STUDIES ON THE LATE-STAGE BIOSYNTHESIS OF THE ANTIBIOTIC MUPIROCIN

by

JACK ANDREW CONNOLLY

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

School of Biosciences
College of Life and Environmental Sciences
University of Birmingham
December 2017

Abstract

There is a clear and well-established need for new antibiotics. The biosynthesis of the antibiotic mupirocin is an important model *trans*-AT polyketide synthase pathway. A key tailoring step is the removal of a hydroxyl from the intermediate pseudomonic acid B (PA-B), which generates active pseudomonic acid A (PA-A). This is proposed to occur on discrete proteins after release from enzyme MmpB.

In this work, a systematic screen for genes required for this key step was developed, which implicated ten genes. On a multi-gene expression plasmid, these ten genes encoded all functions necessary for conversion. This re-identified seven genes from the existing proposed model. A novel requirement for MupM, MupN, and the DNA sequence but not protein function of MupL was uncovered. This represents the first identification of all genes required for this surprisingly convoluted pathway.

With the aim of re-engineering the location of these tailoring steps, domains from thiomarinol enzyme TmpB were inserted into MmpB. Low-level conversion to PA-A was achieved using thiomarinol tailoring enzymes. The TmpB KS^0 domain was used to load the normally free-standing MacpE, and high-level conversion was achieved on the hybrid enzyme. This success opens the future path to further MmpB modifications to generate mupirocin derivatives.

Acknowledgements

Firstly, I would like to thank Chris for all his encouragement, teaching and trust over the last five years, which I think have been instrumental in developing me into an effective, independent researcher. Thank you to my co-supervisor, Peter, for bringing a fresh perspective to our meetings that often spotted new avenues to investigate. Thank you to the MIBTP course for the enjoyable and useful lessons, and for providing a cohort to share this experience with, and to the BBRSC for funding.

Thank you to everyone in S101 for the companionship and for being great colleagues throughout. In particular, thank you to my soon-to-be wife, Yusra, who I will not embarrass here. I would also like to thank Amber for her bright outlook throughout all her hard work on this project, achieving great results, and for being a pleasure to work with. A special thank you to Elton for all your help, it will be very strange to work in a lab without you on the next bay. Thank you to my fellow S101 PhD students, Alessandro, Rabi, and Hadi, for your friendship throughout; and to the post-docs, Tony, Jo, Anand, and Claire, for always being friendly and helpful throughout, and for many great discussions.

Contents

List of Figures

List of Tables

List of Abbreviations

6-dEB	6-deoxyerythronolide B
9-HN	9-hydroxynonanoic acid
AA	Amino acid
ACP	Acyl carrier protein
Amp^R	Ampicillin resistant
AT	Acyl transferase
ATP	Adenosine triphosphate
Bp	Base pairs (DNA)
CDA	Calcium dependent antibiotic
CDS	Coding DNA sequence
CFU	Colony forming unit
CoA	Coenzyme A
dATP	Deoxyadenosine triphosphate
dNTP	Deoxynucleotide triphosphate
DH	Dehydratase
DNA	Deoxyribonucleic acid
DOS	Diversity orientated synthesis
EDTA	Ethylenediaminetetraacetic acid
ER	Enoyl reductase
EtOH	Ethanol
FAS	Fatty acid synthase
HPLC	High performance liquid chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IleRS	Isoleucyl tRNA synthetase
Kan^R	Kanamycin resistant
KR	Ketoreductase
KS	Ketosynthase
KS^0	Non-elongating ketosynthase
LB	Lysogeny broth
LCMS	Liquid chromotography mass spectroscopy
MAT	Malonyl-acetyl transferase
MIC	Minimum inhibitory concentration
MRSA	Methicillin-resistant *Staphylococcus aureus*
MT	Methyl transferase
NAD	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
NHS	National Health Service

NMR	Nuclear Magnetic Resonance
NRP	Non-ribosomal peptide
NRPS	Non-ribosomal peptide synthetase
OR	Oxidoreductase
PA-A	Pseudomonic acid A
PA-B	Pseudomonic acid B
PA-C	Pseudomonic acid C
RBS	Ribosome binding site
PCP	Peptidyl carrier protein
PCR	Polymerase chain reaction
PKS	Polyketide synthase
PP	Phosphopantetheine
PPTase	Phosphopantetheinyl transferase
RFLP	Restriction fragment length polymorphism
RT-PCR	Reverse transcriptase PCR
SDR	Short-chain reductase
SDW	Sterile distilled water
SNAC	N-acetyl-cysteamine
SOB	Super optimal broth
SOC	Super optimal broth with catabolite repression
SSM	Secondary stage media
TAE	Tris, acetate, EDTA buffer
TE	Thioesterase
TraDIS	Transposon directed insertion sequencing
TTC	Triphenyl tetrazolium chloride
UV	Ultraviolet
WT	Wild type

Chapter 1

INTRODUCTION

1.1 The importance of natural products

1.1.1 Natural products have had a key historical role in drug discovery

Natural products have had a key role in both the discovery and production of new medicines throughout history. Drugs have been developed from natural products in a wide variety of compound classes and functions. Different natural product biosynthesis mechanisms combine to yield a large diversity of complex structures. The diverse structures of the examples described in this section are displayed in Figure 1.1.

A range of antibiotics have been identified and derived from natural products, such as the iconic penicillin, discovered in *Penicillium* fungi by Fleming (1929), or streptomycin discovered by Waksman (1953). Both of these discoveries were of huge importance to the scientific and medical communities, and as such both won the Nobel Prize in Physiology or Medicine, in 1945 and 1952 respectively. The first of the aminoglycoside class discovered, streptomycin, is produced by *Streptomyces griseus*. The *Streptomyces* genus has proved particularly fruitful in antibiotic discovery, with more than 4500 antibiotics discovered in *Streptomyces* bacteria (Challinor and Bode, 2015).

Natural products are not limited to antibiotics or microorganisms. For example, the household drug aspirin is named after the plant genus *Spiraea*, the source of salicylic acid from which it is derived; marketed by Bayer AG in 1899 (Sneader, 1997). Another example of high profile natural products are the cholesterol lowering statins, such as lovastatin or simvastatin (Alberts et al., 1980; Mol et al., 1988). Both were identified from the fungus *Aspergillus terreus*, and have become hugely successful natural product derived drugs. Simvastatin was the most commonly prescribed NHS drug in England in 2016, at 30.6 million prescriptions, not including combination treatments (NHS, 2016).

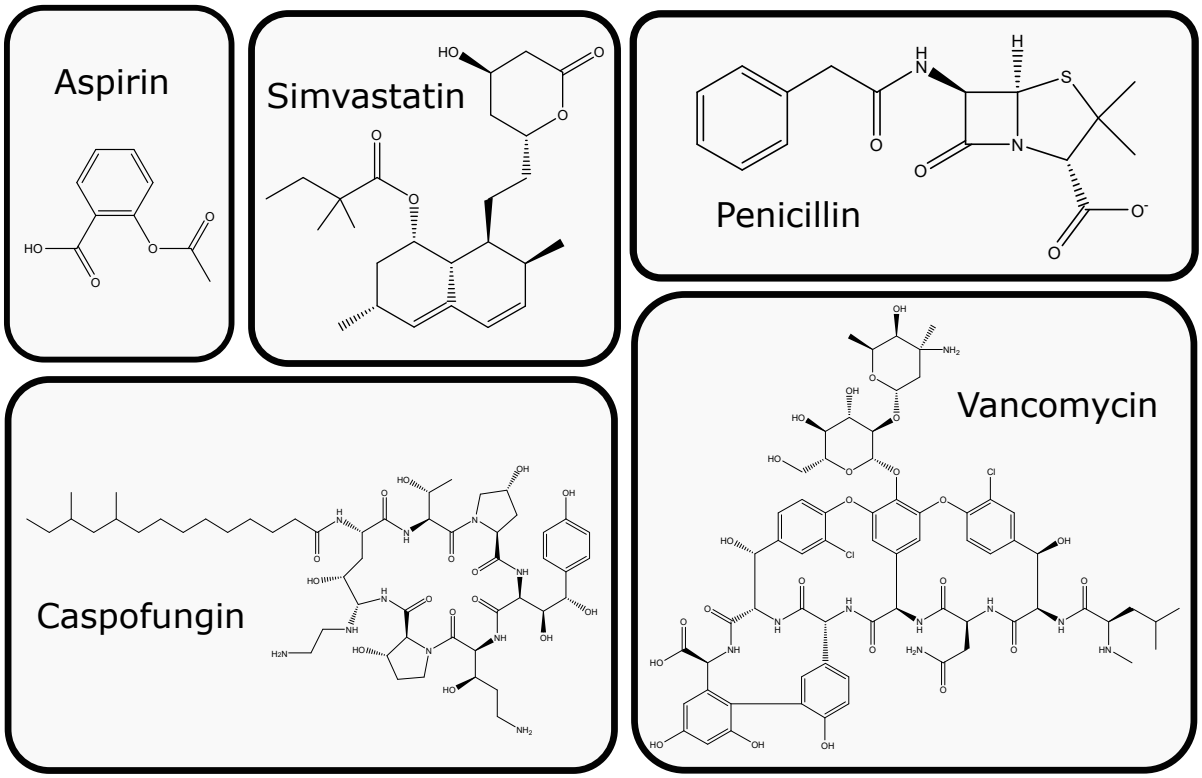


Figure 1.1: Examples of the diversity of natural product structures

Further to the relatively modern examples given so far, there are examples of historical use of natural products for medicinal purposes, such as the anti-parasitic artemisin. Artemisin is produced by the plant *Artemisia annua*, and the herb has been used historically in traditional Chinese medicine for the treatment of malaria. The compound Artemisin was discovered through a screen of more than 2000 Chinese herb preparations, and together with semi-synthetic derivatives, is in current medicinal use (Tu, 2011). Youyou Tu was awarded a joint Nobel Prize in Physiology or Medicine in 2015 for its discovery.

There are also natural product derived anti-cancer and anti-fungal drugs, such as epothilones and caspofungin (Cragg and Newman, 2013). Secondary metabolites, isolated from a variety of both prokaryotes and eukaryotes, have had a key role in the historical development of a wide range of medical drugs.

1.1.2 The urgent need for new antibiotics

There has been a seemingly inexorable rise in antibiotic resistance over the last century, with a positive correlation to antibiotic consumption (Bell et al., 2014). In the classical example of penicillin, resistance was first observed in 1940, before its clinical use (Abraham and Chain, 1940). This is understandable from an evolutionary perspective, as there would be selective pressure on the development of resistance for microbes growing in the same environment as an antibiotic producer. Further to this, microbes that produce antibiotics need some method to avoid action against self, which could provide a source of resistance genes for horizontal transfer. As clinical use of penicillin and its derivatives increased, resistance became more widespread; for example 40% of *Streptococcus pneumoniae* infections in Spain resistant by 1989 (Baquero et al., 1991).

Methicillin, a semi-synthetic derivative of penicillin, was developed by Beechams in 1959. Whilst there was some expectation that its synthetic nature would limit the development of resistance, resistant *Staphyloccus* infections were observed just two years later (Jevons, 1961).

The case study of vancomycin demonstrates that resistance can and will develop against antibiotics with alternative targets. Discovered in 1952, vancomycin was targeted at treating β-lactam resistant *Staphlococcus* infections, although its use was initially limited due to toxicity and more effective alternatives (Geraci et al., 1956; Moellering, 2012). Vancomycin works by binding N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG) of the gram positive cell wall, preventing cross-linking of peptidoglycan (Barna and Williams, 1984). Following an increase in clinical use in the 1980s, vancomycin resistance was observed in *Enterococcus* infections by 1986 (Leclercq et al., 1988). In 2002, highly vancomycin resistant *Staphylococcus aureus* was clinically isolated, carrying resistance gene, *vanA*, presumably aquired by horizontal gene transfer, which encodes an altered vancomycin cell wall binding site (Tenover, 2006).

Table 1.1: Example mechanisms of antibiotic resistance

Mechanism	Example	Reference
Destruction of antibiotic	Hydrolysis of the β-lactam ring by β-lactamases	(R.P.Ambler, 1980)
Modification of antibiotic	Acetylation of aminoglycosides preventing target binding by steric hindrance	(Shaw et al., 1993)
Prevention of target binding	Prevention of vancomycin binding by *vanA*	(Bugg et al., 1991)
Mutation of target	Streptomycin resistance through point mutation of the 16S rRNA sequence	(Finken et al., 1993)
Increased efflux of antibiotic	Tetracycline resistance through efflux pump encoded by *tetA*	(Speer et al., 1992)
Reduced uptake of anibiotic	Reduced aminoglycoside and β-lactam uptake in *Serratia* strain deficient in porins	(Goldstein et al., 1983)

Under the selective pressure of antibiotic use, resistance can arise from a variety of different mechanisms, some examples of which are summarised in Table 1.1 (Blair et al., 2015). The unyielding advance in antibiotic resistance has led to an urgent need for new effective compounds. Large pharmaceutical companies have spent large amounts on drug development; for example GlaxoSmithKline spent £2.3 billion on the search for new pharmaceuticals in 2015 (GlaxoSmithKline, 2016). However, there are policy issues that limit the attractiveness of investment in new antibiotic development by big pharmceutical companies (Piddock, 2012; Sinha and Kesselheim, 2016; Sciarretta et al., 2016). Antibiotics have a lifespan limited by resistance, and are typically prescribed over short term periods, which yields less revenue than long-term drugs such as statins. With the looming threat of resistance, effective new compounds can be shelved by national healthcare providers to provide a last line of defence. Unfortunately, this can allow patents to lapse before significant revenue is achieved by the pharmaceutical company that developed it, which makes antibiotic

drug discovery an unattractive financial prospect. Policy changes are needed to remedy this; some have already been made such as the Generating Antibiotics Incentives Now act in the USA, which guarantees market exclusivity for new antibiotics for 5 further years after the patent has lapsed (Bax and Green, 2014).

1.1.3 Antibiotic development strategies

The strategies used to develop antibiotics has evolved over the last century, largely out of necessity due to antibiotic resistance . Broadly, in the early 20th Century, unmodified natural products such as penicillin proved effective drugs. In the late 20th Century, medicinal chemistry took over, with the development of chemically modified natural products that were often broadly effective natural products (Brown and Wright, 2016). Worryingly, in recent years there has been little discovery of truly novel antibiotics that are as broadly effective as the early generations. Therefore, consideration of the avenues available for antibiotic discovery is of vital importance.

1.1.3.1 Synthetic compound generation

There are historical examples of outright fully synthetic antibiotics, such as nalidixic acid, which was used in the 1960s and 1970s to treat urinary tract infections (Emmerson and Jones, 2003). However, the largest impact chemical synthesis has had on antibiotic development is through the semi-synthesis of derivatives of existing compounds, which has improved the qualities and extended the life of existing classes of antibiotics.

Small synthetic compound library screening

In the late 1990s, there was a push towards the development of small compound libraries through combinatorial chemistry termed combichem. Such libraries were then tested against specific conditions, typically testing for the MIC against a pathogen (Trias, 2001). Whilst results were initially optimistic, in the long term combichem

has failed to have a significant impact on drug development (Lowe, 2014). This is largely as the core chemical scaffolds of the libraries were quite limited, as large natural product-like compounds typically take a large number of synthesis steps. For example, total synthesis of erythromycin A was achieved in 52 steps with less than 0.01% yield (Woodward et al., 1981).

Synthetic library generation has since evolved towards Diversity Orientated Synthesis (DOS), with a focus on generation of a wider range of scaffolds. DOS libraries have not yet achieved wide-scale diversity equivalent to that seen in natural products (Galloway et al., 2009), although efforts are ongoing, such as the syntheses of erythromycin-like macrocycles (Ciardiello et al., 2016; Seiple et al., 2016). A similar approach has been attempted using peptides, with libraries of diverse peptides yielding some hits for antimicrobial activity against *Escherichia coli* (Hilpert et al., 2005). Development of libraries of synthetic compounds, which mimic the diversity of structures found in natural products, remains an attractive goal for future drug development.

Semi-synthetic modification of existing structures

The synthetic modification of existing natural products has historically been hugely successful, yielding new generations of a variety of different clinically important antibiotics. An early example is the semi-synthesis of tetracycline from chlorotetracycline discovered at Pfizer (Wright et al., 2014). Synthetic modification of existing scaffolds has produced new generations of different antibiotic classes, with various potentially improved characteristics. Recently, synthetic modifications to vancomycin have been reported to increase potency 200-fold against resistant pathogens (Okano et al., 2017).

A fully-synthetic example is the quinolone antibiotics, where successive synthetic modifications have been made to yield multiple generations of antibiotics (Figure 1.2). Starting from the accidental generation of nalidixic acid in 1962 (Lesher et al., 1962),

newer generations have some improved characteristics. For example, 3rd generation quinolones such as levoflaxin are active against *Streptococci*, and the 4th generation have an additional mode of action (Andriole, 2005; Walsh and Wencewicz, 2014).

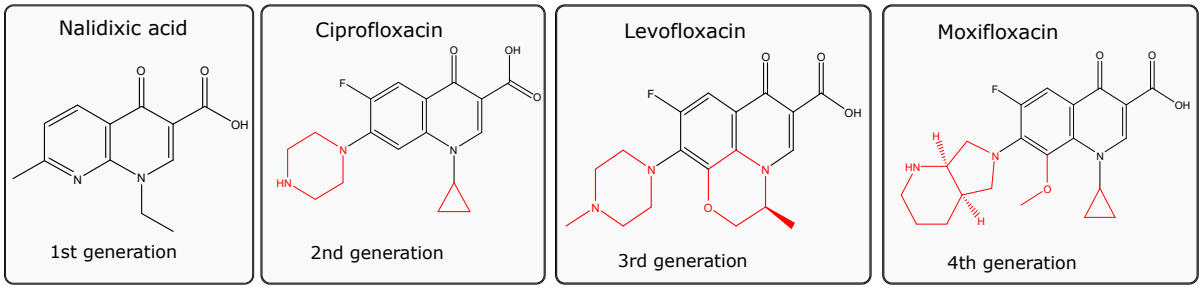


Figure 1.2: The synthetic generations of quinolone antibiotics.
Each generation of the quinolone antibiotics has the core structure (in black). The modifications (in red), have furthered the effectiveness of these antibiotics.

1.1.3.2 Identification of new antibiotic scaffolds

Whilst semi-synthetic approaches have been successful in further developing antibiotics, there is a clear need for new antimicrobial scaffolds, which would essentially open new avenues of modification by medicinal chemistry.

Screening new species

Historically, screening of different species for new antimicrobial compounds has yielded a variety of antibiotics, such as penicillin, erythromycin and streptomycin. However, identification of new antibiotics has faltered as the most accessible, easily culturable, antibiotic producers have been identified. There are a large proportion of organisms that are not conventially culturable in the lab, with estimates at approximately 1% or less of bacteria are lab-culturable (Bérdy, 2012; Clardy et al., 2006; Singh and Pelaez, 2008). Screening still has the potential to provide new scaffolds, but a more modern approach needs to be considered.

An example is the discovery of the non ribosomal peptide synthase (NRPS) derived antibiotic teixobactin, representative of a new class of antibiotic (Ling et al.,

2015). Teixobactin is produced by the previously unidentified gram negative, *Eleftheria terrae*, which was cultured using a modern technology called iChip. In this culturing method, a soil sample is diluted and loaded to approximately one bacterium per channel of the iChip, which is inserted back into the native soil. The device has two semi-permeable membranes, which in effect allow the colony to grow separate from other microbes, yet in its native environment. The screening of previously unculturable bacteria through techniques such as iChip, opens exciting possibilities of the identification of novel natural products.

The authors infer that due to its multiple modes of action, resistance to teixobactin and its future derivatives is unlikely to develop (Ling et al., 2015). Historically, similar claims have proven false, and whilst it seems plausible resistance may develop more slowly, it seems likely to arise eventually (Von Nussbaum and Süssmuth, 2015). Logically, there will have been evolutionary selective pressure on the development of resistance in microorganisms growing in proximity to the teixobactin producer, *E. terrae* (Piddock, 2015). The gram negative *E. terrae* does not have a teixobactin resistance gene, as the compound is only effective again gram positives, a limitation on its potential use as a clinical antibiotic. Therefore, there is not a host resistance gene that other organisms could aquire by horizontal gene transfer (Wright, 2015).

An alternative approach to finding new species, is to look in more niche environments. A recent example is the discovery of a human nose commensal, *Staphylococcus lugdunensis*, which produces the antibiotic lugdunin (Zipperer et al., 2016). This NRPS-derived antibiotic is able to prevent the colonisation of opportunistic pathogen *S. aureus*, with the intriguing future possibility of preventative treatment with the commensal (Lewis and Strandwitz, 2016).

Genomics based approaches

With the advent of successive high-throughput sequencing technologies, the cost of whole genome sequencing has fallen, and there has been an increase in the

volume of genomic data available. Bioinformatic technology has advanced, with servers and algorithms such as antiSMASH facilitating the quick identification of secondary metabolite genome clusters from these data (Medema et al., 2011). After these genomics techniques identify potential silent biosynthetic clusters, molecular approaches can then be undertaken with the aim of turning on the pathway, and identifying new secondary metabolites (Zarins-Tutt et al., 2016).

One approach is to clone the identified biosynthetic clusters to heterologous hosts. An example of this was performed in gram negative *Photorhabdus luminescens*, where 10 unknown biosynthetic gene clusters were cloned to plasmids and mobilised to *E. coli*. This resulted in the discovery of several related new compounds, named luminmycins (Bian et al., 2012).

An alternative is to modify the regulation of gene expression to turn on silent clusters. One successful example of this was performed in the model fungus *Aspergillus nidulans* (Bergmann et al., 2010). A silent NRPS cluster was identified in the genome, flanked by gene encoding a putative regulator *scpR*. Over-expression of this regulator from a plasmid activated expression of a second polyketide synthase (PKS) gene cluster, producing the compound asperfuranone. Similar experiments have been sucessful in turning on silent clusters in *Streptomyces* (Laureti et al., 2011).

Biosynthetic engineering

Biosynthetic engineering is the *in vivo* modification of biosynthesis pathways to yield different secondary metabolites. Whilst knowledge of the mechanisms underpinning such pathways is required, engineering experiments also further that same knowledge. Indeed, an important advantage to these approaches is that insights gained in one pathway can be used to further the redesign of others. In cases such as mupirocin (El-Sayed et al., 2003), where the natural product is industrially produced by fermentation of a producing microbe, re-engineered pathways could have the advantage of being economically viable. Assuming a suitable yield is produced,

the modified compounds could be industrially purified from fermentation broths in a similar fashion to the parent compound.

Genetic engineering of PKS and NRPS biosynthetic pathways, through swapping or mutagenesis of domains, has already yielded new derivative metabolites in a variety of systems (Dunn and Khosla, 2013; Weissman, 2016). Examples include the generation of a library of erythromycin analogs through feeding of fatty acids to a strain containing a hybrid avermectin-erythromycin biosynthesis cluster (Pacey et al., 1998), and the alteration of stereochemistry by KR swaps (Annaval et al., 2015). New epothilone analogs have been generated by engineering of the growth media in a heterologous host (Frykman et al., 2002).

Another example of such an approach was performed on anti-fungal amphotericin B, produced by a PKS pathway in the soil bacterium *Streptomyces nodosus* (Caffrey et al., 2001). When ketoreductase (KR) domain 16 was knocked out, a new amphotericin B derivative also with anti-fungal activity was produced (Power et al., 2008).

A further example involved daptomycin biosynthesis, a lipopeptide antibiotic in clinical use, developed by Cubist Pharmaceuticals. Daptomycin consists of a NRPS produced peptide linked to a lipid and the biosynthesis cluster has been identified in *Steptomyces roseosporus* (Miao et al., 2005). New daptomycin derivatives were isolated following the expression of heterologous NRPS genes in a *dptD* (NRPS) knockout (Table 1.2) (Miao et al., 2006). The heterologous genes expressed were *lptD* from A54145 biosynthesis and *cdaPS3* from calcium dependent antibiotic (CDA) biosynthe-

Table 1.2: Re-engineering daptomycin biosynthesis through expression of heterologous NRPS genes in *S. roseosporus* Δ*dptD* (Miao et al., 2006)

Amino acid position 12	Amino acid position 13	NRPS gene	Gene source
MeGlu	Trp	*dptD*	Daptomycin
Glu	Trp	*cdaPS3*	CDA
Glu	Ile/Val	*lptD*	A54145

sis, both are similar lipopeptide antibiotics produced by *Streptomyces* species. The new derivative compounds demonstrated lower antimicrobial activity against *S. aureus* and *Enterococcus faecalis*. Nevertheless, this remains an exciting demonstration that new natural product derivatives can be generated from biosynthetically engineered hybrid pathways.

1.2 The biosynthesis of natural products

1.2.1 Fatty acid biosynthesis

Describing fatty acid biosynthesis is a sensible prelude to polyketide biosynthesis, as they share similar chemistries and enzymology. However, fatty acid biosynthesis lacks the variety of extender units and non-essential modules seen in polyketide biosynthesis.

1.2.1.1 Enzymology of fatty acid biosynthesis

The fatty acid biosynthesis cycle is summarised in Figure 1.3. Fatty acid biosynthesis begins with the loading of an acetate starter unit from acetyl-CoA onto the ketosynthase (KS) via an acyl carrier protein (ACP), catalysed by a malonyl-acetyl transferase (MAT). This acetate is attached as a thiol to a conserved cysteine on the KS. The ACP is post-translationally modified with a prosthetic flexible phosphopantetheine (PP) arm from CoA, which is held on a conserved serine residue. The PP arm is long and flexible, which aids the ACP in the delivery of intermediates to the various domains. A malonate extender unit is loaded from malonyl-CoA onto the PP arm of the ACP as a thiol, catalysed by the MAT. The ketosynthase catalyses a condensation between the starter and extender unit, adding two carbons with a β-keto group to the starter unit, and leaving the chain on the ACP.

The added β-keto group is then reduced to a hydroxyl catalysed by a ketoreductase (KR) domain, dehydrated by a dehydratase (DH), and further reduced by an

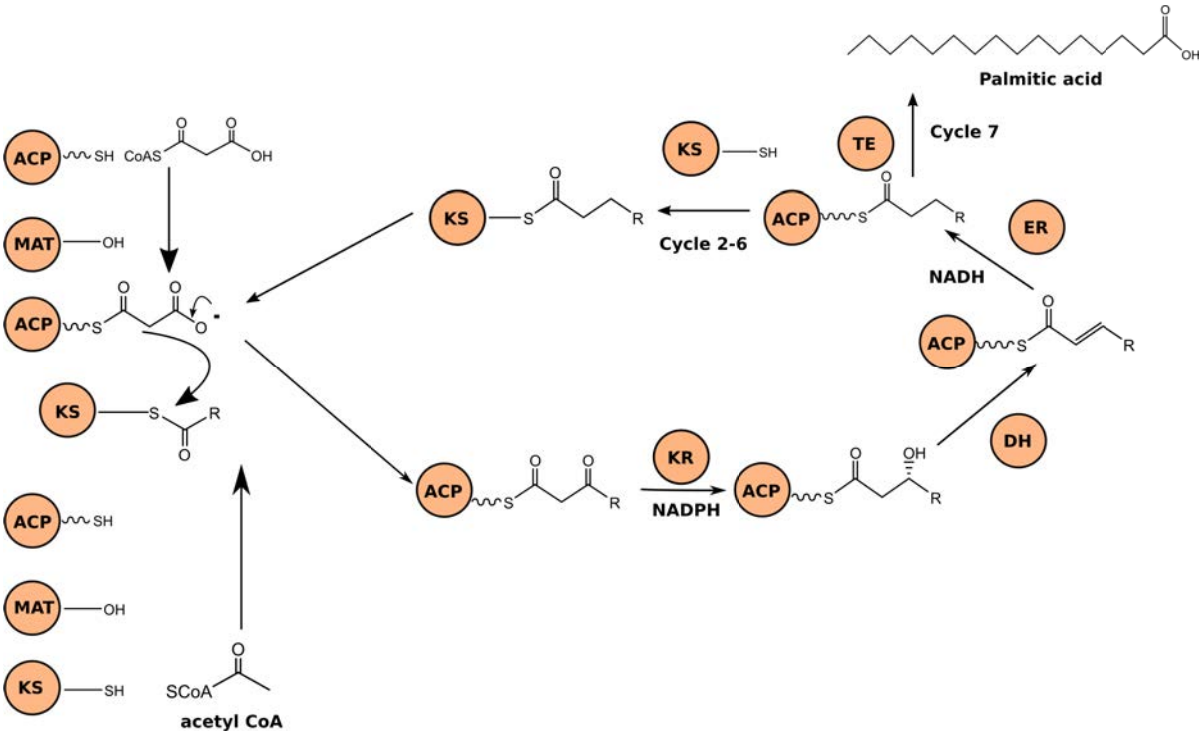


Figure 1.3: The fatty acid biosynthetic pathway.
After 7 cycles of addition and reduction of a β-keto group, the palmitic acid (16 carbons) is released, catalysed by the TE domain. Other length fatty acids are produced by variation in the number of cycles, for example, 8 cycles would generate stearic acid (18 carbons).
Initially, R=CH_3, but it is then increased by two carbons for each biosynthetic cycle.

enoyl reductase (ER). This results in the fully reduced, saturated extension of the fatty acid by two carbons.

The growing chain is transferred to the KS, and a new malonate extender unit loaded onto the ACP by the MAT. Again the KS catalyses a condensation, extending the chain by two carbons, and the KR, DH and ER domains reduce the added β-keto group. After successive cycles, the thioester is cleaved by the thioesterase (TE), releasing the fatty acid product (Staunton and Weissman, 2001; Magnuson et al., 1993).

There are several factors that influence the chain length produced. The MAT can catalyse the transfer between CoA and the ACP/KS in both directions, but favours two and four carbon chains. This prevents the release of the longer FAS-derived acids through transfer to CoA. The rate of elongation rapidly reduces at more than 16 carbons, and the TE domain demonstrates specificity towards fatty acids with 16 or more carbons. The combination of these factors control chain length (Smith et al., 2003).

1.2.1.2 Structural organisation of fatty acid synthases

All these enzymatic domains (KS, MAT, ACP, KR, DH, ER, TE) together form fatty acid synthases (FAS). The organisations of these domains can be classified into two types. In type I FAS, as seen in mammals, these domains are all present on a single polypeptide. In type II FAS, as seen in bacteria, the domains are on separate polypeptides that complex to give the FAS holoenzyme (Campbell and Cronan Jr., 2001).

The structure of the mammalian type I FAS has been solved (Figure 1.4) (Maier et al., 2008). The FAS forms a dimer, with the KS and MAT at the bottom, and the modifying domains (KR, DH, ER) at the top. The ACP and TE domains were not solved in the X-ray structure. This is likely because the ACP domain is allowed to flex by a linker, and therefore shuttle the fatty acid intermediates between the various

domains.

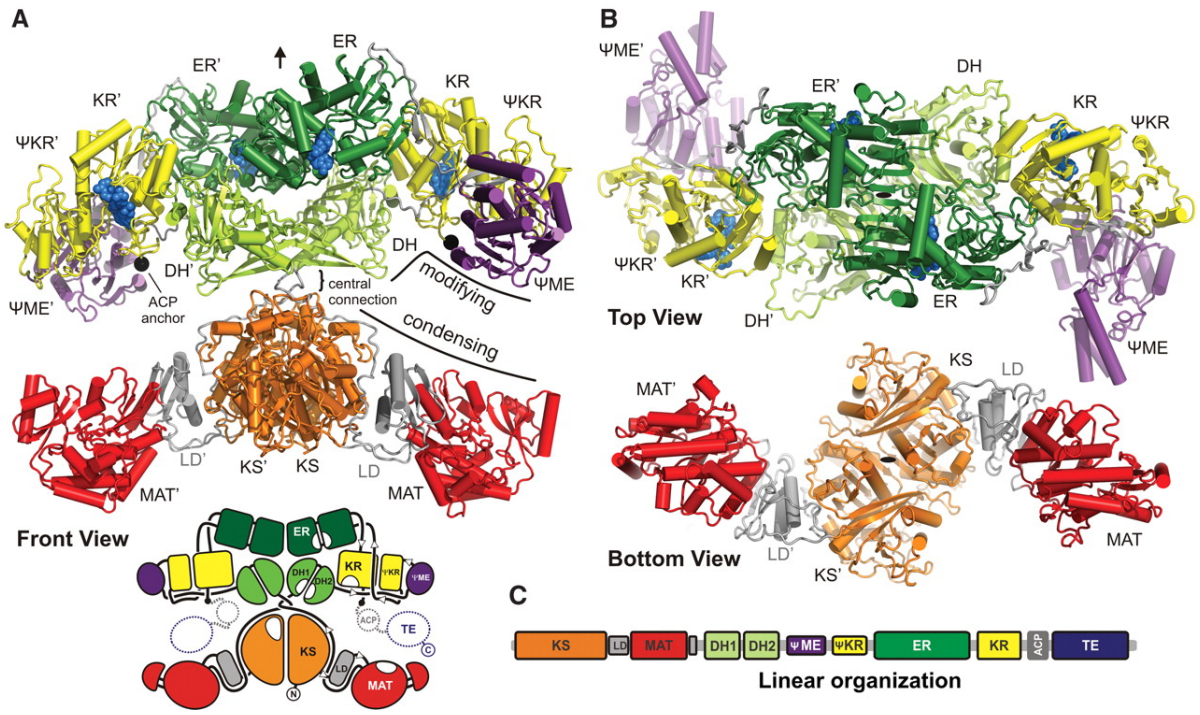


Figure 1.4: The structure of a mammalian type I FAS.
Two non-enzymatic pseudo domains were identified, pseudo-methyltransferase (ΨME) and pseudo-ketoreductase (ΨKR). The ACP and TE domains were not solved in the X-ray structure. The dehydratase domain was discovered to form a dimer of two hot dog folds, labelled DH1 and DH2. Reproduced from Maier et al. (2008).

1.2.2 Polyketide biosynthesis

Polyketide synthase (PKS) enzymes function in a very similar fashion to fatty acid synthases. The core PKS module has an ACP, which holds the extender unit and growing chain, an AT, which loads the ACP, and a KS, which catalyses the condensation. Whereas in fatty acid biosynthesis a KR, DH and ER are always present, in PKS enzymes some or all of these are often missing. PKSs can incorporate a range of starter and extender units, whereas FAS typically only use acetate and malonate respectively. Polyketide pathways often include tailoring enzymes, which further modify the polyketide, before the final compound is produced.

1.2.2.1 The types of polyketide synthase enzymes

There are three broad classes of PKS enzymes. Type I have all their domains covalently linked, such that they are on the same peptide. These can be modular, where there is one module to catalyse each Claisen condensation, as is seen in rapamycin biosynthesis for example (Staunton and Wilkinson, 1997). Type I PKSs can also be iterative, where one ketosynthase catalyses multiple condensations; an example of this is lovastatin biosynthesis, also known for being one of the first natural Diels-Alderases identified (Auclair et al., 2000). Both type I modular and iterative PKSs are used in mupirocin and thiomarinol biosynthesis, the focus of this work. In type II PKSs, each domain is on a discrete peptide, such as in actinorhodin biosynthesis. Finally, type III PKSs lack ACPs, and function iteratively; an example is flavolin biosynthesis (Shen et al., 2002).

Type I PKSs can be further divided into *cis*-AT, where the acyl transferase (AT) domain is positioned behind the KS on the same peptide, and *trans*-AT where the AT is on a separate peptide to the remaining PKS domains (Piel, 2010).

1.2.2.2 Erythromycin biosynthesis

Erythromycin biosynthesis is the best-studied type I modular PKS pathway. The biosynthesis genes were located in 1989 by genomic mapping of non-producer *Saccharopolyspora erythraea* mutants (Vara et al., 1989), and partially sequenced in 1990 by "walking along the chromosome" from the resistance gene *ermE* (Cortes et al., 1990). Three large genes were identified, *eryAI*, *eryAII*, and *eryAIII*; these encode the *cis*-AT PKS enzymes named DEBS1, DEBS2, and DEBS3 respectively. Each protein has two modules, to give six in total (Figure 1.5) (Bevitt et al., 1992; Donadio et al., 1991). Interaction between the three peptides is governed by N and C terminal docking domains (Weissman, 2006).

DEBS1 begins with a loading module, consisting of an acyl transferase (AT) and an ACP domain. The ACP is activated through the transfer of a phosphopantetheine

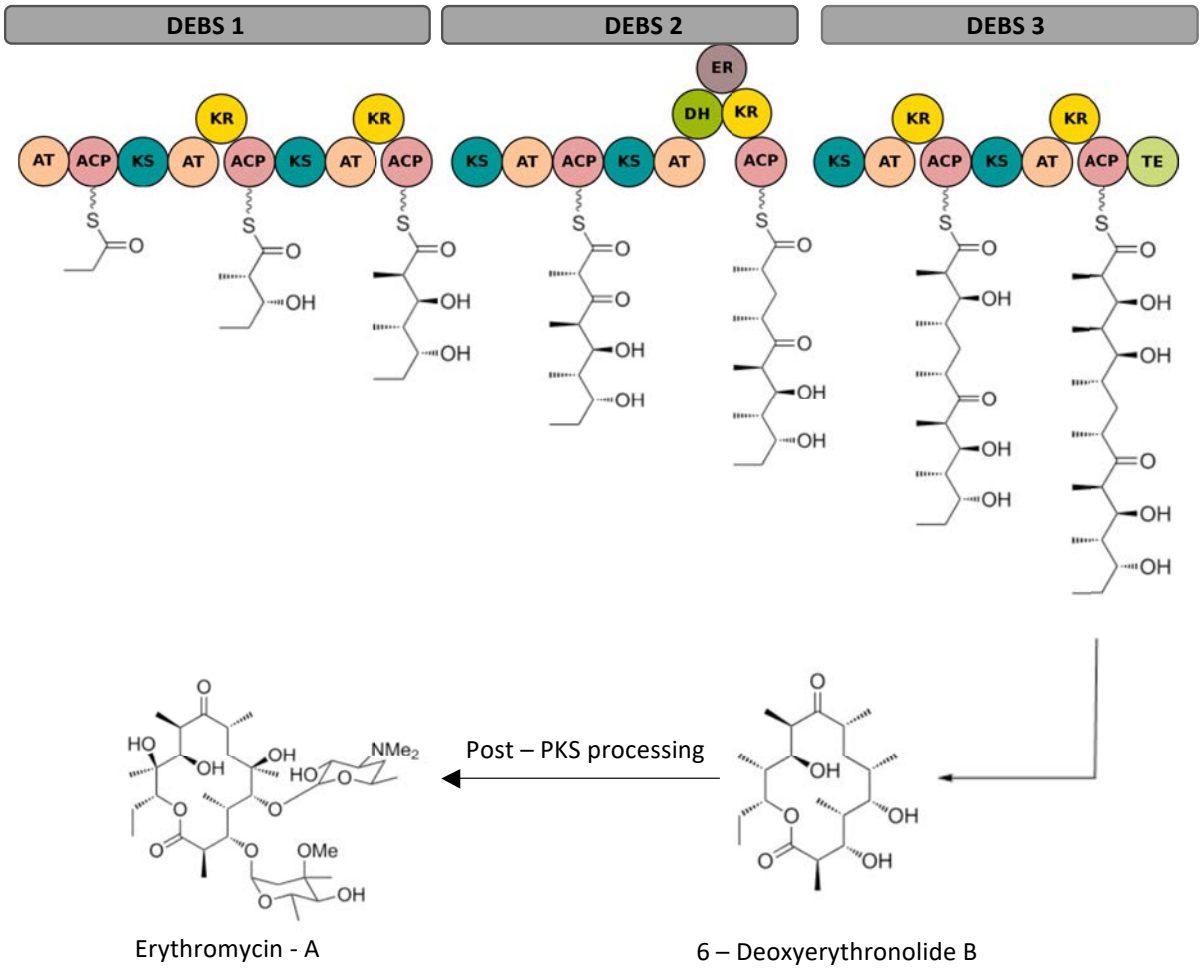


Figure 1.5: The erythromycin biosynthesis pathway.
The PKS pathway consists of six modules spread over three peptides. Each module catalyses a successive Claisen condensation, and the KR, DH and ER domains modify the β-keto groups where present. The thioesterase domain both releases the polyketide and catalyses cyclisation, yielding 6-Deoxyerythronolide B. This is then processed by tailoring enzymes to give erythromycin (Bevitt et al., 1992; Donadio et al., 1991).

(PP) arm from CoA onto a conserved serine on the ACP, catalysed by a phosphopantetheinyltransferase, giving the holoenzyme (Chan et al., 2009). A propionate starter unit is loaded onto the PP arm of the first ACP by the AT domain. In each of the six modules, the KS catalyses the condensation of the growing polyketide with a methylmalonate extender unit.

Each additional β-keto group is modified as dictated by the presence of the optional modules. Module 1 has a KR domain, and therefore the β-keto is reduced to a hydroxy group. Module 3 has none of the optional domains, and therefore the

β-keto is not modified, whereas module 4 has a KR, DH, and ER, and the group is therefore fully reduced, like in fatty acid biosynthesis. After the six condensations, the TE domain catalyses release and cyclisation of the polyketide. As is typical for polyketide natural products, the polyketide released by the PKS is not the final secondary metabolite (Staunton and Weissman, 2001). Tailoring enzymes catalyse further modifications to give the final product, erythromycin.

1.2.2.3 The structure of a type I PKS enzyme

Historically, solving of the structure of a complete PKS has proven difficult due to their large size (Weissman, 2015), although the structure of PKS PikAIII from pikromycin has been solved relatively recently (Dutta et al., 2014). Pikromycin is a macrolide antibiotic produced by *Streptomyces venezuelae*. The biosynthetic cluster was identified in 1998 (Xue et al., 1998); four genes encode four type I modular PKS enzymes with six modules in total. The structure was obtained for PikAIII, which has one module consisting of KS, AT, KR and ACP (Figure 1.6).

Like mammalian FAS, PikAIII forms a dimer, although the organisation of the domains is clearly different. The AT and KR domains are close together, whereas the MAT and KR domains in mammalian FAS are further apart (Figure 1.4). However, some differences are to be expected given the lack of DH and ER domains in PikAIII.

The ACP domains were observed to occupy two positions, one near the KR and one near the KS. Mirroring the assigned role in fatty acid biosynthesis, the ACP is positioned to shuttle intermediates around the catalytic domains.

Whilst the insight into a complete PKS enzyme is extremely valuable, there is significant variability in the domain composition of PKS enzymes. As previously discussed, PKS enzymes can have all, some of, or none of the KR, DH, ER domains. *Trans*-AT systems exist where the AT domain is on a separate peptide (Piel, 2010); indeed mupirocin (Section 1.3) and thiomarinol (Section 1.4) are example *trans*-AT PKS systems. With this variability in the make up of modules in PKS enzymes,

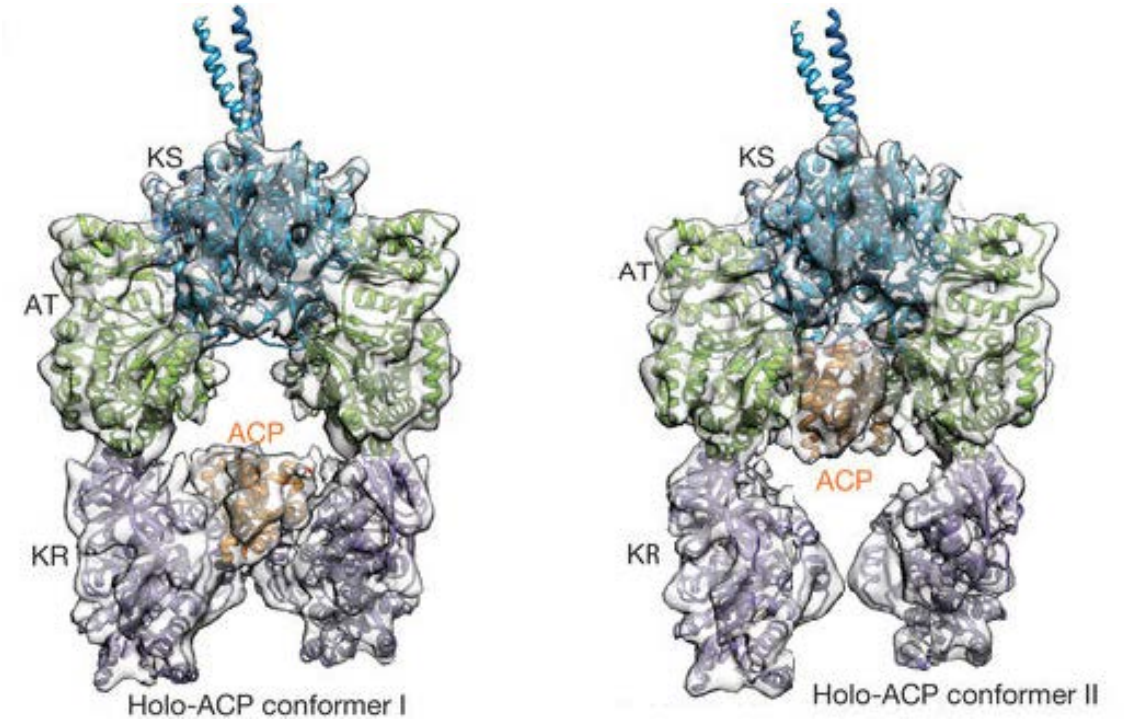


Figure 1.6: The structure of type I modular PKS PikAIII.
In a similar fashion to mammalian FAS, PikAIII was observed as a dimer. However, the AT domains appear closer to the KR side than the equivalent MAT in FAS. The ACP localised in two positions, the one near the KR and one near the KS. This is evidence of the flexibility that would allow the ACP to shuttle intermediates around the various domains. Modified from Dutta et al. (2014).

structural diversity could still be expected. Indeed functionality of PikAIII seems tied to conformational changes of the KR, which could not happen if the KR is not present.

1.2.3 Nonribosomal peptide synthases

Nonribosomal peptides (NRP) are another important class of secondary metabolites, that includes vancomycin. They are synthesised by nonribosomal peptide synthases (NRPS), which are very large enzymes; for example cyclosporin A is 1.6 MDa (Weber et al., 2015). The core NRPS module consists of three domains, adenylation (A), peptidyl carrier protein (PCP) and condensation (C) (Figure 1.7).

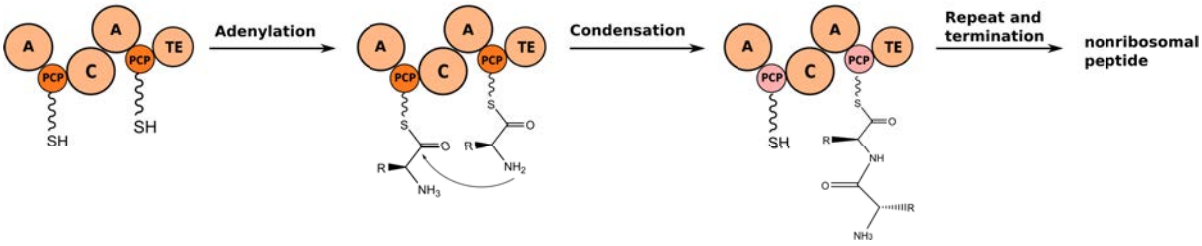


Figure 1.7: Nonribosomal peptide biosynthesis.
The diagram represents a minimal NRPS protein. The adenylation domains specifically select and activates amino acids. These are transferred onto the PP arm of the following PCP domains, where the condensation domain catalyses peptide bond formation. Typically NRPS proteins will have multiple modules, and the growing peptide will be synthesised by successive condensation domains.

Typically, NRP biosynthesis begins with an adenylation domain, which selects the amino acid to be activated. There is a region of approximately 10 amino acids on the A domain responsible for substrate specificity (Schwarzer et al., 2003). This is then transferred onto the following PCP. The PCP holds the peptide on a PP arm taken from CoA, in a similar fashion to the ACP in fatty acid and polyketide biosynthesis.

The extending amino acid is selected by the next A domain, adenylated and loaded onto the next PCP. The condensation domain then catalyses the formation of

the peptide bond. Typically, NRPS pathways include multiple modules, accounting for the large size the enzymes can reach. A TE domain is often present, that catalyses the release of generated peptide. In a similar fashion to polyketides, NRP pathways can include post-NRPS tailoring steps before the final product is generated.

NRPS pathways can be fused with other biosynthesis mechanisms, such as with a PKS in bleomycin or andrimid (Shen et al., 2002; Jin et al., 2006; Magarvey et al., 2008), or with a lipid as in daptomycin (Figure 1.8, 1.9).

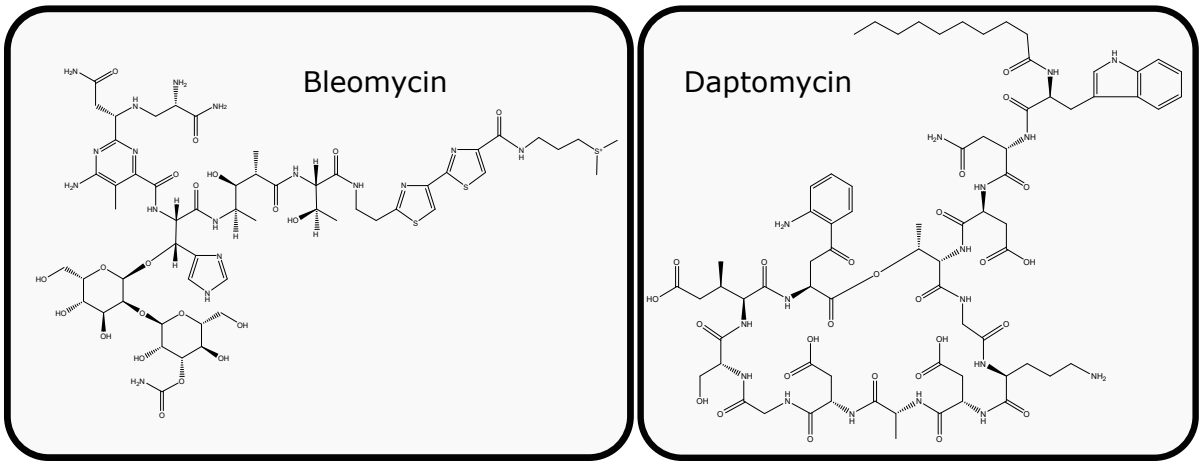


Figure 1.8: Examples of hybrid natural products.
The biosynthesis pathway for the lipopeptide antibiotic daptomycin joins a lipid to a cyclic nonribosomal peptide (Miao et al., 2005). Bleomycin is an anti-cancer drug produced by fused PKS and NRPS pathway components (Shen et al., 2002).

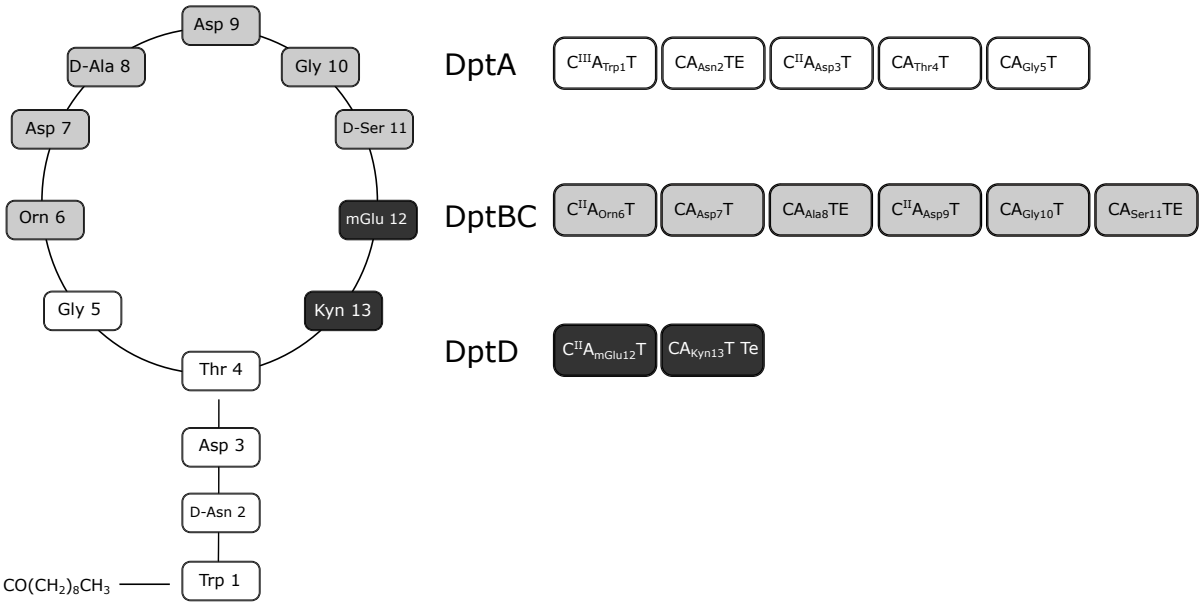


Figure 1.9: Overview of the biosynthesis of the lipopeptide antibiotic, daptomycin.
Three NRPS proteins hold the 13 modules between them to generate the daptomycin NRP. DptA lacks an N-terminal adenylation domain. Instead of beginning with an amino acid, daptomycin biosynthesis is thought to begin with a lipid provided on an ACP, DptF (Miao et al., 2005; Wittmann et al., 2008; Robbel and Marahiel, 2010).

1.3 Mupirocin

1.3.1 The characteristics of mupirocin

The antibiotic mupirocin was discovered in 1971 as a natural product of *Pseudomonas fluorescens* NCIMB10586 (Fuller et al., 1971). *P. fluorescens* is a soil bacterium, that forms part of the rhizosphere around plant roots. Mupirocin is a mixture of pseudomonic acids (Figure 1.10), which consist of monic acid linked by an ester to 9-hydroxynonanoic acid. Pseudomonic acid A (PA-A) makes up approximately 90% of the mixture, and is the primary active form.

Mupirocin is in clinical use, as a topical agent against MRSA, marketed as Bactroban by GlaxoSmithKline. The NHS spent approximately £1.0m on mupirocin prescriptions in England in 2016 (NHS, 2016). Mupirocin is limited to topical use due to hydrolysis of the ester in bodily fluids, and binding to blood albumin (Thomas et al., 2010).

1.3.1.1 Mode of action

The target of mupirocin was identified as isoleucyl tRNA synthetase (IRS); feeding *E. coli* with excess isoleucine was found to alleviate mupirocin inhibition (Hughes and Mellows, 1978). This mode of action was confirmed by the co-crystallisation of mupirocin with its target enzyme (Silvian et al., 1999; Nakama et al., 2001). The monic acid end (C14) of mupirocin appears to mimic the isoleucine side chain; the pyran ring and region of monic acid near the ester fit the ATP binding pocket (Figure 1.11). The 9-hydroxynonanoic acid may stabilise the binding of mupirocin to its target through hydrophobic interactions.

1.3.1.2 Resistance

Relatively low level mupirocin resistance can arise through single point mutations to IRS; for example V588F, G293V and V631F were observed to increase the MIC

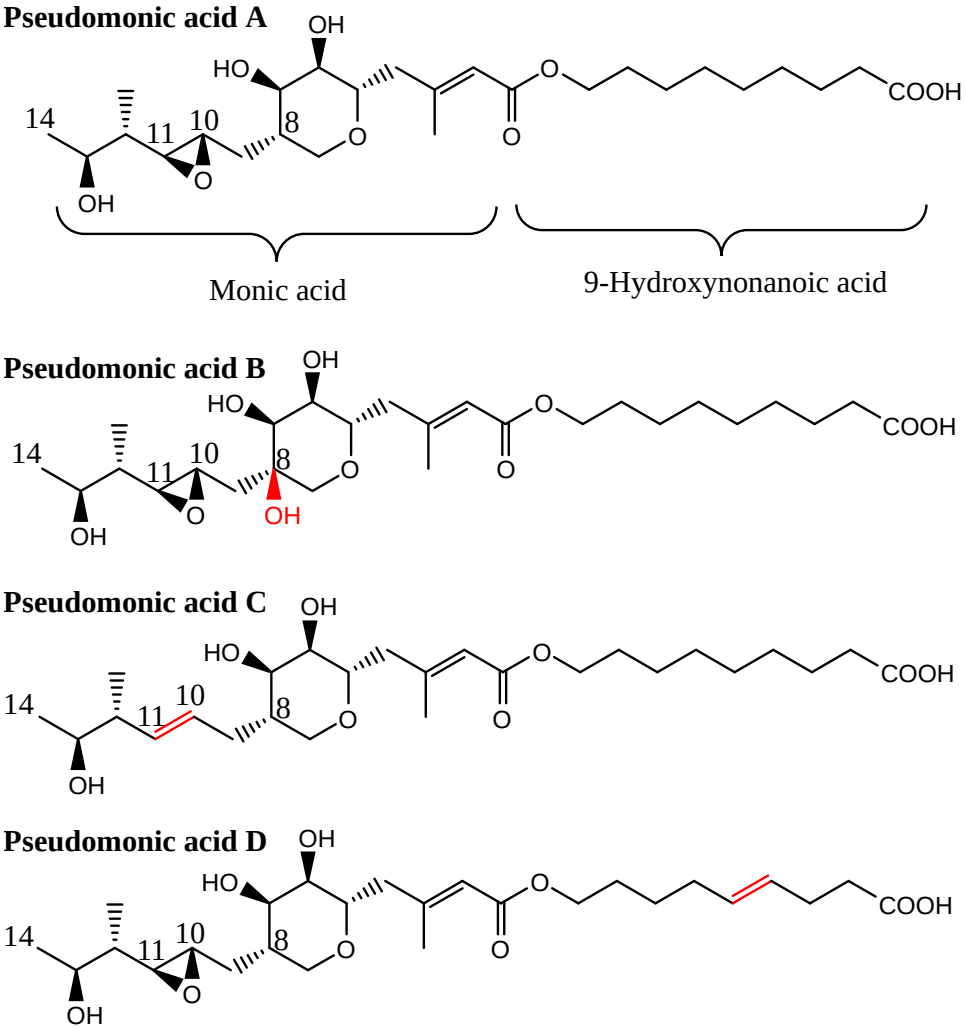


Figure 1.10: The chemical structures of pseudomonic acids.
Pseudomonic acids are comprised of monic acid linked by an ester to 9-hydroxynonanoic acid. The primary active form in mupirocin, pseudomonic acid A (PA-A), has C8-H, where pseudomonic acid B (PA-B) has C8-OH. The differences between the other pseudomonic acids from PA-A are highlighted in red.

to a maximum of 16 μg ml^{-1}. Further secondary point mutations were observed to increase resistance further, reaching up to 128 μg ml^{-1}; however these tripled the doubling time, suggesting a significant fitness cost (Hurdle et al., 2004).

Perhaps more concerning for its clinical use, high level resistance in MRSA and *Staphyloccus epidermis* can arise through acquisition of the eukaryotic like IRS encoded by *mupA*, a gene that has been found on a conjugative plasmid. Note that *mupA* shares no homology with its namesake in mupirocin biosynthesis, but instead the equivalent gene is *mupM*.

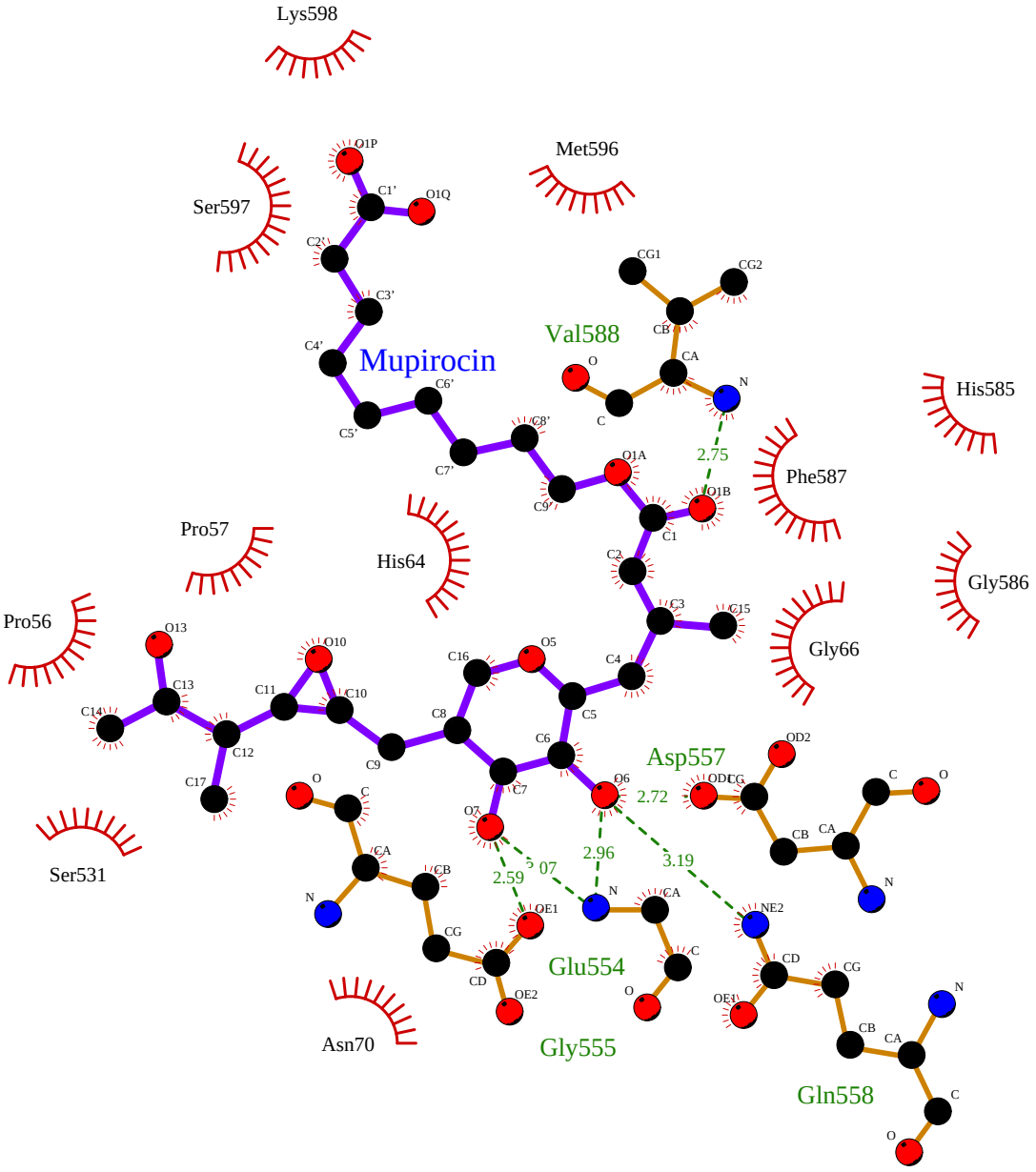


Figure 1.11: Interactions between mupirocin and its target enzyme isoleucyl tRNA synthetase.
Generated by Ligplot for PDB structure 1FFY (Silvian et al., 1999)

1.3.2 The mupirocin biosynthesis gene cluster

Transposon mutagenesis identified a 60 kb region of the *P. fluorescens* NCIMB10586 chromosome required for mupirocin biosynthesis (Whatling et al., 1995). The cluster was later sequenced, with 36 coding DNA sequences (CDS) identified over a 74 kb region (El-Sayed et al., 2003) (Figure 1.12). Genes encoding six large peptides with multiple domains were identified, and these were named *mmpA* through to *mmpF*. The remaining genes were named *mupA* through to *mupW* in chromosome order, with the exception of regulatory genes *mupI*, *mupR*, *mupX*, which were named for their homologs in the model *lux* regulon (El-Sayed et al., 2001). A further putative gene at the start of the cluster was later identified, named *mupZ*. The predicted functions based on bioinformatic homology and the assigned pathway role based on accumulated knowledge, are listed in Table 1.3.

Two large type I modular PKSs are encoded by *mmpA* and *mmpD*. Analysis of the domains reveals four modules in MmpD and two, excluding a loading module, in MmpA. A type I iterative PKS is encoded by *mmpB*, which like MmpD and MmpA, lacks functional AT domains. The AT domains are instead encoded on a separate gene *mmpC*. The remaining 30 genes include *mupM*, which encodes a eukaryotic-like IRS to prevent self action, and regulators *mupI*, *mupR* and *mupX*. The remaining genes either encode proteins accessory to the PKSs, or tailoring enzymes.

The mupirocin cluster is not unique to NCIMB10586, as a near identical cluster was sequenced in *Pseudomonas* sp. BRG-100, and a more divergent cluster with around 60 to 80% amino acid similarity was sequenced in *Pseudomonas pyschrotolerans*. The *mup* genes also share between 35% and 56% similarity with the thiomarinol biosynthesis genes (Section 1.4). A PCR screen of 195 *Pseudomonas* strains for an *mmpD* DNA fragment, identified 5 further strains carrying the cluster (Matthijs et al., 2014).

Table 1.3: Functions of mupirocin biosynthesis cluster genes

Gene	Bioinformatically predicted function	Putative function in pathway
mupZ	Hypothetical protein	Unknown
mupA	Reduced flavin mononucleotide oxygenase	Unknown
mmpA	Type I modular PKS	Monic acid biosynthesis
mupB	3-oxo-ACP-synthase	Unknown
mmpB	Type I iterative PKS	9-hydroxynonanoic acid biosynthesis
mmpC	AT and ER	AT activity for PKSs
mmpD	Type I modular PKS	Monic acid biosynthesis
mupC	NADH oxidase	Pyran ring oxidation state tailoring
macpA	ACP	Unknown
mupD	3-oxo-ACP-synthase	Unknown
mupE	ER	*In trans* reduction at MmpB
macpB	ACP	Unknown
mupF	KR	Pyran ring oxidation state tailoring
macpC	ACP	Site of β branching
mupG	3-oxo-ACP-synthase	HCS cassette (β branching)
mupH	HMG CoA synthase	HCS cassette (β branching)
mupJ	Enoyl CoA hydratase	HCS cassette (β branching)
mupK	Enoyl CoA hydratase	HCS cassette (β branching)
mmpE	KS0, ACP and oxidoreductase	Formation of C10,11 epoxide
mupL	Hydrolase	Unknown
mupM	Isoleucyl-tRNA synthetase	Mupirocin resistance
mupN	Phosphopantetheinyl transferase	Transfer of PP arms to ACPs
mupO	Cytochrome P450	Removal of the C8-OH
mupP	Dehydratase	Removal of the C8-OH
mupQ	Acyl CoA synthase	3-HP starter unit loading
mupS	3-oxo-ACP reductase	3-HP starter unit loading
macpD	ACP	Holds 3-HP starter unit for MmpB
mmpF	KS	Unknown
macpE	ACP	Site of C8-OH removal
mupT	(2Fe-2S) ferredoxin	Electron transport to MupW
mupU	Acyl CoA synthase	Loading onto MacpE
mupV	Oxidoreductase and TE	Release from MacpE
mupW	Rieske dioxygenase	Closure of pyran ring
mupR	Transcriptional activator	Gene expression regulation
mupX	Amidase	Degradation of the lactone
mupI	N-acyl homoserine lactone synthase	Production of quorum sensing molecule

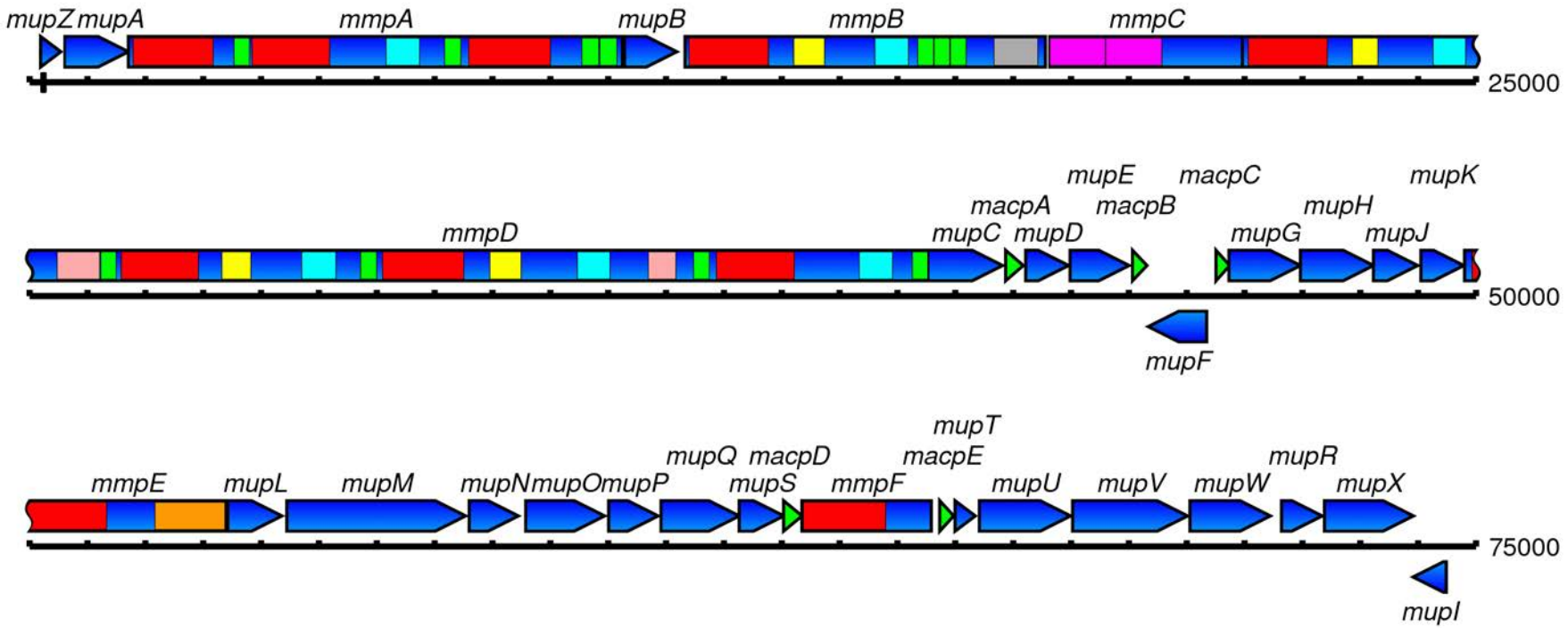


Figure 1.12: Genetic organisation of the mupirocin biosynthetic cluster.
The blue arrows represent the mupirocin biosynthesis genes over the 74 kb region of the chromosome. PKS domains are coloured as follows: red = KS, green = ACP, yellow = KR, blue = DH, grey = TE, orange = OR, pink = AT, peach = MT. Diagram produced by Anthony Haines, University of Birmingham.

1.3.2.1 Gene regulation

Expression of the mupirocin cluster is turned on at late exponential or early stationary phase, based on fusions of *mupA* with the reporter gene *xylE* (El-Sayed et al., 2001). This supports the observation that a significant yield of PA-A is produced after 20 hours (Mantle et al., 2001). Gene expression in the mupirocin cluster is regulated by a quorum sensing system, where small diffusible compounds produced by the organism give a measure of self cell concentration (El-Sayed et al., 2001).

The signalling molecule is homoserine lactone, produced by the homoserine lactone synthase encoded by *mupI*. This molecule is thought to activate the transcription factor MupR, which induces mupirocin gene expression. However, addition of the homoserine lactone or induction of *mupI* expression was not observed to increase mupirocin gene expression (Hothersall et al., 2011). Expression of the putative transcriptional activator MupR increased the expression of the cluster significantly. Finally, *mupX* encodes an amidase, thought responsible for degradation of the signalling lactone.

1.3.3 The mupirocin biosynthesis pathway

Mupirocin biosynthesis can be divided into two broad sections, synthesis of the polyketide backbone by the PKS enzymes, and tailoring of these compounds to yield mupirocin.

1.3.3.1 Synthesis of the polyketide backbone

The monic acid precursor is synthesised by type I modular PKSs, MmpD and MmpA. Before biosynthesis can occur, PP arms are transferred onto the ACP domains, catalysed by phosphopantetheinyl transferase MupN (Shields et al., 2010). The proposed pathway is shown as a cartoon in Figure 1.13.

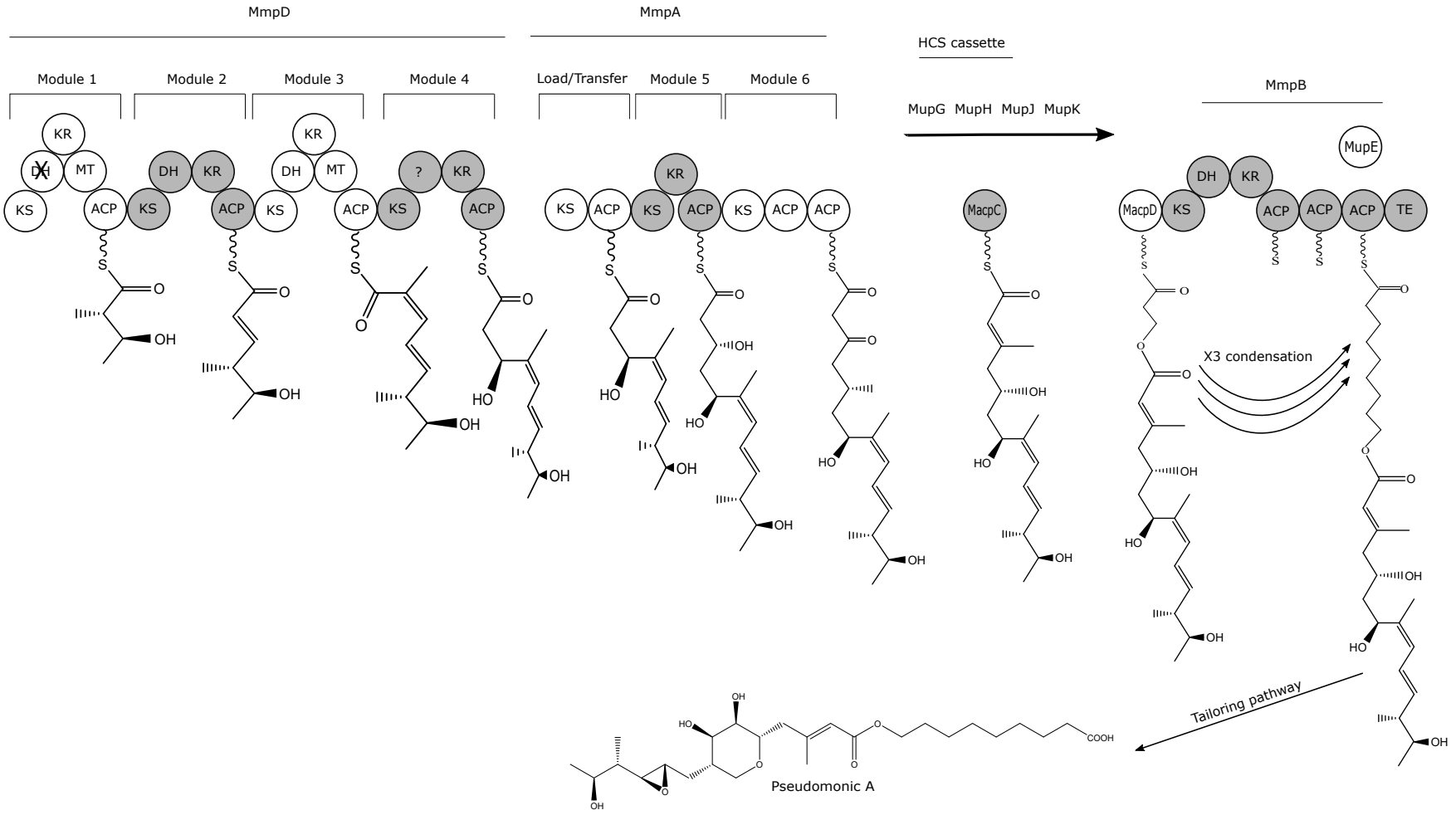


Figure 1.13: Biosynthesis of the mupirocin polyketide backbone.
Type I modular PKS enzymes MmpD and MmpA generate the monic acid precursor. Type I iterative PKS MmpB is thought responsible for 9-HN biosynthesis. The MmpB starter unit, 3-hydroxyhydroxypropionate, is shown joined by an ester bond to the monic acid precursor. This represents the current working hypothesis, although the exact timing and mechanism of ester bond formation has not been determined (El-Sayed et al., 2003; Thomas et al., 2010).

Polyketide biosynthesis begins with loading from acetyl-CoA onto the module 1 KS on MmpD, potentially via the module 1 ACP. The extender unit, malonate, is loaded onto the module 1 ACP. Loading of these units is catalysed by AT domains provided *in trans* on MmpC. The first KS catalyses a Claisen condensation (Heath and Rock, 2002), extending the acetate starter unit by a β-keto group. This is reduced by the KR domain, and an α-methyl group added by the MT.

The growing polyketide is passed along the modules, with each performing a Claisen condensation and modifications where the domains are present. The MmpD modules perform four condensations, before loading the polyketide chain onto MmpA. The first module of MmpA has a non-elongating KS^0, which does not catalyse a condensation, but instead is thought to load the output from MmpD onto MmpA. Two further condensations are performed on MmpA to give the monic acid precursor. The proposed scheme logically fits the bioinformatically predicted domains on the PKS enzymes.

No mupirocin production was observed in mutant strains lacking of the MmpA KR and KS^0 domains, confirming that they are essential to mupirocin biosynthesis (El-Sayed et al., 2003). The nature of the formation of the ester bond between monic acid and 9-HN is not yet clear. The most likely possibility is that monic acid is esterified to the 9-HN starter unit, and the 9-HN is synthesised on monic acid. Another possibility is that the monic acid and 9-HN are produced separately and linked later. However, this seems unlikely, as knockouts of the majority of mupirocin cluster genes have been generated and tested, and no accumulation of individual monic acid and 9-HN has been observed.

The iterative PKS MmpB is proposed responsible for 9-HN biosynthesis. No mupirocin production was observed in a strain lacking the MmpB TE domain (El-Sayed et al., 2003), demonstrating that it is required for mupirocin biosynthesis. Induction of expression of the CoA ligase TmlU in the mupirocin producer NCIMB10586 yielded pseudomonic acids with truncated 9-HN (Omer-Bali, 2013). This suggests

that TmlU interacts with the components responsible for 9-HN biosynthesis in the pathway. An interaction between TmlU and the TE domain of MmpB was detected by the Bacterial Two Hybrid system (Alsammarraie, 2016). Together, this strengthens the logical hypothesis, that MmpB is responsible for 9-HN biosynthesis.

MmpB functions in a iterative fashion, where the KS catalyses three condensations to give the 9-HN product. The starter unit is thought to be 3-hydroxypropionate, and the extender unit malonate. In a similar fashion to MmpA and MmpD, MmpB lack an AT domain, and this is thought to be provided *in trans* by MmpC. MmpB lacks an ER domain, which is required for full reduction of the fatty acid. This could also be provided *in trans*, potentially by ER domains on MupE or MmpC.

MmpB has three ACPs, the need for which is not yet entirely clear. In-frame deletion of each of the ACPs had no significant impact on mupirocin production, deletion of two domains decreased production, whereas deletion of all three abolished production (Rahman et al., 2005). However, no mupirocin production was detected in a strain with a point mutation inactivating ACP5. It seems that the ACP triplet provide a function in parallel, rather than they are all strictly required.

1.3.3.2 Tailoring steps

As is typical in polyketide natural products, there are several tailoring steps that occur before production of the final metabolite mupirocin (Figure 1.14). After biosynthesis of the monic acid precursor by MmpA and MmpD, β-branching is thought to occur, which is the addition of a β-methyl group. The components required have been identified as MupG, MupH, MupJ, MupK and MacpC (Wu et al., 2008; Haines et al., 2013).

MupW and MupT are thought responsible for closure of the pyran ring. No PA-A was produced by strains with deletions of either *mupW* or *mupT* (El-Sayed et al., 2003; Cooper et al., 2005). Production of pseudomonic acids without a closed pyran ring was detected in the Δ*mupW* strain (Gao et al., 2014).

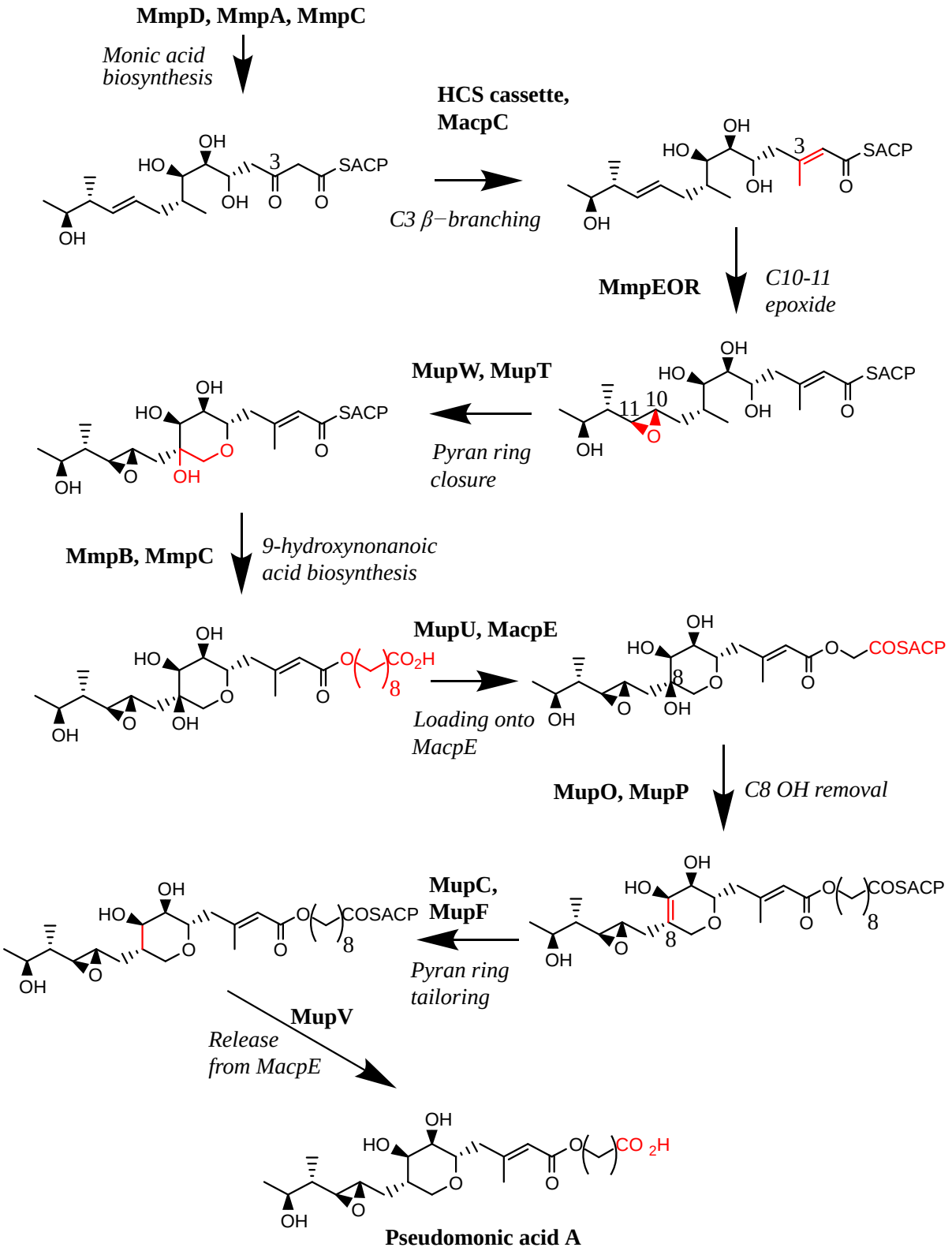


Figure 1.14: The proposed mupirocin tailoring pathway.
Modifications at each step are highlighted in red (Cooper et al., 2005; Hothersall et al., 2007; Gao et al., 2014, 2017).

