

**Design, Synthesis and Biological Evaluation of Novel  
Membrane-Integral Pyrophosphatase Inhibitors**

Aaron Raymond Wilkinson

Submitted in accordance with the requirements for the degree of  
Doctor of Philosophy

The University of Leeds  
School of Chemistry

September 2019

The candidate confirms that the work submitted is his/her own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

© 2019 The University of Leeds and Aaron Raymond Wilkinson

## Acknowledgements

This thesis is dedicated to my Mum who passed away shortly before I started this PhD journey. Without her guidance and support I wouldn't be where I am today. So thank you Mum.

Firstly I would like to thank my supervisor Prof. Colin Fishwick for giving me the opportunity to work on this project. I would also like to thank him for his support over the past three years, which has helped me develop into a better chemist. It's been a pleasure working for you Colin.

I would also like to thank the Goldman group for the collaboration and Keni Vidilaseris for testing my compounds, without whom this project couldn't have happened.

Dr Martin McPhillie deserves a massive thanks for his advice and support over the past few years and for reading the first draft of my thesis.

In addition, I would like to thank my past MChem students (Rob, Pav and Lauren) for their hard work.

I would like to thank my family (Dad, Dave, Alex, Lily and Jacob) for their belief, love and support over the course of my academic career and for putting up with me constantly being stressed.

I wish to thank Ryan Gonciarz, Lewis Turner and Jake Hauser for their help and support when I was in a bad place. You made me realise what's important and helped me a great deal. Glad to have met these life-long friends. Also, for the many memorable times we've had (which I won't get into here!)

A Special mention to Kyle Orritt for being a great guy and helping me through (including sending me some of the greatest memes) the last six months of my PhD, who would of thought that I would have made a friend for life after only six months.

Last but not least I would like to thank the Fishwick group (Chris, Ryan, Heather, Lewis, Ryan, Martin and Kyle) and the Marsden Group (Bobby, Mathew, Dan and Sam) for making the last four years memorable.

## **Abstract**

A larger percentage of the world's population is at risk of developing a protozoan parasitic infection. Current treatments are failing due to the development of anti-parasitic drug resistance and the associated toxicity of these treatments. Therefore, research into new drug targets and drugs is required. The work presented herein aims to identify new drug leads for the treatment of parasitic infections by targeting a membrane-integral pyrophosphatase (M-PPase) using a structure-based drug design approach. Three approaches were employed to identify novel inhibitors of M-PPase, which lead to the identification of three different classes of inhibitors. The first approach involved the SPROUT design of an indole based putative inhibitor. Due to a number of synthetic difficulties this putative inhibitor could not be synthesised. However, simplification of the indole moiety led to the identification of the first inhibitor of a M-PPase that was designed using SBDD. The second approach involved the substrate inspired design of a sulfamide fragment, which had a moderate activity when biologically evaluated against TmPPase. This fragment was designed to mimic phosphate in the hydrolytic centre of M-PPase. Expansion of this fragment using SPROUT led to the identification of a second generation sulfamide fragment, which had an improved potency. Finally, a heterocyclic fragment library was synthesised and biologically evaluated. This library was based upon a heterocyclic fragment identified by collaborators using vHTS. This led to the identification of a potent thiazole fragment. However, the binding site of this fragment could not be identified so further design could not occur.

## Table of Contents

<b>Acknowledgements</b> .....	<b>iii</b>
<b>Abstract</b> .....	<b>iv</b>
<b>Table of Contents</b> .....	<b>v</b>
<b>Abbreviations</b> .....	<b>viii</b>
<b>Chapter 1 Introduction</b> .....	<b>1</b>
1.1 Protozoan Parasitic Infections .....	1
1.2 Toxoplasmosis .....	1
1.2.1 Treatments of Toxoplasmosis .....	3
1.3 Malaria .....	5
1.2.1 Treatments of Malaria .....	6
1.4 Acidocalcisomes .....	8
1.4.1 Functions of Acidocalcisomes in Protozoa .....	9
1.4.2 Hypo-osmotic stress.....	10
1.4.3 Hyper-osmotic stress.....	11
1.4.2 Acidocalcisomes as a Novel Drug Target for Parasitic Infections.....	12
1.5 Membrane-Integral Pyrophosphatases .....	13
1.5.1 Hydrolytic Centre .....	14
1.5.2 Coupling Funnel .....	17
1.5.3 Ion Gate and Exit Channel .....	17
1.5.4 Mechanism of Cation Pumping .....	19
1.5.5 M-PPases as a Novel Drug Target for the Treatment of Parasitic Infections.....	20
1.6 Structure-Based Drug Design .....	21
1.6.1 Substrate-Inspired Design.....	22
1.6.2 <i>De novo</i> Design.....	24
1.6.2.1 SPROUT .....	24
1.7 Project Aims and Objectives .....	21
<b>Chapter 2 <i>De novo</i> design of an Inhibitor of TmPPase</b> .....	<b>5</b>
2.1 SPROUT Design of a Novel Putative Inhibitor of TmPPase .....	25
2.2 Retrosynthetic Analysis of Compound <b>10</b> .....	30
2.3 Attempted Synthesis of Compound <b>10</b> .....	31
2.3.1 Modified Synthetic Route .....	32
2.4 Biological Evaluation of Compounds <b>34</b> and <b>35</b> .....	36
2.5 Alternative Structural Isomer .....	38

2.6 Attempted Synthesis of Compound <b>36</b> .....	39
2.7 Simplification of Core .....	41
2.8 Synthesis of Compound <b>47</b> .....	42
2.9 Biological Evaluation of Compound <b>47</b> .....	43
2.10 Synthesis of Compound <b>52</b> .....	45
2.11 Biological Evaluation of Compound <b>52</b> .....	46
2.12 SAR Exploration of Compound <b>52</b> .....	47
2.13 Analysis of Compound <b>52</b> for Lead-Likeness Using LLAMA .....	48
2.14 SPROUT Expansion of Compound <b>52</b> .....	50
2.15 Retrosynthetic Analysis of Compound <b>60</b> .....	51
2.16 Synthesis of Compound <b>60</b> .....	51
2.17 Biological Evaluation of Compound <b>60</b> .....	52
2.18 Conclusions .....	55
<b>Chapter 3 Substrate-Inspired Design of a Novel Phosphate Mimetic</b> .....	<b>56</b>
3.1 <i>De novo</i> Design of a Novel Sulfamide Fragment .....	56
3.2 Retrosynthetic Analysis of SPROUT Designed Sulfamide Fragment .....	56
3.3 Synthesis of Sulfamide Fragment <b>71</b> .....	57
3.3.1 Optimisation of Chlorination Step and Synthesis of Sulfamide Fragment <b>71</b> .....	58
3.3.2 Alternative Synthesis of <b>71</b> .....	60
3.4 Biological Evaluation of Compound <b>71</b> .....	62
3.5 pK <sub>a</sub> of Sulfamides .....	63
3.6 Synthesis of SAR Library .....	64
3.7 Biological Evaluation of Sulfamide Fragment Library .....	65
3.8 Suitability of Sulfamides for Hit-to-Lead Progression .....	69
3.9 SPROUT Design of Second Generation Sulfamide Fragment .....	71
3.10 Synthesis of Second Generation Sulfamide Fragment .....	72
3.11 Biological Evaluation of Compound <b>112</b> .....	72
3.12 Attempted Synthesis of an Electron Deficient Derivative of <b>112</b> .....	73
3.13 Conclusions .....	75
<b>Chapter 4 Heterocyclic Fragment Series</b> .....	<b>79</b>
4.1 Previous vHTS Campaign .....	79
4.2 Docking of Heterocyclic Fragments <b>121-126</b> .....	80
4.3 Synthesis of Heterocyclic Fragments .....	83

4.3.1 Synthesis of Thiophene <b>122</b> and Thiazole <b>124</b> .....	83
4.3.1 Synthesis of Thiophene <b>122</b> and Thiazole <b>124</b> .....	83
4.3.2 Synthesis of Thiazole <b>123</b> and Oxazole <b>125</b> .....	84
4.3.3 Synthesis of Oxazole <b>126</b> .....	84
4.4 Biological Evaluation of Heterocyclic Fragments.....	86
4.4.1 Biological Evaluation of Methyl Ester <b>130</b> .....	88
4.5 Ligand Efficiencies of Active Heterocyclic Fragments.....	88
4.6 Conclusions.....	89
<b>Chapter 5 Conclusions and Future Work.....</b>	<b>91</b>
5.1 Series 1 .....	91
5.2 Series 2.....	92
5.3 Series 3.....	93
5.4 Overall Conclusions .....	94
5.5 Future Work of Series 1 .....	94
5.6 Future Work of Series 2 .....	96
5.7 Future Work of Series 3 .....	97
<b>Chapter 6 Experimental Section .....</b>	<b>98</b>
6.1 Experimental for High-Throughput TmPPase Activity Assay.....	98
6.1.1 Preparation of Solutions.....	98
6.1.2 Assay .....	98
6.2 Computational Methods .....	100
6.2.1 In Silico Docking Using Glide .....	100
6.2.2 SPROUT .....	101
6.3 General Methods and Instrumentation .....	102
6.3.1 Compound Numbering for NMR.....	103
6.3.2 Synthesis of Series 1 Compounds .....	104
6.3.3 Synthesis of Sulfamide Fragments.....	130
6.3.4 Synthesis of Heterocyclic Fragments .....	140
<b>Chapter 7 References .....</b>	<b>148</b>
<b>Chapter 8 Appendix - NMR Spectra of Novel Final Compounds.....</b>	<b>158</b>

## Abbreviations

**ACT** - Artemisinin Based Combination Therapy  
**AIDS** - Acquired Immune Deficiency Syndrome  
**AQP** - Aquaporin  
**Boc** - tert-butoxy carbonyl  
**cAMP** - Cyclic Adenosine Monophosphate  
**DCM** - Dichloromethane  
**EDC** - 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide  
**Fmoc** - Fluorenylmethyloxycarbonyl  
**HBA** - Hydrogen Bond Acceptor  
**HBD** - Hydrogen Bond Donor  
**HIV** - Human Immunodeficiency Virus  
**IC<sub>50</sub>** - Half maximal Inhibitory Concentration  
**IDP** - Isopentenyl diphosphate  
**LCMS** - Liquid Chromatography - Mass Spectrometry  
**LE** - Ligand Efficiency  
**LLAMA** - Lead-Likeness and Molecular Analysis  
**M-PPase** - Membrane-Integral Pyrophosphatase  
**NBS** - N-bromosuccinimide  
**NMR** - Nuclear Magnetic Resonance Spectroscopy  
**PIFA** - [bis(trifluoroacetoxy)iodo]benzene  
**P<sub>i</sub>** - Phosphate  
**PolyP** - PolyPhosphate  
**PP<sub>i</sub>** - Pyrophosphate  
**PyBrOP** - Bromotripyrrolidinophosphonium hexafluorophosphate  
**RT** - Room Temperature  
**RVD** - Regulatory Volume Decrease  
**RVI** - Regulatory Volume Increase  
**SAR** - Structure Activity Relationships  
**SBDD** - Structure Based Drug Design  
**SEA** - South East Asia  
**TEA** - Triethylamine  
**TM** - Transmembrane Helices

**TmPPase** - *T. maritima* M-PPase

**Tr** -Triphenylmethyl

**V-ATPase** - Vacuolar-type H<sup>+</sup> adenosine triphosphatase

**vHTS** - Virtual Highthroughput Screening

**VTC** - Vacuolar Transport Chaperone

**VrPPase** - *V-radiata* M-PPase

**WHO** - World Health Organisation



## Chapter 1: Introduction

### 1.1 Protozoan Parasitic Infections

Parasitic infections, such as malaria and toxoplasmosis, are caused by protozoan organisms. Protozoa are unicellular eukaryotic organisms and their size typically measures between 1-150  $\mu\text{m}$ . [1] They share a number of organelles with animal cells and use similar metabolic pathways. In addition, they also have a number of unique specialised structures that are not found in other eukaryotes, which are required to meet their particular metabolic and environmental needs. There are four different groups of protozoa that are classified based on their method of locomotion (**Table 1**): the flagellates, which use flagella for locomotion; the amoebae that use pseudopodia; the sporozoans, which do not have any obvious means of locomotion; and the ciliates that use cilia to achieve locomotion. [1]

Group	Means of Locomotion	Example Genus
The Flagella	Flagella	<i>Leishmania</i>
The Amoebae	Pseudopodia	<i>Entamoeba</i>
The Sporozoans	No obvious means	<i>Toxoplasma</i> , <i>Plasmodium</i>
The Ciliates	Cilia	<i>Balantidium</i>

**Table 1:** The classification of protozoa and Genus Examples. [1]

### 1.2 Toxoplasmosis

Toxoplasmosis is a parasitic infection that is caused by *Toxoplasma Gondii*. It is the most widely distributed of all the parasitic infections and infects patients on every continent, with the exception of Antarctica. Toxoplasmosis is also one of the most prevalent human infections, with an estimated 30-50% of people worldwide thought to be infected. [2]

Individuals who have a normal functioning immune system are mainly asymptomatic when infected with toxoplasmosis. However, toxoplasmosis can be harmful to immunocompromised individuals, such as human

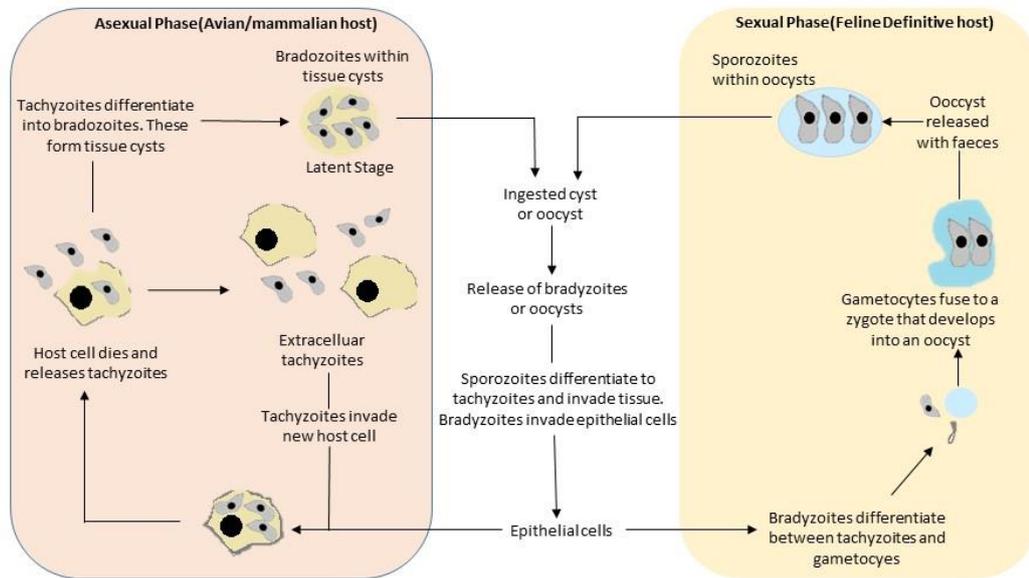
immunodeficiency virus (HIV) and chemotherapy patients, and the unborn foetus. [3] There are two main clinical manifestations of toxoplasmosis in the immunocompromised. Firstly, toxoplasmosis can cause retinal lesions leading to a condition called ocular toxoplasmosis and can cause blindness if not treated. [4] The other main clinical manifestation is toxoplasmic encephalitis (swelling of the brain), which is a leading cause of mortality in acquired immune deficiency syndrome (AIDS) patients. [5] There are a number of symptoms associated with toxoplasmic encephalitis, such as dementia, seizures and headaches. [6]

The life cycle of *T. gondii* consists of a sexual phase and an asexual phase. [7] The sexual phase occurs in the intestine of felines (defined as the definitive host), whilst the asexual phase can occur in any nucleated mammalian or avian cell.

The sexual phase of the life cycle begins when felines ingest prey tissue contaminated with *T. gondii* cysts (Figure 1). Once the cysts enter the feline stomach the slowly dividing form of *T. gondii*, termed bradyzoites, are released. Bradyzoites then invade the epithelial cells of the feline small intestine, where they differentiate into rapidly dividing tachyzoites as well as male and female gametes. These gametes combine to form zygotes that develop into millions of oocysts, which contain thousands of sporozoites. The formation of these oocysts results in the rupture of the feline intestinal epithelial cells, which causes the oocysts to be excreted with the feline faeces and ends the sexual phase of the life cycle. [7]

The asexual phase of the life cycle begins when oocysts are ingested by mammals or birds. When the ingested oocysts reach the stomach of the mammalian or avian host the sporozoites are released and they invade nearby epithelial cells (**Figure 1**). Here, they differentiate into the rapidly replicating tachyzoites. The tachyzoites then replicate inside the cell until it ruptures and they are released into the blood stream, where they can invade nearby cells or move to a different part of the organism. This stage of the infection is called the acute phase and is responsible for the symptoms experienced by immunocompromised patients. The acute phase is followed by the chronic phase, where the tachyzoites differentiate into slowly dividing

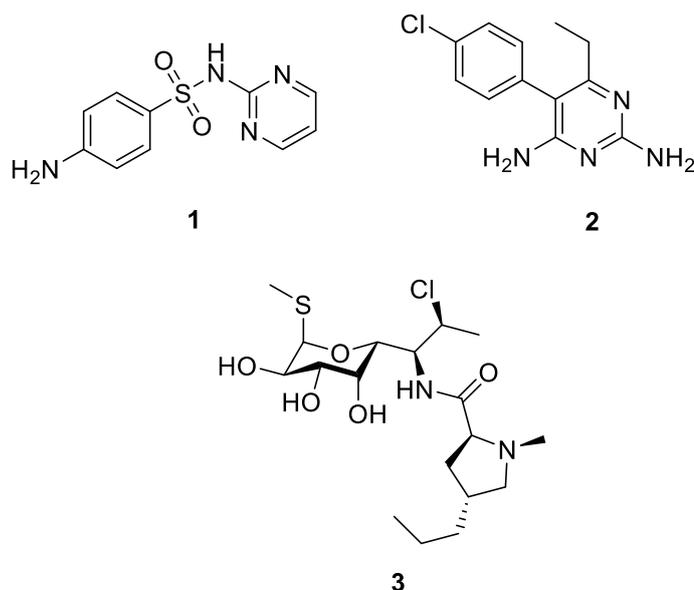
bradyzoites. The bradyzoites form cysts within the hosts' tissue, mainly in the brain and muscle tissue. These tissue cysts can then be transferred back to the feline host and thus beginning the protazoic life cycle again. [7]



**Figure 1;** The lifecycle of *T. Gondii*. There are two phases to the lifecycle; A sexual phase (right) and an asexual phase (left). [8]

### 1.2.1 Treatments for Toxoplasmosis

Current treatments for toxoplasmosis only target the rapidly dividing tachyzoites during the acute phase of infection. [9] One of the most widely used treatments of toxoplasmosis is a combination of a sulfonamide drug, such as sulfadiazine (Compound 1, **Figure 2**), with pyrimethamine (Compound 2, **Figure 2**). [10] This combination sequentially inhibits dihydropterate synthase and dihydropteroate reductase. These enzymes are involved in the synthesis of folic acid compounds that are needed for the growth and survival of *T. gondii*. [11] An alternative treatment involves a combination of a lincosamide antibiotic, such as clindamycin (Compound 3, **Figure 2**), with pyrimethamine. [12] Clindamycin targets the apicoplast in *T. gondii*, which plays a key role in a number of metabolic processes, such as fatty acid synthesis. [13]



**Figure 2;** Current treatments for toxoplasmosis include a combination of sulfadiazine (1) with Pyrimethamine (2) and a combination of Clindamycin (3) with Pyrimethamine.

A study involving AIDS patients diagnosed with toxoplasmosis showed that there is no statistically significant difference in efficacy between sulfadiazene-pyrimethamine and clindamycin-pyrimethamine combinational therapies. There was a complete or partial response in 68% of patients using sulfadiazene-pyrimethamine, compared with 76% for clindamycin-pyrimethamine. [14]

One of the major disadvantages of both combination therapies is toxicity. A study of 115 HIV patients diagnosed with toxoplasmosis encephalitis that were using the sulfadiazene-pyrimethane combination therapy, found some levels of toxicity in 62% of patients. [15] Some of these side effects were severe enough that 44% of patients had to change treatment. The major side effect of this combination therapy is bone marrow suppression. This can lead to a number of complications, such as anaemia and serious infection. Therefore, folic acid is also administered to prevent bone marrow suppression. [16] Another side-effect of this combination therapy is caused by severe allergic reactions to sulfadiazene. [17] In addition, the clindamycin-pyrimethamine treatment regime has also been found to have similar rates of toxicity. [14] Another disadvantage of both these combination therapies is that they only target tachyzoites and not bradyzoites, which means that patients have to undergo lifelong maintenance therapy to prevent relapse.

However, evidence of relapse has been found in patients using both treatments for lifelong maintenance therapy. One study found that 11% of tested AIDS patients relapsed whilst on the sulfadizene-pyrimethamine treatment, whilst 22% of tested patients relapsed on the clindamycin-pyrimethamine treatment. [14]

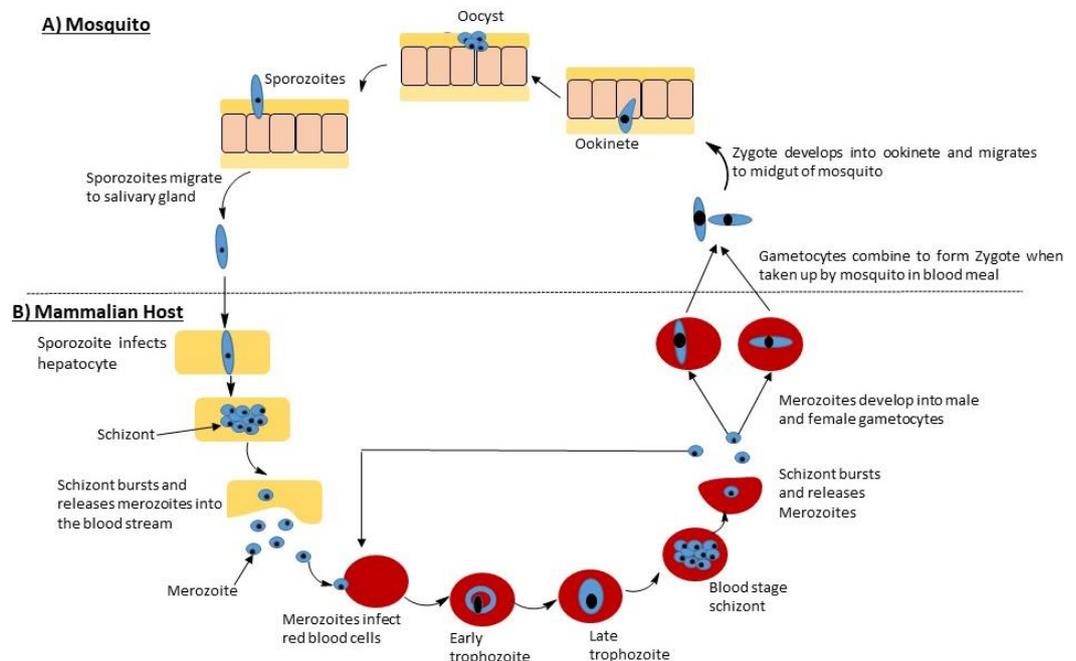
An estimated 10% of patients infected with toxoplasmosis do not respond to current treatments. [10] Resistance of *T.gondii* to current treatments is one explanation as to why these patients do not respond to treatment. However, it is difficult to accurately quantify drug resistance for toxoplasmosis as *T. gondii* is not routinely isolated from patients. It is also difficult to characterise these treatment failures due to a number of issues such as; diagnostic limitations, drug intolerance and non-adherence.[14, 18] However, sulfadiazine resistance has been found in three different strains of toxoplasmosis from across the globe. [19] Resistance of *T. gondii* to clindamycin and pyrimethamine has not been found in clinical isolates to date.

### **1.3 Malaria**

Malaria is a major cause of mortality in tropical regions of the world, particularly in developing countries. There were 216 million reported cases of malaria across the globe in 2016 and 445,000 reported deaths. [20] Malaria is caused by protozoa belonging to the *Plasmodium* genus. While there are five species that can cause malaria in humans, the two that pose the greatest threat are *P. falciparum* and *P. vivax*. *P. falciparum* is the most common cause of malaria infections in sub-Saharan Africa and causes the deadliest form of the disease. Whilst infections caused by *P. Vivax* are not as fatal, it is however the most widely distributed of the plasmodium species and is the largest cause of malaria infections outside Africa. [20] The symptoms of malaria vary depending on the severity of the infection, common symptoms include; shaking chills, high fever, headache, nausea and vomiting. [21]

The *Plasmodium* life cycle consists of an asexual phase, which occurs within a mammalian host, and a sexual phase that occurs within a mosquito (**Figure**

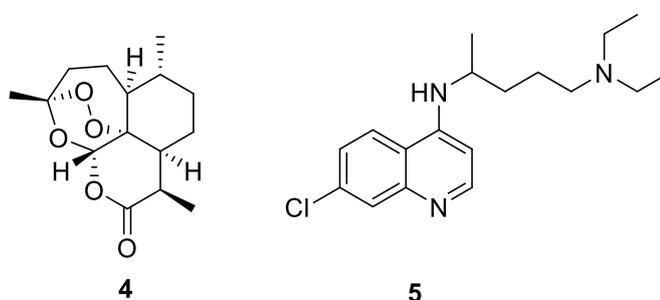
3). [22] The asexual phase of the life cycle begins when *Plasmodium* sporozoites from the saliva of an infected mosquito are injected into the bloodstream of the mammalian host. The sporozoites then migrate to the host liver and invade hepatocytes, thus beginning the hepatic stage of the cycle. Once the sporozoites are taken up by the hepatocytes they differentiate into schizonts that contain thousands of merozoites. The hepatocytes then burst and release the merozoites into the blood stream where they invade and then replicate within red blood cells. This causes the red blood cells to rupture and the merozoites can then invade other red blood cells or differentiate into gametocytes that will transfer back to the mosquito vector. The sexual phase of the life cycle begins within the gut lumen of the mosquito, where the male and female gametocytes combine to form a zygote. [22] The zygote then undergoes meiosis and becomes an ookinete, which then migrates to the midgut wall and transforms into oocysts. After the formation of oocysts mitotic divisions occur and sporozoites are formed. These sporozoites then travel to the salivary gland where they can infect a new mammalian host and begin the cycle again.



**Figure 3:** The life cycle of *Plasmodium*. This life cycle consists of an asexual phase (top) and a sexual phase (bottom). [22]

### 1.3.1 Treatments for Malaria

The World Health Organization (WHO) recommends Artemisinin (Compound **4**, **Figure 4**) based combination therapies (ACT) for the treatment *P. falciparum* based malaria infections. [20] The mechanism of action of artemisinin involves cleavage of the peroxide bond by intraparasitic heme. This generates O centred radicals that rearrange to form C centred radicals. These radicals then damage nearby cell targets by alkylation. [23] Artemisinin is usually given in combination with a drug of another class, such as pyrimethamine. For the treatment of *P. vivax* based malaria infections the WHO recommends chloroquine (Compound **5**, **Figure 4**), which is a derivative of quinine. [20] It is widely accepted that quinine type drugs interfere with the detoxification of heme produced by the degradation of host haemoglobin caused by the parasite. [24] Heme is toxic to the parasite and is normally converted to the non-toxic hemozoin. Quinine type drugs prevent this conversion from occurring and leads to a build-up of heme, which ultimately kills the parasite.



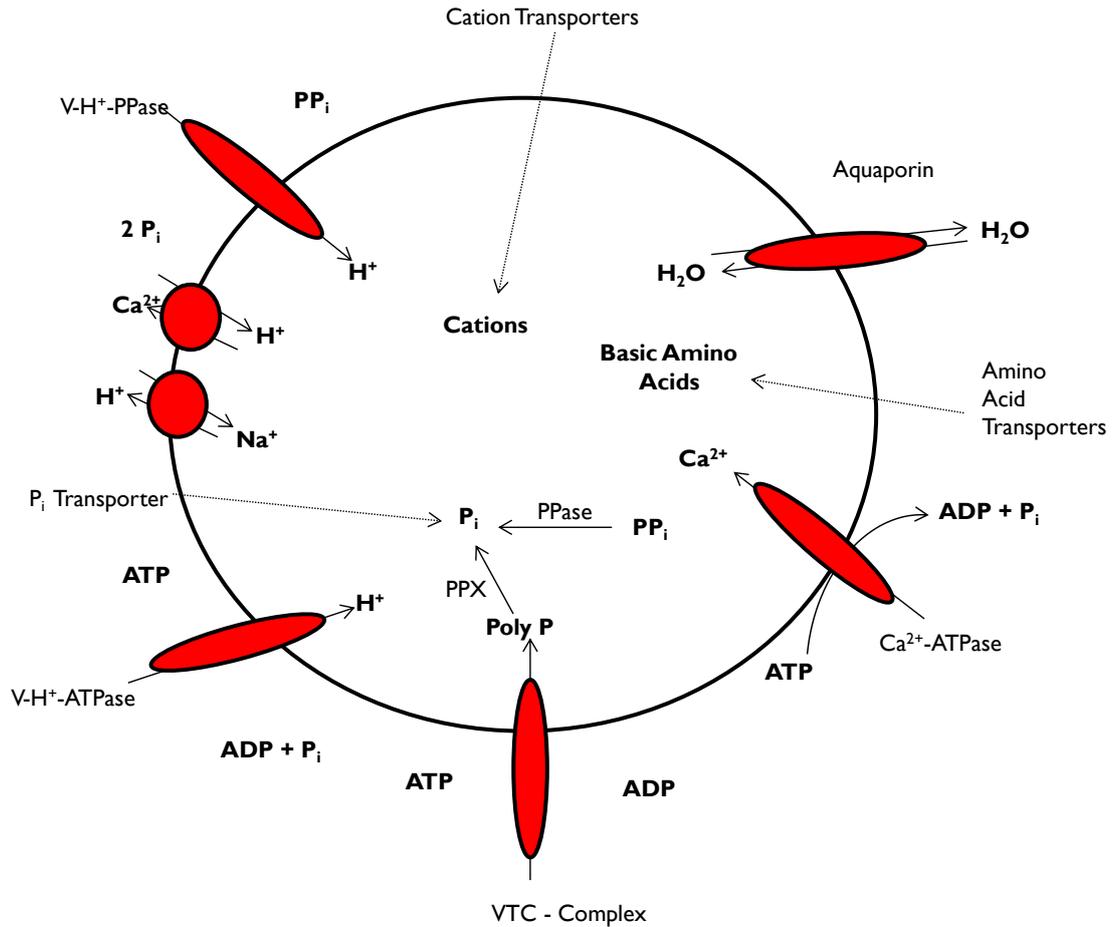
**Figure 4;** The WHO recommends artemisinin (4) based combination therapies for the treatment of *P. falciparum* malaria. Whilst the WHO recommends chloroquine (5) for the treatment of *P. vivax* malaria.

Drug resistance in malaria is quickly becoming a major problem and there is evidence of resistance to all anti-malarial classes currently on the market. [25, 26] ACTs are starting to fail rapidly in South-East Asia (SEA), due to the emergence of artemisinin-resistant *P. falciparum* malaria. [26] This is worrying as ACTs are the first-line treatments for malaria in SEA. Mutations to the kelch13 protein have been found in artemisinin resistant *P. falciparum*. [27] These mutations are thought to mediate artemisinin resistance as some

kelch proteins respond to oxidative stress and artemisinin is a pro-oxidant drug. Resistance of *P. vivax* malaria to chloroquines is also becoming a major problem since it first appeared in 1989. [28] This resistance is thought to be caused by the chloroquine resistance transporter protein (cqr1), which is active in chloroquine resistant *P. Vivax* and *P. falciparum* but not in wild type strains. [29] The cqr1 protein is expressed on the membrane of digestive vacuoles of the parasite and transports chloroquine out of the digestive vacuole, resulting in a lower concentration of chloroquine in the vacuole and rendering it less effective.

#### **1.4 Acidocalcisomes**

Acidocalcisomes are organelles found in protozoa, such as *P. falciparum* and *T. gondii*, they are also found in other eukaryotic organisms. [30] The acidocalcisomes of protozoan parasites are essential for the normal functioning of the cell. They are spherical in shape with average diameters between 0.2-0.6  $\mu\text{m}$ . They are highly acidic and contain a high concentration of calcium ions, which is why they are termed acidocalcisomes. They also contain a high concentration of phosphorous in the form of: phosphate ( $\text{P}_i$ ), polyphosphate (PolyP) and pyrophosphate ( $\text{PP}_i$ ). [30] There are a number of phosphatase enzymes within the acidocalcisomes that hydrolyse the various forms of phosphates into phosphate ions. There is also a vacuolar transporter chaperone (VTC) complex on its membrane that converts cytosolic phosphate ions into acidocalcisome PolyP. In addition to the VTC complex, there are a number of cation pumping and transporting proteins on the acidocalcisome membrane, which are essential to the normal function of the acidocalcisome (**Figure 5**). [31]



**Figure 5:** A schematic representation of the acidocalcisome including the key proteins and pumps. Adapted from [30, 31].

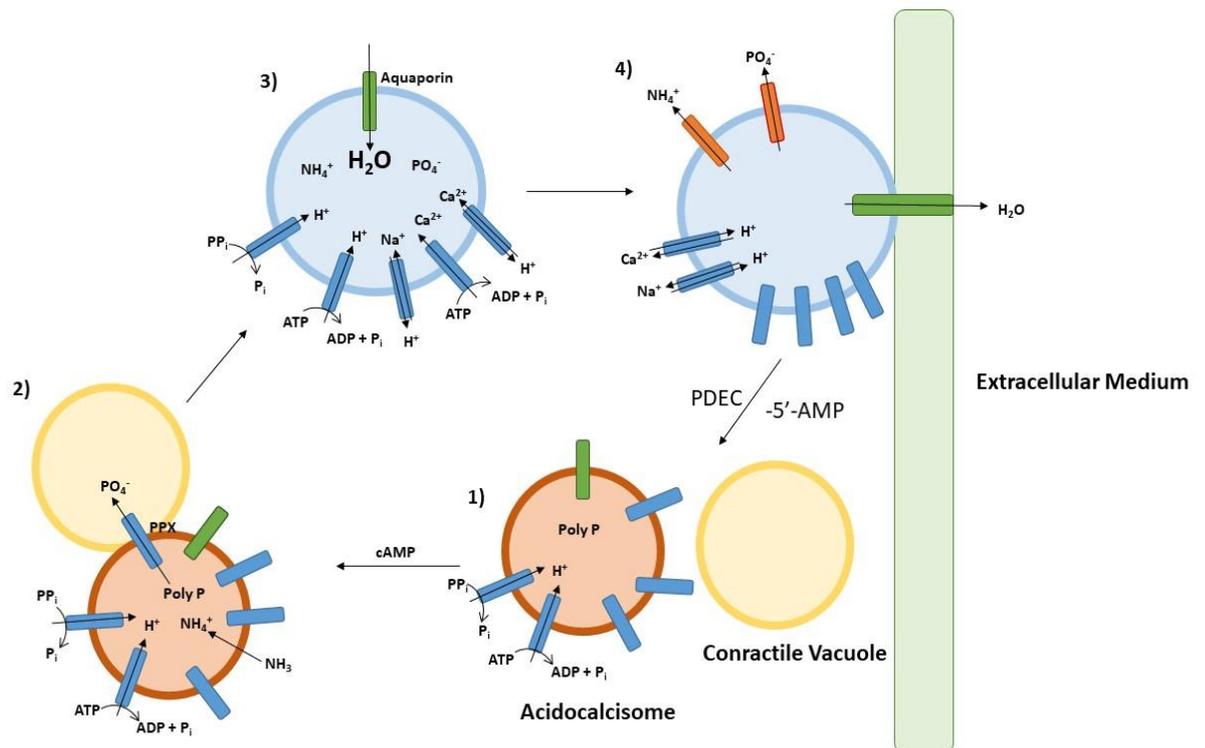
### 1.4.1 Functions of Acidocalcisomes in Protozoa

One of the important functions of the acidocalcisome in protozoan parasites is pH regulation; PolyP plays a key role in this function. When the cell is exposed to alkaline stress PolyP within the acidocalcisome is hydrolysed, which causes the re-acidification of the cell [32]. This is important when the parasite enters the host from the vector as mammalian blood is slightly alkaline.

When a protozoan parasite moves to the host from the vector it encounters osmotic stress. This is due to the difference in osmotic pressure between the host and the vector. The acidocalcisome, together with a contractile vacuole, allows the cell to adapt to these changes in osmotic pressure. The levels of PolyP within the acidocalcisome are also important for osmoregulation [33].

### 1.4.2 Hypo-osmotic Stress

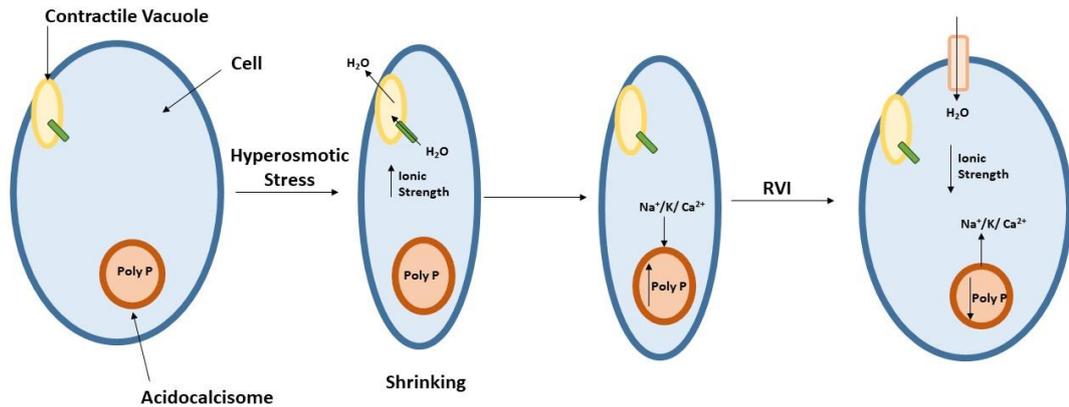
When the cell is exposed to hypo-osmotic stress the protozoan cell swells due to the intake of water caused by a difference in osmotic pressure inside and outside of the cell. During times of hypo-osmotic stress there is a microtubule and cyclic adenosine monophosphate (cAMP) mediated fusion of the acidocalcisome with a contractile vacuole resulting in the translocation of an aquaporin (AQP) from the acidocalcisome to the contractile vacuole (**Figure 6**) [34]. This fusion, together with an accumulation of ammonium ions, causes the hydrolysis of PolyP. [35, 36] This leads to a reduction in the concentration of PolyP in the acidocalcisome. [33] The hydrolysis of PolyP results in the transfer of amino acids,  $P_i$ , cations (osmolytes) and water through the AQP to the contracted vacuole causing it to swell. Water is then released to the extracellular medium and the osmolytes are transferred to the cytoplasm aiding regulatory volume decrease (RVD) [37]. One key piece of evidence for this model is that inhibiting the phosphodiesterase C enzyme, which is involved in the modulation of cAMP, also leads to the inhibition of RVD. [38] This is due to the termination of the cAMP mediated fusion of the acidocalcisome and contractile vacuole. In addition, the fusion of the acidocalcisome and contractile vacuole has also been observed using microscopy. [39]



**Figure 6:** The response of protozoan cells to hypo-osmotic stress. There is a cAMP mediated fusion of the acidocalcisome with the contractile vacuole, which aids RVD. Adapted from [35]

### 1.4.3 Hyper-osmotic Stress

When the cell is exposed to hyper-osmotic stress the protozoan cell shrinks due to the release of water caused by a difference in osmotic pressure inside and outside of the cell. The water is released into the extracellular medium through the aquaporin of the contractile vacuole, causing the cell to shrink (**Figure 7**). [40] During hyper-osmotic stress the concentration of PolyP in the acidocalcisome increases, which complexes to cations in the cytosol. [40] This counteracts the increase in ionic strength caused by the release of water. This model is supported by mutational and inhibitory studies of *T. cruzi* aquaporin (Tcaqp). Inhibition or mutation of Tcaqp results in a decrease in cell shrinkage when exposed to hyper-osmotic stress whilst overexpression of Tcaqp leads to an increase in intensity of cell shrinkage. [40] The cell then undergoes regulatory volume increase (RVI) and the cations return to the cytosol, returning the cell to its pre-hyperosmotic state.



**Figure 7:** The response of protozoa to hyperosmotic stress. The acidocalcisome produces PolyP to help the cell cope with the increase in ionic strength. Adapted from [35].

#### 1.4.4 The Acidocalcisome as a Novel Drug Target for Parasitic Infections

The acidocalcisome plays a key role in the survival of protozoa at times of environmental stress. It therefore offers a novel drug target for parasitic infections. [30] There are a number of proteins located on the membrane of acidocalcisome that offer potential targets for drug design. Research has shown that inhibition of the vacuolar-type H<sup>+</sup> adenosine triphosphatase (V-ATPase) results in a less acidic acidocalcisome, which in turn means that the acidocalcisome is no longer able to function. [41] Huang *et al* also showed that mutation of the V-ATPase in *T.brucei* inhibited growth, as the acidocalcisome was unable to cope with environmental stress.[42] As mentioned earlier, inhibiting phosphodiesterase C prevents cells from undergoing RVD during hypo-osmotic stress and causes cell death. [38] Showing that enzymes involved in RVD are also targets for drug design. Overall, this evidence therefore could suggest that the acidocalcisome is a viable drug target for parasitic infections. This project will target the acidocalcisome in protozoan parasites through the inhibition of membrane-integral pyrophosphatase (M-PPase).

## 1.5 Membrane-Integral Pyrophosphatases (M-PPase)

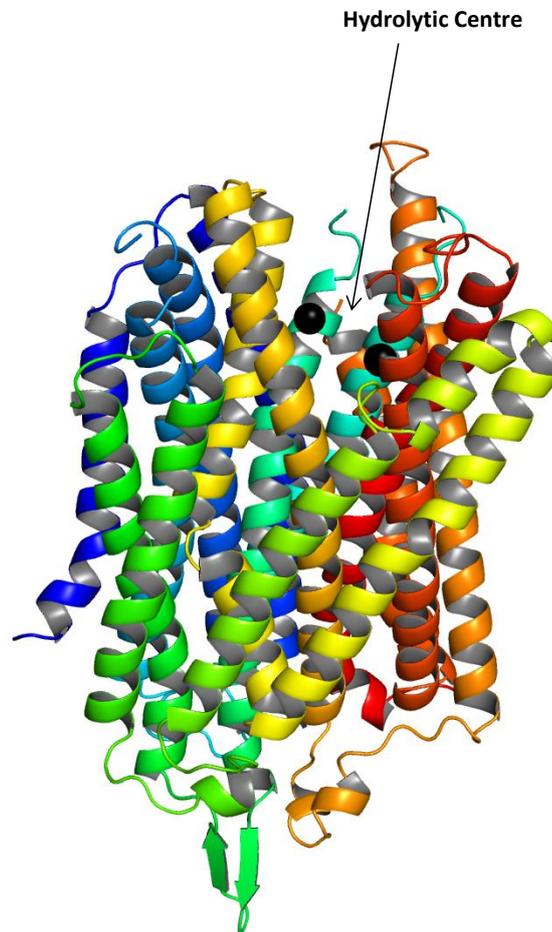
M-PPases are located on the surface of membranes and are found in plants, protozoa and bacteria [43]. They couple  $PP_i$  hydrolysis/synthesis to  $Na^+$  and/or  $H^+$  pumping. There are a number of different sub-types of M-PPase (**Table 2**);  $Na^+$ -pumping  $K^+$ -dependent,  $H^+$ -pumping  $K^+$ -dependent,  $H^+$ -pumping  $K^+$ -independent and  $Na^+/H^+$ -copumping  $K^+$ -dependent. [44] Phylogenetic studies have been carried out on M-PPases and these suggest that  $Na^+$ -pumping M-PPases are likely to be the ancestral proteins, with the other M-PPases evolving from  $Na^+$ -pumping M-PPases later in the evolutionary timeline. [45] M-PPases of protozoa are  $H^+$  pumping and are responsible, along with V-ATPase, for the acidification of the acidocalcisome. [31]

Type of M-PPase	Subtype	Example Organisms
$H^+$ -Pumping	$K^+$ Dependant	<i>Trypanosoma cruzii</i> , <i>Vigna radiata</i>
$H^+$ -Pumping	$K^+$ Independent	<i>Streptomyces coelicolor</i> , <i>Plasmodium falciparum</i>
$Na^+$ -Pumping	$K^+$ Dependant	<i>Thermotoga maritima</i> , <i>Clostridium tetani</i>
$Na^+$ and $H^+$ -Pumping	$K^+$ Dependant	<i>Bacteriodes vulgatus</i>

**Table 2:** The different types and sub-types of M-PPases, as well as example organisms were they are found. [44]

There have been a number of reported crystal structures of M-PPases from *T. maritima* (TmPPase) and *V. radiata* (VrPPase) in various states. The first reported structures were of isopentenyl diphosphate (IDP) bound VrPPase and VrPPase bound with a single phosphate. VrPPase is an  $H^+$ -pumping  $K^+$ -dependant M-PPase. [46] The second reported structures were of resting state TmPPase, IDP bound TmPPase and product bound TmPPase. TmPPase is a  $Na^+$ -pumping  $K^+$ -dependant M-PPase (**Figure 8**). [47] These crystal structures show M-PPases have a single transmembrane domain and

are homodimeric, with each 80 KDa monomer made up 14-17 transmembrane helices (TM). The two crystal structures are highly conserved, showing that the mechanism of hydrolysis is conserved across species.



**Figure 8:** The monomeric structure of TmPPase highlighting the hydrolytic centre. Black spheres are M1 and M2. [47]

The active site of M-PPases are rather unique and are made up of; a hydrolytic centre, a coupling funnel, an ion gate and an exit channel. The hydrolytic centre and coupling funnel are formed from TMs 5-6, 11-12 and 15-16, whilst the gate and exit channel are formed from TMs 5-6, 12 and 16. [46, 47]

### 1.5.1 Hydrolytic Centre

The hydrolytic centre is located at the surface of the pump and is approximately 20 Å above the lipid bilayer. [47] Pyrophosphate binding and hydrolysis takes place at the hydrolytic centre. The key residues that are

involved in metal and pyrophosphate binding in the hydrolytic centre are highly conserved across different species. [48] In the resting state enzyme there are two native  $Mg^{2+}$  cations (termed M1 and M2) that are involved in  $PP_i$  binding and are essential for activity (**Figure 9a**). [47] In the crystal structure of resting state TmPPase M1 and M2 are  $Ca^{2+}$  and  $Mg^{2+}$  respectively. M1 is coordinated by three conserved aspartic acid residues, whilst M2 is coordinated by two conserved aspartic acid residues. The reason  $Ca^{2+}$  is observed in the crystal structure and not  $Mg^{2+}$  is due to the crystallisation conditions. [47]

The hydrolytic centre of IDP bound VrPPase (**Figure 9b**) and IDP bound TmPPase are very similar. [46, 49] These structures show that when substrate binds the volume of the hydrolytic centre is halved, due to the contraction of helices 5, 6, 11, 12, 15 and 16. In addition, the hydrolytic centre is closed due to the closing of a loop between helices 5-6 over the hydrolytic centre.

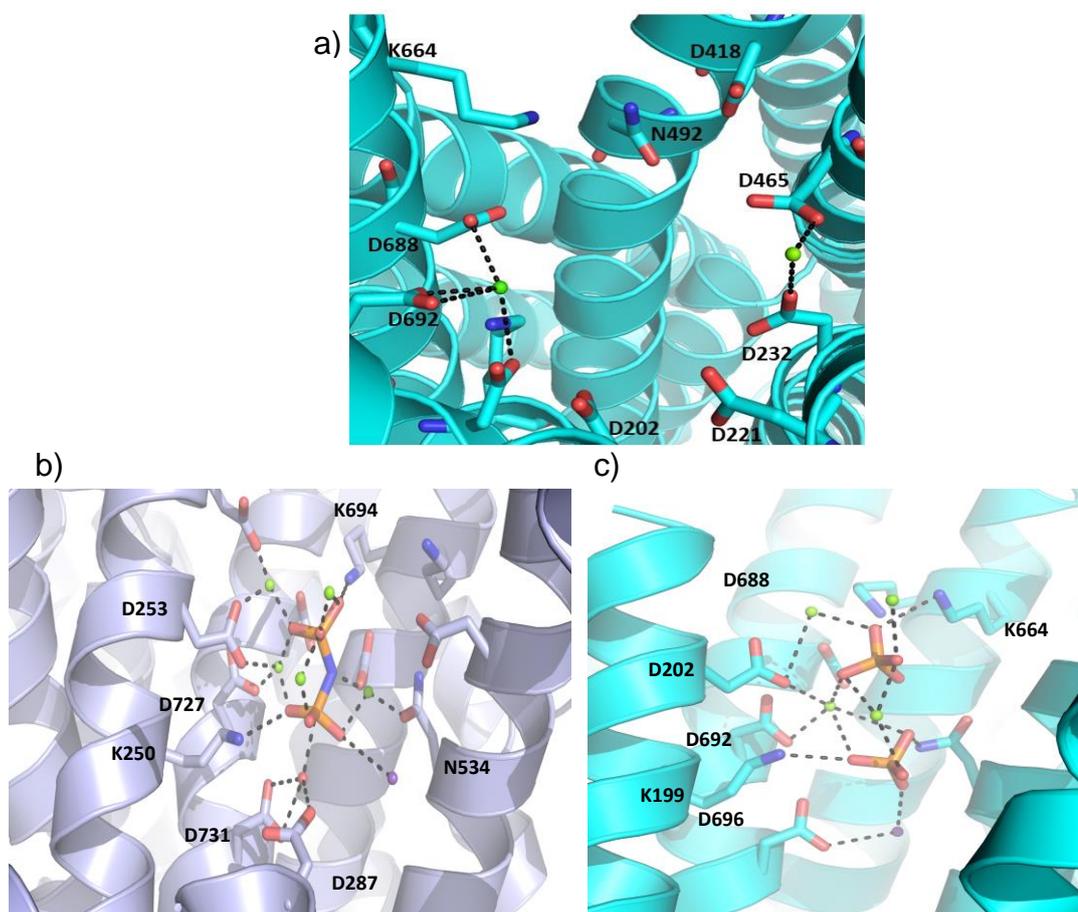
Within the hydrolytic centre of IDP bound VrPPase there are five  $Mg^{2+}$  ions that are coordinated to both IDP and a number of conserved residues. [46] Two of the  $Mg^{2+}$  ions are the native M1 and M2 ions, whilst two of the ions (termed M3 and M4) bind with pyrophosphate in the form  $Mg_2IDP$ . The IDP bound TmPPase crystal structure has four  $Mg^{2+}$  ions, showing that the fifth ion observed in VrPPase is due to the crystallisation conditions. [49]

In both IDP bound structures the leaving group phosphate is coordinated to a conserved lysine residue, making it a better leaving group. In addition, the bottom phosphate group, termed the electrophilic phosphate group, is coordinated to a  $K^+$  ion. Coordination of the bottom phosphate group to  $K^+$  increases the electrophilicity of this group. [46, 49] In  $K^+$  independent M-PPases this electrophilic phosphate group is coordinated by a lysine residue instead, which also increases the electrophilicity. [43]

Positioned underneath the electrophilic phosphate group of both IDP bound structures there is a nucleophilic water molecule that is coordinated by two conserved aspartic acid residues. In the hydrolytic centre of the resting state enzyme one of the aspartic acid residues involved in water binding is

coordinated to a lysine residue, meaning the water molecule is only coordinated by one aspartic acid residue and is not activated for nucleophilic attack. [47] The downwards movement of TM-12 when substrate binds displaces the coordination between the aspartic acid and the lysine, which allows the aspartic acid to coordinate to the water molecule and activate it for nucleophilic attack.

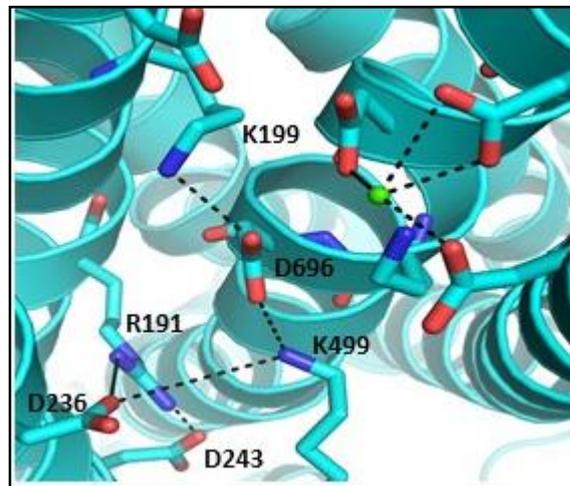
The hydrolytic centre of product bound TmPPase contains two phosphate ions, which are the products of pyrophosphate hydrolysis (**Figure 9c**), and the overall structure is similar to the hydrolytic centre of IDP bound VrPPase and IDP bound TmPPase. [47]



**Figure 9a;** The hydrolytic centre of resting state Tm-PPase. Green spheres donate metal ions. M1 is Mg<sup>2+</sup> and M2 is Ca<sup>2+</sup>. **b)** The hydrolytic centre of Vr-PPase. Green spheres are Mg<sup>2+</sup>, Purple Sphere is K<sup>+</sup>, Red Sphere is nucleophilic water and IDP is shown in orange and blue. **c)** The hydrolytic centre of product bound Tm-PPase. Green spheres are Mg<sup>2+</sup>, purple sphere is K<sup>+</sup> and orange and red sticks are phosphate ions. [46, 47, 49]

### 1.5.2 Coupling Funnel

Below the hydrolytic centre is a coupling funnel that couples pyrophosphate hydrolysis to  $H^+/Na^+$  pumping. [47] This coupling funnel is made up of eight conserved residues and is linked to the ion gate via the Asp residue at the bottom of the funnel (**Figure 10**). There are a number of important interactions between the different residues within the coupling funnel and these form a conformational switch. When substrate binds to the hydrolytic centre TMs; 5, 6, 11, 12, 15 and 16 contract. This contraction causes conformational changes to the coupling funnel that leads to ion pumping. Mutations to any of the conserved residues majorly reduces the pumping activity of the M-PPase. [50]



**Figure 10;** The coupling funnel of Tm-PPase. There are eight conserved residues that make a conformational switch that is dependent on substrate binding. The green sphere is M1 of the hydrolytic centre, shown for perspective. [47]

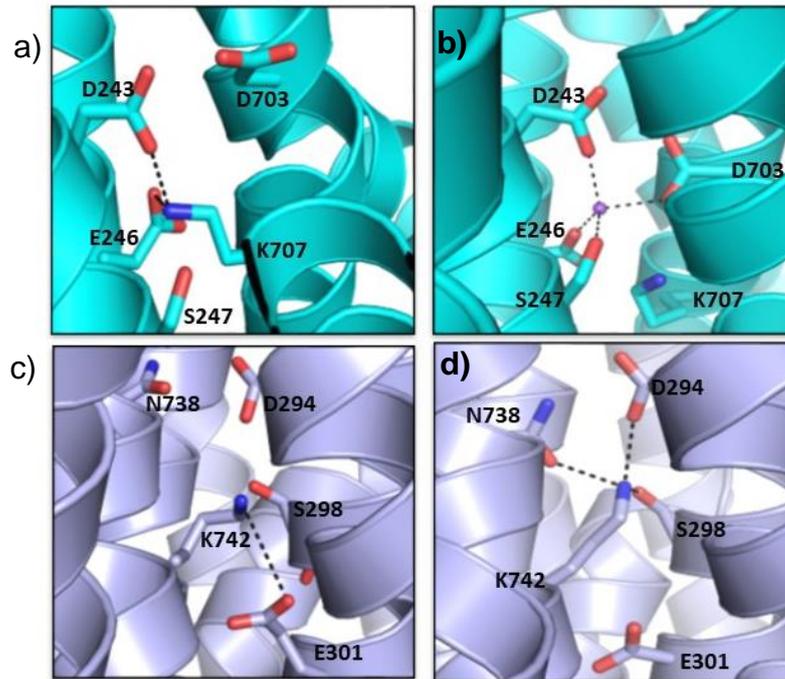
### 1.5.3 Ion Gate and Exit Channel

Below the coupling funnel there is an ion gate where  $Na^+/H^+$  bind. The gate of TmPPase, like all  $Na^+$  pumping M-PPases, is made up of an ion triplet (**Figure 11a**). [47, 49] This ion triplet is made up of: a conserved aspartic acid residue, a conserved lysine residue and a semi-conserved glutamate residue. In TmPPase the exact residues are; Asp243, Lys707 and Glu246. The Asp243 residue is the last residue of the coupling funnel. When substrate binds to TmPPase the salt bridge breaks due to the downwards movement of TMs 11-12 and 15-16, which moves Lys707 out of the  $Na^+$

binding pocket (**Figure 11b**). A sodium ion then coordinates to Asp243 and Glu246, which is supported by coordination to conserved S298. The ion triplet of Na<sup>+</sup> pumping M-PPases are also conserved in H<sup>+</sup>/Na<sup>+</sup> M-PPases and H<sup>+</sup> pumping K<sup>+</sup> independent M-PPases. [43, 45]

In plant K<sup>+</sup> dependant H<sup>+</sup> MPPases, like VrPPase, the ion triplet is replaced with a salt bridge formed between the conserved lysine residue (Lys742) and the semi-conserved glutamate residue (E301) (**Figure 11c**). [46] The Glutamate residue is one helix turn lower than that of TmPPase. When substrate binds to VrPPase, the glutamate residue becomes protonated and is unable to coordinate to the lysine residue and the salt bridge breaks (**Figure 11d**). The lysine residue then coordinates to Asp294 of the coupling funnel. It has been suggest that the extra proton of the protonated glutamate is the one that is pumped upon substrate binding. [43] Analogous to plant M-PPases, the gate of *Flavobacterium johnsonine* type H<sup>+</sup> pumping M-PPases also consists of an ion pair. In these types of M-PPase the conserved glutamate is at an equivalent height to the glutamate in Na<sup>+</sup> pumping M-PPases, but is located on TM5 rather than TM6. [45]

The positioning of the semi conserved glutamate residue seems to be essential for ion specificity. [45-47, 49] The reasoning behind this is the presence of an ion triplet that is observed in Na<sup>+</sup> pumping M-PPases, this ion triplet has an extra formal negative charge compared to the ion doublet observed in certain clades of H<sup>+</sup> pumping M-PPases. The extra negative charge seems to be essential for Na<sup>+</sup> binding. As mentioned earlier there are certain clades of H<sup>+</sup> pumping M-PPases that also have an ion triplet. Here, it is assumed that the glutamate residue is in fact protonated and the pumping mechanism is similar to the plant and *Flavobacterium johnsonine* type M-PPases. Evidence for the importance of the semi conserved glutamate residue for pumping has been obtained using mutagenesis studies. [45] These studies revealed that mutagenesis of the key Glu residue in *Flavobacterium johnsonine* and *Leptospira biflexa* (this is similar to plant M-PPases) type H<sup>+</sup>-PPases uncouples H<sup>+</sup> pumping from PPi hydrolysis, whilst mutagenesis of the key Glu residue in Na<sup>+</sup>-PPases also results in the inhibition of pumping.



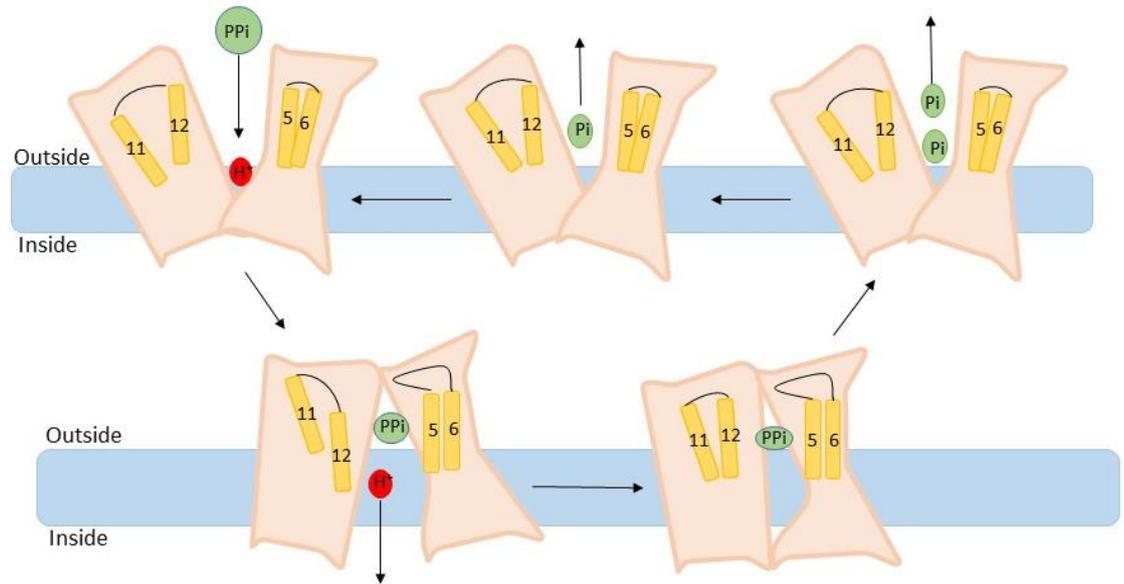
**Figure 11;** (a) The gate of resting state TmPPase consists on an ion triplet made up of the residues D242, K707 and E246. (b) The gate of IDP bound TmPPase. When substrate binds K707 moves out of the Na<sup>+</sup> binding region, which means a Na<sup>+</sup> ion can now bind to D243, E246 and S247. Na<sup>+</sup> is purple. (c) The gate of resting state VrPPase consists of an ion doublet made up of the residues K742 and E301. (d) The gate of IDP bound VrPPase. When substrate bind K742 is no longer able to coordinate to the semi-conserved E301 due to the protonation of E301. [46, 47, 49]

The final feature of the active site is an exit channel for either Na<sup>+</sup> or H<sup>+</sup> ions. This exit channel is the least conserved part of the active site and is closed in all the crystal structures obtained to date. [46, 47, 49]

### 1.6.5 Mechanism of Cation Pumping

All the crystal structures obtained to date and molecular dynamic studies suggest that Na<sup>+</sup>/H<sup>+</sup> pumping is due to pyrophosphate binding to the active site rather than its hydrolysis (**Figure 8**). [46, 47, 49, 51] This binding causes the loop between helices five and six to close over the active site. Contraction of the coupling funnel then occurs and is caused by the downwards movement of TM 5-6, 11-12 and 15-16 toward the centre of the funnel. This allows the ion that is to be pumped to bind at the gate. There is then a large downward movement of TM 12 and a bending of TM 11 at a hinge, which opens the gate and releases the ion. The downward movement of TM 12

also activates the nucleophilic water. Following release of the ion, the gate closes and hydrolysis of PP<sub>i</sub> occurs. This opens the active site and the resultant phosphate ions sequentially leave the active site.

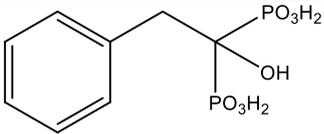
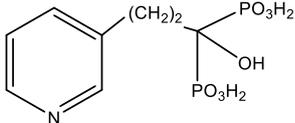
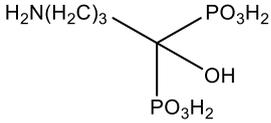
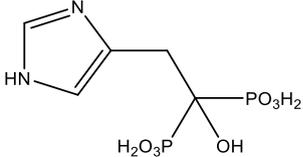
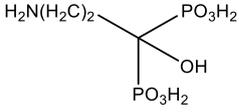


**Figure 12;** The proposed pumping mechanism of M-PPases. Studies suggest substrate binding results in cation pumping and not hydrolysis. Adapted from [47]

### 1.5.5 M-PPases as a Novel Drug Target for the Treatment of Parasitic infections

M-PPases are found on the membrane of acidocalcisomes in protozoa and they play a key role in its acidification and energisation. [30] Therefore, disrupting its function could mean that the acidocalcisome is no longer viable and could provide a therapeutic response. Docampo *et al* mutated a gene coding for M-PPase in *Trypanosoma brucei* and found that the acidity of the mutant acidocalcisomes decreased and also contained a lower concentration of PolyP. [52] This meant that the mutant *T. Brucei* were no longer able to normalise their intracellular pH on exposure to basic conditions, resulting in a slower rate of growth. Liu *et al.* showed that mutation of the gene coding for M-PPase in *T. Gondii* resulted in a slower rate growth both *in vivo* and *in vitro*. [53] They also showed that regulatory volume decrease during hypo-osmotic and hyper-osmotic stress is impaired in these mutants. Bisphosphonates, which are pyrophosphate analogues that inhibit M-PPases, have been shown to inhibit the growth of a number of

different protozoa. [54] In these compounds the central oxygen has been replaced with a carbon, giving a non-hydrolysable phosphoether bond. The central carbon atom can have two other functional groups bonded to it. These bisphosphonates have been shown to inhibit the growth of; *P. falciparum*, *T. brucei*, *T. cruzi* and *Leishmania donovani* (**Table 3**). However the mechanism of action is complex and not widely understood as bisphosphonates also inhibit the enzyme farnesyl pyrophosphate synthase. So, the observed inhibition of growth could be due to inhibition of farnesyl pyrophosphate rather than M-PPase (or inhibition of both). M-PPases are also not found in mammalian cells, which is also beneficial to drug design.

