STUDIES ON BIOSYNTHESIS AND ACTIVITY OF ANTIBIOTIC THIOMARINOL FROM MARINE BACTERIA.

by

HADI HUSSEIN MOHAMMAD

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

School of Biosciences
College of Life and Environmental Sciences
The University of Birmingham
September 2016

ABSTRACT

Mupirocin (Pseudomonic acid A) has long been used against Methicillin Resistant *Staphylococcus aureus* MRSA, yet bacteria have developed resistance, threatening future use. Structurally similar to mupirocin is thiomarinol A, a natural compound produced by the marine bacterium *Pseudoalteromonas* spp, which possesses stronger antibacterial activities. However, it differs from mupirocin by four distinct differences and among these are extra 4-hydroxylation and joining to pyrrothine. Studying these differences should enhance our understanding of the molecular assembly and biosynthesis machinery.

Complementation and mutagenesis studies identified the *tmuB* gene to be responsible for the 4-hydroxylation as a final tailoring step. *In vivo* and *in vitro* studies on purified TmuB revealed that it can hydroxylate diverse pseudomonic acids but is inhibited by molecules with an 8-hydroxyl group, which primarily affects catalysis rather than binding. Molecular modelling plus docking and mutagenesis provides increased understanding of both TmuB potential to modify other substrates and how mupirocin activity can be modulated by 4-hydroxylation. This study also expressed *holA*, purified its gene product, a non-ribosomal polypeptide synthetase (NRPS), and assayed its activity by pyrophosphate release. It presents a proposed pathway for pyrrothine biosynthesis catalysed by HolA, which exhibits the unusual ability to join two cysteine molecules by a single NRPS module.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. Christopher Thomas for his unlimited support, guidance and knowledge within the four years on my PhD. Many thanks to my co-supervisor, Dr. Joanne Hothersall, for her advices, support and assistance whenever required. I would like to express my greatest gratitude to The High Committee For Education Development (HCED) and Ministry Of Higher Education & Scientific Research in Iraq to support my PhD fund.

I wish to acknowledge all the past and present members of S101, particularly Anthony Haines for his help whenever I was stuck in the genetic manipulation issues; Jack Connolly to discuss many relevant scientific aspects; Elton who help me to settle in the lab within the first few months in my study and being a good source of technical advices; Nick Cotton to make me familiar with gel filtration columns till we got our own and Peter Winn for his assistance and guidance to do the bioinformatics analyses.

I wish to thank both Prof. Chris Willis and Dr. Zhongshu Song from the School of Chemistry, University of Bristol for their collaborations and performing the MS and NMR for all the compounds in this study.

Finally, words are unable to express my gratitude to my wife for her supports and sacrifices; my parents, families and friends for their encouragement to travel abroad to do my study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

3-HP	3-hydroxypropionate
9-HN	9-Hydroxynonanoic acid
A	Adenylation domain
ACP	acyl carrier protein
ADME	Absorption, distribution, metabolism and excretion
αKG	Alpha ketoglutarate
Amp^R	Ampicillin resistance
APS	Ammonium persulphate
AT	Acyltransferase
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
C	Condensation domain
CDA	Calcium dependent antibiotic
CDC	The Centres for Disease Control and Prevention
CoA	CoenzymeA
CFU	Colony forming unit
DEBS	6-deoxyerythronolide B synthase
DH	Dehydratase
DHT	Dihydropteroic
DMSO	Dimethyl sulfoxide
DNA	Deoxybribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DSBH	Double strand beta helix
DTT	Dithiothreitol
ECDC	The European Centre for Disease Prevention and Control
EDTA	Ethylenediamine-tetra-acetic-acid
ER/OR	Enoyl reductase/Oxidoreductase
ESAC	The European Surveillance Antimicrobial Consumption
FAS	Fatty acid synthase
HCS	Hydroxymethylglutaryl-CoA synthase
HEPES	4-(2-hydroxyethyl)piperazin-1-ethanesulfonic acid
HPLC	High performance liquid chromatography
IleS	Isoleucyl-tRNA synthase
IPTG	Isopropyl-β-D-thiogalactoside
KanS	Kanamycin sensitive
Kcat	Catalytic constant
Km	Michaelis constant
KR	Ketoreductase
KS	Ketosynthase
L agar	Luria-Bertani agar
L broth	Luria-Bertani broth
LCMS	Liquid chromatography mass spectroscopy

LDD	Loading didomain
MA	Monic acid
mAcp	Type II acyl carrier protein
MDR TB	Multidrug-resistant TB
MIC	Minimum inhibitory concentration
Mmp	Mupirocin multifunctional protein
MPT	malonyl/palmitoyl transferase
MRSA	Methicillin-resistant *Staphylococcus aureus*
MS	Mass spectrometry
MT/ME	Methyltransferase
NCIMB	The National Collection of Industrial, food and Marine Bacteria
NEB	New England Biolabs
Ni-NTA	Nickel-nitriloacetic acid
NMR	Nuclear Magnetic Resonance
NRPS	Non-ribosomal peptide synthase
OD	optical density
ORF	Open reading frame
PA	Pseudomonic acid
PABA	p-aminobenzoic acid
PBP	Penicillin binding protein
PCP	Peptidyl carrier protein
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PKS	Polyketide synthase
PM	Point mutation
PPase	Inorganic pyrophosphatase
PPTase	Phosphopantetheinyl transferase
PSIPRED	Protein Structure Prediction Server
RNA	Ribonucleic acid
RT	Retention time
SAM	S-adenosyl methionine
SANK	Daiichi Sankyo
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SSM	Secondary stage medium
Suc	Sucrose
TAE	Tris-acetate EDTA
TCC	2, 3, 5-triphenyltetrazolium chloride
TE	Thioesterase
TEMED	N, N, N’, N’-tetremethylethylene diamine
Temp	Temperature
TetR	Tetracycline resistance
THF	Tetrahydrofolic acid

Tml	Thiomarinol
Tmp	Thiomarinol multifunctional protein
Tris	tris(hydroxymethyl)aminomethane
Tri-X	Triton-X
UV	Ultraviolet
Vmax	Maximum velocity
WHO	The World Health Organisation
WT	Wild type

CHAPTER ONE

1. INTRODUCTION

1.1. Overview of antibiotics and antibiotic resistance

Antibiotics (*anti: against, bio: life*) are chemical metabolites produced by microorganisms such as bacteria and fungi, and have the ability to kill or inhibit the growth of other microorganisms at low concentration. The noun "*Antibiotic*" was first suggested by soil microbiologist Selman A. Waksman in 1942 (Waksman, 1947).

Before antibiotic development, infectious diseases were the major cause of death. This changed after 1929 when the British microbiologist Alexander Fleming discovered the first antibiotic by detecting the inhibition of the *Staphylococcus aureus* growth around a *Penicillium notatum* mold colony (Torok *et al.,* 2009). However, although Fleming named this active substance penicillin, because of technical difficulties he did not extract it for human administration. Fortunately, about one decade later, Ernst Chain and Howard Florey developed a method to isolate penicillin and used it during the Second World War. This discovery changed the history of medicine and provided the inspiration to others to develop many types of antibiotics. Thus, the Noble Prize in medicine was awarded to Fleming, Chain and Florey in 1945 (Lloyd, 2009).

Since the invention of the first antibiotic, public health has improved dramatically and average life expectancy has witnessed a significant rise (Todar, 2008). However, this improvement did not persist for a long time as a new problem developed which may render antibiotics to lose their efficacy. The first prediction about antibiotic resistance was announced in his 1945 Nobel Prize lecture when Fleming said:

> *"It is not difficult to make microbes resistant to penicillin in the laboratory by exposing them to concentrations not sufficient to kill them, and the same thing has occasionally happened in the body… …and by exposing microbes to non-lethal quantities of the drug make them resistant."*(The Nobel Lecture, 1945).

This prediction was correct when the first antibiotic resistance developed by *S. aureus* against *penicillin* was observed in 1947. During several decades, antibiotic resistance increased steadily and nowadays seventy percent of nosocomial infections are resistant to at least one most commonly used antibiotic (Todar, 2008). The well-known bacteria that are resistant to a wide range of antibiotics and cause high mortality and morbidity are methicillin resistant *Staphylococcus aureus* (MRSA) and Multidrug Resistant Tuberculosis (MDR-TB). According to the Office for National Statistic data in England and Wales, 12% of all deaths in 1993 were due to MRSA and this rate rose significantly to 66% in 2002, although it had fallen back to 16% in 2014 (Office for national statistic, 2004 & 2015). Regarding MDR-TB, a WHO report (2011) points out that this disease causes about one hundred thousand deaths annually. In addition to health risk, the economic impact is another reason to make this problem serious and globally concerning. For example, in the United States, Institution of Medicine report (1998) shows that diseases caused by antibiotic resistant bacteria cost about four billion US dollars annually. Therefore, the antibiotic resistance problem is considered as an issue of global concern and many strategies have been suggested to combat it.

1.1.1. Antibiotic classification

There are several schemes to classify antibiotics into groups according to the type of activity, potency spectrum, chemical structure and the mode of action. Regarding the type of activity, antibiotics can be divided into bactericidals, which kill bacteria, and bacteriostatics, which inhibit bacterial growth. The potency scheme divides antibiotics into the broad spectrum (against a wide range of bacteria) and the narrow spectrum (against few types of bacteria) (Madigan *et al.*, 2012). Based on the chemical structure, antibiotics with similar structure can be classified into groups, for example, β-lactams, macrolides, aminoglycosides, quinolones, tetracyclines (Figure 1.1). Cell wall synthesis, cytoplasmic membrane structure and function, protein synthesis, folic acid metabolism and nucleic acid (DNA and RNA) synthesis are the main targets in the cell that classified antibiotics into groups (Figure 1.2).

The antibiotics which target the bacterial cell wall are highly selective as the cell wall is a unique feature for bacteria. A group of antibiotics targets this part of the cell using different mechanisms to inhibit or disrupt the cell wall. For example, β-lactams antibiotics inhibit cell wall synthesis by binding and reacting with transpeptidase enzymes (penicillin binding proteins) which are responsible for cross-linking of two glycan-linked peptide chains. Vancomycin inhibits cell wall synthesis by blocking transpeptidation via binding to the terminal D-Ala-D-Ala peptide (Madigan *et al*., 2012). Polymyxins act on the bacterial outer membrane and bind with lipid A moiety of the Lipopolysaccharide (LPS) causing destabilization and disruption of the cell membrane (Falagas *et al*., 2010).

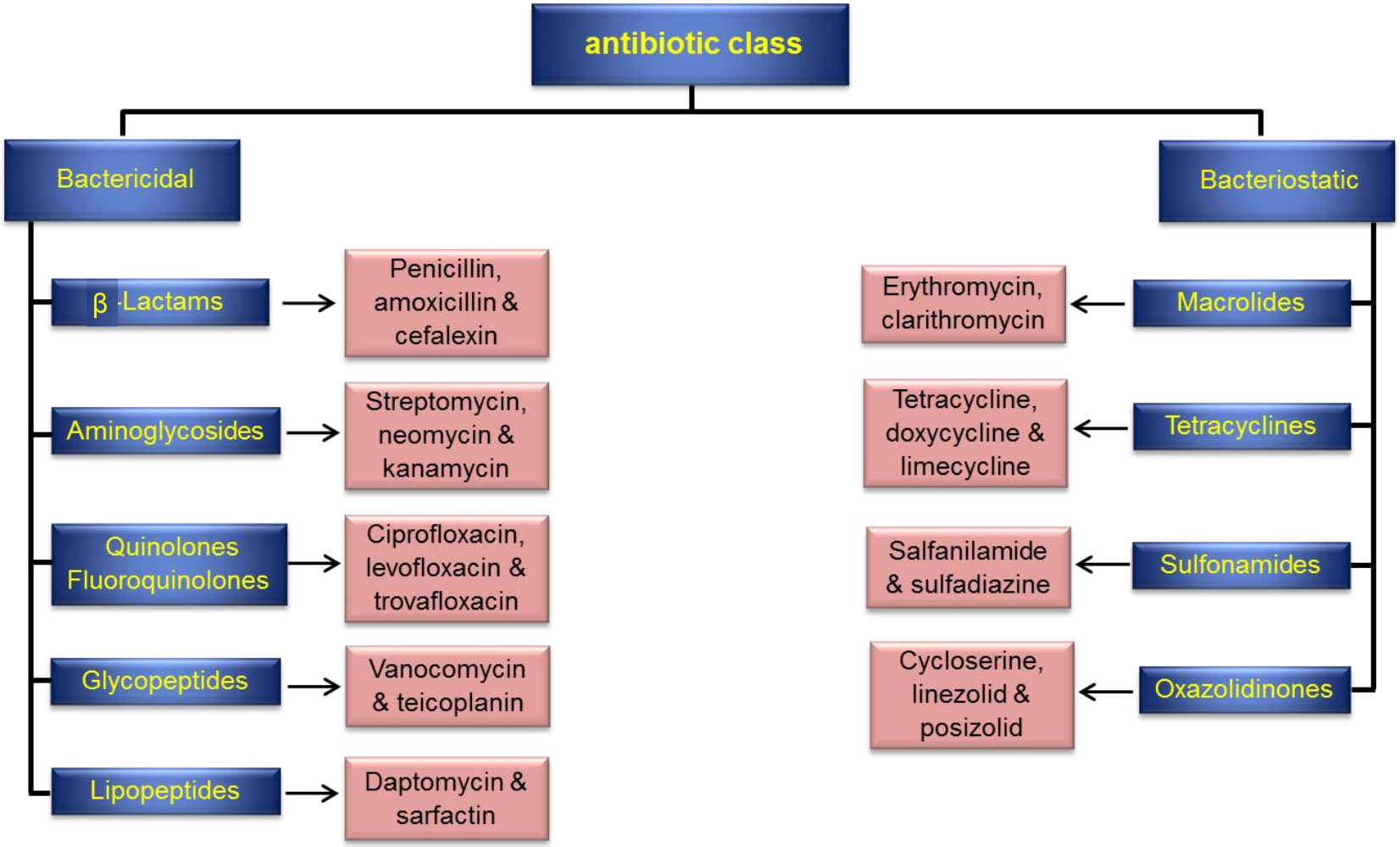


Figure 1.1. Classification of antibiotics according to type of activity and chemical structure. The type of activity divides antibiotic into two groups: bactericidal and bacteriostatic. These can be further divided into subgroup depending on chemical structures.

Daptomycin, a cyclic lipopeptide, targets cytoplasmic membrane, inserts into and forms pores in the lipidbilayer causing rapid depolarization of the membrane (Robbel & Maraheil, 2010). The antibiotics that target protein synthesis can be divided into three subgroups and within the same subgroup, the mechanism of action for each antibiotic is different. Erythromycin, clindamycin and chloramphenicol inhibit protein synthesis by interaction with the 50S ribosomal subunit while gentamycin, kanamycin and tetracycline interact with the 30S ribosomal subunit. Puromycin and mupirocin inhibit protein synthesis in different ways from these two subgroups. Puromycin occupies the A site on the ribosome and therefore the growing chain of the peptide binds to puromycin instead of aminoacylated tRNA causing premature chain termination (Madigan *et al*., 2012). Mupirocin inhibits protein synthesis by binding specifically to isoleucyl-tRNA synthetase and prevents incorporation of isoleucine amino acids into bacterial proteins (Hughes & Mellows, 1978; Eriani *et al.*, 1990). Folic acid in bacteria is produced from p-aminobenzoic acid (PABA) via two enzymatic reactions. Tetrahydrofolic acid (THF) formation is essential for thymidine synthesis in bacteria and inhibition of THF formation means inhibition of bacterial DNA synthesis (Brown, 1962). Both trimethoprim and sulfonamides interfere with folic acid metabolism. Sulfonamides inhibit dihydropteroate synthetase, an enzyme that converts PABA to dihydropteroic (DHT), while trimethoprim inhibits the formation of THF from DHT by binding with the enzyme catalyzing this reaction (Brown, 1962 and Brogden *et al.*, 1982). Quinolones such as ciprofloxacin and nalidixic acid inhibit bacterial DNA gyrase while rifampin affects transcription by binding specifically to the bacterial β-subunit of RNA polymerase (Lebel, 1988 and Madigan *et*

al., 2012). Understanding the mode of action of antibiotics helps us to understand the mechanism of antibiotic resistance and how bacteria acquired resistance.

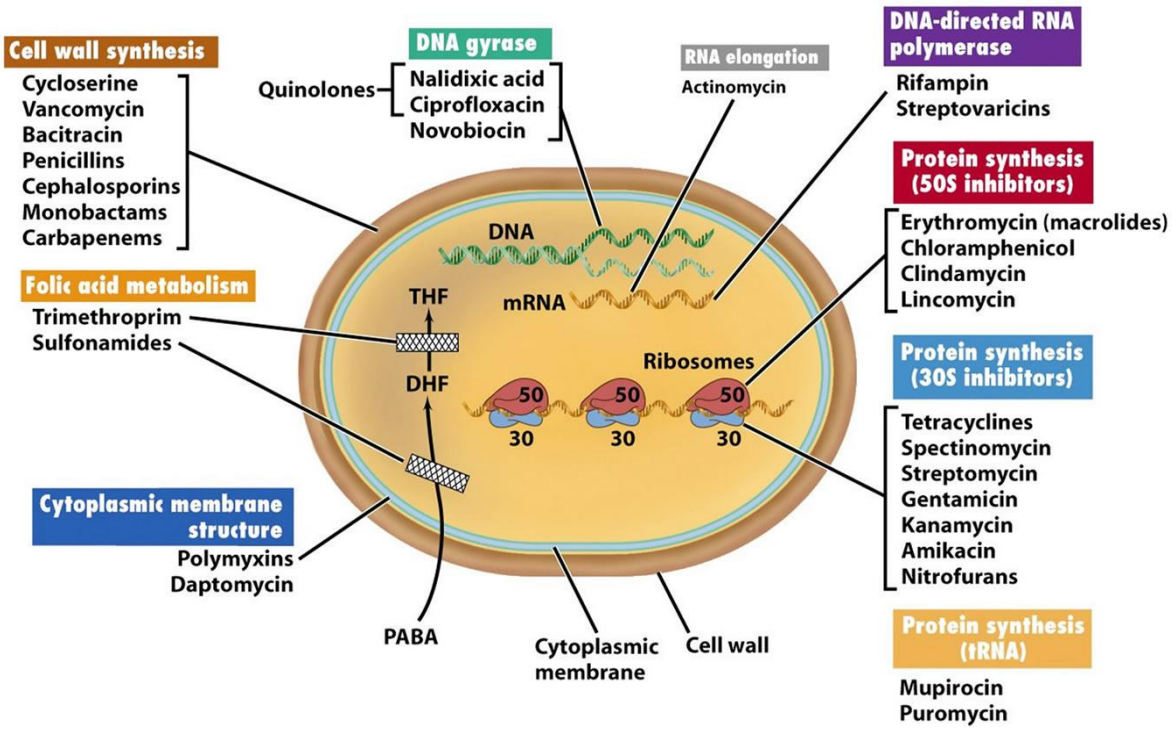


Figure 1.2. Classification of antibiotics according to the target. Each group of antibiotics targets a particular component of the cell. For example, cell wall collects penicillins and cephalosporins in one group (Figure taken from Madigan *et al.*, 2012 after permission).

1.1.2. Mechanisms of antibiotic resistance

Microbes that produce antibiotics must naturally be resistant to the molecule(s) they produce while other microbes acquire resistance by a combination of genetic processes and selective pressure. The origin of the genes that play a role in the resistance could

be chromosomal genes and mutation or they could be acquired by transferrable elements like plasmids. Bacteria use the following mechanisms to avoid being affected by antibiotic activity.

Antibiotic modification and inactivation: resistance to penicillin and cephalosporin is often due to production of β-lactamase enzymes that cleave the essential β-lactam ring rendering the antibiotic inactive. Resistance to chloramphenicol and aminoglycosides is due to R plasmids which encode enzymes that modify the antibiotic molecules through phosphorylation, acetylation and adenylation (Mingeot –Leclercq *et al*., 1999 & Madigan *et al.*, 2012).

Modification of antibiotic targets: mutations in the genes, generally chromosomal, modify the binding site or active site geometry so that the antibiotic no longer inhibits the enzyme. For example, resistance to rifampin can be due to mutation of the RNA polymerase β-subunit (Wehrli, 1983). Modification of lipid A moiety of the Lipopolysaccharide (LPS) in gram-negative bacteria leads to polymyxin resistance (Olaitan *et al.*, 2014).

Alternative biochemical pathway: some bacteria depend on exogenous precursors to build up the essential molecules. For example, a mutation in the thymidylate synthetase, which catalyzes THF conversion to thymine (see above), results in thymine requiring bacteria (thy^-) that depend on an exogenous source of thymine, and such bacteria are resistance to trimethoprim (Brogden *et al.*, 1982). Bacteria that are resistant to

sulfonamides develop a resistant biochemical pathway that depends on exogenous folic acid from the environment (Madigan *et al.*, 2012)

Efflux pumps: Most bacteria use various transport proteins to extrude harmful substances, including antibiotics, out of the cell. These pump systems can transport either specific single molecules or a wide range of antibiotics leading to multidrug resistance (MDR). Resistance to fluoroquinolones due to this mechanism has been reported in a number of clinically important bacteria such as *Staphylococcus aureus*, *Streptococcus pneumoniae* and *Escherichia coli* (Webber & Piddock, 2003 and Sun *et al.*, 2014).

Cell permeability and antibiotic uptake: Bacteria use different techniques to reduce antibiotic uptake from the surrounding environment. Gram-negative bacteria develop permeability barriers to resist β-lactam antibiotics. Some mutations lead to modifications of lipid-mediated or porin-mediated pathways of the outer membrane permeability which reduce antibiotic uptake. Hydrophobic antibiotics such as kanamycin, gentamycin and erythromycin enter the cell through lipid-mediated pathways and modification of this results in resistance against these drugs. Some clinically important pathogens such as *Escherichia coli, Neisseria gonorrhoeae*, *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* develop resistance against hydrophilic antibiotics such as β-lactams and chloramphenicol via porins modification. Other bacteria change proton motive forces to reduce the inner- membrane permeability (Delcour, 2009 and Torok *et al.,* 2009)

1.1.3. Tackling antibiotic resistance

The wrong administration policy of antibiotics is the most important factor that predisposes to antibiotic resistance. Inappropriate and overuse of antibiotics in medicine, agriculture and veterinary medicine are the leading causes for development of antibiotics resistance. In the United States, in one year more than twelve tons had been prescribed for humans and about eighteen tons in agriculture (Torok *et al*., 2009; Todar, 2008). However, this may be different between developed and developing countries (Torok *et al*., 2009). Regarding European countries, the European Surveillance Antimicrobial Consumption (ESAC) report (2009) shows that since 2005, the antimicrobial usage has risen. This may be evidence that antimicrobials still perform a mammoth task in fighting infectious diseases. However, the increasing consumption of antibiotics may be the main cause of antibiotic resistance, a serious problem that will face the future use of antibiotics. In response to the steady increase of antibiotic resistant bacteria and its impact both on health and economic sectors, many governments, health organizations and scientific institutions have taken steps to tackle antibiotic resistant bacteria and many strategies have been suggested. Many well-known organizations such as the World Health Organization (WHO) and Centers for Disease Prevention and Control (CDC) had issued guidelines and policies to tackle this worldwide problem. These include hygienic and preventive measures, public education, awareness about antibiotic usages in heath and agriculture, alternative therapy such as bacteriophages and probiotics, and new antibiotics development (annual reports from WHO, 2011 & CDC 2013).

The strategy to produce new antibiotics following the 1990s has failed as new antibiotic discovery and development had decreased steadily (Figure 1.3). The current situation with a lack of newly licensed antibiotics can-not fulfill the demand created by the fast development of antibiotic resistance. Therefore, many initiatives have been announced by health-care professionals and relevant scientific agencies. An initiative was announced in 2009 by the Infectious Disease Society of America to develop 10 new antibiotics by 2020 (Piddock, 2012). This was followed by British Society for Antimicrobial Chemotherapy action regarding solutions for the lack of new antibiotics (Piddock, 2012).

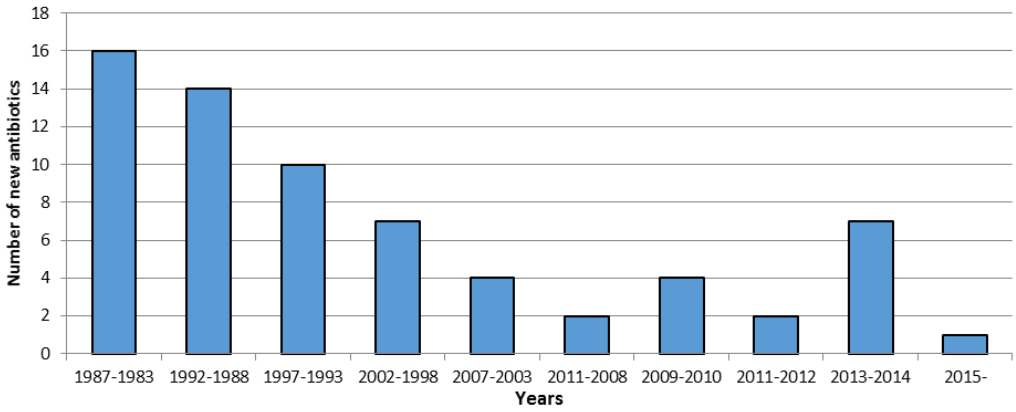


Figure 1.3. Number of new antimicrobials produced over the last three decades (FAD, 2015; CDDEP report, 2015)

New antibiotic analogues production can be often more effective than totally new antibiotics. First, the mechanism of action of the current antibiotics is confirmed. Second, a minor modification does not affect the structure backbone which is critical for activity. Finally, the modification may improve some features of the compounds such as potency,

bioavailability and stability (Madigan *et al.*, 2012). β-hydroxylation of an antimalarial compound by FrbJ enzyme, an Fe(II)/α-ketoglutarate dependent hydroxylase, improves activity and bioavailability (DeSieno *et al.*, 2011).

Because of dramatic developments in genetic techniques and chemical sciences, one can predict the molecular assembly of the antibiotic at the genetic level. This enables us to manipulate the genetic cluster of the producing microorganism to obtain novel derivatives. Many antibiotics are made naturally by bacteria using an assembly line of enzyme complexes like polyketides synthases (PKSs) and non-ribosomal peptide synthetase (NRPS). Understanding the mechanism of enzymatic system and the biosynthesis pathway of products may allow one to rebuild them to produce novel derivatives.

1.2. Polyketides

Polyketides are large classes of biologically active natural compounds produced by a wide range of living organisms such as bacteria, fungi, plants and animals. Because of their diverse pharmaceutical properties and biological activities, polyketides became the subject of many scientific studies and intensive research. They are the most important microbial metabolites in medicinal fields as they are used as antibiotics (for example erythromycin A and rifamycin S), antifungals (for example amphotericin B), anticancer drugs (for example doxorubicin) antiparasitics (for example avermectin) and

immunosuppressants (for example rapamycin) (Figure 1.4) (Weissman & Leadlay, 2005).

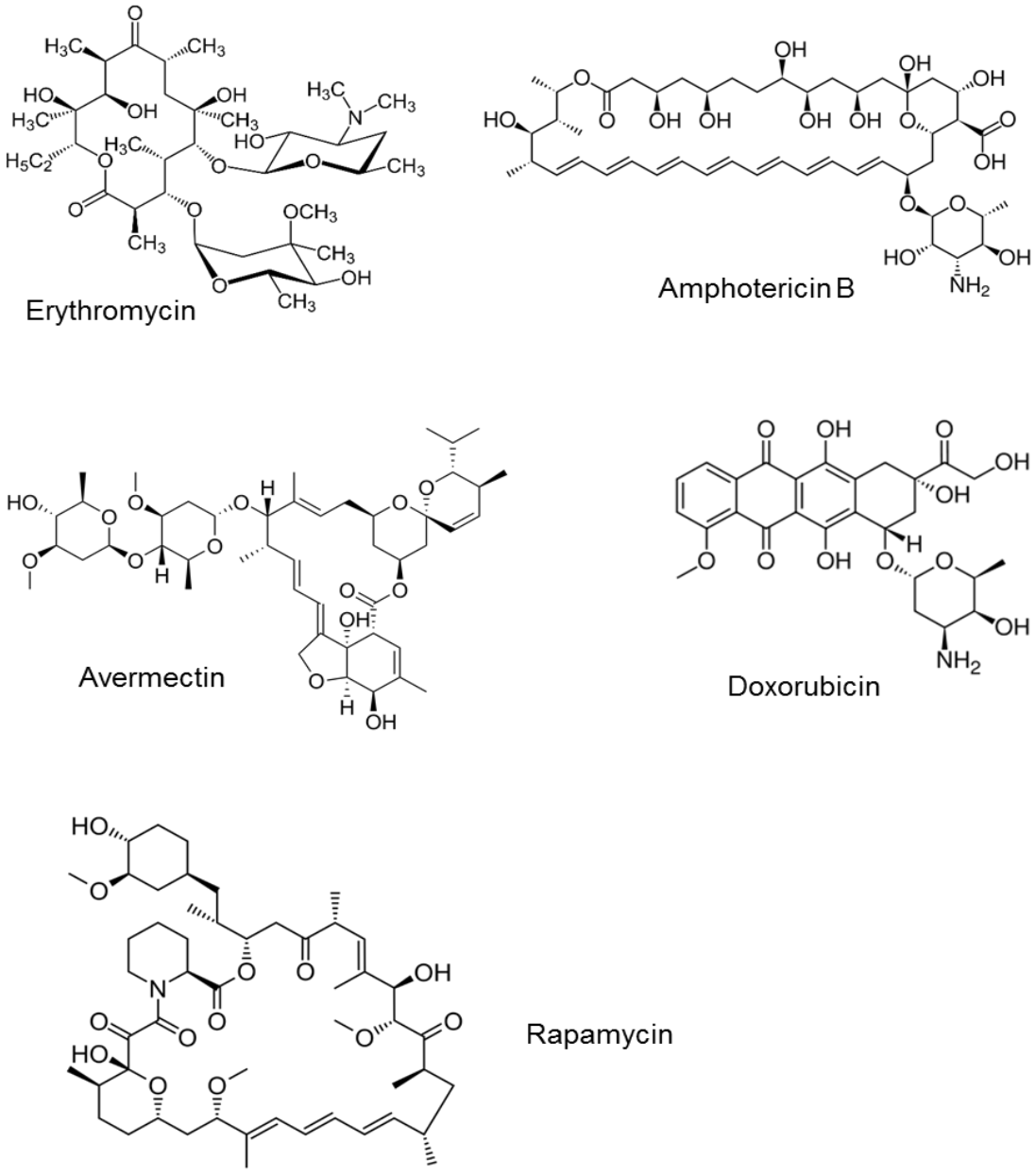


Figure 1.4. Examples of clinically important polyketides (adapted from Weissman & Leadlay, 2005).

1.2.1. Polyketide biosynthesis

Polyketide biosynthesis is catalyzed by multi-domain enzymes called Polyketide Synthase (PKS) which catalyze repetitive decarboxylative condensation reactions in a similar pattern to the fatty acid biosynthesis, fatty acid synthase FAS (Figure 1.5).

Based on protein structure and mode of synthesis, polyketide synthases are classified into three types: Polyketide synthase type I, type II and type III (Table 1.1). Type I PKSs are large multidomain proteins carrying all active sites required for polyketide biosynthesis. According to the mode of action of the enzyme system, Type I PKSs can be subdivided into two types: modular and iterative (Khosla *et al.*, 1999). To carry out one decarboxylative condensation reactions in modular type I PKSs, each module should at least consists of three catalytic domains; Acyltransferase (AT), Acyl carrier protein (ACP) and Ketosynthase (KS). In addition to these three domains, the terminal module houses a Thioesterase (TE) domain which is responsible to release the product (Helfrich & Piel, 2016).

The Acyltransferase (AT) of the PKS system is either located and integrated within each module and referred to as *cis*-AT PKS or as a stand-alone Acyltransferase (AT) domain and referred to as *tran*-AT PKS (Dunn *et al.*, 2014). The ACP domain carries a prosthetic group (4-phosphopatetheine) arm which is attached to highly conserved Serine residue. This arm is a post-translational modification of the ACP domain catalyzed by 4-phosphopantetheinyl transferase (PPTase) (Figure 1.5). This arm works as a flexible crane moving the substrates between catalytic domains.

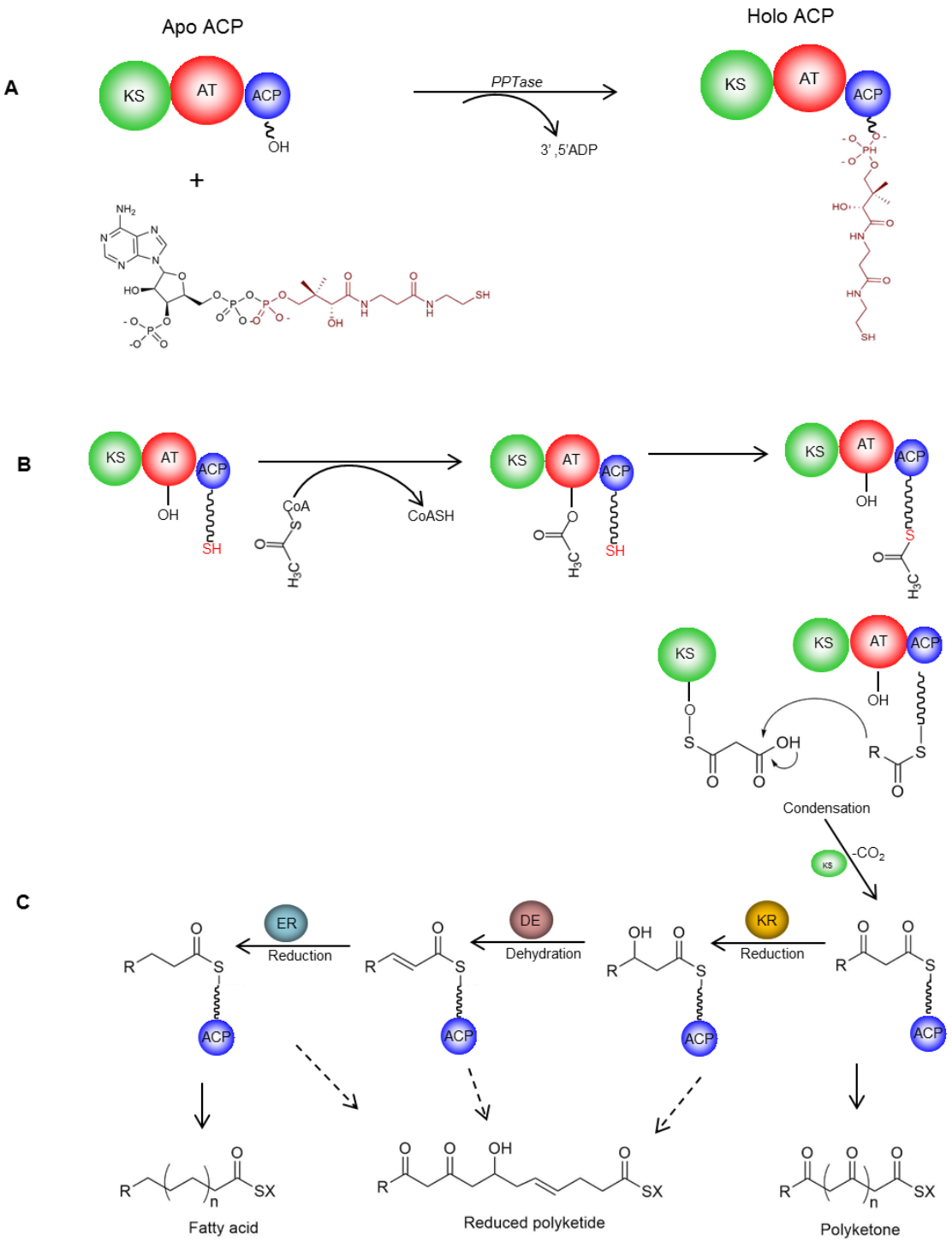


Figure 1.5. Mechanism of polyketide synthetase. A: 4'-phosphopantheinylation of ACP by PPTase. B: a starter unit loading to the AT and PCP domain, and C: the polyketide and fatty acid biosynthesis comparison (Adapted from Peirú *et al*., 2010).

Table 1.1. Types of polyketide synthases (Khosla *et al*, 1999; Watanabe & Ebizuka 2004; Keatinge-Clay, 2012).

PKS types	Structure	Analogy	Organism	Biosynthesis mechanism	Product examples
PKS I	Large multidomain proteins	Analogous to vertebrate fatty acid synthesis	Bacteria: modular type	Linear (use the active sites once	Erythromycin
			Fungi: iterative type	Iterative (reuse the active sites)	Lovastatin
PKS II	Smaller monofuctional active site	Analogous to bacterial fatty acid synthesis	Bacteria	Iterative (reuse the active sites)	Actinorhodin
PKS III	Single protein with multiple modules	Does not rely on ACP	Bacteria and plants	Iterative (reuse the active sites)	Chalcones (CHS) and Stilbenes (STS)

The number of ketide units of the product can be predicted by the number of modules because each module is responsible for the incorporation of one ketide unit (-CH_2-CO-) to the polyketide chain (Khosla *et al.*, 2007; Keatinge-Clay, 2012).

The starter and extender units are activated as Co-A thioesters before loading to the AT domain. Many Co-A derivatives of carboxylic acids can be utilized as starter and extender units. However, the main starter units are acetyl-CoA and propionyl-CoA while malonyl-CoA and methylmalonyl-CoA are the most common extender units (Hranueli *et al.*, 2001). The polyketide carbon chain begins by loading of the starter unit to AT domain.

This catalyzes the transfer of the acyl moiety to the thiol group of the 4-phosphopantetheine arm on the ACP domain. The AT domain loads the extender units to the second module in the same way to ACP. The KS domain catalyzes the condensation reaction between the starter unit and the extender unit to yield a β-ketoacetyl 4C (if the starter is acetate and the extender is malonate). This elongation reaction continues until it reaches the terminal module which releases the forming chain by the TE domain (Hopwood & Sherman, 1990). In addition to the essential domains stated above, PKS modules contain other domains that catalyze extra reactions such as ketoreductase (KR), dehydratase (DH), methyltransferase (MT) and enoylreductase (ER) (Helfrich & Piel, 2016).

Erythromycin, rapamycin and rifamycin are examples of the polyketides that are produced by modular type I PKSs. To understand the mode of enzymatic action, the well characterized and intensively studied example is the 6-Deoxyerythronolide B Synthase (DEBS) which is responsible for biosynthesis of erythromycin (Khosla *et al.*, 1999). The DEBS consists of three large polypeptides (about 350 kDa each) referred to as DEBS1, DEBS2 and DEBS3. Each DEBS unit houses two modules and each module holds a set of domains required to catalyze one round of polyketide chain extension (Staunton & Weissman, 2001) (Figure 1.6).

In addition to the three essential domains (ketosynthase, KS; acyltransferase, AT and acyl carrier protein, ACP) which exist in each DEBS, DEBS 1 contains an extra small module called loading didomain, LDD, which comprises the AT and ACP domains responsible for loading and accepting the starter unit, propionyl-CoA.

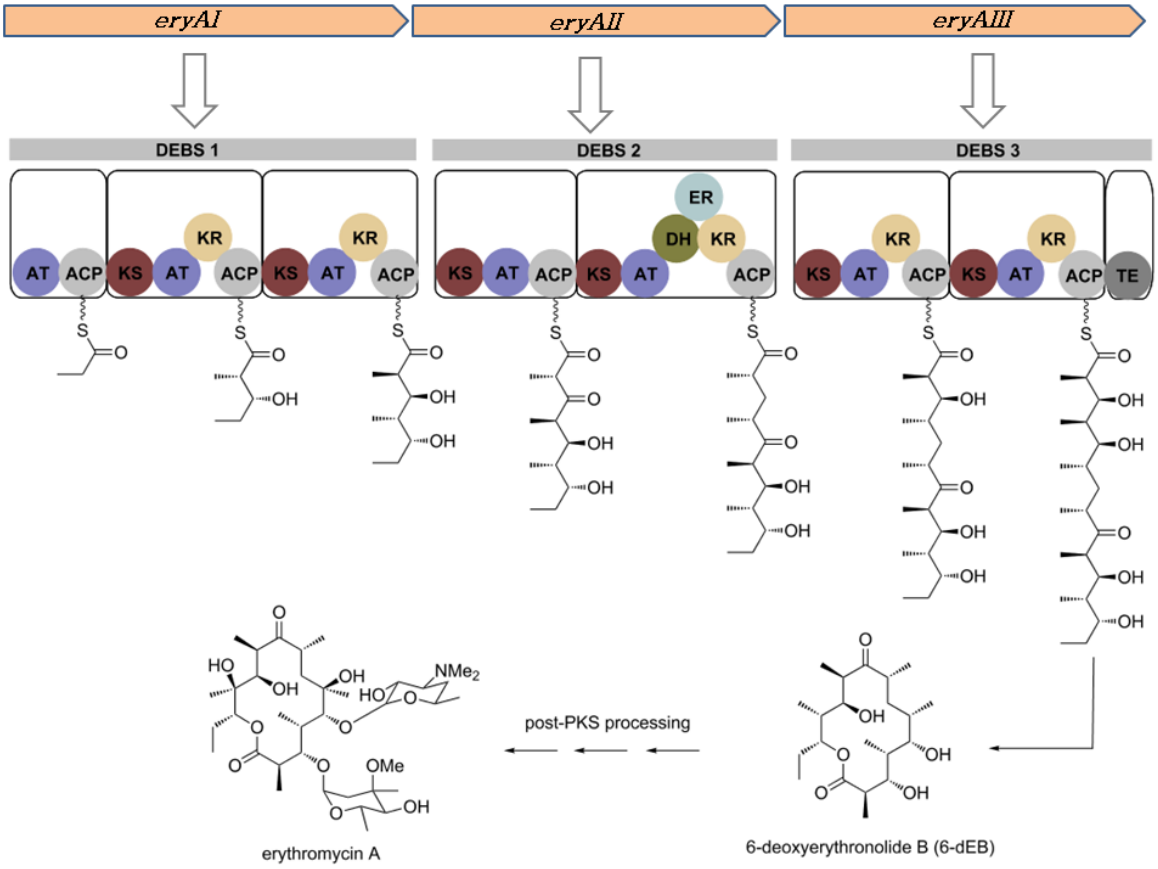


Figure 1.6. The modular type I PKS of 6-deoxyerythronolide B synthase (DEBS) showing the linear organization of the domain, each module houses essential domains KS, AT and ACP for one round extension (Figure adapted from Staunton & Weissman, 2001).

The AT domain in module 1 introduces the extender unit, methylmalonate and transfers it to the prosthetic group (4-phosphopantetheine) arm of ACP domain. The KS domain in the same module accepts the starter unit from the previous module and catalyzes the decarboxylative condensation reaction with the substrate bound to the ACP domain arm. Then, the growing polyketide chain passes from one module to the next until it reaches

the last module which houses the thioesterase TE domain responsible for releasing the final product. While passing from one module to another, the polyketide chain can undergo some extra reactions such as ketoreduction, dehydration and enoyl reduction by ketoreductase KR, dehydratase DH and enoyl reductase ER domains respectively. (Staunton & Weissman, 2001; Khosla *et al.*, 2007).

In contrast to the modular type that contains many active sites which are used only once, the protein structure of iterative PKS is composed of several monofunctional proteins each with only one active site that is re-used repeatedly. After the required number of iteration cycles, the polyketide chain is released by the TE domain (Frandsen, 2010).

1.2.2. Polyketide diversity

The diversity of the products depends on the type and number of the building block units, secondary rearrangement of the product backbone such as cyclization and a broad range of the the tailoring modifications (Figure 1.7).

Although there are many genetic, structural and mechanistic similarities between PKSs and FASs, there are some differences that make polyketides more diverse, active and complex. The polyketides differ from fatty acids by the complete or partial absence of reduction and dehydration reactions that make the latter unfunctionalized alkyl chains (Khosla *et al.*, 1999) Figure 1.5. Low specificity of the enzymes for the starter and

extender units is another cause of polyketide complexity. Studies reported that the loading AT, KR, DH and ER domains in DEBS PKS, which is responsible for erythromycin biosynthesis, accept a wide range of substrates (Khosla *et al.*, 1999). PKSs use a diverse range of starter and extender units which impart structure diversity to the product backbone. The iterative type III PKSs use non- acetate starter units, for example, in the biosynthesis of dihydroxybiphenyl, benzoyl is incorporated as the starter unit. Different malonyl-derivatives have been reported to be used as unusual extender units by modular PKS. In zwittermicin biosynthesis, an antibiotic produced by type I PKS in *Bacillus cereus*, hydroxymalonylate and aminomalonate are incorporated as extender units (Hertweck, 2009). *In vitro* analysis revealed that the methyltransferase (MT) domain of fungal polyketide synthase CazF displays a promiscuity for the substrates and can transfer a non-natural cofactor as an extender unit onto growing polyketide chains imparting the structural diversity of the product (Winter *et al.*, 2013).

Other studies show that the tailoring gene clusters are required to produce active polyketides (Hothersall *et al.*, 2007) which may involve modifications of functional groups such as cyclization and further oxidation-reduction reactions.

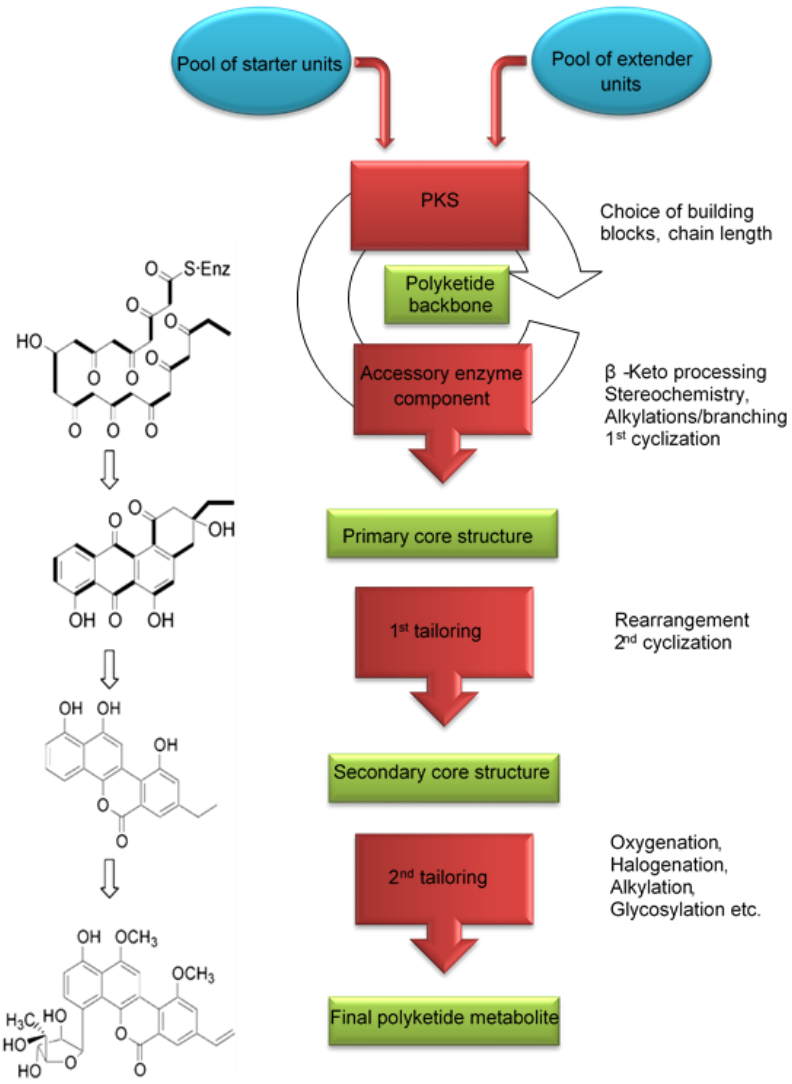


Figure1.7. Gilvocarcin pathway illustrating the diversification of the polyketide product steps (Hertweck, 2009)

1.2.3. Genetic approaches and combinatorial biosynthesis

There are many encouraging factors that push toward the manipulation of the polyketide synthases to generate novel compounds with different activities. The therapeutic importance of polyketide compounds that are already used in clinical and veterinary

medicine and the increasing demand that has been placed on them, and the modular nature of the type I PKSs diversity raises the promise of opportunities to apply genetic engineering to manipulate their structure in ways which are inaccessible by standard synthetic chemical approaches (Kitttendorf & Sherman, 2006).

Many genetic engineering strategies have been applied to modify modular polyketide biosynthesis. These include (i) deletion or inactivation of single domains or mutation that change the substrate specificity; (ii) insertion and substitution of the intact modules; (iii) mutational biosynthesis by feeding of synthetic precursors to mutant phenotype. These manipulations either target the enzymes responsible for backbone building or target the tailoring enzymes that are responsible for post-polyketide processing (Staunton and Weissman, 2001).

Over two hundred new polyketide structures have been produced using two approaches that target the specificity and catalytic power of the PKSs: either by targeting individual catalytic domains via deletion, addition, substitution and inactivation, or by targeting intact modules via replacement, deletion and addition (Kitttendorf & Sherman, 2006). The possible modification of polyketide structures includes chain length modification, incorporating of different building blocks and introducing peripheral moieties (Floss, 2006; Kitttendorf & Sherman, 2006).

There are many studies that involved genetic manipulation of the PKS catalytic domain which leads to incorporation of different starter and extender units. The replacement of the loading didomain (LDD) of the erythromycin cluster with the avermectin loading

domain incorporated new starter units and resulted in erythromycin A, B and D analogs (Khosla, 1999). The alteration of an AT domain by site directed mutagenesis can change the specificity for an extender unit in 6-deoxyerythronolide B synthase (Reeves *et al.*, 2001). AT domain inactivation from module 6 of DEBS can be recovered by malonylCoA:ACP transacylase (MAT) to produce 2-desmethyl-6-deoxyerythronolide B (Kumar *et al.*, 2003). Zhang *et al* (2006) reported that the replacement of reductive domains dehydratase (DH) and ketoreductase (KR) domains of module 2 from the avermectin PKS in *S. avermitilis* Olm73-12 by the DNA fragment encoding the DH, enoylreductase, and KR domains from module 4 of the pikromycin PKS of *Streptomyces venezuelae* produces only avermectin B instead of avermectin A and oligomycin.

A study revealed that the mutant strain of *S. avermitilis* accepts synthetic analogues of the diketide intermediate exogenously added to the culture media as starter units and lead to production of novel analogs of avermectin (Dutton *et al.*, 1994). Some studies focus on the tailoring enzymes that decorate the backbone structure of the polyketides such as oxidation, glycosylation and methylation. These modifications, that are significant for biological activities and the chemical properties of the molecules, thus become the targets for the combinatorial biosynthesis that leads to new polyketides. Examples of these enzymes that exhibited broad substrate tolerance are glycosyltransferases, methyltransferases, P450 mono-oxygenases and nonheme-iron(II) and 2-oxoglutarate-dependent hydroxylases (Weissmann and Leadlay, 2005). Co-expression of glycosyltransferase from the picromycin cluster with a library of plasmids in *S. lividans* host enables addition of a desosamine moiety to the encoded macrolides

and improved antibacterial activities (Tang and McDaniel, 2001). Genetic and biochemical analysis revealed substrate flexibility of P450 hydroxylase (PikC) from *S. venezuelae* to catalyze the hydroxylation of 12- and 14-membered ring macrolides (Xue *et al.*, 1998). Inactivation of nonheme-ketoglutarate-dependent oxygenase in *Glarea lozoyensis* ATCC 20868 abolished the pneumocandin A_0 production and increased the yield of pneumocandin B_0, the semisynthetic precursor of the antifungal drug, caspofungin acetate (Chen *et al.*, 2015). Figure 1.8 shows a number of examples of modified natural products via genetic engineering.

Despite the numerous investigations of polyketide compounds, there are many limitations and difficulties regarding genetic manipulation that must be overcome in order to discover further desired novel molecules. The extraordinary length of the PKS genes from 35 to more than 200 kb, make them difficult to be manipulated. Some of the natural polyketide producers are not amenable to genetic manipulation procedures. Some domains within the same PKS system exhibit a degree of specificity, so that manipulation of the upstream domains may abolish the products. Some relationships and interactions between domains within the same module and between modules are difficult to predict and work in a sequential fashion. Thus, any rational genetic engineering in the gene cluster of the PKS should consider these issues (Menzella *et al.*, 2005; Weissmann and Leadlay, 2005; Kitttendorf & Sherman, 2006).

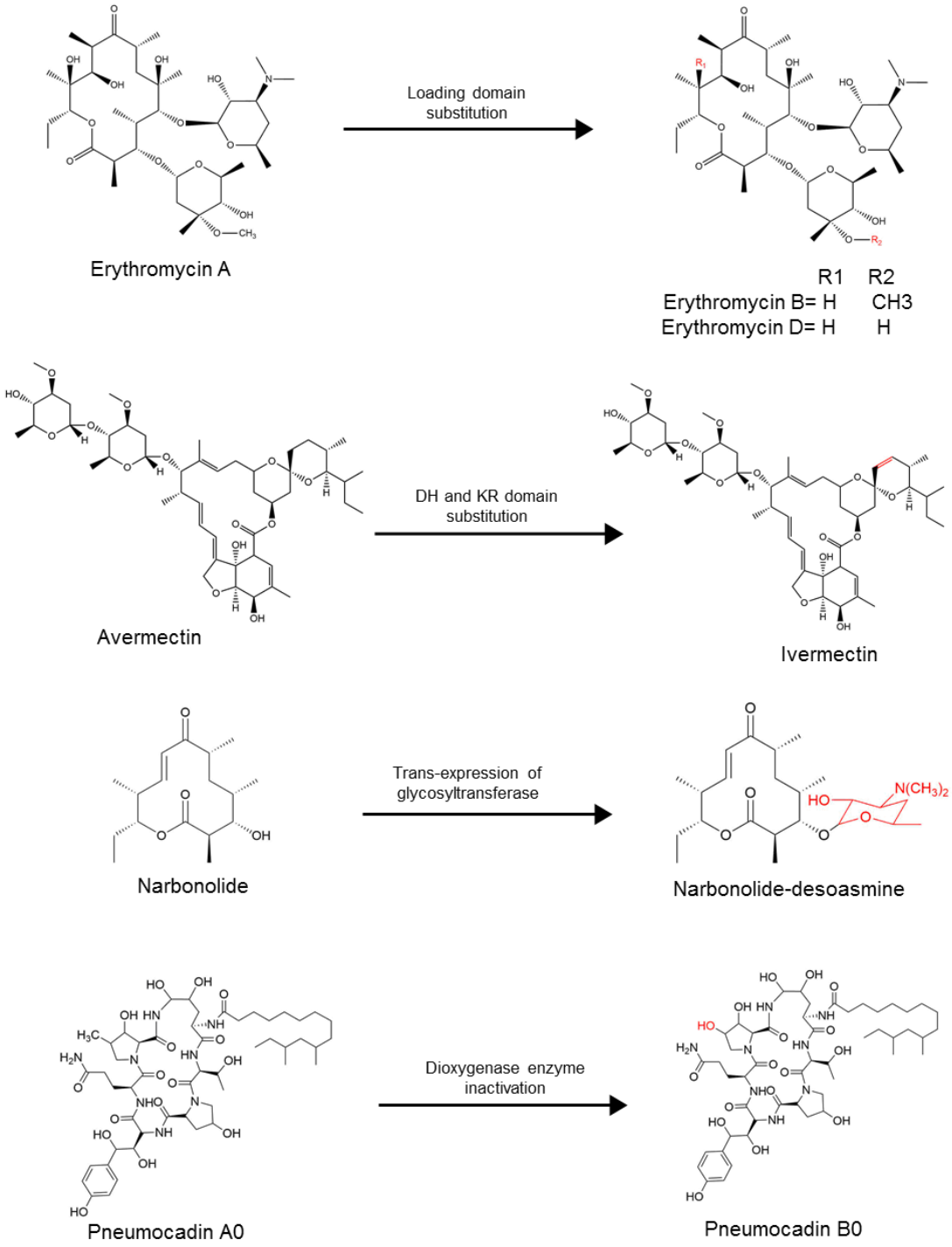


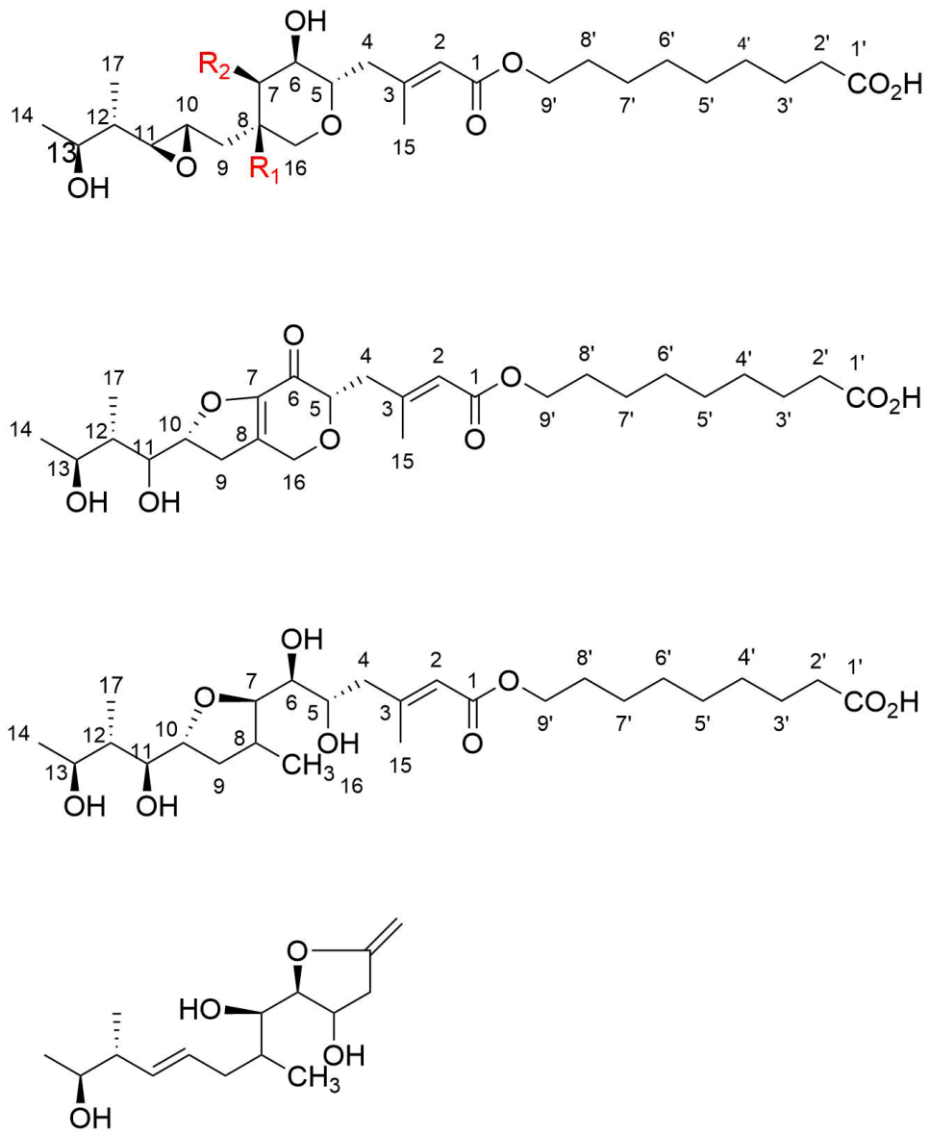
Figure 1.8. Examples of clinically important polyketide structures that have been modified by genetic engineering. The red color represents the modified part.

1.3. Mupirocin

Mupirocin is a polyketide produced by *Pseudomonas fluorescens* NCIMB 10586, a Gram-negative rod-shaped bacteria found in soil living in the vicinity of plant roots. Mupirocin was first described as pseudomonic acid in 1971 by Ernst Chain's laboratory (Fuller *et al.*, 1971). Pseudomonic acid is a mixture of four types (pseudomonic acids A, B, C, and D). Pseudomonic acid A (PA-A) accounts for 90% of the mixture and is composed of two parts: monic acid a C17 molecule containing a pyran ring and saturated fatty acid, 9-hydroxynanoic acid. These two parts are joined by an ester bond. Pseudomonic acid B which accounts (8%) of the mupirocin mixture differs from PA-A by the presence of a hydroxyl group at C8. Pseudomonic acid C and D together account for < 2% of the mixture and differ by a double bond between C10-C11 instead of the epoxide group and unsaturated C-C bond in the fatty acid side chain at C4'-C5' respectively (El- Sayed *et al.*, 2003). Genetic manipulation of the *mup* cluster has generated many derivatives of pseudomonic acids (Figure 1.9).

1.3.1. Mechanism of action

Mupirocin exerts a broad spectrum activity against Gram-positive bacteria, particularly *Staphylococcal* and *Streptococcal* species that cause skin infection, and even against some Gram-negative *Enterobacteriaceae* bacilli particularly at high concentration (Table 1.2). Mupirocin is bacteriostatic at low concentration but bactericidal at high concentration and at lower pH (Wuite, 1983).



	R1	R2	C10-C11	C4'-C5'
Pseudomonic acid A 1	H	OH	Epoxy	Alkane
Pseudomonic acid B 2	OH	OH	Epoxy	Alkane
Pseudomonic acid C 3	H	OH	Alkene	Alkane
Pseudomonic acid D 4	H	OH	Epoxy	Alkene
Mupirocin F 5	H	O	Epoxy	Alkane
Mupirocin C 6				
Mupirocin W 7				
Mupirocin H 8				

Figure 1.9. Pseudomonic acids and derivatives produced by the WT and mutant strains of *Pseudomonas fluorescens* NCIMB 10586.

It is used as a topical antibiotic to treat bacterial skin infections and as a nasal aerosol to clear the nasal passage in patients and hospital staff particularly to eradicate MRSA colonization, which is resistant to a wide range of antibiotics (Cookson *et al.*, 1990).

Table 1.2. Antibacterial activities of mupirocin (Thomas *et al*, 2010)

Organism	Mupirocin MIC ($\mu g.\ ml^{-1}$)
Gram-positive bacteria	
Streptococcus pneumoniae	0.12
Bacillus anthracis	64
Bacillus subtilis	0.12
Clostridium difficile	32
Corynebacterium sp.	> 128
Enterococcus faecalis	64
Listeria monocytogenes	8.0
Staphylococcus aureus	0.25
Staphylococcus epidermidis	0.5
Gram-negative bacteria	
Bacteroides fragilis	> 6,400
Bordetella pertussis	0.02
Escherichia coli	128
Haemophilus influenzae	0.12
Neisseria meningitidis	0.05
Pasturella multocida	0.25
Pseudomonas aeruginosa	6.400

Mupirocin inhibits protein synthesis by binding specifically to isoleucyl-tRNA synthetase (IleRS) which is responsible for the specific aminoacylation of $tRNA^{Ile}$ with isoleucine during ribosomal protein synthesis. IleRS belongs to the class I tRNA synthetases containing an ATP-binding Rossmann fold. As shown in Figure 1.10, in the first step, IleRS catalyzes a covalent linkage between the carboxyl terminus of isoleucine with the 5'-phosphate group of ATP to form an aminoacyl-adenylate intermediate, after which the activated isoleucine is transferred to the 2'OH of the terminal adenosine of $tRNA^{Ile}$. (Hughes & Mellows, 1978; Eriani *et al.*, 1990 and Marion *et al.*, 2009).

The crystal structure revealed that the IleRS structure can be divided into three regions according to the function (Figure 1.11): the Rossmann fold which is responsible for amino acid activation and transferring, the editing domain which removes incorrectly acylated amino acids and the region at the C- and N-terminus that recognizes an unusual anticodon loop conformation in the tRNA (Silvian *et al.*, 1999). IleRS displays a proofreading activity using “double sieve” mechanism depending on the synthetic and editing catalytic sites. First, the synthetic active site refuses amino acids larger than isoleucine such as methionine and phenylalanine, and those with polar sidechains. However, valine, which is smaller by a single methyl group, can fit into the active site and consequently Valyl-AMP or Valyl-$tRNA^{Ile}$ are synthesized. Thus, the second sieve is achieved by the editing site which hydrolyses these misloaded valines (Ibba and Söll, 2000).

Mupirocin acts as a bifunctional inhibitor for IleRS as it occupies both the isoleucine and ATP-binding pockets in active site mimicking the isoleucyl-AMP. The 14-methyl terminus

of the monic acid mimics the side chain of isoleucine while the pyran ring and the region around C1–C3 mimics the ribose and adenine of ATP.

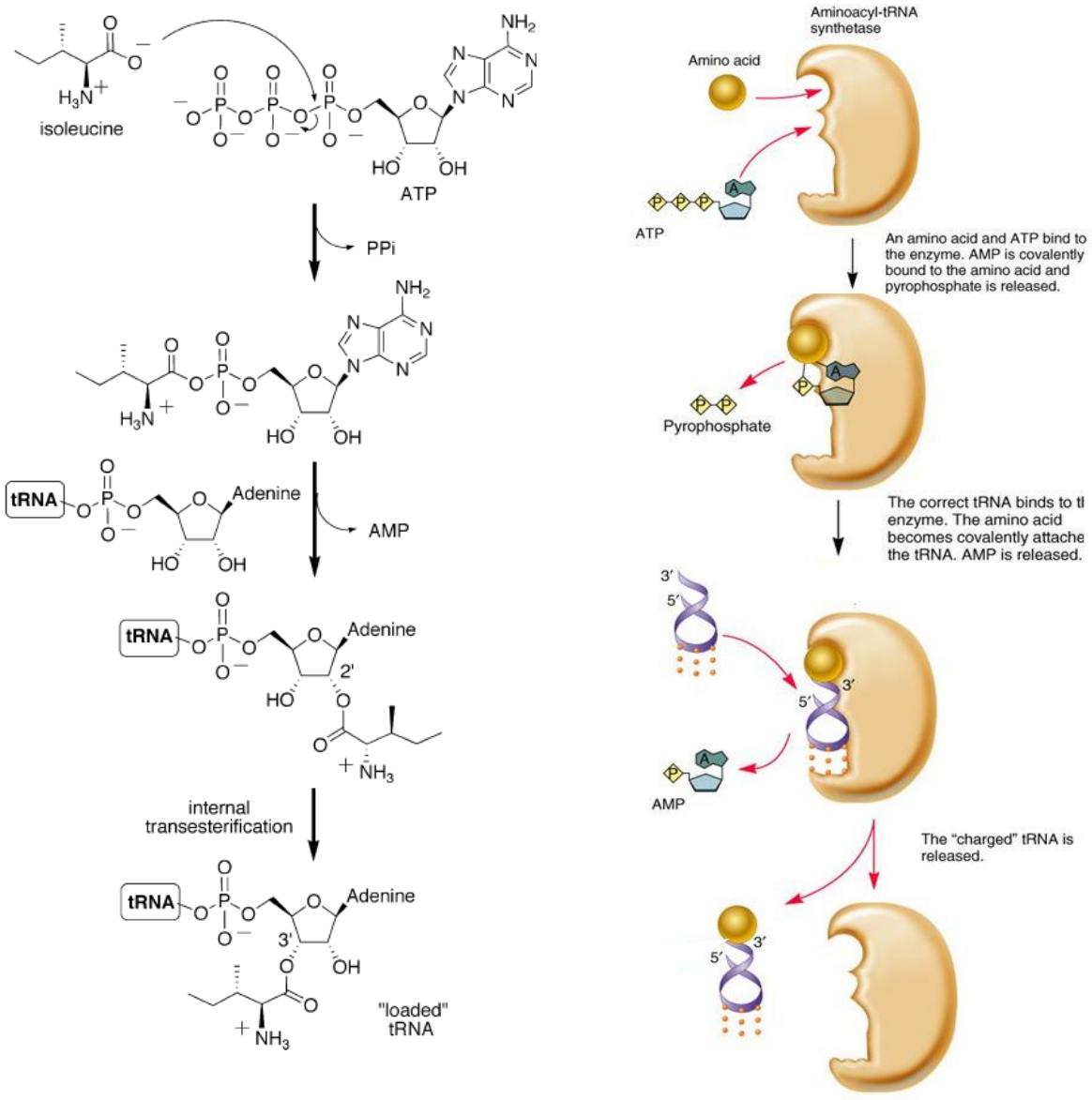


Figure 1.10. Mechanism of action of isoleucyl-tRNA synthetase to activate isoleucine residue (Adapted from Marion *et al.,* 2009 and Madigan *et al.*, 2012).

