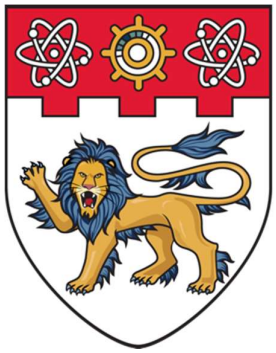


NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

SYNTHESIS OF PEPTIDE ANTIBIOTICS FOR MYCOBACTERIUM TUBERCULOSIS

LIN HONG XUAN

SCHOOL OF CHEMISTRY, CHEMICAL ENGINEERING AND BIOTECHNOLOGY

2023

SYNTHESIS OF PEPTIDE ANTIBIOTICS FOR MYCOBACTERIUM TUBERCULOSIS

LIN HONG XUAN

SCHOOL OF CHEMISTRY, CHEMICAL ENGINEERING AND BIOTECHNOLOGY

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

2023

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

14/08/2023

. . . . . .
Date

. . . . . .
[Lin Hong Xuan]

*please sign on the water mark area.

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

14/08/2023

. . . . . .
Date

. . . . . .
[Roderick Bates]

*please sign on the water mark area.

Authorship Attribution Statement

This thesis **does not** contain any materials from papers published in peer-reviewed journals or from papers accepted at conferences in which I am listed as an author.

14/08/2023

. . . . . .
Date

. . . . . .
[Lin Hong Xuan]

*please sign on the water mark area.

Acknowledgements

First of all, I would like to express my gratitude to my supervisor, Professor Roderick W Bates, for guiding me throughout these four years of my PhD journey. I wanted to thank him for his kind guidance in everything from coming up with new ideas on what to synthesise to how to resolve a difficult synthetic step. Without him, I might not be able to orientate myself in the lab and in our research.

I am also eternally grateful for the support from all my lab colleagues. To all my seniors Ivan, Pat, Weiting, Annabelle, Loi, Yong, Anders, Sherilyn and Sopan, your demonstrations, guidance and even lecturing were invaluable to helping me get used to a research environment and becoming proficient in various techniques in a lab. I would like to thank Pat and Sopan especially for their honest comments and advice in helping me complete this thesis. To all the undergraduates I must thank you all for your companionship and making this lab a batter place.

I also want to thank everyone in Professor Grüber's group in our collaboration and for helping me out with everything related in biology. Special thanks to Dr Priya Ragunatham, who helped me with all the assays, and Dr Harikrishore, who did the computational modelling for my compounds.

To the staff of CCEB, I would like to thank Goh Ee Ling and Zhu Wen Wei for helping me out in troubleshooting problems I encountered with the NMR and MS machines.

Last but not least, I would like to thank my family, my cousins, my partner and my friends for supporting me emotionally all the way, especially during the darkest days of the COVID pandemic.

Table of Contents

Abstract

Many drug candidates against TB fail in an early stage during primary screening. While they were discovered to be effective enzyme inhibitors, this did not always translate to effective inhibitors of bacterial growth. The primary reason for this is that the thick, hydrophobic cell wall of *M. tuberculosis* (*Mtb*) acts as a barrier to the entry of drug molecules. Such was the case for EpMF2, a hit compound that was discovered via virtual screening as an inhibitor of the ε subunit of the F_1F_O ATP synthase. Analogues of EpMF2 have been synthesized that were effective inhibitors of *M. smegmatis* F_1F_O ATP synthase in IMVs but all failed to inhibit growth of live *M. smegmatis.* To get around this problem a prodrug strategy was devised using cholesterol as a carrier. *Mtb* has an affinity to cholesterol and tends to incorporate cholesterol into itself during infection of macrophages. The concept behind the prodrug idea is to induce the bacterium to ingest the cholesterol-EpMF2 conjugate, where EpMF2 would be released in the cytosol after cleavage of the molecule and kill it. Several cholesterol-EpMF2 conjugates were synthesized and underwent assay testing, and some were able to inhibit growth of *M. smegmatis.* This proved the viability of the cholesterol prodrug concept in principle as a method to penetrate the mycobacterial bacterial cell wall, but more work needs to be done to optimize this delivery method, particularly the poor aqueous solubility of the conjugates.

Chapter 1 will provide a brief review of the various types of prodrug chemical structures and activation methods, with a focus on real life examples currently being used or in development. Chapter 2 describes the discovery, synthesis, and assay tests of EpMF2 and its analogues, and an investigation to its structural-activity relationship. Chapter 3 describes the conception of the cholesterol prodrug idea, the synthesis of EpMF2-cholesterol conjugate compounds, and its assay tests. The final chapter will briefly touch upon future work that can be done based on the findings in Chapter 3.

List of Abbreviations

Ac	acetyl
ADC	antibody-drug conjugate
ADEPT	antibody directed enzyme prodrug therapy
ADP	adenosine diphosphate
AMP	antimicrobial polymer
atm	atmospheric pressure
ATP	adenosine triphosphate
BCA	bicinchonic acid
BChE	butylrylcholinesterase
BDQ	bedaquiline
Bn	benzyl
Boc	tert-butyloxycarbonyl
Cbz	benzyloxycarbonyl
CES	carboxylesterase
cLogP	calculated log (partition coefficient)
COVID 19	coronavirus disease 2019
CuAAC	Copper (I) catalysed azide-alkyne cycloaddition
CYP	cytochrome P450
DCC	dicyclohexylcarbodiimide
DCM	dichloromethane
DDQ	2,3-dichloro-5,6-dicyano-1,4-benzoquinone
DIAD	diisopropylazodicarboxylate

DIC	diisopropylcarbodiimide
DIPEA	*N,N*-diisopropylethylamine
DKP	diketopiperazine
DMAP	4-dimethylaminopyridine
DME	dimethoxyethane
DMF	*N,N*-dimethylformaide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
EDCI	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride
EDTA	ethylenediaminetetraacetic acid
ESI-MS	electrospray-ionisation mass spectroscopy
Et	ethyl
fDUMP	2'-deoxy-5-fluorouridine 5'-monophosphate
Fmoc	9-fluorenylmethyloxycarbonyl
GDEPT	gene directed enzyme prodrug therapy
GSH	glutathione
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b] pyridinium 3-oxide hexafluorophosphate
HIV	human immunodeficiency virus
HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxybenzotriazole
HRMS	High-resolution mass spectroscopy
HSQC	Heteronuclear single quantum coherence
IC_{50}	half maximal inhibitory concentration

IMV	inverted membrane vesicles
iPr	isopropyl
KHMDS	potassium hexamethyldisilazide
mAb	mononuclear antibody
MDR-TB	multidrug resistant tuberculosis
Me	methyl
MIC	minimum inhibitory concentration
MOPS	3-morpholinopropane-1-sulfonic acid
Mp	melting point
Mtb	*Mycobacterium tuberculosis*
NADH	nicotinamide adenine dinucleotide
nBu	normal-butyl
NHS	N-hydroxysuccinimide
NMR	nuclear magnetic spectroscopy
NSAID	non-steroidal anti-inflammatory drug
OD_{600}	optical density at 600nm
PEG	poly(ethylene glycol)
Pfp	pentafluorophenyl
Ph	phenyl
Pi	phosphate
psi	pounds-per-square inch
RNA	ribonucleic acid
TACO	tryptophan-aspartate containing coat protein
TB	tuberculosis

tBu	tert-butyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
tPSA	total polar surface area
TPZ	tirapazamine
Troc	2,2,2-trichloroethoxycarbonyl
TS	thymidylate synthase
WHO	World Health Organization
Xantphos	(9,9-Dimethyl-9*H*-xanthene-4,5 diyl) bis(diphenylphosphane)
XDR-TB	extensively drug resistant tuberculosis

CHAPTER 1

PRODRUGS: A REVIEW

1 Prodrugs: A Review

1.1 Introduction

According to the definition provided by the IUPAC, a prodrug is

"*any compound that undergoes biotransformation before exhibiting its pharmaceutical effects*"[1]

The prodrug molecule is not the active drug, but a precursor. Enzymatic reactions in the body then chemically modify it *in vivo*, where the actual active drug is released. One example of a prodrug is heroin **(1-1)**, or diacetylmorphine, which converts to morphine **(1-2)** in vivo. On the heroin molecule, the 2 polar hydroxy groups of morphine were masked by acetyl groups, which makes heroin much more lipophilic compared to morphine. As such, heroin more readily crosses the blood brain barrier, and causes a stronger rush as experienced by users. Heroin itself has a weak affinity to μ-opioid receptors[2], and its effects are mostly due to conversion to morphine in the brain rather than due to heroin, thus making heroin a prodrug.

heroin **(1-1)** (diacetylmorphine) morphine **(1-2)**

Figure 1.1:Heroin **(1-1)** is a prodrug of morphine **(1-2)**, which is responsible for most of its effects.

Prodrugs can be classified in many ways. One method, as proposed by Wu[3], divides prodrugs by where they are converted to the active form: intracellular (Type I) or extracellular (Type II). Subtypes then further specify the location of activation (IA, IB, IIA, IIB, IIC). Other methods include chemical classification by presence of various functional groups of the prodrug.

Why use prodrugs rather than just directly administering the drug directly? Prodrugs improve a drug's pharmacokinetics by modifying its ADME properties (absorption, distribution,

metabolism, and excretion) in a way that maximises the amount of drug reaching its intended target. Prodrugs chemically modify the drug to achieve the above goal. The advantages of prodrugs are that they have better bioavailability, and are more selective, thereby reducing side effects. Because of its advantages, the pharmaceutical industry has shown increasing interest in using prodrugs to improve drug delivery. 10% of all drugs marketed worldwide are prodrugs, and 30 prodrugs have been approved by the Food and Drug Administration from the years 2008-2017[4].

1.2 Advantages of prodrugs

1.2.1 Prodrugs to improve bioavailability

According to Lipinski's observations[5], drugs that are effective when taken orally need to be moderately non-polar in order to effectively permeate lipid bilayer membranes, but also sufficiently water soluble to allow its transport in the aqueous medium found in the human circulatory system. This is elaborated by Ghose' et al [6]who suggested that a compound with a cLogP value between -0.4 to 5.6 is ideal.

When a drug is unable to reach its desired target due to the above issues, it is possible to attach a moiety to alter its cLogP value to a more suitable value.

1.2.1.1 Increasing lipophilicity

When a drug is taken orally, it needs to be absorbed through the digestive tract. All oral drugs must pass through the intestinal epithelial cell wall for absorption before they can enter the bloodstream. There are multiple possible modes of transport. The most common form of passage through the intestinal cells is passive diffusion through the cells (path A), or the transcellular pathway. This requires that the drug be permeable to the lipid bilayer on the surface of such cells, and therefore requires a certain degree of hydrophobicity. Another possible form of transport is between the cells, or the paracellular pathway (path B). Hydrophilic molecules can be transported through this channel, but the presence of tight

junctions between cells severely hampers the transport of anything but the lightest molecules of molecular weight below 200 daltons[7]. Aside from passive diffusion, active transport also occurs through membrane proteins (path C), however other enzymes are also involved in transport in the opposite direction out. (Path D) The transport of xenobiotics and other drugs via active transporters remains an active area of study, but while many active transporters were discovered, there is no transporter effective for all drugs[8].

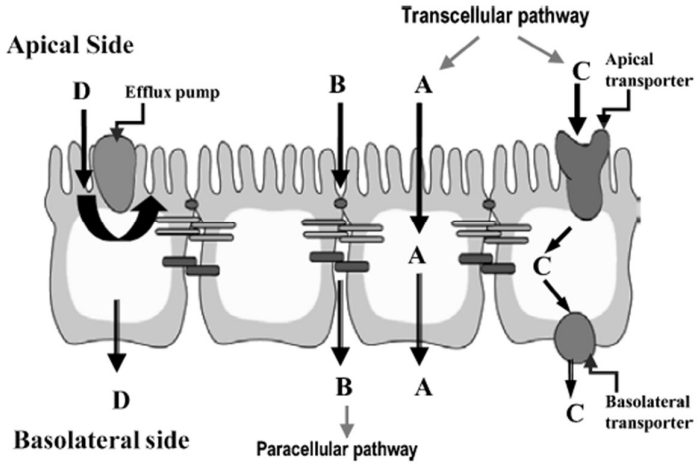
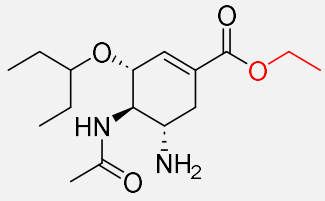
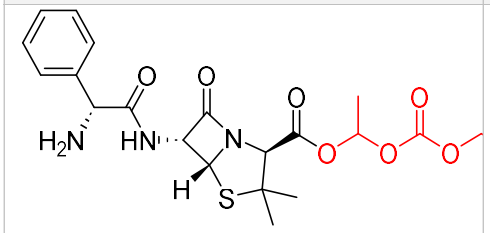
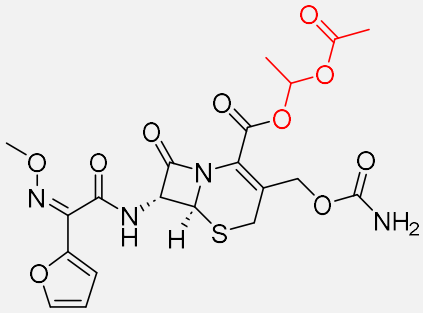


Figure 1.2: An oral drug needs to bypasses the intestinal cell wall before entering circulation. There are multiple possible labelled A to D. A: passive diffusion through epithelial cells. B: Transport through cell junctions. C and D: active transport by membrane proteins.routes. Adapted from Siahaan et al[9].

A drug that is too polar will not be able to diffuse through the intestinal epithelial layer. On the other hand, non-polar drugs, especially those of low to medium molecular mass, are less dependent on active transport into the bloodstream and are mostly absorbed by diffusion. A common method to from prodrugs from polar molecules is to mask polar functional groups such as -OH, -NH_2 and -COOH with non-polar moieties through ester or carbamate groups. The linker groups should ideally be easily hydrolysed to release the active drug in the bloodstream.

Esterification of alcohol and carboxylic acid groups is in fact, one of the most common prodrug strategies. A review of ester prodrugs[10] identified over 30 ester prodrugs where oral bioavailability was improved compared to the unmodified alcohol.

Name and structure of ester prodrug (Red coloured moieties indicate protecting group that is cleaved to form the active drug)	**Type of drug**	**Bioavailability of prodrug**	**Bioavailability of parent alcohol**
 oseltamivir **(1-3)**	Anti-flu Neuraminidase inhibitor	>80	79
 bacampicillin **(1-4)**	Broad spectrum beta lactam antibiotic	83-89	32-55
 cefuroxime axetil **(1-5)**	Cephalosporin antibiotic Broad spectrum beta lactam antibiotic	32-43	0

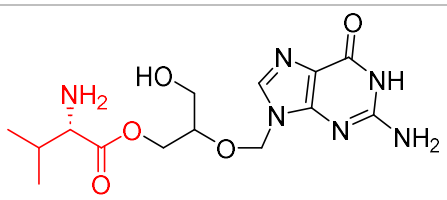
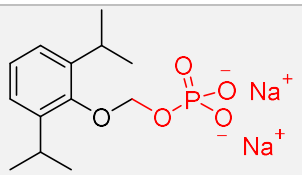
 valganciclovir **(1-6)**	Antiviral Viral DNA polymerase inhibitor	61	6

Table 1.3:A few selected examples of drugs that have increased oral bioavailability orally and their values[5]

Esters are not the only possible functional groups, and many possible functional groups are used in prodrugs. This will be covered in greater detail below.

1.2.1.2 Increasing hydrophilicity

On the other hand, if a drug is too lipophilic, it is difficult to create injectable formulations since they are usually administered as an aqueous solution. A highly lipophilic compound also will have poor distribution kinetics by the circulatory system due to its limited solubility in fat tissue. A common way to increase the aqueous solubility of a drug is to use its salt form, but another solution is to attach ionizable or polar moieties attached to improve its aqueous solubility.

Prodrug (Red coloured moieties indicate protecting group that is cleaved to form the active drug)	**Type of drug**	**Solubility of prodrug**	**Solubility of parent drug**
 fospropofol disodium **(1-7)**	anaesthetic	500 mg mL^{-1}	0.15 mg mL^{-1}

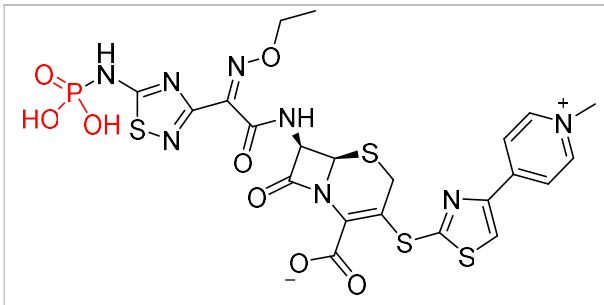
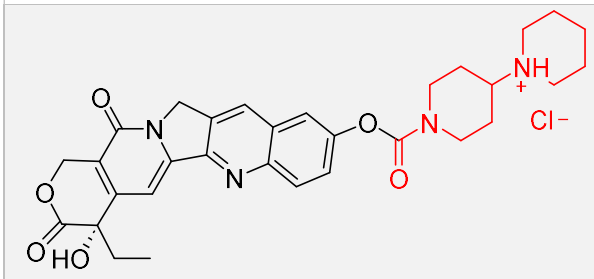
 ceftaroline fosamil **(1-8)**	Cephalosporin Broad spectrum beta lactam antibiotic	>100mg mL^{-1}	2.3 mg mL^{-1}
 irinotecan hydrochloride **(1-9)**	Chemotherapeutic drug Topoisomerase inhibitor	>20 mg mL^{-1}	2-3 µg mL^{-1}

Table 1.4::A few selected examples of drugs that have improved aqueous solubility upon addition of polar moieties[4]

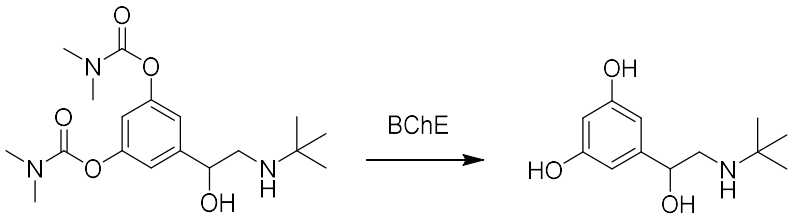
The above examples are a few drugs that are administered intravenously. The addition of polar promoieties significantly increases the aqueous solubility of these prodrugs compared to the parent drugs. This not only allows administration of higher doses by injection, but also simplifies drug formulation by obviating the need to use emulsifiers or other solubilizing additives.

1.2.2 Prodrugs to increase duration of action

A sustained drug release formulation releases a drug not all at once but gradually over a relatively long period of time. This is especially important for fast acting drugs with a short half-life in the body. Another advantage is that the frequency of dosage can be reduced, easing the burden on patients. A common way to make sustained release formulations is to encapsulate a drug in a non-soluble but porous matrix, typically made from a biodegradable polymer. The pores allow the drug to be released in a slow but controlled manner.

However, it is also possible to use a prodrug strategy to create a sustained released formulation of a drug. The prodrug where the bonds linking the promoiety and the drug are rationally chosen so that the promoiety-drug form is converted slowly by enzymes in the body, with the effect that the active drug is leached slowly in the body.

One example of this is bambuterol **(1-10)**. It is a prodrug of terbutaline **(1-11)**, a β adrenergic receptor agonist. Terbutaline is used to treat asthma symptoms. When taken orally, terbutaline has a relatively short half-life of 3.4 hours, which necessitates 3 times dosage daily[11].



Scheme 1.5: Bambuterol **(1-10)** is a prodrug of terbutaline **(1-11)**

In bambuterol, the phenolic hydroxyl groups on terbutaline are modified into carbamate groups. Terbutaline is released when the carbamate moieties are cleaved by butyrylcholinesterase (BChE), an enzyme that normally cleaves the ester bond in neurotransmitter-based choline esters. However, bambuterol also reversibly inhibits cholinesterases. When bambuterol is processed by butyrylcholinesterase, it transfers its carbamoyl groups to the serine residue active site, and the carbamate group on the serine is only hydrolysed slowly. Bambuterol therefore supresses its own hydrolysis, and the end result is that terbutaline is released slowly over an extended period of time.[12] Bambuterol need only be administered only once a day because of this, simplifying the treatment regimen and reducing side effects for patients taking this prodrug over terbutaline.[13]

1.2.3 Prodrugs to increase targeting specificity

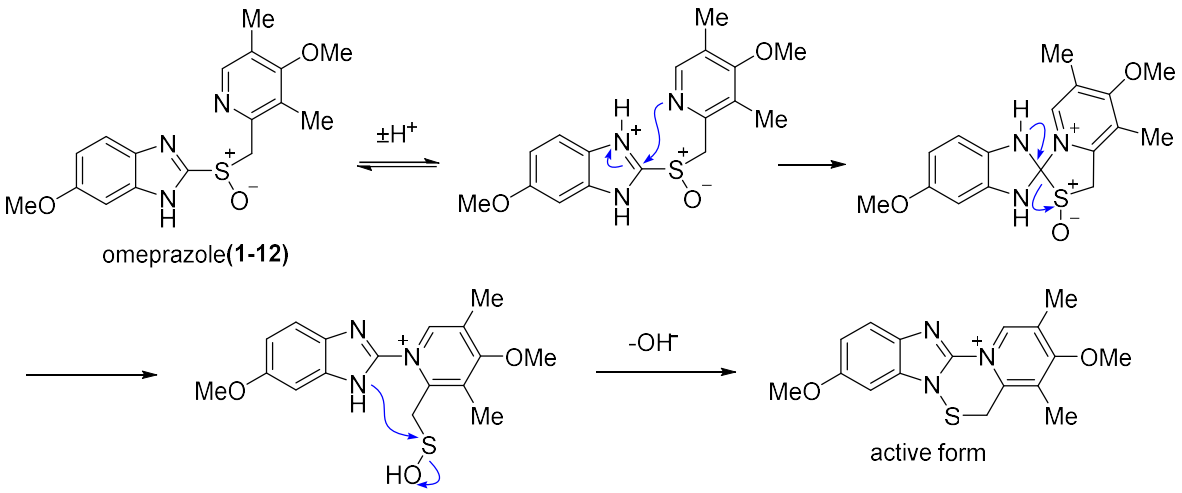
Prodrugs must first be metabolized by enzymes in the body to release the drug and exert their desired effects. Depending on the nature of the promoiety and the nature of the chemical bond

between it and the drug, prodrugs can be designed so that it can only be activated by specific enzymes located in only a few sites in the body. For example, many anticancer drugs (1.3.3) have reducible nitro groups that are stable under physiological conditions, but are reduced in hypoxic cancer cells to release the active drug. Thus prodrugs can increase the specificity of a drug and decrease the incidence of side effects.

1.2.3.1 Specificity from pH change

The pH environment of certain organs, particularly the gastrointestinal tract, can vary from highly acidic in the stomach, to slightly alkaline in the small intestine. Prodrugs targeting this area can preferentially act on certain organs if they are chemically converted into the active drug trigged by changes in pH.

Omeprazole **(1-12)**, used to treat gastroesophageal reflux disease, is one of them. It specifically and irreversibly inhibits H^+/K^+ ATPase found in parietal cells, a type of epithelial cells in the stomach that secrete stomach acid. Omeprazole itself has no bioactivity, but in the highly acidic environment of the canaliculus of the parietal cells, the benzimidazole is protonated, and this triggers a cascade reaction (as shown in below Fig1.6) where omeprazole molecule rearranges to form the sulfenamide, which is the active form.[14] This reacts with the sulfhydryl groups on the cysteine residues on the α-subunit of the ATPase, forming a disulfide bond. The α-subunit catalyses H^+ formation, and formation of the disulfide covalent bond irreversibly inhibits the enzyme, suppressing acid formation.[15]

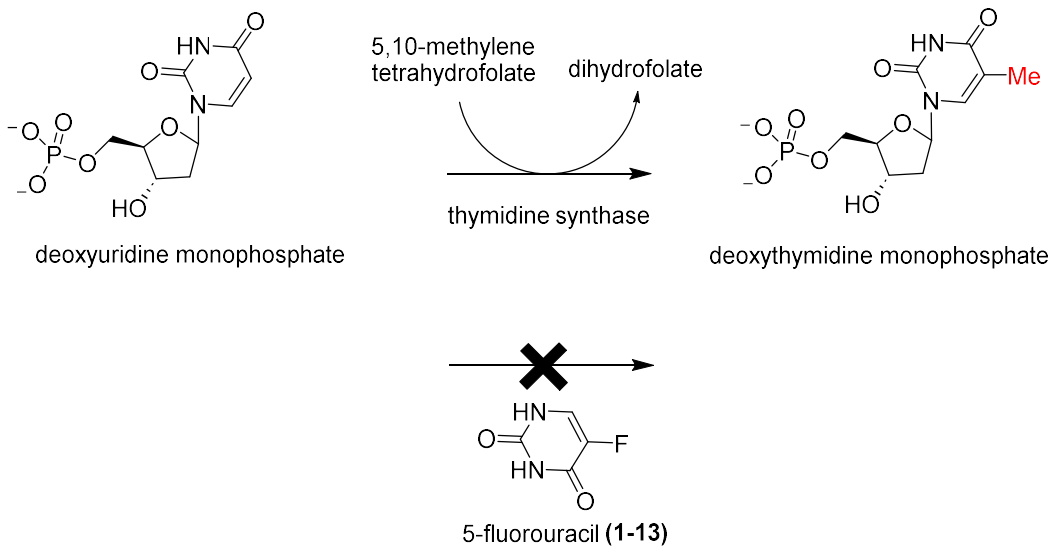


Scheme 1.6: Mechanism of the acidic activation of omeprazole

1.2.3.2 Specificity from increased enzyme expression in target cells

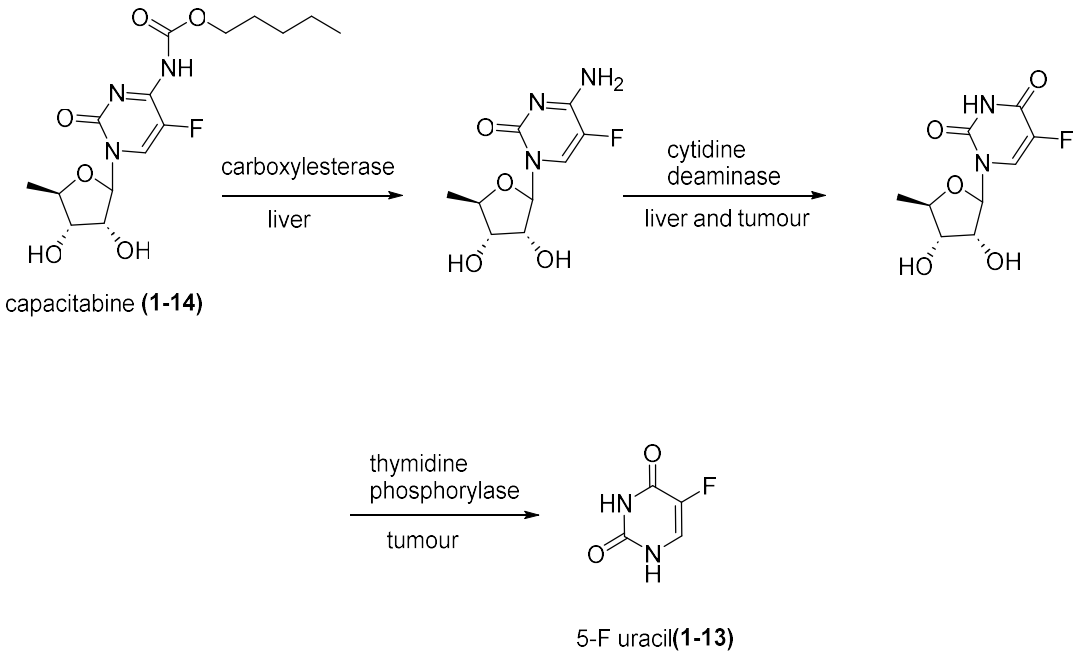
Thymidylate synthase (TS) is an enzyme that catalyses the formation of deoxythymidine monophosphate (dTMP) from deoxyuridine monophosphate (dUMP). dTMP is essential in synthesizing DNA and the processes of DNA repair and replication are dependent on the nucleotides synthesised from TS activity. Bacterial, yeast and mammalian cells that are deprived of dTMP cannot survive and will die[16].

Many anticancer drugs take advantage of the unique biophysical environment of cancer cells to specifically target them. Cancer cells proliferate far more rapidly than non-cancer cells and are therefore much more vulnerable to the effects of TS inhibition. 5-Fluorouracil (5-FU) **(1-13)** is one such inhibitor. It is common administered intravenously to treat colon, oesophageal, pancreatic and stomach cancer, or applied topically for skin cancers. 5-FU is not administered orally because of the first pass effect: it is extensively degraded by dihydropyrimidine dehydrogenase expressed in the liver and intestines. [17]



Scheme 1.7:Action of thymidine synthase (TS), an enzyme overexpressed in cancer cells and a target for anticancer drugs

A prodrug of 5-fluorouracil, capecitabine **(1-14)**, has been developed that can be administered orally. It is metabolized into the active drug in 3 steps, each in a specific organ/tissue to ensure the prodrug is released preferentially in tumour cells. Firstly, liver carboxylesterases hydrolyse the carbamate group of capecitabine and expose the amine group, which is then oxidised to the carbonyl group by cytidine deaminase in the liver or in the tumour cell. Finally, the sugar group is removed by thymidine phosphorylase, which is overexpressed in tumour cells, to release the active drug in cancerous cells. The bioavailability of capecitabine is almost 100%[18], and its concentration in cancerous colon cells is 3.2 times higher concentration than surrounding healthy cells, although in liver cells it is a more modest 1.4 times. [19]



capacitabine **(1-14)**

5-F uracil**(1-13)**

Scheme 1.8: Enzymatic reactions on capecitabine **(1-14)** to form the active drug 5-fluorouracil **(1-13)**.Adapted from Rautio *et al.*[4]

1.2.3.3 Antibodies as the prodrug moiety: ultimate specificity

Nevertheless, there are still limitations in the above methods when it comes to directed drug activation. Enzymes that are localised in certain tissues are few in number and also very substrate specific; they cannot tolerant much variation in prodrug structures, which limits prodrug design. The desire to create a prodrug moiety that can be molecularly fine-tuned to target any chosen part of the human body has led to the development of the use of antibodies in prodrugs. These drug therapies have been an active topic of research particular in the treatment of cancers.

Antibodies have extremely high specificity, due to the presence of hypervariable regions in the antigen binding region, where its amino acid sequence, and the resulting shape, is highly different for each individual immune B/T cell created.

One strategy to create highly specific prodrugs is to use the antibody itself as the carrier. The drug is linked directly to the antibody, forming an antibody-drug conjugate (ADC). Once the ADC is administered into the body, it selectively binds to target cells through antigen binding.

The ADC is then internalised, and the linker selectively cleaved in the tumour cells by degradation by lysozymes, ensuring that the drug is released and specifically kills the target cells while sparing healthy ones.

Five such ADCs have entered market approval, with around a hundred more under development[20]. One recently such developed ADC is brentuximab vedotin **(1-14)**, the prodrug of antitumour drug MMAE (vedotin) **(1-15)** used to treat Hodgkin's lymphoma. The antibody targets the CD30 membrane protein, which is highly expressed in tumour cells[21].

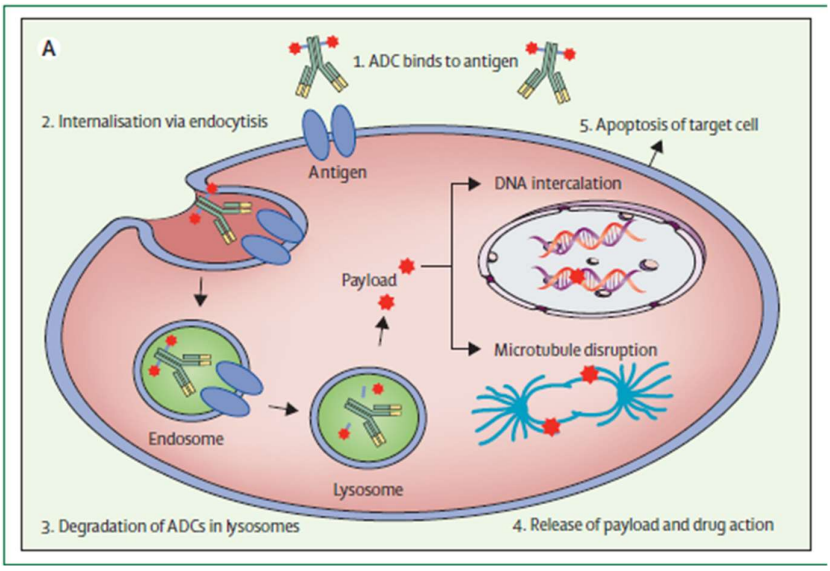
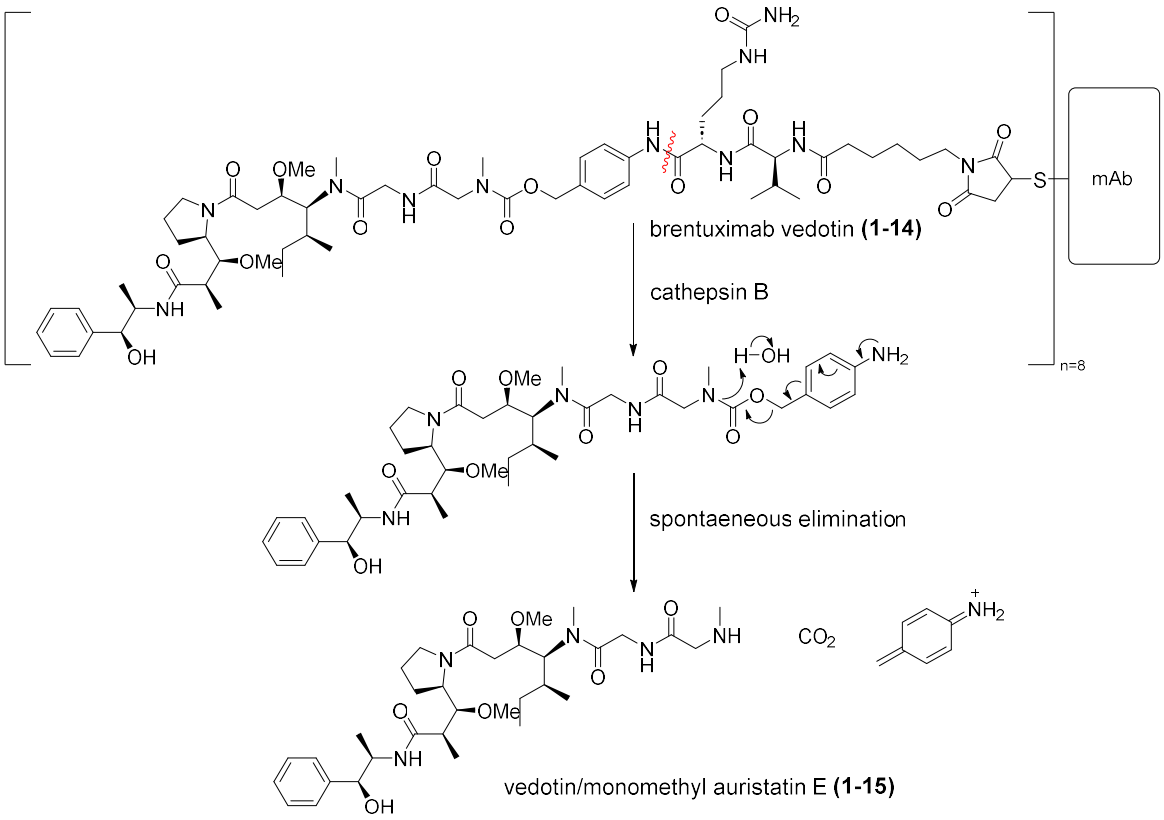


Figure 1.9: Mechanism of action of antibody-drug conjugates. Image adapted from Figg et al.[20]

The antibody-drug conjugate is stable in the human body, but once in the tumour cell, a tumour associated lysosomal enzyme, cathepsin B, cleaves the peptide bond at the C-terminal end of the citrulline-valine dipeptide. This releases an intermediate containing a para-amino benzyloxy carbonyl moiety, which spontaneously degrades to release the actual drug, monomethyl auristatin E(MMAE) **(1-15)**[22]. MMAE kills cells by inhibiting microtubule polymerisation and causing cell cycle arrest.



Scheme 1.10: Mechanism of the release of the antitumour drug MMAE **(1-15)** from the antibody drug conjugate brentuximab vedotin **(1-14)**. Note that the entire structure of the ADC contains 8 units of MMAE per antibody

1.2.4 Antibody directed and gene directed enzyme prodrug therapy

It is worth briefly mentioning the use of antibody directed enzyme prodrug therapy (ADEPT) and gene directed enzyme prodrug therapy (GDEPT). While this topic is straying from the discussions of prodrugs themselves, the treatments fundamentally employ prodrugs as part of its strategy. While still experimental and research is still undergoing, some studies have since progressed to clinical trials. If successful, both treatments can significantly expand the possibilities of drug delivery for the treatment of many diseases, particularly cancers.

In ADEPT, the antibody is conjugated to a prodrug activating enzyme, forming an antibody-enzyme complex. As before, the enzyme-antibody conjugate is localised to target cells through antigen binding. After waiting for the free enzyme to be cleared from the bloodstream, the prodrug is then administered. The prodrug is stable in circulation and is only cleaved to the

active drug in the presence of the enzyme. Due to its localisation of the enzyme in tumour cells the drug is only selectively released and kills cancer cells, sparing healthy cells.

A wide variety of prodrugs can be used in ADEPT therapy. The most successive experiments on ADEPT came from a series of clinical studies on CGP2-CMDA. system to treat gastrointestinal cancers as reviewed[23]. (CGP2 is a bacterial enzyme that cleaves folates and CMDA the prodrug).

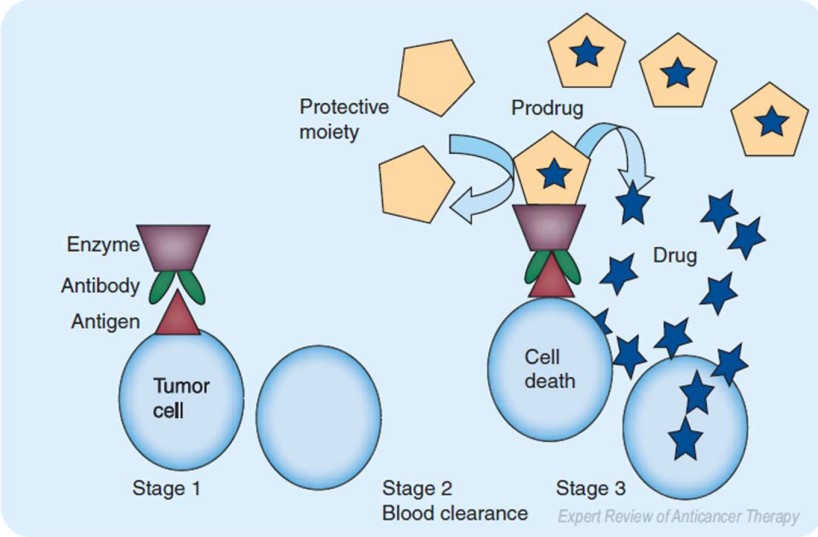


Figure 1.11: Mechanism of action of ADEPT prodrugs. Image adapted from Bagshawe[24]

In the series of clinical trials, certain difficulties are recognised that must be overcome. The antibodies should ideally be humanised (humanised antibodies are antibodies from non-human species whose protein sequence is altered to more resemble a human antibody equivalent) to minimise immune reactions, yet they cannot be identical to that of human enzymes already existing elsewhere in the body, otherwise specificity will be affected. The drug should have a short half-life to minimise diffusion from the tumour cells as much as possible to minimise killing of bystander healthy cells. The clearance of the free enzyme-antibody complex should be rapid and complete, otherwise the activation of the prodrug will not be localised at tumours, making the treatment severely toxic.

In GDEPT, instead of delivery of the enzymes through antibodies, the genes encoding the enzymes are transfected to tumour cells. The tumour cells translate these genes and synthesize the enzymes *in vivo*. The prodrug is then administered, and selectively activated in cells where the introduced gene is expressed.

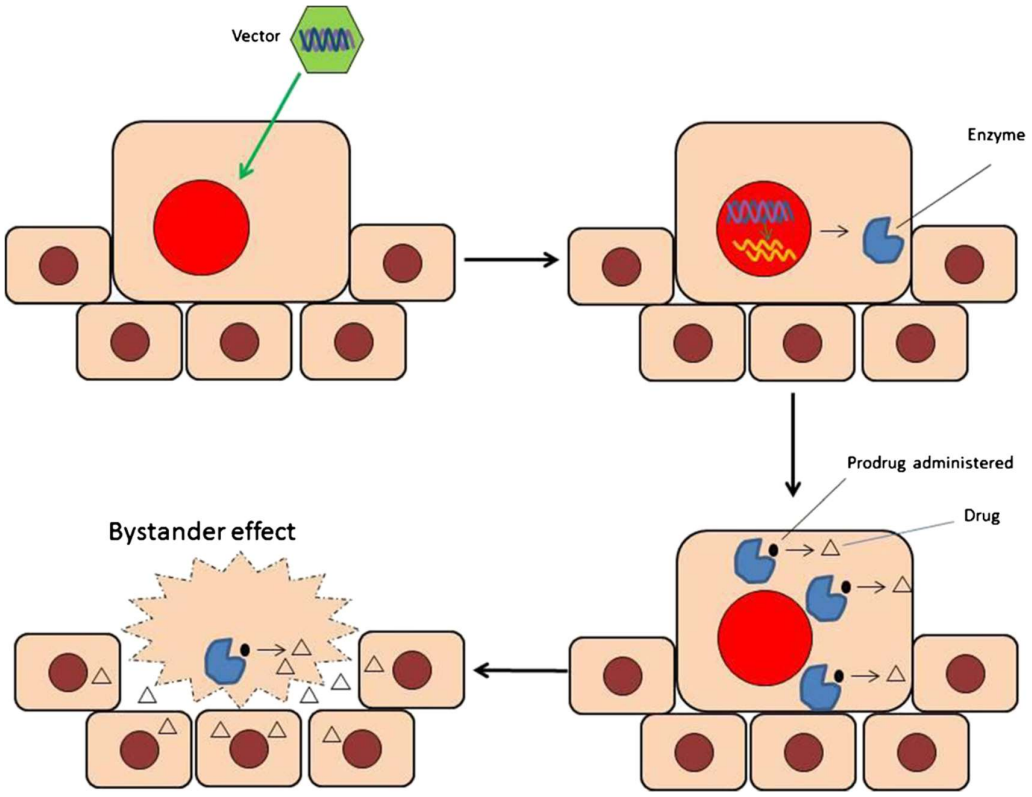


Figure 1.12: Mechanism of GDEPT. Image adapted from Zhang *et al*[25].

Introducing the gene directly as the DNA is uncommon due to poor permeability into cells and rapid degradation, so carriers known as vectors are used. Vectors in introducing drugs include viruses, mesenchymal stem cells, liposomes, and cationic polymers.

Several GEDPT systems are currently undergoing clinical trials. This includes cytochrome P450/ phosphoramide system and nitroreductase/ CB1954 systems, which has been reviewed[25] and will be briefly described later. However, none has yet entered the drug market, as several difficulties are identified that are yet to be fully overcome. Firstly, the rate of gene uptake by tumour target cells is small, at less than 10% GDEPT is therefore reliant on the bystander effect,

a phenomenon where neighboring cells of transfected cells are also killed either by diffusion, or local immune response to the targeted area. However, the bystander effect cannot be too much, or surrounding healthy cells will be killed. Furthermore, there are still safety concerns over the use of virus as vectors. Although these viruses are deactivated and rendered unable to replicate, its safety when it comes to immunogenicity is still not fully established. Alternate non-viral vector methods are less efficient in gene delivery as compared to viral vectors.[25]

In the end, while ADEPT and GDEPT has much potential, there are still optimization issues that need to be overcome in terms of their safety profile and efficiency before they can be deployed for clinical treatments.

1.3 Prodrug activation mechanisms

There are many ways a prodrug is chemically modified to its active from. Morphine, as mentioned earlier is created from the heroin prodrug through hydrolysis of the two ester groups into hydroxyl groups. However, many types of reactions are exploited during prodrug activation, as will be discussed below.

This chapter will attempt to classify prodrugs by mechanism of activation to the active drug. Generally, there are 3 methods, hydrolysis, oxidation, and reduction, which will be elaborated in greater detail in subsequent sections. This is not meant to be an exhaustive list of all the possible prodrug activation mechanisms due to its vast number, so only the more common occurring ones will be described.

1.3.1 Activation by hydrolysis

This is the most common form of activation. In fact, 49% of all prodrugs are activated by this method[26]. Many functional groups in prodrugs that are activated by hydrolysis include ester, phosphates, carbamates, oximes among others.

1.3.1.1 Esters

As mentioned previously, esters are a common motif used in prodrugs to increase the lipophilicity and membrane permeability of many drugs (particularly oral drugs). Esters are hydrolysed to release hydroxyl or carboxylic acid moieties of the active drug. The number of carboxylic acid esters developed is vast, and has been reviewed by Beaumont *et al.* [10]

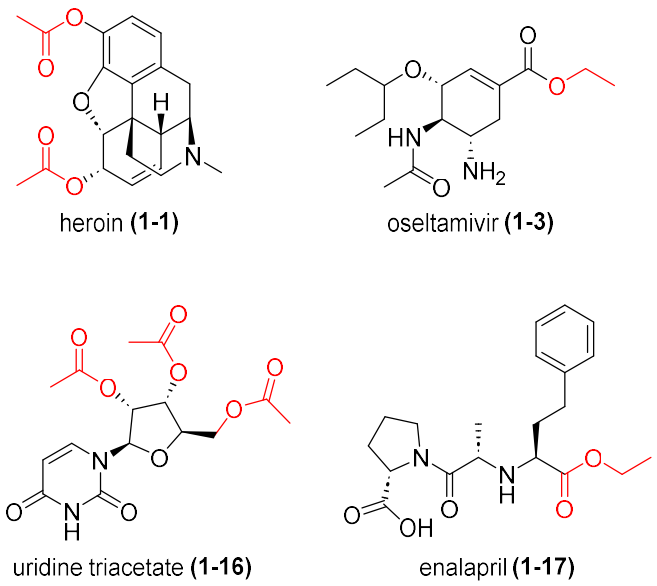
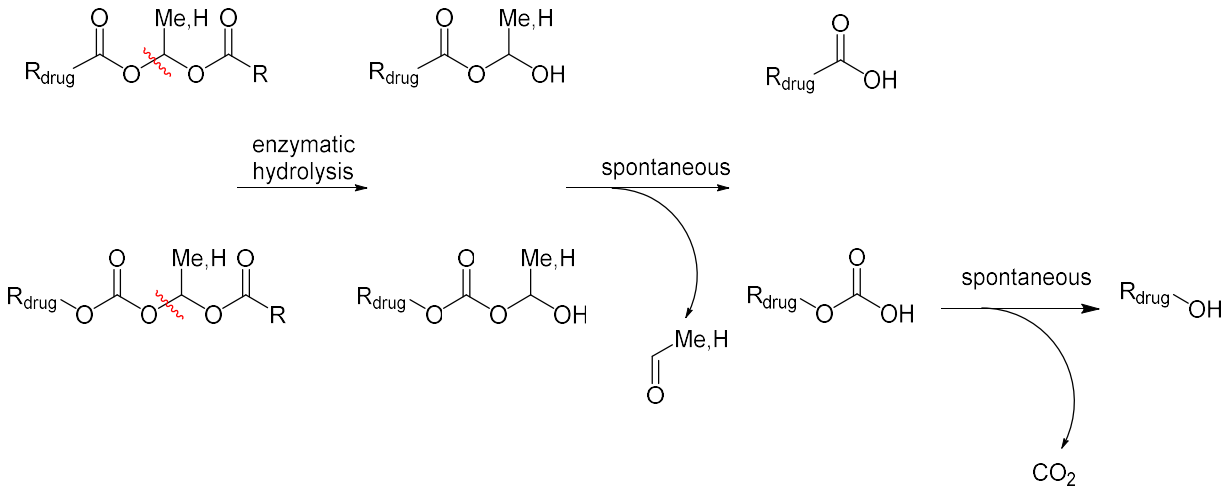


Figure 1.13: A few examples of a few commonly used simple ester prodrugs

Esters are cleaved by a variety of esterases ubiquitous in the human body. However, the assumption that esters are unstable and rapidly cleaved in the human body has been challenged[27], and there are cases where rate and extent of the hydrolysis of esters can be surprisingly limited, depending on the surrounding chemistry. When this happens, a double prodrug activation method can be used. Instead of using simple esters as the prodrug with one hydrolysis step, an α-acyloxyalkyl functional group is used. In this functional group, two esters are employed (carboxylic or carbonate ester), linked by an acetal (-O-CH_2-O-) bridge. (see Fig 1.14) Activation of this group involves multiple step cascade hydrolysis steps, usually the first one enzymatic, and the rest spontaneous. The enzymatic cleavage of one ester (or carbonate) forms an unstable hemiacetal that spontaneously degrades to release the carboxylic acid. If the

drug is linked to a carbonate ester, a carbonic acid is unmasked instead, which undergoes another spontaneous decarboxylation reaction to form the alcohol.



Scheme 1.14: The double ester moiety and the mechanism of the hydrolysis cascade to release the drug.

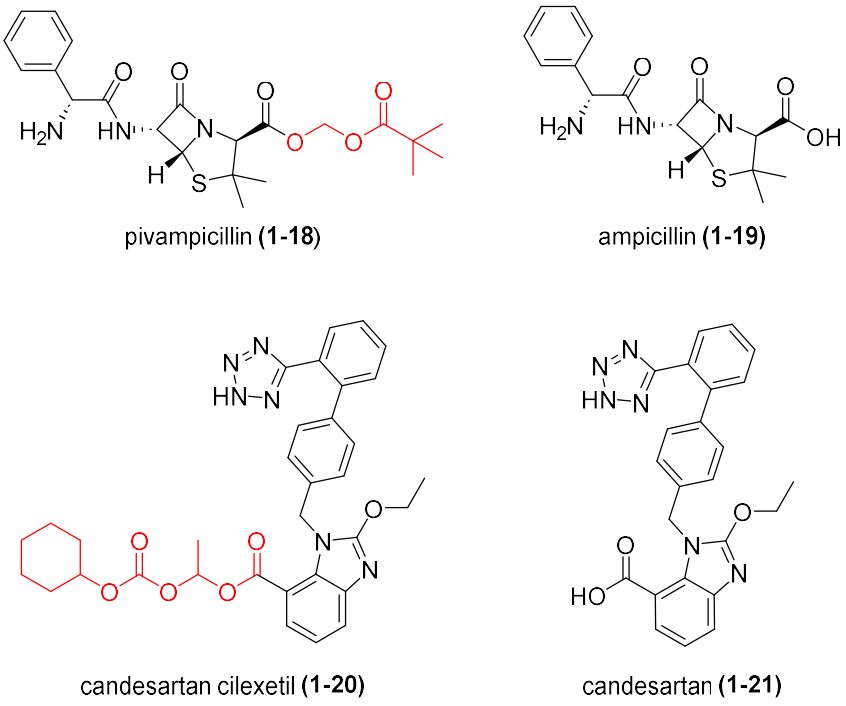


Figure 1.15: Double ester (includes carbonates) prodrugs and their corresponding active drug.

Prodrugs that make use of the double ester strategy include pivampicillin **(1-18)**, the pivaloylmethyl ester of ampicillin **(1-19)**. Compared to the parent drug, the oral bioavailability of the prodrug is increased from 62% to 92%[28]. Another drug is candesartan cilexetil **(1-20)**,

the cyclohexyl 1-hydroxyethyl carbonate ester of the antihypertensive drug candesartan **(1-21)**. Oral administration of this drug results in absorption through the gastrointestinal tract and complete and rapid hydrolysis to the active drug[29].

1.3.1.2 Phosphates and phosphate esters

While carboxylic acid ester prodrugs are more lipophilic than the original drug, phosphates prodrugs are more hydrophilic. The fact that the phosphate group is ionizable gives the phosphate much higher aqueous solubility. This makes phosphate prodrugs suitable for perinatal administration. Phosphates are hydrolysed by alkaline phosphatases, which like esterases, are ubiquitous in the human body.

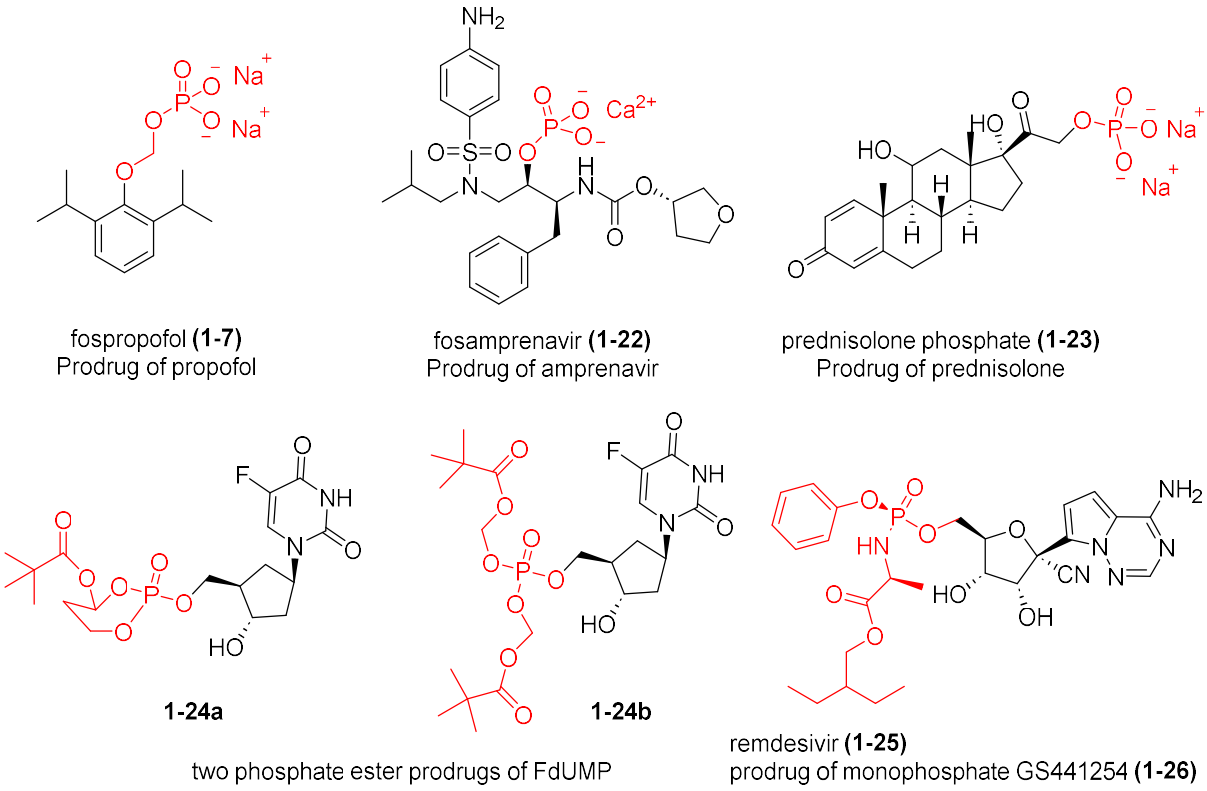


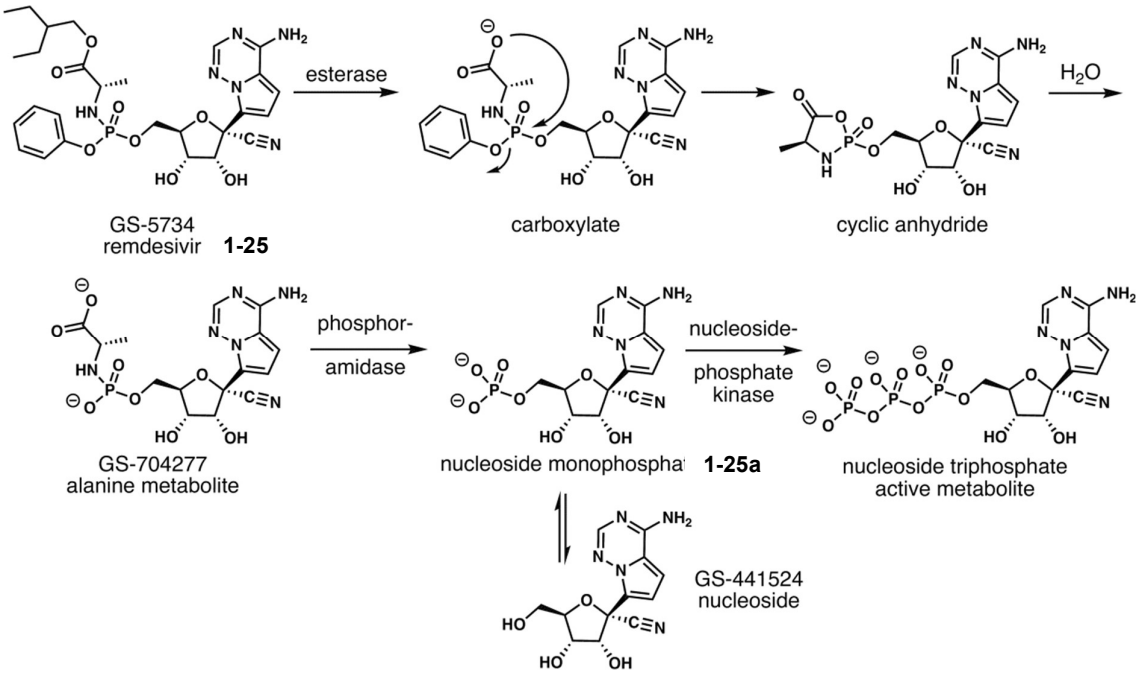
Figure 1.16: A few examples of phosphate, as well as phosphate ester drugs

On the other hand, phosphate containing drugs, mostly nucleoside containing drugs, can be rendered more hydrophobic by alkylation of its oxygen atoms. 2'-deoxy-5-fluorouridine 5'-monophosphate (FdUMP) **(1-24)**, a nucleotide prodrug of the previously mentioned 5-

fluorouridine **(1-13)**, is one such example. The molecule with its phosphate group is highly polar and has poor cell wall membrane penetration. FdUMP is thus further derivatized by alkylation, which gives it significantly more stability and hydrophobicity. However, it turns out that phosphate alkyl esters themselves are exceptionally resistant to enzyme cleavage. Therefore, they are often combined with another group like a carboxylic acid ester to create a cascade multistep hydrolysis activating moiety. Depending on the type of protecting group, the measured half-life in plasma can differ. For example, the bis(pivaloyloxymethyl) ester **(1-24b)** had a two step hydrolysis step with $t_{1/2}$ = 70 min, while the cyclic ester **(1-24a)** showed faster hydrolysis at $t_{1/2}$ = 30min[30].

One most notable example of a phosphate ester drug in current use is remdesivir **(1-25)**. It is an antiviral drug that has been repurposed to treat severe cases of COVID 19. The ester is first cleaved by esterase and the unmasked carboxylic acid displaces the phenoxy group to form a cyclophosphate anhydride intermediate, before another enzyme, phosphoamidase, cleaves the molecule into the monophosphate active compound (**1-25a**).

Scheme

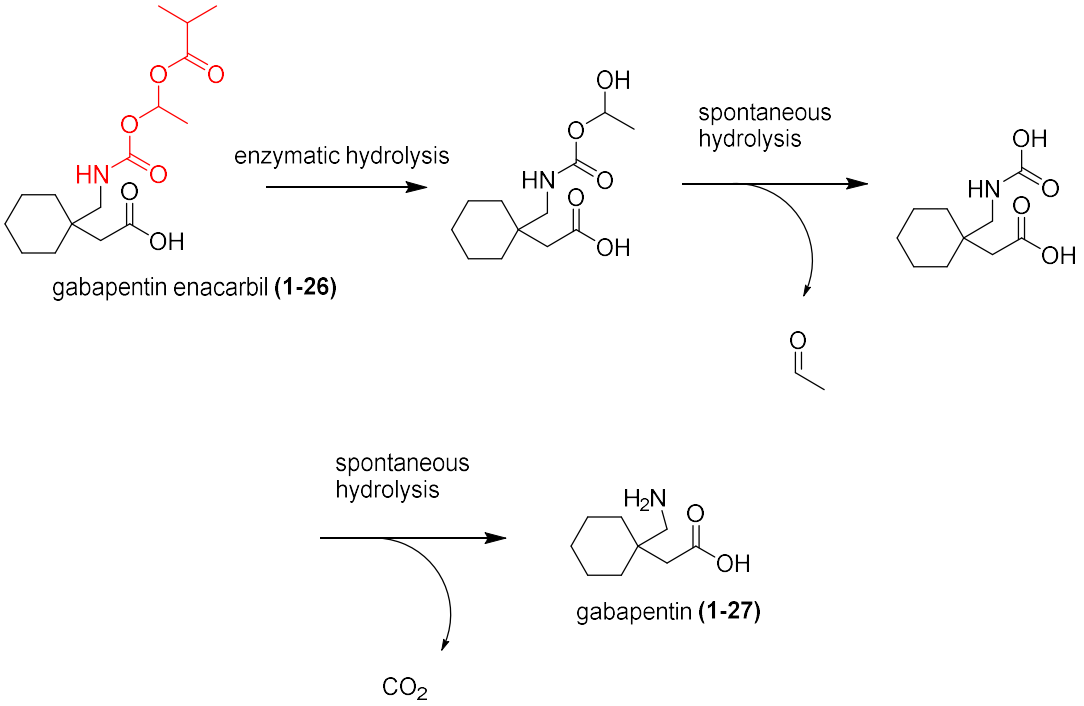


Scheme 1.17: Mechanism of action of remdesivir. Scheme adapted from Hall's[31] *et al.*

1.3.1.3 Carbamates and amides

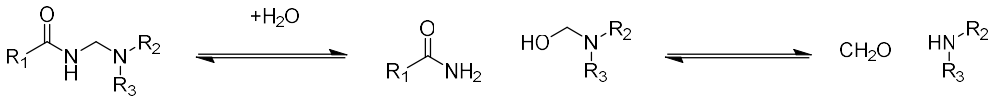
Carbamates have seen occasional use, but as seen in the case of bambuterol **(1-10)** (see section 1.2.2), carbamates are more resistant to hydrolysis than esters and carbonates. Amides are even more resistant. When they do exist, the amides are generally coupled to amino acids or peptides that are recognised by proteases, allowing directed cleavage of the amide bond.

Therefore, carbamates and amides, if used as a prodrug linking moiety, are frequently combined with other functional groups. One method is to combine it with an ester-acetal moiety similar as described in the section 1.3.1.1. One example is gabapentin enacarbil **(1-26)**, the isobutanoyloxyethoxy carbamate of gabapentin **(1-27)**. Its mechanism of activation is identical to candesartan cilexetil as described previously. An enzymatic hydrolysis of the ester is followed by 2 successive spontaneous hydrolysis as 2 unstable moieties are unmasked, finally giving gabapentin.



Scheme 1.18: Mechanism of activation of gabapentin enacarbil **(1-26)**, a carbamate-ester prodrug of gabapentin. **(1-27)**

An alternate method is the use of *N*-Mannich bases. The *N*-terminus of the amide is connected to an aminal (N-CH_2-N) group. This is unstable in the presence of water, and rapidly hydrolyses to release the amide or amine containing prodrug as shown in the mechanism below.



Scheme 1.19: Mechanism of action of N-Mannich prodrugs hydrolysis to release the active amine or amide drug.

One example using this functional group is rolitetracycline **(1-28)**, a prodrug of tetracycline **(1-29)**. The molecule is rendered highly water soluble with the addition of the pyrrolidine moiety. It is also rapidly hydrolysed non-enzymatically, with 50% decomposition within 3 hours in water[32]. A *N*-Mannich base can also be formed intramolecularly, as in hetacillin **(1-30)**, which is formed by the condensation of the beta-lactam antibiotic ampicillin **(1-19)** with acetone. In pH 6.8 aqueous solutions at 37°C it has a half-life of merely 20 minutes[33].

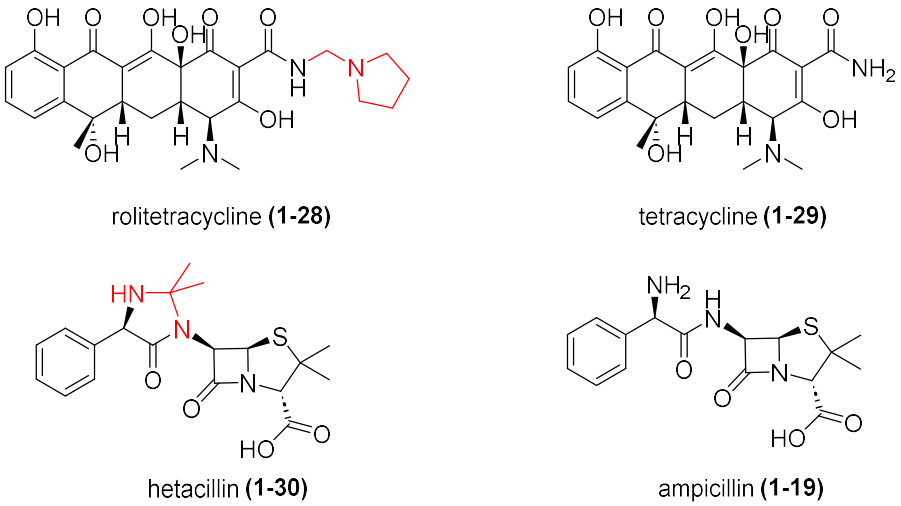


Figure 1.20: Examples of *N*-Mannich prodrugs (left), with their corresponding active drug (right)

1.3.1.4 Schiff bases

While much less common than the rest, oximes have been used as prodrugs. Norgestimate **(1-31)**, a progestrogen, is one example. After oral administration, it is rapidly and fully deacetylated to norelgestromin **(1-32)**, and subsequently also hydrolysed to the ketone levonogestrel **(1-33)**. It may be debatable to call norelgestromin and norgestimate prodrugs, as they do bind to the progesterone receptor, albeit weakly compared to its metabolite. Norelgestromin and norgestimate are both agonists to the progesterone receptor, at 10% to 15% binding affinity compared to related progestogen medication promegestone **(1-34)**. In contrast, the metabolite levonogestrel has a 150% binding affinity[34].

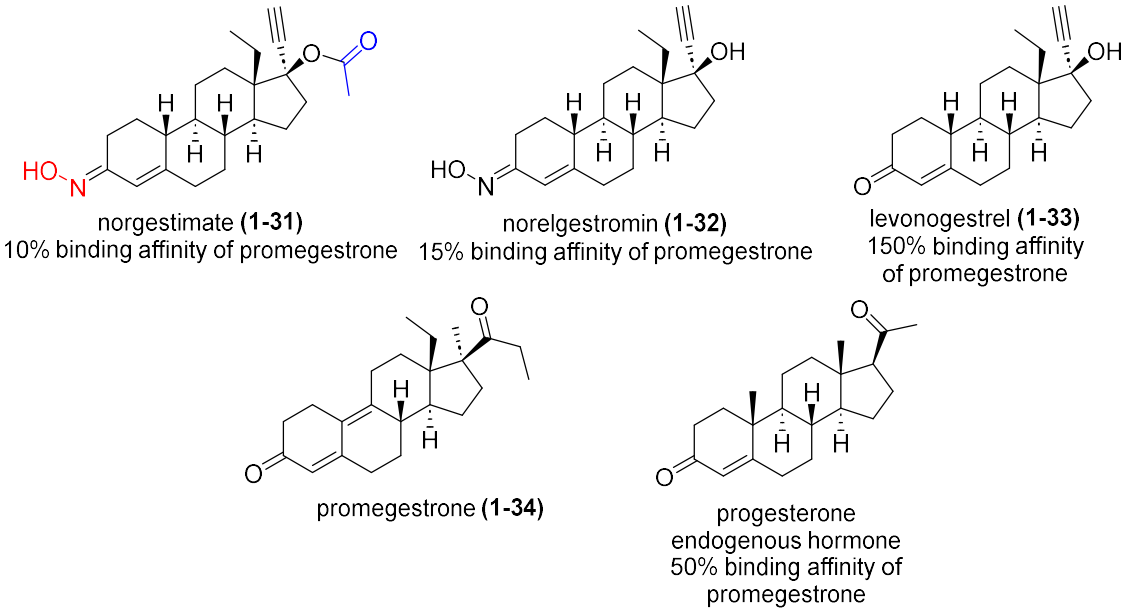
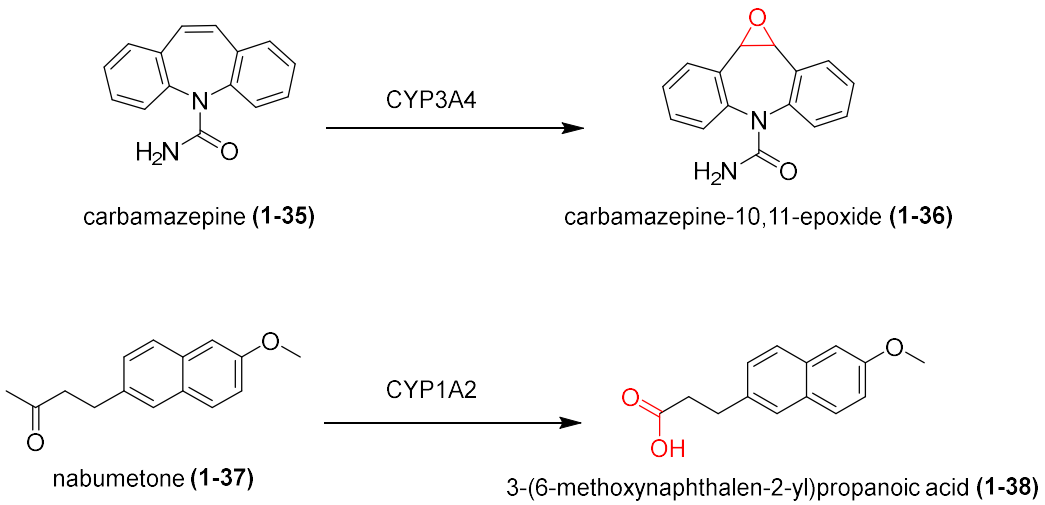


Figure 1.21: Norgestimate **(1-31)** is converted to norelgestromin **(1-32)**, and then to the levonogestrel **(1-33)**, which has a high binding affinity to the progesterone receptor. Structures of progesterone, the endogenous hormone shown as comparison

1.3.2 Activation by oxidation

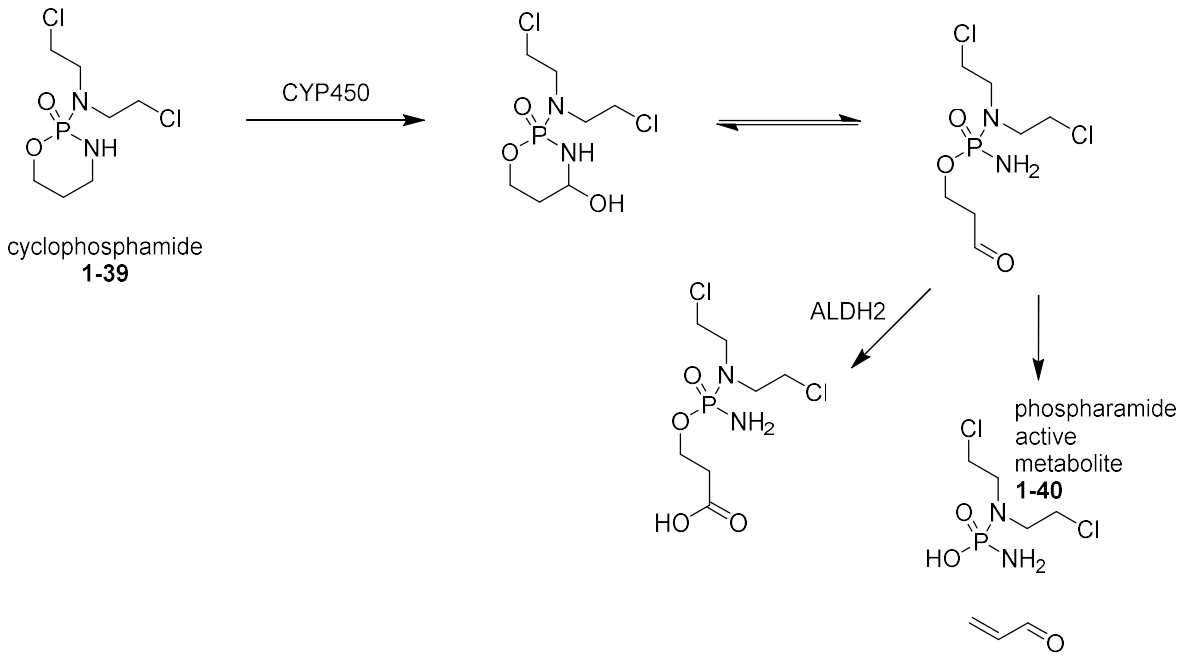
In the human body, the cytochrome P450 enzyme family is responsible for the oxidation of various compounds, including xenobiotics like drugs. It is estimated that 75% of drugs are metabolized by cytochrome P450 enzymes to some extent[35]. While many metabolic steps deactivate drugs, some prodrugs take advantage of this to create the active drug through enzymatic oxidation of the prodrug.

Carbamazepine **(1-35)** is one such example. It is used to treat epilepsy, but the molecule itself has no anticonvulsant effects. However, when the alkene moiety is oxidised to an epoxide by cytochrome P450 enzyme CYP3A4, the active metabolite **(1-36)** is formed[36]. Similarly, nabumetone **(1-37)**, a NSAID, is oxidised at the methyl ketone to form the carboxylic acid metabolite 3-(6-methoxynaphthalen-2-yl)propanoic acid **(1-38)**. [37] Administering nabumetone orally has the added advantage of minimising gastric irritation/ulcers due to its non-acidic nature, unlike the active drug.



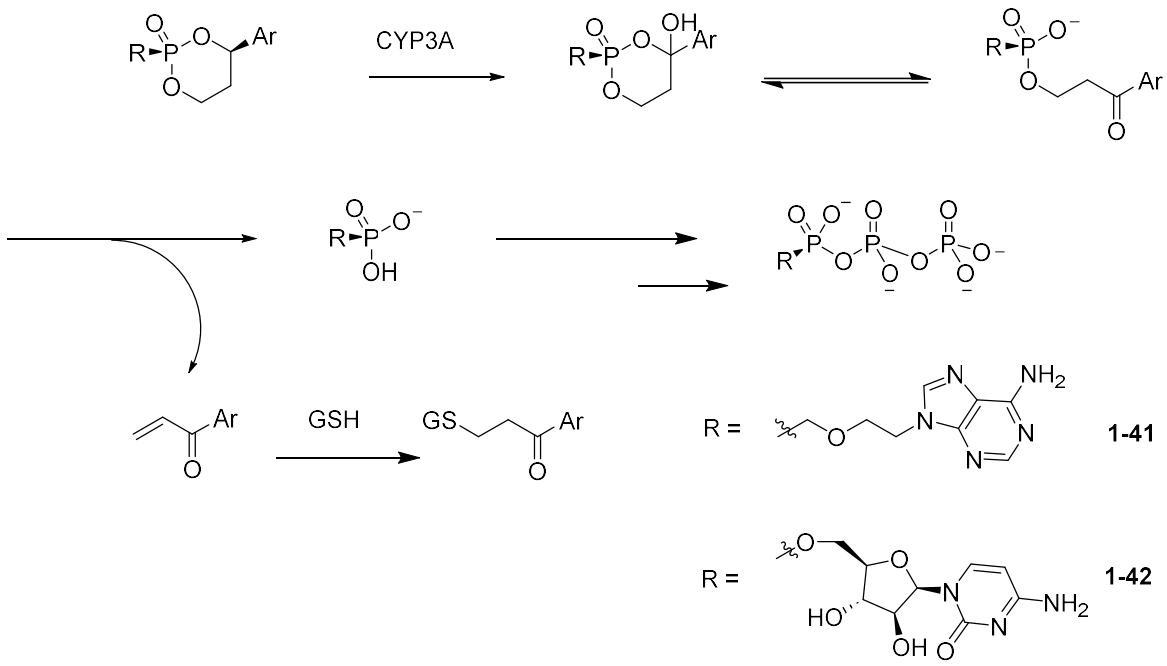
Scheme 1.22: 2 prodrugs that rely on cytochrome activation to release the active drug.

Oxidation can be combined with hydrolysis in prodrug activation, in the case of the chemotherapeutic drug cyclophosphamide **(1-39)**. In hepatocytes, cyclophosphamide is hydroxylated, forming a unstable hemiaminal that undergoes ring opening to from aldophosphamide, which is further hydrolysed to acrolein and the active phospharamide **(1-40)**. [38] A small amount of the aldehyde is oxidised to the inactive carboxylic acid. Unfortunately, the release of acrolein and its low therapeutic index makes this compound highly toxic.



Scheme 1.23: Mechanism of action of cyclophosphamide. Enzymatic hydroxylation is followed by ring opening and further hydrolysis.

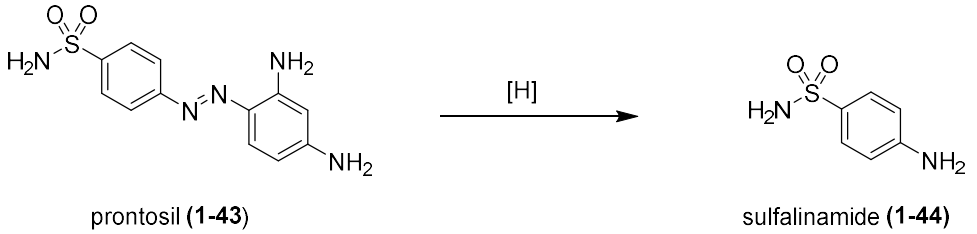
An example of rational prodrug design taking advantage of cytochrome oxidation to achieve targeted delivery to the liver was established by Erion et al. [39]. In their prodrug system, called HepDirect, two drugs, remofovir **(1-41)** and MB07133 **(1-42)** were attached to a cyclic phosphate ester as shown. The mechanism of action is highly similar to phospharamide as described above. The aryl group at position 4 is oxidized by CYP3A to form a hydroxy metabolite, which is unstable and rapidly undergoes a ring opening to form a negatively charged linear phosphate intermediate. Further reactions cleave the phosphate ester to the phosphate, and this monophosphate is converted to the triphosphate, the active drug. Meanwhile the other by product, an aryl vinyl ketone (significantly less toxic than unsubstituted acrolein), is conjugated by glutathione and detoxified. Because the prodrug is only metabolised in the presence of CYP3A, the drug is mostly only formed inside hepatocytes, where the enzyme is mostly located at. The electric charge of the intermediate and drug also prevents its exit through the non-polar cell membrane, causing it to accumulate in the hepatocyte and thus maximising specificity.