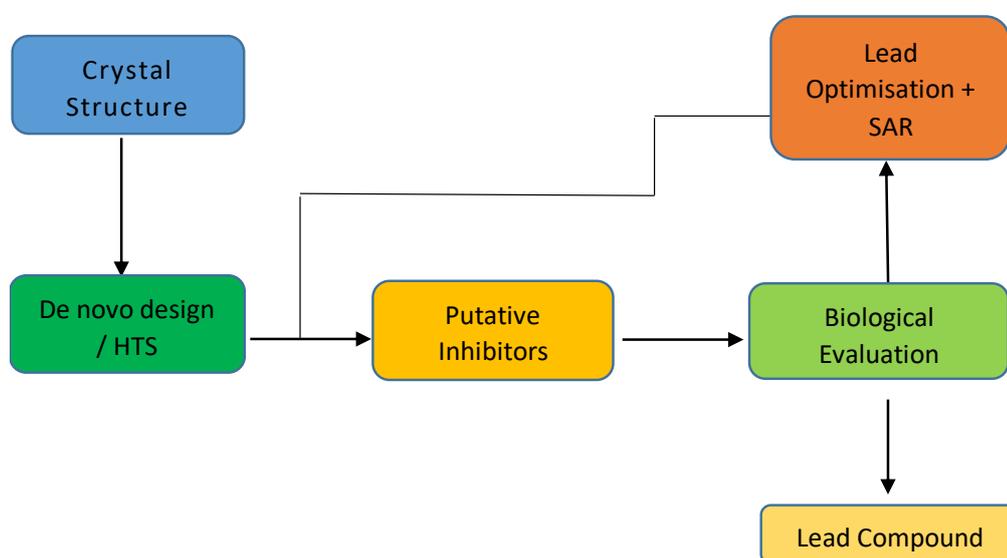
Compound	Protozoa	IC <sub>50</sub>
	<i>P. falciparum</i>	7.7 μm
	<i>T. brucei</i>	1.7 μm
	<i>L. donovani</i>	82.5 μm
	<i>T. cruzi</i>	35 μm
	<i>T. gondii</i>	23 μm

**Table 3**; Examples of Bisphosphonates and their activities against different protozoa [54]

## 1.6 Structure Based Drug Design

In order to develop potential inhibitors of M-PPases, this project will use a structure based drug design (SBDD) approach. SBDD uses the three-dimensional crystal-structure of a protein, as well as various *in silico* techniques to identify potential inhibitors of that protein.[55] There are three

main methods of SBDD; substrate-inspired design, virtual high throughput screening (V-HTS) and *de novo* design of inhibitors.[56] Once potential inhibitors have been identified using one of the SBDD methods, they are synthesised in the laboratory or purchased, if available commercially. This allows biological evaluation of the compounds to take place in order to establish the level of inhibition against the target enzyme. Once 'hit' compounds have been identified, optimisation of the compounds can then take place, both to improve the level of inhibition and the physicochemical properties. Structure activity relationships can also be determined during the lead optimisation stage (**Figure 13**).



**Figure 13:** The Structure based drug design work-flow

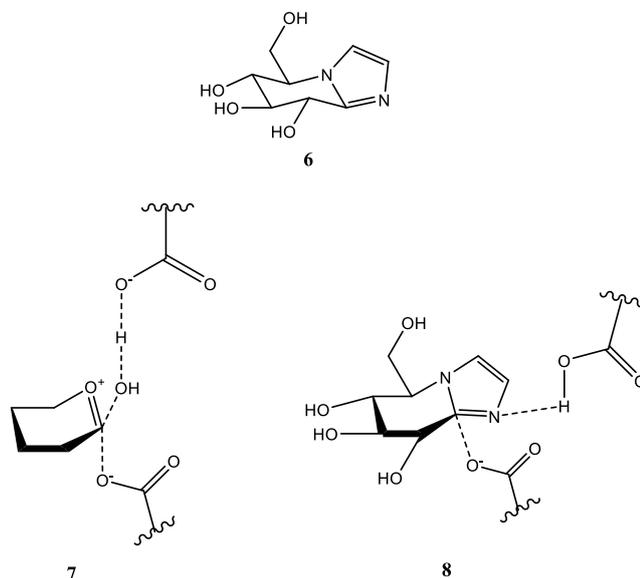
### 1.6.1 Substrate-Inspired Design

Substrate-inspired design uses the crystal structure of the target protein bound with a known inhibitor or substrate. Modifications can then be made to the inhibitor/substrate with the aim of making a more potent inhibitor.[57] An example of how to make an inhibitor more potent is to add additional functional groups to increase the number of interactions between the inhibitor and target protein.

One of the computer programmes that can be used for substrate-inspired design is called Maestro, which allows the user to visualise the three-dimensional crystal structure of the substrate bound protein and examine

any interactions between the two. [58] The programme also allows the user to make changes to the substrate, giving a new putative inhibitor. The new putative inhibitor can then be docked using Glide (grid-based ligand docking with energetics), which is a programme that is built into Maestro. Glide initially uses a series of hierarchical filters that explore the conformational, orientational and positional space of the ligand to be docked. This is then followed by a rough scoring phase. Any poses that are unsuitable for binding to the active site are discarded. The different binding poses that were not discarded are then further refined and the best poses then undergo energy minimisation. Finally, each pose is given a score that is based on how the designed molecule interacts with the protein.[57]

One example of the use of Maestro in substrate inspired design was carried out by Wang *et al*, who developed a series of  $\beta$ -glycosidase inhibitors (**Figure 14**).[59] They designed and synthesised a series of gluco-configured tetrahydroimidazopyridines (Compound **6**, **Figure 14**). These compounds were based on the transition state of the enzyme catalysed reaction (Compound **7** & **8**, **Figure 14**).



**Figure 14:** (6) An example of a tetrahydroimidazopyridine developed by Wang. (7) Transition state of  $\beta$ -glycoside hydrolysis bound to  $\beta$ -glycosidase inhibitors. (8) Proposed mechanism of inhibition of tetrahydroimidazopyridine. Which mimics the transition state of the  $\beta$ -glycosidase catalysed reaction.

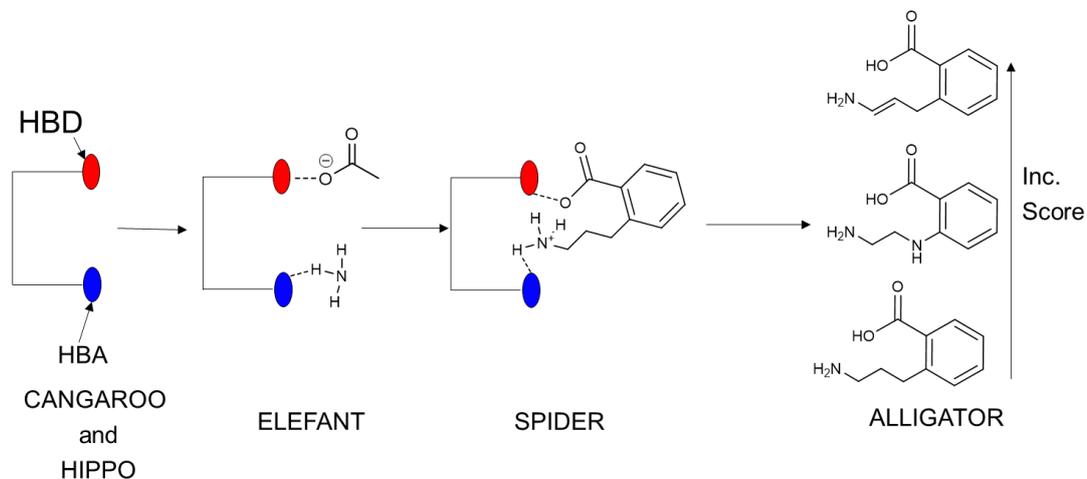
## 1.6.2 *De novo* Design

*De novo* based design was first developed in the early 1990's and uses *in silico* methods to design ligands for a particular target site. Since then its use has become more widespread this is due to improvements in both computer technology and X-ray crystallography.[60] There are a number of steps involved in the *de novo* design process.[57] The first step involves the analysis of the target site, which requires the crystal structure of the protein. Once the target site has been identified, molecular building blocks obtained from a variety of libraries are positioned onto the target site. The building blocks are then connected in a variety of ways to generate a number of putative inhibitors. These putative inhibitors are then assessed and given a score. The user then selects a number of these molecules to synthesise and biologically evaluate.

### 1.6.2.1 SPROUT

SPROUT is a *de novo* design programme that is made up of five different modules, each of which has a different function: CANGAROO (cleft analysis by geometry based algorithm regardless of the orientation), HIPPO (hydrogen bonding interaction site predication as positions with orientations), ELEFANT (election of functional groups and anchoring them to target sites), SPIDER (structure production with interactive design of results) and ALLIGATOR (analyse lots of ligands test and order results) (**Figure 13**). [61] The first module, CANGAROO, defines the receptor site and the ligand cavity site. Once potential binding sites have been identified, HIPPO is used to identify hydrogen bond donor and acceptor sites, as well as any hydrophobic regions. The user then selects some of these sites and then docks small fragments onto those selected sites using the ELEFANT module. These fragments are then connected together in SPIDER using an inbuilt library of 'spacer' fragments to generate a number of different structures. ALLIGATOR then ranks these skeletons according to their binding score. The SPROUT score is the sum of the hydrogen bonding interactions, the hydrophobic interactions, the Van der Waals interactions and the rotatable bonds of the putative inhibitor. The Fishwick group recently applied SPROUT to the design of selective inhibitors of *Plasmodium* Dihydroorotate dehydrogenase,

showing that SPROUT can successfully be applied to the design of novel inhibitors. [62]



**Figure 15:** The SPROUT *de novo* design programme consists of 5 modules that allows the design of novel putative inhibitors of a protein using the crystal structure of that protein.

## 1.11 Project Aims and Objectives

The overall objective of the research presented in this thesis was to identify novel inhibitors of M-PPases using SBDD. To achieve this objective this research had the following aims;

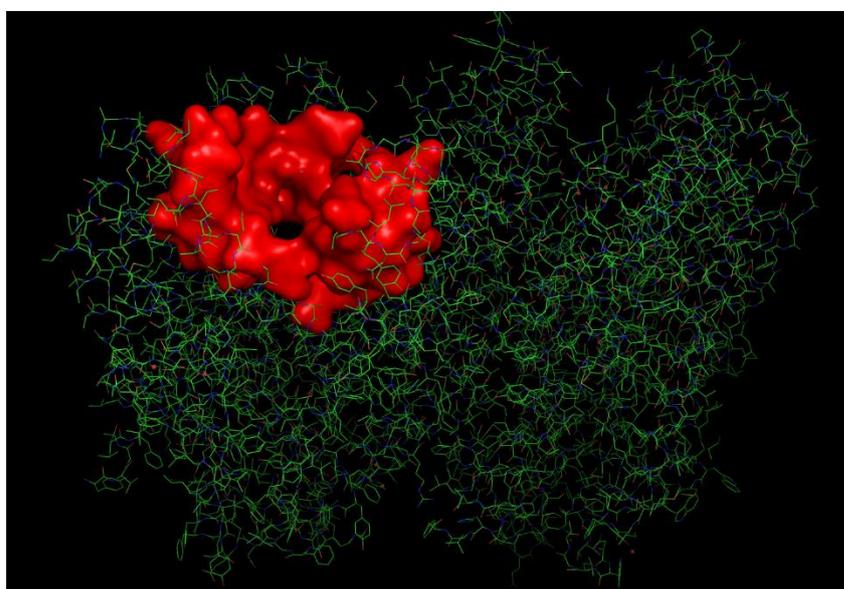
- To use a combination of *de novo* design and substrate inspired design, along with the crystal structure of TmPPase, to identify novel putative inhibitors of TmPPase.
- To develop robust chemical synthetic routes, using modern synthetic techniques, to synthesise any compounds identified during the design stage and to synthesis chemical libraries to determine the SAR of any active fragments.
- To biologically evaluate any synthesised compounds against TmPPase.
- To evaluate the biological results and select suitable candidates for hit-to-lead progression

## Chapter Two – *De Novo* Design of an Inhibitor of TmPPase

### 2.1 SPROUT Design of a Novel Putative Inhibitor of TmPPase

The *de novo* design programme, SPROUT, was used to develop a putative inhibitor of TmPPase and the design process is outlined below.

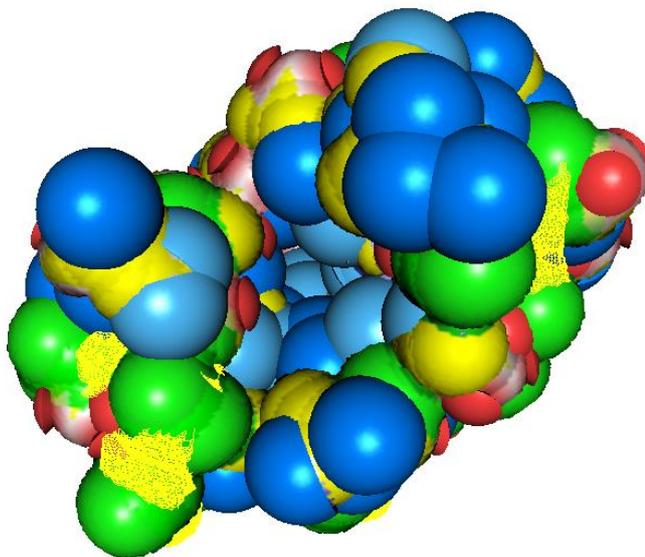
Firstly, the resting state crystal structure of TmPPase (PDB: 4AV3) was loaded into SPROUT and the CANGAROO module was used to make a 10Å cut of the hydrolytic centre (**Figure 16**). [47] This 10Å cut was then redefined as the receptor. [61]



**Figure 16:** The CANGAROO module of SPROUT was used to make a 10Å cut of the hydrolytic centre of TmPPase (PDB:4AV3), which was then redefined as the receptor. [47]

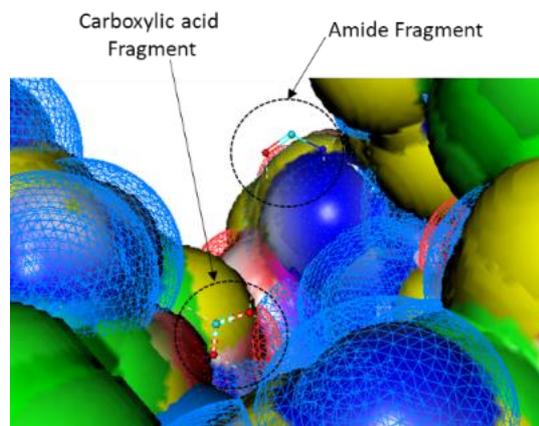
The HIPPO module was then used to identify potential interaction sites, such as hydrogen bond acceptor and donor sites, by analysing the defined receptor for hydrophobic and electrostatic forces (**Figure 17**). HIPPO does this by adding a boundary surface to the protein, which is based upon the Van der Waals radii of each atom. [61] A loci is then created around each polar atom that defines space where a favourable polar contact may be made when a complementary fragment is placed there. These loci are shown as coloured spherical sites, where the colour determines the type of contact that can be made. The colours are; red for hydrogen bond donors (HBD), blue

for hydrogen bond acceptors (HBA) and purple for other contact sites (e.g metal ions). Hydrophobic areas are also shown, these areas are green.



**Figure 17:** HIPPO was used to analyse the defined receptor for potential interaction sites. Hydrogen bond acceptor sites are blue, hydrogen bond donor sites are red and hydrophobic sites are yellow. [61]

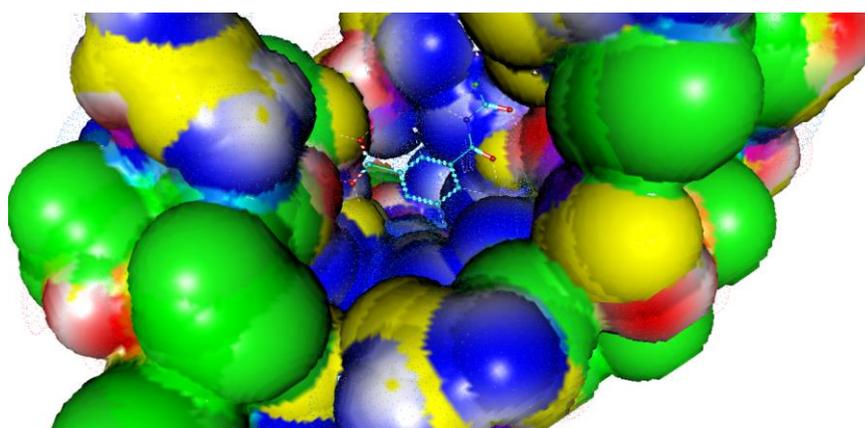
Fragments, which are located within an in built library, were then docked onto two of the previously identified interaction sites using the ELEFANT module (**Figure 18**). The two chosen sites were a hydrogen bond acceptor site of Asp-488 and a hydrogen bond donor site of Lys-695. The chosen fragments that were docked onto the Lys interaction site included; a carboxylic acid, an ester, an alcohol and an amide fragment, whilst an amide and an amine fragment were docked onto the Asp site. [61]



**Figure 18:** The ELEFANT module was used to dock fragments onto the interaction sites of Lys and Asp. The two examples shown are a carboxylate docked onto Lys and an amide docked onto Asp. [61]

The fragments were then connected using a number of different spacer templates within the SPIDER module to generate > 100,000 structures. The spacer templates that were used included; SP3 and SP2 carbon, amide, ester, ether, phenyl, cyclohexyl, five membered aromatic rings, indole, five membered heterocycles and amine linkages.[61]

Once the structures were generated they were scored and ranked using ALIGATOR, the highest scoring fragment is shown below (**Figure 19**). The fragments were then pruned so only the top 100 scoring fragments remained.



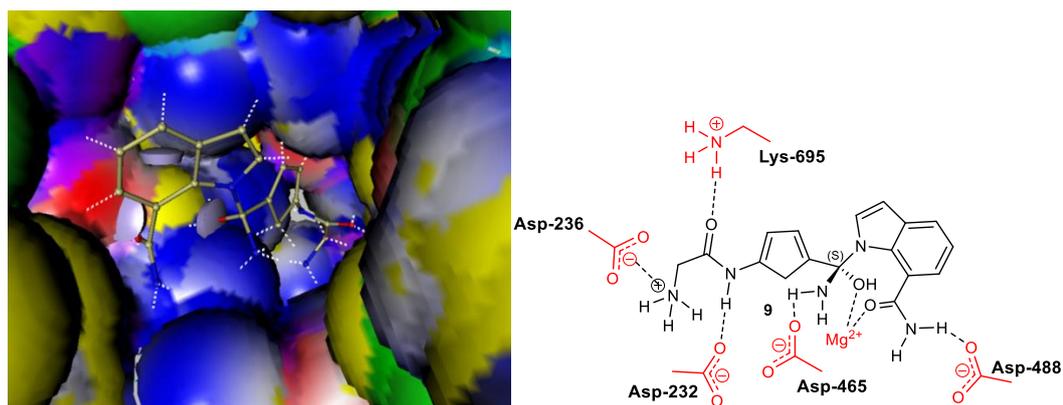
**Figure 19:** The highest scoring fragment (shown in light blue) generated using SPIDER. [61]

Further cycles of design then followed that involved expanding the fragments, with the aim of making further interactions with the receptor. The interaction sites, fragments and spacer templates that were chosen during the iterative cycles of design are summarised below (**Table 4**). [61]

Interaction Site	Fragment Docked	Spacer Templates Used
Asp-465	Amine, Amide	Five membered heterocycle, Indole and SP3 Carbon
Asp-236	Amine, Amide	Amide, Ether, SP3 Carbon, and Benzene
Asp-232	Amine, Amide	Five membered heterocycle and Phenyl

**Table 4:** The chosen interaction sites, fragments and spacer templates used during the cycles of design.

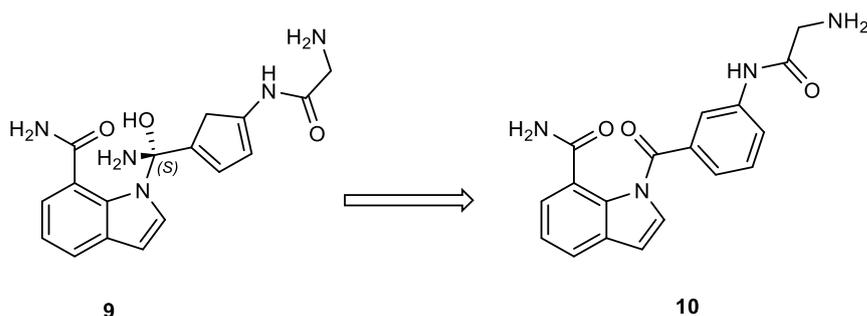
At the end of the design process there were 20 structures generated. The top ten scoring scaffolds were then docked into the hydrolytic centre of TmPPase (PDB: 4AV3) using the Glide docking programme with the Maestro suite. [47, 63] The resultant docking poses were identical to the SPROUT binding poses, which therefore validated the SPROUT binding poses. Visual inspection of these scaffolds was then carried to identify scaffolds that had good docking scores and good synthetic accessibility. This led to the identification of compound **9**, which had a SPROUT score of -5.90 and was selected for further analysis (**Figure 20**).



**Figure 20:** Docking pose of SPROUT designed compound **9** visualised in SPROUT and the 2D representation. The structures highlighted in red for the 2D representation are amino acid residues and metal ions of the hydrolytic centre of TmPPase, whilst dashed lines represent interactions between **9** and the hydrolytic centre. [47, 63]

Compound **9** was predicted to coordinate to the  $Mg^{2+}$  of the hydrolytic centre via the carbonyl of the primary amide and the hydroxyl group of the chiral carbon. In addition, the primary amide was also predicted to form a hydrogen bonding interaction with Asp-488. Hydrogen bonding interactions were predicted to form between the amide of the glycine moiety with Lys-695 and Asp-232. The protonated primary amine group of the glycine moiety was predicted to form an ionic interaction with Asp-236. Lastly, the amine of the chiral carbon was predicted to form a hydrogen bonding interaction with Asp-465.

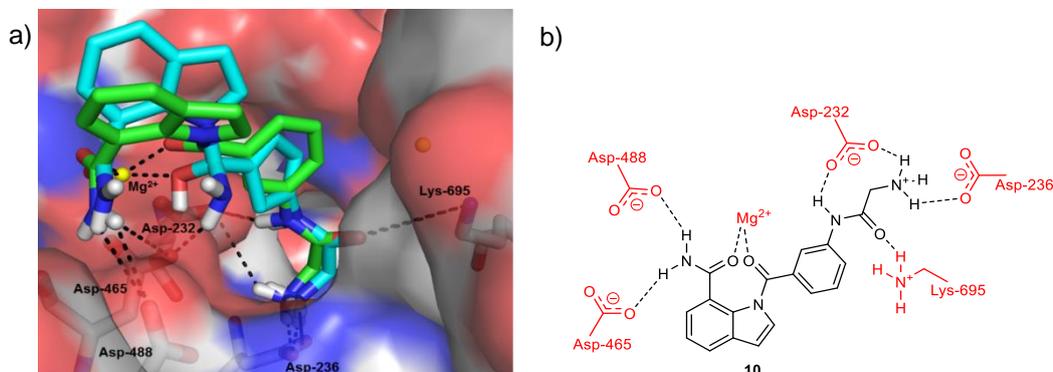
A number of changes were made to compound **9** to make it more synthetically accessible (**Figure 21**). The chiral carbon of compound **9** bears three heteroatoms (including a primary amine and a hydroxyl group), which is hydrolytically unstable and undesirable for a potential drug lead. As mentioned above, the docking pose of **9** predicts that the hydroxyl group coordinates with one of the  $Mg^{2+}$  ions of the hydrolytic centre. In order to retain this interaction and increase the hydrolytic stability, the chiral carbon was replaced with a carbonyl. This resulted in the generation of an *N*-acyl indole. In addition, the cyclopentadiene was replaced with a phenyl ring to make the putative inhibitor more synthetically accessible. These changes lead to the identification of compound **10**.



**Figure 21:** A number of changes were made to **9** to make it synthetically accessible, this lead to the identification of compound **10**.

Compound **10** was docked into the hydrolytic centre of TmPPase and the resultant docking pose was compared to that of compound **9** (**Figure 22**). [63] As predicted the docking poses of both compounds were similar, with similar docking scores (-6.78 for **9** and -6.69 for **10**), and both are predicted

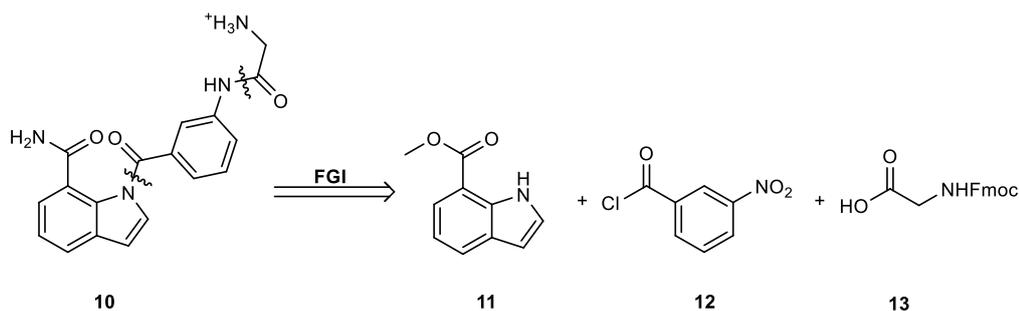
to make similar contacts within the hydrolytic centre. The indole *N*-acyl carbonyl of **10** is predicted to coordinate with one of the  $Mg^{2+}$  ions and there is the loss of a contact with Asp-465 that arises from the loss of the chiral amine group.



**Figure 22:** a) The predicted binding pose of compound **10** (green) and compound **9** (blue) within the hydrolytic centre of TmPPase (PDB : 4AV3). [47] Predicted interactions are shown as dashed black lines. b) The 2-D representation of the docking pose of compound **10**. Residues of the hydrolytic centre are shown in red and interactions are shown as black dashed lines. [64]

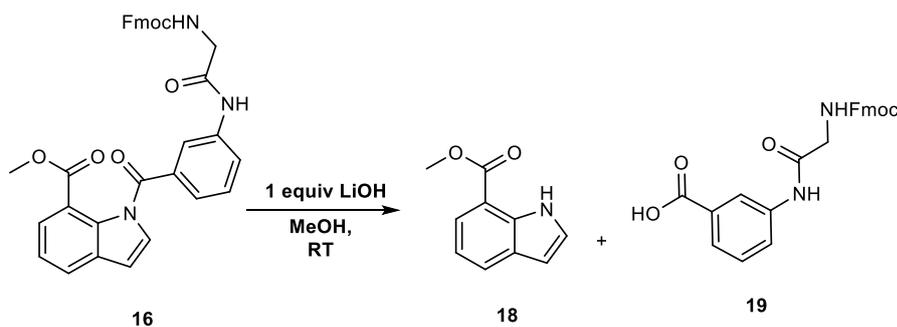
## 2.2 Retrosynthetic Analysis of Compound 10

Retrosynthetic analysis of **10** suggested that this target compound could be synthesised from the commercially available indole **11**, acyl chloride **12**, and *N*-Fmoc glycine **13** (Scheme 1).



**Scheme 1:** Retrosynthetic analysis of SPROUT designed compound **10**.

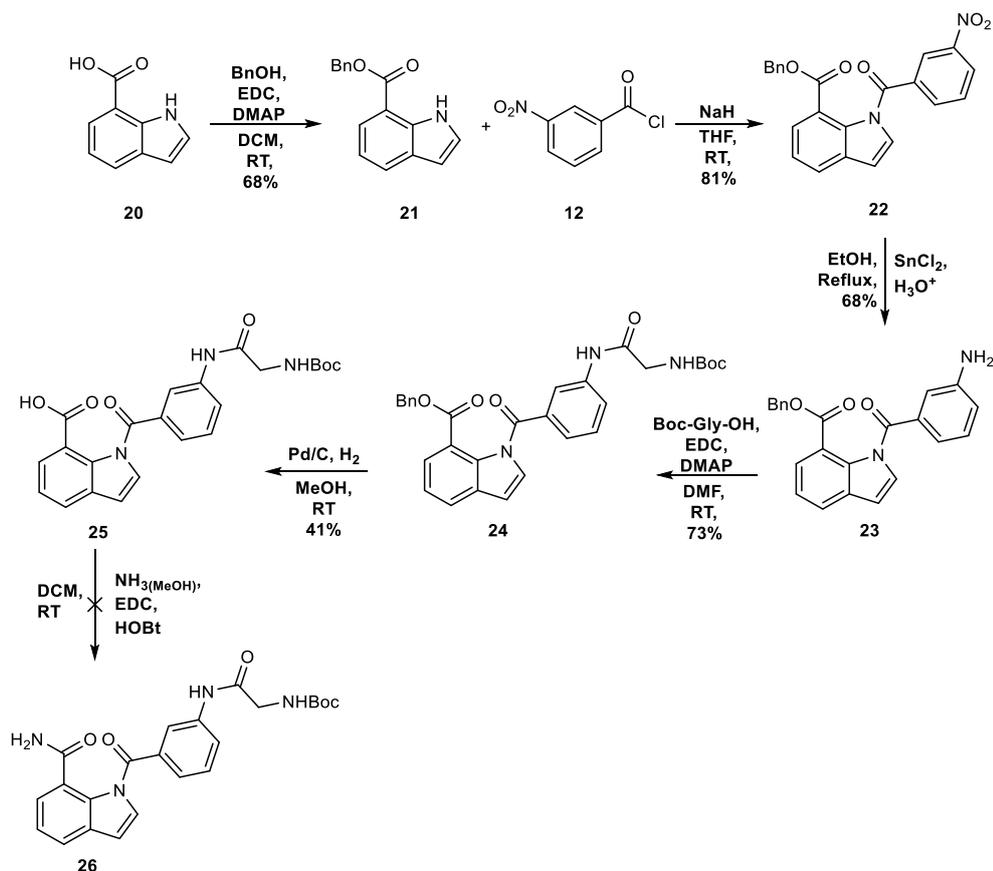




**Scheme 3:** Hydrolysis of Compound **16**.

### 2.3.1 Modified Synthetic Route

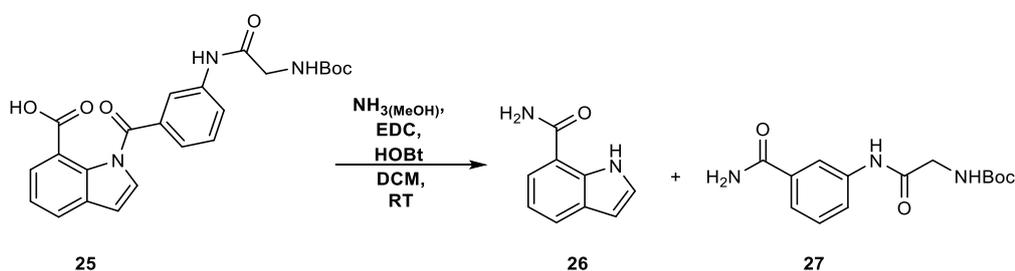
To avoid cleavage of the *N*-acyl bond it was decided to replace the methyl ester of **16** with a benzyl ester, which could be cleaved without the use of basic conditions. In addition, the protecting group of glycine was changed to tert-butyloxycarbonyl (Boc) (*N*-Boc-Gly-OH) to avoid basic deprotection conditions. The modified synthetic route is shown below (**Scheme 4**).



**Scheme 4:** Attempted synthesis of compound **10** using modified synthetic route

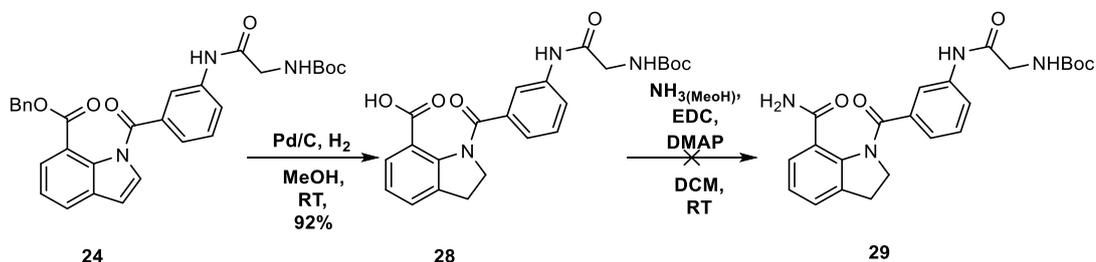
The benzyl ester functionality was introduced to indole **20** through the employment of an EDC mediated coupling reaction with benzyl alcohol

(BnOH), which proceeded with a good yield. *N*-Acylation of compound **21** with acyl chloride **12** gave the corresponding *N*-acyl indole **22** with an improved yield compared to that of the methyl ester **14**. Due to the presence of the benzyl ester the reduction of the nitro group of **22** was achieved using SnCl<sub>2</sub> under acidic conditions, rather than by hydrogenation. This was followed by the EDC mediated amide coupling of amine **23** with *N*-Boc-Gly-OH. Compound **24** was then subjected to hydrogenolysis to afford carboxylic acid **25** in moderate yield. The introduction of the desired amide functionality proved to be unsuccessful due to aminolysis of the indole *N*-acyl bond (**Scheme 5**).



**Scheme 5:** Aminolysis of compound **25**

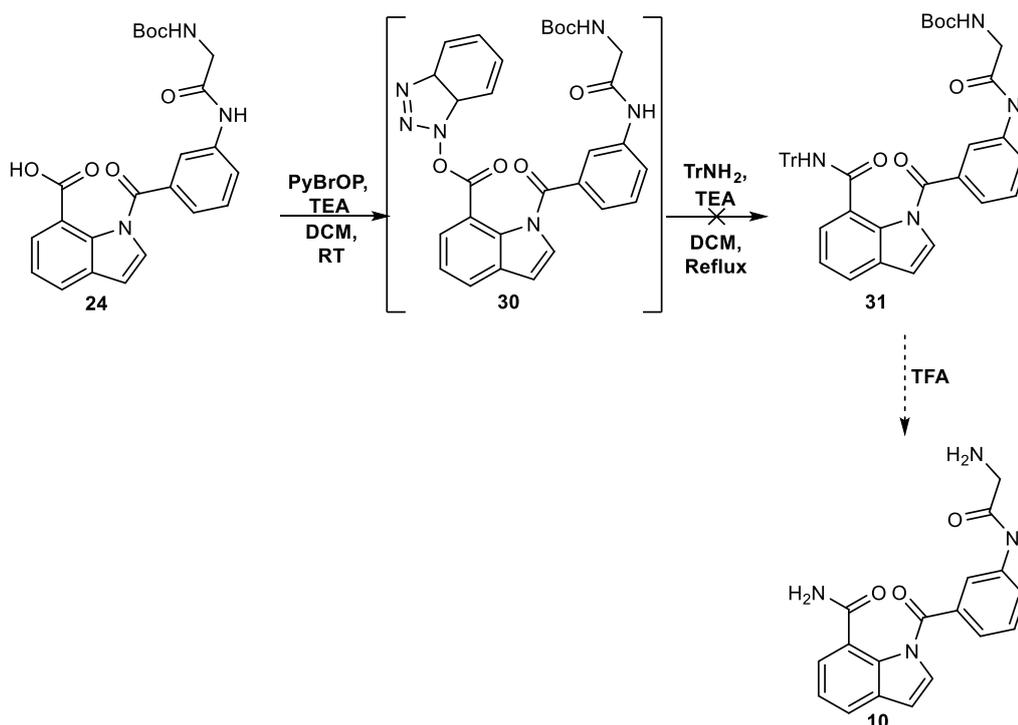
During the hydrogenolysis of compound **24** it was observed that if the reaction was carried out over a longer period the indole was reduced to the corresponding indoline **28**. It was proposed that the *N*-acyl indoline bond of **28** would be more stable towards aminolysis. Therefore, the desired primary amide functionality was targeted using indoline **28** (**Scheme 6**). Surprisingly, the indoline *N*-acyl bond was also labile towards aminolysis and formation of the desired indoline **29** was unsuccessful.



**Scheme 6:** Attempted synthesis of compound **29**.

Based on the lability of both the indole and indoline *N*-acyl bonds to ammonia, it was decided to form the triphenylmethyl protected amide **31** (**Scheme 7**). The protected amide would then be treated with trifluoroacetic

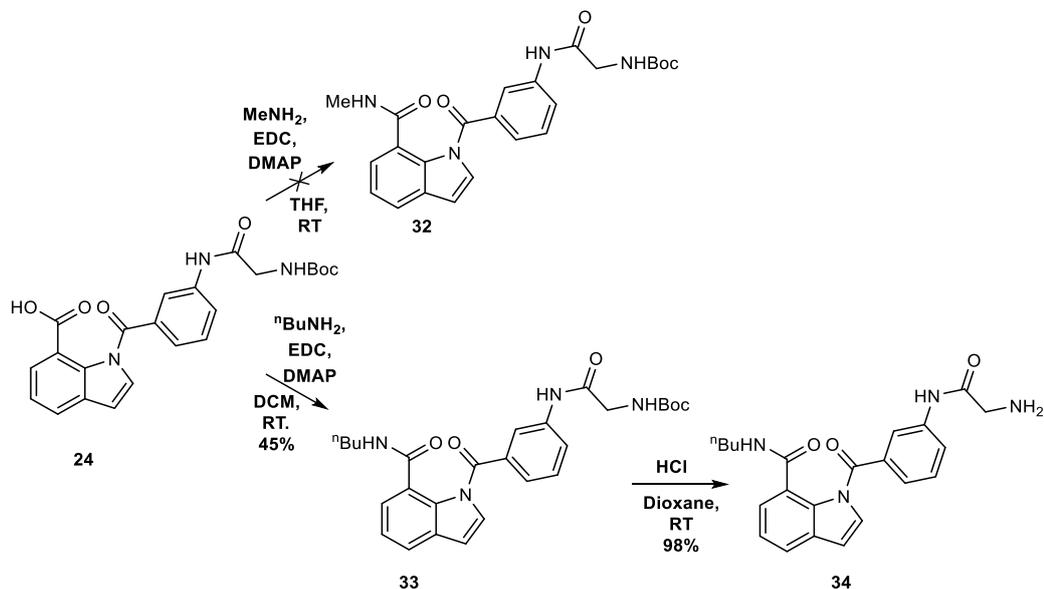
acid (TFA) to form the corresponding primary amide. The attempted synthesis of the protected amide **31** involved a Bromotripyrrolidinophosphonium hexafluorophosphate (PyBrOP) mediated amide coupling between carboxylic acid **24** and triphenylmethyl (Tr) amine. LCMS analysis revealed that the formation of the activated ester intermediate **30** was quantitative, however formation of the desired protected amide **31** was unsuccessful. This could be due to steric hindrance caused by the sterically bulky activated ester intermediate **30**.



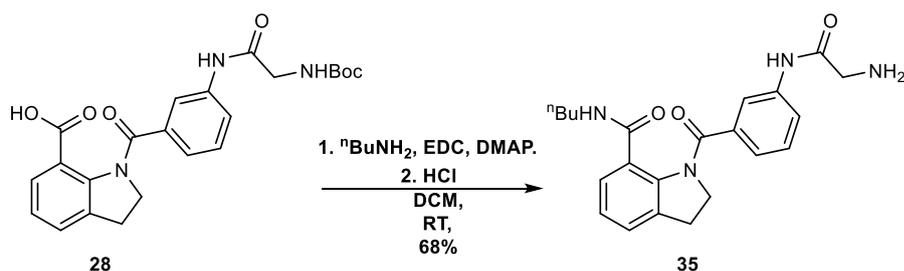
**Scheme 7:** Failed formation of primary amide **10** via formation of triphenylmethyl protected amide **31**.

Due to the synthetic difficulties encountered during the attempted formation of the primary amide it was decided that a simple secondary amide would be targeted using simple alkyl amines. The formation of methyl amide **32**, using an EDC mediated amide coupling reaction between methyl amine and acid **24**, was unsuccessful (**Scheme 8**). However, formation of the corresponding *n*-butyl amide **33** was successful. This amide formation proceeded with a yield of 45%. The low yield of this reaction may be due to the steric hindrance caused by the sterically bulky activated ester intermediate. Boc deprotection of compound **33** then followed to give compound **34** with a near quantitative

yield. The corresponding indoline **35** was also synthesised and comparable yields were obtained (**Scheme 9**).



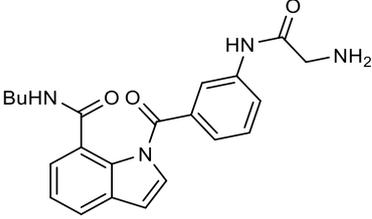
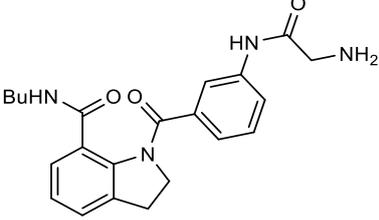
**Scheme 8:** Attempted synthesis of methyl amide **32** and the synthesis of butyl amide **34**.



**Scheme 9** Synthesis of indoline **35**.

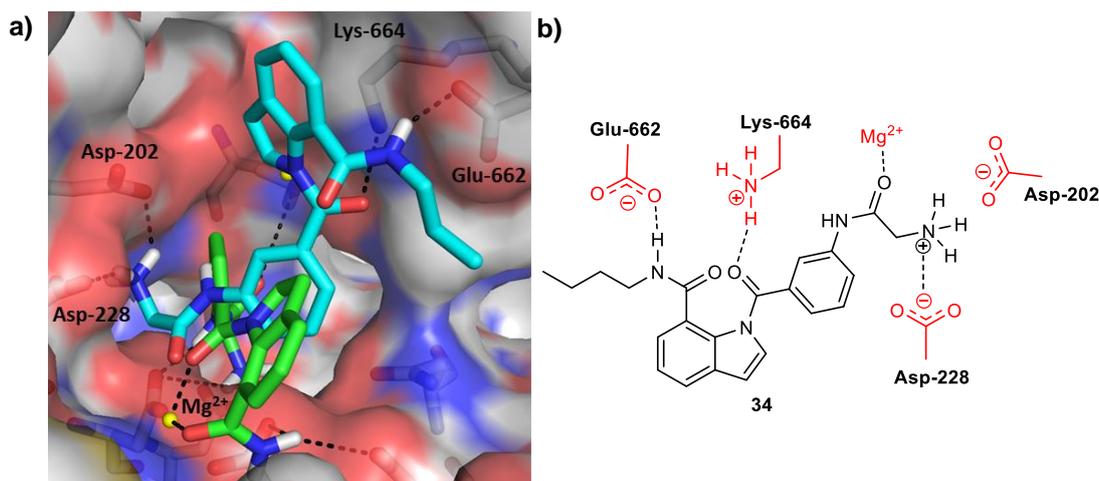
## 2.4 Biological Evaluation of Compounds **34** and **35**

The biological evaluation of compounds **34** and **35** (as well as all future compounds) was carried out by collaborators at the University of Helsinki, Finland. The compounds were screened against TmPPase using a fluorescence-based activity assay, which measured the formation of phosphate produced during the hydrolysis of pyrophosphate (see Experimental Section 6.1 for more details). The results are shown below (**Table 5**).

Compound No.	Structure	% Inhibition (500 $\mu$ M)
34		13 $\pm$ 3.8
35		NA

**Table 5:** Biological evaluation of compounds **34** and **35**. % Inhibition values are given as the mean standard deviation. NA = Not active.

Compound **34** inhibited the activity of TmPPase by 13% at a concentration of 500  $\mu$ M, whilst compound **35** was not active at the concentrations used. Due to the low activities of both compounds  $IC_{50}$  values were not obtained. The low activities observed for these compounds are likely to be caused by the presence of the <sup>n</sup>butyl amide, which could cause **34** and **35** to adopt a less favourable binding pose than was predicted for the primary amide **10**. Therefore, compound **34** was docked into the hydrolytic centre of TmPPase using Glide and the resultant docking pose was compared to the docking pose of the primary amide containing compound **10** (**Figure 23**). [63]

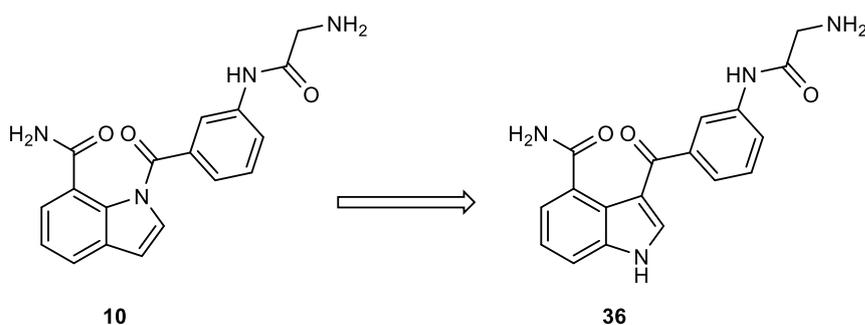


**Figure 23:** a) The predicted docking pose of compound **34** (blue) compared to the predicted docking pose of compound **10** (green). Dashed lines represent interactions between the two compounds and the hydrolytic centre of TmPPase (PDB:4AV3). [47, 63] b) 2-D representation of the predicted docking pose of compound **34**. Structures highlighted in red are amino acid residues of the hydrolytic centre. Dashed lines represent interactions between **34** and the amino acid residues of the hydrolytic centre. [64]

The predicted docking pose of compound **34** was found to be significantly different than that of compound **10**. Compound **34** is also predicted to make less contacts with the hydrolytic centre of TmPPase than **10**, including one less interaction with the  $Mg^{2+}$  ion. This further supports the hypothesis that the butyl amide is making a significant contribution to the poor biological activity observed for **34**.

## 2.5 Alternative Structural Isomer

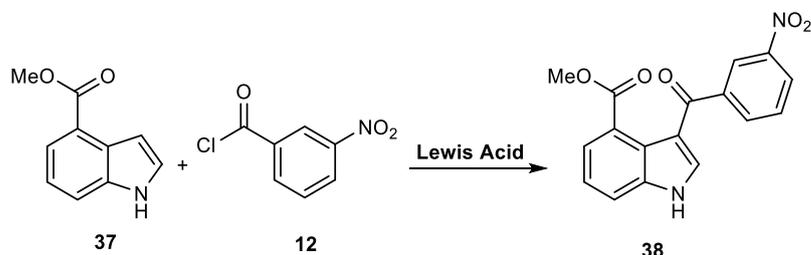
Due to the synthetic difficulties encountered during the synthesis of the *N*-acyl indole **10** it was decided to change the substitution of the indole moiety from 1,7- to 3,4- (Compound **36**, **Figure 24**). It was thought the change in substitution would lead to an increase in hydrolytic stability.



**Figure 24:** The 1,7- substitution pattern of indole **10** was changed to 3,4- of indole **36**.

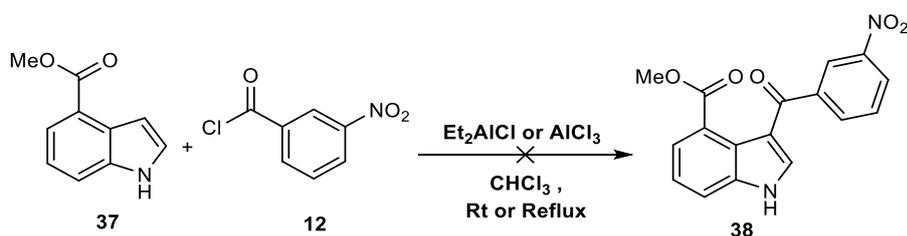
## 2.6 Attempted synthesis of Compound 36

To achieve the desired indole substitution pattern it was planned to use a Friedel-Crafts acylation reaction between indole **37** and acyl chloride **12** (**Scheme 10**).



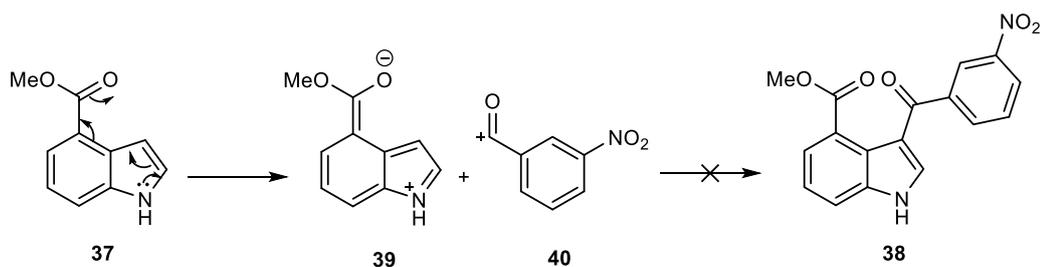
**Scheme 10:** The planned synthesis of indole **38** involved a Friedel-Crafts acylation between indole **37** and acyl chloride **12**.

Initially  $\text{Et}_2\text{AlCl}$  was used as the Lewis acid for the Friedel-Crafts acylation, however no product was observed when the reaction was carried out at both RT and under reflux conditions (**Scheme 11**). [65] The reaction was also attempted using  $\text{AlCl}_3$  as the Lewis acid, which also failed to yield **38**.

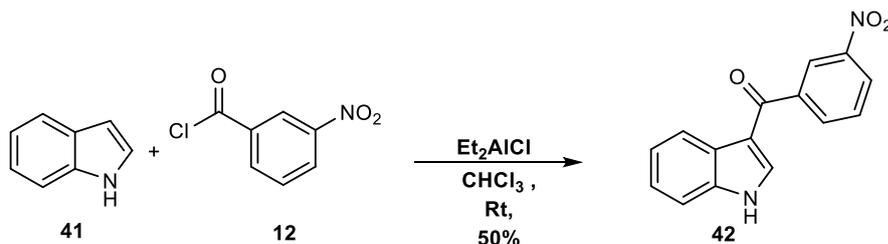


**Scheme 11:** Attempted Synthesis of Compound **38**

One possible reason for the lack of reactivity of indole **37** could be due to the conjugation of the indole *N* lone pair with the carbonyl of the ester at the 4-position (**Scheme 12**). This conjugation would mean that the *N* lone pair is unable to attack acylium ion **40** that is generated by the Lewis acid. To test this hypothesis a Friedel-Crafts acylation was carried out between indole **41** and acyl chloride **12**, using  $\text{Et}_2\text{AlCl}$  as the Lewis acid. This resulted in the formation of compound **42** and proceeded with a good yield (**Scheme 13**), further supporting the role of the carbonyl in reducing the reactivity of indole **37** towards Friedel-Crafts acylation reactions.

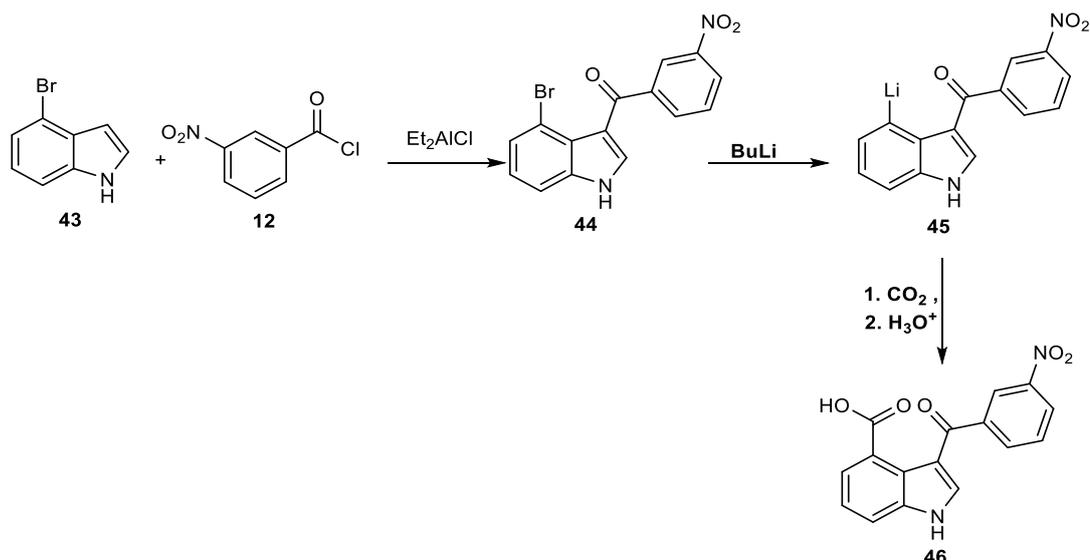


**Scheme 12:** Conjugation of the indole nitrogen lone pair of **37** could prevent it from reacting with acylium ion **40** to give indole **38**.



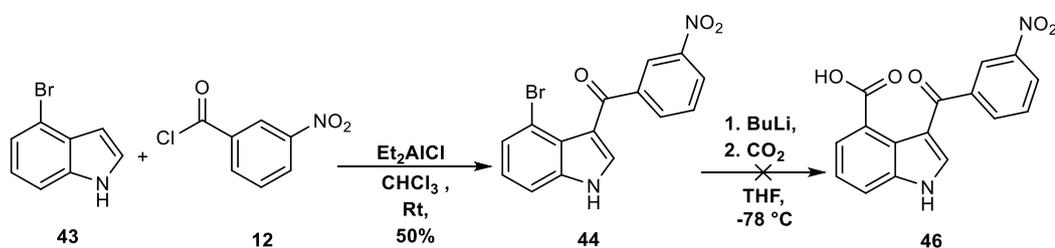
**Scheme 13:** The Friedel-Craft acylation between indole **41** and acyl chloride **12** was successful, supporting the argument that the presence of a carbonyl at the 4- position affects the reactivity of indole towards Friedel-Crafts acylation

An alternative method to achieve the desired substitution of indole was proposed in which the Friedel-Crafts acylation would be carried out using bromo indole **43** (**Scheme 14**). This would then be followed by a bromo-lithium exchange to give the lithiated indole intermediate **45**. Addition of CO<sub>2</sub> would then lead to the desired indole substitution pattern (Compound **46**, **Scheme 14**).



**Scheme 14:** Planned synthesis of indole **46** using Br-Li exchange

The initial Friedel-Crafts acylation reaction between indole **43** and acyl chloride **12** successfully resulted in the formation of compound **44**, however LCMS analysis revealed that the formation of indole **46** proved to be unsuccessful (**Scheme 15**) and a large number of undesired peaks were present. This may have been due to a number of side reactions occurring caused by the high reactivity of the lithiated intermediate **45**.

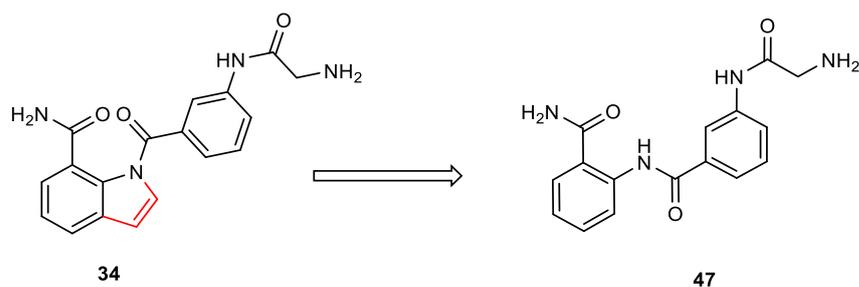


**Scheme 15:** Attempted synthesis of compound **46**

Due to the low biological activity of indole **34** and the synthetic difficulties encountered during the synthesis of both indole **34** and indoline **35**, it was decided to abandon the indole skeleton.

## 2.7 Simplification of Core

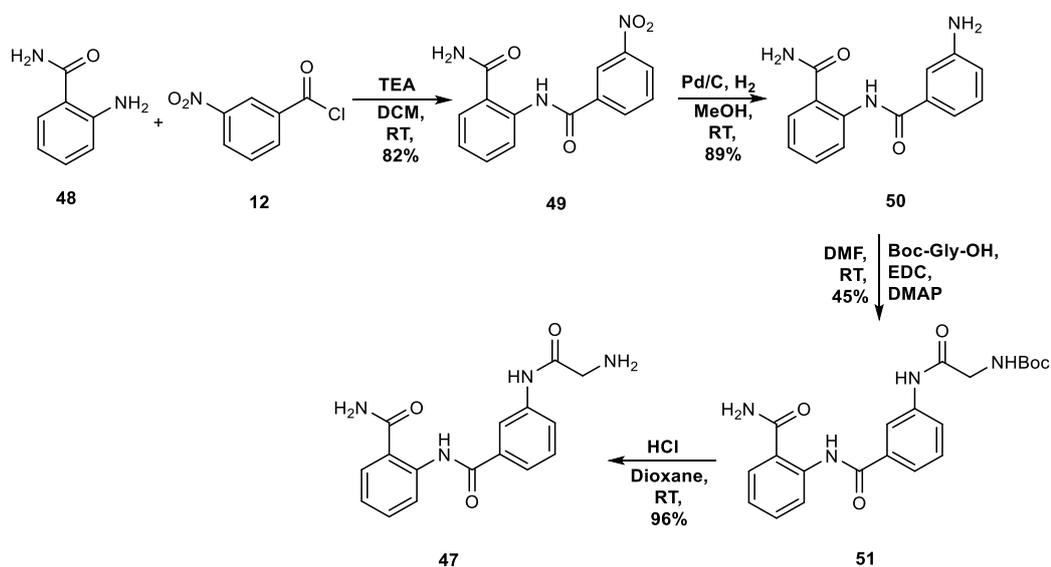
It was decided to simplify the core by removing the indole moiety to give a simple aromatic amide (compound **47**, **Figure 25**). This aromatic amide was predicted to be more hydrolytically stable than the corresponding indole *N*-acyl bond and would therefore allow the desired primary amide to be targeted.



**Figure 25:** Simplification of indole **34** to give compound **47**.

## 2.8 Synthesis of Compound 47

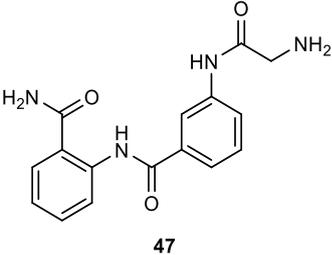
The synthesis of compound **47** is summarised below (**Scheme 16**). The first step involved a nucleophilic addition elimination reaction between 2-aminobenzamide **48** and acyl chloride **12**, which proceeded in good yield. [66] This was followed by the quantitative nitro reduction of compound **49** to give the corresponding aniline **50**. An EDC mediated amide coupling reaction then followed between aniline **50** and Boc-Gly-OH to give intermediate **51** in low yield. The low yield obtained from this reaction was due to the low solubility of aniline **50**. Finally, deprotection of the Boc protecting group by the treatment of compound **51** with hydrochloric acid (HCl) resulted in the formation of compound **47** in good yield.



**Scheme 16:** Synthesis of Compound **47**

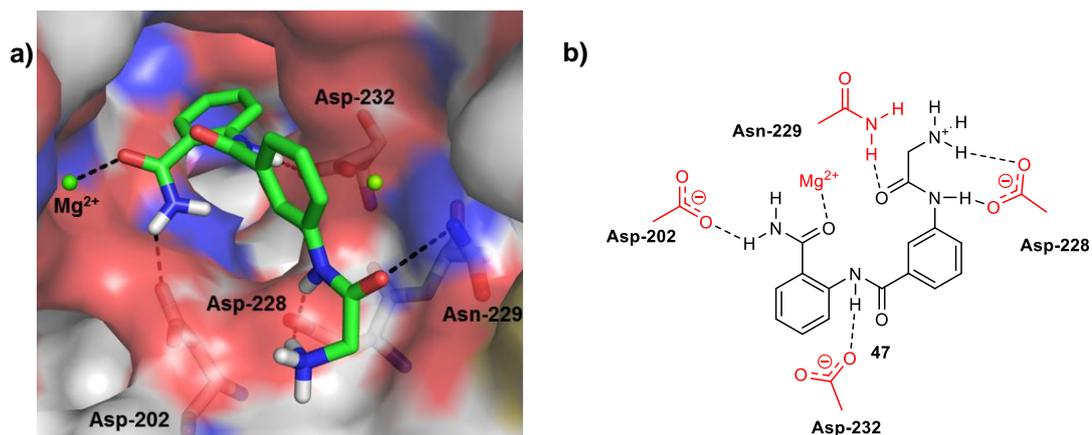
## 2.9 Biological Evaluation of Compound 47

Compound **47** was screened against TmPPase to determine the half maximal inhibitory concentration ( $IC_{50}$ ), the results are summarized below (**Table 6**).

Compound	$IC_{50}$
 47	< 5000 $\mu$ M

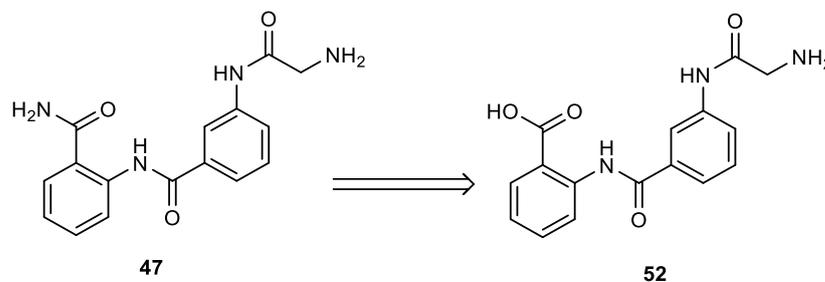
**Table 6:** Biological evaluation of compound **47**

Compound **47** was found to be inactive at the concentrations tested. Therefore, compound **47** was docked into the hydrolytic centre of TmPPase so that alterations could be made in order to improve the potency of the compound (**Figure 26**). [63] Due to the change in geometry arising from the removal of the indole moiety the predicted docking pose of compound **47** was different to that of compound **34**. The primary amide group of **47** was predicted to make a hydrogen bonding interaction with Asp-202 and also coordinate to one of the  $Mg^{2+}$  ions. In addition, the secondary amide of the glycine moiety was predicted to form a hydrogen bonding interaction with Asp-228. The other secondary amide between both phenyl rings was predicted to form one hydrogen bonding interaction with Asp-232. Finally, the primary amine was predicted to form an additional hydrogen bonding interaction with Asp-228.



**Figure 26:** a) Predicted docking pose of simplified SPROUT skeleton **47**, dashed lines represent the predicted interactions between **47** and the hydrolytic centre of TmPPase (PDB: 4AV3). [47, 63] b) The 2D representation of the docking pose. The residues of the hydrolytic centre are highlighted in red and dashed lines represent the predicted interactions between **47** and the hydrolytic centre. [64]

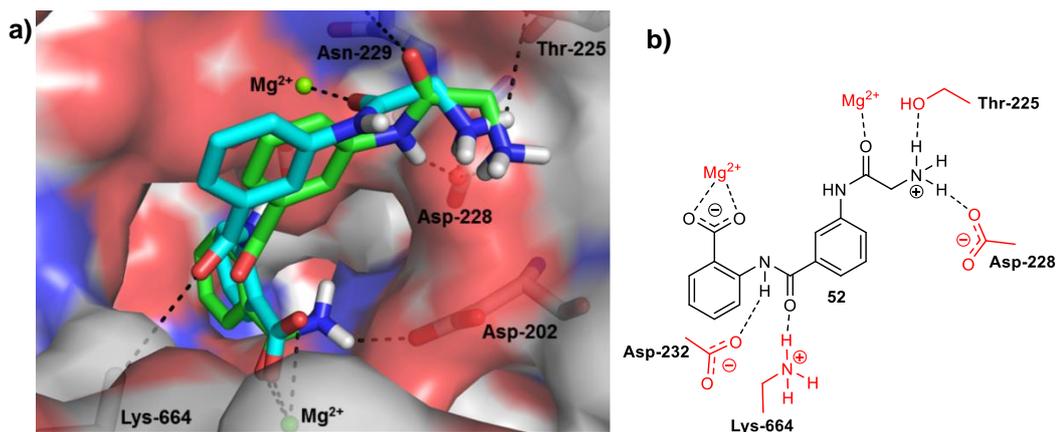
It was predicted that replacing the primary amide of compound **47** with a carboxylic acid (compound **52**, **Figure 27**) would lead to an improvement in activity towards TmPPase. The rationale behind this hypothesis was that the carboxylic acid has a formal negative charge, that is absent on the primary amide, and this would improve the metal coordination potential of the compound.



