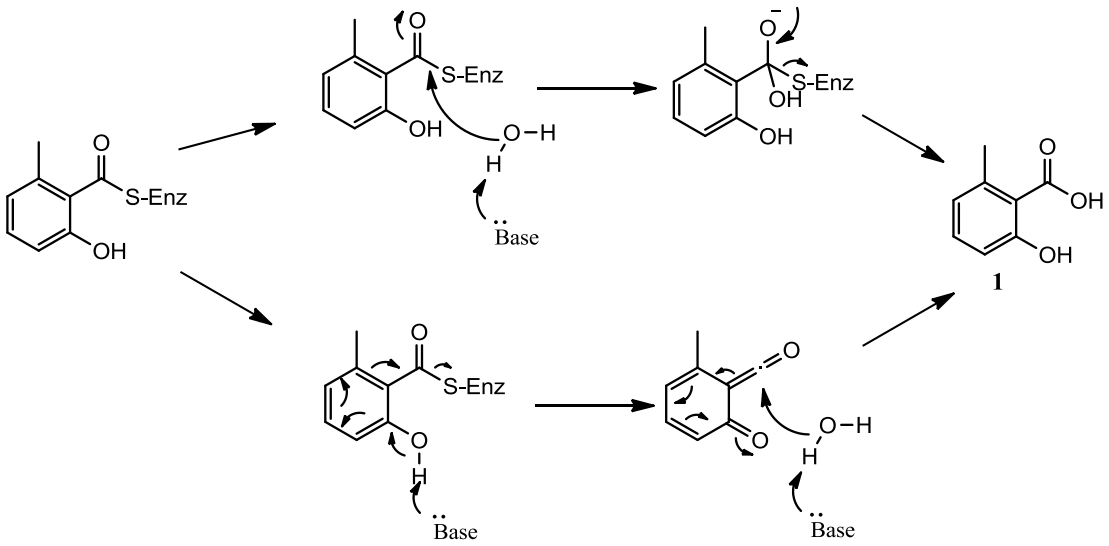


Figure 1.16: Product release from 6-MSAS could occur via a deprotonated water molecule or a ketene intermediate.

Although investigations into 6-MSA biosynthesis have spanned over 50 years, there are still many unanswered questions about the formation of this deceptively simple polyketide. The timing of the dehydration, cyclisation and aromatisation steps remains unclear, as does the mechanism of product offloading. The stereochemistry of the hydroxyl formed by the ketoreductase has also not been determined. This thesis is concerned with applying new methods to this well-studied and archetypal system, in order to shed some light on these questions.

CHAPTER 2

Off-loading of intermediates from 6-MSAS

2.1 Introduction

During polyketide biosynthesis, intermediates remain covalently bound to the polyketide synthase. This presents a problem in establishing the exact biosynthetic pathway, since the identity of enzyme-bound intermediates is difficult to establish. While biosynthetic intermediates are often proposed based on the domain structure of the polyketide synthase and the observed structure of the product, the timing of events such as cyclisation or epoxidation can remain cryptic. In order to examine these enzyme-bound intermediates, a variety of techniques have been employed.

2.1.1 Methods of identifying intermediates in polyketide biosynthesis

2.1.1.1 Labelled precursors

Initial insights into polyketide biosynthesis were gained through the use of radiolabelled ^{14}C-acetate. Incorporation of the radioactive precursor was assayed through a complex series of degradation reactions to known products. Birch was the first to employ this method in his investigation of 6-methylsalicylic acid production.

The radiolabelled 6-methylsalicylic acid (6-MSA, **1**) formed was degraded and the fragments analysed to determine the relative amounts of radiation each contained (Figure 2.1). It was found that ^{14}C was uniformly incorporated at four sites. These corresponded to the sites predicted by Birch in his hypothesis that the biosynthesis proceeded via the formation of a polyketone precursor which underwent a single reduction event before cyclising through an aldol reaction and formation of an aromatic ring.[24]

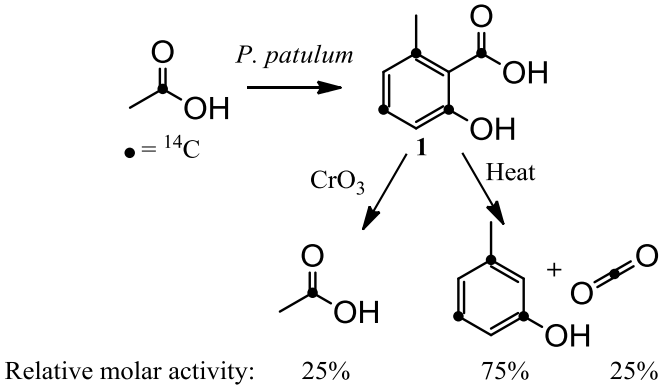


Figure 2.1: Labelling pattern of 6-MSA (**1**) from ^{14}C acetate feeding experiments by Birch and subsequent degradation reactions.

The development of NMR shifted focus from ^{14}C-labelled precursors to ^{13}C-labelling. Singly labelled ^{13}C acetate allowed the pattern of labelling in a polyketide to be quickly established without the need for labour-intensive degradation studies. Doubly labelled ^{13}C acetate greatly added to the depth of knowledge of the field, as for the first time it was possible to observe the incorporation of intact acetate units was able to be measured. This allowed some of the complex structural rearrangements that often occur after the formation of the polyketide backbone to be investigated, for example in the biosynthesis of aflatoxin.[23]

The putative diketide intermediate (**15**) was synthesised by Cane *et al*. The precursor was incorporated into both erythromycin (**4**) and nargenicin A_1 (**16**) retaining both the deuterium and ^{13}C labels, showing that reduction and elimination steps took place after each condensation reaction in a modular PKS, rather than on the

assembled polyketone chain (Figure 2.2).[71] This is referred to as a processive mechanism.

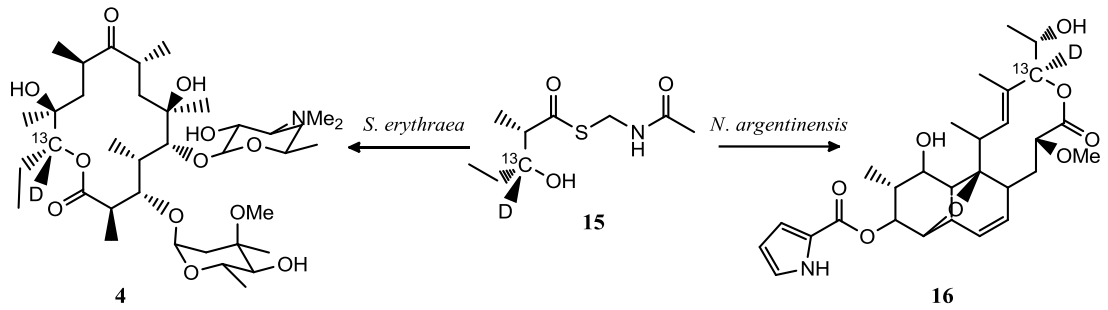


Figure 2.2: Incorporation of the doubly labelled diketide-SNAC (**15**) into erythromycin (**4**) and nargenicin A_1 (**16**) through feeding experiments.

2.1.1.2 Movement of thioesterase domains

In Type I modular polyketide synthases, intermediates can be observed by moving the thioesterase domain to generate a truncated product. In the 6-deoxyerythronolide B synthases (DEBS), the thioesterase (TE) domain (which is normally at the end of the third megasynthase protein, DEBS3) was attached to the end of the first megasynthase, DEBS1 to give DEBS1-TE.[72] This formed a triketide lactone, presumably due to off-loading of the intermediate by TE-catalysed cyclisation. The off-loaded intermediate could then be easily characterised by GC-MS and NMR experiments.

Movement of the TE domain has also enabled the substrate channelling of DEBS modules to be probed using synthetic diketide N-acetylcysteamine (NAC) thioesters to form triketide lactones which can then be analysed.[73] Heterologously expressed recombinant domains were used to establish the stereospecificity of each ketoreductase (KR) domain in DEBS. Incubation of a diketide-NAC ester with methylmalonyl-CoA, an acyl transferase-ketosynthase (AT-KS) didomain, a KR domain and an acyl carrier protein (ACP) led to the production of ACP-bound

triketide intermediates, which were then off-loaded using base hydrolysis or a TE to form triketide lactones.[74]

2.1.1.3 Mass spectrometry

Recently, the development of highly-sensitive mass spectrometers has led to the direct sampling of enzyme-bound intermediates. Using Fourier-transform mass spectrometry (FTMS) Hicks *et al.* were able to detect thioester intermediates bound to EpoC, a polyketide synthase (PKS) involved in epothilone biosynthesis. The starter unit was presented as a NAC-thioester which was loaded on to the synthase and condensed with one unit of methylmalonyl-CoA leaving the product bound to the ACP. EpoC was then digested using a limited proteolysis strategy and fractionated by liquid chromatography before being analysed using FTMS. This was compared to the limited proteolysis chromatograph of the *holo* enzyme and the change in mass of the fragment containing the ACP and phosphopantetheine arm was used to assign the identity of the reaction product.[75]

An alternative strategy has been to utilise the facile ejection of the phosphopantetheine arm under MS^2 conditions to directly assay intermediates. The advantage of this strategy is that it does not require digestion of the intact PKS or NRPS. Instead, the protein is subjected to tandem mass spectrometry (MS/MS) using methods such as collision activated dissociation and infrared multiphoton dissociation to induce ejection of the phosphopantetheine arm as one of two fragments (**17** and **18**, Figure 2.3). These fragments can then be detected directly and the intermediates attached to the thioester characterised.[76] This 'top down' approach has recently been adapted for low-resolution instruments using the characteristic fragmentation patterns of the ejected ion under MS^3 conditions.[77]

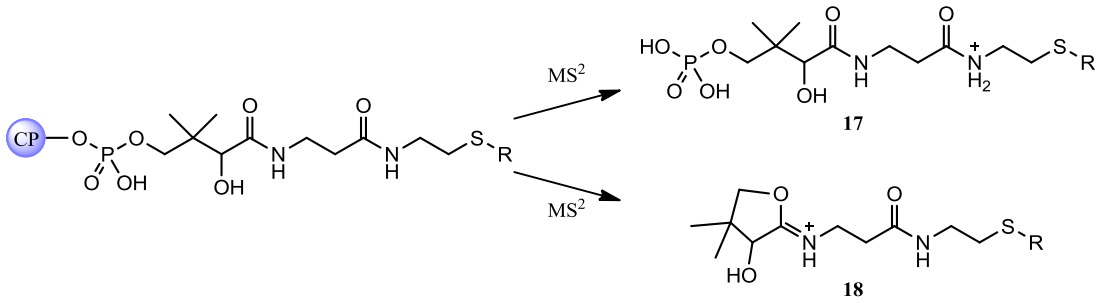


Figure 2.3: The phosphopantetheine arm attached to the carrier protein (CP) is ejected under MS^2 conditions to give both the phosphopantetheine (**17**) and pantetheine (**18**) fragments. R is an intermediate bound as a thioester.

Mass spectrometry has also been used to investigate starter unit specificity in non-ribosomal peptide synthetase (NRPS) systems. The NRPS is incubated with a pool of potential starter amino acids and the correct one is loaded by the adenylation domain. Tandem mass spectrometry on the intact synthase leads to ejection of the phosphopantetheine ion attached to the cognate amino acid.[78]

In the biosynthesis of aflatoxin, a toxin produced by the fungus *Aspergillus parasiticus*, the first discrete precursor is norsolorinic acid, which is synthesised by an iterative PKS, PksA. Norsolorinic acid is formed from a hexanoyl starter unit and seven units of malonyl CoA and the resulting unreduced polyketone chain is cyclised to give three aromatic rings. Crawford *et al.* used the Udwary-Merski algorithm[39] to predict domain boundaries and dissect PksA into its discrete domains. The domains were expressed as separate proteins and recombined *in vitro* and FTMS used to observe intermediates attached to the ACP. Omission of domains, and analysis of the different intermediates produced, allowed the function of each domain to be assigned.[42] The mass spectrometry data also showed that the unreduced polyketone chain was not cyclised during the rounds of chain extension. This was also shown by Ma *et al.* in their study of the interaction between a non-reducing PKS from *Gibberella fujikuroi* and the discrete ketoreductase domain from the actinorhodin Type II PKS. It was shown that the KR performed a reduction exclusively at the C-9 position and so must have access to the complete ACP-bound polyketone.[79]

2.1.2 Non-hydrolysable analogues for off-loading intermediates

The techniques, while successful in their respective systems, either lack a general applicability for all types of polyketide synthase or require access to expensive equipment. It was proposed in our research group that a common feature of all polyketide systems could be exploited to probe enzyme-bound intermediates; namely, the use of malonyl-CoA or its derivatives as an extender unit. It was hypothesised that a malonyl-CoA unit which could decarboxylate and form a carbon-carbon bond with the enzyme-bound intermediate, but which was not itself enzyme bound, could provide a way to off-load intermediates from a megasynthase. To ensure that intermediates were not simply reloaded by reaction with another enzyme-bound malonyl unit, the off-loading agent would have to feature a non-hydrolysable bond in place of the thioester (Figure 2.4).

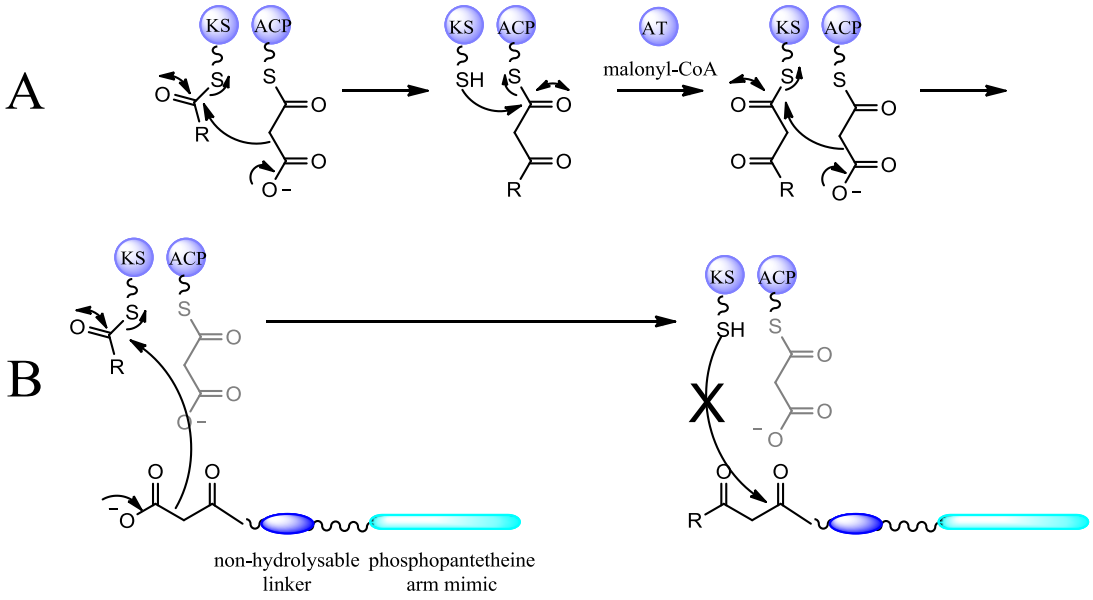


Figure 2.4: A: Polyketide biosynthesis. Decarboxylation of the ACP-bound malonate unit is followed by condensation with the acyl unit on the KS, leading to an extended ACP-bound polyketide chain. The chain is then transferred back to the KS and the ACP is loaded with another unit of malonate. B: Offloading of intermediates. Decarboxylation of the malonyl-unit on the non-hydrolysable analogue is followed by condensation with the acyl unit on the KS leads to off-loading of the KS-bound intermediate. The non-hydrolysable linker does not allow the extended chain to be reloaded onto the enzyme.

Initial studies were performed by Dr Dieter Spiteller and Claire Waterman using a non-hydrolysable analogue (**19**).[80] This features a malonyl unit for reaction with intermediates, but the thioester has been replaced by a thioether. This had the disadvantage of introducing an additional methylene group, which might have interfered with molecular recognition, but this was outweighed by the facile synthesis of **19** in two steps. This analogue was tested with stilbene synthase (STS), a Type III PKS from *Pinus sylvestris*. Both diketide (**23**) and triketide (**22**) intermediates were observed by LC-MS when 4-hydroxyphenylacetyl-CoA (**20**) was used as a starter unit, but no intermediates were detected when cinnamoyl-CoA (**21**) was used (Figure 2.5). This was unexpected as cinnamoyl-CoA is usually elaborated to the tetraketide product pinosylvin. It was proposed that this lack of reactivity was due to the

misalignment of the cinnamoyl-CoA and malonyl-CoA in the KS active site. The 4-hydroxyphenylacetyl unit is a methylene unit shorter, and so may accommodate the longer malonyl analogue more easily.

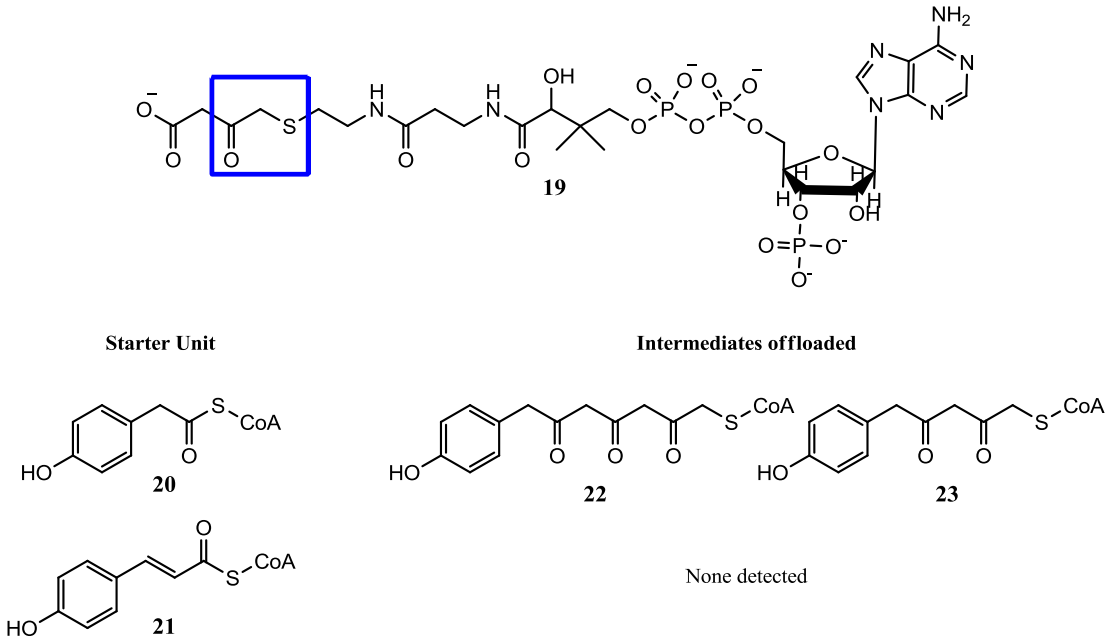


Figure 2.5: Non-hydrolysable malonyl CoA analogue **19** was used to off-load diketide (**23**) and triketide (**22**) intermediates with 4-hydroxyphenylacetyl-CoA (**20**) as a starter unit.[80]

In order to have a better mimic of malonyl-CoA, alternative non-hydrolysable analogues (**24**) were synthesised by Dr Manuela Tosin[81]. These replaced the sulphur of the thioester group with an O or CH_2 to produce a bond that was either less prone or inert to hydrolysis and would be almost isosteric with malonyl-CoA (Figure 2.6). In trials with STS, the oxa(dethia)malonyl-CoA was able to off-load only the starter unit to form a diketide attached to CoA. In contrast, the carba(dethia) analogue was able to off-load diketide, triketide and tetraketide intermediates, including some cyclised and/or dehydrated tetraketides for a variety of starter units. The increased activity of the carba(dethia) analogue over the oxa(dethia) is proposed to be due to its more rapid decarboxylation and because hydrolysis of off-loaded intermediates is not possible.

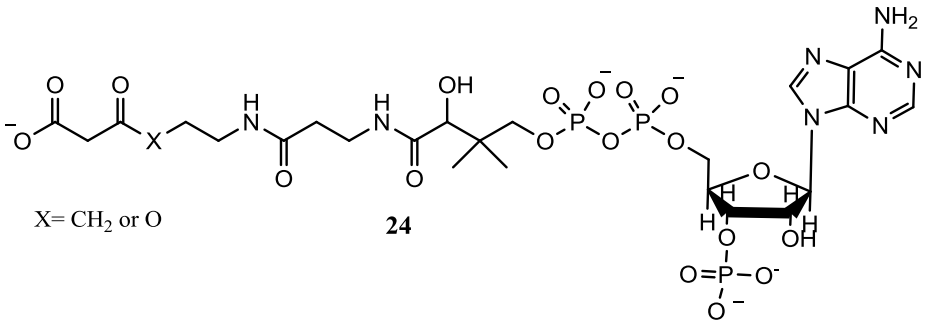


Figure 2.6: Non-hydrolysable malonyl CoA analogue (**24**) synthesised by Tosin *et al.* used to off-load intermediates from stilbene synthase.

Shorter analogues of the non-hydrolysable malonyl-CoA were then investigated.[82] Pantetheine analogues (**27**) were synthesised with either NH, CH_2 or O in place of the thioester sulphur atom, while a carba(dethia)methylmalonyl-NAC analogue (**28**) was also synthesised. The system used was DEBS3, the final polyketide megasynthase used in the biosynthesis of erythromycin. This system normally processes a pentaketide to a heptaketide, but in the absence of the other DEBS proteins it produces triketide lactones (**25** and **26**) from propionyl-CoA as a starter unit followed by two condensations with methylmalonyl-CoA as an extender unit.[83] For both the pantetheine and N-acetylcysteamine analogues, inhibition of triketide lactone production was observed, however no off-loaded intermediates were detected with the pantetheine analogues. By contrast, both malonyl- and methylmalonyl-analogues (**30** and **28**) based on carba(dethia)-N-acetylcysteamine off-loaded intermediates that could be detected successfully. The advantage of the NAC analogues over the coenzyme A analogue is twofold: firstly, the synthesis is greatly simplified; secondly, the off-loaded intermediates are hydrophobic enough to be extracted into organic solvent, allowing the sample to be concentrated before analysis. Using the NAC analogues, off-loaded diketide (**29** and **31**) and triketide (**32**) intermediates were observed for DEBS3 when the synthase was incubated with propionyl-CoA and methylmalonyl-CoA (Figure 2.7).

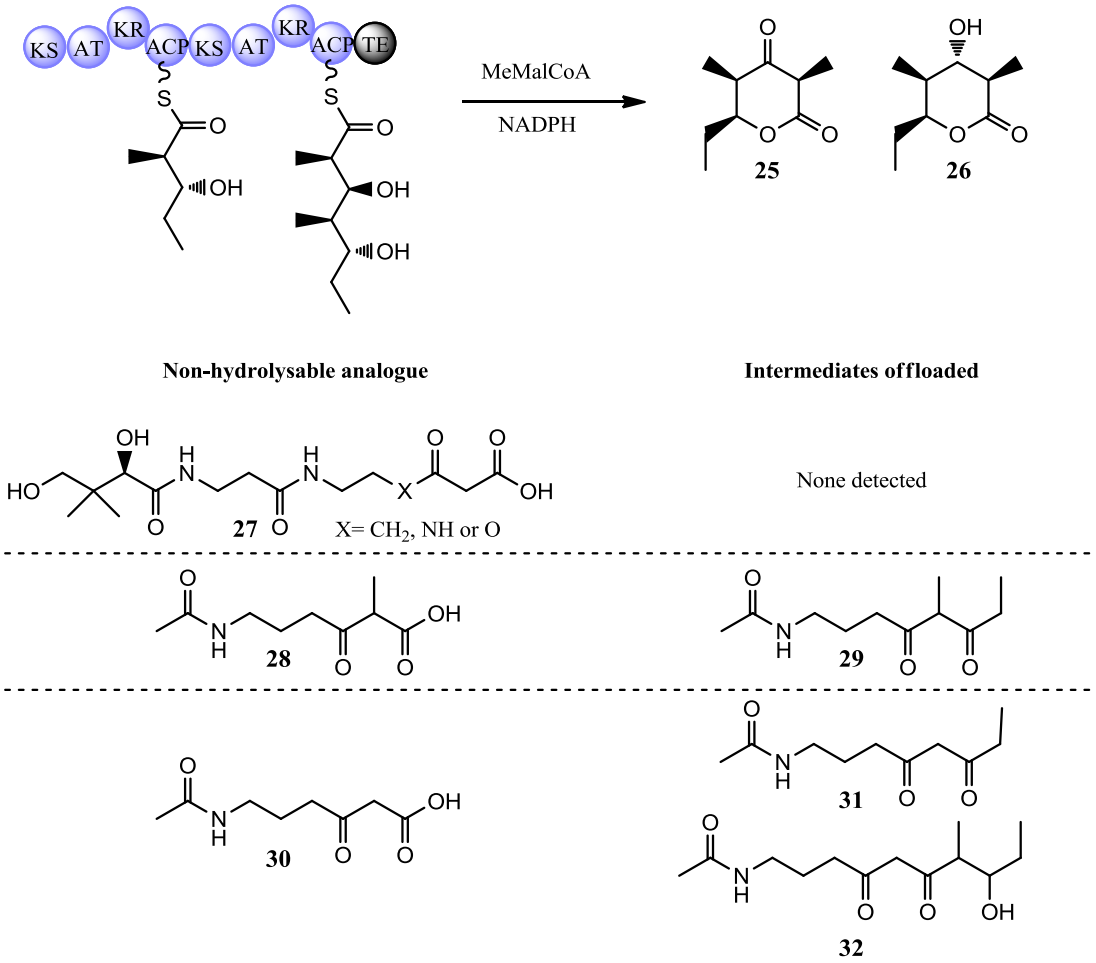


Figure 2.7: Non-hydrolysable pantetheine (**27**) and N-acetylcysteamine (**28** and **30**) analogues were used to off-load intermediates in the biosynthesis of triketide lactones by the megasynthase DEBS 3.

Recent *in vivo* work has been performed on the lasalocid-producing PKS which has had the last two ACPs inactivated by mutation of the phosphopantetheine arm attachment site. Off-loading of late stage intermediates could be observed when non-hydrolysable analogues were fed (Dr M Tosin, personal communication).

2.1.3 Probing intermediates in an iterative synthase

Iterative synthases present a greater challenge to the traditional methods for studying intermediates as they contain only a single set of active sites, so any

mutation in the KS or ACP active sites would destroy the synthase's biosynthetic capacity altogether.

6-methylsalicylic acid biosynthesis has already been extensively studied using labelled acetate[68] and labelled malonate,[69] but these have not helped in the deduction of later stage intermediates, particularly in the timing of the dehydration and cyclisation steps.

2.1.4 Aims

The aim of this section is to use a non-hydrolysable malonyl-NAC analogue to off-load intermediates in the biosynthesis of 6-MSA. This would represent the first use of this methodology on a Type I iterative synthase and could also provide insights into the biosynthesis of this deceptively simple polyketide.

2.2 Results and discussion

2.2.1 Expression of *holo* 6-MSAS

The expression of 6-methylsalicylic acid synthase (6-MSAS) proved very challenging. Initial attempts to express the synthase under standard expression conditions (LB media, 37 °C until $OD_{600} \sim 1$, then induction with 0.4 mM IPTG overnight at 16 °C) did not lead to any high molecular weight bands on SDS-PAGE being seen after ammonium sulphate purification. The method of Kealey *et al.*[59] was used (ATCC 765 minimal media plus 10% glycerol, 37 °C until $OD_{600} \sim 1$, then induction with 0.5 mM IPTG at 30 °C for 24 hours) , followed by ammonium sulphate precipitation, as used by Spencer and Jordan,[56] gave high molecular weight bands. The cell pellet could not be frozen before ammonium sulphate precipitation, as this resulted in the loss of the high molecular weight bands by SDS-PAGE. After a year of unsuccessful attempts, the following method was used to obtain active *holo* 6-MSAS.

Plasmids containing the gene coding for 6-methylsalicylic acid synthase (6-MSAS) (pKOS12-65a) and its expression vector (pKOS007-95) were donated by Kosan Biosciences Ltd. These were digested and the fragments ligated to form the expression plasmid pKOS007-109. This plasmid was transformed into *Escherichia coli* BL21 (DE3) alongside pET28a(+)-MtaA, a plasmid expressing the phosphopantetheinyltransferase (PPTase) MtaA from *Stigmatella aurantiaca*[84] with an N-terminal His_6-tag donated by Dr Fanglu Huang. The co-expression of a PPTase should ensure that the phosphopantetheine arm that is essential for activity is attached *in vivo*.

Previous experiments had found that 6-MSAS did not bind to a nickel affinity column when expressed with N-terminal, C-terminal or both N- and C-terminal His_6-tags (Dr Fanglu Huang, personal communication). In order to avoid any interference with activity from non-natural tags, it was decided to express 6-MSAS in an untagged form and use an alternative method of purification.

The *E. coli* was cultured under the same conditions as previously used for heterologous expression of 6-MSAS by Kealey *et al.*[59] They had found that although expression of the synthase in *Saccharomyces cerevisiae* gave a greater yield of 6-MSA, expression in *E. coli* gave a higher percentage of 6-MSAS in the total cell proteins.

Holo 6-MSAS was purified by ammonium sulphate precipitation, as by Spencer *et al.* for the enzyme from its native producer, *P. patulum.*[56] It was found that 5 mM EDTA and the presence of the protease inhibitors benzamidine and phenylmethylsulphonyl fluoride (PMSF) in the cell resuspension buffer were vital for obtaining active synthase. The 22-40% ammonium sulphate pellet was then either used directly for assays or further purified by FPLC using a Sephadex 200 column and the buffer system used by Richardson *et al.* (Figure 2.8).[85]

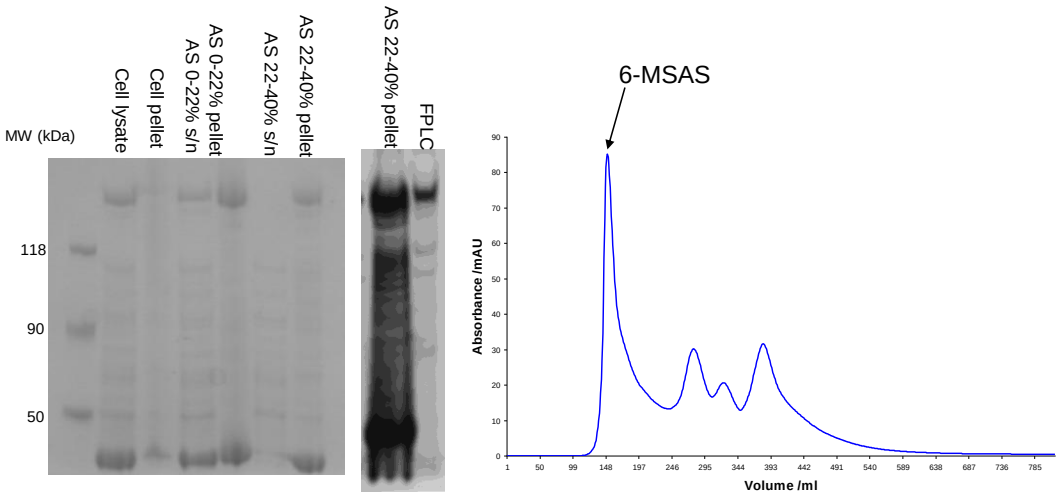


Figure 2.8: SDS-PAGE analysis of ammonium sulphate fractions in the purification of 6-MSAS and gel filtration trace of 6-MSAS.

FPLC purification yielded protein which showed a strong single band of the correct approximate molecular weight (190 kDa calculated) by SDS-PAGE, with faint bands of lower molecular weight. These are probably due to degradation of the protein during purification. This method gave approximately 10 mg of *holo* 6-MSAS per litre of culture.

2.2.2 Activity of *holo* 6-MSAS

The activity of both *holo* and *apo*-6-MSAS was assayed using the fluorimetry method of Spencer and Jordan[56], which monitors the production of 6-MSA directly. The ammonium sulphate pellet was redissolved in assay buffer with 10 mM DTT and incubated in the fluorimeter at 25 °C with acetyl CoA, NADPH and bovine serum albumin (BSA). The reaction was started by addition of malonyl-CoA and the production of 6-MSA was monitored (λ_{ex} = 310 nm, λ_{em} = 390 nm). The specific activity was calculated by comparison with a standard curve of 6-MSA fluorescence. Protein concentration was determined by Bradford assay.

Ammonium sulphate-precipitated *holo* 6-MSAS without EDTA in the resuspension buffer gave a specific activity of 15 m-units/mg protein, while *holo* 6-MSAS resuspended in the presence of EDTA had a specific activity of 65 m-units/mg protein (1 unit is 1 μmole of 6-MSA per minute). This compares extremely favourably with the values obtained by Spencer and Jordan for the native enzyme in *P. patulum*, which had a specific activity of 5.8 m-units/mg protein at the same stage of purification. The low specific activity of the native enzyme was due to a co-purifying FAS.

FPLC purified fractions gave a specific activity of 108 m-units/mg protein under the same assay conditions. This is less than half the specific activity recorded for the purified native synthase of 245 m-units/mg protein. However, this purification method benefits greatly from a simple two-step procedure rather than the more intensive five steps that were previously used. It is, however, comparable to the specific activity observed after four purification steps (135 m-units/mg).

2.2.3 Production of 6-MSA

The production of 6-MSA *in vivo* was assayed using LC-MS. 50 ml of culture supernatant from a 1 litre culture of *holo* 6-MSAS was acidified to pH 1 using hydrochloric acid, then extracted three times using ethyl acetate. The extracts were dried over magnesium sulphate and the solvent removed *in vacuo*. The remaining solid was redissolved in 1 ml of methanol and analysed by LC-MS. The UV traces at 280 nm and 254 nm were monitored and the retention time of 6-MSA was compared to a synthetic standard (Figure 2.9). A peak with a mass of 153 units which corresponded to $[M+H^+]$ was observed.

The UV peak area was used to compute the amount of 6-MSA produced by comparison to the standard peak area.

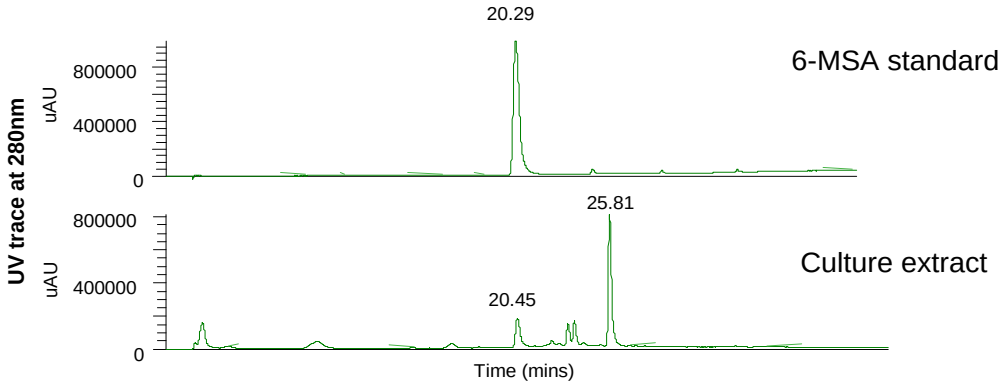


Figure 2.9: LC-MS traces of *holo* 6-MSAS culture extract compared to authentic 6-MSA.

Holo 6-MSAS was calculated to produce 17-25 mg of 6-MSA per litre of culture.

2.2.4 Synthesis of non-hydrolysable analogue

In order to off-load intermediates from 6-MSAS, carba(dethia) malonyl-*N*-acetylcysteamine (**30**) was used since this had been successfully employed in off-loading from DEBS3. The non-hydrolysable analogue (**30**) was synthesised by a procedure designed by Emily Valentine and Dr Manuela Tosin.[86] GABA (**31**) was reacted with acetyl chloride in the presence of triethylamine to form **32**. This was then coupled to Meldrum's acid using EDC and DMAP in THF to give the cyclic adduct **33**. **33** was opened using dry methanol in toluene under reflux conditions to give the methyl ester **34** which was purified by preparative HPLC. The methyl ester was cleaved using pig liver esterase in 6-MSAS assay buffer (Tris-sulphate, pH 7.6) at 37 °C for two hours to yield the active trapping agent **30** (Figure 2.10). As **30** could readily decarboxylate, the protein was precipitated from the reaction mixture using chloroform and the crude aqueous layer was used directly in all off-loading assays.

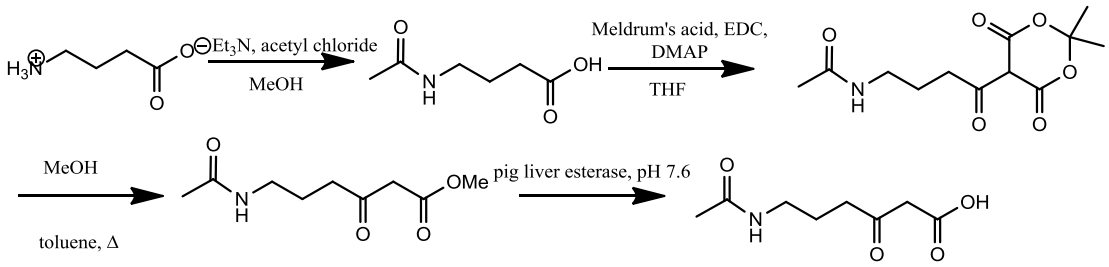


Figure 2.10: The synthesis of non-hydrolysable analogue (**30**) from GABA (**31**).

2.2.5 Starter unit assay: acetyl-CoA

To test the ability of the non-hydrolysable analogue to off-load intermediates from an iterative synthase, it was decided to remove the need for the malonyl analogue to compete with the natural extender unit, malonyl-CoA. By incubating the analogue with acetyl-CoA (the natural starter unit) and 6-MSAS, the acetyl unit should be loaded onto the KS domain by the AT, before being off-loaded by the analogue. This would result in only a single diketide product (**35**) to detect (Figure 2.11).

FPLC-purified *holo* 6-MSAS was incubated overnight at room temperature with acetyl-CoA and carba(dethia)malonyl-NAC (**30**). Concentrations of analogue from 5 mM to 30 mM were used, while concentrations of the other reagents were take from the DEBS3 experiments of Tosin *et al.* (acetyl-CoA 3 mM, 200 μg protein in 100 μl reaction).[82]

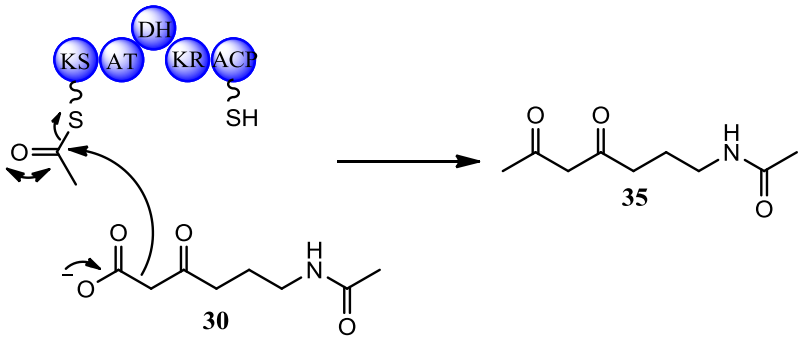


Figure 2.11: Off-loading of acetyl starter unit with non-hydrolysable analogue **30** to form diketide **35**.

Assay mixtures were extracted three times with an equal volume of ethyl acetate which was then evaporated under nitrogen. The residue was then dissolved in a minimal volume of methanol before being analysed on an LTQ-Orbitrap mass spectrometer by Dr Manuela Tosin.

The expected diketide product (**35**) was seen at a retention time of 11.83 mins, with an exact m/z of 186.1126 for all three concentrations of analogue tested (Figure 2.12). This peak did not appear in enzyme-free, analogue-free or acetyl-CoA free controls. MS/MS analysis of this peak gave fragments at m/z = 168.1016 units (loss of water) and m/z = 126.0913 (loss of acetic acid).

Incubation with NADPH did not alter the amount of diketide detected, so it was excluded from further assays. Extraction of the assay mixture after acidification to pH1 with hydrochloric acid also had no effect on the amount of diketide seen. As an additional control, the assay was performed with [$^{13}C_2$]acetyl-CoA. This gave a peak at a retention time of 12.0 minutes but with an increase of two mass units to 188.12, confirming the identity of the diketide (**36**).

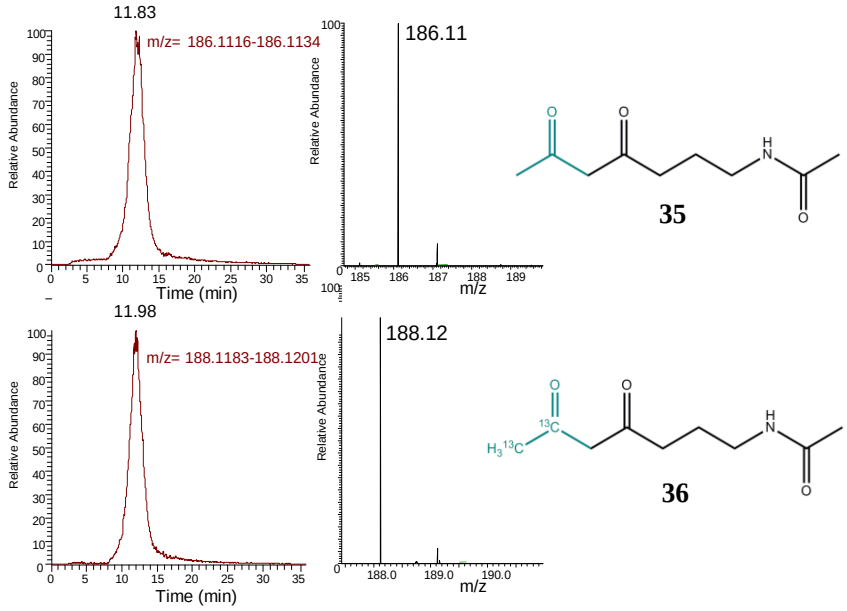


Figure 2.12: HRMS traces of extracts from reactions where 6-MSAS was incubated with acetyl and [$^{13}C_2$]acetyl-CoA and non-hydrolysable analogue **30** to form offloaded diketides (**35** and **36**).

2.2.6 Starter unit assay: acetoacetyl-CoA

6-MSAS can also accept acetoacetyl-CoA as a starter unit in place of acetyl-CoA. Incubating the enzyme with acetoacetyl-CoA and non-hydrolysable analogue should lead to the formation of an off-loaded triketide intermediate (**37**).

This triketide is a potential substrate for the KR domain, which is postulated to act on the enzyme-bound C6 intermediate *in vivo*. If the assay mixture contains NADPH and the KR is able to act on the off-loaded triketide (**37**), then the reduced triketide (**38**) would be formed (Figure 2.13). It has already been reported that 6-MSAS can reduce the model intermediate ethyl 3,5-diketohexanoate in the presence of NADPH.[55]

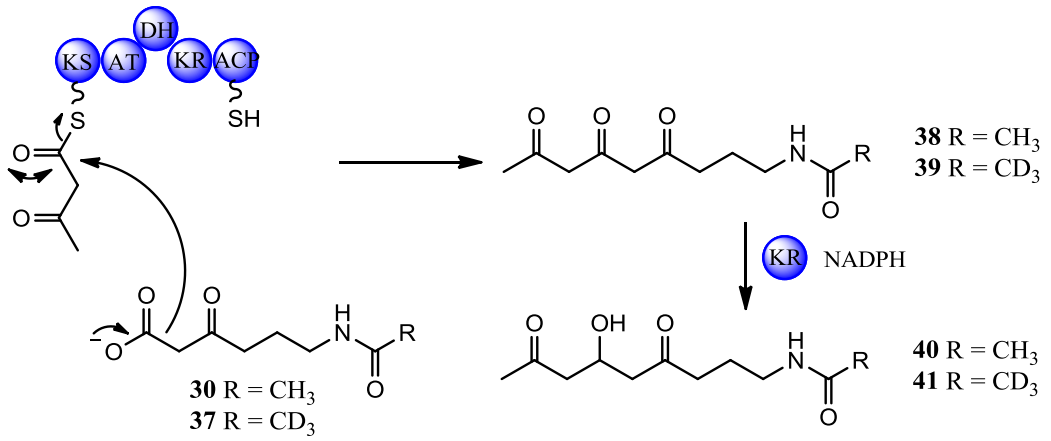


Figure 2.13: Off-loading of acetoacetyl starter unit leads to triketide **38** which can then be further reduced by 6-MSAS to reduced triketide **40** in the presence of NADPH.

FPLC-purified *holo* 6-MSAS was incubated with acetoacetyl-CoA and carba(dethia)malonyl-NAC (**30**) overnight at room temperature. The assay mixture was extracted as before and analysed on an LTQ Orbitrap.

Triketide (**38**) could be detected in the assay mixture with a retention time of 13.67 minutes and an exact mass of 228.1227 units. Repeating the assay using carba(dethia)malonyl-NAC which contained a deuterated methyl group (**37**, donated by Dr Manuela Tosin) gave a peak at the same retention time (13.67 minutes) but with an increase in mass of three units to 231.1417 units (Figure 2.14). MS/MS analysis of the 228.1227 ion gave a peak at 210.1118 units, corresponding to a loss of water from the parent ion, and a peak at 168.1016, corresponding to a loss of acetic acid. MS/MS analysis of the 231.1417 ion, corresponding to the deuterated product, gave a peak at 213.1310 units, which shows the same loss of water observed for the unlabelled molecule.

