Virtual and *in vitro* screening of the Nottingham Managed Chemical Compound Collection *vs Mycobacterium bovis* and *Mycobacterium tuberculosis* cell wall biosynthesis enzyme: InhA.

Thesis submitted to the university of Nottingham for the degree of Doctor of Philosophy

Malcolm Lamont

University ID: 14274569

Supervised by Prof Neil R. Thomas and Dr Cristina de Matteis

DECLARATION

I declare that the content of this thesis has not been submitted nor is currently being submitted for any other degree. The research presented in this thesis is original and is the result of my own investigations. Any research that was conducted by collaboration is clearly indicated and where work of other investigators has been described this has been clearly acknowledged in the text.

Malcolm Lamont

ACKNOWLEDGEMENTS

ABSTRACT

Tuberculosis is a worldwide health concern causing 10 million new cases and 1.4 million deaths annually. Multi-drug resistance is a further issue limiting effective treatment, with the proportion of cases in this category rising each year.

InhA is a key enzyme in the *Mycobacterium tuberculosis* (Mtb) and *Mycobacterium bovis* (Mb) fatty acid synthase II pathway responsible for production of long chain mycolic acids. These acids comprise a high proportion of the Mtb/Mb cell walls. The only first-line, clinically used drug targeting InhA is the prodrug Isoniazid which is losing effectiveness in the wake of growing drug resistance. This resistance is developing at its site of metabolism to the active NAD^+ conjugate, the catalase-peroxidase KatG. By developing direct inhibitors of InhA, resistance due to mutations in KatG can be bypassed, regaining access to a clinically validated anti-TB target.

The Nottingham Managed chemical Compound Collection (NMCCC) was selected as a source of chemical diversity. It contains over 82,000 molecules pre-screened for drug-like properties and a further 3,000 molecules submitted from research projects carried out at the university. The NMCCC was subjected to virtual screening against InhA using the GOLDSuite software package and four compounds were selected for synthesis from among the highest scoring candidates. These four compounds were screened against InhA in single-point isolated enzyme assays, showing minimal activity.

A further 144 high scoring compounds were purchased from the NMCCC and screened against InhA in isolated enzyme assays. Three of these were identified as hits with over 50% inhibition of InhA at 50 µM and taken forward for synthesis. Following synthesis, precipitation of these compounds was observed prior to repeat *in vitro* analysis, requiring extended effort to solvate. Reduced activity against InhA was observed when these compounds were retested, indicating the need for orthogonal screening to confirm any hits moving forward.

There remain 11 NMCCC hits with greater than 30% inhibition of InhA at 50 µM and cLogP below 3 to be investigated and four of these show a common binding mode.

RELATED PUBLICATION

My contribution was the expression and purification of InhA and the undertaking of isolated enzyme assays as described in this thesis to evaluate inhibitors synthesised by Tom Armstrong. The *trans*-2-Octenoyl-CoA enzyme substrate mimic was also provided by Tom Armstrong.

CONTENTS

TABLE OF FIGURES

TABLE OF SCHEMES

TABLE OF TABLES

ABBREVIATIONS

ACP	Acyl Carrier Protein
Alr	Alanine Racemase
AMR	Antimicrobial Resistance
ATP	Adenosine Triphosphate
BCG	Bacillus Calmette-Guérin
BRSM	Based Upon Recovered Starting Material
CoA	Coenzyme A
COVID	Coronavirus Disease
DCC	Dicyclohexylcarbodiimide
DCS	D-Cycloserine
DDN	Deazaflavin (F420)-dependent nitroreductase
DEL	DNA-Encoded Library
DiPEA	*N*,*N*-Diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
DOTS	Directly Observed Treatment, Short course
DS-TB	Drug Susceptible Tuberculosis
E. coli	*Escherichia coli*
ECPR	Enoyl-Acyl Carrier Protein Reductase
EDC	1-Ethyl-3-(3'-dimethylaminopropyl)carbodiimide
ESI	Electrospray Ionisation
FabD32	fatty acyl-AMP ligase
FabH	β-ketoacyl-acyl carrier protein synthase III
FAS	Fatty Acid Synthase
FPLC	Fast Protein Liquid Chromatography
Gal*f*	Galactofuranose

Gal*p*	Galactopyranose
Gln	Glutamine
Glu	Glutamic acid
GOLD	Genetic Optimization for Ligand Docking
Had	β-hydroxyacyl-ACP dehydratase
HCTU	2-(6-Chloro-1*H*-benzotriazole-1-yl)-1,1,3,3,-tetramethylammonium hexafluorophosphate
His	Histidine
HPLC	High-Performance Liquid Chromatography
HTS	High Throughput Screening
HyPs	4-Hydroxy-2-Pyridones
IC_{50}	Half maximal inhibitory Concentration
IFN-γ	Interferon *gamma*
INH	Isoniazid
InhA	Mtb Enoyl-Acyl Carrier Protein Reductase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K_m	Concentration required for 50% maximal enzyme velocity
LB	Lysogeny Broth
LCMS	Liquid Chromatography Mass Spectrometry
LogP	logarithm of partition co-efficient
Lys	Lysine
MabA	keto-acyl-ACP reductase
Mb	*Mycobacterium bovis*
MDR-TB	Multi-Drug Resistant Tuberculosis
Met	Methionine
MIC	Minimum Inhibitory Concentration
Ml	*Mycobacterium leprae*
Mtb	*Mycobacterium tuberculosis*
MNGC	Multinucleated Giant Cell

m/z	Mass to charge ratio
NAD	Nicotinamide Adenine Dinucleotide
NAG	*N*-acetyl glucosamine
NAM	*N*-acetylmuramic acid.
NMCCC	Nottingham Managed Chemical Compound Collection
NMR	Nuclear Magnetic Resonance Spectroscopy
OcCoA	*trans*-2-Octenoyl-CoA
OD	Optical Density
PDB	Protein Databank
PG	Peptidoglycan
Phe	Phenylalanine
PIPES	1,4-Piperazinediethanesulfonic acid
Pks13	Polyketide synthase 13
Pro	Proline
PYR	Pyridomycin
RCSB	Research Collaboratory for structural bioinformatics
RMSD	Root Mean Squared Deviation
RR-TB	Rifampicin Resistant Tuberculosis
rt	Room temperature
SAR	Structure-Activity Relationship
SDS	Sodium Dodecyl Sulfate
SE	Standard Error
TB	Tuberculosis
TBA	Tetrabutylammonium
TCOLF	Tres Cantos Open Lab Foundation
TDZ	Thiadiazoles
TEMED	Tetramethylethylene diamine
TFA	Trifluoroacetic acid
TLC	Thin Layer Chromatography

Tyr	Tyrosine
UDP	Uridine Diphosphate
UGM	UDP-Galactopyranose Mutase
UV	Ultra Violet
V_0	Rate in absence of inhibitor
V_1	Rate in presence of inhibitor
vdw	Van der Waal
Vis	Visible
V_{max}	Maximum enzyme velocity
v/v	Volume per volume
WHO	World Health Organisation
XDR-TB	Extensively Drug Resistant Tuberculosis
XRD	X-Ray diffraction

1.0: Introduction

1.1   Tuberculosis

Tuberculosis (TB) causative agents have been infecting humans since we spread from Africa up to 40,000 thousand years ago.[1,2] The disease is currently responsible for 1.4 million deaths and an estimated 10 million new cases arise annually.[3] It is further estimated that there is a pool of two billion individuals infected with a latent form of the disease, 5-10% of which will progress to the active, transmissible form over their lifetime.[3]

The impact of Coronavirus on TB infections

TB has historically been the highest cause of death by a communicable disease worldwide in the 20th and 21st centuries but this has been complicated in 2020 by the rise of coronavirus.[3] Social distancing measures are expected to have helped somewhat in the spread of both diseases but have had other knock on effects. A drop off has been observed in the effective monitoring, diagnosis and treatment of TB cases, and extra pressure placed on outlets for preventative care. This is expected to have resulted in an increase of 200,000 – 400,000 excess TB deaths in 2020, for a total of 1.6 - 1.8 million deaths compared to the 1.7 million estimated for coronavirus.[3,4] This would be a reversion to 2015 or 2012 TB death rates respectively.[3] In the absence of both economic and health mitigation measures, the outfall from coronavirus is expected to cause an additional 6.3 million TB cases between 2020 and 2025.[3]

Affected areas

Developing countries are especially at risk due to poor therapeutic coverage, poor incidence reporting and the prevalence of HIV which can coinfect and increase mortality with TB.[3] Eight countries; India (26%), Indonesia (8.5%), China (8.4%), the Philippines (6.0%), Pakistan (5.7%), Nigeria (4.4%), Bangladesh (3.6%) and South Africa (3.6%) contain two thirds of global active TB cases, encouraging the development of new and lower cost therapeutics (**Figure 1**).[3]
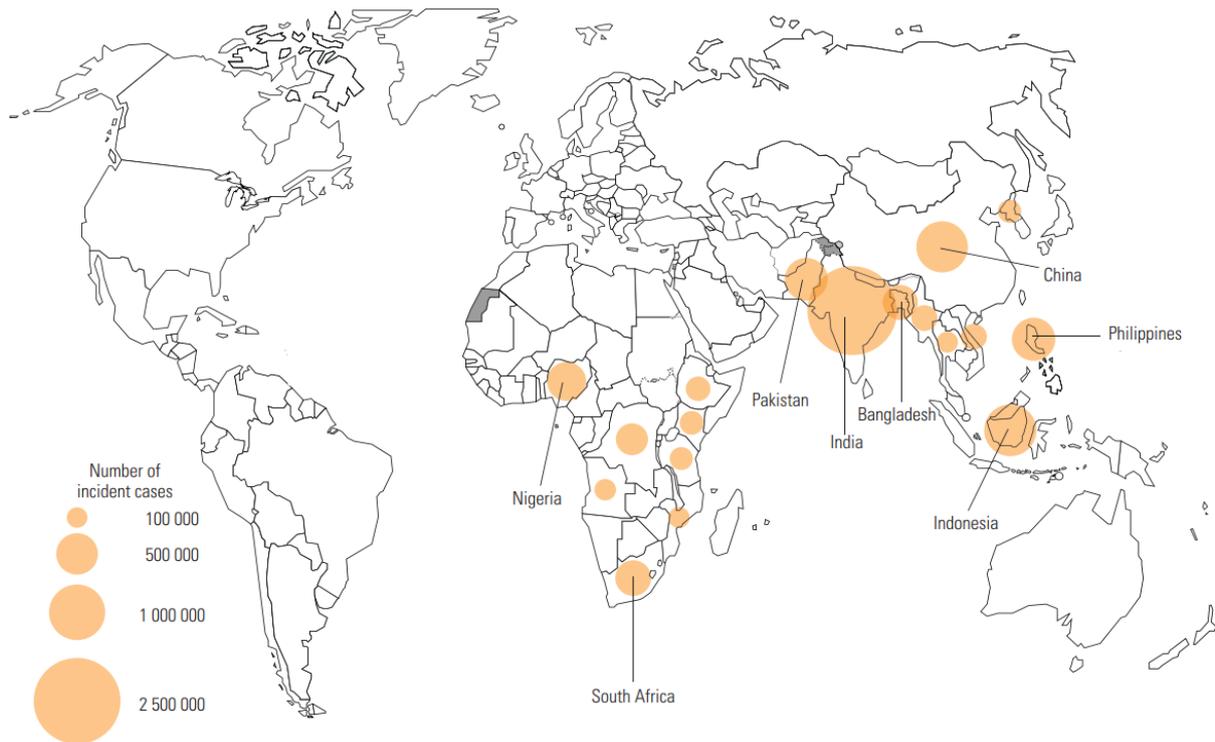
Figure 1: Countries with over 100,000 cases of TB in 2019. The Eight labelled countries comprise two thirds of all global TB cases, image reprinted with permission from the WHO global TB report 2020.[3]

The *Mycobacterium tuberculosis* complex

The *Mycobacterium tuberculosis* complex (MTC) is a pool of 10 related bacteria that cause TB in humans and animals.[5,6] Two well-known members of the complex are *Mycobacterium tuberculosis* (Mtb) and *Mycobacterium bovis* (Mb), the most common causative agents of TB in humans and cattle respectively. Mb can cause zoonosis by crossing species barriers which is a concern in developing countries where milk is often unpasteurised.[7] In 2019, the World Health Organisation (WHO) recorded an estimated 140,000 cases of Mb zoonotic TB, resulting in 11,400 deaths.[3]

Mycobacterium bovis

Mb presents challenges beyond human transmission. Annual testing of the 74,457 cattle herds in Great Britain in 2018 resulted in evidence of bovine TB among 4,378 (6%).[8] The southwest of England in particular is a high risk area, presenting 2,761 of these incidents (**Figure 2**). The north and east of England and central and northern wales are less

2

affected, with fewer than 300 cases per region. Scotland has benefitted from being "officially TB free" since 2009, with 36 incidents observed across the country in 2018.[8]
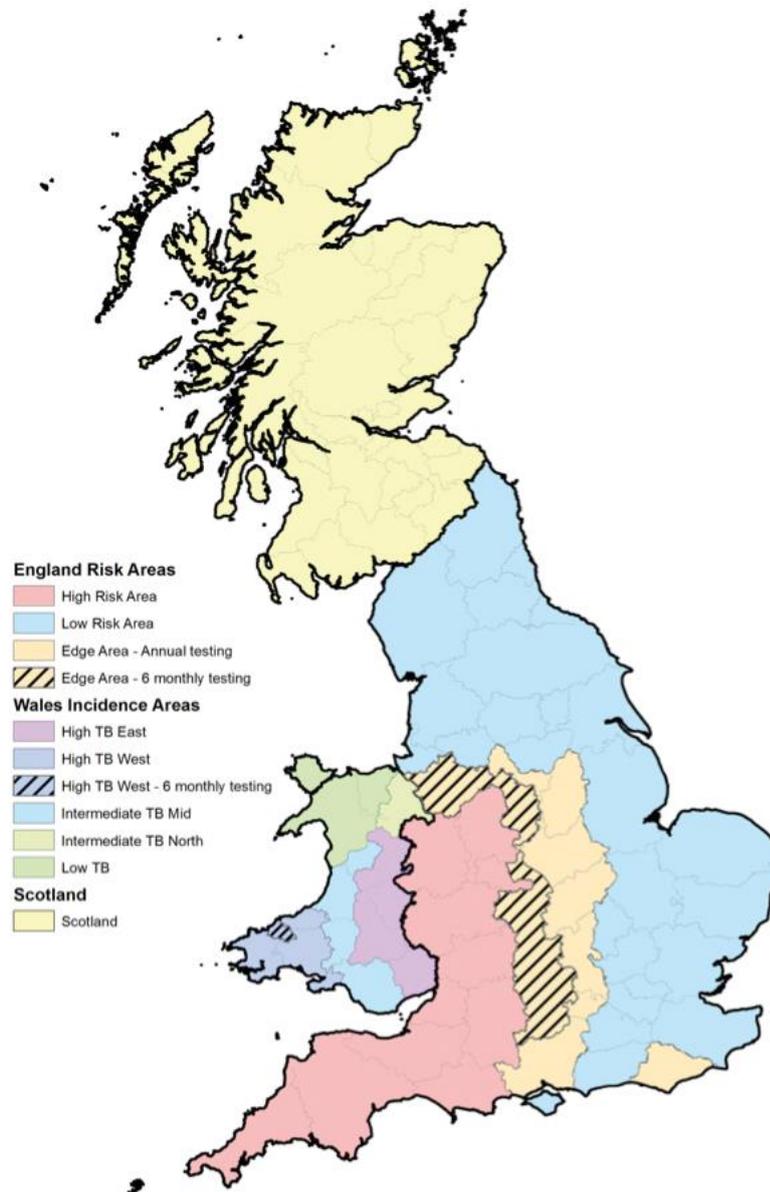
Figure 2: Risk levels of bovine TB in Great British herds by area in 2018, image reused with permission from Duncan *et al*..[8]

Evidence of Mb in a herd results in large scale culling, costing the UK £150M and 30,000 cattle annually, adversely affecting biosafety and food security.[9] Antimicrobial drugs are not currently used to treat bovine TB but work has been ongoing since 1998 to provide an alternative vaccine to the Bacillus Calmette-Guérin (BCG) by 2025. The BCG vaccine causes false positive results during commonly used bovine skin tests and an alternative

would be valuable as part of the government's agenda to reach officially tuberculosis free status for England by 2038.[9]

Pathogenesis

TB infection begins through collision of infected airborne water droplets with the mucous membrane. The pathogen localises in the lungs and is absorbed by alveolar phagocytes such as macrophages.[10] Mtb is capable of preventing the maturation of these cells, freely replicating throughout the phagocyte due to the protection offered by a lipophilic outer cell wall.[10] Once the bacterial load reaches high density, the bacteria is released. In response to foreign objects, uninfected macrophages coalesce and form multinucleated giant cells (MNGCs) through fusion of their outer membranes (**Figure 3**).[10] Other immune cells aggregate around the MNGC, forming lesions called granulomas. Granulomas are formed primarily to 'stalemate' pathogen growth and increase exposure of infected mature phagocytes to T-lymphocytes.[10] If a stable granuloma forms and stalemate is achieved, as in 90% of cases, the infection is classified as "latent" TB. If a stalemate is not established, the granuloma provides a high density of uninfected phagocytic cells to which the pathogen may spread.[10] Bacteria released by infected cells can spread throughout the body and eventually to other patients.

Multinucleated
giant cell

Neutrophil

Macrophage

Dendritic cell

M. tuberculosis

B cell

CD4 T cell

CD8 T cell
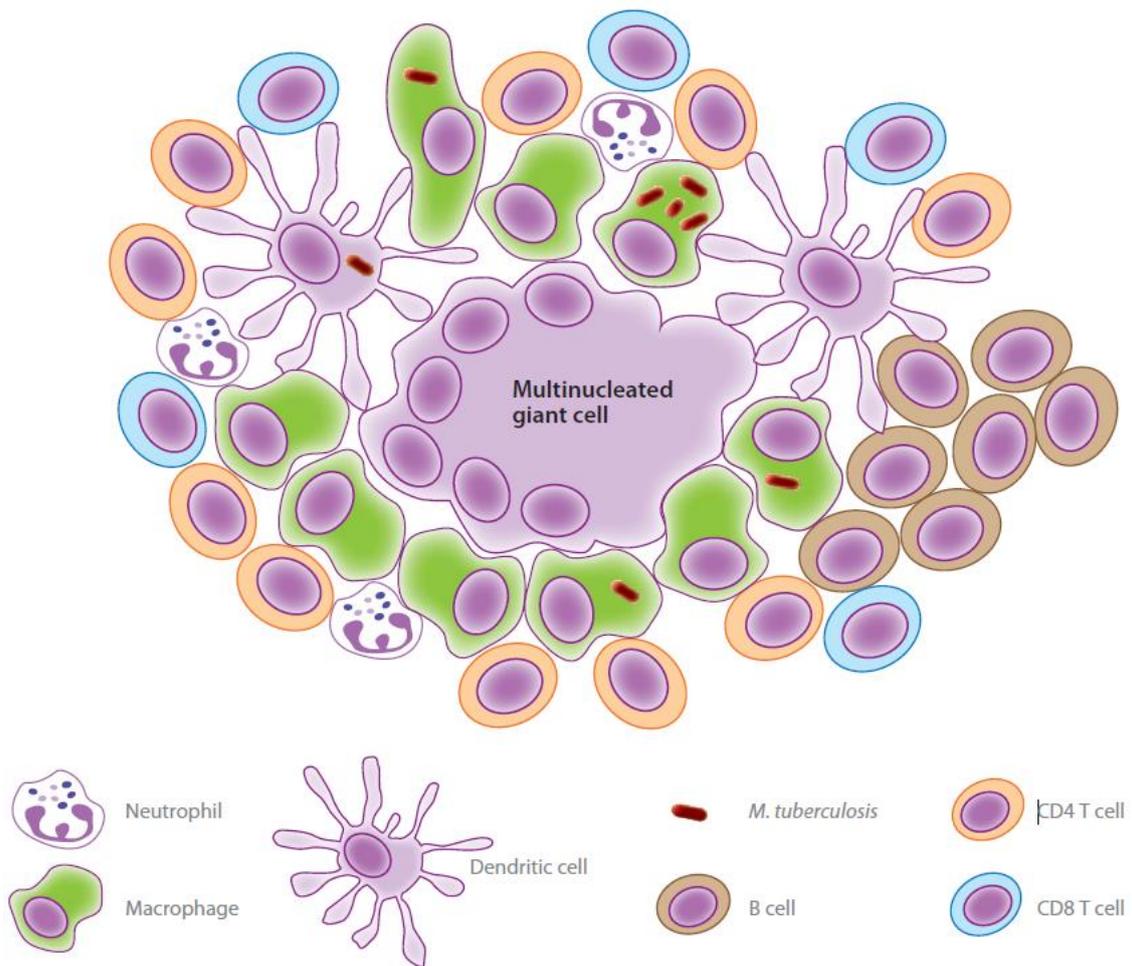
Figure 3: *Mycobacterium tuberculosis* granuloma composition, image reused with permission from Philips and Ernst.[10]

Diagnosis and testing

Testing for TB is generally achieved through two avenues: direct detection or through manipulation of the immune response. Direct detection can be carried out using microscopy, phenotypic screening, mycolic acid analysis or genotypic analysis.

When using microscopy, acid-fast stains are required despite the similarities between the inner cell wall of Mtb and other Gram-positve bacteria (the mycobacterial cell wall is discussed in further detail below). This is due to repulsion of Gram-positive stains by the Mtb outer mycolic acid layer. Instead, acid-fast stains are used which can enter the cell wall with heating and are retained despite the presence of an acid/alcohol wash.[11] Phenotypic screens are wide and varied, observing the growth of *Mycobacteria* in the

presence of established inhibitors and this data can be compared to reference strains to diagnose infection.[11] Both microscopy and phenotypic screening suffer from turnaround times of three to six weeks due to the extended generation time of Mtb.[11] Mycolic acid analysis is usually focused on in research settings as opposed to active population testing and utilises HPLC methods to identify species diversity within *Mycobacteria* according to their outer lipids.[12] Genotypic analysis methods are often a fast and accurate (over 95%) means of obtaining results. Although DNA sequencing usually requires costly equipment and technical expertise to carry out, widespread implementation of the technique has been seen following the advent of qPCR testing for COVID-19.[11,13]

The immune response is tested by observing cell-mediated reaction to mycobacterial antigens. The two most common methods are the tuberculin skin test (TST) and interferon-γ (IFN-γ) release assays (IGRAs). The TST was first developed over 100 years ago as the Mantoux test and involves injecting a purified protein derivative (PPD) intradermally.[14] A positive test is can be read after 48-72 hours by inflammation of the injected area.[11] TSTs have been accepted by the WHO as a standard diagnostic for latent TB infections with a sensitivity of 70-85%.[14] Although valuable tools in tracking TB infection, TSTs have been known to cause false positives in individuals previously inoculated with the BCG vaccine.[14] IGRAs operate by measuring release of IFN-γ following treatment of blood samples with Mtb antigenic material.[14] They offer two advantages over TSTs, they use more specific Mtb antigens and testing can be carried out *in vitro* after one patient visit.[14] IGRAs are, however, more expensive to run due to the requirement for phlebotomy and laboratory access. Sensitivity is variable and ranges from equal to TST (77%) to 98% depending on the commercial test used.[14]

Vaccination
The BCG vaccine is an attenuated vaccine originally developed over 100 years ago from Mb and has been available in 14 different strains since its inception.[15] It is currently the only licensed vaccine for tuberculosis, with no other candidates surpassing it for protection

or sustainability.[16] It has been found to provide 10-20 years of protection in infants but is generally ineffective at preventing the development of pulmonary TB in adults or adolescents.[16] As both TSTs and the BCG vaccine are advocated for use by the WHO, TST screening prior to vaccination is vital to ensure accurate reporting of vaccine effectiveness and coverage due to the chance of BCG inoculation causing false-positive TSTs later in life.[17]

There is a well populated pipeline to produce new TB vaccines, and candidates use one of four methods (**Table 1**). The viral vector vaccines Ad5 Ag85A, ChAdOx1 85A-MVA85A and TB/FLU-04L are all intended to express the MTB Ag58 antigen and TB/FLU-04L has been found to be effective in adults.[16]

Protein adjuvant vaccines AEC/BC02, H56: IC31, ID93 + GLA-SE and M72/AS01E (GSK 692342) contain multiple immune system-activating MTC peptides along with adjuvants intended to boost the effectiveness of the vaccine, especially in the immune compromised.[16] ID93 + GLA-SE and M72/AS01E (GSK 692342) are currently being trialled in TB-HIV co-infected adults and otherwise healthy HIV-infected adults respectively.[16]

The lysate extract vaccines RUTI® and Vaccae™ are both advanced phase III candidates intended to be used in combinatorial treatments with chemotherapy and show some promise.[16] Whole cell vaccines DAR-901 booster and MIP/Immuvac are heat-killed *Mycobacteria*.[16] MIP/Immuvac has already been used to successfully prevent the development of leprosy caused by the MTC related bacteria: *Mycobacterium leprae*.[18]

Finally, MTBVAC is currently the only live candidate based directly on attenuated Mtb in clinical trials.[19] It is intended as a BCG replacement and has been altered by the deletion of two virulence factors.[19]

Table 1: Summary of the current TB vaccine pipeline reconstructed from Li et al.[16]

Type	Candidate	Phase	Sponsor
Viral vector	Ad5 Ag85A	I	McMaster University; Canadian Institutes of Health Research; CanSino
	ChAdOx1 85A-MVA85A	I	Oxford University
	TB/FLU-04L	IIa	Research Institute for Biological Safety Problems; Ministry of Health, Kazakhstan; Research Institute of Influenza, Russia
Protein adjuvant	AEC/BC02	I	AnHui Zhifei Longcom
	H56: IC31	IIb	Statens Serum Institut; Valneva; Aeras
	ID93 + GLA-SE	IIa	Infectious Disease Research Institute; Aeras; Wellcome Trust; International AIDS Vaccine Initiative
	M72/AS01E (GSK 692342)	IIb	GlaxoSmithKline; Aeras
Whole cell/ extract	RUTI®	III	Archivel Farma S.L.
	DAR-901 booster	IIb	Dartmouth University; Global Health Innovative Technology Fund; Aeras
	MIP/Immuvac	III	Indian Council of Medical Research; Cadila Pharmaceuticals
	Vaccae™	III	AnHui Zhifei Longcom
Attenuated	MTBVAC	IIa	University of Zaragoza; Biofabri; Tuberculosis Vaccine Initiative

Chemotherapy and anti-microbial resistance

First line treatment of TB has remained relatively unchanged since the 1970s with all four

members discovered by 1963 (**Figure 4**).[20]

Figure 4: The current first line treatments for TB; isoniazid (**1**), ethambutol (**2**), pyrazinamide (**3**) and rifampicin (**4**).

First line treatments of drug susceptible TB (DS-TB) use combinations of isoniazid (INH,

1), ethambutol (**2**), pyrazinamide (**3**) and rifampicin (**4**) to target a wide range of

Mycobacterial processes.[20,21] INH is a prodrug activated by catalase peroxidase (KatG) and

inhibits one of the mycolic acid biosynthesis enzymes used in construction of the cell wall, InhA.[22] Ethambutol also targets cell wall biosynthesis but by inhibiting production of the arabinogalactan layer *via* arabinosyl transferase.[22] Pyrazinamide, also a prodrug, is hydrolysed in the cytoplasm to pyrazinoic acid by the nicotinamidase PZase.[23] Pyrazinoic acid has been found to be a bifunctional anti-TB therapeutic, acidifying the bacterial cytoplasm whilst inhibiting *trans*-translation *via* the 30S ribosomal subunit.[24,25] Rifampicin is a semi-synthetic antibiotic that readily diffuses through membranes and blocks transcription of messenger RNA in active and latent bacteria.[26,27]

All four drugs are initially taken together daily for 8-12 weeks, after which only INH and rifampicin are taken for a further 18-31 weeks, depending on the treatment regimen.[3,21] INH and rifampicin are also used in preventative therapy for high risk individuals over periods of 3-6 months.[3]

While current first line antibiotics have been sufficient treatment in previous decades, they are becoming less effective due to the emergence of anti-microbial resistance (AMR), first recorded in the 70s.[21] AMR is classified in 3 ways by the WHO: Multi-drug resistant TB (MDR-TB), rifampicin resistant TB (RR-TB) and extensively drug resistant TB (XDR-TB). MDR-TB strains generally tolerate first line treatments and are described by resistance to at least isoniazid and rifampicin.[28] Survival rates can be as low as 50% in patients exhibiting MDR-TB and 206,000 cases were diagnosed in 2019, an increase of 10% from 2018.[3,21] RR-TB has become an important marker in MDR diagnosis as it is almost always accompanied by resistance to other drugs, especially INH.[27,29] XDR-TB combines the traits of MDR-TB with resistance to at least one fluoroquinolone (**5**, **6**), inhibitors of DNA separation and cell replication, and one of the injectable drugs (**7**, **8**), aminoglycoside inhibitors of protein synthesis through blockage of the 30s ribosomal subunit (**Figure 5**).[27] 12,350 cases of XDR-TB were diagnosed in 2019.[3]
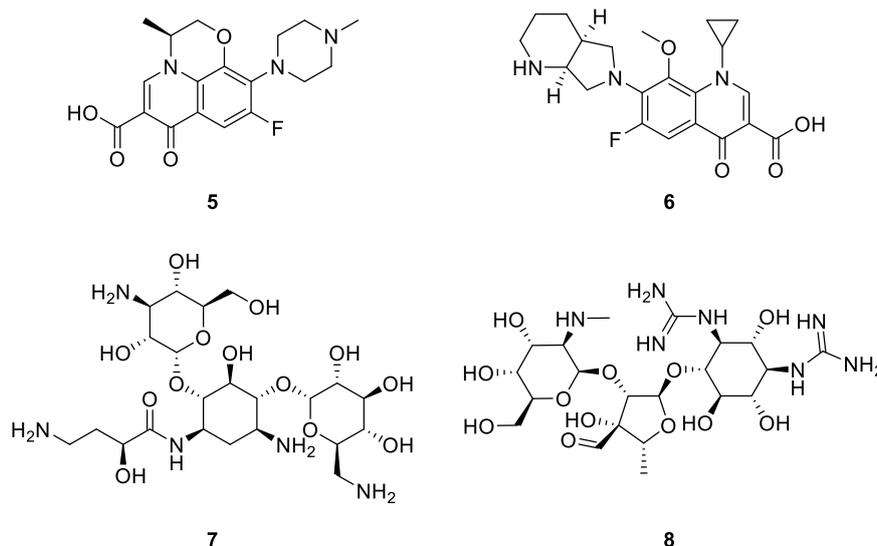
Figure 5: Further drugs involved in the classification of MDR TB. Top: fluoroquinolones levofloxacin (**5**) and moxifloxacin (**6**). Bottom: injectable drugs amikacin (**7**) and streptomycin (**8**).

The production of a treatment plan for MDR/XDR-TB patients from second line therapeutics has been laid out by the WHO in **Table 2**. These drugs generally exhibit greater toxicity, harsher side effects, higher cost or reduced efficacy during their 6 - 12 month regimes.

Table 2: Current second line treatment of drug resistant TB as outlined by the WHO.[29]

Groups and usage	Drug/Class
Group A: Include all three	fluoroquinolones bedaquiline linezolid
Group B: Add one or both	clofazimine cycloserine
Group C: Added to replace members of group A or B to ensure four active drugs are provided.	delamanid β-lactams injectable drugs ethionamide *p*-aminosalicylic acid

MDR-TB treatment plans start with combinatorial therapy of three group A drugs: bedaquiline (**9**, **Figure 6**), linezolid (**10**, **Figure 6**), and a fluoroquinolone and either of the group B drugs: clofazimine (**11**, **Figure 6**) or cycloserine (**13**, **Figure 7**). If resistance to any of these drugs is observed, it is replaced first by the other member of group B and then by any of the members of group C until a therapy containing at least 4 active drugs is achieved. This grouping system is generally constructed so that the more toxic, non-specific, or low efficacy drugs are prescribed as a last resort.

Only four novel drugs have been developed since the 70s: bedaquiline, linezolid, clofazimine and delamanid (**12**) and only linezolid has fully completed phase III trials (**Figure 6**).[3] Bedaquiline is an ATP-synthase inhibitor planned to form the basis of future MDR-TB treatments alongside linezolid within group A.[3,27,29] Linezolid is a broad spectrum oxazolidinone which inhibits protein synthesis through interaction with the 23S ribosomal subunit. Although a keystone treatment of future MDR-TB plans, there is evidence of linezolid increasing risk of nerve damage, decreasing bone marrow activity and decreasing platelet count.[30] Clofazimine's mechanism of action is not yet fully understood but it is expected to disrupt the respiratory chain at the membrane bound protein NADH dehydrogenase II.[31] Originally an anti-leprosy drug, clofazimine has been repurposed to treat MDR-TB. Although the exact target is not yet known, delamanid is a prodrug, activated by deazaflavin (F420)-dependent nitroreductase (DDN) to inhibit production of mycolic acids required as part of the cell wall.[32,33] There is evidence of delamanid prolonging the QTc interval, indicating slow repolarisation of cardiac tissue between heartbeats. As this can lead to life-threatening heart arrhythmia, use of delamanid is controlled as part of group C.[33]
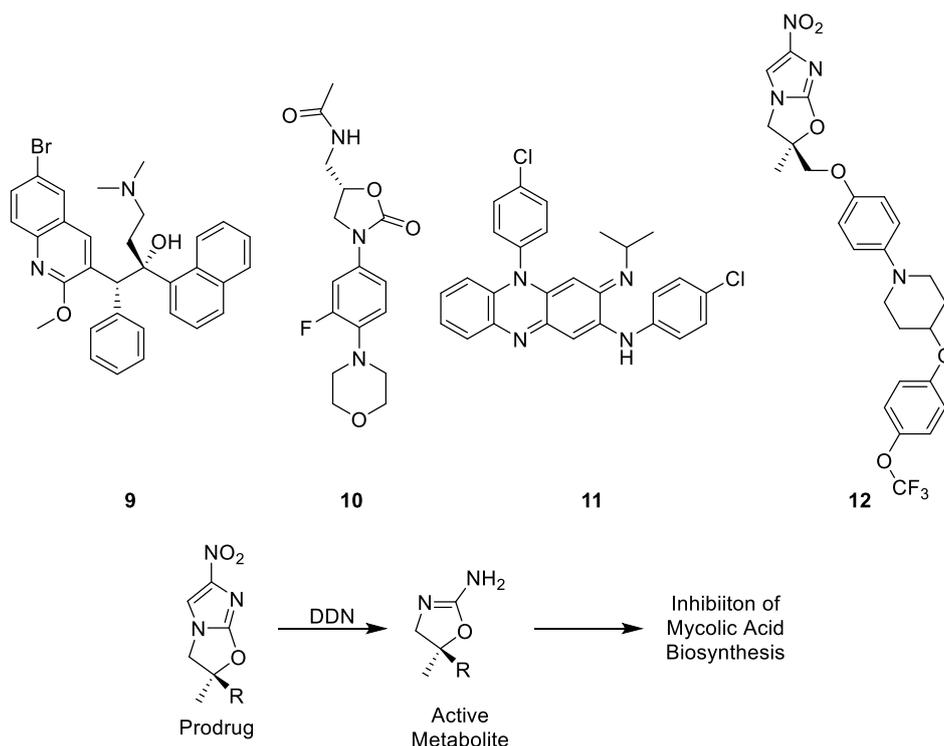
Figure 6: MDR-TB therapies may now also include modern drugs such as bedaquiline (**9**), linezolid (**10**), clofazimine (**11**) and delamanid (**12**). The DDN-mediated transformation of delamanid to its active metabolite is shown below. DDN: deazaflavin (F420)-dependent nitroreductase.

The remaining therapeutics: cycloserine (**13**), β-lactam iminipenem (**14**), ethionamide (**15**) and *p*-aminosalicylic acid (**16**) are pictured in **Figure 7**. Cycloserine is a covalent inhibitor of alanine racemase (Alr), a key enzyme in peptidoglycan (PG) synthesis.[34,35] Psychosis due to partial agonism of the drug with the mammalian neuronal *N*-methyl-D-aspartate receptor relegates its use from frontline treatments.[36,37] β-lactams are broad spectrum antibiotics rarely used in the treatment of Mtb due to the expression of β-lactamases but some examples such as imipenem can prove effective in XDR-TB treatment.[38] By inhibiting bacterial transpeptidases, β-lactams can inhibit production of PG. Although these β-lactams can be used in treatment they still have short half-lives and imipenem must be co-dosed with a dyhydropeptidase inhibitor to reduce breakdown in humans.[20,26,29] Ethionamide is also a prodrug inhibiting InhA through a complex with NAD[+] but is activated instead by the monooxygenase EThA and is occasionally used to replace INH when KatG mutation is observed.[39] Ethionamide treatments are frequently accompanied by gastrointestinal intolerance.[39] *p*-Aminosalicylic acid combats Mtb by

inhibiting the folate pathway responsible for *Mycobacterial* nucleic acid synthesis and has been used as a secondary treatment for TB for decades.[40]
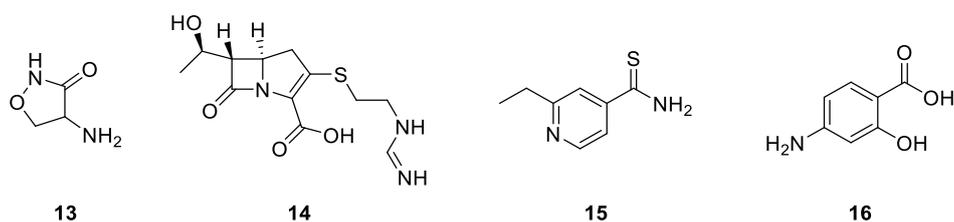
Figure 7: Other second line anti-TB treatments cycloserine (**13**), iminipenem (**14**), ethionamide (**15**) and *p*-aminosalicylic acid (**16**).

Anti-TB chemotherapy requires multiple medications over long periods of time with often unpleasant side effects. This means that patient compliance can be difficult to ensure when symptoms begin to ease. Poor patient compliance is thought to have contributed significantly to the prevalence of AMR in TB.[41] To combat this, the WHO instituted the directly observed treatment, short course (DOTS) program in 1994. The program aimed to standardise treatment regimens and ensure that patients undergo full courses of treatment witnessed by medical personnel. The program was replaced in 1999 by DOTS-Plus, which added treatment guidelines for MDR-TB and HIV/TB coinfection. Although the standardisation and attempted shortening of treatment length has been generally accepted as beneficial, the effectiveness of witnesses being present for administration has been called into question.[29,42]

As resistance to frontline therapeutics accumulates over time, patent protection for new drugs can run out before they see widespread use. This limited window for recouping makes it an unappealing target for drug companies. This lack of industrial support coupled with the difficulty of bypassing the mycobacterial cell wall and a view of TB as a "third world disease" has led to decreased interest in drug production. Currently there are only seven drugs of a new chemical class in the pipeline and those are in early Phase I and II clinical trials.[3] New therapeutics are urgently required to either expand the diversity of druggable targets in TB or replace drugs made ineffectual through resistance.

1.2 Mtb Cell Wall

The structure of the Mtb cell wall presents a significant physical barrier to treatment, increasing the difficulty of accessing potentially sensitive targets within the pathogen (**Figure 8**). The cytoplasmic membrane (CM) is protected by a peptidoglycan (PG) layer, a similarity with Gram-positive bacteria. PG is in turn connected through an acetylglucosamine/rhamnopyranose linker to a series of 30 galactofuranose sugars at 10-12% coverage.[43] Every eighth, tenth and twelfth galactofuranose is in turn connected to arabinofuranose sugars which are further grown and split to form a web.[44] This arabinogalactan (AG) layer provides a dense hydrophilic layer to support the cell wall.[45] A protective mycolic acid (MA) bilayer is then attached to the ends of arabinofuranose chains to form the outer cell wall. Hydrophobic molecules generally enter the cell by passive diffusion through the mycolate bilayer whilst hydrophilic compounds gain access through porins.

Mycolate comprises approximately 60% of the cell wall and 50% of Mtb cell dry weight and provides resistance to dehydration and reactive oxygen molecules.[46] This has, however, translated into long doubling times of around 24h when compared to bacteria such as *Escherichia coli,* which replicates over 20 times faster.[47] Inhibition of mycolate biosynthesis to treat TB has been clinically validated and the frontline drug INH has been a vital component of anti-TB treatment.

Fatty acids are essential building blocks for membrane synthesis in all cells barring *Archea*.[48] They are produced either by the Fatty Acid Synthase I (FAS I, mammals and most animals) or II (FAS II, plants and most bacteria) pathways. The two systems function in broadly the same way in so far that individual enzymes of the FAS II pathway can occasionally be superimposed upon domains of the FAS I superstructure. This has helped in solving the structure of the FAS I counterpart to the FAS II ketoacyl-ACP synthase.[48,49] Mtb is capable of utilising both FAS systems (**Scheme 1**).
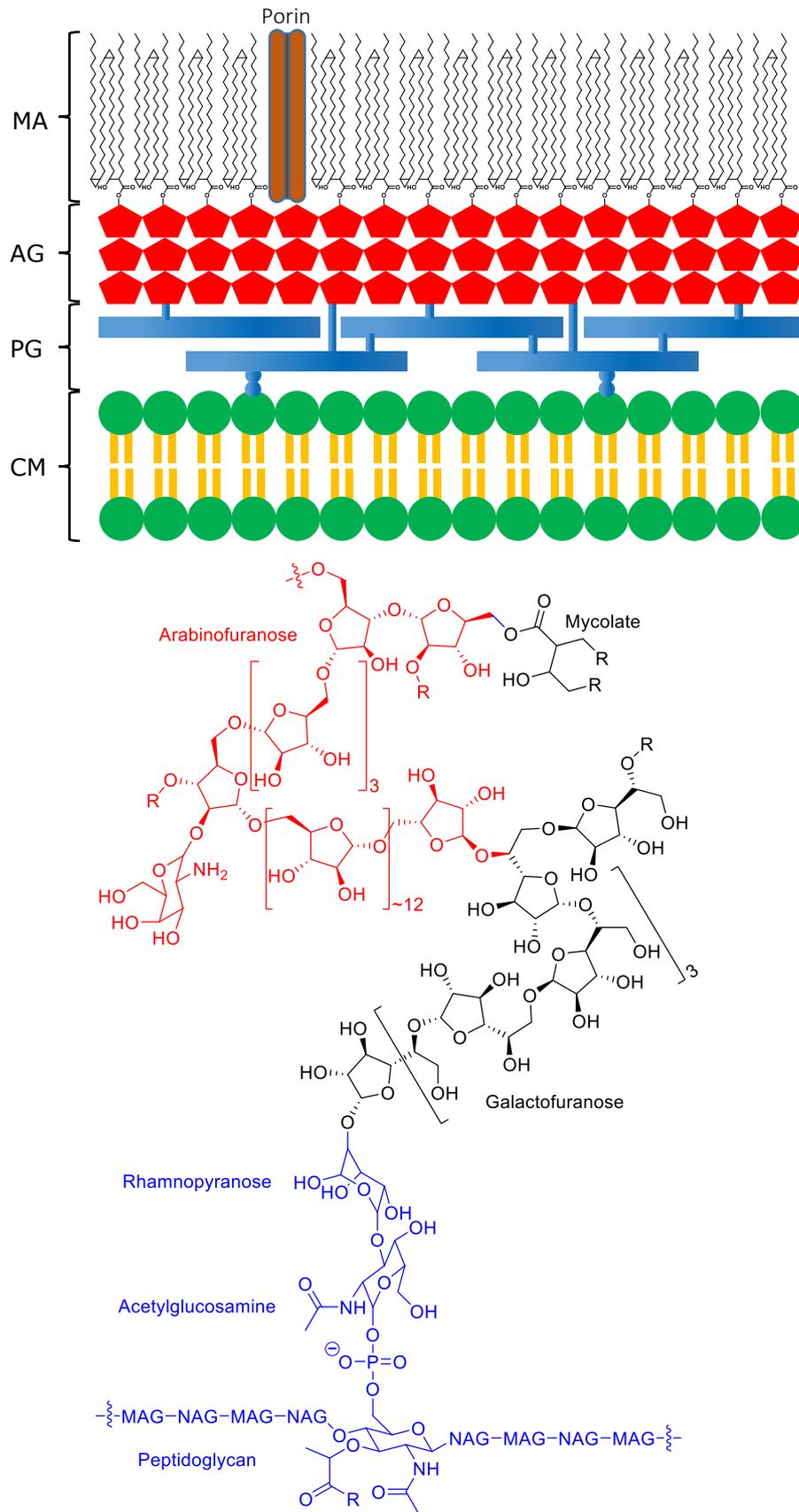
Figure 8: Simplified image of the Mtb cell wall (top) and chemical structure (bottom).[6–8,50] MA: mycolic acid, AG: arabinogalactan, PG: peptidoglycan, CM: cytoplasmic membrane, NAG: *N*-Acetyl Glucosamine, NAM: *N*-Acetylmuramic acid.

15
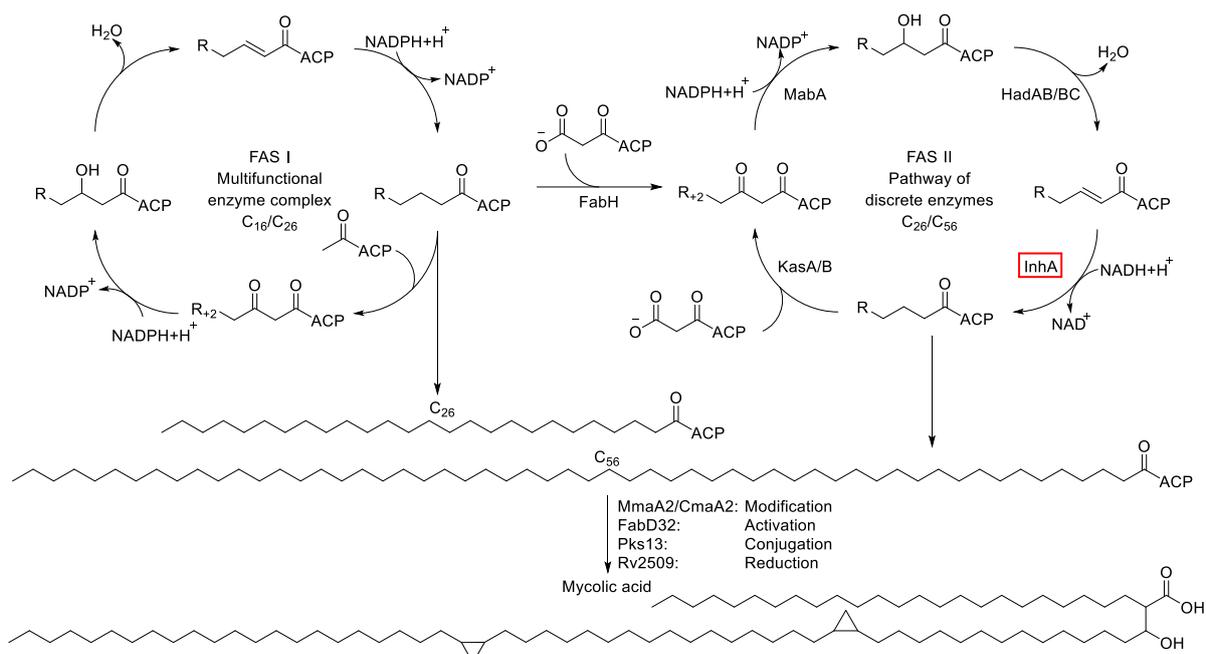
Scheme 1: Mtb mycolic acid biosynthesis *via* the FAS I and FAS II pathways adapted from Abrahams and Besra.[51] ACP: acyl carrier protein.

Short chain fatty acids produced by the FAS-1 enzyme complex can be used to form the α-branch of the finished mycolic acid or go on to be extended through the FAS II cycle to form longer meromycolate chains.[52]

For meromycolate synthesis, the acyl-CoA FAS-I products are fused with an acyl-carrier-protein (ACP) bound dicarbonyl, catalysed by β-ketoacyl-acyl carrier protein synthase III (FabH).[53] The ketone is then reduced with NADPH-dependant keto-acyl-ACP reductase (MabA) prior to loss of water with β-hydroxyacyl-ACP dehydratase (HadAB/BC).[54,55] Saturation of the alkene is carried out by *trans*-enoyl-acyl carrier protein reductase (InhA) to complete the cycle.[56] Further chain extension is facilitated, two carbons per cycle, with the addition of more malonyl-ACP catalysed by the keto synthases KasA (elongating chains to an average of C_{40}) and KasB (elongating chains to an average of C_{54}).[57] Completed mycolic acids can then be modified through cyclopropanation by methyl transferases MmaA2 and CmaA2.[58] The short chain FAS I product is then activated by fatty acyl-AMP ligase FabD32, the two chains condensed by polyketide synthase 13 (Pks13) and the resultant ketone reduced to an alcohol by Rv2509 to afford the final mycolic acid.[59–61]

1.3 InhA
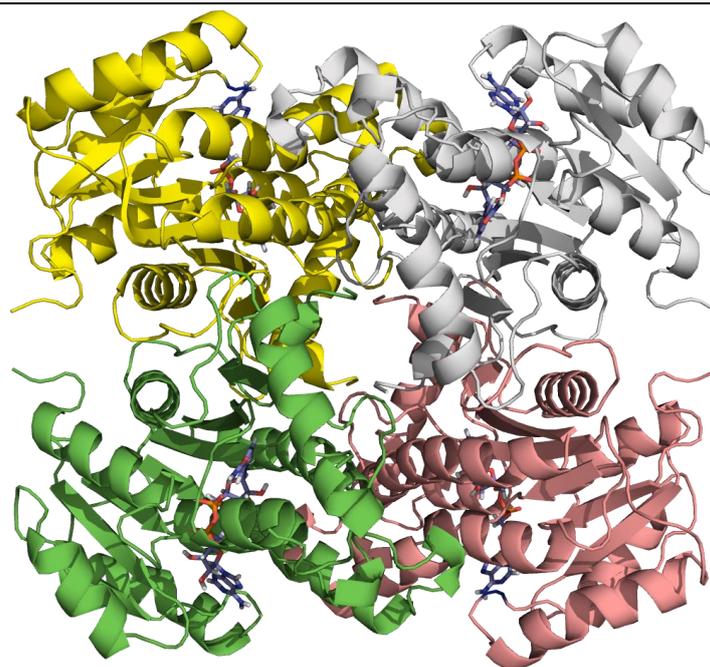
Figure 9: Full tetrameric structure of InhA with NAD$^+$ (PDB: 4U0J, 1.62 Å).[62] Visualised in PyMOL$^©$.[63]

Enoyl-acyl carrier protein reductases (ECPRs) are poorly conserved across those species utilising FAS I or II pathways, making them ideal targets for selective inhibition.[49]

InhA (**Figure 9**) is a 28.5 kDa tetrameric protein present in Mtb and Mb and is comprised of four subunits, each measuring 269 amino acids in length.[62] It catalyses the final step of mycolic acid chain extension in the FAS II pathway through NADH (**17**) dependent saturation of a double bond (**Scheme 2**).
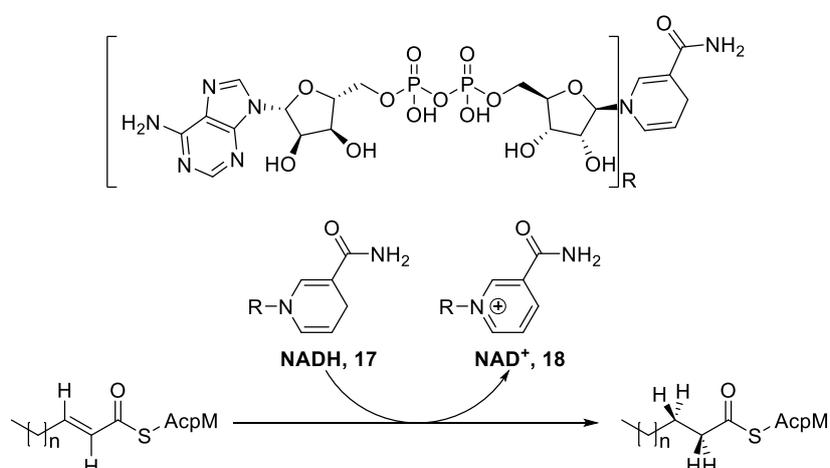
Scheme 2: Reduction of InhA natural substrate with NADH facilitated by InhA[62]

There are multiple key areas in InhA interacting with substrate and cofactor (**Figure 10**). Tyr158, Phe149 and Lys165 are important catalytic residues, correctly orienting cofactor and substrate to facilitate reduction.[64] Phe149 coordinates to the NAD+ nicotinamide ring and C2-3 olefin whilst Lys165 interacts with NADH ribose through interaction with a secondary alcohol. Tyr158 acts to stabilize the enolate intermediate produced after the initial protonation. NADH is bound within the Rossman fold structure of InhA, reminiscent of other short chain dehydrogenase/reductases and is comprised of eight α-helices and seven β-sheets.[65] To facilitate the presence of long, hydrophobic carbon chains, a series of hydrophobic residues from positions 196-219 form the substrate binding loop, acting to stabilise the substrate's long carbon chains.[66] Two conformational changes occur upon ligand binding. The catalytic residue Tyr158 undergoes a 90° rotation of its side chain and the substrate binding loop undergoes a slight shift. Both changes serve to widen the binding pocket and provide a deeper substrate cavity than other dehydrogenases for the comparatively longer fatty acids it processes.[66]
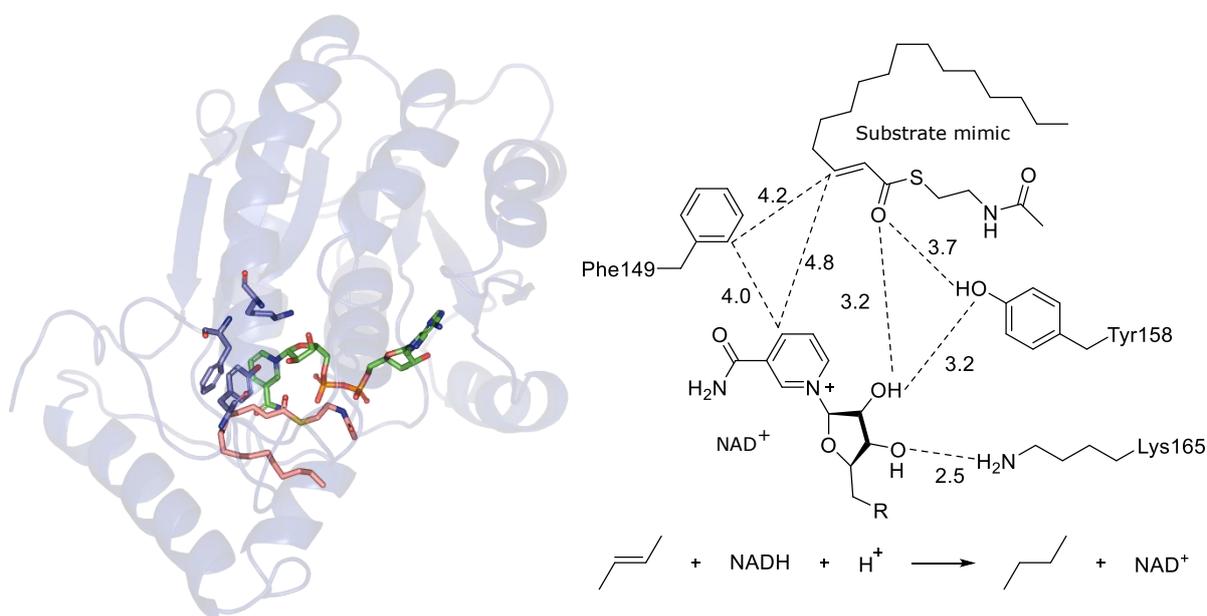
Figure 10: InhA active site displayed in PDB entry: 1BVR (left, 2.80 Å)[66] and in skeletal structure (right). Catalytic residues Phe149, Tyr158 and Lys165 are in blue, NAD+ in green and substrate mimic in beige. Distances are measured in Angstroms and images were prepared in PyMOL©.[64,66,63]

Targeting InhA

One of the best-known InhA inhibitors and a potent front-line TB drug is INH. INH was discovered in 1952 and used almost immediately in combination with streptomycin and *para*-aminosalicylic acid to treat TB.[67,68,69] Combinational therapy was essential as resistance to INH developed rapidly in the clinic.[70] Even with this initial regimen and combination with rifampicin, pyrazinamide and ethambutol, resistance was observed widely in 1985.[22]

The INH prodrug (MIC: 0.03 µg/mL) is initially activated by catalase-peroxidase KatG, an enzyme responsible for protection of actively dividing bacteria from the reactive oxygen species produced by phagocytic action (**Scheme 3**).[71] The isonicotinic acyl radical produced (**19**) is quenched by NAD+ (**18**) to form the active metabolite (**20**).
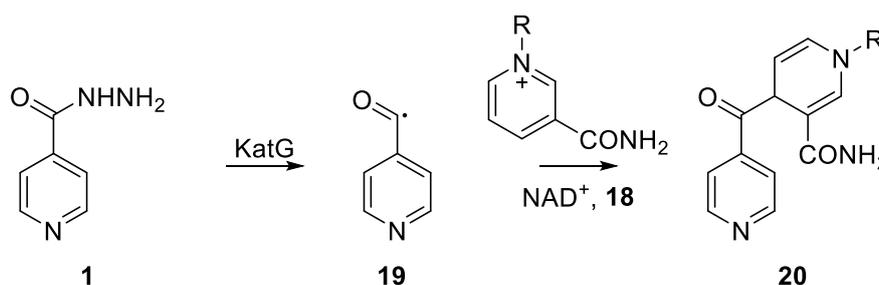
Scheme 3: Isoniazid prodrug activation and attachment to NAD+[72]

INH Resistant strains of Mtb can feature mutation in the genes of both KatG and InhA. The largest contributor by far is KatG with one mutation in particular accounting for up to 94% of resistance in clinical isolates: S315T.[22] In the mutant, a new hydrogen bond is thought to form between Thr315 and INH. This change carries with it a 20-fold reduction in turnover of the prodrug with only a mild drop of catalase-peroxidase activity.[73]

INH resistance arising from InhA mutation is low level, with the three most common being S94A, I21V and I47T.[74] Each one is situated in the NADH binding region and upsets hydrogen bonding, reducing affinity for both NADH and INH-NADH. As each of these mutants are of low prevalence and do not appear to affect the structure of the substrate binding site, it is hoped that new, direct inhibitors of InhA will revitalise this clinically

validated target of which there are none in the current TB drug pipeline for future treatment of DS, MDR or XDR-TB.