**Figure 27:** Replacement of the primary amide on compound **47** with a carboxylic acid to generate compound **52**

Compound **52** was docked into the hydrolytic centre of TmPPase and the resultant docking pose was compared to that of compound **47**. [63] Unsurprisingly, both docking poses were almost identical (**Figure 28**). There is a slight difference in orientation of the phenyl ring of **52** that contains the carboxylic acid. Here, it is orientated so that the carboxylic acid can coordinate to  $Mg^{2+}$  in a bidentate manner. The change in conformation of the phenyl ring containing the carboxylic acid also causes a slight change in

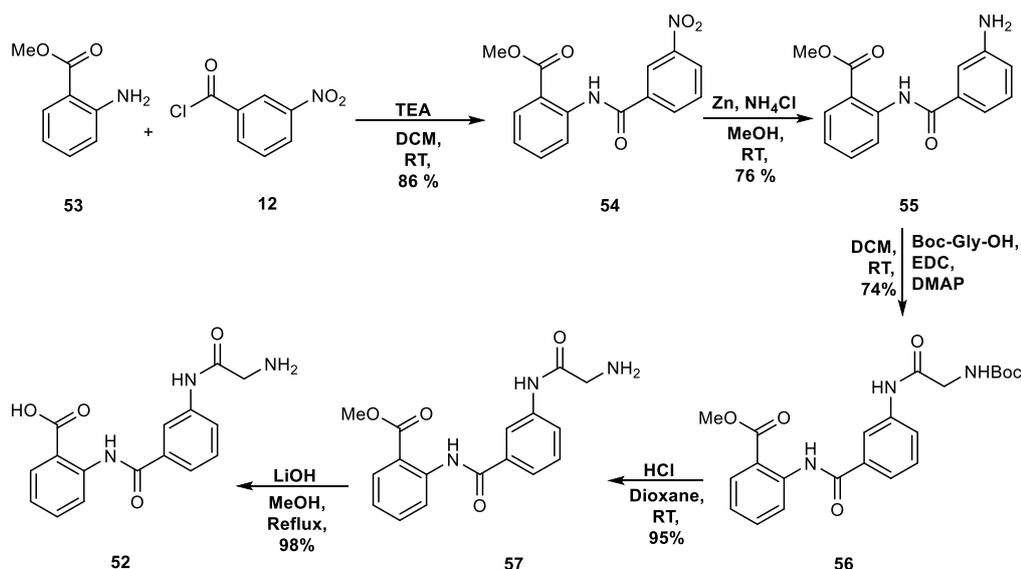
conformation of the other phenyl ring of **52**. The carbonyl of the glycine moiety is now predicted to coordinate to the other  $Mg^{2+}$  ion within the hydrolytic centre. In addition, the carbonyl of the amide between the two aromatic rings is now predicted to make a hydrogen bonding interaction with Lys-664. The final difference in the docking pose of compound **52** is that the primary amine is predicted to make a hydrogen bonding interaction with the hydroxyl group of Thr-225, in addition to Asp-228.



**Figure 28:** a) The docking pose of compound **52** (in blue) compared to that of compound **47** (green). The dashed lines represent interactions between the hydrolytic centre of TmPPase (PDB: 4AV3) and **47** and **52**. [47, 63] b) the 2-D representation of the docking of **52**. The structures in red are the residues of the hydrolytic centre, whilst the dashed lines represent interactions between **52** and the hydrolytic centre. [64]

## 2.10 Synthesis of Compound **52**

The synthesis of compound **52** is summarised below (**Scheme 17**). Firstly, acyl chloride **12** was coupled with aniline **53** to give compound **54** in good yield. [66] This was followed by the quantitative reduction of the nitro group of **54** using zinc and ammonium chloride to yield the corresponding aniline **55**. An EDC mediated coupling reaction between **55** and Boc-Gly-OH resulted in the formation of compound **56**, which proceeded with excellent yield. Deprotection of the Boc group, followed by the hydrolysis of the ester gave the desired compound **52** with near quantitative yield.

Scheme 17: Synthesis of compound **52**

## 2.11 Biological Evaluation of Compound **52**

Compound **52** was screened against TmPPase and the  $IC_{50}$  was compared to that of compound **47**, the results are summarized below (**Table 7**).

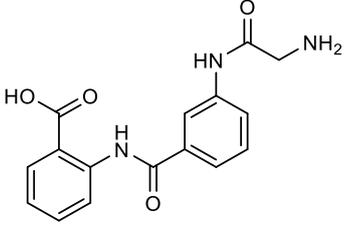
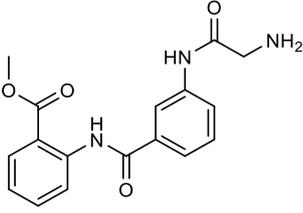
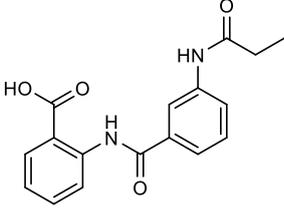
Compound	$IC_{50}$
<p><b>52</b></p>	$179 \pm 22 \mu\text{M}$
<p><b>47</b></p>	$>5000 \mu\text{M}$

Table 7: Biological evaluation of compound **52** and compound **47**

The  $IC_{50}$  of compound **52** was improved when compared with compound **47**, if the docking studies are accurate this could be due to the stronger coordination of the carboxylic acid of **52** with  $Mg^{2+}$ .

## 2.12 SAR Exploration of Compound 52

A library of compounds based on compound **52** was synthesised, using the same route as **Scheme 17**, and screened against TmPPase in order to determine structure activity relationships (SAR) of compound **52** (**Table 8**).

Compound	IC <sub>50</sub>
 <p><b>52</b></p>	179 ± 22 μM
 <p><b>57</b></p>	861 ± 107 μM
 <p><b>58</b></p>	2281 ± 912 μM
 <p><b>59</b></p>	3665 ± 974 μM

**Table 8:** Biological results for compounds **52-59**, which were screened against TmPPase

The methyl ester **57** showed lower activity against TmPPase than the corresponding acid, suggesting that the negative charge of the carboxylate of compound **52** is important for activity. This is unsurprising considering that the carboxylate of **52** is predicated to coordinate with a positively charged Mg<sup>2+</sup>. Compound **58**, which does not contain the glycine primary amine, has a much higher IC<sub>50</sub> than that of compound **52**. This suggests that the primary

amine is essential for potency and the loss of activity observed could be caused by the loss of the hydrogen bonding interactions that the amine is predicted to make with Thr-225 and Asp-228 mentioned above. The amide of the glycine moiety also seems to be important for activity, which can be seen from the drop in activity observed between compound **58** and compound **59**. Loss of the amide carbonyl could result in the loss of the hydrogen bonding interaction that it was predicted to make with Asn-229.

### **2.13 Analysis of Compound 52 for Lead Likeness using LLAMA**

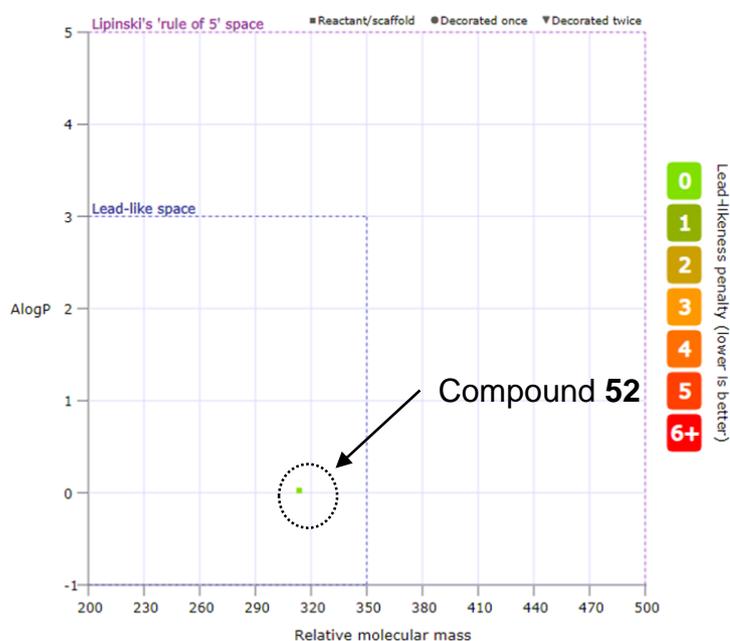
An important factor to consider during the early stages of drug development is the lead-likeness of a compound. Lead-likeness is a qualitative guide for identifying compounds that will likely deliver lead compounds further down the drug discovery pipeline. [67]

The Lead-Likeness and Molecular Analysis (LLAMA) programme determines the lead-likeness of a compound by analysing some of its key predicted properties. [68] The first of these properties is the predicted partition coefficient (ALogP), which is a measure of the predicted lipophilicity of a compound. The lipophilicity of a compound is important because it determines properties such as solubility, drug availability and off target effects. [69] To be considered lead-like by LLAMA the ALogP must be between -1 and 3. In addition, the relative molecular mass (RMM) of the compound must be between 200 and 350 da. If a compound has an ALogP and RMM that is within these ranges then it is said to occupy lead-like space. There are also a number of other properties that LLAMA uses to determine if a compound is lead-like. The heavy atom count must be between 17 and 27 and the compound must have <3 aromatic rings. Furthermore, the compound must not possess any 'bad' functional groups (e.g NO<sub>2</sub>). These bad functional groups are known to be toxic, the nitro group for example is metabolised to give toxic metabolites.

When the LLAMA programme analyses a compound it assigns a "lead-likeness" penalty for properties that are not lead-like. [68] The output score of each compound therefore determines how much the compound deviates from lead-like space. A score of 0 means it does not deviate from lead-like

space, which means that all the properties fall within the accepted ranges discussed above. Whilst, higher scores mean that the compound moves further away from lead like space.

Compound **52** was analysed using LLAMA and the results are shown as both a lead-likeness plot (**Figure 29**) and in tabulated form below (**Table 9**). Overall **52** had a lead-likeness penalty of 0 and a plot of the AlogP against RMM indicates that it occupies lead-like space. The AlogP of **52** is predicted to be 0.03, meaning that is well within the acceptable range for AlogP. In addition, there are 23 heavy atoms (with two aromatic rings) and no unfavourable functional groups. This indicates that compound **52** occupies lead-like space and is worth developing further.



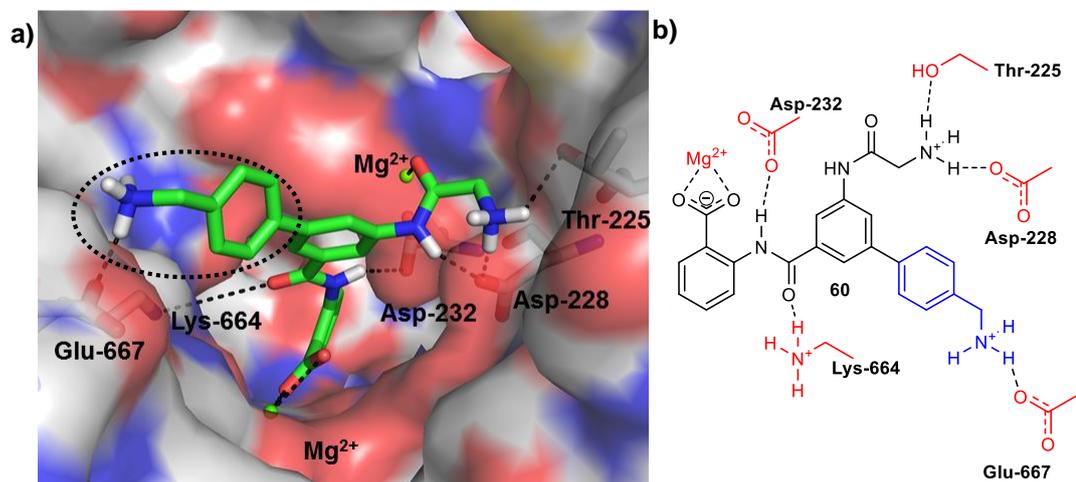
**Figure 29:** Lead-likeness plot of predicted AlogP vs relative molecular mass of compound **52**. The plot indicates that compound **52** occupies lead-like space with no lead-likeness penalties. [68]

Property	Value	Lead-Likeness Penalty
AlogP	0.03	0
No. of Heavy Atoms	23	0
No. of Aromatic Groups	2	0
RMM	313.08	0
No. of 'Bad' functional groups	0	0

**Table 9:** LLAMA predicted properties of compound **52** and their associated lead-likeness penalty. Compound **52** has an overall lead-likeness penalty of 0. [68]

## 2.14 SPROUT Expansion of Compound 52

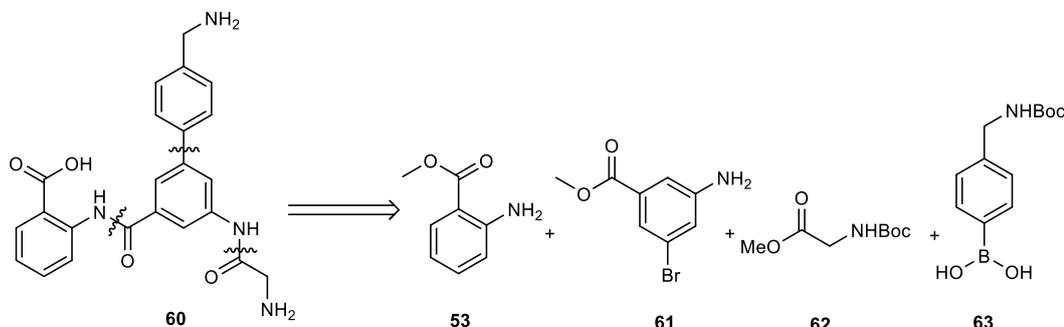
SPROUT was used to expand compound **52**, with the aim of making a more potent inhibitor. [61] This resulted in the generation of compound **60**, the predicted docking pose for compound **60** is shown below (**Figure 30**). The 'pre-expanded' portion of compound **60** has an identical docking pose to compound **52**, whilst the primary amine of the benzylamine moiety is predicted to make an additional hydrogen bonding interaction with Glu-667.



**Figure 30:** a) Predicted docking pose of the SPROUT expanded compound **60** visualised in PyMol highlighting the benzylamine functionality added using SPROUT. Dashed lines represent interactions between **60** and the hydrolytic centre of TmPPase. b) The 2-D representation, the new benzylamine moiety is highlighted in blue. The structures in red are the residues of the hydrolytic centre, whilst the dashed lines represent interactions between **60** and the hydrolytic centre.[64]

## 2.15 Retrosynthetic Analysis of Compound 60

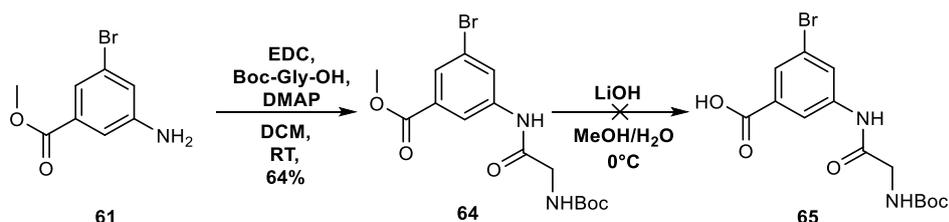
Structure **60** was subjected to a retrosynthetic analysis. It was thought that **60** could be accessed from the commercially available starting materials shown below (**Scheme 18**).



**Scheme 18:** Retrosynthetic analysis of compound **60**, which identified **53**, **61**, **62** and **63** as starting materials

## 2.16 Synthesis of Compound 60

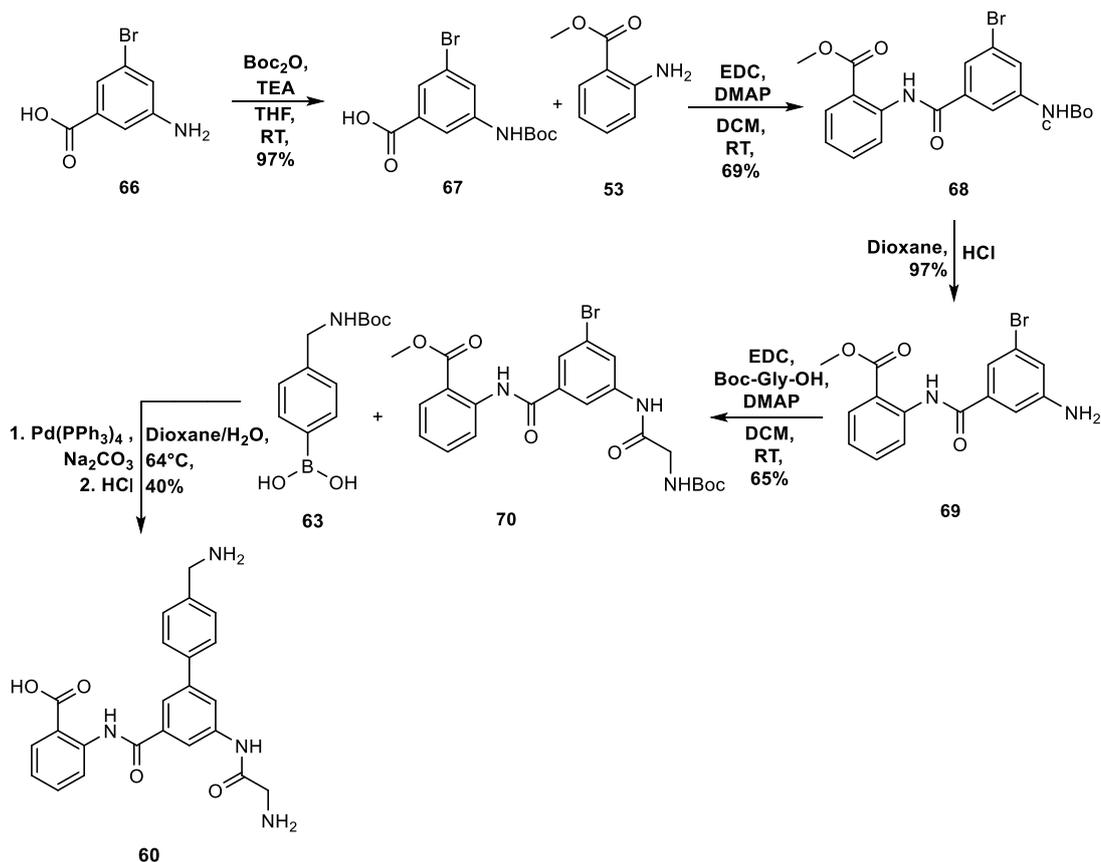
The initial attempted synthesis of compound **65** is shown below (**Scheme 19**). An amide coupling reaction between aniline **53** and Boc-Gly-OH resulted in the generation of compound **64** in good yield. Hydrolysis of the methyl ester was unsuccessful due to the hydrolysis of the amide bond of the glycine moiety.



**Scheme 19:** Attempted synthesis of compound **65**.

An alternative synthetic route was then devised in order to synthesise compound **60** (**Scheme 20**). The first step in the modified synthetic route involved the protection of aniline **66** using di-tert-butyl dicarbonate ( $\text{Boc}_2\text{O}$ ). Intermediate **67** was then coupled with aniline **53** using an EDC mediated amide coupling reaction and resulted in the formation of compound **68**, which proceeded with a moderate yield. Compound **68** was then treated with HCl to form aniline **69**, which then underwent an EDC mediated amide coupling

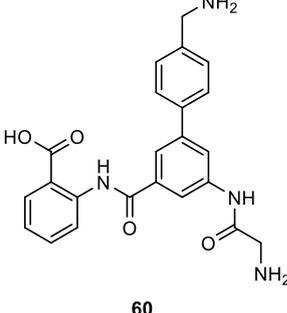
reaction with Boc-Gly-OH to give intermediate **70** in good yield. Boronic acid **63** and intermediate **70** were then coupled using standard Suzuki coupling conditions. [70] Due to the basicity of the Suzuki conditions the methyl ester of compound **70** hydrolysed during the Suzuki reaction to give the corresponding acid, which was treated with HCl to give the desired compound **60**.



Scheme 20: Synthesis of compound **60**

## 2.16 Biological Evaluation of Compound **60**

Compound **60** was screened against TmPPase to determine the  $IC_{50}$ , the results are summarised below (Table 10).

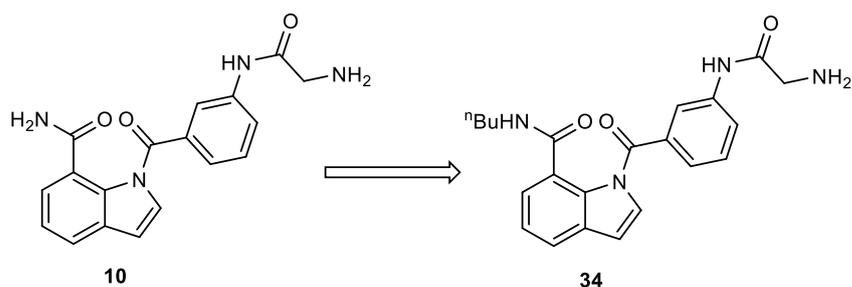
Compound	IC <sub>50</sub>
 <p style="text-align: center;">60</p>	550 ± 94 μM

**Table 10:** Biological evaluation of compound 60

The SPROUT expanded compound **60** had an IC<sub>50</sub> value of 550 μM. This was considerably higher than that of compound **52**, which had an IC<sub>50</sub> value of 179 μM. The observed reduction in potency could be caused by the increased size of **60**. As mentioned earlier (**Section 1.5.1**), when the substrate binds to M-PPases the hydrolytic centre closes and there is a reduction in its volume. This reduced volume could mean that **60** adopts a sub-optimal binding pose that was not predicted by SPROUT. These results therefore suggest that smaller modifications of compound **52** are needed to increase potency.

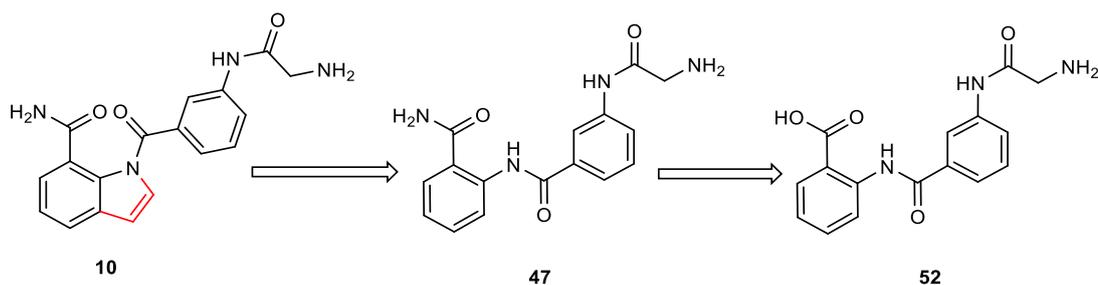
## 2.18 Conclusions

The *de novo* design programme SPROUT was used to develop an indole-based putative inhibitor of TmPPase (compound **10**, Figure **31**), which was designed to bind to the hydrolytic centre. A number of synthetic difficulties were encountered during the synthesis of **10** that were caused by the lability of the indole *N*-acyl bond. Therefore, the *n*-butyl amide **34** was targeted instead. Compound **34**, as well as the corresponding indoline **35**, were shown to have low activity against TmPPase.



**Figure 31:** The SPROUT designed putative inhibitor **10** could not be synthesised, so the corresponding butyl amide **34** was targeted instead. It was found that **34** had a very low potency against TmPPase.

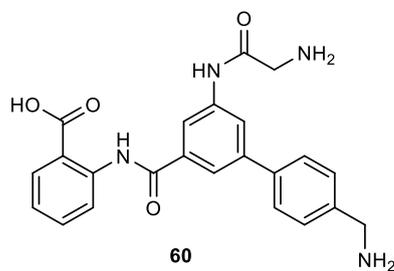
Indole **10** was then simplified to compound **47**, this was due to both the synthetic difficulties encountered during the synthesis of indole **10** and its low activity (**Figure 32**). Compound **47** was found to be inactive against TmPPase at the concentrations tested. In order to make a more potent inhibitor the primary amide of **47** was replaced with a carboxylic acid (compound **52**, **Figure 32**). It was thought that the carboxylate would be a better metal-coordinating group to  $Mg^{2+}$  than the amide. Compound **52** had an  $IC_{50}$  of 179  $\mu M$  against TmPPase and is the first example of an inhibitor of an M-PPase using SBDD. A small library of compounds based on compound **52** was also synthesised to determine the SAR of **52**. All compounds tested were less active than **52**, suggesting that the key functionalities of **52** are important for activity.



**Figure 32:** Simplification of **10** resulted in the generation of compound **47**, which was inactive against TmPPase. Replacing the amide of **47** with a carboxylic acid lead to the identification of **52**, which had an  $IC_{50}$  of 179  $\mu M$ . This is first example of an inhibitor of a M-PPase designed using SBDD.

Compound **52** was then expanded using SPROUT to generate of compound **60** (**Figure 33**), which was synthesised in six steps. The  $IC_{50}$  value of compound **60** was 550  $\mu M$ , which was higher than compound **52** and

therefore less potent. One possible explanation of this lower activity is due to the large size of compound **60**. This large size could mean that compound **60** occupies a sub-optimal binding pose within the closed hydrolytic centre of TmPPase.

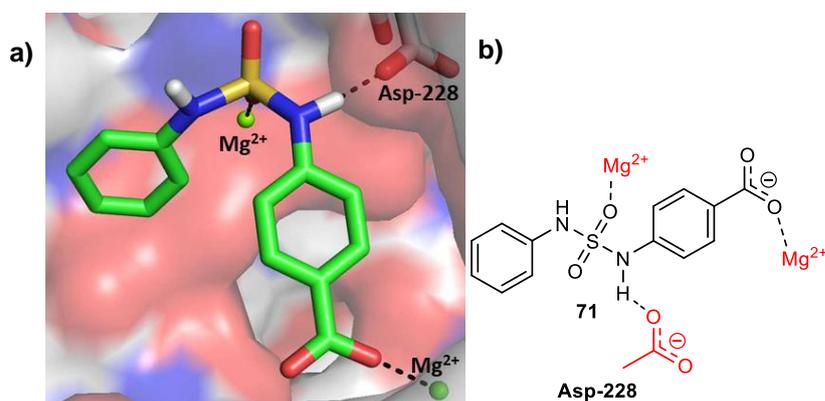


**Figure 33:** SPROUT expansion of **52** lead to the identification of compound **60**, which had an IC<sub>50</sub> value of 550 μM and was less potent than **52**.

## Chapter 3: Substrate-Inspired Design of a Novel Phosphate Mimetic

### 3.1 *De novo* Design of a Novel Sulfamide Fragment.

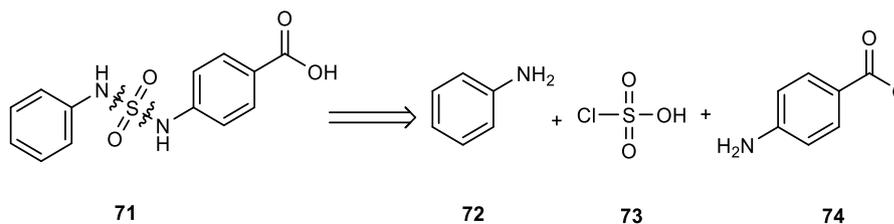
The *de novo* design programme SPROUT was used to develop a novel sulfamide fragment that was designed to mimic phosphate in the hydrolytic centre of TmPPase (compound **71**, **Figure 34**). [47, 61] The design process initially involved docking a sulfonamide group onto one of the  $Mg^{2+}$  ions of the hydrolytic centre, which mimics the coordination of a phosphate ion. In addition, a carboxylic acid group was docked onto the other  $Mg^{2+}$  ion. Linking the sulfonamide and carboxylic acid groups with phenyl rings resulted in the generation of compound **71**, which is a novel sulfamide fragment.



**Figure 34:** a) Predicted docking pose of SPROUT designed sulfamide fragment **71**, visualised using PyMol. Dashed lines represent interactions between sulfamide **71** and the hydrolytic centre of TmPPase. [47, 61] b) The 2D representation. The  $Mg^{2+}$  ions and amino acid residues of the hydrolytic centre are highlighted in red. [64]

### 3.2 Retrosynthetic Analysis of SPROUT Designed Sulfamide Fragment

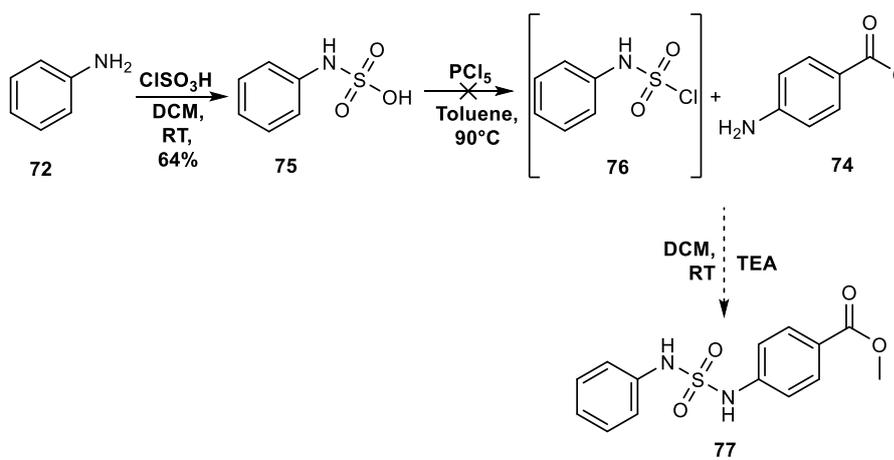
Retrosynthetic analysis of **71** revealed that it could be synthesised using the commercially available; aniline **72**, chlorosulfonic acid **73** and methyl 4-aminobenzoate **74** (**Scheme 21**).



**Scheme 21:** Retrosynthetic analysis of sulfamide **71**.

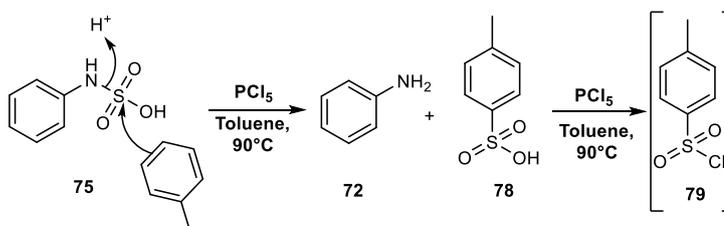
### 3.3 Synthesis of Sulfamide Fragment **71**

The initial synthesis of sulfamide **71** is summarised below (**Scheme 22**).



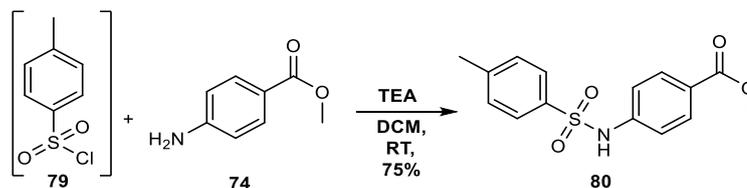
**Scheme 22:** Attempted synthesis of sulfamide **71**.

Formation of sulfamic acid **75** was achieved using the reported method of treating aniline **72** with chlorosulfonic acid, which proceeded with a moderate yield of 64 %. [71] The next stage in the planned synthetic route was the chlorination of **75** using  $\text{PCl}_5$ . [71] However, LC-MS and nuclear magnetic resonance spectroscopy (NMR) analysis showed that formation of sulfamoyl chloride **76** was unsuccessful. This was due to the substitution of toluene (the reaction solvent) with **75** to give the undesired sulfonic acid **78** (**Scheme 23**). Chlorination of **78** then resulted in the formation of sulfonyl chloride **79**.



**Scheme 23:** Substitution between toluene and sulfamic acid **75** resulted in the formation of sulfonic acid **78**.

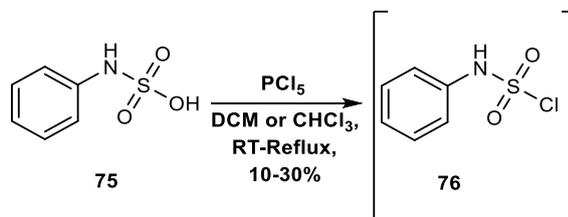
This observation was confirmed by coupling the resultant sulfonyl chloride **79** with **74**, which resulted in the formation of sulfonamide **80** (**Scheme 24**).



**Scheme 24:** Formation of undesired sulfonamide **80**.

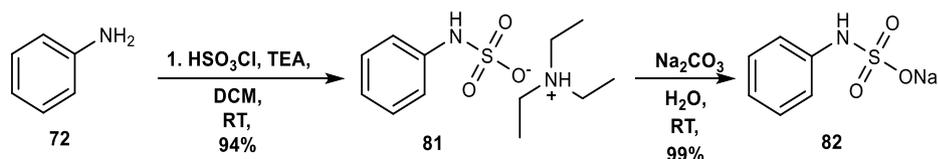
### 3.3.1 Optimisation of Chlorination Step and Synthesis of Sulfamide Fragment **71**

To avoid the substitution of sulfamic acid with toluene, alternative solvents were explored for the chlorination reaction. Both chloroform and dichloromethane (DCM) were initially tested and both led to the successful formation of sulfamoyl chloride **76** (**Scheme 25**) and no side-products were observed. The yield for the chlorination reaction was 10% when DCM was used and 15% when chloroform was used. Heating the reaction mixture led to an increase in yield for both chloroform (30%) and DCM (15%), the remainder of the material in both reactions was **75**. Due to the higher yield obtained, chloroform was chosen as the solvent for the chlorination reaction.



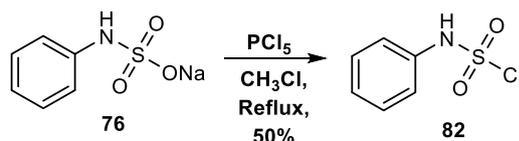
**Scheme 25:** Chlorination of **75**. No side products were observed when DCM or chloroform were used as the solvent. Refluxing the reaction mixture in chloroform gave the highest yield of 30%.

It was thought that the yield of the chlorination reaction could be further improved by using the sodium salt of **75**. The desired sodium sulfamate **82** was obtained through the treatment of aniline **72** with chlorosulfonic acid in the presence of triethylamine (TEA) (**Scheme 26**). [72] This led to the generation of the TEA salt **81**, which was then treated with  $\text{Na}_2\text{CO}_3$  to give the desired sodium sulfamate **82**. [73]



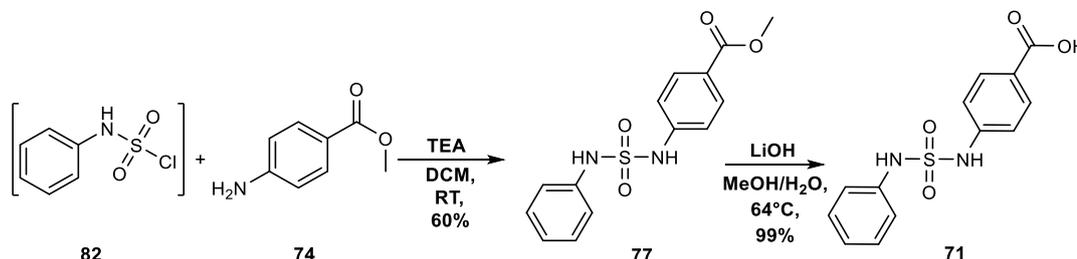
**Scheme 26:** Formation of sodium sulfamate **82**.

Chlorination of sodium sulfamate **82** resulted in the formation of sulfamoyl chloride **76** with an improved yield of ~50 % (Scheme 27).



**Scheme 27:** Formation of sulfamoyl chloride **76** using sodium sulfamate **82**

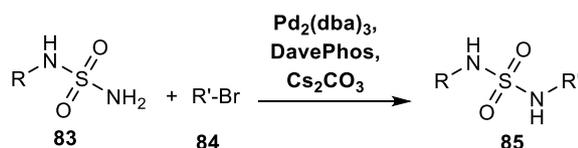
The desired sulfamide fragment was then obtained by coupling sulfonyl chloride **82** with aniline **74** (Scheme 28), which proceeded with a moderate yield of 61%. The overall yield obtained for the formation of **77** starting from aniline was 29%. Finally, ester hydrolysis of **77** gave the desired carboxylic acid **71** with near quantitative yield.



**Scheme 28:** Synthesis of SPROUT designed sulfamide fragment **71** was achieved through the coupling of sulfamoyl chloride **82** and aniline **74**, followed by ester hydrolysis.

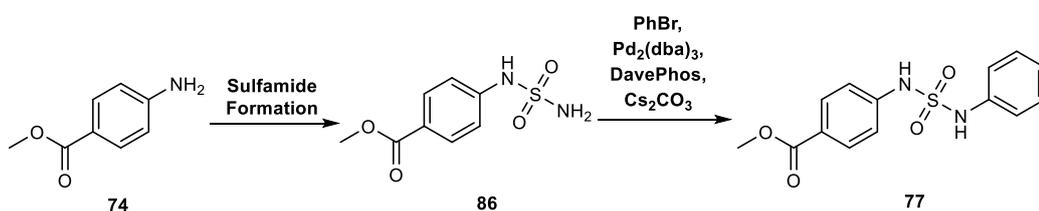
### 3.3.2 Alternative Synthesis of **71**

Due to the low overall yield of **77**, alternative routes for the synthesis of sulfamides were explored. One alternative route involves using a Buchwald-Hartwig coupling reaction to couple a primary sulfamide with an aryl bromide, which produces sulfamides with good yields (**Scheme 29**). [74]



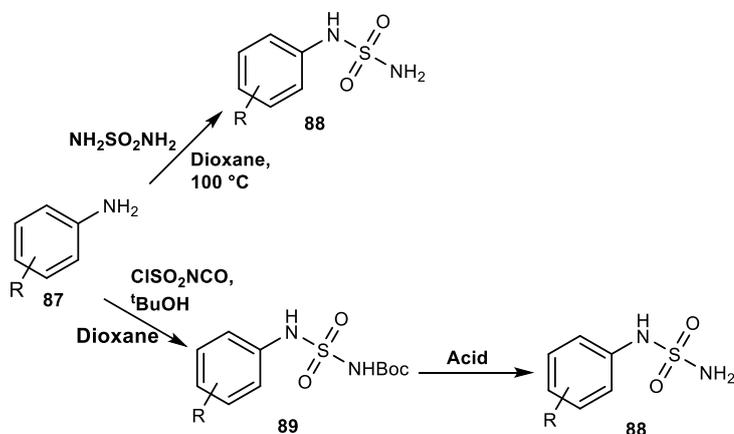
**Scheme 29:** Buchwald coupling of a primary sulfamide with an aryl/vinyl bromide. R = alkyl, aryl or vinyl R'=aryl or vinyl

The planned Buchwald-Hartwig coupling reaction to produce **77** is outlined below (**Scheme 30**). Initially, aniline **74** would be converted to the primary sulfamide **86**. This would then be followed by the Buchwald-coupling reaction between **86** and bromobenzene to give sulfamide **77**. [74]



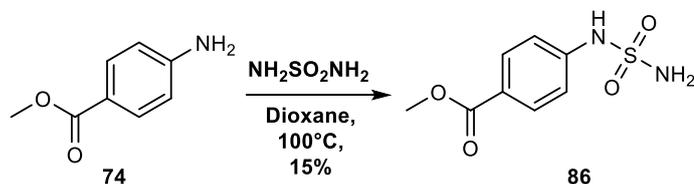
**Scheme 30:** Modified synthetic route of sulfamide **77** involving the formation of primary sulfamide **86**, which would then be coupled with PhBr using the Buchwald reaction

There are a number of different reported methods of forming primary sulfamides. One method involves the addition of a large excess sulfamide ( $\text{H}_2\text{NSO}_2\text{NH}_2$ ) to an aniline **87** and is carried out at high temperatures (**Scheme 31**). [75] An alternative method involves the formation of the Boc-protected sulfamide **89**, which is then deprotected to give the desired primary sulfamide **88** (**Scheme 31**). [76] The Boc-protected sulfamide is formed by the reaction of an aniline with Boc-sulfonyl chloride, which in turn is formed by the addition of chlorosulfonyl isocyanate ( $\text{ClSO}_2\text{NCO}$ ) to  $^t\text{BuOH}$ .



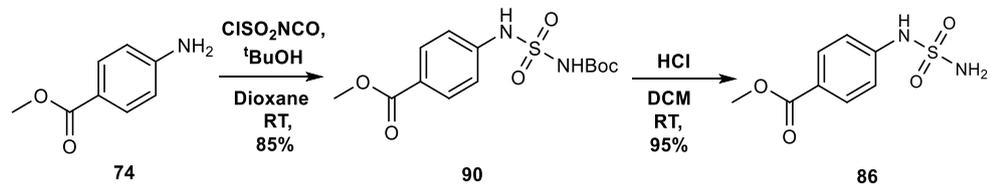
**Scheme 31:** Possible route towards primary sulfamides. [75, 76]

Initial attempts toward the synthesis of primary sulfamide **86** involved the addition of sulfamide to **74** (**Scheme 32**), which lead to the formation of **86** with a poor yield of 15%.



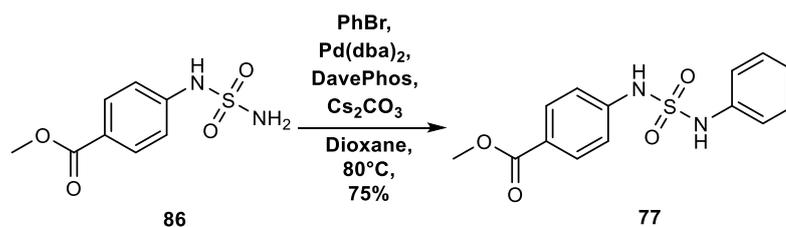
**Scheme 32:** Synthesis of **86** was achieved by heating aniline **74** with a large excess of sulfamide.

Due to the low yield obtained for this reaction it was decided to attempt to target the primary sulfamide using the second method discussed above (**Scheme 33**). Firstly, Boc-sulfamoyl chloride was formed in situ by the addition of chlorosulfonyl chloride to <sup>t</sup>BuOH. This was followed by the addition of **74**, which lead to the formation of intermediate **90** in good yield. Deprotection of **90** under acidic conditions then followed, which lead to the formation of **86** in near quantitative yield. The overall yield for this method was 81% and no purification steps were needed. Therefore, this method was chosen for the formation of primary sulfamides.



**Scheme 33:** The second synthesis of primary sulfamide **86**. Initially, the Boc protected sulfamide **90** was formed. This was followed by Boc deprotection to the corresponding primary sulfamide with an excellent yield.

Primary sulfamide **86** and bromobenzene then underwent a Buchwald-Hartwig coupling reaction, which lead to the formation of sulfamide **77** with a high yield of 75% (**Scheme 34**). [74] The overall yield for the formation of sulfamide **77**, starting from aniline **74**, was 61%. This was much higher than the yield obtained for the previous sulfamide synthesis discussed. Therefore, the Buchwald-Hartwig coupling reaction was used for the synthesis of all future sulfamides.



**Scheme 34:** Buchwald-coupling between **86** and bromobenzene lead to the formation of **77** in good yield.

### 3.4 Biological Evaluation of Compound 71

Compound **71** was screened against TmPPase to determine the % inhibition at a concentration of 100  $\mu\text{M}$  and the  $\text{IC}_{50}$  value (**Table 11**).

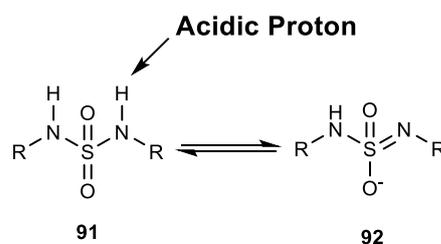
Compound	% Inhibition at 100 $\mu\text{M}$	$\text{IC}_{50}$
<p style="text-align: center;"><b>71</b></p>	$26 \pm 7 \%$	$1353 \pm 173 \mu\text{M}$

**Table 11:** Biological Evaluation of Compound **71**. Results are shown as % inhibition and an  $\text{IC}_{50}$  and Ligand efficiencies.

The SPROUT designed sulfamide fragment inhibited the activity of TmPPase by 26% at a concentration of 100  $\mu\text{M}$  and had an  $\text{IC}_{50}$  value of 1353  $\mu\text{M}$ .

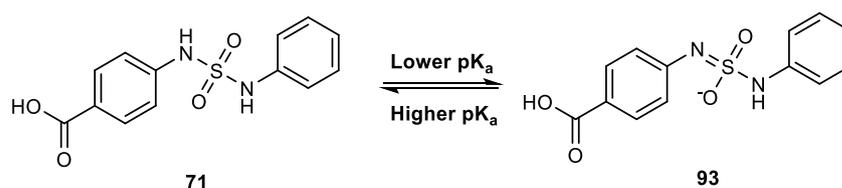
### 3.5 $\text{pK}_a$ of Sulfamides

The sulfamide functionality is acidic and will deprotonate under basic conditions to generate the corresponding conjugate base **92** (**Figure 35**). [77]



**Figure 35:** The sulfamide functionality **91** is acidic and can undergo deprotonation to give the corresponding conjugate base **92**. [77]

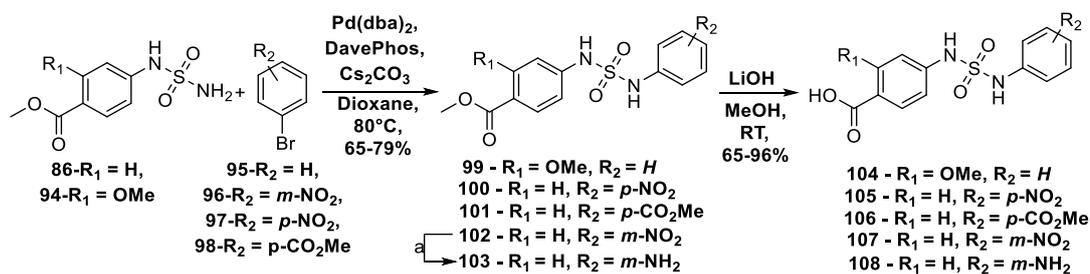
As discussed above, the sulfamide of **71** was predicted to coordinate to one of the  $Mg^{2+}$  ions of the hydrolytic centre of TmPPase. It was therefore hypothesised that the conjugate base of **71** would have a stronger interaction with  $Mg^{2+}$ , due to the formal negative charge. Consequently, if the  $pK_a$  of the sulfamide is decreased it would be expected that there would be a greater proportion of the conjugate base and this would lead to an increase in potency (**Figure 36**). Contrary to this, an increase in  $pK_a$  would be expected to lead to a decrease in potency.



**Figure 36:** It was predicted that if the  $pK_a$  of the sulfamide of **71** was lower, the proportion of conjugate base **93** would be higher and the potency would increase. Whilst, an increase in  $pK_a$  would cause the proportion of **93** to be lower and this would lead to a decrease in potency.

### 3.6 Synthesis of SAR Library

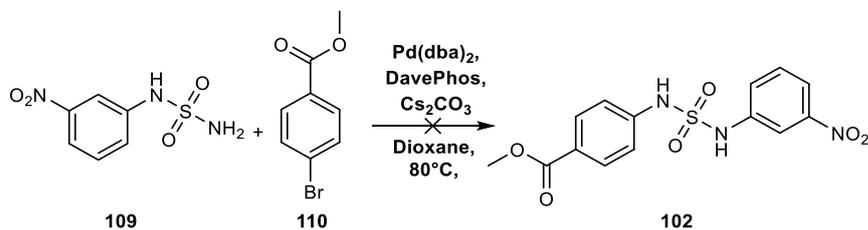
A library of sulfamide fragments were synthesised using the synthetic route shown in **Scheme 35**. [74] This was done to determine both the structure activity relationships of **71** and to determine if the  $pK_a$  of the sulfamide had an effect on the potency of the compounds against TmPPase.



**Scheme 35:** Synthesis of sulfamide fragment library. Conditions for **a** = Pd/C,  $H_2$ , MeOH, RT.

During the synthesis of the fragment library it was found that if the aryl ring of the primary sulfamide contained a strongly electron withdrawing group, it would not undergo the desired Buchwald reaction (**Scheme 36**). The electron withdrawing group will make the aromatic ring less electron rich,

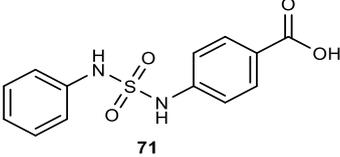
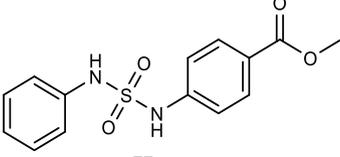
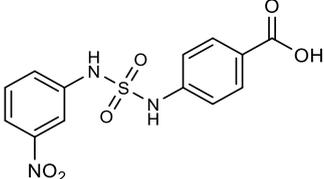
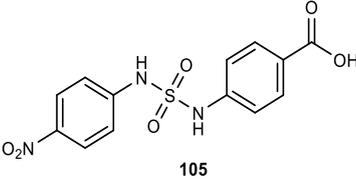
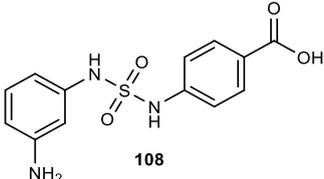
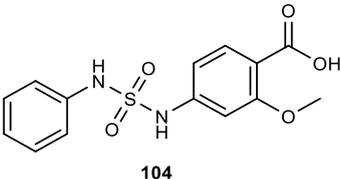
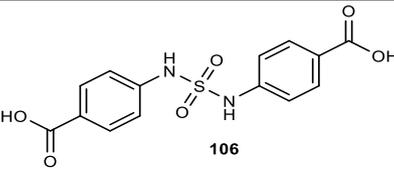
which in turn will make the primary sulfamide group less electron rich. This could mean that the primary sulfamide is unable to coordinate to the Pd catalyst and could account for the failure of the coupling reaction.



**Scheme 36:** The Buchwald reaction to form sulfamides was unsuccessful when the aryl ring of the primary sulfamide contained a strongly electron withdrawing group.

### 3.7 Biological Evaluation of Sulfamide Fragment Library

Sulfamide fragments **104-108** were biologically evaluated against TmPPase and their IC<sub>50</sub> values were determined (**Table 12**). In addition, MarvinSketch was used to predict the sulfamide pK<sub>a</sub> values of each of the fragments to establish any link between pK<sub>a</sub> and potency of the sulfamide fragments. All values were compared to those of the parent sulfamide fragment **71**.

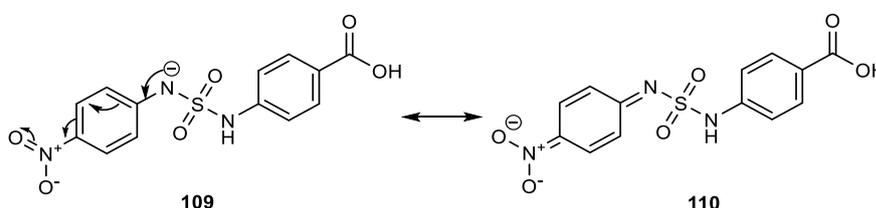
Compound	IC <sub>50</sub>	pK <sub>a</sub> *
 71	1353 ± 173 μM	10.71
 77	<5000 μM	9.41
 107	989 ± 42 μM	8.88
 105	<5000 μM	8.64
 108	<5000 μM	11.03
 104	<5000 μM	10.88
 106	<5000 μM	10.89

**Table 12:** Biological evaluation and pK<sub>a</sub> predicted pK<sub>a</sub> of sulfamide fragments **74-108**. Biological results are shown as IC<sub>50</sub> values. Predicted pK<sub>a</sub> values were determined using MarvinSketch.[78]

Compound **77**, which had ester in place of the carboxylic acid, had a lower  $pK_a$  than **71** and was found not to be active at the concentrations tested. This shows that the negative charge of the acid is essential for the activity of sulfamide **71**. As mentioned above, the carboxylic acid of **71** was predicted to coordinate to one of the  $Mg^{2+}$  ions within the hydrolytic centre. Therefore, the formal negative charge of the acid may be essential for coordination.

Compound **107**, which contained a *m*-NO<sub>2</sub> group on the left-hand aryl ring, had an IC<sub>50</sub> of 989  $\mu$ M and is more potent than **71**. In addition, the sulfamide group of **107** had a predicted  $pK_a$  of 8.88. The nitro group of **107** will make the corresponding aryl ring more electron deficient, which in turn makes the sulfamide group more electron deficient and therefore accounts for the decrease in the predicted  $pK_a$ . The decrease in  $pK_a$  could mean that deprotonation of the sulfamide occurs more readily and a higher percentage of **107** will exist as the conjugate base, accounting for the greater potency.

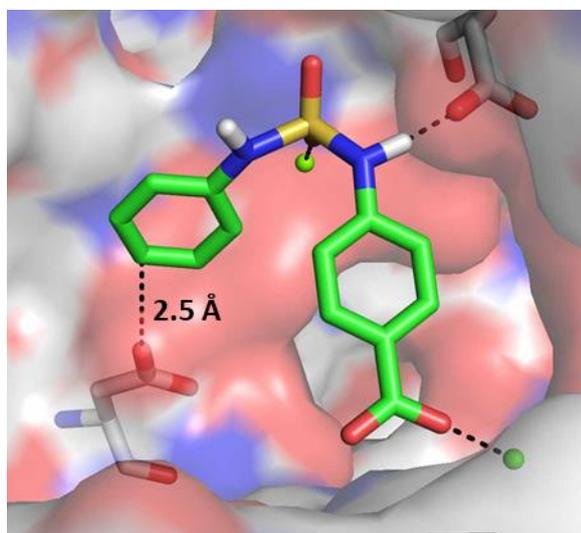
The addition of a *p*-NO<sub>2</sub> (compound **105**) on the left-hand aryl ring resulted in a further decrease in the predicted  $pK_a$  compared to both **71** and **107**. However, **105** was not active against TmPPase. The decreased  $pK_a$  is caused by the conjugation of the sulfamide with the nitro group. This conjugation means that the conjugate base of **105** is more stable due to the resonance of the negative charge (**Figure 37**). This resonance could also mean that the conjugate base is unable to coordinate to the  $Mg^{2+}$  and it is therefore inactive against TmPPase. In addition, the addition of the *p*-NO<sub>2</sub> could result in a steric clash with other residues of the hydrolytic centre causing **105** to adopt an unfavourable binding pose.



**Figure 37:** The negative charge resonance of the conjugate base **109** increases the  $pK_a$  of the sulfamide. This could suggest that the resonance form **110** is unable to coordinate to the  $Mg^{2+}$  of the hydrolytic centre and could account for the poor activity of **105**.

When an electron donating NH<sub>2</sub> group is present on the left-hand aromatic ring of the sulfamide fragment (Compound **108**) the pK<sub>a</sub> is predicted to be higher than that of all the previous fragments. In addition, **108** was found to be inactive against TmPPase. The increased pK<sub>a</sub> could mean that less of **108** will exist as the conjugate base and hence the coordination to Mg<sup>2+</sup> will be weak, resulting in the observed lack of activity towards TmPPase. Similarly, **104** also had an increased predicted pK<sub>a</sub> and was inactive against TmPPase. In this case the electron donating group (methoxy) is present on the same aryl ring as the carboxylic acid.

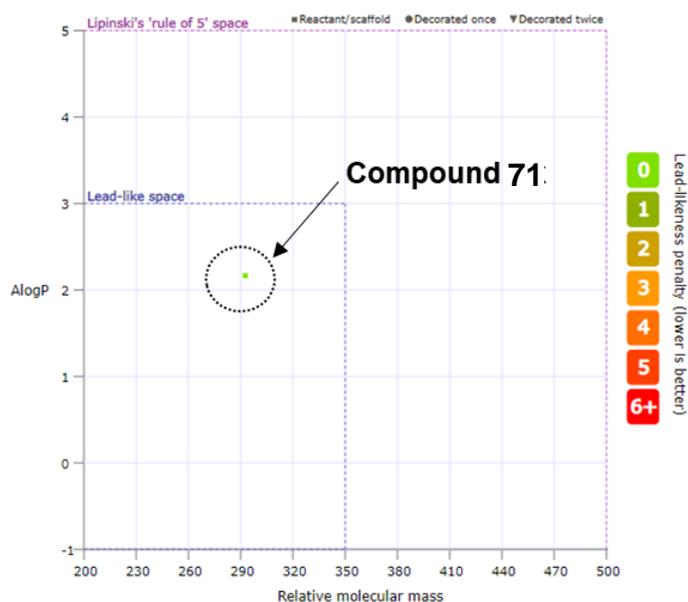
The addition of a carboxylic acid group to the left-hand aromatic ring in the *para*- position had no effect on the predicted pK<sub>a</sub> of the sulfamide and was found to be inactive against TmPPase (compound **106**). One possible explanation for the loss of activity can be obtained by examining the predicted docking pose of **71**. [63] This showed that there was an aspartic acid residue located close to the *para* position of the second aromatic ring (**Figure 38**). The addition of the *para*-carboxylic acid group would therefore lead to a repulsive interaction with aspartic acid residue and could force **106** to adopt an unfavourable binding pose.



**Figure 38:** The predicted docking pose of the parent sulfamide fragment **71**, highlight the close proximity of an aspartic acid residue to the *para*- position of the left-hand aryl ring. [47, 63]

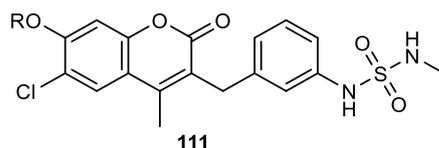
### 3.8 Suitability of Sulfamides for Hit-to-Lead Development

Sulfamide fragment **71** was analysed using LLAMA to determine if it occupied lead-like space (**Section 2.13**). [68] The predicted AlogP of **71** was 2.17, which is within the acceptable AlogP range. Furthermore, a plot of the predicted AlogP vs RMM showed that **71** occupies lead like space, with no lead-likeness penalties (**Figure 39**).



**Figure 39:** Plot of the predicted AlogP vs RMM of compound **71**. The plot indicates that **71** occupies lead-like space.

Extensive studies into the pharmacokinetic properties of sulfamide compounds has been conducted by Aoki *et al.* [79] They studied a series of sulfamide compounds that inhibit both Raf and mitogen-activated protein kinase (**Figure 40**). They showed that these sulfamide compounds undergo demethylation in the presence of mouse and human liver microsomes. However, they also found that the aromatic-sulfamide bond is metabolically stable in the presence of liver microsomes and do not hydrolyse to toxic anilines. In addition, all sulfamides tested had low clearances and high bio availabilities making them ideal candidates for drug design.

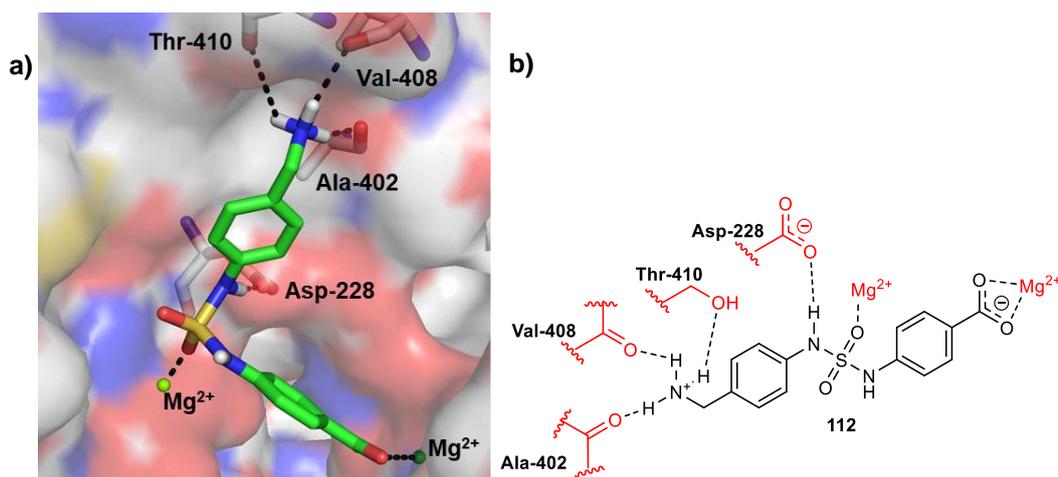


**Figure 40:** Aoki *et al*/found that the aromatic-sulfamide bond of a series of Raf/mek inhibitors based on **111** were metabolically stable, whilst the methyl group of the sulfamide functionality underwent demethylation in the presence of liver microsomes

Both the LLAMA analysis and the metabolic stability of aromatic sulfamides suggest that sulfamide fragment **71** is worth exploring further.

### 3.9 SPROUT Design of a Second-Generation Sulfamide Fragment

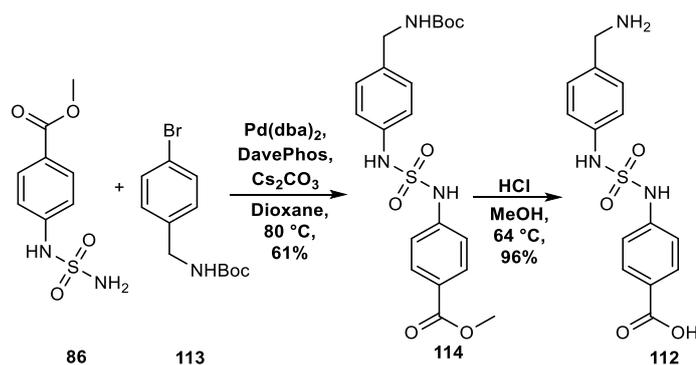
SPROUT was used to expand the parent sulfamide fragment **71**, with the aim of producing a more potent inhibitor. [61] The design process involved docking a primary amine and amide fragment to the Val-408 residue of the hydrolytic centre of TmPPase. To connect **74** with the docked amine and amide fragments a number of spacer templates were selected including; alkyl, phenyl, amide, and ether groups. This resulted in the generation of the second generation sulfamide fragment **112**, which had a *para*- substituted benzylamine moiety in place of the phenyl ring of **71** (Figure 41). The primary amine of **112** was predicted to make hydrogen bonding interactions with Ala-402, Thr-410 and Val-408. There is a slight change in the conformation of the predicted binding pose of **112** compared to that of **71**. The carboxylic acid of **112** was predicted to coordinate to one of the Mg<sup>2+</sup> in a bidentate manner and the sulfamide N-H of the benzyl amine was predicted to make a hydrogen bonding interaction with Asp-228. Compound **112** had a SPROUT score of -5.11 and was therefore predicted to be more potent than **71**, which has a SPROUT score of -4.06.



**Figure 41:** a) Predicted docking pose of SPROUT designed sulfamide fragment **112** visualised in PyMol, dashed lines represent interactions between **112** and the hydrolytic centre of TmPPase. [47, 61, 63] b) 2D representation of the docking pose. Residues of the hydrolytic centre are shown in red.

### 3.10 Synthesis of Second Generation Sulfamide Fragment **112**

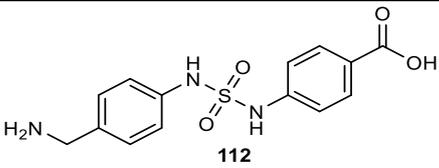
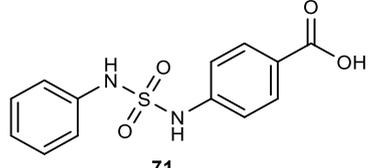
The second generation sulfamide fragment was synthesised using a Buchwald-Hartwig coupling reaction between primary sulfamide **86** and **113** to give intermediate **114**, which proceeded with a moderate yield of 61% (Scheme 37). [74] Ester hydrolysis and Boc-deprotection of **114** lead to the formation of **112** in near quantitative yield.



**Scheme 37:** Synthesis of second generation sulfamide fragment **112**

### 3.11 Biological Evaluation of Compound **112**

The second generation sulfamide fragment **112** was biologically evaluated against TmPPase and the  $IC_{50}$  value was determined and compared to that of **71** (Table 13).

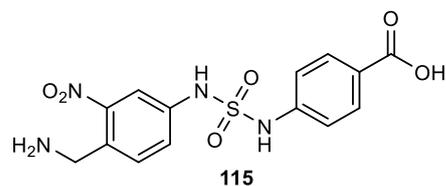
Compound	IC <sub>50</sub>
 <p style="text-align: center;"><b>112</b></p>	716 ± 64 μM
 <p style="text-align: center;"><b>71</b></p>	1353 ± 173 μM

**Table 13:** Biological evaluation of the second generation sulfamide fragment **112**. The IC<sub>50</sub> of **112** was also compared to that of **71**.

Sulfamide **112** had an IC<sub>50</sub> value of 716 μM against TmPPase and was more potent than the parent sulfamide fragment **71**. This suggests that the primary amine of **112** is making additional interactions with the hydrolytic centre of TmPPase, as was predicted by the SPROUT design.

### 3.12 Attempted Synthesis of an Electron Deficient Derivative of **115**

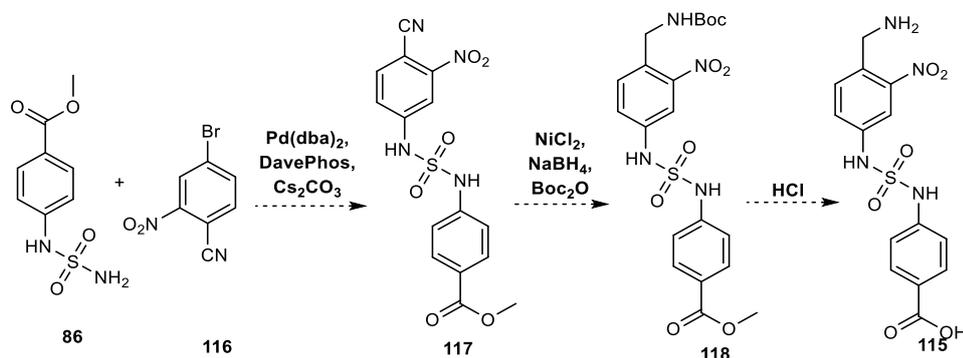
It was hypothesised that the addition of a strongly electron-withdrawing NO<sub>2</sub> group to **112** would lead to an increase in activity, due to the decreased pK<sub>a</sub> of the sulfamide moiety. Therefore, sulfamide **115** was identified as a key compound for synthesis (**Figure 42**).



**Figure 42:** The addition of an electron-withdrawing NO<sub>2</sub> group to the second generation sulfamide fragment lead to the identification of sulfamide **115**.