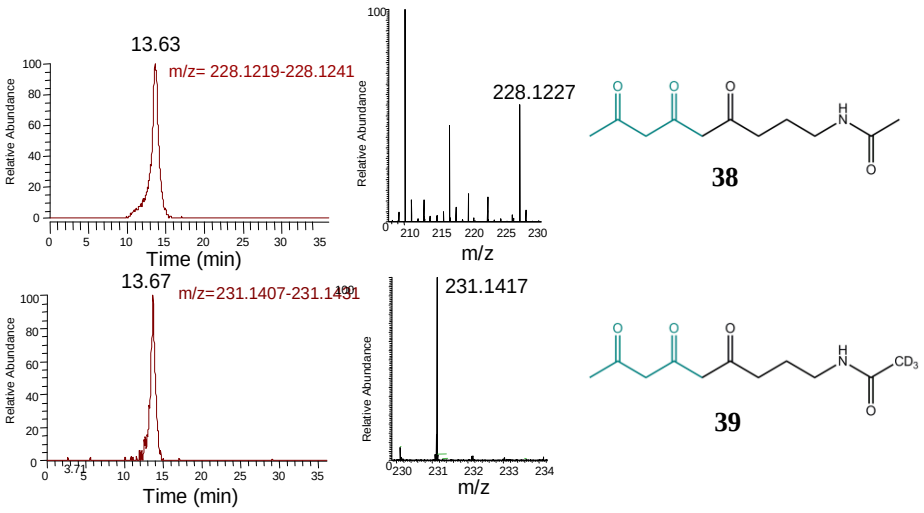


Figure 2.14: HRMS data for extracts of reactions where 6-MSAS and acetoacetyl-CoA were incubated with non-hydrolysable analogue **30** or **37** showed off-loading of triketides **38** or **39**, respectively.

When NADPH was included in the assay mixture, a peak of mass 230.1578 units and retention time 12.38 minutes was observed which did not appear in any of the controls. This mass corresponded to the reduced triketide **40**. A small amount of unreduced triketide was also visible. Incubation with the deuterated non-hydrolysable analogue **37** gave rise to a peak at retention time 12.41 minutes with a corresponding increase in mass to 233.1574 units (Figure 2.15). MS/MS of the peak at 230.1578 units gave rise to a peak at 170.1173 units, which corresponds to loss of acetic acid from the parent molecule.

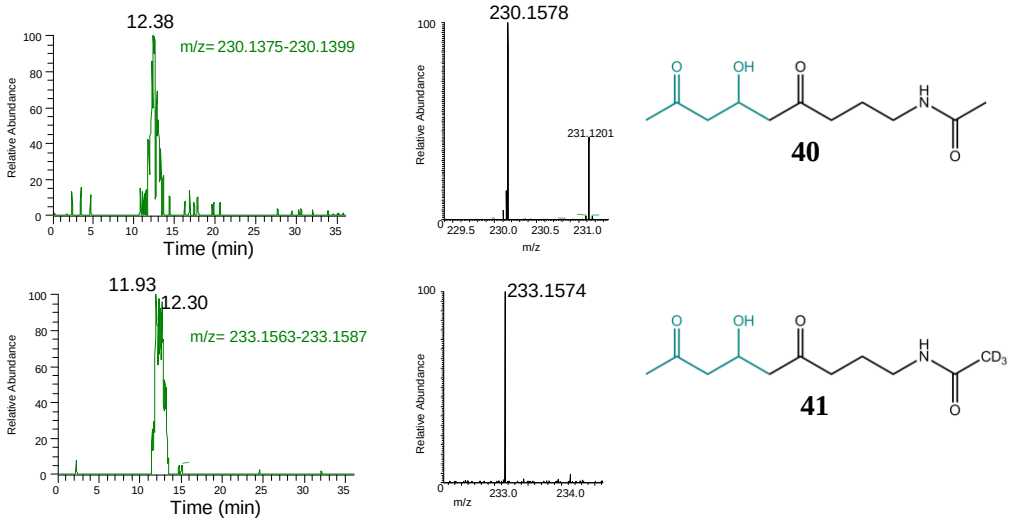


Figure 2.15: HRMS traces of extracts of reactions when 6-MSAS was incubated with acetoacetyl-CoA, NADPH and non-hydrolsable analogue **30** or **37** for detection of the reduced triketides **40** or **41**, respectively**.**

A mass search for the reduced triketide mass minus water yielded two peaks with mass of 212.1392 units, one at 12.4 minutes and one at 14.5 minutes. These peaks were not present in either the acetoacetyl off-loading assay with no NADPH or in control assays (Figure 2.16). For the acetoacetyl assay with no NADPH, a peak corresponding to the mass of triketide with loss of water is visible at the same retention time as the intact triketide, but there are no additional peaks. This suggests that the peak at 12.4 minutes is an artefact of mass spectrometry on the intact reduced triketide, but that the peak at 14.5 minutes may be a genuine dehydrated intermediate. This could be formed by the action of the 6-MSAS ketoreductase on the off-loaded triketide, followed by the 6-MSAS dehydratase domain to form a carbon-carbon double bond. The position of the double bond in the intermediate would be unclear as there are two ketones with which it could be conjugated leading to two possible products (**42** and **43**, Figure 2.16).

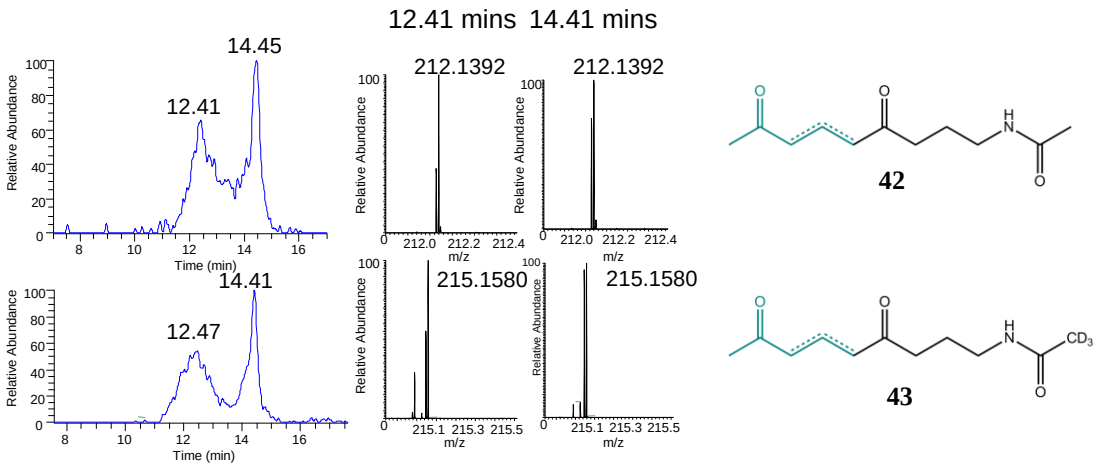


Figure 2.16: HRMS data showing a mass search for reduced triketide with loss of water for non-hydrolysable analogues without label (**42**) and with deuterium label (**43**).

The assay off-loading acetoacetyl using the deuterated analogue (**37**) in the presence of NADPH shows the same pair of peaks when the reduced triketide minus water mass is searched for. Both of these peaks have a mass of 215.1580, corresponding to a 3 mass unit increase due to the deuterated methyl group. This confirms that both peaks are products of a reaction involving the off-loading agent.

MS/MS analysis of the two peaks at 212.1392 units gave rise to an identical fragmentation pattern, with peaks at 194.1173 (loss of water) and 170.1174 (loss of acetic acid). Although this does not confirm the identity of the peaks or any differences between them, it does provide further evidence that they are both compounds which behave in a way consistent with being off-loaded intermediates.

2.2.7 Competitive assays

Theoretically, the non-hydrolysable malonyl-NAC analogue should be able to compete with malonyl-CoA as an extender unit, and so off-load intermediates from every stage in the biosynthesis of 6-MSA (Figure 2.17).

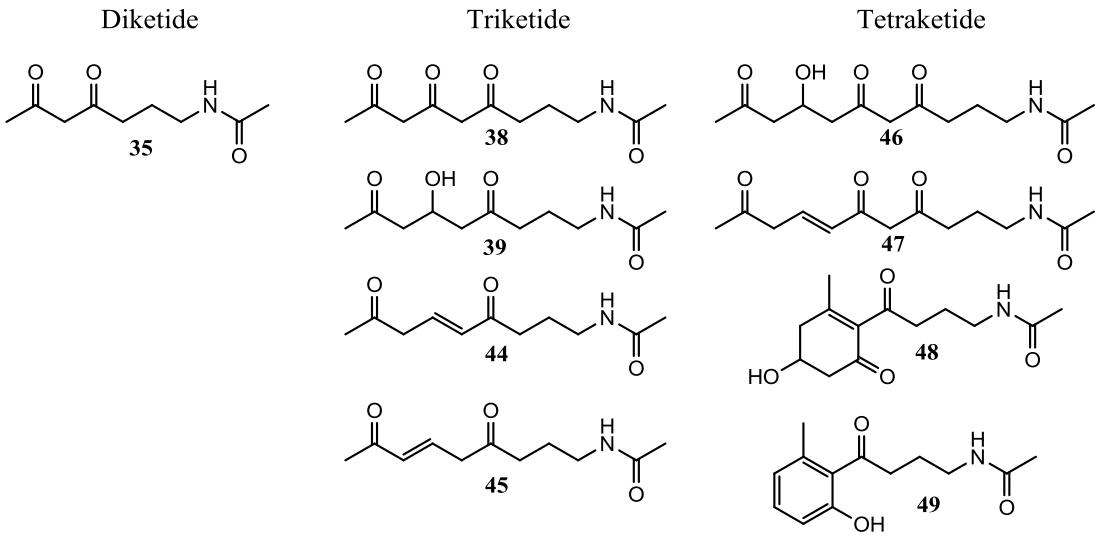


Figure 2.17: Potential off-loaded intermediates in 6-MSA biosynthesis using non-hydrolysable analogue **30**. Theoretically, the isolation of diketide, triketide and tetraketide intermediates could be possible.

Fluorimetry assays with non-hydrolysable analogue, acetyl-CoA, malonyl-CoA and NADPH showed no significant decrease in rate, but the final amount of 6-MSA produced is decreased (Figure 2.18), which is consistent with the off-loading of intermediates. The observation that the rate is not changed suggests that the off-loading occurs before the rate-determining step for 6-MSA formation. Quantitatively, the amount of 6-MSA produced in all cases exceeds the theoretical maximum value, which could be due to interfering fluorescence in the assay mixture. This does, however, make it difficult to determine the reason for the reaction stopping. The reaction could stop when the supply of malonyl-CoA is exhausted, however it is also possible that the reaction stops due to inactivation of the enzyme by the non-hydrolysable analogue. This could be tested by the addition of fresh enzyme or fresh reactants to the assay mixture and observing the fluorimetry traces. If the addition of fresh reactants causes production of 6-MSA to start again, then the amount of malonyl-CoA could be the limiting reagent, but if fresh enzyme is needed, then the enzyme is becoming inactivated by the non-hydrolysable analogue.

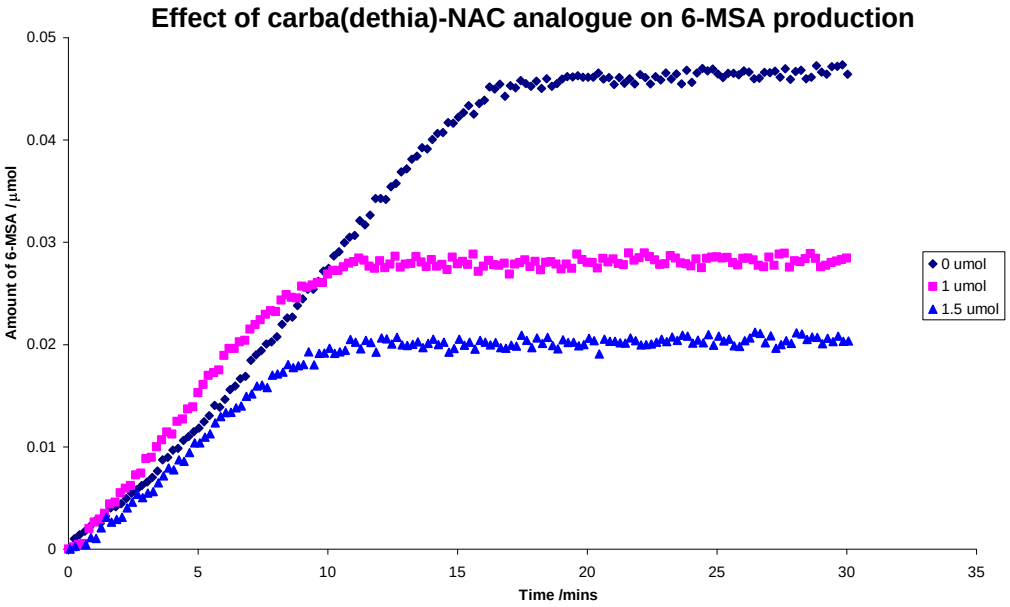


Figure 2.18: Zero-corrected fluorimetry traces of 6-MSAS activity assays (containing 6-MSAS (20 μg ammonium sulphate pellet), acetyl-CoA (0.013 μmol), malonyl-CoA (0.04 μmol) and NADPH (0.013 μmol), BSA (50 μg) in 0.5 ml 6- MSAS assay buffer, pH 7.6, 25 °C) incubated with 0, 1 and 1.5 μmol non-hydrolysable analogue.

2.2.7.1 Acetyl starter unit

Non-hydrolysable analogue **30** was incubated overnight with *holo* 6-MSAS, NADPH, acetyl-CoA and malonyl-CoA. The concentration of malonyl-CoA was varied from 4.9 mM to 12.7 mM for the acetyl assay, while other concentrations were held constant (3 mM starter unit, 10 mM NADPH, 15 mM analogue). The assays were worked up as described previously and analysed on the LTQ Orbitrap for the presence of off-loaded intermediates.

For the acetyl starter unit, both off-loaded diketide (**35**) and triketide (**38**) intermediates were observed (Figure 2.19) but no reduced triketide (**39**) was detected. As the concentration of malonyl-CoA in the assay was increased, the amount of off-loaded intermediates detected fell, suggesting that the analogue did not compete effectively with malonyl-CoA.

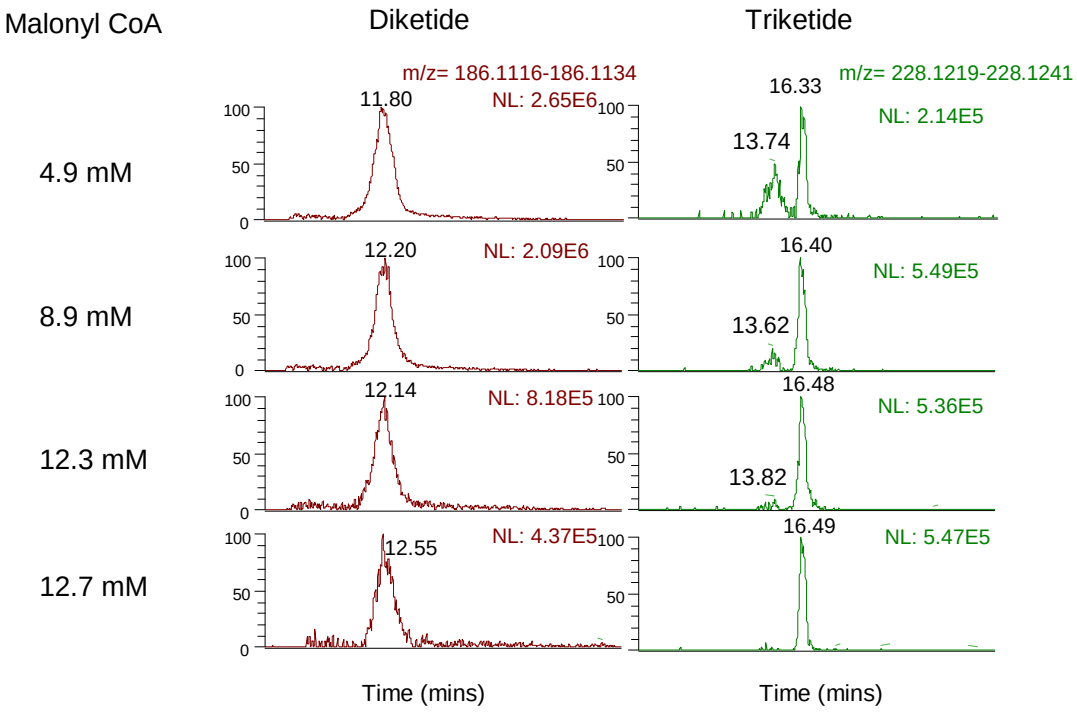


Figure 2.19: HRMS traces of extracts from competitive assays with 6-MSAS using acetyl-CoA as a starter plus malonyl-CoA as an extender unit and **30** as the offloading agent, showing diketide (11.8-12.5 mins) and triketide (13.6-13.8 mins) products for different concentrations of malonyl-CoA. The peak at 16.3-16.5 minutes appears in control samples (FPLC-purified 6-MSAS (100 μg, treated with DTT), **30** (15 mM), acetyl-CoA (3 mM) and malonyl-CoA (1-13 mM) made up to a final volume of 100 μl with 6-MSAS assay buffer, pH 7.6, room temperature, overnight).

In one single assay using 8.9 mM malonyl-CoA, a potential tetraketide peak was seen at 16.40 minutes with a mass of 276.1207 units. This corresponds to the sodium adduct of reduced tetraketide with loss of water. MS/MS analysis of this peak yielded a fragment at 192.0988 units which corresponds to a molecular formula of $C_9H_{15}NaNO_2^+$. Repeat of this experiment with and without a deuterium-labelled analogue did not show this peak again.

From the results of the competitive assays using acetyl-CoA as a starter unit, it was clear that lower concentrations of malonyl-CoA led to better off-loading of

intermediates. Thus, for the acetoacetyl starter unit, malonyl-CoA concentrations of 1-8.9 mM were used.

2.2.7.2 Acetoacetyl starter unit

Acetoacetyl-CoA was incubated with 6-MSAS, malonyl-CoA (1-8.9 mM), NADPH and non-hydrolysable analogue **30**. Millimolar concentrations of malonyl-CoA were used to try to obtain large amounts of offloaded intermediates and to encourage the formation of longer chain intermediates. For 1 mM malonyl-CoA, the unreduced triketide (**38**) was observed at 13.7 minutes. For higher concentrations, very little of this could be seen. A small amount of reduced triketide (**39**) could be seen with 1 mM malonyl-CoA at 12.7 minutes. At higher concentrations of malonyl-CoA, no reduced triketide could be seen (Figure 2.20).

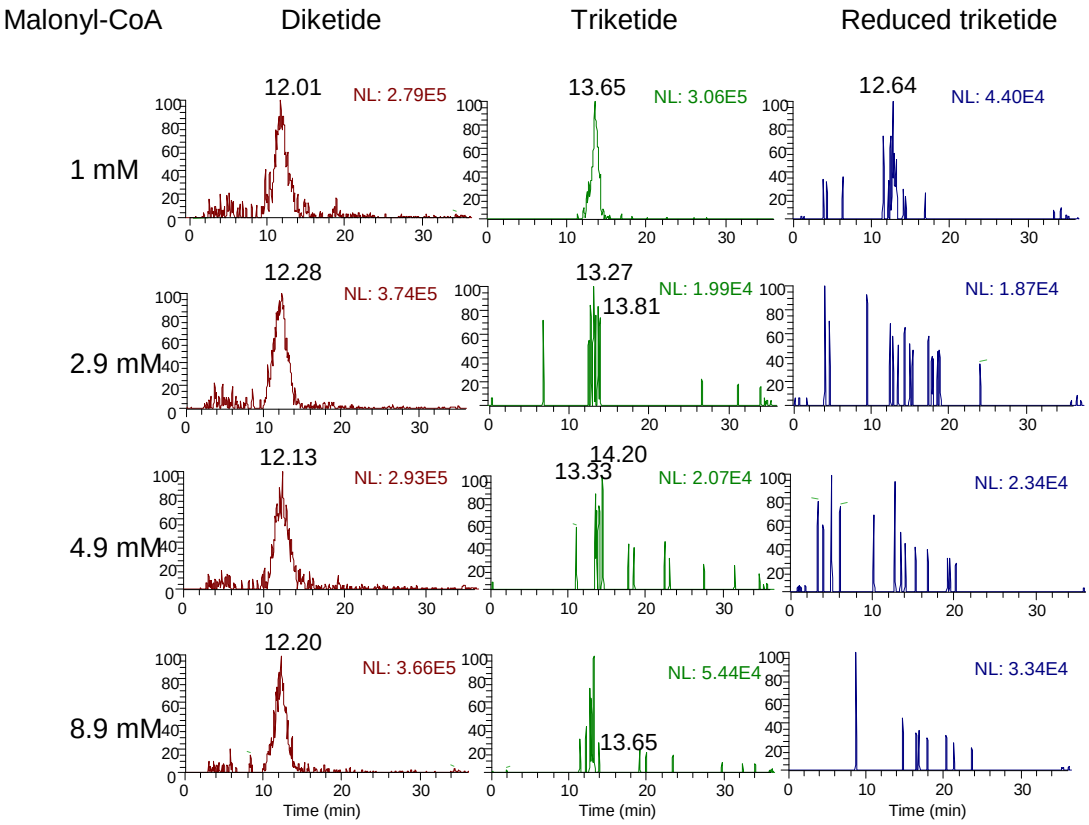


Figure 2.20: HRMS traces of competitive assays with 6-MSAS using acetoacetyl-CoA as a starter unit, malonyl-CoA as an extender unit, NADPH as cofactor and **30** as the offloading agent, showing diketide (12-12.3 mins), triketide (13.7 mins) and reduced triketide (12.6 mins) off-loaded at varying malonyl-CoA concentrations (FPLC-purified 6-MSAS (100 μg, treated with DTT), **30** (15 mM), acetyl-CoA (3 mM), NADPH (10 mM) and malonyl-CoA (1-8.9 mM) made up to a final volume of 100 μl with 6-MSAS assay buffer, pH 7.6, room temperature, overnight).

In all assays a roughly equivalent amount of off-loaded diketide could be seen. This would be from decarboxylation of malonyl-CoA to acetyl-CoA which could then prime 6-MSAS and be off-loaded by the non-hydrolysable analogue. At 1 mM malonyl-CoA, the signal intensity of the diketide peak and unreduced triketide peak are approximately the same. For any higher malonyl-CoA concentrations, the triketide peak is an order of magnitude smaller than the diketide. This implies that the contaminating acetyl-CoA is preferentially loaded on to the enzyme, as the off-loading reaction should proceed at a similar rate. This is not unexpected as 6-MSAS

has a K_m of 10 μM for acetyl-CoA and a V_{max} of 0.25 μmol/min per milligram of protein, while acetoacetyl-CoA has a much higher K_m (65 μM) and a much lower V_{max} (0.012 μmol/min per milligram protein).[56]

2.3 Future work

To obtain tetraketide intermediates, the reduced triketide-CoA (**80**, synthesised in section 4.2.2) could be supplied to the enzyme as a starter unit. Although previous work by Cheung *et al.* implies that the reduced triketide-NAC thioester is not a substrate for the AT domain,[87] if a small amount is loaded on to 6-MSAS then it should be possible to detect off-loaded intermediates. This could provide further information as to the timing of the dehydration, cyclisation and aromatisation steps.

The identity of the tetraketide seen in the single competitive assay could be established by synthesis or the use of a labelled starter unit or off-loading agent.

If the off-loading assays could be scaled up, there is the potential for NMR characterisation of the purified off-loaded intermediates. By using a chiral HPLC column and comparing the results with synthetic standards of known stereochemistry, it could be possible to identify the stereochemistry of the off-loaded reduced triketide. This would rely on the KR exhibiting the same stereospecificity for a small molecule substrate *in vitro* as for the phosphopantetheine-tethered thioester *in vivo*. This method would also allow the stereochemistry and position of the double bond to be established in a dehydrated intermediate.

6-MSAS is an unusual PKS in that it is trameric and features a large core domain between the DH and KR domains which is proposed to be necessary for subunit interaction. A structure would enable this unusual organisation to be elucidated. *Holo* 6-MSAS is a good candidate for crystallisation as the enzyme is available in large quantities and is of reasonable purity. The FPLC-purified 6-MSAS could be further

purified by ion exchange chromatography or gel filtration using a different column to achieve better separation of the intact synthase from contaminating fragments.

Previous attempts to crystallise an intact PKS have been unsuccessful due to the conformational flexibility of the megasynthase, particularly in the ACP region. This could be overcome by use of a non-hydrolysable analogue such as **34** or **50** (Figure 2.21) which could bind to the PKS and introduce a degree of rigidity.

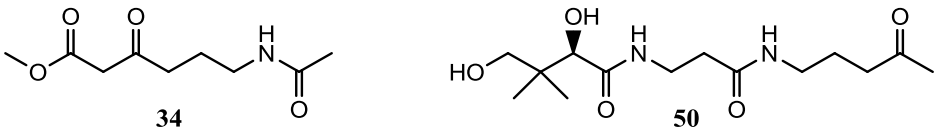


Figure 2.21: Non-hydrolysable analogues which could be used to aid crystallisation of 6-MSAS.

Alternatively, a cross-linking agent, such as one designed by Meier *et al.*, which links ACPs and dehydratases by use of a suicide substrate attached to a phosphopantetheine arm, could be employed.[88] This would require successful expression of the *apo* synthase rather than the *holo* to enable loading of the ACP domain with the substrate using a phosphopantetheinyltransferase.

2.4 Chapter summary

Holo 6-MSAS was successfully expressed in *E. coli*. 6-MSAS was purified in two steps and showed good specific activity when assayed by fluorimetry.

A non-hydrolysable carba(dethia)malonyl-N-acetylcysteamine analogue was synthesised and used to off-load enzyme-bound intermediates. Acetyl and acetoacetyl starter units were successfully off-loaded. These represent the first directly off-loaded intermediates from a type I iterative synthase using the non-hydrolysable analogue methodology. The triketide off-loaded from the acetoacetyl starter unit was found to be a substrate for the ketoreductase domain and possibly

also the dehydratase domain. The latter is particularly interesting in light of the reassignment of the dehydratase domain as a thioester hydrolase.[89] The reduction of the off-loaded triketide strengthens the case for the ketoreductase domain acting on a C_6 intermediate rather than a C_8 chain.

Competitive assays between malonyl-CoA and the non-hydrolysable analogue were carried out using both acetyl- and acetoacety-CoAl starter units. It was found that increasing the malonyl-CoA concentration decreased the amount of off-loaded intermediates. For the acetyl starter unit, diketide and triketide intermediates were seen. For acetoacetyl, there was a small amount of triketide seen, however, a large amount of contaminating diketide was observed due to decarboxylation of malonyl-CoA. In order to see tetraketide intermediates, a careful balance of analogue and malonyl-CoA concentrations may be required to reduce off-loading of the starter unit but to intercept later stage intermediates of the biosynthetic pathway.

CHAPTER 3

The dehydratase/thioester hydrolase domain of 6-MSAS

3.1 Introduction

3.1.1 The dehydratase domain

Dehydratase (DH) domains in fatty acid synthases (FASs) and polyketide synthases (PKSs) catalyse the formation of an α,β–carbon-carbon double bond by elimination of a β-hydroxyl group. While the DH domain in polyketide biosynthesis has not been as extensively studied as the ketoreductase domain, much has been inferred by comparison of the PKS DH with the DH found in FASs.

In mammalian FAS the dehydratase domain forms a pseudodimer from two adjacent regions of the same polypeptide. A single active site is formed at the interface between the two subdomains, comprising a histidine residue from the N-terminal and an aspartate from the C-terminal subdomain.[90] This is in contrast to Type II FAS dehydratases such as FabA and FabZ which are homodimers and contain two equivalent active sites.[91] The DH domain is structurally a 'double hot dog fold', two α-helices with a β-sheet curved around them. Further homodimerisation of the DH pseudodimer domains is not essential for activity as DH domains from FAS monomer have been found to be active.[92]

The mammalian FAS DH dimer (consisting of two double hot dog pseudodimers) has an overall V-shape and an angle of approximately 90° between the two monomers (Figure 3.1).[93]

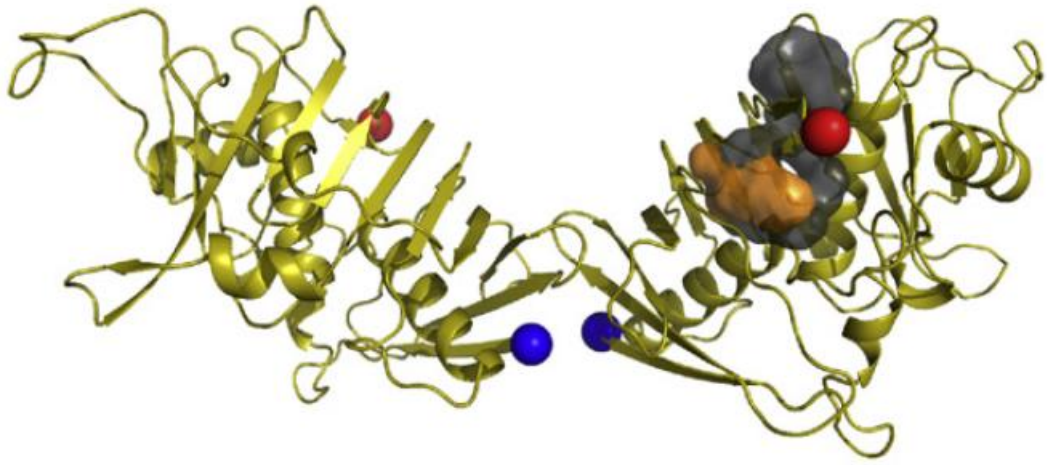


Figure 3.1: Mammalian FAS dehydratase dimer. The substrate tunnel entrance is marked in orange and the substrate tunnel is marked in grey. The N terminus is represented by a blue sphere and the red sphere represents the C terminus (Reprinted from *Structure,* **18**, Akey, D. L.; Razelun, J. R.; Tehranisa, J.; Sherman, D. H.; Gerwick, W. H.; Smith, J. L., Crystal Structures of Dehydratase Domains from the Curacin Polyketide Biosynthetic Pathway, 94-105, Copyright (2010), with permission from Elsevier.) [93]

Crystal structures of the EryDH4 domain[94] (from the 6-deoxyerythronolide B synthase DEBS) and all four DH domains in the curacin biosynthetic pathway[93] confirm that DH domains in PKSs have a high degree of similarity with those in mammalian FAS. The polyketide DHs also consist of a double hot dog fold with the same histidine/aspartate catalytic dyad and form dimers in the crystal lattice. Interestingly all five DH dimers from PKSs are in an extended conformation with approximately 180° between the monomers (Figure 3.2). Although the dimeric structure differs between FAS and PKS DHs, superposition of the two shows that the active site tunnel entrance, N-termini, C-termini and symmetry axes can be

superimposed without altering the angle between monomers.[93] This suggests that despite their structural differences, the DH domains can fulfil similar roles in an mFAS-like assembly line. An alternative is that the observed dimer can hinge between the subunits to form the mFAS configuration in the full PKS.[94] This is less likely in the light of the curacin structures, since identical non-physiological configurations would not be expected for five different DH domains.

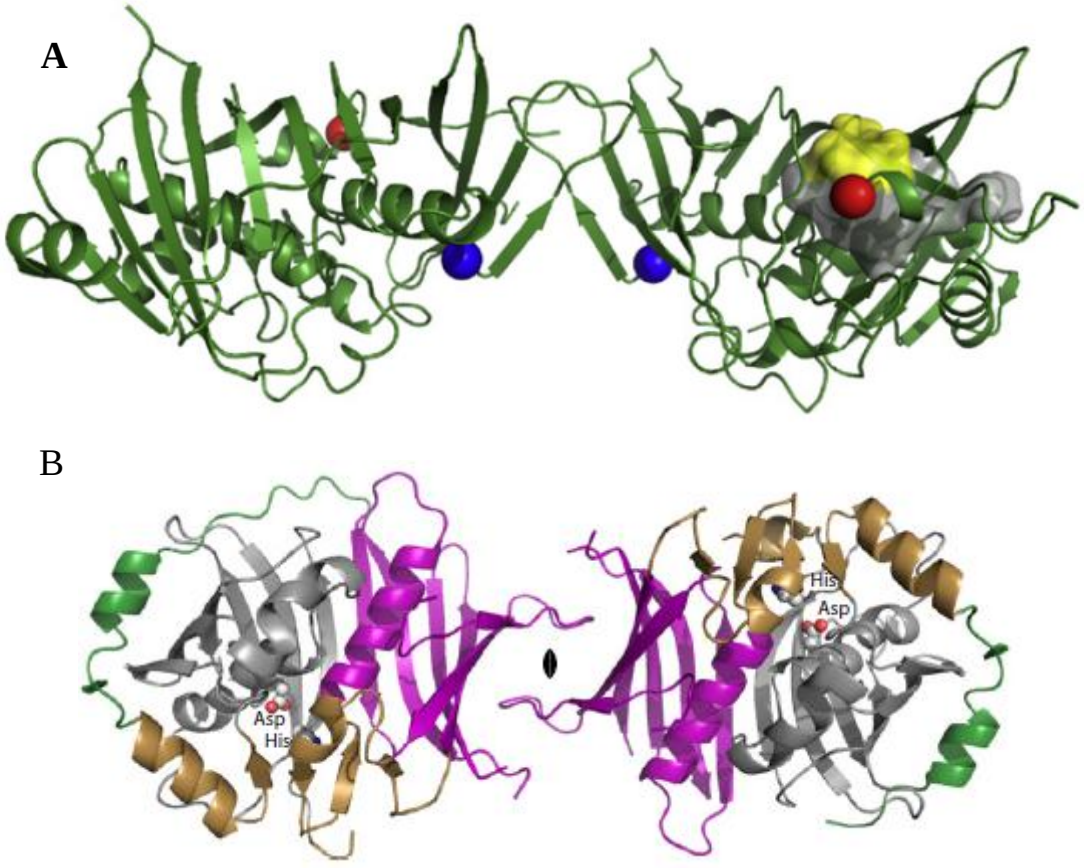


Figure 3.2: A: Dehydratase dimer from the curacin PKS. The substrate tunnel entrance is marked in yellow and the substrate tunnel is marked in grey. The N terminus is represented by a blue sphere and the red sphere represents the C terminus. B: CurK DH dimer showing the Asp and His catalytic dyad (PDB accession number 3KG9). (Reprinted from *Structure,* **18**, Akey, D. L.; Razelun, J. R.; Tehranisa, J.; Sherman, D. H.; Gerwick, W. H.; Smith, J. L., Crystal Structures of Dehydratase Domains from the Curacin Polyketide Biosynthetic Pathway, 94-105, Copyright (2010), with permission from Elsevier.) [93]

Mutation of a nearly conserved arginine to an aspartate in the proposed ACP binding site of EryDH4 caused production of 6-dEB to drop to 15% of the wild type level but no hydrated 6-dEB was observed.[94]

The reaction catalysed by the DH domain proceeds by proton abstraction from the α-position and formation of an enolate intermediate, which is stabilised by hydrogen bonding to a backbone amide. The enolate then reforms the carbon-oxygen double bond before moving the double bond along the molecule to eliminate the hydroxyl group with proton abstraction from the catalytic aspartate (Figure 3.3). The abstracted proton and hydroxyl are proposed to be on the same face of the substrate; i.e. a *syn* elimination, as demonstrated for FabA, a bacterial β-hydroxydecanoyl thioester dehydrase.[95]

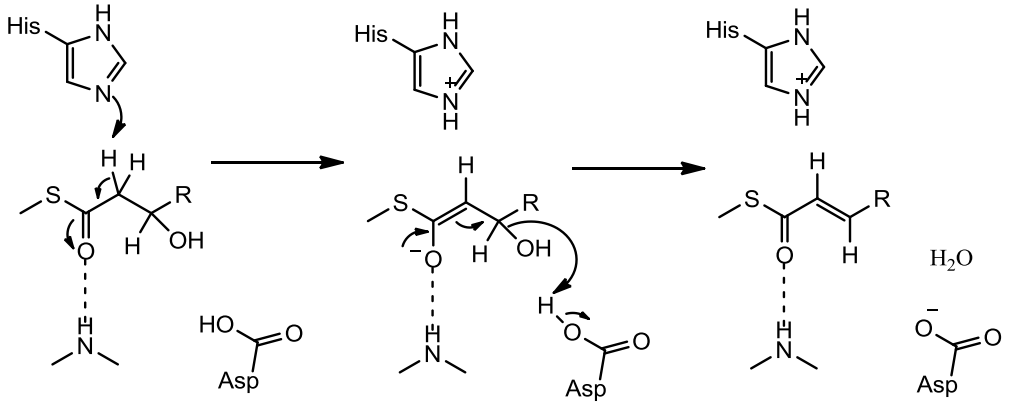


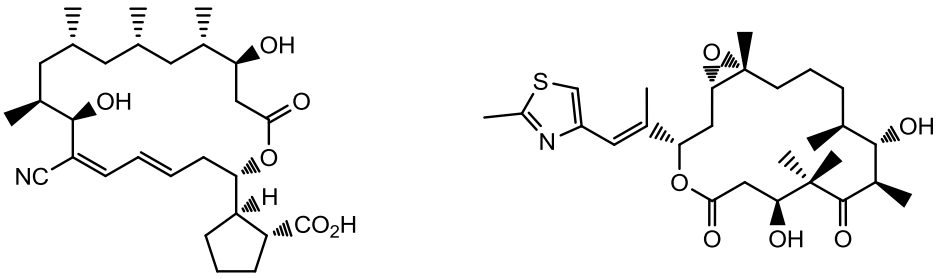
Figure 3.3: The proposed mechanism of double bond formation by the DH domain.[93]

The precise mechanism of dehydratases in polyketide biosynthesis is complicated by the formation of both *cis* and (mainly) *trans* double bonds, while fatty acids only form *trans* 2,3 double bonds, which could then be isomerised to form a *cis* bond if required (as in the isomerisation to a cis-3,4 double bond required for formation of cis-vaccenic acid by *E. coli* FAS). In some cases, such as the phoslactomycins[96], the formation of a *cis* double bond has been shown to be the work of a post-PKS tailoring enzyme. This, however, does not account for all the *cis* bonds found in PKS products. It was proposed after sequence analysis of ketoreductase domains that

trans double bonds could be formed by *syn* elimination from a (3R)-alcohol and *cis* double bonds from *syn* elimination from a (3S)-alcohol.[97]

The recently obtained crystal structures of all four dehydratase domains involved in the formation of curacin by a Type I modular PKS appear to support this hypothesis. Modelling of substrates under the condition that the catalytic aspartate should form a hydrogen bond with the eliminated hydroxyl led to the prediction that the product of an A-type KR would bind in a conformation that led to a *cis* bond, while the product a B-type KR should lead to a *trans* bond. This also suggested that the catalytic histidine would abstract different protons depending on whether the resulting double bond is *cis* (*pro*-S abstracted) or *trans* (*pro*-R abstracted).[93]

However, recent work on the borrelidin (**51**, Figure 3.4) biosynthetic pathway suggests that the relationship between KR specificity and dehydration product may not be as simple as it first appeared. Heterologous expression of DH2 and DH3 from the borrelidin megasynthase and analysis of substrate specificity *in vitro* found that both DH2 and DH3 as isolated domains acted on the (3R)-alcohol, despite DH2 in the context of the borrelidin PKS forming a *trans* bond and DH3 forming a *cis* bond (Dr Jennifer Andexer, personal communication).



51 **52**

Figure 3.4: Borrelidin (**51**) and epothilone A (**52**).

In addition, in myxobacterial PKSs it has been found that some extension modules give rise to a double bond without containing a dehydratase domain. In the case of epothilone (**52**, Figure 3.4) biosynthesis, this is because the product of module 4 is

actually dehydrated by the DH domain present in the adjacent module 5. It is proposed that the hydroxyl-bearing chain is passed from ACP4 to ACP5, where dehydration occurs. The chain is then passed backwards to KS5 for the standard extension carried out by module 5.[98]

3.1.2 The dehydratase domain in 6-MSAS

The DH domain in 6-MSAS is generally proposed to act on the C_6 intermediate, after the action of the ketoreductase. Reaction of 6-MSAS with 3-pentynoyl-N-acetyl cysteamine, an inhibitor which irreversibly modifies the catalytic histidine in DH domains, abolished 6-MSA production but did not affect the rate of NADPH oxidation.[55]

The cyclisation of the tetraketide chain to form 6-MSA (**1**) requires the formation of a *cis* double bond to enable condensation to take place. It is, however, unclear whether this bond is formed directly via dehydration or is the product of a dehydration followed by an allylic isomerisation reaction, such as that catalysed by β-hydroxydecanoyl thioester dehydrase (Figure 3.5).[99] It is also possible that dehydration does not occur until after cyclisation of the tetraketide takes place, which circumvents the problem of *cis* bond formation, but would be unusual in polyketide biosynthesis since tailoring domains generally act before subsequent condensations take place.[52]

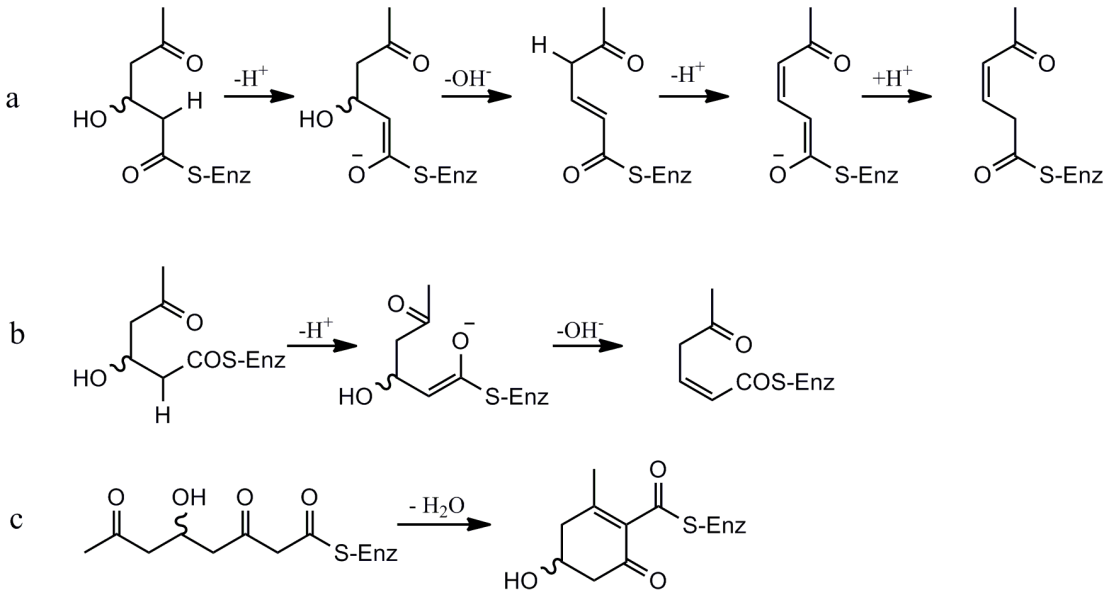


Figure 3.5: Possible routes to the formation of a *cis* double bond in 6-MSA biosynthesis, a) allylic rearrangement of a *trans* double bond formed by the DH, b) direct formation of *cis* double bond by the DH, c) cyclisation of the tetraketide followed by dehydration to form a double bond.

Attempts to synthesise the potential double bond intermediate(s) in the 6-MSA biosynthetic pathway have been unsuccessful. Harris and Harris found that the double bond did not remain localised in conjugation with the thioester, instead forming an equilibrium with the conjugated ketone form.[100]

Within our research group, attempts to synthesise the *cis*- and *trans* 2-enoyl from alkyne **53** were also unsuccessful, instead resulting in the formation of allene **54** upon oxidation of the hydroxyl group (Figure 3.6).[101]

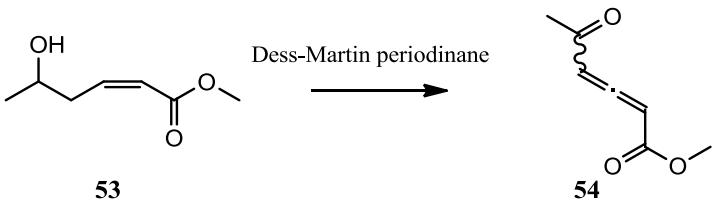


Figure 3.6: Formation of allene **54** after oxidation of alcohol **53**.

3.1.3 The thioester hydrolase domain

Recently, DH domain studies have been performed with the 6-MSAS homologue ATX, a 6-MSA synthase from *Aspergillus terreus*. Mutating the catalytic histidine in the DH domain to an alanine abolished 6-MSA production. As dehydration is proposed to occur on a C_6 intermediate, it was suggested that a triketide intermediate might be off-loaded by the enzyme, in the manner of TAL production, but this was not observed. Coexpression of the dehydratase mutant with ATX mutated in another domain restored 6-MSA production, showing that the DH mutation did not cause a structural change which rendered the enzyme inactive.[102]

A double mutant inactivating the ketoreductase and dehydratase would be expected to produce TAL in the same manner as a previously constructed ketoreductase mutant.[85] Instead neither 6-MSA or TAL was observed in the culture medium.[102]

It was noted that the DH motif HxxxGxxxP was found in bacterial orsellinic acid synthases, despite there being no dehydration involved in this biosynthetic mechanism. It was proposed that the dehydratase acts instead as an isomerase to convert the 2-*trans*-enoyl intermediate to the 3-*cis*-enoyl form or that, in fact, this domain is responsible for release of 6-MSA from the enzyme, since there is no thioesterase domain in the 6-MSAS or ATX genes.[57, 89]

Moriguchi *et al*. expressed ATX containing the His972A mutation (ATX-DHm) which showed no 6-MSA production.[89] After incubation with acetyl-CoA, [2-^{14}C]-malonyl-CoA and NADPH, followed by alkaline hydrolysis it was found that radiolabelled 6-MSA was released. When mutated ATX was incubated with labelled malonyl-CoA and NADPH, then digested with thrombin it was found that the radiolabel was localised on the ACP. These experiments suggest that ATX is able to form the tetraketide intermediate without need for dehydration, but is unable to off-load 6-MSA from the enzyme. Intact ATX was able to interact with labelled ATX-DHm to release labelled 6-MSA.