Direct inhibitors of InhA

Nine different classes of compounds known to directly inhibit InhA will be briefly reviewed, covering natural products, high throughput screening leads and the use of DNA encoded libraries. (**Figure 11**).

The natural product, pyridomicin (**21**, PYR) was isolated and characterised from the bacteria *Streptomyces pyridomyceticus* in 1953.[75] PYR was reported in 2012 to bind InhA from X-ray crystallography and presents a MIC of 0.57 µM.[76] PYR rests in a pocket between both the substrate and cofactor binding sites and has raised interest in exploring this new binding mode through structure-activity relationship analysis (SAR).[76] Attempts have been made to raise PYR-resistant mutants of InhA *in vitro*. Only one resistant mutant was observed at D148G and retains sensitivity to INH whilst PYR is effective against INH-resistant S94A, raising the possibility of PYR/INH combination therapy.[76]

Triclosan (**22**) is a broad-spectrum antibiotic long used in the production of antibacterial soap and toothpaste and bears moderate activity against InhA (MIC: 43 µM). Fragment based drug design has been used to increase its activity against InhA and has produced a number of active compounds, of which triazole **23** had the highest activity (MIC: 5.2 µM).[77,78] Poor solubility and susceptibility towards phase II metabolism is a recurring problem plaguing the series.

4-Hydroxy-2-pyridones (HyPs) are a relatively new series of compounds obtained from whole cell, high throughput phenotypic screening.[79] SAR analysis of the scaffold produced NITD-916 (**24**) as a potent inhibitor of Mtb (MIC: 0.57 µM). HyPs are predicted to block substrate access through interactions with InhA and NADH. H-bonding and π-stacking is observed between inhibitor, NADH pyridine ring and Tyr158. The cyclohexyl moiety makes

use of the hydrophobic substrate binding pocket to further stabilise the interaction. Overall, HyPs show good potency but there is still work required to improve solubility and reduce their distinctively high plasma protein binding.

Compound **25** is an advanced iteration of the methyl-thiazole series developed from a GSK HTS lead. It was obtained through modification of functional groups to the left of the secondary amine and has an MIC of 0.19 µM.[80] They bind to both NADH and InhA through Met98 and force Tyr158 to change orientation in a novel binding mode, pushing the active site into disarray.[80]

In one of the first successful attempts at HTS against InhA in 2003, two active classes were found, the piperazine indoleformamides and the pyrazole derivatives.[81] A total of 500,000 compounds were screened and, following SAR development, produced final compounds **26** and **27**. Compound **26** showed good activity with an IC_{50} of 0.16 µM compared to 2.4 µM for **27** but displayed an MIC in excess of 30 µM during whole cell screening, implying poor access to the target. The pyrazoles appear to retain their activity in whole cell testing with **27** giving a MIC of 2.5 µM.[81] **26** binds to NAD^+ phosphate through its indole amine whilst hydrogen bonding through its carbonyl to both an NAD^+ ribose hydroxyl and Tyr158 hydroxyl. The fluorenyl moiety forms hydrophobic contacts to the substrate pocket. The binding mode of **27** has not yet been resolved.

He *et al.* performed a limited HTS of 30,000 compounds, discovering the pyrrolidine carboxamide series, resulting in the synthesis of **28**.[62] Extensive SAR showed that the alkyl and aryl rings were highly modifiable but no changes were tolerated at the lactam core. This was validated during analysis of the crystal structure where hydrogen bonding between the lactam carbonyl and Tyr158 hydroxyl and NAD^+ ribose 2' hydroxyl were observed, reminiscent of the indoleformamides. The compounds were prepared as racemates but chromatographic resolution of the top three performing compounds elucidated the *R*-enantiomer as the eutomer, bearing 10 times the activity of the distomer.

Although derivatives of the series produced IC_{50} results down to 63 nM, MIC values were significantly higher with the lowest at 62 µM.

The imidazopiperidines are a potent class of inhibitors expressing an essential imidazole core.[82] The two enantiomers of **29** were isolated from each other and, though the stereochemistry of each was not recorded, the eutomer had an IC_{50} of 50 nM compared to 240 nM for the racemate and over 10 µM for the distomer. The imidazopiperazidine binding mode remains unexplored and they suffer from MIC results above 30 µM and high *in vitro* liver toxicity.

With the aid of a more modern method of drug discovery, in 2014 Encinas *et al.* utilised a DNA-encoded library (DEL) to explore the proline series at GSK.[83] Starting with 22 diamino acids, including trans-4-aminoproline, the amines were protected orthogonally and a DNA tag attached through the carboxylic acid. At this stage, one amine could be deprotected and coupled to one of 855 building blocks. Following extension of the DNA tag, the process was repeated with the remaining amine with a series of 857 building blocks. The library encompassed over 16 million molecules and was exposed to immobilised InhA where hits were identified through DNA sequencing. Compound **30** proved to be the culmination of this effort, forming a hydrogen bond to the NADH 2' hydroxyl with an admirable IC_{50} of 4.0 nM and MIC of 0.5 µM but proved ineffective in mouse studies.

InhA inhibitors have been shown to interact with catalytically active residues, NAD^+ and hydrophobic residues of the substrate binding loop, making InhA accepting of a range of chemical diversity. Bulky hydrophobic groups also appear accepted in the substrate binding loop, assuming they are tolerated by the compound's pharmacokinetics. Barring pyridomycin and INH-NAD, inhibitors of InhA appear to target only the substrate binding region without blocking NADH access. This makes the majority of InhA inhibitor classes uncompetitive with respect to NADH despite the use of several large compound libraries during HTS.
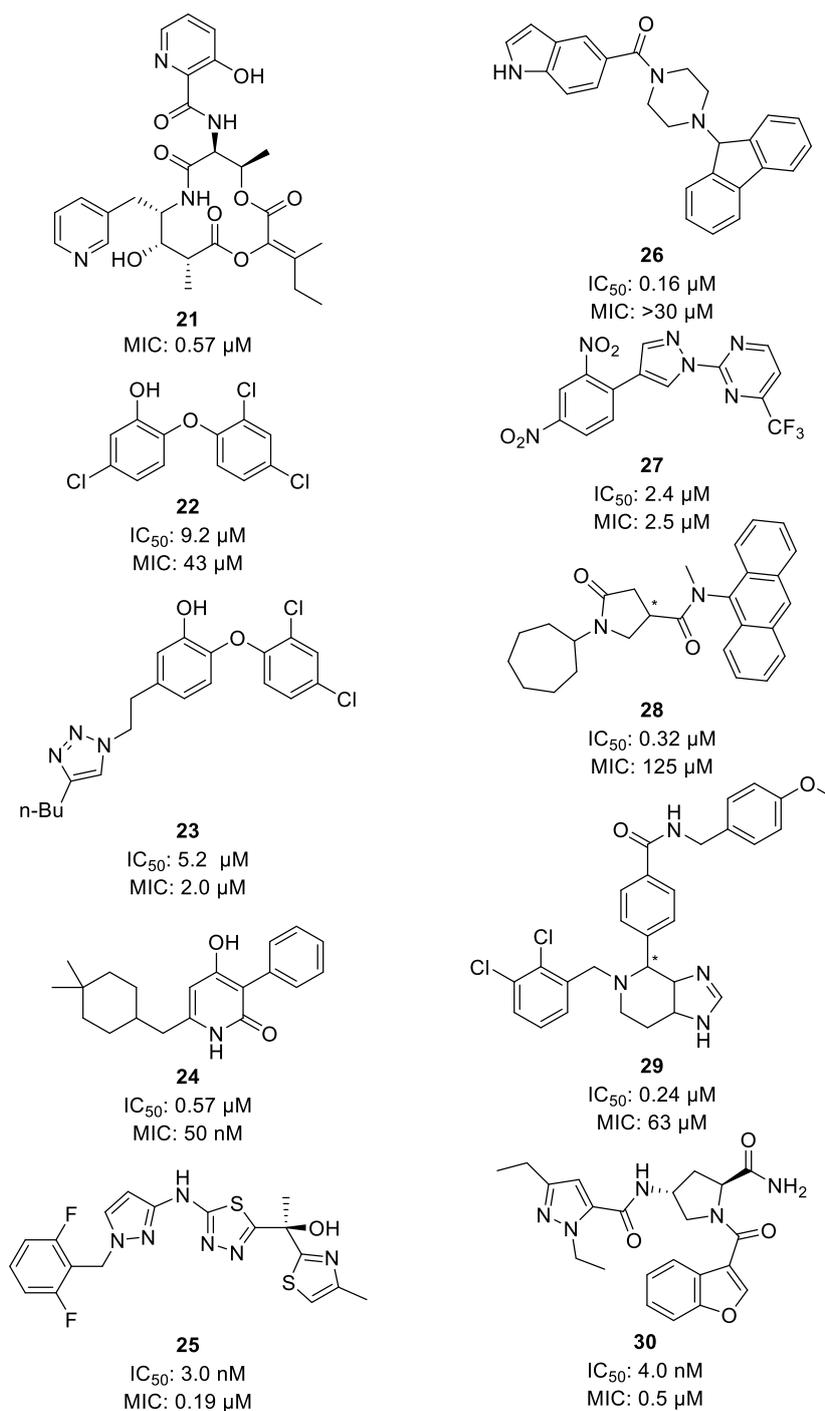
Figure 11: Structures of direct inhibitors of InhA; pyridomycin (**21**),[76] triclosan (**22**),[78] triclosan derivative 6PP (**23**),[77] NITD-916 (**24**),[79] methyl-thiazole (**25**),[80] piperazine indoleformamide (**26**),[81] pyrazole (**27**),[81] pyrrolidine carboxamide (**28**),[62] imidazopiperidine (**29**),[82] and proline (**30**).[83] Racemised chiral centres denoted by *.

Due to the rarity of identifying inhibitors blocking access of NADH to the InhA active site in these projects, attempting to include this region during computational screening may be unfavourable in the identification of new inhibitors.

The results of these studies into direct inhibition have produced a series of potent compounds active against isolated InhA but the difficulties in translating this to the more complex tests of whole cell and animal studies remains apparent. This is evident in the lack of new compounds active against InhA in the current drug development pipeline.[3]

1.4 Aims

InhA is a high priority TB cell wall biosynthesis target with clinical relevance. Currently, treatment regimens for TB include the front line therapeutic INH but this prodrug is losing effectiveness as resistance grows at its site of metabolism to the active form. As InhA is an essential enzyme, structural mutants are low level and direct inhibition remains an attractive target.

This project aims to identify direct inhibitors of InhA. To do this, a virtual library will be screened against InhA to identify high scoring compounds.

This will be satisfied through interrogation of the Nottingham Managed Chemical Compound Collection (NMCCC) using the genetic optimisation for ligand docking (GOLD) software package. The NMCCC is a chemically diverse library containing over 85,000 molecules residing at the University of Nottingham and maintained by the Schools of Pharmacy and Chemistry. The InhA docking structure will be selected from among the highest resolved structures available from the Protein Data Bank.

High scoring compounds will be assessed to ensure reasonable binding poses and then taken forward for screening against InhA in UV/Vis isolated enzyme assays. Any promising candidates will have synthetic methods developed and carried out. These compounds will be retested to confirm their activity with UV/Vis assays and, if possible, orthogonal screening. If any hits are obtained that satisfy these requirements they will then be assessed for chemical derivatisation.

2.0: Virtual screening of the NMCCC

2.1 Introduction

X-ray diffraction (XRD) provides detailed, three-dimensional images of how small molecules interact with protein superstructures. By utilising published XRD data, interactions of other small molecules can be predicted using *in silico* virtual screening methods. For the purposes of structure-activity relationship (SAR) analysis, virtual screening is generally carried out by:

1. Selecting and preparing a receptor
2. Selecting and preparing a ligand set
3. Docking the set into the receptor
4. Post-docking analysis and selection of leads

Receptor selection and preparation

Choosing which receptor structure to use during virtual screening can be impacted by a few considerations: resolution, B-factor and ligand induced conformational changes. The overall resolution of the pool of candidates, especially in well-known biological targets where XRD data has been obtained over decades, is variable and can determine confidence in the protein structure being used. The B-factor is a measure of the rigidity of the site of interest and simplifies binding studies if the site can be considered static. Significant changes occurring upon binding of ligands to the active site can be explored by identifying and comparing available apo- and holo- forms of the enzyme to structures co-crystallised with ligands. Once a candidate has been selected, preparation of the receptor involves:

1. Assigning bond orders and hydrogen atoms
2. Removing non-essential waters
3. Obtaining low energy bond orientations through energy minimisation

Following preparation, a receptor grid can be generated from the structure. The grid holds a set of rules determining how library entries may interact with the receptor. These involve

designating a volume that molecules can dock within the receptor, excluding specific areas within this volume or designating mandatory interactions with the receptor.

Ligand set selection and preparation

Virtual libraries usually store compounds as linear strings in formats such as SMILES or SDF which, when converted into 3D models for use, can then be assigned appropriate stereochemistry, partial charges and ionisation states. Compounds are then minimised to identify their lowest energy state and ready for docking into the receptor grid

Docking

Ligand docking is the most complex part of the process where several key decisions are made. Ligands are first compared with themselves (binding pose prediction) then the remainder of the set (scoring).

Obtaining the most advantageous binding pose of a ligand can be carried out by exploring torsional, translational and rotational structural modifiers, using either a systematic or stochastic conformational search. Systematic conformational searches work by making and comparing small changes to structural parameters, eventually approaching the lowest energy binding pose. Systematic approaches can have poor exploration of the conformational landscape, often being caught in any sufficiently deep local energy well. Stochastic methods take a random structural pose from which a diverse set of poses are developed by randomly changing each parameter. The poses are docked into the receptor grid in successive rounds to find a universal low energy conformer.

Post docking analysis

The low energy binding pose of each library entry can be compared using forcefield, empirical or knowledge-based scoring functions. Forcefields are designed to estimate binding energies by predicting and summing bonded and non-bonded interactions within and between ligand and protein. Empirical scoring functions compare *in situ* results against

binding affinities in the literature. Multiple linear regression analysis is performed on the data set and a model developed against which the experiment can be compared. Compared to forcefield scoring, this method can only be as good as the data used in the standard but is generally faster. Knowledge-based scoring acts similarly to empirical scoring in that data from the experiment is compared to the literature. Instead of binding affinity, interatomic potential energy values are collated and combined by noting how frequently two atoms are found close to each other in the data set. The distances are weighted according to their frequency and the sum of the ligand-receptor potentials produces a score which places between empirical and forcefield methods in terms of speed.

Once scores have been generated for the entire library, the top compounds can be collated for further analysis and confirmed *in vitro*. Binding data can then be used to direct modification of lead compounds to improve their physiochemical properties.

2.2 InhA Structure Analysis and Selection

An InhA protein receptor for use in docking studies would need to fulfill three requirements. It must contain a ligand co-crystallised within the active site to identify the docking area. It must express a low average uncertainty for all atoms within the crystal (high resolution). Finally, it must not display high uncertainty for important residues within the active site (high B-factor).

The Mtb InhA entry on Uniprot references 97 crystal structures available in the RCSB Protein Data Bank (PDB, **Table 3**).[84,85] Filtering for a resolution of below 1.8 Å provided 18 candidates. Of those 18, 16 contained both a NADH cofactor and a separate ligand occupying part of the normal substrate binding site. The number was reduced to 12 entries presenting no other molecules within the structure barring the cofactor, ligand and water. The four most highly resolved complexes remaining were PDB entries: 5G0T (1.54 Å), 4OHU (1.60 Å), 4TZK (1.62 Å) and 4U0J (1.62 Å), the cognate ligands of which (**31** - **34**)

are displayed in **Figure 12**. PDB entry 4U0J was chosen as it has been used during virtual screening to successfully identify inhibitors of InhA in the past.[86]

Table 3: InhA receptor candidate selection.

Selection Criteria	Candidates
All Mtb InhA entries	97
Resolution < 1.8 Å	18
Ligand and NADH	16
Only ligand, NADH, water	12
Four finalists	5G0T,[87] 4OHU,[88] 4TZK,[62] 4U0J[62]
Final candidate	4U0J
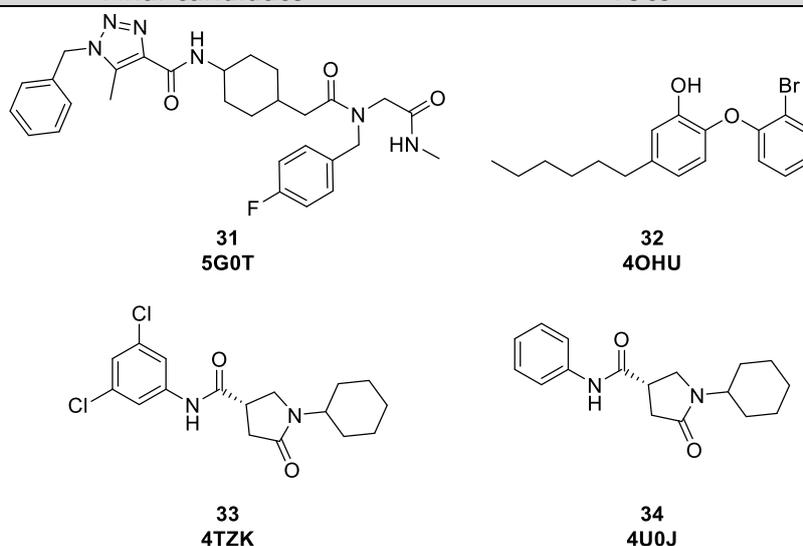
Figure 12: The cognate ligands co-crystallised with entries: 5G0T, 4OHU, 4TZK and 4U0J.

To determine active site stability, 4U0J was coloured by temperature factor to observe the relative uncertainty of each atom (**Figure 13**). Temperature factors below 40 Å^2 were viewed as confident positions whereas those greater than 60 Å^2 were viewed as disordered positions. The catalytic residues in the active site were highly ordered with temperature factors less than 30 Å^2. Sites **A** and **B** correspond to binding loops important for coordination to other InhA subunits and were expected to display slightly higher temperature factors, though still of an acceptable range.[89] Similarly, the substrate binding loop (site **C**) had a slightly higher temperature factor due to its ability to accommodate varying sizes of substrate. The only sites of high disorder were far from the active site, on

the periphery of the subunit corresponding to the N-terminus (**D**) and a short loop including residues 80-90 (**E**) each of which approached the 60 Å² cutoff.
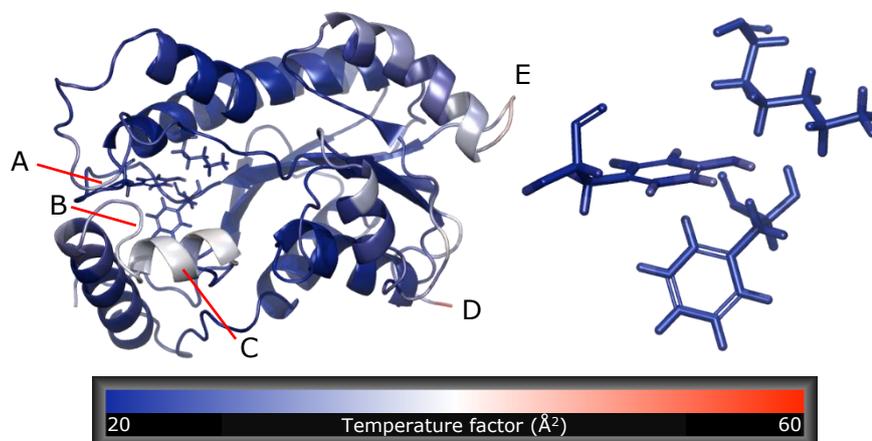
Figure 13: InhA subunit coloured according to temperature factor (PDB: 4U0J, 1.62 Å).[62] Left: entire InhA subunit with active site residues: Phe149, Tyr158 and Lys165. Right: enlarged active site residues. Areas A, B and C are sites of moderate order whereas D and E are of lower order. Structures were explored using PyMOL©.[63]

To observe if any unknown protein-wide changes occur upon binding co-factor or inhibitor, the structures of *holo*-InhA (PDB: 2AQ8, 1.92 Å resolution) and *apo*-InhA (PDB: 4TRM, 1.80 Å resolution) were aligned with 4U0J in place (**Figure 14**).[90,91] *Holo*-InhA and 4U0J were aligned with a score of 0.003 and a root mean squared deviation (RMSD) of 0.284 Å. *Apo*-InhA and inhibited-InhA were aligned with a score of 0.012 and RMSD of 0.546 Å. An alignment score of less than 0.7 is considered a successful alignment. As the larger structure of InhA changes little during occupation of the active site, the protein was considered static during docking analysis.

Figure 14: Left: Aligned structures of *holo*-InhA (PDB: 2AQ8, 1.92 Å, red)[90] and inhibited-InhA (PDB: 4U0J, 1.62 Å, blue).[62] Right: Aligned structures of *apo*-InhA (PDB: 4TRM, 1.80 Å, grey)[91] and inhibited-InhA (PDB: 4U0J, blue). Structures were explored using Maestro®.[92]

2.3 Docking Studies

The Gold suite was chosen for docking as the software has previously been validated with two studies containing 100 and 305 PDB complexes and accuracies of 90 and 70% respectively.[93,94] GOLD has also 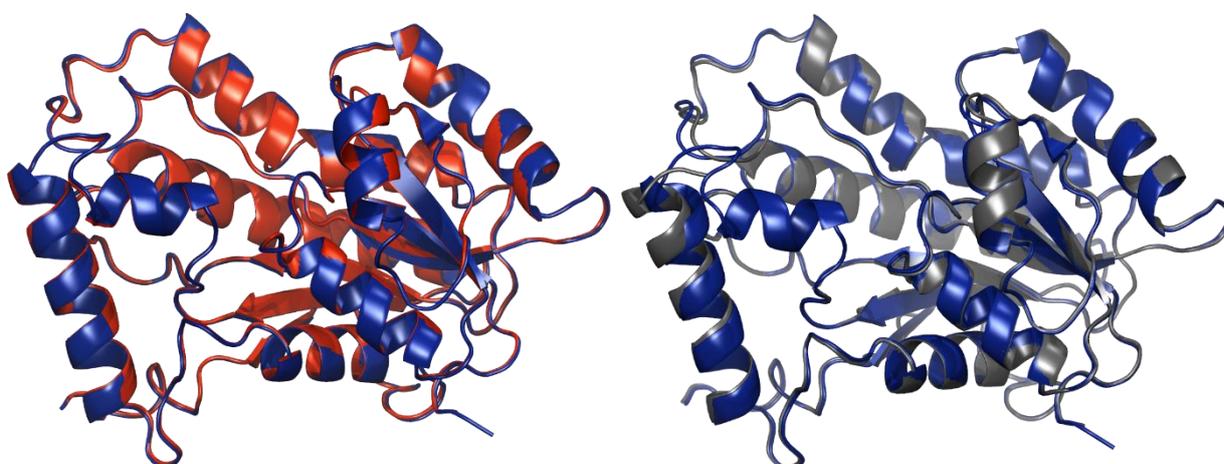been used successfully for screening of compound libraries against InhA in the past.[86,95] GOLD utilizes a stochastic, genetic, conformational search to determine binding poses that can be ranked using the GOLDScore forcefield scoring function (**Equation 1**).

S_{hb_ext}: protein-ligand specific bond energy including ionic, hydrogen and π interactions

S_{vdw_ext}: protein-ligand nonspecific van der Waals energy

S_{hb_int}: ligand internal specific bond energy including ionic, hydrogen and π interactions

S_{vdw_int}: ligand internal nonspecific van der Waals energy and ligand torsional strain energy

$$f = S_{hb\_ext} + 1.3750 \text{ x } S_{vdw\_ext} + S_{hb\_int} + S_{vdw\_int}$$

Equation 1: GoldScore fitness function.

S_{vdw_ext} is multiplied by 1.3750 to increase weighting of hydrophobic interactions during binding and this value has been determined empirically.[96] Docking results were validated for InhA by redocking cognate ligands of the four highest resolved complexes: 5G0T, 4OHU, 4TZK and 4U0J. The results were superimposed and the RMSD calculated (**Table 4**). All four compounds gave accurate docking positions compared to the original structures (**Figure 15**). Only **32** produced an RMSD above 1 Å^2, possibly due to the presence of a six-carbon aliphatic chain attached to one of the aryl rings. If the hexyl-chain is disregarded, alignment produces an RMSD of 0.6875 Å^2.

When each structure was instead docked with 4U0J, **31** and **33** gave GOLDScores within 2 points of the initial value, a reasonable crossover given the stochastic nature of the scoring function. **32**, however, showed a 10-point deficit compared to docking with its home receptor.

Table 4: RMSD and GOLDScores of cognate ligands docked to their own receptors and to 4U0J.

PDB entry	Ligand	RMSD (Å²)	GOLDScore redocked	GOLDScore in 4U0J	Reported IC$_{50}$ (μM)
5G0T	**31**	0.9780	93	91	0.06[87]
4OHU	**32**	1.0228	90	80	0.01[88]
4TZK	**33**	0.7515	81	79	0.39[62]
4U0J	**34**	0.5312	72	-	10.66[62]
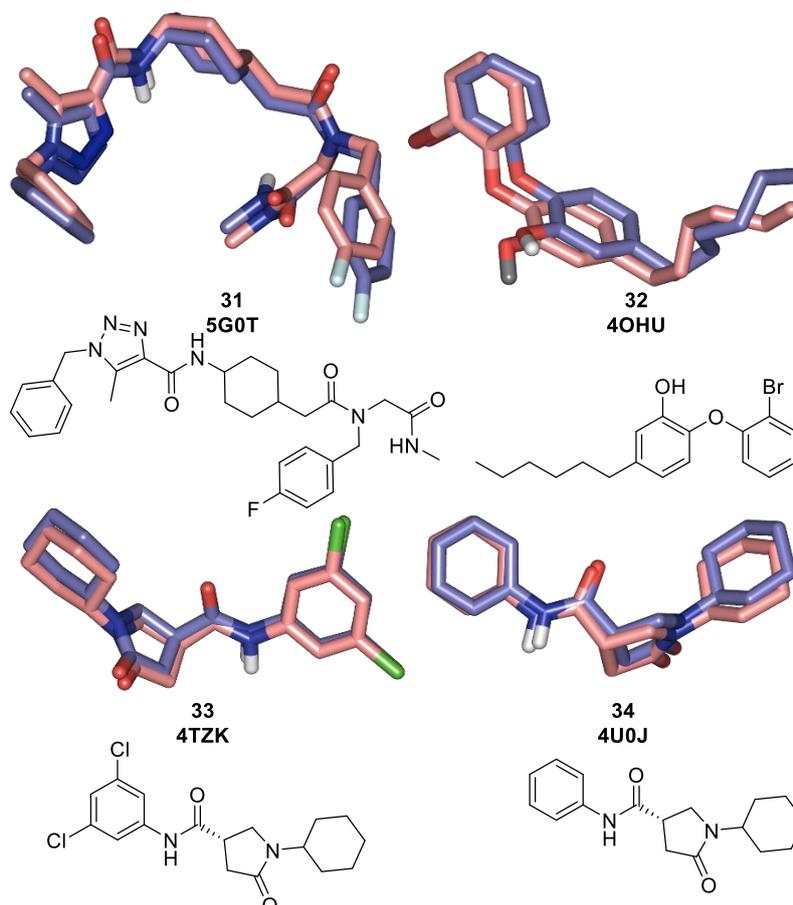
Figure 15: PDB cognate ligands (blue) and alignment with GOLD output pose (beige). Top left: 5G0T, top right: 4OHU, bottom left: 4TZK, bottom right: 4U0J. Structures were explored using Maestro®.

When comparing the fitness functions for both poses, the major component of this anomaly is a large difference in the predicted protein-ligand van der Waals energy (S_{vdw_ext}, **Equation 2**). A significant change in van der Waals energy is not surprising if there is already difficulty in assigning the exact pose for long carbon chains. As exact poses of such flexible chains are evidently hard to predict, it is important to note if any potential leads use them as key spacing groups or in interactions with the receptor.

$$f = S_{hb\_ext} + 1.3750 \text{ x } S_{vdw\_ext} + S_{hb\_int} + S_{vdw\_int}$$

$$\text{Redocked: } 89.66 = 3.87 + 1.3750 \text{ x } 63.23 + 0.00 + \text{-}1.15$$

$$\text{4U0J dock: } 79.83 = 6.00 + 1.3750 \text{ x } 55.26 + 0.00 + \text{-}2.16$$

Equation 2: GOLDScore fitness function for the 4OHU cognate ligand when redocked and when docked to 4U0J.

The University of Nottingham maintains a molecular library of 85,061 entries at the time of writing called the Nottingham Managed Chemical Compound Collection (NMCCC). The library is stored at -20 °C and under a dry, inert atmosphere with HPLC-MS quality control. The NMCCC continuously grows as products from previous drug discovery projects at the University are added, comprising some 3,000 members of the library. Biofocus (now Evotech) provided the remaining entries and have been prescreened to reduce the need for end-user selection:

- No active functional groups as outlined by Brenk *et al.*[97]

- 200 < molecular weight < 500

- AlogP < 4.5

- Aromatic Rings < 4

- Rotatable bonds < 10

- H-bond acceptors < 8

- H-bond donors < 5

- Maximum dissimilarity sampling algorithm ensuring chemical diversity using a feature-connectivity bit string fingerprint for each compound (FCFP4)

The dissimilarity algorithm is a key component of the library, ensuring that compounds occupy a wide area of chemical space despite its relatively small size. The NMCCC has previously been used to successfully identify inhibitors of quorum sensing in *Pseudomonas aeruginosa* and inhibitors of human deadenylase activity.[98,99]

The protein receptor and chemical library were prepared using Schrödinger and docked using GOLD at the fastest setting. The top 2000 compounds by GOLDScore were redocked using the most exhaustive setting and the top 148 compounds ranged from a GOLDScore of 90 – 104 (Supporting information, Inhibition studies and enzyme kinetics, **Table 15**). Four synthetic targets were chosen from among these compounds (**35-38**, **Figure 16**).
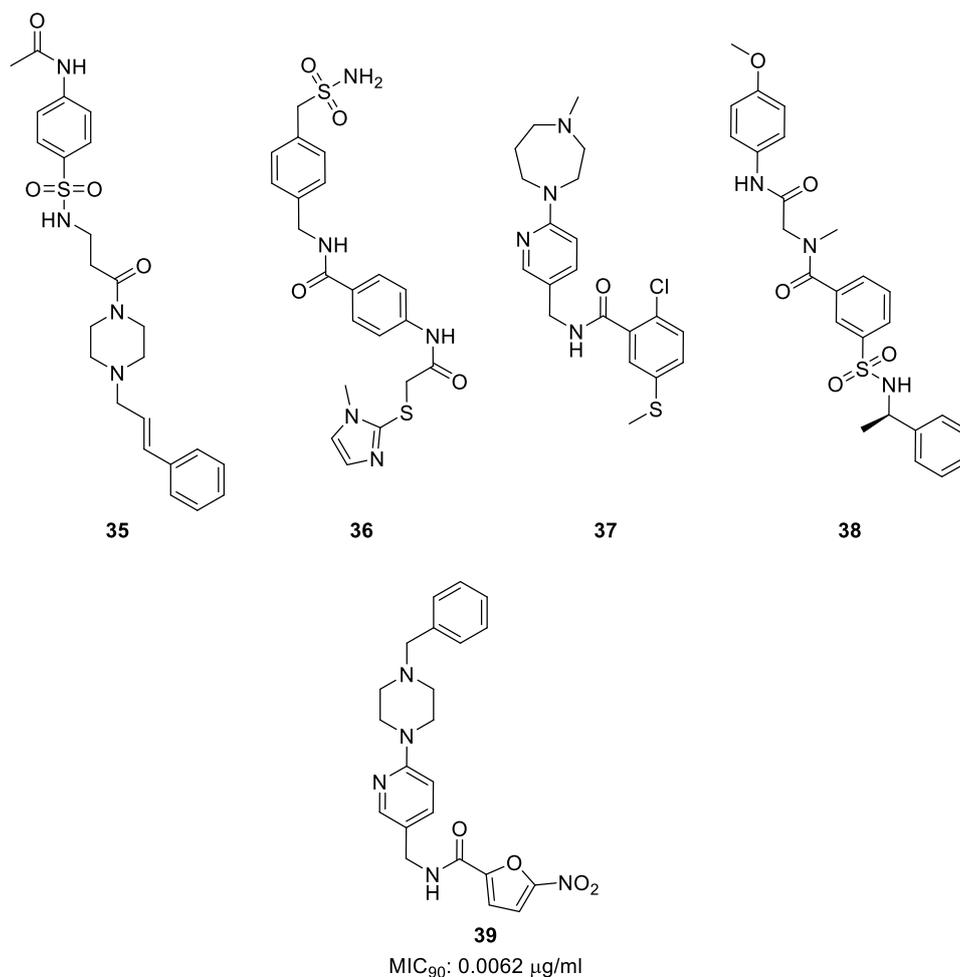
Figure 16: Target compounds chosen for synthesis. Compound **39** is an established whole cell inhibitor of Mtb.[100]

Targets were selected to observe a wide range of binding interactions to the InhA active site including hydrogen bonding, π-π stacking, cation-π and hydrophobic only interactions. All four candidates showed GOLDScores higher than those observed during redocking of cognate ligands (**Table 5**). Finally, three of these compounds were unknown ligands that could be putative inhibitors of InhA whilst **37** showed some similarity to compound **39**, a potent inhibitor of whole cell Mtb which will be discussed in the next chapter.

Table 5: Target compound physiochemical properties.

Compound	Weight	cLogP*	GOLDScore
35	471	1.58	104
36	474	1.16	99
37	405	3.40	96
38	482	3.00	96

*cLogP calculated through the JChem plugin for Office.[101]

2.4 Summary

PDB entry 4U0J was selected and validated for use in virtual screening studies to identify inhibitors of InhA. The NMCCC was used as a source of chemical diversity and docked into 4U0J in place of the cognate ligand. Four candidates were selected for synthesis on the grounds of high GOLDScore, varied originality and a wide range of interactions to the active site. The next chapter will go into the predicted binding modes of each compound, their synthesis and their biological evaluation.

3.0: Synthesis and evaluation of InhA inhibitor candidates identified during virtual screening.

3.1 Compound 35

Compound **35** was the highest scoring of the NMCCC, exhibiting a GoldScore of 104 (**Table 5**) whilst making no predicted hydrogen bond to active site residues (**Figure 17**). The key expected interactions take the form of a double π-stack, orienting the internal benzylsulfonamide ring between two InhA residues (Tyr158 and Phe149). The remaining interactions are predicted to be mostly hydrophobic/van der Waals.
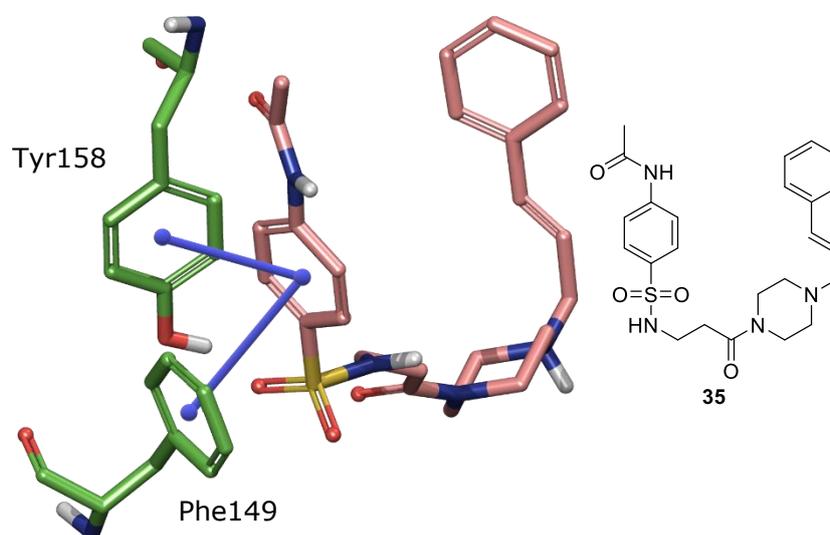

Figure 17: Docking model of **35** (peach), Tyr158 and Phe149 (green, PDB: 4U0J) and skeletal structure of **35**. π-stacking is represented by blue lines.

Synthesis

Due to the modular nature of the target a retrosynthesis proved straightforward to develop, requiring cleavage at the sulfonamide, amide and tertiary amine (**Scheme 4**).

Scheme 4: Compound **35** retrosynthetic analysis.

The synthesis proceeded as expected *via* **Scheme 5** with *boc* protection of piperazine. A reductive amination with cinnamaldehyde facilitated by sodium triacetoxyborohydride then furnished **42**. *Boc* deprotection with TFA produced **43**, allowing amide formation with the aminium based peptide coupling reagent HCTU to produce **44**. Following a second deprotection, a final sulfonamide coupling furnished 24.1 mg of **35** after HPLC purification.

Scheme 5: Forward synthesis of **35**.

Reagents and conditions: i) Boc_2O, CH_2Cl_2, 16 h, 81%. **ii)** Cinnamaldehyde, $NaHB(OAc)_3$, AcOH, 1:3 THF:CH_2Cl_2, 20 h, 95%. **iii)** 1:5 CH_2Cl_2:TFA, 5 h, 99%. **iv)** Fmoc β-alanine, HCTU, K_2CO_3, CH_2Cl_2, 24 h, 96%. **v)** Dimethylamine, EtOH, THF, 1.5 h, 85%. **vi)** 4-acetamidobenzenesulfonyl chloride, NEt_3, DMF, 16 h, 14%.

3.2 Compound 36

Compound **36** produced a GoldScore of 99, the fourth highest predicted binder. The π-interactions visible in **35** are again visible here along with hydrogen bonding to both the essential Tyr158 and to Pro156 (**Figure 18**). At the time of writing there is no evidence of Pro156 being a site of mutation in any clinical strains of TB. There is also a predicted hydrogen bond to the NADH cofactor at the imidazole proton.
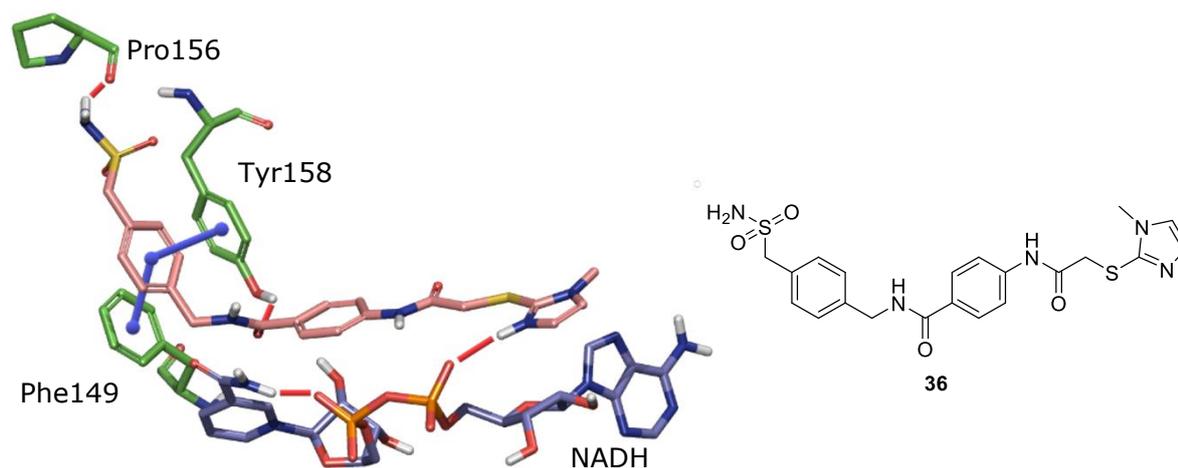

Figure 18: Docking model of **36** (peach), Tyr158 (green), Phe149 (green), Pro156 (green) and NADH (blue, PDB: 4U0J) and skeletal structure of **36**. Hydrogen bonding is represented by red lines, π-stacking is represented by blue lines.
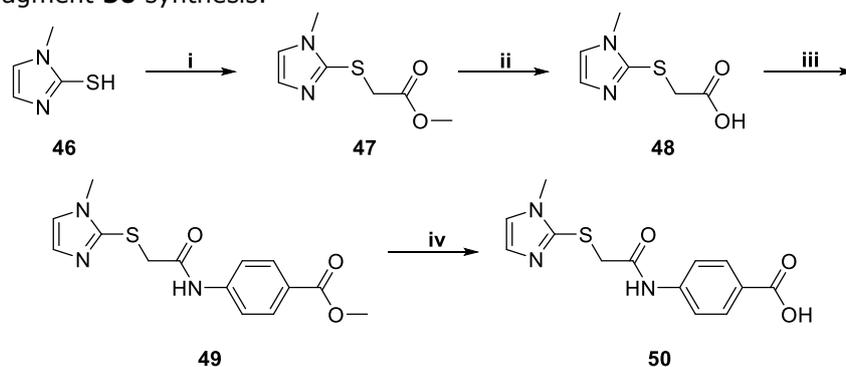
Synthesis

Compound **36** was initially cleaved at an amide bond to facilitate a convergent synthesis (**Scheme 6**). The resulting sulfonamide can be readily prepared from a free thiol in a one-pot synthesis. To go back to a readily available starting material, the thiol would in turn have to be installed in place of a leaving group obtained by treatment of a free alcohol. The alcohol can then be obtained from the readily available *p*-(aminomethyl)benzoic acid. The remaining fragment can then be disconnected at the amide and sulfide bonds.

Scheme 6: Compound **36** retrosynthetic analysis.
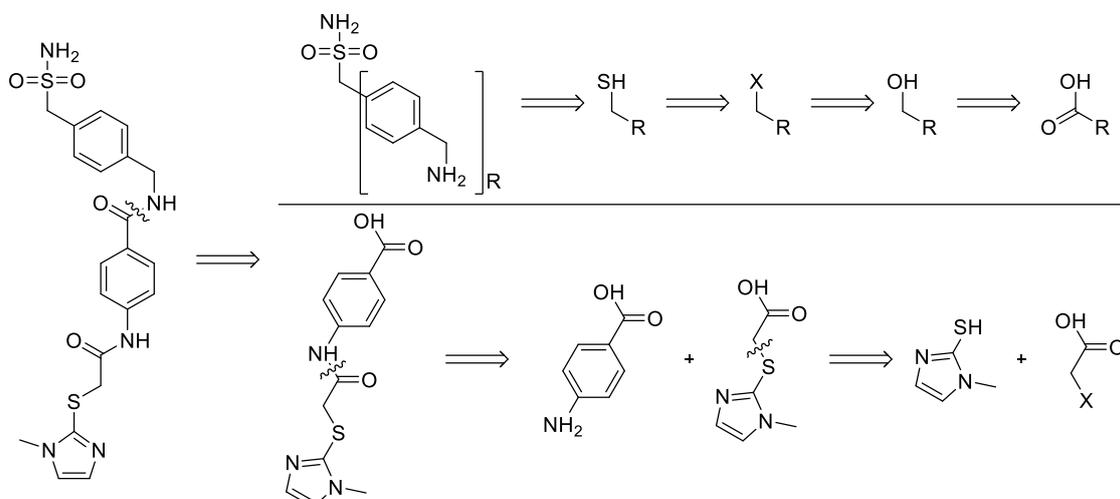
The synthesis of fragment **50** (**Scheme 7**) proceeded readily by initially treating methyl thioimidazole with sodium hydride before introduction to methylbromoacetate to furnish **47**. Hydrolysis of the methyl ester allowed peptide coupling to methyl 4-aminobenzoate, furnishing **49** in modest yield. This is most likely due to the amine's poor lone pair availability and steric bulk. **50** was released by ester cleavage with LiOH.

**Scheme 7:** Fragment **50** synthesis.

Reagents and conditions: i) methylbromoacetate, NaH, THF, 16 h, 64%. **ii)** NaOH, 1:4 H_2O : THF, 2 h, 100%. **iii)** methyl 4-aminobenzoate, EDC.HCl, DMAP, CH_2Cl_2, 16 h, 41%. **iv)** LiOH, H_2O, 16 h, 98%.

Compound **58** was successfully synthesized following **Scheme 8**. Aminomethyl benzoic acid (**51**) was reduced to the corresponding alcohol **52** which, following Cbz protection, was converted to chloride **54** *via* an Appel reaction. The free thiol (**56**) was introduced through displacement of chlorine with thioacetic acid and subsequent cleavage with hydrazine monohydrate. The thiol oxidation and sulfonamide coupling was achieved in a one-pot reaction previously described by Veisi *et al.* to produce **57** with a 44% yield.[102] The proposed mechanism is given in **Scheme 9** in which molecular chlorine is generated by two of the starting materials. This is predicted to allow production of a sulfenyl chloride, creating a target for water to attack and displace chloride which is in turn displaced by a second thiol. The resulting disulfide may undergo dual oxidation facilitated by molecular chlorine to produce an unstable intermediate which collapses to the desired sulfonyl chloride. Introducing the target amine then allows substitution to produce the target sulfonamide **57**. Following Cbz deprotection, **36** was isolated after HCTU coupling of **58** to **50**. HPLC purification provided 10.5 mg material.

Scheme 8: Final synthesis of **36**.
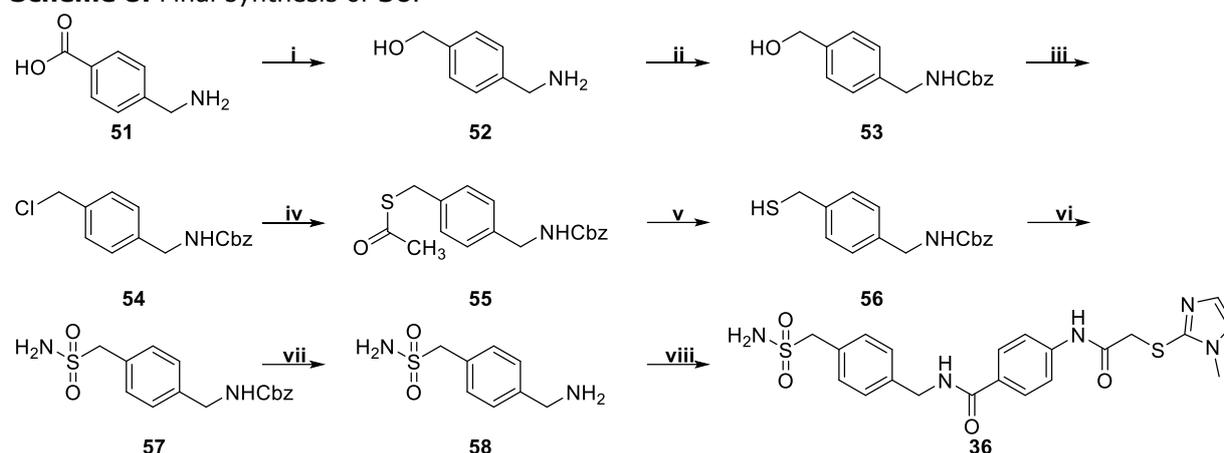
Reagents and conditions: i) p-(aminomethyl) benzoic acid, LiAlH$_4$, THF, 70 °C, 16 h, 83%**. ii)** CbzCl, K$_2$CO$_3$, H$_2$O, 16 h, 90% **iii)** PPh$_3$, CCl$_4$, CH$_2$Cl$_2$, 16 h, 88%. **iv)** Thioacetic acid, NEt$_3$, DMAP, CH$_2$Cl$_2$, 16 h, 61%. **v)** N$_2$H$_2$.H$_2$O, CH$_2$Cl$_2$, 4 h, 35%. **vi)** H$_2$O, TBACl, N-chlorosuccinimide, NH$_4$OH, NH$_4$Cl (sat. sol.), MeCN, 44%. **vii)** Pd(OAc)$_2$, NEt$_3$, TES, CH$_2$Cl$_2$, 59%. **viii)** **50**, HCTU, K$_2$CO$_3$, DMF, 16 h, 5%.
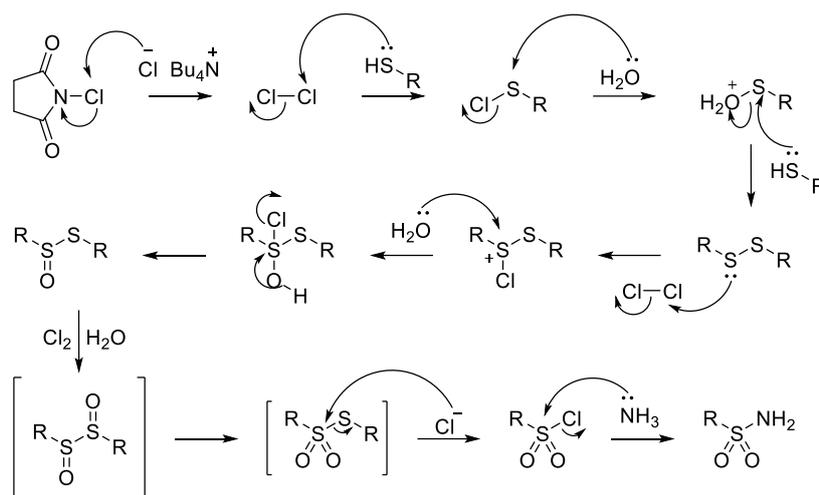
Scheme 9: Proposed mechanism for the one pot functional group interconversion of thiols to sulfonamides as outlined by Veisi *et al*.[102]

3.3 Compound 37

Compound **37** was an attractive target as it not only had a respectable GoldScore of 97, it was also significantly lighter (405 Da) than the three other candidates. This would be beneficial during derivatisation, allowing the addition of more functionality than the other three hits before reaching the Lipinski threshold of 500 Da. Compound **37** is predicted to express one hydrogen bond to the NAD$^+$ cofactor dependent upon the stereochemistry of the protonated amine (**Figure 19**). The *S*- and the *R*-enantiomers were calculated to have GOLDScores of 97 and 86 respectively alongside a corresponding loss of the predicted NAD$^+$ hydrogen bond in the *R*-enantiomer. As rapid pyramidal inversion between these structures will occur at physiological temperatures, they are not truly chiral but comparing the structures may be valuable for the purpose of strengthening the hydrogen bond to NAD$^+$. Replacement of the methyl with hydrogen, for example, or movement of the nitrogen along the ring may place a hydrogen in permanent proximity to NAD$^+$ and allow a persistent hydrogen bond to form.
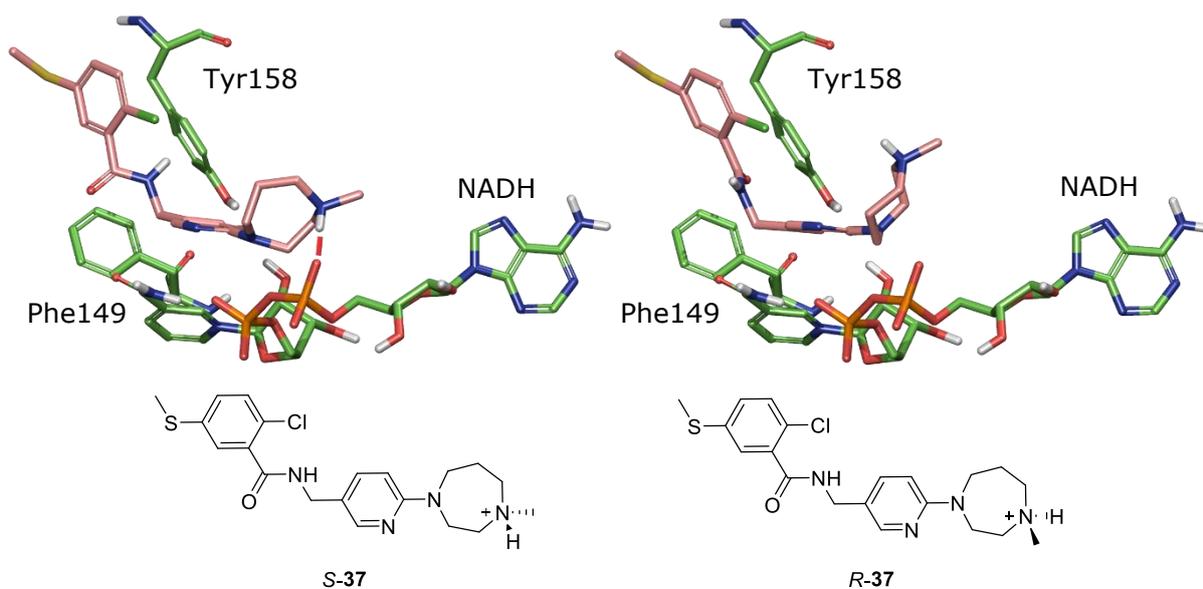
Figure 19: Docking model of protonated **37** *S-* (left) and *R-* (Right) transient enantiomers with Tyr158, Phe149 and NAD$^+$ above enantiomer skeletal structures. Ligands are in peach and InhA residues are in green (PDB: 4U0J).[62] Hydrogen bonding displayed by a red line.