In addition, this binding could be more stabilized by the 9-hydroxynonanoic acid part which sits in a hydrophobic groove (Hughes & Mellows, 1978; Thomas *et al.*, 2010).

The IleRS cannot use the proofreading activity described above to discriminate mupirocin binding, but uses an alternative “shuttle mechanism” in which the incorrect products have to travel between the synthetic and editing active sites in the same manner to DNA polymerase editing (Silvian *et al.*, 1999).

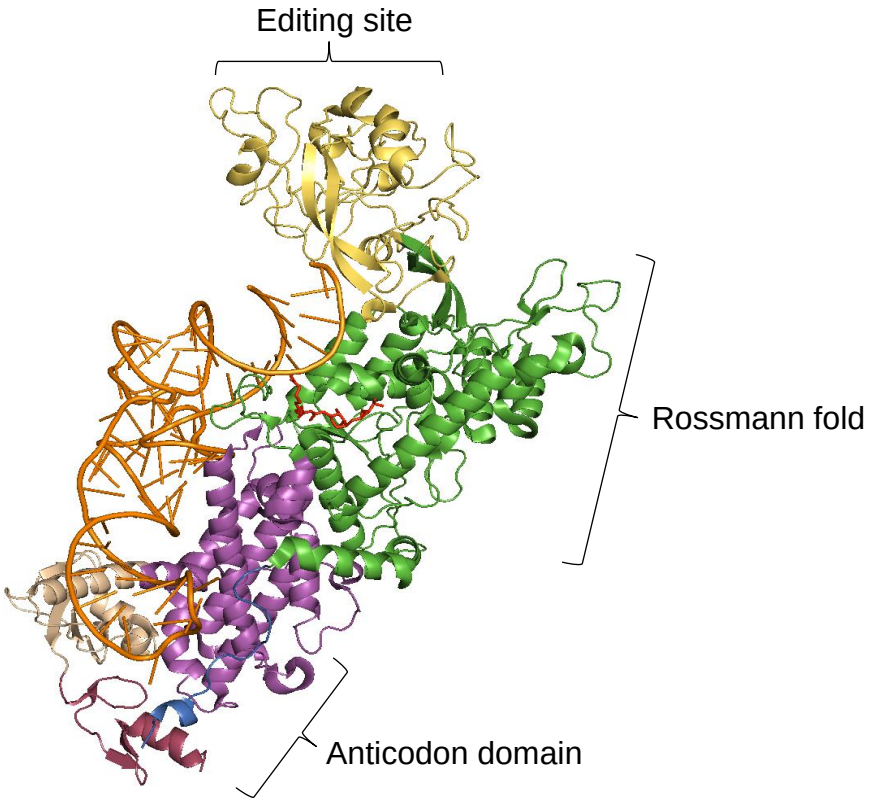


Figure 1.11. Crystal structure of the isoleucyl-tRNA synthetase as a complex with mupirocin. The structure is divided into three regions according to the function; Rossmann fold (synthetic site) (Green), editing domain (site) (Yellow) and N-terminus (anticodon) domain (blue). tRNA is shown as orange sticks and mupirocin as red sticks (Adapted from Silvian *et al.*, 1999).

1.3.2. Limitation and resistance

There are many limitations which allow mupirocin to be used as a topical antibiotic but not as a systemic medication. The ester linkage, which holds both monic acid and 9-hydroxynanoic acid together, is hydrolyzed in body fluids rendering it inactive and the majority binds with serum resulting in insufficient bioavailability. In addition, its biological activity is pH dependent and retained within pH 4-9 and outside these limits, the hydroxyl group at C7 attacks the 10, 11 epoxide producing two inactive cyclic ethers (Clayton *et al.*, 1980; Thomas *et al.*, 2010).

Another problem associated with the use of mupirocin as an antibiotic is the development of resistance by bacteria. Bacteria use more than one mechanism to protect against mupirocin. Resistance was first recorded in 1987 and studies reported that this resistance is due to two types of mechanism. Low-level mupirocin resistance (MIC = 8-56 mg/ml) results from a point mutation in a single nucleotide in the *ileS* gene that encodes the IleRS (Cookson, 1998). The common mutations that occur at the synthetic site of IleRS indentified in resistant *S. aureus* are V588F, V631F and G593V in the vicinity of the KMSKS motif in the Rossmann fold. The substitution of valine 588 with the bulkier residues like phenylalanine distort the Rossmann fold and prevents mupirocin binding (Antonio *et al.*, 2002). Other studies report a similar mutation but in a different residue Q612H which causes considerable disruption in the hydrophobic pocket in the Rossmann fold. Four more mutations (P187F, K226T, F227L and V767D) were identified in clinical isolates of MRSA with low-level mupirocin resistance. These mutations are located outside the Rossmann fold motif and might indirectly affect the

interaction with mupirocin (Yang *et al.*, 2006). Because low-level resistance is non-transferable, it is not currently considered a major threat to clinical use. However, the prevalence of low level resistance in *S. aureus* is more frequent than high-level resistance (Yun *et al.,* 2003).

High-level mupirocin resistance (MIC ≥ 512 mg/ml) is due to the production of a novel IleRS having sequence motifs in its active site similar to the eukaryotic IleRS which is naturally resistant to mupirocin. An evolutionary study identified that this type of IleRS in some bacterial species is acquired by horizontal gene transfer and closely related to archaeal and eukaryotic IleRSs (Brown *et al.*, 2003). This type of IleRS in bacteria is encoded on conjugative plasmids (Hodgson *et al.,* 1994; Yanagisawa & Kawakami, 2003). The prevalence of high-level resistance in *S. aureus* is due to the presence of the *mupA* gene in transferable plasmids and is similar to the *mupM* gene in the mupirocin producer, *Pseudomonas fluorescens,* which is located on the chromosomal DNA (El-Sayed *et al.*, 2003). The prevalence of the high-level mupirocin resistance has been recorded by many studies which is the major threat to clinical use of the antibiotic (Thomas *et al.*, 2010)

Many gram-negative bacteria are resistant to mupirocin because of the effective permeability barrier of the outer membrane against hydrophobic antibiotics including mupirocin (Vaara, 1992).

1.3.3. Mupirocin gene cluster and biosynthesis

The gene cluster that is responsible for mupirocin biosynthesis is known as the *mup* cluster and occupies 74 kb of the *P. fluorescens* NCIMB 10586 chromosome. The cluster contains thirty-five Open Reading Frames (ORFs) divided into two parts (Figure 1.12). The first part includes six large ORFs encoding proteins known as mupirocin multifunctional polypeptides (mmp) from MmpA to MmpF.

The *mmpA* gene encodes one non-elongating module and two elongation modules for monic acid backbone. The *mmpB* plus *mupS, Q, E*, *mAcpD* and *mmpF* encode the necessary domains for 9-hydroxynanoic acid synthesis. The thioesterase domain at C-terminus is responsible for the release of the product from PKS. The *mmpC* gene encodes two acyltransferase domains responsible for starter and extender units loading to KS and ACPs (El-Sayed *et al.*, 2003; Gurney & Thomas, 2011). The *mmpD* gene encodes the four elongation modules for the monic acid backbone. The *mmpE* encodes single KS/OR domains responsible for 10, 11-epoxide formation (Gao *et al.*, 2014). The *mmpF* encodes single KS domain and the function remains unknown, but its deletion aborted PA-A production (Hothersall *et al.*, 2007).

The second part consists of twenty-six individual genes (*mup* and *mAcp*) known as tailoring genes required for the necessary modification of the product backbone. In addition to these two parts, the cluster also includes gene (*mupM*) responsible for self-protection to mupirocin and the genes (*mupX, mupR* and *mupI*) for the regulation of the biosynthesis. (El-Sayed *et al.*, 2003; Gurney & Thomas, 2011).

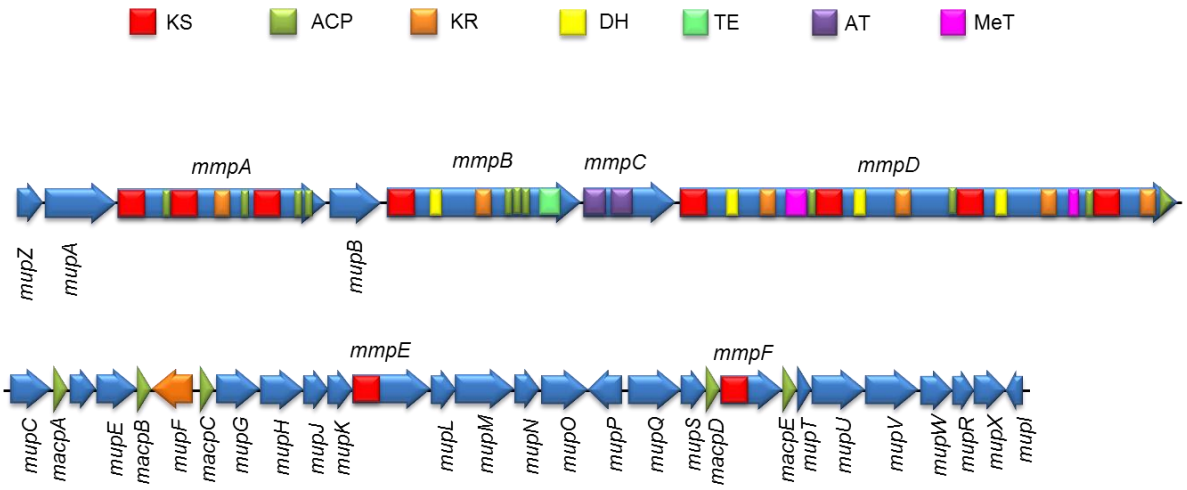


Figure 1.12. Mupirocin biosynthesis gene cluster in *P. flourescens* NCIMB 10586 (Figure taken from El-Sayed *et al*., 2003 after permission).

1.3.3.1. Monic acid biosynthesis

The biosynthesis begins at MmpD which consists of the first four elongation modules for monic acid backbone and followed by MmpA which consists of three modules achieving the fifth and sixth elongation steps. Each module contains KS and ACP domains in a standard manner for type I PKS. The AT domains which are responsible for starter units (acetyl- CoA) activation or the extender unit (Malonyl-CoA) loading to the KS and ACP domains are proposed to be provided by MmpC, which contains tandem acyl hydrolase, in a pattern of trans-acyltransferase class modular PKS. The synthesis may start by transferring the activated acetyl-CoA to the 4'-phosphopantetheinyl arm of ACP in the first module of MmpD, which then transfers to the thiol group of active cysteine of KS, but this is just speculation at present. The extender unit is loaded to the 4'-

phosphopantetheinyl arm of ACP, then the condensation between starter and extender units is catalyzed by KS domain (El-Sayed *et al.*, 2003; Gurney & Thomas, 2011).

The four modules of MmpD perform four condensation reactions before the growing chain is transferred from MmpD to MmpA by the first module of MmpA, which contains KS and ACP, to complete the C14 backbone of monic acid. The KS in the first module of MmpA is inactive in terms of Claisen condensation because it lacks the two key histidines in the active site, so the function of this module is more likely to transfer the intermediate from the last module of MmpD to the second module of MmpA or it may be the location where 6-hydroxylation occurs as the result of a tailoring enzyme, currently suspected to be mupA (Figure 1.13). The first and third modules of MmpD have MT1 and MT2 respectively, which catalyze the addition of the C16 and C17 methyl groups derived from S-adenosyl methionine (El-Sayed *et al.*, 2003; Thomas *et al.*, 2010).

1.3.3.2. Hydroxynanoic acid biosynthesis

The 9-hydroxynonanoic acid biosynthesis is assembled by *mmpB*. The starter unit of the 9-hydroxynonanoic acid is 3-hydroxypropionate, suggested to be produced by *mupQ*, *mupS* and *mAcpD*, combined with the extender unit, malonate, is extended through three rounds of condensations catalyzed by *mmpB* which functions iteratively (Gurney & Thomas, 2011). Studies on mupirocin and thiomarinol, an analogue to mupirocin produced by *Pseudoalteromonas* spp, confirmed the isolation of truncated saturated fatty acyl side chains in mutant and wild type.

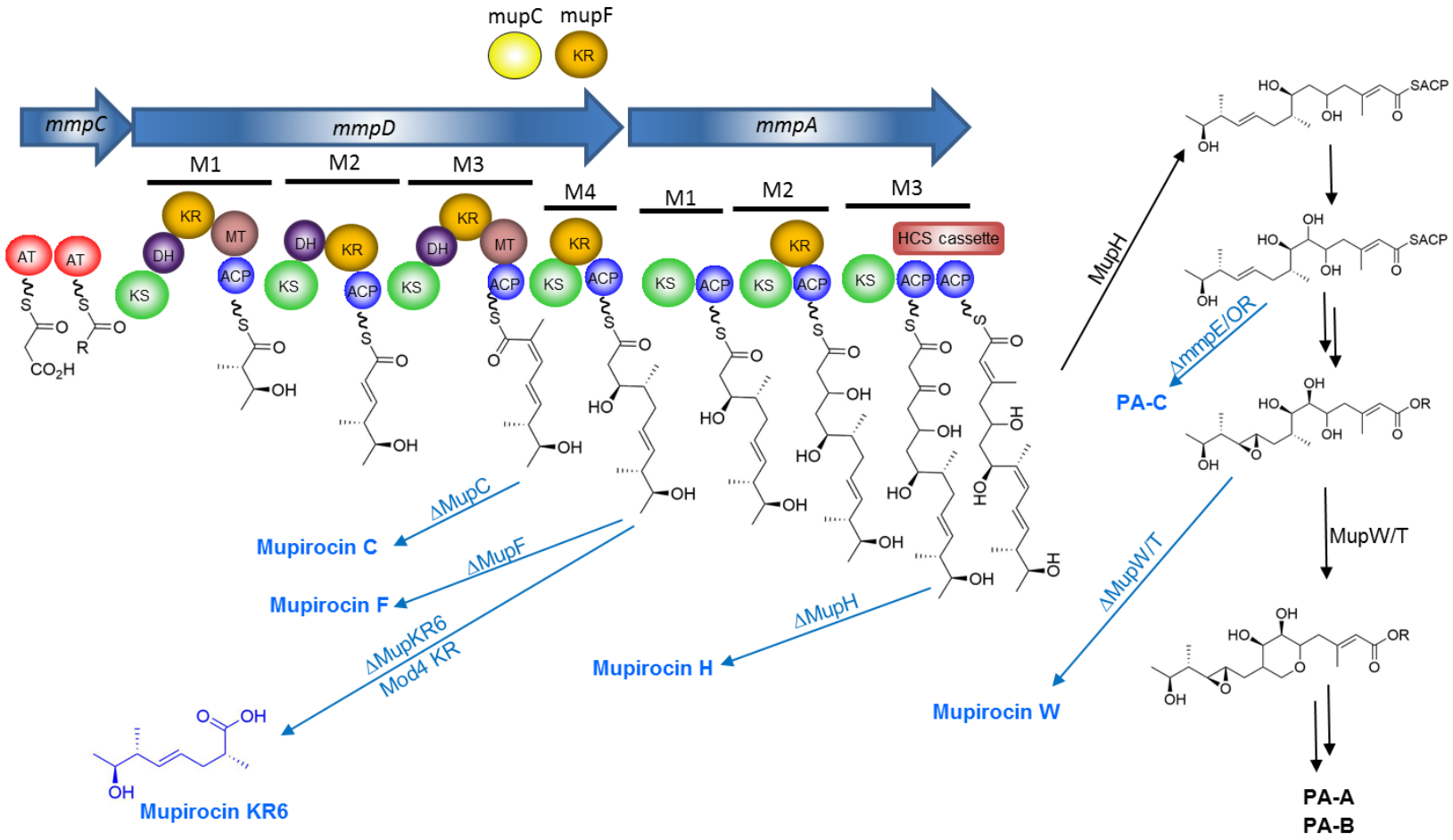


Figure 1.13. Proposed pathway of pseudomonic acid biosynthesis in *P. flourescens* NCIMB 10586. Only the monic acid assembly is shown. The blue compounds produced by gene deletion, see Figure 1.9 for compound structure and Table 1.3 for mutant products.

This suggests that the fatty acyl side chains are built up by successive elongation and not by ligation of fully assembled 9-hydroxynanoic acid (Murphy *et al.*, 2011).

1.3.3.3. Tailoring genes activities

The tailoring genes include twenty-six single open reading frames (*mupC to mupZ, macpA* to *mapcE*). Mutational studies show that all the tailoring genes are necessary to modify the intermediate structure to produce mupirocin. These modifications include pyran ring formation, epoxidation of C10-C11, incorporation of the methyl group at C3 and hydroxylation at C6 (Gurney & Thomas, 2011). The Orf *mupW* is thought to be responsible for the oxidative activation of the C16 methyl group and its deletion produced a novel compound lacking the tetrahydropyran ring known as mupirocin W (Cooper *et al.*, 2005a). Studies showed that MupC, MupF, MupO, MupU, MupV and mAcpE are required for the correct oxidation around the pyran ring and their deletion generates pseudomonic acid B (Table 1.3). Deletion of *mupF* leads to the production of mupirocin analogue with C7 ketone while a *mupC* deletion results in failure to reduce C8-C9. The addition of the C15 methyl group to C3 is introduced by hydroxymethylglutaryl-CoA synthase (HCS) cassette which consists of MupG, MupH, MupJ, MupK and mAcpC. (Cooper *et al.*, 2005b; Hothersall *et al.,* 2007; Haines *et al*., 2013).

Table 1.3. Examples of some products produced by mutant strains of *Pseudomonas flourescens* NCIMB 10586

ORF	Proposed function	Mutant product
mupC	Oxidoreductase	Mupirocin C
mupF	Ketoreductase	Mupirocin F
mupW	Dioxygenase	Mupirocin W
mupV	Oxidoreductase	Pseudomonic acid B
mupO	Cytochrome P450	Pseudomonic acid B
mupU	CoA synthase	Pseudomonic acid B
macpE	Acyl carrier protein	Pseudomonic acid B
mupH	HMG-CoA synthase	Mupirocin H
mmpE	Oxidoreductase (C-terminus part)	Pseudomonic acid C
mupE	Enoylreductase	6'-7' enoyl bond

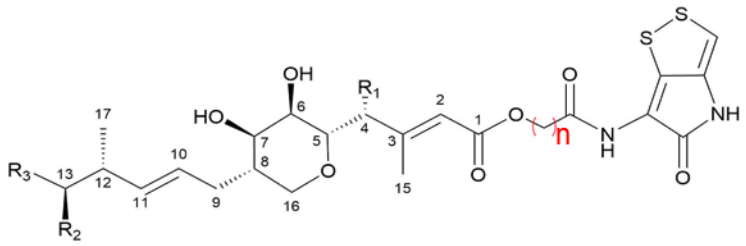
1.4. Thiomarinols

Thiomarinols are natural compounds produced by a marine bacterium, *Pseudoalteromonas* spp SANK 73390 and display antibacterial activity against both Gram positive and Gram negative bacteria (Table 1.4). Thiomarinols consist of two portions of independent antibiotics: marinolic acid, an analogue to pseudomonic acid produced by *Pseudomonas fluorescens,* and the pyrrothine core from holomycin antibiotic linked together via an amide bond (Shiozawa *et al.*, 1993). At least seven thiomarinol structures have been isolated from the wild type *Pseudoalteromonas* spp SANK 73390 (Figure 1.14). Thiomarinol A was isolated as a major product while thiomarinol B, C, D, E, F and G were isolated as the minor products from the culture broth of WT SANK 73390.

Table 1.4. Antibacterial activity of thiomarinol types A-G (Shiozawa *et al*, 1993; 1994 and 1997).

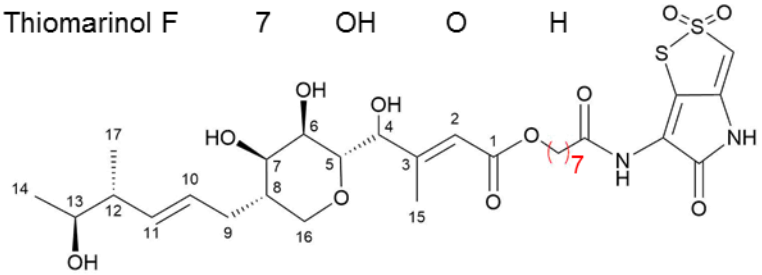
Test Organism	MIC (µg. ml^{-1})						
	A	B	C	D	E	F	G
Staphylococcus aureus MRSA	<0.01	≤0.01	≤0.01	≤0.01	≤0.01	≤0.01	≤0.01
Staphylococcus aureus 209P JC-1	<0.01	≤0.01	≤0.01	≤0.01	≤0.01	≤0.01	≤0.01
Enterobacter cloacae 963	6.25	0.8	3.1	1.5	1.5	12.5	>200
Enterococcus faecalis NCTC775	<0.01	0.05	0.8	0.2	0.05	0.2	3.1
Klebsiella pneumoniae IID685	0.78	0.8	1.5	0.8	0.8	3.1	12.5
Salmonella enteritidis G	3.13	0.4	1.5	0.8	0.8	3.1	25
Serratia marcescens IAM 1184	25	3.1	6.2	6.2	6.2	25	>200
Proteus vulgaris IID874	0.39	0.05	0.2	0.1	0.2	1.5	3.1
Escherichia coli NIHJ JC-2	3.13	0.8	3.1	1.5	1.5	6.2	>200
Morganella morganii IFO3848	12.5	6.2	12.5	12.5	12.5	100	>200
Pseudomonas aeruginosa PA01	0.39	0.8	0.4	0.4	0.4	6.2	12.5

The marinolic acid of thiomarinol A (TMA) is structurally close to PA-C in which the 10,11 epoxide is replaced by an alkene, but it differs by having an extra hydroxyl group at C4 and the fatty acid moiety is shorter by one carbon (Shiozawa *et al.*, 1993). Thiomarinol B (TMB) varies from TMA by having two extra oxygen atoms attached to the pyrrothine moiety, while thiomarinol C (TMC) has one less oxygen atom at C4 (Shiozawa *et al.*, 1995). As shown in Figure 1.14, thiomarinol D (TMD) has one extra methyl group at C14, thiomarinol E (TME) has two more methylenes in the fatty acid chain while in TMF the hydroxyl group at C13 has been replaced by a ketone. Thiomarinol G is similar to PA-C and can be designated as 6-deoxy-8-hydroxymonic acid C (Shiozawa *et al.*, 1997). All these minor products showed variable antibacterial activity against Gram positive and Gram negative bacteria (Table 1.4). Murphy and co-workers isolated more minor products from the WT *Pseudoalteromonas* spp SANK 73390 (Murphy *et al*., 2013).

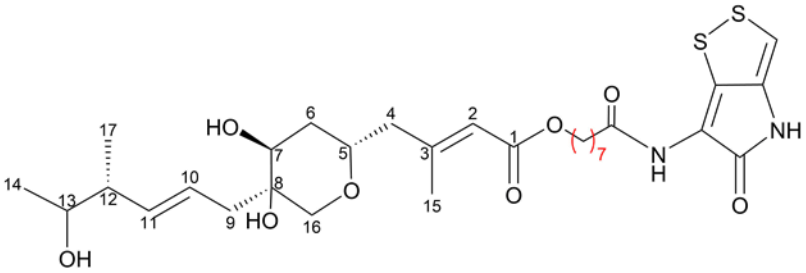


	n	R1	R2	R3
Thiomarinol A	7	OH	OH	H
Thiomarinol C	7	H	OH	H
Thiomarinol D	7	OH	OH	CH3
Thiomarinol E	9	OH	OH	H
Thiomarinol F	7	OH	O	H

Thiomarinol B



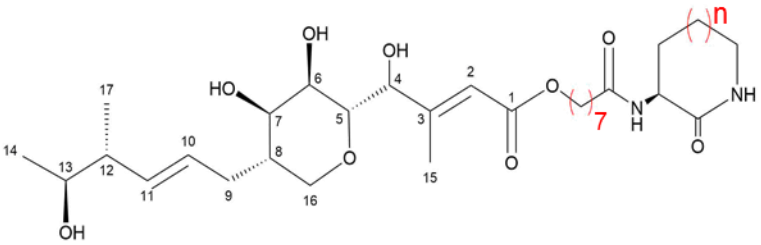
Thiomarinol G



n

Thiomarinol H = anhydroornithine

Thiomarinol J = anhydrolysine



Xenorhabdins

R= different fatty acid lengths

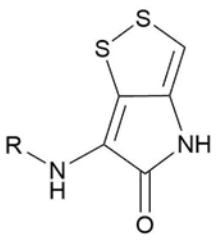


Figure 1.14. Thiomarinols produced by WT and mutant strains of *Pseudoaltermonas spp* SANK.

These include marinolic amide and the range of pyrrothine metabolites in which different fatty acid chain lengths are attached to the pyrrothine molecules and were named xenorhabdins. Later mutation experiments suggested that the marinolic amide is likely a degradation product of thiomarinol A (Murphy *et al.,* 2011). Mutasynthesis experiments produced more thiomarinol derivatives. Thiomarinol H and J (Figure 1.14) were produced by feeding anhydroornithine and α-Aminobutyrolactone respectively to a ∆NRPS strain (Murphy *et al.,* 2011).

1.4.1. Thiomarinol gene cluster and biosynthesis

DNA sequencing of the circular plasmid pTML1 from *Pseudoalteromonas spp* SANK 73390 identified the gene cluster (about 97 kb) responsible for thiomarinol biosynthesis. This gene cluster consists of trans-AT PKSs with forty-five ORFs, twenty-seven of them are responsible for marinolic acid production in a similar way to the *mup* cluster of mupirocin, and seven ORFs encoding production of the pyrrothine core with similarities to NRPS and tailoring enzymes responsible for holomycin biosynthesis in *Streptomyces clavuligerus* (Figure 1.15). Many genes in the thiomarinol gene cluster (*tmp*) are similar to those in the mupirocin (*mup*) cluster (Fukuda *et al.,* 2011). The genes *tmpA* to *tmpD* encode the polyketide synthase (PKS) in a similar way to the *mup* cluster (Figure 1.12).

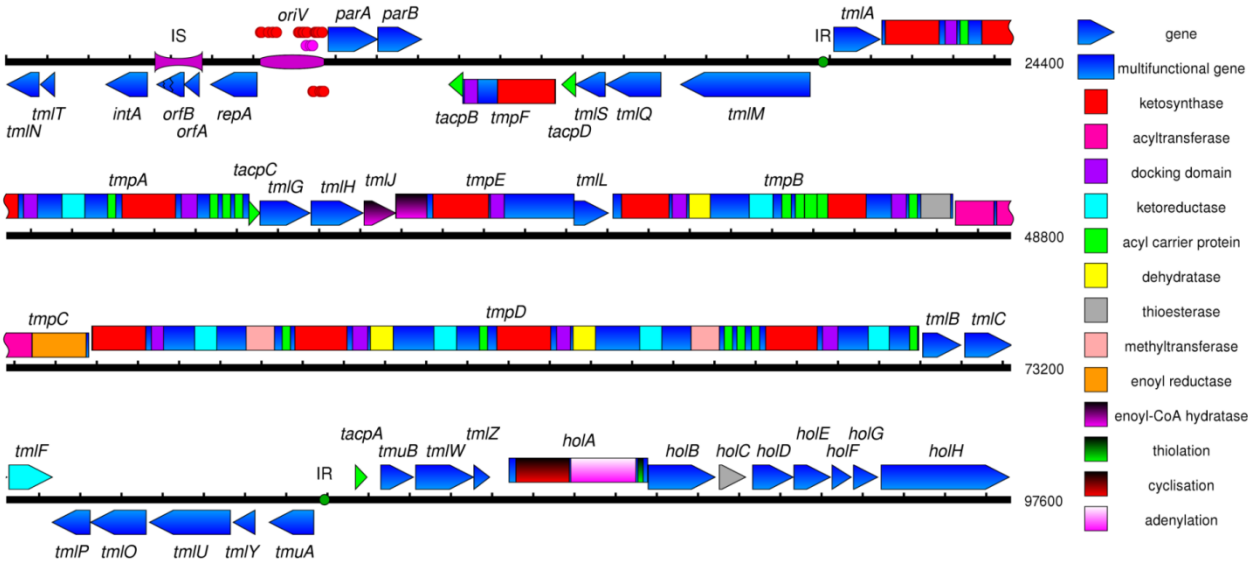


Figure 1.15. Thiomarinol gene cluster on pTML1 (Figure taken from Fukuda *et al.*, 2011 after permission).

In spite of the similarities between the *tmp* and the *mup* clusters, there are many differences. First, the *tmp* possess extra ACP domains in the third module of *tmpD* and in the last module of *tmpA*.

Second, the *tmpB* gene varies from *mmpB* by encoding two modules, one with four ACP domains and one with a single ACP domain. Third, most of the tailoring genes in the *tmp* cluster have a counterpart in the *mup* cluster. However, depending on the direction of transcription and the presence of the *rep/par* interruption, the tailoring genes can be split into five transcription units; *tmlT* to *tmlN*, *tmlM* to *tacpB*, *tmlA* to *tmlF*, *tmuA* to *tmlP* and *tacpA* to *holH* (Fukuda *et al.,* 2011). Fourth, the absence of mAcpE in the *tmp* cluster suggests that all the late tailoring steps are catalyzed by *tmpB* which contains extra KS

and ACP domains. Fifth, the presence of some orfs in *tmp*, but absence in the *mup* cluster such as *tmuB.* Finally, the *tmp* possesses seven orfs (*hol* genes) which encode the non-ribosomal peptide synthetase and tailoring enzymes responsible for pyrrothine molecule assembly (Fukuda *et al.,* 2011).

Suicide mutagenesis revealed that *tmpD* is equivalent to *mmpD* in the *mup* cluster and its deletion aborts thiomarinol and marinolic acid production, while the *holA* knockout, abolishes pyrrothine production and retains marinolic acid. This indicates that the marinolic acid and pyrrothine can be synthesized separately (Fukuda *et al.,* 2011). To confirm joining the pyrrothine moiety marinolic acid as an intact unit, ΔNRPS SANK 73390 strain was fed pyrrothine which resulted in thiomarinol A production. The pyrrothine unit is proposed to be formed from two molecules of cysteine via NRPS biosynthesis before linking to marinolic acid. Isotope labelling studies suggested that the assembly of the octanoate fatty acid of marinolic acid does not follow the standard polyketide pattern in which the acetate units incorporate in head-to-tail pattern, but FAS catalyzes two rounds extension of C_4 precursor or incorporates C_4 precursor to the polyketide and then elongates it by two C_2 extensions (Murphy, *et al.,* 2013). The joining of the marinolic acid and pyrrothine molecule is carried out by *tmlU and holE.* An *in vitro* study revealed that TmlU activates marinolic acid as acyl-CoA while HolE acts as an acyltransferase to join it with pyrrothine. TmlU displays substrate specificity while HolE accepts a broad range of the fatty acyl CoA derivatives which explain the xenorhabdins production (Dunn *et al.,* 2015). Figure 1.16 shows the proposed thiomarinol biosynthesis based on the DNA sequences and mutational analysis.

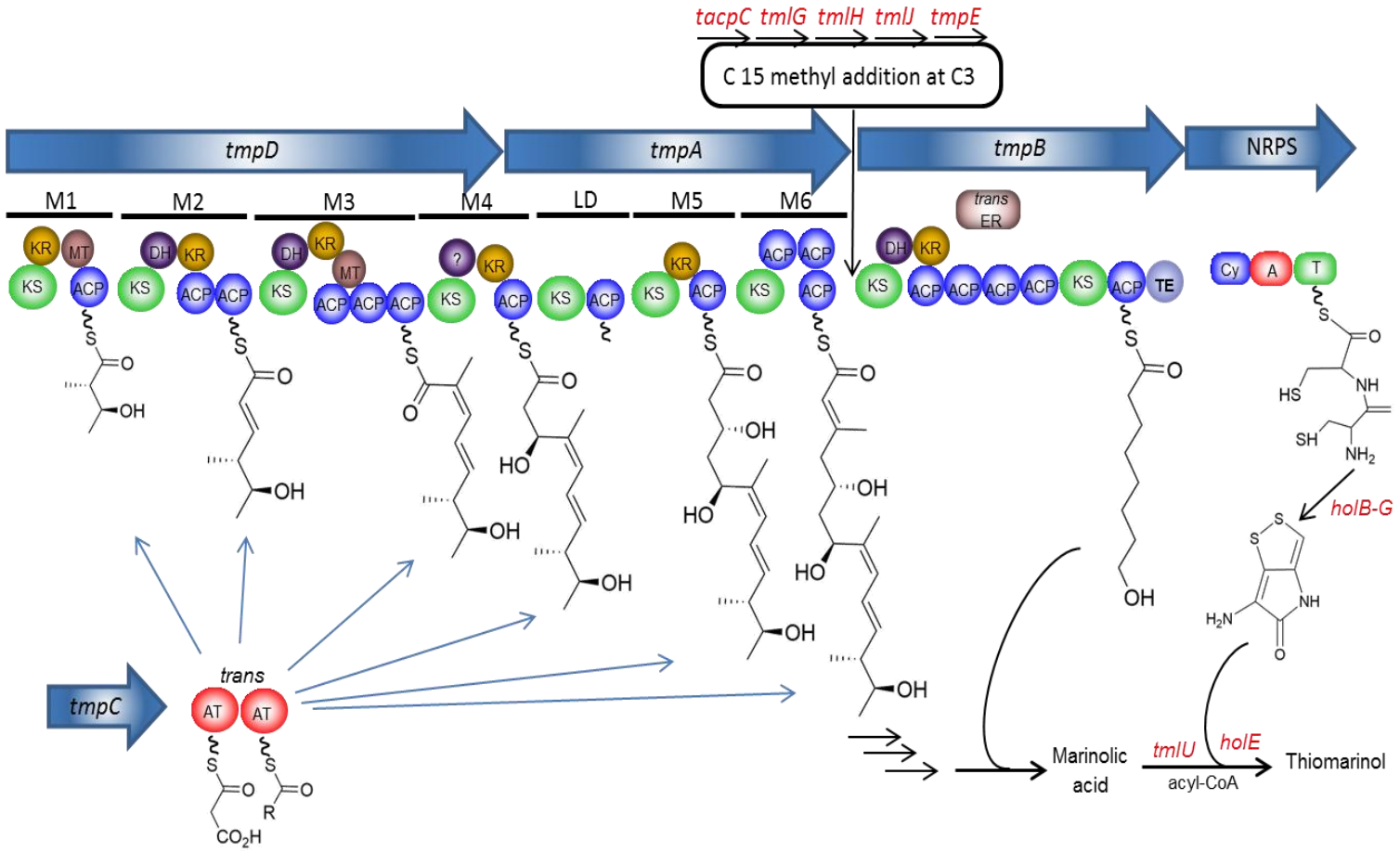


Figure 1.16. Proposed pathway of thiomarinol biosynthesis in *Pseudoalteromonas* SANK. *tmpC* provides tandem AT domains which works *in trans* to supply *tmpD* and *tmpA* with the starter and extender units to produce monic acid with the assist of the tailoring genes, *tmpB* assembles the 8-hydroxy octanoic acid which links to monic acid to form marinolic acid. *tmlU* activates the formed marinolic acid as acyl-CoA and *holE* acts as acyltransferase to join it with pyrrothine.

1.5. Non-ribosomal peptide synthetases

Nonribosomal peptides are natural products produced by multidomain modular enzyme assemblies known as Non-ribosomal peptide synthetase (NRPS). These compounds are structurally diverse with a broad spectrum of biological activities that can be applied as agrochemical agents or in the development of modern medicine including antibiotic, immunosuppressive and anticancer drugs (Finking & Marahiel, 2004; Winn *et al.,* 2016).

To highlight the importance of the NRPS system, some remarkable examples of nonribosomal peptides are shown in Figure 1.17. Vancomycin, a glycopeptide antibiotic active against Gram-positive bacteria including methicillin resistant *Staphylococcus aureus* is made by an NRPS (Wageningen *et al*., 1998). Cyclosporin, the powerful immunosuppressive drugs used to treat autoimmune disorders and in organ transplantation, is a cyclic NRP produced by *Tolypocladium inflatum* fungus (Weber *et al.,* 1994). Bleomycin produced by *Streptomyces verticillus* has anticancer properties (Du *et al*., 2000)

NRPS products characterized by distinctive features differentiate them from ribosomal peptides. Most of the NRPS products are cyclic or branched cyclic containing not only the common 22 amino acids but also non-proteinogenic amino acids, sugar, fatty acids and the unusual modifications in the peptide backbone connecting the amino acids by bonds other than peptide or disulfide bonds (Figure 1.17). These features impart NRPS products with chemical diversity associated with powerful bioactivity (Schwarzer *et al.,* 2003; Challis and Naismith, 2004 and Hur *et al.,* 2012). Fengicin and surfactin, both

antibiotics, contain a different length of fatty acid chain. Non-proteinogenic amino acids such as D-phenylalanine and ornithine are incorporated into the structure of the antibiotic gramicidin S. The immunosuppressive agent cyclosporine A contains unusual amino acids, L-α-amino butyric acid and 2-butenyl-4-methyl-L-threonine. Many NRPs contain heterocyclic rings generated from cysteine, serine and threonine residues and the ring formed is named based on the oxidative state. Bleomycin A contains two thiazol rings generated by cyclization and oxidation of two cysteine residues (Finking & Marahiel, 2004).

1.5.1. NRPS types and biosynthesis

NRPSs are large multimodular megaproteins with repeating catalytic domains. Each module is responsible for incorporating one building block to the product, so that, the number of amino acids in the product can be predicted by the number of NRPS modules. According to the mode of biosynthesis, modular NRPSs are classified into three types: Linear (type A) in which the number and sequence of amino acids are correlated with the number and sequence of the NRPS modules, Iterative (type B), as an analogue to iterative PKSs, in which the catalytic domain is reused repeatedly and Nonlinear (type C), in which the arrangement of NRPS modules does not match the order of amino acids (Figure 1.18) (Mootz *et al.,* 2002; Hur *et al.,* 2012).

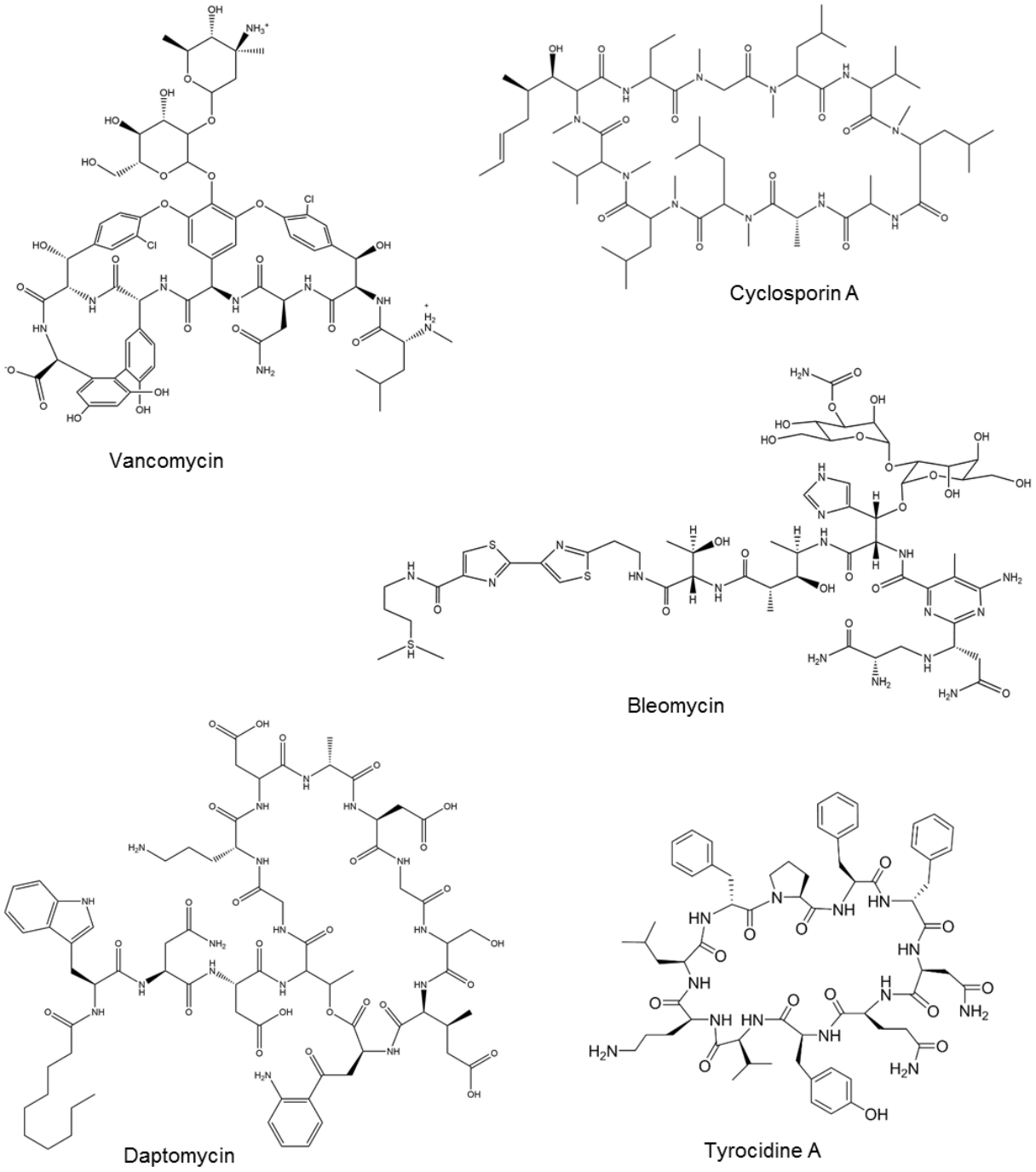


Figure 1.17. Examples of clinically important non-ribosomal peptides (Collected from Hur *et al*., 2012 and Winn *et al*., 2016)

The minimal NRPS elongation module consists of three core domains: Adenylation (A) domain, Condensation (C) domain and Peptidyl Carrier Protein (PCP)(Figure 1.19). By consuming ATP, the Adenylation domain recognizes and activates the amino acid to form an amino-acyl adenylate. This intermediate transfers to and binds covalently to the thiol group of 4'- phosphopantetheinyl arm of the PCP domain. This moiety is approximately 20 Å in length, acts as a flexible arm that helps the bound substrate to shuttle between catalytic sites.

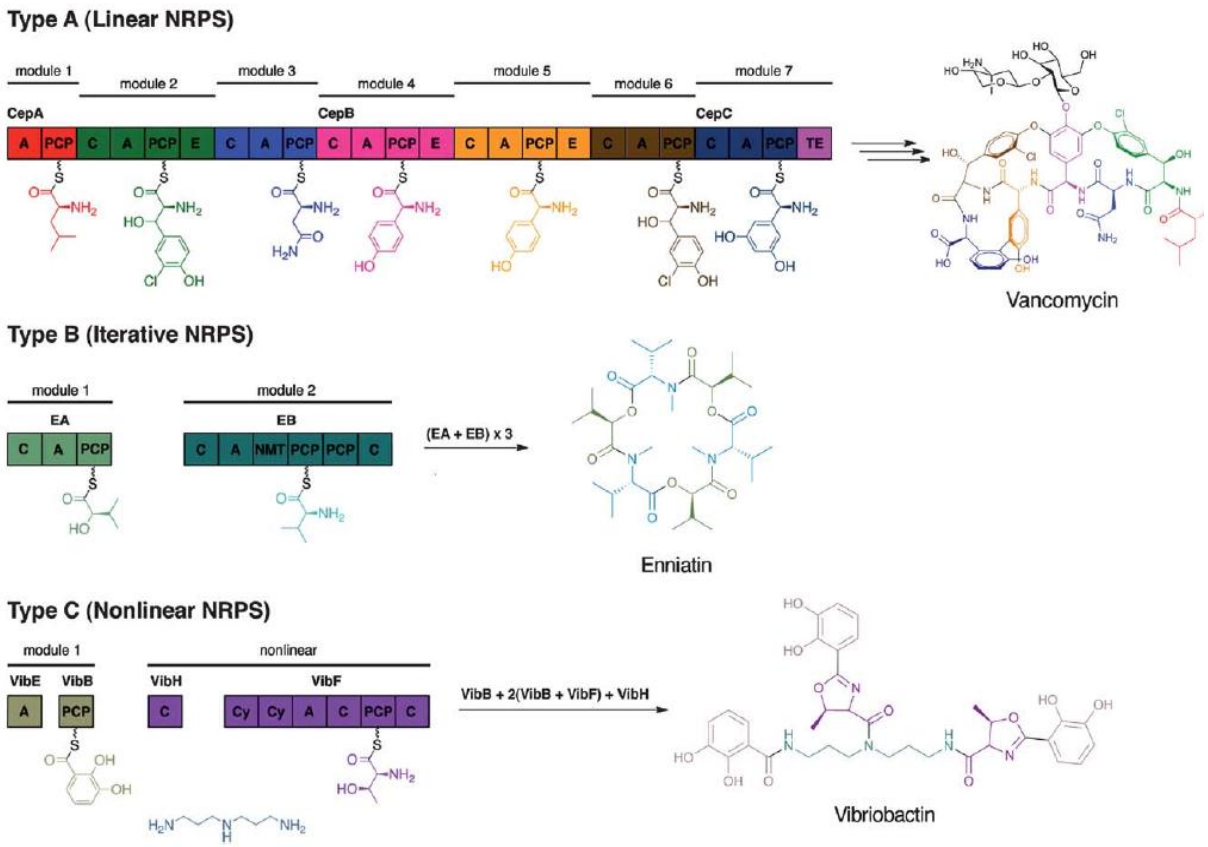


Figure 1.18. Modular types of NRPS; Linear, Iterative and Nonlinear (Figure taken from Hur *et al.*, 2012 after permission).

The condensation domain, which has two similar sub-domains, catalyzes the formation of the peptide bond between the amino acid bound to the phosphopantetheinyl arm of the PCP in the same module with that in the adjacent module and release of the phosphopantetheinyl arm (Figure 1.19) (Condurso and Bruner, 2012).

The function of the first module (Initiation module) is just to activate the substrate because it lacks a C domain, while in the C-terminus the termination module contains the Thioesterase (TE) domain (Figure 1.19), which catalyses the release of the assembled peptide chain from phosphopantetheinyl thioester resulting in either a linear peptide or cyclic peptide using an internal nucleophile (Kohli and Walsh, 2003).

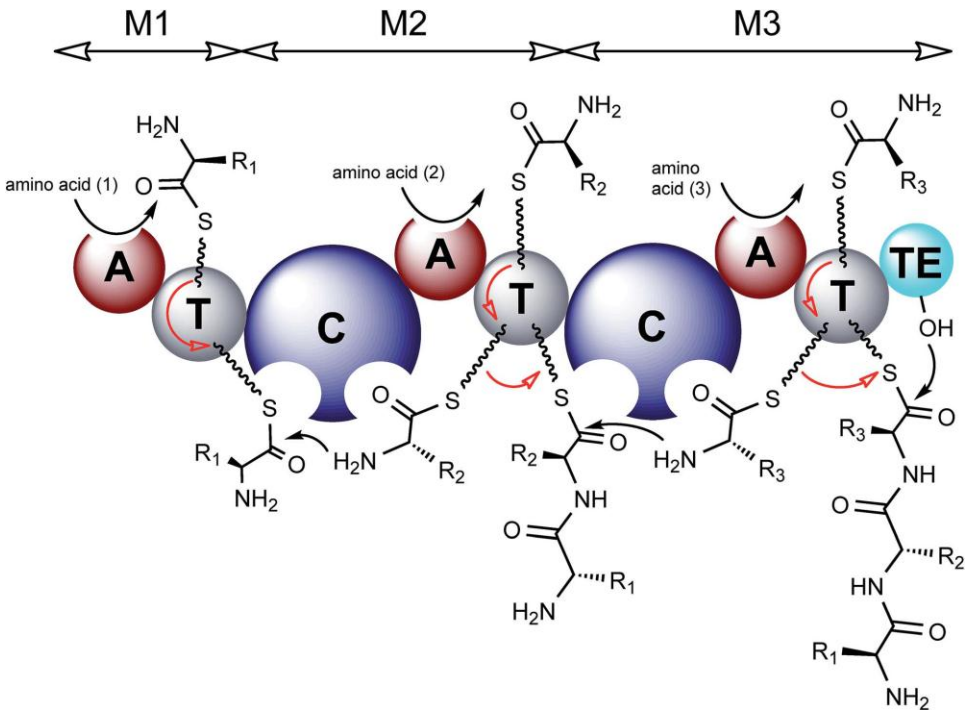


Figure 1.19. The essential domains of NRPS model. A: adenylation domain, T: thiolation domain, C: condensation domain, TE: thioesterase domain (Figure taken from Winn *et al.,* 2016).

In addition to the essential three core domains, there are auxiliary domains which act as tailoring enzymes and impart the NRPS products extra modifications. These modifications occur either to the individual residue before peptide bond formation or after the peptides have been synthesized (Hur *et al.,* 2012). These domains include Epimerization (E), Cyclization (Cy) and Methyltransferase (MT) domains. The epimerization domain located at the C-terminus of PCP is responsible for Cα racemisation. The cyclization domain is responsible for heterocyclization of serine/threonine and cysteine residues. The methyltransferase domain uses S-adenosylmethionine (SAM) as a cofactor to introduce *N*- or *C*-methylation of the building block. These tailoring enzymes either exist as the standalone enzyme, *in trans,* or inserted within the essential enzymes, *in cis* (Walsh *et al.,* 2001).

1.5.2. Biotechnology applications and challenges

Engineered biosynthesis of NRPSs was applied to improve the pharmacokinetics and ADME properties of the NRP. The early work focused on introducing non-natural amino acids into the growing peptides via feeding experiments and mutasynthesis. Cyclosporine A was produced by feeding the WT *Tolypocladium inflatum* with D-serine which replaces the natural D-alanine at position 8. However, the problem associated with this approach is the competition between the natural and alternative substrate consequently leads to a low yield of the desired product (Winn *et al.,* 2016). This problem was overcome by introducing a specific mutation into the biosynthetic pathway

which changes the preference of the catalytic domain toward the substrate. For example, the Lys278Gln mutation of the 10 model adenylation domain of CDA NRPS was fed with synthetic precursor Gly-mGln, and successfully incorporated into the product (Thirlway *et al*. 2012). This technique was not feasible in some NRPS because of the high fidelity of the downstream domains.

The modular nature of the non-ribosomal peptide synthetase became the subject of many studies and intensive investigation to manipulate the biosynthetic pathway. Many approaches have been applied to achieve these purposes including domain swapping, module deletion or swaping and module fusion. Understanding the specificity and structure of the domains and modules is the key tool to generate new templates. In general, four strategies have been suggested to change individual domains to produce new products (Figure 1.20).

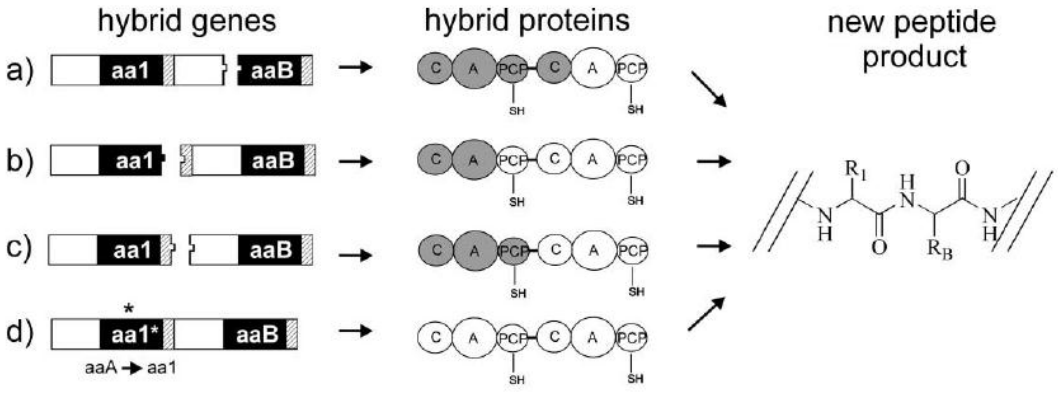


Figure 1.20. Different strategies to create hybrid NRPS, a and b: domains fusion, c: module fusion and d: point mutation (Figure taken from Finking and Marahiel, 2004 after permission).

These strategies include fusion between different domains and modules, and domain specificity alteration. The most remarkable feature that enhances the fusion is the linker regions that connect the modules. These regions consist of about 15 amino acids and display no conserved residues with small and hydrophilic amino acids rendering these regions to be the ideal positions for module fusions. Biochemical studies on the tyrocidine synthetase showed the possibility of module fusion between the second and ninth or tenth modules resulting in hybrid NRPS proteins (Figure 1.20c)(Mootz, Schwarzer & Marahiel, 2000). In surfactin synthetase, deletion of the second module results in a fusion between the first and third modules producing a satisfactory amount of product shortened by one amino acid (Mootz, *et al.,* 2002). The linker region between A and PCP domains gives the possibility to alter a single domain (Doekel & Marahiel, 2000). Point mutations can change the specificity of an A domain (Figure 1.20d) to produce a new compound, but this approach may be limited because of the specificity of the downstream C and TE domains (Eppelmann, Stachelhaus, Marahiel, 2002 ; Challis and Naismith, 2004).

Trans-expression of the tailoring enzymes in the heterologous host is another approach to modify NRP structure. *Trans*-expression of the r*am29* gene, which is responsible for mannosylation of ramoplanin, in eduracidin-producing *Streptomyces fungicidicus* resulted in mannosylation of eduracidins (Winn *et al.,* 2016).

However, the manipulation approaches to reprogramming the NRPS and production of desired compounds is hindered by insufficient knowledge about protein interactions and substrate selectivity. The substrate recognition cannot be predicted by the structure of

the single domain as interdomain interactions play a major role in substrate recognition (Lautru and Challis, 2004). Another challenge is the high specificity of A domain which tolerates very little variation with the consequent reduction in the diversity of the products. Domain alteration and point mutation to change the substrate selectivity have been suggested to solve this problem. This faced another difficulty since the downstream domains, C and TE domains, also show substrate selectivity. Finally, the unknown role of the linker regions between modules and the different conformations of the products which depend on the nature of the TE domain should be investigated by structural and biochemical studies (Schwarzer, Finking and Marahiel, 2003; Challis and Naismith, 2004).

1.6. Objectives of this study

This study aims to understand and characterize PKSs and NRPSs which are responsible for mupirocin and thiomarinol biosynthesis in an attempt to create new derivatives with desired properties. Mupirocin (Pseudomonic acid A) is the topical antibiotic used against MRSA for decades. Resistance against mupirocin has been evolved by bacteria which may threaten its future use. Thiomarinol A is structurally similar to mupirocin but is more potent which makes it a good substitute. Since its discovery, thiomarinol has become the subject of much scientific research trying to bring it into clinical use. Although many studies showed that mutagenesis and mutasynthesis experiments can be used to produce new compounds or derivatives of thiomarinols that

have potent antibiotic properties, there are a number of significant questions related to the biosynthesis of these compounds that need to be answered. This will clarify the role of each related element consequently paving the way to make the desired genetic manipulation which improve the products or increase the yield. This study aims to answer some of these questions in an attempt to add new aspects to the previous studies.

The general backbone structures of mupirocin and thiomarinol are similar (Figure 1.21). The monic acid and the pyran ring are quite similar to those in marinolic acid. The molecular similarities between mupirocin and thiomarinol are consistent with the similarity between the gene clusters as revealed by DNA sequence analysis and their proposed biosynthetic pathways based on this information (Figure 1.22). These similarities between the products and gene clusters may be helpful to predict the function of unknown genes in the *tml*/*tmp* cluster depending on the *mup* gene cluster which already has been characterized and most of the gene function was identified.

Despite similar structures between thiomarinol and mupirocin, there are also four distinct differences between product molecules. These include the absence of 10,11 epoxidation, the presence of 4-hydroxylation, the shorter hydroxy-fatty acid and pyrrothine molecule attached to the hydroxyl group in thiomarinol (Figure 1.21). Mutagenesis and complementation studies identified genes responsible for 10,11 epoxidation and pyrrothine molecule while the presence of 4-hydroxylation and why the fatty acid chain is shorter are still not clear.

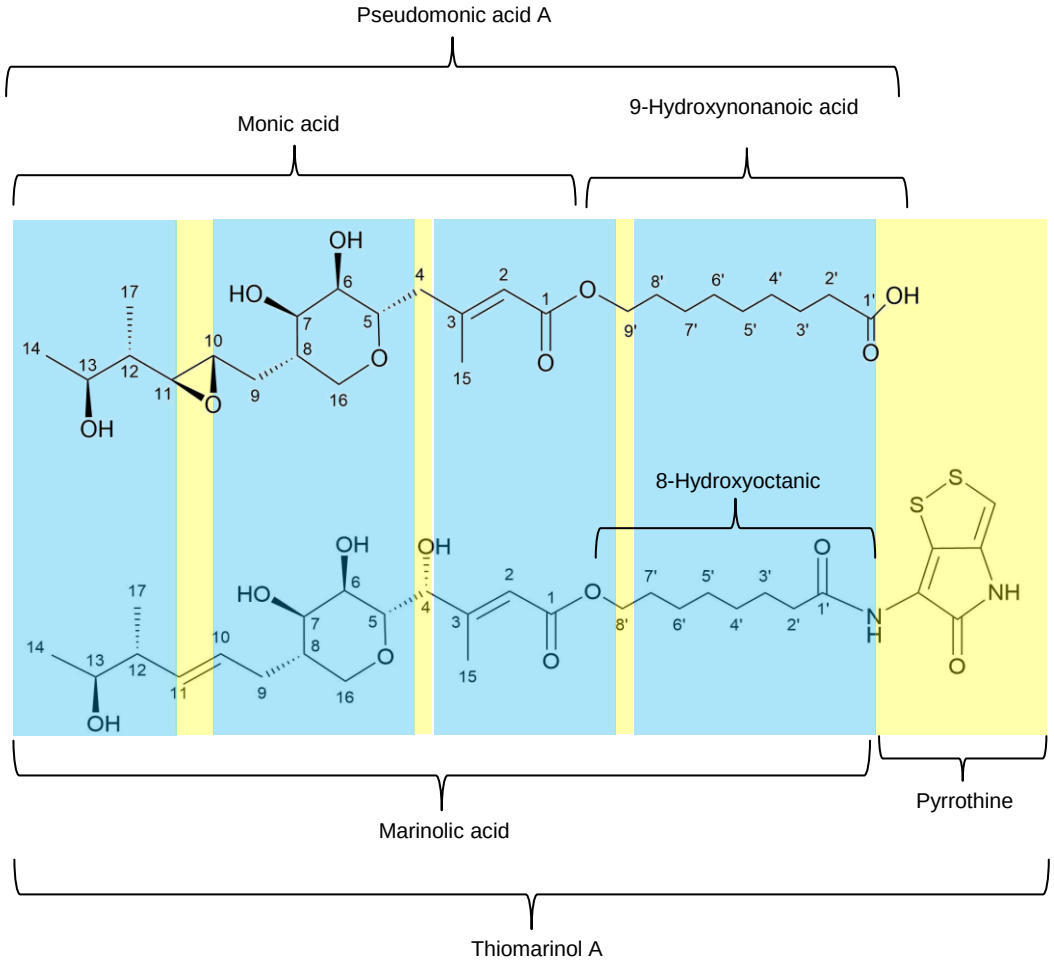


Figure 1. 21. The similarity and differences between mupirocin (PA-A) and thiomarinol A. The similarity and the differences highlighted with blue and yellow shadow respectively. The differences include the absence of 10,11 epoxidation, the presence of 4-hydroxylation, the shorter hydrox-fatty acid and the hydroxyl group is replaced by pyrrothine molecule.

Understanding these differences might allow modification of the core of mupirocin and other molecules in subtle ways which may improve their properties. Outside the pH range of 4-9, the 7-hydroxyl group attacks 10,11 epoxide irreversibly and reduces its activity (Thomas *et al.,* 2010). Mutational analysis revealed that the gene knockout of

Δ*mmpE/OR* results in pseudomonic acid C which lacks the 10,11 epoxide. This makes pseudomonic acid C more stable than pseudomonic acid A and may improve clinical usage. Feeding experiments with *Pseudoalteromonas spp* SANK73390 showed that the main component of mupirocin PA-A could be 4-hydroxylated (Gao *et al*., 2014). However, the enzyme responsible for this hydroxylation was not identified and the significance of thiomarinol hydroxylation was not investigated. Marion and co-workers proposed that the 4-hydroxylation may enable thiomarinol to form more hydrogen bonds with the target, and the addition of this OH to the analogues may improve the antibacterial activity (Marion *et al.,* 2009). Chapter 2 therefore focuses on the investigation of the enzymes that direct the 4-hydroxylation of thiomarinol, studying their specificity and the ability to hydroxylate other analogues including pseudomonic acids via heterologous expression of the encoding genes, and the role of this hydroxyl group in thiomarinol and the hydroxylated version of the analogues activity.

The pyrrothine molecule is a non-ribosomal peptide consisting of two cysteine residues which is assembled by seven orfs designated as *holA* to *holG*. HolA is similar to a non-ribosomal peptide synthetase (NRPS). Pyrrothine imparts potency to thiomarinol and its removal reduced antibacterial activity significantly. The intriguing feature of *holA* gene, which is responsible for pyrrothine biosynthesis, is that it encodes only a single Adenylation (A) domain, Condensation (C) domain and Peptidyl Carrier Protein (PCP) while pyrrothine requires pairs of domains for each.

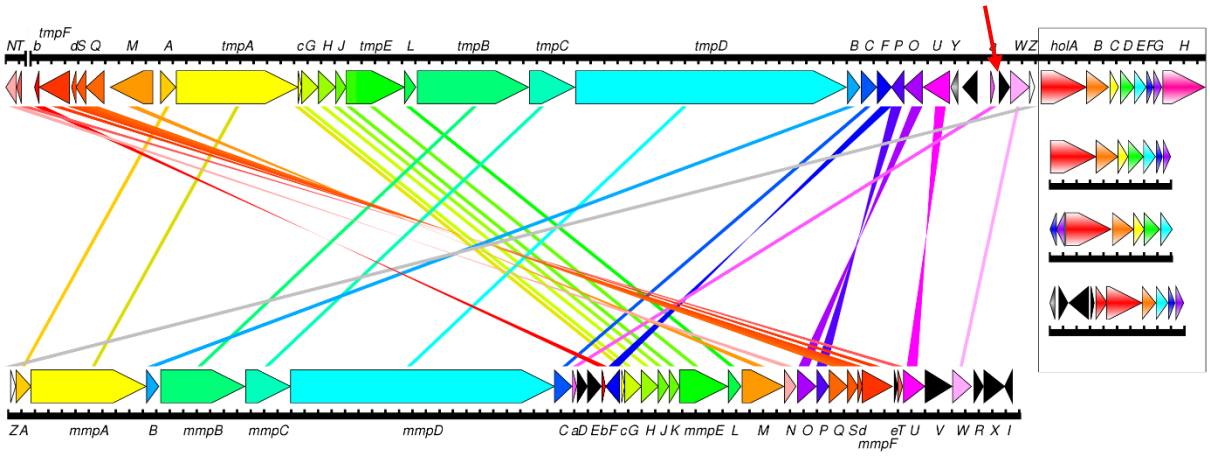


Figure 1.22. The comparison between the gene cluster of thiomarinol (upper line) and the gene cluster of mupirocin (lower line) showing similarity and differences. The *tmuB* gene is indicated by a red arrow which has no counterpart in *mup* cluster. The box to the right represents the related NRPS cluster which is responsible for pyrrothine molecule assembly in *Pseudoalteromonas* sp. SANK 73390 (Adapted from Fukuda *et al*, 2011 after permission). See figure 1.12 and 1.15 for more details about the genes.

This leads to the suggestion that HolA is a homodimer protein and works as an iterative NRPS in a similar manner to fungal NRPS (Fukuda *et al.,* 2011). To investigate more about this kind of NRPS, Chapter 3 presents the *in vivo* and *in vitro* characterizations plus biochemical assays of HolA protein in an attempt to understand the assembly of the pyrrothine molecule by investigating the protein dimerization and activity.

Thiomarinols display cytotoxicity in eukaryotic cells and effort has been made to explore this and to investigate the possible cell components which may be targeted by thiomarinols (Ahmed, 2010 unpublished). Untill recently, it is not clear which part of the thiomarinol is responsible for the loss of selectivity that was seen in PA-A. Thus, more

investigation on 4-hydroxylation and the pyrrothine molecule is needed in terms of the biological activity and cytotoxicity in eukaryotic cells. The overall work in this project includes genetics, bioinformatics and biochemical studies to understand the assembly of the thiomarinol molecule at the genetic level, and possible production of novel derivatives.

CHAPTER TWO

2. INVESTIGATION OF 4-HYDROXYLATION IN THIOMARINOL

2.1. Introduction

Hydroxylation is an essential step in many natural product pathways either as intermediate steps or as tailoring reactions. The hydroxylation of natural products contributes to structural diversification and alters their biological activity profiles. The hydroxylation of pseudomonic acid A at carbon 8 (as in pseudomonic acid B) reduces its anti-bacterial activity by around 60% compared to non-hydroxylated pseudomonic acid A (Cooper *et al.*, 2005b). Other examples of hydroxylation modifying activity include cyclosporine losing its immunosuppressive activity after regio-specific hydroxylation at the 4th N-methyl leucine but retaining its hair growth side effect which might allow it to be used as a hair growth-stimulating compound (Lee *et al.*, 2013).

The common enzymes that introduce hydroxyl groups to the product backbone are cytochrome P450 monooxygenases (CYP450), flavin-dependent mono- and dioxygenases, and Fe(II)/α-Ketoglutarate dioxygenases (Rix *et al*., 2002; Wu *et al.,* 2016). In addition to hydroxylation, some of these enzymes catalyze more than one reaction with different consequences. For example, during clavulanic acid biosynthesis in *Streptomyces clavuligerus*, clavaminate synthase (CAS) performs three separate reactions including hydroxylation of β-Lactam, oxidative cyclization of proclavaminate and desaturation of dihydroclavaminate to produce clavaminate (Lloyd *et al.,* 1999).

As described in Chapter one section 1.6, one prominent difference between mupirocin and thiomarinol is the 4-hydroxylation (Figure 1.21). Comparison of the clusters

identified the *tmuB* gene from the thiomarinol gene cluster, which lacks a homologue in the mupirocin cluster, as a potential candidate to perform the 4-hydroxylation of thiomarinol (Figure 1.15 and 1.22). The *tmuB* gene in the thiomarinol gene cluster encodes for a 29.5 kDa TmuB protein. As an initial prediction about TmuB function, protein-protein BLAST search was carried out which revealed that TmuB possesses sequence similarity to dioxygenase enzymes. Further screening of the thiomarinol cluster plus a protein-protein BLAST search predicted TmlZ (14.7kDa) as monooxygenase enzyme.

The aim of this study is to determine whether these genes in thiomarinol gene cluster are responsible for 4-hydroxlation, to determine the potential role of this hydroxylation in terms of the biological activity of the compound and in terms of the production profile and the physiological character of the producer. It would also be of interest to determine what are the possible analogues that could be hydroxylated by these genes and how this might alter the properties of products. This investigation might thus allow modification of the core of mupirocin and other molecules in subtle ways.