In light of this evidence, Moriguchi *et al*. proposed that the dehydratase domain in ATX and related enzymes is in fact a thioester hydrolase (TH) domain (Figure 3.7). This is related to the standard PKS thioesterase by virtue of also being part of the hot dog fold superfamily, but is unusual in that it is in the middle of the synthase rather than at the end. Also, since ATX and 6-MSAS are not affected by the serine protease inhibitor PMSF, product release cannot occur via an active site serine susceptible to this inhibitor.[56]

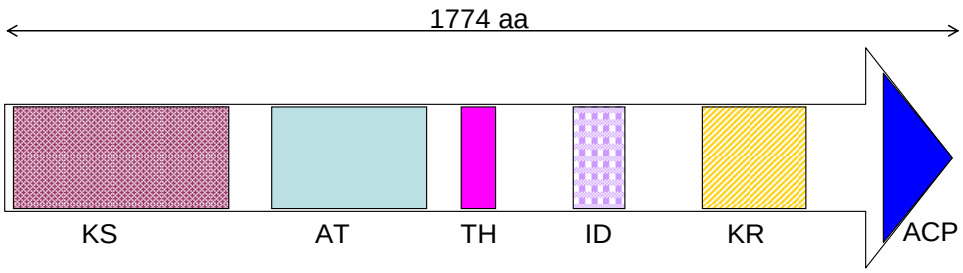


Figure 3.7: Proposed new domain organisation in 6-MSAS replacing the DH with a TH. The interdomain (ID) region is proposed to be critical for subunit interaction.[103]

Despite experiments which showed that ATX-DHm could be complemented with an ATX-KSmATmKRmACPm mutant, expression of the TH domain alone could not complement it. Instead the TH had to be expressed with a region that follows its C-terminus which is reported to contain the interdomain region (ID). This is proposed to be important for subunit binding to form a tetramer.[103] The TH-ID construct was able to complement ATX-DHm.

It was also shown that both intact ATX and TH-ID could hydrolyse 6-MSA-N-acetylcysteamine thioester. The proposed mechanism for the release of enzyme-bound 6-MSA proceeds with deprotonation of water, by the catalytic histidine, which attacks the thioester. The oxyanion is protonated by the catalytic aspartate to form an acetal which can collapse to reform the carbon-oxygen double bond and eliminate the thiol group (Figure 3.8).[89]

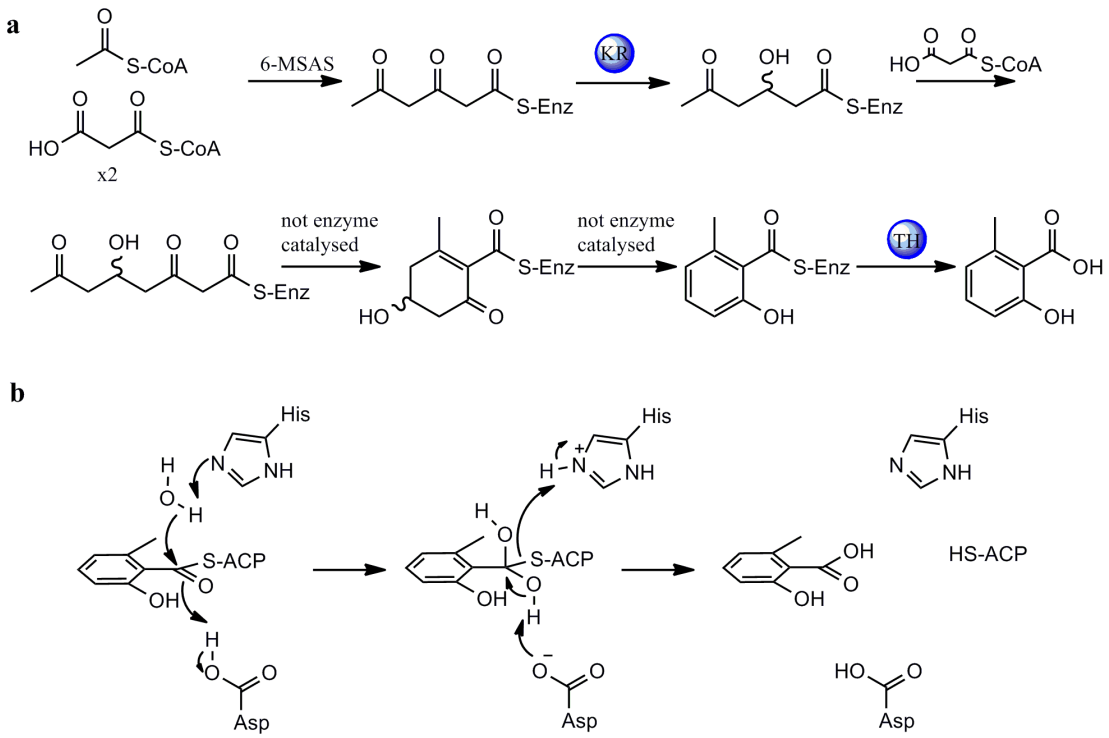


Figure 3.8: A: Revised biosynthetic pathway to 6-MSA with the DH reassigned as a TH. The cyclisation and aromatisation steps are proposed to be non-enzyme catalysed. B: Proposed mechanism of the thioester hydrolase domain.[89]

3.1.4 Aim of chapter

The aim of this chapter is to install the same mutation used in the ATX work into the dehydratase domain of 6-MSAS to first verify that the same results are seen. The 6-MSAS-DHm mutant will be incubated with acetyl-CoA, malonyl-CoA and NADPH and used for off-loading experiments with a non-hydrolysable malonyl-SNAC analogue (**30**). This provides a much milder method for investigating intermediates attached to the mutated synthase, since the alkaline hydrolysis conditions used for ATX would also promote cyclisation and aromatisation of the tetraketide intermediate.

It is worth noting that while the thioester hydrolase activity of the DH/TH domain was confirmed, no experiments were carried out to show that the domain was not able to catalyse a dehydration reaction. By comparison of off-loaded intermediates

from the 6-MSAS-DHm and 6-MSAS, it might be possible to observe a difference which could provide a clue as to whether this domain performs both functions.

3.2 Results and discussion

3.2.1 Expression of 6-MSAS-DHm

3.2.1.1 Construction of plasmid

Primers were designed by Dr Fanglu Huang to amplify the DH domain from the 6-MSAS-containing plasmid pKOS007-109 in two sections. The left-hand section introduced a mutation into the active site of the DH, converting histidine-958 to an alanine residue between a Bsu36I site and a KpnI site. The right-hand section extended the DH domain from a KpnI site to a BamHI site (Figure 3.9).

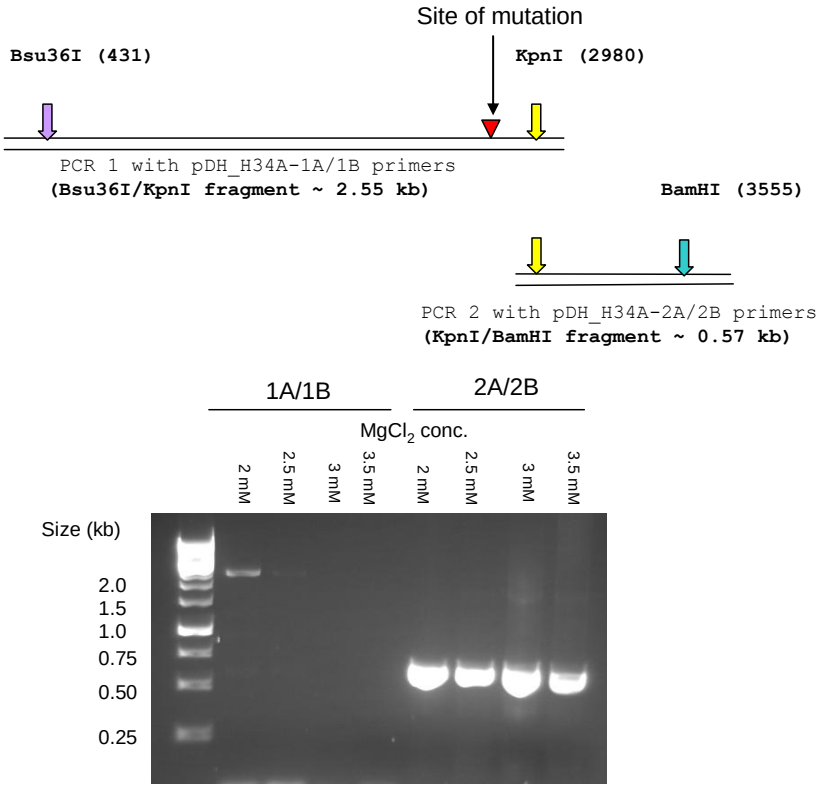


Figure 3.9: PCR scheme for the introduction of the H958A mutation. The intact 6-MSAS plasmid was used as a template for two PCR reactions. The first, using primers 1A/1B introduced the H958A mutation and formed a 2.55 kb fragment containing Bsu36I and KpnI restriction sites. The second, using primers 2A/2B formed a 0.57 kb fragment containing KpnI and BamHI restriction sites. The gel shows PCR products for primer pairs 1A/1B and 2A/2B with varying $MgCl_2$ concentration in the reaction.

PCR was carried out with both sets of primers to obtain the left and right fragments respectively. These were digested with KpnI/Bsu36I (left) and KpnI/BamHI (right) and purified via gel extraction.

pKOS007-109 was digested with Bsu36I and BamHI. Initially a sequential digest was attempted, the DNA being digested by BamHI, purified using a PCR purification kit (Qiagen) and then being digested by Bsu36I. This approach yielded no product, despite both enzymes being able to digest the plasmid separately. It was then decided to avoid the PCR purification step and perform a double digest followed by gel extraction. The three fragments were ligated using T4 ligase overnight at 16°C

using the concentration ratio 1:5:5 (vector:right:left) and the ligation mixture was transformed into competent NovaBlue cells. Colonies were picked and screened by comparison of plasmid size with authentic pKOS007-109, followed by sequencing. Of six plasmids sequenced, one (named DHm1) had the required histidine to alanine mutation. The other five contained a wild-type DH domain, presumably from a small contamination of singly cut plasmid in the doubly-digested sample.

3.2.1.2 Expression

Plasmid DHm1 was transformed into *E. coli* BL21 (DE3) alongside pET28a(+)-MtaA in order to yield mutated *holo* synthase. 6-MSAS-DHm/MtaA was grown under the same conditions as the unmutated holo synthase (section 2.2.1) and 6-MSAS-DHm was purified by ammonium sulphate precipitation.

Analysis of the purifed protein by SDS-PAGE showed a band of the correct size in the ammonium sulphate 22-40% fraction. This fraction could be further purified by FPLC under the same conditions as used previously, which gave a peak at the same elution volume as the intact synthase (Figure 3.10).

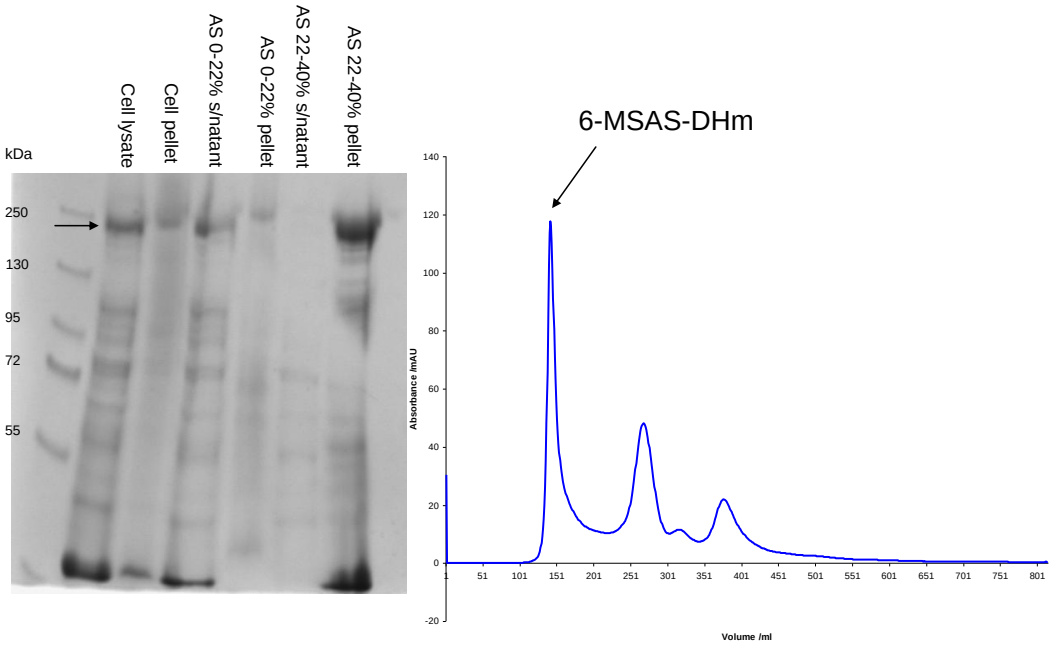


Figure 3.10: SDS-PAGE analysis of ammonium sulphate fractions in the purification of 6-MSAS-DHm and gel filtration trace of 6-MSAS-DHm (eluted with NaH_2PO_4 100 mM, EDTA 1 mM, DTT 2 mM, glycerol 20%, pH 7.1).

3.2.1.3 *In vivo* activity

If the dehydratase has been successfully mutated, then no production of 6-MSA should occur *in vivo*. 50 ml of culture supernatant was acidifed with 1 M HCl and extracted twice with ethyl acetate. The solvent was evaporated and the resulting oil was redissolved in 1 ml of methanol and analysed by LC-MS. By comparison with a standard and a culture expressing unmutated 6-MSAS, it was clear to see that the DHm culture did not produce 6-MSA (Figure 3.11). This confirms that the mutation in the DH domain has the same effect as that observed for ATX in *A. terreus*.

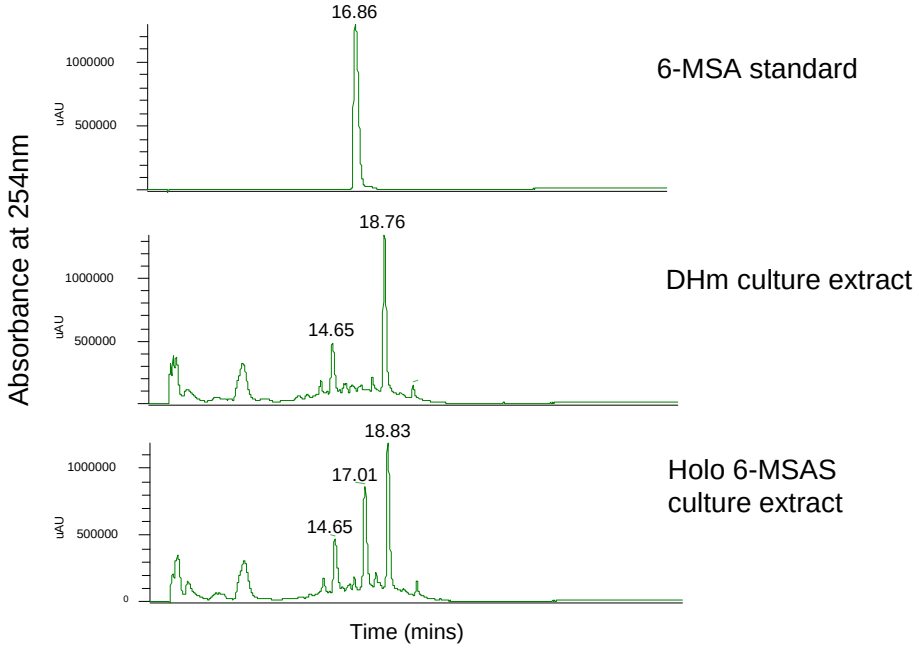


Figure 3.11: LC-MS traces showing culture supernatant extracts from *holo* 6-MSAS-DHm and *holo* 6-MSAS compared with a synthetic 6-MSA standard. The peak corresponding to 6-MSA is present in the *holo* 6-MSAS culture but is absent from the 6-MSAS-DHm culture. Conditions: Synergi 4u Polar-RP 150 x 2.00 mm column (Phenomenex), water/methanol + 0.1% TFA gradient; 0-5 minutes, 0% organic, 5-30 minutes 0-100% organic, 30-38 minutes 100% organic, 38-40 minutes 100%-0% organic.

3.2.1.4 *In vitro* activity

The ammonium sulphate pellet containing 6-MSAS-DHm was incubated overnight with 3 mM acetyl-CoA and 3 mM malonyl-CoA in the absence of NADPH. The assay mixture was then extracted with ethyl acetate, dried under nitrogen and redissolved in methanol before analysis on the LTQ-Orbitrap. A mass search for TAL (**55**) showed a strong peak at 12.0 minutes which had an exact mass of 127.0394 units (Figure 3.12). This showed that 6-MSAS-DHm was capable of producing the shunt product triacetic acid lactone, and therefore that the mutation in the DH domain did not affect the ability of the KS, AT and ACP to interact with each other.

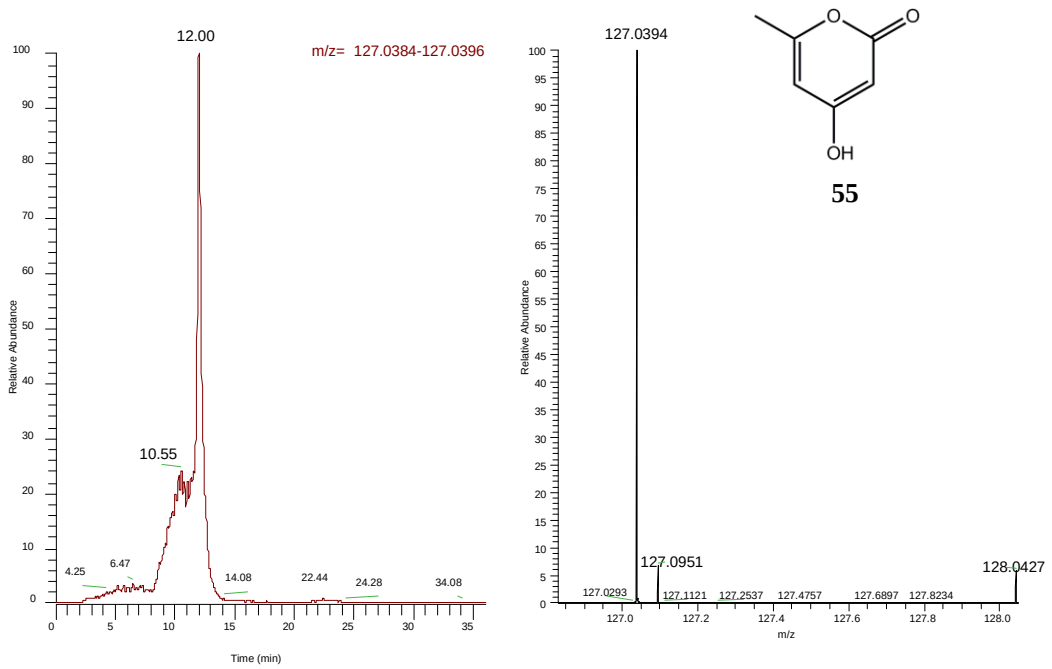


Figure 3.12: A HRMS mass search of an extract of a reaction containing 6-MSAS-DHm, acetyl-CoA, malonyl-CoA and NADPH showed the presence of TAL (**55**), as expected.

3.2.2 Off-loading experiments using 6-MSAS-DHm

3.2.2.1 Starter unit off-loading: Acetyl

6-MSAS-DHm was incubated with acetyl-CoA and non-hydrolysable analogue **30** overnight to see if the starter unit can be off-loaded from the mutated synthase. The assay mixture was extracted with ethyl acetate, dried and redissolved in methanol before analysis on the LTQ Orbitrap (Figure 3.13).

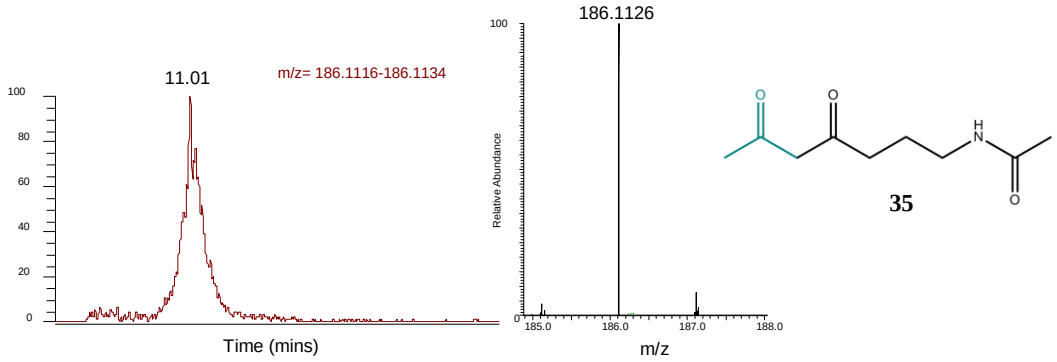


Figure 3.13: HRMS trace of an extract of a reaction containing 6-MSAS-DHm, acetyl-CoA and **30** as an offloading agent shows the presence of offloaded diketide (**35**).

Diketide **35** with an exact mass of 186.1126 units was seen at a retention time of 11.0 minutes. This corresponds with the off-loaded diketide seen when holo 6-MSAS was incubated with acetyl-CoA and non-hydrolysable analogue (section 2.2.5).

3.2.2.2 Competitive assays

For the unmutated synthase, increased malonyl-CoA concentration led to a reduction in the amount of intermediates off-loaded, as the non-hydrolysable analogue did not compete effectively with malonyl-CoA, and so the reaction with malonyl-CoA to form 6-MSA was preferred (section 2.2.7). With the dehydratase mutant unable to remove intermediates from the enzyme, it might be expected that the opposite effect would be observed, i.e. that off-loaded intermediates would increase with malonyl-CoA concentration since a greater amount would be formed on the enzyme. It could also be possible that longer chain intermediates would be isolated at higher malonyl-CoA concentrations since 6-MSAS-DHm could preferentially react with malonyl-CoA until the enzyme is unable to act any longer due to the mutation in the DH/TH domain. This stalled intermediate might then be a better candidate for off-loading.

6-MSAS-DHm was incubated with acetyl or acetoacetyl-CoA as a starter unit alongside varying concentrations of malonyl-CoA, and non-hydrolysable analogue **30** (Figure 3.14).

For the acetyl starter unit a significant amount of off-loaded diketide was seen at 11.9 minutes for all three malonyl-CoA concentrations used (4.7 mM, 8.6 mM and 11.9 mM). A mass search for triketide showed a peak at 12.8 minutes which corresponded to the correct mass, however this retention time is incorrect for the unreduced triketide which is expected at around 13.7 minutes. No reduced triketide could be seen in any assay. Assays using 3 mM malonyl-CoA also showed a peak corresponding to unreduced triketide at 13.7 minutes in samples both with and without NADPH.

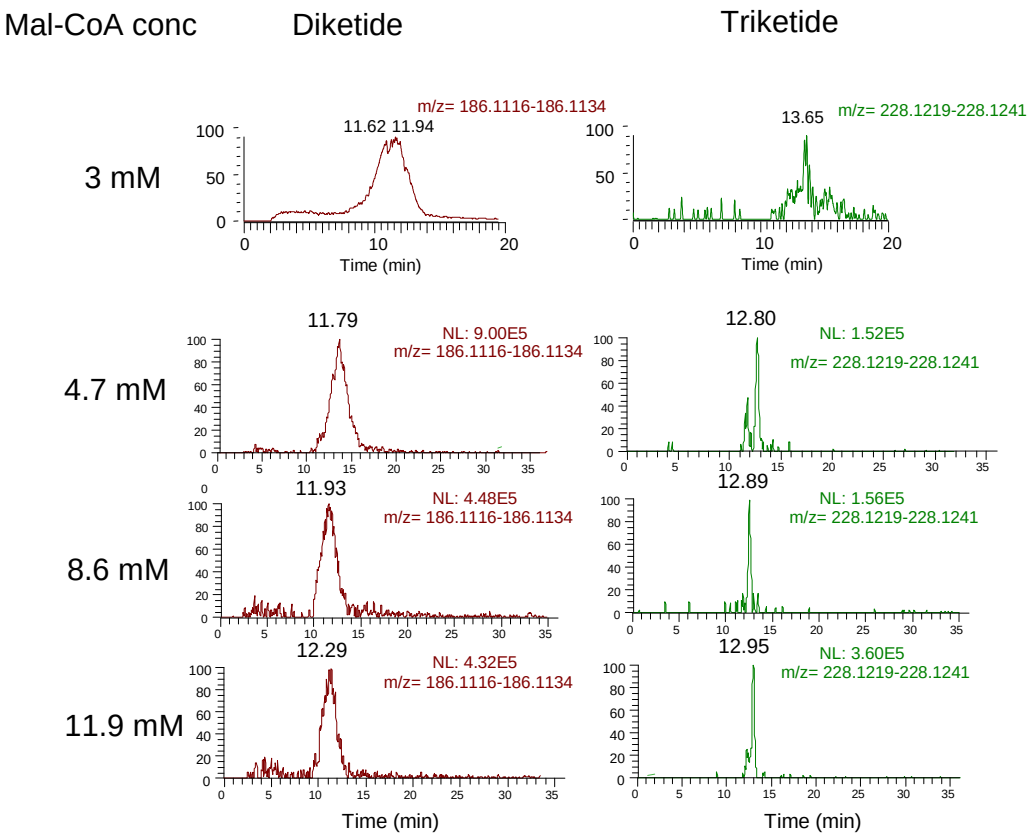


Figure 3.14: HRMS traces of extracts from competitive assays using 6-MSAS-DHm with acetyl-CoA as a starter unit, malonyl-CoA as an extender unit and **30** as an offloading agent showed offloaded diketide and triketide products.

For the acetoacetyl starter unit, no diketide was observed as would be expected. This also means that the decarboxylation of malonyl-CoA to acetyl-CoA which can prime the synthase and leads to off-loading of diketide was not a problem in this experiment. Again, when a mass search for triketide was performed a peak was seen at 12.8 minutes. No reduced triketide was present in the assay mixture.

The lack of off-loaded triketide intermediates is surprising given that these were readily off-loaded from the unmutated synthase and given that the diketide was off-loaded when acetyl-CoA was used as a starter unit.

Since the stalled tetraketide has been shown to be attached to the ACP, it could be that this is not able to be off-loaded by the non-hydrolysable analogue. Since the method of off-loading is dependent on mimicking the action of malonyl-ACP, it implies that all the intermediates seen are off-loaded from the KS domain, rather than the ACP. It is possible that triketide intermediates are not seen since the analogue does not compete well with malonyl-CoA (section 2.2.7), but are instead rapidly extended to tetraketides attached to the ACP which cannot be off-loaded. This would also reduce the total amount of intermediates detectable, since an enzyme which still has tetraketide attached is unable to function further. The loading of acetyl-CoA on to the KS does not involve the ACP and so may occur while the tetraketide chain is stalled on the ACP. This could explain the ease of diketide off-loading. It is also probable that a proportion of triketides (which would form tetraketide intermediates) are removed as TAL (**55**). Mass searches for TAL in initial off-loading assays found peaks at 11.7-11.8 minutes which corresponds to retention time seen when 6-MSAS-DHm acts in the absence of NADPH.

If the work of Moriguchi *et al.* is correct then it might be hoped that a pentaketide (**56**) could be off-loaded, however this would require the 6-MSAS KS to accept an unnaturally large intermediate (Figure 3.15). A mass search yielded no evidence of this pentaketide in any assay.

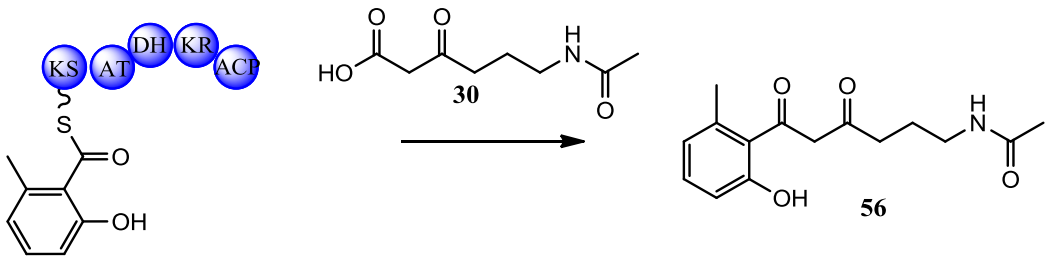


Figure 3.15: Potential pentaketide (**56**) which could be off-loaded from the KS domain.

Since the growing polyketide chain remains stuck on the enzyme in the dehydratase mutant, it was suggested that allowing the reaction between acetyl-CoA and malonyl-CoA on the enzyme before adding off-loading agent might result in the detection of longer chain intermediates rather than a high proportion of off-loaded starter unit as seen for *holo* 6-MSAS.

6-MSAS-DHm was incubated with acetyl-CoA or acetoacetyl-CoA, malonyl-CoA and NADPH for 1 hour before addition of the non-hydrolysable analogue.The assay using an acetyl starter unit gave a strong diketide peak with a small potential peak for unreduced triketide at 13.80 minutes. No reduced triketide, tetraketide or pentaketide was seen. The assay using an acetoacetyl starter unit also gave a strong diketide peak. This must be from decarboxylation of malonyl-CoA and priming of the enzyme with acetyl-CoA which is then off-loaded. No triketide or reduced triketide could be seen, nor any tetraketide or pentaketide.

These results suggest that the preferred enzyme reaction of 6-MSAS-DHm is to extend the starter unit with malonyl-CoA to the enzyme-bound tetraketide which cannot be processed further, or to remove the enzyme-bound triketide before the reduction step to form triacetic acid lactone. Alternatively, if the DH domain is a dehydratase and the mechanism of biosynthesis is not the same as in ATX, it could be that the reduced triketide is stalled on the ACP.

3.3 Future work

The formation of enzyme-bound 6-MSA by 6-MSAS-DHm could be confirmed by alkaline hydrolysis using ^{14}C-labelled acetyl-CoA or malonyl-CoA. Whether the enzyme-bound product can be released using intact 6-MSAS or the heterologously expressed DH domain could also be investigated. These would confirm that 6-MSAS carries out biosynthesis in the same way as ATX.

3.3.1 Starter unit assay: acetoacetyl

A repeat of the acetoacetyl starter unit assay used on the unmutated synthase (section 2.2.6) would be beneficial. This would establish whether off-loaded triketides could be seen when there is no competition from malonyl-CoA. Assays with and without NADPH could assess whether the KR domain is still capable of acting on non-enzyme bound substrates.

Comparison of the reduced triketide minus water peaks for 6-MSAS and 6-MSAS-DHm could give insight into whether the peak at 14.50 minutes seen in other assays (section 2.2.6) is in fact a product of the DH domain acting on the off-loaded intermediate. If the DH does act then the peak should be absent from assays using 6-MSAS-DHm in place of 6-MSAS.

3.3.2 Testing DH activity with synthetic substrates

In order to test the activity of the DH domain on non-enzyme bound substrates, a non-hydrolysable reduced triketide mimic could be synthesised. Protected reduced triketide acid (**57**, synthesised in section 4.2.2) could be coupled to N-(2-aminoethyl)acetamide (**58**) using a coupling reagent such as EDC. The dithiane group could then be removed using methyl iodide and calcium carbonate to give non-hydrolysable analogue **60** (Figure 3.16).

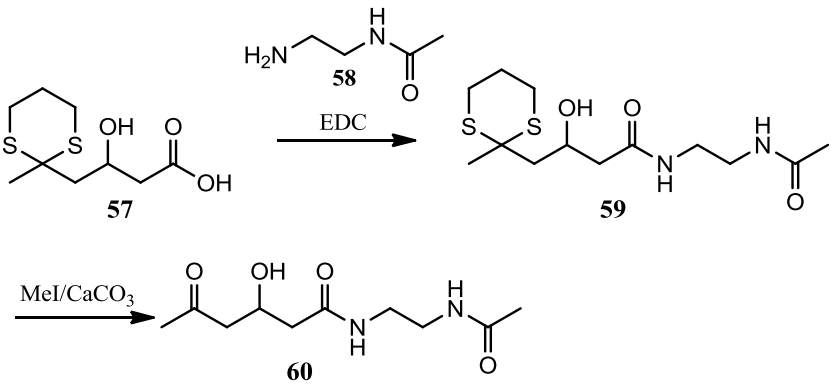


Figure 3.16: Proposed synthesis of non-hydrolysable reduced triketide analogue (**60**).

60 could be incubated with both 6-MSAS and 6-MSAS-DHm to see if dehydration can occur in solution and if the products can be detected by the LTQ-Orbitrap.

If dehydrated intermediates are observed then this is evidence that the DH domain in 6-MSAS may be capable of performing both dehydration and thioester hydrolysis, while lack of dehydration would cement the role of the DH domain as a thioester hydrolase (Figure 3.17).

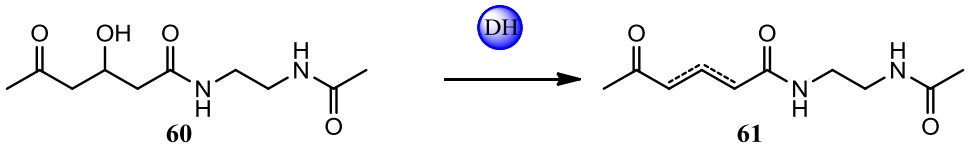


Figure 3.17: Proposed experiment to test the activity of the DH domain using a non-hydrolysable reduced triketide analogue that could form the dehydrated non-hydrolysable reduced triketide analogue **61**.

If the dehydrated triketide is observed, then the stereospecificty of the DH domain could be probed by using enantiomerically resolved acids as starting materials (section 4.2.2).

The same series of experiments could be performed using the commercially available substrate methyl 3,5-diketohexanoate which has previously been shown to be a substrate for the KR domain.[55] This would require the KR activity with the substrate to be confirmed before it was possible to assay the DH.

The non-hydrolysable analogue strategy could also be employed *in vivo*, as has been successful for the lasalocid producer *Streptomyces lasaliensis* (Dr Manuela Tosin, personal communication). Bacterial cultures expressing either 6-MSAS or 6-MSAS-DHm could be supplemented with protected carba(dethia)malonyl-N-acetylcysteamine and grown in the normal way. The culture medium could then be extracted with organic solvent, as could the cell lysate, to retrieve off-loaded intermediates. This approach could lead to more off-loaded intermediates as there would be a greater variety of conditions present in the cell. However, any modifications of the off-loaded intermediates would be difficult to assign to any domain of 6-MSAS due to other enzymes present.

3.4 Chapter summary

The dehydratase domain of 6-MSAS was inactivated by mutation of the catalytic histidine to alanine. This abolished production of 6-MSA in culture medium but the 6-MSAS-DHm mutant was still able to produce TAL when incubated with acetyl-CoA and malonyl-CoA.

Incubation of 6-MSAS-DHm with acetyl-CoA, malonyl-CoA, NADPH and non-hydrolysable analogue carba(dethia)malonyl-NAC led to the off-loading of only diketide intermediates. Using acetoacetyl-CoA did not give any observable triketide intermediates. Incubating the enzyme with substrates before adding the off-loading agent led to off-loading of diketide intermediates for both acetyl and acetoacetyl starter units.

The lack of higher intermediates observed suggests that the off-loading agent does not compete effectively with malonyl-CoA and thus only enzyme-bound tetraketide is available for off-loading. As this would be attached to the ACP, it is unlikely that the off-loading mechanism would be as successful as for KS-bound intermediates. Work by Y. Sun *et al.* has suggested that chain transfer is possible between an intact

PKS module and a discrete *holo* ACP.[104] If the *holo* ACP for 6-MSAS could be heterologously expressed and added in to the reaction mixture, this could enable the enzyme-stalled intermediates to be transferred to the discrete ACP. This small protein could then be monitored by mass spectrometry and the mass of the intermediates attached to it could be determined.

There is still a considerable amount of work to do to assay whether the DH domain in 6-MSAS performs both a dehydration and a hydrolysis reaction in 6-MSA biosynthesis, or whether it is only a thioester hydrolase. This chapter provides a framework for these experiments.

CHAPTER 4

The stereochemistry of ketoreduction by 6-MSAS

4.1 Introduction

4.1.1 Ketoreductase domains

In polyketide synthases the ketoreductase domain transforms the ketone formed by the ketosynthase into a hydroxyl group by addition of a hydride from the reduced form of nicotinamide dinucleotide phosphate (NADPH, **62**). The ketoreductase (KR) is the first step in the reductive loop in fatty acid biosynthesis and is responsible for much of the structural diversity seen in polyketides.

Extensive studies on the first ketoreductase domain (KR1) in the 6-deoxyerythronolide B synthase (DEBS) Type I modular system have been carried out to determine how selective this KR is for stereocentres in the substrate as well as the inherent stereochemistry of reduction. As DEBS utilises methylmalonyl-CoA as an extender unit, there is an additional stereocentre present at C-2 from the methyl group. It was found that KR1 determined the stereochemistry at C-3 by exclusively reducing the ketone to the (3R)-hydroxy group. Additionally the KR set the stereochemistry at C-2 by exclusively recruiting the (2S)-methyl precursor.[105] Surprisingly, experiments with KR5 and KR6 from DEBS showed loss of stereocontrol

when the domains were presented with unnatural substrates which were much smaller.[106] This went against previous results which had shown that DEBS modules engineered to contain a KR from a different module gave products with the natural stereochemistry of the imported KR.[107] Further experiments on excised KR domains also reported loss of stereocontrol when KR domains which act on long chain substrates were presented with short chains, suggesting a fine energetic balance between modes of substrate presentation in the KR active site.[108]

Incubation of DEBS 1 TE and DEBS 3 with chirally deuterated NADPH showed that KRs 1, 2, 5 and 6 all used the *pro*-S hydride to perform ketoreduction.[109] As KR1 generates an (R)-hydroxyl, while KRs 2, 5 and 6 all produce an (S)-hydroxyl, this rules out the KR setting the stereochemistry of its product by using different hydrides from NADPH to form the R and S isomers.

Alignment of KR domains from 17 modular PKSs led to identification of motifs indicative of whether a KR domain formed the R or S hydroxyl, also known as A-type or B-type (Figure 4.1). The A-type was found to have a conserved tryptophan residue at position 142 in the aligned KR sequences, while the B-type lacked this tryptophan, but had a leucine-aspartate-aspartate (LDD) from residues 93-95. D95 was invariant, while D94 was replaced by glutamate in two cases. Leucine-93 was less strictly conserved. B-type KRs also typically had a proline at residue 144 and asparagine at residue 148. The identification of these 'Caffrey motifs' allows the prediction of KR stereochemistry in cases where the reduction product is further processed to remove the hydroxyl.[110]

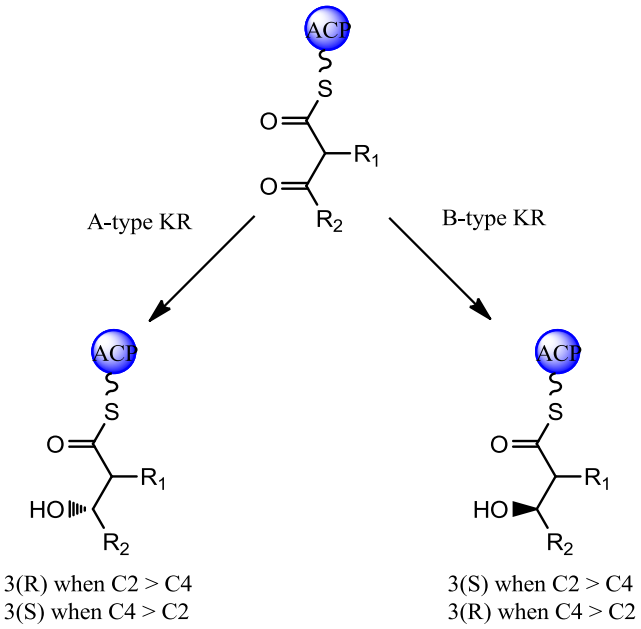


Figure 4.1: Reduction of ketone by A and B type ketoreductases leads to opposing stereochemistry at C3.

A putative catalytic triad consisting of tyrosine, serine and lysine was identified from homology to the short chain dehydrogenase/reductase superfamily. Mutation of each of these three residues abolished activity in the KR6 domain of DEBS. In the proposed mechanism, the ketone is coordinated by hydrogen bonding to tyrosine and serine while hydride attack occurs. A proton is then donated from the tyrosine residue to form the hydroxyl group. This mechanism could allow the formation of opposite stereochemistries of hydroxyl group by opposite orientation of the substrate in the active site (Figure 4.2).[97]

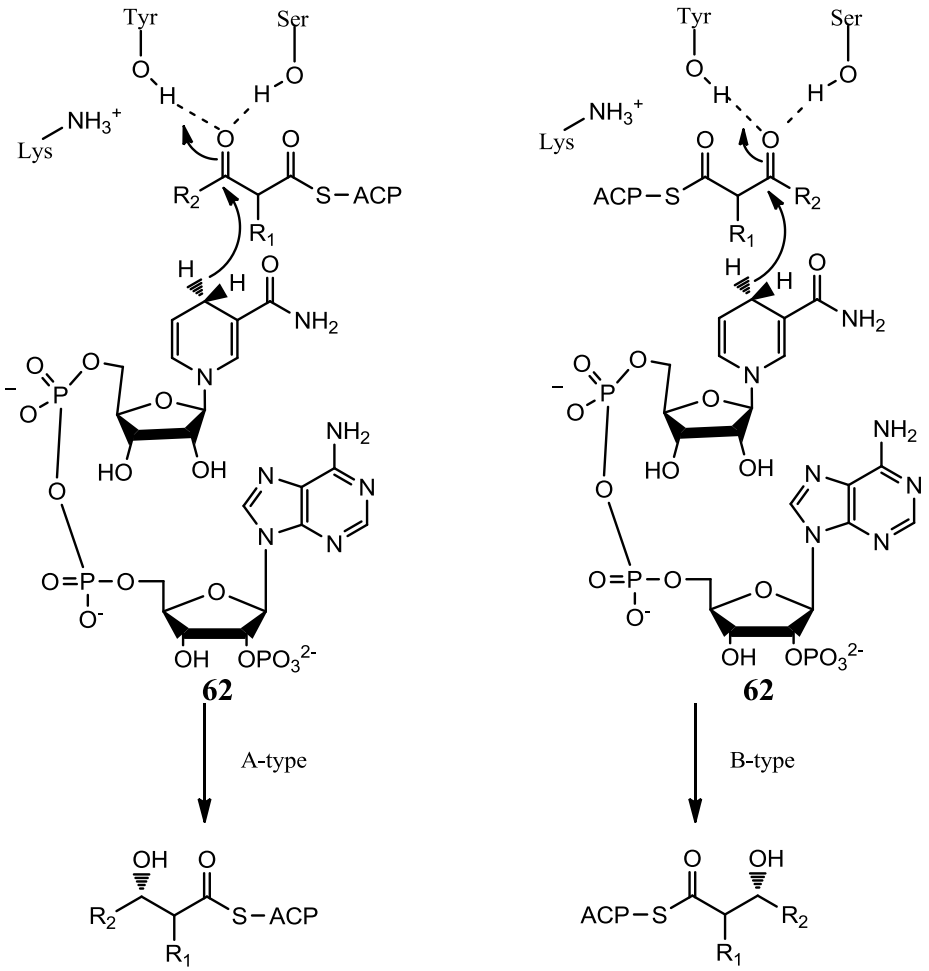


Figure 4.2: Proposed mechanism of reduction in A and B type ketoreductase domains using a hydride from NADPH (**62**).

A KR domain from the actinorhodin Type II PKS was crystallised in the presence of NADPH and NADP+ and confirmed the tyrosine and serine active site residues. The structure also established an extensive proton relay network of water molecules. No conserved Caffrey motifs were found in the sequence, however this is generally true of all aromatic polyketide KRs and may be due to the fact that aromatic polyketide synthases have a different binding motif to modular ones.[111]

The erythromycin KR1 (B-type) was crystallised by Keatinge-Clay and Stroud and revealed an active site groove. If substrates entered from opposite ends of this groove then it would explain the opposing stereoselectivities of A and B type KRs.[112] Comparison with the more ordered structure of the first KR of the tylosin PKS, which is also B-type, led to development of fingerprints for A-type, B-type and non-

functional C-type ketoreductase domains. The conserved tryptophan in A-type KRs makes space for the phosphopantetheine arm to enter the active site from the one side. In B-type KRs the LDD motif seals the lid helix and LDD loop together, positioning the phosphopantetheine arm to enter the KR active site from the opposite side to the A-type (Figure 4.3).[113]

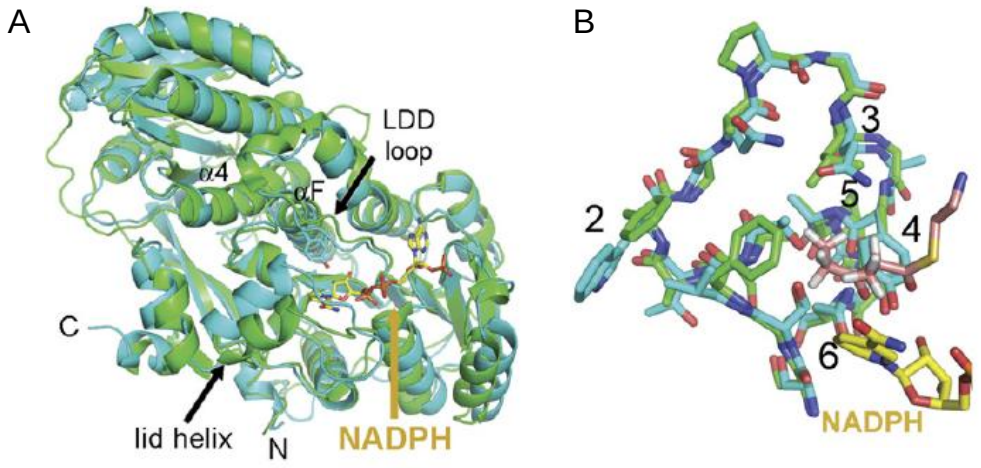


Figure 4.3: A: Superposition of EryKR1 (B type, green) and TylKR1 (A type, blue). B: Alignment of the catalytic regions of EryKR1 (green) and TylKR1 (blue). 2 is the position occupied by tryptophan in an A-type KR. (Reprinted from *Chemistry & Biology,* **113**, Keatinge-Clay, A; A tylosin ketoreductase reveals how chirality is determined in polyketides, 898-908, Copyright (2007), with permission from Elsevier.)[113]

4.1.2 Ketoreduction in 6-MSAS

In 6-MSA biosynthesis, ketoreduction occurs only once in the three cycles of condensation carried out by 6-MSAS. In the absence of NADPH (**62**), the shunt product triacetic acid lactone (TAL, **55**) is formed from acetyl-CoA and two units of malonyl-CoA (Figure 4.4).[54] Incubation of 6-MSAS with acetylenic inhibitors led to reduction of 6-MSA production but did not interfere with NADPH oxidation. Acetylenic inhibitors react irreversibly with the active site histidine present in dehydratase domains, implying that reduction, then dehydration occurs on an intermediate.[55] Mutation of the NADPH binding site in heterologously expressed 6-

MSAS produced TAL in the culture medium, with concomitant loss of 6-MSA production.[85] These pieces of evidence strongly suggest that reduction occurs on a C-6 intermediate rather than on the completed 8 carbon polyketone chain (Figure 4.4).