Compound **37** is also of interest as work has been carried out by Tangallapally *et al.* to produce a similar structure containing a nitro-furan (**39**, **Figure 20**).[100] Compound **39** was developed from an early hit (**59**) found during screening against uridine diphosphate-galactopyranose mutase (UGM).[103] UGM is another Mtb cell wall biosynthesis enzyme involved in production of the galactofuranose component of the cell wall. Only **59** was tested directly against UGM with the remaining compounds analysed through whole cell screening and none of the compounds screened directly against InhA. As a result, any *in vitro* inhibition of InhA with compound **37** would also encourage exploration of the nitro-furan series for this target.
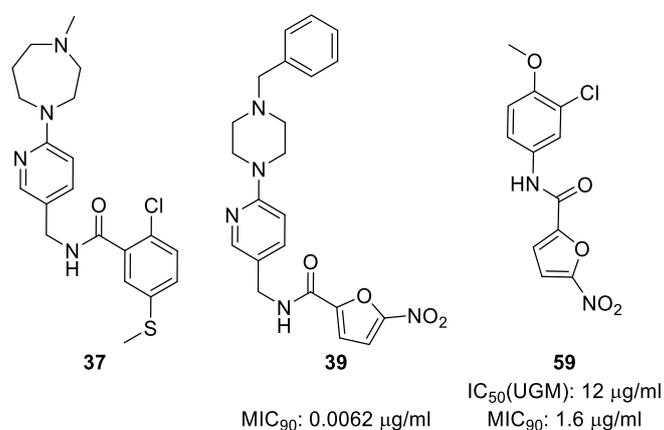

Figure 20: Structures of NMCCC hit (**37**), literature optimised candidate (**39**) and literature initial hit compound (**59**).[100,103]

Synthesis

Retrosynthesis of **37** back to readily available starting materials involved two steps (**Scheme 10**). Cleavage of the central amide reveals commercially available 2-chloro-5-(methylthio)benzoic acid and a methyl-homopiperazenyl-nicotinyl amine. A nucleophilic aromatic substitution (S$_N$Ar) cleavage then provides the methyl-piperazene and nicotinyl amine starting materials.

Scheme 10: Compound **37** retrosynthetic analysis.
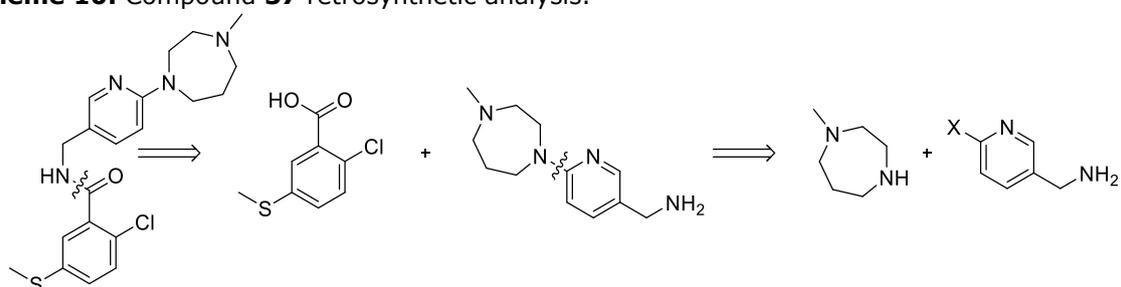
The synthesis of **37** was carried out according to **Scheme 11**. 1-Methylhomopiperazine was installed on 6-chloronicotinamide (**60**) using a S$_N$Ar to afford **61** in modest yield. Due to the volatility of 1-methylhomopiperazine, the reaction was not brought above 70 °C to ensure the amine remained in solution for the duration of the experiment. The reaction was carried out over four days and, if accessible, would likely benefit from microwave irradiation in any future studies. It is expected that use of 6-fluoronicotinamide would provide higher yields as fluoride is generally a better leaving group during S$_N$Ar reactions but the compound proved to be prohibitively expensive. Reduction of the terminal amide of **61** was first attempted with LiAlH$_4$, resulting in degradation of the starting material. BH$_3$.THF was used instead and provided **62** after stirring at 40 °C for five days. The final HCTU-mediated peptide coupling proceeded readily, providing 11.0 mg of **37** after HPLC purification.

Scheme 11: Synthesis of **37**.
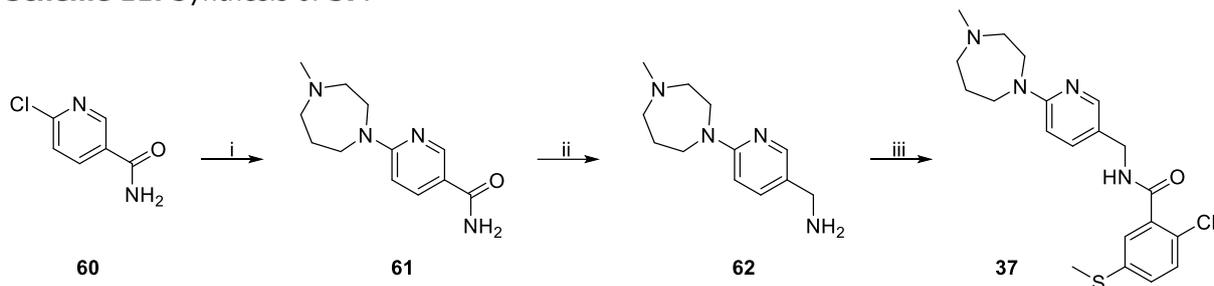
Reagents and conditions: i) 6-chloronicotinamide, DiPEA, 1-methylhomopiperazine, MeCN, 70 °C, 96 h, 33%. **ii)** BH$_3$.THF, THF, 0-40 °C, 120 h, 26%. **iii)** 2-Chloro-5-(methylthio)benzoic acid, HCTU, NEt$_3$, DMF, 16 h, 5%.

3.4 Compound 38

Compound **38** is predicted to have two weak interactions with the active site of InhA (**Figure 21**). The first is an edge-face π-stack between the terminal benzene and InhA Phe97 over 5.16 Å offset by 77.5°. The second is a cation-π interaction occurring over 6.47 Å between the central ring and catalytic triad member Lys165. Although these are both within the ranges recognised by Maestro (5.5 Å for edge-face stacking and 6.6 Å for the cation interaction) it should be noted that they are very close to the outer limits. The docking score for **38** is, however, still appropriately high at 96 and offers an interaction with Lys165 not commonly seen in the data set.
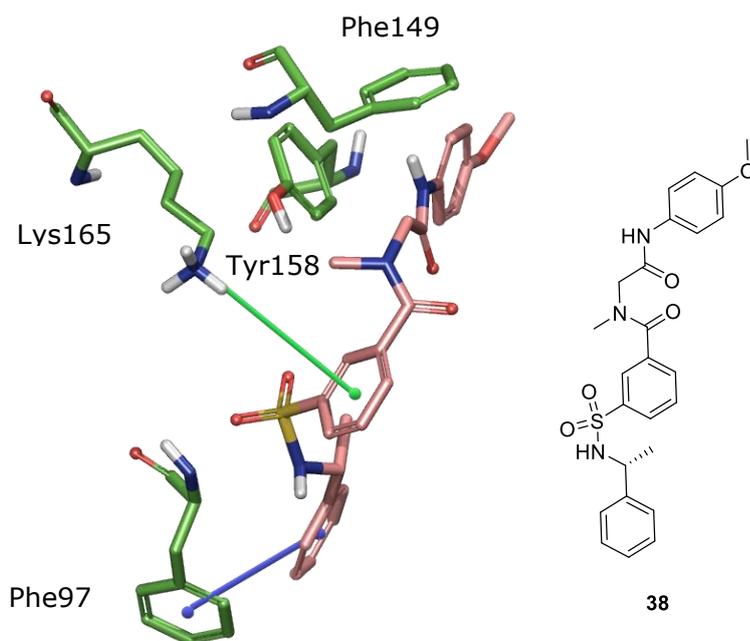
Figure 21: Compound **38** docking image and skeletal structure. **38** in peach, Tyr158, Phe149, Phe97 and Lys165 in green. Green line represents cation-π interaction and blue line represents edge-face stacking (PDB: 4U0J).

Synthesis

Compound **38** can be disconnected along its amides and sulfonamide (**Scheme 12**) with alternatives to each of the fragments released proving commercially available.

Scheme 12: Compound **38** retrosynthetic analysis.
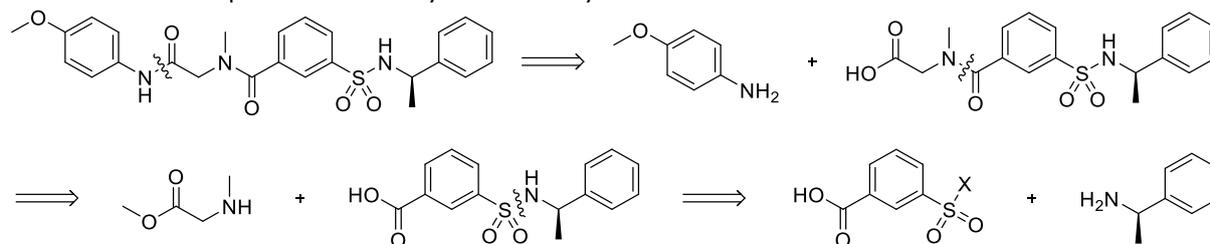
Synthesis began with esterification of 3-(chlorosulfonyl)benzoic acid (**63**) to form **64** according to literature precedent.[104] Following an initial sulfonamide coupling in moderate yield, the plan repeated an ester cleavage – HCTU coupling sequence until collection of **38**. Of note is the poor yield of **67** (28%), possibly due the sterically hindered nature of the starting materials. Compounds **67**, **68** and **38** formed rotamers observed during NMR experiments. The final structure of **38** was confirmed by repeating NMR measurements at 75 °C, at which point split peaks coalesced (Supporting information, Final compound data, **Figure 49**). Compound **38** was purified by HPLC to provide 24.1 mg material.

Scheme 13: Forward synthesis of **38**.
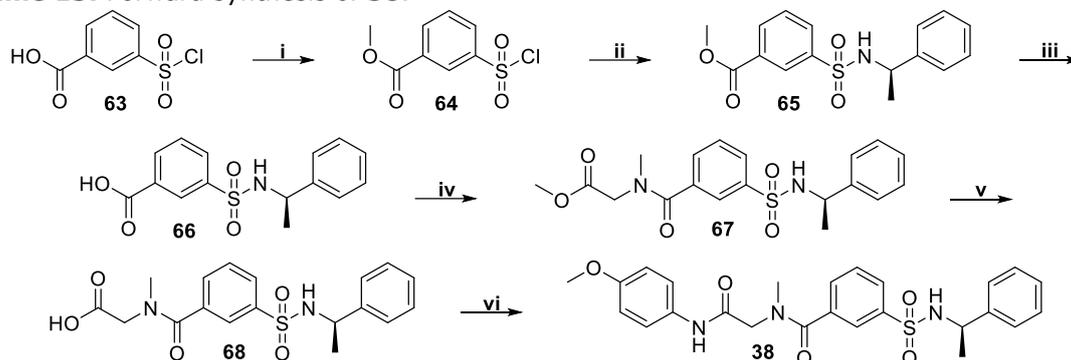
Reagents and conditions: i) 3-(chlorosulfonyl)benzoic acid, oxalyl chloride, MeOH, DMF, CH_2Cl_2, 1.5 h, 33%. **ii)** (R)-1-phenylethan-1-amine, K_2CO_3, THF:H_2O 1:1, 16 h, 70%. **iii)** LiOH, H_2O, 3 h, 77%. **iv)** sarcosine methyl ester, HCTU, K_2CO_3, DMF, 16 h, 28%. **v)** LiOH, H_2O, 2.5 h, 50%. **vi)** 4-methoxyaniline, HCTU, K_2CO_3, DMF, 16 h, 17%.

3.5 Single Point Inhibition Studies

Two positive controls were used during isolated enzyme assays: Triclosan and **NITD-564**.

Triclosan is readily available for purchase and **NITD-564** was synthesised according to literature precedent (**Scheme 14**).[105] The top 148 compounds identified during virtual screening were purchased from the NMCCC and underwent single point inhibition testing at 50 µM (Supporting information, Inhibition studies and enzyme kinetics, **Table 15**).

Scheme 14: Synthesis of hydroxypyridone reference (**NITD-564**).[105,106]
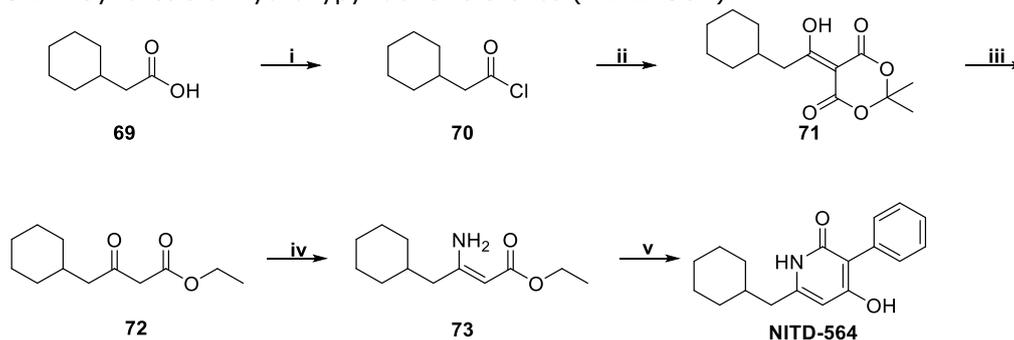
Reagents and conditions: i) oxalyl chloride, DMF, $CHCl_3$, CH_2Cl_2, 0-50 °C, 3 h. 99%. **ii)** DMAP, Meldrum's acid, CH_2Cl_2, 2.5 h. 97%. **iii)** EtOH, reflux 16 h. 98%. **iv)** 30% aq. NH_4OH, EtOH, 16 h. 30%. **v)** diethyl 2-phenylmalonate, 220 °C, 45 mins then NaOH, 140 °C, 150 W, 1 hr, 6%.

Unfortunately, all four of the compounds chosen for synthesis displayed low single point inhibition at 50 µM (**35**: 7%, **36**: 2%, **37**: 15%, **38:** 11%) and were discontinued. Of the remaining 144 compounds purchased from the NMCCC tested against InhA, 24 showed

inhibition of 30 – 50%, one at 51 – 60% and four at greater than 60% (**74**, **75**, **76** and
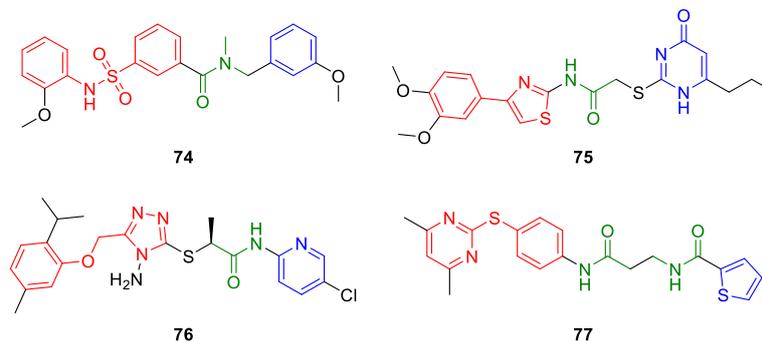
77, Figure 22).

Figure 22: Four compounds exhibiting greater than 60% inhibition of InhA at 50 µM. **74**: 63%, **75**: 63%, **76**: 70%, **77**: 70%. Compounds share the general characteristics of a terminal aryl ring (blue), an amide bridge (green) and closely linked second and third ring (red).

The compounds generally present with a lone, terminal aryl ring (blue) connected by an amide bridge (green) to a second and third aryl ring joined by a shorter linker (red). Although it is tempting to assume a similar binding orientation across all four compounds, Gold suite predicts a division between how **74**/**76** and **75**/**77** are predicted to dock (**Figure 23**). With lone aryl rings (blue) oriented towards the catalytic centre in the case of **74**/**76** and away from the catalytic centre in the case of **75**/**77**. Compounds **74** - **76** are all predicted to dock in the U-shaped geometry favoured by the substrate whereas **77** appears to generate a flatter pose away from the catalytic centre.
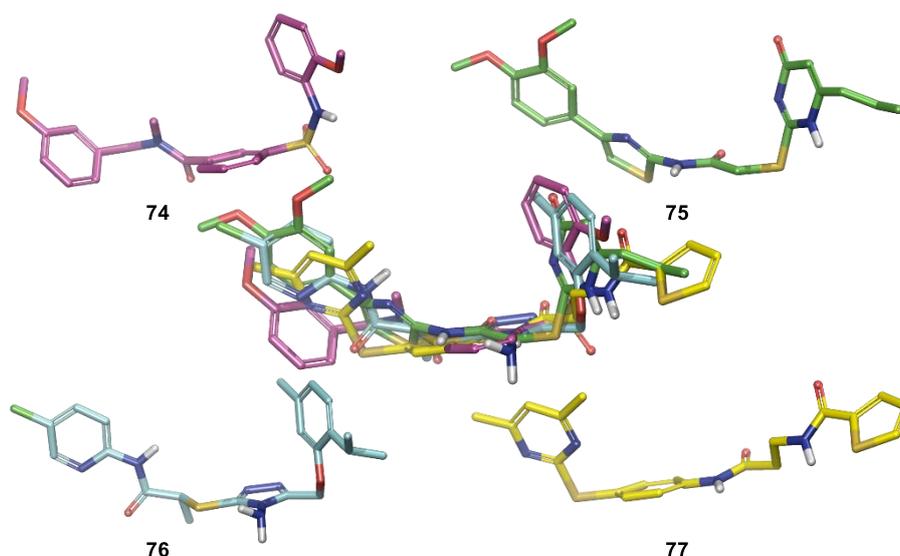
Figure 23: Predicted binding orientations of the top 4 NMCCC hits individually and superimposed in the center.

3.6 IC$_{50}$ Studies

The four compounds of interest had IC$_{50}$ results investigated along with **NITD-564** and triclosan (**Table 6**, **Figure 24**). **NITD-564** produced an IC$_{50}$ of 760 nM which was within three-fold of the published 560 nM.[79] Triclosan also behaved as expected with an experimental IC$_{50}$ of 9.2 µM compared to the previously reported 12.5 µM.[77] **74** and **77** produced results of 12.0 and 16.2 µM respectively. **75** and **76** did not display inhibition during IC$_{50}$ analysis and were discounted as false positives. Compounds were obtained directly from the NMCCC and any solution instability can be difficult to observe in 100 µL volumes.

Table 6: Inhibitory data for **74**, **77**, **NITD-564** and **triclosan** ± standard error (SE, the standard deviation of a sample mean from the population mean). Percentage inhibition testing was carried out in duplicate and IC$_{50}$ testing in triplicate.

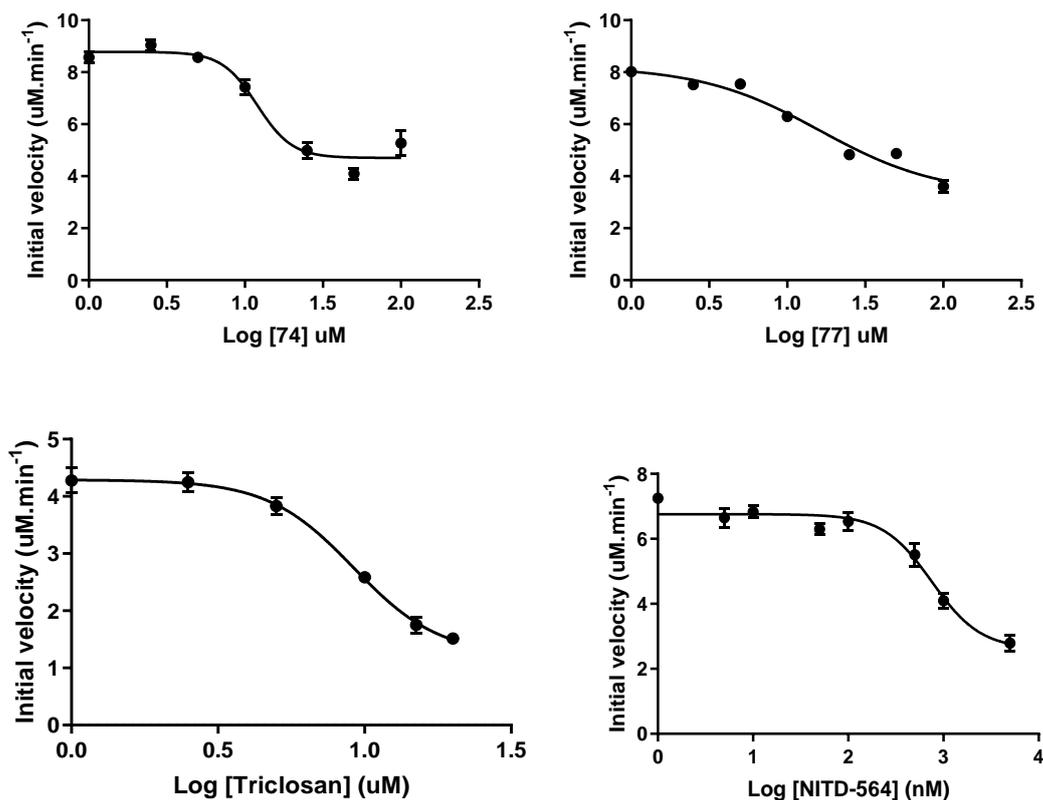| Compound	% inhibition at 50 µM	IC$_{50}$ (µM) ± SE
74	63	12.0 ± 1.1
77	70	16.2 ± 1.3
NITD-564[79]	N/A	0.8 ± 0.1
Triclosan	N/A	9.2 ± 1.1

Figure 24: IC_{50} data for **74**, **77**, triclosan and **NITD-564**. Experiments were carried out in triplicate with InhA (150 nM), NADH (100 µM), trans-*2*-octenoyl-CoA (400 µM) in PIPES buffer (30 mM, pH: 6.8) at 30 °C.

With hits **74** and **77** confirmed by IC_{50} analysis, they were taken forward for synthesis. At this point, university closure due to the coronavirus pandemic came into effect. Syntheses for both compounds along with the next highest scoring candidate from single point inhibition testing (**78**) were developed in preparation for regaining access to chemistry labs (**Figure 25**). As the NMCCC was put out of commission at this time, an IC_{50} for **78** could not be determined.
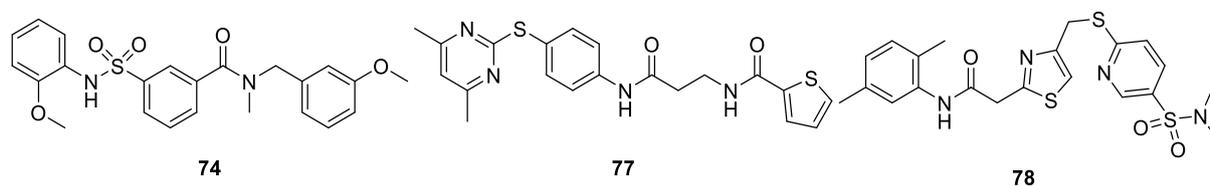
Figure 25: Target compounds showing activity *in vitro*, **74**, **77** and **78**.

Upon examination of the docking modes for each structure, explored further in the next chapter, only **78** was expected to exhibit a formal interaction with the active site of InhA: a π-stack with Phe149. Further data on the candidates can be seen in **Table 7**.

Table 7: Target compound physiochemical properties.

Compound	Molecular weight	cLogP*	Gold Score	Inhibition (%, 50 μM)	IC$_{50}$ (μM)
74	441	3.68	94	63	12.0 ± 1.1
77	413	3.32	91	70	16.2 ± 1.3
78	476	3.66	93	55	-
NITD-564[79]	283	4.47	-	-	0.76 ± 0.12
Triclosan	290	5.53	-	-	9.2 ± 1.1

*cLogP calculated through the JChem plugin for Office.[101]

3.7 Summary

In silico screening of the NMCCC allowed the selection of four compounds for synthesis (**35**, **36**, **37** and **38**). Each was successfully synthesised and purified by HPLC prior to point inhibition testing against InhA at 50 μM. As all candidates proved to exhibit inhibition below 30% at 50 μM, attention was instead moved to four compounds with inhibition of greater than 60%. These were finally narrowed down to two compounds which produced viable IC$_{50}$ curves (**74** and **77**). Due to closure of university facilities in the wake of the coronavirus pandemic, one further compound was planned for synthesis (**78**). The next chapter will focus on the synthesis and further analysis of these molecules.

4.0: Synthesis and evaluation of InhA inhibitor candidates identified during preliminary isolated enzyme assays.

4.1 Compound 74

Compound **74** displayed inhibition of 63% at 50 µM, an IC_{50} of 12.0 ± 1.1 µM and a GoldScore of 94. The distinct similarity between **74** and **38**, explored in the previous chapter, is also interesting. Both compounds are composed of a terminal methoxyaniline, central 1,3-sulfonylbenzamide core and a third, phenyl-based, substituent (**Figure 26**). Compound **74** is not expected to form any formal interactions with InhA but the larger **38** was anticipated to form two weak interactions through Lys165 and Phe97, a cation-π and π-π stack respectively. In the case of Lys165, the cationic side chain is only 5.6 Å from the closest phenyl carbon on **74** but approaches the edge of the ring, precluding a π-cation interaction. In the case of Phe97, the rings are poorly oriented towards each other, and no interaction is expected.
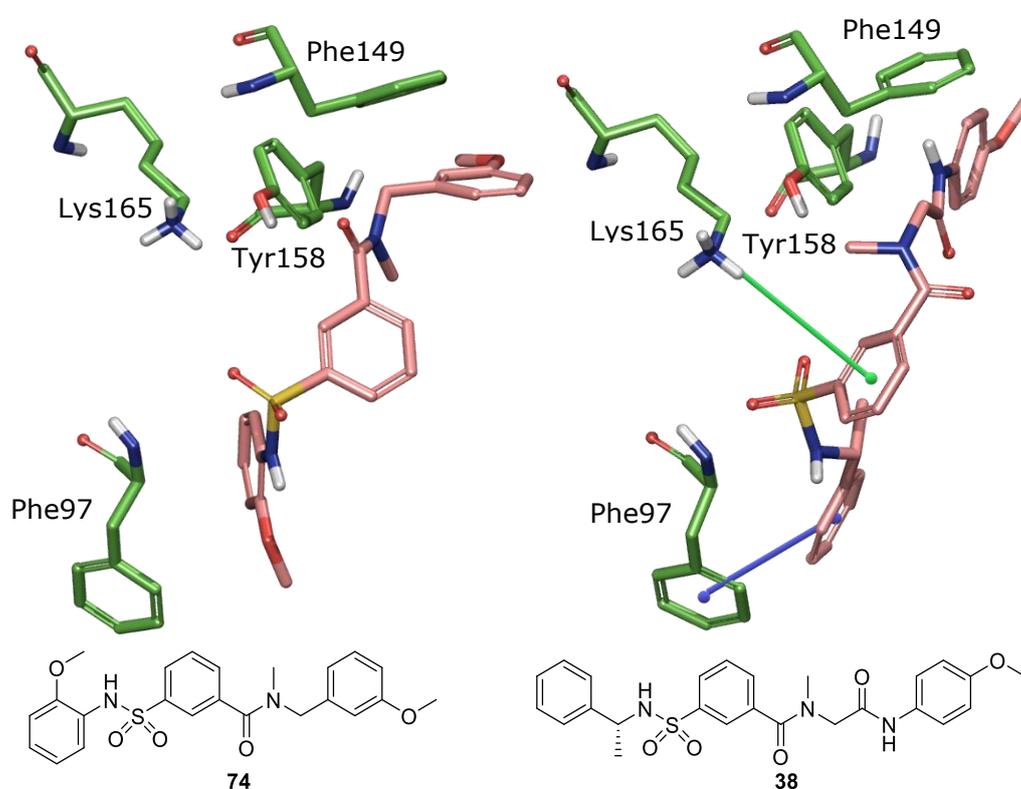
Figure 26: Docking model and structure of **74** (left, peach) and **38** (right, peach) in the presence of InhA residues: Tyr158, Phe149, Lys165 and Phe97 (green). Cation-π interaction is represented by a green line, π-stacking is represented by a blue line.

Synthesis

Retrosynthesis of **74** was possible through the cleavage of sulfonamide and amide bonds to furnish three building blocks (**Scheme 15**).

Scheme 15: Compound **74** retrosynthetic analysis.
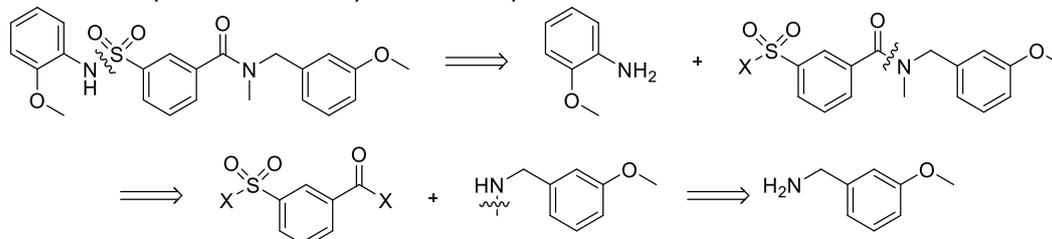
The forward synthesis proceeded according to **Scheme 16**. Secondary amine **80** was prepared from **79** by initial formation of a formamide and subsequent reduction to the amine with lithium aluminium hydride according to an adjusted literature procedure.[107] Coupling of **80** to 3-(chlorosulfonyl)benzoic acid with DCC proceeded with a moderate yield of 64%, despite the presence of a terminal sulfonyl chloride and furnished **81**. Correct attachment of **80** to the carboxylic acid of to 3-(chlorosulfonyl)benzoic acid was confirmed when NMCCC candidate **74** was produced by simple treatment of **81** with *o*-anisidine under basic conditions and purified *via* HPLC to provide 61.1 mg material. Compounds **81** and **74** were observed to form rotamers during NMR spectroscopy experiments. When spectroscopy was repeated at 75 °C, split peaks for **74** were observed to coalesce (Supporting information, Final compound data, **Figure 54**).

Scheme 16: Forward synthesis of **74**.
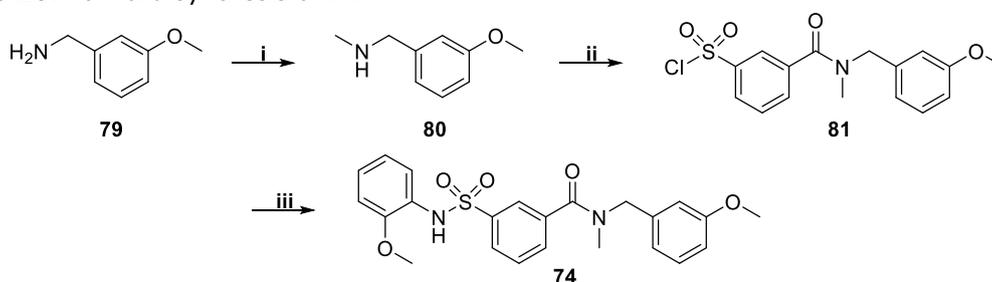
Reagents and conditions: i) HCO$_2$Et, THF, 16 h, 47 °C then LiAlH$_4$, THF, 4.5 h, 0-70 °C, 43%. **ii)** 3-(chlorosulfonyl)benzoic acid, DCC, DMAP, CH$_2$Cl$_2$, 16 h, 43%. **iii)** o-anisidine, NEt$_3$, CH$_2$Cl$_2$, 16 h, 17%.

4.2 Compound 77

Compound **77** was recorded with an inhibition of 70% at 50 µM, an IC$_{50}$ of 16.2 ± 1.3 µM and a GoldScore of 91. Like **74**, **77** is not expected to form any significant polar interactions with InhA. Some portions of **77** come into close contact with two InhA catalytic residues, however (**Figure 27**). The thioether bridge is within 4 Å of the Phe149 aromatic ring and the benzene ring of **77** is near Tyr158. When an *R*-primary amine is attached to carbon in place of the thioether (**77-A**), a cation-π interaction is predicted to form. By installing a hydroxyl group *ortho*- to the thioether (**77-B**), it is possible to predict hydrogen bonding to Tyr158. Both interactions could prove interesting places to start derivatization.
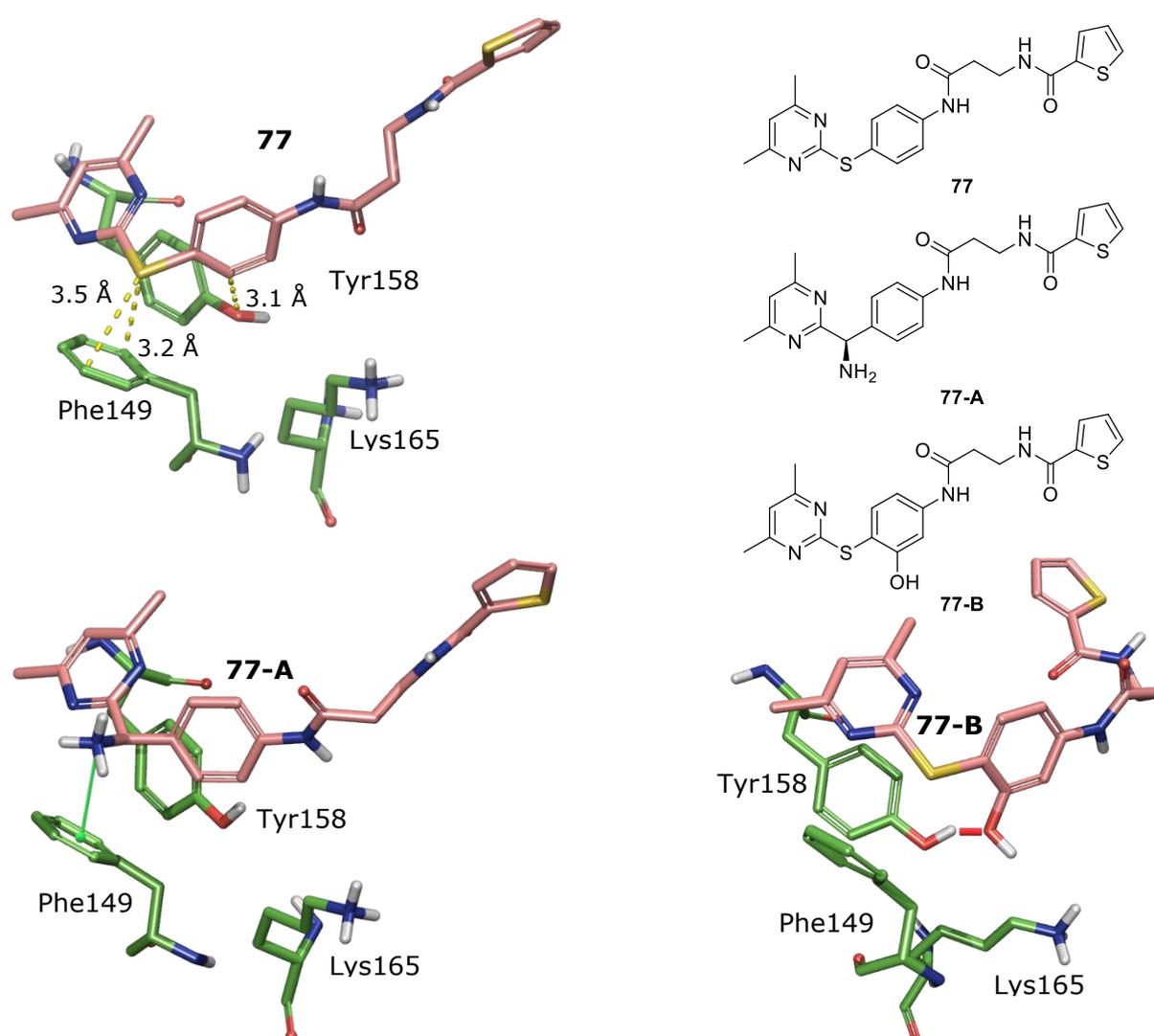
Figure 27: Top left: docking model of **77** with no formal interactions with InhA. Top right: skeletal structures of **77** and proposed derivatives **77-A** and **77-B**. Bottom Left: docking model of **77-A**. Bottom right: docking model of **77-B**. Ligands are colored in peach and InhA residues in green. Distances are represented by dashed lines, a cation-π interaction by a green line and hydrogen bonding by a red line.

Synthesis

Compound **77** was initially cleaved at the central amide to furnish a diarylsulfane and thiophene-amide (**Scheme 17**). Two of the four starting materials were obtained through S$_N$Ar cleavage of the diarylsulfane at the sulfur-pyrimidine bond. Finally, another amide cleavage affords the two other starting materials.

Scheme 17: Compound **77** retrosynthetic analysis
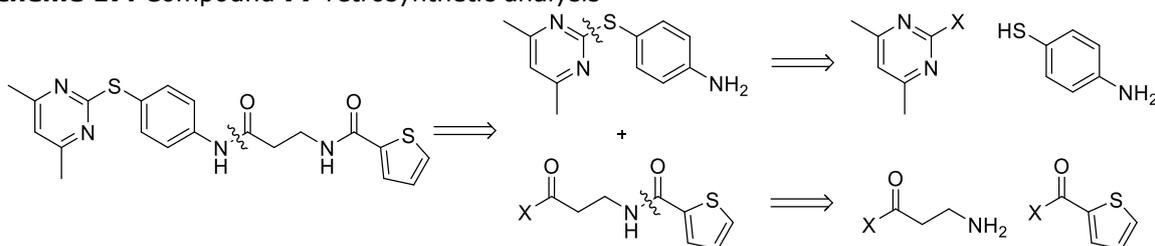
The forward synthesis began with a rapid S$_N$Ar reaction modified from the literature in which equimolar amounts of 4-aminobenzenethiol and pyrimidine **82** were mixed at 100 °C for two hours (**Scheme 18**).[108] The reaction proceeded and provided **83** in good yield. The aryl rings were confirmed to have coupled through the sulfur by proton NMR spectroscopy where two aniline peaks were observed at 5.45 ppm. Compound **85** was afforded though a DCC mediated peptide coupling of methyl 3-aminopropanoate and thiophene-2-carboxylic acid. Unmasking of the terminal carboxylic acid to obtain **86** using lithium hydroxide then facilitated a final peptide coupling with **83** and completed the synthesis of **77**, providing 130 mg material after HPLC purification.

Scheme 18: Forward synthesis of **77**.
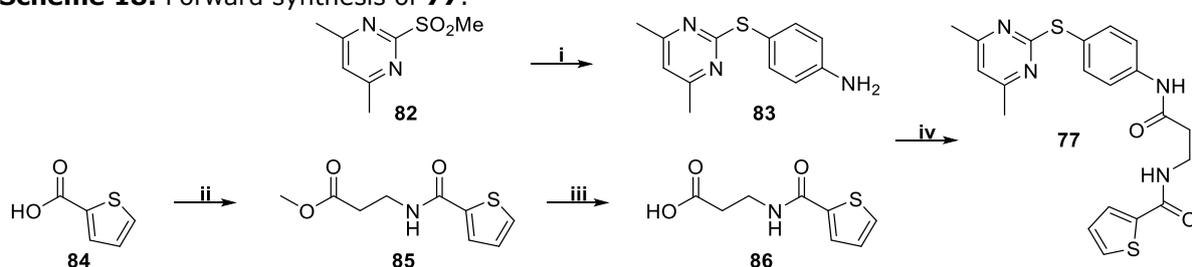
Reagents and conditions: i) 4-aminobenzenethiol, K$_2$CO$_3$, DMF, 2 h, 100 °C, 89%. **ii)** methyl 3-aminopropanoate, DCC, DMAP, CH$_2$Cl$_2$, 18 h, 71%. **iii)** LiOH, NEt$_3$, 1:1 DMF:H$_2$O, 3 h, 88%. **iv)** **83**, **86**, DCC, DMAP, CH$_2$Cl$_2$, 16 h, 8%.

4.3 Compound 78

Compound **78** produced an inhibition of 55% at 50 µM and a GoldScore of 93. A synthesis was planned and undertaken before obtaining IC_{50} data, during a period of closure for facilities required to carry out assays. Compound **78** was the fifth most potent candidate observed during single point *in vitro* assay of the top 148 NMCCC compounds produced from virtual screening. Unlike **74** and **77**, **78** is expected to produce an offset π-stack with InhA through Phe149 (**Figure 28**).
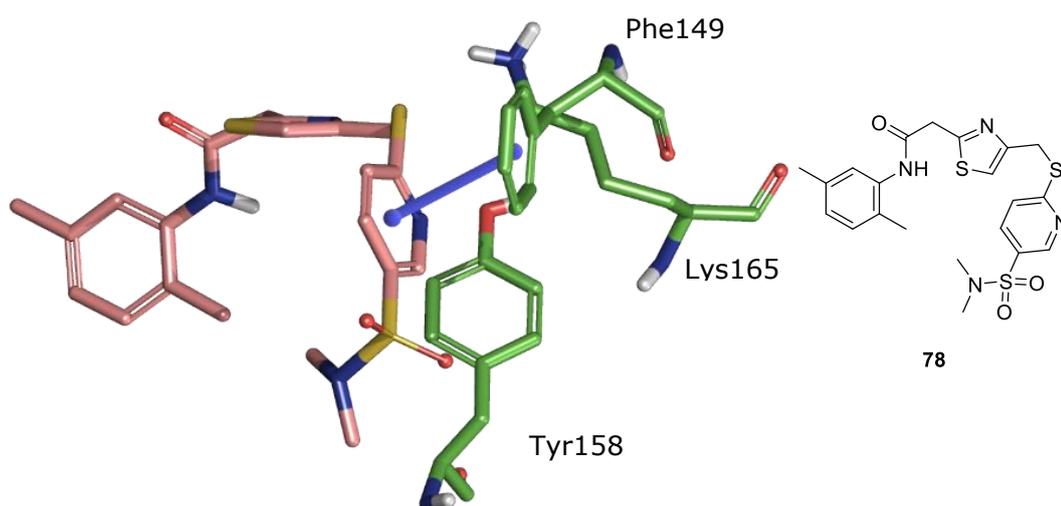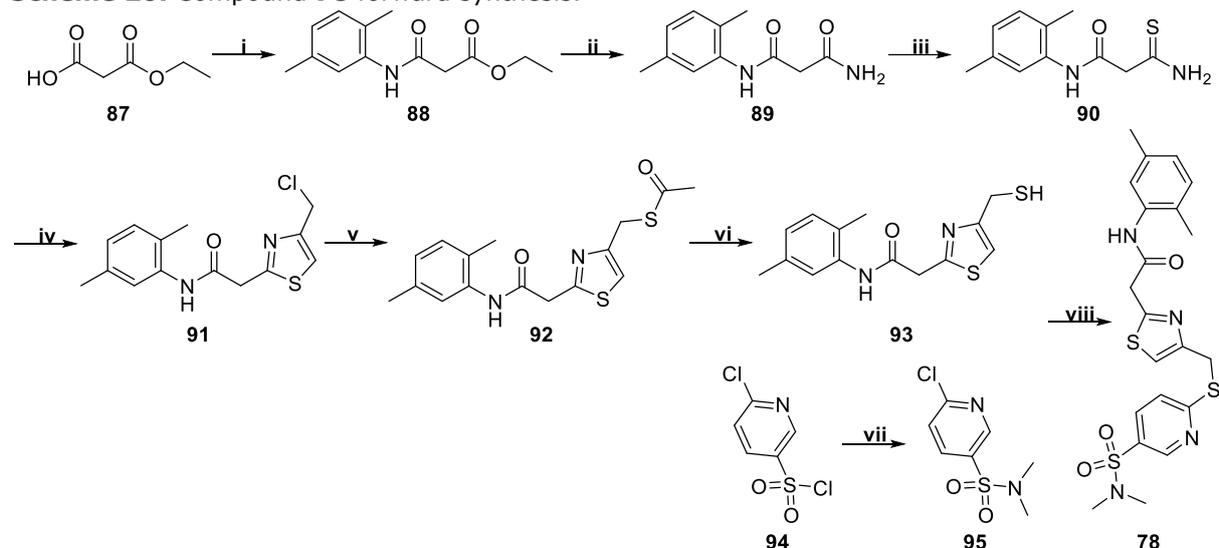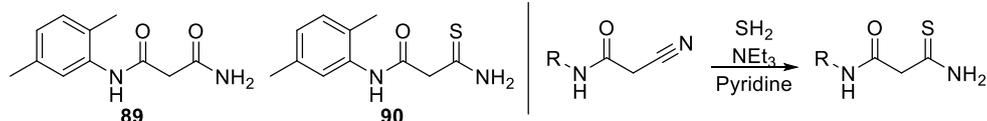

Figure 28: Docking model of **78** in the presence of InhA residues: Tyr158, Phe149 and Lys165 and the skeletal structure of **78**. **78** is coloured peach with the InhA residues in green and an offset π-stack with Phe149 represented by a blue line.

Synthesis

Retrosynthetic analysis of compound **78** began by cleaving the product for a S_NAr at the electron poor pyridine ring (**Scheme 19**). The released heterocycle was further simplified by cleavage of the sulfonamide. The remaining fragment was then opened at the thiazole, providing a terminal thioamide. The amide bond was then cleaved to provide a commercially available dimethyl aniline and 3-amino-3-thioxopropanoic acid derivative. This derivative would then be developed from a 1,3-dicarbonyl.

Scheme 19: Compound **78** retrosynthetic analysis.

The forward synthesis of **78** began with a peptide coupling between 2,5-dimethylaniline and 3-ethoxy-3-oxopropionic acid (**87**) in good yield (**Scheme 20**). Following this, a terminal amide was installed by treatment of **88** with methanolic ammonia solution and a distillation workup. In anticipation of installing the thiazole, the terminal amide of **89** was converted to thioamide **90** using Lawesson's reagent in the lowest yielding step of the scheme at 48%. Completion of the reaction was confirmed by proton NMR, where a high change in shift of the -NH$_2$ protons was observed following conversion. This behaviour is supported in the literature with similar products achieved through an orthogonal route (**Scheme 21**).[109] Formation of the thiazole was achieved through exposure of **90** to 1,3-dichloroacetone and the resultant chloride displaced in favour of thioacetone to provide **92**. Cleavage of the thioester was carried out using hydrazine monohydrate to provide the first portion of **78**, **93**. The final fragment, **95** was prepared from dimethyl amine and **94** which in turn participated in the final S$_N$Ar to provide 15.8 mg of **78** after HPLC purification.

Scheme 20: Compound **78** forward synthesis.

Reagents and conditions: i) 2,5-dimethylaniline, DCC, DMAP, NEt$_3$, CH$_2$Cl$_2$, 16 h, 90%. **ii)** NH$_3$, MeOH, 16 h, 96%. **iii)** Lawesson's reagent, 1,4-dioxane, 2 h, 60 °C, 48%. **iv)** 1,3-dichloroacetone, toluene, 2 h, 80 °C, 54 %. **v)** thioacetic acid, DMAP, NEt$_3$, CH$_2$Cl$_2$, 3 h, 84%. **vi)** N$_2$H$_2$.H$_2$O, CH$_2$Cl$_2$, 2.5 h, 78%. **vii)** NHMe$_2$, NEt$_3$, THF, 2 h, 87%. **viii)** **93**, **95**, K$_2$CO$_3$, DMF, 16 h, 8%.

Scheme 21: Comparative NMR chemical shifts for amide **89**, thioamide **90** (left) and a range of thioamides characterized by Dyachenko and Vovk in DMSO (*d*-6, right).[109]

Compound	δ (CXNH$_2$)
89	7.19, 7.59
90	9.41, 9.67

R	δ (CSNH$_2$)
PhCH$_2$	9.29, 9.55
Cyclo-C$_3$H$_5$	9.28, 9.58
Ph	9.33, 9.60
2-MeC$_6$H$_4$	9.45, 9.70
Furan-2-ylmethyl	9.31, 9.61

4.4 Inhibition studies

The three candidates were evaluated as inhibitors at 50 μM, where very low activities were observed (**74**: 13%, **77**: 8% and **78**: 0%). This was compared to initial testing carried out on ligand samples purchased from the NMCCC, where inhibition of InhA by each candidate was more than 60% at 50 μM. This change in behaviour can be rationalised with experiences gathered during sample preparation. All three compounds were successfully synthesised on larger scale than provided by the NMCCC. As such, preparation of 1 mL stock solutions made observation considerably easier than in the 20 μL ligand solutions made in 96 well plates during the initial 148 compound screen. Upon dilution in aqueous

buffer, all three synthesised compounds precipitated out to some degree. This was successfully reversed with heating and sonication over an hour. Although effective in this case, these conditions were not possible during the preliminary 148 compound screen due to the 96-well plate format. As a result, presumed instability of ligand solutions is expected to be responsible for the apparent inhibition obtained during preliminary screening. This is possibly due to a reduction in transmittance over time in the sample well as the ligand precipitated out, mimicking actual inhibition of the enzyme and associated reduction of NADH turnover. This is similar to the disparity observed when single point inhibition results were not replicated during IC_{50} testing for the two other compounds exhibiting inhibition over 60% at 50 µM, **75** and **76**. Unfortunately due to decommissioning of the NMCCC, revisiting the stock samples was not possible.

4.5 Summary

Two candidates (**74** and **77**) were selected from *in vitro* IC_{50} screening of the NMCCC. The synthesis of these along with a third compound (**78**) was planned and undertaken during restricted access to equipment necessary for IC_{50} testing. Solvation of these samples required extensive effort prior to carrying out isolated enzyme assays. These conditions were not available during previous testing and all three candidates showed low activity at 50 µM. At this stage, an orthogonal method of *in vitro* screening would be required to confirm further compound activities. These methods will be described in the next chapter.

5.0: Conclusions and Future Prospects

5.1 Conclusions

This project has followed the screening of the Nottingham Managed Chemical Compound Collection against the Mtb and Mb cell wall biosynthesis enzyme, InhA. An initial *in silico* screen facilitated the selection and synthesis of four compounds predicted to interact with InhA (**35**-**38**, **Figure 29**, **Table 8**). These high scoring compounds did not display activity against InhA in isolated enzyme assays. Two other compounds showed activity against InhA with IC$_{50}$ values in the low micromolar range (**74**, **77**). One other compound showed over 50% inhibition at 50 µM (**78**). These three compounds did not display inhibition of InhA following resynthesis and required extended effort to solvate prior to isolated enzyme assays. A further 24 compounds have shown greater than 30% InhA inhibition at 50 µM in initial isolated enzyme studies.
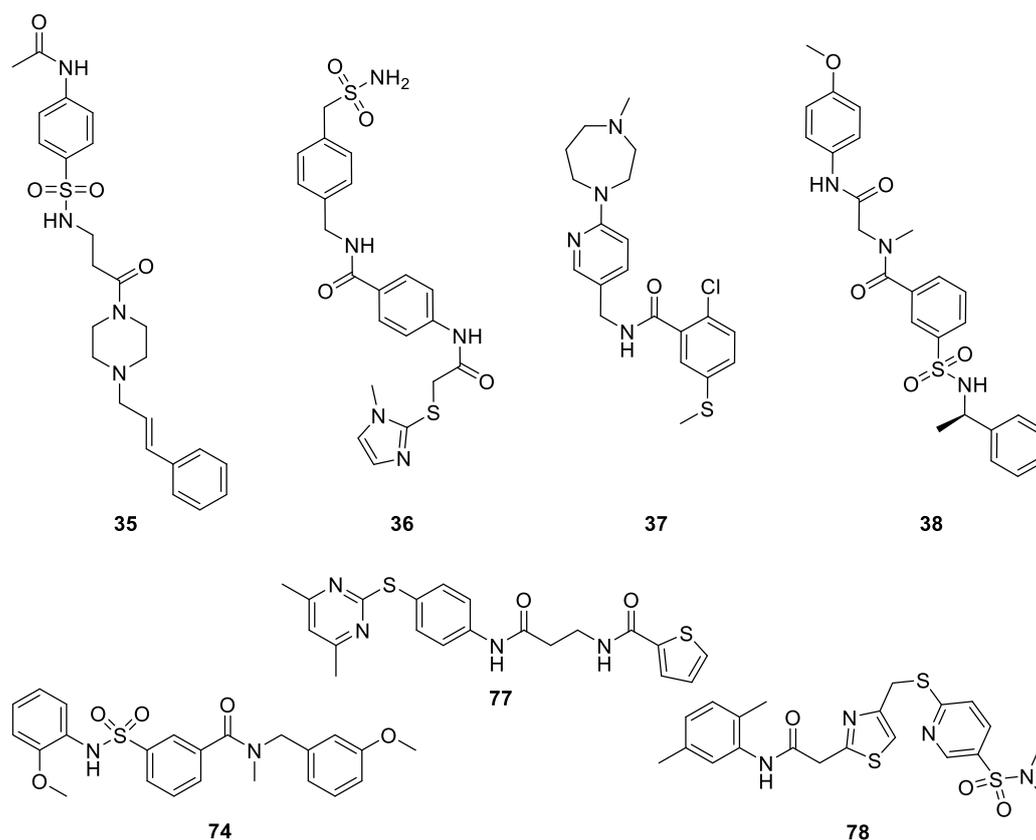
Figure 29: Hit compounds selected and synthesised from virtual screen (35-38) and hits selected from initial point inhibition testing (**74**, **77** and **78**).

Table 8: Project synthetic targets and their activity against InhA in isolated enzyme assays. Initial point inhibitions were taken from preliminary screening of ligands obtained from the NMCCC and final point inhibitions taken following resynthesis. Results ± SE.

Compound	GoldScore	Point Inhibition Initial (%, 50 μM)	IC$_{50}$ (μM)	Point Inhibition Final (%, 50 μM)
35	104	13	-	-
36	99	7	-	-
37	97	14	-	-
38	96	11	-	-
74	94	63	12.0 ± 1.1	13
77	91	70	16.2 ± 1.3	8
78	93	55	-	0

Due to the extended effort required to solvate compounds **74**, **77** and **78**, it is possible that initial inhibition results were false positives arising as a function of solution instability, a complicating factor in several InhA studies.[110–112] When observing cLogP data, the first set of compounds (**35**-**38**) generally have lower values than the second (**74**-**78**) and all barring **77** express more hydrophobic contact (S$_{vdw\_ext}$) with InhA relative to the four known inhibitors explored during validation of the structure (**Table 9**). The top performing candidates from single point inhibition studies have, therefore, been comprised of a subset of lipophilic compounds resistant to solvation and with little activity towards InhA.