In an attempt to answer these questions, TmuB and TmlZ were subjected to complementation and mutational analyses followed by a combination of *in vivo* studies and *in vitro* biochemical characterization plus protein bioinformatics and structural modelling. The results shed new light on proteins of the phytanoyl-CoA dioxygenase family and specifically add to the toolbox for building and manipulating novel polyketide biosynthetic pathways.

2.2. Materials and methods

2.2.1. Bacterial strains and plasmids

The wild type thiomarinol producer *Pseudoalteromonas sp* SANK 73309 was used for DNA amplification of the *tmuB* and *tmlZ* genes, suicide mutagenesis and thiomarinol extraction. Wild type and mutant strains of *Pseudomonas fluorescens* NCIMB 10586 (Whatling *et al.*, 1995) were used for pseudomonic acids production and as the heterologous hosts. For plasmids transformation and gene cloning competent cell of *E.coli* DH5α (Hanahan, 1983) was used. *E.coli* S17-1(Simon *et al.*, 1983) was used to mobilize expression vectors that hold the target genes to the wild type and mutant strains. Bacterial strains and plasmids used in this study are listed in Tables 2.1 and 2.2 respectively.

Table 2.1. Bacterial strains used in this study.

Bacterial Strains	Genotype/Phenotype	Usage	Source
Bacillus subtilis 1604	TrpC2	Bioassay and minimal inhibitory concentration tests	Moir *et al*., 1979
Escherichia coli BL21	F^- ompT gal *dcm* lon *hsdS*$_B$(r_B^- m_B^-) λ(DE3 [*lacI* lacUV5-T7 gene 1 ind1 sam7 nin5])	Over-expression of protein by pET28a vector	Novagen
Escherichia coli BTH101	F-, *cya-99*, *araD139, galE15, galK16, rpsL1 (Str r)*, *hsdR2, mcrA1, mcrB1.*	Used as a detector bacterium to test protein-protein dimerization *in vivo*	Karimova & Ladant, 2005

Escherichia coli DH5α	*F⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96* deoR nupG Φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17(r_K^- m_K^+), λ–	Competent cell preparation and constructed vectors propagation	Gibco BRL
Escherichia coli DHM1	F-, *cya-854*, *recA1, endA1, gyrA96 (Nal r)*, *thi1, hsdR17*, *spoT1*, *rfbD1, glnV44(AS).*	Used as a detector bacterium to test protein-protein dimerization *in vivo*	Karimova *et al*, 2005
Escherichia coli S17-1	TpR Sm^R recA, thi, pro, hsdR-M^+RP4: 2-Tc:Mu: Km Tn7 λpir.	Competent cell preparation and bacterial conjugation	Simon *et al.*, 1983
NCIMB 10586 Δ*macpE*	*P. fluorescens* with *mupV* gene deletion	Heterologus hosts and pseudomonic acid B production	Cooper *et al.*, 2005b
NCIMB 10586 Δ*mmpE/OR*	*P. fluorescens* with *mmpE* gene deletion	Heterologus hosts and pseudomonic acid C production	Gao *et at*., 2014
NCIMB 10586 Δ*mupC*	*P. fluorescens* with *mupC* gene deletion	Heterologus hosts and mupirocin C production	Hothersall *et al.*, 2007
NCIMB 10586 Δ*mupF*	*P. fluorescens* with *mupF* gene deletion	Heterologus hosts and mupirocin F production	Hothersall *et al.*, 2007
NCIMB 10586 Δ*mupO*	*P. fluorescens* with *mupO* gene deletion	Heterologus hosts and pseudomonic acid B production	Cooper *et al.*, 2005b
NCIMB 10586 Δ*mupU*	*P. fluorescens* with *mupU* gene deletion	Heterologus hosts and pseudomonic acid B production	Cooper *et al.*, 2005b
NCIMB 10586 Δ*mupV*	*P. fluorescens* with *mupV* gene deletion	Heterologus hosts and pseudomonic acid B production	Cooper *et al.*, 2005b
NCIMB 10586 Δ*mupW*	*P. fluorescens* with *mupW* gene deletion	Heterologus hosts and mupirocin W production	Cooper *et al.*, 2005b
Pseudoalteromonas sp SANK 73390	Wild Type: Thiomarinol producing bacteria	Thiomarinol extraction, gene amplification and manipulation	G.T. Banks
Pseudomonas fluorescens NCIMB 10586	Wild type mupirocin producer	Heterologus hosts and pseudomonic acid A production	Whatling *et al*., 1995
Staphylococcus aureus MRSA NCTC 12493	mecA methicillin resistant strain used standardly by BSAC for susceptibility test	Used as a detector bacterium for minimal inhibitory concentration	Heartlands Hospital

2.2.2. Growth of bacterial strains and culture conditions

Marine agar and marine broth for *Pseudoalteromonas* SANK 73390: provided by Difico^{TM} were used to *Pseudoalteromonas* SANK 73390 propagation. These media contain all the nutritional requirements for *Pseudoalteromonas* SANK 73390 growth. The media were prepared according to manufacture instructions and sterilized by autoclaving (121 ^{o}C at a pressure of 15 pounds per square inch for 15 min).

Minimal medium (M9 medium) for *Pseudomonas fluorescens* NCIMB 10586: This medium was used as a selective medium to inhibit *E.coli* S17-I growth after conjugal mating with *Pseudomonas flourescens* strains. To prepare 400 ml of this medium, 200 ml of 3% agar and 200 ml of M9 x2 medium were mixed and 400 µl of 0.1 M $CaCl_2$, 400 µl of 1 M thiamine-HCl, 400 µl of 1 M $MgSO_4$, 400 µl of Tetracycline (100 mg/ml) and 2ml of 40% glucose were added and poured into petri dishes.

Secondary stage medium (SSM) was used to produce mupirocin from *Pseudomonas fluorescens* strains, per liter, 25g soya flour, 2.5g spray dried corn liquor, 5g $(NH_4)_2$ SO_4, 0.5g $MgSO_4$ $7H_2O$, 1g Na_2HPO_4, 1.5g KH_2PO_4, 6.25g $CaCO_3$ were mixed with 850 ml distilled water, dissolved by stirring and heating, the pH adjusted to 7.5 by NaOH, the volume completed to 1 liter and autoclaved. Prior to setting up cultures, 40% sterile glucose was added.

Luria (L) Broth and L agar were used to culture *E.coli* DH5α and *E.coli* S17-1. One liter of LB broth consists of: 5g yeast extract, 10g tryptone, 5g NaCl and (12g agar in case of

agar media). The media was sterilized by autoclaving (121 ^{o}C at a pressure of 15 pounds per square inch for 15 min).

Table 2.2. Plasmids used and constructed in this study.

Plasmid's name	Features	Usage	Source
pAKE604	Amp^{R}, Kan^{R}, Bsac, LacZα, ori, MCS	Suicide mutagenesis	El-Sayed *et al.*, 2001
pGEM-T Easy	Amp^{R}, LacZα PCR cloning vector	Sequencing of the recombinant DNA	Promega
pET28a	T7 promotor, His tag, Kan^{R}, LacZα, lacI,MCS,Ori	Protein expression	Novagen
pJH10	Low copy number, $tetracycline^{r}$,	Gene expression in hetrologues strains	El-Sayed *et al.,* 2003
pHHM01*	*tmuB* gene cloned to pGEM-T Easy	For sequencing purposes	This study
pHHM02	*tmlZ* gene cloned to pGEM-T Easy	For sequencing purposes	This study
pHHM03	*tmuB* gene cloned to pET28a	To overexpress TmuB protein as N-terminus His tag	This study
pHHM04	*tmuB* gene cloned to pJH10	To express TmuB protein in hetrologues strains	This study
pHHM05	*tmlZ* gene cloned to pJH10	To express TmlZ protein in hetrologues strains	This study
pHHM06	*tmuB* gene with point mutation I109N cloned to pJH10	To express TmuB protein with this mutation in *P. fluorescens* strains	This study
pHHM07	*tmuB* gene with point mutation I109N cloned to pAKE604	To introduce point mutation I109N in WT SANK	This study
pHHM08	*tmuB* gene with point mutation I109N cloned to pET28a	To overexpress TmuB protein with this mutation as N-terminus His tag	This study
pHHM09	*tmuB* gene with point mutation I109V cloned to pJH10	To express TmuB protein with this mutation in *P. fluorescens* strains	This study

pHHM10	*tmuB* gene with point mutation L141N cloned to pJH10	To express TmuB protein with this mutation in *P. fluorescens* strains	This study
pHHM11	*tmuB* gene with point mutation L141V cloned to pJH10	To express TmuB protein with this mutation in *P. fluorescens* strains	This study
pHHM12	*tmuB* gene with point mutation L141S cloned to pJH10	To express TmuB protein with this mutation in *P. fluorescens* strains	This study
pHHM13	*tmuB* gene with point mutation R69K cloned to pJH10	To express TmuB protein with this mutation in *P. fluorescens* strains	This study
pHHM14	*tmuB* gene with point mutation K105E cloned to pJH10	To express TmuB protein with this mutation in *P. fluorescens* strains	This study
pHHM15	*tmuB* gene with double point mutation R69K&L141V cloned to pJH10	To express TmuB protein with this mutation in *P. fluorescens* strains	This study
pKT25 and pKNT25	Low copy number, pSU40, lac promoter, Kan^r, MCS, T25	In vivo protein-protein interaction	Euromedex
pUT18 and pUT18C	high copy number, pUC19, lac promoter, Amp^r, MCS, T18	In vivo protein-protein interaction	Euromedex
pKT25-zip	Low copy number, pSU40, lac promoter, Kan^r, MCS, T25 fused to leucine zipper	Used as a positive control in B2H	Euromedex
pUT18-zip	high copy number, pUC19, lac promoter, Amp^r, MCS, T18 fused to leucine zipper	Used as a positive control in B2H	Euromedex
pHHM16	*tmuB* gene cloned to pKT25	To test protein dimerization *in vivo*	This study
pHHM17	*tmuB* gene cloned to pUT18	To test protein dimerization *in vivo*	This study

pHHM18	*tmuB* gene cloned to pKNT25	To test protein dimerization *in vivo*	This study
pHHM19	*tmuB* gene cloned to pUT18C	To test protein dimerization *in vivo*	This study

* All the mutants of *tmuB* gene were cloned to pGEM-T Easy for sequencing purposes.

Table 2.3. antibiotics used in this study.

Antibiotic name	Stock concentration mg/ml	Working concentration µg/ml	Solvent
Ampicillin	100 mg/ml	100 µg/ml	Water
Kanamycin	50 mg/ml	50 µg/ml	Water
Tetracycline	15 mg/ml	15 µg/ml	70% Ethanol
Streptomycin	100 mg/ml	100 µg/ml	Water
Nalidixic acid	30 mg/ml	30 µg/ml	Water, 0.3M NaOH

2.2.3. Polymerase Chain Reaction PCR

Primers were designed manually and online software Netprimer and OligoAnalyzer 3.1 were used to check primer quality. Restriction sites and extra three bases were added to the 5' end of the primers to facilitate digestion and recombination. The primers were synthesized by Alta Bioscience in the University of Birmingham. The standard PCR reaction was carried out using high fidelity DNA polymerase from (Velocity provided by BIOLINE and Q5 High Fidelity DNA polymerase by NEB). The DNA template from wild

type bacterium was prepared using boiled cell preps. The primer sequences are listed in Table 2.4.

Table 2.4. Primers used in this study.

Name	Function	Sequence 5'⟶ 3'	Restriction site	Tm*
TmlZF	TmlZ expression on pJH10	TCAGAATTCATGTTGGTCAAGAGTGAAACTT	*EcoRI*	53.28
TmlZR		TCGTCTAGATGCTTAATCAAACTCTAACAACTC	*XbaI*	53.8
TmuBF	To clone *tmuB* into pJH10	GCAGAATTCATGGATAGTTTGCAGTCATTTA	*EcoRI*	52.4
TmuBR		CTGTCTAGAGAATCGTTGCTAATGCCA	*XbaI*	52.19
FTmuB-Exp	To clone *tmuB* into pET28a	CTACATATGATGGATAGTTTGCAGTCATTTA	*NdeI*	52.4
RTmuB-Exp		TAGCTCGAGGAATCGTTGCTAATGCCA	*XhoI*	52.19
FSalTmuB	To clone *tmuB* into pAKE604	GCAGTCGACATGGATAGTTTGCAGTCATTTA	*SalI*	52.4
TmuBR		CTGTCTAGAGAATCGTTGCTAATGCCA	*XbaI*	52.19
I109V	Reverse primer to create insert with point mutation I109V in *tmuB*	GGAAAAATTTCTACTCCCACAGACTTATC	-	65
FI109N	Forward primer to create insert with point mutation I109N in *tmuB*	ATCAAATAGATAAGTCTGTGGGAAACGAAATTTTTCCAGGAGAAAGCAGG	-	82
TmuBPMF2	To create insert with point mutation L141S in *tmuB*	GTTCAGCTGTTCGGTGGCATTGG	-	68
TmuBPMR1		CCAATGCCACCGAACAGCTGAAC	-	68

FL141V	To create insert with point mutation L141V in *tmuB*	AGGCATGAATTTAATGTTCA GCTGTGTGGTGGCATTGGA	-	82
RL141V		TCCAATGCCACCACACAGC TGAACATTAAATTCATGCCT	-	82
FL141N	To create insert with point mutation L141N in *tmuB*	GGCATGAATTTAATGTTCAG CTGTAACGTGGCATTGGAT GATTTTACAAGC	-	85
RL141N		GCTTGTAAAATCATCCAATG CCACGTTACAGCTGAACATT AAATTCATGCC	-	85
R69K	Reverse primer to create insert with point mutation R69K in *tmuB*	ATACTCGAACAGCTTTAAAG CTCTCCG	-	68
K105ER	Reverse primer to create insert with point mutation K105E in *tmuB*	ATTCCCACAGACTCATCTAT TTGAT	-	56
TmuB-pKT25	Forward primer to clone into pKT25	ACGCTGCAGGGATGGATAG TTTGCAGTCATTTA	*PstI*	58.38
TmuB-pUT18C	Forward primer to clone into pUT18C	ACGCTGCAGGATGGATAGT TTGCAGTCATTTA	*PstI*	54.5
TmuB-pUT18	Reverse primer to clone into pUT18 & pKT25	CAGTCTAGAACCAGGATAA CGTCTGTCTG	*XbaI*	51.94

*The Tm was calculated by online software NetPrimer from primer Biosoft. The PCR annealing temperature is lower by 2 ^{o}C.

The PCR reaction was carried out using SensoQuest thermocycler and set up with the following program:

Denaturation step: 98 ^{o}C for 2 minutes 1 cycle

Denaturation step: 98 ^{o}C for 30 sec

Annealing step: 50 ^{o}C for 30 sec* } 35 cycles

Extension step: 72 ^{o}C for 30 sec*

Extension step: 72 ^{o}C for 7 minutes 1 cycle

* The annealing temperature and extension time varies according to primers, and DNA template length and G-C content.

2.2.4. DNA isolation and manipulation

2.2.4.1. DNA and plasmid extraction

Chromosomal DNA was isolated using the boiled cell miniprep method and used as a DNA template in PCR reaction. About 10 colonies of the experiment bacterium were picked up by sterile pipette tip and resuspended in 50µl of sterile distilled water. The mixture was then boiled for 10 minutes and centrifuged for 5 minutes (1300 xg, at 4 ^{o}C). The resulting supernatant was transferred to a new sterile tube and kept on ice.

Plasmid was isolated using the ISOLATE II Plasmid mini Kit provided by BIOLINE. A single colony was picked up and inoculated into 5 ml L-broth and incubated at 37 ^{o}C for 16 hours with shaking at 200rpm. The principles include alkaline SDS lysis method (Birnboim and Doly 1979). The isolation procedure was carried out according to the user's manual provided with the kit. The general outline of the procedure includes cell

lysis to release the DNA material, DNA binding to silica membrane, washing the column to eliminate salts and endonucleases, and finally collecting the DNA by elution buffer.

2.2.4.2. Restriction enzyme digestion

The restriction enzymes were purchased from New England Biolab (NEB). The total volume of the digestion reaction was 20µl-50µl depending on the downstream application. The amount of enzymes added to the reaction was 1µl per 1µg of DNA in accordance with the manufacture's instructions. The mixture was incubated at 37 ^{o}C in a water bath for 2-3 hours. The restriction enzymes were deactivated either by heating or by gel purification depending on the type of enzyme.

2.2.4.3. DNA ligation

The ligation between DNA fragments and vectors was carried out using T4 DNA ligase provided by Invitrogen. The total volume of the ligation reaction was 20 µl, using 0.5 µl of the enzyme. The insert : vector ratio was 3:1 using W_I /S_I : W_V /S_V formula (W: DNA weight in ng, S: DNA size) to find the appropriate amount of DNA to add to the ligation reaction. The mixture was incubated overnight at 4 ^{o}C, then different amounts were used to transform the competent bacteria, *E.coli* DH5α.

2.2.4.4. Agarose gel electrophoresis

Agarose gel electrophoresis was used to examine the size of DNA (PCR product, digested DNA and plasmids) and for purification purposes. 1% agarose gel was prepared by dissolving one gram of agarose powder in 100 ml 1x TAE buffer (40 mM tris-acetate and 1mM EDTA pH8) with heating. The mixture was cooled down and 2 µl of ethidium bromide was added to enable DNA visualization under UV light. The gel tray was prepared, suitable combs were inserted and the gel mixture was poured gently, and left at room temperature to solidify. Before loading the DNA sample onto the gel, loading buffer (0.25% w/v bromophenol blue and 15% w/v ficoll in water) was added in a 1:5 ratio of the DNA sample. 5 µl of marker DNA (1 kb) was loaded in the lane next to the samples to estimate the size of DNA samples. The gel was run for 45 minutes at 110 V in 1x TAE buffer.

2.2.4.5. DNA extraction from gel

GFX PCR and Gel Band Purification Kit from GE Healthcare was used to extract and purify DNA from the agarose gel using supplied solutions. The correct bands of DNA on the gel were cut with a scalpel under ultraviolet light and placed into a 1.5 ml microfuge tube. The gel slice was dissolved by adding 10 µl of capture buffer to each 10 µg of the gel slice and incubated at 60 ^{o}C in water bath for 15-30 minutes with mixing by inversion every 5 minutes. The capture buffer sample was transferred to a GFX microspin column and collection tube, incubated at room temperature for 1 minute and centrifuged for 1

minute at 16000 x g. The flow through discarded, the GFX microspin column washed using 500 µl by wash buffer and dried by centrifugation for 1 minute at 16000 x g. The GFX microspin column was transferred to a clean 1.5 microfuge tube, 30 µl of elution buffer was added to the column membrane, incubated at room temperature for 2 minutes and centrifuged for 1 minute at 16000 x g. The GFX microspin column removed and the eluted sample stored at -20 ^{o}C.

2.2.5. Preparation of the competent cells

E. coli $DH_{5\alpha}$ and S17-1 were used to prepare competent cells according to the calcium chloride method (Cohen *et al.*, 1972). The *E. coli* was plated out on the LB agar and then a single colony was picked up and inoculated into 5 ml L-broth, incubated at 37 ^{o}C / 200 rpm for 16 hours. From this, 1 ml was used to inoculate 100 ml of L-broth, incubated at 37 ^{o}C / 200 rpm. The growth was monitored and measured by spectrophotometer untill the cell density reached OD_{600}= 0.4-0.6. The culture was centrifuged (5000 rpm, 7 minutes at 4^{o}C). The cell pellet was resuspended using chilled 100 mM calcium chloride (2 ml per 5 ml culture) then incubated on ice for 30 minutes. The resuspeneded cells were centrifuged again (as stated above) and the cell pellet resuspended in 100 mM $CaCl_2$ 15% glycerol. 100 µl of this was aliquoted into sterile 1.5ml microfuge tubes, flashed with liquid nitrogen and stored at -80 ^{o}C.

2.2.6. DNA transformation

The recombinant vector was transformed to the competent cells according to the standard heat-shock protocol (Cohen *et al.*, 1972). 5 µl to 10 µl of ligation mixture were added to the 100 µl of competent cells in sterile 1.5ml microfuge tube and mixed gently. This mixture was incubated on ice for 30 minutes, then heat shocked at 42 ^{o}C in a water bath for 2 minutes. This enhances the entry of the recombinant vector into the competent cells. To recover the transformed cells, 1 ml of L-broth was added, gently mixed and incubated at 37 ^{o}C for 1-2 hrs. 100-150 µl of the recovered cells were plated out on selective media and incubated overnight at 37 ^{o}C.

2.2.7. Bacterial conjugation

E.coli S17-1 (donor) was used to mobilize the constructed vector to the recipient bacteria. One ml of overnight culture of each of the donor and recipient were mixed in universal bottle. This mixture was thenpassed through a sterile filter by syringe then the filter was placed on the appropriate agar plate without selection and incubated overnight at appropriate temperature. The bacteria from the filter were resuspended in one ml of (saline or broth according to the recipient bacteria) and serial dilutions were made from 10^{-1} to 10^{-5}. From each dilution 100 µl was plated out on appropriate selective media and incubated for 2-3 days.

2.2.8. A-Tailing of PCR product

The blunt-end PCR product was ligated with pGEMT-Easy vector for sequencing purposes. This vector is sold as linear molecules with T-overhang, so an A-overhang was added to the PCR product by mixing 6µl DNA, 1µl $MgCl_2$ (50mM), 1µl dATP(2mM), 1µl 10x buffer (200 Tris-HCl, pH 8.4 and 500mM KCl) and 1µl of Invitrogen Taq DNA polymerase (5U/µl) in a clean 1.5 ml microfuge tube and incubated at 70 ^{o}C for 30 minutes. 4µl of this was mixed with 1µl of pGEMT-Easy vector provided by Promega, 6µl buffer and 1µl ligase enzyme, and incubated overnight at 4 ^{o}C. This was transformed to competent *E.coli* DH5α and plated on selective media with IPTG and X-gal .

2.2.9. DNA sequencing

DNA sequencing was carried out by the Functional Genomics lab in the University of Birmingham using an ABI 3730 DNA analyzer. The procedure is based on the dideoxynucleotide chain termination method invented by Sanger *et al.,* in 1977. The total volume of the sample was 10 µl (200-500 ng DNA, 3.2 pmol primer, SDW to 10 µl) and 10 µl BigDye Reaction Mixture (Big Dye^{T} Terminator v 3.1 Cycle sequencing kit) was added. The results of the DNA sequencing were visualized using Chromas software and checked by BLAST alignment to compare the base sequence with original DNA bases.

2.2.10. Gene cloning and heterologous expression

The *tmuB* and *tmlZ* genes were amplified by PCR (as describe previously) with primers *TmuBF&TmuBR* and *TmlZF/TmlZR* respectively (Table 2.4) using genomic DNA from the WT *Pseudoalteromonas sp* SANK 73309 as a template. The genes were cloned into pGEM-T Easy vector (Promega) as A-tailed inserts and transformed into competent DH5α as described previously. The DNA samples were sent for sequencing in Functional Genomic lab in the University of Birmingham using Big Dye Terminator kit (PE-ABI). The target genes with the identical sequences were ligated into the expression vector, pJH10, as *EcoRI/XbaI* inserts. These constructed vectors were moved to the heterologous hosts via conjugation using *E.coli* S17-1.

2.2.11. Samples preparation for HPLC

Pure single colony of transformed *Pseudomonas flourescens* with constructed vector was set up into L-broth supplemented with appropriate antibiotics and incubated overnight at 25 C^{o} with shaking at 200 rpm. Secondary stage medium (SSM) was inoculated with 5% of the overnight culture (1.25 ml to 25 ml of SSM) and incubated at 22 C^{o} for 40-60 hours with shaking at 200 rpm. The culture was centriguged at 7000 x g for 7 min and the supernatant was transferred to 50 ml volume falcon tube. The pH was adjusted to 4.5 with HCl and filtered through a 0.2 µm filter.

2.2.12. HPLC analysis

HPLC was used to analyze products of *Pseudomonas flourescens* strains. Both solvent A (HPLC grade water) and solvent B (HPLC grade acetonitrile) were supplemented with 0.01% formic acid and degassed properly. Gilson® 321 pump was used to pump the solvent and the flow rate was 1 ml/min. Unipoint LC system software was used to perform HPLC. 100 µl of the sample was injected into C18 column (15cm x 4.6mm, 5µm) and 233 nm UV was used to detect the samples.

2.2.13. Mass spectrometry and NMR

The MS and NMR were done at the School of Chemistry, University of Bristol. 500 ml to 1 L of culture was incubated for target product purification in the same manner as for HPLC analysis. The cells were removed by centrifugation at 22,000 ×g for 20 min. The supernatant was then extracted by ethyl acetate (1:1) once, followed by an extra ethyl acetate extraction after the aqueous was acidified to pH 5.0. The two extracts were combined and ethyl acetate was evaporated in vacuo. The residue was collected by methanol for LCMS analysis and all the samples were dissolved in 3.0 ml methanol. Analytical samples were prepared by 10-fold dilution with methanol and analysed by LCMS using a Waters HPLC system (Waters 2545 Binary Gradient Module and Waters SFO System Fluidics Organizer). Detection was achieved by UV between 200 and 400 nm using a Waters 2998 diode array detector, and by simultaneous electrospray (ES) mass spectrometry using a Waters Quattro Micro™ API spectrometer

detecting between 150 and 600 m/z units. Chromatography (flow rate 1 mL·min-1) was achieved using Phenomenex Kinetex column (5 µm, C18, 100 Å, 4.6 × 250 mm). Solvents were: A, HPLC grade water containing 0.05% formic acid; B, HPLC grade acetonitrile containing 0.045% formic acid. Analytical gradients were as follows: 0 min, 5% B; 22 min, 60% B; 24 min, 95% B; 26 min, 95% B; 27 min, 5% B; 30 min, 5% B. LCMS purification of target compounds were performed on prep Phenomenex Kinetex column (5 µm, C18, 100 Å, AXIA 21.2 × 250 mm). Collection of target peaks was triggered by mass. Both positive (PI) and negative (NI) mode were employed for the characterization of the target compounds together with the UV absorption pattern. Comparative yields of target compounds were measured by ELS peaks.

2.2.14. Bioinformatic analysis

Bioinformatic analysis: DNA and protein sequences of *tmuB* were obtained from the GenBank database at NCBI (accession no. FN689524 & protein id: CBK62743). HHpred (http://toolkit.tuebingen.mpg.de/hhpred) (Soding, 2005) was used to identify homologous resolved protein structures. The first four structures with the highest identity score were PDB ID 4NAO (Havemann *et al.*, 2014), 4NMI (Widderich *et al.*, 2014), 4MHR (Hoppner *et al.*, 2014) and 2AX1 (McDonough *et al.*, 2005) with identity score (20%, 19%, 17% and 15%) respectively. DSSP (Joosten *et al.*, 2011) was used to determine the secondary structure of the templates. The secondary structure prediction server PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred) (Buchan *et al.*, 2013) and Jpred3

(http://www.compbio.dundee.ac.uk/jpred) (Cole, 2008) were used to predict TmuB secondary structure. ClustalX (Jeanmougin *et al.*, 1998) was used to set up the initial multiple alignments and the software Seaview (Galtier *et al.*, 1996) was used to manually refine the alignment guided by the known and predicted secondary structure. Modeller version 9.12 (Narayanan *et al.*, 2006) was used to generate a homology model of TmuB on the basis of the crystal structure of EasH (4NAO). The overall geometric and stereo-chemical qualities of the final modelled structure of TmuB were evaluated using multiple programs and online servers such as PROSESS (Wishart Lab, http://www.prosess.ca/index.php), WHAT_CHECK (Hooft *et al.*, 1997), and RAMPAGE (Lovell *et al.*, 2003). Evolutionary Trace Annotation (ETA) servers (Amin *et al*., 2013; Erdin *et al.,* 2013) and ESPript3 (Robert and Gouet, 2014) were used to identify the important residues for the protein catalytic function and those which are conserved or diverse and to find the suitable residues for mutation.

To predict the catalytic reaction of TmuB on the substrates, AutoDock4 and AutoDockTools 1.5.6 (Morris *et al.*, 2009; Sanner, 1999) were used to dock the substrates into the active site of the modelled TmuB and to analyze the interaction between the protein and the ligands. Rigid docking was performed using the genetic algorithm with default docking parameters. AutoGrid was used to define the docking searching area in active site and the grid size set to 54 X 52 X 52 points, grid spacing of 0.375 Å and grid center 71.5 X 30.09 X 34.89. Before running the AutoGrid and AutoDock, manually +2.0 charge was added to the Fe ion. Finally, the program PyMOL

(http://www.pymol.org) was used to modify the substrate (PA-A to PA-B) and to visualize the predicted structure of TmuB protein.

2.2.15. Point mutation using overlap extension

The mutagenesis overlap extension method (Vallejo *et al.*, 1994; Mergulhao *et al.,* 1999) was used to create point mutations. A pair of overlapping primers was designed in which the point mutation was located in the middle of the primers *RP1* 5'CCAATGCCACCGAACAGCTGAAC and *FP2* 5'GTTCAGCTGTTCGGTGGCATTGG. These two primers were used with previous external primers (Table 2.4) to generate two fragments each with point mutation at one end. To fuse these two fragments, a second round of PCR was run using *TmuBF* as a forward primer, downstream fragment as a reverse primer and the WT DNA genome as a template (The procedure illustrated in Figure 2.1). In this study, a modified method was used by designing one primer with the point mutation in the middle. This was used with normal forward or reverse primers to create a piece of DNA with the desired mutation. This DNA product was used with the forward or reverse primer to amplify the whole gene with the desired mutation (the procedure illustrated in Figure 2.2). The *tmuB* with the point mutation was cloned into the expression vectors and mobilized to *P. fluorescens* strains via conjugal mating using *E.coli* S17-1.

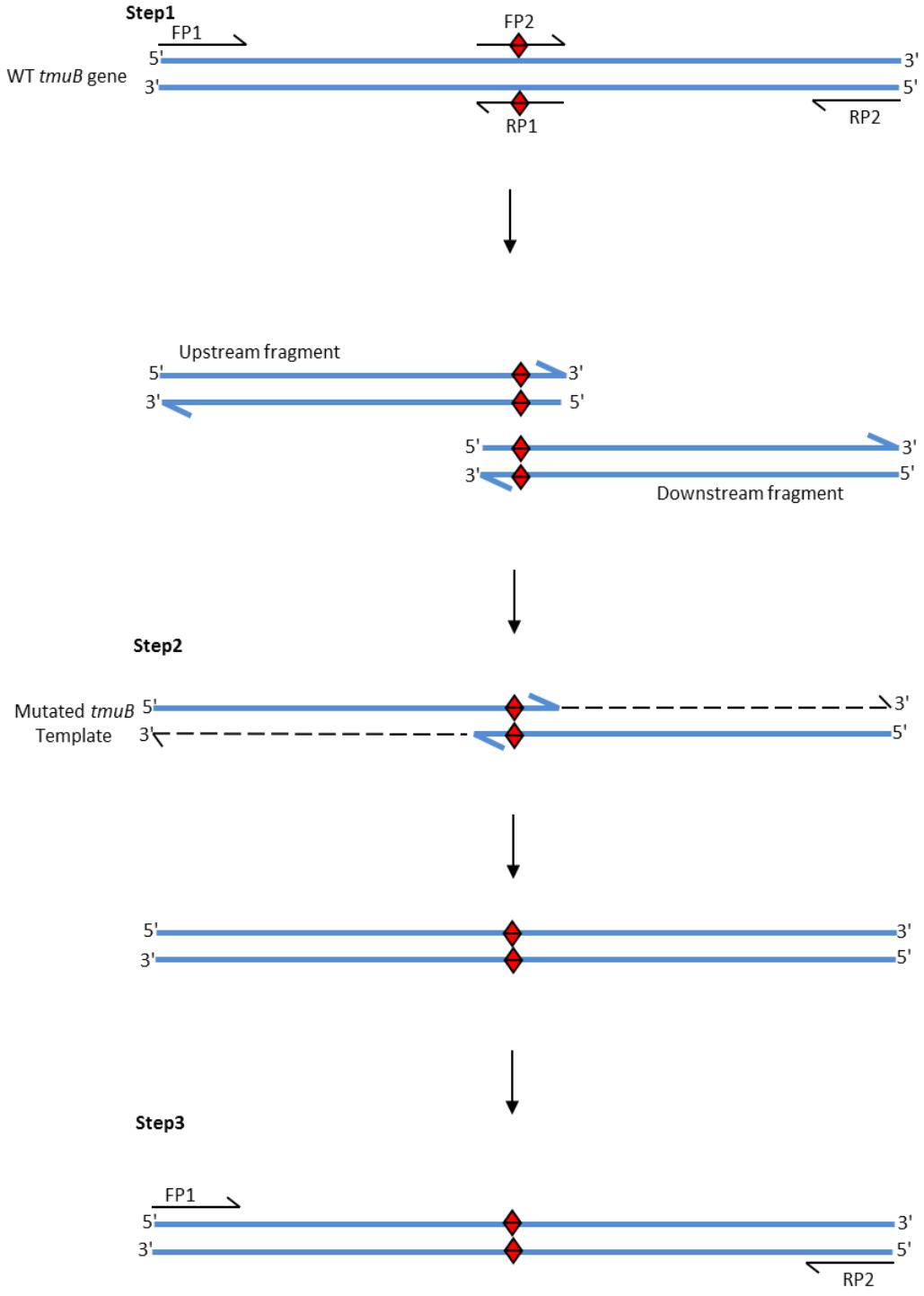


Figure 2.1. Overlap extension to create point mutation: step1, amplifying two overlapping fragments; step2, extension and joining the overlapped fragment to create the template with desired mutation; and step3, PCR amplifying using external primers

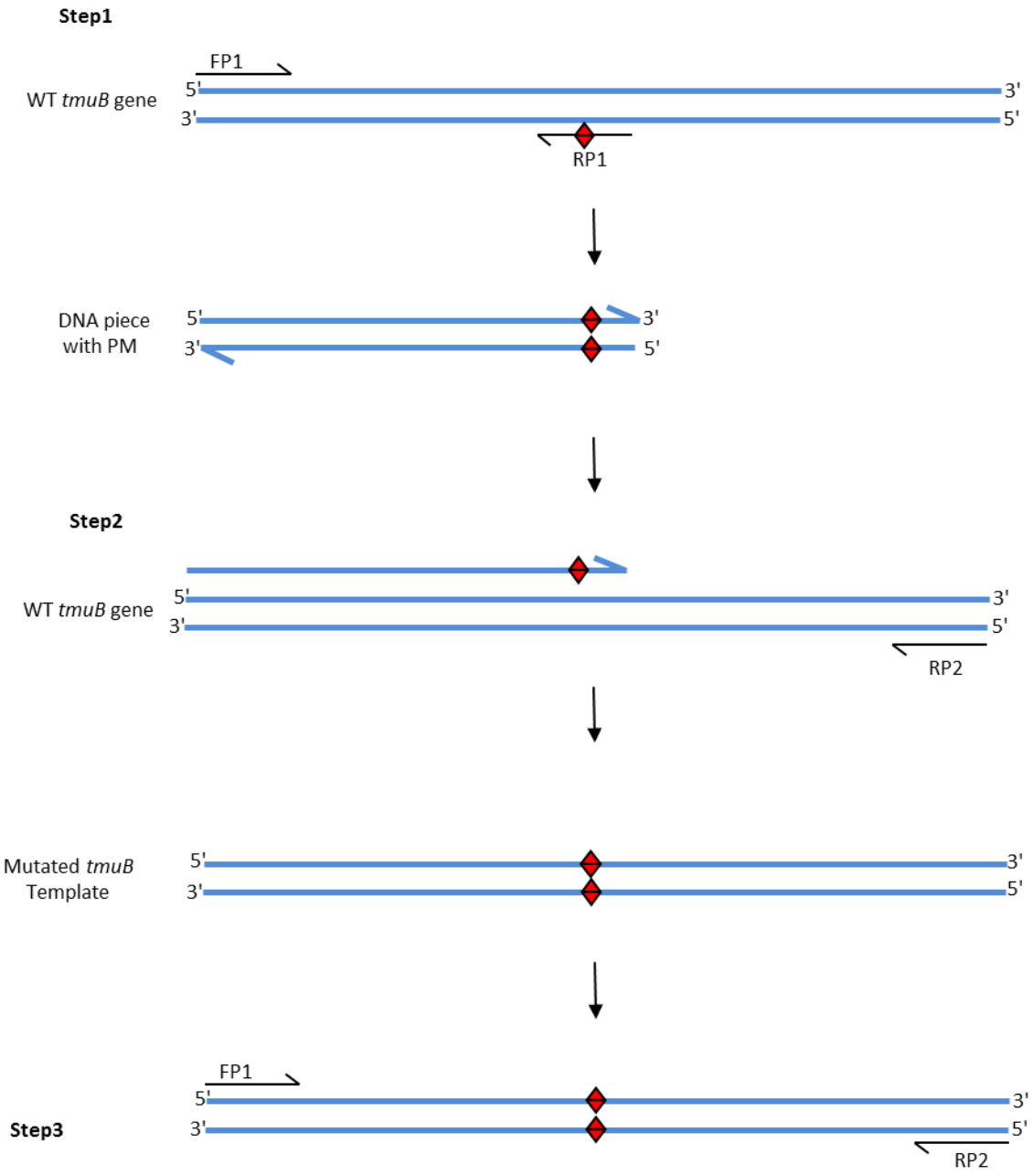


Figure 2.2. Modified overlap extension to create point mutation: step1, amplifying a piece of the DNA template with desired mutation; step2, using this piece of DNA as a primer to amplify the whole gene; and step3, PCR amplifying using external primers.

2.2.16. Protein overexpression using pET28a

The target gene was expressed using pET28a vector which carries an N-terminal His-Tag and thrombin Figure 2.3. Both pET28a vector and the sequenced gene on pGEM-T Easy vector were digested with *NdeI* and *XhoI* restriction enzymes and purified by gel electrophoresis. The ligation between the gene and the vector was carried out as described in section 2.2.4.3 and transformed into competent *E.coli* DH5α and plated out on LB agar plus kanamycin (50µg/ml). The recombinant vector was isolated from grown colonies and was checked by restriction enzyme digestion and gel electrophoresis. The recombinant vector with the correct insert was transformed into competent *E.coli* BL21 (DE3) strain and plated out onto LB agar plus kanamycin (50µg/ml).

2.2.16.1. Small scale expression using IPTG induction

Small scale expression was done to find an optimum temperature and IPTG concentration to produce soluble protein. Universal bottles of L-broth plus kanamycin (50µg/ml) were inoculated with single a colony of *E.coli* BL21 (DE3) strain which carries the recombinant pET28a vector and incubated for 16 hours at 37 ^{o}C with shaking at 200 rpm. 25 ml of fresh LB- broth in 250 ml volume flask seeded with1:100 overnight culture, antibiotic was added and incubated at 37 ^{o}C /200 rpm till the OD_{600} reached 0.4-0.6.

To induce protein expression, different concentrations of IPTG (0.0mM, 0.1mM, 0.5mM and 1mM) were added to different flasks incubated for 2, 4, 6 and 16 hours at different

temperature (37 ^{o}C, 25 ^{o}C, 18 ^{o}C). From each flask, 1ml was taken and centrifuged for 10 minutes at 13000 x g / 4 ^{o}C. The supernatants were discarded and the pellets were washed with STE buffer (10mM Tris-HCl pH 8.0, 0.1 M NaCl and 1 mM EDTA) and analyzed by SDS-PAGE.

2.2.16.2. Large scale expression using autoinduction approach

Autoinduction was used to produce large amounts of soluble protein by preparing half strength L- broth. To prepare 400 ml of half strength L- broth, 4g tryptone, 2g yeast extract, 1.42g Na_2HPO_4, 1.36g KH_2PO_4 and 1.07g NH_4Cl were added, dissolved by stirring and heating and autoclaved. To set up the autoinduction culture, to each 400 ml of half strength L- broth, 3.2ml of 60% glycerol, 20ml of 1% glucose/ 4% lactose/ and 40mM $MgSO_4$, 50µg/ml kanamycin and 1ml of trace metal solution were added, mixed, poured into sterile 2L volume flask, and 4ml of overnight culture was inoculated. This was incubated at 37 ^{o}C for 3.5 hours then moved to 18 ^{o}C for 16 hours. The cells were harvested by centrifugation for 10 minutes at 14000 x g at 4 ^{o}C using 500ml volume bottles in Beckman centrifuge using F10BA rotor. The supernatants were discarded and the pellets were washed with STE buffer and stored at -20 ^{o}C.

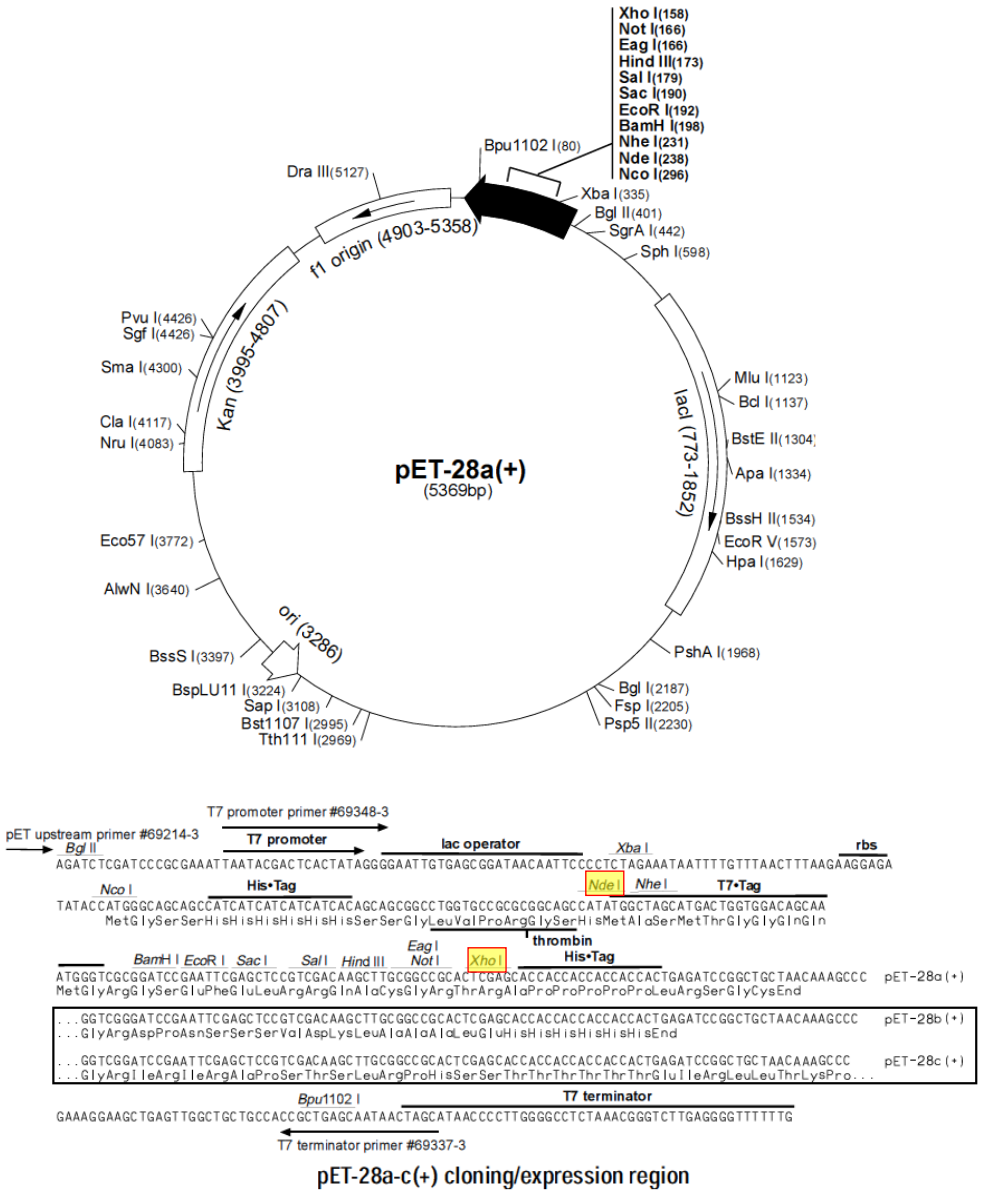


Figure 2.3. General features of pET28a vector and multiple cloning sites (From Novagen user manual, 2002). The *tmuB* gene was cloned between *NdeI* and *XhoI* restriction sites.

2.2.17. Cell lysis

Different methods were used to lyse the cells and separate the soluble and insoluble protein fractions depending on the downstream application.

2.2.17.1. Sonication

Sonication was used to disrupt the cells and check the solubility of the protein in small scale expression. The pellet was resuspended by lysis buffer (20 mM Tris-HCl pH7.5) and for each 1 gram of pellet 10 ml of lysis buffer was added. One tablet of EDTA-free protease inhibitor cocktail provided by Roche was added to each 50 ml of cell suspension. The cell suspension was kept on ice for 1 hour and sonicated for 4 x 10 seconds at 10 micron.

2.2.17.2. Bugbuster® Master Mix

Bugbuster® Master Mix provided by Novagen was used to disrupt the cells and check the solubility of the protein in small scale expression. For each 1 gram of the pellet 5 ml of Bugbuster solution was added and incubated at room temperature for 60 minutes to allow complete lysis of the cells. The soluble and insoluble fractions were separated by centrifugation at 20000 xg for 10 minutes at 4 ^{o}C. The supernatant was carefully transferred to a new tube and treated as soluble protein while the pellet was resuspended with 1% SDS and treated as insoluble protein.

2.2.17.3. French press

The French press is considered to be the first choice to obtain pure and intact protein from large scale expression. The pellet was resuspended in resuspension buffer (20 mM Tris-HCl pH7.5) and for each 1 gram of pellet 10-15 ml of lysis buffer was added. One tablet of EDTA-free protease inhibitor cocktail provided by Roche was added to each 50 ml of cell suspension. The French press cell body and compartments were kept in fridge for 1 hour, the sample loaded (35 ml per run) and the cells were disrupted under 12000 psi pressure. The soluble and insoluble fractions were separated by Beckman centrifuge at 30.000 x g for 20 minutes at 4 ^{o}C using AJ25 rotor. The supernatant was carefully transferred to new tube and treated as soluble protein and ready to be purified.

2.2.18. Protein purification

2.2.18.1. Affinity chromatography

The soluble fraction of protein was purified using HisPur Ni-NTA Resin® provided by Thermo Scientific and following the manufactures instructions. To propylene column (Qiagen), an appropriate amount of the resin was added, left for 15-30 min to settle the nickel particles and the storage buffer was drained. The resin was washed three time with two resin-bed volumes of equilibration buffer (20mM Tris-HCl, 500mM sodium chloride, 10mM imidazole; pH 7.5). The protein sample was loaded on the column gently and the flow-through fraction was collected and reapplied again to maximize protein binding with the beads. The column was washed three times three three resin-bed

volumes with wash buffer (20mM imidazol, 20mM Tris-HCl, 500mM NaCl pH7.5) to elute the unbound proteins. The target protein was then eluted with elution buffer (20mM Tris-HCl, 500mM NaCl, pH7.5) in steps containing different concentrations of imidazole (50mM, 100mM, 150mM, 200mM and 250mM). All fractions from flow through to final elution were collected separately and analyzed by SDS-PAGE to check the purification process.

2.2.18.2. Gel filtration

The protein purified from the affinity chromatography step was further purified by gel filtration using a Superdex 200 26/60 column with 300 ml capacity. ÄKTAprime plus was used to inject the buffer and the sample. The running buffer (20mM Tris-HCl, 500mM NaCl pH7.5) was filtered and degassed using a 0.2µm filter under vacuum. The buffer flow rate was 3 ml/min with pressure 0.5 MPa, after 600 ml of buffer running, the flow rate was reduced to 1ml/min and 2 ml of protein sample (5 mg/ml) was injected. The absorbance scale was 280nm and the fractions (4ml/tube) were collected in 10 ml volume polystyrene tubes. The fractions within the displayed peaks were analyzed by SDS-PAGE.

2.2.19. Sodium dodecyl sulfate- polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were analyzed using SDS-PAGE. 12% resolving gel (lower part) and 5% stacking gel (upper part) were prepared as follows:

12% Resolving gel	10 ml	5% Stacking gel	4 ml
Distilled water	3.3 ml	Distilled water	2.2 ml
30% acrylamide	4.0 ml	30% acrylamide	670 µl
1.5 M Tris (pH 8.8)	2.5 ml	0.5 M Tris (pH 6.8)	1.0 ml
10% SDS	100 µl	10% SDS	40 µl
10% ammonium persulfate	100 µl	10% ammonium persulfate	40 µl
TEMED	4 µl	TEMED	4 µl

A BioRad gel electrophoresis kit was used to set up the gel, the glass plates were fixed on the casting frame and checked for leaking by adding distilled water. The components of the resolving gel were mixed and poured between glass plates and the upper part was filled with isopropanol to level up the gel surface and eliminate air bubbles. After gel polymerization, the isopropanol was removed, washed three times with distilled water and dried with Whatman filter paper. The upper part between glass plates was filled with stacking gel mixture, the desired size comb was inserted and left to polymerize.

The protein samples were prepared by adding an equal amount of loading dye (50 mM Tris-HCl pH 6.8, 2% SDS, 0.1% Bromophenol blue and 10% glycerol), mixed well and

boiled for 5 minutes. The glass plates were assembled using clamping frame and electrode assembly, and placed into the running tank. The running tank was filled with 1X Tris- glycine electrophoresis buffer (5X buffer: 25 mM Tris pH 8.3, 250 mM glycine and 0.1% SDS). The plastic combs were removed and 10 µl of protein sample and prestained protein marker were loaded into the gel wells. The electrodes were connected to power supply and run at 120V / 500mA for 2 hours. The gels were removed from the glass plates and stained with Coomassie Brilliant blue stain (5 tablets of Phastgel Blue R dissolved in one liter of 1:1 distilled water and methanol) for at least 3 hours. The gels were destained by De-staining solution (30% methanol and 10% glacial acetic acid) untill the gels became clear and the protein bands were obvious. Finally, the gels were washed and equilibrated for 5 minutes in distilled water and fixed with cellophane sheets.

2.2.20. Protein dialysis, concentration and storage

The imidazole in the protein sample was eliminated either by gel filtration or dialysis. Dialysis buffer was prepared as the wash buffer used for affinity chromatography (section 2.2.18) but without imidazole or any other buffers were used according to the downstream application. The dialysis tubing Specta/Pro® supplied by SpectrumLabs with 6,000-8,000 MWCO was used and prepared according to the manufacture's manual. The tube was soaked in distilled water for 30 minutes, washed with 0.5% sodium sulphite solution at 80 C^o for 1 minute, washed with distilled water at 60 C^o for 2 minutes

followed by 2 ml of 0.4% of sulphuric acid per 50 ml of total solution for 1 minute and finally washed thoroughly with distilled water. To remove the heavy metal ions, the tube was soaked in 0.5 M of EDTA for 30 minutes and washed several times with distilled water to remove all EDTA and ions. One end of the tube was closed with the plastic clip and the sample was loaded by pipetting and the other end closed by clip. The sample to buffer ratio was 1:200 to maximize buffer exchange and the sample was dialysed with stirring at room temperature for 4 hours and then replaced with fresh buffer and dialysed for overnight at 4 ^{o}C. Amicon® Ultra Centrifugal Filter with 100,000 MWCO was used to concentrate the protein sample. The protein concentration was measured using a Nanodrop 1000 spectrophotometer. 25% (v/v) of glycerol was added to the protein sample, 250µl was aliquoted in 1.5 ml microfuge tube, flushed by liquid nitrogen and was stored at -20 ^{o}C.

2.2.21. *In vivo* protein-protein interactions

Bacterial Adenylate Cyclase Two-Hybrid System Kit was used to test protein-protein interactions *in vivo*. To amplify the *tmuB* gene, forward and reverse primers were designed with *PstI* and *XbaI* restriction sites respectively (Table 2.4). This allowed the gene to be cloned into the Two-Hybrid System vectors.

pKT25, pKNT25, pUT18 and pUT18C vectors (Table 2.2) were used to create a fusion protein of the T25 and T18 domains of adenylate cyclase with the target protein. Control vectors provided with the kit were pKT25-zip and pUT18-zip in which the leucine zipper

is fused to T25 and T18 domains respectively. The *tmuB* gene was sub-cloned in frame into the pKT25, pKNT25, pUT18 and pUT18C vectors using standard techniques. The recombinant plasmids were transformed into *E.coli* DH5α using the standard protocol and plated out on LB agar with the appropriate antibiotics (Figure 2.4). The recombinant vectors were checked for correct insertion by restriction enzyme digestion and gel electrophoresis. The two compatible recombinant vectors were co-transformed to the reporter strains, *E.coli* BTH101 and DHM1 (Table 2.1), which are adenylate cyclase deficient (*cya*) and were plated out on both MacConkey agar containing 1% maltose, kanamycin and ampicillin, and LB containing IPTG, Xgal, kanamycin and ampicillin (Table 2.3). The plates were incubated at 30 ^{o}C for 2-3 days.

The suspected colonies were further characterized by measuring the β-galactosidase activity using o-nitrophenol-β-galactoside (ONPG) as a substrate. The protocol procedure was carried out according to the manufactures instructions (Euromedex BACTH system kit) and described in detail by (Miller, 1992 and Karimova, 2005). Single pure colony was inoculated in 5ml LB broth supplemented with 0.5mM IPTG and the appropriate antibiotics, and incubated overnight at 30 ^{o}C. This was diluted 1:5 with M63 medium and the optical density (OD) was measured at 600nm. To permeabilized bacterial cells, to each 5ml of the diluted bacterial culture, 60µl of each toluene and 0.1%SDS were added in a glass tube, plugged slightly with cotton, vortexed for 10 seconds and incubated at 37 ^{o}C with shaking for 40 minutes. Aliquots of 0.1 to 0.5ml of the permeabilized cells were added to 1ml of PM2 assay buffer (70mM Na_2HPO_4**.**$12H_2O$, 30mM NaH_2PO_4. H_2O, 1mM $MgSO_4$, 0.2mM $MnSO_4$, pH 7.0 and

100mM β-mercaptoethanol was added before use). The enzymatic reaction was set up by adding 0.1 ml of the permeabilized cells to 0.9 ml of PM2 buffer (in a 5 ml glass tube) and incubated in a water bath at 28 °C for 5 min. This was followed by adding 0.25ml of the substrate solution (4% ONPG in PM2 medium). A yellowish color developed as a result of the enzymatic activity and the reaction was stopped by adding 0.5 ml of the stop solution (1M Na_2CO_3). The OD for each tube was measured at 420nm. The β-galactosidase activity is measured in units/mg of the bacterial dry weight (one unit equivalents to 1 nmol of ONPG hydrolyzed per min at 28°C).

The following equation was used to calculate the enzyme activity, E (in Units/ml):

E = 200 x OD_{420} / min of incubation) x Df.

200: Is the inverse of the absorption coefficient of o-nitrophenol.

Df: The dilution factor for the permeabilized cells which was 10.

The result was converted to Units/mg, by considering that 1ml of culture at OD600 =

1 equivalent to 300 µg dry weight bacteria.

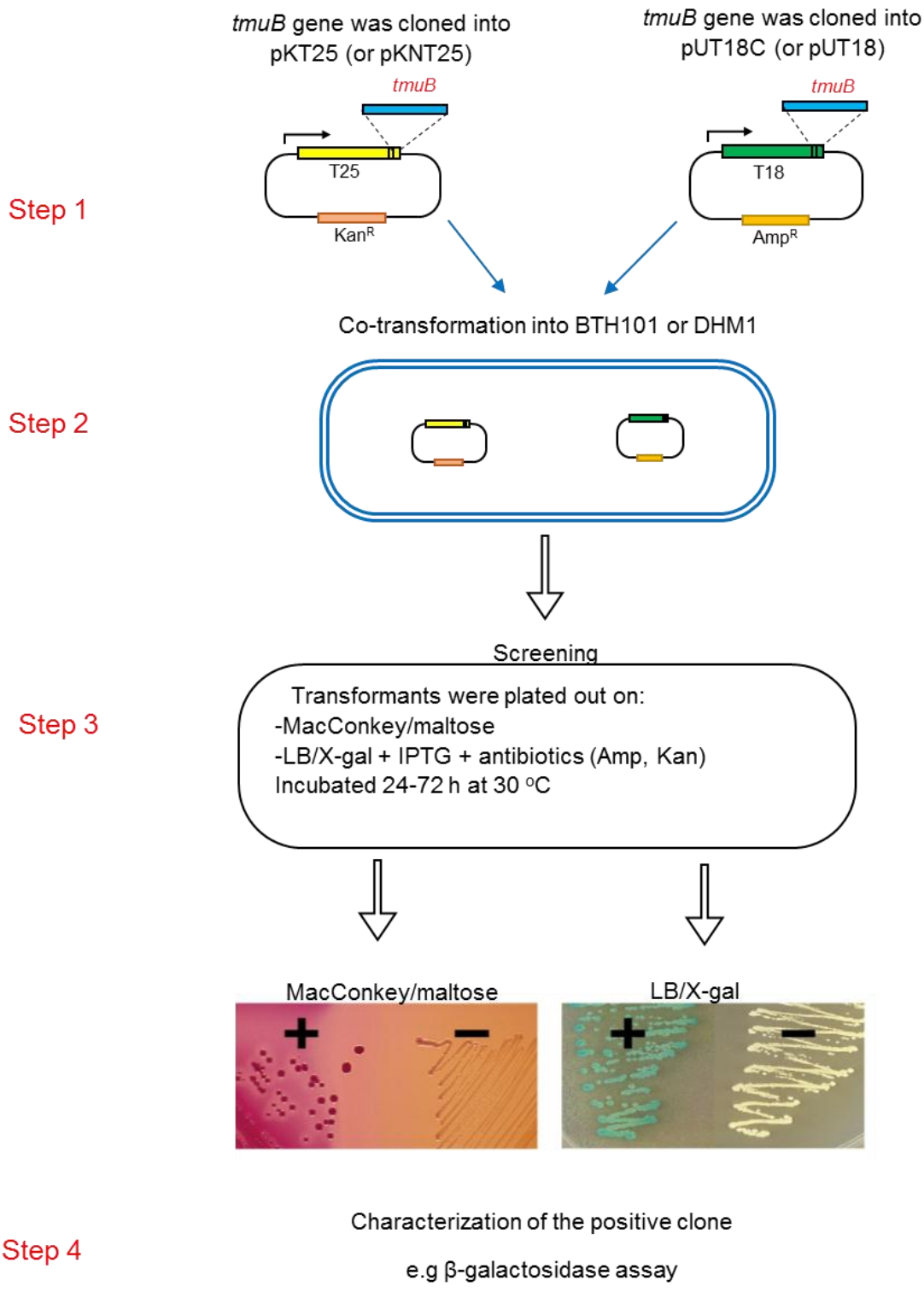


Figure 2.4. General steps for identification *in vivo* protein-protein interactions using the bacterial two hybrid BACTH system.

2.2.22. *In vitro* enzyme activity assay

The initial hydroxylase TmuB activity was measured at 23 ^{o}C in 500 µl reaction volume which consists of 35µM TmuB protein, 0.25mM $FeSO_4$ (Co-factor), 0.5mM α-keto glutaric acid (co-substrate) and 50µM pseudomonic acid A in 20mM Tris-HCl buffer, pH7.5 and 500mM NaCl. The reaction was set up in different time course and the reaction was stopped by adding an equal volume of 100% acetonitrile. The mixture was left for 5 minutes at room temperature, centrifuged at 13000 x g for 5 minutes, the supernatant was separated and analyzed by HPLC. Various conditions were used to reach the optimal reaction and to find the initial velocity of the enzyme activity. These variations included different temperature, time, and different concentrations of TmuB protein, the substrate, co-factor and co-substrate. Each assay point was triplicates to determine the mean standard error. The data were fitted to a Michaelis-Menton model using Prism graphpad software (http://www.graphpad.com/ scientific-software/prism).

2.2.23. Suicide mutagenesis in *Pseudoalteromonas sp.* SANK 73390

To inactivate TmuB activity in *Pseudoalteromonas sp*. SANK 73390, isoleucine 109 was changed to Asparagine via suicide mutagenesis. The *tmuB* gene with point mutation I109N was cloned into suicide vector pAKE604 as a *SalI/XbaI* insert. This was moved to the WT *Pseudoalteromonas sp*. SANK 73390 via conjugation using *E.coli* S17-1. 1ml of overnight culture of the donor and recipient were mixed, passed through 0.45µm filter, placed on marine agar and incubated overnight at 23 ^{o}C. The filter was resuspended in

1 ml of marine broth, followed by serial dilutions (10^{-1} to 10^{-5}), 200µl of each dilution being spread on marine agar supplemented with 100µg ml^{-1} kanamycin and incubated at 23 ^{o}C for 3-4 days. Pure colonies (cointegrant clone) were selected and incubated at 23 ^{o}C in marine broth without antibiotic to isolate strains in which the suicide plasmid had excised from the chromosomal DNA. From this, serial dilutions were plated on marine agar supplemented with 5% sucrose and the grown colonies were tested for kanamycin sensitivity. The colonies that were Kan^{s} and Suc^{r} were selected and checked by PCR and DNA sequencing. Figure 2.5 illustrates the strategy of suicide mutagenesis.

2.2.24. Antibiotic extraction and purification

2.2.24.1. Thiomarinol from *Pseudoalteromonas sp* SANK 73309: 100 ml of marine broth was inoculated with 5 ml of overnight culture and incubated at 23 ^{o}C for 48 hours. To stabilize the thiomarinol compound, the pH was adjusted to 6.0 by HCl. An equal volume (100 ml) of acetone was added to the culture, mixed and incubated at 4 ^{o}C for 2 hours to decompose the cell wall.

The acetone solvent was evaporated using a vacuum aspirator in the fume hood. An equal volume (100 ml) of ethyl acetate was added, mixed by shaking for 5 -10 minutes and transferred to separating funnel. The mixture was separated into two phases: solvent phase containing thiomarinol in the upper part and the water phase in the lower part. The latter was discarded by opening the funnel outlet slightly and the solvent phase transferred to the rotary evaporator. The dried extracted powder on the flask wall was recovered by an appropriate amount of methanol or DMSO.

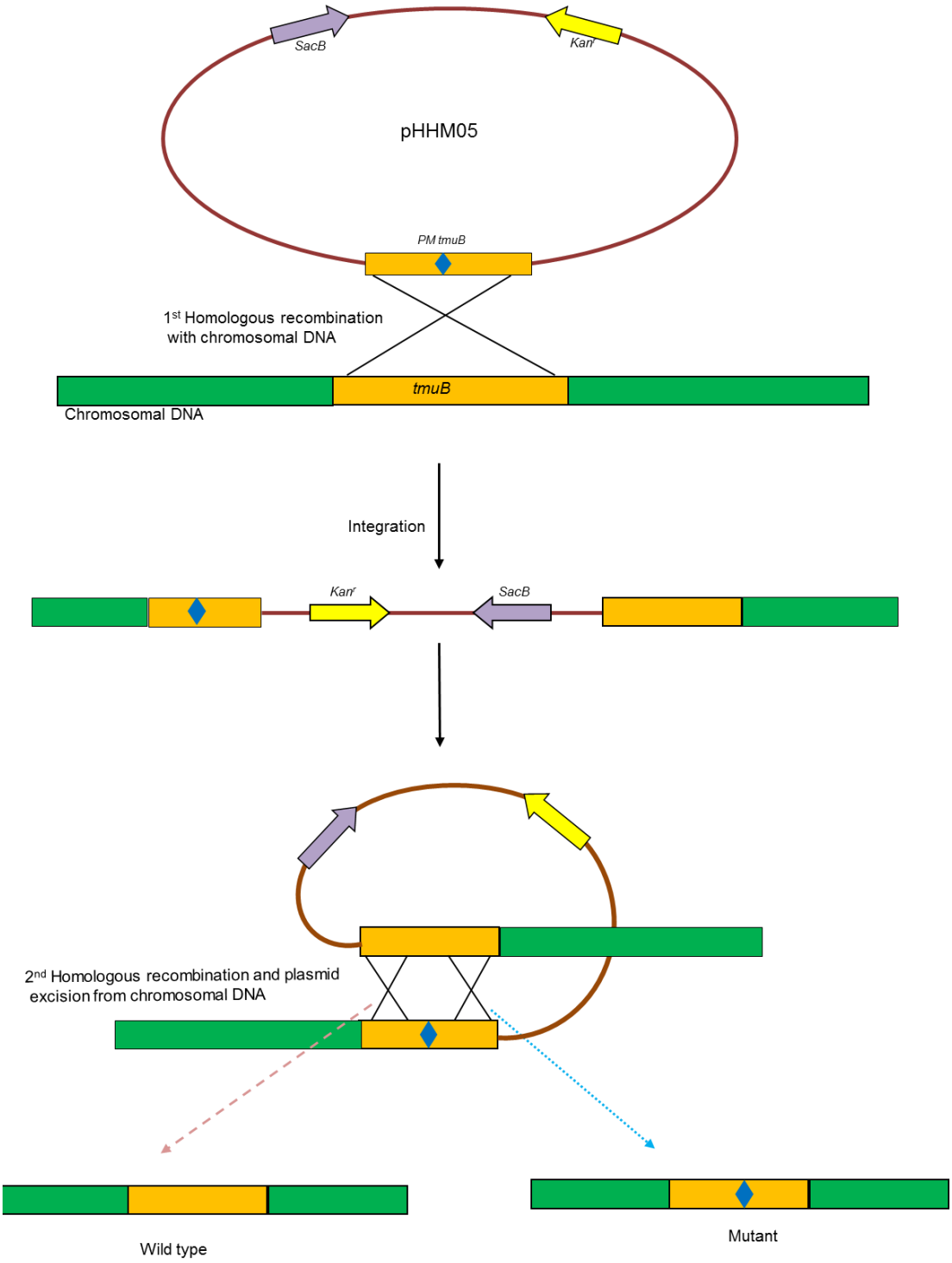


Figure 2.5. The suicide mutagenesis technique using pAKE604 plasmid. This technique used to create point mutation changing isoleucine 109 to asparagine.

The product was analyzed by HPLC and the target peak was isolated using a fraction collector.

2.2.24.2. Mupirocin from *Pseudomonas fluorescens* NCIMB 10586: 100 ml of secondary stage media (SSM) was inoculated with 5 ml of overnight culture and incubated at 22 C^{o} for 40-60 hours with shaking at 200 rpm. The culture was transferred to 50ml Falcon tubes, centrifuged at 8,000 x g, the supernatant separated and the pH was adjusted to 4.5 by HCl. An equal volume (100 ml) of ethyl acetate was added, mixed by shaking for 5 -10 minutes and transferred to a separating funnel. The rest of steps were done as described in the previous section.

2.2.25. Bioassay test

2.2.25.1. Plate bioassay for mupirocin activity was carried out by spotting 10µl of *Pseudomonas flourescens* overnight culture onto 20 ml L-agar plates and incubating overnight at 30 ^{o}C. This was overlaid with 15 ml of molten agar at 45 ^{o}C seeded with overnight culture of *Bacillus subtilis* 1064 (40µl/ml) and 0.5 ml of triphenyl tetrazolium chloride TTC (5%). After incubation for 18h at 37 ^{o}C, the clear zones of inhibition around the bacterial spots were measured.

2.2.25.2. Plate bioassays for thiomarinol using a disc diffusion assay to test the potency of thiomarinol compounds. The plate was prepared by pouring of 15 ml of LB agar into the Petri dish and leaving it to solidify before overlaying with 10 ml of the detection layer

which was also left to solidify. The detection layer was prepared by mixing 4.75ml of melted LB agar 45 ^{o}C, 4.75ml pre warmed 45 ^{o}C LB broth, 100µl of TTC 5% and 400µl of the overnight culture of *Bucillus subtilis* 1064 in a Falcon tube. The discs paper with thiomarinol compounds were prepared by adding 40µl of the thiomarinol stock slowly, leaving it to dry and placing it over the overlaid layer. After incubation for 18h at 37 ^{o}C, the clear zones of inhibition around the paper disc were measured.