Scheme 1.24: Hepdirect prodrug cytochrome P450 induced mechanism of action from Erion *et al*'s paper.

1.3.3 Activation by reduction

One of the earliest discovered antibacterial drugs, prontosil **(1-43)**, was discovered to be a prodrug. The azo moiety (-N=N-) in prontosil is reductively hydrolysed to release the active amine metabolite sulfanilamide **(1-44)**.



Scheme 1.25: Prontosil is reduced to the active drug sulfalinamide

Today, reduction-activated prodrugs as a topic have received a lot of attention in drug research, because they are commonly used in the treatment of cancers. Many tumours have a highly reducing microenvironment and are particularly affected by hypoxia. Due to their rapid uncontrolled growth, they often outgrow its blood supply and receive insufficient available oxygen. This hypoxic microenvironment is associated with treatment resistance and poor

outcomes[40], but the unique tumour microenvironment offers an opportunity to try a new prodrug strategy, one where the prodrug is activated specifically by the reducing conditions of tumour cells. This allows of a creation of a new series of highly tumour selective chemotherapeutic drugs, which has often suffered from serious side effects due to its poor selectivity and therapeutic index.

1.3.3.1 Mitomycins and quinones

Mitomycins were first discovered in the 1950s from the bacteria *Streptomyces capitosus*. It is used to treat various cancers as a chemotherapeutic agent. It works by crosslinking the two strands in DNA, which prevents the separation of the strands needed for DNA replication and RNA transcription. It is effective enough that a single crosslink in a bacterial genome is lethal to the bacterium[41].

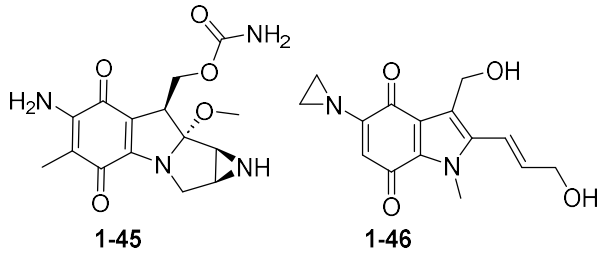
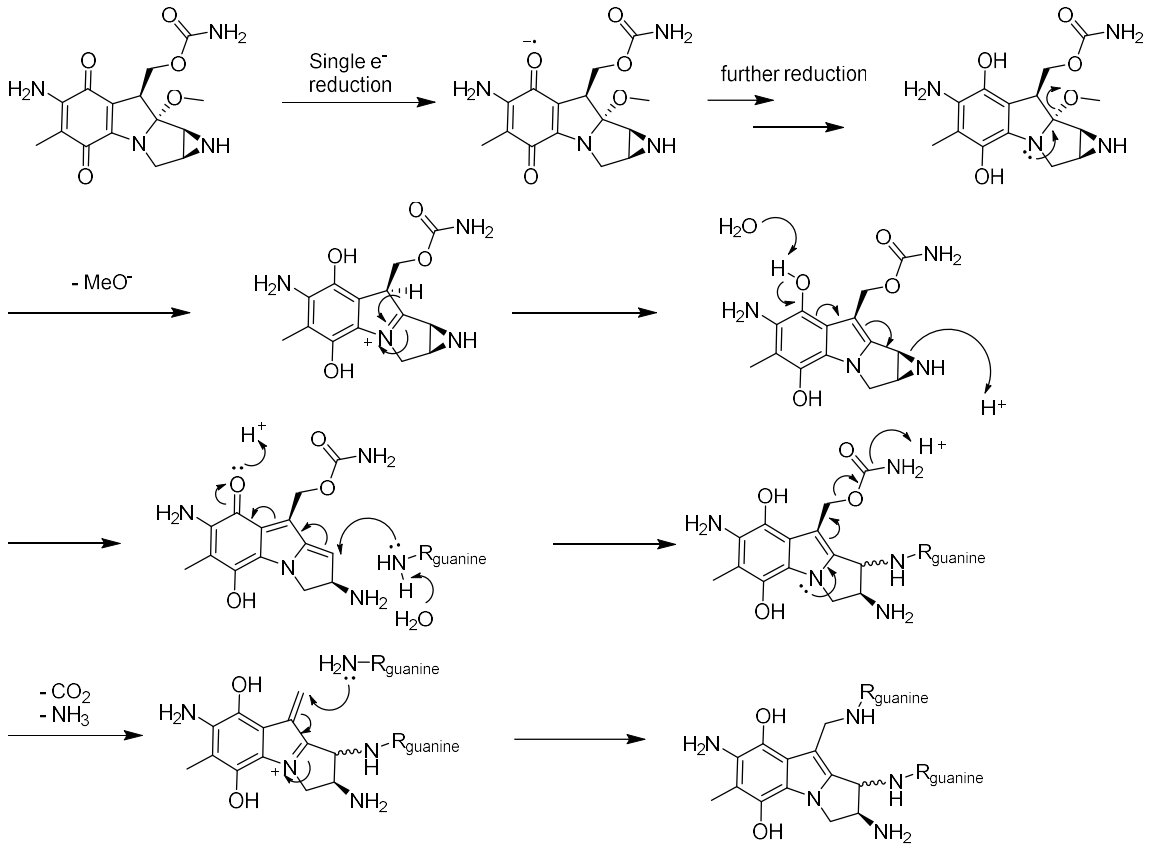


Figure 1.26: Left: Mitomycin C (MMC) **(1-45)**, representative compound of the mitomycins. Right: EO9 **(1-46)**, a synthetic mitomycin analog.

All mitomycins contain an aziridine as well as a benzoquinone group, both which are crucial in its mechanism of action. Mitomycin itself has negligible activity, but it is activated by reduction of the quinone moiety, which causes a cascade reaction leading to DNA alkylation (Scheme 1.27) Firstly, the quinone is reduced to the hydroquinone (this can be through various reduction enzymes[42]) inside the hypoxic microenvironment of the tumour cell. This subsequently leads to the expulsion of the methoxy group, followed by the opening of the aziridine ring. The intermediate enamine formed after ring opening is a potent alkylating agent and forms a covalent bond with a guanine base from DNA to form a complex, further expelling

ammonia and carbon dioxide. The resultant exocyclic alkene that is formed is also electrophilic can further react with another guanine base, forming the bis-guanidine crosslinked complex.



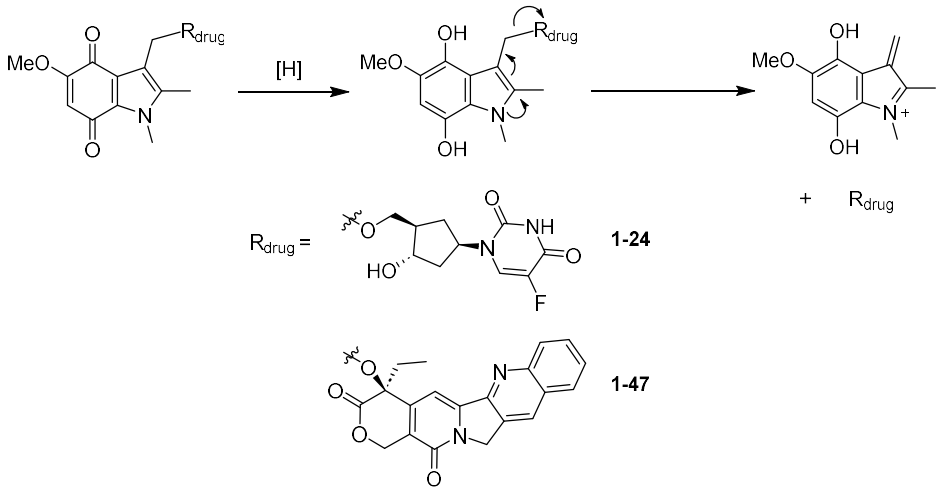
Scheme 1.27: Mechanism of crosslinking of mitomycin C to two DNA guanine bases activated by reduction.

What is remarkable is that mitomycins are completely specific for the DNA sequence CpG•CpG[41] (a specific dinucleotide sequence of cytosine adjacent to a guanine linked with a phosphate group). Crosslinks only occur between guanines on opposite DNA strands, which require specifically the CpG•CpG sequence. Crosslinking by mitomycin does not occur on other nucleotides. Most DNA alkylating agents, while having a preferred target DNA sequence, are not absolutely specific in comparison.

Analogues of mitomycins have been synthesized, one notable one being EO9 **(1-46)** (see Fig. 21). While it showed promising antitumour activity in animal models[43], it was therapeutically ineffective in humans, due to its poor cell wall penetration and short half-life inhibiting delivery of the drug[44]. Development of the drug stalled until the last decade, where it was successfully

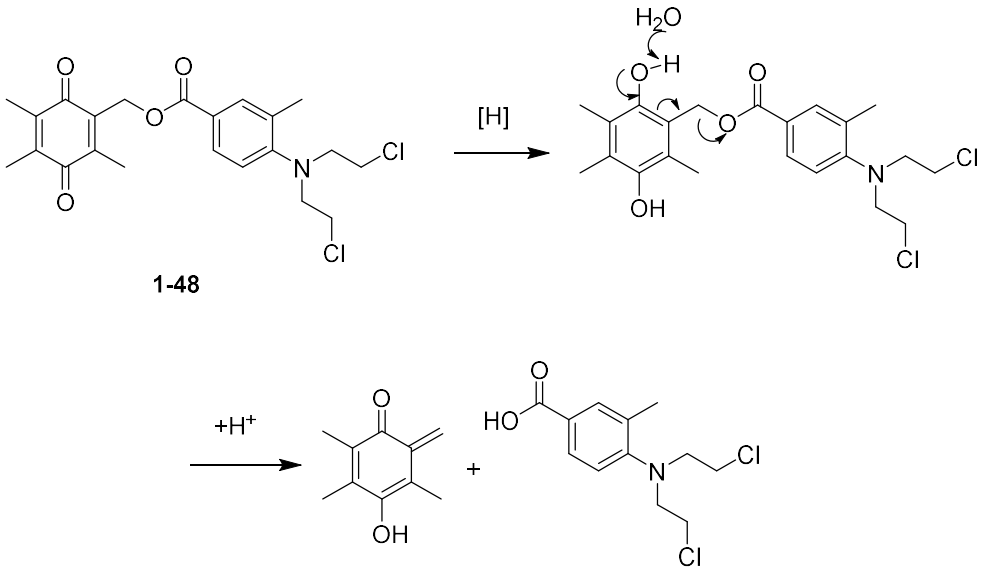
repurposed to treat bladder cancer through intravesical administration (directly into the bladder), bypassing its poor pharmacokinetic profile.[45]

The concept of benzoquinone reduction triggering a cascade reaction to create an active cytotoxic agent has inspired the design of other prodrugs that rely on a similar mechanism to release the active drug in reducing conditions. Indolequinones have been investigated as one example of a reduction activated prodrug moiety, with a mechanism of activation like that of mitomycins as shown.



Scheme 1.28: Mechanism of action of the release of drugs in an indolequinone reduction-activated prodrug system with examples 5-deoxyuridine and camptothecin.

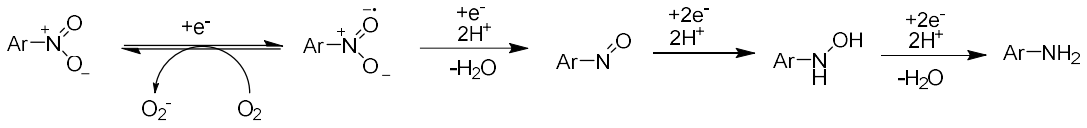
Drugs that have been tested *in vitro* using this strategy include: 5-deoxyuridine[46] and camptothecin[47]. Benzoquinone itself has also been used as the reduction activated prodrug, one example being a nitrogen mustard releasing compound **(1-48)** as shown[48].



Scheme 1.29: Mechanism of drug release in a benzoquinone reduction activated prodrug system

1.3.3.2 Nitrophenyl compounds

The nitro group attached to a phenyl ring can be reduced in *vivo* by flavoproteins in the body through successive one electron reductions as shown.

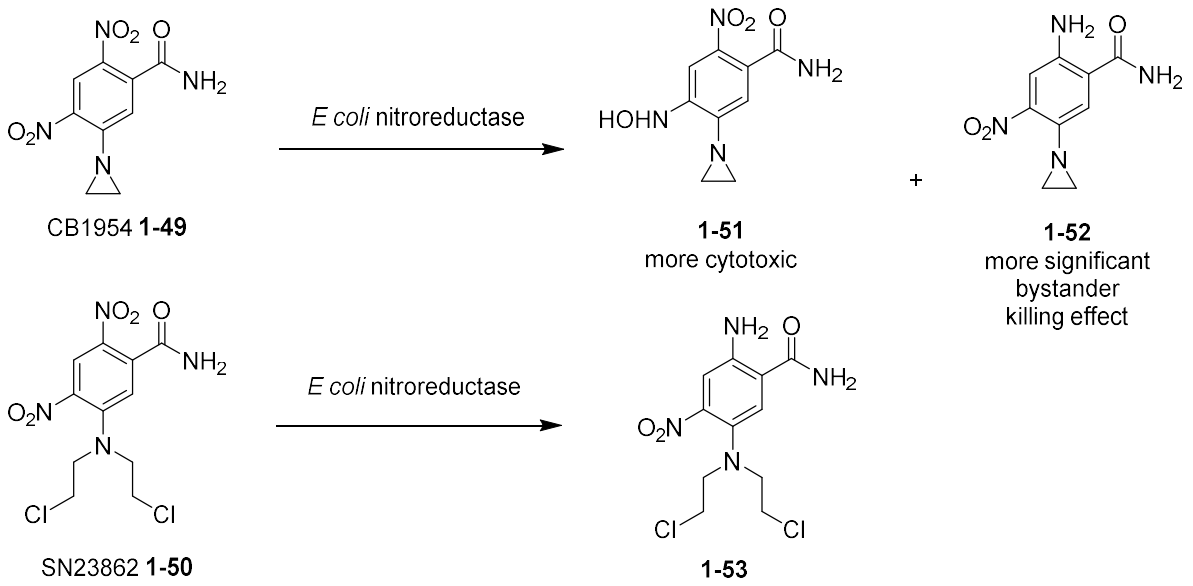


Scheme 1.30: Successive reduction of nitroaromatic compounds into aryl amines. Adapted from Chen *et al*[49].

Under aerobic conditions, the presence of oxygen rapidly reoxidises most of the nitro radical anions that are formed from one electron reduction of aromatic nitro compounds. In the hypoxic microenvironment of tumour cells, lower levels of cellular oxygen mean these compounds persist for longer and are more likely to be reduced to the corresponding hydroxylamine or amine. This can be exploited in a prodrug strategy by designing a nitro containing prodrug that is stable in normal physiological conditions, but the corresponding reduced compound is unstable and disintegrates to release the drug or alkylating agent.

One class of nitroaromatics under investigation are nitrobenzyl nitrogen mustards. Two examples that were under investigation of this were CB1954 **(1-49)** and SN23862 **(1-50)**. When

the electron withdrawing group was reduced to an electron donating amine group, the aziridine/ nitrogen mustard group was significantly activated to alkylate DNA.

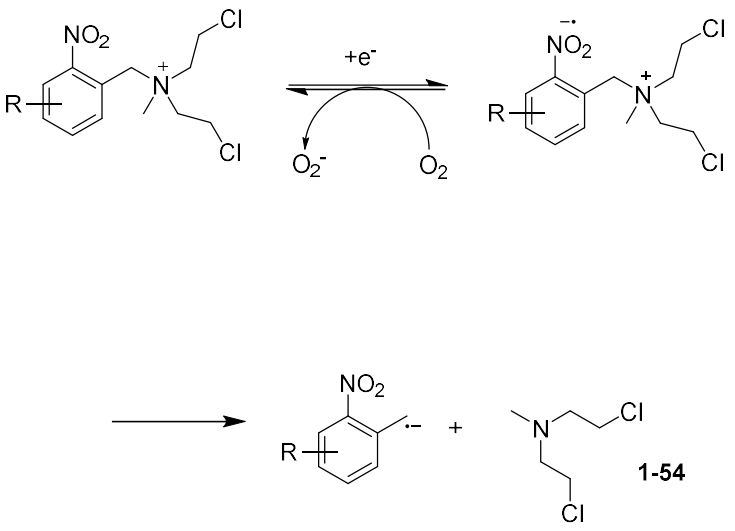


Scheme 1.31: Two reduction-activating prodrugs CB1954**(1-49)** and SN23862**(1-50)**. In both cases, reduction of a nitro group by *E coli* nitroreductase to the amine or hydroxylamine significantly increases the alkylating activity of the aziridine of nitrogen mustard moiety respectively.

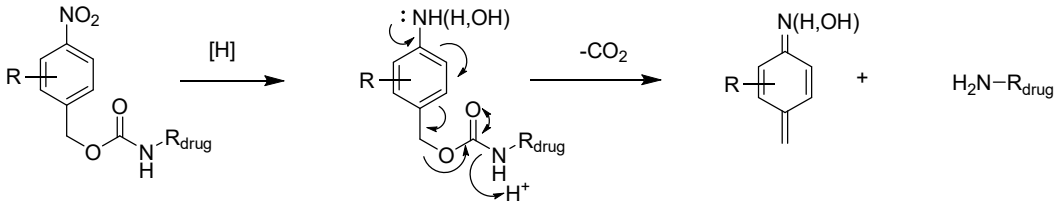
CB1954 was not efficiently reduced by human enzyme DT-diaphorase but was rapidly reduced by *E coli* nitroreductase to form a mixture of 2 and 4 reduced compounds. It was found that the 4-hydroxylamine metabolite **(1-51)** was the most potent in DNA crosslinking, but the 2-amino metabolite **(1-52)** was also significant, as it has a greater bystander killing effect[50]. SN23862, *E. coli* nitroreductase exclusively reduced the 2-nitro group to the amine **(1-53)** and was discovered to have superior pharmacokinetics in terms of bystander effect killing compared to CB1954. GEDPT trials with SN23862 and *E coli* reductase are still under investigation.

More generally, nitroaromatics have been used as a prodrug moiety that is cleaved under reducing conditions to release the drug. One early investigated class of drugs are nitrobenzyl quaternary ammonium salts, as represented by the molecular structure below. They are believed to exert their effect by one-electron reduction to the radical anion followed by spontaneous lysis with the release of the active compound mechlorethamine **(1-54)**. While the

cytotoxicity is selective for hypoxic cells, its extent of selectivity is not correlated to reductive potential, and selectivity significantly varies from none to 3000 fold depending on the substituents on the phenyl ring due to undetermined reasons[51]. It should be noted that reduction potential only partially contributes to reductive selectivity, because as an enzymatic reaction, enzyme binding is also a very important factor. Another investigation by the same authors on nitro heterocycles (including pyrrole, imidazole, thiophene and pyrazoles) also showed a similar lack of correlation between reduction potential and hypoxic selectivity[52] There is also the additional difficulty in its reported instability and tendency to release the drug prematurely, which would have led to extreme toxicity *in vivo*.



Scheme 1.32: Mechanism of action of nitrobenzyl quaternary ammonium salts, which release nitrogen mustard **1-54** as the active drug.



Scheme 1.33: Mechanism of action of 4-nitrobenzyl carbamate prodrugs

The 4-nitrobenzyl carbamate group is another class of drugs that utilises a similar mechanism of drug release as shown above. Drugs including mechlorethamine **(1-54)** (note that it is linked to a carbamate, as opposed to with a quaternary ammonium compound) and mitomycin C **(1-45)** as previously mentioned, as well as doxorubicin **(1-55)**, have been described[53] that use this prodrug moiety. In all cases, these prodrugs are reduced by nitroreductase, and the enzyme was proposed to be introduced either in ADEPT or GDEPT.

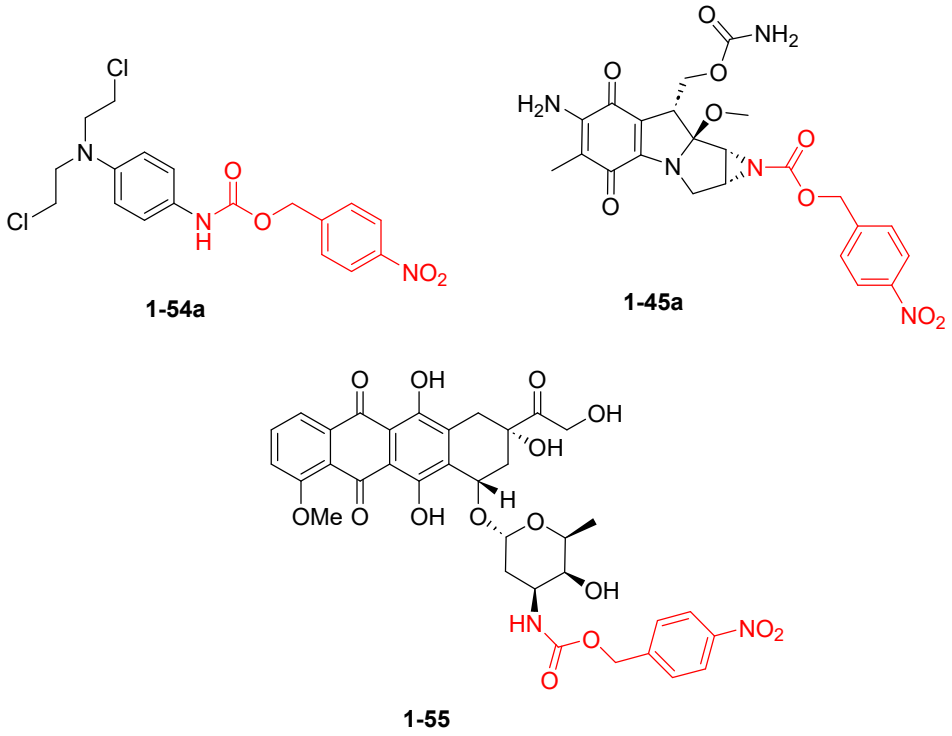
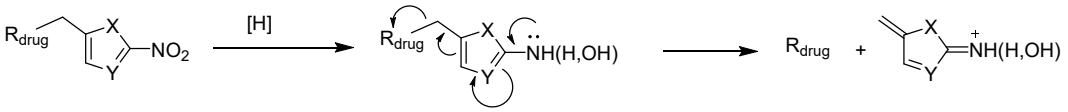


Figure 1.34: Some examples of 4-nitrobenzyl carbamate prodrugs that have been tested. They include prodrugs of mechlorethamine, mitomycin C, and doxorubicin.

1.3.3.3 Nitroheterocyclic compounds

Aside from nitrophenyl compounds, nitro groups attached to heterocyclic rings have also been used as anti-cancer prodrugs. Their mechanism of action is like that of nitrophenyl compounds; the nitro group is reduced to the amine which alters the electronic property of the ring and induces expulsion of the drug.



	X	Y
N-methyl imidazole	NMe	N
Furan	O	CH
Thiophene	S	CH

Scheme 1.35: Mechanism of action of nitro heterocyclic prodrugs

Threadgill *et al*. synthesised and investigated the use of a 2-nitro 5-*N*-methyl imidazolyl[54] group and nitrofuryl[55] groups as a prodrug moiety. The moiety was linked to a known poly(ADP-ribose)polymerase inhibitor 5-bromoisoquinoline as a test compound. The reduction of the nitro group and subsequent expulsion of the quinoline drug was shown to be possible, but this was done through chemical reduction. No bioassays of this compounds were done, but other biological assays done on 1-methyl-2-nitroimidazolyl linkers of drugs phenyleneamine mustards **(1-56)** and 5-amino1-(chloromethyl)-3-[(5,6,7-trimethoxyindol-2-yl)carbonyl]1,2-dihydro-3H-benz[e]indoline **(1-57)** failed to show selective antitumour activity[56]. There are some successes however, with the use of this moiety linked to a phospharamide drug TH-302 **(1-58)**, which showed 400-fold cytotoxic activity against H640 cells under hypoxic over normal conditions[57].

Nitrothiophenes have been used as a prodrug moiety of the anti-tubulin drug combretastatin A4 **(1-59)**. The prodrug has very weak cytotoxicity against A549 cells, but under hypoxic conditions was able to induce formation of the active drug *in vitro*, rendering them cytotoxic[58].

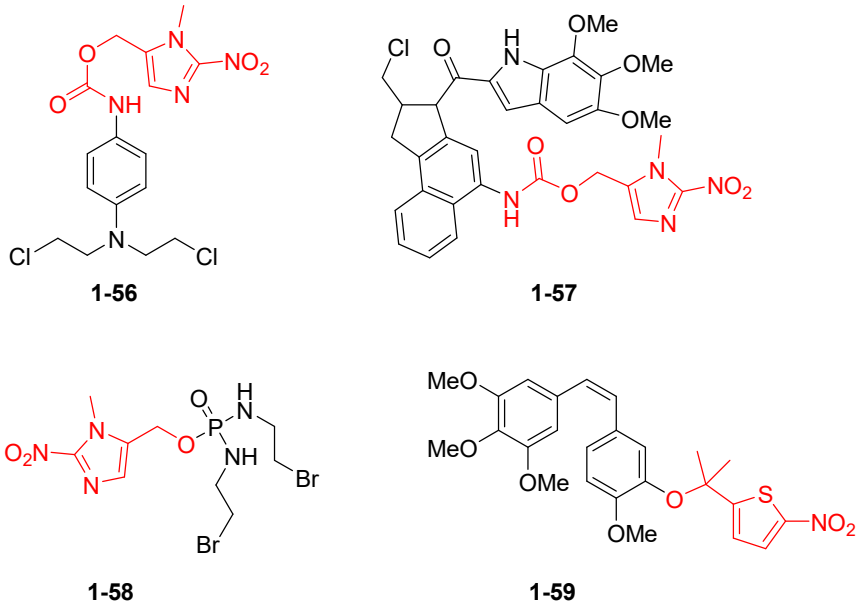
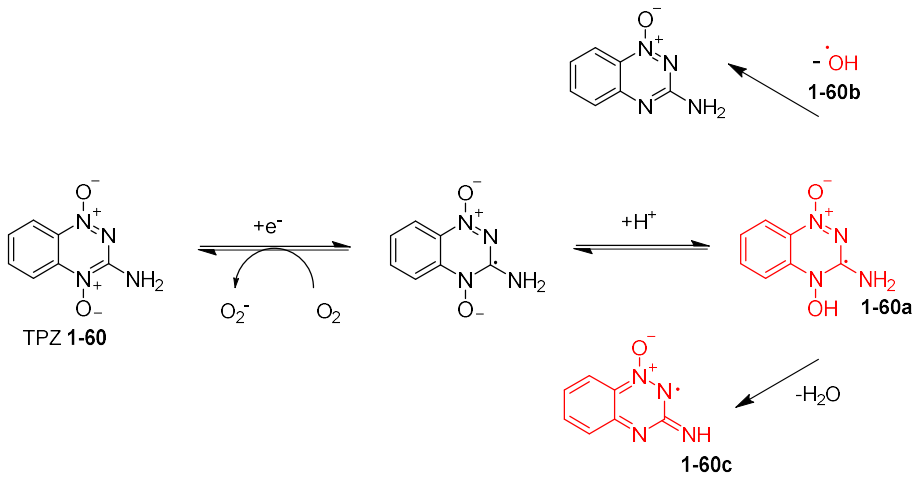


Figure 1.36: Examples of nitro heterocyclic prodrugs that have been tested. Examples include 1-methyl-2-nitroimidazolyl prodrugs of phenylene mustards **(1-56)**, 5-amino1-(chloromethyl)-3-[(5,6,7-trimethoxyindol-2-yl)carbonyl]1,2-dihydro-3H-benz[e]indoline **(1-57)**, TH-302 **(1-58)** and the nitrothiophenyl prodrug of combretastatin A4 **(1-59)**.

1.3.3.4 N-oxides

In 1985, Zeman et al. discovered that the *N*-oxide compound SR-4233, now called tirapazamine (TPZ) **(1-60)**, is highly cytotoxic specifically to hypoxic cells. In mouse SCC VII cells, hypoxic cells can be killed by it TPZ at concentrations 75 to 200 lower than in oxygenated cells. The mechanism of action is similar to that of previous classes of drugs[59]. The prodrug itself has negligible cytotoxic activity, but single electron reduction forms radicals that are potent DNA alkylating agents. The rapid oxidation of the radicals by oxygen minimises its cytotoxicity in oxygenated cells. The actual radical species that have the DNA damaging effect has not been established with certainty. Hypotheses include the *N*-hydroxy neutral radical species **(1-60a)**, the hydroxy radical released from the aforementioned neutral radical**(1-60b)**[60], or the dehydrated benzotriazine 1-oxide radical**(1-60c)**[61].



Scheme 1.37: The mechanism of activation of tirapazamine **(1-60)**, and hypothesized active species **(1-60a-c)**.

Despite promising results *in vitro*, clinical trials have shown that TPZ provides minimal additional benefit to patients when added to a standard regimen of cisplatin for the treatment of lung, head, or neck cancers[62]. Reasons suggested include the limited killing range of TPZ due to the instability of the radical intermediate and failure to consider the low proportion of cancerous cells that are hypoxic in some patients before trials.

Another *N*-oxide undergoing clinical trials is AQ4N or banoxantrone **(1-61)**, a double *N*-oxide. It is reduced to the corresponding amine AQ4 **(1-62)** by NADPH: cytochrome C reductase, which is an analogue of the topoisomerase II inhibitor mitoxantrone. The amine moiety aids in cellular uptake and binding to DNA, properties which do not exist in the amine oxide[63]. AQ4 is indiscriminate in its cytotoxicity, but AQ4N is nontoxic, and is only selectively cytotoxic in hypoxic tumour cells through enzymatic reduction to AQ4. *In vivo* tests in mice showed the successful localisation of AQ4 upon administration of AQ4N. It also demonstrated that while AQ4N has minimal effect on tumour growth in xenografts, co-administration of AQ4N with mitoxantrone was more effective in tumour growth suppression than either drug alone[64].

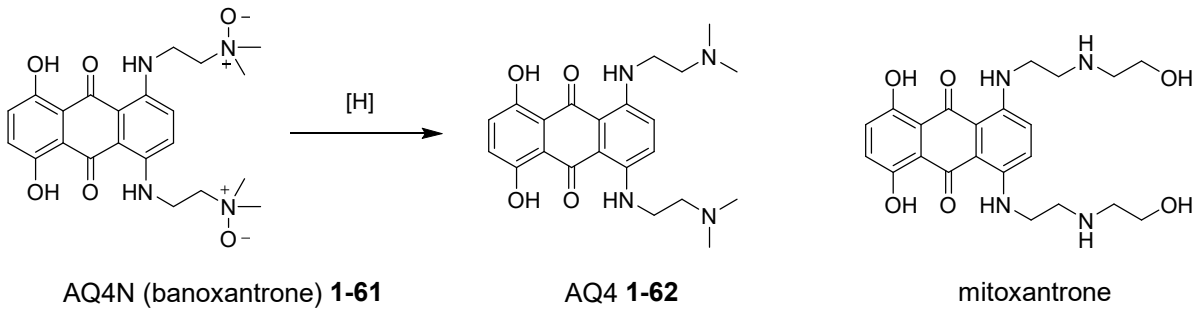


Figure 1.38: Structure of AQ4N **(1-61)** and its reductive activation to AQ4 **(1-62)**. The structure of the analogue mitoxantrone is also shown

1.4 Challenges

The ideal prodrug should have the following properties: it should be non-toxic in the body and stable, releasing the drug only at the target site. The conversion of the prodrug to the active drug should be complete, while the side products should be harmless and easily excreted. It should be sufficiently hydrophilic to allow rapid circulation in the bloodstream, yet hydrophobic enough that it can bypass the cell membrane of the desired target cells. In reality, no compound can meet all of these requirements perfectly, and prodrug discovery must take into account all of these factors in order to maximise beneficial properties. This has been the source of some difficulties encountered during clinical testing of prodrugs, as explained below.

1.4.1 Unnecessary prodrugs/wasted cost

Sometimes, a prodrug may not turn out to be significantly better than the parent drug in clinical trials. Hetacillin **(1-30)** (section 1.3.1.3), the *N*-Mannich analogue of ampicillin **(1-19)**, was originally developed as a distinct drug analogue of ampicillin. It was later found to have negligible activity by itself, and its action was purely due to its rapid hydrolysis to ampicillin, making it a prodrug[33]. A trial showed no difference in peak blood levels of the 2 compounds after administration, which suggests that hetacillin has no advantages over ampicillin[65]. This calls into question the need for further development of the drug, and the drug is no longer marketed for human use.

More recently, there has been questions on the use of remdesivir **(1-25)** in the treatment of COVID-19, as the benefit of using it over the nucleoside active drug GS-421524 **(1-25a)** was much less significant than anticipated. The synthesis of remdesivir is very capital intensive, requiring 25 steps and 12 months of production, although it was later reduced to 6 months[66]. In addition, while the phosphate ester was supposed to stabilise remdesivir, a pharmacokinetic study revealed that after IV administration, it was rapidly hydrolysed in the bloodstream, and that the active drug was the main metabolite reaching the lungs, not the prodrug[67]. This appears to render the whole purpose of the prodrug moiety pointless. However, it should be kept in mind that remdesivir was a repurposed drug and was originally designed to treat hepatitis C, therefore its kinetics were not optimised for delivery to the lungs.

When developing a prodrug, it is important to determine that the prodrug moiety, during both in vivo and in vitro trials, serves its purpose in improving drug delivery to the site, and that the prodrug is metabolised to the active drug selectively.

1.4.2 Unpredictable differences in metabolism between species

Drugs are tested on animal models first to establish their safety profile, and then once satisfactory results are given, human clinical trials can proceed. However, what works well on mice does not necessary translate into the same results on humans. This is due to differences in enzyme structure between species, and its expression in various tissues and organs.

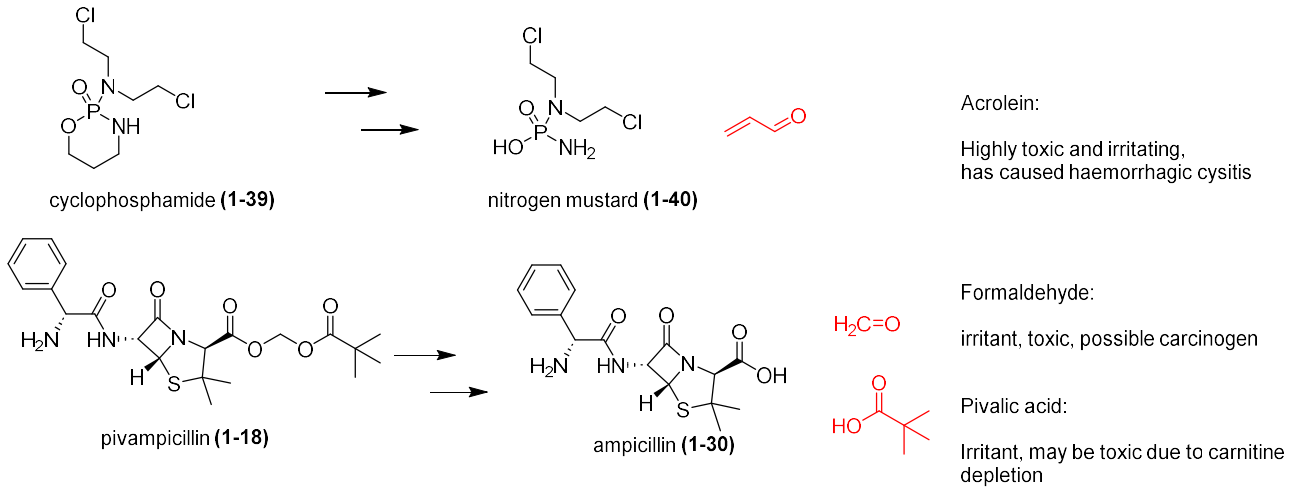
Oseltamivir **(1-3)**, an ethyl ester prodrug of anti-influenza drug GS4071, is a neuraminidase inhibitor. The esterification of the drug increases its lipophilicity and improves its oral bioavailability from 2-4% in the carboxylic acid to 35% for oseltamivir in humans. However, in mice it is 30%, in ferrets, 11%, and in dogs, 71%[68]. The same study also showed that in dogs, the prodrug persisted in plasma for much longer than in other species. A 30-minute incubation of the prodrug showed complete conversion in rat plasma, 10% and 15% conversion for ferret and human respectively, and negligible conversion in dogs[68].

This observation can be explained by the fact that the enzymes that hydrolyses esters, CES (carboxylesterase) enzymes, varies in expression significantly by species. Rat CES is highly expressed in blood, but not in other mammalian species. Humans express CES1 and CES2 highly in the liver, but only CES1 in the intestines. However, cynomolgus monkeys express both CES1 and CES2 in the intestines. [69, 70]

Care must be taken when extrapolating results from non-human animal models to humans, as the same drug may have significantly different pharmacokinetics between animal species due to the reasons described above.

1.4.3 Non-innocent side productions from prodrug conversion

When it comes to safety evaluations of prodrugs, not only should the toxicity of the prodrug and active drug be considered, but also the side products that are formed from the metabolization of prodrug moieties.



Scheme 1.39: Toxic metabolites of prodrug breakdown of 2 compounds of 2 compounds and their side effects

Cyclophosphamide **(1-39)** releases highly toxic acrolein on metabolism to the nitrogen mustard. It attacks the bladder epithelium and can cause haemorrhagic cystitis (bleeding from the bladder)[71]. As such, patients taking the drug are advised to take the drug intermittently to minimise the accumulation of acrolein.

Some prodrug functional groups have attracted concern over potential release of toxic molecules. The use of the α-acyloxyalkyl functional group, as shown from pivampicillin, has shown concern through the release of formaldehyde, a known irritant and carcinogen, despite reassurances on the contrary[72], given that the body produces formaldehyde endogenously and has ways to metabolise it.

Pivampicillin **(1-18)** is also a pivaloyl ester, another commonly used prodrug moiety of concern. It releases pivalic acid which then reacts with coenzyme A to form pivaloyl CoA, in a similar manner to how other carboxylic acids form acyl-CoAs. Catabolism of acyl-CoAs involves β-oxidation of acyl-CoA leading to CO_2 elimination and shortening of the carboxylic acid chain by two carbons. However, the tertiary β-carbon of pivalic acid means pivaloyl-CoA cannot be oxidised in this manner. Instead, the acyl group is transferred from coenzyme A to carnitine to form pivaloylcarnitine, which is then excreted in urine. As carnitine is synthesized and accumulated in the human body slowly, there are concerns that excessive consumption of high doses of pivampicillin (and other pivalic acid releasing prodrugs) can cause carnitine depletion, which can cause muscle and liver problems. However, animal and human studies have shown no symptoms resulting from high doses of pivampicillin administration in the short term, and any side effects that might occur on long term administration were nullified by taking carnitine supplements[73].

Many other prodrug functional groups have not been tested for its safety profile, and its possible effects on the human body remain unknown.

1.5 Conclusion

Many drugs that are biologically active to their targets *in vitro* have little to no activity *in vivo*, because in their inability to reach the desired targets due to unsuitable pharmacokinetic properties. Therefore, the prodrug approach has been successfully to overcome these obstacles. These prodrugs also have additional benefits such as lower toxicity, lower required dosage, and

easier administration. Instead of seeing prodrugs as a last-minute salvage attempt to improve a poorly performing drug, prodrugs are now often considered at an early stage of drug design, taking into account knowledge of the biological properties of the target enzyme as well as the environment around it, both intracellular and extracellular.

However, designing prodrugs is more intensive than other drugs, as additional investigation is required to examine the biochemistry of not just the drug itself, but also on how the promoieties undergo metabolism to release the active drug. In addition, manufacturing prodrugs require additional synthetic steps that add costs and time in development. Failure to consider these factors thoroughly will lead to the problems of high cost and a higher rate of failure in human clinical trials. Nevertheless, due to the immense potential in treating many difficult conditions, prodrugs as a strategy will remain. The success of future prodrugs will depend on improving our knowledge of pharmacology, biology and chemistry surrounding these drugs and targets.

The use of a prodrug strategy will be relevant in the development and investigation of a potential antitubercular drug, EpMF2, in our group. EpMF2 is an ATP synthase inhibitor and has shown notable nanomolar activity *in vitro* studies on inverted membrane vesicles, but had no inhibitory effect *in vivo* on live *M.smegmatis* bacteria. It was hypothesized that this was due to the inability of EpMF2 to penetrate the cell membrane of live bacteria to reach its target in the ATP synthase, which was located inside the cytoplasm. The use of a cholesterol-based moiety as a prodrug to penetrate the cell wall of mycobacteria will be elaborated in Chapter 3.

1.6 References

1. Wermuth, C. G.; Ganellin, C. R.; Lindberg, P.; Mitscher, L. A., *Pure and Applied Chemistry* **1998,** *70* (5), 1129-1143.

2. Inturrisi, C. E.; Schultz, M.; Shin, S.; Umans, J. G.; Angel, L.; Simon, E. J., *Life Sciences* **1983,** *33*, 773-776.

3. Wu, K. M., *Pharmaceuticals (Basel)* **2009,** *2* (3), 77-81.

4. Rautio, J.; Meanwell, N. A.; Di, L.; Hageman, M. J., *Nat Rev Drug Discov* **2018,** *17* (8), 559-587.

5. Lipinski, C. A., *Drug Discov Today Technol* **2004,** *1* (4), 337-341.

6. Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J., *J Comb Chem* **1999,** *1* (1), 55-68.

7. Egan, W. J.; Lauri, G., *Adv Drug Deliv Rev* **2002,** *54* (3), 273-289.

8. Estudante, M.; Morais, J. G.; Soveral, G.; Benet, L. Z., I *Adv Drug Deliv Rev* **2013,** *65* (10), 1340-1356.

9. Binghe Wang, T. S., Richard Solterro, *Drug Delivery - Principles and Applications*. John Wiley & Sons: Hoboken, New Jersey, 2005.

10. Beaumont, K.; Webster, R.; Gardner, I.; Dack, K., *Curr Drug Metab* **2003,** *4* (6), 461-485.

11. DailyMed LABEL: TERBUTALINE SULFATE tablet. https://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid=382dcebb-6e41-4335-818d-b5822e3fb884 (accessed 20 Dec).

12. A. Tunek, L. A. S., *Drug Metabolism and Disposition* **1988,** *16* (5), 759-764.

13. Persson, G.; Baas, A.; Knight, A.; Larsen, B.; Olsson, H., *Eur Respir J* **1995,** *8* (1), 34-39.

14. Lindberg, P.; Brändström, A.; Wallmark, B., *Trends in Pharmacological Sciences* **1987,** *8* (10), 399-402.

15. Maton, P. N., *N Engl J Med* **1991,** *324* (14), 965-975.

16. Ahmad, S. I.; Kirk, S. H.; Eisenstark, A., *Annu Rev Microbiol* **1998,** *52*, 591-625.

17. Longley, D. B.; Harkin, D. P.; Johnston, P. G., *Nat Rev Cancer* **2003,** *3* (5), 330-338.

18. Walko, C. M.; Lindley, C., *Clin Ther* **2005,** *27* (1), 23-44.

19. Reigner, B.; Blesch, K.; Weidekamm, E., *Clin Pharmacokinet* **2001,** *40* (2), 85-104.

20. Chau, C. H.; Steeg, P. S.; Figg, W. D., *Lancet* **2019,** *394* (10200), 793-804.

21. Younes, A.; Yasothan, U.; Kirkpatrick, P., *Nat Rev Drug Discov* **2012,** *11* (1), 19-20.

22. Doronina, S. O.; Toki, B. E.; Torgov, M. Y.; Mendelsohn, B. A.; Cerveny, C. G.; Chace, D. F.; DeBlanc, R. L.; Gearing, R. P.; Bovee, T. D.; Siegall, C. B.; Francisco, J. A.; Wahl, A. F.; Meyer, D. L.; Senter, P. D., *Nat Biotechnol* **2003,** *21* (7), 778-784.

23. Sharma, S. K.; Bagshawe, K. D., *Adv Drug Deliv Rev* **2017,** *118*, 2-7.

24. Bagshawe, K. D., *Expert Rev Anticancer Ther* **2006,** *6* (10), 1421-1431.

25. Zhang, J.; Kale, V.; Chen, M., *AAPS J* **2015,** *17* (1), 102-110.

26. Ettmayer, P.; Amidon, G. L.; Clement, B.; Testa, B., *J Med Chem* **2004,** *47* (10), 2393-2404.

27. Lavis, L. D., *ACS Chem Biol* **2008,** *3* (4), 203-206.

28. Ehrnebo, M.; Nilsson, S. O.; Boreus, L. O., *J Pharmacokinet Biopharm* **1979,** *7* (5), 429-451.

29. Easthope, S. E.; Jarvis, B.,. *Drugs* **2002,** *62* (8), 1253-1287.

30. Schultz, C., *Bioorg Med Chem* **2003,** *11* (6), 885-898.

31. Eastman, R. T.; Roth, J. S.; Brimacombe, K. R.; Simeonov, A.; Shen, M.; Patnaik, S.; Hall, M. D., *ACS Cent Sci* **2020,** *6* (5), 672-683.

32. Hughes, D. W.; Wilson, W. L.; Butterfield, A. G.; Pound, N. J., *J Pharm Pharmacol* **1974,** *26* (1), 79-80.

33. Faine, S.; Harper, M., *Antimicrob Agents Chemother* **1973,** *3* (1), 15-18.

34. Kuhl, H., *Climacteric* **2005,** *8 Suppl 1*, 3-63.

35. Guengerich, F. P., *Chem Res Toxicol* **2008,** *21* (1), 70-83.

36. Bertilsson, L.; Tomson, T., *Clin Pharmacokinet* **1986,** *11* (3), 177-198.

37. Varfaj, F.; Zulkifli, S. N.; Park, H. G.; Challinor, V. L.; De Voss, J. J.; Ortiz de Montellano, P. R., *Drug Metab Dispos* **2014,** *42* (5), 828-38.

38. Boddy, A. V.; Yule, S. M., *Clin Pharmacokinet* **2000,** *38* (4), 291-304.

39. Erion, M. D.; van Poelje, P. D.; Mackenna, D. A.; Colby, T. J.; Montag, A. C.; Fujitaki, J. M.; Linemeyer, D. L.; Bullough, D. A., *J Pharmacol Exp Ther* **2005,** *312* (2), 554-560.

40. Jing, X.; Yang, F.; Shao, C.; Wei, K.; Xie, M.; Shen, H.; Shu, Y., *Mol Cancer* **2019,** *18* (1), 157.

41. Tomasz, M., *Chem Biol* **1995,** *2* (9), 575-9.

42. Bass, P. D.; Gubler, D. A.; Judd, T. C.; Williams, R. M., *Chem Rev* **2013,** *113* (8), 6816-6863.

43. Casner, M. L.; Remers, W. A.; Bradner, W. T., *J Med Chem* **1985,** *28* (7), 921-926.

44. Phillips, R. M.; Loadman, P. M.; Cronin, B. P., *Br J Cancer* **1998,** *77* (12), 2112-2119.

45. Phillips, R. M.; Hendriks, H. R.; Peters, G. J.; *Br J Pharmacol* **2013,** *168* (1), 11-18.

46. Tanabe, K.; Makimura, Y.; Tachi, Y.; Imagawa-Sato, A.; Nishimoto, S., *Bioorg Med Chem Lett* **2005,** *15* (9), 2321-2324.

47. Zhang, Z.; Tanabe, K.; Hatta, H.; Nishimoto, S., *Org Biomol Chem* **2005,** *3* (10), 1905-1910.

48. de Groot, F. M.; Damen, E. W.; Scheeren, H. W., *Curr Med Chem* **2001,** *8* (9), 1093-1122.

49. Chen, Y.; Hu, L., *Med Res Rev* **2009,** *29* (1), 29-64.

50. Helsby, N. A.; Ferry, D. M.; Patterson, A. V.; Pullen, S. M.; Wilson, W. R., *Br J Cancer* **2004,** *90* (5), 1084-1092.

51. Tercel, M.; Wilson, W. R.; Anderson, R. F.; Denny, W. A., *J Med Chem* **1996,** *39* (5), 1084-1094.

52. Tercel, M.; Lee, A. E.; Hogg, A.; Anderson, R. F.; Lee, H. H.; Siim, B. G.; Denny, W. A.; Wilson, W. R., *J Med Chem* **2001,** *44* (21), 3511-3522.

53. Denny, W. A., *Curr Pharm Des* **2002,** *8* (15), 1349-1361.

54. Parveen, I.; Naughton, D. P.; Whish, W. J.; Threadgill, M. D., *Bioorg Med Chem Lett* **1999,** *9* (14), 2031-2036.

55. Berry, J. M.; Watson, C. Y.; Whish, W. J. D.; Threadgill, M. D., *Journal of the Chemical Society, Perkin Transactions 1* **1997,** (8), 1147-1156.

56. Hay, M. P.; Anderson, R. F.; Ferry, D. M.; Wilson, W. R.; Denny, W. A., *J Med Chem* **2003,** *46* (25), 5533-5345.

57. Duan, J. X.; Jiao, H.; Kaizerman, J.; Stanton, T.; Evans, J. W.; Lan, L.; Lorente, G.; Banica, M.; Jung, D.; Wang, J.; Ma, H.; Li, X.; Yang, Z.; Hoffman, R. M.; Ammons, W. S.; Hart, C. P.; Matteucci, M., *J Med Chem* **2008,** *51* (8), 2412-2420.

58. Thomson, P.; Naylor, M. A.; Everett, S. A.; Stratford, M. R.; Lewis, G.; Hill, S.; Patel, K. B.; Wardman, P.; Davis, P. D., *Mol Cancer Ther* **2006,** *5* (11), 2886-2894.

59. Zeman, E. M.; Brown, J. M.; Lemmon, M. J.; Hirst, V. K.; Lee, W. W., *Int J Radiat Oncol Biol Phys* **1986,** *12* (7), 1239-1242.

60. Daniels, J. S.; Gates, K. S., *Journal of the American Chemical Society* **1996,** *118* (14), 3380-3385.

61. Anderson, R. F.; Shinde, S. S.; Hay, M. P.; Gamage, S. A.; Denny, W. A., *J Am Chem Soc* **2003,** *125* (3), 748-756.

62. Reddy, S. B.; Williamson, S. K.,. *Expert Opin Investig Drugs* **2009,** *18* (1), 77-87.

63. Smith, P. J.; Blunt, N. J.; Desnoyers, R.; Giles, Y.; Patterson, L. H., *Cancer Chemother Pharmacol* **1997,** *39* (5), 455-61.

64. Tredan, O.; Garbens, A. B.; Lalani, A. S.; Tannock, I. F., *Cancer Res* **2009,** *69* (3), 940-947.

65. Smith, J. T.; Hamilton-Miller, J. M., *Chemotherapy* **1970,** *15* (6), 366-378.

66. Langreth, R., All Eyes on Gilead. *Bloomberg Businessweek* 2020.

67. Yan, V. C.; Muller, F. L., *ACS Med Chem Lett* **2020,** *11* (7), 1361-1366.

68. Li, W.; Escarpe, P. A.; Eisenberg, E. J.; Cundy, K. C.; Sweet, C.; Jakeman, K. J.; Merson, J.; Lew, W.; Williams, M.; Zhang, L.; Kim, C. U.; Bischofberger, N.; Chen, M. S.; Mendel, D. B., *Antimicrob Agents Chemother* **1998,** *42* (3), 647-653.

69. Hosokawa, M., *Molecules* **2008,** *13* (2), 412-431.

70. Nishimuta, H.; Houston, J. B.; Galetin, A., *Drug Metab Dispos* **2014,** *42* (9), 1522-1531.

71. Emadi, A.; Jones, R. J.; Brodsky, R. A.,. *Nat Rev Clin Oncol* **2009,** *6* (11), 638-647.

72. Dhareshwar, S. S.; Stella, V. J., *J Pharm Sci* **2008,** *97* (10), 4184-4193.

73. Brass, E. P., *Pharmacol Rev* **2002,** *54* (4), 589-598.

CHAPTER 2

Discovery and investigation of antitubercular drug candidate EpMF2

2 Discovery and investigation of antitubercular drug candidate EpMF2

2.1 Abstract

In 2012, bedaquiline (BDQ) was approved for use in the treatment of multidrug-resistant tuberculosis. The compound has a novel mechanism through inhibition of the ATP synthase. The compound binds primarily to the c subunit which inhibits the rotational movement necessary for ATP synthesis, but BDQ also binds to the ε subunit which enhances the inhibitory effect on ATP generation. However, there were concerns over potential severe side effects, and such worries have led to the push for discovery of safer analogues of BDQ. The compound EpMF2 was discovered as a potential hit through *in silico* screening as an inhibitor of the ε subunit of mycobacterial ATP synthase. Synthesis of the compound confirmed its inhibitory activity of ATP synthase on inverted membrane vesicles (IC_{50} IMV = 432nM), but it did not appear to be effective on inhibiting live mycobacterial growth. Describe within this chapter was the synthesis of analogues of EpMF2 to investigate its structure activity relationship. While some analogues have comparable IC_{50} activity on IMVs, all of them suffer the same problem from a lack of activity against living *M.smegmatis*.

2.2 Introduction

2.2.1 Introduction to tuberculosis

Tuberculosis (TB) is a disease primarily affecting primarily the lungs. Symptoms include coughing (sometimes with blood), weight loss, fatigue, and chest pain, among others. It is caused by the bacterium *Mycobacterium tuberculosis* (*Mtb*), a non-motile, aerobic, and obligate parasitic species. In 2019, there were 7.1 million reported cases of TB, falling to 5.8 million cases in 2020 before rebounding to 6.4 million in 2021.[1] It is highly certain that the fall in numbers were due to underreporting as a result from the COVID-19 pandemic. The WHO

estimated in 2022 that there were 10.9 million actual cases in 2021, along with 1.6 million deaths.[1] This makes tuberculosis the second leading cause of death from an infectious pathogen after COVID-19 in 2021 (with 3.5 million deaths that year).[2] The COVID pandemic has in fact negatively impacted the state of TB treatment, with evidence to suggest TB and COVID co-infected patients suffer from worse outcomes than COVID only patients.[3]

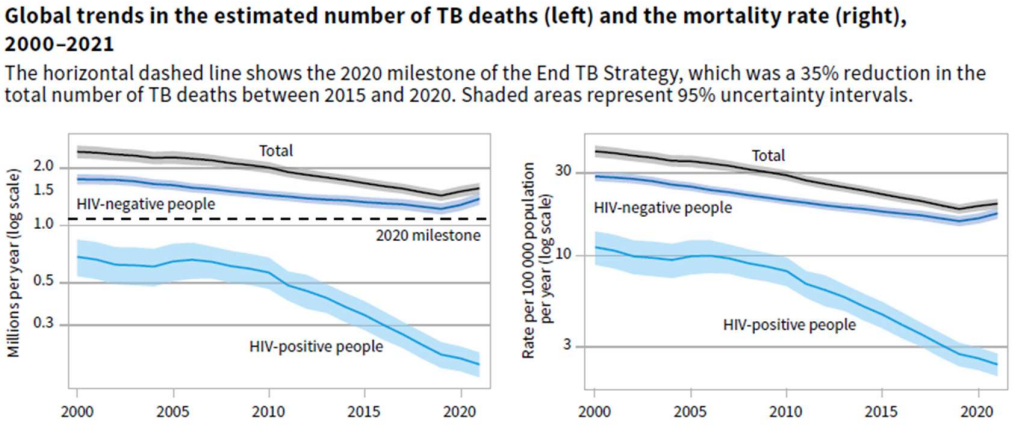


Figure 2.1: Numbers and proportion of TB cases worldwide from 2000 to 2020. The slight increase in numbers near 2020 is evident. Adapted from the WHO Tuberculosis report 2022[1]

Mycobacteria species, including *Mtb*, are characterized by their unusual cell wall. Their cell walls consist of mycolic acids, very long chain fatty acids from 60 to 90 carbon atoms long, and frequently containing ketone, methoxy and cyclopropane moieties.[4] This gives the Mycobacterial cell wall a thick, waxy character and this is believed to contribute to its hardiness and resistance towards dehydration and chemical attacks from immune cells.[5] This makes treatment of TB particularly difficult.

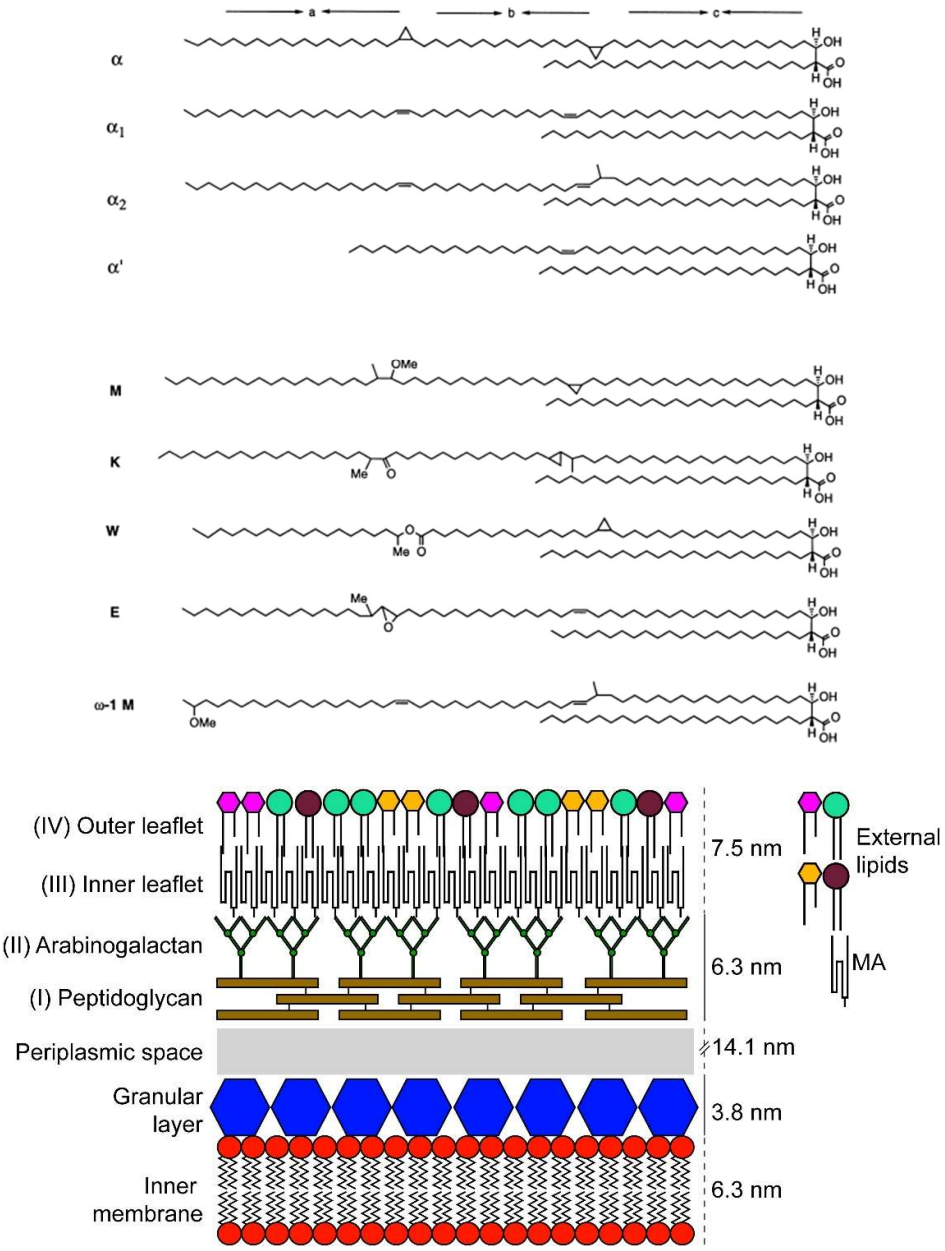


Figure 2.2: Top: A few examples of mycolic acid structures identified from Mycobacteria as described by Barry *et al.*[4] Bottom: Structure of the mycobacterial cell wall. In addition to mycolic acids (labelled MA), the cell wall also contains arabinogalactans and peptidoglycans common to other bacteria. Image adapted from Veyrier *et al.*[6]

Tuberculosis is treatable with antibiotics but requires a long regimen of multiple drugs. For drug susceptible TB, the recommended treatment as prescribed by WHO is the daily intake of a combination of isoniazid **(2-1)**, rifampicin **(2-2)**, pyrazinamide**(2-3)**, and ethambutol**(2-4)** for two months, followed by isoniazid and rifampicin daily for four more months.[7] These four drugs are known as first-line drugs.

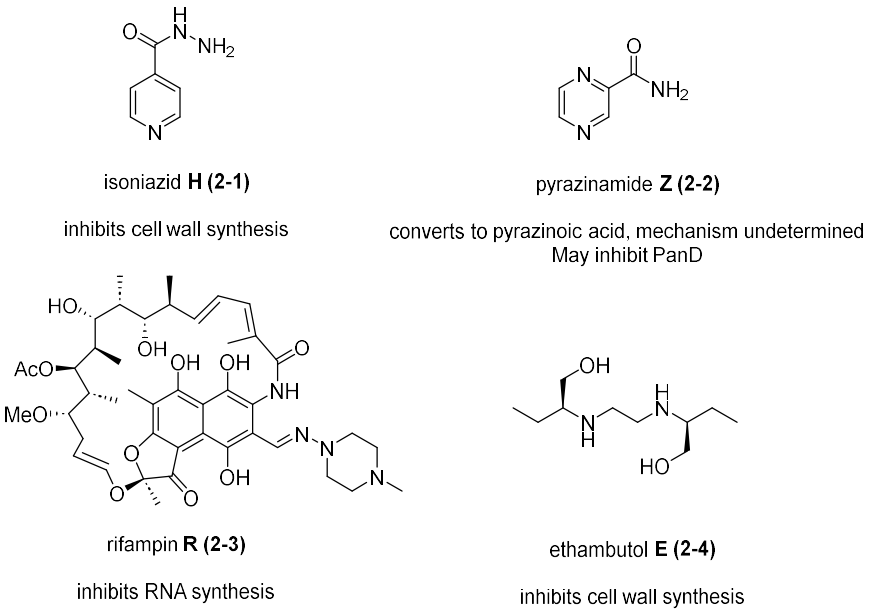


Figure 2.3: Molecular structures of the four first line drugs for the treatment of drug sensitive TB. Letters besides the drug names represent abbreviations used in medicine.

However not all patients complete their course of medication, due to its long length and in some developing countries, limited availability and high cost of medication. This contributes to antimicrobial resistance, and multidrug resistant TB (MDR-TB) is becoming a pervasive problem worldwide. MDR-TB is defined as strains that are resistant to at least rifampicin or isoniazid.[8]

If MDR-TB is suspected, second line drugs are administered instead. The WHO classifies these drugs into 3 groups depending on their mechanism of action. They include polypeptides and aminoglycosides (Group 2), fluoroquinolones (Group 3), and cycloserine and thioamides (Group 4). Compared to first line drugs, treatment of MDR-TB with second line drugs require a much greater duration at 18 months, and come with more serious side effects.[9]

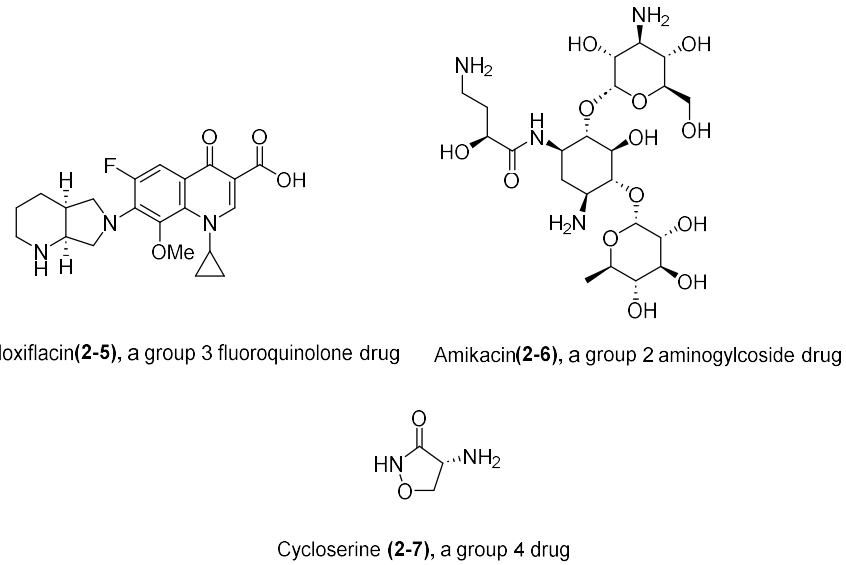


Figure 2.4: Molecular structures of some second line drugs used against multidrug resistant TB

If the strain shows resistance even to second-line drugs, it is labelled extensively drug resistant TB (XDR-TB). These strains require treatment with third line drugs (group 5 as defined by WHO), many of which are insufficiently tested, but may still be used if the use of first- and second-line drugs have failed. Until recently, development of these third line drugs were extremely limited, and it was only in the last decade that new drugs of this type were developed and approved for use for XDR-TB, including bedaquiline by the US Food and Drug Administration in 2012[10] and delamanid by the European Medicines Agency in 2014.[11] Below sections will focus exclusively on bedaquiline.

2.2.2 Bedaquiline, a novel drug for multidrug resistant TB

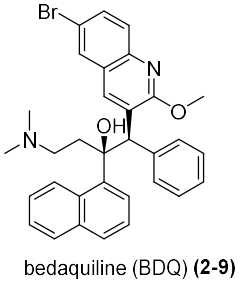


Figure 2.5: Chemical structure of bedaquiline (BDQ) **(2-9)**

Bedaquiline (**2-9**), or BDQ, belongs to a class of drugs known as the diarylquinolines. The compound was first described under the name R207910 by Andres *et al.*.[12] During their search

for diarylquinoline drugs, the compound was identified as a lead compound that is highly bactericidal against drug-sensitive and drug-resistant *Mtb* in vitro (MIC of 0.06μg/mL).The drug was highly specific to mycobacterial species, and appeared to be well tolerated in human and mice subjects. Through genomic analysis between BDQ vulnerable and resistant strains, they identified the F_o subunit of the mycobacterial ATP synthase as the target of the compound. This was a breakthrough in the discovery of new antimycobacterial drugs, as it kills the bacteria using a novel mode of action not used by currently existing drugs and is therefore effective against even MDR-TB and XDR-TB strains.

Further studies have established that BDQ targets the c subunit of the enzyme, through the interaction of the basic dimethylamine moiety of BDQ to the acidic carboxylate group of glutamate 61 of subunit c.[13] This hinders its proton motive force driven rotation of the c subunits along with the central stalk subunit γ. The rotation of the subunit γ induces conformational changes on the $\alpha^3\beta^3$ subunits that catalyses synthesis of ATP from ADP and phosphate. Therefore, the binding of BDQ to the enzyme effectively halts synthesis of ATP, and the bacteria dies.

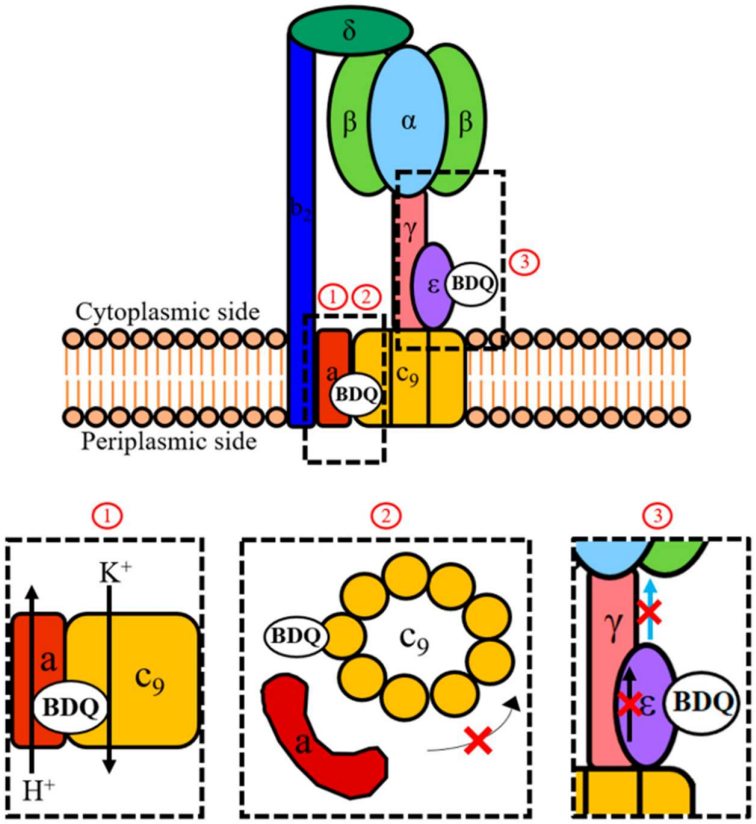


Figure 2.6: 3 Proposed mechanisms of BDQ inhibition of ATP synthase as outlined in Dick's paper: 1. Binding of BDQ to the c subunit stalls its rotation. 2. Binding to the ε subunit uncouples rotational movement of the c9γ subunits to ATP synthesis at α3β3. 3. BDQ acts as an ionophore to eliminate the proton motive force. The experiments performed by the authors suggests the dominant role of 1 and 2, while 3 appeared to play a minor role.

A more recent study from Dick et al. [14] summarized experiments uncovering the mechanism of BDQ in more detail. Grüber *et al*, discovered that BDQ not only exerts its inhibitory effects by binding to subunit c, but also to subunit ε.[15] The binding of BDQ to subunit ε furthers suppresses ATP formation by disrupting the ability to transfer rotational movement of the c subunit to conformational changes of the $\alpha^3\beta^3$ subunits required for ATP synthesis. Although the binding affinity for BDQ to subunit ε is relatively low (K_d = 19.1μM, compared to 500nM for subunit c), the binding of both subunits synergises each other in its inhibitory ability. A study by Noji[16] used the compound dicyclohexylcarbodiimide (DCC) on *E coli*. to inhibit the c subunit rotation in a similar manner to BDQ on *Mtb*. He found that rotation was not completely stopped, and transient rotations are still possible. If a similar phenomenon happens in BDQ affected *Mtb*, it was hypothesized that the binding of the ε subunit uncouples the

occasional c subunit rotation from ATP synthesis at the $\alpha^3\beta^3$ subunit. Therefore, binding of both c and ε subunits of ATP synthase are required for effective inhibition.

BDQ was also hypothesized to inhibit mycobacterial ATP synthesis through acting as a H^+/K^+ ionophore from experiments conducted by Cook *et al*.[17] Ionophores are substances that complexes to ions and aid its transport across the cell lipid membrane. They act to eliminate the concentration difference of protons on both sides of the membrane (ΔpH), in effect abolishing the proton motive force that is necessary for the activity of ATP synthase. The authors hypothesize that BDQ transports H^+ by bypassing the cell wall in its protonated form using its basic amine moiety, and transports K^+ in the opposite direction by forming a poorly defined chelation complex separate from the amine H^+ binding site. However, Dick *et al's* experiments with a related analogue TBAJ-876 did not show similar activity on *Mtb* ATP synthase, [14] likely because TBAJ-876 is more polar in structure and is reluctant to pass the non-polar cell membrane. This suggests that this mode of action is not fundamental to BDQ's bactericidal activity.

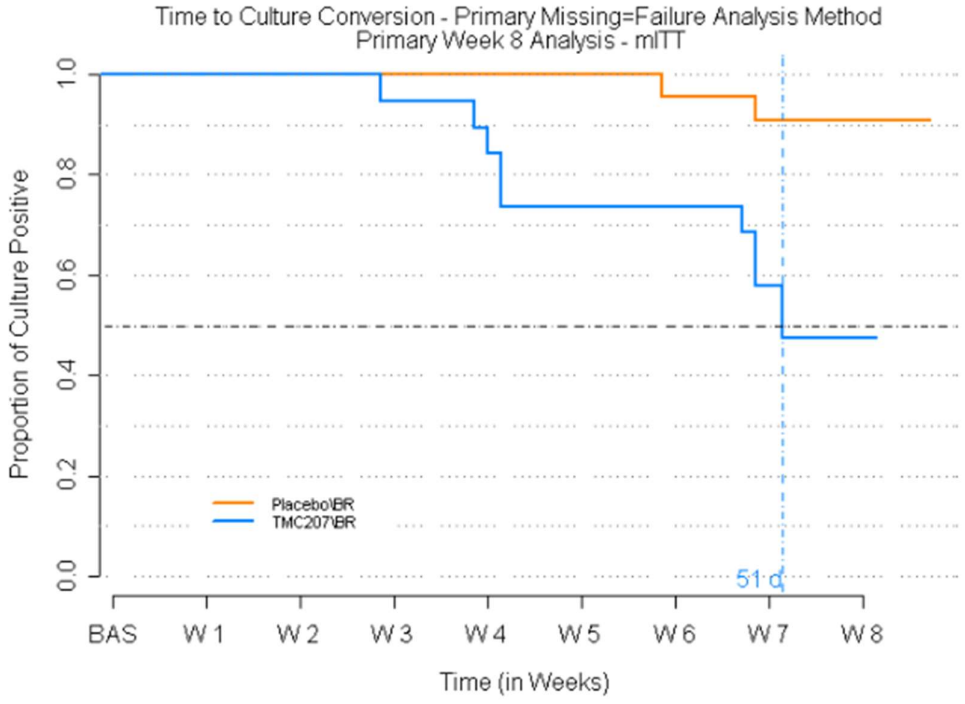
2.2.3 *In vivo* trials of BDQ

A study by Ordway *et al*. [18]on the use of BDQ in the treatment of *Mtb* infected guinea pigs showed some promising results. Guinea pigs treated with a combined regimen of BDQ, rifampicin and pyrazinamide showed undetectable bacterial loads in the lungs 8 weeks after treatment. The lungs of treated animals also showed some healing of lesions, as compared to those of untreated animals, where lesions were more numerous and showed signs of necrosis. However, relapse occurred in both the BDQ taking and non BDQ taking animals a year after infection at similar levels.

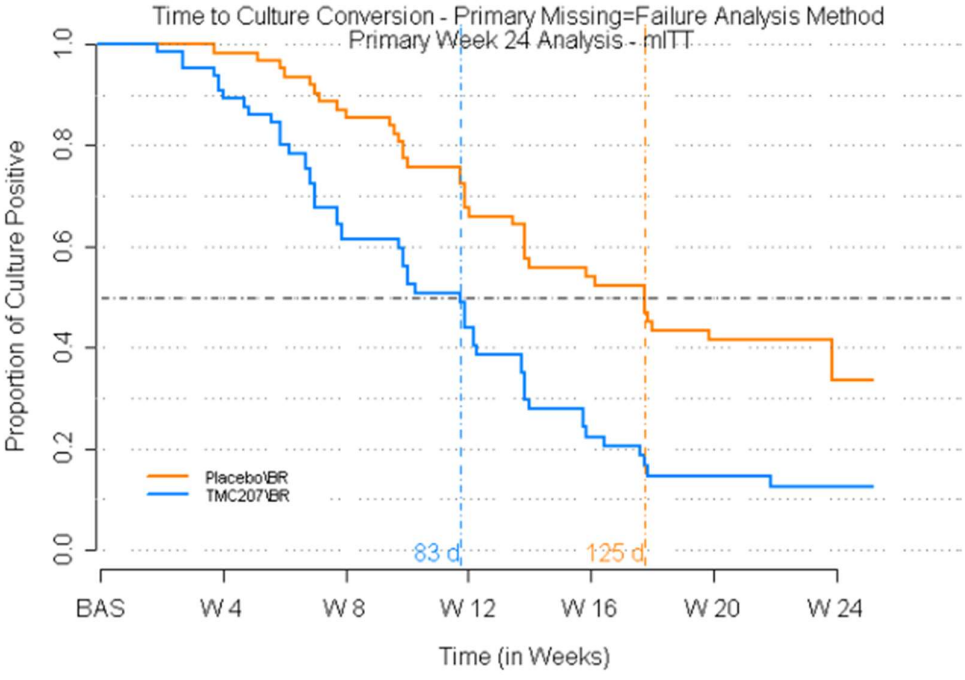
On the use on humans, two phase IIb studies have been conducted on humans to determine the clinical efficiency of BDQ. The first, labelled C208, was a 2-stage phase of double-blind trials where patients taking BDQ along with a background regimen of second-line drugs were

compared with patients taking placebo with the same background treatment.[19] These patients were newly infected (not caused by a relapse) with MDR-TB with known resistances to isoniazid and rifampicin. In the first stage, lasting for 8 weeks, patients in the BDQ group showed statistically significant higher rates of sputum culture conversion (defined as two consecutive negative tests for *Mtb* taken at least 30 days apart) at 47.6% than patients in the placebo group at 8.7%.[20] In the second stage, lasting 24 weeks, the median time to sputum culture conversion was reduced from 125 days in the placebo group compared to 83 days in the BDQ group. Patients in the BDQ group showed statistically significant higher rates of sputum culture conversion at 79%, as compared to those on the placebo group at 58%. [21]

During the 120-week duration period after administration for the stage 2 trial, there was a statistically significant increase in QT prolongation on the BDQ group compared to the placebo group. Some classes of drugs can cause QT prolongation,[22] where the heart takes longer to charge up between beats. This can increase the risk of tachycardia (fast heart rate) which can even lead to complications like cardiac arrest in severe cases. More worrying, the morality rate was higher for the BDQ group (10 out of 79, 12.7%) compared to the placebo group (2 out of 81, 2.5%.[21] However, none of the deaths could be attributed to any side effects of the drug, and was no correlation between QTcF duration between patients who died and those who did not.



# Placebo\BR	23	23	23	22	22	21	20	4
# TMC207\BR	21	19	18	16	14	14	11	1



# Placebo\BR	61	53	40	30	22	5
# TMC207\BR	58	37	25	12	7	3

Figure 2.7: Top: Proportion of culture positive patients over 8 weeks in the 1st stage of clinical trial C208. Orange represents those on placebo and background regimen, and blue those on BDQ (labelled TMC207) and background regimen. Vertical line at 51 days represents mean time to sputum culture conversion for the BDQ group. Bottom: Proportion of culture positive patients over 24 weeks in the 2nd stage of clinical trial C208. Orange represents those on placebo and background regimen, and blue those on BDQ (labelled TMC207) and background regimen. Vertical line at 83 and 125 days respectively represent mean time to sputum culture conversion for the BDQ and the background group. Image adapted from Janssen Pharmaceutical Company's report. [19]

The other trial, labelled C209, was a single arm, open label study of 233 patients with MDR and XDR TB. After taking BDQ, the observed median time to sputum culture conversion was 57 days, and 80% or patients achieved sputum culture conversion by week 24.[13] For patients with MDR-TB, sputum culture conversion rates were higher at 95% (88 out of 93). For pre XDR-TB, the rate was 81% (36 out of 44) while XDR-TB patients had even lower, but still impressive, rates at 69% (25 out of 36).[19] Morality rate over the study was at 7% (16 out of 233).