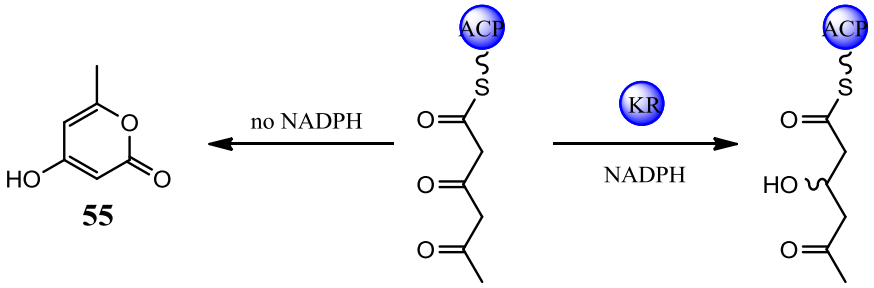


Figure 4.4: Ketoreduction occurs on the C_6 intermediate in 6-MSA biosynthesis. In the absence of NADPH, TAL (**55**) is formed.

As ketoreduction occurs on a C_6 intermediate followed by elimination of the resulting hydroxyl, the stereochemistry of the reduced triketide produced is cryptic. In order to investigate this, the enantiomerically resolved reduced triketides **63** and **64** were synthesised as NAC thioesters by Cheung *et al.* (Figure 4.5).[87] It was hoped that the intermediate with the correct stereochemistry would be selected and loaded on to 6-MSAS, where it would be elaborated into 6-MSA. Unfortunately, on incubation with purified 6-MSAS neither intermediate was able to form 6-MSA at a rate greater than that observed with malonyl-CoA alone. Both intermediates did serve as inhibitors for the enzyme.

Figure 4.5: Enantiomerically resolved reduced triketide NAC esters **63** and **64** synthesised by Cheung *et al.*[87]

The failure of this work could be due to the presence of intermediates as NAC thioesters when the natural substrate for the AT domain is a CoA thioester. Alternatively, the AT domain could be unable to identify and load the triketide on to the KS domain.

Recent work on the bacterial 6-methylsalicylic acid synthase ChlB1, from the chlorothricin pathway in *Streptomyces antibioticus*, has found that mutation of the NADPH binding site from GxGxxG via a single or triple mutation of glycine residues abolished 6-MSA production, but did not give rise to production of TAL. Ding *et al.* postulate that in the bacterial MSAS, NADPH binding induces a conformational change which is essential for activity. Instead, mutation of the active site tyrosine to phenylalanine was carried out in order to eliminate the reductive ability of the KR. Intriguingly, this led to the formation of orsellinic acid (**2**), albeit at a significantly reduced level, which suggests that the ketosynthase is not specific for the reduced triketide intermediate in bacterial MSAS (Figure 4.6).[114]

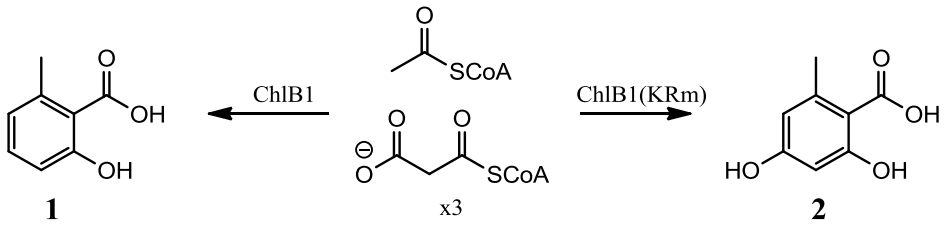


Figure 4.6: Mutation of the ketoreductase domain in MSAS ChlB1 resulted in the formation of orsellinic acid (**2**) rather than 6-MSA (**1**).

4.1.3 Aim

The aim of this section is to synthesise the diastereomerically resolved reduced triketide-CoAs **65** and **66**. These will be loaded on to *apo* 6-MSAS using a phosphopantetheinyltransferase. This can then be incubated with ^{13}C-labelled malonyl-CoA and the 6-MSA produced analysed by LC-MS (Figure 4.7). 6-MSA produced from the triketide will only incorporate one unit of labelled malonate, while any 6-MSA formed from decarboxylation of labelled malonate to form acetate which can prime the synthase will be labelled at every carbon.

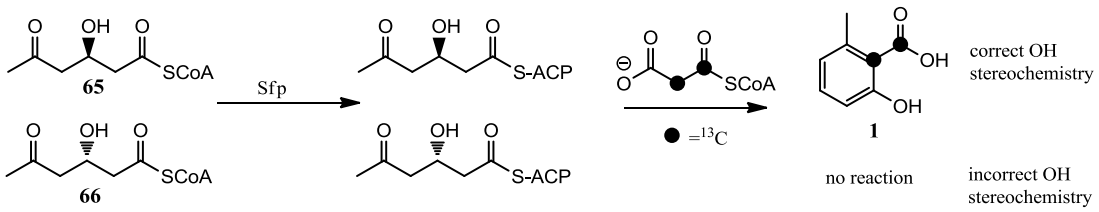


Figure 4.7: Synthesised reduced triketides **65** and **66** can be loaded on to 6-MSAS. The correct OH stereochemistry should be elaborated into 6-MSA (**1**).

4.2 Results and discussion

4.2.1 Caffrey motif analysis.

Analysis of the protein sequence of 6-MSAS reveals that it contains two phenylalanine-serine-serine (FSS) motifs, which contain the active site serine in ketoreductase domains. The first covers residues 1256 to 1258, and the second, residues 1567 to 1569. The glycine-rich region proposed for the NADPH binding site is ^{1198}GTPGVSG for the first potential KR domain and ^{1418}GGLGVLG for the second.

Both 6-MSAS KR sequences were aligned with the KR1 and KR2 domains from DEBS using ClustalW.[115] DEBS KR1 is a B-type reductase while KR2 is an A-type. This showed that the first KR domain lacked the catalytic tyrosine close to the FSS motif, while the second KR domain contains tyrosine-1572 which could fulfil this role (Figure 4.8). This second domain is the one which is identified as the KR in the annotation of the sequence, from homology with FAS.[57]

```
DEBS_KR1        GTVLVTGGTGGVGGQIARWLARRGAPHLLLVSRSGPDADGAGELVAE-------------
DEBS_KR2        GTILVTGGTAGLGAEVARWLAGRGAEHLALVSRRGPDTEGVGDLTAE-------------
C-terminal      ------GGLGVLGLEVADFLVEKGARRLLLISRRALPPRRTWDQVSEDLQPTIA------
N-terminal      GTPGVSGSMESLVHQIAWPPATPAEEPLSIETVILVSPDATTRALYAASLPTRVNSFQFS
                      *.   :  ::*   .  .   * : :     .  .

DEBS_KR1        ------------LEALGARTTVAACDVTDRESVRELLGG-IGDDV-PLSAVFHAAATLDD
DEBS_KR2        ------------LTRLGARVSVHACDVSSREPVRELVHGLIEQGD-VVRGVVHAAGLPQQ
C-terminal      --------KIRLLESRGASVHVLPLDITKPDAVEQLTTALDRLSLPSVQGVVHAAGVLDN
N-terminal      STQEFFSNASSLPLEKGTVVTYIPGEVASLAEVPAASESFTWNLLELIKFTVNGSLPIKV
                              *: .   . :::.   *     .        :  ..:.:   .

DEBS_KR1        GTVDTLTGER-------------IERASRAKVLGARNLHELTRELDLTAFVLFSSFASAF
DEBS_KR2        VAINDMDEAA-------------FDEVVAAKAGGAVHLDELCSDAEL--FLLFSSGAGVW
C-terminal      ELVMQTTRDA-------------FNRVLAPKIAGALALHEVFPPKSVDFFVMFSSCGNLV
N-terminal      FTLTANIGEGQTPTALAQSPLYGLARVIASEHPDLGTLIDVEEPVIPLSTMRYIQGADII
                  :                     : ..  .:  .   * ::         : : . ..

DEBS_KR1        ----GAPGLGGYA--PGNAYLDGLAQQRRSDGLPATAVAWGTWAG---------------
DEBS_KR2        ----GSARQGAYA--AGNAFLDAFARHRRGRGLPATSVAWGLWAA---------------
C-terminal      ----GFTGQASYG--SGNAFLDTLATHRARLGDAAVSFQWTSWRGLGMGASTDFINAELE
N-terminal      RINDGIARTSRFRSLPRNKLLPASEGPRLLPRPEGTYLITGGLGVLGLEVADFLVEKGAR
```

Figure 4.8: Alignment of the N-terminal and C-terminal potential KR domains in 6-MSAS with KR1 (B-type) and KR2 (A-type) from DEBS. Glycine-rich regions, the FSS motif, the catalytic tyrosine, the B-type LDD/LDN motif (KR1 and C-terminal KR from 6-MSAS) and the A-type tryptophan (KR2) are marked.

The C-terminal domain features an LDN motif between residues 1515 and 1517, which is similar to the LDD B-type Caffrey motif. This aligns with the LDD motif in the DEBS KR1 domain and suggests that the 6-MSAS KR is a B-type and will thus produce a product with an (R)-hydroxyl group (Figure 4.9). There is no tryptophan residue in proximity to the FSS motif and none which aligns with the tryptophan in DEBS KR2, which suggests that it is not an A-type KR.

The assignment of the KR as B-type is in opposition to the assignment of the KR from the MSAS ChlB1, which was assigned as an A-type from sequence analysis as it lacked both the His residue in the active site and the LDD motif.[114]

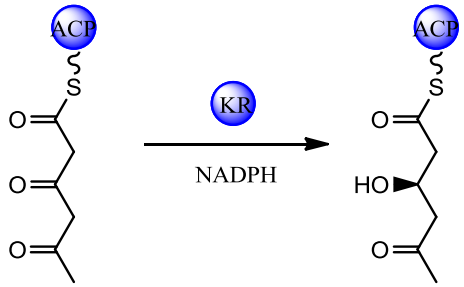


Figure 4.9: Predicted stereochemistry of reduced triketide intermediate in 6-MSAS with a B-type KR domain.

4.2.2 Synthesis

Initially it was decided to synthesise the dithiane-protected reduced triketide acid (**57**) using the same method as Cheung *et al.* using (S)-camphanoyl chloride (**75**) to derivatise the hydroxyl group and enabling separation of the resulting diastereomers (Figure 4.10).[87] This could then be deprotected to form a reduced triketide which could then be couple to coenzyme A. It was decided to perform the deprotection before coupling to CoA as the reactivity of methyl iodide, used in deprotection of the dithiane, with coenzyme A was unknown and side reactions could occur.

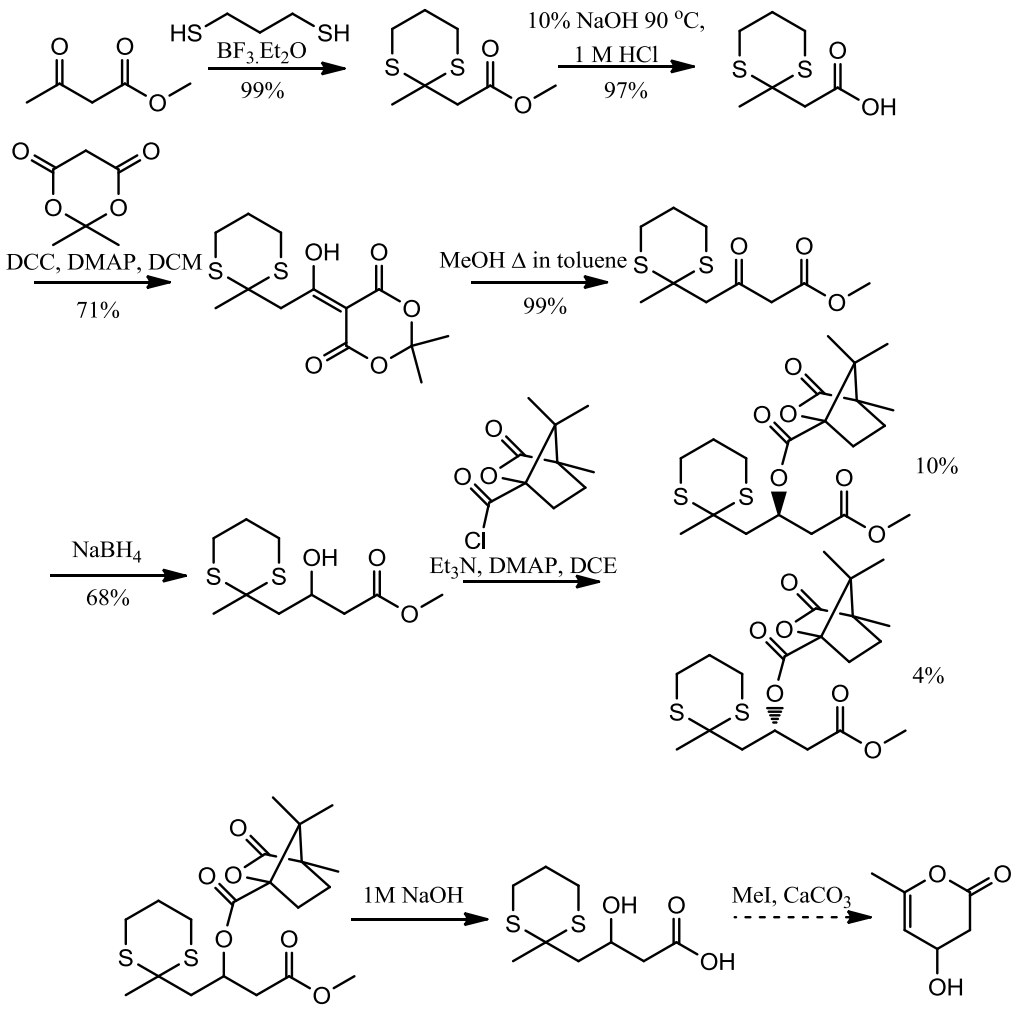


Figure 4.10: Synthesis of the resolved diastereomers **76** and **77**.

Methyl acetoacetate (**67**) was reacted with with 1,3-propanedithiol (**68**) and boron trifluoride dietherate at room temperature for 1 hour to form the protected species (**69**) in 99% yield. The methyl ester was removed by reaction with sodium hydroxide at 90 °C for 1 hour. The acid (**70**) was coupled with Meldrum's acid (**71**) using DCC and DMAP in dry DCM to form the cyclic adduct (**72**). **72** was opened up using dry methanol in toluene to give the β–ketoester (**73**). Reaction with sodium borohydride reduced the ketone to a racemic alcohol (**74**) which was then coupled to (S)-camphanic chloride (75) using triethylamine and DMAP in dichloroethane. The diastereomers formed (**76** and **77**) were separated by column chromatography on silica gel using a gradient of petroleum ether 40-60 and ethyl acetate (100:1 to 5:1) (Figure 4.10).

Deprotection of the camphanic and methyl esters by the acid and base hydrolysis method given in the paper of Cheung *et al.* did not lead to isolation of any product. Instead a simple base hydrolysis was used to cleave both esters simultaneously and the product (**57**) could be obtained by preparative TLC in low yield.

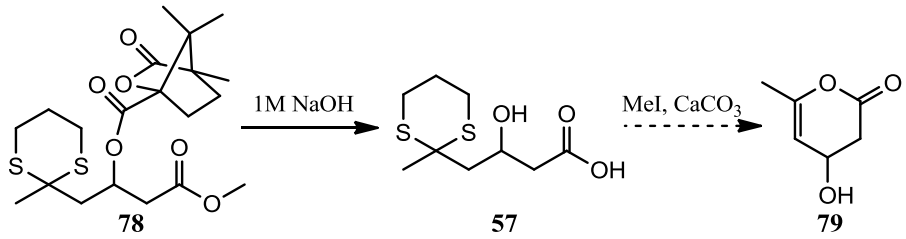


Figure 4.11: Cleavage of esters to form **57** and deprotection of dithiane, which could potentially form the pyrone **79** .

Attempts to deprotect **57** using methyl iodide and calcium carbonate did not yield the deprotected reduced triketide. This could be due to the formation of pyrone **79** (Figure 4.11). After several attempts it was decided that an alternative strategy might be more successful.

In the synthesis of non-hydrolysable malonyl-CoA analogues, Tosin *et al.* utilised the enzymes which form the pathway to coenzyme A in *E. coli* to elaborate synthetic pantetheine compounds to the corresponding CoA.[81] These enzymes had been shown to have broad substrate specificity for pantothenate derivatives.[116]
It was postulated that the reduced triketide-CoA (**80**) could be formed by the action of pantothenate kinase (PK), phosphopantetheine adenyltransferase (PT) and de-phosphocoenzyme A kinase (DK) on the reduced triketide-pantetheine (**81**). This could be formed by deprotection of the dithiane group after coupling of the previously synthesised acid (**57**) with pantetheine (**83**) (Figure 4.12).

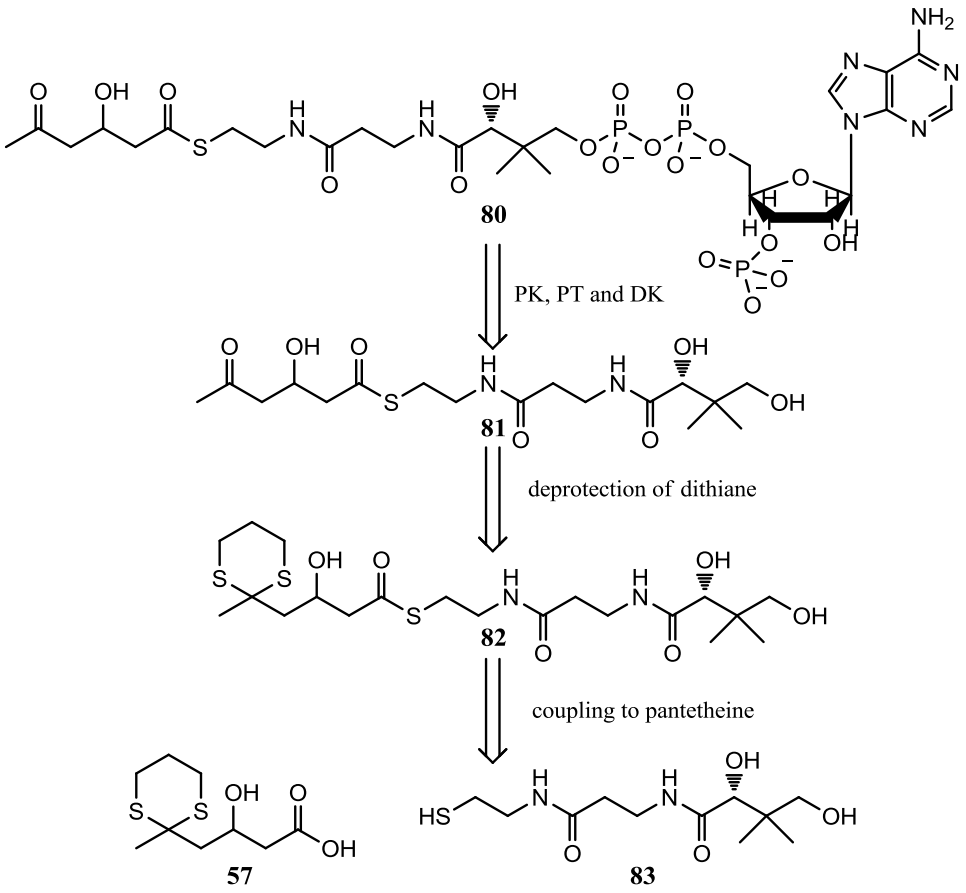
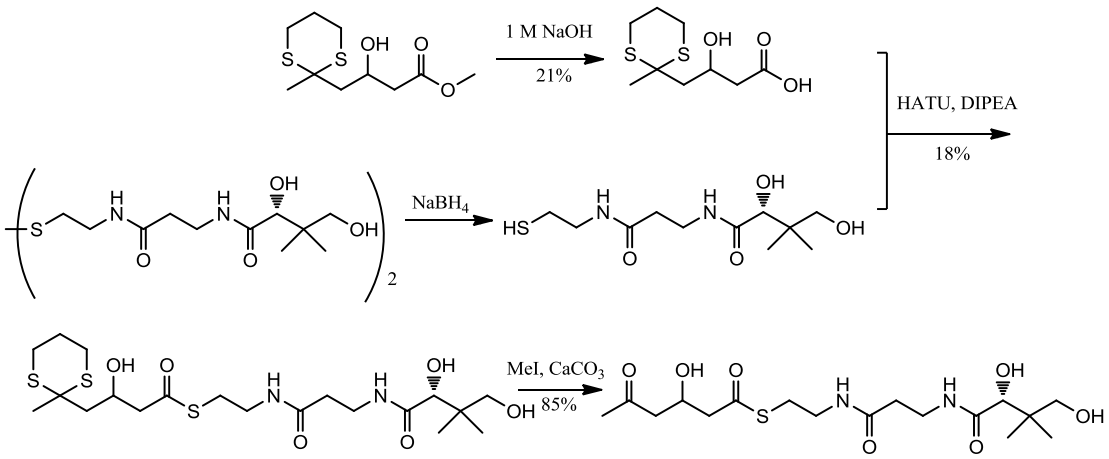


Figure 4.12: Retrosynthetic analysis of racemic reduced triketide CoA (**80**) to protected racemic reduced triketide **57** and pantetheine (**83**).

To test this synthetic route previously synthesised racemic ester (**74**) was converted to the racemic acid (**57**) using 1 M sodium hydroxide at room temperature and purified by column chromatography. D-Pantetheine (**83**) was obtained by reduction of D-pantethine (**84**) with sodium borohydride. Coupling of acid **57** to pantetheine was carried out using HATU as coupling reagent in the presence of DIPEA to give the thioester (**82**) in 18% yield after purification by preparative HPLC. The dithiane protecting group was removed using methyl iodide and calcium carbonate and the product purified again using preparative HPLC to give ketone (**81**) in 85% yield (Figure 4.13).



Scheme 4.13: Synthesis of reduced triketide-pantetheine **81**.

E. coli stocks containing plasmids coding for the hexa-histidine- tagged enzymes PK, PT and DK were kindly donated by Dr Manuela Tosin. All three enzymes were expressed in liquid culture with induction by IPTG and purified by nickel affinity chromatography.

The pantetheine compound (**81**) was incubated overnight with PK, PT and DK in the presence of ATP. After precipitation of the proteins with chloroform, the aqueous layer was analysed using LC-MS but no CoA derivative was detected. This could be due to an old sample of pantetheine-triketide being used.

4.2.3 *Apo* 6-MSAS expression and purification

The 6-MSAS plasmid pKOS007-109 previously constructed in section 2.2.1 was used in this study. This plasmid was transformed into *E. coli* BL21 (DE3) for expression. It was found that the conditions used to obtain the *holo* synthase, namely induction with IPTG at an OD_{600} of 1 then incubation at 30 °C for 24 hours, did not give any high molecular weight bands in any fraction observed in purification. Lower incubation temperatures after induction were tried, along with the addition of EDTA (5 mM final concentration) to the cell lysis buffer to combat the activity of metalloproteinases. Induction at both 16 °C and 20 °C gave strong bands of the correct molecular weight on SDS-PAGE, but the 16 °C 22-40 % ammonium sulphate fraction contained fewer

additional bands. 16 °C was then used as the induction temperature for all future experiments.

Apo 6-MSAS was purified by ammonium sulphate precipitation as for the *holo* synthase in section 2.2.1 (Figure 4.14). The 22-40 % pellet was then used directly for assays or further purified by FPLC.

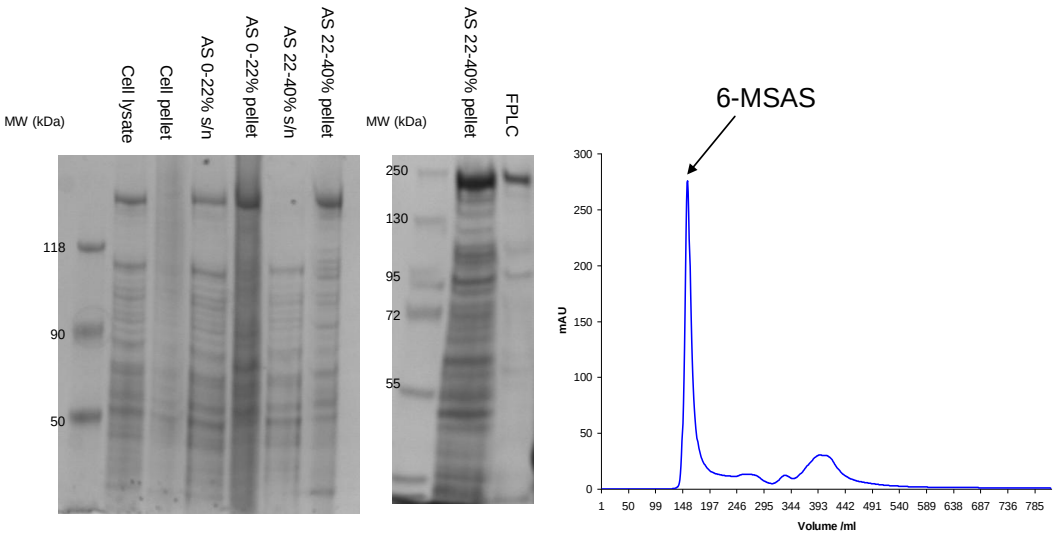


Figure 4.14: SDS-PAGE showing the expression of *apo* 6-MSAS (approx. 190 kDa) in the cell lysate, and the presence of 6-MSAS in fractions from ammonium sulphate purification and FPLC purification and a gel filtration trace of *apo* 6-MSAS.

Gel filtration of *apo* 6-MSAS on Sephadex 200 using the buffer system of Richardson *et al.*[85] gave a single peak at elution volume 157 ml. This was collected and concentrated to give approximately 1 mg of 6-MSAS from 110 mg of ammonium sulphate pellet. Scaling up, this gives an approximate yield of 10 mg of *apo* 6-MSAS per litre of culture.

4.2.4 Activity of *apo* 6-MSAS

Apo 6-MSAS was assayed for activity by looking for 6-MSA production in the culture medium and by *in vitro* fluorimetry assay of the purified protein.

50 ml of harvested culture medium was acidified to pH 1 using hydrochloric acid and extracted using ethyl acetate. The ethyl acetate extracts were evaporated and the residue dissolved in methanol for LC-MS analysis. 6-MSA production was determined with regard to an authentic 6-MSA standard of known concentration.

Surprisingly, production of approximately 2 mg /litre of 6-MSA was observed, compared to 17-25 mg/l for the *holo* synthase. This is unexpected as *E. coli* phosphopantetheinyltransferases do not generally act upon heterologously expressed PKSs. This suggests that a very small amount of 6-MSAS is converted to the active *holo* form *in vivo*. Production of 6-MSA during heterologous expression of the 6-MSAS homologue ATX was also observed for the *apo* enzyme.[89]

Flurorimetry assays which directly measure the production of 6-MSA were carried out under the same conditions used by Spencer and Jordan for the native synthase[56]. No production of 6-MSA could be seen for either the ammonium sulphate pellet or the FPLC-purified protein over the control run which did not contain malonyl-CoA. This is as expected for an *apo* synthase, which lacks the phosphopantetheine arm which tethers the growing polyketide to the enzyme, but contradicts the *in vivo* result.

4.2.5 Addition of a phosphopantetheine arm to *apo* 6-MSAS

In order to attach the synthesised triketide-CoA on to 6-MSAS for elaboration into 6-MSA, it is necessary to use a phosphopantetheinyltransferase (PPTase). This will transfer the phosphopantetheine section of CoA on to the active site serine in the ACP domain. The PPTase Sfp from *B. subtilis* has been extensively used to transfer acyl-CoAs on to ACP and PCP domains.[117]

Apo 6-MSAS was incubated with Sfp (donated by Dr Yuhui Sun) and CoASH under standard conditions used within our group (pH 6.5, 1 hour at 37 °C). 6-MSA production was assayed using fluorimetry and no additional activity could be seen

over the control which contained no PPTase. This was unexpected as Sfp has been used extensively to phosphopantetheinylate ACPs *in vitro* and has also been coexpressed with 6-MSAS *in vivo* to yield active *holo* synthase.[59] Additionally, recent work by Lee *et al.* has shown that a fluorescent CoA derivative is transferred on to heterologously expressed *apo* 6-MSAS using Sfp, although no activity studies have been performed using this method.[118]

4.2.5.1 MtaA

It was decided to use MtaA[84], the PPTase from *S. auriantica* which had been successfully coexpressed with 6-MSAS in *E. coli* to give active synthase (see section 2.2.1). MtaA was cloned into the pET28a (+) vector by Dr Fanglu Huang and the resultant plasmid was transformed into *E. coli* BL21 Codon Plus RP. His_6-tagged MtaA was expressed in LB medium under IPTG induction before purification by nickel affinity chromatography. Essentially pure His-MtaA was obtained in the 200 mM imidazole elute in a yield of 27 mg/litre of culture. Purified His-MtaA was analysed by LC-MS. The spectrum revealed a single peak of retention time 14.7 minutes, which was deconvoluted to give a protein of mass 33 777 Da. This is in good agreement with the predicted mass for MtaA + His_6 tag with loss of N-terminal methionine then N-terminal gluconylation[119] of 33782 Da (Figure 4.15).

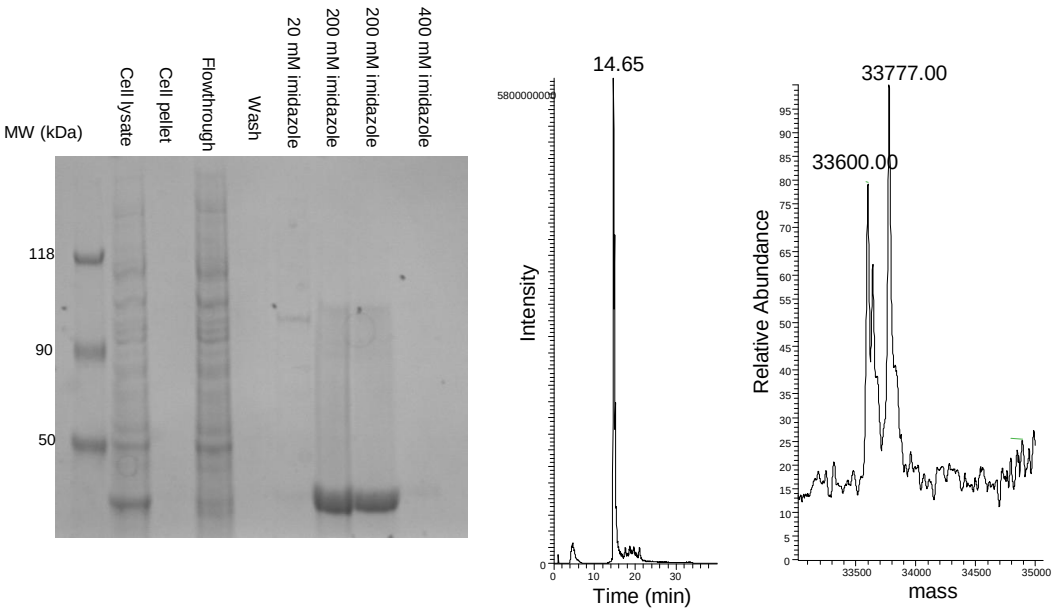


Figure 4.15: SDS-PAGE analysis of His-MtaA purification fractions and LCMS trace of His-MtaA that deconvolutes to a mass of 33 777 Da.

In order to check that the N-terminal His-tag did not interfere with activity, the PPTase activity of purified His-MtaA was assayed by mass spectrometry.

The discrete ACPs ChlD2 (from the chlorothricin gene cluster of *Streptomyces antibioticus*) and Con2* (from the concanamycin gene cluster of *Streptomyces neyagawaensis*) were donated by Dr Yuhui Sun (Department of Biochemistry). They are readily phosphopantetheinylated by Sfp and the reaction can be monitored by LC-MS. They have previously been successfully phosphopantetheinylated using Sfp and glyceryl-CoA.[120] The required ACP was incubated with MtaA at room temperature for one hour in conditions used for its reaction with Sfp (10 mM $MgCl_2$, 200 mM sodium phosphate, pH 7.5), and the conditions used by Lee *et al.* (10 mM $MgCl_2$, 75 mM Tris-HCl, pH 8.8). The assay mixture was then analysed by LC-MS.

In the phosphate buffer, Con2* showed a loss of a peak at 18.3 mins, that deconvoluted to a mass of 12,102 Da, and a gain of a peak at 17.2 mins, which deconvoluted to a mass of 12,442 Da. This gain of 340 mass units is due to attachment of the phosphopantetheine arm. Under the conditions used by Lee *et al.,*

the *apo* ACP eluted at 25.8 mins, while the *holo* ACP had a retention time of 17.4 mins. Again, a gain of 340 mass units was observed after deconvolution (Figure 4.16).

ChlD2 in phosphate buffer showed a no change in retention time of ACP peaks to the control experiment lacking MtaA, but deconvolution showed that the MtaA experiment gave a mass of 10,471 Da, while the control experiment gave a mass of 10,131 Da. This difference of 340 Da indicates that a phosphopantetheine arm has been added. At pH 8.8, ChlD2 exhibited a small change in retention time between the putative *apo* form (R_t = 19.5mins) and the putative *holo* form (R_t = 19.0 mins). Deconvolution showed a mass gain of 341 Da, from 10,130 Da to 10,471 Da, which again indicates attachment of the phosphopantetheine arm (Figure 4.16).

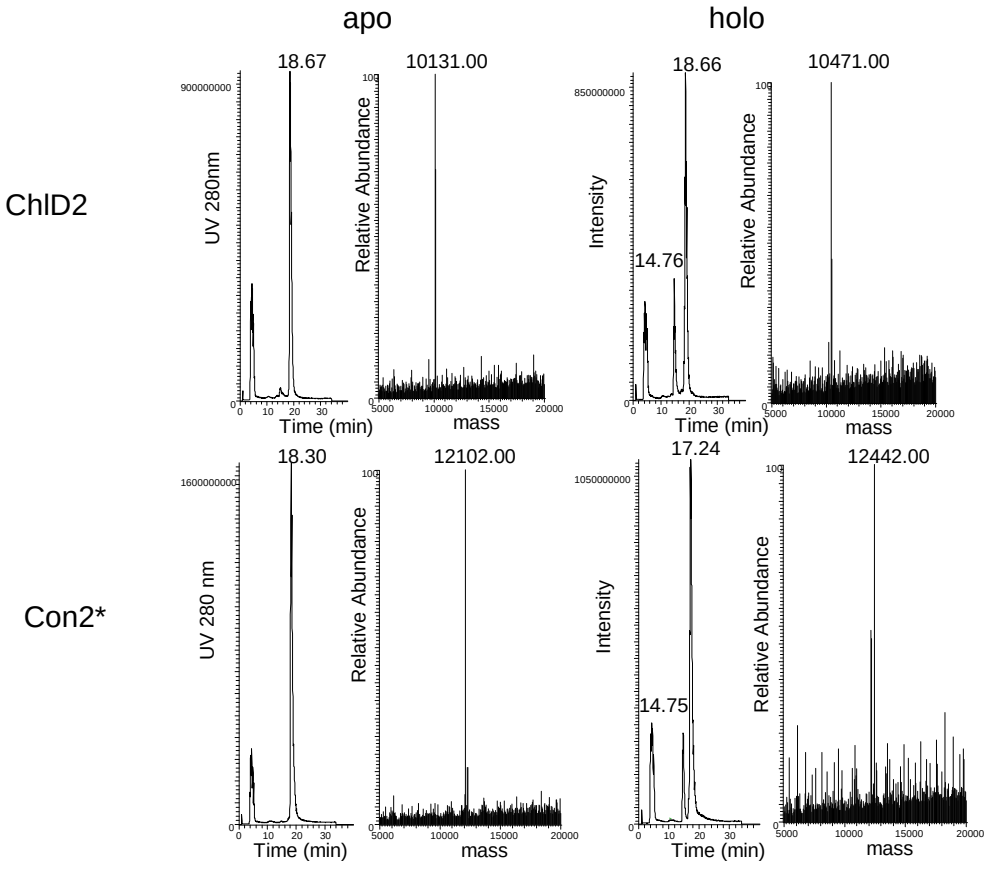


Figure 4.16: LC-MS traces for ChlD2 and Con2* before and after incubation with His-MtaA (6 μg), $MgCl_2$ (5 mM) and CoASH (0.55 mM) in phosphate buffer (100 mM sodium phosphate, pH 7.5).

Both Con2* and ChlD2 were completely converted from the *apo* form to the *holo* form as determined by mass spectrometry, proving that His-MtaA was active *in vitro.*

To ascertain whether the His-tag would interfere with the *in vivo* activity of MtaA the plasmid encoding for His_6-tagged MtaA and kanamycin resistance (pET28a(+)-MtaA), and the plasmid encoding for 6-MSAS and carbenicillin resistance (pKOS007-109) were cotransformed into *E. coli* BL21 and coexpressed in LB medium with IPTG induction.

Holo 6-MSAS was harvested and purified from a 1 litre culture as in section 2.2.1. The activity was assayed by fluorimetry and showed similar specific activity to the *holo* 6-MSAS that had been co-expressed with non-His-tagged MtaA.

Purified His-MtaA was used to phosphopantetheinylate *apo* 6-MSAS *in vitro*. MtaA was incubated with *apo* 6-MSAS (as the ammonium sulphate pellet), $MgCl_2$ and CoASH under conditions that had been used within the group for phosphopantetheinylation of an ACP using Sfp, and the conditions of Lee *et al*. The reaction mixture was used directly in the fluorimetry assay to ascertain if activity had been conferred. In both cases, no activity was seen over the control samples which lacked MtaA.

In order to ascertain whether high concentrations of ammonium sulphate inhibited the reaction with 6-MSAS, *apo* 6-MSAS was desalted using a PD-10 column, concentrated and then incubated with MtaA and CoASH as before. No activity could be seen when the enzyme was assayed by fluorimetry.

As MtaA has been shown to be active in the conditions used for phosphopantetheinylating 6-MSAS, but no active 6-MSAS was found from its incubation with intact *apo* 6-MSAS, this leads to two possibilities. Either MtaA is incapable of phosphopantetheinylating 6-MSAS *in vitro*, or MtaA does transfer the phosphopantetheine arm to 6-MSAS but the enzyme is misfolded or unfolds during phosphopantetheinylation and is thus still inactive.

4.2.5.2 SePptII

To rule out the possibility that MtaA was unable to phosphopantetheinylate 6-MSAS *in vitro* due to its inability to activate an intact megasynthase, it was decided to express and test a PPTase which is known to perform this role *in vitro*. SePptII from *S. erythraea* transfers the phosphopantetheine arms onto the ACPs present in the DEBS megasynthases. It had previously been identified and expressed by Weissman *et al.* and used to phosphopantetheinylate the ACP-TE didomain and DEBS 3 from the DEBS assembly line *in vitro*.[121]

pKJW152, which encodes for SePptII with an N-terminal His tag from pET28b(+), was obtained from Shilo Dickens (Department of Biochemistry) and transformed into *E. coli* BL21 Codon Plus RP.

Expression of SePptII was carried out according to literature procedure[121] and SePptII was purified by nickel affinity chromatography. The protein eluted was of approximately 60% purity as estimated by SDS-PAGE gel (Figure 4.17).

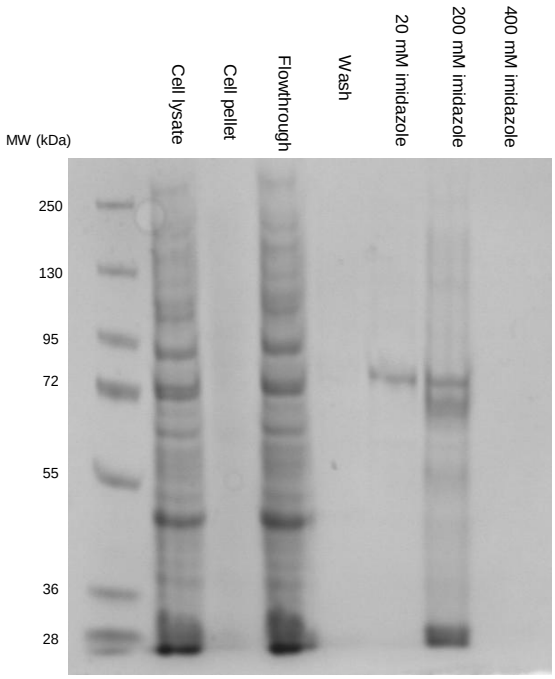


Figure 4.17: SDS-PAGE of SePptII (approx. 27 kDa) purification fractions.

SePptII was assayed for its ability to transfer a phosphopantetheine arm from CoASH to the discrete ACPs Con2* and ChlD2. After incubation at room temperature for 2 hours, the assay mixture was analysed by LC-MS. It was found that SePptII was able to phosphopantetheinylate Con2* but no *holo* ChlD2 was observed under the same conditions.

To examine whether SePptII could phosphopantetheinylate 6-MSAS *in vivo*, pKJW152 was cotransformed into *E. coli* BL21 (DE3) with pKOS007-109 which codes for the 6-MSAS gene. 6-MSAS was purified by ammonium sulphate precipitation as previously described and activity was assayed by fluorimetry. On incubation with acetyl-CoA, malonyl-CoA and NADPH, 6-MSA production was observed at a specific activity which was comparable to that of 6-MSAS coexpressed with MtaA.

Incubation of *apo* 6-MSAS with SePptII and CoASH at both pH7.5 and pH 8.8 did not lead to any activity as seen by fluorimetry.

4.2.4.3 Co-sonication

As attempted phosphopantetheinylation of *apo* 6-MSAS with three different PPTases had been unsuccessful, it invited the possibility that *apo* 6-MSAS could be denaturing during purification. In order to assess whether the instability of the *apo* synthase was due to an inherent misfolding of the protein during its production or an artefact of purification, co-sonication of *apo* 6-MSAS and MtaA was performed. If purification were the cause of the instability, it would be expected that the MtaA would be able to phosphopantetheinylate the 6-MSAS and produce active *holo* synthase.

Apo 6-MSAS and MtaA were expressed, as previously described, in separate cultures. After resuspension of the cell pellets, the two cultures were combined before sonication. Purification of 6-MSAS was then carried out as before. The ammonium sulphate pellet was assayed for activity using fluorimetry. Under these conditions, no production of 6-MSAS could be detected, which suggests that the protein is misfolded *in vivo* rather than becoming inactive during the purification steps. As successful phosphopantetheinylation was observed by Lee *et al.* with heterologous expression of 6-MSAS in *S. cerevisiae*,[118] this might be an improved system for obtaining the *apo* synthase.

4.2.6 Limited proteolysis

Since *in vitro* phosphopantetheinylation of *apo* 6-MSAS was unsuccessful, it was decided to investigate the structural differences between the *holo* and the *apo* forms using limited proteolysis to see if this showed any significant difference between the two.

4.2.6.1 Hydroxylamine

Hydroxylamine cleaves peptide bonds between asparagine and glycine residues by cyclisation of the asparaginyl residue with a backbone amide of the neighbouring

glycine residue.[122] Cleavage of 6-MSAS by hydroxylamine is predicted to give six cutting sites, yielding eight fragments.

Both *holo* and *apo* FPLC-purified 6-MSAS were incubated with 1.8 M hydroxylamine at 37 °C for four hours, with samples taken at intervals and analysed by SDS-PAGE (Figure 4.18). Although both synthases show a reduction in the amount of intact enzyme over time, there are no new bands appearing with lower molecular weight. This may be due to a lack of reaction with the hydroxylamine. Normally cleavage is carried out under denaturing conditions (6 M guanidinium chloride) but this would not allow the structural differences between the *holo* and *apo* synthases to be probed. It may also be that 6-MSAS unfolds under the conditions at which the reaction occurs, so that the two synthases are structurally the same and no differences in the structure of the ACP can be observed.

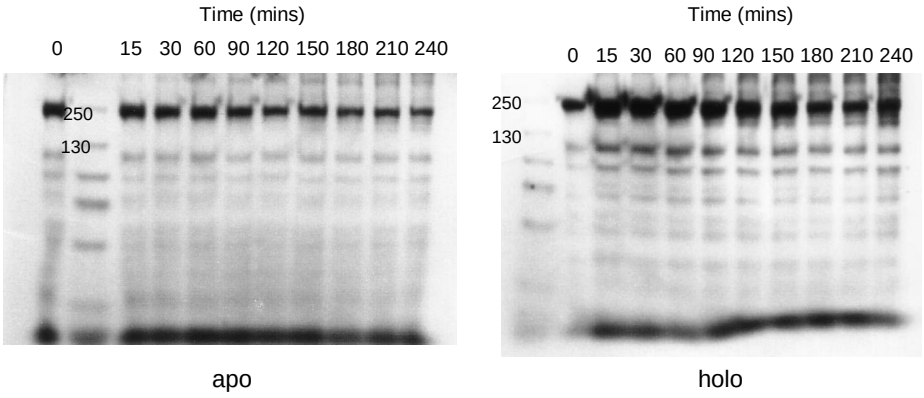


Figure 4.18: SDS-PAGE analysis of 6-MSAS incubated with hydroxylamine over 4 hours.

4.2.6.2 Trypsin

Trypsin was previously used to investigate the structural organisation of 6-MSAS isolated from *P. patulum*, although the experimental details were not available for this study.[52] In limited proteolysis studies on DEBS megasynthases it was found that cleavage occurred in the linker regions between domains.[123]

Sequence analysis of 6-MSAS predicted 146 cleavage sites. Incubation of *holo* and *apo* 6-MSAS with trypsin (100:1 wt/wt) showed a significant difference between the two protein forms. The *holo* enzyme formed a band at around 150 kDa slowly over 45 minutes as the 190 kDa band disappeared. Surprisingly, the *apo* enzyme formed a band at the same molecular weight within 1 minute of addition of trypsin, with concomitant loss of the 190 kDa band. The 150 kDa band in both cases was resistant to further proteolysis over the 60 minute timescale of the reaction (Figure 4.19).

Decreasing the concentration of trypsin to 250:1 and 500:1 did not slow the observed rate of cleavage of the *apo* enzyme. In both cases there was essentially full formation of the 150 kDa band within one minute. Increasing the concentration of trypsin to 50:1 did not lead to increased cleavage of the 150 kDa band, which remained resistant to further proteolysis.