2.2.26. Minimal inhibitory concentration

Three flasks of L- broth were inoculated (1:100) with overnight cultures of detector bacteria (*Bacillus subtilis or E.coli)* and incubated at 37 ^{o}C with shaking untill the OD_{600} reached to 0.4 - 0.6. This was diluted with L-broth to obtain OD_{600} 0.08-0.10 (0.5 McFarland standard) (McFarland, 1907), which contains approximately 10^{8} CFU/ ml. This was diluted 1:100 to the working dilution (10^{6} CFU/ml). A Costar® 96 well flat bottom plate was used to set up the tests. To each well, 100 µl of bacterial suspension and 100 µl of antibiotic dilutions were added. The plates were incubated at 37 ^{o}C with shaking at 200 rpm for 18-20hrs. A SPECTROstar Nano microplate reader was used to read the OD at 600 nm as raw data, and three combination wavelengths (600, 900,977)nm were used to standardize the water content in the well. These data were analyzed using MARS software.

2.3. Results

2.3.1. Blast search to identify putative hydroxylase

One prominent difference between mupirocin and thiomarinol structure is the extra hydroxyl group at C4 in monic acid (Figure 2.6). As mentioned in the Introduction to this Chapter, the alignment between the mupirocin and thiomarinol gene clusters, which were described previously (Fukuda *et al*., 2011) identified some genes in thiomarinol gene cluster that are absent from the *mup* cluster and might be responsible to catalyse 4-hydroxylation of thiomarinol. To identify the potential candidate, the predicted amino acid sequences of these genes were used to perform BLAST searches against the NCBI GenBank protein database (Altschul *et al*., 1997). This identified TmuB as a putative phytanoyl-CoA dioxygenase with the identity score ranging from 42% to 33% and E-value 1e-60 to 5e-45 respectively (Figure 2.7A) and TmlZ as putative monooxygenase with the identity score 46% and E-value 5e-34 (Figure 2.7B). Based on this, *tmuB* and *tmlZ* have been chosen as candidates that may encode the hydroxylase.

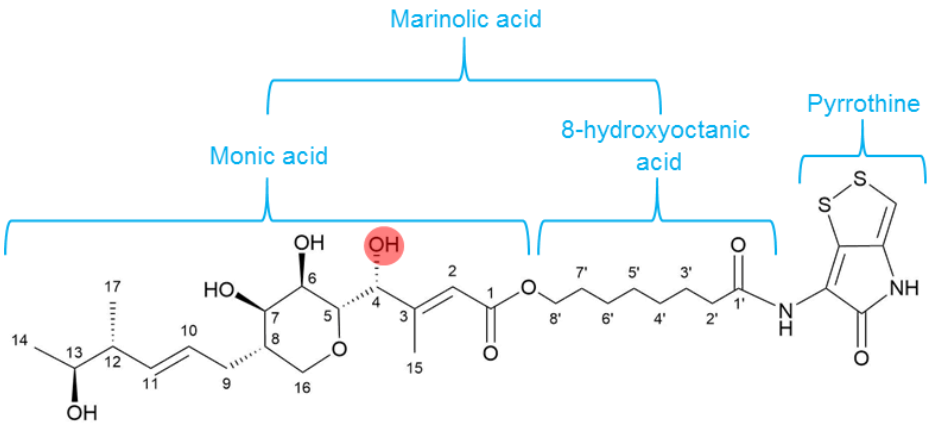


Figure 2.6. Thiomarinol structure showing the 4-hydroxylation of monic acid.

A

Description	Max score	Total score	Query cover	E value	Ident
phytanoyl-CoA dioxygenase [Roseovarius sp. 217] >phytanoyl-CoA dioxygenase [Roseovarius sp	202	202	90%	1e-66	42%
phytanoyl-CoA dioxygenase [Roseovarius sp. TM1035] >	200	200	91%	7e-66	41%
phytanoyl-CoA dioxygenase [Roseovarius mucosus] >gi|699744541|gb|KGM86687.1| Protein invc	199	199	92%	3e-65	40%
phytanoyl-CoA dioxygenase [Aestuariivita boseongensis]	198	198	91%	6e-65	41%
phytanoyl-CoA dioxygenase [Luminiphilus syltensis] >phytanoyl-CoA dioxygenase [Luminiphilus_	184	184	97%	2e-59	39%
phytanoyl-CoA dioxygenase [Parvibaculum lavamentivorans] >Phytanoyl-CoA dioxygenase [Parv	166	166	95%	5e-52	37%
phytanoyl-CoA dioxygenase [Tepidicaulis marinus] >phytanoyl-CoA dioxygenase [Tepidicaulis ma	164	164	95%	1e-51	38%
phytanoyl-CoA dioxygenase [Caenispirillum salinarum] >Phytanoyl-CoA dioxygenase [Caenispiril	162	162	93%	1e-50	36%
phytanoyl-CoA dioxygenase [alpha proteobacterium AAP81b] >phytanoyl-CoA dioxygenase [alph	159	159	95%	1e-49	35%
phytanoyl-CoA dioxygenase [uncultured proteobacterium]	157	157	89%	8e-49	38%
phytanoyl-CoA dioxygenase [gamma proteobacterium HIMB55] >gi|374303988|gb|EHQ58172.1| p	156	156	95%	1e-48	38%
phytanoyl-CoA dioxygenase [Candidatus Phaeomarinobacter ectocarpi] >gi|610513160|emb|CDC	150	150	98%	3e-46	35%
phytanoyl-CoA dioxygenase [marine gamma proteobacterium HTCC2080] >gi|119459397|gb|EAV	147	164	95%	3e-45	35%
phytanoyl-CoA dioxygenase [Acinetobacter venetianus] >phytanoyl-CoA dioxygenase [Acinetoba	147	147	95%	5e-45	33%

B

Description	Max score	Total score	Query cover	E value	Ident
antibiotic biosynthesis monooxygenase [Methylophilus methylotrophus]	123	123	96%	5e-34	46%
hypothetical protein A2157_13515 [Deltaproteobacteria bacterium RBG_16_47_11]	123	123	94%	5e-34	45%
MULTISPECIES: hypothetical protein [Methylophilus]	123	123	91%	6e-34	47%
hypothetical protein [Methylophilus sp. OH31]	122	122	94%	1e-33	47%
hypothetical protein [Methylophilus sp. Q8]	122	122	94%	1e-33	47%
hypothetical protein [Microcystis aeruginosa]	122	122	97%	1e-33	47%
hypothetical protein [Microcystis aeruginosa]	121	121	97%	2e-33	47%
hypothetical protein [Thiothrix disciformis]	121	121	95%	2e-33	45%
hypothetical protein [Microcystis aeruginosa]	121	121	97%	3e-33	47%
hypothetical protein [Methylophilus sp. 5]	120	120	88%	5e-33	48%
hypothetical protein [Microcystis aeruginosa]	120	120	97%	6e-33	46%
hypothetical protein [Candidatus Marithrix sp. Canyon 246]	120	120	94%	6e-33	44%
hypothetical protein [Leptolyngbya sp. PCC 7376]	120	120	93%	6e-33	45%
hypothetical protein imdm_45 [gamma proteobacterium IMCC2047]	120	120	94%	7e-33	45%

Figure 2.7. The result of Blast (non-redundant protein sequence) search for the candidate genes. A: TmuB and B: TmlZ.

2.3.2. Gene amplification and recombinant vector construction

The *tmuB* gene (795bp) was amplified by standard PCR and the product was run on the agarose gel which showed about 800bp. The purified DNA was cloned to pGEMT-Easy vector and was digested by relevant restriction enzymes to check that it is the right recombinant. The identical sequence insert was cloned to the expression vectors, pJH10 and pET28a (Figure 2.8). The same procedures were carried out to clone *tmlZ* into pJH10 plasmid.

2.3.3. HPLC analysis of *in trans* expression of TmuB and TmlZ in heterologous hosts

The constructed vectors (pHHM04 and pHHM05) were transferred to wild type and mutants of *P. fluorescens* NCIMB 10586 via conjugation and the culture supernatants were analyzed by HPLC to investigate the effect of genes expression. HPLC analysis showed that *tmuB* expression in WT *P. fluorescens* NCIMB10586 shifted the retention time (RT) of the pseudomonic acid A peak from 20.2min (PA-A) to give a new peak at 18.2 min (Figure 2.9) while *tmlZ* expression did not show any activity. LCMS followed by NMR confirmed that this new product was the result of 4-hydroxylation.

TmuB was expressed in the mutant strains that produce other derivatives rather than PA-A. *P. fluorescens* mutants *ΔmacpE, ΔmupO, ΔmupU and ΔmupV* were created previously (Cooper *et al*., 2005b) and produce pseudomonic acid B. TmuB production in

these mutants did not show any change in metabolite profile of PA-B (RT 19.3min) (Figure 2.9). TmuB was expressed in other mutants of *P. fluorescens* which produce different pseudomonic acid derivatives: *ΔmmpE/OR, ΔmupF* and *ΔmupC* which produce PA-C, mupirocin F and mupirocin C respectively and *ΔmupT* and *ΔmupW* which produce mupirocin W. The HPLC profile for the mutant strains revealed that TmuB expression yielded some new major and minor peaks.

As illustrated in Figure 2.9, expressing *tmuB* in the *ΔmmpE/OR* mutant gave a novel peak eluting at 20.2 min. Likewise, *tmuB* expression in the *ΔmupF* mutant yielded a new major peak at 19.6 min while the mupirocin F peak at 22.3 min increased by two-fold (Figure 2.9). However, there were no significant differences in the metabolite profile after TmuB expression in the strains producing mupirocin C (22.4 min) or mupirocin W (22.0 min). The structures of the new derivatives were determined using LC-ESI-MS and NMR analysis. LC-ESI-MS identified the molecular weights of 516Da, 500Da and 514Da from WT, *ΔmmpE/OR* and *ΔmupF* respectively. The structures of these novel products were elucidated using NMR and were identified as 4-hydroxy PA-A, 4-hydroxy PA-C and 4-hydroxy mupirocin F respectively (Figure 2.10, Table 2.5 and Appendix A, B and C).

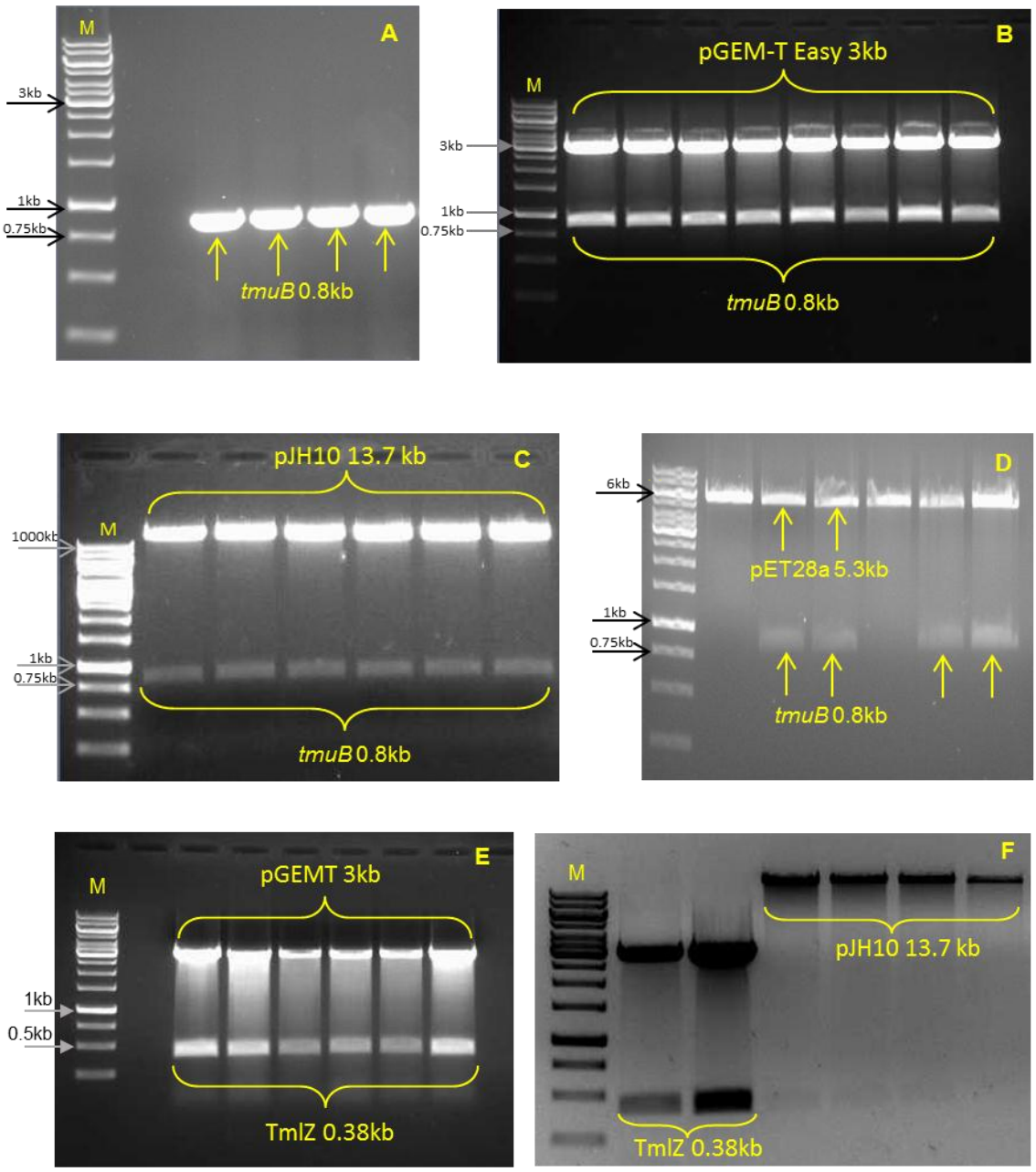


Figure 2.8. The *tmuB* and *tmlZ* genes amplification and cloning. PCR product of *tmuB* (A), The *tmuB* gene (0.8kb) cloned to pGEM-T Easy as A-tailed insert (B), pJH10 vector as *EcoRI*/*XbaI* insert (C) and pET28a vector as *NdeI*/*XhoI* insert (D), The *tmlZ* gene (0.38kb) cloned to pGEM-T Easy as A-tailed insert (E) and into pJH10 vector as *EcoRI*/*XbaI* insert (F). Constructed plasmids were digested with relevant restriction enzymes to release the target inserts, M: 1kb DNA marker.

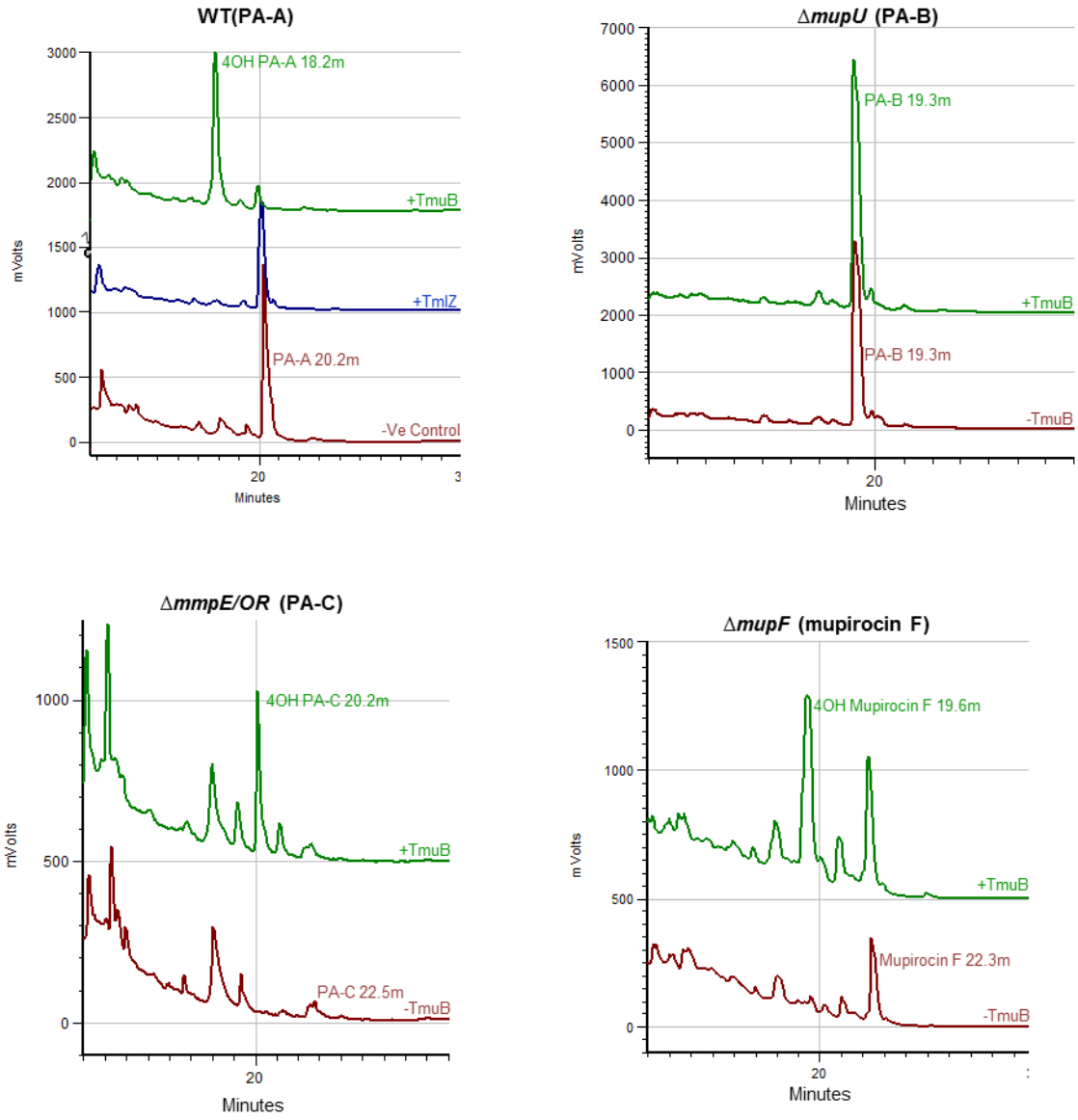


Figure 2.9. HPLC analysis of culture supernatant from *Pseudomonas fluorescens* strains. TmuB modifies pseudomonic acid A and C but not B, that it has a hydroxyl at C8. Mupirocin F with a keto- rather than hydroxyl at C7 is also 4-hydroxylated. TmlZ (blue trace) could not modify PA-A.

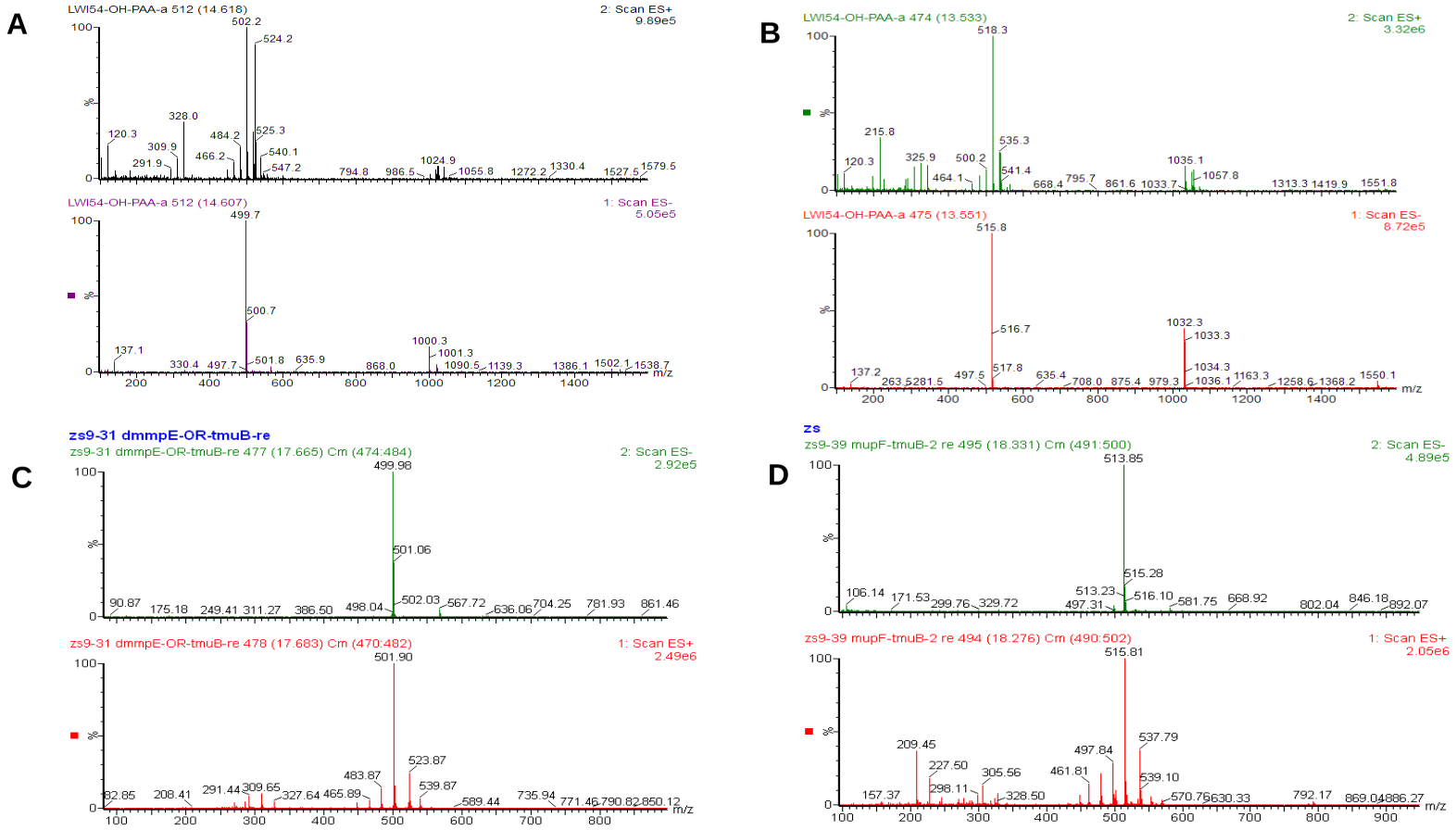


Figure 2.10. Mass spectrometry analysis of hydroxylated version of pseudomonic acid derivatives. A & B: PA-A (MW 500) and 4OH PA-A (MW 516) respectively. C. 4OH PA-C (MW 500). and D: 4OH mupirocin F (MW 514).

Table 2.5. Characterization of the new metabolites produced by *P. fluorescens* strains in presence of TmuB expression.

***P. fluorescens* strains transformed with pJH10**	**-TmuB HPLC RT/min**	**+TmuB HPLC RT/min**	**-TmuB MS**	**+TmuB MS**	**NMR**
WT	20.2	18.2	500	516	4-hydroxy PA-A
ΔmacpE	19.35	19.35	516	516	PA-B and trace of PA-A changed to 4-hydroxy PA-A
ΔmupO	19.35	19.35	516	516	PA-B and trace of PA-A changed to 4-hydroxy PA-A
ΔmupU	19.35	19.35	516	516	PA-B and trace of PA-A changed to 4-hydroxy PA-A
ΔmupV	19.35	19.35	516	516	PA-B and trace of PA-A changed to 4-hydroxy PA-A
ΔmmpE/OR	22.5	20.22	484	500	4-hydroxy PA-C
ΔmupW	22	22	503	503	Mupirocin W
ΔmupC	22.4	22.4	496	496	Mupirocin C
ΔmupF	22.35	19.6	498	514	4-hydroxy mupirocin F

2.3.4. Bioinformatics analysis

As described in the Materials and Methods section 2.2.14, the HHpred online program was used to identify the crystal structures of homologous proteins to be used as templates to create a homology model of TmuB. It has been found that all the templates belong to the nonheme-iron(II)/α-ketoglutarate(αKG)-dependent superfamily. Some enzymes in this group use α-ketoglutarate as a co-substrate to incorporate oxygen atom into the substrate (Hausinger, 2004) as shown in Figure 2.11.

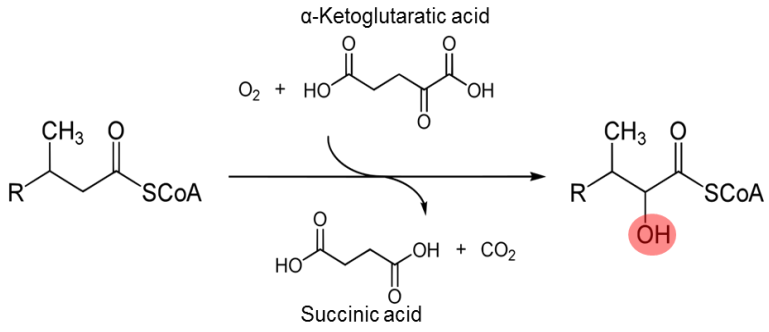


Figure 2.11. The hydroxylation reaction catalyzed by some enzymes belonging to the nonheme-iron(II)/α-ketoglutarate(αKG)-dependent superfamily.

2.3.4.1. Modelling

TmuB secondary structure of TmuB was predicted using the secondary structure prediction server PSIPRED and Jpred3. This was aligned with the sequence and secondary structure of the templates in Seaview and refined manually (Figure 2.12).

Modeller version 9.12 was used to generate a homology model of TmuB on the basis of the crystal structure of EasH (pdb 4NAO). Up to 10 models were generated, and were evaluated using a set of programs to choose the best model (see below).

As illustrated in Figure 2.13, the core of the homology model of the TmuB protein contains a double stranded beta-helix (DSBH) or jelly-roll fold which is the characteristic feature that unifies the structures of the nonheme-iron(II)/2-oxoglutarate-dependent dioxygenase superfamily. This jelly-roll strucuture is surrounded by extra beta sheets and alpha helixes. The β-strand II and VII are disordered while the other strands are well determined. The cofactor (Fe^{+2} ion) is located in the center of the active site in the homology model and bound by three residues, His121 and Asp123 are located at the end of β-strand II and His191 is located on β-strand VII. The co-substrate αKG binds to Fe^{+2} ion in a bidentate manner via C1- carboxylate and C2-keto group while the C5-carboxylate forms hydrogen bonds with Gln118 and Thr154 and electrostatic interaction with Arg202 (Figure 2.14). These residues are highly conserved among the enzyme family. The surface mode of the protein shows that the active site is identified as a long groove leading to the pocket where the cofactor and the co-substrate are bound. The docking process shows that this groove and pocket in the active site are consistent with the size and conformation of the substrate (Figure 2.19).

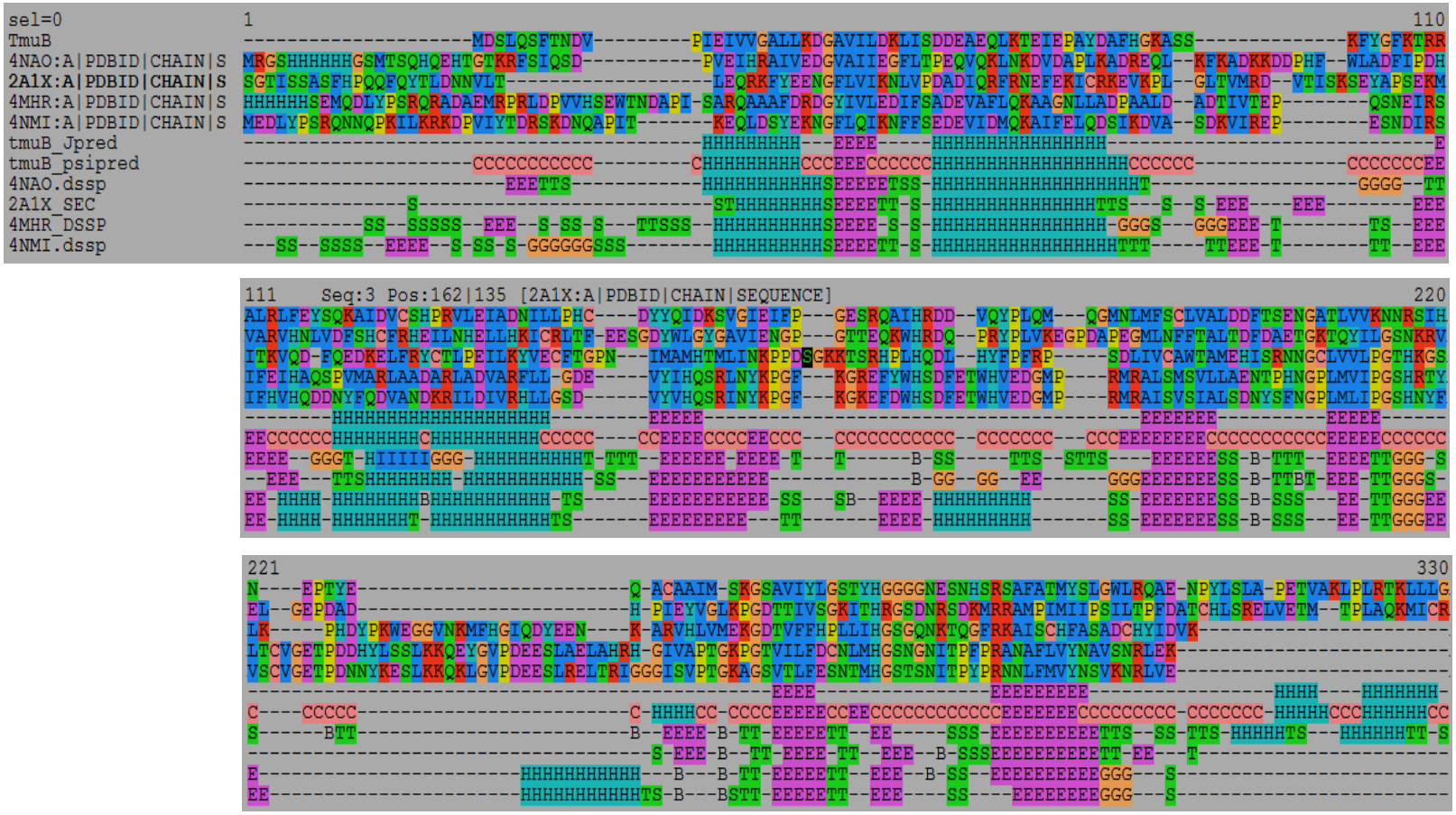


Figure 2.12. Secondary structure alignment of TmuB with the previously resolved protein templates 4NAO, 2A1X, 4MHR and 4NMI. DSSP software was used to generate the secondary structures and Seaview software was used to visualise and to refine the alignments. The TmuB model obtained was based on the 4NAO structure.

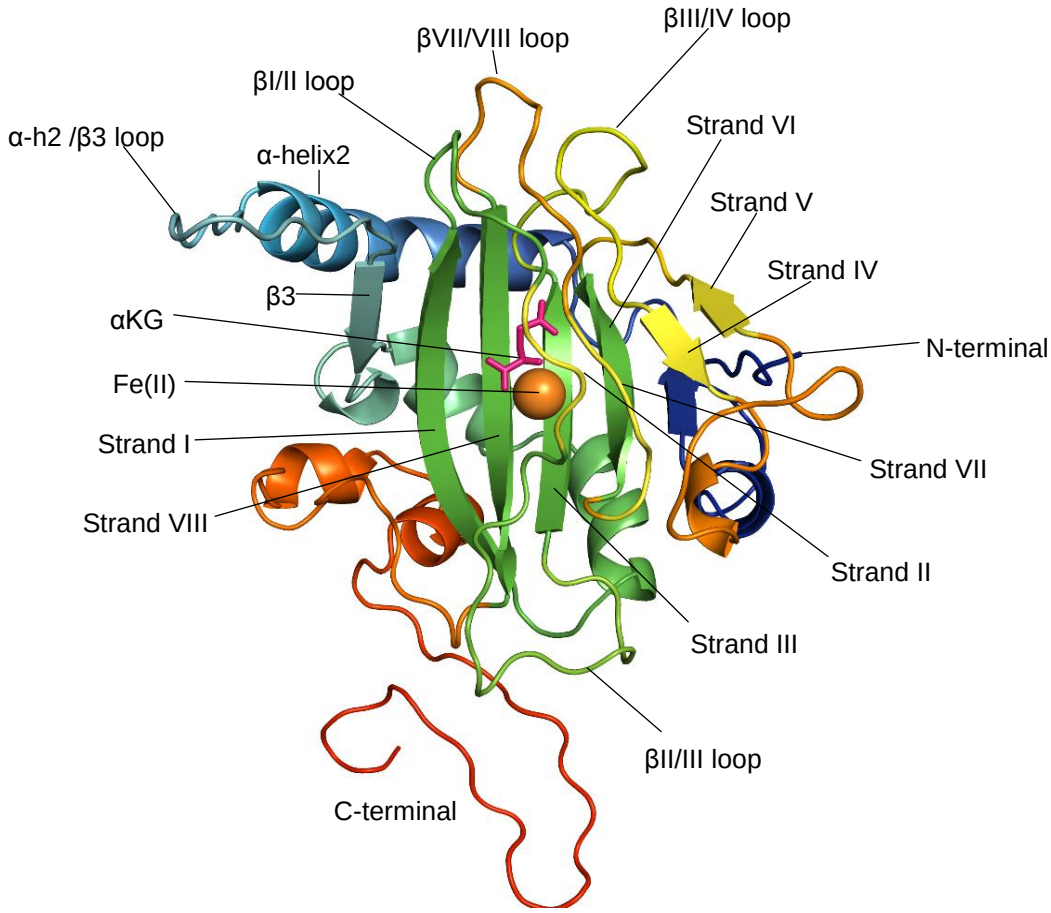


Figure 2.13. Homology model of TmuB (8^{th}) illustrating the jelly-roll fold in the active site. The parts of the structure named based on previous crystal structures. Fe^{+2} (orange sphere) and α-keto-glutarate (pink stick) are located in the centre of the jelly-roll structure. The jelly-roll consists of eight β-strands arranged into four at each side. The strands I, VIII, III, and VI are in green color at one side while the strands II, VII, IV, and V in yellow color on the opposite side. Note that both strand II and VII are disordered. N-terminal in blue and C-terminal in red color. See Figure 2.21 for amino acid sequence and secondary structure location.

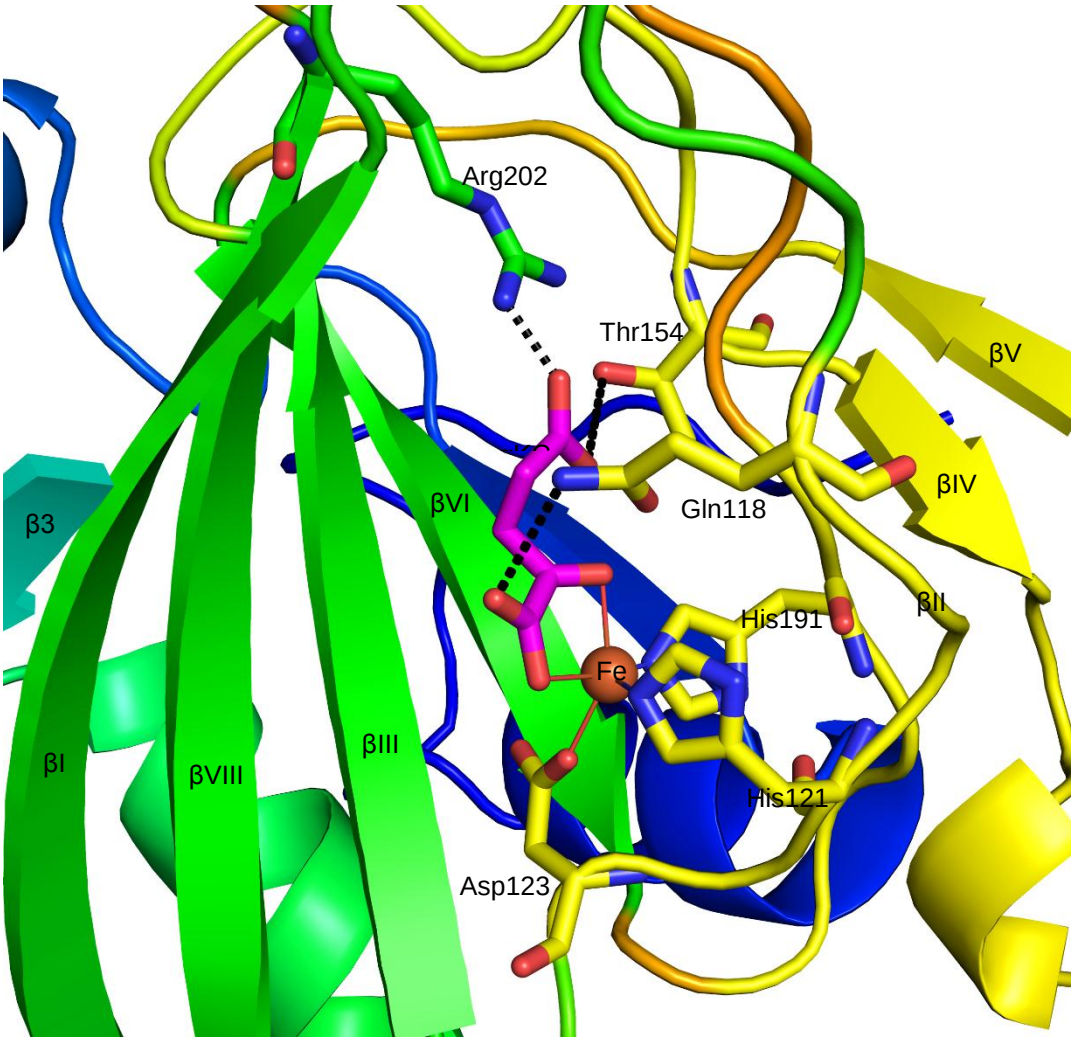


Figure 2.14. Close insight into the active site of TmuB model. Fe^{+2} (orange sphere) is located in the centre of the jelly-roll structure and octahedrally coordinated by three conserved residues His121, Asp123 (at the end of strand II) and His19 (on strand VII), and the co-substrate α-keto-glutarate (pink stick) which is held by Gln118 (at beginning of strand II), Thr154 (on βIII/IV loop) and Arg202 (on beginning of strand VIII). The black dash lines represent hydrogen bonds and electrostatic interaction.

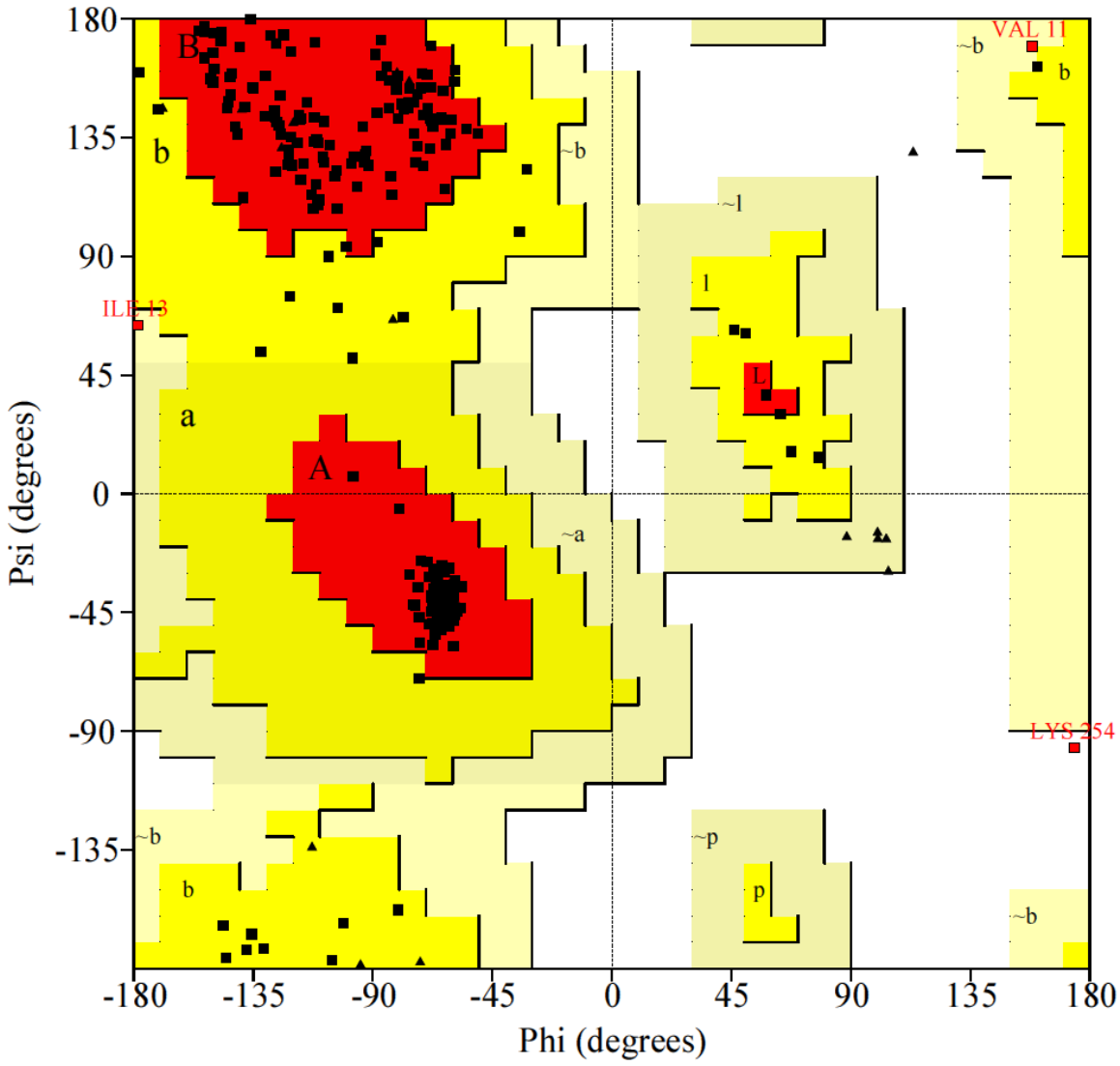
2.3.4.2. Homology model structure assessment

The overall geometric and stereo-chemical qualities of the TmuB modelled structures were evaluated. WHAT_CHECK shows the residues packing score and the models quality, and the overall quality score of the models obtained by PROSESS (Table 2.6). All the models were checked and those with the highest score were selected.

The Ramachandran plot for 8th model shows that all the residues around the active site are located within the favoured region. Only three residues (Val11, Ile13 and Lys254) are located within outlier regions which are far from the active site and located on the extreme N- and C-termini (Figure 2.15) which may not affect the catalytic activity. However, all models showed that the α-helix2/β3 loop has not been folded properly because it was missing from the crystal structure of the templates. Recently, two more proteins have been determined and were identified by HHpred with higher identity to TmuB, 21% for AsqJ (Bräuer *et al.,* 2016) (PDB 5daw) and 23% for FtmOx1 (Yan *et al*., 2015) (PDB 4y5t). It can be noticed that α-helix2/β3 loops in these proteins have been resolved and appear to be properly folded (Figure 2.16). Therefore, TmuB has been remodelled using these two proteins as the primary templates. The secondary structure of the TmuB was aligned with the templates and manually refined using Seaview (Figure 2.17) and the TmuB structure was predicted according to AsqJ. Up to 6 models of TmuB were generated and were evaluated as described previously, and the quality scores were obtained (Table 2.6). As illustrated in Figure 2.18, in this model (15th) the α-helix2/β3 loop has been folded and in combination with the βII/III loop covers the entrance to the active site.

Table 2.6. WHAT-CHECK and PROSESS evaluation for the homology structures of TmuB comparing with the templates, 5DAW and 4NAO. The Z score out of (0 to -3) range is unacceptable for all criteria except the last two rows.

Criteria	Mod 8	Mod 10	Mod 12	Mod 15	5DAW	4NAO
Resolution read from PDB file	-1.000Å	-1.000Å	-1.000Å	-1.000Å	1.600Å	1.650Å
1st generation packing quality	-3.032	-2.932	-2.491	-2.764	-0.709	-0.921
2nd generation packing quality	-3.195	-3.035	-1.988	-2.939	-1.519	-0.589
Ramachandran plot appearance	-0.550	-0.775	-4.531	-0.556	-0.513	0.024
chi-1/chi-2 rotamer normality	-1.099	-1.240	-4.063	-3.361	-1.477	0.206
Backbone conformation	-1.320	-2.022	-3.253	2.944	0.260	0.742
Inside/Outside distribution	1.159	1.179	1.122	1.208	1.032	1.058
PROSESS total score	4.5	3.5	3.0	3.5	7.5	6.5



Plot statistics

Residues in most favoured regions [A,B,L]	199	86.9%
Residues in additional allowed regions [a,b,l,p]	27	11.8%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.9%
Residues in disallowed regions	1	0.4%
Number of non-glycine and non-proline residues	229	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	20	
Number of proline residues	13	
Total number of residues	264	

Figure 2.15. The geometric and stereo-chemical evaluation of 8th homology model of TmuB protein according to Ramachandran plot.

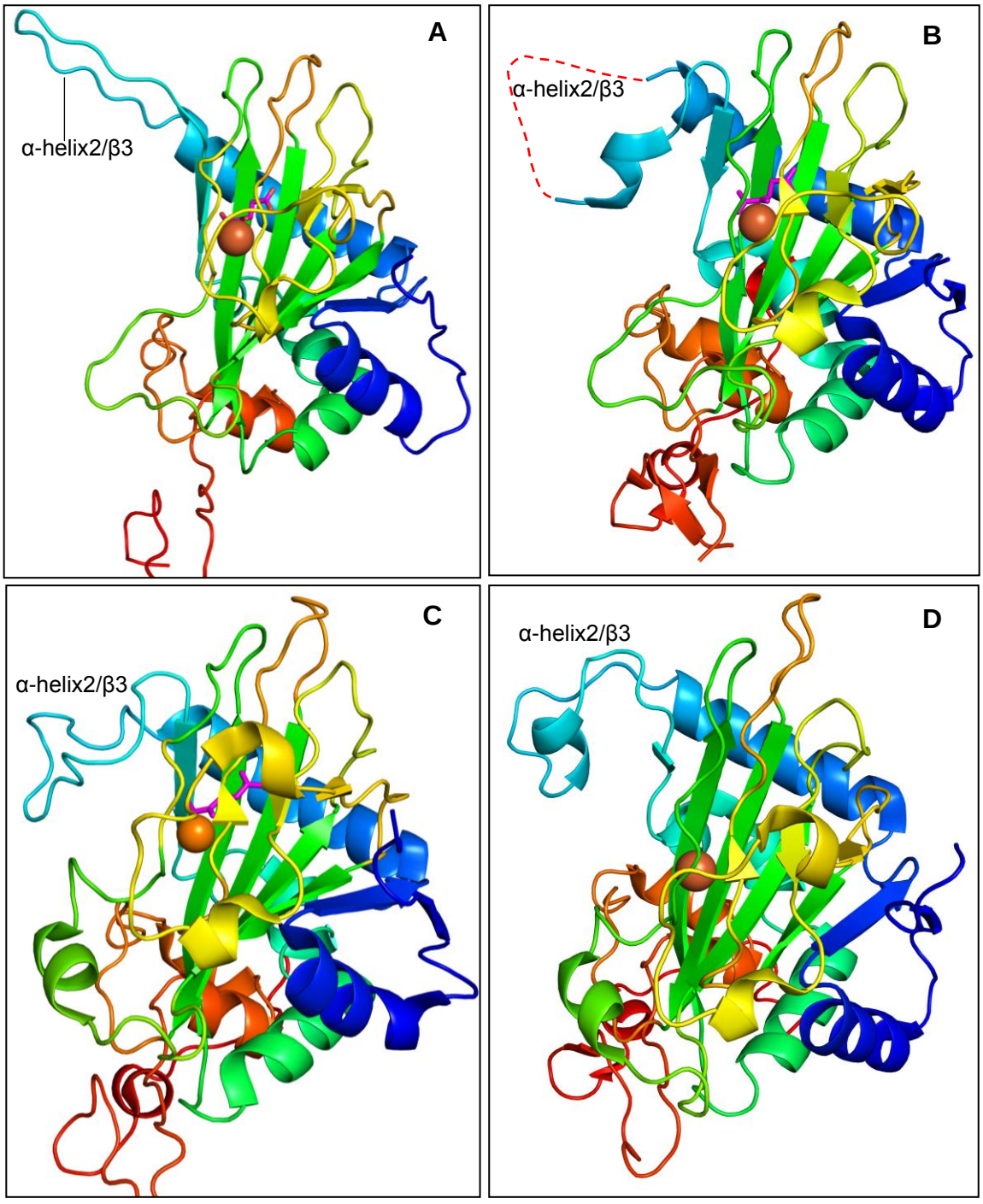


Figure 2.16. The structural model of TmuB and the crystal structures of the templates. Critical importance is the potentially mobile α-helix2/β3 loop that is missing from the EasH structure and is presents in AsqJ and FtmOx1 protein. A. TmuB model generated according to the structure of EasH protein (B) in which the loop is missing in the crystal structure (depicted as dash line). C and D. AsqJ and FtmOx1 respectively showing the loop folding over the active site (Yellow).

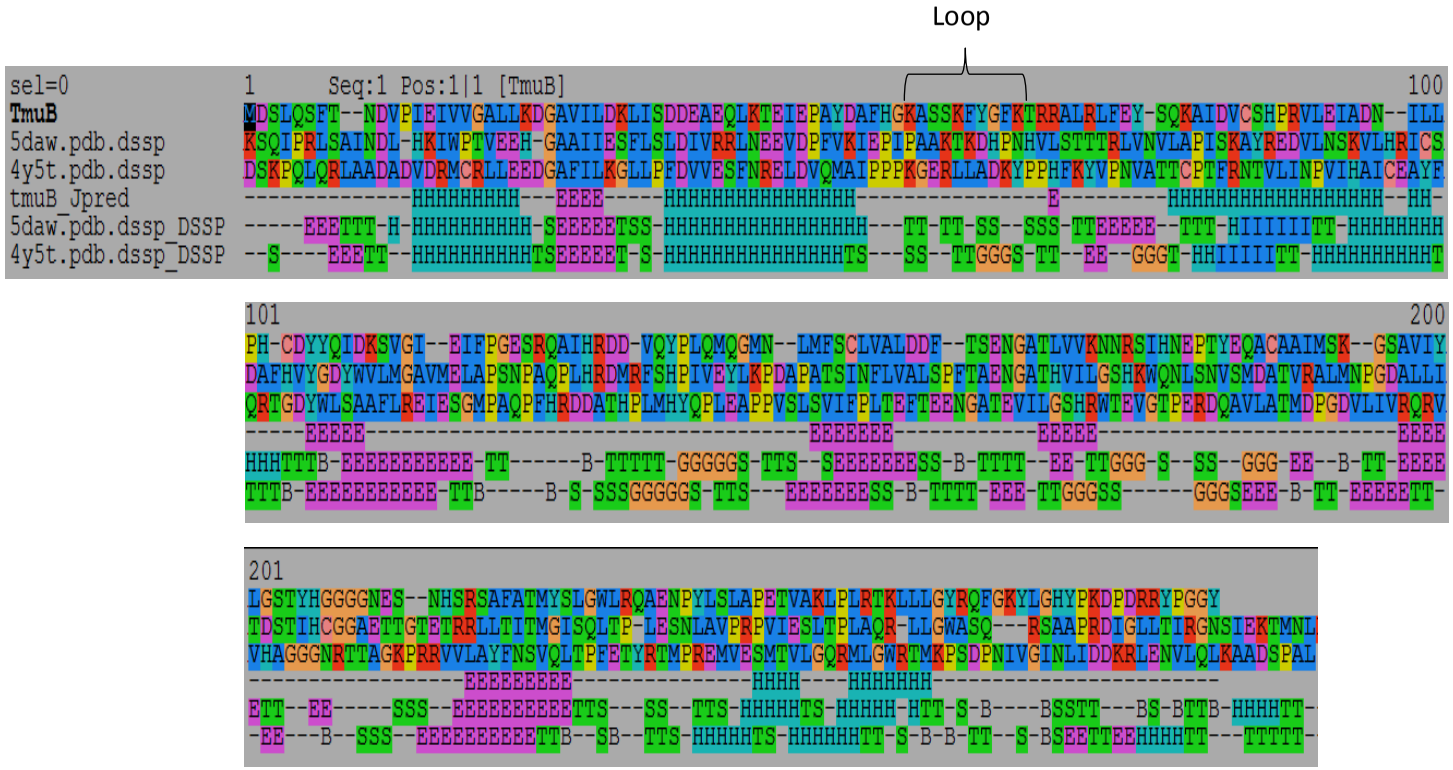


Figure 2.17. Secondary structure alignment of TmuB with the previously resolved protein templates AsqJ (5daw) and FtmOx1 (4y5t). DSSP software was used to generate the secondary structures and Seaview software was used to visualise and to refine the alignments. The TmuB model was obtained based on the AsqJ structure.

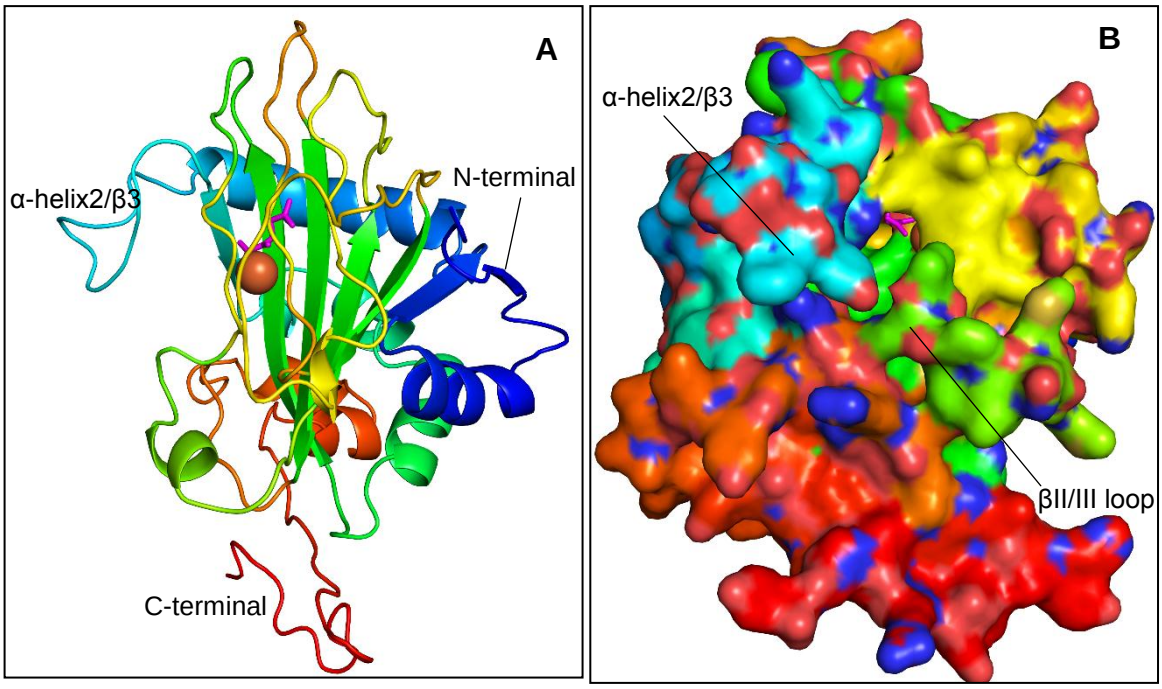


Figure 2.18. The structure of the TmuB model (15th) showing the α-helix2/β3 loop packing. A. The model was generated according to the structure of the AsqJ protein. B. The same model rotated 90^{o} degree to right to show the active site. Note that the α-helix2/β3 loop and the βII/III loop partially cover the entrance of the active site.

2.3.4.3. Receptor-ligand Docking

The substrate PA-A was docked to the TmuB model to investigate the residues which should interact with the substrate, particularly around the pyran ring where the 8OH is. The 8th and 15th of the TmuB models were used as a receptor to dock the ligand, PA-A. The docking runs were repeated with different parameters to determine the most

repeated and dominant conformations and orientations, and each run of the docking resulted in 50 docking conformations. With the 8th model, the majority of the docking results revealed that the monic acid of PA-A fits into the active site pocket while the 9-hydroxyoctanic acid fits into the external groove (Figure 2.19 A&B). The target site, C4, is within 4.38 Å of the Fe ion in presence of the cofactor (Fe) and the cosubstrate (αKG) while in absence of the latter, the distance is 3.4 Å. The pyran ring is fitted at the entrance of the pocket surrounded by the residues Arg69 (on β1), Lys105 and Ile109 (on βI) and Met208 (on βVIII). However, the residues, which are involved in the interaction with the substrate around the pyran ring, include Arg69, Lys105 and Ile109, Leu141 (on βIII) and Ala206 on (βVIII) (Figure 2.19C&D). The hydrophobic groove, which accommodates the fatty acid chain, is formed by residues from βII/III loop, end of βVIII and the turn following the βVIII.

The docking run was repeated using 15th model to test the effect of the loops and turns around the active site. These loops include the α-helix2/β3 loop, βII/III loop and the turn following the βVIII and are arranged as a triangle covering the active site. As a result, the docking conformation of the ligand to this model is quite differed from the 8th model and PA-A cannot fit into the active site, particularly the fatty acid chain. However, the location of the pyran ring near the pocket entrance and the interacted residues remained the same. The conformations obtained in this docking also suggest residues on α-helix2/β3 loop to be involved into the interaction with the fatty acid chain rather than fitting into the hydrophobic groove made by the turn following βVIII (Figure 2.20).

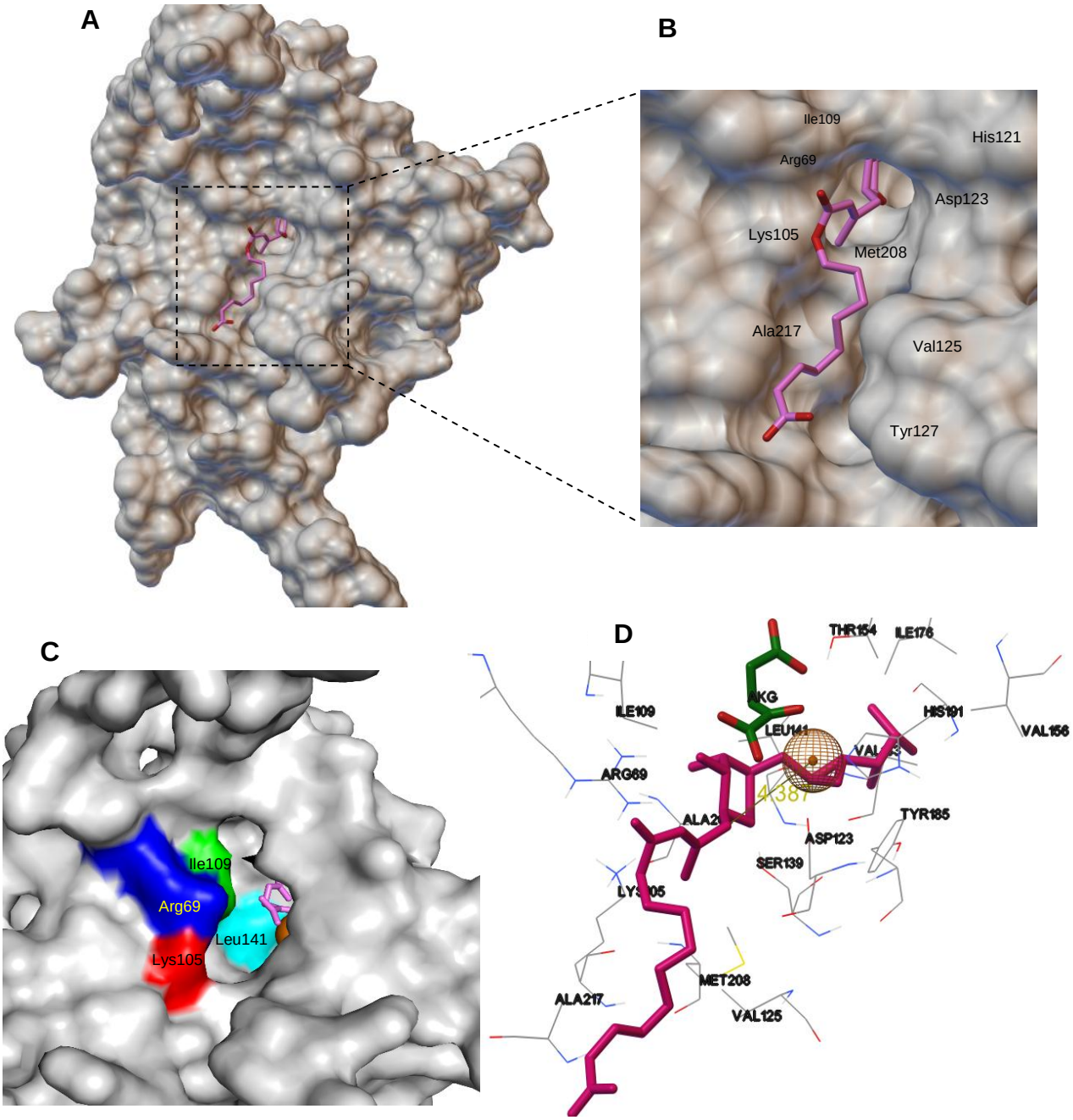


Figure 2.19. A&B. The 8th homology model of TmuB showing the PA-A docked into the active site. The fatty acid chain fits into groove along with interacting residues while the monic acid is inserting into the pocket, C. The residues in the active site which surround the pyran ring and were selected to be replaced, D. Homology model of TmuB showing the interaction between PA-A (Pink stick) and the residues close to the different substrate sectors in the active site. The target site C4 is 4.38 Å from the Fe II (Orange sphere).

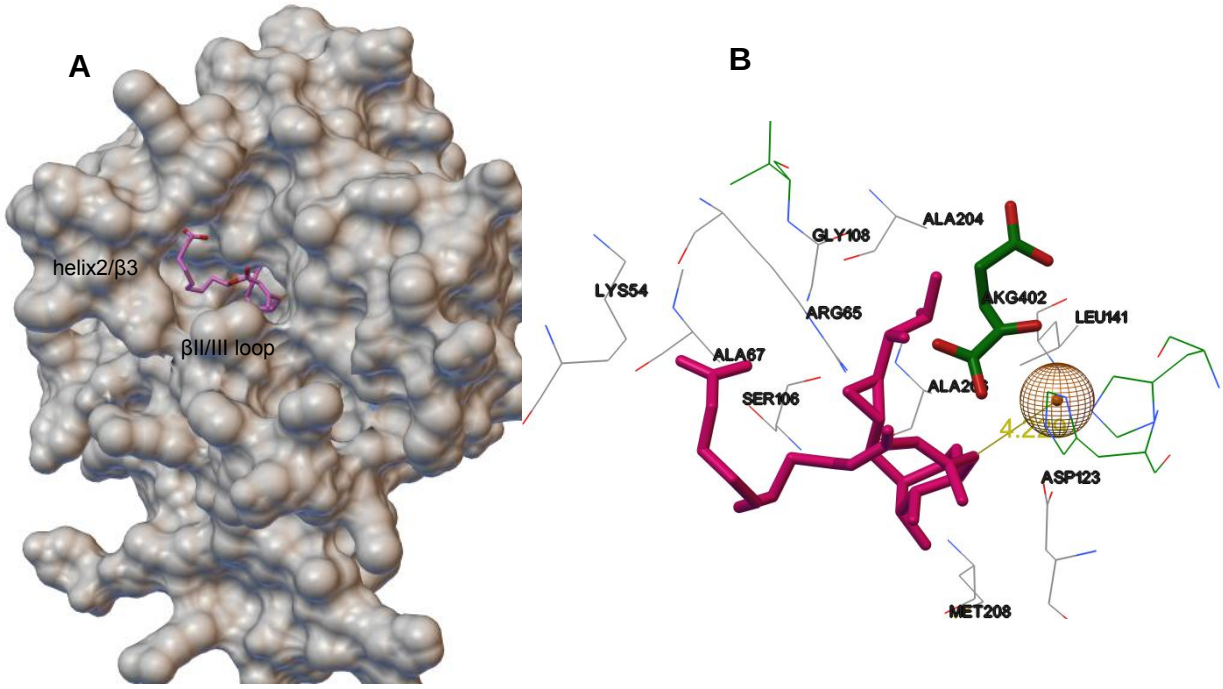


Figure 2.20. The 15th homology model of TmuB showing the PA-A docked into the active site in presence of αKG (Green stick) and Fe. A. The surface mode showing the fatty acid chain interacting with the residues Lys54 and Ala67 on helix2/β3 loop. B. The residues interacting PA-A (Pink stick) and the residues close to the different substrate sectors in the active site. The target site C4 is 4.2 Å from the Fe II (Orange sphere).

2.3.4.4. *In silico* mutagenesis

Arg69, Lys105, Ile109 and Leu141 residues were selected for mutation to smaller and more hydrophilic amino acids to create more space to accommodate a bulkier hydroxylated substrate (like PA-B) and interact with it. To find appropriate amino acids to substitute the current residues, the TmuB sequence was aligned against the sequences of nonheme-iron(II)/2-oxoglutarate-dependent dioxygenase superfamily. It has been found that Ile109 is conserved among these enzymes and might play an important role

in enzyme activity while amino acids Arg69, Lys105 and Leu141 were at less conserved positions (Figure 2.21 & 2.22). Therefore, these residues were replaced with appropriate amino acids to create mutated model of TmuB and used as a receptor in the docking process.

2.3.4.5. Receptor-ligand Docking (mutated TmuB)

The mutated model of TmuB was used as a receptor to dock the ligands PA-A and PA-B using the same parameters used previously for PA-A with native TmuB model (Figure 2.23A&B shows an example of the mutant TmuB models which were generated *in silico*). First, the ligand PA-A was docked to these mutated model to obtain conformations which are similar to those with WT model. Comparing the predicted conformations allowed us to choose mutations that might allow similar conformations to be achieved with PA-B. As illustrated in Figure 2.23C&D the docking of PA-B to the native and mutated TmuB model (I109N) was compared to obtain similar conformations to those with PA-A and native TmuB.