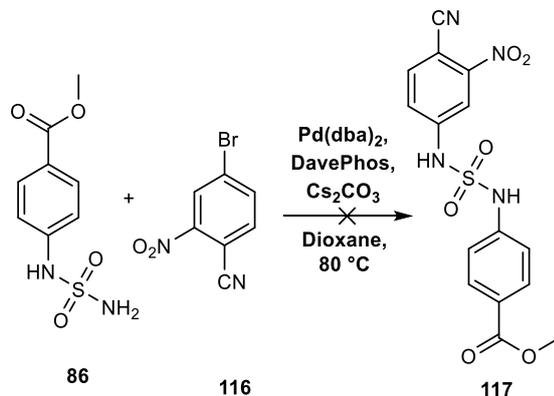
The attempted formation of **115**, summarised below (Scheme **38**), involved the Buchwald-Hartwig coupling of primary sulfamide **86** with **116**. This would give intermediate **117**, which would be followed by the conversion of the nitrile group of **117** to the corresponding Boc-protected amine **118**. Boc-

deprotection and ester hydrolysis would have then lead to the formation of the desired sulfamide fragment **115**.

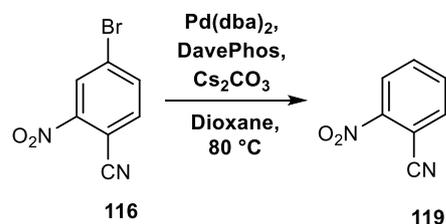


**Scheme 38:** Synthetic strategy for the formation of sulfamide fragment **115**.

LCMS analysis revealed that the formation of **117** was unsuccessful due to the rapid debromination of **116** (**Scheme 39** and **Scheme 40**). The debromination is likely to have occurred due to the presence of two strongly electron withdrawing groups on aryl bromide **116**, thus making a highly electron deficient aryl bromide.



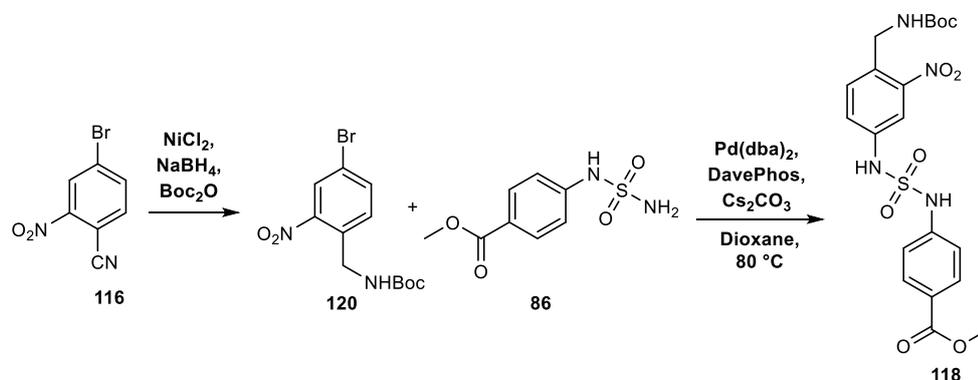
**Scheme 39:** LCMS analysis revealed that the formation of sulfamide **117** was unsuccessful



**Scheme 40:** Debromination of **116** to give **119**. This was likely caused by the presence of two strongly electron withdrawing groups on **116**.

One alternative route towards **118** involves the conversion of the nitrile group of **116** to the Boc-protected amine **120** (**Scheme 41**). This would make the

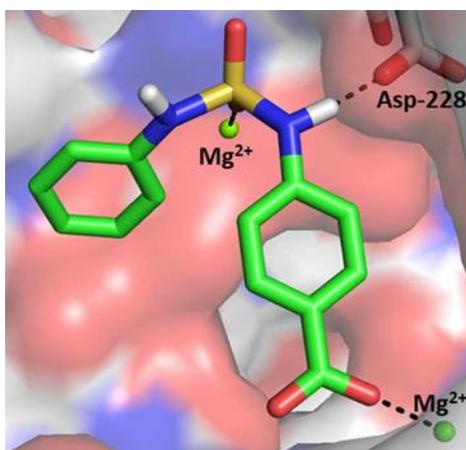
resultant aryl bromide **120** more electron rich and will reduce the rate of debromination and allow the Buchwald coupling reaction between **120** and **86** to occur.



**Scheme 41:** Proposed alternative synthesis of **118**.

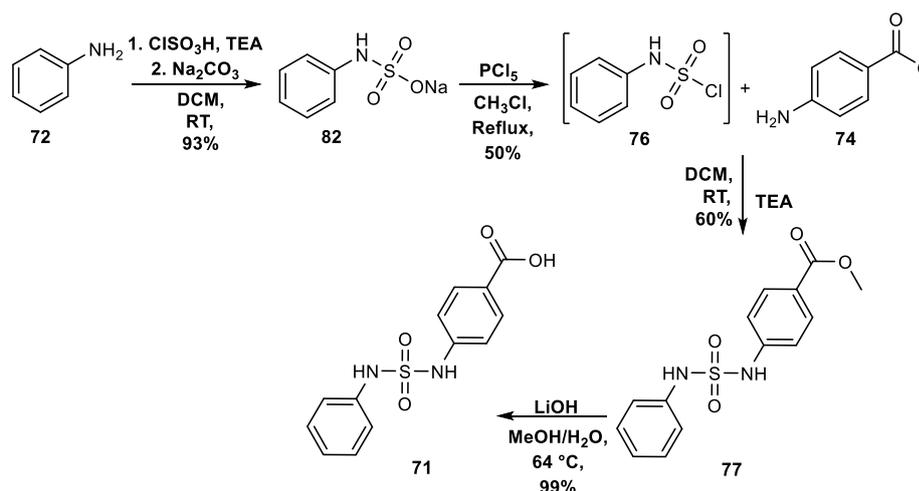
### 3.13 Conclusions

The *de novo* design programme SPROUT was used to design the novel sulfamide fragment **71**, which was designed to mimic phosphate and bind to the hydrolytic centre of TmPPase (**Figure 43**).



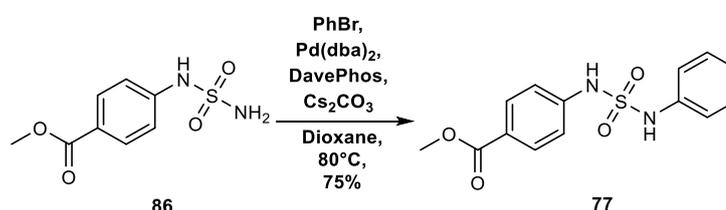
**Figure 43:** SPROUT designed sulfamide fragment **71**, which was designed to mimic phosphate and bind to the hydrolytic centre of TmPPase

The initial synthesis of **71** involved the formation of sodium sulfamate **82**, which was then chlorinated to give sulfamoyl chloride **76** (**Scheme 42**). This was followed by the coupling of **76** with aniline **74** to give **77** with a poor overall yield of 29%.



**Scheme 42:** The initial synthesis of **71**. The insertion of the sulfamide moiety proceeded with a poor overall yield of 29%.

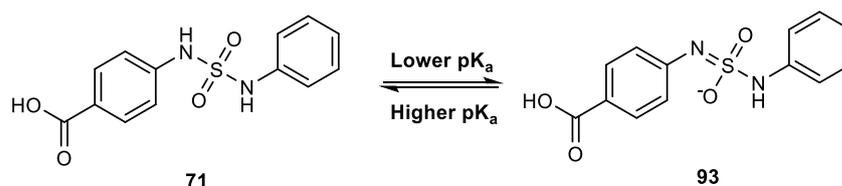
Due to the low yield of the previously used strategy for the formation of sulfamide **77**, an alternative strategy was explored. This strategy involved the formation of primary sulfamide **86**, which was then coupled with phenylbromide using a Buchwald-Hartwig coupling reaction to give **77** (**Scheme 43**). The overall yield for this strategy was 61%. Therefore, this strategy was selected for the future synthesis of sulfamides.



**Scheme 43:** Buchwald-coupling between **86** and bromobenzene lead to the formation of **77** in good yield.

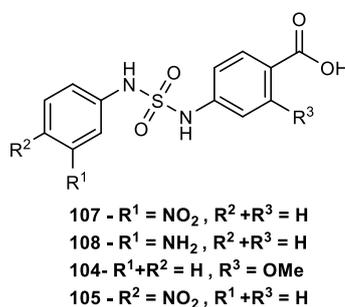
Sulfamide **71** was biologically evaluated against TmPPase and was found to have an  $IC_{50}$  value of 1353  $\mu$ M. The corresponding ester **77** was also tested and was found to be inactive against TmPPase. This supported the binding pose that was predicted by SPROUT and highlights the importance of the formal negative charge of the acid.

It was hypothesised that decreasing the  $pK_a$  of the sulfamide moiety would lead to an increase in potency against TmPPase, due to the increased proportion of the conjugate base **93** (**Figure 44**). To test this hypothesis a library of sulfamide compounds were synthesised (**Scheme 35**).



**Figure 44:** It was hypothesized that lowering the pK<sub>a</sub> of the sulfamide of **71** would lead to an increase in potency due to the increased coordination potential of **93**.

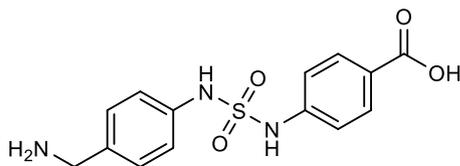
Compounds **104-107** were biologically evaluated against TmPPase. It was found that decreasing the pK<sub>a</sub> of the sulfamide moiety (compound **107**) led to an increase in potency, whilst an increase in the pK<sub>a</sub> (compounds **104** and **108**) led to the loss of activity. One exception to the trend was compound **105**, which contained a *para*-NO<sub>2</sub> group on the second aromatic ring. This was predicted to have the lowest pK<sub>a</sub> of the series, but it was inactive against TmPPase. It was thought that this was due to the conjugation of the lone pair of electrons of its conjugate base with the nitro group. This would mean that the lone pair is unable to coordinate to the Mg<sup>2+</sup> ion of the hydrolytic centre.



**Figure 45:** Biological evaluation of **104-108** suggested that the pK<sub>a</sub> of the sulfamide moiety is important for activity against TmPPase.

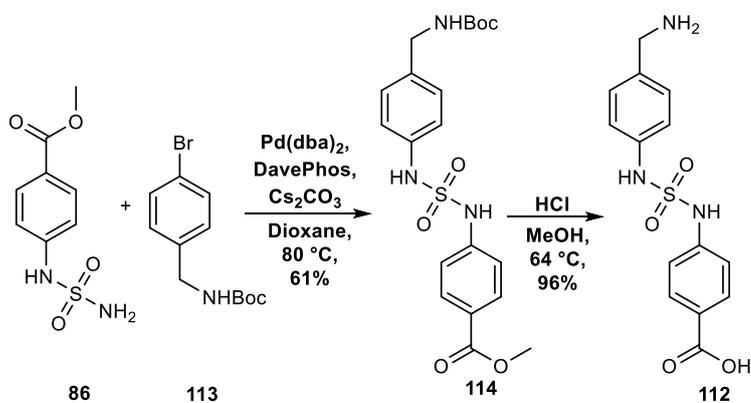
LLAMA analysis of the original sulfamide fragment **71** revealed that it occupies lead-like space and had a lead-likeness penalty score of 0 (**Figure 39**). In addition, Aoki *et al* showed that sulfamides have favourable properties and that aromatic-sulfamide bonds are not metabolically labile. This suggested that **71** was worth exploring further.

SPROUT was then used to expand the initial sulfamide fragment **71**, which resulted in the generation of compound **112** (**Figure 45**).



**Figure 46:** The second generation sulfamide fragment **112**, which contained a benzylamine moiety on the left hand aromatic ring.

The second generation inhibitor **112** was synthesised according to the scheme shown below (**Scheme 44**) and was then biologically evaluated against TmPPase. The  $IC_{50}$  was found to be 716  $\mu$ M, which means that it is more potent than the original sulfamide fragment **71**.

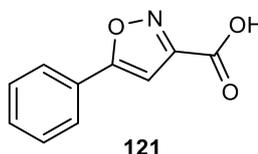


**Scheme 44:** Synthesis of second generation sulfamide fragment **112**

## Chapter Four - Heterocyclic Fragment Series

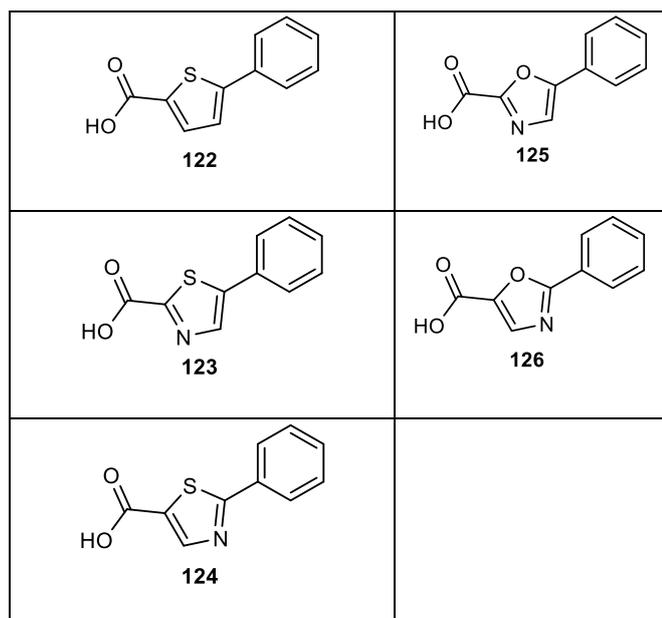
### 4.1 Previous vHTS Campaign

A virtual-highthroughput screening (vHTS) campaign has previously been conducted by collaborators at the University of Helsinki. [80] This involved docking small molecule fragments, from an in-house virtual library, into the hydrolytic centre of TmPPase using Glide. The compounds that gave the highest docking scores were then biologically evaluated against TmPPase. This led to the identification of isoxazole **121** (**Figure 47**), which inhibited the activity of TmPPase by 30% at a concentration of 500  $\mu\text{M}$ . An  $\text{IC}_{50}$  value could not be obtained at the concentrations tested.



**Figure 47:** A fragment screening campaign carried out by collaborators led to the identification of Isoxazole **121**, which inhibited the activity of TmPPase by 30% at a concentration of 500  $\mu\text{M}$ .

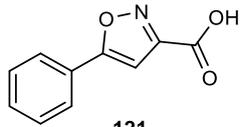
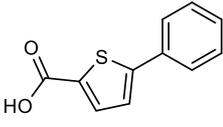
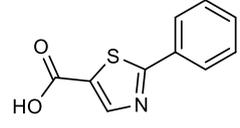
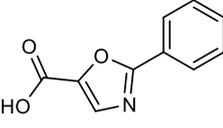
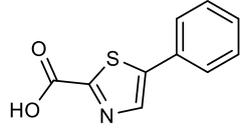
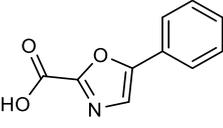
To determine if the activity of **121** could be improved with alternative heterocycles, a number of heterocyclic fragments were targeted (**Table 14**). These fragments were to be docked into the hydrolytic centre of TmPPase and biologically evaluated against TmPPase. If the biological results supported the docking models the *de novo* design programme SPROUT would be used to expand any moderately potent fragments.



**Table 14:** A number of different heterocyclic fragments were targeted for synthesis and biological evaluation.

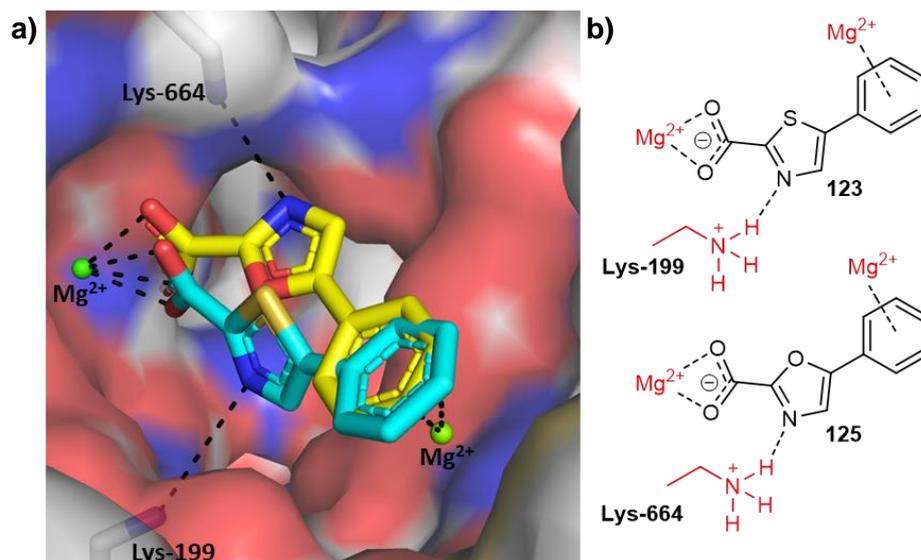
## 4.2 Docking of Heterocyclic Fragments 121-126

The heterocyclic fragments **121-126** were docked into the hydrolytic centre of TmPPase (PDB: 4AV3) using Glide. [47, 63] Overall, the predicted binding poses and docking scores (**Table 15**) were similar for each of the fragments, with only minor differences observed. This suggests that the fragments would bind to the hydrolytic centre of TmPPase in a similar manner and have similar potencies. The docking poses for each fragment are discussed in detail below.

Compound	Docking Score	Compound	Docking Score
 121	-6.11	 122	-5.10
 124	-5.81	 126	-5.98
 123	-5.86	 125	-6.08

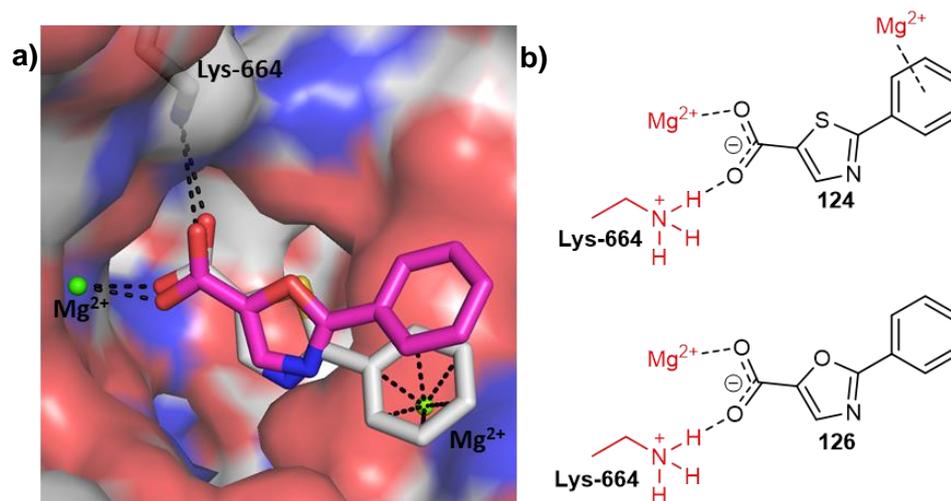
**Table 15:** The Glide docking scores of heterocyclic fragments **121-126**

Firstly, the predicted binding poses of oxazole **125** and thiazole **123** are shown below (**Figure 48**). The carboxylic acid groups of both **125** and **123** were predicted to coordinate to the same  $Mg^{2+}$  ion of the hydrolytic centre in a bidentate manner. In addition, the phenyl rings of both **125** and **123** were predicted to form  $\pi$ -cation stacking interactions with the other  $Mg^{2+}$  ion. Finally, the heterocyclic nitrogen of **123** was predicted to form a hydrogen bonding interaction with Lys-199. Whilst, the heterocyclic nitrogen of **125** was predicted to make a similar interaction with Lys-664.



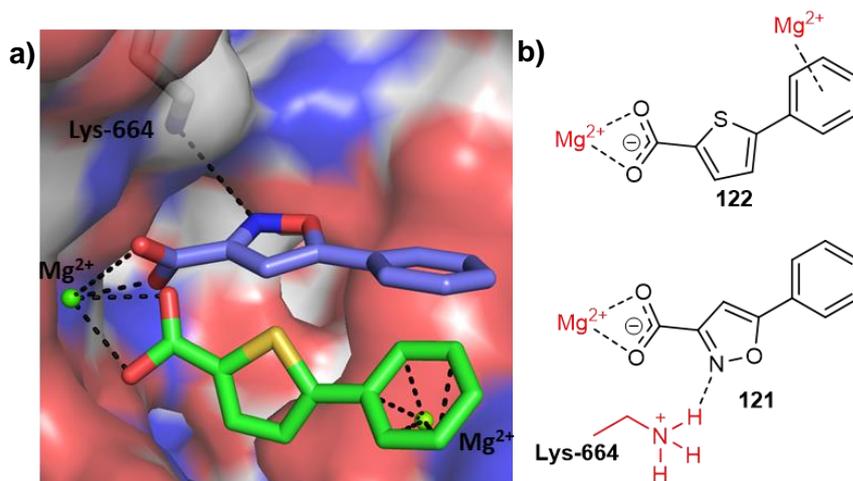
**Figure 48:** a) The predicted docking poses of thiazole **123** and oxazole **125** within the hydrolytic centre of TmPPase, visualised in PyMol (PDB: 4AV3). Dashed lines represent interactions between the fragments and the hydrolytic centre. [47, 63] b) The 2D representations of the docking poses of **123** and **125**. The residues of the hydrolytic centre are shown in red.

The predicted binding poses of thiazole **124** and oxazole **126** are shown below (**Figure 49**). The carboxylic acid groups of both **124** and **126** were predicted to make a monodentate interaction with one of the Mg<sup>2+</sup> ions, as well as a hydrogen bonding interaction with Lys-664. In addition, the phenyl ring of **124** was predicted to form a  $\pi$ -cation stacking interaction with the other Mg<sup>2+</sup> ion. No other interactions were predicted for **126**.



**Figure 49:** a) The predicted docking poses of thiazole **124** and oxazole **126** within the hydrolytic centre of TmPPase, visualised in PyMol (PDB: 4AV3). The dashed lines represent the predicted interactions between the fragments and the hydrolytic centre. [47, 63] b) The 2D representation of the predicted docking poses of **124** and **126**. The residue of the hydrolytic centre are shown in red.

Finally, the predicted binding poses of thiophene **122** and isoxazole **121** are shown below (**Figure 50**). The carboxylic acid groups of both **122** and **121** were predicted to coordinate with one of the  $Mg^{2+}$  ions of the hydrolytic centre in a bidentate fashion. A  $\pi$ -cation stacking interaction was also predicted between the phenyl ring of **122** and the other  $Mg^{2+}$  ion. Whilst, the heterocyclic nitrogen of **121** was predicted to make a hydrogen bonding interaction with Lys-664.

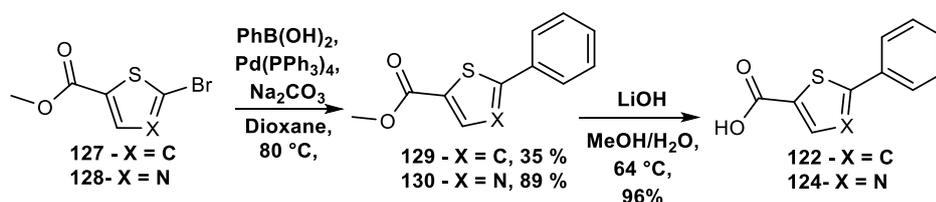


**Figure 50:** a) The predicted docking poses of thiophene **122** and oxazole **121** within the hydrolytic centre of TmPPase, visualised in PyMol (PDB: 4AV3). Dashed lines represent interactions between the fragments and the hydrolytic centre. [47, 63] b) The 2D representation of the docking posed of **122** and **121**. The residues of the hydrolytic centre are shown in red.

### 4.3 Synthesis of Hererocyclic Fragments

#### 4.3.1 Synthesis of Thiophene **122** and Thiazole **124**

Thiophene **122** and thiazole **124** were synthesised according to the scheme outlined below (**Scheme 45**). Initially, a Suzuki coupling reaction between the appropriate heterocyclic bromide (**127** and **128**) and phenylboronic acid lead to the formation of intermediates **129** and **130**. [70] Ester hydrolysis of both **129** and **130** lead to the formation of the desired compounds **122** and **124** with near quantitative yields.

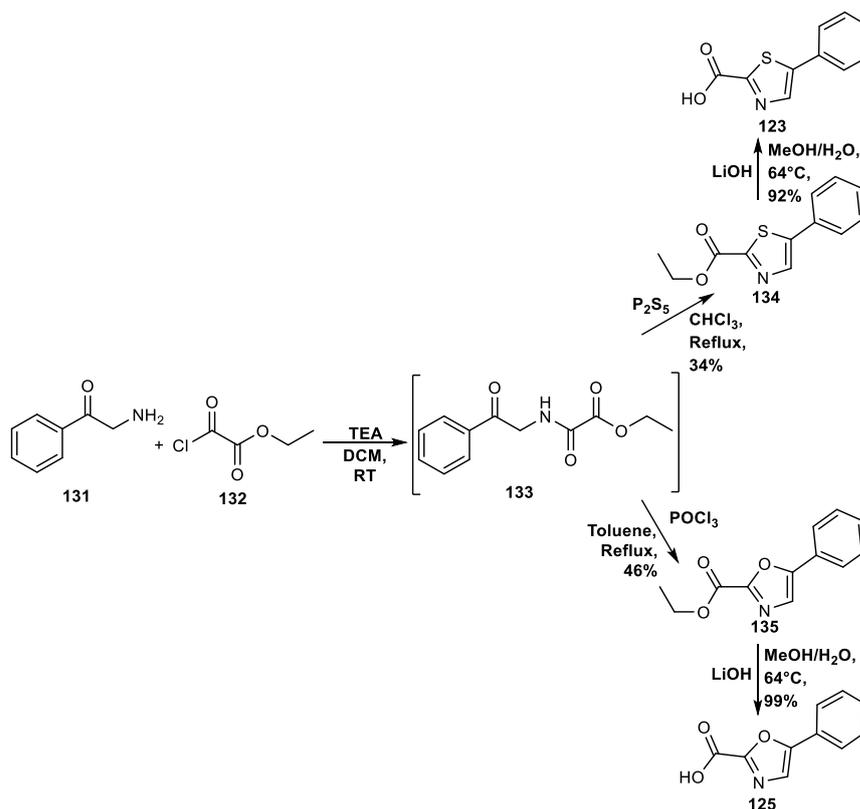


**Scheme 45:** Synthesis of thiophene **122** and thiazole **124**

#### 4.3.2 Synthesis of Thiazole **123** and Oxazole **125**

The synthesis of both thiazole **123** and oxazole **125** involved the cyclisation of intermediate **133**, which was obtained through the coupling of 2-aminoacetophenone **131** with ethylchlorooxoacetate **132** (**Scheme 46**).

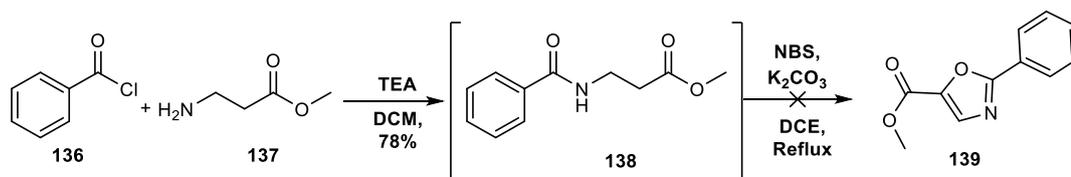
Heating intermediate **133** in the presence of phosphoryl chloride lead to the formation of oxazole **135**, whilst heating **133** in the presence of phosphorous pentasulfide lead to the formation of thiazole **134**. [81, 82] Both the cyclization reactions had moderate yields of 34% and 46% respectively. Ester hydrolysis of both **134** and **135** lead to the formation of **123** and **125** in near quantitative yields.



**Scheme 46:** Synthesis of thiazole **123** and oxazole **125**

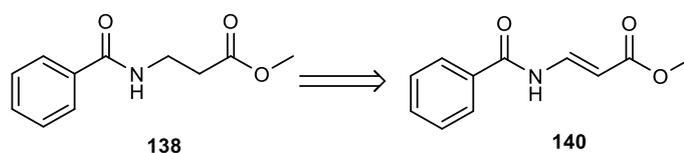
#### 4.3.3 Synthesis of Oxazole **126**

The synthetic strategy that was used to target oxazole **126** is summarised below (**Scheme 47**). [83] Firstly, benzoyl chloride **136** was coupled with amine **137** to give intermediate **138** with a moderate yield of 77%. The cyclisation of **138**, which involved heating **138** in the presence of *N*-bromosuccinimide (NBS) and  $K_2CO_3$ , was unsuccessful. LCMS and NMR analysis revealed a complex mixture and that there was no formation of **139**. One possible reason for this was the thermal decomposition of intermediate **138**. The reaction was therefore repeated at room temperature (RT) and used recrystallized NBS. However, the same complex mixture was observed by LCMS analysis.



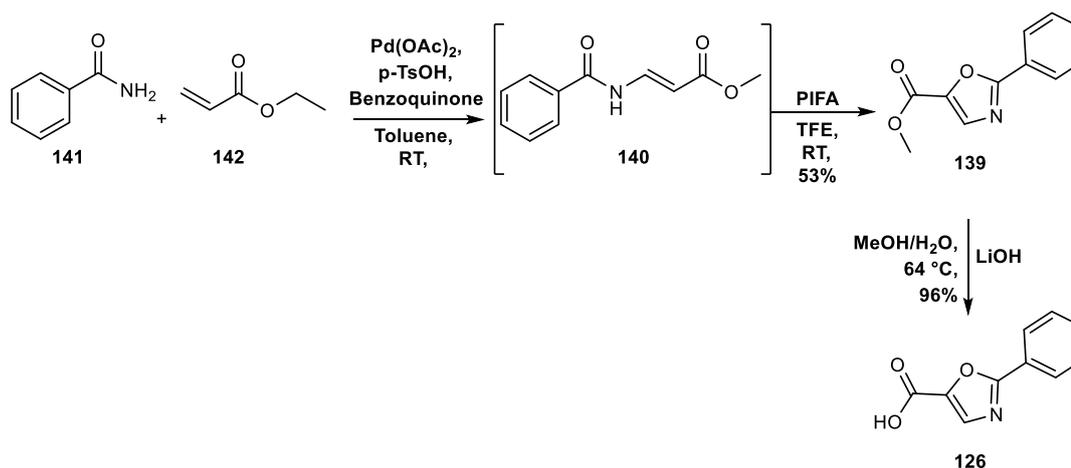
**Scheme 47:** Attempted synthesis of oxazole **139**. The cyclisation of intermediate **138** was unsuccessful, which may have been caused by the decomposition of **138**.

It was hypothesised that if intermediate **138** was replaced with the unsaturated derivative **140** (**Figure 51**), then the cyclisation would be more favourable.



**Figure 51:** Intermediate **140** was predicted to undergo the desired cyclisation reaction more readily than the saturated intermediate **138**.

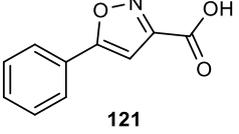
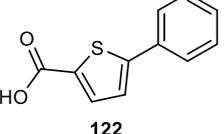
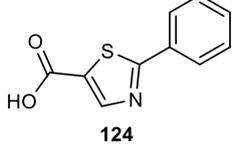
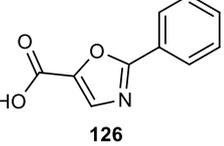
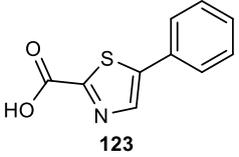
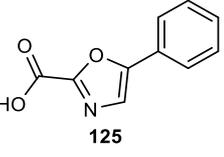
Intermediate **140** was formed through the Pd(OAc)<sub>2</sub> catalysed coupling of benzamide **141** with methyl acrylate **142** (**Scheme 48**). [84] This was followed by the cyclisation of **140** using [bis(trifluoroacetoxy)iodo]benzene (PIFA), which lead to the formation of **139** with a moderate yield of 53%. Finally, ester hydrolysis of **139** lead to the formation of the desired oxazole **126** in near quantitative yield.



**Scheme 48:** Synthesis of oxazole **126**

#### 4.4 Biological evaluation of Heterocyclic Fragments

Heterocycles **122-126** were biologically evaluated against TmPPase and their IC<sub>50</sub> values were determined (**Table 16**).

Compound	IC <sub>50</sub>	Compound	IC <sub>50</sub>
 121	>5000 μM	 122	>5000 μM
 124	22 ± 24 μM	 126	2025 ± 550 μM
 123	789 ± 36 μM	 125	>5000 μM

**Table 16:** Biological evaluation of heterocycles **121-126** against TmPPase.

Thiazole **124** had an IC<sub>50</sub> of 22 μM and was more potent than **123**, which had an IC<sub>50</sub> of 789 μM. A similar pattern was observed between oxazoles **126** and **125**. Oxazole **126** had the highest potency with an IC<sub>50</sub> of 2025 μM, whilst **125** was not active at the concentrations tested. In addition, thiophene **122** was not active at the concentrations tested. These results suggest that the presence of a heterocyclic nitrogen is essential for activity and the preferred position of this nitrogen is adjacent to the phenyl ring. When the nitrogen is in this position it could be close enough to an amino acid residue that it can partake in a hydrogen bonding interaction. Conversely, when the nitrogen is adjacent the carboxylic acid it is unable to partake in a hydrogen bonding interaction. This is the contrary to what was predicted by the docking studies, which predicted hydrogen bonding interactions between the heterocyclic nitrogen of **123** and **125** with Lys-664 and Lys-199 respectively. Whilst, the heterocyclic nitrogen of **124** and **126** were not predicted to make any interactions.

Both thiazole **124** and oxazole **126** were more potent than isoxazole **121**, which has an oxygen atom adjacent to the phenyl ring. It has been observed that heterocyclic oxygen atoms are poor HBAs, this therefore means that any hydrogen bonding interactions it partakes in are much weaker than those of

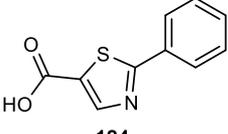
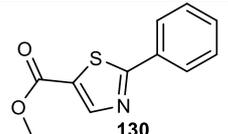
the corresponding nitrogen atoms of **124** and **126** and could account for the poor potency of **121**.

Comparing the IC<sub>50</sub> values of thiazole **124** with oxazole **126** and **123** with **125** suggests that thiazoles are preferred over oxazoles. There are a number of potential reasons for this. Firstly, the thiazole nitrogen is more basic than the oxazole nitrogen and a number of studies suggest that the thiazole nitrogen is a more effective HBA. [85, 86] Therefore, the thiazole nitrogen could be forming a stronger HB interaction with an amino acid residue of TmPPase. Finally, the thiazole ring has more aromatic character than the oxazole ring. This would be beneficial to the activity if the heterocycle was involved in a  $\pi$ -stacking interaction. [87]

Overall, the results suggest that the heterocycles are not binding as was predicted by the docking studies. This means that they could be binding to a different part of the hydrolytic centre or they are binding to an allosteric site. Therefore, a co-crystal structure of **124** and TmPPase is needed before SPROUT can be used to expand the fragment. Work is currently ongoing to obtain this crystal structure.

#### 4.4.1 Biological Evaluation of Methyl Ester **130**

The methyl ester of **124** (compound **130**, **Table 17**) was biologically evaluated against TmPPase and the IC<sub>50</sub> value was determined and compared to that of **124**.

Compound	IC <sub>50</sub>
 <p style="text-align: center;"><b>124</b></p>	22 ± 24 $\mu$ M
 <p style="text-align: center;"><b>130</b></p>	>5000 $\mu$ M

**Table 17:** Biological evaluation of methyl ester **130**. The IC<sub>50</sub> value was compared to that of the corresponding acid **124**

It was found that methyl ester **130** was not active at the concentrations tested, suggesting that the carboxylic acid is essential for activity.

#### 4.5 Ligand Efficiencies of Active Fragments

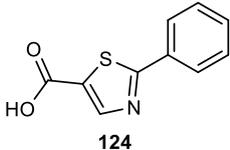
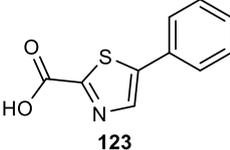
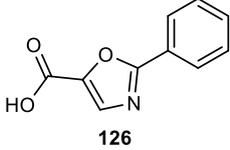
The ligand efficiency is defined as the binding affinity per heavy atom of a ligand that is binding to a particular protein. The determination of the ligand efficiency of a fragment is often used to determine if a fragment is worth further development. [88]

The equation used to determine the ligand efficiency is shown below (**Equation 1**). [89] The closer the calculated ligand efficiency of a compound is to 1, the more efficiently that compound binds to the protein. Generally, if a compound has a calculated ligand efficiency of <0.2 it is considered poor, if it is between 0.2-0.4 it is considered moderate and if it is >0.4 it is considered good.

$$LE = 1.4(-\log_{10} IC_{50})/N$$

**Equation 1:** The equation used to determine the ligand efficiency (LE) of a fragment. N = No. of heavy atoms.

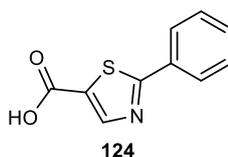
Highlighted below are the ligand efficiencies of the three active heterocyclic fragments **123**, **124** and **126** (**Table 18**). Unsurprisingly, **124** was found to have an excellent ligand efficiency of 0.47. This suggests that **124** is worth exploring further once a co-crystal structure has been obtained.

Compound	Ligand Efficiency
 124	0.47
 123	0.31
 126	0.27

**Table 18:** Calculated ligand efficiencies of **123**, **124** and **126**. LE was calculated using equation 1.

## 4.6 Conclusions

A small library of heterocyclic fragments were synthesised and biologically evaluated against TmPPase. These fragments were based on isoxazole **121**, which was identified by collaborators using vHTS. The results indicated that; the heterocycle, the presence of a heterocyclic nitrogen and the position of the nitrogen are important for biological activity. The most active fragment was thiazole **124** that had an  $IC_{50}$  of 22  $\mu$ M (**Figure 51**), which also had an excellent ligand efficiency of 0.47.



**Figure 51:** The most active fragment was thiazole **124**.

Each of the fragments were docked into the hydrolytic centre of TmPPase. Overall, the observed trends for the predicted docking scores and predicted docking poses did not follow the trends for the biological results. This suggests that the fragments are either binding to the hydrolytic centre in a different manner than was predicted by the docking models, or they are binding to an allosteric site. Therefore, SPROUT could not be used to

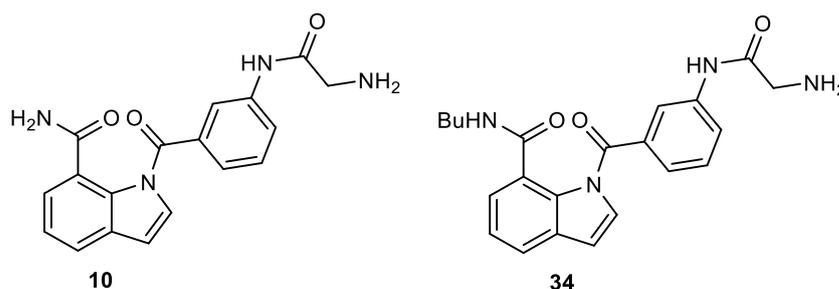
develop a second generation inhibitor. However, the excellent ligand efficiency of **124** suggests that it is worth exploring further. Therefore, work is currently ongoing to obtain a crystal structure of TmPPase with **124** bound. If a crystal structure is obtained SPROUT can be used to further develop the fragments.

## Chapter 5 - Conclusions and Future Work

The aim of this study was to identify small molecule inhibitors of a M-PPase using structure based drug design. Three approaches were employed to achieve this aim, which are evaluated below.

### 5.1 Series 1

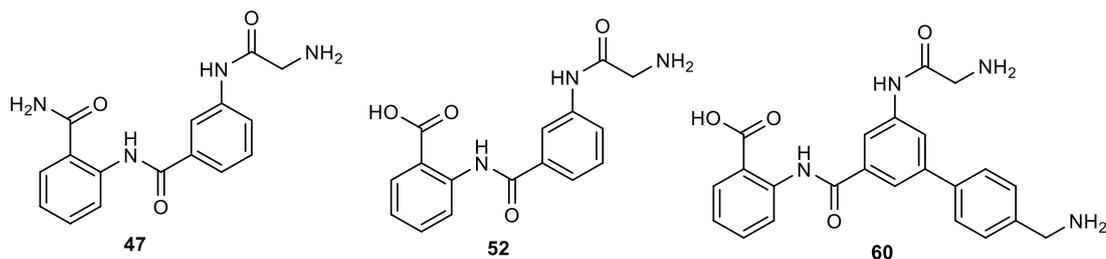
The first approach involved the SPROUT design of putative inhibitor **10** (Sections 2.1 and 2.2) (Figure 52). A number of synthetic challenges were encountered during the synthesis of **10**, which were caused by the lability of the indole *N*-acyl bond (Schemes 2-4). Therefore, the *n*-butyl amide **34** was targeted instead (Figure 52). It was found that **34** was not active when biologically evaluated against TmPPase, this was likely caused by the bulky nature of the butyl group (Table 4).



**Figure 52:** Putative inhibitor **10** was designed using SPROUT. The synthesis of **10** was unsuccessful due to the lability of the indole *N*-acyl bond. Due to the lack of synthetic accessibility of **10**, butyl amide **34** was targeted and found not to be active against TmPPase.

The poor potency and synthetic accessibility of **34** lead to the replacement of the indole moiety of **34** with a simple carboxamide, resulting in the identification of **47** (Section 2.7) (Figure 53). Whilst **47** was synthetically more accessible than **34**, it also showed poor potency against TmPPase (Table 5). Examination of the predicted binding pose of **47** showed that the primary amide of **47** was predicted to coordinate to one of the Mg<sup>2+</sup> ions of the hydrolytic centre (Section 2.9) (Figure 26). Therefore, the amide was replaced with a carboxylic acid (compound **52**, Figure 53). It was hypothesised that the carboxylic acid would increase the strength of the coordination with Mg<sup>2+</sup> ion, which in turn would lead to an increase in potency. The IC<sub>50</sub> value of **52** was found to be 179 μM and is the first example

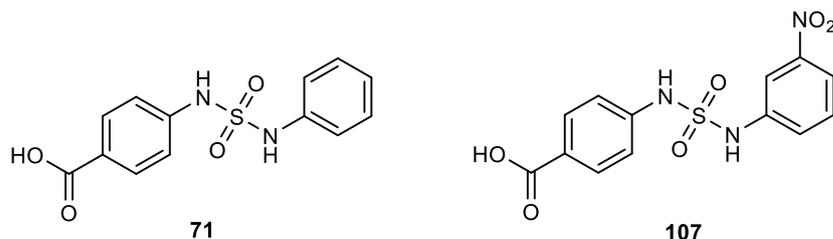
of an inhibitor of an M-PPase using SBDD (**Table 6**). Finally, SAR analysis of **52** supported the predicted binding pose as removal of key binding functionalities lead to decreases in activity (**Table 7**). Expansion of **52** using SPROUT failed to lead to an increase in potency, which may have been caused by the large size of the second generation compound **60** (**Section 2.13**) (**Table 9**) (**Figure 53**). This observation will be used to guide future design.



**Figure 53:** The simplified core **47** was inactive against TmPPase. Replacement of the amide of **47** with a carboxylic acid lead to the identification of **52**, which had an IC<sub>50</sub> of 179  $\mu$ M against TmPPase and is the first inhibitor of a TmPPase designed using SBDD. SPROUT expansion of **52** lead to the identification of **60**, which was less potent than **52**.

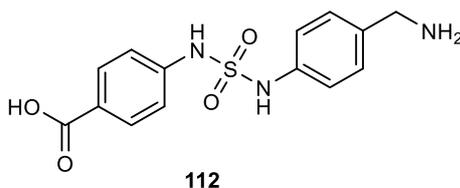
## 5.2 Series 2

The second approach involved the SPROUT aided design of the novel sulfamide fragment **71** (**Section 3.1**) (**Figure 54**), which was designed to mimic phosphate within the hydrolytic centre of TmPPase. This fragment was found to have a moderate potency against TmPPase, with an IC<sub>50</sub> value of 1353  $\mu$ M (**Table 10**). An efficient synthetic route towards **71** was also developed, involving a Buchwald-Hartwig coupling of the appropriate primary sulfamide with the appropriate aryl bromide (**Scheme 34**, **Section 3.5**). This route allowed the synthesis of a small library of sulfamide fragments (**Scheme 35**). The most potent fragment in this series was **107**, which had an IC<sub>50</sub> value of 989  $\mu$ M (**Table 11**). This suggests that the presence of groups that decrease the pK<sub>a</sub> of the sulfamide moiety are beneficial for activity.



**Figure 54:** Sulfamide **71** was designed to mimic phosphate within the hydrolytic centre of TmPPase and it was found to have a moderate IC<sub>50</sub> of 1353  $\mu$ M. Fragment **107** was the most potent fragment in a series of sulfamide fragments, which may be due to the decreased pKa of the sulfamide moiety.

Expansion of **71** using SPROUT lead to the identification of the second generation fragment **112** (**Section 3.10**) (**Figure 55**), which had an improved IC<sub>50</sub> value of 716  $\mu$ M (**Table 12**). The moderate potency of **112**, its synthetic accessibility and the favourable predicted properties of the sulfamide functionality suggest that it is a suitable candidate for hit-to-lead progression.

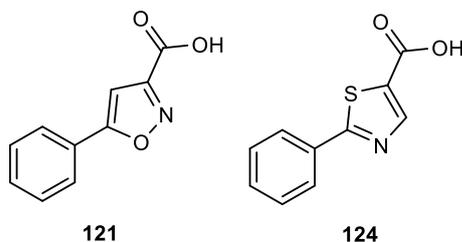


**Figure 55:** Expansion of the original fragment lead to the identification of **112**, which was more potent than **71**.

### 5.3 Series 3

Finally, a library of heterocyclic fragments were synthesised and biologically evaluated against TmPPase (**Section 4.1-4.4**). These fragments were based on the isoxazole fragment **121** (**Figure 56**), which was identified by collaborators during a vHTS campaign. Each of the fragments within the series, including isoxazole **121**, was docked into the hydrolytic centre of TmPPase (**Section 4.2**). The docking poses and docking scores were similar for each of the fragments. However, the fragments had a wide variety of potencies when tested against TmPPase. The most active fragment was thiazole **124** (**Figure 56**)(**Table 15**). This fragment had an IC<sub>50</sub> value of 22  $\mu$ M and had an excellent ligand efficiency. The biological results suggested that the fragments could be binding to the hydrolytic centre in a different manner to what was predicted by the docking studies, or they are binding to

an allosteric site. Due to the unknown binding mode of these fragments, SPROUT could not be used to guide further design. Therefore, work is currently ongoing to crystallise these fragments with TmPPase. However, thiazole **124** is a good candidate to take forward for further design due to its excellent activity and ligand efficiency.



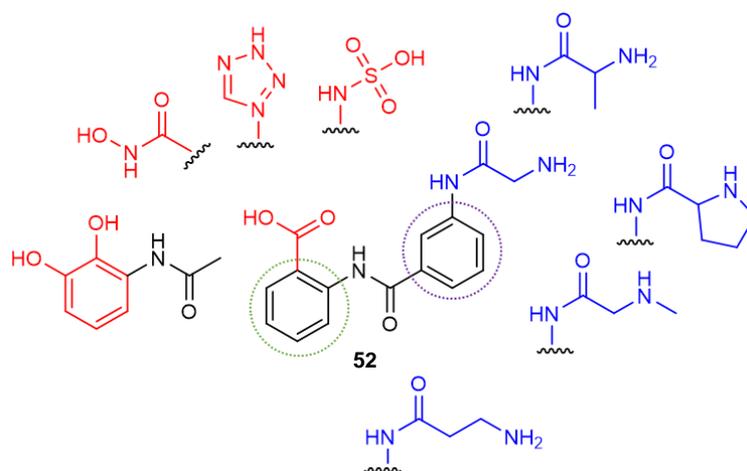
**Figure 56:** A number of heterocyclic fragments that were based on isoxazole **121** were synthesised and biologically evaluated against TmPPase. The most active fragment was thiazole **124**.

## 5.4 Overall Conclusions

There are currently a number of problems associated with current treatments of parasitic infections, including resistance and negative side-effects. Therefore, research into new treatments and new drug targets is required. This thesis has presented three series' of novel M-PPase inhibitors that are suitable candidates for hit-to-lead progression and could ultimately provide a good starting point for the development of the first anti-parasitic drug that targets a M-PPase. In addition, two of the series' presented herein are the first examples of M-PPase inhibitors designed using SBDD and suggests that SBDD can be applied to the discovery of novel M-PPase inhibitors.

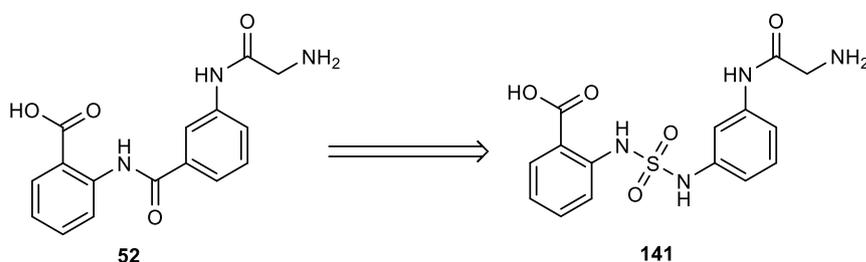
## 5.5 Future Work for Series 1

In the short term it is planned to get a better understanding of the SAR of compound **52**. This will be achieved by exploring three different regions of **52** (**Figure 57**). Firstly, the carboxylate of **52** will be substituted with different metal binding groups e.g hydroxamic acid. In addition, the effect of changing the substitution pattern of both aromatic rings will be explored. Finally, the glycine moiety will be replaced with alternative amino acid type moieties.



**Figure 57:** Further SAR analysis of compound **52** is planned. Alternative metal binding groups, shown in red, will be explored. In addition, SAR of the glycine moiety will also be explored, shown in blue. Finally, differing substitution patterns of both aromatic rings (circled) will also be explored.

It has been widely reported that amides derived from anilines are metabolically labile. Therefore, the aromatic amide of **52** (highlighted below, **Figure 58**) will need to be replaced with a more metabolically stable group. One potential candidate is a sulfamide group (compound **141**), which has been shown to be metabolically stable.

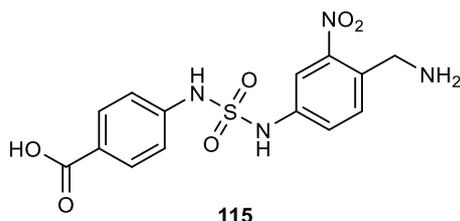


**Figure 58:** The metabolically labile amide of **52** will be replaced with a sulfamide (compound **141**), with the aim of making a more metabolically stable compound

The longer term goal is to improve the potency of **52**. To achieve this further cycles of SPROUT aided design will be used. As discussed above, large expansions are not tolerated so the best approach will be to make smaller sequential expansions. It is hoped that this approach will eventually lead to the identification of a lead compound.

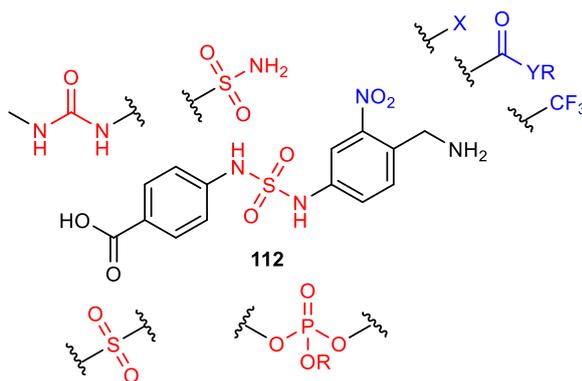
## 5.6 Future Work of Series 2

One of the short term aims of this series is to synthesis sulfamide **115** using the modified synthetic route discussed in **Section 3.13 (Figure 59)**. It is hypothesised that this fragment will be more potent than the parent fragment **112** due to the decreased  $pK_a$  of the sulfamide moiety.



**Figure 59:** Sulfamide **115** will be targeted. It is hypothesised that the decreased  $pK_a$  of the sulfamide moiety will lead to an increase in potency.

An additional aim is to further explore the SAR of **112 (Figure 60)**. This will include; exploring the effect of changing the substitution pattern of both the aromatic rings; changing the size and type of aromatic rings and exploring other metal binding moieties in place of the sulfamide (such as urea). The use of alternative electron withdrawing groups that decrease the  $pK_a$  of the sulfamide moiety will also be explored. One example is CF<sub>3</sub>, which is a strongly electron withdrawing group that doesn't have the toxicity associated with NO<sub>2</sub>.



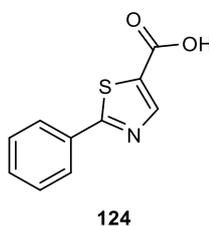
**Figure 60:** SAR analysis of **112** will involve exploring other metal binding groups in place of sulfamide, these are highlighted in red. Electron withdrawing groups on the benzylamine moiety will also be explored, these are highlighted in blue. Finally, different aromatic groups and substitution patterns will be explored

In the longer term further cycles of SPROUT aided design will occur, with the aim of increasing the potency of fragment **112**. The designed compounds will

continue to be synthesised using the synthetic strategy discussed above. It is hoped that this will allow the quick progression of this series and will lead to the identification of a more potent inhibitor.

### 5.6 Future Work of Series 3

Work is currently ongoing to obtain a crystal structure of TmPPase bound with thiazole **124** (**Figure 61**). This will confirm the binding site of **124**, which will mean that SPROUT could then be used to expand **124** to further increase potency. Due to the simplistic nature of **124** and its synthetic accessibility, further derivatives can be easily accessed.



**Figure 61:** Work is ongoing to obtain a crystal structure of **124** bound to TmPPase. This will be used to guide further design

## Chapter 6 - Experimental Section

### 6.1 Experimental for High-Throughput TmPPase activity assay [90]

#### 6.1.1 Preparation of Solutions

Solution A: Ascorbic acid (0.3 g, 1.70 mmol) was dissolved in ice cold 0.5 M HCl (10 mL) and the solution was kept at 0 °C. Solution B: Ammonium heptamolybdate (70 mg, 0.06 mmol) was dissolved in ice cold, fresh double distilled water (1 mL) and the solution was kept at 0 °C. Preparation of Solution C: To 10 mL of solution A was added solution B (1 mL).

Preparation of sodium arsenite solution: Sodium arsenite (5 g, 38.49 mmol), trisodium citrate dihydrate (5 g, 17.00 mmol) and ice cold acetic acid (5 mL, 87.43 mmol) was dissolved in water (245 mL).

Preparation of RX mixture: To a solution of water (280 mL) was added 1M Tris-HCl pH 8.0 (300 mL), 40 mM aqueous MgCl<sub>2</sub> (300 mL), 1M aqueous KCl (500 mL) and 1M aqueous NaCl (100mL).

Preparation of TmPPase Solution: To a solution of 20% DDM (22.5 µL) was added 30 mg mL<sup>-1</sup> soybean lecithin (40 µL). The mixture was then heated to 55 °C for 15 min before being cooled to ambient temperature. Buffer (36.5 µL) (20 mM MES pH 6, 3.5% glycerol, 2 mM DTT, 0.5% NM) was added and the resultant solution was stirred for 5 min before 13 mg mL<sup>-1</sup> TmPPase (1.5 µL) was added. 20 µL of the resulting solution was dissolved in the previously prepared RX mixture (1480 µL) to produce the desired TmPPase solution.

Preparation of Compound Solution: A 25 nM stock solution of each compound in DMSO was prepared before being diluted with water to various concentrations

#### 6.1.2 Assay

200 µL tube strips (8 tubes per strip) were used for the assay. To each tube was added 25 µL of compound solution followed by 15 µL of TmPPase solution. The tube strips were then sealed and incubated at 71 °C for 5 min.

Following this, 2 mM Na<sub>2</sub>PP<sub>i</sub> solution (10 µL) was added to each tube and the tubes were resealed and incubated at 71 °C for a further 5 min. The tubes were then cooled to 0 °C for a period of 10 mins (after 8 min the tubes were centrifuged and any water drops under the seal were decanted off). Solution C (60 µL) was then added to each tube and the tubes were again cooled to 0 °C for 10 min. Sodium arsenite solution (90 µL) was then added to each tube and the tubes were stored at ambient temperature for 1 h. Each tube was then emptied into a clear, PS Micro Plate 96 Well and the absorbance was measured at 860 nm using a MultiSkan Go.

## **6.2 Computational Methods**

All computational aspects of this research was carried out on a LINUX computer using the CentOS operating system.

### **6.2.1 *In silico* Docking Using Glide**

The Glide docking programme within the Maestro suite was used for the docking studies mentioned throughout this work. All docking procedures used the resting-state crystal structure of TmPPase (PDB: 4AV3), which was obtained from the RSCB PDB databank. All compounds and ligands were energy minimised using the ligand preparation tool of Maestro prior to docking. This used the OPLS3 force field and charges were assigned to ionisable groups using a pH range of 5.0-9.0.

The crystal structure of TmPPase was prepared using the protein preparation tool within Maestro, which included the assignment of protonation and deprotonation sites of amino acid residues at a pH of 7.4 and removal of water molecules. On completion, all recommended changes to the crystal structure were accepted. The prepared protein was then energy minimised using the OPLS3 force field. This was followed by the generation of a Glide grid, which was a 10x10x10 box around the hydrolytic centre of the TmPPase crystal structure. The centroid of this box was the two Mg<sup>2+</sup> ions of the hydrolytic centre.

The minimised compounds were then docked into the previously defined Glide grid using the Glide docking programme of Maestro, the standard

precision settings were used. For each round of docking 5 poses were generated and each generated pose was subjected to post-docking minimisation. Each of the docking poses were then inspected by eye.

The resultant docking poses were then loaded into PyMol to visualise the predicted contacts. A screen capture of the docking poses was taken from PyMol and used for the docking images reported in this thesis. TmPPase was shown as a surface and the ligands were shown as sticks.

### **6.2.2 SPROUT**

Prior to carrying out SPROUT design, the TmPPase crystal structure that was used (PDB: 4AV3) was minimised within Maestro using the methodology described above. The minimised crystal structure was exported as a PDB file and imported into SPROUT.

The PDB crystal structure was initially loaded into CANGAROO and the receptor and cavity site were defined using a 10Å cut of the hydrolytic centre. This cut was made by selecting the two Mg<sup>2+</sup> ions of the hydrolytic centre and selecting the residues within a 10Å radius. The hydrogen bond acceptor and donor sites, along with any hydrophobic sites, were then explored using HIPPO. The ELEFANT module was then used to dock fragments onto these sites using fragments within SPROUTs built-in fragment library. The first round of design only involved docking to two sites. These fragments were then connected using SPIDER and used the spacer templates within an in-built library to generate a number of fragments. The fragments were scored and the top 10% were pruned using the ALLIGATOR module. The process was repeated and with each cycle of design one further binding site was chosen to expand to.

At the end of the design process the top 10 scoring ligands were selected for further analysis. These selected ligands were docked into the hydrolytic centre of TmPPase using Maestro and the methodology described above to improve the confidence of the proposed docking pose predicted by SPROUT. The re-docked compounds were then visually inspected and the most synthetically accessible compounds with the highest docking scores were chosen to synthesise.

### 6.3 General Experimental Information and Instrumentation

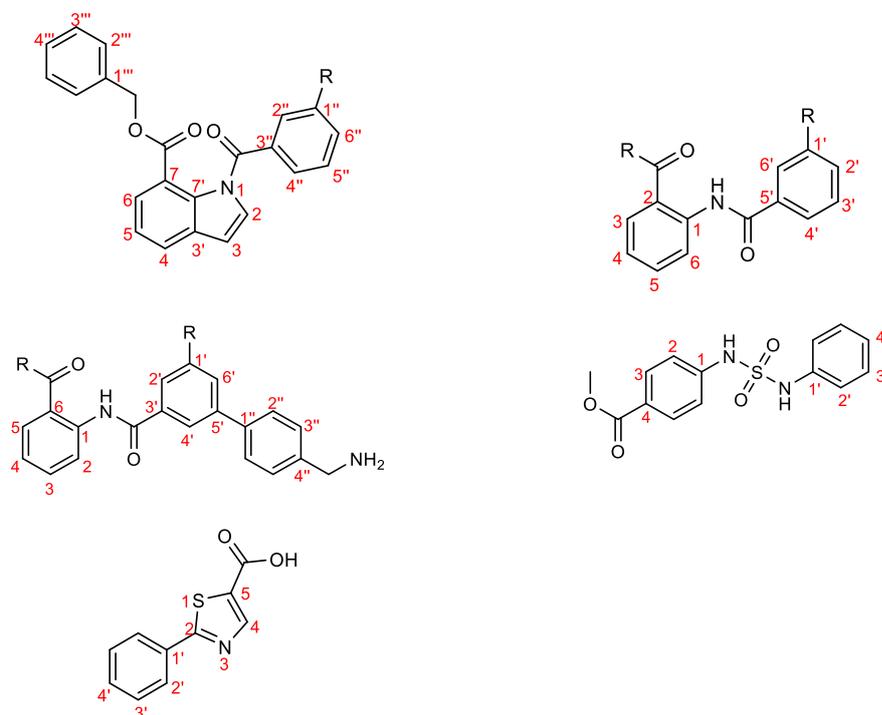
All reactions, unless otherwise stated, were carried out under an inert nitrogen atmosphere with anhydrous solvents and using a magnetic stirrer with a stirrer-hotplate. All reagents were obtained from commercial sources and were used without further purification. Anhydrous solvents were obtained from a PureSolv MD6 solvent purifications system. Thin layer chromatography (TLC) was performed on aluminium pre-coated silica gel plates (254  $\mu\text{m}$ ) supplied by Merck chemicals and visualised by UV-light (254 nm). Solvents were removed under reduced pressure using a Buchi rotary evaporator and a high vacuum pump was used for the removal of residual solvent. Silica flash chromatography was performed using Geduran silica gel 60. Reverse phase column chromatography was performed on a Biotage® Isolera Four with a Biotage® Ultra C18 12 g column (mobile phase: 0.1% formic acid in water + 0.1% formic acid in acetonitrile). Normal phase automated column chromatography was carried out using a Biotage® Isolera Four with Thomson pre-packed silica cartridges (various sizes) at various flow rates.

LCMS analysis was performed on a Bruker Daltronics instrument running on a gradient of increasing acetonitrile in water containing 0.1% formic acid on a 50  $\times$  20 mm C18 reverse phase column. Runs were conducted using a flow rate of 1 mL per min and samples dissolved in MeOH. High-resolution mass spectrometry was carried out using a Bruker MaXis Impact Time of Flight spectrometer using electron spray ionisation, with samples dissolved in MeOH. The reported masses correct to four decimal places. Unless stated otherwise, Proton and Carbon NMR spectra were recorded at 400 and 500 MHz (at room temperature) on a Bruker Advance 400 or 500 Fourier Transform spectrometer. Samples were dissolved in  $\text{CDCl}_3$ ,  $\text{DMSO}-d_6$  or  $\text{CD}_3\text{OD}$ . Chemical shifts are reported in ppm. Multiplicities are reported with coupling constants and are given to the nearest 0.1 Hz. The following abbreviations were used when displaying NMR data; s = singlet, d = doublet, t = triplet, q = quartet, sxt = sextet, dd = doublet of doublets, ddd = doublets of doublets of doublets, m = multiplet and app = apparent. Where needed,

two-dimensional correlation spectroscopy (2D-COSY), heteronuclear single quantum coherence spectroscopy (HSQC) and distortionless enhancement by polarisation transfer (DEPT) were used to aid assignment. IR spectra were recorded in solid phase on a Bruker Alpha Platinum ATR FTIR spectrometer with vibrational frequencies given in  $\text{cm}^{-1}$ . Analytical high performance liquid chromatography (HPLC) was conducted on an Agilent 1290 infinity series equipped with a UV detector and a Hyperclone  $\text{C}_{18}$  reverse phase column eluting with an increasing MeCN-Water gradient and 0.1% TFA over a run time of 5 minutes and a flow rate of  $0.5 \text{ mL min}^{-1}$ . HPLC analysis was conducted on compounds that were biologically evaluated only and samples were dissolved in MeOH. Melting points were measured on a Stuart SMP30 for crystalline solids only.

### 6.3.1 Compound Numbering for NMR

Listed below (**Table 19**) are the numbering systems used for the NMR assignments of aromatic H and C atoms.

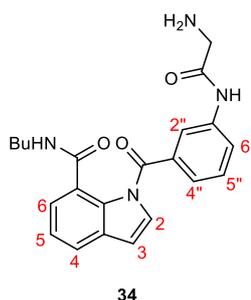


**Table 19:** The numbering systems used in the reporting of a compounds NMR spectrum.

### 6.3.1.1 Assignment of NMR Spectra

The NMR Spectra of each compound reported in this thesis were assigned using the methodology described below, using compound **34** as an example.