Table 9: GOLDScore fitness functions for redocked compounds during validation and hit compounds synthesised during the study.

	f	=	S$_{hb\_ext}$	+	1.3750	x	S$_{vdw\_ext}$	+	S$_{vdw\_int}$	cLogP*
5G0T	93.36	=	4.46	+	1.3750	x	65.02	+	-0.50	3.07
4OHU	89.66	=	3.87	+	1.3750	x	63.23	+	-1.15	6.58
4TZK	80.68	=	3.30	+	1.3750	x	56.47	+	-0.26	2.87
4U0J	71.75	=	2.92	+	1.3750	x	50.17	+	-0.15	1.76
35	104.27	=	10.23	+	1.3750	x	72.79	+	-6.05	1.58
36	99.09	=	12.79	+	1.3750	x	66.70	+	-5.41	1.16
37	95.83	=	3.30	+	1.3750	x	67.50	+	-0.28	3.40
38	95.83	=	2.41	+	1.3750	x	73.89	+	-8.18	3.00
74	93.64	=	1.65	+	1.3750	x	67.99	+	-1.49	3.68
77	90.91	=	6.00	+	1.3750	x	63.80	+	-2.82	3.32
78	92.88	=	0.42	+	1.3750	x	71.74	+	-6.19	3.66

*cLogP calculated through the JChem plugin for Office.

This may be in part due to the nature of the enzyme itself and the highly hydrophobic, large, mycolic acid precursors it processes. In **Figure 30**, the active site surface has been coloured according to relative hydrophobicity of the associated peptide side chains.[113] The substrate binding pocket is almost entirely red, indicating a region of high lipophilicity. The

only major hydrophilic sites of interest are at the catalytic centre, containing Lys165 and Tyr158 (**A**), the lip of the binding pocket (Gln100, **B**) and two smaller regions depressed from the remainder of the pocket (Glu219, **C** and Gln214, **D**). The substrate pocket's extensive hydrophobicity may have the side effect of over-selection for hydrophobic and potentially insoluble or aggregative ligands during virtual screening.
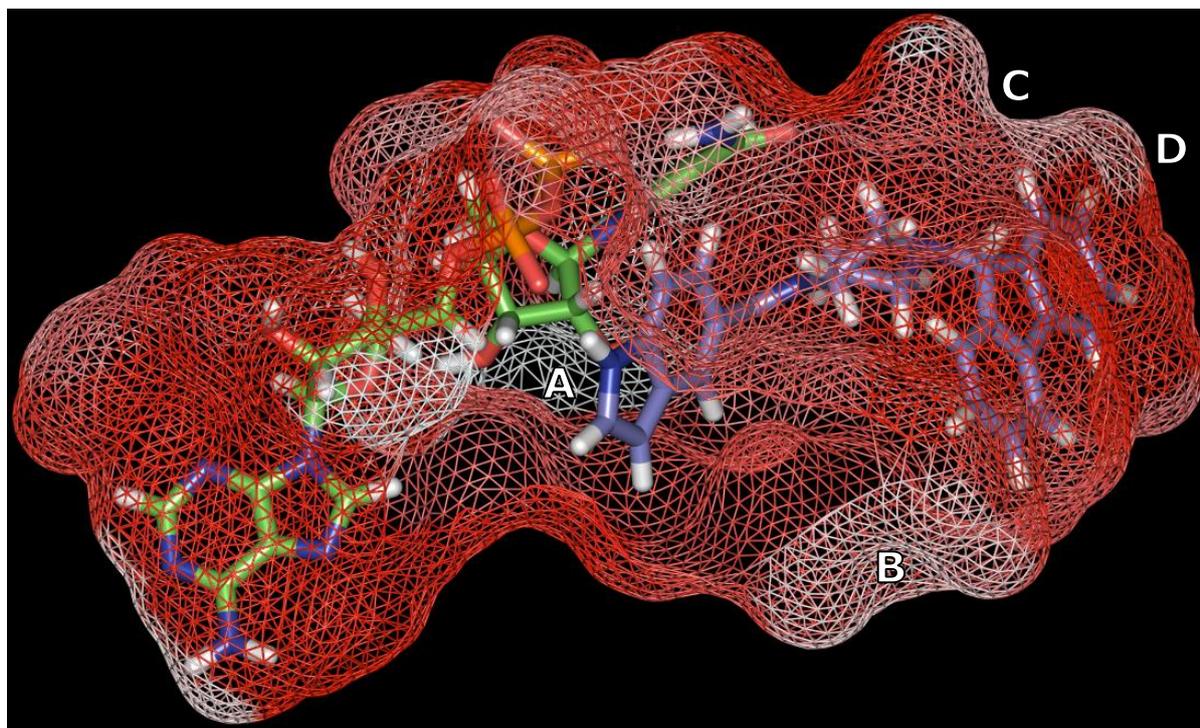
Figure 30: Binding cavity of InhA (PDB: 4U0J, 1.62 Å)[62] coloured according to side chain hydrophobicity as described by Eisenburg *et al.*[113] Hydrophobicity is on a scale where white describes hydrophilic residues and red describes hydrophobic residues. NADH cofactor is coloured in green and cognate ligand in blue. Lettered regions indicate hydrophilic sites of interest.

5.2 Future Prospects

When interrogating any remaining NMCCC compounds with apparent inhibition of InhA, it would be beneficial to pay close attention to available parameters relating to lipophilicity. As cLogP is only a measure of relative solubility between the organic and aqueous phase and does not account for a molecule's overall solvation limit, a solubility parameter for these compounds was also computed.

The Scrödinger suite includes a software package called QikProp which can calculate predicted physiochemical properties based on compound structure. As the NMCCC was

constructed to contain molecules expected to be soluble in common aqueous media, QikProp was not used for compound selection during the preliminary 148 compound screen. It is evident that at least some of the compounds of interest screened are not soluble during InhA experiments carried out, however, and QikProp could be a valuable tool in the future of the project. Specifically, Qikprop contains a prediction of the aqueous logarithm of solubility (Log S_{QP}) in moles *per* litre which has been shown to have reasonable correlation to experimentally derived parameters for a series of 429 compounds.[114] The study also included a statistical distribution of 476 orally bioavailable drugs where 80% fell between Log S_{QP} of -7 and 0, with a peak of 25% of the compounds in the range of -4 to -3. As a result, any candidates with Log S_{QP} greater than -4 would prove encouraging for future testing.

Remaining compounds of interest within the NMCCC

Compounds showing greater than 30% inhibition of InhA at 50 µM and a cLogP of less than 3 are listed in **Table 10** along with calculated Log S_{QP} values. All 11 showed a tight range of GOLDScores between 91 and 93 and point inhibition values from 31-39% at 50 µM. Two of the candidates, **99** and **106** expressed Log S_{QP} values greater than -4. Unfortunately, at the time of writing (February 2021), the NMCCC is out of commission and so a follow up study has not been possible.

Table 10: NMCCC hits with greater than 30% inhibition at 50 µM and cLogP less than 3. Log S calculated using Maestro QikProp.[92]

#	Structure	cLogP	GScore	Point Inhibition (50 µM)	LogS$_{QP}$	Active Site Contacts
96		1.14	93	33	-4.4	Phe149 π-stack
97		1.39	93	31	-5.4	Phe97 π-stack
98		1.05	93	33	-5.1	Hydrophobic only
99		1.73	92	36	-3.9	Tyr158/Phe149 double π-stack
100		1.34	92	37	-4.7	Tyr158 H-bond, Phe97 π-stack
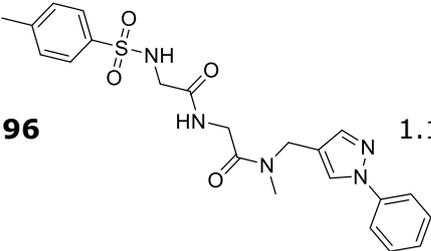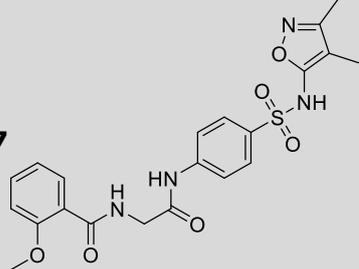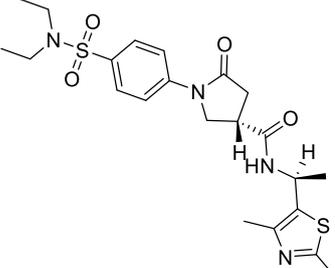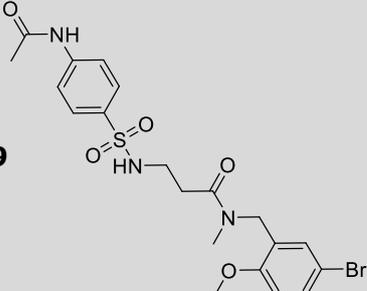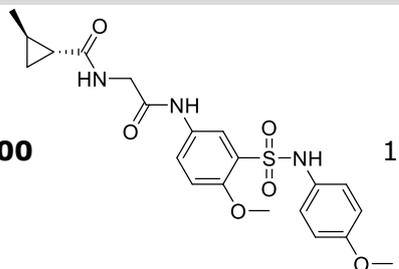
#	Structure	cLogP	GScore	Point Inhibition (50 μM)	LogS$_{QP}$	Active Site Contacts
101		2.25	92	36	-5.1	Hydrophobic only
102		1.38	92	37	-6.3	Tyr158 H-bond
103		2.38	91	34	-5.8	Tyr158/Phe149 double π-stack
104		0.45	91	31	-5.2	Tyr158 H-bond Tyr158/Phe149 double π-stack
105		2.47	91	39	-5.0	Tyr158 H-bond Tyr158/Phe149 double π-stack
106		2.57	91	32	-2.0	Met98 H-bond, Tyr158 π-stack
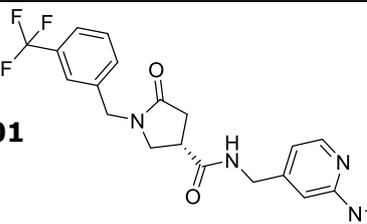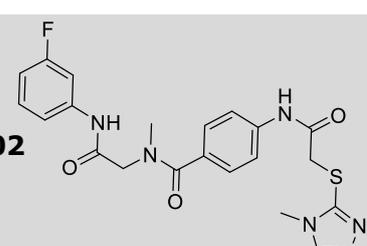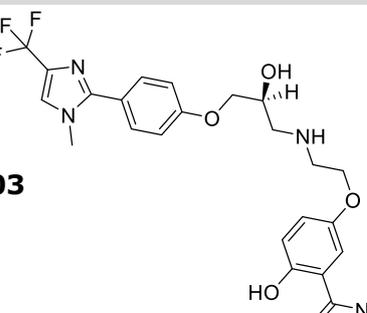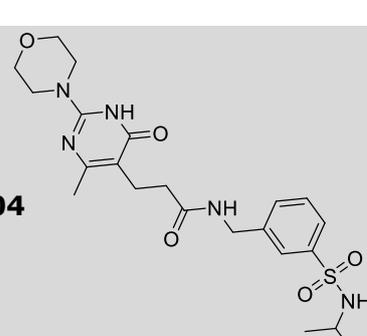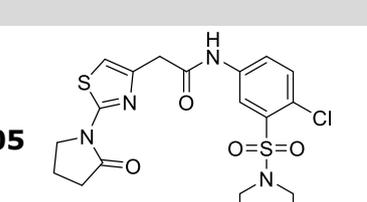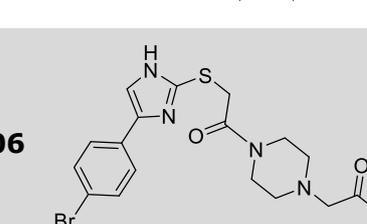
Four compounds are expected to form a double π-stack to InhA Tyr158 and Phe149 (**99**, **103**, **104** and **105**, **Figure 31**). Compounds **104** and **105** are also predicted to show hydrogen bonding to Tyr158. This repetition of binding behaviour, low cLogP and, barring **103**, predicted solubility better than **74** (-5.6), **77** (-7) and **78** (-6.8) makes them encouraging starting points. Although this binding mode did not seem to translate to potency with compounds **35** and **36**, its accompanying activity during *in vitro* studies deserves further research. The binding mode appears reasonable in **99**, **103** and **104** but in **105** a thiazole nitrogen presents its lone pair towards the edge of the Tyr158 ring. This interaction may be unfavourable due to interference with the Tyr158 ring π-electron cloud above the plane. This is different to orientations seen in **99**, **103** and **104**, where hydrogen is presented to the ring instead and a more likely prediction of the correct binding pose.
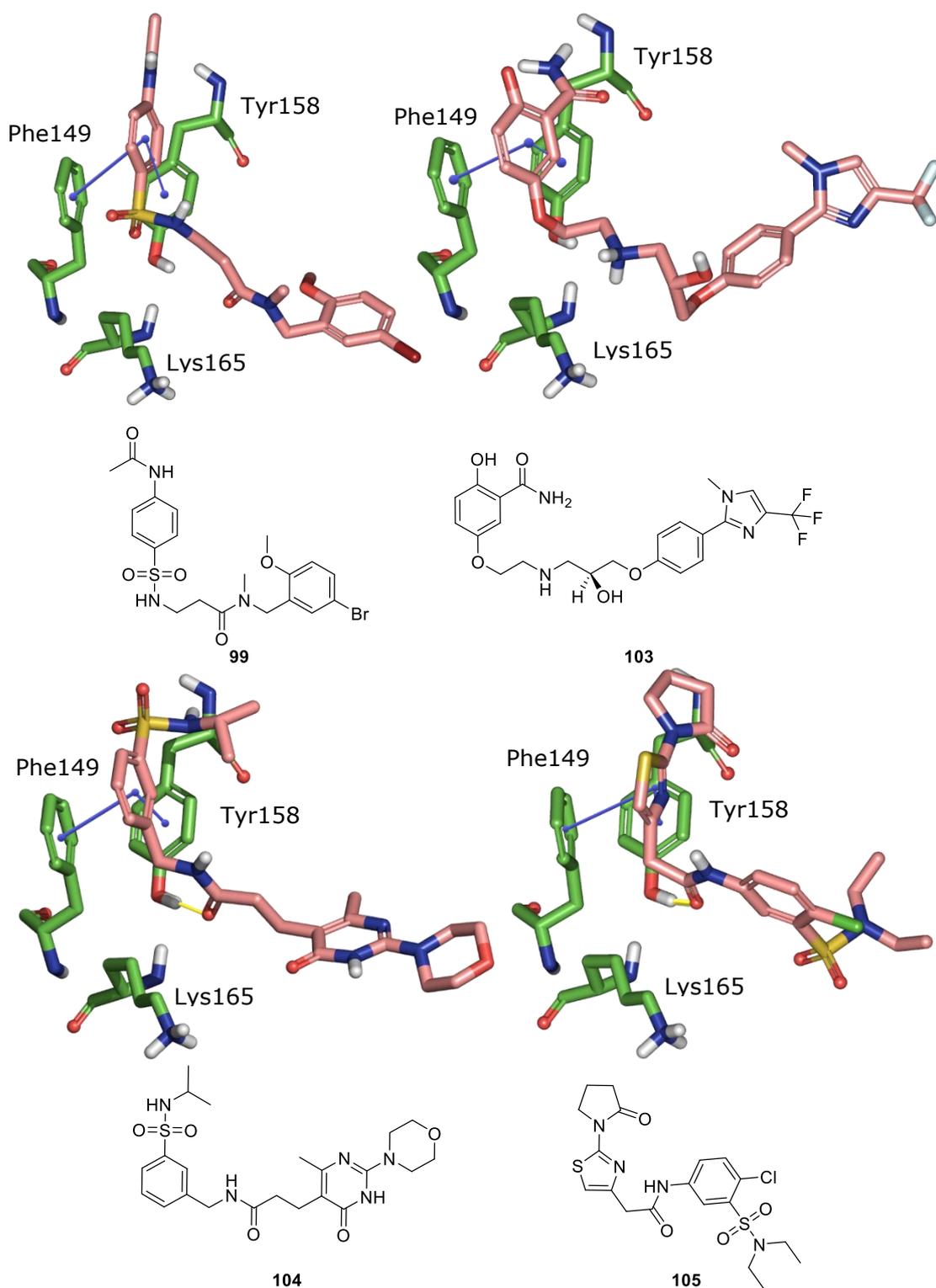
Figure 31: Predicted binding modes and skeletal structures of **99** (top left), **103** (top right), **104** (bottom left) and **105** (bottom right) to InhA (PDB: 4U0J). π-stacking is indicated with blue lines and hydrogen bonding with yellow lines.

The double π-stack has been predicted in the literature where a different series is expected

to interact with Tyr158 and Phe149 through the imidazole of **107** and the methyl benzene

of **108** (**Figure 32**).[115] The interaction was observed using docking program Glide from the Schrodinger suite, lending further weight to the interaction as it appears to be viable across compound series and docking programs. The compounds proceeded successfully to whole cell testing with MICs of 26.3 µM and 6.03 µM respectively.

107   108
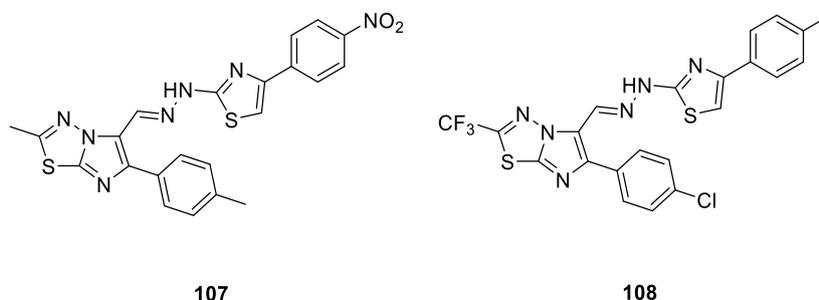
Figure 32: Compounds previously predicted to form double π-stack interactions with Tyr158 and Phe149.

As compound **99** shares a common predicted binding mode with three other candidates, of which it is the predicted most soluble, it is currently the highest priority target.

It is evident from previous screening difficulties that further analysis of any compounds of interest will be required and so an orthogonal screen with inbuilt quality control with regards to compound precipitation and aggregation would be valuable.

Orthogonal Screening

Saturation transfer difference (STD) NMR spectroscopyexperiments selectively saturate protein hydrogens outside the range of ligand hydrogens.[116] This saturation can transfer to any ligand binding to the protein, partially saturating the ligand bulk.[116] Binding ligands express attenuated peaks in the presence of protein and the difference *versus* the normal spectrum creates a positive peak in the STD result. This can be carried out in the presence of a cocktail of different ligands.[116] STD NMR is also able to identify false positives from aggregation by running an STD NMR experiment in the absence of protein and discounting any "active" compounds.[116] Limited determination of ligand binding sites can also be achieved through competition assays with known inhibitors.[117]

STD NMR screening has been used successfully to identify lead compounds active against InhA as well as cooperative binding with NADH.[117,118] Although this process is usually used to screen fragment libraries, it might also be used for interrogation of NMCCC leads. The chief concern associated with using NMR screening for larger compounds is poor identification of hits too potent to dissociate from the protein across NMR timescales. This occurs when inhibitors bind to a target with long residence times, resulting in poor exchange with the ligand bulk over the course of an experiment. Only a small proportion of free ligand is saturated by the protein target and provides either a very weak or complete lack of signal in the STD NMR spectrum. As mentioned in chapter 2, the NMCCC was designed as a small library of diverse chemistry to identify starting points for development. These compounds are not predicted to be initial nM inhibitors of InhA, supported by their modest point inhibition scores, and so generation of false negatives in this way are not expected. Any leads corroborated by orthogonal assays would be considered InhA inhibitors with a high degree of confidence and may proceed to whole-cell testing. Organisations such as the not-for-profit Tres Cantos Open Lab Foundation (TCOLF) provide a wide range of facilities including labs capable of carrying out whole cell testing against Mtb and Mb. This would be a valuable resource for testing any confirmed hits.

Compound development

Following identification of any hits active against InhA and whole cell Mtb and Mb, development would be required. As many of the structures identified as potential hits from screening of the NMCCC are well developed, they do not fit fully within the substrate pocket of the active site. For example, **99**, **103**, **104** and **105** all contain solvent exposed residues predicted to extend out of the pocket (**Figure 33**). These portions could either be pruned to increase compound efficiency or as modifiable handles for future physiochemical development. Although these compounds are predicted to be soluble, development within the extremely lipophilic substrate binding loop is likely to carry an increase in LogP. Solvent

exposed regions free of the loop may allow tolerated modification to increase hydrophilicity and ensure solubility throughout series development.
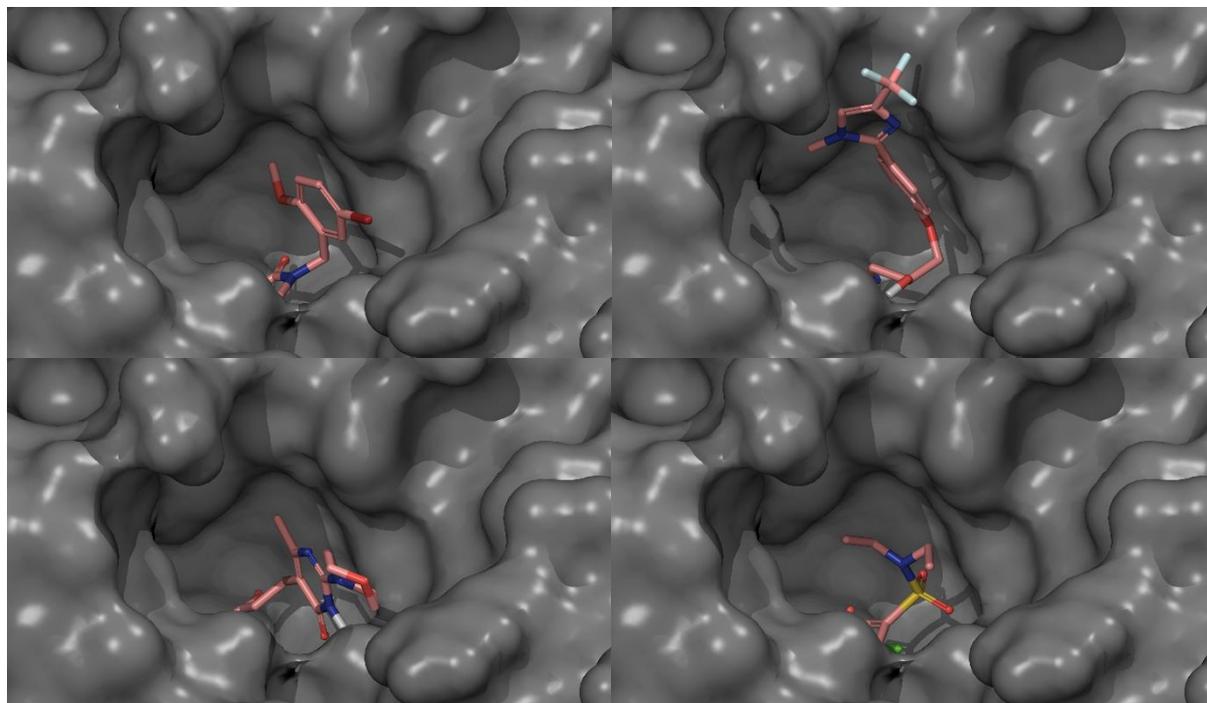
Figure 33: Compounds **99**, **103**, **104** and **105** occupying the active site of InhA (PDB:4U0J, 1.62 Å)[62] and displaying solvent exposed functionality. All resides and cofactor have been treated as the protein surface.

5.3 Summary

In conclusion, the Nottingham Managed Chemical Compound Collection has been screened computationally and a subset of 148 compounds screened in isolated enzyme assays for the presence of inhibitors of InhA. A set of seven NMCCC candidates have been synthesised and retested for activity against InhA with minor activity observed. A further 24 compounds have shown activity in isolated enzyme assays, with **99** displaying a cLogP below 3 and good predicted solubility. A similar binding mode has been observed in four compounds and future prospects laid out to further expand the project. Unfortunately, due to the emergence of Covid-19 and the closure of many university facilities, some intended experiments could not be completed. These include the development of an orthogonal compound screen and synthesis of further compounds of interest. This does mean, however, that interesting work remains to be carried out in the future.

6.0: Experimental Section

6.1 *In silico* screening

Virtual Screening was carried out using RCSB PDB entries 4U0J and 2AQ8.

Schrödinger 2018-1 and Maestro v11.5 were used for the following:

Protein superimposition/alignment. Structures were superimposed in place and aligned to calculate root mean squared deviation using the protein structure alignment tool.

Protein preparation. The Maestro protein preparation wizard was used to prepare a single monomer of InhA. Bond orders and hydrogen atoms were assigned automatically, and ionization states formed using Epik at pH: 7.0. When more than one ionization state was possible, the candidate with the lowest ionization penalty was chosen. Hydrogen bonding was assigned using PROPKA at pH 7. The OPLS3 forcefield was used to minimize heavy atom locations partially to a RMSD of below 0.3 Å and hydrogen atom locations fully. NADH was treated as part of the protein structure and the minimised protein was superimposed with the original entry to ensure that neither key residues nor the larger structure had changed significantly during minimisation. All water molecules were removed, and the protein exported to GOLD for docking studies as a Mol2 file.

Ligand preparation. The NMCCC SDF file was loaded into Maestro and a SMILES search carried out to ensure no structures submitted following the library's inception expressed Michael acceptors, esters or molecular weight exceeding 500 Da. The Epik and LigPrep modules were used to prepare ligand SDF files, using the OPLS3 forcefield for minimization and a pH of 7 ± 1 to obtain 3D structures of each candidate with any appropriate ionized or tautomeric states. Where chiral compounds were not explicitly stated by the incoming SDF file or were generated by protonation during ligand preparation, all possible variations were prepared and included within the screen.

Solvent Prediction. The QikProp module was used to predict solubility data at pH 7 ± 1.

GOLD v5.2.2 and Hermes 1.6.2 were used for the following:

Docking. The protein Mol2 file was loaded into GOLD and the active site identified by selecting all protein atoms within 10 Å of the 4U0J cognate ligand. The ligand was then

removed from the structure. Docking was carried out on a 4 – core processor running at 3.6 GHz, 16 Gigabytes RAM and a Quadro FX 5000 GPU. Candidates were docked with a maximum of 10 runs each and the top scoring pose for each retained. Individual runs took approximately 10 seconds, with each ligand taking up to 100 seconds if all 10 runs were carried out. The docking run in its entirety took six weeks. Poses were ranked according to the GoldScore Fitness function.

PyMOL v2.1 was used for the following:

Post docking analysis. Data was visualized to ensure no atypical ligand binding such as atomic overlap or occupying areas outside defined binding site prior to isolated enzyme studies.

Temperature factoring. Structure 4U0J was coloured by temperature factor to observe the relative uncertainty of each atom. Temperature factors below 40 Å^2 were viewed as confident positions, whereas > 60 Å^2 were viewed as disordered positions.

6.2 Enzyme Isolation

Plasmid. A pET 15b plasmid containing an 871-base sequence InhA insert between NcoI and BamHI was donated to the group by Prof. Peter Tonge (State University of New York, Stoney Brook) exhibiting a thrombin His tag cleavage site (Supporting information, **Figure 35**) and stored at -20 °C. Methods for carrying out transformation, protein expression and purification were developed by Aneesa Ahmed and utilised for this study.[119]

Transformation. Sample prepared from 1 µl plasmid and 20 µl calcium-chloride competent BL21-DE3 cells and incubated on ice for 30 mins. The transformation mixture was shocked at 42 °C for 45 secs and placed back on ice for 2 mins. Lysogeny broth (LB, 250 µl) was added and incubated at 37 °C for 45 mins before being plated on ampicillin treated agar and incubated overnight at 37 °C.

DNA sequencing. Sequence alignment was confirmed by dosing LB (20 mL) with 100 µg/mL ampicillin then treated with a single colony from a Petri dish and incubated for 12 hours at 37 °C. The cells were centrifuged at 30,000 g at 4 °C for 15 mins and treated to

extract samples for sequencing using a New England BioLabs Monarch Plasmid Miniprep Kit.

Protein expression. LB (20 mL) was inoculated with a single colony and grown at 37 °C with 180 rpm shaking overnight. This was then used to inoculate 1 L LB and left shaking at 37 °C, 180 rpm until an OD_{600} of 0.5 - 0.8 was reached. IPTG was added to an overall concentration of 1 mM and the culture left for 6 hours at 37 °C, 180 rpm. All work involving the use of bacteria up until this point was carried out using autoclaved bacteria and under sterile conditions. This involved cleaning of workstations with *iso*-propyl alcohol prior to the commencement of work and the use a Bunsen burner near any bacteria-containing open vessels. The culture was centrifuged at 30,000 *g*, 4 °C for 15 mins to obtain the cell pellet. The pellet was suspended in 20 mM Tris, 300 mM NaCl, pH 7.5, 20 mM Imidazole (binding buffer) and sonicated on ice using an MSE Soniprep 150 at an amplitude of 15 µm. The supernatant was filtered through a Minisart 0.45 µm unit.

Purification. The sample was purified on a Fast Protein Liquid Chromatography (FPLC) setup (AKTA prime). A nickel affinity column (HisTrap) was equilibrated with 10 column volumes of binding buffer at 1 mL/min and the sample loaded at 1 mL/min until protein flow through finished and a steady baseline reached. Protein was eluted on a gradient of 20 – 500 mM imidazole over 100 mL using elution buffer (200 mM Tris, 300 mM NaCl, pH 7.5, 500 mM imidazole, Supporting information, **Figure 36**). Absorbance was measured at 280 nm (ε: 37,300 M^{-1} cm^{-1})[120] and InhA protein containing fractions confirmed using SDS-PAGE (Supporting information, **Figure 37**) and Mass Spectrometry kindly carried out by Faadil Fawzy on a Thermo LTQ FT-Ultra and analysed using Water SYNAPT G2-Si HDMS software (Supporting information, **Figure 38**). SDS-PAGE solutions were used for gels composed of solutions laid out in **Table 11**. Resolving gel was added to the mould and cast over 30 minutes followed by the stacking gel, again left for 30 minutes. FPLC fractions were mixed with loading dye and heated to 90 °C for 5 minutes before being added to the gel along with PageRuler Plus prestained protein ladder. Gels were submerged in running buffer and run at 120 V until the dye front travelled the length of the gel. The Gel was

stained using Coomassie Blue over 30 minutes with shaking and de-stained in water overnight with shaking to allow visualisation under UV light.

Table 11: SDS gel components (12%).

Resolving Gel – 10 mL	3.4 mL water, 4 mL 30% bisacrylamide, 2.5 mL TRIS-HCl pH 8.8, 50 µL 20% SDS, 100 µL 10% APS, 10 µL TEMED
Stacking Gel – 4 mL	2.7 mL water, 0.8 mL 30% bisacrylamide, 0.5 mL TRIS-HCl pH 6.8, 20 µL 20% SDS, 40 µL 10% APS, 4 µL TEMED
Running Buffer	25 mM TRIS-HCl, 250 mM glycine, 0.1% SDS, pH 8.3
Loading dye	50 mM TRIS-HCl pH 6.8, 2% SDS, 10% Glycerol, 0.01% Bromophenol blue, 2 mM DTT

Purified InhA fractions were moved from elution buffer to storage buffer (30 mM PIPES, pH 6.8) with spin dialysis and glycerol added to a final concentration of 10%. The protein concentration was determined using a Nanodrop ND-1000 UV spectrophotometer at 280 nm (Σ: 37,300 $M^{-1}cm^{-1}$).[120] Protein stocks were divided into 0.5 mL aliquots, flash frozen in liquid N_2 and stored at -80 °C as single use samples for up to six months. A 1 L culture provided 9 - 13 mg purified protein.

6.3 UV/Vis screening

Materials. All materials were purchased from Sigma-Aldrich except the following compounds: InhA prepared as stated in the previous section. NADH disodium salt purchased from PanReac AppliChem and a 10 mM stock solution freshly prepared in 30 mM PIPES buffer at pH 6.8 (assay buffer) and stored on ice. Samples were diluted

accordingly and used within the same day. *trans*-2-Octenoyl CoA (OcCoA, **Figure 34**) was prepared by Tom Armstrong and provided for use in this study.[121]
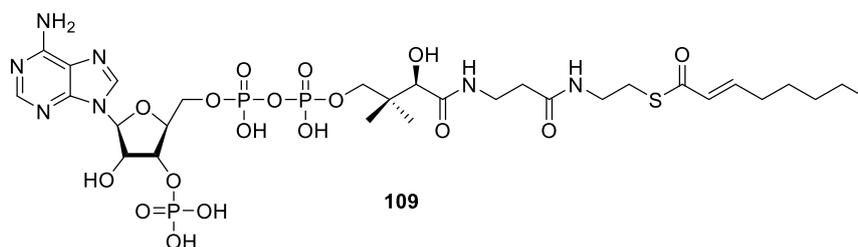
Figure 34: Assay substrate: *trans-2*-octenoyl-CoA (**109**).

Aliquots were thawed on ice on the day of testing and dissolved in 800 μL assay buffer to provide a 5 mM stock solution. Concentrations were confirmed prior to testing with a TECAN Spark for both NADH (340 nm) and OcCoA (254 nm) with extinction coefficients of $6220 \ M^{-1}cm^{-1}$ and $20,400 \ M^{-1}cm^{-1}$ respectively and diluted accordingly. Assays were carried out using Corning® 96-well clear flat bottom UV-transparent microplates with a path length of 0.29 cm.

All NMCCC ligands were purchased and provided by Lubna Hashmi as 10 mM stocks in DMSO in 96 well plates. Volumes of 2 uL and 10 uL per compound were purchased for percentage inhibition and IC_{50} testing respectively and used without further manipulation. Triclosan and **NITD-564** were taken as reference ligands during testing. Triclosan had to be prepared differently to those obtained from the NMCCC. As Triclosan precipitates upon DMSO-water dilution, it was first dissolved in 1 M NaOH to a concentration of 10 mM. NMCCC candidates and Triclosan were then diluted in assay buffer to a range of concentrations (0.1 μM – 100 μM) required for testing, ensuring a final percentage of 0.5% (v/v) DMSO or 1 M NaOH (final assay volume 100 μL). Negative control solutions containing only 0.5% (v/v) DMSO or 1 M NaOH were carried out alongside each row.

Enzyme Kinetics. Methods for carrying out enzyme kinetics and inhibition studies were developed by Aneesa Ahmed and utilised for this study.[119] Measurements were carried out using a TECAN Sunrise. To obtain K_m data for InhA, NADH absorbance was observed at

340 nm every 20 seconds for 15 cycles (5 min) at 30 °C (± 0.5 °C, Supporting information, **Figure 39**). The NADH concentration was varied (0 – 120 µM) whilst holding the OcCoA concentration at 400 µM then the OcCoA concentration was varied (0 – 415 µM) whilst holding the NADH concentration at 100 µM (Supporting information, **Table 12**, **Table 13**, **Table 14 Figure 40** and **Figure 41**). A row of 12 wells were charged with OcCoA and assay buffer with a 12-channel pipette and an initial reading taken. NADH was added to the wells and InhA added to the row directly underneath (final concentration of 150 nM). The top row was transferred to the bottom and mixed three times to initiate the 100 µL reaction. Initial absorbances were subtracted from each measurement cycle and the resultant values transformed into NADH concentration through use of the Beer – Lambert law (ε: 6220 $M^{-1}cm^{-1}$, l: 0.29 cm). The first two-minute linear slope was taken as the initial rate and GraphPad Prism 8 used to plot OcCoA and NADH concentrations against reaction rate to determine K_m constants using the non-linear regression tool and Michaelis - Menten enzyme kinetics analysis. Constants were within three-fold of three independent studies.

Enzyme Inhibition Studies. Enzyme kinetics methods were repeated for use in inhibition studies with the addition of ligand stock solutions to the wells prior to measurement of initial absorbance. Assay components were: InhA (5 µL, 150 nM), NADH (10 – 11 µL, 100 µM), OcCoA (9 – 11 µL, 400 µM), Ligand (10 µL, 0.1 – 100 µM) made up to 100 µL with pH: 6.8, 30 mM PIPES buffer ensuring a final concentration of 0.5% (v/v) DMSO or 1M NaOH in the reaction. Percentage inhibition studies were carried out at 50 µM in duplicate and results obtained by calculating loss of activity vs negative control (**Equation 3**). Negative controls were run once per row, positive controls once per plate and each compound assessed in duplicate.

$$\left(\frac{([V_0 - V_1])}{Vo} \right) * 100$$

Equation 3: Calculation of percentage inhibition where V_0: rate in absence of inhibitor and V_1: rate in presence of inhibitor (µM.min^{-1}).

IC$_{50}$ values were determined by plotting reaction rates against the log of concentrations from 0 – 50 µM in triplicate. GraphPad Prism four-parameter model: log (inhibitor) vs response – (variable slope) then provided the appropriate IC$_{50}$ (**Equation 4**).

$$Y = Bottom + \frac{(Top - Bottom)}{\left(1 + 10^{((\log IC_{50} - X) * HillSlope)}\right)}$$

Equation 4: IC$_{50}$ four parameter model. Top and Bottom: velocity plateaus, Hillslope: incline of the slope, IC$_{50}$: concentration at which the response is halfway between Top and Bottom.

Positive controls took the form of triclosan and hydroxypyridone **46** (NITD-564) with reported IC$_{50}$ values of 12.5 and 0.59 µM respectively, within threefold of the values obtained during this study (**Table 14**).[77,79] Error recorded as standard error (SE).

6.3 Chemical Synthesis

Reagents were commercial or HPLC grade obtained from Merck, Alfa Aesar, Fisher or Fluorochem and used as received unless indicated otherwise.

Solvents were obtained from Fisher Scientific for common use and used as received. Deuterated and anhydrous solvents were obtained from Merck and used as received.

Reactions were conducted under a nitrogen atmosphere unless otherwise stated and monitored by TLC using Merck Kieselgel 60 F254 plates with detection by UV absorption (254 nm) and staining with $KMnO_4$ (1.58 mg/mL) or Vanillin (60 mg/mL). Column chromatography was performed on silica gel from Merck or Fluorochem (0.040-0.063 nm). Microwave irradiation was carried out using a CEM Discover-S system.

HPLC purification was carried out using two systems. System 1: an Agilent 1200 series equipped with an Agilent 1260 Infinity Variable Wavelength Detector VL measuring at 215 nm. System 2: a Shimadzu Prominence UFLC system equipped with an SPD-20A Detector measuring at 220 nm. Purification was carried out on a semi-preparative column: XDB-C18 eclipse, 9.4 x 250mm, 5µm pore size, 2 mL/min. Runs were carried out using 0.1% v/v formic acid in water as solvent A and 0.1% v/v formic acid in MeCN as solvent B.

LCMS analysis was performed on a Shimadzu UFLCXR system coupled to an applied biosystems API2000 using a Phenomenex Gemini-NX 3mm-110A C18 column at a flow rate of 0.5 mL/min and UV detection at 220 nm. Samples were equilibrated at 5% solvent B in solvent A for one minute then method run: 5 to 98% solvent B over 2 mins, 98% solvent B for 2 mins, 98 to 5% solvent B over 0.5 min then 5% solvent B for one min. LCMS reference spectra for the purpose of displaying purity have been provided for all final compounds with retention times noted (Supporting information, Final compound data).

Optical rotation was measured using a Bellingham and Stanley ADP 220 Polarimeter.

HRMS was recorded on a BrukerTM micrOTOF, an orthogonal Time Of Flight (TOF) instrument with electrospray ionisation (ESI, positive and negative ion). Mass to charge ratio (m/z) given to four decimal places.

¹H- and ¹³C- NMR spectra were recorded (at 400 and 100 MHz respectively) on Bruker AV-400, AV(III) 400 or DPX 400 spectrometers running at ambient temperature unless otherwise stated. Chemical shifts are given in parts per million (ppm, δ) and *J* values in Hertz (Hz) relative to solvent residual peaks as an internal standard (chloroform-*d*: 7.26ppm and 77.1ppm, methanol-*d*$_4$: 3.31ppm and 49.00ppm, dimethylsulfoxide-*d*$_6$: 2.50ppm and 39.52ppm and deuterium oxide-*d*$_2$: 4.79ppm). Coupling constants are given in Hz. ¹³C- spectra were proton decoupled. Proton assignment was assisted through use of 2D COSY and HSQC spectra. Multiplets are designated by the following notations: apparent (app, indicating the formation of one of the following multiplets arising from interaction with more than the expected number of nuclei) singlet (s), broad signal (br) doublet (d), triplet (t), quartet (q), pentet (p), doublet doublet (dd), doublet triplet (dt) or multiplet (m). ¹H NMR reference spectra for the purpose of displaying purity have been provided for all final compounds (Supporting information, Final compound data).

All final compounds were synthesised and purified to at least 95% by LCMS.

boc-Piperazine (41)[122]
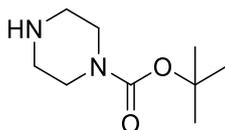
Piperazine (861 mg, 10.0 mmol) was dissolved in CH_2Cl_2 (15 mL) and stirred. *boc*-Anhydride (1.09 g, 5.00 mmol) in CH_2Cl_2 (15 mL) was then added dropwise and the reaction stirred for 16 hours. The reaction was filtered and washed with CH_2Cl_2 before concentration under reduced pressure. The crude product was re-dissolved in water and filtered before being saturated with K_2CO_3 and extracted three times with Et_2O. The combined organic layers were dried over $MgSO_4$, filtered, and concentrated under reduced pressure to yield the pure product as a clear oil (Yield: 754 mg, 4.05 mmol, 81%).

HRMS (ESI) required for $C_9H_{19}N_2O_2^+$: 187.1441, found 187.1444.

¹H NMR (400 MHz, chloroform-*d*) δ 3.44 – 3.31 (m, 4H, C**H₂**), 2.92 – 2.74 (m, 4H, C**H₂**), 2.29 (s, N**H**), 1.45 (s, 9**H**).

¹³C NMR (101 MHz, chloroform-*d*) δ 154.9(**C**O), 79.8 (**C**), 45.9 (**C**H₂), 28.6 (**C**H₃).

Note: ¹H NMR consistent with reported values

4-Cinnamylpiperazine-1-*boc* (42)
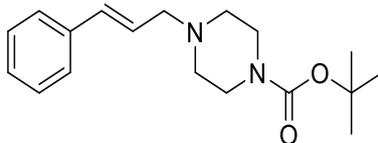
Compound **41** (1.00 g, 5.40 mmol) was dissolved in 1:3 THF:CH_2Cl_2 (40 mL) with cinnamaldehyde (800 mg, 5.90 mmol), sodium triacetoxyborohydride (1.50 g, 7.00 mmol) and two drops of AcOH and stirred for 20 hours. The mixture was partitioned between $CHCl_3$ and $NaHCO_3$ aq. sat. sol. and the organic layer dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The crude product was purified by column chromatography (50-75% EtOAc/hexane) to produce a white solid (Yield: 1.60 g, 5.15 mmol, 95%).

HRMS (ESI) required for $C_{18}H_{27}N_2O_2^+$: 303.2073, found 303.2088.

^{1}H NMR (400 MHz, chloroform-*d*) δ 7.41 – 7.34 (m, 2H, Ar**H**), 7.31 – 7.21 (m, 2H, Ar**H**), 7.26 – 7.20 (m, 1H, Ar**H**), 6.52 (d, *J* = 15.6, 1H, C**H**), 6.26 (dt, *J* = 15.7, 6.8 Hz, 1H, C**H**), 3.47 (m, 4H, C**H₂**), 3.18 (d, *J* = 6.8, 2H, C**H₂**), 2.46 (m, 4H, C**H₂**), 1.46 (s, 9H, C**H₃**).

^{13}C NMR (101 MHz, chloroform-*d*) δ 154.9 (**C**O), 136.9 (Ar**C**), 133.7 (**C**H), 128.7 (Ar**C**), 127.8 (**C**H), 126.5 (Ar**C**), 126.0 (Ar**C**), 79.8 (**C**), 61.2 (**C**H₂), 53.1 (**C**H₂), 43.2(**C**H₂), 28.6 (**C**H₃).

Note: ^{1}H NMR consistent with reported values.[123]

4-Cinnamylpiperazine (43)[123]
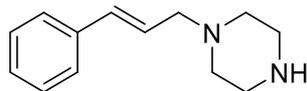
Compound **42** (1.60 g, 5.15 mmol) was dissolved in 1:5 TFA:CH_2Cl_2 (50 mL) and stirred for 5 hours producing a turquoise solution. The solution was concentrated to yield the crude product as a beige solid which was split between CH_2Cl_2 and 1 M NaOH and extracted with CH_2Cl_2 a further two times. The solution was dried over $MgSO_4$, filtered, and concentrated under reduced pressure to afford a pale-yellow oil (Yield: 1.03 g, 5.07 mmol, 99%) which was used without further purification.

HRMS (ESI) required for $C_{13}H_{19}N_2^+$: 203.1543, found 203.1546.

¹H NMR (400 MHz, methanol-d_4) δ 7.53 − 7.46 (m, 2H, Ar**H**), 7.40 − 7.28 (m, 3H, Ar**H**), 6.96 (d, J = 15.6 Hz, 1H, C**H**), 6.33 (dt, J = 15.5, 7.5 Hz, 1H, CHC**H**), 4.03 (d, J = 7.6, 2H, C**H₂**), 3.60 (s, 8H, C**H₂**).

¹³C NMR (101 MHz, methanol-d_4) δ 141.7 (**C**H), 137.2 (Ar**C**), 131.0 (Ar**C**), 129.2 (Ar**C**), 127.7 (Ar**C**), 116.7 (**C**H), 60.0 (**C**H₂), 49.1 (**C**H₂), 42.2 (**C**H₂).

Note: ¹H NMR consistent with reported values.

(9*H*-Fluoren-9-yl)methyl (3-(4-cinnamylpiperazin-1-yl)-3-oxopropyl)carbamate (44)
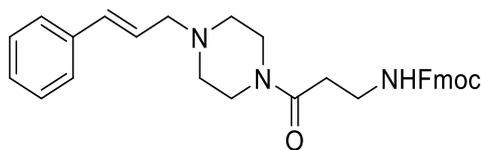
Compound **43** (2.00 g, 9.89 mmol) was dissolved in CH_2Cl_2 (100 mL) with HCTU (6.10 g, 14.8 mmol), Fmoc-β-alanine (3.39 g, 10.9 mmol) and K_2CO_3 (4.10 g, 29.7 mmol) and stirred for 24 hours. The suspension was diluted to half concentration in CH_2Cl_2 and treated successively with $NaHCO_3$ and brine. The crude was dried over $MgSO_4$, filtered, and concentrated under reduced pressure before purification by column chromatography to produce a clear, viscous oil (1-3% MeOH/CH_2Cl_2) (Yield: 4.75 g, 9.58 mmol, 96%).

HRMS (ESI) required for $C_{31}H_{34}N_3O_3^+$: 496.2595, found 496.2621.

¹H NMR (400 MHz, chloroform-*d*) δ 7.76 (d, *J* = 7.5 Hz, 2H, Ar**H**), 7.59 (d, *J* = 7.5 Hz, 2H, Ar**H**), 7.43 – 7.35 (m, 4H, Ar**H**), 7.35 – 7.27 (m, 4H, Ar**H**), 7.23 (m, 1H, Ar**H**), 6.52 (d, *J* = 15.9 Hz, 1H, C**H**), 6.24 (dt, *J* = 15.9, 6.8 Hz, 1H, C**H**), 5.61 (t, *J* = 6.4 Hz, 1H, N**H**), 4.35 (d, *J* = 7.2 Hz, 2H, C**H₂**), 4.20 (t, *J* = 7.2 Hz, 1H, C**H**), 3.75 – 3.62 (m, 2H, C**H₂**), 3.58 – 3.49 (m, 2H, C**H₂**), 3.49 – 3.40 (m, 2H, C**H₂**), 3.17 (d, *J* = 6.8 Hz, 2H, C**H₂**), 2.54 (t, *J* = 5.5 Hz, 2H, C**H₂**), 2.51 – 2.37 (m, 4H, C**H₂**).

¹³C NMR (101 MHz, chloroform-*d*) δ 170.0 (**C**O), 156.6 (**C**O), 144.1 (Ar**C**), 141.4 (Ar**C**), 136.8 (Ar**C**), 133.7 (**C**H), 128.8 (Ar**C**), 127.8 (Ar**C**), 127.8 (Ar**C**), 127.2 (Ar**C**), 126.5 (Ar**C**), 125.9 (**C**H), 125.3 (Ar**C**), 120.1 (Ar**C**), 66.9 (**C**H₂), 61.0 (**C**H₂), 53.2 (**C**H₂), 52.9 (**C**H₂), 47.4 (**C**H), 45.4 (**C**H₂), 41.7 (**C**H₂), 36.9 (**C**H₂), 33.3 (**C**H₂).

3-Amino-1-(4-cinnamylpiperazin-1-yl)propan-1-one (45)
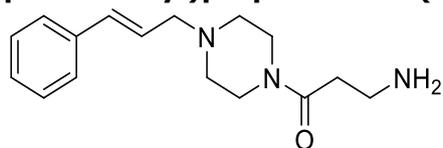
Compound **44** (500 mg, 1.01 mmol) was dissolved in THF (1.8 mL) with dimethylamine in EtOH (1.8 mL, 5.6 M, 10.00 mmol) and stirred for 1.5 hours. The crude product was concentrated in vacuo then co-evaporated with MeCN. The product was purified by column chromatography using a short silica plug (5-25% MeOH/CH_2Cl_2) to produce a clear, viscous oil (Yield: 232 mg, 850 µmol, 85%).

HRMS (ESI) required for $C_{16}H_{24}N_3O^+$: 274.1914, found 274.1925.

¹H NMR (400 MHz, dimethylsulfoxide-d_6) δ 7.51 – 7.38 (m, 2H, Ar**H**), 7.38 – 7.27 (m, 2H, Ar**H**), 7.27 – 7.14 (m, 1H, Ar**H**), 6.55 (d, J = 16.0 Hz, 1H, C**H**), 6.30 (dt, J = 16.0, 6.6 Hz, 1H, C**H**), 3.57 – 3.36 (m, 4H, C**H₂**), 3.11 (d, J = 6.6 Hz, 2H, C**H₂**), 2.72 (t, 2H, J = 6.5 Hz, C**H₂**), 2.45 – 2.17 (m, 6H, C**H₂**), 1.42 (br, 2H, N**H₂**).

¹³C NMR (101 MHz, dimethylsulfoxide-d_6) δ 170.4 (**C**O), 137.1 (Ar**C**), 132.7 (**C**H), 129.1 (Ar**C**), 127.9 (Ar**C**), 127.2 (**C**H), 126.7 (Ar**C**), 60.5 (**C**H₂), 53.4 (**C**H₂), 52.9 (**C**H₂), 45.3 (**C**H₂), 41.3 (**C**H₂), 38.6 (**C**H₂), 36.9 (**C**H₂).

N-(4-(N-(3-(4-Cinnamylpiperazin-1-yl)-3-oxopropyl)sulfamoyl)phenyl) acetamide (35)
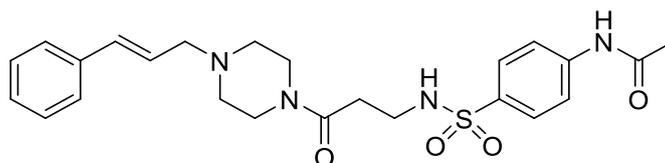
Compound **45** (115 mg, 370 μmol) was dissolved in DMF (2 mL) with NEt$_3$ (94.0 μl, 730 μmol) at 0 °C using an ice water bath. 4-Acetamidobenzenesulfonyl chloride (94.0 mg, 400 μmol) was dissolved in DMF (2 mL) and added dropwise and stirred for ten minutes before allowing the reaction to warm to room temperature and stirring for 16 hours. The reaction was concentrated under reduced pressure, providing 240 mg crude material. Reverse-phase HPLC purification (10% B over 8 mins, 10-25% B over 1 min, 25% B over 10 mins, 25-100% B over 1 min, 100% B over 5 mins) of 60.0 mg crude material provided a white powder (Yield: 24.1 mg, 50.0 μmol, 14%).

HRMS (ESI) required for C$_{24}$H$_{31}$N$_4$O$_4$S$^+$: 471.2066, found 471.2057.

^{1}H NMR (400 MHz, dimethylsulfoxide-d_6) δ 10.32 (s, 1H, CON**H**), 8.15 (s, 1H, N**H**), 7.79 – 7.69 (m, 4H, Ar**H**), 7.46 – 7.39 (m, 2H, Ar**H**), 7.35 – 7.29 (m, 2H, Ar**H**), 7.26 – 7.20 (m, 1H, Ar**H**), 6.53 (d, J = 16.0 Hz, 1H, C**H**), 6.29 (dt, J = 16.0, 6.6 Hz, 1H, C**H**), 3.45 – 3.31 (m, 4H, C**H₂**), 3.10 (d, J = 6.6 Hz, 2H, C**H₂**), 2.93 (app q, J = 6.8 Hz, 2H, C**H₂**), 2.43 (t, J = 6.8 Hz, 2H, C**H₂**), 2.39 – 2.20 (m, 4H, C**H₂**), 2.08 (s, 3H, C**H₃**).

^{13}C NMR (101 MHz, dimethylsulfoxide-d_6) δ 169.0 (**C**O), 168.4 (**C**O), 142.8 (Ar**C**), 136.6 (Ar**C**), 134.0 (Ar**C**), 132.3 (**C**H), 128.6 (Ar**C**), 127.7 (Ar**C**), 127.5 (Ar**C**), 126.6 (**C**H), 126.2 (Ar**C**), 118.6 (Ar**C**), 60.0 (**C**H₂), 52.6 (**C**H₂), 52.2 (**C**H₂), 44.7 (**C**H₂), 41.0 (**C**H₂), 38.9 (**C**H₂), 32.5 (**C**H₂), 24.1 (**C**H₃).

Methyl 2-((1-methyl-1*H*-imidazol-2-yl)thio)acetate (47)
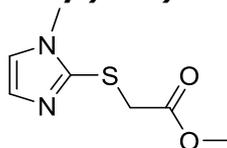
1-Methyl-1H-imidazole-2-thiol (1.00 g, 8.76 mmol) was dissolved in THF (18 mL) at 0 °C using an ice water bath and NaH (60%, 526 mg, 13.1 mmol) added portion-wise to form a slurry which was left to stir for 30 mins at 0 °C. Methyl 2-bromoacetate (0.8 mL, 8.76 mmol) was added dropwise at 0 °C, allowed to warm to room temperature and stirred for 16 hours. The reaction was acidified to pH 3 with 1M HCl, extracted with CH_2Cl_2 and washed successively with $NaHCO_3$ aq. sat. sol. and brine. The organic layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure to produce a yellow oil (Yield: 768 mg, 5.64 mmol, 64%).