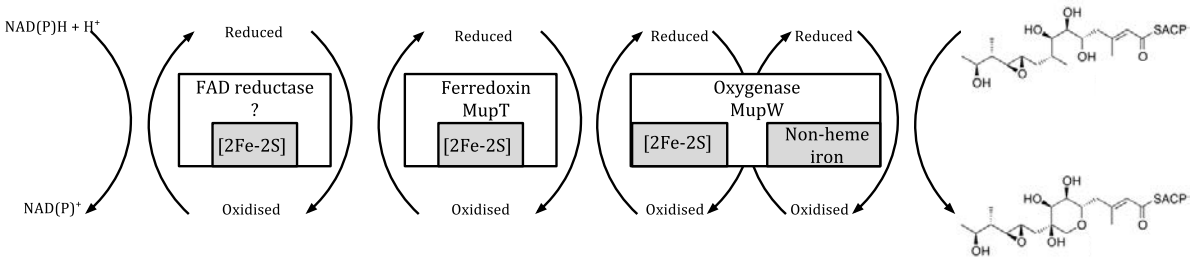


Figure 1.15: Proposed functions of MupT and MupW.
MupT and MupW may form an electron transport chain starting with an unidentified reductase enzyme. This provides an electron, reducing MupT, which in turn reduces the [2Fe-2S] of MupW, which then reduces the catalytic non-heme iron.

MupT is a Rieske ferredoxin, a small protein that contains an [2Fe-2S] cluster. It is likely that MupT and MupW function like the characterised Rieske dioxygenases they show homology to (Ferraro et al., 2005; Barry and Challis, 2013). These form an electron transport chain, which begins with the reduction of the [2Fe-2S] cluster of MupT by an unidentified reductase (Figure 1.15). MupT acts as a shuttle, in turn reducing the [2Fe-2S] of the Rieske dioxygenase, MupW. MupW likely forms a homotrimer, such that the [2Fe-2S] is placed in close proximity to the active site mononuclear iron. MupW can then catalyse pyran ring closure, though the exact chemical mechanism of this is not yet clear.

There are also tailoring steps thought to occur after 9-HN biosynthesis on MmpB. A switch to PA-B production was observed in strains with deletions to either *mupO*, *mupU*, *mupV*, and *macpE* (Cooper et al., 2005). PA-B, which is inactive, has C8-OH, whereas the highly active PA-A has C8-H, this step is explored in more detail in Section 3.1. Strains with deletions of *mupC* and *mupF* were observed to generate compounds where the tailoring of the oxidation state of the pyran ring had not been completed (Scott et al., 2011; Gao et al., 2014).

1.4 Thiomarinol

Thiomarinol was discovered in 1993, as a secondary metabolite produced by *Pseudoalteromonas sp.* SANK73390, which was cultured from Japanese seawater (Shiozawa et al., 1993). The chemical structure of thiomarinol is similar to mupirocin, and in particular to PA-C (Figure 1.16). Thiomarinol, like PA-C, does not have the C-10,11 epoxide. Thomarinol also has one less carbon in the fatty acid, and C4-OH not found on PA-A.

The biosynthesis cluster has been sequenced, and is found on a 98 kb plasmid name pTML1 (Fukuda et al., 2011) (Figure 1.17). The largest difference between the mupirocin and thiomarinol clusters is the presence of the 12 kb NRPS *hol* genes on pTML1, thought responsible for pyrrothine biosynthesis. Otherwise, the two clusters share strong homology, with the thiomarinol cluster having homologs to all of the *mmp* genes present (named *tmp*).

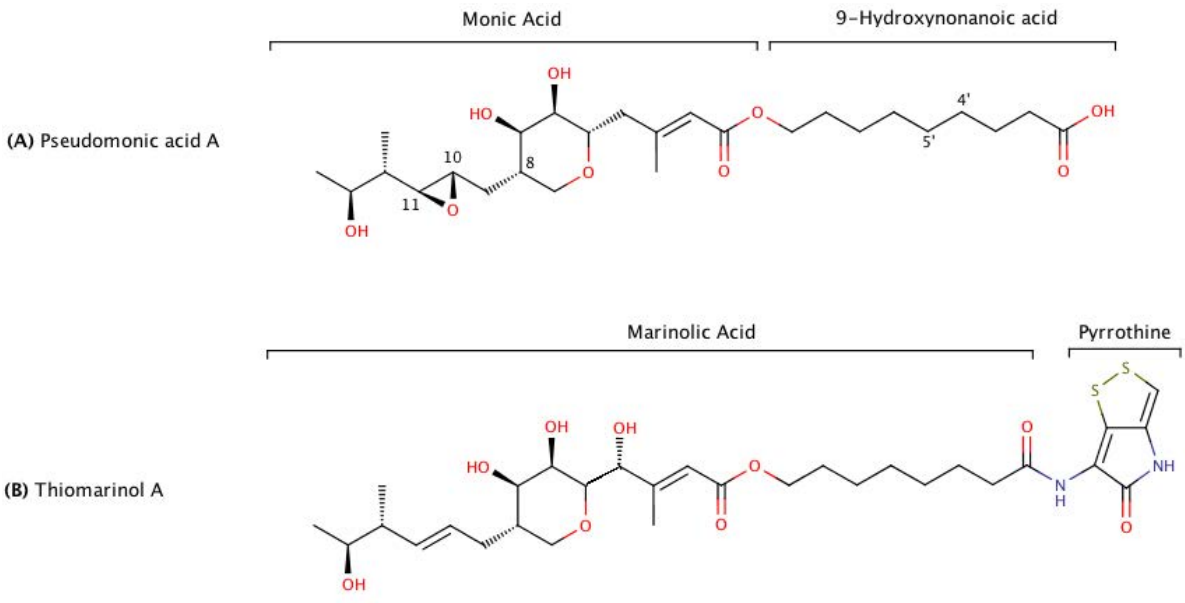


Figure 1.16: The structures of pseudomonic acid A and thiomarinol.
Thiomarinol is comprised of marinolic acid joined by an amide to pyrrothine. Marinolic acid is similar to pseudomonic acid; however marinolic acid lacks the C-10,11 epoxide, has an additional hydroxyl at C-4, and has one less carbon in the fatty acid.

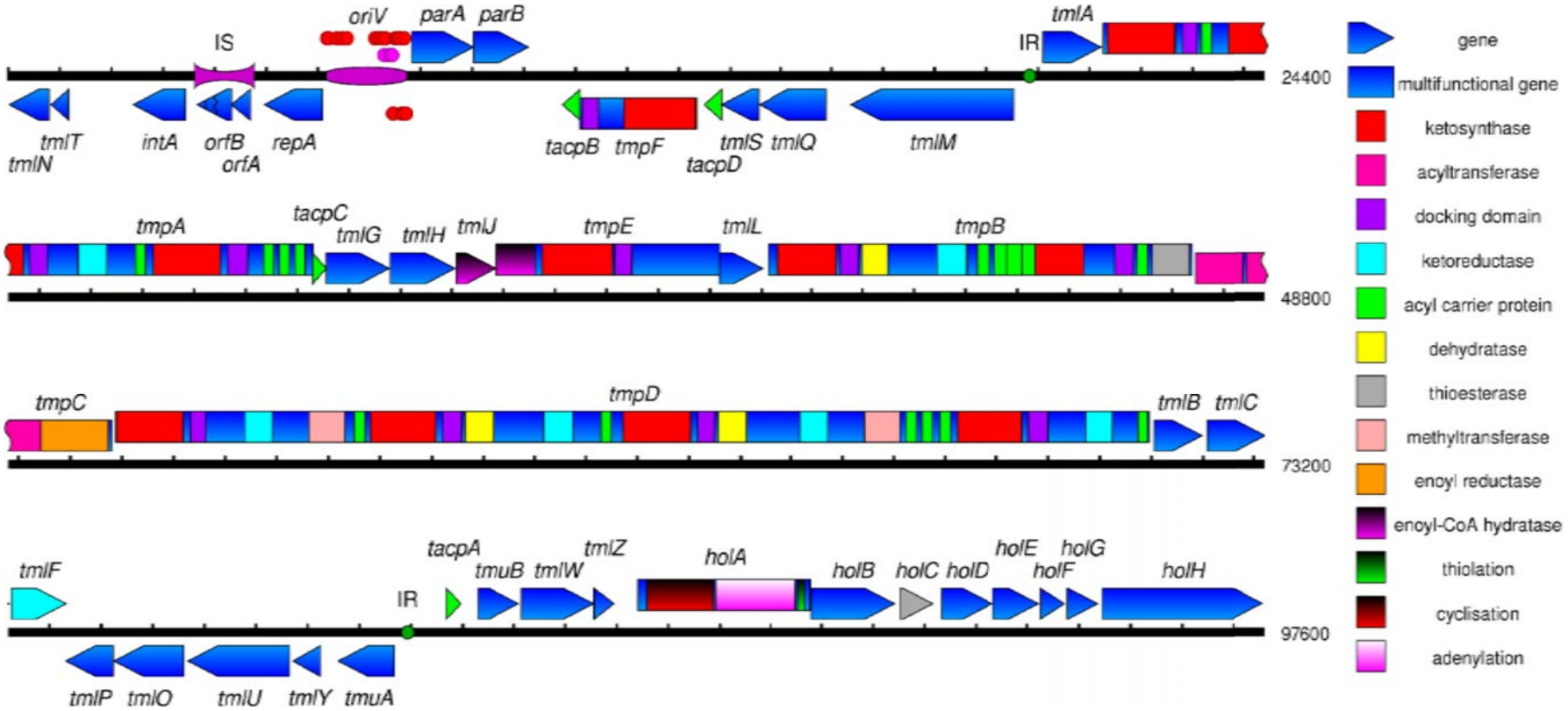


Figure 1.17: The genetic organisation of the thiomarinol cluster.
Approximately 88 kb of the 98 kb pTML1 encodes biosynthesis genes, which includes the 12 kb NRPS cluster thought responsible for holomycin (pyrrothine) biosynthesis. The plasmid replicon occupies approximately 5.3 kb, and the remaining 3.5 kb region seems to be the remnants of a transposon. Reproduced from Fukuda et al. (2011).

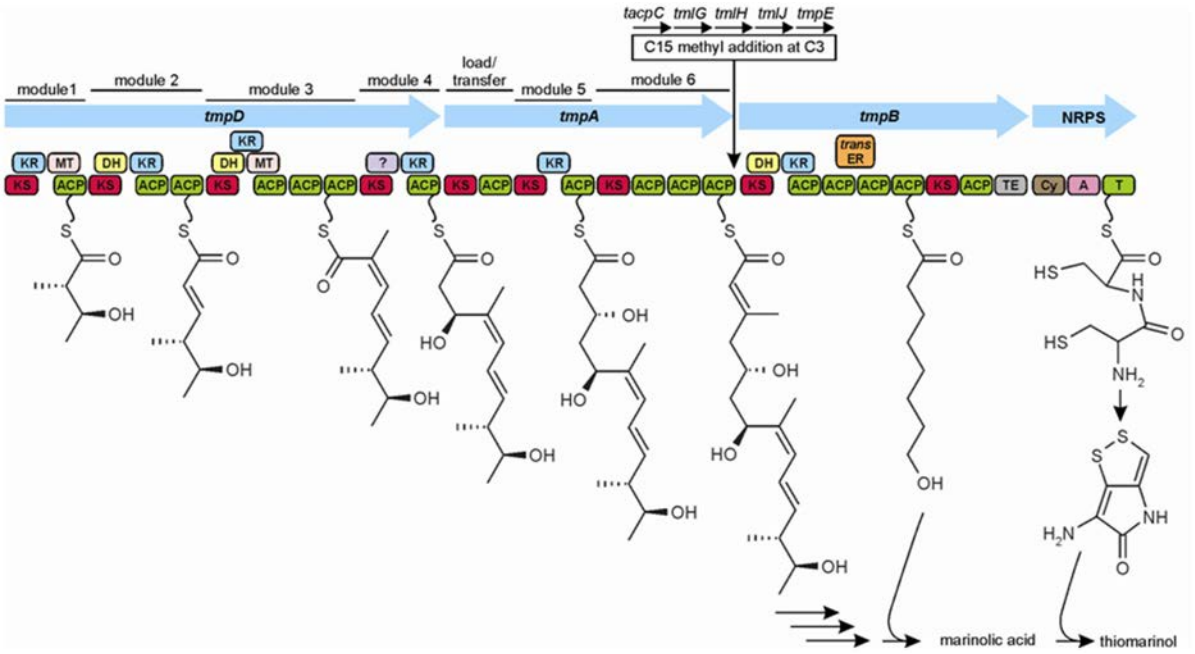


Figure 1.18: Proposed thiomarinol PKS-NRPS biosynthesis pathway.
TmpD and TmpA are type I module PKS enzymes proposed responsible for monic acid biosynthesis. TmpB is an iterative type I PKS that is proposed responsible for 8-hydroxyoctanoic acid. HolA is proposed to be an iterative NRPS that, together with HolB, C, D, F-H, is responsible for pyrrothine biosynthesis. Reproduced from Fukuda et al. (2011).

Biosynthesis of thiomarinol is thought to occur in a similar manner to mupirocin biosynthesis (Figure (1.18). Six modules across type I *trans*-AT modular PKS enzymes TmpD and TmpA are proposed responsible for biosynthesis of monic acid. The tailoring of the monic acid precursor in thiomarinol is different to in mupirocin, as the former has C10,11-alkene where the latter has C10,11-epoxide. Installation of this epoxide is proposed to be catalysed by the oxidoreductase (OR) domain of MmpE of mupirocin biosynthesis, which is bioinformatically predicted to hold a non-elongating KS^0, ACP and OR domains (Gao et al., 2014). Intriguingly, thiomarinol has a homolog to this enzyme, TmpE, despite the epoxide not being present in thiomarinol. TmpE also has an extra predicted hydratase domain at the N-terminal side.

The fatty acid component, 8-hydroxyoctanoic acid (8-HO), is proposed to be synthesised by the type I iterative PKS TmpB. This has an additional ACP in module 1 and an additional second module compared with its mupirocin homolog, MmpB.

A logical route to biosynthesis of 8-HO would be an acetate starter unit with three condensations of malonate extender units, similarly to fatty acid biosynthesis. However, radiolabelling experiments indicated the starter unit is 4C, and proposed to be 4-hydroxybutyrate, with two condensations to give the 8C fatty acid (Murphy et al., 2014). TmpB is discussed in more detail in Section 4.1.

There are some additional genes on pTML1 with no homolog present in the mupirocin cluster: *tmlY*, *tmuA*, *tmuB*. The function of *tmlY* is unknown, *tmuA* appears to encode a DNA binding protein, and the *tmuB* product has been implicated in C4-OH formation (Mohammad, 2017). TmlU has been demonstrated to have a different function to its mupirocin homolog, MupU. In mupirocin biosynthesis, a switch to PA-B production was observed with deletion of *mupU* (Cooper et al., 2005). However, in thiomarinol, mutation of *tmlU* caused the buildup of marinolic acid and pyrrothine (Fukuda et al., 2011). TmlU was therefore first assigned as an amide ligase; however *in vitro* studies have since demonstrated it is a CoA ligase that activates marinolic acid as a CoA thioester (Dunn et al., 2015). The formation of the amide bond between marinolic acid-CoA and holomycin is then thought to be catalysed by HolE.

The similarities between the mupirocin and thiomarinol clusters could facilitate the fusion of elements from both pathways. A mupirocin-pyrrothine derivative has been generated by feeding PA-A to a Pks- SANK73390 mutant, although the compound demonstrated reduced activity compared to both mupirocin and thiomarinol (Murphy et al., 2011). However, reverse genetic investigations of the thiomarinol pathway have been hindered by the high failure rate of mutagenesis methods on pTML1 in SANK73390.

Investigations of mupirocin biosynthesis in *P. fluorescens* NCIMB10586 have made extensive use of pAKE604 for the generation of scar-less chromosomal mutations, such as in-frame deletions (Cooper et al., 2005; Hothersall et al., 2007; Gao et al., 2014, 2017) or point mutations (Alsammarraie, 2016; Mohammad, 2017). Further examples of both are included in this thesis (Sections 3.2.1.2, 3.2.4.1, 4.2.4.2, 4.2.5.3), using an

improved pAKE604 generated in this work, pJC70. Both plasmids have an *E. coli* pMB1 origin of replication from pUC18, which does not function in *Psuedoalteromonas* or *P. fluorescens*. For generation of scar-less mutations, two arms of typically 500 bp DNA are designed flanking the desired mutational sequence in pAKE604. Formation of a cross-over between one of these arms and the homologous region of the 10586 chromosome, mediates insertion of the plasmid into the desired chromosomal sequence. Excision of the plasmid mediated by a second cross-over, which can occur with equal frequency in the same arm causing a reversion to WT, or in the other arm, which removes the plasmid sequence leaving the desired mutation.

Despite the wide-spread successful use of this method in 10586, it has largely failed in mutagensis of pTML1 in SANK73390, with multiple unsuccessful unpublished mutagenesis attempts made by the Thomas group. A *tmuB* mutant was successfully generated, although required much replication and additional screening compared with the relative ease in 10586 (Mohammad, 2017). The three mutants described in Fukuda et al. (2011) as PKS-, NRPS- and TmlU-, are actually all insertion mutants. These were generated using a single homologous DNA arm, which allowed for insertion of the mutagenesis plasmid into the targeted gene, disrupting its function. However, with no excision method, this leaves the entire plasmid of approximately 8 kb inside the mutated gene. This is likely to have polar effects on neighbouring genes on pTML1, and as such it is not surprising that complementation of the TmlU- mutant failed (Alsammarraie, 2016). The cause for the high failure rate of double recombination style mutants in 73390 remains cryptic, especially as the same technique using plasmids carrying the same core features has been reported successful in other *Pseudoalteromonas* strains (Wang et al., 2014).

1.5 Context and objectives of this work

1.5.1 General Context

As discussed in Section 1.1, there is a clear and immediate need for new antibiotics, in particular new chemical scaffolds. One avenue towards discovering and generating these scaffolds is the modification of existing natural products to generate new derivatives. Development of a mupirocin derivative conjugated to a non-ribosomal peptide (NRP) could provide immediately useful clinical compounds. Furthermore, the methods used to develop such a derivative would add to the "toolbox" of technologies available for engineering polyketide and NRP natural products in general.

The antibiotic thiomarinol consists of marinolic acid, which is similar to pseudomonic acid C, conjugated to pyrrothine (Section 1.4). As thiomarinol is a polyketide conjugated to NRP already, it provided a sensible starting point for considering how mupirocin NRP derivatives could be generated.

1.5.1.1 The generation of mupirocin pyrrothine derivatives catalysed by thiomarinol enzymes

Unlike *P. fluorescens*, the thiomarinol producing organism *Pseudoalteromonas sp.* SANK73390 is not readily genetically tractable, although three mutants were achieved in Fukuda et al. (2011). One was an insertion mutation to the first KS domain of TmpD, which is the equivalent type I modular PKS to MmpD. The thiomarinol producer carrying this *tmpD* mutation, SANK73390 PKS^{-}, did not produce the polyketide component of thiomarinol, marinolic acid. Mupirocin conjugated to pyrrothine was generated by mutasynthesis with this strain, through the feeding of PA-A to SANK73390 PKS- (Figure 1.19) (Murphy et al., 2011). Unfortunately, this derivative proved less active against *B. subtilis* than mupirocin or thiomarinol, with plate bioassay zone of inhibition diameters of 70 and 68% respectively. However, activity against MRSA remained at approximately the same level.

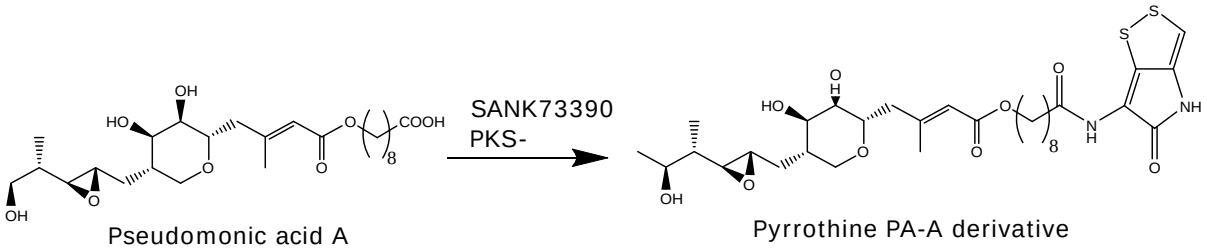


Figure 1.19: Generation of a PA-A-pyrrothine derivative by mutasynthesis (Murphy et al., 2011)

Encouragingly, these mutasynthesis experiments demonstrated that the thiomarinol enzymes responsible for conjugation of marinolic acid to pyrrothine are able to process PA-A. This suggested there was scope for introducing these thiomarinol enzymes into the mupirocin producer *P. fluorescens* NCIMB10586, and generating such derivatives *in vivo*. Catalysis of the amidation of marinolic acid and pyrrothine was originally attributed to TmlU, based on the lack of amide bond formation in SANK73390 Δ*tmlU* (Fukuda et al., 2011). However, the elegant *in vitro* experiments reported in Dunn et al. (2015) demonstrated that TmlU acts as a CoA ligase, which generates a marinolic acid-CoA thioester, prior to amide bond formation catalysed by a second enzyme, HolE (acetyl transferase).

Before the discovery of the requirement of HolE, *tmlU* was introduced on an expression plasmid into the mupirocin producer 10586, with the goal of later generating *in vivo* mupirocin derivatives (Omer-Bali, 2013). Surprisingly, induction of *tmlU* expression in 10586 proved inhibitory to mupirocin biosynthesis, and inactive metabolites were produced with truncated fatty acid and C8-OH (as in PA-B). In the same work, relief of this inhibition was reported in a strain with ACP6 of MmpB deleted; however this effect proved unreproducible (Alsammarraie, 2016).

The inhibition of mupirocin biosynthesis by TmlU represents a significant barrier to its use for generation of *in vivo* derivatives. *In vitro* experiments with HolE indicated it is fairly selective as to which amines it will accept. Therefore, even if the problem of TmlU inhibition were solved, it is likely that its use in mupirocin biosynthesis could only yield mupirocin conjugated to pyrrothine, which has already

been generated by mutasynthesis and demonstrated to be less active (Murphy et al., 2011).

1.5.1.2 The long term ambition of generating a library of mupirocin nonribosomal peptide derivatives

With the various problems associated with use of thiomarinol enzymes TmlU and HolE, an alternative objective was conceived where mupirocin derivatives would be generated *in vivo* through the splicing of NRPS from other systems to a mupirocin PKS.

The designated PKS was MmpB, as MmpB is proposed to offload pseudomonic acid B, having undergone all necessary Claisen condensations. Many PKS/NRPS hybrid natural products exist in nature, such as bleomycin (Shen et al., 2001), haliamide (Sun et al., 2016) and andrimid (Jin et al., 2006). Daptomycin (Miao et al., 2005) and the similar calcium dependent antibiotic (Hojati et al., 2002) were primary considerations, largely due to their relatively well studied coupling of fatty acids into NRPS biosynthesis (Bloudoff et al., 2013; Kraas et al., 2012). The daptomycin freestanding ACP, DptF, is proposed to hold decanoic fatty acid, prior to condensation with the first amino acid (Trp) by the first condensase on DptA (Wittmann et al., 2008). One possibility for generation of an *in vivo* mupirocin-NRP hybrid would be to load an ACP such as DptF fused to the mupirocin PKS MmpB, after which the daptomycin NRPS modules could potentially add amino acids.

There are significant hurdles to be overcome before a successful *in vivo* mupirocin daptomycin hybrid could be achieved, including the need for a mechanism to load DptF on MmpB without disrupting normal MmpB function. Also, the daptomycin NRPS enzymes need to show sufficiently relaxed specificity to accept this substrate, and in particular the first condensation domain of DptA. Instead it seems sensible to attempt to load a variety of ACPs from PKS (or fatty acid) NRPS hybrid systems. Hopefully, a library of mupirocin NRP derivatives could be generated, which could

then be screened for any potential clinical usefulness. More importantly, this accomplishment would add to our understanding of how to re-engineer PKS and NRPS systems in general, some examples of which were discussed in Section 1.1.3.2.

Regardless of the specific NRPS enzymes used, the proposed site of fusion to mupirocin biosynthesis is the iterative PKS MmpB. However, at the point of release from MmpB, the intermediate is proposed to be the inactive PA-B, with tailoring to PA-A only occurring after release (Figure 1.14). Therefore any derivatives would likely be generated with the inactive PA-B form.

1.5.2 Objectives

The core objective of the work described in this thesis was to re-engineer mupirocin biosynthesis so that PA-B to PA-A conversion occurs on MmpB. As covered in more detail in Section 3.1, the inactive PA-B is proposed to be released from MmpB, before loading to freestanding MacpE and conversion to the active PA-A. Conversion of PA-B to PA-A requires removal of the C8 hydroxl, and occurs by a surprisingly complicated process, with six proposed enzymatic steps (Gao et al., 2017). The successful re-engineering of this conversion process, to occur on the PKS MmpB, could unlock the possibility for generation of PA-A NRP *in vivo* derivatives through MmpB-NRPS hybrids in the future.

The main aim of the work presented in Chapter 3 was to characterise what genes are required for conversion of PA-B to PA-A in mupirocin biosynthesis. This was to be accomplished in two parts: firstly the systematic screening of mutants of all mupirocin tailoring genes for ability to convert PA-B to PA-A, followed by cloning of a set of genes to generate an expression plasmid that encodes all necessary functions. Our understanding of how this complicated conversion occurs has historically become more complex (Section 3.1.1), and this work aimed to comprehensively demonstrate what functions are required. Understanding how PA-B to PA-A conversion works is an important prelude to designing MmpB modifications to re-engineer how it occurs.

Chapter 4 then covers the core objective: re-engineering the timing or position of PA-B to PA-A conversion such that it occurs on the PKS MmpB. This was to be accomplished through fusions with a non-elongating KS^0 domain from thiomarinol biosynthesis, and initially through provision of thiomarinol tailoring enzymes *in trans*. However, later this transitioned into use of the thiomarinol KS^0 domain to load the normally freestanding MacpE whilst fused to MmpB. This would also establish the usefulness of this KS^0 domain for loading foreign ACPs, which in the future could be a potential method for loading DptF.

This work should hopefully add to the knowledge pool for engineering of polyketide biosynthesis in general, and could unlock the generation of *in vivo* mupirocin-NRP hybrids in the future.

Whilst undertaking these objectives, the methods used historically to investigate mupirocin biosynthesis in *P. fluorescens* NCIMB10586 have been appraised, and where possible improved. An alternative transconjugant selection procedure was developed that saves 8-12 days on the minimal media selection procedure used previously (Omer-Bali, 2013; Alsammarraie, 2016; Gurney, 2012). The suicide mutagenesis vector pJC70 was developed in this work to confer this same time-saving to the 10586 mutagenesis process.

Chapter 2

MATERIALS AND METHODS

2.1 Bacterial strains, plasmids and growth conditions

2.1.1 Bacterial strains and plasmids

Table 2.1: Bacterial strains used and generated in this study

Strain	Description and genotype	Source / reference
Bacillus subtilis 1604	Used as a gram positive test organism for mupirocin activity, *trpC2*	Moir et al. (1979)
Escherichia coli DH5α	Used for cloning and DNA propagation. *F- endAI thi-1 lacZ*Δ*M15*	Gibco BRL Hanahan (1985)
E. coli ER2925	MG1655 *endA1* derivative. *Dam- Dcm-* Cm^R Sm^R. Used for preparation of DNA lacking Dam methylation	New England Biolabs Ltd.
E. coli S17-1	Used for RP4 mediated conjugation. Chromosomal RP4 *tra*, thi, pro, $hsdR^-$, $hsdM^+$, *recA*, Sm^R	Simon et al. (1983)
Pseudoalteromonas sp. SANK 73390	Wild type thiomarinol producer	Daichi Sankyo Co. Ltd. Shiozawa et al. (1993)
Pseudomonas fluorescens SBW25	Reference strain without a mupirocin cluster, the genome has been sequenced	Rainey and Bailey (1996)
P. fluorescens NCIMB 10586	The wild type mupirocin producer	Fuller et al. (1971)
P. fluorescens Δ*mup-cluster*	10586 with a deletion of all mupirocin cluster genes, with the exception of transcriptional regulators *mupRXI*	Miller, C. (Unpublished)
P. fluorescens Δ*mupI mupA::xylE*	10586 with an in-frame deletion of *mupI* and reporter gene *xylE* fused to *mupA*	El-Sayed et al. (2001)
P. fluorescens Δ*macpA*	10586 with an in-frame deletion of *macpA*	Hothersall et al. (2007)
P. fluorescens Δ*macpB*	10586 with an in-frame deletion of *macpB*	Hothersall et al. (2007)
P. fluorescens Δ*macpC*	10586 with an in-frame deletion of *macpC*	Hothersall et al. (2007)

Continued on next page

Table 2.1 – *Continued*

Strain	**Description and genotype**	**Source / reference**
P. fluorescens Δ*macpD*	10586 with an in-frame deletion of *macpD*	Hothersall et al. (2007)
P. fluorescens Δ*macpE*	10586 with an in-frame deletion of *macpE*	Hothersall et al. (2007)
P. fluorescens Δ*mmpA*	10586 with an in-frame deletion of *mmpA*	El-Sayed et al. (2003)
P. fluorescens Δ*mmpB*	10586 with an in-frame deletion of *mmpB*	Hothersall, J. (unpublished)
*P. fluorescens mmpB*ΔACP567	10586 with an in-frame deletion of the *mmpB* ACP triplet	Rahman et al. (2005)
*P. fluorescens mmpB*ΔTE	10586 with an in-frame deletion of the *mmpB* TE domain	El-Sayed et al. (2003)
*P. fluorescens mmpE*ΔOR	10586 with an in-frame deletion of the oxidoreductase domain of *mmpE*	Gao et al. (2014)
P. fluorescens Δ*mupA*	10586 with an in-frame deletion of *mupA*	Hothersall, J. (unpublished)
P. fluorescens Δ*mupB*	10586 with an in-frame deletion of *mupB*	Hothersall, J. (unpublished)
P. fluorescens Δ*mupC*	10586 with an in-frame deletion of *mupC*	Hothersall et al. (2007)
P. fluorescens Δ*mupD*	10586 with an in-frame deletion of *mupD*	Hothersall et al. (2007)
P. fluorescens Δ*mupE*	10586 with an in-frame deletion of *mupE*	Hothersall et al. (2007)
P. fluorescens Δ*mupF*	10586 with an in-frame deletion of *mupF*	Hothersall et al. (2007)
P. fluorescens Δ*mupG*	10586 with an in-frame deletion of *mupG*	Hothersall et al. (2007)
P. fluorescens Δ*mupH*	10586 with an in-frame deletion of *mupH*	Hothersall et al. (2007)
P. fluorescens Δ*mupI*	10586 with an in-frame deletion of *mupI*	El-Sayed et al. (2001)

Continued on next page

Table 2.1 – *Continued*

Strain	**Description and genotype**	**Source / reference**
P. fluorescens Δ*mupJ*	10586 with an in-frame deletion of *mupJ*	Hothersall et al. (2007)
P. fluorescens Δ*mupK*	10586 with an in-frame deletion of *mupK*	Hothersall et al. (2007)
P. fluorescens Δ*mupL*	10586 with an in-frame deletion of *mupL*	Hothersall et al. (2007)
P. fluorescens Δ*mupM*	10586 with an in-frame deletion of *mupM*	Hothersall et al. (2007)
P. fluorescens Δ*mupN*	10586 with an in-frame deletion of *mupN*	Hothersall et al. (2007)
P. fluorescens Δ*mupO*	10586 with an in-frame deletion of *mupO*	Cooper et al. (2005)
P. fluorescens Δ*mupP*	10586 with an in-frame deletion of *mupP*	Gao et al. (2017)
P. fluorescens Δ*mupR*	10586 with an in-frame deletion of *mupR*	El-Sayed et al. (2001)
P. fluorescens Δ*mupS, Q* Δ*macpD* Δ*mmpF*	10586 with in-frame deletions of *mupS, mupQ, macpD, mmpF*	Yadav (2017)
P. fluorescens Δ*mupU*	10586 with an in-frame deletion of *mupU*	Cooper et al. (2005)
P. fluorescens Δ*mupU* Δ*mupL*	10586 with in-frame deletions of *mupU* and *mupL*	Section 3.2.4.1
P. fluorescens Δ*mupV*	10586 with an in-frame deletion of *mupV*	Cooper et al. (2005)
P. fluorescens mupV Y167F	10586 with a point mutation to domain 1 of *mupV*	Wilson, A. (Section 3.2.1)
P. fluorescens mupV V581A	10586 with a point mutation to domain 2 of *mupV*	Section 3.2.1
P. fluorescens mupV V581A H631A	10586 with a point mutations to domain 2 of *mupV*	Section 3.2.1
P. fluorescens Δ*mupW P. fluorescens* Δ*mupT*	10586 with in-frame deletions of *mupW* and *mupT*	Hothersall, J. (unpublished)
P. fluorescens Δ*mupX*	10586 with an in-frame deletion of *mupX*	Cooper et al. (2005)

Continued on next page

Table 2.1 – *Continued*

Strain	**Description and genotype**	**Source / reference**
P. fluorescens mupZ p.m.	Point mutation to *mupZ*	Mohammad (2017)
P. fluorescens Δ*mmpB* Δ*macpE*	10586 with in-frame deletions of *mmpB* and *macpE*	Section 4.2.5.3
*P. fluorescens*Δ*mmpB* Δ*macpE* Δ*mupU*	10586 with in-frame deletions of *mmpB*, *macpE* and *mupU*	Stephens, E. (unpublished)
*P. fluorescens mmpE*ΔOR	10586 with an in-frame deletion of the *mmpE* oxidoreductase domain	Gao et al. (2014)
P. fluorescens mmpB.JC50	10586 with MmpB/TmpB hybrid 1 in place of MmpB	Section 4.2.4.2
P. fluorescens mmpB.JC58	10586 with MmpB/TmpB hybrid 2 (full TE) in place of MmpB	Section 4.2.4.2
P. fluorescens mmpB.JC59	10586 with MmpB/TmpB hybrid 2 (short TE) in place of MmpB	Section 4.2.4.2
P. fluorescens Δ*mupU mmpB*.JC50	10586 Δ*mupU* with MmpB/TmpB hybrid 1 in place of *mmpB*	Section 4.2.4.2
P. fluorescens Δ*mupU mmpB*.JC58	10586 Δ*mupU* with MmpB/TmpB hybrid 2 (full TE) in place of *mmpB*	Section 4.2.4.2
P. fluorescens Δ*mupU mmpB*.JC59	10586 Δ*mupU* with MmpB/TmpB hybrid 2 (short TE) in place of *mmpB*	Section 4.2.4.2
P. fluorescens Δ*mupV mmpB*.JC50	10586 Δ*mupV* with MmpB/TmpB hybrid 1 in place of *mmpB*	Section 4.2.4.2
P. fluorescens Δ*mupV mmpB*.JC58	10586 Δ*mupV* with MmpB/TmpB hybrid 2 (full TE) in place of *mmpB*	Section 4.2.4.2
P. fluorescens Δ*mupV mmpB*.JC59	10586 Δ*mupV* with MmpB/TmpB hybrid 2 (short TE) in place of *mmpB*	Section 4.2.4.2
*P. fluorescens mmpE*ΔOR *mmpB*.JC50	10586 *mmpE*ΔOR with MmpB/TmpB hybrid 1 in place of *mmpB*	Section 4.2.4.2
*P. fluorescens mmpE*ΔOR *mmpB*.JC58	10586 *mmpE*ΔOR with MmpB/TmpB hybrid 2 (full TE) in place of *mmpB*	Section 4.2.4.2
*P. fluorescens mmpE*ΔOR *mmpB*.JC59	10586 *mmpE*ΔOR with MmpB/TmpB hybrid 2 (short TE) in place of *mmpB*	Section 4.2.4.2

Table 2.2: Plasmids used and generated in this study

Plasmid	Size (kb)	Description	Source
pAKE604	7.2	Suicide mutagenesis vector, Amp^R, Kan^R, *sacB*, pUC18 *ori*, RK2 *oriT*, *lacZα*	El-Sayed et al. (2001)
pJC70	6.6	Suicide mutagenesis vector, a derivative of pAKE604 with *bla* removed	Section 2.2.12
pJH10	13.8	Expression vector, RSF1010 *ori*, Tet^R, *lacI*, *oriT*	El-Sayed et al. (2003)
pUC18	2.6	Cloning vector, Amp^R, *lacZα*, pMB1 replicon	Yanisch-Perron et al. (1985)
pGEM-T	3.0	T overhang, linearised cloning vector, Amp^R, *lacZα*, pMB1 replicon	Promega
pAKE604/V.Y167F	8.2	MupV Y167F mutagenesis vector, derived from a 1 kb *Mfe*I/*Bam*HI insert cloned to pAKE604.	Wilson, A. (Section 3.2.1)
pAKE604/Δ*mupL*	8.2	Mutagenesis vector for in-frame deletion of *mupL*, derived from pAKE604	Hothersall et al. (2007)
pAKE604/Δ*mupU*	8.2	Mutagenesis vector for In-frame deletion of *mupU*, derived from pAKE604	Cooper et al. (2005)
pJH2	14.5	Gene *mupR* in pJH10	Hothersall et al. (2007)
pJH10/*tmlOPCF*	18.2	Genes *tmlO*, *tmlP*, *tmlC*, *tmlF* cloned using *Eco*RI/*Kpn*I to pJH10	Alsammarraie, Y., Wilson, A. (unpublished)
pJH10/*tmpYOPCF*	1.9	Genes *tmlY*, *tmlO*, *tmlP*, *tmlC*, *tmlF* cloned *Eco*RI/*Kpn*I to pJH10	Wilson, A. (unpublished)
pUC/*mmpB*CT2kb	4.7	Final 2.0 kb of *mmpB* cloned *Kpn*I to *Xba*I into pUC18	Section 4.2.1
pUC/*mmpB*	8.9	Front 4.2 kb of *mmpB* cloned into pUC/*mmpB*CT2kb, giving the full *mmpB* gene in pUC18	Section 4.2.1
pJH10/*mmpB*	20.0	Derived from pJH10 with *mmpB* cloned *Eco*RI/*Xba*I, Tet^R	Section 4.2.1
pJC41	6.0	Second module of *tmpB* cloned into pGEM-T	Section 4.2.2.1

Continued on next page

Table 2.2 – *Continued*

Plasmid	Size (kb)	Description	Source
pJC46	6.3	Second module of *tmpB* cloned from pJC41 (*Pci*I/*Xba*I) to pUC.2kb.*mmpB* (*Nco*I/*Xba*I)	Section 4.2.2.1
pJC50	21.6	MmpB/TmpB hybrid I in pJH10. Derived from a 3.6 kb *Kpn*I/*Xba*I fragment from pJC46 cloned into pJH10/MmpB	Section 4.2.2.1
pJC50.C1824A	21.6	MmpB/TmpB hybrid I in pJH10, with C1824A point mutation to KS^0 active site cysteine	Section 4.2.3.3
pJC56	6.4	MmpB/TmpB hybrid II (full length TE) pUC18 based construction vector. Derived from cloning PCR product *Bam*HI/*Xba*I into pJC46	Section 4.2.3.1
pJC57	6.3	MmpB/TmpB hybrid II (short length TE) pUC18 based construction vector	Section 4.2.3.1
pJC58	21.7	MmpB/TmpB hybrid II (full length TE) in pJH10. Derived from cloning a 3.8 kb *Kpn*I/*Xba*I fragment from pJC56 into pJC50	Section 4.2.3.1
pJC58.C1824A	21.7	MmpB/TmpB hybrid II (full length TE) in pJH10, with C1824A point mutation to KS^0 active site cysteine	Section 4.2.3.3
pJC59	21.6	MmpB/TmpB hybrid II (short length TE) in pJH10. Derived from cloning a 3.6 kb *Kpn*I/*Xba*I fragment from pJC57 into pJC50	Section 4.2.3.1
pJC59.C1824A	21.6	MmpB/TmpB hybrid II (short length TE) in pJH10, with C1824A point mutation to KS^0 active site cysteine	Section 4.2.3.3
pJC74.50	10.7	MmpB/TmpB hybrid I pJC70 based mutagenesis vector	Section 4.2.4.1
pJC74.58	10.0	MmpB/TmpB hybrid II (full length TE) pJC70 based mutagenesis vector	Section 4.2.4.1
pJC74.59	9.9	MmpB/TmpB hybrid II (short length TE) pJC70 based mutagenesis vector	Section 4.2.4.1
pJC99	4.0	MupV H631A mutagenesis fragment AT cloned into pGEM-T	Section 3.2.1.2

Continued on next page

Table 2.2 – *Continued*

Plasmid	Size (kb)	Description	Source
pJC102	7.6	MupV H631A mutagenesis vector, derived from 1.0 kb *Eco*RI/*Sal*I fragment cloned into pJC70	Section 3.2.1.2
pJC118	21.7	MmpB/TmpB hybrid generation III: TmpB ACP replaced by MacpE, followed by the TmpB TE domain	Section 4.2.5.1
pJC119	20.7	MmpB/TmpB hybrid generation III: TmpB ACP replaced by MacpE, without a TE domain	Section 4.2.5.1 (Connolly, J., Wilson, A.)
pJC120	21.7	MmpB/TmpB hybrid generation III: TmpB ACP replaced by MacpE, followed by the MmpB TE domain	Section 4.2.5.1 (Connolly, J., Wilson, A.)
pMMH6	21.4	Derived from pJH10, with *mupO,U,V,C,F,macpE* cloned *Kpn*I/*Xba*I	Macioszek (2009)
pJC124	22.4	Genes *mupO*, *mupP*, *macpE* cloned *Kpn*I/*Afl*II to pMMH6	Section 3.2.3.1
pJC128	7.1	Genes *mupM* and *mupN* AT cloned into pGEM-T	Section 3.2.3.2
pJC129	8.1	Genes *mupL*, *mupM*, and *mupN* AT cloned into pGEM-T	Section 3.2.3.2
pJC132	24.9	Genes *mupM* and *mupN* cloned *Xba*I/*Bsa*I to pJC124 cut *Xba*I/*Not*I. Expression vector carrying *mupO, P, macpE, U, V, C, F, M, N*	Section 3.2.3.2
pJC133	26.0	Genes *mupL*, *mupM*, and *mupN* cloned *Xba*I/*Bsa*I to pJC124 cut *Xba*I/*Not*I. Expression vector carrying *mupO, P, macpE, U, V, C, F, L, M, N*	Section 3.2.3.2
pJC133.H256A	26.0	Plasmid pJC133 with point mutation H256A in *mupL*	Section 3.2.4.2

2.1.2 Bacterial Growth

All *Escherichia coli* strains and *Bacillus subtilis* were routinely incubated at 37 °C, and *Pseudomonas fluorescens* at 30 °C. Liquid cultures were incubated with shaking at

200 rpm.

2.1.2.1 Media

Lysogeny broth (LB)

LB was used for the general culturing of *E. coli*, *B. subtilis*, and *P. fluorescens* strains. To prepare LB, 10 g bactotryptone, 5 g yeast extract, and 10 g NaCl were mixed. The pH was then adjusted to 7.0, and the mixture made up to 1 L dH_2O. LB was prepared by Central Services (School of Biosciences, University of Birmingham), and was autoclaved before use. For a typical liquid culture, 5 ml of LB in a sterile glass universal bottle was inoculated with bacteria.

Super optimal broth (SOB)

SOB was used for the recovery of bacteria following electroporation and chemical transformation where high efficiency was required. To prepare SOB, 20 g tryptone, 5 g yeast extract, 0.5 g NaCl, and 10 ml of 250 mM KCl were mixed. The pH was then adjusted to 7.0, and the mixture made up to 1 L dH_2O, then stirred and autoclaved. Before use 1 ml of 1 M $MgCl_2$ was added per 100 ml SOB (Hanahan, 1983).

Super optimal broth with catabolite repression (SOC)

SOC was used for the recovery of bacteria following electroporation and chemical transformation, where high efficiency was required. SOC is SOB with added glucose; to generate SOC, 1 ml of 40% (w/v) glucose was added per 100 ml SOB before use (Hanahan, 1983).

Secondary stage media (SSM)

SSM was used for mupirocin production cultures prior to HPLC analysis. To prepare SSM, a mixture of 25 g soya flour, 2.5 g spray dried corn liquor, 5 g $(NH_4)_2SO_4$, 0.5 g $MgSO_4.7H_2O$, 1 g Na_2HPO_4, 1.5 g K_2HPO_4, 1 g KCl, and 6.25 g $CaCO_3$ was generated. The pH of the mixture was adjusted to 7.5, and made up to 1 L dH_2O

and autoclaved. Before use 40% (w/v) glucose was added to the SSM to give a final concentration of 4%.

Mupirocin production media (MPM)

MPM was used for *xylE* assays. *P. fluorescens* yields less mupirocin on MPM than SSM. To prepare MPM, a mixture of 2.3 g yeast extract, 2.6 g Na_2HPO_4, 2.4 g K_2HPO_4, 5 g $(NH_4)_2SO_4$ was generated. The pH of the mixture was adjusted to 7.5, made up to 1 L dH_2O, and autoclaved. Before use 550 μl of 40% (w/v) glucose was added per 200 ml MPM media.

Marine broth

Marine broth was used for culture of *Pseudoalteromonas sp.* SANK 73390. To prepare marine broth a mixture of 33.33 g sea salt, 5 g bacteriological peptone, 0.1 g ferrous sulphate, and 0.8 g yeast extract was generated. The mixture was made up to 1 L dH_2O, and was autoclaved before use (Shiozawa et al., 1993).

M9 minimal media

M9 minimal media was used to select for *P. fluorescens* against a background of *E. coli* S17-1 in pAKE604 mating experiments. To prepare this media, a mixture of 200 ml 3 % agar, 200 ml M9 salts, 400 μl of 1M $MgSO_4$, 400 μl of 1M thiamine HCl, 400 μl of 100 mM $CaCl_2$, and 2 ml of 40 % glucose was generated (El-Sayed et al., 2003). The 3 % agar, M9 salts, $CaCl_2$, and 40 % glucose were individually autoclaved prior to use. The 1M $MgSO_4$ and the 1M thiamine HCl were filter sterilised before use.

Agar

To generate solid media, 15 g of agar was added per litre of the relevant broth recipe.

2.1.2.2 Antibiotics and other media additives

Table 2.3: Antibiotics used in this work

Antibiotic	Dissolved in	Stock concentration (mg ml^{-1})	Working concentration (μg ml^{-1})
Ampicillin	dH_20	100	100
Chloramphenicol	100 % EtOH	35	35
Kanamycin	dH_20	50	50
Streptomycin	dH_20	50	50
Tetracycline	70 % (v/v) EtOH	20	20

2.1.2.3 Other media additives

Isopropyl β-D-1-thiogalactopyranoside (IPTG)

IPTG was used for the induction of the *tac* and *lac* promoters. IPTG was filter sterilised and stocked at 1 M in dH_2O. The typical working concentration of IPTG used in experiments was 0.5 mM.

X-Gal

X-gal was used for blue-white screening through detection of disrupted *lacZ*α when cloning into pUC18, pGEM-T, pAKE604, pJC70 and their derivatives. X-gal was stocked at 20 mg ml^{-1} in dimethyl formamide. A working concentration of 20 μg ml^{-1} was used.

2.2 Manipulation of DNA

2.2.1 Transformation

Competent cells

To generate chemically competent cells, 250 μl of overnight culture was used to innoculate 25 ml of LB medium. The culture was incubated at 37 °C 200 rpm until it reached an OD_{600} of 0.4 to 0.6. Cultures were cooled on ice for 5 minutes, before centrifugation at 3,000 x g, 4 °C for 7 minutes. The supernatant was discarded and pellet resuspended in 10 ml of ice cold 100 mM $CaCl_2$. The resuspended pellet was incubated on ice for more than 20 minutes before centrifugation at 3,000 x g, 4 °C for 7 minutes. The supernatant was then discarded and the pellet gently resuspended in 2.5 ml of ice cold 100 mM $CaCl_2$.

Transformation

For transformation, between 0.5 μl and 5 μl DNA was added to 100 μl competent cells, and then the mixture was incubated on ice for 30 minutes. The mixture was heat shocked at 42 °C for 90 seconds, before the addition of 1 ml LB or SOC, and recovery at 37 °C for between 1 and 2 hours. Cells were then plated onto media with suitable antibiotics to select transformants.

2.2.2 Conjugation

Conjugation was used to deliver plasmids containing *oriT*, such as pJH10 and pJC70, to *P. fluorescens*. The plasmids were first transformed to *E. coli* S17-1, which has the RP4 *tra* regions integrated into the chromosome (Simon et al., 1983).

For conjugation experiments expected to occur with high efficiency, such as transfer of pJH10 derivatives, 10 μl overnight *E. coli* S17-1 culture carrying the *oriT* plasmid was spotted with 10 μl recipient overnight culture onto the L-agar. The mix of cells was grown for between 2 and 16 hours at conditions suitable for the recipient strain.

An inoculum was then taken from the mixed spots and streaked to single colonies on medium selective for both recipient and plasmid. For transfer of pJH10-based plasmids to *P. fluorescens* tetracycline (for the plasmid) and ampicillin (for *P. fluorescens*) was used. If donor contamination was suspected then a further streak onto selective media was performed.

Conjugations of mutagenesis plasmids, such as pJC70 derivatives, were performed using an alternative method, mixing more cells. This is necessary as plasmids need not only to conjugate to donor cells, but also integrate into the target sequence by homologous recombination as well. For such conjugations, 1 ml of *E. coli* S17-1 culture with *oriT* containing plasmid was mixed with 1 ml of recipient culture. The mixed culture was pushed through a sterile 0.45 μl milipore filter, such that the cells were concentrated on the surface of the filter. The filter was then placed on an L-agar plate, and incubated at conditions suitable for the recipient. After incubation overnight, the filters were placed into universal flasks with 1 ml of 0.85 % saline, and cells resuspended by vortexing. Serial dilutions to 10^{-5} in 0.85 % saline were performed, and spread onto solid medium selective for both recipient and plasmid.

2.2.3 Plasmid miniprep

The ISOLATE II Plasmid Mini Kit purchased from Bioline Reagents Ltd. was used for small scale plasmid extractions from overnight cultures, following the included protocol.

2.2.4 Gel electrophoresis

Tris, acetate, EDTA buffer (TAE)

TAE buffer was used for the preparation of agarose gels and as the running buffer for agarose gel electrophoresis. Firstly, 50 ml of 0.5 M ethylenediaminetetraacetic acid (ETDA) was made up: 9.306 g EDTA free acid and 1 g NaOH were dissolved in 40

ml dH_2O, pH adjusted to 8.0 with NaOH, and made up to 50 ml with dH_2O. A stock of 50x TAE was prepared: 121 g tris base, 29 ml glacial acetic acid, 50 ml of 0.5 M EDTA were mixed and made up to 500 ml dH_2O.

Agarose gel preparation

Agarose gels were prepared between 0.5% and 2% (w/v) agarose depending on DNA fragment size to be resolved. In the gels pictured in the results sections, 1% (w/v) agarose was used unless otherwise stated. The agarose was added to a 250 ml flask with the required amount of 1x TAE buffer for desired gel thickness (typically 100 ml). The mixture was heated in the microwave until the agarose dissolved, cooled to hand temperature, and 2 μl of 10 mg ml^{-1} ethidium bromide per 100 ml gel added.

Running agarose gels

Gels were submerged in 1x TAE in tanks. DNA samples were prepared by adding 6x loading dye from either NEB Ltd. or Thermo Fisher Scientific Inc. For a typical sample, a suitable volume of DNA was made up to 10 μl with dH_2O, 2 μl 6x loading dye added, and the solution mixed by pipetting up and down. To run a larger volume of DNA (x μl), typically for purification, $x/5$ μl of 6x loading dye was added. For size verification, 5 μl of DNA ladders were used, typically GeneRuler 1 kb from Thermo Fisher Scientific Inc. Where small DNA fragments were visualised, the 100 bp ladder from NEB Ltd. was used.

2.2.5 DNA purification by gel electrophoresis

For size purification of DNA fragments, DNA was run on an agarose gel and desired band cut out with a razor under UV light. DNA was purified from the gel using the Illustra GFX PCR DNA and Gel Band purification kit from GE Healthcare, following the included protocol.

2.2.6 Ethanol precipitation

For the purification of DNA from a 20 μl solution, 5 μl of 3 M NaOAc and 110 μl of 100% EtOH were mixed, and incubated at -20 °C for 20 minutes. The DNA pellet was collected by centrifugation at 13,000 g for 15 minutes, and washed with 70% ethanol. The pellet was dried at 48 °C for 20 minutes and resuspended in sterile distilled water (SDW).

2.2.7 Restriction Digests

Restriction enzymes were purchased from NEB Ltd. Digests were performed using the recommended NEB protocol, typically in a 20 μl or 50 μl reaction volume with incubation for 1 to 3 hours. For restriction digest by *Xba*I, plasmids were propagated through *dam* negative strain *E. coli* ER2925 where the sequence indicated methylation would block digestion.

2.2.8 Ligations

DNA was typically prepared for ligation by digest and size purification (Section 2.2.5). Approximately 50 ng of vector DNA was mixed with insert DNA to give an approximate molar ratio of 1:3. Ligations were performed using T4 DNA ligase from Invitrogen or NEB and incubated overnight at 4 °C, before transformation.

2.2.8.1 Cloning to pGEM-T-Easy

Fragments were cloned in pGEM-T-Easy for sequencing and propagation of DNA, prior to subcloning to other vectors. An advantage pGEM-T-Easy offers over direct cloning is that fragments gel purified from pGEM must be cut by restriction enzyme on both sides. With direct cloning, typically 3 to 5 nucleotides were included in primers such that when digested, around 6 to 8 nucleotides would be cleaved from either flank of the product. Therefore digestion of the PCR product would not normally

yield a visible size change by agarose gel electrophoresis, as products were often of the order of kilobases. Cloning in pGEM-T-Easy also allows for identification and quality control of the desired fragment by sequencing, after which it can be subcloned into the suitable vector.

Gel purified (Section 2.2.5) DNA was A-tailed using *Taq* polymerase from Invitrogen (Table 2.4). A-tailed DNA was ligated to the linear pGEM-T-Easy using the included T4 ligase following the manufacturer's instructions.

Table 2.4: Mix for A-tailing PCR products

Component	**Volume** (μl)
10x PCR buffer	1
DNA (typically PCR product)	6
50 mM $MgCl_2$	1
2 mM dATP	1
Taq polymerase	1

2.2.9 Molecular design software

Table 2.5: Software used for visualising and design of molecular approaches

Software	Description of use	URL
Chromas lite	DNA sequence chromatogram visualisation.	http://technelysium.com.au/wp/chromas/
Jpred	Protein secondary structure prediction.	http://www.compbio.dundee.ac.uk/jpred/
Ligplot+	Visualisation of ligand-protein interactions.	https://www.ebi.ac.uk/thornton-srv/software/LigPlus/
Netprimer	Prediction of PCR primer Tm's and quality.	http://www.premierbiosoft.com/netprimer/
OriginPro 2016	Generation of graphs. HPLC chromatogram baseline correction and peak integration.	http://www.originlab.com/
Psipred	Protein secondary structure prediction.	http://bioinf.cs.ucl.ac.uk/psipred/
SeaView	Visualisation of DNA and protein alignments.	http://doua.prabi.fr/software/seaview
Snapgene Viewer	Visualisation of plasmid DNA sequences and molecular designs. Generation of plasmid maps.	http://www.snapgene.com/
WebLogo	Generation of sequence logo diagrams.	http://weblogo.berkeley.edu

2.2.10 DNA sequencing

All DNA sequencing was performed by the Genomics Service at the University of Birmingham, run on a 3730 DNA Analyser. For plasmid sequencing, 200 to 500 ng of template DNA was mixed with 3.2 pmol primer, up to 10 μl in dH_2O, and provided to the Genomics Service. For PCR sequencing, approximately 10 ng of template DNA was used, following the guidelines from the Genomics service. Chromatograms were visualised using Chromas (Table 2.5).

2.2.11 PCR

Preparation of template DNA

For PCRs from chromosomal template DNA, a boil preparation was performed. An area of approximately 1 cm^2 of culture was scraped from the surface of solid media, and resuspended in 50 μl dH_2O. The mixture was boiled at 100 °C for 10 minutes, before centrifugation for 3 minutes at 11,000 g. The supernatant was used as template DNA for PCR. Where suitable, plasmid or PCR product template DNA was used instead.

High fidelity

All PCRs destined for cloning were performed using high fidelity enzymes. These proofreading enzymes have 3′ to 5′ exonuclease activity that can excise incorrectly incorporated nucleotides. The majority of such PCRs were performed using Velocity DNA polymerase from Bioline Reagents Ltd., while a minority were performed using Q5 DNA polymerase from NEB Ltd. PCR reactions were performed to the respective manufacturers protocols. Melting temperatures were estimated using Netprimer (Table 2.5).

Low fidelity

The non-proofreading *Taq* polymerase from Invitrogen was used for applications where fidelity was not required, such as screening for inserts or mutations. Mix and cycling conditions are listed in Table 2.6. Alternatively, GoTaq DNA polymerase from Promega was used for screening potential mutants, following the manufacturer's protocol using the green buffer with DNA loading dye included.

Table 2.6: PCR with Invitrogen *Taq*

(a) Mix for a 50 μl reaction

Component	Volume (μl)
10x PCR buffer	5
Template DNA	5
50 mM $MgCl_2$	1
10 mM each dNTPs	1
50% glycerol	10
100 μM primer 1	0.3
100 μM primer 2	0.3
Taq polymerase	1
dH_2O	up to 50μl

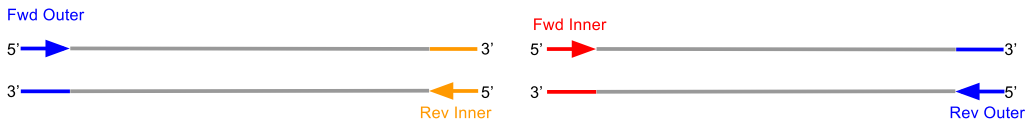
(b) Cycling conditions

Step	Time (seconds)	Temp. (°C)	Cycles
Initial denaturation	120	94	1
Denaturation	15	94	
Annealing	30	Tm	30
Extension	60 per kb	72	
Final extension	300	72	1

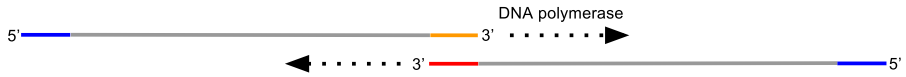
2.2.11.1 Overlap extension PCR

Overlap extension PCR was used broadly in this work to join DNA segments, whilst introducing designed changes such as point mutations, deletions or insertions. The method is explained in Figure 2.1.

(a) Initial PCR amplification.



(b) Filling in to generate full length template DNA.



(c) Amplification of full length product.

Figure 2.1: Overlap extension PCR, used to generate primer dictated modifications to DNA.
(a) Two PCR products are amplified using pairs of primers. The Inner primers, shown in red and orange, are designed to overlap such that they are complementary to each other. Design of this overlap is flexible, and can introduce new insertions, deletions or point mutations.
(b) The PCR products are purified to remove any unincorporated primers and any non-specific products. They now act as template for each other in a second, overlap extension PCR. Typically, 3 PCR cycles were performed without primers, this allows the polymerase to extend the 3' end of DNA fragments using the other fragment as template, shown by the dotted arrow.
(c) Now full length product with the desired modification is available in the reaction mix. The outer primers are added, and 27 PCR cycles performed to amplify the full length product. This is typically performed at a higher Tm than possible for the overlap, which is why the primers are not included initially.

2.2.11.2 Colony PCR

Colony PCR was used to screen colonies for presence of insert where a large proportion of religated empty vector was suspected, and blue white screening not available. A PCR master mix containing all components was made up (without template DNA), and dispensed into 20 μl aliquots. Each colony to be screened was picked up by pipette tips, and suspended into a PCR aliquot by pipetting up and down.

To propagate live cells from each colony, 1 μl of each PCR aliquot was transferred to 100 μl LB. PCR was performed and products checked by agarose gel electrophoresis. Where the desired size insert was present, the 100 μl LB storage cultures were used to inoculate 5 ml LB. These were incubated overnight for plasmid preparations.

2.2.12 Suicide mutagenesis

Construction of plasmids

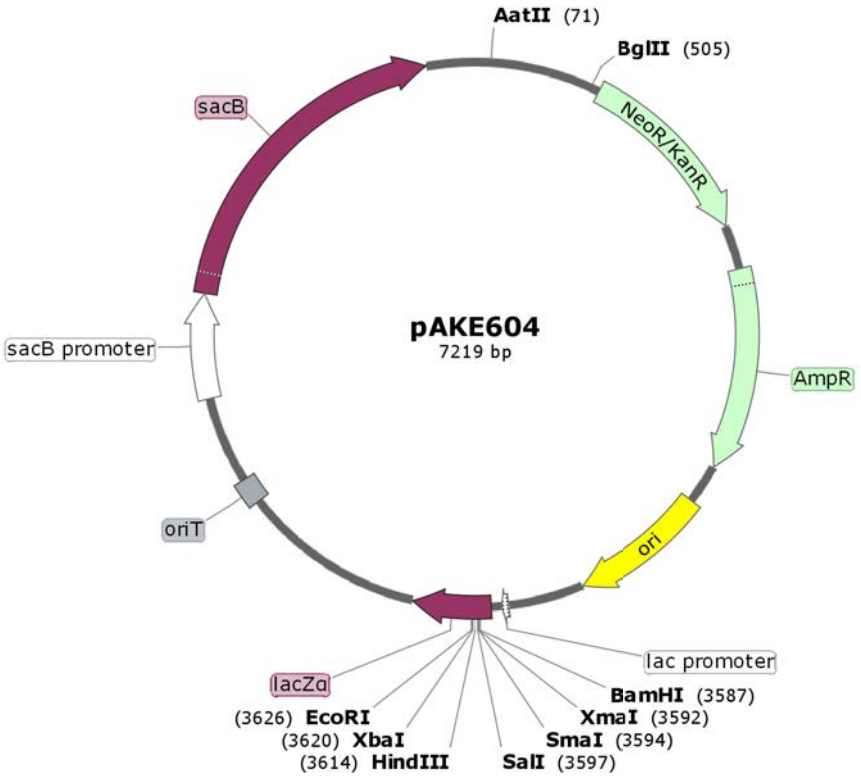
Primers were designed to generate two PCR fragments or arms with the desired overlap, such that they could be stitched together by overlap extension PCR (Section 2.2.11.1). Each of these arms was typically around 500 bp long. The design of overlapping primers defined the mutation to be made. In the scenario of a point mutation, the overlapping primers are designed to carry the desired point mutation.

Overlapped mutagenesis products were cloned into the multiple cloning site of pJC70. Mutagenesis vector pJC70 (Figure 2.2b) was derived in this work from pAKE604 (Figure 2.2a), through the use of PCR to remove the majority of the ampicillin resistance gene. The advantage of pJC70 is it allows selection of *P. fluorescens* through its chromosomally encoded Amp^R after conjugation. With pAKE604 this is not possible, as the plasmid encoded *bla* gene renders the donor *E. coli* Amp^R. Selection of *P. fluorescens* is instead achieved using minimal media, where the bacterial colonies can take 2-3 extra days incubation to grow to a suitable size.

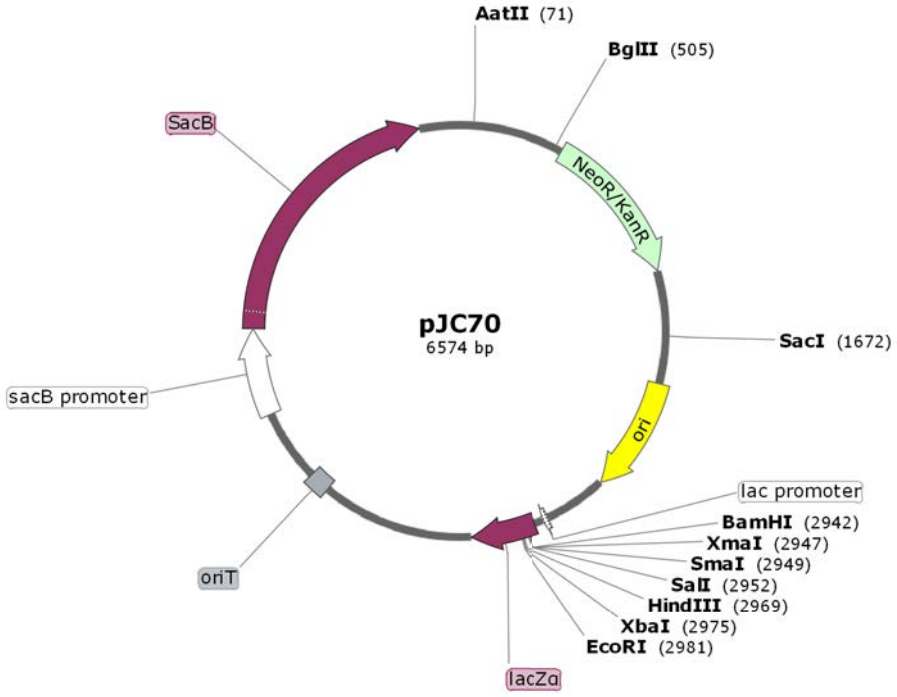
Generation of mutations

Derivatives of pJC70 carrying a mutagenesis insert derived from overlap extension PCR were transformed to *E. coli* S17-1. The vector includes the RK2 *oriT* which allows RK2 mediated conjugation from S17-1 to *P. fluorescens*. Transconjugants were selected for using kanamycin and ampicillin, which select for the plasmid and recipient *P. fluorescens* respectively.

The vector is referred to as suicide as its pMB1 replicon is not functional in *P.*



(a)



(b)

Figure 2.2: Maps of suicide mutagenesis vectors pAKE604 and pJC70.

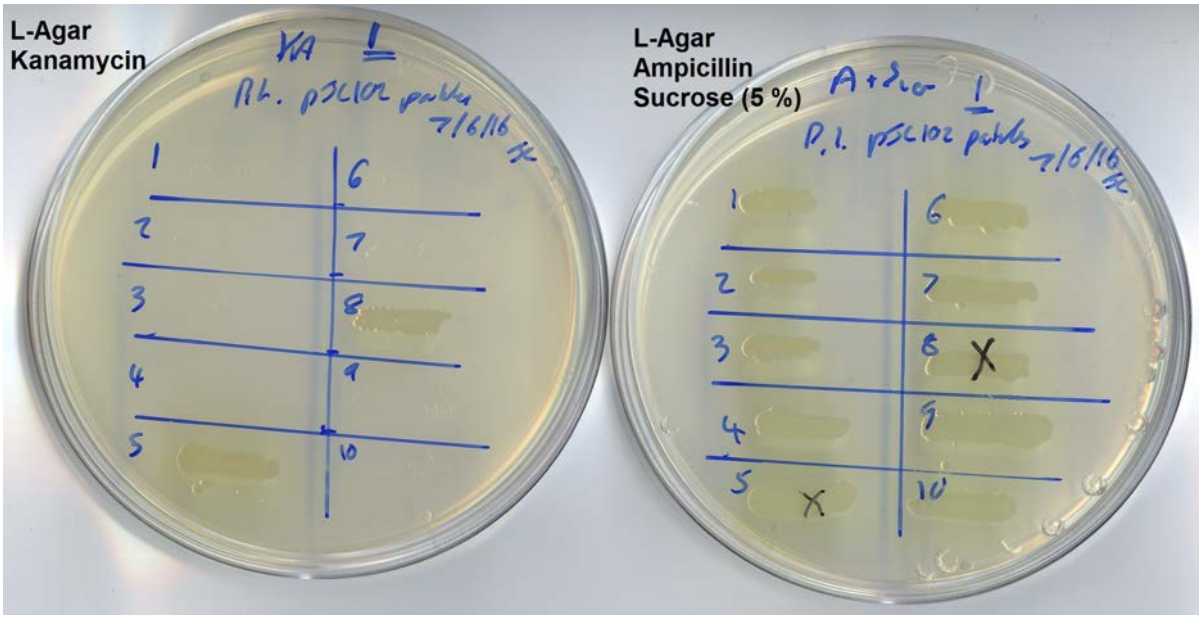


Figure 2.3: Example patch plates to check for kanamycin resistance.
Each streak is derived from a single sucrose resistant colony, that are putative mutatgenesis plasmid excisants. Isolates 5 and 8 are both kanamycin and sucrose resistant, this suggests integration of the suicide plasmid and inactivation of the *sacB* gene rather than excision.

fluorescens. Therefore kanamycin selection forces integration of the plasmid into the chromosome, mediated by homologous recombination in one of the 500 bp arms. *P. fluorescens* strains with chromosomal pJC70 integrants were grown overnight in LB without kanamycin selection. This allows a second recombination to occur, where the plasmid excises.

The vector includes *sacB*, which encodes levansucrase from *Bacillus subtilis*, which confers a lethal periplasmic sucrose polymerisation in gram negative bacteria (Gay et al., 1985). Serial dilutions of these cultures were plated onto L-agar 5% sucrose to select for strains that had lost *sacB*. Sucrose resistant colonies were then patched onto L-agar with and without kanamycin, to check for plasmid loss (Figure 2.3). Typically around 5 to 10 % of sucrose resistant strains were observed to remain kanamycin resistant, indicative of *sacB* inactivation rather than plasmid excision.

If the second recombination occurs in the arm alternative to that of the first recombination, then the desired mutant sequence is left in the chromosome. Alternatively,

if the second recombination event occurs in the same arm as the first, then the WT sequence remains. Both scenarios should occur with equal frequency, and in both excised plasmid is non-replicative and lost through dilution. Mutant strains were then detected by PCR, restriction digest and sequencing.

2.3 Assays

2.3.1 High performance liquid chromatography

Sample preparation

P. fluorescens 5 ml LB seed cultures, with suitable antibiotics, were inoculated in triplicate from single colonies on solid media, and incubated for 16 hours at 30 °C. Seed culture (200 μl) was used to inoculate 5 ml SSM media (Section 2.1.2.1) in 50 ml conical flasks. Suitable antibiotics were used to maintain any plasmids, and IPTG added to the SSM medium where required (typical final concentration 0.5 mM). In some later experiments, prednisone was used as an internal standard (typical final concentration 50 μM). Prednisone was added to the flasks after incubation, such that any discrepancies in HPLC sample preparation could be detected by comparison of prednisone peaks on the chromatograms. SSM cultures were incubated at 22 °C 200 r.p.m. for between 40 and 64 hours.

Samples were centrifuged at 13,000 g and supernatants collected. The supernatants were filtered through 0.2 μm millipore filters prior to injection.

HPLC analysis

HPLC was performed on a Gilson system using a reverse phase C18 column (15 cm x 4.6 mm). Mobile phase ran on a water:acetonitrile gradient, starting 95:5 (t = 0 minutes) to 30:70 (t = 30 minutes). Both the HPLC grade water and acetonitrile were made up with 0.01 % formic acid. Compounds were detected by UV absorption at 233 nm.

Peak integration

To generate figures and quantify peak area, raw data was exported from the Gilson software and imported to OriginPro 2016. The OriginPro Peak Analyzer tool was used to determine peak area, with asymmetric least squares smoothing to determine baseline.

2.3.2 Plate bioassay

Plate bioassays were used to determine the antibiotic activity of a strain, through measuring the area of inhibition of growth of susceptible *B. subtilis* 1604. Overnight cultures of *P. fluorescens* were grown with suitable antibiotics; 10 μl of these were spotted onto 20 ml measured L-agar plates. If IPTG induction was required it was included in the L-agar. Plates were incubated for 16 hours at 20 °C.

B. subtilis overnight culture (8 ml) was mixed with 1 ml 5 % triphenyl tetrazolium chloride (TTC) in 200 ml L-agar. A variety of dehydrogenase enzymes in live cells can reduce the white TTC to red formazan (Brodie and Gots, 1952; Rodriguez et al., 1992). *B. subtilis* is a mupirocin sensitive gram positive; when stained with TTC the zone of inhibition of mupirocin and any other antimicrobials produced by *P. fluorescens* can be detected. Each plate was overlaid with 15 ml of the *B. subtilis* containing mixture. Plates were incubated for 16 hours at 37 °C, and diameters of zones of inhibition measured.

2.3.3 Chromogenic *xylE* reporter assay

XylE reporter assays were used to determine the approximate level of *mupA* expression. These used the reporter strain 10586 Δ*mupI mupA::xylE*, which has *xylE* inserted downstream of the putative *mupA* promoter (El-Sayed et al., 2001). To perform the assays, overnight LB cultures of the reporter and relevant test strains were streaked onto MPM agar plates (Section 2.1.2.1), which were incubated for 22 hours

at 20 °C. Plates were then mist-sprayed with 1% catechol (w/v in H_2O), and colour change observed after 5 minutes. Colourless catechol is converted to yellow 2- hydroxymuconic semialdehyde, catalysed by XylE (Zukowski et al., 1983), allowing for detection of expression from the *mupA* promoter.

2.3.4 Ethyl acetate extraction of PA-B

Ethyl acetate extraction was used to concentrate PA-B for use in feeding experiments. This was adapted from a method for chlorothricin extraction, experienced first hand in the Manuela Tosin lab at the University of Warwick (Parascandolo et al., 2016). Plasmid pJH2, which carries *mupR* was mobilised to PA-B producing strain 10586 Δ*mupU*, as *mupR* expression increases metabolite yield (Hothersall et al., 2011). Seed cultures were grown overnight at 30 °C in LB. SSM cultures (200 ml) in 2 L conical flasks, were inoculated with 2 ml saturated seed culture, and incubated at 22 °C for 40 to 44 hours.

Cultures were aliquoted into 50 ml Falcon conical centrifuge tubes and centrifuged at 3000 xg for 10 minutes at room temperature. The supernatants were collected and pH adjusted to 4.5, before being split into aliquots in 50 ml Falcon tubes, approximately half full (25 ml). Each tube was topped up with approximately 25 ml ethyl acetate (EtOAc), mixed thoroughly by inversion and vortexing approximately ten times over half an hour. Where emulsions were formed, these were resolved to layers by centrifugation (3000 x g for 5 minutes). The clear top layer was collected using a glass pipette. Each tube was refilled with EtOAc, and the process repeated, again collecting the top layer. The collected solvent was evaporated on a rotovap, until less than 500 μl oily yellow liquid remained. Methanol (2 to 4 ml) was added and any precipitated solids dissolved by pipetting. The purity and quantity of the extract was appraised by HPLC.

2.3.5 Determination of minimum inhibitory concentration of mupirocin

LB (100 μl) was aliquoted to all wells of 96-well microplates (12 columns by 8 rows). A stock of mupirocin was prepared at 200 mM in DMSO from commercial powder. This was diluted to 16 mM mupirocin in LB in a petri dish, and 100 μl was added to the wells of column 1, diluting them to 8 mM mupirocin in LB. Serial dilutions were performed by pipetting 100 μl from each column to the next, with pipetting up and down to mix. This results in 100 μl of LB with mupirocin dilutions in columns 1 to 11, and 200μl of LB 4 μM mupirocin in column 12; 100 μl was removed from column 12 and discarded.

Saturated overnight cultures (grown 16 hours in LB) were diluted 1 in 50 in a petri dish, 9.8 ml LB with 200 μl culture. These diluted cultures were used to inoculate the microplates, adding 100 μl 1 in 50 culture dilution to each row. This resulted in an effective 1 in 100 dilution of each culture in LB, with mupirocin concentrations ranging from 4 mM to 2 μM across the columns.

Plates were incubated in a shaking incubator at 200 RPM with the lid on, at 37 °C for *E. coli* and 22 °C for *P. fluorescens*. After 16 hours, bacterial growth was both appraised by eye, and OD_{600} absorbance readings recorded on a BMG Labtech SPECTROstar Nano plate reader, adjusted for a non-incolulated blank row.

2.3.6 Determination of colony forming units

To determine the number of colony forming units (CFUs) within *P. fluorescens* SSM cultures, serial dilutions were performed in 1.5 ml eppendorf tubes. Saline (0.85%, 900 μl) was aliquoted to each tube, 100 μl SSM culture was added to the first, mixed by vortex, and serial 100 μl dilutions were performed to 10^{-6}. L-agar plates were divided into six sections, and 20 μl of each dilution deposited in each section as microdrops (Thomas et al., 2015). Colonies were counted after incubation at 30 °C for 16 hours, an example plate is pictured in Figure 2.4.

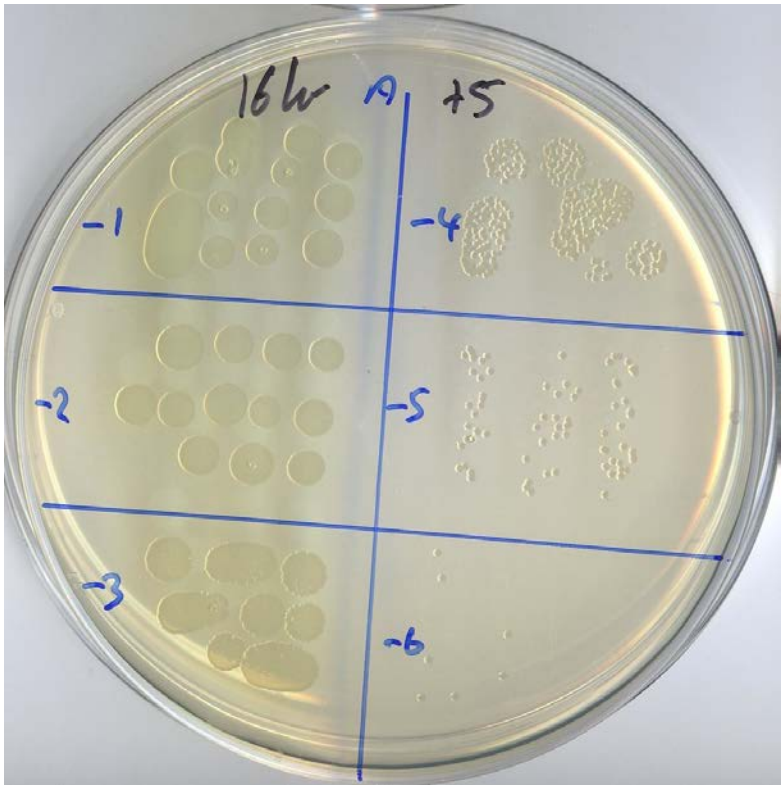


Figure 2.4: Example CFU determination plate.
Each section contains 20 μl of diluted culture, deposited as microdrops.

2.4 PCR primers

All primers used in this work were purchased from Altabioscience Ltd, and are listed by chapter in the following tables.