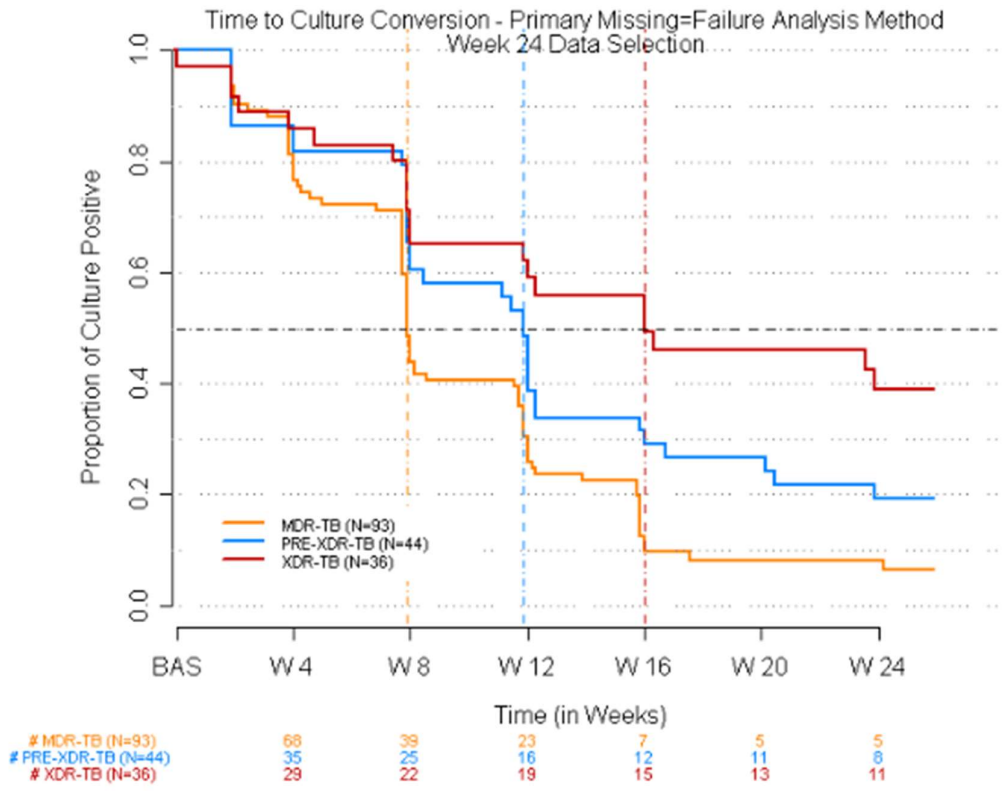


Figure 2.8: Proportion of culture positive patients over 24 weeks in the clinical trial C209. Orange represents MDR-TB cases, blue pre-MDR-TB, and red XDR-TB. Image adapted from Janssen Pharmaceutical Company's report. [19]

Study	Mortality (BDQ group)	Cause of death (BDQ group)	Mortality (placebo group)	Cause of death (placebo group)
C208 Stage 1	2/23 (8.7%)	Acute myocardial infarction, pulmonary tuberculosis	2/24 (8.3%)	TB-related illness, pulmonary tuberculosis
C208 Stage 2	10/79 (12.7%)	Alcohol poisoning, tuberculosis (2 patients), hepatitis/hepatic cirrhosis, septic shock/peritonitis, cerebrovascular accident, motor vehicle accident, TB-related illness (3 patients)	2/81 (2.5%)	Haemoptysis, TB-related illness
C209	16/233 (6.9%)	Unknown	NA	NA

Table 2.9: Mortality of patients during clinical phase trials C208 and C209 using BDQ. The cause of deaths of these patients were also ascertained. Although the BDQ group in the 2nd stage of study C208 had a relatively high mortality rate, none of the deaths could be directly attributed to side effects of the drug.

Considering concerns over the potential cardiotoxicity of BDQ (as well as other unpleasant side effects reported by patients in the above studies), analogues of BDQ with a better safety profile are being investigated and synthesized, as elaborated below.

2.2.4 The discovery of EpMF2 as a ε subunit inhibitor of *Mtb* ATP synthase

As mentioned previously, BDQ exerts its effects not only through binding of subunit c, but also subunit ε. NMR titration experiments of the subunit ε with BDQ by Grüber *et al*, have demonstrated that the molecule's most important interaction with the enzyme is to tryptophan 16 of the subunit near the *N*-terminal.[23] This binding interaction was further confirmed through tryptophan fluorescence spectroscopy. The measured emission spectrum of the tryptophan moiety (which is the only one in the subunit) showed that upon addition of BDQ, the emission maximum reduced in intensity and shifted to a longer wavelength.[24] Using computational-based modelling with AutoDock Vina, a binding model of BDQ to subunit ε was suggested as briefly described [23]:

- The naphthalene moiety of BDQ faces the aromatic ring of tryptophan 16 and π-π interactions creates a strong interaction between them. There is an additional stabilizing force from adjacent serine 17.
- The nitrogen atom in the quinoline moiety has a hydrogen bond interaction with arginine 13 while the ring has hydrophobic interactions with alanine 84.
- The phenyl moiety has an off-set face to face interaction with tryptophan 16 and edge-face interactions with phenylalanine 86.
- The nitrogen atom in the dimethylamino moiety has a hydrogen bond interaction with asparagine 14.
- The tertiary hydroxyl group has a hydrogen bond interaction with isoleucine 15.

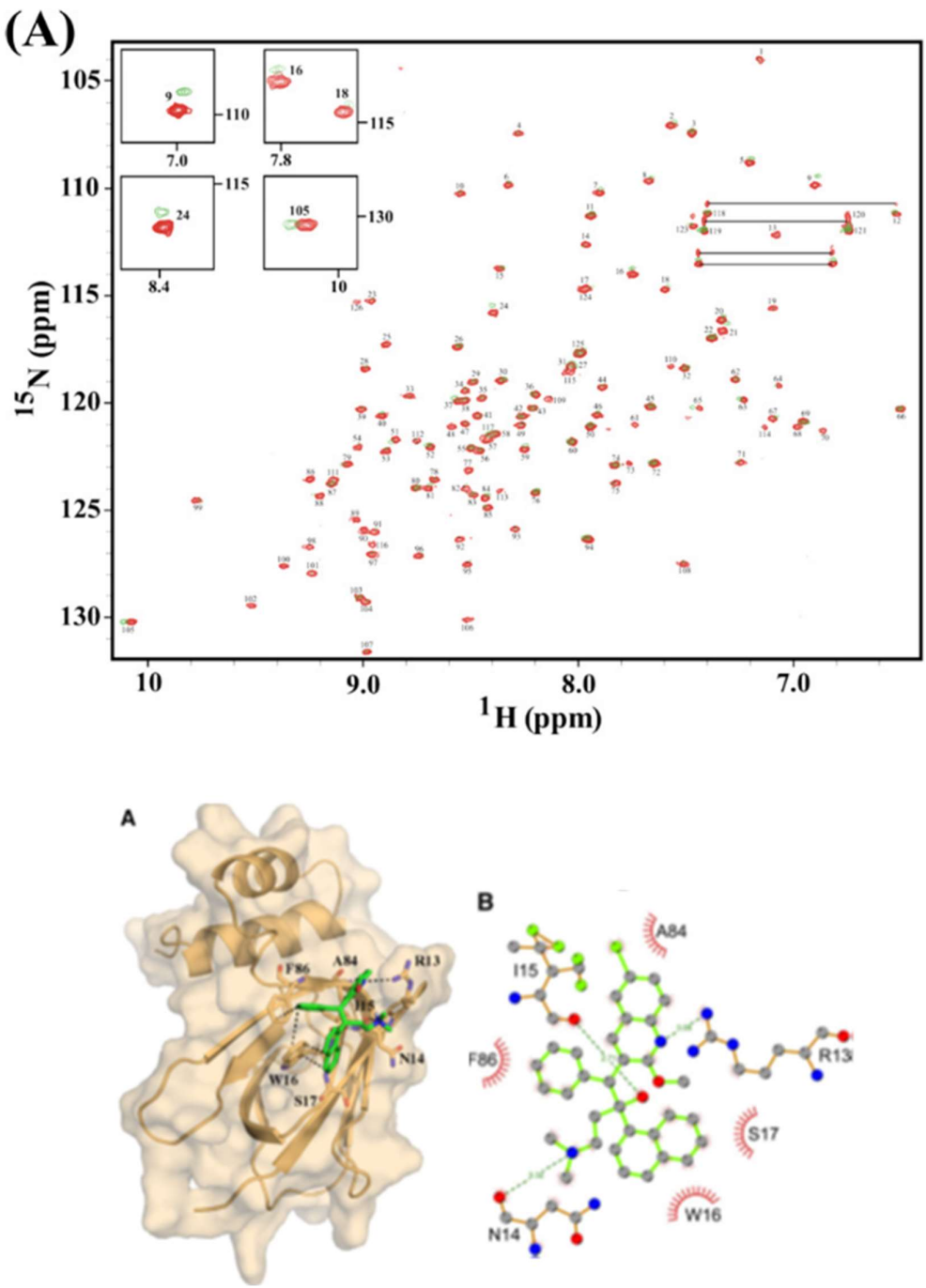
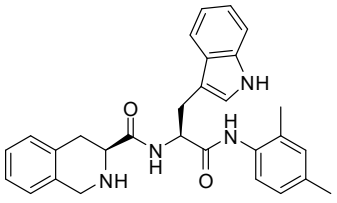


Figure 2.10: Top: 2D 1H-15N HSQC spectra of subunit ε without added BDQ (red) and with 2 equivalents of BDQ (green). Each peak represents a N-H amide bond corresponding to one amino acid. Spots with significant change in chemical shifts (corresponding to amino acids 9, 16, 18, 24 and 105)between red and green indicate regions with significant binding to the drug. Image adapted from Grüber et al[24] Bottom: Computer aided modelling of BDQ binding to subunit ε based on data from above NMR titration spectroscopy. A is a space filling model of enzyme with BDQ bound to the active site. B represents schematically specific interactions of important amino acids to various moieties in BDQ. Image adapted from same author on another paper[23]

Subsequent computational chemistry work was done by Dr Harikrishore, also in Grüber's group. Based on ligand based pharmaceutical modelling derived from the above model of BDQ binding to subunit ε, a structural-based virtual screen was conducted to discover other molecules with a similar binding profile. A library of compounds was created following this in silico study.

Recent advances with computing technology have allowed potential compounds to be screened virtually through computational analysis. The use of virtual screening, as opposed to testing with actual compounds, allowed a much wider variety of chemical structures to be tested compared to the past. The contemporary method of real screening, high throughput screening, requires a library of large amounts of compounds, and is highly labour intensive. On the other hand, the effectiveness of virtual screening methods is limited by the accuracy of the models used to simulate ligand-enzyme binding.

One compound, labelled EpMF2 was discovered as a hit. It is a dipeptide, with a non-proteogenic amino acid tetrahydroisoquinoline 3-carboxylic acid at the *N*-terminus, followed by a tryptophan. To the *C*-terminus was attached a 2,4-dimethylphenyl moiety through another amide bond.



EpMF2 **(2-10)**
IC_{50} M. smegmatis IMV ATP synthase = 432nM

Figure 2.11: Molecular structure of EpMF2, discovered by virtual screening and its inhibitory activity on the ε subunit mycobacterial ATP synthase

The above compound is calculated to bind to similar amino acids as BDQ in the *Mtb* ε subunit.

- Hydrogen bonding between the tetrahydroisoquinoline N-H and adjacent amide N-H with the peptide main chain amide oxygen on isoleucine 15.

- π- π interactions between 2,4 dimethyl phenyl ring and the side chain of phenylalanine 86.
- Hydrophobic interaction between the indole of tryptophan and isoleucine 15.
- Hydrophobic interaction between the carbon atoms of tetrahydroisoquinoline and tryptophan 16.

Through a collaboration, the compound was synthesized in our group by Dr. Patcharaporn Sae-Lao, then the analogues were tested in an ATP synthesis assay by Dr Priya Ragunathan from Grüber's group.

The assay was condcucted on micro well plates, each containing a stock solution containing ADP, phosphate, NADH and buffer along with varying concentrations of the EpMF2 compound. A solution of *M. smegmatis* IMVs in fixed concentration containing the ATP synthase was added and incubated for 30 minutes, followed by addition of CellTiter-Glo reagent, and then further incubated in the dark for 10 minutes. In the presence of ATP, the luciferin in the CellTiter-Glo reagent will emit light in a bioluminescent reaction, and its intensity is proportional to the concentration of ATP present. Thus, the intensity of emitted light is correlated with the level of inhibition of ATP synthase. The intensities of emitted lights were measured and then converted to the equivalent in ATP synthase inhibition. By measuring inhibition across a range of concentrations, the mean value IC_{50}, which is the concentration of drug where half of ATP synthase was inhibited, was calculated.

EpMF2 was found to inhibit ATP synthase of a related mycobacterial species *M. smegmatis.* The IC_{50} value as measured on *M. smegmatis* inverted membrane vesicles was 432nM. However, it was found to have no inhibitory activity against live *M. smegmatis.* up to 2μM.

2.2.5 Previous work of syntheses of EpMF2 analogues

Further work was done by Dr. Patcharaporn Sae-Lao to synthesize analogues of EpMF2 to determine its structural activity relationship. These analogues can be broadly divided into 3

groups. Series 1 analogues involve altering the right most 2,4-dimethylphenyl ring. Series 2 analogues involve altering the left side tetrahydroisoquinoline moiety and Series 3 the middle tryptophan moiety. There are several analogues that do not fit neatly into the above 3 categories and were classified in a miscellaneous group.

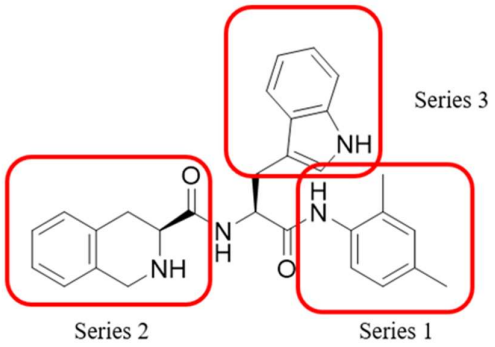
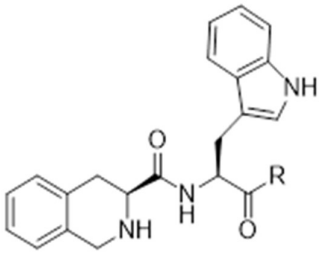
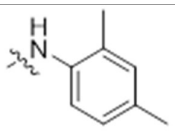
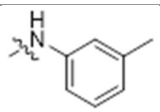
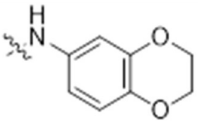
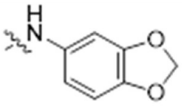
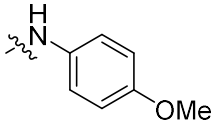
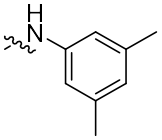
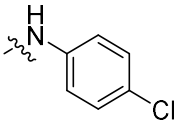
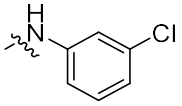
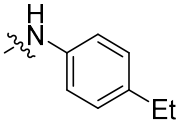
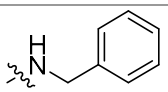
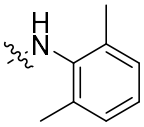
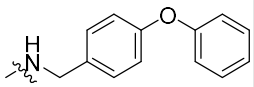
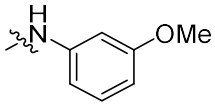
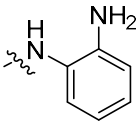
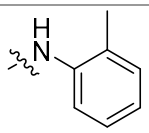
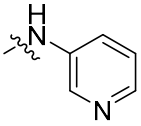
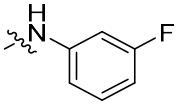
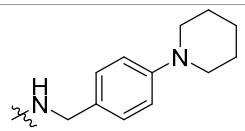
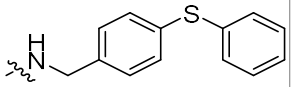


Figure 2.12: Previous work was done to synthesise 3 different sets of analogues of EpMF2 by altering different parts of the molecule (highlighted in red).

SERIES 1 ANALOGUES



Analogue	R	IC_{50} (μM)	Analogue	R	IC_{50} (μM)
(EpMF2) **2-10**		0.432	**2-10a**		0.542
2-10b		0.532	**2-10c**		1.40

2-10d		0.464	**2-10e**		1.80
2-10f		0.547	**2-10g**		0.347
2-10h		0.628	**2-10i**		3.40
2-10j		no inhibition	**2-10k**		0.482
2-10l		10.4	**2-10m**		6.10
2-10n		14.9	**2-10o**		5.30
2-10p		1.76	**2-10q**		2.26
2-10r		1.90	**2-10s**	-OH	no inhibition
2-10t	-OMe	26.9			

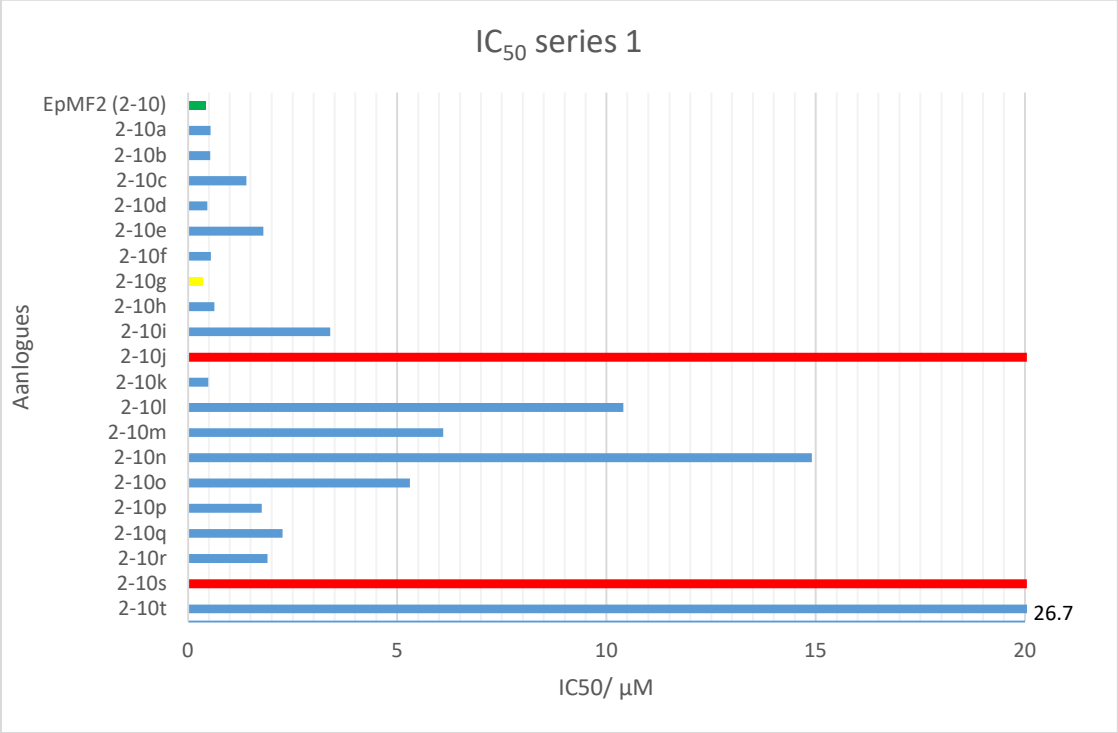
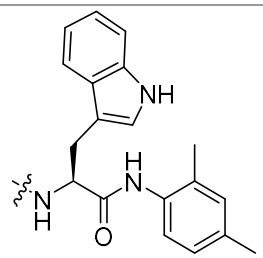


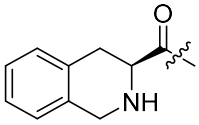
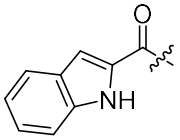
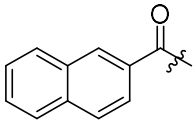
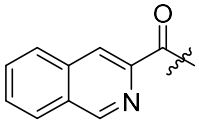
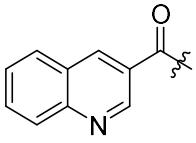
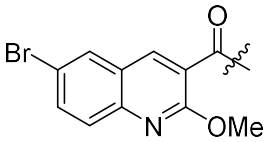
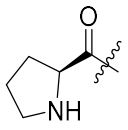
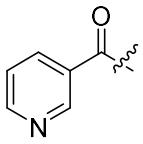
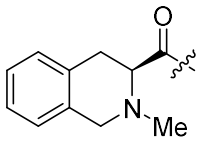
Table 2.13: IC_{50} Activity of series 1 analogues on M. smegmatis ATP synthase. On the graph, green represents EpMF2. Red bars represent no measurable inhibition up to 2mM. Yellow indicates the best performing analogue in its series.

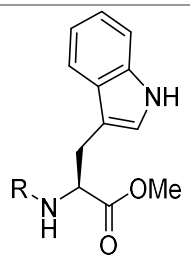
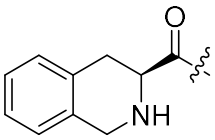
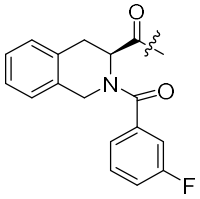
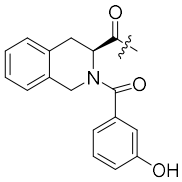
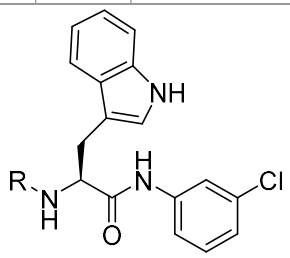
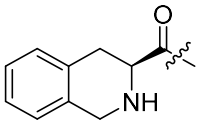
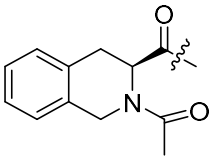
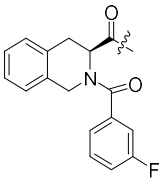
From Series 1, a wide variety of substituents were tolerated on the phenyl ring. Electron donating groups such as alkyl and ether groups (**2-10a, 2-10e, 2-10h**) showed a similar magnitude of activity as electron withdrawing halogen groups **(2-10f, 2-10g, 2-10p**). It appeared that substituents on the meta and para positions (**2-10a, 2-10h**) are associated with good binding activity regardless of the electronic nature of the substituent. Even bulky bicyclic substituents (**2-10k, 2-10q, 2-10r**) had decent activity. The only exception was the meta methoxy substituent **2-10l**, which showed only moderate activity at around 10µM. However, analogues substituted only at the ortho positions (**2-10m, 2-10n**) or with no substitution (**2-10i, 2-10o**) gives much weaker binding activity. The phenyl ring itself is important to activity as replacing the phenyl ring with a carboxylic acid (**2-10s**) eliminated its activity, and if replaced with a methyl ester (**2-10t**), increased IC_{50} to 26.9µM (Amongst all the series 1 analogues, the

best results come from analogue **2-10g** with a 3-chloro substituent, at IC_{50} = 307nM, which in fact outperforms the representative compound EpMF2 at 432nM.

SERIES 2 ANALOGUES



Analogue	R	IC_{50} (μM)	Analogue	R	IC_{50} (μM)
EpMF2 (2-10)		0.432	**2-10u**	-H	1.60
2-10v		1.23	**2-10w**		1.40
2-10x		0.627	**2-10y**		0.220
2-10z		1.10	**2-10aa**		no inhibition
2-10ab		32.5	**2-10ac**		no inhibition

					
2-10t		26.9	**2-10ad**		0.481
2-10ae		1.90			
					
2-10g		0.347	**2-10af**		9.10
2-10ag		1.70			

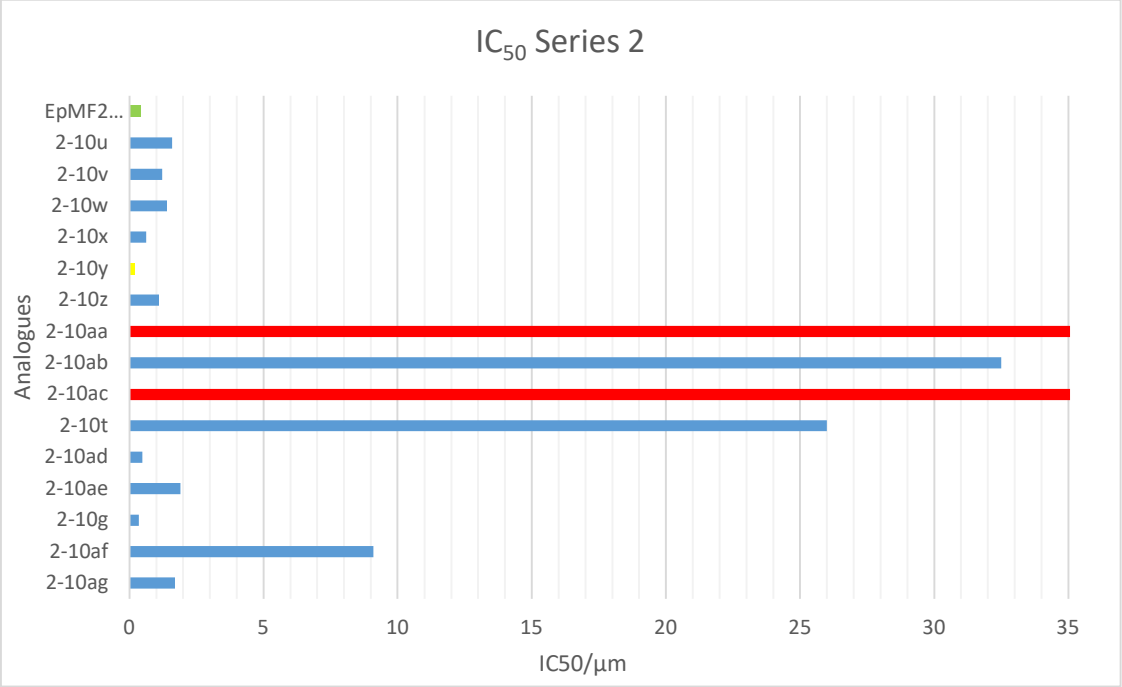
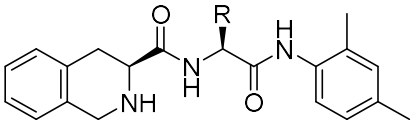


Table 2.14: IC_{50} Activity of series 2 analogues on M. smegmatis ATP synthase. On the graph, green represents EpMF2. Red bars represent no measurable inhibition up to 2mM. Yellow indicates the best performing analogue in its series.

Compared to Series 1 analogues, Series 2 analogues were not so tolerant to various functional groups. It appears that a bicyclic ring moiety, both aromatic, is necessary (**2-10v-2-10z**). The presence of a basic nitrogen atom seemed important in its activity, as substituents with a non-basic N atom (**2-10v, 2-10w**) decreased binding to above 1µM. Replacing it with a single ring significantly reduced (**2-10aa**) and even abolished (**2-10ab**) its activity. Adding substituents to the bicyclic ring, to make it more resemble the original structure of BDQ, actually decreased activity to above 1µM (**2-10z**). Replacement of a single ring with a saturated ring also led to loss of activity (**2-10ac**). An exception comes from **2-10u**, where the moiety was completely removed to leave a terminal amine, but activity was weaker than the bicyclic aromatic ring moieties.

Another set of Series 2 analogues, based on **2-10t**, was synthesized. While **2-10t** by itself was not a potent inhibitor, intriguingly, attaching a meta substituted phenyl moiety (**2-10ae, 2-10af**) significantly increased its activity. (In the case of **2-10ae**, almost equal to that of the original EpMF2 molecule.

Because of the outstanding performance of 3-chloro analogue **2-10g**, another set of Series 2 analogues was synthesized based on **2-10g**. Just as before, attaching a 3-fluorobenzoyl moiety at the N atom of the tetrahydroisoquinoline enhanced its activity (**2-10ag**). The enhancement was not present if the substituent is an acetyl group (**2-10af**), suggesting the phenyl ring is also somehow involved in binding.

SERIES 3 ANALOGUES 

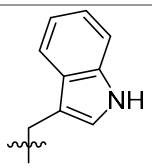
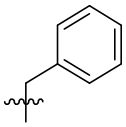
Analogue	R	IC_{50} (μM)	Analogue	R	IC_{50} (μM)
EpMF2 (2-10)		0.432	**2-10ah**		no inhibition
2-10ai		9.50			

Table 2.15: IC_{50} values of Series 3 analogues.

In series 3, it was established that the tryptophan moiety was important to its activity, replacement of the 3-indoyl moiety with a phenyl group (effectively substituting tryptophan with phenylalanine) increase its IC_{50} more than twentyfold. (**2-10ah**) Removing any substituent (effectively replacing the tryptophan with glycine) abolished its activity. (**2-10ai**)

Analogue	Structure	IC_{50} (μM)

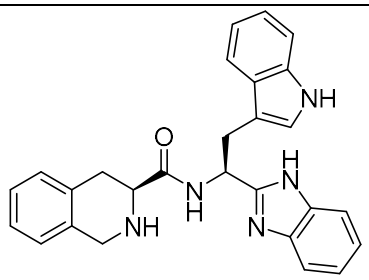
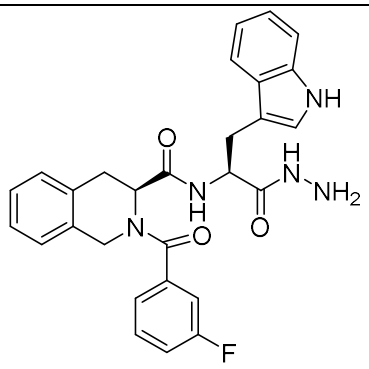
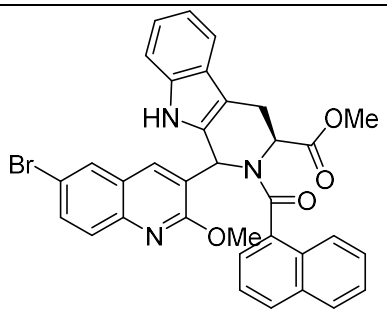
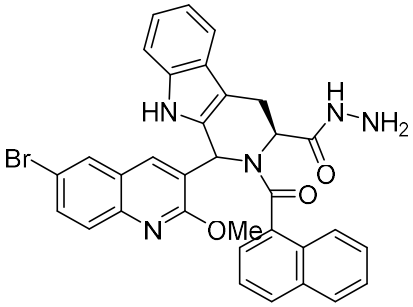
2-10aj		58.5
2-10ak		0.659
2-10al		0.635
2-10am		1.23

Table 2.16: IC_{50} values of miscellaneous analogues.

2-10aj was synthesized as an attempt to investigate what happens when the amide in Series 1 analogues was replaced with a benzimidazole. The functional group functions as an isostere that is more rigid than the corresponding amide due to its fixed cyclic confirmation. The activity

was significantly affected negatively to an IC_{50} value of 58.5μM. Perhaps the confirmational flexibility of the right-side group is important in its activity.

2-10ai was synthesized to investigate the replacement the methyl ester of **2-10ae** with a hydrazine. Interestingly, the activities of the two are comparable (0.481 and 0.659 μM). It appears as far as these two analogues were concerned, the rightmost moiety is not very important to its activity.

2-10al and **2-10am** were synthesized as analogues that closely mimic BDQ in chemical structure. The naphthalene moiety and 2-methoxy-6-bromo quinoline moiety are also found in BDQ. However, there was no improvement, however, in activity over BDQ (and EpMF2 for that matter).

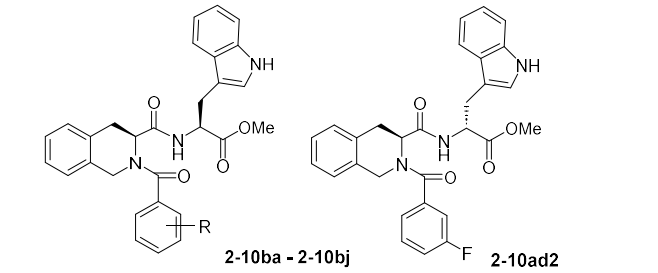
Regardless of its inhibitory activity against ATP synthase, all EpMF2 analogues have a serious problem; they lack activity against live mycobacteria. Only one analogue has any measurable MIC activity against live mycobacteria, **2-10o**, and even then, it is only a 30% inhibition at the maximum concentration of 2mM.

2.3 Results and Discussion

2.3.1 Syntheses of Series 4 analogues based on **2-10ad**.

Setting aside the problems of its lack of MIC_{50} activity at the moment, this section will describe the synthesis, by the author, of more analogues expanding on work done previously above.

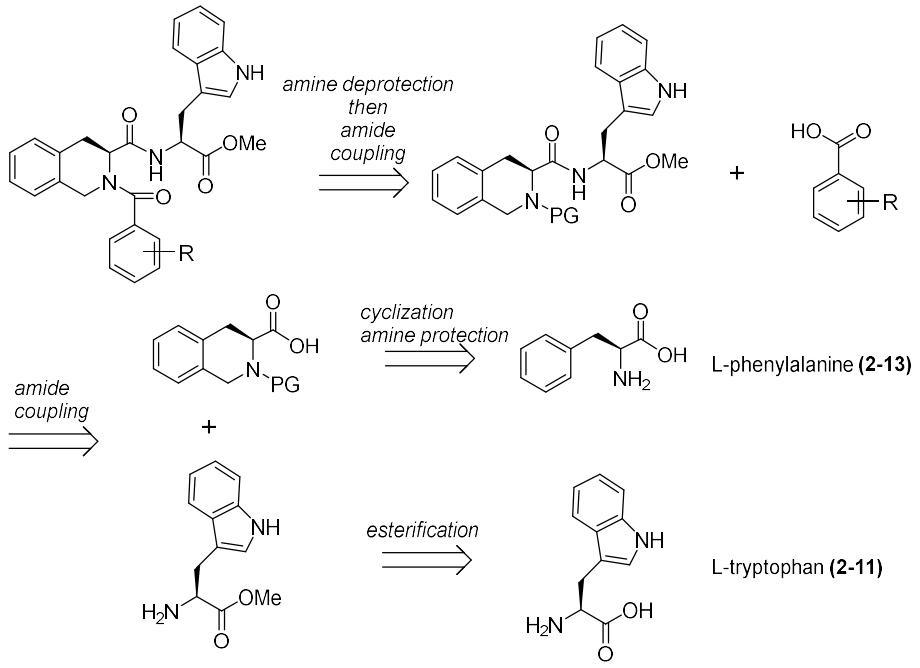
We were curious by the restoration of inhibitory activity when a benzoyl group was attached to the N atom of the tetrahydroisoquinoline component in analogues **2-10ad** and **2-10ae**. To extend from these analogues, other compounds with different substitution patterns were on the bottom most phenyl ring were synthesized to create a new series of analogues, Series 4. One analogue, **2-10ad2**, was also synthesized with inverted stereochemistry at the tryptophan moiety (R instead of S) to investigate the influence of stereoisomerism on the structural activity relationship of EpMF2.



2-10ba R = 4-F **2-10bb R** = 2-F **2-10bc** R = 3-Cl **2-10b4** R = 3-Br
2-10be R = 4-Br **2-10bf** R = 4-OH **2-10bg** R = 2,3,4-OH
2-10bg2 R = 2,3,4-OBn
2-10bh R = 4-Et **2-10bi** no substitution
2-10ad R = 3-F **2-10ae** R = 3-OH **2-10ae2** R = 3-OAc

2-10ad2 (D-tryptophan, 3-F benzoyl)

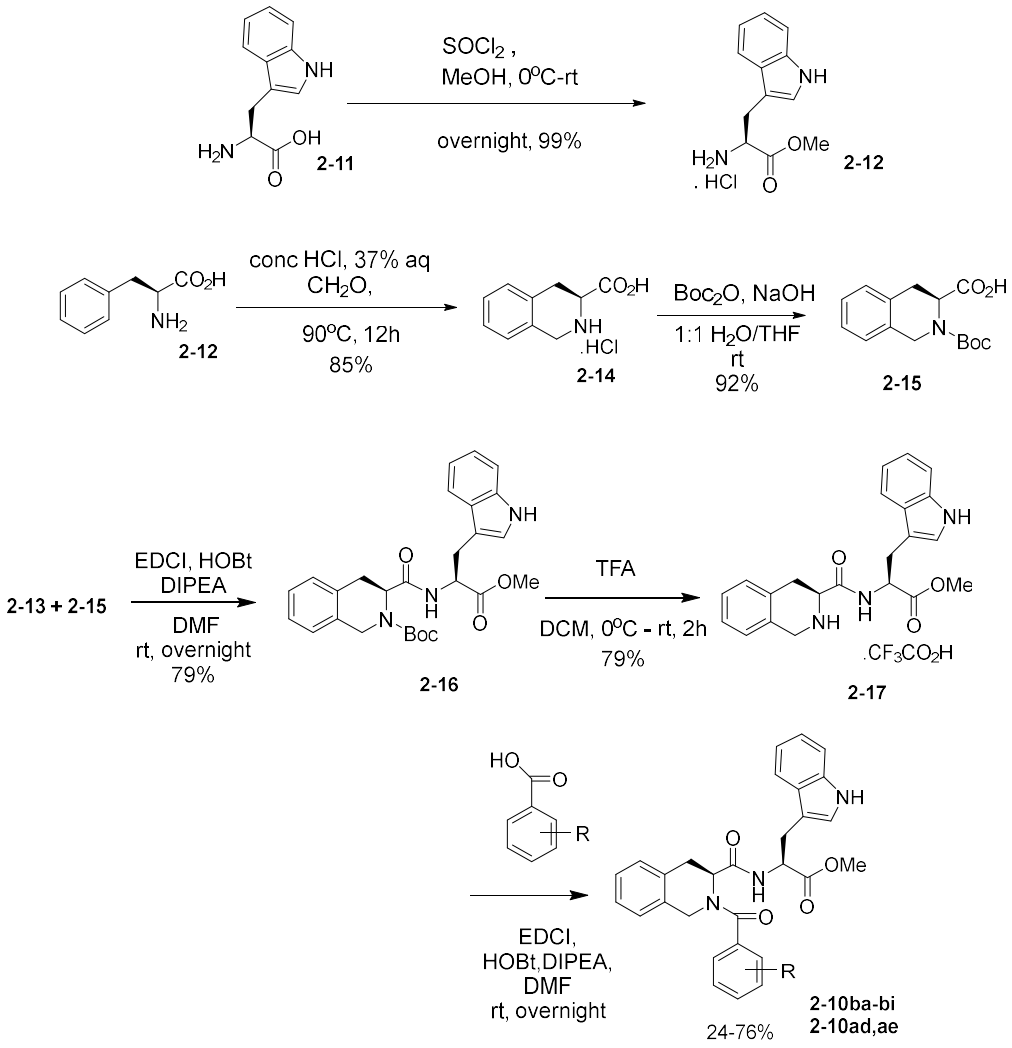
Figure 2.17: Series 4 analogues that were synthesized. **2-10ad** and **2-10ae** were previously synthesized and was resynthesized in this work for comparison purposes



Scheme 2.18: Retrosynthesis of Series 4 analogues

The retrosynthesis of Series 4 analogues was simple, and the key lies in the amide moieties. The common proteinogenic amino acids L-Phenylalanine and L-Tryptophan were used as starting materials. Tryptophan would be esterified to its methyl ester, while phenylalanine would be cyclized to a tetrahydroisoquinolinyl carboxylic acid through the Pictet-Spengler reaction with formaldehyde, then the amine protected. Coupling of the amine of the tryptophan moiety with the carboxylic acid of the tetrahydroisoquinoline derivative forms the first amine

bond. After deprotection, the exposed amine will undergo a second coupling with the benzoic acid derivative to give rise to the final analogue. For analogue **2-10ad2** D-Tryptophan was used as starting material instead.

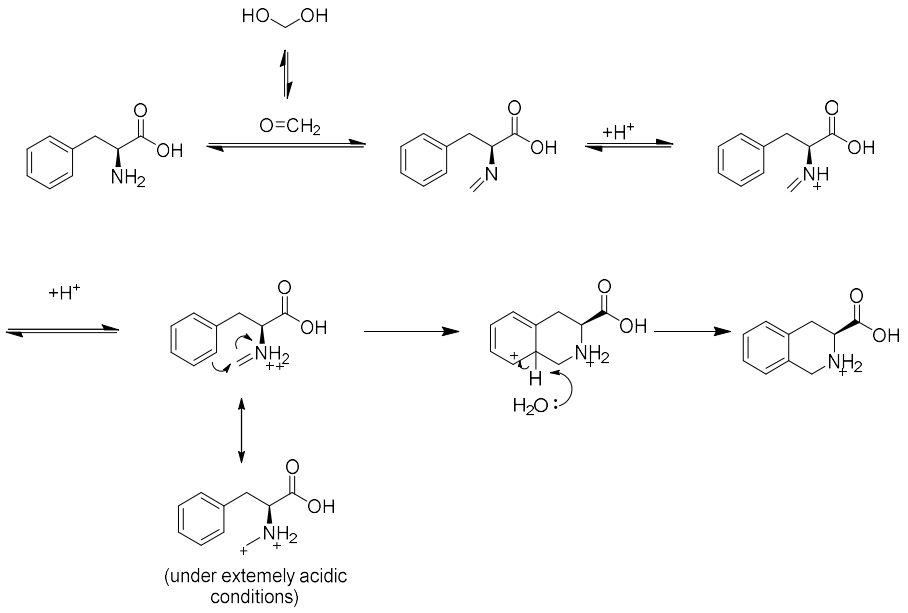


Scheme 2.19: Synthetic route to Series 4 analogues

The syntheses of series 4 analogues began when L-Tryptophan **2-11** was first converted to its methyl ester hydrochloride **2-12** in 99% yield with methanol and thionyl chloride. Methanol is the solvent and also as the alcohol component of the esterification reaction. The Cl atoms in thionyl chloride are easily displaced by MeOH and when used in excess in the case of this reaction, forms the sulfite ester. This intermediate is a potent alkylating agent and transfers the methyl group to the carboxylic acid of tryptophan to form the methyl ester.[25] The side products

that are formed are SO_2 and HCl, which are gaseous and evaporates easily, simplifying workup to just evaporation of solvent. The amine moiety of tryptophan reacts with one equivalent of HCl to form the hydrochloride salt.

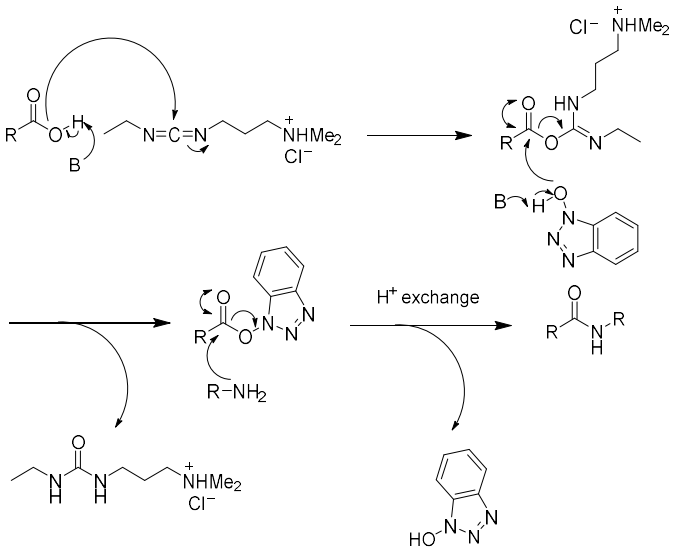
The tetrahydroisoquinoline component was derived from L-phenylalanine **2-13**. This amino acid underwent the Pictet-Spengler reaction where the amine was condensed with formaldehyde under acidic conditions followed by cyclization to form tetrahydroisoquinoline 3-carboxylic acid hydrochloride **2-14** in 85% yield. Mechanistic studies on β-indoyl ethylamines revealed that the formaldehyde first reacts with the amine moiety to form the imine, which is protonated to the iminium ion intermediate under acid conditions. The protonated iminium ion is highly electrophilic, and the neighbouring indole ring undergoes an addition reaction to form a spirocycle. Alkyl migration then deprotonation follows to furnish the final product.[26] In this variant of the reaction, the phenyl ring of the amino acid is less electron rich than indole. In order to induce the phenyl ring to undergo cyclization, forcing conditions were required to create the extremely electrophilic *N,N*-diprotonated imine,[27] thus necessitating the use of reflux in concentrated HCl for this reaction.



Scheme 2.20: Mechanism of the Pictet-Spengler reaction for β-phenylamines. This reaction requires extremely acidic conditions due to the relative electron poor nature of the unsubstituted phenyl ring.

The amine of tetrahydroisoquinoline derivative **2-14** was protected with di-tert-butyl dicarbonate (Boc_2O) to form the Boc-amine **2-15** under basic conditions in aqueous THF under 92% yield.

2-12 and **2-15** were coupled together using the reagent 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride EDCI along with HOBt hydrate to from **2-16** at a 79% yield. The carbon atom of the carbodiimide moiety of EDCI is fairly electrophilic and reacts with the carboxylic acid of **2-15** to form an activated ester intermediate. In the presence of HOBt a substitution reaction occurs on the intermediate to form another benzotriazole ester intermediate. Both intermediates can react with amines to form the amide product. HOBt, while strictly not necessary for the reaction, helps in the reaction by preventing the formation of side products and also minimizes the racemization of the amino acid.[28] Some side product was observed, which was determined to be *N*-Boc tryptophan methyl ester. The reason for this was unknown.

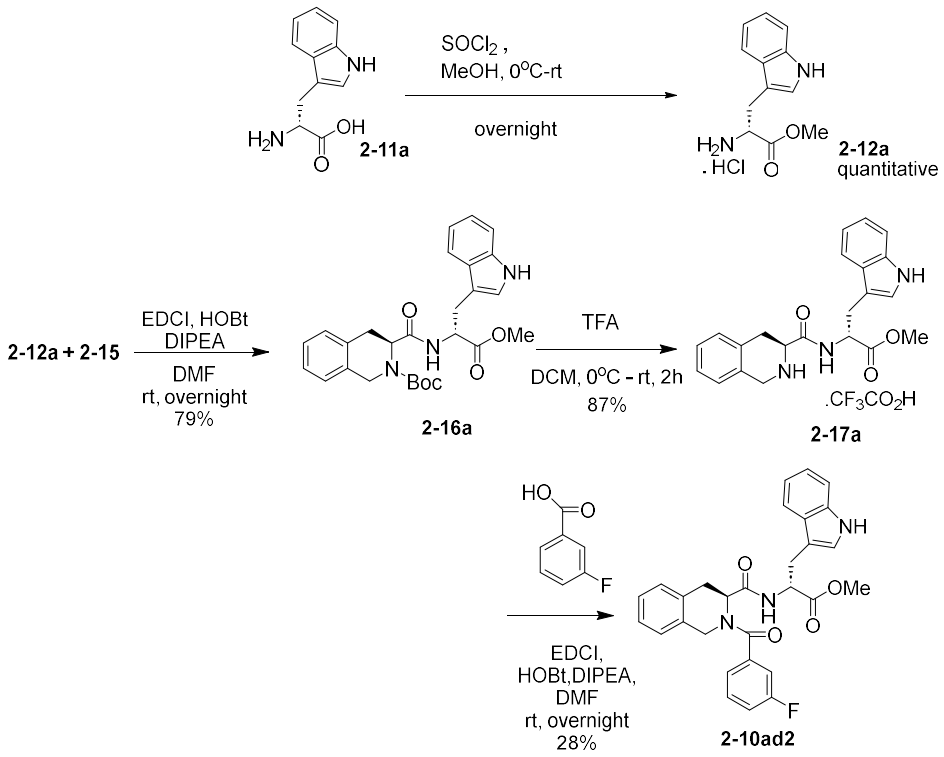


Scheme 2.21: Peptide coupling mechanism using EDCI and HOBt.

The N-Boc coupled product **2-16** was deprotected under acidic conditions using trifluoroacetic acid (TFA) to give the amine **2-17** at 79% yield, which was isolated as its trifluoroacetate salt. The reaction is driven by the release by the relatively stable t-butyl cation. The coupling

reaction was then repeated between **2-17** and the corresponding benzoic acids using EDCI in identical conditions to form the final compounds. The yields varied from 24 to 75%.

The synthesis of D-tryptophan containing analogue **2-10ad2**, a diastereomer of **2-10ad**, was done under identical conditions as the synthesis of **2-10ad**, the only difference being that D-tryptophan **(2-11a)** was used as the precursor. The methyl ester hydrochloride was formed from tryptophan in methanol and $SOCl_2$ in quantitative yield, and this was coupled to Boc-protected tetrahydroisoquinoline 3-carboxylic acid **2-15** with EDCI/HOBt to form **2-16a** a in 73% yield. The Boc group was deprotected by TFA to **2-17a** in 87% yield, and a second coupling reaction was done with 3-fluorobenzoic acid to give the final product **2-10ad2** in a moderate 28% yield.



Scheme 2.22: Synthesis of analogue **2-10ad2**, a diastereomer of **2-10ad** using D-tryptophan.

For the final coupling step for hydroxy analogue **2-10bf**, using 4-hydroxybenzoic acid led to significantly reduced yield at 24%, and purification on column chromatography led to trace amounts of doubly coupled product (8%) where the hydroxyl group of **2-10be1** was coupled to another 4-hydroxybenzoic acid molecule. Mass spectroscopy even suggested the presence

of triply coupled product **2-10be2**. The hydroxyl group competes with the amine on **2-17** on coupling to the carboxylic acid moiety, which explained the poor yields and side products. Subsequent reactions with hydroxyl containing benzoic acids were thus done with protection of these -OH groups.

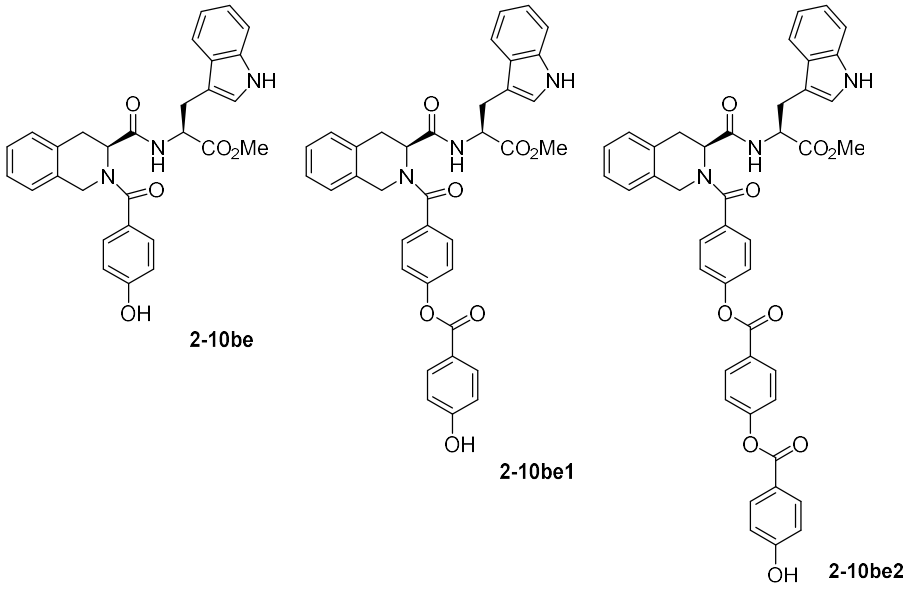
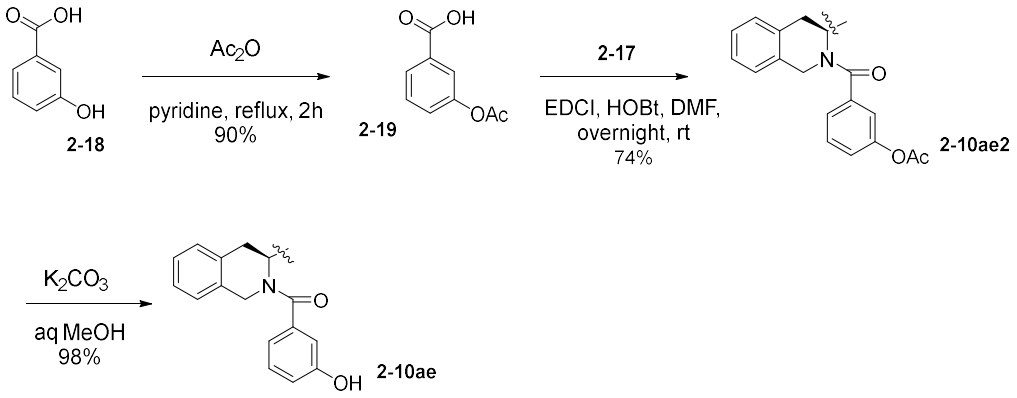


Figure 2.23: Coupling of compound **2-17** with unprotected hydroxyl benzoic acids led to poor yields of the desired product (left)**2-10be1** and small amounts of diester (medium). LR-MS even suggested the presence of triester **2-10be2**(right)

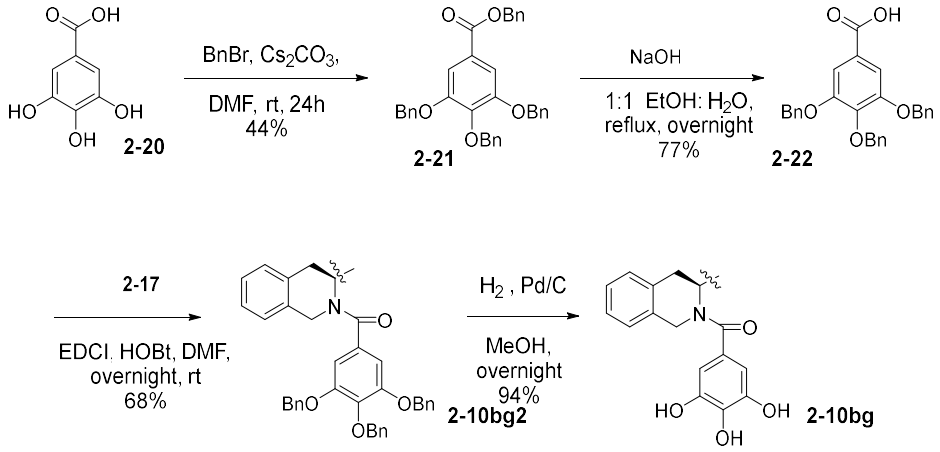
For the subsequent reaction with 3-hydroxybenzoic acid **2-18**, the hydroxyl moiety was first protected by acetylation through refluxing with acetic anhydride in pyridine to furnish 3-acetylbenzoic acid compound **2-19** in 90% yield. After the subsequent coupling step, purification by column chromatography isolated the acetyl compound **2-10ae2** in 74% yield, along with some deacetylated product **2-10ae** at 20%. The remaining **2-10ae2** was then deacetylated under mild conditions using K_2CO_3 in aqueous methanol to complete the synthesis of 3-hydroxy analogue **2-10ae**.



Scheme 2.24: Protection of 3-hydroxy benzoic acid before coupling to **2-17** and its subsequent deprotection.

The use of the 2,3,4 hydroxy analogue **2-10bg**, derived from 2,3,4-hydroxybenzoic acid or gallic acid, was of interest to us due to a paper by Karlscheuer *et al*. describing a series of 3-methoxy gallic acid ester analogues that showed antitubercular activity against *Mtb*.[29] They determined that its antibacterial activity was caused by inhibition of stearoyl dehydrogenase DesA3, which introduces unsaturation into fatty acids. This ultimately results in the inhibition of oleic acid synthesis, which is believed to be a precursor in the synthesis of mycolic acids found in the *Mtb* cell wall. Although this was a different mechanism from ATP synthase inhibition of EpMF2 analogues, we were still interested in seeing whether the gallate moiety might confer any additional inhibitory activity on mycobacteria.

Like other hydroxy containing analogues, gallic acid **2-20** was first protected, this time with the benzyl group, by stirring with excess benzyl bromide and caesium carbonate in DMF to obtain the quadruply benzylated complex **2-21** at 47% yield. The benzyl ester of **2-21** was then cleaved selectively with NaOH in aqueous ethanol to form the triply *O*-benzylated compound **2-22** at 77% yield, or 36% over 2 steps. After coupling to **2-17** to give intermediate **2-10bg2** in 64% yield, the benzyl groups of the benzylated analogue **2-10bg2** were cleaved by hydrogenolysis by H_2 gas with a Pd catalyst to obtain the final product **2-10bg** in 94% yield.



Scheme 2.25: Protection of 3,4,5-hydroxy benzoic acid (gallic acid) before coupling to **2-17** and its subsequent deprotection.

2.3.2 Syntheses of Series 5 analogues *(n*-butyl ester and carboxylic acid derivatives)

From the paper from Karlscheuer *et al*. about gallic acid analogues with anti-tubercular activity,[29] the analogues showed improved potency when long chain alkyl esters of gallic acid are used instead of the carboxylic acid. While 3-methoxy gallic acid itself has low MIC_{90} activity at 100 μM, replacement of the carboxylic acid with a long chain butyl, hexyl, and octyl esters significantly increased its activity to a MIC_{90} value of just 6.25μM. This was notable given esters are often cleaved *in vivo*.

From this results came the idea of replacing the methyl ester of **2-10ad** and **2-10bg** with longer chain alkyl groups. The n-butyl ester was chosen, and it was also decided to synthesize the carboxylic acid equivalents as a comparison. It would then be possible to investigate any relation between the analogue's lipophilicity and its activity, and also if increasing the hydrophobicity of the analogue might give it MIC_{50} activity. Thus the syntheses of Series 5 analogues was set.

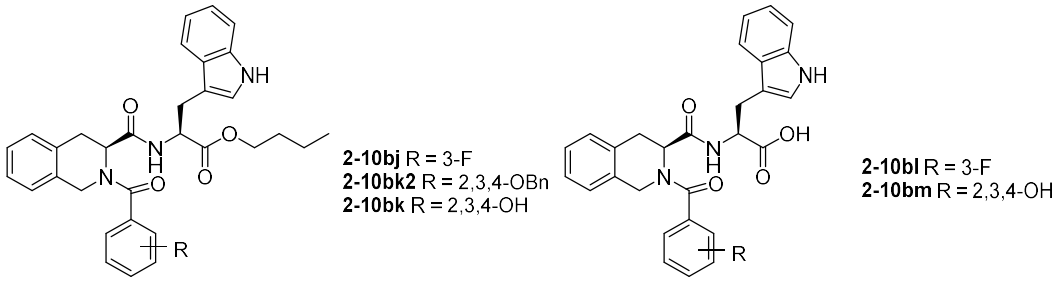
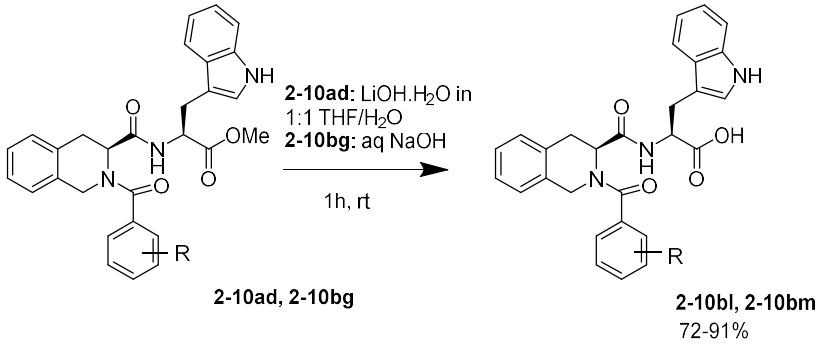


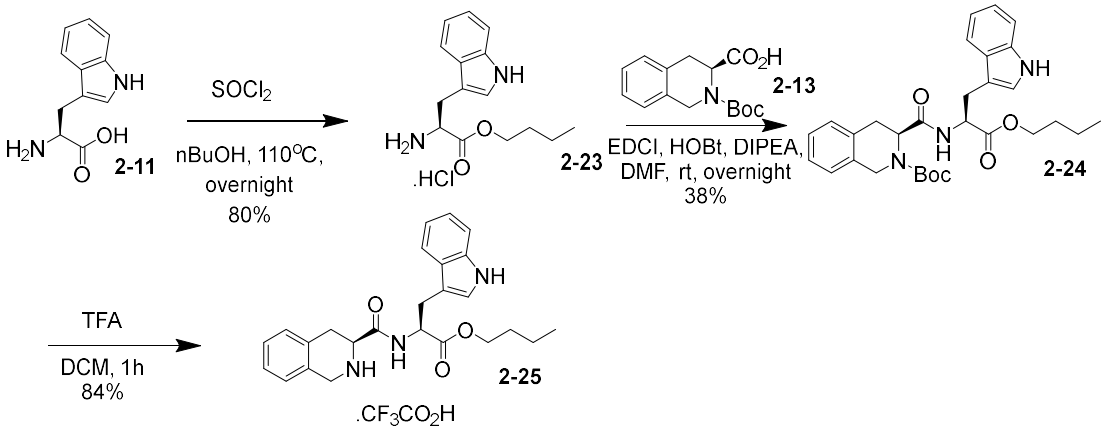
Figure 2.26: Structures of Series 5 analogues

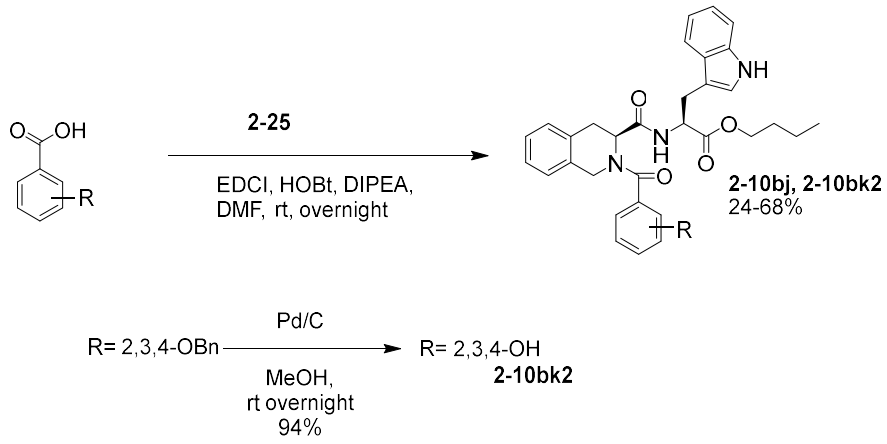
The carboxylic acid analgoues **2-10bl** and **2-10bm** were synthesized in one step from hydrolyzing the corresponfing methyl esters **2-10ad** and **2-10bg** with LiOH monohydrate in aqeuous THF and aq. NaOH solution respectively at 91% and 72% yields respectively.



Scheme 2.27: Synthesis of carboxylic acid analogues in Series 5

The n-butyl esters were synthesized in the same synthetic route that gave rise to the methyl esters in Series 4, except that the esterification step at the beginning had the solvent changed from methanol to n-butanol.





Scheme 2.28: Synthetic route to n-butyl esters in Series 5.

Tryptophan **2-11** was esterified to its butyl ester hydrochloride **2-23** under reflux in n-BuOH and $SOCl_2$ at 80% yield. The amine of **2-23** was coupled to carboxylic acid **3** with EDCI and HOBt to obtain compound **2-24** in a modest 38% yield, after which the Boc group was deprotected with trifluoroacetic acid to its trifluoroacetyl salt of amine **2-25** in 84% yield. The final coupling step to the benzoic acid derivative led to 3-fluorobenzoyl analogue **2-10bj** in 24% yield and the 2,3,4-O-benzyl analogue **2-10bk2** in 68% yield respectively. **2-10bk2** was then debenzylated by hydrogenolysis over Pd to form the 2,3,4-hydroxy analogue **2-10bk** in 94% yield.

2.3.3 Synthesis of miscellaneous analogues **2-10bn:** Best of both worlds

Syntheses of analogues previously from Series 1 and 2 analogues revealed that compounds **2-10g** and **2-10y** were the most active in ATP synthase inhibition on M. smegmatis.

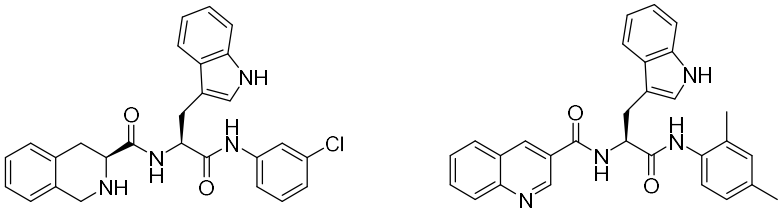
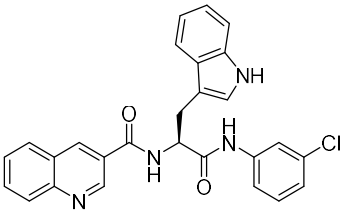


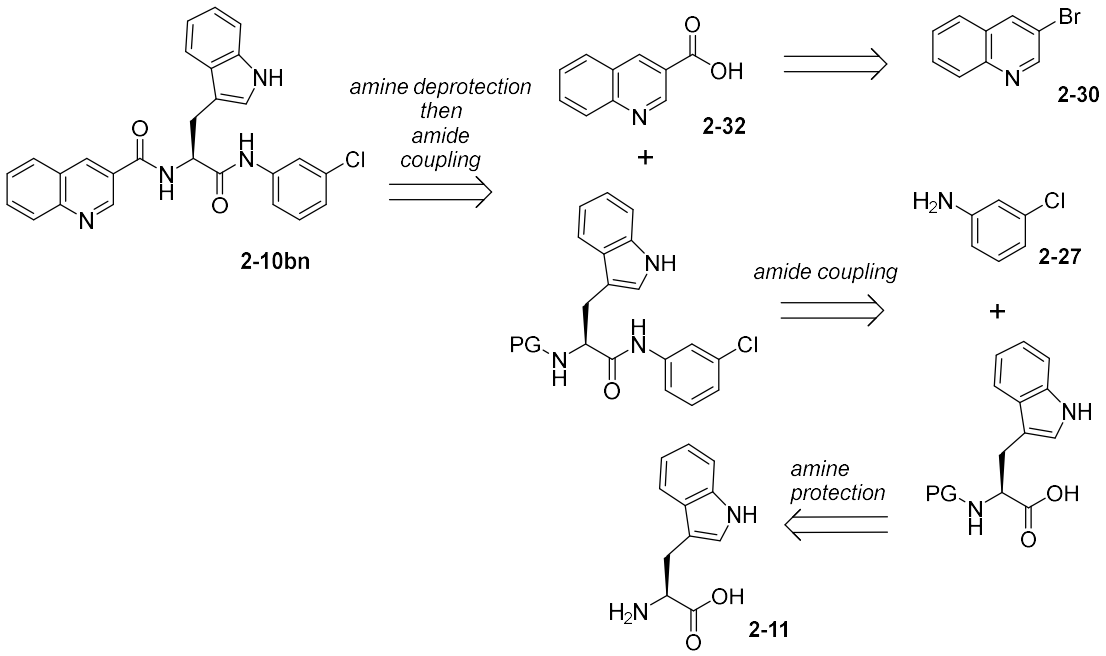
Figure 2.29: Compound **2-10g** (left) is the most active among all series 1 analogues of EpMF2 with IC_{50}=307nM, and compound EpMF2 **2-10y** the most active of series 2 analogues with IC_{50}=220nM

The idea thus came to combine the most active left-side moiety (quinolinyl) with the most active right-side moiety (3-chlorophenyl) to form the miscellaneous compound **2-10bn**.



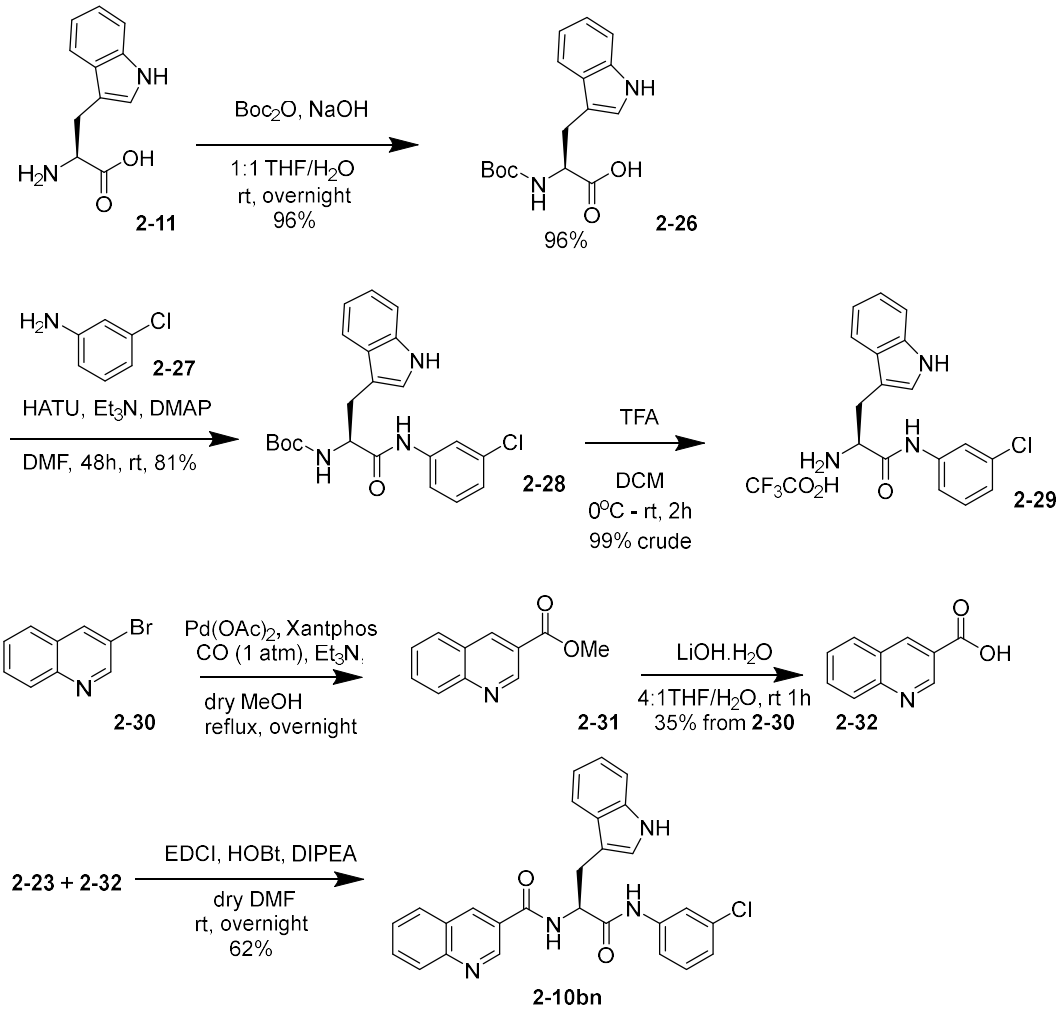
2-10bn

Figure 2.30: Chemical structure of potential lead molecule **2-10bn**



Scheme 2.31: Retrosynthesis of analogue **2-10bn**

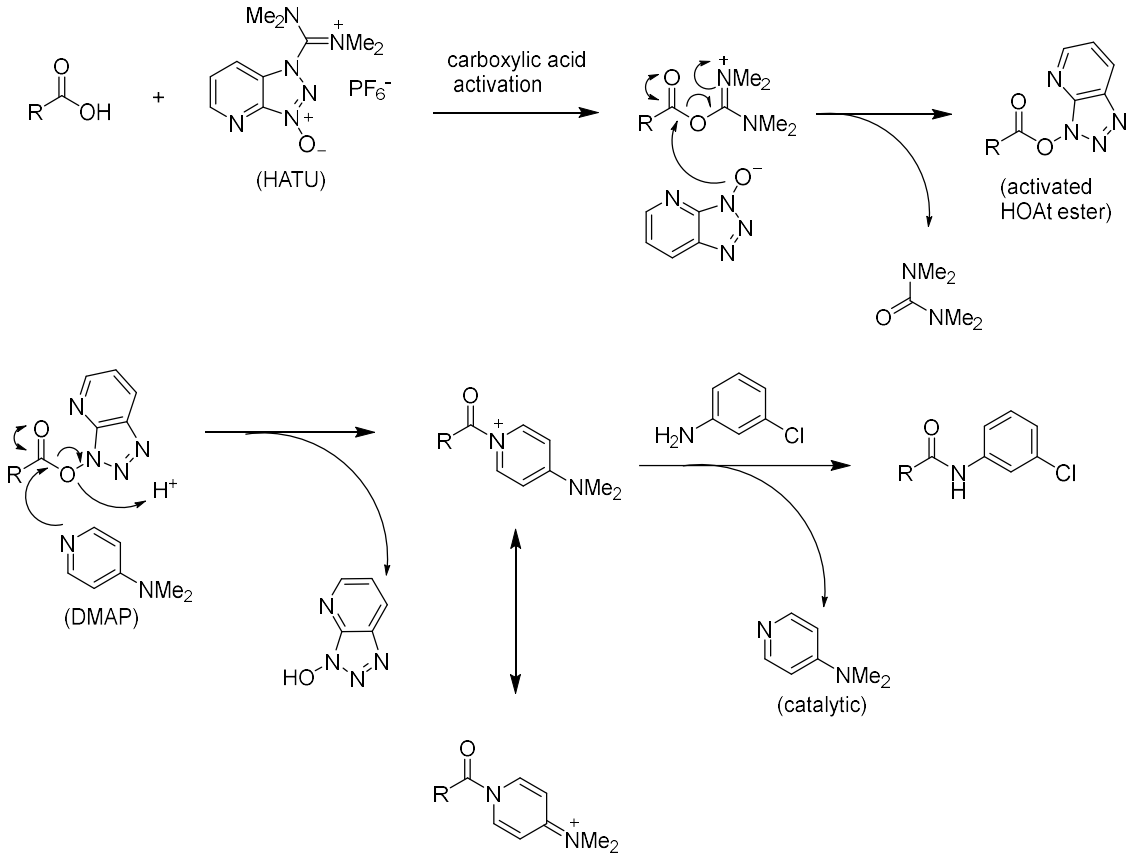
The retrosynthesis relies on the sequential coupling of 2 amide bonds. The amine moiety of tryptophan **2-11** would be protected, followed by the first coupling of the carboxylic acid with the amine of 3-chloroaniline **2-27**. The amine was then deprotected, and the exposed amine would then be coupled to quinoline 3-carboxylic acid **2-32** to from the final product. **2-32** was not immediately available in our group, but it was easily derivatized in two steps from the precursor 3-bromoquinoline **2-30**.



Scheme 2.32: Synthetic route to compound **2-10bn**

The synthesis began by protecting the amine moiety of tryptophan **2-11** with the Boc group by Boc_2O to give the Boc derivative **2-26** in 96% yield. Attempts to couple **2-26** with 3-chloroaniline **2-27** using EDCI led to very slow, incomplete reaction after 24 hours. This arises from the weak nucleophilicity of the amine. The lone pair of the amine is delocalised into the aromatic ring. The presence of electron withdrawing meta-chloro substituent further exacerbated the problem. This was resolved by replacing EDCI with HATU and the addition of 10% DMAP as a catalyst. HATU is a compound consisting of an azabenzotriazole moiety attached to an iminium moiety. It reacts with the carboxylic acid to form an azabenzotriazole ester similar to that of HOBt activated esters. The key to the reaction, however, lies with the

displacement of the HOAt ester by DMAP to form a highly activated acetylpyrimidium moiety. Finally, the amine underwent nucleophilic acyl substitution with the aryl amine to form the final product. The use of DMAP, originally used for coupling of alcohols to carboxylic acids in the Steglich esterification reaction,[30] was also applicable to this reaction.

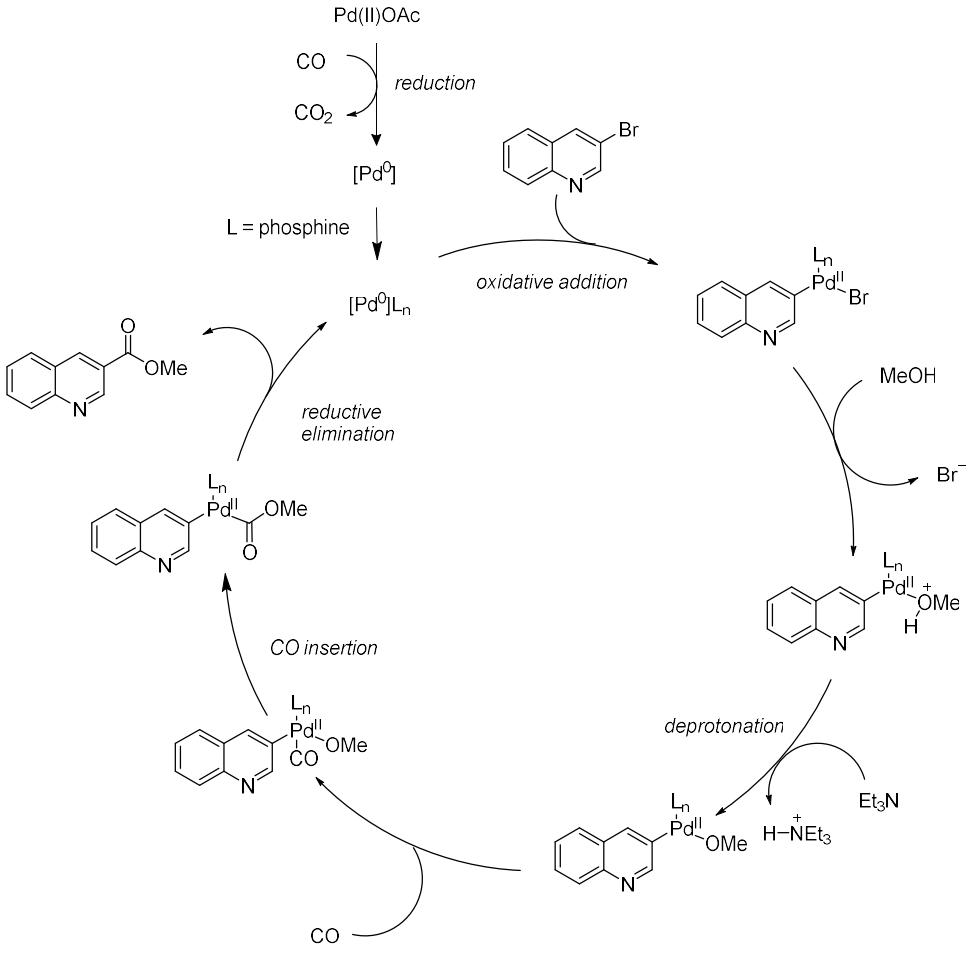


Scheme 2.33: The mechanism of the DMAP catalyzed amide formation from the poorly nucleophilic aniline derivative.

With the change in conditions, the reaction was pushed to completion forming compound **2-28** in 81% yield. **2-28** was then deprotected with DCM in TFA to form the trifluoroacetate salt **2-29** with 99% crude product obtained.

Isoquinoline-3-carboxylic acid **2-32** was synthesized from 3-bromoisoquinoline **2-30** using a variant of a palladium catalyzed carbonylation reaction developed by Heck[31]. Xantphos was used as the phosphorous ligand. Its bidentate nature and wide bide angle allow trans-chelation of square planar Pd complexes, which in this case aided in the reaction. In the reaction, the

bromine atom was displaced by the methyl ester, where the methoxy group originated from methanol, and carbonyl group added by subsequent carbonylation (scheme 2.34). The ester intermediate **2-31** was then hydrolysed under basic conditions to of the corresponding carboxylic acid **2-32** at 35% yield. **2-29** and **2-32** were then coupled together using EDCI/ HOBt to form the final compound **2-10bn** in a 63% yield.