HRMS (ESI) required for $C_7H_{11}N_2O_2S^+$: 187.0536, found 187.0543.

¹H NMR (400 MHz, chloroform-*d*) δ 7.07 (d, *J* = 1.3 Hz, 1H, Ar**H**), 6.93 (d, *J* = 1.3 Hz, 1H, Ar**H**), 3.86 (s, 2H, C**H₂**), 3.70 (s, 3H, C**H₃**), 3.67 (s, 3H, C**H₃**).

¹³C NMR (101 MHz, chloroform-*d*) δ 169.8 (**C**O), 140.1 (Ar**C**), 129.4 (Ar**C**), 122.8 (Ar**C**), 52.9 (**C**H₃), 36.4 (**C**H₂), 33.6 (**C**H₃).

2-((1-Methyl-1*H*-imidazol-2-yl)thio)acetic acid (48)
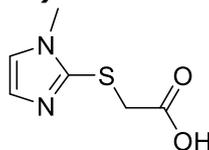
Compound **47** (1.00 g, 5.37 mmol) was dissolved in 4:1 THF:2M aqueous NaOH (54 mL) and left to stir for 2 hours. THF was removed by rotary evaporation and the reaction neutralised with 1 M HCl. The suspension was extracted three times with EtOAc and the combined organic layers dried over MgSO$_4$, filtered, and concentrated under reduced pressure to produce a yellow solid with no further purification (Yield: 930 mg, 5.39 mmol, 100%).

HRMS (ESI) required for C$_6$H$_9$N$_2$O$_2$S$^+$: 173.0379, found 173.0387.

^{1}H NMR (400 MHz, deuterium oxide-d_2) δ 7.59 (d, J = 2.0 Hz, 1H, Ar**H**), 7.54 (d, J = 2.0 Hz, 1H, Ar**H**), 3.95 (s, 3H, C**H$_3$**), 3.91 (s, 2H, C**H$_2$**).

^{13}C NMR (101 MHz, deuterium oxide-d_2) δ 172.2 (**C**O), 133.4 (Ar**C**), 125.5 (Ar**C**), 120.9 (Ar**C**), 36.5 (**C**H$_2$), 35.2 (**C**H$_3$).

Note: ^{1}H NMR consistent with reported values.[124]

Methyl 4-(2-((1-methyl-1*H*-imidazol-2-yl)thio)acetamido)benzoate (49)
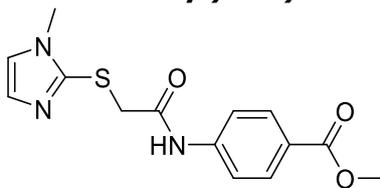
Compound **48** (925 mg, 5.37 mmol) was dissolved in CH_2Cl_2 (25 mL) with methyl 4-aminobenzoate (812 mg, 5.37 mmol), EDC.HCl (1.13 g, 5.91 mmol) and DMAP (131 mg, 1.07 mmol) and left to stir for 16 hours. NH_4Cl aq. sat. sol. was added and the suspension extracted with CH_2Cl_2 three times before being washed with H_2O and Brine. The combined organic layers were dried over $MgSO_4$, filtered, and concentrated before purification by column chromatography (50% EtOAc/hexane) to produce a white solid (Yield: 674 mg, 2.21 mmol, 41%).

HRMS (ESI) required for $C_{14}H_{14}N_3O_3S^-$: 304.0761, found 304.0754.

^{1}H NMR (400 MHz, chloroform-*d*) δ 11.58 (s, 1H, N**H**), 8.04 – 7.91 (m, 2H, Ar**H**), 7.75 – 7.62 (m, 2H, Ar**H**), 7.16 (d, *J* = 1.5 Hz, 1H, Ar**H**), 6.98 (d, *J* = 1.5 Hz, 1H, Ar**H**), 3.90 (s, 2H, C**H₂**), 3.88 (s, 3H, C**H₃**), 3.63 (s, 3H, C**H₃**).

^{13}C NMR (101 MHz, chloroform-*d*) δ 167.7 (**C**O), 166.8 (**C**O), 143.0 (Ar**C**), 142.7 (Ar**C**), 130.7 (Ar**C**), 127.3 (Ar**C**), 125.2 (Ar**C**), 122.9 (Ar**C**), 118.8 (Ar**C**), 52.0 (**C**H₃), 37.4 (**C**H₂), 33.6 (**C**H₃).

4-(2-((1-Methyl-1*H*-imidazol-2-yl)thio)acetamido)benzoic acid (50)
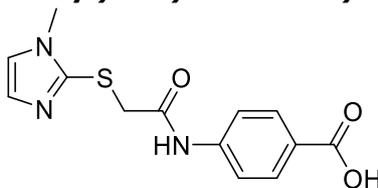
Compound **49** (305 mg, 1.00 mmol) was dissolved in H_2O (2 mL) with LiOH (59.9 mg, 2.5 mmol) and left to stir for 16 hours. The solution was acidified to pH 1 and extracted with EtOAc three times before being dried over $MgSO_4$, filtered, and concentrated to produce a pale-yellow solid which was taken forward without further purification (Yield: 286 mg, 0.98 mmol, 98%).

HRMS (ESI) required for $C_{13}H_{12}N_3O_3S^-$: 290.0605, found 290.0601.

¹H NMR (400 MHz, dimethylsulfoxide-d_6) δ 11.32 (s, 1H, N**H**), 7.93 – 7.79 (m, 2H, Ar**H**), 7.77 – 7.68 (m, 2H, Ar**H**), 7.68 – 7.60 (m, 1H, Ar**H**), 6.69 – 6.62 (m, 1H, Ar**H**), 4.31 (s, 2H, C**H₂**), 3.85 (s, 3H, C**H₃**).

¹³C NMR (101 MHz, dimethylsulfoxide-d_6) δ 167.3 (**C**O), 166.6 (**C**O), 143.1 (Ar**C**), 139.7 (Ar**C**), 131.6 (Ar**C**), 130.8 (Ar**C**), 121.5 (Ar**C**), 119.0 (Ar**C**), 114.4 (Ar**C**), 37.2 (**C**H₂), 35.6 (**C**H₃).

(4-(Aminomethyl)phenyl)methanol (52)[125]
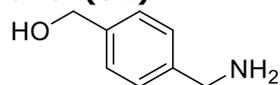
Lithium aluminium hydride (2.4 M in THF, 44 mL, 106 mmol) was diluted in THF (30 mL) at 0 °C using an ice water bath. 4-(aminomethyl)benzoic acid (4 g, 26.5 mmol) was dissolved in THF (10 mL) and added dropwise, forming a white suspension and left to stir for 30 mins. The reaction was heated to reflux slowly over 1 hour, turning pale-blue at 30 °C then purple at 45 °C. The suspension was left to reflux at 70 °C for 16 hours, forming a white suspension. The reaction was quenched at 0 °C with MeOH before being filtered through Celite and washed with EtOAc. The product was concentrated and co-evaporated with toluene to produce a yellow oil which was taken forward without further purification (Yield: 3.02 g, 22.0 mmol, 83%).

HRMS (ESI) required for $C_8H_{12}NO^+$: 138.0913, found 138.0916.

^{1}H NMR (400 MHz, methanol-d_4) δ 7.31 (s, 4H, Ar**H**), 4.58 (s, 2H, C**H₂**), 3.77 (s, 2H, C**H₂**).

^{13}C NMR (101 MHz, methanol-d_4) δ 142.7 (Ar**C**), 141.4 (Ar**C**), 128.4 (Ar**C**), 128.2 (Ar**C**), 65.0 (**C**H₂), 46.5 (**C**H₂).

Note: ^{1}H and ^{13}C NMR consistent with reported values.

Benzyl (4-(hydroxymethyl)benzyl)carbamate (53)[126]
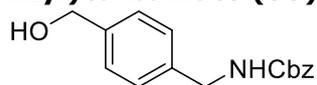
Compound **52** (18.2 g, 132 mmol) was suspended in H_2O (1.0 L) with K_2CO_3 (36.6 g, 264 mmol) and left to stir for 10 mins until solvated. CbzCl (32 mL, 225 mmol) was added dropwise and stirred for 16 hours. The suspension was extracted three times with CH_2Cl_2, dried over $MgSO_4$, filtered, and concentrated prior to purification by column chromatography (30-100% EtOAc/hexane) to provide a white solid (Yield: 32.2 g, 119 mmol, 90%).

HRMS (ESI) required for $C_{16}H_{17}NO_3Na^+$: 294.1101, found 294.1109.

¹H NMR (400 MHz, chloroform-d) δ 7.48 – 7.12 (m, 9H, Ar**H**), 5.22 - 4.98 (m, 3H, C**H₂** N**H**), 4.67 (s, 2H, C**H₂**), 4.37 (d, J = 5.6 Hz, 2H, C**H₂**), 1.77 (s, 1H, O**H**).

¹³C NMR (101 MHz, chloroform-d) δ 156.6 (**C**O), 140.4 (Ar**C**), 138.0 (Ar**C**), 136.6 (Ar**C**), 128.7 (Ar**C**), 128.3 (Ar**C**), 128.3 (Ar**C**), 127.9 (Ar**C**), 127.4 (Ar**C**), 67.0 (**C**H₂), 65.1 (**C**H₂), 45.0 (**C**H₂).

Note: ¹H and ¹³C NMR consistent with reported values.

Benzyl (4-(chloromethyl)benzyl)carbamate (54, adapted procedure)[127]
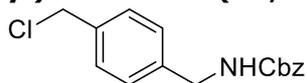
Compound **53** (32.2 g, 119 mmol) was dissolved in CH_2Cl_2 (450 mL) with PPh_3 (46.7 g, 178 mmol) and CCl_4 (90 mL, 920 mmol) and left to stir for 16 hours before being concentrated and purified by column chromatography (100% $CHCl_3$) to produce a white solid (Yield: 30.2 g, 104 mmol, 88%).

HRMS (ESI) required for $C_{16}H_{16}ClNO_2Na^+$: 312.0767, found 312.0769.

¹H NMR (400 MHz, chloroform-d) δ 7.53 – 7.12 (m, 9H), 5.24 – 5.06 (m, 3H), 4.60 (s, 2H), 4.41 (d, J = 6.0 Hz, 2H).

¹³C NMR (101 MHz, chloroform-d) δ 156.4 (**C**O), 138.8 (Ar**C**), 136.8 (Ar**C**), 136.4 (Ar**C**), 129.0 (Ar**C**), 128.6 (Ar**C**), 128.2 (Ar**C**), 128.2 (Ar**C**), 127.9 (Ar**C**), 67.0 (**C**H$_2$), 45.9 (**C**H$_2$), 44.8 (**C**H$_2$).

S-(4-(((((Benzyloxy)carbonyl)amino)methyl)benzyl)ethanethioate (55)
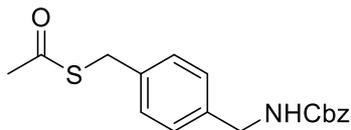
Compound **54** (30.2 g, 104 mmol) was dissolved in CH_2Cl_2 (420 mL) with thioacetic acid (10 mL, 146 mmol), NEt_3 (20 mL, 146 mmol) and DMAP (1.27 g, 10.4 mmol) and stirred for 16 hours. The reaction was washed with $NaHCO_3$ and H_2O before being dried over $MgSO_4$, filtered, concentrated and purified by column chromatography (10-30% EtOAc/hexane) to provide a yellow oil (Yield: 21.0 g, 63.8 mmol, 61%).

HRMS (ESI) required for $C_{18}H_{19}NO_3SNa^+$: 352.0983, found 352.0990.

^{1}H NMR (400 MHz, chloroform-*d*) δ 7.41 – 7.15 (m, 9H, Ar**H**), 5.22 – 4.99 (m, 3H, C**H$_2$**, N**H**), 4.35 (d, *J* = 6.0 Hz, 2H, NHC**H$_2$**), 4.09 (s, 2H, C**H$_2$**), 2.34 (s, 3H, C**H$_3$**).

^{13}C NMR (101 MHz, chloroform-*d*) δ 195.2 (**C**O), 156.5 (Ar**C**), 137.6 (Ar**C**), 137.1 (Ar**C**), 136.6 (Ar**C**), 129.2 (Ar**C**), 128.6 (Ar**C**), 128.3 (Ar**C**), 127.9 (Ar**C**), 67.0 (**C**H$_2$), 44.9 (**C**H$_2$), 33.2 (**C**H$_2$), 30.4 (**C**H$_3$).

Benzyl (4-(mercaptomethyl)benzyl)carbamate (56, adapted procedure)[128]
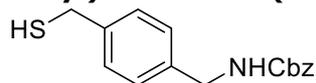
Compound **55** (21.0 g, 63.8 mmol) was dissolved in CH_2Cl_2 (450 mL) at 0 °C using an ice water bath and hydrazine monohydrate (31 mL, 178 mmol) added dropwise. The reaction warmed to room temperature and stirred for 4 hours. The reaction was diluted to 1 l in CH_2Cl_2 and washed with water and sat. aq. NH_4Cl before being dried over $MgSO_4$, filtered, concentrated and purified by column chromatography (15-100% EtOAc/hexane) to provide an off-white solid (Yield: 6.4 g, 22.2 mmol, 35%).

HRMS (ESI) required for $C_{16}H_{17}NO_2SNa^+$: 310.0872, found 310.0885.

¹H NMR (400 MHz, chloroform-*d*) δ 7.44 – 6.90 (m, 9H, Ar**H**), 5.14 – 4.93 (m, 2H, C**H₂**, N**H**), 4.28 (d, *J* = 6.0 Hz, 2H, C**H₂**), 3.64 (d, *J* = 7.5 Hz, 2H, SHC**H₂**), 1.66 (t, *J* = 7.5 Hz, 1H, S**H**).

¹³C NMR (101 MHz, chloroform-*d*) δ 156.4 (**C**O), 140.5 (Ar**C**), 137.3 (Ar**C**), 136.5 (Ar**C**), 129.7 (Ar**C**), 128.6 (Ar**C**), 128.4 (Ar**C**), 128.2 (Ar**C**), 127.9 (Ar**C**), 66.9 (**C**H₂), 44.8 (**C**H₂), 28.6 (**C**H₂).

Benzyl (4-(sulfamoylmethyl)benzyl)carbamate (57, adapted procedure)[102]
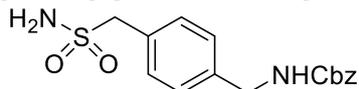
Compound **56** (2.00 g, 6.96 mmol) was dissolved in MeCN (60 nL) with H_2O (0.31 mL, 17.3 mmol) and tetrabutylammonium chloride (5.80 g, 21.0 mmol). *N*-chlorosuccinimide (2.80 g, 20.8 mmol) was added in portions over 5 mins resulting in slight heat evolution. After stirring for 110 mins, NH_4OH (15 M in water, 2.0 mL, 29.1 mmol) was added dropwise and stirred for 10 mins followed by addition of NH_4Cl aq. sat. sol (8.0 mL). After stirring for 50 mins, the solution was extracted with EtOAc three times and the combined organic layers washed with aq. 1M NaOH, H_2O and brine before being dried over $MgSO_4$, filtered, and concentrated under reduced pressure to provide a white powder (Yield: 1.01 g 3.03 mmol, 44%).

HRMS: required for $C_{16}H_{18}N_2O_4SNa^+$: 357.0879, found 357.0892.

^{1}H NMR (400 MHz, dimethylsulfoxide-d_6) δ 7.84 (t, *J* = 6.2 Hz, 1H, N**H**), 7.42 – 7.19 (m, 9H, Ar**H**), 6.81 (s, 2H, N**H₂**), 5.04 (s, 2H, C**H₂**), 4.26- 4.16 (m, 4H, C**H₂**).

^{13}C NMR (101 MHz, dimethylsulfoxide-d_6) δ 156.4 (**C**O), 139.5 (Ar**C**), 137.2 (Ar**C**), 130.7 (Ar**C**), 129.3 (Ar**C**), 128.4 (Ar**C**), 127.8 (Ar**C**), 127.8 (Ar**C**), 126.9 (Ar**C**), 65.4 (**C**H₂), 59.9 (**C**H₂), 43.5 (**C**H₂).

(4-(Aminomethyl)phenyl)methanesulfonamide (58)
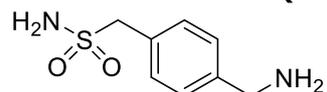
Compound **57** (563 mg, 1.68 mmol) was dissolved in CH_2Cl_2 (20 mL) with $Pd(OAc)_2$ (20.3 mg, 0.09 mmol), NEt_3 (36.0 µl, 0.260 mmol) and triethylsilane (434 µl, 2.72 mmol) and stirred for 48 hours. The reaction was filtered through Celite and stirred in sat. aq. NH_4Cl for 15 mins before the layers were separated and the aqueous phase extracted twice with EtOAc. The combined organic layers were dried over $MgSO_4$, filtered, concentrated and purified by column chromatography (5-20% MeOH/CH_2Cl_2) to provide a white solid (Yield: 200 mg, 1.00 mmol, 62%).

HRMS (ESI) required for $C_8H_{13}N_2O_2S^+$: 201.0692, found 201.0698.

¹H NMR (400 MHz, dimethylsulfoxide-d_6) δ 7.48 – 7.13 (m, 4H), 6.77 (br, 2H), 4.23 (s, 2H), 3.75 (s, 2H), 1.84 (s, 1H).

¹³C NMR (101 MHz, dimethylsulfoxide-d_6) δ 142.6 (Ar**C**), 130.5 (Ar**C**), 128.9 (Ar**C**), 127.1 (Ar**C**), 60.0 (**C**H₂), 44.9 (**C**H₂).

4-(2-((1-Methyl-1*H*-imidazol-2-yl)thio)acetamido)-*N*-(4-(sulfamoylmethyl) benzyl)benzamide (36)
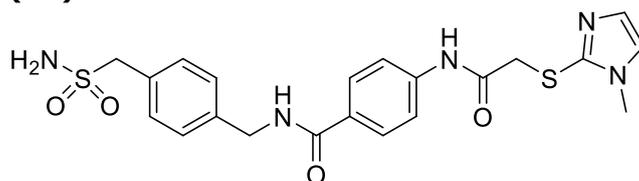
Compound **58** (79.6 mg, 0.45 mmol) was dissolved in DMF (5 mL) with compound **50** (157 mg, 0.54 mmol), K_2CO_3 (186 mg, 1.34 mmol) and HCTU (278 mg, 0.67 mmol) and stirred for 16 hours. The reaction was diluted four-fold in CH_2Cl_2 and washed successively with $NaHCO_3$ and brine then dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The reaction provided 140 mg crude material, 100 mg of which was purified by reverse-phase HPLC purification (0% B over 5 mins, 0-15% B over 1 min, 15% B over 10 mins, 15-100% B over 1 min, 100% B over 6 mins) to provide a white powder (10.5 mg, 22.2 µmol, 5%).

HRMS (ESI) required for $C_{21}H_{24}N_5O_4S_2^+$: 474.1264, found 474.1267.

¹H NMR (400 MHz, dimethylsulfoxide-d_6) δ 10.61 (s, 1H, N**H**), 8.95 (t, *J* = 6.1, 1H, N**H**), 7.85 (m, 2H, Ar**H**), 7.63 (m, 2H, Ar**H**) 7.31 (s, 4H, Ar**H**), 7.25 (d, *J* = 1.3, 1H, Ar**H**), 6.96 (d, *J* = 1.3, 1H, Ar**H**), 6.81 (s, 2H, N**H₂**), 4.46 (d, *J* = 6.1, 2H, C**H₂**), 4.23 (s, 2H, C**H₂**), 3.89 (s, 2H, C**H₂**), 3.60 (s, 3H, C**H₃**).

¹³C NMR (101 MHz, dimethylsulfoxide-d_6) δ 167.45 (**C**O), 166.23 (**C**O), 141.92 (Ar**C**), 139.97 (Ar**C**), 139.91 (Ar**C**), 131.14 (Ar**C**), 129.60 (Ar**C**), 129.48 (Ar**C**), 129.03 (Ar**C**), 128.66 (Ar**C**), 127.61 (Ar**C**), 124.09 (Ar**C**), 118.80 (Ar**C**), 60.33 (**C**H₂), 42.77 (**C**H₂), 38.84 (**C**H₂), 33.46 (**C**H₃).

6-(4-Methyl-1,4-diazepan-1-yl)nicotinamide (61)
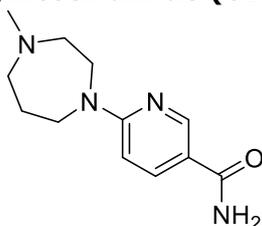
1-Methylhomopiperazine (2.0 mL, 15.8 mmol) was dissolved in MeCN (60 mL) with 6-chloronicotinamide (2.35 g, 15.0 mmol) and DiPEA (7.84 mL, 45.0 mmol) and stirred at 80 °C for 48 hours. DMF (10 mL) was added, and the reaction stirred for a further 48 hours. The reaction was concentrated under reduced pressure and diluted with 1 M NaOH before extraction with EtOAc three times. The organic phases were combined, dried over MgSO₄, filtered, and concentrated under reduced pressure. The product was purified using a silica plug flushed successively with CH₂Cl₂, EtOAc and MeOH to produce the desired product as a yellow solid (1.17 g, 5.01 mmol, 33%).

HRMS (ESI) required for $C_{12}H_{19}N_4O^+$: 235.1553, found 235.1567.

¹H NMR (400 MHz, dimethylsulfoxide-d_6) δ 8.58 (d, J = 2.3 Hz, 1H, Ar**H**), 7.91 (dd, J = 9.0, 2.3 Hz, 1H, Ar**H**), 7.70 (br, 1H, N**H₂**), 7.06 (br, 1H, N**H₂**), 6.62 (d, J = 9.0 Hz, 1H, Ar**H**), 3.80 – 3.70 (m, 2H, C**H₂**), 3.70 – 3.57 (m, 2H, C**H₂**), 2.61 – 2.54 (m, 2H, C**H₂**), 2.47 – 2.40 (m, 2H, C**H₂**), 2.24 (s, 3H, C**H₃**), 1.96 – 1.76 (m, 2H, C**H₂**).

¹³C NMR (101 MHz, dimethylsulfoxide-d_6) δ 166.9 (**C**O), 159.0 (Ar**C**), 148.4 (Ar**C**), 136.6 (Ar**C**), 117.0 (Ar**C**), 104.3 (Ar**C**), 57.2 (**C**H₂), 56.6 (**C**H₂), 46.3 (**C**H₂), 46.1 (**C**H₂), 46.0 (**C**H₃), 26.8 (**C**H₂).

(6-(4-Methyl-1,4-diazepan-1-yl)pyridin-3-yl)methanamine (62)
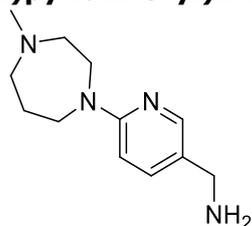
Compound **61** (1.17 g, 5.01 mmol) was dissolved in dry THF (5 mL) and cooled to 0 °C using an ice water bath. BH$_3$.THF (20 mL, 1.0 M) was added dropwise and the reaction heated to 40 °C and left to stir for 5 days. After observing completion by TLC, the reaction was cooled to 0 °C and quenched with 5 M HCl. The mixture was stirred at room temperature for 1 hour then at 60 °C for 20 mins. The solvent was evaporated under reduced pressure and the product extracted from 1 M NaOH with EtOAc three times. The organic phases were combined, dried over MgSO$_4$, filtered, and concentrated under reduced pressure to afford the product as a pale-yellow oil (Yield: 281 mg, 1.27 mmol, 26%).

HRMS (ESI) required for C$_{12}$H$_{21}$N$_4^+$: 221.1761, found 221.1764.

^{1}H NMR (400 MHz, dimethylsulfoxide-d_6) δ 7.96 (d, J = 2.4 Hz, 1H, Ar**H**), 7.45 (dd, J = 8.8, 2.4 Hz, 1H, Ar**H**), 6.54 (d, J = 8.8 Hz, 1H, Ar**H**), 3.85 – 3.65 (m, 2H, C**H$_2$**), 3.62 – 3.49 (m, 4H, C**H$_2$**), 2.64 – 2.53 (m, 2H, C**H$_2$**), 2.45 – 2.38 (m, 2H, C**H$_2$**), 2.25 – 2.21 (m, 5H, N**H$_2$**, C**H$_3$**), 1.98 – 1.75 (m, 2H, C**H$_2$**).

^{13}C NMR (101 MHz, dimethylsulfoxide-d_6) δ 157.4 (Ar**C**), 146.8 (Ar**C**), 137.5 (Ar**C**), 126.5 (Ar**C**), 105.4 (Ar**C**), 58.0 (**C**H$_2$), 57.2 (**C**H$_2$), 46.6 (**C**H$_3$), 46.5 (**C**H$_2$), 46.3 (**C**H$_2$), 43.1 (**C**H$_2$), 27.5 (**C**H$_2$).

2-Chloro-*N*-((6-(4-methyl-1,4-diazepan-1-yl)pyridin-3-yl)methyl)-5-(methylthio)benzamide (37)
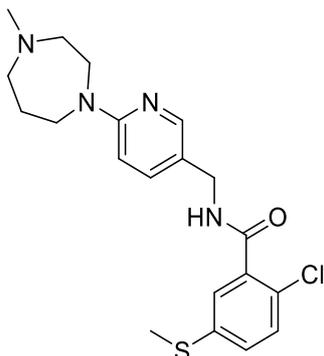
2-Chloro-5-(methylthio)benzoic acid (122 mg, 0.600 mmol) was dissolved in DMF (1.5 mL) with HCTU (414 mg, 1.00 mmol) and NEt$_3$ (0.20 mL, 1.50 mmol) for 10 mins, producing a light brown solution. **62** (110 mg, 0.50 mmol) was added and the solution stirred for 16 hours. The reaction was split between EtOAc and 1 M NaOH and the organic layer washed with water and brine before being dried over MgSO$_4$, filtered, and concentrated under reduced pressure to afford 220 mg crude product. 105 mg was purified using reverse-phase HPLC (15% B over 3 mins, 15-25% B over 12 min, 25-100% B over 1 min, 100% B over 9 mins) to produce a white powder (Yield: 11.0 mg, 27.2 μmol, 5%).

HRMS (ESI) required for C$_{20}$H$_{26}$ClN$_4$OS$^+$: 405.1510, found 405.1528.

^{1}H NMR (400 MHz, dimethylsulfoxide-*d$_6$*) δ 8.85 (t, *J* = 5.8, 1H, N**H**), 8.02 (d, J = 2.4 Hz, 1H, Ar**H**), 7.48 (dd, *J* = 8.8, 2.4 Hz, 1H, Ar**H**), 7.41 (d, *J* = 8.4 Hz, 1H, Ar**H**), 7.31 (dd, *J* = 8.4, 2.4 Hz, 1H, Ar**H**), 7.25 (d, *J* = 2.4 Hz, 1H, Ar**H**), 6.60 (d, *J* = 8.8 Hz, 1H, Ar**H**), 4.28 (d, *J* = 5.8, 2H, C**H$_2$**), 3.73 (m, 2H, C**H$_2$**), 3.56 (t, *J*= 6.2 Hz, 2H, C**H$_2$**), 2.67 – 2.60 (m, 2H, C**H$_2$**), 2.29 (s, 3H, C**H$_3$**), 1.94 – 1.85 (m, 2H, C**H$_2$**).

^{13}C NMR (101 MHz, dimethylsulfoxide-*d$_6$*) δ 165.7 (**C**O), 157.1 (Ar**C**), 146.8 (Ar**C**), 137.6 (Ar**C**), 137.3 (Ar**C**), 137.1 (Ar**C**), 130.0 (Ar**C**), 127.8 (Ar**C**), 126.3 (Ar**C**), 125.7 (Ar**C**), 121.4 (Ar**C**), 105.1 (Ar**C**), 57.3 (**C**H$_2$), 56.5 (**C**H$_2$), 46.1 (**C**H$_2$), 45.8 (**C**H$_3$), 45.3 (**C**H$_2$), 39.9 (**C**H$_2$), 26.6 (**C**H$_2$), 14.7 (**C**H$_3$).

Methyl 3-(chlorosulfonyl)benzoate (64)[104]
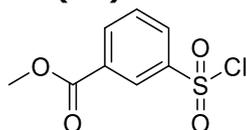
3-(Chlorosulfonyl)benzoic acid (8.80 g, 40.0 mmol) was dissolved in CH_2Cl_2 (400 mL) with DMF (0.20 mL). Oxalyl chloride (5.2 mL 60.0 mmol) was added dropwise and stirred 1.5 hours. MeOH (8.1 mL, 200 mmol) was added dropwise and left to stir for a further 15 mins. The reaction was concentrated under reduced pressure and purified by column chromatography (20% EtOAc/hexane) to provide a pale-yellow solid (Yield: 3.10 g, 13.2 mmol, 33%).

HRMS (ESI) required for $C_8H_8ClO_4S^+$: 234.9826, found 234.9833.

^{1}H NMR (400 MHz, chloroform-*d*) δ 8.78 – 8.58 (m, 1H, Ar**H**), 8.47 – 8.35 (m, 1H, Ar**H**), 8.28 – 8.15 (d, 1H, Ar**H**), 7.74 (app t, *J* = 7.9 Hz, 1H, Ar**H**), 4.00 (s, 3H, C**H₃**).

^{13}C NMR (101 MHz, chloroform-*d*) δ 164.6 (**C**O), 144.8 (Ar**C**), 135.9 (Ar**C**), 132.1 (Ar**C**), 130.7 (Ar**C**), 130.1 (Ar**C**), 128.1 (Ar**C**), 52.9 (**C**H₃).

Note: ^{1}H NMR consistent with reported values.

Methyl (*R*)-3-(*N*-(1-phenylethyl)sulfamoyl)benzoate (65)
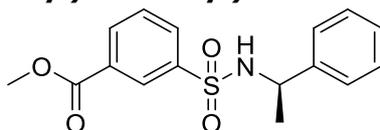
Compound **64** (551 mg, 2.35 mmol) was suspended in 1:1 THF:H_2O (10 mL) with K_2CO_3 (974 mg, 7.05 mmol) and (*R*)-1-phenylethan-1-amine (948 mg, 2.58 mmol) and stirred for 16 hours. THF was removed under reduced pressure and the resulting suspension acidified with 1 M HCl and extracted twice with EtOAc. After being washed with $NaHCO_3$ and brine, the product was dried over $MgSO_4$, filtered, concentrated under reduced pressure, and purified by column chromatography (20% EtOAc/hexane) to provide a clear oil (Yield: 528 mg, 1.65 mmol, 70%).

[α]$^{20}_D$: +36.0 (c 1.0, $CHCl_3$).

HRMS (ESI) required for $C_{16}H_{16}NO_4S^-$: 318.0806, found 318.0811.

^{1}H NMR (400 MHz, chloroform-*d*) δ 8.32 – 8.28 (m, 1H, Ar**H**), 8.11 (d, *J* = 7.8 Hz, 1H, Ar**H**), 7.84 (d, *J* = 7.8 Hz, 1H, Ar**H**), 7.42 (t, *J* = 7.8 Hz, 1H, Ar**H**), 7.12 (m, 3H, Ar**H**), 7.08 – 7.01 (m, 2H, Ar**H**), 5.08 (d, *J* = 6.9 Hz, 1H, N**H**), 4.55 (app p, *J* = 6.9 Hz, 1H, C**H**), 3.92 (s, 3H, C**H₃**), 1.45 (d, *J* = 6.9 Hz, 3H, C**H₃**).

^{13}C NMR (101 MHz, chloroform-*d*) δ 165.6 (**C**O), 141.5 (Ar**C**), 141.5 (Ar**C**), 133.2 (Ar**C**), 131.1 (Ar**C**), 131.0 (Ar**C**), 129.1 (Ar**C**), 128.7 (Ar**C**), 128.3 (Ar**C**), 127.7 (Ar**C**), 126.3 (Ar**C**), 54.1 (**C**H), 52.6 (**C**H₃), 23.8 (**C**H₃).

(*R*)-3-(*N*-(1-Phenylethyl)sulfamoyl)benzoic acid (66)
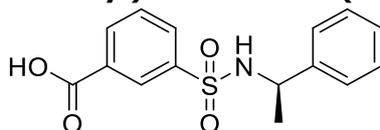
Compound **65** (528 mg, 1.65 mmol) was suspended in H_2O (35 mL) with LiOH (98.9 mg, 4.13 mmol) and stirred for 3 hours. The reaction was acidified to pH 1 with 1 M HCl and extracted three times with CH_2Cl_2. The product was dried over $MgSO_4$, filtered, and concentrated under reduced pressure to yield a white solid (Yield: 386 mg, 1.26 mmol, 77%).

[α]$^{20}_D$: +32.0 (c 1.0, $CHCl_3$).

HRMS (ESI) required for $C_{15}H_{14}NO_4S^-$: 304.0649, found 304.0644.

^{1}H NMR (400 MHz, methanol-d_4) δ 8.28 – 8.14 (m, 1H, Ar**H**), 8.05 (d, *J* = 7.8 Hz, 1H, Ar**H**), 7.82 (d, *J* = 7.8 Hz, 1H, Ar**H**), 7.44 (t, *J* = 7.8 Hz, 1H, Ar**H**), 7.18 – 6.95 (m, 5H, Ar**H**), 4.46 (q, *J* = 7.0 Hz, 1H, C**H**), 1.37 (d, *J* = 7.0 Hz, 3H, C**H₃**).

^{13}C NMR (101 MHz, methanol-d_4) δ 168.0 (**C**O), 143.7 (Ar**C**), 143.5 (Ar**C**), 133.8 (Ar**C**), 132.7 (Ar**C**), 131.8 (Ar**C**), 130.1 (Ar**C**), 129.3 (Ar**C**), 129.2 (Ar**C**), 128.1 (Ar**C**), 127.2 (Ar**C**), 55.1 (**C**H), 24.1 (**C**H₃).

Methyl (*R*)-*N*-methyl-*N*-(3-(*N*-(1-phenylethyl)sulfamoyl)benzoyl)glycinate (67)
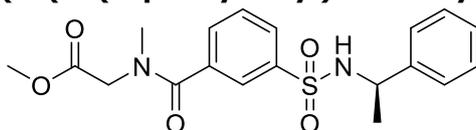
Compound **66** (386 mg, 1.30 mmol) was dissolved in DMF (20 mL) with sarcosine methyl ester (194 mg, 1.39 mmol), K_2CO_3 (698 mg, 5.05 mmol) and HCTU (784 mg, 1.89 mmol) and stirred for 16 hours to form a pale-yellow suspension. The reaction was concentrated under reduced pressure and diluted with EtOAc before being washed with 1 M HCl, $NaHCO_3$ aq. sat. sol., and brine. The crude product was dried over $MgSO_4$, filtered, concentrated under reduced pressure, and purified by column chromatography (50% EtOAc/hexane) to produce an off white solid (Yield: 140 mg, 388 μmol, 28%).

[α]$^{20}_D$: +20.0 (c 1.0, CHCl$_3$).

HRMS (ESI) required for $C_{19}H_{23}N_2O_5S^+$: 391.1354, found 391.1367.

^{1}H NMR (400 MHz, chloroform-*d*) δ 7.90 – 7.77 (m, 1H, Ar**H**), 7.75 - 7.65 (m, 1H, Ar**H**), 7.64 – 7.48 (m, 1H, Ar**H**), 7.45 – 7.31 (m, 1H, Ar**H**), 7.21 – 7.04 (m, 5H, Ar**H**), 5.22 – 5.02 (m, 1H, N**H**), 4.64 – 4.42 (m, 1H, C**H**), 4.26 (s, 1H), 3.87 (s, 1H), 3.81 – 3.73 (m, 3H), 3.12 (s, 1H), 2.95 (s, 2H), 1.44 (d, *J* = 6.7 Hz, 3H, CHC**H₃**).

^{13}C NMR (101 MHz, chloroform-*d*) δ 170.5 (**C**O), 169.3 (**C**O), 141.8 (Ar**C**), 141.4 (Ar**C**), 136.3 (Ar**C**), 131.3 (Ar**C**), 129.2 (Ar**C**), 128.7 (Ar**C**), 128.4 (Ar**C**), 127.8 (Ar**C**), 126.3 (Ar**C**), 125.9 (Ar**C**), 54.0 (**C**H), 52.5 (**C**H₃), 49.3 (**C**H₂), 38.8 (**C**H₃), 23.8 (**C**H₃).

Note: Rotamers observed, peak assignments confirmed in final product through high temperature ^{1}H NMR.

(*R*)-*N*-Methyl-*N*-(3-(*N*-(1-phenylethyl)sulfamoyl)benzoyl)glycine (68)
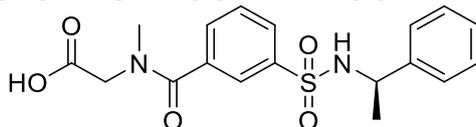
Compound **67** (143 mg, 366 μmol) was suspended in H_2O (5 mL) with LiOH (30.0 mg, 1.25 mmol) and stirred for 16 hours. The reaction was acidified to pH 1 with 1 M HCl and extracted three times with CH_2Cl_2. The product was dried over $MgSO_4$ and concentrated under reduced pressure to produce a viscous clear oil (Yield: 68.3 mg, 181 μmol, 50%).

[α]^{20}$_D$: +16.0 (c 1.0, CHCl$_3$).

HRMS (ESI) required for $C_{18}H_{19}N_2O_5S^-$: 375.1020, found 375.1031.

^{1}H NMR (400 MHz, methanol-d_4) δ 7.80 – 7.66 (m, 2H), 7.58 – 7.53 (m, 1H), 7.49 – 7.35 (m, 1H), 7.17 – 7.04 (m, 5H), 4.59 – 4.37 (m, 1H), 4.25 (s, 1H), 3.92 (s, 1H), 3.10 (s, 1H), 2.92 (s, 2H), 1.42 - 1.31 (m, 3H).

^{13}C NMR (101 MHz, methanol-d_4) δ 172.8 (**C**O), 172.3 (**C**O), 172.0 (**C**O), 143.9 (Ar**C**), 143.5 (Ar**C**), 143.4 (Ar**C**), 137.5 (Ar**C**), 137.3 (Ar**C**), 131.5 (Ar**C**), 130.9 (Ar**C**), 130.3 (Ar**C**), 130.2 (Ar**C**), 129.3 (Ar**C**), 129.3 (Ar**C**), 128.2 (Ar**C**), 127.2 (Ar**C**), 126.5 (Ar**C**), 126.1 (Ar**C**), 54.9 (**C**H), 50.1 (**C**H$_2$), 39.3 (**C**H$_3$), 34.9 (**C**H$_3$), 24.2 (**C**H$_3$), 24.1 (**C**H$_3$).

Note: Rotamers observed, peak assignments confirmed in final product through high temperature ^{1}H NMR.

(*R*)-*N*-(2-((4-Methoxyphenyl)amino)-2-oxoethyl)-*N*-methyl-3-(*N*-(1-phenyl ethyl)sulfamoyl)benzamide (38)
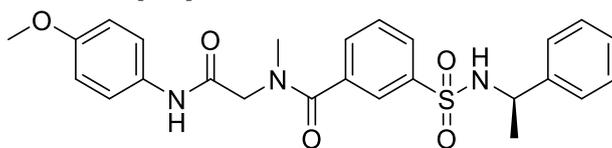
Compound **68** (114 mg, 304 µmol) was dissolved in DMF (5 mL) with *p*-anisidine (52.4 mg, 425 µmol), K_2CO_3 (210 mg, 1.52 mmol) and HCTU (226 mg, 547 µmol) and stirred for 16 hours. The reaction was concentrated and dissolved in EtOAc then washed with 1 M HCl, $NaHCO_3$ aq. sat. sol. and brine. The product was dried over $MgSO_4$, filtered, and concentrated under reduced pressure to afford 118 mg crude material. 60 mg crude material was purified using reverse-phase HPLC (42% B over 5 mins, 42-60% B over 1 min, 60% B over 10 mins, 60-100% B over 1 min, 100% B over 6 mins) to produce a white powder (Yield: 24.1 mg, 50.3 µmol, 17%).

HRMS: (ESI) required for $C_{25}H_{26}N_3O_5S^-$: 480.1599, found 480.1592.

[α]^{20}$_D$: +19.9 (c 1.0, $CHCl_3$).

^{1}H NMR (400 MHz, dimethylsulfoxide-d_6) δ 10.12 – 9.86 (m, 1H, N**H**), 8.50 – 8.21 (m, 1H, N**H**), 7.87 – 7.63 (m, 2H, Ar**H**), 7.63 – 7.39 (m, 4H, Ar**H**), 7.28 – 7.01 (m, 5H, Ar**H**), 7.01 – 6.73 (m, 2H, Ar**H**), 4.48 – 4.18 (m, 2H, C**H**, C**H₂**), 3.95 (s, 1H, C**H₂**) 3.80 – 3.65 (m, 3H, C**H₃**), 2.99 (s, 1H, C**H₃**), 2.90 (s, 2H, C**H₃**), 1.25 (d, *J* = 7.0 Hz, 2H, CHC**H₃**) 1.14 (d, *J* = 7.0 Hz, 1H, CHC**H₃**).

^{13}C NMR (101 MHz, dimethylsulfoxide-d_6) δ 170.4 (**C**O), 169.7 (**C**O), 166.5 (**C**O), 166.4 (**C**O), 155.9 (Ar**C**), 155.6 (Ar**C**), 143.5 (Ar**C**), 143.4 (Ar**C**), 142.2 (Ar**C**), 142.2 (Ar**C**), 137.3 (Ar**C**), 137.0 (Ar**C**), 132.5 (Ar**C**), 132.0 (Ar**C**), 130.7 (Ar**C**), 130.4 (Ar**C**), 129.5 (Ar**C**), 128.5 (Ar**C**), 127.7 (Ar**C**), 127.6 (Ar**C**), 127.3 (Ar**C**), 127.2 (Ar**C**), 126.4 (Ar**C**), 125.3 (Ar**C**), 125.2 (Ar**C**), 121.4 (Ar**C**), 121.0 (Ar**C**), 114.3 (Ar**C**), 55.6 (**C**H₃), 54.4 (**C**H₃), 53.4 (**C**H), 53.3 (**C**H₂), 50.8 (**C**H₂), 39.1 (**C**H₃), 34.4 (**C**H₃), 24.1 (**C**H₃), 23.8 (**C**H₃).

Note: Rotamers observed, peak assignments confirmed through high temperature ^{1}H NMR (**Figure 49**).

2–Cyclohexylacetyl chloride (70, modified procedure)[106]
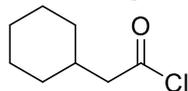
To a yellow solution of cyclohexylacetic acid (4.00 g, 28.1 mmol) in chloroform (60 mL) was added CH_2Cl_2 (50 mL) and DMF (0.5 mL) and the reaction cooled to 0 °C sing an ice water bath. Oxalyl chloride (7.2 mL, 84.4 mmol) was added dropwise over 10 mins, turning the solution colourless with gas evolution. The solution was warmed to 50 °C and stirred for 3 hours before being concentrated under reduced pressure, producing a brown solid. The product was taken forward without further purification (Yield: 4.45 g, 27.7 mmol, 99%).

¹H NMR (400 MHz, chloroform-*d*) δ 2.75 (d, *J* = 7.0 Hz, 2H, COC**H₂**), 1.99 – 1.83 (m, 1H, C**H**), 1.81 – 1.60 (m, 4H, C**H₂**), 1.38 – 1.08 (m, 4H, C**H₂**), 1.07 – 0.90 (m, 2H, C**H₂**).

¹³C NMR (101 MHz, chloroform-*d*) δ 173.1 (**C**O), 54.6 (**C**H₂), 35.2 (**C**H), 32.6 (**C**H₂), 26.0 (**C**H₂), 25.9 (**C**H₂).

Note: ¹H and ¹³C NMR consistent with reported values (spectra not recorded by title paper).[129]

5-(2-Cyclohexyl-1-hydroxyethylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione (71, modified procedure)[105,106]

To a solution of **70** (4.45 g, 27.7 mmol) in CH_2Cl_2 (30 mL) at 0 °C in an ice water bath was added DMAP (7.18 g, 59.0 mmol) and Meldrum's acid (4.04 g, 28.1 mmol). The solution was stirred at room temperature for 2.5 hours. The solution was washed with 5% aq. potassium bisulfate three times and brine once then dried over $MgSO_4$, filtered, and concentrated under reduced pressure to yield an off-white solid (Yield: 7.20 g, 26.7 mmol, 97%).

HRMS (ESI) required for $C_{14}H_{19}O_5^-$: 267.1238, found 267.1236.

¹H NMR (400 MHz, chloroform-*d*) δ 15.27 (s, 1H, O**H**), 2.97 (d, *J* = 7.0 Hz, 2H, C**H₂**), 1.94 – 1.79 (m, 1H, C**H**), 1.78 – 1.59 (m, 10H, C**H₂**, C**H₃**), 1.35 – 0.97 (m, 6H, C**H₂**).

¹³C NMR (101 MHz, chloroform-*d*) δ 197.6 (**C**OH), 170.5 (**C**O), 104.7 (**C**), 92.0 (**C**), 42.6 (**C**H₂), 36.7 (**C**H), 33.1 (**C**H₂), 26.8 (**C**H₃), 26.1 (**C**H₂), 26.0 (**C**H₂).

Note: ¹H NMR consistent with reported values (spectra not recorded by title paper).[130]

Ethyl 4-cyclohexyl-3-oxobutanoate (72, modified procedure)[105,106]
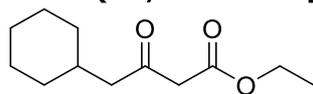
Compound **71** (7.20 g, 26.7 mmol) was dissolved in EtOH (70 mL) and stirred at reflux for 16 hours. The solution was concentrated under reduced pressure and taken forward without further purification as a viscous clear oil (Yield: 5.56 g, 26.2 mmol, 98%).

HRMS (ESI) required for $C_{12}H_{20}O_3Na^+$: 235.1305, found 235.1330.

^{1}H NMR (400 MHz, chloroform-d) δ 4.17 (q, J = 7.1 Hz, 2H, OC**H₂**), 3.39 (s, 2H, C**H₂**), 2.38 (d, J = 6.8 Hz, 2H, C**H₂**), 1.91 – 1.76 (m, 1H, C**H**), 1.76 – 1.56 (m, 5H, C**H₂**), 1.32 – 1.19 (m, 5H, C**H₃,** C**H₂**), 1.18 – 1.05 (m, 1H, C**H₂**), 0.98 – 0.84 (m, 2H, C**H₂**).

^{13}C NMR (101 MHz, chloroform-d) δ 202.5 (**C**O), 167.2 (**C**O), 61.3 (**C**H₂), 50.6 (**C**H₂), 49.8 (**C**H₂), 33.5 (**C**H), 33.0 (**C**H₂), 26.1 (**C**H₂), 26.0 (**C**H₂), 14.1 (**C**H₃).

Note: ^{1}H and ^{13}C NMR consistent with reported values.[131]

Ethyl 3-amino-4-cyclohexylbut-2-enoate (73, modified procedure)[105,106]
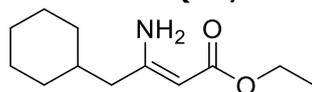
To a solution of **72** (5.60 g, 26.2 mmol) in EtOH (13 mL) cooled with an ice water bath at 0 °C was added 30% aq NH$_4$OH (30 mL, 210 mmol). The solution was left to warm to room temperature and stirred for 16 hours. The solution was concentrated under reduced pressure. The crude product was purified by column chromatography (0-6% MeOH/CH$_2$Cl$_2$) to provide a pale-yellow oil (Yield: 1.62 g, 7.79 mmol, 30%).

HRMS (ESI) required for C$_{12}$H$_{22}$NO$_2{}^+$: 212.1645, found 212.1656.

^{1}H NMR (400 MHz, chloroform-*d*) δ 4.46 (s, 1H, C**H**), 4.07 (q, *J* = 7.1 Hz, 2H, C**H$_2$**), 1.94 (d, *J* = 7.3 Hz, 2H, C**H$_2$**), 1.77 – 1.57 (m, 5H, C**H$_2$**), 1.50 – 1.40 (m, 1H, C**H**), 1.25 – 1.05 (m, 6H, C**H$_3$**, C**H$_2$**), 0.96 – 0.81 (m, 2H, C**H$_2$**).

^{13}C NMR (101 MHz, chloroform-*d*) δ 170.3 (**C**O), 162.7 (**C**NH$_2$), 84.1 (**C**H), 58.5 (**C**H$_2$), 44.6 (**C**H$_2$), 37.0 (**C**H), 33.1 (**C**H$_2$), 26.3 (**C**H$_2$), 26.1 (**C**H$_2$), 14.6 (**C**H$_3$).

6-(Cyclohexylmethyl)-4-hydroxy-3-phenylpyridin-2(1*H*)-one (NITD-564)[105]
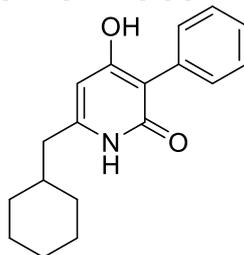
Compound **73** (99.2 mg, 0.47 mmol) was dissolved in diethyl 2-phenylmalonate (100 µL, 0.47 M) and heated to 220 °C for 45 mins. The crude product was then dissolved in aq. 2 M NaOH and microwave irradiated at 140 °C, 150 W in a sealed tube for 1 hour. The solution was cooled to room temperature and acidified to pH 6 with aq. 1 M HCl. The precipitated product (230 mg) was collected, and 106 mg purified by reversed phase HPLC (30-95% B over 25 mins) to provide a white solid (8.0 mg, 28.2 µmol, 6%).

HRMS (ESI) required for $C_{18}H_{20}NO_2^-$: 282.1500, found 282.1501.

¹H NMR (400 MHz, dimethylsulfoxide-d_6) δ 11.02 (s, 1H), 10.46 (s, 1H), 7.42 – 7.36 (m, 2H, Ar**H**), 7.29 – 7.25 (m, 2H, Ar**H**), 7.18 – 7.12 (m, 1H, Ar**H**), 5.78 (s, 1H, Ar**H**), 2.26 (d, J = 7.0 Hz, 2H, C**H₂**), 1.77 – 1.45 (m, 6H, C**H**), 1.32 – 1.02 (m, 3H, C**H₂**, C**H**), 1.02 – 0.77 (m, 2H, C**H₂**).

¹³C NMR (101 MHz, dimethylsulfoxide-d_6) δ 163.6 (**C**OH), 163.4 (**C**O), 146.8 (Ar**C**), 134.4 (Ar**C**), 130.9 (Ar**C**), 127.1 (Ar**C**), 125.7 (Ar**C**), 108.5 (Ar**C**), 98.6 (Ar**C**), 37.0 (**C**H₂), 32.4 (**C**H), 26.0 (**C**H₂), 25.7 (**C**H₂).

Note: ¹H and ¹³C NMR consistent with reported values.

1-(3-Methoxyphenyl)-*N*-methylmethanamine (80, adapted procedure)[107]
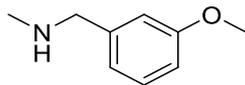
(3-Methoxyphenyl)methanamine (3.00 g, 21.9 mmol) was dissolved in THF (32 mL) with ethyl formate (7.3 mL, 98.4 mmol) and stirred 47 °C for 16 h. The mixture was concentrated under reduced pressure before being dissolved in THF (10 mL) and cooled to 0 °C using an ice water bath. A solution of LiAlH$_4$ (66.0 mmol, 2.4 M in THF) in THF (6M) was added dropwise at 0 °C using an ice water bath before being warmed slowly to 70 °C over 1 hour and refluxed for 3.5 h. The completed reaction was cooled to 0 °C and quenched with 10% aq. KOH solution, filtered and extracted twice with CH$_2$Cl$_2$. The solution was dried over MgSO$_4$ and concentrated under reduced pressure to provide a white solid (Yield: 2.61 g, 17.3 mmol, 79%) which was taken forward without further purification.

HRMS (ESI) required for C$_9$H$_{14}$NO$^+$: 152.1070, observed: 152.1080.

¹H NMR (400 MHz, dimethylsulfoxide-*d*$_6$) δ 7.21 (app t, *J* = 7.9 Hz, 1H, Ar**H**), 6.99 – 6.84 (m, 2H, Ar**H**), 6.84 – 6.65 (m, 1H, Ar**H**), 3.73 (s, 3H, C**H₃**), 3.63 (s, 2H, C**H₂**), 3.20 (s, 1H, N**H**), 2.26 (s, 3H, C**H₃**).

¹³C NMR (101 MHz, dimethylsulfoxide-*d*$_6$) δ 159.7 (Ar**C**), 142.3 (Ar**C**), 129.6 (Ar**C**), 120.7 (Ar**C**), 113.9 (Ar**C**), 112.7 (Ar**C**), 55.4 (**C**H$_2$), 55.2 (**C**H$_3$), 35.7 (**C**H$_3$).

3-((3-Methoxybenzyl)(methyl)carbamoyl)benzenesulfonyl chloride (81)
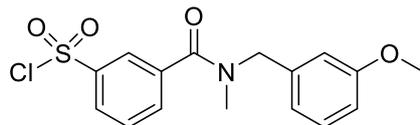
Compound **80** (454 mg, 3.00 mmol) was dissolved in CH_2Cl_2 (10 mL) with DCC (1.86 g, 4.50 mmol) and DMAP (41.0 mg, 300 µmol) then stirred at room temperature for 10 mins. 3-(chlorosulfonyl)benzoic acid (0.66 g, 3.00 mmol) was then added and stirred for 16 h. The contents were filtered then concentrated under reduced pressure. Following flash column chromatography (0-50% EtOAc/hexane) 681 mg of a pale-yellow solid was obtained, co-eluting with a 34 mol% impurity of dicyclohexylurea (calculated product yield: 452 mg, 1.28 mmol, 43%).

HRMS (ESI) required for $C_{16}H_{16}NO_4SClNa^+$: 376.0381, observed: 376.0387.

¹H NMR (400 MHz, chloroform-*d*) δ 8.03 – 7.84 (m, 2H, Ar**H**), 7.55 – 7.37 (m, 2H, Ar**H**), 7.36 – 7.15 (m, 2H, Ar**H**), 7.04 – 6.52 (m, 3H, Ar**H**), 4.73 (s, 1H), 4.48 (s, 1H), 3.95 - 3.68 (m, 3H), 3.54 (m, 1H, dicyclohexylurea), 3.07 (s, 1H), 2.91 (s, 2H), 1.96 – 0.97 (m, 8H, dicyclohexylurea).

¹³C NMR (101 MHz, chloroform-*d*) δ 171.86 (**C**O), 160.00 (Ar**C**), 157.00 (Ar**C**), 144.41 (Ar**C**), 144.29 (Ar**C**), 136.76 (Ar**C**), 133.23 (Ar**C**), 130.16 (Ar**C**), 130.05 (Ar**C**), 129.31 (Ar**C**), 128.45 (Ar**C**), 124.80 (Ar**C**), 120.41 (Ar**C**), 119.15 (Ar**C**), 114.10 (Ar**C**), 113.43 (Ar**C**), 113.29 (Ar**C**), 112.80 (Ar**C**), 55.57 (**C**H$_3$), 55.30 (**C**H$_3$), 53.52 (**C**H$_2$), 51.86 (**C**H$_2$), 34.13 (**C**H$_3$), 32.34 (**C**H$_3$).