Table 2.7: Primers used in Chapter 3

Name	Sequence (5′ to 3′)	Purpose
M13 (-40)	GTTTTCCCAGTCACGAC	Sequencing of inserts in *lacZα* cloning sites
M13 (-48)	AGCGGATAACAATTTCACACAGGA	
F outer Y167F MupV	AGACAATTGAGCAGGCCCGTCATTTAC	Construction of MupV Y167F mutagenesis vector
R inner Y167F MupV	TGGTTTCTTCGAAATAGTTATTGAATTCAGTGC	
F inner Y167F MupV	CAATAACTATTTCGAAGAAACCAAATGTGCG	
R outer Y167F MupV	AGAGGATCCGTTGTGGATGTAGGGGTTGTAAAAC	
P17 F outer H631A MupV SalI	GGAATCGTCGACAAAGCACATCTGATCGGCT	Construction of MupV H631A mutagenesis vector
P17 R inner H631A MupV	CAAACCAAATGCATCCATGCCAGCGATTTC	
P17 F inner H631A MupV	GGCATGGATGCATTTGGTTTGTATAACCACCTGG	
P17 R outer H631A MupV EcoRI	GCACTGGAATTCTGACGACAACTGTTGTAGTAGGC	
P25 F mupOP KpnI SpeI	GAACGGTACCTACTAGTGGGAGAGCAACGATGACATC	Cloning of *mupO*, *mupP* and *macpE*
P25 R mupOP overlap	GCCTCCTGTGGTGTGCTCCTGTTGCTGTTTG	
P25 F macpE overlap	ACAGGAGCACACCACAGGAGGCGTAGATGC	
P25 R macpE AflII	AGACTTAAGTCACTGCTCACGTTGGGC	
P25 F mupLMN	CCACTCTAGACAGTGAAGAGAGGGACACACAGC	Cloning of *mupL*, *mupM* and *mupN*
P25 R mupLMN	GGTCTCAGGCCGCGATCCAAACCTATAAGCACTG	
P25 F mupMN	TAGCTCTAGACTGACAGGTGTGACTGATGAGTACG	

Continued on next page

Table 2.7 – *Continued*

Name	**Sequence (5′ to 3′)**	**Purpose**
P31 F mupL H256A	GATTGGGGGGCCTACACGCTTTTCTCTGATACACAAGAG	Generation of *mupL*
P31 R mupL H256A	AGCGTGTAGGCCCCCCAATCGTCCATGATG	H256A in pJC133
mupL1F	GGATCCGCGTTCTACAGCAGCACAC	Screening for *mupL*
mupL2R	GAATTCCTTGAGCACCAGCTCTCTG	deletion
mupU1F	GGATCCGACCAGATCGTCCAGTGGTTC	Screening for *mupU*
mupU2R	GAATTCGTGCTCTGGTGCAGGCATTC	deletion

Table 2.8: Primers used in Chapter 4

Name	Sequence (5′ to 3′)	Purpose
M13 (-40)	GTTTTCCCAGTCACGAC	Sequencing of inserts in
M13 (-48)	AGCGGATAACAATTTCACACAGGA	*lacZα* cloning sites
MmpBF1	CTGGAATTCATGGTCAAAGACTTCGACAGC	Cloning of *mmpB*
MmpBF2	GAGTCGATGAGTTCGATCCGC	
MmpBF3	GCACCTACACACTCGAACTGG	
MmpBR1	GCGGATCGAACTCATCGACTC	
MmpBR2	CTGTCTAGATCAGAACTCGCCTGCGCGTC	
MmpBR3	TCGGATGCTCAAACAGCG	
MmpBF5	TCCACGCAGCTCAATCGC	Sequencing of pJH10/MmpB
MmpBF6	ACGACAGATCACGGCAGG	
MmpBF7	GTGGCAGCTTTGCCTTCG	
MmpBR4	ACAATTGACACACCAATGCC	
TmpBF1	GGCACATGTCGGAGACCCTACGTCTGAGCCATC	Construction of MmpB/TmpB hybrid I
TmpBF2	TGGCGGGGATCAATGCATTTGGTTTTGGTGGCGTCAATGCGCATGTCATC	
TmpBF3	GGCAAAATATGTATCAGGGAGAAAC	
TmpBR1	CTCTCTAGATTACTCATACGTCACCTCATCGCTA	
TmpBR2	AAATGCATTGATCCCCGCCATTCTTGGTCTGTTATCCAGAGTCGGCTC	

Continued on next page

Table 2.8 – *Continued*

Name	**Sequence (5′ to 3′)**	**Purpose**
TmpBR3	CCCTGATACATATTTTGCCATGC	
F pJC46 ACPfrag	TCTTTGGAATTGCTGCTATGGC	
R pJC46 ACPfrag	TCCCGGTGCATCTTCATCTG	Construction of
F mmpB full TE	CAGATGAAGATGCACCGGGAGATTGGCGTTCGCTGGCG	MmpB/TmpB hybrid
F mmpB short TE	CAGATGAAGATGCACCGGGAGGGCTCTTCTGTTTCCCCG	generation II
R mmpB TE	ACATCTAGAAGATCAGAACTCGCCTGCGC	
pJH10 Seq F	GCCGACATCATAACGGTTCTGGC	Sequencing of inserts
pJH10 Seq R	CTCCTGCCAGTTGATGACCTCG	in pJH10
R tmpB C1824A	AACTGCAGCCTCTAGTTCATTGGTCGCTTTGC	Generation of C1824A
F tmpB C1824A	ATGAACTAGAGGCTGCAGTTTCAACCGGCTCACTC	mutant hybrids
P16 F 2.6kb	AATCGCACCTACACACTCGAACTG	
P16 R 2.6kb	GGCGAGTGATTTGTGACGAGTCTTTTGGTTGCTC	
P16 F macpE	CTCGTCACAAATCACTCGCCAGACCGTCCAG	Construction of
P16 R macpE stop SacI	CAGAGAGCTCCTCACTGCTCACGTTGGGCG	MmpB/TmpB hybrid
P16 R macpE TEoverlap	TTATCGCTGTTCTGCTCACGTTGGGCGAACC	generation III
P16 F TE revised	GAACGTGAGCAGAACAGCGATAAAAATGAGCA	
P16 R TE SacI	ATTGAGCTCGTTTTCCCAGTCACGACGTTGTAA	

Continued on next page

Table 2.8 – *Continued*

Name	**Sequence (5′ to 3′)**	**Purpose**
F pJC5x XbaI	CTCTCTAGACTGTTTGAGCATCCGACGC	Construction of the pJC74 series of hybrid 10586 chromosomal mutagenesis vectors
R pJC5x SalI	TGTGGGGAAGACTGCTTGC	
F hybriddownarm overlap	CGTATGAGTAAAAGGTGGAGCGGCAAGGCT	
R pJC50 overlap	GCTCCACCTTTTACTCATACGTCACCTCATCGC	
R hybrid downarm SalI	CAGGTCGACAGCCGCCAATGCAGAGACTGTC	
MmpB 1F Jo	GGATCCGGACCTGGAACAGGTGGTG	Screening for *mmpB* deletion

Chapter 3

INVESTIGATION OF THE CONVERSION OF PA-B TO PA-A IN MUPIROCIN BIOSYNTHESIS

3.1 Introduction

The mixture of pseudomonic acids produced by *P. fluorescens* NCIMB 10586 is comprised of 90% PA-A and approximately 5% PA-B. These molecules differ by one group at carbon 8, -H in PA-A, and -OH in PA-B. This modification is required for antibiotic function, PA-A is active and PA-B is inactive (Figure 3.1). The X-ray structure of PA-A complexed with IleRS (Silvian et al., 1999) revealed 4.2 Å between carbon 8 and the closest IleRS amino acid, His67 (Figure 3.2). With the addition of the hydroxyl, this distance could close to 2.8 Å. It is therefore possible that the C8-OH of PA-B causes a steric hinderance that prevents target IleRS binding.

PA-B is thought to be a precursor to PA-A, with the biosynthetic removal of the C8-OH occurring in a suprisingly convoluted fashion (Gao et al., 2017). The current model was proposed based on reverse genetic approaches and intermediate feeding experiments, which refuted earlier hypotheses of PA-B being generated from PA-A or parallel biosynthetic pathways (Chain and Mellows, 1977; El-Sayed et al., 2003; Cooper et al., 2005).

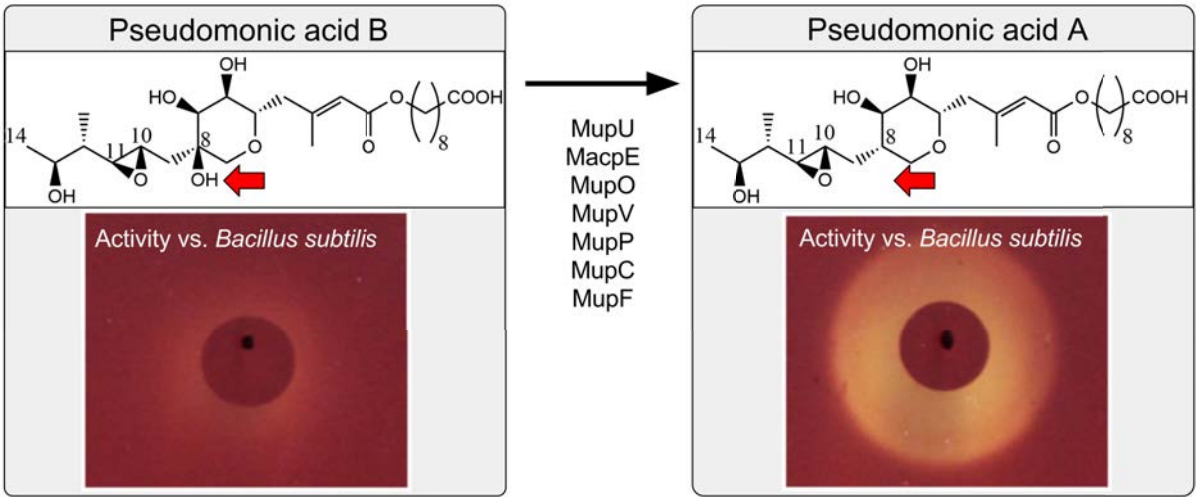


Figure 3.1: Conversion of PA-B to PA-A in mupirocin biosynthesis summary. Processing to PA-A is required for activation of the antibiotic. Example bioassay plates are shown, taken from the experiment in Section 3.2.3.3. Prior to the work presented in this chapter, the seven listed enzymes had been implicated as required for conversion (Gao et al., 2017).

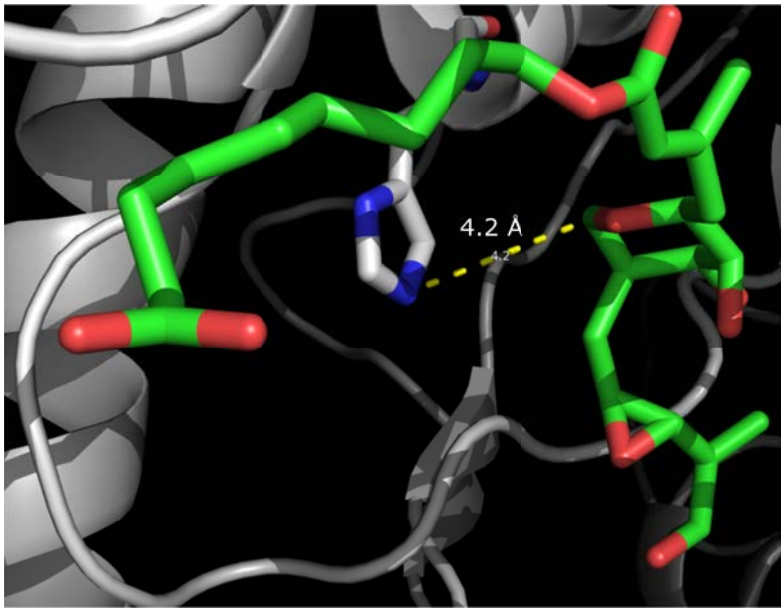


Figure 3.2: Position of PA-A C8 when bound to isoleucyl t-RNA synthetase. The closest amino acid on the *S. aureus* IleRS is His67, which is 4.2 Å from carbon 8 of PA-A. Generated from PDB entry 1QU2 (Silvian et al., 1999). Mupirocin is shown in green, with oxygens in red, IleRS is in grey with nitrogens in blue, and only His67 is shown as sticks for clarity.

3.1.1 Historical evidence and models of the biosynthesis of PA-B and PA-A

After the discovery of mupirocin in 1971 (Fuller et al., 1971), the structure of PA-B was first determined in 1977 (Chain and Mellows, 1977). Chain and Mellows hypothesised that PA-B could be generated by enzymatic hydroxylation of PA-A. Later experiments with radiolabelled acetate did not support this hypothesis; incorporation of radiolabelled precursors occurred differentially between PA-B and PA-A (Mantle et al., 2001). Mantle and colleagues expressed difficulty interpreting their results, with limited information available on the biosynthesis pathway at this time.

This changed with the identification and sequencing of the mupirocin biosynthetic cluster (Whatling et al., 1995; El-Sayed et al., 2003). In this work, the genes and enzymes we now know are involved in PA-B processing were named and analysed by homology searches, but their precise functions were not known. Repeating the hypothesis of Chain, El-Sayed et al. (2003) suggested that PA-B may arise from further

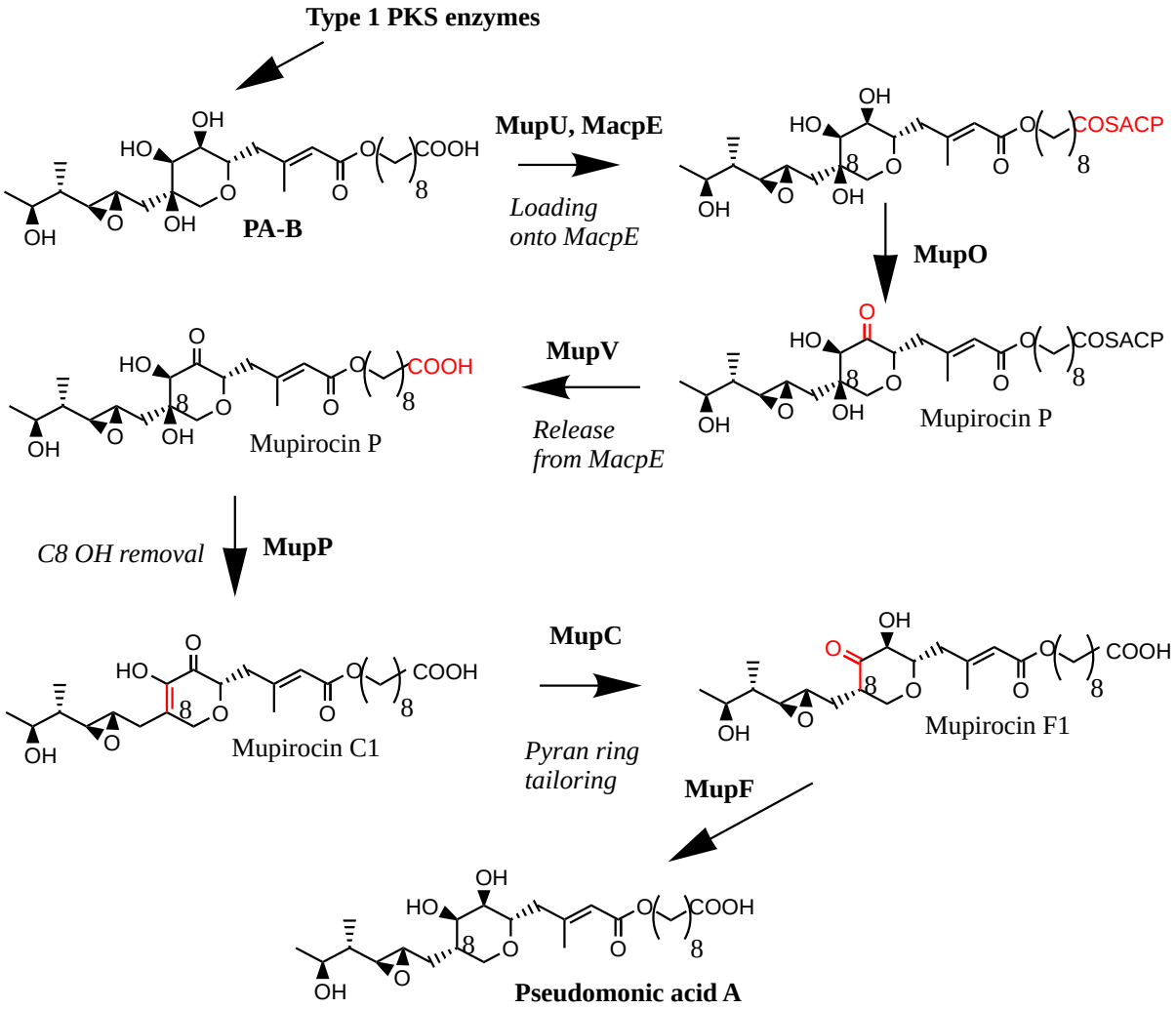


Figure 3.3: Biosynthetic pathway from PA-B to PA-A proposed in Gao et al. (2017).

oxidative modification of PA-A.

A key breakthrough came with the observation that deletion of *mupO*, *mupU*, *mupV* or *macpE* in the producing organism *P. fluorescens* NCIMB 10586 caused the accumulation of PA-B (Cooper et al., 2005). A branched biosynthesis pathway was proposed in Cooper et al. (2005), where action of these enzymes occurs before pyran ring closure, after which the pathways to PA-B and PA-A diverge.

3.1.2 Current model: PA-B is a biosynthetic precursor to PA-A

The current model is outlined in Figure 3.3. Contradictory to the historical hypotheses, PA-B is now proposed to be a precursor to PA-A, first described in Hother-

sall et al. (2007). Late stage intermediates mupirocin C1 and mupirocin F1 were isolated from 10586 Δ*mupC* and 10586 Δ *mupF* respectively. Slight elevation in PA-B production was observed in 10586 Δ*mupP*, but not at the scale of the other PA-B producing mutants, and therefore a role for MupP was not initially proposed in Hothersall et al. (2007).

Intermediate mupirocin P was later detected from 10586 Δ*mupP* by NMR, but proved unstable (Gao et al., 2017). Further evidence for its existence was uncovered when the stable desepoxy mupirocin P was isolated from 10586 *mmpE*ΔOR Δ*mupP* (Gao et al., 2017). The MmpE oxidoreductase domain is thought responsible for generation of the C10,11 epoxide, and therefore 10586 *mmpE*ΔOR produces pseudomonic acid C, a PA-A variant lacking the epoxide. When a double mutant was generated with Δ*mupP*, the stable compound desepoxy mupirocin P was observed.

When 10586 Δ*mupW*, a strain deficient in the earlier pathway that does not produce PA-A, was fed with exogenous PA-B, 35% conversion was observed to PA-A, providing strong further evidence for the current model (Gao et al., 2014). This model also fits the observation of PA-B accumulation in 10586 Δ*mupU*, 10586 Δ*macpE*, 10586 Δ*mupO* and 10586 Δ*mupV* (Cooper et al., 2005).

3.1.3 Objectives

Aim 1: Determine if both domains of MupV are enzymatically active

Bioinformatic re-analysis of MupV has indicated it is a dual domain protein, with an α/β hydrolase in addition to the first oxidoreductase domain previously reported. In-frame deletion of *mupV* in 10586 yields a complete switch to PA-B production (Cooper et al., 2005), but this could be caused by the deletion of either or both of these domains. MupV is proposed to catalyse release of pseudomonic acid from MacpE in PA-B to PA-A tailoring, and the discovery of a second active domain could indicate a secondary function.

Prior to the work described in this chapter, an approach was devised where *mupV*

carrying mutations to either domain would be provided on an expression plasmid to 10586 Δ*mupV*. However, expression of WT *mupV* from plasmid pJH10 proved unable to complement PA-A production in 10586 Δ*mupV* (unpublished work by Stephens, E. and Di Martino, M.).

Given this setback, an alternative approach is described in this chapter, where point mutations to *mupV* were introduced to the 10586 chromosome, to help determine the functionality of both domains.

Aim 2: Develop and perform a systematic screen to identify all genes required for conversion of PA-B to PA-A

As detailed in Section 3.1.1, the models of conversion of PA-B to PA-A have trended towards increased complexity as our understanding has improved. The most recent example of this is the demonstration of the function of MupP (Gao et al., 2017). Whilst the majority of mupirocin biosynthetic enzymes have been assigned functions, there remain a number with unknown functions such as: MupZ (hypothetical), MupA (reduced flavin mononucleotide oxygenase), MacpA (ACP), MacpB (ACP), MupL (hydrolase), and MmpF (ketosynthase).

In this chapter, a systematic screen is described which aimed to determine the ability of mutants of every mupirocin tailoring gene to convert PA-B to PA-A, with the aim of re-inforcing our current pathway model and identifying or eliminating the role of any further cluster genes.

Aim 3: Generate an expression plasmid that provides all functions required for conversion of PA-B to PA-A

As a counterpart to Aim 2, generation of a plasmid that carries all necessary genes could demonstrate that all required functions had been determined. Such a plasmid should restore conversion of exogenous fed PA-B in *P. fluorescens* 10586 Δ*mup-cluster*, a strain with the mupirocin biosynthetic cluster deleted.

This objective has previously been undertaken in the theses of Macioszek (2009)

and Yadav (2017); however plasmid encoded conversion was not achieved, as we now know the generated plasmids were missing required genes. Macioszek (2009) generated pJH10-based expression vector pMMH6, which holds *mupO*, *macpE*, *mupU*, *mupV*, *mupC* and *mupF*, and is able to complement individual chromosomal 10586 knockouts of each gene. However, it is unable to provide all the functions for conversion of PA-B to PA-A, which is expected as *mupP* was not included, and is now thought required. Plasmid pMMH6 is also lacking the resistance gene *mupM*, so it is possible any conversion to PA-A would be self inhibitory.

Aim 3 of this work was to build an expression vector that provides all functions, starting from pMMH6, using the knowledge gained in Aim 2. These new insights would also inform the redesign of the pathway described in Chapter 4.

Aim 4: Investigation of the putative promoter within *mupL* and its requirement for conversion of PA-B to PA-A

The final aim of this chapter arose from discoveries made during Aims 2 and 3 centered around *mupL*. This was not an initial aim as there was no knowledge of a requirement for *mupL* or a promoter within it prior to this work. Investigation of this putative promoter led to the discovery of some intriguing potential interactions with the resistance gene *mupM*, which were then investigated.

3.2 Results

3.2.1 The role of MupV in mupirocin biosynthesis

MupV was originally bioinformatically predicted to be an oxidoreductase, and *P. fluorescens* Δ*mupV* was observed to produce PA-B (El-Sayed et al., 2003; Cooper et al., 2005).

3.2.1.1 The bioinformatic re-analysis of MupV reveals two predicted domains

Re-analysis of the MupV amino acid sequence revealed two predicted domains, both the previously identified oxidoreductase and a second, α/β hydrolase domain. All prediction methods used agreed there are two domains in MupV. NCBI conserved domains predicted domain 1 as a short chain dehydrogenase/reductase (SDR) (E-value $1.46E^{-91}$), and domain 2 as an α/β hydrolase (E-value $9.12E^{-4}$). A HH-pred search of the Pfam database revealed domain 1 shows homology to the NAD binding domain of the male sterility protein (E-value $8.6E^{-37}$), and to the 3-beta hydroxysteroid dehydrogenase/isomerase family (E-value $2.1E^{-36}$). By the same search, domain 2 showed homology to an α/β hydrolase fold (E-value $1.5E^{-23}$), and also to a thioesterase domain (E-value $3.4E^{-10}$).

The hypothesis that MupV acts as a thioesterase fits with our proposed pathway, where it is thought to catalyse release of mupirocin P from MacpE (Figure 3.3). However, this does not explain why two domains have evolved. One possibility is that both are required for effective thioesterase activity, another is that MupV has a dual enzymatic function, and a third is that one domain is functionally inactive. Point mutations to conserved residues in both domains were designed to test whether the function of both are required for conversion of PA-B to PA-A. It was decided not to carry out these experiments by deletion of each domain, as this could disrupt the structure of the other domain, preventing the attribution of any phenotypic change to a single domain.

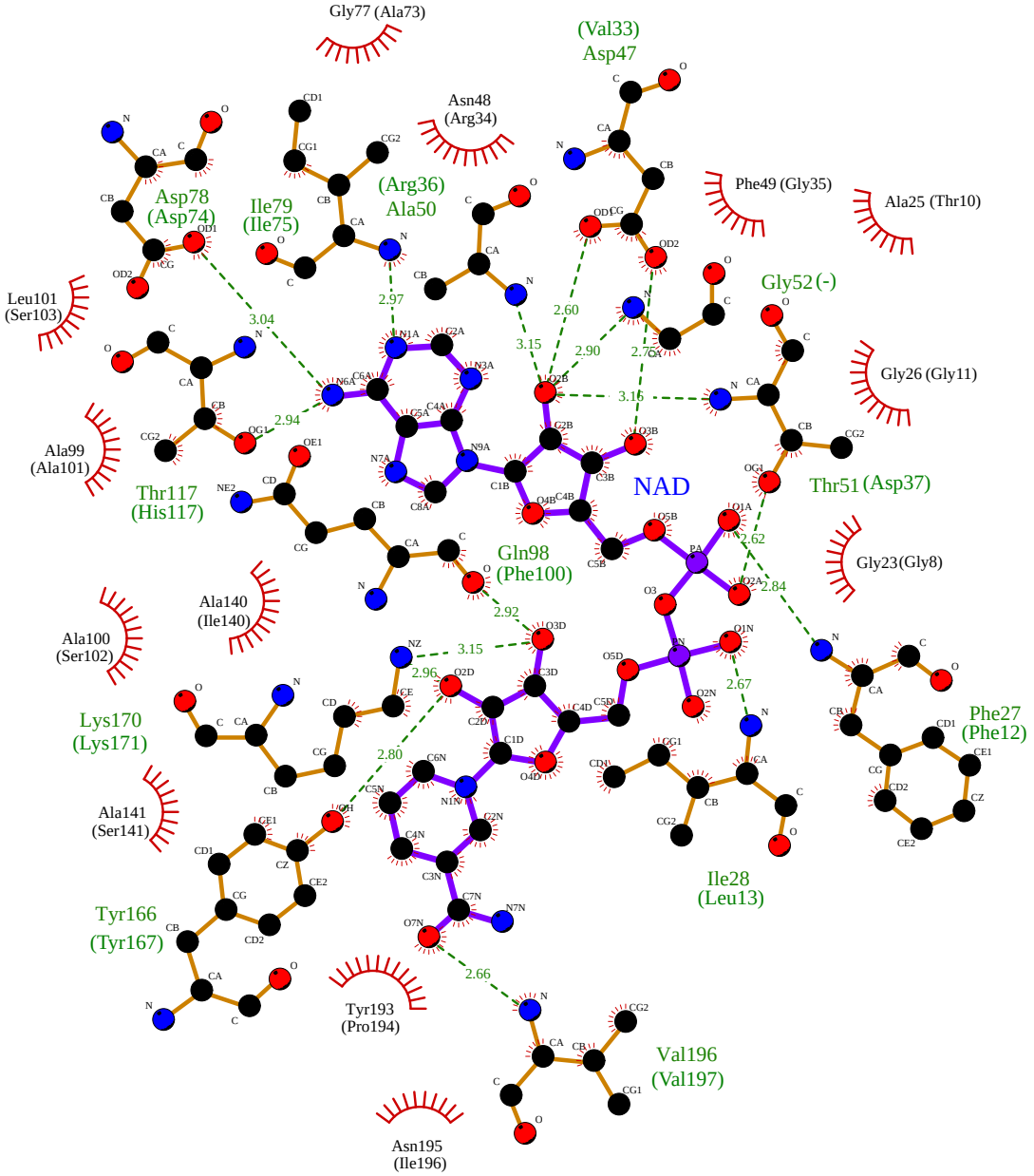


Figure 3.4: Active site interactions of MupV domain 1 homolog UDP-N-acetylglucosamine 4-epimerase from *Pseudomonas aeruginosa*.
Residues without brackets refer to 1SB8, whereas residues with brackets are the MupV equivalent. Tyr166, equivalent to MupV (Tyr167), forms a hydrogen bond with NAD in the active site. Generated by Ligplot for X-ray structure 1SB8 (Ishiyama et al., 2004).

Domain 1 (SDR) has a Rossman fold with characterised active site and NAD binding residues, allowing for the straightforward identification of key residues for point mutation. The residue chosen was the Tyr of the active site YxxxK motif, with the point mutation designed to change it to Phe, Y167F (Kavanagh et al., 2008).

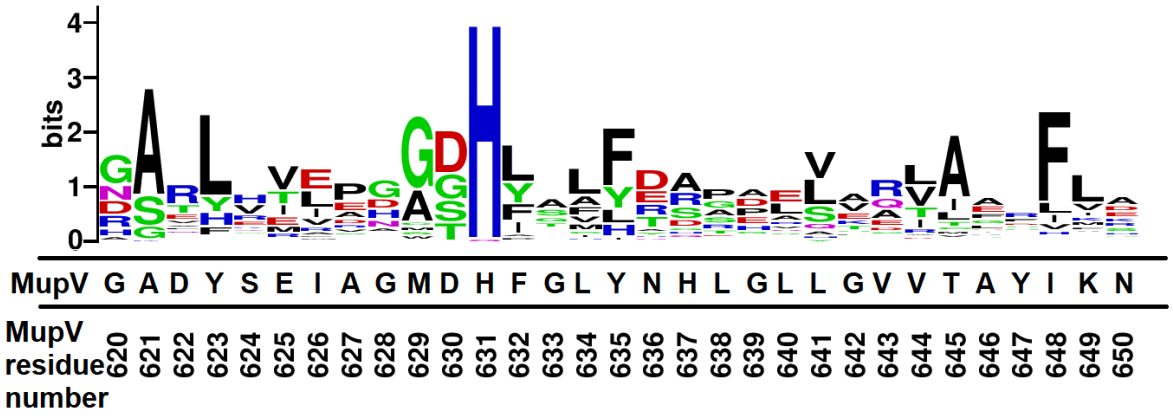


Figure 3.5: Sequence logo of MupV domain 2.
PSI-BLAST was used to identify a set of 55 homologs manually curated for diversity. The sequence logo represents the frequency of each amino acid at each position in the alignment. Underneath the logo the MupV residue at each position is listed. His631 is highly conserved, and was therefore chosen for mutagenesis.

Phenylalanine lacks the hydroxyl that forms the hydrogen bond with NAD in the homolog, but is otherwise the same as tyrosine, minimising the chance of structural changes to the enzyme.

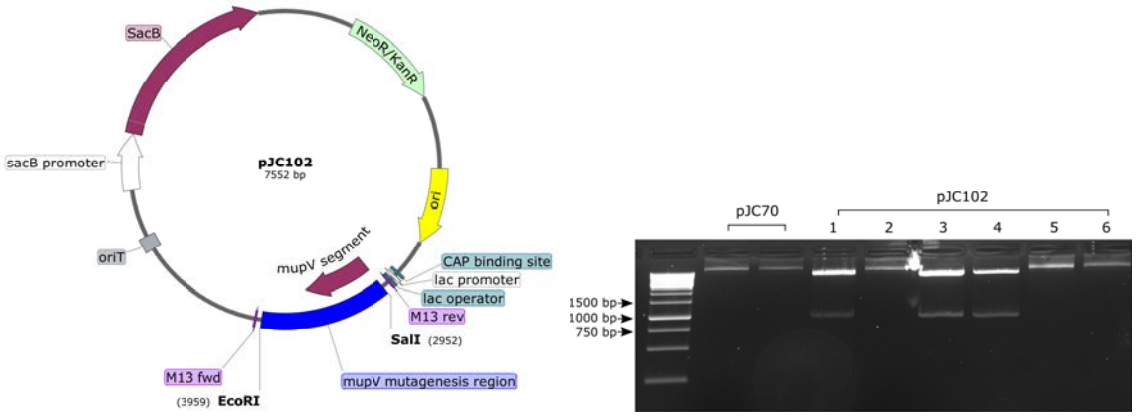
Domain 2 (α/β hydrolase) does not show homology to proteins with well-characterised active site domains. Instead, a point mutation was designed on the basis of degree of amino acid conservation. Alignment of 55 related sequences revealed a strongly conserved His. The point mutation was designed to change it to Ala, H631A (Figure 3.5).

3.2.1.2 Generation of *mupV* point mutations

As discussed in Section 3.1.3, complementation of *P. fluorescens* NCIMB10586 Δ*mupV* with *mupV* provided *in trans* on expression vector pJH10 was not achieved. Therefore a chromosomal mutagenesis strategy was pursued, where point mutations to both domains would be introduced into the 10586 chromosome.

Generation of the MupV domain 2 H631A mutagenesis plasmid, pJC102

Primers P17_F_outer_H631A_MupV and P17_R_inner_H631A_MupV were used



(a) **(b)** *Eco*RI/*Sal*I digests.

Figure 3.6: MupV H631A mutagenesis vector, pJC102.
(a) Map of pJC102, a pJC70 derivative carrying V581A, H631A MupV mutagenesis insert.
(b) Screening digest of potential pJC102 isolates. Isolates 1, 3 and 4 contain a 1 kb insert, and this was confirmed as the correct fragment by DNA sequencing. Isolates 2, 5 and 6 lack an insert and appear to be re-ligated empty pJC70.

to amplify a 546 bp fragment, and P17_F_inner_H631A_MupV and P17_R_outer_-H631A_MupV to amplify a 500 bp fragment. All primer sequences used are listed in Table 2.7. The inner primers were designed to generate a 21 bp overlap, including the H631A mutation (DNA codon CAT to GCA). The products were overlapped by PCR to yield 1025 bp, which was AT-cloned in pGEM-T-easy, yielding pJC99. Insert presence was confirmed by blue-white screening, followed by plasmid preparation and diagnostic *Eco*RI digest. Sequencing using primers M13(-40) and M13(-48) revealed that all isolates carried a secondary mutation, V581A. Whilst unintentional, V581A is also a point mutation to domain 2 of MupV, and therefore deemed acceptable.

The fragment was cloned using *Sal*I/*Eco*RI into mutagenesis vector pJC70, yielding plasmid pJC102. The presence of correct insert was screened by *Eco*RI/*Sal*I digest (Figure 3.6b), and sequencing with primers M13(-40) and M13(-48).

Generation of the MupV domain 1 Y167F mutagenesis plasmid

Following a similar strategy to domain 2, primers F_outer_Y167F_MupV and R_inner_Y167F_MupV were used to amplify a 555 bp upstream arm, and primers

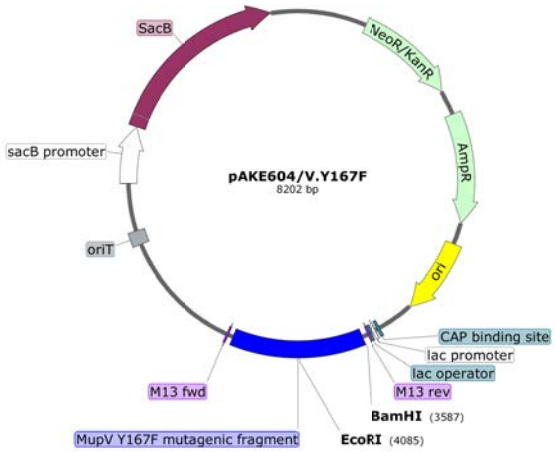


Figure 3.7: Map of pAKE604/V.Y167F.

F_inner_Y167F_MupV and R_outer_Y167F_MupV a 502 bp downstream arm. The inner primers were designed with a 23 bp overlap, including the Y167F mutation (DNA codon TAC to TTC). The products were overlapped by PCR to yield 1034 bp, which was cloned *Mfe*I/*Bam*HI to pAKE604 cut *Eco*RI/*Bam*HI. The resulting mutagenesis plasmid, pAKE604/V.Y167F (Figure 3.7), was checked by *Eco*RI digest and sequencing with the M13(-40) and M13(-48) primers[1].

Point mutations were generated to both domains of MupV

Plasmids pJC102 and pAKE604/V.Y167F were mobilised to the mupirocin producer *P. fluorescens* 10586, and mutants generated as described in section 2.2.12. As this suicide methodology yields strains reverted to WT and mutant at a 50:50 frequency, resulting strains were screened to identify correct mutants. For Y167F, no restriction change was included in the design of the mutation. Mutants were instead detected by sequencing of 1 kb PCR products from primers F_outer_Y167F_MupV and R_outer_Y167F_MupV.

As plasmid pJC102 included both the intended H631A and unintended V581A mutation, screening for both was necessary. Mutagenesis with this plasmid should yield both individual mutations and the dual mutation, depending on the random

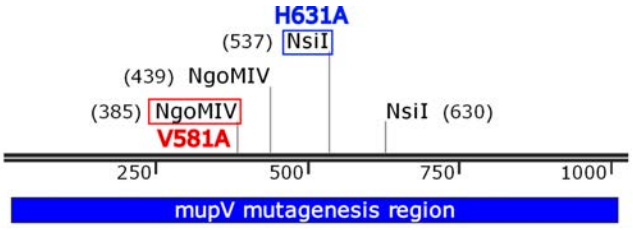
[1]Design and construction of pAKE604/V.Y167F, and subsequent mutagensis of 10586, were performed by undergraduate project student Amber Wilson.

position of the crossovers during the process. After pJC102 integration and excision into 10586 WT, 20 isolates were screened for loss of kanamycin resistance. Three were still kan^R, suggestive of inactivation of SacB rather than plasmid excision. The remaining 17 were screened by PCR of 1 kb using primers P17_F_outer_H631A_MupV and P17_R_outer_H631A_MupV, followed by *Ngo*MIV and *Nsi*I digests. These digests indicated that strains carrying both V581A and H631A together, and V581A individually, had been isolated (Figure 3.8). Two of each of these strains were confirmed by sequencing of the PCR products with the outer primers, and carried forward for assays.

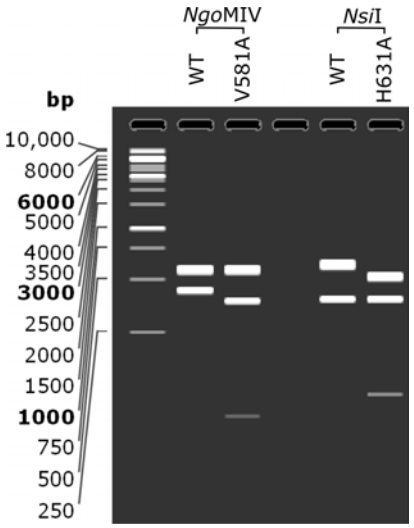
3.2.1.3 Both domains of MupV are required for conversion of PA-B to PA-A

The three MupV mutant strains were assayed for active mupirocin production against *Bacillus subtilis* as described in Section 2.3.2. The Y167F and H631A (with V581A) mutations decreased the zones of inhibition compared with WT, whereas the unintended V581A mutation displayed no effect (Figure 3.9). This demonstrated that enzymatic activity of both domains of MupV is required for full production of active mupirocin.

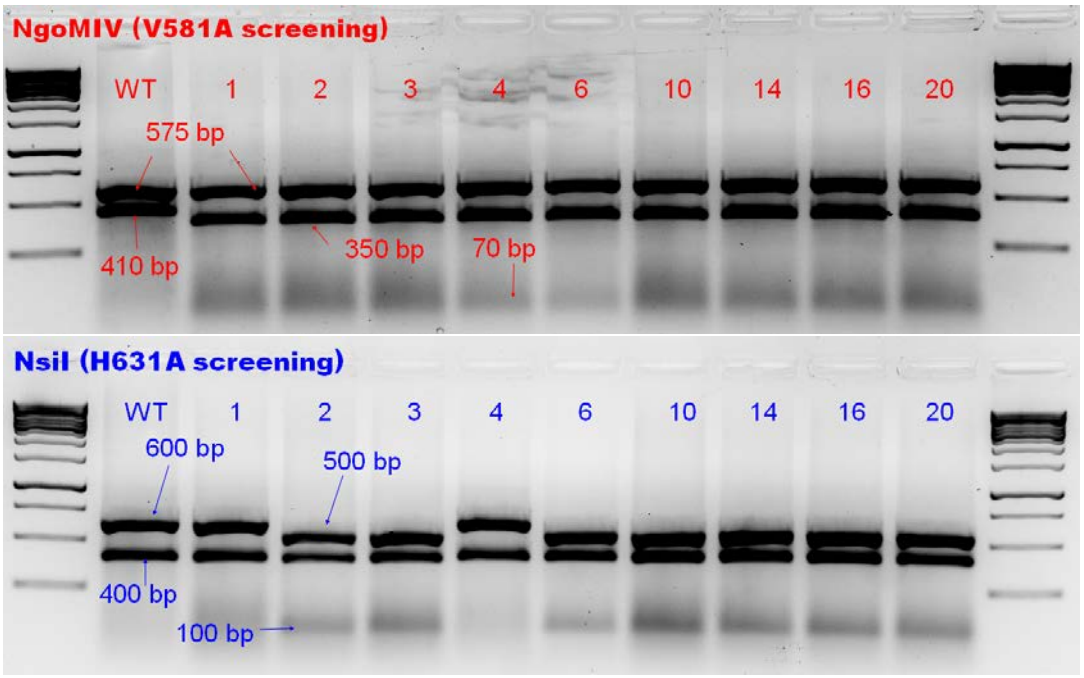
P. fluorescens 10586 Δ*mupV*, carrying an in-frame deletion of *mupV*, has been previously demonstrated to produce PA-B as its primary metabolite (Cooper et al., 2005). HPLC analysis of the point mutations demonstrated that both domains are required for conversion of PA-B to PA-A (Figure 3.10). This observation of accumulation of PA-B was confirmed by our collaborators at University of Bristol by LC-MS and ^{1}H-NMR (C. Willis and L. Wang 2017, personal communication, 7 June). As these are point mutations to key residues rather than deletions, the deficiency is likely due to a lack of enzymatic function. This suggests that either both domains are involved in the putative role in intermediate release from MacpE, or that MupV has a further function in PA-B to PA-A conversion.



(a) Map of *mupV* domain 2 PCR product.

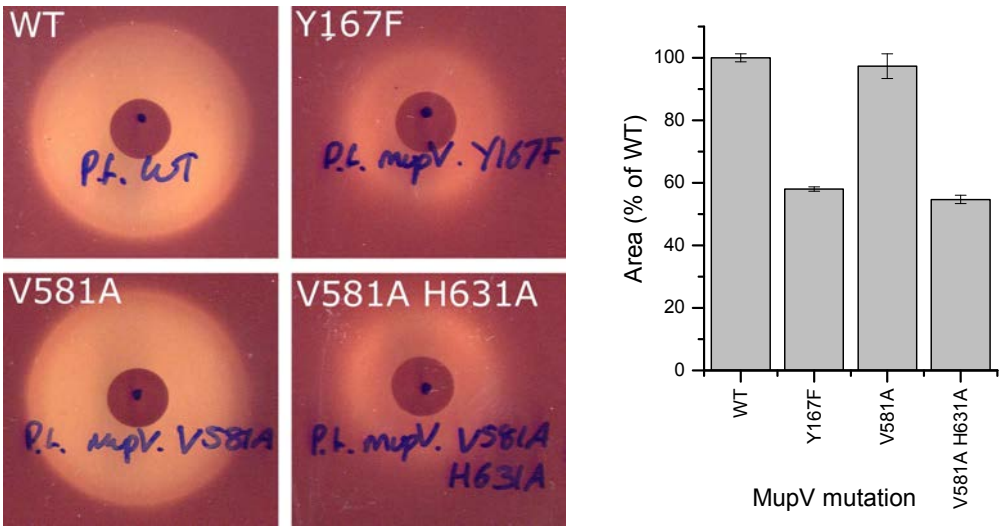


(b) Simulated restriction digest gel.



(c) Restriction digests, 2% agarose gel

Figure 3.8: Screening for *mupV* domain 2 mutations.
Map **(a)** and simulated digest **(b)** show the extra *Ngo*MIV and *Nsi*I sites introduced by the V581A and H631A mutations respectively.
(c) Of the 9 PCR products from potential mutant strains shown, all contain the V581A mutation, with the 350 bp band in place of 410 bp in the *Ngo*MIV digest. Strains 1 and 4 have the WT *Nsi*I restriction profile, suggesting they are not mutant at H631. All other strains carry the H631A restriction profile.
These digests suggest strains 1 and 4 are V581A single mutants, and all others are dual V581A and H631A mutants. This conclusion was confirmed by sequencing of the PCR products.



(a) **(b)**

Figure 3.9: Point mutation to either domain of MupV decreases mupirocin activity.
(a) Mutants Y167F and V581A+H631A show deficient zones of clearing compared with WT. No defect is present for unintended mutant V581A.
(b) Areas of inhibition, expressed as % of WT. Bioassays were performed independently in duplicate, error bars are standard deviation.

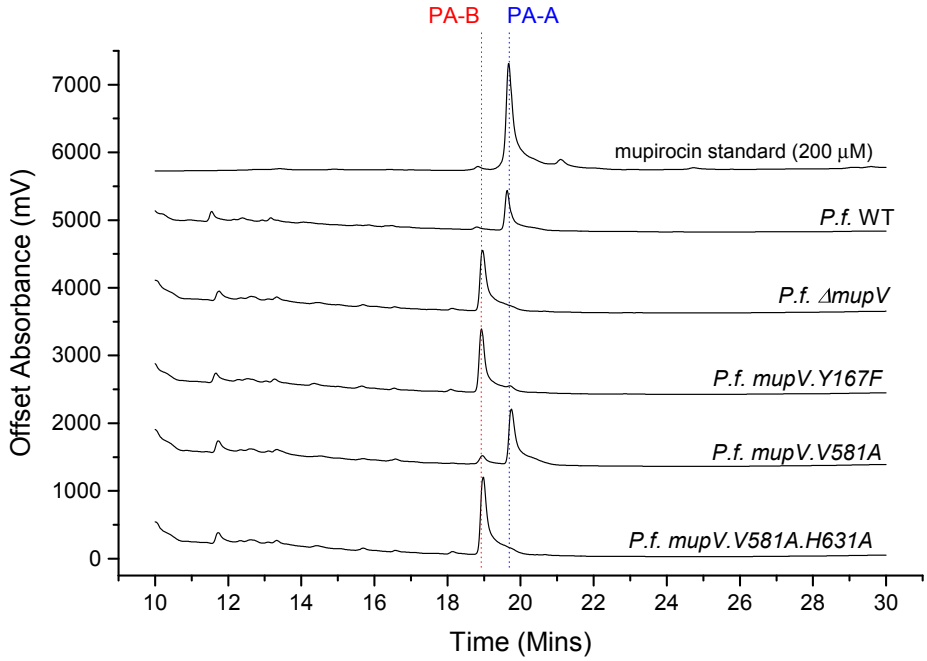


Figure 3.10: HPLC of MupV point mutants.
In-frame deletion of *mupV*, 10586 Δ*mupV*, switches production from PA-A to PA-B. Point mutation of domain 1 (Y167F) and domain 2 (H631A) reproduces this switch to PA-B. This suggests that enzymatic function of both domains of MupV is required for conversion of PA-A to PA-B. Analyses were performed independently in triplicate, and representative chromatograms shown.

3.2.2 Screening tailoring gene mutants for the ability to convert PA-B to PA-A

Deletion of *mupU*, *macpE*, *mupV*, or *mupO* causes accumulation of PA-B, directly implicating these genes as required for conversion to PA-A (Cooper et al., 2005). Deletion of *mupC* and *mupF* cause accumulation of later stage intermediates mupirocin C and F respectively (Hothersall et al., 2007). Recently, requirement for further genes or functions has been demonstrated: *mupP* (Gao et al., 2017), and both domains of *mupV* (in this work). Given the historical identification of further genes required for conversion to PA-A, together with the several mupirocin cluster genes without assigned functions, an approach was devised to systematically screen all mupirocin tailoring genes for ability to convert PA-B to PA-A.

This approach aimed to identify all required genes by feeding PA-B to individual chromosomal mutations of each tailoring gene. Firstly, the logical preliminary objective was to observe conversion of fed PA-B to PA-A in a positive control strain.

This was initially attempted by co-inoculation of SSM media with saturated LB culture of 10586 Δ*mmpA* and 10586 Δ*mupU*. The Δ*mupU* mutant is known to produce PA-B, whereas MmpA is a PKS in part responsible for monic acid biosynthesis, and therefore 10586 Δ*mmpA* does not produce any pseudomonic acid but is expected to be able to convert PA-B to PA-A. Theoretically, in co-culture the Δ*mupU* mutant could generate PA-B that the Δ*mmpA* strain could convert to PA-A. However, no pseudomonic acid production was detected by HPLC from these co-cultures.

3.2.2.1 A Pks^{-} strain, deficient in pseudomonic acid production, is able to convert fed PA-B to PA-A

With the failure of the co-culturing approach, an alternative was pursued where PA-B would be extracted and fed to test cultures. To decrease the amount of PA-B required, it was necessary to downscale the 25 ml SSM production cultures used

in previous experiments with 10586. Good mupirocin yields were observed from 5 ml SSM cultures in 50 ml conical flasks, whereas no yield was observed from 5 ml cultures in vented 50 ml Falcon tubes or 30 ml universal containers.

Plasmid pJH2, carrying the gene for transcriptional activator MupR, was mobilised to 10586 Δ*mupU*. Expression of *mupR in trans* has been observed to increase mupirocin metabolite production (Hothersall et al., 2011). To generate PA-B, 200 ml SSM cultures in 2 L flasks were inoculated with 10586 Δ*mupU* [pJH2]. One possibility for purifying PA-B was to fraction collect by HPLC, however this would not have efficiently generated the scale of PA-B necessary for these experiments.

Instead, ethyl acetate extractions were performed: the ethyl acetate top layer was collected and evaporated to less than 500 μl yellow oily liquid per 200 ml starting culture, using the method described in Section 2.3.4. To aid resuspension of any solid matter, 2 to 4 ml of methanol was added to each. The approximate concentration of PA-B present in each extract was analysed by comparison to mupirocin standards by HPLC. For the first two batches of PA-B generated, a working concentration of 1% (v/v) was used in subsequent experiments (very approximately 200 μM). The third batch contained a higher concentration of PA-B, and a working concentration of 0.5% (v/v) was used, also approximately 200 μM.

As a preliminary experiment, SSM cultures of Pks$^-$ strain 10586 Δ*mmpA* were fed with PA-B in triplicate. Gratifyingly, 42.2% $\pm$4.8 (s.d.) of the fed PA-B was converted to PA-A, measured by integrated peak area (Figure 3.11). No PA-A was detected in negative control 10586 Δ*mup-cluster*, a strain with all mupirocin biosynthesis genes deleted with the exception of transcriptional regulators *mupR*, *mupX* and *mupI*. As expected, deletion of genes known to be required for conversion such as *mupU* as part of the cluster deletion, prevents conversion from occuring. This demonstrated that this method can screen for mutants that are deficient in conversion.

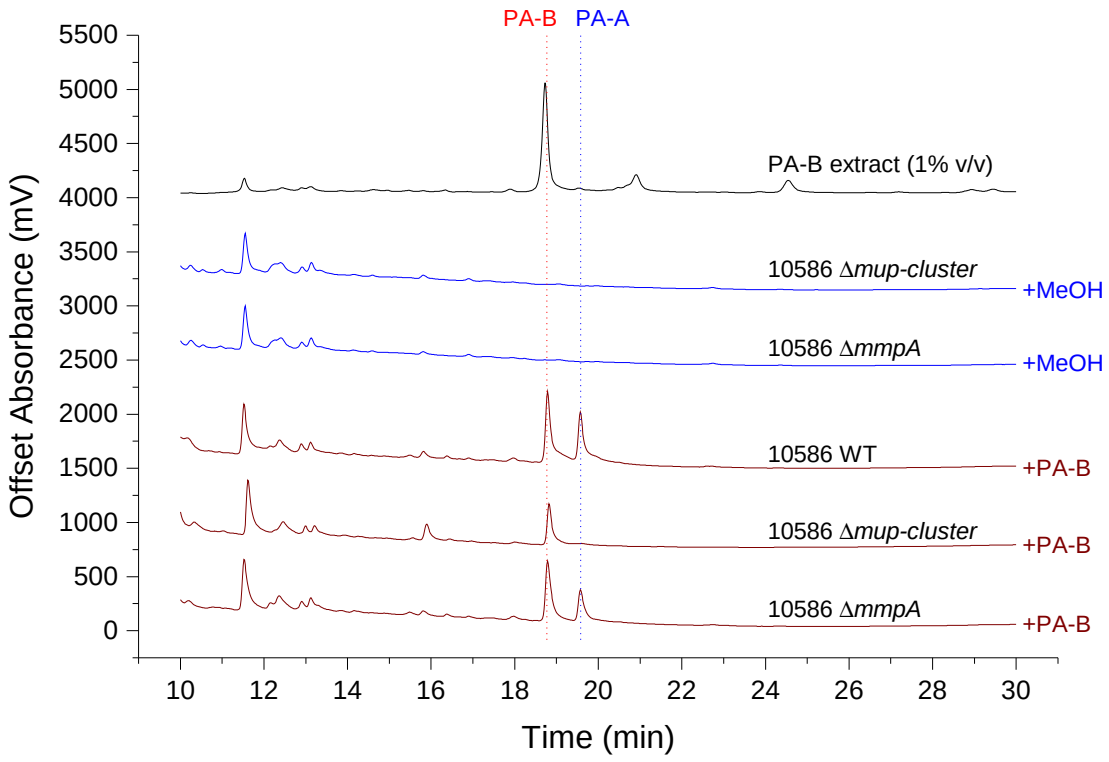


Figure 3.11: HPLC: conversion of fed PA-B to PA-A by 10586 Δ*mmpA*.
Strains 10586 Δ*mmpA* and Δ*mup-cluster* both produce no pseudomonic acids in the methanol (1% v/v) control. When fed with PA-B extract (1% v/v) (shown in red), conversion is observed in Δ*mmpA* but not Δ*mup-cluster*.

3.2.2.2 Identification of tailoring genes required for conversion of PA-B to PA-A

Continuing from the successful preliminary experiment, the same method was used to screen individual mutations to 35 mupirocin tailoring genes. SSM cultures were grown with 0.5% methanol or 0.5% PA-B, initially in duplicate, and supernatants analysed by HPLC. The amount of PA-A and PA-B present in each sample was determined by integration of HPLC peak areas.

To determine the percentage of conversion of fed PA-B to PA-A, samples were first adjusted for their methanol controls, to account for any inherent production of either compound by the strains. The calculations are listed below, where **A** = area of PA-A, **B** = area of PA-B, x_{PA-B} = from sample fed with PA-B, x_{MeOH} = from control sample fed with methanol:

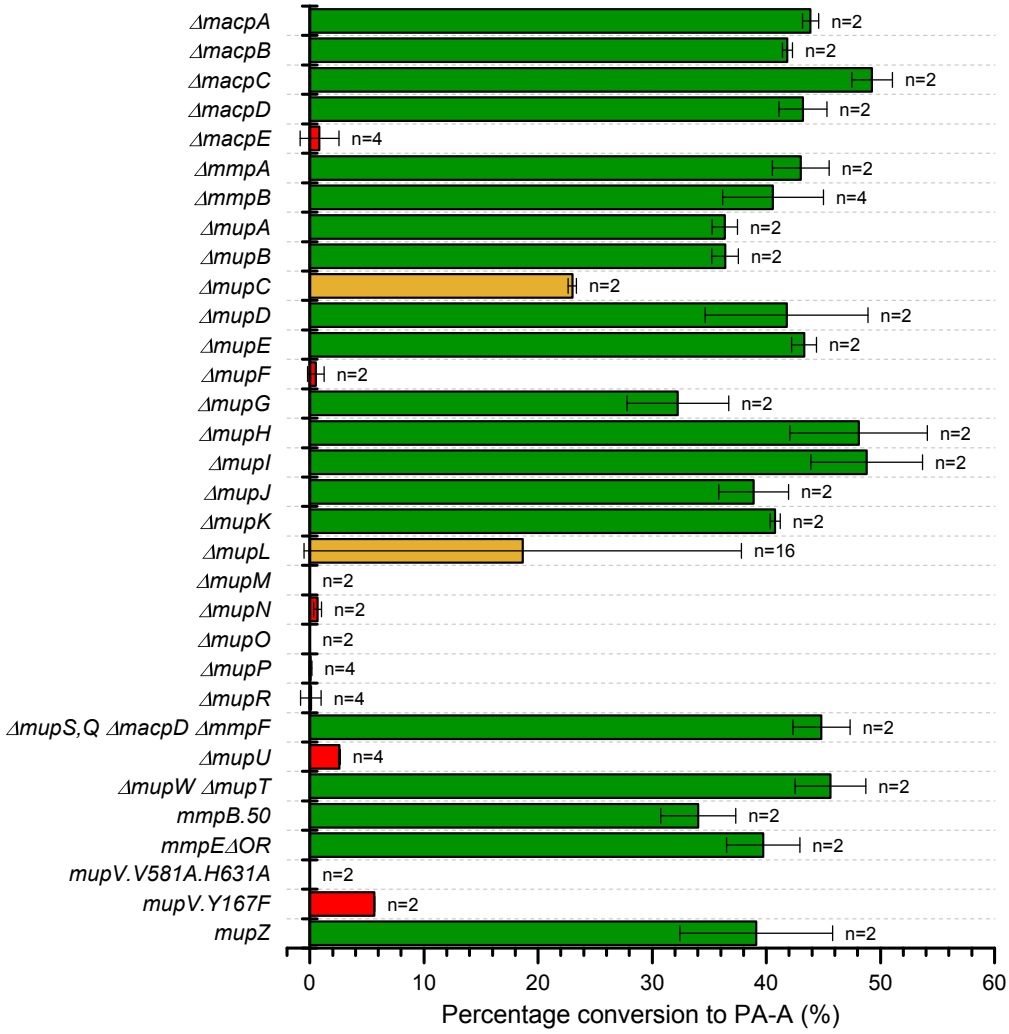
$$\mathbf{A} = A_{PA-B} - A_{MeOH}$$

$$\mathbf{B} = B_{PA-B} - B_{MeOH}$$

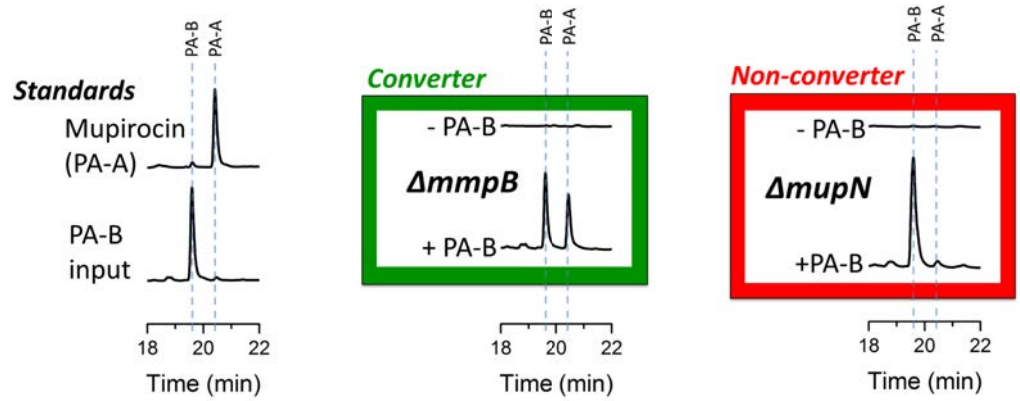
$$Conversion = \frac{\mathbf{A} \times 100}{\mathbf{A} + \mathbf{B}}$$

This screening method proved successful, with a distinct phenotype between converters and non-converters for the majority of genes (Figure 3.12). Mutants of genes *macpE, mupF, mupU, mupO, mupP* and both domains of *mupV* were deficient in conversion, all of which were already thought required in the current model (Section 3.1). Mutants of genes with characterised functions earlier in the pathway, such as *mmpA* or *mupW* and *mupT*, were capable of conversion as expected.

Interestingly, deletions of multiple further genes were deficient in conversion: Δ*mupM*, Δ*mupN* and Δ*mupR*. These are logical requirements, *mupM* encodes the mupirocin resistance gene, and this requirement is explored further in Section 3.2.4. Gene *mupN* encodes a phosphopantetheinyl transferase, and this is the first evidence that it is presumably obligately required for generation of the MacpE holoenzyme. MupN has previously been demonstrated to activate MacpA, MacpC and MacpD



(a) Conversion calculated by integrated HPLC peak areas



(b) Example HPLC chromatograms

Figure 3.12: Screening of 10586 mutants for ability to convert PA-B to PA-A. Mutations to *macpE*, *mupF*, *mupM*, *mupN*, *mupO*, *mupP*, *mupR*, *mupU* and *mupV* were unable to convert fed PA-B to PA-A. The Δ*mupL* mutant was observed to variably convert or not convert. The *mupC* mutant is able to produce some PA-A as well as mupirocin C.

Table 3.1: Genes implicated by screening experiment as required for conversion

Gene	Putative encoded function
macpE	Acyl carrier protein
mupC	Enoate reductase
mupF	Ketoreductase
mupL	Hydrolase
mupM	Isoleucyl-tRNA synthetase (self resistance)
mupN	Phosphopantetheinyl transferase
mupO	Cytochrome P450
mupP	Dehydratase
mupR	Transcriptional activator
mupU	Acyl CoA synthase
mupV	Oxidoreductase and thioesterase

in vitro, whereas soluble MacpE was not achieved in the experiment (Shields et al., 2010). Thirdly, *mupR* encodes the transcriptional activator known to be required for mupirocin gene expression (Hothersall et al., 2011).

The ability of the non-mupirocin producing, Δ*mupI* mutant to convert was suprising, and this is explored in the following Section 3.2.2.3.

Two further mutants demonstrated an intermediate phenotype, Δ*mupC* and Δ*mupL*. The *mupC* mutant has been previously observed to produce a lower amount of PA-A than WT, as well as intermediate metabolites mupirocin C1 and C2 (Gao et al., 2014). The *mupL* mutant was observed to be variably able to convert PA-B to PA-A (n=16). This variability was also observed and explored further in experiments described in Section 3.2.4.

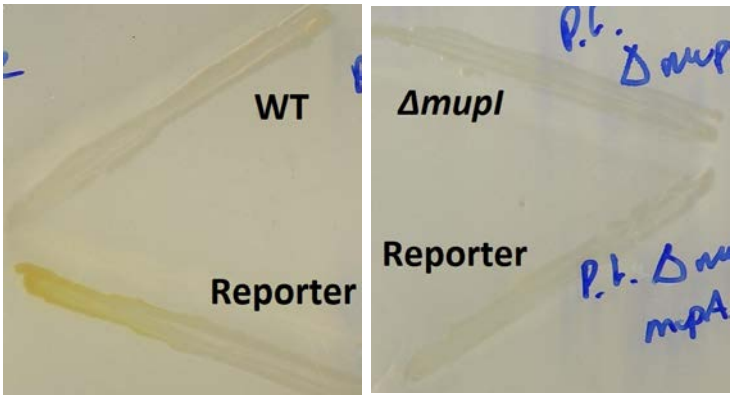
3.2.2.3 The extracted PA-B includes the signalling lactone, facilitating the conversion of PA-B to PA-A by *P. fluorescens* Δ*mupI*

The 10586 Δ*mupI* knockout mutant has previously been observed to be deficient in pseudomonic acid production (El-Sayed et al., 2001). MupI is an N-acylhomoserine lactone synthase, responsible for production of the quorum sensing signalling molecule (Hothersall et al., 2007). This lactone is required for activation of transcription of the *mupA* promoter, and likely other promoters in the mupirocin biosynthesis cluster (El-Sayed et al., 2001). The observation in Section 3.2.2.2 that the Δ*mupI* mutant is able to convert fed PA-B to PA-A is therefore unexpected. In this mutant, there should not be expression of other genes required for conversion, such as *macpE* or *mupO*.

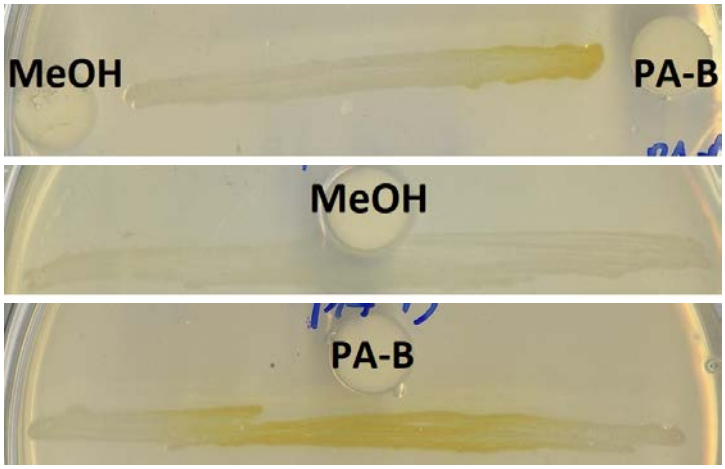
One hypothesis that explains this observation is that the extract is complementing the lack of MupI. As the PA-B extract was generated from the supernatant of a Δ*mupU* strain without any selective purification, it could have contained the MupI product, the signalling lactone. This hypothesis was tested using a *xylE* reporter strain generated in El-Sayed et al. (2001), 10586 Δ*mupI mupA::xylE*. This strain lacks MupI, and therefore does not produce homoserine lactone, and has chromogenic reporter gene *xylE* fused to *mupA* downstream of its promoter. This allows for detection of induction of *mupA* expression using catechol, and the method is described in Section 2.3.3.

This reporter strain can detect the presence of the signalling lactone. When grown close to 10586 WT on solid media, the WT produces the lactone, which diffuses to the reporter and activates *xylE* expression. This is demonstrated by the positive control, the reporter next to 10586 WT turns yellow in presence of catechol (Figure 3.13a). As expected, no colour change is observed in the reporter next to the negative control test 10586 Δ*mupI*, as this strain does not generate the lactone (Figure 3.13b).

To test the PA-B extract for lactone presence, 20 μl PA-B extract, or methanol as a negative control, were spotted onto paper discs and dried overnight. The discs were placed onto agar plates with the reporter strain streaked in close proximity. When



(a) (b)



(c)

Figure 3.13: Transcriptional *xylE* assay demonstrating presence of signalling lactone in PA-B extract.
The reporter strain, 10586 Δ*mupI mupA::xylE*, expresses *xylE* in the presence of the signalling homoserine lactone, that it does not produce itself (El-Sayed et al., 2001). Lactone presence activates *xylE* expression, the product XylE catalyses the conversion of catechol to the yellow compound 2- hydroxymuconic semialdehyde (Zukowski et al., 1983).
(a) 10586 WT produces the signalling lactone that induces expression of *xylE* in the reporter strain, n=2.
(b) 10586 Δ*mupI* does not produce the lactone, therefore no *xylE* expression or colour change is observed, n=2.
(c) The PA-B extract is able to induce *xylE* expression in the reporter (n=9) as it contains the lactone, whereas no colour change is observed in the methanol negative control.

sprayed with catechol, the reported clearly turned yellow in the proximity of the PA-B extract and not in proximity of the negative control (Figure 3.13). This result neatly explains how the 10586 Δ*mupI* mutant is able to convert fed PA-B, because the extract also contains the signalling lactone.

3.2.3 Demonstration of the minimum set of genes required for conversion of PA-B to PA-A

The counterpart to screening for all genes required, is to clone them all to build a multi-gene expression plasmid, then test if it can provide all functions necessary for conversion. This idea was first explored in the PhD thesis of Macioszek (2009), who generated pJH10 based expression plasmid pMMH6, carrying: *mupO*, *macpE*, *mupU*, *mupV*, *mupC* and *mupF*. This plasmid was demonstrated to complement the mupirocin production deficiency of individual in-frame deletions of each included gene in 10586, detected by bioassay. However, pMMH6 did not include all genes required, and as such was unable to catalyse conversion of PA-B to PA-A.

Later mutagenesis and feeding experiments using PA-C demonstrated *mupP* is also required for C8-OH removal (Gao et al., 2017). Insertion of *mupP* into pMMH6 was attempted in the PhD work of Yadav (2017), but due to difficulties with cloning, this was not achieved. The systematic approach undertaken in Section 3.2.2.2 has also implicated *mupM* and *mupN*, as well as a variable phenotype for *mupL*. Therefore an aim of this study became to clone *mupP*, *L*, *M*, *N* into pMMH6 already carrying *mupO,U,V,C,F,* and *macpE*.

3.2.3.1 Construction of expression vector carrying *mupOPUVCF* and *macpE*

Plasmid pMMH6 proved a very useful starting point, as it already carried six of the desired ten genes, and they had been confirmed functional by bioassay. However, the restriction cloning opportunities were fairly limited, and the stop codon of *mupO*

was missed in the design of pMMH6, resulting in an additional 27 random amino acids on MupO before a stop codon was reached. As such, the planned approach was to clone *mupO*, *mupP*, *macpE* on a *Kpn*I/*Afl*II fragment, and then insert *mupL*, *mupM*, *mupN Xba*I/*Not*I, at the back of the cloned operon.

A schematic and maps of the constructions are shown in Figure 3.14. Primers P25_F_mupOP_KpnI_SpeI and P25_R_mupOP_overlap were used to amplify a 2371 bp fragment carrying *mupO* and *mupP*, which neighbour each other on the chromosome. These were designed to include the ribosome binding sites (RBS) of both genes. Primers P25_F_macpE_overlap and P25_R_macpE_AflII were used to amplify a 278 bp fragment carrying *macpE*, again with the native RBS. These were designed to overlap, and SOEing PCR performed to generate a 2627 bp fragment carrying *mupO*, *mupP* and *macpE*. This fragment was cloned *Kpn*I/*Afl*II into pMMH6, resulting in pJC124, after effectively replacing *mupO*/*macpE* with *mupO*/*mupP*/*macpE*. A unique *Spe*I site was included next to *Kpn*I to maintain future cloning flexibility.

Plasmid DNA from transformants that could potentially contain the planned pJC124 were checked by diagnostic *Kpn*I/*Afl*II digests (Figure 3.15), and sequenced using primer P25_R_macpE_AflII (partial coverage). The *Xba*I site of pMMH6 and consequently pJC124 is blocked by Dam methylation, and therefore pJC124 was transformed and prepped from Dam$^-$ *E. coli* ER2925. Further diagnostic digests were performed, both repeating the *Kpn*I/*Afl*II digests, and using *Xba*I/*Not*I to check agreement with the restriction map prior to the next step (not shown).

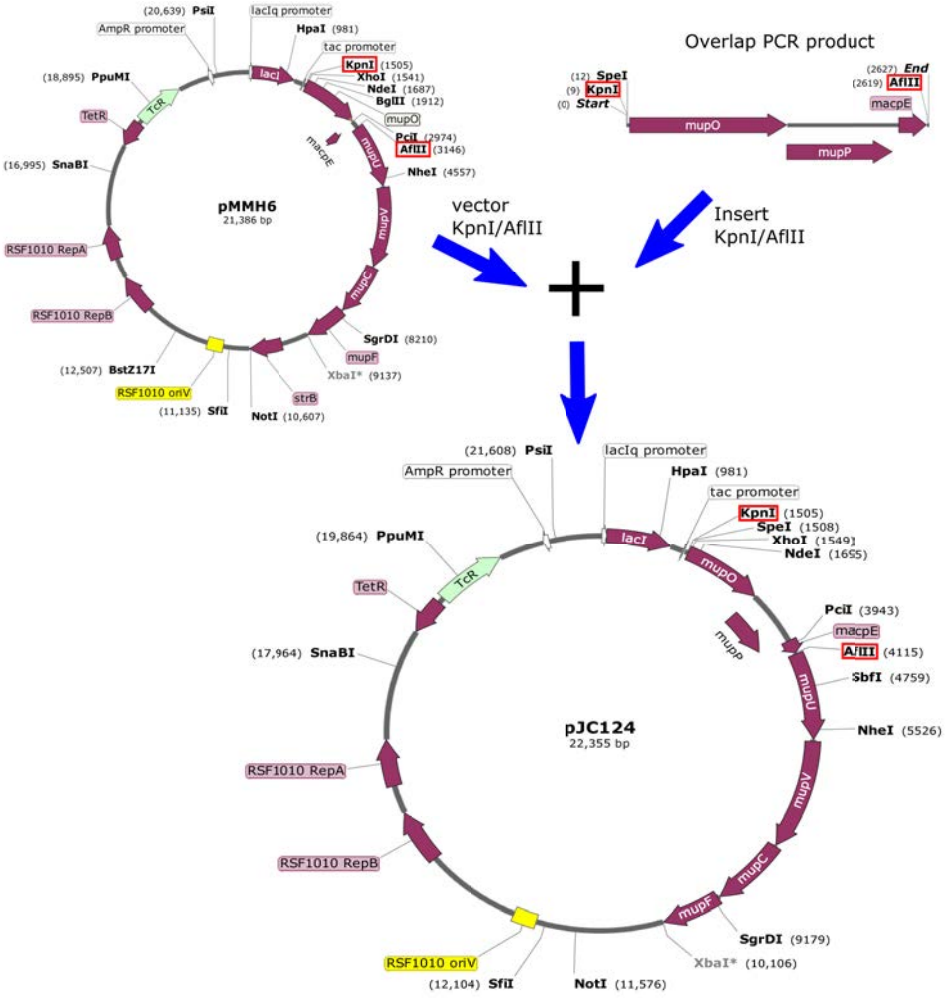


Figure 3.14: Map and construction of pJC124.
A 2.6 kb PCR product carrying *mupO*, *mupP* and *macpE* was cloned *Kpn*I/*Afl*II to pMMH6, yielding pJC124.

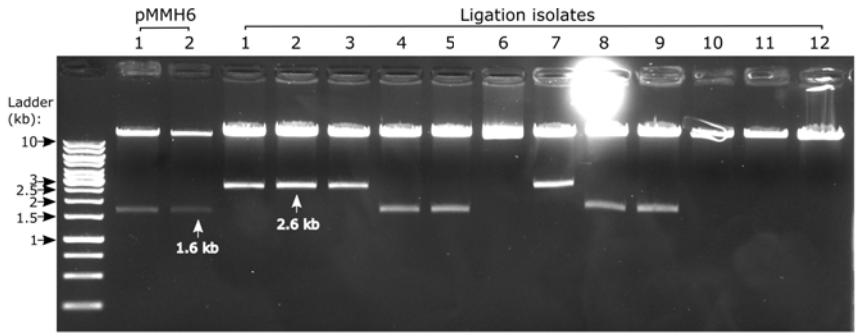


Figure 3.15: Diagnostic *Kpn*I/*Afl*II digest of pJC124.
Isolates 1, 2, 3 and 7 have the 2.6 kb fragment indicative of size increase due to *mupP*. Isolates 3, 4, 8, and 9 appear to be recircularised pMMH6, whereas isolates 6, 10, 11 and 12 appear to be recircularised vector that has lost the *Kpn*I/*Afl*II fragment.

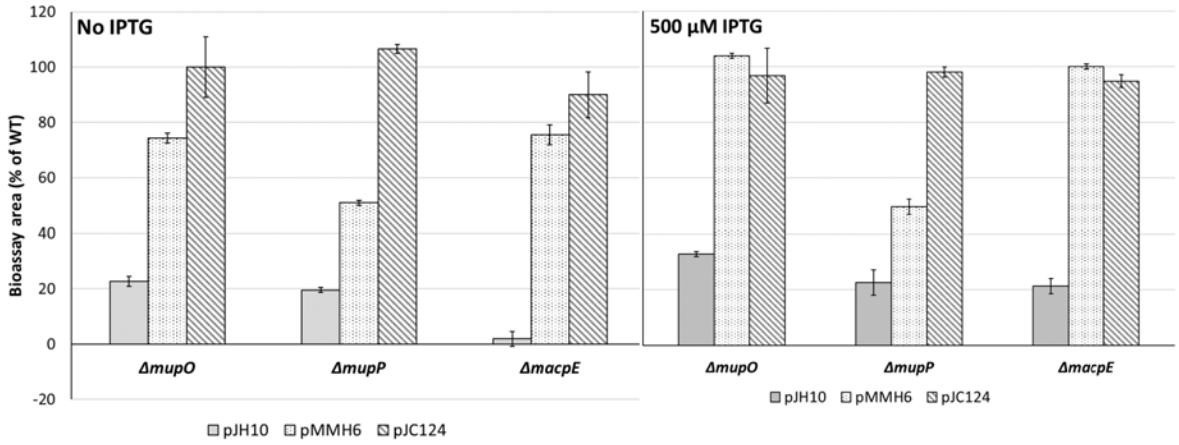


Figure 3.16: Complementation analysis to check that genes present in pJC124 are functional.
Areas are average zones of clearance of *B. subtilis* expressed as a percentage of 10586 WT, n=4 for pJC124, n=2 for pJH10 and pMMH6; error bars show standard deviation.

Complementation of 10586 Δ*mupO*, 10586 Δ*mupP*, and 10586 Δ*macpE* by pJC124

In generating pJC124, *mupO*, *mupP* and *macpE* were derived from overlap PCR products, which could therefore contain sequence errors. Sanger sequencing with primer P25_R_macpE_AflII yielded relatively short reads of approximately 500 bp, leaving the majority of the cloned 2.6 kb unsequenced. Instead of performing further sequencing, pJC124 was tested for its ability to complement individual chromosomal mutations of each of the three genes.

Empty expression vector pJH10 and tailoring gene plasmids pMMH6 and pJC124 were mobilised into 10586 Δ*mupO*, 10586 Δ*mupP*, and 10586 Δ*macpE*. Bioassays were performed against the susceptible test strain, *B. subtilis* (Figure 3.16). With the empty vector, all three mutations are inhibitory to active mupirocin production, as predicted. Plasmid pMMH6 is able to complement 10586 Δ*mupO* and Δ*macpE* as reported by Macioszek (2009). Suprisingly, intermediate complementation of Δ*mupP* was also observed, despite pMMH6 not carrying *mupP*. This could perhaps indicate that over-expression of the related pathway genes, *mupO,U,V,C,F, macpE* on pMMH6 is able to overcome some of the Δ*mupP* deficit.

Table 3.2: Ribosome binding site of *mupO*

Scenario	Sequence (5′ to 3′)
	RBS +1
Consensus	NNNAGGAGGNNNNNNATG
10586 chromosome	TCCGGGAGAGCAACGATG
pMMH6	TACCGGAGAGCAACGATG
pJC124	AGTGGGAGAGCAACGATG

Plasmid pJC124 was observed to complement all three mutations to WT levels, both with and without IPTG induction. Plasmid pMMH6 only complemented 10586 Δ*mupO* and 10586 Δ*macpE* fully with IPTG induction. This could be caused by the missed stop codon on *mupO* in pMMH6 having translational effects on downstream genes; in pMMH6 *macpE* begins 61 bp before the artificial stop codon of *mupO*. There is also a difference in the design of *mupO* cloning between the plasmids, with the RBS in pMMH6 beginning with C compared with G in pJC124 and the 10586 chromosome (Table 3.2). However, neither of these match the *E. coli* consensus at this position (Chen et al., 1994). Another possibility is that insertion of *mupP* in pJC124 has downstream translational effects.

The full complementation of knockout mutants of each of the three genes, indicated that pJC124 carried functional copies of each, an important conclusion for later experiments.

3.2.3.2 Addition of *mupLMN* to expression vector

The last step towards building an expression vector carrying all necessary genes identified in the systematic screening experiment, was cloning *mupL*, *mupM* and *mupN* into pJC124. As a variable *mupL* phenotype was observed in Section 3.2.2.2, versions with and without *mupL* were designed. As these genes lie together on the 10586 chromosome, *mupL, M, N* were amplified as a 5102 bp PCR product, using

primers P25_F_mupLMN and P25_R_mupLMN. A 4074 bp PCR product carrying *mupM* and *mupN* was amplified using primers P25_F_mupMN and P25_R_mupLMN. These were designed to include the native RBS for each gene.

There is a *Not*I site internal to *mupN*, and cloning *Xba*I/*Not*I was one of the very few available possibilities in pJC124. This site could have been removed through overlap PCR by introducing silent point mutation, but instead a quicker *Bsa*I approach was used, detailed in Figure 3.18. The *mupMN* and *mupLMN* PCR products were AT ligated into cloning vector pGEM-T, yielding plasmids pJC128 and pJC129 respectively. Correct isolates were identified using *Sac*I and *Bsa*I/*Xba*I diagnostic digests, and by sequencing using primers M13(-40) and M13(-48).

The *mupMN* and *mupLMN* inserts were released from pJC128 and pJC129 respectively by *Xba*I/*Bsa*I digest. These were ligated to *Xba*I/*Not*I digested pJC124 isolated from *dam*$^{-}$ *E. coli* ER2925, as the *Xba*I site overlapped a Dam methylation site. This yielded plasmid pJC133, carrying *mupO, P, macpE, mupU, V, C, F, L, M, N*, and pJC132 carrying the same genes without *mupL* (Figure 3.17). Isolates were checked by *Sac*I diagnostic digest (Figure 3.19).

3.2.3.3 Expression plasmid pJC133 can complement individual chromosomal deletions of each of the ten genes carried

In sequencing of pGEM-T cloning vectors used to construct pJC132 and pJC133, only coverage of *mupL* and *mupN* was achieved. The larger *mupM* was not covered by the sequencing reads. Instead, the ability of pJC133 and pJC132 to complement 10586 Δ*mupM* would be tested. The Δ*mupM* has a dual phenotype: defective mupirocin production and decreased mupirocin resistance (Hothersall et al., 2007). Complementation of the resistance phenotype is later explored in Section 3.2.4.3.

For completeness, both plasmids were mobilised to individual 10586 chromosomal in-frame deletions of each of the ten (nine for pJC132) mupirocin genes carried, along with the empty expression vector, pJH10. Bioassays were performed on each inde-

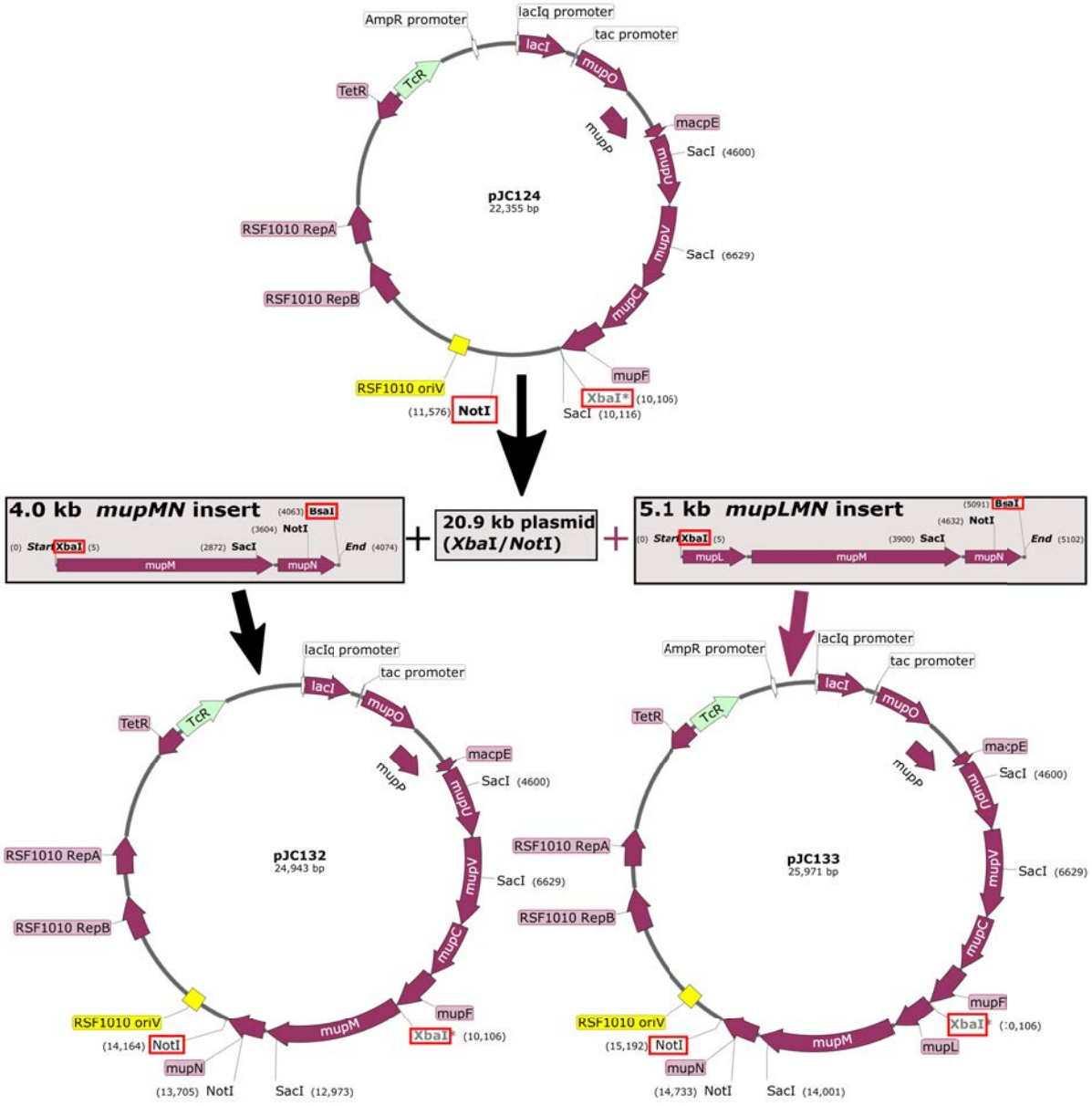


Figure 3.17: Maps and construction of pJC132 and pJC133.
PCR products carrying *mupM, N* and *mupL, M, N* were amplified, cloned in pGEM-T, then digested *Xba*I/*Bsa*I. These were ligated to pJC124 *Xba*I/*Not*I, to yield expression vectors pJC132 and pJC133 respectively.