Analysis of the 100:1 mixture by SDS-PAGE on an 18% acrylamide gel revealed a band of ~10 kDa was formed by digestion of both *apo* (after 1 minute) and *holo* synthase (after 5 minutes).

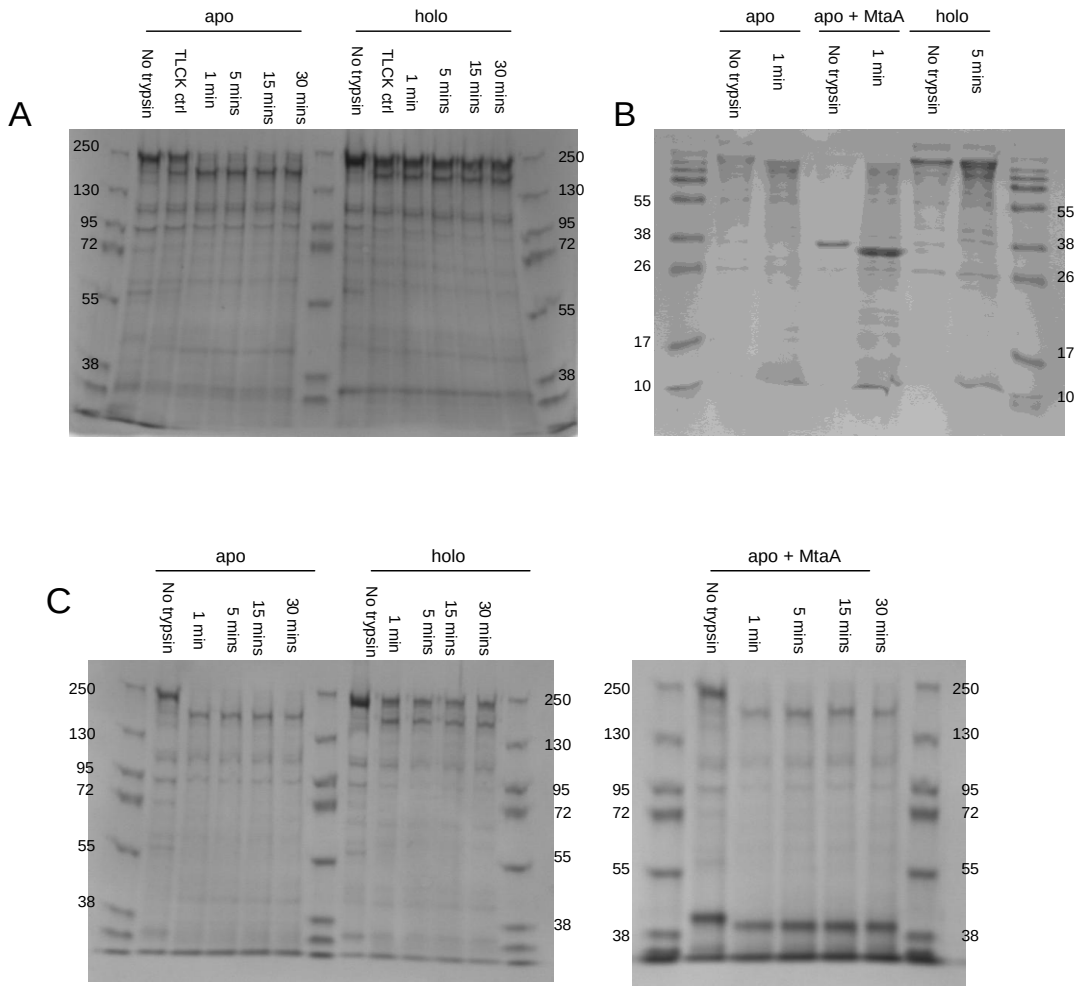


Figure 4.19: A: 6-MSAS/trypsin (100:1 wt/wt) digestion over 30 minutes. The lane labelled TLCK ctrl contains 6-MSAS/trypsin with TLCK added immediately. B: 6-MSAS/trypsin (100:1 wt/wt) after 1 minute of reaction (*apo* and *apo* + Sfp) or 5 minutes of reaction (*holo*) analysed on an 18% gel. C: Digest of *holo*, *apo* and *apo*+MtaA 6-MSAS with trypsin (50:1 wt/wt) over 30 minutes. All reactions performed in NaH_2PO_4 100 mM, EDTA 1 mM, DTT 2 mM, glycerol 50%, pH 7.1 and aliquots at each time point were quenched with TLCK (1 mM, final concentration).

To investigate the phosphopantetheinylation of *apo* 6-MSAS, *apo* synthase was incubated as before with MtaA for 1 hour and the mixture digested with trypsin (100:1) over 1 hour. SDS-PAGE analysis at time points revealed that the mixture acted in the same way as the *apo* synthase, with a fragment cleaved within 1 minute.

This suggests that phosphopantetheinylation does not take place in the reaction mixture.

Both the 150 kDa fragment and the 10 kDa fragment were submitted for N-terminal sequencing. The 150 kDa fragment corresponded to the N-terminus of 6-MSAS, although missing the first 12 amino acids after cleavage adjacent to lysine-13, for *apo*, *apo*+MtaA and *holo* 6-MSAS. As the missing N-terminal sequence cannot account for the large change in apparent mass of the band, there must also be cleavage at the C-terminus of a more substantial nature.

The 10 kDa band did not give any distinct sequence information for any of the three bands submitted and it was suggested that it was in fact a mixture of sequences rather than the product of a single cleavage.

In the *apo*+MtaA mixture, a strong band at 40 kDa was seen. N-terminal sequencing revealed this to be from the cleavage of His-MtaA in the N-terminal His-tag (predicted mass 31 852 Da).

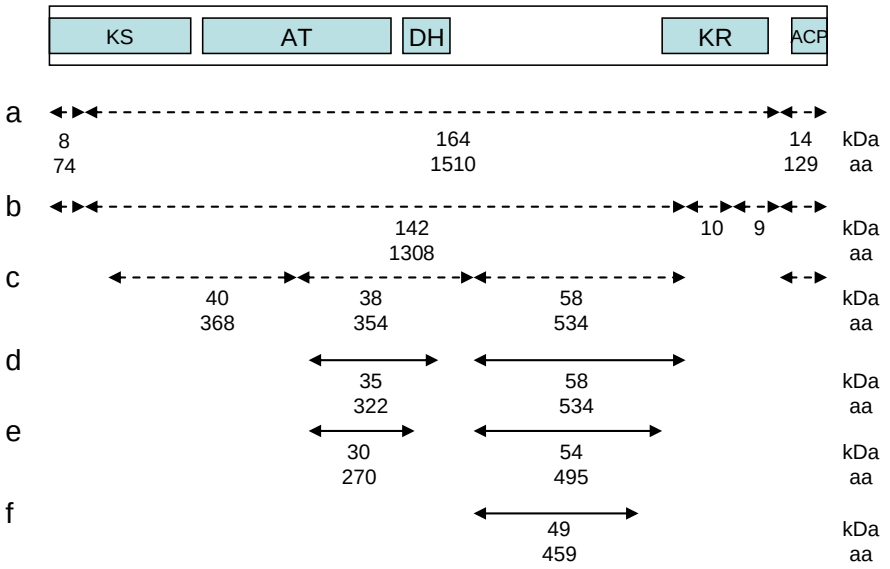


Figure 4.20: Partial peptide map of 6-MSAS from *P. patulum* after trypsin digestion. Dashed peptides are formed transiently within 60 s (a, b and c). After 2 minutes, further fragments are observed (d and e). After 5 minutes only one major fragment remains (f). Figure redrawn from Shoolingin-Jordan and Campuzano, in *Comprehensive Natural Products*, 345-365.[52]

The reported limited proteolysis results for 6-MSAS from *P. patulum* are shown in figure 4.20. These show rapid cleavage of C-terminal fragments including the ACP, before further digestion of the central domains.

The results broadly concur with those observed for 6-MSAS expressed in *E. coli*, which rapidly loses C-terminal fragments. The mixed fragments proposed for the 10 kDa band could be from the 10, 9 and 14 kDa bands. The 6-MSAS expressed in *E. coli* is not, however, digested further in the large central fragment, which is unexpected given the results from the native enzyme. As experimental conditions for the limited proteolysis of the *P. patulum* enzyme were unavailable, it is unknown whether this is due to different trypsin concentrations or reaction conditions.

The difference between *apo* and *holo* 6-MSAS is striking. The *apo* enzyme loses the C-terminal fragment(s) within 1 minute, while the *holo* enzyme is stable for up to 45 minutes under the same conditions. This implies a greater degree of conformational flexibility at the C-terminus of the *apo* enzyme, suggesting that the ACP is either

unfolded or very flexible. This could be why the phosphopantetheinylation is unsuccessful as the PPTases used might not be able to recognise the ACP domain.

4.2.7 ACP expression

As MtaA and Sfp were unable to phosphopantetheinylate *apo* 6-MSAS *in vitro*, it was decided to simplify the problem by cloning out the ACP from 6-MSAS. This ACP could be assayed for phosphopantetheinylation *in vitro* using LC-MS.

The cloned ACP could also be used with either a truncated 6-MSAS which was missing its ACP or *apo* 6-MSAS to reform active enzyme, and so circumvent the problem of phosphopantetheinylation of the megasynthase *in vitro*.

4.2.7.1 N-terminal His-tagged ACP

To clone out the ACP, PCR was carried out on pKOS007-109 using primers designed by Dr Fanglu Huang. The PCR product was purified by preparative agarose gel electrophoresis and digested with NdeI and EcoRI. The digested DNA was purified again by preparative agarose gel electrophoresis. This product was ligated with pET28a(+) which had previously been digested with NdeI and EcoRI, using T4 ligase and incubated overnight. The ligation mixture was transformed into NovaBlue competent cells and plated out onto LB/Kanamycin plates. Colonies were screened for the ACP insert using colony PCR. Of the 12 colonies screened, all showed a band at ~300 bp which corresponded to a successful ligation (Figure 4.21).

The identity of the plasmid was confirmed by sequencing and the plasmid was transformed into *E. coli* BL21 (DE3). The His_6-tagged ACP was expressed in LB medium with IPTG induction and the cell lysate was purified by nickel affinity chromatography.

Initially, no *apo* ACP could be seen on an SDS-PAGE gel, so co-expression with MtaA was attempted in case the *apo* ACP was misfolded and proteolysed *in vivo*. *E. coli* BL21 was transformed with both pET-28a(+) containing the ACP and pSU-MtaA containing untagged MtaA. Conditions were repeated for the expression of the intact synthase: ATCC-765 medium + 10% glycerol followed by sonication with lysozyme, benzamidine and PMSF.

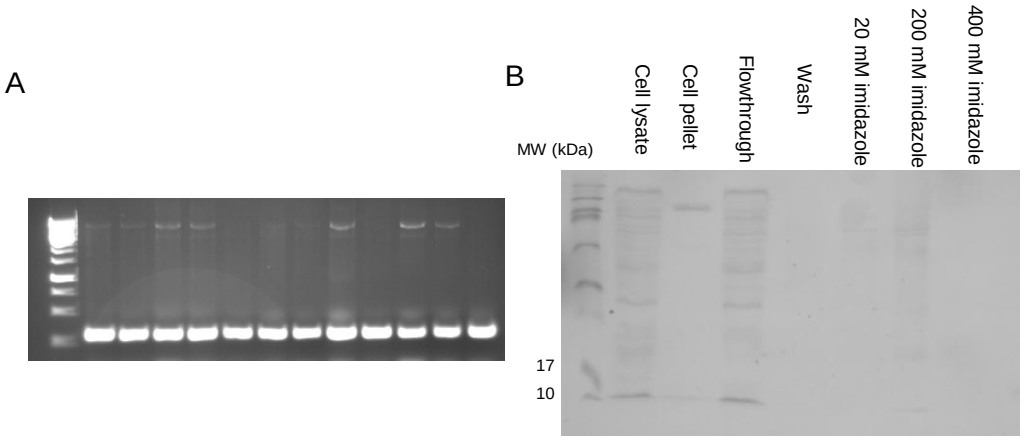


Figure 4.21: A: Colony PCR of ACP tranformants shows the presence of the plasmid containing the ACP. B: SDS-PAGE of nickel affinity chromatography fractions from purification of *holo* N-His-ACP (predicted mass 11 kDa).

On analysis by SDS-PAGE gel, it was found that neither *holo* nor *apo* N-His-ACP bound significantly to the His-bind column, instead being mainly found in the flowthrough.

4.2.7.2 C-terminal His-tagged ACP

The same forward primer as for the N-terminal ACP was used, but a new reverse primer was designed which incorporated an EcoRI site and changed the stop codon to a glycine residue, to enable the C-terminal tag to be incorporated. PCR was carried out on pKOS-109 as before and the PCR product was found to be of the correct size. The purified insert was digested with NdeI and EcoRI and ligated into pET-20b which had been digested with NdeI and EcoRI. The plasmid was transformed into NovaBlue competent cells and colonies were picked and screened by colony PCR.

pET20b(+)-ACP was transformed into *E. coli* BL21 and expressed in LB medium under IPTG induction. Purification was carried out by nickel affinity chromatography. Analysis of purification fractions by SDS-PAGE gel revealed that, like the ACP bearing the N-terminal His-tag, most of the ACP was found in the column flowthrough (Figure 4.22).

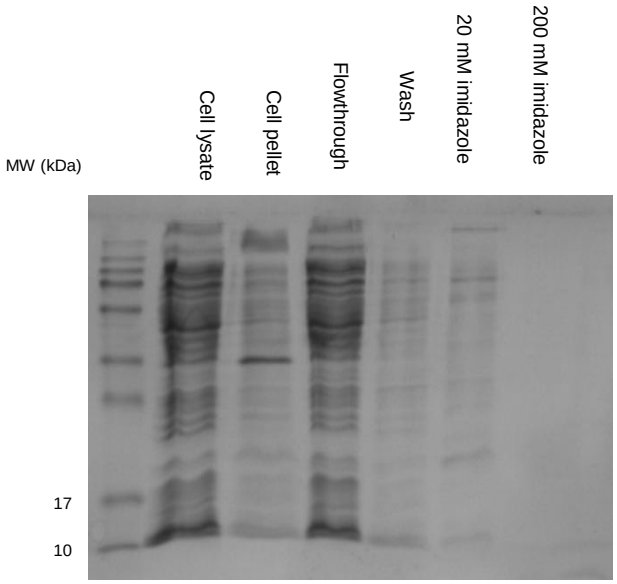


Figure 4.22: SDS-PAGE analysis of nickel affinity chromatography fractions for *apo* C-His-ACP (predicted mass 11 kDa).

4.2.7.3 N-terminal and C-terminal His-tagged ACP

Since no single His-tag successfully bound to the nickel affinity column, it was decided to express the 6-MSAS ACP domain with both the N- terminal and C-terminal His_6-tags. This is a similar approach to that taken with the ACP domain from the fungal PKS nosolorinic acid synthase.[124]

The plasmid coding for the doubly tagged ACP can be constructed from pET20b(+)-ACP and pET28a(+) (Figure 4.23). Digestion of pET28a(+) with BglII and NdeI gave a 150 base pair fragment which codes for the T7 promoter and N-terminal His-tag. Digestion of pET20b(+)-ACP with BglII and NdeI excised the T7 promoter. The T7 and

His-tag containing fragment from pET28a(+)-ACP was then ligated with the digested pET20b(+)-ACP to give pHKP1 containing the 6-MSAS ACP flanked by N-terminal and C-terminal His-tags. The ligation mixture was transformed into NovaBlue cells and transformants screened by sequencing.

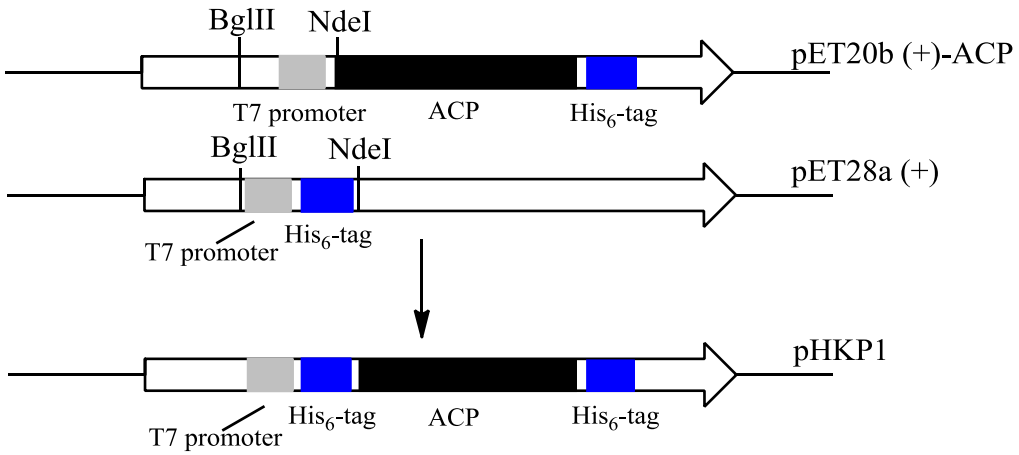


Figure 4.23: Formation of pHKP1 from pET20b(+)-ACP and pET28a(+) by digestion with BglII and NdeI, followed by ligation.

pHKP1 was transformed into *E. coli* BL21 and expressed in liquid culture with IPTG induction. The cell lysate was purified by nickel affinity chromatography and the fractions analysed by SDS-PAGE. A greater amount of ACP was seen in the 200 mM imidazole column eluate than ACP expressed with a single His-tag, however a substantial amount was still present in the flowthrough and wash steps (Figure 4.24). The 200 mM eluate also contained a large amount of non-specifically bound proteins.

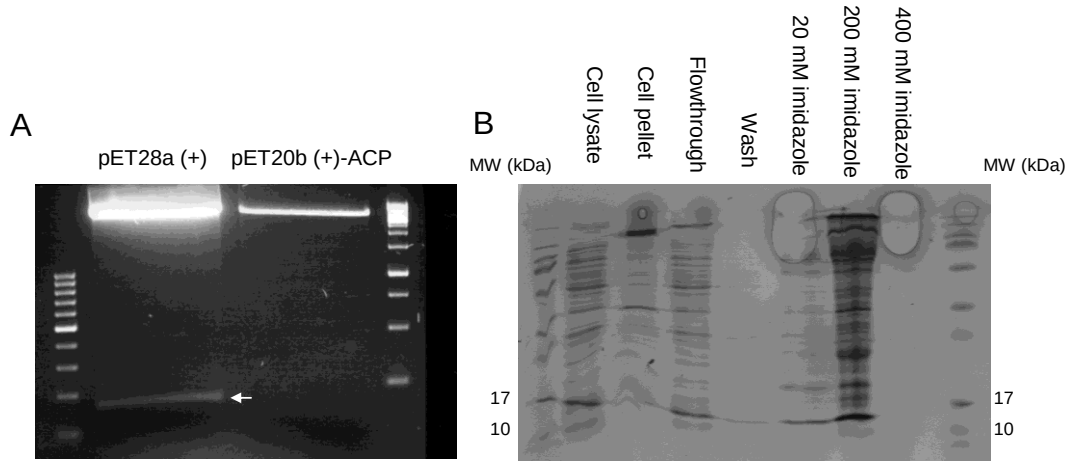


Figure 4.24: A: pET28a(+) and pET20b(+)-ACP digested with BglII and NdeI. B: SDS-PAGE of doubly His-tagged ACP (predicted mass 13 kDa).

The 200 mM elute could be purified by FPLC to remove proteins of larger molecular weight and the resulting fractions analysed by LC-MS to confirm the presence of doubly-tagged ACP.

4.3 Future work

The experiment of Lee *et al.* using a fluorescent-CoA analogue could be repeated on *apo* 6-MSAS expressed in *E. coli*. This would provide a direct measure of whether the phosphopantetheinylation reaction takes place *in vitro*. If it does not, then expression in an alternative expression organism, such as *S. cerevisiae* could be considered.

Alternative tags such as the glutathione-S-transferase (GST) tag or the maltose binding protein (MBP) tag could be used to express the ACP, with the tag being cleaved before any assay of ACP activity. The purified ACP could be incubated with a phosphopantetheinyltransferase and CoASH. The assay mixture could be easily analysed by LC-MS, and the ability of PPTases to act on the heterologously expressed ACP domain investigated. If the reaction with CoASH is successful, then the reduced triketide CoA (**80**) could be loaded on to the ACP and the product triketide-ACP observed by LC-MS.

Holo ACP could react in trans with *apo* 6-MSAS. Varying concentrations of *holo* ACP could be used in the standard fluorimetry assay for 6-MSA production to see if the separate ACP domain was able to successfully replace the inactive one linked to the enzyme.

It is possible that the link of the *cis* ACP to the rest of the megasynthase makes reaction in trans difficult. One way of getting around this would be to express a truncated version of 6-MSAS which lacks the terminal ACP (procedure designed by Dr Fanglu Huang).

pKOS007-109, which codes for the 6-MSAS gene could be digested with AflII which cuts in the linker region between the KR and ACP domains. The sticky end produced could be filled in to produce a blunt end which can be ligated in to an EcoRV site. The plasmid would then be digested again with NdeI to produce a DNA fragment containing the 6-MSAS gene lacking the ACP. This could then be ligated in to pET20b or pET29a which had been digested with NdeI and EcoRV (Figure 4.25). The plasmid could then be transformed in to NovaBlue for verification of the gene sequence and BL21 (DE3) for protein expression.

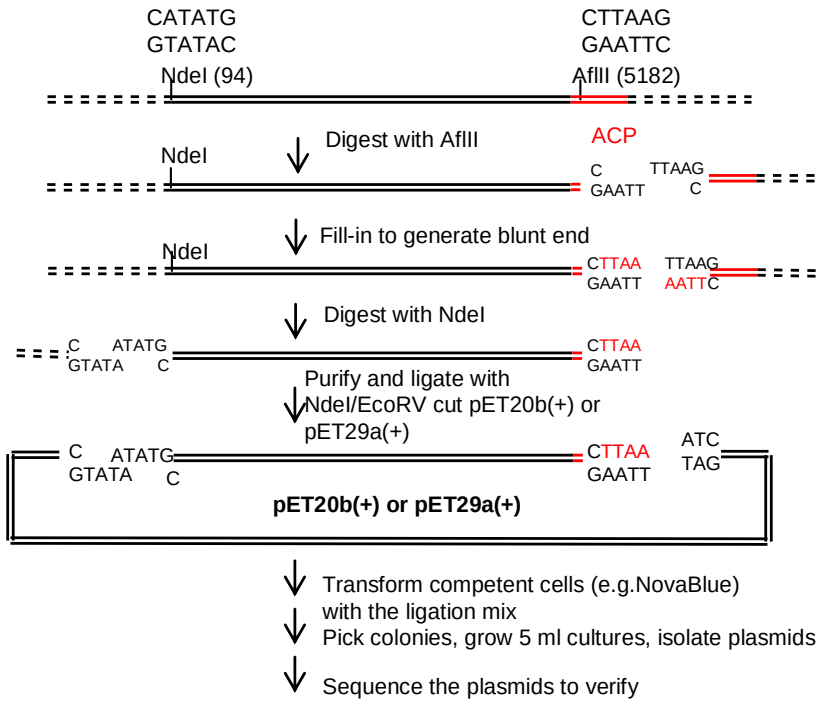


Figure 4.25: Proposed method for generating truncated 6-MSAS lacking the ACP domain.

The truncated 6-MSAS would be assayed with varying concentrations of *holo* ACP, malonyl CoA, acetyl CoA and NADPH to optimise the reaction conditions. One advantage of using the discrete ACP rather than the intact *apo* synthase would be that it would remove the single turnover nature of the proposed labelling experiment. The amount of triketide reacted would be dependent only on the concentration of loaded ACP, which could be many times the concentration of synthase.

4.4 Chapter Summary

Steps have been made toward the synthesis of the resolved reduced triketide CoAs **65** and **66**. The synthetic diastereomeric intermediates **76** and **77** have been resolved by column chromatography. The racemic mixture of these intermediates has been

coupled with pantetheine and elaborated to the triketide-coenzyme A by the action of three enzymes.

Apo 6-MSAS was heterologously expressed in *E. coli* and purified in a two step procedure. Attempts to phosphopantetheinylate the *apo* synthase using three different PPTases were unsuccessful, although all three enzymes were shown to be active *in vitro*. Co-sonication of the *apo* culture with a PPTase culture did not yield active synthase, implying that the ACP was misfolded *in vivo*.

Limited proteolysis studies on *apo* and *holo* 6-MSAS with hydroxylamine and thrombin found no difference between the two protein forms. Studies with trypsin however showed that the *apo* enzyme contained a C-terminal fragment which was rapidly cleaved, while the *holo* enzyme was cleaved at a much slower rate, although N-terminal sequencing was unable to identify the fragment in question.

The ACP from 6-MSAS was heterologously expressed in *E. coli* with N-terminal, C-terminal and N- and C-terminal His_6-tags. It was found that purification by nickel affinity chromatography was difficult as the ACP did not bind well to the column. Attempts to determine the stereochemistry of ketoreduction in 6-MSAS are ongoing, and may benefit from the use of the discrete 6-MSAS ACP with either the *apo* synthase or a truncated version.

CHAPTER 5

Materials and Methods

5.1 Chemicals

Ethyl 6-methylsalicylate was purchased from TCI Europe. All other chemicals were purchased from Sigma-Aldrich and were of analytical grade. HPLC-grade organic solvents were from Fisher Scientific, while HPLC water was from Rathburn. Dry solvents were from in-house stills. All water was deionised using a Milli-Q Gradient A10 Ultrapure water system.

5.2 Enzymes

Pfu DNA polymerase, *DreamTaq* polymerase, T4 DNA ligase, all restrictions enzymes, BSA, all associated buffers and dNTPs were purchased from Fermentas. Chymotrypsin, trypsin and lysozyme were purchased from Sigma-Aldrich.

5.3 Biological reagents

NADPH, dithiothreitol (DTT), isopropyl-β-D-thiogalactopyranoside (IPTG), kanamycin and carbenicillin were purchased from Melford. Agarose and Bradford reagent were purchased from Sigma-Aldrich. Luria-Bertani (LB) medium was purchased from Fisher Scientific. 30% acrylamide/bisacrylamide mix was purchased from Severn Biotech Ltd. DNA and protein molecular weight markers were purchased from

Fermentas. Chemically competent cells were donated by Dr Fanglu Huang and El-Habib Sahraoui. pET28a(+) and pET20b(+) were obtained from Novagen. pKOS012-65a and pKOS007-95 containing the 6-MSAS gene and appropriate vector were donated by Kosan Biosciences Ltd. pSU-MtaA was kindly donated by Dr Yanyan Li. pET28a(+)-MtaA, encoding MtaA with an N-terminal His_6 tag was kindly donated by Dr Fanglu Huang. pKJW152 containing the gene for SePptII with an N-terminal His_6-tag was obtained from Shilo Dickens, Department of Biochemistry, University of Cambridge. Cells containing the plasmids coding for the CoA biosynthetic enzymes PK, PT and DK were kindly donated by Dr Manuela Tosin. GeneJET PCR purification kits, GeneJET Gel Extraction kits and GeneJET Plasmid Preparation kits were purchased from Fermentas.

5.4 Bacterial strains

Escherichia coli Novablue (Novagen) was used for general cloning and *Escherichia coli* BL21(DE3) was used for protein expression.

5.5 Apparatus

PCR was carried out using an Eppendorf Mastercycler Gradient thermocycler with 0.5 ml PCR tubes. Agarose gels were run in a Mini-Sub Gel GT gel electrophoresis tank attached to a Model BRL 250EX power supply (Gibco). SDS-polyacrylamide gels were cast using an 80 mm by 70 mm 4 gel caster from Hoefer. SDS-PAGE was carried out using an EPS Model 301 (Amersham Pharmacia Biotech). Gels were visualised with a White/UV Transilluminator (UVP) and photographs taken with a GelCam (Polaroid). Sterile work was carried out in a BioMAT2 laminar flow hood (Thermo Electron Corporation). Buffers and media were sterilised in a KaVoKlave 2100 autoclave from Prestige Medical. pH was measured using a Philips PW9410 pH meter attached to a Mettler Toledo pH electrode. LB-agar plates were incubated in a Raven incubator oven (LTE Scientific) and liquid cultures in a Multitron II incubator shaker (Infors). Cells were lysed using a Vibra-Cell sonicator. Centrifugation was carried out in an Avanti J-26 XPI (Beckman Coulter) or a 5804R benchtop centrifuge (Eppendorf). Microcentrifugation was performed in a Biofuge Pico centrifuge

(Heraeus). Samples were lyophilised using a Drywinner MicroModulyo (Jencons-Plus) lyophiliser attached to a RV8 pump (Edwards). TLC analysis was performed on glass sheets precoated with silica gel 60 and plates were visualised using UV and a potassium permanganate dip. Flash column chromatography was carried out using silica gel 60 (0.040-0.630 mm, Merck). UV/Vis spectroscopy was carried out using a Cary 100 Bio UV-Visible Spectrophotometer (Varian). Fluorimetric assays were performed using a Cary Eclipse fluorimeter (Varian). The iBlot Dry Blotting system from Invitrogen was used to transfer protein from acrylamide gels to PVDF membrane.

5.6 Computational analysis and Bioinformatics

DNA sequencing data was viewed using FinchTV v1.4.0 developed by Geospiza, Inc. Restriction sites within DNA sequences were determined using NEBcutter v2.0 (http://tools.neb.com/NEBcutter2/).[125] Proteolytic sites in proteins were found using Peptide Cutter (http://expasy.org/tools/peptidecutter/). Sequence alignment was performed using ClustalW[115] accessed at EBI (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Protein sequences were accessed from UniProt Protein Knowledge Base (http://www.uniprot.org/)[126] and BLAST searches were also performed using this site. Other protein tools were used at the ExPASy proteomics server (http://expasy.org/tools/#proteome). Domain searches within a PKS sequence were carried out using the PKS/NRPS analysis at http://nrps.igs.umaryland.edu/nrps/.[127]

5.7 Analytical techniques

5.7.1 LC-MS

LC-MS analysis was performed using a Hewlett-Packard HPLC 1100 series coupled to a Finnigan MAT LCQ ion trap mass spectrometer fitted with an ESI source using Xcalibur™ Version 1.1.

MS of synthetic samples was carried out on a Luna 5u C18 250 x 2.00 mm column (Phenomenex) or a Synergi 4u Polar-RP 150 x 2.00 mm column (Phenomenex). Both columns were run at 0.3 ml/min in either water/acetonitrile + 0.1% TFA or water/methanol + 0.1% TFA gradient; 0-5 minutes, 0% organic, 5-30 minutes 0-100% organic, 30-38 minutes 100% organic, 38-40 minutes 100%-0% organic.

MS of protein samples was carried out on a Jupiter 5u C4 250 x 2.00 mm column (Phenomenex) run at 0.2 ml/min in water/acetonitrile + 0.1% TFA. The gradient was 0-3 minutes, 5-40% MeCN; 3-30 minutes, 40-95% MeCN; 30-37 minutes, 95% MeCN; 37-40 minutes, 95-5% MeCN.

5.7.2 HPLC

Semi-preparative HPLC was carried out on an Agilent 1100 HPLC fitted with a Synergi 4u Polar-RP 250 x 10.00 mm column (Phenomenex) using either a water/acetonitrile + 0.1% TFA or water/methanol + 0.1% TFA gradient from 0% organic solvent to 100% over 40 minutes at 2.5 ml/min. UV traces were observed at 210 nm, 254 nm and 280 nm.

Preparative HPLC was carried out on an Agilent 1200 HPLC fitted with a Luna C18 10u 250 x 21.20 mm column (Phenomenex) using a water/acetonitrile + 0.1% formic acid or water/methanol + 0.1% formic acid gradient from 0% organic solvent to 100% over 40 minutes at 20 ml/min. UV traces were observed at 210 nm, 254 nm and 280 nm.

5.7.3 LTQ Orbitrap

HPLC-HR-ESI-MS analyses of off-loading assays were performed on a Thermo Electron LTQ-Orbitrap. Samples were injected onto a Dionex Acclaim C18 PepMap 100 column (150 mm x 1.0 mm, 3 Dm), eluting with a linear gradient of 0% to 100% B in 28 min with a flow rate of 50 μl/min (A: 98% H_2O, 2% MeCN, 0.1% formic acid, B: 90% MeCN, 10% H_2O, 0.1% formic acid). The mass spectrometer was run in positive

ionization mode, scanning from m/z 100 to 1800, with the FTMS analyser resolution set at 60K; the collision energy was set to 30-35% for mass fragmentation.

5.8 Synthetic Chemistry

5.8.1 Synthesis of 6-MSA (1) standard[128]

Ethyl 6-methysalicylate (100 mg, 0.55 mmol) was refluxed in 1M NaOH (5 ml) for 2 hours. The solution was cooled and 10% sulphuric acid was added to precipitate a white solid. The mixture was filtered and the solid freeze dried. 6-Methylsalicylic acid was obtained as a white solid (41.4 mg, 50%). LCMS: R_t = 26.0 mins, $[M+H^+]$ = 152.94. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 10.94 (s, 1H, COO*H*), 7.27 (t, 1H, J = 8 Hz), 6.80 (d, 1H, J = 8 Hz), 6.70 (d, 1H, J = 8 Hz), 2.56 (s, 3H). R_f = 0.50 (3:1 ethyl acetate: methanol, spot blue under UV).

5.8.2 Synthesis of non-hydrolysable analogue 30[86]

5.8.2.1 Synthesis of 4-acetamidobutanoic acid (32)

γ-Aminobutyric acid (0.5 g, 4.85 mmol) and triethylamine (1.35 ml, 9.70 mmol) were suspended in dry methanol (30 ml) and cooled to 0 °C. Acetyl chloride (0.379 ml, 5.36 mmol) was added dropwise and the mixture stirred at 0 °C for 1 hour, then at room temperature overnight. The resulting colourless solution was evaporated under reduced pressure to leave crude **32** as a white solid (704 mg).

5.8.2.2 Synthesis of N-(4-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-yl)-4-oxobutyl)acetamide (33)

32 (704 mg, 4.85 mmol), Meldrum's acid (770 mg, 5.34 mmol) and DMAP (1.50 g, 7.28 mmol) were stirred in dry THF (100 ml) and cooled to 0 °C. EDC (1.02 g, 5.34 mmol) was added and the mixture stirred overnight at room temperature. The yellow solution was evaporated under reduced pressure and the resulting oil was redissolved in DCM. The organic layer was washed with 1 M aq. HCl, dried with

magnesium sulphate and evaporated under reduced pressure to give crude **33** as a yellow oil (925 mg).

5.8.2.3 Synthesis of methyl 6-acetamido-3-oxohexanoate (34)

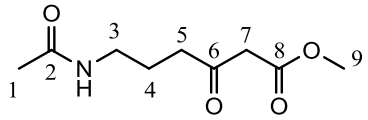


Figure 5.1: Numbering of **34** for NMR assignment.

33 (925 mg, 3.41 mmol) was dissolved in dry methanol (0.65 ml) and dry toluene (20 ml) and heated under reflux for 3 hours. The solvents were removed under reduced pressure to give **34** as a yellow oil.

Crude **34** (18 mg) was purified by preparative HPLC using a methanol/water + 0.1% formic acid gradient (0% methanol to 100% methanol over 40 minutes) to give **34** as a yellow oil (4.2 mg, 23%). R_t = 12.4 mins. LCMS: R_t = 15.1 mins, $[M+H^+]$ = 201.89. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.66 (br s, 1H, NH), 3.73 (s, 3H, *H-9*), 3.46 (s, 2H, *H-7*), 3.25 (q, 2H, J = 6 Hz, *H-3*), 2.61 (t, 2H, J = 6 Hz, *H-5*), 1.95 (s, 3H, *H-1*), 1.81 (qn, 2H, J = 7 Hz, *H-4*). ^{13}C (100MHz, $CDCl_3$): δ = 202.6 (*C-6*), 170.3 (*C-2*), 167.7 (*C-8*), 52.4 (*C-9*), 48.9 (*C-7*), 40.3 (*C-3*), 38.8 (*C-5*), 23.3, 23.2 (*C-4, C-1*).

5.8.2.4 Formation of 6-acetamido-3-oxohexanoic acid (30)

34 was dissolved in 6-MSAS assay buffer (*e.g.* 10 μl) to a final concentration of 60 mM. An equal volume of pig liver esterase solution (*e.g.* 10 μl, ammonium sulphate suspension, ≥150 units/mg protein) was added and the mixture was incubated at 37 °C for 2 hours. Chloroform (*e.g.* 20 μl) was added and the mixture was vortexed and spun at 13,000 rpm for 5 minutes. The crude aqueous layer was used directly in assays.

5.8.3 Synthesis of reduced triketide

5.8.3.1 Synthesis of methyl 2-(2-methyl-1,3-dithian-2-yl)acetate (69)[129]

Methyl acetoacetate (0.43 ml, 4 mmol) was stirred in dry DCM (15 ml). 1,3-Propanedithiol (0.60 ml, 6 mmol) and boron trifluoride dietherate (1.01 ml, 8 mmol) were added and the mixture stirred at room temperature for 1 hour. The mixture was diluted with DCM (20 ml) and 1 M NaOH (30 ml). The organic layer was washed with brine and dried ($MgSO_4$). The solvent was removed under reduced pressure and the residue was freeze dried to give **69** as a white solid (1.29 g, 99%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.68 (s, 3H, -OCH_3), 3.03 (s, 2H, CH_2), 2.97-3.06 (m, 2H, CH_2), 2.69-2.76 (m, 2H, CH_2), 2.03-2.12 (m, 1H, CH), 1.80-1.91 (m, 1H, CH), 1.75 (s, 3H, CH_3)

5.8.3.2 Synthesis of 2-(2-methyl-1,3-dithian-2-yl)acetic acid (70)[87]

69 (0.5 g, 2.42 mmol) was dissolved in 1.0M aq. NaOH (10 ml) and stirred at 90 °C for 30 minutes. The mixture was allowed to cool, acidified to pH 1 using 1 M HCl and extracted with ethyl acetate. The solvent was removed under reduced pressure and the product lyophilised. **70** was obtained as a white solid (0.45 g, 97%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.09 (s, 2H, CH_2), 2.98-3.07 (m, 2H, CH_2), 2.71-2.79 (m, 2H, CH_2), 2.05-2.14 (m, 1H, CH), 1.80-1.91 (m, 1H, CH), 1.77 (s, 3H, CH_3).

5.8.3.3 Synthesis of 5-(1-hydroxy-2-(2-methyl-1,3-dithian-2-yl)ethylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione (72)[87]

70 (0.289 g, 1.50 mmol), DMAP (0.275 g, 2.25 mmol) and Meldrum's acid (0.237 g, 1.65 mmol) were dissolved in dry DCM (10 ml) at 0 °C. DCC (0.34 g, 1.65 mmol) in dry DCM (5 ml) was added dropwise over 1 hour. The mixture was allowed to warm to room temperature overnight with stirring. The mixture was filtered, and the filtrate washed with $NaHSO_4$ (10 ml), brine (10 ml) and water (10 ml). The organic layer was dried ($MgSO_4$), the solvent removed under reduced pressure and freeze dried to give **72** as a white powder (0.337 g, 71%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.99 (s, 2H, CH_2), 3.07-3.16 (m, 2H, CH_2), 2.64-2.71 (m, 2H, CH_2), 2.06-2.14 (m, 1H, CH), 1.78-1.86 (m, 1H, CH), 1.76 (s, 6H, 2xCH_3), 1.74 (s, 3H, CH_3).

5.8.3.4 Synthesis of methyl 4-(2-methyl-1,3-dithian-2-yl)-3-oxobutanoate (73)[87]

72 (0.337 g, 1.06 mmol) was dissolved in dry toluene (10 ml) with dry methanol (0.213 ml) and heated under reflux for 3 hours. The solvents were removed under reduced pressure to give **73** as a white solid (0.292 g, 99%). ^{1}H NMR (400 MHz, $CDCl_3$): [keto] δ = 3.74 (s, 3H, CH_3), 3.60 (s, 2H, CH_2), 3.22 (s, 2H, CH_2), 2.88-2.99 (m, 2H, CH_2), 2.78-2.85 (m, 2H, CH_2), 1.98-2.10 (m, 1H, CH), 1.85-1.96 (m, 1H, CH), 1.73 (s, 3H, CH_3). [enol] δ = 12.00 (s, 1H, OH), 5.08 (s, 1H, CH), 3.73 (s, 3H, CH_3), 3.60 (s, 2H, CH_2), 3.22 (s, 2H, CH_2), 2.88-2.99 (m, 2H, CH_2), 2.78-2.85 (m, 2H, CH_2), 1.98-2.10 (m, 1H, CH), 1.85-1.96 (m, 1H, CH), 1.73 (s, 3H, CH_3).

5.8.3.5 Synthesis of methyl 3-hydroxy-4-(2-methyl-1,3-dithian-2-yl)butanoate (74)[87]

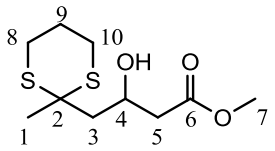


Figure 5.2: Numbering of **74** for NMR assignment.

73 (270 mg, 1.09 mmol) was dissolved in dry methanol (5 ml) and cooled to 0 °C. Sodium borohydride (123 mg, 3.26 mmol) was added portionwise over 15 minutes and the mixture stirred at room temperature for 1 hour. Saturated aq. citric acid (5 ml) was added and the mixture extracted twice with ether (2x 20 ml). The organic layer was washed with water, dried with magnesium sulphate and evaporated under reduced pressure to give crude **74** as a colourless oil (279 mg, 102%).^{1}H NMR (400 MHz, $CDCl_3$): ^{1}H NMR (400 MHz, $CDCl_3$): δ 4.38 (1H, m, *H-4*), 3.70 (3H, s, *H-7*), 3.64 (1H, d, J = 2 Hz, O*H*), 2.89-3.03 (2H, m, *H-8/H-10*), 2.72-2.82 (2H, m, *H-8/H-10*), 2.56 (1H, dd, J = 8, 16 Hz, *H-5*), 2.46 (1H, dd, J = 5, 16 Hz, *H-5*), 2.37 (1H, dd, J = 9, 15 Hz, *H-3*), 1.99-2.08 (1H, m, *H-9*), 1.95 (1H, dd, J = 2, 15 Hz, *H-3*), 1.92-1.81 (1H, m, *H-9*), 1.66 (3H, s, *H-1*). ^{13}C (100 MHz, $CDCl_3$): δ 172.2 (C-6), 65.8 (C-4), 51.8 (C-7), 47.5 (C-2), 46.7 (C-3), 42.0 (C-5), 28.4 (C-1), 26.8 (C-8/C-10), 26.6 (C-8/C-10), 24.6 (C-9)

5.8.3.6 Synthesis of (1R,4R)-4-methoxy-1-(2-methyl-1,3-dithian-2-yl)-4-oxobutan-2-yl 4,7,7-trimethyl-3-oxo-2-oxabicyclo[2.2.1]heptane-1-carboxylate (76 and 77)[87]

To a solution of **73** (253 mg, 0.34 mmol), triethylamine (0.14 ml, 1.01 mmol) and DMAP (41.5 mg, 0.34 mmol) in dichloroethane (10 ml) was added (S)-camphanoyl chloride (219 mg, 1.01 mmol) in dichloroethane (2 ml) and the mixture was heated under reflux for 24 hours. The mixture was cooled and the organic layer washed with water (20 ml), saturated aq. sodium bicarbonate (20 ml), saturated aq. citric acid (20 ml) and brine (20 ml). The organic layer was dried with magnesium sulphate and evaporated under reduced pressure. Crude **78** was obtained as a brown oil and purified by column chromatography on silica gel using a gradient from 50:1 (Petroleum ether 40-60: ethyl acetate) to 2:1.

The (R,S) diastereomer (**77**) was isolated as a white solid (15 mg , 10%). R_f = 0.33 (2:1 petroleum ether 40-60: ethyl acetate. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.48 (m, 1H), 3.63 (s, 3H), 2.86 (m, 2H), 2.68 (m, 4H), 2.33 (m, 3H), 1.95-1.80 (m, 4H), 1.61 (m, 1H), 1.57 (s, 3H), 1.03 (s, 3H), 1.01 (s, 3H), 0.90 (s, 3H).

The (S,S) diastereomer (**76**) was isolated as a white solid (6 mg, 4%). R_f = 0.41 (2:1 petroleum ether 40-60: ethyl acetate. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.49 (m, 1H), 3.61 (s, 3H), 2.91 (m, 2H), 2.67 (m, 4H), 2.52-2.23 (m, 3H), 1.92-1.76 (m, 4H), 1.61 (s, 1H), 1.56 (s, 3H), 1.04 (s, 3H), 1.00 (s, 3H), 0.88 (s, 3H).

5.8.3.7 Removal of esters to form 3-hydroxy-4-(2-methyl-1,3-dithian-2-yl)butanoic acid (57)

78 (20 mg, 0.05 mmol) was stirred in 1M NaOH (20 ml) overnight. The pH was adjusted to 1 with HCl and the solution extracted with ethyl acetate. The organic layer was dried with magnesium sulphate and evaporated under reduced pressure to give crude **57** as a yellow oil (14.2 mg). **57** was purified by preparative TLC run in ethyl acetate. A very small amount was recovered with R_f = 0.46. LC-MS (C18, MeCN/H_2O + 0.1% TFA) found 236.96 $[M+H]^+$ at 16.2 minutes.

5.8.3.8 Deprotection of 3-hydroxy-4-(2-methyl-1,3-dithian-2-yl)butanoic acid (79)[87]

57 (15 mg, 0.05 mmol), methyl iodide (0.1 ml) and calcium carbonate (10 mg) were stirred in acetonitrile/water (9:1, 1 ml) at 50 °C. The reaction was monitored by TLC every 30 minutes and additional portions of methyl iodide and calcium carbonate added until no further changes in number of spots or R_f were observed. The mixture was evaporated under reduced pressure, suspended in DCM and filtered. The filtrate was evaporated under reduced pressure. A brown oil was produced which showed the loss of the dithiane signals in NMR, but no structure could be assigned.

5.8.3.9 Synthesis of 3-hydroxy-4-(2-methyl-1,3-dithian-2-yl)butanoic acid (57)

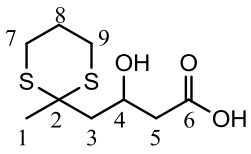


Figure 5.3: Numbering of **57** for NMR assignment.