Scheme 2.34: Possible catalytic cycle mechanism of the carbomethylation of the bromoquinoline to form the methyl ester.

2.3.4 Synthesis of miscellaneous analogue **2-10bo**: a diketopiperazine derivative

Analogue **2-10bo** was serendipitously synthesized during an unsuccessful attempt to synthesize *N*-methylated amide derivatives. By removing a source of hydrogen bond donor moiety, it was desired to observe its effects on its activity. In addition, *N*-methyl amino acid containing

peptides were known to be biologically active compounds, including the immunosuppressant ciclosporin, the antibiotic echinomycin and insecticide destruxin A.[32]

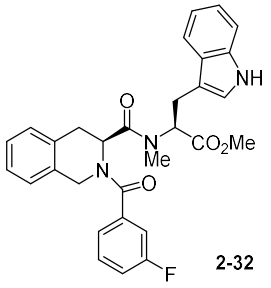
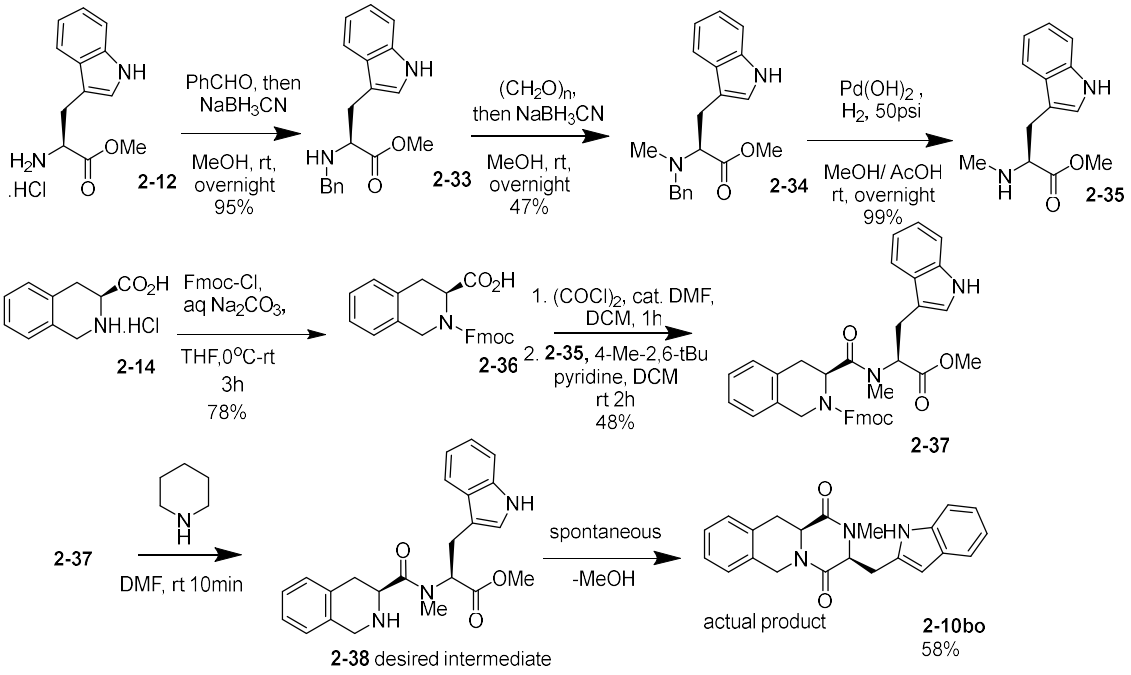


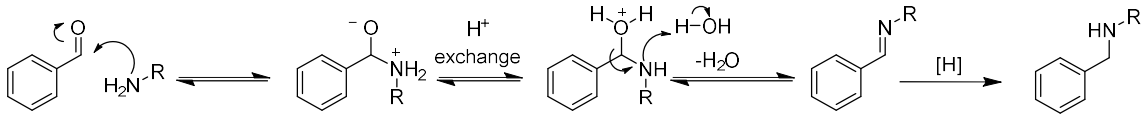
Figure 2.35: An attempt to synthesize this N-methylated amide analogue **2-32** was not succesful.



Scheme 2.36: Synthetic route to the diketopiperazine compound **2-10bo**

The synthesis required the formation of *N*-methyl tryptophan **(2-35)** as an intermediate. A common method to methylate amines is the Eschweiler-Clarke reaction,[33] however under these conditions, the primary amine of tryptophan would be converted to the *N,N*-dimethylamine and it was not possible to selectively halt the reaction at the mono *N*-methyl intermediate. An attempt to submit tryptophan methyl ester **2-12** under Eschweiler-Clarke conditions with formic acid and formaldehyde led to a complex mixture of mono and dimethylated amines, in addition to *N*-formyl products.

Therefore, an indirect route based on Konopelski *et al*.[34] was adopted where the amine in starting compound **2-12** first underwent *N*-benzylation with benzaldehyde through a reductive amination reaction to form secondary amine **2-33** in 95% crude yield. The condensation of the amine with benzaldehyde led to the formation of an imine intermediate *in situ*, which is reduced to the amine with the mild reducing agent sodium cyanoborohydride. However, the condensation is an equilibrium reaction between benzaldehyde and its imine and prolonged reaction did not lead to a complete conversion of benzaldehyde. This was remedied by the addition of 5% anhydrous $MgSO_4$ by mass, which absorbed the water and shifted the equilibrium towards complete conversion to the imine (and subsequent reduction to the benzylamine). The crude product **2-33** was subjected to a second reductive amination reaction with paraformaldehyde to attach the second methyl group to the amine using identical conditions to form the tertiary amine **2-34** at 47% yield. The benzyl group was selectively removed under hydrogenolysis with palladium hydroxide under a medium pressure H_2 atmosphere to form the monomethyl amine **2-35** in 99% yield.

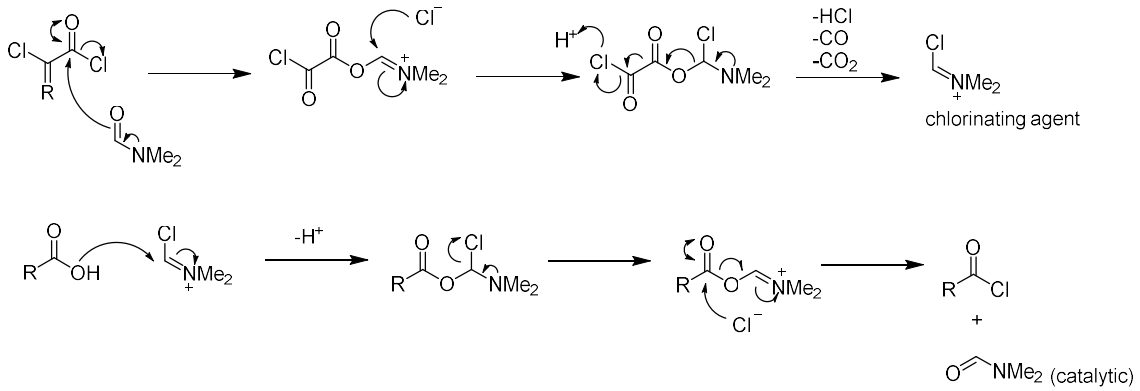


Scheme 2.37: Mechanism of reductive amination

However, many attempts to couple the *N*-methylamino acid **2-35** with tetrahydroisoquinoline 3-carboxylic acid **2-14** proved unsuccessful. The presence of an alkyl group on the amine increases its steric hindrance and this presented extreme difficulties during coupling with carboxylic acids. The use of EDCI, HATU and even more exotic peptide coupling agents like PyBOP did not lead to the formation of any product that could be identified as the coupled peptide. Literature on the peptide couplings of *N*-alkyl amino acids were few, and often require nonstandard conditions. To date there has been no procedure found that can be universally applied to the coupling of *N*-alkyl amino acids with carboxylic acids. One of the few success

came from a paper from Jung where he was able to successfully couple resin bound N-methyl amino acids using triphosgene to generate acyl chlorides in situ from corresponding carboxylic acids.[35] He demonstrated the procedure to be superior to other reagents used, including DIC/HOAt and tetramethylfluoroformamidinium hexafluorophosphate (TFFA), a acyl fluoride forming reagent.

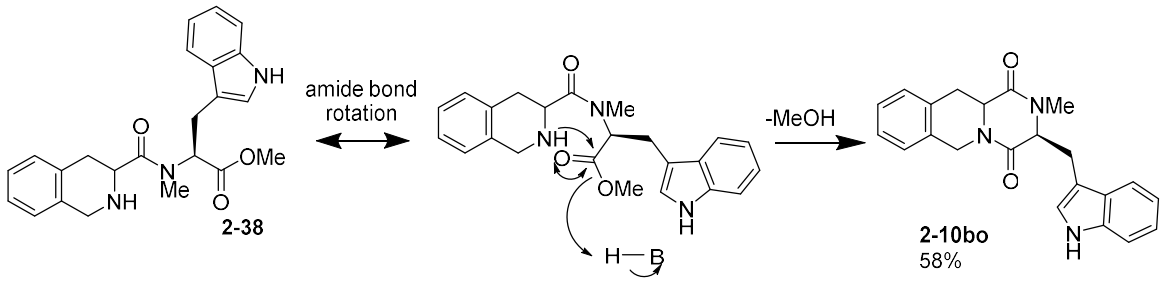
Based on his reports, an attempt was made to couple the *N*-methyl amine by formation of an acyl chloride from the tetrahydroisoquinoline carboxylic acid **2-14**. The amine of **2-14** was protected with the Fmoc group, as it is stable to the acidic conditions used for acyl chloride formation while the corresponding Boc group would decompose. The Fmoc protected compound **2-36** was prepared by adding FmocCl to a THF/water solution of **2-14** at 78% yield. The carboxylic acid of **2-36** was then converted into the acyl chloride *in situ* with oxalyl chloride, catalyzed by DMF. The DMF is converted by oxalyl chloride to an imidoyl chloride, which is the active chlorinating agent. Once the formation was complete, the intermediate acyl chloride was immediately added to a solution of N-methyl amine **2-35** in dry DCM. Pleasingly, the coupling was successful and the coupled product **2-37** was formed in 48% yield after purification by column chromatography.



Scheme 2.38: Mechanism of the formation of the imidoyl chloride as the active chlorinating from the addition of catalytic DMF to oxalyl chloride and the subsequent formation of acyl chlorides from carboxylic acids

The deprotection of the Fmoc group in **2-37** led to unexpected results. Instead of the expected product **2-38**, a compound was isolated that did not provide the expected NMR data. The proton

NMR singlet peak at δ = 3.7ppm, corresponding to the hydrogens on the methyl ester, was missing, additionally, ESI-mass spectroscopy revealed that its molar mass was 359, 31 lower than the expected structure which gives 370, suggesting the loss of a methanol equivalent. It was concluded that the expected compound was quantitatively converted to its diketopiperazine (DKP) side product **2-10bo** after deprotection according to the proposed mechanism below. DKP formation was a long known side product arising from deprotection of Fmoc protected peptides, and *N*-methyl amino acids were particularly susceptible to this side reaction.[36]



Scheme 2.39: The unexpected cyclization side reaction that arose from Fmoc deprotection to **2-38**

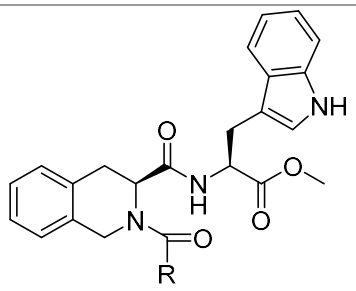
Despite this setback, it was decided to submit this compound anyway for assay testing, as DKPs are widely known to be bioactive natural products.[37] Due to the insurmountable difficulties of coupling *N*-methylated peptides, the synthesis of further such analogues were abandoned.

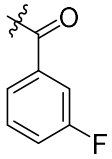
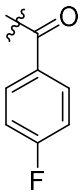
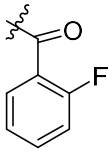
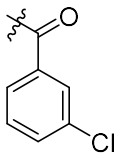
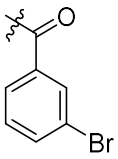
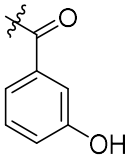
2.3.5 Biological assays of Series 4 and 5 analogues

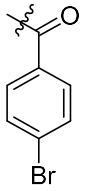
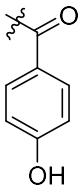
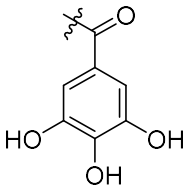
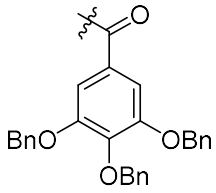
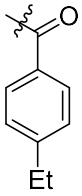
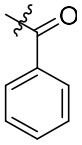
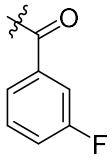
The compounds synthesized in series 4 and 5 were subjected to ATP synthesis assays against *M. smegmatis* inverted membrane vesicles IMV. These were prepared by lysing the bacteria's cell membranes and subjecting them to high-speed centrifugation, which inverts the bilayer membrane and creates inverted membranes after resealing. The ε subunit of the ATP synthase, which was previously in the cytosol, is now on the outside layer and thus exposed for subsequent inhibitory assay tests. For practical reasons, *M. smegmatis* was chosen over *Mtb* as M. *smegmatis*, is fast growing and non-pathogenic, allowing experimental work to be done in a biosafety level one laboratory.

The assay was done by adding different concentrations of the analogue molecule to a mixture of IMV, ADP, phosphate and NADH. The amount of ATP synthesized was measured by adding CellTiter-glo reagent, which contains a mixture of luciferin and the enzyme luciferase. In the presence of ATP and oxygen, a bioluminescent reaction occurs where the luciferase enzyme oxidises luciferin to an excited state, which emits light. The intensity of this luminescence is proportional to ATP concentration, which is correlated with the activity of the ATP synthase. If the intensity of the light decreases upon addition of the EpMF2 analogue, it indicates that ATP synthase was being inhibited. The weaker the luminescence for the same concentration of analogue added, the greater its inhibitory activity against ATP synthase.

SERIES 4 ANALOGUES



R	**IC_{50} / µM**	**R**	**IC_{50} / µM**	**R**	**IC_{50} / µM**
 **2-10ad (previously synthesized)**	0.432	 **2-10ba**	No inhibition	 **2-10bb**	No inhibition
 **2-10bc**	No inhibition	 **2-10bd**	No inhibition		1.90

				2-10ae previously synthesized)	
 **2-10be**	No inhibition	 **2-10bf**	3.60	 **2-10bg**	10.0
 **2-10bga**	No inhibition	 **2-10bh**	No inhibition	 **2-10bi**	No inhibition
 **2-10ad2 (D-Trp)**	No inhibition				

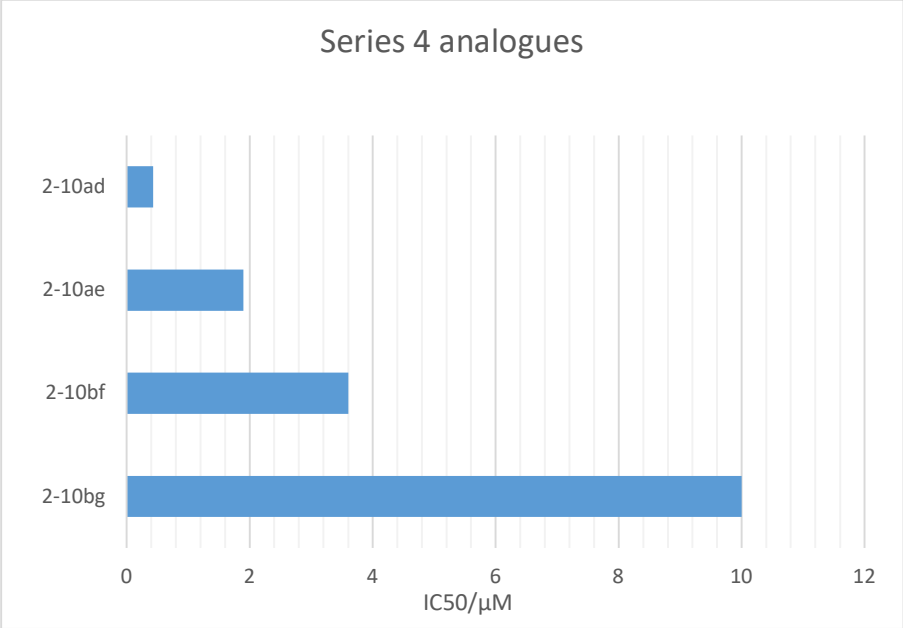
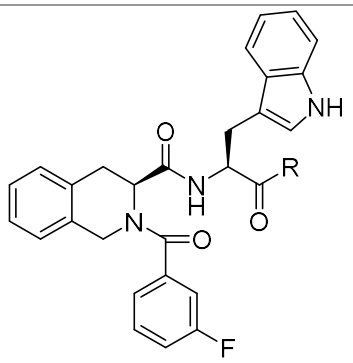
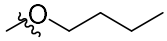
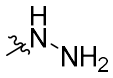


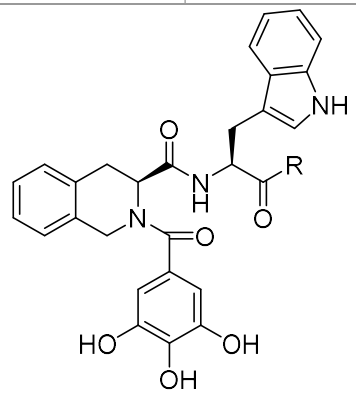
Table 2.40: IC_{50} values of series 4 analogues, compared to the two base compounds **2-10ad** and **2-10ae** as comparison. Analogues with measurable inhibition above 2mM were not included.

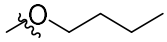
The only analogues with any measurable inhibitory activity are substituents with fluoro or hydroxyl groups(s). It is notable that these moieties were hydrogen bond acceptors or donors, other analogues with no hydrogen bond capability (**2-10bc-5,2-10bga, 2-10bh**) have no measurable activity whatsoever. For the fluoro substituted analogues, only the meta substituent **2-10ad** is active, and ortho and para isomers showed no activity (**2-10ba-2**). For the hydroxyl πsubstituted analogues, activity was retained, but both 4-hydroxy and the gallate analogues (**2-10bf-7**) showed weaker activity than the 3-hydroxy analogue. It can only be concluded that some hydrogen bond interaction of the bottommost phenyl ring is essential of its activity. As for the D-Tryptophan derivative **2-10ad2**, no measurable inhibition was also detected. Therefore, the stereochemistry at the tryptophan is also crucial, and replacement by its diastereomer altered the shape of the molecule enough to prevent formation of crucial non-polar π-π interactions between the indoyl ring and the target.



SERIES 5 ANALOGUES

R	IC_{50}(µM)	R	IC_{50}(µM)	R	IC_{50}(µM)
 **2-10ad (previously synthesized)**	0.485	 **2-10bj**	No inhibition	 **2-10bl**	25.0
 **2-10ak (previously synthesized)**	0.659				



R	IC_{50}(µM)	R	IC_{50}(µM)	R	IC_{50}(µM)
 **2-10bg**	10.0	 **2-10bk**	8.8	 **2-10bm**	50.0

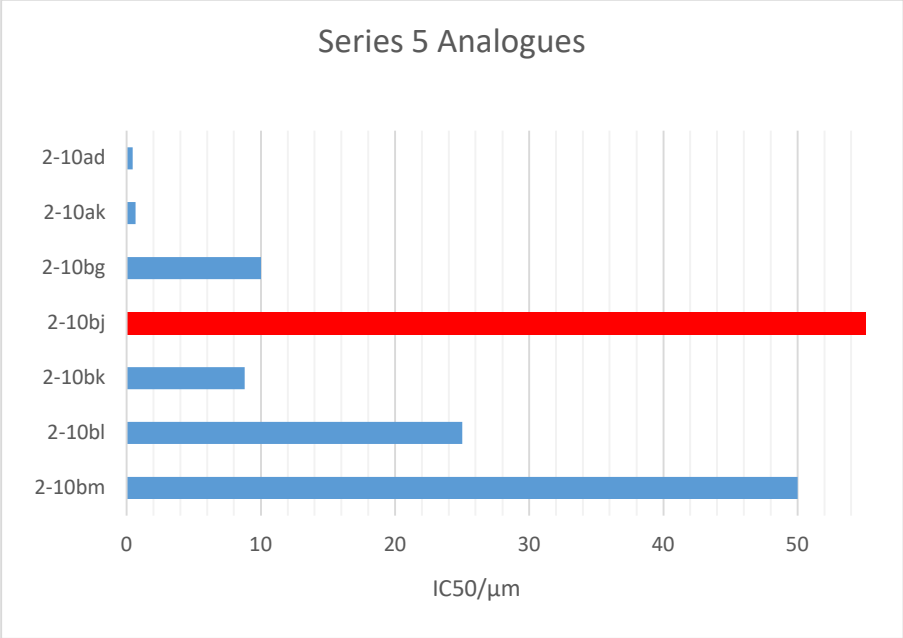


Table 2.41: IC_{50} values of series 5 analogues, compared to the two base compounds 2-10ad and 2-10bg as comparison. Red labels indicate no activity measured up to 2mM.

Among series 5, 3-fluoro phenyl analogues, the carboxylic acid derivative **2-10bl** (25 µM) has poorer activity than the methyl ester **2-10ad** (485nM) as expected, but puzzlingly, the n-butyl ester **2-10bj** showed no inhibition. The reason for this anomaly is uncertain, but it could be that the binding pocket of 3-F analogues and gallate analogues are different, and steric hindrance between the butyl ester and the binding pocket of 3-flouoro analogues may be a factor in the loss of activity. As for the gallic acid analogues ATP synthase inhibitory activity was also shown, although IC_{50} values were higher than the methyl ester. However, the n-butyl analogue **2-10bk** showed slightly greater activity (8.8µM) compared to the methyl ester analogue **2-10bg** (10µM), and the carboxylic acid analogue **2-10bm** had significantly lower activity (50µM), which was in line of expectations. It is interesting to note that replacement of the methyl group with a hydrazine moiety only slightly reduced its activity (659nM) despite being polar. Perhaps steric factors for the fluoro-substituted analogue may have some effect.

When it came to miscellaneous analogues, **2-10bn** had an IC_{50} of 223nM, which was better than the original molecule EpMF2 structure as well as **2-10x**. The diketopiperazine molecule

2-10bo, however, had no measurable IC_{50} activity, which was not surprising as it has a vastly different structure compared to other analogues.

Regardless of its IC_{50} activity, none of the analogues from Series 4 or 5, as well as the miscellaneous analogues, has any inhibitory effect on live *M. smegmatis*. It seemed unlikely that further synthesis of more analogues would lead to any resolution of this problem. A new approach to this problem was required. This will be further elaborated in the next chapter.

2.4 Conclusion

EpMF2 and its analogues was discovered to be an effective inhibitor of mycobacterial ATP synthase, as shown from IMV assays. From the syntheses of Series 4 and 5 analogues, the structural activity relationship was determined to be as follows:

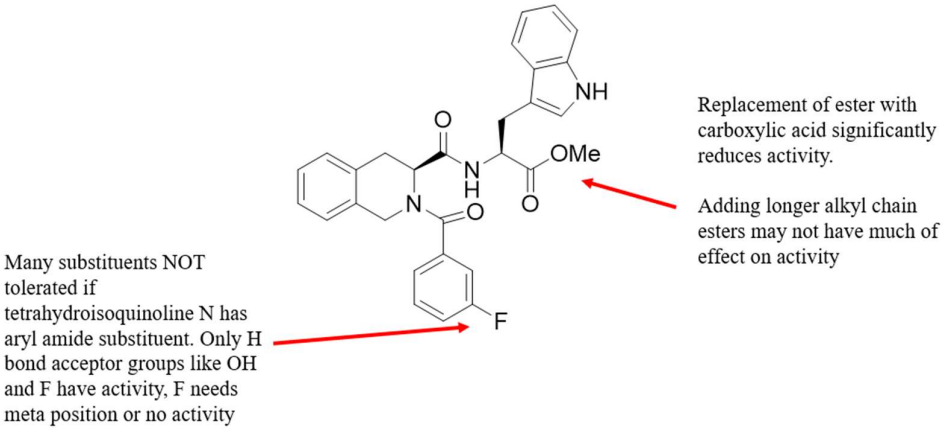


Figure 2.42: Structure-activity relationship of Series 4 and 5 analogues of EpMF2 discovered from testing.

Whatever the efficacy of EpMF2 in binding to the intended target in the ATP synthase, the inability by the compound to inhibit living bacteria remains a pressing issue. If EpMF2 is capable of inhibiting ATP synthase *in vitro* on IMVs, but not in living bacteria, it means the drug is not present in the cytosol at the inner membrane where the ε subunit is located. We hypothesized that EpMF2 was unable to pass through the thick hydrophobic cell wall that makes up mycobacterial species. A method to resolve this problem with a prodrug strategy will be elaborated on the next chapter.

2.5 Experimental

2.5.1 Chemical syntheses of EpMF2 analogues 4 and 5

All chemical compounds are obtained from commercial suppliers and are used as received. When anhydrous solvent was specified, in addition to using dried solvent, oven dried glassware was used (heated at 120°C for at least 2 hours), and the reaction was conducted under a nitrogen/argon atmosphere using Schlenk line techniques.

DCM was dried with CaH_2, then distilled and freshly used. DMF was dried with 4Å molecular sieves for overnight, then distilled at reduced pressure at 20 Torr. The distillation temperature was kept below 70 °C to prevent its decomposition. After distillation, DMF was then stored under a new batch of 4Å molecular sieves. THF, diethyl ether and toluene were dried with sodium-benzophenone until the solution turned blue from the formation of benzophenone ketyl radical, then distilled and freshly used. Methanol was first pre-dried with 3Å molecular sieves overnight, then treated with magnesium and iodine (0.5g Mg and 0.05g I_2 per 100mL of solvent), allowed to auto-reflux from the own heat generated, then thereafter distilled and stored under 3Å molecular sieves.

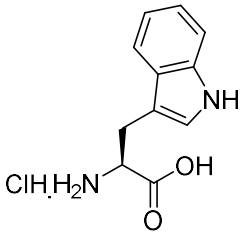
^1H NMR and ^{13}C NMR spectra were obtained on a 400MHz JEOL ECA400 machine. All spectra were recorded at 25°C. Chemical shifts were reported in parts per million (ppm) relative to tetramethylsilane (δ = 0 ppm) and were calibrated with respect to the residual solvent peak. Coupling constants were reported in Hertz. Coupling splitting patterns were labelled depending on multiplicity. (s for singlet, br s for broad singlet d for doublet, t for triplet, m for multiplet). When multiple peaks arising from the presence of interconverting rotamers exist, it will be indicated.

Nominal mass spectra were captured with a Thermo-Fischer Scientific LTQ XL linear ion trap mass spectrometer by the electrospray ionization method. Whether positive or negative mode was used was indicated.

High resolution mass spectra were captured with a Xevo® **G2-XS QTof** by the electrospray ionization method.

Melting points were measured with a Stuart SMP3 apparatus and were uncorrected.

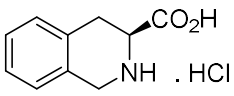
2.5.1.1 Synthesis of series 4 analogues: methyl esters



Methyl L-tryptophanate hydrochloride 2-12

A suspension of L-tryptophan **2-11** (24.5mmol, 5.00g) in dry methanol (50mL) was cooled down to 0°C. $SOCl_2$ (2eq, 49.0mmol, 3.55mL) was then added dropwise. The suspension gradually dissolved to a clear solution and was further left to stir for 6h at room temperature. The solvent was evaporated in vacuo, to obtain the product as a white solid. The compound was isolated as its hydrochloride salt **2-12**. (6.24g, 99%)

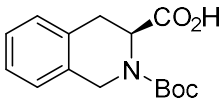
White solid; 99% yield; Mp 204-206°C (with decomposition); ^{1}H NMR (400MHz, DMSO-d6) δ = 11.12 (s, 1H), 8.26 (3H, br s), 7.50 (d, J=7.8 Hz, 1H), 7.37 (d, J=8.2 Hz, 1H), 7.25 (s, 1H), 7.09 (t, J= 8.2Hz, 1H), 7.01 (t, J= 8.2Hz, 1H), 4.21 (m, 1H), 3.65(s, 3H), 3.35-3.29 (m, 2H); ^{13}C NMR (100MHz, DMSO-d6) δ = 169.7, 136.2, 126.9, 125.0, 121.1, 118.6, 118.0, 111.5, 106.3, 52.6, 39.6, 26.1. HRMS Calculated for $C_{12}H_{15}N_2O_2$ $[M+H]^+$ 219.1134. Found 219.1128.



(S)-1,2,3,4-tetrahydroisoquinoline 3-carboxylic acid hydrochloride 2-14

L-phenylalanine **2-13** (24.2mmol, 4.00g) was suspended in concentrated HCl (37% m/v, 40mL) and stirred. To it was then added 37% m/v formaldehyde solution (2eq, 3.6mL) dropwise. The mixture was then heated at reflux until all the solids had dissolved and further stirred for 12h under reflux. The product was precipitated over the course of the reaction. After heating, the mixture was left to cool to room temperature, and the precipitate isolated by filtration. The solid was washed with cold water followed by cold acetone, then dried in vacuum to obtain the product as its hydrochloride salt **2-14** as a white solid (4.864g, 85%).

White solid; 85% yield; Mp 275-278°C (with decomposition) ^{1}H NMR (400MHz, DMSO-d_6) δ = 10.1 (br s, 2H), 7.27-7.24 (m, 4H), 4.38 (dd, J= 4.6, 10.6Hz, 1H), 4.31 (s, 2H), 3.31 (dd, J= 5.0, 17.0Hz, 1H), 3.14 (dd, J= 11.0, 17.0Hz, 1H) ; ^{13}C NMR (100MHz, DMSO-d_6) δ = 169.8, 130.8. 128.8, 128.5, 127.4, 126.9, 126.5, 53.0, 43.7, 28.1. HRMS Calculated for $C_{10}H_{12}N_2O_2$ $[M+H]^+$ 178.0868. Found 178.0864.

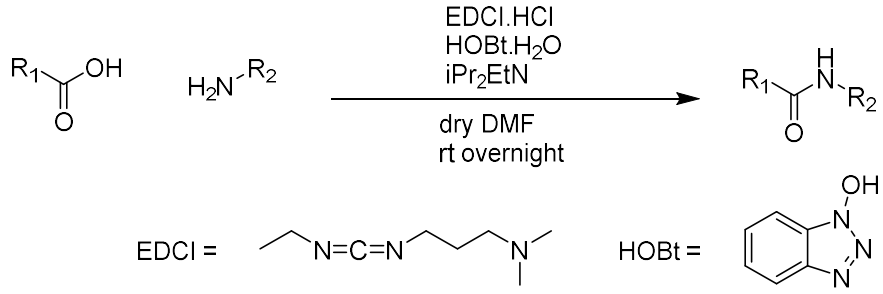


(S)-2-(tert-butoxycarbonyl)-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid 2-15

A suspension of (S)-1,2,3,4-tetrahydroisoquinoline 3-carboxylic acid hydrochloride salt **2-14** (14.04mmol, 3.00g) in a 1:1 H_2O/THF solvent mixture (45mL) was cooled to 0°C. To it was added NaOH (2.2eq, 1.24g) and stirred until all the solid dissolved into a clear solution, followed by di-tert-butyl dicarbonate (1.2eq, 3.87mL) dropwise. The mixture was left to warm to room temperature and stirred overnight. The mixture was then neutralised with addition of saturated $NaHSO_4$ solution until the pH is around 4. The mixture was then extracted with EtOAc 3 times. The combined organic layers were then washed with brine and dried with anhydrous $MgSO_4$. The solvent was then removed in vacuo to obtain the product as a colourless gummy solid (3.89g, quantitative) and used without further purification.

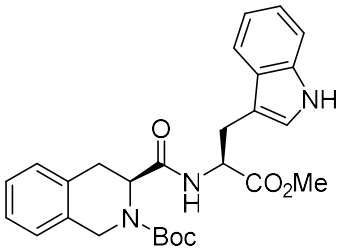
Colourless gummy solid; quantitative yield; ^{1}H NMR (400MHz, $CDCl_3$) (1.5:1 mixture of rotamers) δ =7.20-7.09 (m, 4H), 5.10 (br s, 0.6H), 4.75 (br s, 0.6H), 4.74-4.64 (m, 0.9H), 4.50-4.43 (m, 0.9H), 3.26-3.10 (m, 2H), 1.51(s, 5.4H), 1.41 (s, 3.6H) ; ^{13}C NMR (100MHz, $CDCl_3$) (mixture of rotamers) δ = 177.6, 176.6, 146.9, 134.0, 133.0, 131.9, 128.7, 128.0, 127.1, 127.0, 126.6, 126.4, 81.2, 54.3, 52.6, 44.7, 44.1, 31,5, 31,0, 28.6, 27.6. ES-MSI (+): m/z = 300.14 $[M+Na]^{+}$. HRMS calculated for $C_{15}H_{20}NO_4$ $[M+H]^{+}$. 278.1392. Found 278.1365

General procedure 1: Amide formation in peptide coupling



A mixture of carboxylic acid (1eq), EDCI (1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide)(1.2 eq) and HOBt monohydrate (1.2eqv) was dissolved in anhydrous DMF and stirred for 15 min at room temperature. The amine (1eq) was then added, followed by DIPEA (*N,N*-diisopropylethylamine) (2 eq, 3eq if the amine is in its protonated salt form), then the mixture was left to further stir overnight at room temperature.

The mixture was then diluted to 5 times its volume with water, then extracted with EtOAc 3 times. The organic layers were then washed with H_2O 4 times, then brine once, then dried with $MgSO_4$. Evaporation of the solvent in vacuo led to the crude residue which was further purified by column chromatography if necessary.

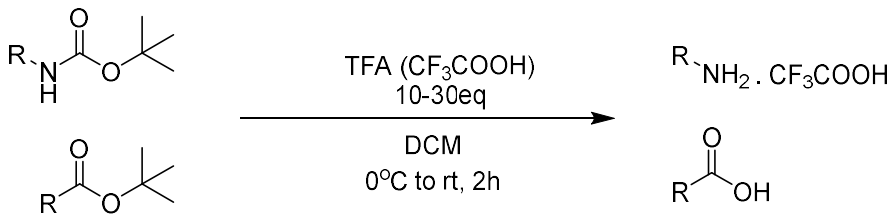


tert-butyl (S)-3-(((S)-3-(1H-indol-3-yl)-1-methoxy-1-oxopropan-2-yl)carbamoyl)-3,4-dihydroisoquinoline-2(1H)-carboxylate 2-16

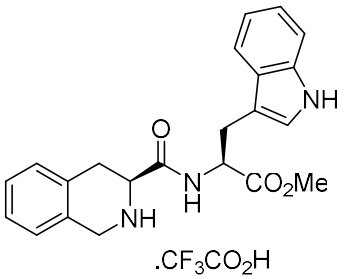
A mixture of N-Boc protected carboxylic acid **2-15** (13.3mmol, 3.69g), was reacted with L-tryptophan methyl ester hydrochloride **2-12** (13.3mmol, 2.91g) following general procedure 1 to obtain the crude residue, which was further purified by column chromatography (gradient eluent 40:60 to 50:50 hexane/EtOAc) to obtain the product as an off-white solid. (4.20g, 66%). Some side product, N-Boc tryptophan methyl ester, was also isolated (1.09g, 27%)

Off white solid; 42% yield; Mp: 62-65°C. ^{1}H NMR (400MHz, $CDCl_3$) (2.5: 1 mixture of rotamers) δ = 8.10-7.82 (m, 1H), 7.49-7.31 (m, 2H), 7.18-7.00 (m, 6H), 6.32 (s, 1H), 4.97-4.46 (m, 3H), 4.28 (app t), 3.60 (s, 2.1H), 3.51 (br s, 0.3H), 3.31-2.99 (m, 3H), 1.59 (br s, 6.4H), 1.41 (br s, 2.6H) ;^{13}C NMR (100MHz, $CDCl_3$), (mixture of rotamers) δ = 172.0, 171.5 (br s), 155.1 (br s), 136.1, 133.7, 128.6, 128.1, 127.6, 126.8, 126.3, 123.1, 122.8, 122.3, 119.7, 118.5, 111.3, 109.6, 81.5(br s), 56.3 (br s), 52.4, 44.6 (br s), 31.7 (br s), 28.3, 27.7 ; MS-ESI(+): m/z = 478.38 $[M+H]^+$ HRMS calculated for $C_{27}H_{32}N_3O_5$ $[M+H]^+$ 478.2342. Found 478.2339.

General procedure 2: Tert-butyl carbamate and ester deprotection



A solution of N-Boc (tert-butyloxycarbonyl) compound or tert-butyl ester (1eq) in DCM was cooled to 0°C and neat trifluoroacetic acid (10eq for Boc-amines, 30eq for tert-butyl esters) was then added dropwise into the mixture. The mixture was then stirred for 2h at room temperature. The volatiles were removed in vacuo, then the residue triturated with diethyl ether or hexane. The undissolved solid was then isolated with filtration, further washed with cold diethyl ether/hexane, and dried under vacuum to obtain the amine trifluoroacetate salt or carboxylic acid.



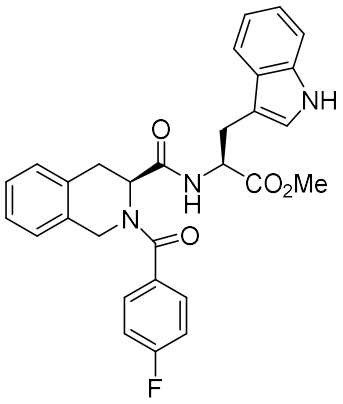
Methyl ((S)-1,2,3,4-tetrahydroisoquinoline-3-carbonyl)-L-tryptophanate trifluoroacetate 2-17

A solution of N-Boc protected compound **2-16** (4.10g, 8.59mmol) in DCM (40mL) was deprotected using general condition 2. The residue was triturated with diethyl ether and the undissolved solid was then isolated with filtration, then further washed with cold diethyl ether, and dried under vacuum to obtain **2-17** as the trifluoroacetate salt as an orange solid. (3.22g, 79%)

Orange solid; 78% yield; Mp 201-205°C (with decomposition); ^{1}H NMR (400MHz, methanol-d_4) (3:1 mixture of rotamers); δ = 7.59 (d, 0.25H, J= 7.8Hz), 7.55 (d, 0.75H, J= 8.3Hz), 7.39-7.01 (m, 9H), 4.96 (m, 0.25H), 4.90 (m, 0.75H), 4.37 (s, 1.5H), 4.32 (s, 0.5H), 4.12 (dd, 0.75H, J= 4.8, 12.0Hz), 4.06 (dd, 0.25H, J= 4.8, 12.0Hz), 3.77(s, 0.75H), 3.72 (s, 2.25H), 3.47-3.09 (m, 4.5H), 2.80 (dd, 0.25H, J= 4.8, 17.0Hz), 2.49 (dd, 0.25H, J= 12.0, 17.0Hz); ^{13}C NMR (100MHz, methanol-d_4), (mixture of rotamers) δ = 173.4, 169.6, 131.8, 130.0, 129.9, 129.4, 129.2, 128.9, 128.7, 128.6, 128.5, 127.6, 127.5, 124.8, 124.6, 122.6, 119.9, 119.1, 112.4, 110.5, 56.3, 55.0, 54.9, 52.9, 45.4, 30.6, 30.5, 28.6, 28.5. MS-ESI(+): m/z = 477.36 $[M]^+$ HRMS calculated for $C_{22}H_{24}N_3O_3[M^+]$ 378.1818. Found 378.1803.

Synthesis of analogues (2-10ba-2-10bf,2-10bh,2-10ad)

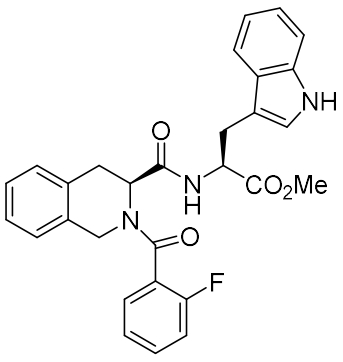
Note: Most of the analogues in series 4, excepting the ortho substituted ones, show significant peak broadening in 1H and 13C spectra due to presence of rotamers in the 2 amide groups and also arising from the different conformations of the 6-member piperidine ring.



2-10ba

2-17 was reacted with 4-fluorobenzoic acid according to general procedure 1. The crude compound was purified on column chromatography (eluent 40:60 hexane/EtOAc) to obtain the product as an off-white solid. (31.8mg, 63%).

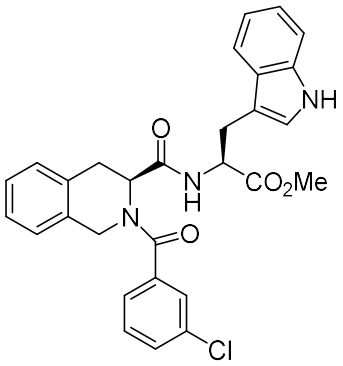
Off-white solid, 63% yield. Mp: 89-92°C. ^{1}H NMR (400Hz, $CDCl_3$), (complex mixture of rotamers) δ = 8.11, 8.06 (overlapping br s, 1H), 7.54-6.36 (m, 14H), 5.37-4.31 (m, 3H), 4.16 (app d, 1H), 3.69, 3.65 (overlapping br s, 3H), 3.36-3.10 (m, 4H); ^{13}C NMR (100MHz, $CDCl_3$) (complex mixture of rotamers), δ = 172.3, 171.4, 171.1, 170.3, 170.0, 165.0, 162.5, 136.3, 133.3, 131.5, 129.5, 129.1, 128.8, 128.6, 127.7, 126.8, 125.5 (br s),123.5 (br s), 122.5, 120.0, 118.7, 116.0, 115.8, 111.5, 109.8, 53.9, 52.9, 52.6, 48.3, 29.8, 27.5. ;ESI-MS(+): m/z = 500.11 $[M+H]^+$. HRMS calculated for $C_{29}H_{27}N_3O_4F$ $[M+H]^+$ 500.1986. Found 500.1977.



2-10bb

2-17 was reacted with 2-fluorobenzoic acid according to general procedure 1. The crude compound was purified on column chromatography (eluent 40:60 hexane/EtOAc) to obtain the product as an off-white solid. (37.1mg, 74%).

Off-white solid, 74% yield. Mp: 88-91°C, ^{1}H NMR (400Hz, $CDCl_3$), (complex mixture of rotamers) δ = 8.11-7.94 (m, 1H), 7.52-7.05 (m, 12.5H), 6.85-6.34 (2.5H), 5.51 (app t, 0.21H), 5.29 (m, 0.54H), 5.08 (app d, 0.26H), 4.91-4.87 (m, 0.74H), 4.60-4.56 (m, 0.40H), 4.49-4.40 (m, 0.48H), 4.38-4.20 (m, 1.37 H). 3.66-3.62 (m, 3H), 3.46-2.98 (m, 4H); ^{13}C NMR (100MHz, $CDCl_3$) (complex mixture of rotamers), δ = 172.3, 172.0, 170.0, 169.1, 167. 4 (br s), 136.5, 136.3, 136.2, 133.7, 133.3, 132.8, 132.0 (br s), 131.7, 131.2, 129.2, 128.9, 128.6, 127.9, 127.7, 127.5, 127.3, 127.2, 127.0, 126.7, 126.5, 125.7, 125.1, 124.1, 123.9, 123.7, 123.3, 123.0, 122.7, 122.2, 120.1, 119.7, 118.7, 118.6, 116.2, 116.0, 111.6, 111.3, 109.9, 109.6 (2 peaks), 53.7, 53.1, 52.8-52.5 (3 overlapping peaks), 47.2, 47.0, 30.9, 30.0, 29.1, 27.8 (2 peaks), 27.5. (multi ESI-MS(+): m/z = 500.14 $[M+H]^+$. HRMS calculated for $C_{29}H_{27}N_3O_4F$ $[M+H]^+$ 500.1986. Found 500.1978.

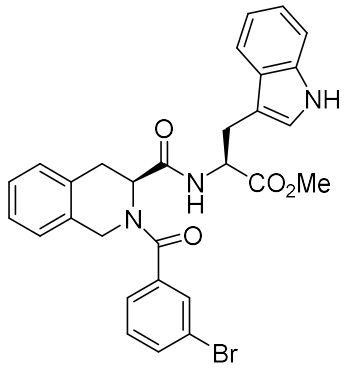


2-10bc

2-17 was reacted with 3-chlorobenzoic acid according to general procedure 1. The crude compound was purified on column chromatography (eluent 50:50 hexane/EtOAc) to obtain the product as an off-white solid. (27.9mg, 54%)

Off white solid, 54% yield; Mp: 96-98°C ,^{1}H NMR (400Hz, CDCl3), (complex mixture of rotamers) δ = 8.10 (br s, 1H), 7.55-6.30 (m, 14H), 5.42-4.27 (m, 4H), 3.70 (br s, 3H), 3.36-2.90

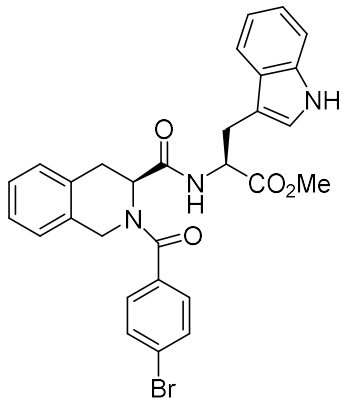
(m, 4H). ^{13}C NMR (100MHz, $CDCl_3$), δ = 172.4, 170.5, 170.3, 137.3, 136.3, 134.8, 133.8, 133.3, 130.5, 130.2, 128.8, 128.5, 128.0, 127.7, 127.3, 126.9, 125.5, 125.2, 123.5, 122.7, 122.4, 119.9, 118.7, 111.5. 109.9, 53.9, 53.0, 52.7, 29.9, 27.6.; MS-ESI(+): m/z = 516.30 [^{35}Cl M+H]$^+$,100 , 518.27 [^{37}Cl M+H]$^+$,34. HRMS calculated for $C_{29}H_{27}N_3O_4Cl$ [^{35}Cl M+H]$^+$ 516.1690. Found 516.1693.



2-10bd

2-17 was reacted with 3-bromobenzoic acid according to general procedure 1. The crude compound was purified on column chromatography (eluent 50:50 hexane/EtOAc) to obtain the product as an off-white solid. (27.5mg, 49%)

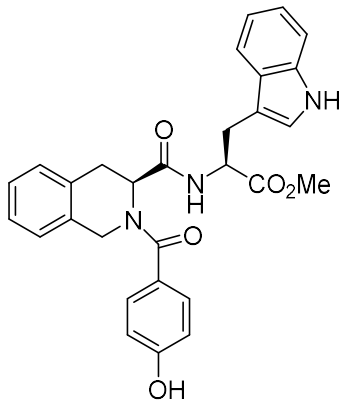
Off white solid, 54% yield; Mp: 94-96°C ,^{1}H NMR (400Hz, $CDCl_3$), (complex mixture of rotamers) δ = 8.08, 8.04 (overlapping br s, 1H), 7.59-6.29 (m,14H), 5.37-4.56 (m, 3H), 4.41-4.15 (m, 1H), 3.71, 3.66 (overlapping br s, 3H), 3.37-2.98 (m, 4H); ^{13}C NMR (100MHz, $CDCl_3$), δ = 172.4, 170.6, 170.3, 137.5, 136.3, 133.8, 133.4, 130.4, 130.1, 128.7, 128.4, 128.0, 127.7, 126.8, 126.7, 125.7, 125.5, 123.5, 122.8, 122.4, 119.9, 118.7, 111.5, 109.9, 109.7, 53.9, 53.0, 52.7, 48.2, 29.9, 27.6. MS-ESI(+): m/z = 560.27 [^{79}Br M+H,]$^+$ 21, 562.31[^{81}Br M+H]$^+$, 18. HRMS calculated for $C_{29}H_{27}N_3O_4{}^{79}Br$ [^{79}Br M+H]$^+$ 560.1185. Found 560.1180.



2-10be

2-17 was reacted with 4-bromobenzoic acid according to general procedure 1. The crude compound was purified on column chromatography (eluent 50:50 hexane/EtOAc) to obtain the product as an off-white solid. (42.1mg, 75%).

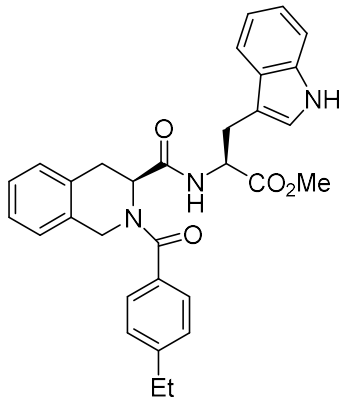
Off-white solid, 75% yield. Mp: 98-100°C, ^{1}H NMR (400Hz, $CDCl_3$) (complex mixture of rotamers), δ = 8.09, 8.03 (overlapping br s, 1H), 7.52-6.32 (m, 14H), 5.37-4.27 (m, 3H), 4.16 (app d, 1H), 3.70, 3.65 (overlapping br s, 3H), 3.37-2.95 (m, 4H); ^{13}C NMR (100MHz, $CDCl_3$) (complex mixture of rotamers), δ = 172.4, 171.4, 171.1, 170.3, 136.3, 134.3, 134.0, 133.3, 132.1, 128.8, 128.6, 128.0, 127.7126.8, 126.7, 125.5, 124.7, 123.4, 122.5, 120.0, 118.7, 111.5, 109.9, 109.7, 53.8, 53.0, 52.6, 48.2, 29.8, 27.6. ;ESI-MS(+): m/z = 582.32 [^{79}Br M+Na]$^{+}$,100 , 584.32 [^{81}Br M+Na]$^{+}$,88. HRMS calculated for $C_{29}H_{27}N_3O_4Br$ [^{79}Br M+H]$^{+}$ 560.1185. Found 560.1166.



2-10bf

2-17 was reacted with 4-hydroxybenzoic acid according to general procedure 1. The crude compound was purified on column chromatography (eluent 40:60 hexane/EtOAc) to obtain the product as an off-white solid. (12.0mg, 24%). A diester was also isolated (4.0mg, 8%) (phenolic -OH forming ester bond with an additional carboxylic acid of 4-hydroxybenzoic acid).

Off white solid. 24% yield. ^{1}H NMR (400Hz, acetone d_6)(complex mixture of rotamers) δ = 10.10 (br s, 1H), 8.83 (br s, 1H), 7.55-6.78 (m, 14H), 5.35-4.10 (m, 4H), 3.62, 3.59 (overlapping br s, 3H), 3.28-3.06 (m, 4H); ^{13}C NMR (100MHz, acetone d_6)(mixture of rotamers) δ = 173.3, 172.9, 172.1, 171.0, 137.5, 134.2 (br s), 133.2, 132.7, 132.4, 131.9, 131.6, 130.3, (br s), 129.1, 128.4, 127.6 (br s), 127.1 (br s), 126.5 (br s), 124.3(br s), 123.0, 122.8, 122.2, 119.7, 119.2, 119.1, 116.4, 115.8, 112.2 (2 overlapping peaks), 110.8, 110.5, 60.5, 53.8, 53.6, 52.3, 28.9. (;ESI-MS(-): m/z = 496.41 $[M-H]^-$. HRMS calculated for $C_{29}H_{28}N_3O_5$ $[M+H]^+$ 498.2029. Found 498.2014.

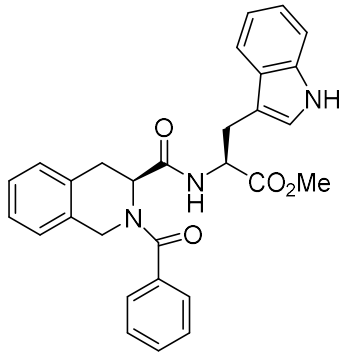


2-10bh

2-17 was reacted with 4-ethylbenzoic acid according to general procedure 1. The crude compound was purified on column chromatography (eluent 50:50 hexane/EtOAc) to obtain the product as an off-white solid. (27.5mg, 54%)

Off white solid, 54% yield; Mp: 83-86°C, ^{1}H NMR (400Hz, CDCl3), (complex mixture of rotamers) δ = 8.27, 8.20 (overlapping br s, 1H), 7.33-6.46 (m, 14H), 5.40-4.40 (m, 3H), 4.16

(app d, 1H), 3.67 (s, 3H), 3.35-2.90 (m, 4H), 2.67 (br q, may not be visible at lower temp, J= 7.2Hz, 2H), 1.26 (br t, may not be visible at lower temp, J= 7.2Hz, 3H) ^{13}C NMR (100MHz, $CDCl_3$), (complex mixture of rotamers) δ = 172.2, 170.4 (overlapping broad peaks), 146.6, 135.7 (br s), 133.3 (br s), 132.5, 128.5-127.1 (overlapping broad preaks), 126.7, 126.5, 125.2, 123.3, 122.1, 119.6, 118.4, 111.3, 53.5 (br s), 52.7 (br s), 52.4, 48.0 (br s), 29.7 (br s), 28.7, 27.3, 15.3. MS-ESI(+): m/z = 510.29$[M+H]^+$ HRMS calculated for $C_{31}H_{32}N_3O_4$ $[M+H]^+$, 510.2393. Found 510.2390.

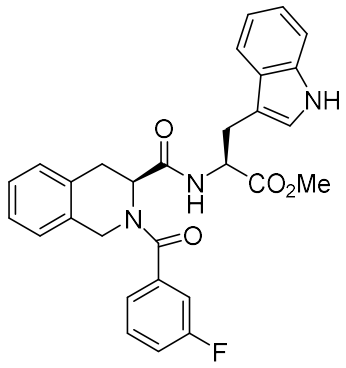


2-10bi

A solution of **2-17** (0.100mmol, 49.1mg**)** in dry DCM (1mL) was added triethylamine (2eq, 28μL), then cooled to 0°C. Benzoyl chloride (1.1eqv, 13μL) was added dropwise, and the mixture warmed to room temperature and further stirred for 2 hours. When TLC shows completion of reaction, the reaction was diluted with H_2O (5mL) and extracted with DCM twice. The combined organic layers were washed with H_2O, then 1M HCl followed by saturated $NaHCO_3$ solution respectively. The solvent was then dried with $MgSO_4$. Filtration and evaporation gave the crude product as a brown gum, which upon purification by column chromatography (eluent 60:40 hexane/EtOAc) afforded the product as a white solid (24.6mg, 51%)

White solid, 51% yield. Mp: 86-89°C, ^{1}H NMR (400Hz, $CDCl_3$), (complex mixture of rotamers) δ = 8.18-8.05 (m, 1H), 7.53-6.44 (m, 15H), 5.44-4.34 (m, 3H), 4.16 (app d, 1H), 3.68, 3.64

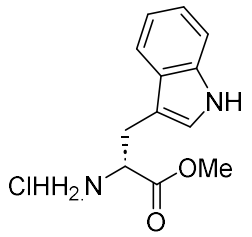
(overlapping br s, 3H), 3.48-2.90 (m, 4H); ^{13}C NMR (100MHz, $CDCl_3$), (complex mixture of rotamers) δ = 172.4, 172.2, 170.5, 170.1, 136.3, 135.5, 133.9 (br s), 133,4 (br s), 130.3, 128.8 (br s), 127.8, 127.7, 127.1, 126.7 (br s), 125.5, 123.7, 123.5, 122.8 (br s), 122.4, 119.9, 118.7, 111.5, 109.9 (br s), 109.7(br s), 53.7, 53.0, 52.6, 29.8, 27.6. ESI-MS(+): m/z = 482.35 $[M+H]^+$. HRMS calculated for $C_{29}H_{28}N_3O_4$ $[M+H]^+$ 482.2080. Found 482.2079.



2-10ad (resynthesis)

2-17 (1000mg, 2.04mmol) was reacted with 3-fluorobenzoic acid (285.1mg 1.0eq) according to general procedure 1. The crude compound was purified on column chromatography (eluent 50:50 hexane/EtOAc) to obtain the product as an off-white solid. (772mg, 79%)

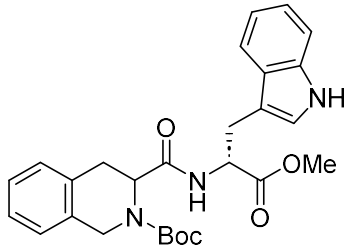
Off-white solid. 79% yield. Mp: 97-100°C. ^{1}H NMR (400Hz, $CDCl_3$), (complex mixture of rotamers) δ = 8.02 (br s, 1H), 7.53-6.31 (m, 13H), 5.42-4.15 (m, 4H), 3.70, 3.66 (overlapping br s, 3H), 3.45-2.96 (m, 4H); ^{13}C NMR (100Hz, $CDCl_3$), (mixture of rotamers) δ = 172.4, 170.6(br s), 164.0, 161.4, 137.5, 136.3, 133.8, 133.3, 130.6, 128.8, 128.5, 128.0, 127.7, 127.2 (br s), 126.8 (2 overlapping peaks), 125.5 (2 overlapping peaks),123.5, 122.8, 122.4, 119.9, 118.7, 117.5, 117.3, 114.5, 114.3 (br s), 111.5, 109.8, 109.7, 53.9, 52.6, 48.1, 29.9, 27.5. (ESI-MS(+): m/z = 500.38 $[M+H]^+$. HRMS calculated for $C_{29}H_{27}N_3O_4F$ $[M+H]^+$ 500.1986. Found 500.1985.



Methyl D-tryptophanate hydrochloride 2-12a

D-tryptophan **2-11a** (1.00g, 4.90mmol) was reacted with $SOCl_2$ (2eq, 9.79mmol, 0.71mL) in anhydrous methanol (15mL) in identical conditions as L-tryptophan to obtain the product as hydrochloride salt white solid (1.250g, quantitative yield).

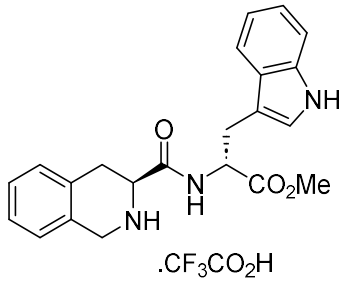
White solid, quantitative yield; ^{1}H NMR (400MHz, DMSO-d6) δ = 11.12 (s, 1H), 8.26 (3H, br s), 7.50 (d, J=7.8 Hz, 1H), 7.37 (d, J=8.2 Hz, 1H), 7.25 (s, 1H), 7.09 (t, J= 8.2Hz, 1H), 7.01 (t, J= 8.2Hz, 1H), 4.21 (m, 1H), 3.65(s, 3H), 3.35-3.29 (m, 2H). (^{13}C NMR and MS data were not obtained and were assumed to be identical to **2-12**)



tert-butyl 3-(((R)-3-(1H-indol-3-yl)-1-methoxy-1-oxopropan-2-yl)carbamoyl)-3,4-dihydroisoquinoline-2(1H)-carboxylate 2-16a

A mixture of *N*-Boc protected carboxylic acid **2-15** (1.19mmol, 304.0mg), was reacted with L-tryptophan methyl ester hydrochloride **2-12a** (1.19mmol, 279.0mg) following general procedure 1 to obtain the crude residue, which was further purified by column chromatography (gradient eluent 40:60 to 50:50 hexane/EtOAc) to obtain the product as a foamy off-white solid. (415mg, 79%).

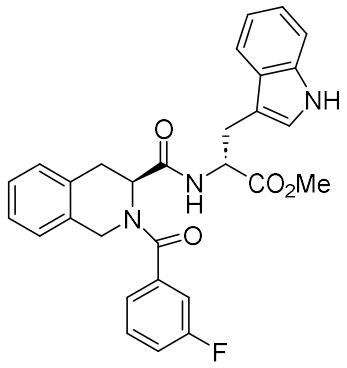
Off-white solid, 79%.[1]H NMR (400MHz, $CDCl_3$),(mixture of rotamers) δ = 8.12-8.08 (m, 1H), 7.48 (d, J= 8.0Hz, 1H), 7.36 (d, J= 8.0Hz, 1H), 7.17-6.22 (m, 5H), 4.97-4.59 (m, 4H), 4.33-4.29 (apparent d, 1H), 3.51 (s, 3H), 3.28-2.99 (m, 4H), 1.47-1.36 (two overlapping s, 9H).



Methyl ((S)-1,2,3,4-tetrahydroisoquinoline-3-carbonyl)-D-tryptophanate trifluoroacetate 2-17a

A solution of N-Boc protected compound **2-16a** (400mg, 0.838mmol) in DCM (4mL) was deprotected using general condition 2. The residue was triturated with diethyl ether and the undissolved solid was then isolated with filtration, then further washed with cold diethyl ether, and dried under vacuum to obtain the trifluoroacetate salt as an orange solid. (358mg, 87%)

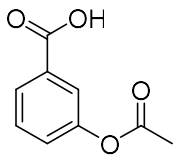
Orange solid, 87%; [1]H NMR (500MHz, methanol-d_4); δ = 8.66 (d, J= 8.4Hz, 0.6H), 7.59 (d, J= 8.0Hz, 1H), 7.38 (d, J= 8.0Hz, 1H), 7.26-7.23 (m, 2H), 7.18-7.13 (m, 3H), 7.06 (t, J= 8.0Hz, 1H), 7.02-7.01 (m, 1H), 4.98-4.94 (m, 1H), 4.32 (s, 2H), 4.06 (dd, J= 5.0, 12.0Hz, 1H), 3.77 (s, 3H), 3.42 (dd, J= 5.0, 14.6Hz, 1H), 3.20 (dd, J= 5.0, 9.6Hz, 1H), 2.79 (dd, J= 5.0, 16.8Hz, 1H), 2.49 (dd, J= 12.0, 16.8Hz), [13]C NMR (125MHz, methanol-d_4), (mixture of rotamers) δ = 173.4, 169.3, 138.1, 131.6, 129.9, 129.2, 128.9, 128.6, 128.5, 127.5, 124.7, 122.6, 119.9, 119.3, 112.4, 110.6, 56.2, 54.7, 52.9, 45.3, 30.6, 28.6.



2-10ad2

2-17 (51.3mg, 0.104mmol) was reacted with 3-fluorobenzoic acid (20.0mg 1.0eq) according to general procedure 1. The crude compound was purified on column chromatography (eluent 50:50 hexane/EtOAc) to obtain the product as an off-white solid. (20mg, 28%)

Off-white solid. 28% yield. ^{1}H NMR (500MHz, $CDCl_3$), (complex mixture of rotamers) δ = 8.13 (s, 1H), 7.53-7.47 (m, 1H), 7.34-7.33 (m, 2H), 7.20-7.04 (m, 7.6H), 6.96-6.88 (m, 2.4H), 6.84-6.78 (m, 1H), 6.25 (0.2H, br s), 5.43-5.41 (m, 0.8H), 5.17-5.15 (m, 0.2H), 4.88-4.87 (m, 0.8H), 4.71 (br s, 0.2H), 4.52-4.42 (m, 2H), 3.66 (s with shoulder peak at 3.56, 3H), 3.43-3.05 (m, 4H); ^{13}C NMR (100MHz, $CDCl_3$), (mixture of rotamers) δ = 172.4, 170.9, 169.9, 164.0, 161.5, 137.5, 137.4, 136.4, 133.3, 132.2, 130.7, 128.8, 127.8, 127.6, 126.7, 125.6, 123.4, 122.7, 122.5, 120.0, 118.7, 117.5, 117.3, 114.5, 114.3, 111.5, 109,8, 52.7, 52.6, 28.7, 27.6.

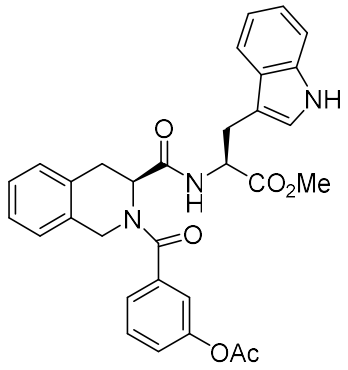


3-acetoxybenzoic acid 2-19

A solution of 3-hydroxybenzoic acid **2-18** (1500mg, 10.9mmol) in pyridine (10mL) was cooled to 0°C and to it added acetic anhydride (2.06 mL, 16.4mmol, 2eq). The mixture was then heated to reflux for 2 hours. After that, the mixture was poured to ice cold water (30mL) and acidified with concentrated HCl (12M) until the pH turned acidic, then extracted with Et_2O twice. The combined organic extracts are washed with H_2O, then brine, dried with $MgSO_4$, and volatiles

evaporated in vacuo to obtain the product as a white solid (1.75g, 90%), which was analytically pure enough to not require further purification.

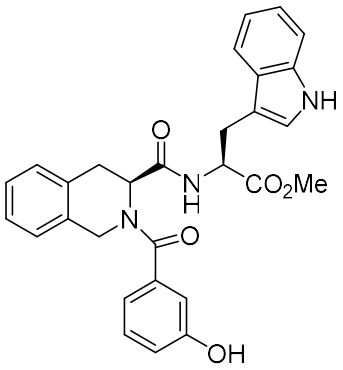
White solid. 90% yield. Mp 128-130°C. ^{1}H NMR (400Hz, $CDCl_3$), δ = 7.99 (d, 1H, J =7.6Hz), 7.83, (t, 1H, J =1.6Hz), 7.50 (t, 1H, 8.0Hz), 7.37(dt, 1H, J= 1.6Hz, 8.0Hz), 2.33 (s, 3H). ^{13}C NMR (100Hz, $CDCl_3$) δ = 171.3, 169.4, 150.8, 131.0, 129.8, 127.8, 127.4, 123.6, 21.2. HRMS calculated for $C_9H_7O_4$ $[M-H]^-$ 179.0344. Found 179.0341.



2-10ae2

2-17 (2.50g,4.63mmol) was reacted with 3-acetoxybenzoic acid (916.5mg, 4.63mmol) according to general procedure 1. The crude compound was purified by column chromatography (eluent 50:50 hexane/EtOAc) to obtain the product as an off-white solid. (878mg, 74%). A side product of deacetylated compound was also isolated (507mg, 20%).

Off white solid, 80% yield. ^{1}H NMR (400Hz, $CDCl_3$), (complex mixture of rotamers) δ = 8.30-8.10 (overlapping s, 1H), 7.54-6.63 (m, 15H), 5.17-4.65 (m, 2H), 4.36-3.97 (apparent dd, 1H), 3.71 (s, 3H), 3.41-3.05 (m, 4H), 2.36 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$)(complex mixture of rotamers) δ = 172.4, 170.9, 170.3, 169.8, 150.8, 136.6, 136.4, 133.9, 133.2, 131.9, 129.9, 128.5, 127.9, 127.5, 126.7, 125.5, 124.7,123.6, 123.5, 122.3, 121.2, 119.8, 118.7, 111.4, 109.8, 53.7, 52.8, 52.6, 48.0, 29.5, 27.7, 21.3, 14.4. ESI-MS(+): m/z = 540.15$[M+H]^+$ HRMS for $C_{31}H_{30}N_3O_6$ $[M+H]^+$ 540.2135. Found 540.2144.

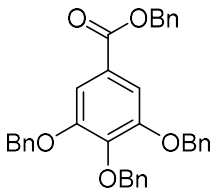


2-10ae (resynthesis)

To a solution of **2-10ae2** (2.00g, 3.71mmol) in MeOH (30mL) was added a solution of K_2CO_3 (563.5mg, 1.1eq) in water (15mL). The mixture was then stirred vigorously for 30 min at room temperature. When TLC showed completion of reaction the mixture was evaported in vacuo until approximately half its volume was left, then the mixture partitioned between saturated aq. NH_4Cl (40mL) and EtOAc. (40mL). The layers were seperated and the aqueous layer further extracted with EtOAc twice. The combined organic layers were washed with brine, dried with $MgSO_4$, and evaporated in vacuo to form the product **2-10ae** as an off-white solid, (1.807g, 98%) which was analytically pure enough to not require further purification.

Off-white solid. 98% yield. Mp 174-178°C. ^{1}H NMR (400Hz, $CDCl_3$), (complex mixture of rotamers) δ = 8.54-8.50 (overlapping s, 1H), 7.50-6.42 (m, 13H), 5.02-4.11 (m, 4H), 3.69 (s, 3H), 3.25-2.84 (m, 4H). ^{13}C NMR (100Hz, DMSO-d6) δ = 172.0 (apparent d), 170.2 (apparent d), 157.4, 137.4 (apparent d), 136.1, 133.6, 132.8, 131.8, 129.6, 127.9 (apparent d), 127.0, 126.3, 125.4, 123.8, 123.2, 120.9, 118.4, 118.0, 117.4, 116.6 (br s), 113.6 (apparent d), 111.4, 109.4, 56.6, 52.8 (br s), 51.7, 48.5 (apparent d), 43.4, 31.9, 30.2, 26.7. ESI-MS(+): m/z = 498.16$[M+H]^+$ HRMS calculated for $C_{29}H_{28}N_3O_5$ $[M+H]^+$ 498.2029. Found 498.2029.

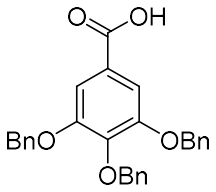
Synthesis of gallic acid analogues for Series 4 and 5



Benzyl 3,4,5-tris(benzyloxy)benzoate 2-21

To a mixture of gallic acid (2,3,4 trihydroxybenzoic acid) monohydrate **2-20** (1.00g, 5.32mmol) and Cs_2CO_3 (5eq, 8.66g) was added dry DMF (10mL). To the suspension was then added benzyl bromide (8eq, 5.06mL), and left to stir for 24h. When TLC indicated completion of reaction, the mixture was then quenched with water (100mL) and extracted with Et_2O twice. The organic layers were combined and washed with brine, then dried with $MgSO_4$. The solvent was removed in vacuo to obtain the crude residue, which was then used for the next reaction without prior purification. (1.334g crude, 47%)

White solid; 47% yield; ^{1}H NMR (400MHz, $CDCl_3$) δ = 7.43-7.31 (m, 20H), 7.28-7.24 (m, 2H), 5.33 (s, 2H), 5.13 (s, 4H), 5.12 (s, 2H).; ^{13}C NMR (100MHz, $CDCl_3$), δ = 165.9, 152.7, 142.5, 137.6, 136.8, 136.3, 128.8, 128.7, 182.4, 182.3, 128.2, 182.1, 127.7, 125.3, 109.5, 75.3, 71.4, 66.9. HRMS calculated for $C_{35}H_{31}O_5$ $[M+H]^+$ 532.2171. Found 531.2156.

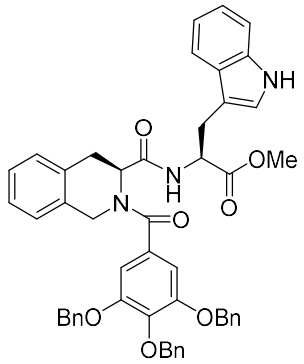


3,4,5-tris(benzyloxy)benzoic acid 2-22

A 5M solution of aqueous NaOH (5.00g in 25mL) was mixed with ethanol (25mL). To this solvent mixture was added (from the previous reaction) crude benzyl 3,4,5-tris(benzyloxy)benzoate **2-21** (1.334g). The mixture was heated until the solid dissolved into a clear solution, then further left to stir for 6h under reflux. After heating, the mixture was cooled to room temperature, and the organic volatiles evaporated in vacuo. The remaining solution

was then acidified by dropwise addition of concentrated HCl until the pH is about 2. The solid that was precipitated during the acidification was isolated by filtration and thoroughly washed with water until the pH of the wash was neutral. The solid was then purified by recrystallisation with MeOH to obtain the product **2-22** as a white solid. (920mg, 36% over 2 steps).

White solid; 36% yield; Mp: 187-190°C. ^{1}H NMR (400MHz, DMSO-d_6) δ = 7.47-7.27 (m, 17H), 5.16 (s, 4H), 5.02 (s, 2H); ^{13}C NMR (100MHz, DMSO-d_6), δ = 176.3, 151.8, 140.3, 137.5, 136.9, 128.4, 128.2, 128.1, 127.8, 127.5, 108.1, 74.2, 70.1. MS-ESI(-): m/z = 439.29 $[M-H]^-$ HRMS calculated for $C_{28}H_{27}O_5$ $[M-H]^-$ 439.1545. Found 439.1943

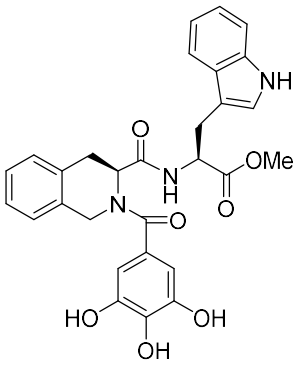


2-10bg2

2-17 (80.0mg, 0.163mmol) was reacted with 3,4,5-tris(benzyloxy)benzoic acid **2-22** (1eq, 71.7mg) according to general procedure 1. The crude compound was purified on column chromatography (eluent 50:50 hexane/EtOAc) to obtain the product as a pale yellow solid. (87.0mg, 67%).

Pale yellow solid, 67% yield. Mp: 76-81°C. ^{1}H NMR (400Hz, $CDCl_3$), (complex mixture of rotamers) δ = 8.12-7.81 (m, 1H), 7.47-6.22 (m, 26H), 5.46-4.82 (m, 8H), 4.65-3.98 (m, 2H), 3.68-3.61 (m, 3H), 3.47-2.70 (m, 4H); ^{13}C NMR (100MHz, $CDCl_3$)(mixture of rotamers) δ = 172.4 (br s), 171.5 (br s), 170.5, 169.6, 153.0 (br s), 140.1, 137.7, 136.8, 136.3, 133.7 (br s), 131.7 (br s), 128.7, 128.3, 128.1, 127.6, 126.7, 125.7, 123.6 (br s), 122.2 (br s), 119.8, 118.6, 111.4, 109.6, 107.3, 106.2, 75.3, 71.3, 58.0, 54.2, 53.1, 52.6, 48.2, 43.7, 30.1, 27.6, 21.2. ESI-

MS(+): m/z = 800.12 $[M+H]^+$. HRMS calculated for $C_{50}H_{45}N_3O_7Na$ $[M+Na]^+$ 822.3155. Found 822.3178.

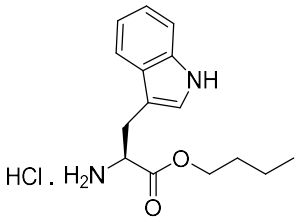


2-10bg

A solution of the substance **2-10bg2** (50.0mg, 0.06mmol) in methanol (1.5mL) was deoxygenated by bubbling with N_2 gas for 30 minutes. Palladium on carbon was then added (5.0mg, 10% mass), and the reaction vessel flushed with H_2 and then stirred overnight in a H_2 atmosphere (1 atm from balloon). When TLC shows completion of reaction, the reaction vessel was flushed with N_2 gas for 30min, the Pd/C removed by filtration through a pad of celite, and organic volatiles evaporated in vacuo to obtain the crude product, which was further purified by column chromatography (eluent 90:9:1 DCM/MeOH/AcOH) to afford the purified product **2-10bg** as an off white solid.(31mg, 94%).

Off-white solid, 94% yield. Mp: 122-146°C. ^{1}H NMR (400Hz, acetone d_6), (complex mixture of rotamers) δ = 10.01 (m, 1H), 8.17-6.90 (m, 10H), 6.58-6.55(m, 2H), 5.20-4.24 (m, 4H), 3.60, 3.59 (overlapping s, 3H), 3.26-3.05 (m, 4H); ^{13}C NMR (100MHz, acetone d6)(mixture of rotamers) δ = 173.0, 172.3, 171.3, 146.3, 137.6, 135.6, 134.6(br s), 129.1, 128.4, 127.7 (br s), 127.2, 126.7 (br s), 124.6(br s), 122.2, 119.6, 119.0, 112.2, 110.4, 107.8, 53.9, 52.3, 28.0. ESI-MS(-): m/z = 528.43. HRMS calculated for $C_{29}H_{27}N_3O_7Na$ $[M+Na]^+$ 552.1747, found 552.1740.

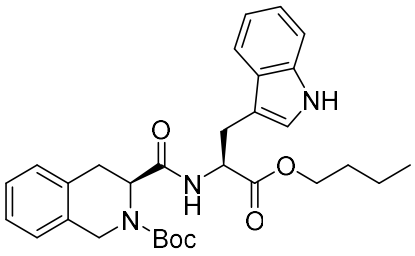
2.5.1.2 Synthesis of series 5 analogues: n-butyl and carboxylic acid analogues



L-tryptophan butyl ester hydrochloride 2-23

A suspension of L-tryptophan (9.8mmol, 2.00g) in n-butanol (40mL) was cooled down to 0°C. $SOCl_2$ (2eq, 4.90mmol, 1.42mL) was then added dropwise. After that, the mixture was heated to 110°C until the solid completely dissolved, and further left to stir for 12h. The product was precipitated upon cooling to room temperature and isolated by filtration. The precipitate was then washed with ice cold n-butanol, followed by diethyl ether and cold water. Drying under vacuum afforded the product as its hydrochloride salt, a white solid (2.292g, 80%)

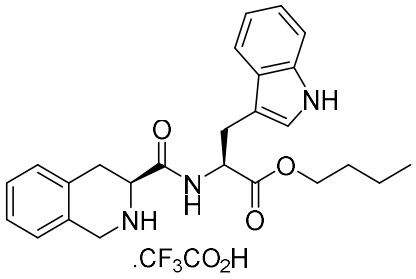
White solid, 80% yield. Mp: 220-223°C (with decomposition). ^{1}H NMR (400MHz, DMSO-d6) δ = 11.05 (s, 1H), 8.43 (br s, 3H), 7.50 (d, J= 7.8Hz, 1H), 7.37 (d, J= 8.2Hz, 1H), 7.22 (s, 1H), 7.09 (t, J= 7.8Hz, 1H), 7.01(t, J= 7.3Hz, 1H), 4.23 (m, 1H), 4.02 (m, 2H), 3.25 (m, 2H), 1.40 (m, 2H), 1.16 (m, 2H), 0.80 (t, J= 7.3 Hz, 3H) ;^{13}C NMR (100MHz, DMSO-d6), δ = 169.4, 136.2, 126.9, 124.7, 121.3, 118.5, 117.9, 111.5, 106.4, 65.2, 52.7, 29.7, 26.3, 18.2, 13.4. HRMS calculated for $C_{15}H_{21}N_2O_2$ $[M]^+$ 261.1603. Found 261.1598.



tert-butyl (S)-3-(((S)-1-butoxy-3-(1H-indol-3-yl)-1-oxopropan-2-yl)carbamoyl)-3,4-dihydroisoquinoline-2(1H)-carboxylate 2-24

N-Boc protected carboxylic acid **2-15** (2.00mmol, 555mg), was reacted with L-tryptophan butyl ester hydrochloride **2-23** (2.00mmol, 594mg) as in general procedure 1. The crude residue isolated was further purified by column chromatography (eluent 80:20 hexane/EtOAc to 50: 50 hexane/EtOAc)) to obtain the product as an off-white solid. (398g, 38%). Some side product, *N*-Boc tryptophan butyl ester, was also isolated (mass not determined).