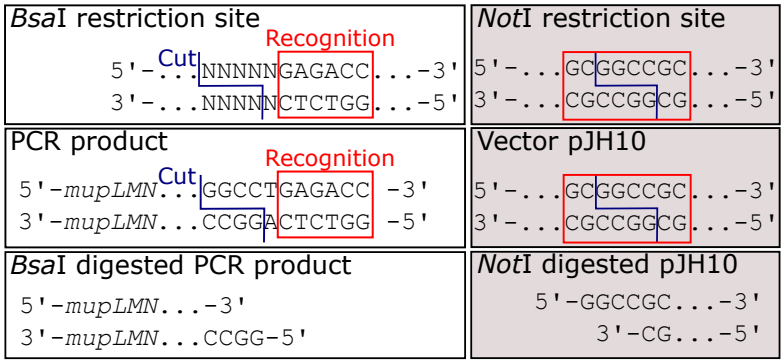


Figure 3.18: Generation of a *Not*I overhang using *Bsa*I.
As the *mupLMN* PCR product would contain an internal *Not*I site, a *Bsa*I site was used instead. *Bsa*I cuts away from its recognition site, such that a 5′ overhang of any four nucleotides can be generated. *Not*I is a more typical 8-cutter, that cuts in the middle of its recognition palindrome.

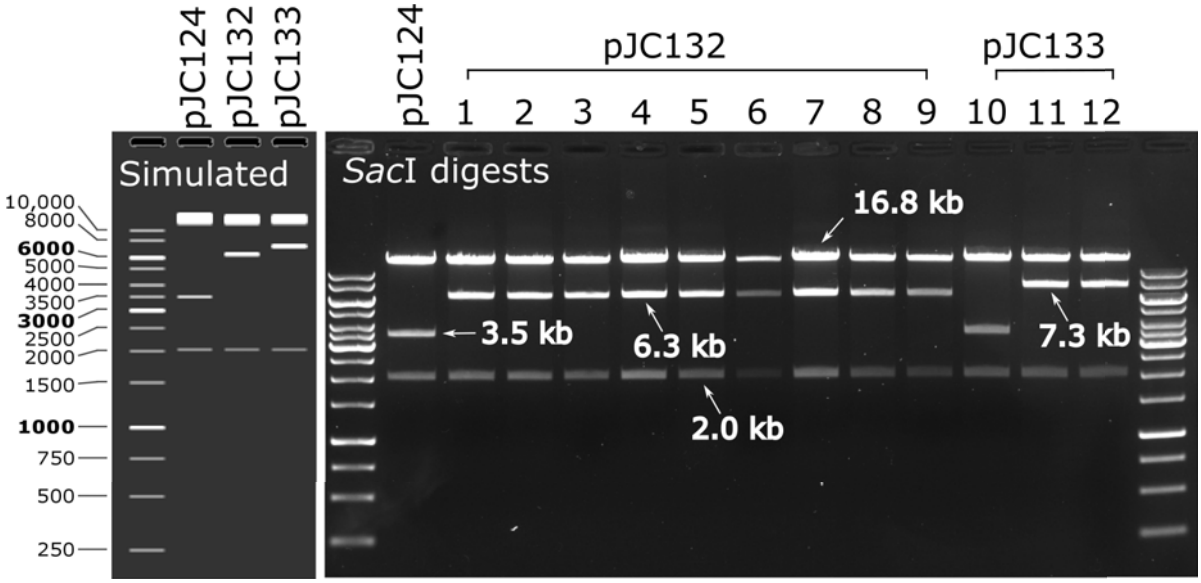


Figure 3.19: *Sac*I diagnostic digests of pJC132 and pJC133.
As digest by *Xba*I would require passage through a *dam*$^-$ strain, a *Sac*I diagnostic digest was performed instead. As seen in the simulated gel on the left, the size of the middle fragment is expected to increase when *mup(L)MN* is cloned in pJC124. All of the potential pJC132 isolates (1 to 9) have the correct 6.3 kb fragment for *mupMN* insertion. Isolates 11 and 12 have the correct restriction profile for pJC133, whereas 10 appears to be recircularised pJC124.

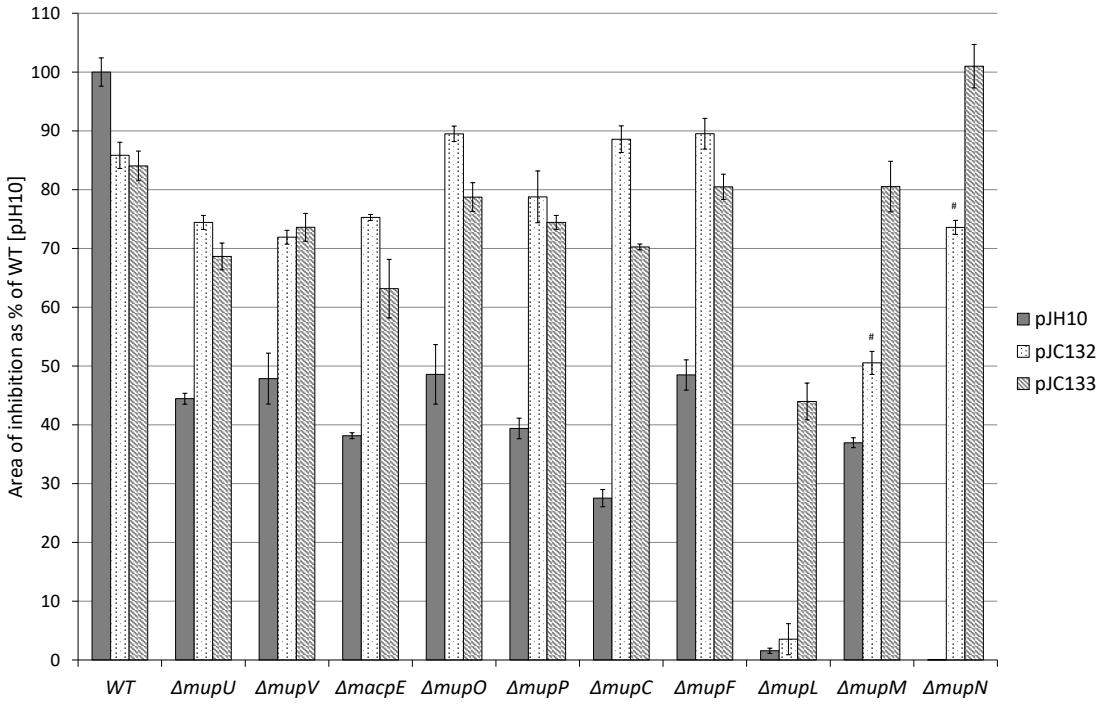


Figure 3.20: Complementation of chromosomal knockouts by pJC132 and pJC133. Bioassay areas against *B. subtilis*, expressed as percentage WT area. All combinations were performed in triplicate with 500 μM and no IPTG induction. Error bars represent standard deviation. The graph shows the best performing induction condition in each case, which was 0 IPTG for all, except Δ*mupM* [pJC132] and Δ*mupN* [pJC132], which were 500 μM IPTG (marked #).

pendently in triplicate, both with 500 μM IPTG induction and without IPTG (Figure 3.20). Complementation was observed of each knockout to approximately 65 to 90% of WT area. There were some exceptions to this, as Δ*mupL* was not complemented by pJC132, as expected since the plasmid lacks *mupL*. Complementation of the *mupL* mutation by pJC133 was observed at the lower level of 44% of WT. Plasmid pJC132 complemented the *mupM* and *mupN* mutants to a lower level than pJC133; this is putatively due to a promoter present within *mupL* in pJC133, and this is explored in Section 3.2.4.3.

Complementation of *mupL*, *mupM* and *mupN* was also confirmed by HPLC (Figure 3.21). Mirroring the bioassay results, pJC133 was observed to complement individual

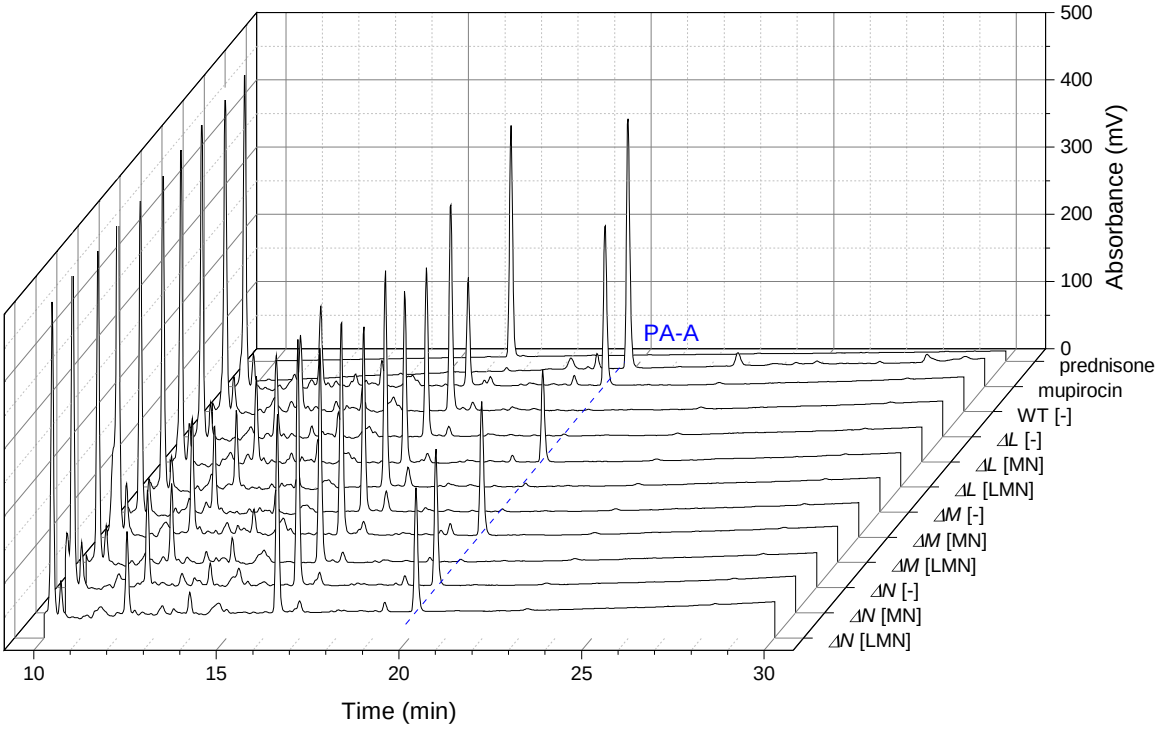


Figure 3.21: HPLC: plasmid pJC133 carries functional copies of *mupL*, *mupM* and *mupN*.
Representative HPLC chromatograms, analyses performed independently in quadruplicate without IPTG induction. Prednisone is included as an internal HPLC standard. Deletion of *mupL*, *mupM* or *mupN* is inhibitory to mupirocin production, no pseudomonic acid was observed with the negative control pJC124 [-], which is the expression plasmid prior to *mupLMN* cloning. Plasmid pJC133 [LMN] was observed to complement all three mutant strains, providing evidence that functional copies of each gene are carried on pJC133. Plasmid pJC132 [MN] was observed to complement 10586 Δ*mupN* but not Δ*mupL* or Δ*mupM*.

10586 chromosomal deletions of each of these genes. By contrast pJC132, which lacks *mupL*, did not complement the Δ*mupL* or Δ*mupM* strains, without IPTG induction.

The demonstration of complementation of the three new cloned genes *mupL*, *mupM* and *mupN* by HPLC, together with the improvement in zone of inhibition for all mutants carrying pJC133, provides evidence that functional copies of all genes are carried on the plasmid. These plasmids could now be tested for ability to provide all functions necessary for conversion of PA-B to PA-A.

3.2.3.4 Conversion of PA-B to PA-A by functions provided *in trans* on plasmid pJC133 achieved

To test if plasmid encoded conversion could be achieved, pJC132 and pJC133, along with empty vector pJH10, were mobilised to 10586 Δ*mup-cluster*. This 10586 strain has all mupirocin biosynthesis genes deleted, with the exception of the transcriptional regulatory genes *mupR*, *mupX* and *mupI*, and cannot therefore convert PA-B to PA-A (Figure 3.11). SSM cultures of 10586 Δ*mup-cluster* carrying plasmids pJH10, pJC132 and pJC133, fed with 1% (v/v) PA-B, both with and without IPTG induction, were performed in quadruplicate. SSM with 1% (v/v) methanol negative control cultures were performed in duplicate.

In HPLC analysis of supernatants of negative methanol controls (without fed PA-B), no peaks were observed at either the PA-B or PA-A positions. When PA-B was fed to 10586 Δ*mup-cluster* carrying the empty vector pJH10, no conversion to PA-A was observed (Figure 3.22). These results mirror the observations of Section 3.2.2.1 (Figure 3.11).

Gratifyingly, HPLC revealed conversion of 42% ±1.7 (s.d.) PA-B to PA-A in Δ*mup-cluster* [pJC133] without IPTG induction (Figure 3.22). No conversion was observed with the empty pJH10, or pJC132, which lacks *mupL*; this is explored in the following Section 3.2.4. No conversion was observed in any combination of strain and plasmid with 500 μM IPTG induction (not shown). This indicates that all mupirocin cluster functions required for conversion are included on plasmid pJC133, with the exception of the transcriptional regulators still present in 10586 Δ*mup-cluster*.

3.2.3.5 Further assorted short experiments with expression plasmid pJC133

With the significant milestone of *in trans* encoded conversion achieved, plasmid pJC133 facilitated several short further experiments.

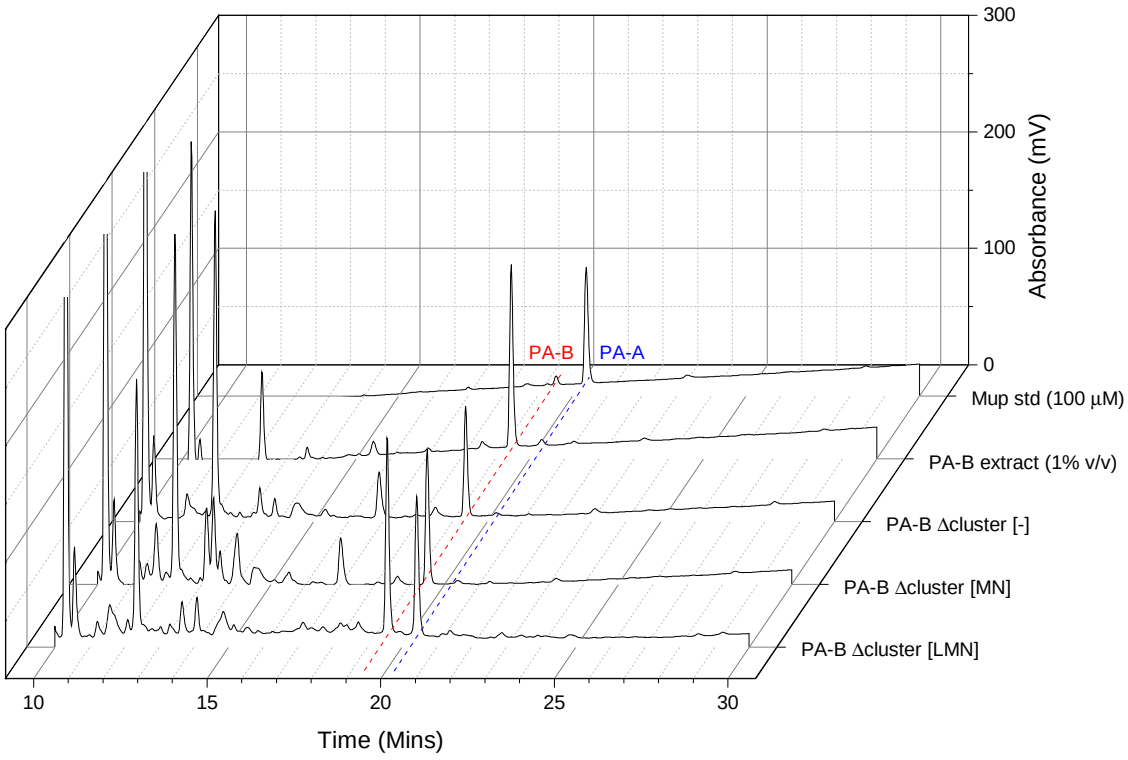


Figure 3.22: HPLC: conversion of PA-B to PA-A via plasmid encoded functions. Representative chromatograms shown, 0 IPTG, analysis performed independently in quadruplicate. Each sample was fed with 1% PA-B extract at the point of inoculation. No PA-A was observed in 10586 Δ*mup-cluster* carrying the empty vector, as it is unable to produce PA-A or convert fed PA-B to PA-A. With plasmid pJC132 *in trans* [MN], no conversion is observed. With plasmid pJC133 [LMN], conversion of fed PA-B to PA-A is restored, indicating all necessary biosynthetic genes from the mupirocin cluster have been cloned in pJC133.

Conversion does not require the genes *mupR*, *mupX* or *mupI*

As the regulatory genes *mupR*, *mupX* and *mupI* are still present in 10586 Δ*mup-cluster*, it was still possible that these could be required for pJC133 based conversion. However, this seemed unlikely, as cloned gene transcription in pJC133 is driven by the *lac*/*trp* fusion *tac* promoter, which is completely unrelated to mupirocin regulation. To eliminate this possibility, pJC133 and empty vector pJH10 were mobilised to 10586 Δ*mupR*, 10586 Δ*mupX* and 10586 Δ*mupI*. HPLC of SSM culture supernatants fed with and without PA-B, without IPTG induction, were performed in triplicate (Figure 3.23).

As expected, plasmid pJC133 was observed to provide conversion functions in

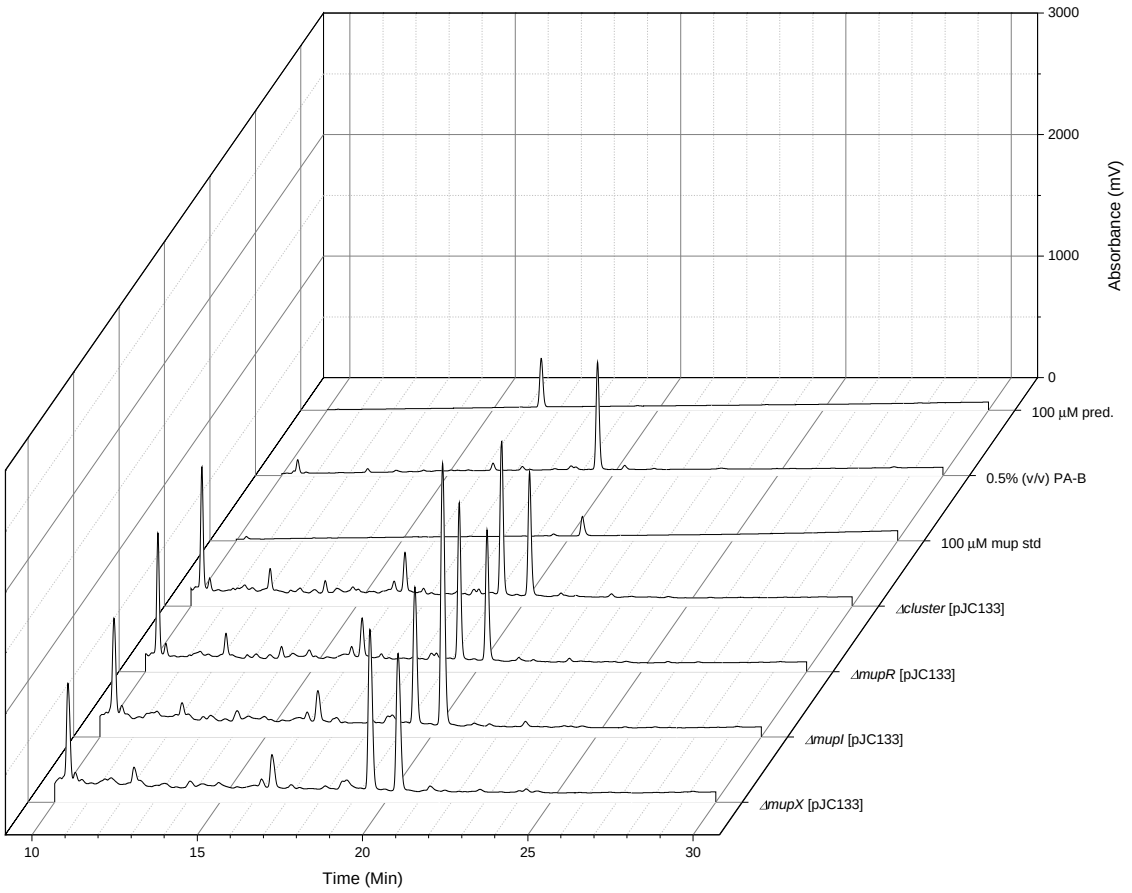


Figure 3.23: HPLC: conversion of PA-B to PA-A in *mupI*, *mupR* and *mupX* knock-outs with pJC133.
Analysis was performed without IPTG induction, independently in triplicate. Prednisone was used as an internal standard in biological samples. Each sample was fed with 0.5% (v/v) PA-B at the point of inoculation. Plasmid pJC133 can provide all functions necessary for conversion without *mupR*, *mupX* or *mupI* individually. With the exception of unlikely redundancy, this demonstrates all mupirocin cluster conversion functions are present on pJC133.

Δ*mupR* and 10586 Δ*mupX*. This demonstrates that *mupR* and *mupX* are not required for pJC133 based conversion. As detailed in Section 3.2.2.3, the Δ*mupI* mutation is complemented by the lactone present in the PA-B extract, and is therefore already able to convert to PA-A, as well as likely produce PA-A from its mupirocin pathway. In this case, pJC133 was observed to increase the proportion of PA-A from 53% ±4 (s.d.) in Δ*mupI* [pJH10] 0.5% PA-B, to 70% ±3 (s.d.) in Δ*mupI* [pJC133] 0.5% PA-B. This increase is significant, by unpaired two-tailed *t* test ($p=0.0036<0.05$).

Table 3.3: Mupirocin MICs (μg ml^{-1})

Strain	[pJH10]	[pJC133]
P. fluorescens 10586	>3000	>3000
P. fluorescens SBW25	1500	>3000
E. coli DH5α	94	188

Plasmid based conversion not achieved in alternative strains

The ability of pJC133 to provide conversion functions was also assayed in alternative strains, *E. coli* DH5α and *P. fluorescens* SBW25, both of which do not contain mupirocin clusters. LB cultures with 0.5% PA-B were inoculated with DH5α [pJC133] in triplicate, but no conversion was observed (not shown). In DH5α supernatants, a compound was observed in the negative controls at the PA-A position, which absorbed at 233 nm on the UV HPLC detection. This meant that low levels of conversion would not have been detected, but certainly there was no conversion to PA-A of the same order as in 10586 (approximately 40%). Conversion was also tested in *P. fluorescens* SBW25 [pJC133] using SSM media cultures, but no conversion was observed.

The mupirocin resistance of both *P. fluorescens* SBW25 and *E. coli* DH5α with pJH10 (empty vector) and pJC133, was determined in duplicate as detailed in Section 2.3.5. An increase in resistance was observed in both strains in the presence of pJC133, indicative of *in trans* MupM expression. However the DH5α mupirocin MIC remained low at 188 μg ml^{-1} with pJC133, which is of a similar order of magnitude to the concentration of PA-B included in feeding experiments. Therefore, if DH5α were able to convert to PA-A, it seems likely that the antibiotic would be self-inhibitory. The poor performance of *mupM* expression in *E. coli* may suggest that the other nine genes are unlikely to all function efficiently as well.

Expression plasmid pJC133 is not immediately suitable for *in vitro* protein experiments

To test if pJC133 might prove suitable for *in vitro* protein experiments, *E. coli* DH5*α* [pJC133] and *E. coli* BL21 [pJC133] cultures were grown at various temperatures and induction conditions. Cells were lysed by sonication, and run on SDS-PAGE gels with coomassie staining. DH5*α* [pJC133] was compared with DH5*α* [pJH10] for each condition, and no difference in banding was observed in any case (not shown). As no difference was observed to the empty vector, there was no indication of high scale protein expression. To explore *in vitro* protein experiments in the future, a sensible approach would be to clone the individual genes into a generic *E. coli* protein expression vector such as pET28a.

3.2.4 Investigation of the roles of *mupL* and *mupM* in conversion of PA-B to PA-A

In screening 10586 mutants for conversion of PA-B to PA-A, the *mupL* mutant demonstrated a reproducibly variable phenotype (Figure 3.24). Three independently generated 10586 Δ*mupL* strains were included in the conversion tests, performed by PA-B feeding and monitoring by HPLC (as described in Section 3.2.2.2). For each of these independent strains, it was observed that some cultures would convert PA-B to PA-A, but others would not. This demonstrates that the phenotype is unlikely to be caused by a random secondary mutation to another mupirocin gene. Therefore, the *mupL* gene is needed to allow consistent conversion of PA-B to PA-A.

Similarly, plasmid pJC133, carrying ten genes including *mupL*, was able to provide all functions required for conversion (Section 3.2.3.4), whereas pJC132, a sister expression plasmid carrying nine genes without *mupL*, failed to restore conversion in 10586 Δ*mup-cluster*, n=6. This indicated that the *mupL* gene is strictly required for *in trans* conversion in this configuration. However, this did not agree logically with the screening experiments, wherein approximately 50% of the cultures conversion would

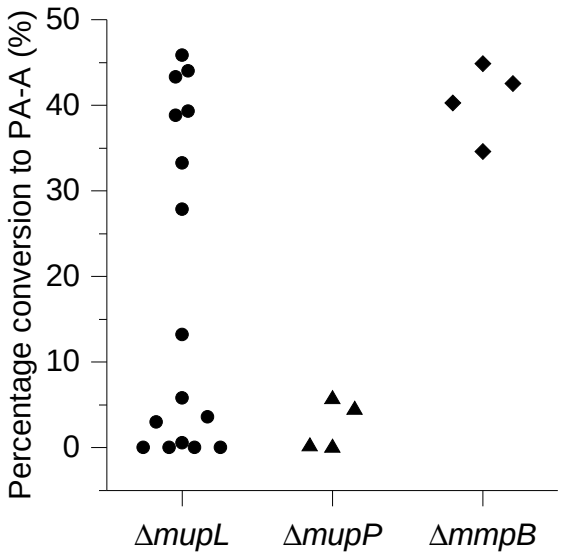


Figure 3.24: Variable levels of conversion observed in the Δ*mupL* mutant. Conversion to PA-A measured by integrated HPLC peak area, as detailed in Section 3.2.2.2. Each data point represents an independent culture and analysis; X-axis variation is simply to prevent overlap of data points. The Δ*mupP* mutant is known to be deficient in conversion (Gao et al., 2017), whereas the Δ*mmpB* mutant is included as a positive converter.

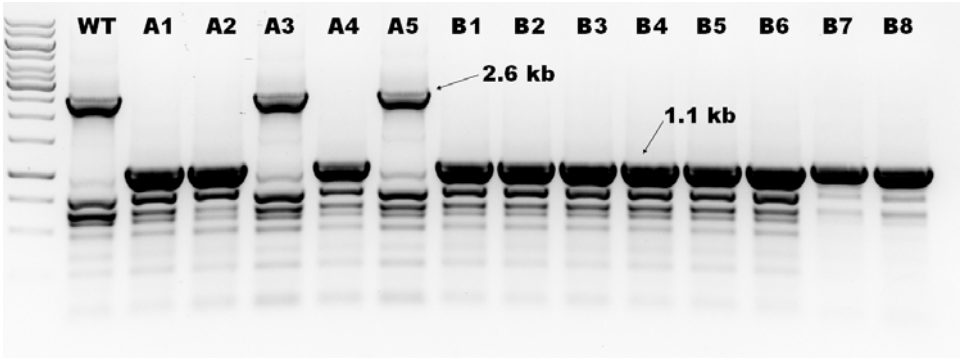
occur in the absence of *mupL*.

This led to investigation of whether MupL enzymatic activity is required for conversion, and whether there were effects on expression of downstream gene *mupM*.

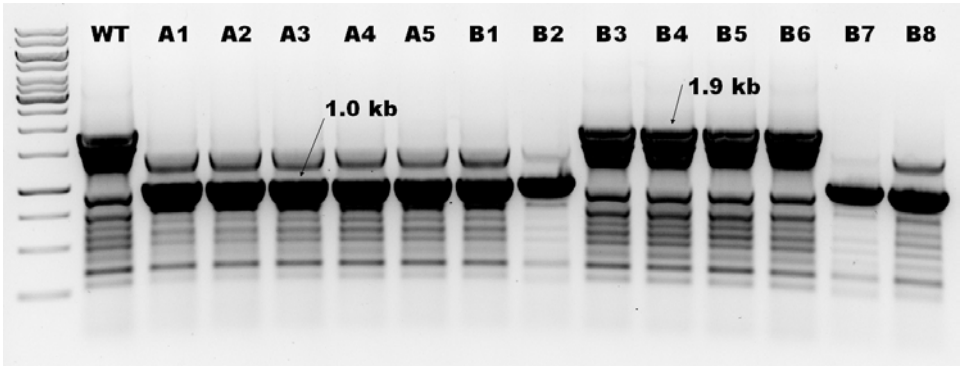
3.2.4.1 MupL first acts before MupU in mupirocin biosynthesis

The relative order of mupirocin biosynthesis enzyme action has previously been explored by dual mutations. Starting with two individual mutants with different mupirocin production phenotypes, if a dual mutant mirrors one of these phenotypes, it indicates that the mutated enzyme acts first in the pathway. This approach was used to demonstrate MupU acts before MupC (Hothersall et al., 2007).

A similar experiment was conceived to determine whether MupL or MupU acts first in mupirocin biosynthesis. MupU (Acyl CoA synthase) putatively loads MacpE with PA-B prior to conversion; 10586 Δ*mupU* produces PA-B as the primary metabo-



(a) *mupU* PCR



(b) *mupL* PCR

Figure 3.25: Screen for double *mupL* and *mupU* knockout 10586 mutants. Strains A1-5 are 10586 Δ*mupL* mutants that have undergone suicide mutagenesis with p604/Δ*mupU*. Stains B1-8 are the inverse, starting Δ*mupU*, screening for *mupL* mutation. Hence all of B1-8 have the 1.1 kb deleted *mupU* size PCR product, compared with 2.6 kb WT *mupU*. This PCR screening revealed strains A1, A2, A4, B1, B2, B7 and B8 are all the desired double mutants.

lite. The role of MupL (putative hydrolase) in mupirocin biosynthesis is unknown.

To generate a double knockout mutant, individual mutants 10586 Δ*mupL* (Hothersall et al., 2007) and 10586 Δ*mupU* (Cooper et al., 2005) were used as starting points. The second knockout was generated in each strain using the mutagenesis plasmids pAKE604/Δ*mupU* and pAKE604/Δ*mupL* respectively. Suicide mutagenesis was performed as described in Section 2.2.12. After mutagenesis plasmid excision, mutants were identified by PCR screening with primers mupL1F, mupL1R, and mupU1F, mupU1R (Figure 3.25).

To test the phenotype of the double mutant, 10586 Δ*mupL* Δ*mupU*, SSM cultures and HPLC analysis were performed in quadruplicate (isolates A1, A4, B1, B8). For

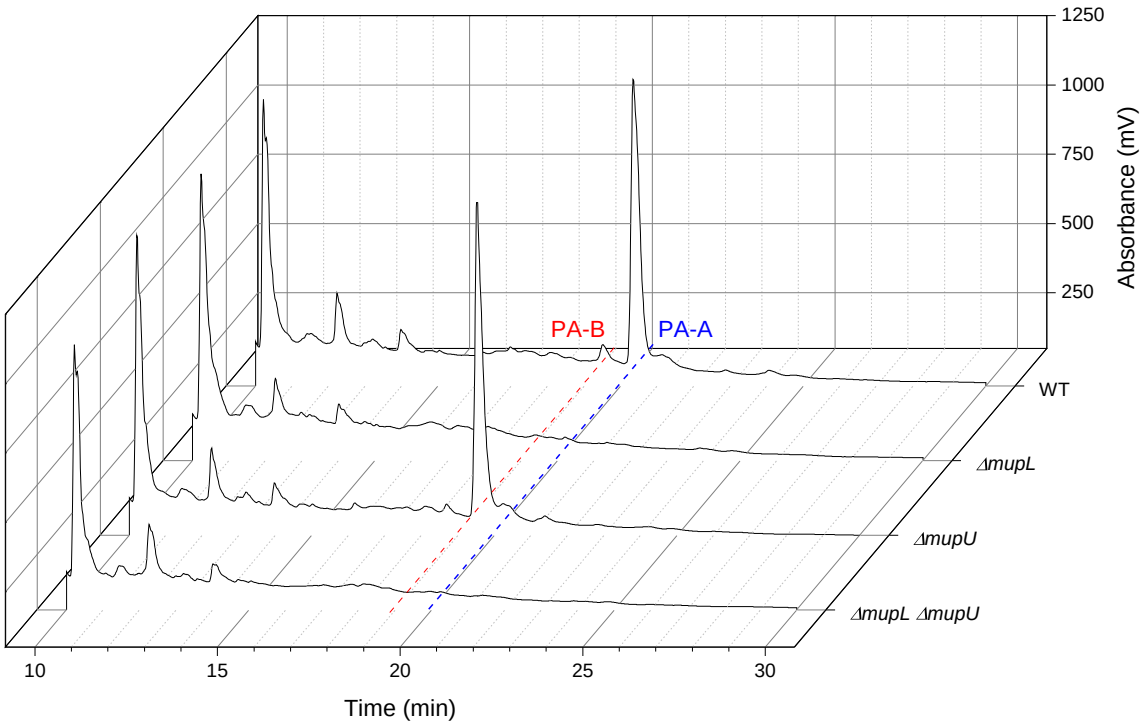
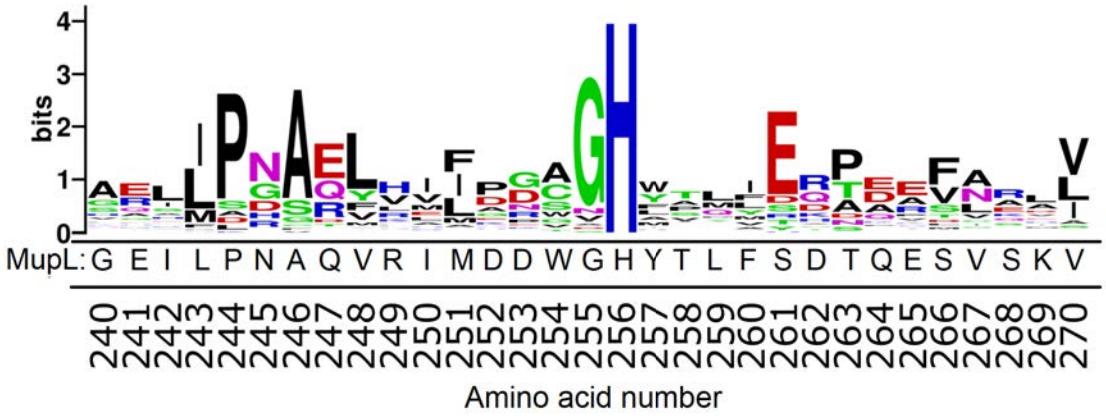


Figure 3.26: HPLC: 10586 Δ*mupU* Δ*mupL* does not produce pseudomonic acid B. Representative chromatograms shown, experiment performed independently in triplicate, except Δ*mupL*Δ*mupU* that was performed in quadruplicate. Single mutant 10586 Δ*mupL* does not produce pseudomonic acid, whereas 10586 Δ*mupU* primarily produces PA-B. The double mutant 10586 Δ*mupL* Δ*mupU* does not produce pseudomonic acid, mirroring the Δ*mupL* phenotype.

comparison, 10586 WT, 10586 Δ*mupL* and 10586 Δ*mupU* were performed in triplicate. HPLC analysis revealed that the double mutant phenotype mirrored the Δ*mupL* phenotype, with no pseudomonic acid production observed (Figure 3.26). This suggested that MupL acts first before MupU in the pathway. Therefore, MupL would have to act multiple times in mupirocin biosynthesis if it were to also act after MupU in the conversion of PA-B to PA-A.



(a) MupL homolog amino acid sequence logo

```
#HMM     epkavvvlvHGlaehsgryaelaeaLaaqGfavyalDhrGhGrsegkrghvpsfddyvdDlaafvekvkeehpeaplfllGhSmGGliallyllrypdkvaglils.aPalkikaklk.appllkllkl
#MATCH   +p ++vv+v+ l+      + +a+aL+a+G+ ++++  rG       r +   +dd v+Dl+ ++ +++      p++l +  +G  +al+y+ r p+++++l l+ +P++    +   a  +   ++++
#PP      5778999999999999999**********************9999***************99987777....559************************777763677766655444444444444555
#SEQ     APGGTVVFVNALGIGNAVTQGVAQALQAAGIGFVTWERRGLPGAYDDRFREYALDDQVRDLEHLMVMLA----PGPVVLAAWCTGIHTALAYARRWPARLEALLLFnSPNFFNSRYSAvAGDAIGKVSRI

#HMM     lgkllpklqv.........snnlkadslsrdpevvaayeadplv.hgkitvrllllellleamkrleeraakit..lPllllhGeaDkvvdpeasrelyerlgskdktlklypglyHelln
#MATCH   l +   kl +         ++++ ++      ++++ + + p+  + +  +r ++ +  a++  ++r++ +   +P+l++ G +D+ v+ ++s  l e l+  + ++++++ + H +l+
#PP      5444444444888888887666677777777888999999999988888888888888888888887654455**************************..6999999999998774
#SEQ     LVEDESKLDFmyhtifannTEQTATRLTGLADARIRGLVEAPFKsGPQALLRYMYLIQSAAEPEVDRQWCAPisVPVLVIGGRQDTMVSFQDSLSLGEILP--NAQVRIMDDWGHYTLF
```

(b) MupL Pfam output

Figure 3.27: Analysis of MupL homologs.
(a) The sequence logo shows the amino acids near the His256 chosen for mutation to alanine. The numbering is relative to MupL.
(b) The Pfam output shows the level of conservation of the full MupL protein sequence. The lines are labelled as follows: #HMM is the hidden Markov model consensus, #MATCH represents the matches between the HMM and MupL sequences, #PP is a confidence degree (higher is better), and #SEQ is the MupL sequence.

3.2.4.2 Enzymatic activity of MupL is not required for conversion

To test whether MupL enzymatic activity is required for conversion, versions of pJC133 carrying point mutations to *mupL* were generated. MupL is a predicted α/β hydrolase that does not show strong homology to any well characterised enzyme. Consequently, sequence conservation was used to choose mutation targets, rather than relying on experimental evidence from homologs, as was the case with MupV domain 1.

An alignment of 42 proteins homologous to MupL was generated, manually curated for diversity. This revealed a histidine is completely conserved within the alignment at position 256 in MupL, which was chosen as a mutation target to alanine (Figure 3.27).

Construction of the *mupL* point mutant expression plasmid

The mutation was generated in a similar fashion to that used to insert *mupL, mupM,* and *mupN* into pJC124 to generate pJC133 (Section 3.2.3.2). Primers P25_F_-mupLMN and P31_R_mupL_H256A were used to amplify an 809 bp fragment. Primers P31_F_mupL_H256A and P25_R_mupLMN were used to amplify a 4313 bp fragment. These were designed to overlap, carrying the H256A mutation (DNA codon CAC mutated to GCC). The two fragments were stitched by SOEing PCR using the outer primers P25_F_mupLMN and P25_R_mupLMN, to generate a 5102 bp fragment.

This fragment carried *mupL* (with H256A mutation), *mupM* and *mupN*, with *Xba*I and *Bsa*I (*Not*I overhang) restriction sites. The overlapped fragment was AT cloned in pGEM-T, yielding intermediate plasmid pJC145. Correct insert size was confirmed by *Xba*I/*Bsa*I digest, and sequencing performed using primers M13(-40) and M13(-48).

The insert was released by *Xba*I/*Bsa*I digest, and cloned into pJC124 *Xba*I/*Not*I, isolated from *dam*$^{-}$ *E. coli* ER2925. This resulted in expression plasmid pJC133.H256A, carrying *mupO, P, macpE, mupU, V, C, F, L, M, N*, with point mutation H256A to *mupL*. The presence of insert of the correct size was detemined by *Sac*I digest (Figure

3.28). The construction technique used exactly mirrored the cloning technique used previously to generate pJC133 (Figure 3.17).

The point mutation H256A is disruptive to MupL enzymatic function

Plasmid pJC133.H256A, together with pJC133 and empty pJH10 as controls, were mobilised to 10586 Δ*mupL* to check for disruption of MupL enzymatic function. As *mupM* and *mupN* had been derived from PCR product in construction of pJC133.H256A, the plasmids were also mobilised to 10586 Δ*mupM* and Δ*mupN*, to check for ability to complement, therefore detecting any significant PCR errors.

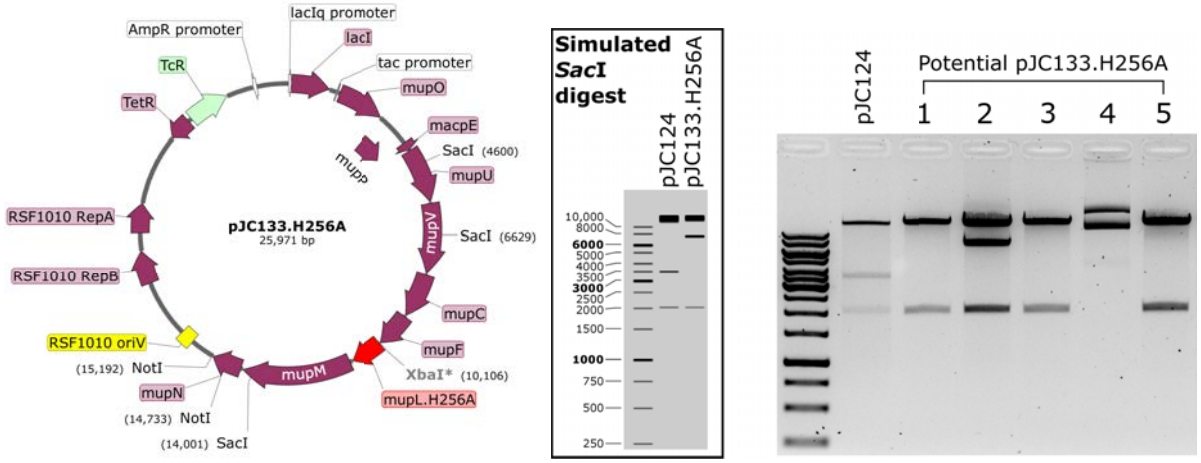
Antibiotic activity produced by all combinations of strain and plasmid were tested by plate bioassay against *B. subtilis* (Figure 3.29). As expected, 10586 Δ*mupN*, 10586 Δ*mupM*, and 10586 Δ*mupL* were deficient in active mupirocin production, as indicated by the decreased zones observed with empty vector pJH10. Both pJC133 and pJC133.H256A were observed to complement the Δ*mupM* and Δ*mupN* mutations, which demonstrated that pJC133.H256A carries functional copies of both of these genes.

Plasmid pJC133 complemented the Δ*mupL* mutation to 70% of WT area, whereas no complementation was observed for pJC133.H256A. This provided evidence that mutant version of MupL is not enzymatically functional, either because the conserved His256 is required for function or because the mutant version no longer folds correctly.

Plasmid pJC133.H256A can still provide all functions required for conversion

Plasmid pJC133.H256A, along with controls pJH10 and pJC133, were mobilised to *P. fluorescens* Δ*mup-cluster*. SSM cultures were grown with 0.5% PA-B, and the supernatants analysed by HPLC (Figure 3.30). Conversion to PA-A was still observed with inactivated MupL, at the same level as with functional MupL present. This clearly indicated that MupL enzymatic function is not required for conversion of PA-B to PA-A.

The observed conversion phenotypes for the different *mupL* experiments are sum-



(a) Plasmid map **(b)** *Sac*I digest gel

Figure 3.28: Map and digest of plasmid pJC133.H256A, carrying point mutation to *mupL*.
Isolate 2 has the correct *Sac*I restriction profile for pJC133.H256A. Isolates 1, 3, and 5 appear to have lost a *Sac*I site, most likely recircularised *Xba*I/*Not*I digested pJC124. Isolate 4 has an anomalous profile.

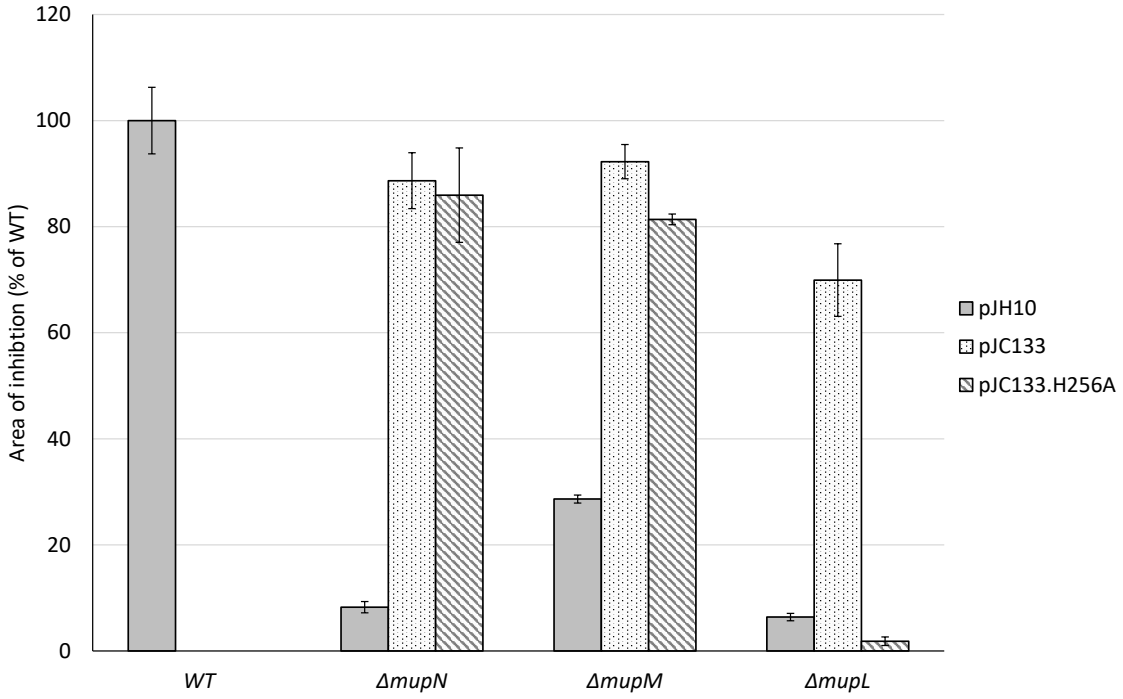


Figure 3.29: The H256A mutation results in loss of MupL function.
Bioassay zones of inhibition against *B. subtilis*, expressed as percentage area of WT [pJH10]. WT, Δ*mupN* and Δ*mupM* experiments were performed independently in triplicate, for Δ*mupL* n=9. Error bars represent the standard error of the mean.

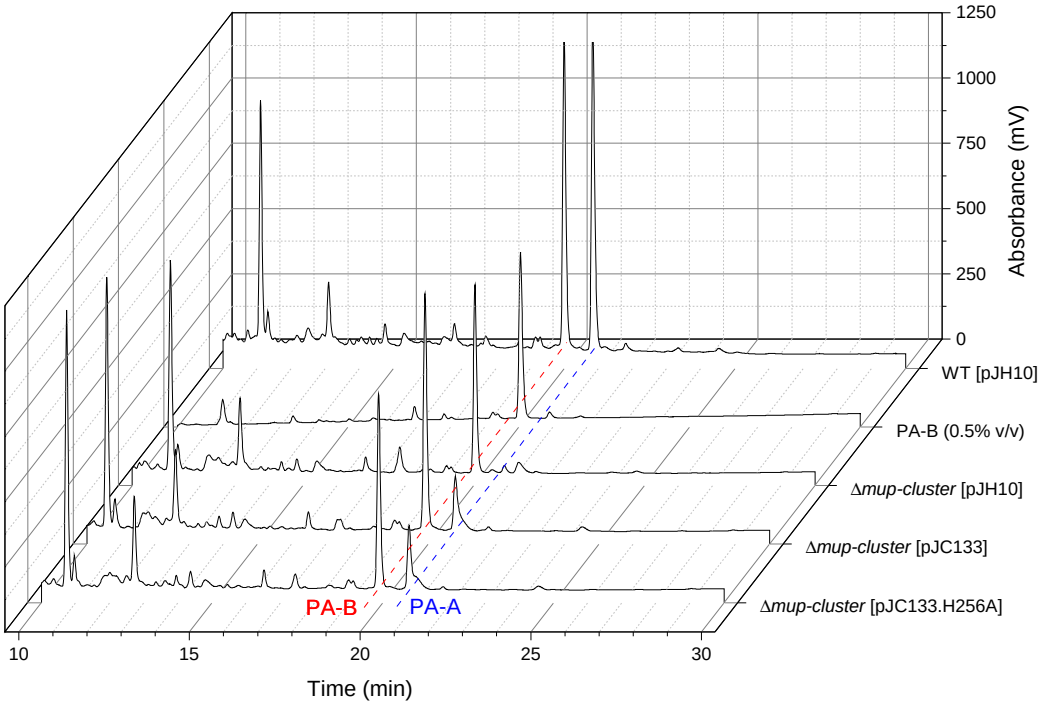


Figure 3.30: HPLC: pJC133.H256A can still provide all functions required for PA-B to PA-A conversion.
Representative HPLC chromatograms shown, analysis performed independently in quadruplicate. Each sample was fed with 0.5% (v/v) PA-B at time of inoculation. No conversion to PA-A was observed in 10586 Δ*mup-cluster* with empty vector, pJH10. Plasmid pJC133.H256A, carrying an enzymatically inactive version of MupL, was still able to restore conversion to the same level as seen in pJC133.

marised in Figure 3.31.

3.2.4.3 Deletion of *mupL* decreases *mupM* expression

The experimental evidence presented so far has indicated plasmid pJC133 requires the presence of *mupL* for conversion to occur, but does not require MupL enzymatic function. MupL first acts before MupU in the biosynthetic pathway, and when *mupL* is deleted in the 10586 chromosome, a variable conversion phenotype is observed. A potential explanation is that *mupL* contains a promoter that is required to increase downstream *mupM* and *mupN* expression (Figure 3.32). There is a putative promoter sequence within *mupL*, `5'-TACACA-18nt-TATGCT-3'`, positioned 200 bp before the *mupM* start.

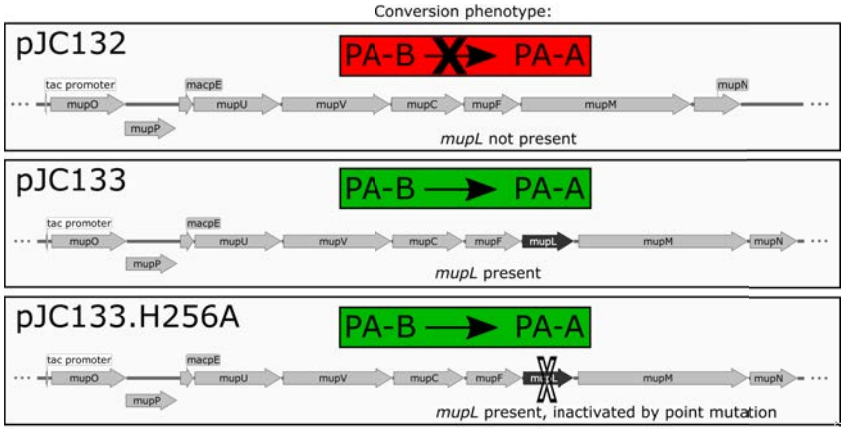


Figure 3.31: Summary of plasmid based *mupL* PA-B conversion phenotypes. Plasmid pJC132, which lacks *mupL*, is unable to convert fed PA-B to PA-A, whereas pJC133, with WT *mupL*, is able to provide conversion functions. When MupL is inactivated by point mutation on pJC133.H256A, conversion still occurs. These results suggest that for conversion to occur in this configuration, *mupL* must be present, but MupL enzyme function is not necessary.

In the *P. fluorescens* 10586 cluster, there is an 82 bp gap between *mupL* and *mupM*, whereas in the *P. psychrotolerans* NS383 cluster, *mupL* and *mupM* overlap (Figure 3.33). There is an 87 bp open reading frame in the gap in 10586, out of frame with *mupL*, that shows some homology at the C-ter of NS383 *mupL*. The C-ter is not conserved in the thiomarinol homolog *tmlL* either, or in distant homologs. The conserved functional structure may not continue to the end of the NS383 *mupL*, which would facilitate the genetic divergence observed.

The potential promoter within *mupL* could be detected by comparison of MupM levels in plasmids with *mupL* (pJC133) and without *mupL* (pJC132). If *mupL* contains a promoter that is functional in the culturing conditions used, then pJC133 should provide more MupM than pJC132. This could be detected by complementation of either of the Δ*mupM* phenotypes, decreased mupirocin resistance and production.

Plasmid pJC133 can complement PA-A production in 10586 Δ*mupM* without IPTG induction, observed by HPLC (Figure 3.34a) and previously by bioassay (Figure 3.20). However, IPTG induction is required for complementation by pJC132 (Figure 3.34b). This is consistent with the previous bioassay results in Figure 3.20. These results

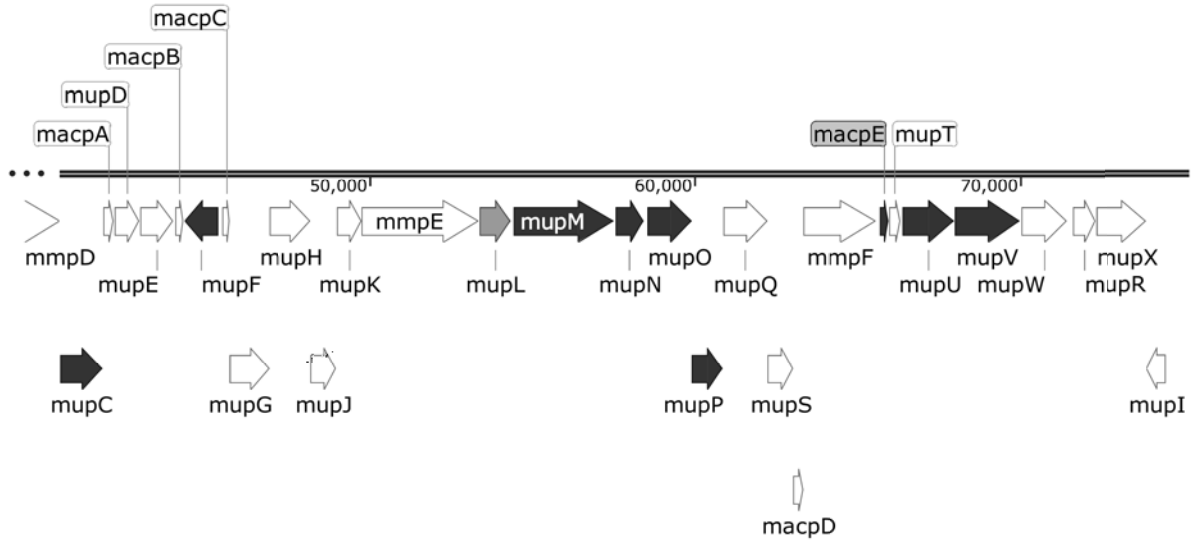


Figure 3.32: Map of the chromosomal 10586 mupirocin gene cluster region of the conversion genes.
The nine genes required for conversion of PA-B to PA-A are shaded in black, *mupL* is shown in grey. The putative promoter within *mupL* would start transcription into *mupM* then *mupN*. The *mupL, M, N* operon is also in place in pJC133.

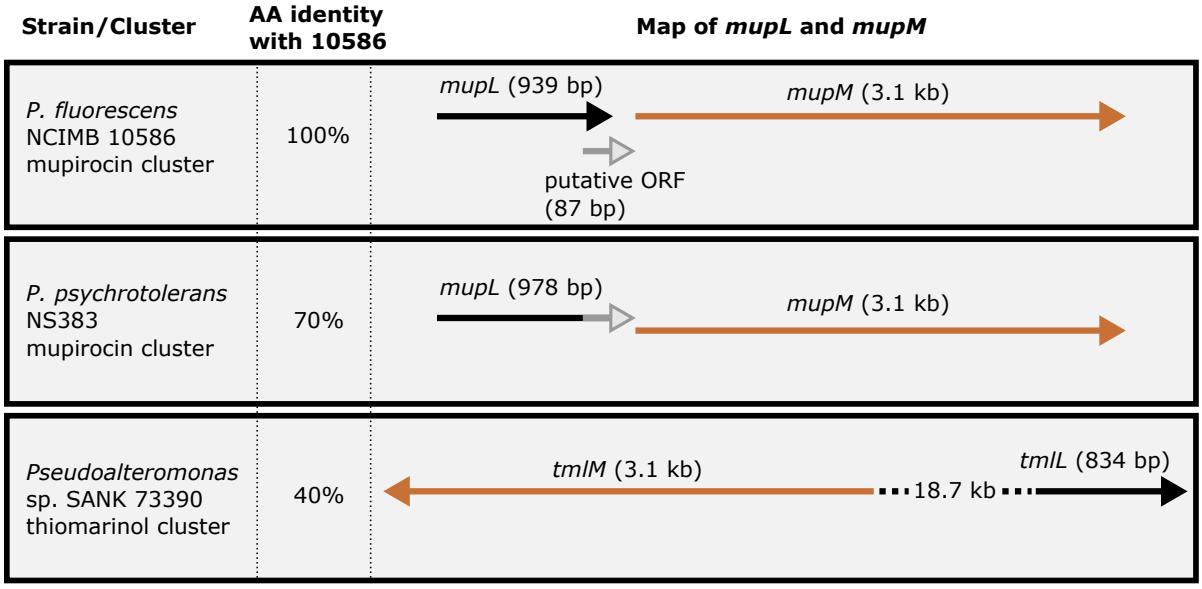
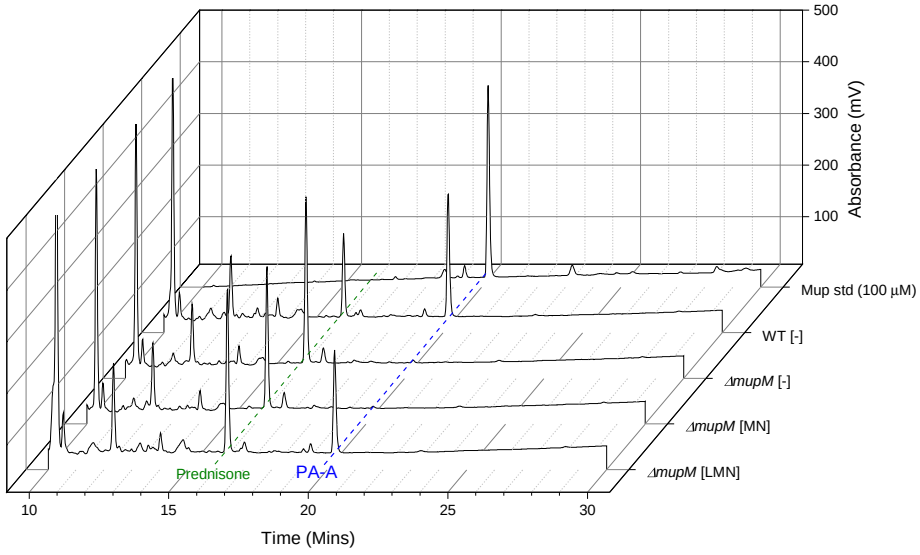
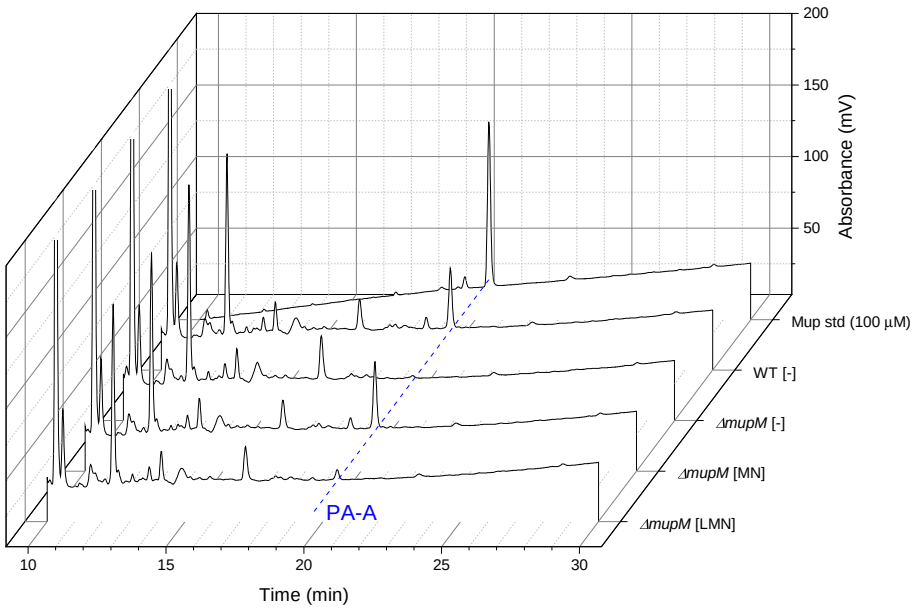


Figure 3.33: Relative genetic position of *mupL*, *mupM*, and their homologs.



(a) No IPTG



(b) 500 μM IPTG

Figure 3.34: HPLC: IPTG induction is required for pJC132 to complement 10586 Δ*mupM*.
Parts **(a)** and **(b)** were performed separately on different HPLC columns, with different injection volumes: **(a)** 50 μl, **(b)** 25 μl. Part **(a)** is a subset of the data from Figure 3.21, included for comparison purposes, prednisone was used as an internal standard. Representative chromatograms shown, **(a)** analysis performed independently in quadruplicate, **(b)** independently in duplicate.
Plasmid pJC132 [MN], can complement 10586 Δ*mupM* with IPTG induction, but cannot without (Figure 3.21). However, pJC133 [LMN], which does complement Δ*mupM* without IPTG, complements less well with induction.

support the hypothesis that *mupL* is required for conversion because of an internal promoter that drives *mupM* expression. When *mupL* is not present, IPTG induction is needed to increase *mupM* expression to the required level.

To test for complementation of the Δ*mupM* increased mupirocin sensitivity phenotype, minimum inhibitory concentrations (MIC) were recorded of mupirocin against 10586 Δ*mupM* strains, carrying plasmids pJC132, pJC133 and negative control pJC124 (precursor plasmid that lacks *mupL*, *mupM* and *mupN*) (Figure 3.35).

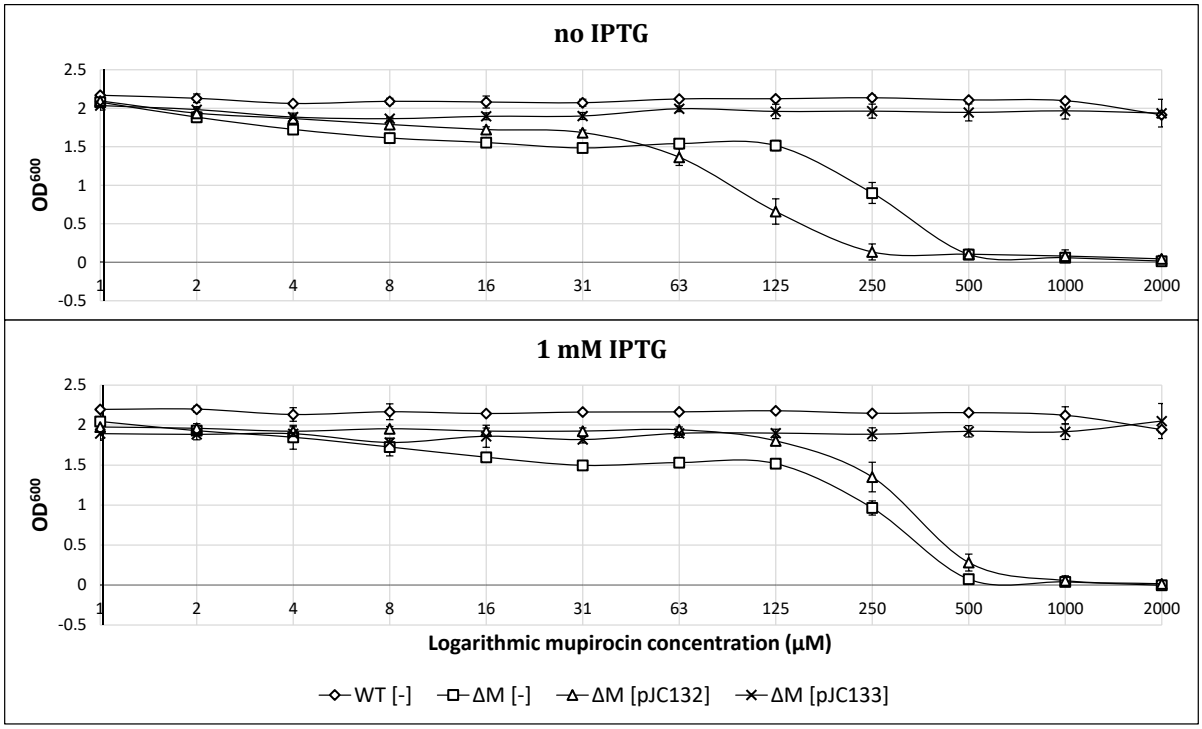


Figure 3.35: Presence of *mupL* is required for complementation of the decreased mupirocin MIC of 10586 Δ*mupM*.
Absorbance readings of 10586 LB cultures grown in a 96-well plate for 16 hours, with serial dilutions of mupirocin. The Δ*mupM* experiments were performed in quadruplicate, and the WT in duplicate. Error bars represent standard deviation. Plasmid pJC133 was observed to complement mupirocin resistance of 10586 Δ*mupM* to WT levels, with or without IPTG induction. Plasmid pJC132, which lacks *mupL*, is unable to complement Δ*mupM*.

The WT strain with negative control plasmid pJC124, was fully resistant to the highest concentration of mupirocin tested, 2 mM, which was limited by solubility in LB media. The Δ*mupM* strain with control pJC124, demonstrated decreased resistance, with an MIC of 500 μM. Plasmid pJC133 complemented this deficiency to the WT

level of >2 mM, both with and without IPTG induction; whereas pJC132 suprisingly decreased the MIC to 250 μM without IPTG induction, and remained the same at 500 μM with induction.

Considering the mupirocin production deficiency phenotype, IPTG induction can increase MupM production in the absence of *mupL* to the required levels. However, complementation of the MIC phenotype obligately required the presence of *mupL*. This could be because of how the different experiments are performed: with the MIC determination mupirocin is introduced immediately to the cultures, whereas when testing for production, the culture media initially contains no mupirocin. In this case, mupirocin is only present when production switches on at stationary phase, allowing time for expression of sufficient *mupM* from pJC132 with IPTG.

10586 Δ*mupL* is deficient in growth at high mupirocin concentrations

As a counterpart to the plasmid based experiments, the effect of a chromosomal deletion of *mupL* on MupM production was assayed through analysis of mupirocin resistance. The MIC of 10586 Δ*mupL* was observed to remain at the WT level of >2 mM, with testing of higher concentrations limited by solubility (Figure 3.36). However, at 2 mM mupirocin the OD^{600} of 10586 Δ*mupL* was 1.4 $\pm$ 0.2 (s.d.), compared with 2.4 $\pm$ 0.1 in WT.

This could be explained by the deletion of the promoter within *mupL* decreasing constitutive *mupM* expression. However, we would still expect *mupM* expression at stationary phase once the mupirocin cluster is switched on, driven by the upstream mupirocin regulated promoters of the *mmpE* operon. A futher test would be to repeat this experiment whilst monitoring over a time course, as the model predicts that Δ*mupL* will exhibit a lack of initial growth in the presence of mupirocin that could later recover once MupR regulated *mupM* expression is turned on.

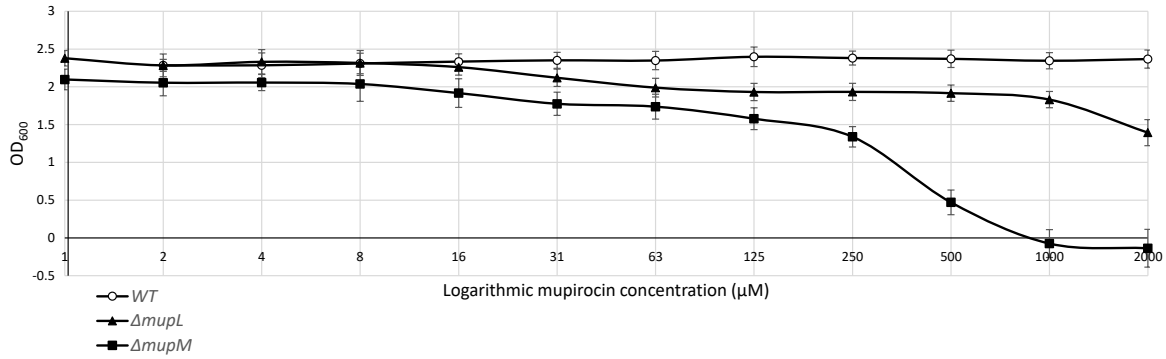


Figure 3.36: Chromosomal deletion of *mupL* decreases level of growth in presence of mupirocin.
Absorbance readings of 10586 LB cultures grown in a 96-well plate for 18 hours, with serial dilutions of mupirocin. Experiments were performed in quadruplicate. Error bars represent standard deviation.

3.2.4.4 Non-converting Δ*mupL* cultures display loss of culturability at 44 hours

When 10586 Δ*mupL* is grown with exogenous PA-B, it was consistently observed to variably convert this to PA-A (Figure 3.24). However, the chromosomal Δ*mupL* Δ*mupU* dual mutant (Section 3.2.4.1) and *in trans mupL* point mutant experiments (Section 3.2.4.2) demonstrated that MupL enzymatic function is not required for conversion of PA-B to PA-A. In the previous Section 3.2.4.3, evidence was presented for the presence of a promoter within *mupL* that drives expression of *mupM*. This could be the cause of the variable conversion phenotype observed in 10586 Δ*mupL*.

To test this hypothesis, growth of 10586 Δ*mupL* with exogenous PA-B was analysed. Monitoring the absorbance of cultures was considered unsuitable, as *P. fluorescens* 10586 produces mupirocin at stationary phase, and the production media (SSM) is opaque with solid precipitates. Instead, viable cells were analysed by colony forming unit (CFU) count, using the method described in Section 2.3.6. SSM cultures were inoculated with 10586 Δ*mupL*, either with 0.5% PA-B or 0.5% methanol as a negative control (n=6). CFUs were counted at 16 and 44 hours (Table 3.4). At 16 hours, all cultures yielded CFUs of between the order of 10^8 and 10^9 cells ml^{-1}, indicative of growth to near saturation. At 44 hours, half of the PA-B cultures had collapsed to between 10^3 and 10^5 CFU ml^{-1}, whereas all other cultures were saturated at 10^{10}.

Table 3.4: CFU estimates of 10586 Δ*mupL* cultures with exogenous PA-B

(a) CFU count (cells ml^{-1}) + MeOH (0.5% v/v)

	#1	#2	#3	#4	#5	#6
16 hours	5.3×10^9	2.2×10^9	1.8×10^9	5.6×10^9	4.4×10^9	2.5×10^9
44 hours	3.1×10^{10}	3.4×10^{10}	4.0×10^{10}	3.8×10^{10}	3.0×10^{10}	4.8×10^{10}

(b) CFU count (cells ml^{-1}) + PA-B (0.5% v/v)

	#1	#2	#3	#4	#5	#6
16 hours	1.3×10^9	7.5×10^8	1.4×10^9	8.5×10^8	3.5×10^8	3.0×10^8
44 hours	2.9×10^{10}	2.2×10^{10}	2.7×10^{10}	2.5×10^3	4.0×10^5	2.0×10^3

MeOH and PA-B samples are not paired

At 44 hours, cultures were imaged by phase contrast microscopy to analyse cell morphology and numbers (Figure 3.37). Despite a more than five orders of magnitude decrease in CFUs, no difference in cell counts were observed between the low and high CFU groups. This suggests the low CFU group of cultures are not colony forming rather than low in absolute cell count.

The percentage conversion of fed PA-B to PA-A in supernatants of the six cultures was measured by HPLC (Figure 3.38). Half the cultures were observed to convert, averaging 37% $\pm$5 (s.d.), and these were observed to contain high CFUs at 44 hours (Table 3.4). The low CFU cultures failed to convert (1.3% $\pm$1.5).

Given the evidence for the Δ*mupL* mutation decreasing self-resistance gene *mupM* expression, it seems likely that the decreased level of MupM is the cause of the collapse in cell culturability. However, why this would vary between cultures and why this would occur on a culture-wide level is yet unclear. This is discussed, together with the extensive potential future experimentation, in Section 3.3.3.3.

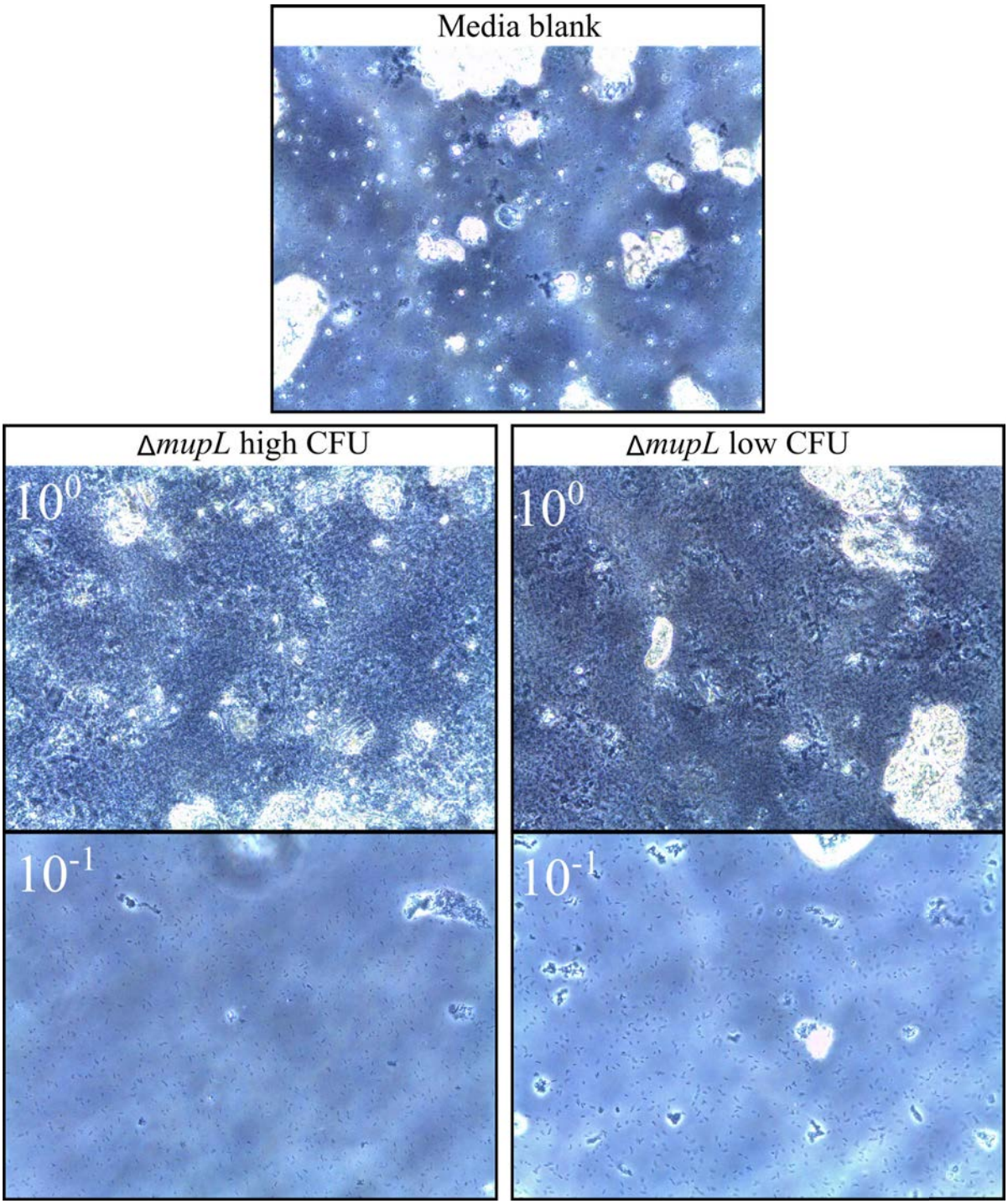


Figure 3.37: Microscopy of 10586 Δ*mupL* cultures with exogenous PA-B.
Images are 40^x phase contrast. No difference was apparent between the 10586 Δ*mupL* cultures yielding low and high CFUs. 10^0 and 10^{-1} represent undiluted and a ten-fold dilution of cultures respectively.

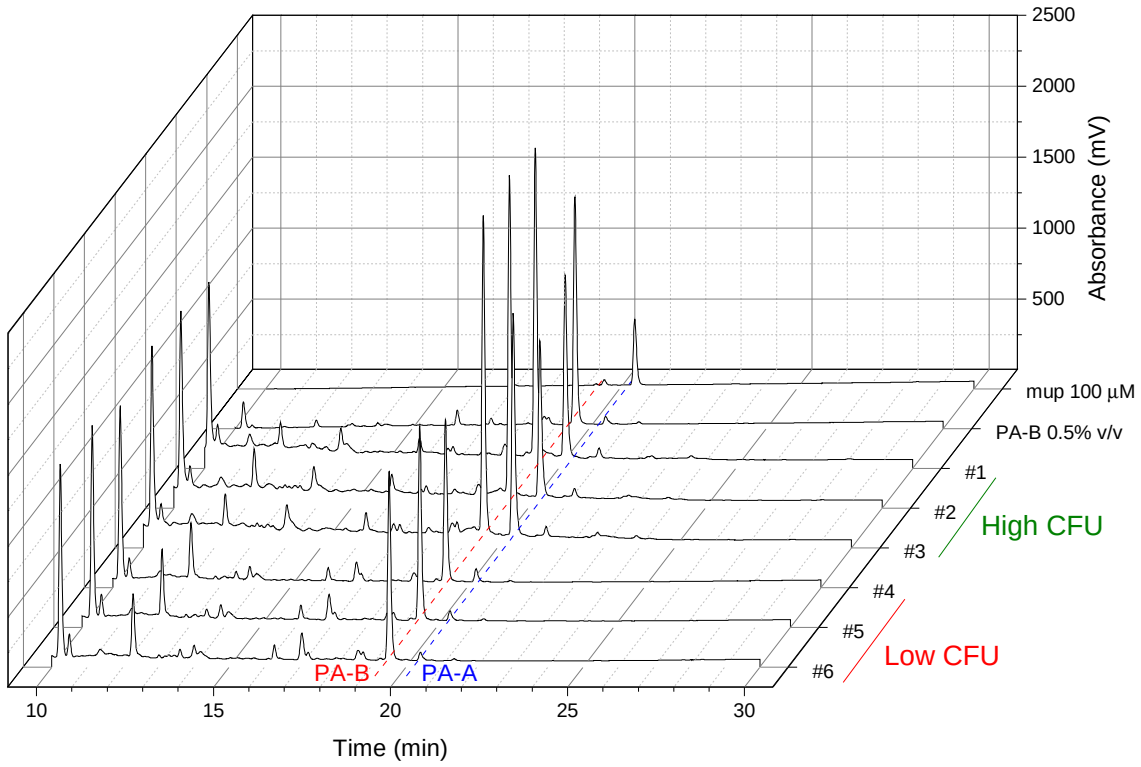


Figure 3.38: HPLC: conversion of PA-B to PA-A by 10586 Δ*mupL* correlates with CFU count.
Samples labelled #1 to #6 are 10586 Δ*mupL* with 0.5% PA-B; these are the same samples from the CFU count experiment (Table 3.4).

Investigation of the growth phenotype of 10586 Δ*mupM*

When 10586 Δ*mupM* was first tested in Hothersall et al. (2007), our research group noted the decreased PA-A production, and hypothesised that either there is a polar effect or MupM forms part of a multi-enzyme complex required for biosynthesis. Self-killing was not considered a possibility as the cultures were observed to be viable; they reached saturation in LB liquid media, and cell growth in SSM was not monitored. If the Δ*mupL* decrease in cell culturability after stationary phenotype is caused by a lack of MupM, then 10586 Δ*mupM* should also exhibit a growth or viability deficit.