Firstly, an easily identifiable signal was chosen to start the assignment. In the case of **34** this signal was an apparent singlet at 8.16 ppm (see appendix for spectrum), which was assigned to 2''-H (**Figure 62**). The COSY spectrum of **34** showed that the signal for 2''-H coupled to two protons. One of these protons had a signal at 7.93 ppm, whilst the other had a signal at 7.52 ppm. Therefore, the signal at 7.93 ppm was assigned to 4''-H as it was ortho- to a carbonyl and would be the most de-shielded. Thus, the signal at 7.52 ppm was assigned to 6''-H. In addition, both the signals at 7.93 ppm and 7.52 ppm were found to couple to a triplet at 7.56 ppm that was assigned to 5''-H. One other triplet was observed at 7.33 ppm, which was assigned to 5-H of the indole. This was found to couple to signals at 7.78 ppm and 7.40 ppm, which were assigned to 6-H and 4-H respectively. 6-H was the most de-shielded as it was ortho- to a carbonyl. The two final aromatic protons were assigned as 2-H (at 7.46 ppm) and 3-H (at 6.78 ppm), these were simple to assign as 2-H is next to the electron withdrawing N of the indole. The aliphatic protons were also assigned using the COSY spectrum. The CH<sub>2</sub> of the Gly only coupled to one of the N-H signals, whilst all of the protons of the butyl group coupled with the aliphatic protons they were adjacent to.



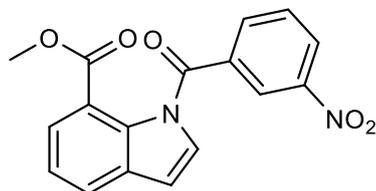
**Figure 62:** The numbering system used in the reporting of the NMR spectra of **34**.

The <sup>1</sup>H NMR spectral assignments were then used to assign the <sup>13</sup>C spectrum using HSQC and DEPT experiments.

This methodology was used for the assignment of all Spectra reported herein.

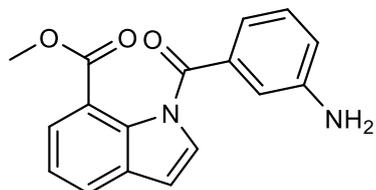
### 6.3.2 Synthesis of Series One Compounds

#### Synthesis of methyl 1-[(3-nitrophenyl)methyl]-1H-Indole-7-carboxylate (14)



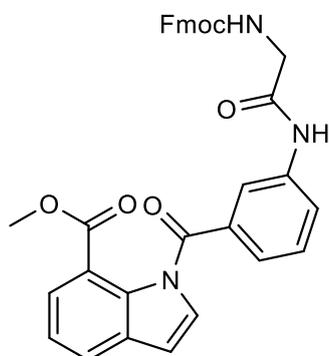
To a stirred solution of methyl indole-7-carboxylate (1.00 g, 5.70 mmol) in THF (30 mL) was added NaH (0.342 g, 14.25 mmol) portion wise at 0 °C. The reaction mixture was then stirred at 0 °C for 15 mins and then at rt for a further 15 mins. 3-Nitrobenzoyl chloride (1.58 g, 8.55 mmol) was then added and the reaction mixture was then stirred at RT overnight. The reaction mixture was then quenched with water (20 ml) and extracted with EtOAc (3 x 25 mL). The organic layers were then combined, dried with MgSO<sub>4</sub>, filtered and the solvent removed to yield a crude yellow oil. The yellow oil was purified using column chromatography (20% EtOAc-Petroleum Ether) to afford the title compound as a colourless solid (1.33 g, 72%). *R<sub>f</sub>* 0.71 (30% EtOAc-Petroleum Ether).  $\delta_{\text{H}}$  (500 MHz, DMSO) 8.59 (app d, 1H, *J* = 8.3 Hz, 6''-H), 8.56 (app s, 1H, 2''-H), 8.30 (app d, 1H, *J* = 8.3 Hz, 4''-H), 7.94-7.92 (m, 2H, 6-H and 5''-H), 7.65 (app d, 1H, *J* = 7.0 Hz, 4-H), 7.60 (app d, 1H, *J* = 3.6 Hz, 2-H), 7.46 (app t, 1H, *J* = 7.0 Hz, 5-H), 6.90 (app d, 1H, *J* = 3.6 Hz, 3-H), 3.69 (s, 3H, CH<sub>3</sub>).  $\delta_{\text{C}}$  (500 MHz, DMSO) 166.8 (C=O), 166.3 (C=O), 148.0 (1''-C), 135.8 (3''-C), 135.7 (4''-C), 134.0 (7'-C), 132.2 (3'-C), 132.0 (5''-C), 130.9 (2-C) 130.0 (6''-C), 127.8 (6-C), 125.0 (4-C), 124.3 (5-C), 123.6 (2''-C), 120.6 (7-C), 109.6 (3-C), 52.4 (OCH<sub>3</sub>).  $\nu_{\text{max}}/\text{cm}^{-1}$  (ATR); 3084, 1720, 1693, 1532, 1422, 1277. *m/z* (ESI) (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 325.0825. C<sub>17</sub>H<sub>12</sub>N<sub>2</sub>O<sub>5</sub> requires MH<sup>+</sup>, 325.0826).

### Synthesis of methyl 1-[(3-aminophenyl)methyl]-1H-Indole-7-carboxylate (15)



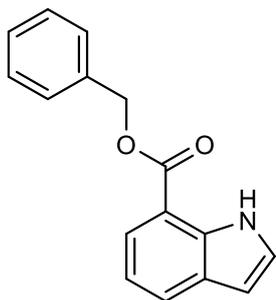
To a stirred solution of SnCl<sub>2</sub> (0.235 g, 1.24 mmol) in EtOH was added concentrated HCl (4.5 mL) at RT. The resultant cloudy solution was then heated to 65 °C and stirred for 15 mins. To this solution was added methyl 1-[(3-nitrophenyl)methyl]-1H-Indole-7-carboxylate **14** (0.100 g, 0.31 mmol). The reaction mixture was then stirred at 65 °C for 45 mins and then cooled to rt. The reaction mixture was then neutralised using 2M NaOH and the resultant solution was extracted with EtOAc (3 x 15 mL). The combined organic layers were then washed with brine (30 mL), dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a yellow residue. The residue was taken up in EtOAc (20 mL) and washed with a further amount of brine (20 mL). The organic layer was again dried over MgSO<sub>4</sub>, filtered and the solvent removed to give the title compound as a yellow solid. (0.074 g, 81%) R<sub>f</sub> 0.58 (30% EtOAc-Petroleum Ether). δ<sub>H</sub> (500 MHz, DMSO) 7.90 (app d, 1H, J = 7.8 Hz, 6-H), 7.59 (app d, 1H, J = 7.8 Hz, 4-H) 7.48 (app d, 1H, J = 3.6 Hz, 2-H), 7.39 (m, 1H, 5-H), 7.25 (m, 1H, 5''-H), 6.95 (s, 1H, 2''-H), 6.92-6.90 (m, 2H, 4''-H and 6''-H), 6.80 (app d, 1H, J = 3.6 Hz, 3-H), 5.53 (s, 3H, CH<sub>3</sub>). δ<sub>C</sub> (500 MHz, DMSO) 169.3 (C=O), 167.4 (C=O), 149.7 (1''-C), 133.5 (7'-C), 132.5 (3''-C), 132.4 (3'-C), 130.7 (2-C), 129.9 (5''-C), 125.2 (6-C), 125.0 (4-C), 123.5 (5-C), 121.0 (7-C), 119.1 (6''-C), 117.4 (4''-C), 114.9 (2''-C), 107.5 (3-C), 52.1 (CH<sub>3</sub>). ν<sub>max</sub>/cm<sup>-1</sup> (ATR); 3464, 3358, 1707, 1693, 1421, 1339. m/z (ESI) (100%, MH<sup>+</sup>); (Found MNa<sup>+</sup>, 317.0899. C<sub>17</sub>H<sub>14</sub>N<sub>2</sub>O<sub>3</sub> requires MNa<sup>+</sup>, 317.0897).

**Synthesis of methyl 1-{3-[2-({[(9H-fluoren-9-yl)methoxy]carbonyl}amino)acetamido]benzoyl}-1H-indole-7-carboxylate (16)**



To a stirred solution of methyl 1-[(3-aminophenyl)methyl] - 1 H - Indole-7-carboxylate **15** (0.500 g, 1.70 mmol) in DCM (15 mL) was added; Fmoc-Gly-OH (0.607 g, 2.04 mmol), EDC.HCl (0.391 g, 7.04 mmol) and DMAP (0.025 g, 0.2 mmol) at RT. The reaction mixture was then stirred for 12 hours at RT and then diluted with DCM (15 mL). The resultant solution was then washed with water (15 mL) and brine (15 mL), dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude yellow oil. The crude product was then purified using column chromatography (20% EtOAc-Petroleum Ether) to give the title compound as a crystalline colourless solid. (0.618 g, 63%) m.p 126-128 °C. R<sub>f</sub> 0.34 (2% MeOH-DCM). δ<sub>H</sub> (500 MHz, DMSO) 10.31 (s, 1H, NH amide) 8.21 (s, 1H, 2''-H), 7.88-7.95 (m, 4H, 3-H Fmoc, 6-H and 4''-H), 7.74 (app d, 2H, J = 7.4 Hz, 6-H Fmoc), 7.67 (app t, 1H, J = 6.1 Hz, Fmoc-NH), 7.61 (m, 2H, 4-H and 6''-H), 7.55 (m, 1H, 5''-H), 7.52, (d, 1H, J = 3.6 Hz, 2-H) 7.42 (m, 3H, 5-H and 5-H Fmoc), 7.32 (dd, 2H, J = 10.7, 4.2, 4-H Fmoc), 6.84, (d, 1H, J = 3.6 Hz, 3-H), 5.52 (s, 3H, OCH<sub>3</sub>), 4.32 (m, 2H, CH<sub>2</sub> Fmoc), 4.25 (m, 1H, CH Fmoc), 3.84 (d, 2H, J = 6.1Hz, CH<sub>2</sub> Glycine). δ<sub>C</sub> (500 MHz, DMSO) 169.2 (C=O), 168.4 (C=O), 166.9 (C=O), 156.4 (C=O), 143.8 (1-C Fmoc), 140.7 (2-C Fmoc), 139.4 (1''-C), 133.0 (7'-C), 132.1 (3''-C), 132.0 (3'-C), 130.1 (2-C), 129.6 (4''-C), 127.6 (4-C Fmoc), 127.0 (5-C Fmoc), 125.2 (6-C Fmoc), 124.9 (6-C), 124.7 (4-C), 124.5 (5''-C), 123.7 (5-C), 123.2 (6''-C), 121.3 (7-C) 120.4 (2''-C), 120.1 (4-C Fmoc), 107.4 (3-C), 65.7 (CH<sub>2</sub> Fmoc), 51.7 (CH<sub>3</sub>), 46.5 (CH Fmoc), 44.0 (CH<sub>2</sub> Gly). ν<sub>max</sub>/cm<sup>-1</sup> (ATR); 1686, 1534, 1433, 1417, 1318, 1277. *m/z* (ESI) (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 574.1966. C<sub>34</sub>H<sub>27</sub>N<sub>3</sub>O<sub>6</sub> requires MNa<sup>+</sup>, 574.1973).

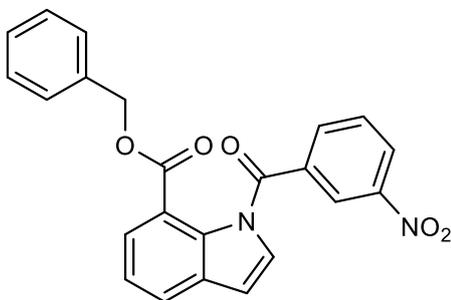
### Synthesis of benzyl indole-7-carboxylate (**21**)



To a stirred solution of indole-7-carboxylic acid (0.300 g, 1.86 mmol) in DCM (15 mL) was added; benzyl alcohol (0.39 mL, 3.72mmol), EDC.HCl (0.713g, 3.72 mmol) and DMAP (0.454 g, 3.72 mmol) at RT. The reaction mixture was then stirred for 12 hours at RT. The reaction mixture was then diluted with DCM (15

mL). The resultant solution was then washed with water (3 x 15 mL) and brine (15 mL). The organic layer was then dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude yellow oil. The crude product was then purified using column chromatography (2% MeOH- DCM) to give the title compound as a colourless oil (0.317g, 68%). R<sub>f</sub> 0.92 (2% MeOH-DCM). δ<sub>H</sub> (500 MHz, DMSO) 9.87 (bs, 1H, NH), 7.89 (app d, 1H, J = 7.60 Hz, 6-H), 7.83-7.81 (dd, 1H, J = 7.60 Hz, 4-H), 7.51 (app d, 2H, J = 7.36 Hz, 2''-H), 7.44-7.41 (m, 3H, 2-H and 3'''-H), 7.36-7.34 (m, 1H, 4'''-H), 7.13 (app t, 1H, J = 7.60 Hz, 5-H), 6.59-6.58 (m, 1H, 3-H), 5.47 (s, 2H, OCH<sub>2</sub>). δ<sub>C</sub> (500 MHz, DMSO) 166.4 (C=O), 137.1 (7'-C), 136.6 (1'''-C), 130.0 (3'-C), 129.0 (3'''-C), 128.5 (4'''-C), 128.4 (2'''-C), 127.6 (2-C), 126.8 (6-C), 124.2 (4-C), 119.0 (5-C), 112.8 (7-C), 102.2 (3-C), 66.1 (OCH<sub>2</sub>) ν<sub>max</sub>/cm<sup>-1</sup> (ATR); 3416, 1689, 1586, 1486, 1338, 1266 m/z (ESI) (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 252.1022. C<sub>16</sub>H<sub>13</sub>NO<sub>2</sub> requires MH<sup>+</sup>, 252.1019).

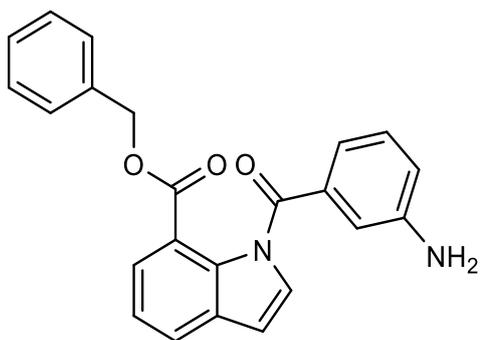
### Synthesis of benzyl 1-[(3-nitrophenyl)methyl]-1H-indole-7-carboxylate (**22**)



To a stirred solution of benzyl indole-7-carboxylate **21** (0.300g, 1.19 mmol) in THF (10 mL) was added NaH (0.071 g, 2.98 mmol) portion wise at 0 °C. The reaction mixture was then stirred at 0°C for 15 minutes. The reaction mixture was then warmed to RT and stirred at that temperature for 15 minutes. 3-nitrobenzoyl chloride (0.331 g, 1.79 mmol) was then added and the resultant solution was then stirred at RT for 18 hours. The reaction mixture was then quenched with water (10 mL) and the resultant solution was extracted with EtOAc (3 x 25

mL). The organic layers were then combined, washed with brine (20 mL), dried over MgSO<sub>4</sub>, filtered and the solvent removed to yield a crude yellow oil. The crude product was purified by column chromatography (20% EtOAc-Petroleum ether) to afford the title compound as a crystalline yellow solid (0.390 g, 81%). R<sub>f</sub> 0.69 (20% EtOAc-Petroleum ether). m.p 83.8-84.5°C. δ<sub>H</sub> (500 MHz, DMSO) 8.56 (ddd, 1H, J = 8.3, 2.2, 0.9 Hz, 6''-H), 8.25 (app s, 1H, 2''-H), 8.15 (dd, 1H, J = 8.3, 2.2 Hz, 4''-H) 7.94 (dd, 1H, J = 7.7, 1.0 Hz, 6-H), 7.86 (app t, 1H, J = 8.3 Hz, 5''-H), 7.71 (m, 1H, 4-H), 7.52 (app d, 1H, J = 3.7 Hz, 2-H), 7.46 (app t, 1H, J = 7.7 Hz, 5-H), 7.06-7.17 (m, 5H, 2'''-H, 3'''-H and 4'''-H), 6.86 (app d, 1H, J = 3.7 Hz, 3-H), 5.12 (s, 2H, OCH<sub>2</sub>). δ<sub>C</sub> (500 MHz, DMSO) 166.8 (C=O), 166.8 (C=O), 148.18 (1''-C) 136.4 (3''-C), 135.5 (7'-C), 134.3 (4''-C), 132.8 (1'''-C) 132.4 (3'-C), 131.3 (5''-C), 130.5 (2-C), 129.0 (2'''-C), 128.5 (3'''-C), 128.4 (4'''-C), 128.1 (6''-C), 125.8 (4-C), 125.6 (6-C), 124.7 (2''-C), 124.3 (5-C), 121.04 (7-C), 108.7 (3-C), 67.1 (OCH<sub>2</sub>). ν<sub>max</sub>/cm<sup>-1</sup> (ATR); 3113, 1710, 1597, 1422, 1315, 1276. m/z (ESI) (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 423.0984. C<sub>23</sub>H<sub>16</sub>N<sub>2</sub>O<sub>5</sub> requires MNa<sup>+</sup>, 423.0957).

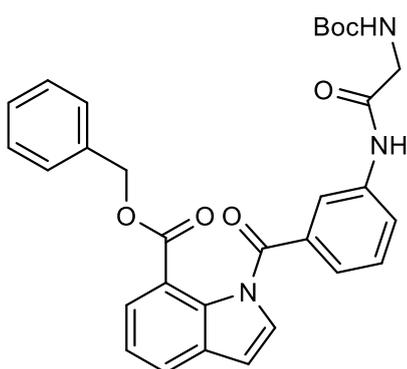
### Synthesis of benzyl 1-[(3-aminophenyl)methyl]-1H-indole-7-carboxylate (**23**)



To a solution of SnCl<sub>2</sub> (0.659 g, 3.48 mmol) in EtOH (23 mL) was added concentrated HCl (4.5 mL). The resultant solution was then heated to 65°C for 15 mins. To this solution was added benzyl 1-[(3-nitrophenyl)methyl]-1H-indole-7-carboxylate **22** (0.348 g, 0.87 mmol). The reaction mixture was then stirred at 65 °C for 45 mins and then cooled to rt. The solution was then neutralised using 4M NaOH. The resultant solution was extracted with EtOAc (3 x 15 mL). The organic layers were then combined, washed with brine (20 mL), dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a yellow residue. The residue was taken up in EtOAc (15 mL) and was washed with a further amount of brine (20 mL). The organic layer was again dried over MgSO<sub>4</sub>, filtered and the solvent removed to afford

the title compound as a yellow oil (0.189 g, 68%).  $R_f$  0.54 (50% EtOAc-Petrol)  $\delta_H$  (500 MHz, DMSO) 7.91 (dd, 1H,  $J = 7.8, 1.1$  Hz 6-H), 7.63 (dd, 1H,  $J = 7.6, 1.1$  Hz, 4-H), 7.51 (app d, 1H,  $J = 3.7$  Hz, 2-H), 7.41 (m, 1H, 5-H) 7.22-7.40 (m, 6H, 2'''-H, 3'''-H, 4'''-H and 5'''-H) 7.06 (app s, 1H, 2''-H) 6.91 (m, 2H, 4''-H and 6''-H) 6.82 (app d, 1H,  $J = 3.7$  Hz, 3-H), 5.53 (bs, 2H, NH<sub>2</sub>), 5.14 (s, 2H, OCH<sub>2</sub>).  $\delta_C$  (500 MHz, DMSO) 169.2 (C=O), 166.9 (C=O), 146.7 (1''-C), 135.8 (7'-C), 133.5 (3'''-C), 132.5 (1'''-C) 132.3 (3'-C) 130.7 (2-C), 129.84 (5''-C), 128.9 (2''-C), 128.7 (3'''-C), 128.5 (4'''-C), 125.3 (6-C), 125.2 (4-C), 123.6 (5-C), 121.1 (7-C), 119.1 (6''-C), 117.6 (4''-C), 115.1 (2''-C), 107.6 (3-C), 66.8 (OCH<sub>2</sub>).  $\nu_{max}/cm^{-1}$  (ATR); 3417, 3033, 1692, 1336, 1316, 1272.  $m/z$  (ESI) (100%, MH<sup>+</sup>); (Found MH<sup>+</sup>, 371.1399. C<sub>23</sub>H<sub>18</sub>N<sub>2</sub>O<sub>3</sub> requires MH<sup>+</sup>, 371.1390).

**Synthesis of benzyl 1-[3-(2-{tert-butoxy}carbonyl)amino}acetamido)benzoyl]-1H-indole-7-carboxylate (24)**

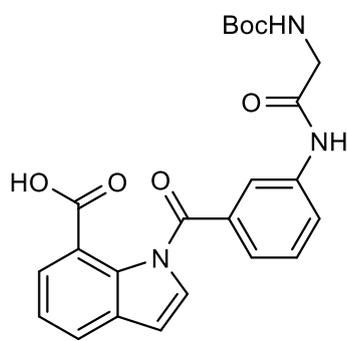


To a stirred solution of benzyl 1-[(3-aminophenyl)methyl]-1H-indole-7-carboxylate **23** (0.725 g, 1.95 mmol) in DCM (20 mL) was added; Boc-Gly-OH (0.410 g, 2.34 mmol), EDC.HCl (0.449 g, 2.34 mmol) and DMAP (0.030g, 0.2 mmol) at RT. The reaction mixture was then stirred at RT for 12 hours.

The reaction mixture was then diluted with DCM (20 mL). The resultant solution was then washed with water (2 x 20 mL) and brine (2 x 20mL). The organic layer was then dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude colourless solid. The crude solid was purified by column chromatography (2% MeOH-DCM) to give the title compound as yellow solid (1.03 g, 73%).  $R_f$  0.41 (2% MeOH-DCM).  $\delta_H$  (500 MHz, DMSO), 10.21 (s, 1H, amide NH), 8.10 (app s, 1H, 2''-H), 7.91-7.93 (m, 2H, 6-H and 4''-H), 7.65 (app d, 1H,  $J = 7.8$  Hz, 4-H), 7.54-7.52 (m, 1H, 5''-H), 7.49 (d, 1H,  $J = 3.8$  Hz, 2-H), 7.47-7.44 (m, 1H, 6''-H), 7.43 (app t, 1H,  $J = 7.8$  Hz, 5-H), 7.21-7.22 (m, 1H, 4'''-H), 7.17-7.18 (m, 4H, 2'''-H and 3'''-H), 7.10 (t, 1H,  $J = 6.1$  Hz, Gly-NH), 6.83 (d, 1H,  $J = 3.8$  Hz, 3-H), 5.11 (s, 2H, OCH<sub>2</sub>), 3.75 (d, 2H,  $J = 6.1$  Hz, Gly-CH<sub>2</sub>), 1.40 (s, 9H, Boc-CH<sub>3</sub>).  $\delta_C$  (500 MHz, DMSO) 169.2

(C=O), 168.5 (C=O), 166.9 (C=O), 156.4 (C=O), 139.9 (1''-C), 135.6 (7'-C), 133.3 (3''-C), 132.5 (1'''-C), 132.4 (3'-C), 130.6 (5''-C), 130.0 (2-C), 128.8 (2'''-C), 128.7 (3'''-C), 128.5 (4'''-C), 125.4 (6-C and 4''-C), 125.0 (4-C), 124.1 (6''-C), 123.9 (5-C), 121.0 (7-C), 120.6 (2''-C), 108.0 (3-C), 78.6 (C-Boc), 66.9 (OCH<sub>2</sub>), 44.3 (Gly-CH<sub>2</sub>), 28.7 (Boc-CH<sub>3</sub>)  $\nu_{\max}$  / cm<sup>-1</sup> (solid): 3354, 1709, 1687, 1590, 1533, 1271;  $m/z$  (ESI) (100%, MNa<sup>+</sup>); (Found MH<sup>+</sup>, 550.1952. C<sub>30</sub>H<sub>29</sub>N<sub>3</sub>O<sub>6</sub> requires MNa<sup>+</sup>, 550.1949).

**Synthesis of 1-[3-(2-[(tert-butoxy)carbonyl]amino)acetamido)benzoyl]-1H-indole-7-carboxylic acid (25)**

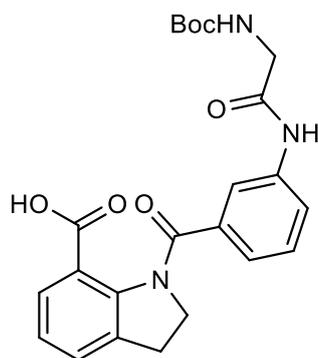


To a stirred solution of benzyl 1-[3-(2-{tert-butoxy)carbonyl] amino)acetamido)benzoyl]-1H-indole-7-carboxylate **24** (0.527 g, 1.00 mmol) in MeOH (20 mL) was added 10% Pd/C (0.150 g). The resultant slurry was then stirred at RT under an atmosphere of hydrogen (1 atmosphere) for 12 mins. The reaction mixture was then diluted with

MeOH (20 mL) and the resultant solution was filtered through Celite. The filtrate was collected and the solvent was removed to give a crude purple solid. The crude solid was then purified by reverse phase ACC (gradient 0-60% MeCN-H<sub>2</sub>O in 0.1 % formic acid) to give the title compound as a colourless crystalline solid (0.179 g, 41%).  $R_f$  0.24 (2% MeOH-DCM). m.p 187-191 °C.  $\delta_H$  (500 MHz, DMSO), 10.35 (s, 1H, amide NH), 8.18 (app s, 1H, 2''-H), 7.94 (app d, 1H, J = 8.0 Hz, 4''-H), 7.89 (dd, 1H, J = 7.8 and 0.9 Hz, 6-H), 7.64 (dd, 1H, J = 7.8 Hz and 0.9 Hz, 4-H) 7.56 (app t, 1H, J = 8.0 Hz, 5''-H), 7.50 (m, 2H, 6''-H and 2-H), 7.40 (app t, 1H, J = 7.8 Hz, 5-H), 7.09 (t, 1H, J = 6.1 Hz, Gly N-H), 6.81 (d, 1H, J = 3.63 Hz, 3-H), 3.74 (d, 2H, J = 6.1 Hz, Gly-CH<sub>2</sub>), 1.39 (s, 9H, Boc-CH<sub>3</sub>).  $\delta_C$  (500 MHz, DMSO) 169.2 (C=O), 168.4 (C=O), 168.2 (C=O), 156.4 (C=O), 139.4 (1''-C), 133.7 (7'-C), 132.9 (3''-C), 132.4 (3'-C), 130.8 (5''-C), 130.4 (2-C), 125.2 (4-C), 125.0 (6''-C), 124.9 (6-C), 124.1 (4''-C), 123.4 (5-C), 122.0 (7-C), 120.6 (2''-C), 107.5 (3-C), 78.6 (Boc-C), 44.3 (Gly-CH<sub>2</sub>), 28.6 (Boc-CH<sub>3</sub>).  $\nu_{\max}$  / cm<sup>-1</sup> (solid): 3344,

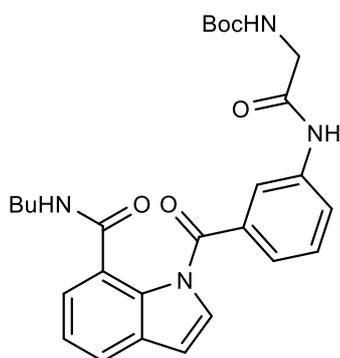
1683, 1645, 1542, 1477, 1368, 1215.  $m/z$  (ESI) (100%, M-H); (Found M-H, 436.1513.  $C_{23}H_{23}N_3O_6$  requires M-H, 436.1509).

**Synthesis of 1-[3-(2-[(tert-butoxy)carbonyl]amino)acetamido)benzoyl]-2,3-dihydro-1H-indole-7-carboxylic acid (28)**



To a stirred solution of benzyl 1-[3-(2-[(tert-butoxy)carbonyl] amino)acetamido)benzoyl]-1H-indole-7-carboxylate **24** (0.549 g, 1.00 mmol) in MeOH (20 mL) was added 10% Pd/C (0.150 g). The resultant slurry was then stirred at RT under an atmosphere of hydrogen (1 atmosphere) for 12 hours. A further amount of 10% Pd/C (0.050 g) was then added and the reaction mixture was stirred under an atmosphere of hydrogen (1 atmosphere) for a further 2 hours. The reaction mixture was then filtered through Celite, the filtrate was collected and the solvent removed to give a crude yellow solid. The crude solid was then purified by reverse phase ACC (gradient 0-30% MeCN-H<sub>2</sub>O in 0.1 % formic acid) to give the title compound as a colourless solid (0.404 g, 92%).  $R_f$  0.20 (2% MeOH-DCM).  $\delta_H$  (500 MHz, DMSO), 10.15 (s, 1H, amide N-H), 7.94 (app s, 1H, 2''-H), 7.78 (app d, 1H, J = 8.1 Hz, 4''-H), 7.47-7.42 (m, 3H, 6-H, 6''-H and 5''-H), 7.34 (app d, 1H, J = 7.5 Hz, 4-H), 7.15 (app t, 1H, J = 7.5 Hz, 5-H), 7.09 (t, 1H, J = 6.0 Hz, Gly Boc-NH), 4.08 (t, 2H, J = 7.8 Hz, 2-H), 3.74 (d, 2H, J = 6.0 Hz, Gly-CH<sub>2</sub>), 3.07 (t, 2H, J = 7.8 Hz, 3-H), 1.40 (s, 9H, Boc-CH<sub>3</sub>). Exchangeable proton for acid not observed.  $\delta_C$  (500 MHz, DMSO) 169.3 (C=O), 169.1 (C=O), 168.1 (C=O), 156.5 (C=O), 141.0 (7'-C), 139.5 (1''-C), 136.7 (4-C), 135.8 (3''-C), 129.4 (6-C), 127.9 (3'-C), 126.9 (5''-C) 127.5 (4''-C), 124.6 (5-C), 123.3 (6''-C), 122.0 (2''-C), 119.2 (7-C), 78.5 (Boc-C), 53.2 (2-C), 44.3 (Gly-CH<sub>2</sub>), 29.4 (3-C), 28.3 (Boc-CH<sub>3</sub>)  $\nu_{max}/cm^{-1}$  (ATR); 3304, 2975, 1688, 1608, 1368, 1160. (ESI) (100%, M-H); (Found M-H, 438.1668.  $C_{23}H_{25}N_3O_6$  requires M-H, 438.1665).

**Synthesis of tert-butyl N-[(3-[7-(butylcarbamoyl)-1H-indole-1-carbonyl]phenyl)carbamoyl]methyl]carbamate (33)**

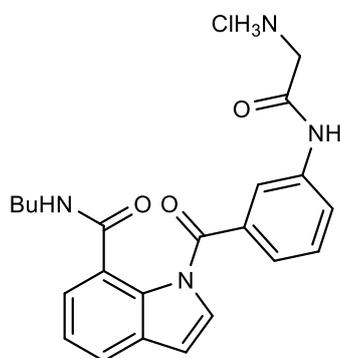


To a stirred solution of 1-[3-(2-[[tert-butoxy]carbonyl]amino)acetamido]benzoyl]-1H-indole-7-carboxylic acid **25** (0.725 g, 1.95 mmol) in DCM (20 mL) was added; butylamine (0.23 mL, 2.34 mmol), EDC.HCl (0.449 g, 2.34 mmol) and DMAP (0.030g, 0.2 mmol) at RT. The reaction mixture was then stirred at RT for 3

hours. The reaction mixture was then diluted with DCM (20 mL) and the resultant solution was then washed with water (2 x 20 mL) and brine (2 x 20mL). The organic layer was then dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude colourless solid. The crude solid was purified by column chromatography (2% MeOH-DCM) to afford the title compound as colourless solid (0.441 g, 46%). R<sub>f</sub> 0.37 (2% MeOH-DCM). δ<sub>H</sub> (500 MHz, DMSO), 10.22 (s, 1H, amide NH), 8.35 (t, 1H, J = 5.6 Hz, butyl amide NH), 8.11 (app s, 1H, 2''-H), 7.91 (app d, 1H, J = 8.0 Hz, 4''-H), 7.77 (dd, 1H, 7.6 Hz and 1.6 Hz, 6-H), 7.52 (app t, 1H, J = 8.0 Hz, 5''-H), 7.47-7.43 (m, 2H, 6''-H and 2-H), 7.39-7.37 (m, 1H, 4-H), 7.33 (app t, 1H, J = 7.6 Hz, 5-H), 7.09 (t, 1H, J = 6.1 Hz, Boc-NH), 6.77 (d, 1H, J = 3.6 Hz, 3-H), 3.74 (d, 2H, J = 6.1 Hz, Gly-CH<sub>2</sub>), 3.08 (q, 2H, J = 6.6 Hz, butyl amide CH<sub>2</sub>NH), 1.39 (m, 11H, Boc-CH<sub>3</sub> and butyl amide CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>), 1.28 (sxt, 2H, J = 7.4 Hz butyl amide CH<sub>3</sub>CH<sub>2</sub>) 0.83 (t, 3H, J = 7.4 Hz, butyl amide CH<sub>3</sub>) δ<sub>C</sub> (500 MHz, DMSO) 169.1 (C=O), 167.8 (C=O), 167.6 (C=O), 156.4 (C=O), 139.8 (1''-C), 134.0 (7''-C), 132.4 (3''-C), 132.1 (3'-C), 129.9 (5''-C), 129.7 (2-C), 125.9 (4-C), 125.1 (6''-C), 123.8 (6-C), 123.6 (4''-C), 123.1 (5-C), 120.7 (7-C), 120.1 (2''-C), 107.1 (3-C), 78.6 (Boc-C), 44.3 (Gly-CH<sub>2</sub>), 39.0 (Butyl amide CH<sub>2</sub>NH<sub>2</sub>), 31.6 (Butyl amide CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>), 28.6 (Boc-CH<sub>3</sub>), 20.2 (Butyl amide CH<sub>3</sub>CH<sub>2</sub>), 14.2 (Butyl amide CH<sub>3</sub>). ν<sub>max</sub> / cm<sup>-1</sup> (solid): 3271, 2931, 1688, 1529,

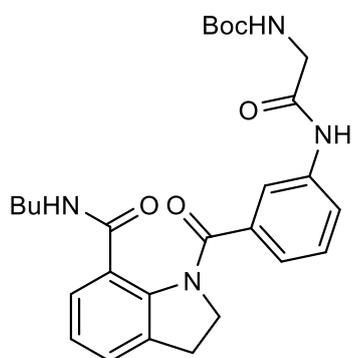
1280, 1189.  $m/z$  (ESI) (100%,  $MNa^+$ ); (Found  $MNa^+$ , 515.2270.  $C_{27}H_{32}N_4O_5$  requires  $MNa^+$ , 515.2270).

**Synthesis of 3-{3-[7-(butylcarbamoyl)-1H-indole-1-carbonyl]phenyl}-2-oxopropan-1-aminium chloride (34)**



To a RBF charged with tert-butyl *N*-[({3-[7-(butylcarbamoyl)-1H-indole-1-carbonyl]phenyl} carbamoyl)methyl]carbamate **33** (0.330 g, 0.67 mmol) was added 4M HCl in dioxane (5 mL) at RT. The reaction mixture was then stirred at RT for 2 hours. The resultant ppt was then collected by vacuum filtration and was washed with DCM (5 mL) to afford the title compound as a colourless solid (0.280 g, 98%).  $R_f$  0.10 (10% methanolic ammonia-DCM).  $m.p$   $>250$  °C  $\delta_H$  (500 MHz, DMSO), 11.13 (s, 1H, amide NH), 8.39 (t, 1H,  $J = 6.0$  Hz, butyl amide NH), 8.32 (bm, 3H,  $NH_3^+$ ), 8.16-8.15 (m, 1H, 2''-H), 7.93-7.92 (m, 1H, 4''-H), 7.78-7.76 (m, 1H, 6-H), 7.56 (app t, 1H,  $J = 7.9$  Hz, 5''-H), 7.52-7.50 (m, 1H, 6''-H), 7.46 (d, 1H,  $J = 3.6$  Hz, 2-H), 7.40-7.38 (m, 1H, 4-H), 7.33 (app t, 1H,  $J = 7.6$  Hz, 5-H), 6.78 (d, 1H,  $J = 3.6$  Hz, 3-H), 3.83 (q, 2H,  $J = 5.7$  Hz, Gly- $CH_2$ ), 3.08 (q, 2H,  $J = 6.0$  Hz, butyl amide  $CH_2NH$ ), 1.41 (p, 2H  $J = 6.0$  Hz, butyl amide  $CH_2CH_2NH_2$ ) 1.29 (sxt, 2H,  $J = 6.5$  Hz butyl amide  $CH_3CH_2$ ), 0.83 (t, 3H,  $J = 6.5$  Hz, butyl amide  $CH_3$ ).  $\delta_C$  (500 MHz, DMSO) 167.6 (C=O), 167.6 (C=O), 165.7 (C=O), 139.2 (1''-C), 134.2 (7'-C), 132.4 (3''-C), 132.1 (3'-C), 130.0 (5''-C), 129.9 (2-C), 125.9 (4-C), 125.8 (6''-C), 124.0 (6-C), 123.7 (4''-C), 123.2 (5-C), 123.1 (7-C), 120.8 (2''-C), 107.3 (3-C), 41.5 (Gly- $CH_2$ ), 39.0 (Butyl amide  $CH_2NH_2$ ), 31.6 (Butyl amide  $CH_2CH_2NH_2$ ), 20.1 (Butyl amide  $CH_3CH_2$ ), 14.2 (Butyl amide  $CH_3$ )  $\nu_{max} / cm^{-1}$  (solid): 2930, 1693, 1555, 1414, 1397, 1372  $m/z$  (ESI) (100%,  $MNa^+$ ); (Found  $MNa^+$ , 415.1736.  $C_{22}H_{24}N_4O_3$  requires  $MNa^+$ , 415.1746). HPLC (RT: 1.81 min(100% relative area)).

**Synthesis of tert-butyl N-[(3-[7-(butylcarbamoyl)-2,3-dihydro-1H-indole-1-carbonyl]phenyl)carbamoyl]methyl]carbamate**

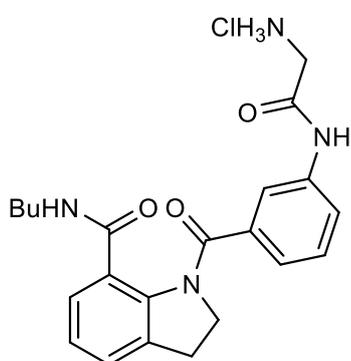


To a stirred solution of 1-[3-(2-[(tert-butoxy)carbonyl]amino)acetamido)benzoyl]-2,3-dihydro-1H-indole-7-carboxylic acid **28** (0.725 g, 1.95 mmol) in DMF (9 mL) was added; butylamine (0.23 mL, 2.34 mmol), EDC.HCl (0.449 g, 2.34 mmol) and DMAP (0.030g, 0.2 mmol) at RT. The reaction mixture was then stirred at RT for 12

hours. A further amount of *n*-BuNH<sub>2</sub> (0.12 mL, 1.17 mmol) and the reaction mixture was then stirred at 80 °C for 2 hours. The reaction mixture was then cooled to RT and then diluted with water (20 mL). The resultant solution was then extracted with EtOAc (3 x 20 mL). The organic layers were then combined and washed with water (2 x 20 mL) and brine (2 x 20mL). The organic layer was then dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude colourless solid. The crude solid was recrystallized from DCM to give the title compound as a colourless crystalline solid (0.539 g, 56%). m.p 160-165 °C. R<sub>f</sub> 0.30 (2% MeOH-DCM). δ<sub>H</sub> (500 MHz, DMSO), 10.13 (s, 1H, amide NH), 7.94-7.91 (m, 2H, 2''-H and butyl amide NH), 7.74 (app d, 1H, J = 7.9 Hz, 4''-H), 7.48-7.36 (m, 3-H, 6-H, 6''-H and 5''-H), 7.29 (app d, 1H, J = 7.5 Hz, 4-H), 7.18-7.06 (m, 2H, 5-H and Gly NH), 4.06 (t, 2H, J = 7.7 Hz, 2-H), 3.74 (d, 2H, J = 5.8 Hz, Gly CH<sub>2</sub>), 3.05-3.02 (m, 4H, 3-H and butyl amide CH<sub>2</sub>NH), 1.37-1.25 (m, 11H, Boc-CH<sub>3</sub> and butyl amide CH<sub>2</sub>CH<sub>2</sub>NH), 1.20 (sxt, 2H, J = 7.3 Hz, butyl amide CH<sub>3</sub>CH<sub>2</sub>CH<sub>2</sub>), 0.73 (t, 3H, J = 7.3 Hz, butyl amide CH<sub>3</sub>). δ<sub>C</sub> (500 MHz, DMSO) 169.0 (C=O). 168.8 (C=O), 167.0 (C=O), 156.4 (C=O), 140.1 (7'-C), 139.5 (1''-C), 136.9 (4-C), 135.5 (3''-C), 129.0 (6-C), 128.1 (3'-C), 126.7 (4''-C), 126.4 (5''-C), 124.5 (5-C), 123.4 (6''-C), 121.9 (2''-C), 119.5 (7-C), 78.5 (Boc C), 53.1 (2-C), 44.3 (Gly CH<sub>2</sub>), 39.0 (Butyl amide CH<sub>2</sub>NH), 31.7 (butyl amide CH<sub>2</sub>CH<sub>2</sub>NH), 29.5 (3-C), 28.7 (Boc CH<sub>3</sub>), 20.2 (butyl amide CH<sub>3</sub>CH<sub>2</sub>CH<sub>2</sub>) 14.3 (butyl amide CH<sub>3</sub>). ν<sub>max</sub>/cm<sup>-1</sup>

(ATR); 3296, 1634, 1543, 1440, 1380, 1208. *m/z.* (ESI) (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 517.2433. C<sub>27</sub>H<sub>34</sub>N<sub>4</sub>O<sub>5</sub> requires MNa<sup>+</sup>, 517.2421)

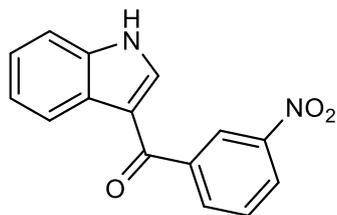
**Synthesis of ({3-[7-(butylcarbamoyl)-2,3-dihydro-1H-indole-1-carbonyl]phenyl}carbamoyl)methanaminium chloride (35)**



To a solution of tert-butyl *N*-[({3-[7-(butylcarbamoyl)-2,3-dihydro-1H-indole-1-carbonyl]phenyl}carbamoyl)methyl]carbamate (0.330 g, 0.67 mmol) in dioxane (5 mL) was added 4M HCl in dioxane (5 mL) at RT. The reaction mixture was then stirred at 50 °C for 2 hours. The reaction mixture was then cooled to

RT. The resultant ppt was then collected by vacuum filtration and was washed with DCM to afford the title compound as a colourless solid (0.285 g, 99%). *R<sub>f</sub>* 0.10 (10% methanolic ammonia-DCM). *m.p* >250 °C  $\delta_{\text{H}}$  (500 MHz, DMSO), 11.00 (s, 1H, amide NH), 8.31 (bs, 3H, NH<sub>3</sub><sup>+</sup>), 7.95 (m, 2H, 2''-H and butyl amide NH), 7.78 (app d, 1H, *J* = 7.7 Hz, 4''-H), 7.41-7.30 (m, 3H, 6-H, 6''-H and 5''-H), 7.21 (app d, 1H, *J* = 7.5 Hz, 4-H), 7.13 (app t, 1H, 7.5 Hz, 5-H), 4.07-4.05 (m, 2H, 2-H), 3.83-3.81 (m, 2H, Gly CH<sub>2</sub>), 3.04-3.02 (m, 4H, 3-H and butyl amide CH<sub>2</sub>NH), 1.32-1.28 (m, 2H, butyl amide CH<sub>2</sub>CH<sub>2</sub>NH), 1.23-1.17 (m, 2H, butyl amide CH<sub>3</sub>CH<sub>2</sub>CH<sub>2</sub>). 0.72-0.70 (m, 3H, butyl amide CH<sub>3</sub>).  $\delta_{\text{C}}$  (500 MHz, DMSO) 168.6 (C=O), 167.2 (C=O), 165.4 (C=O), 140.1 (7'-C), 138.8 (1''-C), 137.0 (4-C), 135.5 (3''-C), 129.3 (6-C), 128.1 (3'-C), 126.7 (4''-C), 126.4 (5''-C), 124.6 (5-C), 123.7 (6''-C), 122.0 (2''-C), 119.5 (7-C), 53.0 (2-C), 46.6 (Gly CH<sub>2</sub>), 39.0 (butyl amide CH<sub>2</sub>NH), 31.7 (butyl amide CH<sub>2</sub>CH<sub>2</sub>NH), 29.1 (3-C), 20.2 (butyl amide CH<sub>3</sub>CH<sub>2</sub>CH<sub>2</sub>), 14.1 (butyl amide CH<sub>3</sub>)  $\nu_{\text{max}}/\text{cm}^{-1}$  (ATR); 2957, 1695, 1632, 1585, 1556, 1487, 1433. *m/z.* (ESI) (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 417.1893. C<sub>22</sub>H<sub>26</sub>N<sub>4</sub>O<sub>3</sub> requires MNa<sup>+</sup>, 417.1897). HPLC (RT: 1.64 min(100% relative area)).

### Synthesis of 3-[(3-nitrophenyl)carbonyl]-1H-Indole (42)



To a solution of Indole (0.500 g, 4.34 mmol) in DCM (19 mL) was added  $\text{Et}_2\text{AlCl}$  (1M in hexane, 6.6 mL, 6.60 mmol) dropwise at  $-78\text{ }^\circ\text{C}$ . The reaction mixture was then stirred at  $-78\text{ }^\circ\text{C}$  for 15 mins. To this solution was added 3-nitrobenzoyl chloride (1.20 g, 6.51 mmol) in DCM (15 mL). The reaction mixture was then warmed to  $0\text{ }^\circ\text{C}$  and stirred at that temperature for three hours. The reaction mixture was then quenched with pH 7 buffer solution (20 mL). The resultant ppt was filtered and washed with water (10 mL). The crude product was purified using column chromatography (2% MeOH/DCM) to give the title compound as a yellow crystalline solid (0.335g, 29%), m.p.  $217\text{-}218\text{ }^\circ\text{C}$ ,  $\delta_{\text{H}}$  (500 MHz, DMSO); 12.24 (s, 1H, NH), 8.50 (app s, 1H, 2''-H), 8.45 (m, 1H, 6''-H), 8.28 (app d, 1H,  $J = 7.1\text{ Hz}$ , 4-H), 8.23 (dd, 1H,  $J = 7.6\text{ Hz}$ ,  $0.9\text{ Hz}$ , 4''-H), 8.10 (s, 1H, 2-H) 7.84 (app t, 1H,  $J = 7.6\text{ Hz}$ , 5''-H), 7.56 (dd, 1H,  $J = 6.7$ ,  $1.6\text{ Hz}$ , 7-H), 7.30 (m, 2H, 5-H and 6-H),  $\delta_{\text{C}}$  (500 MHz, DMSO); 187.4 (C=O), 147.7 (1''-C), 141.6 (7'-C), 136.8 (2-C), 136.7 (3''-H) 134.6 (4-C), 133.2 (4''-H) 130.3 (5''-C), 126.1 (3'-C), 125.5 (6''-C), 123.5 (6-C), 123.0 (5-C), 121.4 (2''-C), 114.6 (3-C), 112.5 (7-C);  $\nu_{\text{max}}/\text{cm}^{-1}$  (ATR); 3114, 1596, 1530, 1514, 1424;  $m/z$  (ESI) (100%,  $\text{MH}^+$ ); (Found  $\text{MH}^+$ , 267.0764.  $\text{C}_{15}\text{H}_{10}\text{N}_2\text{O}_3$  requires  $\text{MH}^+$ , 267. 0764).

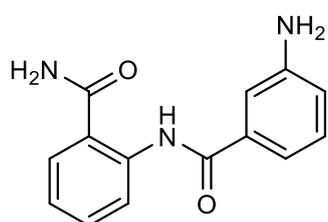
### Synthesis of 2-(3-nitrobenzamido)benzamide (49)



To a stirred solution of methyl 2-aminobenzamide (0.450g, 3.31 mmol) in THF (20 mL) was added  $\text{Et}_3\text{N}$  (0.7 mL, 4.97 mmol) and 3-nitrobenzoyl chloride (0.614 g, 3.31 mmol) at RT. The reaction mixture was then stirred at RT for 12 hours and the resultant ppt was collected by vacuum filtration and the solid was washed with water (10 mL) and DCM (10 mL). The crude solid was then recrystallized from EtOH to afford the title compound as a colourless solid (0.774 g, 82%). m.p  $178\text{-}180\text{ }^\circ\text{C}$ .  $R_f$  0.46 (10% MeOH-DCM).  $\delta_{\text{H}}$  (500 MHz, DMSO) 13.31 (s, 1H, amide NH), 8.73-8.72 (m, 1H, 2'-H), 8.69-8.67 (m, 1H, 3-H), 8.51 (bs, 1H, amide NH), 8.49-8.46 (m, 1H, 6'-H), 8.36-8.35 (m, 1H, 3'H), 7.98 (bs,

1H, amide NH), 7.97-7.95 (dd, 1H, J = 1.4, 7.9 Hz, 6-H), 7.90 (app t, 1H, J = 8.0 Hz, 5'-H), 7.63-7.59 (m, 1H, 4-H), 7.25-7.21 (m, 1H, 5-H)  $\delta_c$  (500 MHz, DMSO) 171.6 (C=O), 162.7 (C=O), 148.6 (1'-C), 140.2 (1-C), 136.5 (3'-C), 133.6 (4'-C), 133.2 (4-C), 131.2 (5'-C), 129.3 (6-C), 127.0 (6'-C), 123.7 (5-C), 122.2 (2'-C), 120.7 (3-C), 119.9 (2-C)  $\nu_{max}$  /  $cm^{-1}$  (solid): 3414, 3152, 1678, 1511, 1312.  $m/z$  (ESI); (100%,  $MH^+$ ); (Found  $MH^+$ , 308.0647.  $C_{14}H_{11}N_3O_4$  requires  $MNa^+$ , 308.0647).

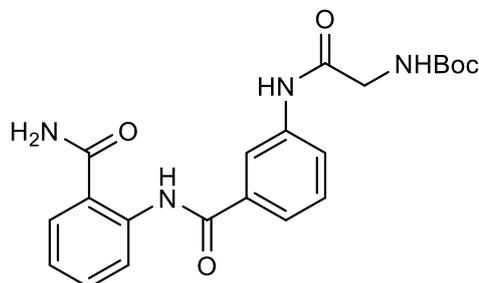
### Synthesis of 2-(3-aminobenzamido)benzamide (50)



To a solution of 2-(3-nitrobenzamido)benzamide **49** (0.285 g, 1 mmol) in MeOH (20 mL) was added 10% Pd/C (0.150 g). The resultant slurry was then stirred at RT under an atmosphere of hydrogen (1 atmosphere) for 12 hours. The reaction mixture was

then filtered through Celite and the solvent removed to give a crude purple solid. The crude solid was then recrystallized from ethanol to afford the title compound as a colourless solid (0.228 g, 89%). m.p 194-195 °C.  $R_f$  0.21 (10% MeOH-DCM).  $\delta_H$  (500 MHz, DMSO), 12.78 (s, 1H, amide NH), 8.70 (app d, 1H, J = 8.3 Hz, 3-H), 8.41 (bs, 1H, primary amide NH), 7.89 (dd, 1H, 8.3 and 1.3 Hz, 6-H), 7.80 (bs, 1H, primary amide NH), 7.57-7.54 (m, 1H, 4-H), 7.20-7.14 (m, 3H, 5-H, 5'-H and 2'-H), 7.05 (app d, 1H, J = 7.6 Hz, 4'-H), 6.79-6.77 (m, 1H, 6'-H), 5.41 (bs, 2H, aniline-NH<sub>2</sub>).  $\delta_c$  (500 MHz, DMSO) 171.6 (C=O), 165.7 (C=O), 149.8 (1'-C), 140.7 (1-C), 136.0 (3'-C), 132.9 (4-C), 129.7 (5'-C), 129.2 (6-C), 122.8 (5-C), 120.5 (3-C), 119.6 (2-C), 117.7 (6'-C), 114.0 (4'-C), 113.0 (2'-C)  $\nu_{max}/cm^{-1}$  (ATR); 3215, 2601, 1581, 1521, 1395, 1313.  $m/z$  (ESI); (100%,  $MH^+$ ); (Found  $MH^+$ , 256.1077.  $C_{14}H_{13}N_3O_2$  requires  $MH^+$ , 256.1080)  $R_f$  0.21 (10% MeOH-DCM)

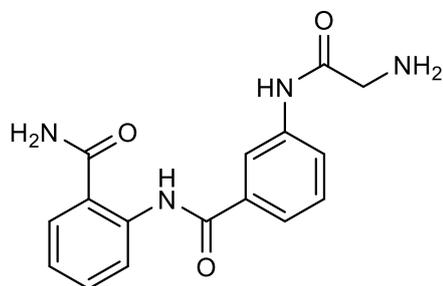
**Synthesis of tert-butyl N-[(3-[(2-carbamoylphenyl)carbamoyl]phenyl)carbamoyl]methyl]carbonate (51)**



To a stirred solution of 2-(3-aminobenzamido)benzamide **50** (0.214 g, 0.84 mmol) in DMF (10 mL) was added; Boc-Gly-OH (0.178 g, 1.01 mmol), EDC.HCl (0.194 g, 1.01 mmol) and DMAP (0.010g, 0.08 mmol) at RT.

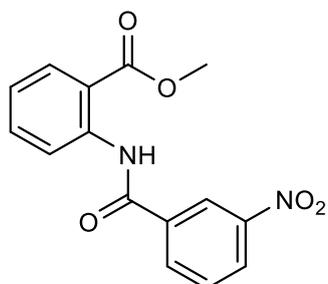
The reaction mixture was then stirred at 100 °C for 12 hours. A further amount of EDC.HCl (0.097 g, 0.05 mmol) and Boc-Gly-OH (0.089 g, 0.05 mmol) was added and the reaction mixture was stirred at 100°C for a further 3 hours. The reaction mixture was then cooled to RT and then diluted with DCM (20 mL). The resultant solution was then washed with water (2 x 20 mL) and brine (2 x 20mL). The organic layer was then dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude colourless solid. The crude solid was then recrystallized from ethanol to afford the title compound as a colourless solid (0.156 g, 45%). m.p 198-206 °C. R<sub>f</sub> 0.15 (10% MeOH-DCM). δ<sub>H</sub> (500 MHz, DMSO), 12.93 (s, 1H, amide NH), 10.20 (s, 1H, amide NH), 8.70 (app d, 1H, J = 8.3 Hz, 3-H), 8.42 (bs, 1H, primary amide N-H), 8.20 (app s, 1H, 2'-H, 7.91-7.85 (m, 3H, 6-H, 4'-H and primary amide N-H), 7.60-7.59 (m, 2H, 4-H and 6'-H), 7.51 (app t, 1H, J = 7.86, 5'-H), 7.19 (app t, 1H, J = 8.1 Hz, 5-H), 7.08 (t, 1H, J = 5.83, Boc-NH), 3.76 (d, 2H, J = 5.83, Gly-CH<sub>2</sub>) 1.45 (s, 9H, Boc-CH<sub>3</sub>) δ<sub>C</sub> (500 MHz, DMSO) 171.9 (C=O), 169.0 (C=O), 164.8 (C=O), 156.4 (C=O), 140.5 (1'-C), 140.1 (1-C), 135.8 (3'-C), 133.1 (4-C), 129.8 (5'-C), 129.2 (6-C), 123.2 (5-H), 122.9 (4'-C), 121.5 (6'-C), 120.6 (3-C), 119.6 (2-C), 118.6 (2'-C), 78.6 (Boc-C), 46.2 (Gly-CH<sub>2</sub>), 28.5 (Boc-CH<sub>3</sub>). ν<sub>max</sub>/cm<sup>-1</sup> (ATR); 3195, 1662, 1543, 1520, 1297 m/z (EI): (Found M+Na, 435.1642 C<sub>21</sub>H<sub>24</sub>N<sub>4</sub>O<sub>5</sub> requires M+Na, 435.1644)

**Synthesis of  $\{3-[(2\text{-carbamoylphenyl})\text{carbamoyl}]phenyl\}\text{carbamoyl}\}$ methanaminium chloride (47)**



To a stirred solution of tert-butyl  $N-[\{3-[(2\text{-carbamoylphenyl})\text{carbamoyl}]phenyl\}\text{carbamoyl}\}$ methylcarbonate **51** (0.200 g, 0.84 mmol) in DCM (2 mL) was added 4M HCl in dioxane (2 mL) at RT. The resultant reaction mixture was then heated at reflux for 1 hour. The reaction mixture was then cooled to RT and the resultant ppt was collected by vacuum filtration and washed with DCM (10mL) to afford the title compound as colourless crystalline solid (0.280 g, 96%).  $R_f$  0.16 (10% methanolic ammonia-DCM). m.p >250 °C.  $\delta_H$  (500 MHz, DMSO), 12.97 (s, 1H, amide NH), 11.02 (s, 1H, amide NH), 8.70 (dd, 1H,  $J = 8.4$  Hz and 0.97 Hz, 3-H), 8.47 (bs, 1H, primary amide N-H), 8.32 (bs, 3H,  $NH_3^+$ ) 8.25 (app s, 1H, 2'-H), 7.93 (dd, 1H,  $J = 7.86$  Hz and 1.35 Hz, 6-H), 7.89-7.87 (m, 1H, 4'-H), 7.83 (bs, 1H, primary amide N-H), 7.65-7.64 (m, 1H, 6'H) 7.60-7.55 (m, 2H, 4-H and 5'-H), 7.12-7.18 (m, 1H, 5-H), 3.84 (d, 2H,  $J = 5.48$ , Gly-CH<sub>2</sub>)  $\delta_C$  (500 MHz, DMSO) 171.5 (C=O), 165.6 (C=O), 164.6 (C=O), 140.5 (1'-C), 139.4 (1-C), 136.0 (3'-C), 133.0 (4-C), 130.0 (5'-C), 129.3 (6-C), 123.2 (5-H), 123.0 (4'-C), 122.2 (6'-C), 120.6 (3-C), 119.6 (2-C), 118.7 (2'-C), 66.8 (Gly CH<sub>2</sub>).  $\nu_{max}/cm^{-1}$  (ATR); 3190, 2892, 1692, 1590, 1524, 1474.  $m/z$  (EI): (Found  $M+Na$ , 335.1115  $C_{16}H_{16}N_4O_3$  requires  $M+Na$ , 335.1115) HPLC (RT: 1.45 min(98 % relative area)).

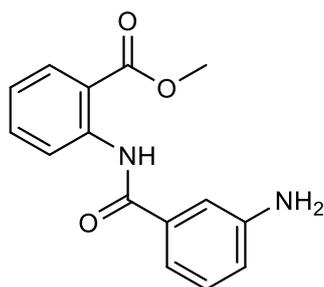
**Synthesis of methyl 2-(3-nitrobenzamido)benzoate (54)**



To a stirred solution of methyl 2-aminobenzoate (0.500g, 3.31 mmol) in THF (20 mL) was added Et<sub>3</sub>N (0.7 mL, 4.97 mmol) and 3-nitrobenzoyl chloride (0.614 g, 3.31 mmol) at RT. The reaction mixture was then stirred at RT for 12 hours. The reaction mixture was then diluted with EtOAc (25 mL). The resultant solution was then washed with water (2 x 25 mL) and

brine (25 mL). The organic layer was then dried over  $\text{MgSO}_4$ , filtered and the solvent removed to give a crude yellow solid. The crude solid was recrystallized from EtOAc to afford the title compound as a colourless solid (0.854 g, 86%).  $R_f$  0.85 (20% EtOAc-Petrol). m.p 85-87 °C  $\delta_H$  (500 MHz, DMSO), 11.60 (s, 1H, amide NH), 8.80-8.79 (m, 1H, 2'-H) 8.56-8.53 (ddd, 1H,  $J = 8.1, 2.3$  and 1.0 Hz, 6'-H), 8.45-8.40 (m, 2H, 4'-H and 6-H), 8.06-8.04 (dd, 1H,  $J = 1.6, 7.9$  Hz, 3-H), 7.99 (app t, 1H,  $J = 8.1$  Hz, 5'-H), 7.78-7.76 (m, 1H, 5-H), 7.38-7.35 (m, 1H, 4-H), 3.94 (s, 3H,  $\text{OCH}_3$ ).  $\delta_C$  (500 MHz, DMSO) 168.2 (C=O), 163.2 (C=O), 148.5 (1'-C), 139.5 (1-C), 136.3 (3'-C), 134.5 (5-C), 133.8 (4'-C), 131.2 (3-C), 131.1 (5'-C), 127.1 (6'-C), 124.7 (4-C), 122.6 (2'-C), 122.5 (6-C), 119.6 (2-C), 53.1 ( $\text{OCH}_3$ ).  $\nu_{\text{max}} / \text{cm}^{-1}$  (solid): 3248, 1693, 1671, 1526, 1304.  $m/z$ . (ESI) (100%,  $\text{MH}^+$ ); (Found  $\text{MH}^+$ , 301.1168.  $\text{C}_{15}\text{H}_{12}\text{N}_2\text{O}_5$  requires  $\text{MH}^+$ , 301.1160).

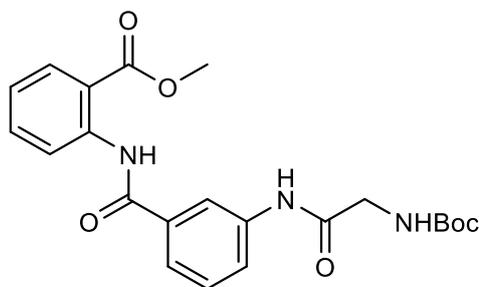
### Synthesis of methyl 2-(3-aminobenzamido)benzoate (55)



To a stirred solution of methyl 2-(3-nitrobenzamido)benzoate **54** (0.330 g, 1.10 mmol) in MeOH (20 mL) was added zinc powder (0.719 g, 11.00 mmol) at 0 °C. Ammonium chloride (11.00 mmol) was then added portionwise over a period of 5 minutes. The reaction mixture was then warmed to RT and stirred at that temperature for 1 hour. The reaction mixture was then filtered through Celite. The filtrate was then collected and the solvent removed to give a crude residue. The crude residue was then taken up in EtOAc (30 mL). The resultant solution was washed with water (20 mL) and brine (20 mL). The organic layer was dried over  $\text{MgSO}_4$ , filtered and the solvent removed to afford the title compound as a yellow solid (0.226 g, 76%).  $R_f$  0.64 (20% EtOAc-Petrol).  $\delta_H$  (500 MHz, DMSO), 11.60 (s, 1H, amide NH), 8.68-8.67 (dd, 1H,  $J = 8.54, 1.02$  Hz, 3-H), 8.09-8.07 (m, 1H, 6-H), 7.75-7.71 (m, 1H, 4-H), 7.29-7.24 (m, 2H, 5-H and 5'-H), 7.22-7.21 (m, 1H, 2'-H), 7.13-7.12 (m, 1H, 4'-H), 6.87-6.83 (m, 1H, 6'-H), 5.50 (bs, 2H,  $\text{NH}_2$ ), 3.95 (s, 3H,  $\text{OCH}_3$ ).  $\delta_C$  (500 MHz, DMSO) 168.6 (C=O), 166.00 (C=O), 149.8 (1'-C), 141.2 (1-C), 135.7 (3'-C), 134.9 (5-C), 131.2 (3-C), 129.8 (5'-C), 123.5 (4-C), 120.9 (6-C), 117.9 (6'-C), 116.7 (2-C), 114.0 (4'-C), 113.0

(2'-C), 53.1 (OCH<sub>3</sub>)  $\nu_{\max}$  / cm<sup>-1</sup> (solid): 3470, 1667, 1631, 1601, 1301.  $m/z$  (ESI) (100%, MNa<sup>+</sup>); (Found MNa<sup>+</sup>, 293.1221. C<sub>15</sub>H<sub>14</sub>N<sub>2</sub>O<sub>3</sub> requires MNa<sup>+</sup>, 293.1220).