Off white solid; 42% yield; ^{1}H NMR (400MHz, $CDCl_3$) (2.5: 1 mixture of rotamers) δ = 7.96 (br s, 0.3H), 7.81 (br s, 0.7 H), 7.52-7.05 (m, 9H), 6.36 (br s, 1H), 4.76-4.68 (m, 0.7H), 4.66 (app d, 0.7H), 4.49(app d, 0.7H), 4.32 (app d, 0.3H), 4.25 (app d, 0.7H), 4.05-3.87(m, 2H), 3.32-3.00 (m, 2.9H), 1.52-1.35(m, 11H), 0.88-0.82 (m, 3H); MS-ESI(+): m/z = 520.15 $[M+H]^+$ HRMS calculated for $C_{30}H_{38}N_3O_5$ $[M+H]^+$ 520.2811, Found 520.2794.

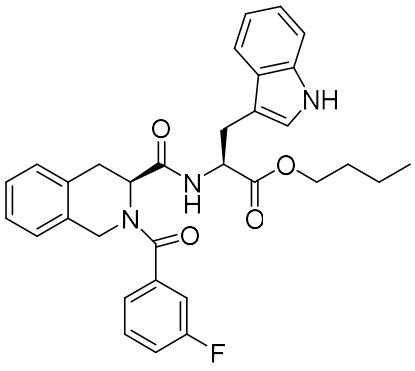


butyl ((S)-1,2,3,4-tetrahydroisoquinoline-3-carbonyl)-L-tryptophanate trifluoroacetate salt 2-25

N-Boc protected compound **2-24** (380mg, 0.796mmol) was deprotected using general procedure 2. The residue was triturated with hexane, and the undissolved solid was then isolated with filtration. It was then further washed with cold hexane and dried under vacuum to obtain the trifluoroacetate salt as an orange solid. (321mg, 81%)

Orange solid; 81% yield; Mp 165-168°C ^{1}H NMR (400MHz, methanol-d_4) (6:1 mixture of rotamers); δ = 7.59-7.55 (m, 2H), 7.35-7.03 (m, 8H), 4.38 (s, 1.7 H), 4.32 (s, 0.3H), 4.16-4.13 (m, 1.3H), 4.09 (t, J= 5.2Hz, 1.7H), 3.42-3.17 (m, 4.7H), 1.55-1.52 (m, 2H), 1.31-1.26 (m, 2H), 0.95-0.88 (overlapping t, 3H); ^{13}C NMR (100MHz, methanol-d_4), (mixture of rotamers) δ = 173.2, 169.6, 138.2, 131.7, 130.0, 129.4,128.8, 128.6, 127.7, 124.5, 122.6, 119.9, 119.1, 112.4,

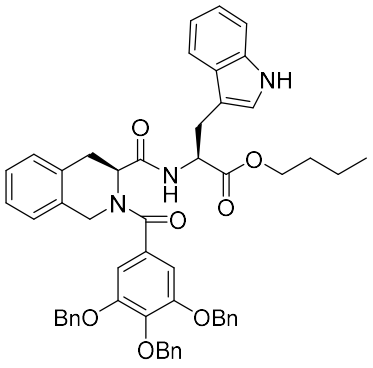
110.5, 66.3, 56.3, 55.1, 45.3, 31.6, 30.9, 28.6, 20.1, 14.0. MS-ESI(+): m/z = 420.32 $[M]^+$ HRMS calculated for $C_{25}H_{30}N_3O_3$ $[M+H]^+$. 420.2287. Found 420.2291.



2-10bj

2-25 was reacted with 3-fluorobenzoic acid according to general procedure 1. The crude compound was purified on column chromatography (eluent 60:40 hexane/EtOAc) to obtain the product as an off-white solid. (26.7mg, 24%)

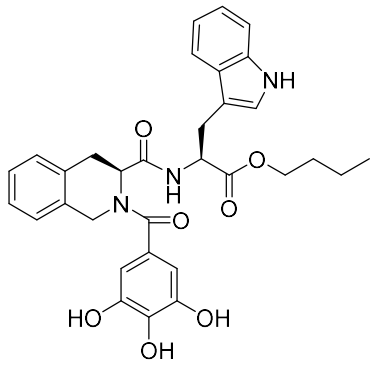
Off-white solid. 24% yield. Mp: 70-72°C. ^{1}H NMR (400Hz, $CDCl_3$), (complex mixture of rotamers) δ = 8.16, 8.10 (overlapping s, 1H), 7.53-6.36 (m, 14H), 5.37-4.06 (m, 6H), 3.37-2.96 (m, 4H), 1.55-1.43 (m, 2H), 1.36-1.25 (m, 2H), 0.91-0.87 (m, 3H); ^{13}C NMR (100MHz, $CDCl_3$)(mixture of rotamers) δ = 172.0, 170.5, 170.3, 164.0, 137.6, 137.5, 136.3, 133.8, 133.2, 130.5, 128.8, 128.5, 127.9, 127.7, 126.8, 125.5, 123.4, 122.8, 122.4, 119.8, 118.7, 117.4, 117.2, 114.5, 114.3, 111.5, 109.9, 65.7, 53.9, 53.1, 48.1, 30.6, 29.9, 27.6, 19.2, 13.8. ;ESI-MS(+): m/z = 542.31$[M+H]^+$. HRMS calculated for $C_{32}H_{33}N_3O_{4F}$ $[M+H]^+$, 542.2455. Found 542.2451.



2-10bk2

2-25 (80.0mg, 0.150mmol) was reacted with with 3,4,5-tris(benzyloxy)benzoic acid **2-22** (1eq, 66.1mg) according to general procedure 1. The crude compound was purified on column chromatography (eluent 60:40 hexane/EtOAc) to obtain the product as a pale yellow solid. (86.0mg, 68%).

Pale yellow solid. 68% yield. Mp: 64-66°C. ^{1}H NMR (400Hz, $CDCl_3$), (complex mixture of rotamers) δ = 8.16-7.88 (m, 1H), 7.52-6.29 (m, 27H), 5.47-4.00 (m, 12H), 3.40-2.68 (m, 4H), 1.70-1.50 (m, 2H), 1.39-1.22 (m, 2H), 0.90-0.81 (m, 3H); ^{13}C NMR (100MHz, $CDCl_3$), (complex mixture of rotamers) δ = 171.9 (br s), 170.4 (br s), 169.6 (br s), 168.9 (br s), 153.0, 140.0 (apparent d), 137.7, 136.9, 135.2, 133.9 (br s), 130.7, 128.7, 128.4, 128.2, 127.7, 126.8 (br s), 125.6 (br s), 123.5 (br s), 123.0 (br s), 122.4, 119.8, 118.6 (apparent d), 111.4, 109.8, 109.5, 107.4, 106.8, 106.2, 75.4, 71.4, 65.7, 54.1, 53.1, 52.5, 30.6, 30.0, 27.7, 19.2, 13.8. ESI-MS(+): m/z = 842.13 $[M+H]^+$. HRMS calculated for $C_{53}H_{52}N_3O_7$ $[M+H]^+$, 842.3805. Found 842.3819

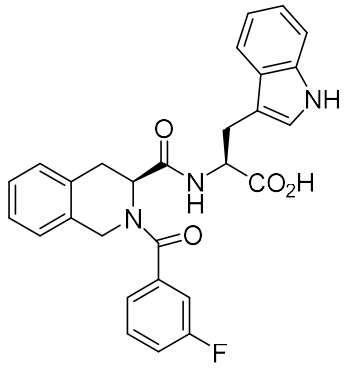


2-10bk

Compound **2-10bk2** (75.0mg, 0.089mmol) was in methanol (1.5mL) was deoxygenated by bubbling with N_2 gas for 30 minutes. Palladium on carbon was then added (7.50mg, 10% mass), and the reaction vessel flushed with H_2 and then stirred overnight in a H_2 atmosphere (1 atm from balloon). When TLC shows completion of reaction, the reaction vessel was flushed with N_2 gas for 30min, the Pd/C removed by filtration through a pad of celite, and organic volatiles evaporated in vacuo to obtain the crude product. The crude compound was further purified by

column chromatography (eluent 25:74:1 hexane/EtOAc/AcOH) to obtain the product as an off-white solid. (47.1mg, 93%).

Off-white solid, 93% yield. Mp: 72-76°C. ^{1}H NMR (400Hz, $CDCl_3$), (complex mixture of rotamers) δ = 10.02 (m, 1H), 8.16-6.90 (m, 11m), 6.58, 6.56 (overlapping s, 1H), 5.20-4.20 (m, 2H), 4.01-3.97 (m, 2H), 3.29-3.03 (m, 2H), 1.59-1.40 (m, 2H), 1.29-1.21 (m, 2H), 0.92-0.82 (m, 3H); ^{13}C NMR (100MHz, acetone d6), (complex mixture of rotamers) δ = 172.6, 172.2, 171.1, 146.3, 146.0, 137.6, 135.6, 134.4 (br s), 129.0, 128.4, 127.6 (br s),127.1, 126.5 (br s), 124.5, 122.2, 119.6, 119.1, 112.2, 110.5, 110.1, 107.8, 65.4, 54.0, 31.3, 28.2, 19.6, 13.9. ESI-MS(-): m/z =570.55 $[M-H]^-$. HRMS mass calculated for $C_{32}H_{33}N_3O_7$ $[M+H]^+$, 572.2397. Found 572.2411.

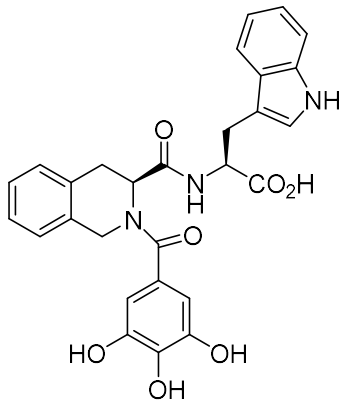


2-10bl

2-10ad (100.0mg, 0.200mmol) was dissolved in a 4:1 THF/H_2O v/v solvent mixture and stirred. To the mixture was then added LiOH.H_2O (3eq, 20.0mg) and stirring continued for 1h. When TLC indicates completion of reaction, the organic volatiles were removed in vacuo until the volume was reduced to half, and to the remaining solution was neutralised with addition of saturated $NaHSO_4$ solution until the pH was about 3. The mixture was then extracted with EtOAc twice and the combined organic layers dried with $MgSO_4$, filtered, and evaporated in vacuo to obtain the product as a white solid. (89mg, 91%)

White solid, 91H% yield. Mp: 221-225°C (with decomposition). ^{1}H NMR (400Hz, acetone-d_6), (complex mixture of rotamers) δ = 10.15 (br s, 1H), 7.61-6.85 (m, 14H), 5.28-4.13 (m, 4H),

3.29-3.00 (m, 4H); ^{13}C NMR (100Hz, acetone d6) (peak broadening from rotamer exchange with relatively low solubility of compound means poor signal-noise ratio and not all peaks may be detected) δ = 178.0, 173.3, 170.7, 131.4, 129.5, 127.6, 127.3, 126.4, 123.9, 122.1, 119.6, 112.1, 110.6, 53.7, 48.3, 27.8. ESI-MS(-): m/z = 484.37 $[M-H]^-$. HRMS calculated for $C_{28}H_{23}N_3O_4F$ $[M-H]^-$ 484.1673. Found 484.1679.

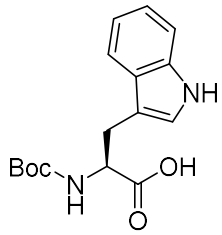


2-10bm

Compound **2-10bg** (30.0mg, 0.056mmol) was dissolved in aqueous 2M NaOH solution and stirred for 2h. When TLC indicated completion of reaction, 2M HCl solution until the pH is about 2. The mixture was then extracted with EtOAc, (5mL), and the organic layer washed with brine (5mL) and dried with $MgSO_4$. Filtration and evaporation in vacuo afforded the product, which was purified by recrystallization from hot MeOH/H_2O to obtain the product as a brown solid. (21.0mg, 72%).

Brown solid, 72% yield. Mp: 235-241°C (decomposition without melting) ^{1}H NMR (400Hz, acetone-d_6), (complex mixture of rotamers) δ = 10.02 (s, 1H), 7.63-6.79 (m, 11H), 6.57 (app d, 1H), 5.30-4.10 (m, 4H), 3.50-3.02 (m, 4H). ^{13}C NMR (100Hz, acetone d6) (peak broadening from rotamer exchange with relatively low solubility of compound means poor signal-noise ratio and not all peaks may be detected) δ = 172.9 (br s), 137.5, 130.6, 129.5, 128.9, 128.4, 128.1, 127.7, 127.0, 124.2 (br s), 122.0, 119.6, 119.1, 112.1, 110.2, 53.6, 23.2. ESI-MS(-): m/z = 514.37 $[M-H]^-$. HRMS calculated for $C_{28}H_{24}N_3O_7$ $[M-H]^-$ 514.1614. Found 514.1606.

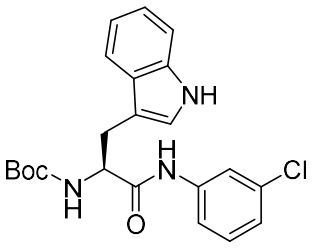
2.5.1.3 Synthesis of **2-10bn**



N-(tert-butoxycarbonyl)-L-tryptophan 2-26

Tryptophan **2-11** (2.45mmol, 500mg) was suspended in a THF/H_2O (1:1 by volume) (8mL) mixture. The mixture was cooled to 0°C and to it was added NaOH (2.2eq, 215.4mg) and stirred until all the solid dissolved to form a clear solution, followed by di-tert-butyl dicarbonate (1.2eq, 0.68mL) dropwise. The mixture was left to warm to room temperature and stirred overnight. The mixture was then neutralised with addition of saturated $NaHSO_4$ solution until the pH is around 4. The mixture was then extracted with EtOAc 3 times. The combined organic layers were then washed with brine and dried with anhydrous $MgSO_4$. The solvent was then removed in vacuo to obtain the product as a white solid. (713mg, 96%)

White solid; 96% yield; Mp: 139-141°C (with decomposition) ^{1}H NMR (400MHz, DMSO-d_6) δ = 12.52 (br s, 1H), 10.82 (s, 1H), 7.52 (d, J= 8.0Hz, 1H), 7.33 (d, J= 8.0Hz, 1H), 7.14 (s, 1H), 7.06 (t, J= 7.5Hz), 6.97 (t, J=8.0Hz, 1H), 4.16-4.12 (m, 1H), 3.14-3.11 (m, 1H), 2.99-2.95 (m 1H), 1.32 (br s, 7.6H), 1.21 (br s, 1.4H); ^{13}C NMR (100MHz, DMSO-d_6) δ = 174.0, 155.4, 136.1, 127.2, 123.7, 120.9, 118.4, 118.2, 111.4, 110.2, 78.0, 54.5, 28.2, 26.8. HRMS calculated for $C_{16}H_{21}N_2O_4$ $[M+H]^+$ 305.1501. Found 305.1503.

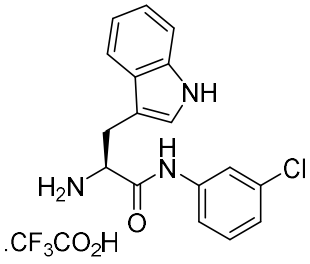


tert-butyl (S)-(1-((3-chlorophenyl)amino)-3-(1H-indol-3-yl)-1-oxopropan-2-yl)carbamate 2-28

A mixture of *N*-Boc tryptophan **2-26** (1.31mmol, 400mg), HATU (1.5eq, 750mg), and DMAP (0.1 eq, 16mg) was dissolved in anhydrous DMF (6mL) and stirred for 30min at room temperature. 3-chloroaniline (1eq, 168mg) was then added, followed by DIPEA (2 eq, 458µL), then the mixture was left to further stir overnight at room temperature.

The mixture was then diluted with 30mL of water, then extracted with EtOAc twice. The organic layers were then washed with H_2O, four times, then brine once, then dried with $MgSO_4$. The solvent was removed in vacuo to obtain the crude residue, which was further purified by column chromatography (eluent 70:30 hexane/EtOAc) to obtain the product as an off-white solid. (418mg, 77%).

Off-white solid, 77% yield. Mp 77-80°C. ^{1}H NMR (100MHz, $CDCl_3$) δ = 8.11 (s, 1H), 7.76 (br s,1H), 7.68(d, J= 7.5 Hz), 7.41-7.39(m, 2H), 7.22 (t, J= 7.5Hz, 1H), 7.17-7.03 (m, 5H), 5.21 (br s, 1H), 4.58 (br s, 1H), 3.41-3.35 (m, 1H), 3.27-3.23 (m, 1H), 1.44 (s, 9H); ^{13}C NMR (100MHz, $CDCl_3$) δ = 170.6, 156.2 (br s), 138.7, 136.4, 134.6, 129.9, 127.4, 124.5, 123.5, 122.6, 120.2, 120.1, 118.9, 118.1, 115.1, 110.5, 80.7 (br s), 56.1 (br s), 28.4. MS-ESI(+): m/z = 414.05 $[^{35}Cl\ M+H]^+$ HRMS calculated for $C_{22}H_{25}N_3O_3Cl$ $[^{35}Cl\ M+H]^+$ 414.1584. Found 414.1572.

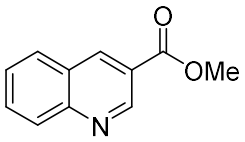


(S)-2-amino-N-(3-chlorophenyl)-3-(1H-indol-3-yl)propenamide trifluoroacetate salt 2-29

Compound **2-28** (200mg, 0.482mmol) was dissolved in DCM (4mL) and cooled to 0°C and neat trifluoroacetic acid (10eq, 550mg, 0.37mL) was then added dropwise into the mixture.

The mixture was then stirred for 2h at room temperature. The volatiles were removed in vacuo, then the residue triturated with hexane. The undissolved solid was then isolated with filtration, further washed with cold hexane, and dried under vacuum to obtain the product as a brown solid. (198mg, 96%)

Brown solid, 96% yield, Mp:131-134°C (with decomposition) ^{1}H NMR (400MHz, methanol-d_4) δ = 7.65-7.62 (m, 2H), 7.38 (d, J= 8.0Hz,1H), 7.31-7.26 (m, 2H), 7.20 (s, 1H), 7.14-7.11 (m, 1H), 7.01 (t, J= 7.5Hz, 1H), 4.21 (t, 7.3 Hz, 1H), 3.49-3.45 (m, 1H), 3.37-3.33 (m, 1H). ^{13}C NMR (100MHz, methanol-d_4) δ = 168.7, 140.3, 138.2, 135.4, 131.1, 128.3, 125.6, 125.6, 122.9, 121.3, 120.3, 119.4, 119.1, 112.5,107.8, 55.8, 28.8. MS-ESI(+): m/z = 314.20 $[M]^+$ HRMS calculated for $C_{17}H_{16}N_3OCl$ $[M]^+$ 314.1060. Found 314.1058.

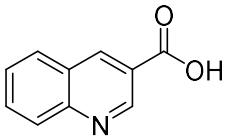


Methyl quinoline-3-carboxylate 2-31

A mixture of 3-bromoquinoline **2-30** (208.1mg, 1.00mmol), Xantphos (0.2 eq, 115.7mg) and palladium (II) acetate (0.1 eq, 22.5mg) in dry MeOH (4 mL) was subjected to a CO atmosphere (1 atm from a balloon), then added Et_3N (1.5eq, 151.8mg, 0.21mL). The reaction mixture is heated to reflux under CO overnight. When TLC indicated completion of reaction, the mixture was cooled to room temperature and filtered through a pad of celite, and the solvent evaporated in vacuo. The residue was then redissolved in Et_2O (10mL) and washed with H_2O twice. The aqueous layer was extracted with Et_2O twice, and the combined organic extracts washed with brine and dried with $MgSO_4$. The solvent was then evaporated in vacuo to obtain the crude product as a yellow solid (104.8mg, 56% crude), which was immediately used for the next reaction without further purification.

Yellow solid. 56% yield (crude). ^{1}H NMR (400MHz, $CDCl_3$) δ = 9.46 (d, J= 2.0Hz, 1H), 8.86 (d, J= 2.0Hz, 1H), 8.17 (d, J= 8.0Hz, 1H), 7.95 (d, J= 8.0Hz), 7.84 (dt, J= 2.0, 7.0Hz, 1H), 7.63

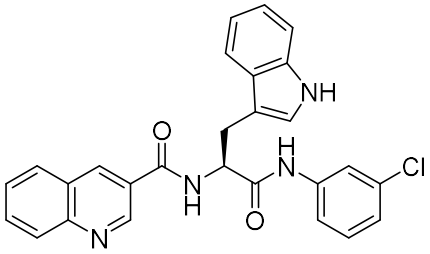
(t, J= 7.0Hz, 1H), 4.03 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 166.0, 150.1, 149.8, 139.0, 129.5, 129.3, 127.6, 127.0, 123.2, 52.6. MS-ESI (+) m/z = 188.03 $[M+H]^+$. HRMS calculated for $C_{11}H_{10}NO_2$ $[M+H]^+$ 188.0712. Found 188.0727.



Quinoline 3-carboxylic acid 2-32

Crude methyl quinoline-3-carboxylate **2-31** (104.8mg, 0.56mmol) was dissolved in a 4:1 THF/H_2O solution and stirred. To it was added $LiOH.3H_2O$ (3eq, 126.1mg), then stirred at room temperature for 1 hour. When TLC indicated completion of reaction, the organic volatiles were removed in vacuo, and to the remaining aqueous solution was neutralised with addition of saturated $NaHSO_4$ solution until the pH was about 3. The mixture was then extracted with EtOAc twice and the combined organic layers dried with $MgSO_4$, filtered, and evaporated in vacuo to obtain the crude residue, which was purified by recrystallization from hot MeOH to obtain the product as a pale yellow solid, (60mg, 63%)

Pale yellow solid, 63%. Mp: 260-263°C (with decomposition) ^{1}H NMR (400MHz, DMSO-d6) δ = 9.31 (s,1H), 8.97 (s, 1H), 8.20 (d, J= 8.4Hz, 1H), 8.10 (d, J= 8.4Hz, 1H), 7.91 (t, J= 8.0Hz, 1H), 7.71 (t, J= 8.0Hz, 1H) ^{13}C NMR (100MHz, DMSO-d6) δ = 166.3, 149.8, 149.1, 138.4, 129.5, 128.7, 127.4, 126.6, 123.7. HRMS calculated for $C_{10}H_6NO_2$ $[M-H]^-$. 172.0399. Found 172.0396.



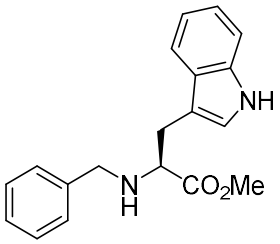
2-10bn

Quinoline-3-carboxylic acid **2-32** (30.0mg, 0.173mmol) was subjected to general procedure 1 with trifluoroacetate salt **2-29** (74.1mg, 0.173mmol).

The crude product was purified with column chromatography (eluent 40:60 hexane/EtOAc) to obtain product as an off -white solid (50mg, 62%)

Off-white solid. 62% yield. Mp 228-232°C ^{1}H NMR (400MHz, acetone-d_6) δ = 10.08 (s, 1H), 9.60 (s, 1H), 9.30 (d, J= 2.3Hz, 1H), 8.75 (d, J= 1.8 Hz, 1H), 8.23 (d, J= 7.8Hz, 1H), 8.09 (d, J= 8.7 Hz, 1H), 8.02 (d, J= 8.2Hz, 1H) 7.89 (t, J= 1.8 Hz, 1H), 7.85 (dt, J= 1.8, 6.9Hz, 1H), 7.74 (d, J= 7.8Hz, 1H), 7.67 (dt, J= 0.9, 6.9Hz, 1H), 7.48 (td, J= 0.9, 6.9Hz, 1H), 7.37 (d, J= 7.8 Hz, 1H), 7.32-7.28 (m, 2H), 7.11-7.07 (m, 2H), 7.01 (dt, J= 0.9, 8.2Hz, 1H), 5.15-5.09(m, 1H), 3.57-3.51 (m, 1H), 3.45-3.39 (m, 1H). ^{13}C NMR (100MHz, acetone-d_6) δ = 171.4, 166.3, 150.1, 149.8, 141.2, 137.6, 136.3, 134.6, 131.8, 130.9, 130.0, 129.8, 128.6, 128.1, 127.6, 124.6, 124.2, 122.2, 120.3, 119.6, 119.4, 118.7, 112.2, 111.1, 56.5, 28.7. MS-ESI(+): m/z = 469.32 $[M+H]^+$ HRMS calculated for $C_{27}H_{21}N_4O_2Cl$ $[M+H]^+$ 469.1431. Found 469.1417.

2.5.1.4 Chemical synthesis of **2-10bo**

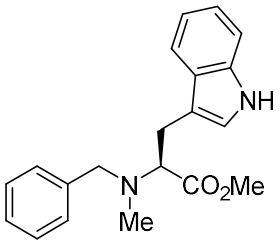


Methyl N$^{\alpha}$-benzyl L-tryptophanate 2-33

Tryptophan methyl ester hydrochloride **2-12** (1500mg, 5.89mmol), was dissolved in a minimal amount of MeOH and neutralized to its freebase amine by stirring with Amberlite IRA400 (OH form) for 15 minutes until the colour of the solution changed from brown to orange. The resin was removed by filtration and the solvent evaporated to obtain the freebase amine. The freebase form was redissolved in dry MeOH (40mL) and to it added benzaldehyde (1.1 eq, 6.48mmol, 687.5mg, 660µL), and $MgSO_4$ as water removing agent (5% mass, 75mg). The mixture was

stirred for 1 hour at room temperature and sodium cyanoborohydride (2eq, 11.78mmol, 740.1mg) was then added followed by overnight stirring. When TLC showed completion of reaction, the solvent was evaporated in vacuo, the residue redissolved in EtOAc (30mL), and then washed with 5% aqueous NH_3 solution (25mL). The aqueous layer was extracted with EtOAc twice and the combined organic extracts washed with brine (30mL) and dried with $MgSO_4$. Filtration then evaporation of the solvent in vacuo afforded the product as a white solid. (1805mg, 95%), which was used without further purification.

White solid, 95% yield. Mp: 85-89°C. ^{1}H NMR (400MHz, $CDCl_3$) δ = 8.01 (s, 1H), 7.58 (d, J= 7.8Hz, 1H), 7.35 (J= 8.1 Hz, 1H), 7.26-7.17 (m, 6H), 7.12 (t, J= 7.4 Hz, 1H), 7.04 (s, 1H), 3.84-3.81 (m, 1H), 3.68-3.64 (m, 1H), 3.63 (s, 1H), 3.23-3.12 (m, 2H); ^{13}C NMR (100MHz, $CDCl_3$) δ = 175.5, 139.9, 136.4, 128.6, 127.7, 127.2, 122.9, 122.3, 119.7, 119.0, 111.7, 111.3, 61.4, 52.3, 51.9, 29.5. ESI-MS(+): m/z = 309.20 $[M+H]^+$. HRMS calculated for $C_{19}H_{21}N_2O_2$ $[M+H]^+$ 309.1603. Found 309.1594.

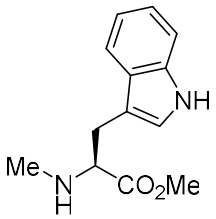


Methyl N$^{\alpha}$-benzyl-N$^{\alpha}$-methyl-L-tryptophanate 2-34

To a solution of compound **2-33** (1500mg, 4.86mmol) in dry MeOH (20mL) was added paraformaldehyde (1.5eq, 7.30mmol, 219mg) and $MgSO_4$ (5% mass, 75mg) as drying agent, then stirred for 1 hour at room temperature. Sodium cyanoborohydride (1.5eq, 7.30mmol, 458.5mg) was then added and the mixture left to stir overnight. When TLC showed completion of reaction, the solvent was evaporated in vacuo, the residue redissolved in EtOAc (30mL), and then washed with 5% aqueous NH_3 solution (25mL). The aqueous layer was extracted with EtOAc twice and the combined organic extracts washed with brine and dried with $MgSO_4$.

Filtration then evaporation of the solvent in vacuo afforded the product as a pale yellow oil. (738mg, 47%), which was used without further purification.

Yellow oil, 47% yield. ^{1}H NMR (400MHz, $CDCl_3$) δ = 7.97 (s, 1H), 7.47 (d, J= 8.2 Hz, 1H), 7.35-7.20 (m, 6H), 7.15 (dt, J= 1.4, 7.4Hz, 1H), 7.05 (dt, J= 0.9, 8.2 Hz, 1H), 6.99 (d, 2.3Hz), 3.84 (d, J= 13.72Hz, 1H), 3.74-3.70 (m, 1H), 3.63 (s, 3H), 3.62 (d, J= 14.7Hz, 1H), 3.35-3.29 (m, 2H), 3.11-3.06 (m, 2H), 2.35 (s, 3H) ^{13}C NMR (400MHz, CDCl3) δ = 172.8, 139.5, 136.3, 129.0, 128.4, 127.7, 127.2, 122.8, 122.1, 119.5, 118.9, 112.6, 111.2, 66.5, 59.0, 51.1, 38.4, 25.9.; ESI-MS(+): m/z = 323.22 $[M+H]^+$. HRMS calculated for $C_{20}H_{23}N_2O_2$ $[M+H]^+$ 323.1760. Found 323.1759.

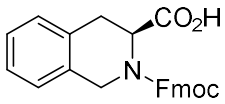


N$^{\alpha}$-methyl methyl-L-tryptophanate 2-35

Compound **2-34** (400mg, 1.22mmol), was dissolved in a 19:1 MeOH/AcOH solvent mixture (10mL) and flushed with inert gas. The solution was transferred to a high-pressure reactor with palladium hydroxide on carbon (10% mass, 40mg) and stirred under a 50psi H_2 atmosphere at room temperature overnight. After TLC showed completion, the solvent was purged again with inert gas, and reaction mixture was filtered through a pad of celite. The solvent was evaporated in vacuo and the residue was then redissolved in EtOAc (10mL), then washed with saturated aqueous $NaHCO_3$.The aqueous layer was extracted with EtOAc 2 times, and the combined organic layers washed with brine and dried with $MgSO_4$. Filtration and evaporation of the solvent in vacuo afforded the product as a yellow oil (285mg, 99%)

Yellow oil, 99% yield, ^{1}H NMR (400MHz, $CDCl_3$) δ = 8.09 (s, 1H), 7.60 (d, J= 7.8Hz, 1H), 7.35 (d, J= 8.2 Hz, 1H), 7.19 (t, J= 8.2Hz, 1H), 7.12 (t, J= 8.2Hz, 1H), 7.07 (d, J= 2.3Hz, 1H), 3.66 (s, 3H), 3.54 (t, J= 6.4Hz, 1H), 3.22-3.09 (m, 2H), 2.37 (s, 3H); ^{13}C NMR (100MHz,

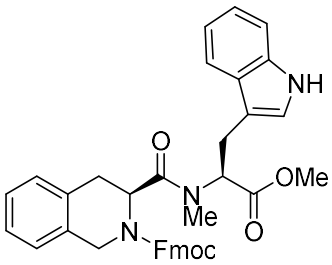
CDCl3) δ = 175.2, 136.4, 128.7, 127.5, 123.1, 119.6, 118.8, 111.4, 63.9, 51.9, 34.9, 29.2. ESI-MS(+): m/z = 233.14. $[M+H]^+$. HRMS calculated for $C_{13}H_{17}N_2O_2$ $[M+H]^+$ 233.1290. Found 233.1285.



(S)-2-(((9H-fluoren-9-yl)methoxy)carbonyl)-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid 2-36

A suspension of **2-14** (200mg, 0.936mmol) in deionised H_2O (10mL) was stirred vigorously and its pH adjusted to about 9 by the addition of saturated $NaHCO_3$ solution. The mixture was then cooled to 0°C and Fmoc-Cl (1.1eq, 266mg) in THF (10mL) was added over 30 minutes using a syringe pump. The mixture was then warmed to room temperature and further stirred for another 3h. When TLC shows completion of reaction, the solution was acidified by addition of 2M HCl and diluted with H_2O (50mL). The mixture was then extracted with EtOAc 3 times, and the combined organic layers washed with brine and dried with $MgSO_4$. The solution was filtered and evaporated in vacuo to afford the product as a white solid. (292mg, 78%)

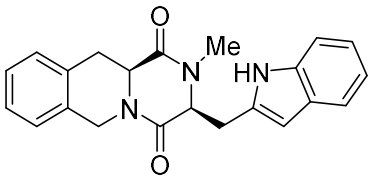
White solid, 78% yield. ^{1}H NMR (400MHz, $CDCl_3$) (mixture of rotamers) δ = 7.79-7.48 (m, 4.3H), 7.40-7.10 (m. 7.7H), 5.12-5.07 (m, 0.45H), 4.72 (app d, 1.25H), 4.60-4.05 (m, 2.3H), 3.28-3.09 (m, 2H). ^{13}C NMR (100MHz, $CDCl_3$) (mixture of rotamers), δ = 176.1, 156.5, 155.6, 144.5, 144.1, 144.0, 141.7, 133.0, 132.5, 131.7, 131.3, 128.7, 128.3, 127.9, 127.8, 127.3, 126.6, 126.4, 125.2 (apparent d), 124.9 (apparent d), 120.2 (apparent t), 68.3, 68.0, 65.4, 53.4, 53.2, 50.5, 47.4, 44.6, 31.5, 31.0. HRMS calculated for $C_{25}H_{20}NO_4$ $[M-H]^-$ 398.1392. Found 398.1397.



(9H-fluoren-9-yl)methyl (S)-3-(((S)-3-(1H-indol-3-yl)-1-methoxy-1-oxopropan-2-yl)(methyl)carbamoyl)-3,4-dihydroisoquinoline-2(1H)-carboxylate 2-37

To a solution of compound Fmoc protected compound **2-36** (200mg, 0.50mmol) in DCM (4mL), was added 1 drop of DMF and then oxalyl chloride (1.1eq, 69.9mg, 50μL) dropwise. The mixture was stirred vigorously until no more effervescence was observed, then the solvent evaporated in vacuo to obtain the crude acyl chloride. The acyl chloride was dried in vacuum. In another reaction flask, N^{α}-methyl methyl-L-tryptophanate **2-35** (1eq, 116.3mg) and 4-methyl-2,6-tertbutylpyridine (2eq, 1.00mmol, 205.6mg) was dissolved in dry DCM (4mL). The acyl chloride was then added to this mixture and stirred for 2h at room temperature. When TLC indicates complete reaction, the mixture was washed with 1M HCl (10mL). The aqueous layer was extracted with DCM twice and the combined organic extracts washed with brine and dried with $MgSO_4$. The mixture was filtered and solvent evaporated in vacuo to obtain the residue, which was purified by column chromatography (eluent 60:40 hexane/EtOAc) to afford the product as an off-white solid (147mg, 48%)

Off-white solid, 48% yield. ^{1}H NMR (400MHz, $CDCl_3$) (mixture of rotamers) δ = 7.79-6.74 (m, 18H), 5.40-5.36 (m, 0.5H), 5.16-5.13(m, 0.5H), 4.89-4.23 (m, 4H), 3.97-3.64 (m,5H), 3.49-2.48 (m, 7H). MS-ESI(+): m/z = 614.23 $[M+H]^+$



2-10bo

A solution of compound **2-37** (145mg, 0.236mmol) in DMF (4mL) was added 20% volume piperidine (0.8mL) and stirred at room temperature for 15 minutes. After TLC showed completion of reaction the solution was then diluted with H_2O (50mL) and extracted with EtOAc 4 times. The combined organic extracts were thoroughly washed with H_2O 4 times, then brine once, and dried with $MgSO_4$. The mixture was filtered, and solvent evaporated in vacuo to obtain the product, which was purified by column chromatography (eluent 50:50 hexane/EtOAc) to obtain the product **2-10bo** as a white solid. (49mg, 58%)

White solid. 58% yield. Mp 205-207°C. ^{1}H NMR (400MHz, $CDCl_3$) δ = 7.98 (s, 1H), 7.43 (d, J= 8.0Hz, 1H), 7.15 (d, J= 8.5 Hz), 7.07 (t, J= 7.5Hz, 1H), 6.99 (t, J= 7.5Hz), 6.95 (d, J= 7.5Hz, 1H), 6.89, (d, J= 2.5Hz, 1H), 6.85 (t, J= 7.5Hz), 6.75 (t, J= 7.5Hz, 1H), 6.62 (d, J= 7.5 Hz, 1H), 5.31 (d, J= 15.0 Hz, 1H), 4.32 (t, J= 4.0Hz, 1H), 3.96 (d, J= 17.0Hz, 1H), 3.91 (dd, J= 4.0,13.0 Hz, 1H), 3.63 (dd, J= 2.5, 15.0 Hz, 1H) 3.32 (dd, J= 4.5, 15.0Hz), 3.15 (s, 3H), 2.62 (dd, J= 4.5, 17.0Hz, 1H), 0.69 (dd, J= 13.0, 17.0Hz); ^{13}C NMR (100MHz, $CDCl_3$) δ = 165.6, 164.2, 135.8, 132.0, 130.4, 128.5, 127.5, 126.4, 126.3, 123.7, 122.5, 119.9, 118.6, 110.9, 108.7, 63.0, 55.6, 43.7, 32.5, 32.4, 27.6. MS-ESI(+): m/z = 360.27 $[M+H]^+$ HRMS calculated for $C_{22}H_{22}N_3O_2$ $[M+H]^+$, 360.1712. Found 360.1697.

2.5.2 Biological Assays

All biological experiments and assays are credited to Ragunathan Priya.

2.5.2.1 Preparation of inverted membrane vesicles from M. smegmatis

About 5 g (wet weight) of *M. smegmatis* were resuspended in 20 ml membrane preparation buffer (50 mM MOPS, 2 mM $MgCl_2$, pH 7.5) containing EDTA-free protease inhibitor cocktail (1 tablet in 20 ml buffer, Roche-USA) and 1.2 mg/ml lysozyme. The suspension was stirred at room temperature for 45 min and additionally supplemented with 300 µl 1 M $MgCl_2$ and 50 µl DNase I (Thermo Fischer, USA), and continued stirring for another 15 min at room temperature. All subsequent steps were performed on ice. Cells were broken by three passages through an

ice precooled Model M-110L Microfluidizer processor (M-110L) at 18,000 psi. The suspension containing lysed cells was centrifuged at 4,200 x g at 4 °C for 20 min. The supernatant containing membrane fraction was further subjected to ultracentrifugation 45,000 x g at 4 °C for 1 h. The supernatant was discarded, and the precipitated membrane fraction was resuspended in membrane preparation buffer containing 15% glycerol, aliquoted, snap frozen and stored at -80 °C. The concentrations of the proteins in the vesicles were determined by the BCA method.

2.5.2.2 ATP Synthesis Assay

ATP synthesis was measured in flat bottom white microtiter 96 well plates (Corning USA). The reaction mixture was made in assay buffer (50 mM MOPS, pH 7.5, 10 mM $MgCl_2$) containing 10 µM ADP, 250 µM Pi and 1 mM NADH. Concentration of Pi was adjusted by adding 100 mM KH_2PO_4 salt dissolved in the assay buffer. ATP synthesis was started by adding inverted vesicles of *M. smegmatis* to a final protein concentration of 5 µg/ml. The reaction mixture was incubated at room temperature for 30 min before addition of 50 µl of the CellTiter-glow reagent, and the mixture was incubated for another 10 min in dark at room temperature. Emitted luminescence, which is correlated to the synthesized ATP, was measured by a Tecan plate reader Infinite 200 Pro (Tecan USA), using the following parameter: luminescence, integration time of 500 ms and no attenuation.

2.5.2.3 Agar and broth MIC

Initial stock solutions of the test compounds were made in 90% DMSO to a concentration of 10 mM. Bedaquiline was used as a positive control and the vehicle DMSO was used as negative control. In the first approach, the compounds were tested on microbial cultures at a fixed concentration of 50 µM. Each of the above strains were cultured at 37 °C in Middlebrook 7H9 broth supplemented with 0.2% glycerol and 10% ADC (Albumin Dextrose Catalase) until logarithmic growth was achieved (OD_{600} 0.4 - 0.6). The test inoculum was obtained by diluting the

suspensions to OD_{600} 0.1 to a final volume of 1 ml in the test tubes and were incubated at 37 °C for 24 hours.

Broth micro dilution method was used to determine the minimum inhibitory concentration (MIC) of the test compounds. Two-fold serial dilutions of the compounds were made from 200 µM to 0.05 µM. Each concentration was assayed in triplicates. The diluted test inoculum was added to all the wells (final volume 200 µl) in the microplate and was incubated at 37 °C for 24 hours (*M. smeg.*) and 5 days (BCG). The final OD of the cultures in the plate was measured by Tecan Infinite 200 PRO plate reader. The MIC_{50} was defined as the drug concentration that inhibited 50% of the bacterial growth when compared to the growth in the drug-free medium.

2.6 References

1. *Global Tuberculosis Report 2022*. World Health Organization: Geneva, 2022.

2. WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int/ (accessed 5 Jun 2023).

3. Song, W. M.; Zhao, J. Y.; Zhang, Q. Y.; Liu, S. Q.; Zhu, X. H.; An, Q. Q.; Xu, T. T.; Li, S. J.; Liu, J. Y.; Tao, N. N.; Liu, Y.; Li, Y. F.; Li, H. C., *Front Med (Lausanne)* **2021,** *8*, 657006.

4. Barry, C. E., 3rd; Lee, R. E.; Mdluli, K.; Sampson, A. E.; Schroeder, B. G.; Slayden, R. A.; Yuan, Y., *Prog Lipid Res* **1998,** *37* (2-3), 143-179.

5. Barry, C., *Trends in Microbiology* **2001,** *9* (5), 237-241.

6. Vincent, A. T.; Nyongesa, S.; Morneau, I.; Reed, M. B.; Tocheva, E. I.; Veyrier, F. J., *Front Microbiol* **2018,** *9*, 2341.

7. *Treatment of Tuberculosis Guidelines*. 4 ed.; World Health Organization: Geneva, 2010.

8. Multidrug-Resistant Tuberculosis (MDR TB) Fact Sheet. https://www.cdc.gov/tb/publications/factsheets/drtb/mdrtb.htm (accessed 6 Jun 2023).

9. *WHO Treatment Guidelines for Drug Resistant Tuberculosis*. World Health Organization: Geneva, 2016.

10. FDA NEWS RELEASE. https://web.archive.org/web/20161219184517/https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm333695.htm (accessed 6 Jun 2023).

11. European Medicines Agency recommends two new treatment options for tuberculosis. https://www.ema.europa.eu/en/news/european-medicines-agency-recommends-two-new-treatment-options-tuberculosis (accessed 6 Jun 2023).

12. Andries, K.; Verhasselt, P.; Guillemont, J.; Gohlmann, H. W.; Neefs, J. M.; Winkler, H.; Van Gestel, J.; Timmerman, P.; Zhu, M.; Lee, E.; Williams, P.; de Chaffoy, D.; Huitric, E.; Hoffner, S.; Cambau, E.; Truffot-Pernot, C.; Lounis, N.; Jarlier, V., *Science* **2005,** *307* (5707), 223-227.

13. Worley, M. V.; Estrada, S. J., *Pharmacotherapy* **2014,** *34* (11), 1187-1197.

14. Sarathy, J. P.; Grüber, G.; Dick, T., *Antibiotics (Basel)* **2019,** *8* (4), 261.

15. Kundu, S.; Biukovic, G.; Grüber, G.; Dick, T., *Antimicrob Agents Chemother* **2016,** *60* (11), 6977-6979.

16. Toei, M.; Noji, H., *J Biol Chem* **2013,** *288* (36), 25717-25726.

17. Hards, K.; McMillan, D. G. G.; Schurig-Briccio, L. A.; Gennis, R. B.; Lill, H.; Bald, D.; Cook, G. M., *Proc Natl Acad Sci U S A* **2018,** *115* (28), 7326-7331.

18. Shang, S.; Shanley, C. A.; Caraway, M. L.; Orme, E. A.; Henao-Tamayo, M.; Hascall-Dove, L.; Ackart, D.; Lenaerts, A. J.; Basaraba, R. J.; Orme, I. M.; Ordway, D. J., *Antimicrob Agents Chemother* **2011,** *55* (1), 124-131.

19. *Anti-Infective Drugs Advisory Committee Meeting Briefing Document*

TMC207 (bedaquiline) Treatment of Patients with MDR-TB; Janssen Pharmaceurical Company: 2012.

20. Diacon, A. H.; Pym, A.; Grobusch, M.; Patientia, R.; Rustomjee, R.; Page-Shipp, L.; Pistorius, C.; Krause, R.; Bogoshi, M.; Churchyard, G.; Venter, A.; Allen, J.; Palomino, J. C.; De Marez, T.; van Heeswijk, R. P.; Lounis, N.; Meyvisch, P.; Verbeeck, J.; Parys, W.; de Beule, K.; Andries, K.; Mc Neeley, D. F., *N Engl J Med* **2009,** *360* (23), 2397-2405.

21. Diacon, A. H.; Pym, A.; Grobusch, M. P.; de los Rios, J. M.; Gotuzzo, E.; Vasilyeva, I.; Leimane, V.; Andries, K.; Bakare, N.; De Marez, T.; Haxaire-Theeuwes, M.; Lounis, N.; Meyvisch, P.; De Paepe, E.; van Heeswijk, R. P.; Dannemann, B.; Group, T. C. S., *N Engl J Med* **2014,** *371* (8), 723-732.

22. Drug Induced QT Prolongation. https://www.uspharmacist.com/article/drug-induced-qt-prolongation (accessed 8 Jun 2023).

23. Joon, S.; Ragunathan, P.; Sundararaman, L.; Nartey, W.; Kundu, S.; Manimekalai, M. S. S.; Bogdanovic, N.; Dick, T.; Grüber, G., *FEBS J* **2018,** *285* (6), 1111-1128.

24. Biukovic, G.; Basak, S.; Manimekalai, M. S.; Rishikesan, S.; Roessle, M.; Dick, T.; Rao, S. P.; Hunke, C.; Gruber, G., *Antimicrob Agents Chemother* **2013,** *57* (1), 168-176.

25. van Woerden, H. F., *Chemical Reviews* **2002,** *63* (6), 557-571.

26. Jackson, A. H.; Smith, A. E., *Tetrahedron* **1968,** *24* (1), 403-413.

27. Yokoyama, A.; Ohwada, T.; Shudo, K., *The Journal of Organic Chemistry* **1998,** *64* (2), 611-617.

28. Konig, W.; Geiger, R., *Chem Ber* **1970,** *103* (3), 788-798.

29. Rehberg, N.; Omeje, E.; Ebada, S. S.; van Geelen, L.; Liu, Z.; Sureechatchayan, P.; Kassack, M. U.; Ioerger, T. R.; Proksch, P.; Kalscheuer, R., *Antimicrob Agents Chemother* **2019,** *63* (9), e00136-19.

30. Neises, B.; Steglich, W., *Angewandte Chemie International Edition in English* **1978,** *17* (7), 522-524.

31. Schoenberg, A.; Bartoletti, I.; Heck, R. F., *The Journal of Organic Chemistry* **1974,** *39* (23), 3318-3326.

32. Di Gioia, M. L.; Leggio, A.; Malagrino, F.; Romio, E.; Siciliano, C.; Liguori, A., *Mini Rev Med Chem* **2016,** *16* (9), 683-690.

33. Eschweiler, W., *Chemiste Berichte* **1905,** *38*, 880-882.

34. White, K. N.; Konopelski, J. P., *Org Lett* **2005,** *7* (19), 4111-4112.

35. Thern, B.; Rudolph, J.; Jung, G., *Tetrahedron Letters* **2002,** *43* (28), 5013-5016.

36. Pedroso, E.; Grandas, A.; de las Heras, X.; Eritja, R.; Giralt, E., *Tetrahedron Letters* **1986,** *27* (6), 743-746.

37. Borthwick, A. D., *Chem Rev* **2012,** *112* (7), 3641-3716.

CHAPTER 3

Resolving the problem of lack of inhibitory activity of EpMF2

3. Resolving the problem of lack of inhibitory activity of EpMF2

3.1 Abstract

EpMF2 analogues were synthesized and shown to be effective inhibitors of the F_1F_O mycobacterial ATP synthase, but they were ineffective against living *M. smegmatis*. We hypothesized that the reason for this is that EpMF2 molecule was unable to bypass the thick, hydrophobic cell wall of the mycobacteria where it can bind to its target inside the cytosol. To circumvent this, a prodrug strategy was formulated where the EpMF2 would be conjugated to a cholesterol molecule. *Mtb* is known to have a high affinity of cholesterol and tends to incorporate it into itself. The concept behind the prodrug idea is that the bacterium would be "tricked" into bringing the EpMF2 molecule into its cytosol as it ingests cholesterol. Enzymes in the bacteria would then cleave the linker and release EpMF2 inside the cytosol, thus bypassing the cell wall. Synthesis and viability assays of EpMF2-cholesterol prodrugs showed mixed results, with some analogues successfully inhibiting growth of living *M. smegmatis,* but not others. Still, it is a proof of concept that demonstrated the viability in the use of cholesterol to bypass the defences of *Mtb* against many drugs. A second prodrug proposal, where EpMF2 was linked to an antimicrobial polymer, is at a preliminary stage, with chemical syntheses of the polymer-EpMF2 conjugate underway.

3.2 Introduction

3.2.1 Good target inhibition does not mean effective growth inhibition.

Many analogues of EpMF2 have been synthesized and tested on *M. smegmatis* IMVs to determine its inhibitory activity on the ATP synthase. While some analogues have good inhibitory ability on *M. smegmatis* ATP synthase, none of them had any inhibitory activity on live *M. smegmatis*. *M. smegmatis* (and *Mtb*) is an aerobic species and the functioning of its ATP synthase was shown to be necessary for cell growth.[1] If the growth of living *M. smegmatis* was not inhibited by the addition of EpMF2 analogues, it follows that its ATP synthase was

not inhibited and the compound is not present at the target binding site. In live bacteria, the ε subunit is in the inner membrane facing the cytosol, and any introduced EpMF2 molecules must bypass the cell wall to reach its target. Based on these facts, there are two proposed hypotheses to explain why EpMF2 was not inhibiting the growth of living mycobacteria:

- The compound is not penetrating the cell wall of live mycobacteria in the first place. (Figure 3.1 middle)
- The compound did penetrate the cell wall, but it was quickly pumped out through drug efflux. (Figure 3.1 right)

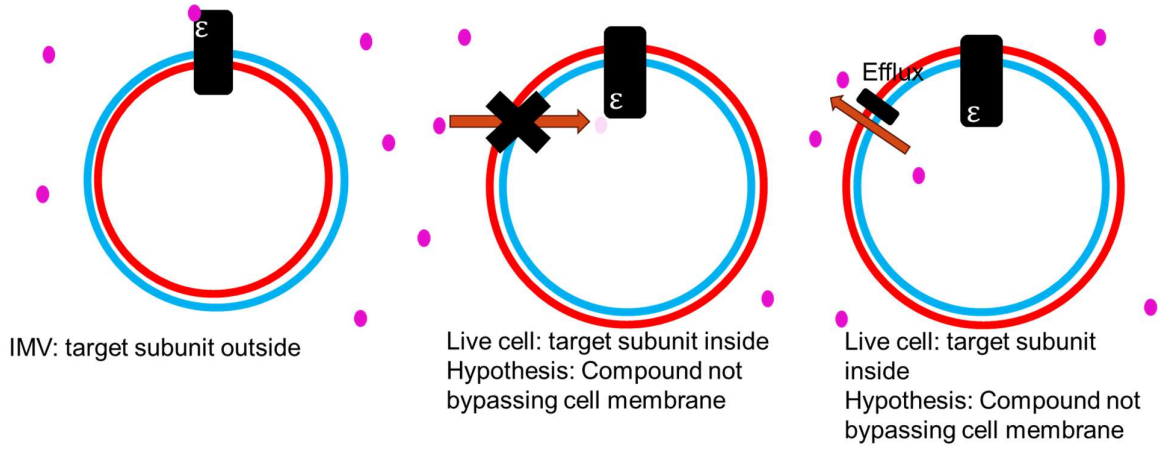


Figure 3.1: The ε subunit is in the interior of the membrane bilayer in mycobacteria. If EpMF2 is not inhibiting live mycobacterial growth, a hypothesis is that EpMF2 is not entering the cell interior to bind to ε (middle), or that it is entering the cell but was removed by pump efflux (right).

3.2.2 Lack of activity is not caused by drug efflux.

The second hypothesis can be tested with verapamil. If a drug is being expelled through efflux pumps, the addition of verapamil, which is an efflux inhibitor, will retain the drug in the cell, therefore reducing its MIC. Verapamil has been shown to enhance the bactericidal effect of BDQ, where the addition of 50μg/mL of the drug reduced the MIC activity of BDQ on *Mtb* by 8 to 16 times.[2]

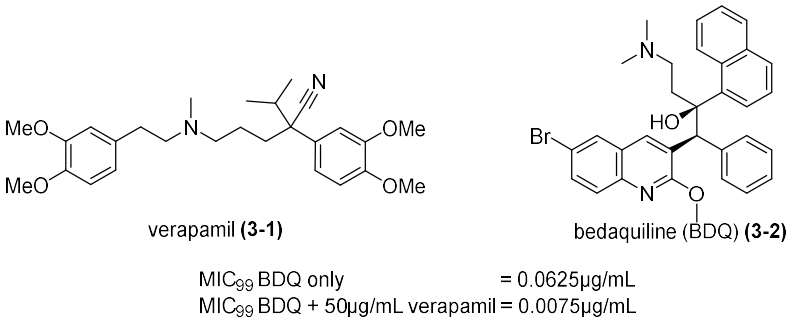


Figure 3.2: Molecular structure of verapamil and BDQ. The addition of verapamil increases Mtb's vulnerability to BDQ by reducing MIC values by 8.

In the control reaction with BDQ using the broth dilution method, there was a 3-fold decrease in MIC_{50} activity on *M. smegmatis*. when sub-inhibitory concentrations of verapamil (8 µM) were added along with BDQ. However, a similar test with verapamil (8 µM) and EpMF2 as well as related analogue EpMF3 (both synthesized by Dr. Patcharaporn) did not lead to any measurable inhibition of growth (as shown by the mostly flat lines in Figure 3.3), whether verapamil was added or not. As the addition of an efflux pump inhibitor did not cause EpMF2 to gain any inhibitory effect, this shows that drug efflux is not the main reason for the inability of EpMF2 to suppress *M. smegmatis* growth.

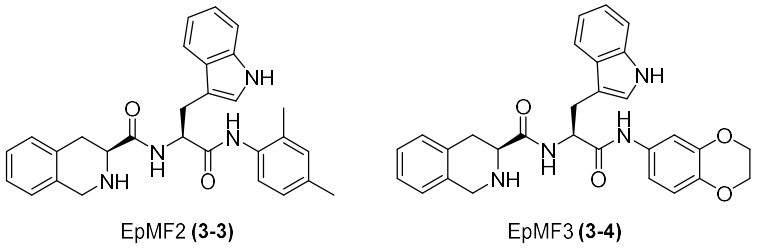


Figure 3.3: Chemical structures of compounds EpMF2 **(3-3**) and EpMF3 **(3-4)**

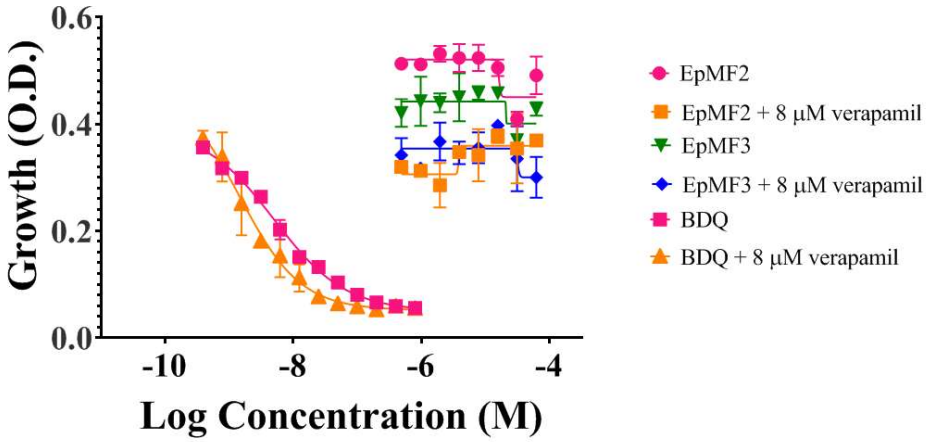


Figure 3.4: The addition of verapamil did not change the lack of inhibitory activity of various EpMF1 analogues on M. smegmatis. Level of inhibition was measured by any changes in the absorbance at wavelength 660nm, which is proportional to concentration of bacteria. DF was a control experiment without any added drug.

3.2.3 Compounds that penetrate the bacterial cell wall.

Thus, we went back to investigate the first hypothesis. To understand how compounds were able to penetrate the bacterial cell membrane into the cytosol, some understanding of its composition is needed. The bacterial membrane typically contains high concentrations of negatively charged phospholipids, particularly phosphatidylglycerol **(3-5)**, cardiolipin **(3-6)**, and other anionic phosphate groups. In contrast, the membrane of mammalian cells tend to have predominantly neutral zwitterionic lipids, such as phosphatidylethanolamine **(3-7)** and phosphatidylcholine **(3-8)**.[3]

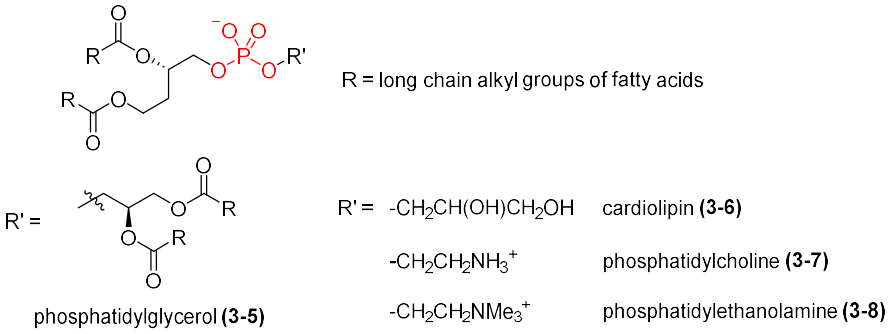


Figure 3.5: The bacterial membrane contains a high concentration of negatively charged lipids such as phosphatidylglycerol and cardiolipin., whose structure are as shown on the top row. Human cells on the other hand tend to have more neutral, zwitterionic lipids such as phosphatidylcholine and phospatidylethanolamine (bottom row).

A class of compounds that are effective in penetrating and killing bacterial cell membrane are known as antimicrobial peptides (AMPs). In nature, AMPs were discovered in both

prokaryotes and eukaryotes as a natural defence mechanism against foreign pathogens.[4] AMPs have diverse structures, but they generally have both hydrophobic moieties and hydrophilic moieties in their structure, making them amphiphilic. The non-polar hydrophobic component aids insertion of the AMP into the non-polar lipid bilayer, and the charged hydrophilic moieties, most carrying cationic functional groups, increases the binding of the compound onto the lipid through attractive electrostatic forces between it and the negatively charged phosphates in the phospholipids. Most common bacterial pathogenic species are vulnerable to the killing effects of these amphiphiles, including gram-positive species like *S. aureus*, and gram-negative species such as *E. coli*. [5] AMPs also bind weakly to lipids containing predominantly zwitterionic lipids found on human cells, therefore making these compounds selective against bacterial species over human cells, a factor important for its potential as antimicrobials.[6] Amphiphilic compounds need not be limited to natural peptides, as recent examples have also expanded to include synthetic peptides,[3] polymers[7],and amphiphilic small molecules.[8]

Among amphiphilic molecules examples have been discovered to be bactericidal against *Mtb*, as are shown in figure 3.6 below. Several classes of small molecules have been found to act as antitubercular drugs, and all have an amphiphilic structure. Dick *et al*, synthesized a series of amphiphilic *N*-alkyl indoyl compounds **3-9**[9] with MIC_{50} values as low as 2μM against *Mtb*. The lipophilic portion came from the long chain alkyl group extending from the nitrogen atom of indole, and the hydrophilic cationic moiety from the nitrogen-containing 7-member ring azepane. As seen in **3-9** and **3-10**, many basic amines can be regarded as cationic assuming they were protonated *in vivo*. SQ109 **(3-10)** is an antitubercular drug currently under development. Its pharmacophore was derived from the 1,2-ethylenediamine structure of ethambutol.[10] It contains a geranyl and an adamantyl lipophilic moiety, with the ethylenediamine moiety as the hydrophilic cationic portion. The compound inhibits *Mtb* with

a MIC_{50} value of 1.56 µM. The compound has been tested in phase I and II clinical trials and appeared to be well tolerated.[11, 12]

However, the positive charge need not come from an amine, as in Cook's compounds[13] **(3-11)**, a phenothiazine derivative where the non-polar portion was derived from the two phenyl rings and the positive charge from a tertiary phosphonium moiety. Phenothiazines have been discovered to be bactericidal against drug-sensitive and MDR TB strains.[14] The addition of the phosphonium moiety served to significantly decrease its inhibitory concentration on *Mtb* from >64µg/mL in the parent phenothiazine to 1 µg/mL in **3-11**. Notably, BDQ **(3-2)** itself is also amphiphilic; it has a polar dimethylamino moiety that is protonated under physiological conditions, and non-polar phenyl and naphthyl moieties. This property might be one of the contributory factors to its activity towards *Mtb* as its cationic and lipophilic nature allows it to translocate H^+ ions across the cell membrane and dissipate the proton motive force.[15]

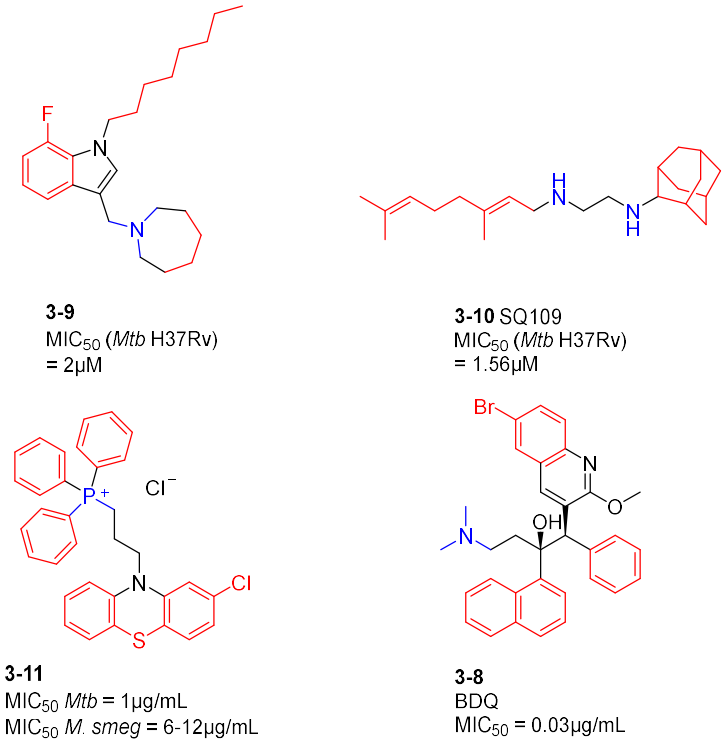
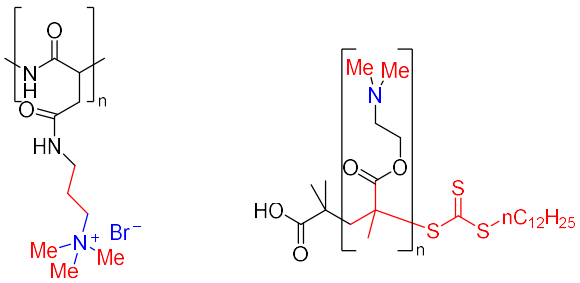


Figure 3.6: Examples of compounds that are capable of existing in a positively charged form and were shown to be bactericidal against Mtb. Top Left: Dick et al's [9] indoyl amphiphilic small molecule **(3-9)**. Top Right: Protopova's SQ109, **(3-10)** a geranyl amphiphilic small molecule.[10]. Bottom left: Cook's triphenylphosphonium compound[13] **(3-11)** Bottom right: BDQ **(3-2)**, which contains a polar dimethylamino moiety and non-polar phenyl and naphthyl rings. Blue areas represent polarpositively charged moieties under physiological conditions. Red areas represent non-polar moieties.

Amphiphilic compounds also include polymers in addition to small molecules. These peptides were designed to mimic AMPs in terms of bactericidal activity. This has been the subject of several reviews.[16, 17]

One such polymer that was demonstrated to be effective against *Mtb* was synthesized by Prof Aasheesh Srivastava's group at IISER Bhopal.[18] It was a polysuccinimide derived polyamide with a tertiary ammonium moiety at each end of the side chain. This moiety gave the polymer its amphiphilic nature derived from its combination of positive charge and non-polar alkyl substituents. It was highly effective against Mtb, with analogue **3-12** (Figure 3.7) completely inhibiting growth at a concentration of 40µg/mL. Another polymer, from Gibson's group, was a poly(dimethylaminoethyl methacrylate) polymer **3-13**.The tertiary dimethyl amino group contributed to its amphiphilicity by utilising the cationic nature of the amine group tempered by the attachment of 2 non-polar methyl groups. The polymer completely supressed growth (MIC_{99}) of *Mtb* at 31.25 µg/mL concentration.



3-12
MIC_{99} (*Mtb* H37Rv)
= 40µg/mL

3-13
MIC_{99} (*M. smegmatis*)
= 31.25µg/mL

Figure 3.7: A few examples of synthetic antitubercular polymers. Left: Asheesh's polypeptide.[18]**(3-12)** Right. Gibson's poly(dimethylaminoethyl methacrylate) polymer.[19]**(3-13).** Blue areas indicate polar cationic moieties, red areas indicate hydrophobic moieties.

For polymers, a net cationic charge is essential for antimicrobial activity, but the hydrophilicity of the polar moieties must be balanced with its hydrophobicity of its alkyl moieties.[3] For example, Srivastava's polypeptides were found to be less active when one of the methyl groups on the tertiary ammonium group was replaced with a longer alkyl chain and made more hydrophobic. On the other hand, Gibson's polymer became less selective on Mtb over human

cells and its MIC_{50} value increased when the dimethylamino group was replaced with an amine group and made more hydrophilic.

However, EpMF2 analogues such as **3-3a-d** do have basic amine moieties yet still have no MIC_{50} activity, thus the lack of amphiphilic nature of EpMF2 is not the sole reason for its lack of activity on living bacteria either. Perhaps EpMF2 lacked large enough hydrophobic components. It might be possible to make the compound more hydrophobic by attaching a long alkyl chain, but if the alkyl chain is too large the steric hindrance may decrease its binding ability to the ε subunit and also creates solubility problems under aqueous physiological conditions.

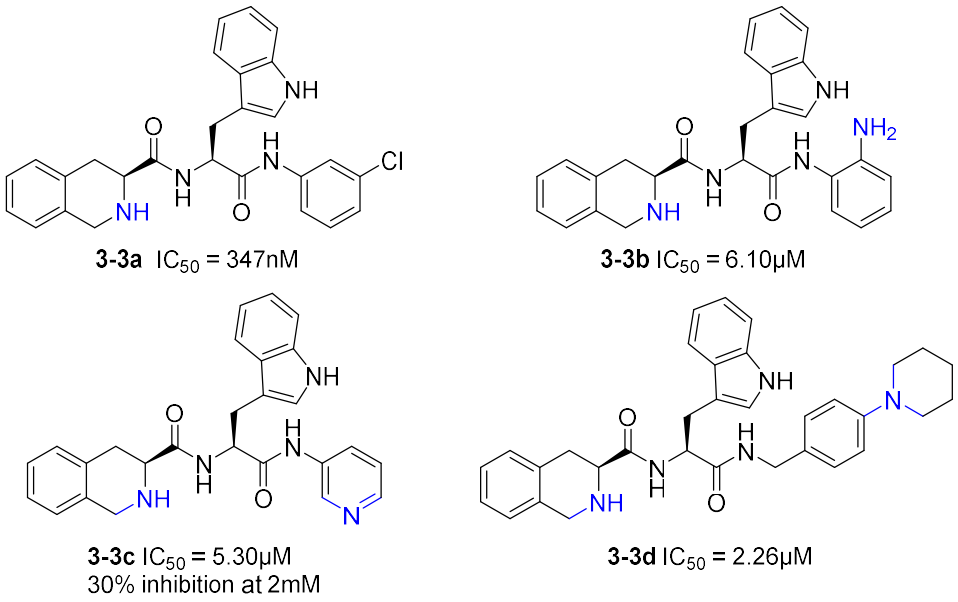


Figure 3.8: However, EpMF2 analogues with amine moieties are not necessarily active. The sole EpMF2 analogue with inhibitory activity **3-3c** has only 30% inhibition at a very high concentration of 2mM. Therefore, the membrane penetrating activity of a compound is not just a matter of having cationic moieties.

3.2.4 *Mtb* and cholesterol intake

Another method to bypass the cell wall, unique to mycobacteria, may be conjugation to cholesterol. *Mtb* contains an unusually large number of genes involved with cholesterol metabolism and intake[20, 21], despite not being able to synthesize cholesterol by itself. In addition, *Mtb* has a high affinity for cholesterol molecules, and it appeared that the presence of cholesterol on the macrophage lipid membrane facilitated the entry and infection of *Mtb* of

immune cells.[22] When the *Mtb* bacterium was phagocytized by a macrophage, it recruits a protein called TACO (tryptophan-aspartate-containing coat protein) onto the phagosome. TACO has a high affinity to cholesterol. The presence of TACO on the phagosome membrane prevents its acidification and fusion to the lysosome.[23] Far from killing it, *Mtb* turns the tables on the immune response by converting the phagosome into a shelter. In the later stages, a mass of immune cells will surround infected cells to form a granuloma, and some of the infected macrophages become foamy due to accumulation of lipids around the bacteria. The bacteria in foamy macrophages became dormant but continue to survive. In these lipid bodies, Mtb can incorporate cholesterol from the host and use it as its carbon and energy source.[24, 25] The main enzyme responsible for cholesterol intake was determined to be Mce4, a multi-subunit transporter protein that imports cholesterol molecules across its membrane.[26] A secondary membrane bound protein, LucA (lipid uptake coordinator A), also assists in cholesterol intake and metabolism.[27] Cholesterol thus may be a significant factor for the bacteria's virulence as it contributes to its ability to survive and persist even after phagocytosis by macrophages and by hiding in the macrophage, prevents the immune system from eliminating it completely. [26, 28]

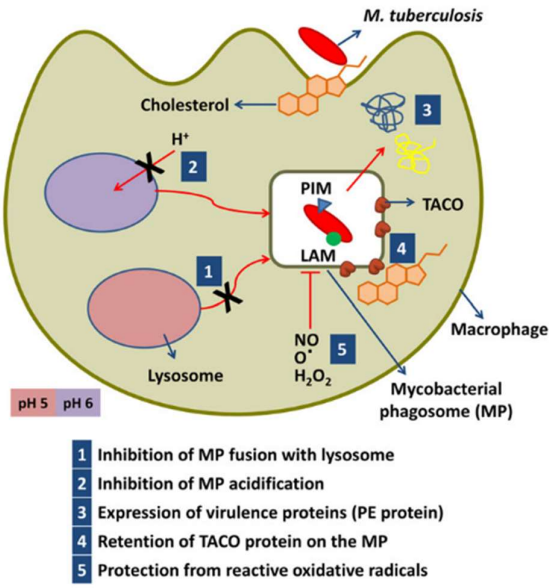


Figure 3.9: How Mtb uses cholesterol to survive phagocytosis by macrophages. Key to this is cholesterol, which aids in entry and the suppression of phagosome maturation, allowing the bacteria to survive inside. Adapted from Abuhammad's paper[24]

3.2.5 Two strategies of mycobacterial cell wall bypass

Based on the above studies on bacterial cell wall penetration came an idea on how to possibly bring the EpMF2 analogue through the mycobacterial membrane. Since EpMF2 alone cannot enter the cell, it could be conjugated with a carrier that can enter the cell, and once in the cytosol, the linker between EpMF2 and the carrier will be cleaved to release EpMF2 *in vivo*. This effectively means creating a prodrug from EpMF2. The combined EpMF2-carrier compound itself should not have any binding affinity to the ε subunit due to significant steric hindrance brought by the carrier, but the carrier helps in terms of pharmacokinetics in bringing the EpMF2 into the cytosol.

Two possible types of carriers have been conceived based on the above studies: one uses cholesterol as the carrier, and one uses an antimicrobial polymer as the carrier. The main point of discussion would be on the former cholesterol strategy, but the second polymer strategy will be briefly touched upon as well.

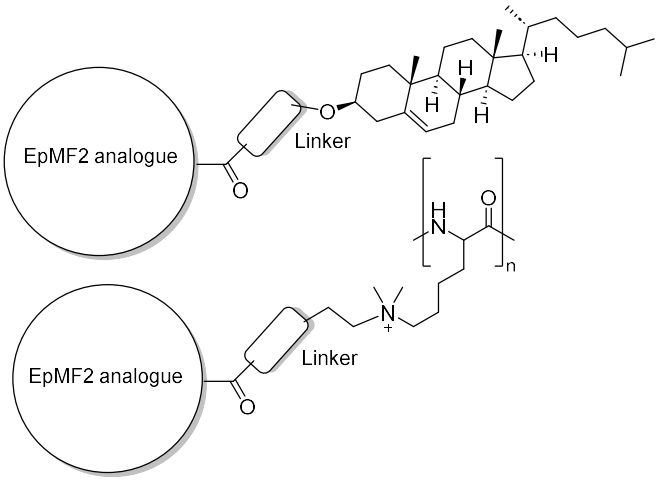


Figure 3.10: Proposed structure of two prodrug ideas of EpMF2. One is an EpMF2-cholesterol conjugate prodrug. The other is an EpMF2-antimicrobial polymer conjugate

The use of cholesterol to aid in drug delivery has some precedents. Cholesterol has been experimented as a component in drug encapsulating micelles and polysomes and covalently linked to peptides and nucleosides to potentially treat various diseases, including viral and bacterial infections as well as cancers. This has been covered in a review by Aderibigbe *et al*.[29] However, to the best of our knowledge, there has not been any study of a drug making use of cholesterol as a carrier to treat specifically tuberculosis by taking advantage of *Mtb*'s propensity to accumulate this lipid. The cholesterol-EpMF2 strategy might thus be a novel use of cholesterol as a carrier and prodrug for anti TB drugs.

3.3 Results and Discussion.

3.3.1 Selecting EpMF2 analogues for the cholesterol prodrug synthesis

The cholesterol conjugate idea was first adopted due to the relatively simple small molecular structure of cholesterol and its commercial availability. A decision needed to be made on choosing which EpMF2 analogues were to be used, and how the chemical structure of the linker

be like, and what parts of EpMF2 molecule and cholesterol molecule can the linker be attached to. Eventually, EpMF2 analogues **3-3e** and **3-3f** were chosen.

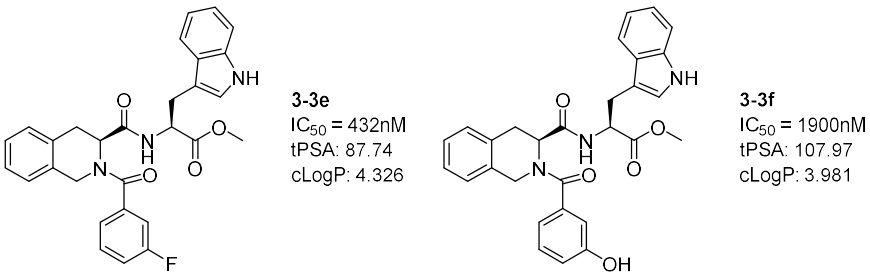


Figure 3.11: Chemical structure of the analogue **3-3e** and **3-3f**, analogues chosen for the prodrug strategy. The IC50 value against M. smegmatis IMVs and its tPSA and cLogP values were included.

One major reason for the choice of the analogues above were its high activity. **3-3e** had one of the greatest inhibitory activities on the *M. smegmatis* ε subunit at IC_{50} = 432nM among Series 1 to 3 analogues. **3-3f** was chosen as a structurally similar analogue. Although its activity was lower at IC_{50} = 1.9μM, it was the most active of the hydroxylated analogues. Having a hydroxy group would be ideal to simplify the attachment of the linker as will be elaborated below. Both compounds exhibit the appropriate polarity (the cLogP and tPSA values being in the appropriate range) appropriate for a drug-like structure. Another reason for the choice of analogues is that the synthetic route to **3-3e** and **3-3f** were also already established previously, so scale-up to multigram quantities would be simple, and that would provide ample starting material for the further synthesis of the prodrugs.

The location of attachment of the linker to compounds **3-3e** and **3-3f**, and the type of moiety to coupler the linker and drug required some consideration. The nature of the functional group between the linker to the EpMF2 molecule depends on the atom where the linker is attached to. This functional group should be stable enough during delivery but would be easily cleaved once inside the bacterium. Attachment of the linker to a nitrogen atom could be done through an amide or carbamate, but they are generally highly resistant to hydrolysis, (as shown with the case from bambuterol in the previous chapter)[30] unless specific amidases were present. Attachment of the linker to an oxygen atom through an ester or carbonate is more suitable as

these functional groups are still stable but easily hydrolysed by *Mtb* enzymes. *Mtb* contains multiple esterases and lipases important for its metabolism from cholesterol and other lipids taken from the host during its dormant stage in foamy macrophages. [31] For **3-3f**, there is a hydroxyl group to attach an ester linker to, on the other hand for **3-3e** the only feasible ester attachment would be to convert the methyl ester to the carboxylic acid and attach the linker through there. This was not ideal as the carboxylic acid analogue has a reduced activity of IC_{50} = 25μM. (Chapter 2).

In the end, it was decided to attach the linker to **3-3e** through its indole >NH to form an indole amide. The attachment of a non-polar group through the nitrogen atom of indole was inspired by Dick *et al*'s alkylated indole compound **3-9** which were active inhibitors of *Mtb*. The synthesis of amides using indoles presented some challenges, as the nitrogen atom in indoles are non-nucleophilic owing to delocalisation of the lone pairs of electrons from the nitrogen to the aromatic rings. Common peptide coupling agents such as EDCI or DCC would not work for indoles, and an alternate synthetic method needed to be found.[32] The chemical property of the linker also needed consideration. 2 types of linkers were proposed, one with a non-polar hydrocarbon chain and one polar linker containing ether groups within the hydrocarbon chain. For **3-3f**, the attachment point would be through the formation of the ester from the phenol moiety. Unlike the formation of the indoyl amide, the synthesis of this linker was much more straightforward. As in **3-3e**, a variety of linkers would be synthesized and tested.

3.3.2 Proposed structure of indolyl amide EpMF2-cholesterol prodrug conjugates and its retrosynthesis

Two EpMF2-cholesterol conjugates were proposed and synthesized from **3-3f** to test our hypothesis, The linker strcture of **3-14** is relatively non-polar hydrocarbon chain with methylene groups, while **3-15** has a more polar linker with ether oxygens within the chain. The incorporation of ether moieties in **3-15** was inspired by the related process of PEGlyation,

where drugs were encapsulated or conjugated with poly(ethylene glycol) to render it more soluble and imporve its pharmacokinetics.[33] The two types of linkers would be tested to determine which linker would provide a better MIC_{50} activity.