74 (141 mg, 0.56 mmol) was stirred in 1 M NaOH (25 ml) overnight. The solution was acidified to pH1 using hydrochloric acid and extracted with diethyl ether. The organic layer was dried with magnesium sulphate and evaporated under reduced pressure. The residue was purified by column chromatography on silica gel using ethyl acetate/methanol (3:1) to give **57** as a colourless oil (130 mg, 0.51 mmol, 93 %).^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.40 (m, 1H, H-4), 2.98 (m, 2H, H-7/H-8), 2.78 (m, 2H, H-7/H-8), 2.57 (m, 2H, H-5), 2.45 (dd, 1H, J = 9, 15 Hz, H-3), 2.05 (m, 1H, H-9), 1.94 (dd, 1H, J = 1.4, 15 Hz, H-3), 1.88 (m, 1H, H-9), 1.67 (s, 3H, H-1). ^{13}C (100 MHz, $CDCl_3$): δ = 175.9 (C-6), 65.7 (C-4), 47.2 (C-2), 46.4 (C-3), 41.8 (C-5), 28.4 (C-1), 26.8 (C-8/C-10), 26.6 (C-8/C-10), 24.5 (C-9). R_f: 0.53 (3:1 ethyl acetate/methanol).

5.8.3.10 Synthesis of D-pantetheine (83)[130]

D-pantethine (226 mg, 0.41 mmol) was dissolved in dry isopropanol (5 ml) and degassed with argon for 40 minutes. Sodium borohydride (108 mg, 2.86 mmol) was added and the mixture heated under reflux for 18 hours. The mixture was allowed to cool then methanol (7.5 ml) and glacial acetic acid (2.5 ml) were added and the

solution stirred for 30 minutes. The solvents were removed under reduced pressure. Crude **83** was purified by column chromatography on silica gel using ethyl acetate/methanol (3:1). D-pantetheine (**83**) was obtained as a colourless oil (70 mg, 31%).^{1}H NMR (400 MHz, D_2O): δ = 3.92 (s, 1H), 3.47 (m, 2H), 3.44 (s, 1H), 3.35 (s, 1H), 3.31 (m, 2H), 2.60 (t, 2H, J = 7Hz), 2.46 (t, 2H, J= 7Hz), 0.86 (s, 3H), 0.83 (s, 3H). R_f: 0.57 (3:1 ethyl acetate/methanol).

5.8.3.11 Synthesis of S-(2-(3-((R)-2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl) 3-hydroxy-4-(2-methyl-1,3-dithian-2-yl)butanethioate (82)

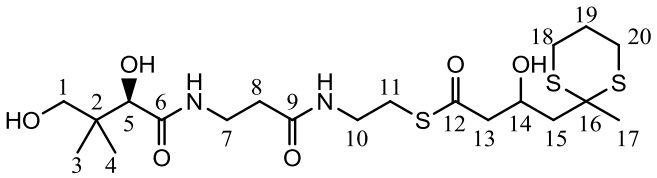


Figure 5.4: Numbering of 82 for NMR assignment.

57 (103 mg, 0.44 mmol) was placed in a flask which was flushed with argon and cooled to 0 °C. D-pantetheine (64 mg, 0.23 mmol) dissolved in dry THF (8 ml) was added and the flask flushed with argon again. DIPEA (0.043 ml, 0.25 mmol) and HATU (175 mg, 0.46 mmol) were added and the mixture stirred at 0 °C for three hours before being stirred at room temperature overnight. The mixture was evaporated under reduced pressure and the yellow oil was redissolved in ethyl acetate and water. The aqueous layer was extracted with ethyl acetate. The combined organic layers were dried with magnesium sulphate and evaporated under reduced pressure. **82** was obtained as a yellow oil (58.4 mg). Purification of **82** was carried out by preparative HPLC with a C18 column using an acetonitrile/water gradient (0% acetonitrile to 100% acetonitrile over 40 minutes) to give a colourless oil (7.5 mg, 0.015 mmol, 7 %). R_t = 14.5 mins. HRMS: 497.1810 found ($[M+H]^+$, calcd 497.1808, Δ = 0.42 ppm).^{1}H NMR (400 MHz, D_2O): δ 4.29 (m, 1H, *H-14*), 3.89 (s, 1H, *H-5*), 3.37-3.47 (m, 1H + 2H, *H-1, H-7*), 3.24-3.34 (m, 1H + 2H, *H-1, H10*), 2.97 (td, 2H, J = 6.6, 1.9 Hz, *H-11*), 2.92-2.74 (m, 2H + 2H + 2H, *H-13, H-18, H-20*), 2.38 (t, 2H, J = 6.6 Hz, *H-8*), 2.17 (dd, 1H, J = 15.2, 2.2 Hz, *H-15*), 1.99 (dd, 1H, J = 15.2, 8.2 Hz, *H-15*),

1.86-1.95 (m, 1H, *H-19*), 1.79-1.81 (m, 1H, *H-19*), 1.56 (s, 3H, *H-17*), 0.82 (s, 3H, *H-3/H-4*), 0.79 (s, 3H, *H-3/H-4*). ^{13}C NMR (100 MHz, D_2O): δ 201.0 (*C-12*), 175.1 (*C-9*), 174.0 (*C-6*), 75.8 (*C-5*), 68.4 (*C-1*), 66.2 (*C-14*), 51.7 (*C-13*), 47.4 (*C-16*), 46.4 (*C-15*), 38.6 (*C-10*), 38.6 (*C-2*), 35.5 (*C-7*), 35.3 (*C-8*), 28.2 (*C-11*), 27.7 (*C-17*), 26.1 (*C-18/C-20*), 25.9 (*C-18/C-20*), 24.3 (*C-19*), 20.5 (*C-3/C-4*), 19.1 (*C-3/C-4*).

5.8.3.12 Synthesis of S-(2-(3-((R)-2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl) 3-hydroxy-5-oxohexanethioate (81)

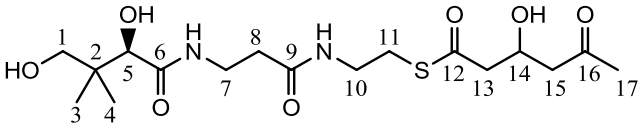


Figure 5.5: Numbering of **81** for NMR assignment.

82 (3 mg, 0.006 mmol), methyl iodide (150 μl) and calcium carbonate (15 mg) were stirred in MeCN/H2O (9:1, 15 ml) at 50 °C for 24 hours. The solvent was removed under reduced pressure and the solid (23.4 mg) redissolved in methanol for purification by semi-preparative HPLC using a MeCN/H_2O + 0.1% formic acid gradient (0-100% MeCN over 40 minutes). **81** was obtained as a colourless oil (0.9 mg, 0.002 mmol, 33%). R_t = 13.5 mins. HRMS: 407.1837 found ($[M+H]^+$, calcd 407.1846, Δ = -2.2 ppm). ^{1}H NMR (500 MHz, D_2O): δ 4.47 (q, 1H, J = 6 Hz, *H-14*), 3.91 (s, 1H, *H-5*), 3.37-3.47 (m, 2H+1H, *H-1 + H-7*), 3.28-3.34 (m, 2H+1H, *H-1 + H-10*), 2.99 (dt, 2H, J = 2, 6 Hz, *H-11*), 2.78 (dd, 2H, J = 6, 2 Hz, *H-13*), 2.72 (d, 2H, J = 7 Hz, *H-15*), 2.40 (t, 2H, J = 7 Hz, *H-8*), 2.16 (s, 3H, *H-17*), 0.84 (s, 3H, *H-3/H-4*), 0.81 (s, 3H, *H-3/H-4*). ^{13}C NMR (125 MHz, D_2O): δ 212.9(*C-16*), 200.7 (*C-12*), 175.0 (*C-6*), 174.0 (*C-9*), 75.7 (*C-5*), 68.3 (*C-1*), 64.4 (*C-14*), 49.9 (*C-13*), 49.2 (*C-15*), 38.5 (*C-2*), 38.4 (*C-10*), 35.3 (*C-8*), 35.2(*C-7*), 29.9 (*C-17*), 28.1 (*C-11*), 20.4 (*C-3/C-4*), 19.0 (*C-3/C-4*).

5.8.3.13 Formation of 3-hydroxy-5-oxohexanoyl CoA (80)[81]

81 (1 mg, 0.002 mmol) was dissolved in buffer (Tris-HCl (50 mM), KCl (20 mM), $MgCl_2$ (10 mM), pH 7.5) to a final volume of 1 ml. ATP (5 mg, 0.01 mmol) was added and the pH was adjusted to 7 with 4 M NaOH. Pantothenate kinase (800 μg),

phosphopantetheine adenyltransferase (900 μg) and dephosphocoenzyme A kinase (948 μg) were added and the mixture was incubated at room temperature overnight. Chloroform (1 ml) was added and the mixture was centrifuged at 13,000 rpm for 5 minutes. The crude aqueous layer was lyophilised and redissolved in water for analysis by LC-MS, but no peak for 3-hydroxy-5-oxohexanoyl CoA could be found.

5.9 Molecular biology methods

5.9.1 Polymerase Chain Reaction (PCR)

5.9.1.1 Primers

Name	Sequence
pMSAS_ACP-for*	5'-CAGCGACGCG GTGCATATGT CTGGACCGGA GCTTAAGGCT-3'
pMSAS_ACP-rev*	5'-GCCGTTCTTA TGAATTCAAT CATTTGGCAA-3'
MSAS_ACP-rev-CHis	5' –GCCGTTCTTA TGAATTCAAT CCTTTGGCAA-3'
pDH_H34A_1A*	5'-GGCATCTCCCCTAAGGAGGCCGAGCAGATG-3'
pDH_H34A_1B*	5'-CTCGGTACCATGGAGAGGAGCGCTGCCCGG-3'
pDH_H34A_2A*	5'-CAGCCACCCTCTCCATGGTACCGAG ATCGT-3'
pDH_H34A_2B*	5'-GATCTTGGGC GGATCCTGCG AGGTGAAGAC-3'

***Designed by Dr Fanglu Huang**

5.8.1.2 Plasmid PCR

Plasmid DNA (1 μl), 10x Pfu buffer (5 μl), dNTP mix (2 mM each NTP, 5 μl), forward and reverse primers (1 μl each), Pfu DNA polymerase (1 μl) and DMSO (1 μl) were made up to 50 μl with water.

The mixture was heated to 95 °C for 3 minutes initially. Then 25-30 cycles of the following conditions were carried out: 95 °C for 30 seconds, 58 °C for 30 seconds, 72 °C for 2 minutes. The mixture was then held at 72 °C for 10 minutes before being cooled to 4 °C.

5.9.1.3 Colony PCR

A single colony was picked from an LB-agar plate and resuspended in water (30 μl) using a sterilised toothpick. 10x Dream Taq buffer (5 μl), dNTP mix (2 mM each NTP, 5 μl), forward and reverse primers (1 μl each), Dream Taq polymerase (1 μl) and DMSO (1 μl) were added and the total volume made up to 50 μl with water.
The PCR reaction was carried out as for plasmid PCR.

5.9.2 Digestion of DNA with restriction endonucleases

Plasmid DNA (30 μl), appropriate 10x buffer (5 μl), BSA (2 μl) and restriction endonucleases (1 μl each) were made up to 50 μl with water and incubated at 37 °C overnight. Digested DNA was purified by preparative agarose gel electrophoresis.

5.9.3 Ligation

5.9.3.1 Two piece ligation

Digested vector (*e.g.* x μl) and insert (*e.g.* y μl) were incubated with 10x reaction buffer (1.5 μl) and T4 DNA ligase (1.5 μl) in a total reaction volume of 15 μl at 16°C overnight. The relative concentrations of vector and insert were established either by analysing the fluorescence of each on an agarose gel or by analysis using a nanodrop. The amounts of vector and insert (x and y) were adjusted to give a final concentration ratio of 2:1 (insert:vector). 7.5 μl of the ligation mixture was transformed into *E. coli* NovaBlue cells as described in section 5.9.7. Colonies were screened for the presence of insert using colony PCR as previously described. Colonies containing the correct insert then had the plasmid extracted and sent for sequencing. All constructs in this study were sequenced by the staff of the

sequencing facility of the Department of Biochemistry, University of Cambridge. pET28a(+) and pET20b(+) constructs were sequenced using T7 and T7T primers.

5.9.3.2 Three piece ligation

pKOS007-109 that had been digested with Bsu36I and BamHI (*e.g.* x μl), the left insert digested with Bsu36I and KpnI (*e.g.* y μl) and the right insert digested with KpnI and BamHI (*e.g.* z μl) were added to 10x reaction buffer (1.5 μl) containing T4 ligase (1.5 μl) in a total volume of 15 μl. The amounts x, y and z were calculated to give a final concentration ratio of (1:5:5 vector:left:right). The mixture was incubated overnight at 16 °C and 7.5 ml of ligation mixture was transformed into NovaBlue and plated onto LB/Cb plates. Colonies were screened by plasmid preparation and comparison of plasmid size to the original pKOS007-109 plasmid. Sequencing of plasmids was performed using the DHm-1A and DHm-2B primers.

5.9.4 Plasmid preparation

NovaBlue cells containing the required plasmid were grown in LB medium (10 ml) overnight at 37 °C, 220 rpm. The culture was centrifuged at 5000 rpm for 5 minutes. Plasmid was extracted from the cell pellet using the GeneJET Plasmid Purification kit (Fermentas) according to the manufacturer's protocol.

5.9.5 Agarose gel electrophoresis

Agarose (electrophoresis-grade) was dissolved in TAE buffer to final concentration of 0.7-1.5% as required with microwave heating. Ethidium bromide solution (2 μl) was added and the agarose poured into a horizontal caster. DNA samples were mixed with 5:1 DNA loading dye before being loaded onto the gel. Gels were run at room temperature in TAE buffer at 100V for 25 minutes and then visualised on a UVI-Doc ultraviolet transilluminator.

TAE buffer: 40 mM Tris-acetate, 1 mM EDTA.

5.9.6 DNA purification

PCR products and digested DNA were purified on a preparative 1% agarose gel. The appropriate band was excised and the DNA extracted using a GeneJET Gel extraction kit (Fermentas) according to the manufacturer's protocol.

5.9.7 Transformation of plasmid

E. coli NovaBlue or BL21(DE3) competent cells were thawed on ice, then recombinant plasmid was added (1 μl). The mixture was incubated on ice for ten minutes, then heated to 42 °C for 35 seconds before being incubated on ice for 5 minutes. SOC medium (80 μl) was added and the culture was incubated at 37 °C, 220 rpm for 1 hour. The culture was then spread onto an LB-agar plate containing an appropriate antibiotic (50 μg/ml) and incubated overnight at 37 °C.

SOC medium: Bacto-tryptone, 20 g; bacto-yeast extract, 5 g; NaCl, 0.5 g; KCl, 0.186 g; $MgCl_2$, 0.952 g; glucose, 3.603 g; ddH_2O to 1 litre, pH 7.0.
LB-agar: LB granules 2.5 g/100 ml, agar 2 g/100 ml, ddH_2O.

5.9.8 Gene expression

5.9.8.1 6-MSAS[59]

A single colony from an LB-agar plate was inoculated into ATCC 765 media (10 ml) containing 100 μg/ml carbenicillin (*apo*) or 100 μg/ml carbenicillin and 50 μg/ml kanamycin (*holo*). The culture was grown for 8 hours at 37 °C with shaking at 220 rpm and then added to 1L of ATCC 765 medium containing 10% glycerol (v/v), 100 μg/ml carbenicillin (*apo*) or 100 μg/ml carbenicillin and 50 μg/ml kanamycin (*holo*) and grown at 37 °C, 220 rpm until an OD_{600} of 1 was reached. Protein expression was induced with IPTG (0.5 mM). *Holo* 6-MSAS was then incubated at 30 °C for 24 hours, while *apo* 6-MSAS cultures were incubated at 16 °C for 24 hours.

Cells were harvested by centrifugation at 7000 rpm for 10 minutes at 4 °C, resuspended in resuspension buffer with the addition of benzamidine (20 mg), PMSF

(20 mg) and lysozyme (20 mg), and sonicated on ice for 24 minutes (2 seconds on, 10 seconds off, 4 minutes total on, 80% amplitude). The cell lysate was cleared by centrifugation at 17,000 rpm and 4 °C for 30 minutes.

ATCC 765 medium: potassium dihydrogen phosphate, 4.49 g; dipotassium hydrogen phosphate, 11.67 g; ammonium sulphate, 1.98 g; iron (II) sulphate heptahydrate, 0.5 mg; magnesium sulphate, 0.25 g; glycerol, 20.25 ml; casamino acids, 11.0 g; dd H_2O, 1 litre.

5.9.8.2 Other proteins

A single colony from an LB-agar plate was inoculated into LB medium (10 ml) containing an appropriate antibiotic (50-100 μg/ml). The culture was grown overnight at 37 °C, 220 rpm and then added to 1L of LB medium containing the appropriate antibiotic (50-100 μg/ml). The culture was grown at 37 °C, 220 rpm until an OD_{600} of 0.5-0.9 was reached and protein expression was induced using IPTG (0.2 mM-0.5 mM). The culture was then incubated overnight at 16°C with shaking at 220 rpm. Cells were harvested by centrifugation at 7000 rpm for 10 minutes at 4 °C and resuspended in binding buffer, before being sonicated on ice (2 seconds on, 10 seconds off, 4 minutes total on, 80% amplitude). The cell lysate was cleared by centrifugation at 17,000 rpm and 4 °C for 30 minutes.

Luria-Bertani medium: LB granules, 25 g; distilled water, 1 litre.

5.9.9 Protein purification

5.9.9.1 Ammonium sulphate precipitation of 6-MSAS

Solid ammonium sulphate was added to cleared cell lysate at 4 °C to give a final concentration of 22%. The mixture was left to equilibrate at 4 °C for 30 minutes before centrifugation at 11000*g* and 4 °C for 15 minutes. The supernatant was decanted and solid ammonium sulphate added at 4 °C to give a final concentration

of 40%. The mixture was left to equilibrate at 4 °C for 1 hour and centrifuged at 5000*g* and 4 °C for 30 minutes. The pellet was stored at -80 °C.

5.9.9.2 FPLC of 6-MSAS

Gel filtration was carried out on an ÄKTA Explorer FPLC system with a HiLoad 16/60 Superdex 200 Prep Grade column. The mobile phase was that used by Richardson *et al.*[85] 6-MSAS ammonium sulphate pellet (100 mg/ml in assay buffer, 1 ml) was injected on to the column. The fractions containing the desired peak were collected and concentrated in a PVP-treated spin concentrator with a 100 kDa cut-off (Amicon). The glycerol content of concentrated fractions was adjusted to 50% and the protein was stored at -20 °C.

FPLC buffer: NaH_2PO_4 100 mM, EDTA 1 mM, DTT 2 mM, glycerol 20%, pH 7.1.

5.9.9.3 Purification of His_6-tagged proteins

His-bind resin (2 ml) was washed with water (12 ml) and charged with 50 mM $NiCl_2$ (10 ml), before being washed with binding buffer (6 ml). Cleared cell lysate was passed through the column and the column washed with binding buffer (20 ml). Fractions were eluted with wash buffer (6 ml), elution buffer (6 ml) and 400 mM elution buffer (6 ml). The elution buffer fraction was concentrated in an Amicon spin concentrator of appropriate size and the buffer exchanged for storage buffer. Protein samples were stored at -20 °C.

Binding buffer: 20 mM Tris.HCl, 0.5 M NaCl, 5 mM imidazole, 10% glycerol, pH 7.9.
Wash buffer: 20 mM Tris.HCl, 0.5 M NaCl, 20 mM imidazole, 10% glycerol, pH 7.9.
Elution buffer: 20 mM Tris.HCl, 0.5 M NaCl, 200 mM imidazole, 10% glycerol, pH 7.9.
400 mM elution buffer: 20 mM Tris.HCl, 0.5 M NaCl, 400 mM imidazole, 10% glycerol, pH 7.9.
Storage buffer: 50 mM HEPES, 200 mM NaCl, 50% glycerol, pH 7.6.

5.9.10 Sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Premade 10% gels and 8-16% gradient Precise Protein gels were purchased from Thermo Scientific. 18% gels were prepared according to standard protocols with a 5 % acrylamide stacking gel. Protein samples were mixed 1:1 with SDS loading buffer before loading on to the gel. Premade gels were run at 150 V, 400 mA for 45 minutes in Tris-HEPES buffer. 18% gels were run in Tris-glycine buffer, 250 V, 20 mA for 2 hours. PageRulerPlus protein marker (Fermentas) was used for reference. After running, gels were stained with Coomassie Blue staining solution for 10 minutes and then in destaining solution overnight. Gels were visualised on a UVI-Doc ultraviolet transilluminator.

Tris-HEPES buffer: Tris base, 12.1 g; HEPES, 23.8 g; SDS, 1.0 g; ddH_2O 1 litre.
Tris-glycine buffer: Tris 25 mM, glycine 250 mM, 0.1% SDS, pH 8.3.
SDS loading buffer: Tris-HCl 50 mM, DTT 100 mM, 2 % SDS, 0.1% bromophenol blue, 10% glycerol, pH 6.8.
Staining solution: 2.5 g/L Coomassie Brilliant Blue in MeOH:H_2O:acetic acid (4.5:4.5:1)

Destaining solution: MeOH:H_2O:acetic acid (4.5:4.5:1)

5.9.11 Protein concentration

Protein concentration was determined by the Bradford method.[131] 1-5 μl of protein stock was mixed with Bradford reagent (900 μl) in a final volume of 1 ml at room temperature. The mixture was allowed to stand for 5 minutes before the absorbance at 595 nm was determined. This was compared with a standard curve measured using known concentrations of BSA to calculate protein concentration.

5.9.12 Limited proteolysis

5.9.12.1 Digestion of 6-MSAS with hydroxylamine

FPLC-purified 6-MSAS (600 μg) and hydroxylamine (34 μl, 1.8 M final concentration) in a final volume of 300 μl were incubated at 37 °C. A sample of 50 μg was taken every 30 minutes and quenched with 1 drop of 10% HCl before freezing. Samples were analysed using SDS-PAGE.

5.9.12.2 Digestion of 6-MSAS with trypsin

FPLC-purified 6-MSAS (100 μg) and trypsin (x μg, where desired ratio is 100:x wt/wt 6-MSAS:trypsin) were incubated at room temperature. Samples were taken at timed intervals and the reaction was stopped by addition of TLCK (1 mM final concentration). Samples were analysed using SDS-PAGE.

5.9.12.3 N-terminal sequencing

SDS-PAGE gels containing the bands of interest were blotted onto PVDF membrane using the iBlot dry blotting system. The PVDF membrane was placed on top of the nitrocellulose membrane in the supplied cartridges and the rest of the blotting apparatus was assembled according to the manufacturer's instructions. The blot was stained using Coomassie Blue staining solution and destained with destaining solution until bands were visible, before being allowed to air dry. N-terminal sequencing was carried out by the Protein and Nucleic Acid Chemistry Facility (Department of Biochemistry).

5.10 Enzyme assays

5.10.1 Production of 6-MSA in culture

Culture supernatant (50 ml) was acidified to pH 1 with 1M HCl and extracted three times with ethyl acetate. The combined organic extracts were dried with magnesium sulphate and evaporated under reduced pressure. The residue was dissolved in methanol (1 ml) and 15 μl of this was injected onto the LC-MS. The 6-MSA peak was

identified by comparison with authentic 6-MSA (10 μl of 1 mg/ml) run in the same session. The amount of 6-MSA was calculated based on a comparison of the peak area between the extract and the standard in the UV traces at 254 nm and 280 nm.

5.10.2 Activity of 6-MSAS *in vitro* [56]

6-MSAS, either as the ammonium sulphate pellet (40 mg/ml, 50 μl), or FPLC-purified (3 mg/ml, 25 μl) was incubated with DTT (10 mM final concentration) at room temperature for 15 minutes. 6-MSAS was then added to a mixture of acetyl-CoA (0.1 μmol), NADPH (0.1 μmol) and bovine serum albumin (0.5 mg) in assay buffer with a final volume of 0.5 ml. Malonyl-CoA (0.1 μmol) was added to start the reaction. Production of 6-MSA was monitored using a Cary fluorimeter (Varian) with λ_{ex} = 310 nm and λ_{em} = 390 nm at 25 °C for 30 minutes. Activity of 6-MSAS was calculated using a standard curve of 6-MSA concentrations.

6-MSAS assay buffer: 100 mM Tris sulphate, 1 mM EDTA, 15% glycerol, pH 7.6.

5.10.3 Phosphopantetheinylation of 6-MSAS

Apo 6-MSAS (50 μl, 40 mg/ml pellet), PPTase (1 μl, 6 mg/ml), CoASH (1 μl, 30 mM stock) and DTT (10 mM final concentration) were incubated at room temperature or 37 °C for 1 hour in Sfp, CoA or phosphate buffer with addition of $MgCl_2$ (5 mM final concentration) with a final volume of 55 μl.

Sfp buffer: 50 mM HEPES, 1 mM MgCl2, pH 6.5.
CoA buffer:[118] 75 mM Tris.HCl, 10 mM $MgCl_2$, 25 mM DTT, pH 8.8
Phosphate buffer: 100 mM sodium phosphate, pH 7.5.

5.10.4 Phosphopantetheinylation of Con2* and ChlD2

Apo ACP (1.25 μl, 10 mg/ml stock), PPTase (0.7 μl, 3 mg/ml stock) and CoASH (1 μl, 30 mM stock) were made up to a final volume of 25 μl with either CoA buffer or phosphate buffer with addition of $MgCl_2$ (10 mM final concentration). The mixture

was incubated for 2 hours at room temperature before injection of 15 μl for analysis by LC-MS.

5.10.5 Off-loading of intermediates from 6-MSAS

5.10.5.1 Starter unit

A typical assay contained FPLC-purified 6-MSAS (100 μg, treated with DTT), **30** (15 mM), acetyl-CoA or acetoacetyl-CoA (3 mM), with NADPH (10 mM) if required, made up to a final volume of 100 μl with 6-MSAS assay buffer. Assays were incubated at room temperature overnight. The aqueous layer was extracted twice with ethyl acetate (150 μl) and the organic extracts were dried under N_2. The residue was dissolved in methanol (20 μl) and 10 μl was analysed by HRMS.

5.10.5.2 Competitive assays

A typical assay contained FPLC-purified 6-MSAS (100 μg, treated with DTT), **30** (15 mM), acetyl-CoA or acetoacetyl-CoA (3 mM), NADPH (10 mM) and malonyl-CoA (1-13 mM) made up to a final volume of 100 μl with 6-MSAS assay buffer. Assays were incubated at room temperature overnight. The aqueous layer was extracted twice with ethyl acetate (150 μl) and the organic extracts were dried under N_2. The residue was dissolved in methanol (20 μl) and 10 μl was analysed using the LTQ Orbitrap.

5.10.5.3 Delayed off-loading from 6-MSAS-DHm

The assay mixture containing FPLC-purified 6-MSAS-DHm (100 μg, treated with DTT), acetyl-CoA or acetoacetyl-CoA (3 mM), malonyl-CoA (5 mM) and NADPH (10 mM) made up to a final volume of 100 μl with 6-MSAS assay buffer was incubated at room temperature for 1 hour. Non-hydrolysable analogue **30** was added to a final concentration of 15 mM and the reaction mixture was incubated overnight. The aqueous layer was extracted twice with ethyl acetate (150 μl) and organic extracts were dried under N_2. The residue was dissolved in methanol (20 μl) and analysed directly using the LTQ Orbitrap.

ADDENDUM

Additional experiments in 6-methylsalicylic acid biosynthesis and synthetic approaches towards 5-amidinyl teicoplanin

A1 Alternative syntheses of reduced triketide intermediates

A1.1 Introduction

The biosynthesis of 6-MSA is proposed to proceed via an enzyme-bound reduced triketide.[52] The hydroxyl group at C-4 is produced by reduction of a ketone by the ketoreductase domain found in 6-MSAS. The hydride group is provided by NADPH.

To probe the stereochemistry of this reduction, it was proposed to synthesise the enantiomerically resolved reduced triketides **65** and **66** , after Cheung *et al*. [87] These could be provided to the 6-MSAS enzyme as CoA thioesters, which could potentially be loaded on to the *apo* enzyme by a phosphopantetheinyltransferase enzyme, such as Sfp. The reduced triketide that had the correct hydroxyl configuration could be elaborated to 6-MSA by dehydration, condensation with a unit of malonyl CoA, cyclisation and aromatisation, followed by release of the product. The reduced triketide that has the incorrect stereochemistry should not be a substrate for the dehydratase domain and so no production of 6-MSA should be observed from incorporation of the reduced triketide (Figure A1).

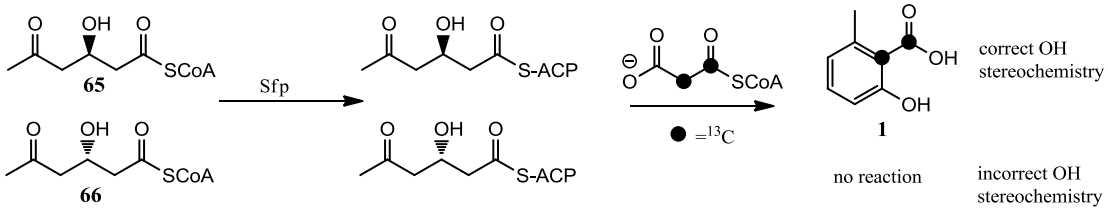


Figure A1: Resolved reduced triketides **65** and **66** could be loaded on to 6-MSAS using Sfp. The triketide with the correct hydroxyl stereochemistry should be extended to form 6-MSA (**1**).

As malonyl-CoA can decarboxylate to form acetyl-CoA, which can prime 6-MSAS, it would be expected that there would be a significant background rate of formation of 6-MSA. In order to differentiate 6-MSA formed through the background reaction from that formed through reaction with the enantiomerically resolved reduced triketides, it was necessary to include an isotopic label in the reduced triketide. This would allow 6-MSA formed from the triketides to be easily distinguished in a mass spectrum.

A1.2 Synthesis of singly-D labelled reduced triketide

The simplest method of incorporating a label into the reduced triketide CoA that would be retained in 6-MSA is to replace the sodium borohydride used in the reduction of the central ketone in **73** with sodium borodeuteride. This would add a single deuterium label at C-4 and cause the mass of the 6-MSA produced to increase by 1 unit.

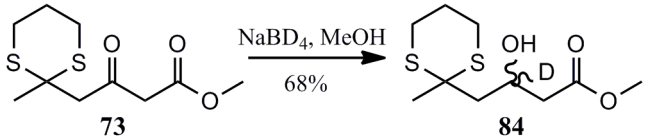


Figure A2: Synthesis of protected singly-deuterated reduced triketide **84** from protected triketide **73**.

Protected triketide **73** was dissolved in methanol and cooled to 0 °C. Sodium borodeuteride was added portionwise over 15 minutes and the solution was stirred

at room temperature for 1 hour. Citric acid was added, the mixture was extracted with diethyl ether and the organic layer washed with water. The organic layer was dried with magnesium sulphate and removed under reduced pressure to give **84** as a yellow oil in 68% yield (Figure A2).

84 was easily prepared and could potentially be deprotected and coupled with CoASH to form the desired reduced triketide-CoA (section 4.2.2). However the +1 mass shift in the 6-MSA produced would coincide with the +1 peak observed for the incorporation of ^{13}C into 6-MSA, which should account for a peak height of 1.1% of that for unlabelled 6-MSA. As the loading of the reduced triketide on to 6-MSAS was a single turnover event, while the background reaction forming 6-MSA from acetyl- and malonyl-CoA would turn over many times, the unlabelled 6-MSA peak would be much larger than the labelled 6-MSA peak. This meant that the ^{13}C [M+1] peak would be a significant contributor to any peak at seen at m/z = 154 and so the incorporation of the labelled reduced triketide would be very difficult to detect unambiguously.

A1.3 Synthesis of multiply D-labelled reduced triketide

To avoid the problem of overlapping peaks, it was decided that it would be best to label the reduced triketide with deuterium at multiple sites. The position with the best potential retention of deuterium label throughout the synthesis is the methyl group at C-1. This is the least acidic site throughout the synthesis, as it is adjacent to a dithiane, and allows the label to be installed in the very first step.

Methyl acetoacetate, the starting point for the synthesis of the reduced triketide-CoA (see section 4.2.2), can be deuterated using freshly prepared deuterated sodium methoxide in deuterated methanol (Figure A3).[132] The first protons to be replaced would be the two between the ketone and ester groups as this position is the most acidic. The next most acidic position is the methyl group adjacent to the ketone and these protons would then be replaced to produce the fully labelled acetoacetate

moiety. It is also possible that the methyl ester would be replaced by a deuterated methyl ester.

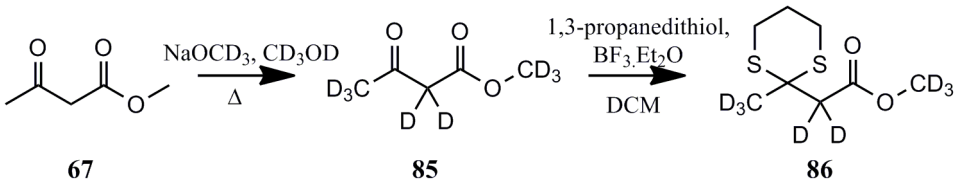


Figure A3: Deuterium labelling of methyl acetoacetate at multiple sites, followed by dithiane protection of the ketone to produce **86**.

Attempts to purify the deuterated product by Kugelrohr distillation produced very little labelled methyl acetoacetate. The distillation product obtained was protected with 1,3-propanedithiol to install the dithiane at the C-2.[129] This should prevent the exchange of the deuterium atoms attached to C-1 and also provide an internal standard for comparison of the extent of labelling at positions C-1 and C-3 by NMR.

It was found that a shorter reaction time produced a slightly higher yield of methyl acetoacetate, but that the labelling extent was reduced (Table A1). An overnight reaction had little advantage over a 5 hour reaction time.

Table A1: The extent of deuterium labelling at different positions in **86** for different deuteration reaction times, as determined by ^{1}H NMR.

		Extent of deuteration at position		
Time	Yield (%)	-CH_3	-CH_2-	-OCH_3
3 hours	20.5	67	50	83
5 hours	17.7	83	50	100
Overnight	17	83	50	100

The low yields of the labelled methyl acetoacetate meant that it would be challenging to carry a useful amount of labelled material through the rest of the synthesis to form labelled reduced triketide-CoA. Instead, it was decided to use unlabelled reduced triketide-CoA and labelled malonyl-CoA for this experiment (see section 4.2.2). This would produce a peak for 6-MSA from the reduced triketide at m/z = 155 $[M+H]^+$ while the background reaction would produce fully deuterated 6-MSA at m/z = 160 $[M+H]^+$.

A1.4 Alternative synthetic route using chymotrypsin

As the cleavage of the camphanic ester in the original synthesis was proving difficult (see section 4.2.2), an alternative synthetic route was sought.

Diethyl 3-hydroxyglutarate that has been protected at the central hydroxyl group with an acetate group (**87**) is a substrate for the enzyme chymotrypsin.[133] This cleaves one of the ethyl esters to preferentially form the (R)-hydroxy chiral compound **88**. The presence of the acetate group increases the enantioselectivity of the reaction.

The carboxylic acid formed could be reacted with Meldrum's acid to form the cyclic adduct **89**. Heating **89** under reflux in water should result in the opening of the cyclic adduct and the decarboxylation of the β-ketoacid formed to give the protected reduced triketide **90** (Figure A4). This could be reacted with sodium hydroxide to remove the ester protecting groups before coupling with CoASH to form the (R)-reduced triketide CoA.

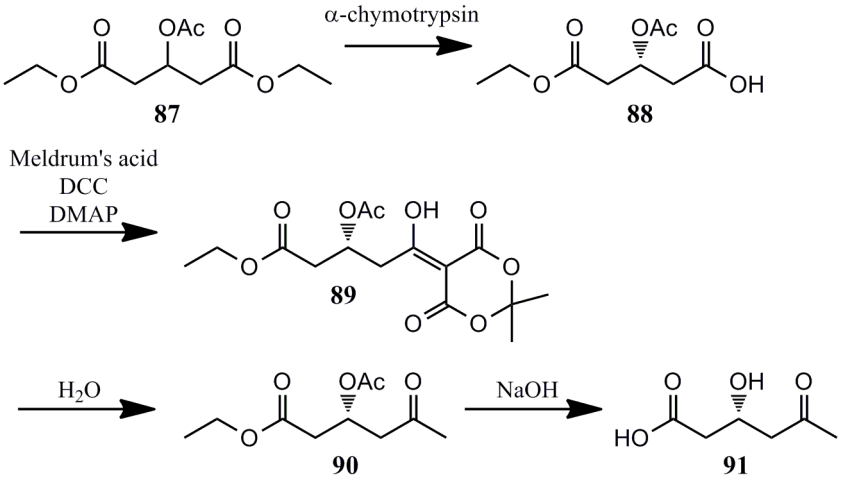


Figure A4: Proposed synthesis of the (R)-reduced triketide **91** from the product of the a-chymotrypsin reaction via coupling to Meldrum's acid and ring opening/decarboxylation.

The (S)-hydroxyl could be formed by the action of pig liver esterase (PLE) on diethyl 3-hydroxyglutarate in moderate ee.[134] However, the addition of the acetyl protecting group does not increase the enantioselectivity of the reaction. Instead, the specificity of the PLE is switched from the (S)-hydroxyl to the (R)-hydroxyl.[135] The formation of the (S)-reduced triketide CoA presents a more significant synthetic challenge, as the separation of the (S)- and (R)-enantiomers of the ester cleavage reaction would be difficult. Alternatively, the (S)-reduced triketide could be formed from the (R)-product. The acid moiety would have to be protected, for example with a 2-alkyl-1,3-oxazoline while the esters were removed and a methyl group installed on the newly revealed acid. This could be done using a Weinreb amide and methyl magnesium iodide. The original carboxylic acid moiety could then be deprotected and coupled to CoASH (Figure A5).

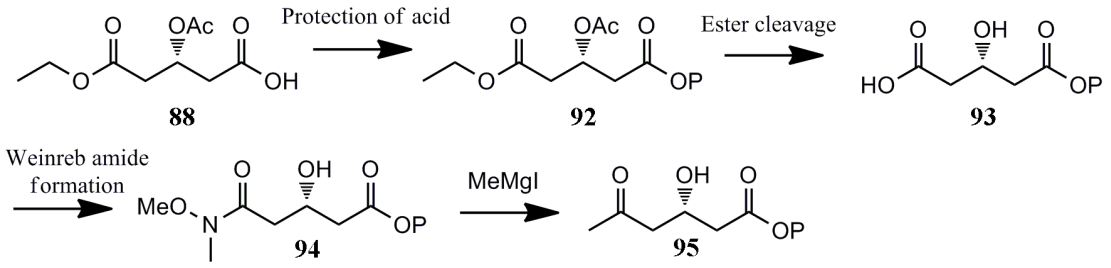


Figure A5: Proposed synthesis of the (S)-reduced triketide from the (R)-ester cleavage product.

Diethyl 3-hydroxymethylglutarate (**96**) was protected with acetic anhydride in pyridine to give the acetylated species **87** in 92% yield. The acetylated species was dissolved in 1 M KH_2PO_4 and the pH was adjusted to 7.8 with 1 M NaOH. Chymotrypsin was suspended in distilled water and added to the buffered solution. The pH was adjusted back to 7.8 using 1 M NaOH and kept constant throughout the reaction by further addition of NaOH. After 2 hours, the reaction was acidified with 1 M HCl and extracted with ether. The organic layer was treated with ammonia water and the mixture was separated. The aqueous layer was acidified with 1 M HCl and extracted with ether. The organic layer was dried with magnesium sulphate and the solvent was removed under reduced pressure to give the acid **88** (Figure A6).

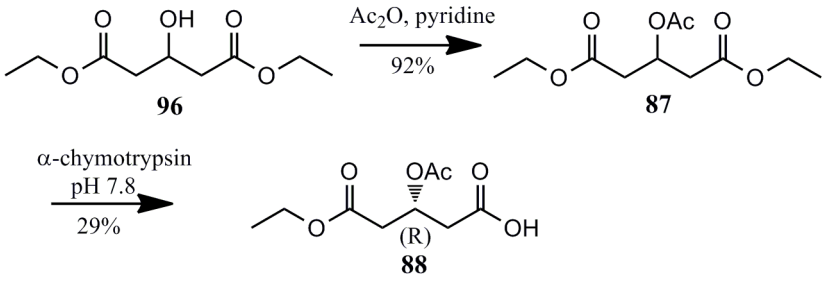


Figure A6: Protection of diethyl 3-hydroxyglutarate with acetic anhydride in pyridine, followed by ester cleavage using chymotrypsin leads to the chiral molecule **88**.

To determine the enantiomeric excess (ee) of **88**, the acid was derivatised by coupling with (R)-(+)-2-phenylethylamine.[136] This enabled the stereochemistry of the protected hydroxyl group to be measured by the difference in chemical shift of the

2-phenylethylamine methyl group and comparing them to the literature values for these compounds.

Acid **97** was coupled with (R)-(+)-2-phenylethylamine using oxalyl chloride to give **98** as a crude white solid (Figure A7). Comparison of the signals in the methyl region of the NMR spectrum of **98** with the literature values gave the (R)-(R) isomer at d = 1.97 and (S)-(R) isomer at δ = 1.95.[135] The ratio of the peak integrals gives an ee of 77% for the (R)-(R) isomer. The yield was not determined, but the crude product contained mainly starting material.

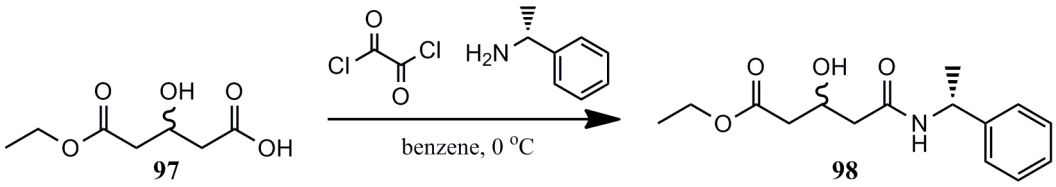


Figure A7: Coupling of the chymotrypsin reaction product **97** with (R)-phenylethylamine for the determination of the ee by NMR.

The chymotrypsin reaction was challenging since the pH had to be carefully controlled throughout the reaction, as a wider pH range would result in a reduced ee.[136] The yield was also rather low at 29%, compared with 95% reported in the literature.[135]

A1.5 Coupling to Meldrum's acid

The crude product from the chymotrypsin reaction was coupled with Meldrum's acid using DCC and DMAP in DCM, conditions that were used for a similar reaction in section 4.2.2. Under these conditions, no product was seen. Using THF as a solvent or EDC as a coupling reagent also did not result in any product, although in all cases the reaction turned blue, then brown.

It was decided that the protecting group might be causing problems with the coupling reaction, and so a different protecting group for the hydroxyl was investigated.

A1.6 Change of protecting group

It has been shown that a variety of protecting groups are tolerated by chymotrypsin in the cleavage reaction.[137] A methoxymethyl (MOM) ether protecting group due to its ease of installation and cleaveage, as well as the fact that the NMR data for the MOM-protected 3-hydroxyglutarate compound was reported in the literature, which would allow calculation of the ee.

Diethyl 3-hydroxyglutarate was protected using MOM chloride and DIPEA in DCM to give **99** in nearly quantitative yield.[138] Chymotrypsin-catalysed cleavage of the ethyl ester was carried under the same conditions that were used for the acetyl-protected version. The cleavage product was obtained in a yield of approximately 10%. No ee was determined since the viability of the route would depend on whether the coupling to Meldrum's acid was successful.

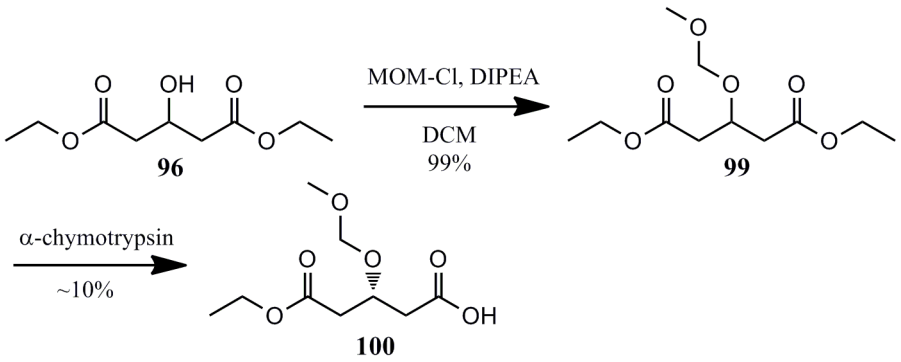


Figure A8: Protection of the hydroxyl group in diethyl 3-hydroxyglutarate (**96**) with MOM-Cl, followed by ester cleavage with chymotrypsin.

100 was coupled with Meldrum's acid using DCC and DMAP, or EDC and DMAP, in DCM. In both cases, no product could be seen.

A1.7 Reasons for discontinuing work

The difficulty in obtaining a good yield from the enzymatic ester cleavage, the fact that only one stereoisomer of the reduced triketide should be formed and the subsequent synthesis required to obtain both forms of the reduced triketide, and the difficulty of the coupling reaction with Meldrum's acid all contributed to this synthetic route being abandoned. Instead, it was decided to pursue the original synthesis but to couple the protected reduced triketide with a pantetheine derivative rather than CoA. This should allow elaboration to the CoA form without any of the associated problems with CoA reactivity (section 4.2.2).