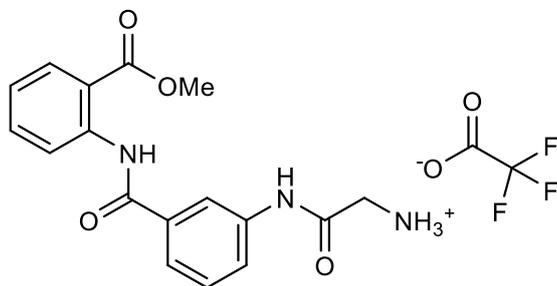
**Synthesis of methyl 2-[3{[(tert-butoxy)carbonyl]amino}acetamido]benzamido]benzoate (56)**



To a stirred solution of methyl 2-(3-aminobenzamido)benzoate **55** (0.283 g, 1.05 mmol) in DCM (20 mL) was added; Boc-Gly-OH (0.193 g, 1.01 mmol), EDC.HCl (0.211 g, 1.10 mmol) and DMAP (0.012 g, 0.10 mmol) at RT. The

reaction mixture was then stirred at RT for 12 hours. The reaction mixture was then diluted with DCM (20 mL). The resultant solution was then washed with water (2 x 20 mL) and brine (2 x 20 mL). The organic layer was then dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude colourless solid. The crude solid was then recrystallized from ethanol to afford the title compound as a colourless solid (0.331 g, 74%).  $R_f$  0.51 (2% MeOH-DCM). m.p 133-138 °C.  $\delta_H$  (500 MHz, DMSO), 11.57 (s, 1H, amide NH), 10.21 (s, 1H, amide NH), 8.56 (app d, 1H,  $J = 7.7$  Hz, 3-H), 8.22 (app s, 1H, 2'-H), 8.01 (app d, 1H,  $J = 7.7$  Hz, 6-H), 7.86 (app d, 1H,  $J = 7.4$  Hz, 4'-H), 7.70 (app t, 1H,  $J = 7.7$  Hz, 4-H), 7.63 (app d, 1H,  $J = 7.4$  Hz, 6'-H), 7.55 (app t, 1H,  $J = 7.4$  Hz, 5'-H), 7.26 (app t, 1H,  $J = 7.7$  Hz, 4-H), 7.10 (m, 1H, Boc-NH), 3.90 (s, 3H, OCH<sub>3</sub>), 3.76 (d, 2H,  $J = 5.5$  Hz, Gly-CH<sub>2</sub>), 1.41 (s, 9H, Boc-CH<sub>3</sub>).  $\delta_C$  (500 MHz, DMSO) 169.1 (C=O), 168.5 (C=O), 165.2 (C=O), 140.6 (1-C), 140.1 (1'-C), 135.5 (4-C), 134.8 (3'-C), 131.2 (6-C), 129.9 (5'-C), 123.9 (5-C), 123.0 (4'-C), 121.6 (6'-C), 121.4 (3-C), 118.7 (2'-C), 117.6 (2-C), 78.6 (Boc-C), 53.1 (O-CH<sub>3</sub>), 44.3 (Gly-CH<sub>2</sub>), 28.6 (Boc-CH<sub>3</sub>);  $\nu_{\max}$  / cm<sup>-1</sup> (solid): 3423, 1692, 1671, 1556, 1274;  $m/z$  (EI): (Found MNa<sup>+</sup>, 450.2135 C<sub>22</sub>H<sub>25</sub>N<sub>3</sub>O<sub>6</sub> requires MNa<sup>+</sup>, 450.2128).

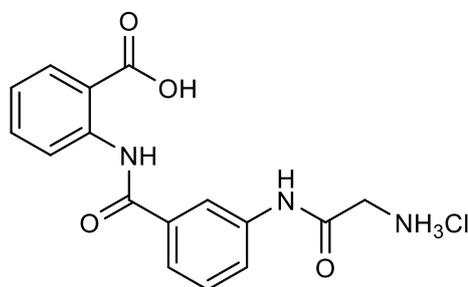
**Synthesis of [(3-{[2-(methoxycarbonyl)phenyl]carbamoyl}phenyl)carbamoyl]methanaminium trifluoroacetate (57)**



To a stirred solution of methyl 2-[3{[(tert-butoxy)carbonyl]amino}acetamido]benzoate **56** (0.214 g, 0.50 mmol) in DCM (2 mL) was added TFA (2 mL) at RT.

The resultant reaction mixture was then stirred at RT for 1 hour. The solvent was then removed to give a crude colourless solid. The crude solid was then triturated with DCM (5 mL) to afford the title compound as a colourless amorphous solid (0.211 g, 96%).  $R_f$  0.47 (10% Methanolic ammonia-DCM). m.p >250 °C.  $\delta_H$  (500 MHz, DMSO) 11.59 (s, 1H, amide NH), 10.72 (s, 1H, amide NH) 8.56-8.54 (dd, 1H,  $J = 8.14, 1.14$  Hz, 3-H), 8.21-8.20 (m, 1H, 2'-H), 8.17 (bs, 3H,  $NH_3^+$ ), 8.02-8.01 (dd, 1H,  $J = 8.14, 1.14$  Hz, 6-H), 7.85-7.83 (dd, 1H,  $J = 7.99, 1.14$  Hz, 4'-H), 7.71-7.67 (m, 2H, 4-H and 6'-H), 7.59 (app t, 1H,  $J = 7.99$  Hz, 5'-H), 7.27-7.24 (m, 1H, 5-H), 3.88 (s, 3H,  $OCH_3$ ), 3.84 (d, 2H,  $J = 5.33$ , Gly- $CH_2$ )  $\delta_C$  (500 MHz, DMSO) 168.5 (C=O), 165.7 (C=O), 165.0 (C=O), 159.2-158.3 (q, TFA C=O), 140.6 (1-C), 139.3 (1'-C), 135.8 (4-C), 134.9 (3'-C), 131.2 (6-C), 130.2 (5'-C), 124.0 (5-C), 123.1 (4'-C), 122.3 (6'-C), 121.4 (3-C), 120.1-113.1 (q, TFA  $CF_3$ ), 118.7 (2'-C), 117.8 (2-C), 53.2 ( $OCH_3$ ), 41.6 (Gly  $CH_2$ );  $\nu_{max}$  /  $cm^{-1}$  (solid): 3128, 1781, 1680, 1586, 1173;  $m/z$  (EI): (Found  $MNa^+$ , 350.1485  $C_{17}H_{17}N_3O_4$  requires  $MNa^+$ , 350.1479). HPLC (RT: 1.77 min(100 % relative area)).

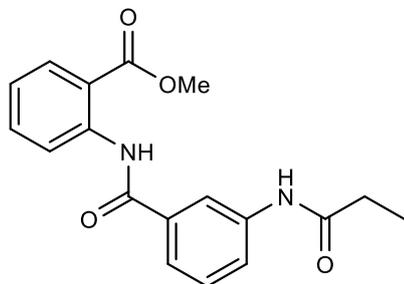
**Synthesis of ((3-[(2-carboxyphenyl)carbamoyl]phenyl)carbamoyl)methanaminium chloride (52)**



To a stirred solution of [(3-[(2-(methoxycarbonyl)phenyl)carbamoyl]phenyl)carbamoyl]methanaminium trifluoroacetate **57** (0.200 g, 0.45 mmol) in MeOH (2 mL) and H<sub>2</sub>O (2 mL) was added LiOH·H<sub>2</sub>O (0.084 g, 2.00 mmol) at

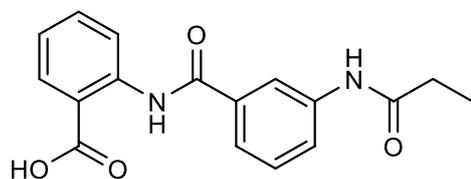
RT. The resultant reaction mixture was then stirred at RT for 12 hours. The reaction mixture was then concentrated and acidified with 2M HCl. The resultant precipitate was then collected by vacuum filtration and washed with H<sub>2</sub>O (5 mL) and DCM (5 mL) to afford the title compound as a colourless solid (0.157 g, 95%). *R<sub>f</sub>* 0.05 (10% methanolic ammonia-DCM). *m.p.* >250 °C.  $\delta_{\text{H}}$  (500 MHz, DMSO) 12.19 (s, 1H, amide NH), 11.12 (s, 1H, amide NH), 8.71 (app d, 1H, *J* = 8.3 Hz, 3-H), 8.37 (bs, 3H, primary amine 3 x N-H), 8.31-8.30 (m, 1H, 2'-H), 8.07 (dd, 1H, *J* = 8.0 and 1.7 Hz, 6-H), 7.88-7.87 (m, 1H, 4'-H), 7.70-7.66 (m, 2H, 4-H and 6'-H), 7.58 (app t, 1H, *J* = 7.9 Hz, 5'-H), 7.24-7.21 (m, 1H, 5-H), 3.85-3.84 (m, 2H, Gly-CH<sub>2</sub>).  $\delta_{\text{C}}$  (500 MHz, DMSO) 170.1 (C=O), 165.6 (C=O), 164.4 (C=O), 141.5 (1-C), 139.5 (1'-C), 135.0 (4-C), 134.9 (3'-C), 131.8 (6-C), 130.1 (5'-C), 123.5 (5-C), 123.1 (4'-C), 122.2 (6'-C), 120.4 (3-C), 118.7 (2'-C), 117.0 (2-C), 41.5 (Gly-CH<sub>2</sub>).  $\nu_{\text{max}}$  / cm<sup>-1</sup> (solid): 2958, 1677, 1586, 1528, 1477, 1445 *m/z* (EI): (Found M-H, 312.0987 C<sub>16</sub>H<sub>15</sub>N<sub>3</sub>O<sub>4</sub> requires M-H, 312.0990). HPLC (RT: 1.72 min(98 % relative area)).

### Synthesis of methyl 2-(3-propanamidobenzamido)benzoate



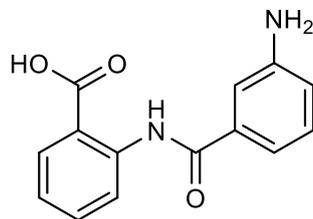
To a stirred solution of methyl 2-(3-aminobenzamido)benzoate **55** (0.230 g, 0.84 mmol) in DCM (20 mL) was added; propionic acid (0.076 mL, 1.01 mmol), EDC.HCl (0.194 g, 1.01 mmol) and DMAP (0.010g, 0.08 mmol) at RT. The reaction mixture was then stirred at RT for 12 hours. The reaction mixture was then diluted with DCM (20 mL). The resultant solution was then washed with water (2 x 20 mL) and brine (2 x 20mL). The organic layer was then dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude colourless solid. The crude solid was then recrystallized from DCM to afford the title compound as a colourless amorphous solid (0.189 g, 69%). R<sub>f</sub> 0.66 (2% MeOH-DCM). δ<sub>H</sub> (500 MHz, DMSO) 11.57 (s, 1H, amide NH), 10.14 (s, 1H, amide NH), 8.56 (app d, 1H, J = 8.0 Hz, 3-H), 8.22 (s, 1H, 2'-H), 8.02 (dd, 1H, J = 8.0, 1.5 Hz, 6-H), 7.89-7.88 (m, 1H, 4'H), 7.71-7.67 (m, 1H, 4-H), 7.61-7.59 (m, 1H, 6'-H), 7.52 (app t, 1H, J = 7.90 Hz, 5'-H), 7.27-7.24 (m, 1H, 5-H), 3.90 (s, 3H, OCH<sub>3</sub>), 2.37 (q, 2H, J = 7.5 Hz, propyl-CH<sub>2</sub>), 1.11 (t, 3H, J = 7.5 Hz, propyl-CH<sub>3</sub>). δ<sub>C</sub> (500 MHz, DMSO) 172.8 (C=O), 168.5 (C=O), 165.2 (C=O), 140.7 (C-1), 140.5 (1'-C), 135.5 (4-C), 134.8 (3'-C), 131.2 (6-C), 129.8 (5'-C), 123.9 (5-C), 122.9 (4'-C), 121.3 (6'-C), 121.2 (3-C), 118.6 (2'-C), 117.5 (2-C), 53.1 (O-CH<sub>3</sub>), 30.0 (propyl-CH<sub>2</sub>), 10.0 (propyl-CH<sub>3</sub>). ν<sub>max</sub>/cm<sup>-1</sup> (ATR); 3255, 1697, 1587, 1536, 1488, 1235 m/z (ESI) (100%, M+Na); (Found M+Na, 349.1168. C<sub>18</sub>H<sub>18</sub>N<sub>2</sub>O<sub>5</sub> requires M+Na, 349.1164).

### Synthesis of 2-(3-propanamidobenzamido)benzoic acid **58**



To a stirred solution of methyl 2-(3-propanamidobenzamido)benzoate (0.137 g, 0.42 mmol) in MeOH (2 mL) and H<sub>2</sub>O (2 mL) was added LiOH·H<sub>2</sub>O (0.084 g, 2.00 mmol) at RT. The resultant reaction mixture was then stirred at 64 °C for 12 hours. The reaction mixture was then cooled to RT, concentrated in vacuo and acidified with 2M HCl. The resultant precipitate was then collected by vacuum filtration and the solid was washed with H<sub>2</sub>O (5 mL) and DCM (5 mL) to afford the title compound as a colourless solid (0.119 g, 92%). *R*<sub>f</sub> 0.19 (10% MeOH-DCM). m.p >250 °C. δ<sub>H</sub> (500 MHz, DMSO) 12.10 (s, 1H, NH), 10.14 (s, 1H, NH), 8.74 (dd, 1H, *J* = 8.4, 0.8 Hz, 3-H), 8.28-8.27 (m, 1H, 2'-H), 8.07 (dd, 1H, *J* = 8.2, 1.5 Hz, 6-H), 7.86-7.84 (m, 1H, 4'H), 7.69-7.66 (m, 1H, 4-H), 7.61-7.59 (m, 1H, 6'-H), 7.51 (app t, 1H, *J* = 7.9 Hz, 5'-H), 7.23-7.20 (m, 1H, 5-H), 2.36 (q, 2H, *J* = 7.6 Hz, propyl-CH<sub>2</sub>), 1.11 (t, 3H, *J* = 7.6 Hz, propyl-CH<sub>3</sub>). δ<sub>C</sub> (500 MHz, DMSO) 172.8 (C=O), 170.5 (C=O), 165.1 (C=O), 141.6 (1-C), 140.5 (1'-C), 135.6 (4-C), 134.8 (3'-C), 131.8 (6-C), 129.8 (5'-C), 123.4 (5-C), 122.9 (4'-C), 121.4 (6'-C), 120.4 (3-C), 118.4 (2'-C), 116.8 (2-C), 30.0 (propyl-CH<sub>2</sub>), 10.0 (propyl-CH<sub>3</sub>) *v*<sub>max</sub>/cm<sup>-1</sup> (ATR); 3265, 1633, 1586, 1524, 1486, 1234 (ESI) (100%, M-H); (Found M-H, 307.1034. C<sub>17</sub>H<sub>16</sub>N<sub>2</sub>O<sub>4</sub> requires M-H, 307.1037).

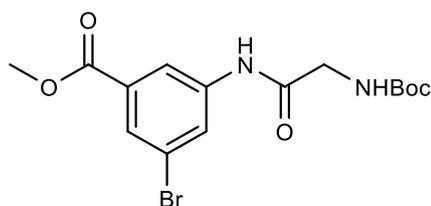
### Synthesis of 2-(3-aminobenzamido)benzoic acid (**59**)



To a stirred solution of methyl 2-(3-aminobenzamido)benzoate **55** (0.270 g, 1.00 mmol) in MeOH (2 mL) and H<sub>2</sub>O (2 mL) was added LiOH·H<sub>2</sub>O (0.168 g, 4.00 mmol) at RT. The resultant reaction mixture was then stirred at 64 °C for 12 hours. An additional amount of LiOH·H<sub>2</sub>O (0.042 g, 1.00 mmol) was then added and the reaction mixture was stirred at 64 °C for a further 2 hours. The reaction mixture was then cooled to RT, concentrated in vacuo and acidified with 2M HCl. The resultant precipitate was then collected by vacuum filtration

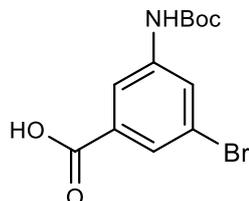
and washed with H<sub>2</sub>O (5 mL) and DCM (5 mL) to afford the title compound as a colourless solid (0.251 g, 98%). R<sub>f</sub> 0.25 (10% MeOH-DCM). m.p >250 °C. δ<sub>H</sub> (500 MHz, DMSO) 12.18 (s, 1H, amide NH), 8.69-8.68 (m, 1H, 3-H), 8.06 (dd, 1H, J = 7.9 Hz and 1.6 Hz, 6-H), 7.81-7.80 (m, 2H, 2'-H and 4'-H), 7.70-7.67 (m, 1H, 5-H), 7.63 (app t, 1H, J = 8.0 Hz, 5'-H), 7.51-7.50 (m, 1H, 6'-H), 7.25-7.22 (m, 1H, 4-H). δ<sub>C</sub> (500 MHz, DMSO) 170.4 (C=O), 164.3 (C=O), 141.3 (1-C), 136.4 (1'-C), 134.9 (5-C), 131.7 (6-C), 130.9 (5'-C), 125.6 (6'-C), 124.1 (4'-C), 123.7 (4-C), 121.2 (2'-C), 120.9 (3'-C), 120.6 (3-C), 117.4 (2-C). ν<sub>max</sub> / cm<sup>-1</sup> (solid): 3379, 2811, 2590, 1667, 1607, 1524, 1257. m/z (EI): (Found M-H, 255.0770 C<sub>14</sub>H<sub>12</sub>N<sub>2</sub>O<sub>3</sub> requires M-H, 255.0775). HPLC (RT: 1.67 min(100 % relative area)).

**Synthesis of methyl 3-bromo-5-(2-[(tert-butoxy)carbonyl]amino)acetamido)benzoate (64)**



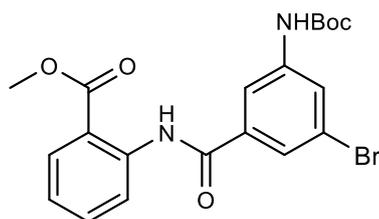
To a stirred solution of methyl 3-amino-5-bromobenzoate (0.230 g, 1.00 mmol) in DCM (20 mL) was added; Boc-Gly-OH (0.193 g, 1.10 mmol), EDC.HCl (0.211 g, 1.10 mmol) and DMAP (0.013 g, 0.10 mmol) at RT. The reaction mixture was then stirred at RT for 12 hours and then diluted with DCM (20 mL). The resultant solution was then washed with water (2 x 20 mL) and brine (2 x 20 mL). The organic layer was then dried over MgSO<sub>4</sub>, filtered and the solvent removed to afford the title compound as a colourless oil (0.247 g, 64%). R<sub>f</sub> 0.64 (30% EtOAc-Petroleum ether). δ<sub>H</sub> (500 MHz, DMSO) 10.34 (app s, 1H, amide NH), 8.17 (app s, 2H, 2-H and 4-H), 7.73 (app s, 1H, 6-H), 7.16-7.11 (m, 1H, Gly-NH), 3.88 (s, 3H, OCH<sub>3</sub>), 3.74-3.68 (m, 2H, Gly-CH<sub>2</sub>), 1.44 (s, 9H, Boc-CH<sub>3</sub>). δ<sub>C</sub> (500 MHz, DMSO) 169.5 (C=O), 165.3 (C=O), 156.4 (C=O), 141.3 (3-C), 132.5 (6-C), 126.3 (5-C), 125.9 (2-C), 122.2 (1-C), 118.9 (4-C), 78.6 (Boc-C), 53.1 (OCH<sub>3</sub>), 43.3 (Gly-CH<sub>2</sub>), 28.7 (3 x Boc-CH<sub>3</sub>) ν<sub>max</sub> / cm<sup>-1</sup> (solid): 2956, 1678, 1586, 1524, 1438, 1264. m/z (ESI) (100%, MNa<sup>+</sup>): (Found MNa<sup>+</sup>, 409.0368. C<sub>15</sub>H<sub>19</sub>BrN<sub>2</sub>O<sub>3</sub> requires MNa<sup>+</sup>, 409.0375).

### Synthesis of 3-bromo-5-[[tert-butoxy)carbonyl]amino}benzoic acid (67)



To a solution of 3-amino-5-bromobenzoic acid (2.16 g, 10 mmol) in dioxane (50 mL) and water (25 mL) was added triethylamine (2 mL, 15 mmol) and di-tert-butyl dicarbonate (3.27 g, 15 mmol). The reaction mixture was then stirred at RT for 12 hours. A further amount of di-tert-butyl dicarbonate (0.218 g, 1.00 mmol) was added and the reaction mixture was stirred at RT for a further 2 hours. The reaction mixture was then then concentrated and cooled to 0 °C. 2M HCl was then added dropwise and the resultant precipitate was collected by vacuum filtration. The solid was then washed with water (5 mL) to afford the title compound as a colourless solid (3.06 g, 97%).  $R_f$  0.42 (5% MeOH-DCM).  $\delta_H$  (500 MHz, DMSO) 9.78 (s, 1H, amide NH), 8.07 (app s, 1H, 4-H), 7.92 (app s, 1H, 2-H), 7.62 (app s, 1H, 6-H), 1.49 (s, 9H, Boc-CH<sub>3</sub>).  $\delta_C$  (500 MHz, DMSO) 166.9 (C=O), 153.0 (C=O), 141.9 (3-C), 133.7 (5-C), 125.4 (6-C), 124.4 (1-C), 122.0 (2-C), 118.1 (4-C), 80.4 (Boc-C), 28.9 (Boc-CH<sub>3</sub>).  $\nu_{max}/cm^{-1}$  (ATR); 3341, 1691, 1455, 1404, 1249, 1156.  $m/z$  (ESI) (100%, M-H); (Found M-H, 314.0028. C<sub>12</sub>H<sub>14</sub>BrNO<sub>4</sub> requires M-H, 314.0031).

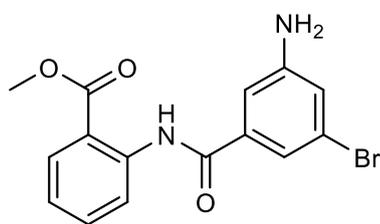
### Synthesis of methyl 2-(3-bromo-5-[[tert-butoxy)carbonyl]amino}benzamido)benzoate (68)



To a stirred solution of 3-bromo-5-[[tert-butoxy)carbonyl]amino}benzoic acid **67** (0.315 g, 1.00 mmol) in DCM (20 mL) was added; methyl 4-aminobenzoate (0.166 g, 1.10 mmol), EDC.HCl (0.211 g, 1.10 mmol) and DMAP (0.012 g, 0.10 mmol) at RT. The reaction mixture was then stirred at RT for 12 hours. The reaction mixture was then diluted with DCM (20 mL). The resultant solution was then washed with water (2 x 20 mL) and brine (2 x 20 mL). The organic layer was then dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude colourless solid. The crude solid was then purified by column chromatography (2% MeOH in DCM) to give the title compound as a colourless oil (0.311 g, 69%).  $R_f$  0.68 (2% MeOH-DCM).  $\delta_H$  (500 MHz,

DMSO) 11.42 (s, 1H, amide NH), 9.86 (s, 1H, amide NH), 8.40-8.38 (m, 1H, 3-H), 8.04-8.03 (m, 1H, 4'-H), 8.00-7.98 (m, 2H, 6-H and 2'-H), 7.70-7.67 (m, 2H, 6'-H and 4-H), 7.29-7.26 (m, 1H, 5-H), 3.88 (s, 3H, CH<sub>3</sub>), 1.5 (s, 9H, Boc-CH<sub>3</sub>).  $\delta_c$  (500 MHz, DMSO), 168.2 (C=O), 164.0 (C=O), 153.0 (C=O), 142.2 (3'-C), 139.5 (1-C), 137.5 (5'-C), 134.6 (4-C), 131.1 (6-C), 124.3 (6'-C), 123.8 (1'-C), 123.1 (5-C), 122.3 (2'-C), 122.0 (3-C), 118.7 (4'-C), 116.0 (2-C), 80.4 (Boc-C), 46.0 (CH<sub>3</sub>), 28.1 (Boc-CH<sub>3</sub>).  $\nu_{max}$  / cm<sup>-1</sup> (solid): 3318, 1677, 1587, 1519, 1443, 1258  $m/z$  (ESI) (100%, M+H); (Found M+H, 451.0680. C<sub>20</sub>H<sub>21</sub>BrN<sub>2</sub>O<sub>5</sub> requires M+H, 451.0610).

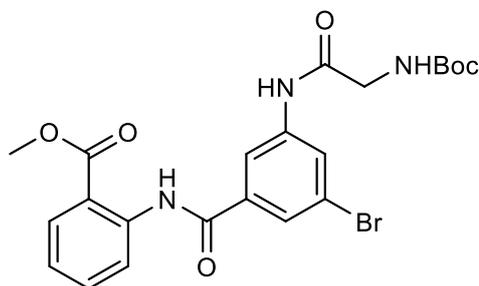
**Synthesis of 3-bromo-5-[[2-(methoxycarbonyl)phenyl]carbamoyl]anilinium chloride (69)**



To a RBF charged with methyl 2-(3-bromo-5-[[tert-butoxy]carbonyl]amino)benzamido benzoate **68** (0.675 g, 1.50 mmol) was added 4M HCl in dioxane (5 mL) at RT. The reaction mixture was then stirred at RT for 2 hours. The

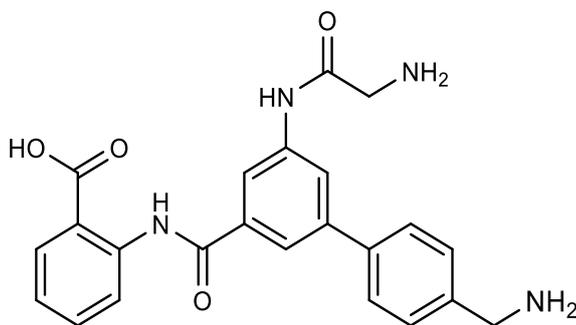
resultant ppt was then collected by vacuum filtration and was washed with DCM (5 mL) to afford the title compound as a colourless solid (0.475 g, 91 %),  $R_f$  0.68 (20% EtOAc-Petroleum ether).  $\delta_H$  (500 MHz, DMSO) 11.38 (s, 1H, amide NH), 8.41 (app d, 1H,  $J = 8.3$  Hz, 3-H), 7.99 (dd, 1H,  $J = 8.2$  Hz and 1.5 Hz, 6-H), 7.70-7.67 (m, 1H, 4-H), 7.42 (app s, 1H, 6'-H), 7.33 (app s, 1H, 4'-H), 7.28-7.25 (m, 1H, 5-H), 7.22 (app s, 1H, 2'-H), 6.59 (bs, 2H, NH<sub>2</sub>), 3.89 (s, 3H, CH<sub>3</sub>)  $\delta_c$  (500 MHz, DMSO) 168.3 (C=O), 164.2 (C=O), 147.0 (3'-C), 140.1 (1-C), 137.8 (5'-C), 134.6 (4-C), 131.1 (6-C), 124.2 (5-C), 122.8 (1'-C), 122.1 (2'-C), 121.9 (3-C), 119.9 (6'-C), 118.5 (2-C), 114.8 (4'-C), 53.2 (CH<sub>3</sub>)  $\nu_{max}$  / cm<sup>-1</sup> (solid): 2844, 1678, 1590, 1536, 1431, 1282.  $m/z$  (ESI) (100%, M+Na); (Found M+Na, 371.0002. C<sub>15</sub>H<sub>13</sub>BrN<sub>2</sub>O<sub>3</sub> requires M+Na, 371.0007).

**Synthesis of methyl 2-[3-bromo-5-(2-[[tert-butoxy]carbonyl]amino)acetamido]benzamido]benzoate (70)**



To a stirred solution 3-bromo-5-[[2-(methoxycarbonyl)phenyl]carbamoyl]anilinium chloride **69** (0.348 g, 1.00 mmol) in DCM (20 mL) was added; Boc-Gly-OH (0.193 g, 1.10 mmol), EDC.HCl (0.211 g, 1.10 mmol) and DMAP (0.012 g, 0.10 mmol) at RT. The reaction mixture was then stirred at RT for 12 hours. An additional amount of Boc-Gly-OH (0.045 g, 0.25 mmol) was then added and the reaction mixture was stirred at RT for an additional 3 hours. The reaction mixture was then diluted with DCM (20 mL). The resultant solution was then washed with water (2 x 20 mL) and brine (2 x 20mL). The organic layer was then dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude colourless solid. The crude solid was then recrystallized from ethanol to give the title compound as a colourless solid (0.328 g, 65%). R<sub>f</sub> 0.53 (2% MeOH-DCM). m.p 145-149 °C. δ<sub>H</sub> (500 MHz, DMSO) 11.38 (s, 1H, amide NH), 10.40 (s, 1H, amide NH), 8.37 (app d, 1H, J = 8.0 Hz, 3-H), 8.20-8.19 (m, 1H, 4'-H), 8.11-8.10 (m, 1H, 2'-H), 8.00 (dd, 1H, J = 8.0 and 1.5 Hz, 6-H), 7.78-7.77 (m, 1H, 6'-H), 7.71 (m, 1H, 4-H), 7.30-7.25 (m, 1H, 5-H), 7.13 (t, 1H, J = 6.1 Hz, Gly-NH), 3.89 (s, 3H, OCH<sub>3</sub>), 3.76 (d, 2H, J = 6.1 Hz, Gly-CH<sub>2</sub>), 1.41 (s, 9H, Boc-CH<sub>3</sub>) δ<sub>C</sub> (500 MHz, DMSO) 169.5 (C=O), 168.4 (C=O), 163.8 (C=O), 156.4 (C=O), 141.2 (3'-C), 140.1 (1-C), 137.4 (5'-C), 134.5 (4-C), 131.1 (6-C), 124.8 (6'-C), 124.5 (1'-C), 124.2 (5-C), 122.4 (2'-C), 122.2 (3-C), 119.1 (4'-C), 117.6 (2-C), 78.6 (Boc-C), 45.4 (OCH<sub>3</sub>), 44.3 (Gly-CH<sub>2</sub>), 28.7 (Boc-CH<sub>3</sub>) ν<sub>max</sub>/cm<sup>-1</sup> (ATR); 3075, 1670, 1600, 1547, 1440, 1266 (ESI) (100%, M-H); (Found M-H, 504.0759. C<sub>22</sub>H<sub>24</sub>BrN<sub>3</sub>O<sub>6</sub> requires M-H, 504.0775)

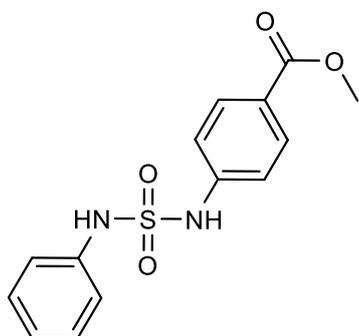
**Synthesis of {[4'-(azaniumylmethyl)-5-[(2-carboxyphenyl)carbamoyl]-[1,1'-biphenyl]-3-yl]carbamoyl}methanaminium dichloride (60)**



To a solution of methyl 2-[3-bromo-5-(2-[[tert-butoxy carbonyl]amino]acetamido)benzamido]enzoate **70** (0.505 g, 1.00 mmol) in dioxane were added; Pd(PPh<sub>3</sub>)<sub>4</sub> (0.058 g, 0.05mmol), 4-(N-Boc-aminomethyl phenylboronic acid (0.276 g, 1.10 mmol) and 2M Na<sub>2</sub>CO<sub>3</sub> (1.5 mL, 3.00 mmol). The reaction mixture was then heated at 80 °C for 4 hours. An additional amount of Pd(PPh<sub>3</sub>) (0.029 g, 0.03 mmol) was then added and the reaction mixture was heated at 80 °C for a further 16 hours. The reaction mixture was then cooled and water was then added (20 mL). The resultant solution was then washed with DCM (3 x 15 mL) and chloroform (3 x 15 mL). The aqueous layer was then acidified to pH 1 and stirred at RT for 30 mins. The solvent was then removed to give a crude brown solid. The solid was then purified by reverse phase ACC (gradient 0-40% MeCN-H<sub>2</sub>O in 0.1 % formic acid) to give the title compound as an off-white solid (0.195 g, 40%). R<sub>f</sub> 0.05 (10% Methanolic ammonia-DCM). m.p >250 °C. δ<sub>H</sub> (500 MHz, DMSO), 10.91 (s, 1H, amide NH), 8.72 (app d, 1H, J = 8.0 Hz, 3-H), 8.32-8.27 (m, 7H, 2 x NH<sub>3</sub><sup>+</sup> and 2'-H), 8.11 (app s, 1H, 4'-H), 8.08 (app d, 1H, J = 7.9 Hz, 6-H), 7.98 (app s, 1H, 6'-H), 7.75 (app d, 2H, J = 8.3 Hz, 2''-H), 7.67-7.61 (m, 3H, 3''-H and 4-H), 7.22 (app t, 1H, J = 7.8 Hz, 5-H), 4.13 (s, 2H, CH<sub>2</sub>), 3.88 (s, 2H, Gly CH<sub>2</sub>) δ<sub>C</sub> (500 MHz, DMSO), 170.6 (C=O), 165.8 (C=O), 164.6 (C=O), 141.5 (4''-C), 141.4 (5'-C), 140.0 (1-C), 139.8 (1''-C) 137.0 (1'-C), 134.6 (4-C), 134.3 (3'-C), 131.9 (6-C), 130.4 (3''-C), 127.5 (2''-C), 123.5 (5-C), 121.0 (4'-C), 120.8 (6'-C), 120.3 (3-C), 118.8 (2'-C), 116.5 (2-C), 42.5 (CH<sub>2</sub>), 41.6 (Gly CH<sub>2</sub>). ν<sub>max</sub>/cm<sup>-1</sup> (ATR) 2920, 1666, 1587, 1517, 1450, 1130. m/z (ESI) (100%, M-H); (Found M-H, 417.1600 C<sub>23</sub>H<sub>22</sub>N<sub>5</sub>O<sub>3</sub> requires M-H, 417.1693) HPLC (RT: 1.65 min(100 % relative area)).

### 6.3.3 Synthesis of Sulfamide Fragments

#### Synthesis of methyl 4-[(phenylsulfamoyl)amino]benzoate (**77**)

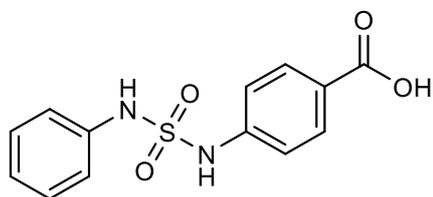


**Method 1:** To a stirred solution of sodium *N*-phenylsulfamate (0.200 g, 0.79 mmol) in  $\text{CHCl}_3$  (10 mL) was added  $\text{PCl}_5$  (0.340 g, 1.60 mmol) at RT. The reaction mixture was then refluxed for 4 hours. The reaction mixture was then cooled to RT, filtered and the solvent removed to give the crude sulfonyl chloride. The crude sulfonyl chloride was then dissolved in DCM (10 mL). This was followed by the addition of methyl 4-aminobenzoate (0.061 g, 0.41 mmol) and trimethylamine (0.061 mL, 0.44 mmol) at RT. The reaction mixture was then stirred at RT for 1 hour. The reaction mixture was then diluted with DCM (10 mL). The resultant solution was then washed with water (2 x 20 mL) and brine (2 x 20 mL). The organic layer was then dried over  $\text{MgSO}_4$ , filtered and the solvent removed to give a crude colourless solid. The crude solid was purified by column chromatography (2% MeOH-DCM) to afford the title compound as yellow solid (0.070 g, 29%).

**Method 2:** To a round-bottom flask charged with methyl 4-(sulfamoylamino)benzoate **86** (0.230 g, 1.00 mmol),  $\text{Cs}_2\text{CO}_3$  (0.456 g, 1.4 mmol), XPhos (0.036 g, 0.075 mmol) and  $\text{Pd}_2(\text{dba})_3$  was added dioxane (5 mL) and bromobenzene (0.105 mL, 1.00 mmol) at RT. The reaction mixture was then refluxed for 12 hours. The reaction mixture was then cooled to RT and the solvent removed to give a crude solid. The crude solid was then purified by reverse phase ACC (gradient 0-40% MeCN- $\text{H}_2\text{O}$  in 0.1 % formic acid) to afford the title compound as a colourless solid (0.230 g, 75%).  $R_f$  0.45 (2% MeOH-DCM)  $\delta_{\text{H}}$  (500 MHz, DMSO), 10.76 (s, 1H, sulfamide N-H), 10.46 (s, 1H, sulfamide N-H), 7.85 (app d, 2H, J = 8.7 Hz, 3-H), 7.27-7.21 (m, 4H, 2'-H and 3'-H) 7.10 (app d, 2H, J = 8.7 Hz, 2-H), 7.00 (app t, 1H, J = 7.38 Hz, 4'-H), 3.80 (s, 3H,  $\text{CH}_3$ )  $\delta_{\text{C}}$  (500 MHz, DMSO), 166.2 (C=O), 143.2 (1-C), 138.1 (1'-C), 130.9 (3-C), 129.5 (3'-C), 123.7 (4-C), 123.6 (4'-C), 119.1 (2'-C), 117.0 (2-C), 52.3 ( $\text{CH}_3$ )  $\nu_{\text{max}}/\text{cm}^{-1}$  (ATR); 3242, 1693, 1605, 1429, 1315, 1158.  $m/z$  (ESI) (100%, M-H); (Found M-H,

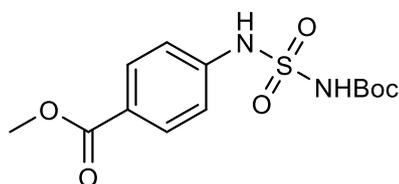
305.0594. C<sub>14</sub>H<sub>14</sub>N<sub>2</sub>O<sub>4</sub>S requires M-H, 305.0596) HPLC (RT: 1.72 min(100 % relative area)).

### Synthesis of 4-[(phenylsulfamoyl)amino]benzoic acid (71)



To a stirred solution of methyl 4-[(phenylsulfamoyl)amino]benzoate **77** (0.306 g, 1.00 mmol) in MeOH (2 mL) and H<sub>2</sub>O (2 mL) was added LiOH·H<sub>2</sub>O (0.168 g, 4.00 mmol) at RT. The resultant reaction mixture was then stirred at RT for 12 hours. The reaction mixture was then concentrated in vacuo and acidified with 2M HCl. The resultant precipitate was then collected by vacuum filtration and washed with H<sub>2</sub>O (5 mL) and DCM (5 mL) to afford the title compound as a colourless solid (0.289 g, 99%). R<sub>f</sub> 0.22 (10% MeOH-DCM). m.p >250 °C δ<sub>H</sub> (500 MHz, DMSO) 10.28 (s, 1H, sulfamide NH), 10.20 (s, 1H, sulfamide NH), 7.85-7.82 (app d, 2H, J = 8.8 Hz, 3-H), 7.27-7.24 (m, 2H, 3'-H), 7.22 (app d, 2H, J = 8.8 Hz, 2-H), 7.13-7.10 (m, 2H, 2'-H), 6.99 (app t, 1H, J = 7.4 Hz, 4'-H) δ<sub>C</sub> (500 MHz, DMSO), 167.4 (C=O), 143.2 (1-C), 138.1 (1'-C), 131.1 (3-C), 129.5 (3'-C), 124.8 (4-C), 123.7 (4'-C), 119.1 (2'-C), 117.0 (2-C). ν<sub>max</sub>/cm<sup>-1</sup> (ATR); 3291, 2679, 1680, 1607, 1321, 1151. *m/z* (ESI) (100%, M-H); (Found M-H, 291.0444. C<sub>13</sub>H<sub>12</sub>N<sub>2</sub>O<sub>4</sub>S requires M-H, 291.0445). HPLC (RT: 2.07 min(99 % relative area)).

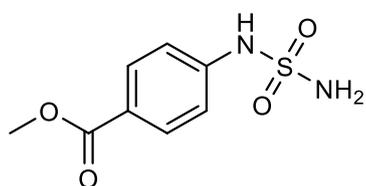
### Synthesis of methyl 4-[(tert-butoxycarbonyl)sulfonyl]amino]benzoate (90)



To a solution of chlorosulfonyl isocyanate (0.61 mL, 7.07 mmol) in DCM (20 mL) was added di-tert-butyl dicarbonate (1.54 g, 7.07 mmol) at 0 °C. The reaction mixture was then stirred at 0 °C for 15 mins, this was followed by the dropwise addition of a solution methyl 4-aminobenzoate (1.07 g, 7.07 mmol) and triethylamine (1.47 mL, 10.61 mmol) in DCM (10 mL) at 0 °C. The reaction mixture was then warmed to RT and stirred at that temperature for two hours. The reaction mixture was then diluted with DCM (10 mL). The resultant solution was then washed with 0.1 M HCl (20 mL), water (20 mL)

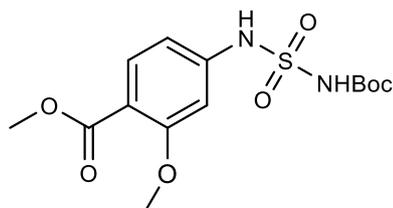
and brine (20 mL). The organic layer was dried over  $\text{MgSO}_4$ , filtered and the solvent removed to afford the title compound as an off-white solid (1.98 g, 85%).  $R_f$  0.53 (2% MeOH-DCM).  $\delta_H$  (500 MHz, DMSO) 11.56 (s, 1H, sulfamide NH), 10.90 (s, 1H, sulfamide NH), 7.93-7.91 (m, 2H, 3-H), 7.28-7.25 (m, 2H, 2-H), 3.82 (s, 3H,  $\text{OCH}_3$ ), 1.32 (s, 9H, Boc- $\text{CH}_3$ )  $\delta_C$  (500 MHz, DMSO) 166.3 (C=O), 154.0 (C=O), 142.7 (1-C), 130.9 (3-C), 124.5 (4-C), 117.3 (2-C), 82.4 (Boc-C), 52.4 (ester  $\text{CH}_3$ ), 28.0 (Boc- $\text{CH}_3$ ).  $\nu_{\text{max}}/\text{cm}^{-1}$  (ATR) 3226, 1690, 1606, 1520, 1436, 1146.  $m/z$  (ESI) (100%, M-H); (Found M-H, 329.0806.  $\text{C}_{13}\text{H}_{18}\text{N}_2\text{O}_6\text{S}$  requires M-H, 329.0813).

### Synthesis of methyl 4-(sulfamoylamino)benzoate (86)



To a RBF charged with methyl 4-[[[(tert-butoxy)carbonyl]amino]sulfonyl]amino]benzoate **90** (0.500 g, 1.51 mmol) was added 4M HCl in dioxane (5 mL) at RT. The resultant solution was then stirred at RT for 2 hours. The solvent was then removed to give a crude solid, which was triturated with DCM to afford the title compound as a colourless solid (0.366 g, 95%).  $R_f$  0.29 (2% MeOH-DCM). m.p 134-136 °C.  $\delta_H$  (500 MHz, DMSO) 10.14 (s, 1H, sulfamide NH), 7.89-7.87 (m 2H, 3-H), 7.36 (bs, 2H, sulfamide  $\text{NH}_2$ ), 7.24-7.22 (m, 2H, 2-H), 3.82 (s, 3H,  $\text{OCH}_3$ )  $\delta_C$  (500 MHz, DMSO), 166.4 (C=O), 142.7 (1-C), 130.8 (3-C), 124.4 (4-C), 116.7 (2-C), 52.3 ( $\text{CH}_3$ ).  $\nu_{\text{max}}/\text{cm}^{-1}$  (ATR) 3256, 1697, 1607, 1437, 1290, 1146.  $m/z$  (ESI) (100%, M-H); (Found M-H, 229.0284.  $\text{C}_8\text{H}_{10}\text{N}_2\text{O}_4\text{S}$  requires M-H, 229.0289).

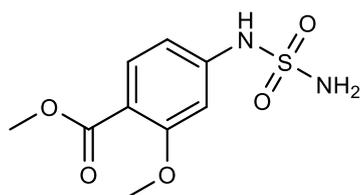
### Synthesis of Methyl 4-[[[(tert-butoxy)carbonyl]amino]sulfonyl]amino]-2-methoxybenzoate



To a solution of chlorosulfonyl isocyanate (0.61 mL, 7.07 mmol) in DCM (20 mL) was added di-tert-butyl dicarbonate (1.54 g, 7.07 mmol) at 0 °C. The reaction mixture was then stirred at 0 °C for 15 mins, this was followed by the dropwise addition of a solution methyl 4-amino-2-methoxybenzoate (1.27 g, 7.07 mmol) and triethylamine (1.47 mL, 10.61 mmol) in DCM (20

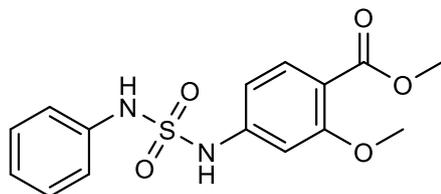
mL) at 0 °C. The reaction mixture was then warmed to RT and stirred at that temperature for 2 hours. The reaction mixture was then diluted with DCM (20 mL) and washed with 0.1 M HCl (20 mL), water (20 mL) and brine (20 mL). The organic layer was dried over MgSO<sub>4</sub>, filtered and the solvent removed to give the title compound as an off-white solid (2.27 g, 89%). R<sub>f</sub> 0.63 (2% MeOH-DCM) δ<sub>H</sub> (500 MHz, DMSO) 11.35 (s, 1H, sulfamide NH), 10.56 (s, 1H, sulfamide NH), 7.47 (app d, 1H, J = 8.5 Hz, 5-H), 6.71-6.70 (m, 1H, 2-H), 6.60-6.58 (m, 1H, 4-H). 3.59 (s, 3H, ester CH<sub>3</sub>), 3.53 (s, 3H, OMe), 1.13 (s, 9H, Boc CH<sub>3</sub>) δ<sub>C</sub> (500 MHz, DMSO) 165.9 (C=O), 160.0 (1-C), 150.4 (C=O), 143.3 (3-C), 132.2 (5-C), 114.5 (6-C), 110.0 (4-C), 102.4 (2-C), 82.3 (Boc C), 56.2 (OMe), 55.6 (ester CH<sub>3</sub>), 27.8 (Boc CH<sub>3</sub>) ν<sub>max</sub>/cm<sup>-1</sup> (ATR) 3268, 3086, 1702, 1609, 1330, 1140 *m/z* (ESI): (Found M-H, 359.0919 C<sub>14</sub>H<sub>20</sub>N<sub>2</sub>O<sub>7</sub>S requires M-H, 359.0913).

#### Synthesis of methyl 2-methoxy-4-(sulfamoylamino)benzoate (94)



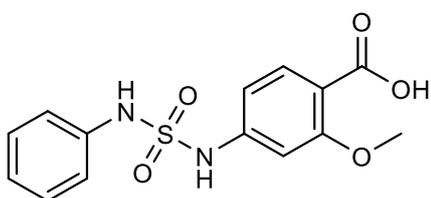
To a RBF charged with Methyl 4-[[[[(tert-butoxy)carbonyl]amino]sulfonyl]amino]-2-methoxy benzoate (0.500 g, 1.38 mmol) was added 4M HCl in dioxane (5 mL) at RT. The reaction mixture was then stirred at RT for 2 hours. The resultant ppt was collected by vacuum filtration and the solid was washed with water (5 mL) and DCM (5 mL) to afford the title compound as a colourless solid (0.327 g, 91%). R<sub>f</sub> 0.41 (2% MeOH-DCM). δ<sub>H</sub> (500 MHz, DMSO) 9.91 (s, 1H, sulfamide NH), 7.55 (app d, 1H, J = 8.5 Hz, 5-H), 7.27 (bs, 2H, sulfamide NH<sub>2</sub>), 6.78 (app d, 1H, J = 1.9 Hz, 2-H), 6.66 (dd, 1H, J = 8.5 Hz and 1.9 Hz, 4-H). 3.70 (s, 3H, ester CH<sub>3</sub>), 3.64 (s, 3H, OMe). δ<sub>C</sub> (500 MHz, DMSO) 165.9 (C=O), 160.0 (1-C), 145.3 (3-C), 132.8 (5-C), 112.5 (6-C), 108.9 (4-C), 101.1 (2-C), 56.1 (OMe), 51.9 (ester CH<sub>3</sub>). ν<sub>max</sub>/cm<sup>-1</sup> (ATR) 3346, 3239. 1676, 1599, 1339, 1146. *m/z* (ESI): (Found M-H, 259.0392 C<sub>9</sub>H<sub>12</sub>N<sub>2</sub>O<sub>5</sub>S requires M-H, 259.0381)

### Synthesis of methyl 2-methoxy-4-[(phenylsulfamoyl)amino]benzoate (99)



To a round-bottom flask charged with methyl 2-methoxy-4-(sulfamoylamino)benzoate **94** (0.260 g, 1.00 mmol), Cs<sub>2</sub>CO<sub>3</sub> (0.456 g, 1.4 mmol), XPhos (0.036 g, 0.075 mmol) and Pd<sub>2</sub>(dba)<sub>3</sub> (0.029 g, 0.05 mmol) was added dioxane (5 mL) and bromobenzene (0.105 mL, 1.00 mmol). The reaction mixture was then refluxed for 12 hours. The reaction mixture was then cooled and the solvent removed to give a crude solid. The crude solid was then purified by reverse phase ACC (gradient 0-40% MeCN-H<sub>2</sub>O in 0.1 % formic acid) to give the title compound as a colourless solid (0.265 g, 79%). R<sub>f</sub> 0.51 (2% MeOH-DCM) δ<sub>H</sub> (500 MHz, DMSO) 10.64 (s, 1H, sulfamide NH), 10.48 (s, 1H, sulfamide NH), 7.60 (app d, 1H, J = 8.5 Hz, 5-H), 7.28-7.25 (m, 2H, 3'-H), 7.13-7.11 (m, 2H, 2'-H), 7.03-7.00 (m, 1H, 4'-H), 6.82 (app d, 1H, J = 1.9 Hz, 2-H), 6.73 (dd, 1H, J = 8.5 and 1.9 Hz, 4-H), 3.77 (s, 3H, ester CH<sub>3</sub>), 3.72 (s, 3H, OMe) (500 MHz, DMSO) 165.8 (C=O), 160.2 (1-C), 143.9 (3-C), 138.2 (1'-C), 132.8 (5-C), 129.9 (3'-C), 123.7 (4'-C), 119.2 (2'-C), 113.5 (6-C), 109.2 (4-C), 101.6 (2-C), 55.2 (OMe), 52.1 (ester CH<sub>3</sub>) ν<sub>max</sub>/cm<sup>-1</sup> (ATR) 3276, 3167, 1699, 1331, 1153 *m/z* (ESI): (Found M-H, 335.0706 C<sub>15</sub>H<sub>16</sub>N<sub>2</sub>O<sub>5</sub>S requires M-H, 335.0706).

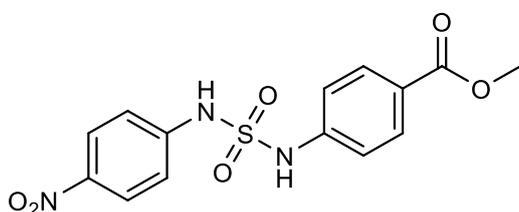
### Synthesis of 2-methoxy-4-[(phenylsulfamoyl)amino]benzoic acid (104)



To a stirred solution of 2-methoxy-4-[(phenylsulfamoyl)amino]benzoic acid **99** (0.150 g, 0.45 mmol) in MeOH (2 mL) and H<sub>2</sub>O (2 mL) was added LiOH·H<sub>2</sub>O (0.076 g, 1.80 mmol) at RT. The resultant reaction mixture was then stirred at reflux for 12 hours. The reaction mixture was then cooled to RT, concentrated in vacuo and acidified with 2M HCl. The resultant precipitate was then collected by vacuum filtration and the solid was washed with H<sub>2</sub>O (5 mL) and DCM (5 mL) to afford the title compound as a colourless solid (0.101 g, 70%). R<sub>f</sub> 0.25 (2% MeOH-DCM). m.p >250 °C. δ<sub>H</sub> (500 MHz, DMSO) 10.60 (s, 1H, sulfamide NH), 10.47 (s, 1H, sulfamide NH),

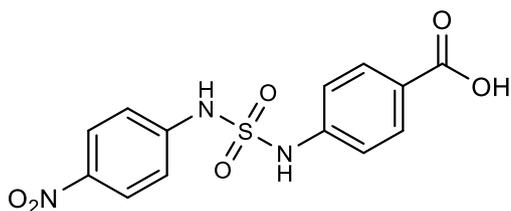
7.61 (app d, 1H, J = 8.6 Hz, 5-H), 7.28-7.25 (m, 2H, 3'-H), 7.13-7.11 (m, 2H, 2'-H), 7.03-7.00 (m, 1H, 4'-H), 6.82 (app d, 1H, J = 2.0 Hz, 2-H), 6.73 (dd, 1H, J = 2.0 and 8.6 Hz, 4-H), 3.76 (s, 3H, OMe)  $\delta_c$  (500 MHz, DMSO) 166.8 (C=O), 160.2 (1-C), 143.5 (3-C), 138.2 (1'-C), 132.9 (5-C), 129.5 (3'-C), 123.7 (4'-C), 119.1 (2'-C), 114.6 (6-C), 109.2 (4-C), 101.6 (2-C), 56.2 (OMe)  $\nu_{\max}/\text{cm}^{-1}$  (ATR) 3208, 1707, 1605, 1345, 1143  $m/z$  (ESI): (Found M-H, 321.0548  $\text{C}_{14}\text{H}_{14}\text{N}_2\text{O}_5\text{S}$  requires M-H, 321.0551) HPLC (RT: 2.03 min(100 % relative area)).

### Synthesis of methyl 4-[[4-(4-nitrophenyl)sulfamoyl]amino]benzoate (100)



To a round-bottom flask charged with methyl 4-(sulfamoylamino) benzoate **86** (0.230 g, 1.00 mmol),  $\text{Cs}_2\text{CO}_3$  (0.456 g, 1.4 mmol), XPhos (0.036 g, 0.075 mmol), 1-bromo-4-nitrobenzene (0.202 g, 1.00 mmol) and  $\text{Pd}_2(\text{dba})_3$  (0.029 g, 0.05 mmol) was added dioxane (5 mL). The reaction mixture was then refluxed for 12 hours. A further amount of 1-bromo-4-nitrobenzene (0.101 g, 0.50 mmol) and  $\text{Pd}_2(\text{dba})_3$  (0.029 g, 0.05 mmol) was then added and the reaction mixture was refluxed for a further 5 hours. The reaction mixture was then cooled to RT and the solvent removed to give a crude solid. The crude solid was then purified by reverse phase ACC (gradient 0-40% MeCN- $\text{H}_2\text{O}$  in 0.1 % formic acid) to afford the title compound as a colourless solid (0.228 g, 65%).  $R_f$  0.31 (2% MeOH-DCM).  $\delta_H$  (500 MHz, DMSO) 11.47 (s, 1H, sulfamide NH), 11.27 (s, 1H, sulfamide NH), 8.26-8.22 (m, 2H, 2'-H), 7.93-7.89 (m, 2H, 3-H), 7.38-7.34 (m, 2H, 3'-H) 7.29-7.26 (m, 2H, 2-H), 3.85 (s, 3H,  $\text{CH}_3$ )  $\delta_c$  (500 MHz, DMSO) 166.2 (C=O), 144.6 (4'-C), 142.7 (1-C), 142.4 (1'-C), 131.1 (3-C), 125.8 (3'-C), 125.7 (4-C), 117.5 (2-C), 111.3 (2'-C), 52.4 ( $\text{CH}_3$ )  $\nu_{\max}/\text{cm}^{-1}$  (ATR); 3126, 1699, 1609, 1439, 1289, 1144  $m/z$  (EI): (Found M-H, 350.0450  $\text{C}_{14}\text{H}_{13}\text{N}_3\text{O}_6\text{S}$  requires M-H, 350.0452)

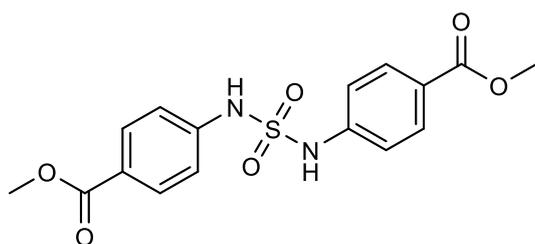
### Synthesis of 4-[[4-(nitrophenyl)sulfamoyl]amino]benzoic acid (**105**)



To a stirred solution of methyl 4-[[4-(4-nitrophenyl)sulfamoyl]amino]benzoate **100** (0.200 g, 0.57 mmol) in MeOH (2 mL) and H<sub>2</sub>O (2 mL) was added LiOH·H<sub>2</sub>O (0.096 g, 2.28 mmol) at RT.

The resultant reaction mixture was then stirred at RT for 12 hours. The reaction mixture was then concentrated and acidified with 2M HCl. The resultant precipitate was then collected by vacuum filtration and washed with H<sub>2</sub>O (5 mL) to afford the title compound as a colourless solid (0.125 g, 65%). *R<sub>f</sub>* 0.16 (10% MeOH-DCM). m.p 234-238 °C  $\delta_{\text{H}}$  (500 MHz, DMSO) 11.38 (s, 1H, sulfamide NH), 11.15 (s, 1H, sulfamide NH), 8.20-8.17 (m, 2H, 2'-H), 7.88-7.83 (m, 2H, 3-H), 7.37-7.29 (m, 2H, 3'-H), 7.21-7.19 (m, 2H, 2-H).  $\delta_{\text{C}}$  (500 MHz, DMSO) 167.3 (C=O), 144.6 (4'-C), 142.5 (1-C), 142.0 (1'-C), 131.0 (3-C), 125.8 (3'-C), 125.2 (4-C), 117.5 (2-C), 117.2 (2'-C).  $\nu_{\text{max}}$ /cm<sup>-1</sup> (ATR); 3157, 1685, 1607, 1522, 1346, 1153 *m/z* (ESI): (Found M-H, 336.0304 C<sub>13</sub>H<sub>11</sub>N<sub>3</sub>O<sub>6</sub>S requires M-H, 336.0296) HPLC (RT: 2.25 min(100 % relative area)).

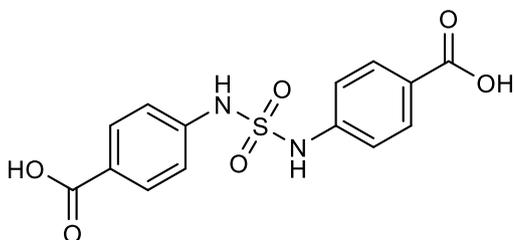
### Synthesis of methyl 4-([4-(methoxycarbonyl) phenyl] sulfamoyl) amino benzoate (**101**)



To a round-bottom flask charged with methyl 4-(sulfamoylamino) benzoate **86** (0.230 g, 1.00 mmol), Cs<sub>2</sub>CO<sub>3</sub> (0.456 g, 1.4 mmol), XPhos (0.036 g, 0.075 mmol), methyl 4-bromobenzoate (0.215 g, 1.00 mmol) and Pd<sub>2</sub>(dba)<sub>3</sub> (0.029 g, 0.05 mmol) was added dioxane (5 mL). The reaction mixture was then refluxed for 4 hours. The reaction mixture was then cooled to RT and the solvent removed to give a crude solid. The crude solid was then purified by reverse phase ACC (gradient 0-40% MeCN-H<sub>2</sub>O in 0.1 % formic acid) to afford the title compound as a colourless solid (0.277 g, 76%). *R<sub>f</sub>* 0.46 (2% MeOH-DCM).  $\delta_{\text{H}}$  (500 MHz, DMSO) 11.08 (s, 2H, sulfamide NH), 7.85 (app d, 4H, J = 8.8 Hz, 3-H), 7.23 (app d, 4H, J = 8.8 Hz, 2-H), 3.78 (s, 6H, 2 x OMe)  $\delta_{\text{C}}$  (500

MHz, DMSO) 166.2 (C=O), 142.8 (1-C), 131.1 (3-C), 124.0 (4-C), 117.3 (2-C), 52.4 (OMe).  $\nu_{\max}$  /  $\text{cm}^{-1}$  (solid): 3292, 1605, 1679, 1294, 1115.  $m/z$  (ESI): (Found M-H, 363.0670  $\text{C}_{16}\text{H}_{16}\text{N}_2\text{O}_6\text{S}$  requires M-H, 363.0656).

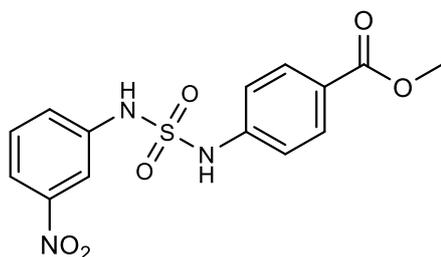
### Synthesis of 4-[[4-(4-carboxyphenyl)sulfamoyl]amino]benzoic acid (**106**)



To a stirred solution of methyl 4-([4-(methoxycarbonyl) phenyl] sulfamoyl) amino) benzoate **101** (0.250 g, 0.69 mmol) in MeOH (2 mL) and H<sub>2</sub>O (2 mL) was added LiOH·H<sub>2</sub>O (0.116 g, 2.76 mmol) at RT. The resultant

reaction mixture was then stirred at reflux for 12 hours. The reaction mixture was then cooled to RT, concentrated and acidified with 2M HCl. The resultant precipitate was then collected by vacuum filtration and washed with H<sub>2</sub>O (5 mL) and DCM (5 mL) to give the title compound as a colourless solid (0.223 g, 96%).  $m.p$  >250 °C.  $\delta_{\text{H}}$  (500 MHz, DMSO) 10.95 (s, 2H, 2 x sulfamide NH), 7.83 (app d, 4H,  $J = 8.9$  Hz, 3-H), 7.20 (app d, 4H,  $J = 8.9$  Hz, 2-H)  $\delta_{\text{C}}$  (500 MHz, DMSO) 167.3 (C=O), 142.4 (1-C), 131.2 (3-C), 125.2 (4-C), 117.2 (2-C).  $\nu_{\max}$  /  $\text{cm}^{-1}$  (solid): 3210, 2948, 1688, 1606, 1291, 1151  $m/z$  (ESI): (Found M-H, 335.0345  $\text{C}_{14}\text{H}_{12}\text{N}_2\text{O}_6\text{S}$  requires M-H, 335.0343). HPLC (RT: 1.79 min(100 % relative area)).

### Synthesis of methyl 4-[[3-(3-nitrophenyl)sulfamoyl]amino]benzoate **102**

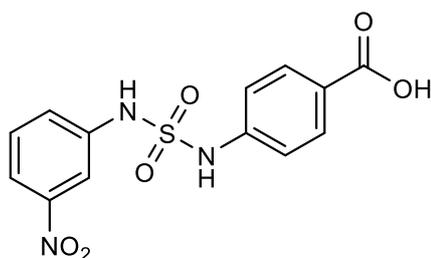


To a round-bottom flask charged with methyl 4-(sulfamoylamino)benzoate **86** (0.230 g, 1.00 mmol), Cs<sub>2</sub>CO<sub>3</sub> (0.456 g, 1.4 mmol), XPhos (0.036 g, 0.075 mmol), 1-bromo-3-nitrobenzene (0.202 g, 1.00 mmol) and Pd<sub>2</sub>(dba)<sub>3</sub> (0.029 g, 0.05 mmol)

was added dioxane (5 mL). The reaction mixture was then refluxed for 12 hours. A further amount of 1-bromo-3-nitrobenzene (0.101 g, 0.50 mmol) and Pd<sub>2</sub>(dba)<sub>3</sub> (0.029 g, 0.05 mmol) was then added and the reaction mixture was refluxed for a further 6 hours. The reaction mixture was then cooled to RT. The reaction mixture was then filtered through a pad of Celite and the

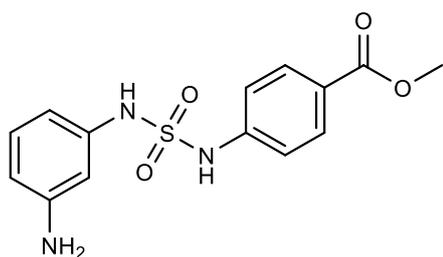
solvent was then removed to give a crude solid. The crude solid was then purified by reverse phase ACC (gradient 0-40% MeCN-H<sub>2</sub>O in 0.1 % formic acid) to afford the title compound as a colourless solid (0.242 g, 69%). *R<sub>f</sub>* 0.32 (2% MeOH-DCM).  $\delta_{\text{H}}$  (500 MHz, DMSO) 11.18 (bs, 1H, sulfamide N-H), 11.13 (bs, 1H, sulfamide N-H), 7.98-7.97 (m, 1H, 2'-H), 7.87-7.85 (m, 3H, 3-H and 6'-H), 7.56 (app t, 1H, *J* = 8.1 Hz, 5'-H), 7.52-7.50 (m, 1H, 4'-H), 7.23 (app d, 2H, *J* = 8.9 Hz, 2-H), 3.80 (s, 3H, OCH<sub>3</sub>).  $\delta_{\text{C}}$  (500 MHz, DMSO), 166.2 (C=O), 148.7 (3'-C), 142.7 (1-C), 140.0 (1'-C), 131.1 (5'-C), 131.0 (3-C) 124.5 (6'-C), 124.1 (C-4), 118.1 (4'-C), 117.4 (2-C), 112.6 (2'-C), 54.1 (OCH<sub>3</sub>).  $\nu_{\text{max}}/\text{cm}^{-1}$  (ATR); 3224, 1703, 1610, 1510, 1144. *m/z* (EI): (Found M-H, 350.0459 C<sub>14</sub>H<sub>13</sub>N<sub>3</sub>O<sub>6</sub>S requires M-H, 350.0452).