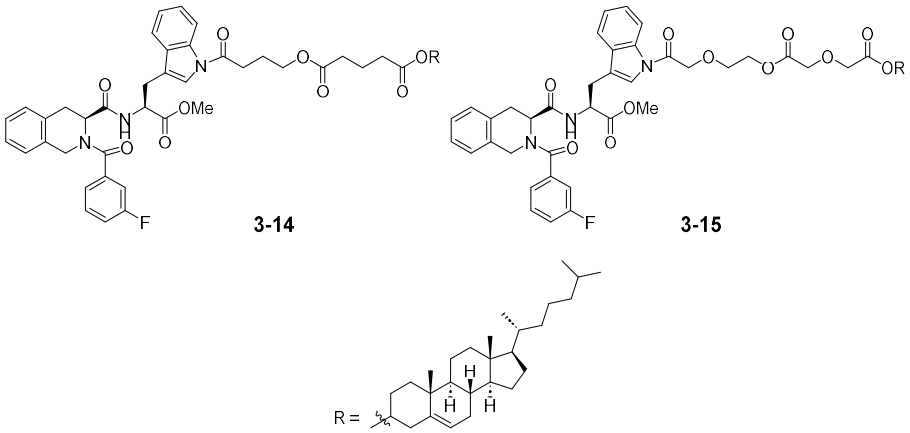
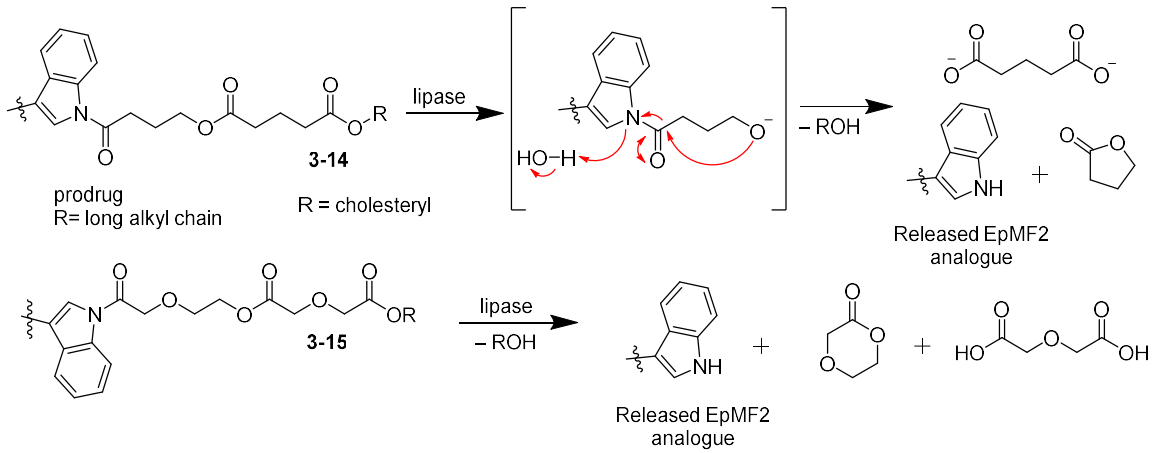


Figure 3.12:These two EpMF2-cholesterol conjugates with indole amide-based linker attachments were synthesized.

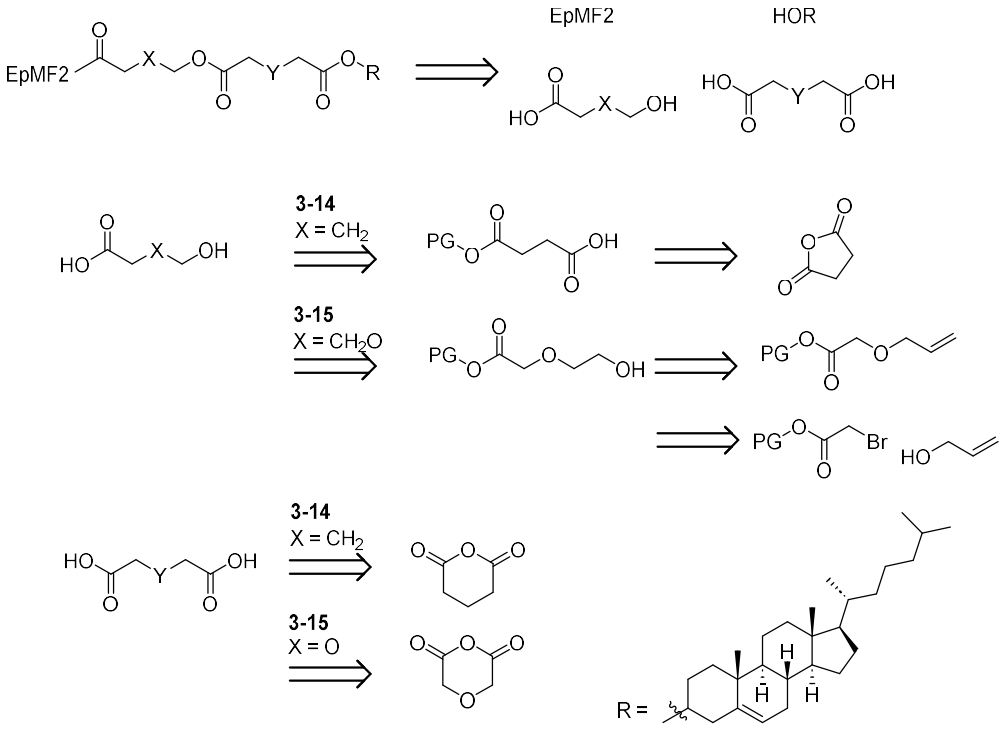
The central ester moiety of the linkers in **3-14** and **3-15** served to activate the hydrolysis of the indoyl amide (otherwise resistant to hydrolysis) through a non-enzymatic mechanism as described in Scheme 3.13: the ester moieties would be hydrolysed by *Mtb* lipases and unmask a hydroxyl end side chain from the indole amide. The hydroxyl group will then spontaneously undergo an intermolecular nucleophilic substituion to expel a cyclic lactone and release the free EpMF2 compound.



Scheme 3.13: Proposed structure of the linker attachment through the indolyl moiety and its hydrolysis mechanism to release the drug after hydrolysis by lipases

The retrosynthesis of **3-14** and **3-15** both depend on piecing the fragments of the linker together, along with cholesterol and EpMF2 analogue **3-3e**. By breaking apart the ester bonds, the components of the linker **3-14** were to be synthesized from 4-hydroxybutyric acid, and glutaric acid. 4-Hydroxybutyric acid can then be derived from partial reduction of the monoester of succinic acid, which in turn would be derived from glutaric anhydride. Glutaric acid would be derived from glutaric anhydride.

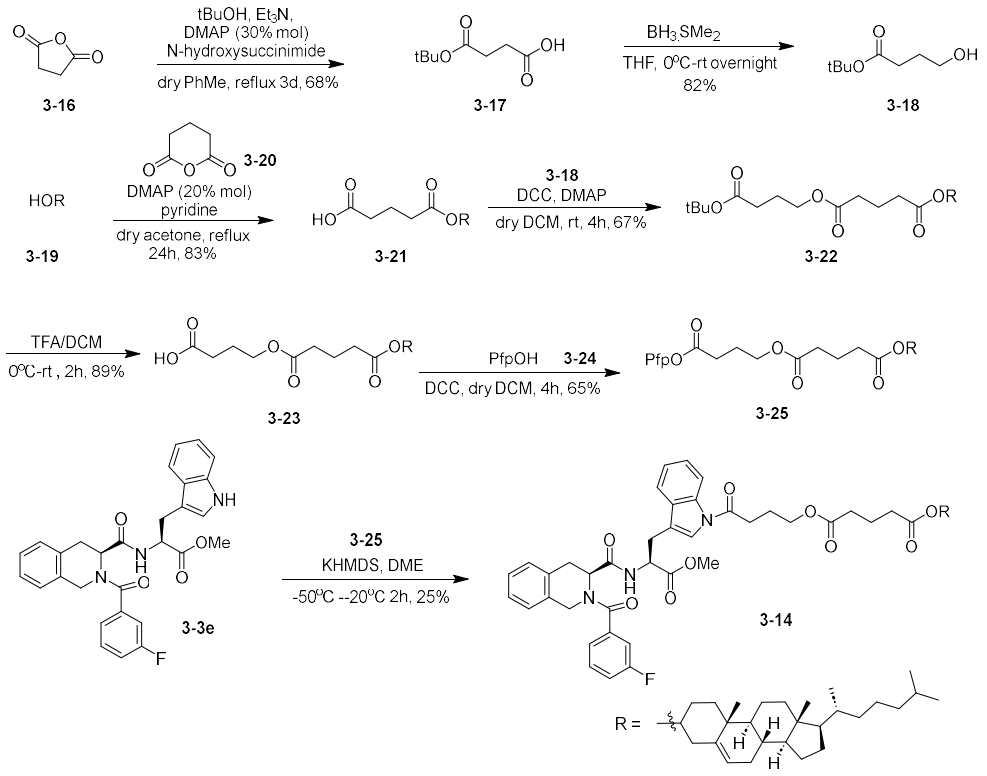
For **3-15**, the linker, after breaking down of the ester bonds, gave diglycolic acid, and 3-oxo-5-hydroxypentanoic acid. The end-hydroxy group was then derived from the alkene through ozonolysis followed by reduction. The ether was then derived from nucleophilic substitution of the bromoacetic acid with allyl alcohol.



Scheme 3.14: Retrosynthesis of EpMF2-cholesterol conjugates **3-14** and **3-15**

3.3.3 Synthesis of indolyl amide EpMF2-cholesterol prodrug conjugate **3-14**

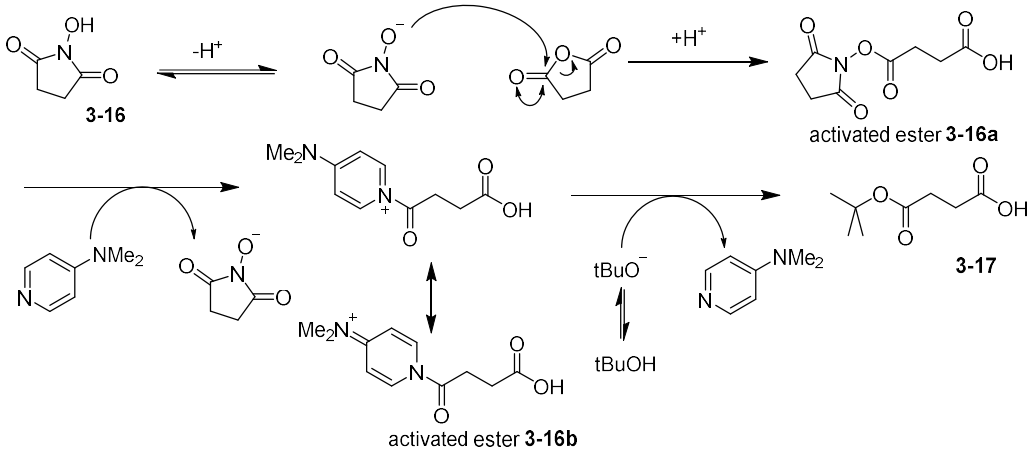
The synthesis of EpMF2-cholesterol conjugate **3-14** with a non-polar linker proceeded as follows in the below scheme 3.14.



Scheme 3.15: Synthesis of EpMF2-cholesterol conjugate prodrug **3-14**.

The synthesis of the first linker fragment began with succinic anhydride **3-14**. Based on the reaction as reported by Guzzo *et al*,[34] the cyclic anhydride was opened up to a tert-butyl monoester **3-16** in modest 68% yield, with DMAP and NHS (*N*-hydroxysuccinimide) as the catalyst. The mechanism was not elaborated in the original paper, but it likely involved nucleophilic acyl substitution at the anhydride and subsequent formation of an activated *N*-hydroxysuccinic ester **3-16a**, which reacts further with DMAP to form an even more active acylpyridinium intermediate **3-16b**, which undergoes a final nucleophilic substitution by tert-butyl alcohol to form the tert-butyl ester **3-17**. (This is a similar mechanism of the activation of carboxylic acids in the Steglich esterification; see below). However, because tert-butanol was a highly hindered alcohol, the subsequent substitution reaction was very sluggish, and an initial reaction with 1.1 equivalents of alcohol after 24 hours of reflux showed incomplete conversion to **3-17** with significant traces of *N*-hydroxysuccinic ester intermediate **3-16a**. This

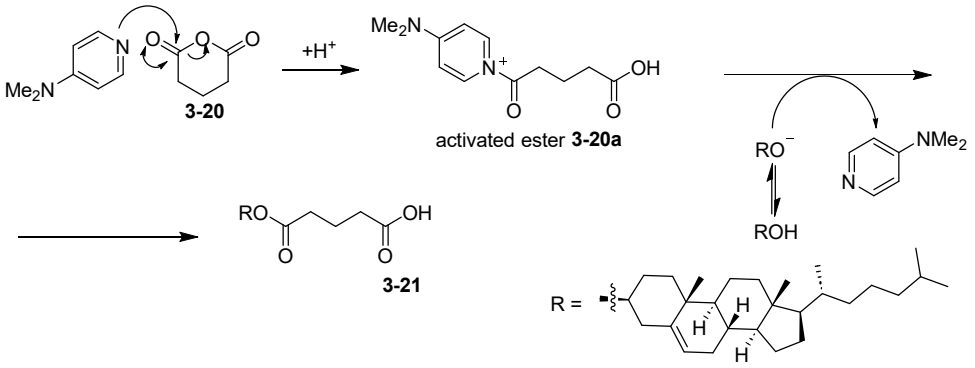
was remedied by increasing the amount of tert-butanol in the reaction to 2 equivalents and extending the heating time to 72 hours.



Scheme 3.16: Proposed Reaction mechanism of the reaction of **3-14** to **3-15.**

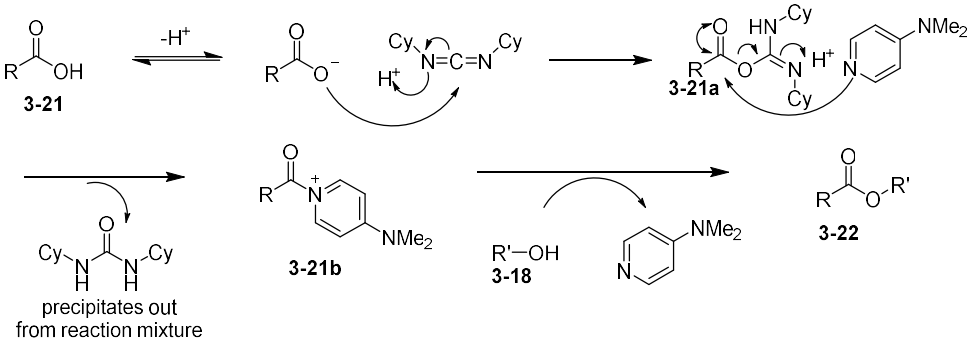
Once **3-17** was formed the carboxylic acid moiety was selectively reduced to the alcohol **3-18** in the presence of ester through borane reduction. An initial attempt using borane-THF resulted in its decomposition to tri(n-butyl) borate without any conversion of **3-17**. However, borane forms a weak complex with THF, can only be stored as a dilute solution (1.0M) in THF, and is known to be somewhat unstable long term when stored above 5°C, even with the addition of $NaBH_4$ as stabiliser.[35] Therefore, the dimethyl sulfide adduct ($Me_2S{\bullet}BH_3$), was used, which is much more stable and can be stored as a neat liquid (10M).[36] Reduction with this complex smoothly gave the alcohol **3-18** in 82% yield.

Meanwhile, the other end of the linker construction began when cholesterol **3-19** was reacted with glutaric anhydride **3-20** and DMAP to form the monoester **3-21** in 83% yield. As before, the reaction likely involved nucleophilic acyl substitution on the anhydride by DMAP to form an activated ester **3-20a** which subsequently undergoes a second substitution reaction to form the cholesteryl monoester **3-21**. A small amount of side product, which might be the diester, was formed. However, being highly non-polar, it was removed by trituration with acetone, in which it was insoluble.



Scheme 3.17: Proposed Reaction mechanism of **3-20** to **3-21** via activated ester formation

The two linker components **3-18** and **3-21** were then coupled together with DCC and DMAP in the Steglich esterification[37] to form the ester **3-22** in 67% yield. The centre carbon atom in the diimide moiety of DCC is slightly electrophilic and reacts with the carboxylic acid of **3-21** to form an activated ester **3-21a**. (*O*-acylurea). Further acyl substitution with DMAP forms a reactive *N*-acylpyridinium intermediate **3-21b**, which reacts with the alcohol to form the ester **3-22** as the final product. The dicyclohexylurea side product has limited solubility in the solvent and can be removed by precipitation at low temperatures.

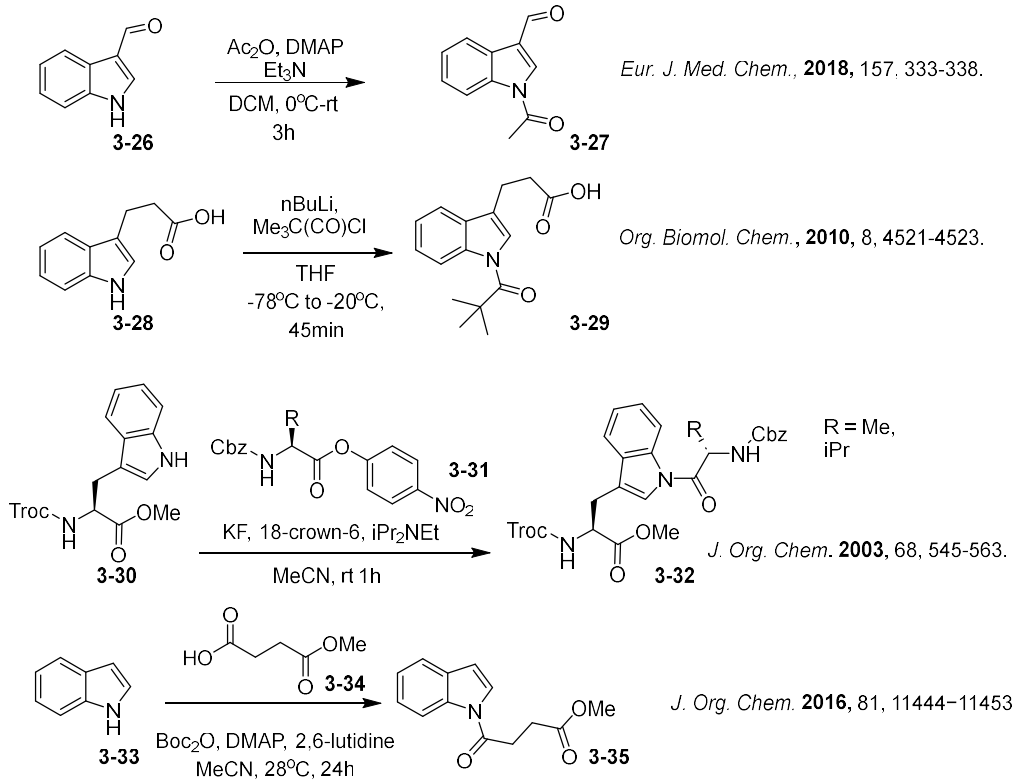


Scheme 3.18: Steglich esterification mechanism using DCC as coupling reagent.

Once **3-22** was isolated, the tert-butyl ester was deprotected with trifluoroacetic acid into the corresponding carboxylic acid **3-23** in 89% yield. The final step involves the coupling of the carboxylic acid to the indole NH of the EpMF2 analogue **3-3e**. Unlike other amines, the indole NH is non-basic and non-nucleophilic, given that the lone pair on N are delocalised into the aromatic ring system. A brief literature search on the *N*-acylations of indoles would be needed to determine a suitable synthetic method.

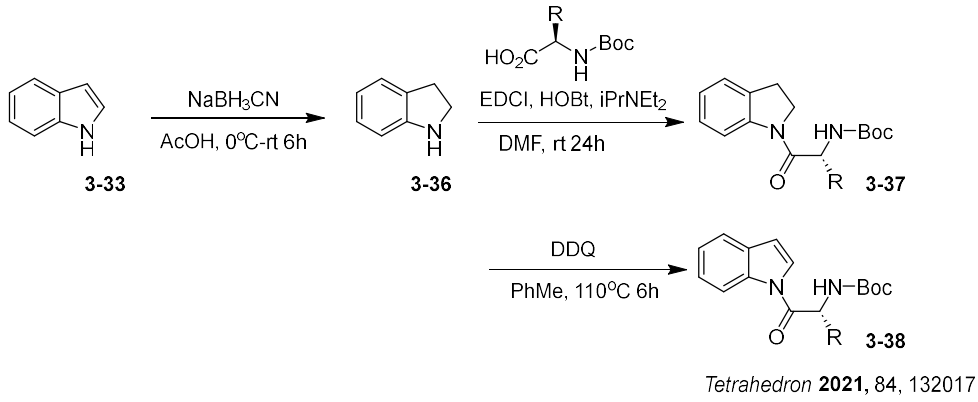
3.3.4 Investigation of amide coupling reaction between indoyl amine and carboxylic acid

In general, literature on acylations of the N atom in indoles (as well as other non-basic nitrogen heterocyclic systems such as pyrroles, pyrazoles etc) is limited. To induce *N*-acyl substitution of the indole, the NH was deprotonated with a base to form a nucleophilic anionic species, which was then coupled to an activated ester of the carboxylic acid. The nature of the activated ester included anhydrides to the synthesis of **3-27** from **3-26**, [38] and acyl chlorides for the synthesis of **3-29**.[39] The use of an electron poor nitro-substituted phenyl ester **(3-31)** to couple the carboxylic acid of an amino acid onto the indoyl nitrogen of tryptophan derivative **3-30** to from an indoyl amide **3-32** has been reported.[40] A rare example (perhaps the only one to the best of our knowledge) of a direct amide coupling between an indole NH **(3-33)** and an underivatized carboxylic acid **(3-34)** to form indoyl amide **3-35** was reported by Tokuyama *et al*, using 2,6-lutidine as base,Boc_2O as additive and DMAP as catalyst.[32]



Scheme 3.19: A few examples of reaction involving the direct amide coupling between the NH of an indolyl compound and a carboxylic acid (as its activated ester). An exception is in the last entry where the carboxylic acid was directly used.

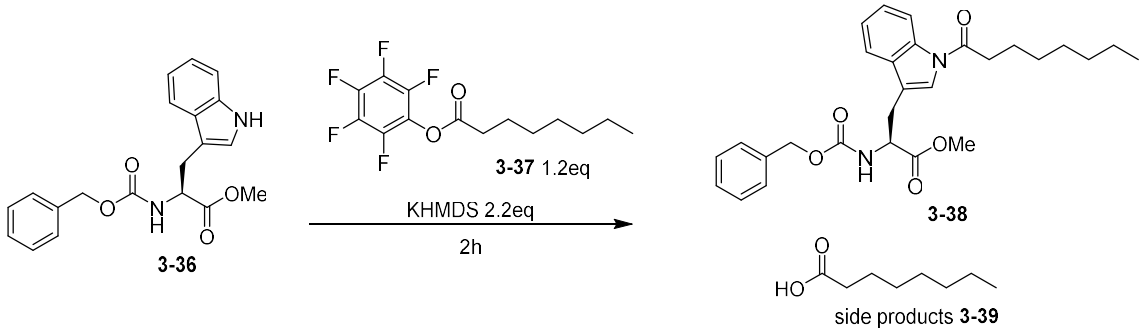
An alternate approach is to use a 3-step indirect coupling method using a saturated indoline intermediate, as done by Chandra *et al*.[42] during their synthesis of pyrazinoindoles. They first reduced indole **3-33** into indoline **3-36** with $NaBH_3CN$. After the N atom was rendered nucleophilic due to its saturation, an amide coupling reaction was done with EDCI and HOBt to form **3-37**. Afterwards, the indoline moiety was oxidised back to the indole **3-38** using DDQ. This route was eventually not considered due to concerns of possible incompatibility of DDQ oxidation with the tetrahydroisoquinolyl moiety in the EpMF2 analogue.



Scheme 3.20: An indirect synthetic method of forming N-indoyl amides via formation of an indoline intermediate.

It was then decided that the coupling reaction between **3-3e** and the cholesterol-linker intermediate **3-23** be done via the first method using a variant of Snider's reaction. Instead of a *p*-nitrophenyl ester, pentafluorophenyl (Pfp) esters would be used as the activated ester moiety. The 5 electron withdrawing F atoms would make the carbonyl ester highly electrophilic and amendable to substitution by the indole. The base used would be potassium hexamethyldisilazide (KHMDS), a sterically bulky base (pK_a = 26)[43] that is strong enough to deprotonate the indoyl NH moiety (pK_a of indole = 21)[44] but bulky enough to leave the methyl ester of **3-3e** intact. A screening of test substrates pentafluorophenyl octanoate **(3-36)** and N_α-benzyloxycarbonyltryptophan **(3-37)** was performed in three solvents, THF, toluene and DME (dimethoxyethane). The choice of solvent is important as the extent of aggregation of the

KHDMS molecule (whether it exists as a monomer, dimer or polymer) varies by solution, which can influence its reactivity.[45]



Scheme 3.21: Top: General scheme of the N-indole amide forming reaction. Bottom: Test reaction of indole NH amide coupling with pentafluorophenyl esters.

Entry	Solvent	Temperature / C°	Starting material: crude product molar ratio based on ^{1}H NMR	Comments
1	THF	-78	2.6: 1	Significant hydrolysis of ester to carboxylic acid
2	PhMe	-78	1.7 : 1	Minor hydrolysis of ester to carboxylic acid
3	DME	-55 (melting point of solvent = -56°C)	0.4 : 1	Minor hydrolysis of ester to carboxylic acid

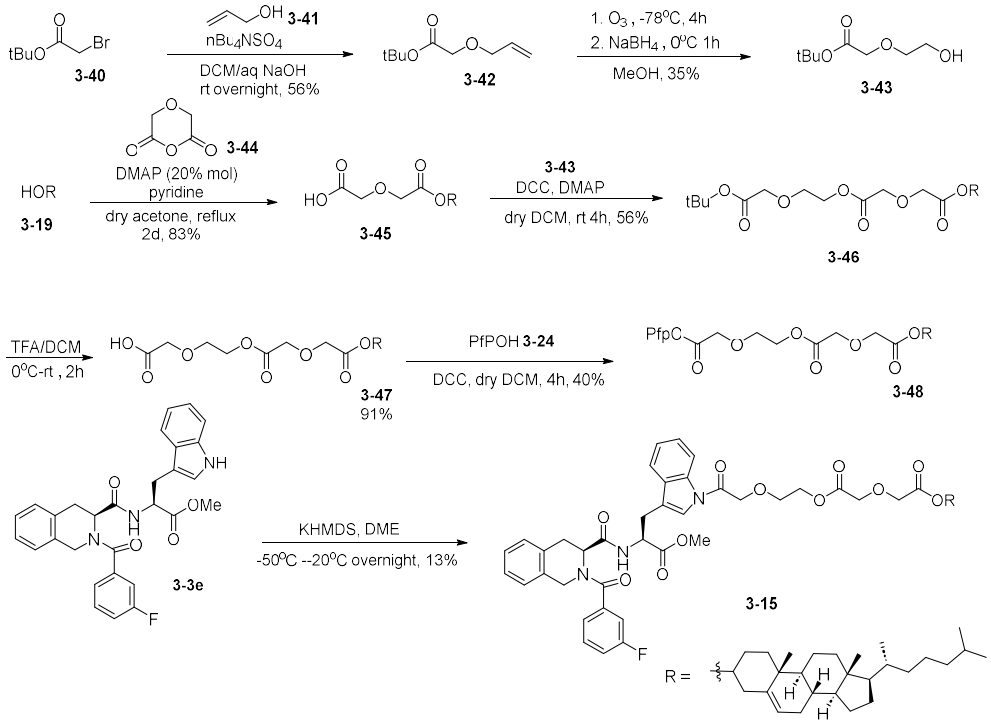
Table 3.22: Indoyl amide test coupling and optimization of **3-36** with **3-37** under various conditions

In all 3 test reactions the indoyl substrate **3-36** was partially converted to the coupled amide **3-38**. There was a side reaction where some Pfp ester **3-37** was hydrolyzed presumably from traces of water, to give the carboxylic acid **3-39** and pentafluorophenol **3-24**, even in very dry freshly distilled solvent. DME was determined to be the most suitable solvent as it had the greatest conversion to the desired product with minimal side reaction of ester hydrolysis.

The above reaction conditions were thus replicated for the EpMF2 analogue **3-3e**. Firstly, the carboxylic acid **3-23** was esterified under Steglich conditions with pentafluorophenol **3-24** to form the Pfp ester **3-25** in a moderate 65% yield. Coupling of this ester with **3-3e** gave the final product **3-14** in a low 25% yield due to poor conversion to the product. The final coupling reaction proved to be challenging and difficult to replicate. Just as before in the model reaction, significant amounts of hydrolysed pentafluorophenol and carboxylic acid **3-23** were found in the crude ^{1}H NMR spectrum, even after thorough drying of solvent. An excess of Pfp ester is required (2eq) to completely react with **3-3e.** The reaction appeared to be extremely moisture intolerant, as the use of non-chromatographically purified Pfp ester (**3-25**), or non-freshly distilled DME, led to complete hydrolysis of **3-25** into **3-23** without any conversion of **3-3e.**

3.3.5 Synthesis of indolyl amide EpMF2-cholesterol prodrug conjugate **3-15**

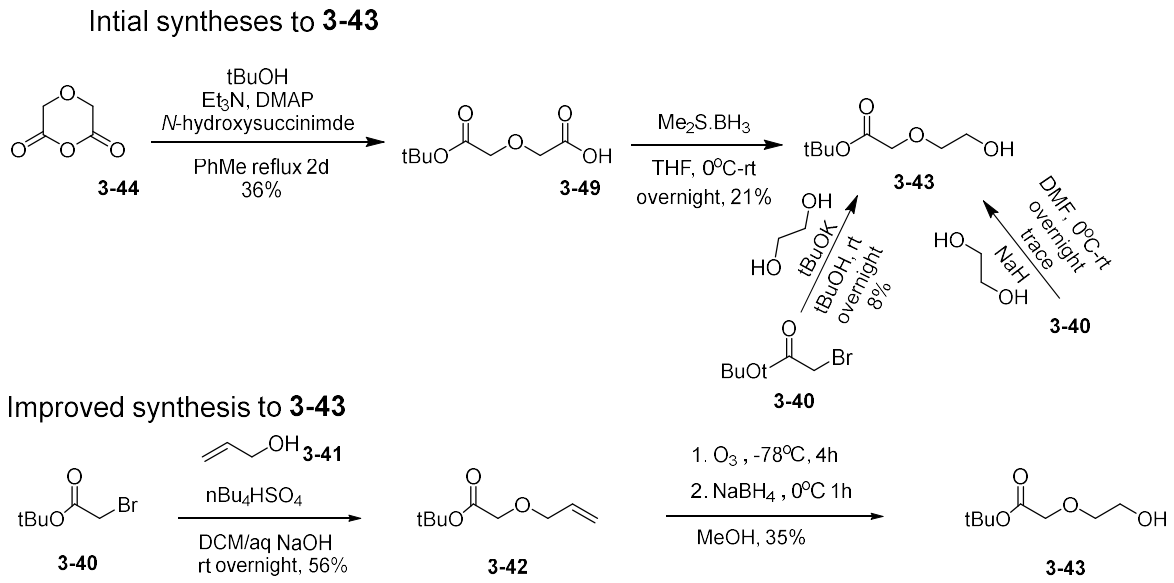
The synthesis of EpMF2-cholesterol conjugate **3-15** with a polar ether-containing linker was done through a synthetic route as described in Scheme 3.23



Scheme 3.23: Synthetic route to EpMF2-cholesterol conjugate **3-13**

The other cholesterol-EpMF2 conjugate, **3-15**, has ether groups in the linkers to increase its water solubility. The use of ether moieties is similar to the related process of PEGylation where drugs were coupled to, or encapsulated in poly(ethylene glycol) to improve the pharmacokinetics of various drugs by increasing aqueous solubility.[33]

The first fragment linker **3-43** was originally prepared from diglycolic anhydride **3-44** in a similar synthetic route as the alkyl linker fragment **3-18** during the previous synthesis of EpMF2-cholesterol conjugate **3-14**. Diglycolic anhydride **3-44** was converted to its t-butyl monoester **3-49** using NHS and DMAP in a mediocre 36% yield. Subsequent reduction of the carboxylic acid moiety in **3-49** by borane-dimethylsulfide to the alcohol **3-43** led to poor yield of 21%. The poor yields in the two reactions were hypothesized to arise from the high water solubility of compounds **3-49** and **3-43** due to the presence of the ether group, leading to high losses during aqueous extractive workup.



Scheme 3.24: Initial synthesis of linker **3-43** led to low yields due to losses in extractive workup of **3-43**. An alternate synthesis (below) was devised to circumvent this and avoid extractive workup.

An alternative synthetic step to **3-43** was devised from tert-butyl bromoacetate **3-40** and ethylene glycol in one step, based on Sasaki *et al.*[46] The ethylene glycol was deprotonated by a strong base and underwent a nucleophilic substitution reaction with the bromine atom of **3-**

40 to form the final product **3-43**. An attempt to use 1.1 equivalents of NaH as the base with 1.5 equivalents of anhydrous ethylene glycol only led to low conversion of **3-40** to the product **3-43**. If the reagents are used in great excess (5eq NaH and 10eq ethylene glycol), conversion was complete, but the product **3-43** became difficult to separate from excess ethylene glycol. (Extraction only led to the product settling in the ethylene glycol rich aqueous layer). An alternate base, tBuOK in dry tBuOH was used, which allowed for complete conversion of **3-40** with only 1.5 equivalents of ethylene glycol, but the difficulty in separating excess ethylene glycol from the product remained, and an attempt at separation by column chromatography led to a poor 8% yield.

Since any liquid-liquid extraction procedure will lead to high losses of **3-43**, An alternative route was devised where **3-43** was synthesized indirectly through the alkene intermediate **3-42** in a reaction that required no extractive work-up. **3-41** underwent a nucleophilic substitution reaction where the bromine atom was displaced by allyl alcohol **3-41** to form the allyl ether **3-42** in 56% yield. This reaction was conducted under biphasic conditions (DCM/water) with NaOH as the base and the use of nBu_4NHSO_4 as a phase transfer reagent to transport the hydroxide ion base to the organic DCM layer where **3-41** is located and substitution can occur. Vigorous stirring of the biphasic mixture is necessary to ensure sufficient mixing to the two phases.

The alkene in **3-42** was then converted to the alcohol **3-43** through reductive ozonolysis at -78°C. A trioxolane was formed as an intermediate[47] which upon further reduction with $NaBH_4$ furnished the alcohol. The workup is simple; evaporation of the methanol was followed by redissolving in EtOAc to precipitate excess borate side products (from $NaBH_4$), followed by filtration of the solids and evaporation of EtOAc gave pure **3-43**.

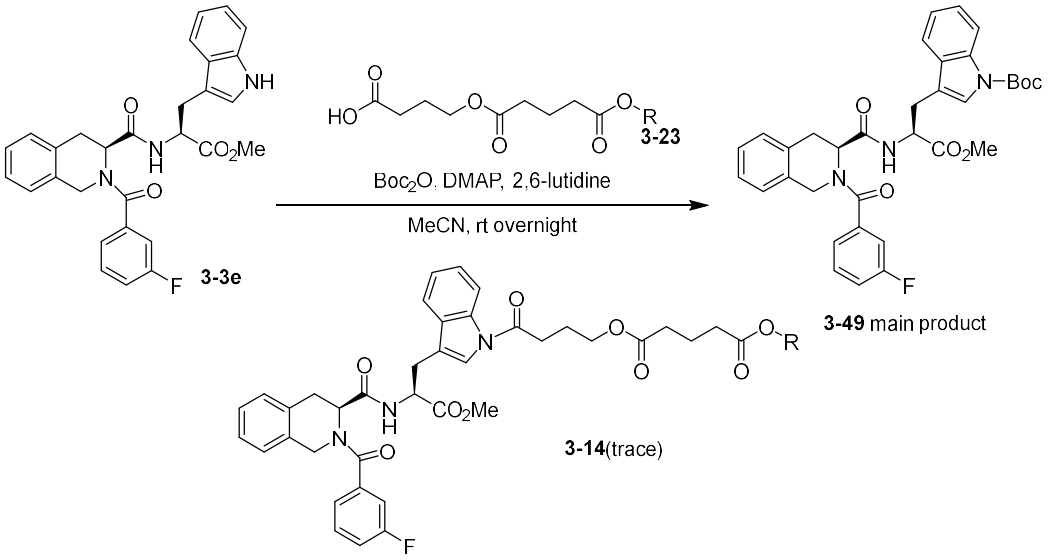
The rest of the synthesis was done in identical conditions as those used in the synthesis of **3-12**. Cholesterol **3-19** reacted with diglycolic anhydride **3-44** with DMAP and pyridine to form

monoester **3-45** in 83% yield, which subsequently reacted with **3-43** under Steglich esterification conditions to form coupled linker compound **3-46** in 56% yield. Deprotection of the tert-butyl ester of **3-46** with TFA gave carboxylic acid **3-47** after 1 hour at 91% yield, and the Pfp ester **3-48** was formed from **3-47** under Steglich esterification conditions with pentafluorophenol **3-24** in 40% yield.

Coupling of **3-48** to EpMF2 analogue **3-3e** proved to be even more difficult than the corresponding non-ether linker cholesterol-linker **3-23**. Firstly, as before, significant amounts of carboxylic acid **3-47** and pentafluorophenol by-product was formed. However, **3-47** had the same R_f value as the final product **3-15** in TLC, thus **3-15** was never obtained pure, but contaminated with **3-47** along with some other unidentifiable peaks. Even then, the yield of this crude product was a paltry 13%.

In addition, the linker of carboxylic acid **3-47** and Pfp ester **3-48** appeared vulnerable to acidic hydrolysis. When the TFA deprotection reaction from t-butyl ester **3-46** to **3-47** was done overnight, a mixture of expected compound **3-47** was obtained along with **3-45,** indicating that the middle ester in the linker was broken. Column chromatography of **3-47** and **3-48** with 1% AcOH as the eluent led to isolation of only **3-45**, suggesting complete decomposition from acid. In addition, the coupling reaction was successfully done only once, and subsequent efforts to replicate this experiment failed, with the only reactions observed being Pfp ester hydrolysis.

The use of alternate conditions for this difficult coupling step was attempted. Reacting **3-14** under Tokuyama's coupling conditions[32] with Boc_2O and DMAP only gave traces of the coupled product with almost quantitative amount of indolyl *N*-Boc side product being formed instead. While other methods were yet to be tried (i.e. indoline intermediate coupling), it was decided that the effort required to figure out a reliable synthetic method for the formation of indoyl amides was not worth the time, and our attention turned to more easily synthesized, ester coupled analogues instead.



Scheme 3.25: An attempt to use Tokugawa's conditions in indoyl amide coupling only resulted in the installation of the Boc group on the indole nitrogen of **3-3e** to form side product **3-49**

3.3.6 Proposed structure of phenyl ester EpMF2-cholesterol prodrug conjugates

Our attention then turned to EpMF2 analogue **3-3f**, with a hydroxy group. Attachment of the linker can be done from the hydroxy group through an ester group. As esters were easier to synthesize than indoyl amides as in compounds **3-14** and **3-15**, more analogues were proposed and successfully synthesized as shown in the below scheme 3.26.

As esters were easier to hydrolyse by lipases than amides, the linker structure did not require an internal ester moiety as seen in **3-14** and **3-15**. The polyether linker structure as in **3-15** was abandoned due to issues over its stability during its synthesis.

Analogues synthesized were group into 3 types based on the structure of the linker. One type is a simple alkyl chain linker as seen in **3-50.** The second type as in **3-51** contains a lysine incorporated within the linker. The third type, seen in **3-52** and **3-53a-e** has a 1,2,3-triazole moiety within the linker.

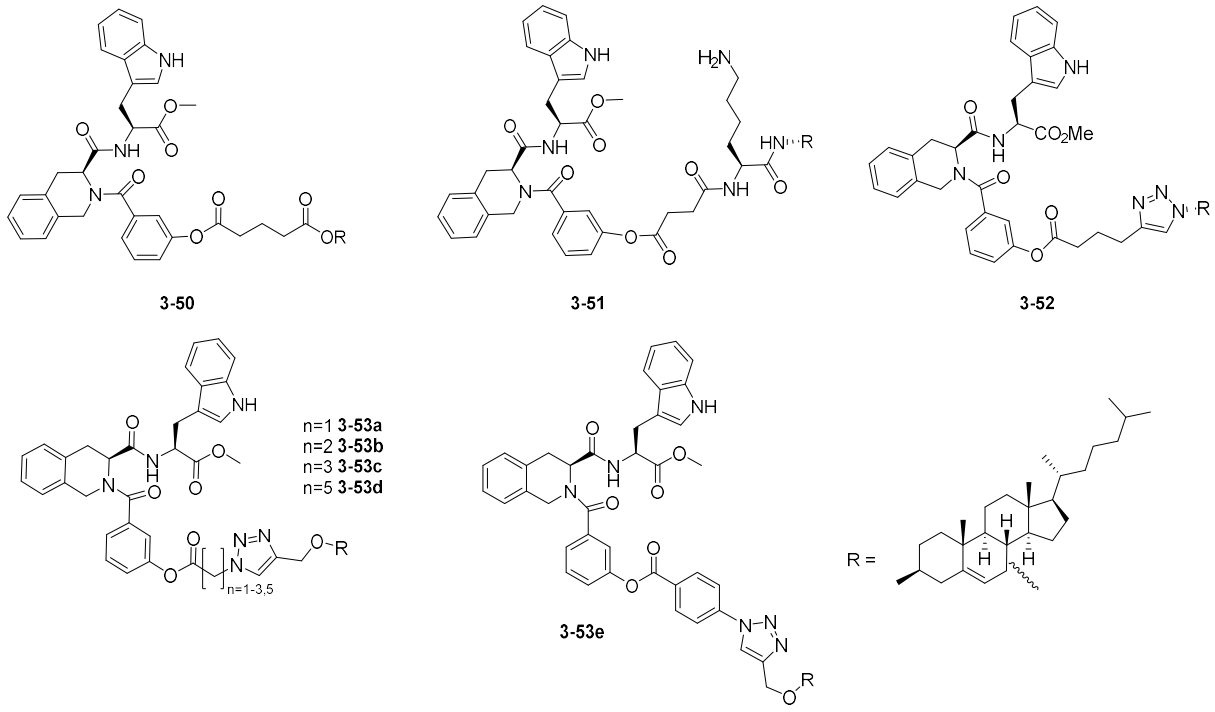
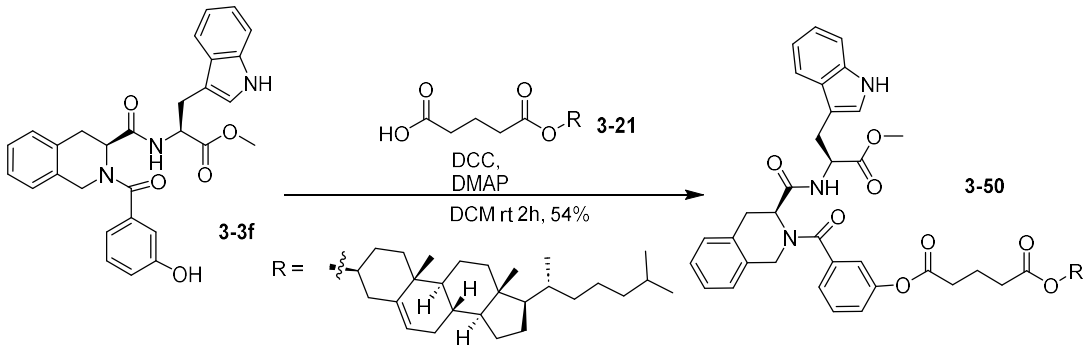


Figure 3.26: EpmF2-cholesterol conjugate compounds with ester linkers derived from EpMF2 analogue **3-3f** that were synthesized

3.3.7 Synthesis of phenyl ester EpMF2-cholesterol prodrug conjugate **3-50**

Compound **3-50** has the simplest linker structure, with the hydroxy groups of **3-3f** and cholesterol linked by ester groups to a glutaric acid moiety.



Scheme 3.27: Synthesis of EpMF-2-cholesterol conjugate **3-50**

3-50 was the first ester linker analogue to be synthesized due to the presence of available starting materials **3-3f** and **3-21** already synthesized in the previous sections and chapters. A single step ester coupling under Steglich conditions provided **3-50** in 54% yield after purification.

3.3.8 Retrosynthesis and synthesis of phenyl ester EpMF2-cholesterol prodrug conjugate **3-51**

Compound **3-51** has a lysine incorporated into the linker chain, to imitate the structure of cationic antibacterial polymers. This structure was inspired by Salunke's synthesis of lysine-cholic acid amphiphilic conjugate molecules **(3-54)**, which have potent killing activity against a variety of bacterial species, including *S. aureus*, and *E. coli,* as well as fungal species like *C. albicans*.[48]

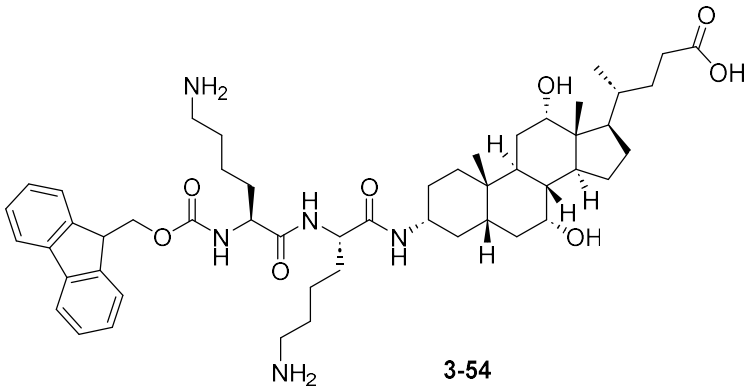
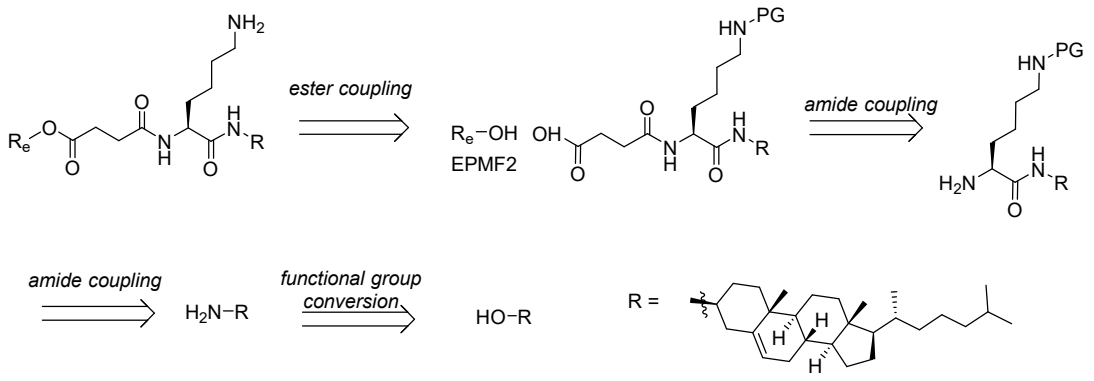


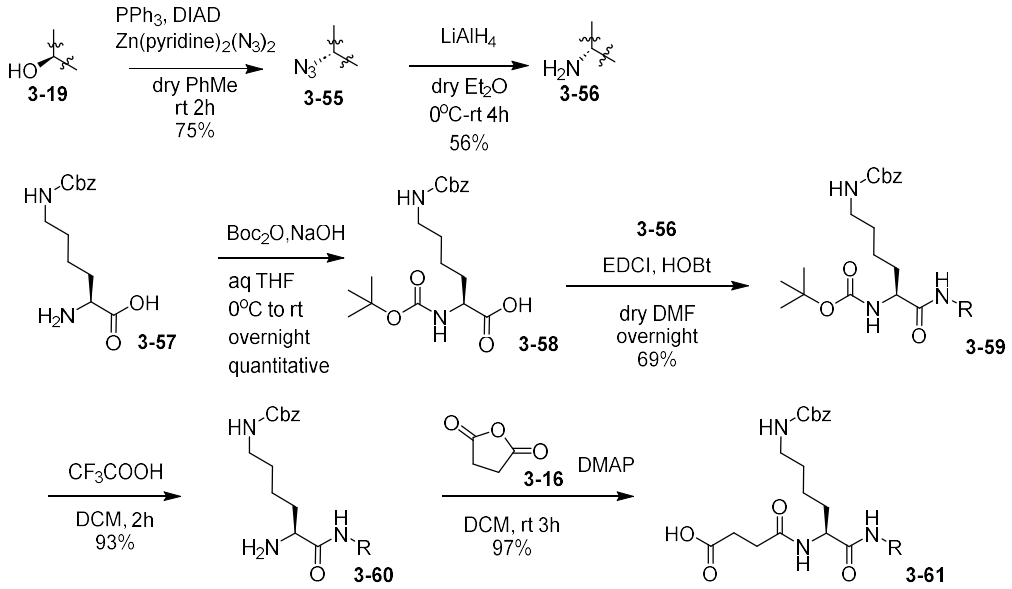
Figure 3.28: Structure of one representative compound of the cholic acid-lysine conjugates **3-54** synthesized by Salunke et al, and its activities against bacterial and fungal species.

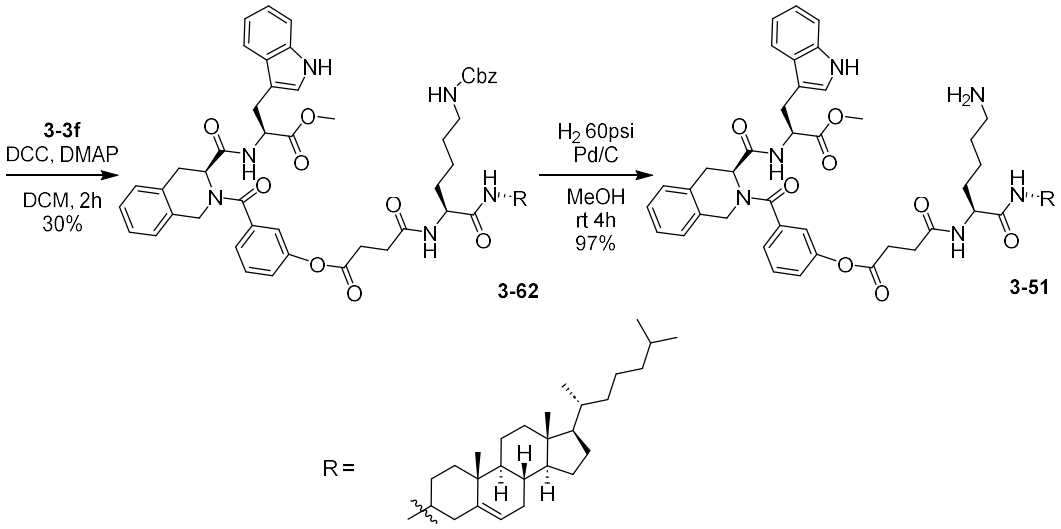
Compound **3-51** has a more complicated linker structure, but its retrosynthesis was still simple. The ester link to the EpMF2 and the linker was first disconnected to the corresponding carboxylic acid. Next, disconnection of the middle amide would separate the succinic acid moiety and reveal a lysine-cholesteryl conjugate. The final connection would separate the far-right amide to the lysine and amino-substituted cholestene derivative. The ε-nitrogen on the lysine side chain would need to be protected during ester/amide coupling steps to prevent interference during ester/amide coupling of the linker. The amino-substituted cholestene can then be derived from cholesterol.



Scheme 3.29: Retrosynthesis of the linker in EpMF2-cholesterol conjugate compound **3-51**

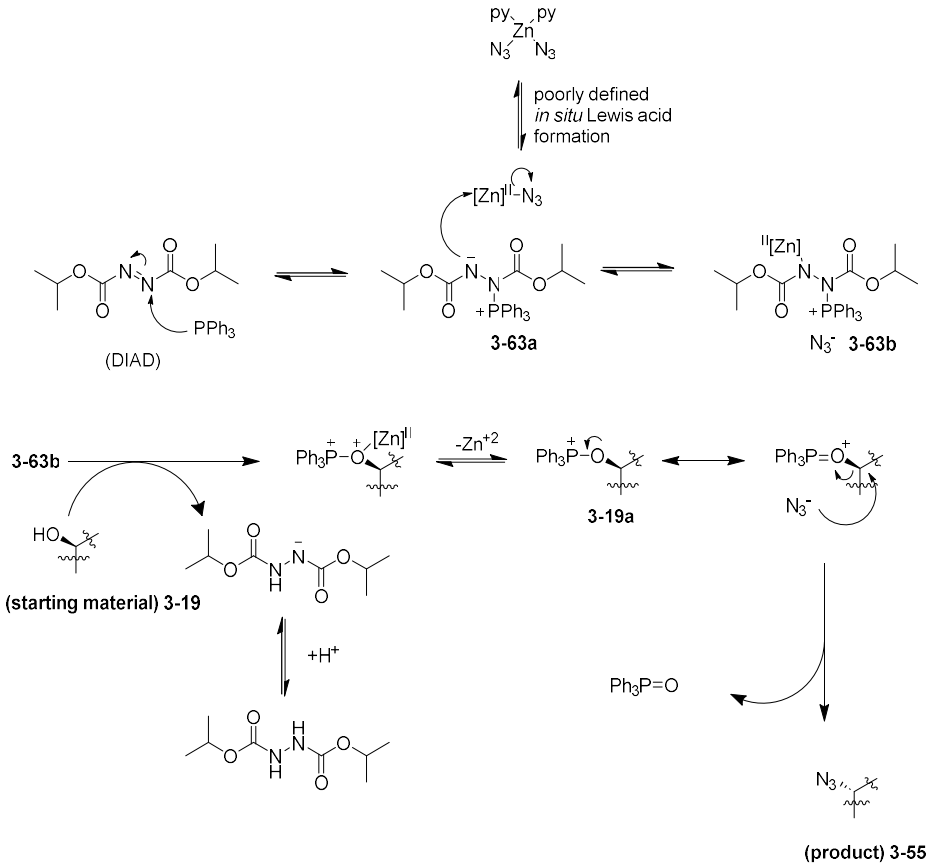
Based on the retrosynthesis scheme above, the synthesis of compound **3-51** was done through the following synthetic steps as described in Scheme 3.30.





Scheme 3.30: Synthesis of EpMF2-cholesterol conjugate analogue **3-51**

To begin, cholesterol **3-19** was converted to its corresponding amine **3-56** through the azide intermediate **3-55**. The hydroxyl group underwent nucleophilic substitution with the azide compound $Zn(pyridine)_2(N_3)_2$ aided by DIAD, in a variant of the Mitsunobu reaction as reported by Rollin *et al*.[49] The azide **3-55** was isolated in 75% yield after 2 hours of reaction time, approximately similar to the 80% as was described by the authors. As expected for Mitsunobu reactions, there was a complete inversion of configuration at the alpha carbon. However, the stereochemistry of the 3-carbon in cholesterol is not expected to affect its penetration ability, given that cholesterol serves as a carrier rather than a binding site.



Scheme 3.31: Mechanism of action of the Mitsunobu reaction using azides as the nucleophile.

The reaction began when the P atom of triphenylphosphine undergoes nucleophilic attack on the azo moiety of DIAD to form a phosphonium intermediate **3-63a**. The subsequent step is not fully defined. The intermediate **3-63a** has a formal negative charge on the carbamate and rapidly abstracts either a proton from *in situ* formed HN_3 or a Zn^{2+} ion to **3-63b**, depending on whether a Brønsted or Lewis acid is involved in the reaction. (The presence of the Zn(II) complex suggests the latter is more likely. The paper by Rollin did not propose a mechanism.) The key step involves the transfer of the PPh_3 moiety from the **3-63b** to the hydroxyl group of cholesterol **(3-19)**, forming the reduced hydrazine by-product and the oxyphosphonium compound **(3-19a**). **3-19a** then underwent a nucleophilic substitution by the azide to form the secondary alkyl azide **3-55** as the final product, while also forming one equivalent of triphenylphosphine oxide. This substitution is a bimolecular S_N2 reaction, and indeed, the azide compound **3-55** was isolated solely as the inverted 3α substituted azide from the 3β substituted

cholesterol. However, the stereochemistry of the 3-carbon of the cholest-5-ene derivative is not expected to be important as the cholesterol is only meant to be used as a carrier.

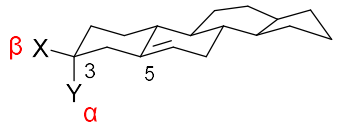
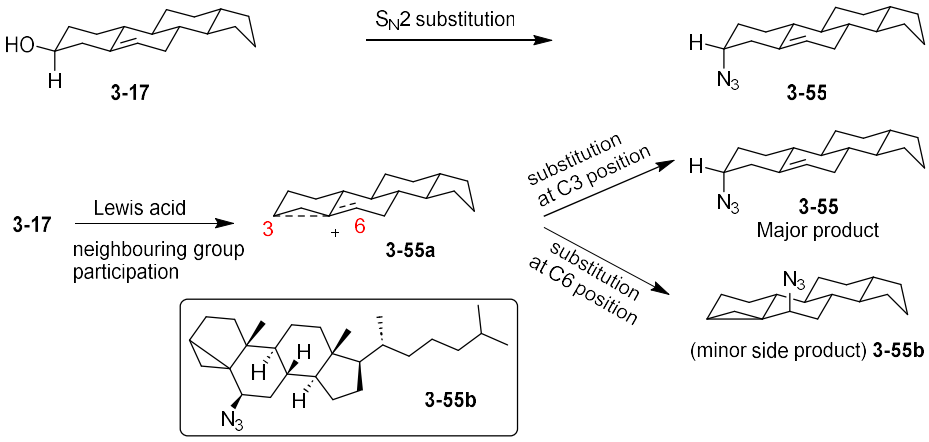


Figure 3.32: The alpha and beta positions of substituents from the 3-carbon of cholest-5-ene derivative. Natural cholesterol is also known as cholest-5-ene-3-ol. (Other alkyl substituents omitted for clarity)

Although not reported by Rollin, the ^{1}H NMR spectrum of **3-55** had several unexplained peaks which was found to be small amounts (~10% mole ratio) of chromatographically inseparable side product **3-55b**. According to reports by Peterson and El Sayed [50, 51] in the presence of Lewis acid a delocalised carbocation **3-55a** can be formed arising from neighbouring group participation of the alkene bond. (Scheme 3.33) Substitution at the position 3 or position 6 of this intermediate would lead to the desired product **3-55** and side product **3-55b** respectively. The only plausible source of Lewis acid in the reaction might come from *in situ* hydrolysis of the bis(pyridine)zinc azide complex. In any case, using freshly distilled solvent minimized this side reaction.



Scheme 3.33: Chemical structure of side product **3-55b** and its possible formation from neighbouring group participation of the alkene. Other substituents not shown for simplicity.

While it is possible to use additional PPh_3 to directly convert the cholesterol **3-19** to its amine **3-56** through a one-pot Staudinger reduction, this was not done. The main reason was ease of workup. Both the Mitsunobu reaction and Staudinger reduction creates stoichiometric amounts

of triphenylphosphine oxide as a side product, which must be removed from the mixture. Due to the highly non-polar nature of **3-55** compared to **3-56**, it was easy to separate **3-55** from triphenylphosphine oxide by eluting with pure hexane in chromatography. However, this is much more difficult for **3-56** as it tends to co-elute with triphenylphosphine oxide.

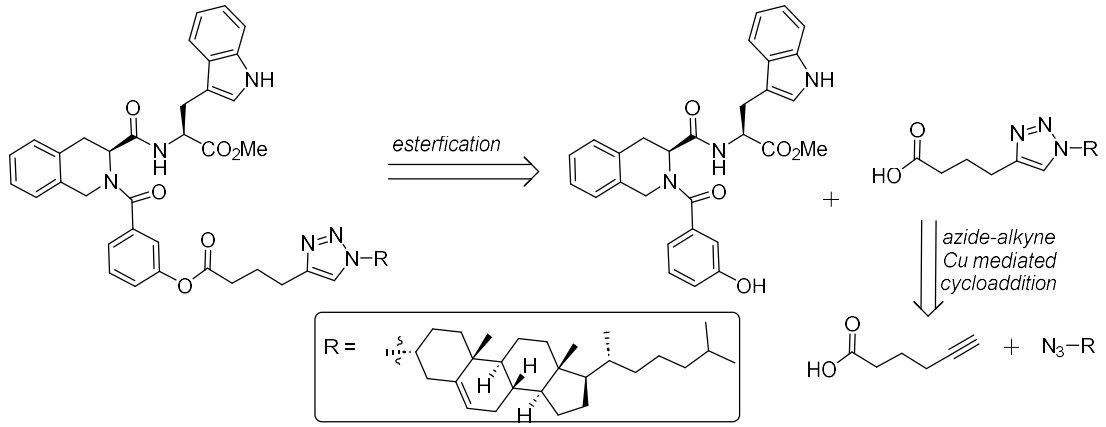
Therefore, azide **3-55** was reduced to the amine **3-56** using $LiAlH_4$ in anhydrous Et_2O. This furnished the compound in 56% yield, with no purification required.

It was now possible to attach the cholesterol-derived amine to the linker. Continuing on scheme 3.30 above, commercially available N_ε-Cbz L-Lysine **3-57** was protected at the α nitrogen to the Boc group with Boc_2O in aqueous THF and NaOH to form **3-58** at quantitative yield. The carboxylic acid of **3-58** was then coupled to the amine of **3-56** with EDCI/HOBt to form **3-59** in 69% yield. The Boc group of **3-59** was then deprotected with TFA to **3-60** in 93% yield. This compound then underwent nucleophilic acyl substitution reaction with succinic anhydride **3-14** catalysed by DMAP to furnish the compound **3-61** smoothly in 93% yield, completing the linker synthesis. This was coupled to EpMF2 analogue **3-3f** under Steglich esterification conditions with DCC to form the penultimate compound **3-62** in a modest 30% yield. The final step was to deprotect the Cbz group on the lysine side chain via hydrogenolysis under a medium pressure H_2 atmosphere (60psi) in 4 hours to unmask the amine in the final product **3-51** in an almost quantitative 97% yield.

3.3.9 Retrosynthesis and synthesis of phenyl ester EpMF2-cholesterol prodrug conjugate **3-52**

Compounds **3-52** and **3-53a-e** contain a 1,2,3-triazole moiety. Triazoles are common motifs in many bioactive drugs, including antiviral drug TSAO and anticancer drugs carboxyamido-triazole.[52] The related 1,2,4-triazoles also widely occurs in antifungal drugs such as fluconazole.[52] More importantly, 1,2,3-triazoles were easily prepared from their corresponding alkyne and azide starting materials through a Huisgen copper(I) catalysed azide-alkyne

cycloaddition (CuAAC) reaction pioneered by Sharpless[53] and Meldal[54]. The ease of the synthesis and workup, its tolerance towards water, the high atom economy, and the extreme specificity of the reaction between the alkyne and azide has led to this kind of reactions to be labelled as "click" chemistry.



Scheme 3.34: Retrosynthesis of EpMF2-cholesterol conjugate **3-52**.

The retrosynthesis of analogue **3-52** containing a 1,2,3-triazole moiety would start from the breakage of the ester to the carboxylic acid **3-65** and EpMF2 analogue **3-3f**. **3-65** would be synthesized from the azide **3-55** and hex-5-ynoic acid **3-64** through a Cu(I) mediated formal [2+3] cycloaddition reaction (CuAAC) as will be described in the below scheme 3.35.

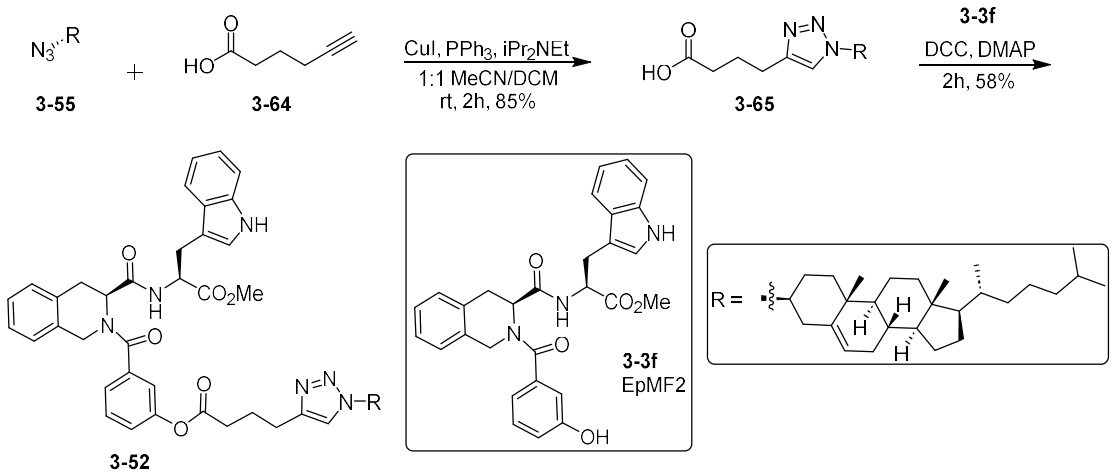


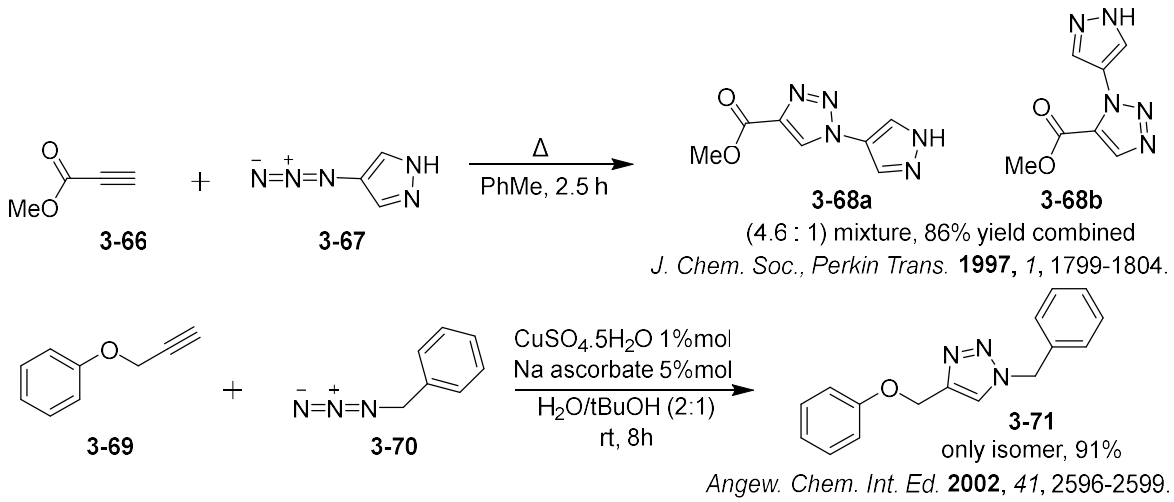
Figure 3.35: Synthetic route to EpMF2-cholesterol conjugate **3-52**

The Cu(I) catalysed azide-alkyne cycloaddition (CuAAC) reaction, since its initial report by Sharpless in 2002,[53] had more than 14 thousand citations and has become highly popular in

synthetic organic chemistry as well as biochemistry. Reviews of its applications can be found in a diverse variety of fields such as pharmaceutical science,[55] materials[56] and polymer science,[57] supramolecular chemistry[58], and in biosciences including its use in bioligation,[59] nucleotide modification[60] and carbohydrate chemistry[61], among others. The reason for its popularity can be ascribed to characteristics of its high atom economy, extremely simple workup, high reaction specificity only between alkynes and azides and tolerance of aqueous conditions including under physiological conditions. These characteristics generally fall under the label of "click" chemistry, and the CuAAC has been named as a premier example of such click reactions by Sharpless himself.[62]

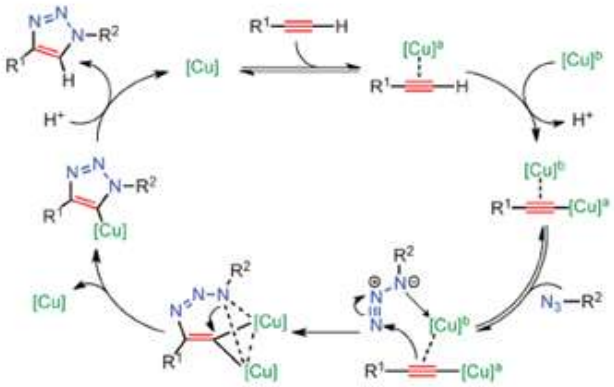
The original alkyne azide cycloaddition reaction as reported by Huisgen were uncatalyzed reactions that proceed through a direct pericyclic cycloaddition mechanism.[63] The uncatalyzed reactions required high temperatures, which led to competing azide decomposition.[64] The regioselectivity of the reactions were also poor, with 1,4 and 1,5 substituted triazoles being formed, although electron withdrawing groups on the alkyne would favour 1,4 formation.[64] For example, the reaction between methyl propiolate **(3-66)** and 4-azido-1H-pyrazole **(3-67)** under heating gave a 4.6:1 mixture of 1,4 and 1,5 regioisomers **(3-68a-b)** respectively. (Scheme 3.36) The reaction was revolutionised by Sharpless[53] and Meldal[54] who discovered that the addition of a Cu(I) compound significantly accelerated the reaction, allowing it to be done at room temperature. The reaction gives exclusively the 1,4-substituted triazole. It also works with a wide variety of azides and alkynes (alkyl and aryl, electron withdrawing and donating substituents, primary, secondary and tertiary azides), and tolerant to almost all functional groups. The only restriction required is that the use of terminal acetylenes is necessary. As an example, the Cu(I) catalysed reaction between phenyl propargyl ether **(3-69)** and benzyl azide **(3-70)** furnished 1,4-regioisomer product **3-71** exclusively. (Scheme 3.36) The reaction can be

done at room temperature, and workup was very simple: dilution with water to precipitate the product followed by filtration to obtain pure compound.



Scheme 3.36: Examples of azide-alkyne cycloaddition reactions. The uncatalyzed condition (top) gave a mixture of regioisomers, and Cu(I) catalysed conditions (bottom) formed only one isomer under mild, aqueous conditions. Cu(I) was formed in situ by ascorbate reduction of Cu(II) sulfate.

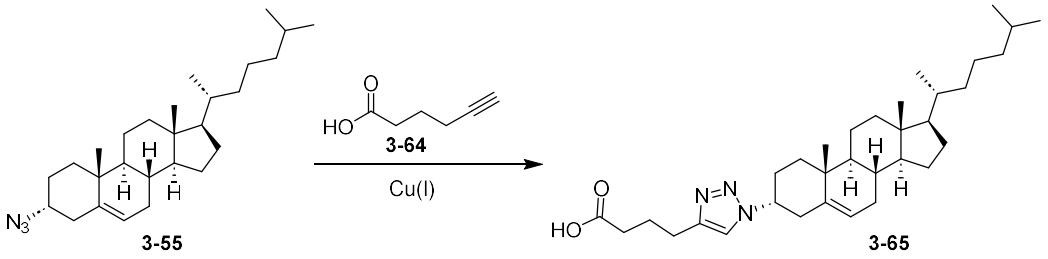
The reaction mechanism of the Cu(I) catalysed cycloaddition is still under elucidation, but it is not a concerted pericyclic addition as in the uncatalyzed variant. The mechanism is fairly complex involving multiple steps that begins with the attachment of the copper(I) atom to the alkyne to form a copper (I) acetylide. Subsequent isotope labelling studies have suggested that formation of a dinuclear two-copper atom intermediate may be the key to the cyclisation.[65, 66]



Scheme 3.37: Mechanism of the CuAAC as proposed by Fokin[65]and whose paper was this image adapted from. Fokin also authored the initial report of CuAAC.

The source of Cu(I) usually come from two types of copper compounds. The most common variant, also used by Sharpless in the original paper, is to use a Cu(II) salt and reduce it *in situ* using the reducing agent ascorbate, then conduct the reaction in a water-alcohol solvent mixture. This has the advantage of allowing the reaction to be done under air without the need to specifically exclude oxygen. Although Cu(I) species are highly unstable to disproportionation in water and oxygen, the use of ascorbate in large excess relative to the Cu(II) precatalyst consumes any oxygen and protects the Cu(I) species from unwanted oxidation. After the reaction, workup is a simple matter of adding water and filtering the precipitated product, which is usually pure enough to not require further chromatographic steps.

The second, less common variant is to use the copper(I) compound directly as catalyst. Some compounds that have been used for these types of reactions are CuI, $Cu(PPh_3)_3Br$ and $Cu(MeCN)_4BF_4$ amongst many others. This allows for reactions to be done in purely organic solvent, but usually must be done in an inert atmosphere and requires the addition of a base. They are also more prone to side reactions, especially if oxygen is present.[53]



Scheme 3.38: CuAAC reaction between azide **3-55** and alkyne **3-63** to form 1,2,3-triazole **3-64**

Entry	Cu(I) source	Conditions	Results
1	$Cu(II)SO_4$ 1% mol + Na ascorbate 5%mol	1:1 tBuOH/H_2O RT overnight	Both azide and alkyne starting materials recovered
2	$Cu(II)SO_4$ 10% mol + Na ascorbate	1:1 tBuOH/H_2O RT overnight	Both azide and alkyne starting materials recovered

	50%mol		
3	$Cu(II)SO_4$ 10% mol + Na ascorbate 50%mol	1:1 tBuOH/H_2O, buffered to pH = 8 RT overnight	Both azide and alkyne starting materials recovered
4	CuI 10% mol	DCM/MeCN 1:1 RT overnight	Both azide and alkyne starting materials recovered
5	CuI 10% mol + 2eq iPr_2NEt	DCM/MeCN 1:1 RT overnight	4:5 product:azide from crude ^1H NMR
6	CuI 20% mol + 2eq iPr_2NEt	DCM/MeCN 1:1 RT overnight	9:1 product:azide from crude ^1H NMR
7	$Cu(I)(MeCN)_4BF_4$ + 1eq iPr_2NEt	DCM/MeCN 1:1 2h	1:9 product:azide from crude ^1H NMR
8	CuI 10%mol + PPh_3 30%mol + 2eq iPr_2NEt	DCM/MeCN 1:1 2h	Complete conversion, 42% isolated yield after column chromatography

Table 3.39: Click reaction optimization between **3-55** and **3-64** under various conditions.

Synthesis of **3-64** using CuAAC from 3α-azido cholest-5-ene **3-55** and hex-5-ynoic acid **3-63** was attempted under a variety of conditions. First the Sharpless conditions were attempted with copper (II) sulfate and sodium ascorbate in aqueous tBuOH. Unexpectedly, no reaction was detectable (entry 1), and the starting materials were recovered after overnight stirring, even after increasing the catalyst loading (entry 2) and adjusting the pH to 8 through the addition of a buffer (entry 3). It seemed that the azide, with was attached to a large non-polar steroid

hydrocarbon moiety, is so insoluble in tBuOH and water that its concentration in solution was almost negligible, hindering the progress of the reaction. This was contrary to Sharpless's reports that the reaction can proceed even when the substrate is not soluble in the reaction solvent.[53]

The CuAAC reaction was thus modified to use a Cu(I) reagent in organic solvent, where the azide would be more soluble. CuI was initially chosen because it is a commonly used Cu(I) compound and also partially soluble in organic solvents used in the reaction. The reaction did not proceed without addition of a base (entry 4) but adding 2 equivalents of DIPEA this time did result in progress in the reaction. However, CuAAC reaction specifically for the synthesis of **3-64** had difficulties not encountered in the click reaction of subsequent 1,2,3-triazole analogues. The cyclization reaction was only partially complete (4:5 product/azide) after overnight stirring with 10% mol of catalyst (entry 5). Increased catalyst loading to 20% increased the conversion to 9:1 product/azide after overnight. (entry 6). CuI exists as a polymeric compound and has limited solubility in organic solvents, which may explain the low rate of conversion. Substitution with another organic solvent-soluble Cu(I) complex $Cu(MeCN)_4BF_4$ was also not successful and led to only 10% conversion of the azide after 2 hours. (Entry 7).

However, an attempt to add PPh_3 and CuI form $CuI(PPh_3)_3$ in situ also caused the complex to fully dissolve and the reaction went to complete conversion in less than 2 hours (entry 8). However, this condition led to the formation of small amounts of triphenylphosphine oxide, which required purification of the product via chromatography to obtain the final product in 42% yield, which calls into question the "click" nature of this conditions.

The requirement of a base to activate CuAAC reactions in organic solvents directly using Cu(I) compounds was consistent in reported literature.[53, 67] The mechanism of the CuAAC begins with the deprotonation of the terminal alkyne to form of activated Cu(I) acetylides. Terminal

alkynes are ordinarily very weak acids, but a Cu atom that forms a π-complex with the alkyne triple bond makes the terminal H much more acidic. In the Sharpless condition in alcohol/water, the solvent (as well as ascorbate) can abstract the H^+ to aid formation in the acetylide. But in organic solvents, which in our case used DCM/MeCN, the solvent molecules are not proton acceptors, which necessitates the addition of DIPEA as the base to do the job.