Note: Rotamers formed, peak assignments were confirmed in final product through high temperature ¹H NMR.

N-(3-Methoxybenzyl)-3-(N-(2-methoxyphenyl)sulfamoyl)-N-methylbenzamide (74)
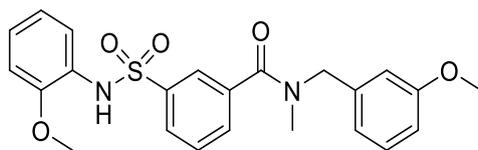
Compound **81** (431 mg, 800 μmol) was dissolved in CH_2Cl_2 (4 mL) with *o*-anisidine (110 μl, 960 μmol) and NEt$_3$ (120 μl, 1.60 mmol) and stirred for 16 h. The reaction was diluted four-fold with CH_2Cl_2 and washed with 1 M HCl, ½ sat. NaHCO$_3$ and brine. The solution was dried over MgSO$_4$, filtered, and concentrated under reduced pressure to afford 186 mg crude product. Following reverse-phase HPLC purification (30-70% B over 1 min, 70% B over 2 mins, 70-71% B over 2 mins, 71% B over 2 mins, 71-95% B over 0.5 mins, 95% B over 3 mins, 95-30% 0.5 mins, 30% B over 4 mins.) of 125 mg crude material, a white powder was obtained (Yield: 61.1 mg, 139 μmol, 17%).

HRMS (ESI) required for $C_{23}H_{23}N_2O_5S^-$: 439.1333, observed: 439.1335.

¹H NMR (400 MHz, dimethylsulfoxide-d_6) δ 7.83 – 7.47 (m, 4H, Ar**H**), 7.45 – 7.17 (m, 2H, Ar**H**), 7.17 – 7.03 (m, 1H, Ar**H**), 6.97 – 6.76 (m, 4H, Ar**H**), 6.76 – 6.44 (m, 1H, Ar**H**), 4.64 (s, 1H, C**H₂**), 4.31 (s, 1H, C**H₂**), 3.84 – 3.64 (m, 3H, C**H₃**), 3.58 – 3.27 (s, 3H, C**H₃**), 2.88 (s, 1H, C**H₃**), 2.72 (s, 2H, C**H₃**).

¹³C NMR (101 MHz, dimethylsulfoxide-d_6) δ 169.74 (**C**O), 160.0 (Ar**C**), 153.1 (Ar**C**), 141.2 (Ar**C**), 139.2 (Ar**C**), 137.1 (Ar**C**), 131.2 (Ar**C**), 130.9 (Ar**C**), 130.2 (Ar**C**), 129.6 (Ar**C**), 128.0 (Ar**C**), 127.4 (Ar**C**), 126.4 (Ar**C**), 125.5 (Ar**C**), 125.4 (Ar**C**), 120.9 (Ar**C**), 120.2 (Ar**C**), 119.4 (Ar**C**), 113.8 (Ar**C**), 113.1 (Ar**C**), 112.2 (Ar**C**), 55.7 (**C**H₃), 55.4 (**C**H₃), 54.2 (**C**H₂), 50.3 (**C**H₂), 37.2 (**C**H₃), 33.5 (**C**H₃).

Note: Rotamers observed, peak assignments were confirmed through high temperature ¹H NMR (**Figure 54**).

4-((4,6-Dimethylpyrimidin-2-yl)thio)aniline (83, modified procedure)[108]
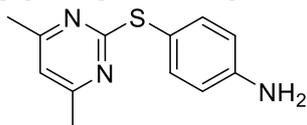
4,6-Dimethyl-2-(methylsulfonyl)pyrimidine (1.86 g, 10.0 mmol) was dissolved in DMF (25 mL) with 4-aminobenzenethiol (1.25 g, 10.0 mmol) and K_2CO_3 (2.76 g, 20.0 mmol) then stirred at 100 °C for 2 h. The contents were poured into 200 mL water then extracted three times with EtOAc. The organic layers were washed with water and brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The crude product was purified *via* flash column chromatography (33-50% EtOAc/hexane) to provide a yellow solid (Yield: 2.05 g, 8.86 mmol, 89%).

HRMS (ESI) required for $C_{12}H_{14}N_3S^+$: 232.0903, observed: 232.0913.

¹H NMR (400 MHz, dimethylsulfoxide-d_6) δ 7.26 – 7.03 (m, 2H, Ar**H**), 6.89 (s, 1H, Ar**H**), 6.65 – 6.47 (m, 2H, Ar**H**), 5.45 (s, 2H, N**H₂**), 2.25 (s, 6H, C**H₃**).

¹³C NMR (101 MHz, Dimethylsulfoxide-d_6) δ 171.7 (Ar**C**), 166.9 (Ar**C**), 149.9 (Ar**C**), 136.6 (Ar**C**), 116.0 (Ar**C**), 114.2 (Ar**C**), 112.9 (Ar**C**), 23.4 (**C**H₃).

Methyl 3-(thiophene-2-carboxamido)propanoate (85)
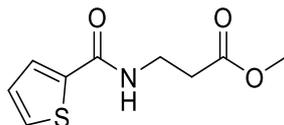
Thiophene-2-carboxylic acid (3.84 g, 30.0 mmol) was dissolved in CH_2Cl_2 (100 mL) with methyl 3-aminopropanoate (3.09 g, 30.0 mmol), DCC (9.28 g, 45.0 mmol) and DMAP (422 mg, 3.00 mmol) then stirred at room temperature for 18 h. The contents were filtered then concentrated under reduced pressure. The crude product was purified *via* flash column chromatography (20-50% EtOAc in hexane) to provide a white solid (Yield: 4.56 g, 21.4 mmol, 71%).

HRMS (ESI) required for $C_9H_{11}NO_3SNa^+$: 236.0352, observed: 236.0359.

¹H NMR (400 MHz, dimethylsulfoxide-d_6) δ 8.57 (t, *J* = 5.6 Hz, 1H, N**H**), 7.77 – 7.69 (m, 2H, Ar**H**), 7.13 (dd, *J* = 5.0, 3.7 Hz, 1H, Ar**H**), 3.60 (s, 3H, C**H₃**), 3.50 – 3.41 (m, 2H, C**H₂**), 2.58 (t, *J* = 7.0 Hz, 2H, NHCH₂C**H₂**).

¹³C NMR (101 MHz, Dimethylsulfoxide-d_6) δ 171.7 (**C**O), 161.2 (**C**O), 139.8 (Ar**C**), 130.7 (Ar**C**), 128.1 (Ar**C**), 127.9 (Ar**C**), 51.4 (**C**H₃), 35.3 (**C**H₂), 33.6 (**C**H₂).

3-(Thiophene-2-carboxamido)propanoic acid (86)
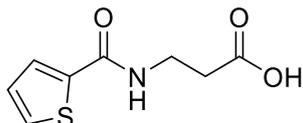
Compound **85** (1.89 g, 8.86 mmol) was dissolved in 1:1 DMF:water (10 mL) with LiOH (587 mg, 24.6 mmol) and stirred for 3 h. The contents were poured into 1 M NaOH and washed with EtOAc. The aqueous layer was acidified with 6 M HCl and extracted three times with EtOAc. The organic layers were combined, dried over MgSO$_4$, filtered, and concentrated under reduced pressure to afford a yellow oil (Yield: 1.56 g, 7.83 mmol, 88%).

HRMS (ESI) required for C$_8$H$_8$NO$_3$S$^-$: 198.0230, observed: 198.0234.

^{1}H NMR (400 MHz, dimethylsulfoxide-d_6) δ 12.24 (s, 1H, CO$_2$**H**), 8.54 (t, J = 5.6 Hz, 1H, N**H**), 7.84 – 7.63 (m, 2H, Ar**H**), 7.13 (dd, J = 4.9, 3.8 Hz, 1H, Ar**H**), 3.47 – 3.38 (m, 2H, C**H$_2$**), 2.52 – 2.48 (m, 2H, NHCH$_2$C**H$_2$**).

^{13}C NMR (101 MHz, dimethylsulfoxide-d_6) δ 173.3 (**C**O), 161.6 (**C**O), 140.4 (Ar**C**), 131.1 (Ar**C**), 128.5 (Ar**C**), 128.3 (Ar**C**), 35.9 (**C**H$_2$), 34.3 (**C**H$_2$).

N-(3-((4-((4,6-Dimethylpyrimidin-2-yl)thio)phenyl)amino)-3-oxopropyl)thiophene-2-carboxamide (77)
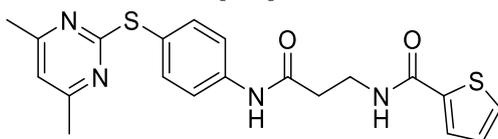
Compound **83** (910 mg, 3.93 mmol) was dissolved in CH_2Cl_2 (40 mL) with **86** (784 g, 3.93 mmol), DCC (1.22 g, 5.90 mmol) and DMAP (55.0 mg, 0.40 mmol) then stirred at room temperature for 16 h. The contents were filtered then washed with 33% aqueous NH_4Cl solution (w/w) and 1 M NaOH before being dried over $MgSO_4$ and concentrated under reduced pressure to provide 871 mg crude product. Following reverse-phase HPLC purification (30-54% B over 5 mins, 54-60% B over 12 mins, 60-95% B over 0.5 mins, 95% B over 2.5 mins, 95-30% B over 0.5 mins, 30% B over 1 min) of 177.1 mg crude material, a white powder was obtained (Yield: 130.0 mg, 0.32 mmol, 8%).

HRMS (ESI) required for $C_{20}H_{21}N_4O_2S_2^+$: 413.1100, observed: 413.1110.

^{1}H NMR (400 MHz, dimethylsulfoxide-d_6) δ 10.19 (s, 1H, N**H**), 8.66 (t, *J* = 5.6 Hz, 1H, N**H**), 7.77 – 7.71 (m, 2H, Ar**H**), 7.70 – 7.65 (m, 2H, Ar**H**), 7.50 – 7.46 (m, 2H, Ar**H**), 7.16 – 7.09 (m, 1H, Ar**H**), 6.96 (s, 1H, Ar**H**), 3.53 (m, 2H, C**H₂**), 2.65 (t, *J* = 6.9 Hz, 2H, C**H₂**), 2.27 (s, 6H, C**H₃**).

^{13}C NMR (101 MHz, dimethylsulfoxide-d_6) δ 170.3 (**C**O), 169.8 (**C**O), 167.2 (Ar**C**), 161.3 (Ar**C**), 140.0 (Ar**C**), 140.0 (Ar**C**), 135.7 (Ar**C**), 130.6 (Ar**C**), 128.0 (Ar**C**), 127.9 (Ar**C**), 122.6 (Ar**C**), 119.5 (Ar**C**), 116.5 (Ar**C**), 36.4 (**C**H₂), 35.7 (**C**H₂), 23.4 (**C**H₃).

Ethyl 3-((2,5-dimethylphenyl)amino)-3-oxopropanoate (88)
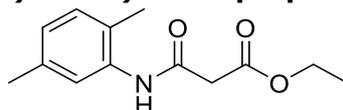
Monoethyl malonate (12 mL, 102 mmol) was dissolved in CH_2Cl_2 (100 mL) with DCC (31.5 g, 152 mmol), NEt_3 (40 mL, 305 mmol) and DMAP (138 mg, 1.00 mmol). The reaction was stirred for 10 mins prior to addition of 2,5-dimethylaniline (15 mL, 112 mmol) and then stirred for 16 hours. The reaction was filtered, the filtrate diluted in EtOAc and treated successively with 1M NaOH, 1M HCl and brine before being dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The product was purified by flash column chromatography (20-40% EtOAc/hex) to produce a white solid (Yield: 10.8 g, 45.7 mmol, 90%).

HRMS (ESI) required for $C_{13}H_{17}NO_3Na^+$: 258.1101, observed: 258.1104.

^{1}H NMR (400 MHz, dimethylsulfoxide-d_6) δ 9.47 (s, 1H, N**H**), 7.22 (s, 1H, Ar**H**), 7.09 (d, J = 7.5 Hz, 1H, Ar**H**), 6.91 (d, J = 7.5 Hz, 1H, Ar**H**), 4.13 (q, J = 7.1 Hz, 2H, C**H$_2$**), 3.46 (s, 2H, C**H$_2$**), 2.24 (s, 3H, C**H$_3$**), 2.15 (s, 3H, C**H$_3$**), 1.21 (t, J = 7.1 Hz, 3H, C**H$_3$**).

^{13}C NMR (101 MHz, dimethylsulfoxide-d_6) δ 168.4 (**C**O), 164.5 (**C**O), 136.2 (Ar**C**), 135.5 (Ar**C**), 130.6 (Ar**C**), 129.0 (Ar**C**), 126.5 (Ar**C**), 125.8 (Ar**C**), 61.1 (**C**H$_2$), 43.5 (**C**H$_2$), 21.0 (**C**H$_3$), 17.8 (**C**H$_3$), 14.5 (**C**H$_3$).

N-(2,5-Dimethylphenyl)malonamide (89)
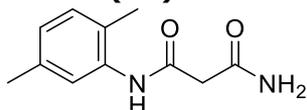
Compound **88** (2.48 g, 10.5 mmol) was dissolved in 7M NH_3 solution in methanol (15 mL, 105 mmol). The reaction was stirred for 16 hours and concentrated under reduced pressure to produce a white powder (Yield: 2.08 g, 10.1 mmol, 96%) which was taken forward without further purification.

HRMS (ESI) required for $C_{11}H_{14}N_2O_2Na^+$: 229.0947, observed: 229.0947.

^{1}H NMR (400 MHz, dimethylsulfoxide-d_6) δ 9.64 (s, 1H), 7.59 (s, 1H), 7.42 (s, 1H), 7.19 (s, 1H), 7.07 (d, J = 7.7 Hz, 1H), 6.87 (d, J = 7.7 Hz, 1H), 3.26 (s, 2H), 2.24 (s, 3H), 2.16 (s, 3H).

^{13}C NMR (101 MHz, dimethylsulfoxide-d_6) δ 170.0 (**C**O), 165.9 (**C**O), 136.5 (Ar**C**), 135.5 (Ar**C**), 130.5 (Ar**C**), 127.5 (Ar**C**), 125.8 (Ar**C**), 124.5 (Ar**C**), 43.7 (**C**H$_2$), 21.2 (**C**H$_3$), 17.7 (**C**H$_3$).

3-Amino-*N*-(2,5-dimethylphenyl)-3-thioxopropanamide (90, adapted procedure)[132]
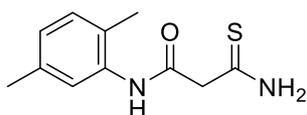
Compound **89** (6.30 g, 30.5 mmol) was suspended in 1,4-dioxane (50 mL) and heated to 60 °C. Lawesson's reagent (6.20 g, 15.3 mmol) was added and stirred for two hours to form a clear yellow solution which was cooled and concentrated under reduced pressure. The crude material was diluted with 0.5 M NaHCO$_3$ and extracted with EtOAc three times before being dried over MgSO$_4$, filtered, and concentrated under reduced pressure. The product was purified by flash column chromatography (35-100% EtOAc/hex) to produce a pale-yellow solid (3.25 g, 14.6 mmol, 48%).

HRMS (ESI) required for C$_{11}$H$_{13}$N$_2$OS$^-$: 221.0754, observed: 221.0750.

¹H NMR (400 MHz, dimethylsulfoxide-*d$_6$*) δ 9.67 (s, 1H), 9.45 (s, 1H), 9.41 (s, 1H), 7.26 (s, 1H), 7.09 (d, *J* = 7.8 Hz, 1H), 6.90 (d, *J* = 7.8 Hz, 1H), 3.70 (s, 2H), 2.25 (s, 3H), 2.18 (s, 3H).

¹³C NMR (101 MHz, dimethylsulfoxide-*d$_6$*) δ 200.6 (**C**S), 165.8 (**C**O), 136.3 (Ar**C**), 135.4 (Ar**C**), 130.6 (Ar**C**), 128.8 (Ar**C**), 126.4 (Ar**C**), 125.6 (Ar**C**), 52.5 (**C**H$_2$), 21.1 (**C**H$_3$), 17.9 (**C**H$_3$).

2-(4-(Chloromethyl)thiazol-2-yl)-*N*-(2,5-dimethylphenyl)acetamide (91)
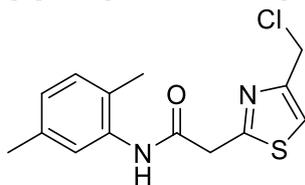
Compound **90** (3.25 g, 14.6 mmol) was suspended in DMF (20 mL) with 1,3-dichloroacetone (2.12 g, 16.8 mmol) and heated at 80 °C for 2 hours at which point it dissolved. EtOAc was added to the reaction and treated with aq. sat. sol. NaHCO$_3$ and brine before being dried over MgSO$_4$, filtered, and concentrated under reduced pressure. The product was purified by flash column chromatography (35-100% EtOAc/hex) to produce a white solid (2.35 g, 7.97 mmol, 54%).

HRMS (ESI) required for C$_{14}$H$_{16}$N$_2$OSCl$^+$: 295.0666, observed: 295.0665.

^{1}H NMR (400 MHz, dimethylsulfoxide-d_6) δ 9.69 (s, 1H), 7.68 (s, 1H), 7.24 (s, 1H), 7.09 (d, *J* = 7.8 Hz, 1H), 6.91 (d, *J* = 7.8 Hz, 1H), 4.81 (s, 2H), 4.17 (s, 2H), 2.24 (s, 3H), 2.16 (s, 3H).

^{13}C NMR (101 MHz, dimethylsulfoxide-d_6) δ 166.3 (**C**O), 164.5 (Ar**C**), 150.7 (Ar**C**), 135.7 (Ar**C**), 135.0 (Ar**C**), 130.7 (Ar**C**), 128.5 (Ar**C**), 126.1 (Ar**C**), 125.4 (Ar**C**), 120.0 (Ar**C**), 40.9 (**C**H$_2$), 40.2 (**C**H$_2$), 20.6 (**C**H$_3$), 17.4 (**C**H$_3$).

S-((2-(2-((2,5-Dimethylphenyl)amino)-2-oxoethyl)thiazol-4-yl)methyl) ethanethioate (92)
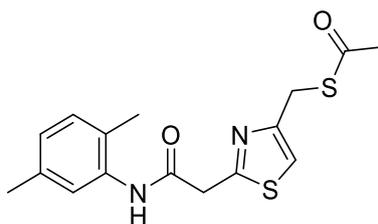
Compound **91** (2.35 g, 7.97 mmol) was dissolved in CH_2Cl_2 (50 mL) with thioacetic acid (1.00 mL, 12.1 mmol), NEt_3 (1.00 mL, 12.1 mmol) and DMAP (100 mg, 0.90 mmol) and stirred for 3 hours. The reaction was treated with aq. sat. sol. $NaHCO_3$ and water before being dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The product was purified by flash column chromatography (10-30% EtOAc/hex) to produce a white powder (2.25 g, 6.73 mmol, 84%).

HRMS (ESI) required for $C_{16}H_{19}N_2O_2S_2^+$: 335.0882, observed: 335.0885.

¹H NMR (400 MHz, chloroform-*d*) δ 9.30 (s, 1H, N**H**), 7.82 (s, 1H, Ar**H**), 7.17 (s, 1H, Ar**H**), 7.05 (d, *J* = 7.9 Hz, 1H, Ar**H**), 6.87 (d, *J* = 7.9 Hz, 1H, Ar**H**), 4.24 (s, 2H, C**H₂**), 4.09 (s, 2H, C**H₂**), 2.35 (s, 3H, C**H₃**), 2.32 (s, 3H, C**H₃**), 2.25 (s, 3H, C**H₃**), .

¹³C NMR (101 MHz, chloroform-*d*) δ 194.6 (**C**O), 164.7 (**C**O), 164.3 (Ar**C**), 152.4 (Ar**C**), 136.6 (Ar**C**), 135.8 (Ar**C**), 130.3 (Ar**C**), 125.8 (Ar**C**), 125.4 (Ar**C**), 122.8 (Ar**C**), 116.8 (Ar**C**), 40.9 (**C**H₂), 30.5 (**C**H₃), 29.0 (**C**H₂), 21.3 (**C**H₃), 17.8 (**C**H₃).

N-(2,5-Dimethylphenyl)-2-(4-(mercaptomethyl)thiazol-2-yl)acetamide (93)
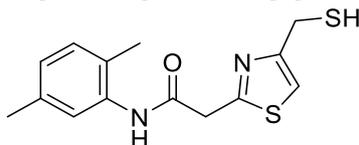
Compound **92** (2.25 g, 6.73 mmol) was dissolved in CH_2Cl_2 (50 mL) with hydrazine monohydrate (2.90 mL, 60.0 mmol) and stirred for 2.5 hours. The reaction was diluted with CH_2Cl_2 and treated with aq. sat. sol. $NaHCO_3$ and water before being dried over $MgSO_4$ and concentrated under reduced pressure. The product was purified by flash column chromatography (50-100% EtOAc/hex) to produce a pale-yellow oil (Yield: 1.54 g, 5.28 mmol, 78%).

HRMS (ESI) required for $C_{14}H_{17}N_2OS_2^+$: 293.0777, observed: 293.0779.

^{1}H NMR (400 MHz, dimethylsulfoxide-d_6) δ 9.67 (s, 1H, N**H**), 7.39 (s, 1H, Ar**H**), 7.24 (s, 1H, Ar**H**), 7.09 (d, J = 7.6 Hz, 1H, Ar**H**), 6.91 (d, J = 7.6 Hz, 1H, Ar**H**), 4.13 (s, 2H, COC**H₂**), 3.82 (d, J = 7.6 Hz, 2H, SC**H₂**), 2.81 (t, J = 7.6 Hz, 1H, S**H**), 2.24 (s, 3H, ArC**H₃**), 2.16 (s, 3H, ArC**H₃**).

^{13}C NMR (101 MHz, dimethylsulfoxide-d_6) δ 166.9 (**C**O), 164.2 (Ar**C**), 155.6 (Ar**C**), 137.6 (Ar**C**), 135.5 (Ar**C**), 130.6 (Ar**C**), 128.9 (Ar**C**), 127.2 (Ar**C**), 125.8 (Ar**C**), 116.4 (Ar**C**), 40.7 (**C**H₂), 23.1 (**C**H₂), 20.6 (**C**H₃), 18.1 (**C**H₃).

6-Chloro-*N,N*-dimethylpyridine-3-sulfonamide (95)
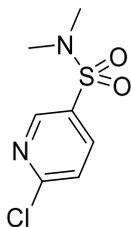
2-Chloro-5-pyridinesulfonyl chloride (1.50 g, 7.07 mmol), THF (7 mL), NEt$_3$ (1.12 mL, 8.49 mmol) and NHMe$_2$ (1.40 mL, 5.6 M in EtOH) were combined and stirred for 2 hours. The solution was concentrated under reduced pressure and split between CH$_2$Cl$_2$ and water. The organic layer was dried over MgSO$_4$, filtered, and concentration under reduced pressure, the product was purified by flash column chromatography (20–50% EtOAc/hexane) to produce a pale-yellow solid (Yield: 1.31 g, 6.18 mmol, 87%).

HRMS (ESI) required for C$_7$H$_{10}$ClN$_2$O$_2$S$^+$: 221.0146, observed: 221.0142.

^{1}H NMR (400 MHz, dimethylsulfoxide-*d$_6$*) δ 8.78 (dd, *J* = 2.6, 0.7 Hz, 1H, Ar**H**), 8.21 (dd, *J* = 8.4, 2.6 Hz, 1H, Ar**H**), 7.81 (dd, *J* = 8.4, 0.7 Hz, 1H, Ar**H**), 2.68 (s, 6H, C**H$_3$**).

^{13}C NMR (101 MHz, dimethylsulfoxide-*d$_6$*) δ 154.34 (Ar**C**), 148.46 (Ar**C**), 138.90 (Ar**C**), 131.06 (Ar**C**), 125.31 (Ar**C**), 37.35 (**C**H$_3$).

N-(2,5-Dimethylphenyl)-2-(4-(((4-(*N*,*N*-dimethylsulfamoyl)phenyl)thio)methyl)thiazol-2-yl)acetamide (78)
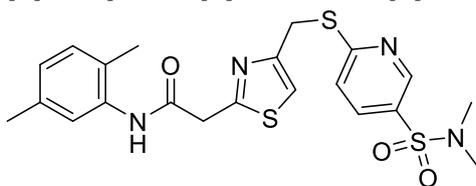
Compound **93** (127 mg, 434 µmol) was dissolved in DMF (5 mL) with compound **95** (96.0 mg, 434 µmol) and K_2CO_3 (180 mg, 1.30 mmol) and stirred for 16 hours, forming a dark brown solution. The reaction was diluted with EtOAc and treated with 1M NaOH, 1 M HCl and brine before being dried over $MgSO_4$, filtered, and concentrated under reduced pressure to afford 70 mg crude material. The product was purified by reverse-phase HPLC purification (50% B over 5 min, 50-60% B over 1 min, 60-80% B over 10 mins, 80-100% B over 1 min, 100% B over 8 mins) to provide a white powder (Yield: 15.8 mg, 33.2 µmol, 8%).

HRMS (ESI) required for $C_{21}H_{25}N_4O_3S_3^+$: 477.1083, observed: 477.1075.

¹H NMR (400 MHz, dimethylsulfoxide-d_6) δ 9.66 (s, 1H, N**H**), 8.76 (dd, J = 0.9, 2.4, 1H, Ar**H**), 7.95 (dd, J = 2.4, 8.5 Hz, 1H, Ar**H**), 7.61 (dd, J = 0.9, 8.5 Hz, 1H, Ar**H**), 7.54 (s, 1H, Ar**H**), 7.23 (s, 1H, Ar**H**), 7.09 (d, J = 7.7 Hz, 1H, Ar**H**), 6.91 (d, J = 7.7 Hz, 1H, Ar**H**), 4.60 (s, 2H, C**H₂**), 4.13 (s, 2H, C**H₂**), 2.65 (s, 6H, C**H₃**), 2.24 (s, 3H, C**H₃**), 2.16 (s, 3H, C**H₃**).

¹³C NMR (101 MHz, dimethylsulfoxide-d_6) δ 166.8 (**C**O), 164.5 (Ar**C**), 164.4 (Ar**C**), 150.8 (Ar**C**), 148.2 (Ar**C**), 136.2 (Ar**C**), 135.9 (Ar**C**), 135.5 (Ar**C**), 130.6 (Ar**C**), 128.9 (Ar**C**), 127.8 (Ar**C**), 126.6 (Ar**C**), 125.8 (Ar**C**), 122.2 (Ar**C**), 118.4 (Ar**C**), 39.7 (**C**H₂), 37.9 (**C**H₂), 29.8 (**C**H₃), 21.0 (**C**H₃), 17.9 (**C**H₃).

7.0: Bibliography

1   T. Wirth, F. Hildebrand, C. Allix-Béguec, F. Wölbeling, T. Kubica, K. Kremer, D. Van Soolingen, S. Rüsch-Gerdes, C. Locht, S. Brisse, A. Meyer, P. Supply and S. Niemann, *PLoS Pathog.*, 2008, **4**, 1–10.

2   I. Hershkovitz, H. D. Donoghue, D. E. Minnikin, H. May, O. Y. C. Lee, M. Feldman, E. Galili, M. Spigelman, B. M. Rothschild and G. K. Bar-Gal, *Tuberculosis*, 2015, **95**, S122–S126.

3   World Health Organisation (Global Tuberculosis Programme), *Global tuberculosis report 2020*, 2020.

4   World Health Organization, Weekly epidemiological update - 29 December 2020.

5   R. C. Huard, M. Fabre, P. De Haas, L. C. O. Lazzarini, D. Van Soolingen, D. Cousins and J. L. Ho, *J. Bacteriol.*, 2006, **188**, 4271–4287.

6   J. van Ingen, Z. Rahim, A. Mulder, M. J. Boeree, R. Simeone, R. Brosch and D. van Soolingen, *Emerg. Infect. Dis.*, 2012, **18**, 653–655.

7   F. Vayr, G. Martin-Blondel, F. Savall, J.-M. Soulat, G. Deffontaines and F. Herin, *PLoS Negl. Trop. Dis.*, 2018, **12**, e0006208.

8   D. Duncan, A. Brouwer, K. A. Harris, J. R. Lawes, R. Avigad, J. Dale and P. A. Upton, *Vet. Rec.*, 2020, **186**, 373–380.

9   DEFRA, *Next steps for the strategy for achieving bovine tuberculosis free status for England*, London, 2020.

10   J. A. Philips and J. D. Ernst, *Annu. Rev. Pathol. Mech. Dis.*, 2012, **7**, 353–384.

11   A. C. Whitelaw and W. A. Sturm, in *Tuberculosis*, ed. S. H. Schaaf, 2009, pp. 169–178.

12   W. R. Butler and L. S. Guthertz, *Clin. Microbiol. Rev.*, 2001, **14**, 704–726.

13   World Health Organization, *Diagnostic testing for SARS-CoV-2: Interim guidance 11 September 2020*, 2020.

14   D. Menzies, in *Tuberculosis*, ed. H. S. Schaaf, 2009, pp. 179–196.

15   S. Mostowy, A. G. Tsolaki, P. M. Small and M. A. Behr, *Vaccine*, 2003, **21**, 4270–

4274.

16   J. Li, A. Zhao, J. Tang, G. Wang, Y. Shi, L. Zhan and C. Qin, *Eur. J. Clin. Microbiol. Infect. Dis.*, 2020, **39**, 1405–1425.

17   I. Abubakar, L. Pimpin, C. Ariti, R. Beynon, P. Mangtani, J. Sterne, P. Fine, P. Smith, M. Lipman, D. Elliman, J. Watson, L. Drumright, P. Whiting and E. Vynnycky, *Systematic review and meta-analysis of the current evidence on the duration of protection by bacillus Calmette-Guérin vaccination against tuberculosis*, 2013, vol. 17.

18   P. Sharma, R. Mukherjee, G. P. Talwar, K. G. Sarathchandra, R. Walia, S. K. Parida, R. M. Pandey, R. Rani, H. Kar, A. Mukherjee, K. Katoch, S. K. Benara, Tulsi and P. Singh, *Lepr. Rev.*, 2005, **76**, 127–143.

19   D. Marinova, J. Gonzalo-Asensio, N. Aguilo and C. Martin, *Expert Rev. Vaccines*, 2017, **16**, 565–576.

20   A. Zumla, P. Nahid and S. T. Cole, *Nat. Rev. Drug Discov.*, 2013, **12**, 388–404.

21   M. S. Vasava, M. N. Bhoi, S. K. Rathwa, M. A. Borad, S. G. Nair and H. D. Patel, *Indian J. Tuberc.*, 2017, **64**, 252–275.

22   C. Vilchèze and W. R. Jacobs JR., *Microbiol. Spectr.*, 2014, **2**, MGM2-0014–2013.

23   K. Konno, F. M. Feldmann and W. McDermott, *Am. Rev. Respir. Dis.*, 1967, **95**, 461–469.

24   Y. Zhang, A. Scorpio, H. Nikaido and Z. Sun, *J. Bacteriol.*, 1999, **181**, 2044–2049.

25   W. Shi, X. Zhang, X. Jiang, H. Yuan, J. S. Lee, C. E. Barry, H. Wang, W. Zhang and Y. Zhang, *Science (80-.).*, 2011, **333**, 1630–1632.

26   L. A. Kayukova and E. A. Berikova, *Pharm. Chem. J.*, 2020, **54**, 555–563.

27   S. Chetty, M. Ramesh, A. Singh-Pillay and M. E. S. Soliman, *Bioorganic Med. Chem. Lett.*, 2017, **27**, 370–386.

28   M. A. Espinal, S. J. Kim, P. G. Suarez, K. M. Kam, A. G. Khomenko, G. B. Migliori, J. Baéz, A. Kochi, C. Dye and M. C. Raviglione, *JAMA*, 2000, **283**, 2537–2545.

29   World Health Organization, *WHO consolidated guidelines on drug-resistant tuberculosis treatment*, 2019.

30   A. Di Paolo, P. Malacarne, E. Guidotti, R. Danesi and M. Del Tacca, *Clin. Pharmacokinet.*, 2010, **49**, 439–447.

31   T. Yano, S. Kassovska-Bratinova, J. Shin Teh, J. Winkler, K. Sullivan, A. Isaacs, N. M. Schechter and H. Rubin, *J. Biol. Chem.*, 2011, **286**, 10276–10287.

32   Y. Shimokawa, K. Sasahara, N. Koyama, K. Kitano, M. Shibata, N. Yoda and K. Umehara, *Drug Metab. Dispos.*, 2015, **43**, 1277–1283.

33   Y. Liu, M. Matsumoto, H. Ishida, K. Ohguro, M. Yoshitake, R. Gupta, L. Geiter and J. Hafkin, *Tuberculosis*, 2018, **111**, 20–30.

34   G. a Prosser and L. P. S. de Carvalho, *FEBS J.*, 2013, **280**, 1150–66.

35   B. Badet and C. Walsh, *Biochemistry*, 1985, **24**, 1333–1341.

36   R. R. Rando, *Biochem. Pharmacol.*, 1975, **24**, 1153–1160.

37   J. S. Kass and W. X. Shandera, *CNS Drugs*, 2010, 24, 655–667.

38   S. R. Levine and K. E. Beatty, *ACS Infect. Dis.*, 2021, **7**, 461–470.

39   S. Thee, A. J. Garcia-Prats, P. R. Donald, A. C. Hesseling and H. S. Schaaf, *Tuberculosis*, 2016, **97**, 126–136.

40   J. Zheng, E. J. Rubin, P. Bifani, V. Mathys, V. Lim, M. Au, J. Jang, J. Nam, T. Dick, J. R. Walker, K. Pethe and L. R. Camacho, *J. Biol. Chem.*, 2013, **288**, 23447–23456.

41   WHO, *WHO Tuberculosis Programme: framework for effective tuberculosis control*, 1994.

42   J. Karumbi and P. Garner, *Directly observed therapy for treating tuberculosis*, 2015.

43   L. J. Alderwick, J. Harrison, G. S. Lloyd and H. L. Birch, *Cold Spring Harb. Perspect. Med.*, 2015, **5**, 1–16.

44   L. J. Alderwick, E. Radmacher, M. Seidel, R. Gande, P. G. Hitchen, H. R. Morris, A. Dell, H. Sahm, L. Eggeling and G. S. Besra, *J. Biol. Chem.*, 2005, **280**, 32362–32371.

45   D. Kaur, M. E. Guerin, H. Kovierov, P. J. Brennan and M. Jackson, *Adv. Appl. Microbiol.*, 2009, **69**, 23–78.

46   C. Ghazaei, *J. Res. Med. Sci.*, 2018, **23**, 1–10.

47   S. Srivastava, S. P. Van Rijn, A. M. A. Wessels, J. W. C. Alffenaar and T. Gumbo,

Antimicrob. Agents Chemother., 2016, **60**, 3193–3195.

48   G. Gago, L. Diacovich, A. Arabolaza, S.-C. Tsai and H. Gramajo, *FEMS Microbiol. Rev.*, 2011, **35**, 475–497.

49   R. P. Massengo-Tiassé and J. E. Cronan, *Cell. Mol. Life Sci.*, 2009, 66, 1507–1517.

50   Y. Wu and A. Zhou, *Chem. Commun.*, 2009, **45**, 7021–7023.

51   K. A. Abrahams and G. S. Besra, *Parasitology*, 2018, **145**, 116–133.

52   S. T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead and B. G. Barrell, *Nature*, 1998, **393**, 537–544.

53   K. H. Choi, L. Kremer, G. S. Besra and C. O. Rock, *J. Biol. Chem.*, 2000, **275**, 28201–28207.

54   H. Marrakchi, S. Ducasse, G. Labesse, H. Montrozier, E. Margeat, L. Emorine, X. Charpentier, M. Daffé and A. Quémard, *Microbiology*, 2002, **148**, 951–960.

55   E. Sacco, A. S. Covarrubias, H. M. O'Hare, P. Carroll, N. Eynard, T. A. Jones, T. Parish, M. Daffé, K. Bäckbro and A. Quémard, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 14628–14633.

56   A. Banerjee, E. Dubnau, A. Quemard, V. Balasubramanian, K. S. Um, T. Wilson, D. Collins, G. De Lisle and W. R. Jacobs, *Science (80-.).*, 1994, **263**, 227–230.

57   R. A. Slayden and C. E. Barry, *Tuberculosis*, 2002, **82**, 149–160.

58   D. Barkan, V. Rao, G. D. Sukenick and M. S. Glickman, *J. Bacteriol.*, 2010, **192**, 3661–3668.

59   O. A. Trivedi, P. Arora, V. Sridharan, R. Tickoo, D. Mohanty and R. S. Gokhale, *Nature*, 2004, **428**, 441–445.

60   D. Portevin, C. De Sousa-D'Auria, C. Houssin, C. Grimaldi, M. Chami, M. Daffé and C. Guilhot, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 314–319.

61 A. Bhatt, A. K. Brown, A. Singh, D. E. Minnikin and G. S. Besra, *Chem. Biol.*, 2008, **15**, 930–939.

62 X. He, A. Alian, R. Stroud and P. R. Ortiz De Montellano, *J. Med. Chem.*, 2006, **49**, 6308–6323.

63 *Schrodinger, LLC. 2010. The PyMol Molecular Graphics System, Version 1.8*, .

64 A. F. Bell, C. F. Stratton, X. Zhang, P. Novichenok, A. A. Jaye, P. A. Nair, S. Parikh, R. Rawat and P. J. Tonge, *J. Am. Chem. Soc.*, 2007, **129**, 6425–6431.

65 A. Dessen, A. Quémard, J. S. Blanchard, W. R. Jacobs and J. C. Sacchettini, *Science (80-.).*, 1995, **267**, 1638–1641.

66 D. A. Rozwarski, C. Vilchèze, M. Sugantino, R. Bittman and J. C. Sacchettini, *J. Biol. Chem.*, 1999, **274**, 15582–15589.

67 H. L. Bernstein, J.; Lott, W. A.; Steinberg, B. A.; Yale, *Am. Rev. Tuberc. Pulm. Dis.*, 1952, **65**, 357–364.

68 J. Crofton, *Br. Med. J.*, 1959, **1**, 1610–1614.

69 L. B. Heifets, P. J. Lindholm-Levy and M. Flory, *Am. Rev. Respir. Dis.*, 1991, **143**, 268–70.

70 G. Simon, *Br. Med. J.*, 1952, **2**, 735–746.

71 A. Ghodousi, E. Tagliani, E. Karunaratne, S. Niemann, J. Perera, C. U. Köser and D. M. Cirillo, *Antimicrob. Agents Chemother.*, 2019, **63**, e00092-19.

72 E. J. North, M. Jackson and R. E. Lee, *Curr. Pharm. Des.*, 2014, **20**, 4357–78.

73 L. Jena, P. Waghmare, S. Kashikar, S. Kumar and B. C. Harinath, *Int. J. Mycobacteriology*, 2014, **3**, 276–282.

74 M. Seifert, D. Catanzaro, A. Catanzaro and T. C. Rodwell, *PLoS One*, 2015, **10**, e0119628.

75 K. Maeda, H. Kosaka, Y. Okami and H. Umezawa, *J. Antibiot. (Tokyo).*, 1953, **6**, 140.

76 R. C. Hartkoorn, C. Sala, J. Neres, F. Pojer, S. Magnet, R. Mukherjee, S. Uplekar, S. Boy-Röttger, K. H. Altmann and S. T. Cole, *EMBO Mol. Med.*, 2012, **4**, 1032–1042.

77 J. Stec, C. Vilchèze, S. Lun, A. L. Perryman, X. Wang, J. S. Freundlich, W. Bishai,

W. R. Jacobs and A. P. Kozikowski, *ChemMedChem*, 2014, **9**, 2528–2537.

78   S. L. Parikh, G. Xiao and P. J. Tonge, *Biochemistry*, 2000, **39**, 7645–7650.

79   U. H. Manjunatha, S. P. S. Rao, R. R. Kondreddi, C. G. Noble, L. R. Camacho, B. H. Tan, S. H. Ng, P. S. Ng, N. L. Ma, S. B. Lakshminarayana, M. Herve, S. W. Barnes, W. Yu, K. Kuhen, F. Blasco, D. Beer, J. R. Walker, P. J. Tonge, R. Glynne, P. W. Smith and T. T. Diagana, *Sci. Transl. Med.*, 2015, **7**, 269ra3.

80   P. S. Shirude, P. Madhavapeddi, M. Naik, K. Murugan, V. Shinde, R. Nandishaiah, J. Bhat, A. Kumar, S. Hameed, G. Holdgate, G. Davies, H. McMiken, N. Hegde, A. Ambady, J. Venkatraman, M. Panda, B. Bandodkar, V. K. Sambandamurthy and J. A. Read, *J. Med. Chem.*, 2013, **56**, 8533–8542.

81   M. R. Kuo, H. R. Morbidoni, D. Alland, S. F. Sneddon, B. B. Gourlie, M. M. Staveski, M. Leonard, J. S. Gregory, A. D. Janjigian, C. Yee, J. M. Musser, B. Kreiswirth, H. Iwamoto, R. Perozzo, W. R. Jacobs, J. C. Sacchettini and D. A. Fidock, *J. Biol. Chem.*, 2003, **278**, 20851–20859.

82   M. D. Wall, M. Oshin, G. A. C. Chung, T. Parkhouse, A. Gore, E. Herreros, B. Cox, K. Duncan, B. Evans, M. Everett and A. Mendoza, *Bioorganic Med. Chem. Lett.*, 2007, **17**, 2740–2744.

83   L. Encinas, H. O'Keefe, M. Neu, M. J. Remuiñán, A. M. Patel, A. Guardia, C. P. Davie, N. Pérez-Macías, H. Yang, M. A. Convery, J. A. Messer, E. Pérez-Herrán, P. A. Centrella, D. Álvarez-Gómez, M. A. Clark, S. Huss, G. K. O'Donovan, F. Ortega-Muro, W. McDowell, P. Castañeda, C. C. Arico-Muendel, S. Pajk, J. Rullás, I. Angulo-Barturen, E. Álvarez-Ruíz, A. Mendoza-Losana, L. Ballell Pages, J. Castro-Pichel and G. Evindar, *J. Med. Chem.*, 2014, **57**, 1276–1288.

84   A. Bateman, M. J. Martin, C. O'Donovan, M. Magrane, E. Alpi, R. Antunes, B. Bely, M. Bingley, C. Bonilla, R. Britto, B. Bursteinas, H. Bye-AJee, A. Cowley, A. Da Silva, M. De Giorgi, T. Dogan, F. Fazzini, L. G. Castro, L. Figueira, P. Garmiri, G. Georghiou, D. Gonzalez, E. Hatton-Ellis, W. Li, W. Liu, R. Lopez, J. Luo, Y. Lussi, A. MacDougall, A. Nightingale, B. Palka, K. Pichler, D. Poggioli, S. Pundir, L. Pureza, G. Qi, S. Rosanoff, R. Saidi, T. Sawford, A. Shypitsyna, E. Speretta, E. Turner, N. Tyagi, V.

Volynkin, T. Wardell, K. Warner, X. Watkins, R. Zaru, H. Zellner, I. Xenarios, L. Bougueleret, A. Bridge, S. Poux, N. Redaschi, L. Aimo, G. ArgoudPuy, A. Auchincloss, K. Axelsen, P. Bansal, D. Baratin, M. C. Blatter, B. Boeckmann, J. Bolleman, E. Boutet, L. Breuza, C. Casal-Casas, E. De Castro, E. Coudert, B. Cuche, M. Doche, D. Dornevil, S. Duvaud, A. Estreicher, L. Famiglietti, M. Feuermann, E. Gasteiger, S. Gehant, V. Gerritsen, A. Gos, N. Gruaz-Gumowski, U. Hinz, C. Hulo, F. Jungo, G. Keller, V. Lara, P. Lemercier, D. Lieberherr, T. Lombardot, X. Martin, P. Masson, A. Morgat, T. Neto, N. Nouspikel, S. Paesano, I. Pedruzzi, S. Pilbout, M. Pozzato, M. Pruess, C. Rivoire, B. Roechert, M. Schneider, C. Sigrist, K. Sonesson, S. Staehli, A. Stutz, S. Sundaram, M. Tognolli, L. Verbregue, A. L. Veuthey, C. H. Wu, C. N. Arighi, L. Arminski, C. Chen, Y. Chen, J. S. Garavelli, H. Huang, K. Laiho, P. McGarvey, D. A. Natale, K. Ross, C. R. Vinayaka, Q. Wang, Y. Wang, L. S. Yeh and J. Zhang, *Nucleic Acids Res.*, 2017, **45**, D158–D169.

85   H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.

86   T. Kinjo, Y. Koseki, M. Kobayashi, A. Yamada, K. Morita, K. Yamaguchi, R. Tsurusawa, G. Gulten, H. Komatsu, H. Sakamoto, J. C. Sacchettini, M. Kitamura and S. Aoki, *J. Chem. Inf. Model.*, 2013, **53**, 1200–1212.

87   H. H. Soutter, P. Centrella, M. A. Clark, J. W. Cuozzo, C. E. Dumelin, M. A. Guie, S. Habeshian, A. D. Keefe, K. M. Kennedy, E. A. Sigel, D. M. Troast, Y. Zhang, A. D. Ferguson, G. Davies, E. R. Stead, J. Breed, P. Madhavapeddi and J. A. Read, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, E7880–E7889.

88   P. Pan, S. E. Knudson, G. R. Bommineni, H. J. Li, C. T. Lai, N. Liu, M. Garcia-Diaz, C. Simmerling, S. S. Patil, R. A. Slayden and P. J. Tonge, *ChemMedChem*, 2014, **9**, 776–791.

89   E. K. Schroeder, L. A. Basso, D. S. Santos and O. N. De Souza, *Biophys. J.*, 2005, **89**, 876–884.

90   J. S. Oliveira, J. H. Pereira, F. Canduri, N. C. Rodrigues, O. N. de Souza, W. F. de Azevedo, L. A. Basso and D. S. Santos, *J. Mol. Biol.*, 2006, **359**, 646–666.

91   A. Chollet, L. Mourey, C. Lherbet, A. Delbot, S. Julien, M. Baltas, J. Bernadou, G. Pratviel, L. Maveyraud and V. Bernardes-Génisson, *J. Struct. Biol.*, 2015, **190**, 328–337.

92   *Schrödinger Release 2018-1: Maestro, Schrödinger, LLC, New York, NY, 2018*, .

93   G. Jones, P. Willett, R. C. Glen, A. R. Leach and R. Taylor, *J. Mol. Biol.*, 1997, **267**, 727–48.

94   J. W. M. Nissink, C. Murray, M. Hartshorn, M. L. Verdonk, J. C. Cole and R. Taylor, *Proteins Struct. Funct. Genet.*, 2002, **49**, 457–471.

95   I. Pauli, R. N. Dos Santos, D. C. Rostirolla, L. K. Martinelli, R. G. Ducati, L. F. S. M. Timmers, L. A. Basso, D. S. Santos, R. V. C. Guido, A. D. Andricopulo and O. Norberto De Souza, *J. Chem. Inf. Model.*, 2013, **53**, 2390–2401.

96   *Guide, GOLD User Suite, A Component of the CSD-Discovery 2, 2019 CSD Release Update*, 2019.

97   R. Brenk, A. Schipani, D. James, A. Krasowski, I. H. Gilbert, J. Frearson and P. G. Wyatt, *ChemMedChem*, 2008, **3**, 435–444.

98   F. Soukarieh, R. Liu, M. Romero, S. N. Roberston, W. Richardson, S. Lucanto, E. V. Oton, N. R. Qudus, A. Mashabi, S. Grossman, S. Ali, T. Sou, I. Kukavica-Ibrulj, R. C. Levesque, C. A. S. Bergström, N. Halliday, S. N. Mistry, J. Emsley, S. Heeb, P. Williams, M. Cámara and M. J. Stocks, *Front. Chem.*, 2020, **8**, 204.

99   M. Maryati, I. Kaur, G. P. Jadhav, L. Olotu-Umoren, B. Oveh, L. Hashmi, P. M. Fischer and G. S. Winkler, *Nucleic Acids Res.*, 2014, **42**, e30.

100   R. P. Tangallapally, R. E. B. Lee, A. J. M. Lenaerts and R. E. Lee, *Bioorganic Med. Chem. Lett.*, 2006, **16**, 2584–2589.

101   *ChemAxon, JChem for excel, 2013*, .

102   H. Veisi, R. Ghorbani-Vaghei, S. Hemmati and J. Mahmoodi, *Synlett*, 2011, **16**, 2315–2320.

103   R. P. Tangallapally, R. Yendapally, R. E. Lee, K. Hevener, V. C. Jones, A. J. M. Lenaerts, M. R. McNeil, Y. Wang, S. Franzblau and R. E. Lee, *J. Med. Chem.*, 2004, **47**, 5276–5283.

104   A. Freitag, P. Prajwal, A. Shymanets, C. Harteneck, B. Nürnberg, C. Schächtele, M. Kubbutat, F. Totzke and S. A. Laufer, *J. Med. Chem.*, 2015, **58**, 212–221.

105   P. S. Ng, U. H. Manjunatha, S. P. S. Rao, L. R. Camacho, N. L. Ma, M. Herve, C. G. Noble, A. Goh, S. Peukert, T. T. Diagana, P. W. Smith and R. R. Kondreddi, *Eur. J. Med. Chem.*, 2015, **106**, 144–156.

106   US 7,396,936 B1, 2008, 21.

107   J. Di Huang, X. P. Hu, Z. C. Duan, Q. H. Zeng, S. B. Yu, J. Deng, D. Y. Wang and Z. Zheng, *Org. Lett.*, 2006, **8**, 4367–4370.

108   C. Jin, Y. J. Liang, H. He and L. Fu, *Eur. J. Med. Chem.*, 2011, **46**, 429–432.

109   I. V. Dyachenko and M. V. Vovk, *Russ. J. Org. Chem.*, 2013, **49**, 259–267.

110   C. W. am Ende, S. E. Knudson, N. Liu, J. Childs, T. J. Sullivan, M. Boyne, H. Xu, Y. Gegina, D. L. Knudson, F. Johnson, C. A. Peloquin, R. A. Slayden and P. J. Tonge, *Bioorganic Med. Chem. Lett.*, 2008, **18**, 3029–3033.

111   S. Pajk, M. Živec, R. Šink, I. Sosič, M. Neu, C. W. Chung, M. Martínez-Hoyos, E. Pérez-Herrán, D. Álvarez-Gómez, E. Álvarez-Ruíz, A. Mendoza-Losana, J. Castro-Pichel, D. Barros, L. Ballell-Pages, R. J. Young, M. A. Convery, L. Encinas and S. Gobec, *Eur. J. Med. Chem.*, 2016, **112**, 252–257.

112   A. L. Perryman, W. Yu, X. Wang, S. Ekins, S. Forli, S. G. Li, J. S. Freundlich, P. J. Tonge and A. J. Olson, *J. Chem. Inf. Model.*, 2015, **55**, 645–659.

113   D. Eisenberg, E. Schwarz, M. Komaromy and R. Wall, *J. Mol. Biol.*, 1984, **179**, 125–142.

114   L. Ioakimidis, L. Thoukydidis, A. Mirza, S. Naeem and J. Reynisson, *QSAR Comb. Sci.*, 2008, **27**, 445–456.

115   J. Ramprasad, N. Nayak, U. Dalimba, P. Yogeeswari and D. Sriram, *Medchemcomm*, 2016, **7**, 338–344.

116   D. Baldisseri, *Practical Aspects of Fragment-Based Screening Experiments in TopSpin*, Bruker, 2018.

117   F. Prati, F. Zuccotto, D. Fletcher, M. A. Convery, R. Fernandez-Menendez, R. Bates, L. Encinas, J. Zeng, C. W. Chung, P. De Dios Anton, A. Mendoza-Losana, C.

Mackenzie, S. R. Green, M. Huggett, D. Barros, P. G. Wyatt and P. C. Ray, *ChemMedChem*, 2018, **13**, 672–677.

118 M. Sabbah, V. Mendes, R. G. Vistal, D. M. G. Dias, M. Záhorszká, K. Mikušová, J. Korduláková, A. G. Coyne, T. L. Blundell and C. Abell, *J. Med. Chem.*, 2020, **63**, 4749–4761.

119 A. Ahmed, *Evaluation and identification of novel inhibitors of mycobacterium tuberculosis inha, University of Nottingham*, 2016.

120 A. Quémard, J. C. Sacchettini, A. Dessen, C. Vilcheze, R. Bittman, W. R. Jacobs and J. S. Blanchard, *Biochemistry*, 1995, **34**, 8234–8241.

121 T. Armstrong, M. Lamont, A. Lanne, L. J. Alderwick and N. R. Thomas, *Bioorg. Med. Chem.*, 2020, **28**, 115744.

122 C. F. Lu, C. Xie, Z. X. Chen and G. C. Yang, *Russ. J. Org. Chem.*, 2009, **45**, 788-791.

123 S. Dei, M. Coronnello, G. Bartolucci, D. Manetti, M. N. Romanelli, C. Udomtanakunchai, M. Salerno and E. Teodori, *Eur. J. Med. Chem.*, 2018, **147**, 7-20.

124 US 5,747,513 A, 1998, 16.

125 L. Hoffer, Y. V. Voitovich, B. Raux, K. Carrasco, C. Muller, A. Y. Fedorov, C. Derviaux, A. Amouric, S. Betzi, D. Horvath, A. Varnek, Y. Collette, S. Combes, P. Roche and X. Morelli, *J. Med. Chem.*, 2018, **61**, 5719-5732.

126 W. Zeinyeh, Z. Mahiout, S. Radix, T. Lomberget, A. Dumoulin, R. Barret, C. Grenot, L. Rocheblave, E. L. Matera, C. Dumontet and N. Walchshofer, *Steroids*, 2012, **12**, 1177-1191.

127 C. Yu and J. W. Taylor, *Bioorganic Med. Chem.*, 1999, **1**, 161-175.

128 Y. G. Li, J. X. Wang, G. N. Zhang, M. Zhu, X. F. You, Y. C. Wang and F. Zhang, *Bioorganic Med. Chem. Lett.*, 2020, **7**, 126969.

129 J. Nyhlén, L. Eriksson and J. E. Bäckvall, *Chirality*, 2008, **20**, 47–50.

130 M. Einsiedler, C. S. Jamieson, M. A. Maskeri, K. N. Houk and T. A. M. Gulder, *Angew. Chemie - Int. Ed.*, 2021, **15**, 8297-8302.

131 E. Lager, J. Nilsson, E. Østergaard Nielsen, M. Nielsen, T. Liljefors and O. Sterner, *Bioorganic Med. Chem.*, 2008, **16**, 6936–6948.

132 WO 2009/006,404 A2, 2009.

133 K. F. Geoghegan, H. B. F. Dixon, P. J. Rosner, L. R. Hoth, A. J. Lanzetti, K. A. Borzilleri, E. S. Marr, L. H. Pezzullo, L. B. Martin, P. K. Lemotte, A. S. McColl, A. V. Kamath and J. G. Stroh, *Anal. Biochem.*, 1999, **267**, 169–184.