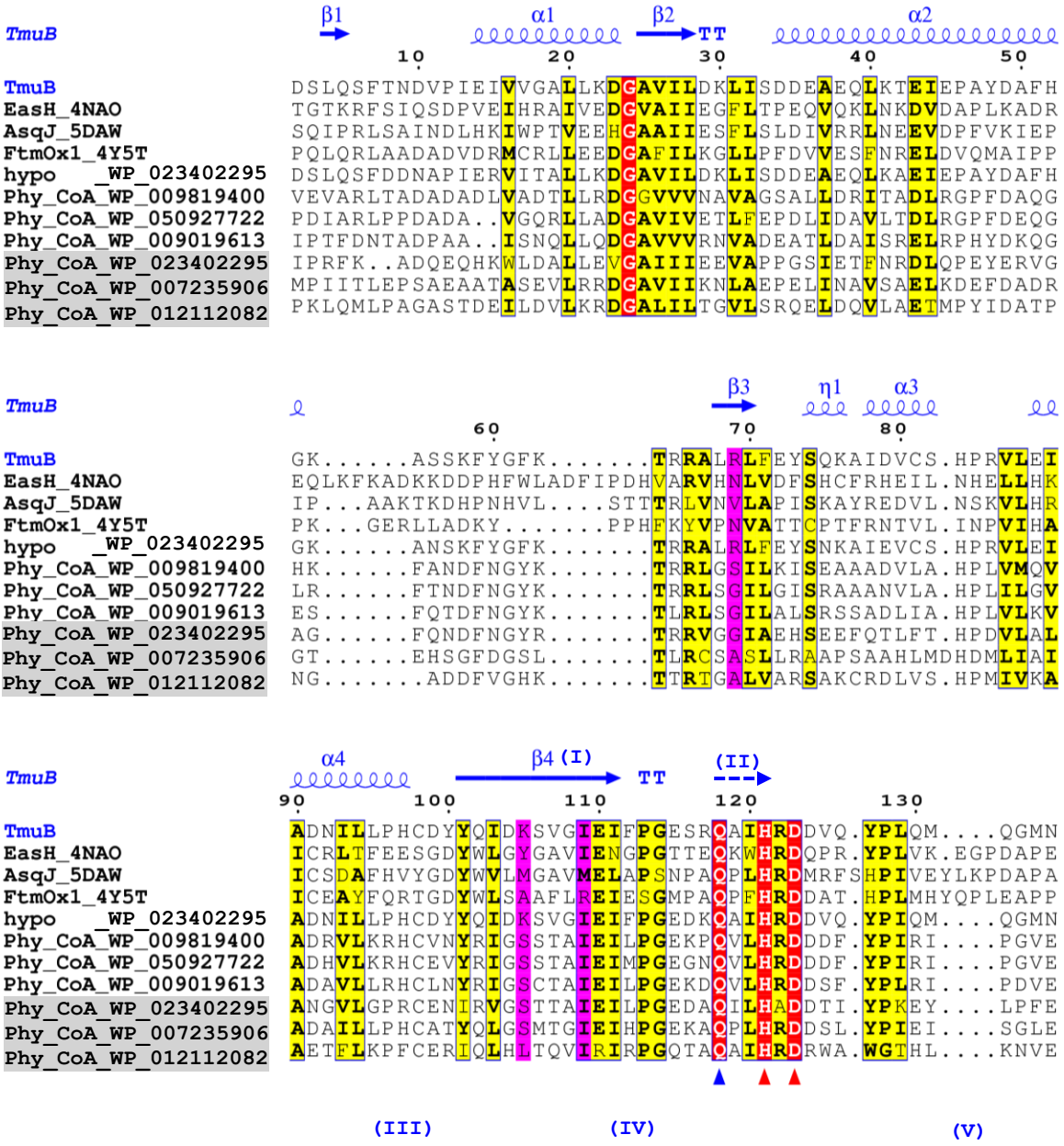
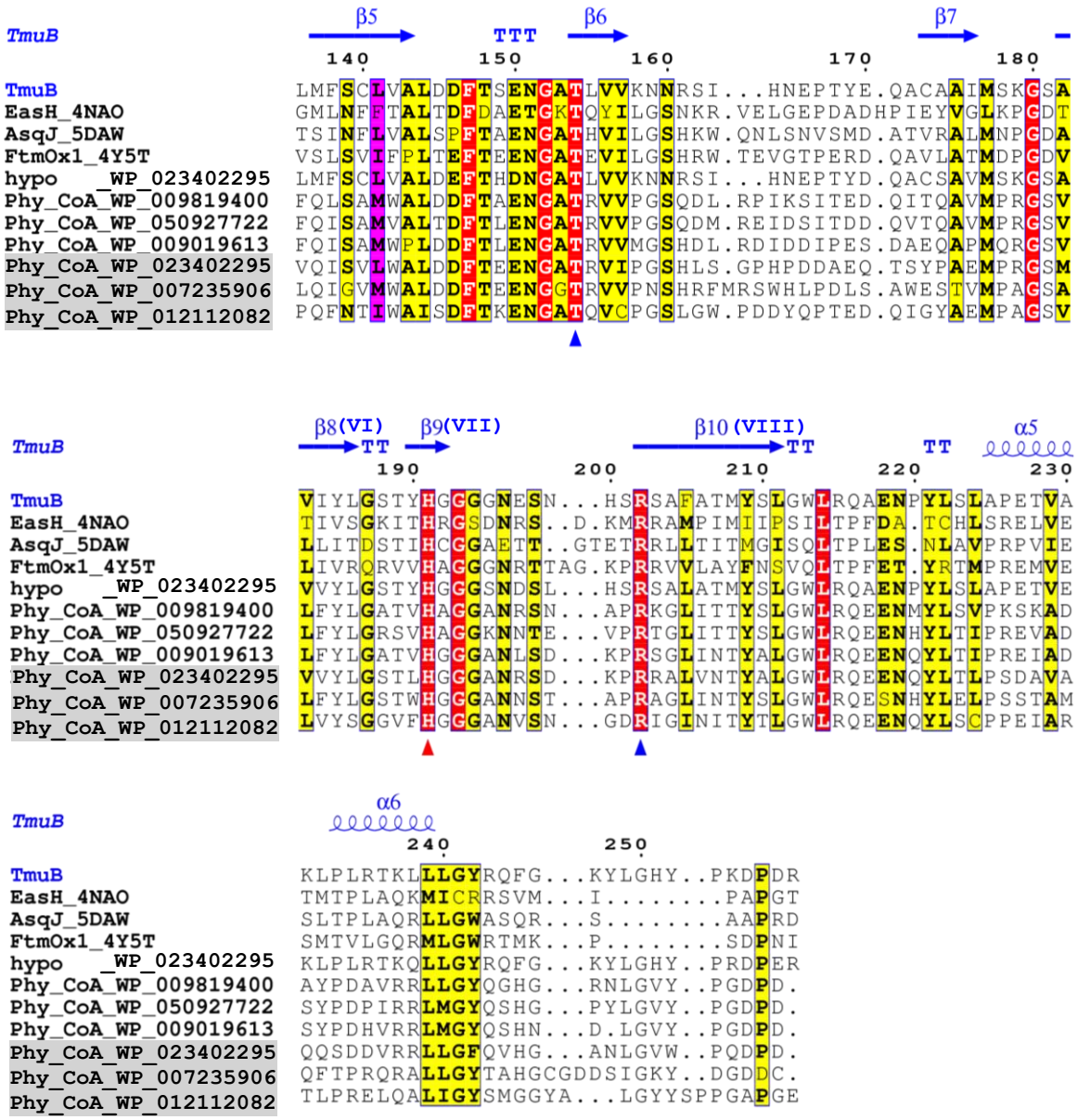


Figure 2.21. The multiple sequence alignment of TmuB protein with homologue proteins. The conserved residues in the active site HXD/H which bind with Fe ion (red arrows) and those which bind with 2KG (blue arrows), and the residues that were mutated (pink highlight), R69, I109, K105 and L141…. To be continued….



Continue of Figure 2.21. The secondary structure of TmuB is shown on the top of the sequences. The beta strands of the jelly-roll depicted by Roman numerical (**I** to **VIII**). ESPript3 was used to create this figure.

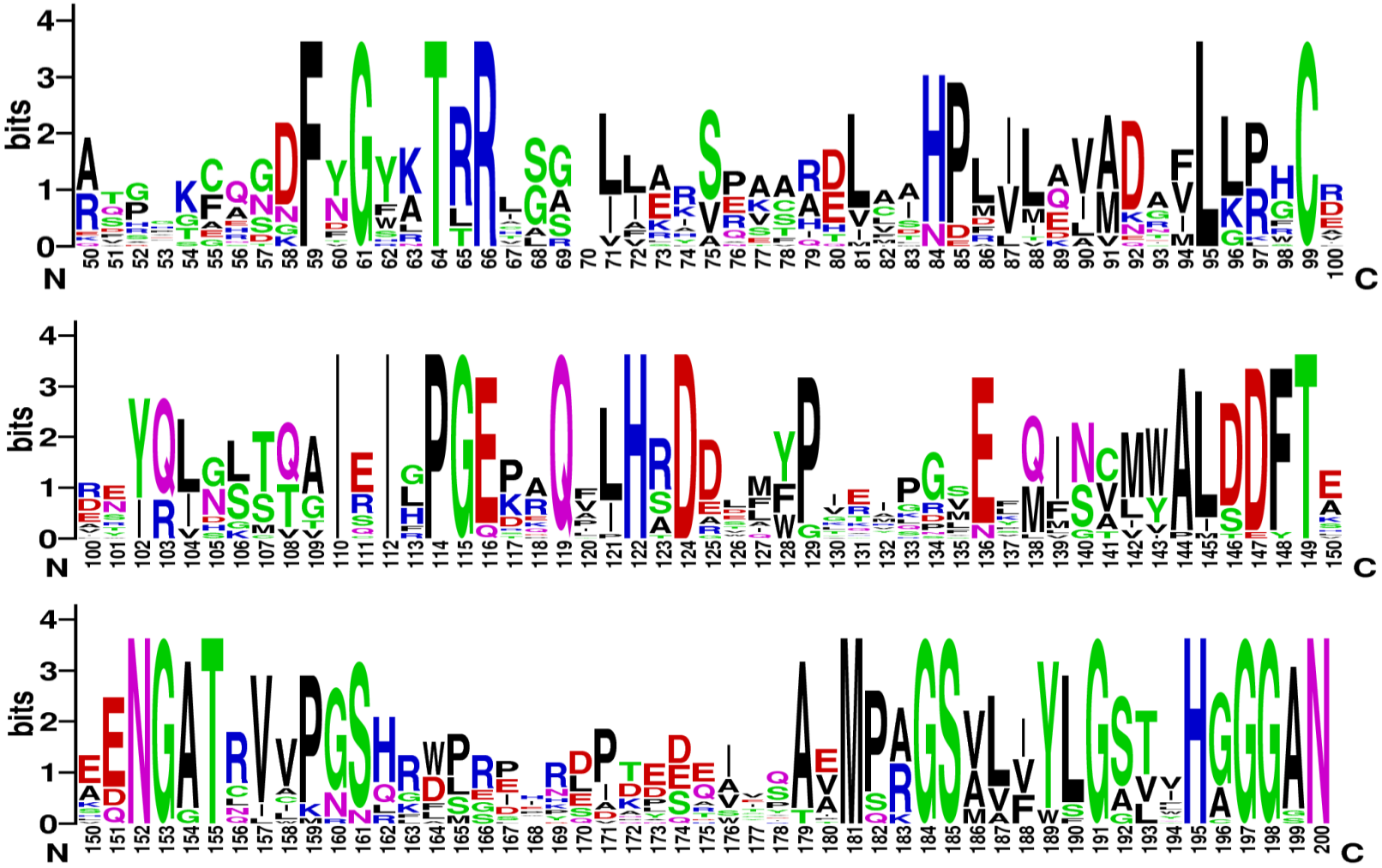


Figure 2.22. Amino acid sequence alignment logo showing the conserved residues in the active site of TmuB (H122XD124/H195) and the residues which were selected to be mutated (Arg69, Lys106, Il110 and Leu142). Note that the residues number is added by (+1) or more as gaps were generated in the alignment process.

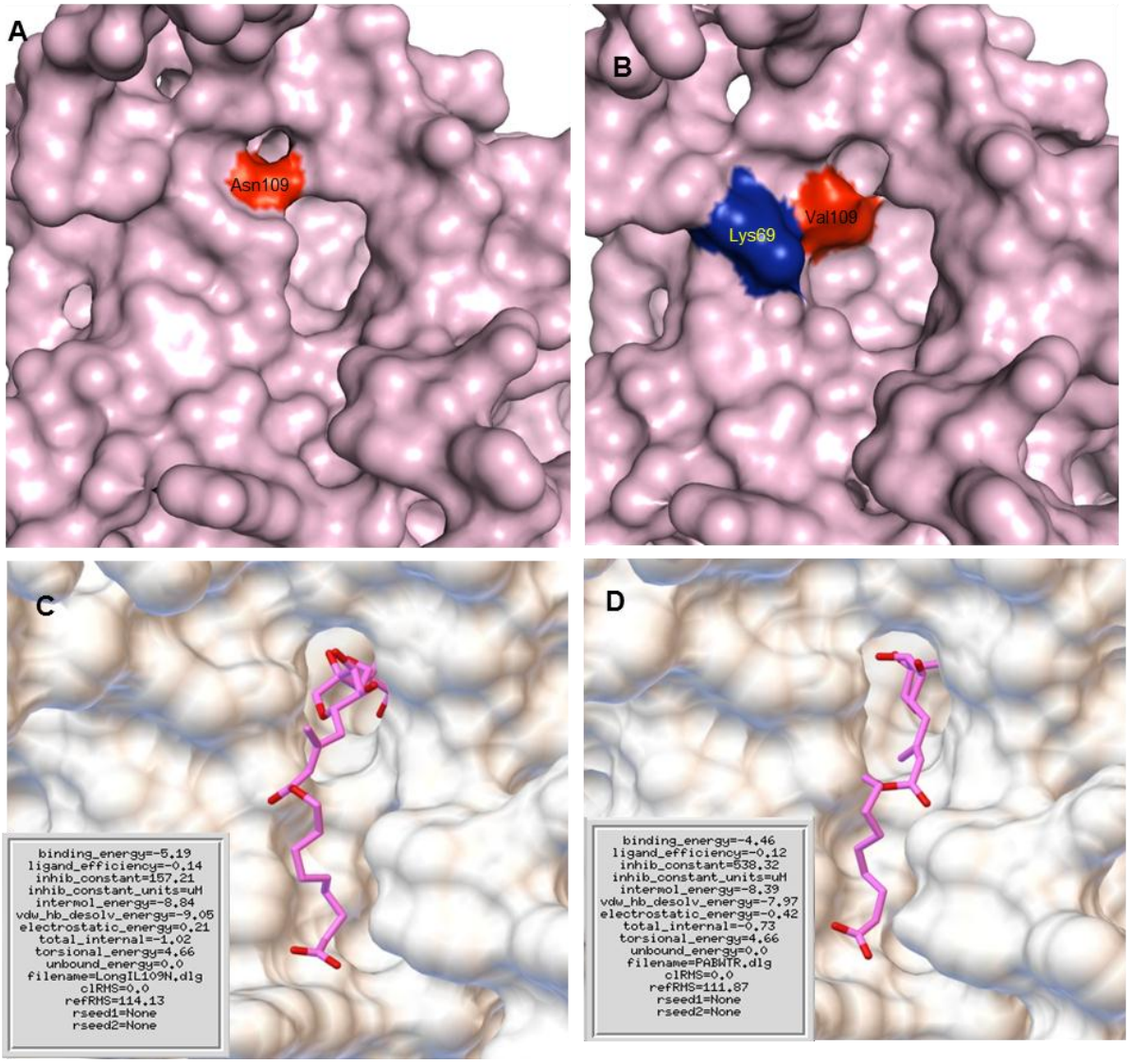


Figure 2.23. Mutant model of TmuB showing the replaced residues in the active site and how it changed the shape of the pocket. A. The residue isoleucine 109 was replaced with asparagine and this mutation inactivated TmuB both *in vivo* and *in vitro*. B. Double mutation in the active site, the residues arginine 69 and isoleucine 109 were replaced with lysine and valine respectively and this mutation did not change TmuB activity. C.PA-B (pink stick) docked into the native TmuB model and D.PA-B docked into the mutated TmuB (I109N) model. Criteria for the conformations are listed including binding energy and inhibition constant. The mutant criteria are more close to the PA-A docking to the native TmuB model.

2.3.5. TmuB mutagenesis

Modified overlap extension mutagenesis was used to create single and double point mutations in *tmuB*, cloned into expression vectors (Figure 2.24) and mobilized to *P. fluorescens* strains to test for hydroxylation of PA-B.

The four candidate residues were replaced with a variety of amino acids (Table 2.7) and the metabolite profile was analyzed by HPLC. None of these mutations changed the specificity, no appearance of any ability to hydroxylate PA-B was observed and the ability to hydroxylate PA-A was decreased. The single mutations R69K and L141V did not change the product profile while L141S and I109V reduced the 4-hydroxylated product of PA-A by 25% and L141N by 75%. Both I109N and K105E mutations abolished 4-hydroxylation of PA-A completely. A double mutation R69K - L141V was created in an attempt to accommodate the bulky substrate like PA-B but again HPLC revealed no change in enzyme activity on either substrate (Figure 2.25).

Table 2.7. The residues mutation in the active site of TmuB protein

Origin residues	Mutant	4OH product
Arginine69	Lysine	No change
Leucine141	Serine	Reduced by 25%
	Valine	No change
	Asparagine	Reduced by 75%
Isoleucine109	Valine	Reduced by 25%
	Asparagine	Inactivated
Double mutation Arg 69 Leu 141	Lysine Valine	No change
Lysine105	Glutamic acid	Inactivated

Figure 2.24. Strategy to create point mutation I109N in *tmuB* gene. The previously cloned pGME-T Easy (pHHM01) vector was used as a DNA template. The same strategy was used to create the rest of the point mutations.

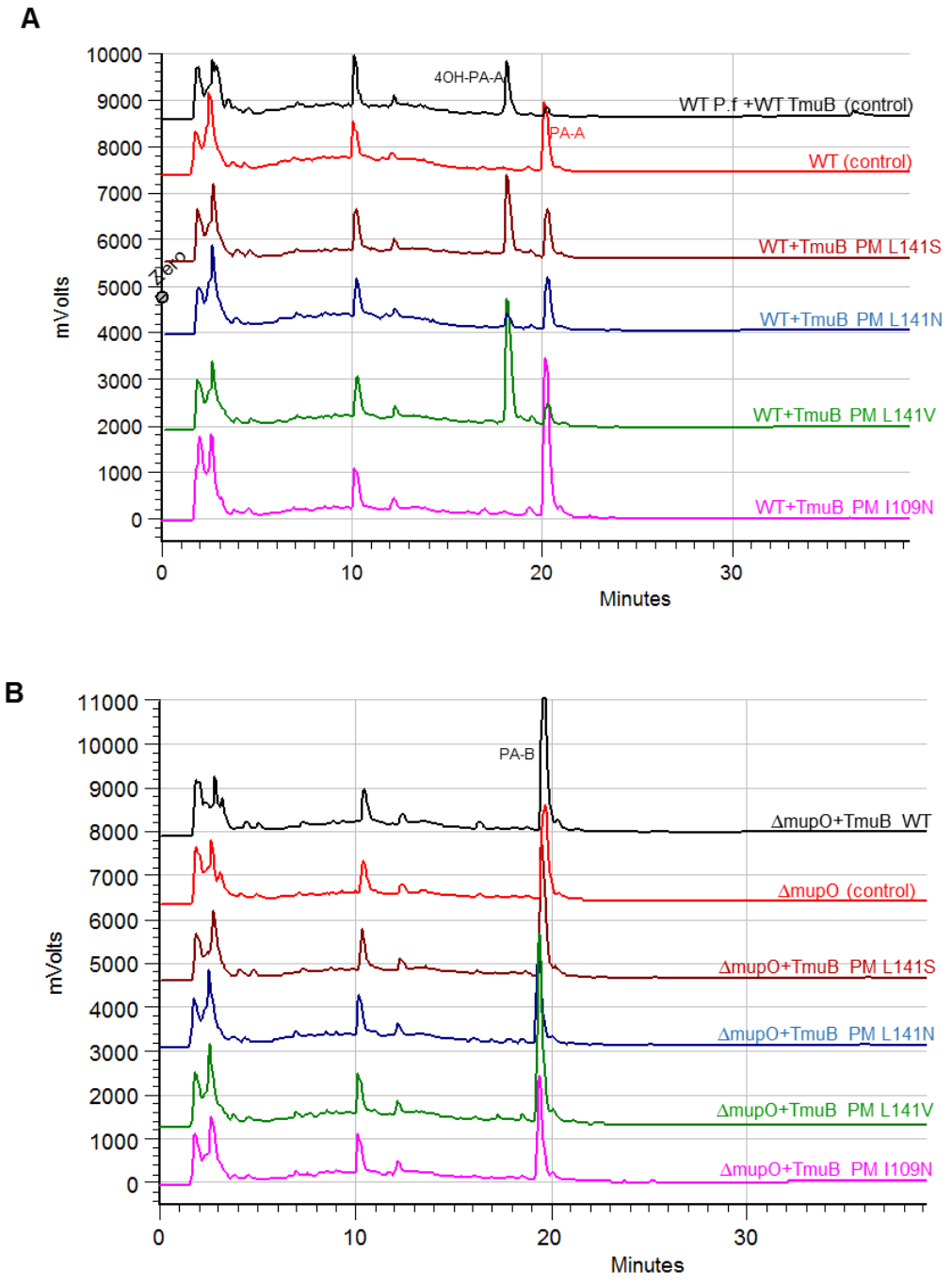


Figure 2.25. HPLC analysis of the culture supernatant of *P. fluorescens* transformed with point mutation (PM) TmuB in pJH10 vector. A: wild type *P.fluorescens* which produce PA-A (retention time 20.1min) and 4-hydroxylated in the presence of active TmuB protein (retention time 18.17min), B: Δ*mupO P.fluorescens* which produces PA-B (retention time 19.3min) and could not be hydroxylated (*in vivo*) by WT and mutant TmuB shown above.

2.3.6. TmuB protein expression and purification

To purify TmuB the *tmuB* orf was inserted into pET28a, and expressed in *E.coli* BL21 (DE3) with an N-terminal His-tag. For the point mutation (I109N) TmuB, the mutated insert which was created previously and cloned into the pJH10 vector, was used as a DNA template. *Escherichia coli* BL21 DE3 was cultivated at different temperatures and IPTG concentrations to check the initial expression. In small scale expression trials all the temperatures and IPTG concentrations induced sufficient expression of TmuB protein when analyzed by SDS-PAGE. However, it was found that the optimum expression of soluble protein can be achieved at 25 ^{o}C and with 0.5 mM of IPTG (Figure 2.26). Therefore, this optimum growth condition was used for large scale expression.

After large scale expression, the protein was purified from other protein components by affinity chromatography depending on the 6xHis-tag technique. The principle of this technique depends on the specific interaction of 6xHis-tag fusion protein with Nickel ion on agarose beads. This retains the tagged protein within the column while untagged protein can be washed out with buffer. The elution buffer contains imidazole which competes with histidine for nickel and releases the tagged protein. Buffers with gradient concentrations of imidazole (from 50 mM to 250 mM) were used to achieve this process. It was found that the most tagged protein was eluted at between 100 mM to 250 mM imidazole concentration. To collect as much pure protein as possible, the 100mM fraction was repurified again (Figure 2.27A).

Despite the repeat purification by nickel affinity chromatography, the SDS-PAGE shows a secondary band. Therefore, gel filtration was performed using HiLoad 16/600 Superdex 200pg column, the flow rate 1ml/min and 4ml per fraction, a single peak of TmuB protein comes out and the fractions from 43 to 50 were analyzed by SDS-PAGE (Figure 2.27B and figure 2.28).

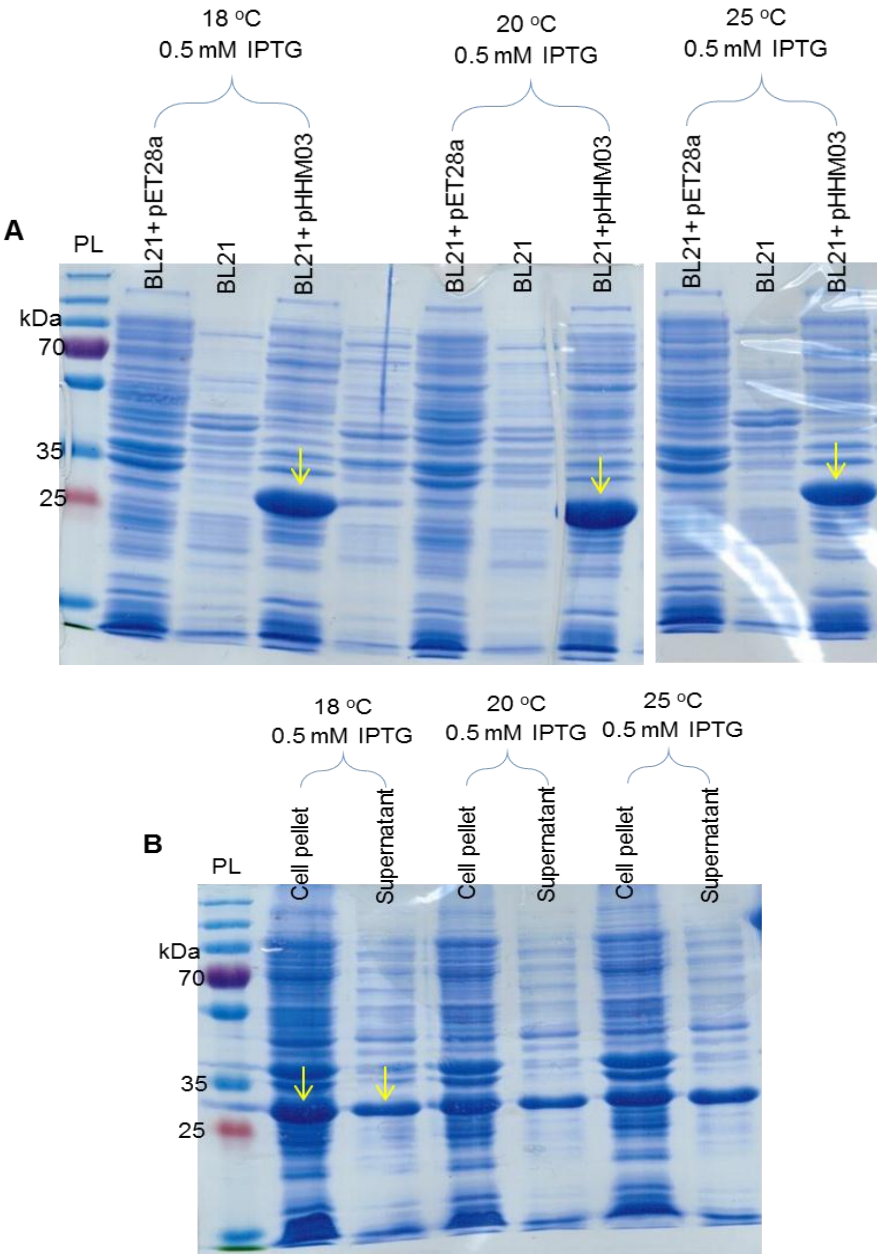


Figure 2.26. SDS-PAGE for TmuB protein overexpression to find the optimum temperature for soluble protein expression. 25ml cultures were induced by 0.5mM IPTG for 18 hours at different temperature. A: the cell pellet of the negative controls (Only BL21 and BL21 plus empty pET28a) and for the recombinant vector (pHHM03) all temperatures induced TmuB expression, B: The cell opened by sonication and the supernatant (soluble protein) separated from the cell pellet (insoluble). 25 ^{o}C and 0.5mM IPTG was appointed as the optimum condition for soluble protein expression. TmuB protein (29.5 kDa) is indicated by yellow arrow. PL: protein ladder

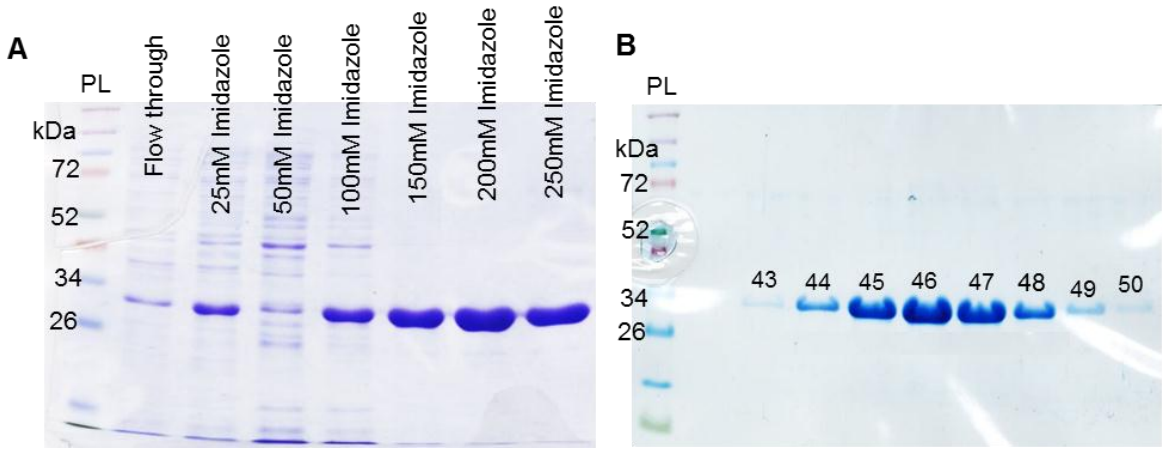
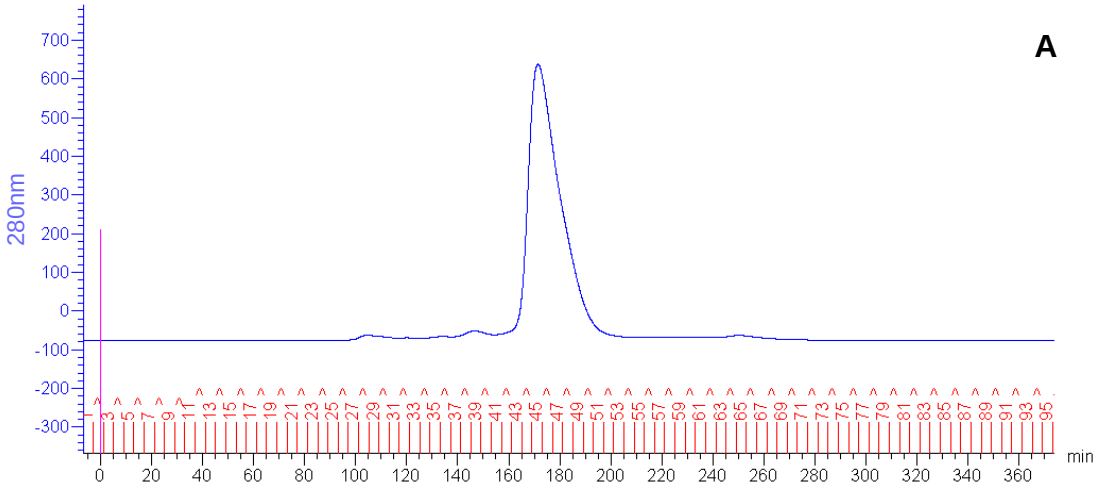


Figure 2.27. SDS-PAGE analysis for TmuB protein purification. A: purification by Ni-NTA affinity chromatography, the protein was eluted by different imidazole concentrations as shown above. The last four fractions (100mM to 250mM) were collected and repurified by gel filtration. B: SDS-PAGE for gel filtration fractions (43 to 50).

2.3.7. TmuB dimerization *in vivo*

The Bacterial Adenylate Cyclase Two-Hybrid System was used to test protein-protein interaction *in vivo*. The TmuB protein was fused to both C- and N-terminus of the adenylate cyclase domains T25 and T18 as described in 2.2.21 section. The colonies of the positive control of the *E.coli* BTH101 and DHM1 were red and blue on the MacConkey and LB/X-gal agars respectively, while the negative control colonies were pale color or white on both plates. The *E.coli* BTH101 and DHM1 colonies, which contain constructed vectors, were darker than the negative control but brighter than the positive control (Figure 2.29). These colonies were tested to measure the β-galactosidase enzyme activity to confirm the results. The level of β-galactosidase activity



Protein/compound	M_r (10^3)	log(M_r)	V_{el} (ml)
Thyroglobulin	670	5.826	119.8
Ferritin	440	5.643	132.9
Catalase	232	5.365	158.7
gamma-Globulin	158	5.199	158.0
Aldolase	158	5.199	164.0
Ovalbumin	44	4.643	204.0
Myoglobin	17	4.230	239.9
Blue Dextran	>2000		112.5
Vitamin B-12	1.35		293.1

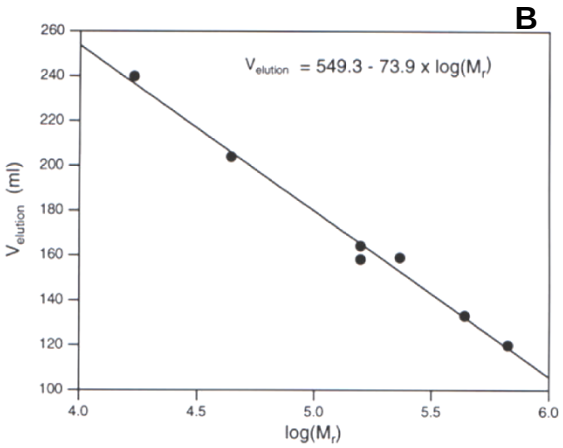


Figure 2.28. Gel filtration for TmuB protein purification. A: The protein was eluted at 180 ml. B: Calibration curve to estimate protein size. The size of TmuB (plus His tag) = 31121.16 Da (Log=4.49) which should be eluted around 220ml, while elution at 180ml (log 62242.32=4.79) of the shown peak refer it as a dimeric state.

was below the positive limit which may reveal that this technique provides no evidence for protein-protein interaction *in vivo* (Table 2.8). However, this may contradict the gel filtration chromatography which shows the TmuB as a dimer protein (Figure 2.28)

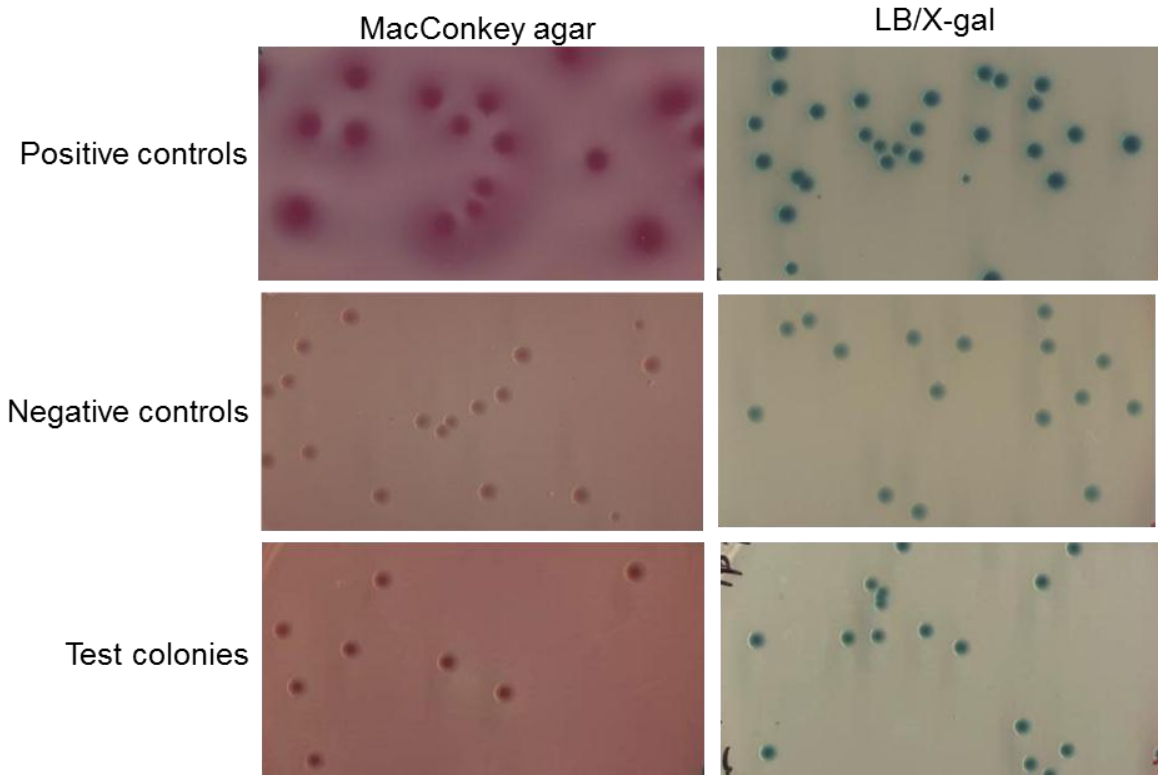


Figure 2.29. The *in vivo* protein-protein interaction assay using Bacterial-Two Hybrid. pUT18zip and pKT25zip vectors were transformed into BTH101 and used as a positive control. The empty pUT18 and pKT25 vector in BTH101 or empty BTH101 were used as a negative control. TmuB was inserted to pUT18 and pKT25 to test the dimerization.

Table 2.8. The result for *in vivo* protein-protein interactions using Bacterial Two Hybrid system.

Plasmid 1	Plasmid 2	Phenotype on MacConkey agar	β-galactosidase activity (unit/mg)
pKT25-zip	pUT18-zip	Red (positive control)	1583
pKT25	pUT18	Pale red (negative control)	56.7
pHHM16	pHHM17	Pale red	51.7
pHHM18	pHHM19	Pale red	56.2
pHHM16	pHHM19	Pale red	60.74
pHHM18	pHHM17	Pale red	58.5

2.3.8. TmuB activity *in vitro*

Both the native and mutant I109N TmuB protein were tested on PA-A in the presence of the co-substrate (α-ketoglutaric acid) and co-factor (Fe^{+2}). As an initial test, the reaction was incubated overnight at 25^{o}C. When all the reaction components were included in the reaction mixture, the native TmuB generated a new peak at 18.2 min that was well separated from the PA-A (20.2min) (Figure 2.30A). The same reaction was repeated with the mutant I109N, which showed no activity (Figure 2.30B). To confirm the hydroxylation of the product, large-scale reaction was set up, the product peak was purified by HPLC, and analyzed by MS and NMR at School of Chemistry, University of Bristol, confirming that the new peak is 4-hydroxyl PA-A. This is consistent with the *in vivo* experiments in which the native TmuB showed activity while the mutant I109N did not.

To test the ability of TmuB protein to hydroxylate PA-B *in vitro*, the same reaction with the same conditions except using PA-B as substrate was repeated. A new peak generated from TmuB activity on PA-B was detected by HLPC (RT=17.2min) and around 43% of the PA-B was converted to the product at 25 ^{o}C after overnight incubation. The new peak was purified by HPLC and MS detected a peak of MW 532 (Figure 2.31). The location of the hydroxylation was elucidated by NMR which confirmed it as 4-hydroxyl PA-B. In attempt to convert all the substrate in the reaction to 4OH product, the reaction was repeated at 23 ^{o}C, 25 ^{o}C, 30 ^{o}C and 37 ^{o}C. The proportion of PA-B hydroxylated was found to vary, but none of these temperatures enable TmuB to hydroxylate PA-B completely (Figure 2.32).

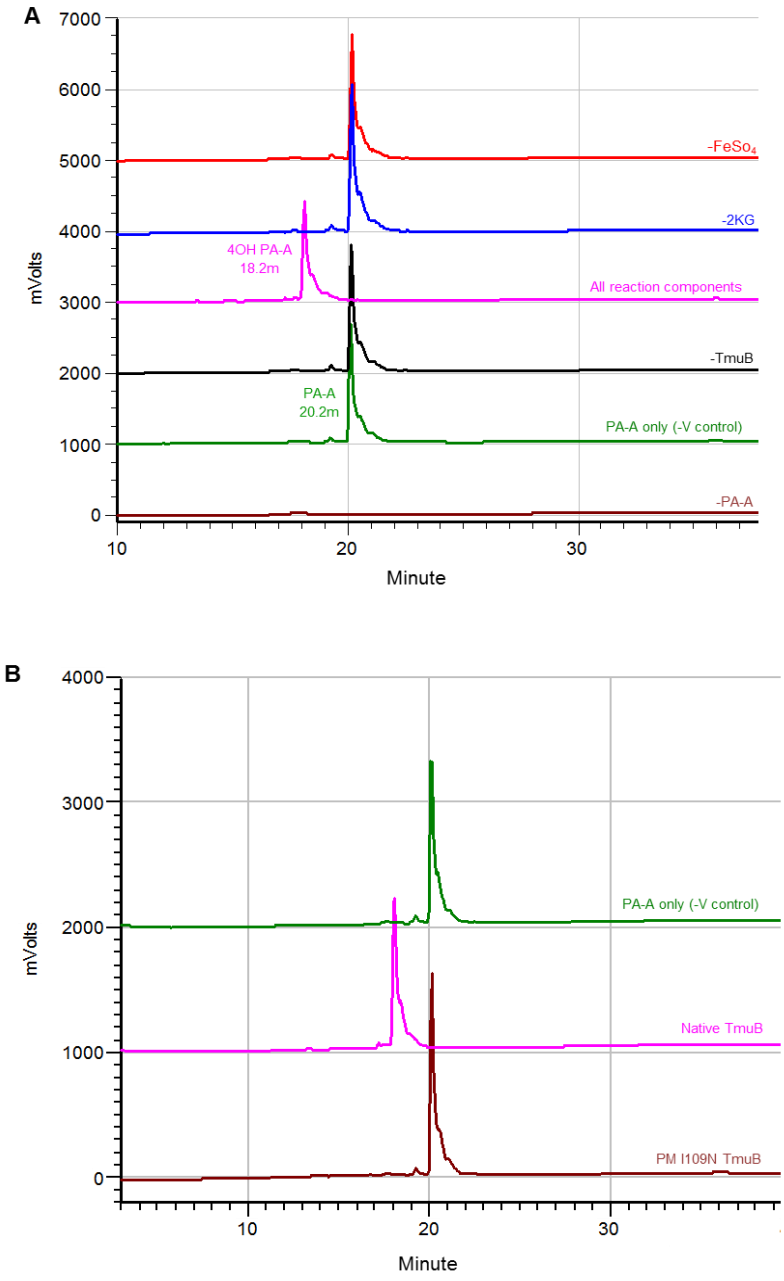


Figure 2.30. HPLC analysis to test the activity of purified TmuB protein *in vitro*. The reaction set up in 500µl reaction volume (35µM TmuB protein, 0.2mM $FeSO_4$, 0.5mM α-keto glutaric acid and 50µM pseudomonic acid A in 20mM Tris-HCl buffer, pH7.5 and 500mM NaCl) at 23 ^{o}C overnight. A: Activity of purified TmuB on PA-A is dependent on all reaction components. When all components included (pink line) the PA-A peak shifted from 20.1 min to 18.17 min. B: PM I109N TmuB protein showed no activity. (Green line) buffer contains PA-A only as a negative control.

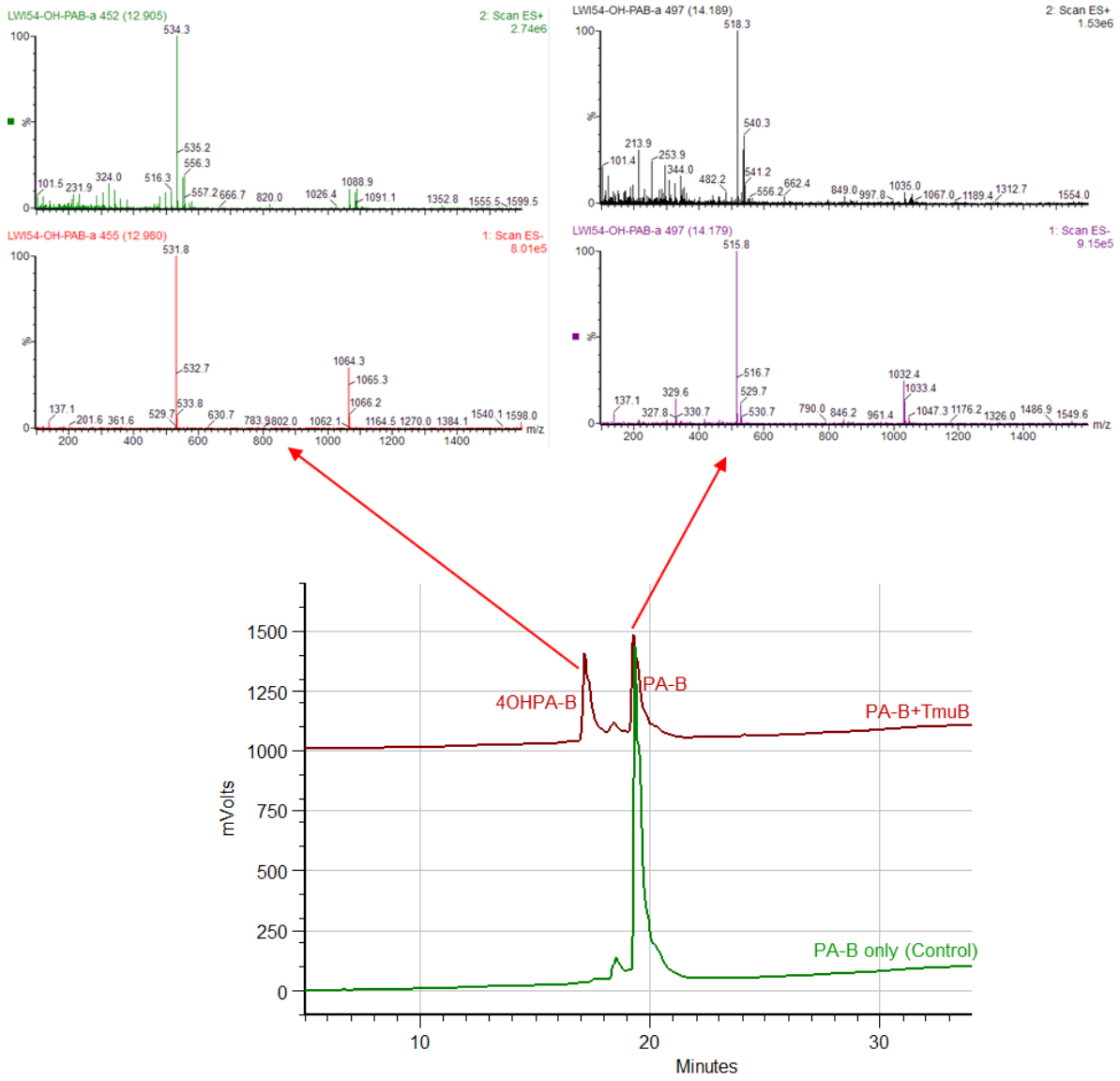


Figure 2.31. HPLC and MS analysis of pseudomonic acid B hydroxylated *in vitro* by TmuB activity. The retention time of PA-B shifting from 19.3min to 17.4min and MW from 516 to 532 indicates the hydroxylation of PA-B.

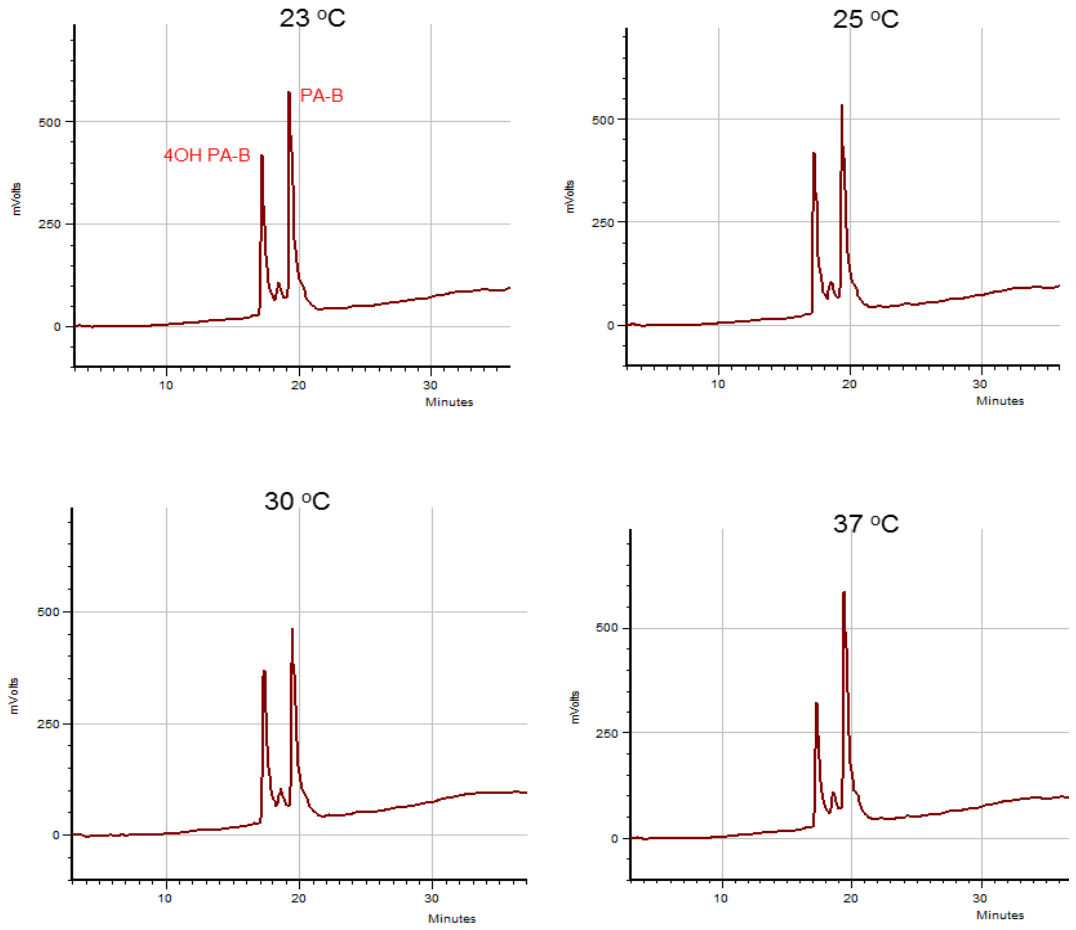


Figure 2.32. HPLC analysis of products of TmuB activity using pseudomonic acid B (PA-B) at different temperature. About 41% of the substrate hydroxylated at 23 ^{o}C, 43% at 25 ^{o}C, 45% at 30 ^{o}C and 36% at 37 ^{o}C. The retention time of PA-B and 4OH PA-B are 19.3 min and 17.6 min respectively.

2.3.8.1. Reaction condition optimization

Various reaction conditions were used to optimize TmuB activity. These variations included temperature, TmuB concentration, concentrations of the co-factor and co-substrate, and the incubation time. Reactions were set up at 23^{o}C, 25^{o}C, 30^{o}C and 37^{o}C and followed for up to 180 minutes showing that the rate of reaction increased with temperature (Figure 2.33). Different concentrations of TmuB protein (17µM, 35 µM,

50µM and 85 µM) were used with different concentrations of the substrate (Table 2.9), and the reaction was incubated for 1 min at 23 ^{o}C. Different concentrations of cosubstrate (α-keto glutaric acid) and cofactor ($FeSO_4$) were used. The product peaked at 0.5mM and 0.25mM of cosubstrate and cofactor concentration respectively (Figure 2.34).

Table 2.9. The effect of TmuB protein concentration on the reaction. The table shows the amount of the product (µM) by each concentration. The reaction set up at 23 ^{o}C for 1min.

Substrate concentration (µM)	17µM TmuB	35µM TmuB	50µM TmuB	85µM TmuB
3.75	0.45	2.75	3.6	3.75
7.5	1.98	5.7	6.5	7.4
15	5.1	10	11.64	13.4
30	11.0	17	19.75	24
60	18.2	21	28.95	36.9
120	20.1	22	22.7	55.7
240	20.2	22.5	23	56.8

The condition that converts small amount of the substrate to the product was identified to find the early linear phase of the reaction. The 17µM TmuB protein for 1 min at 23^{o}C hydroxylated about 30% of the substrate. From these experiments the optimum condition for TmuB activity was identified and used to analyse the enzyme kinetics of TmuB using different substrates. A range of PA-A concentrations were run by HPLC to plot the calibration curve and calculate the amount of the product in the enzymatic reaction (Figure 2.35).

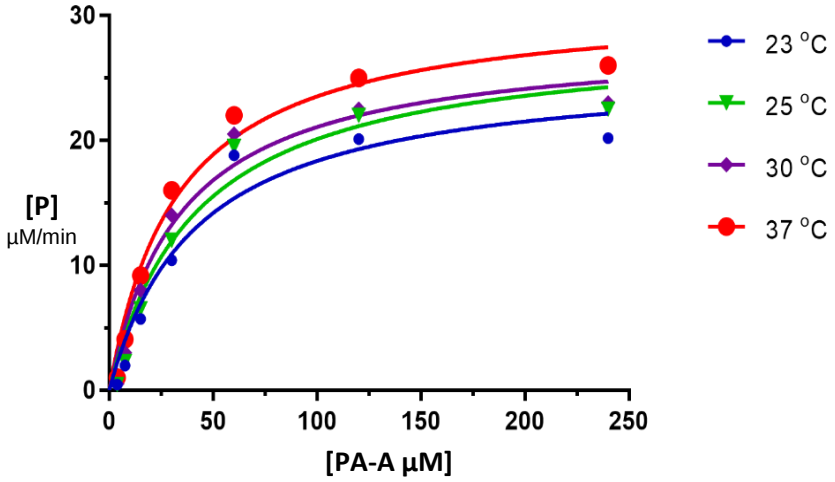


Figure 2.33. The products of TmuB activity at different temperature using PA-A as a substrate. The reaction set up in 500µl reaction volume containing 17µM TmuB, 0.25mM $FeSO_4$ (Co-factor), 0.5mM α-keto glutaric acid (Co-substrate) and various substrate as shown above in buffer (20mM Tris-HCl, pH7.5 and 500mM NaCl).

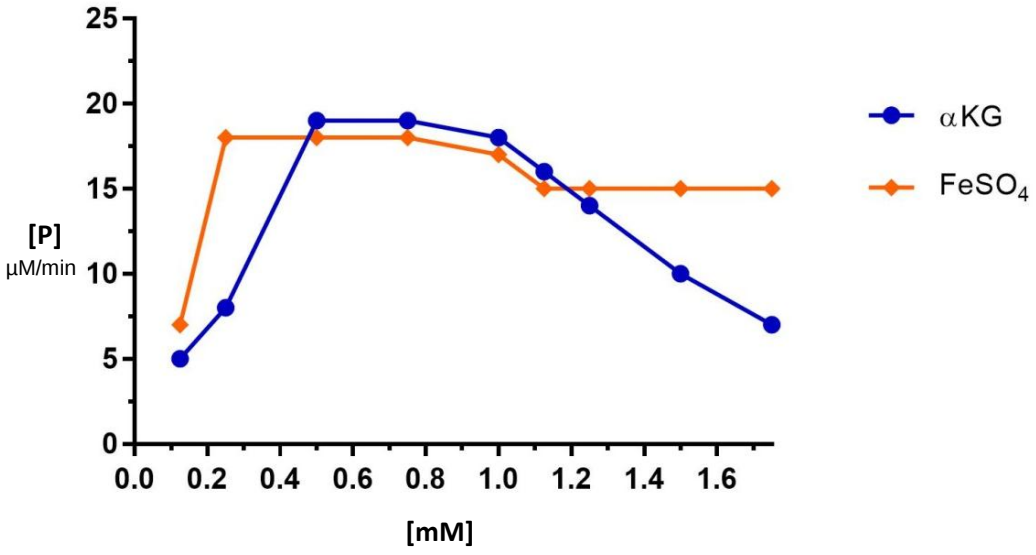


Figure 2.34. The effect of co-substrate (αKG) and co-factor ($FeSO_4$) on the product. The reaction set up for 1 min at 23 °C containing 17µM TmuB protein plus 60µM substrate PA-A. The product reached the peak at 0.5mM αKG and 0.25mM $FeSO_4$.

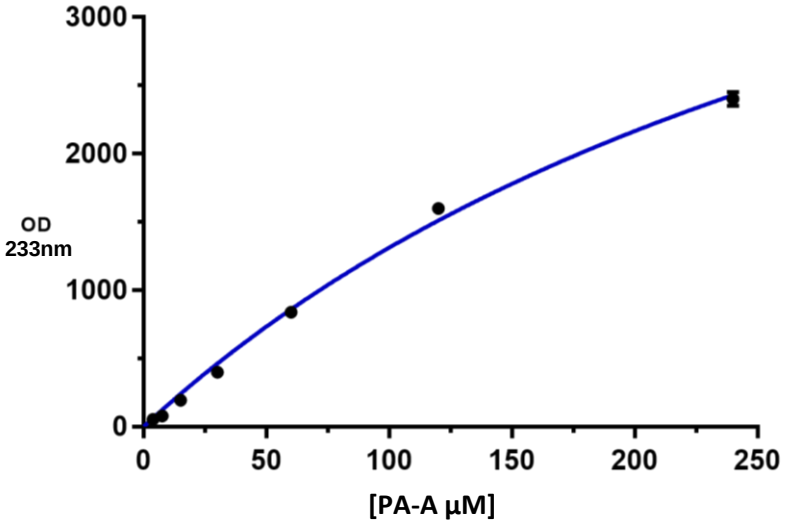
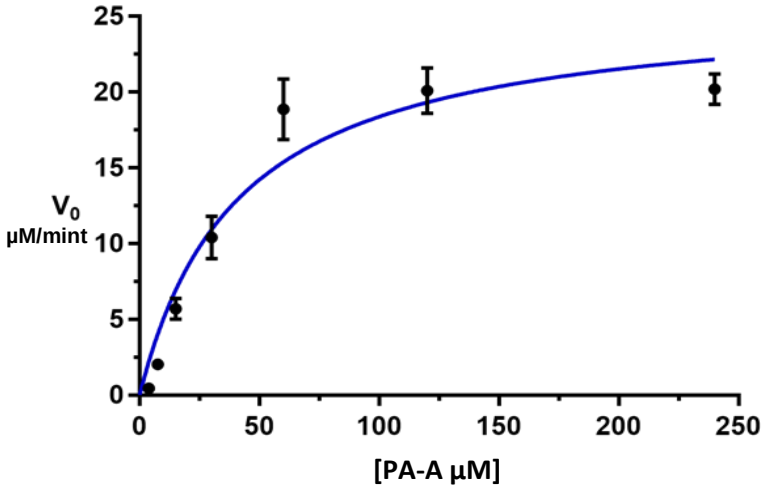
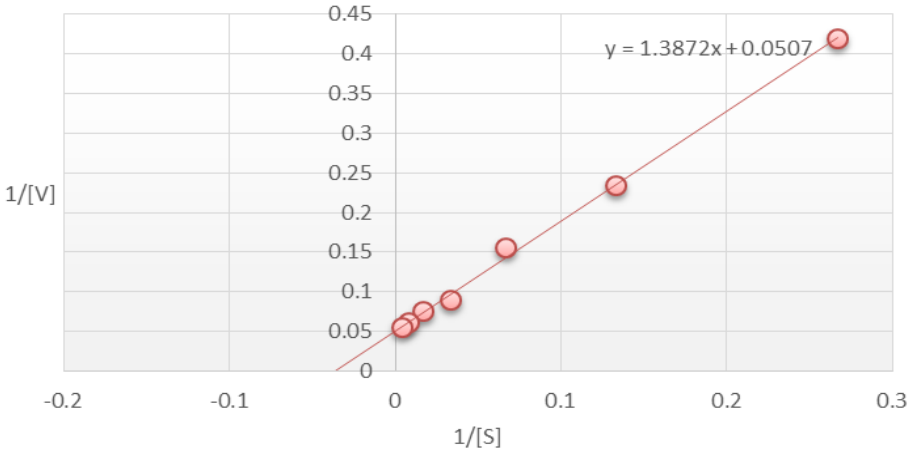


Figure 2.35. HPLC calibration curve for PA-A concentrations.

2.3.8.2. Enzyme kinetics

The enzyme kinetics of the TmuB activity was identified according to Michaelis–Menten and Lineweaver-Burk analysis. A range of substrate concentrations (3.75–240 µM) was added to the reaction mixture (17µM TmuB, 0.5mM αKG, 0.25mM $FeSO_4$ and buffer made up to 500µl) incubated at 23 ^{o}C for 1 minute. The initial velocity (V_0) of the reaction was calculated by dividing the amount of product formed by the time of incubation. The initial velocity was plotted versus substrate concentrations [S] to construct a Michaelis–Menten model. V_{max} and K_m values for TmuB with PA-A as a substrate were 20.2 µM/min and 28 µM respectively (Figure 2.36).





Intercept Y=1/Vmax
0.0507=1/Vmax
Intercept X= -1/Km
0.0366=-/Km

Figure 2.36. Enzyme kinetics for TmuB using PA-A as substrate. (Top) Michaelis- Menten model, V_{max}=20.2 µM/min and K_m= 28 µM and (Bottom) Lineweaver-Burk model, V_{max}=19.7 µM/min and K_m=27.3 µM.

Reactions were set up to obtain the similar V_{max} and K_m values for TmuB using PA-B as substrate. Short time incubation reactions did not show any product, therefore reactions were incubated for longer time. After 6 hours of incubation, V_{max} and K_m were 0.039 µM/min and 85 µM respectively (Figure 2.37).

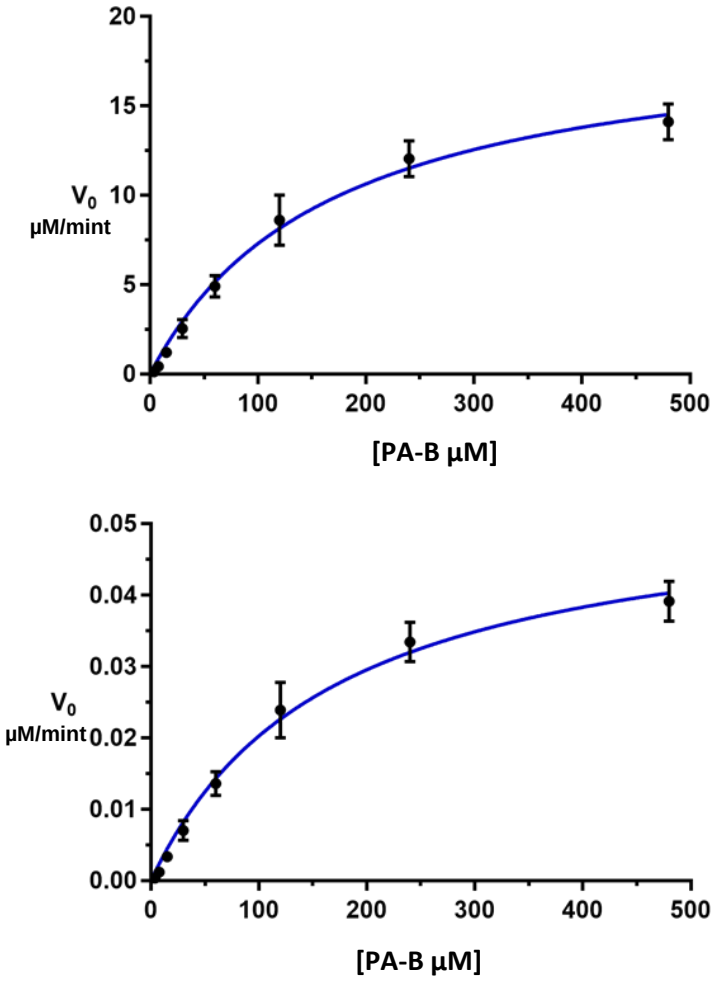


Figure 2.37. TmuB kinetics using PA-B as substrate. (Top)The reaction set up for 6 hours at 23 ^{o}C and (Bottom) The product divided by 360 to find out the product within one minute. V_{max} = 0.039 µM/min and K_m= 85 µM.

The K_{cat} value for TmuB using PA-A as a substrate was 1986 s^{-1} while using PA-B as a substrate the K_{cat} for TmuB was 3.83 s^{-1}, an approximately 500-fold difference. The catalytic efficiency (K_{cat}/K_m) of TmuB for these two substrates are 70.9 and 0.045 respectively (Table 2.10).

Table 2.10. TmuB kinetics using different substrates.

Substrates	V_{max} (μM/min)	K_m (μM)	K_{cat} (s^{-1})	K_{cat}/K_m (μM^{-1} s^{-1})
PA-A	20.2	28	1986	70.9
PA-B	0.039	85	3.83	0.045
TMC	22.5	29.5	2212.5	75

It appears that PA-B still binds reasonably to TmuB but is not a good substrate, so it should be a competitive inhibitor of the TmuB reaction on PA-A. The reaction on PA-A was therefore carried out at a range of PA-B concentrations (30μM, 60μM and 120μM). The result is consistent with this hypothesis which revealed that the K_m increased with PA-B increasing while V_{max} remained unchanged (Figure 2.38).

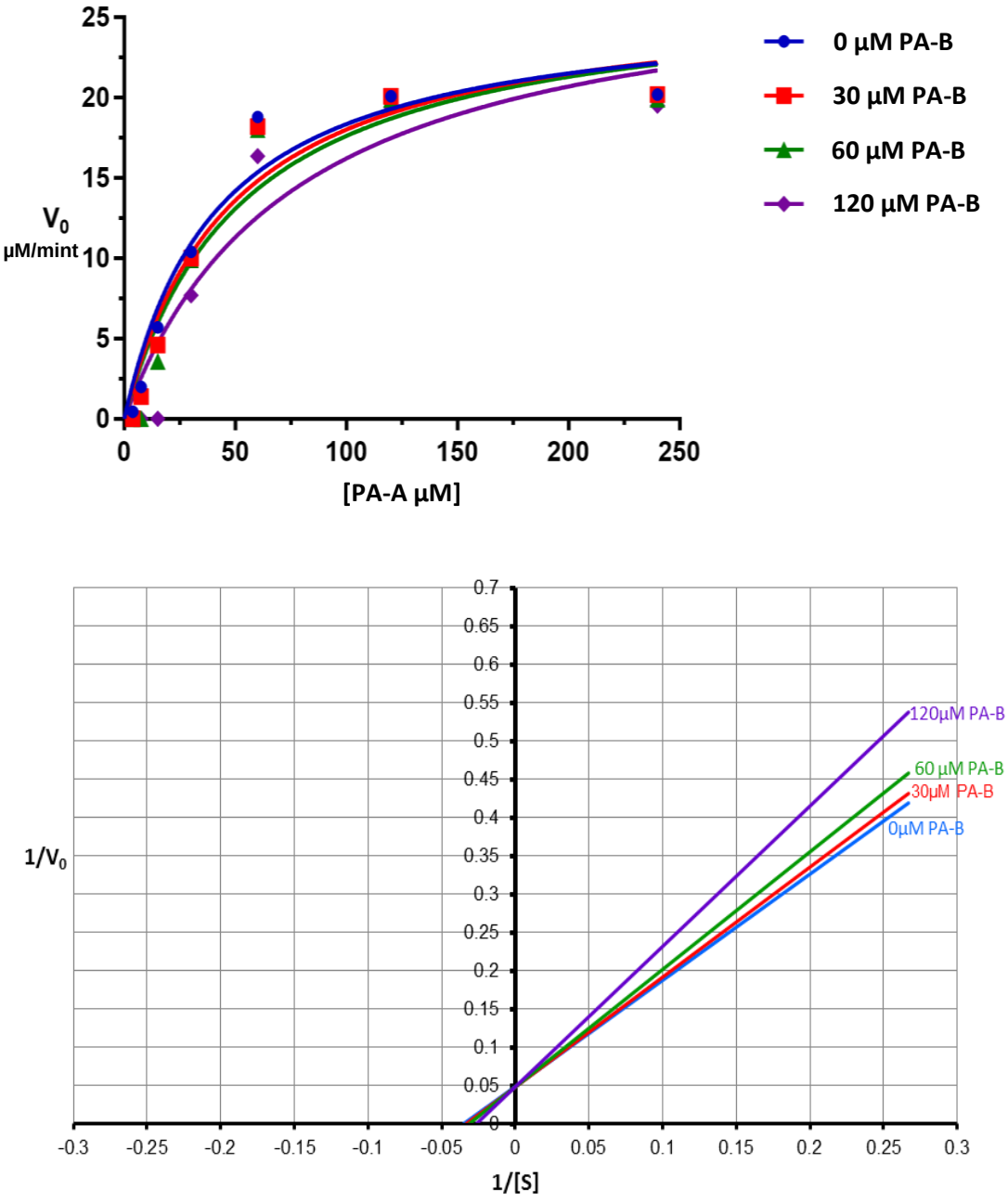


Figure 2.38. The influence of PA-B on TmuB activity while catalysing PA-A hydroxylation. The K_m increased with the inhibitor concentration increasing (as shown above) and V_{max} remained unchanged. With adding of 30µM, 60 µM and 120 µM of PA-B, the K_m increased to 28.9 µM, 31 µM and 37 µM respectively.