It was found that the order of addition of reagents was important in ensuring the success of this specific reaction. Dissolving the Cu(I), azide and alkyne in solvent, before adding base, led to no product formation whatsoever and isolation of only starting materials. However, with the correct order of addition, the reaction was rapid. The addition of Cu(I) reagent (predissolved in MeCN) must be done last, after the dissolution of azide, alkyne, and base. The reason for this curious phenomenon remains to be fully explained. It might be that the addition of copper(I) directly to the solution of alkyne without base activation resulted in the formation of inactive acetylide-Cu aggregates that resist reaction even after subsequent addition of base. It was known that isolated Cu acetylides do not react with azides to form triazoles.[68]

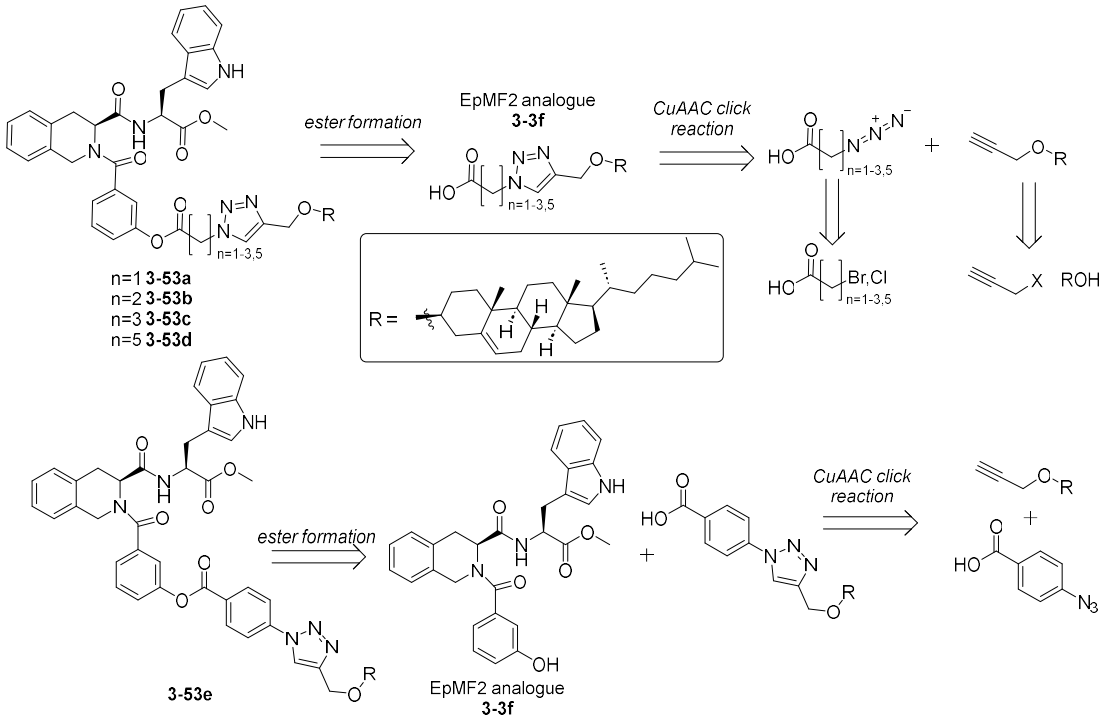
While CuAAC reactions in organic solvent using Cu(I) compounds directly was reported to be prone to side product formation,[53, 68] no side products were observed in our case in the synthesis of **3-65**, as care was taken to conduct the reactions in an oxygen-free atmosphere including purging of solvents with inert gas before reaction.

With the synthesis of **3-65** complete, coupling with **3-3f** to the final product **3-52** using Steglich esterification conditions gave the compound in 58% yield.

3.3.10 Retrosynthesis and synthesis of phenyl ester EpMF2-cholesterol prodrug conjugates **3-53a-e**

3-53a-e also contain a triazole moiety, but opposite in orientation. From the retrosynthesis as devised in scheme 3.42, the azide is now derivatized from the EpMF2 side of the linker, and the alkyne closer to the cholesterol. Fragmentation of the 1,2,3-triazole leaves an azido

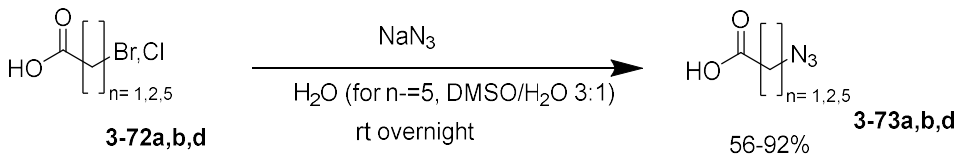
carboxylic acid, where the azide could be derived from the corresponding bromide or chloride. On the other end is a propargyl ether of cholesterol, which could be derived from a propargyl halide and cholesterol. For **3-53e**, fragmentation of the triazole would give the 4-azidobenzoic acid, which was obtained from commercial sources.



Scheme 3.40: Retrosynthesis of EpMF2-cholesterol conjugates **3-53a-e**

The alkyl azido carboxylic acids were first prepared from their corresponding alkyl halide carboxylic acids. (scheme 3.41) This was done by stirring the halide with a fivefold molar excess of NaN_3. An initial reaction of 6-bromohexanoic acid **3-72d** in DMSO gave the corresponding azide **3-73d** in 66% yield but reacting 2-bromoacetic acid **3-72a** in similar conditions gave only trace amounts of product **3-73a** after extractive workup. It appeared that the short chain acids were highly polar, and would not partition into Et_2O during extraction if DMSO was present in the aqueous layer. However, replacing the solvent with H_2O resolved the issue, and in fact gave faster reactions and higher yields. Running the reaction in water for **3-72a** allowed isolation of **3-73a** in 92% yield, and using a 3:1 DMSO:H_2O solution for the

longest chain acid **3-72d** gave higher yields at 88%. Subsequent reaction with 3-chloropropionic acid **3-72b** and **3-73b** in water gave the compound in 56% yield.



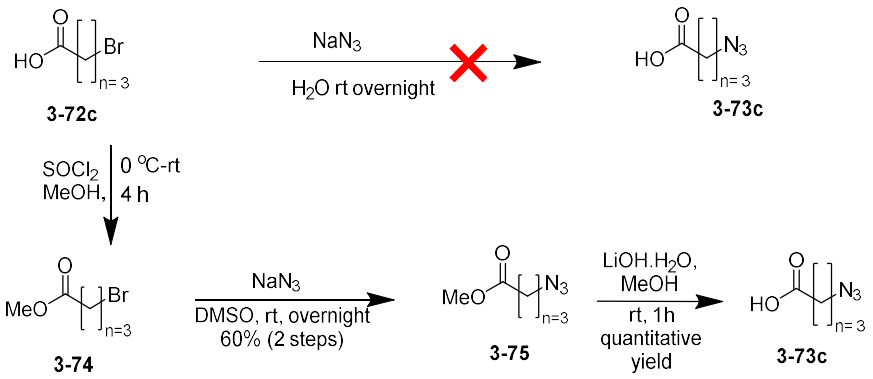
Scheme 3.41: Synthesis of azide precursors **3-72a,b** and **d** to be used for the construction of EpMF2-cholesterol conjugates

It was found that during extraction it was also necessary to acidify the aqueous layer to protonate all acid to encourage dissolution of the product in the organic layer, otherwise yields would be poor. Acidification of azide-containing solutions is usually not recommended due to the risk of forming the highly toxic, volatile, and explosive hydrazoic acid. This was not possible in this case, but as an additional precaution, the discarded aqueous layer after extraction has sodium nitrite added to destroy remaining hydrazoic acid before disposal.

For 4-bromobutyric acid **3-72c**, an unexpected side reaction happened. The ^{1}H NMR spectrum of the product did not agree with literature. Based on comparison of the chemical shifts with other literature it was determined at the compound was cyclised product γ-butyrolactone. It seemed that under aqueous conditions when the carboxylic acid moiety was deprotonated to from the more nucleophilic carboxylate, intramolecular substitution by carboxylate in **3-72c** was kinetically favoured over azide substitution. This was only possible in **3-59c** because the length of its alkyl chain was just right to form a kinetically favoured 5-membered lactone ring. Cyclisation did not happen for **3-72a** and **3-72b** because cyclization will form the highly strained and unstable acetolactone and β-propiolactone respectively. **3-72d** can theoretically form a 7-member ring, but formation of rings larger than 6 atoms is highly kinetically unfavoured over nucleophilic substitution by azide.

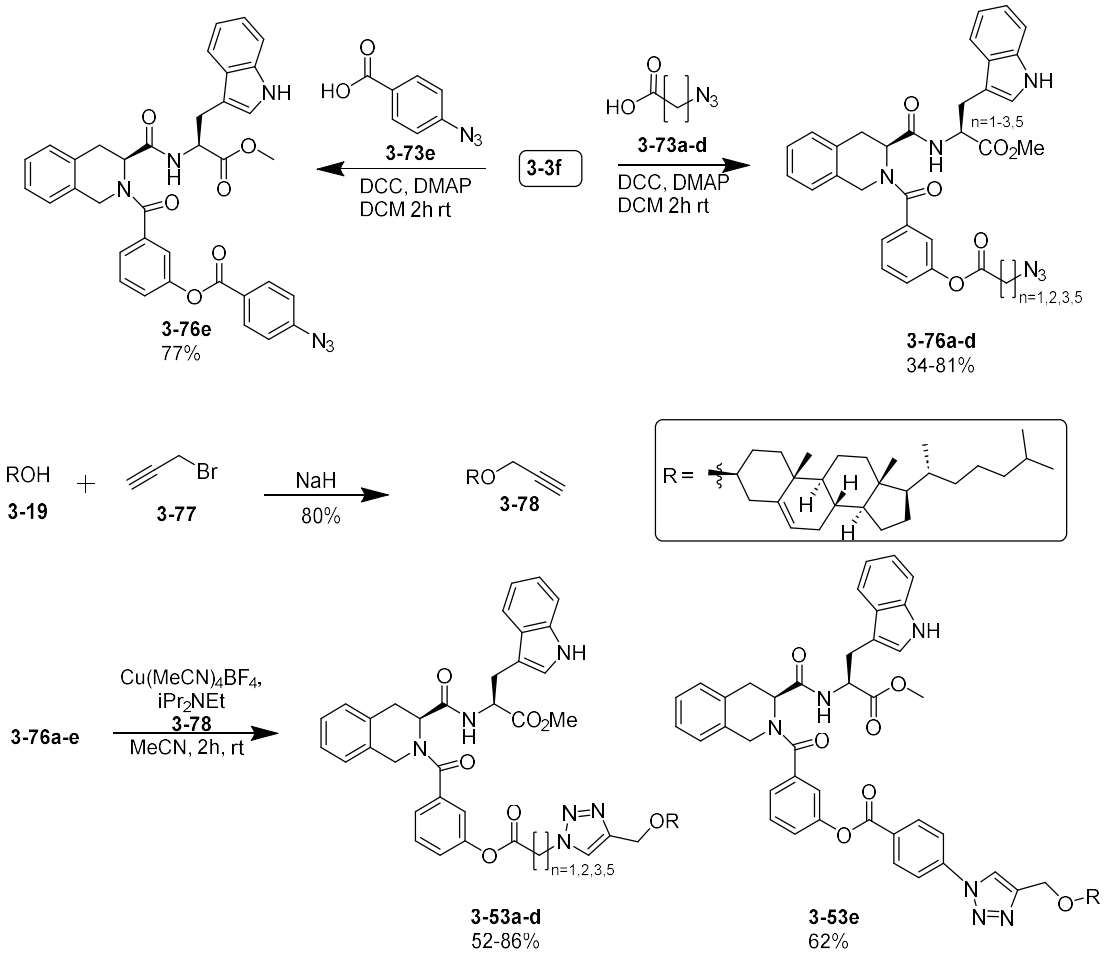
The carboxylic acid of **3-72c** was thus protected to the methyl ester **3-74** using $SOCl_2$ and methanol. Subsequent reaction by sodium azide in DMSO (with the ester now water-insoluble)

gave the corresponding azide **3-75** in 60% over 2 steps. Base hydrolysis with NaOH gave the desired 4-azidobutyric acid without any conversion to lactone **3-72c**.



Scheme 3.42: Synthesis of azide precursors **3-72c** to be used for the construction of EpMF2-cholesterol conjugates. Protection to its methyl ester **3-74** was needed to avoid the formation of side reaction.

The final synthesis of conjugates **3-53a-e** was done as described in scheme 3.45 below.



Scheme 3.43: Synthesis of EpMF2-cholesterol conjugates **3-53a-e**

With the synthesis of alkyl azide carboxylic acids done, the next step was to couple the carboxylic acids to the phenol in **3-3f** under Steglich esterification conditions to form the esters **3-76a-e** in varying yields from 34 to 81% yields. This completed the azide portion of the conjugate.

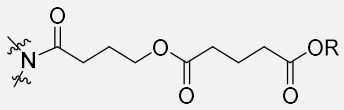
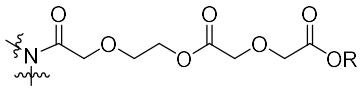
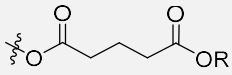
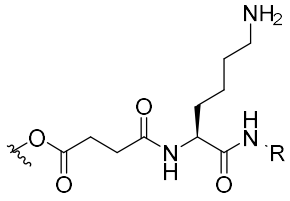
Meanwhile, cholesterol **3-19** was deprotonated with the strong base NaH and reacted with propargyl bromide **3-77** in a nucleophilic substitution reaction to form propargyl cholesteryl ether **3-78** in 80% yield after purification.

Propargyl cholesteryl ether **3-78** and azides **3-76a-e** were then coupled together with CuAAC for the final step. An initial attempt with **3-76d** and **3-78** under Sharpless conditions with aqueous tBuOH and $CuSO_4$ /ascorbate only gave trace amounts of product after overnight stirring. A test reaction of with 4-propargyl benzoic acid was able to give partial reaction after 24 hours, again suggesting the lack of reaction was due to the extreme insolubility of the highly non-polar alkyne **3-78** in aqueous solutions. This was an identical problem to azide **3-55** where its insolubility in tBuOH/H_2O hindered reactivity in CuAAC using water as co-solvent. Thankfully, as before, the use of organic solvents DCM:acetonitrile using a Cu(I) complex as catalyst alleviated the problem. Unlike the case for the synthesis of **3-64**, the use of $Cu(MeCN)_4(BF_4)$ allowed the reaction to proceed smoothly with full conversion in under 2 hours to give the final product conjugates **3-53a-e** in yields ranging from 52 to 86% yields.

3.3.11 Biological assays of EpMF2-cholesterol conjugates

The following assays were done by Dr. Priya Ragunathan and Ms Logeshwari from Professor Grüber's group. Broth dilution assays were carried out on *M. smegmatis* and *M. abscessus* up to 1mM and inhibition was measured by measuring the optical density of samples after incubation of the compound. As a control test, ATP synthesis inhibition assays were carried out on the complete EpMF2-cholesterol conjugate to confirm our hypothesis that the full EpMF2-cholesterol conjugate should have no inhibitory activity.

3-14, 3-53, 3-51 and 3-52 showed some measurable inhibition on the growth of *M. smegmatis*, although no discernible inhibition was observed for *M. abscessus*. No other measurable inhibition was observed for the other conjugate compounds. Although its inhibitory activity was still weak, it did show that the cholesterol prodrug concept, was successful in enhancing the antimycobacterial activity of otherwise MIC inactive compounds.

Compound	Linker Structure	Growth inhibition at 1 mM *M. smegmatis*	Growth inhibition at 1 mM *M. abscessus*	ATP synthesis inhibition in *M. smegmatis* IMV's
3-14		**20% (at 2mM)**	**N.D.**	**N.D.**
3-15		**Not measured due to precipitation**	**N.D.**	**N.D.**
3-50		**1st attempt 3 % 2nd attempt: No inhibition**	**No inhibition**	**No inhibition**
3-51		**1st attempt 36 % 2nd attempt No inhibition**	**No inhibition**	**No inhibition**

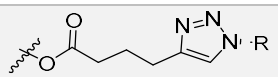
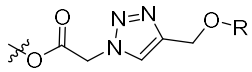
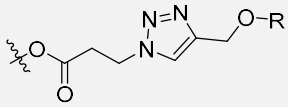
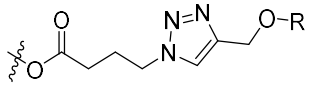
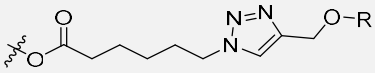
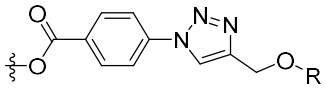
3-52		**1st attempt 41 % 2nd attempt 34%**	**No inhibition**	**No inhibition**
3-53a		**No inhibition**	**No inhibition**	**0.5 µM**
3-53b		**No inhibition**	**No inhibition**	**No inhibition**
3-53c		**No inhibition**	**No inhibition**	**12 µM**
3-53d		**No inhibition**	**N.D.**	**No inhibition**
3-53e		**No inhibition**	**N.D.**	**12 µM**

Table 3.44: Inhibitory activities of various EpMF2-cholesterol conjugates. N.D. Represents experiments that were not done.

No obvious correlation between linker structure and activity can be deduced from the few examples tested. One common difficulty encountered in all assays is the tendency of the compound to precipitate in solution. Given the high lipophilicity of the cholesterol moiety, it might not be unexpected that its attachment might lead to a significant decrease in aqueous solubility of the EpMF2 conjugate under the conditions of the assay. It is very likely that the lack of measured inhibition of the other compounds might simply be due to this fact. It seemed that this broth dilution inhibition assay test, done in an aqueous solution, might not be suitable for type of compound and further types of viability assays should be investigated.

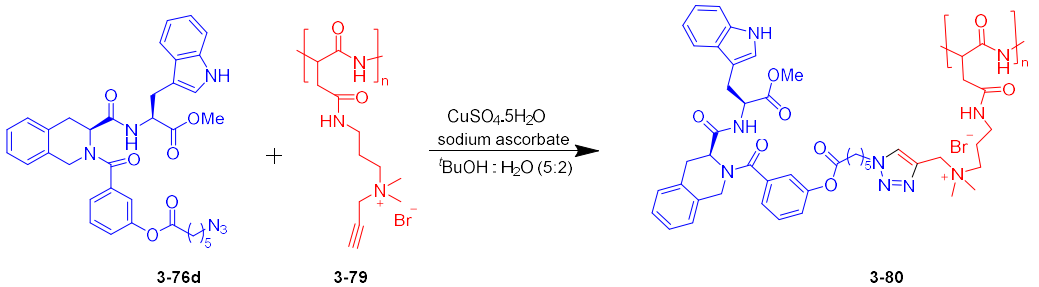
Unexpectedly, some compounds such as **3-47a, 3-47c** and **3-47e** showed inhibition of ATP synthase at very high millimolar concentrations. This observation is not consistent for all conjugates; thus, we hypothesise that any ATP synthase inhibitory activity was not caused by

the whole conjugate molecule itself, but rather from *in situ* hydrolysis of the linker and release of EpMF2 analogue that was responsible for the inhibition. It might be possible that hydrolysis was caused by presence of lipases, which were known be localised on the cell membrane of *M. smegmatis*.[31] If that was the case, it is still a promising observation as it is proof that the linker can indeed be potentially degraded in living bacteria to ultimately release EpMF2 *in vivo*.

Still, the fact that **3-14** and **3-52** were able to inhibit *M. smegmatis* reliably, showed that the cholesterol-intake hypothesis is valid, and that this concept is worth further investigating. Should investigations continue, problems that need to be resolved include finding a more suitable linker to resolve the poor aqueous solubility of the complex. This will be discussed in the final chapter.

3.3.12 Synthesis and characterization of polymer-EpMF2 analogues

A collaborative work with Professor Srivastava's group is at an early preliminary stage where his antimicrobial polymers[18] would be covalently attached to our EpMF2 compounds via triazole formation using click chemistry. As mentioned previously, the polymers carry a positively charged quaternary ammonium moiety that favours its insertion and lysis into the bacterial membrane. It was hypothesized that the combination of the 2 compounds will have a synergetic effect with each other to increase the membrane permeability and antimycobacterial effect of both compounds combined compared to each individual component.



Scheme 3.45: Preliminary click reaction between EpMF2-azide conjugate **3-76d** and alkyne attached polymer **3-79**

A preliminary reaction was attempted by EpMF2 compound **3-61d,** which has an alkyl azide attached as an ester. A variant of Srivastava's polymer **3-79**, containing an alkynyl moiety

attached to the tertiary ammonium moiety, was used for the reaction. The azide and alkynyl moieties would be linked via CuAAC click reaction.

Sharpless conditions were done under aqueous tBuOH with copper (II) sulfate as precatalyst and sodium ascorbate as reducing agent. Unlike the previous reactions where the non-polar nature of the cholesterol containing component hindered its solubility and progress of the reaction, the click reaction for this case proceeded smoothly.

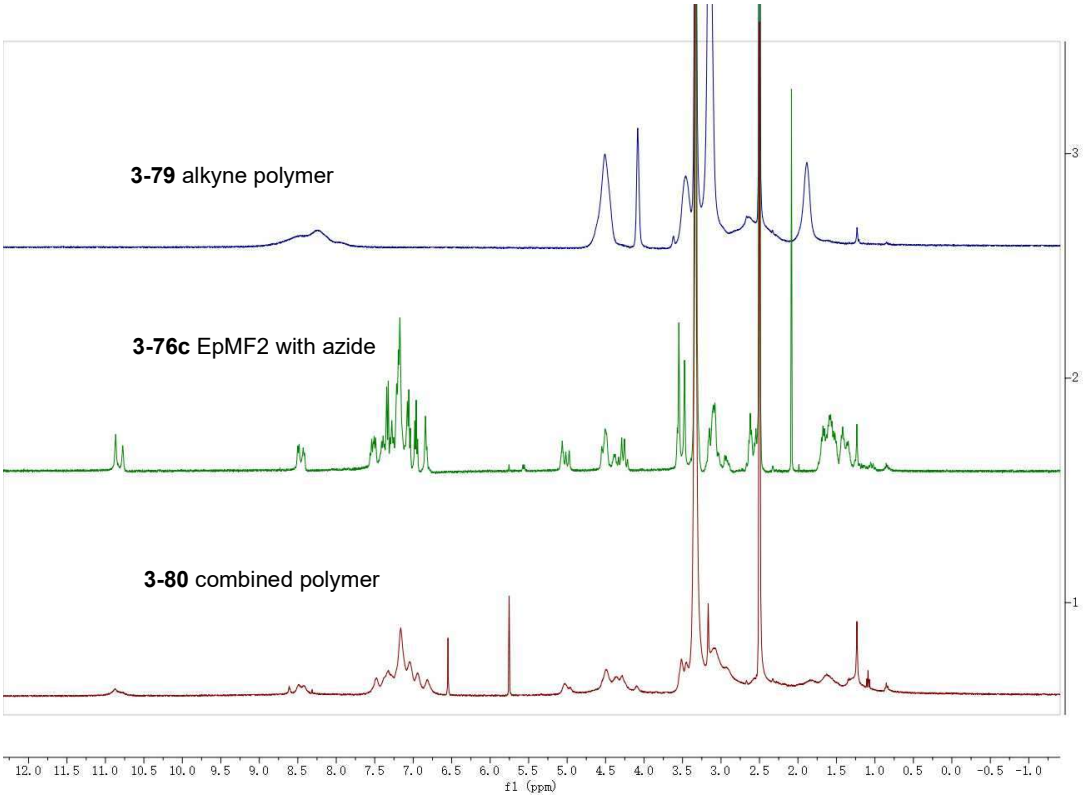
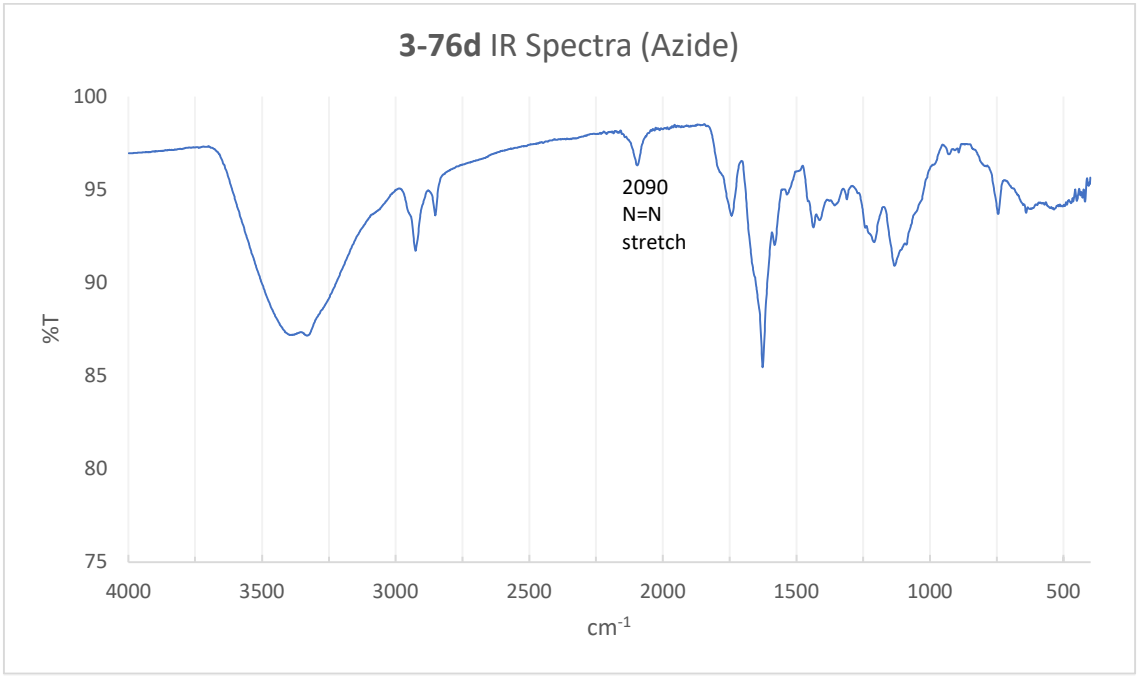
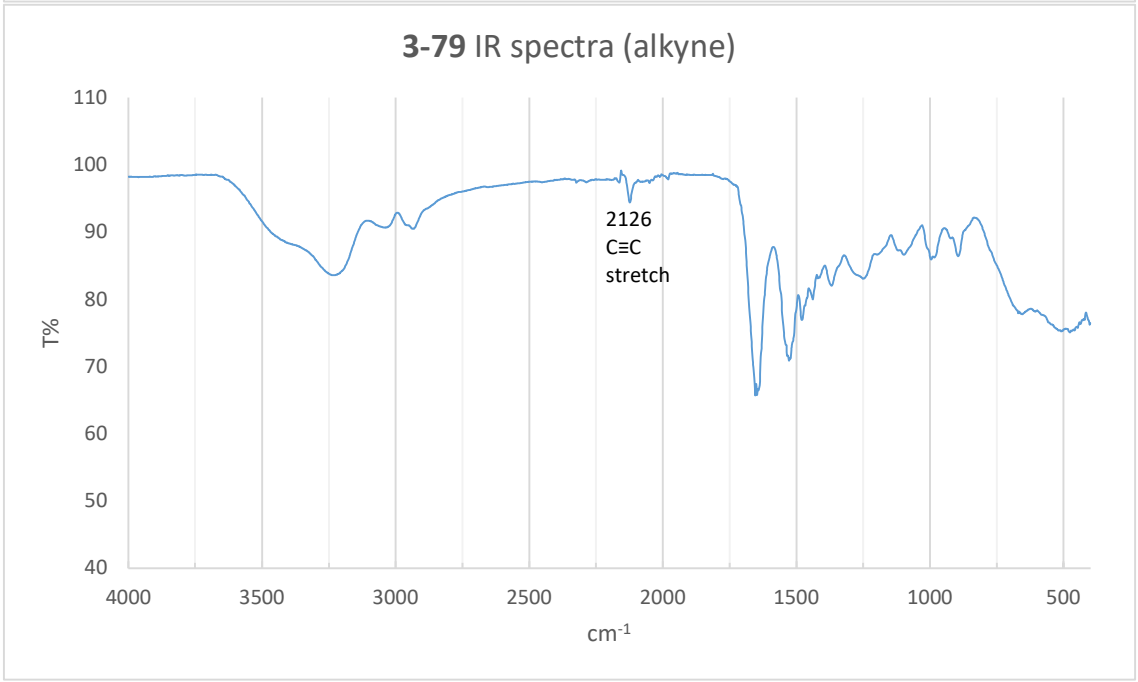


Figure 3.46: ^{1}H NMR spectra of (top) the polymer **3-79**, the EpMF2 analogue (middle)**3-76d**, and the product (bottom)**3-80** after CuAAC. Note the broadening of peaks at the bottom spectra characteristic of polymers. All spectra were recorded in DMSO-d6.


3-76d IR Spectra (Azide)
100
95
90
85
80
75
%T
2090
N=N
stretch
4000
3500
3000
2500
2000
1500
1000
500
cm-1


3-79 IR spectra (alkyne)
110
100
90
80
70
60
50
40
T%
2126
C≡C
stretch
4000
3500
3000
2500
2000
1500
1000
500
cm-1

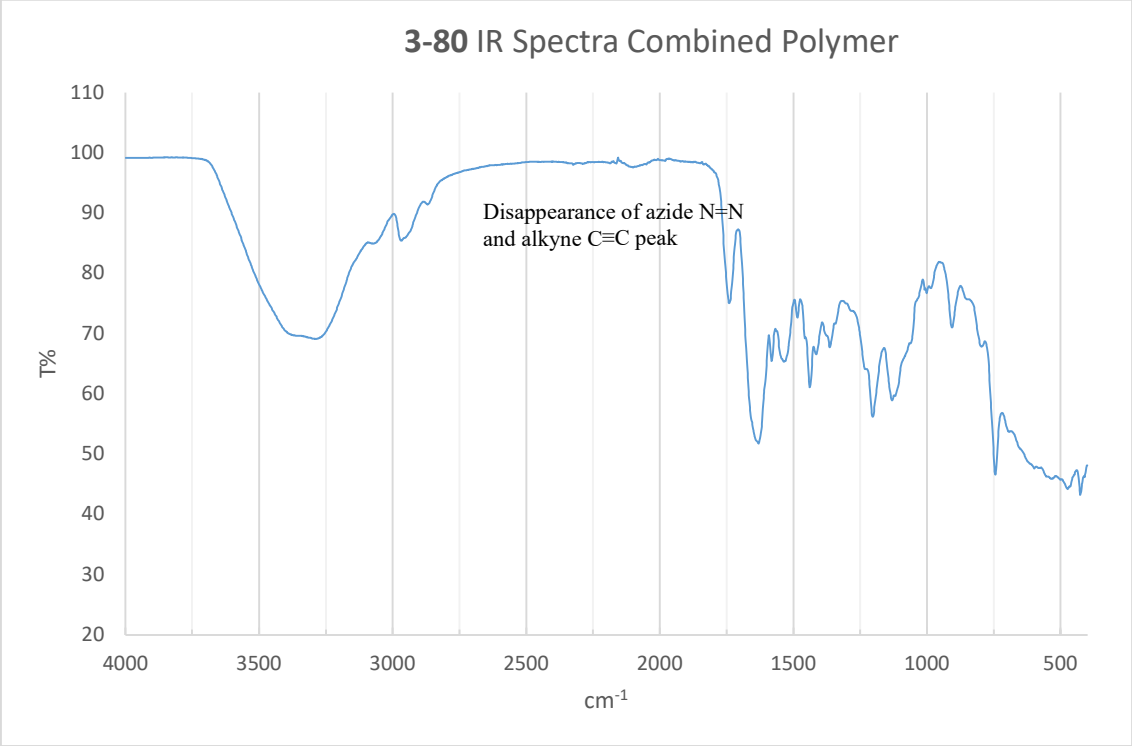


Figure 3.47: 1R spectra of the polymer (left)**3-79**, EpMF2 analogue (middle)**3-76d**, and product (right)**3-80**. Note the disappearance of the alkyne C≡C peak at v = 2126cm^{-1} and azide N=N peak at v = 2090 cm^{-1} in the final product

Evidence of a chemical change was provided by IR and NMR spectra of the product. The ^{1}H NMR of the coupled product showed peaks from the original polymer **3-79** combined with peaks from **3-76d** in the aromatic regions corresponding to δ = 7-11 ppm. This is not just a simple overlay, as the peaks showed broadening, which was characteristic of H atoms from polymers, which showed that the EpMF2 component was chemically linked to the polymer, and not just loosely encapsulated through non-covalent interactions.

IR spectra provided additional evidence as the peak characteristic of alkyne C≡C stretching (v = 2126cm^{-1}) and azide N=N stretching (v = 2090cm^{-1}) found in their respective components were not found in the final product, indicative of completion of the CuAAC reaction with these two functional groups converted into a triazole moiety.

Many other analogues remain to be synthesized, but some issues need to be resolved given the polymeric nature of the drug. First of all, the extent of alkyne conversion to triazole needs to be quantified. Although IR suggest a completion of reaction, it is not quantitative, and the

presence of leftover free alkyne was a possibility. It is important to the determine ratio of EpMF coupling to the alkyne as affects the calculated effective concentration of EpMF2 used.

Secondly, the issue of ensuring complete Cu removal from the peptide needs to be considered. Copper ions are known to be potent antimicrobials due to the oligodynamic effect. Thus even low concentrations of leftover Cu can give false positives during viability assays. For click reactions of small molecules Cu can be removed by filtering a solution of the product in organic solvent through celite, but in a water-soluble macromolecule other methods need to be done. This could include ion-exchange or addition of chelating agents, but more investigations need to be done on these.

3.4 Conclusion

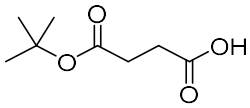
While the effect is still weak, the prodrug concept of attaching cholesterol to a drug to induce its entry into mycobacteria followed by release in vivo is shown to be possible as shown by its success in inhibition of growth of bacteria for some compounds. However, the major obstacle encountered for this strategy is still the high insolubility of such a prodrug in aqueous, physiological conditions as expected from the lipophilic nature of the cholesterol molecule. Therefore, any future drug making use of this cholesterol prodrug concept needs further optimization of the linker chemistry compared to what was synthesized here. More details will be elaborated in the conclusion chapter.

The potential of an antimicrobial peptide-EpMF2 conjugate is still being realized as collaboration efforts are still at an early stage, and it would be difficult now to estimate the success of the strategy. However, given the high aqueous solubility of the peptide compared to cholesterol there will not be issues in solubility. However, further work needs to be done in optimisation of synthesis and selection of analogues.

3.5 Experimental

All chemical compounds are obtained from commercial suppliers and are used as received. When anhydrous solvent was specified, in addition to using dried solvent, oven dried glassware was used (heated at 120°C for at least 2 hours), and the reaction was conducted under a nitrogen/argon atmosphere using Schlenk line techniques. DCM was dried with CaH_2, then distilled and freshly used. DMF was dried by standing with 4Å for overnight, then distilled at reduced pressure at 20 Torr. The distillation temperature was kept below 70 °C to prevent its decomposition. After distillation, DMF was then stored under 4Å molecular sieves. THF, diethyl ether and toluene were dried with sodium-benzophenone until the solution turned blue from the formation of benzophenone ketyl radical, then distilled and freshly used. Methanol was first pre-dried by standing with 3Å molecular sieves overnight, then treated with magnesium and iodine (0.5g Mg and 0.05g I_2 per 100mL of solvent), allowed to auto-reflux from the own heat generated, then thereafter distilled and stored under 3Å molecular sieves. ^1H NMR and ^{13}C NMR spectra were obtained on a 400MHz JEOL ECA400. All spectra were recorded at 25°C. Chemical shifts were reported in parts per million (ppm) relative to tetramethylsilane (δ = 0 ppm) and were calibrated with respect to the residual solvent peak. Coupling constants were reported in Hertz. Coupling splitting patterns were labelled depending on multiplicity. (s for singlet, br s for broad singlet d for doublet, t for triplet, m for multiplet). When multiple peaks arising from the presence of interconverting rotamers exist, it will be indicated. Nominal mass spectra were captured with a Thermo-Fischer Scientific LTQ XL linear ion trap mass spectrometer by the electrospray ionization method. Whether positive or negative mode was used was indicated. High resolution mass spectra were captured with a Xevo® **G2-XS QTof** by the electrospray ionization method. Melting points were measured with a Stuart SMP3 apparatus and were uncorrected.

3.5.1 Synthesis of EpMF2-cholesterol conjugate: Towards Indolyl amide linker **3-14**

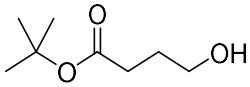


4-(tert-butoxy)-4-oxobutanoic acid 3-17

To a suspension of succinic anhydride **3-16** (500.4mg, 5.00mmol), in toluene (10mL) was added DMAP (61.1mg, 0.3 eq), *N*-hydroxysuccinimide (172.6mg, 0.3eq), tert-butyl alcohol (1.90mL, 2eq), and Et_3N (0.2mL, 0.3 eq), then heated to reflux for 3 days.

After the reaction, the mixture was cooled to room temperature, and the mixture was diluted with EtOAc (20mL), and washed with 2M HCl twice, then by brine once, followed by drying with $MgSO_4$. The solvent was then evaporated in vacuo to obtain the product as a pale yellow oil that solidified upon standing to a pale brown solid (592mg, 68%). The compound was used in the next step without further purification.

Pale brown solid, 68%, Mp: 40-43 °C. ^{1}H NMR (400MHz, $CDCl_3$) δ = 2.62 (t, J= 7.0Hz, 2H), 2.53 (t, J= 7.0Hz, 2H), 1.45 (s, 9H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 177.2, 171.6, 81.2, 30.3, 29.2, 28.2. FTIR (DCM, cm^{-1}) ν_{max} = 3441, 2980, 2934, 1735, 1717, 1153. HRMS calculated for $C_8H_{14}O_4Na$ $[M+Na]^+$ 197.0790. Found 197.0780.

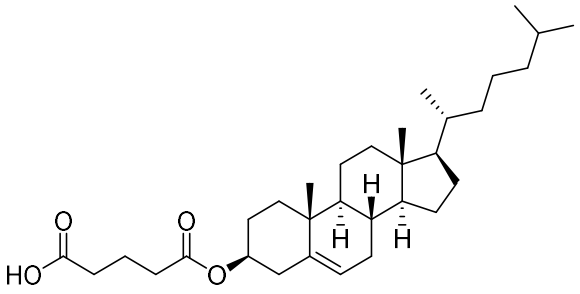


Tert-butyl 4-hydroxybutanoate 3-18

A solution of **3-17** (560mg, 3.21mmol) in dry THF (6mL) was cooled to 0°C and to it added borane-dimethyl sulfide (neat 10M, 268.6mg, 0.35mL, 1.1eq) in an inert atmosphere. The mixture was then warmed to room temperature and stirred overnight. After that, the mixture was then partitioned between EtOAc (20mL) and H_2O (20mL). The layers were separated, and the aqueous layer further extracted with EtOAc twice. The combined organic layers were

washed with brine, dried with $MgSO_4$, and evaporated in vacuo to obtain a pale yellow oil. (423mg, 82%), which was used in the next step without further purification.

Pale yellow oil, 82% yield. ^{1}H NMR (400MHz, $CDCl_3$) δ = 3.67 (t, J= 6.5Hz, 2H), 2.34 (t, J= 7.0Hz, 2H), 2.37-2.31 (apparent quintet, 2H), 1.44 (s, 9H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 173.6, 80.7, 62.5, 32.6, 28.3, 28.0. HRMS calculated for $C_8H_{17}O_3$ $[M+H]^+$ 161.1178. Found 161.1176.



Cholesteryl hemiglutarate 3-21

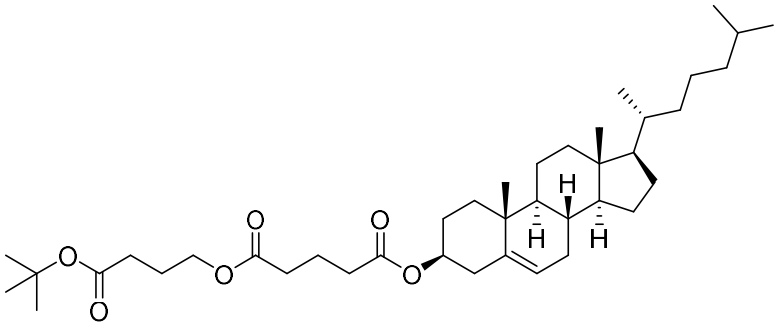
To a suspension of cholesterol **3-19** (2.00g, 5.17mmol) in dry acetone (50mL) was added glutaric anhydride **3-20** (885mg, 7.76mmol, 1.5eq), followed by DMAP (63.2mg, 0.517mmol, 0.1eq) and pyridine (613.7mg, 0.62mL ,7.76mmol, 1.5eq). The whole mixture was heated to reflux for 24 hours. When TLC showed completion, the mixture was cooled to room temperature, and any precipitate (which was the diester side product) was removed by filtration. The solvent was evaporated in vacuo, and the residue redissolved in EtOAc (30mL). Any undissolved precipitate was filtered, and the residue was washed with 2M HCl followed by brine, and subsequent drying with $MgSO_4$. The solvent was again evaporated in vacuo to obtain the crude product as a brown solid. The solid was recrystallised by dissolving in a minimal amount of hot ethanol, then cooling to -20°C and isolation of the precipitated white solid. The mother liquor was further evaporated until more product is precipitated, before storing again at -20°C and isolating a second batch of white solid. (Combined 2.15g, 83%)

White solid, 83%. Mp: 119-121°C ^{1}H NMR (400MHz, $CDCl_3$) δ = 5.37 (m, 1H), 4.64-4.57 (m, 1H), 2.43 (t, J= 7.6Hz ,2H), 2.37 (t, J= 7.6Hz, 2H), 2.32 (d, J= 7.6Hz, 2H), 2.03-1.79 (m, 7H),

1.73-0.85 (m, 31H), 0.67 (s, 3H). ^{13}C NMR (100MHz, CDCl3) δ = 178.1, 172.5, 149.5, 139.8, 122.9, 74.3, 56.9, 56.3, 50.2, 42.5, 39.9, 39.7, 38.3, 37.2, 36.8, 36.4, 36.0, 33.7, 33.0, 28.4, 28.2, 28.0, 24.5, 24.0, 23.0, 22.8, 21.2, 20.1, 19.5, 18.9, 12.0. MS-ESI(-): m/z = 545.34 $[M+formate]^-$, 999.54 $[2M-H]^-$. HRMS calculated for $C_{32}H_{51}O_4$ $[M-H]^-$ 499.3787. Found 499.3781.

General procedure 3: Steglich esterification

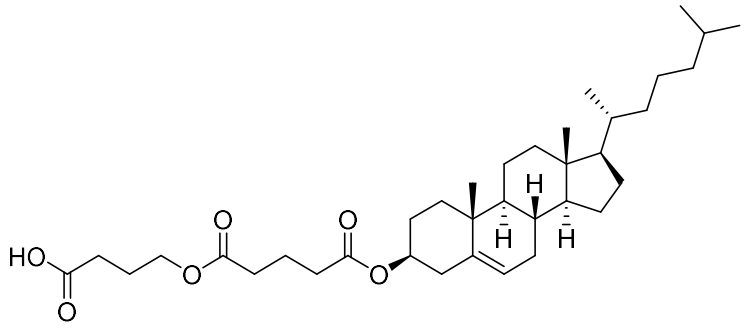
The carboxylic acid (1.0eq) and alcohol (1.0eq) were dissolved in dry DCM and to it added DMAP (0.2 eq) followed by DCC (1.2eq). The solution was stirred at room temperature for 2 hours, and a precipitate of dicyclohexylurea (DCU) formed over the course of the reaction. When TLC showed completion, the DCU precipitate was removed by filtration, and the solvent washed with 2M HCl followed by brine, and then dried with $MgSO_4$. The solvent was evaporated in vacuo and the residue was redissolved in a minimum amount of acetonitrile and stored overnight at -20°C, precipitating more DCU. This was then again removed by filtration and the solvent again evaporated in vacuo to obtain the product, which was purified by column chromatography if necessary.



cholesteryl 4-tert-butoxy-4-oxobutyl glutarate 3-22

3-18 (176.0mg, 1.1eq) and **3-21** (500.0mg, 0.998mmol) were reacted according to general procedure 3. The residue was purified by column chromatography (10:90 EtOAc:hexane) to obtain the product as an off-white solid. (430mg, 67%)

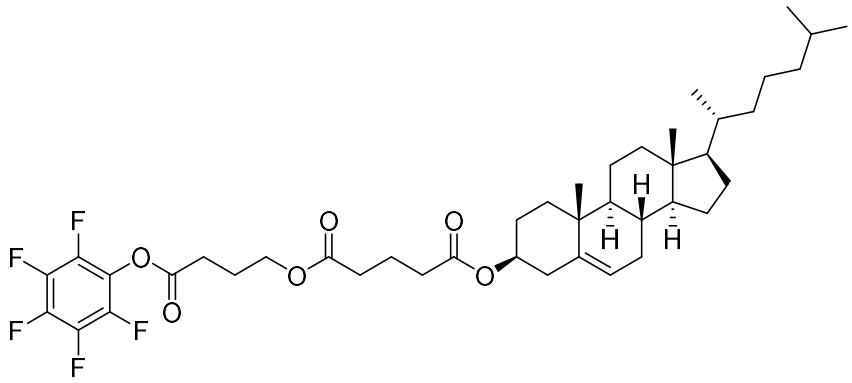
Off-white solid. 67%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 5.38-5.36 (m, 1H), 4.62-4.60 (m, 1H), 4.10 (t, J= 6.5Hz, 2H), 2.39-2.28 (m, 8H), 2.03-1.83 (m, 9H), 1.59-0.85(m, 42H) (includes notable peak at 1.45 (s, 9H)), 0.67 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 173.1, 172.5, 172.3, 139.8, 122.9, 80.7, 74.2, 63.7, 56.9, 56.3, 50.2, 42.5, 39.9, 39.7, 38.3, 37.2, 36.8, 36.4, 36.0, 33.9, 33.4, 32.2, 32.0, 28.9, 28.2, 28.0, 24.5, 24.3, 24.0, 23.0, 22.8, 21.2, 20.4, 19.5, 18.9, 12.0. HRMS calculated for $C_{40}H_{66}O_6Na$ $[M+Na]^+$ 665.4757. Found 665.4753.



4-((5-(cholesteryloxy)-5-oxopentanoyl)oxybutanoic acid 3-23

A solution of **3-22** (321.5mg, 0.50mmol) in DCM (5mL) is deprotected according to general procedure 2 (described in chapter 2). The solvent was then removed in vacuo to obtain the crude oily product, which was redissolved in hexane. The hexane was evaporated in vacuo and the process repeated once more to obtain the crude product as a beige solid. (261mg ,89%).

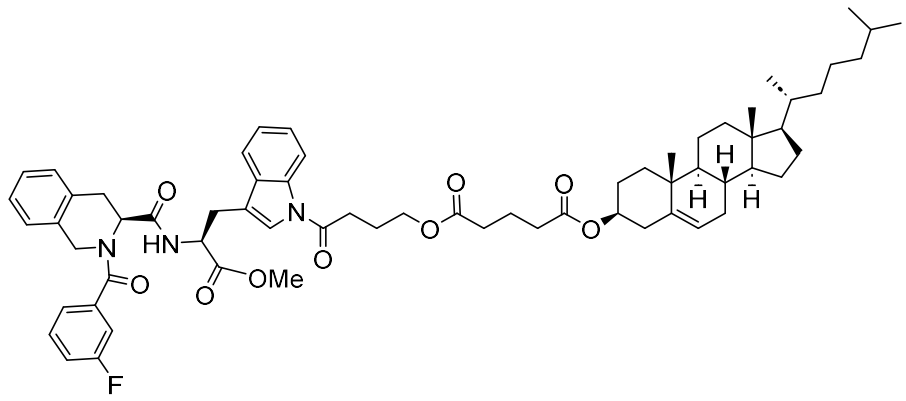
Beige solid. 89%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 5.38-5.36 (m, 1H), 4.64-4.58 (m, 1H), 4.15 (t, J= 6.0Hz, 2H), 2.45 (t, J= 6.0Hz, 2H), 2.37-2.28 (m, 6H), 2.02-1.79 (m, 9H), 1.60-0.85 (m, 33H), 0.67 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 178.5, 173.1, 172.7, 139.7, 122.9, 74.3, 63.5, 56.9, 56.3, 50.2, 49.2, 42.5, 39.9, 39.7, 38.3, 37.1, 36.4, 36.0, 33.8, 33.4, 32.1, 32.0, 30.7, 28.4, 28.2, 27.9, 24.5, 24.0, 23.0, 22.7, 21.2, 20.4, 19.5, 18.9, 12.0. HRMS calculated for $C_{36}H_{58}O_6$ $[M-H]^-$ 585.4155. Found 585.4147.



Cholesteryl (4-oxo-4-(perfluorophenoxy)butyl) glutarate 3-25

3-23 (176.1mg, 0.30mmol) and pentafluorophenol **3-24** (110.5mg, 2eq) was reacted according to general procedure 3. During workup, the organic layer was additionally washed with saturated $NaHCO_3$ to remove excess pentafluorophenol. The residue was purified by column chromatography (10:90 EtOAc:hexane) to obtain the product as a colourless gum (146.8mg, 65%.)

Colourless gum, 65%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 5.37-5.36 (m, 1H), 4.65-4.58 (m, 1H), 4.19 (t, J= 7.0Hz, 2H), 2.78 (t, J= 7.0Hz, 2H), 2.40-2.32 (m, 6H), 2.13-2.07 (m, 2H), 2.03-1.93 (m, 4H), 1.88-1.79 (m, 3H), 1.60-0.85 (m, 33H), 0.67 (s, 3H). ^{19}F NMR (376MHz, $CDCl_3$) δ = -152.6 (d, J= 17.7Hz), -157.7 (t, J= 21.8Hz), -162.1(dt, J= 17.8, 21.8 Hz).



3-14

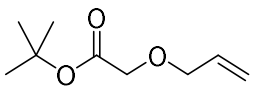
A solution of **3-3e** (30.0mg, 0.060mmol) in freshly distilled dry DME (1.5mL) under an inert atmosphere was cooled to -50°C and to it added potassium bis(trimethylsilyl)amide (1.0M in

THF, 200µL, 3.0eq) solution. The mixture was stirred for 5 minutes at -50°C, and then a solution of **3-25** (66.4mg, 2.0eq) in dry DME (1mL) was added by cannula. Thereafter, the temperature was increased to -20°C and the mixture stirred for 1 hour. After TLC showed completion, the mixture was warmed to room temperature and partitioned between Et_2O (8mL) and 2M HCl (8mL). The layers were separated, and the aqueous layer extracted with Et_2O twice. The combined organic layers were washed with saturated $NaHCO_3$, then brine, then dried with $MgSO_4$. The solvent was evaporated in vacuo to obtain the crude residue, which was purified by column chromatography (eluent 20:80 EtOAc/hexane) to obtain the product as a white solid (25.0mg, 37%).

Note: Due to reproducibility issues of this type of reaction, the data obtained was limited.

White solid, 37%, ^{1}H NMR (400MHz, $CDCl_3$) δ = 8.44 (d, J= 8.0Hz, 1H), 7.63-6.85 (m, 13H), 5.36-5.35 (m, 1H), 5.01-4.95 (2H), 4.60-4.53 (m, 1H), 4.38-4.22 (apparent dd, 2H), 4.12-4.06 (m, 2H), 3.72 (s, 3H), 3.32-2.84 (m, 6H), 2.34-2.27 (m, 6H), 2.06-1.79 (m, 9H), 1.56-0.86 (m, 33H), 0.67 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 173.1, 172.5, 171.9, 170.9, 170.7, 170.4, 139.8, 137.5, 136.0, 134.0, 133.8, 130.6, 130.2, 128.1, 127.1, 125.3, 124.6, 123.8, 122.8, 118.6, 117.6, 117.4, 117.0, 116.2, 114.5, 114.3, 74.2, 63.6, 56.9, 56.3, 54.7, 52.8, 52.2, 50.2, 48.3, 42.5, 39.8, 39.7, 38.2, 37.2, 36,8, 36.4, 36.0, 33.7, 33.4, 32.1, 32.0, 30.9, 30.3, 29.7, 28.4, 28.2, 28.0, 27.7, 24.7, 23.9, 23.7, 22.9, 22.7, 21.4, 20.5, 19.4, 18.9, 12.0. MS-ESI(+): m/z = 1068.31 $[M+H]^+$ HRMS calculated for $C_{65}H_{83}N_3O_9F$ $[M+H]^+$, 1068.6113. Found 1068.6136.

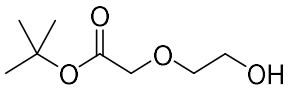
3.5.2 Synthesis of EpMF2-cholesterol conjugate: Towards Indolyl amide linker **3-15**



Tert-butyl 2-(allyloxy)acetate 3-42

A solution of allyl alcohol **3-41** (893.2mg, 1.05mL, 15.38mmol, 1.5eq) in DCM (30mL) was combined with an aqueous solution (30mL) of NaOH (820.4mg, 20.51mmol, 2eq). The biphasic mixture was stirred vigorously and to it added tetra n-butyl ammonium hydrogen sulfate (696.2mg, 0.2eq). The mixture was cooled to 0°C and to it added tert-butyl bromoacetate **3-40** (2000.0mg, 10.25mmol) dropwise. The vigorous stirring of the mixture was continued overnight at room temperature. After TLC (visualised with $KMnO_4$) showed completion, hexane (10mL) was added to the mixture and the layers were separated. The aqueous layer was extracted with Et_2O twice, and the combined organic layers washed with brine, dried with $MgSO_4$, and evaporated gently in vacuo at 30°C to obtain the product as a colourless oil, which was used without further purification. (1235mg ,70%)

Colourless oil, 70%, ^1H NMR (400MHz, $CDCl_3$) δ = 5.96-5.89 (m, 1H), 5.30 (dd, J= 1.4, 17.4Hz, 1H), 5.21 (dd, J= 1.4Hz, 10.4Hz), 3.96 (s, 2H), 1.48 (s, 9H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 169.8, 134.2, 118.1 81.8, 72.4, 67.9, 28.3. FTIR (neat, cm^{-1}) ν_{max} = 3385, 2978, 2934, 1748, 1647, 1369, 1227.

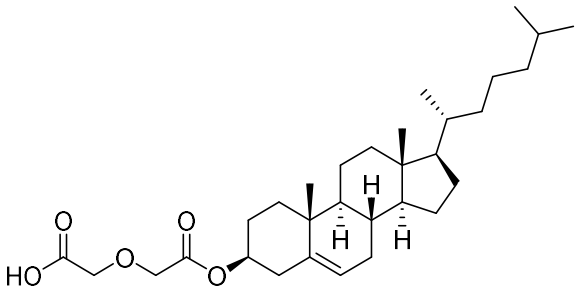


Tert-butyl 2-(2-hydroxyethoxy)acetate 3-43

A solution of **3-42** (776mg, 4.51mmol), in MeOH (25mL) was cooled to -78°C and to it bubbled O_3 gas (through passing pure O_2 through an ozone generator) for 2.5 hours until the solution turned very pale blue. The temperature of the mixture was raised to 0°C and $NaBH_4$ (217.89mg, 9.01mmol, 2eq) was then added slowly portion wise into the mixture to prevent excessive effervescence. The mixture was then stirred for another 2.5 hours at 0°C, then the solvent was removed in vacuo. The residue was again redissolved in EtOAc, and any borate side products were precipitated out, which were then removed by filtration through celite. The

solvent was then again removed in vacuo to obtain the product as a colourless oil. (329mg, 41%).

Colourless oil, 41%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 4.02 (s, 2H), 3.74 (t, J= 4.0 Hz), 3.67 (t, J= 4.0Hz), 1.48 (s, 9H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 170.7, 82.4, 73.7, 69.0, 61.9, 28.3. FTIR (neat, cm^{-1}) ν_{max} = 3404, 2965, 2876, 1744, 1233, 1142.

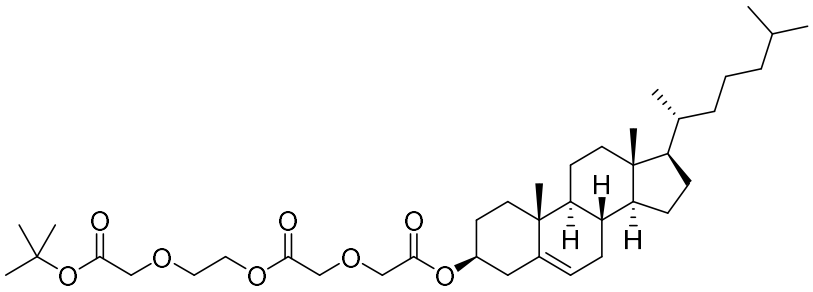


2-(2-cholesteryloxy-2-oxoethoxy)acetic acid 3-45

To a suspension of cholesterol **3-19** 1000mg, 2.59mmol) in dry acetone (20mL) was added diglycolic anhydride **3-44** (600.4mg, 5.17mmol, 2eq), followed by DMAP (31.6mg, 0.259mmol, 0.1eq) and pyridine (409.2mg, 0.44mL, 5.17mmol, 2eq). The whole mixture was heated to reflux for 24 hours. When TLC showed completion, the mixture was cooled to room temperature, and any precipitate (which was the diester side product) was removed by filtration. The solvent was evaporated in vacuo, and the residue redissolved in EtOAc (20mL) This was followed by washing with 2M HCl followed by brine, and subsequent drying with $MgSO_4$. The solvent was again evaporated in vacuo to obtain the crude product as a brown solid. The solid was purified by recrystallisation with hot ethanol, then cooling to -20°C and isolation of the precipitated white solid. The mother liquor was further evaporated until more product is precipitated, before storing again at -20°C and isolating a second batch of white solid. (Combined 951mg, 74%)

White solid. 74%. Mp 158-161°C ^{1}H NMR (400MHz, $CDCl_3$) δ = 5.41-5.39 (m, 1H), 4.77-4.70 (m, 1H), 4.25 (s, 2H), 4.23 (s, 2H), 2.39-2.30 (m, 2H), 2.04-1.78 (m, 5H), 1.67-0.85 (m, 33H), 0.68 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 171.9, 170.5, 139.2 123.5, 76.1, 69.7, 69.6,

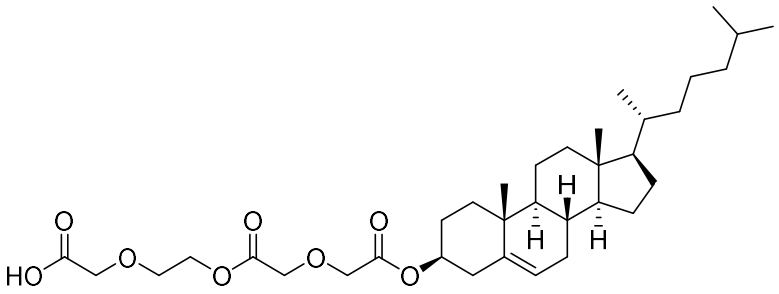
50.2, 42.5, 39.9, 39.7, 38.1, 37.0, 36.7, 36.4, 36.0, 32.1, 32.0, 28.4, 28.2, 27.9, 24.5, 24.0, 23.0, 22.7,21.2, 19.5, 18.0, 12.0. MS-ESI(-): m/z = 547.16 $[M+formate]^-$. HRMS calculated for $C_{31}H_{50}O_5$ $[M-H]^-$ 501.3580. Found 501.3586.



2-(2-(tert-butoxy)-2-oxoethoxy)ethyl 2-((2-cholesteryloxy)-2-oxoethoxy)acetate 3-46

Carboxylic acid **3-45** (500mg, 0.995mmol) and alcohol **3-43** (192.8mg, 1.1eq) were reacted according to general procedure 3. The residue was purified by column chromatography (gradient eluent 20:80 to 50:50 EtOAc:hexane) to obtain the product as a colourless gummy solid. (366mg, 56%)

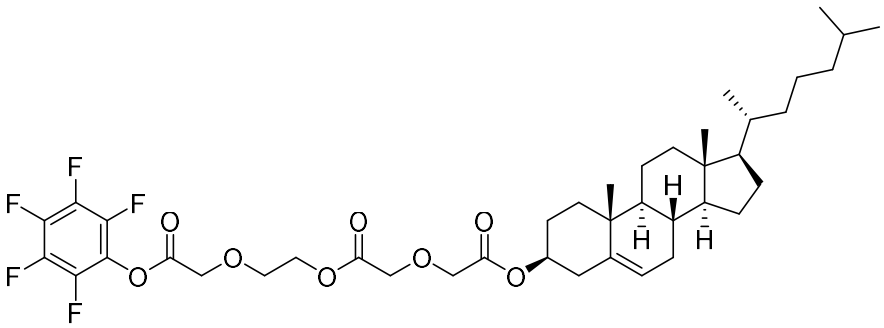
Colourless solid, 56%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 5.38-5.37 (m,1H), 4.75-4.66 (m, 1H), 4.35 (t, J= 3.2Hz, 2H), 4.29 (s, 2H), 4.21 (s, 2H), 3.99 (s, 2H), 3.78 (t, J= 3.2Hz, 2H). 2.34 (apparent d, 2H), 2.07-1.78 (m, 5H), 1.66-0.85 (m, 42H) (includes prominent peak at 1.48 (s, 9H)), 0.68 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 170.0, 169.4, 169.3, 139.5, 123.1, 82.0, 75.1, 69.2, 69.0, 68.5, 68.2, 64.1. 56.9, 56.3, 50.2, 42.5, 39.9, 39.7, 38.3, 37.2, 36.8, 36.4, 36.0, 32.2, 28.3, 28.2, 27.7, 24.4, 24.123.0, 22.7, 20.9, 19.5, 18.9, 12.0. FTIR (DCM, cm^{-1}) ν_{max} = 2936, 2904, 2401, 1748, 1368, 1207, 1136. ESI-MS(+): m/z = 683.57 $[M+Na]^+$. HRMS calculated for $C_{39}H_{64}O_8Na$ $[M+Na]^+$ 683.4499. Found 683.4482.



2-(2-(2-(2-((cholesteryloxy-2-oxoethoxy)acetoxy)ethoxy)acetic acid 3-47

A solution of **3-46** (337mg, 0.5mmol) in DCM (5mL) was deprotected according to general procedure 2. The solvent was removed in vacuo to obtain the crude oily product, which was redissolved in hexane. The hexane was evaporated in vacuo and the process repeated once more to obtain the crude product as a beige solid. (280mg, 91%)

Beige solid, 91%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 5.39-5.38 (m, 1H), 4.72-4.67 (m, 1H), 4.37 (t, J= 4.8Hz, 2H), 4.28 (s, 2H), 4.20 (s, 2H), 4.16 (s, 2H), 3.82 (t, J= 4.8Hz, 2H), 2.34 (apparent d, 2H), 2.03-1.79 (m, 5H), 1.61-0.85 (m, 33H), 0.68 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 173.3, 170.0, 169.5, 139.4, 123.2, 75.3, 69.7, 68.5, 68.3, 68.2, 63.9, 60.7, 50.2, 42.5, 39.9, 39.8, 39.7, 38.1, 37.1, 36.7, 36.4, 36.0, 32.1, 32.0, 28.4, 28.2, 27.9, 24.5, 24.0, 23.0, 22.7, 21.2, 19.5, 18.9, 12.0.FTIR (neat, cm^{-1}) ν_{max} = 3437, 2936, 2905, 2359, 1744, 1672, 1217, 1138. ESI-MS (-): m/z = 603.47 $[M-H]^-$ HRMS calculated for $C_{35}H_{55}O_8$ $[M-H]^-$ 603.3897. Found 603.3902.

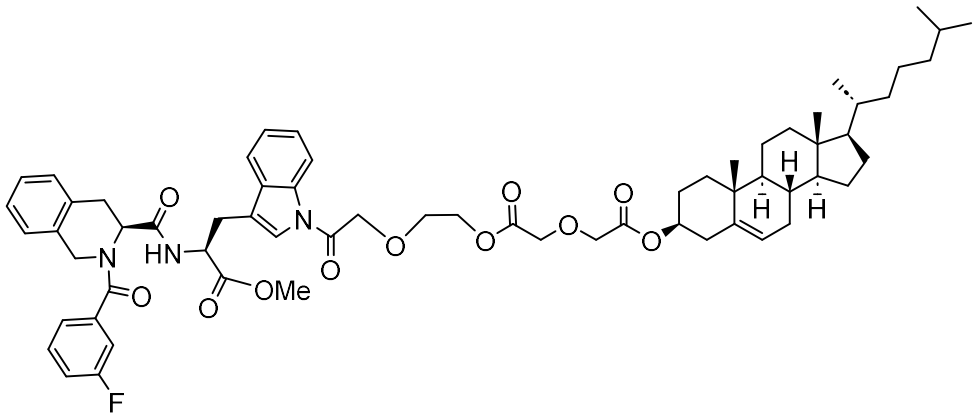


cholesteryl 2-(2-oxo-2-(2-(2-oxo-2-(perfluorophenoxy)ethoxy)ethoxy)ethoxy)acetate 3-48

3-47 (280.0mg, 0.463mmol) was reacted with pentafluorophenol **3-24** (170.4mg, 2eq) according to general procedure 3. During workup, the organic layer was additionally washed

with saturated $NaHCO_3$ to remove excess pentafluorophenol. The residue was purified by column chromatography (20:80 EtOAc:hexane) to obtain the product as a colourless gum (142mg, 40%).

Colourless gum, 40%. ^1H NMR (400MHz, $CDCl_3$) δ = 5.38-5.37 (m, 1H), 4.72-4.65 (m, 1H), 4.50 (s, 2H), 4.40 (t, J= 4.6Hz, 2H), 4.29 (s, 2H), 4.20 (s, 2H), 3.89 (t, J= 4.6Hz), 2.33 (d, J= 7.8Hz, 2H), 2.04-0.85 (m, 38H), 0.68 (s, 3H). 19F NMR (376MHz, $CDCl_3$) δ = -152.3 (dt, J= 17.3 Hz, 2F), -157.0 (t, J= 21.8Hz, 1F), -161.6 (t, J= 17.3, 21.8Hz, 2F).



3-15

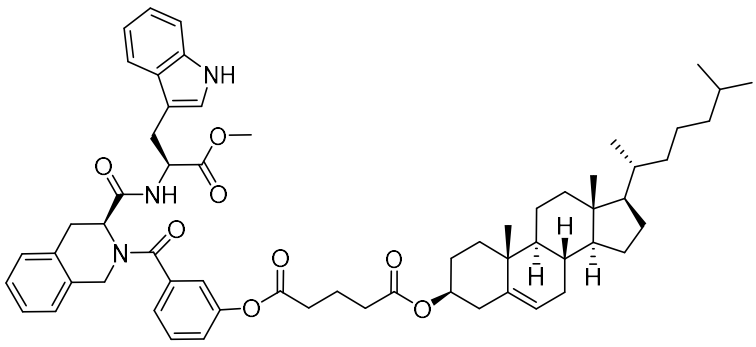
A solution of **3-3e** (50.0mg, 0.100mmol) in freshly distilled dry DME (1.5mL) under an inert atmosphere was cooled to -50°C and to it added potassium bis(trimethylsilyl)amide (1.0M in THF, 300μL, 3.0eq) solution. The mixture was stirred for 5 minutes at -50°C, and then a solution of **3-48** (154.3mg, 2.0eq) in dry DME (1mL) was added by cannula. Thereafter, the temperature was increased to -20°C and the mixture stirred for 1 hour. After TLC showed completion, the mixture was warmed to room temperature and partitioned between Et_2O (8mL) and 2M HCl (8mL). The layers were separated, and the aqueous layer extracted with Et_2O twice. The combined organic layers were washed with saturated $NaHCO_3$, then brine, then dried with $MgSO_4$. The solvent was evaporated in vacuo to obtain the crude residue, which

was purified by column chromatography (eluent 307 EtOAc/hexane) to obtain the product as a white solid (14.0mg, 13%), along with some inseparable **3-47** side product.

Note: Due to reproducibility issues of this type of reaction, the data obtained was limited.

Off white solid, 13%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 8.44-8.42 (m, 1H), 8.07-8.04 (m, 1H), 7.55-6.88 (m, 13H), 5.39-5.37 (m, 1H), 4.96-4.88 (m, 2H), 4.70-4.18 (m, 11H), 3.72-3.65 (m, 5H), 3.37-3.17 (m, 5H), 2.34-2.32 (apparent d, 2H), 2.09-1.83 (m, 5H), 1.56-0.86 (m, 31H), 0.68 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) (poor signal noise ratio means not all peaks were identifiable) δ = 170.0, 169.7, 139.6, 127.2, 125.5, 123.1, 118.7, 117.0, 75.0, 68.5, 68.2, 50.2, 42.5, 39.9, 39.7, 38.2, 37.1, 36.8, 36.4, 36.0, 32.1, 32.0, 29.9, 28.2, 27.9, 24.5, 24.0, 23.0, 22.8, 21.2, 19.5, 18.9, 12.0.

3.5.3 Synthesis of EpMF2-cholesterol conjugate: phenyl ester **3-50**



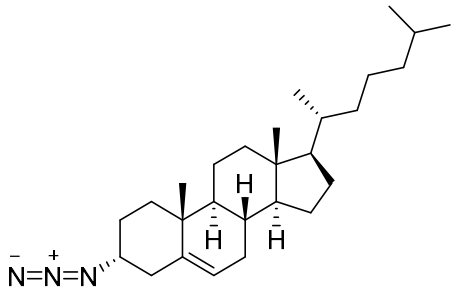
3-50

3-3f (0.402mmol, 200.0mg) and **3-21** (1.1eq, 221.4mg) were reacted according to general procedure 3. The product was purified by column chromatography (Eluent 40:60 EtOAc: hexane) to obtain product L3 as an off-white solid. (214mg, 54%).

Off white solid, 54%. Mp: 116-120°C ^{1}H NMR (400MHz, $CDCl_3$) δ (complex mixture of rotamers) = 8.36 (s, 1H), 7.54-6.79 (m, 13H), 5.38-5.37 (m, 1H), 5.16 (apparent t, 1H), 4.90-4.88 (m, 1H), 4.66-4.60 (m, 1H), 4.33 (apparent d, 0.5H), 3.98 (apparent d, 0.5H), 3.71 (s,

3H),3.42-3.04 (m, 4H), 2.70 (apparent t, 2H), 2.46 (m, 2H), 2.43 (apparent d, 2H), 2.09-1.79 (m, 7H), 1.59-0.85 (m, 35H), 0.67 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 172.4, 171.9, 170.9, 170.2, 150.8, 139.8, 136.6, 136.4, 133.9, 133.2, 131.9, 129.9, 128.5, 127.9, 127.5, 126.7, 125.5, 124.8, 123.6, 123.0, 122.3, 121.2, 119.8, 118.8, 111.5, 109.8, 74.5, 56.9, 56.4, 53.7, 52.8, 52.6, 48.0, 42.5, 39.9, 39.7, 38.3, 37.2, 36.8, 36.3, 36.0, 29.3, 28.4, 28.2, 28.0, 27.7, 24.5, 24.0, 23.0, 22.7, 21.2, 20.3, 19.5, 18.9, 12.1. HRMS calculated for $C_{61}H_{78}N_3O_8$ $[M+H]^+$ 980.5789. Found 980.5797.

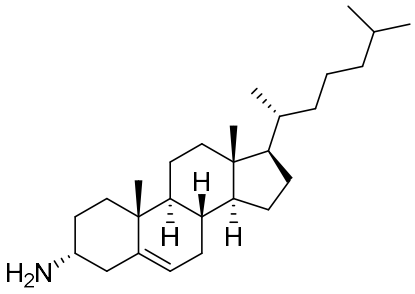
3.5.4 Synthesis of EpMF2-cholesterol conjugate: phenyl ester **3-51**



3α-azidocholest-5-ene 3-55

To a solution of cholesterol **3-19** (1000mg, 2.59mmol) in freshly distilled dry toluene (20mL) was added triphenylphosphine (2eq, 1357mg) and $Zn(N_3)_2(pyridine)_2$ (1eq, 857.7mg) and the suspension stirred rapidly. The mixture was then cooled to 0°C and to it added DIAD (2eq, 1046mg, 1.02mL) dropwise. After that, stirring continued for 2h at room temperature. When TLC indicated completion, the heterogenous mixture was filtered through a pad of celite. The solvent was evaporated in vacuo and the crude compound purified by column chromatography (Eluent 100% hexane) to obtain the compound as a white solid. (794mg, 75%). There is also a small amount of inseparable non-polar impurity, tentatively identified as the rearrangement product 6β-azido-3α,5-cyclo-5α-cholestane **3-55a** (see section 3.3.8) arising from neighbouring group participation of the double bond.

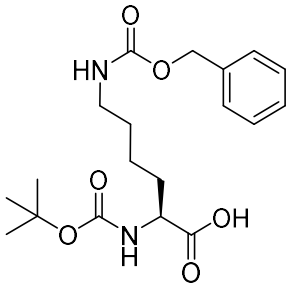
White solid, 75%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 5.41-5.39 (m, 1H), 3.88-3.87 (m, 1H), 2.54-2.50 (m, 1H), 2.21-2.17 (m, 1H), 2.03-0.86 (m, 38H), 0.68 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 138.3, 123.4, 58.5, 56.9, 56.3, 50.1, 42.5, 39.9, 39.7, 37.3, 36.4, 36.0, 33.8, 32.0, 28.4, 28.2, 26.3, 24.5, 24.0, 23.0, 22.7, 20.9, 18.9, 12.0. ESI-MS(+): m/z = 434.11 $[M+Na]^+$. HRMS calculated for $C_{27}H_{48}N$ $[M-N_2+H_3]^+$, 386.3787. Found 386.3792. (No peak corresponding to M was found. Presumably, the azide underwent fragmentation with expulsion of N_2 followed by protonation of the remaining N atom to the amine)



3α-amino-cholest-5-ene 3-56

A solution of **3-55** (400.0mg, 0.972mmol) in anhydrous Et_2O (10mL) was stirred in an inert atmosphere at 0°C. $LiAlH_4$ (368.8mg, 10eq) was added in portions slowly enough to prevent excess effervescence. After addition, the mixture was then stirred vigorously for 4h at room temperature. When TLC indicated completion, the mixture was cooled to 0°C and to it added bench Et_2O (10mL) followed by careful dropwise addition of H_2O (5mL), causing more effervescence. When no more bubbles are formed, 2M NaOH solution was added and stirred until the aluminate precipitate turned white and clumped up. The precipitate was removed by decantation and the two layers separated. The aqueous layer was extracted with Et_2O twice and the combined organic extracts washed with brine, dried with $MgSO_4$, and evaporated in vacuo to give the crude compound as a white solid, which was used without further purification. (208mg, 56%).

White solid, 56%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 5.36-5.35 (m, 1H), 3.16 (br s, 1H), 2.59-2.56 (apparent d, 1H), 2.04-0.85 (m, 43H) 0.68 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 139.2, 123.4, 57.0, 56.3, 50.7, 47.1, 42.5, 40.3, 40.0, 39.7, 37.6, 36.4, 36.0, 33.3, 32.2, 32.0, 29,8, 28.4, 28.2, 24.5, 24.0, 23.0, 22.7, 21.0, 19.0, 18.9,12.0. ESI-MS(+): m/z = 386.39 $[M+H]^+$. HRMS calculated for $C_{27}H_{48}N$ $[M+H]^+$, 386.3787. Found 386.3788.

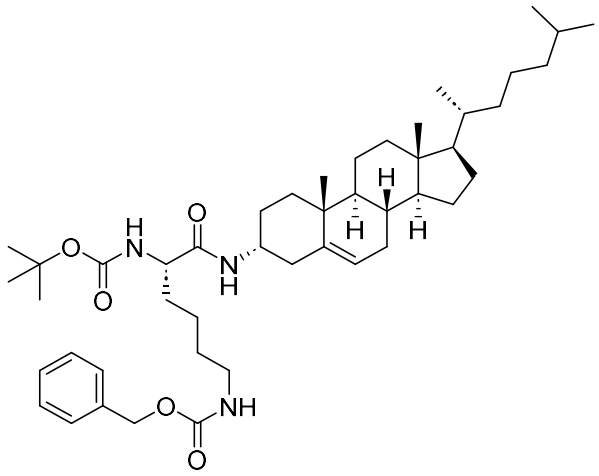


$N_ε$-((benzyloxy)carbonyl)-$N_α$-(tert-butoxycarbonyl)-L-lysine 3-58

A suspension of $N_ε$-(benzyloxycarbonyl)-L-lysine **3-57** (500.0mg, 1.78mmol) in a 1:1 THF/H_2O solvent mixture (15mL) was cooled to 0°C. To the mixture was added NaOH (157.0mg, 2.2eq) and stirred until the solid dissolves into a clear solution, followed by addition of di-tert butyl dicarbonate (467.1mg, 0.49mL, 1.2eq). The solution was further stirred overnight at room temperature. After TLC showed completion, the mixture was then neutralised with addition of saturated $NaHSO_4$ solution until the pH is around 4. The mixture was then extracted with EtOAc 3 times. The combined organic layers were then washed with brine and dried with anhydrous $MgSO_4$. The solvent was then removed in vacuo to obtain the crude product as a colourless gummy solid. (678mg, quantitative yield) and used without further purification.

Colourless gum, 99%. ^{1}H NMR (400MHz, $CDCl_3$) (mixture of rotamers) δ = 7.33-7.30 (m, 5H), 6.27 (br s, 0.6H), 5.24-4.94 (m, 3.4H), 4.28 (br s, 0.7H), 4.10 (br s, 0.3H), 3.20-3.17 (m, 2H), 1.84-1.66 (m, 2H), 1.53-1.43 (m, 13H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 176.4, 156.9,

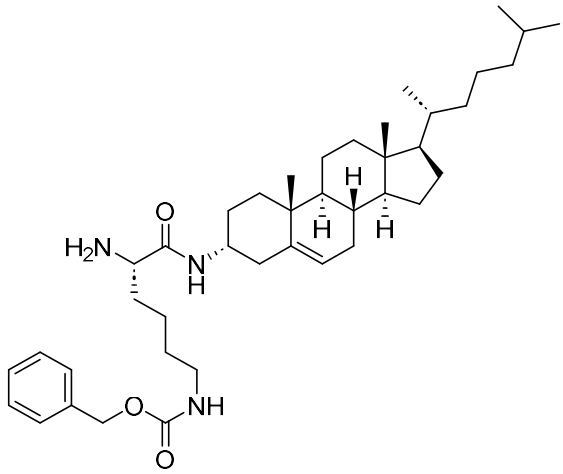
156.0, 136.7, 128.7, 128.3, 128.1, 80.4, 66.9, 53.3, 40.7, 32.0, 29.5, 29.4, 28.5, 22.5. HRMS calculated for $C_{19}H_{28}N_2O_6Na$ $[M+Na]^+$, 403.1845. Found 403.1835.



benzyl tert-butyl (((S)-6-(cholest-5-enyl)amino)-6-oxohexane-1,5-diyl)dicarbamate 3-59

3-58 (84.1mg, 0.519mmol) and **3-56** (197.3mg 1,0eq) were reacted under general conditions 1. (described in chapter 2) (The residue was purified by column chromatography to obtain the product as an off-white gummy solid. (268mg, 69%).

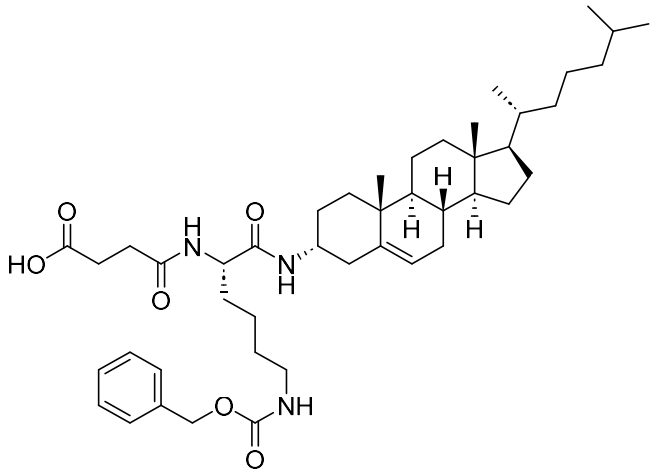
Off white gum, 69%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 7.36-7.29 (m, 5H), 6.08-6.06 (m, 1H), 5.41-5.39 (m, 1H), 5.09 (s, 2H), 4.82 (br s, 1H), 3.99-.3.96 (m, 1H),3.20-3.16 (apparent q, 2H), 2.60-2.56 (apparent d, 1H), 2.04-0.86 (m, 58H) (including notable peak at 1.44 (s, 9H)), 0.67 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 171.0, 156.7, 138.8, 136.8, 128.7, 128.3, 124.0, 66.8, 56.9, 56.4, 42.5, 39.7, 37.6, 36.4, 36.0, 32.0, 28.6, 28.2, 24.1, 23.0, 22.7, 19.0, 18.9, 12.0. ESI-MS (+): m/z = 748.52 $[M+H]^+$. HRMS calculated for $C_{46}H_{74}N_3O_5$ $[M+H]^+$, 748.5628. Found 748.5613.



benzyl ((S)-5-amino-6-(cholest-5-enyl)amino)-6-oxohexyl)carbamate 3-60

To a solution of **3-59** (180.0mg, 0.241mmol), in DCM (3mL), was deprotected using general procedure 2. After TLC showed completion of reaction, the solvent was washed with 2M NaOH 2 times, followed by brine 1 time. The solution was then dried with $MgSO_4$ and evaporated in vacuo to obtain the product as an off-white solid as freeform base. (171mg, 93%).