To investigate this possibility, SSM 10586 Δ*mupM* cultures were grown in quadruplicate, with 10586 WT in duplicate as a control. Note that these cultures did not contain PA-B as in the previous CFU experiments. At 44 hours, the CFU count

was recorded and cultures examined by microscopy. Each of the Δ*mupM* cultures had CFUs >2 to 4 orders of magnitude lower than WT (Table 3.4). A significantly decreased number of cells was observed by microscopy, with visible clumping (Figure 3.39). It therefore seems likely that self-inhibition is the cause of the low levels of mupirocin production observed in 10586 Δ*mupM* cultures.

Table 3.5: 10586 Δ*mupM* demonstrates decreased CFUs in SSM compared with WT

	CFUs at 44 hours (cells ml^{-1})			
10586 WT	$> 5 \times 10^9$	$> 5 \times 10^9$	-	-
10586 Δ*mupM*	5×10^5	5×10^6	2×10^7	1×10^6

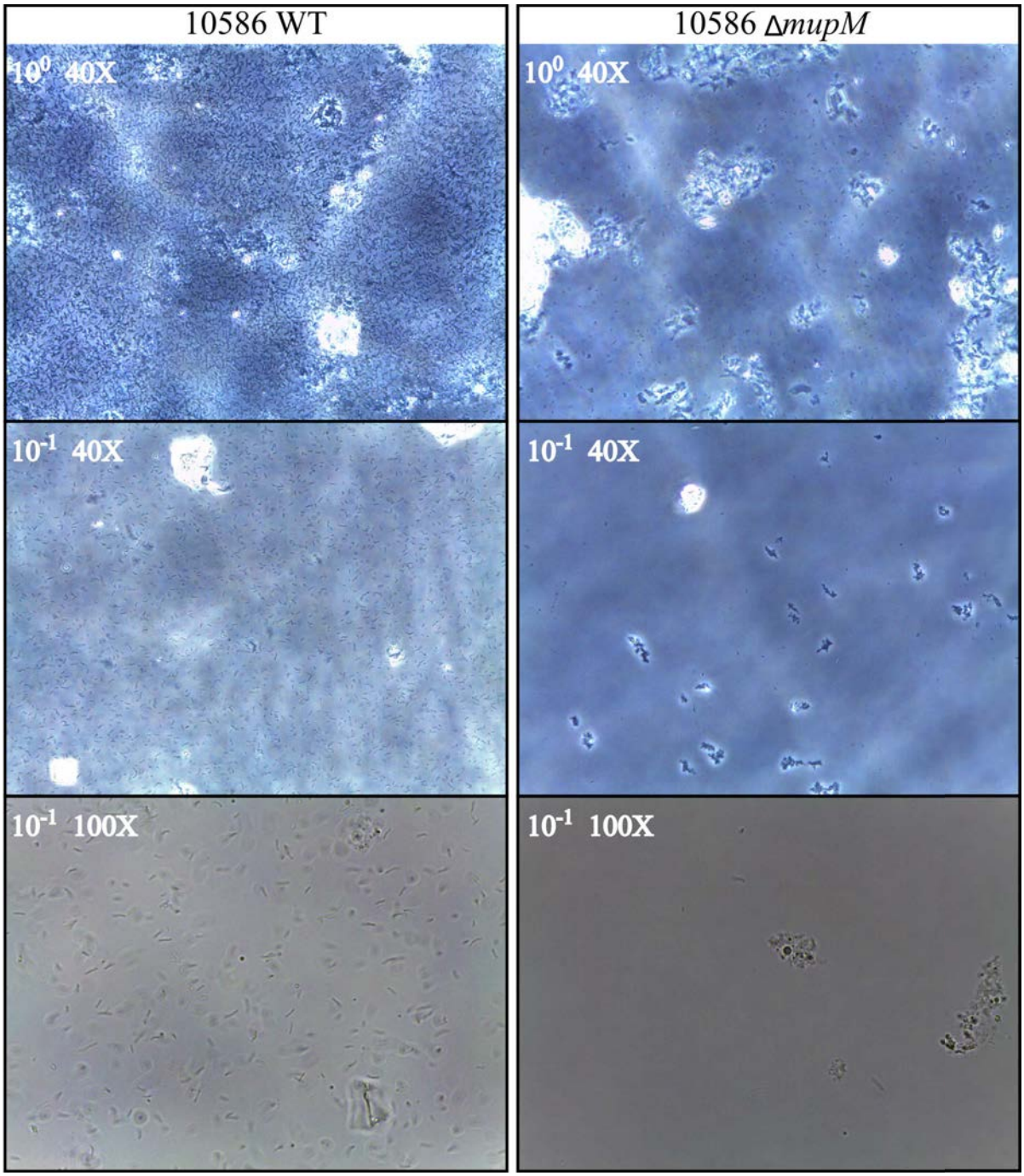


Figure 3.39: Microscopy of 10586 Δ*mupM* cultures.
The SSM cultures were imaged after 44 hours incubation. 10^0 and 10^{-1} represent undiluted and a ten-fold dilution of cultures respectively.

3.3 Discussion

In this chapter, nine mupirocin cluster enzymes were identified and demonstrated sufficient for conversion of PA-B to PA-A. This increased and re-inforced our understanding of the mupirocin tailoring pathway, and an updated putative model of conversion is presented. As the variable *mupL* mutant phenotypes were of particular interest, a model is presented to explain this observation, linking it to the role of MupM in biosynthesis.

3.3.1 Nine mupirocin cluster genes are required and sufficient for conversion of PA-B to PA-A in *P. fluorescens* 10586

The systematic screening of mutants for ability to process PA-B implicated ten genes as required. This informed the design of a series of expression plasmids, which ultimately revealed nine enzymes are required for conversion of PA-B to PA-A.

3.3.1.1 Systematically screening 10586 mutants for conversion of exogenous PA-B proved successful, yielding ten candidate genes

The systematic screening for ability of 10586 mutants to convert PA-B to PA-A, reported in Section 3.2.2.2, proved successful. Ten candidate mutants deficient in conversion were identified, *mupO, mupP, mupU, mupV, mupC, mupF, mupL, mupM, mupN* and *macpE*. Importantly, this ruled out the involvement of mupirocin genes of unknown function, such as *mupZ*, *mupA*, *mupB*, *mupD*, *macpA*, and *macpB*. This information guided the choice of thiomarinol homologs used in the experiments of Chapter 4. All of the genes we have reported previously as required for conversion were re-discovered (Cooper et al., 2005; Hothersall et al., 2007; Gao et al., 2017). Three further genes were implicated for the first time: *mupL* (enzyme not required), *mupM* and *mupN*. Without this knowledge, the plasmid based conversion experiments would not have succeeded.

The initial design for this experiment called for wide scale successive deletions of the mupirocin cluster. The generation of 10586 Δ*mup-cluster* demonstrated that large deletions were readily achieveable with our mutagenesis methodology. Starting from a deletion strain that is able to convert PA-B to PA-A, such as 10586 Δ*mmpA*, successive large deletions could be made to arrive at a minimal set of genes required for conversion of PA-B to PA-A. However, we decided instead to systematically screen existing mutants of each individual gene. This approach proved successful and very efficent, taking advantage of the considerable 10586 mutant library available.

A disadvantage of the systematic approach is that it would not detect redundancy of function, if two genes were able to substitute for each other then deletion of either would still give conversion. Given the dissimilarity of the predicted function of the various genes tested, with the exception of the five ACPs MacpA-E, functional redundancy seems unlikely. Redudancy between these ACPs is not likely, as each individual *macpA-E* knockout strain is defective in PA-A production, and complementation experiments have demonstrated that these do not substitute for each other (Shields et al., 2010).

This method could potentially be adapted to investigate other points of the pathway, although the majority of the earlier biosynthesis pathway is thought to occur on PKS enzymes, which would likely limit the point of pathway entry for exogenous substrates. No other pseudomonic acid intermediate is produced on the scale of PA-B, so any alternative substrates would likely need to be chemically synthesised.

3.3.1.2 Plasmid encoded conversion of PA-B to PA-A demonstrated that nine functional mupirocin enzymes are required and sufficient

The systematic screening experiments informed the design of a series of multiple mupirocin tailoring gene plasmids summarised in Figure 3.40. The ability of pJC133 to provide all necessary functions for conversion of exogenous PA-B to PA-A in 10586 Δ*mup-cluster* is the first demonstration that all necessary mupirocin cluster genes

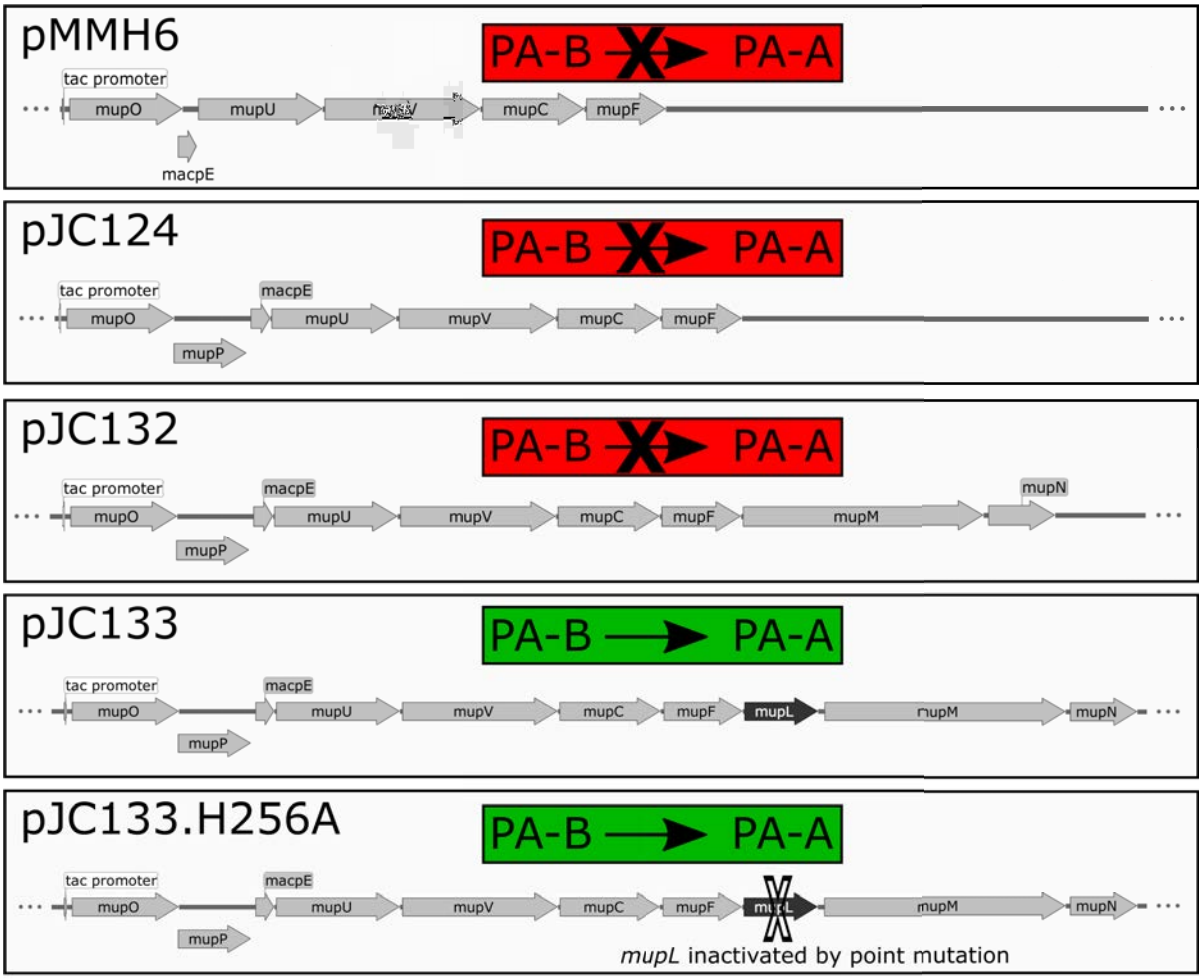


Figure 3.40: Summary of multiple tailoring gene plasmids generated.
Starting from pMMH6 (Macioszek, 2009), additional tailoring genes were inserted into the expression plasmids as guided by the screening experiments. This culminated with pJC133.H256A, which demonstrated that nine mupirocin cluster enzymes can provide all necessary functions for conversion: MupO, MupP, MacpE, MupU, MupV, MupC, MupF, MupM and MupN.

have been identified.

Together with the screening experiments, this represents the first evidence that conversion of MacpE to the holoenzyme obligately requires MupN (phosphopantetheinyl transferase), and therefore also that the native *P. fluorescens* PPTase cannot activate MacpE. MupN has been observed to convert MacpA, MacpC and MacpD to their holoenzyme *in vitro*, however MacpE was not tested due to solubility issues (Shields et al., 2010).

Plasmids pJC132 and pJC133 were observed to complement individual 10586

knockouts of *mupU*, *mupV*, *macpE*, *mupO*, *mupP*, *mupC* and *mupF* to between 65 and 90% of WT antibiotic activity (Figure 3.20). Earlier plasmids pMMH6 and pJC124 complemented to full 100% WT antibiotic activity (Figure 3.16) (Macioszek, 2009). As summarised in Figure 3.40, the difference between pJC133 and pJC124 is the addition of *mupL*, *mupM* and *mupN*. That pJC132 also confers the lower activity and lacks *mupL* logically suggests the addition of either *mupM* or *mupN* to the plasmids decreases mupirocin production. Which of these causes this effect could readily be tested by plasmid based expression of *mupM* and *mupN* individually in 10586 WT, and observing biological activity with and without IPTG induction.

Intriguingly, no conversion of exogenous PA-B to PA-A was observed in *P. fluorescens* SBW25 carrying pJC133. One possible explanation is that the biosynthesis genes carried on pJC133 are not expressed in SBW25, as it lacks the mupirocin regulatory genes. However, this seems unlikely as the cloned genes are under the control of the *tac* promoter, which is unrelated to mupirocin biosynthesis. The putative promoter within *mupL* from the 10586 chromosome does not appear to be under MupR regulation, as pJC133 provides conversion functions in 10586 Δ*mupR*. Indeed, the observed increase in mupirocin MIC from 1500 μg ml^{-1} in SBW25 [pJH10] to >3000 in SBW25 [pJC133] is evidence that *mupM* is expressed from pJC133 in SBW25, and that the *mupL* promoter is still functional.

An alternative explanation is that there are further genes required outside of the mupirocin biosynthetic cluster in the 10586 genome, that are missing in SBW25. Elsewhere in the mupirocin biosynthesis pathway, functional analysis of MupW (Rieske dioxygenase) and MupT (Ferredoxin) revealed they belong to a Rieske oxygenase system (Connolly, 2014 - unpublished), which characteristically require a third reductase component (Ferraro et al., 2005). The putative reductase of this electron transport system has not been identified, and is proposed to be present elsewhere on the genome. As pJC133 was observed to provide the conversion functions in 10586 Δ*mupR* (Figure 3.23), we can predict that if any genes are required elsewhere on the

10586 chromosome, they will not be under MupR regulation.

Future work: bioinformatic screening for any genes required outside of the mupirocin cluster

Bioinformatic analyses could be performed to identify any putative further genes required for mupirocin biosynthesis outside of the identified gene cluster. Five genomes could be compared: the first identified mupirocin producer *P. fluorescens* 10586 (Haines, unpublished), *P. fluorescens* BRG100, *P. synxantha* A342, *P. pyschrotolerans* NS383, and *P. fluorescens* SBW25. BRG100 and A342 contain >99% (A.A.) similar mupirocin clusters to 10586, and it is therefore possible the genomes would be too similar to yield any useful comparisons. NS383 is a sequenced *P. fluorescens* group strain carrying a mupirocin cluster with approximately 65-70% similarity to 10586. It was sequenced as part of a meta-genomics study largely at the species level, which did not focus or identify individual clusters, and as such it has not been confirmed to produce mupirocin (Midha et al., 2016).

Any genes required for mupirocin biosynthesis that are located outside of the cluster that are responsible for the lack of PA-B conversion in SBW25 [pJC133], should logically be present in the chromosomes of 10586, BRG100, A342 and NS383, but not in SBW25. The first step would be developing a software search for such genes, which could then be appraised for likelihood of involvement in biosynthesis, and reverse genetic experiments later on. It is conceivable that this would generate too large a list to be useful for identifying potential missing conversion genes. However, this could still be useful in identifying components where the encoded enzyme type could be predicted, such as the predicted third component of the MupW, MupT electron transport chain.

3.3.1.3 The relative level of *mupV* expression may be essential to complementation *in trans*

As discussed in Section 3.1.3, two attempts to complement 10586 Δ*mupV* with independently constructed pJH10/*mupV* expression plasmids failed (Di Martino, M., Stephens, E., *unpublished*). Given this negative result, an alternative chromosomal point mutatagenesis strategy was pursued in this study (Section 3.2.1). However, the failure of these *mupV* expression plasmids to complement is intriguing, as this complementation has been observed in alternative scenarios. Complementation of 10586 Δ*mupV* was observed by multi-gene expression plasmids pJC124, pJC132 and pJC133 in this work (Section 3.2.3.3), as well as by their parent, pMMH6 (Macioszek, 2009). Complementation was also observed by an alternative design of pJH10/*mupV* constructed during the thesis of Cooper (2003).

In these different designs, the 5′ side of *mupV* has been cloned differently, summarised in Table 3.6. Multiple gene expression plasmids pJC124, pJC133 and pMMH6 all use the 10586 chromosomal DNA sequence between the end of the preceding *mupU* and *mupV*. However, the different pJH10/*mupV* versions were constructed to make use of a RBS already present in expression vector pJH10, upstream of the multiple cloning site. The pJH10/*mupV* designed by Maura Di Martino was cloned using *Mfe*I into the *Eco*RI, with 9 bp between translation start and the RBS, compared with 7 bp in the consensus. This plasmid was observed inhibitory to mupirocin production in 10586 WT. Expression of *mmpB* was achieved using this exact same 5′ setup, reported in Section 4.2.1 of this work. The pJH10/*mupV* designed by Sian Cooper was cloned using *Kpn*I, which follows *Eco*RI in the multiple cloning site, which resulted in 15 bp between translation start and the RBS.

It is possible that the Di Martino plasmid is inhibitory to mupirocin production because the relative expression of *mupV* is too high. The 10586 chromosomal *mupV* is single copy and under MupR regulation, whereas the plasmid is 12 copy (Frey et al., 1992) and the cloned *mupV* expressed by the IPTG regulated, strong *taq* promoter.

Table 3.6: Ribosome binding site of *mupV* in different experiments

Scenario	Sequence (5′ to 3′)	Complementation of 10586 Δ*mupV*
	RBS +1	
Consensus	AGGAGGNNNNNNATG	N/A
10586 chromosome	AGGAAGGCGGACATG	N/A
pJC124, pJC132, pJC133, pMMH6	AGGAAGGCGGACATG	+
pJH10/*mupV* Di Martino/Stephens	AGGAAACAGAATTGATG	-
pJH10/*mupV* Cooper	AGGAAACAGAATTCGGTACCATG	+

In this explanation, the Cooper plasmid is able to complement 10586 Δ*mupV* as the expression is decreased by the increased spacing between the RBS and start of 15 bp. An equivalent increase in RBS spacing in *E. coli* was observed to reduce reporter gene expression more than 6-fold (Chen et al., 1994). In the pMMH6 derived plasmids, a near consensus optimally spaced RBS is present, but the plasmid also includes the genes encoding enzymes that could logically interact with MupV: MupU, MupO and MacpE. The expression of *mupV* may need to at a similar level to one or more of these related genes, and that the Di Martino pJH10/*mupV* is inhibitory because the relative expression of *mupV* is too high. This hypothesis could be tested by replacement of *mupV* with reporter gene *lacZ* in each scenario to assay expression, and by addition of *mupU*, *mupO* or *macpE* to the Di Martino pJH10/*mupV* to check for relief of inhibition.

3.3.2 Models of PA-B to PA-A conversion

3.3.2.1 Both domains of MupV are required for conversion

Re-analysis of *mupV* revealed two predicted domains: short chain reductase and hydrolase. Chromosomal point mutation of either of these domains in 10586 fully switched production to PA-B (Section 3.2.1). This suggested that both domains are

required for conversion of PA-B to PA-A, Aim 1 of this chapter. When these mutant strains were fed with exogenous PA-B no conversion to PA-A was observed (Section 3.2.2.2), supporting this conclusion. It is possible MupV has a dual function, or that both domains are required for its proposed function of polyketide offloading from MacpE.

3.3.2.2 Putative schematic pathway of conversion

The results discussed in this chapter support the current model of conversion of PA-B to PA-A, most recently proposed in Gao et al. (2017). This is presented in Figure 3.41, with slight modification to the timing of release from MacpE, based on the hypothesis described in the next Section 5.2.1.1. This conversion schematic is logically based on historical reverse genetics experiments.

The functions of MupP, MupC and MupF have been characterised by accumulation of late stage intermediates mupirocin P, C1, and F in their respective chromsomal knockout mutants (Hothersall et al., 2007; Gao et al., 2017). Deletion of either *macpE* (ACP), *mupO* (cytochrome P450), *mupU* (acyl CoA synthase) or *mupV* (oxidoreductase and TE) causes a near complete switch to PA-B production (Cooper et al., 2005). It is logical that MacpE (acyl carrier protein) is the site of at least some of the conversion pathway, and the evidence in Sections 3.2.2 and 3.2.3 suggests that MupN (phosphopantetheinyl transferase) catalyses formation of the MacpE holoenzyme. MupU (acyl CoA synthase) is the most likely candidate for loading PA-B onto MacpE in an ATP dependent reaction. The second domain of MupV (thioesterase) could catalyse release of pseudomonic acids from MacpE. This leaves the first domain of MupV (oxidoreductase) and MupO (cytochrome P450). These could both function together to generate mupirocin P, or MupV may require action of both domains for release from MacpE.

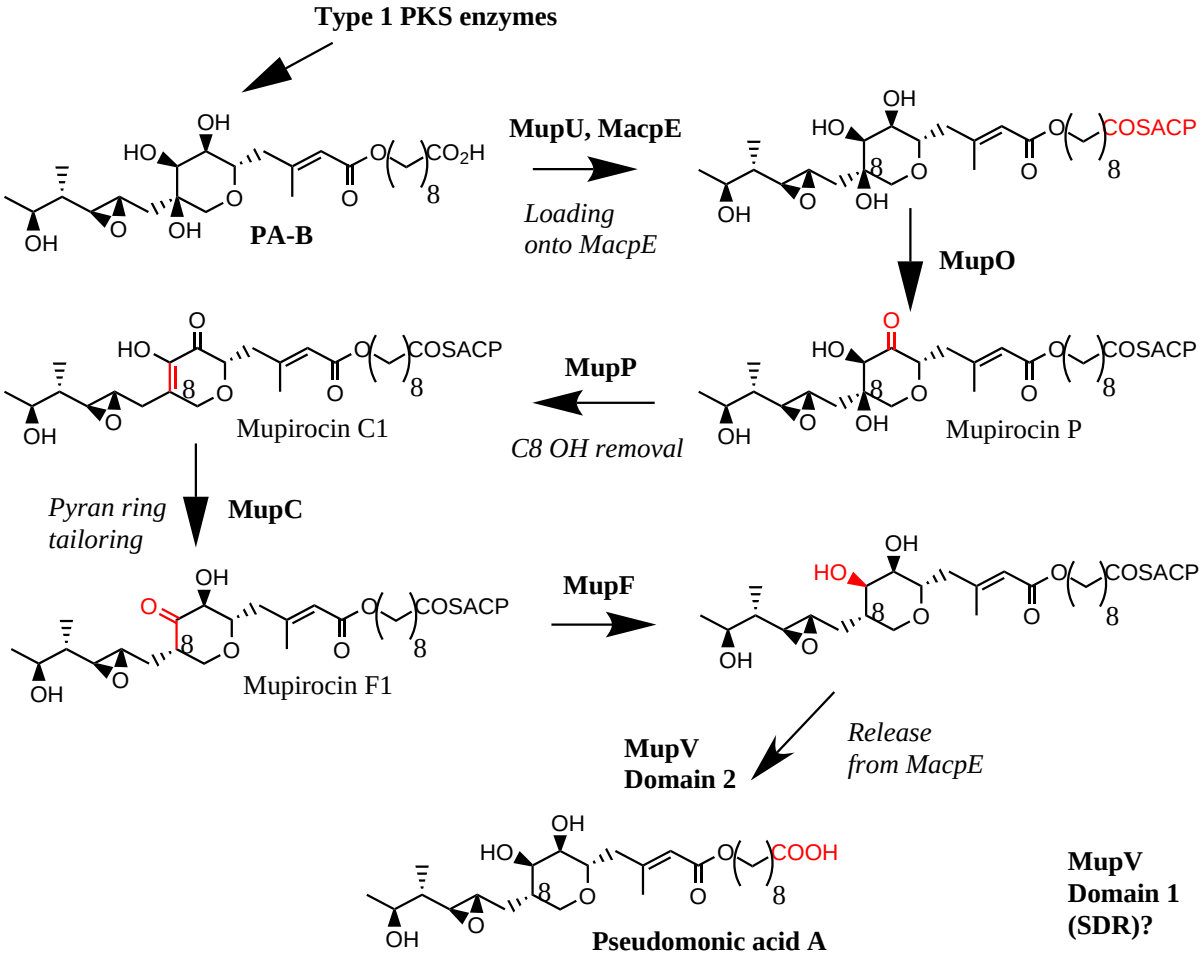


Figure 3.41: Putative PA-B to PA-A conversion schematic.
Enzymatic order and steps proposed by Gao et al. (2017), MacpE release timing modified as discussed in Section 5.2.1.1.

3.3.3 Gene *mupL* contains a putative promoter that could be key for cell viability

In Section 3.2.4, various evidence was presented that supports presence of a putative promoter within *mupL*, which later led to intriguing discoveries regarding interactions with downstream *mupM*.

3.3.3.1 Evidence for a promoter within *mupL*

As summarised in Figure 3.40, multi-gene expression plasmid pJC132 without *mupL* failed to provide PA-B to PA-A conversion functions, whereas pJC133 with *mupL* does. Inactivation of the MupL enzyme by point mutation demonstrated that

its activity is not required for conversion. Analysis of the complementation of both the mupirocin production and decreased resistance phenotypes of 10586 Δ*mupM* by pJC132 and pJC133 suggested improved *mupM* expression in pJC133. This supports the hypothesis that *mupL* contains a promoter that drives transcription of downstream genes *mupM* and *mupN*.

As the putative *mupL* promoter and *mupM* are required for plasmid based conversion, and conversion can occur in 10586 Δ*mupR* [pJC133], it is likely the putative *mupL* promoter is not MupR regulated.

Future work

The experimental evidence presented in this work all supports presence of a promoter within *mupL*, however it would be sensible to demonstrate this using transcriptional assays. There are various options to accomplish this: quantitative reverse-transcriptase PCR (RT-PCR) of a product in *mupM* with and without *mupL* presence, β-galactosidase assays of the putative promoter fused to *lacZ*, or RNA sequencing of the 10586 transcriptome with and without *mupL*.

3.3.3.2 The promoter within *mupL* may have evolved to increase *mupM* expression before or after mupirocin biosynthesis

Upon discovery of the putative promoter, we hypothesised that its role was to express *mupM* prior to the pathway being switched on, such that sufficient MupM is already present in the cytoplasm before mupirocin production begins, preventing self-action. However, there is an alternative hypothesis; the promoter may have evolved to prevent self action after mupirocin production finishes. There are two isoleucyl tRNA synthetases (IleRS) in the 10586 chromosome, the eukaryotic like IleRS-R2 (MupM) that confers high mupirocin resistance, and the more susceptible IleRS-R1 (Yanagisawa and Kawakami, 2003), which explains why 10586 Δ*mupM* is viable.

A prediction for how *mupM* expression levels could vary between in 10586 WT

compared with 10586 Δ*mupL* is presented in Figure 3.42. It is conceivable that the constitutive promoter within *mupL* is required to provide sufficient MupM after cluster switch off, if PA-A levels in the cell were to remain high. This possibility is supported by the observation that mupirocin concentrates in sensitive *B. subtilis* at nine times the extracellular concentration (Capobianco et al., 1989). Therefore it is likely that 10586 intra-cellular PA-A concentration remains higher than extracellular, due to binding to the more susceptible IleRS-R1.

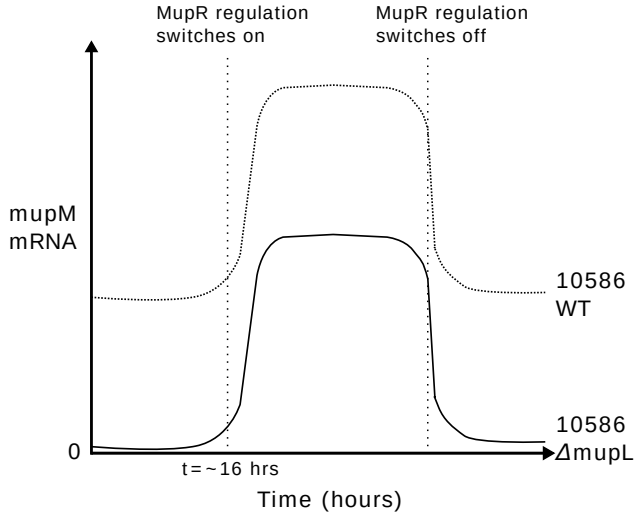


Figure 3.42: Prediction of how deletion of the promoter within *mupL* could change *mupM* expression.

Alternatively, the promoter may have evolved to provide cells where the cluster is switched off with constitutive mupirocin resistance, to guard against mupirocin produced by neighbouring *P. fluorescens* cells in the native soil rhizosphere. However, this is contradicted by the quorum sensing regulation of mupirocin production (Hothersall et al., 2011). We might expect *P. fluorescens* cells close in space to either both be on or off together with respect to mupirocin biosynthesis regulation.

3.3.3.3 10586 Δ*mupL* is variably unable to convert PA-B to PA-A, correlating with a decrease in cell culturability

Providing an explanation of the reproducibly variable conversion of exogenous PA-B to PA-A in 10586 Δ*mupV* and corresponding decrease in cell culturability is

difficult, as an explanation must account for the culture wide effect observed. Several conclusions, and clues as to what could be responsible can be deduced from the experimental evidence.

The observation that 10586 Δ*mupM* grows poorly in SSM, with low CFUs and clumps visible by light microscopy (Section 3.2.4.4), logically suggests that produced PA-A is self-inhibitory in this strain. Furthermore, the decreased mupirocin production observed in 10586 Δ*mupM* could then be explained by a lack of cell density.

The key observation is that 10586 Δ*mupL* SSM cultures fed with PA-B will grow to saturation, before a proportion of the culture's CFUs collapse, and correspondingly in these cultures, no PA-A conversion is observed. These different observed phenotypes are summarised in Figure 3.43.

In the previous Section 3.3.3.1, evidence was discussed for the presence of a promoter within *mupL* that drives *mupM* expression. Given that PA-A production seems inhibitory in a *mupM* knockout, and the Δ*mupL* mutant may lack a promoter that drives *mupM* expression (Section 3.3.3.1), the observed collapse in CFUs in PA-B SSM 10586 Δ*mupL* cultures could be caused by self inhibition due to lack of MupM. Whilst this provides a potential mechanism for CFU collapse, it is more difficult to explain how this could occur variably, on a culture-wide basis, after the cultures reach stationary phase.

Mupirocin is regulated by quorum sensing, with a signalling homoserine lactone produced by MupI (Hothersall et al., 2011). This provides a possible mechanism for a culture wide phenomenom to occur, further complicated by the presence of the lactone within the fed PA-B extract (Section 3.2.2.3).

If MupR regulation were to switch off through MupI signalling in some saturated 10586 Δ*mupL* SSM PA-B cultures, then any PA-A already produced could prove self inhibitory as *mupM* expression would switch off. Although it is unclear how or why this would occur.

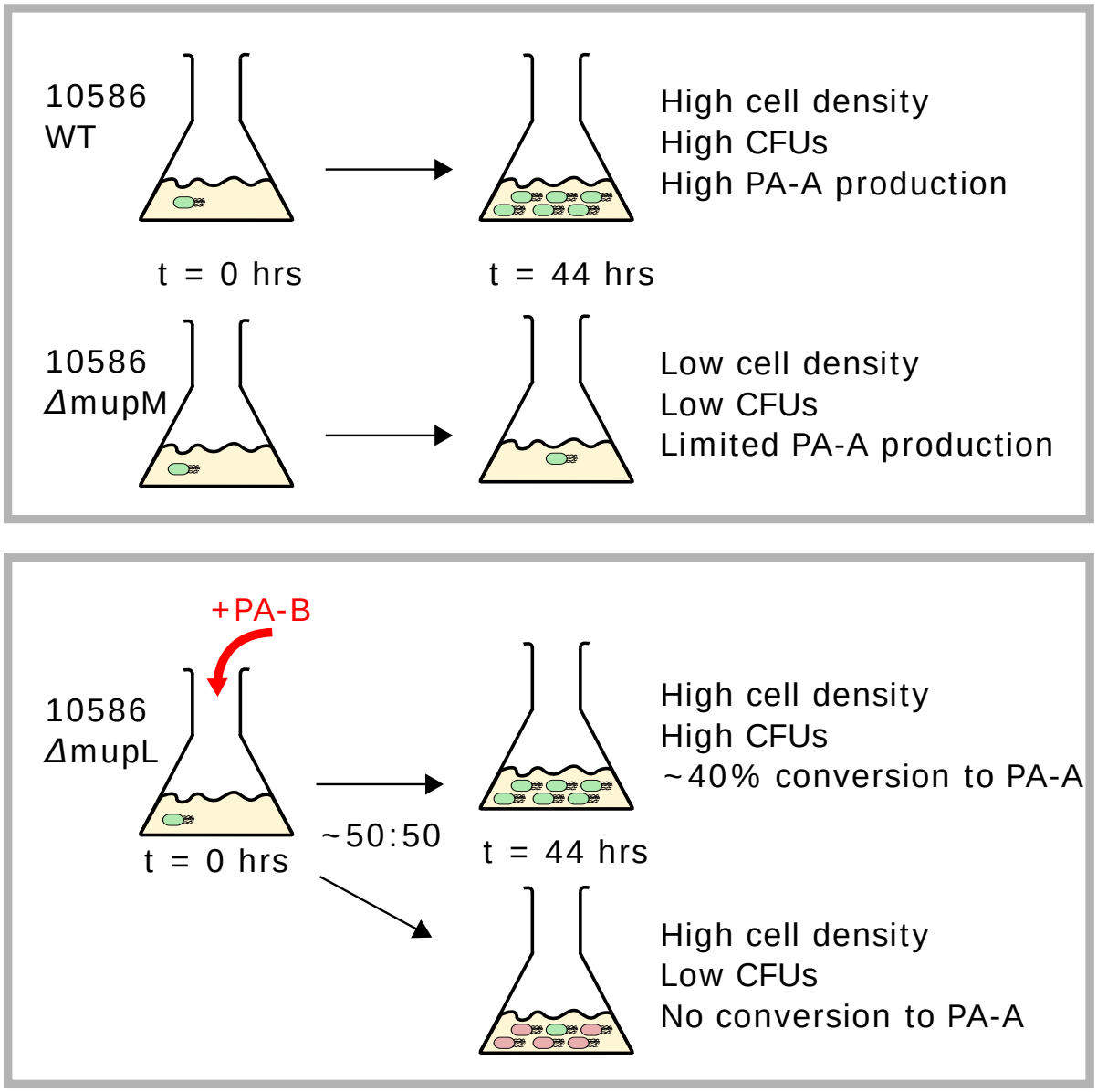


Figure 3.43: Summary of growth and PA-A conversion or production phenotypes of *mupL* and *mupM* mutant SSM cultures.

Future work

The culture wide collapse in CFUs in some 10586 Δ*mupL* SSM cultures was an unexpected result, which due to time contraints has not yet been fully explored. The immediate experiments should aim to reproduce this effect. From there, various exploratory approaches could be undertaken, such as recording the culture CFUs at more frequent time points over a wider range. It would be very interesting to test where the decision is made if a culture is split at a given timepoint, will all the aliquots present with the same phenotype?

It would be worthwhile assaying the level of lactone in low and high CFU Δ*mupL* SSM cultures; this could be performed by *xylE* assay as in Section 2.3.3 (El-Sayed et al., 2001). Whilst it is not clear why the lactone level would fall in some cultures, it is a culture-wide signal that could potentially drive such a culture-wide phenotype. More advanced microscopy could be performed; of particular interest would be live-dead staining (such as propidium iodide) of the high cell density, low CFU cultures.

Further reverse genetic investigation is already underway, with the design and generation of a 10586 *mupL* nonsense mutation, with MupL curtailed by an early stop codon, leaving the putative promoter intact. An in-frame deletion of the majority of *mupL* that leaves the putative promoter intact has also been designed. We would expect these new *mupL* mutants to be deficient in mupirocin production, because of the lack of MupL, and its unknown requirement in the mupirocin pathway. However, they should be able to reliably convert fed PA-B to PA-A, if the hypothesis is correct that the failure to convert and low CFU effect is due to MupM levels.

Chapter 4

RE-ENGINEERING OF PA-B TO PA-A CONVERSION *IN VIVO*

4.1 Introduction

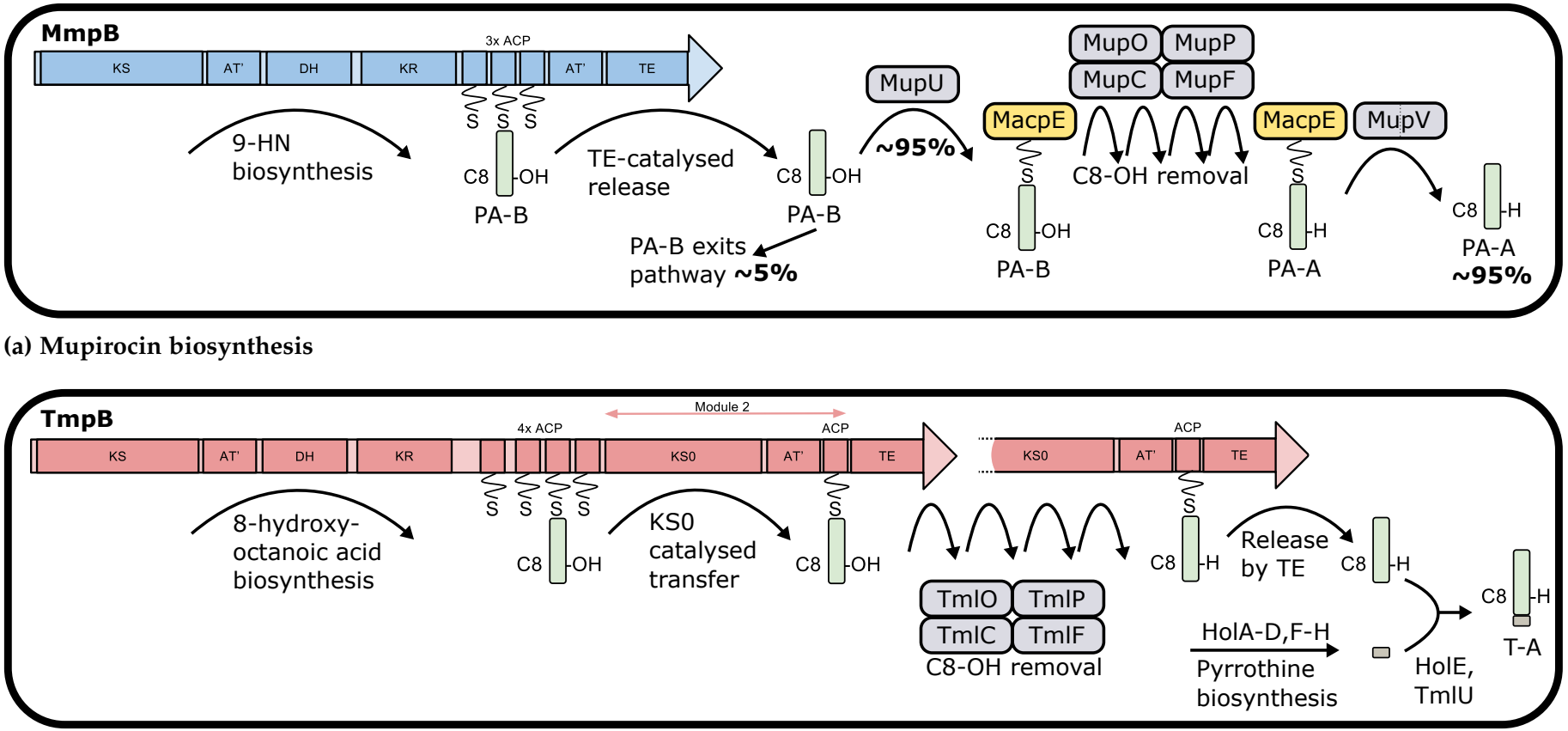
As outlined in Section 1.5.1, a long term goal of the Thomas research group is to generate *in vivo* mupirocin-NRP derivatives through fusion of additional modules to MmpB. However, the inactive PA-B is proposed to be released from MmpB, and converted to active PA-A on the discrete enzyme MacpE (Section 3.1). This established the key goal pursued in this work: the re-engineering of mupirocin biosynthesis to generate PA-A "*in cis*" on MmpB.

A previous attempt to achieve this is described in the PhD thesis of Omer-Bali (2013). Omer-Bali reported that an *mupU-macpE* fusion on an expression plasmid could complement 10586 Δ*macpE* Δ*mupU*. This fusion was then inserted into the chromosomal *mmpB* gene of 10586 Δ*macpE* Δ*mupU*, however this proved completely disruptive to pseudomonic acid production. The *mupU-macpE* fusion was also inserted into *mmpA* (PKS) of 10586 Δ*macpE* Δ*mupU*, and the resulting strain was observed to produce PA-B. This suggests that this insertion did not disrupt MmpA function, but that MupU-MacpE do not function in this position as PA-A was not produced.

In the work presented in this Chapter, an alternative strategy is pursued using enzymatic domains from the biosynthesis of the similar antibiotic, thiomarinol.

4.1.1 Two strategies for C8-OH removal found in nature

In mupirocin biosynthesis PA-B (C8-OH) is proposed to undergo TE-catalysed release from MmpB (Section 3.1). MupU (acyl CoA synthase) is proposed to then load PA-B onto MacpE (ACP), before successive action by MupO (cytochrome P450) and MupP (dehydratase) to remove the C8-OH. Tailoring of the pyran ring is then completed by MupC (NADH oxidase) and MupF (ketoreductase). Release from MacpE is proposed to be catalysed by the second domain of MupV (TE), and there is a further unknown role for the first domain of MupV (SDR) (Section 3.2.1).



(a) Mupirocin biosynthesis

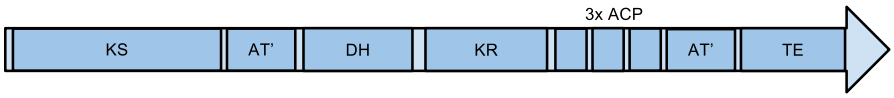
(b) Thiomarinol biosynthesis

Figure 4.1: Overview of the mupirocin and thiomarinol late-stage biosynthesis pathways.
(a) In mupirocin biosynthesis, PKS MmpB is proposed to synthesise 9-hydroxynonanoic acid (9-HN) to generate PA-B with C8-hydroxy. The inactive PA-B is released from the PKS and then undergoes a multi-step pathway to remove the C8-OH and form active PA-A, proposed to occur on MacpE.
(b) In thiomarinol biosynthesis, TmpB is proposed to synthesise 8-hydroxyoctanoic acid, generating marinolic acid B (C8-OH). This intermediate may then undergo KS^0-mediated transfer to the module 2 ACP (the fifth ACP of TmpB). The equivalent steps that may occur on MacpE are proposed to occur on this module 2 ACP, removing the C8-OH. The C8-H marinolic acid A could be released from TmpB, where it then undergoes amidation with pyrrothine to generate thiomarinol A.

Intriguingly, the equivalent step in thiomarinol biosynthesis may be accomplished in an alternative fashion. A putative pathway is proposed in Figure 4.1b, based on bioinformatic analysis and logical comparisons with mupirocin (Fukuda et al., 2011).

Homologs to *mupO*, *mupP*, *mupC* and *mupF* are present in the thiomarinol gene cluster, with 43%, 30%, 55% and 33% amino acid similarity respectively, named *tmlO* and so forth. However, there are no thiomarinol homologs to *mupV* or *macpE*. The *mupU* homolog *tmlU* is one of the few genes that have been studied by knockout in thiomarinol biosynthesis, and it has been demonstrated to have an alternative function to MupU: the TmlU and HolE-catalysed amidation of pyrrothine to marinolic acid (Section 1.5.1.1) (Dunn et al., 2015). Therefore thiomarinol biosynthesis is missing the mupirocin genes that encode MacpE, and the enzymes proposed to load and offload it.

(a) MmpB (2077 amino acids)



(b) TmpB (2751 amino acids)

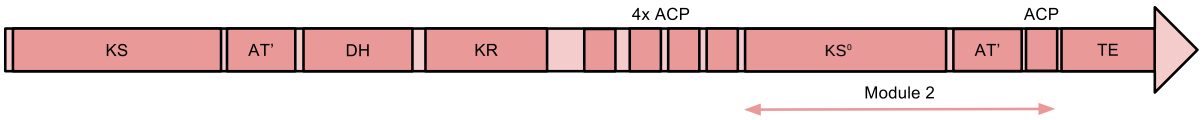


Figure 4.2: Comparison of the domains of MmpB and TmpB
(a) MmpB is the type I iterative PKS responsible for 9-hydroxynonanoic acid biosynthesis in the mupirocin pathway. The KS, DH, KR, 3 ACPs work together with the *trans*-AT on MmpC to perform three successive Claisen condensations, then the C8-OH intermediate is released, catalysed by the TE.
(b) The thiomarinol homolog, TmpB, has an additional module, consisting of a non-elongating KS^0, KS-AT linker (AT'), and ACP. Removal of the C8-OH is hypothesised to occur on this extra module. The reason for the additional ACP in module one is not known.

TmpB has an extra module not present in MmpB, consisting of a non-elongating ketosynthase (KS^0), a KS-AT linker domain (AT') and acyl carrier protein (ACP) (Figure 4.2). The KS^0 domain lacks the conserved histidine of the HGTGT motif that is required for elongation, and it instead has QGTGT (Robbins et al., 2016). However,

this domain still has the conserved cysteine (Cys1824) required for chain transfer, and therefore may be able to transfer intermediates, but not catalyse condensations. This extra module could be the site of C8-OH removal in thiomarinol biosynthesis, performing an equivalent role to MacpE in mupirocin. In the proposed hypothesis, the first KS and associated domains catalyse two iterative Claisen condensations to generate the C8-OH intermediate. The KS^0 then acts equivalently to MupU, loading the C8-OH intermediate onto the following ACP on TmpB. Whilst on this ACP, TmlO, TmlP and possibly TmlC and TmlF could catalyse the equivalent tailoring steps to their *mup* counterparts. Finally the thioesterase would release the C8-H polyketide, equivalent to MupV action.

This model of C8-OH removal in thiomarinol is supported by the observation that no C8-OH thiomarinol is detected by LC-MS analysis of cultures of its producer SANK73390 (Fukuda et al., 2011; Murphy et al., 2014). In mupirocin biosynthesis, approximately 5% of the pseudomonic acids produced are PA-B with C8-OH. It is reasonable that there would be some leak of PA-B in the transfer from MmpB to MacpE in mupirocin biosynthesis, but none of the equivalent thiomarinol intermediate, as it does not leave the PKS.

Recent bioinformatic searches of MmpB have revealed what appears to be a remnant KS-AT linker located between ACP7 and the TE domain, which has been named AT' in the diagrams in this work. This domain shows around 30% amino acid similarity to a variety of putative Bacillus subtilis PKS and NRPS enzymes, and 14% similarity to AT domains with solved X-ray structures, with HHPRED identifying the *M. tuberculosis* Pks13 AT domain (PDB: 3TZY) as the closest hit.

This putative KS-AT linker in MmpB is not positioned in the canonical *cis*-AT domain position, where it would follow the KS domain (Staunton and Weissman, 2001). It is not known if this domain has a function within MmpB, but with a length of 150 amino acids it is likely to have an effect on the 3D positioning of the following TE domain. The existence of this putative domain was not known when the work

presented in this chapter was performed.

Similar KS-AT linkers are present following the KS domains on MmpA, MmpB and MmpD. These are positioned in the correct in *cis* position and have been hypothesised to be a potential position for the docking of the *trans*-acting AT domains of MmpC. Although the position of *trans*-acting AT domain interactions with the KS is not known (Gay et al., 2014), and could not be the same for all systems. In some previous work on the mupirocin biosynthesis, these KS-AT linkers have been referred to as docking domains (Gurney and Thomas, 2011; Gurney, 2012). However, in the wider field of polyketide research, in particular of the model erythromycin, the term docking domains has been used to describe docking between PKS polypeptides (Broadhurst et al., 2003; Weissman, 2006; Zeng et al., 2016). Given the lack of direct evidence for the hypothesised *trans*-AT docking role, and the potential for confusion with the inter-peptide function, in this study these domains have been termed AT'.

4.1.2 Objectives

The key objective of the work described in this chapter was to re-engineer C8-OH removal in mupirocin biosynthesis so that it occurs on MmpB, in a similar manner to that proposed in thiomarinol biosynthesis.

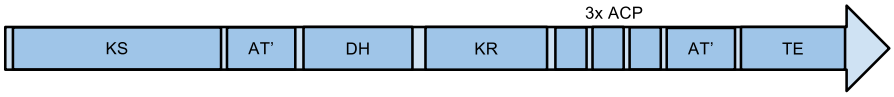
Aim 1: Complement 10586 Δ*mmpB* with *mmpB* provided *in trans*

The first aim was to generate an *mmpB* expression plasmid, then mobilise to 10586 Δ*mmpB* and test for complementation. This would establish the basis for *in trans* expression of *mmpB* so that later work on *mmpB-tmpB* derivatives could also be performed *in trans*.

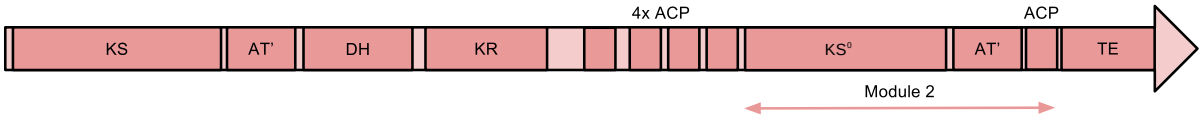
Aim 2: Insert the second module of TmpB into MmpB and test for function

The next aim was to generate a functional MmpB/TmpB hybrid through insertion of the second module of TmpB into MmpB (Figure 4.3). In reality, this took some

(a) MmpB



(b) TmpB



(c) MmpB/TmpB hybrid

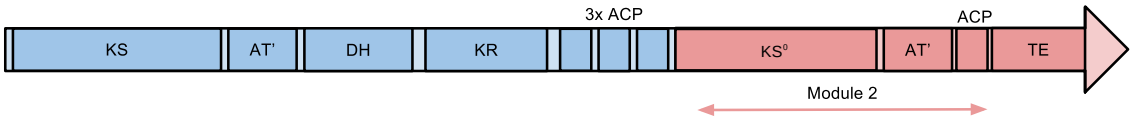


Figure 4.3: Aim 2: Construction of an MmpB/TmpB hybrid protein
Blue represents MmpB (mupirocin) domains and red TmpB (thiomarinol) domains. A hybrid of MmpB and TmpB was designed, with the front end of MmpB, up to and including the third of its ACP triplet, ACP7. After this, the second module of TmpB would be fused on, ending with the TmpB TE domain.

design iterations for a functional hybrid to be achieved. Furthermore, an aim was to inactive the transferase activity of the KS^0 domain of the hybrid by point mutation and test function. Inactivation of the KS^0 domain should decrease throughput if it is actively processing intermediates.

Aim 3: Provide the candidate thiomarinol tailoring enzymes and attempt to convert PA-B to PA-A on the MmpB/TmpB hybrids

With a functional MmpB/TmpB hybrid generated, the next aim was to provide the thiomarinol tailoring enzymes putatively involved in C8-OH removal, and test for C8-OH removal on the hybrid (Figure 4.4). Logistically, this required insertion of the various hybrid iterations into the chromosome of 10586 strains deficient in the mupirocin C8-OH removal pathway. Then, informed by the work described in Chapter 3, a selection of thiomarinol tailoring enzymes could be provided on a plasmid with the aim of restoring PA-A production.

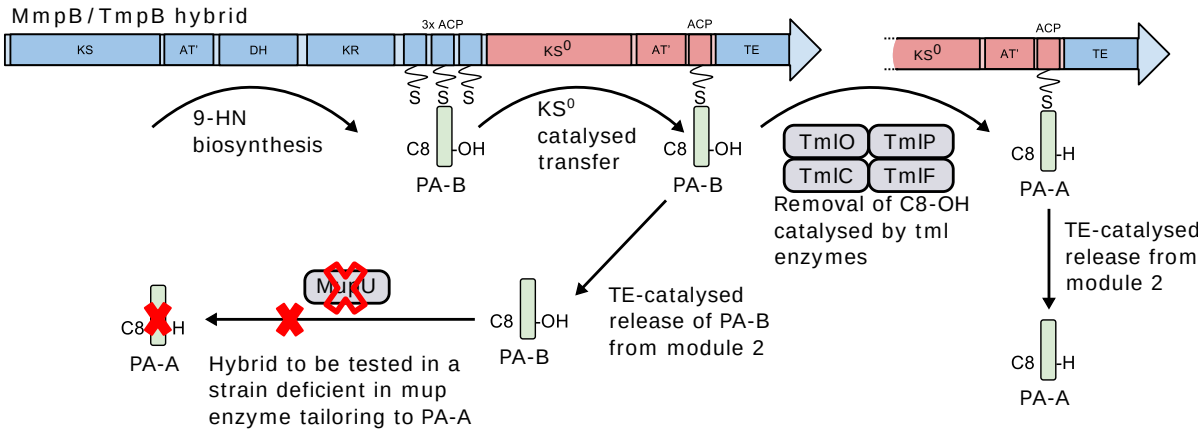


Figure 4.4: Aim 3: Conversion of PA-B to PA-A on an MmpB/TmpB hybrid catalysed by *tml* enzymes.
A later iteration of the MmpB/TmpB hybrid enzymes is shown, which has the MmpB TE domain (in blue).

Aim 4: Use the TmpB KS^0 to load MacpE fused to MmpB and convert PA-B to PA-A on the MmpB/TmpB hybrids

At the outset of this work, Aim 3 was the envisaged path to "*in cis*" PA-B to PA-A conversion on MmpB/TmpB hybrid enzymes. However, with the successful splicing of the KS^0 domain in MmpB, an alternative possibility was considered. The KS^0 domain could be used to re-engineer the native mupirocin tailoring pathway to occur on MmpB. This would require splicing of the normally free-standing, MacpE into the MmpB/TmpB hybrids, and testing for conversion to PA-A on the hybrid (Figure 4.5).

Together success in Aims 3 and 4 could unlock the potential for coupling of NRPS modules on MmpB/TmpB, to generate PA-A-NRP *in vivo* derivatives in the future.

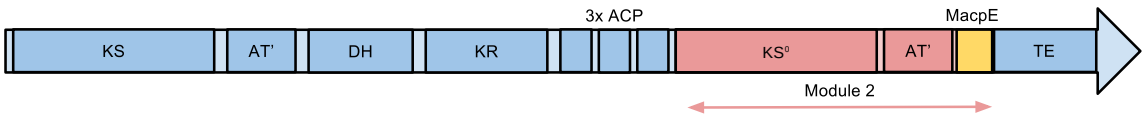


Figure 4.5: Aim 4: Construction of a hybrid protein where the TmpB KS^0 domain is positioned to load MacpE.
The TmpB module 2 ACP would be replaced by the normally free-standing MacpE, which could then be loaded by the transferase activity of the KS^0 domain. Conversion to PA-A could then occur on the hybrid through the native *mup* tailoring pathway on this fused MacpE domain.

4.2 Results

4.2.1 Complementation of *P. fluorescens* Δ*mmpB*

Two approaches were considered for modifying *mmpB* to introduce new modules into MmpB: replacement of the 10586 chromosomal copy with modified versions, or provision of modified *mmpB* on an expression plasmid to *P. fluorescens* NCIMB10586 with *mmpB* deleted. Although a mixture of both approaches were eventually pursued, initially the latter *in trans* approach was used. To establish the feasibility of this, an important preliminary aim was to test if 10586 Δ*mmpB* could be complemented by WT *mmpB* provided on an expression vector.

4.2.1.1 Construction of pJH10/*mmpB*

Amplification of *mmpB* was achieved in three parts, both to alleviate difficulties amplifying the full length gene and to remove several internal *Eco*RI sites, to facilitate its use for cloning. The first 270 bp of the gene were amplified using primers MmpBF1 and MmpBF2, followed by a 4 kb fragment amplified with MmpBF2 and MmpBR3. All of the primers used in this chapter are listed in Table 2.8. The PCR strategy and generation of intermediate construction plasmids is summarised in Figure 4.6.

These two fragments were designed to overlap, with a silent mutation (Glu92 GAA to GAG) to remove an internal *Eco*RI site. Overlap extension PCR was performed (Section 2.2.11.1) with primers MmpBF1 and MmpBR3, generating a 4.3 kb fragment. This fragment reached from the gene start to its native unique *Kpn*I site.

The final 2 kb segment of *mmpB*, from its native *Kpn*I site to the stop codon, was amplified from the 10586 chromosome using primers MmpBF2 and MmpBR3. This end segment was cloned in high copy vector pUC18 *Kpn*I/*Xba*I, yielding pUC/-*mmpB*CT2kb.

The 4.3 kb fragment containing the front of the gene was cloned in pUC/*mmpB*-CT2kb *Eco*RI/*Kpn*I, yielding the full length *mmpB* in pUC18. This was named

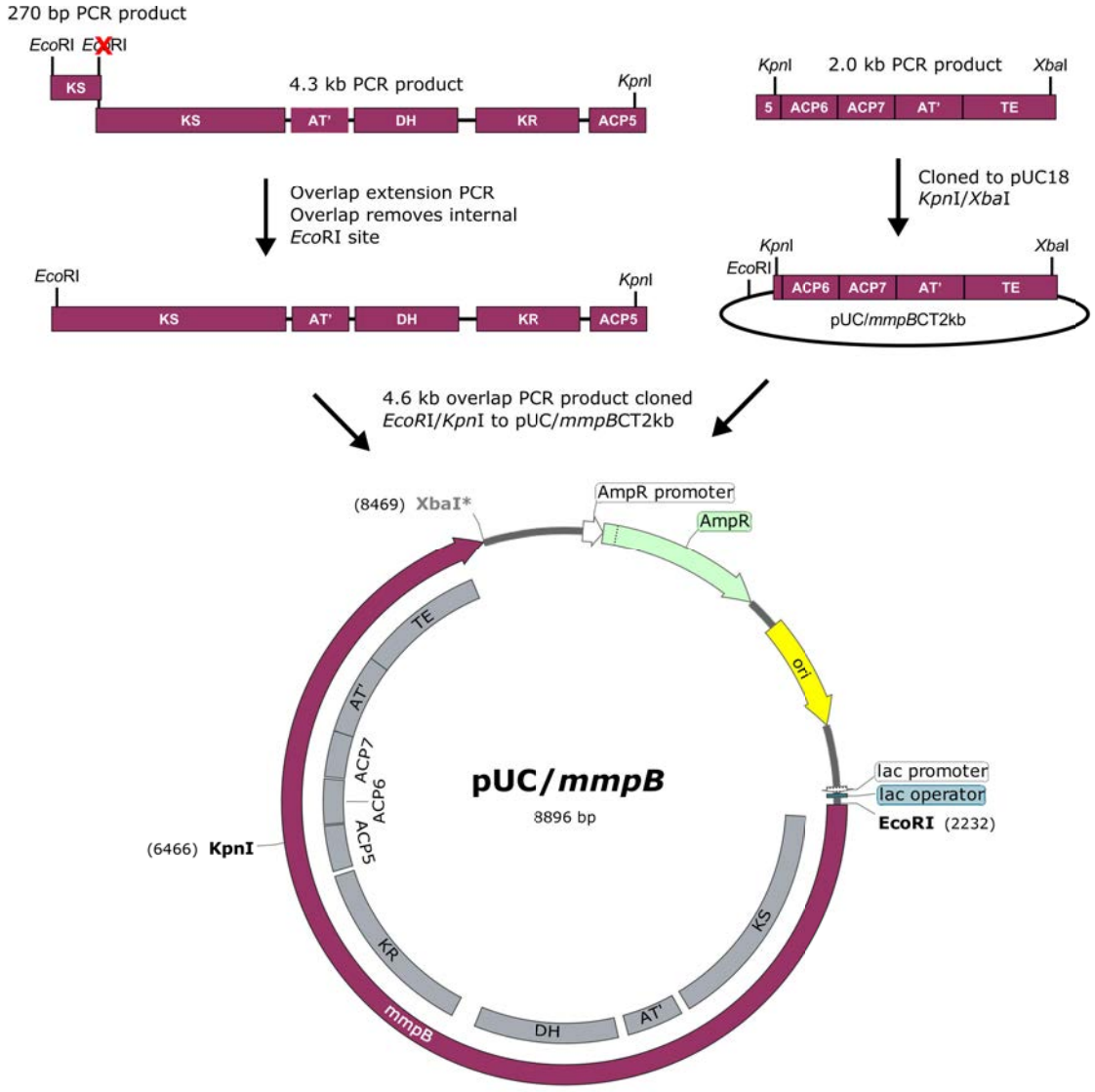
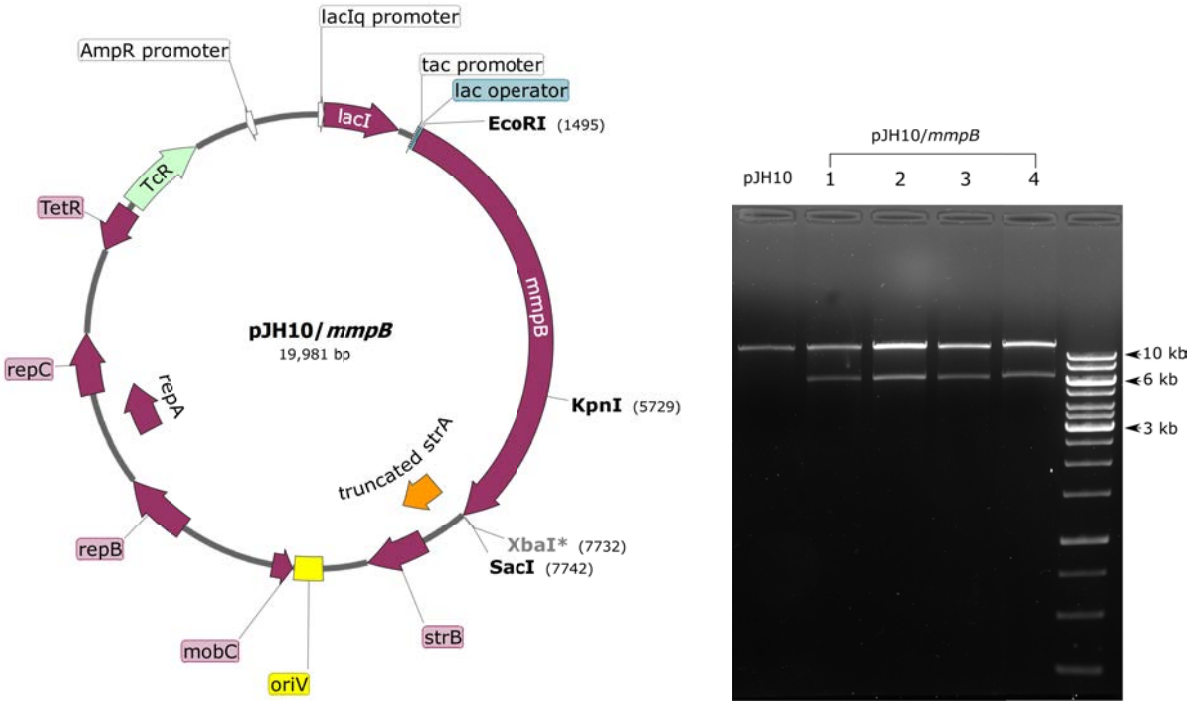


Figure 4.6: Cloning of *mmpB*.
The *mmpB* gene in pUC18 was confirmed by sequencing. The correct full gene was then subcloned *Eco*RI/*Xba*I to broad host range expression vector pJH10, to yield pJH10/*mmpB*.



(a) Plasmid map **(b)** *Eco*RI/*Sac*I digest

Figure 4.7: Construction of pJH10/*mmpB*.
(a) Map of pJH10/*mmpB*, *mmpB* is cloned *Eco*RI/*Xba*I downstream of the *tac* promoter. **(b)** *Eco*RI/*Sac*I diagnostic digests. Digest of pJH10/*mmpB* plasmid isolates yielded two bands corresponding to the 14 kb plasmid backbone and the 6.2 kb *mmpB* insert.

pUC/*mmpB*, and contained the full *mmpB* flanked by unqiue *Eco*RI and *Xba*I sites[1].

The *Xba*I site of pUC/*mmpB* is blocked by Dam methylation, therefore plasmid isolates were passaged through *dam*$^-$ *E. coli* ER2925. Potential pUC/*mmpB* isolates were checked by *Eco*RI/*Xba*I digest, and sequenced with M13(-40), M13(-48) and internal primers mmpBF5, mmpBF6, mmpBF7, and mmpBR4. Plasmids containing the correct *mmpB* sequence were identified, and the gene subcloned in expression vector pJH10 *Eco*RI/*Xba*I, yielding pJH10/*mmpB* (Figure 4.7a). Diagnostic digests with *Eco*RI and *Sac*I confirmed the presence of *mmpB* (Figure 4.7b) (*Sac*I was used in place of *Xba*I to avoid the need for passaging through a *dam*$^-$ strain).

The IncQ plasmid pJH10 has several key features that make it a suitable expression vector. It is derived from RSF1010, with its broad host range *oriV* for replication, and conjugation components *repA*, *repB* and *oriT* (Scholz et al., 1989), which allow

[1]Primer design and *mmpB* PCRs were performed by undergraduate project student Alicia Nash.

its conjugative transfer to *P. fluorescens* from donor strain *E. coli* S17-1. *E. coli* S17-1 provides the RP4 conjugative machinery from its chromosome (Simon et al., 1983). Transcription of cloned genes is driven by the *lac-trp* hybrid promoter, *tac*, and a ribosome binding site (RBS) included upstream of the *Eco*RI site. The vector also holds *lacI*, which is not present on the 10586 chromosome. LacI represses the *tac* promoter through binding of the *lac* operator, also allowing for IPTG induction (Amann et al., 1983).

4.2.1.2 *P. fluorescens* NCIMB 10586 Δ*mmpB* is complemented by pJH10/*mmpB*

Plasmids pJH10/*mmpB* and an empty pJH10 control were mobilised to *P. fluorescens* NCIMB 10586 WT and 10586 Δ*mmpB*. The Δ*mmpB* mutant has an in-frame chromosomal deletion of the majority of *mmpB*, removing 1991 amino acids (AA), leaving 30 AA at the N-terminus and 54 AA at the C-terminus. Complementation of 10586 Δ*mmpB* by pJH10/*mmpB* was observed by plate bioassay (Figure 4.8) and HPLC (Figure 4.9). On the bioassays, the deletion of *mmpB* was observed to prevent production of active mupirocin, and provision of *mmpB in trans* restored production to WT levels. HPLC analysis confirmed the bioassay observations; no detectable PA-A is produced by the Δ*mmpB* strain, and full complementation was observed.

Complementation of MmpB domain mutants 10586 *mmpB*ΔACP567 (Figure 4.10b) and 10586 *mmpB*ΔTE (not shown) was also observed. These mutations remove all acyl carrier proteins and the thioesterase domain, respectively, both resulting in no production of mupirocin.

Each of these experiments was initially performed with 500 μM IPTG induction and without induction. Complementation to WT level activity was observed without induction, whereas no complementation was observed with 0.5 mM IPTG induction. These observations held across all the *mmpB* mutants tested.

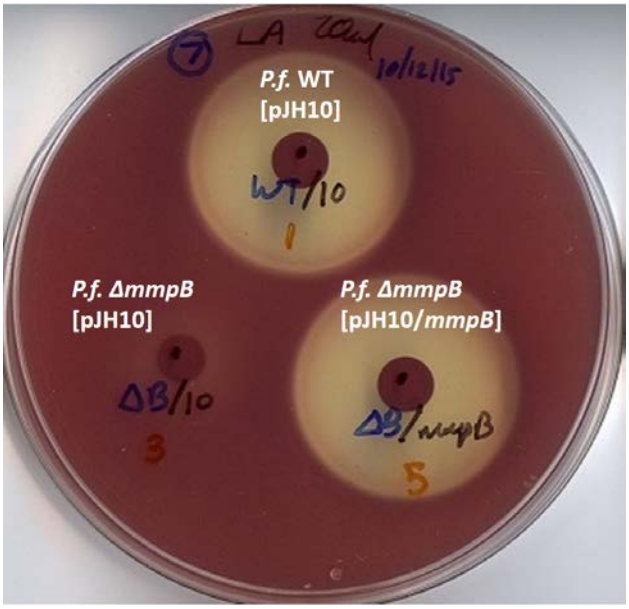


Figure 4.8: Bioassay demonstrating complementation of 10586 Δ*mmpB*.
Assay performed independently in triplicate, a representative plate is shown. No zone of inhibition is observed in the *P. fluorescens* 10586 Δ*mmpB* mutant, as MmpB is required for mupirocin production. When pJH10/*mmpB* is provided *in trans*, a similar zone of inhibition to WT was observed.

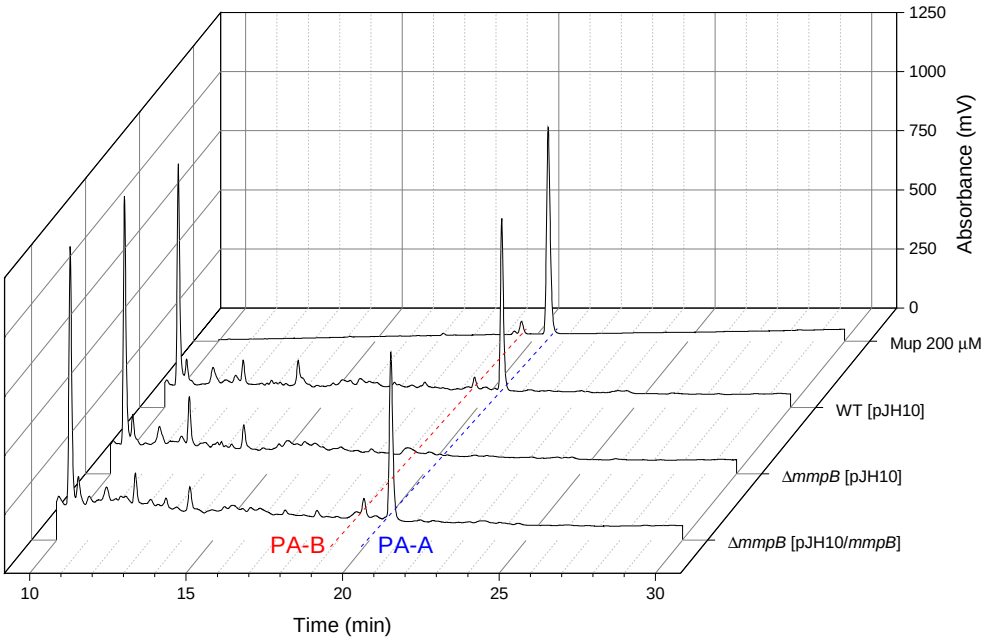


Figure 4.9: HPLC: complementation of 10586 Δ*mmpB* by pJH10/*mmpB*
Pseudomonic acid A elutes at approximately 20 minutes, as indicated by the standard, n=4. Complementation was observed to WT level, confirming the bioassay result.

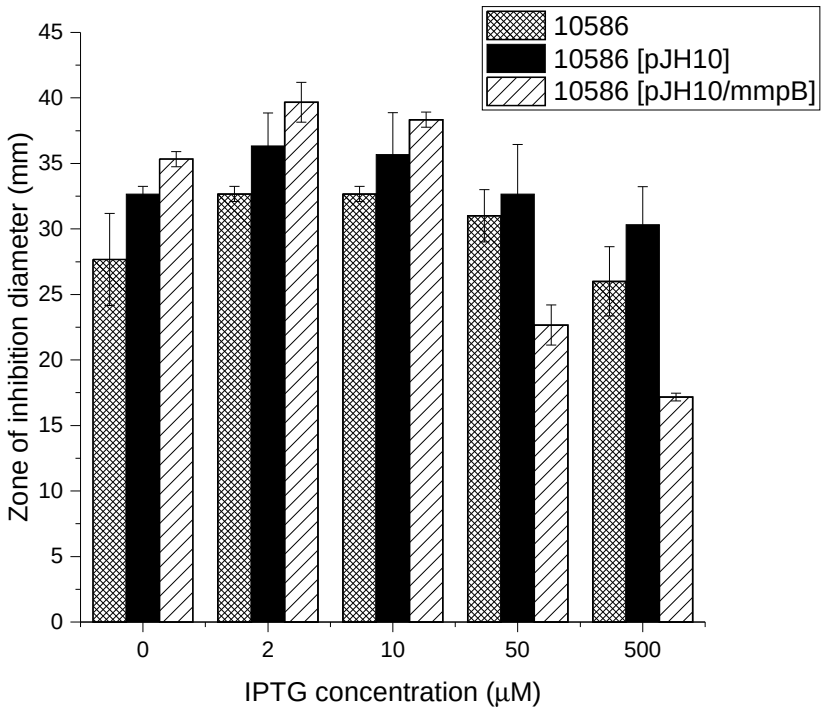
4.2.1.3 Induction of *mmpB* expression is inhibitory to mupirocin biosynthesis

In pJH10/*mmpB*, *mmpB* is positioned downstream of the *tac* promoter and *lac* operator. With IPTG induction, LacI repression is relieved, and the near consensus *tac* promoter will direct transcription of high levels of *mmpB*. A high level of IPTG induction (500 μM) was inhibitory to pJH10/*mmpB* complementation, whereas full complementation was observed without IPTG. However, there remains a wide space of induction concentrations to test between these, and it was conceivable that a lower level of induction could improve mupirocin production levels.

To investigate this, bioassays were performed at a range of IPTG concentrations on *P. fluorescens* WT and *mmpB*ΔACP567 (Figure 4.10). In the WT without plasmid, or with empty pJH10, IPTG concentration was observed to have no significant effect at any concentration (Figure 4.10a). However, at 500 μM IPTG induction, pJH10/*mmpB* was inhibitory to antibiotic production of the WT, whereas 50 μM had an intermediate effect. In 10586 *mmpB*ΔACP567 without a plasmid, or with empty pJH10, small zones corresponding to no mupirocin production were observed across the range of IPTG concentrations. Induction of pJH10/*mmpB* to complement this mutation appeared to work best at approximately 2 to 10 μM IPTG. Again the inhibitory effect was observed with high induction levels, with zones corresponding to the non-complemented mutant size observed at 500 μM IPTG.

To summarise, full complementation of *P. fluorescens* NCIMB 10586 Δ*mmpB* was observed without IPTG induction. Overexpression of *mmpB* on the plasmid was inhibitory towards mupirocin production in the WT producing strain, and correspondingly overexpression prevented complementation of *mmpB* mutant strains. These results opened the possibility of expressing modified versions of *mmpB in trans*. Strong IPTG induction proved inhibitory in all of the plasmid expression experiments performed with *mmpB* and its derivatives. Therefore, for all of the remaining plasmid *mmpB*, and *mmpB* derivatives, expression results presented in this chapter were performed without IPTG induction.

(a) 10586 WT



(b) 10586 *mmpB*ΔACP567

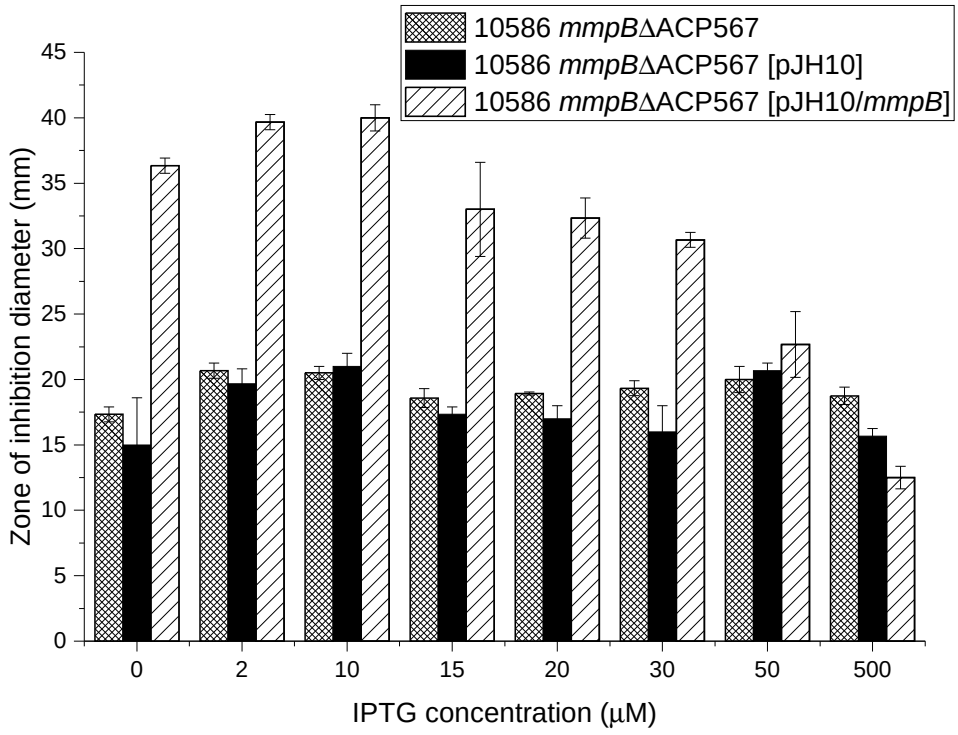


Figure 4.10: Bioassay of pJH10/*mmpB* across an IPTG induction gradient.
Assays were peformed independently in triplicate; error bars represent standard deviations. **(a)** In 10586 WT, pJH10/*mmpB* was inhibitory to antibiotic activity with high IPTG induction of 500 μM. **(b)** Complementation of 10586 *mmpB*ΔACP567 by pJH10/*mmpB* was observed to work best at 2-10 μM IPTG.

4.2.2 MmpB/TmpB hybrid generation I: Insertion of *tmpB* module 2 into *mmpB*

As discussed in Section 4.1.2, a key objective pursued in this work was to attempt to re-engineer the *mup* pathway to convert PA-B to PA-A *in cis* as may occur in the *tml* pathway (Figure 4.1). The second module of TmpB is comprised of the non-elongating KS^0 and an ACP, on which we proposed that PA-B is converted to PA-A.

Generation I consisted of a single MmpB/TmpB hybrid, formed by splicing the second module of TmpB (KS^0-ACP-TE) after ACP7 of MmpB, as summarised in Figure 4.11. The amino acid sequence of the junction between MmpB and TmpB sequence is shown in Figure 4.12. This was designed to show whether MmpB can still function with the extra module, and whether the TE domain of TmpB would function in the mupirocin system. The longer term goal was to re-engineer the pathway to convert PA-B to PA-A on the hybrid polypeptide, facilitated by these extra TmpB domains.

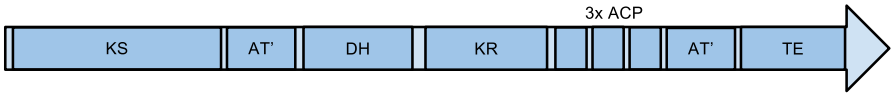
4.2.2.1 Molecular construction of MmpB/TmpB hybrid I

The second module of *tmpB* was amplified as three fragments using pTML1 as template DNA: primers TmpBF1 and TmpBR2 amplified 1354 bp, TmpBF2 and TmpBR3 amplified 499 bp, and TmpBF3 and TmpBR1 amplified 1217 bp [2]. These and all other primers that were used in the work described in this chapter are listed in Table 2.8. TmpBF2 and TmpBR2 were designed to overlap, whilst making silent mutations to remove a *Pci*I and an *Xba*I site. TmpBF3 and TmpBR3 were also designed to overlap, making a further silent mutation to remove a *Pci*I site.

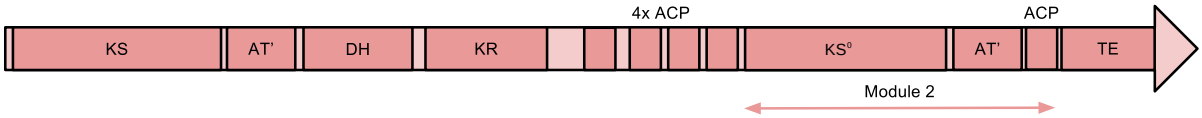
The three fragments were spliced by overlap extension PCR, using outer primers TmpBF1 and TmpBR1, which contained a *Pci*I and an *Xba*I site respectively. This yielded a 3031 bp fragment flanked by unique *Pci*I and *Xba*I sites at either end. The

[2]Primer design and initial *tmpB* PCRs were performed by undergraduate project student Wenqian Wang

(a) MmpB (2077 amino acids)



(b) TmpB (2751 amino acids)



(c) MmpB/TmpB hybrid generation I: JC50 (2620 amino acids)

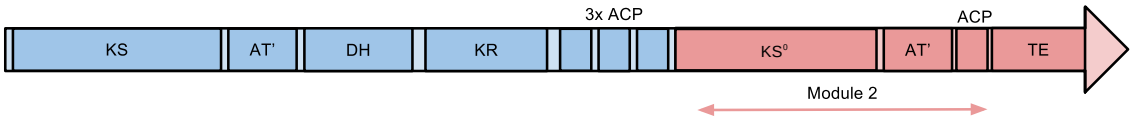


Figure 4.11: Domain map of MmpB, TmpB, and MmpB/TmpB hybrid I
Blue represents MmpB (mupirocin) domains and red TmpB (thiomarinol) domains.
(a) MmpB does not have a second module, PA-B is thought to be released from MmpB, and processing to PA-A, C8-OH removal, occurs on MacpE.
(b) TmpB has a second module, consisting of KS^0 and ACP, where the equivalent C8-OH removal putatively occurs.
(c) MmpB/TmpB hybrid generation I (JC50) was designed to have the front end of MmpB, up to and including the third of its ACP triplet, ACP7. After this, the second module of TmpB was fused on, ending with the TmpB TE domain.

fragment was AT cloned into cloning vector pGEM-T, yielding pJC41, and with the insert size checked by *Eco*RI, *Xba*I, and *Spe*I/*Aat*II digests (not shown). Two correct plasmids were confirmed by sequencing with primers M13 (-40), M13 (-48), TmpBF2 and TmpBR3.