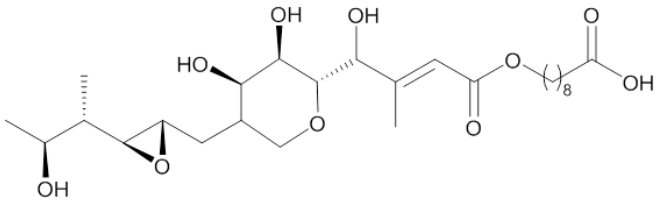
2.3.9. Antibacterial activity of pseudomonic acid derivatives

The new derivatives of pseudomonic acid (Figure 2.39) produced by TmuB activity *in vivo* and *in vitro* were assessed for antibacterial activity. Plate bioassay and minimal inhibitory concentration (MIC) were achieved using *Bacillus subtilis* 1064, *E.coli* DH5α and *S. aureus* MRSA as indicator bacteria. Plate bioassay showed that the antibacterial activity of the 4-hydroxy version of pseudomonic acids was reduced against *B. subtilis* 1064 when performed on plates supplemented with 0.5 mM IPTG. However, without IPTG induction the inhibition zones were bigger than the controls (Figure 2.40). This is consistent with the HPLC results which show bigger peaks in the presence of TmuB. The 4-hydroxylated versions of PA-A and PA-B were purified by HPLC and the fractions were collected and dried using a spin vacuum concentrator. Minimum inhibitory concentration was determined and the results revealed a reduction in antibacterial activity of the 4-hydroxy version compounds. Against *B. subtilis*, the MICs for PA-A and PA-B were 0.12 µg ml^{-1} and 8 µg ml^{-1} respectively while the 4OH PA-A and 4OH PA-B MIC were 0.5 µg ml^{-1} and 16 µg ml^{-1} respectively (Table 2.11).

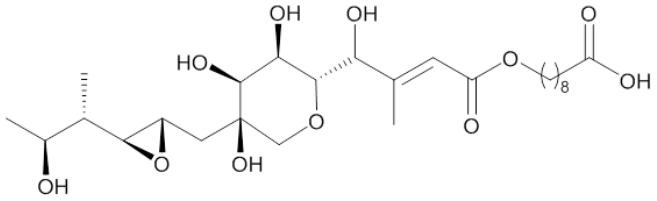
Table 2.11. Minimal inhibitory concentration (MIC) of derivatives produced in this study.

Derivatives	Minimal inhibitory concentration(MIC)		
	B. subtilis	*E.coli* DH5α	*S. aureus* MRSA
Pseudomonic acid A	0.125 µg. ml^{-1}	128 µg.ml^{-1}	0.25 µg. ml^{-1}
Pseudomonic acid B	8 µg. ml^{-1}	256 µg.ml^{-1}	16 µg. ml^{-1}
4OH-pseudomonic acid A	0.5 µg. ml^{-1}	256 µg.ml^{-1}	1µg. ml^{-1}
4OH-pseudomonic acid B	16 µg. ml^{-1}	256 µg.ml^{-1}	16 µg. ml^{-1}
Thiomarinol A	15.6 ng. ml^{-1}	4 µg.ml^{-1}	7.8ng. ml^{-1}
Thiomarinol C	31.25 ng. ml^{-1}	8 µg.ml^{-1}	15.6ng. ml^{-1}

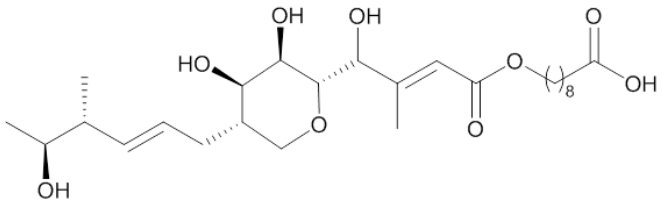
1- Name: 4OH PA-A
Producer: WT *P.flourescens* +*tmuB* gene *in trans*
MS: 516
HPLC RT: 18.2 min
MIC: 0.5 µg/ml



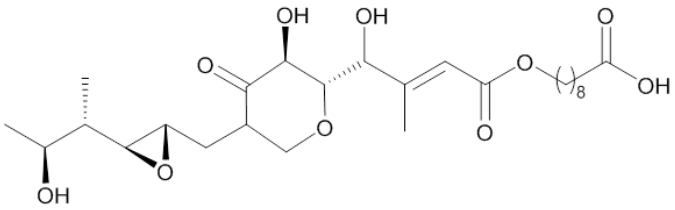
2-Name: 4OH PA-B
Producer: Δ*mupO P.flourescens* and hydroxylated by TmuB protein *in vitro*
MS: 532
HPLC RT: 17.3 min
MIC: 16 µg/ml



3-Name: 4OH PA-C
Producer: Δ*mmpE/OR P.flourescens* +*tmuB* gene *in trans*
MS: 500
HPLC RT: 20.2 min



4-Name: 4OH mupirocin F
Producer: Δ*mupF P.flourescens* +*tmuB* gene *in trans*.
MS: 514
HPLC RT: 19.5 min



5-Name: Thiomarinol C
Producer: deactivated *tmuB* gene in SANK 73390
MS: 624
HPLC RT: 19.70 min
MIC: 31.25 ng/ml

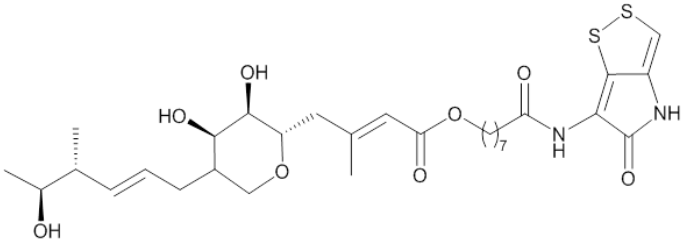


Figure 2.39. Pseudomonic acid and thiomarinol derivatives produced in this study. The MIC of each compound was determined against *Bacillus subtilis* 1064.

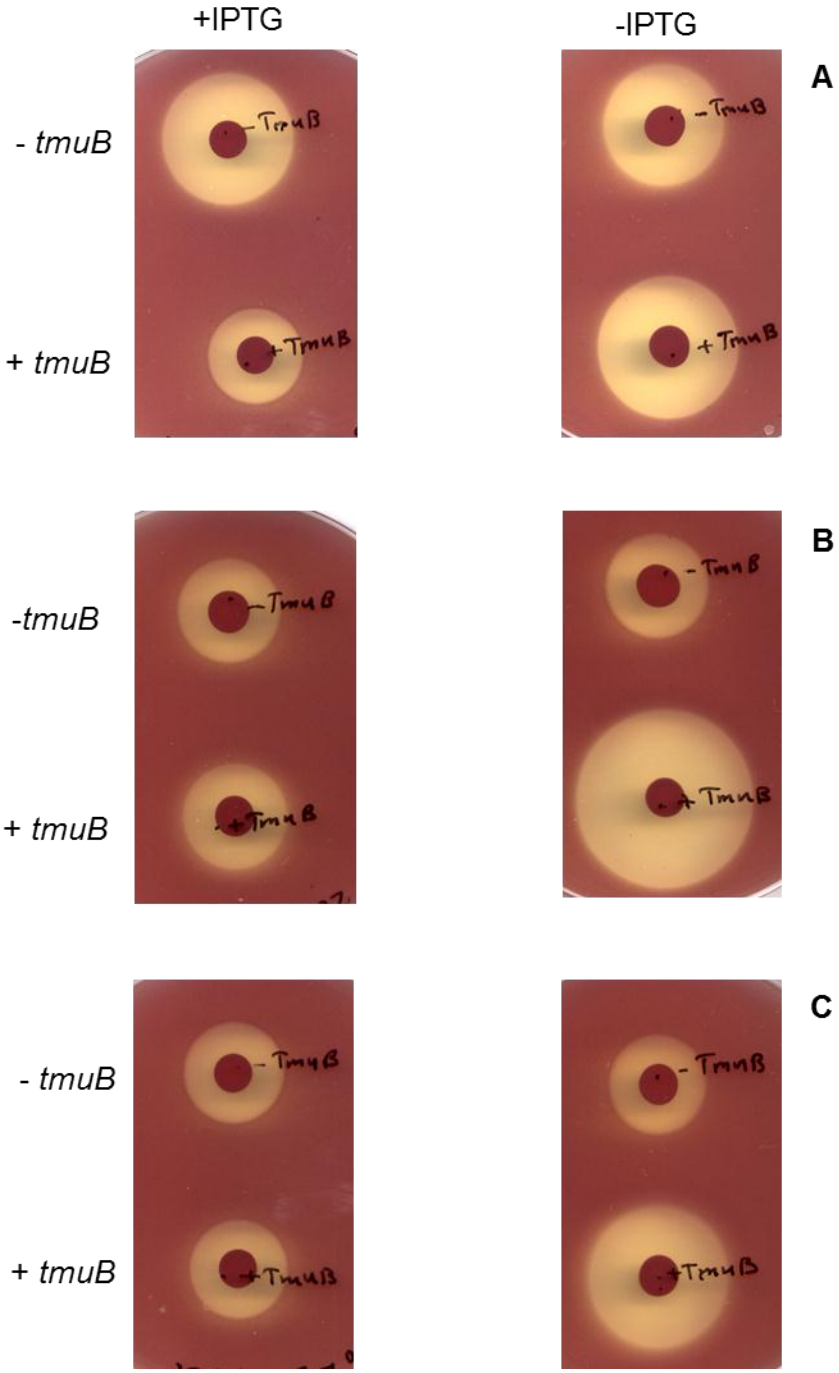


Figure 2.40. Bioassay for the *P. fluorescens* strains expressing TmuB to test the antibacterial activity of the 4-OH version of the products against *B.subtilis* 1064. A: WT *P. fluorescens* ± *tmuB*, B: *ΔmmpE/OR P. fluorescens* ± *tmuB* and C: Δ*mupF P. fluorescens* ± *tmuB*. The left hand plates were supplemented with 0.5 mM IPTG.

2.3.10. Site directed mutagenesis to deactivate TmuB in *Pseudoalteromonas spp* SANK

2.3.10.1. Construction of suicide vector

To create point mutation I109N in TmuB, the pHHM06 plasmid (Table 2.2), used as a DNA template and a pair of primers with *SalI/XbaI* restriction sites were used to amplify by standard PCR as stated in the Materials and Methods section. The size of the PCR product was checked by gel electrophoresis. The correct size band was cut out, the DNA purified digested with appropriate enzymes and then ligated with suicide vector pAKE604 (Figure 2.24). The ligation mixture was transformed to *E.coli* DH5α for propagation. The correct clone was checked by digestion followed by sequencing and the correct clone transformed to *E.coli* S17-1.

The suicide mutagenesis strategy begins by moving the suicide vector (pHHM07) to wild type of *Pseudoalteromonas* SANK 73390. This type of vector is unable to replicate independently in the wild type *Pseudoalteromonas* SANK 73390 unless it integrates into host the chromosome via homologous recombination. This vector contains two selection markers the Kan^r gene, resistant to kanamycin antibiotic, and *SacB* gene which encodes levansucrase hydrolyzing sucrose in the media producing levan which is lethal to the cells. The homologous recombination events, which results in either mutant or wild type bacterium, can be investigate in two steps. The first step is to investigate the conjugation and integration of the vector to the bacterial chromosome by plating on the marine agar containing kanamycin. This allows only the bacteria that obtain the integrated vector to

grow and prevents the growth of the sensitive bacterium (without transformed or integrated vector). However, in our lab, the standard kanamycin concentration (50µg/ml) was used to select the bacteria that have been transformed with target vectors. In the current mutation procedure, a negative control was set up by plating WT SANK on the marine agar supplemented with kanamycin (50µg/ml) and tetracycline (15µg/ml). The incubated plates showed reasonable growth of colonies which means that the WT is resistant at these concentrations. Therefore, higher concentration of kanamycin (100-200µg/ml) was used to prevent the growth of the non-transformed bacteria.

The second recombination event between homologous regions of the DNA results in excision of the vector from the host genome. This will produce resistant cells to sucrose as the *sacB* gene is unable to encode levansucrase any more. This can be achieved by plating the cells on marine agar containing 5% sucrose. The sucrose resistant colonies were checked finally by making patches on marine agar containing kanamycin to confirm the plasmid excision. The homologous recombination and plasmid excision generate two types of cell with equal probability, either the wild type or mutant of the target gene. PCR was used to amplify the target gene (mutant phenotype) using outer primers, and the product was sent for sequencing (Figure 2.24 & 2.41). The bases ATA which encodes isoleucine were changed to AAC to encode asparagine.

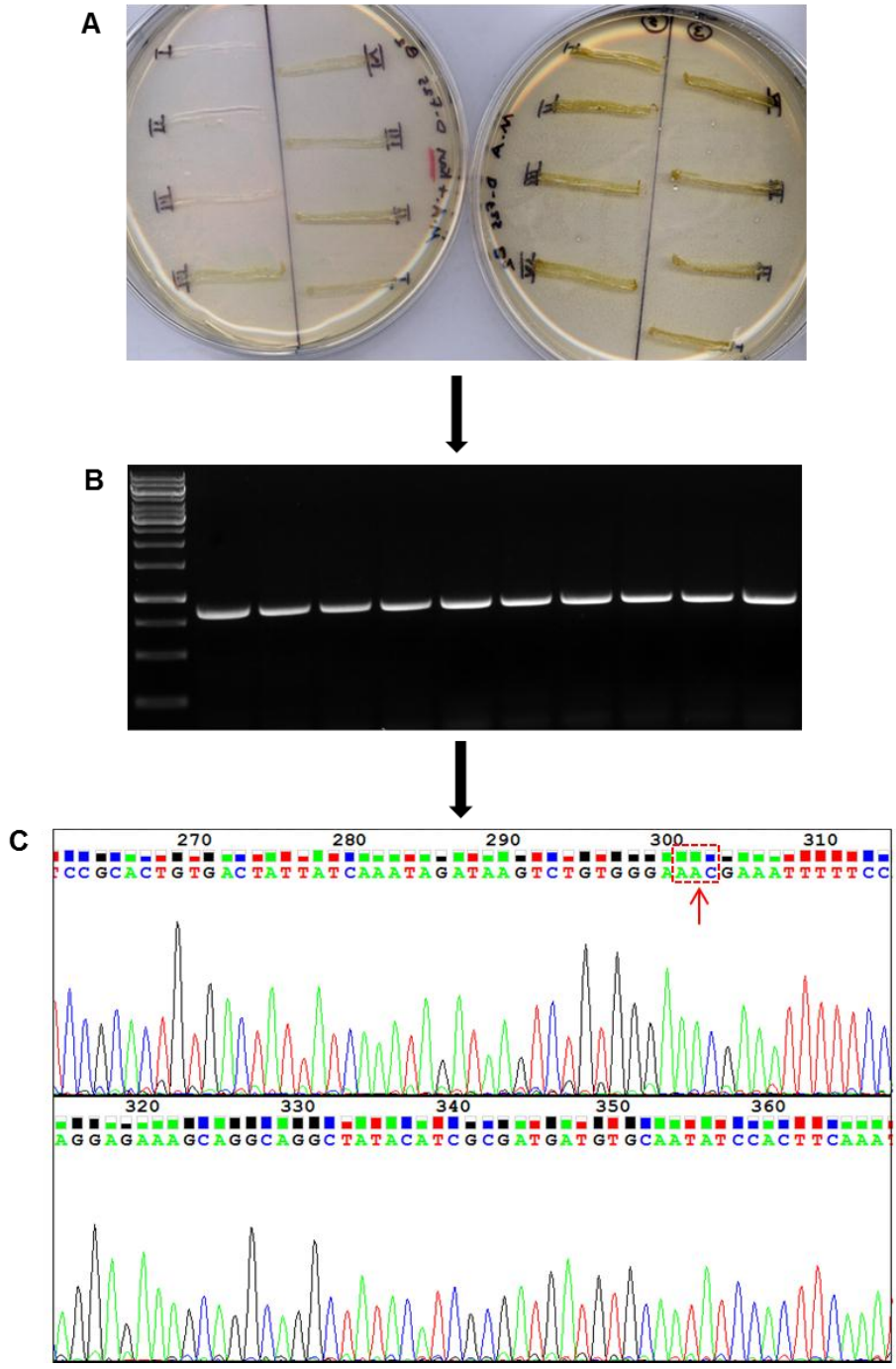


Figure 2.41. Screening the point mutation I109N in SANK. (A) The Kan^{s} and Suc^{r} patches were used as a DNA template for PCR amplification, (B) purified PCR fragments were sent for sequencing, and (C) The sequence result in Chromas showing the bases ATA were changed to AAC (pointed with red arrow)

2.3.10.2. Product characterization from mutant *Pseudoalteromonas* SANK

The culture was set up for the I109N mutant in marine broth and the antibiotic was extracted as described in material and method section. The extract from this mutant was analyzed by HPLC which showed a shifted peak of thiomarinol A from 19.77 min to 21.91 min. This new peak was purified and subjected to LC-MS and NMR which confirm no hydroxylation of thiomarinol at carbon 4 (Figure 2.42)

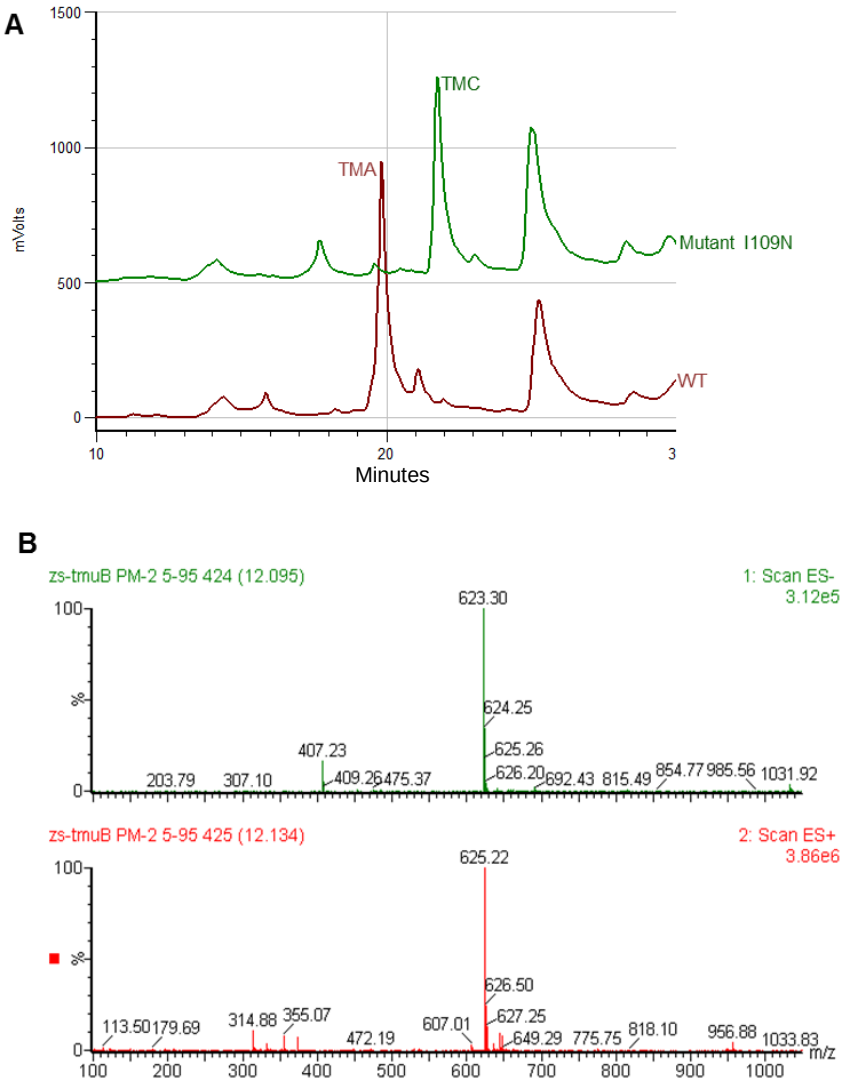


Figure 2.42. Product characterization of the mutant I109N *Pseudoalteromonas* SANK. A. HPLC analysis of the extract from mutant I109N *Pseudoaltromonas* SANK showed the peak of thiomarinol A (WT) (19.77min) shifted to (21.91min). B. Mass spectrophotometer of the new peak MW 624 for mutant which is has a mass 16 Da less than thiomarinol A 640.

2.3.10.3. Growth and product rate

To investigate the role of the *tmuB* gene in the thiomarinol gene cluster, both the growth rate and product profile were monitored. Cultures for the WT and the mutant strain were set up in marine broth. This was used to set up a fresh culture with equal OD_{600} and 200 µl was aliquoted into 96 well microplates and the growth rate of the WT and the mutant strain were monitored overnight at 23 ^{o}C (Figure 2.43A).

To compare the product profile of the mutant and the WT, fresh 25 ml of marine broth was seeded with overnight culture, incubated at 23 ^{o}C /200 rpm and the supernatants were analyzed by HPLC (without opening the cells). The result showed that the thiomarinol C product amount is much less than thiomarinol A (Figure 2.43B).

2.3.11. TmuB activity with thiomarinol C

Enzymatic reactions were set up using thiomarinol C as a substrate to test the activity of TmuB protein and its ability to re-hydroxylate it *in vitro*. The reactions were set up as described previously for PA-A replacing the substrate with thiomarinol C. The data obtained were fitted to Michaelis–Menten model (Figure 2.44). The V_{max} and K_m were calculated accordingly which showed similar activity of TmuB using PA-A as a substrate (Table 2.10).

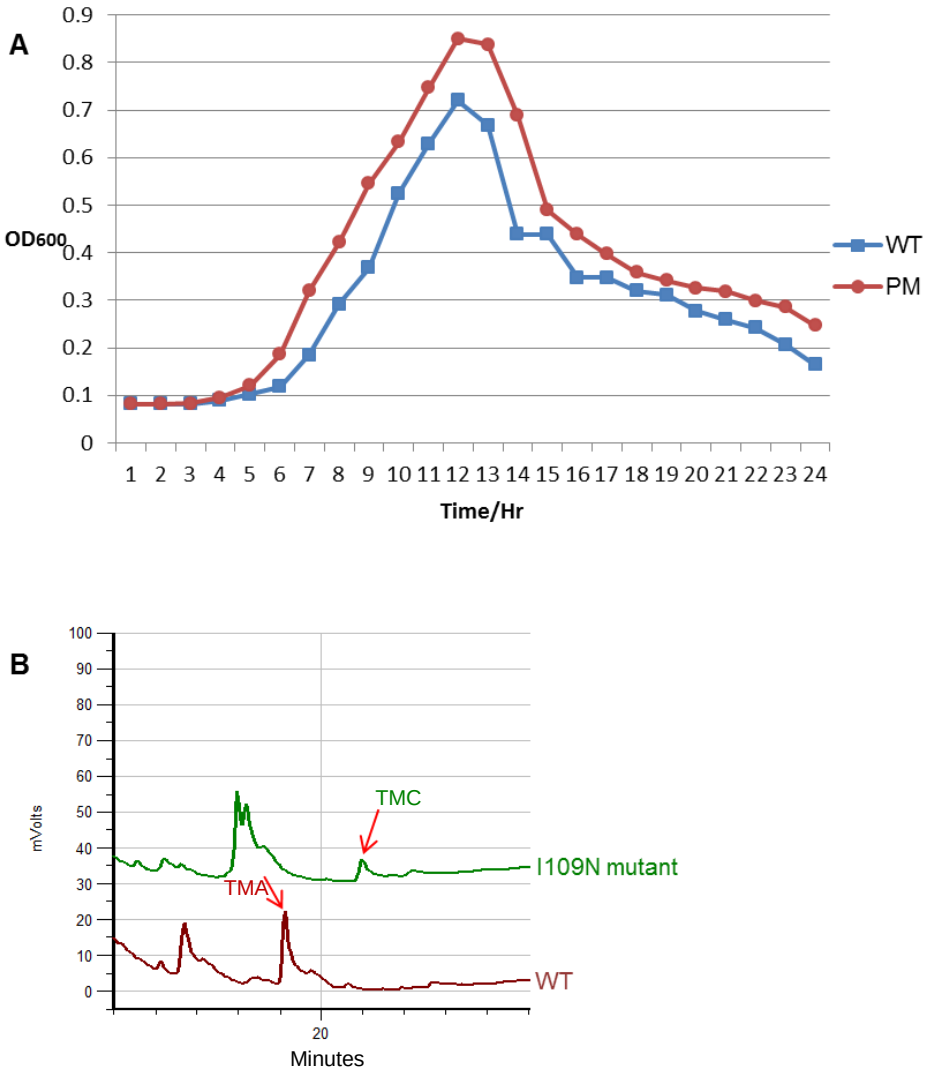


Figure 2.43. Growth rate and product profile of the WT and PM I109N TmuB *Psuedoalteromonas sp.* SANK. The overnight culture seeds were diluted in fresh marine broth to obtain the equal OD_{600} for both strains. A. Growth rate was monitored by aliquoting 200 µl into 96 a well microplate and incubated at 23 ^{o}C for 24 hrs in microplate reader, and B. HPLC analysis of the culture supernatant of WT and mutant I109N SANK 73390. The samples were from the supernatant (not the cell extract) hence the peak size is small for both and the I109N is much smaller.

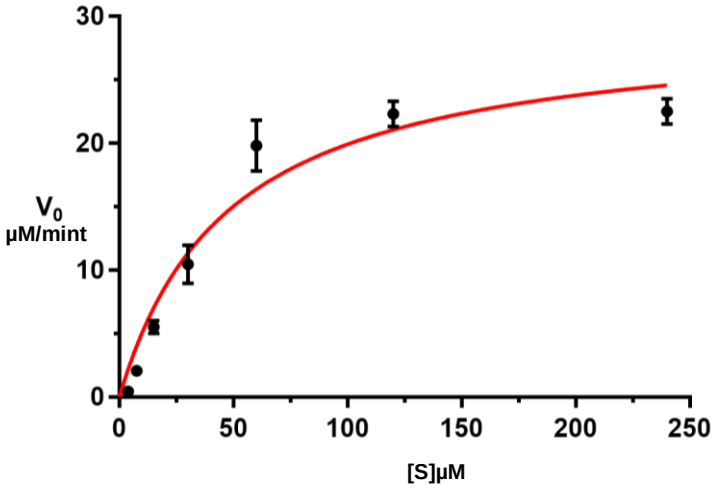


Figure 2.44. Enzyme kinetics for TmuB using thiomarinol C as a substrate. V_{max}=22.5 µM/min and K_m= 29.5 µM. The reaction conditions were similar to the reaction carried out for PA-A in Figure 2.29.

2.3.12. Antibacterial activity of thiomarinol C

The antibacterial activity of the thiomarinol C produced by the mutant strain was tested using a plate bioassay and by minimal inhibitory concentration (MIC). The bacterial strains *Bacillus subtilis*, *Escherichia coli* DH5α and *Staphylococcus aureus* MRSA were used as indicator strains. Thiomarinol A was used as a control to compare the antibacterial activity. The experiments for the bioassay test for each bacterium were carried out separately.

The amount of antibiotics added to the discs for testing against different bacteria were varied from one bacteria to another, for example for *B. subtilis*, 100µg of thiomarinol A and C were added to the discs, while for *E.coli* and *S. aureus* only 60 µg was added

hence the inhibitory zones look bigger. The aim for the test is to compare the potency between thiomarinol A and C against each bacterium not which bacterium is more sensitive as this has been done in previous studies.

Plate bioassay showed that the antibacterial activity of thiomarinol C was reduced compared to thiomarinol A (Figure 2.45). However, from the inhibitory zone, it seems that the difference is not significant in Gram positive bacteria while in Gram negative bacteria, thiomarinol C did not show detectable inhibitory zone around the discs. To exclude any difference being due to the diffusion properties of the tested compounds, MIC test was carried out with similar bacterial strains.

Minimal inhibitory concentration test results are consistent with bioassay plate test regarding the comparison between thiomarinol A and C potency. As can be noted in the Table 2.11, against *B. subtilis* the MIC of thiomarinol A is 15.6 ng. ml^{-1} while for thiomarinol C MIC = 31.25 ng. ml^{-1}. The same figure is applied for *E.coli* and *S. aureus* MRSA.

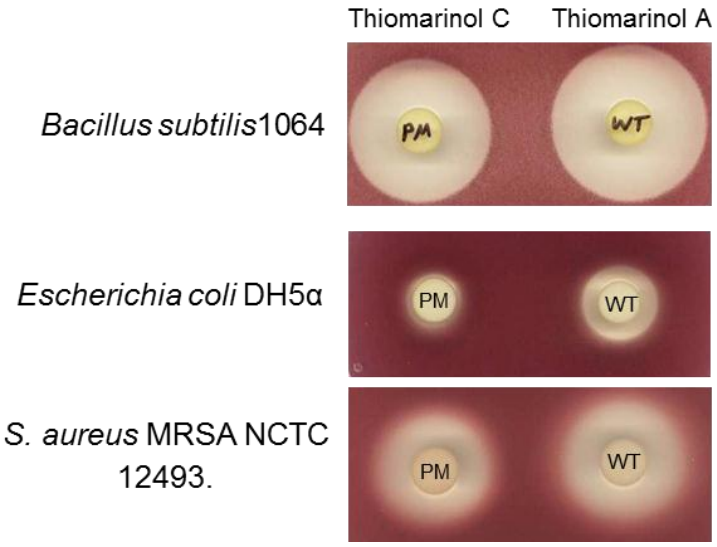


Figure 2.45. Plate bioassay to test the antibacterial activity of thiomarinol A and C from WT and PM I109N SANK respectively. The discs were soaked with purified thiomarinol A and C and tested against *B. subtilis* 1064, *E.coli* DH5α and *S. aureus* MRSA NCTC 12493.

2.4. Discussion

The candidate *tmuB* and *tmlZ* genes which were identified as a putative phytanoyl-CoA dioxygenase and monooxygenase respectively were expressed in the WT and mutant strains of *P. fluorescens* NCIMB 10586. The results showed that only TmuB can hydroxylate thiomarinol analogues (pseudomonic acid derivatives) and display a specificity to the substrates. The results showed that TmuB action appears to classify pseudomonic acid derivatives into two structurally distinct groups. The first group, which can be hydroxylated by TmuB, comprises PA-A, PA-C and Mupirocin F while the second group, which cannot be modified (or modified very slowly), comprises PA-B, Mupirocin C and Mupirocin W suggesting that TmuB does not prefer a substrate with extra groups attached to pyran ring (Figure 1.9).

These results raised a number of questions. Why does TmuB display this kind of specificity? Which residues in the active site are responsible for this specificity? Does the replacement of these residues modify TmuB specificity and allow it to modify more substrates? Therefore, bioinformatic tools were used to predict the protein structure and explore how the protein might interact with the substrates and the possible reaction mechanism.

The modelling and docking process identified the possible interactions between the protein and the substrate and it's fitting into the active site (Figure 2.19). As can be noticed, the residues around the pyran ring are likely to be responsible for TmuB specificity. However, the side of the pocket, which is formed by βII and βVII, is lined by

highly conserved residues, His121, Asp123 and His191 coordinating the Fe ion. Gln 118, Thr154 and Arg202 are located on I/II loop, III/IV loop and βVIII holding the co-substrate. These six residues have been identified by the Evolutionary Trace Annotation (ETA) server as important residues for protein function and are highly conserved. It has been reported that mutation most of these residues inactivates the protein (McDonough *et al.*, 2005).

So to modulate substrate specificity while retaining function the only residues that interact with substrates can be mutated. These residues are located on βI, βIII and βVIII, and form about 50% of the pocket. To explore which residues might be involved in the specificity, a hydroxyl group was added manually to the docked PA-A at carbon 8. This revealed that the residues Arg69, Lys105, Ile109 and Leu141 are located at the entrance to the active site (Figure 2.19C) and potentially close enough to this hydroxyl group to generate steric repulsion. These residues were selected for mutation to smaller and more hydrophilic amino acids to create more space to accommodate a bulkier hydroxylated substrate and interact with it.

The mutagenesis results suggest that it may not be possible to manipulate the catalytic specificity of TmuB or the *in vivo* hydroxylation of PA-B or may hydroxylate it but the 4-hydroxylated product may affect thioesterase activity to release it from the PKS, hence the product cannot be detected in the culture supernatant. However, it is also possible that analogues may be lethal to the producer. To clarify these possibilities, the enzyme activity of the WT and mutants TmuB was analysed *in vitro* followed by product characterization.

As mentioned in the Results section, the enzyme activity revealed that TmuB could hydroxylate PA-B *in vitro* but at a low rate in comparison with PA-A. The ability of TmuB to hydroxylate PA-B *in vitro* as a free molecule excludes the hypothesis that TmuB may work on the substrate while tethered to ACP domain and the 4-hydroxylated product cannot be released by thioesterase from the PKS. In addition, the low antibacterial activity of the 4OH PA-B refutes the possibility that it may be lethal to the producer. Thus, the mutants of TmuB that did not hydroxylate the PA-B *in vivo* are more likely not to hydroxylate it *in vitro*. To confirm this, initially one mutant (I109N) was tested *in vitro* and the result showed no activity which is consistent with the *in vivo* result. The inability of TmuB to hydroxylate PA-B and similar derivatives *in vivo* may be due to the slow reaction rate (K_{cat} = 3.83 s^{-1} and K_{cat}/K_m = 0.045), so the product may be released from the cell before hydroxylation happens. Alternatively, it may be that the very low level of hydroxylated product is just below the HPLC detection limit.

The docking result showed clearly the difference between PA-A and PA-B binding with TmuB model. The docking conformations were compared with the homology crystal structures to estimate the distance between the target site (C4) on the substrate and Fe^{+2} ion in the active site of the protein which may be necessary for the catalytic process. Although the conformation binding energy and the inhibitory constant of the PA-A and PA-B docking are similar (but not identical) the distances of C4 to the active site (Fe^{+2}) are different (4.38 Å and 6.3 Å respectively, see Figure 2.19 and 2.23C). This may explain the *in vitro* enzyme kinetics where the catalytic efficiency (K_{cat}/K_m) for PA-A is much higher than PA-B while there is not such a big difference in K_m values.

Although many mutations were carried out based on modelling and docking results, TmuB specificity was not modified in favour of PA-B binding. This may be due to some uncertainties about this kind of protein which make the docking result unreliable or misleading. First, the loops around the active site (particularly helix2/β3 loop in TmuB) may play a substantial role in substrate binding. The active sites of phytanoyl CoA dioxygenase enzymes are surrounded by flexible elements which undergo conformation changes and induced fit mechanism to ensure the binding of the substrates in the right position (Clifton *et al*, 2006). For example, upon the substrate binding, a finger-like loop N-terminal to DSBH in PHD2 enzyme folds to form a stable loop that encloses the substrate (Chowdhury *et al*, 2009).

In addition, the missing of this loop in some crystal structures which have been resolved without the substrates may show it to be a flexible loop flapping over the active site and may cover the substrate after fitting into the active site. Furthermore, the sequence alignment (Figure 2.12 and 2.17) and evolutionary traces (Figure 2.46) identified the diversity of the residues in this loop among the homologous proteins which may be consistent with the different shape and other prospectives of the substrates.

It is not clear whether or not this loop plays a role in the substrate binding and catalytic activity in TmuB protein. To investigate how the folding of helix2/β3 loop affects the docking process, the 15th TmuB model was used as a receptor to dock PA-A and PA-B. The conformations that were obtained from this docking did not show similarity to the previous conformations which were obtained using the model with the unfolded loop. However, the pyran ring location in the active site remained unchanged as the C4

should stay near the Fe(II) (Figure 2.20). In such case, it is difficult to predict *in silico* the substrate binding with the protein and the role of this loop as the loop is more likely to fold over the active site after substrate binding.

Second, the oligomeric state of the protein may play a role in substrate binding. Both N-terminal and C-terminal regions can be involved in protein dimerization and essential in catalytic activity. It has been reported that C-terminal helical region in FIH protein involved in dimerization which is essential for substrate hydroxylation (Lancaster *et al*, 2004). The gel filtration chromatography showed TmuB protein as a dimer protein which contradicts the bacterial-tow hybrid result. The fusion of T18 or T25 protein into N-terminal or C-terminal may prevent the protein-protein interaction leading to false negative results.

Thus, for further characterization and to figure out the most accurate structure of TmuB and the way in which it interacts with the substrates, the crystal structure is needed. The crystal structure study of Fumitremorgin B endoperoxidase (FtmOx1) from *Aspergillus fumigatus* revealed that the replacing of the single residues (Y224A or Y224F) diverts the FtmOx1 catalysis from endoperoxide formation to hydroxylation (Yan *et al*., 2015).

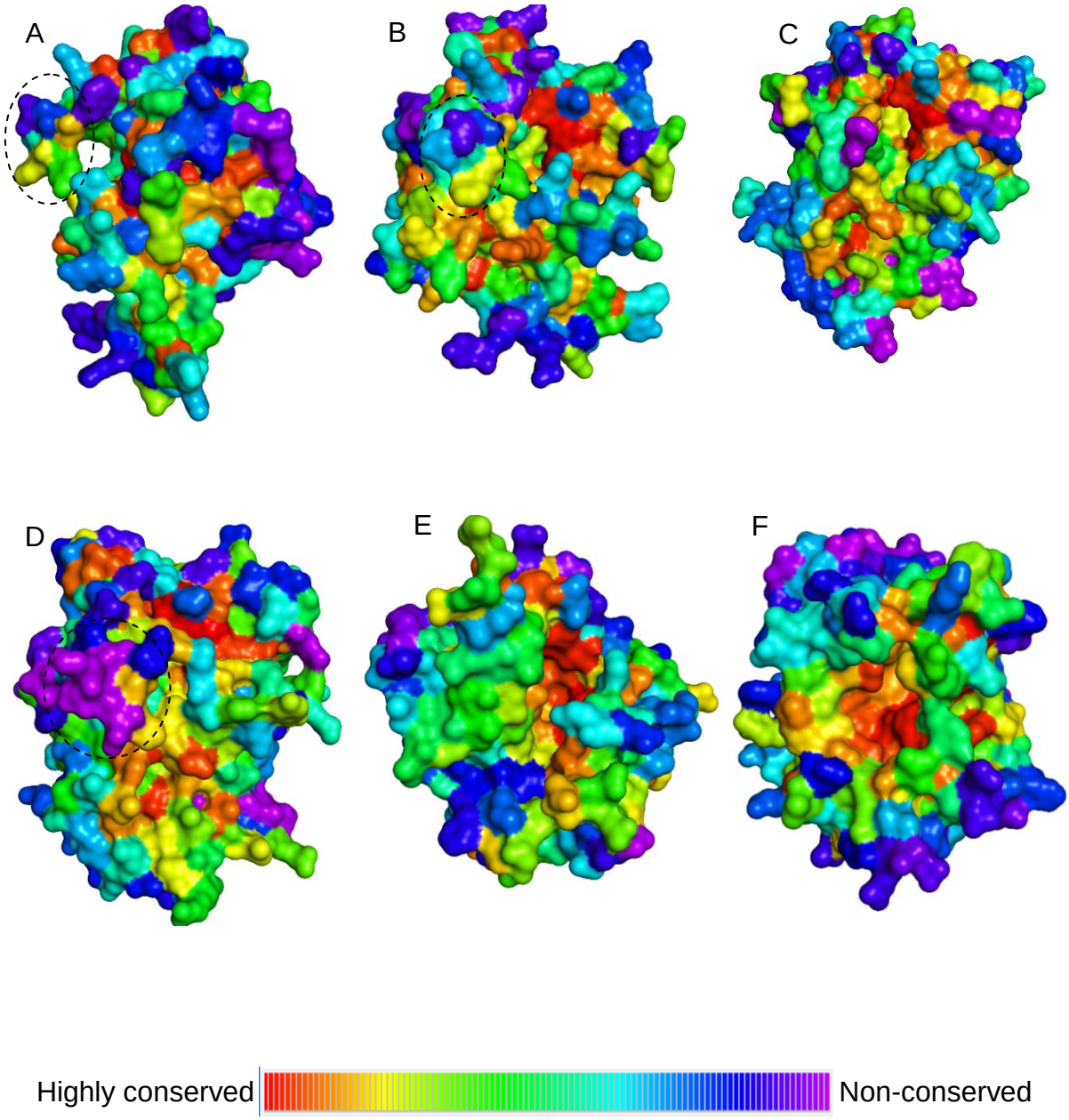


Figure 2.46. The Evolutionary Trace analysis of the TmuB model and the templates showing the conserved residues as red and non-conserved as blue. A and B. Lateral and front view of TmuB model respectively, C. EasH, D. AsqJ, E. FAHX and F. EctD. The residues in the loop (Dashed circle) are vary in these resolved proteins while the six residues in the active site are highly conserved.

The inactivation of TmuB (Point mutation I109N) in thiomarinol producer, *Pseudoalteromonas sp.* SANK 73390, shifts the thiomarinol C, which lacks OH group at C4, as a minor product to be the major product in the mutant strain leaving no doubt about the role of TmuB in 4-hydroxylation.

The growth rate and product profile of the mutant were examined to address the significance of the *tmuB* gene in thiomarinol cluster. The result showed that the mutant strain grows faster than the WT strain and the thiomarinol secretion outside the cell is less than the WT strain (Figure 2.43 A&B). However, when the amount of the products from both strains were detected from the cell extract (not culture supernatant), there was no a significant difference (Figure 2.42 A). In addition, when TmuB was expressed in the heterologous hosts, the amount of the products from the culture supernatant was much higher than those without TmuB. This may demonstrate that the 4-hydroxylation may enhance the ability of the product to pass through the cell membrane, particularly in the Gram-negative bacteria. It may reveal that *tmuB* in the thiomarinol cluster gives a sort of balance between the product formation and the growth rate. However, in terms of biological interaction, it is difficult to clarify the significance of the *tmuB* gene to the surroundings as the producer is a marine bacterium and there is not enough knowledge about the habitat environment and how it interacts with or what is the beneficial effect of the product 4-hydroxylation to the surrounding living organisms.

Rehydroxylation of TMC at C4 by TmuB *in vitro* with a similar catalytic efficiency to PA-A allowed three significant conclusions. First, it appears to confirm the orientation of substrate docking in which the monic acid fits into the pocket while the fatty acid fits in

the external groove (see Figure 2.19). If this were not the case, the active site pocket should not be able to accommodate the bulky moieties like the pyrrothine which is attached to the fatty acid chain. Second, it suggests that TmuB can catalyse 4-hydroxylation either as a final step before pyrrothine joining to marinolic acid or after joining to create thiomarinol C. Third, except around the pyran ring, the modification of the substrates (presence of pyrrothine molecule and shorter fatty acid chain in thiomarinol C) does not affect TmuB catalytic activity. This gives a possibility to use TmuB as a tool to hydroxylate a broad range of substrate analogues. Thus, the modification of the hydroxylase enzyme specificity could enable us to recruit it to modify a wide range of analogues *in vitro*. Furthermore, *trans* expression of such proteins in other analogue biosynthesis pathways may result in significant products.

Based on the crystal structure of the IleRS that was resolved as a complex with mupirocin, Marion and co-workers proposed that the 4-hydroxylation may enable mupirocin to form extra hydrogen bonds with His64 or Asp557, and may improve the antibacterial activity (Marion *et al.,* 2009). However, the results of this study showed a lower antibacterial activity of the hydroxylated PA-A and derivatives.

As an attempt to explain this result, the 4OH PA-A was docked to the crystal structure of the isoleucyl-tRNA synthetase (IleRS) (1FFY PDB) (Silvian *et al.,* 1999). First, PA-A was docked to the IleRS to determine the docking parameters which should give the conformation similar to that in the crystal complex (Figure 2.47).

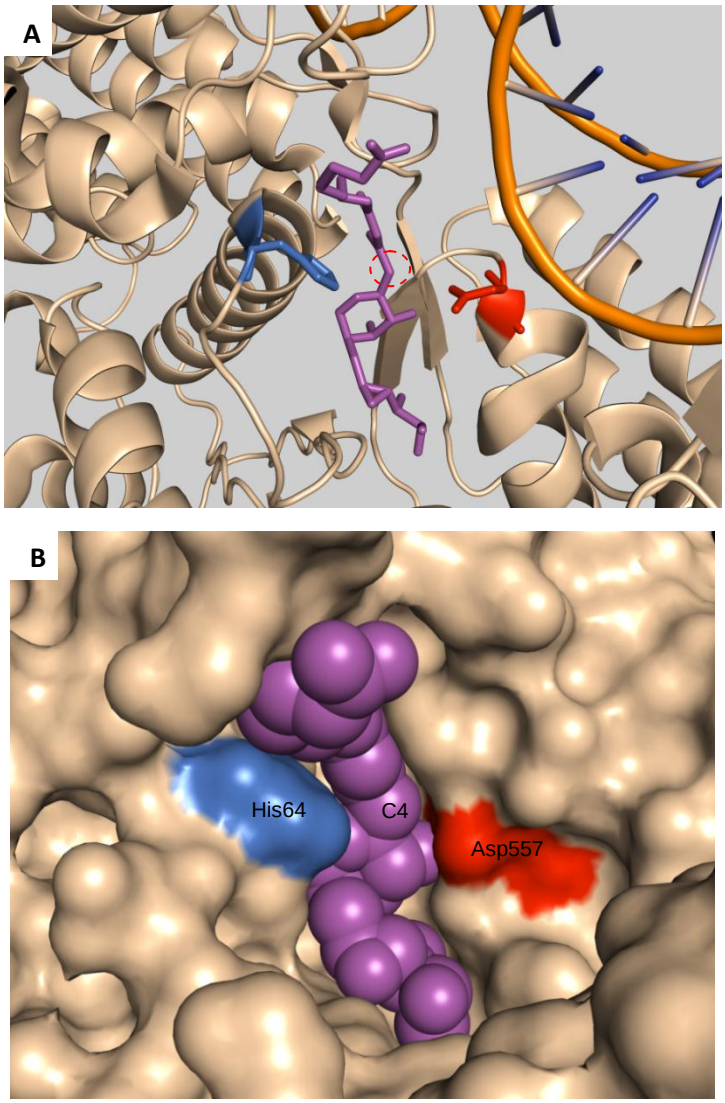


Figure 2.47. Crystal structure of isoleucyl-tRNA synthetase complex with PA-A. The PDB:1FFY was visualized using PyMol, A. Cartoon mode and B. surface mode. PA-A (purple stick), 4C in red circle, Asp 557 (red stick) and His64(Blue stick).

The same parameters were used to dock 4OHPA-A, but the similar conformation could not be obtained. It is not clear whether or not the hydroxyl group on carbon 4 is close

enough to the Asp557 residue to cause a steric repulsion and prevents the PA-A fitting into the target site. Alternatively, this extra hydroxyl group may impart PA-A more hydrophilicity to interact somewhere else on the protein surface or other proteins inside the cell. It was reported that amino acid changes in the active site of IleRS with much bulkier residues resulted in low level resistant against mupirocin. The mutation of valine 588 to phenylalanine in the active site of *S. aureus* IleRS distorts and fills the pocket, and prevents mupirocin binding (Antonio *et al*., 2002; Hurdle *et al*., 2004).

Based on the antibacterial activity test and docking results, it seems “theoretically” that the lack of hydroxylation of thiomarinol at 4C (thiomarinol C) may make it a more potent compound. However, the tests showed that this is not the case as thiomarinol C has lower potency compared with thiomarinol A (Table 2.10) and (Figure 2.38). A study in our lab proposed that thiomarinol may target some other cell components rather than IleRS and display cytotoxicity in eukaryotic cells (Ahmed, 2010 unpublished). The result also showed that thiomarinol C is still more potent than PA-A. This raises the question whether or not the 4-dehydroxylation of thiomarinol A reduces the cytotoxicity in eukaryotic cells which could be confirmed by further analysis. Hydroxylation of the analogues at C4 and lack of 4-hydroxylation of thiomarinol A has changed the antibacterial activities and may give them novel biological activities.

According to the results presented here, it can be concluded that TmuB works as a final tailoring step to direct 4-hydroxylation of thiomarinol and displays specificity for substrates. However, this specificity is mainly restricted to the pyran ring as the catalytic efficiency (K_{cat}/K_m) for PA-A and TMC are similar. This means that TmuB may be able to

convert a wide range of substrate analogous either *in vitr*o or *in vivo* by expressing it *in trans* in other analogue biosynthesis pathways. Furthermore, the mutagenesis trials in the active site of enzymes belonging to the nonheme-iron(II)/2-oxoglutarate-dependent dioxygenase superfamily may not change just the specificity but also diverts the catalysis function as was reported recently with fumitremorgin B endoperoxidase from *Aspergillus fumigatus* (Yan *et al*., 2015). This approach may participate in producing new diverse derivatives with significant biological activities. The hydroxylation and dehydroxylation of the natural products is a vital step in terms of biological activities of the products and metabolism profile of the producers. The antibacterial activities of the products formed as a result of TmuB expression or inactivation have changed but might possess news interesting biological activities which can be investigated by further studies.

CHAPTER THREE

3. CHARACTERIZATION OF NON-RIBOSOMAL PEPTIDE SYNTHETASES IN THE THIOMARINOL CLUSTER

3.1. Introduction

Non-ribosomal peptide synthetases (NRPSs) are responsible for most peptide bonds in secondary metabolism. As described in Chapter one, NRPSs are multi modular proteins and normally each module is responsible for adding one amino acid to the product backbone. Some NRPSs do not follow this general rule in NRP biosynthesis, iterative NRPSs re-use the catalytic domain repeatedly and add more than one building block. (Mootz *et al.*, 2002; Hur *et al.*, 2012).

Congocidine is a pyrrole-amide antibiotic produced by an iterative NRPS in *Streptomyces ambofaciens* (Juguet *et al.*, 2009). However, the iterative NRPS types are common in many fungal species. Siderophores, high affinity iron chelators, and Enniatins, cyclohexadepsipeptides with various biological activities, are among the compounds that are produced by iterative NRPS in fungi (Bushley,Ripoll and Turgeon, 2008; Sy-Cordero, Pearce , Oberlies, 2012).

Thiomarinol produced by *Pseudoalteromonas* bacteria contains pyrrothine consisting of two cysteine molecules which are proposed to be produced by the NRPS in the thiomarinol gene cluster (Figure 3.1). The intriguing feature of the *holA* gene, which encodes the NRPS responsible for pyrrothine biosynthesis, is that encodes only one Adenylation (A) domain, Condensation (C) domain and Peptidyl carrier protein (PCP) that is a single module, while pyrrothine should require two modules since it is made by

joining two amino acids (Fukuda *et al*., 2011). This leads to the suggestions that either HolA is a dimeric protein or it works as an iterative NRPS in a similar manner to fungal NRPS. Further investigations of this unusual NRPS will provide more data that should help us to understand the assembly of the products and provide a better understanding of biosynthesis by related NRPs.

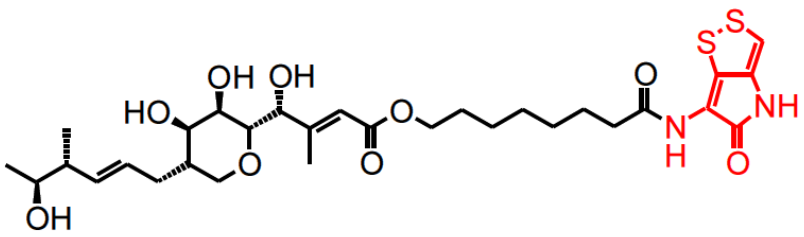


Figure 3.1. The thiomarinol structure. The red part is the pyrrothine moiety produced by the NRPS

The aim of this chapter, therefore, is to characterize the Non-Ribosomal Peptide Synthetase in the thiomarinol gene cluster and to understand the assembly of the pyrrothine molecule. The *holA* gene possesses 3369 nucleotides encoding 1122 amino acid residues (125KDa). The residues of this protein are arranged as a tri-domain module consisting of three putative domains, Condensation, Adenylation and Peptidyl Carrier Protein (PCP). Both *in vitro* and *in vivo* experiments were used to investigate the oligomeric state of HolA. The native and mutant purified proteins were subjected to a biochemical assay to check the *in trans* activity and the amount of the substrate that can be activated which may give a clue about the oligomeric state of the protein or the iterative behavior of the domain.

3.2. Materials and Methods

3.2.1. *holA* gene amplification and plasmid construction

A set of primers (Table 3.1) were used to amplify *holA* and to construct plasmids using the standard molecular biology techniques as described in Chapter 2. Forward and reverse primers were designed with *NdeI* and *XohI* restriction sites respectively. This enabled the gene to be cloned into the expression vector pET28a. The gene was amplified using a high fidelity Q5 polymerase from New England Biolabs. The following program was set up to run PCR:

Denaturation step: 98 ^{o}C for 2 minutes
Denaturation step: 98 ^{o}C for 30 sec
Annealing step: 63 ^{o}C for 30 sec
Extension step: 72 ^{o}C for 2.5 min
} 2 cycles

Denaturation step: 98 ^{o}C for 30 sec
Annealing step: 65 ^{o}C for 30 sec
Extension step: 72 ^{o}C for 2.5 min
} 33 cycles

Extension step: 72 ^{o}C for 7 minutes 1 cycle

The purified PCR product was cloned into pGEM-T Easy vector for DNA sequencing. In addition to the forward and reverse universal primers, two primers were designed to

cover the middle sequence of the gene during the sequencing process (Table 3.1). The correct sequence of *holA* was inserted into pET28a and the protein expression and purification were achieved as described in Chapter 2. For the bacterial strains and plasmids used in this study see Table 2.1and 2.2 in Chapter 2.

Table 3.1. Primers designed and used in this study.

Name	Function	Sequence	Rest. site	Tm*
HolAF	*holA* amplification to express by pET28a	5'GTACATATGAACATGGATGCATTTAAGC 3'	*NdeI*	63
HolAR		5'GAACTCGAGTCACACATCCTGACGTTCCAC 3'	*XohI*	64
HolAF2	*holA* sequencing	5' CGAATACTGGGAGAAGAAGC 3'		54
HolAR2		5' CGGTCACAGGATGGTCAA 3'		54
HAB2HF	*holA* amplification for Bacterial Two-Hybrid vectors	5'GTATCTAGAGATGAACATGGATGCATTTAAGC 3'	*Xbal*	59
HAB2HR		5'GAAGGTACCTCACACATCCTGACGTTCCAC 3'	*Kpnl*	58.5
HAADF	HolA adenlylation domain amplification for Bacterial Two-Hybrid vectors	5'GTATCTAGAGAGTCGTCAAGGTAGTGTGGTCG 3'	*Xbal*	61
HAADR		5'GAAGGTACCTGGGTAATTCATTG TGGCTG 3'	*Kpnl*	57
HACYR	Reverse primer for C domain in B2H	5'GAAGGTACCTTGACGACTTAAGA GCTGTG 3'	*Kpnl*	63
PCPF	Forward primer for PCP domain in B2H	5'GTATCTAGAGACAGAGGCGGCA TTATTGG 3'	*Kpnl*	66
CPMr	Reverse primer to create PM H202A in C domain	5'CGAGGTACACACTAGCCATATCGACAA3'		70
APMf	Forward primer to create PM K515T in A domain	5'CGAGCAATGGAACAATCAGCCACAAT3'		68

PCPPMf	Forward primer to create PM S29A in PCP domain	5'TGGAGGTGATGCTCTACATGCGGTG3'		70
TmlNF	*tmlN* gene amplification to express by pJH10	5'AGCGAATTCATGAATATTGAGCGTCAATCTATAC3'	*EcoRI*	54
TmlNR		CTGTCTAGATATTTAAGCAAATGAGTGAGGG	*Xbal*	54.5

*The Tm was calculated by online software NetPrimer from primer Biosoft. The PCR annealing temperature is lower by 2 ^{o}C.

3.2.2. *In vitro* protein-protein interactions

3.2.2.1. Cross linking with Glutaraldehyde

Chemical cross linking was used to test protein dimerization *in vitro*. 50 µl of different concentrations (mg/ml) of HolA protein sample in 20mM HEPES buffer, 500mM NaCl, pH 7.5 was incubated with different concentrations of Glutaraldehyde (2%, 1%, 0.1% to 0.5% and 0.01% to 0.05%) at 37 ^{o}C for five minutes. The reaction was terminated by adding 10 µl of 1M Tris pH 8.0. The protein was precipitated by adding 4 volumes of 1:1 methanol and acetone and incubated at -20 ^{o}C for 1 hour. The protein pellet was collected by centrifugation at 14000 xg for 20 minutes at 4 C^{o} and the supernatant was removed carefully by pipetting. The pellet was resuspended in 20 µl of SDS loading buffer and boiled for five minutes and analysed by SDS-PAGE.

3.2.2.2. Analytical Ultra Centrifugation (AUC)

The protein samples were purified by affinity chromatography, dialysed overnight in buffer (50mM NaH_2PO_4, 200mM NaCl, pH 7.5) to eliminate imidazole and analysed by SDS-PAGE before Analytical Ultra Centrifugation (AUC). 400 µl of three samples (0.65 mg/ml) were loaded into the cells and 420 µl of the dialysed buffer was used as a blank. The test was run at 40.000 rpm / 4 ^{o}C and the absorbance scanned at 280 nm every 6 minutes using Proteome Lab Xli machine and the An-50 Ti rotor. The sedimentation velocity data were analysed using Sedfit version 15.01b software (http://www.analyticalultracentrifugation.com) and 180 scans were analysed for each sample.

3.2.3. *In vivo* protein-protein interactions

The Bacterial Adenylate Cyclase Two-Hybrid System Kit was used to test protein-protein interactions *in vivo*. To amplify the *holA* gene, forward and reverse primers were designed with *Xbal* and *Kpnl* restriction sites (Table 3.1) respectively. This enabled the gene to be cloned in the Two-Hybrid System vectors. In addition, primers were designed to clone each of Condensation (C), Adenylation (A), and Peptidyl Carrier Protein (PCP) separately to the system vectors (Table 3.2). The details of the procedure were described in Chapter2.

3.2.4. Domain inactivation by point mutation

The modified overlap extension technique, which was described in Chapter 2, section 2.2.15, was used to create point mutations in HolA domains. The target residues in the active site were determined depending on the analogues protein sequence alignment and activity. New primers were designed (Table 3.1) to introduce the point mutations using the recombinant *holA*-pET28a as a DNA template. Standard molecular biology techniques were used to clone the mutant *holA* gene to pET28a and the protein expression and purification were achieved as described previously.

Table 3.2. Plasmids constructed in this study. For the plasmids features see Table 2.2.

Plasmid's name	Features	Usage	Source
pHHM20	*holA* cloned to pGEM-T easy as A-tail insert	Sequencing of the *holA* gene	This study
pHHM21	*holA* cloned to pET28a as *NdeI/XohI* insert	HolA expression in BL21	This study
pHHM22	*holA* cloned to pKT25 as *XbaI/KpnI* insert	To test protein-protein interactions	This study
pHHM23	*holA* cloned to pUT18 as *XbaI/KpnI* insert	To test protein-protein interactions	This study
pHHM24	*holA* cloned to pKNT25 as *XbaI/KpnI* insert	To test protein-protein interactions	This study
pHHM25	*holA* cloned to pUT18C as *XbaI/KpnI* insert	To test protein-protein interactions	This study
pHHM26	C domain cloned to pKT25 as *XbaI/KpnI* insert	To test protein-protein interactions	This study
pHHM27	C domain cloned to pUT18 as *XbaI/KpnI* insert	To test protein-protein interactions	This study
pHHM28	A and PCP domains cloned to pKT25 as *XbaI/KpnI* insert	To test protein-protein interactions	This study

pHHM29	A and PCP domains cloned to pUT18 as *XbaI/KpnI* insert	To test protein-protein interactions	This study
pHHM30	*holA* with PM H202A in C domain cloned to pET28a	Mutant HolA expression in BL21	This study
pHHM31	*holA* with PM K1003T in A domain cloned to pET28a	Mutant HolA expression in BL21	This study
pHHM32	*holA* with PM S1072A in PCP domain cloned to pET28a	Mutant HolA expression in BL21	This study
pHHM33	*holA* with double PM in A and PCP domains cloned to pET28a	Mutant HolA expression in BL21	This study
pHHM34	*TmlN* gene cloned to pJH10 as *EcoRI/XbaI* insert	Works as phosphopantetheine transferase	This study

3.2.5. *In vivo* phosphopantetheinylation of HolA

The phosphopantetheine moiety was added *in vivo* to the serine residues of PCP domain by co-expression of *tmlN* gene with *holA* gene. This enzyme converts the apo-PCP to holo-PCP by transferring the 4- phosphopantetheine group from CoenzymeA. The *tmlN* gene was cloned into pJH10 and co-transformed with constructed pHHM21 into BL21 *E.coli*. The phosphopantetheinylation of HolA was checked by mass spectrometry in the genomic lab, School of Bioscience, University of Birmingham. Three types of HolA protein samples were sent for mass spectrometry to check for post translational addition of the phosphopantetheine arm by PPTase. These include: WT HolA protein expressed in the absence of TmlN (apo form), WT HolA protein co-

expressed with TmlN (holo form) and the mutant HolA protein in which the serine residue in the active site was replaced by alanine in the PCP domain.

3.2.6. ATPase assay

Malachite green ATPase assay was used to test protein activity depending on the pyrophosphate (PPi) that is released by the enzymatic activity. The adenylation domain activates the amino acid through a reaction with ATP to form an aminoacyl-AMP intermediate. This reaction releases inorganic pyrophosphate (PPi) from ATP. In the assay, inorganic pyrophosphatase (PPase) (provided by New England Biolabs) was added to catalyse the hydrolysis of PPi to two phosphate ions which can be detected by Malachite green.

To prepare malachite green solution, 340mg of Malachite green crystal was dissolved in 75ml of distilled water and 10.5g of Ammonium molybdate was dissolved in 250ml of 4N HCl. These two were mixed and the volume was made up to 1000ml, stirred on ice for 1 hour and filtered through Whatman filter paper. To prepare the assay solution, 50 ml of malachite solution was mixed with 250 µl of 20% Triton X-100. The initial reaction was set up by adding 40 µl of different concentrations of protein mg/ml (0.09, 0.187, 0.375, 0.75 and 1.5), 2 µl of 100mM L-cysteine, 2 µl of (50mM ATP, 100mM magnesium acetate) and 10 µl of PPase (100units/ml). The mixture was incubated at room temperature for 1, 5, 10 and 20 minutes, 800 µl of assay solution was added, incubated at room temperature for 1 minute followed by addition of 100µl of 34% of citric acid, and

incubated at room temperature for 40 minutes. Different amount of HolA and PPase protein were used to optimumize the reaction. A calibration curve was made using different concentrations of KH_2PO_4 (0.0 µM, 50 µM, 100 µM, 150 µM and 200 µM) to measure the amount of pyrophosphate released by protein activity. The optical density (OD) was measured at 640nm using a spectrophotometer. A set of controls were set up to determine the base line of the optical density.

3.3. Results

3.3.1. HolA-His tag expression and purification

To express HolA protein as an N-terminal His-tag fusion, a pET28-*holA* (pHHM21) vector was constructed (Figure 3.2) as described in Chapter 2. *Escherichia coli* BL21 DE3 was used to over express the target protein at different temperatures and IPTG concentrations to check the initial expression. In small scale expression, all the temperatures and IPTG concentrations induced sufficient production of HolA protein when analysed by SDS-PAGE. However, it was found that the optimum expression of the soluble protein can be achieved at 18 ^{o}C with 1mM IPTG (Figure 3.3). Therefore, this optimum growth condition was used for large scale expression.

After expression, the protein was purified by affinity chromatography based on 6xHis-tag protein as described in Chapter 2. It was found that the majority of tagged protein was eluted at 100mM, 200mM and 250mM imidazole. To obtain as much pure protein as possible, the last three fractions of the first purification were re-purified again (Figure 3.4)

Despite the repeat purification by affinity chromatography, the SDS-PAGE shows secondary bands. Therefore, to obtain homogenous pure protein and eradicate imidazole, gel filtration was performed which separates the molecules according to their size. The absorbance at 280nm showed three peaks after running 120 ml of the volume: from 120 to 140, from 140 to 160 and from 160 to 180. According to the calibration curve, the estimated sizes are 770-680KD, 250 KDa and 125-128KDa respectively

(Figure 3.5). These variations were either due to the different oligomeric states of the protein or different shape and conformation because of the protein folding state. Therefore, for simplicity these three peaks were named according to the peak order as "1st Peak, 2nd Peak, and 3rd Peak" and used in the rest of this study. The fractions for different peaks were collected separately and analysed by SDS-PAGE.

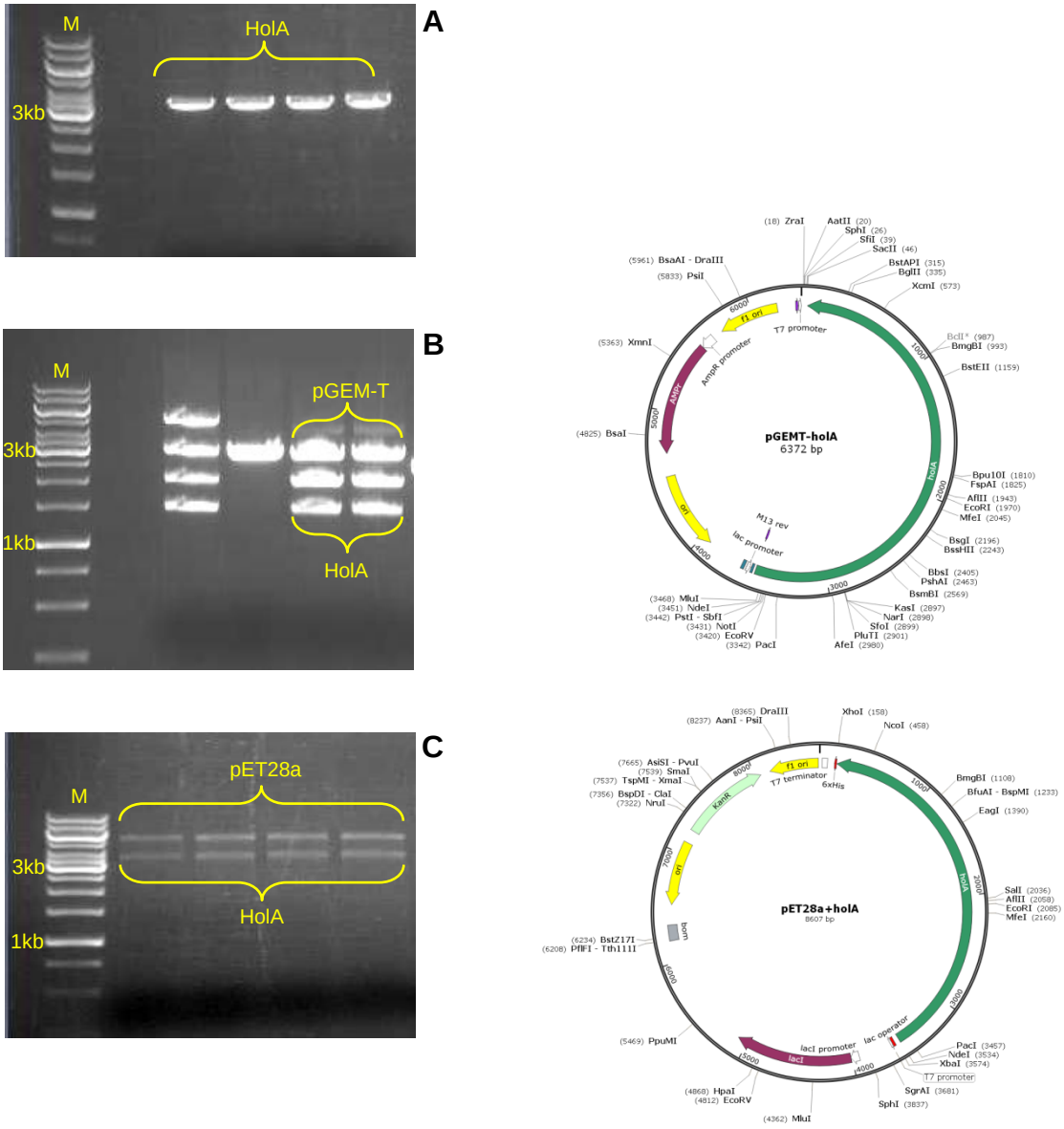


Figure 3.2. Construction of pET28a-*holA*. (A) PCR amplification of *holA* gene 3.4kb, (B) recombinant pGEMT+*holA* gene digested with *EcoRI* restriction enzyme which cuts the *holA* gene into 2kb & 1.4kb bands and 3kb for the vector. (C) Recombinant pET28a vector digested by *NdeI* & *XhoI* restriction enzymes and released *holA* from the vector 5.3kb

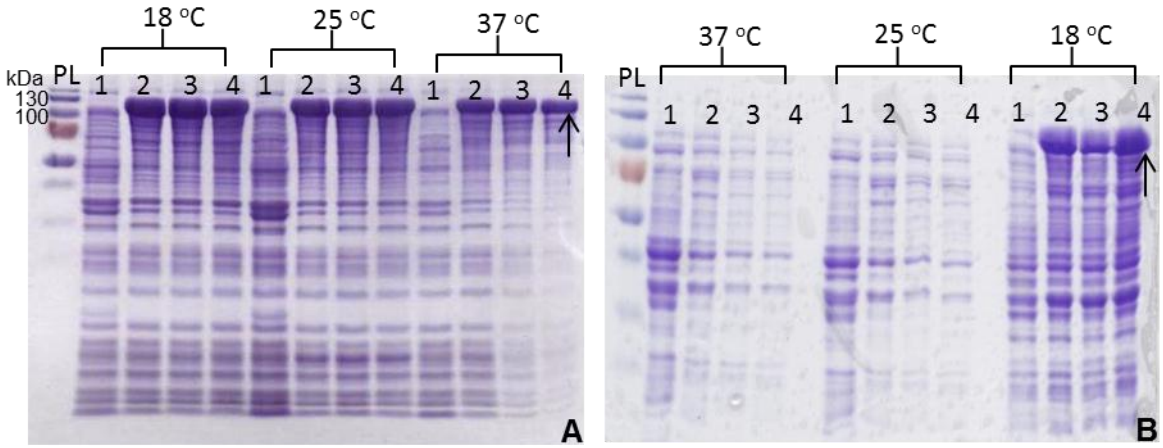


Figure 3.3. SDS-PAGE analysis of small scale expression of HolA protein. PL: Prestained Protein Marker 170kDa. 1,2,3 & 4: IPTG concentrations 0 Mm, 0.1 mM, 0.5 mM & 1 mM respectively. A: Insoluble protein, B: Soluble protein. The soluble HolA protein (125 kDa) was seen at 18 ^{o}C.