### Synthesis of 4-[[[3-nitrophenyl)sulfamoyl]amino]benzoic acid **107**



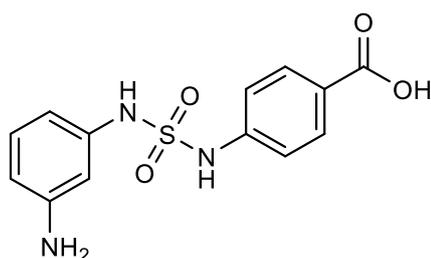
To a stirred solution of methyl 4-[[[3-nitrophenyl)sulfamoyl]amino]benzoate **102** (0.120 g, 0.34 mmol) in MeOH (2 mL) and H<sub>2</sub>O (2 mL) was added LiOH·H<sub>2</sub>O (0.058 g, 1.36 mmol) at RT. The resultant reaction mixture was then stirred at RT for 12 hours. A further amount of LiOH·H<sub>2</sub>O (0.042 g, 1.00 mmol) was then added and the reaction mixture was stirred at RT for a further 3 hours. The reaction mixture was then concentrated in vacuo and acidified with 2M HCl. The resultant precipitate was then collected by vacuum filtration and washed with H<sub>2</sub>O (5 mL) and DCM (5 mL) to afford the title compound as a colourless solid (0.103 g, 90%). *R<sub>f</sub>* 0.14 (10% MeOH-DCM). *m.p.* >250 °C.  $\delta_{\text{H}}$  (500 MHz, DMSO) 11.11 (s, 1H, sulfamide N-H), 11.05 (s, 1H, sulfamide NH), 7.98-7.97 (m, 1H, 2'-H), 7.87-7.83 (m, 3H, 3-H and 6'-H), 7.57 (app t, 1H, *J* = 8.2 Hz, 5'-H), 7.51-7.49 (m, 1H, 4'-H), 7.20 (app d, 2H, *J* = 8.9 Hz, 2-H)  $\delta_{\text{C}}$  (500 MHz, DMSO) 167.2 (C=O), 148.7 (3'-C), 142.3 (1-C), 139.6 (1'-C), 131.2 (5'-C), 131.1 (3-C), 125.4 (6'-C), 124.4 (4-C), 118.0 (4'-C), 117.9 (2-C), 112.0 (2'-C)  $\nu_{\text{max}}/\text{cm}^{-1}$  (ATR); 3157, 1685, 1607, 1522, 1346, 1102 *m/z* (ESI): (Found M-H, 336.0308 C<sub>13</sub>H<sub>11</sub>N<sub>3</sub>O<sub>6</sub>S requires M-H, 336.0296) HPLC (RT: 2.28 min (98 % relative area)).

### Synthesis of methyl 4-[[[3-aminophenyl)sulfamoyl]amino]benzoate (103)



To a solution of 4-[[[3-nitrophenyl)sulfamoyl]amino]benzoate **102** (0.176 g, 0.5 mmol) in MeOH (20 mL) was added 10% Pd/C (0.125 g). The resultant slurry was then stirred at RT under an atmosphere of hydrogen (1 atmosphere) for 2 hours. The reaction mixture was then filtered through a bed of Celite and the solvent was then removed to give a crude purple solid. The crude solid was then purified by reverse phase ACC (gradient 0-40% MeCN-H<sub>2</sub>O in 0.1 % formic acid) to give the title compound as a colourless solid (0.135 g, 84%). *R<sub>f</sub>* 0.41 (2% MeOH-DCM)  $\delta_{\text{H}}$  (500 MHz, DMSO) 7.84 (app d, 2H, *J* = 8.7 Hz, 3-H), 7.23 (app d, 2H, *J* = 8.7 Hz, 2-H), 6.85 (app t, 1H, *J* = 7.9 Hz, 5'-H), 6.32-6.29 (m, 2H, 2'-H and 6'-H), 6.22 (app d, 1H, *J* = 7.9 Hz, 4'-H), 5.04 (bs, 2H, NH<sub>2</sub>), 3.80 (s, 3H, CH<sub>3</sub>)  $\delta_{\text{C}}$  (500 MHz, DMSO) 166.7 (C=O), 149.7 (3'-C), 143.7 (1-C), 138.7 (1'-C), 130.8 (3-C), 129.7 (5'-C), 123.2 (4-C), 116.9 (2-C), 110.1 (4'-C), 107.2 (6'-C), 105.3 (2'-C), 52.3 (CH<sub>3</sub>).  $\nu_{\text{max}}/\text{cm}^{-1}$  (ATR); 3229, 1693, 1606, 1489, 1283, 1145 *m/z* (ESI): (Found M-H, 320.0715 C<sub>14</sub>H<sub>15</sub>N<sub>3</sub>O<sub>4</sub>S requires 320.0709)

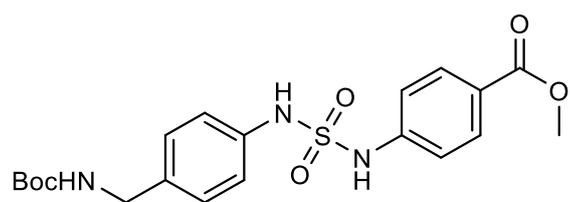
### Synthesis of 4-[[[3-aminophenyl)sulfamoyl]amino]benzoic acid (108)



To a stirred solution of methyl 4-[[[3-aminophenyl)sulfamoyl]amino]benzoate **103** (0.100 g, 0.31 mmol) in MeOH (2 mL) and H<sub>2</sub>O (2 mL) was added LiOH·H<sub>2</sub>O (0.052 g, 1.24 mmol) at RT. The resultant reaction mixture was then stirred at reflux for 12 hours. The reaction mixture was then cooled to RT and the resultant solution was concentrated and acidified with 2M HCl. The resultant precipitate was then collected by vacuum filtration and washed with H<sub>2</sub>O (5 mL) and DCM (5 mL) to afford the title compound as a colourless solid (0.080 g, 84%). *R<sub>f</sub>* 0.22 (10% MeOH-DCM). *m.p* >250 °C.  $\delta_{\text{H}}$  (400 MHz, DMSO) 7.80 (app d, 2H, *J* = 8.7 Hz, 3-H), 7.18 (app d, 2H, *J* = 8.7 Hz, 2-H), 6.85 (app t, 1H, *J* = 7.8 Hz, 5'-H), 6.36-6.32 (m, 2H, 2'-H and 6'-H), 6.21 (app d, 1H, *J* = 7.8 Hz, 4'-H)  $\delta_{\text{C}}$  (400 MHz, DMSO)

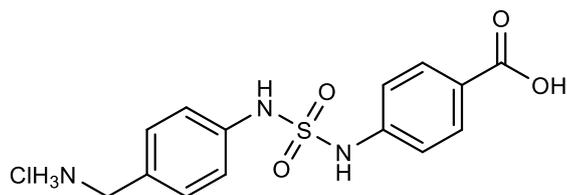
168.5 (C=O), 149.7 (3'-C), 142.3 (1-C), 139.0 (1'-C), 130.7 (3-C), 129.6 (5'-C), 128.0 (4-C) 117.0 (2-C), 109.8 (4'-C), 107.2 (6'-C), 105.2 (2'-C)  $\nu_{\max}/\text{cm}^{-1}$  (ATR); 3267, 3086, 1702, 1610, 1584, 1241, 1144  $m/z$  (ESI): (Found M-H, 306.0558 C<sub>13</sub>H<sub>11</sub>N<sub>3</sub>O<sub>4</sub>S requires M-H, 306.0549) HPLC (RT: 1.10 min(100 % relative area)).

### Synthesis of Methyl 4-({[4-({[(tert-butoxy)carbonyl]amino}methyl)phenyl]sulfamoyl}amino)benzoate (114)



A round-bottom flask was charged with methyl 4-(sulfamoylamino)benzoate **86** (0.230 g, 1.00 mmol), Cs<sub>2</sub>CO<sub>3</sub> (0.456 g, 1.4 mmol), XPhos (0.036 g, 0.075 mmol), tert-butyl 4-bromobenzylcarbamate (0.286 g, 1.00 mmol) and Pd<sub>2</sub>(dba)<sub>3</sub> (0.029 g, 0.05 mmol) was added dioxane (5 mL). The reaction mixture was then refluxed for 6 hours. An additional amount of tert-butyl 4-bromobenzylcarbamate (0.143 g, 0.50 mmol) and Pd<sub>2</sub>(dba)<sub>3</sub> (0.029 g, 0.05 mmol) was then added and the reaction mixture was refluxed for an additional 12 hours. This was again followed by a further addition of tert-butyl 4-bromobenzylcarbamate (0.143 g, 0.50 mmol) and Pd<sub>2</sub>(dba)<sub>3</sub> (0.029 g, 0.05 mmol) and the reaction mixture was refluxed for an additional 6 hours. The reaction mixture was then cooled to RT and the solvent removed to give a crude solid. The crude solid was then purified by reversephase ACC (gradient 0-40% MeCN-H<sub>2</sub>O in 0.1 % formic acid) to afford the title compound as a colourless solid (0.265 g, 61%).  $R_f$  0.28 (2% MeOH-DCM)  $\delta_H$  (500 MHz, DMSO) 10.71 (s, 1H, sulfamide N-H), 10.40 (s, 1H, sulfamide N-H), 7.85 (app d, 2H, J = 8.7 Hz, 3-H), 7.30 (t, 1H, J = 6.1 Hz, Boc N-H), 7.24 (app d, 2H, J = 8.4 Hz, 3'-H), 7.10 (app d, 2H, J = 8.7 Hz, 2-H), 7.04 (app d, 2H, J = 8.4 Hz, 2'-H), 4.02 (d, 2H, J = 6.1 Hz, CH<sub>2</sub>), 3.80 (s, 3H, OCH<sub>3</sub>), 1.37 (s, 9H, Boc-CH<sub>3</sub>)  $\delta_C$  (500 MHz, DMSO) 166.2 (C=O), 156.2 (C=O), 143.2 (1-C), 136.6 (1'-C), 135.7 (3-C), 130.9 (3'-C), 128.2 (4'-C), 123.6 (4-C), 119.4 (2'-C), 117.0 (2-C), 78.2 (Boc-C), 52.3 (OCH<sub>3</sub>), 43.3 (CH<sub>2</sub>), 28.7 (Boc-CH<sub>3</sub>)  $\nu_{\max}/\text{cm}^{-1}$  (ATR); 3258, 1682, 1636, 1607, 1421, 1186  $m/z$  (ESI): (Found M-H, 434.1373 C<sub>20</sub>H<sub>25</sub>N<sub>3</sub>O<sub>6</sub>S requires M-H, 434.1364)

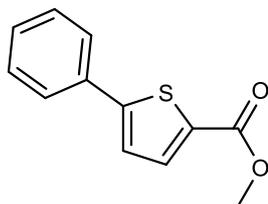
### Synthesis of (4-(((4-carboxyphenyl)sulfamoyl)amino)phenyl) methanaminium chloride (**112**)



To a solution of Methyl 4 -(((4-(((tert-butoxy)carbonyl)amino) methyl)phenyl)sulfamoyl)amino benzoate **114** (0.100 g, 0.23 mmol) in MeOH (10 mL) was added conc. HCl (4 mL) and the resultant mixture was refluxed for 12 hours. The reaction mixture was then cooled to RT and the solvent was then removed to give a crude colourless solid. The crude solid was purified by reverse phase ACC (gradient 0-30% MeCN-H<sub>2</sub>O in 0.1 % formic acid) to give the title compound as a colourless solid (0.061 g, 75%). *R<sub>f</sub>* 0.09 (10% Methanolic ammonia-DCM) m.p >250 °C.  $\delta_{\text{H}}$  (500 MHz, DMSO) 10.84 (bs, 1H, sulfamide NH), 10.65 (bs, 1H, sulfamide NH), 8.33 (bs, 3H, NH<sub>3</sub><sup>+</sup>), 7.83 (app d, 2H, J = 8.7 Hz, 3-H), 7.38 (app d, 2H, J = 8.4 Hz, 3'-H), 7.23 (app d, 2H, J = 8.7 Hz, 2-H), 7.16 (app d, 2H, J = 8.4 Hz, 2'-H), 3.89 (s, 2H, CH<sub>2</sub>)  $\delta_{\text{C}}$  (500 MHz, DMSO) 167.3 (C=O), 142.7 (1-C), 138.4 (1'-C), 131.0 (3-C), 130.3 (3'-C), 128.9 (4'-C), 124.9 (4-C), 118.3 (2'-C), 116.9 (2-C), 68.8 (CH<sub>2</sub>)  $\nu_{\text{max}}/\text{cm}^{-1}$  (ATR); 3242, 1694, 1606, 1430, 1287, 1114. *m/z* (ESI): (Found M-H, 320.0715 C<sub>14</sub>H<sub>15</sub>N<sub>3</sub>O<sub>4</sub>S requires M-H, 320.0705) HPLC (RT: 1.12 min(100 % relative area)).

## 6.3.4 Synthesis of Heterocyclic Fragments

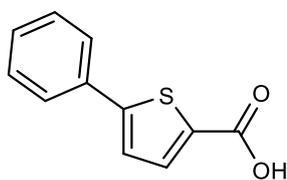
### Synthesis of methyl 5-phenylthiophene-2-carboxylate (**129**) [91]



To a RBF charged with methyl 5-bromo-2-thiophenecarboxylate (0.200 g, 0.90 mmol), phenylboronic acid (0.122 g, 1.00 mmol) and Pd(PPh<sub>3</sub>)<sub>4</sub> (0.070 g, 0.045 mmol) was added DMF (10 mL) and 2M Na<sub>2</sub>CO<sub>3</sub> (1.35 mL, 2.70 mmol). The reaction mixture was then stirred at 80 °C for 18 hrs. The reaction mixture was then cooled to RT and diluted with EtOAc (30mL). The resultant solutions was then washed with; water (2 x 30 mL), sat. Na<sub>2</sub>HCO<sub>3</sub> (2 x 30 mL) and brine (30 mL). The organic layer was then dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude orange solid. The crude solid was then purified column chromatography (Eluent; DCM/Petrol 25:75) to give the tile compound as a colourless

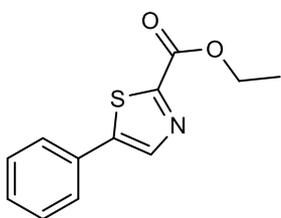
crystalline solid (0.070 g, 35%). Experimental data matched that reported in [91]. m.p: 97-98°C.  $R_f$  0.95 (20% EtOAc-Petrol).  $\delta_H$  (500 MHz, DMSO) 7.82 (d, 1H,  $J = 4.0$  Hz, 3-H), 7.78-7.75 (m, 2H, 2'-H), 7.63 (d, 1H,  $J = 4.0$  Hz, 4-H), 7.49-7.47 (m, 2H, 3'-H), 7.43-7.40 (m, 1H, 4'-H), 3.85 (s, 3H, OCH<sub>3</sub>)  $\delta_C$  (500 MHz, DMSO) 162.2 (C=O), 150.8 (5-C), 135.4 (3-C), 133.0 (1'-C), 131.7 (2-C), 129.8 (3'-C), 129.6 (4'-C), 126.4 (2'-C), 125.2 (4-C), 52.8 (CH<sub>3</sub>)  $\nu_{max}$  /  $cm^{-1}$  (solid): 2931, 1707, 1551, 1431, 1258, 1224.  $m/z$  (ESI) (100%, M+H); (Found M+H, 219.0472. C<sub>12</sub>H<sub>10</sub>O<sub>2</sub>S requires M+H, 219.0479).

### Synthesis of 5-phenylthiophene-2-carboxylic acid (122) [92]



To a solution of methyl 5-phenylthiophene-2-carboxylate **129** (0.050 g, 0.23 mmol) in MeOH (5 mL) and H<sub>2</sub>O (1 mL) was added LiOH·H<sub>2</sub>O (0.038 g, 0.92 mmol). The reaction mixture was then stirred at RT for 18 hrs. The reaction mixture was then concentrated and acidified with 2M HCl and the resultant ppt was collected by vacuum filtration to afford the title compound as a colourless solid (0.015 g, 32%).  $R_f$  0.55 (5% MeOH-DCM).  $\delta_H$  (500 MHz, DMSO) 7.75-7.72 (m, 3H, 3-H and 2'-H), 7.58 (d, 1H,  $J = 3.8$  Hz, 4-H), 7.46 (app t, 2H,  $J = 7.5$  Hz, 3'-H), 7.40 (app t, 1H,  $J = 7.5$  Hz, 4'-H),  $\delta_C$  (500 MHz, DMSO) 163.3 (C=O), 150.2 (5-C), 134.7 (3-C), 133.9 (1'-C), 133.3 (2-C), 129.8 (3'-C), 129.6 (4'-C), 126.1 (2'-C), 125.2 (4-C)  $\nu_{max}$  /  $cm^{-1}$  (solid): 2914, 2552, 1647, 1517, 1448, 1293.  $m/z$  (ESI) (100%, M+H); (Found M+H, 205.0317. C<sub>11</sub>H<sub>8</sub>O<sub>2</sub>S requires M+H, 205.0323). HPLC (RT: 2.81 min(100 % relative area)).

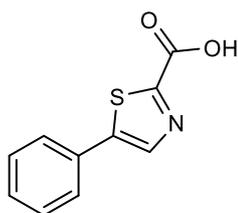
### Synthesis of ethyl 5-phenyl-1,3-thiazole-2-carboxylate (134) [82]



To a stirred solution of 2-aminoacetophenone hydrochloride (1.00 g, 5.83 mmol) in DCM (20 mL) was added triethylamine (2.40 g, 17.49 mmol) and ethyl oxalyl chloride (0.65 mL, 5.83 mmol) at 0°C. The reaction mixture was then warmed to RT and stirred for 18 hours. The reaction mixture was then quenched with water (45 mL) and the resultant solution was extracted with ethyl acetate (2 x 15 mL). The combined organic layers were then washed with water (3 x 30 mL) and brine

(30 mL). The organic layer was dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude orange residue. The crude residue was dissolved in chloroform (6 mL) and phosphorus decasulfide (0.95 g, 4.16 mmol) was added. The reaction mixture was then heated at reflux for 18 hours. A further amount of phosphorus decasulfide (0.475 g, 2.08 mmol) was then added and the reaction mixture was refluxed for a further 3 hours. The reaction mixture was then cooled to RT and quenched with water (20 mL). The resultant solution was then extracted with chloroform (2 x 10 mL). The organic layers were then combined and washed with water (20 mL) and brine (20 mL). The organic layer was then dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude orange residue. The crude residue was purified using column chromatography (15% EtOAc-Petrol) to give the title compound as a crystalline yellow solid (0.489 g, 36%). R<sub>f</sub> 0.87 (15% EtOAc-Petrol). m.p: 49.8-51.8°C δ<sub>H</sub> (400 MHz, DMSO), 8.52 (s, 1H, 4-H), 7.81 (d, 2H, J = 7.0 Hz, 2'-H), 7.51-7.43 (m, 3H, 3'-H and 4'-H), 4.40 (q, 2H, J = 7.1 Hz, OCH<sub>2</sub>), 1.34 (t, 3H, J = 7.1 Hz, CH<sub>3</sub>) δ<sub>C</sub> (400 MHz, DMSO), 159.8 (2-C), 156.1 (C=O), 145.42 (4-C), 141.58 (1'-C), 130.2 (5-C), 130.1 (4'-C), 130.0 (3'-C), 127.5 (2'-C), 62.7 (OCH<sub>2</sub>), 14.54 (CH<sub>3</sub>) ν<sub>max</sub> / cm<sup>-1</sup> (solid): 3077, 2981, 1700, 1412, 1291 *m/z* (ESI):(molecular ion not observed)

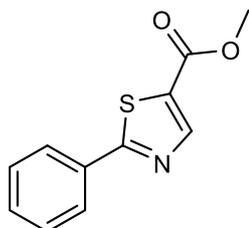
### Synthesis of 5-phenyl-1,3-thiazole-2-carboxylic acid (**123**) [82]



To a stirred solution of Ethyl 5-phenyl-1,3-thiazole-2-carboxylate **134** (0.170 g, 0.73 mmol) in methanol (5 mL) and water (2 mL) was added lithium hydroxide (0.120 g, 2.92 mmol). The reaction mixture was stirred at room temperature for 18 hours. A further amount of LiOH (0.120 g, 2.92 mmol) was then added and the reaction mixture was heated to reflux and stirred at temperature for two hours. The reaction mixture was then cooled to RT and concentrated in vacuo. The resultant residue was dissolved in water (3 mL) and 2M HCl was added dropwise. The resultant ppt was then collected by vacuum filtration and washed with water (5 mL) and DCM (5 mL) to give the title compound as a yellow solid. R<sub>f</sub> 0.48 (2% MeOH-DCM) δ<sub>H</sub> (400 MHz, DMSO), 8.44 (s, 1H, 4-H), 7.79 (d, 2H, J = 1.4 Hz, 2'-H), 7.51-7.43 (m, 3H, 3'-H and 4'-H) δ<sub>C</sub> (400 MHz, DMSO), 161.4 (2-

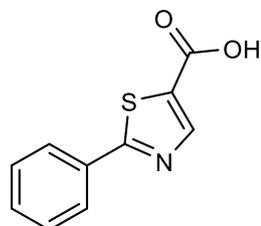
C), 154.1 (C=O), 141.1 (4-C), 139.9 (1'-C), 130.6 (5-C), 129.9 (3'-C), 129.8 (4'-C), 127.4 (2'-C)  $\nu_{\text{max}} / \text{cm}^{-1}$  (solid): 2379, 1707, 1629, 1362, 1259 1224.  $m/z$  (ESI) (100%, M+H); (Found M+H, 206.0267.  $\text{C}_{10}\text{H}_7\text{NO}_2\text{S}$  requires M+H, 206.0275). HPLC (RT: 2.40 min (100 % relative area)).

### Synthesis of methyl 2-phenyl-1,3-thiazole-5-carboxylate (130) [93]



To a stirred suspension of methyl 2-bromothiazole-5-carboxylate (0.500 g, 2.3 mmol), phenylboronic acid (410 mg, 3.4 mmol) and  $\text{Pd}(\text{PPh}_3)_4$  (260 mg, 0.23 mmol) in DMF (10 mL) was added an aqueous 2M solution of  $\text{Na}_2\text{CO}_3$  (4.7 mL, 9.5 mmol) and the reaction mixture was stirred at 85 °C for 18 h. The reaction mixture was then cooled to RT and diluted with EtOAc (20 mL). The resultant solution was then filtered through Celite and washed with;  $\text{H}_2\text{O}$  (20 mL) 1M NaOH (20 mL) and brine (20 mL). The organic layer was then dried over  $\text{MgSO}_4$ , filtered and concentrated to give a crude black residue. The crude product was purified using column chromatography (eluent; 10% EtOAc in Petroleum ether) to afford the title compound as a white solid (0.450 g, 89%).  $R_f$  0.87 (15% EtOAc-Petrol).  $\delta_{\text{H}}$  (500 MHz, DMSO), 8.52 (s, 1H, 4-H), 8.05-8.02 (m, 2H, 2'-H), 7.63-7.59 (m, 3H, 3'-H and 4'-H), 3.76 (s, 3H,  $\text{OCH}_3$ ).  $\delta_{\text{C}}$  (500 MHz, DMSO) 172.9 (C=O), 161.6 (2-C), 149.8 (4-C), 132.9 (5-C), 132.1 (1'-C), 129.9 (3'-C), 128.8 (4'-C), 127.2 (2'-C), 53.1 ( $\text{OCH}_3$ )  $\nu_{\text{max}} / \text{cm}^{-1}$  (solid): 2944, 1705, 1247, 1091.  $m/z$  (ESI) (100%, M+H); (Found M+H, 220.0423.  $\text{C}_{11}\text{H}_9\text{NO}_2\text{S}$  requires M+H, 220.0427) HPLC (RT: 3.18 min (100 % relative area)).

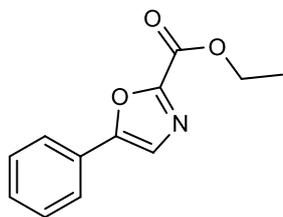
### Synthesis of 2-phenyl-1,3-thiazole-5-carboxylic acid (124) [94]



To a stirred solution of ethyl 2-phenyl-1,3-oxazole-5-carboxylate **130** (0.100 g, 0.46 mmol) in methanol (5 mL) and water (1 mL) was added  $\text{LiOH}\cdot\text{H}_2\text{O}$  (0.08 g, 1.84 mmol). The reaction mixture was then refluxed for 18 hours. The reaction mixture was then cooled to RT and the solution was then concentrated in vacuo. The resultant residue was then taken up in water (3 mL) and 2M HCl was added. The resultant ppt was collected by vacuum filtration and the solid was then washed with water (5

mL) and DCM (5 mL) to afford the title compound as a colourless solid (0.057 g, 61%).  $R_f$  0.39 (10% MeOH-DCM)  $\delta_H$  (400 MHz, DMSO) 8.30 (s, 1H, 4-H), 7.98-7.97 (m, 2H, 2'-H), 7.50-7.56 (m, 3H, 3'-H and 4'-H).  $\delta_C$  (400 MHz, DMSO) 172.7 (C=O), 163.4 (2-C), 147.6 (4-C), 132.9 (5-C), 132.6 (1'-C), 130.9 (3'-C), 128.9 (4'-C), 126.4 (2'-C).  $\nu_{max}$  /  $cm^{-1}$  (ATR); 3037, 2872, 1653, 1244, 1095.  $m/z$  (ESI) (100%, M+H); (Found M+H, 206.0221.  $C_{10}H_7NO_2S$  requires M+H, 206.0197). HPLC (RT: 2.39 min(100 % relative area)).

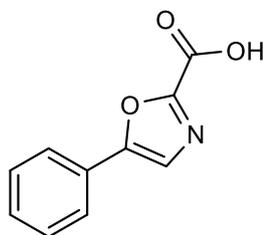
### Synthesis of ethyl 5-phenyl-1,3-oxazole-2-carboxylate (135) [81]



To a stirred solution of 2-aminoacetophenone hydrochloride (1.00 g, 5.83 mmol) in DCM (20 mL) was added triethylamine (2.40 g, 17.49 mmol) and ethyl oxalyl chloride (0.65 mL, 5.83 mmol) in DCM (2 mL), at 0°C. The reaction mixture was then warmed to RT and stirred for 18 hours. The reaction mixture was then quenched with water (45 mL). The resultant solution was then extracted with ethyl acetate (2 x 15 mL). The combined organic layers were then washed with water (3 x 30 mL) and brine (30 mL). The organic layer was dried over  $MgSO_4$ , filtered and the solvent removed to give a crude orange residue. The crude residue was dissolved in toluene (10 mL) and  $POCl_3$  (0.95 g, 4.16 mmol) was added. The reaction mixture was heated at reflux for 18 hours. An additional amount of  $POCl_3$  (0.95 g, 4.16 mmol) was then added and the reaction mixture was refluxed for an additional 18 hours. The reaction mixture was then cooled to 0 °C and water (20 mL) was slowly added. The resultant solution was then warmed to RT and stirred for 30 mins. The solution was then extracted using EtOAc (3 x 20 mL). The organic layers were then combined, washed with water (20 mL) and brine (20 mL). The organic layer was then dried over  $MgSO_4$ , filtered and the solvent removed give a crude orange oil. The crude product was purified using column chromatography (10% EtOAc:Petrol) to afford the title compound as a white solid. (0.521 g, 41%).  $R_f$  0.92 (20% EtOAc-Petrol)  $\delta_H$  (500 MHz, DMSO) 8.00 (s, 1H, 4-H), 7.83-7.80 (m, 2H, 2'-H), 7.56-7.52 (m, 2H, 3'-H), 7.49-7.46 (m, 1H, 4'-H), 4.40 (q, 2H,  $J = 6.8$  Hz,  $OCH_2CH_3$ ), 1.35 (t, 3H,  $J = 6.8$  Hz,  $OCH_2CH_3$ )  $\delta_C$  (500 MHz, DMSO) 155.6 (C=O), 153.7 (2-C), 151.7 (5-C), 130.3 (1'-C), 129.8 (4'-C), 126.8 (3'-C),

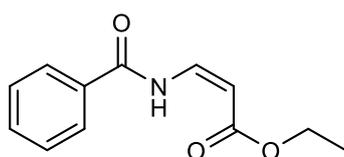
125.3 (5-C), 125.2 (2'-C), 62.6 (OCH<sub>2</sub>), 14.5 (CH<sub>3</sub>).  $\nu_{\max}$  / cm<sup>-1</sup> (solid): 2928, 1722, 1445, 1173  $m/z$  (ESI) (100%, M+Na); (Found M+Na, 240.0639. C<sub>12</sub>H<sub>11</sub>NO<sub>3</sub> requires M+Na, 240.0631).

### Synthesis of 5-phenyl-1,3-oxazole-2-carboxylic acid (**125**) [81]



To a stirred solution of ethyl 2-phenyl-1,3-oxazole-5-carboxylate **135** (0.100 g, 0.46 mmol) in methanol (5 mL) and water (1 mL) was added LiOH·H<sub>2</sub>O (0.08 g, 1.84 mmol). The reaction mixture was then refluxed for 18 hours. The reaction mixture was then cooled to RT and concentrated in vacuo. The resultant residue was dissolved in water (3 mL) and 2M HCl was added dropwise. The resultant ppt was collected by vacuum filtration and washed with water (5 mL) and DCM (5 mL) to afford the title compound as a white solid (0.075 g, 86%).  $\delta_{\text{H}}$  (500 MHz, DMSO) 7.87 (s, 1H, 4-H), 7.74-7.72 (m, 2H, 2'-H), 7.46-7.43 (m, 2H, 3'-H), 7.39-7.36 (m, 1H, 4'-H),  $\delta_{\text{C}}$  (500 MHz, DMSO) 157.0 (C=O), 153.4 (2-C), 152.9 (5-C), 130.1 (1'-C), 129.8 (4'-C), 127.1 (3'-C), 125.2 (5-C), 125.1 (2'-C).  $\nu_{\max}$  / cm<sup>-1</sup> (solid): 3130, 2335, 1753, 1450, 1164  $m/z$  (ESI) (100%, M-H); (Found M-H, 188.0345. C<sub>10</sub>H<sub>7</sub>NO<sub>3</sub> requires M-H, 188.0348). HPLC (RT: 2.63 min(100 % relative area)).

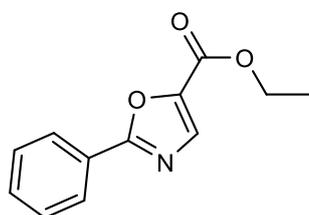
### Synthesis of ethyl (2Z)-3-(phenylformamido)prop-2-enoate (**140**) [84]



To a stirred solution of benzamide (1.20 g, 10.00 mmol) in toluene (50 mL) was added benzoquinone (1.80 g, 16.60 mmol), p-Toluenesulfonic acid (1.00 g, 5.30 mmol) and palladium (II) acetate (0.200 g, 0.90 mmol) at RT. The reaction mixture was then stirred for 5 minutes at RT before methyl acrylate (12.50 mL, 115.00 mmol) was added. The reaction mixture was then stirred at RT for 72 hours. The reaction mixture was then filtered through Celite and the filtrate was washed with water (30 mL) and brine (30 mL). The organic layer was dried over MgSO<sub>4</sub>, filtered and the solvent removed to give a crude yellow residue. The crude residue was purified using column chromatography (eluent 8:1 hexane: EtOAc) to afford the title compound as a colourless crystalline solid (0.635 g, 29%). m.p: 73-

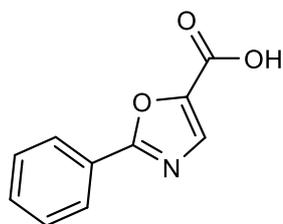
75 °C.  $R_f$  0.71 (15% EtOAc-Petrol).  $\delta_H$  (400 MHz, DMSO), 7.90 (app d, 2H,  $J = 7.8$  Hz, 2'-H), 7.75-7.68 (m, 2H, 4'-H and 2-H), 7.62 (app t, 2H,  $J = 7.8$  Hz, 3'-H), 5.34 (d, 1H,  $J = 8.8$  Hz, 3-H), 4.20 (q, 2H,  $J = 7.1$  Hz, 5-H), 1.26 (t, 3H,  $J = 7.1$  Hz, 6-H)  $\delta_C$  (400 MHz, DMSO), 169.3 (C=O), 164.1 (C=O), 139.3 (2-C), 133.7 (1'-C), 132.3 (4'-C), 129.7 (3'-C), 127.8 (2'-C), 97.6 (3-C), 60.6 (5-C), 14.6 (6-C)  $\nu_{max}$  /  $cm^{-1}$  (solid): 3293, 2967, 1655, 1595, 1457, 1362, 1207  $m/z$  (ESI) (100%, M+Na); (Found M+Na, 242.0786.  $C_{12}H_{13}NO_3$  requires M+Na, 242.0793).

### Synthesis of ethyl 2-phenyl-1,3-oxazole-5-carboxylate (139) [84]



To a stirred solution of ethyl (2Z)-3-(phenylformamido)prop-2-enoate **140** (0.250 g, 1.25 mmol) in TFE (15 mL) was added PIFA (0.590 g, 1.35 mmol). The reaction mixture was stirred at room temperature for 18 hours. The solvent was then removed to give a crude residue. The crude residue was purified using column chromatography (10% EtOAc:Petrol) to afford the title compound as a crystalline colourless solid (0.210 g, 77%). m.p: 68-69°C.  $R_f$  0.80 (10% EtOAc-Petrol).  $\delta_H$  (400 MHz, DMSO), 8.11 (s, 1H, 4-H), 8.10-8.05 (m, 2H, 2'-H), 7.63-7.57 (m, 3H, 3'-H and 4'-H), 4.36 (q, 2H,  $J = 7.1$  Hz, OCH<sub>2</sub>), 1.34 (t, 3H,  $J = 7.1$  Hz, CH<sub>3</sub>).  $\delta_C$  (400 MHz, DMSO), 163.7 (2-C), 157.6 (C=O), 142.4 (5-C), 136.0 (4-C), 132.4 (4'-C), 129.9 (3'-C), 127.2 (1'-C), 126.3 (2'-C), 61.7 (OCH<sub>2</sub>), 14.6 (CH<sub>3</sub>)  $\nu_{max}$  /  $cm^{-1}$  (solid): 2966, 1707, 1569, 1517, 1466, 1207, 1138.  $m/z$  (ESI) (100%, M+Na); (Found M+Na, 240.0627.  $C_{12}H_{11}NO_3$  requires M+Na, 240.0637).

### Synthesis of 2-phenyl-1,3-oxazole-5-carboxylic acid (126) [84]



To a stirred solution of ethyl 2-phenyl-1,3-oxazole-5-carboxylate **139** (0.100 g, 0.46 mmol) in methanol (5 mL) was added a solution of lithium hydroxide (0.08 g, 1.84 mmol) in water (1 mL). The reaction mixture was then refluxed for 18 hours. An additional amount of LiOH (0.080g, 1.84 mmol) was then added and the reaction mixture was refluxed for an additional 4 hours. The reaction mixture was then cooled to

RT and concentrated in vacuo. The resultant residue was dissolved in water (3 mL) and 2M HCl was added. The resultant ppt was collected by vacuum filtration and the solid was washed with water (5 mL) and DCM (5 mL) to afford the title compound as a white solid (0.082 g, 93%).  $R_f$  0.51 (10% MeOH-DCM)  $\delta_H$  (400 MHz, DMSO), 8.06-8.04 (m, 2H, 2'-H), 8.02 (app s, 1H, 4-H), 7.61-7.57 (m, 3H, 3'-H and 4'-H)  $\delta_C$  (400 MHz, DMSO), 163.4 (C=O), 158.9 (2-C), 143.4 (5-C), 135.5 (4-C), 132.3 (4'-C), 129.8 (3'-C), 127.1 (1'-C), 126.5 (2'-C)  $\nu_{max}$  /  $cm^{-1}$  (solid): 2810, 2509, 1724, 1517, 1414, 1293, 1138.  $m/z$  (ESI) (100%, 2M+H); (Found 2M+H, 383.0865.  $C_{10}H_7NO_3$  requires 2M+H, 383.0865). HPLC (RT: 2.03 min(100 % relative area)).

## Chapter Seven - References

1. Cox, F.E.G., *A Textbook of Parasitology*. 1982: Blackwell Scientific Publications.
2. Flegr, J., et al., *Toxoplasmosis--a global threat. Correlation of latent toxoplasmosis with specific disease burden in a set of 88 countries*. PLoS One, 2014. **9**(3): p. e90203.
3. Wang, Z.D., et al., *Toxoplasma gondii Infection in Immunocompromised Patients: A Systematic Review and Meta-Analysis*. Front Microbiol, 2017. **8**: p. 389.
4. Park, Y.H. and H.W. Nam, *Clinical features and treatment of ocular toxoplasmosis*. Korean J Parasitol, 2013. **51**(4): p. 393-399.
5. Luft, B.J., et al., *Toxoplasmic encephalitis in patients with the acquired immunodeficiency syndrome. Members of the ACTG 077p/ANRS 009 Study Team*. N Engl J Med, 1993. **329**(14): p. 995-1000.
6. Basavaraju, A., *Toxoplasmosis in HIV infection: An overview*. Trop Parasitol, 2016. **6**(2): p. 129-135.
7. Black, M.W. and J.C. Boothroyd, *Lytic cycle of Toxoplasma gondii*. Microbiol Mol Biol Rev, 2000. **64**(3): p. 607-623.
8. Robert-Gangneux, F. and M.L. Darde, *Epidemiology of and diagnostic strategies for toxoplasmosis*. Clin Microbiol Rev, 2012. **25**(2): p. 264-296.
9. Palencia, A., et al., *Targeting Toxoplasma gondii CPSF3 as a new approach to control toxoplasmosis*. EMBO Mol Med, 2017. **9**(3): p. 385-394.
10. Alday, P.H. and J.S. Doggett, *Drugs in development for toxoplasmosis: advances, challenges, and current status*. Drug Des Devel Ther, 2017. **11**: p. 273-293.
11. Schoondermark-van de Ven, E., et al., *In vitro effects of sulfadiazine and its metabolites alone and in combination with pyrimethamine on*

- Toxoplasma gondii*. Antimicrob Agents Chemother, 1995. **39**(3): p. 763-765.
12. Rolston, K.V., *Treatment of acute toxoplasmosis with oral clindamycin*. Eur J Clin Microbiol Infect Dis, 1991. **10**(3): p. 181-183.
  13. Camps, M., G. Arrizabalaga, and J. Boothroyd, *An rRNA mutation identifies the apicoplast as the target for clindamycin in Toxoplasma gondii*. Mol Microbiol, 2002. **43**(5): p. 1309-1318.
  14. Katlama, C., et al., *Pyrimethamine-clindamycin vs. pyrimethamine-sulfadiazine as acute and long-term therapy for toxoplasmic encephalitis in patients with AIDS*. Clin Infect Dis, 1996. **22**(2): p. 268-275.
  15. Porter, S.B. and M.A. Sande, *Toxoplasmosis of the central nervous system in the acquired immunodeficiency syndrome*. N Engl J Med, 1992. **327**(23): p. 1643-1648.
  16. Van Delden, C. and B. Hirschel, *Folinic acid supplements to pyrimethamine-sulfadiazine for Toxoplasma encephalitis are associated with better outcome*. J Infect Dis, 1996. **173**(5): p. 1294-1295.
  17. de la Hoz Caballer, B., et al., *Management of sulfadiazine allergy in patients with acquired immunodeficiency syndrome*. J Allergy Clin Immunol, 1991. **88**(1): p. 137-138.
  18. Dannemann, B., et al., *Treatment of toxoplasmic encephalitis in patients with AIDS. A randomized trial comparing pyrimethamine plus clindamycin to pyrimethamine plus sulfadiazine. The California Collaborative Treatment Group*. Ann Intern Med, 1992. **116**(1): p. 33-43.
  19. Meneceur, P., et al., *In vitro susceptibility of various genotypic strains of Toxoplasma gondii to pyrimethamine, sulfadiazine, and atovaquone*. Antimicrob Agents Chemother, 2008. **52**(4): p. 1269-1277.

20. WHO. *World Malaria Report 2017*. 2017 [cited 22/07/2018; Available from: <http://apps.who.int/iris/bitstream/handle/10665/259492/9789241565523-eng.pdf;jsessionid=FAEE8F1128F0A46B660ED0B182F41982?sequence=1>].
21. Bartoloni, A. and L. Zammarchi, *Clinical aspects of uncomplicated and severe malaria*. *Mediterr J Hematol Infect Dis*, 2012. **4**(1): p. e2012026.
22. Wells, T.N., P.L. Alonso, and W.E. Gutteridge, *New medicines to improve control and contribute to the eradication of malaria*. *Nat Rev Drug Discov*, 2009. **8**(11): p. 879-891.
23. Meunier, B. and A. Robert, *Heme as trigger and target for trioxane-containing antimalarial drugs*. *Acc Chem Res*, 2010. **43**(11): p. 1444-1451.
24. Fry, M. and M. Pudney, *Site of action of the antimalarial hydroxynaphthoquinone, 2-[trans-4-(4'-chlorophenyl) cyclohexyl]-3-hydroxy-1,4-naphthoquinone (566C80)*. *Biochem Pharmacol*, 1992. **43**(7): p. 1545-1553.
25. White, N.J., *Antimalarial drug resistance*. *J Clin Invest*, 2004. **113**(8): p. 1084-1092.
26. Noedl, H., D. Socheat, and W. Satimai, *Artemisinin-resistant malaria in Asia*. *N Engl J Med*, 2009. **361**(5): p. 540-541.
27. Arieu, F., et al., *A molecular marker of artemisinin-resistant Plasmodium falciparum malaria*. *Nature*, 2014. **505**(7481): p. 50-53.
28. Rieckmann, K.H., D.R. Davis, and D.C. Hutton, *Plasmodium vivax resistance to chloroquine?* *Lancet*, 1989. **2**(8673): p. 1183-1184.
29. Martin, R.E., et al., *Chloroquine transport via the malaria parasite's chloroquine resistance transporter*. *Science*, 2009. **325**(5948): p. 1680-1682.

30. Docampo, R. and S.N. Moreno, *Acidocalcisomes*. Cell Calcium, 2011. **50**(2): p. 113-119.
31. Docampo, R., et al., *Acidocalcisomes - conserved from bacteria to man*. Nat Rev Microbiol, 2005. **3**(3): p. 251-261.
32. Pick, U. and M. Weiss, *Polyphosphate Hydrolysis within Acidic Vacuoles in Response to Amine-Induced Alkaline Stress in the Halotolerant Alga Dunaliella salina*. Plant Physiol, 1991. **97**(3): p. 1234-1240.
33. Ruiz, F.A., C.O. Rodrigues, and R. Docampo, *Rapid changes in polyphosphate content within acidocalcisomes in response to cell growth, differentiation, and environmental stress in Trypanosoma cruzi*. J Biol Chem, 2001. **276**(28): p. 26114-26121.
34. Montalvetti, A., P. Rohloff, and R. Docampo, *A functional aquaporin co-localizes with the vacuolar proton pyrophosphatase to acidocalcisomes and the contractile vacuole complex of Trypanosoma cruzi*. J Biol Chem, 2004. **279**(37): p. 38673-38682.
35. Docampo, R., et al., *New insights into roles of acidocalcisomes and contractile vacuole complex in osmoregulation in protists*. Int Rev Cell Mol Biol, 2013. **305**: p. 69-113.
36. Rohloff, P. and R. Docampo, *Ammonium production during hypo-osmotic stress leads to alkalization of acidocalcisomes and cytosolic acidification in Trypanosoma cruzi*. Mol Biochem Parasitol, 2006. **150**(2): p. 249-255.
37. Lander, N., et al., *Polyphosphate and acidocalcisomes*. Biochem Soc Trans, 2016. **44**(1): p. 1-6.
38. King-Keller, S., et al., *Chemical validation of phosphodiesterase C as a chemotherapeutic target in Trypanosoma cruzi, the etiological agent of Chagas' disease*. Antimicrob Agents Chemother, 2010. **54**(9): p. 3738-3745.

39. Niyogi, S., et al., *Rab32 is essential for maintaining functional acidocalcisomes, and for growth and infectivity of Trypanosoma cruzi*. J Cell Sci, 2015. **128**(12): p. 2363-2373.
40. Li, Z.H., et al., *Hyperosmotic stress induces aquaporin-dependent cell shrinkage, polyphosphate synthesis, amino acid accumulation, and global gene expression changes in Trypanosoma cruzi*. J Biol Chem, 2011. **286**(51): p. 43959-43971.
41. Miranda, K., et al., *P-type proton ATPases are involved in intracellular calcium and proton uptake in the plant parasite Phytomonas francai*. J Eukaryot Microbiol, 2005. **52**(1): p. 55-60.
42. Huang, G., et al., *Proteomic analysis of the acidocalcisome, an organelle conserved from bacteria to human cells*. PLoS Pathog, 2014. **10**(12): p. e1004555.
43. Kajander, T., J. Kellosoalo, and A. Goldman, *Inorganic pyrophosphatases: one substrate, three mechanisms*. FEBS Lett, 2013. **587**(13): p. 1863-1869.
44. N. R. Shah, K.V., H. Xhaard, A. Goldman, *Integral membrane pyrophosphatases: a novel drug target for human pathogens?* AIMS Biophysics, 2016. **3**(1): p. 171-194.
45. Luoto, H.H., et al., *Na<sup>+</sup>-translocating membrane pyrophosphatases are widespread in the microbial world and evolutionarily precede H<sup>+</sup>-translocating pyrophosphatases*. J Biol Chem, 2011. **286**(24): p. 21633-21642.
46. Lin, S.M., et al., *Crystal structure of a membrane-embedded H<sup>+</sup>-translocating pyrophosphatase*. Nature, 2012. **484**(7394): p. 399-403.
47. Kellosoalo, J., et al., *The structure and catalytic cycle of a sodium-pumping pyrophosphatase*. Science, 2012. **337**(6093): p. 473-476.
48. Maeshima, M., *Vacuolar H(+)-pyrophosphatase*. Biochim Biophys Acta, 2000. **1465**(1-2): p. 37-51.

49. Li, K.M., et al., *Membrane pyrophosphatases from Thermotoga maritima and Vigna radiata suggest a conserved coupling mechanism*. Nat Commun, 2016. **7**: p. 13596.
50. Gaxiola, R.A., M.G. Palmgren, and K. Schumacher, *Plant proton pumps*. FEBS Lett, 2007. **581**(12): p. 2204-2214.
51. Shah, N.R., et al., *Insights into the mechanism of membrane pyrophosphatases by combining experiment and computer simulation*. Struct Dyn, 2017. **4**(3): p. 032105.
52. Lemercier, G., et al., *A vacuolar-type H<sup>+</sup>-pyrophosphatase governs maintenance of functional acidocalcisomes and growth of the insect and mammalian forms of Trypanosoma brucei*. J Biol Chem, 2002. **277**(40): p. 37369-37376.
53. Liu, J., et al., *A vacuolar-H(+) -pyrophosphatase (TgVP1) is required for microneme secretion, host cell invasion, and extracellular survival of Toxoplasma gondii*. Mol Microbiol, 2014. **93**(4): p. 698-712.
54. Martin, M.B., et al., *Bisphosphonates inhibit the growth of Trypanosoma brucei, Trypanosoma cruzi, Leishmania donovani, Toxoplasma gondii, and Plasmodium falciparum: a potential route to chemotherapy*. J Med Chem, 2001. **44**(6): p. 909-916.
55. Simmons, K.J., I. Chopra, and C.W. Fishwick, *Structure-based discovery of antibacterial drugs*. Nat Rev Microbiol, 2010. **8**(7): p. 501-510.
56. Cunningham, F., et al., *An in silico structure-based approach to anti-infective drug discovery*. Parasitology, 2014. **141**(1): p. 17-27.
57. Cain, R., et al., *Applications of structure-based design to antibacterial drug discovery*. Bioorg Chem, 2014. **55**: p. 69-76.
58. *Maestro, version 9.7, Schrödinger, LLC, New York, NY, 2014*. 2012.
59. Li, T., et al., *Structure-activity relationships in a series of C2-substituted gluco-configured tetrahydroimidazopyridines as  $\beta$ -glucosidase inhibitors*. Bioorg Med Chem, 2011. **19**(7): p. 2136-2144.

60. Kutchukian, P.S. and E.I. Shakhnovich, *De novo design: balancing novelty and confined chemical space*. Expert Opin Drug Discov, 2010. **5**(8): p. 789-812.
61. Gillet, V.J., et al., *SPROUT: recent developments in the de novo design of molecules*. J Chem Inf Comput Sci, 1994. **34**(1): p. 207-217.
62. Bedingfield, P.T., et al., *Factors influencing the specificity of inhibitor binding to the human and malaria parasite dihydroorotate dehydrogenases*. J Med Chem, 2012. **55**(12): p. 5841-5850.
63. S.Schrodinger, *Glide, LLC, New York, NY, 2018*. Release 2018-2.
64. *The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC*.
65. Wynne, J.H., et al., *3-Acyloindoles via a one-pot, regioselective Friedel-Crafts reaction*. Synthesis-Stuttgart, 2004(14): p. 2277-2282.
66. Hoegberg, T.U.T.F.O.R.E.K.T., *Preparation of substituted benzamides as CRTH2 receptor ligands*, in *PCT Int. Appl.*, D. 7TM Pharma A/S, Editor. 2005: Denmark.
67. Hughes, J.P., et al., *Principles of early drug discovery*. Br J Pharmacol, 2011. **162**(6): p. 1239-1249.
68. Colomer, I., et al., *A divergent synthetic approach to diverse molecular scaffolds: assessment of lead-likeness using LLAMA, an open-access computational tool*. Chem Commun (Camb), 2016. **52**(45): p. 7209-7212.
69. Waring, M.J., *Lipophilicity in drug discovery*. Expert Opin Drug Discov, 2010. **5**(3): p. 235-248.
70. Miyaura, N.Y., K. Suzuki, A., *A new stereospecific cross-coupling by the palladium-catalyzed reaction of 1-alkenylboranes with 1-alkenyl or 1-alkynyl halides*. Tetrahedron Letters, 1979. **20**(36): p. 3437-3440.
71. Olson, R.E., et al., *Orally active isoxazoline glycoprotein IIb/IIIa antagonists with extended duration of action*. J Med Chem, 1999. **42**(7): p. 1178-1192.

72. Li, L., et al., *Vicinal Diamination of Arenes with Domino Aryne Precursors*. *Org Lett*, 2016. **18**(15): p. 3726-3729.
73. Kelly, D.P., W.J. Spillane, and J. Newell, *Development of structure-taste relationships for monosubstituted phenylsulfamate sweeteners using classification and regression tree (CART) analysis*. *J Agric Food Chem*, 2005. **53**(17): p. 6750-6758.
74. Alcaraz, L., et al., *Novel N-aryl and N-heteroaryl sulfamide synthesis via palladium cross coupling*. *Org Lett*, 2004. **6**(16): p. 2705-2708.
75. Klinger, A.L., et al., *Inhibition of carbonic anhydrase-II by sulfamate and sulfamide groups: an investigation involving direct thermodynamic binding measurements*. *J Med Chem*, 2006. **49**(12): p. 3496-3500.
76. Winum, J.Y., et al., *N-(tert-butoxycarbonyl)-N-[4-(dimethylazaniumylidene)-1,4-dihydropyridin-1-ylsulf onyl]azanide: a new sulfamyolating agent. Structure and reactivity toward amines*. *Org Lett*, 2001. **3**(14): p. 2241-2243.
77. Trogolo, D., et al., *Molecular mechanism of NDMA formation from N,N-dimethylsulfamide during ozonation: quantum chemical insights into a bromide-catalyzed pathway*. *Environ Sci Technol*, 2015. **49**(7): p. 4163-4175.
78. *Calculator Plugins were used for structure property prediction and calculation, Marvin n.n.n (18.10), 201n (2018), ChemAxon.*
79. Aoki, T., et al., *The sulfamide moiety affords higher inhibitory activity and oral bioavailability to a series of coumarin dual selective RAF/MEK inhibitors*. *Bioorg Med Chem Lett*, 2013. **23**(23): p. 6223-6227.
80. Niklas G. Johansson, A.T., Keni Vidilaseris, Daniel Ayuso Pérez, Aaron Wilkinson, Matti Tamminen, Yuezhou Zhang, Alexandros Kiriazis, Evgeni Grazhdankin, Colin Fishwick, Jari Yli-Kauhaluoma, Adrian Goldman, Gustav Boije af Gennäs<sup>1</sup>, Henri Xhaard, *Discovery*

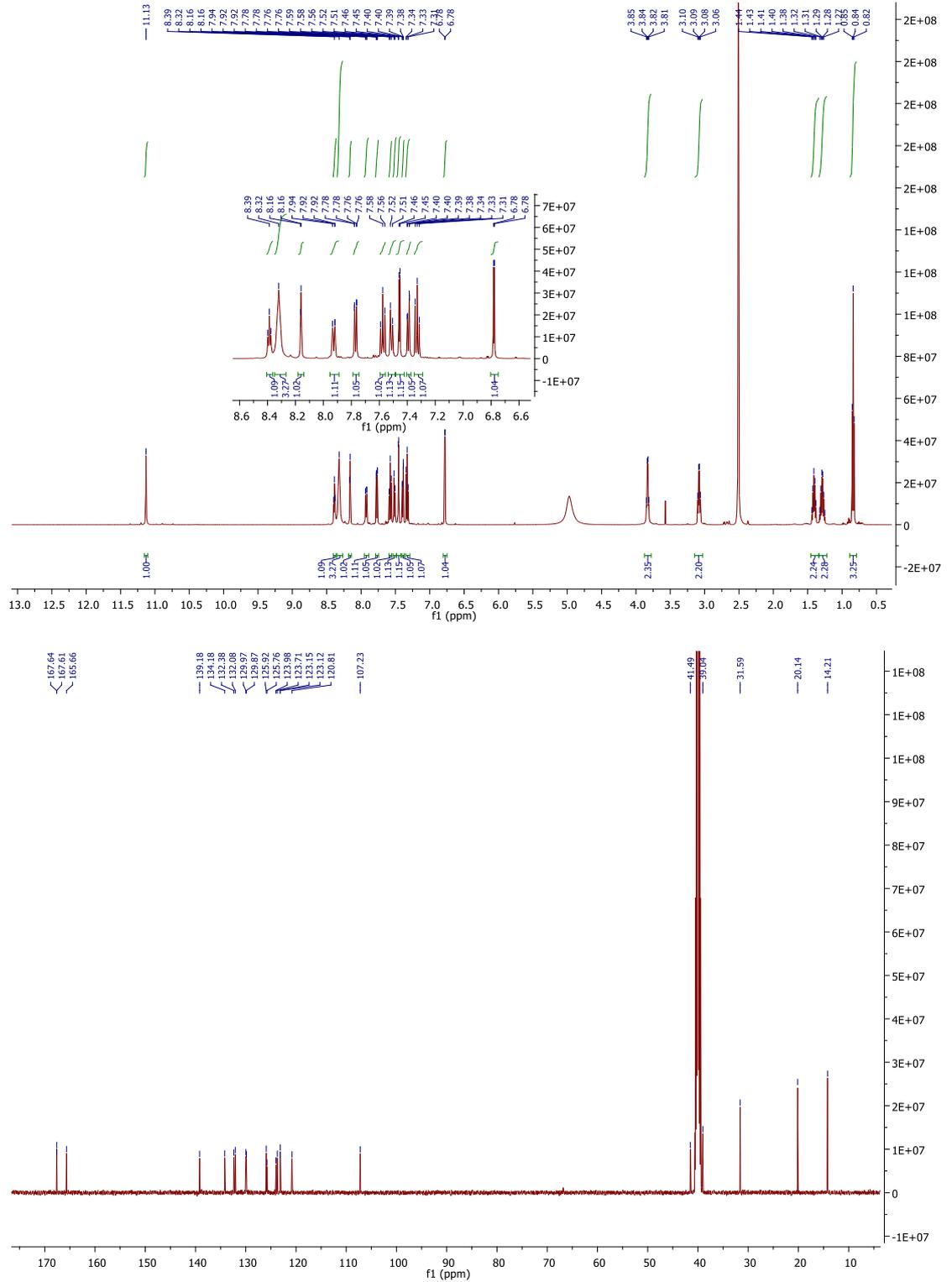
*of membrane-bound pyrophosphatase inhibitors.* J. Med. Chem Manuscript in preparation, 2019.

81. Tanaka, C.N., Keiko; Yamamoto, Noriko; Shibata, Megumi, *Pyrolysis of benzyl 2-oxazolecarbamates and benzyl 4-alkylallophanates.* Chemical & Pharmaceutical Bulletin, 1982. **30**(11): p. 4195-4198.
82. Krasavin, M., et al., *Probing the 'bipolar' nature of the carbonic anhydrase active site: aromatic sulfonamides containing 1,3-oxazol-5-yl moiety as picomolar inhibitors of cytosolic CA I and CA II isoforms.* Eur J Med Chem, 2015. **101**: p. 334-347.
83. Bathula, S.R.R., Muktapuram Prathap; Viswanadham, K. K. Durga Rao; Sathyanarayana, Pochampalli; Reddy, Maddi Sridhar, *Access to Di- and Trisubstituted Oxazoles by NBS-Mediated Oxidative Cyclisation of N-Acyl Amino Acid Derivatives.* European Journal of Organic Chemistry, 2013. **2013**(21): p. 4552-4557.
84. Kamiya, M.S., Motohiro; Tanimori, Shinji, *A rapid access to substituted oxazoles via PIFA-mediated oxidative cyclization of enamides.* Tetrahedron 2017. **73**(9): p. 1247-1254.
85. Walsh, C.T., S.J. Malcolmson, and T.S. Young, *Three ring posttranslational circuses: insertion of oxazoles, thiazoles, and pyridines into protein-derived frameworks.* ACS Chem Biol, 2012. **7**(3): p. 429-442.
86. Kim, S.J., et al., *A structure-activity relationship study on multi-heterocyclic molecules: two linked thiazoles are required for cytotoxic activity.* Medchemcomm, 2013. **4**(2): p. 406-410.
87. Horner, K.E. and P.B. Karadakov, *Shielding in and around Oxazole, Imidazole, and Thiazole: How Does the Second Heteroatom Affect Aromaticity and Bonding?* J Org Chem, 2015. **80**(14): p. 7150-7157.
88. Sabine Schultes, C.d.G., Eric E.J.Haaksma, Iwan J.P.de Esch, Rob Leurs, Oliver Krämer, *Ligand efficiency as a guide in fragment hit selection and optimization.* Drug Discovery Today: Technologies, 2010. **7**(3): p. e157-e162.