A2 Synthetic approaches to 5-amidinyl teicoplanin

A2.1 Introduction

Teicoplanin is a glycopeptide antibiotic, which ares used as a "drugs of last resort" in the treatment of methicillin-resistant *Staphylococcus aureus* (MRSA) infections. It is marketed under the name Targocid, and has been shown to be effective in the treatment of pseudomembranous colitis and *Clostridium difficile*-associated diarrhoea.[139]

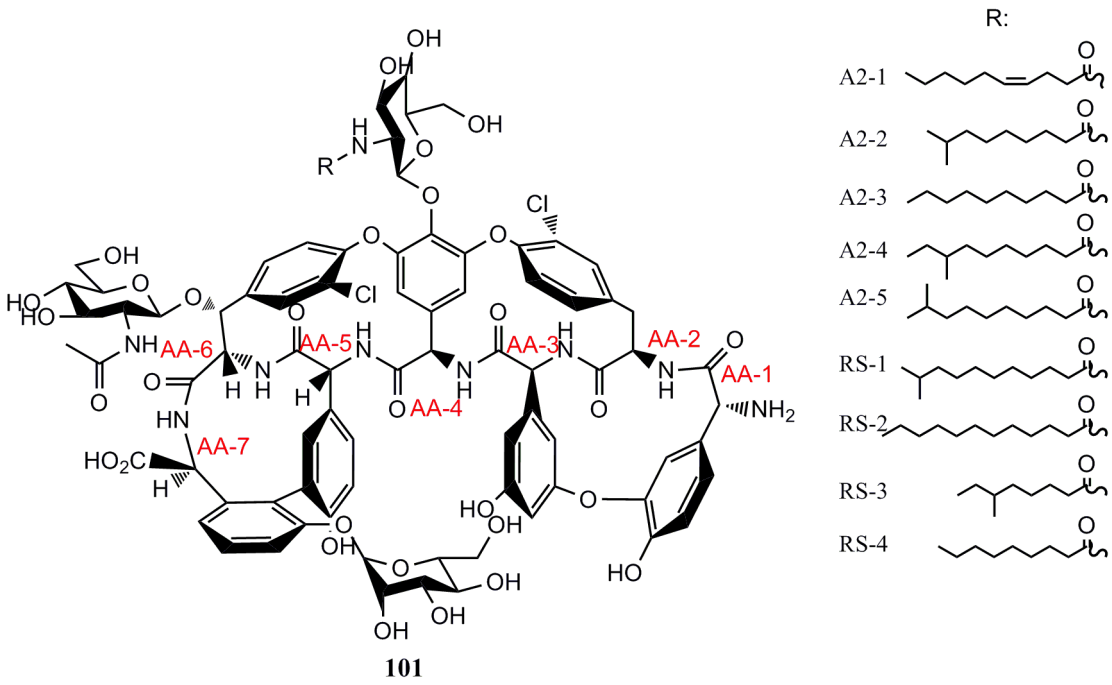


Figure A9: The structure of the glycopeptide antibiotic, teicoplanin **101**. The fatty acid side chain can vary, giving rise to different forms.

Teicoplanin (**101**) is made up of a heptapeptide core, which follows the consistent stereochemical motif for glycopeptides of D-D-L-D-D-L-L. Not only does the chain contain 4 amino acids with the unnatural D configuration, it also incorporates the non-proteinogenic amino acids 4-D-hydroxyphenylglycine (hpg) at positions 1, 4, and 5, (S)-3,5-dihydroxyphenylglycine (Dhpg) at postions 3 and 7 and β-hydroxy-*m*-chlorotyrosine at postion 6. The structure of teicoplanin is identical to that of vancomycin at the C terminus, but the 3 amino acids at the N terminus are different (Figure A9).

The amino acids are cross-linked by ether formation between the aromatic rings on residues 1 and 3, 2 and 4, 4 and 6, and by a C-C bond formation between rings at residues 5 and 7. The aglycon is then coupled to three sugars – two N-acetyl glucosamine molecules and a mannose. A fatty acyl chain is coupled to the N of the glucosamine attached to amino acid 4.

A2.1.1 Biosynthesis of Teicoplanin

The teicoplanin gene cluster in *A. teichomyceticus* was sequenced previously within the group[140] and by Sosio *et al*. [141] It spans about 89 kb and contains 49 putative open reading frames (ORFs), most of which have been assigned via homology to characterised proteins (Figure A10).

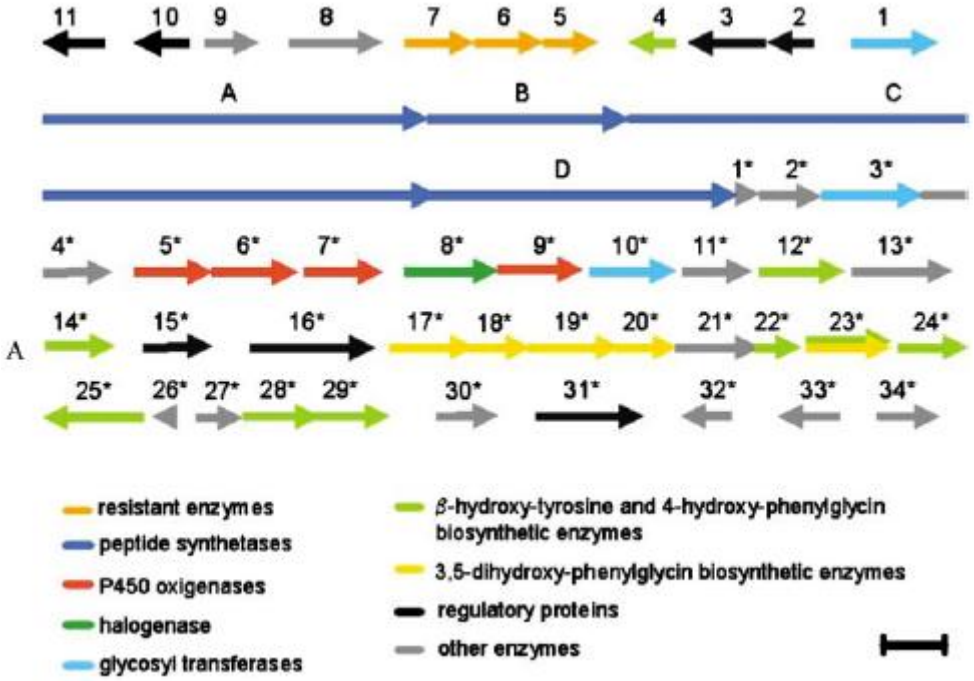


Figure A10: The gene cluster encoding all the enzymes required for teicoplanin biosynthesis. (Reprinted from *Chemistry & Biology,* **11**, Li, T. L.; Huang, F. L.; Haydock, S. F.; Mironenko, T.; Leadlay, P. F.; Spencer, J. B., Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: Characterization of two glycosyltransferases and the key acyltransferase, 107-119, Copyright (2004), with permission from Elsevier.)[140]

A2.1.1.1 Amino acids

Teicoplanin includes the non-proteinogenic amino acids D-hydroxyphenylglycine, L-dihydroxyphenylglycine and β-hydroxytyrosine.

D-hydroxyphenylglycine (**107**) is synthesised from tyrosine by four enzyme catalysed steps. Initially chorismate is produced by the shikimate primary metabolic pathway and is then converted into prephenate (**102**) by Orf14*. The prephenate undergoes decarboxylation by Orf 24* to form 4-hydroxyphenylpyruvate (**103**) which acts as a substrate for Orf28* to form (S)-4-hydroxymandelic acid (**104**). This is then oxidised by Orf29* and transaminated by Orf23*, using tyrosine (**106**) as the amine donor to produce HPG (**107**). The deaminated tyrosine can then be used directly as a substrate for Orf28* (Figure A11).

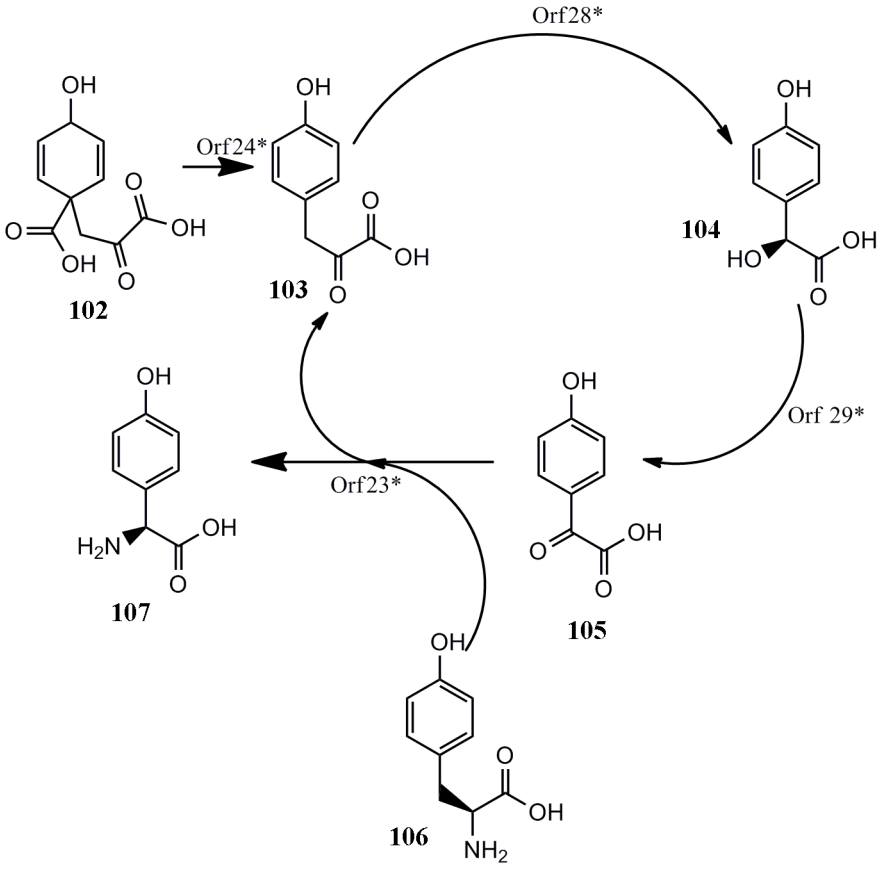


Figure A11: Biosynthesis of D-hydroxyphenylglycine (**107**) from tyrosine, via (S)-4-hydroxymandelic acid (**104**).

Despite their structural similarity, L-dihydroxyphenylglycine (DHPG, **109**) is biosynthesised by a very different pathway. Feeding experiments with acetate showed that DHPG was formed by a polyketide synthase. Experiments in the glycopeptide chloroeremomycin have elucidated a 4 enzyme pathway to DHPG.[142] A Type III PKS forms the cyclised tetraketide (**108**) from four malonyl-CoA units,

followed by dehydration and aromatisation by two enzymes and the addition of oxygen by a mono-oxygenase. The product is then transaminated by HpgT to form DHPG (**109**) (Figure A12).

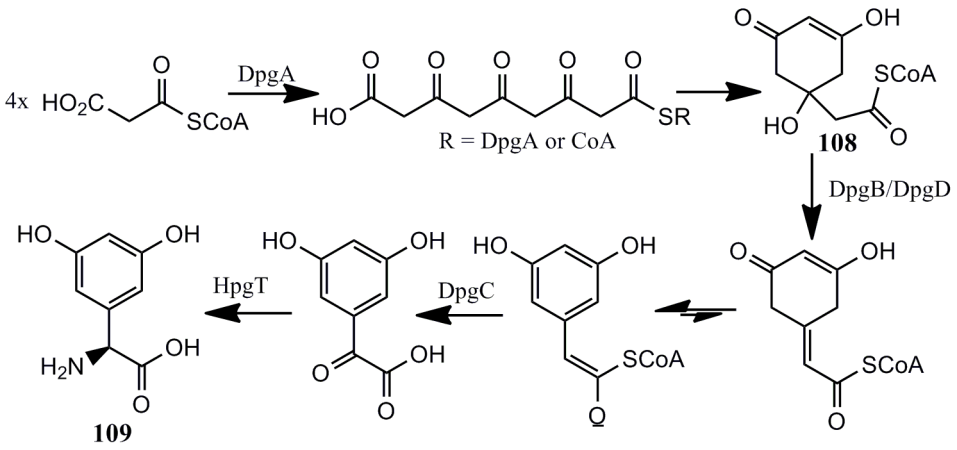


Figure A12: The unnatural amino acid L-dihydroxyphenylglycine (**109**) is biosynthesised by a Type III PKS, DpgA, encoded in the biosynthetic gene cluster for teicoplanin.

The biosynthesis of β-hydroxytyrosine remains cryptic. Feeding studies in the glycopeptide vancomycin indicated that β-hydroxylation occurred before the assembly of the heptapeptide chain and a 3 gene operon was identified for this process.[141] There are no homologues of these genes in the teicoplanin gene cluster and so it is probable that an alternative mechanism for the biosynthesis is present. Gene inactivation studies in the glycopeptide A40926 suggest that Obv28 (Orf12* in the *teic* cluster) is the β-hydroxylase.[143]

A2.1.1.2 Non-ribosomal peptide synthetases

Non-ribosomal peptide synthetases (NRPSs) are molecular assembly lines that are analogous to polyketide synthases. NRPSs catalyse the stepwise condensation of both proteinogenic and specially biosynthesised amino acids.

The adenylation (A) domain (analogous to the acyltransferase domain in PKSs) selects and activates the amino acid. The thiolation (T) domain or peptidyl carrier

protein (PCP, analogous to the ACP in PKSs) is the site of binding for the 4'-phosphopantetheine cofactor and the attachment of the substrate. The condensation (C) domain (analogous to the ketosynthase domain in PKSs) catalyses the formation of the peptide bond between the growing peptide chain attached to the PCP and the activated amino acid attached to the adenylation domain. An additional epimerisation (E) domain is also present in modules that require the incorporation of D- rather than L-amino acids. The first module in the teicoplanin NRPS contains no E domain, despite the amino acid having a D configuration in the final product (Figure A13). This is consistent with other glycopeptide NRPS systems.

In the teicoplanin gene cluster, the genes *teicA-D* encode for four subunits of the NRPS. The NRPS has seven modules in total (one per amino acid incorporated) distributed 2:1:3:1 between the subunits. Each module catalyses a single condensation between an amino acid and the growing peptide chain. A thioesterase domain at the end of the final module cleaves the completed heptapeptide chain from the megasynthase.

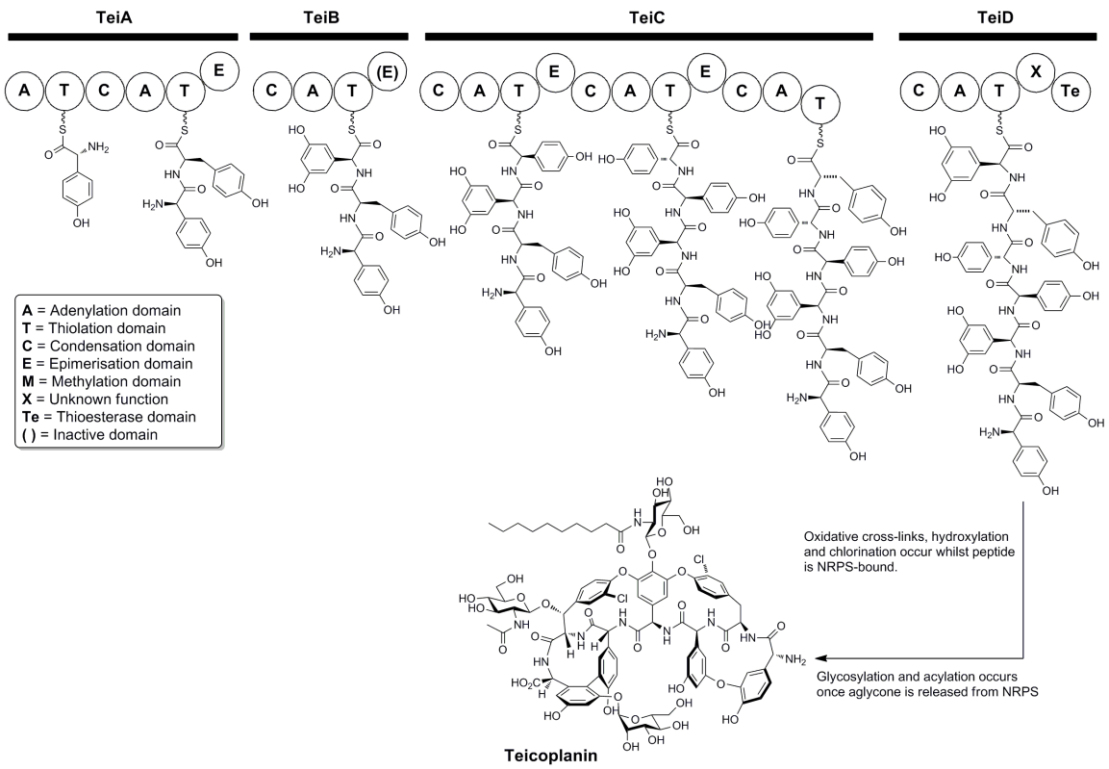


Figure A13: The biosynthesis of the heptapeptide core of teicoplanin by the NRPS proteins TeiA-D.

A2.1.1.3 Post NRPS tailoring

Teicoplanin has two chlorinated amino acids at positions 2 and 6, m-chlorotyrosine and m-chloro-β-hydroxytyrosine, respectively. Feeding experiments have suggested that chlorination occurs after the incorporation of the amino acid into the peptide chain.[144] The chlorination of both residues is proposed to be carried out by a single halogenase, Orf8*.[140]

Gene knockout work in the balhimycin cluster has suggested that the oxidative coupling to form the ring systems is carried out by three oxygenases, OxyA, OxyB and OxyC.[145]

Work on the balhimycin cluster knocking out the Dpg gene for formation of DHPG led to the isolation of a cyclised hexapeptide, implying that the C-O-D ring closure catalysed by OxyB took place while the peptide chain was still bound to the NRPS.[146]

Additional work has also suggested that OxyB may be able to act on both the hexa- and heptapeptide chains.[147]

By homology, is proposed that the cross-linking in teicoplanin is carried out by Orfs 5*, 6*, 7* and 9*.[141]

The teicplanin aglycon is decorated with three sugar moieties, a mannose and two N-acyl-glucosamine residues.

N-acetyl-glucosamine is transferred onto the amino acids at postions 4 and 6 glycosyltransferases Orf1 and Orf 10*. The substrate for both enzymes is UDP-(N-acetyl)-glucosamine.[140]

It is proposed that the mannose is attached by Orf 3*, a putative glycosyltransferase that shows sequence motifs that are characteristic of mannosyl-transferases.[140]
The N-acetyl-glucosamine at amino acid 4 must be deacetylated before the fatty acyl chain can be attached. Work within the group has shown that Orf2*, originally characterised as a nucleotidyltransferase is in fact a deacetylase.[148] The acyl chain is then attached to the free amine by Orf11*.[140]

A2.1.1.4 Mode of action

The bacterial cell wall consists of peptidoglycan, a rigid polymer made up of alternating N-acetyl-glucosamine and N-acteylmuramic acid units. The polymer is cross-linked via short peptide chains. The precursor disaccharide-pentapeptide monomers are synthesised inside the cell and translocated across the cell membrane as a C55 lipid carrier complex. The polymer is then cross-linked outside the cell by transglycosylases and transpeptidases.[149]

Glycopeptides act by disrupting the cross-linking of the peptidoglycan chains that form the cell wall in Gram-positive bacteria. This lack of cross-linking means that the

chains no longer have the tensile strength to withstand the osmotic pressure within the cell and the cell lyses.

Glycopeptides bind to the D-Ala-D-Ala terminus of the extracellular peptidoglycan pentapeptide (N-acetyl-muramyl-L-Ala-D-Glu-Lys-D-Ala-D-Ala) via hydrogen bonds and disrupt both transpeptidase activity (responsible for cross linking the chains) and transglycosylase activity (preventing growth of the peptidoglycan chain).[150] This mode of action explains the glycopeptides' specificity for Gram-positive bacteria. Such large molecules would be unable to pass through the outer lipopolysaccharide membrane that protects the peptidoglycan layer.[151]

Most glycopeptides form a head-to-tail, back-to-back dimer *in vivo.*[152] There are two hypotheses as to why this would be favourable. The first is that the dimer enables the binding of the second glycopeptides to D-Ala-D-Ala to be effectively an intramolecular binding event, and so greatly decreases the entropic cost of binding. The second is that the formation of a dimer allows the amide bonds to become more polarised, thus enhancing the binding to D-Ala-D-Ala.[153,154]

The exception is teicoplanin which, unusually, binds as a monomer. It is believed that the lipid chain enables teicoplanin to anchor itself in the lipid bilayer, postioning itself correctly for binding to peptidoglycan and so making any binding to D-Ala-D-Ala effectively intramolecular again.[152]

A2.1.1.5 Antibiotic resistance

With the rise of bacterial strains resistant to traditional antibiotics such as penicillins, macrolides and cephalosporins, vancomycin became a drug of last resort for treatment of infections with multiply resistant bacteria. Vancomycin-resistance was observed in enterococci (VRE) in 1986 and it was demonstrated in a laboratory environment that this resistance could be transferred to *S. aureus*.[155] In Japan in 1996 the first *S. aureus* strains showing intermediate resistance (VISA) (MIC ≈ 8

ug/ml) was discovered.[156] The mechanism of resistance in VISA is different to that in enterococci, with strains characterised by thickened cell walls instead of a transposable DNA element.[157] Vancomycin-resistant *S. aureus* (VRSA) have been reported since 2002 carrying the transposon responsible for resistance in enterococci.[158]

There are seven different phenotypes (vanA-G) currently thought to be responsible for vancomycin resistance in enterococci, with the most prevalent clinically being vanA and vanB. These disrupt the binding of the antibiotic to the peptidoglycan cell wall by replacing D-Ala-D-Ala with D-Ala-D-Lac (Figure A14). This simple change results in a loss of approximately 4 kcal/mol of binding energy and results in a thousand-fold decrease in binding affinity.[151]

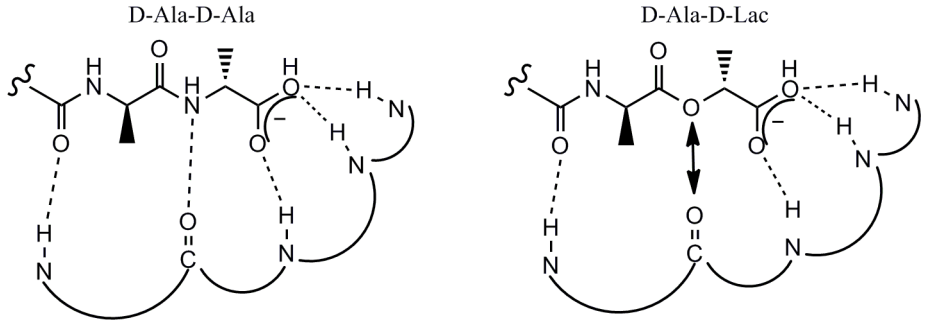


Figure A14: Schematic of the binding of glycopeptides to D-Ala-D-Ala (left) and D-Ala-D-Lac (right). The binding occurs via 5 hydrogen bonds. The substitution of the NH for O leads to the exchange of a hydrogen bond for an electrostatic repulsion (double headed arrow).

Teicoplanin is still active against strains that exhibit low level vanB resistance but the *vanZ* gene in the vanA cassette confers teicoplanin resistance.[159]

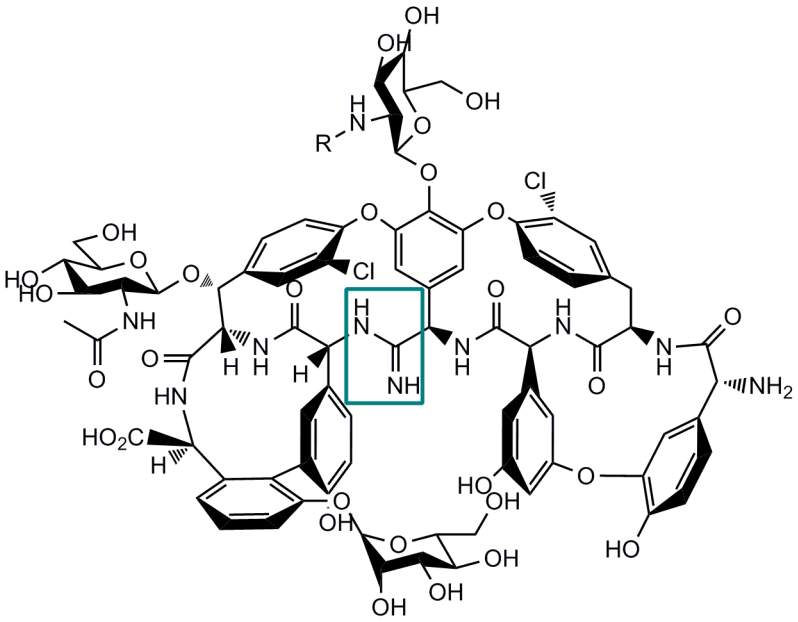
A synthetic teicoplanin analogue synthesised by Crowley and Boger, where the carbonyl from residue 4 was replaced with a methylene group. The analogue exhibited a 40-fold increase in affinity for D-Ala-D-Lac over the aglycon, but showed a 35-fold reduction in binding to D-Ala-D-Ala. Although the reduced binding would

hinder the activity of the antibiotic against non-resistant *S. aureus*, the analogue exhibited antimicrobial activity against a vanA-resistant organism (MIC of 31 μg/ml).[160]

A2.1.1.6 5-amidinyl teicoplanin

The aim of this project is to produce a teicoplanin analogue that can bind to D-Ala-D-Lac whilst retaining binding to D-Ala-D-Ala, by a chemoenzymatic method. It is proposed to disrupt the NRPS assembly line to halt the biosynthesis of the heptapeptide core. A novel peptide could be synthesised chemically then loaded onto the NRPS where it will be extended, cyclised and tailored into a novel teicoplanin analogue.

It is proposed to introduce an amidine linkage between residues 4 and 5 of the teicoplanin heptapeptide core (Figure A15). This has both hydrogen bond donor and hydrogen bond acceptor characteristics and so should bind to D-Ala-D-Lac while still binding to D-Ala-D-Ala.



110

Figure A15: Proposed structure of 5-amidinyl teicoplanin **110**, with the amidine linkage between residues 4 and 5.

Work by Robinson *et al.* has shown that is possible to attach a synthetically constructed peptide chain to the NRPS using Sfp and that it will continue to be enzymatically elaborated into the finished glycopeptide.[161]

If the appropriate peptidyl carrier protein (PCP) were knocked out on the NRPS it could be possible to feed a chemically synthesised peptide chain incorporating the amidine linkage. This could be incorporated into the assembly line and extended into the altered teicoplanin. The knockout of the PCP would remove the potential for background formation of wild-type teicoplanin. It would be most efficient to synthesise only the minimum chain length that is required for the amidine to be installed between amino acids 4 and 5 (Figure A16).

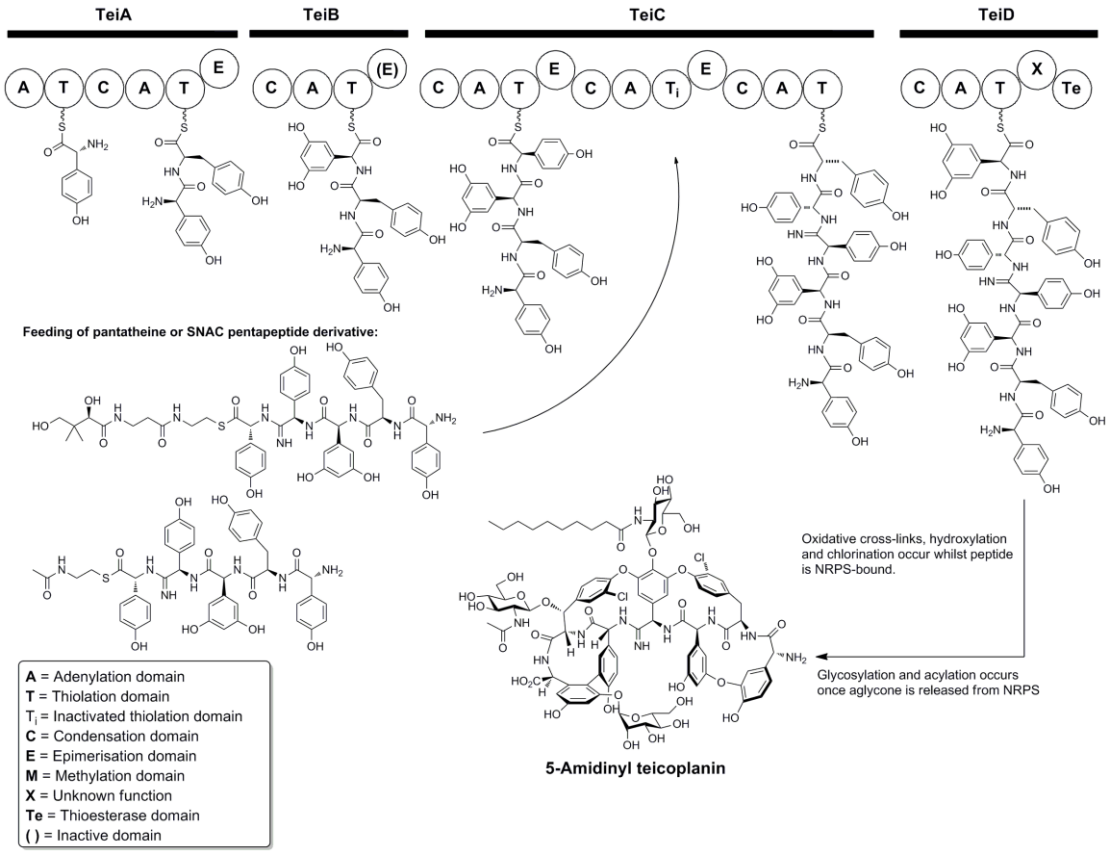


Figure A16: The proposed chemoenzymatic synthesis of 5-amidinylteicoplanin. The teicoplanin assembly line would be disrupted at the thiolation domain in module 5. A pentapeptide incorporating the amidine linkage could be fed as an SNAC derivative and extended to the altered teicoplanin.

A2.2 Retrosynthetic analysis

The SNAC or pantetheine derivatives could be formed from the pentapeptide chain that contains an amidine linkage. Disconnecting the pentapeptide (**111**) along the bonds indicated leads to two dipeptide fragments and the commercially available (S)-dihydroxyphenylglycine (Dhpg, **109**). As Dhpg is only available in small quantities it would be best to reduce the number of coupling steps by synthesising the 1-2 (**113**) and 4-5 (**112**) building blocks separately, rather than building up the peptide chain in a sequential manner (Figure A17).

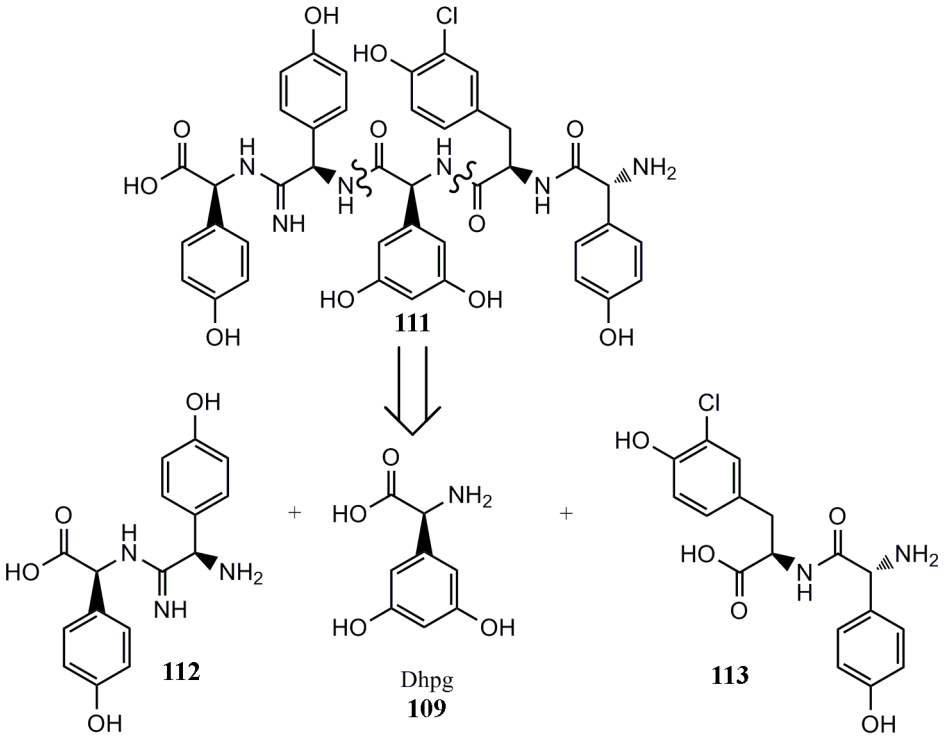


Figure A17: Disconnection of the pentapeptide to two dipeptide units and the commercially available dihydroxyphenylglycine.

The 1-2 building block could be disconnected into commercially available D-hydroxyphenylglycine (**107**) and m-chloro-D-tyrosine (**114**) formed from commercially available D-tyrosine (**115**) and sulfuryl chloride (Figure A18).

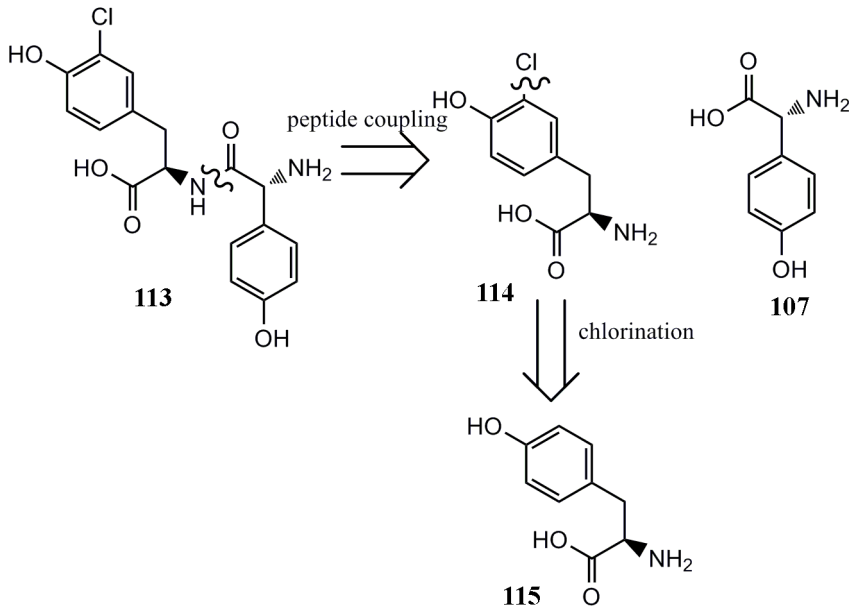


Figure A18: Disconnection of amino acids 1 and 2 to D-hydroxyphenylglycine and m-chlorotyrosine.

It was envisaged that the amidine could be disconnected into D-hydroxyphenylglycine (**107**) and a nitrile (**116**). The nitrile could be formed by the oxidation of D-hydroxyphenylglycinol (**117**) in the presence of ammonia (Figure A19).

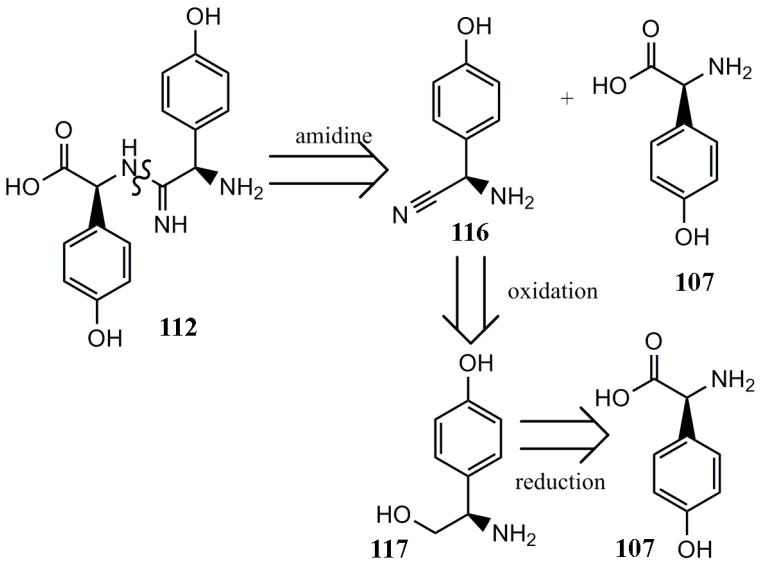


Figure A20: Disconnection of the amidine linking amino acids 4 and 5 into a nitrile and an amine.

A2.3 Synthesis of 1-2 dipeptide

A2.3.1 Chlorination

The chlorination of D-Tyrosine (**115**) was performed in ether. The reaction proceeded either with 50% conversion to the mono-chlorinated derivative (by ^{1}H NMR) using 3 equivalents of sulfuryl chloride or with complete conversion to the di-chlorinated amino acid (by ^{1}H NMR) using 10 equivalents. It was not possible to separate the chlorinated and unchlorinated tyrosines by TLC, and it was decided that the separation would be best performed on the coupled product.

A2.3.2 Coupling of amino acids

Boc-protected hpg was coupled to the mixture of m-chloro-D-tyrosine (**114**) and D-tyrosine (**115**) using DCC, HOBt and DMAP. The resulting mixture of compounds was purified by column chromatography and the fractions analysed by LC-MS. One fraction was found to contain the correct compound in a 12% yield. For repeating the experiment it was suggested that using EDC as a coupling reagent would give better results. The synthesis was not been repeated due to issues with protecting groups.

A2.3.3 Protecting groups

The synthesis of the 1-2 dipeptide (**113**) was hindered by lack of appropriate protecting groups. The N of hpg was easily protected as both an Fmoc- and Boc-derivative. The carboxylic acid of D-tyrosine proved more difficult to protect, with repeated attempts to install a ^{t}Bu ester failing. A methyl ester was successfully installed using thionyl chloride and methanol, but the conditions for removal of this group could result in racemisation of the peptide chain. A benzyl ester protecting group was successfully installed and this strategy could be carried forward using N-Boc protection.

A2.4 Amidine trial synthesis

On examination of the literature it became apparent that although the formation of terminal amidines was well documented, the formation of substituted amidines was underreported.

The synthesis of terminal amidines using a mild method under neutral conditions was reported by Lange *et al.*,[162] using N-acetylcysteine as a catalyst (Figure A21).

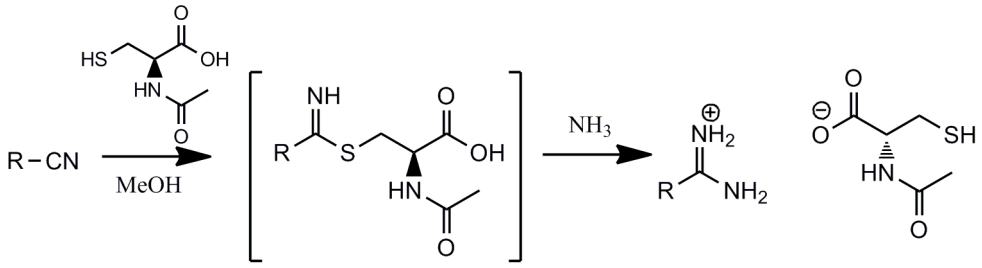


Figure A21: A mild method for the synthesis of amidines using N-acetyl cysteine as a catalyst.[162]

It was proposed that replacing ammonia with a substituted amine should lead to the production of the substituted amidine.

A2.4.1 Trial reactions

The formation of substituted amidines was tried using 3-(trifluoromethyl)benzonitrile (**118**), which had been used successfully in the Lange paper (Figure A22). The nitrile was stirred with triethylamine, isopropylamine or benzylamine and N-acetylcysteine for four days in dry methanol at 50˚C (33˚C for isopropylamine). The product was then analysed via ^{13}C NMR and LC-MS.

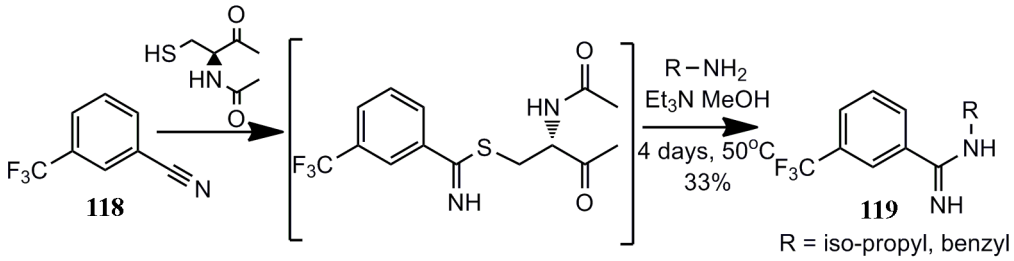


Figure A22: Trial synthesis of amidines using N-acetyl cysteine as a catalyst and iso-propylamine and benzylamine as amine donors.

The ^{13}C NMR showed the loss of the carbon signal at 117 ppm assigned to the nitrile carbon, and the introduction of a signal at 138 ppm that was assigned to the carbon bearing the amidine group. No starting material signals were visible in the product NMR.

It was decided to trial the amidine-forming ability of the amino group of hpg with the successfully used 3-(trifluoromethyl)benzonitrile. This would suggest conditions that could be used to form the final amidine product with the nitrile from hpg. D-hydroxyphenylglycine methyl ester was stirred at 50°C with N-acetylcysteine and triethylamine in methanol. The solution turned bright yellow and an orange oil was extracted. Unfortunately no amidine was detected by either ^{13}C NMR or direct injection mass spectrometry. This could be due to the reduced nucleophilicity of the amino acid amine, in which case a longer reaction time may lead to product formation.

A2.4.2 Amidine stability

In order to assist with the choice of protecting groups for the overall synthesis, the stability of the amidine group was investigated. **119** was stirred overnight in water at 37°C, for 24 hours in TFA and overnight in piperidine and the results analysed by ^{13}C NMR.

There was no difference in the ^{13}C NMR for the amidine (**119**) before and after treatment with water, TFA and piperidine. In particular there was no reformation of the nitrile group by acid cleavage of the amidine or substitution of the benzylamine by piperidine. This enables both Boc and Fmoc protecting groups to be used. The water incubation was used to assess the stability of the amidine in conditions that would be used in future enzymic steps and ensuring that the pentapeptide analogue would be sufficiently stable to be loaded on to the NRPS.

A2.4.3 Nitrile formation

The oxidation of a primary alcohol to a nitrile would be a simple route to providing the correct stereochemistry for the nitrile.

Protected hpg was reduced using lithium aluminium hydride to produce the alcohol (**120**). Both Boc- and Fmoc-protected amino acids were reduced successfully, although Fmoc protection was less stable under the reduction conditions.

Both Boc-D-hydroxyphenylglycinol and Fmoc-D-hydroxyphenylglycinol were stirred with iodine in ammonia water at 60°C for 24 hours (Figure A23).[163] In both cases no product was isolated by NMR. It would be worth repeating the experiment with a protecting group on the phenolic OH in case this interfered with nitrile formation.

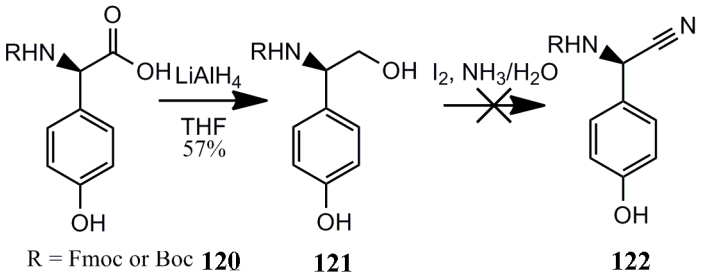


Figure A23: Attempted synthesis of D-hydroxyphenylglycine-nitrile (**122**) from D-hydroxyphenylglycine.

A2.4.4 Synthesis of nitrile

Due to the difficulty in synthesising the enantiomerically resolved nitrile, it was decided to proceed with the racemic nitrile, which could be formed via the Strecker synthesis from 4-hydroxybenzaldehyde (**123**).

The hydroxyl group was protected with a MOM-ether using MOM-Cl and DIPEA in dichloromethane, to give **124** in a 99% yield. The protected aldehyde **124** was reacted with potassium cyanide, ammonium chloride and ammonia in methanol, to give protected nitrile **125** in 76% yield. **125** could be deprotected by stirring in methanol and 2 M HCl to give the crude nitrile **126**.

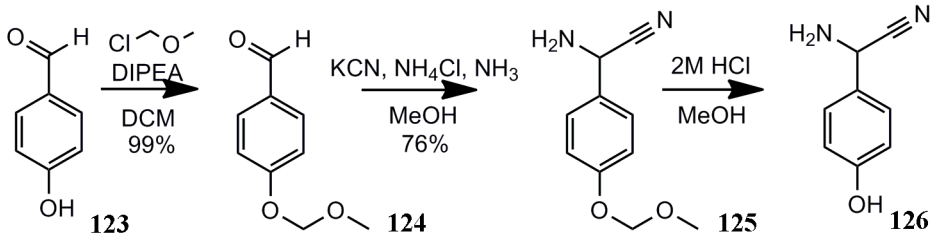


Figure A24: Synthesis of racemic nitrile **126** from 4-hydroxybenzaldehyde.

A2.5 Feeding of nitrile

Nitriles have been observed to be inhibitors of cysteine proteases due to the formation of a thioimidate. This implies that an activated thiol group, such as the one on the PCP could form a covalent bond with the nitrile. This might then undergo a condensation reaction with the growing peptide chain to form an amidine *in vivo*.

This could result in the formation of 5-amidinyl teicoplanin by the NRPS. As dihydroxyphenylglycine is incoporated at postions 1, 4, and 5 in the teicoplanin backbone, the formation of amidine linkages between amino acids 1 and 2, and positions 5 and 6 would also be possible, and separation of these compounds would be challenging.

Crude deprotected nitrile **126** (15 mg total) was fed to cultures of the teiocoplanin-producing strain *A. orientalis* at 36, 40, 44 and 48 hours. Growth was continued until 110 hours, as for normal teicoplanin production. The cultures were spun down and the supernatant was filtered and analysed by LCMS. No peaks of the correct mass could be seen for amidinyl teicoplanin, while production of teicoplanin was observed in both control and fed cultures. A bioassay for teicoplanin production after 96 hours using activity against *B. subtilis* showed that approximately equal amounts of antibiotic was produced by both control and fed cultures. This implies that the nitrile did not inhibit the biosynthesis, and so probably did not form a covalent intermediate with active thiols present on the NRPS.

A2.6 Reasons for discontinuing work

The difficulties encountered in the synthesis of the amidine, along with the inherent challenges in synthesising the pentapeptide analogue, attaching it to the NRPS, elaborating it to 5-amidinyl teicoplanin and finally detection of the product, meant that this approach had an extremely limited chance of success. It was therefore decided to concentrate on the 6-MSA project.