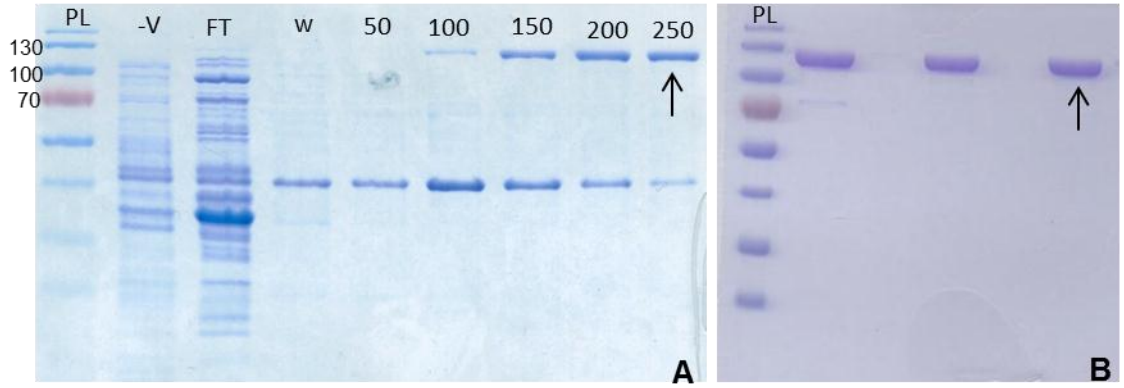
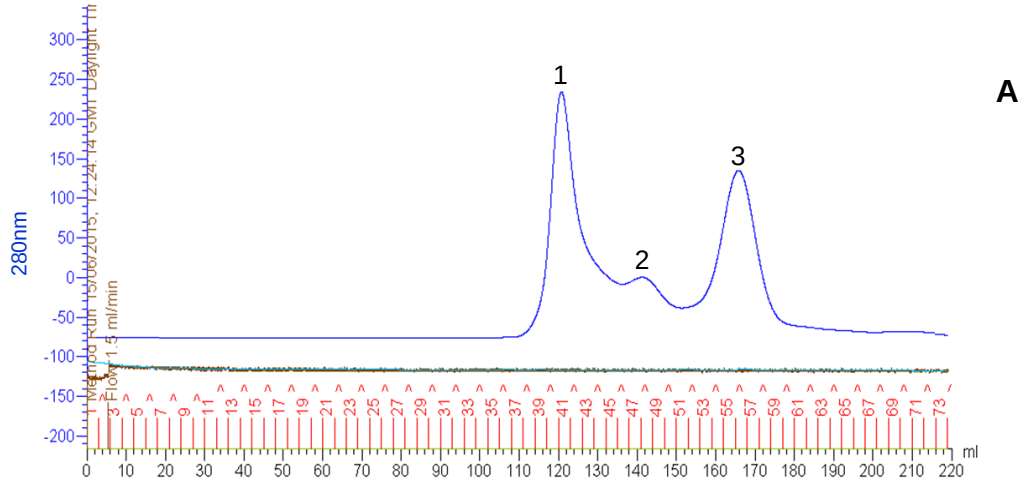
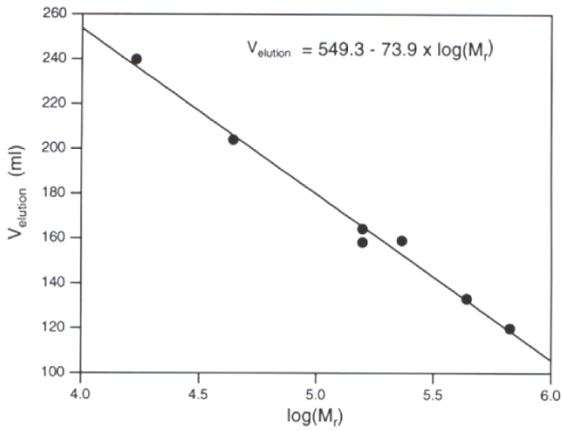


Figure 3.4. SDS-PAGE analysis of HolA protein purified by affinity purification. PL: Prestained Protein Marker 170kDa. –V: Negative control, FT: Flow through, W: Wash buffer, 50 to 250: Imidazole concentrations. A: First purification, B: Repeated purification by gel filtration of the last three fractions. The 125kDa of HolA protein is indicated by the arrow



B

Protein/compound	M_r (10^3)	log(M_r)	V_{el} (ml)
Thyroglobulin	670	5.826	119.8
Ferritin	440	5.643	132.9
Catalase	232	5.365	158.7
gamma-Globulin	158	5.199	158.0
Aldolase	158	5.199	164.0
Ovalbumin	44	4.643	204.0
Myoglobin	17	4.230	239.9
Blue Dextran	>2000		112.5
Vitamin B-12	1.35		293.1



C

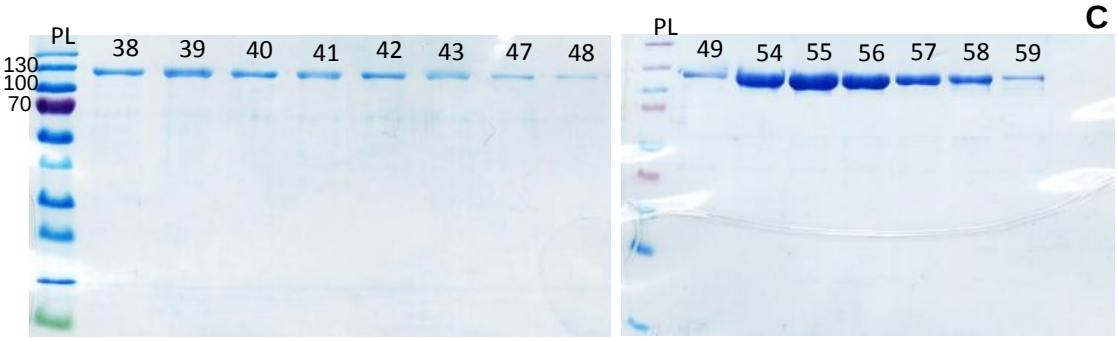


Figure 3.5. Gel filtration for HolA protein purification. A: The protein was separated into three peaks: 1:1st Peak eluted with void volume at 120 ml, 2: 2nd Peak at 140 ml and 3: 3rd Peak at 170. B: Calibration curve to estimate protein size, and C: SDS-PAGE analysis of the fractions.

3.3.2. *In vitro* protein-protein interactions

3.3.2.1. Cross linking with glutaraldehyde

HolA protein sample was tested before and after gel filtration *in vitro* using glutaraldehyde. The result for the sample purified only by affinity purification revealed that the protein molecules appeared to interact with each other at low concentration of glutaraldehyde. The SDS-PAGE showed a large molecule of around 650-700 kDa at 0.1-0.5% while at higher concentrations (1-2%), the protein made aggregrates and stuck in the wells and could not run even through stacking gel (Figure 3.6A). It looks like the slowly moving band disappears at the lower protein concentration which might be non-specific cross-linking – and this would not be surprising at 1 mg/ml (Figure 3.6B).

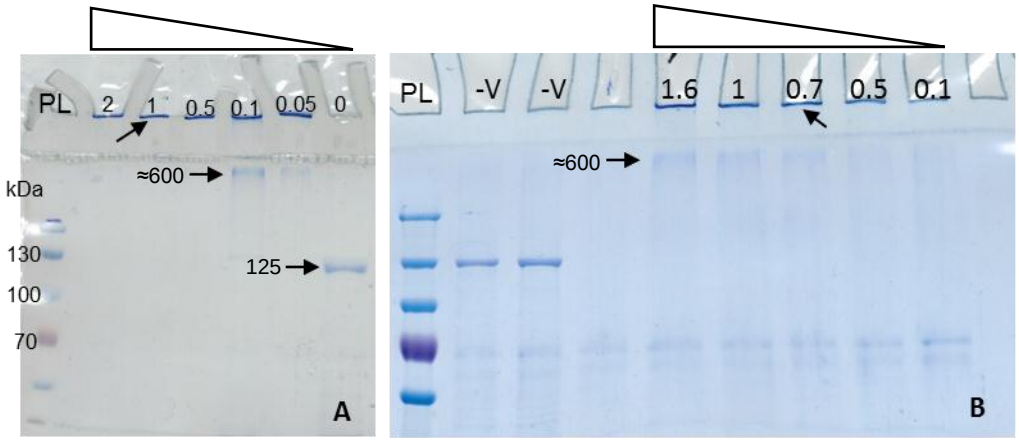


Figure 3.6. Cross linking of HolA with glutaraldehyde before gel filtration. A: Different concentrations of glutaraldehyde (%) with (1mg/ml) protein, and B: different concentrations of protein (mg/ml) with 0.5% of glutaraldehyde. PL: protein ladder, -V: negative control. The pointed arrows show the estimated MW in kDa.

After gel filtration, the test was repeated using the 1st and the 3rd Peaks of protein (0.5mg/ml) (see Figure 3.5) treated with a range of glutaraldehyde (concentrations from 0.01% to 1%). For 3rd Peak protein, the SDS-PAGE showed that the protein was able to cross linked as a dimer, particularly from 0.03% to 0.5% of glutaraldehyde. At higher concentration of the cross linking reagent (0.1%, 0.5% and 1%) larger molecules were formed which stuck in the well and consumed all the protein (Figure 3.7A). For the sample from the 1st Peak of gel filtration, the reaction produced complexes protein and couldn't run through the gel and the result was consistent with sample reaction before gel filtration (Figure 3.7B).

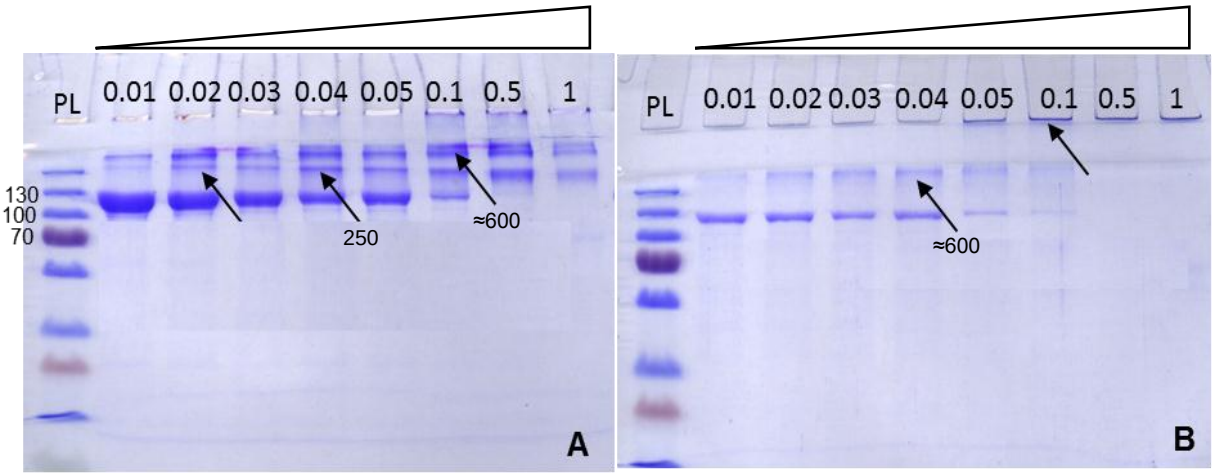


Figure 3.7. Cross linking of HolA with glutaraldehyde after gel filtration. A: 3rd Peak protein (0.5mg/ml) with different concentrations of glutaraldehyde % shown. The proposed dimer (250kDa) protein bands indicated by arrows. B: The 1st Peak protein with different concentration of glutaraldehyde % shown. With the glutaraldehyde increasing, the protein mobility is reduced and sticks in the wells. PL: protein ladder. The pointed arrows show the estimated MW in kDa.

3.3.2.2. Analytical Ultra Centrifugation (AUC)

The data were analysed using sedfit software at friction ratio 2.39 and rmsd 0.008. The result showed HolA protein in a monomer state as a single peak. As illustrated in Figure 3.8, the data fitting result showed the majority of HolA as a monomer protein with MW 125KDa. However, there were some minor peaks which were estimated to be in the higher oligomeric state or due to the sample impurity. The result appears to contradict with the gel filtration chromatograph but this could be due to different conformations of the protein interacting differently with the gel filtration matrix.

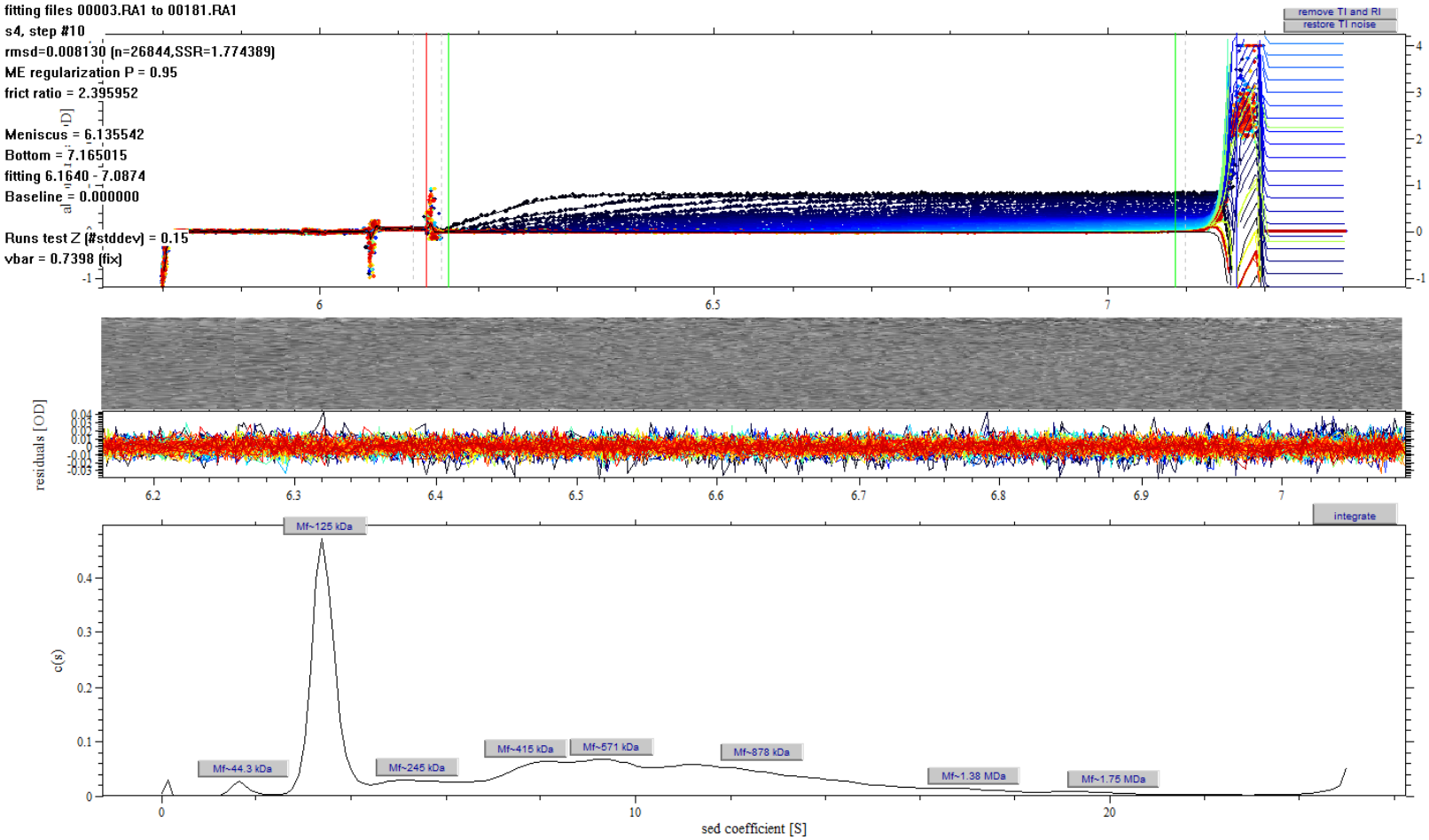


Figure 3.8. The analytical ultracentrifugation result for HolA protein using Sedfit software. The fitting parameters are shown on the upper left side. The analysed data in the graph (the lower part) shown HolA protein as monomer (125kDa).

3.3.3. *In vivo* protein-protein interactions

The Bacterial-Two hybrid system was used to test protein-protein interactions *in vivo* as described in Chapter2. HolA protein was fused with the C-terminus of T25 using pKT25 plasmid, and with N-terminus and C-terminus of T18 using pUT18 and pUT18C plasmids respectively. The two compatible recombinant vectors were co-transformed into competent *E.coli* BTH101(-*cya*) and plated out on MacConkey plus 1% maltose and on LB agar plus IPTG and X-gal. After three days of incubation at 30 ^{o}C, the result showed red colonies for positive controls on MacConkey and blue colonies on X-gal LB plates while the test protein was negative (pale colonies) and more similar to the negative controls as shown in Figure 3.9. Thus the bacterial two-hybrid experiments suggest that this protein cannot interact *in vivo* or at least cannot reconstitute a functional adenylate cyclase. The large size of the protein, which consists of three domains, may prevent the physical contact between the two domains (T18 & T25) of adenylate cyclase (CyaA) resulting in false negative. To exclude this possibility, individual domains of C, A, and PCP were fused separately with T25 & T18 as stated previously. The result showed that no interaction between the single domains can be detected by this method *in vivo* either.

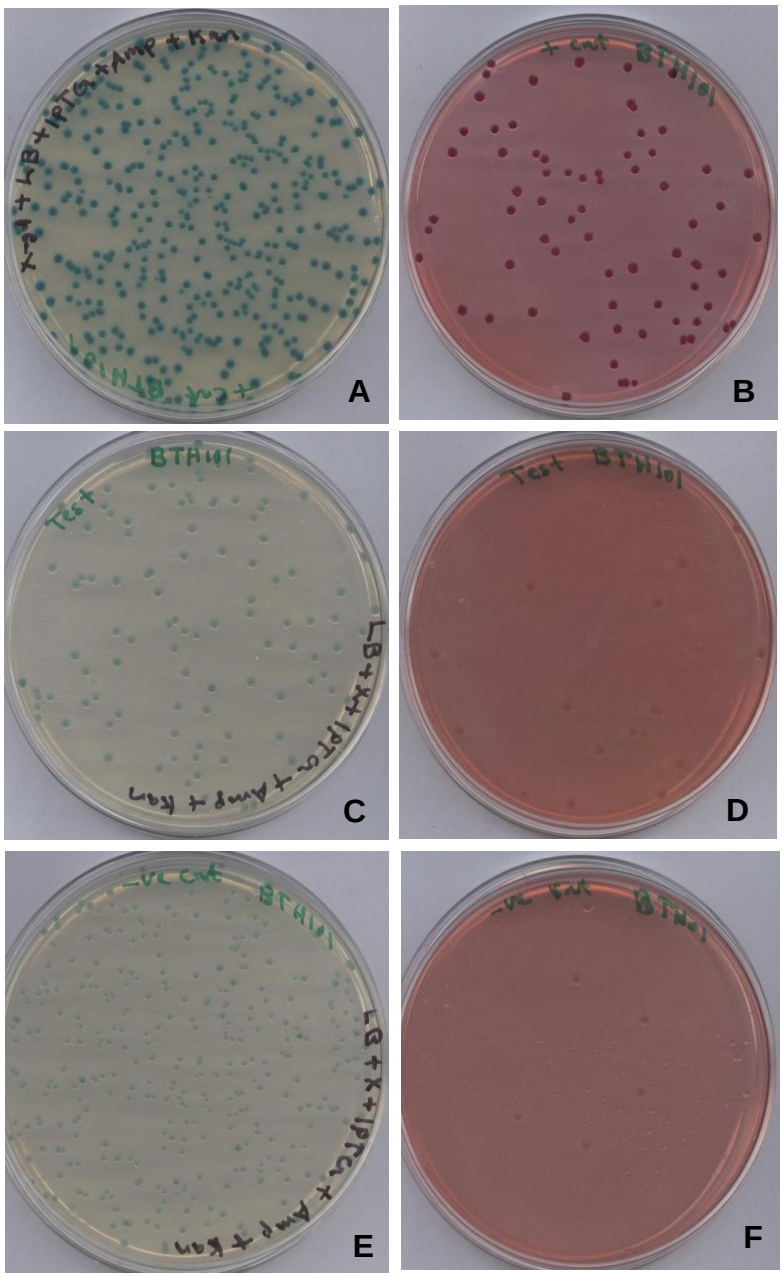


Figure 3.9. Protein-protein interactions *in vivo* using Bacterial Two-Hybrid system. BTH101 colonies on MacConky (Right) and X-gal LB agar (left). A&B: positive controls, C&D: Hybrid proteins (pHHM22 & pHHM25) and E&F: Negative controls.

3.3.4. Domain inactivation

An additional way to test the function of the protein was to determine the consequences of inactivating the individual domains both *in vivo* and *in vitro*. The predicted amino acid sequence of HolA protein was aligned against the previously resolved and characterized protein structures using COBALT multiple alignment tool from NCBI to identify the residues in the active site which are essential for catalytic activity. As illustrated in Figure 3.10, His202 residue in the condensation domain active site is well conserved which may play a key role in the catalytic activity. This was replaced with alanine by site-directed mutation by converting the CAT to GCT. With respect to the adenylation domain, the residue Lys515 was replaced with threonine by converting AAA to ACA. In the PCP domain, Ser29 is the only serine which is highly conserved so this was replaced with alanine (TCT to GCT) to prevent the phosphopantetheinylation of this domain. Two mutant HolA proteins were produced, HolA A with a single mutation K515T in A domain and HolA C-T with double mutation H202 and S29A in the C and PCP (Thiolation) domains respectively.

Figure 3.10. Multiple sequence alignment of HolA domains with homologous proteins. The target residues H202 in C domain, Lys515 in A domain and Ser29 in PCP domain were replaced by alanine, threonine and alanine respectively.

3.3.5. *In vivo* phosphopantetheinylation of HolA

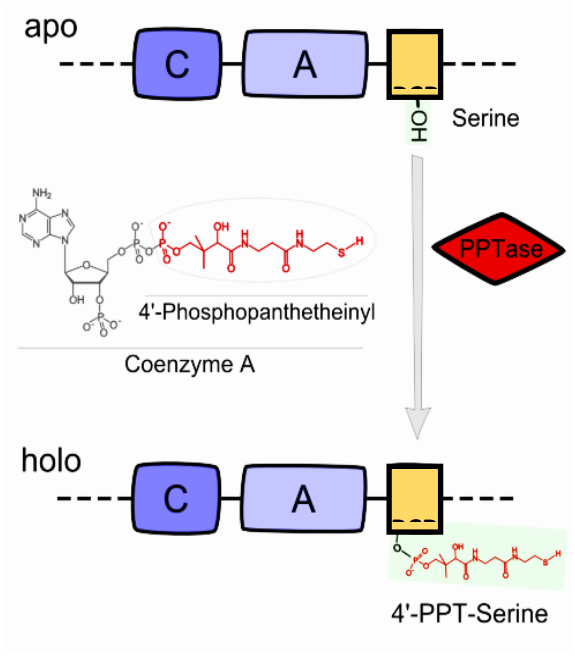
The *tmlN* gene is proposed to work as phosphopantetheine transferase (PPTase) in the thiomarinol cluster. The *tmlN* gene possesses 780bp encoding 259 residues with high identity sequence to phosphopantetheine transferases (PPTase) (55%, 2e-88) including MupN from the mupirocin cluster which has been shown to be essential for mupirocin biosynthesis and to phosphopantetheinylate to a number of the *mup* ACPs (Shields *et al*., 2010). Thus, HolA was co-expressed in the presence of TmlN in an attempt to obtain the holo form of the protein. As a result, the phosphopantetheine transferase should add a 4- phosphopantetheine moiety to the highly conserved serine residue in the peptidyl carrier protein (PCP) domain as illustrated in Figure 3.11A. Prior to MS analysis, the

protein samples were subjected to trypsin digestion (Figure 3.11B). Therefore, the molecular weight of the digested peptide of PCP domain was calculated to see the modification of the three samples (see Section 3.2.4). The theoretical MW of each peptide was calculated by ExPASy (Compute pI/Mw tool) and compared with the result. The native apo PCP domain is 3255.63 Da and the MW of 4'-phosphopantetheine is 358.348282 Da. This means that the MW of the digested peptide of the holo PCP domain should be 3595.963 Da. A similar signal (3595.43 Da) was detected by the MS analysis for holo form protein only. The same signal was missed for the other two samples, the apo form and mutated PCP (Table 3.3).

Table 3.3. The MS result of HolA phosphopantetheinylation showing the MW of digested PCP domain.

HolA protein samples	TmlN	TMW* of PCP	AMW* of PCP
WT HolA protein	-	3255.63 Da	3254.66 Da
WT HolA protein	+	3595.963 Da	3595.43 Da
Mutant HolA protein S29A	+	3239.64 Da	3238.86 Da

*TMW: Theoretical molecular weight.
*AMW: Analytical molecular weight.



A

B

PCP domain seq: AALLEIWIKQLQHSSLDVTSGFFDIGGDSLHAVALIGTIRERFG

Trypsin digestion

MW = 3255.63 Da. QLQHSSLDVTSGFFDIGGDSLHAVALIGTIR

PPTase

MW = 3595.963 Da. QLQHSSLDVTSGFFDIGGDSLHAVALIGTIR

Figure 3.11. Phosphopantetheinylation of HolA by TmlN. A: The process catalysed by PPTase, converting the apo protein to holo form. B: The trypsin digestion process prior to MS analysis, the green highlighted residues are the site for trypsin digestion, the yellow highlighted serine is the active site of the PCP domain and the location for the 4- phosphopantetheine attachment.

3.3.6. ATPase assay

The initial assay was set up for 20 minutes to measure the endpoint of the enzymatic activity by reading the optical density for different protein concentrations. The Holo and Apo forms of 2nd and 3rd Peaks of HolA protein, which are shown in Figure 3.5, were tested to check the amount of the PPi released by each form. Four controls were set up to measure the natural hydrolysis of ATP, auto dissociation of PPi and to show that the PPi release by HolA protein depends on the presence of substrate and ATP (Figure 3.12A). The results showed that the ATP auto hydrolysis and PPi dissociation (Control 1 and 2 respectively) increased slightly over time. The amount of the phosphate released by the protein activity was calculated by subtracting the controls OD_{640nm}, then comparing the OD with the calibration curve of standard concentration of KH_2PO_4 (Figure 3.12B).

The test shows that in the presence of cysteine and ATP each 10µmole of Apo form of the 3rd Peak protein can release about 20µmole of phosphate. This indicates that the Adenylation domain is functional and can activate the substrate releasing one PPi molecule (Figure 3.13). The assay was repeated for the 2nd Peak and the result showed that there was not a significant difference between the two Peaks of HolA protein. As showing in these Figures, the optical density declined with the lowering of protein concentration which means that the PPi release depends on protein concentration.

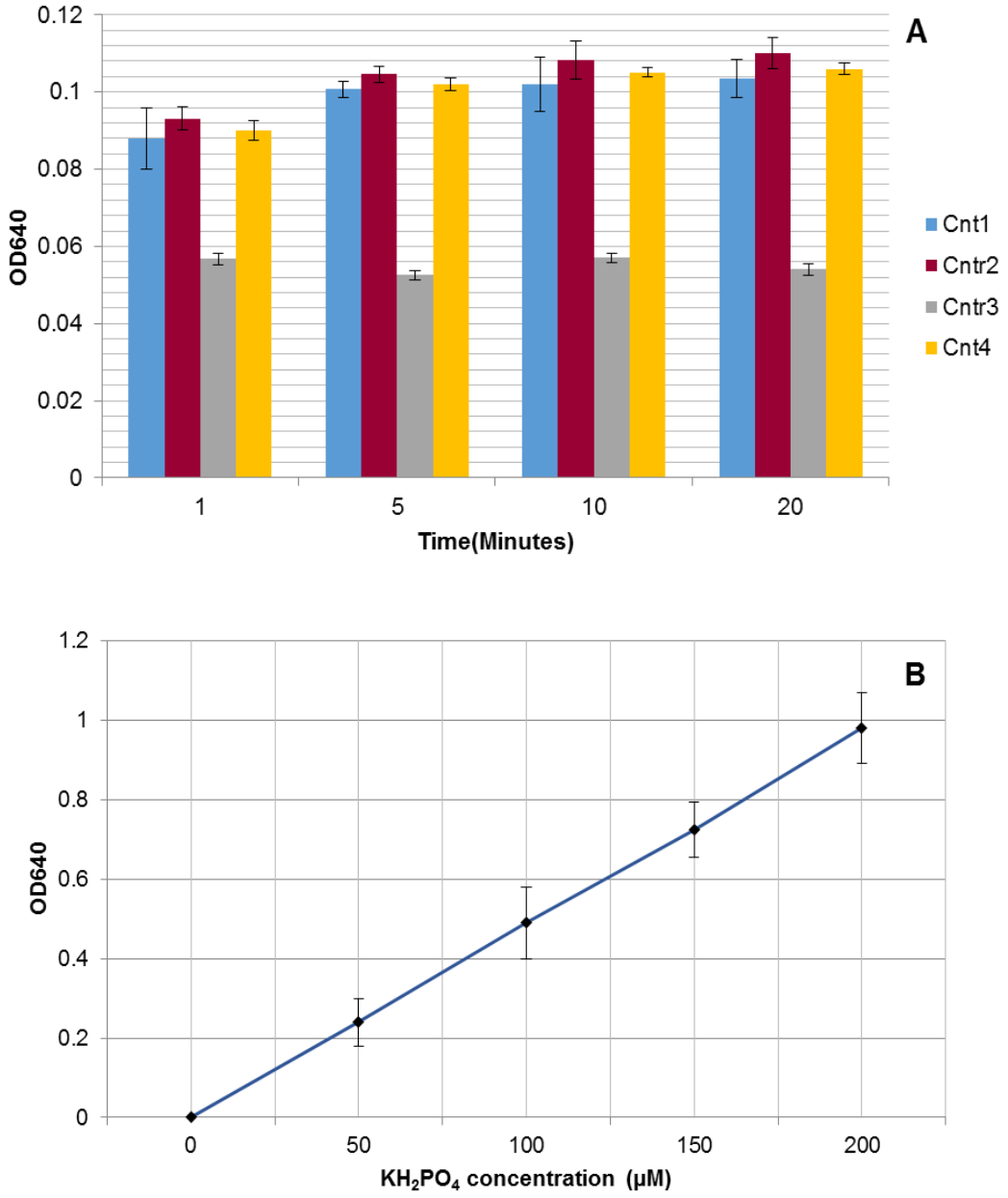
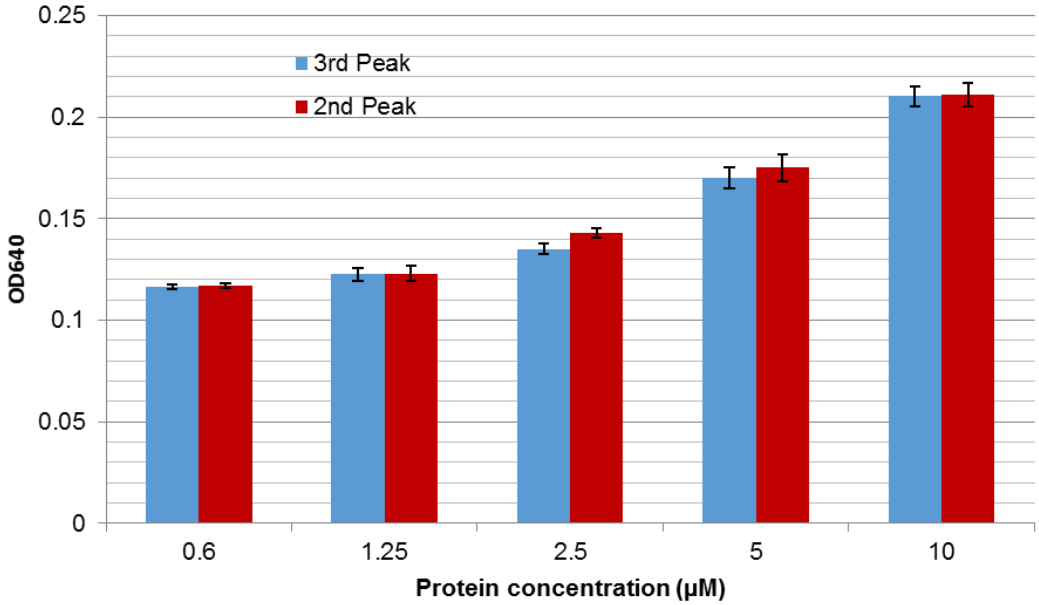


Figure 3.12. Controls and calibration curve for malachite green ATPase. A. Four controls at different time in which one or more components were omitted, Cnt1(-HolA & PPase), Cnt2(-PPase), Cnt3(-ATP) and Cnt4(-Cys); B. Calibration curve.



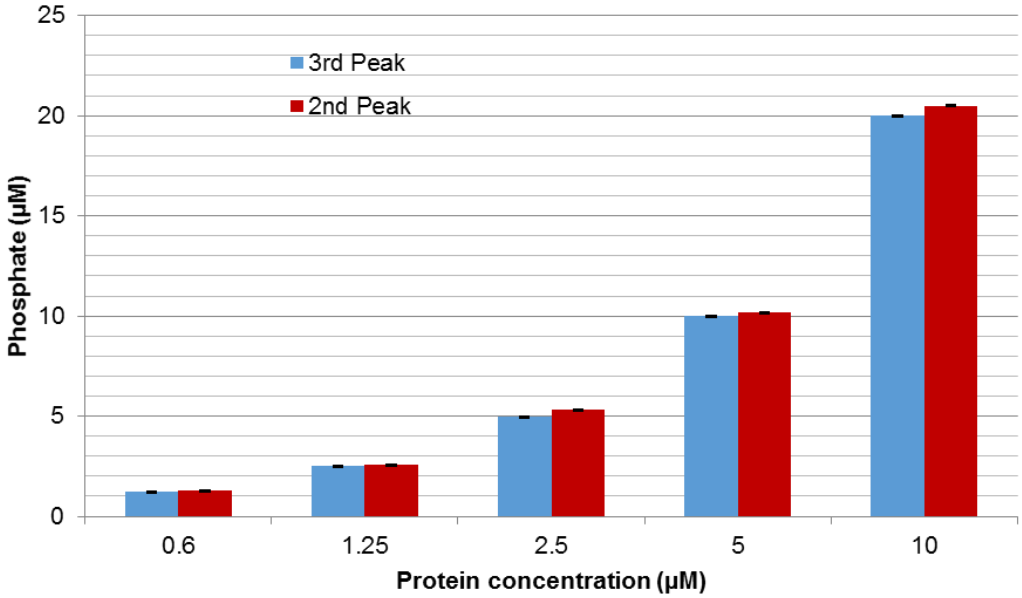


Figure 3.13. Malachite green ATPase test for Apo HolA protein. A. The total OD_{640} for each concentration of HolA protein, and B. The amount of phosphate (μM) released by each concentration was calculated by subtracting the total OD_{640} from the control OD_{640} and comparing the net OD_{640} with the calibration curve of standard concentration of KH_2PO_4 in Figure 3.12B.

The test was repeated using the Holo form protein which interestingly showed about two-fold higher OD reading. According to the calibration curve, 40µmole of phosphate can be released by 10 µmole of Holo form protein (i.e. two PPi molecules were released) as shown in Figure 3.14.

The assay was repeated using the Holo form protein concentration (10µM) at different time course (1, 5 and 10 minutes) to check the auto dissociation of PPi and to confirm that the OD reading is due to the HolA activity, not the PPase activity. The HolA activity is nearly flat after 10-20 minutes which means that HolA reached the maximum activity and the measured OD is due to its activity, not the PPase (Figure 3.15).

The activity of the mutant protein, HolA A and HolA C-T (see section 3.3.4) was tested by this assay. The HolA A showed reduced OD by 75%, while the OD of the HolA C-T was similar to the Apo form protein (Figure 3.16). To test whether or not the adenylation domain of HolA C-T can complement the HolA A and recover the OD as in the Holo form, both with equal amount were mixed and tested by ATPase assay. The OD could not be recovered to the Holo form level which may explain that the proteins are not able to complement each other *in trans*.

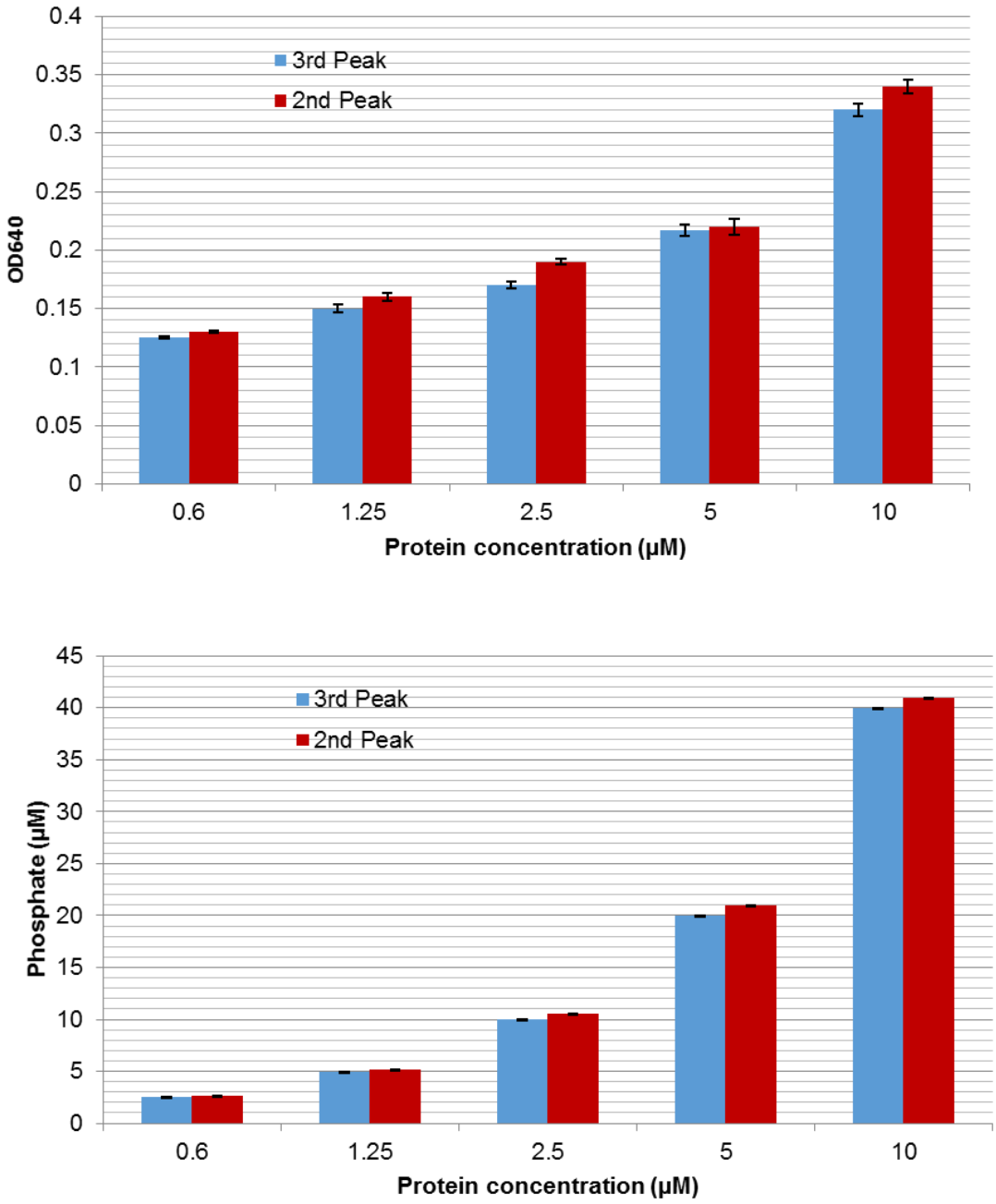
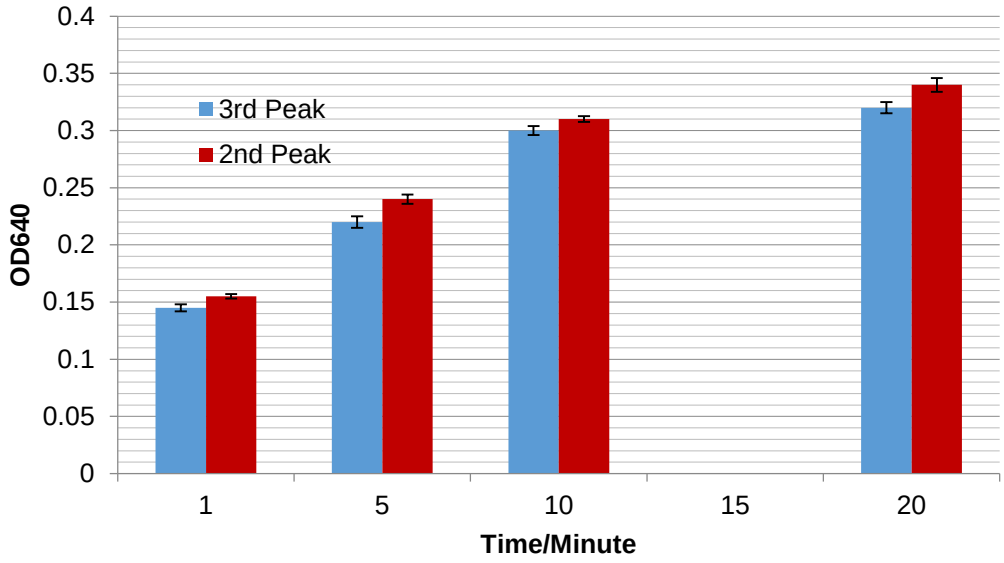


Figure 3.14. Malachite green ATPase test for Holo HolA protein. A. The total OD_{640} for each concentration of HolA protein, and B. The amount of phosphate (µM) released by each concentration was calculated by subtracting the total OD_{640} from the control OD_{640} and comparing the net OD with the calibration curve of standard concentration of KH_2PO_4 in Figure 3.12B.



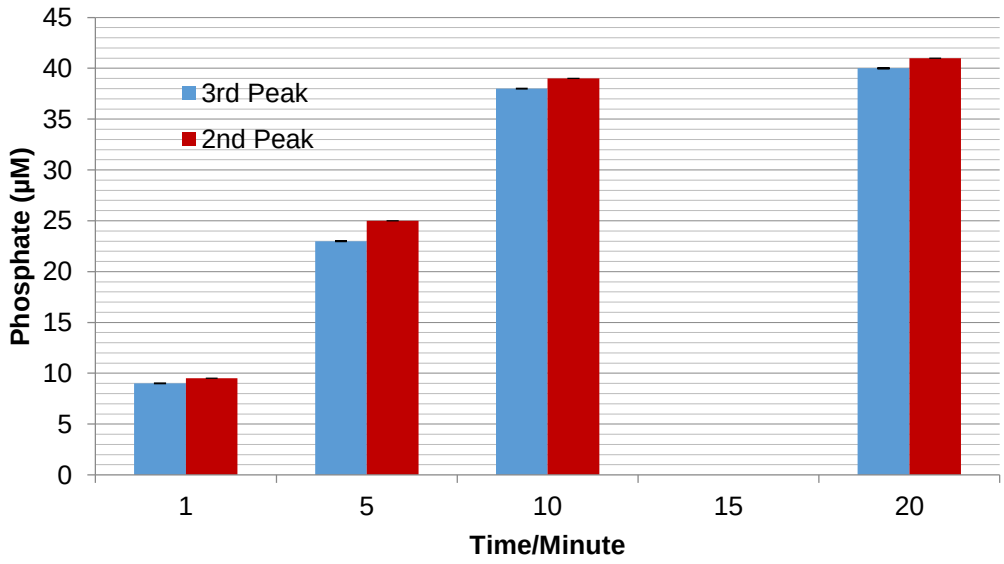


Figure 3.15. Malachite green ATPase test for Holo HolA protein at different time. A. The total OD_{640} for 10µM HolA protein within the time shown, and B. The amount of phosphate (µM) released was calculated by subtracting the total OD_{640} from the control OD_{640} and comparing the net OD_{640} with the calibration curve of standard concentration of KH_2PO_4 in Figure 3.12B.

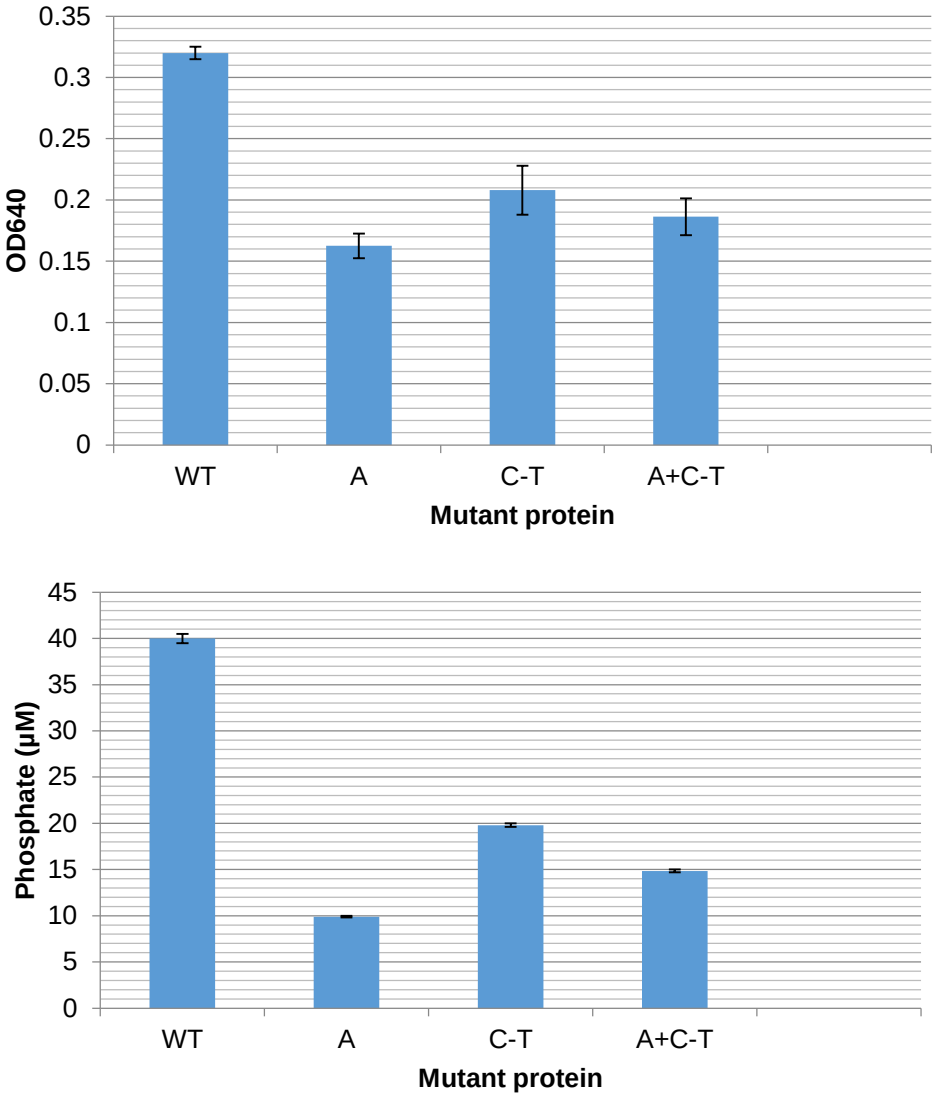


Figure 3.16. Malachite green ATPase test for mutant HolA proteins (A & C-T). Top: The total OD_{640} for 10µM HolA protein within 20 min, and Bottom: The amount of phosphate (µM) released was calculated by subtracting the total OD_{640} from the control OD_{640} and comparing the net OD_{640} with the calibration curve of standard concentration of KH_2PO_4 in Figure 3.12B. The proteins were tested separately as shown in A and C-T column. In A+C-T column, equal amount of each mutant was mixed.

3.4. Discussion

The gel filtration chromatograph showed three distinct peaks which might be three different oligomeric forms of the protein (Figure 3.5A). However, when the 1st and 2nd Peaks of the protein were rerun separately through the gel filtration column using the buffer with different salt concentrations, the protein could not reform the same three peaks and the vast majority kept the same conformations (Figure 3.17). This may mean that the protein samples in these peaks have different conformations or different folding state rather than the different oligomeric state. The elution of the protein with the void volume may be due to the extended protein molecules which have a linear shape being unable to diffuse into the gel pores. However, the globular proteins may penetrate the pores to varying degrees based on their size and therefore elute more slowly. The different conformation of the protein might either due to the natural conformational change of the domains to catalyse the joining of two cysteines or might due to improper folding of the protein. The *holA* gene from the marine bacteria SANK may need specific chaperones which might not be provided by *E.coli* BL21 (DE3) leading to improper folding of the protein.

This framework of hypotheses may explain the glutaraldehyde crosslinking results. The protein from the first peak interacts aggressively producing complexes that couldn't run through the gel. This may be that the extended protein exposes more amine groups on the surface which can be interconnected by glutaraldehyde making a sort of lattice sheet (Figure 3.18). The properly folded (globular) protein may expose fewer amine groups on the surfaces and the majority may be embedded between the domains. Structural study

on three different NRPS proteins revealed large interfaces between the NRPS domains particularly between C and A domains. The interface area ranges from 780Å to 1,097Å, forming a catalytic platform (Drake *et al*., 2016). Thus, the third peak protein from the gel filtration may be a properly folded (globular) monomer protein. These suggestions can be supported by the analytical ultracentrifugation results which showed the protein as a monomer protein rather than dimer or hexamer.

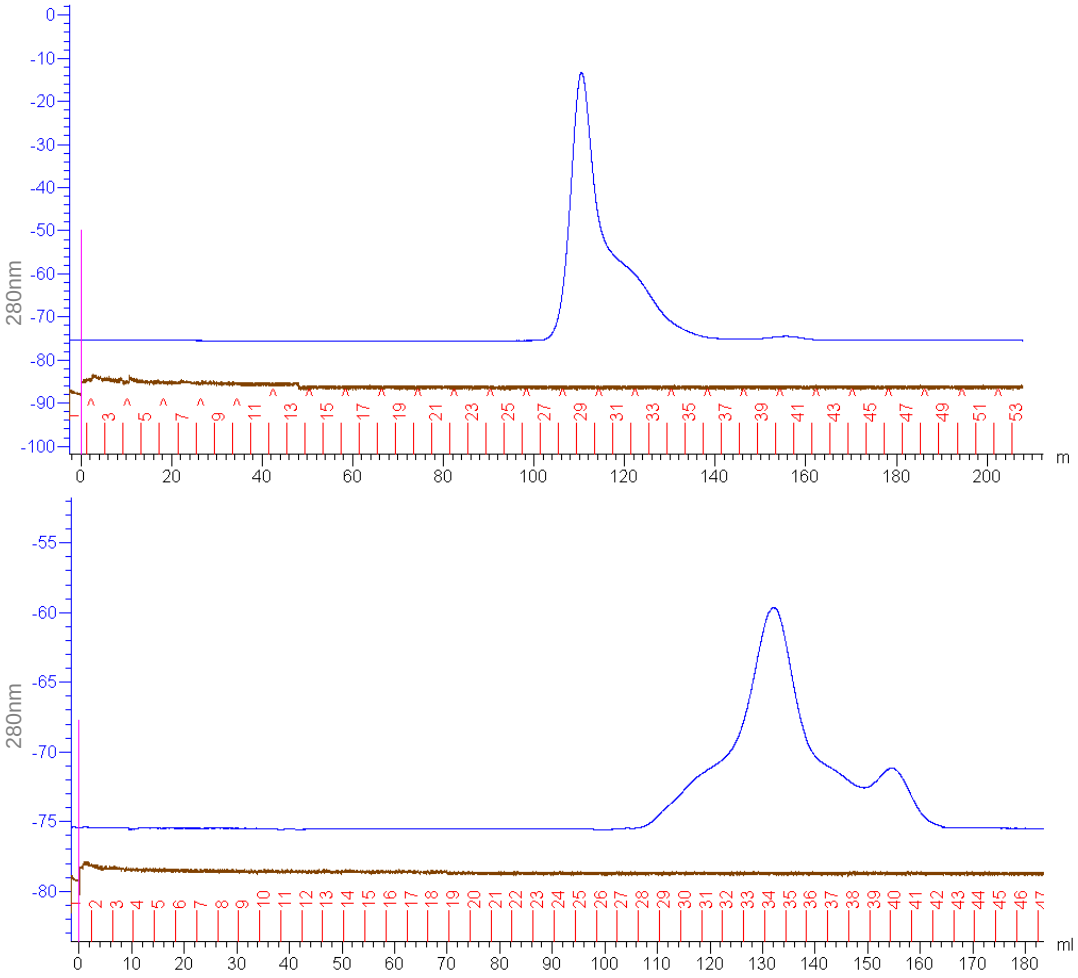


Figure 3.17. Gel filtration rerun for the separate 1st and 2nd Peaks of HolA protein. The peaks in Figure 3.5 were collected separately and reapplied to the gel filtration column. Top: The 1st Peak was eluted with void volume as the first time run, Bottom: The 2nd Peak eluted around 140ml. Note that the peaks location may be slightly different depending on the injection time. The generated shoulders of the peaks may be due the peaks overlapping of the first run or slow conformational change.

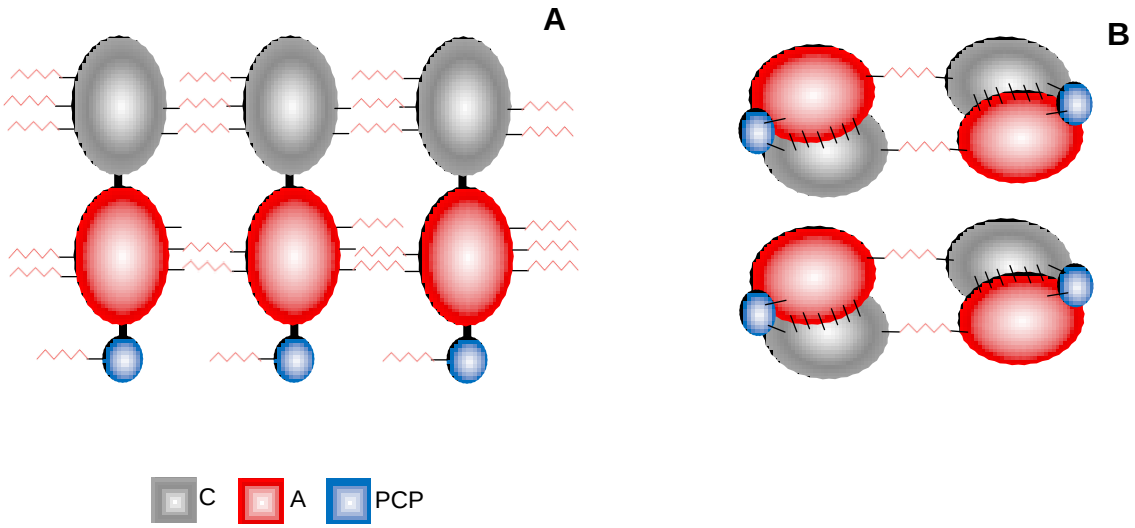


Figure 3.18. Proposed reaction of HolA proteins with glutaraldehyde. A. The extended shape of the protein (1st Peak) increases exposure of the amine groups on the surface of the protein interacting with the crosslinking reagent (Red zigzag line) and forming a lattice sheet and big particles. B. The globular protein exposes less amine groups forming smaller particles (2nd Peak).

The bacterial two hybrid system showed that HolA protein interaction cannot be detected *in vivo*. There are some possibilities which may lead to this result. The size of the HolA protein, which consists of three domains and about 125 KDa, may prevent the physical contact of adenylate cyclase fragments T25 (fused to N-terminus of HolA protein) and T18 (fused to the C-terminus of HolA protein) resulting in a negative outcome. This was addressed by fusing HolA protein to the same terminus for both T25 and T18 fragments using pUT18C vector which enables fusion of T18 to the N-terminus of HolA protein. Another solution for this problem is to fuse a short fragment or single

domain of HolA protein to T25 and T18. The results for both of these approaches were also negative confirming that the protein size should not be the problem. One of the Two Hybrid system drawback is the high possibility of false positive and false negative results *(*Deane *et al*, 2002). The fusion of T25 and T18 with the target protein may alter its properties leading to false results. In addition, this approach cannot detect weak protein-protein interactions *in vivo*. Finally, the improper post-translational modification of protein when expressed in different species is among the possibilities that result in false negative results (Stynen *et at,* 2012). In spite of these facts, however, this result is consistent with gel filtration, AUC and ATPase assays (see below).

The protein samples purified by gel filtration were used to test the catalytic activity of the adenylation domain. The malachite green ATPase assay was used to test both the protein of the 3rd and 2nd Peaks. The results showed that the phosphate precipitation by malachite green depends on protein concentration and there is no significant difference between the protein from the "3rd" and "2nd" Peaks. After calibration with KH_2PO_4, the assay shows that in the presence of cysteine each 10 µmole of Apo protein can release about 20µmole phosphate. This indicates that the Adenylation domain is active and activates only one cysteine molecule releasing one molecule of PPi. The HolA protein was co-expressed with TmlN to ensure the 4-phosphopantetheinylation of the PCP domain (Holo protein) as the *E.coli* PPTase is not guaranteed. When the test was repeated with Holo protein, the phosphate released was twice as much as released by the Apo protein. This may mean that the A domain activates two molecules of cysteine releasing two molecules of the PPi. In case of Holo protein, A domain activates the

cysteine as cysteinyl-adenylate followed by nucleophilic attack of the 4-phosphopantetheine arm leaving A domain to activate the second cysteine (i.e. A domain works in an iterative way). This unique feature of an iterative adenylation domain was reported in congocidine assembly by the cognate NRPS in *Streptomyces ambofaciens* (Juguet *et al*., 2009) and the adenylation domain from *Acinetobacter baumannii* (Drake *et al*., 2016). In Apo protein, A domain activates only one cysteine as the 4-phosphopantetheine arm is missing.

Based on the amount of the PPi released by A domain, the ATPase assay could be used to determine the oligomeric state of the protein and to test the ability of the monomers to interact with each other *in trans*. Thus, mutations in the active sites of the domains were created to produce two HolA proteins (see Section 3.3.4), HolA A with inactive A domain (i.e active C and PCP) and HolA C-T with inactive C and PCP domains (active A). Both these mutant proteins were tested by ATPase separately which showed a reduction in the amount of PPi released by each. As shown in Figure 3.16 HolA A releases about 25% of the PPi compared to the native Holo protein. This means that the point mutation K515T does not inactivate the A domain completely. This conserved Lys515 is located in the active site and according to the alignment with homologous proteins (Conti *et al*., 1997), it interacts with both the substrate and the ATP. However, the binding site for the substrate may vary depending on the substrate shape and size while the binding site for ATP is highly conserved in these proteins. Previous study identified the conserved residues (SGTTGxPKG) in ATP binding site and showed that the replacement of the lysine residue in this motif with threonine in

tyrocidine synthetase I (TycA) from *Bacillus brevis* leads to complete loss of activity (Gocht and Maraheil, 1994). Thus, the conserved residues SGSTGEPKG in the ATP binding site in HolA protein would be a better target to inactivate A domain and the replacement of Lys180 in HolA is more likely to inactivate A domain (Figure 3.19).

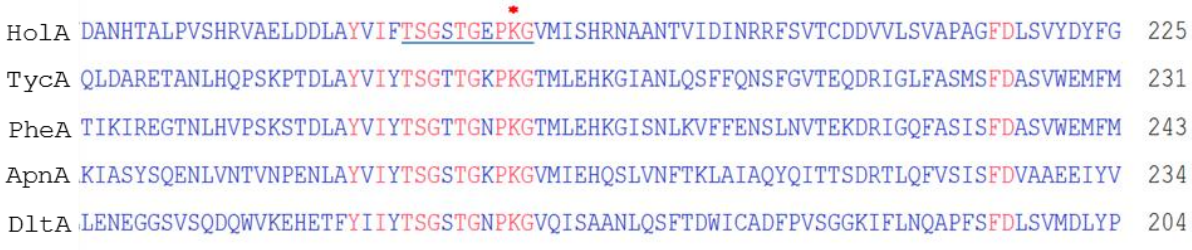


Figure 3.19. Multiple sequence alignment of A domain with homologous proteins. The highly conserved residues in the ATP binding site are shown. The Lys180 in HolA protein is marked by red asterisk.

HolA C-T exhibited similar activity to the Apo protein which means that HolA lost the 4-phosphopantetheine arm because the serine residue in the active site of PCP was replaced with alanine. HolA A and HolA C-T were mixed to test the ability of the A domain of HolA C-T to load cysteine to the PCP domain of HolA A *in trans*. As shown in Figure 3.16, this mixture could not restore the activity to the same level as the WT Holo HolA protein. This may mean that the monomers cannot interact with each other and the ability of the native Holo HolA protein to release 40µM phosphate is due to internal monomer reactions (i.e. A domain activates two cysteine). Thus, the ATPase assay result is consistent with other assays that showed HolA as a monomer protein.

According to the results presented here, some specific features of HolA protein can be outlined which enable assembly of the pyrrothine molecule by one module of the NRPS in the thiomarinol gene cluster. The A domain may work iteratively and introduce two cysteine molecules to the C and PCP domains. Phylogenetic analysis subdivided the C domain in NRPS into six different functional subtypes: Cyclization, starter Condensation, $^{L}C_{L}$, $^{D}C_{L}$, E, dual E/C domains (Rausch *et al*., 2007). The conserved residues in the active site are different depending on the catalytic function. The C domain possesses the His-motif HHXXXDG in the active site that catalyses peptide bond formation. The Cy domain catalyses the cyclodehydration of its own generated peptide bond and the His motif replaced by DXXXD (Grünewald and Marahiel, 2013). The multiple sequence alignment revealed that the C domain in HolA belongs to Cy subtype and possesses a unique conserved sequence, DLIFVD. In congocidine biosynthesis from *Streptomyces ambofaciens,* one of its C domain accepts CoA- activated substrate rather than substrate introduced by the PCP-Ppant arm (Juguet *et al*., 2009).

The linker region between Cy and A domain in HolA consists of twelve residues which facilitate the conformational change during the catalytic process between these two domains creating a platform to deliver the substrate as reported in the recent structural study (Drake *et al*., 2016). The linker region between A and PCP domain consists of fifteen residues which may allow more flexibility to move and make different conformations to visit both C and A catalytic sites.

These features suggest possible initial steps of pyrrothine biosynthesis by HolA as summarized in Figure 3.20. The process starts with the A domain which activates the

first cysteine and transfers it to the 4-phosphopantetheine arm on the PCP domain. The activated cysteine is covalently bound as a thioester to the free thiol of 4-phosphopantetheine arm. The next step involves the activation of the second cysteine either by A domain or may be by a stand-alone Acyl-CoA synthetase provided either by pTML1 or chromosomal DNA as reported in congocidine biosynthesis. The cysteine activation as a CoA-thiolester depends on the C domain ability to accept such substrates as reported in the congocidine system. To catalyse the peptide bond formation and cyclization, all the three domains should undergo a conformational change to guarantee the delivery of the substrates to the catalytic site of the C domain which accepts the substrates from two sides.

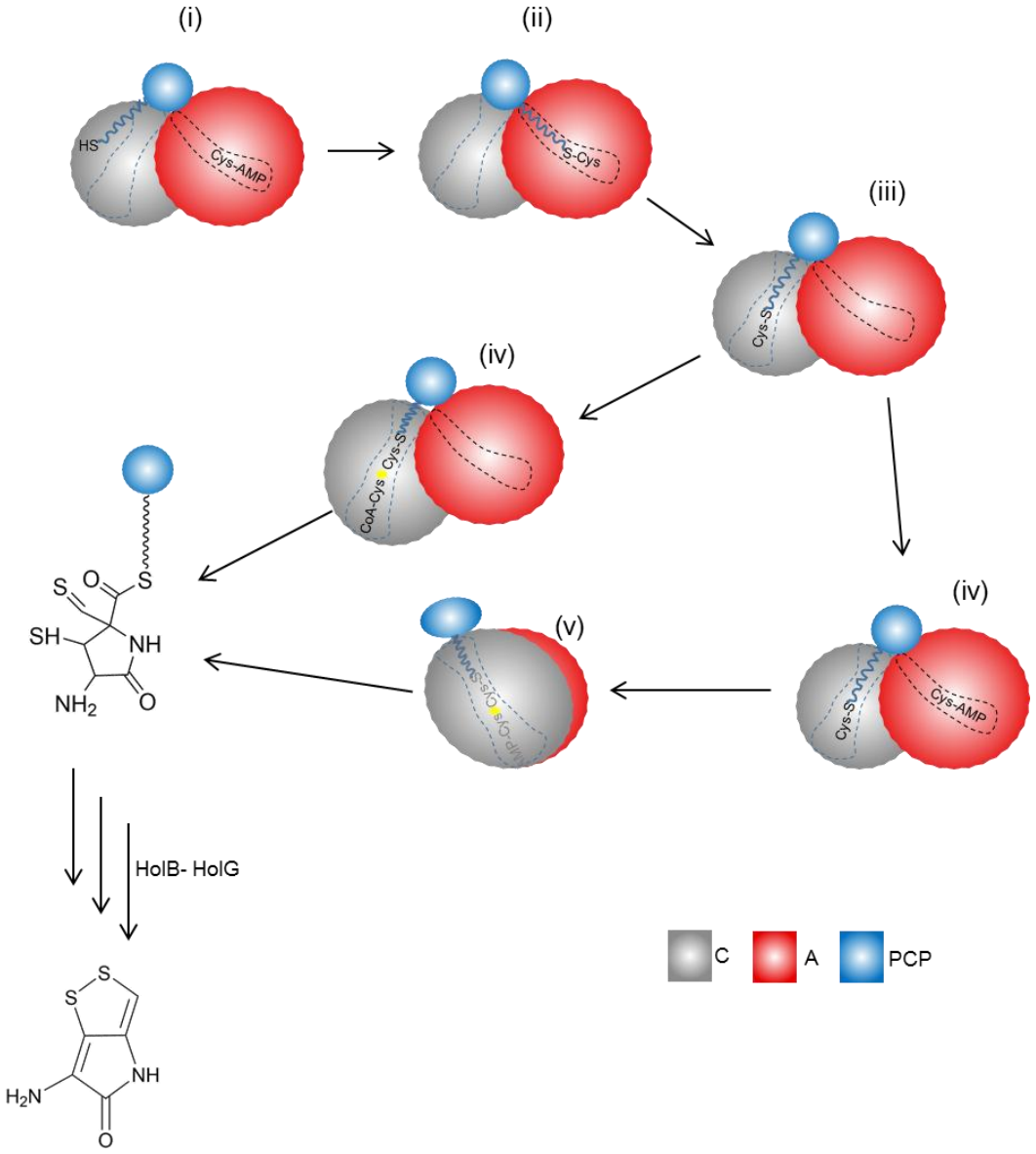


Figure 3.20. The proposal pathway for pyrrothine biosynthesis. (i) Activation of the 1st cysteine as AMP by A domain; (ii) Transferring of the activated cysteine to 4-Ppnt (wave line) on PCP; (iii) The loaded Ppnt moves to C domain active site (dash line); (iv) Either C domain accept activated CoA-cysteine or A domain activates the 2nd cysteine in the iterative manner and (v) Conformational change of the domains to deliver both cysteines into the active site of C domain to achieve peptide bond (yellow star) formation followed by cyclization.