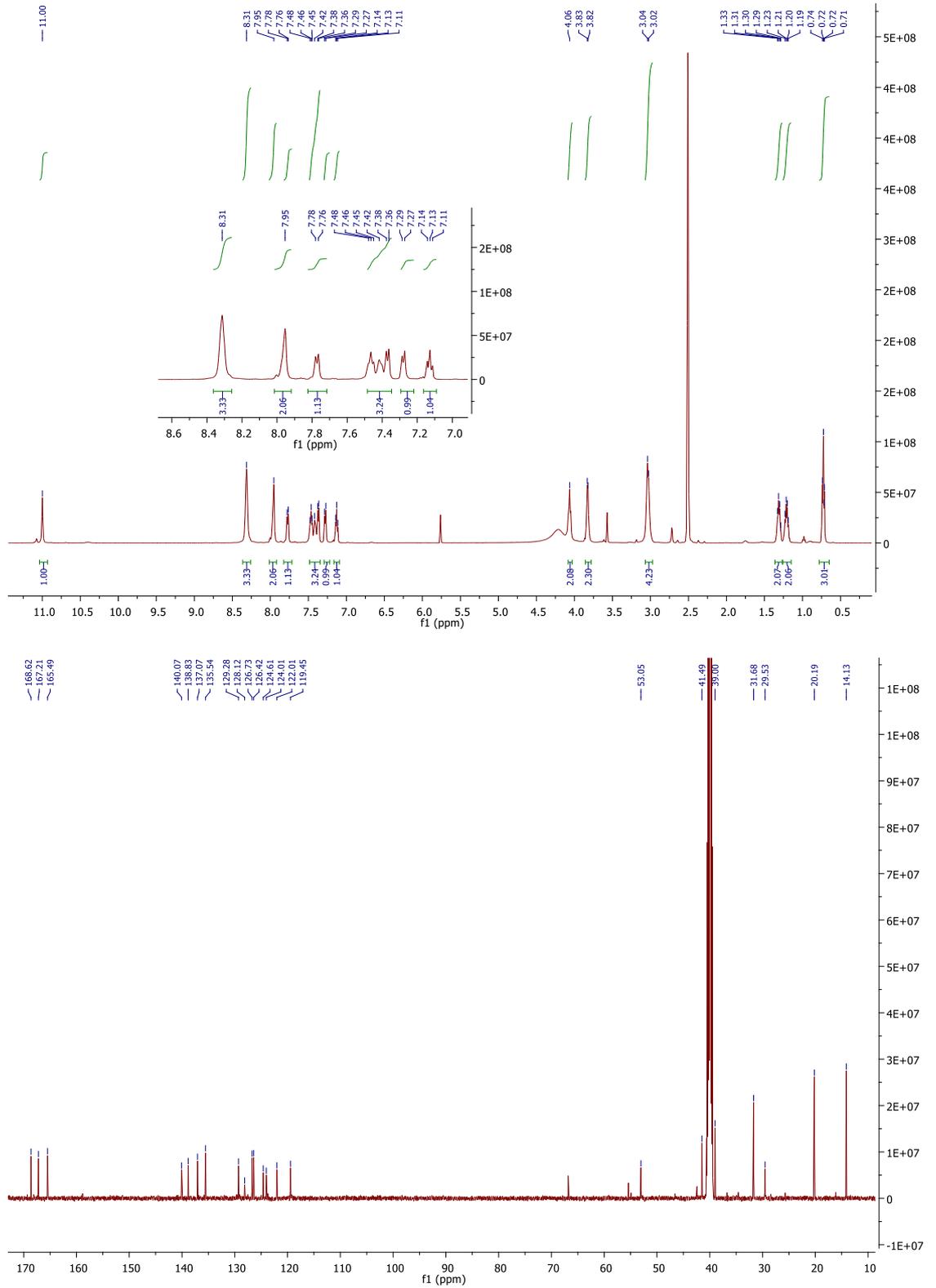
89. Shultz, M.D., *Setting expectations in molecular optimizations: Strengths and limitations of commonly used composite parameters*. *Bioorg Med Chem Lett*, 2013. **23**(21): p. 5980-5991.
90. Vidilaseris, K., J. Kellosalo, and A. Goldman, *A high-throughput method for orthophosphate determination of thermostable membrane-bound pyrophosphatase activity*. *Analytical Methods*, 2018. **10**(6): p. 646-651.
91. McAllister, L.A., Hixon, M.S., Kennedy, J.P., Dickerson, T.J., Janda, K.D, *Superactivation of the botulinum neurotoxin serotype A light chain metalloprotease: a new wrinkle in botulinum neurotoxin*. *J. Am. Chem. Soc.*, 2006, **128**(13): p 4176-4177
92. Zhao, S., *et. al.*, *Design, synthesis and evaluation of aromatic heterocyclic derivatives as potent antifungal agents*. *Eur. J. Med. Chem.*, 2017, **137**: p. 96-107
93. Han, S.J., Repub. Korean Kongkae Taeho Kongbo, *Novel aromatic compound as electroluminescent material for organic optoelectronic device*, *KR 2017016734*, 2017, Repub. Korea.
94. Keenan, M., *Selection and optimization of hits from a high-throughput phenotypic screen against Trypanosoma cruzi*, *Future Med. Chem.*, 2013, **5**(15): p. 1733-1752

### Chapter 8 – Appendix; NMR Spectra of Novel Final Compounds

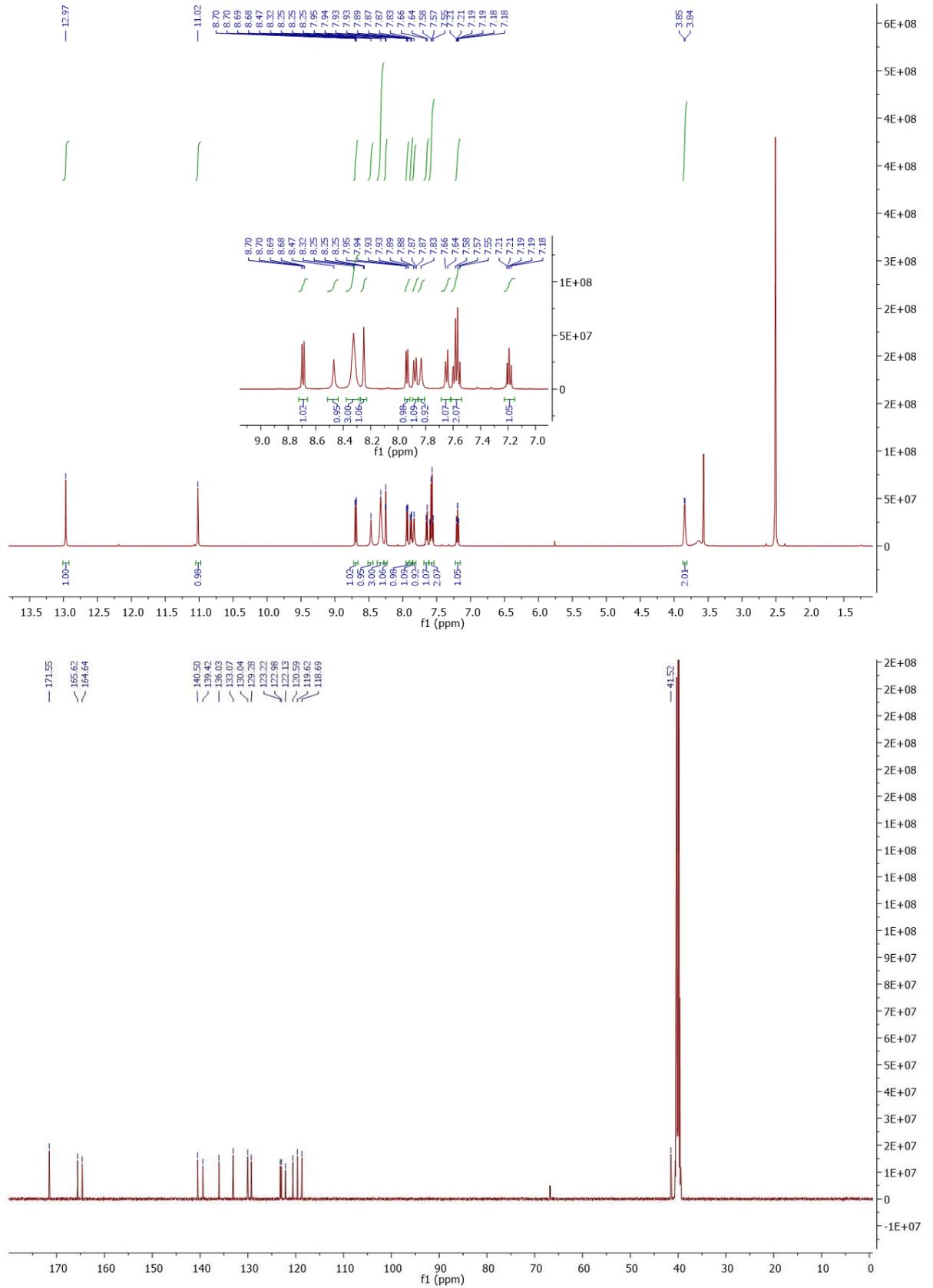
## NMR Spectra of 3-{3-[7-(butylcarbamoyl)-1H-indole-1-carbonyl]phenyl}-2-oxopropan-1-aminium chloride (34)



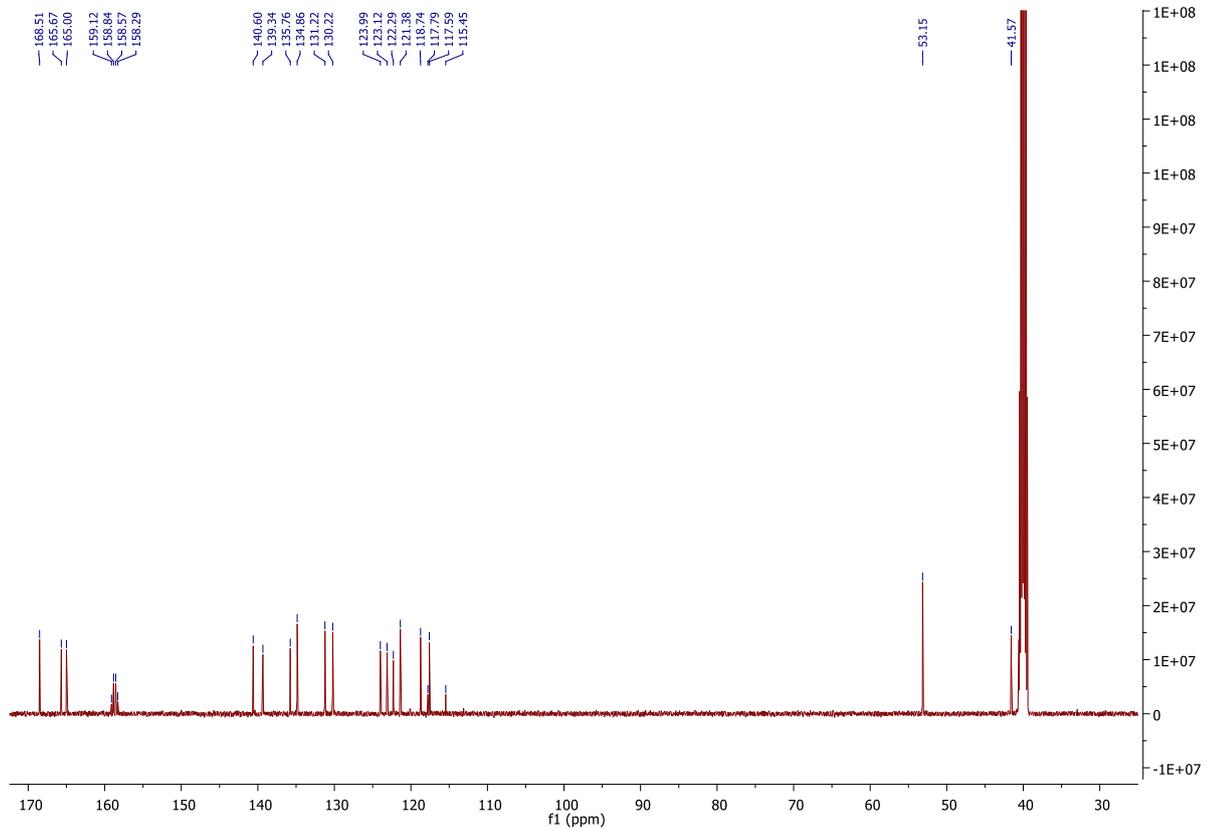
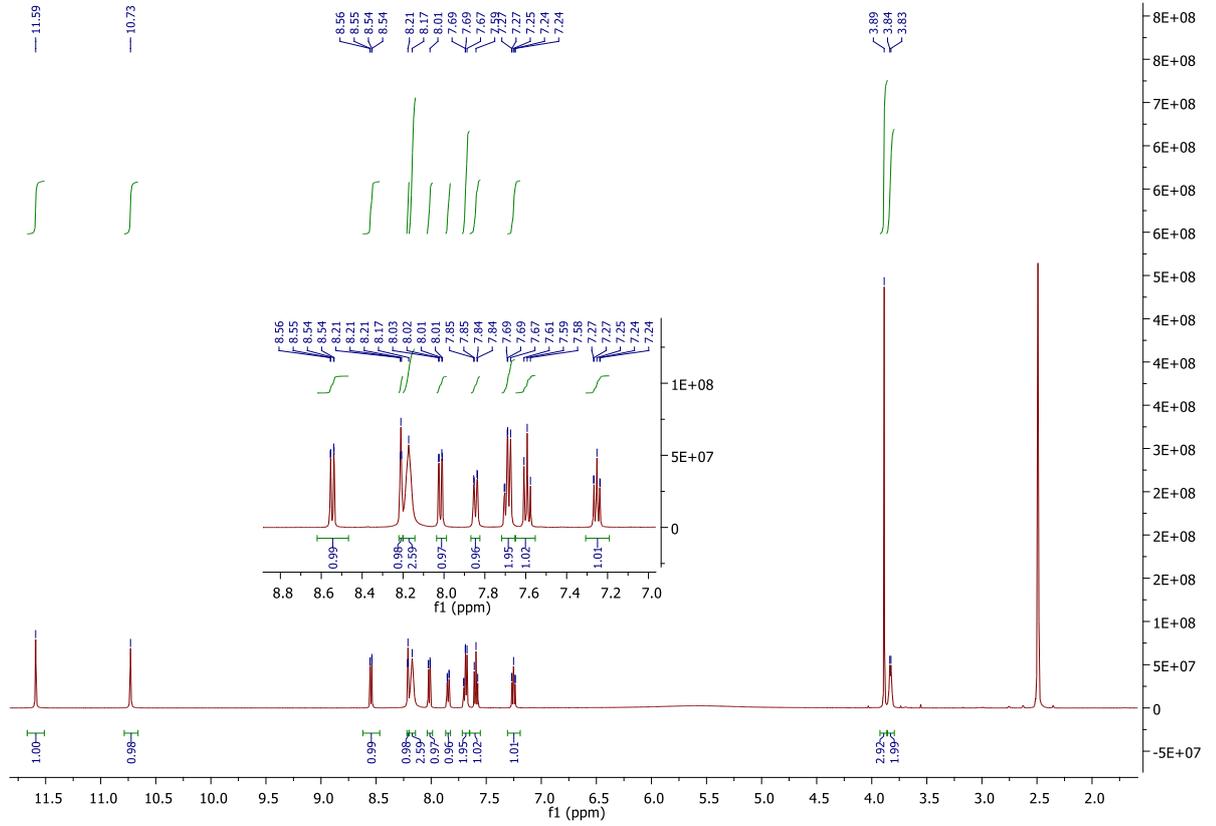
### NMR Spectra of ({3-[7-butylcarbamoyl)-2,3-dihydro-1H-indole-1-carbonyl]phenyl}carbamoyl) methanaminium chloride (35)



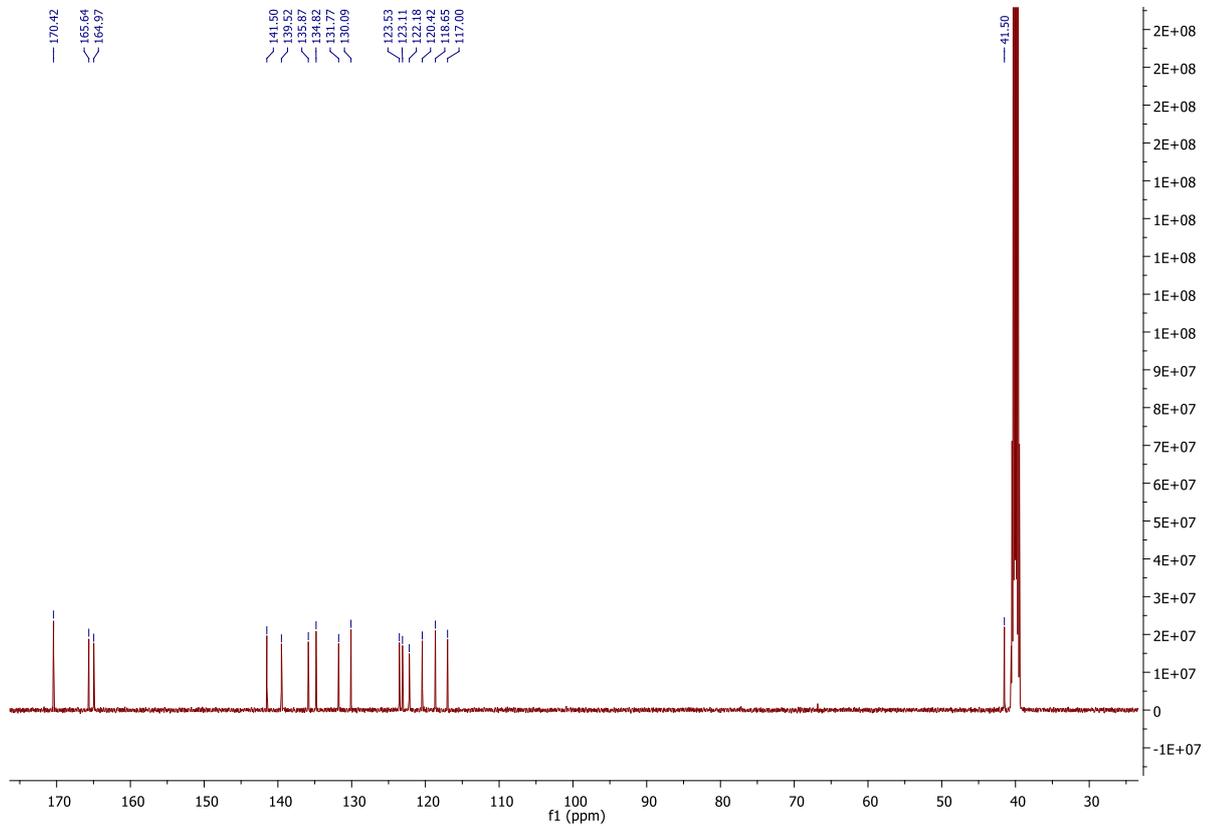
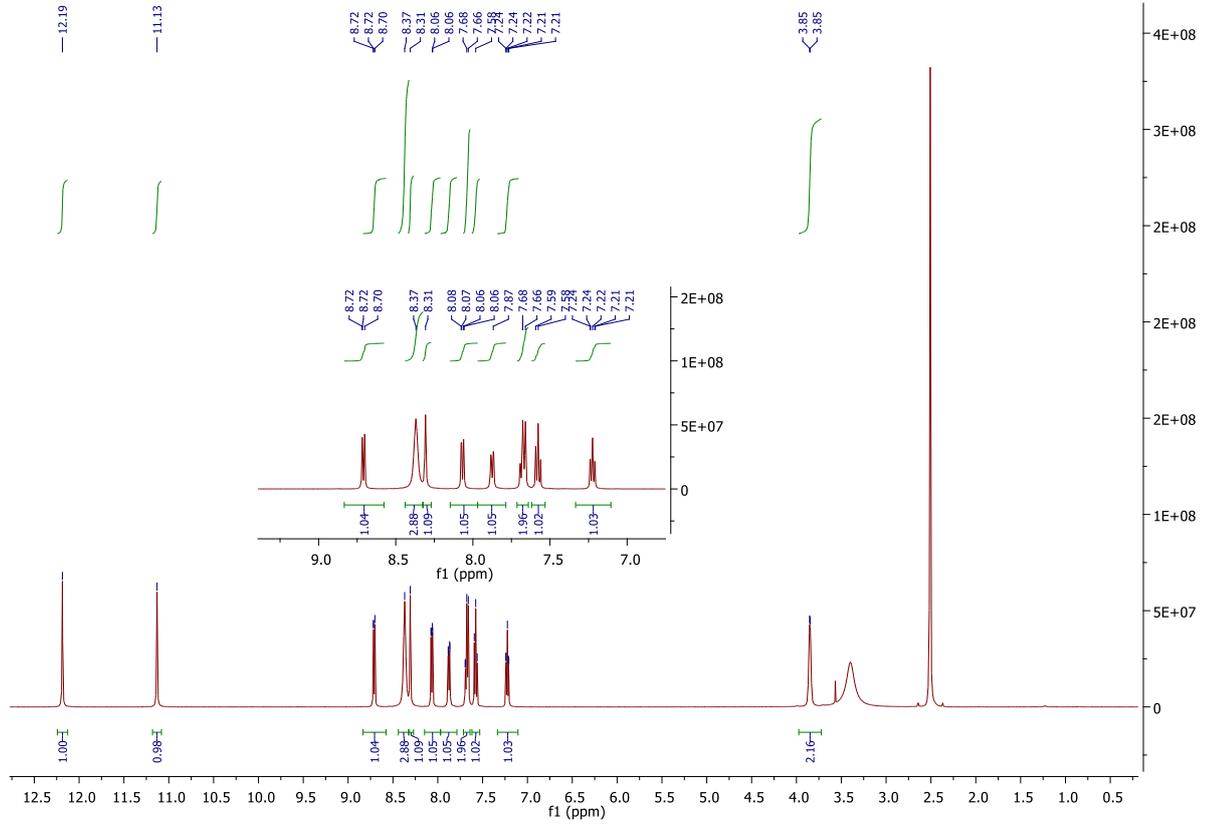
### NMR Spectra of ({3-[(2-carbamoylphenyl)carbamoyl]phenyl} carbamoyl)methanaminium chloride (47)



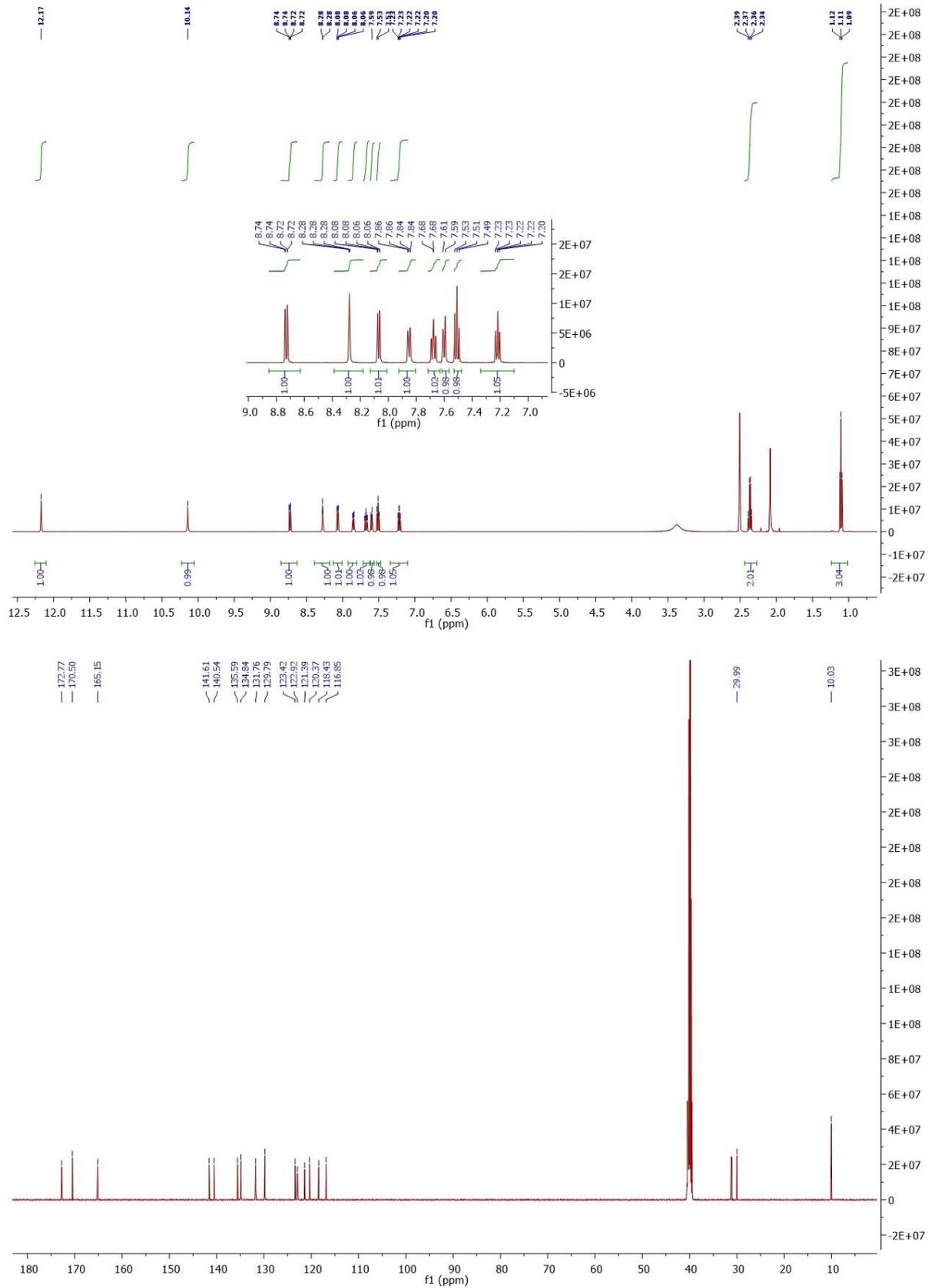
### NMR Spectra of [(3-{[2-(methoxycarbonyl)phenyl]carbamoyl}phenyl)carbamoyl]methanaminium trifluoroacetate (57)



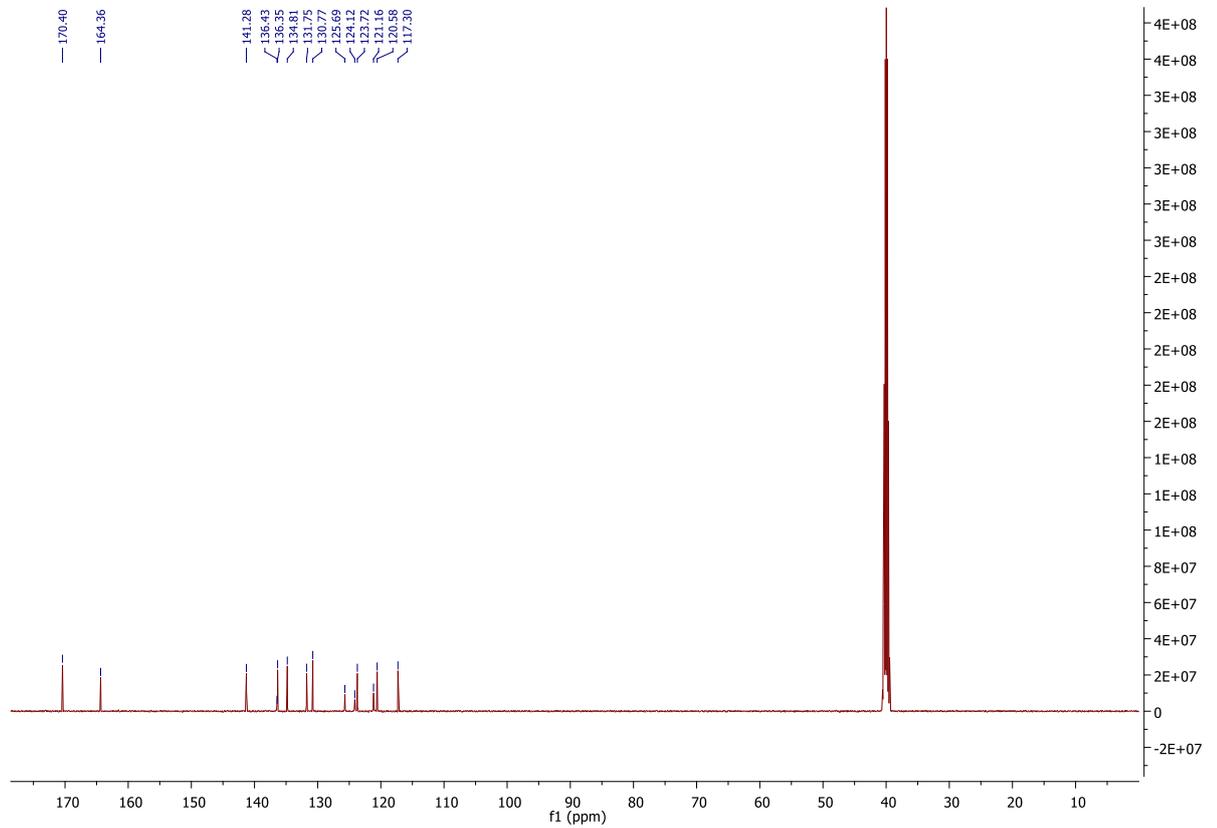
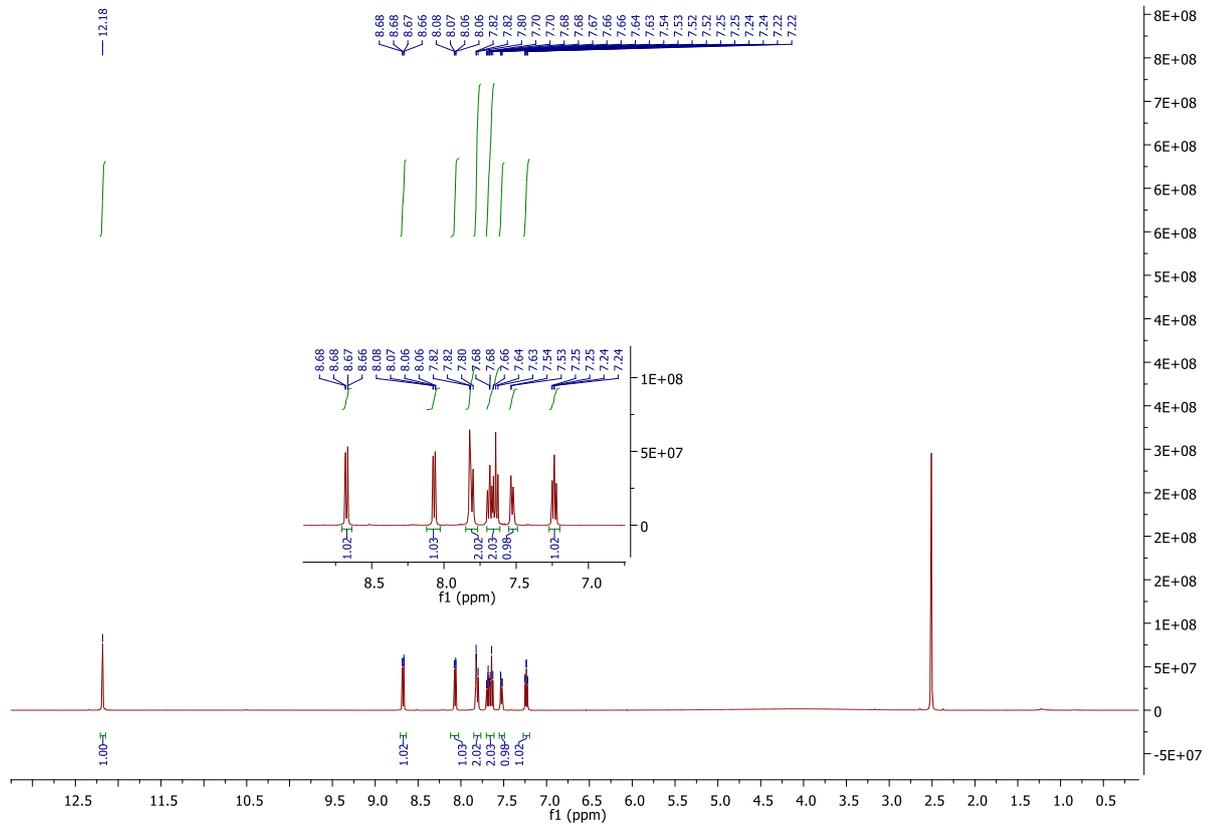
### NMR Spectra of ({3-[(2-carboxyphenyl)carbamoyl]phenyl}carbamoyl) methanaminium chloride (52)



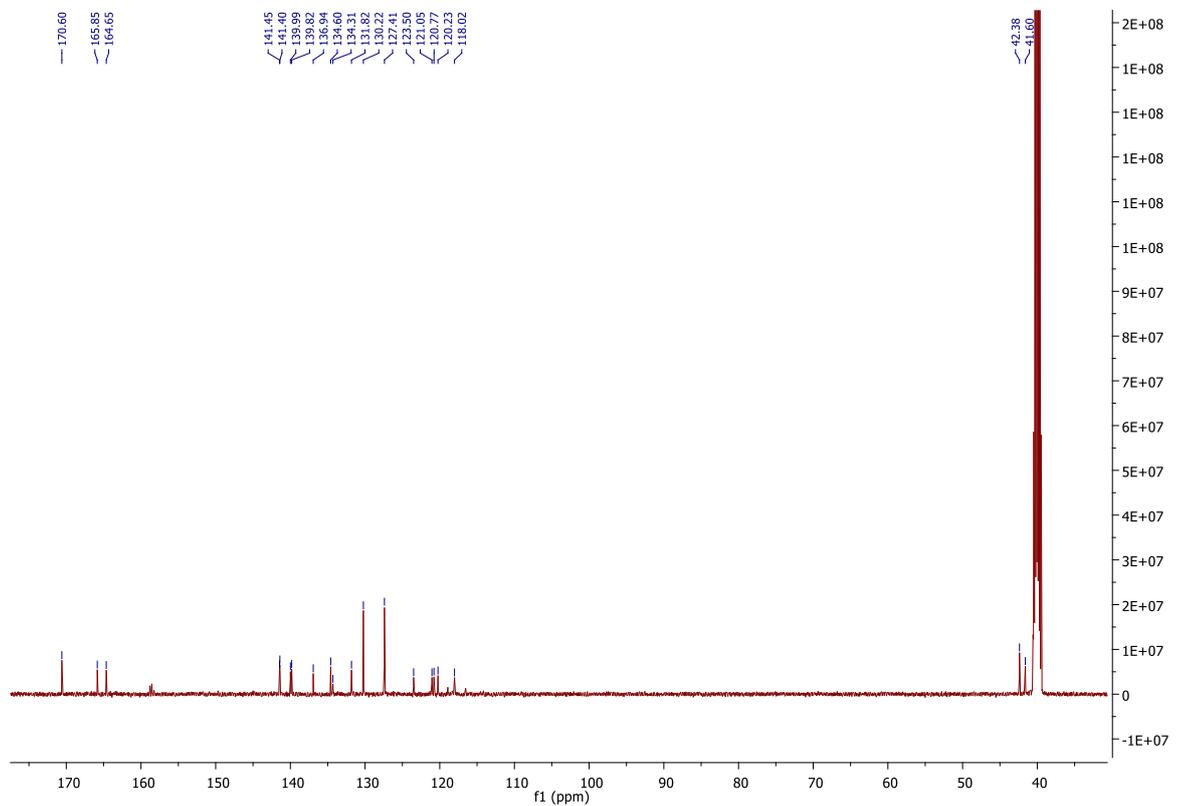
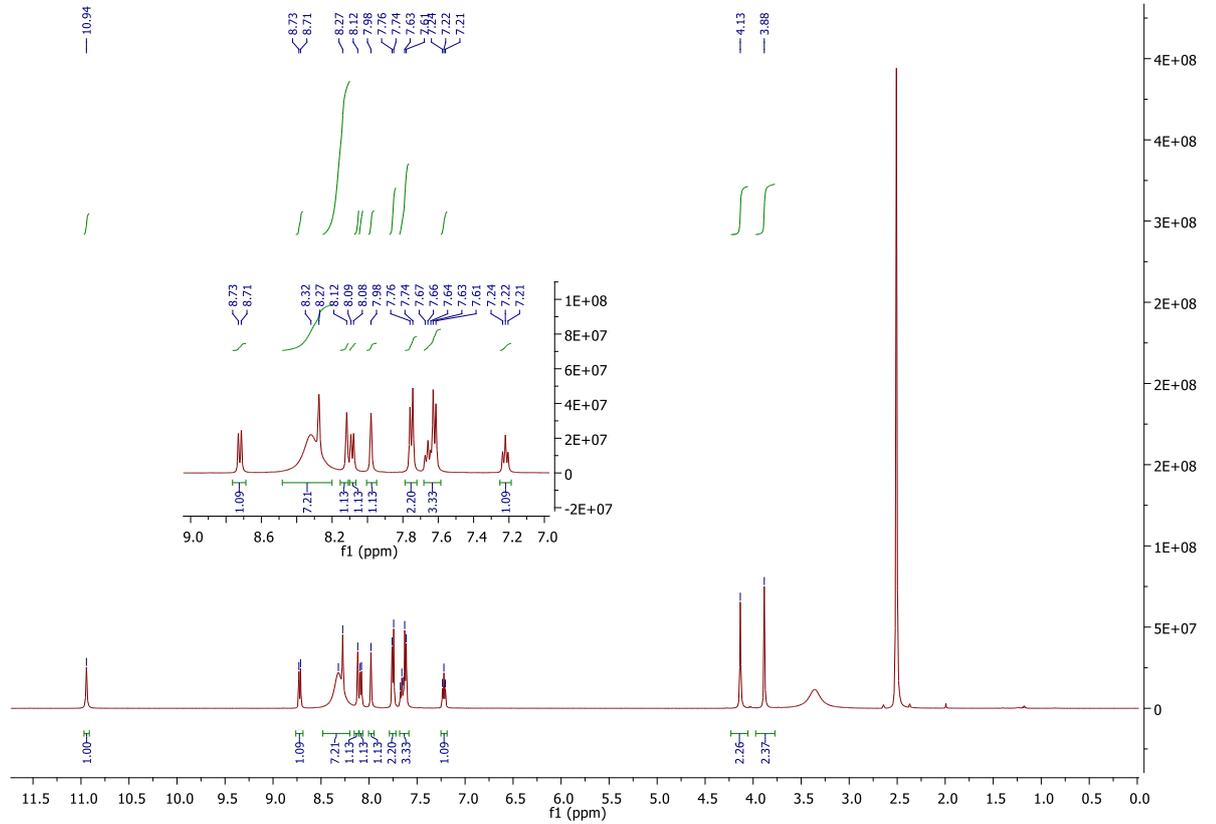
### NMR Spectra of 2-(3-propanamidobenzamido)benzoic acid (58)



### NMR Spectra of 2-(3-aminobenzamido)benzoic acid

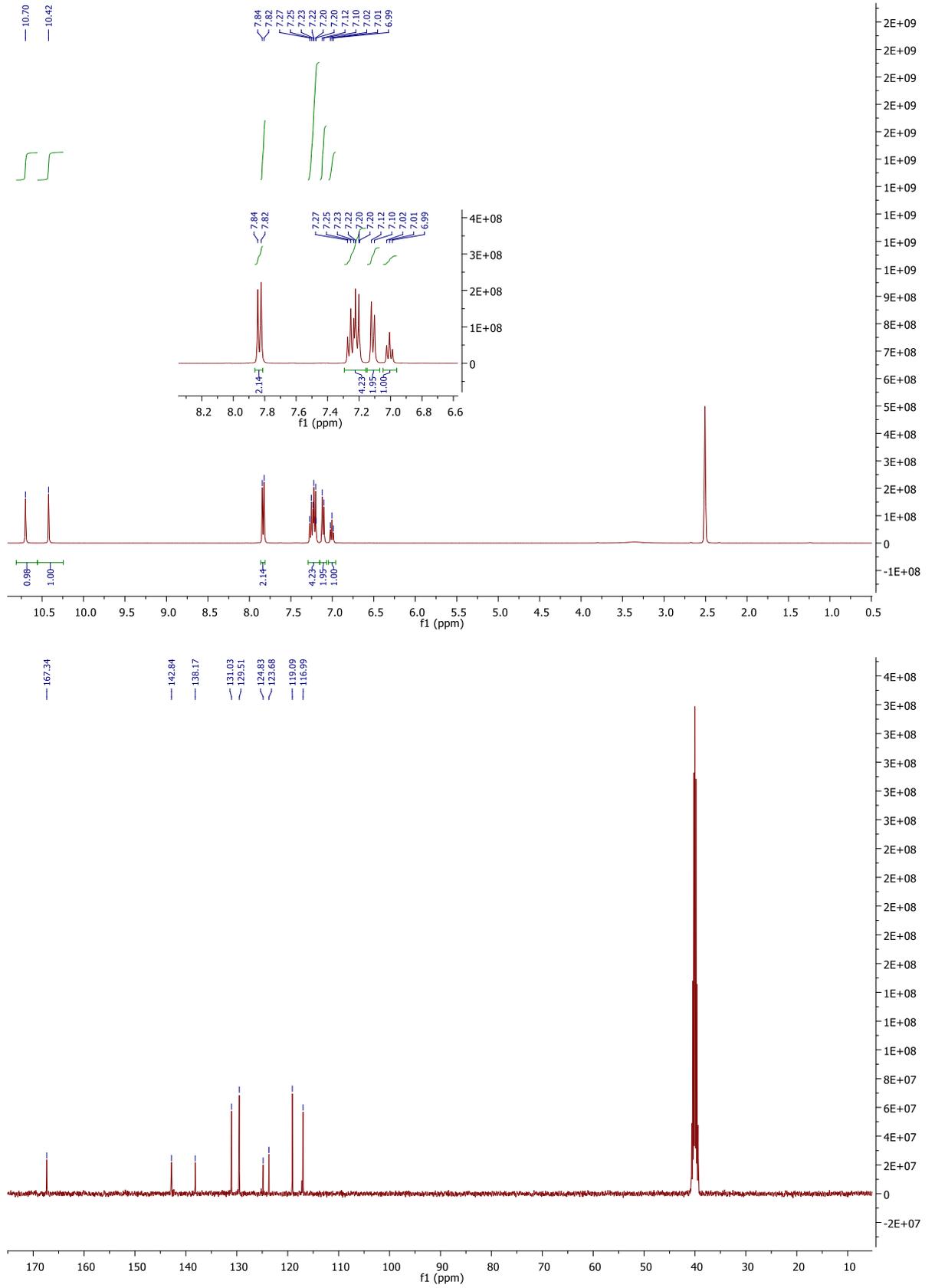


**NMR Spectra of {[4'-(azaniumylmethyl)-5-[(2-carboxyphenyl) carbamoyl]-[1,1'-biphenyl]-3 -yl]carbamoyl}methanaminium dichloride (60)**

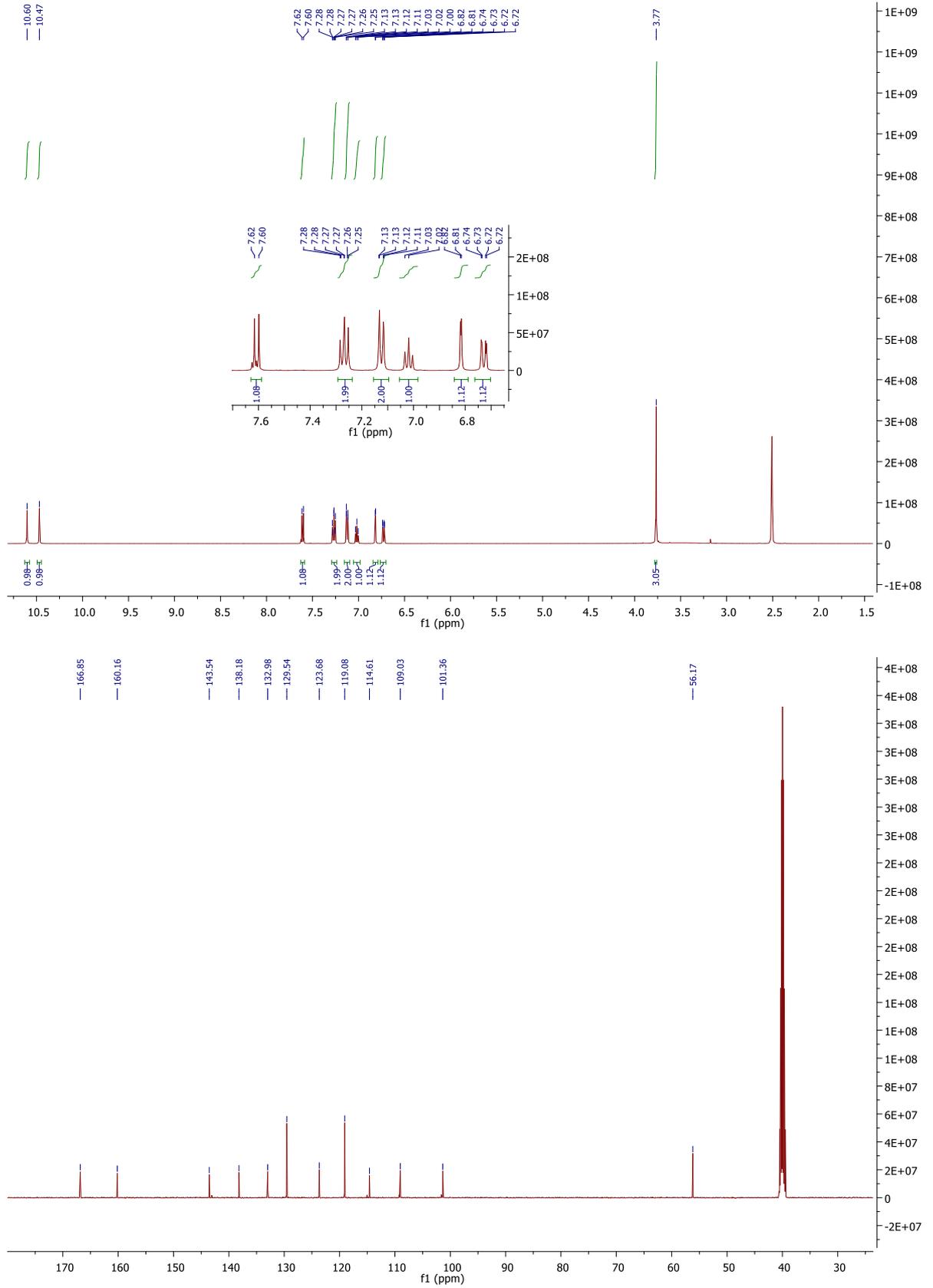




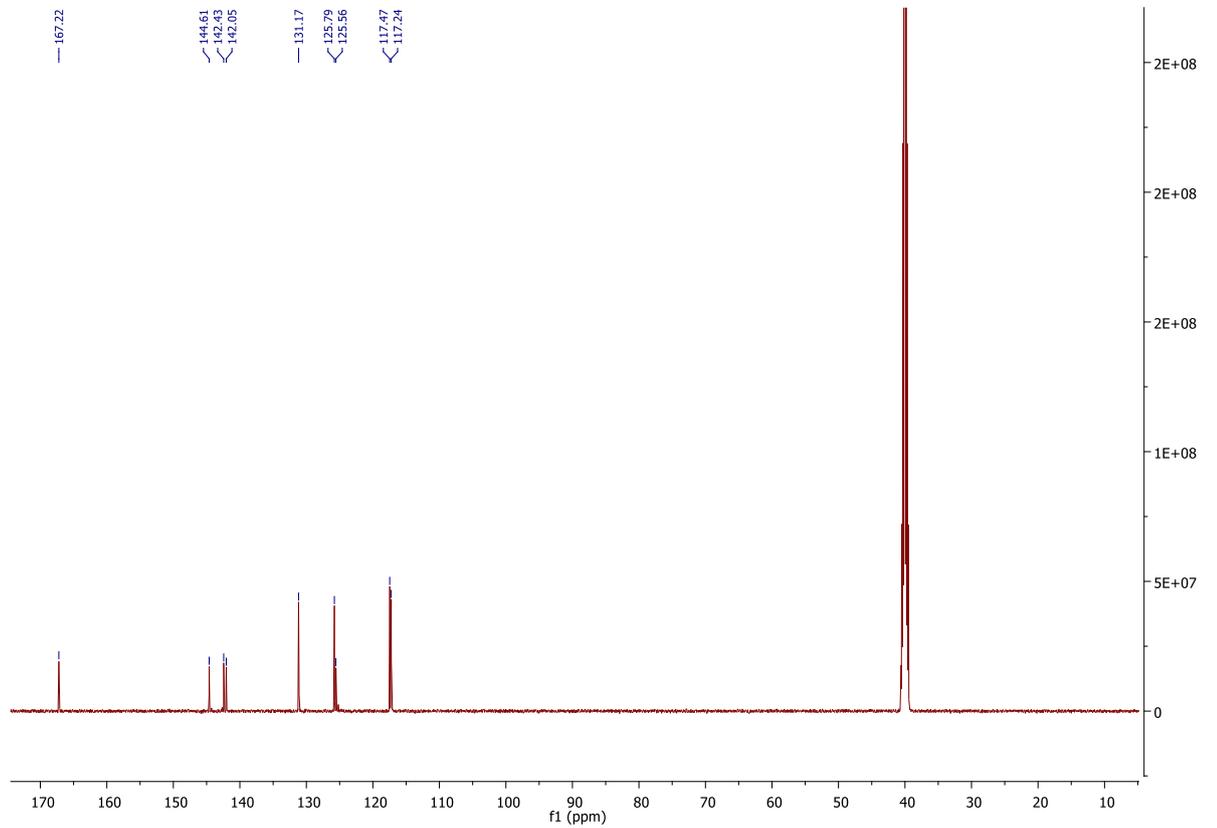
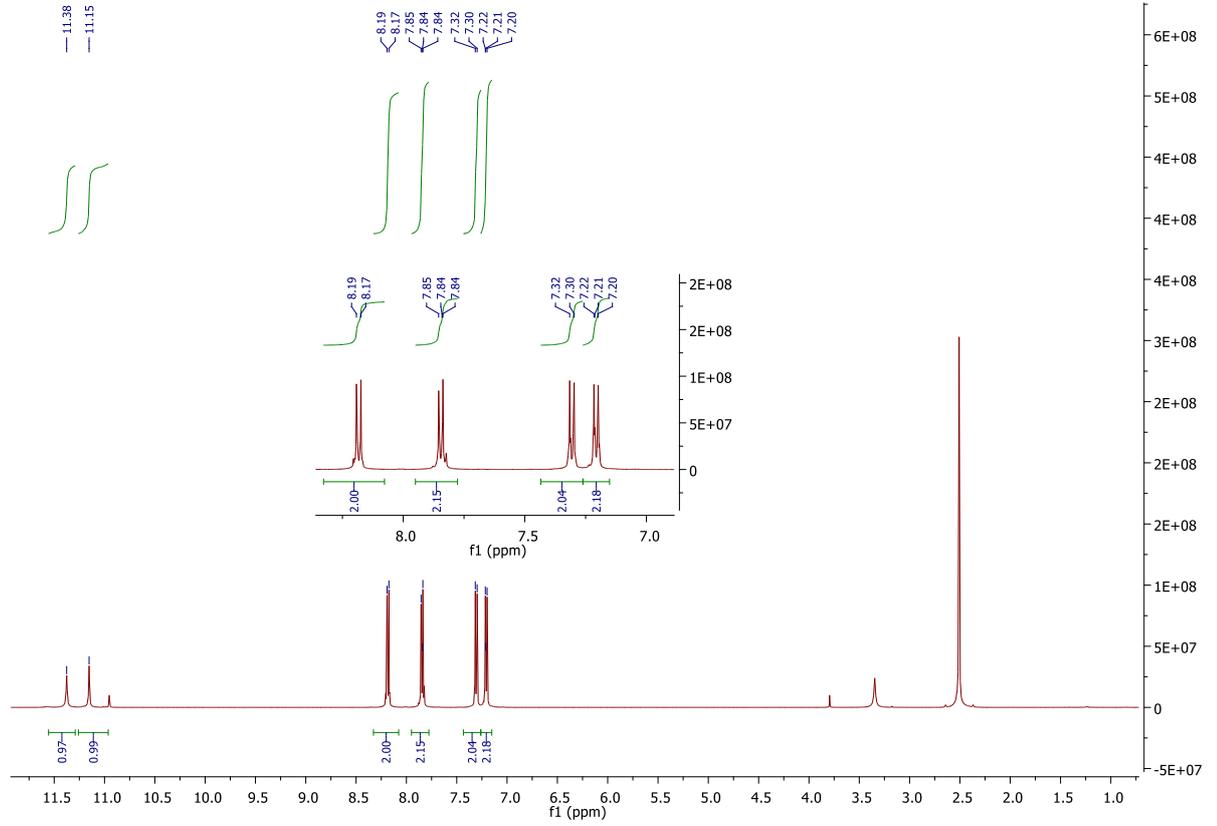
### NMR Spectra of 4-[(phenylsulfamoyl)amino]benzoic acid (71)



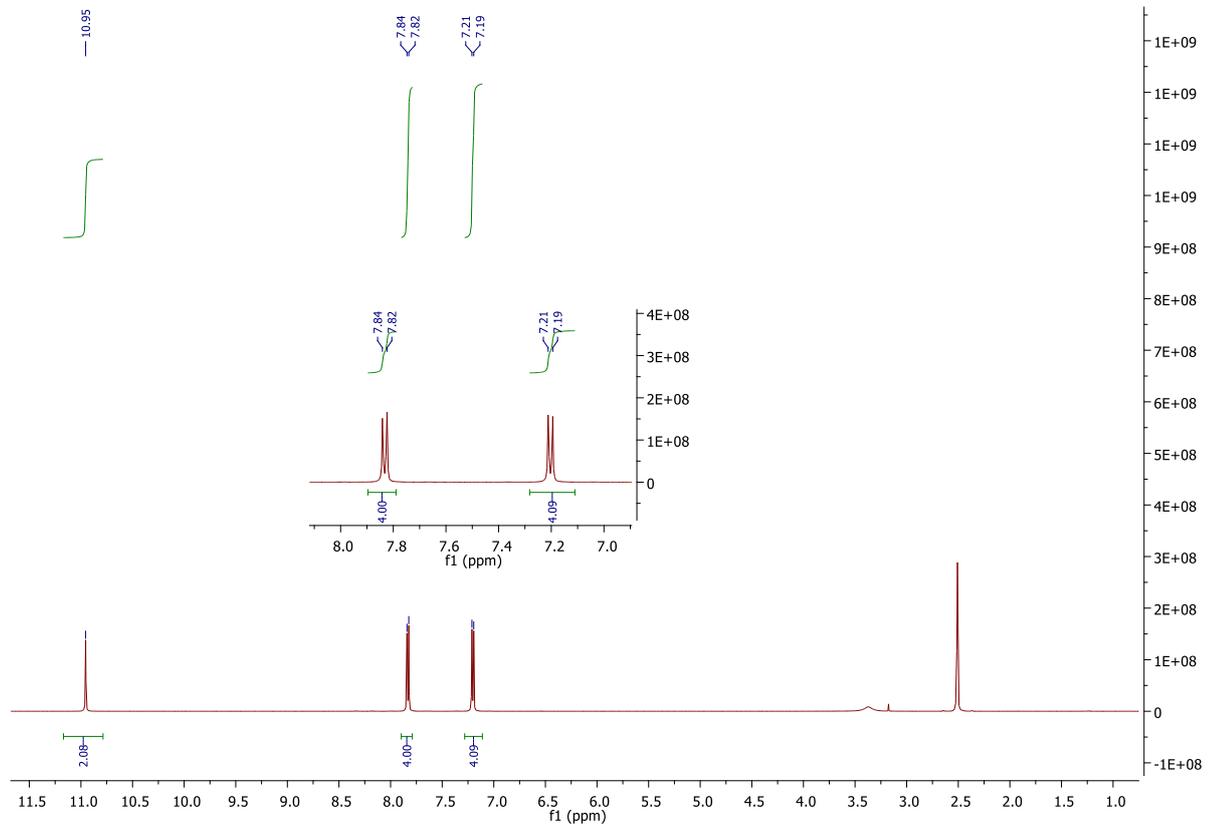
### NMR Spectra of 2-methoxy-4-[(phenylsulfamoyl)amino]benzoic acid (104)

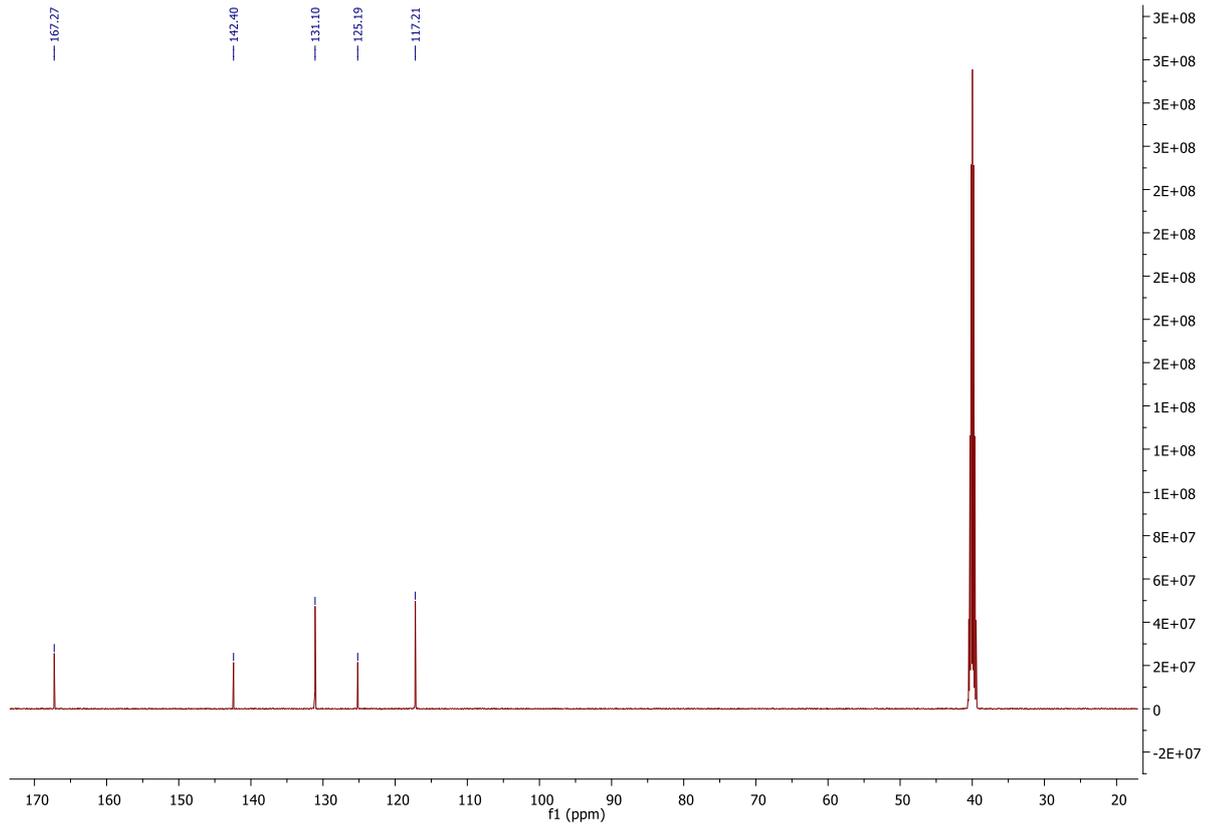


### NMR Spectra of 4-[(4-nitrophenyl)sulfamoyl]amino}benzoic acid (105)

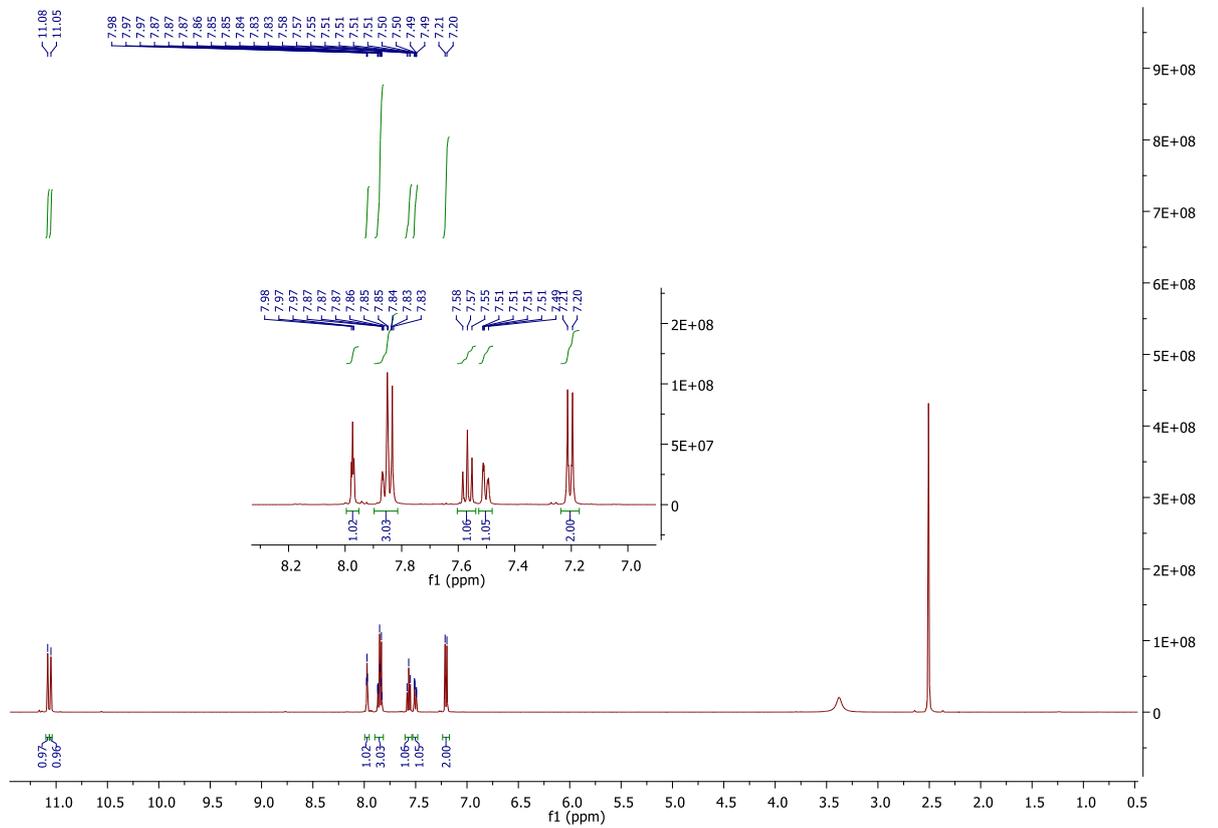


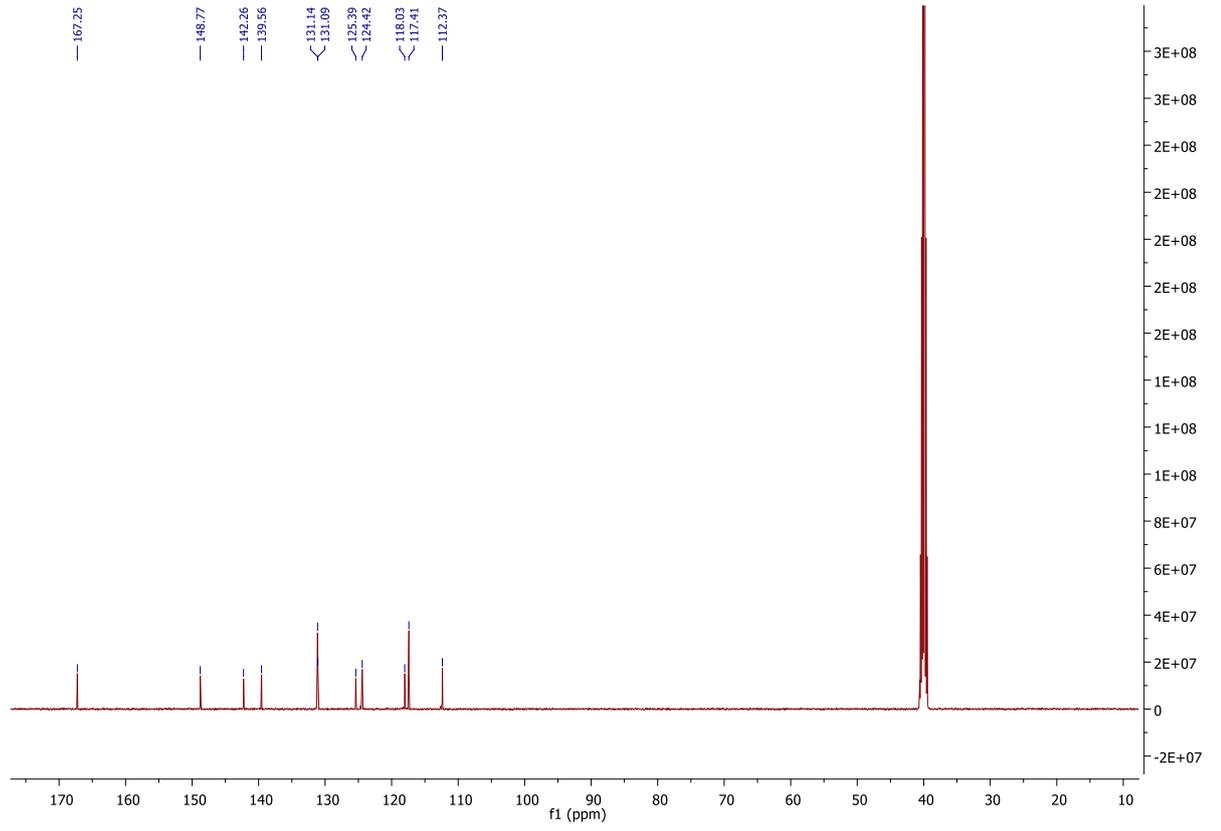
### NMR Spectra of 4-[(4-carboxyphenyl)sulfamoyl]amino}benzoic acid (106)



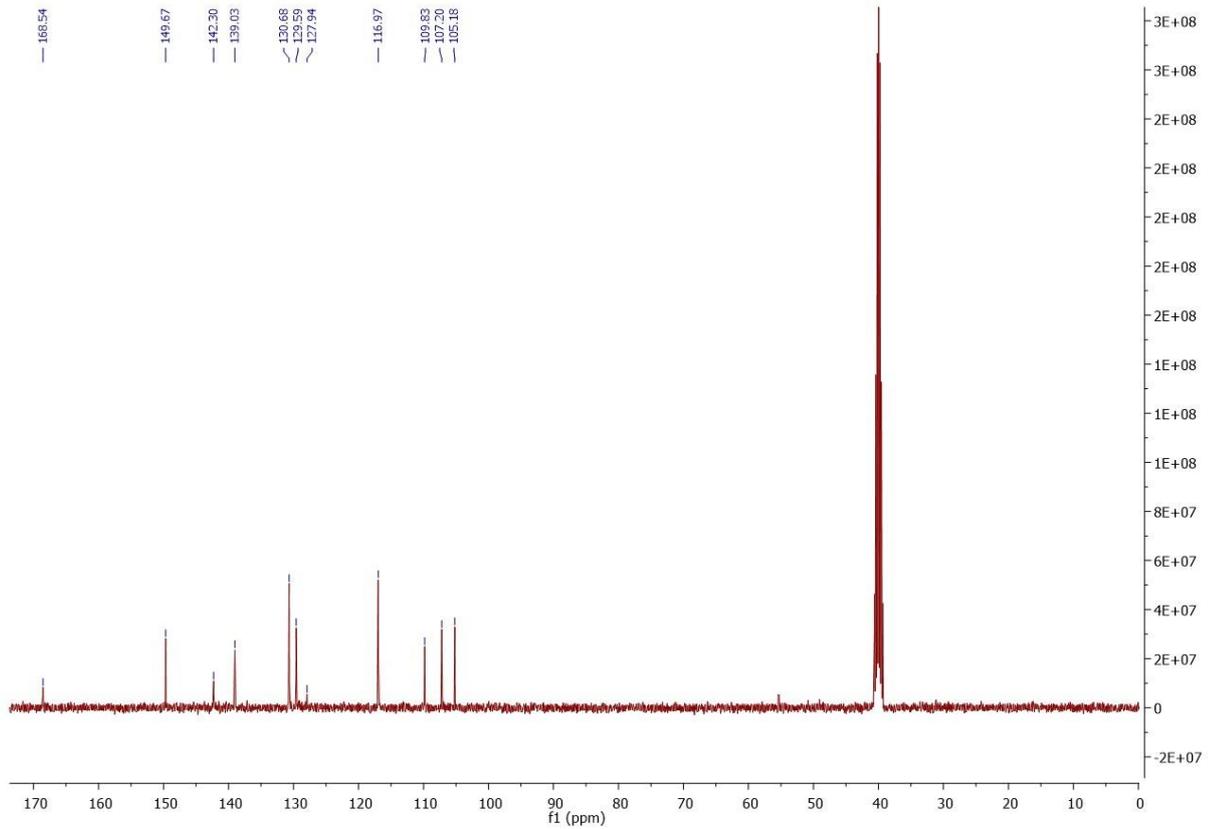
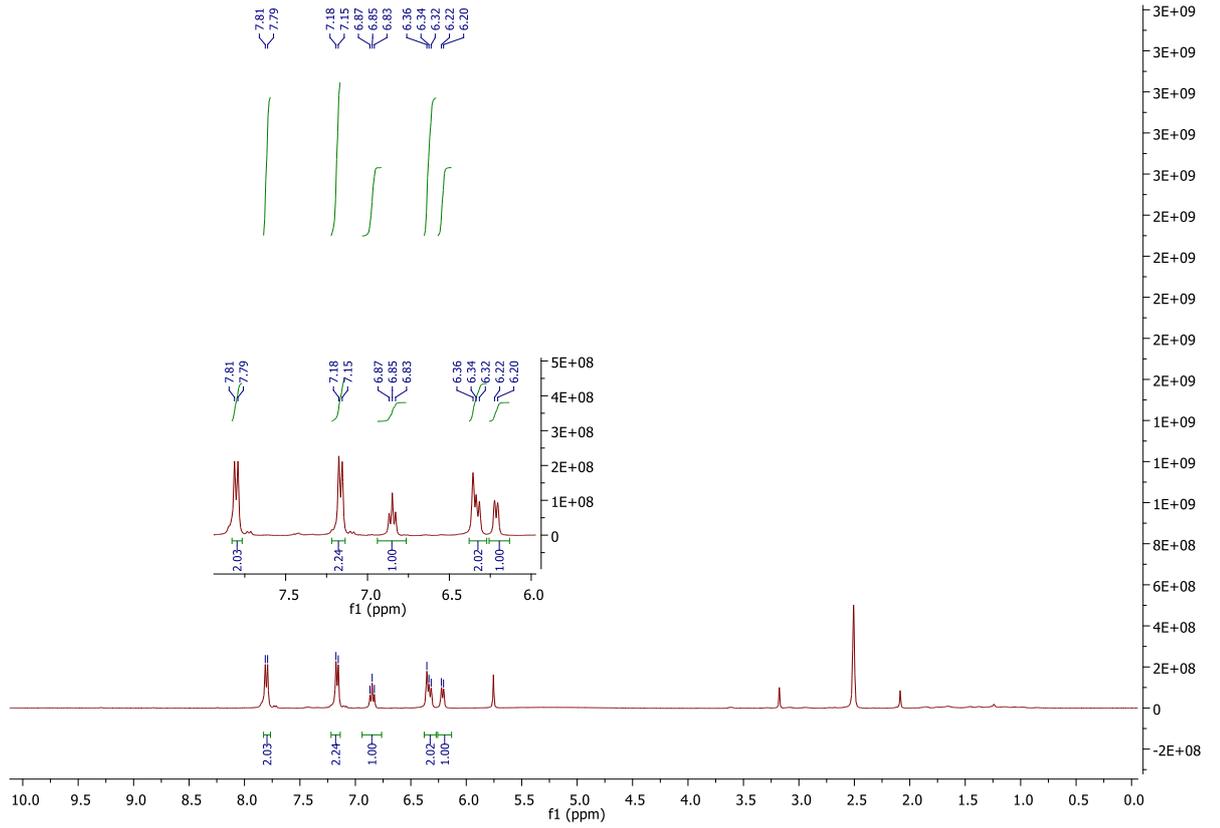


### NMR Spectra of 4-[[3-nitrophenyl)sulfamoyl]amino}benzoic acid (107)





**NMR Spectra of 4-((3-aminophenyl)sulfamoyl)amino)benzoic acid (108)**



### NMR Spectra of (4-[[4- carboxyphenyl)sulfamoyl]amino}phenyl) methaninium chloride (112)