Off white solid, 93%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 7.35-7.20 (m, 5H), 5.38-5.37 (m, 1H), 5.08 (s, 2H), 4.89 (br s, 1H), 4.13-4.09 (m, 0.8H), 3.81-3.80 (m, 0.2H), 3.20-3.19 (m, 2H), 2.75-2.53 (apparent d, 1H), 2.04-0.86 (m, 48H), 0.67 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 156.7, 136.8, 128.7, 128.2, 123.7, 66.8, 60.6, 56.9, 56.4, 50.8, 42.5, 40.8, 40.0, 39.7, 37.5, 37.2, 36.4, 36.0, 34.6, 32.2, 32.0, 30.1, 28.4, 28.2, 26.6, 24.5, 24.1, 23.0, 22.7, 19.1, 12.0. HRMS calculated for $C_{41}H_{66}N_3O_3$ $[M+H]^+$, 648.5104. Found 648.5112.

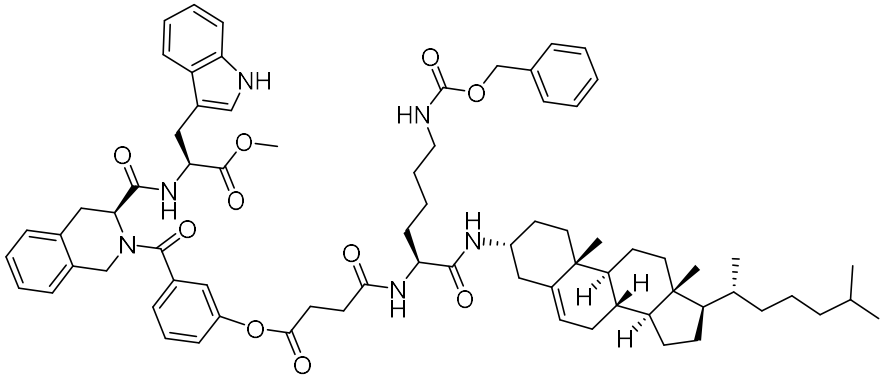


4-(((S)-6-(((benzyloxy)carbonyl)amino)-1-((((cholest-5-enyl)amino)-1-oxohexan-2-yl)amino)-4-oxobutanoic acid 3-61

To a solution of **3-60** (150.0mg, 0.197mmol), in DCM (3mL) was added succinic anhydride **3-16** (29.6mg, 1.5eq), DIPEA (56.0mg, 75µL, 2.2eq), and DMAP (4.8mg, 0.1eq), and the mixture stirred for 3 hours. When TLC showed completion, 2.0M HCl was added (3mL), the mixture shaken, and the layers separated. The aqueous layer was extracted with DCM twice, and the combined organic extracts washed with brine, dried in $MgSO_4$, and evaporated in vacuo to obtain the crude product as an off-white foamy solid (241mg, 97%), which was used without further purification.

Off white solid, 97%. ^{1}H NMR (400MHz, $CDCl_3$) (complex mixture of rotamers) δ = 7.35-7.30 (m, 5H), 6.84-6.55 (m, 1.2H), 6.45-6.02 (m, 0.8H), 5.39-5.37 (m, 0.5H), 5.14-5.08 (m, 2.5H), 4.40-3.67 (m, 4H), 3.36-3.10 (m, 2H), 2.90-2.35 (m, 4H), 2.04-0.85 (m, 45H), 0.67 (m, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 172.6, 171.1, 170.8, 158.9, 157.0, 141.1, 138.6, 138.5, 136.6, 136.2, 128.7, 128.4, 128.3, 128.1, 124.3, 72.0, 67.5, 67.0, 56.9, 56.3, 53.4, 53.0, 50.8, 50.3, 47.6, 46.2, 46.0, 42.5, 40.8, 39.8, 39.7, 37.5, 37.0, 36.7, 36.4, 36.0, 34.4, 32.1, 32.0, 31.8, 31.1, 29.9, 29.7, 28.6, 28.4, 28.2, 26.2, 24.5, 24.0, 23.0, 22.7, 22.5, 22.0, 21.2, 20.9, 19.0, 18.9,

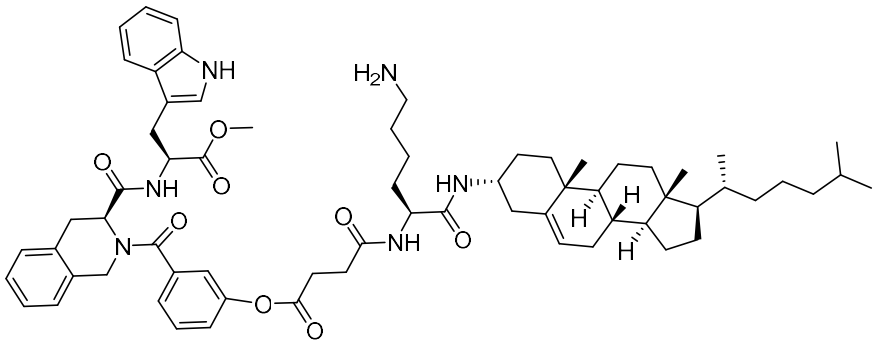
12.0. ESI-MS(-): m/z = 746.83 $[M-H]^-$. HRMS calculated with $C_{45}H_{70}N_3O_6$ $[M+H]^+$, 748.5265. Found 748.5280.



3-62

3-61 (230.0mg, 0.308mmol) was reacted with **3-3f** (0.308mmol, 153.0mg) according to general procedure 3. The crude product was purified by column chromatography (75:25 EtOAc/hexane) to obtain the product as a white foamy solid. (110mg, 29%).

White solid, 29%. ^{1}H NMR (400MHz, $CDCl_3$)(mixture of rotamers) δ = 8.99 (br s, 0.7H), 8.83(br s, 0.3H). 7.52-6.52(m, 18H), 5.93-5.91(m, 1H), 5.34-5.33(m, 1H), 5.21-5.19 (m, 0.5H), 5.08-5.06 (m, 2.5H), 4.88-4.85, (m, 1H), 4.34-4.05 (m, 1H), 3.73 (s, 3H), 3.38-3.35 (m, 1H), 3.24-3.18 (apparent dd, 1H), 3.03-2.53 (m, 7H), 2.03-0.86 (m, 53H), 0.66 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 172.6, 172.4, 171.8, 171.5, 170.7, 170.3, 156.8, 150.8, 138.5, 136.7, 136.5, 133.7, 132.9, 131.8, 129.7, 128.4, 128.2, 128.1, 127.7, 127.4, 126.6, 125.5, 124.6, 125.1, 123.7, 123.6 122.0, 121.2 119.5, 118.6, 111.7, 109.3, 66.7, 56.8, 56.3, 53.4, 52.6, 50.7, 47.7, 46.1, 42.4, 40.4, 39.8, 39.6, 37.4, 36.8, 36.3, 35.9, 34.3, 32.3, 32.0, 31.9, 30.7, 29.7, 29.5, 28.3, 28.1, 26.1, 24.4, 24.0, 23.0, 22.7, 22.5, 20.8, 18.9, 18.8, 12.0. ESI-MS(+): m/z = 1227.36 $[M+H]^+$. HRMS calculated for $C_{74}H_{95}N_6O_{10}$ $[M+H]^+$, 1227.7110. Found 1227.7134.

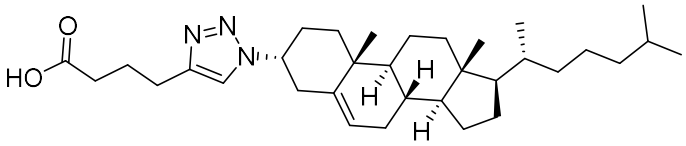


3-51

A solution of **3-62** (100.0mg, 0.0815mmol) in methanol (2mL) was flushed with an inert gas for 5 minutes. It was then transferred to a high-pressure reactor with 30.0mg palladium on carbon (30% by mass) and stirred under 60psi of H_2 for 6 hours. After that, the vessel was again flushed with an inert gas, and the mixture filtered through celite to remove the palladium. The solvent was then evaporated in vacuo to obtain the product as an off-white solid. (86.0mg, 97%).

Off-white solid, 97%. ^{1}H NMR (400MHz, $CDCl_3$)(mixture of rotamers) δ = 9.12 (br s, 0.6H), 8.63 (br s, 0.4H), 7.47-6.30 (m, 13H), 5.37-5.36 (m, 1H), 4.96-4.26 (m, 4H), 4.15-4.05 (m, 2H), 3.72 (s, 3H), 3.29-2.44 (m, 11H), 2.10-0.86 (m, 50H), 0.68 (m, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 174.4, 173.5, 172.6, 172.4, 172.0, 171.8, 171.2, 171.0, 170.8, 157.8, 138.7, 136.4, 136.1, 133.9, 133.5, 132.0, 131.5, 130.6, 130.0, 128.2, 127.9, 127.6, 127.2, 127.0, 126.8, 126.5, 125.6, 124.0, 123.8, 123.3, 122.6, 122.1, 114.4, 111.9, 109.2, 108.9, 56.9, 56.3, 54.3, 53.8, 53.4, 52.8, 52.6, 52.0, 50.7 (2 overlapping peaks), 50.4, 48.3, 46.1, 42.5, 39.9, 39.7, 37.5, 37.0, 36.4, 36.0, 34.3, 32.1, 31.9, 30.8, 30.3, 29.3, 28.4, 28.2, 27.3, 27.1, 26.4, 26.2, 24.5, 24.1, 23.0, 22.7, 20.9, 19.0, 18.9, 12.0. ESI-MS(+): m/z = 1093.46 $[M+H]^+$. HRMS calculated for $C_{66}H_{89}N_6O_8$ $[M+H]^+$, 1093.6742. Found 1093.6766.

3.5.5 Synthesis of EpMF2-cholesterol conjugate: phenyl ester **3-52**

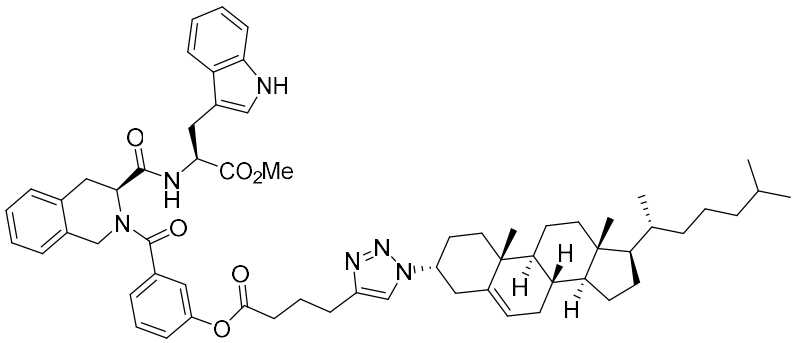


4-(1-(3α-cholest-5-eneyl)-1H-1,2,3-trizol-4-yl)butanoic acid 3-65

Azide **3-55** (244.0mg, 0.593mmol), DIPEA (153.2mg, 0.21mL, 2.0eq), and hex-5-ynoic acid **3-64** (66.5mg, 1.0eq) were dissolved in DCM (3mL). In another vessel, copper (I) iodide (11.3mg, 0.1eq) and triphenylphosphine (46.4mg, 0.3eq) was added to MeCN (3mL) and stirred until complete dissolution. The 2 solutions were combined, and the combined mixture flushed with inert gas for 30 mins. After that, the mixture was left to stir under an inert atmosphere overnight.

When TLC showed completion, the mixture was partitioned between EtOAc (15mL) and 2M HCl (15mL). The layers were separated, and the aqueous layer extracted with EtOAc twice. The combined organic layers were decolourised by shaking with a dilute solution of aqueous NaS_2O_3, (100mg in 10mL) then washed with brine and dried with $MgSO_4$. The solvent was evaporated in vacuo to obtain the crude mixture, which was purified by column chromatography to obtain the product as an off-white solid (130.4mg, 42%)

Off white solid, 42% ^{1}H NMR (400MHz, $CDCl_3$) δ = 7.56 (s, 1H), 5.49-5.48 (m, 1H), 4.89-4.88 (m, 1H), 2.97-2.93 (app d), 2.79 (t, J= 9.2 Hz, 2H), 2.51-0.85 (m, 43H) 0.68 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 176.6, 138.3, 132.3, 128.7, 124.5, 56.7, 56.3, 50.2, 42.5, 39.7 (2 overlapping peaks), 37.2, 36.3, 36.0, 35.7, 33.0, 32.1, 31.9, 28.4, 28.2, 27.3, 24.9, 24.4, 24.0, 23.0, 22.7, 20.8, 19.4, 19.1, 18.9, 12.0. ESI-MS(+): m/z = 524.32 HRMS calculated for $C_{33}H_{52}N_3O_2$ $[M-H]^-$ 522.4060. Found 522.4052.



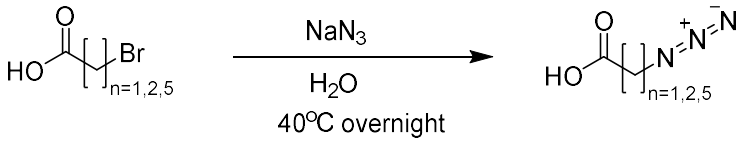
3-52

3-3f (0.221mmol, 110.0mg) and **3-65** (1.1eq, 127.4.4mg) were reacted according to general procedure 3. The product was purified by column chromatography (Eluent 50:50 EtOAc: hexane) to obtain product as an off-white solid. (128mg, 58%).

Off white solid, 58%. ^{1}H NMR (400MHz, $CDCl_3$) (complex mixture of rotamers) δ = 8.62 (br s, 1H), 7.56-6.97 (m, 13H), 5.46-5.45 (m, 1H), 5.16-5.13 (m, 1H). 4.52-3.96 (m, 4H), 3.69-3.59 (m, 1H), 3.35-2.45 (m, 11H), 2.15-0.85 (m, 42H), 0.65 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 175.7, 172.5, 170.4, 150.7, 137.2, 136.4, 133.8, 132.3, 132.2, 129.9, 128.8, 128.7, 128.4, 127.9, 127.4, 127.1, 126.7, 125.5, 125.3, 123.8 (br s), 123.3, 122.4, 122.1, 121.2, 120.6 (br s), 119.5, 118.6, 118.2, 111.8, 111.6, 109.4, 58.9, 56.7, 56.6, 56.3, 53.9, 52.8, 52.7, 52.6, 50.1, 48.1, 42.4, 39.6, 38.6, 37.2, 36.3, 35.9, 35.3, 33.5, 33.3, 33.2, 32.9, 32.0,31.8, 29.7, 28.3, 28.1, 27.6, 27.0, 26.8, 25.5, 24.9, 24.3, 24.0, 23.4, 23.0, 22.7, 21.1, 20.8, 19.4, 18.8, 12.0. HRMS calculated for $C_{62}H_{79}N_6O_6$ $[M+H]^+$ 1003.6061. Found 1003.6089.

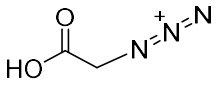
3.5.6 Synthesis of EpMF2-cholesterol conjugate: phenyl estesrs **3-53a-e**

General procedure 4: Azide substitution of bromo-substituted carboxylic acid



To a solution of the bromo-substituted carboxylic acid (1eq) in water (0.5mL per 100mg) was added NaN_3 (5eq). The mixture was then stirred overnight at 40°C. After that, the mixture was

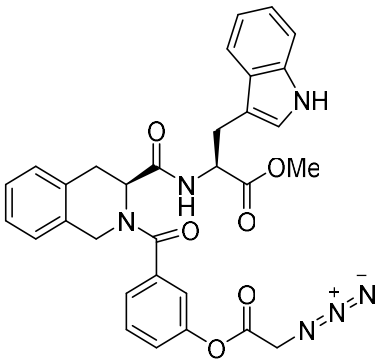
acidified by concentrated HCl to around pH 1 (Caution: formation of hydrazoic acid) and extracted with Et_2O three times. The combined organic extracts were washed with brine, dried with MgSO4, and evaporated in vacuo (the water bath temperature of the rotavap not set above 30°C) to obtain the product.



2-azidoacetic acid 3-73a

2-bromoacetic acid **3-72a** (500.0mg, 3.60mmol) was reacted according to general procedure 4. The product was obtained as a colourless oil. (137mg, 45%).

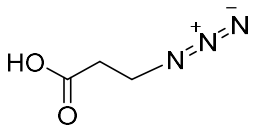
Colourless oil, 45%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 3.96 (s, 2H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 172.7, 50.1. FTIR (DCM, cm^{-1}) V_{max} = 3474, 2116, 1732.



methyl ((S)-2-(3-(2-azidoacetoxy)benzoyl)-1,2,3,4-tetrahydroisoquinoline-3-carbonyl)-L-tryptophanate 3-76a

Azide **3-73a** (44.7mg, 0.442mol, 1.1eq), was reacted with **3-11f** (200.0mg, 0.402mmol) according to general procedure 3. The compound was purified by column chromatography (Eluent 50:50 EtOAc/hexane) to obtain the product as an off-white foamy solid. (123mg, 53%). Off-white solid, 53%. 1H NMR (400MHz, $CDCl_3$)(complex mixture of rotamers) δ = 8.19 (br s, 1H), 7.55-6.32 (m, 13H), 5.37-4.59, (m, 3H), 4.47-4.09 (m, 4H), 3.70 (s, 3H), 3.38-3.04 (m,

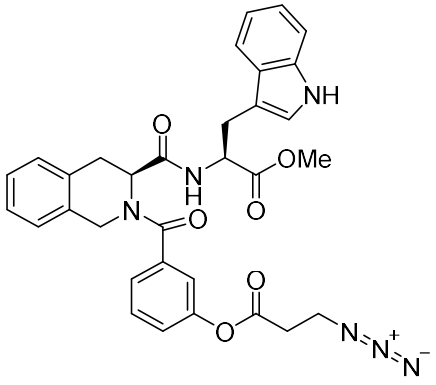
4H). ^{13}C NMR (100MHz, CDCl3) δ = 172.4, 170.5, 170.2, 169.6, 167.2, 150.1, 136.9, 136.3, 133.8, 133.3, 131.7, 130.4, 130.2, 128.4, 128.0, 127.6, 127.2, 127.0, 126.8, 125.5, 125.2, 123.5, 123.2, 122.3, 120.7, 120.1, 119.8, 118.7, 118.4, 111.7, 111.4, 109.8, 109.5, 53.9, 52.9, 52.6, 50.6, 48.1, 31.7, 29.7, 27.6, 27.1, 25.1, 14.4. ESI-MS(+): m/z = 581.03 $[M+H]^+$. HRMS calculated for $C_{31}H_{29}N_6O_6$ $[M+H]^+$ 581.2149. Found 581.2164.



3-azidopropanoic acid 3-73b

3-choloropropanoic acid **3-72b** (500.0mg, 4.61mmol) was reacted according to procedure 4. The product was obtained as a colourless oil. (295mg, 56%).

Colourless oil, 56%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 3.59 (t, J= 6.4Hz, 2H), 2.64 (t, J= 6.4Hz, 2H). FTIR (DCM, cm^{-1}) ^{13}C NMR (100MHz, $CDCl_3$) δ = 176.2, 46.7, 33.9. V_{max} = 3443, 2104, 1715.

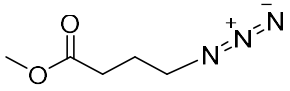


methyl ((S)-2-(3-((3-azidopropanoyl)oxy)benzoyl)-1,2,3,4-tetrahydroisoquinoline-3-carbonyl)-L-tryptophanate 3-76b

Azide **3-73b** (50.9mg, 0.442mmol, 2eq) was reacted with **3-3f** (200.0mg, 0.402mmol) according to general procedure 3. The reaction time was extended to 4 hours for this reaction.

The compound was purified by column chromatography (Eluent 50:50 EtOAc/hexane) to obtain the product as an off-white foamy solid. 82mg, 34%).

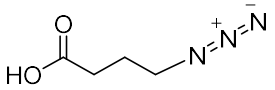
Off-white solid, 34%. ^{1}H NMR (400MHz, $CDCl_3$)(complex mixture of rotamers) δ = 8.25 (br s, 1H), 7.55-6.30 (m, 13H), 5.33-4.63 (m, 3H), 4.35-3.99 (m, 2H), 3.70-3.62 (m, 5H), 3.39-3.00 (m, 4H), 2.89-2.83 (m, 2H).



Methyl 4-azidobutyrate 3-75

4-bromobutyric acid **3-72c (**100mg, 5.99mmol) was dissolved in dry methanol (10mL) and cooled to 0°C. $SOCl_2$ (1424mg, 0.87mL, 2eq) was then added dropwise, and the mixture was warmed to room temperature and further stirred for 4 hours. The solvent was evaporated in vacuo to obtain the intermediate crude methyl 4-bromobutyric acid **3-74** as a brown oil. The intermediate compound was again redissolved in bench DMSO (10mL) and NaN_3 (778.6mg, 2eq) was added. This suspension was then stirred overnight at 40°C. After reaction, the mixture was diluted with H_2O (40mL) and extracted with Et_2O three times. The combined organic extracts were washed with brine, dried with MgSO4, and evaporated in vacuo (the water bath temperature of the rotavap not set above 30°C) to obtain the product **3-75** as a pale orange oil and used without further purification. (518mg, 60% over 2 steps)

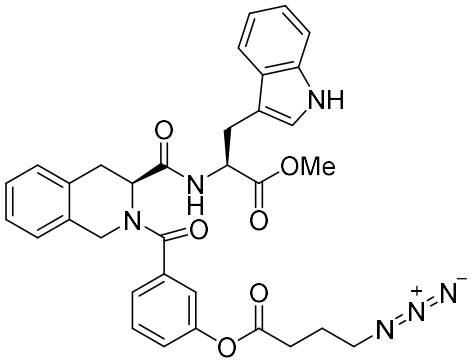
Pale orange oil, 60%, ^{1}H NMR (400MHz, $CDCl_3$) δ = 3.69 (s, 3H), 3.36 (t, J= 6.5Hz ,2H), 2.42 (t, J= 7.5Hz, 2H), 1.93 (tt, J= 6.5, 7.5Hz, 2H).



4-azidobutyric acid 3-73c

To a solution of methyl 4-azidobutyrate **3-75** (518.0mg, 3.62mmol) in MeOH (5mL) was added aqueous 2M NaOH (289.5mg, 3.62mL, 2.0eq), and the mixture was stirred at room temperature for 2 hours. After reaction, the mixture was partitioned between Et_2O (10mL) and H_2O (10mL). and layers separated. The aqueous layer was acidified to pH 1 with concentrated HCl, and further extracted with Et_2O twice. The combined organic extracts were washed with brine, dried with $MgSO_4$, and evaporated in vacuo (the water bath temperature of the rotavap not set above 30°C) to obtain the product as a pale orange oil. (468mg, quantitative yield).

Pale orange oil, 100%, 1H NMR (400MHz, $CDCl_3$) δ = 3.38 (t, J= 7.0Hz, 2H), 2.48 (t, J= 6.8Hz), 1.92 (quintet, J= 7.0Hz, 2H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 177.9, 50.7, 31.0, 24.2. FTIR (DCM, cm^{-1}) V_{max} = 3046, 2100, 1709.

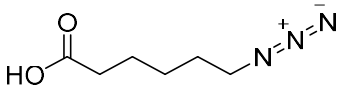


3-((S)-3-(((S)-3-(1H-indol-3-yl)-1-methoxy-1-oxopropan-2-yl)carbamoyl)-1,2,3,4-tetrahydroisoquinoline-2-carbonyl)phenyl 4-azidobutanoate 3-76c

Azide **3-73c** (142.7mg, 1.11mol, 1.1eq), was reacted with **3-3f** (500.0mg, 1.00mmol) according to general procedure 3. The compound was purified by column chromatography (Eluent 50:50 EtOAc/hexane) to obtain the product as an off-white foamy solid. (356mg, 58%).

Off-white solid, 58%. ^1H NMR (400MHz, $CDCl_3$)(complex mixture of rotamers) δ = 8.28 (br s, 1H), 7.54-6.30 (m, 13H), 5.16-4.64 (m, 3H), 4.36-4.01 (m, 2H), 3.70 (s, 3H), 3.47-2.98 (m, 6H), 2.73-2.64 (m, 2H), 2.07-2.02 (broadened t, J= 7.0Hz, 2H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 172.4, 171.6, 170.8, 170.3, 150.7, 136.6, 136.3, 133.8, 133.2, 131.8, 130.0, 128.4, 127.9,

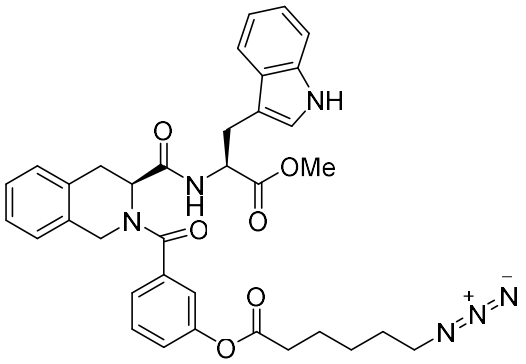
127.5, 126.7, 125.5, 124.8, 123.5, 122.2, 121.1, 119.7, 118.7, 111.4, 109.7, 53.8, 52.8, 52.6, 50.7, 48.0, 31.5, 29.6, 27.6, 24.3. ESI-MS(+): m/z = 609.08 $[M+H]^+$. HRMS calculated for $C_{33}H_{33}N_6O_6$ $[M+H]^+$ 609.2462. Found 609.2476.



6-azidohexanoic acid 3-73d

6-bromohexanoic acid **3-72d** (804.0mg, 4.12mmol) was reacted according to procedure 4 but substituting the solvent with 5mL of DMSO/H_2O mixture (3:1 v/v) and stirring at room temperature. The product was obtained as a colourless oil. (571mg, 88%).

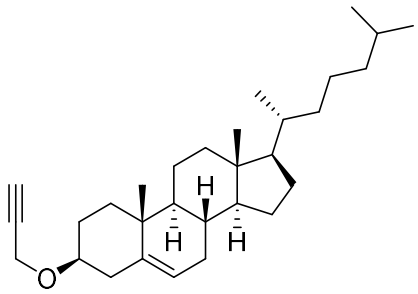
Colourless oil, 66%. ^{1}H NMR (400MHz, $CDCl_3$) δ = 3.28 (t, J= 6.9Hz, 2H), 2.38 (t, J= 7.3Hz, 2H), 1.71-1.40 (m, 4H), 1.44 (tt, J= 6.9, 7.3 Hz, 2H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 178.1, 51.4, 33.7, 28.7, 26.4, 24.4. FTIR (DCM, cm^{-1}) V_{max} = 3410, 2102, 1627, 1265.



3-((S)-3-(((S)-3-(1H-indol-3-yl)-1-methoxy-1-oxopropan-2-yl)carbamoyl)-1,2,3,4-tetrahydroisoquinoline-2-carbonyl)phenyl 6-azidohexanoate 3-76d

Azide **3-73d** (0.442mmol, 1.1eq, 69.5mg) was reacted with **3-3f** (0.402mmol, 200.0mg) according to general procedure 3. The compound was purified by column chromatography (Eluent 40:60 EtOAc/hexane) to obtain the product as an off-white foamy solid. (166mg, 65%)

Off-white solid, 65%. ^{1}H NMR (400MHz, $CDCl_3$)(complex mixture of rotamers)(at 50°C) δ = 8.21 (br s, 1H), 7.53-6.82 (m, 13H), 5.17-4.05 (m, 4H), 3.69 (s, 3H), 3.33-3.04 (m, 7H), 2.61 (t broadened to almost a s, J= 7.0Hz, 2H), 1.84-1.50 (m, 6H). ^{13}C NMR (100MHz, $CDCl_3$)(complex mixture of rotamers) δ = 127.3 (br s), 170.8, 170.3, 136.6, 136.3, 133.7, 133.1, 129.9, 128.4, 127.9, 126.7, 125.5, 123.5, 122.2, 121.6, 119.7, 118.7, 111.4, 109.6, 53.7, 52.8, 52.6, 51.3, 48.0, 34.3, 29.5, 28.7, 27.6, 26.4, 24.5. FTIR (DCM, cm^{-1}) ν_{max} = 3464, 3410, 3053, 2943, 2099, 1744, 1680, 1265. ESI-MS(+): m/z= 636.71 $[M+H]^+$. HRMS calculated for $C_{35}H_{37}N_6O_6$ $[M+H]^+$ 637.2775. Found 637.2776.

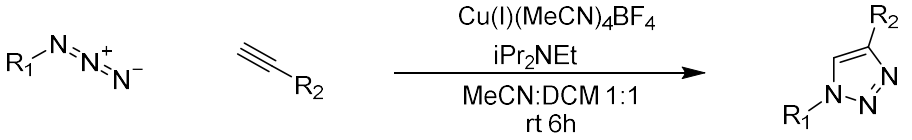


Cholesteryl propargyl ether 3-78

Sodium hydride (60% w/w in mineral oil, 206.9mg total, 5.17mmol, 2eq) was suspended in dry THF and cooled to 0°C. To this mixture was added cholesterol **3-19** (1000mg, 2.59mmol) followed by dropwise addition of propargyl bromide **3-77** (80% w/w solution in toluene, 576.9mg total, 3.88mmol ,1.5eq). The solution was warmed to room temperature and stirred for another 48 hours. When TLC showed completion, the mixture was quenched by slow dropwise addition of MeOH, and effervescence was observed. When no more bubbling occurs on further addition of MeOH, the solvent was evaporated in vacuo. The residue was partitioned between Et_2O and 2M HCl, and the layers separated. The aqueous layer was extracted with Et_2O twice, and the combined organic layers washed with brine, dried with MgSO4, and evaporated in vacuo to obtain the compound, which was purified by column chromatography (Eluent 10:90 EtOAc:hexane) to obtain the product as a pale yellow solid (876mg, 80%).

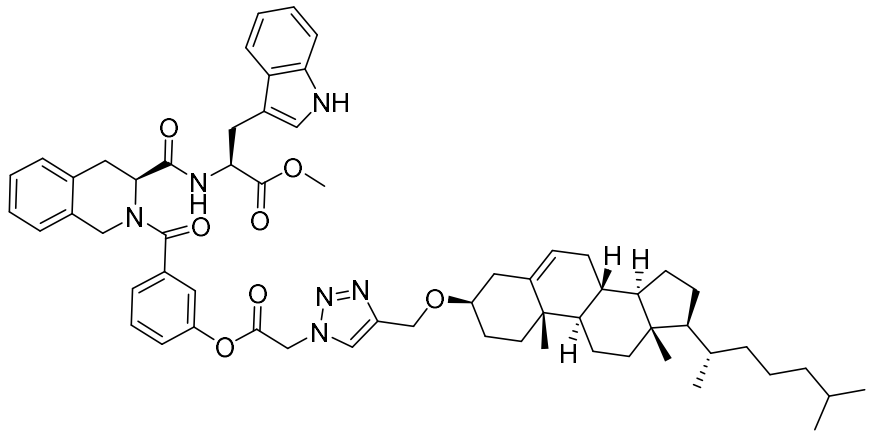
Pale yellow solid, 80%. Mp: 110-114°C. ^{1}H NMR (400MHz, $CDCl_3$) δ = 5.37-5.35 (m, 1H), 4.19 (d, J= 3.0Hz, 1H), 3.42-3.35 (m, 1H), 2.40-2.37 (m, 2H), 2.24-2.18 (m, 1H), 2.03-1.79 (m, 5H), 1.58-0.86 (m, 34H), 0.68 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$) δ = 140.8, 122.1, 80.7, 78.4, 73.9, 57.0, 56.4, 55.3, 50.4, 42.5, 40.0, 39.7, 38.9, 37.3, 36.4, 36.0, 32.1 (2 overlapping peaks), 28.4, 28.3, 28.2, 24.5, 24.0, 23.0, 22.8, 21.3, 19.5, 18.9 ,12.1. HRMS calculated for $C_{30}H_{49}O$ $[M+H]^-$ 425.3783. Found 425.3800.

General procedure 5: Copper mediated azide/alkyne cycloaddition



The azide (1eq) and alkyne were dissolved in DCM. In another vessel, tetrakis(acetonitrile) copper (I) tetrafluoroborate (0,2eq) was dissolved in an equal volume of MeCN. The 2 solutions were combined, and the combined mixture flushed with inert gas for 15 mins. After that, DIPEA (1.5eq) was added to the mixture was left to stir for 6 hours.

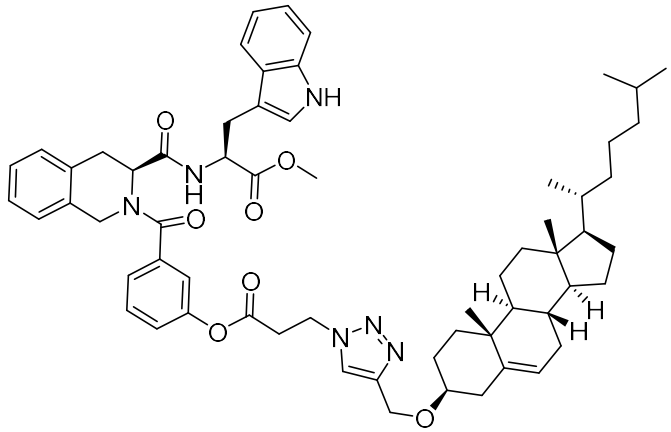
When TLC showed completion, the mixture was partitioned between EtOAc and water. The layers were separated, and the aqueous layer extracted with EtOAc twice. The combined organic layers were then washed with brine, dried with $MgSO_4$, and filtered through celite. The solvent was evaporated in vacuo to obtain the crude mixture, which was purified by column chromatography if necessary to obtain the product.



3-53a

3-76a (70.0mg, 0.121mmol) was reacted with alkyne **3-78** (51.2mg, 0.121mmol) under general condition 5. The crude mixture was washed with hexane to obtain analytically pure compound as an off-white solid (63mg, 52%)

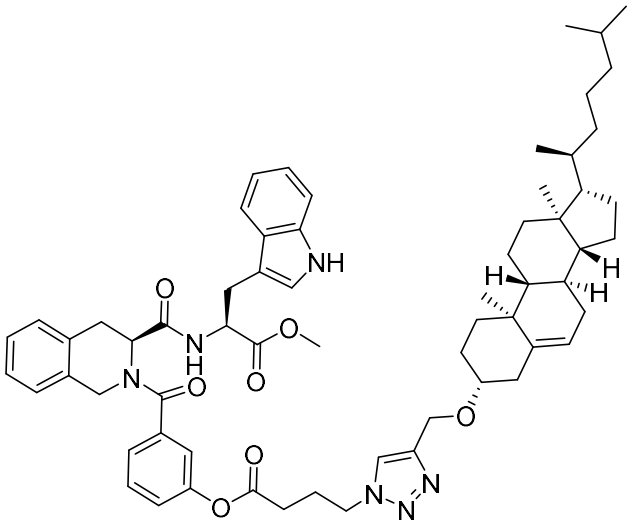
Off-white solid. 52%. Mp: 109-115°C. ^{1}H NMR (400MHz, $CDCl_3$)(complex mixture of rotamers) δ = 8.40 (s, 1H), 7.76 (s, 1H), 7.63-6.35 (m, 13H), 5.43 (s, 1H), 5.34-5.33 (m, 1H), 5.11-4.89 (m, 1.5H), 4.7 (s, 1.5H), 4.55-4.04 (m, 2H), 3.69 (s, 3H), 3.35-2.88 (m, 5H), 2.43-2.40 (m, 1H), 2.26-2.18 (m, 1H), 1.99-1.78 (m, 5H), 1.58-0.85 (m, 36H), 0.67 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$)(complex mixture of rotamers), δ = 172.4, 170.6, 170.5, 170.3, 165.1, 149.9, 147.0 (br s), 140.7, 136.9, 136.3, 133.8, 133.2, 131.7 (br s), 131.6, 131.2, 128.4, 128.0, 127.7, 127.3, 126.8, 125.5, 124.1, 123.5, 123.1, 122.2, 122.1, 120.6, 120.2, 119.7, 118.7, 118.3, 111.7 (br s), 111.5, 109.6, 79.3, 61.7, 56.9, 56.3, 54.0, 52.9, 52.7, 51.0, 50.3, 48.1, 42.5, 39.9, 39.7, 39.1, 37.3, 37.0, 36.4, 32.1, 32.0, 29.8 (br s), 28.5, 28.4, 28.2, 27.6, 27.1, 24.5, 24.0, 23.0, 22.7, 21.2, 19.5, 19.0, 12.0. ESI-MS(+): m/z = 1005.47 $[M+H]^+$ HRMS calculated for $C_{61}H_{76}N_6O_7Na$ $[M+Na]^+$ 1027.5673. Found 1027.5680.



3-53b

3-76b (54.0mg,0.091mmol) was reacted with alkyne **3-78** (38.6mg, 0.091mmol) under general condition 5. The crude mixture was washed with hexane to obtain analytically pure compound as an off-white solid (80mg, 86%)

Off-white solid. 86%. ^{1}H NMR (400MHz, $CDCl_3$)(complex mixture of rotamers) δ = 8.50 (2 overlapping br s, 1H), 7.66-6.36 (m, 14H), 5.34 (s, 1H), 5.14-4.66 (m, 6H), 4.33-4.02 (m, 1H), 3.70 (br s, 3H), 3.32-2.97 (m, 7H), 2.41-2.20 (m, 2H), 2.10-0.86 (m, 40H), 0.67 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$)(complex mixture of rotamers), δ = 170.3, 170.2, 169.4, 150.4, 146.4, 140.8, 136.8, 136.4, 133.9, 132.2, 130.3, 130.0, 128.5, 127.9, 127.6, 126.8, 125.5, 125.0, 123.4, 123.3, 122.3 (2 peaks), 122.1, 119.8, 118.7, 111.5, 109.7, 79.3, 61.7, 57.0, 56.4, 53.8, 52.6, 50.4, 45.5, 42.5, 40.0, 39.7, 39.2, 37.3, 37.0, 36.4, 36.0, 34.9, 32.1 29.8, 28.5 (2 peaks), 28.2, 24.5, 24.0, 23.0, 22.8, 21.3, 19.6, 18.9, 12.1. ESI-MS(+): e/z = 1019.18 $[M+H]^+$. HRMS calculated for $C_{62}H_{79}N_6O_7$ 1019.6010. Found 1019.6027.

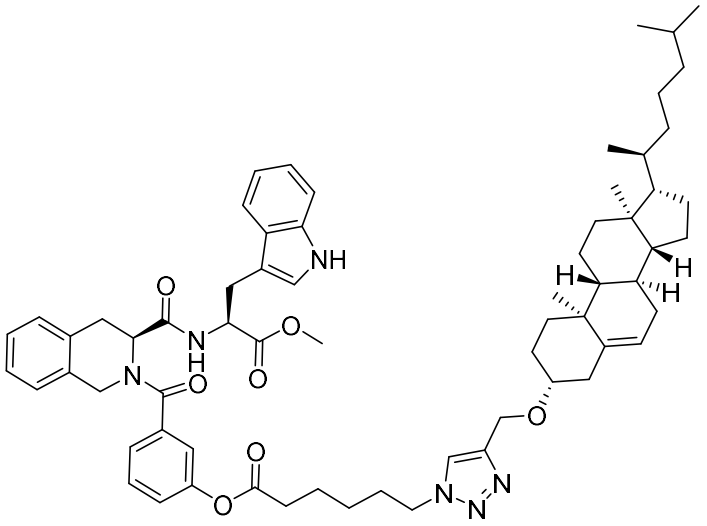


3-53c

3-76c (71.7mg,0.118mmol) was reacted with alkyne **3-78** (50.0mg, 0.118mmol) under general condition 5. The crude mixture was washed with hexane to obtain analytically pure compound as an off-white solid (107mg, 88%)

Off-white solid. 88%. Mp:114-118°C. ^{1}H NMR (400MHz, $CDCl_3$)(complex mixture of rotamers) δ = 8.43 (br s with shoulder peak, 1H), 7.58-6.28 (m, 13H), 5.36-5.34 (m, 1H), 5.16-5.05 (m, 1H), 4.90-4.50 (m, 5H), 4.35-3.98 (m, 1H), 3.70 (s with shoulder peak, 3H), 3.37-3.03 (m, 5H), 2.68-2.59 (m, 2H), 2.40-2.22 (m, 2H), 1.60-0.86(m, 41H), 0.67 (s, 3H). ^{13}C NMR

(100MHz, $CDCl_3$)(complex mixture of rotamers), δ = 172.4, 170.7 (br s), 170.3 (br s), 150.6, 140.8, 136.7, 136.4, 133.9, 133.2, 130.0, 128.4, 128.3, 127.9, 127.8, 127.6, 126.7, 125.5, 123.6, 123.5, 122.2, 122.0, 121.0, 119.7, 118.7, 111.5, 109.6, 79.3, 61.8, 56.9, 56.3, 53.9, 52.9, 52.6, 50.3, 49.1, 48.1, 42.5, 40.0, 39.7, 39.1, 37.3, 37.0, 36.4, 35.9, 32.1 (2 overlapping peaks), 30.9, 28.5, 28.4, 28.2, 25.4, 24.5, 24.0, 23.0, 22.7, 21.2, 19.5, 18.9, 12.0. ESI-MS(+): m/z = 1033.51 $[M+H]^+$. HRMS calculated for $C_{63}H_{81}N_6O_7$ 1033.6167. Found 1033.6180.

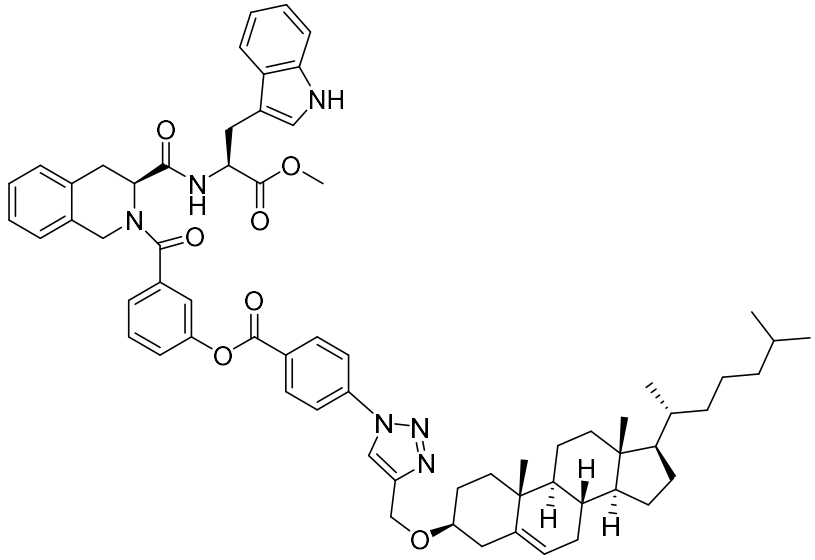


3-53d

3-76d (0.042mmol, 27.0mg) was reacted with alkyne **3-78** (18.0mg, 0.042mmol) under general procedure 5. The crude mixture was washed with hexane to obtain analytically pure compound as an off-white solid (34.0mg, 76%).

Off-white solid. 76%. Mp: 91-95°C; ^1H NMR (400MHz, $CDCl_3$)(complex mixture of rotamers) δ = 8.40 (1H, broad s), 7.52-6.31(m, 15H), 5.35 (m, 1H), 5.16-4.88 (m, 1.6H), 4.69-4.67 (overlapping s and br s, 2.4H), 4.36-4.34 (m, 2.4H), 4.03-3.99 (apparent d, 0.6H), 3.70-3.63 (s with satellite peak, 3H), 3.51-2.90 (m, 6H), 2.63-2.57 (m, 2H), 2.43-2.39 (m, 1H), 2.28-2.20 (m, 1H), 2.03-0.85 (m, 45H), 0.67 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$))(complex mixture of rotamers) δ = 172.4 (br s), 172.1 (br s) 170.2 (br s), 170.2 (br s), 150.8, 146.3, 140.9, 136.6, 136.4, 133.9, 133.2, 129.9, 128.5, 127.5, 126.7, 125.5, 124.7, 123.5, 122.3, 122.0, 121.2, 119.8,

118.8, 111.4, 79.2, 61.9, 57.0, 56.4, 53.7, 52.6, 50.4, 48.0, 42.5, 40.0, 39.7, 39.3, 37.03, 37.0, 36.4, 36.0, 35.1, 32.1, 30.2, 29.4, 28.5, 28.4, 28.2, 27.7, 25.2, 24.5, 24.0, 23.0, 22.7, 21.3, 19.6, 12.1. ,ESI-MS(+): m/z = 1061.57 $[M+H]^+$. HRMS calculated for $H_{65}H_{85}N_6O_7$ 1061.6480. Found 1061.6486.



3-53e

3-76e (32.9mg,0.051mmol) was reacted with alkyne **3-78** (21.7mg, 0.051mmol) under general condition 5. The crude mixture was washed with hexane to obtain analytically pure compound as an off-white solid (39mg, 77%)

Off-white solid. %. ^{1}H NMR (400MHz, $CDCl_3$)(complex mixture of rotamers) δ = 8.41-8.33 (m, 3H), 8.11 (br s, 1H), 7.98-7.95 (m, 2H), 7.53-6.40 (m, 13H), 5.39-5.38 (m, 1H), 5.17-4.61 (m, 4H), 4.32-4.04 (m, 4H), 3.70-3.52 (br s with "satellite" peaks) 3H), 3.43-3.06 (m, 5H), 2.47-2.44 (m, 1H), 2.33-2.26 (m, 1H), 2.05-1.77 (m, 6H), 1.56-0.86 (m, 31H), 0.68 (s, 3H). ^{13}C NMR (100MHz, $CDCl_3$))(complex mixture of rotamers) δ = 172.4, 171.1, 170.7(br s), 170.3, 150.6, 140.8, 136.7, 136.4, 133.9, 133.2, 129.9, 128.4, 127.9, 127.5, 126.7, 125.5, 124.8, 123.5 (2 peaks), 122.2, 122.0, 121.0, 119.7, 118.7, 111.5, 109.6, 79.3, 61.8, 56.9, 56.3, 53.9/ 52.9, 52.8,50.3, 49.5, 48.6, 42.5, 40.2, 40.0, 39.6, 37.2, 37.0, 36.4, 36.1, 32.1, 31.2, 29.8, 28.4, 28.2,

25.6, 24.7, 24.3, 23.0, 22.7, 21.8, 19.4, 18.9, 12.0 . ESI-MS(+): e/z = 1067.25 $[M+H]^+$. HRMS calculated for $C_{66}H_{97}N_6O_7$ $[M+H]^+$. 1067.6010. Found 1067.6008.

3.5.7 Biological assays of EpMF2-cholesterol conjugates

All biological experiments and assays are credited to Ragunathan Priya and Logeshwari

Preparation of inverted membrane vesicles from *M. smegmatis*

About 5 g (wet weight) of *M. smegmatis* were resuspended in 20 ml membrane preparation buffer (50 mM MOPS, 2 mM $MgCl_2$, pH 7.5) containing EDTA-free protease inhibitor cocktail (1 tablet in 20 ml buffer, Roche-USA) and 1.2 mg/ml lysozyme. The suspension was stirred at room temperature for 45 min and additionally supplemented with 300 µl 1 M $MgCl_2$ and 50 µl DNase I (Thermo Fischer, USA), and continued stirring for another 15 min at room temperature. All subsequent steps were performed on ice. Cells were broken by three passages through an ice precooled Model M-110L Microfluidizer processor (M-110L) at 18,000 psi. The suspension containing lysed cells was centrifuged at 4,200 x g at 4 °C for 20 min. The supernatant containing membrane fraction was further subjected to ultracentrifugation 45,000 x g at 4 °C for 1 h. The supernatant was discarded, and the precipitated membrane fraction was resuspended in membrane preparation buffer containing 15% glycerol, aliquoted, snap frozen and stored at -80 °C. The concentrations of the proteins in the vesicles were determined by the BCA method.

ATP Synthesis Assay

ATP synthesis was measured in flat bottom white microtiter 96 well plates (Corning USA). The reaction mixture was made in assay buffer (50 mM MOPS, pH 7.5, 10 mM $MgCl_2$) containing 10 µM ADP, 250 µM Pi and 1 mM NADH. Concentration of Pi was adjusted by adding 100 mM KH_2PO_4 salt dissolved in the assay buffer. ATP synthesis was started by adding inverted vesicles of *M. smegmatis* to a final protein concentration of 5 µg/ml. The

reaction mixture was incubated at room temperature for 30 min before addition of 50 µl of the CellTiter-glow reagent, and the mixture was incubated for another 10 min in dark at room temperature. Emitted luminescence, which is correlated to the synthesized ATP, was measured by a Tecan plate reader Infinite 200 Pro (Tecan USA), using the following parameter: luminescence, integration time of 500 ms and no attenuation.

Agar and broth MIC

Initial stock solutions of the test compounds were made in 90% DMSO to a concentration of 10 mM. Bedaquiline was used as a positive control and the vehicle DMSO was used as negative control. In the first approach, the compounds were tested on microbial cultures at a fixed concentration of 50 µM. Each of the above strains were cultured at 37 °C in Middlebrook 7H9 broth supplemented with 0.2% glycerol and 10% ADC (Albumin Dextrose Catalase) until logarithmic growth was achieved (OD_{600} 0.4 - 0.6). The test inoculum was obtained by diluting the suspensions to OD_{600} 0.1 to a final volume of 1 ml in the test tubes and were incubated at 37 °C for 24 hours.

Broth micro dilution method was used to determine the minimum inhibitory concentration (MIC) of the test compounds. Two-fold serial dilutions of the compounds were made from 200 µM to 0.05 µM. Each concentration was assayed in triplicates. The diluted test inoculum was added to all the wells (final volume 200 µl) in the microplate and was incubated at 37 °C for 24 hours (*M. smeg.*) and 5 days (BCG). The final OD of the cultures in the plate was measured by Tecan Infinite 200 PRO plate reader. The MIC_{50} was defined as the drug concentration that inhibited 50% of the bacterial growth when compared to the growth in the drug-free medium.

3.6 References

1. Tran, S. L.; Cook, G. M., *J Bacteriol* **2005,** *187* (14), 5023-5028.

2. Gupta, S.; Cohen, K. A.; Winglee, K.; Maiga, M.; Diarra, B.; Bishai, W. R., *Antimicrob Agents Chemother* **2014,** *58* (1), 574-576.

3. Yeaman, M. R.; Yount, N. Y., **2003,** *55* (1), 27-55.

4. Bahar, A. A.; Ren, D., *Pharmaceuticals (Basel)* **2013,** *6* (12), 1543-1575.

5. Findlay, B.; Zhanel, G. G.; Schweizer, F., *Antimicrob Agents Chemother* **2010,** *54* (10), 4049-4058.

6. Matsuzaki, K., *Biochim Biophys Acta* **2009,** *1788* (8), 1687-1692.

7. Timofeeva, L.; Kleshcheva, N., *Appl Microbiol Biotechnol* **2011,** *89* (3), 475-492.

8. Hoque, J.; Konai, M. M.; Samaddar, S.; Gonuguntala, S.; Manjunath, G. B.; Ghosh, C.; Haldar, J., *Chem Commun (Camb)* **2015,** *51* (71), 13670-13673.

9. Yang, T.; Moreira, W.; Nyantakyi, S. A.; Chen, H.; Aziz, D. B.; *J Med Chem* **2017,** *60* (7), 2745-2763.

10. Protopopova, M.; Hanrahan, C.; Nikonenko, B.; Samala, R.; Chen, P.; Gearhart, J.; Einck, L.; Nacy, C. A., *J Antimicrob Chemother* **2005,** *56* (5), 968-974.

11. Sacksteder, K. A.; Protopopova, M.; Barry, C. E., 3rd; Andries, K.; Nacy, C. A., *Future Microbiol* **2012,** *7* (7), 823-837.

12. Borisov, S. E.; Bogorodskaya, E. M.; Volchenkov, G. V.; Kulchavenya, E. V.; Maryandyshev, A. O.; Skornyakov, S. N.; Talibov, O. B.; Tikhonov, A. M.; Vasilyeva, I. A., *Tuberculosis and lung diseases* **2018,** *96* (3), 6-18.

13. Dunn, E. A.; Roxburgh, M.; Larsen, L.; Smith, R. A.; McLellan, A. D.; Heikal, A.; Murphy, M. P.; Cook, G. M., *Bioorg Med Chem* **2014,** *22* (19), 5320-5328.

14. Amaral, L.; Kristiansen, J. E.; Viveiros, M.; Atouguia, J., *J Antimicrob Chemother* **2001,** *47* (5), 505-511.

15. Hards, K.; McMillan, D. G. G.; Schurig-Briccio, L. A.; Gennis, R. B.; Lill, H.; Bald, D.; Cook, G. M., *Proc Natl Acad Sci U S A* **2018,** *115* (28), 7326-7331.

16. Santos, M. R. E.; Fonseca, A. C.; Mendonca, P. V.; Branco, R.; Serra, A. C.; Morais, P. V.; Coelho, J. F. J., *Materials (Basel)* **2016,** *9* (7), 599.

17. Jain, A.; Duvvuri, L. S.; Farah, S.; Beyth, N.; Domb, A. J.; Khan, W., *Adv Healthc Mater* **2014,** *3* (12), 1969-1985.

18. Yavvari, P. S.; Gupta, S.; Arora, D.; Nandicoori, V. K.; Srivastava, A.; Bajaj, A., *Biomacromolecules* **2017,** *18* (7), 2024-2033.

19. Phillips, D. J.; Harrison, J.; Richards, S. J.; Mitchell, D. E.; Tichauer, E.; Hubbard, A. T. M.; Guy, C.; Hands-Portman, I.; Fullam, E.; Gibson, M. I., *Biomacromolecules* **2017,** *18* (5), 1592-1599.

20. Ouellet, H.; Johnston, J. B.; de Montellano, P. R., *Trends Microbiol* **2011,** *19* (11), 530-539.

21. Chang, J. C.; Miner, M. D.; Pandey, A. K.; Gill, W. P.; Harik, N. S.; Sassetti, C. M.; Sherman, D. R., *J Bacteriol* **2009,** *191* (16), 5232-5239.

22. Munoz, S.; Rivas-Santiago, B.; Enciso, J. A., *Scand J Immunol* **2009,** *70* (3), 256-263.

23. Gatfield, J.; Pieters, J., *Science* **2000,** *288* (5471), 1647-1650.

24. Abuhammad, A., *Br J Pharmacol* **2017,** *174* (14), 2194-2208.

25. Peyron, P.; Vaubourgeix, J.; Poquet, Y.; Levillain, F.; Botanch, C.; Bardou, F.; Daffe, M.; Emile, J. F.; Marchou, B.; Cardona, P. J.; de Chastellier, C.; Altare, F., *PLoS Pathog* **2008,** *4* (11), e1000204.

26. Pandey, A. K.; Sassetti, C. M., *Proc Natl Acad Sci U S A* **2008,** *105* (11), 4376-4380.

27. Nazarova, E. V.; Montague, C. R.; La, T.; Wilburn, K. M.; Sukumar, N.; Lee, W.; Caldwell, S.; Russell, D. G.; VanderVen, B. C., *Elife* **2017,** *6*, e26969.

28. Lee, W.; VanderVen, B. C.; Fahey, R. J.; Russell, D. G., *J Biol Chem* **2013,** *288* (10), 6788-6800.

29. Ruwizhi, N.; Aderibigbe, B. A., *Molecules* **2020,** *25* (18), 4330.

30. A. Tunek, L. A. S., *Drug Metabolism and Disposition* **1988,** *16* (5), 759-764.

31. Cao, J.; Dang, G.; Li, H.; Li, T.; Yue, Z.; Li, N.; Liu, Y.; Liu, S.; Chen, L., *PLoS One* **2015,** *10* (9), e0138151.

32. Umehara, A.; Ueda, H.; Tokuyama, H., *J Org Chem* **2016,** *81* (22), 11444-11453.

33. Knop, K.; Hoogenboom, R.; Fischer, D.; Schubert, U. S., *Angew Chem Int Ed Engl* **2010,** *49* (36), 6288-6308.

34. Guzzo, P. R.; Miller, M. J., Catalytic, *The Journal of Organic Chemistry* **2002,** *59* (17), 4862-4867.

35. Potyen, M.; Josyula, K. V. B.; Schuck, M.; Lu, S.; Gao, P.; Hewitt, C., *Organic Process Research & Development* **2007,** *11* (2), 210-214.

36. Lane, C. F., *Aldrichmica Acta* **1975,** *8*, 20-23.

37. Neises, B.; Steglich, W., S*Angewandte Chemie International Edition in English* **1978,** *17* (7), 522-524.

38. Dai, J.; Dan, W.; Li, N.; Wang, J., *Eur J Med Chem* **2018,** *157*, 333-338.

39. Ruiz, M.; Lopez-Alvarado, P.; Menendez, J. C., *Org Biomol Chem* **2010,** *8* (20), 4521-4523.

40. Snider, B. B.; Zeng, H., *J Org Chem* **2003,** *68* (2), 545-563.

41. Langer, P.; Wu, X.-F.; Oschatz, S.; Sharif, M., *Synthesis* **2015,** *47* (17), 2641-2646.

42. Singh, A.; Singh, S.; Sewariya, S.; Singh, N.; Singh, P.; Kumar, A.; Bandichhor, R.; Chandra, R., *Tetrahedron* **2021,** *84*.

43. KHMDS. https://www.chembk.com/en/chem/KHMDS (accessed 31 July 2023).

44. R., B., Essentials of heterocyclic chemistry-I. chemistry1-2009.pdf, E. o. h., Ed. Scripps Research Institute: San Diego, 2009.

45. Ojeda-Amador, A. I.; Martinez-Martinez, A. J.; Robertson, G. M.; Robertson, S. D.; Kennedy, A. R.; O'Hara, C. T., *Dalton Trans* **2017,** *46* (19), 6392-6403.

46. Murase, H.; Wakisaka, G.; Noguchi, T.; Sasaki, S., *Bioorg Med Chem* **2020,** *28* (20), 115730.

47. Criegee, R., *Angewandte Chemie International Edition in English* **1975,** *14* (11), 745-752.

48. Singla, P.; Kaur, M.; Kumari, A.; Kumari, L.; Pawar, S. V.; Singh, R.; Salunke, D. B., *ACS Omega* **2020,** *5* (8), 3952-3963.

49. Viaud, M. C.; Rollin, P., *Synthesis* **1990,** *1990* (02), 130-132.

50. Schumacher, C.; Ward, J. S.; Rissanen, K.; Bolm, C.; Aly, M., *Beilstein J Org Chem* **2023,** *19*, 91-99.

51. Sun, Q.; Cai, S.; Peterson, B. R., *Org Lett* **2009,** *11* (3), 567-570.

52. Matin, M. M.; Matin, P.; Rahman, M. R.; Ben Hadda, T.; Almalki, F. A.; Mahmud, S.; Ghoneim, M. M.; Alruwaily, M.; Alshehri, S., *Front Mol Biosci* **2022,** *9*, 864286.

53. Rostovtsev, V. V.; Green, L. G.; *Angew Chem Int Ed Engl* **2002,** *41* (14), 2596-2599.

54. Tornoe, C. W.; Christensen, C.; Meldal, M., *J Org Chem* **2002,** *67* (9), 3057-3064.

55. Hein, C. D.; Liu, X. M.; Wang, D., *Pharm Res* **2008,** *25* (10), 2216-2230.

56. Nandivada, H.; Jiang, X.; Lahann, J., *Advanced Materials* **2007,** *19* (17), 2197-2208.

57. Lutz, J. F.;. *Angew Chem Int Ed Engl* **2007,** *46* (7), 1018-1025.

58. Hanni, K. D.; Leigh, D. A., *Chem Soc Rev* **2010,** *39* (4), 1240-1251.

59. Lallana, E.; Riguera, R.; Fernandez-Megia, E., *Angew Chem Int Ed Engl* **2011,** *50* (38), 8794-8804.

60. Amblard, F.; Cho, J. H.; Schinazi, R. F., *Chem Rev* **2009,** *109* (9), 4207-4220.

61. Dedola, S.; Nepogodiev, S. A.; Field, R. A., *Org Biomol Chem* **2007,** *5* (7), 1006-1017.

62. Kolb, H. C.; Sharpless, K. B., *Drug Discovery Today* **2003,** *8* (24), 1128-1137.

63. Huisgen, R., *Angewandte Chemie International Edition in English* **1963,** *2* (11), 633-645.

64. Clarke, D.; Mares, R. W.; McNab, H., *Journal of the Chemical Society, Perkin Transactions 1* **1997,** (12), 1799-1804.

65. Worrell, B. T.; Malik, J. A.; Fokin, V. V., *Science* **2013,** *340* (6131), 457-460.

66. Zhu, L.; Brassard, C. J.; Zhang, X.; Guha, P. M.; Clark, R. J., *Chem Rec* **2016,** *16* (3), 1501-1517.

67. Meldal, M.; Tornoe, C. W., *Chem Rev* **2008,** *108* (8), 2952-3015.

68. Hein, J. E.; Fokin, V. V., *Chem Soc Rev* **2010,** *39* (4), 1302-1315.

CHAPTER 4

Future Work and Conclusion

4. Conclusion and future work

4.1 Future Work

4.1.1 EpMF2 detection in cells.

Going by the results of assays from Chapter 3, a few possible avenues of future research can be investigated.

As mentioned in chapter 3, the hypothesis was that EpMF2 failed to bypass the mycobacterial cell wall membrane. There are additional methods to test this that were not described in the previous chapter due to time constraints.

Confocal microscopy is one possible option to prove the existence of EpMF2 in the cell by visualising fluoresced light coming from the molecule, or a dye conjugated to the drug. A confocal microscope uses a pinhole to block out-of-focus light from the sample, significantly increasing the image resolution by focusing on a thin slice of the sample. By adjusting the illumination point over the sample, scans over multiple slices could be done and a 3D image could be obtained. Most confocal microscopes are fluorescence microscopes, and detect light emitted from fluorescent dyes or other molecules. In the context of biology and life science, there are many types of stains and dyes that bind to specific components of the cell, allowing its visualisation of locations of interest by looking at emitted light in a dark background.

To undergo confocal microscopy, the dye, or object of interest, must be fluorescent. A vast variety of fluorophores are commercially available for this purpose. Such dyes include small molecules as well as proteins. Oftentimes, they are conjugated to an antibody, peptide or protein that binds to specific cell organelles, allowing its visualisation under fluorescence microscopy.

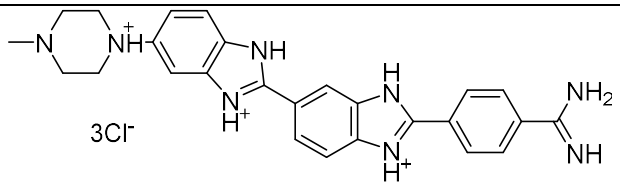
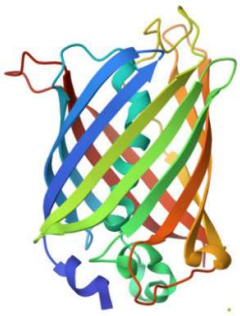
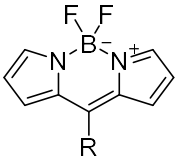
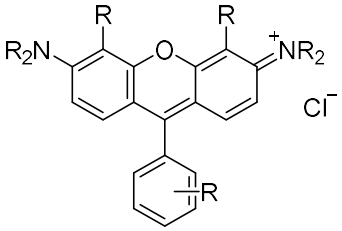
Dye	Structure	Target
Hoechst Dyes		DNA
GFP (Green fluorescent protein)	 Adapted from [1]	Tagged to protein of interest
BODIPY Dyes		lipids
Rhodamines		Tagged to other peptides, antibodies etc.

Table 4.1: Molecular structure of common classes of dyes used in biological staining.

EpMF2 is mildly fluorescent as it contains a tryptophan moiety. Among all amino acids, tryptophan is the most fluorescent. While phenylalanine and tyrosine do fluoresce, they do so much more weakly. Therefore, in proteins, tryptophans contribute the most to its fluorescence. This characteristic has been exploited in tryptophan fluorescence spectroscopy to study protein structures.[2] However, direct imaging of cells using tryptophan itself as the fluorophore is very

limited. The main reason is that tryptophan absorbs in the UV region, and rapidly undergoes bleaching after a few excitation cycles. However, an attempt to excite tryptophan using a two-photon excitation gave some limited success in the imaging of spider hemocyanins and avidin-coated spheres, both proteins containing hundreds of tryptophan residues per molecule.[3] Tryptophan-fluorescence microscopy has also been done for imaging leukocytes,[4](two-photon excitation method), non-melanoma skin cancer cells[5] antimicrobial peptides[6](single-photon). Therefore, it should be possible to visualise EpMF2 via in-vivo fluorescence microscopy. The emissions wavelengths from EpMF2 tryptophans should be distinguishable from emissions arising from tryptophans from bacterial proteins, due to their different chemical environment. Tryptophan in EpMF2, which is exposed to the aqueous environment, have higher emission wavelengths of 350nm compared to tryptophans deep in hydrophobic residues of proteins that fluorescence around 310nm.[2]

An alternative method to visualise EpMF2 would be to conjugate it with a fluorophore. As an example, Lavilla and Vendrall *et al*, have conjugated a highly fluorescent but compact BODIPY moiety onto the 2-position of tryptophan **(4-1)** in peptides to render them highly fluorescent and demonstrated its use in imaging of fungal pathogens.[7] However, given the small molecular mass of EpMF2-cholesterol conjugate, attachment of the moiety will still add a significant percentage of molar mass adversely affecting its pharmacokinetic properties that may not be representative of the original structure.

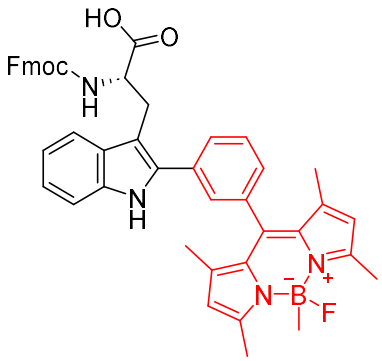


Figure 4.2: Lavilla and Vendrall's fluorescent BODIPY conjugate tryptophan **(4-1)**

Future work needs to be done to establish ideal conditions for viewing EpMF2 under fluorescent confocal microscopy. The above two imaging methods can be used as a starting point, but the difficulties of both methods should be kept in mind.

4.1.2 Expanding EpMF2-polymer conjugate analogue synthesis

As mentioned in the previous chapter, work in the synthesis of the EpMF2-polymer conjugate is still at an early stage. Three stages of future work are expected. Firstly, the synthesis of more analogues and optimization of the click reaction. Aside from the original starting polymer **4-2**, variants such as 4-3 can be used with an extended alkyl chain to the alkyne to make it more hydrophobic to see its effects on antimicrobial activity. A PEGylated variant of the polymer with a gallate moiety at the end at was being synthesized in the group as a more hydrophilic analogue. Attachment of propargyl group as in **4-4** would be another idea. On the other hand, the positions of the alkyne and azide groups can be swapped so that the azide would lie on the polymer, as shown in **4-5**. This would require synthesis of an alkyne containing derivative of EpMF2 analogue such as **4-6.**

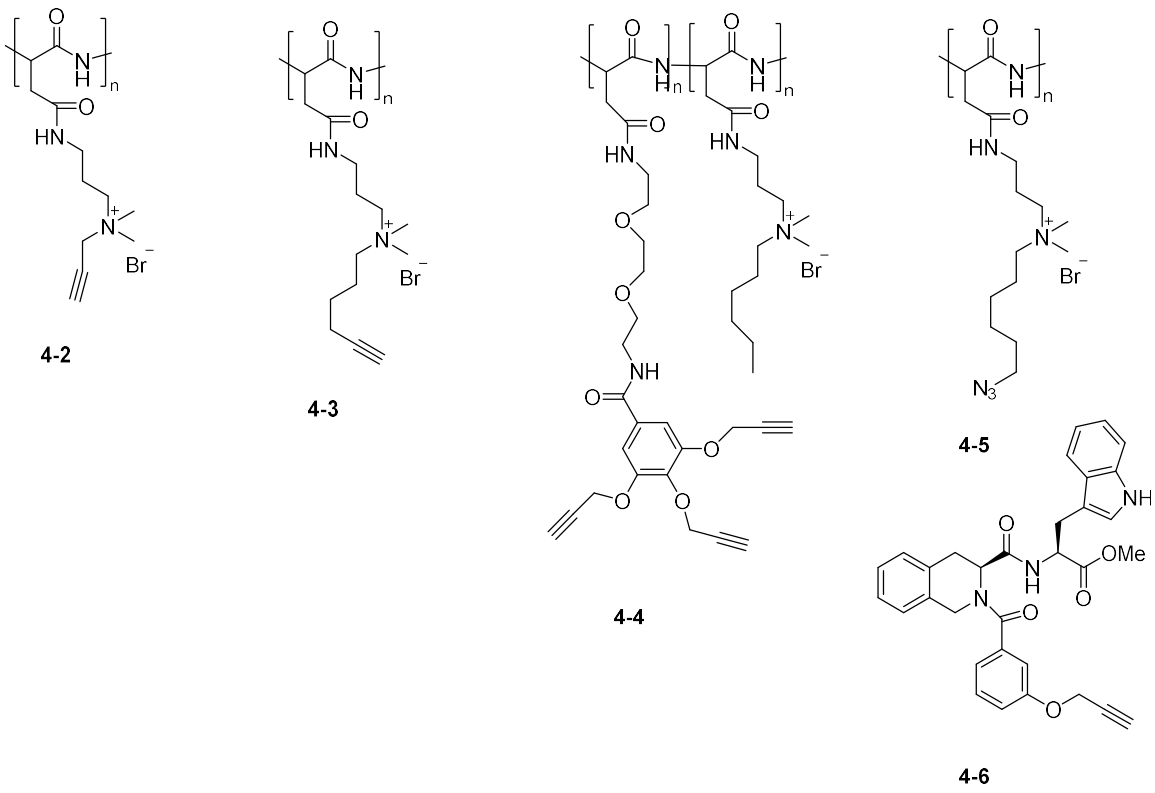
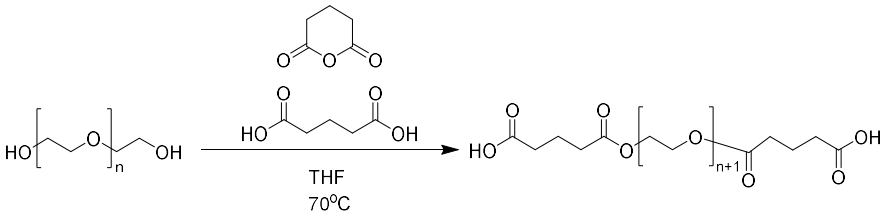


Figure 4.3: Possible future polymer analogues to be synthesized for subsequent click coupling.

Secondly, viability assays should be done to ascertain that there is increased effectiveness in the bactericidal activity of the EpMF2-polymer conjugate compared to just the polymer alone. The viability assay test can be done with unattached polymer as a control experiment, as well as a blank as a negative control. There are some complications in calculating MIC values that should be noted. MIC values for polymers are generally measured by micrograms per mL, as opposed to milli or micromoles per mL, due to the variable molecular weight of such high mass polymers. To accurately determine the average molecular mass M_n (number average molecular weight) and M_w (mass average molecular weight) of the polymer, the fraction of alkynes in the polymer attached to the EpMF2 moiety must be known. Also important was the exclusion of Cu ions that might be present from the click reaction. This may be done via repeated dialysis of the polymer solution, or perhaps addition of chelating agents like EDTA.

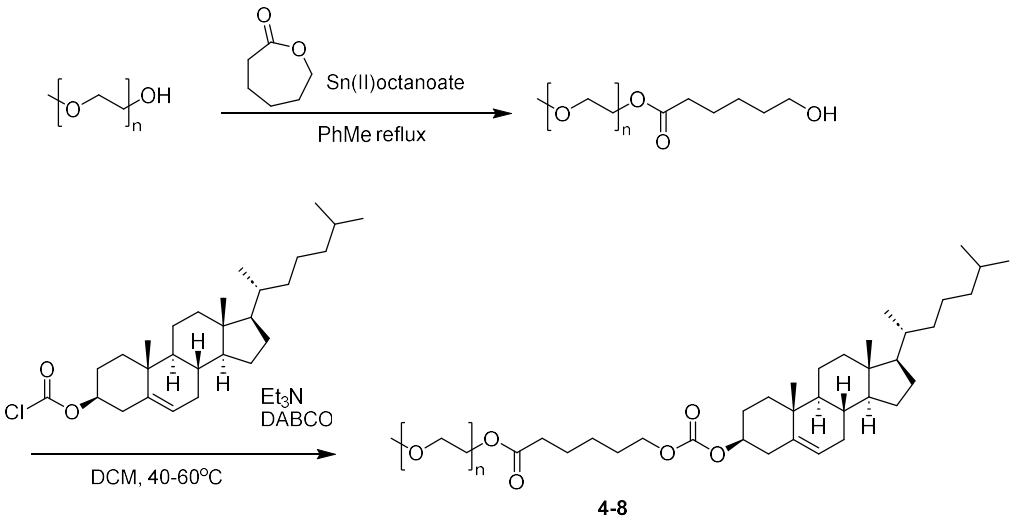
4.1.3 Looking at other EpMF2-cholesterol carriers

One major drawback the EpMF2-cholesterol conjugates synthesised in the previous chapter is its low aqueous solubility hindering dissolution and delivery of the compound to bacteria. An attempt to synthesise a PEGylated like linker as compound **3-11** (as describe in the previous chapter) had problems in stability. However, the linker was likely too short to significantly impact the poor aqueous solubility of the complex. A much longer PEGyated linker is required. This means synthesis of the linker directly from poly(ethylene glycol). Kao et al, has provided a synthesis of a poly(ethylene glycol) derivative **4-7** with carboxylic acid moieties on both ends,[8] where the hydroxyl groups of the cholesterol and EpMF2 analogue could be attached to. However, if this path is pursued work needs to be done to characterise these compounds through HPLC and mass spectroscopy in addition to ^{1}H NMR, to ascertain the extent of coupling by molar percentage, due to variable masses of these high mass polymers.



Scheme 4.4: Synthesis of a PEG linker derivative **4-7** with carboxylic acids at both ends suitable for coupling to cholesterol and EpMF2 analogues

Alternatively, cholesterol might not need to be conjugated to EpMF2 through covalent bonds. Instead, EpMF2 drug can be encapsulated in a micelle or polysome containing a cholesterol moiety. For example, Perez, *et al*. synthesized a cholesterol conjugated polymer **4-8**, that was able to spontaneously form micelles in water and encapsulate amphotericin B, a large molecular weight drug otherwise difficult to capture by micelles, at 60% efficiency.[9]



Scheme 4.5: Formation of cholesteryl mPEG-conjugate **4-8**, which spontaneously forms micelles that can encapsulate drug molecules in the interior.

4.2 Conclusion

Drug discovery for TB remains a challenge even to this day. This is made all the more urgent with rising cases of MDR-TB and XDR-TB.

In Chapter 1, the concept of prodrugs was reviewed, with particular focus on drugs and drug candidates that were marketed or tested as prodrugs. We examine how the use of prodrugs aid in the delivery of the medicine to the desired target, and this will be the main idea behind Chapter 3 in an attempt to improve the delivery of a potential antitubercular compound called EpMF2 to its desired target in *M.smegmatis*.

Chapter 2 describes the discovery and SAR analysis of EpMF2. To develop safer analogues of the antitubercular drug BDQ (a mycobacterial specific ATP synthase inhibitor), EpMF2 was discovered through *in silico* screening of compounds with a similar mechanism of action as BDQ. However, while it was a good inhibitor of ATP synthase *in vitro*, it was inactive against live mycobacteria *in vivo*.

One major obstacle in the development of antitubercular drugs is its inability to penetrate the thick, hydrophobic cell wall layer in *in vivo* assays. To be able to do this the molecule needs a certain level of amphiphilicity like BDQ. However, this may not be possible for drug molecules

whose structure is fined tuned to inhibit specific enzymes. Therefore, in Chapter 3, we took inspiration of the ideas and concepts discussed in Chapter 1 to propose a cholesterol-linker prodrug solution to allow EpMF2 to penetrate the mycobacterial cell-wall. There are a few positive results, which proved the viability of the cholesterol prodrug method, however, this avenue needs further optimisation to deal with the low aqueous solubility of cholesterol conjugates. In addition, the combination of potential drugs with cationic anti-bacterial polymers is also another potential solution still under investigation. It is hoped that discovering an answer to the membrane bypass problem will help expedite the discovery of more drugs to add in the arsenal of tools against this persistent disease.

4.3 References

1. King T, M. P. Green Fluorescent Protein. A molecular tag that can be inserted into genes to make animals and plants glow green. https://www.chm.bris.ac.uk/motm/GFP/GFPh.htm (accessed 20 July 2023).

2. Hellmann, N.; Schneider, D., *Methods Mol Biol* **2019,** *1958*, 379-401.

3. Lippitz, M.; Erker, W.; Decker, H.; van Holde, K. E.; Basche, T., *Proc Natl Acad Sci U S A* **2002,** *99* (5), 2772-2777.

4. Li, C.; Pastila, R. K.; Pitsillides, C.; Runnels, J. M.; Puoris'haag, M.; Cote, D.; Lin, C. P., *Opt Express* **2010,** *18* (2), 988-999.

5. Brancaleon, L.; Durkin, A. J.; Tu, J. H.; Menaker, G.; Fallon, J. D.; Kollias, N., *Photochemistry and Photobiology* **2007,** *73* (2), 178-183.

6. Pages, J. M.; Kascakova, S.; Maigre, L.; Allam, A.; Alimi, M.; Chevalier, J.; Galardon, E.; Refregiers, M.; Artaud, I., *ACS Med Chem Lett* **2013,** *4* (6), 556-559.

7. Mendive-Tapia, L.; Subiros-Funosas, R.; Zhao, C.; Albericio, F.; Read, N. D.; Lavilla, R.; Vendrell, M., *Nat Protoc* **2017,** *12* (8), 1588-1619.

8. Li, J.; Kao, W. J., *Biomacromolecules* **2003,** *4* (4), 1055-1067.

9. Villamil, J. C.; Parra-Giraldo, C. M.; Pérez, L. D., *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **2019,** *572*, 79-87.