CHAPTER FOUR

4. GENERAL DISCUSSION, CONCLUSION AND FUTURE WORKS

4.1. Overview

Polyketides are structurally diverse natural compounds produced in many bacteria and fungi by polyketide synthases (PKSs). Because of the pharmacological activities, polyketides have occupied a significant area in the drugs industry and become a subject for intensive research and studies to produce novel derivatives with improved properties. Many approaches of genetic engineering have been applied to manipulate the enzyme machinery responsible for polyketides assembly. In general, these approaches either target the enzymes responsible for the backbone assembly or target tailoring enzymes which direct post-assembly modifications. Post-PKS enzymes add significant functional groups to the polyketides backbone such as glycosyl, methyl and hydroxyl groups. These impart the produced compounds structural diversity and alter biological activities (Rix *et al.,* 2002; Weissman and Leadlay, 2005).

The enzymes that introduce hydroxyl groups to the product backbone have received considerable interest and become the subject of many studies. Mutation and heterologous expression experiments have been used to exploit such enzymes to generate novel derivatives (Rix *et al.,* 2002; Wu *et al.,* 2016). Site-directed mutagenesis of Ala245 to Thr in the active site of 6-hydroxyerythonolide B hydroxylase altered the substrate specificity of this enzyme (Xiang *et al.*, 2000). Inactivation of nonheme-ketoglutarate-dependent oxygenase in *Glarea lozoyensis* ATCC 20868 abolished the

pneumocandin A_0 production and increased the yield of pneumocandin B_0, the semisynthetic precursor of the antifungal drug, caspofungin acetate (Chen *et al.,* 2015).

The characterization of tailoring enzymes not only contributes to the generation of new products with altered biological activities but also provides a significant tool to study the biosynthetic pathways. Many genetic manipulation studies in our lab which targeted the *mup* cluster and the studies that targeted the tailoring enzymes revealed quite important facts about the biosynthesis of mupirocin and genetic manipulation allowed the production of many derivatives (Gurney and Thomas, 2011; Gao *et al*., 2014). However, quite a few such studies were carried out about the thiomarinol cluster because of the genetic manipulation difficulties (See below), and the majority of the proposed genes function and biosynthesis pathway is based on the sequence similarities of the mupirocin and thiomarinol clusters (Fukuda, *et al*., 2011). As an alternative strategy, the *in trans* expression of the genes between the heterologous hosts and the characterization *in vitro* of purified enzymes have been carried out to investigate gene function. For example, *in vitro* analysis revealed that both TmlU and HolE are necessary to join pyrrothine molecule with marinolic acid (Dunn *et al*., 2015). The same strategy was used to investigate TmuB responsibility in the thiomarinol cluster.

In modular type I PKSs such as deoxyerythronolide B synthase (DEBS) which catalyses erythromycin biosynthesis, the gene order is consistent with the biosynthetic order (Donadio *et al*., 1991). However, the *mup* cluster doesn't follow this rule as revealed by mutational analyses (some of examples are shown in: Hothersall *et al*., 2007, Cooper *et al*., 2005a and 2005b). The location of TmuB at the downstream end of the thiomarinol

cluster is consistent with its function to hydroxylate the product as a final step as revealed by the experiments in this study. Some of the tailoring enzymes in the *mup* cluster work together to achieve single tailoring step and deletion of one of these abolishes the PA-A production (Cooper *et al*., 2005a). Contrariwise, TmuB independently catalyses the product hydroxylation and gene inactivation didn't block the biosynthesis machinery.

Both thiomarinol C and G (Figure 1.14) were isolated as minor products from the WT producer (Shiozawa, *et al*., 1997). Thiomarinol C lacks a hydroxyl group at C4, while thiomarinol G lacks hydroxyl groups at C4 and C6 with an extra hydroxyl group at C8. From these facts, one can suggest two important points about these hydroxylations in the thiomarinol system. First, the presence of hydroxyl group at C8, prevents both the C4 and C6 hydroxylation which means that the enzyme responsible for 6-hydroxlation might display the same substrate specificity as TmuB does. Second, the un-hydroxylated intermediate precursor of thiomarinol G could be accepted by all the enzymatic assembly line through the thiomarinol cluster which might mean that the 6-hydroxylation is a final tailoring step either before or after 4-hydroxylation.

The inactivation of TmuB (Point mutation I109N) in the thiomarinol producer, *Pseudoalteromonas sp.* SANK 73390, is among the few successful mutations that were achieved in this bacterium in our Lab. The site directed mutagenesis strategy used in our lab to introduce mutations has worked properly in other bacteria such as *Pseudomonas fluorescens* NCIMB 10586. However, the same strategy doesn't seem to work efficiently in *Pseudoalteromonas sp.* SANK 73390. This prevents its use in

mutagenesis and complementation analyses to characterize and manipulate many genes in the thiomarinol cluster as was done in the *mup* cluster in *P. fluorescens* NCIMB 10586.

Pseudoalteromonas spp SANK 73390 is a marine bacteria and many features have not been characterized yet which might make the genetic modifications difficult. These include the efflux pump system which makes the bacteria resistant to antibiotics. This feature is a very critical issue as antibiotic selection steps during mutation procedure is an important step to distinguish between transconjugant and non-transconjugant phenotypes (Higher antibiotic concentration was used in this study, see below). The second feature is the barriers mechanism such as surface exclusion and restriction systems that could be used by bacteria against foreign DNA which might reduce the transformation efficiency.

Previous studies reported many successful mutagenesis experiments in *Pseudoalteromonas* species, but all the mutations in all these species were targeted chromosomal genes (Wang *et al*., 2015). However, the thiomarinol gene cluster in *Pseudoalteromonas spp* SANK 73390 is encoded by circular plasmid pTML1 rather than chromosomal genome which may make the mutation difficult. As an attempt to overcome these difficulties, many strategies have been used to mobilise pTML1 from SANK, but unfortunately, none were successful due to unclear reasons (Omer-Bali, 2013).

Our enthusiasm for the TmuB story, encouraged us to carry out a mutational experiment in the original producer, *Pseudoalteromonas spp* SANK 73390, to deactivate *tmuB* as described with required modifications. Four remarkable points were the differences between the current mutation (TmuB I109N) with the previous successful and unsuccessful trials. First, the current mutation is a point mutation replacing only two bases (Figure 2.41) while the previous experiments involved gene knockout. Second, in the previous experiments, the constructed vector was carried out by cloning two separate inserts (500bp) flanking the target gene, while in TmuB mutation about 800bp was cloned into the suicide vector which might give more chance to integrate into the target DNA. Third, the yellow pigment of the colonies worked as an indicator for the integration and excision of the suicide vector. This narrowed down the search to focus on particular colonies with a high possibility of successful plasmid integration. The integration of the suicide vector upstream of *holA-H* in the thiomarinol cluster, interrupted the formation of the pyrrothine molecule and generated white colonies. When the suicide vector has excised, the yellow pigment restored and the massive screening of the colonies identified the mutant phenotype from these colonies. Fourth, a high concentration of antibiotics, which might be the most important point, was used for transconjugant selection to overcome any possibility of antibiotic resistance problem (see Section 2.3.10.1 in Chapter 2).

According to the principle of the mutation strategy, the proportion of the generated mutant phenotype should be equal to the wild type (i.e. 50% of the growing colonies). This was noticed in the previous mutation studies with *Pseudomonas fluorescens*

NCIMB 10586, while with *Pseudoalteromonas spp* SANK 73390, the vast majority of the transconjugated cells reverted to the wild type. In conclusion, it is really not clear yet why the gene manipulation in this species is quite difficult. Therefore, more studies are required to discover the mysterious features of the producer or to characterize the pTLM1.

Since non-ribosomal peptides are structurally diverse products, the biosynthesis system responsible for the product assembly should possess the same diversity. The product diversity either comes from the unusual building units or from the unusual features of the system. Therefore, more biochemical and structural studies are required to predict the dynamic interactions between modules and domains that mediate the enzymatic processes. Characterization of more NRPS proteins from diverse origin might broaden our knowledge and understanding of non-ribosomal peptide biosynthesis and exploit this complex machinery. Investigating the NRPS from marine bacteria such as HolA protein which exhibited some unusual behavior might add new aspect in this regard.

4.2. Conclusion

Within the context of the above general points the key conclusions that can be drawn from the work described in this Thesis are:

1- TmuB in the thiomarinol cluster catalyses C-4 hydroxylation of thiomarinol as a tailoring step. In addition to C9-10 epoxide and pyrrothine, this uncovers one more difference between mupirocin and thiomarinol structure assembly at the genetic level.

2- TmuB expression in heterologous hosts catalyses analogous substrates efficiently except those with a different pyran ring (with extra groups attached). The mutagenesis based on the modelling and docking results couldn't change the specificity of the TmuB to catalyse such substrates.

3- *In vitro* analysis revealed that TmuB is able to catalyse analogues with extra groups on the pyran ring but with much lower efficiency. However, apart from the pyran ring the analogues with different C-terminal structure such as thiomarinol C doesn't affect TmuB catalytic efficiency.

4- The hydroxylation or absence of hydroxylation of the natural product backbone plays a substantial role in antibacterial potency and may give it new biological activities.

5- TmuB works independently to add a hydroxyl group to thiomarinol in the sense that its inactivation doesn't cause a defect in the rest of the biosynthesis system suggesting that the hydroxylation occurres as a final step either before or after the pyrrothine joining to the marinolic acid.

6- The successful of the mutation experiment in *Pseudoalteromonas spp* SANK 73390 give us a prospect of success and paves the way to carry out more experiments with

required modifications. On the other hand, more research is required to uncover the unknown features of the producer which are needed to achieve successful mutagenesis.

7- HolA, which assembles two cysteines to form the pyrrothine molecule, possesses some unusual features which facilitate the joining and cyclization of two cysteine molecules. The protein produced in heterologous bacteria appear to exhibit three distinct conformations or folding states. The results showed HolA as monomer protein and the conformational change of the domains may be necessary to achieve the catalytic activity.

4.3. Future work

4.3.1. TmuB project

In an attempt to change the TmuB specificity, many mutations of the residues in the active site were carried out depending on the bioinformatic analyses which predict the homology model and substrate docking. However, the homology model of TmuB was generated depending on the other resolved proteins which catalyse substrates that are structurally different from PA-A. Therefore, the orientation of residues in the active site may slightly differ from reality. In addition, the interactions between the substrates and the protein structures in the docking process depend on the parameters that have been used (i.e. different parameters resulted in different conformations). Finally, the correct packing of the loops around the active site might prevent the substrate docking in the

right position. Therefore, to obtain a correct and accurate structure of the protein and identify the critical residues in the active site, the protein should be crystalized as a complex with the substrate. Many studies have identified the crystal structure of nonheme-iron(II)/2-oxoglutarate-dependent hydroxylase successfully (Hausinger, 2004). Mutagenesis experiments based on a determined crystal structure can not only facilitate specificity changes but could also allow modification of the catalysis of the protein (Yan *et al*., 2015). Therefore, the future work for TmuB specificity should include crystallography trials followed by mutagenesis experiments according to the obtained crystal structure. Work carried out by colleagues in Bristol, specifically in the laboratory of Paul Race did obtain crystals but apparently the diffraction data obtained did not allow structure determination and so this was not included in this Thesis. The pseudomonic acid derivatives produced in this study should be subject for further studies to explore the biological activities which might be useful. Thiomarinol C was produced as a major product from the mutant (I109N) SANK and the antibacterial potency was lower than thiomarinol A but still higher than PA-A. Therefore, the cytotoxicity test of this compound in eukaryotic cells might explain the role of 4-hydroxylation.

4.3.2. HolA project

As an initial test to check the activity of the A domain, the malachite green ATPase test was used which is the indirect method to measure the amount of the pyrophosphate (PPi) released by A domain. Prior to adding reagent, pyrrophosphatase (PPase) was

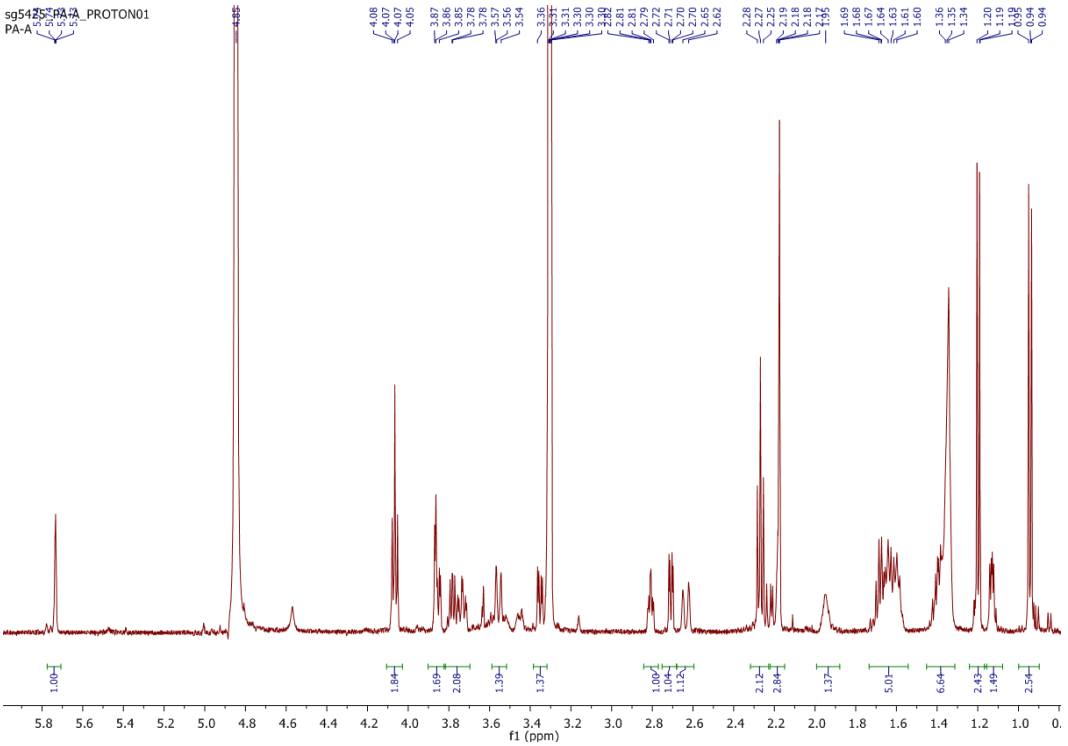
added to convert PPi to phosphate ion which can be detected by the test. Therefore, the amount of the phosphate measured by this test may represent the activity of the second enzyme (PPase) rather than HolA A domain. To overcome this possibility, a higher concentration of PPase was used and incubated for a longer time to allow conversion of all of the PPi released by the A domain into phosphate ions. A more accurate test could be used to directly measure pyrophosphate released by the A domain such as Pyrophosphate Assay Kit (Fluorometric) which provides the most robust spectrophotometric method.

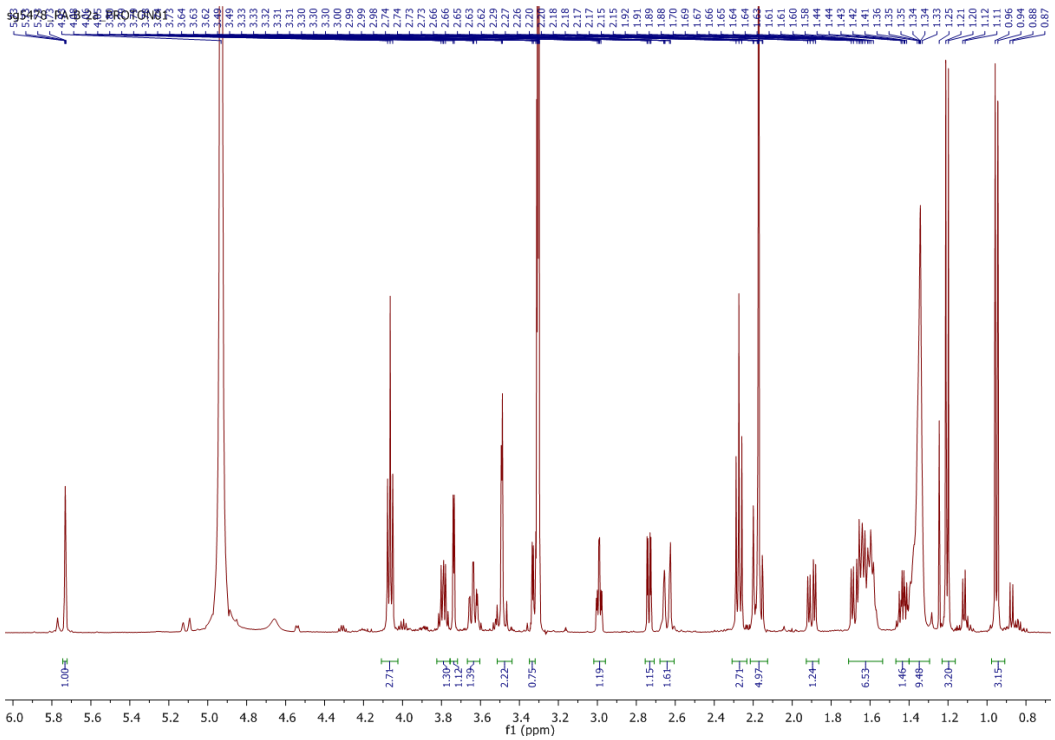
Mass spectrometry (MS) experiment could be used to test the ability of HolA to both activate and condense cysteine. To confirm that these two are loaded as a cys-cys dimer on the PCP domain, the protein should be subjected to proteolytic digest and the molecular mass of the peptide should increase by the expected amount. Both pyrophosphate assay kit and MS could be used with the mutant proteins to check the oligomeric state of HolA protein as described in Chapter 3. For this purpose, the complete inactivation of individual domains should be created by targeting the critical residues in the active site as determined in Chapter 3.

Structural studies revealed intermolecular interactions and the conformational changes between NRPS domains which are necessary to achieve the catalytic activity. Therefore, the crystal structure of HolA protein might tell us the significance of the different peaks shown on the gel filtration chromatography in terms of catalytic activity and this might provide fundamental insights into the way in which NRPSs work which is an area of great potential for the future.

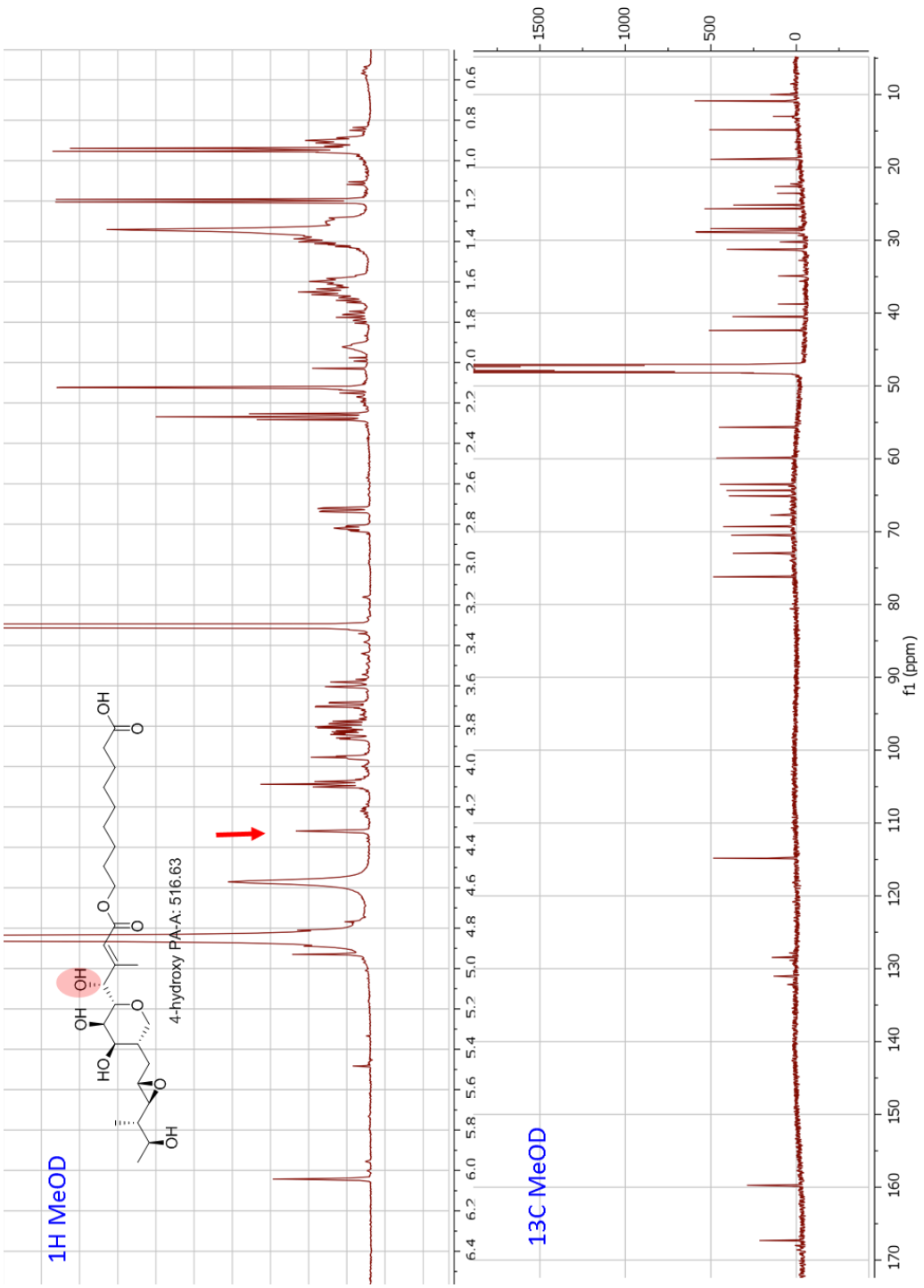
APPENDIX

Appendix A. NMR for PA-A (Top) and PA-B (Bottom

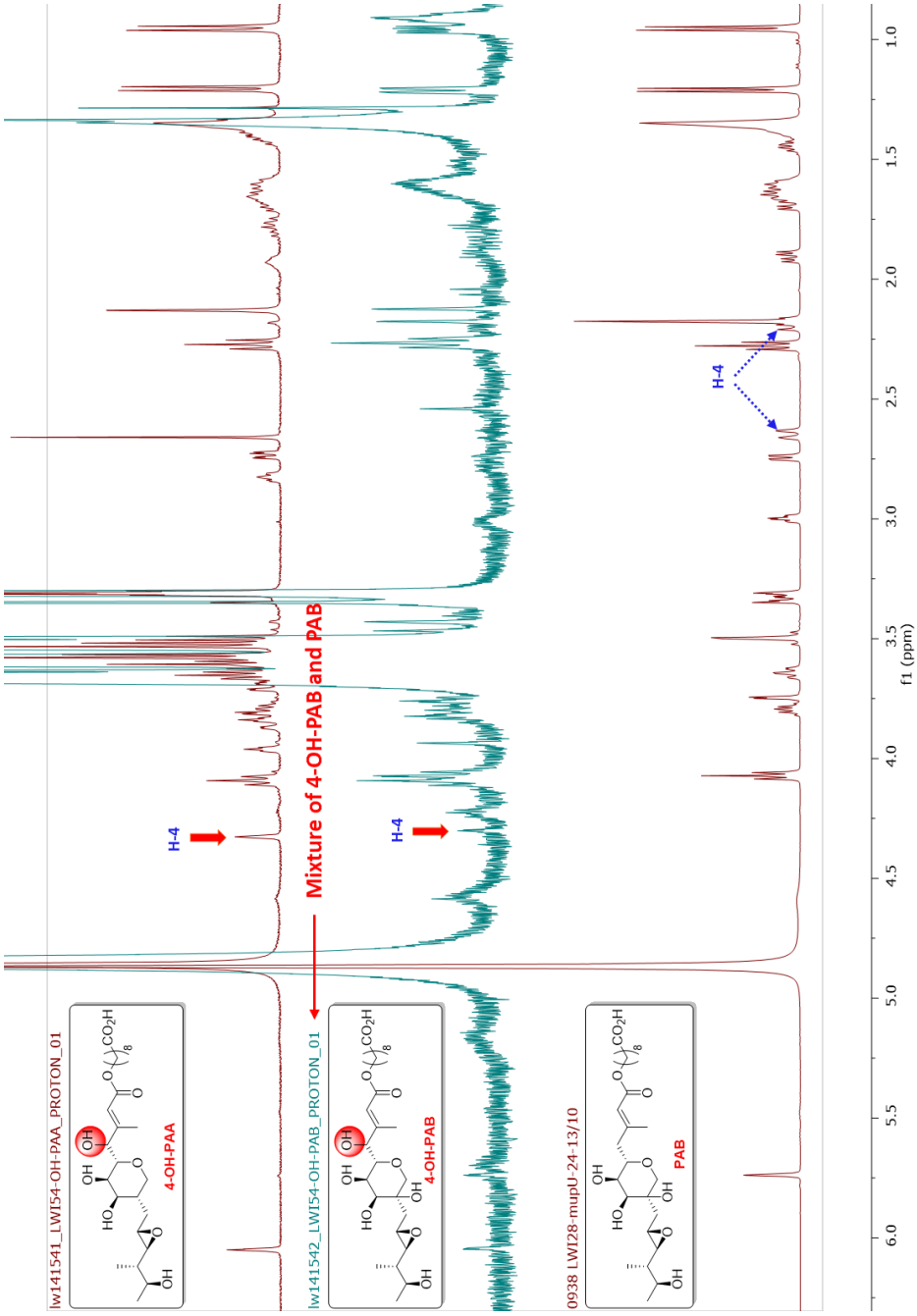




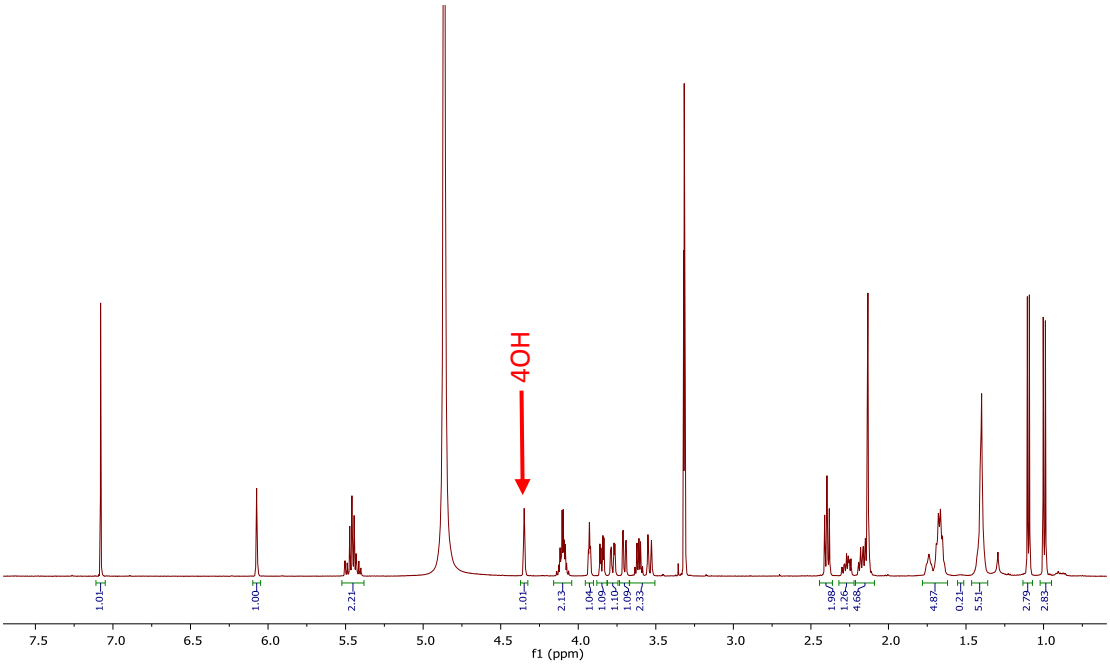
Appendix B. The NMR for 4-hydroxylated PA-A

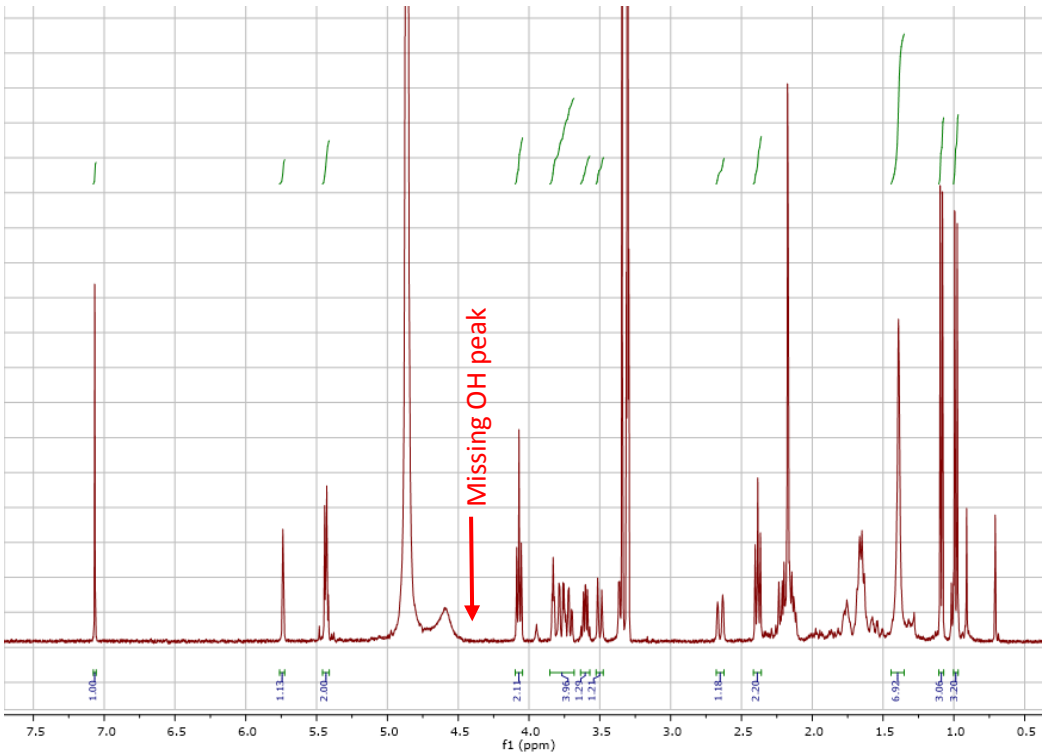


Appendix C. The NMR for, PA-B, 4-hydroxylated PA-A and PA-B



Appendix D. The NMR for thiomarinol A (Top) and thiomarinol C (Bottom)





REFERENCES

6. References

Ahmed A. (2010) “Determining the toxicity of Thiomarinol and related novel compounds.”, Unpublished, a research project report submitted to The University of Birmingham.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, and David J.L(1997) “Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.” *Nucleic Acids Res*., 25, 3389-3402.

Amin S.R., Erdin S., Ward R.M., Lua R.C., Lichtarge O. (2013) “Prediction and experimental validation of enzyme substrate specificity in protein structures.” *Proc Natl Acad Sci*., 110 (45): E4195–E4202.

Antonio M., McFerran, N. and Pallen, M. J. (2002) “Mutations Affecting the Rossman Fold of Isoleucyl-tRNA Synthetase Are Correlated with Low-Level Mupirocin Resistance in *Staphylococcus aureus*.” *Antimicrobial Agents and Chemotherapy*, 46, (2): 438-422.

Birnboim H.C. and Doly J. (1979) “A rapid alkaline extraction procedure for screening recombinant plasmid DNA.” *Nucleic Acids Research*, 7, 1513-1527.

Bräuer A., Beck P., Hintermann L., & Groll M. (2016) “Structure of the Dioxygenase AsqJ: Mechanistic Insights into a One-Pot Multistep Quinolone Antibiotic Biosynthesis.” *Angew Chem Int Ed Engl*, 55(1): 422-426.

Brogden R.N., Carmine AA., Heel R.C., Speight T.M. and Avery G.S. (1982) “Trimethoprim: a review of its antibacterial activity, pharmacokinetics and therapeutic use in urinary tract infections”, *Drugs*, 23, 405-430.

Brown Gene M. (1962) “The biosynthesis of folic acid.” *The Journal of Biological Chemistry*, 237 (2): 536-540.

Brown J.R., Gentry D., Becker J.A., Karen Ingraham K., Holmes D.J. and Stanhope M.J. (2003) “Horizontal transfer of drug-resistant aminoacyl-transfer-RNA synthetases of anthrax and Gram-positive pathogens.” *EMBO reports,* 4, (7): 692-698.

Buchan D.W., Minneci F., Nugent T.C., Bryson K., Jones D.T. (2013) “Scalable web services for the PSIPRED Protein Analysis Workbench.” *Nucleic Acids Research*, 41, W349-57.

Bushley K. E., Ripoll D. R., & Turgeon B. G. (2008). “Module evolution and substrate specificity of fungal nonribosomal peptide synthetases involved in siderophore biosynthesis.” *BMC Evol Biol*, 8, 328.

Center for Disease Dynamics, Economics & Policy (CDDEP), (2015) “The state of the world’s antibiotics 2015.” Washington, DC 20005, USA.

Centers for Diseases Control and Prevention (CDC) report, (2013) Antibiotic resistance threats in the United State.” April, 2013.

Challis G.L. and Naismith, J.H. (2004) “Structural aspects of non-ribosomal peptide biosynthesis.” *Current Opinion in Structural Biology*. 14, 748-756.

Chen L., Yue Q., Li Y., Niu X., Xiang M., Wang W., Gerald F., Bills GF., Liu X., An Z. (2015) “Engineering of *glarea lozoyensis* for exclusive production of the pneumocandin b0 precursor of the antifungal drug caspofungin acetate.” *Applied and Environmental Microbiology*, 81, 1550 –1558.

Chowdhury R, McDonough MA, Mecinovic J, Loenarz C, Flashman E, Hewitson KS, Domene C, Schofield CJ (2009) “Structural basis for binding of hypoxia-inducible factor to the oxygensensing prolyl hydroxylases.” *Structure*, 17:981-989.

Clayton J.P., O'Hanlon P.J. and Rogers N.H. (1980) “The Structure and Configuration of Pseudomonic Acid C.” *Tetrahedron Letters*, 21, 881-884.

Clifton IJ, McDonough MA, Ehrismann D, Kershaw NJ, Granatino N, Schofield CJ (2006) “Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins.” *Journal of inorganic biochemistry*, 100:644-69.

Cohen S.N., Chang A.C.Y. and Hsu L. (1972) “Non chromosomal antibiotic resistance in bacteria; genetic transformation of *Escherichia coli* by R factor DNA.” *Proceedings-National Academy of Sciences USA*., 69(8): 2110-2114.

Cole C., Barber J.D., Barton G.J. (2008) “The Jpred 3 secondary structure prediction server.” *Nucleic Acids Research*, 36, W197-W201.

Condurso H.L., Bruner S.D. (2012) “Structure and noncanonical chemistry of nonribosomal peptide biosynthetic machinery.” *Natural Product Report,* 29(10):1099-110.

Conti E., Stachelhaus T., Marahiel M.A., & Brick P. (1997) “Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S.” *EMBO J*, 16(14): 4174-4183.

Cookson B.D. (1998) “The emergence of mupirocin resistance: a challenge to infection control and antibiotic prescribing practice.” *Journal of Antimicrobial Chemotherapy*, 41, 11–18.

Cookson B.D., Lacey R.W., Noble W.C., Reefves D.S., Wise R., Redhead R.J. (1990) “Mupirocin-resistant *Staphylococcus aureus*.” *Lancet,* 335,1095–1096.

Cooper S.M., Cox R.J., Crosby J., Crump M.P., Hothersall J., Laosripaiboon W., Simpson T.J. and Thomas C. M. (2005a) “Mupirocin W, a novel pseudomonic acid produced by targeted mutation of the mupirocin biosynthetic gene cluster.” *The Royal Society of Chemistry Chemical Communications*, 1179–1181.

Cooper S.M., Laosripaiboon W., Rahman A.S., Hothersall J., El-Sayed A.K., Winfield C., Crosby J., Cox R.J., Simpson T.J. and Thomas C.M. (2005b) “Shift to Pseudomonic Acid B Production in *P. fluorescens* NCIMB10586 by Mutation of Mupirocin Tailoring Genes *mupO, mupU, mupV,* and *macpE.” Chemistry and Biology*, 12, 825-833.

Deane C.M. Deane, Ski U.S, Xenarios L, and Eisenberg D. (2002) “Protein Interactions: Two methods for assessment of the reliability of high throughput observations. *Molecular & Cellular Proteomics,* 1(5): 349-56.

Delcour AH. (2009) “Outer membrane permeability and antibiotic resistance.” *Biochimica et Biophysica Acta*, 1794, 808–816.

DeSieno M.A., van der Donk W.A., Zhao H. (2011) “Characterization and application of the fe(II) and α-ketoglutarate dependent hydroxylase FrbJ.” *Chemical Communications,* 47(36):10025-10027.

Doekel S., Marahiel M.A. (2000) “Dipeptide formation on engineered hybrid peptide synthetases.” *Chemical Biololgy*, 7(6): 373-84.

Donadio S., Staver M.J., Mcalpine J.B., Swanson S.J., and Lowden P.A.S., Bohm G.A., Metcalfe S., Staunton J., and Katz L. (1991) “Modular organization of genes required for polyketide biosynthesis.” *Science*, *252*, 675–679.

Drake E.J., Miller B.R., Shi C., Tarrasch J.T., Sundlov J.A., Allen C.L., Gulick A. M. (2016) “Structures of two distinct conformations of holo-non-ribosomal peptide synthetases.” *Nature*, 529(7585): 235-238.

Du L.C., Sanchez C., Chen M., Edwards D.J., Shen B. (2000) “The biosynthetic gene cluster for the antitumor drug bleomycin from Streptomyces verticillus ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase.” *Chemistry & Biology*, 7, 623–642.

Dunn B.J., Watts K.R., Robbins T., Cane D.E., and Khosla C. (2014) “Comparative analysis of the substrate specificity of trans- versus cis-acyltransferases of assembly line polyketide synthases.” *Biochemistry*, 53, 3796−3806.

Dunn Z.D., Wever W.J., Economou N.J., Bowers A.A., and Li B. (2015) “Enzymatic Basis of “Hybridity” in Thiomarinol Biosynthesis.” *Angewandte Chemie International Edition,* 54, 1 – 6.

Dutton C.J., Hooper A.M., Leadlay P.F., Staunton J. (1994) “Avermectin biosynthesis. Intact incorporation of a diketide chain-assembly intermediate into the polyketide macrocyclic ring.” *Tetrahedron Letters,* 35(2): 327–330.

El-Sayed, A. K., Hothersall, J., & Thomas, C. M. (2001). Quorum-sensing-dependent regulation of biosynthesis of the polyketide antibiotic mupirocin in *Pseudomonas fluorescens* NCIMB 10586. *Microbiology*, 147(Pt 8): 2127-2139.

El- Sayed A.K., Hothersall J., Cooper S.M., *et al* (2003) “Characterization of the mupirocin biosynthesis gene cluster from *Pseudomonas fluorescens* NCIMB 10586.” *Chemistry & Biology*, 10, 419–430.

Eppelmann K., Stachelhaus T., Marahiel M.A. (2002) “Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics.” *Biochemistry,* 41 (30): 9718-26.

Erdin S., Venner E., Lisewski A.M., Lichtarge O. (2013) “Function prediction from networks of local evolutionary similarity in protein structure.” *BMC Bioinformatics*, 14(Suppl 3): S6.

Eriani G., Delarue M., Poch J. Gangloff and Moras D. (1990) “Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs”, *Nature*, 347, 203–206.

European Surveillance Antimicrobial Consumption (ESAC) report (2009), *ESAC yearbook*, ISBN number: 9789057283307.

Falagas M., Rafailidis P., Matthaiou D. (2010) “Resistance to polymyxins: Mechanisms, frequency and treatment options” *Drug Resistance Updates*, 13, 132–138.

Finking R. and Marahiel M.A., (2004) “Biosynthesis of nonribosomal peptides.” *Annual Review of Microbiology*, 58, 453–88.

Floss H.G. (2006) “Combinatorial biosynthesis--potential and problems.” *J Biotechnol*. 124 (1):242-57.

Food and Drug Administration (FAD) “Novel drugs 2015 summary” USA.

Frandsen R. (2010) “Polyketide synthases” [online] Available from: http://www. rasmusfrandsen.dk/polyketide_synthases.htm [Accessed 20 Dec 2012].

Fukuda D., Haines A.S., Song Z., Murphy A.C., Hothersall J., Stephens E.R., Gurney R., Cox R.J., Crosby J., Willis C.L., Simpson T.J. and Thomas C.M. (2011) “A Natural Plasmid Uniquely Encodes Two Biosynthetic Pathways Creating a Potent Anti-MRSA Antibiotic.” *PLoS ONE*, 6(3): e18031.

Fuller A.T., Mellows G., Woolford M., Banks G.T., Barrow K.D., Chain E.B. (1971) "Pseudomonic acid: an antibiotic produced by *Pseudomonas fluorescens*." *Nature* 234, 416–417.

Galtier N., Gouy M., Gautier C. (1996) "SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny." *Computer applications in the biosciences, CABIOS*, 12, 543-8.

Gao S., Hothersall J., Wu J., Murphy A.C., Song Z., Stephens E.R., Thomas C.M., Crump M.P., Cox R.J., Simpson T.S. and Willis C.L. (2014) "Biosynthesis of mupirocin by *pseudomonas fluorescens* NCIMB 10586 involves parallel pathways." *Journal of the American Chemical Society*, 136, 5501−5507.

Gocht M., & Marahiel M. A. (1994) "Analysis of core sequences in the D-Phe activating domain of the multifunctional peptide synthetase TycA by site-directed mutagenesis." *J Bacteriol*, 176(9): 2654-2662.

***Grünewald* J*. and Marahiel* M.A.** (*2013*) "Nonribosomal Peptide Synthesis." *Handbook of Biologically Active Peptides*. Chapter 21. S. 138-149. Academic Press.

Gurney R.S. and Thomas C.M. (2011) "Mupirocin: biosynthesis, special features and applications of an antibiotic from a Gram-negative bacterium." *Appl. Microbiol. Biotechnol*., 90, 11-21.

Haines A., Dong X., Song Z., *et al*. (2013) "A conserved motif flags acyl carrier proteins for b-branching in polyketide synthesis." *Nature Chemical Biology*, 9(11):685-92.

Hanahan D. (1983) "Studies on transformation of *Escherichia coli* with plasmids." *J Mol Biol*., 166, 557-580.

Hausinger R.P. (2004) **"**Fe(II)/*α*-Ketoglutarate-Dependent Hydroxylases and Related Enzymes." *Critical Reviews in Biochemistry and Molecular Biology*, 39, 21–68.

Havemann J., Vogel D., Loll B., Keller U. (2014) "Cyclolization of D-lysergic acid alkaloid peptides." *Chemistry & Biology*, 21,146-55.

Helfrich EJ., and Piel J. (2016) "Biosynthesis of polyketides by trans-AT polyketide synthases." *Natural Product Report*, 33, 231-316.

Hertweck C. (2009) "Biosynthetic logic of polyketide diversity." *Angewandte Chemie International Edition*., 48, 4688 – 4716.

Hodgson J.E., Curnock S.P., Dyke K.G., Morris R., Sylvester D.R. and gross M.S. (1994) "Molecular characterization of the gene encoding high-level mupirocin resistance in *Staphylococcus aureus* J2870." *Antimicrob. Agents Chemother*., 38(5): 1205-1208.

Hooft R., Sander C. and Vriend G. (1997) “Objectively judging the quality of a protein structure from a Ramachandran plot.” *CABIOS*, 13(4): 425-430.

Hoppner A., Widderich N., Lenders M., Bremer E., Smits S.H. (2014) “Crystal structure of the ectoine hydroxylase, a snapshot of the active site.” *The Journal of Biological Chemistry*, 289, 29570-83.

Hopwood D.A. and Sherman D.H. (1990) “Molecular genetics of polyketides and its comparison to fatty acid biosynthesis.” *Annu. Rev. Genet*., 24, 37-66.

Hothersall J., Wu J., Rahman A.S., *et al*. (2007) “Mutational analysis reveals that all tailoring region genes are required for production of polyketide antibiotic mupirocin by *Pseudomonas fluorescens*.” *The Journal of Biological Chemistry*, 282(21): 15451–15461.

Hranueli D., Perić N., Borovička B., Bogdan S., Cullum J., Peter G. Waterman P.G., and Hunter I.S. (2001) “Molecular Biology of Polyfketide Biosynthesis.” *Food Technol. Biotechnol., 39* (3): 203–213.

Hughes J. and Mellows G. (1978) “Inhibition of Isoleucyl-Transfer Ribonucleic Acid Synthetase in *Escherichia coli* by Pseudomonic Acid.”, *Biochemical Journal*, 179, 305-318.

Hur G.H., Vickery C.R. and Burkart M.D. (2012) “Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology.” *Natural Product Report*, 29, 1074-1098.

Hurdle J.G., O'Neill A.J., Ingham E., Fishwick C., & Chopra I. (2004) “Analysis of mupirocin resistance and fitness in *Staphylococcus aureus* by molecular genetic and structural modeling techniques.” *Antimicrob Agents Chemother*, 48(11): 4366-4376.

Ibba M., and Söll D. (2000) “Aminoacyl-tRNA synthesis.” *Annual Review Biochemistry*., 69, 617-650.

Institution of Medicine (1998) “Antimicrobial drug resistance: issues and options.” Workshop report. Washington: D.C., National Academy Press.

Jeanmougin F., Thompson J.D., Gouy M., Higgins D.G., Gibson T.J. (1998) “Multiple sequence alignment with Clustal x.” *Trends Biochem Sci,* 23, 403-5.

Jhstton C.J., Hooper A.M., Leadlap P.F., and St.aamtonb J. (1994) “Avermectin biosynthesis. intact incorporation of a diketide chain-assembly intermediate into the polyketide macrocyclic Ring.” *Tefrahcdmn Latas,* 35 (2): 327-330.

Joosten R.P., te Beek T.A., Krieger E., Hekkelman M.L, Hooft R.W, Schneider R., et al. (2011) “A series of PDB related databases for everyday needs.” *Nucleic Acids Research*, 39, D411-9.

Juguet M., Lautru S., Francou F.X., Nezbedova S., Leblond P., Gondry M., & Pernodet J.L. (2009) “An iterative nonribosomal peptide synthetase assembles the pyrrole-amide antibiotic congocidine in *Streptomyces ambofaciens*.” *Chem Biol*, 16(4): 421-431.

Karimova G. and D. Ladant. (2005) “A bacterial two-hybrid system based on a Cyclic AMP signaling cascade. Chap. 26, pp 499-515. Protein-Protein Interactions.” A Molecular Cloning Manual 2nd Edition. Cold Spring Harbor Laboratory Press. Edited by E. Golemis. Cold Spring Harbor, New York.

Karimova G., N. Dautin and D. Ladant (2005) “Interaction network among *Escherichia coli* membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis.” *J. Bact.,* 187, 2233-2243.

Karimova G., Ullmann A., and D. Ladant (2001) “Protein-protein interaction between *Bacillus stearothermophilus* tyrosyl-tRNA synthetase subdomains revealed by a bacterial two-hybrid system.” *J. Mol. Microbiol. Biotechnol.*, 3, 73-82.

Keatinge-Clay A.D. (2012) “The structures of type I polyketide synthases.” *The Royal Society of Chemistry*. *Natural Product Report*, 29, 1050–1073

Khosla C., Gokhale R.S., Jacobsen J.R. and Cane D.E. (1999) “Tolerance and specificity of polyketide synthases.” *Annual Review Biochemistry,* 68, 219–253.

Khosla C., Tang, Y., Chen A.Y., Schnarr N.A. and Cane D.E. (2007) “Structure and Mechanism of the 6-Deoxyerythronolide B Synthase.” *Annu. Rev. Biochem*., 76, 195–221.

Kitttendorf J.D. and Sherman D.H. (2006) “Developing tools for engineering hybrid polyketide synthetic pathways.” *Current Opinion in Biochemistry*, 17, 519-605.

Kohli R.M. and Walsh C.T. (2003) “Enzymology of acyl chain macrocyclization in natural product biosynthesis.” *Chemical Communication,* 7(3): 297-307.

Kumar P., Koppisch A.T., Cane D.E., Khosla C. (2003) “Enhancing the modularity of the modular polyketide synthases: transacylation in modular polyketide synthases catalyzed by malonyl-CoA: ACP transacylase.” *Journal of American Chemical Society*,125, 14307-14312.

Lancaster DE, McNeill LA, McDonough MA, Aplin RT, Hewitson KS, Pugh CW, Ratcliffe PJ, Schofield CJ (2004) “Disruption of dimerization and substrate phosphorylation inhibit factor inhibiting hypoxiainducible factor (FIH) activity.” *Biochemical Journal*, 383:429-437.

Laskowski R.A., Macarthur M.W., Moss D.S., Thornton J.M. (1993) "Procheck - a Program to Check the Stereochemical Quality of Protein Structures." *J Appl Crystallogr*, 26, 283-91.

Lautru S. and Challis G. L. (2004) "Substrate recognition by nonribosomal peptide synthetase Multi-enzymes." *Microbiology*, 150**,** 1629-1636.

Lebel M. (1988) "Ciprofloxacin: Chemistry, Mechanism of Action, Resistance, Antimicrobial Spectrum, Pharmacokinetics, Clinical Trials, and Adverse Reactions. Pharmacotherapy." *The Journal of Human Pharmacology and Drug Therapy,* 8, 3-30.

Lee M.J., Kim H.B., Yoon Y.J., Han K., & Kim E.S. (2013) "Identification of a cyclfosporine-specific P450 hydroxylase gene through targeted cytochrome P450 complement (CYPome) disruption in *Sebekia benihana*." *Appl Environ Microbiol*, 79(7): 2253-2262.

Lloyd E. (2009) "Fleming, Florey, & Chain: The Discovery and Development of Penicillin." available from: http://www.brighthub.com/science/medical/articles/ 12679. aspx [accessed 23/08/2014].

Lloyd M.D., Merritt K.D., Lee, V., Sewell T.J., Wha-Son B., Baldwin J.E., Nicholson N.H. (1999) "Product-substrate engineering by bacteria: Studies on clavaminate synthase, a trifunctional dioxygenase." *Tetrahedron*, 55(33): 10201-10220.

Lovell S.C., Davis I.W., Arendall W.B., de Bakker P.I., Word J.M., Prisant M.G., et al. (2003) "Structure validation by Calpha geometry: phi, psi and Cbeta deviation." *Proteins*, 50, 437-50.

Madigan M., Martinko J., Stahl D., and Clark D. (2012) "Brock Biology of Microorganisms." San Francisco, 13th Edition p795-810, Pearson Education Inc.

Marion O., Gao X., Marcus S., and Hall DG. (2009) "Synthesis and preliminary antibacterial evaluation of simplified thiomarinol analogs." *Bioorganic & Medicinal Chemistry*, 17, 1006–1017.

McDonough M.A., Kavanagh K.L., Butler D., Searls T., Oppermann U., Schofield C.J. (2005) "Structure of human phytanoyl-CoA 2-hydroxylase identifies molecular mechanisms of Refsum disease." *The Journal of Biological Chemistry*, 280, 41101-10.

McFarland J. (1907) "The nephelometer - An instrument for estimating the number of bacteria in suspensions used for calculating the opsonic index and for vaccines." *J Amer Med Assoc,* 49,1176-8.

Menzella, H.G., Reid R., Carney J.R., *et al,* (2005) "Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes." *Nature Biotechnology*, 23, 1171- 1176.

Mergulhao F.J., Kelly A.G., Monteiro G.A., Taipa M.A., Cabral J.M. (1999) "Troubleshooting in gene splicing by overlap extension: a step-wise method." *Molecular Biotechnology,* 12, 285-7.

Miller J.H. (1992) "A short course in bacterial genetics." Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.f

Mingeot-leclercq M. Glupczynski Y., Tulkens P., (1999) "Aminoglycosides: Activity and Resistance." A*ntimicrobial Agents and Chemotherapy*, 43(4): 727–737.

Moir A., Lafferty E. and Smith D.A. (1979) "Genetic analysis of spore germination mutants of *Bacillus subtilis* 168: the correlation of phenotype with map location." *Journal of General Microbiology*, 111, 165-180.

Mootz H.D., Kessler N., Linne U., Eppelmann K., Schwarzer D. and Marahiel M.A. (2002) "Decreasing the ring-size of a nonribosomal peptide antibiotic by in frame module deletion in the biosynthetic genes." *Journal of American Chemical Society*, 124, 10980-10981.

Mootz H.D., Schwarzer D. and Marahiel M.A. (2002) "Ways of Assembling Complex Natural Products on Modular Nonribosomal Peptide Synthetases" *Chembiochem.*, 3, 490-504.

Mootz H.D., Schwarzer D., Marahiel M A.(2000) "Construction of hybrid peptide synthetases by module and domain fusions." *PNAS*, 97(11): 5848–5853.

Morris G.M., Huey R., Lindstrom W., Sanner M.F., Belew R.K., Goodsell D.S., et al. (2009) "AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility." *Journal of Computational Chemistry*, 30, 2785-91.

Murphy A., Fukuda D., Song Z., *et al*, (2011) "Engineered Thiomarinol Antibiotics Active against MRSA Are Generated by Mutagenesis and Mutasynthesis of *Pseudoalteromonas* SANK 73390." *Angewandte Chemie International Edition*, 50(14): 3271–3274.

Murphy A.C., Gao S., Han L.C., Carobene S., Fukuda D., Song S., Hothersall J., Cox R.J., Crosby J., Crump M.P., Thomas C.M., Willis C.L. and Simpson T.J. (2013) "Biosynthesis of thiomarinol A and related metabolites of *Pseudoalteromonas sp*. SANK 73390." *Chemical Science*, 5, 397-402.

Narayanan Eswar B.W., Marc A. Marti-Renom M.S. Eramian M.D., Min-yi Shen, Ursula Pieper, and Sali A. (2006) "Comparative Protein Structure Modeling Using Modeller." *Curr Protoc Bioinformatics*, 5.6.

Office for National Statistic (2004) "Trend in MRSA in England and Wales: analysis of morbidity and mortality data for 1993-2002." *Health Statistic Quarterly* 21.

Olaitan A., Morand S. and Rolain JM. (2014) “Mechanisms of polymyxin resistance :acquired and intrinsic resistance in bacteria.” *Frontiers in Microbiology*, 26(5): 1-18.

Omer-bali, AM (2013) “Studies on key steps controlling biosynthesis of antibiotics thiomarinol and mupirocin.” Unpublished, PhD thesis submitted to the school of bioscience, The University of Birmingham.

Peirú S., Gramajo H. and Menzella H. (2010) “Recombinant approaches to large polyketide molecules as potential drugs.” *Drug Discovery Today*: Technologies, 7(2): e105–e113.

Piddock L.J.V. (2012) “The crisis of no new antibiotics—what is the way forward?” *The Lancet Infectious Disease*, 2, 249–53.

Rausch C., Hoof I., Weber T., Wohlleben W., & Huson D. H. (2007) “Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution.” *BMC Evol Biol*, 7, 78.

Reeves C.D., Murli S., Ashley G.W., Piagentini M., Hutchinson C.R., McDaniel R. (2001) “Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations.” *Biochemistry*, 40,15464-15470.

Rix U., Fischer C., Remsing L.L., & Rohr J. (2002) “Modification of post-PKS tailoring steps through combinatorial biosynthesis.” *Nat Prod Rep*., 19(5): 542-580.

Robbel L., and Marahiel M., (2010) “Daptomycin, a bacterial lipopeptide synthesized by a nonribosomal machinery.” *Journal of Biological Chemistry,* 285(36): 27501–27508.

Robert, X. and Gouet, P. (2014) “Deciphering key features in protein structures with the new ENDscript server”. Nucl. Acids Res. 42(W1), W320-W324 - doi: 10.1093/nar/gku316 (freely accessible online).

Sanger F., Nicklen S. and Coulson A.R. (1977) “DNA sequencing with chain terminating inhibitors.” *Proc*. *Natl*. *Acad*. *Sci*. *U.S.A.,* 74, 5463-5467.

Sanner M.F. (1999) “Python: a programming language for software integration and development.” *Journal of Molecular Graphics & Modelling*, 17, 57-61.

Schwarzer D., Finking R. and Marahiel M.A. (2003) “Nonribosomal peptides: from genes to products.” *Natural Product. Report*, 20, 275–287.

Shields J.A., Rahman A.S., Arthur C.J., Crosby J., Hothersall J., Simpson T.J., & Thomas C.M. (2010) “Phosphopantetheinylation and specificity of acyl carrier proteins in the mupirocin biosynthetic cluster.” *Chembiochem*, 11(2): 248-255.

Shiozawa H., Kagasaki T., Kinoshita T., Haruyama H., Domon H., Utsui Y., Kodama K. and Takahashi S. (1993) "Thiomarinol, a new hybrid antimicrobial antibiotic produced by a marine bacterium." *The Journal of Antibiotics*, 46(12): 1834-1842.

Shiozawa H., Kagasaki T., Torikata A., Tanaka N., Fujimoto K., Hata T., Furukawat Y. and Takahashi S. (1995) "Thiomarinols B and C, New Antimicrobial Antibiotics Produced by a Marine Bacterium." *The Journal of Antibiotics*, 48(8): 907-909.

Shiozawa H., Shimada A. and Takahashi S. (1997) "Thiomarinols D, E, F and G, NewHybrid Antimicrobial Antibiotics Produced by a Marine Bacterium; Isolation, Structure, and Antimicrobial Activity." *The Journal of Antibiotics*, 50(5): 449-452.

Silvian L.F., Wang J. and Steitz T.A. (1999) "Insights into Editing from an lLe-tRNA Synthetase Structure with tRNAILe and Mupirocin" *Science*, 285, 1074-1077.

Simon R., Priefer V. and Puhler A. (1983) "A broad host range mobilisation system for *in vivo* genetic engineering: transposon mutagenesis in Gram-negative bacteria." In: *'Biotechnology*, 784-791.

Sir Alexander Fleming (1945) "Nobel Lecture: Penicillin", Nobelprize.org. Nobel Media AB 2014. <http://www.nobelprize.org/nobel_prizes/medicine/laureates/1945/ fleming-lecture.html>, Accessed 20th April, 2016.

Soding J. (2005) "Protein homology detection by HMM-HMM comparison." *Bioinformatics*, 21,951-60.

Staunton J. and Weissman K.J. (2001) "Polyketide biosynthesis: a millennium review." *Natual Product Report*, 18, 380–416.

Stynen B., Tourno H., Tavernier J., and Van Dijck P., (2012) "Mammalian Split-Luciferase System the Yeast Two-Hybrid System to the Protein-Protein Interaction Studies: from Methods for In Vivo Diversity in Genetic." *Microbiol. Mol. Biol. Rev*., 76(2):331.

Sun J., Deng Z., and Yan A. (2014) "Bacterial multidrug efflux pumps: Mechanisms, physiology and pharmacological exploitations." *Biochemical and Biophysical Research Communications,* 453, 254–267.

Sy-Cordero A.A., Pearce C.J., & Oberlies N.H. (2012) "Revisiting the enniatins: a review of their isolation, biosynthesis, structure determination and biological activities." *J Antibiot* (Tokyo), 65(11): 541-549.

Tang L., and McDaniel R. (2001) "Construction of desosamine containing polyketide libraries using a glycosyltransferase with broad substrate specificity." *Chemistry & Biology,* 8, 547-555.

The Noble Lecture (1945), “Sir Alexander Fleming - Nobel Lecture: Penicillin” available from:http://www.nobelprize.org/nobel_prizes/medicine/laureates/1945/ fleming-lecture. html _ [accessed 25/08/2014]

The World Health Organization WHO (2011) “Antimicrobial resistance: Fact sheet”, World Health day, 7th of April 2011.

Thirlway J., Lewis R., Nunns L., Al Nakeeb M., Styles M., Struck A.W., Smith C.P., and Micklefield J. (2012) “Introduction of a non-natural amino acid into a Non-ribosomal peptide antibiotic by modification of adenylation domain specificity.” *Angewandte Chemie International Edition*, 51, 7181 –7184.

Thomas C.M., Hothersall J., Willis C.L. and Simpson T.J. (2010) “Resistance to and synthesis of the antibiotic mupirocin.” *Nature Reviews Microbiology*, 8, 281- 289

Todar K., (2008) “Bacterial Resistance to Antibiotics.” Todar’s Online Textbook of Bacteriology. www.textbookofbacteriology.net.

Torok E., Moran E. and Cooke F. (2009) “Oxford Handbook of Infectious Diseases and Microbiology.” Oxford University Press, ISBN-13: 9780198569251.

Vaara M. (1992) “The outer membrane as the penetration barrier against mupirocin in Gram-negative enteric bacteria.” *Journal of Antimicrobial Chemotherapy*, 29(2): 221-2.

Vallejo A.N., Pogulis R.J., Pease L.R. (1994) “In vitro synthesis of novel genes mutagenesis and recombination by PCR.” *PCR Methods and Applications*, 4, S123-30.

Wageningen M.A.A, Kirkpatrick P.N., Williams D.H., Harris B.R., Kershaw J.K., Lennard N.J., Jones M., Jones S.J.M. and Solenberg P.J. (1998) “Sequence analysis of genes involved in the biosynthesis a vancomycin group antibiotic.” *Chemistry & Biology*, 5(3): 155-162.

Waksman S. Selman (1947) “What is an antibiotic or antibiotic substance” *Mycologia*, 39(5): 565-569.

Walsh C.T., Chen H., Keating T.A., Hubbard B.K., Losey H.C., Luo L., Marshall C.G., Miller D.A. and Patel H.M. (2001) “Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines.” *Current Opinion in Chemical Biology*, 5, 525–534.

Wang P., Yu Z., Li B., Cai X., Zeng Z., Chen X., & Wang X. (2015) “Development of an efficient conjugation-based genetic manipulation system for *Pseudoalteromonas*.” *Microb Cell Fact*, 14, 11.

Watanabe A. and Ebizuka Y. (2004) “Unprecedented Mechanism of Chain Length Determination in Fungal Aromatic Polyketide Synthases.” *Chemistry & Biology*,11, 1101–1106.

Webber, M.A. and Piddock, L.J.V. (2003) “The importance of efflux pumps in bacterial antibiotic resistance.” *Journal of Antimicrobial Chemotherapy*, 51, 9–11.

Weber G., Schorgendorfer K., Schneiderscherzer E., Leitner E. (1994) “The peptide synthetase catalyzing cyclosporine production in tolypocladium-niveum is encoded by a giant 45.8-kilobase open reading frame.” *Current Genetics,* 26,120–125.

Wehrli W. (1983) “Rifampin: mechanisms of action and resistance.” *Rev Infect Dis.* S407-11. *Clinical Infectious Diseases*, 5 (Supplement 3): S407-S411.

Weissman K.J. and Leadlay P.F. (2005) “Combinatorial biosynthesis of reduced polyketides.” *Nature Reviews Microbiology*, 3: 925- 933.

Whatling C., Hodgson J.E., Burnham M.K.R., Clarke N.J., Franklin C.H. and Thomas C.M. (1995) “Identification of a 60kb region of the chromosome of *Pseudomonas fluorescens* NCIB 10586 required for the biosynthesis of pseudomonic acid (mupirocin).” *Microbiology*, 141, 973-982.

Widderich N., Hoppner A., Pittelkow M., Heider J., Smits S.H.J., Bremer E. (2014) “Biochemical Properties of Ectoine Hydroxylases from Extremophiles and Their Wider Taxonomic Distribution among Microorganisms.” *PloS one*,14,9.

Winn M., Fyans J.K., Zhuo Y. and Micklefield J. (2016) “Recent advances in engineering nonribosomal peptide assembly lines.” *Natural Product Reports,* 33, 317-347.

Winter J.M., Chiou G. Bothwell I.R., Xu W., Garg N.K., Luo M., and Tang Y. (2013) “Expanding the structural diversity of polyketides by exploring the cofactor tolerance of an inline methyltransferase domain.” *Organic Letters*, 15(14): 3774–3777.

Wu L.F., Meng S., Tang GL. (2016) “Ferrous iron and α-ketoglutarate-dependent dioxygenases in the biosynthesis of microbial natural products.” *Biochimica Biophysica Acta*,1864(5):453-70.

Wuite J., Davies B.I., Go M., Lambers J., Jackson D. and Mellows G. (1983) “Pseudomonic Acid: A New Topical Antimicrobial Agent.” *The Lancet*, 322(8346): 394.

Xiang H., Tschirret-Guth, R.A., & Ortiz De Montellano P.R. (2000) “An A245T mutation conveys on cytochrome P450eryF the ability to oxidize alternative substrates.” *J Biol Chem,* 275(46): 35999-36006.

Xue Y., Zhao L., Liu H.W. and Sherman D. H. (1998) “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, 12111-12116.

Yan W., Song H., Song F., Guo Y., Wu C. H., Sae Her A., Zhang Y. J. (2015) “Endoperoxide formation by an alpha-ketoglutarate-dependent mononuclear non-haem iron enzyme.” *Nature*, 527(7579), 539-543.

Yanagisawa T. & Kawakami M. (2003) “How does *Pseudomonas fluorescens* avoid suicide from its antibiotic pseudomonic acid? Evidence for two evolutionarily distinct isoleucyl-tRNA synthetases conferring self-defense.” *J. Biol. Chem.* 278, 25887–25894.

Yang J.A., Park D.W., Sohn J.W., Yang I.S, Kim K.H. and Kim M.J. (2006) “Molecular analysis of isoleucyl-trna synthetase mutations in clinical isolates of methicillin-resistant *Staphylococcus aureus* with low-level mupirocin resistance.” *Journal of Korean Medical Science, 21, 827-32.*

Yun H.J., Lee S.W., Yoon G.M., Kim S.Y., Choi S., Lee Y.S., Choi E.C. and Kim S. (2003) “Prevalence and mechanisms of low- and high-level mupirocin resistance in *staphylococci* isolated from a Korean hospital.” *Journal of Antimicrobial Chemotherapy,51*, 619–623.

Zhang X., Chen Z., Li M., Wen Y., Song Y., Li J. (2006) “Construction of ivermectin producer by domain swaps of avermectin polyketide synthase in *Streptomyces avermitilis*.” *Bioresource Technology*, 101, 9228–9235.