The sub-cloning strategy used to generate the MmpB/TmpB hybrid expression plasmid is summarised in Figure 4.14. The 3 kb *tmpB* fragment was cut from pJC41 *Pci*I/*Xba*I, and cloned *Nco*I/*Xba*I to pUC/*mmpB*CT2kb, yielding pJC46. This made use of the compatible sticky ends of *Pci*I and *Nco*I, and as a result destroyed both sites in the product. The partial *mmpB tmpB* module 2 hybrid was then released from pJC46 by *Kpn*I/*Xba*I digestion and cloned *Kpn*I/*Xba*I to pJH10/*mmpB* to yield pJC50 (Figure 4.14).

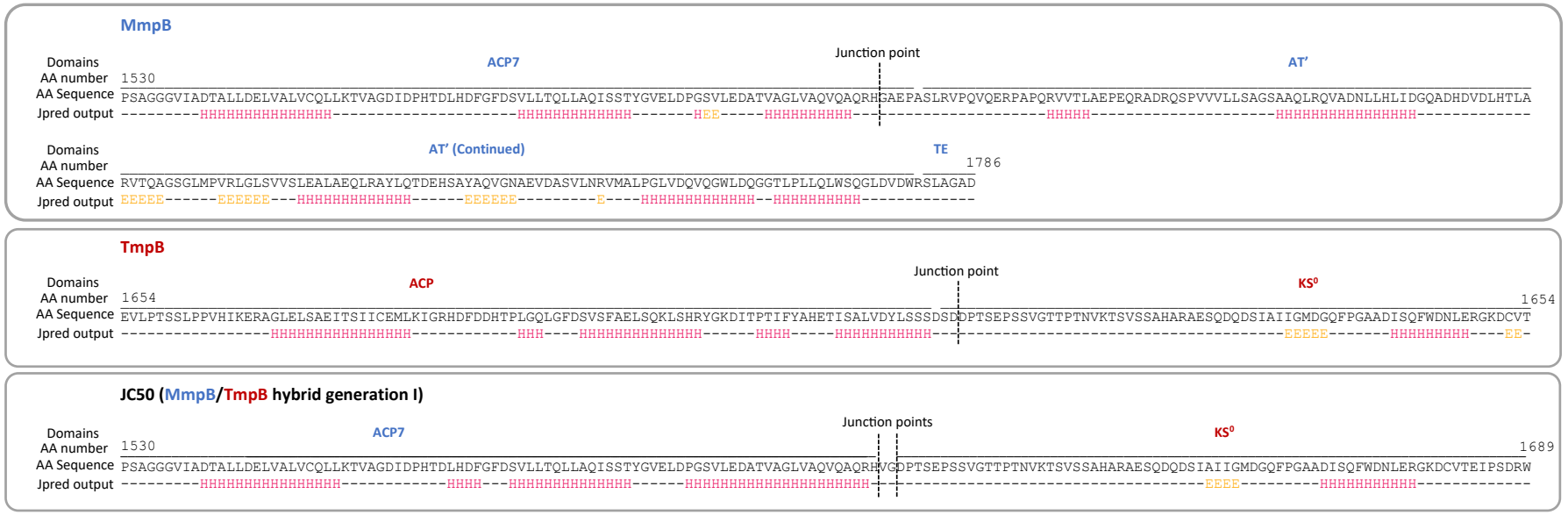


Figure 4.12: Amino acid sequence of the junction point between MmpB and TmpB domains in JC50.
The top line shows an approximate annotation of where each domain is positioned, the second line shows the amino acid number at both the start and end of the sequence displayed, respective to each protein. The third line shows the amino acid sequence, and the fourth line is the predicted secondary structure from Jpred (H is α-helix and E is β-strand). The valine and glycine between the junction points in JC50 were introduced as a by-product of the molecular construction strategy used, and do not originate from either MmpB or TmpB.

Potential pJC50 plasmids were checked by *Eco*RI/*Sac*I restriction digest (Figure 4.13) and sequencing with primers pJH10seqF, pJH10seqR, TmpBF2 and TmpBR3. This resulted in MmpB/TmpB hybrid generation I, consisting of a single expression plasmid, pJC50. Given the variation in motility of the smallest band from *Eco*RI/*Sac*I digestion, and some of the negative complementation phenotypes discussed in the following section, construction of pJC50 was repeated from independent PCR products. All the phenotypes reported were reproducible with both independently constructed replicates of pJC50.

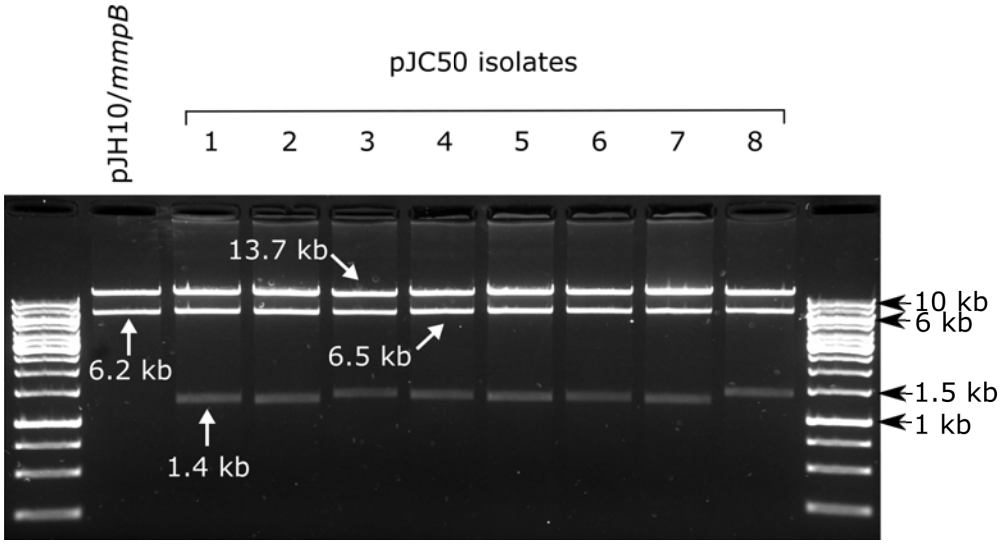


Figure 4.13: Diagnostic *Eco*RI/*Sac*I digest of pJC50.
All of the 8 potential pJC50 isolates contain a band at around 1.4 kb band indicative of presence of insert, due to the extra *Eco*RI site in *tmpB* module 2. Isolate 1 was confirmed correct by sequencing. The reason for the slight variations in the motility of the smallest band between potential pJC50 isolates is not known, and was not explored further.

4.2.2.2 The TmpB TE domain in the pJC50-encoded MmpB/TmpB hybrid does not off-load PA-B

To test if the first generation hybrid was functional, pJC50 was mobilised to 10586 Δ*mmpB*. No complementation of PA-A production was observed by HPLC or bioassay (not shown). A summary of the phenotypes of all MmpB/TmpB hybrids tested in this chapter is included at the end of this results section (Figure 4.49).

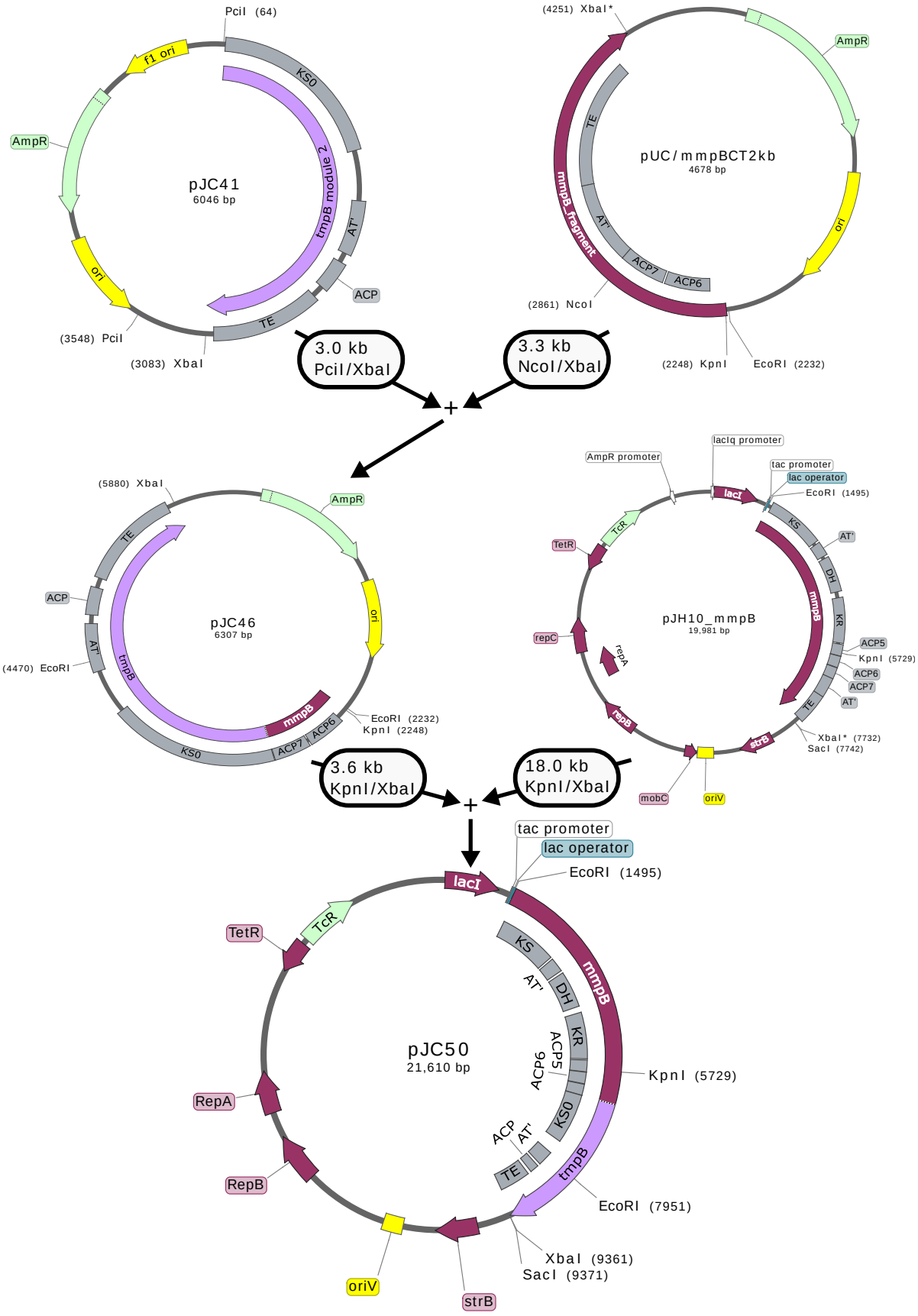


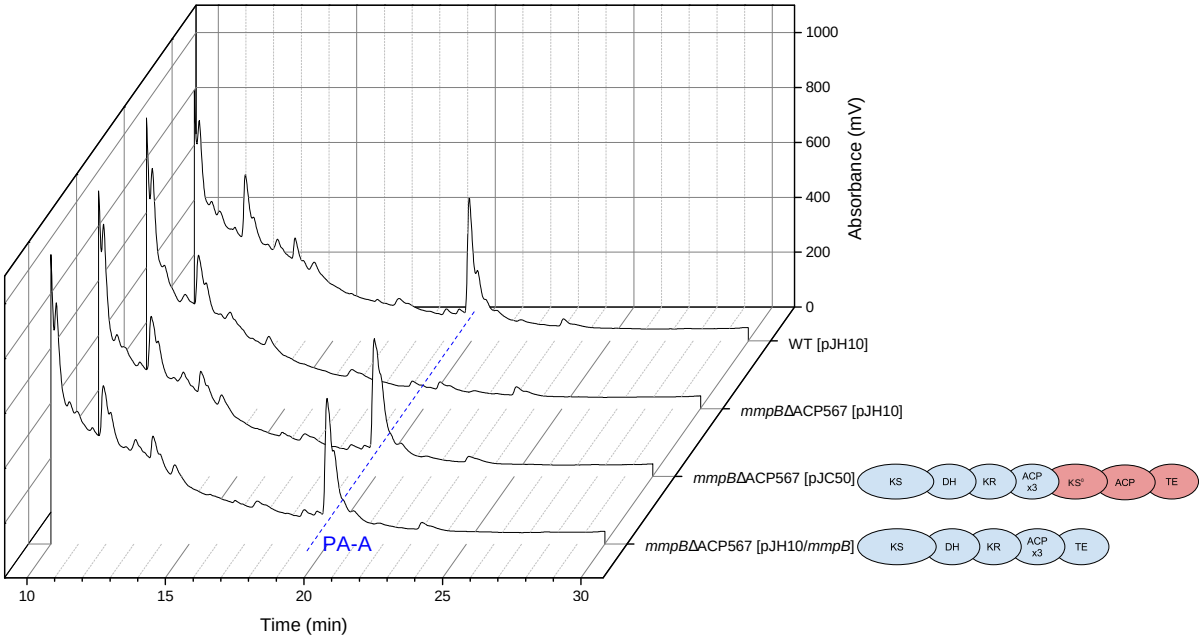
Figure 4.14: Maps of pJC50 and plasmids used in its construction.

This could have been due to the failure of one or more individual domains to complement, or insertion of the extra module could have disrupted the structure of MmpB, or the hybrid protein may not be expressed from pJC50. Plasmid pJC50 was mobilised to the two available *mmpB* domain deletion strains: 10586 *mmpB*ΔACP567, and 10586 *mmpB*ΔTE. Mupirocin production was determined by HPLC of SSM culture supernatants.

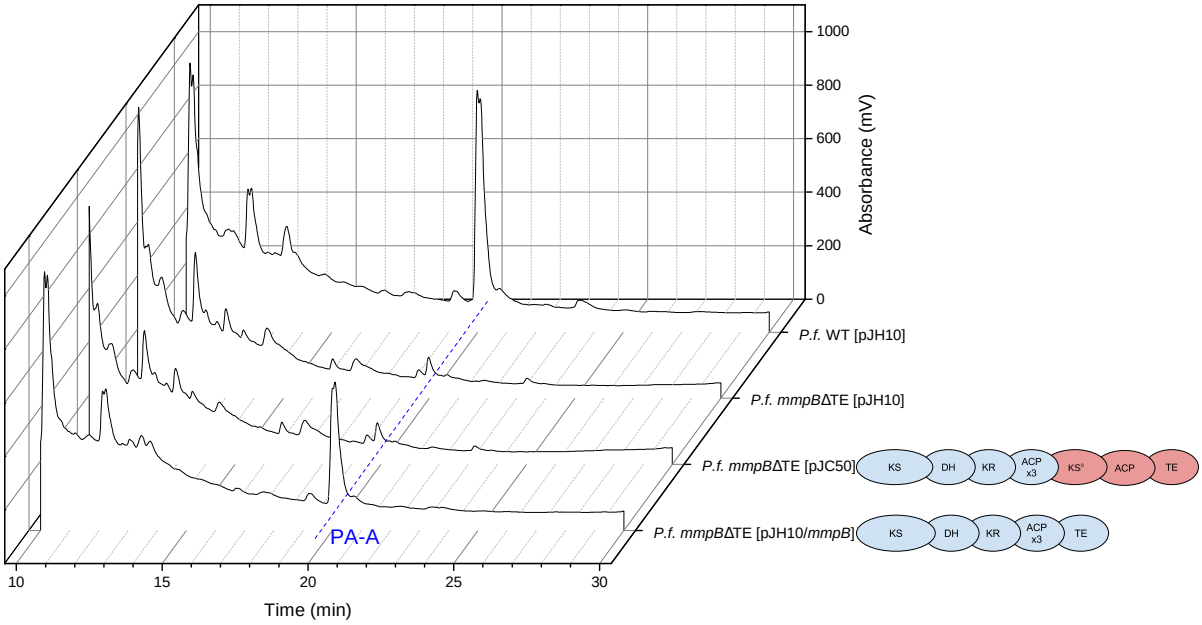
P. fluorescens 10586 *mmpB*ΔACP567 carries an in-frame deletion of the ACP triplet in MmpB. As expected, without any ACPs MmpB is non-functional and this strain does not produce PA-B or PA-A. Suprisingly, pJC50 was observed to complement the PA-A production of 10586 *mmpB*ΔACP567 (Figure 4.15a). This logically suggests that pJC50 the failure to complement 10586 Δ*mmpB* is not due to a lack of expression of the hybrid, or a wide scale disruption of its protein structure.

However, no complementation of 10586 *mmpB*Δ*TE* was observed (Figure 4.15b). This could suggest that the TmpB TE domain carried on hybrid I on pJC50 is not functional in mupirocin biosynthesis.

As was the case for pJH10/*mmpB* (Section 4.2.1.3), induction of expression of the hybrid from pJC50 by 0.5 mM IPTG proved inhibitory to mupirocin biosynthesis in 10586 WT, and complementation of 10586 *mmpB*ΔACP567 (not shown).



(a) Production of MmpB/TmpB hybrid I encoded by pJC50 in 10586 *mmpB*ΔACP567



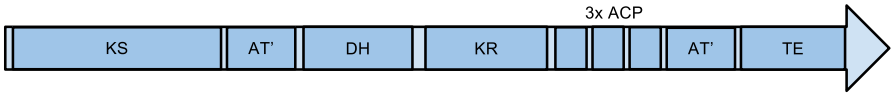
(b) Production of MmpB/TmpB hybrid I encoded by pJC50 in 10586 *mmpB*ΔTE

Figure 4.15: HPLC: the TmpB TE domain in hybrid I may not be functional. Representative chromatograms shown, n=6. The ellipses on the right are a reminder of MmpB variants, blue represents MmpB domains, and red TmpB domains.
(a) Complementation of PA-A production in 10586 *mmpB*ΔACP567 was observed to WT levels by pJC50 and pJH10/*mmpB*.
(b) Complementation of PA-A production in 10586 *mmpB*ΔTE was observed by pJH10/*mmpB*, whereas no complementation was observed for pJC50.

4.2.3 MmpB/TmpB hybrid generation II: insertion of the MmpB thioesterase domain into hybrid I

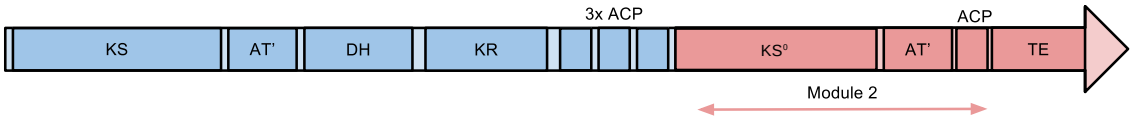
The failure of MmpB/TmpB hybrid I (pJC50) to complement 10586 *mmpB*ΔTE suggested the TmpB TE domain may not be functional in mupirocin biosynthesis. Therefore hybrid generation II was designed, where the TmpB TE was replaced with the native MmpB TE (Figure 4.16).

(a) MmpB (2077 amino acids)



(b) TmpB (2751 amino acids)

(c) MmpB/TmpB hybrid generation I: JC50 (2620 amino acids)



(d) MmpB/TmpB hybrid generation II: JC58, JC59 (2658, 2619 amino acids)

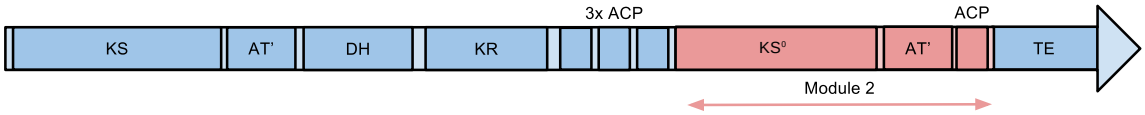


Figure 4.16: Domain map of MmpB, TmpB and MmpB/TmpB hybrid generations I and II.
Blue represents MmpB (mupirocin) domains and red TmpB (thiomarinol) domains. **(d)** MmpB/TmpB hybrid generation II has the MmpB TE domain spliced after module 2 from TmpB. Generation II hybrid JC58 has the full length MmpB TE domain, whereas JC59 has an N-terminal truncation to the TE domain to match the TmpB TE size.

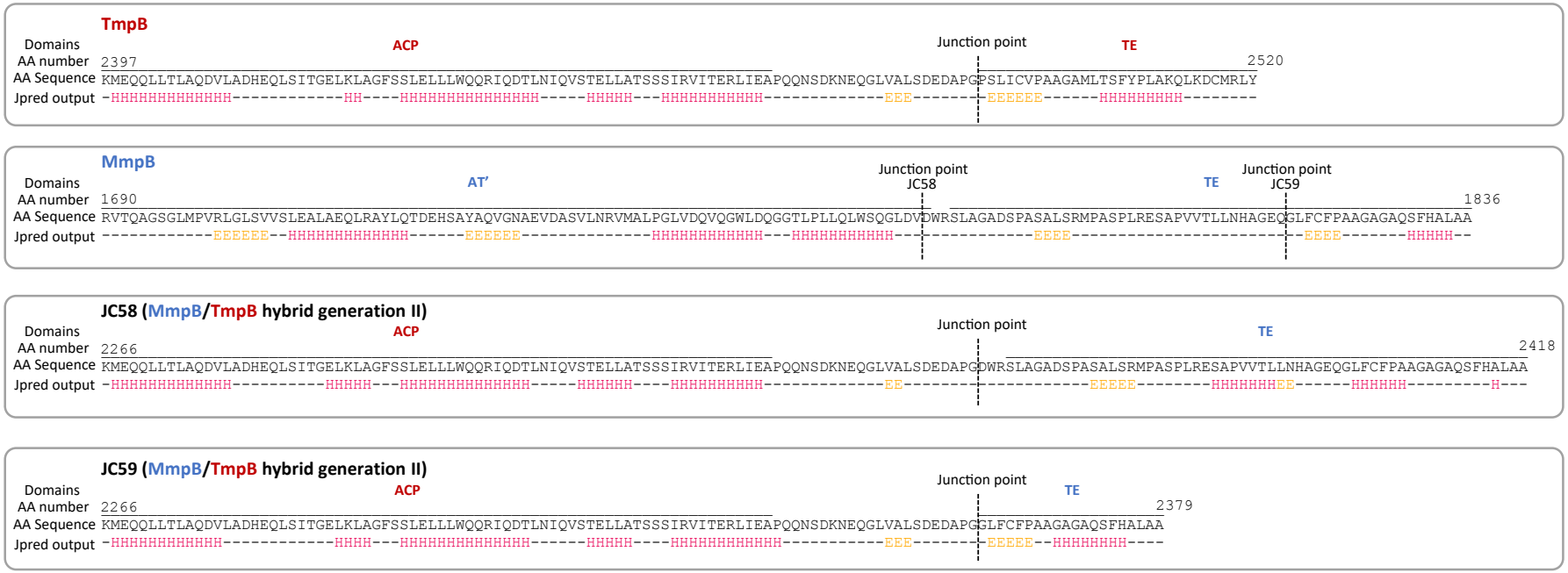


Figure 4.17: Amino acid sequence of the junction point between the TmpB ACP and MmpB TE in MmpB/TmpB hybrid generation II.

The top line shows an approximate annotation of where each domain is positioned, the second line shows the amino acid number at both the start and end of the sequence displayed, respective to each protein. The third line shows the amino acid sequence, and the fourth line is the predicted secondary structure from Jpred (H is α-helix and E is β-strand).

The junction between the module 2 ACP and MmpB TE domain was decided based on an alignment of MmpB and TmpB TE domains. The amino acid sequence of the junction is shown in Figure 4.17. The domains have 33% amino acid identity, yet align reasonably well without gaps, supported by the secondary structure prediction (Figure 4.18). Interestingly, the MmpB TE domain is 35 AA longer at the N-terminal side than TmpB. This seemed a real difference rather than an annotation inaccuracy, as if the TmpB TE domain were annotated to the same length, it would cross the preceding ACP domain. To investigate the significance of this region, two versions were designed in MmpB/TmpB hybrid generation II: one with full length MmpB TE and one truncated to the TmpB length.

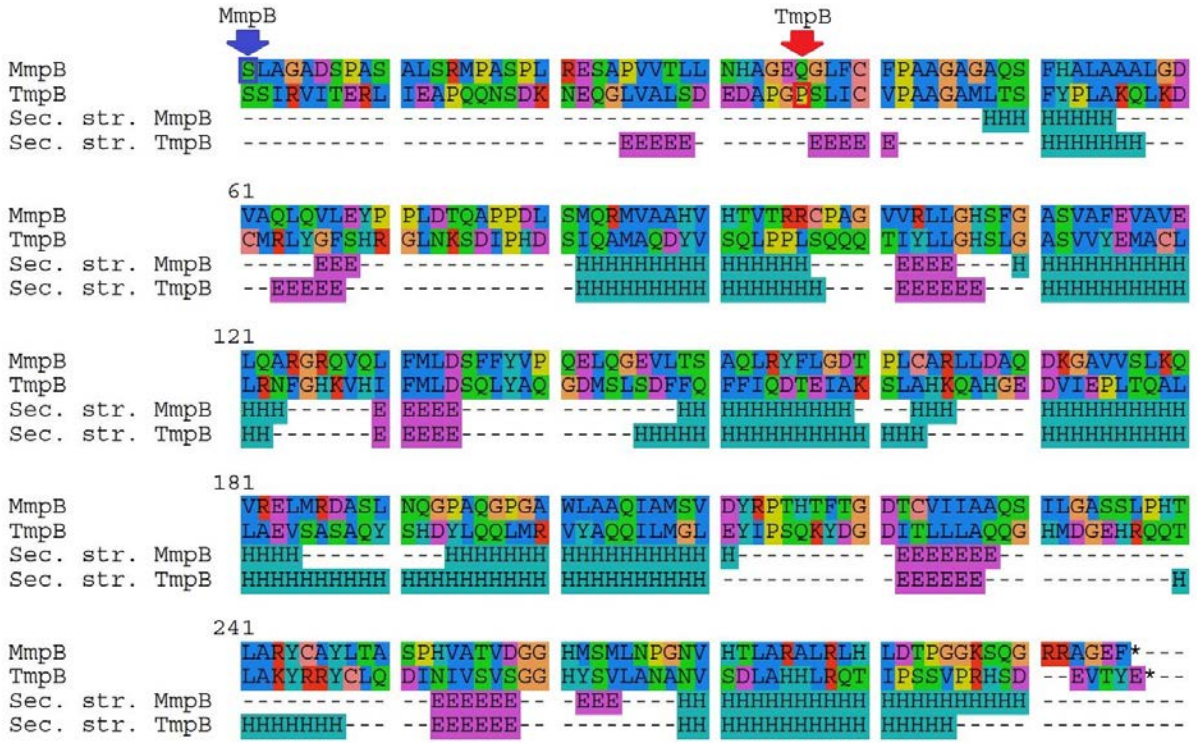


Figure 4.18: Alignment of MmpB and TmpB TE domains.
The first two lines are the amino acid sequence of MmpB and TmpB respectively. Jpred secondary structure predictions are included underneath, E is extended (β-strand), H is α-helix.

4.2.3.1 Molecular construction of MmpB/TmpB hybrid generation II

Construction of MmpB/TmpB hybrid generation II plasmids was performed in two stages. Firstly, overlap PCR was used to generate the desired hybrid sequences from the *Kpn*I site within ACP5 to the *Xba*I site following the hybrid, which was cloned in pUC18 (Figure 4.19). Secondly, these *Kpn*I/*Xba*I fragments were subcloned in pJC50 (hybrid I in expression vector pJH10), to give hybrid generation II expression vectors (Figure 4.20).

Stage 1: Construction of *Kpn*I/*Xba*I hybrid generation II fragments

A 186 bp fragment was amplified from pJC46 using primers F_pJC46_ACPfrag and R_pJC46_ACPfrag (Figure 4.19). This provided a bridge from the native *Bam*HI site within the second module ACP to the desired junction point, the start of the TE domain. This fragment was designed to overlap with two *mmpB* fragments, one full length and one truncated to *tmpB* TE length. The full length *mmpB* TE fragment was amplified from pJH10/*mmpB* using primers F_mmpB_full_TE and R_mmpB_TE, whereas the short length used F_mmpB_short_TE and R_mmpB_TE.

The ACP bridging fragment and the two different length TE fragments were joined by overlap extension PCR with primers F_pJC46_ACPfrag and R_mmpB_TE. The resulting fragments were digested *Bam*HI/*Xba*I (*Xba*I was included in R_mmpB_TE) and cloned into pJC46. This yielded pJC56 and pJC57, pUC18 based vectors carrying the full length, and short length TE fragments respectively (Figure 4.19).

Stage 2: Construction of hybrid generation II expression vectors

The hybrid fragments were released from both pJC56 and pJC57 by *Kpn*I/*Xba*I digest, and cloned in pJC50, yielding pJC58 and pJC59 respectively (Figures 4.20 and 4.21). The resulting plasmids are pJH10 derivatives with MmpB/TmpB hybrid II

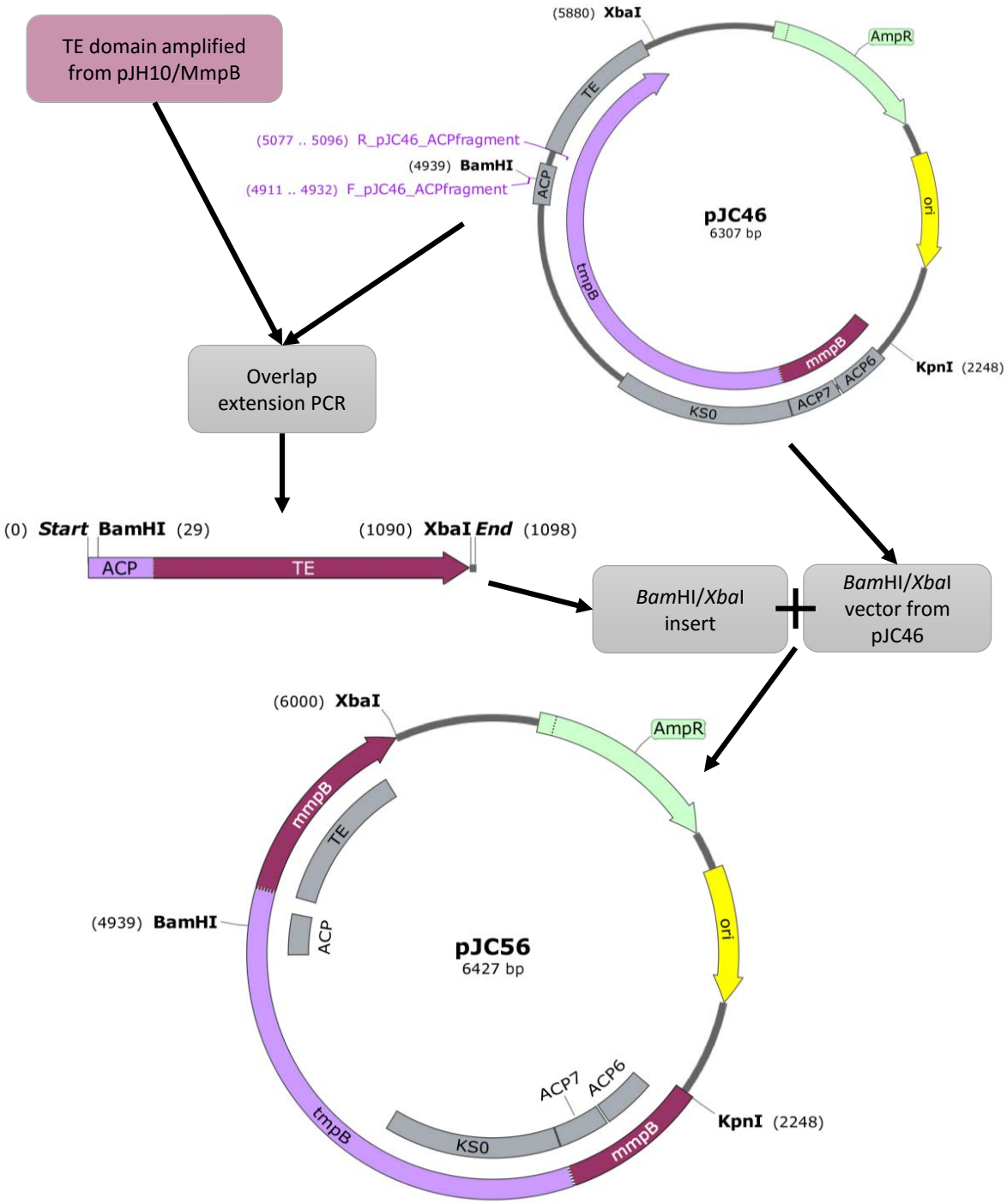


Figure 4.19: Stage 1: Construction of MmpB/TmpB hybrid II intermediate vectors pJC56 and pJC57.
To build MmpB/TmpB hybrid 2, PCR product of the *mmpB* TE domain was overlapped with a short linker in the ACP domain amplified from pJC46. This was cloned into pJC46 *Bam*HI/*Xba*I, yielding pJC56. The same steps were also performed with a shorter version of the *mmpB* TE, truncated to *tmpB* TE length, yielding pJC57 (not shown). The resulting pJC56 and pJC57 are pUC18 derivatives carrying the TE alteration ready for cloning *Kpn*I/*Xba*I.

positioned under control of the *taq* promoter. These were confirmed by restriction digest (Figure 4.22), and sequencing with primer pJH10_seq_R, which gave good reads through the TE swaps up to the ACP domain of module 2.

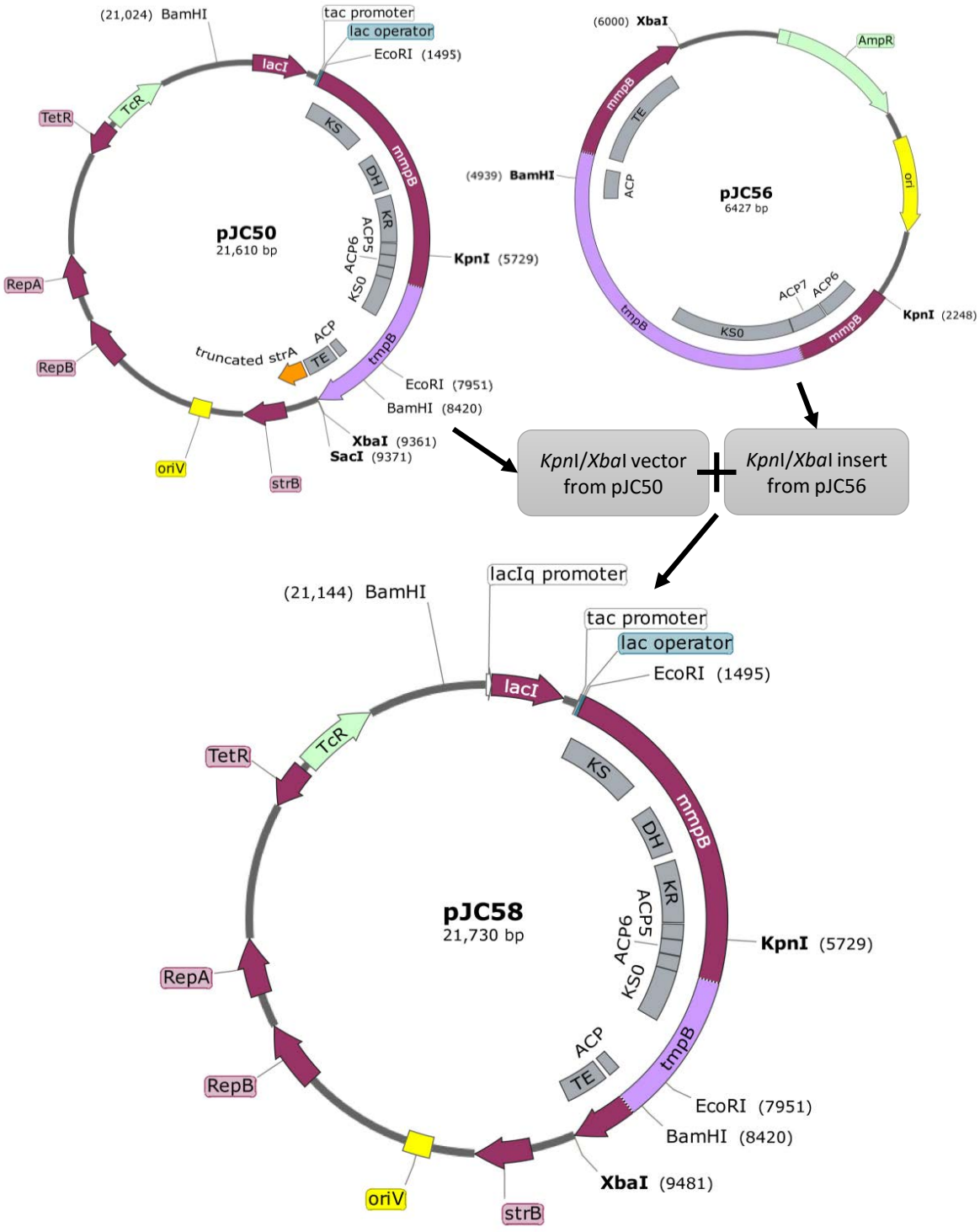


Figure 4.20: Stage 2: Construction of MmpB/TmpB hybrid II expression vectors pJC58 and pJC59.
The full length and truncated swapped TE domain fragments were excised using *Kpn*I/*Xba*I from pJC56 and pJC57 respectively. These were cloned into pJC50, which is a pJH10 derivative carrying MmpB/TmpB hybrid I.
The resulting plasmid, pJC58, is a pJH10 derivative carrying MmpB/TmpB hybrid II positioned downstream of the *tac* promoter. The truncated TE equivalent plasmid is pJC59.

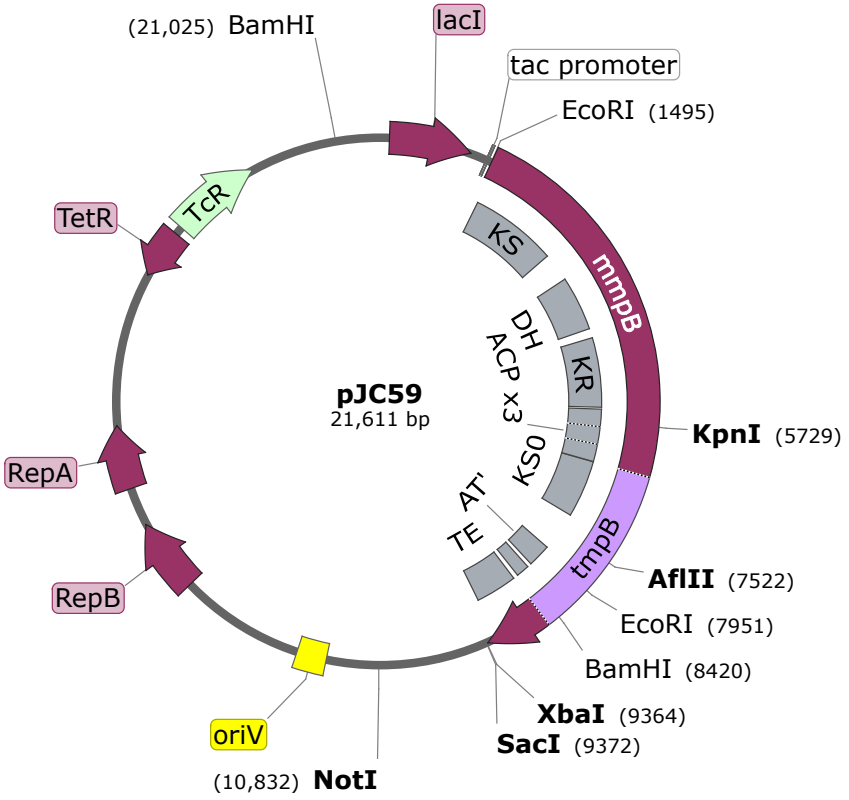


Figure 4.21: Map of pJC59, MmpB/TmpB hybrid generation II expression plasmid with truncated TE.

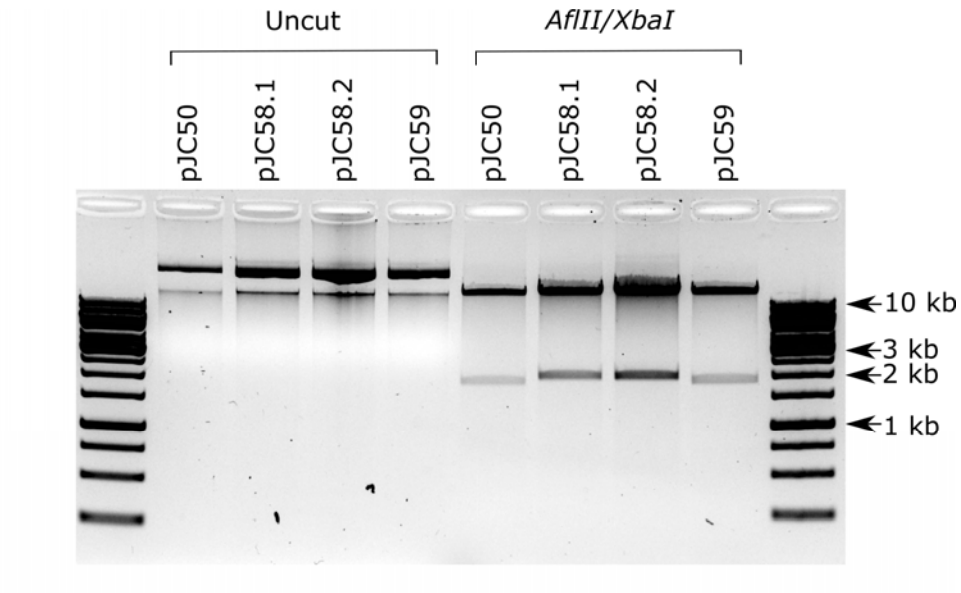


Figure 4.22: Diagnostic digests of pJC58 and pJC59.
*Afl*II/*Xba*I digest released the same 19.7 kb fragment from all plasmids. A smaller fragment corresponding to 1.84 kb was released from pJC50 (MmpB/TmpB hybrid I) and pJC59 (MmpB/TmpB hybrid II, with truncated TE). A larger 1.96 kb fragment was released from pJC58, due to the longer TE domain. Plasmids pJC58 and pJC59 were confirmed to have the correct *mmpB* TE domain DNA by sequencing.

4.2.3.2 MmpB/TmpB hybrid II is functional in *P. fluorescens* NCIMB10586

MmpB/TmpB hybrid generation II expression vectors pJC58 and pJC59 were mobilised to *P. fluorescens* mutants: Δ*mmpB*, *mmpB*ΔTE, and *mmpB*ΔACP567. Complementation was tested by plate bioassay against *B. subtilis* and HPLC of SSM culture supernatants.

As expected, deletion of the ACP triplet, the TE domain or the entire *mmpB* disrupted mupirocin production with a decrease in bioassay inhibition zone to 10% of WT, when carrying empty vector pJH10 (Figure 4.23). Both of the Mmp/TmpB hybrid II expression plasmids were observed to complement these mutations to approximately 70 to 75% of WT zone area. The complementation of PA-A production was confirmed by HPLC analysis (Figure 4.24). These data suggest that both variants MmpB/TmpB hybrid generation II can perform normal MmpB function, thought to be the biosynthesis of 9-hydroxynonanoic acid.

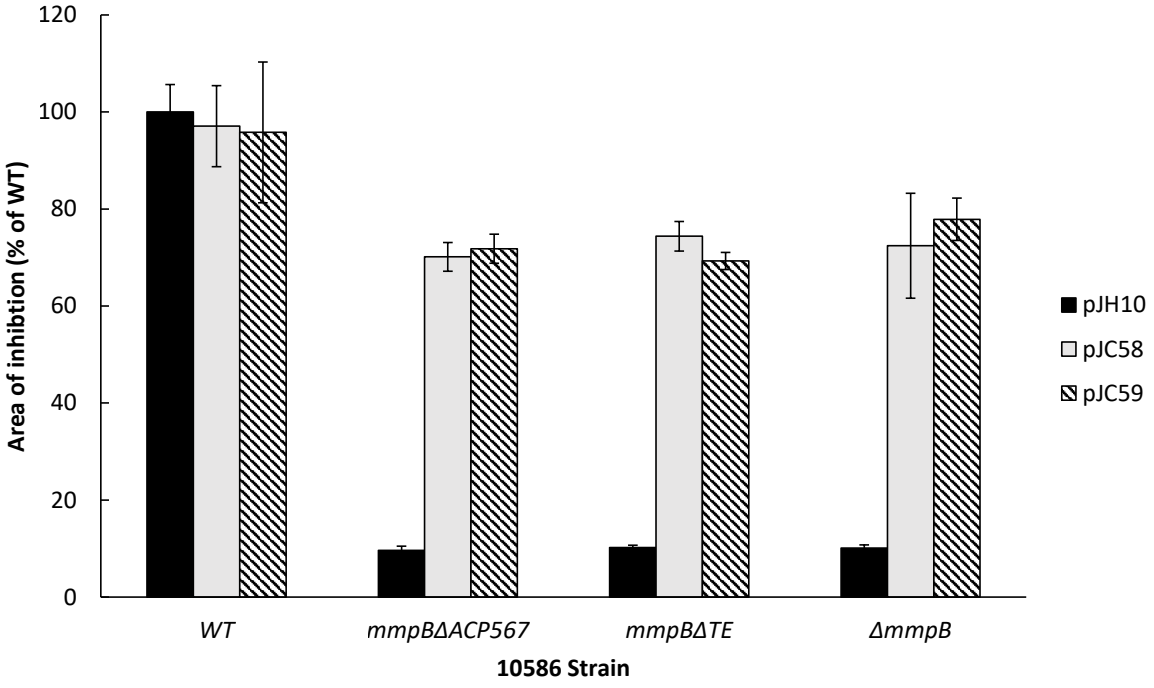
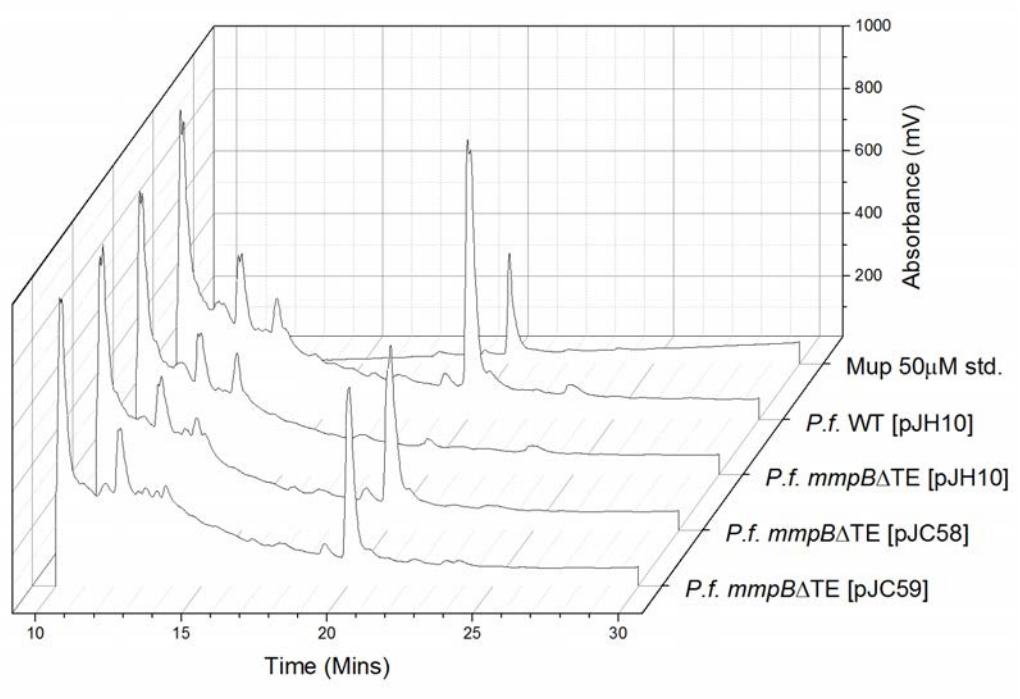
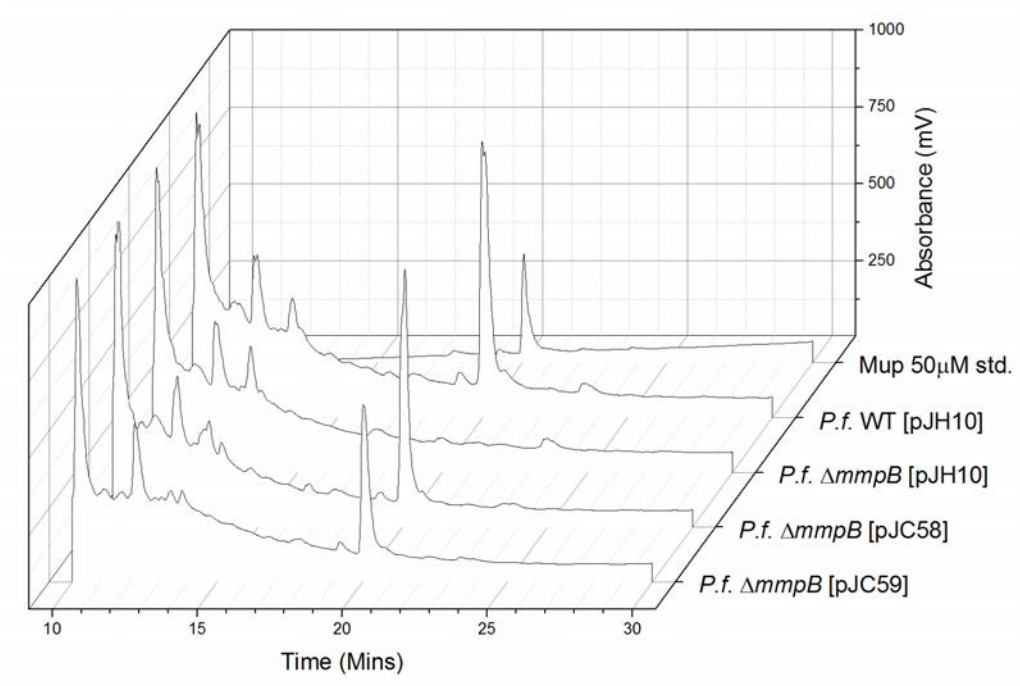


Figure 4.23: Complementation of 10586 *mmpB* mutants by MmpB/TmpB hybrid generation II plasmids pJC58 and pJC59.
Plate bioassays against *B. subtilis*, performed independently in duplicate, error bars are standard deviation.



(a) MmpB/TmpB hybrid generation II in 10586 *mmpB*ΔTE



(b) MmpB/TmpB hybrid generation II in 10586 Δ*mmpB*

Figure 4.24: HPLC: MmpB/TmpB hybrid II is functional.
In this experiment, PA-A eluted at approximately 20 minutes, as demonstrated by the standard. Experiments were performed independently in triplicate, representative chromatograms are shown.

Hybrid I (TmpB TE) failed to complement 10586 *mmpB*ΔTE and 10586 Δ*mmpB*, whereas hybrid generation II (MmpB TE) complemented these mutants. This suggests that the reason for the non-function of Hybrid I (pJC50) is that the TmpB TE domain is not offloading PA-B in this scenario.

Through integration of HPLC peak areas, full length TE domain plasmid pJC58 restored an average of 189 mV^2 PA-A peak, compared with 149 mV^2 in the truncated TE version, pJC59. However, a two-tailed unpaired *t* test indicated that this difference is not significant at the 95% confidence level, $p=0.24>0.05$, n=3. Indeed, no difference between the versions was observed when used as controls in later experiments. Therefore, the additional 35 amino acids at the N-terminus of the MmpB TE that are not present in the TmpB TE do not seem to be essential for function.

The key conclusion from this experiment is that MmpB can remain structured and functional after insertion of the second module of TmpB. This result is a key step towards the development of a hybrid on which PA-B can be converted to PA-A, in an *in cis* manner as it may occur on TmpB. However, these data do not yet demonstrate that the second module KS^0 is actively functional in the hybrids.

4.2.3.3 Inactivation of the second module KS^0 in MmpB/TmpB hybrid generations I and II by point mutation

To test whether the TmpB KS^0 inserted into the hybrids is functional, a point mutation was designed to the active site cysteine. If the KS^0 function is required, this should disrupt the complementation phenotype of hybrid generation II. The mutation was also performed in hybrid I for completeness. The active site cysteine was readily identified by NCBI conserved domains as Cys1824, and confirmed by an alignment with BLASTp hits. A mutation was designed to change this to alanine, C1824A. The equivalent mutation to DEBS KS1, C211A, completely disrupted catalytic activity (Robbins et al., 2016).

Molecular construction of C1824A mutants

The construction strategy used is overviewed in Figure 4.25. A 1.3 kb 5′ fragment was amplified using pJC58 as template, using primers MmpBF3 and R_tmpB_C1824A. This spanned from the *Kpn*I site within ACP5 to the C1824A mutation within the KS^0, and this fragment was the same for all hybrids.

A second fragment was generated spanning from C1824A to the end of the protein; this fragment was different for each hybrid. For MmpB/TmpB hybrid I, a 2.5 kb fragment was amplified from pJC50 using primers F_tmpB_C1824A and pJH10_Seq_R. For MmpB/TmpB hybrid generation II a 2.6 or 2.5 kb fragment was amplified with the same primers from pJC58 (full length TE) and pJC59 (truncated TE) respectively.

Each of these 2.5 or 2.6 kb fragments were spliced to the shared 5′ 1.3 kb fragment by overlap extension PCR, with the overlap designed to introduce the C1824A mutation. Each were AT-cloned in pGEM-T, checked by *Aat*II/*Pst*I digest, and sequenced using primers TmpBF1, M13(-40) and M13(-48). The three fragments were released by *Kpn*I/*Xba*I digest, and cloned in pJC58. This yielded plasmids pJC50.C1824A, pJC58.C1824A and pJC59.C1824A, each the KS^0 mutant variant of their parent namesake. The C1824A mutation introduced an additional *Pst*I site, which facilitated detection by digestion (Figure 4.26).

4.2.3.4 Inactivation of the second module KS^0 is inhibitory to MmpB/TmpB hybrid II activity

The KS^0 C1824A mutant hybrid expression plasmids pJC50 (generation I), pJC58 and pJC59 (generation II) were mobilised to 10586 Δ*mmpB* to determine their effect on mupirocin production. Mupirocin production was determined by plate bioassay (Figure 4.27) and HPLC of SSM culture supernatants (Figure 4.28). As in Section 4.2.2.2, no complementation of 10586 Δ*mmpB* by hybrid I expressed from pJC50 or its KS^0 mutant version pJC50.C1824A was observed by bioassay or HPLC (not shown).

For hybrid generation II, pJC59 (truncated TE) was observed to complement 10586

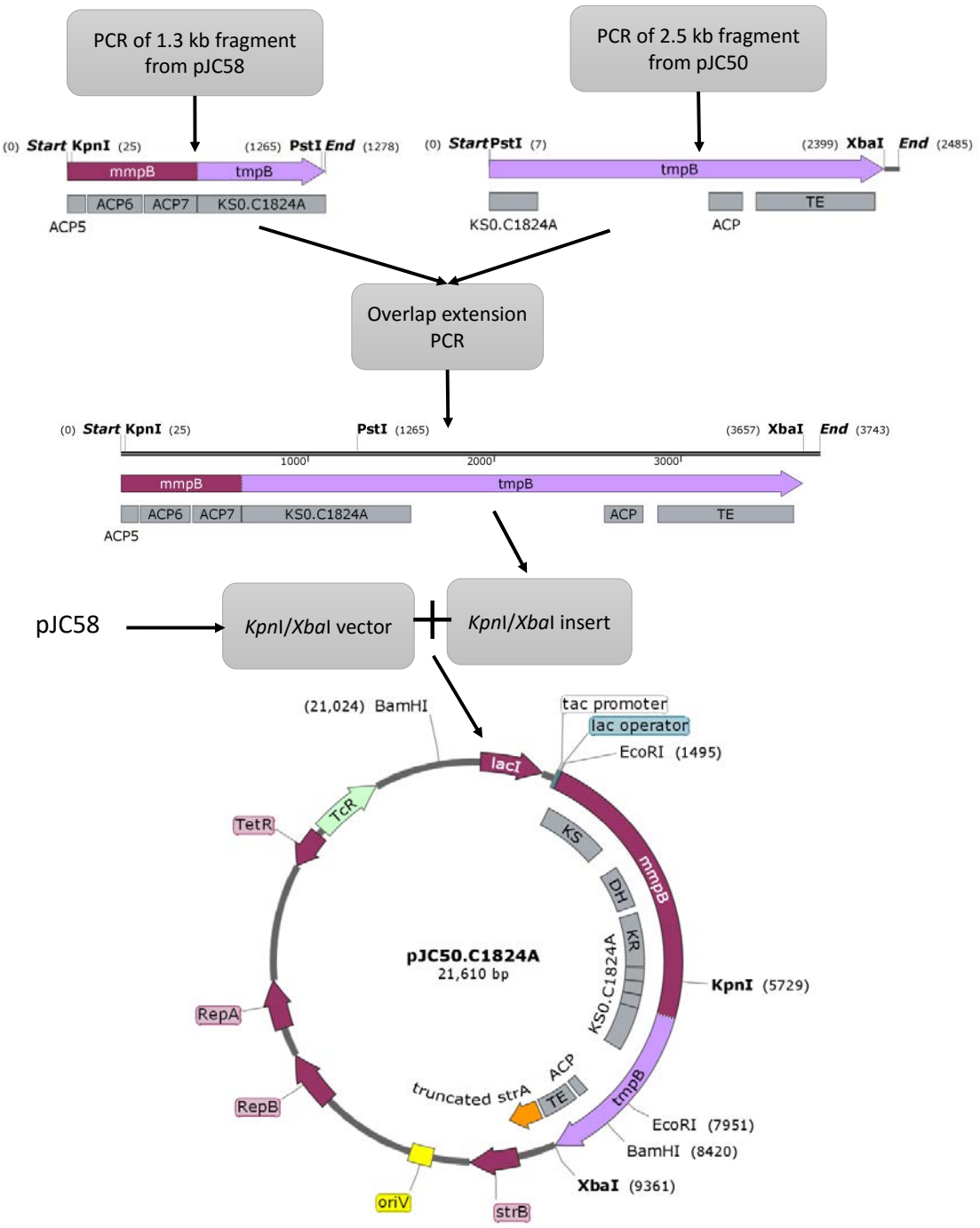
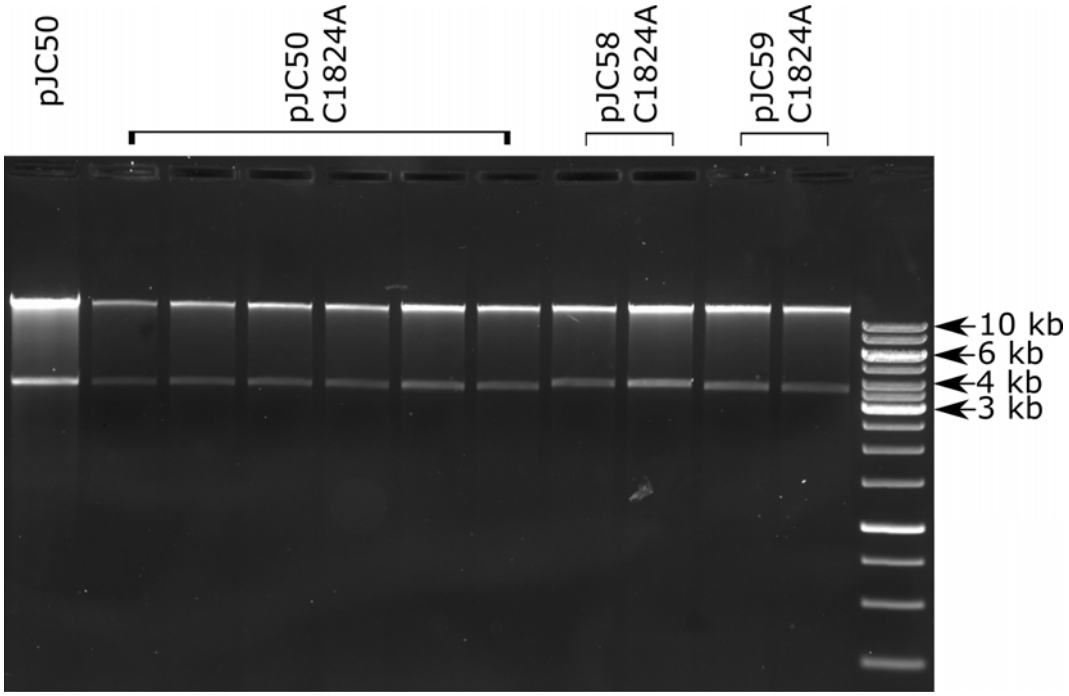
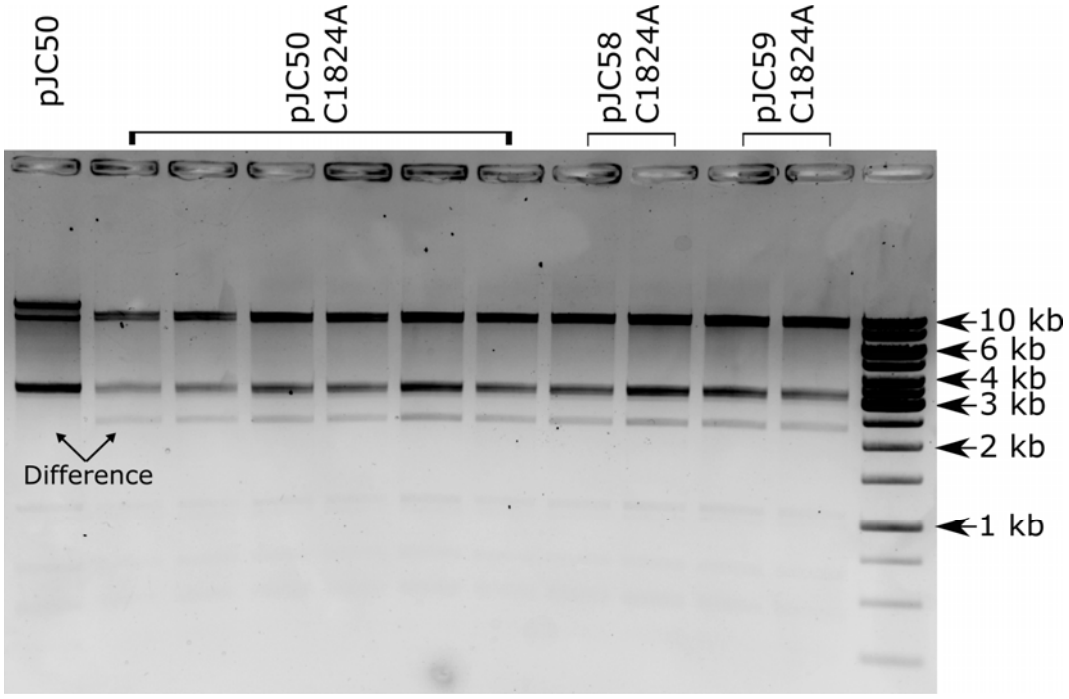


Figure 4.25: Construction of C1824A MmpB/TmpB hybrid expression vectors. A 1.3 kb PCR product was overlapped with a 2.5 kb product to yield a 3.8 kb *Kpn*I/*Xba*I fragment carrying the KS^0 mutation, C1824A. This overlapped fragment was cloned *KpnI*/*Xba*I to pJC58, the MmpB/TmpB hybrid II expression vector. The diagram shows construction of pJC50.C1824A, mutants of the other hybrids were constructed the same way, through use of respective template DNA for the 2.5 kb inital PCR.

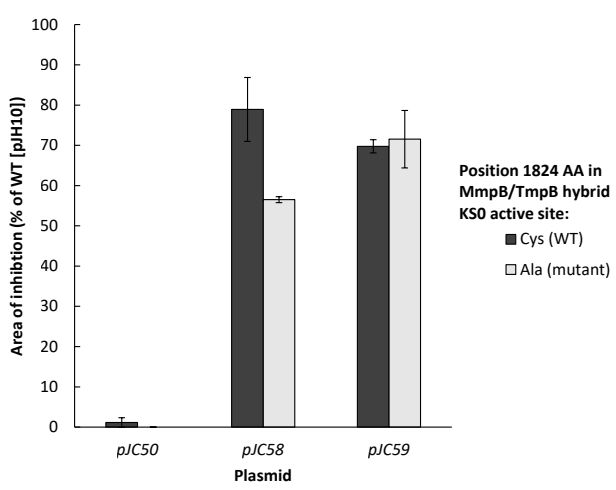


(a) ***Xba*I/*Kpn*I digests**

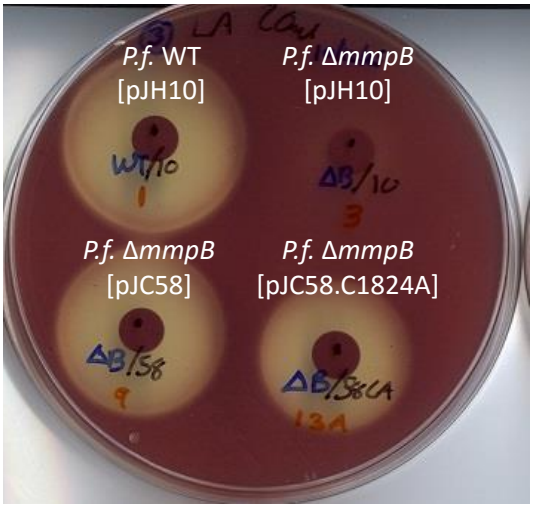


(b) ***PstI* digests**

Figure 4.26: Diagnostic digests of C1824A mutant hybrids in expression vectors.
(a) *Xba*I/*Kpn*I digest releases a fragment from ACP5 to the end of the gene: 3.6 kb in *tmpB* TE length hybrids (pJC50 and pJC59) and 3.8 kb in the *mmpB* TE length variant (pJC58).
(b) The C1824A mutation introduced a *Pst*I site. In pJC50, without the mutation, there is a 12.3 kb band, which is cleaved into 10.0 kb and 2.3 kb in C1824A mutants. The extra 2.3 kb band indicative of C1824A mutation is highlighted by the arrows.



(a) Average bioassay areas, *P.f.* 10586 Δ*mmpB*



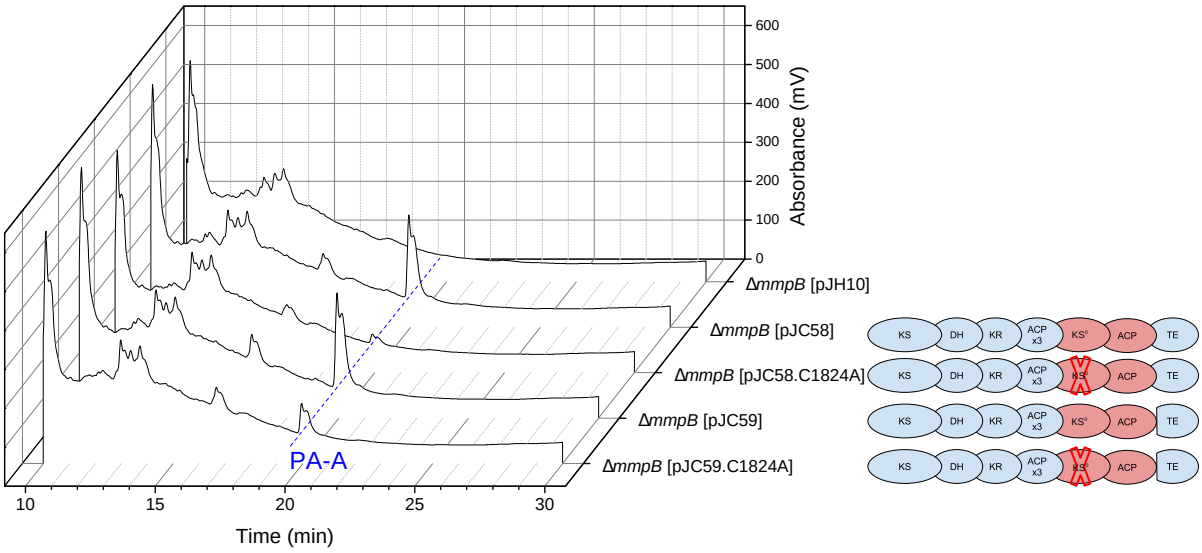
(b) Example bioassay plate

Figure 4.27: Bioassay: MmpB/TmpB hybrids with point mutation to KS^0.
Plate bioassay against *B. subtilis*, n=3, error bars are standard deviations.

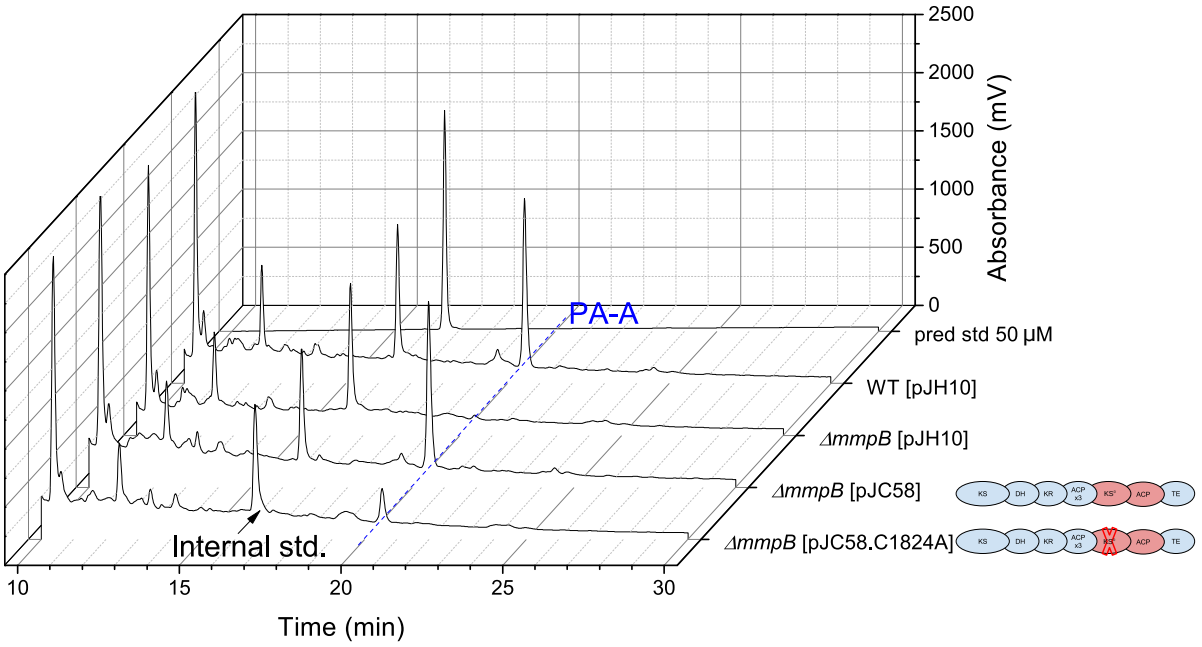
Δ*mmpB* to approximately 70% of the area of 10586 WT, which agrees with the data presented in the previous section. The KS^0 mutant variant of pJC59 complemented to a similar level, suggesting that the hybrid from pJC59 can generate PA-B without KS^0 activity. With the other hybrid from generation II, pJC58 (full length TE), a drop from 79% (% area of WT) to 57% was observed with the mutation. This difference was significant at 95% confidence level by unpaired two-tailed *t* test, $p=0.008<0.05$, n=3.

HPLC analysis revealed a more definitive phenotypic difference in the complementation of hybrids with the C1824A mutation (Figure 4.28). Complementation of 10586 Δ*mmpB* was again observed by both hybrid II expression plasmids. For pJC59 (truncated TE), point mutation of the KS^0 decreased PA-A peak area to 29% of the non-mutant, whereas for pJC58 (full length TE), C1824A decreased the peak area to 10%. This difference for hybrid II, pJC58, was significant by unpaired two-tailed *t* test of integrated peak areas, $p=0.0005<<0.05$, n=5. Excitingly, this suggests that the KS^0 enzymatic activity contributes to the function of hybrid II, and that intermediates may be actively passed through the extra module.

However, the mutation did not completely abolish mupirocin production, with



(a) Initial HPLC analysis



(b) Repeat HPLC analysis of pJC58 and pJC58.C1824A

Figure 4.28: HPLC: inactivation of the KS^0 domain of MmpB/TmpB hybrid generation II by point mutation decreases function.
(a) Representative chromatogram for comparison of pJC58 and pJC59, n=3.
(b) Repeat analysis of 10586 Δ*mmpB* [pJC58] and [pJC58.C1824A], n=2. **(b)** was performed independently from **(a)**, from the point of conjugation of plasmid DNA. This repeat analysis is included to demonstrate that the split peaks in **(a)** are due to the condition of the HPLC column and not a characteristic of the produced metabolites.

small PA-A peaks detected by HPLC and large zones of inhibition present on bioassays. This discrepency between bioassay zones and production detected by HPLC has been observed previously in other experiments performed on mupirocin biosynthesis, such as in the PA-C producer *mmpE*ΔOR (Gao et al., 2014) or in later generation III tests of pJC119 (Section 4.2.5.3). This could be due to the difference in conditions of each experiment, as HPLC was performed on supernatants of liquid SSM cultures incubated at 22 °C with shaking, whereas bioassays were performed on solid LB medium incubated at 20 °C then 37 °C on the second day. These assays measure different aspects of production: the HPLC method measures the final concentration of PA-A in the extra-celullar liquid culture, whereas bioassay measures inhibition of the growth of *B. subtilis*, which is likely influenced by production level, timing of production, diffusion rate, and compound antibiotic activity.

To summarise, these data provide evidence that the second module of TmpB in MmpB/TmpB hybrid generation II is enzymatically active. This represents an important step towards the goal of re-engineering the C8-hydroxy removal tailoring step to occur on the hybrid.

4.2.4 Conversion of PA-B to PA-A on MmpB/TmpB hybrids using thiomarinol enzymes

Now that a hybrid with a seemingly functional second domain had been constructed, the next goal was to express the thiomarinol tailoring enzymes with the hybrid, and determine whether PA-A can be generated on the second module of the hybrid. The primary candidate tailoring enzymes identified were TmlO, TmlP, TmlC and TmlF. All are homologs of their mupirocin namesakes, with 43%, 30%, 55%, 33% AA identity respectively. TmlY (hypothetical membrane protein), which lacks a mupirocin homolog, was also considered as a secondary, more speculative target [3].

[3] Design and construction of thiomarinol tailoring enzyme expression plasmids were performed by Yusra Alsammarraie (pJH10/*tmlOPCF*) and Amber Wilson (pJH10/*tmlOPCF* and pJH10/*tmlYOPCF*).

One option considered was cloning these *tml* genes into the expression vectors carrying *mmpB/tmpB* hybrids, such as pJC58, and expressing all genes at once *in trans*. However, cloning into an increasingly large pJH10 based plasmid (starting at 21.7 kb) may have proved a difficult molecular construction. Also, the simplest way of building such a vector would be to place all genes behind the *taq* promoter. However, this would prevent use of IPTG induction, due to the inhibitory effect of *mmpB* over expression.

An alternative option was to provide *tmlOPCF* on a second expression vector. The vector would need to be of a different compatibility group to pJH10, and be broad host range, ruling out the *E. coli* replicon plasmids. Co-expression from IncP vectors with the IncQ pJH10 has previously been attempted, but proved inconsistent (Shields, 2008).

Therefore, the most immediate way to express these proteins together was to insert the *mmpB/tmpB* hybrids into the *P. fluorescens* chromosome. This would allow for the expression of *tmlOPCF* from the reliable pJH10 vector.

4.2.4.1 Construction of suicide vectors for insertion of *mmpB/tmpB* hybrids into the *P. fluorescens* NCIMB10586 chromosome

The suicide mutagenesis strategy used requires homology arms for recombination flanking the desired mutational sequence. Fortunately, MmpB/TmpB hybrid II already has such a setup, with module 2 of *tmpB* flanked on both sides by *mmpB* sequence. A 3.6 kb PCR product was amplified with pJC58 as template, using primers F_pJC5x_XbaI and R_pJC5x_SalI. This was cloned *Xba*I/*Sal*I into mutagenesis vector pJC70, yielding pJC74.58 (Figure 4.29). The equivalent steps were performed for MmpB/TmpB hybrid II with truncated TE domain, using pJC59 as template and yielding pJC74.59.

The construction of the mutagenesis vector for MmpB/TmpB hybrid I required one further step, as there is no downstream homology arm already in place. A 3.6

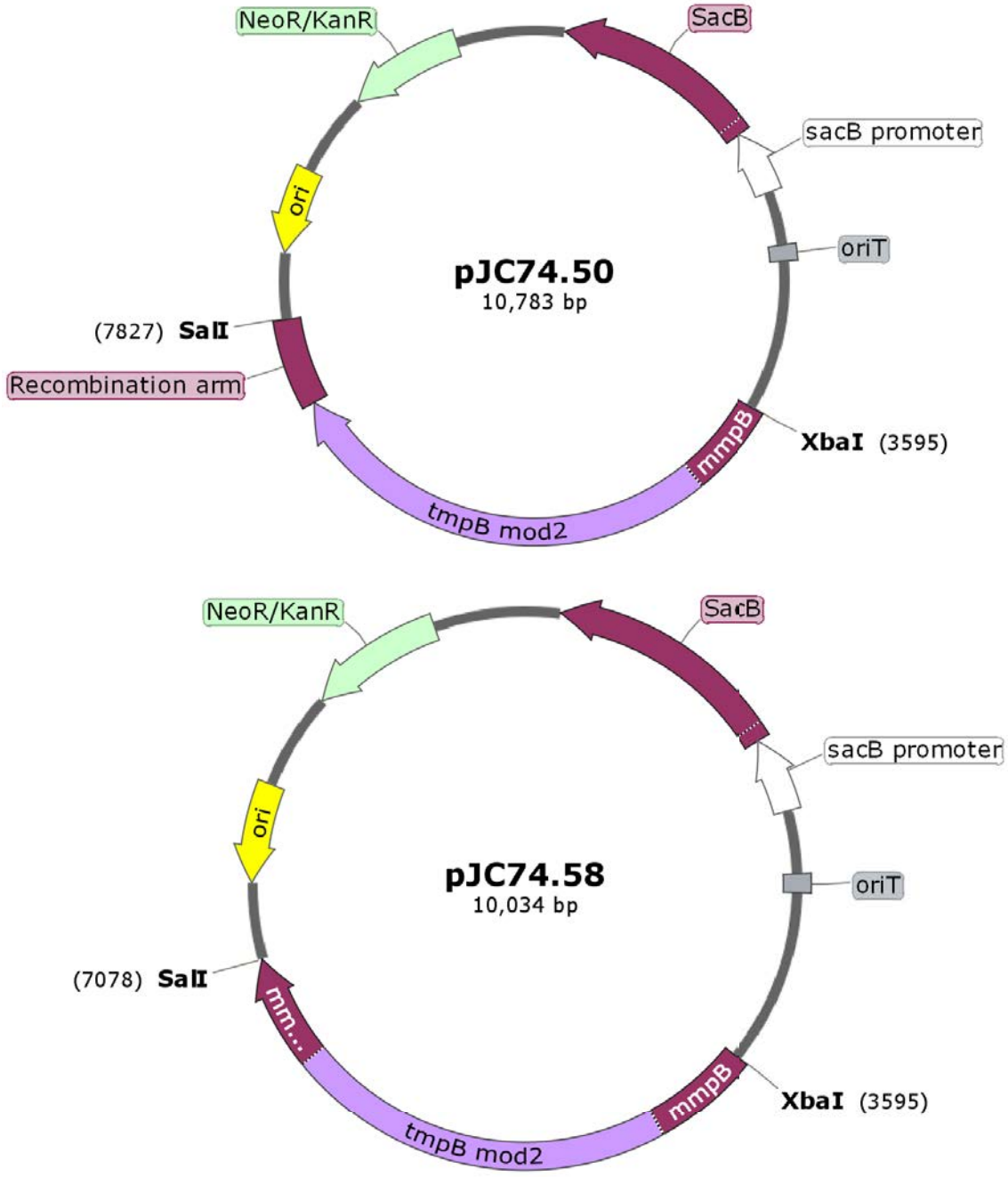


Figure 4.29: Maps of pJC74 suicide mutagenesis vectors.
Plasmid pJC74.50 was constructed for insertion of hybrid I into the 10586 chromosome, named for the equivalent hybrid expression plasmid pJC50.
Plasmids pJC74.58 and pJC74.59 were constructed for insertion of hybrid generation II into the chromosome, both full TE length (JC58) and truncated TE (JC59).

kb product was amplified with pJC50 as template using primers F_pJC5x_p604_XbaI and R_pJC50_overlap. This product has a 600 bp region of *mmpB*, followed by the second module and TE of *tmpB*, to the stop codon. To generate the downstream homology arm, the 600 bp region following *mmpB* on the 10586 chromosome was amplified using primers F_hybriddownarm_overlap and R_hybriddownarm_SalI. The two products were spliced by overlap extension PCR, and the resulting 4.2 kb product cloned *Xba*I/*Sal*I to pJC70, yielding pJC74.50 (Figure 4.29).

The resulting suicide vectors (pJC74.50, pJC74.58, pJC74.59) were screened by *Xba*I/*Sal*I digest to check for presence of insert. Correct DNA sequence was confirmed by sequencing with primers M13(-40), M13(-48), and F_tmpB_C1824A. Plasmid pJC74.50 was further sequenced with primers TmpBF3 and R_tmpB_S2298A to achieve full coverage.

4.2.4.2 Replacement of the *P. fluorescens* NCIMB10586 chromosomal *mmpB* with *mmpB/tmpB* hybrids

Mutagenesis plasmids pJC74.50, pJC74.58 and pJC74.59 were used to replace the 10586 chromosomal copy of *mmpB* with their respective hybrids, using the method described in Section 2.2.12. Mutant strains were PCR screened using primers F_pJC5x-_p604 and R_hybriddownarm_SalI (Figure 4.30).