CHAPTER 6

References

1. Baker, D. D.; Chu, M.; Oza, U.; Rajgarhia, V., The value of natural products to future pharmaceutical discovery. *Nat. Prod. Rep.* **2007,** *24*, 1225-1244.

2. Davies, J., In praise of antibiotics. *ASM News* **1999,** *65*, 304-310.

3. Schatz, A.; Bugie, E.; Waksman, S. A., Streptomycin, a substance exhibiting antibiotic activity against gram positive and gram-negative bacteria. *Proc. Soc. Exp. Biol. Med.* **1944,** *55*, 66-69.

4. Weinstein, M. J.; Rosselet, J. P.; Herzog, H. L.; Black, J.; Luedemann, G. M.; Wagman, G. H.; Charney, W.; Marquez, J. A.; Coniglio, C. T.; Oden, E. M., Gentamicin, a new antibiotic comples from Micromonspora. *J. Med. Chem.* **1963,** *6*, 463-464.

5. Boothe, J. H.; Morton, J.; Petisi, J. P.; Wilkinson, R. G.; Williams, J. H., Tetracycline. *J. Am. Chem. Soc.* **1953,** *75*, 4621-4621.

6. Vannucchi, V., Clinical study of a new antibiotic; erythromycin. *Riv. Crit. Clin. Med.* **1952,** *52*, 128-36.

7. Peacock, J. E.; Moorman, D. R.; Wenzel, R. P.; Mandell, G. L., Methicillin-resistant Staphylococcus aureus - microbiological characteristics, antimicrobial susceptibilities, and assessment of virulence of an epidemic strain. *J. Infect. Dis.* **1981,** *144*, 575-582.

8. Uttley, A. H. C.; Collins, C. H.; Naidoo, J.; George, R. C., Vancomycin-resistant Enterococci. *Lancet* **1988,** *1*, 57-58.

9. Shin, S. S.; Keshavjee, S.; Gelmanova, I. Y.; Atwood, S.; Franke, M. F.; Mishustin, S. P.; Strelis, A. K.; Andreev, Y. G.; Pasechnikov, A. D.; Barnashov, A.; Tonkel, T. P.; Cohen, T., Development of extensively drug-resistant tuberculosis during multidrug-resistant tuberculosis treatment. *Am. J. Respir. Crit. Care Med.* **2010,** *182*, 426-432.

10. Oura, M.; Sternberg, T. H.; Wright, E. T., A new antifungal antibiotic, amphotericin B. *Antibiot. Annu.* **1955,** *3*, 566-73.

11. Sehgal, S. N.; Baker, H.; Vezina, C., Rapamycin (AY-22,989), a new antifungal antibiotic .2. Fermentation, isolation and characterization. *J. Antibiot.* **1975,** *28*, 727-732.

12. a) Brown, A. G.; Smale, T. C.; King, T. J.; Hasenkamp, R.; Thompson, R. H., Crystal and molecular structure of compactin, a new antifungal metabolite from Penicillium brevicompactum. *J. Chem. Soc., Perkin Trans. 1* **1976,** 1165-1173;

b) Alberts, A. W.; Chen, J.; Kuron, G.; Hunt, V.; Huff, J.; Hoffman, C.; Rothrock, J.; Lopez, M.; Joshua, H.; Harris, E.; Patchett, A.; Monaghan, R.; Currie, S.; Stapley, E.; Albersschonberg, G.; Hensens, O.; Hirshfield, J.; Hoogsteen, K.; Liesch, J.; Springer, J., Mevinolin - a highly potent competitive inhibitor of hydroxymethylglutaryl-coenzyme-A reductase and a cholesterol-lowering agent. *Proc. Nat. Acad. Sci. U.S.A.* **1980,** *77*, 3957-3961.

13. Breinbauer, R.; Vetter, I. R.; Waldmann, H., From protein domains to drug candidates - Natural products as guiding principles in the design and synthesis of compound libraries. *Angew. Chem., Int. Ed.* **2002,** *41*, 2879-2890.

14. Dixon, N.; Wong, L. S.; Geerlings, T. H.; Micklefield, J., Cellular targets of natural products. *Nat. Prod. Rep.* **2007,** *24*, 1288-1310.

15. Nudelman, I.; Rebibo-Sabbah, A.; Cherniavsky, M.; Belakhov, V.; Hainrichson, M.; Chen, F. Q.; Schacht, J.; Pilch, D. S.; Ben-Yosef, T.; Baasov, T., Development of Novel Aminoglycoside (NB54) with Reduced Toxicity and Enhanced Suppression of Disease-Causing Premature Stop Mutations. *J. Med. Chem.* **2009,** *52*, 2836-2845.

16. Kuhajda, F. P.; Pizer, E. S.; Li, J. N.; Mani, N. S.; Frehywot, G. L.; Townsend, C. A., Synthesis and antitumor activity of an inhibitor of fatty acid synthase. *Proc. Natl. Acad. Sci. U. S. A.* **2000,** *97*, 3450-3454.

17. Weissman, K. J., Introduction to polyketide biosynthesis. *Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates,* **2009,** *459*, 3-16.

18. Galm, U.; Hager, M. H.; Van Lanen, S. G.; Ju, J. H.; Thorson, J. S.; Shen, B., Antitumor antibiotics: Bleomycin, endiynes, and mitomycin. *Chem. Rev.* **2005,** *105*, 739-758.

19. Murata, M.; Naoki, H.; Matsunaga, S.; Satake, M.; Yasumoto, T., Structure and partial stereochemical assignments for maitotoxin, the most toxic and largest natural non-biopolymer. *J. Am. Chem. Soc.* **1994,** *116*, 7098-7107.

20. Collie, J.; Myers, W., The formation of orcinol and other condensation products from dehydracetic acid. *J. Chem. Soc., Trans.* **1893,** *63*, 122-128.

21. Collie, J.; Chrystall, E., The production of orcinol derivatives from the sodium salt of ethyl acetoacetate by the action of heat. *J. Chem. Soc.* **1907**, 1802-1806.

22. Collie, J., Derivatives of the multiple keten group. *J. Chem. Soc.* **1907**, 1806-1813.

23. Staunton, J.; Weissman, K., Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* **2001**, 380-416.

24. Birch, A.; Massywestropp, R.; Moye, C., Studies in relation to biosynthesis .7. 2-Hydroxy-6-methylbenzoic acid in Penicillium griseofulvum Dierckx. *Aust. J. Chem.* **1955**, 539-544.

25. Chan, D. I.; Vogel, H. J., Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochem. J.* **2010,** *430*, 1-19.

26. Larrabee, A.; McDaniel, E.; Bakerman, H.; Vagelos, P., Acyl carrier protein .V. Identification of 4'-phosphopantetheine bound to a mammalian fatty acid synthetase preparation. *Proc. Natl. Acad. Sci. U. S. A.* **1965,** *54*, 267-273.

27. Misra, A.; Sharma, S. K.; Surolia, N.; Surolia, A., Self-acylation properties of type II fatty acid biosynthesis acyl carrier protein. *Chem. Biol.* **2007,** *14*, 775-783.

28. White, S. W.; Zheng, J.; Zhang, Y. M.; Rock, C. O., The structural biology of type II fatty acid biosynthesis. *Annu. Rev. Biochem.* **2005,** *74*, 791-831.

29. Jenni, S.; Leibundgut, M.; Maier, T.; Ban, N., Architecture of a fungal fatty acid synthase at 5 angstrom resolution. *Science* **2006**, 1263-1267.

30. Jenni, S.; Leibundgut, M.; Boehringer, D.; Frick, C.; Mikolasek, B.; Ban, N., Structure of fungal fatty acid synthase and implications for iterative substrate shuttling. *Science* **2007**, 254-261.

31. Leibundgut, M.; Maier, T.; Jenni, S.; Ban, N., The multienzyme architecture of eukaryotic fatty acid synthases. *Curr. Opin. Struct. Biol.* **2008**, 714-725.

32. a) Maier, T.; Jenni, S.; Ban, N., Architecture of mammalian fatty acid synthase at 4.5 angstrom resolution. *Science* **2006**, 1258-1262;

b) Maier, T.; Leibundgut, M.; Ban, N., The crystal structure of a mammalian fatty acid synthase. *Science* **2008**, 1315-1322.

33. Hertweck, C., The Biosynthetic Logic of Polyketide Diversity. *Angew. Chem., Int. Ed.* **2009**, 4688-4716.

34. a) Bevitt, D. J.; Cortes, J.; Haydock, S. F.; Leadlay, P. F., 6-Deoxyerythronolide-B synthase-2 from Saccharopolyspora erythraea - cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme. *Eur. J. Biochem.* **1992,** *204*, 39-49;

b) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L., Modular organization of genes required for complex polyketide biosynthesis. *Science* **1991,** *252*, 675-679.

35. Sutherland, A.; Auclair, K.; Vederas, J. C., Recent advances in the biosynthetic studies of lovastatin. *Curr. Opin. Drug. Discov. Devel.* **2001,** *4*, 229-36.

36. Auclair, K.; Sutherland, A.; Kennedy, J.; Witter, D. J.; Van den Heever, J. P.; Hutchinson, C. R.; Vederas, J. C., Lovastatin nonaketide synthase catalyzes an intramolecular Diels-Alder reaction of a substrate analogue. *J. Am. Chem. Soc.* **2000,** *122*, 11519-11520.

37. Campbell, C. D.; Vederas, J. C., Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. *Biopolymers* **2010,** *93*, 755-763.

38. Brown, M. P.; Brown-Jenco, C. S.; Payne, G. A., Genetic and molecular analysis of aflatoxin biosynthesis. *Fungal Genet. Biol.* **1999,** *26*, 81-98.

39. Udwary, D.; Merski, M.; Townsend, C., A method for prediction of the locations of linker regions within large multifunctional proteins, and application to a type I polyketide synthase. *J. Mol. Biol.* **2002**, 585-598.

40. Crawford, J.; Korman, T.; Labonte, J.; Vagstad, A.; Hill, E.; Kamari-Bidkorpeh, O.; Tsai, S.; Townsend, C., Structural basis for biosynthetic programming of fungal aromatic polyketide cyclization. *Nature* **2009**, 1139-1143.

41. Crawford, J.; Dancy, B.; Hill, E.; Udwary, D.; Townsend, C., Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 16728-16733.

42. Crawford, J.; Thomas, P.; Scheerer, J.; Vagstad, A.; Kelleher, N.; Townsend, C., Deconstruction of iterative multidomain polyketide synthase function. *Science* **2008**, 243-246.

43. Cox, R.; Simpson, T., Fungal type I polyketide synthases. *Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates* **2009**, 49-78.

44. Metz, J. G.; Roessler, P.; Facciotti, D.; Levering, C.; Dittrich, F.; Lassner, M.; Valentine, R.; Lardizabal, K.; Domergue, F.; Yamada, A.; Yazawa, K.; Knauf, V.; Browse, J., Production of polyunsaturated fatty acids by polyketide synthases in both prokaryotes and eukaryotes. *Science* **2001,** *293*, 290-293.

45. Arthur, C. J.; Szafranska, A.; Evans, S. E.; Findlow, S. C.; Burston, S. G.; Owen, P.; Clark-Lewis, I.; Simpson, T. J.; Crosby, J.; Crump, M. P., Self-malonylation is an

intrinsic property of a chemically synthesized type II polyketide synthase acyl carrier protein. *Biochemistry* **2005,** *44*, 15414-15421.

46. Hertweck, C.; Luzhetskyy, A.; Rebets, Y.; Bechthold, A., Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat. Prod. Rep.* **2007,** *24*, 162-190.

47. Weissman, K. J.; Leadlay, P. F., Combinatorial biosynthesis of reduced polyketides. *Nat. Rev. Microbiol.* **2005,** *3*, 925-936.

48. a) Reeves, C. D.; Ward, S. L.; Revill, W. P.; Suzuki, H.; Marcus, M.; Petrakovsky, O. V.; Marquez, S.; Fu, H.; Dong, S. D.; Katz, L., Production of hybrid 16-membered macrolides by expressing combinations of polyketide synthase genes in engineered Streptomyces fradiae hosts. *Chem. Biol.* **2004,** *11*, 1465-1472;

b) Gaisser, S.; Kellenberger, L.; Kaja, A. L.; Weston, A. J.; Lill, R. E.; Wirtz, G.; Kendrew, S. G.; Low, L.; Sheridan, R. M.; Wilkinson, B.; Galloway, I. S.; Stutzman-Engwall, K.; McArthur, H. A. I.; Staunton, J.; Leadlay, P. F., Direct production of ivermectin-like drugs after domain exchange in the avermectin polyketide synthase of Streptomyces avermitilis ATCC31272. *Org. Biomol. Chem.* **2003,** *1*, 2840-2847;

c) Rowe, C. J.; Bohm, I. U.; Thomas, I. P.; Wilkinson, B.; Rudd, B. A. M.; Foster, G.; Blackaby, A. P.; Sidebottom, P. J.; Roddis, Y.; Buss, A. D.; Staunton, J.; Leadlay, P. F., Engineering a polyketide with a longer chain by insertion of an extra module into the erythromycin-producing polyketide synthase. *Chem. Biol.* **2001,** *8*, 475-485.

49. Kennedy, J., Mutasynthesis, chemobiosynthesis, and back to semi-synthesis: combining synthetic chemistry and biosynthetic engineering for diversifying natural products. *Nat. Prod. Rep.* **2008,** *25*, 25-34.

50. Weissman, K. J., Mutasynthesis - uniting chemistry and genetics for drug discovery. *Trends Biotechnol.* **2007,** *25*, 139-142.

51. Denoya, C. D.; Fedechko, R. W.; Hafner, E. W.; McArthur, H. A. I.; Morgenstern, M. R.; Skinner, D. D.; Stutzmanengwall, K.; Wax, R. G.; Wernau, W. C., A 2nd branched chain alpha-keto acid dehydrogenase gene cluster (BKDFGH) from Streptomyces avermitilis - its relationship to avermectin biosynthesis and the construction of a BKDF mutant suitable for the production of novel antiparasitic avermectins. *J. Bacteriol.* **1995,** *177*, 3504-3511.

52. Shoolingin-Jordan, P.; Campuzano, I., in *Comprehensive Natural Products*. ed. U. Sankawa, Elsevier: Oxford, 1999, 345-365.

53. Arihood, S.; Light, R., Inhibition of 6-methylsalicylic acid synthesis in Penicillium patulum. *Nature* **1966**, 629-630.

54. Dimroth, P.; Walter, H.; Lynen, F., Biosynthesis of 6-methylsalicylic acid. *Eur. J. Biochem.* **1970,** *13*, 98-110.

55. Scott, A.; Beadling, L.; Georgopa, N.; Subbaray, C., Biosynthesis of polyketides - purification and inhibition studies of 6-methylsalicylic acid synthase. *Bioorg. Chem.* **1974**, 238-248.

56. Spencer, J.; Jordan, P., Purification and properties of 6-methylsalicylic acid synthase from Penicillium patulum. *Biochem. J.* **1992**, 839-846.

57. Beck, J.; Ripka, S.; Siegner, A.; Schiltz, E.; Schweizer, E., The multifunctional 6-methylsalicylic acid synthase gene of Penicillium patulum - its gene structure relative to that of other polyketide synthases. *Eur. J. Biochem.* **1990**, 487-498.

58. Dimroth, P.; Ringelmann, E.; Lynen, F., 6-Methylsalicylic acid synthetase from Penicillium patulum - some catalytic properties of enzyme and its relation to fatty acid synthetase. *Eur. J. Biochem.* **1976**, 591-596.

59. Kealey, J.; Liu, L.; Santi, D.; Betlach, M.; Barr, P., Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, 505-509.

60. Bedford, D.; Schweizer, E.; Hopwood, D.; Khosla, C., Expression of a functional fungal polyketide synthase in the bacterium Streptomyces coelicolor A3(2). *J. Bacteriol.* **1995**, 4544-4548.

61. Wattanachaisaereekul, S.; Regueira, T.; Lantz, A.; Nielsen, M.; Nielsen, J., Hetrologous expression of 6-methylsalicylic acid synthase in Saccharomyces cerevisiae and Aspergillus sp. *J. Biotechnol.* **2005**, S126-S126.

62. Wattanachaisaereekul, S.; Lantz, A.; Nielsen, M.; Nielsen, J., Production of the polyketide 6-MSA in yeast engineered for increased malonyl-CoA supply. *Metab. Eng.* **2008**, 246-254.

63. Yalpani, N.; Altier, D.; Barbour, E.; Cigan, A.; Scelonge, C., Production of 6-methylsalicyclic acid by expression of a fungal polyketide synthase activates disease resistance in tobacco. *Plant Cell* **2001**, 1401-1409.

64. Child, C. J.; Spencer, J. B.; Bhogal, P.; ShoolinginJordan, P. M., Structural similarities between 6-methylsalicylic acid synthase from Penicillium patulum and vertebrate type I fatty acid synthase: Evidence from thiol modification studies. *Biochemistry* **1996,** *35*, 12267-12274.

65. Child, C.; Shoolingin-Jordan, P., Inactivation of the polyketide synthase, 6-methylsalicylic acid synthase, by the specific modification of Cys-204 of the beta-ketoacyl synthase by the fungal mycotoxin cerulenin. *Biochem. J.* **1998**, 933-937.

66. a) Yalpani, M.; Willecke, K.; Lynen, F., Triacetic acid lactone a derailment product of fatty acid biosynthesis. *Eur. J. Biochem.* **1969,** *8*, 495;

b) Nixon, J. E.; Putz, G. R.; Porter, J. W., Synthesis of triacetic acid lactone by pigeon liver fatty acid synthetase complex. *J. Biol. Chem.* **1968,** *243*, 5471.

67. Abell, C.; Staunton, J., The use of H-2 NMR-spectroscopy and beta-isotopic shifts in the C-13 NMR-spectrum to measure deuterium retention in the biosynthesis of the polyketide 6-methylsalicylic acid. *J. Chem. Soc., Chem. Comm.* **1981**, 856-858.

68. Abell, C.; Staunton, J., Biosynthesis of 6-methylsalicylic acid - the combined use of mono-deuteriated and tri-deuteriated acetate precursors to investigate the degree of stereocontrol in the aromatization sequence. *J. Chem. Soc., Chem. Comm.* **1984**, 1005-1007.

69. a) Jordan, P.; Spencer, J., Stereospecific manipulation of hydrogen atoms with opposite absolute orientations during the biosynthesis of the polyketide 6-methylsalicylic acid from chiral malonates in Penicillium patulum. *J. Chem. Soc., Chem. Commun.* **1990**, 238-242;

b) Spencer, J.; Jordan, P., Use of chiral malonates to determine the absolute configuration of the hydrogen atoms eliminated during the formation of 6-methylsalicylic acid by 6-methylsalicylic acid synthase from Penicillium patulum. *J. Chem. Soc., Chem. Commun.* **1990**, 1704-1706.

70. Spencer, J.; Jordan, P., Investigation of the mechanism and steric course of the reaction catalyzed by 6-methylsalicylic acid synthase from Penicillium patulum using (R)-[1-C-13 2-H-2]malonate and (S)-[1-C-13 2-H-2]malonate. *Biochemistry* **1992**, 9107-9116.

71. Cane, D. E.; Prabhakaran, P. C.; Tan, W.; Ott, W. R., Macrolide biosynthesis. 6 Mechanism of polyketide chain elongation. *Tet. Lett. 32*, 5457-5460.

72. Cortes, J.; Wiesmann, K. E. H.; Roberts, G. A.; Brown, M. J. B.; Staunton, J.; Leadlay, P. F., Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage. *Science* **1995,** *268*, 1487-1489.

73. Wu, N.; Cane, D. E.; Khosla, C., Quantitative Analysis of the Relative Contributions of Donor Acyl Carrier Proteins, Acceptor Ketosynthases, and Linker Regions to Intermodular Transfer of Intermediates in Hybrid Polyketide Synthases. *Biochemistry* **2002,** *41*, 5056-5066.

74. Castonguay, R.; He, W.; Chen, A. Y.; Khosla, C.; Cane, D. E., Stereospecificity of Ketoreductase Domains of the 6-Deoxyerythronolide B Synthase. *J. Am. Chem. Soc.* **2007,** *129*, 13758-13769.

75. Hicks, L.; O'Connor, S.; Mazur, M.; Walsh, C.; Kelleher, N., Mass spectrometric interrogation of thioester-bound intermediates in the initial stages of epothilone biosynthesis. *Chem. Biol.* **2004**, 327-335.

76. Dorrestein, P.; Bumpus, S.; Calderone, C.; Garneau-Tsodikova, S.; Aron, Z.; Straight, P.; Kolter, R.; Walsh, C.; Kelleher, N., Facile detection of acyl and peptidyl intermediates on thiotemplate carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry. *Biochemistry* **2006**, 12756-12766.

77. Meluzzi, D.; Zheng, W.; Hensler, M.; Nizet, V.; Dorrestein, P., Top-down mass spectrometry on low-resolution instruments: Characterization of phosphopantetheinylated carrier domains in polyketide and non-ribosomal biosynthetic pathways. *Bioorg. Med. Chem. Lett.* **2008**, 3107-3111.

78. Dorrestein, P. C.; Blackhall, J.; Straight, P. D.; Fischbach, M. A.; Garneau-Tsodikova, S.; Edwards, D. J.; McLaughlin, S.; Lin, M.; Gerwick, W. H.; Kolter, R.; Walsh, C. T.; Kelleher, N. L., Activity screening of carrier domains within nonribosomal peptide synthetases using complex substrate mixtures and large molecule mass spectrometry. *Biochemistry* **2006,** *45*, 1537-1546.

79. Ma, S. M.; Zhan, J.; Xie, X.; Watanabe, K.; Tang, Y.; Zhang, W., Redirecting the cyclization steps of a fungal polyketide synthase. *J. Am. Chem. Soc.* **2008,** *130*, 38-39.

80. Spiteller, D.; Waterman, C.; Spencer, J., A method for trapping intermediates of polyketide biosynthesis with a nonhydrolyzable malonyl-coenzyme A analogue. *Angew. Chem., Int. Ed.* **2005**, 7079-7082.

81. Tosin, M.; Spiteller, D.; Spencer, J., Malonyl carba(dethia)- and malonyl oxa(dethia)-coenzyme A as tools for trapping polyketide intermediates. *ChemBioChem* **2009**, 1714-1723.

82. Tosin, M.; Betancor, L.; Stephens, E.; Li, W.; Spencer, J.; Leadlay, P., Synthetic chain terminators off-load intermediates from a type I polyketide synthase. *ChemBioChem* **2010**, 539-546.

83. Roberts, G. A.; Staunton, J.; Leadlay, P. F., Heterologous expression in Escherichia coli of an intact multienzyme component of the erythromycin-producing polyketide synthase. *Eur. J. Biochem.* **1993,** *214*, 305-311.

84. Gaitatzis, N.; Hans, A.; Muller, R.; Beyer, S., The mtaA gene of the myxothiazol biosynthetic gene cluster from Stigmatella aurantiaca DW4/3-1 encodes a phosphopantetheinyl transferase that activates polyketide synthases and polypeptide synthetases. *J. Biochem.* **2001**, 119-124.

85. Richardson, M. T.; Pohl, N. L.; Kealey, J. T.; Khosla, C., Tolerance and specificity of recombinant 6-methylsalicyclic acid synthase. *Metab. Eng.* **1999,** *1*, 180-187.

86. Valentine, E., Synthetic approaches towards a fluoroacetyl CoA analogue as an inhibitor of FlK. Part III project, 2007.

87. Cheung, K.; Coles, S.; Hursthouse, M.; Johnson, N.; Shoolingin-Jordan, P., The determination of the absolute configurations of diastereomers of (S)camphanoyl 3-hydroxy-5-oxohexanoic acid derivatives by X-ray crystallography. *Angew. Chem., Int. Ed.* **2002,** *41*, 1198-1202.

88. Meier, J. L.; Haushalter, R. W.; Burkart, M. D., A mechanism based protein crosslinker for acyl carrier protein dehydratases. *Bioorg. Med. Chem. Lett.* **2010,** *20*, 4936-4939.

89. Moriguchi, T.; Kezuka, Y.; Nonaka, T.; Ebizuka, Y.; Fujii, I., Hidden function of catalytic domain in 6-methylsalicylic acid synthase for product release. *J. Biol. Chem.* **2010**, 15637-15643.

90. Pasta, S.; Witkowski, A.; Joshi, A. K.; Smith, S., Catalytic residues are shared between two pseudosubunits of the dehydratase domain of the animal fatty acid synthase. *Chem. Biol.* **2007,** *14*, 1377-1385.

91. a) Kimber, M. S.; Martin, F.; Lu, Y. J.; Houston, S.; Vedadi, M.; Dharamsi, A.; Fiebig, K. M.; Schmid, M.; Rock, C. O., The structure of (3R)-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from Pseudomonas aeruginosa. *J. Biol. Chem.* **2004,** *279*, 52593-52602;

b) Leesong, M.; Henderson, B. S.; Gillig, J. R.; Schwab, J. M.; Smith, J. L., Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: Two catalytic activities in one active site. *Structure* **1996,** *4*, 253-264.

92. Smith, S.; Witkowski, A.; Joshi, A. K., Structural and functional organization of the animal fatty acid synthase. *Prog. Lipid Res.* **2003,** *42*, 289-317.

93. Akey, D. L.; Razelun, J. R.; Tehranisa, J.; Sherman, D. H.; Gerwick, W. H.; Smith, J. L., Crystal Structures of Dehydratase Domains from the Curacin Polyketide Biosynthetic Pathway. *Structure* **2010,** *18*, 94-105.

94. Keatinge-Clay, A., Crystal Structure of the Erythromycin Polyketide Synthase Dehydratase. *J. Mol. Biol.* **2008,** *384*, 941-953.

95. Schwab, J. M.; Habib, A.; Klassen, J. B., A thorough study of the stereochemical consequences of the hydration dehydration reaction catalyzed by beta-hydroxydecanoyl thioester dehydrase. *J. Am. Chem. Soc.* **1986,** *108*, 5304-5308.

96. Palaniappan, N.; Alhamadsheh, M. M.; Reynolds, K. A., cis-Delta(2,3)-double bond of phoslactomycins is generated by a post-PKS tailoring enzyme. *J. Am. Chem. Soc.* **2008,** *130*, 12236.

97. Reid, R.; Piagentini, M.; Rodriguez, E.; Ashley, G.; Viswanathan, N.; Carney, J.; Santi, D. V.; Hutchinson, C. R.; McDaniel, R., A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. *Biochemistry* **2003,** *42*, 72-79.

98. Tang, L.; Ward, S.; Chung, L.; Carney, J. R.; Li, Y.; Reid, R.; Katz, L., Elucidating the mechanism of cis double bond formation in epothilone biosynthesis. *J. Am. Chem. Soc.* **2004,** *126*, 46-47.

99. Schwab, J. M.; Klassen, J. B., Steric course of the allylic rearrangement catalyzed by β-hydroxydecanoylthioester dehydrase. Mechanistic Implications. *J. Am. Chem. Soc.* **1984,** *106*, 7217-7227.

100. Harris, C.; Harris, T., Synthesis of 5-oxohexenoic acid. *J. Org. Chem.* **1971,** *36*, 2181-2182.

101. Taster, G., Synthetic approaches towards potential intermediates in the 6-methyl salicylic acid synthase biosynthetic pathway. Part III project, 2008.

102. Moriguch, T.; Ebizuka, Y.; Fujii, I., Domain-domain interactions in the iterative type I polyketide synthase ATX from Aspergillus terreus. *ChemBioChem* **2008**, 1207-1212.

103. Moriguchi, T.; Ebizuka, Y.; Fujii, I., Analysis of subunit interactions in the iterative type I polyketide synthase ATX from Aspergillus terreus. *ChemBioChem* **2006**, 1869-1874.

104. Sun, Y. H.; Hahn, F.; Demydchuk, Y.; Chettle, J.; Tosin, M.; Osada, H.; Leadlay, P. F., In vitro reconstruction of tetronate RK-682 biosynthesis. *Nature Chemical Biology* **2010,** *6*, 99-101.

105. Holzbaur, I. E.; Harris, R. C.; Bycroft, M.; Cortes, J.; Bisang, C.; Staunton, J.; Rudd, B. A. M.; Leadlay, P. F., Molecular basis of Celmer's rules: the role of two ketoreductase domains in the control of chirality by the erythromycin modular polyketide synthase. *Chem. Biol.* **1999,** *6*, 189-195.

106. Holzbaur, I.; Ranganathan, A.; Thomas, I.; Kearney, D.; Reather, J.; Rudd, B.; Staunton, J.; Leadlay, P., Molecular basis of Celmer's rules: role of the ketosynthase domain in epimerisation and demonstration that ketoreductase domains can have altered product specificity with unnatural substrates. *Chem. Biol.* **2001**, 329-340.

107. Kao, C. M.; McPherson, M.; McDaniel, R. N.; Fu, H.; Cane, D. E.; Khosla, C., Alcohol stereochemistry in polyketide backbones is controlled by the beta-ketoreductase domains of modular polyketide synthases. *J. Am. Chem. Soc.* **1998,** *120*, 2478-2479.

108. Siskos, A. P.; Baerga-Ortiz, A.; Bali, S.; Stein, V.; Mamdani, H.; Spiteller, D.; Popovic, B.; Spencer, J. B.; Staunton, J.; Weissman, K. J.; Leadlay, P. F., Molecular basis of Celmer's rules: Stereochemistry of catalysis by isolated ketoreductase domains from modular polyketide synthases. *Chem. Biol.* **2005,** *12*, 1145-1153.

109. a) McPherson, M.; Khosla, C.; Cane, D. E., Erythromycin biosynthesis: The beta-ketoreductase domains catalyze the stereospecific transfer of the 4-pro-S hydride of NADPH. *J. Am. Chem. Soc.* **1998,** *120*, 3267-3268;

b) Yin, Y. F.; Gokhale, R.; Khosla, C.; Cane, D. E., Erythromycin biosynthesis. The 4-pro-S hydride of NADPH is utilized for ketoreduction by both module 5 and module 6 of the 6-deoxyerythronolide B synthase. *Bioorg. Med. Chem. Lett.* **2001,** *11*, 1477-1479.

110. Caffrey, P., Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketicle synthases. *ChemBioChem* **2003,** *4*, 654-657.

111. Korman, T. P.; Hill, J. A.; Vu, T. N.; Tsai, S. C., Structural analysis of actinorhodin polyketide ketoreductase: Cofactor binding and substrate specificity. *Biochemistry* **2004,** *43*, 14529-14538.

112. Keatinge-Clay, A. T.; Stroud, R. M., The structure of a ketoreductase determines the organization of the beta-carbon processing enzymes of modular polyketide synthases. *Structure* **2006,** *14*, 737-748.

113. Keatinge-Clay, A. T., A tylosin ketoreductase reveals how chirality is determined in polyketides. *Chem. Biol.* **2007,** *14*, 898-908.

114. Ding, W.; Lei, C.; He, Q.; Zhang, Q.; Bi, Y.; Liu, W., Insights into bacterial 6-methylsalicylic acid synthase and its engineering to orsellinic acid synthase for spirotetronate generation. *Chem. Biol.* **2010**, 495-503.

115. Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G., Clustal W and clustal X version 2.0. *Bioinformatics* **2007,** *23*, 2947-2948.

116. Nazi, I.; Koteva, K.; Wright, G., One-pot chemoenzymatic preparation of coenzyme A analogues. *Anal. Biochem.* **2004**, 100-105.

117. Quadri, L.; Weinreb, P.; Lei, M.; Nakano, M.; Zuber, P.; Walsh, C., Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **1998**, 1585-1595.

118. Lee, K. K. M.; Da Silva, N. A.; Kealey, J. T., Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in Escherichia coli and Saccharomyces cerevisiae. *Anal. Biochem.* **2009,** *394*, 75-80.

119. Geoghegan, K. F.; Dixon, H. B. F.; Rosner, P. J.; Hoth, L. R.; Lanzetti, A. J.; Borzilleri, K. A.; Marr, E. S.; Pezzullo, L. H.; Martin, L. B.; LeMotte, P. K.; McColl, A. S.; Kamath, A. V.; Stroh, J. G., Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in Escherichia coli: The cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* **1999,** *267*, 169-184.

120. Sun, Y.; Hong, H.; Gillies, F.; Spencer, J.; Leadlay, P., Glyceryl-S-acyl carrier protein as an intermediate in the biosynthesis of tetronate antibiotics. *ChemBioChem* **2008**, 150-156.

121. Weissman, K.; Hong, H.; Oliynyk, M.; Siskos, A.; Leadlay, P., Identification of a phosphopantetheinyl transferase for erythromycin biosynthesis in Saccharopolyspora erythraea. *ChemBioChem* **2004**, 116-125.

122. Bornstein, P.; Balian, G., Specific nonenzymatic cleavage of bovine ribonuclease with hydroxylamine. *J. Biol. Chem.* **1970,** *245*, 4854-4856.

123. a) Aparicio, J. F.; Caffrey, P.; Marsden, A. F. A.; Staunton, J.; Leadlay, P. F., Limited proteolysis and active-site studies of the first multienzyme component of the erythromycin-producing polyketide synthase. *J. Biol. Chem.* **1994,** *269*, 8524-8528;

b) Hong, H.; Appleyard, A.; Siskos, A.; Garcia-Bernardo, J.; Staunton, J.; Leadlay, P., Chain initiation on type I modular polyketide synthases revealed by limited proteolysis and ion-trap mass spectrometry. *FEBS J.* **2005,** *272*, 2373-2387.

124. Wattana-Amorn, P.; Williams, C.; Ploskon, E.; Cox, R.; Simpson, T.; Crosby, J.; Crump, M., Solution Structure of an Acyl Carrier Protein Domain from a Fungal Type I Polyketide Synthase. *Biochemistry* **2010,** *49*, 2186-2193.

125. Vincze, T.; Posfai, J.; Roberts, R. J., NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Res.* **2003,** *31*, 3688-3691.

126. Apweiler, R.; Martin, M. J.; O'Donovan, C.; Magrane, M.; Alam-Faruque, Y.; Antunes, R.; Barrell, D.; Bely, B.; Bingley, M.; Binns, D.; Bower, L.; Browne, P.; Chan, W. M.; Dimmer, E.; Eberhardt, R.; Fedotov, A.; Foulger, R.; Garavelli, J.; Huntley, R.; Jacobsen, J.; Kleen, M.; Laiho, K.; Leinonen, R.; Legge, D.; Lin, Q.; Liu, W. D.; Luo, J.; Orchard, S.; Patient, S.; Poggioli, D.; Pruess, M.; Corbett, M.; di Martino, G.; Donnelly, M.; van Rensburg, P.; Bairoch, A.; Bougueleret, L.; Xenarios, I.; Altairac, S.; Auchincloss, A.; Argoud-Puy, G.; Axelsen, K.; Baratin, D.; Blatter, M. C.; Boeckmann, B.; Bolleman, J.; Bollondi, L.; Boutet, E.; Quintaje, S. B.; Breuza, L.; Bridge, A.; deCastro, E.; Ciapina, L.; Coral, D.; Coudert, E.; Cusin, I.; Delbard, G.; Doche, M.; Dornevil, D.; Roggli, P. D.; Duvaud, S.; Estreicher, A.; Famiglietti, L.; Feuermann, M.; Gehant, S.; Farriol-Mathis, N.; Ferro, S.; Gasteiger, E.; Gateau, A.; Gerritsen, V.; Gos, A.; Gruaz-Gumowski, N.; Hinz, U.; Hulo, C.; Hulo, N.; James, J.; Jimenez, S.; Jungo, F.; Kappler, T.; Keller, G.; Lachaize, C.; Lane-Guermonprez, L.; Langendijk-Genevaux, P.; Lara, V.; Lemercier, P.; Lieberherr, D.; Lima, T. D.; Mangold, V.; Martin, X.; Masson, P.; Moinat, M.; Morgat, A.; Mottaz, A.; Paesano, S.; Pedruzzi, I.; Pilbout, S.; Pillet, V.; Poux, S.; Pozzato, M.; Redaschi, N.; Rivoire, C.; Roechert, B.; Schneider, M.; Sigrist, C.; Sonesson, K.; Staehli, S.; Stanley, E.; Stutz, A.; Sundaram, S.; Tognolli, M.; Verbregue, L.; Veuthey, A. L.; Yip, L. N.; Zuletta, L.; Wu, C.; Arighi, C.; Arminski, L.; Barker, W.; Chen, C. M.; Chen, Y. X.; Hu, Z. Z.; Huang, H. Z.; Mazumder, R.; McGarvey, P.; Natale, D. A.; Nchoutmboube, J.; Petrova, N.; Subramanian, N.; Suzek, B. E.; Ugochukwu, U.;

Vasudevan, S.; Vinayaka, C. R.; Yeh, L. S.; Zhang, J.; UniProt, C., The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res.* **2010,** *38*, D142-D148.

127. Bachmann, B. O.; Ravel, J., Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data. *Complex Enzymes in Microbial Natural Product Biosynthesis, Part a: Overview Articles and Peptides* **2009,** *458*, 181-217.

128. Mazzini, F.; Galli, F.; Salvadori, P., Vitamin E metabolites: Synthesis of [D-2]- and [D-3]-gamma-CEHC. *Eur. J. Org. Chem.* **2006**, 5588-5593.

129. Sui, Z.; Furth, P. S.; Devoss, J. J., Ring expansion of halo dithiolanes and dithianes - a facile synthesis of medium-ring dithiacycloalkenes. *J. Org. Chem.* **1992,** *57*, 6658-6662.

130. Kwan, D. H., The stereochemistry of reduction in modular polyketide synthases. PhD Thesis, 2009.

131. Bradford, M. M., Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Anal. Biochem.* **1976,** *72*, 248-254.

132. Rampe, D.; Hake, P. W.; Borretzen, B.; Holm, K. H.; Skattebol, L., Deuterated analogs Of verapamil and nifedipine - synthesis and biological-activity. *Eur. J. Med. Chem.* **1993,** *28*, 259-263.

133. Cohen, S. G.; Khedoui, E., Requirements for Stereospecificity in Hydrolysis by α-Chymotrypsin. IV. The Hydroxyl Substituent. Absolute Configurations. *J. Am. Chem. Soc.* **1961,** *83*, 4228-4233.

134. Baader, E.; Bartmann, W.; Beck, G.; Bergmann, A.; Fehlhaber, H. W.; Jendralla, H.; Kesseler, K.; Saric, R.; Schussler, H.; Teetz, V.; Weber, M.; Wess, G.,

Syntheses of 4(R)-silyloxy-6(S)-iodomethyl-tetrahydropyran-2-one and its enantiomer, building-blocks for HMG-CoA reductase inhibitors. *Tetrahedron Lett.* **1988,** *29*, 2563-2566.

135. Santaniello, E.; Chiari, M.; Ferraboschi, P.; Trave, S., Enhanced and reversed enantioselectivity of enzymatic-hydrolysis by simple substrate modifications - the case of 3-hydroxy glutarate diesters. *J. Org. Chem.* **1988,** *53*, 1567-1569.

136. Rosen, T.; Watanabe, M.; Heathcock, C. H., A convenient assay for the optical purity of monomethyl 3-hydroxypentanedioate. *J. Org. Chem.* **1984,** *49*, 3657-3659.

137. Roy, R.; Rey, A. W., Chemoenzymatic synthesis of a C5-chiral building block - a substrate modification approach. *Tetrahedron Lett.* **1987,** *28*, 4935-4938.

138. Lion, C. J.; Matthews, C. S.; Stevens, M. F. G.; Westwell, A. D., Synthesis, antitumor evaluation, and apoptosis-inducing activity of hydroxylated (E)-stilbenes. *J. Med. Chem.* **2005,** *48*, 1292-1295.

139. Nicolaou, K. C.; Boddy, C. N. C.; Brase, S.; Winssinger, N., Chemistry, biology, and medicine of the glycopeptide antibiotics. *Angew. Chem., Int. Ed.* **1999,** *38*, 2097-2152.

140. Li, T. L.; Huang, F. L.; Haydock, S. F.; Mironenko, T.; Leadlay, P. F.; Spencer, J. B., Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: Characterization of two glycosyltransferases and the key acyltransferase. *Chem. Biol.* **2004,** *11*, 107-119.

141. Sosio, M.; Kloosterman, H.; Bianchi, A.; de Vreugd, P.; Dijkhuizen, L.; Donadio, S., Organization of the teicoplanin gene cluster in Actinoplanes teichomyceticus. *Microbiology-Sgm* **2004,** *150*, 95-102.

142. Sandercock, A. M.; Charles, E. H.; Scaife, W.; Kirkpatrick, P. N.; O'Brien, S. W.; Papageorgiou, E. A.; Spencer, J. B.; Williams, D. H., Biosynthesis of the di-meta-hydroxyphenylglycine constituent of the vancomycin-group antibiotic chloroeremomycin. *Chem. Commun.* **2001,** 1252-1253.

143. Stinchi, S.; Carrano, L.; Lazzarini, A.; Feroggio, M.; Grigoletto, A.; Sosio, M.; Donadio, S., A derivative of the glycopeptide A40926 produced by inactivation of the beta-hydroxylase gene in Nonomuraea sp. ATCC39727. *FEMS Microbiol. Lett.* **2006,** *256*, 229-235.

144. Puk, O.; Bischoff, D.; Kittel, C.; Pelzer, S.; Weist, S.; Stegmann, E.; Sussmuth, R. D.; Wohlleben, W., Biosynthesis of chloro-beta-hydroxytyrosine, a nonproteinogenic amino acid of the peptidic backbone of glycopeptide antibiotics. *J. Bacteriol.* **2004,** *186*, 6093-6100.

145. Bischoff, D.; Pelzer, S.; Bister, B.; Nicholson, G. J.; Stockert, S.; Schirle, M.; Wohlleben, W.; Jung, G.; Sussmuth, R. D., The biosynthesis of vancomycin-type glycopeptide antibiotics - The order of the cyclization steps. *Angew. Chem., Int. Ed.* **2001,** *40*, 4688-4691.

146. Bischoff, D.; Bister, B.; Bertazzo, M.; Pfeifer, V.; Stegmann, E.; Nicholson, G. J.; Keller, S.; Pelzer, S.; Wohlleben, W.; Sussmuth, R. D., The biosynthesis of vancomycin-type glycopeptide antibiotics - A model for oxidative side-chain cross-linking by oxygenases coupled to the action of peptide synthetases. *ChemBioChem* **2005,** *6*, 267-272.

147. Woithe, K.; Geib, N.; Zerbe, K.; Li, D. B.; Heck, M.; Fournier-Rousset, S.; Meyer, O.; Vitali, F.; Matoba, N.; Abou-Hadeed, K.; Robinson, J. A., Oxidative phenol coupling reactions catalyzed by OxyB: A cytochrome p450 from the vancomycin producing organism. Implications for vancomycin biosynthesis. *J. Am. Chem. Soc.* **2007,** *129*, 6887-6895.

148. Truman, A. W.; Robinson, L.; Spencer, J. B., Identification of a deacetylase involved in the maturation of teicoplanin. *ChemBioChem* **2006,** *7*, 1670-1675.

149. Allen, N. E.; Nicas, T. I., Mechanism of action of oritavancin and related glycopeptide antibiotics. *FEMS Microbiol. Rev.* **2003,** *26*, 511-532.

150. Walsh, C. T.; Fisher, S. L.; Park, I. S.; Prahalad, M.; Wu, Z., Bacterial resistance to vancomycin: Five genes and one missing hydrogen bond tell the story. *Chem. Biol.* **1996,** *3*, 21-28.

151. Pootoolal, J.; Neu, J.; Wright, G. D., Glycopeptide antibiotic resistance. *Ann. Rev. Pharmacol. Toxicol.* **2002,** *42*, 381-408.

152. Mackay, J. P.; Gerhard, U.; Beauregard, D. A.; Westwell, M. S.; Searle, M. S.; Williams, D. H., Glycopeptide antibiotic-activity and the possible role of dimerization - a model for biological signaling. *J. Am. Chem. Soc.* **1994,** *116*, 4581-4590.

153. Mackay, J. P.; Gerhard, U.; Beauregard, D. A.; Maplestone, R. A.; Williams, D. H., Dissection of the contributions toward dimerization of glycopeptide antibiotics. *J. Am. Chem. Soc.* **1994,** *116*, 4573-4580.

154. Williams, D. H.; Maguire, A. J.; Tsuzuki, W.; Westwell, M. S., An analysis of the origins of a cooperative binding energy of dimerization. *Science* **1998,** *280*, 711-714.

155. Noble, W. C.; Virani, Z.; Cree, R. G., Co-transfer of vancomycin and other resistance genes from Enterococcus faecalis NCTC 12201 to Staphylococcus aureus. *FEMS Microbiol. Lett.* **1992,** *72*, 195-198.

156. Hiramatsu, K.; Hanaki, H.; Ino, T.; Yabuta, K.; Oguri, T.; Tenover, F. C., Methicillin-resistant Staphylococcus aureus clinical strain with reduced vancomycin susceptibility. *J. Antimicrob. Chemother.* **1997,** *40*, 135-136.

157. Hiramatsu, K., Vancomycin resistance in Staphylococci. *Drug Resistance Updates* **1998,** *1*, 135-150.

158. Van Bambeke, F., Glycopeptides in clinical development: pharmacological profile and clinical perspectives. *Curr. Opin. Pharmacol.* **2004,** *4*, 471-478.

159. Arthur, M.; Depardieu, F.; Molinas, C.; Reynolds, P.; Courvalin, P., The vanZ Gene of TN1546 from enterococcus-faecium BM4147 confers resistance to teicoplanin. *Gene* **1995,** *154*, 87-92.

160. Crowley, B. M.; Boger, D. L., Total synthesis and evaluation of psi CH2NH Tpg(4) vancomycin aglycon: Reengineering vancomycin for dual D-Ala-D-Ala and D-Ala-D-Lac binding. *J. Am. Chem. Soc.* **2006,** *128*, 2885-2892.

161. Li, D. B.; Robinson, J. A., An improved solid-phase methodology for the synthesis of putative hexa- and heptapeptide intermediates in vancomycin biosynthesis. *Org. Biomol. Chem.* **2005,** *3*, 1233-1239.

162. Lange, U. E. W.; Schafer, B.; Baucke, D.; Buschmann, E.; Mack, H., A new mild method for the synthesis of amidines. *Tetrahedron Lett.* **1999,** *40*, 7067-7070.

163. Mori, N.; Togo, H., Direct oxidative conversion of primary alcohols to nitriles using molecular iodine in ammonia water. *Synlett* **2005**, 1964-1964.