134 S. Khan, S. N. Nagarajan, A. Parikh, S. Samantaray, A. Singh, D. Kumar, R. P. Roy, A. Bhatt and V. K. Nandicoori, *J. Biol. Chem.*, 2010, **285**, 37860–37871.

8.0: Supporting information

8.1 InhA Purification

5' CATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGACAGGACTGCTGGACGGCAAACGGATTCTGGTTAGC
3' GTACCCGTCGTCGGTAGTAGTAGTAGTAGTGTCGTCGCCGGACCACGGCGCGCCGTCGGTATACTGTCCTGACGACCTGCCGTTTGCCTAAGACCAATCG
100

GGAATCATCACCGACTCGTCGATCGCGTTTCACATCGCACGGGTAGCCCAGGAGCAGGGCGCCCAGCTGGTGCTCACCGGGTTCGACCGGCTGCGGCTGA
CCTTAGTAGTGGCTGAGCAGCTAGCGCAAAGTGTAGCGTGCCCATCGGGTCCTCGTCCCGCGGGTCGACCACGAGTGGCCCAAGCTGGCCGACGCCGACT
200

TTCAGCGCATCACCGACCGGCTGCCGGCAAAGGCCCCGCTGCTCGAACTCGACGTGCAAAACGAGGAGCACCTGGCCAGCTTGGCCGGCCGGGTGACCGA
AAGTCGCGTAGTGGCTGGCCGACGGCCGTTTCCGGGGCGACGAGCTTGAGCTGCACGTTTTGCTCCTCGTGGACCGGTCGAACCGGCCGGCCCACTGGCT
300

GGCGATCGGGGCGGGCAACAAGCTCGACGGGGTGGTGCATTCGATTGGGTTCATGCCGCAGACCGGGATGGGCATCAACCCGTTCTTCGACGCGCCCTAC
CCGCTAGCCCCGCCCGTTGTTCGAGCTGCCCCACCACGTAAGCTAACCCAAGTACGGCGTCTGGCCCTACCCGTAGTTGGGCAAGAAGCTGCGCGGGATG
400

GCGGATGTGTCCAAGGGCATCCACATCTCGGCGTATTCGTATGCTTCGATGGCCAAGGCGCTGCTGCCGATCATGAACCCCGGAGGTTCCATCGTCGGCA
CGCCTACACAGGTTCCCGTAGGTGTAGAGCCGCATAAGCATACGAAGCTACCGGTTCCGCGACGACGGCTAGTACTTGGGGCCTCCAAGGTAGCAGCCGT
500

TGGACTTCGACCCGAGCCGGGCGATGCCGGCCTACAACTGGATGACGGTCGCCAAGAGCGCGTTGGAGTCGGTCAACAGGTTCGTGGCGCGCGAGGCCGG
ACCTGAAGCTGGGCTCGGCCCGCTACGGCCGGATGTTGACCTACTGCCAGCGGTTCTCGCGCAACCTCAGCCAGTTGTCCAAGCACCGCGCGCTCCGGCC
600

CAAGTACGGTGTGCGTTCGAATCTCGTTGCCGCAGGCCCTATCCGGACGCTGGCGATGAGTGCGATCGTCGGCGGTGCGCTCGGCGAGGAGGCCGGCGCC
GTTCATGCCACACGCAAGCTTAGAGCAACGGCGTCCGGGATAGGCCTGCGACCGCTACTCACGCTAGCAGCCGCCACGCGAGCCGCTCCTCCGGCCGCGG
700

CAGATCCAGCTGCTCGAGGAGGGCTGGGATCAGCGCGCTCCGATCGGCTGGAACATGAAGGATGCGACGCCGGTCGCCAAGACGGTGTGCGCGCTGCTGT
GTCTAGGTCGACGAGCTCCTCCCGACCCTAGTCGCGCGAGGCTAGCCGACCTTGTACTTCCTACGCTGCGGCCAGCGGTTCTGCCACACGCGCGACGACA
800

CTGACTGGCTGCCGGCGACCACGGGTGACATCATCTACGCCGACGGCGGCGCGCACACCCAATTGCTTTAG   3'
GACTGACCGACGGCCGCTGGTGCCCACTGTAGTAGATGCGGCTGCCGCCGCGCGTGTGGGTTAACGAAATC   5'
871

Figure 35: Nucleotide sequence for plasmid InhA insert kindly donated to the group by Peter Tonge (State university of New York at Stoney Brook).
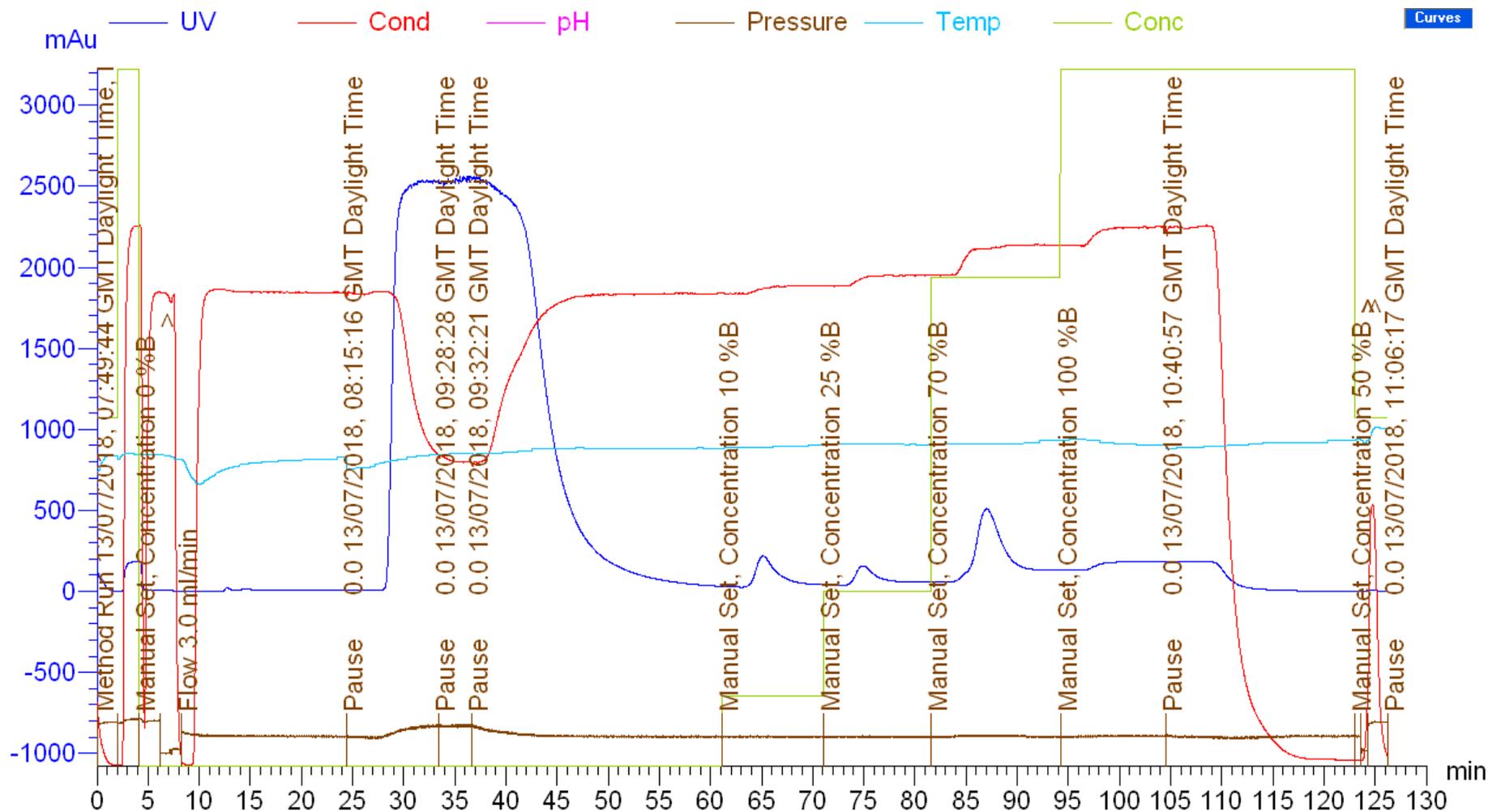
Figure 36: Typical FPLC chromatogram for InhA purification. Peaks: 3-5 min (50% buffer 2), 28-55 min (flowthrough), 63-70 min (10 % Buffer 2), 70-82 min (25 % Buffer 2), 85-95 min (InhA, 70 % Buffer 2).
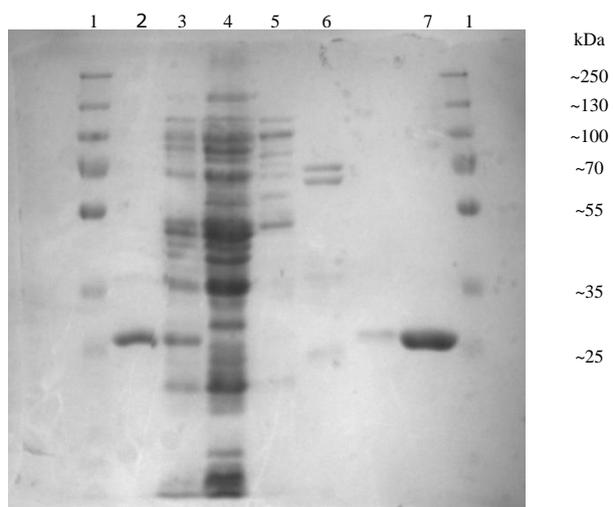
Figure 37: SDS-PAGE gel of FPLC purification. Numbering: 1 PageRuler Plus prestained protein ladder, 2-Insoluble fraction, 3-soluble fraction, 4-FPLC flowthrough, 5-FPLC peak at 10% imidazole, 6-FPLC peak at 25% imidazole, 7-FPLC peak at 70% imidazole containing InhA.
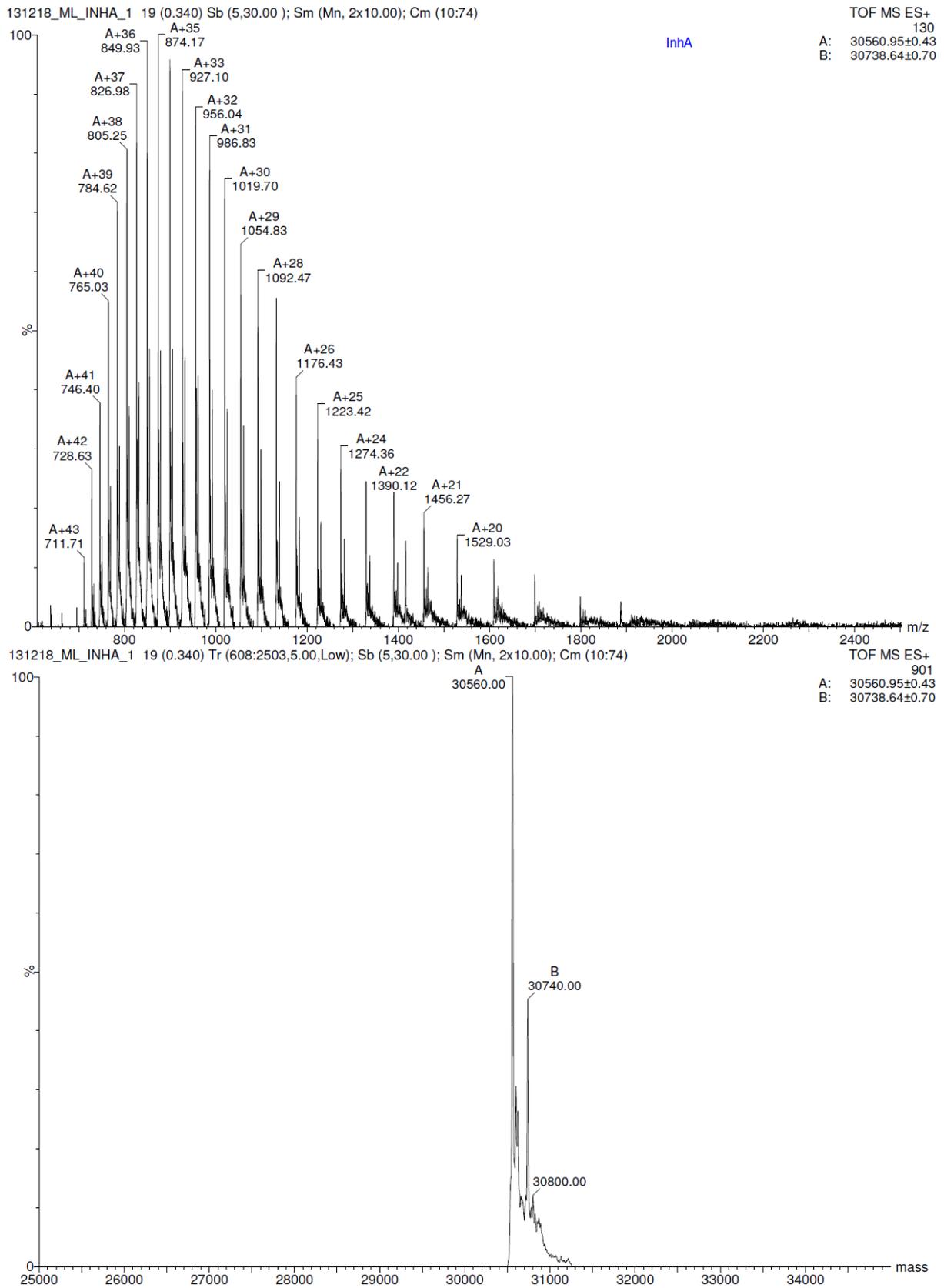
Figure 38: Protein Mass spectrometry result for InhA (target mass -N-terminal Met: 30,559.97 Da). Top: raw data, bottom: deconvolution. Secondary peak at 30,740.00 Da consisted with partial alpha-N-6-phosphogluconylation of his-tag.[133]

8.2 Inhibition studies and enzyme kinetics
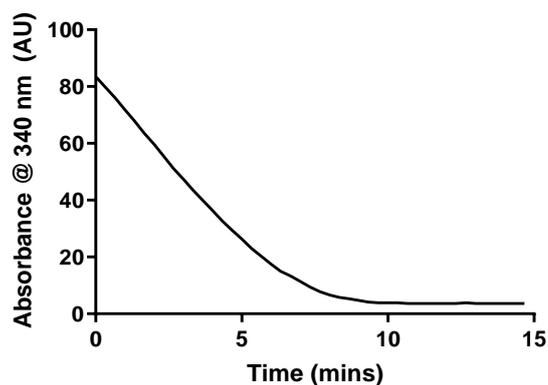
Figure 39: Standard reaction curve for InhA over 15 mins. Conditions: 100 µM NADH, 400 µM OcCoA, 150 nm InhA, 30 mM PIPES buffer, 30 °C, pH 6.8.

Table 12: Initial velocities for varying NADH concentration, results taken in duplicate.

[NADH] (µM)	Average initial velocity ± SE (µM.min^{-1})
120	10.22 ± 0.09
100	9.30 ± 0.12
75	9.59 ± 0.10
50	7.87 ± 0.17
20	5.79 ± 0.06
10	3.86 ± 0.22
0	0
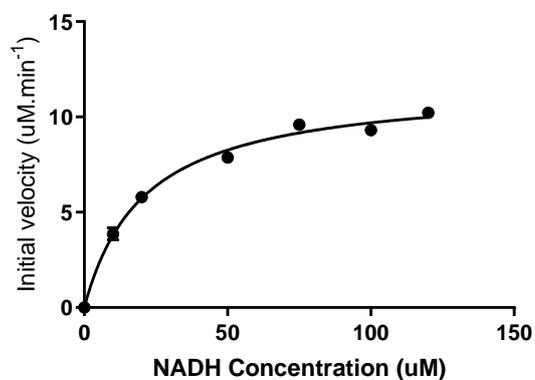
Figure 40: Michaelis Menton plot for InhA rate *vs* NADH concentration.

Table 13: Initial velocities for varying OcCoA concentration, results taken in duplicate.

[OcCoA] (µM)	Average initial velocity ± SE (µM.min^{-1})
415	9.09 ± 0.20
400	8.29 ± 0.15
300	7.48 ± 0.08
200	6.24 ± 0.10
150	4.81 ± 0.18
100	3.92 ± 0.14
50	2.61 ± 0.16
0	0.24 ± 0.10
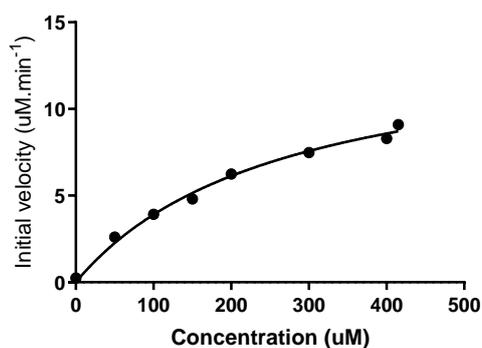
Figure 41: Michaelis Menton plot for InhA rate *vs* OcCoA concentration.

Table 14: NADH K_m, OcCoA K_m and Triclosan IC_{50} results compared with other studies. Obtained K_m results are within three-fold of each study, error in SE.

Study	NADH K_m (µM)	OcCoA K_m (µM)	V_{max} (µM.min^{-1})	Triclosan IC_{50} (µM)
Current	21.0 ± 1.9	267 ± 31	14.3 ± 0.62	9.17 ± 1.3
Ahmed[119]	26.0 ± 4.5	428 ± 195	19.0 ± 0.73	6.14 ± 1.1
Quemard[120]	7.60 ± 0.5	467 ± 90	3.60 ± 0.50	-
Khan[134]	19.1 ± 5.4	528 ± 17	15.3 ± 2.4	-

Table 15: Collated information of the top 148 hits taken from screening of the NMCCC library and their associated GoldScore and point inhibition data at 50 µM. Compounds synthesised in chapter 3 are highlighted in blue. Compounds synthesised in chapter 4 are highlighted in orange. The remaining 11 compounds discussed in the conclusion are highlighted in black. Assays were run in duplicate at pH 6.8 in 30 mM PIPES buffer at 30 °C using 150 nM InhA, 100 µM NADH and 400 µM OcCoA with a final DMSO concentration of 0.5%.
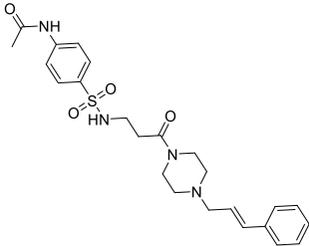
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
35		NCC-00020625	471	1.58	104.27	7
S1		NCC-00008099	446	4.32	100.86	12
S2		NCC-00017189	461	3.15	100.46	18
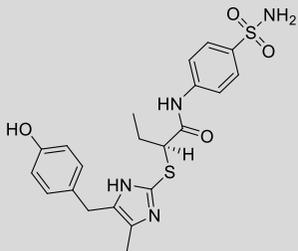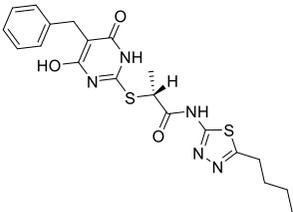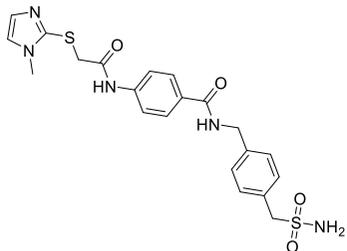
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
36		NCC-00078392	474	1.16	99.09	< 5
S3		NCC-00044634	463	1.32	97.75	19
S4		NCC-00019782	429	2.57	97.63	< 5
S5		NCC-00013648	438	4.38	97.33	36
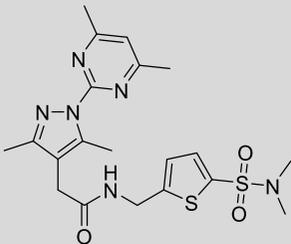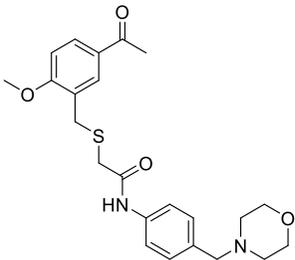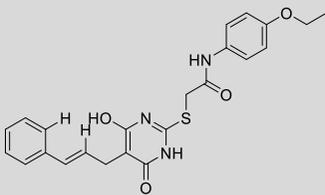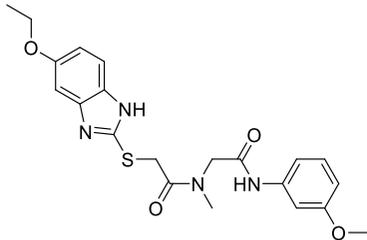
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
S6		NCC-00061384	429	2.33	97.15	< 5
S7		NCC-00069051	464	2.09	96.57	8
S8		NCC-00003129	430	2.95	96.55	14
75		NCC-00015764	447	3.35	96.42	70
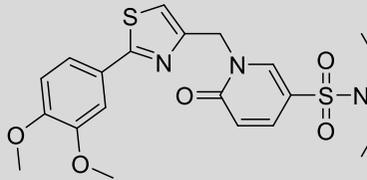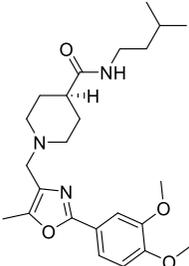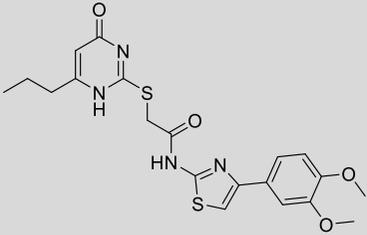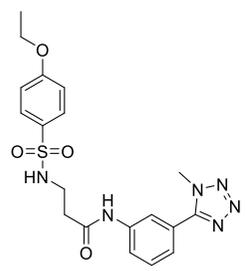
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
S9		NCC-00049062	430	1.56	96.37	< 5
S10		NCC-00013343	487	2.05	96.37	26
S11		NCC-00003112	427	3.29	96.32	< 5
S12		NCC-00044159	475	2.77	96.18	< 5
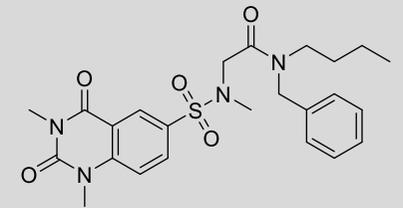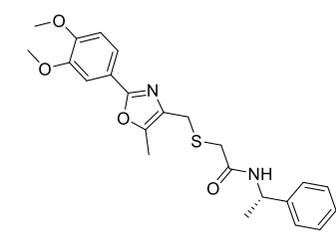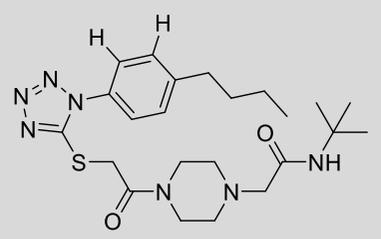
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
S13		NCC-00044711	422	2.43	95.84	< 5
37		NCC-00043102	405	3.40	95.83	15
38		NCC-00038339	482	3.00	95.83	11
S14		NCC-00017216	458	3.59	95.75	< 5
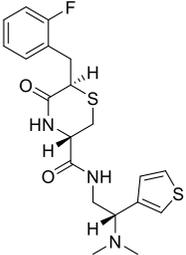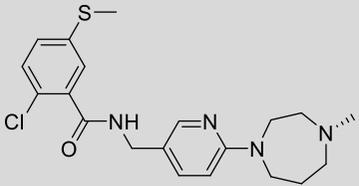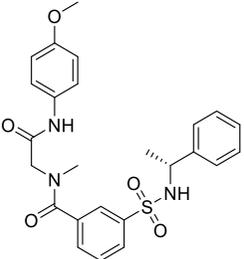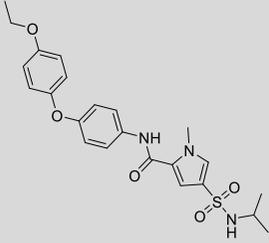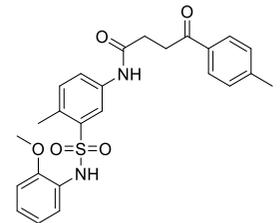
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S15		NCC-00062250	467	4.15	95.59	20
S16		NCC-00084480	467	3.25	95.51	19
S17		NCC-00076700	417	4.73	95.36	38
S18		NCC-00060655	370	2.74	95.05	21
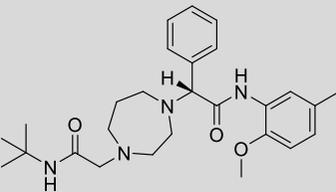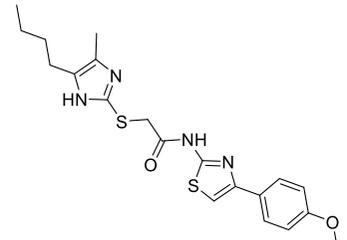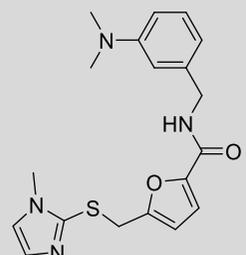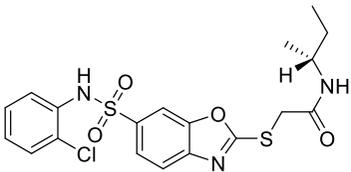
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
S19		NCC-00083041	454	3.60	95.02	6
S20		NCC-00031492	474	2.08	94.76	< 5
S21		NCC-00043752	451	2.48	94.51	< 5
S22		NCC-00035527	442	4.71	94.32	16
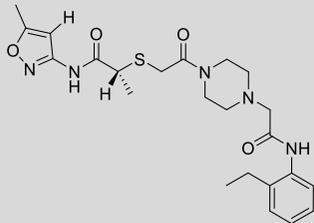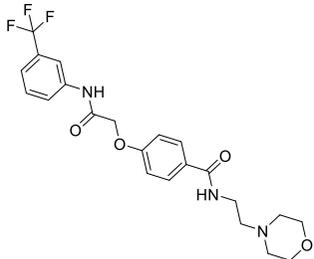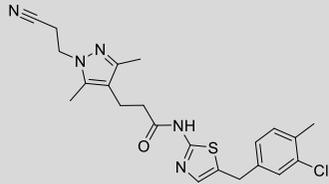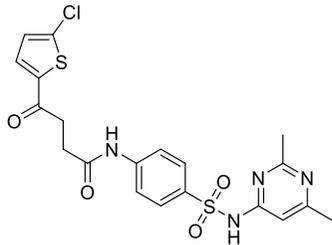
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S23		NCC-00035200	479	3.05	94.31	30
S24		NCC-00050564	480	2.10	94.24	23
S25		NCC-00046305	489	1.66	94.22	< 5
S26		NCC-00029161	439	4.50	94.21	23
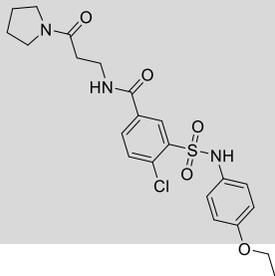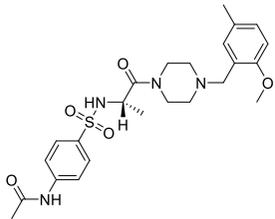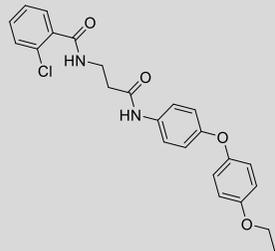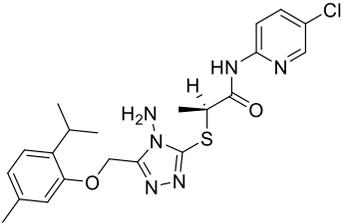
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
76		NCC-00030190	461	3.79	94.09	63
S27		NCC-00031433	460	3.24	94.08	< 5
74		NCC-00035138	441	3.68	93.64	63
S28		NCC-00019669	459	3.89	93.44	10
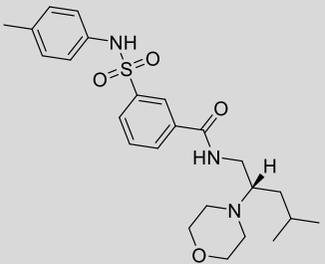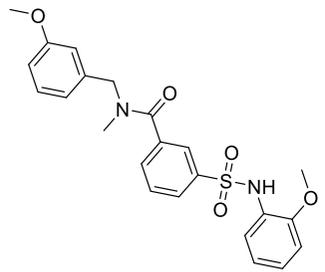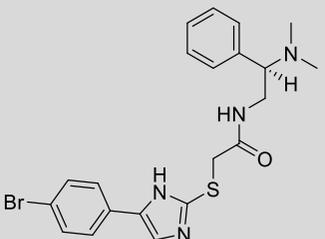
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S29		NCC-00073832	452	2.20	93.44	14
S30		NCC-00061756	422	4.21	93.42	< 5
S31		NCC-00078453	435	3.66	93.15	16
S32		NCC-00064307	342	2.08	93.14	< 5
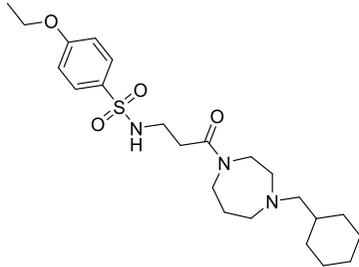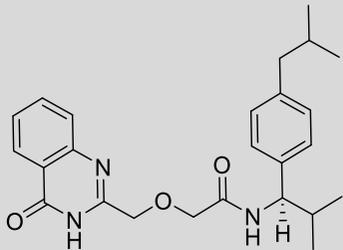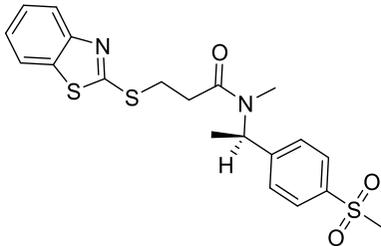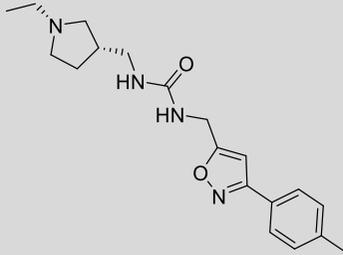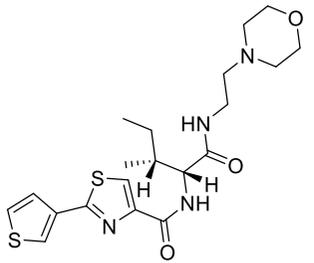
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S33		NCC-00032290	437	2.49	93.14	< 5
S34		NCC-00078455	423	2.57	93.13	10
S35		NCC-00072289	432	3.36	93.09	6
S36		NCC-00035976	461	3.97	93.07	14
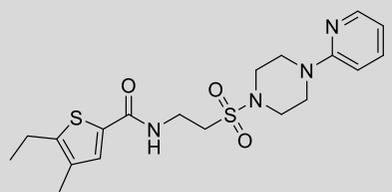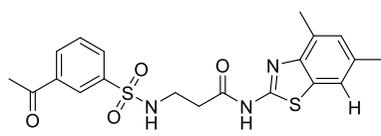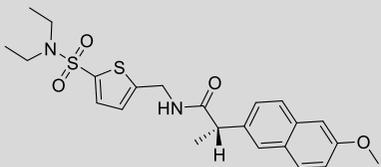
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S37		NCC-00061375	474	3.81	93.07	31
96		NCC-00038513	456	1.14	93.07	33
S38		NCC-00084383	495	-0.40	93.06	8
S39		NCC-00001131	459	2.30	93.05	8
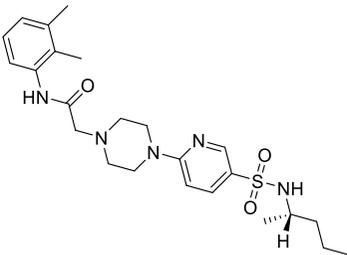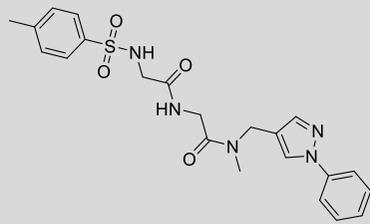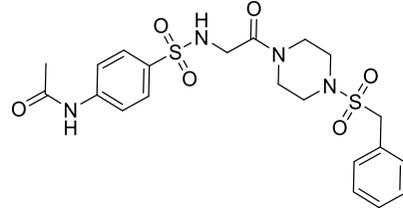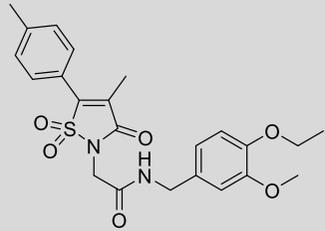
154

#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S40		NCC-00079144	461	3.01	92.95	12
S41		NCC-00083453	498	3.42	92.91	15
S42		NCC-00063300	457	4.29	92.89	6
78		NCC-00027115	477	3.66	92.88	55
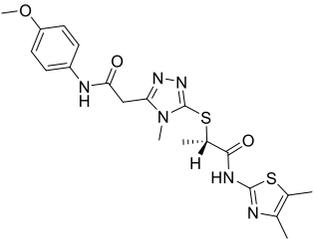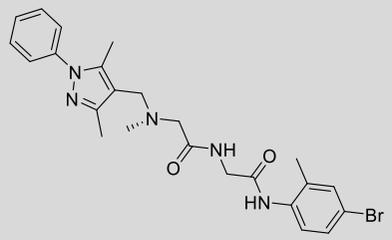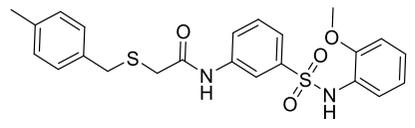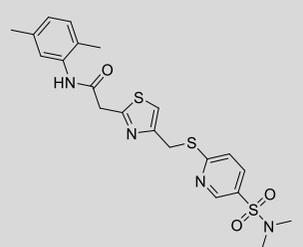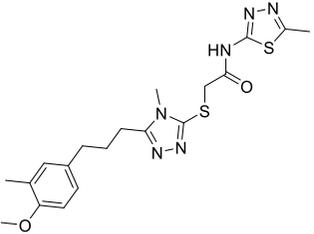
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
S43		NCC-00076135	433	2.82	92.83	< 5
S44		NCC-00014722	473	3.51	92.79	< 5
S45		NCC-00090465	482	6.60	92.79	< 5
S46		NCC-00073657	490	2.53	92.76	18
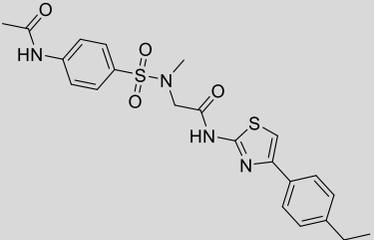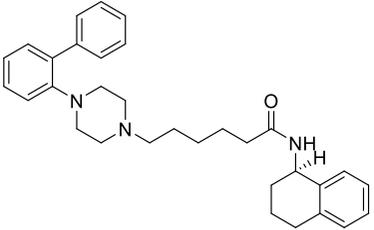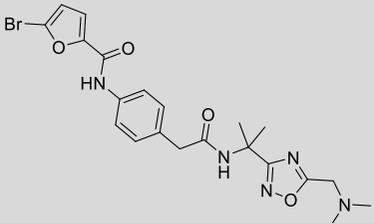
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
97		NCC-00035467	458	1.39	92.75	31
S47		NCC-00018829	487	3.68	92.71	16
S48		NCC-00036484	417	0.92	92.69	< 5
S49		NCC-00043875	445	2.83	92.66	< 5
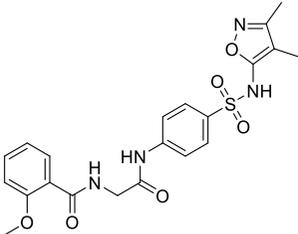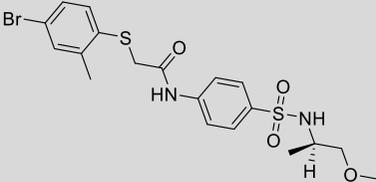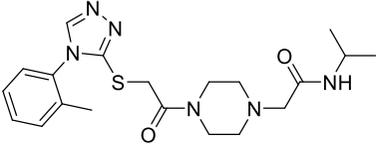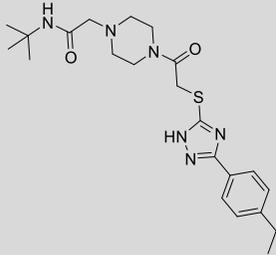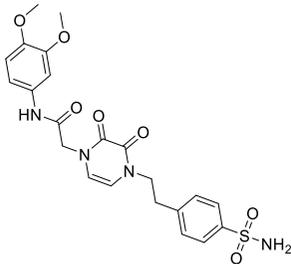
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S50		NCC-00011675	489	0.42	92.56	16
S51		NCC-00082137	433	3.67	92.56	9
S52		NCC-00043364	406	2.46	92.54	< 5
98		NCC-00062869	479	1.05	92.51	33
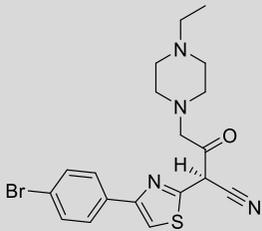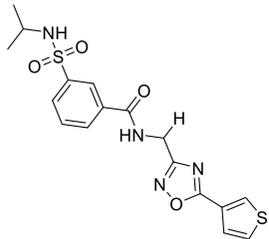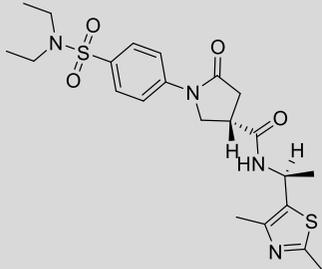
158

#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
S53		NCC-00043164	486	4.23	92.46	7
S54		NCC-00034272	441	0.80	92.42	< 5
S55		NCC-00057181	445	1.76	92.41	13
S56		NCC-00052224	451	2.32	92.41	5
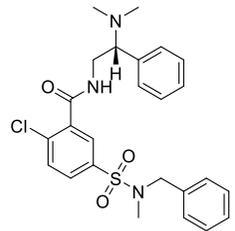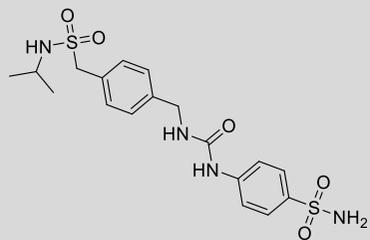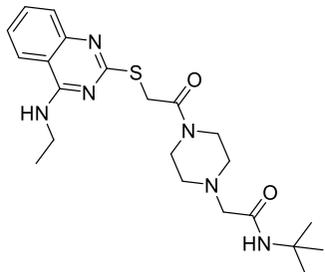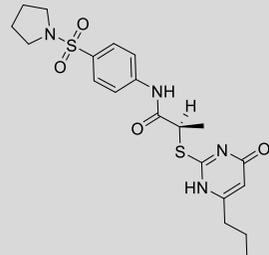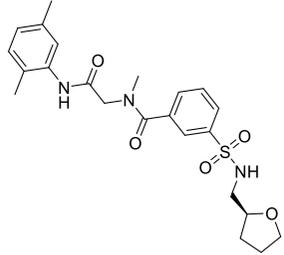
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S57		NCC-00015791	460	2.46	92.40	18
S58		NCC-00027573	432	1.97	92.37	< 5
S59		NCC-00061342	439	4.37	92.37	28
S60		NCC-00063795	442	2.83	92.31	26
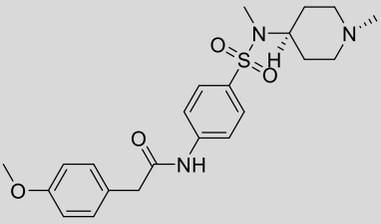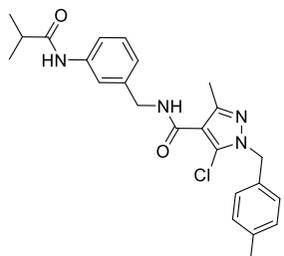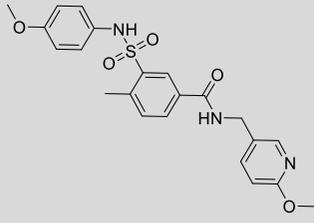
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
99		NCC-00084054	498	1.73	92.23	36
S61		NCC-00011206	471	3.35	92.19	< 5
S62		NCC-00049847	456	3.15	92.13	42
S63		NCC-00047309	476	2.60	92.09	9
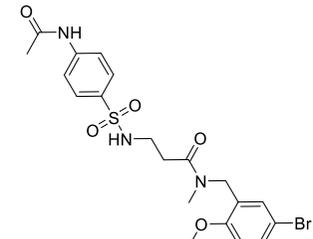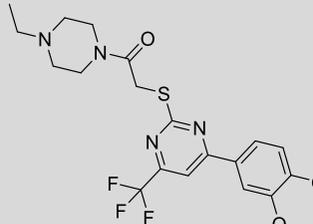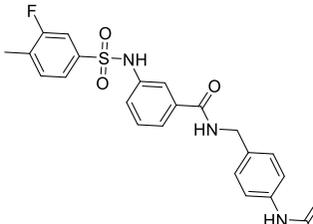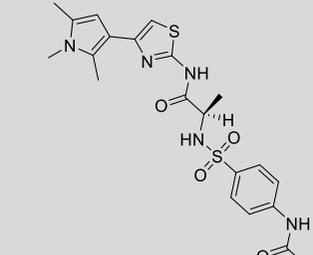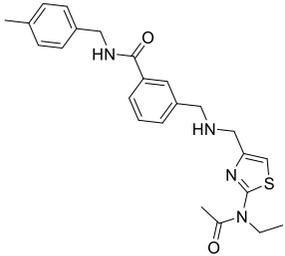
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
S64		NCC-00047455	437	3.45	91.99	33
100		NCC-00031435	448	1.34	91.98	37
S65		NCC-00003134	422	0.13	91.94	< 5
S66		NCC-00027221	475	1.72	91.93	21
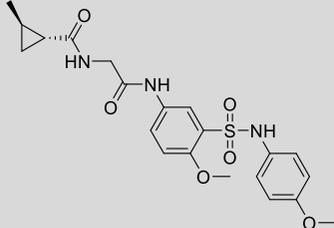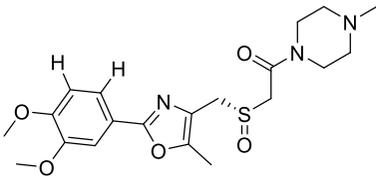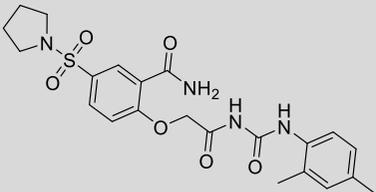
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S67		NCC-00007254	439	3.77	91.89	8
S68		NCC-00090135	418	7.46	91.87	35
101		NCC-00058480	420	2.25	91.85	36
S69		NCC-00077911	432	0.92	91.73	< 5
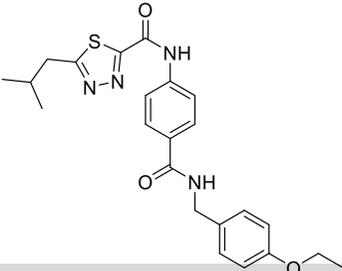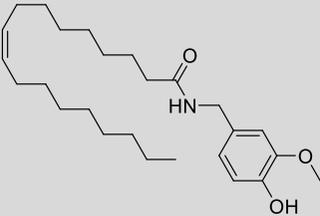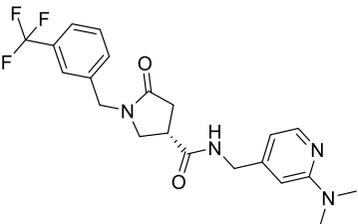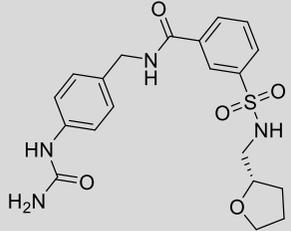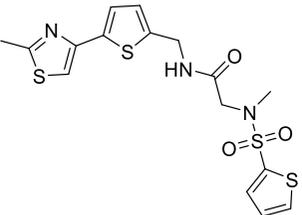
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
S70		NCC-00033199	428	2.32	91.70	6
S71		NCC-00059611	443	2.37	91.70	10
S72		NCC-00054702	376	3.68	91.70	< 5
S73		NCC-00046915	392	2.35	91.67	< 5
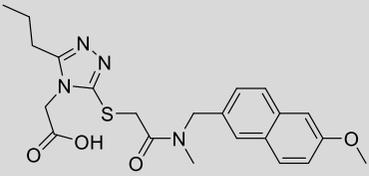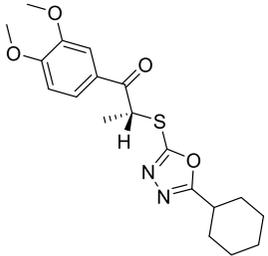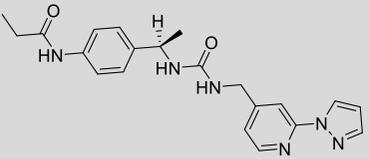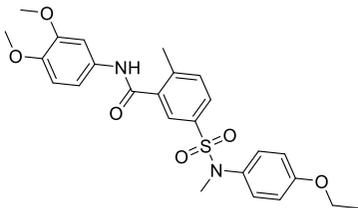
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S74		NCC-00084953	485	4.17	91.66	35
S75		NCC-00049697	470	1.26	91.66	10
S76		NCC-00082665	451	2.87	91.65	11
102		NCC-00027648	456	1.38	91.60	37
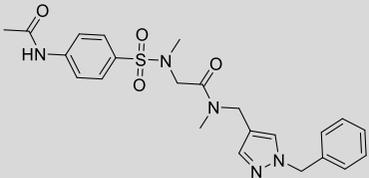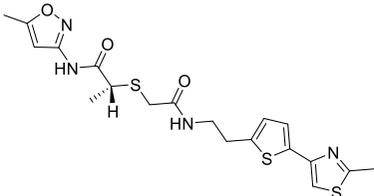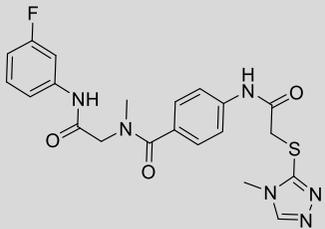
165

#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S77		NCC-00083759	480	4.20	91.55	35
S78		NCC-00044189	432	4.06	91.52	< 5
S79		NCC-00012159	412	2.84	91.51	15
S80		NCC-00088118	426	4.94	91.50	39
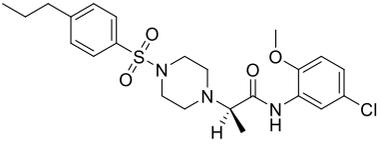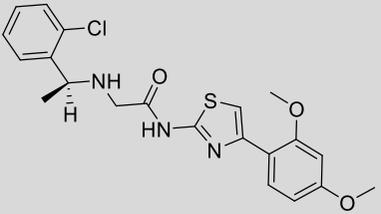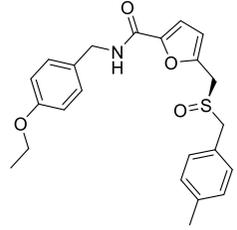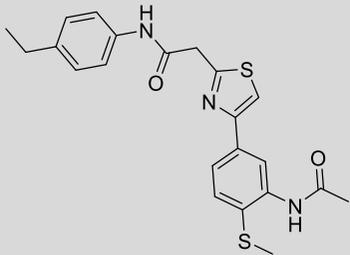
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S81		NCC-00043187	475	3.95	91.50	13
S82		NCC-00009861	458	4.14	91.43	9
S83		NCC-00083427	463	2.68	91.42	< 5
S84		NCC-00061296	419	1.47	91.41	< 5
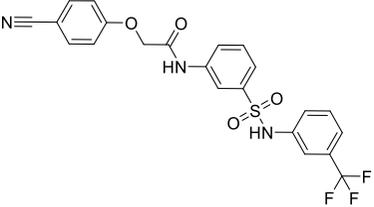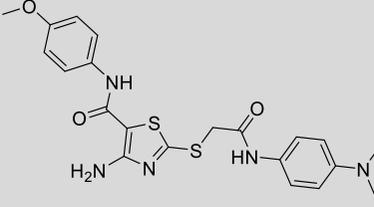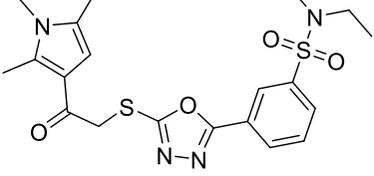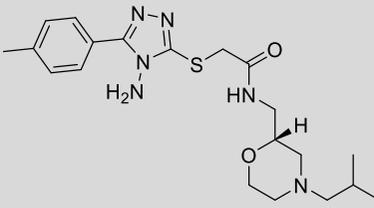
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
S85		NCC-00027947	480	3.24	91.38	18
S86		NCC-00077924	466	2.88	91.33	25
103		NCC-00089833	494	2.38	91.32	34
S87		NCC-00032159	496	1.66	91.29	25
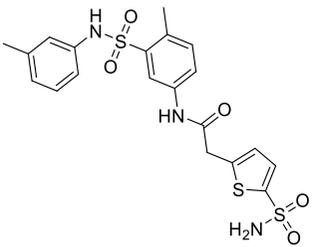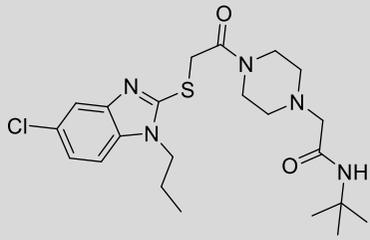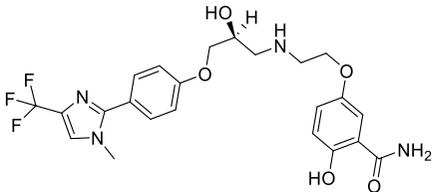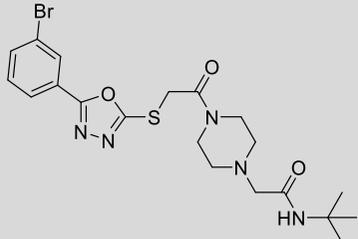
168

#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S88		NCC-00033682	409	2.12	91.28	13
S89		NCC-00017770	476	3.18	91.27	12
S90		NCC-00073713	437	2.87	91.25	9
S91		NCC-00066273	388	3.13	91.22	< 5
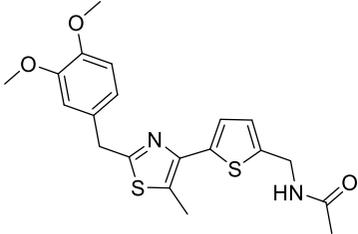
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S92		NCC-00032054	403	3.81	91.18	< 5
S93		NCC-00073582	439	4.62	91.17	< 5
S94		NCC-00034329	461	1.14	91.14	< 5
S95		NCC-00030345	488	3.01	91.12	13
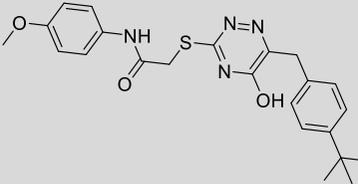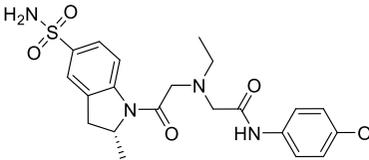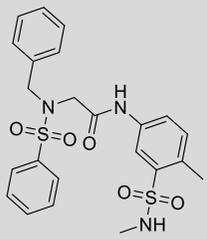
170