Each of the three hybrid mutations were introduced to four *P. fluorescens* NCIMB-10586 strains: 10586 WT, 10586 Δ*mupU*, 10586 Δ*mupV*, and 10586 *mmpE*ΔOR. The *mupU* and *mupV* mutant strains were included as these are deficient in the native mupirocin PA-B to PA-A tailoring pathway. In the absence of MupU or MupV, restoration of PA-A production in these strains through a combination of MmpB/TmpB hybrids and thiomarinol enzymes would indicate C8-OH removal by the thiomarinol enzymes.

The oxidoreductase domain of MmpE is required for formation of the C-10,11 epoxide, and therefore the *mmpE*ΔOR produces PA-C as the primary metabolite (Gao

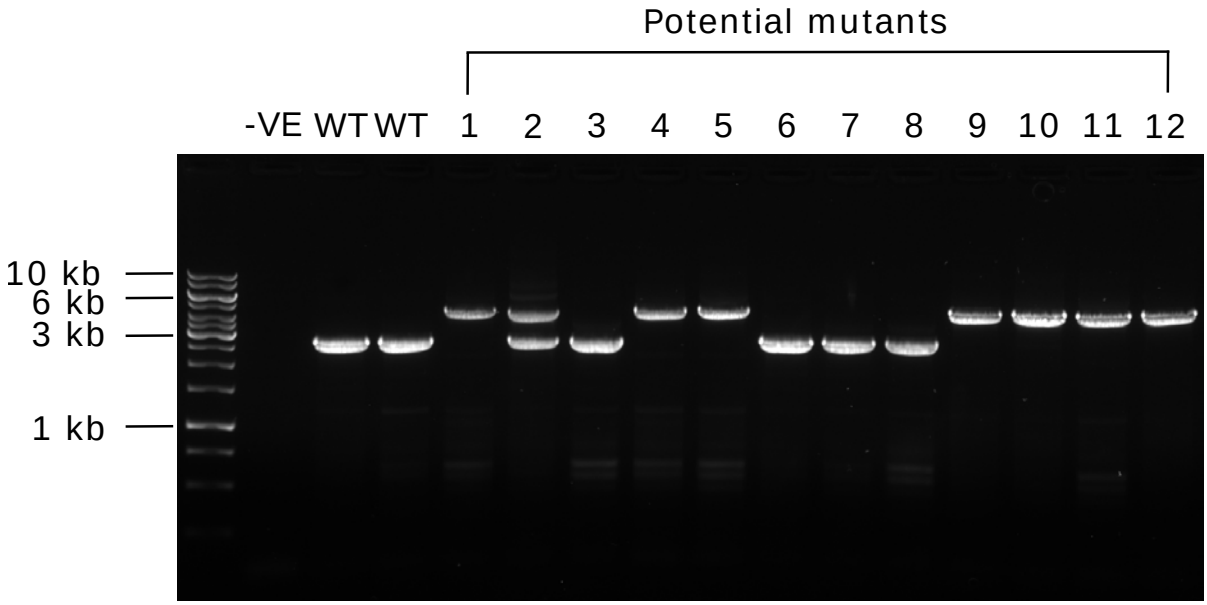


Figure 4.30: Example of PCR screening for MmpB/TmpB hybrid chromosomal mutations
The lower bands correspond to 2.6 kb, the size of the WT *mmpB* fragment amplified. The larger bands correspond to 4.2 kb, the size with *tmpB* module 2 hybrid. Sample number 2 has bands for both WT and mutant; this pattern is likely due to incorporation of the mutation elsewhere in the chromosome.

et al., 2014). Thiomarinol also lacks the epoxide, therefore it was conceivable the *tml* TE domain in hybrid I was showing specificity towards desepoxy intermediates. The hybrids were therefore introduced to 10586 *mmpE*ΔOR to test this hypothesis.

Mupirocin production of the 12 generated *mmpB*/*tmpB* mutant strains was tested by bioassay (Figure 4.31). In 10586 WT, similarly to the plasmid-based results presented so far in this chapter, PA-A was still produced after replacement of MmpB with either TE variant of MmpB/TmpB hybrid generation II. MmpB/TmpB hybrid I again proved non-functional.

No restoration of PA-A production was observed in the PA-B producing strains 10586 Δ*mupU* and 10586 Δ*mupV*. The PA-C producer, 10586 *mmpE*ΔOR was included to test whether the TmpB TE was demonstrating specificity towards des-epoxy polyketide intermediates. However, no clearing zone was observed for *mmpE*ΔOR with MmpB/TmpB hybrid I (pJC50).

As the hybrid I mutants were not functional, pJH10/*mmpB* was mobilised to these

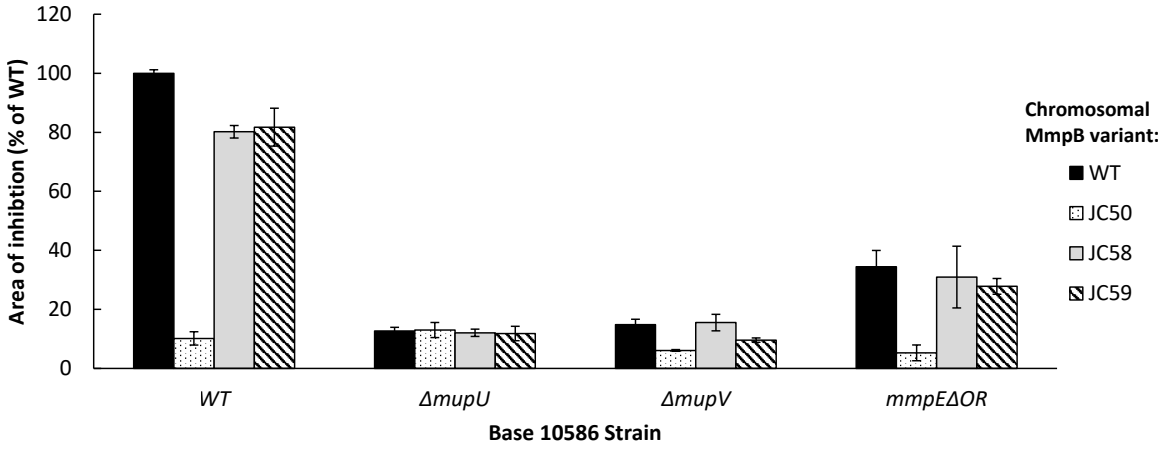


Figure 4.31: Bioassay zones of inhibition for MmpB/TmpB hybrid chromosomal replacements
Average inhibition zone area against *B. subtilis*, performed independently in duplicate, error bars represent standard deviation.

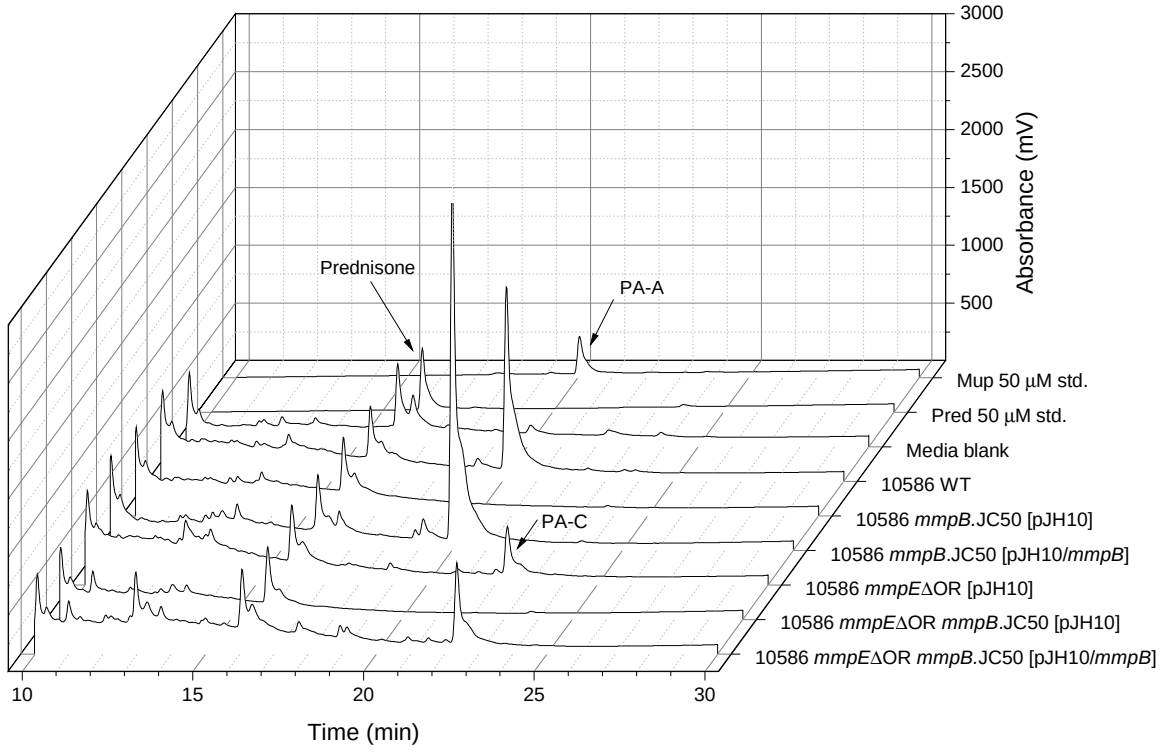


Figure 4.32: Complementation of MmpB/TmpB hybrid I chromosomal mutants by pJH10/*mmpB*.
Representative chromatogram shown, analyses performed independently in duplicate. Chromosomal mutant 10586 *mmpB*.JC50 (hybrid I) is fully deficient in PA-A production, and was complemented by pJH10/*mmpB*. Similarly, PA-C production was complemented in 10586 *mmpE*ΔOR *mmpB*.JC50 by pJH10/*mmpB*.

strains to check for complementation. HPLC analysis confirmed that all MmpB/TmpB hybrid I strains have a complementable phenotype (Figure 4.32), which suggests that the phenotype is caused by lack of hybrid I function rather than any polar effect.

4.2.4.3 Small scale conversion of PA-B to PA-A was achieved on MmpB/TmpB hybrids with plasmid encoded *tmlO*, *tmlP*, *tmlC* and *tmlF*

Now that a suite of 10586 MmpB/TmpB hybrid mutants had been generated, the candidate thiomarinol genes could be provided on expression plasmids. Two expression plasmids were constructed, pJH10/*tmlOPCF* (Figure 4.33) and pJH10/*tmlYOPCF* (Alsammarraie, Y. and Wilson, A., unpublished)[4].

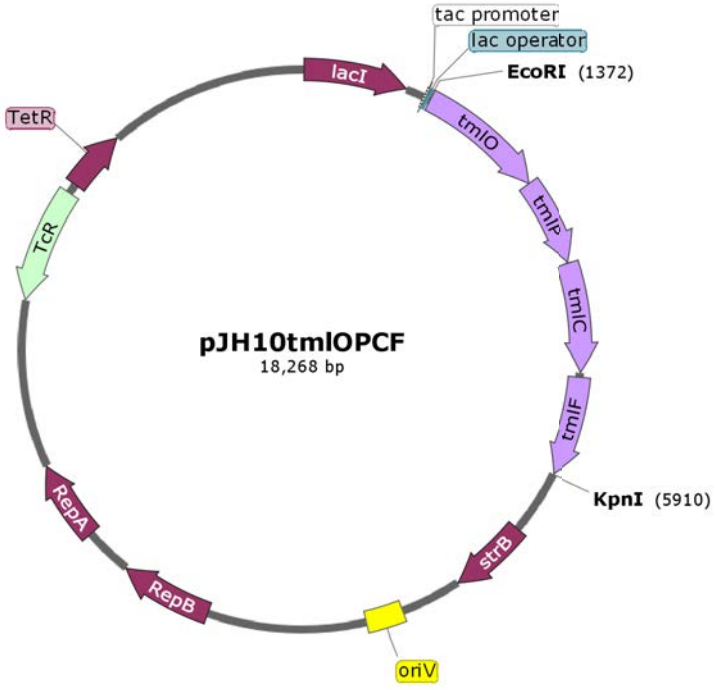


Figure 4.33: Map of pJH10/*tmlOPCF*.
Plasmid pJH10/*tmlYOPCF* has the same gene order, although with *tmlY* inserted upstream of *tmlO* [4].

The *tmlOPCF* and *tmlYOPCF* plasmids, together with empty vector pJH10, were mobilised to WT and mutants of 10586 carrying Δ*mupU*, and Δ*mupV* plus their derivatives with *mmpB*.JC50, JC58 and JC59. Mupirocin production of each combination of

[4] Design and construction of thiomarinol tailoring enzyme expression plasmids were performed by Yusra Alsammarraie (pJH10/*tmlOPCF*) and Amber Wilson (pJH10/*tmlOPCF* and pJH10/*tmlYOPCF*).

strain and plasmid was tested by plate bioassay, both with 0.5 mM and without IPTG induction (Figure 4.34).

As first reported in Cooper et al. (2005), deletion of either *mupU* or *mupV* causes a switch to production of the inactive PA-B, and a corresponding deficiency in clearing zone was observed. None of the *mmpB*.JC50 strains (carrying MmpB/TmpB hybrid I) displayed any antibiotic activity, as was the case in every test reported in this chapter where hybrid I was the sole *mmpB* version available to the bacteria.

Gratifyingly, restoration of 35% $\pm$6.6 (n=4) of the WT zone was observed in 10586 Δ*mupU mmpB*.JC58 (MmpB/TmpB hybrid II) when pJH10/*tmlOPCF* was present. Without *mupU* the native mupirocin pathway is only able to produce the inactive metabolite PA-B. This partial complementation of antibiotic activity suggests restoration of PA-A production mediated by some or all of the thiomarinol enzymes. Plasmid pJH10/*tmlYOPCF* complemented to a similar level. As complementation worked as well without *tmlY*, it seems unlikely that *tmlY* is involved or required for this part of the pathway. IPTG Induction of *tml* gene expression from the plasmids was observed completely inhibitory to antibiotic production (Figure 4.34b).

To further investigate these promising results, HPLC analysis of SSM culture supernatants was performed with 10586 Δ*mupU*, 10586 Δ*mupU mmpB*.JC58, carrying plasmids pJH10 or pJH10/*tmlOPCF* (Figure 4.35). As was originally reported in Cooper et al. (2005), deletion of *mupU* switches production to the inactive PA-B. When the *mupU* deletion is combined with *mmpB*.JC58, MmpB/TmpB hybrid II can substitute for normal MmpB function, and only PA-B is produced. When pJH10/*tmlOPCF* is provided to 10586 Δ*mupU mmpB*.JC58, some PA-A production is restored. By integration of peak areas, this was measured as 12% $\pm$3.0 (n-4) PA-A as a proportion of total pseudomonic acids. A lower level of PA-A production was observed in 10586 Δ*mupU* [pJH10/*tmlOPCF*], at 5.9% $\pm$1.2 PA-A as a proportion of total pseudomonic acid.

This restoration of PA-A production with TmlO, TmlP, TmlC, TmlF and MmpB/TmpB

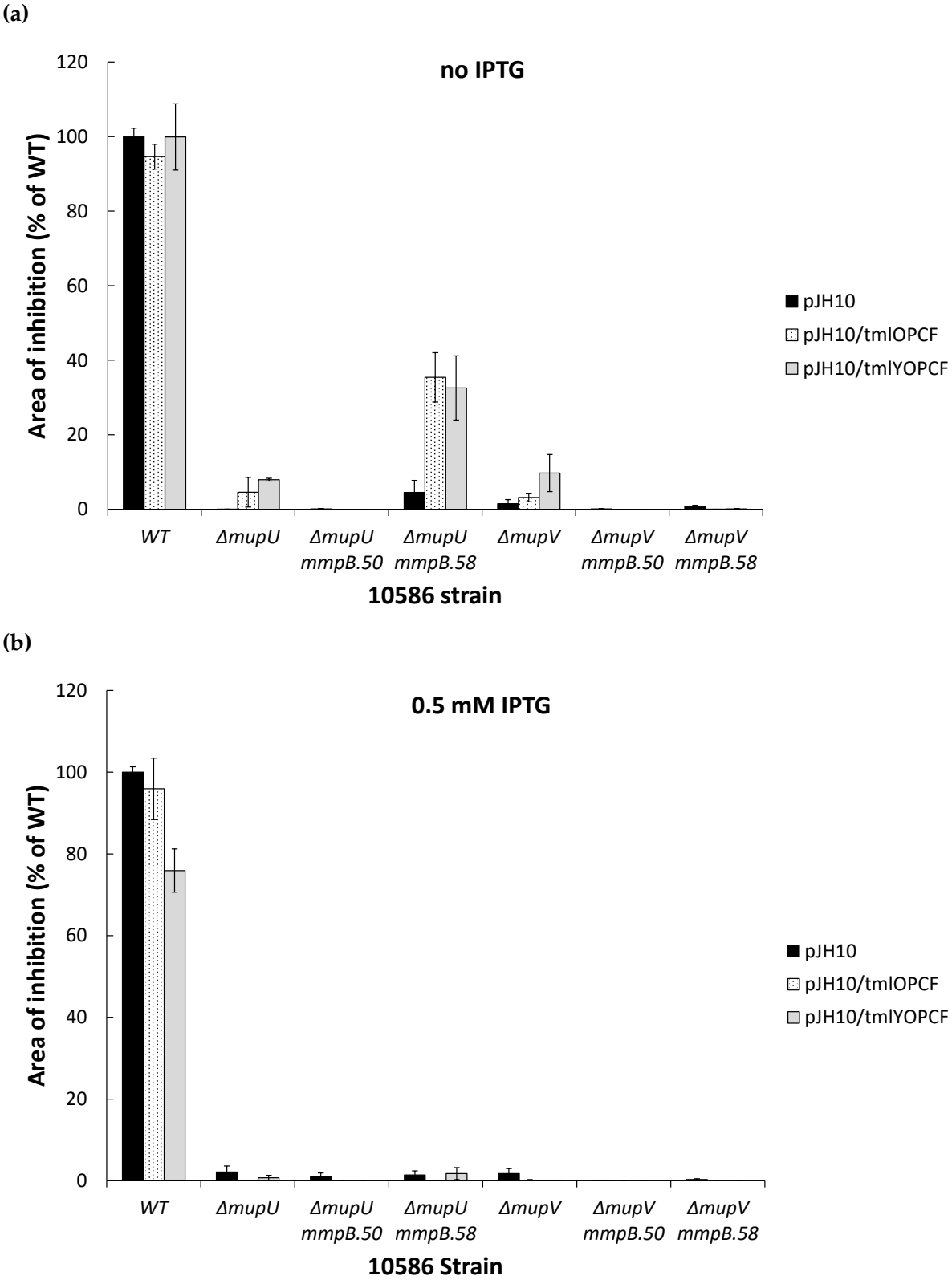


Figure 4.34: Bioassay: pJH10/*tmlOPCF* can partially complement mupirocin production in 10586 Δ*mupU* with MmpB/TmpB hybrid II, JC58.
Areas of inhibition of *B. subtilis*, analyses performed independently in quadruplicate; error bars are standard deviation.

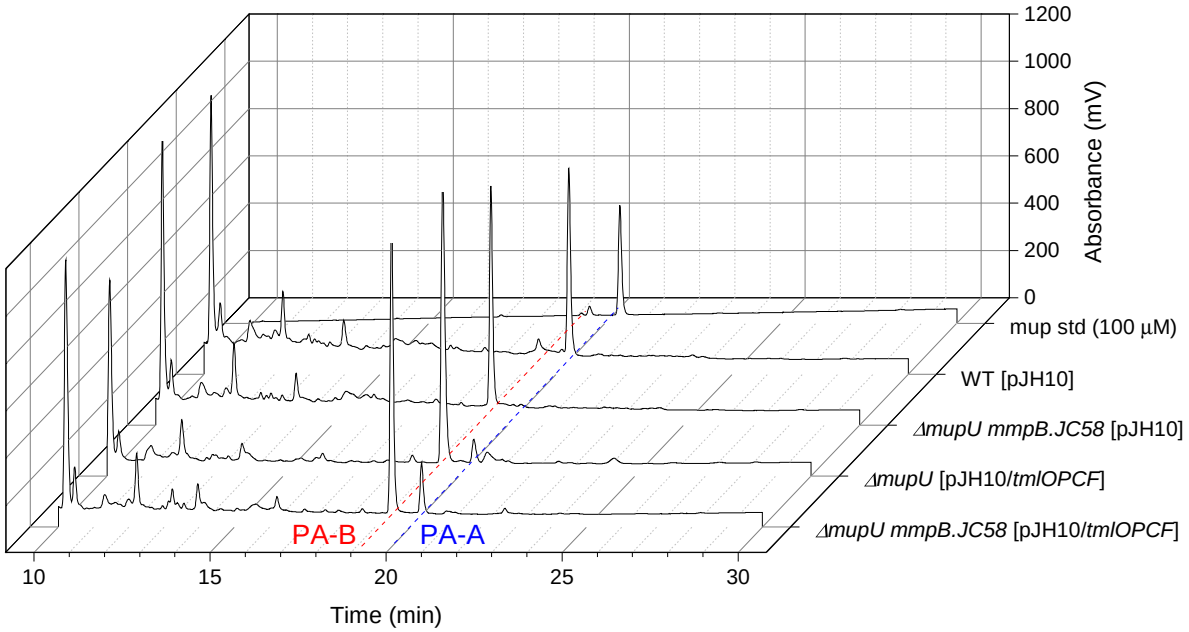


Figure 4.35: HPLC: partial restoration of PA-A production in 10586 Δ*mupU* with MmpB/TmpB hybrid I and pJH10/*tmlOPCF*.
Representative chromatogram shown, no IPTG induction was used, analyses were performed independently in quadruplicate, with the exception of WT [pJH10] that was performed in duplicate.

hybrid II, in the absence of MupU, agrees with the bioassay data (Figure 4.34a). This could suggest successful re-engineering of the pathway to make use of the thiomarinol enzymes, and convert PA-B to PA-A on the hybrid version of MmpB. However, the HPLC data indicated the hybrid is not obligately required for TmlOPCF mediated restoration of PA-A production in the absence of MupU, as some PA-A was observed in 10586 Δ*mupU* [pJH10/*tmlOPCF*]. TmlOPCF were obligately required; no PA-A was observed in the Δ*mupU* mutant strains without pJH10/*tmlOPCF*.

Regardless of whether C8-OH removal is occurring on MmpB/TmpB hybrid II or not, the PA-A yields achieved using the thiomarinol enzymes were much lower than usual. An alternative strategy was conceived where the native MmpB enzymes would be re-engineered to work on the MmpB/TmpB hybrids (Section 4.2.5).

4.2.5 MmpB/TmpB hybrid generation III: splicing of MacpE behind the thiomarinol KS^0

MmpB/TmpB hybrid generation III consisted of three constructions, each of which have MacpE fused behind the TmpB KS^0 in place of the TmpB ACP of generations I and II. The structures of each of the hybrids is outlined in Figure 4.36.

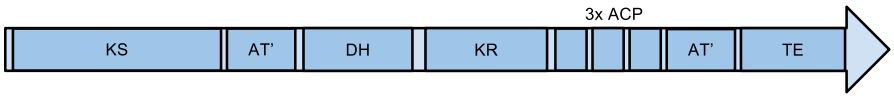
The first construction of hybrid generation III, JC118, was designed with the TmpB TE domain following MacpE, in a similar fashion to hybrid I. As discussed in Section 4.2.2.2, the TmpB TE domain was not functional in hybrid I. However, this may be due to the TmpB TE domain demonstrating specificity towards polyketide intermediates that had lost the C8-OH, which would be logical given that this step is proposed to occur on TmpB in thiomarinol biosynthesis. If mupirocin enzyme-catalysed C8-OH removal were to occur on the fused MacpE of JC118, then it is possible that the TmpB could act as a gatekeeper for PA-A release.

The second construction of generation III, JC119, was designed without a TE domain. As covered in Chapter 3, the release of late biosynthesis stage pseudomonic acids is proposed to be catalysed by MupV. If C8-OH removal were to occur on the fused MacpE of JC119, then MupV could potentially catalyse release of PA-A.

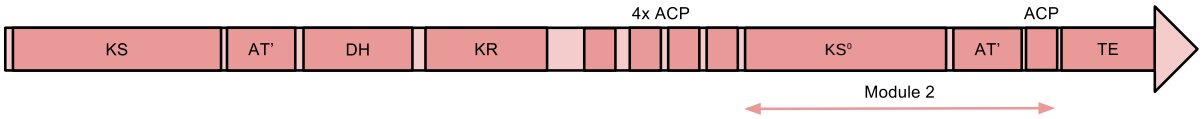
Thirdly, JC120 was designed with the MmpB TE domain following MacpE, in a similar fashion to hybrid generation II. This was designed as hybrid generation II had proven functional.

An important decision in the design of hybrid generation III was the positioning of the junction of MacpE and MmpB/TmpB. To aid this, an amino acid alignment was generated of the TmpB module 2 ACP to be replaced, MacpE from *P. fluorescens* NCIMB10586, and MacpE from *P. pyschrotolerans* NS383 (Figure 4.38). These aligned without gaps, supported by the Jpred predicted secondary structure. The junction was designed with MacpE beginning from its third amino acid (Ile) at position 2264 in the hybrids, such that the MacpE PP-carrying serine was positioned in the same place as the TmpB ACP serine previously (position 2298).

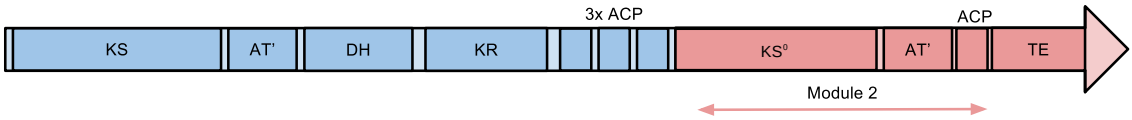
(a) MmpB (2077 amino acids)



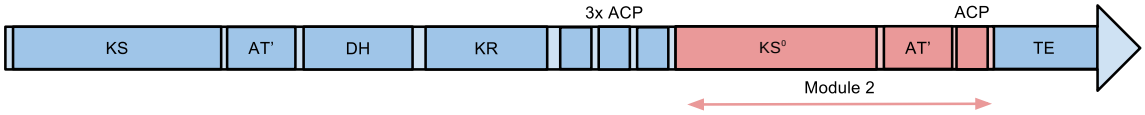
(b) TmpB (2751 amino acids)



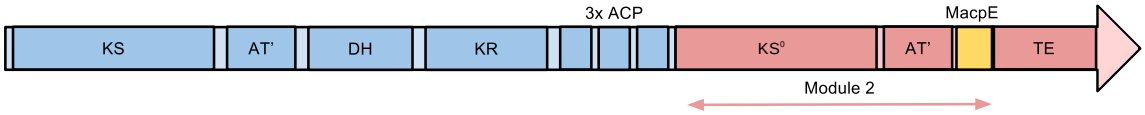
(c) MmpB/TmpB hybrid generation I: JC50 (2620 amino acids)



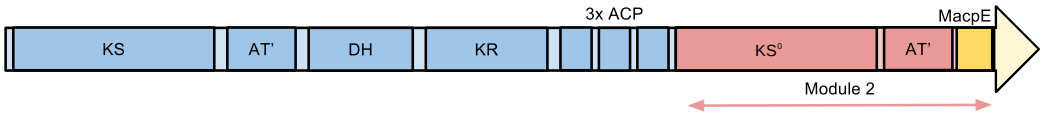
(d) MmpB/TmpB hybrid generation II: JC58, JC59 (2658, 2619 amino acids)



(e) MmpB/TmpB hybrid generation III: JC118 (2620 amino acids)



(f) MmpB/TmpB hybrid generation III: JC119 (2341 amino acids)



(g) MmpB/TmpB hybrid generation III: JC120 (2659 amino acids)

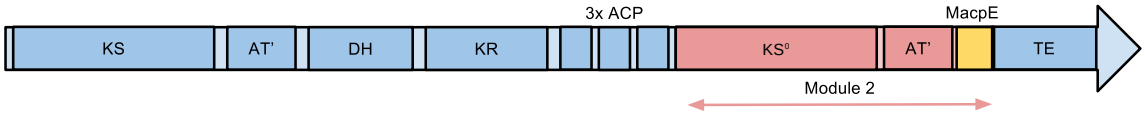


Figure 4.36: Domain maps of MmpB, TmpB and MmpB/TmpB hybrid generations I, II and III.
Blue represents MmpB (mupirocin) domains, and red TmpB (thiomarinol) domains. The natively freestanding mupirocin ACP, MacpE, is shown in yellow.

Figure 4.37: Amino acid sequence of the junction points between the TmpB derived sequence in the hybrid and MacpE in JC120.

The top line shows an approximate annotation of where each domain is positioned, the second line shows the amino acid number at both the start and end of the sequence displayed, respective to each protein. The third line shows the amino acid sequence, and the fourth line is the predicted secondary structure from Jpred (H is α-helix and E is β-strand).

Figure 4.38: Amino acid alignment of MacpE with the TmpB module 2 ACP. The first line is the ACP from TmpB module 2, which is the fifth ACP of TmpB and the fourth ACP in the MmpB/TmpB hybrids. Lines two and three are MacpE from *P. fluorescens* NCIMB10586 and *P. pyschrotolerans* NS383. Jpred secondary structure predictions are included underneath, E is extended (β-strand), H is α-helix.

4.2.5.1 Molecular construction of MmpB/TmpB hybrid generation III expression plasmids: pJC118, pJC119 and pJC120

To construct the generation III hybrids, DNA fragments were generated by overlap PCR spanning from the *Kpn*I site within ACP5 to a *Sac*I site following the stop codon[5]. Each of these fragments were then cloned *Kpn*I/*Sac*I to pJC58 (although any hybrid expression plasmid or pJH10/*mmpB* would have been suitable), to give full MmpB/TmpB hybrid generation III expression plasmids. This molecular construction approach is summarised in Figure 4.39 for the version without a TE domain, and Figure 4.40 for the versions with TE domains.

A 2599 bp fragment was PCR amplified from pJC58 using primers P16_F_2.6kb and P16_R_2.6kb, which spanned from the *Kpn*I site within ACP5 to the *macpE* junction point, and was used in the construction of all three generation III variants. To construct the variant without a TE domain, a 260 bp *macpE* fragment was amplified from the 10586 chromosome using primers P16_F_macpE and P16_R_macpE_stop_SacI. This included the native *macpE* stop codon, which would become the stop codon of this hybrid version. A 2837 bp fragments was generated by overlap PCR of the 2599 and 260 bp fragments, using primers P16_F_2.6kb and P16_R_macpE_stop_SacI. This was cloned *Kpn*I/*Sac*I to pJC58, yielding pJC119 (Figure 4.39).

To create the generation III variant with the *mmpB* TE domain (JC120), a 255 bp *macpE* containing fragment was amplified from the 10586 chromosome using primers

[5]Molecular construction and assay of the hybrid generation III plasmids was performed by both this author and Amber Wilson.

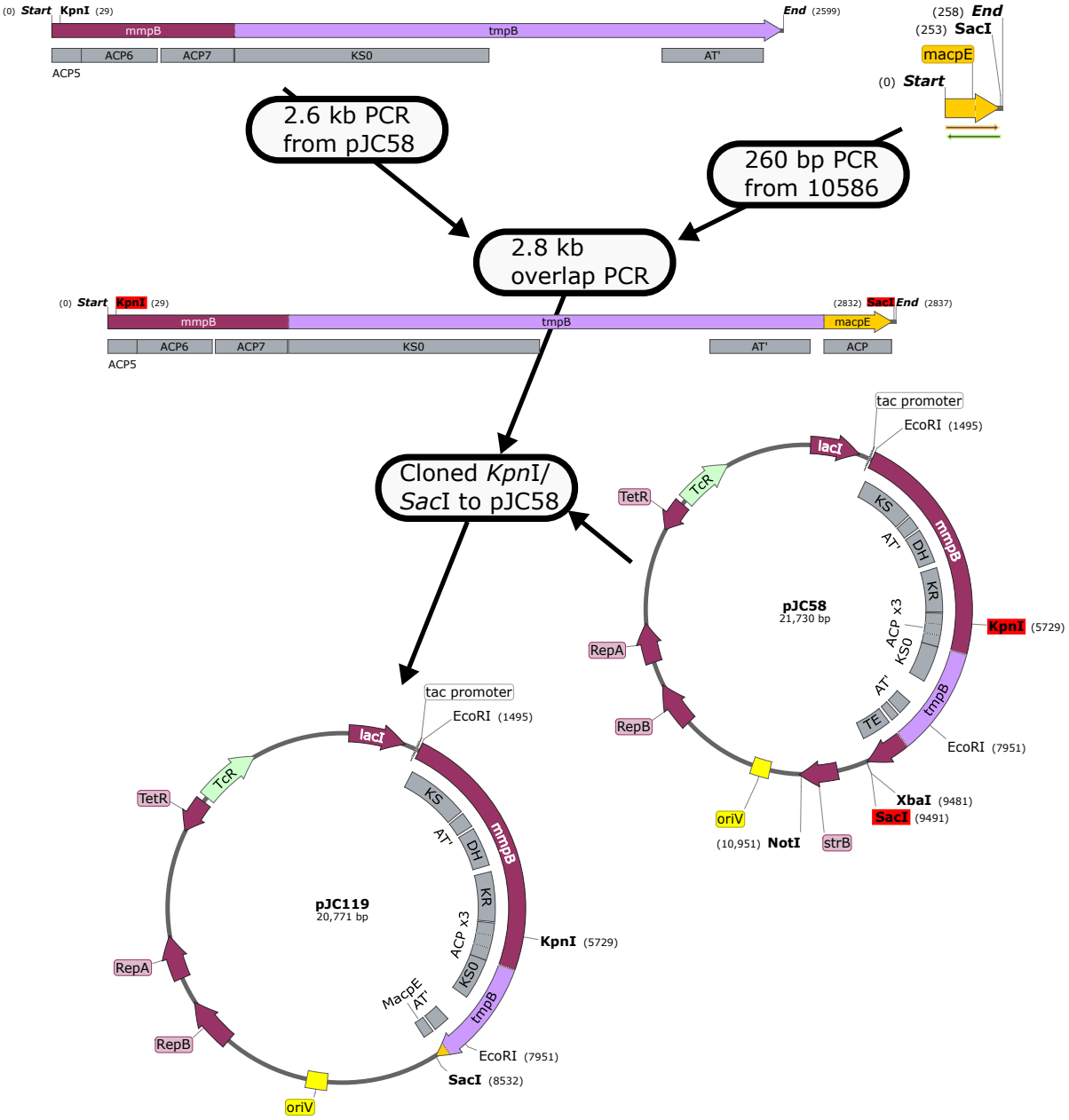


Figure 4.39: MmpB/TmpB hybrid generation III: construction overview of pJC118, without a TE domain.

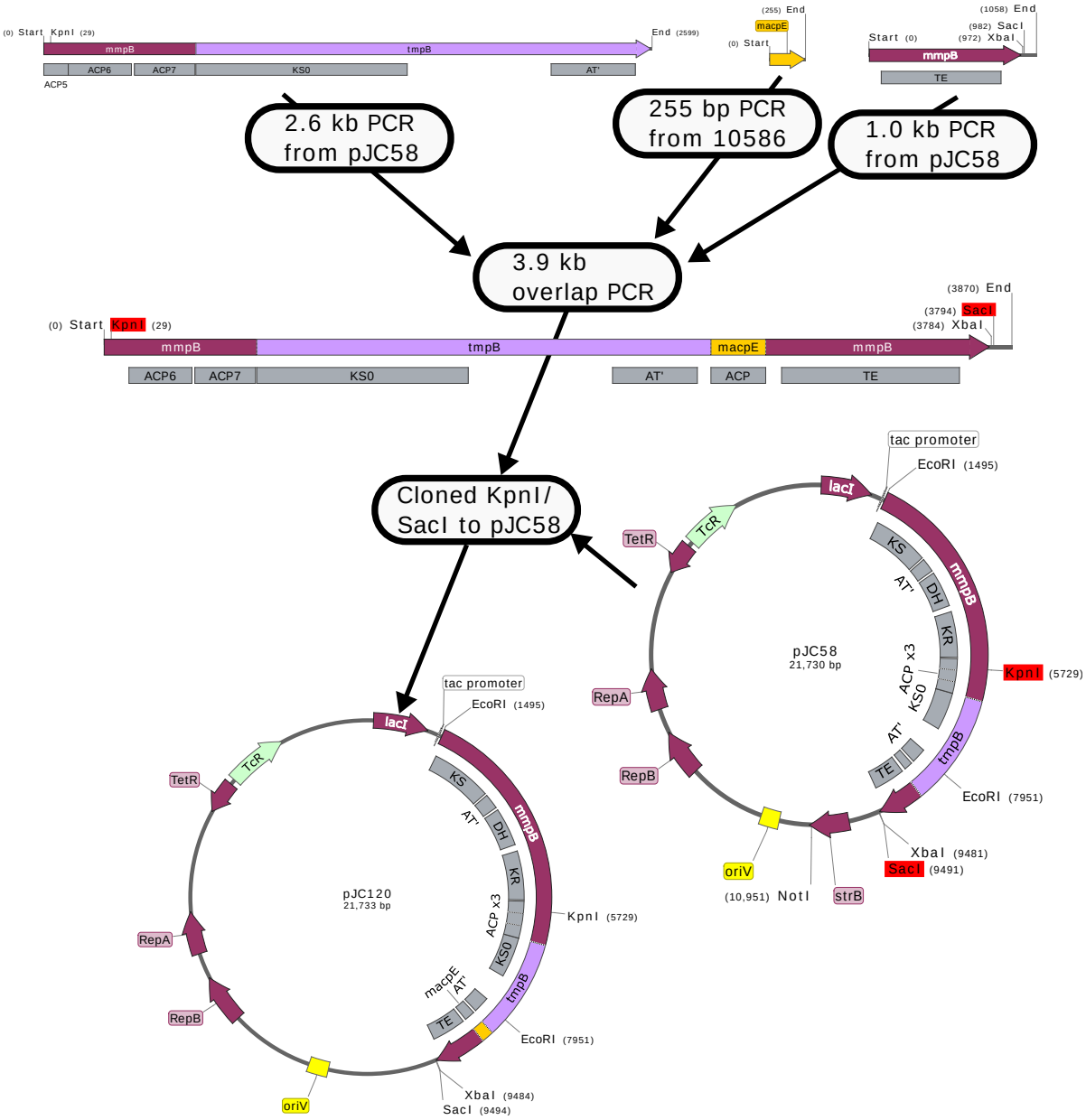


Figure 4.40: MmpB/TmpB hybrid generation III: construction overview of pJC120, with the MmpB TE domain.
The TmpB TE domain version, pJC119, was generated in a similar fashion, but with the TE domain amplified from pJC50 (hybrid I) instead of pJC58 (hybrid II).

P16_F_macpE and P16_R_macpE_TEoverlap. This omitted the *macpE* stop codon so that translation could continue into the TE domain of the hybrid. A 1058 bp fragment was amplified from pJC58 using primers P16_F_TE_revised and P16_R_TE_SacI, designed to carry the *mmpB* TE domain through to the hybrid stop codon. A 3870 bp fragment was generated by overlap PCR of the 2599, 255, and 1058 bp fragments, using primers P16_F_2.6kb and P16_R_TE_SacI. This was cloned *Kpn*I/*Sac*I to pJC58, yielding pJC120 (Figure 4.40).

To generate the *tmpB* TE variant, a 931 bp TE-containing fragment was amplified from pJC50 (hybrid I), using primers P16_F_TE_revised and P16_R_TE_SacI. In a similar fashion to the *mmpB* TE variant, a 3742 bp fragment was generated through overlap PCR of the 2599, 255 and 931 bp fragments. This was cloned *Kpn*I/*Sac*I to pJC58, yielding pJC118.

Replacement of the *tmpB* module 2 ACP with *macpE* was confirmed by *Bam*HI digest of the plasmids, since there is a *Bam*HI site in the *tmpB* ACP not present in *macpE*. The correct sequence from ACP5 to the end of the hybrid was confirmed in pJC118, pJC119 and pJC120 by sequencing with primers pJH10_seq_R, P16_F_macpE and R_pJC5x_p604_SalI.

4.2.5.2 Generation III hybrid expression plasmids pJC119 and pJC120 are able to produce PA-B

Generation III MmpB/TmpB hybrid expression plasmids pJC118, pJC119 and pJC120, and control plasmids pJH10 and pJC58, were mobilised to 10586 Δ*mmpB* and 10586 Δ*mmpB* Δ*macpE* Δ*mupU* to test function. Pleasingly, plate bioassay of 10586 Δ*mmpB* with pJC120 indicated complementation to 124% ±13.9 (n=3) of the WT zone of inhibition (Figure 4.41). This indicated that pJC120 is able to restore MmpB function, that is production of 9-HN and therefore PA-B, which can then be processed by the native enzymes to PA-A. However, no significant complementation of zones of inhibition was observed for plasmids pJC118 and pJC119. This is reminiscent of the

generation I and II results, where the MmpB TE domain was required for function.

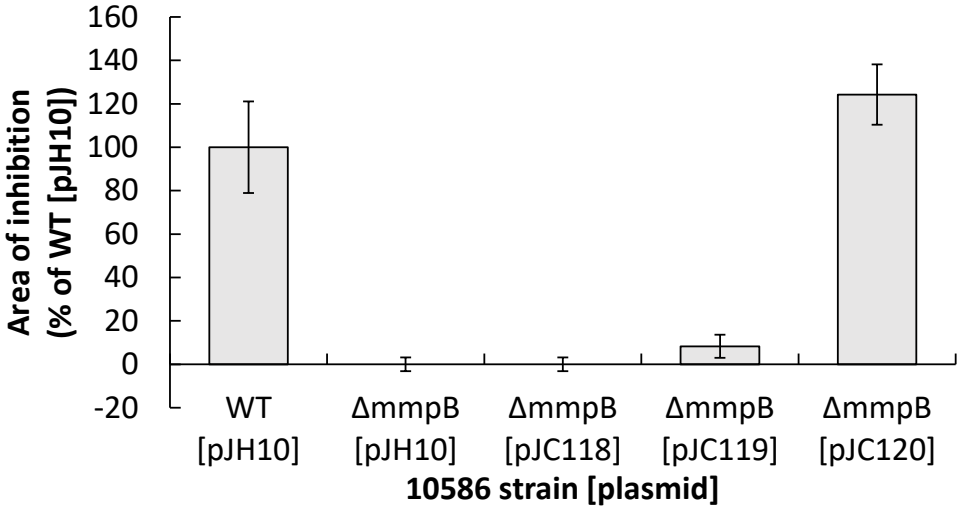


Figure 4.41: Bioassay: MmpB/TmpB generation III hybrid pJC120 can complement 10586 Δ*mmpB*.
Analyses performed independently in triplicate. Bars represent average areas. Error bars are standard deviation.

HPLC of SSM culture supernatants confirmed that 10586 Δ*mmpB* with pJC120 produces significant PA-A, 272% ±7 (n=3) of the integrated peak area of WT [pJH10] (Figures 4.44 and 4.42). As with the bioassay, no complementation was observed by the TmpB TE version, pJC118. Unlike the bioassay, pJC119 without a TE domain was observed to complement PA-A production of 10586 Δ*mmpB*, to 65% ±8 (n=3) WT [pJH10] peak area.

The generation III hybrids were also tested in triple mutant 10586 Δ*mmpB* Δ*macpE* Δ*mupU*, the rationale being to ensure that the native pathway of MupU-catalysed loading onto free-standing MacpE cannot proceed. No bioassay zones of inhibition were observed in this triple mutant carrying any of the plasmids tested: pJH10 (negative control), pJC58 (hybrid II), pJC118, pJC119 or pJC120 (not shown). HPLC analysis revealed that pJC120 could restore PA-B production in the triple mutant, and pJC119 to a lesser extent (Figure 4.43). This is logical as the hybrids can complement the *mmpB* knockout, which leaves a deficiency in *macpE* and *mupU*. 10586 Δ*macpE* and 10586 Δ*mupU* both primarily produce PA-B (Cooper et al., 2005). Disappointingly,

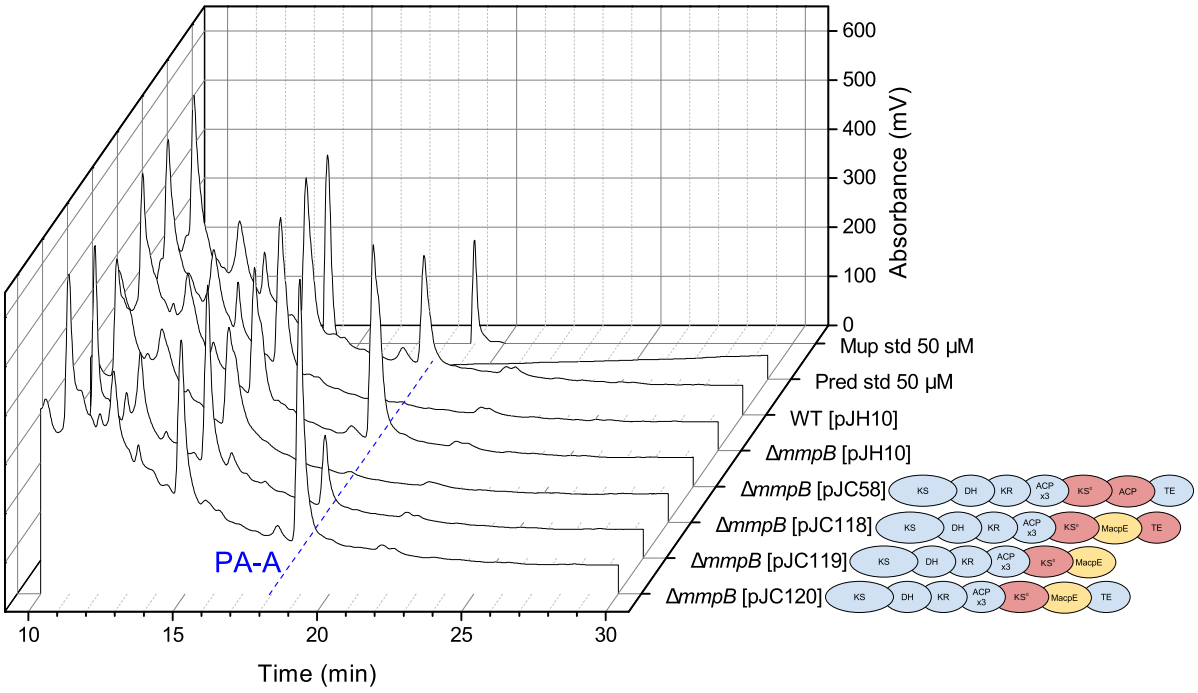


Figure 4.42: HPLC: MmpB/TmpB generation III hybrids pJC119 and pJC120 complement PA-A production in 10586 Δ*mmpB*.
Representative chromatograms are shown, analyses were performed independently in triplicate. In the hybrid domain overview reminders on the right, blue represents MmpB domains, red TmpB and yellow MacpE.

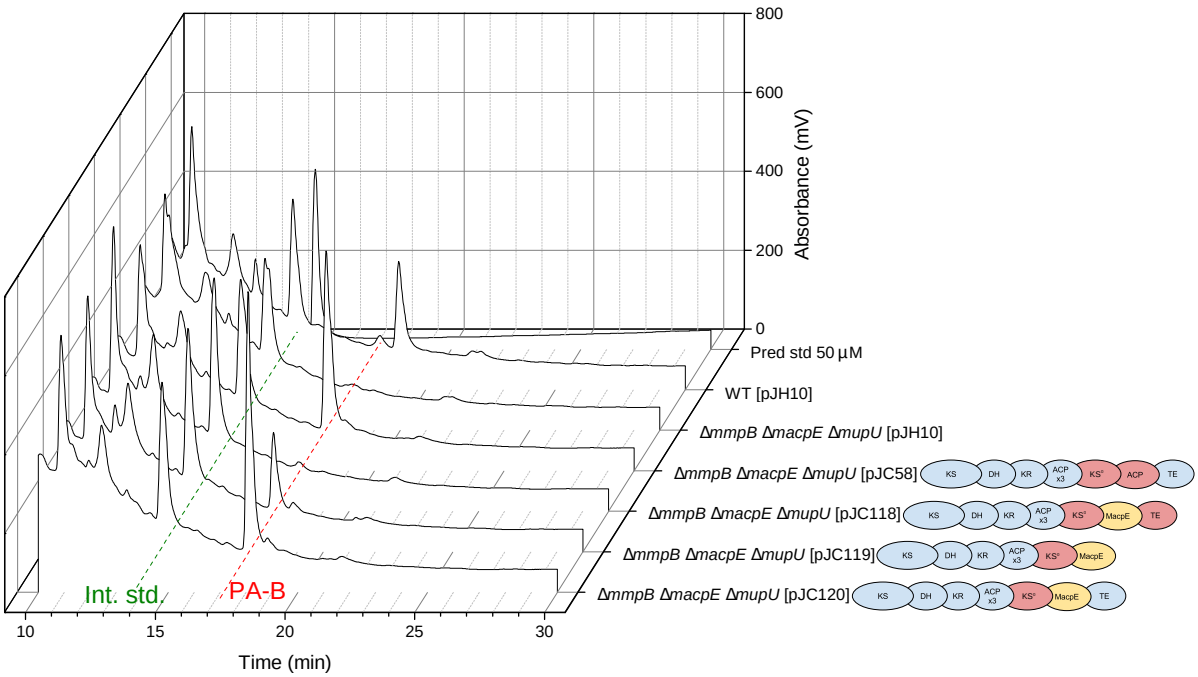


Figure 4.43: HPLC: MmpB/TmpB hybrid generation III cannot restore PA-A production in the absence of MupU and MacpE.
Representative chromatograms are shown, analyses were performed independently in triplicate.

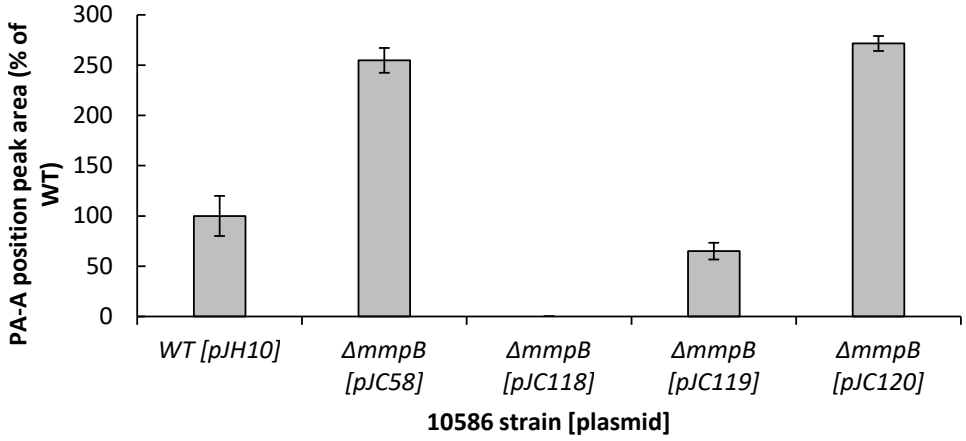


Figure 4.44: HPLC: Integrated PA-A peak areas of 10586 Δ*mmpB* complementation by generation III hybrid expression plasmids.
Example chromatograms shown in Figure 4.42, analyses performed independently in triplicate, error bars are standard deviation.

this suggested that conversion to PA-A does not occur on MmpB/TmpB generation III hybrids in the absence of MupU.

4.2.5.3 Conversion to PA-A achieved on hybrids JC119 and JC120 in the presence of MupU

MupU is thought to be responsible for loading of PA-B onto MacpE, therefore the equivalent function in the hybrids should be performed by the KS^0 domain. However, if a complex of tailoring enzymes (such as MupU, MupV, MupO and MupP) were to form on MacpE, it is conceivable that the absence of MupU could disrupt the formation of this complex. This would prevent conversion of PA-B to PA-A on the generation III hybrids. It was therefore desirable to test the hybrids on 10586 Δ*mmpB* Δ*macpE*, but this strain had not yet been generated.

To produce this double mutant, mutagenesis plasmid pAKE604/Δ*mmpB* (Hothersall, J., unpublished) was used to make an in-frame deletion of *mmpB* in 10586 Δ*macpE* (Cooper et al., 2005). Mutagenesis was performed as described in Section 2.2.12, and *mmpB* deletion strains were identified by PCR with primers MmpB_1F_Jo and R_hybriddownarm_SalI (Figure 4.45). This amplified 1280 bp in *mmpB* dele-

tions, whereas no band was amplified in WT *mmpB* strains as the PCR was not optimised for the expected 7.2 kb. Correct in-frame deletion of *mmpB* and generation of 10586 Δ*mmpB* Δ*macpE* was confirmed by sequencing of the PCR product with primer MmpB_1F_Jo.

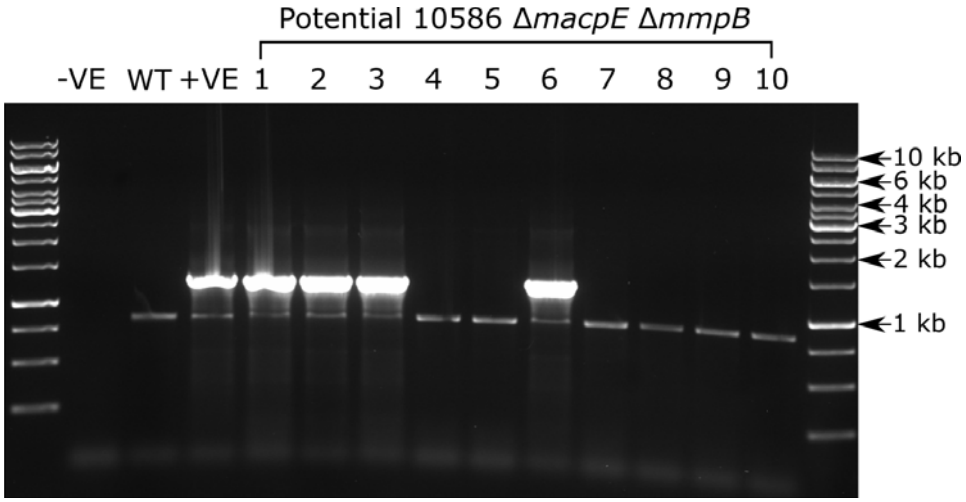


Figure 4.45: PCR screening for the *mmpB* deletion in the generation of 10586 Δ*mmpB* Δ*macpE*.
PCR of *mmpB* using primers MmpB_1F_Jo and R_hybriddownarm_SalI, expected to amplify 1280 bp in the in-frame deletion of *mmpB* and 7.2 kb in WT. Each lane differs by template DNA used, the first is a no template negative control, the second has 10586 WT, and the third has 10586 Δ*mmpB* (known mutant). The remainder are the potential double mutants being tested: isolates 1, 2, 3, and 6 have the 1.3 kb band indicative of *mmpB* deletion.

Generation III hybrid plasmids pJC118, pJC119 and pJC120, alongside pJH10 and pJC58 controls, were mobilised to the newly generated 10586 Δ*mmpB* Δ*macpE*. Mupirocin production was tested by bioassay (Figure 4.46), and HPLC of SSM culture supernatants (Figure 4.48). As logically expected, 10586 Δ*mmpB* Δ*macpE* with empty vector pJH10 did not generate a clearing zone on plate bioassays, or any pseudomonic acids as detected by HPLC. When generation II MmpB/TmpB hybrid pJC58 was introduced to 10586 Δ*mmpB* Δ*macpE*, complementation of PA-B production was observed by HPLC, with no restoration of clearing zone on bioassay. This can be rationalised by complementation of the *mmpB* deletion, restoring production of the inactive PA-B, but without MacpE no conversion to PA-A can occur.

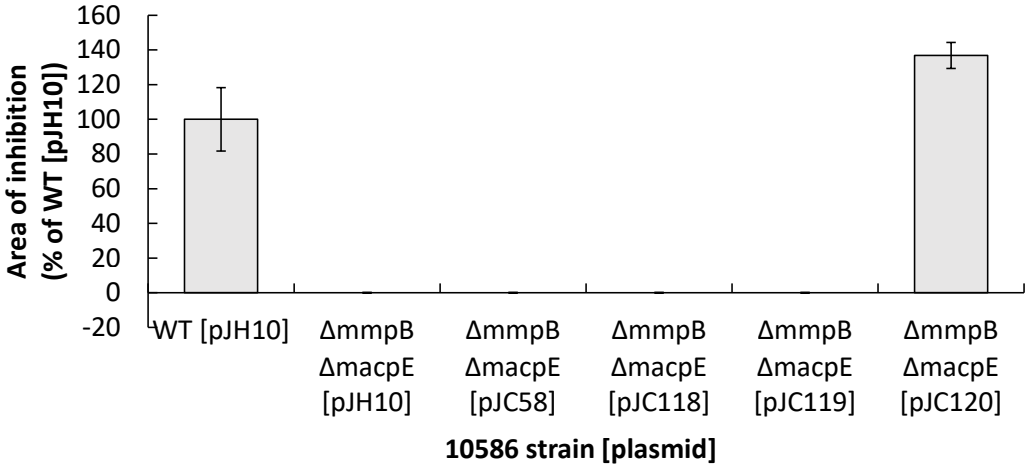


Figure 4.46: Bioassay: MmpB/TmpB generation III hybrid pJC120 can complement 10586 Δ*mmpB* Δ*macpE*.
Analyses performed independently in duplicate, bars represent average areas, error bars are standard deviation.

Gratifyingly, generation III hybrid pJC120 was observed to complement active antibiotic production of 10586 Δ*mmpB* Δ*macpE* to 156% ±12 (n=2) of the WT [pJH10] clearing zone. A corresponding restoration of PA-A production was observed by HPLC, with 10586 Δ*mmpB* Δ*macpE* [pJC120] producing 124% ±10 (n=3) PA-A peak area compared with WT [pJH10] (Figure 4.47). PA-B production of a similar order to PA-A was observed in 10586 Δ*mmpB* Δ*macpE* [pJC120]. Encouragingly, these results suggest that at least some of the tailoring to PA-A is occurring on hybrid pJC120, as the only copy of MacpE, which is mandatory for C8-OH removal, is located within the hybrid.

As has been the case in every hybrid with the *tml* TE domain reported on in this chapter, no pseudomonic acid production was observe with hybrid pJC118. The generation III hybrid without TE domain, pJC119, was not observed to complement the clearing zone of 10586 Δ*mmpB* Δ*macpE* by plate bioassay (Figure 4.46). However, this is expected, as it previously also failed to complement the clearing zone of 10586 Δ*mmpB* (Figure 4.41). Suprisingly, 10586 Δ*mmpB* Δ*macpE* with pJC119 was observed by HPLC to produce 40% ±1.8 (n=3) PA-A peak area of WT [pJH10]. Unlike pJC120, no PA-B production was detected with pJC119.

To summarise, these results represent the accomplishment of a key milestone towards the MmpB directed *in vivo* re-engineering of mupirocin biosynthesis. The C8-OH removal tailoring steps that would natively occur on free-standing MacpE have been re-engineerred to occur on MmpB derivatives, taking advantage of the thiomarinol KS^0 domain.

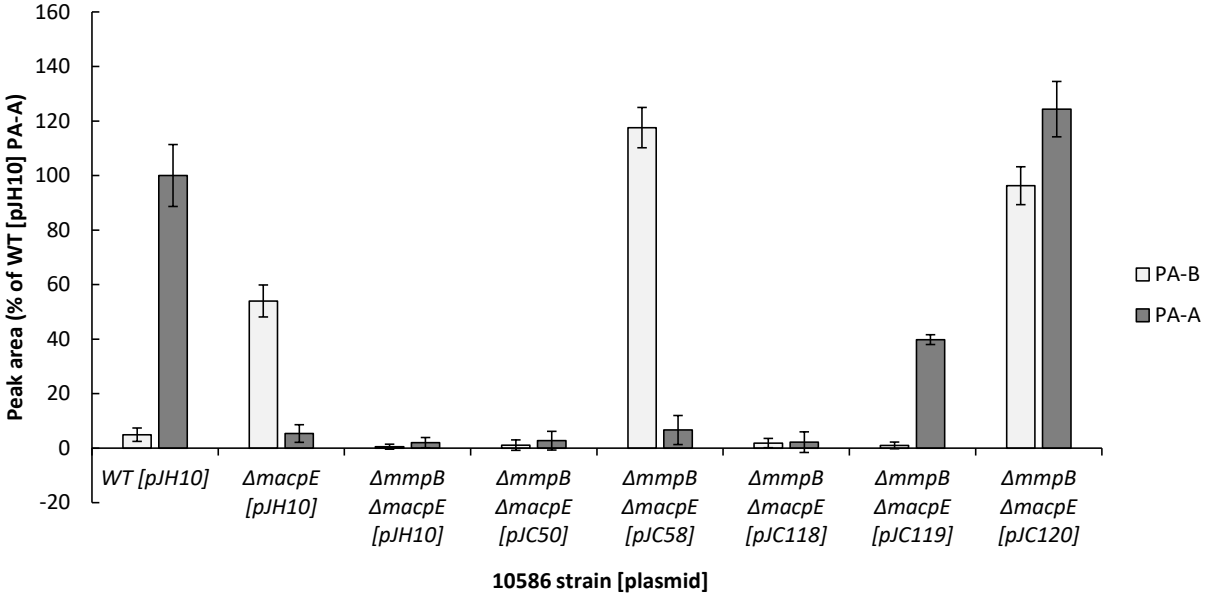


Figure 4.47: Integrated HPLC peak areas: C8-OH removal occurs on hybrid generation III constructs pJC119 and pJC120.
Example chromatograms are shown in Figure 4.48, analyses performed independently in triplicate, error bars are standard deviation. Bars represent average peak area eluting at the same time as PA-B and PA-A standards, 19.7 and 20.6 min respectively, and are expressed as a % of the WT PA-A area.

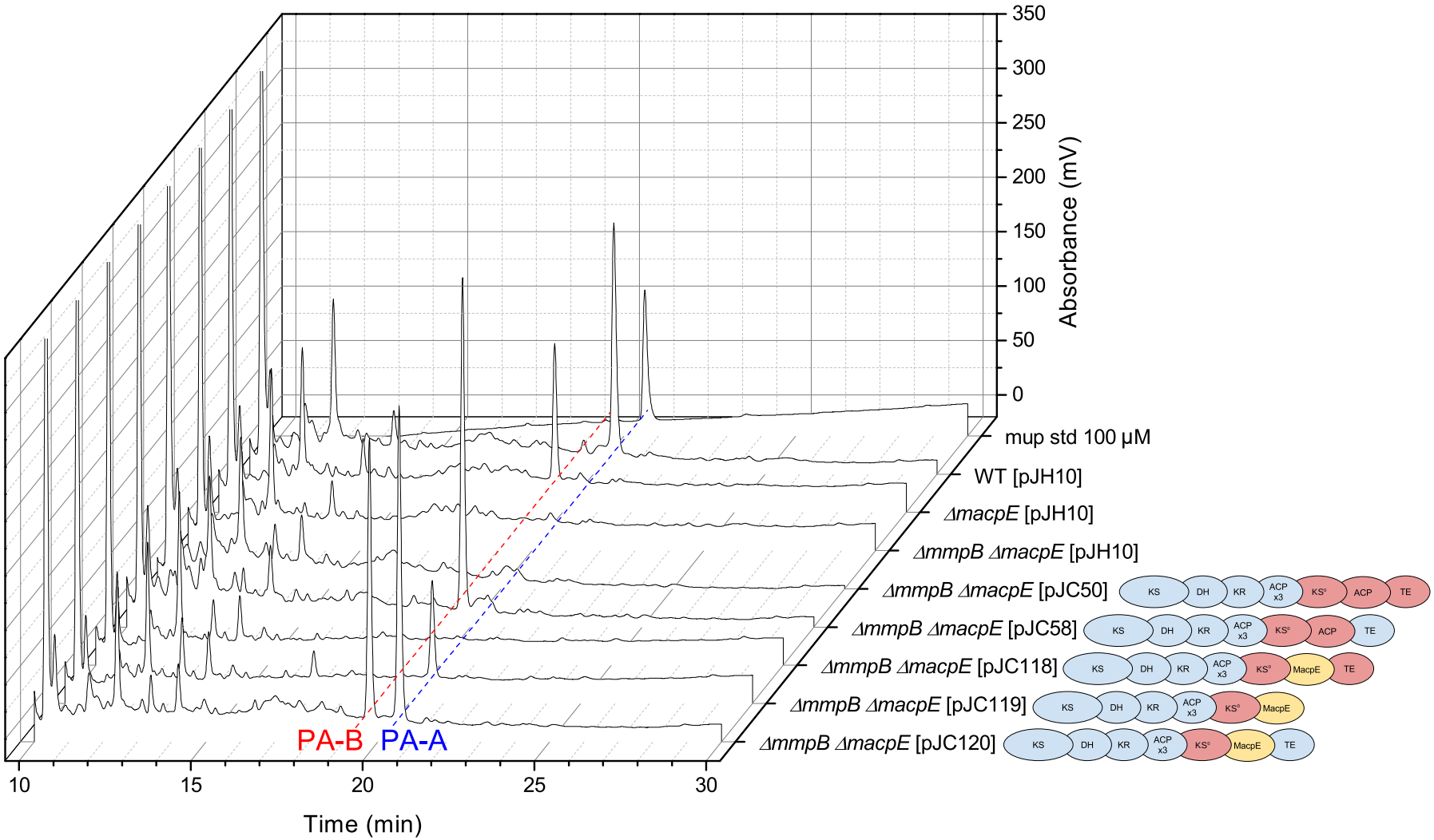
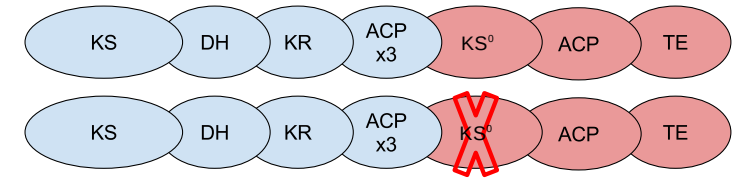


Figure 4.48: HPLC: C8-OH removal occurs on hybrid generation III constructs pJC119 and pJC120.
Representative chromatograms are shown, analyses performed independently in triplicate without IPTG induction.



Hybrid generation	Name	Domain summary	P. fluorescens NCIMB 10586 strain: mmpB ΔACP567	mmpB ΔTE	ΔmmpB	ΔmmpB ΔmacpE	ΔmmpB ΔmacpE ΔmupU
-	pJH10	*Empty vector*	-	-	-	-	-
I	JC50	KS DH KR ACP x3 KS⁰ ACP TE	+++ PA-A	-	-	-	**
I	JC50 C1824A	KS DH KR ACP x3 KS⁰ ACP TE	**	**	-	**	**
II	JC58	KS DH KR ACP x3 KS⁰ ACP TE	+++ PA-A	+++ PA-A	+++ PA-A	+++ PA-B	+++ PA-B
II	JC59	KS DH KR ACP x3 KS⁰ ACP TE	+++ PA-A	+++ PA-A	+++ PA-A	**	**
II	JC58 C1824A	KS DH KR ACP x3 KS⁰ ACP TE	**	**	+ PA-A	**	+ PA-B
II	JC59 C1824A	KS DH KR ACP x3 KS⁰ ACP TE	**	**	+ PA-A	**	**
III	JC118	KS DH KR ACP x3 KS⁰ MacpE TE	**	**	-	-	-
III	JC119	KS DH KR ACP x3 KS⁰ MacpE	**	**	+ PA-A	+ PA-A	+ PA-B
III	JC120	KS DH KR ACP x3 KS⁰ MacpE TE	**	**	+++ PA-A	+++PA-A +++PA-B	

Figure 4.49: Summary of MmpB/TmpB hybrid tests reported in this chapter.
Legend: +++ represents higher production (>50% of WT), + lower production (<50% of WT), ** is not tested, red dash represents no PA-A or PA-B.

4.3 Discussion

In the work presented in this chapter, the timing of C8-OH removal in mupirocin biosynthesis has been re-engineered, drawing on the knowledge gained in Chapter 3. A series of MmpB/TmpB hybrids were generated that demonstrate the capability of the KS^0 to load following ACPs, including the non-cognate MacpE. In the course of these experiments, further insights have been gained into the function of MmpB, TmpB and the PA-B to PA-A tailoring pathways.

4.3.1 The TmpB KS^0 domain may transfer PA-B onto the following ACP

The successful complementation of 10586 Δ*mmpB* by MmpB/TmpB hybrid generation II demonstrated that the PKS could tolerate the presence of the foreign extra module, whilst still performing the normal MmpB function of 9-hydroxynonanoic acid (9-HN) biosynthesis. The TmpB KS^0 is lacking the conserved histidine of the HGTGT motif, which is required for chain elongation, but instead it has QGTGT (starting from Gln2090 in TmpB). The KS^0 domain does have the conserved active site cysteine, Cys1824, the equivalent of which has been demonstrated to be absolutely required for turnover in DEBS (Robbins et al., 2016). As predicted if the inserted module is functional, inactivation of the KS^0 domain by point mutation (C1824A) decreased the throughput of pseudomonic acid on the generation II hybrids.

However, there was still some throughput with the mutation, indicated by complementation of PA-A production in 10586 Δ*mmpB* by pJC58.C1824A and pJC59.C1824A at 10% and 29% of the non-mutant versions respectively. This suggests that either offloading is occurring directly from the MmpB ACP triplet, as would occur in WT MmpB, or that PA-B is somehow transferred to the module 2 ACP at a lower rate. This could perhaps occur through the simultaneous docking of two ACPs to the KS, and spontaneous transfer between the ACPs. Without a functioning KS^0 domain,

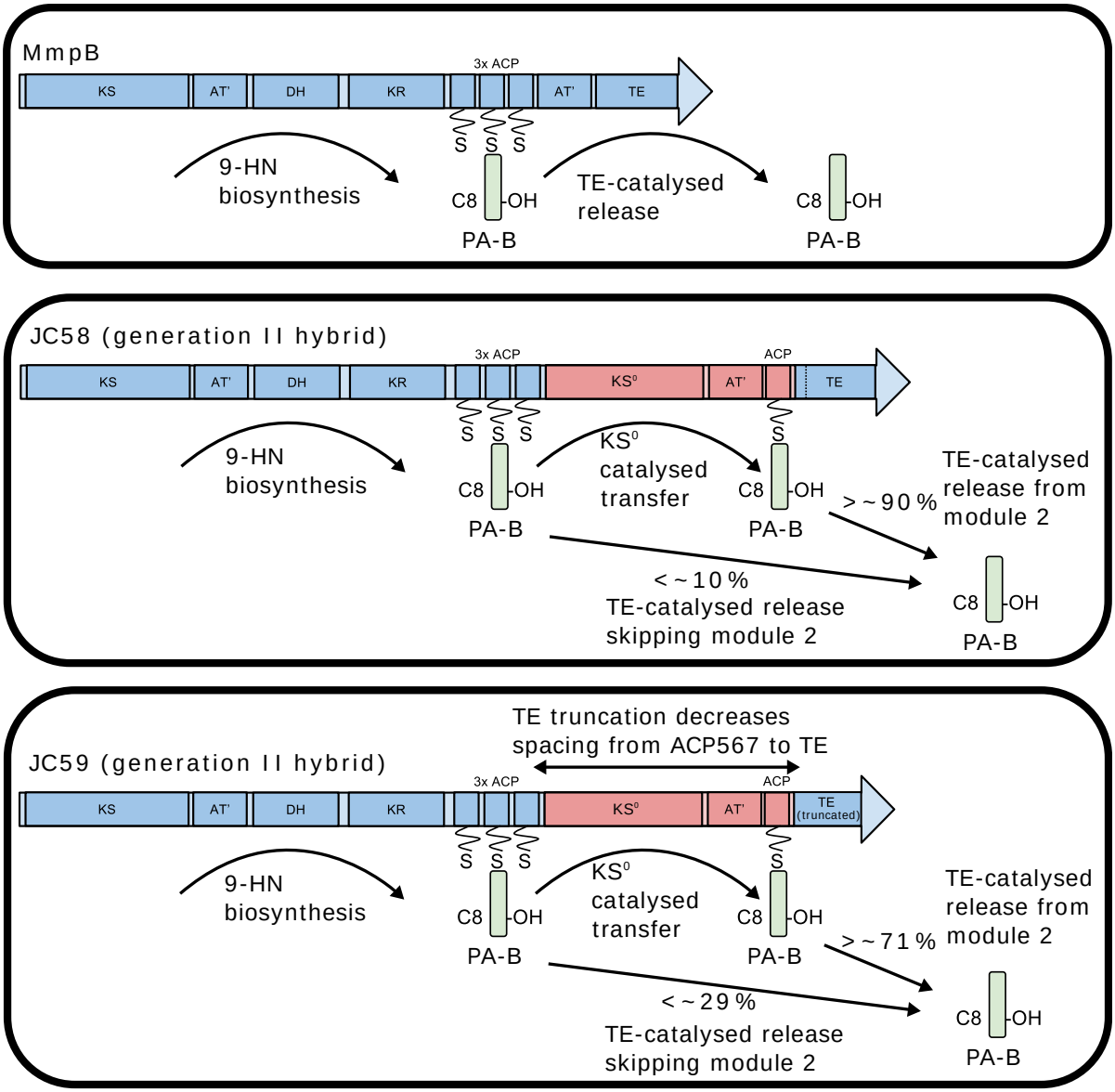


Figure 4.50: The thioesterase skipping hypothesis: the MmpB TE may catalyse release of PA-B from its native ACP triplet in the MmpB/TmpB hybrids.
The proportion of PA-B released directly from the ACP triplet is inferred from the level of PA-B production in the KS^0 mutants. It is possible that the functional KS^0 could decrease the proportion of PA-B released by the proposed skipping mechanism compared with the KS^0 mutants, due to competing rates of reaction.

the latter seems unlikely. Instead the former hypothesis seems more likely: the thioesterase can effectively skip the second module, albeit with lower efficiency, and directly offload PA-B from ACP5, ACP6 or ACP7 (Figure 4.50).

This is supported by the observation that the proposed skipping occurs more efficiently in hybrid JC59 (29%) than JC58 (10%). In JC59, the N-terminus of the thioesterase domain is truncated by 35 AA, which decreases the number of amino acids between the ACP triplet and TE domain. This could potentially aid contact between the TE and ACP triplet.

To test this proposed model, a point mutation to alanine could be made to the phosphopantetheine (PP) carrying serine of the module 2 ACP, Ser2298. The TE skipping module would predict that this would yield the same phenotype as the KS^0 mutation. However, if the 10% or 29% PA-B still produced by the KS^0 mutant hybrids was somehow due to transfer to the module 2 ACP, then S2298A should abolish this entirely.

These experiments have demonstrated the ability of this thiomarinol KS^0 to process PA-B, and its usefulness as a tool for re-engineering mupirocin biosynthesis. It is also possible that it could prove useful for the *in vivo* re-engineering of other PKS pathways, especially as the hybrid generation III results indicated it can load a foreign ACP. However, we do not know what, if any, substrate specificity the KS^0 has, beyond that it accepts PA-B and the equivalent thiomarinol intermediate (marinolic acid B). It is possible that the KS^0 could display gatekeeper activity towards correct PKS chain length, but it is not known what controls the number of condensations the iterative KS performs.

4.3.2 *In vivo* re-engineering of the MacpE tailoring steps to occur on MmpB was successful

The long term ambition of *in vivo* generation of mupirocin NRPS hybrids, through re-engineering of MmpB, requires the significant milestone of removing the C8-OH on

MmpB. In the native mupirocin pathway, this is thought to occur on the freestanding acyl carrier protein, MacpE. In this work, the C8-OH removal steps that would normally occur on freestanding MacpE were successfully re-engineered to occur on MmpB/TmpB hybrids in two different ways.

4.3.2.1 Conversion on MmpB/TmpB hybrid generation II mediated by thiomarinol enzymes

Firstly, conversion was re-engineered to occur on generation II hybrid JC58 with thiomarinol tailoring enzymes TmlO, TmlP, TmlC, and TmlF. Chromosomal hybrid mutant 10586 *mmpB*.JC58 Δ*mupU* produces only PA-B (Section 4.2.4.3), as without MupU, PA-B is not loaded onto MacpE or converted to PA-A. When TmlO, P, C, and F were provided on a plasmid to this strain, 12% of pseudomonic acid production shifted to PA-A, despite the absence of MupU (Figure 4.51). In this scenario, MupO, P, C, F and V are all also present, so it is not clear which combination of these enzymes is required for conversion on hybrid II. However, at least one of TmlO, P, C or F is required, as no PA-A was produced by 10586 *mmpB*.JC58 Δ*mupU* without the four-gene expression plasmid.

This method of conversion proved inefficient, with a ratio of 88:12, PA-B:PA-A, near the inverse of WT 10586. This could be explained by the lack of TE specificity, as summarised in Figure 4.51. In normal mupirocin biosynthesis, the MmpB TE domain is thought to release PA-B. This means that after KS^0 mediated transfer of PA-B from the MmpB ACP triplet to the module 2 ACP, the MmpB TE could catalyse immediate release of PA-B. In addition to this, we can predict that the skipping effect described previously will occur, and approximately 10% of PA-B will be putatively released from the MmpB ACP triplet before transfer by the KS^0 domain. There are also further alternative possible release mechanisms that cannot be ruled out: an AT domain in pederin biosynthesis has been reported to have editing activity (Jensen et al., 2012), as has thioesterase II in pikromycin biosynthesis (Kim et al., 2002). Any

PA-B released by either mechanism should not be converted to PA-A due to the absence of MupU.

As the TE domain cannot demonstrate any gatekeeping activity towards PA-A, there could be a competition between the thiomarinol or mupirocin tailoring enzymes and the TE domain. The low efficiency of conversion may be due to the majority of PA-B being released by the TE domain before tailoring enzyme mediated conversion on the second module ACP can occur.

To continue this work, there are multiple further scenarios that could be tested. Firstly, if the presented model is correct, then the observed conversion phenotype should not require MacpE. 10586 *mmpB*.JC58 Δ*mupU* Δ*macpE* could be generated by further mutagensis of the double mutant, and should still produce 10% PA-A. Secondly, insertion of the KS^0 mutant JC.58.C1824A into the chromosome should abolish any conversion. These models predict that 10586 *mmpB*.JC58.C1824A Δ*mupU* would produce only a small amount of PA-B mediated by "skipping" release from the ACP triplet, and no PA-A.

If the hypothesis that the low efficiency of PA-A production is due to premature offloading of PA-B by the TE domain is correct, then partially disruptive point mutation to the TE domain could increase the relative proportion of PA-A. A range of point mutations could be introduced into JC58 TE domain amino acids in 10586 *mmpB*.JC58 Δ*mupU*. TE mutants that were defective in release of PA-B, but still functional, could be identified by HPLC of SSM culture supernatants. If TmlOPCF were provided to such a strain, then the model would predict an increase in PA-A production relative to PA-B, as the TE turnover of PA-B would be lower allowing more time for conversion to occur. However, this would likely only increase the proportion of PA-A, and not the yield, as the tailoring process should still occur at the same rate.

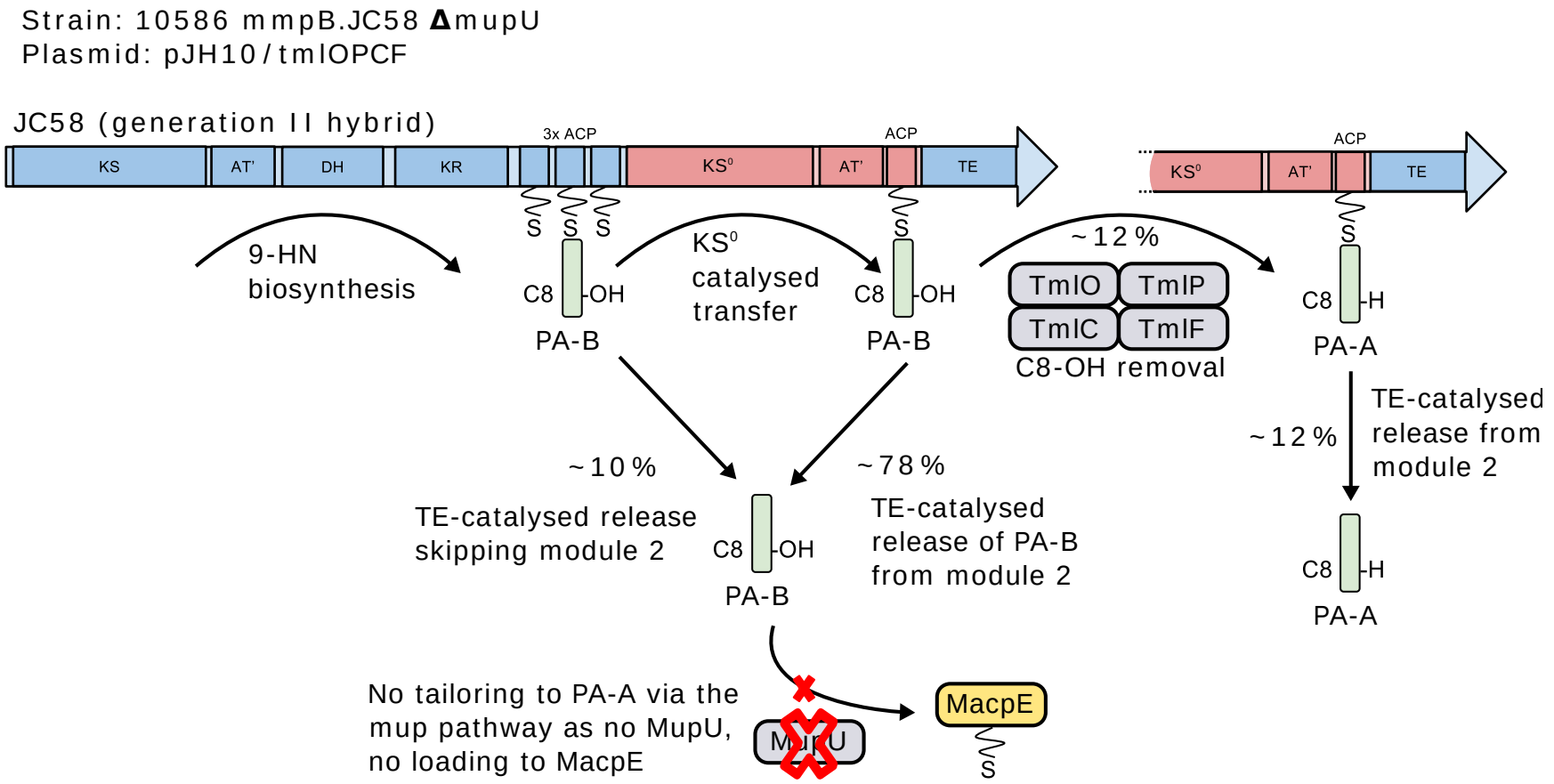


Figure 4.51: Proposed model of thiomarinol enzyme mediated conversion on MmpB/TmpB hybrid II.
Any PA-B released from the hybrid will not be processed further as no MupU is present. As the MmpB TE domain will offload PA-B or PA-A, thiomarinol enzyme mediated conversion needs to occur before TE catalysed release. This provides a putative explanation for the 12:88 (PA-A:PA-B) proportion: the TE domain offloads the majority of PA-B before the tailoring enzymes manage to convert it to PA-A.

As mentioned previously, we can deduce that at least one of TmlO, P, C, or F is required for this conversion. As MupO, P, C, F, and V are also present in the test, they could still be functioning in this re-engineered conversion pathway. Experiments could be performed where successive deletions were made of each of these domains to determine which combination(s) are required.

These results represent the first evidence for the proposed hypothesis that some or all of the TmlOPCF-catalysed C8-OH removal steps occur on TmpB module 2. Ultimately, the alternative approach of re-purposing the native mupirocin tailoring enzymes to generate PA-A on MmpB/TmpB hybrid generation III immediately proved more efficient than using thiomarinol enzymes.

4.3.2.2 Conversion on MmpB/TmpB hybrid generation III through the re-purposing of MacpE

As covered throughout Chapter 3, conversion of PA-B to PA-A is thought to occur at least partially on the free-standing acyl carrier protein, MacpE. MmpB/TmpB generation III consisted of three hybrid constructs with MacpE positioned behind the KS^0 domain. Hybrids JC119 (no TE domain) and JC120 (MmpB TE domain) were both observed to restore PA-A production in 10586 Δ*mmpB* Δ*macpE*. The proposed models outlining a mechanism for the hybrid III observations are outlined in Figure 4.52.

Hybrid JC120 proved an effective MmpB replacement, able to complement 10586 Δ*mmpB* to 124 ±14% of WT bioassay zone, and 272 ±7% of WT PA-A HPLC peak area. In the double mutant 10586 Δ*mmpB* Δ*macpE*, JC120 was observed to restore both PA-B and PA-A to more than WT PA-A levels (Figure 4.48). Complementation of the double mutant was also observed by plate bioassay, to 156 ±12% of WT area. These data demonstrate that JC120 can effectively perform the native MmpB function of 9-HN and therefore PA-B biosynthesis, but also act as the site of the PA-B to PA-A tailoring steps that natively occur on MacpE. It is possible that each step catalysed

by MupO, P, C and F occurs on MacpE, this hypothesis is explored in Section 5.2.1.1.

10586 Δ*mmpB* Δ*macpE* [pJC120] was observed to produce approximately equal proportions of PA-B and PA-A. If the TE skipping hypothesis (Figure 4.50) were correct, then a proportion of this PA-B could be attributed to direct offloading from the MmpB ACP triplet by the TE. However, in generation II hybrid JC58, 10% of the PA-B produced appears to be offloaded before KS^0 processing. This therefore does not explain the large production of both PA-B and PA-A by JC120. Instead, it seems most likely that TE catalysed release of PA-B could also occur from MacpE on JC120 (Figure 4.52). This model corresponds with the previous *tml* enzyme model (Section 4.3.2.1), that C8-OH removal must occur quickly on the second module ACP or the TE will release PA-B. Conversion appears to occur much more efficiently on JC120 through the relocated native pathway than on JC58 in the thiomarinol style, which is perhaps not suprising.

In 10586 Δ*mmpB* Δ*macpE*, JC119 was observed to only offload PA-A, albeit at a much lower throughput than JC120. This supports the proposed models: no PA-B is released prematurely as the TE domain is not present. Without the PKS TE domain, the most likely candidate for release of the observed PA-A is MupV (SDR and TE). In the putative native PA-B to PA-A tailoring pathway, MupV is thought to catalyse release from MacpE. It would make sense if MupV had evolved a specificity towards releasing PA-A (or mupirocin C/F) and not releasing unprocessed PA-B, to avoid leakage of PA-B from the pathway. This agrees with the 10586 Δ*mmpB* Δ*macpE* [pJC119] phenotype of only PA-A production, but is contradicted by 10586 Δ*mmpB* Δ*mupU* Δ*macpE* [pJC119] PA-B production phenotype.

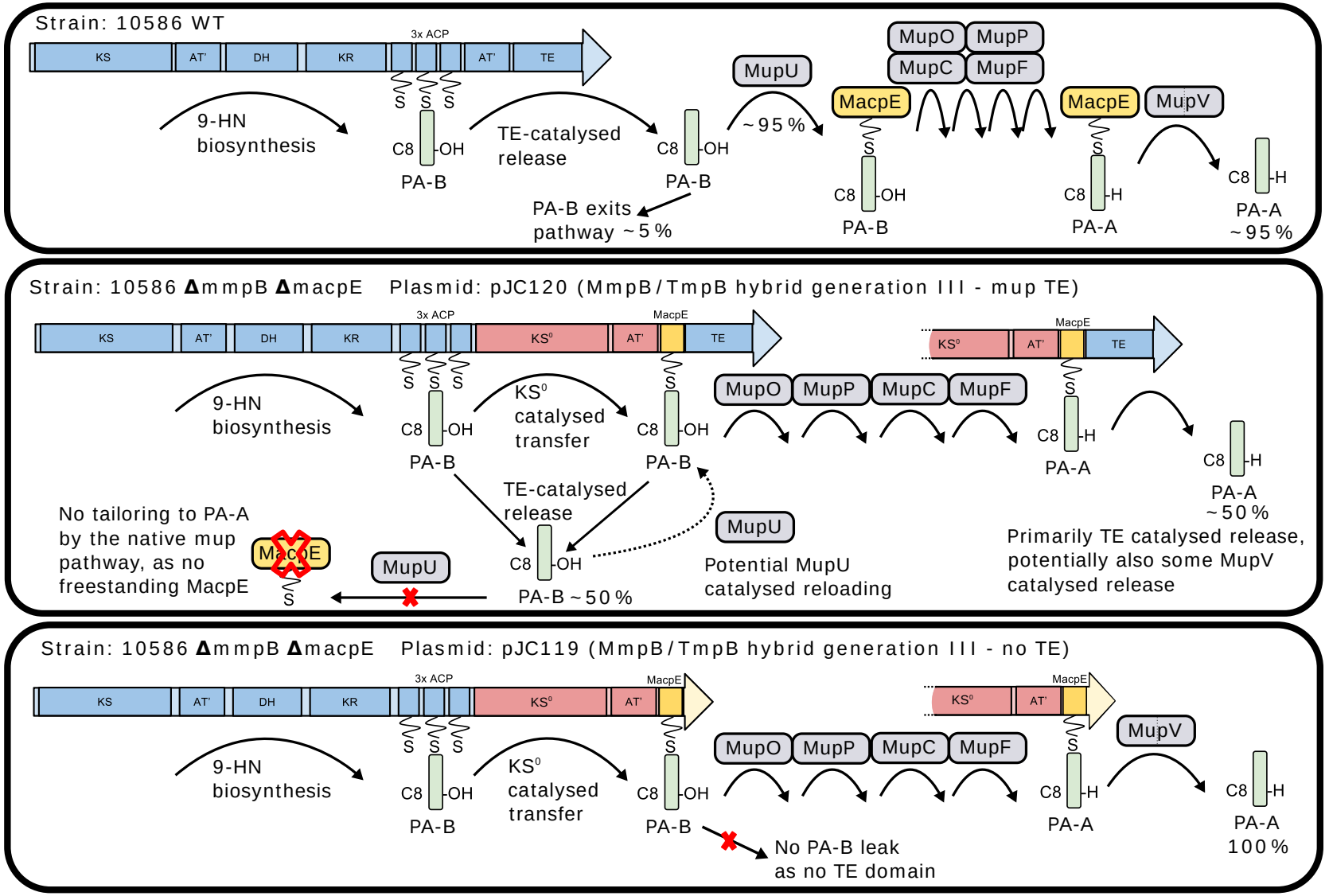


Figure 4.52: Proposed model of PA-A production by MmpB/TmpB generation III hybrids JC119 and JC120.

That JC119 releases less PA-A and PA-B than JC120 suggests that regardless of the release mechanism in the absence of the TE domain, it occurs at a slower rate. An alternative explanation for 10586 Δ*mmpB* Δ*macpE* [pJC119] only producing PA-A is that this slower release allows full conversion of PA-B to PA-A on the in-line MacpE. These hypotheses could be tested by deletion of *mupV* in 10586 Δ*mmpB* Δ*macpE*. If MupV were responsible for release from JC119, then its deletion should abolish the observed PA-A production. Further to this, the MupV point mutations described in Section 3.2.1.2 could be used, and this may yield insights into the function of the two domains.

4.3.3 Insights from the lack of activity of MmpB/TmpB hybrid I

Although the objective of re-engineering C8-OH tailoring was only achieved in later hybrid iterations, the non-functional MmpB/TmpB hybrid I, in particular its TmpB TE domain, yielded some interesting insights.

4.3.3.1 The MmpB and TmpB thioesterase domains may have different substrate specificity

MmpB/TmpB hybrid I (JC50), incorporating the TmpB TE, failed to complement 10586 *mmpB*ΔTE and Δ*mmpB* (Section 4.2.2.2). The protein hybrid must be correctly expressed, as it was observed to complement ΔACP567. MmpB/TmpB generation II hybrids (JC58 and JC59), with the MmpB TE instead, were able to complement both *mmpB*ΔTE and Δ*mmpB*. These results suggest the TmpB TE was not functional in the hybrids.

One explanation could be failure of the TmpB TE domain to recognise the preceding ACP in the PKS. However, the hybrid has been constructed with the TmpB module 2 ACP preceding the TE domain, exactly as it is in TmpB. Therefore, the TE domain should be able to interact with its native ACP, and it seems unlikely this is the issue. The experiments with MmpB/TmpB hybrid generation II and III indicated that

the MmpB TE can interact and offload from the TmpB module 2 ACP, also suggesting this explanation is unlikely.

A more likely explanation is that the TmpB TE domain is demonstrating substrate specificity. The erythromycin TE has been reported to demonstrate some sterochemical specificity towards hydrolysis of different chain length linear polyketides (Gokhale et al., 1999). Conversely, in similar experiments with the same DEBS TE and the pimaricin TE, with a more diverse range of *N*-acetyl-cysteamine (SNAC) substrates, very little substrate specificity was observed (Sharma and Boddy, 2007). In biosynthesis of the NRPS-synthesised β-lactam nocardicin, the TE domain was demonstrated to have stringent specificity for β-lactam-containing substrates (Gaudelli and Townsend, 2014). It therefore seems possible that TE domains could vary in substrate specificity (Horsman et al., 2016), particularly in a scenario like TmpB, where there could be evolutionary pressure on selectivity. Assuming the proposed model of C8-OH removal on TmpB is correct, it would make sense if the TE domain was specific to C8-H intermediates, thereby only releasing after conversion to the A form.

The TmpB TE domain could also be demonstrating substrate specificity towards other parts of the molecule. One hypothesis, that the TmpB TE demonstrates specificity towards des-epoxy intermediates as would be seen in thiomarinol, was tested and rejected in Section 4.2.4.2, as it was unable to offload PA-C.

In the future, it would be interesting and potentially impactful to the polyketide field in general, to both determine if the TmpB TE domain is demonstrating specificity, and if so what the structural basis for it is. This investigation could begin with the cloning, expression and affinity purification of the TmpB TE domain from a suitable vector such as pET28a. Then, its ability to catalyse the release of different mupirocin and thiomarinol related compounds as *N*-acetylcysteamine (SNAC) thioesters could be tested.

There are no close homologs to the either the MmpB or TmpB TE domains in the PDB of solved structures. HHPRED searches of the PDB identified the best structures

as NRPS thioesterase domains from tyrocidine synthetase 3 (4ZXI) and enterobactin synthase F (3TEJ), at 16 and 17% amino acid similarity. This similarity is too low for homology modelling to give any useful insight into the specificity differences. However, it makes the solving of the X-ray structures of these thioesterase domains a useful and impactful objective.

4.3.3.2 MmpB/TmpB hybrid I may form a heterodimer with MmpB

It is well established that PKS enzymes form homodimers, and attempts have been made to identify the regions responsible for this dimerisation. NMR analysis of DEBS enzymes revealed regions of hydrophobicity thought to be responsible for stabilising dimerisation (Broadhurst et al., 2003), whereas a dimersation motif before the KR domain has been reported in spinosyn (Zheng et al., 2013).

The results with MmpB/TmpB hybrid I provide evidence for heterodimer formation with WT MmpB (Figure 4.53). The TmpB TE in hybrid I proved non-functional, and was unable to complement a chromosomal TE in-frame deletion mutant. However, hybrid I was observed to complement *mmpB*ΔACP567. In this scenario, the *in trans* hybrid I lacks a functional TE domain, whereas the chromosomal MmpB lacks ACP domains. To generate PA-A as observed, domains from both the plasmid and chromosome must function on the same molecule. One explanation is the TE domain is able to function *in trans* on the chromosomal copy lacking ACPs. Another possibility is that these form a heterodimer, using the ACP domains from the chromosomal copy and the TE from the hybrid.

This seems the most likely explanation, as PKS enzymes are known to form homodimers. Similar heterodimers have been reported in pikromycin PKS, PikAIV, where a chromosomal TE deletion was complemented by inactive hybrids carrying a functional TE, with the conclusion that the ketosynthase domain is responsible for dimerisation (Chen et al., 2000).

It could be possible to identify regions necessary for dimerisation of MmpB by

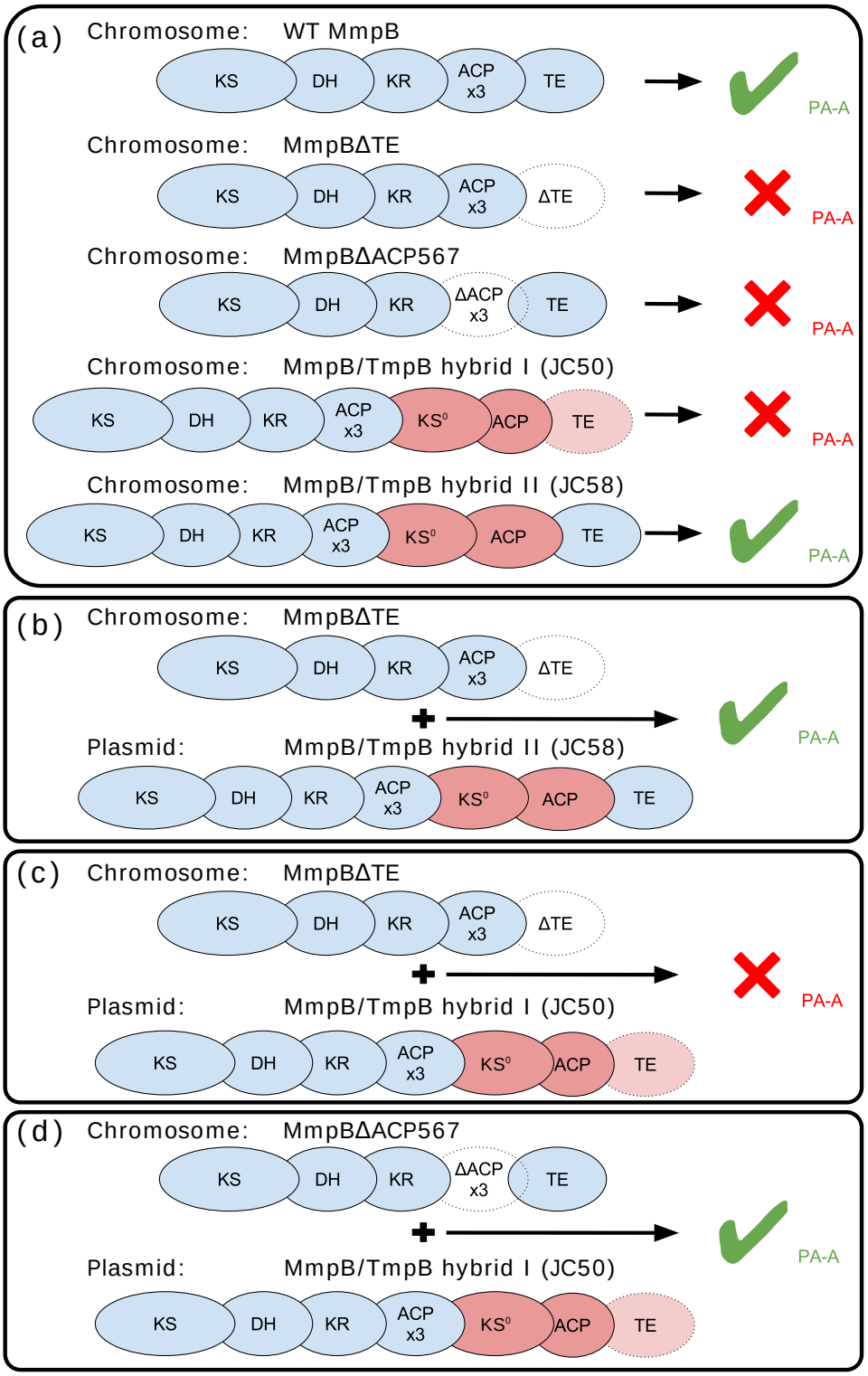


Figure 4.53: Evidence of heterodimer formation between MmpB variants.
MmpB domains are shown in blue, TmpB in red.
(c) Hybrid I is unable to complement a chromosomal MmpB TE deletion, which suggests the TmpB TE domain is not functional.
(d) Hybrid I is able to complement a chromosomal ACP deletion, yet neither peptide has all the domains required for function. This could be due to formation of heterodimers between the ACP mutant MmpB and hybrid I.

providing differing truncations *in cis* and *in trans*. A series of chromosomal MmpB N-terminal truncations could be constructed alongside a series of C-terminal truncations on pJH10. The ability of these to complement each other would yield insight into regions responsible for dimerisation. Similar experiments have been performed in fungal type I PKS ATX (Moriguchi et al., 2006).

4.3.4 Induction of MmpB expression by IPTG is inhibitory to production of mupirocin

Throughout all the experiments reported in this chapter, induction of expression of MmpB and its hybrid derivatives have proven inhibitory to mupirocin production. IPTG is used to induce the *tac* promoter in pJH10 derivatives, where it relieves LacI repression at the *lac* operator. This effect is not observed with empty pJH10 under IPTG induction, and therefore must be due to the MmpB expression.

The effect was best characterised in Section 4.2.1.3, where MmpB was expressed from pJH10/*mmpB* in *P. fluorescens* WT. Induction at 50 μM IPTG was partially inhibitory, whereas 500 μM IPTG was observed completely inhibitory to detectable mupirocin production.

One possible explanation is that MmpB needs to be expressed to a specific level relative to the later tailoring enzymes, assuming MmpB interacts with these enzymes. In such a scenario, the majority of MmpB enzymes would not carry intermediates, and these unloaded proteins could interact with MupU, effectively sequestering the loaded MmpB from interacting with its downstream partners. However, this would presumably generate PA-B, with a phenotype similar to *P. fluorescens* Δ*mupU*. Induction of MmpB expression in 10586 WT yielded less PA-A with more IPTG, but no increase in PA-B as detected by HPLC analysis.

A more likely explanation is that overexpression of MmpB sequesters MmpC, therefore inhibiting the activity of MmpA and MmpD. MmpC holds the *trans* acting AT domains, without which no extender units should be loaded. This hypothesis

would predict no pseudomonic acid production with MmpB overexpression, and as such seems the most likely explanation for this effect. This could be tested by increasing MmpC expression together with MmpB.

Finally, the transition from no inhibition to significant inhibition occurred over the range of 10 to 50 μM IPTG (Figure 4.10b). Whilst there is no guarantee this will hold true for all genes expressed from pJH10, it is a useful starting point for other experiments. Typically, most experiments involving expression from pJH10 have been performed with no IPTG induction, 0.1 mM or 0.5 mM (Hothersall et al., 2007; Omer-Bali, 2013; Gurney and Thomas, 2011; Alsammarraie, 2016). These results demonstrate that there is a middle ground to be explored.

Chapter 5

GENERAL DISCUSSION AND FUTURE WORK

5.1 General context

There is a well-established current need for new clinical antibiotics, and in particular new compound scaffolds for drug discovery. Historically, natural products have proven a major source of clinical antibiotics, often with chemical modification to improve the drugs in various ways. One pathway to the identification of new antibiotics is the modification of existing clinical antibiotics such as mupirocin.

The work reported in this thesis was undertaken with the aim of contributing to a wider goal of the Thomas group, the generation of *in vivo* mupirocin non ribosomal peptide (NRP) derivatives. These derivatives would be generated on MmpB, which in the mupirocin biosynthesis pathway is proposed to carry the inactive PA-B. Therefore, the main goal of this work arose: *in vivo* re-engineering of late-stage mupirocin biosynthesis so that conversion to active PA-A occurs on MmpB. This could facilitate the *in vivo* generation of NRP-conjugated derivatives of the active PA-A in the future.

5.2 Key conclusions and future work

5.2.1 All mupirocin biosynthetic cluster genes required for conversion of PA-B to PA-A have been identified

In the work described in Chapter 3, experiments were performed with the aim of characterising the surprisingly convoluted conversion pathway between PA-B and PA-A. A systematic screen was developed for the ability of *P. fluorescens* NCIMB10586 mutants to convert PA-B to PA-A. This screen identified 11 genes which, when mutated, were deficient in conversion (Figure 3.12). A series of multi-gene expression plasmids were developed, culminating in pJC133, which carried the 10 mupirocin biosynthesis genes identified by the screen (transcriptional regulator *mupR* was not included). This plasmid encoded all functions necessary for conversion of PA-B to PA-A in 10586 Δ*mup-cluster*, and this represented the first evidence that all mupirocin

biosynthetic cluster genes responsible for conversion of PA-B to PA-A have been identified. However, as discussed in Section 3.3.1.2, no conversion was observed in *P. fluorescens* SBW25 with pJC133.

In the screening experiments, a reproducibly variable defect in PA-B conversion was observed in 10586 Δ*mupL*. Inactivation of *mupL* in pJC133 by point mutation demonstrated its enzymatic activity is not required for conversion of PA-B to PA-A. However, no conversion was observed without *mupL* in the multi-gene expression plasmid. Fortunately, the gene downstream of *mupL* is the self-resistance encoding *mupM*, which allowed for assay of expression levels by determination of mupirocin resistance. These experiments supported the presence of a non-MupR regulated promoter within *mupL*, which increases expression of *mupM* and *mupN*.

The results of both the systematic screen and pJC133-encoded conversion supported the current proposed PA-B to PA-A conversion pathway (Figure 3.41) (Gao et al., 2017). These results provided evidence of a novel requirement for conversion for MupM (IleRS, self-resistance) and MupN (phosphopantetheinyl transferase). As the only ACP required is MacpE, this also represents the first evidence that MacpE is completely reliant on MupN for activation, whereas other mupirocin ACPs are not all reliant on MupN (Shields et al., 2010).

Historically, MupV has been annotated as a single domain enzyme (SDR), thought responsible for catalysing release of late-stage mupirocin intermediates from MacpE (Thomas et al., 2010; Gurney and Thomas, 2011; Gao et al., 2014; Cooper et al., 2005; El-Sayed et al., 2003). However, bioinformatic re-analysis revealed a second domain (TE). In this work, inactivation of either of these domains by point mutation demonstrated the enzymatic activity of both is required for conversion of PA-B to PA-A. It is not clear what additional function MupV could perform, or whether both domains are required for its proposed role in release from MacpE.

5.2.1.1 Proposed model: Tailoring from PA-B to PA-A could occur entirely on MacpE

The Thomas-Simpson partnership have previously hypothesised that conversion to mupirocin P occurs on MacpE, after which MupV catalysed release occurs. MupP, MupC and MupF are then proposed to act on free standing intermediates (Gao et al., 2017). This is supported by the conversion of fed mupirocin C1 to PA-A observed in 10586 Δ*macpE*. However, this author believes it is more likely all tailoring from PA-B to PA-A occurs on MacpE; a comparison of these hypotheses is presented in Figure 5.1. In 10586 WT, PA-A (90%) and PA-B (5-8%) are produced but not mupirocin C1, P or F. If mupirocin P, C1 or F were released from MacpE, as proposed in Gao et al. (2017), then there should be evidence of production of mupirocin P, C1 or F at a readily detectable scale in 10586 WT, as PA-B is. It is possible, and would be more functionally efficient that MupV has a gatekeeping function, specifically catalysing the release of PA-A from MacpE. Indeed, this would explain the production of only PA-A and no PA-B by generation III MmpB/TmpB hybrid JC119, which lacks a thioesterase domain (Section 4.2.5.3).

It is important to note that in this proposed model, PA-B generated on MmpB would be converted to the final PA-A on MacpE, but not obligately so for all steps. The conversion of PA-B to mupirocin P obligately requires MacpE, but externally fed Mupirocin C1 has been observed to be processed to PA-A in the absence of MacpE (Gao et al., 2014).

In Thomas et al. (2010), published before the MupP function was uncovered, the MupC- and MupF-catalysed steps are pictured occurring on an unknown ACP, labelled ACPX. This hypothesis follows a similar logic to **B)**, that these steps would occur on a carrier protein to facilitate efficient conversion. We have now putatively assigned all the ACPs in the mupirocin pathway to functions, so it seems the best candidate is for MacpE to perform all the conversion steps.

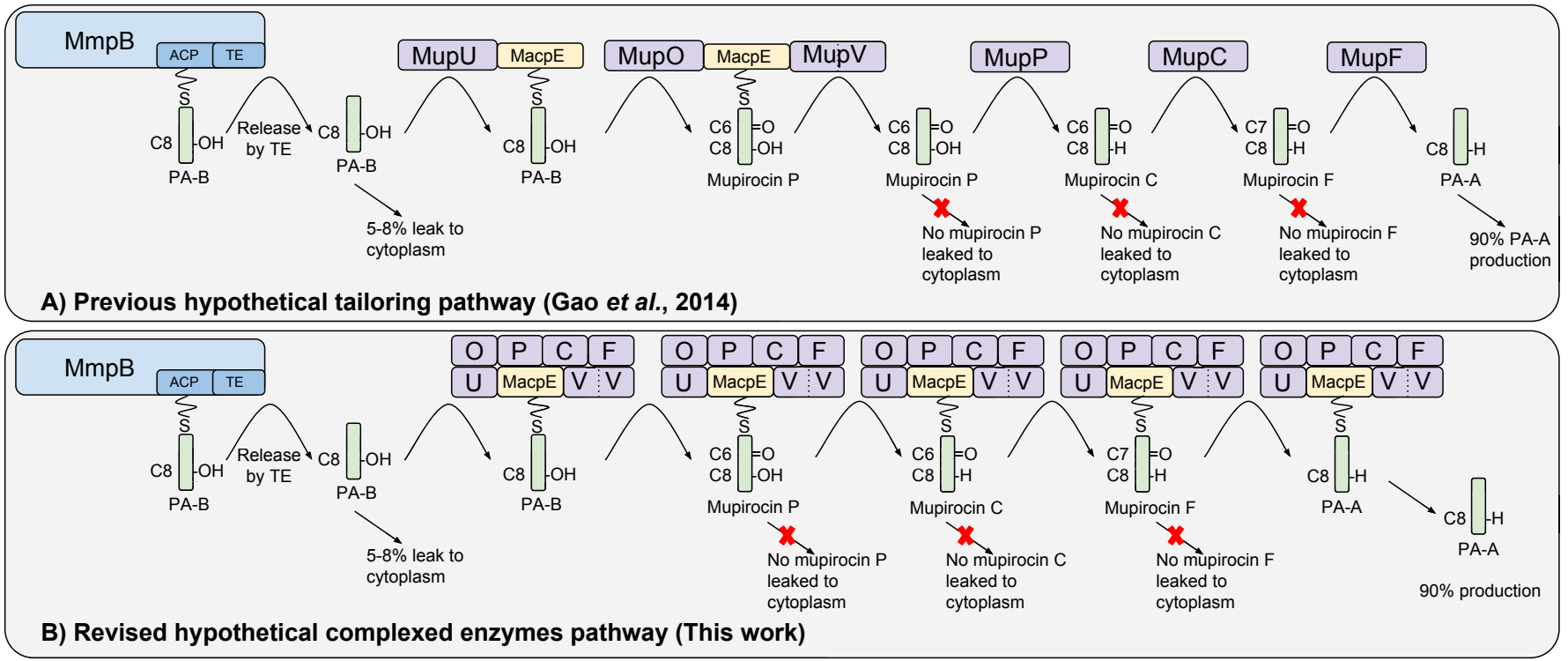


Figure 5.1: Proposed models of PA-B to PA-A conversion on MacpE.
A) Our recent publications on this part of the pathway have depicted conversion from mupirocin P onwards occuring after release from MacpE (Gao et al., 2014, 2017). However no detectable mupirocin P, C or F is produced by 10586 WT.
B) In this work, an alernative hypothesis is proposed where all tailoring steps from PA-B to PA-A occur whilst the intermediates are still on MacpE.
The key evidence in support of model **B)** instead of **A)** is lack of mupirocin P, C or F produced by 10586 WT, whereas PA-B and PA-A are produced at significant scales. Further to this, it is possible the tailoring enzymes could form a complex with MacpE as shown, although there is no direct evidence for or against this.

A further prediction: MupU could interact with the TE of MmpB

The thiomarinol homolog of MupU, TmlU, has been demonstrated to bind the MmpB thioesterase domain by the bacterial two hybrid system (Alsammarraie, 2016). Furthermore, expression of *tmlU* in 10586 caused a switch to C8-OH intermediates with truncated fatty acids, which indirectly also supports a TmlU-MmpB interaction (Omer-Bali, 2013; Alsammarraie, 2016). In mupirocin biosynthesis, 90% of all pseudomonic acid produced by the endogenous pathway is converted to PA-A, whereas only around 40% of exogenous fed PA-B is converted to PA-A (Section 3.2.2.2). It is possible MupU also interacts with MmpB, which would tether PA-B production and conversion to PA-A in physical space in the cell. In this model a complex of these tailoring enzymes would form, allowing for the efficient production of PA-A observed in the cell to occur. This is supported by the need for MupU for PA-A conversion on generation III MmpB/TmpB hybrid JC120, where the normal MupU function should be provided by the KS^0 domain (Section 4.2.5.3).

5.2.1.2 Future work: *In vitro* investigation of the mupirocin PA-B to PA-A tailoring pathway

In vitro analysis of this tailoring pathway is an attractive prospect that could answer the remaining questions and appraise the current model. Each of MupO, MupP, MacpE, MupC, MupF, MupU, MupV, and MupN could be cloned individually to a suitable protein expression vector such as pET28a. In this example vector, each protein would be expressed from *E. coli* BL21, and nickel affinity purified using terminal His tags.

The logical starting objective would be the generation of *holo*-MacpE followed by its loading with PA-B. The expressed *apo*-MacpE is inactive as it lacks the phosphopantetheine arm on which it holds substrates. Activation could be achieved by co-incubation of apo-MacpE, MupN (or Sfp), and coenzyme A. MupN (phosphopantetheinyl transferase) should catalyse the transfer of 4'-phosphopantetheine from CoA

to MacpE, generating the active *holo* form. Co-incubation of *holo*-MacpE, MupU, PA-B and ATP would hopefully result in MupU-catalysed loading of PA-B onto MacpE, which could be monitored by HPLC or LC-MS.

Once MacpE carrying PA-B had been generated, co-incubation with each of the other enzymes could be performed to confirm the pathway proposed in Gao et al. (2017), generating mupirocin P, C, F then PA-A successively. In Chapter 4, conversion was observed to occur with MmpB/TmpB hybrid generation III only when MupU was present, although the MupU function should be performed by the KS^0. Therefore, it would be intriguing to test whether conversion of PA-B on MacpE to PA-A could proceed without MupU present.

The proposed MupV role of catalysing release from MacpE could also be investigated. MacpE loaded with either PA-B, mupirocin P, mupirocin C, mupirocin F and PA-A could be co-incubated with MupV to test its role and determine any specificity. Furthermore, expression of the *mupV* point mutants could determine the role of each domain.

Finally, protein-protein interaction experiments, such as co-immunoprecipitation, could be used to test for interactions between these enzymes. For example, MupU could be generated with a FLAG affinity tag, and mixed with His-tagged MacpE. Nickel bead purification of MacpE should also capture MupU if there is an interaction, and this could be detected by Western blots using FLAG antibodies conjugated to a reporter.

5.2.2 Re-engineering of late-stage mupirocin biosynthesis has been achieved

In Chapter 4, successive generations of hybrids between MmpB and TmpB were generated with the aim of achieving the key goal of this work: re-engineering of mupirocin biosynthesis such that C8-OH removal occurs on MmpB. In the PA-B producer 10586 Δ*mupU*, restoration of 12% PA-A as a proportion of total pseudomonic

acids was achieved with generation II hybrid, JC58, with *tmlO*, *tmlP*, *tmlC* and *tmlF* expressed *in trans* (Section 4.2.4.3). MmpB/TmpB hybrid generation II has the TmpB module 2 ACP positioned behind the KS^0 domain as it is in TmpB. Without MupU, no PA-B should be loaded onto MacpE, and no MupO-catalysed conversion to mupirocin P should occur. Instead, it seems likely that the equivalent steps have been achieved on the module 2 ACP using the thiomarinol enzymes. This represents the first experimental evidence supporting the proposed hypothesis of C8-OH removal occurring on TmpB in thiomarinol biosynthesis.

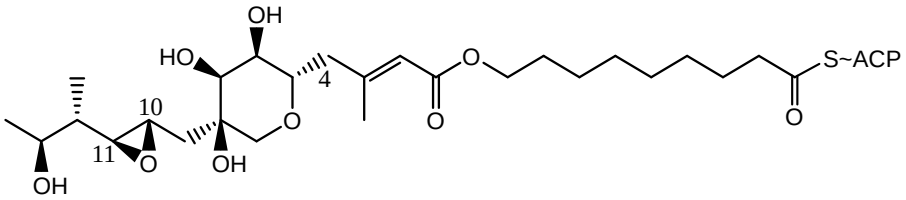
Restoration of conversion was also achieved with MmpB/TmpB hybrid generation III, which have MacpE in the second module. In 10586 Δ*mmpB* Δ*macpE*, generation III hybrids JC119 (no TE) and JC120 (MmpB TE) were observed to restore 40% and 124% PA-A respectively as a proportion of 10586 WT PA-A (Section 4.2.5.3). Without a chromosomal copy of *macpE*, all PA-B to PA-A conversion steps that would normally occur on MacpE have been successfully re-engineered to occur on the generation III hybrids. As proposed in Section 5.2.1.1, it is possible all conversion to PA-A occurs on MacpE, and that PA-A is released from hybrid generation III. However, release of an earlier intermediate, such as mupirocin P, and subsequent tailoring off of the hybrid cannot be ruled out. Regardless, these results represent promising progress towards the longer term goal of generation of a PA-A NRP derivative through coupling of NRPS domains to MmpB.

5.2.2.1 The thiomarinol KS^0 proved a useful tool for re-engineering mupirocin biosynthesis

The TmpB KS^0 domain from thiomarinol biosynthesis has proven an effective tool for the transfer of mupirocin intermediates to an extra ACP. KS^0-mediated transfer was demonstrated when the inactivation of the KS^0 by point mutation decreased the PA-B throughput on JC58 to 10% (Section 4.2.3.4). Furthermore, the successful re-engineering of tailoring steps to occur on the TmpB ACP in generation II hybrids,

and MacpE in generation III hybrids, indirectly demonstrates the KS^0 function in loading these ACPs.

(a) Pseudomonic acid B



(b) Marinolic acid B (hypothetical)

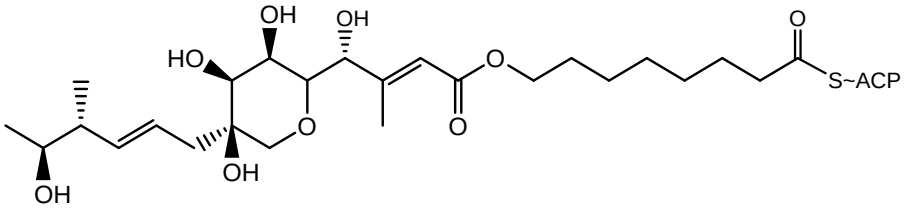


Figure 5.2: Substrates putatively accepted by the TmpB KS^0 domain.

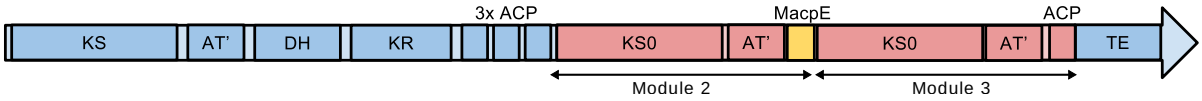
The KS^0 appears to have sufficiently relaxed substrate specificity to accept PA-B, although PA-B is fairly similar to its thiomarinol equivalent, marinolic acid B. The structure of marinolic acid B can be logically predicted from the proposed thiomarinol biosynthesis pathway (Figure 5.2). Marinolic acid B has C10,11-alkene, whereas PA-B has C10,11-epoxide, and the fatty acid chain is one carbon longer in PA-B. Marinolic acid B may also have C4-OH, although the timing of installation of this hydroxyl by TmuB is not yet known (Mohammad, 2017).

The TmpB KS^0 does not seems to demonstrate specificity towards the ACPs with which it interacts. In generation II hybrid JC58, the KS^0 is proposed to transfer intermediates from MmpB ACPs to the TmpB ACP that the KS^0 would natively load in thiomarinol biosynthesis. However, in generation III hybrids JC119 and JC120, the KS^0 may transfer from MmpB ACPs to MacpE, neither of which the KS^0 domain has co-evolved with. This KS^0 domain therefore demonstrates some promise as a useful tool for manipulating PKS pathways *in vivo*.

5.2.2.2 Ongoing and Future work on coupling NRPS biosynthesis to MmpB/TmpB hybrids

Now that conversion to PA-A has seemingly been achieved on MmpB/TmpB hybrids, the next step towards coupling to NRPS biosynthesis is to install a third module. We decided to generate an MmpB/TmpB hybrid generation IV, where module 1 synthesises 9-HN, module 2 is the site of C8-OH removal, and module 3 would later be the coupling point to NRPS biosynthesis (Figure 5.3). Hybrid generation IV was designed primarily by Amber Wilson[1], and re-used the thiomarinol KS^0-ACP module amino acid sequence, whilst taking advantage of degeneracy to change the DNA sequence.

(a) JC150



(b) JC151

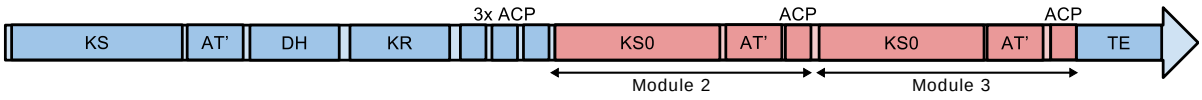


Figure 5.3: Domain overview of MmpB/TmpB hybrid generation IV.

Preliminary results with these hybrids have been promising, both are able to complement MmpB function, and conversion to PA-A was observed in 10586 Δ*macpE* with JC150. These new generation IV hybrids now need to be thoroughly tested as generations II and III were. If the model of competing rates of reaction between tailoring enzymes and the TE proposed in Section 4.3.2.1 is correct, then addition of the third module could increase the proportion of PA-A conversion before TE-catalysed release. A point mutation could be introduced to the KS^0 of either module 2 or 3. With an inactivated module 2 KS^0, the TE-skipping hypothesis (Figure 4.50) would predict a more severe decrease in PA-B release than seen in the generation II

[1]MmpB/TmpB hybrid generation IV was designed primarily by Amber Wilson, in collaboration with this author, Chris Thomas and Anthony Haines. Molecular construction was performed by both Amber Wilson and this author.

mutants.

Next a point mutation could be introduced to the active site serine of the module 3 ACP, and this would hopefully cause a decrease in PA-B release from the hybrid. If this mutant did generate a phenotype, then the module 3 ACP could be replaced with a variety of ACPs from other systems, and their functionality in the hybrids assayed by PA-B production. The ACPs used would be taken from PKS-NRPS junctions in nature, such as DptF from daptomycin (fatty acid-NRPS) (Wittmann et al., 2008). Of particular interest would be NRPS pathways that generate chromogenic products, such as indigoidine (Reverchon et al., 2002).

Significant challenges remain before a successful *in vivo* mupirocin-NRP is generated, such as the manner of release from hybrids with NRPS components. Constructions need to be designed in such a manner that the function, or lack of function, of any components added can be appraised. The NRPS C domain used will need to accept pseudomonic acid as a substrate, and be able to interact with its native ACP fused to MmpB/TmpB. It is also even possible that generation of a new derivative compound with alternate mode of action would be self-inhibitory to the 10586 host, which would lack evolved resistance. Generation of *in vivo* mupirocin-NRP derivatives remains an ambitious goal, although the work reported in this thesis demonstrates there is some scope for modifying mupirocin biosynthesis *in vivo*.

References

E. P. Abraham and E. Chain. An enzyme from bacteria able to destroy penicillin. *Reviews of infectious diseases*, 10(4):677–678, 1940.

A. W. Alberts, J. Chen, G. Kuron, V. Hunt, J. Huff, C. Hoffman, J. Rothrock, M. Lopez, H. Joshua, E. Harris, A. Patchett, R. Monaghan, S. Currie, E. Stapley, G. Albers-Schonberg, O. Hensens, J. Hirshfield, K. Hoogsteen, J. Liesch, and J. Springer. Mevinolin: a highly potent competitive inhibitor of hydroxymethylglutaryl-coenzyme A reductase and a cholesterol-lowering agent. *Proceedings of the National Academy of Sciences of the United States of America*, 77(7):3957–3961, 1980.

Y. Alsammarraie. *Expanding the potential of mutasynthetic approaches for pseudomonic acids*. PhD thesis, University of Birmingham, 2016.

E. Amann, J. Brosius, and M. Ptashne. Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in Escherichia coli. *Gene*, 25(2–3): 167–178, 1983.

V. T. Andriole. The Quinolones : Past, Present, and Future. *Clinical Infectious Diseases*, 41(Suppl 2):113–119, 2005.

T. Annaval, C. Paris, P. F. Leadlay, C. Jacob, and K. J. Weissman. Evaluating ketoreductase exchanges as a means of rationally altering polyketide stereochemistry. *ChemBioChem*, 16(9):1357–1364, 2015.

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

F. Baquero, J. Martínez-Beltrán, and E. Loza. A review of antibiotic resistance patterns of Streptococcus pneumoniae in Europe. *The Journal of antimicrobial chemotherapy*, 28:31–38, 1991.

C. Barna and D. H. Williams. The structure and mode of action of glycopeptide antibiotics of the vancomycin group. *Ann.Rev.Microbiol.*, 38:339–57, 1984.

S. M. Barry and G. L. Challis. Mechanism and Catalytic Diversity of Rieske Non-Heme Iron-Dependent Oxygenases. *ACS catalysis*, 3(10), 2013.

R. Bax and S. Green. Antibiotics: The changing regulatory and pharmaceutical industry paradigm. *Journal of Antimicrobial Chemotherapy*, 70(5):1281–1284, 2014.

B. G. Bell, F. Schellevis, E. Stobberingh, H. Goossens, and M. Pringle. A systematic review and meta-analysis of the effects of antibiotic consumption on antibiotic resistance. *BMC Infectious Diseases*, 14(1):13, 2014.

J. Bérdy. Thoughts and facts about antibiotics: Where we are now and where we are heading. *The Journal of Antibiotics*, 65(8):441–441, 2012.

S. Bergmann, A. N. Funk, K. Scherlach, V. Schroeckh, E. Shelest, U. Horn, C. Hertweck, and A. A. Brakhage. Activation of a silent fungal polyketide biosynthesis pathway through regulatory cross talk with a cryptic nonribosomal peptide synthetase gene cluster. *Applied and Environmental Microbiology*, 76(24):8143–8149, 2010.

D. J. Bevitt, J. Cortes, S. F. Haydock, and P. F. Leadlay. 6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme. *European Journal of Biochemistry*, 204(1):39–49, 1992.

X. Bian, A. Plaza, Y. Zhang, and R. Muller. Luminmycins A-C, cryptic natural products from Photorhabdus luminescens identified by heterologous expression in Escherichia coli. *Journal of Natural Products*, 75(9):1652–1655, 2012.

J. Blair, M. Webber, A. Baylay, D. Ogbolu, and L. Piddock. Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology*, 13(1):42–51, 2015.

K. Bloudoff, D. Rodionov, and T. M. Schmeing. Crystal structures of the first condensation domain of CDA synthetase suggest conformational changes during the synthetic cycle of nonribosomal peptide synthetases. *Journal of molecular biology*, 425(17):3137–50, 2013.

R. W. Broadhurst, D. Nietlispach, M. P. Wheatcroft, P. F. Leadlay, and K. J. Weissman. The structure of docking domains in modular polyketide synthases. *Chemistry and Biology*, 10(8):723–731, 2003.

A. F. Brodie and J. S. Gots. The reduction of tetrazolium salts by an isolated bacterial flavoproteins. *Science (New York, N.Y.)*, 116(3022):588–9, 1952.

E. D. Brown and G. D. Wright. Antibacterial drug discovery in the resistance era. *Nature*, 529(7586):336–343, 2016.

T. D. Bugg, G. D. Wright, S. Dutka-Malen, M. Arthur, P. Courvalin, and C. T. Walsh. Molecular basis for vancomycin resistance in Enterococcus faecium BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA. *Biochemistry*, 30(43):10408–10415, 1991.

P. Caffrey, S. Lynch, E. Flood, S. Finnan, and M. Oliynyk. Amphotericin biosynthesis in Streptomyces nodosus: Deductions from analysis of polyketide synthase and late genes. *Chemistry and Biology*, 8(7):713–723, 2001.

J. W. Campbell and J. E. Cronan Jr. Bacterial fatty acid biosynthesis: targets for antibacterial drug discovery. *Annu.Rev.Microbiol.*, 55(0066-4227):305–332, 2001.

J. O. Capobianco, C. C. Doran, and R. C. Goldman. Mechanism of mupirocin transport into sensitive and resistant bacteria. *Antimicrobial Agents and Chemotherapy*, 33 (2):156–163, 1989.

E. B. Chain and G. Mellows. Pseudomonic acid. Part 3. Structure of pseudomonic acid B. *Journal of the Chemical Society, Perkin Transactions 1*, 0(3):318–322, 1977.

V. L. Challinor and H. B. Bode. Bioactive natural products from novel microbial sources. *Annals of the New York Academy of Sciences*, 1354(1):82–97, 2015.

Y. A. Chan, A. M. Podevels, B. M. Kevany, and M. G. Thomas. Biosynthesis of polyketide synthase extender units. *Natural product reports*, 26(1):90–114, 2009.

H. Chen, M. Bjerknes, R. Kumar, and E. Jay. Determination of the optimal aligned spacing between the shine - dalgarno sequence and the translation initiation codon of escherichia coli m RNAs. *Nucleic Acids Research*, 22(23):4953–4957, 1994.

S. Chen, Y. Xue, D. H. Sherman, and K. A. Reynolds. Mechanisms of molecular recognition in the pikromycin polyketide synthase. *Chemistry and Biology*, 7(12): 907–918, 2000.

J. J. Ciardiello, W. R. J. D. Galloway, C. J. O'Connor, H. F. Sore, J. E. Stokes, Y. Wu, and D. R. Spring. An expedient strategy for the diversity-oriented synthesis of macrocyclic compounds with natural product-like characteristics. *Tetrahedron*, 72 (25):3567–3578, 2016.

J. Clardy, M. a. Fischbach, and C. T. Walsh. New antibiotics from bacterial natural products. *Nature biotechnology*, 24(12):1541–50, 2006.

S. Cooper, W. Laosripaiboon, A. Rahman, J. Hothersall, A. El-Sayed, C. Winfield, J. Crosby, R. Cox, T. Simpson, and C. Thomas. Shift to Pseudomonic Acid B Production in P. fluorescens NCIMB10586 by Mutation of Mupirocin Tailoring Genes mupO, mupU, mupV, and macpE. *Chemistry & Biology*, 12(7):825–833, 2005.

S. M. Cooper. *Auxiliary genes involved in the biosynthesis of mupirocin by Pseudomonas Fluorescens NCIMB10586*. PhD thesis, University of Birmingham, 2003.

J. Cortes, S. F. Haydock, G. A. Roberts, D. J. Bevitt, and P. F. Leadlay. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea. *Nature*, 348(6297):176–8, 1990.

G. M. Cragg and D. J. Newman. Natural products: A continuing source of novel drug leads. *Biochimica et Biophysica Acta - General Subjects*, 1830(6):3670–3695, 2013.

S. Donadio, M. J. Staver, J. B. Mcalpine, S. J. Swanson, and L. Katz. Modular organization of genes required for complex polyketide biosynthesis. *Science*, 252 (5006):675–679, 1991.

B. J. Dunn and C. Khosla. Engineering the acyltransferase substrate specificity of assembly line polyketide synthases. *Journal of the Royal Society, Interface / the Royal Society*, 10(85):20130297, 2013.

Z. D. Dunn, W. J. Wever, N. J. Economou, A. a. Bowers, and B. Li. Enzymatic Basis of "Hybridity" in Thiomarinol Biosynthesis. *Angewandte Chemie International Edition*, 54(17):5137–5141, 2015.

S. Dutta, J. R. Whicher, D. A. Hansen, W. A. Hale, J. A. Chemler, G. R. Congdon, A. R. H. Narayan, K. Håkansson, D. H. Sherman, J. L. Smith, and G. Skiniotis. Structure of a modular polyketide synthase. *Nature*, 510(7506):512–7, 2014.

A. K. El-Sayed, J. Hothersall, and C. M. Thomas. Quorum-sensing-dependent regulation of biosynthesis of the polyketide antibiotic mupirocin in Pseudomonas fluorescens NCIMB 10586. *Microbiology-Sgm*, 147:2127–2139, 2001.

A. K. El-Sayed, J. Hothersall, S. M. Cooper, E. Stephens, T. J. Simpson, and C. M. Thomas. Characterization of the Mupirocin Biosynthesis Gene Cluster from Pseudomonas fluorescens NCIMB 10586. *Chemistry & Biology*, 10(5):419–430, 2003.

A. M. Emmerson and A. M. Jones. The quinolones: decades of development and use. *The Journal of antimicrobial chemotherapy*, 51:13–20, 2003.

D. J. Ferraro, L. Gakhar, and S. Ramaswamy. Rieske business: structure-function of Rieske non-heme oxygenases. *Biochemical and biophysical research communications*, 338(1):175–90, 2005.

M. Finken, P. Kirschner, A. Meier, A. Wrede, and E. C. Böttger. Molecular basis of streptomycin resistance in Mycobacterium tuberculosis: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot. *Molecular microbiology*, 9(6):1239–1246, 1993.

A. Fleming. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. *Bulletin of the World Health Organization*, 79(8):780–790, 1929.

J. Frey, M. M. Bagdasarian, and M. Bagdasarian. Replication and copy number control of the broad-host-range plasmid RSF1010. *Gene*, 113(1):101–106, 1992.

S. Frykman, H. Tsuruta, J. Lau, R. Regentin, S. Ou, C. Reeves, J. Carney, D. Santi, and P. Licari. Modulation of epothilone analog production through media design, 2002.

D. Fukuda, A. S. Haines, Z. Song, A. C. Murphy, J. Hothersall, E. R. Stephens, R. Gurney, R. J. Cox, J. Crosby, C. L. Willis, T. J. Simpson, and C. M. Thomas. A natural plasmid uniquely encodes two biosynthetic pathways creating a potent anti-MRSA antibiotic. *PLoS One*, 6(3), 2011.

A. T. Fuller, G. Mellows, M. Woolford, G. T. Banks, K. D. Barrow, and E. B. Chain. Pseudomonic acid - an antibiotic produced by Pseudomonas fluorescens. *Nature*, 234(5329):416, 1971.

W. R. J. D. Galloway, A. Bender, M. Welch, D. R. Spring, C. Walsh, G. Wright, P. Fernandes, and E. Al. The discovery of antibacterial agents using diversity-oriented synthesis. *Chemical Communications*, 105(18):2446, 2009.

S.-S. Gao, J. Hothersall, J. Wu, A. C. Murphy, Z. Song, E. R. Stephens, C. M. Thomas, M. P. Crump, R. J. Cox, T. J. Simpson, and C. L. Willis. Biosynthesis of mupirocin by Pseudomonas fluorescens NCIMB 10586 involves parallel pathways. *Journal of the American Chemical Society*, 136(14):5501–7, 2014.

S.-S. Gao, L. Wang, Z. Song, J. Hothersall, E. R. Stevens, J. Connolly, P. J. Winn, R. J. Cox, M. P. Crump, P. R. Race, C. M. Thomas, T. J. Simpson, and C. L. Willis. Selected Mutations Reveal New Intermediates in the Biosynthesis of Mupirocin and the Thiomarinol Antibiotics. *Angewandte Chemie International Edition*, 129(14): 3988–3992, 2017.

N. M. Gaudelli and C. A. Townsend. Epimerization and substrate gating by a TE domain in β-lactam antibiotic biosynthesis. *Nature chemical biology*, 10(4):251–8, 2014.

D. C. Gay, G. Gay, A. J. Axelrod, M. Jenner, C. Kohlhaas, A. Kampa, N. J. Oldham, J. Piel, and A. T. Keatinge-Clay. A close look at a ketosynthase from a trans-acyltransferase modular polyketide synthase. *Structure (London, England : 1993)*, 22 (3):444–51, 2014.

P. Gay, D. L. Coq, M. Steinmetz, T. Berkelman, and C. I. Kado. Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. *Journal of*, 164(2):918–921, 1985.

J. E. Geraci, F. R. Heilman, D. R. Nichols, G. T. Ross, and W. E. Wellman. Some laboratory and clinical experiences with a new antibiotic, vancomycin. *Proc Staff Meet Mayo Clin*, 31(21):564–582, 1956.

GlaxoSmithKline. About us - pharmaceuticals http://www.gsk.com/en-gb/about-us/what-we-do/pharmaceuticals/ Accessed 01/07/16, 2016.

R. S. Gokhale, D. Hunziker, D. E. Cane, and C. Khosla. Mechanism and specificity of the terminal thioesterase domain from the erythromycin polyketide synthase. *Chemistry and Biology*, 6(2):117–125, 1999.

F. W. Goldstein, L. Gutmann, R. Williamson, E. Collatz, and J. F. Acar. In vivo and vitro emergence of simultaneous resistance to both beta-lactam and aminoglycoside antibiotics in a strain of Serratia marcescens. *Annales de l'Institut Pasteur. Microbiologie*, 134(3):329–337, 1983.

R. Gurney. *Biosynthesis of the antibiotic mupirocin by Pseudomonas fluorescens NCIMB 10586*. PhD thesis, University of Birmingham, 2012.

R. Gurney and C. M. Thomas. Mupirocin: biosynthesis, special features and applications of an antibiotic from a Gram-negative bacterium. *Applied Microbiology and Biotechnology*, 90(1):11–21, 2011.

A. S. Haines, X. Dong, Z. Song, R. Farmer, C. Williams, J. Hothersall, E. Płoskoń, P. Wattana-Amorn, E. R. Stephens, E. Yamada, R. Gurney, Y. Takebayashi, J. Masschelein, R. J. Cox, R. Lavigne, C. L. Willis, T. J. Simpson, J. Crosby, P. J. Winn,

C. M. Thomas, and M. P. Crump. A conserved motif flags acyl carrier proteins for β-branching in polyketide synthesis. *Nature chemical biology*, 9(11):685–92, 2013.

D. Hanahan. Studies on transformation of Escherichia coli with plasmids. *Journal of Molecular Biology*, 166(4):557–580, 1983.

D. Hanahan. Techniques for transformation of E. coli. In *DNA cloning: a practical approach. Vol. 1.*, pages 109–135. Oxford, United Kingdom: IRL Press, 1985. ISBN 9780947946197.

R. J. Heath and C. O. Rock. The Claisen condensation in biology. *Nat. Prod. Rep.*, 19 (5):581–596, 2002.

K. Hilpert, R. Volkmer-Engert, T. Walter, and R. E. W. Hancock. High-throughput generation of small antibacterial peptides with improved activity. *Nature biotechnology*, 23(8):1008–12, 2005.

Z. Hojati, C. Milne, B. Harvey, L. Gordon, M. Borg, F. Flett, B. Wilkinson, P. J. Sidebottom, B. A. M. Rudd, M. A. Hayes, C. P. Smith, and J. Micklefield. Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from Streptomyces coelicolor. *Chemistry and Biology*, 9(11):1175–1187, 2002.

M. E. Horsman, T. P. A. Hari, and C. N. Boddy. Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: logic gate or a victim of fate? *Nat. Prod. Rep.*, 33(2):183–202, 2016.

J. Hothersall, J. Wu, A. S. Rahman, J. A. Shields, J. Haddock, N. Johnson, S. M. Cooper, E. R. Stephens, R. J. Cox, J. Crosby, C. L. Willis, T. J. Simpson, and C. M. Thomas. Mutational analysis reveals that all tailoring region genes are required for production of polyketide antibiotic mupirocin by Pseudomonas fluorescens: Pseudomonic acid B biosynthesis precedes pseudomonic acid A. *Journal of Biological Chemistry*, 282(21):15451–15461, 2007.

J. Hothersall, A. C. Murphy, Z. Iqbal, G. Campbell, E. R. Stephens, J. Wu, H. Cooper, S. Atkinson, P. Williams, J. Crosby, C. L. Willis, R. J. Cox, T. J. Simpson, and C. M. Thomas. Manipulation of quorum sensing regulation in Pseudomonas fluorescens NCIMB 10586 to increase mupirocin production. *Applied Microbiology and Biotechnology*, 90(3):1017–1026, 2011.

J. Hughes and G. Mellows. Inhibition of isoleucyl-transfer ribonucleic acid synthetase in Escherichia coli by pseudomonic acid. *The Biochemical journal*, 176(1):305–318, 1978.

J. G. Hurdle, A. J. O'Neill, E. Ingham, C. Fishwick, and I. Chopra. Analysis of mupirocin resistance and fitness in Staphylococcus aureus by molecular genetic and structural modeling techniques. *Antimicrobial Agents and Chemotherapy*, 48(11): 4366–4376, 2004.

N. Ishiyama, C. Creuzenet, J. S. Lam, and A. M. Berghuis. Crystal structure of WbpP, a genuine UDP-N-acetylglucosamine 4-epimerase from Pseudomonas aeruginosa:

Substrate specificity in UDP-hexose 4-epimerases. *Journal of Biological Chemistry*, 279(21):22635–22642, 2004.

K. Jensen, H. Niederkrüger, K. Zimmermann, A. L. Vagstad, J. Moldenhauer, N. Brendel, S. Frank, P. Pöplau, C. Kohlhaas, C. A. Townsend, M. Oldiges, C. Hertweck, and J. Piel. Polyketide proofreading by an acyltransferase-like enzyme. *Chemistry and Biology*, 19(3):329–339, 2012.

M. Jevons. "Celbenin"- Resistant Staphylococci. *British Medical Journal*, 1(5219): 124–125, 1961.

M. Jin, M. A. Fischbach, and J. Clardy. A biosynthetic gene cluster for the acetyl-CoA carboxylase inhibitor andrimid. *Journal of the American Chemical Society*, 128(33): 10660–1, 2006.

K. L. Kavanagh, H. Jörnvall, B. Persson, and U. Oppermann. Medium- and short-chain dehydrogenase/reductase gene and protein families: The SDR superfamily: Functional and structural diversity within a family of metabolic and regulatory enzymes, 2008.

B. S. Kim, T. A. Cropp, B. J. Beck, D. H. Sherman, and K. A. Reynolds. Biochemical evidence for an editing role of thioesterase II in the biosynthesis of the polyketide pikromycin. *Journal of Biological Chemistry*, 277(50):48028–48034, 2002.

F. I. Kraas, T. W. Giessen, and M. A. Marahiel. Exploring the mechanism of lipid transfer during biosynthesis of the acidic lipopeptide antibiotic CDA. *FEBS letters*, 586(3):283–8, 2012.

L. Laureti, L. Song, S. Huang, C. Corre, P. Leblond, G. L. Challis, and B. Aigle. Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in Streptomyces ambofaciens. *Proceedings of the National Academy of Sciences of the United States of America*, 108(15):6258–6263, 2011.

R. Leclercq, E. Derlot, J. Duval, and P. Courvalin. Plasmid-Mediated Resistance to Vancomycin and Teicoplanin in Enterococcus Faecium. *New England Journal of Medicine*, 319(3):157–161, 1988.

G. Y. Lesher, E. J. Froelich, M. D. Gruett, J. H. Bailey, and R. P. Brundage. 1,8-Naphthyridine Derivatives. A New Class of Chemotherapeutic Agents. *Journal of Medicinal and Pharmaceutical Chemistry*, 5(5):1063–1065, 1962.

K. Lewis and P. Strandwitz. Microbiology: Antibiotics right under our nose. *Nature*, 535(7613):501–502, 2016.

L. L. Ling, T. Schneider, A. J. Peoples, A. L. Spoering, I. Engels, B. P. Conlon, A. Mueller, D. E. Hughes, S. Epstein, M. Jones, L. Lazarides, V. A. Steadman, D. R. Cohen, C. R. Felix, K. A. Fetterman, W. P. Millett, A. G. Nitti, A. M. Zullo, C. Chen, and K. Lewis. A new antibiotic kills pathogens without detectable resistance. *Nature*, 517(7535):455–459, 2015.

D. B. Lowe. Drug discovery: Combichem all over again. *Nature Chemistry*, 6(10): 851–852, 2014.

M. Macioszek. *Biosynthesis of mupirocin by Pseudomonas fluorescens NCIMB 10586*. PhD thesis, University of Birmingham, 2009.

N. A. Magarvey, P. D. Fortin, P. M. Thomas, N. L. Kelleher, and C. T. Walsh. Gatekeeping versus promiscuity in the early stages of the andrimid biosynthetic assembly line. *ACS chemical biology*, 3(9):542–54, 2008.

K. Magnuson, S. Jackowski, C. Rock, and J. E. Cronan. Regulation of Fatty Acid Biosynthesis in Escherichia coli. *Microbiological reviews*, 57(3):522–542, 1993.

T. Maier, M. Leibundgut, and N. Ban. The Crystal Structure of a Mammalian Fatty Acid Synthase. *Science*, 321(5894):1315–1322, 2008.

P. Mantle, M. De Langen, and V. K. Teo. Differentiating the Biosynthesis of Pseudomonic Acids A and B developed1 V. *The Journal of Antibiotics*, 54(2):166–174, 2001.

S. Matthijs, C. Vander Wauven, B. Cornu, L. Ye, P. Cornelis, C. M. Thomas, and M. Ongena. Antimicrobial properties of pseudomonas strains producing the antibiotic mupirocin. *Research in Microbiology*, 165(8):695–704, 2014.

M. H. Medema, K. Blin, P. Cimermancic, V. De Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano, and R. Breitling. AntiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Research*, 39(SUPPL. 2):W339–46, 2011.

V. Miao, M.-F. Coëffet-Legal, P. Brian, R. Brost, J. Penn, A. Whiting, S. Martin, R. Ford, I. Parr, M. Bouchard, C. J. Silva, S. K. Wrigley, and R. H. Baltz. Daptomycin biosynthesis in Streptomyces roseosporus: cloning and analysis of the gene cluster and revision of peptide stereochemistry. *Microbiology (Reading, England)*, 151(Pt 5): 1507–23, 2005.

V. Miao, M. F. Coeffet-Le Gal, K. Nguyen, P. Brian, J. Penn, A. Whiting, J. Steele, D. Kau, S. Martin, R. Ford, T. Gibson, M. Bouchard, S. K. Wrigley, and R. H. Baltz. Genetic Engineering in Streptomyces roseosporus to Produce Hybrid Lipopeptide Antibiotics. *Chemistry and Biology*, 13(3):269–276, 2006.

S. Midha, K. Bansal, S. Sharma, N. Kumar, P. P. Patil, V. Chaudhry, and P. B. Patil. Genomic resource of rice seed associated bacteria. *Frontiers in Microbiology*, 6(JAN): 1551, 2016.

R. C. Moellering. MRSA: the first half century. *Journal of Antimicrobial Chemotherapy*, 67(1):4–11, 2012.

H. H. Mohammad. *Studies on biosynthesis and activity of antibiotics thiomarinol from marine bacteria*. PhD thesis, University of Birmingham, 2017.

A. Moir, E. Lafferty, and D. Smith. Genetic analysis of spore germination mutants of Bacillus subtilis 168: the correlation of phenotype with map location. *Microbiology*, 111(165):180, 1979.

M. J. T. M. Mol, D. Willem Erkelens, J. A. Gevers Leuven, J. A. Schouten, and A. F. H. Stalenhoef. Simvastatin (MK-733): a potent cholesterol synthesis inhibitor in heterozygous familial hypercholesterolaemia. *Atherosclerosis*, 69(2-3):131–137, 1988.

T. Moriguchi, Y. Ebizuka, and I. Fujii. Analysis of Subunit Interactions in the Iterative Type I Polyketide Synthase ATX from Aspergillus terreus. *Chembiochem*, 7(12): 1869–1874, 2006.

A. C. Murphy, D. Fukuda, Z. Song, J. Hothersall, R. J. Cox, C. L. Willis, C. M. Thomas, and T. J. Simpson. Engineered Thiomarinol Antibiotics Active against MRSA Are Generated by Mutagenesis and Mutasynthesis of Pseudoalteromonas SANK73390. *Angewandte Chemie International Edition*, 50(14):3271–3274, 2011.

A. C. Murphy, S.-S. Gao, L.-C. Han, S. Carobene, D. Fukuda, Z. Song, J. Hothersall, R. J. Cox, J. Crosby, M. P. Crump, C. M. Thomas, C. L. Willis, and T. J. Simpson. Biosynthesis of thiomarinol A and related metabolites of Pseudoalteromonas sp. SANK 73390. *Chemical Science*, 5(1):397, 2014.

T. Nakama, O. Nureki, and S. Yokoyama. Structural Basis for the Recognition of Isoleucyl-Adenylate and an Antibiotic, Mupirocin, by Isoleucyl-tRNA Synthetase. *Journal of Biological Chemistry*, 276(50):47387–47393, 2001.

NHS. Prescription cost analysis http://www.content.digital.nhs.uk/catalogue/PUB23631 Accessed 13/9/17, 2016.

A. Okano, N. A. Isley, and D. L. Boger. Peripheral modifications of [Ψ[CH 2 NH]Tpg 4]vancomycin with added synergistic mechanisms of action provide durable and potent antibiotics. *Proceedings of the National Academy of Sciences*, page 201704125, 2017.

A. Omer-Bali. *Studies on key steps controlling biosynthesis of antibiotics thiomarinol and mupirocin*. PhD thesis, University of Birmingham, 2013.

M. S. Pacey, J. P. Dirlam, R. W. Geldart, P. F. Leadlay, H. A. McArthur, E. L. McCormick, R. A. Monday, T. N. O'Connell, J. Staunton, and T. J. Winchester. Novel erythromycins from a recombinant Saccharopolyspora erythraea strain NRRL 2338 pIG1. I. Fermentation, isolation and biological activity. *The Journal of antibiotics*, 51 (11):1029–34, 1998.

J. S. Parascandolo, J. Havemann, H. K. Potter, F. Huang, E. Riva, J. Connolly, I. Wilkening, L. Song, P. F. Leadlay, and M. Tosin. Insights into 6-Methylsalicylic Acid Bio-assembly by Using Chemical Probes. *Angewandte Chemie - International Edition*, 55(10):3463–3467, 2016.

L. J. V. Piddock. The crisis of no new antibiotics-what is the way forward? *The Lancet Infectious Diseases*, 12(3):249–253, 2012.

L. J. V. Piddock. Teixobactin, the first of a new class of antibiotics discovered by ichip technology? *Journal of Antimicrobial Chemotherapy*, 70(10):2679–2680, 2015.

J. Piel. Biosynthesis of polyketides by trans-AT polyketide synthases. *Natural Product Reports*, 27(7):996, 2010.

P. Power, T. Dunne, B. Murphy, L. N. Lochlainn, D. Rai, C. Borissow, B. Rawlings, and P. Caffrey. Engineered Synthesis of 7-Oxo- and 15-Deoxy-15-Oxo-Amphotericins: Insights into Structure-Activity Relationships in Polyene Antibiotics. *Chemistry and Biology*, 15(1):78–86, 2008.

A. S. Rahman, J. Hothersall, J. Crosby, T. J. Simpson, and C. M. Thomas. Tandemly duplicated acyl carrier proteins, which increase polyketide antibiotic production, can apparently function either in parallel or in series. *Journal of Biological Chemistry*, 280(8):6399–6408, 2005.

P. B. Rainey and M. J. Bailey. Physical and genetic map of the Pseudomonas fluorescens SBW25 chromosome. *Molecular microbiology*, 19(3):521–33, 1996.

S. Reverchon, C. Rouanet, D. Expert, and W. Nasser. Characterization of indigoidine biosynthetic genes in Erwinia chrysanthemi and role of this blue pigment in pathogenicity. *Journal of Bacteriology*, 184(3):654–665, 2002.

L. Robbel and M. A. Marahiel. Daptomycin, a bacterial lipopeptide synthesized by a nonribosomal machinery, 2010.

T. Robbins, J. Kapilivsky, D. E. Cane, and C. Khosla. Roles of Conserved Active Site Residues in the Ketosynthase Domain of an Assembly Line Polyketide Synthase. *Biochemistry*, 55(32):4476–4484, 2016.

G. G. Rodriguez, D. Phipps, and K. Ishiguro. Use of a Fluorescent Redox Probe for Direct Visualization of Actively Respiring Bacteria. *Applied and Environmental Microbiology*, 58(6):1801–1808, 1992.

R.P.Ambler. The structure of beta-lactamases. *Phil. Trans. R. Soc. Lond.*, B289(1036): 321, 1980.

P. Scholz, V. Haring, B. Wittmann-Liebold, K. Ashman, M. Bagdasarian, and E. Scherzinger. Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010. *Gene*, 75(2):271–288, 1989.

D. Schwarzer, R. Finking, and M. A. Marahiel. Nonribosomal peptides: from genes to products. *Natural product reports*, 20(3):275–287, 2003.

K. Sciarretta, J.-A. Røttingen, A. Opalska, A. J. Van Hengel, and J. Larsen. Economic Incentives for Antibacterial Drug Development: Literature Review and Considerations From the Transatlantic Task Force on Antimicrobial Resistance: Table 1. *Clinical Infectious Diseases*, 63(11):1470–1474, 2016.

R. W. Scott, A. C. Murphy, J. Wu, J. Hothersall, R. J. Cox, T. J. Simpson, C. M. Thomas, and C. L. Willis. Mupirocin F: structure elucidation, synthesis and rearrangements. *Tetrahedron*, 67(27-28):5098–5106, 2011.

I. B. Seiple, Z. Zhang, P. Jakubec, A. Langlois-Mercier, P. M. Wright, D. T. Hog, K. Yabu, S. R. Allu, T. Fukuzaki, P. N. Carlsen, Y. Kitamura, X. Zhou, M. L. Condakes, F. T. Szczypiński, W. D. Green, and A. G. Myers. A platform for the discovery of new macrolide antibiotics. *Nature*, 533(7603):338–345, 2016.

K. K. Sharma and C. N. Boddy. The thioesterase domain from the pimaricin and erythromycin biosynthetic pathways can catalyze hydrolysis of simple thioester substrates. *Bioorganic and Medicinal Chemistry Letters*, 17(11):3034–3037, 2007.

K. J. Shaw, P. N. Rather, R. S. Hare, and G. H. Miller. Molecular genetics of aminoglycoside resistance genes and familial relationships of the aminoglycoside-modifying enzymes. *Microbiological reviews*, 57(1):138–163, 1993.

B. Shen, L. Du, C. Sanchez, D. J. Edwards, M. Chen, and J. M. Murrell. The biosynthetic gene cluster for the anticancer drug bleomycin from Streptomyces verticillus ATCC15003 as a model for hybrid peptide-polyketide natural product biosynthesis. *Journal of Industrial Microbiology and Biotechnology*, 27(6):378–385, 2001.

B. Shen, L. Du, C. Sanchez, D. J. Edwards, M. Chen, and J. M. Murrell. Cloning and Characterization of the Bleomycin Biosynthetic Gene Cluster from Streptomyces v erticillus ATCC15003 1. *Journal of Natural Products*, 65(3):422–431, 2002.

J. A. Shields. *Biosynthesis of the polyketide antibiotic mupirocin by Pseudomonas fluorescens NCIMB10586*. PhD thesis, Unversity of Birmingham, 2008.

J. A. Shields, A. S. Rahman, C. J. Arthur, J. Crosby, J. Hothersall, T. J. Simpson, and C. M. Thomas. Phosphopantetheinylation and Specificity of Acyl Carrier Proteins in the Mupirocin Biosynthetic Cluster. *Chembiochem*, 11(2):248–255, 2010.

H. Shiozawa, T. Kagasaki, T. Kinoshita, H. Haruyama, H. Domon, Y. Utsui, K. Kodama, and S. Takahashi. Thiomarinol, a new hybrid antimicrobial antibiotic produced by a marine bacterium - fermentation, isolation, structure, and antimicrobial activity. *Journal of Antibiotics*, 46(12):1834–1842, 1993.

L. F. Silvian, J. M. Wang, and T. A. Steitz. Insights into Editing from an Ile-tRNA Synthetase Structure with tRNAIle and Mupirocin. *Science*, 285(5430):1074–1077, 1999.

R. Simon, U. Priefer, and A. Pühler. A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Bio/Technology*, 1(9):784–791, 1983.

S. B. Singh and F. Pelaez. Biodiversity, chemical diversity and drug discovery, 2008.

M. S. Sinha and A. S. Kesselheim. Regulatory Incentives for Antibiotic Drug Development: A Review of Recent Proposals. *Bioorganic and Medicinal Chemistry*, 24(24):6446–6451, 2016.

S. Smith, A. Witkowski, and A. K. Joshi. Structural and functional organization of the animal fatty acid synthase, 2003.

W. Sneader. The discovery of aspirin. *Pharmaceutical Journal*, 259(6964):614–617, 1997.

B. S. Speer, N. B. Shoemaker, and A. A. Salyers. Bacterial resistance to tetracycline: Mechanisms, transfer, and clinical significance, 1992.

J. Staunton and K. J. Weissman. Polyketide biosynthesis: a millennium review. *Natural Product Reports*, 18(4):380–416, 2001.

J. Staunton and B. Wilkinson. Biosynthesis of Erythromycin and Rapamycin. *Chemical reviews*, 97(7):2611–2630, 1997.

Y. Sun, T. Tomura, J. Sato, T. Iizuka, R. Fudou, and M. Ojika. Isolation and Biosynthetic Analysis of Haliamide, a New PKS-NRPS Hybrid Metabolite from the Marine Myxobacterium Haliangium ochraceum. *Molecules (Basel, Switzerland)*, 21(1): 59, 2016.

F. C. Tenover. Mechanisms of antimicrobial resistance in bacteria. *The American journal of medicine*, 119(6 Suppl 1):S3–10; discussion S62–70, 2006.

C. M. Thomas, J. Hothersall, C. L. Willis, and T. J. Simpson. Resistance to and synthesis of the antibiotic mupirocin. *Nature Reviews Microbiology*, 8(4):281–289, 2010.

P. Thomas, A. C. Sekhar, R. Upreti, M. M. Mujawar, and S. S. Pasha. Optimization of single plate-serial dilution spotting (SP-SDS) with sample anchoring as an assured method for bacterial and yeast cfu enumeration and single colony isolation from diverse samples. *Biotechnology Reports*, 8:45–55, 2015.

J. Trias. The role of combichem in antibiotic discovery. *Current Opinion in Microbiology*, 4(5):520–525, 2001.

Y. Tu. The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. *Nature Medicine*, 17(10):1217–1220, 2011.

J. Vara, M. Lewandowska-Skarbek, Y. G. Wang, S. Donadio, and C. R. Hutchinson. Cloning of genes governing the deoxysugar portion of the erythromycin biosynthesis pathway in Saccharopolyspora erythraea (Streptomyces erythreus). *Journal of Bacteriology*, 171(11):5872–5881, 1989.

F. Von Nussbaum and R. D. Süssmuth. Multiple attack on bacteria by the new antibiotic teixobactin. *Angewandte Chemie - International Edition*, 54(23):6684–6686, 2015.

S. a. Waksman. Streptomycin: background, isolation, properties, and utilization. *Science (New York, N.Y.)*, 118(3062):259–266, 1953.

C. T. Walsh and T. A. Wencewicz. Prospects for new antibiotics: a molecule-centered perspective. *The Journal of Antibiotics*, 67(1):7–22, 2014.

H. Wang, D. P. Fewer, L. Holm, L. Rouhiainen, and K. Sivonen. Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. *Proceedings of the National Academy of Sciences of the United States of America*, 111(25):9259–64, 2014.

T. Weber, P. Charusanti, E. M. Musiol-Kroll, X. Jiang, Y. Tong, H. U. Kim, and S. Y. Lee. Metabolic engineering of antibiotic factories: New tools for antibiotic production in actinomycetes. *Trends in Biotechnology*, 33(1):15–26, 2015.

K. J. Weissman. The structural basis for docking in modular polyketide biosynthesis. *ChemBioChem*, 7(3):485–494, 2006.

K. J. Weissman. The structural biology of biosynthetic megaenzymes. *Nature Chemical Biology*, 11(9):660–670, 2015.

K. J. Weissman. Genetic engineering of modular PKSs: from combinatorial biosynthesis to synthetic biology. *Nat. Prod. Rep.*, 33(2):203–230, 2016.

C. A. Whatling, J. E. Hodgson, M. K. R. Burnham, N. J. Clarke, F. Christopher, H. Franklin, and C. M. Thomas. Identification of a 60 kb region of the chromosome of Pseudomonas fluorescens NClB 10586 required for the biosynthesis of pseudomonic acid (mupirocin). *Microbiology*, 141:973–982, 1995.

M. Wittmann, U. Linne, V. Pohlmann, and M. A. Marahiel. Role of DptE and DptF in the lipidation reaction of daptomycin. *FEBS Journal*, 275(21):5343–5354, 2008.

R. B. Woodward, E. Logusch, K. P. Nambiar, K. Sakan, D. E. Ward, B. W. Au-Yeung, P. Balaram, L. J. Browne, P. J. Card, and C. H. Chen. Asymmetric total synthesis of erythromycin. 3. Total synthesis of erythromycin. *Journal of the American Chemical Society*, 103(11):3215–3217, 1981.

G. Wright. An irresistible newcomer. *Nature*, 517(7535):442–443, 2015.

P. M. Wright, I. B. Seiple, and A. G. Myers. The evolving role of chemical synthesis in antibacterial drug discovery. *Angewandte Chemie (International ed. in English)*, 53 (34):8840–69, 2014.

J. Wu, J. Hothersall, C. Mazzetti, Y. O'Connell, J. A. Shields, A. S. Rahman, R. J. Cox, J. Crosby, T. J. Simpson, C. M. Thomas, and C. L. Willis. In vivo Mutational Analysis of the Mupirocin Gene Cluster Reveals Labile Points in the Biosynthetic Pathway: the "Leaky Hosepipe" Mechanism. *Chembiochem*, 9(9):1500–1508, 2008.

Y. Xue, L. Zhao, H. W. Liu, and D. H. Sherman. A gene cluster for macrolide antibiotic biosynthesis in Streptomyces venezuelae: architecture of metabolic diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 95(21): 12111–6, 1998.

M. Yadav. *Investigations on the biosynthesis of antibiotics mupirocin and thiomarinol*. PhD thesis, University of Birmingham, 2017.

T. Yanagisawa and M. Kawakami. How does Pseudomonas fluorescens avoid suicide from its antibiotic pseudomonic acid? Evidence for two evolutionarily distinct isoleucyl-tRNA synthetases conferring self-defense. *Journal of Biological Chemistry*, 278(28):25887–25894, 2003.

C. Yanisch-Perron, J. Vieira, and J. Messing. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. *Gene*, 33(1):103–119, 1985.

J. S. Zarins-Tutt, T. T. Barberi, H. Gao, A. Mearns-Spragg, L. Zhang, D. J. Newman, and R. J. M. Goss. Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products. *Nat. Prod. Rep.*, 33(1):54–72, 2016.

J. Zeng, D. T. Wagner, Z. Zhang, L. Moretto, J. D. Addison, and A. T. Keatinge-Clay. Portability and Structure of the Four-Helix Bundle Docking Domains of trans-Acyltransferase Modular Polyketide Synthases. *ACS Chemical Biology*, 11(9): 2466–2474, 2016.

J. Zheng, C. D. Fage, B. Demeler, D. W. Hoffman, and A. T. Keatinge-Clay. The missing linker: A dimerization motif located within polyketide synthase modules. *ACS Chemical Biology*, 8(6):1263–1270, 2013.

A. Zipperer, M. C. Konnerth, C. Laux, A. Berscheid, D. Janek, C. Weidenmaier, M. Burian, N. A. Schilling, C. Slavetinsky, M. Marschal, M. Willmann, and H. Kalbacher. Human commensals producing a novel antibiotic impair pathogen colonization. *Nature*, 535(7613):511–516, 2016.

M. M. Zukowski, D. F. Gaffney, D. Speck, M. Kauffmann, A. Findeli, A. Wisecup, and J. P. Lecocq. Chromogenic identification of genetic regulatory signals in Bacillus subtilis based on expression of a cloned Pseudomonas gene. *Proceedings of the National Academy of Sciences of the United States of America*, 80(4):1101–5, 1983.