#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S96		NCC-00071344	485	3.40	91.12	< 5
S97		NCC-00005976	418	1.87	91.11	27
S98		NCC-00073967	446	2.43	91.11	25
S99		NCC-00073772	418	0.51	91.10	< 5
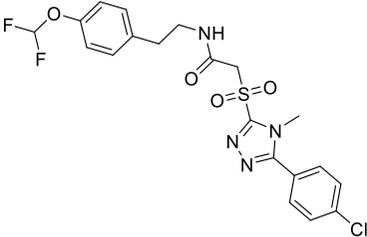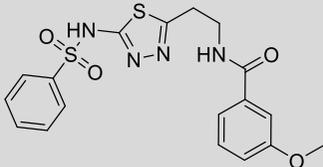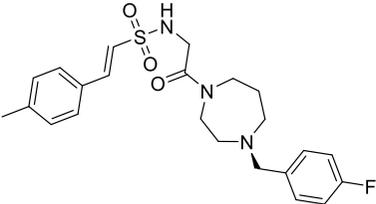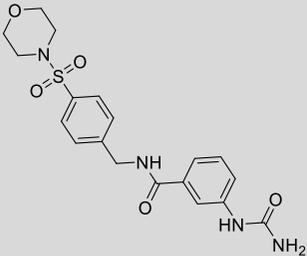
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S100		NCC-00011426	480	2.44	91.10	< 5
S101		NCC-00061952	456	3.43	91.09	10
S102		NCC-00019925	451	3.28	91.08	< 5
S103		NCC-00078768	419	2.30	91.00	19
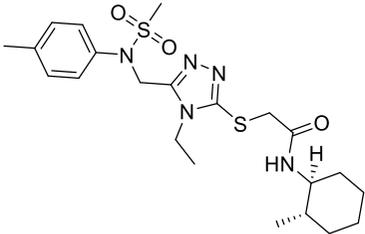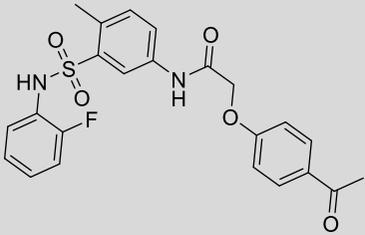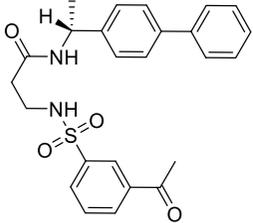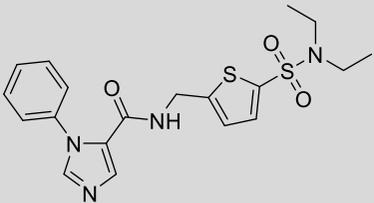
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
S104		NCC-00031976	474	1.82	91.00	19
104		NCC-00078613	478	0.45	90.98	31
S105		NCC-00016448	431	0.07	90.97	16
S106		NCC-00068742	473	5.39	90.93	38
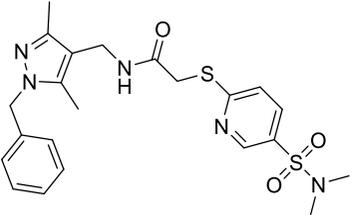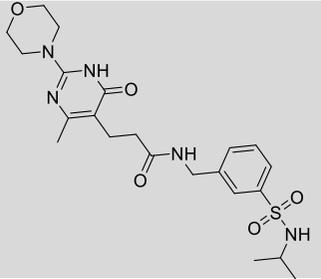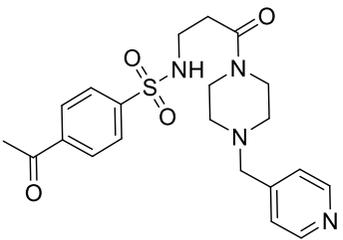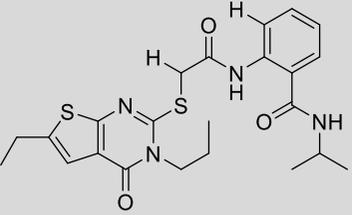
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S107		NCC-00061383	491	2.92	90.91	< 5
S108		NCC-00017269	435	2.74	90.91	6
77		NCC-00050259	413	3.32	90.91	70
S109		NCC-00035209	482	1.69	90.88	< 5
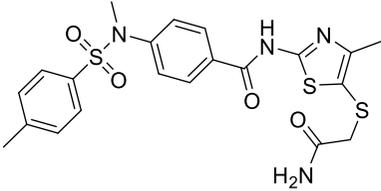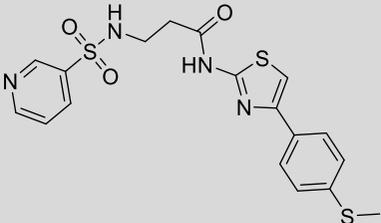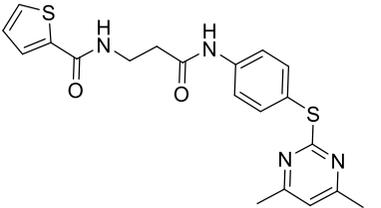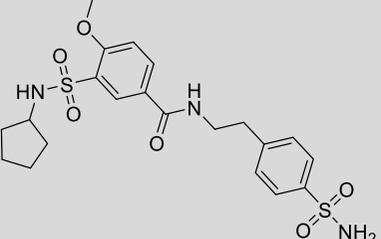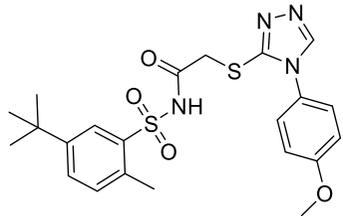
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S110		NCC-00038104	475	3.95	90.87	27
S111		NCC-00032818	397	2.82	90.87	6
S112		NCC-00086630	426	3.48	90.86	< 5
S113		NCC-00070517	436	3.74	90.85	< 5
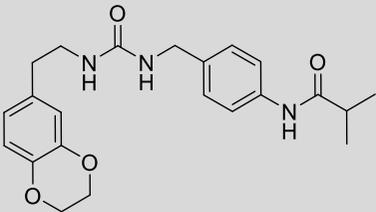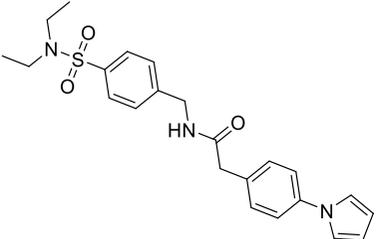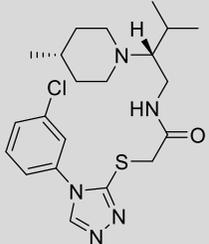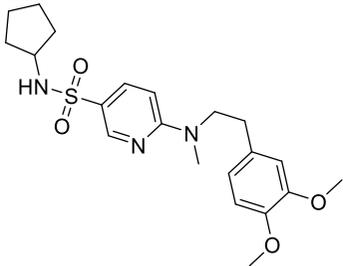
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S114		NCC-00088401	420	3.34	90.85	15
S115		NCC-00049492	485	4.04	90.84	21
S116		NCC-00067003	473	3.96	90.84	50
S117		NCC-00012351	413	3.20	90.84	15
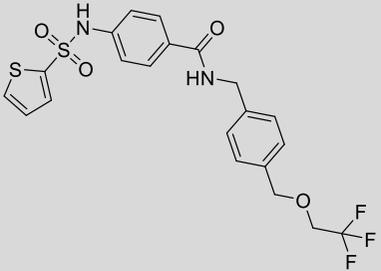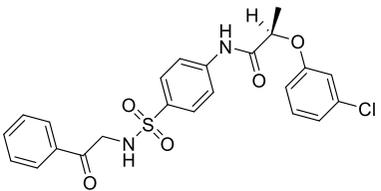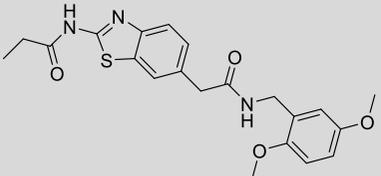
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S118		NCC-00014759	455	2.61	90.84	< 5
105		NCC-00033787	471	2.47	90.81	39
S119		NCC-00066289	462	3.60	90.81	< 5
S120		NCC-00019630	403	4.40	90.81	< 5
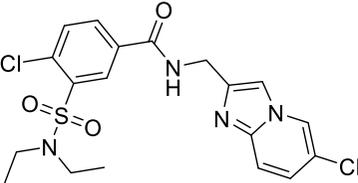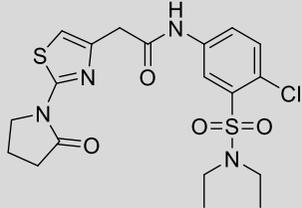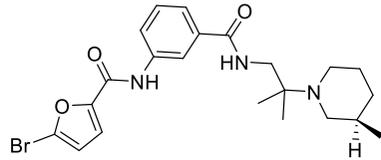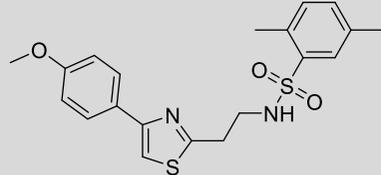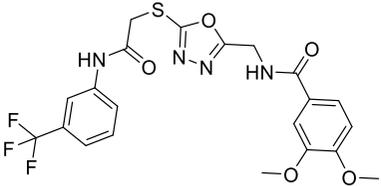
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 µM)
S121		NCC-00003261	496	2.30	90.79	8
S122		NCC-00043343	494	4.09	90.79	< 5
S123		NCC-00012485	414	0.38	90.78	< 5
S124		NCC-00012642	438	1.65	90.77	27
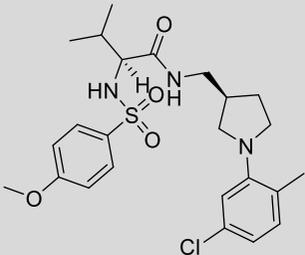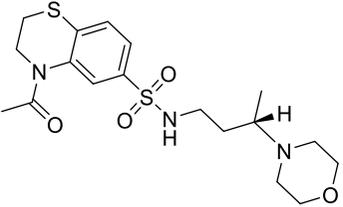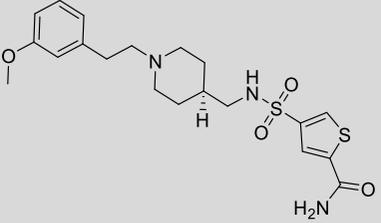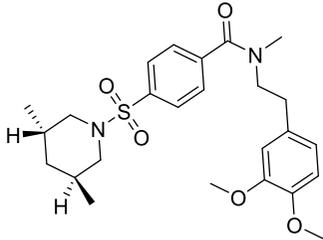
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
S125		NCC-00081729	475	3.60	90.77	< 5
S126		NCC-00034296	434	2.73	90.76	< 5
S127		NCC-00018863	446	4.54	90.75	35
S128		NCC-00059162	424	2.69	90.73	5
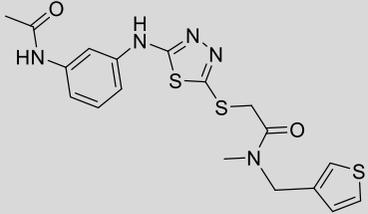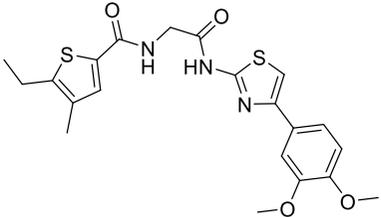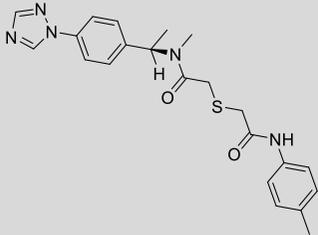
#	Structure	Sample Id	MW	cLogP	GScore	Point Inhibition (%, 50 μM)
106		NCC-00072471	494	2.57	90.69	32

8.3 Final compound data
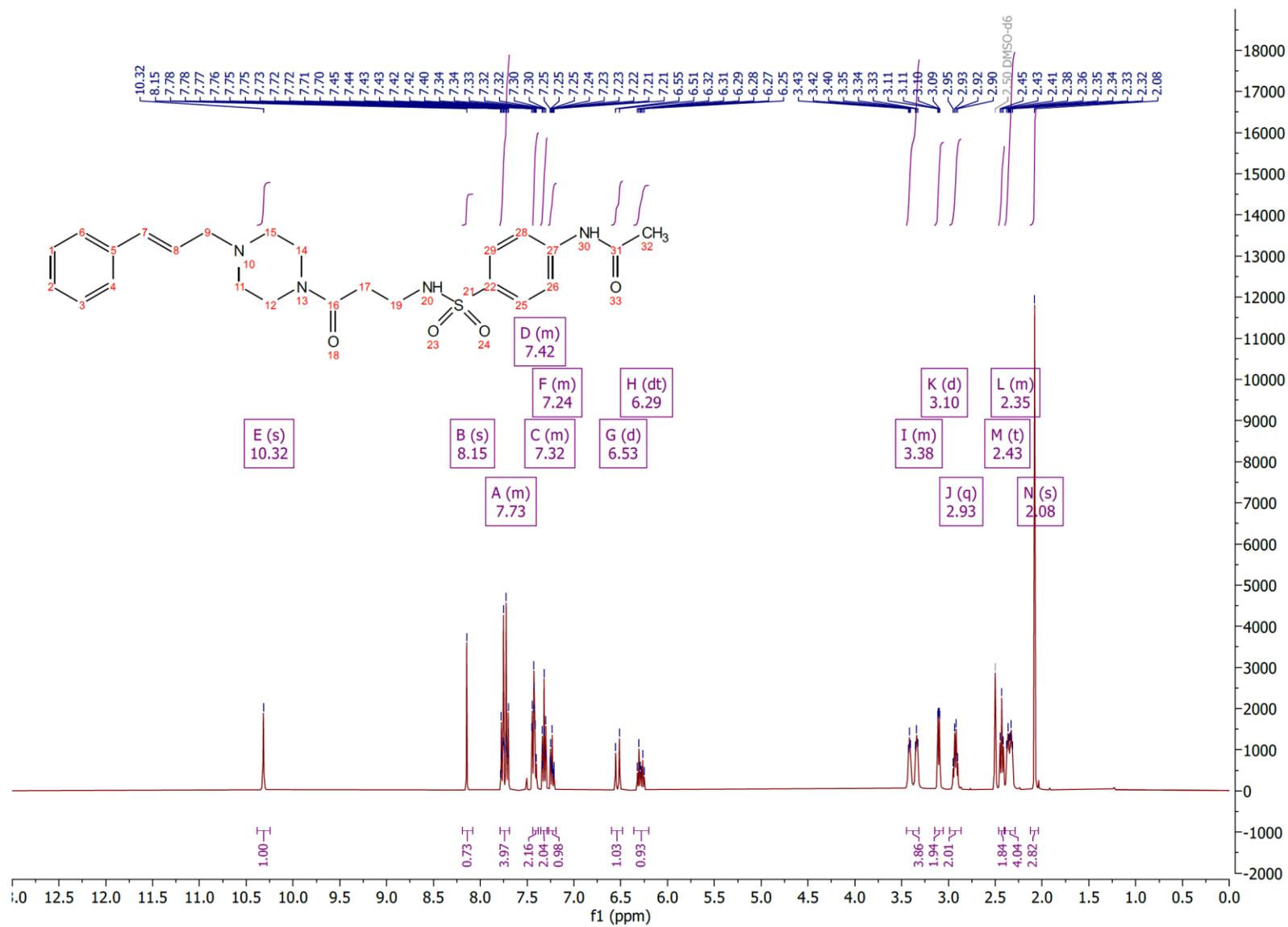
Figure 42: ^{1}H NMR spectra of compound **35** (400 MHz, dimethylsulfoxide-d_6).

■ Shimadzu LC Controller Detector A, Channel 1 from Sample 13 (336) of DataML2021.wiff   Max. 3.6e5

■ +Q1: 2.239 to 2.406 min from Sample 13 (336) of DataML2021.wiff (Turbo Spray), Baseline ...   Max. 2.6e6 cps

Figure 43: LCMS result for compound **35**, with a retention time of 2.30 min ($C_{24}H_{31}N_4O_4S^+$: 471.2, found 471.1).

Figure 44: ^{1}H NMR spectra of compound **36** (400 MHz, dimethylsulfoxide-d_6).

■ Shimadzu LC Controller Detector A, Channel 1 from Sample 14 (332) of DataML2021.wiff   Max. 3.2e5

■ +Q1: 2.039 to 2.239 min from Sample 14 (332) of DataML2021.wiff (Turbo Spray)   Max. 1.9e6 cps
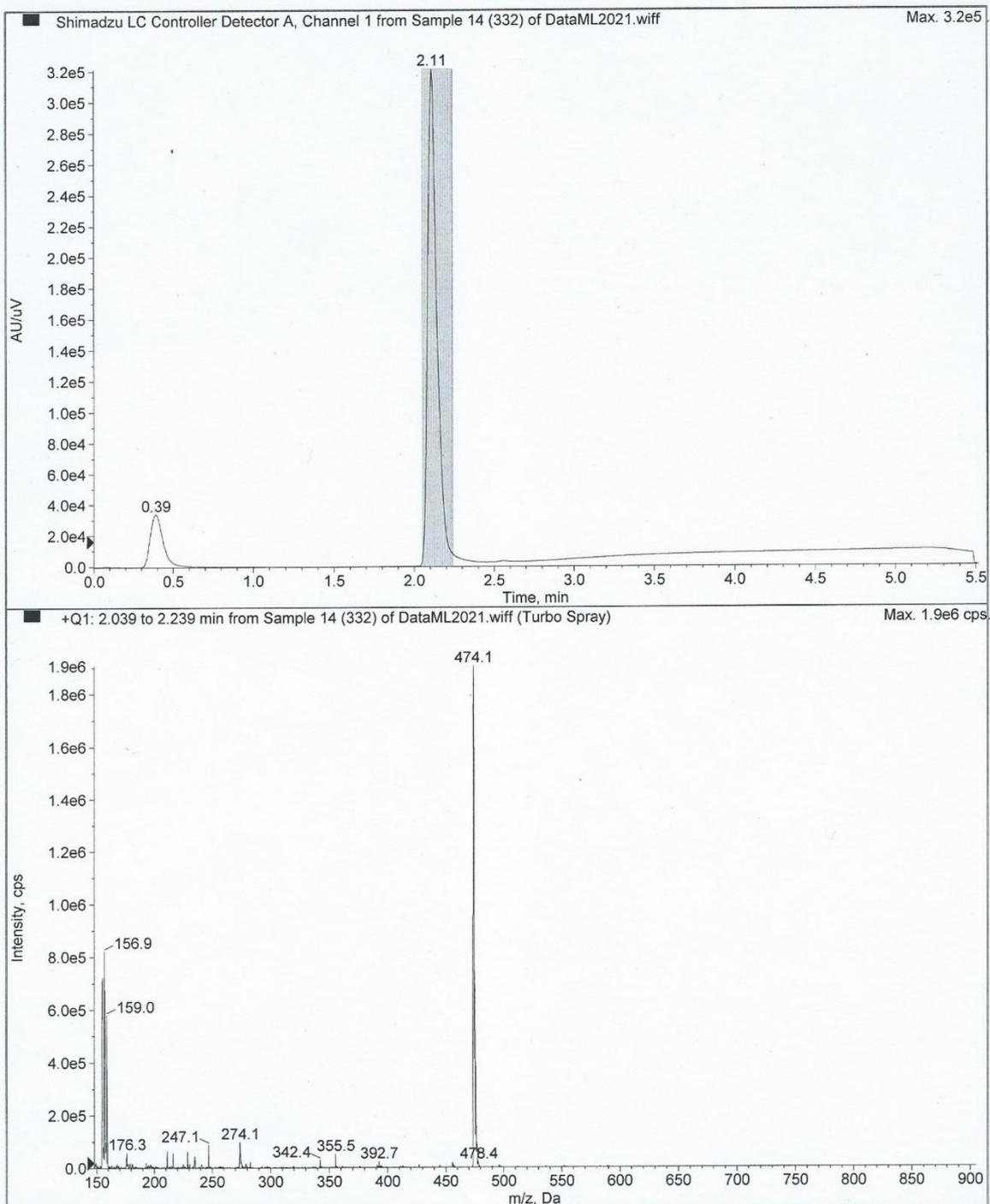
Collected by: API2000\hplc_user
Electronic Signature: no
Operator: hplc_user

Figure 45: LCMS result for compound **36** with a retention time of 2.11 min ($C_{21}H_{24}N_5O_4S_2^+$: 474.1, found 474.1).
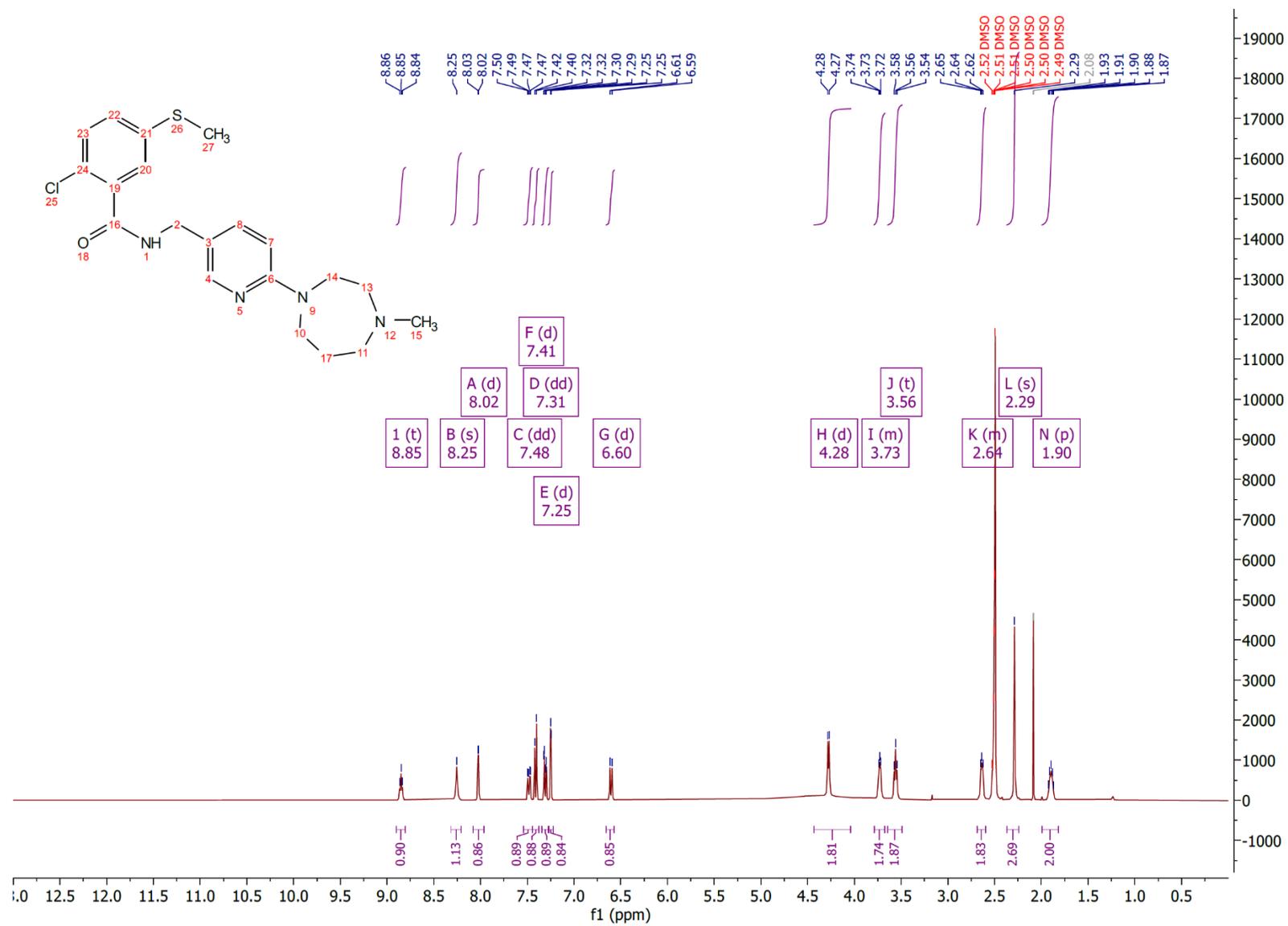
Figure 46: ^{1}H NMR spectra of compound **37** (400 MHz, dimethylsulfoxide-d6).

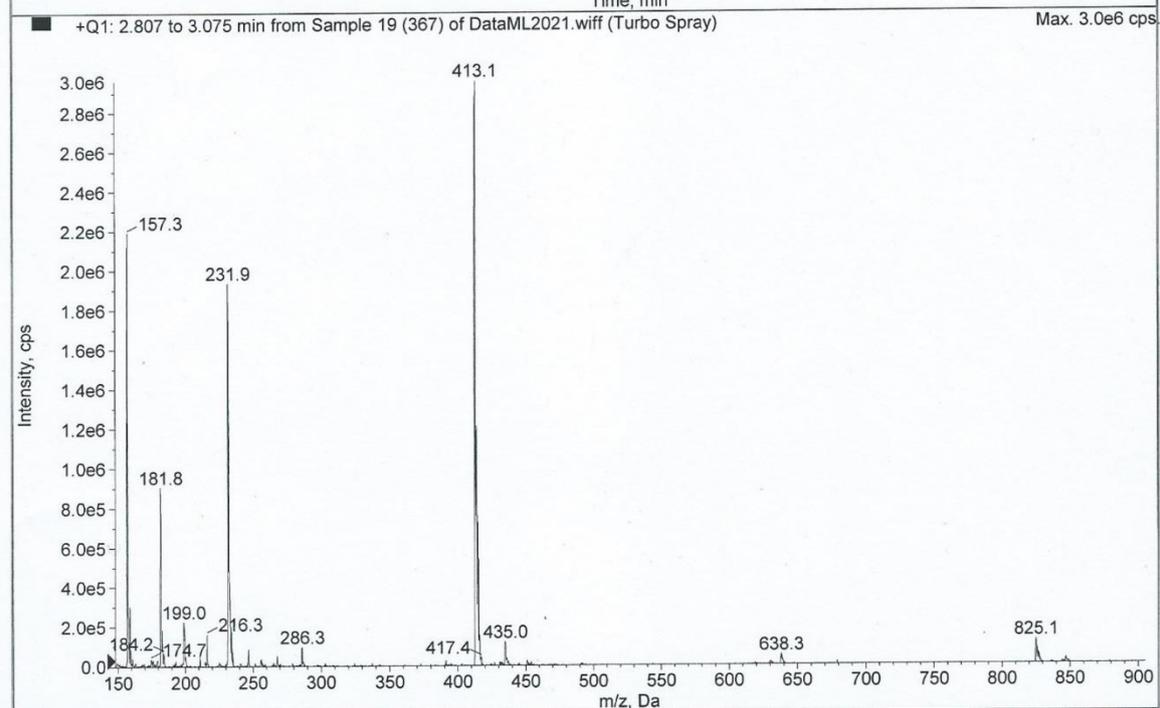■ Shimadzu LC Controller Detector A, Channel 1 from Sample 22 (3753) of DataML2021.wiff   Max. 1.6e5

■ +Q1: 1.872 to 2.172 min from Sample 22 (3753) of DataML2021.wiff (Turbo Spray)   Max. 2.1e6 cps
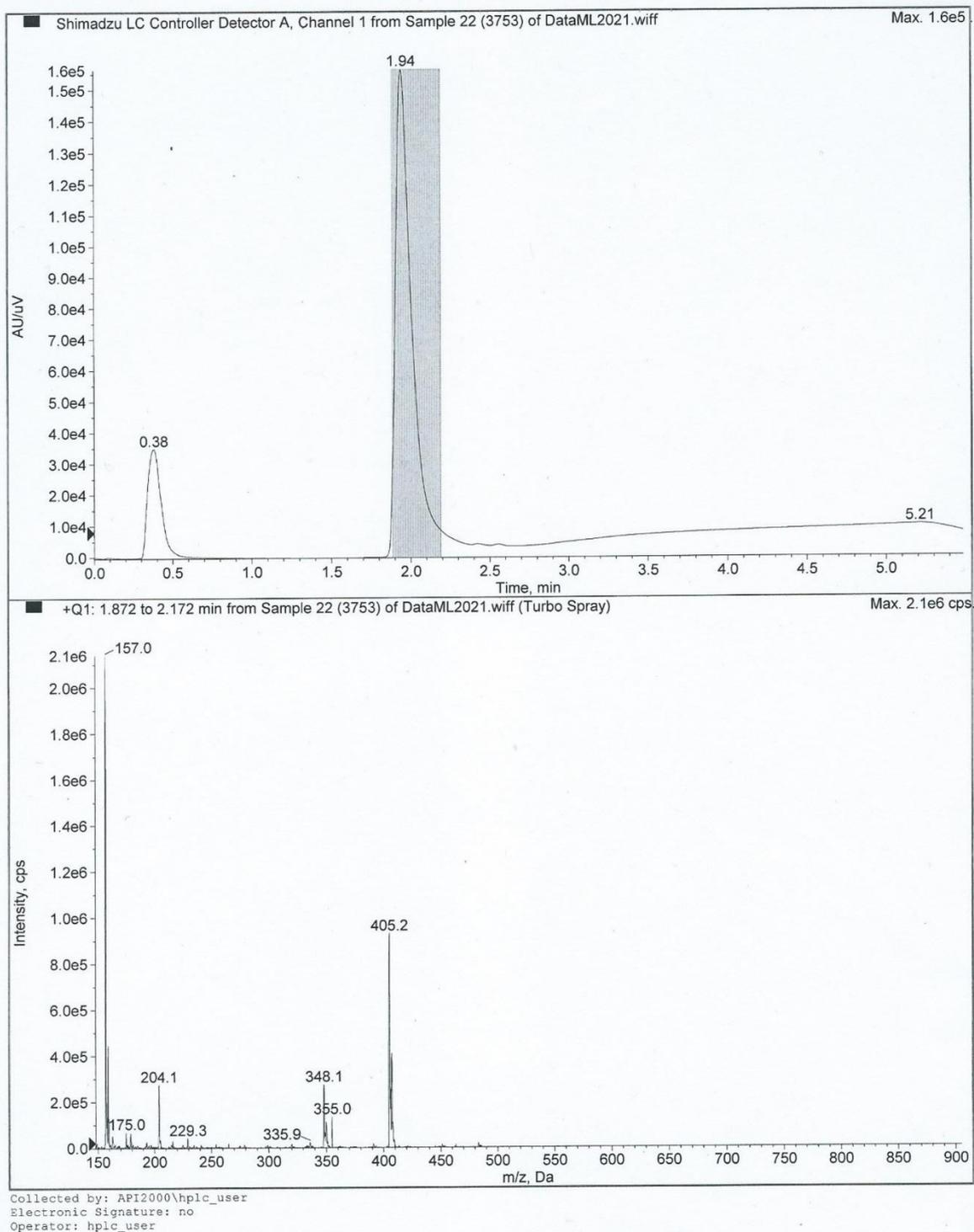
Figure 47: LCMS result for compound **37** with a retention time of 1.94 min ($C_{20}H_{26}ClN_4OS^+$: 405.2, found 405.2).
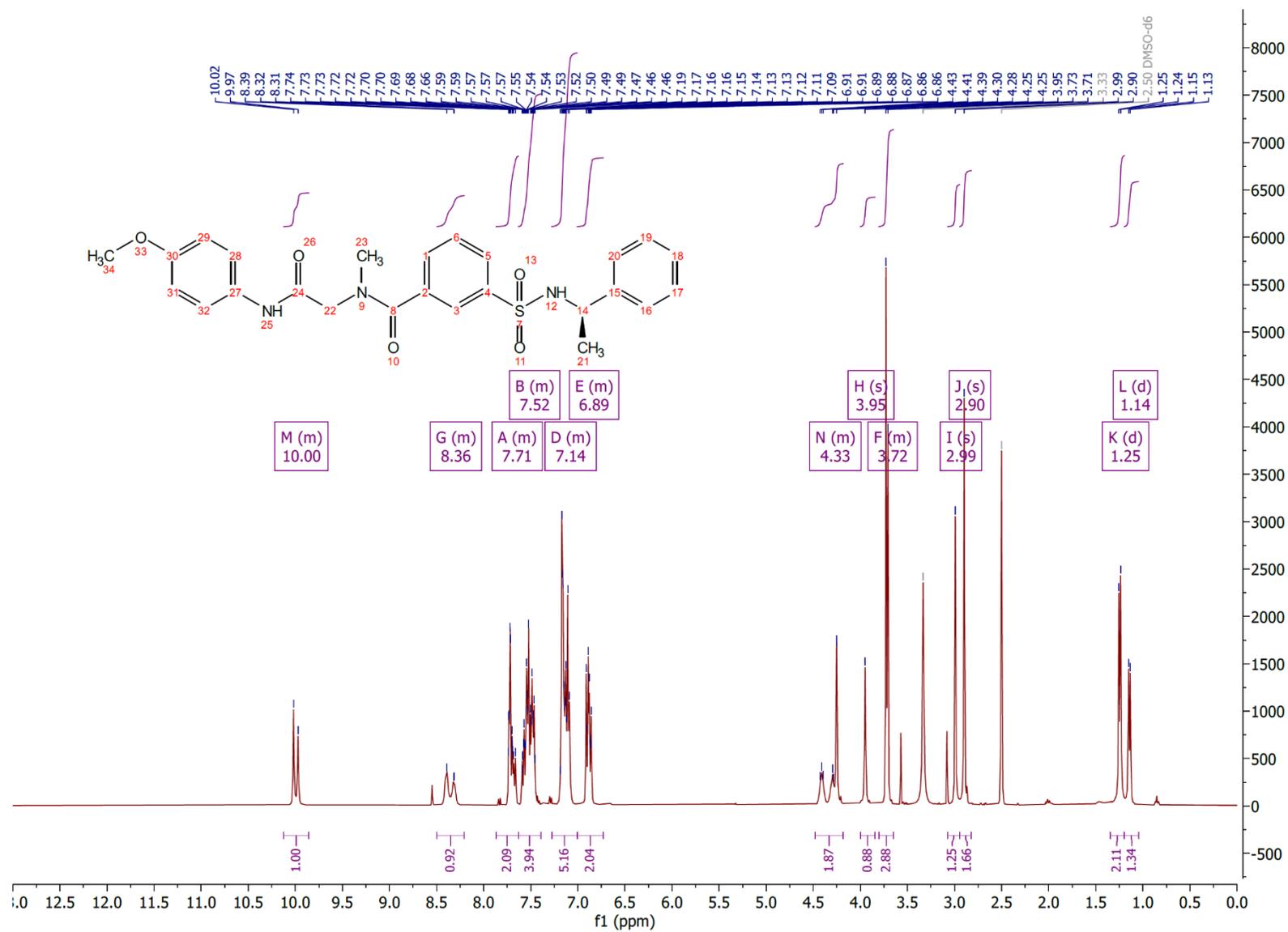
Figure 48: ^{1}H NMR spectra of compound **38** (400 MHz, dimethylsulfoxide-d6, 25 °C).
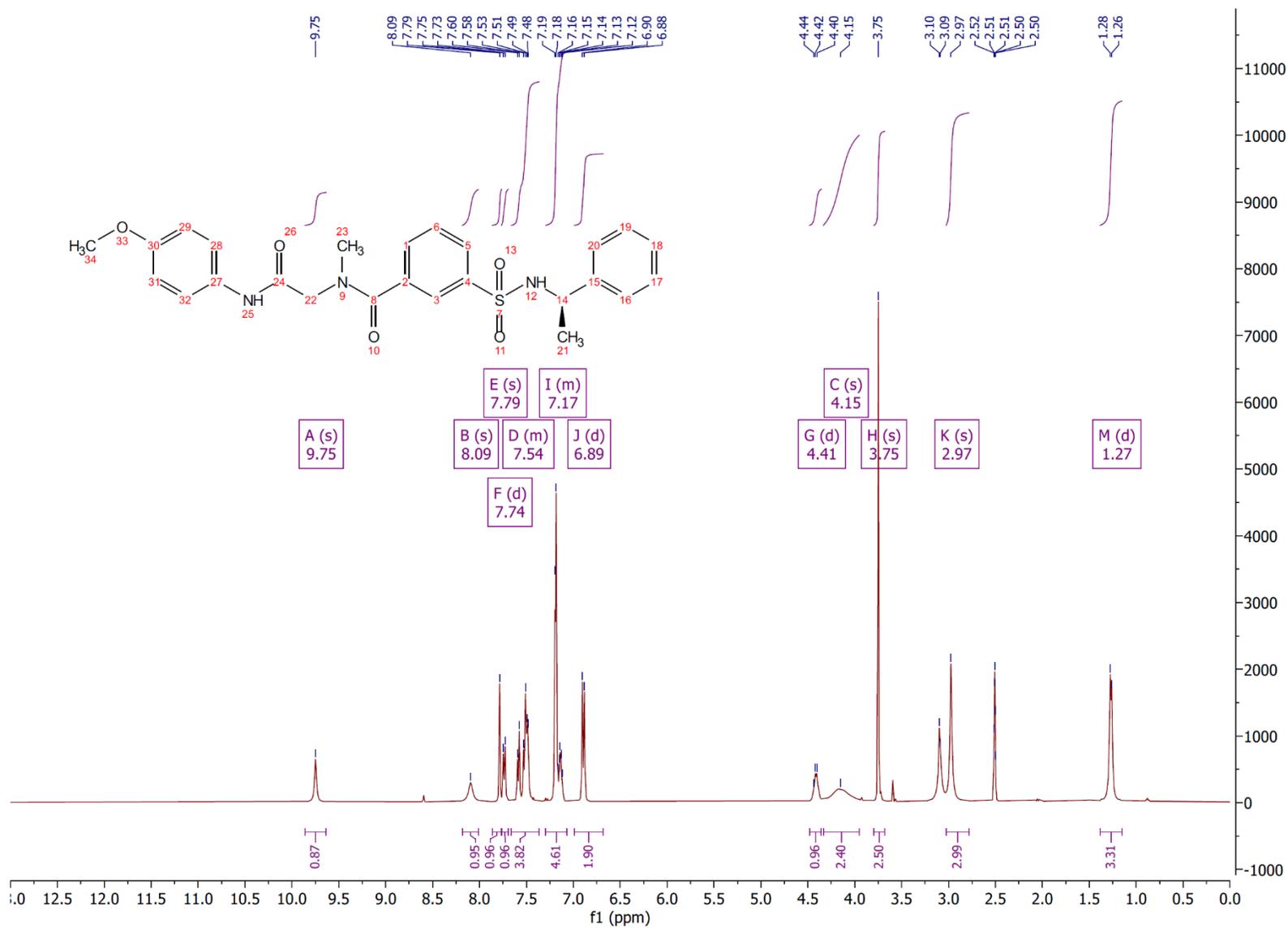
Figure 49: ¹H NMR spectra of compound **38** (400 MHz, dimethylsulfoxide-d6, 75 °C).
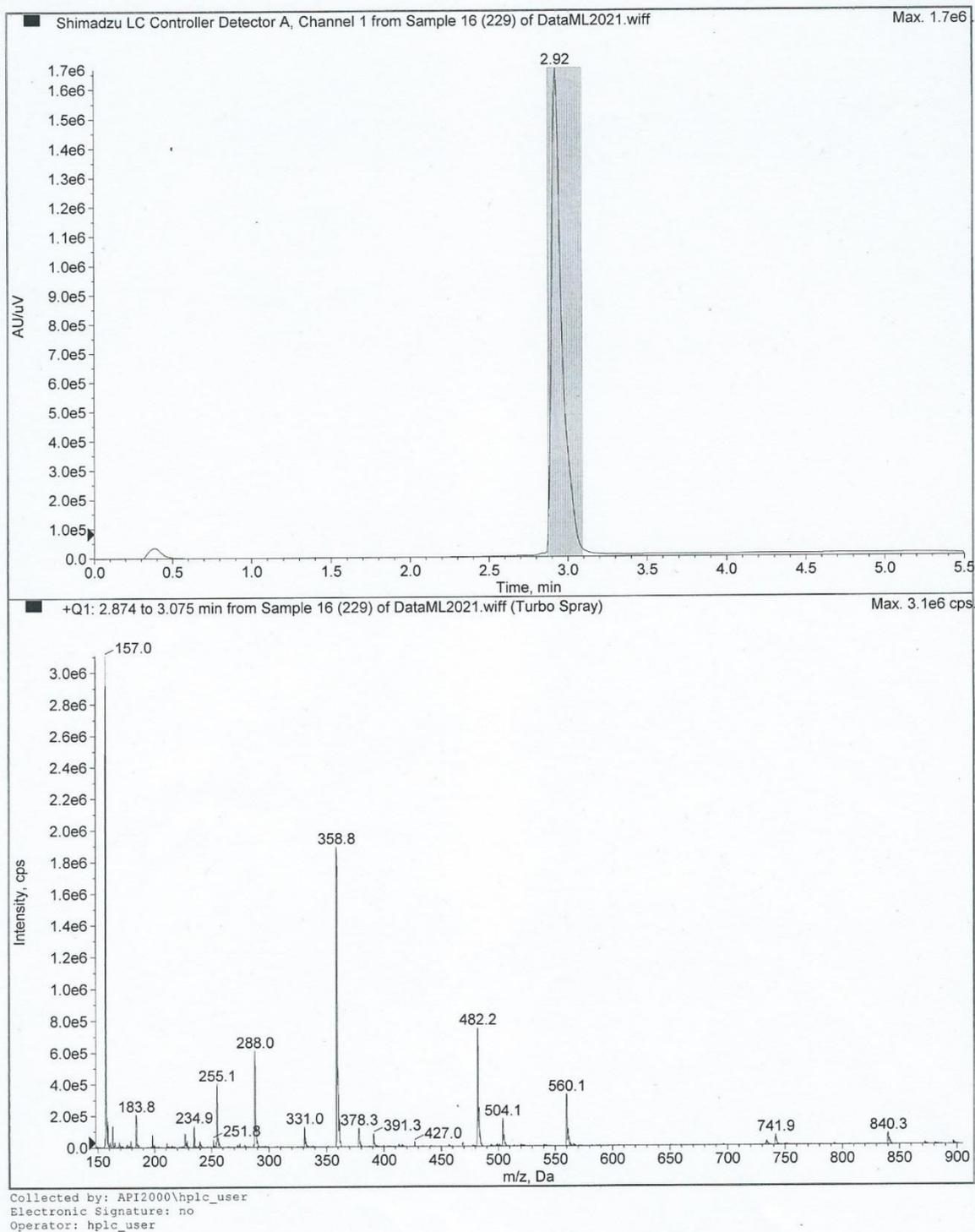
Figure 50: LCMS result for compound **38** with a retention time of 2.92 min ($C_{25}H_{28}N_3O_5S^+$: 482.2, found 482.2).
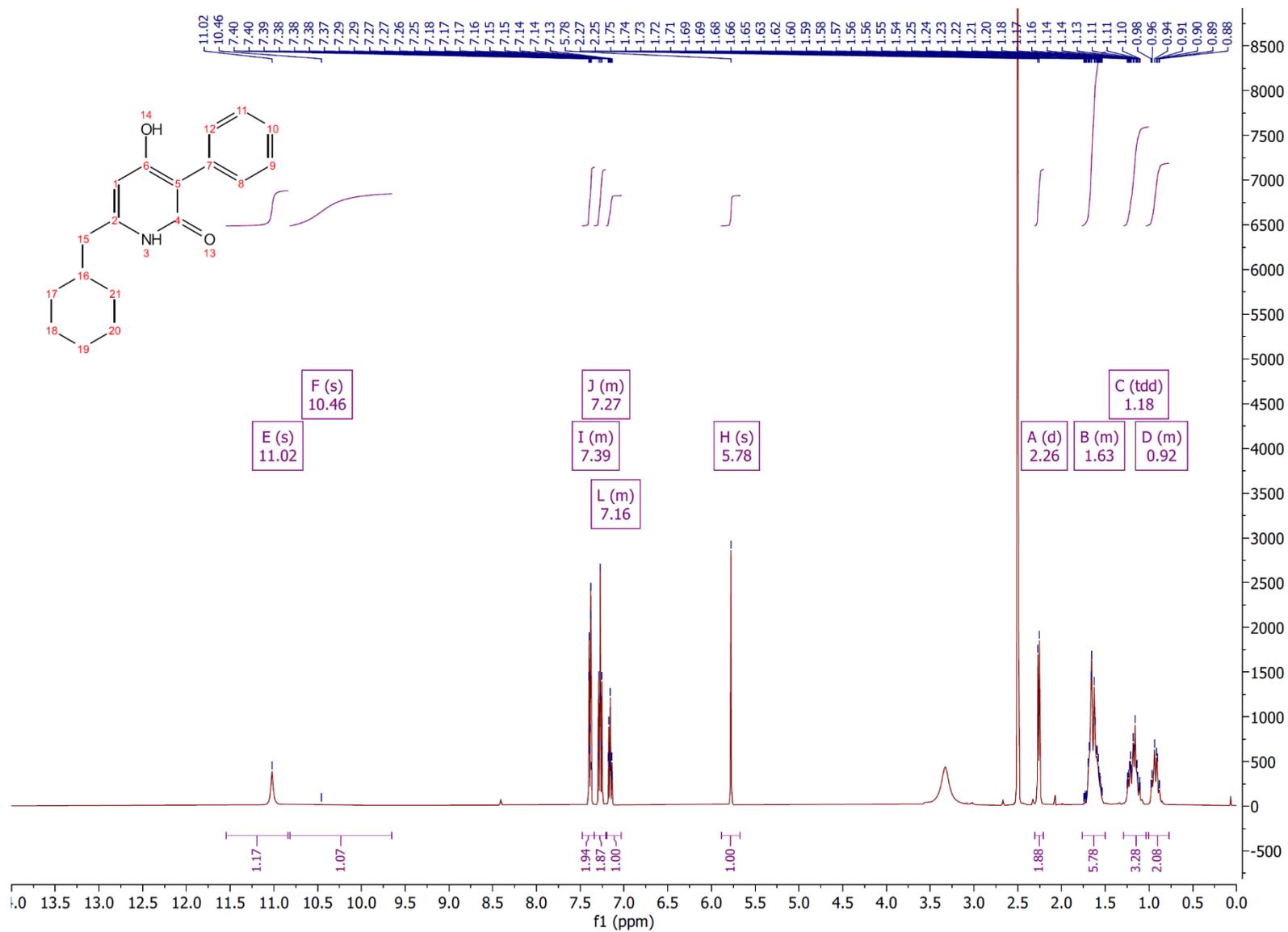
Figure 51: ^{1}H NMR spectra of **NITD-564** (400 MHz, dimethylsulfoxide-d_6).
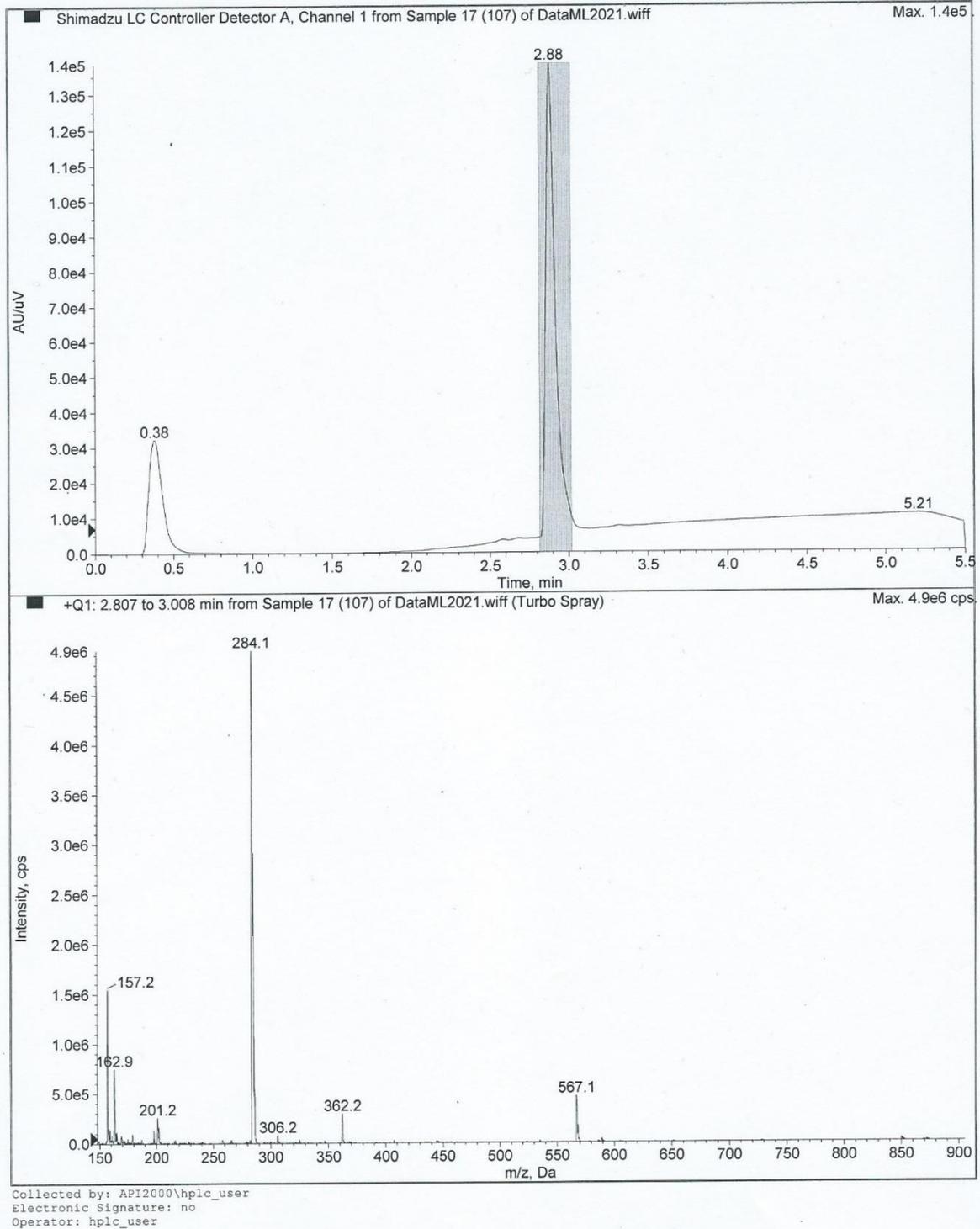
Figure 52: LCMS result for **NITD-564** with a retention time of 2.88 min ($C_{18}H_{22}NO_2^+$: 284.2, found 284.1).

192
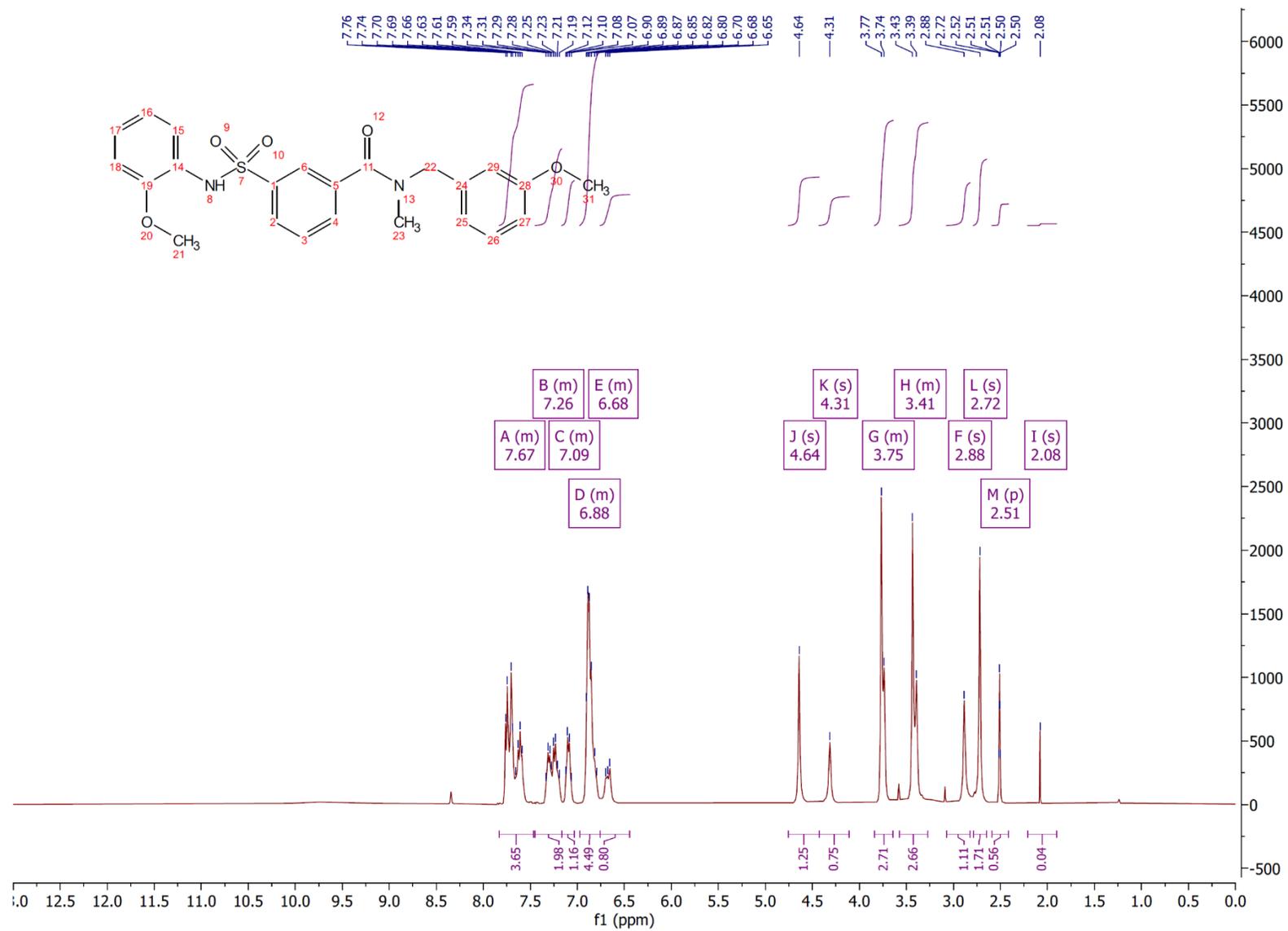
Figure 53: ^{1}H NMR spectra of compound **74** (400 MHz, dimethylsulfoxide-d6, 25 °C).
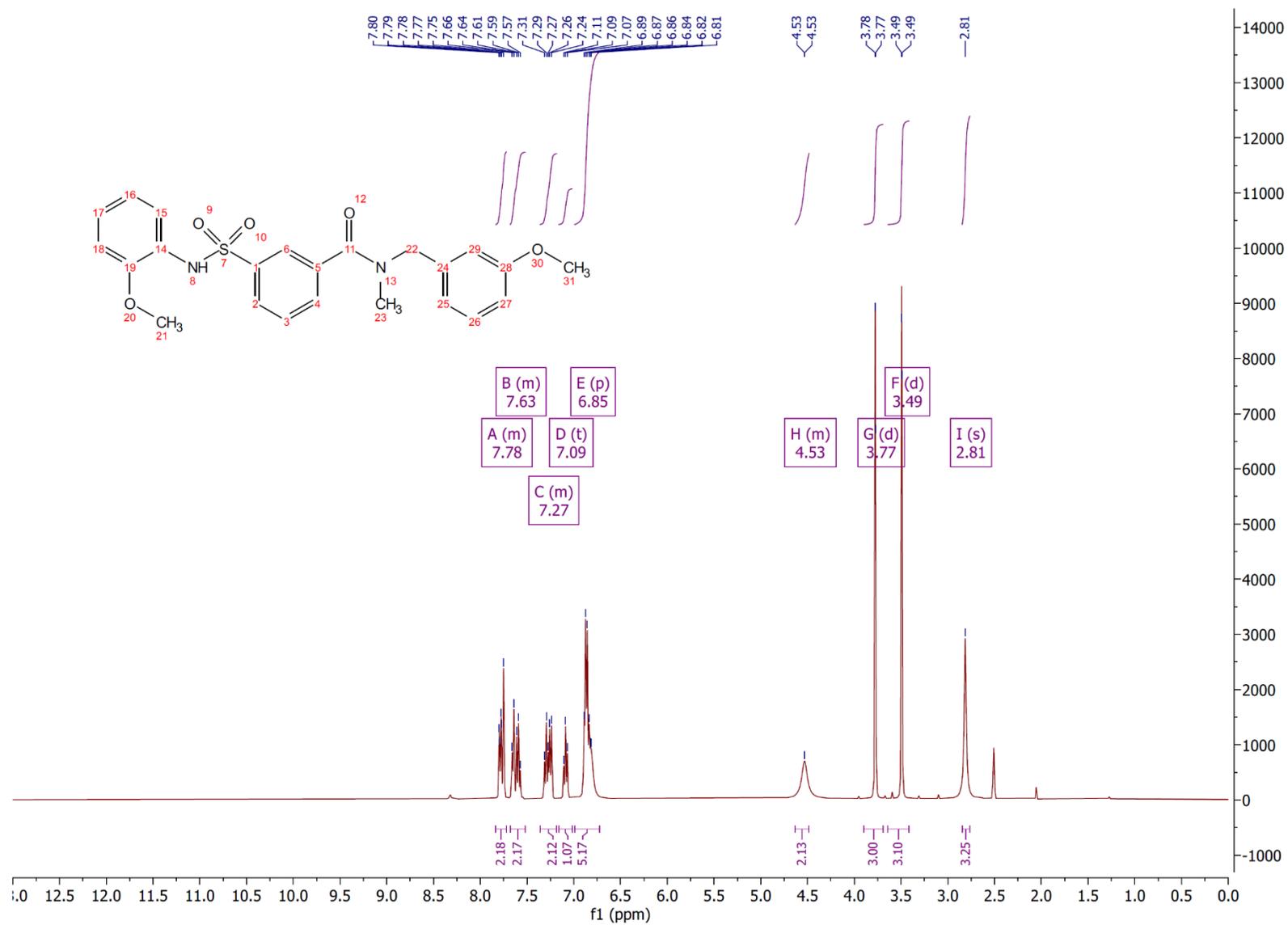
Figure 54: ¹H NMR spectra of compound **74** (400 MHz, dimethylsulfoxide-d6, 75 °C).

Figure 55: LCMS result for compound **74** with a retention time of 3.04 min ($C_{23}H_{25}N_2O_5S^+$: 441.1, observed: 441.0).
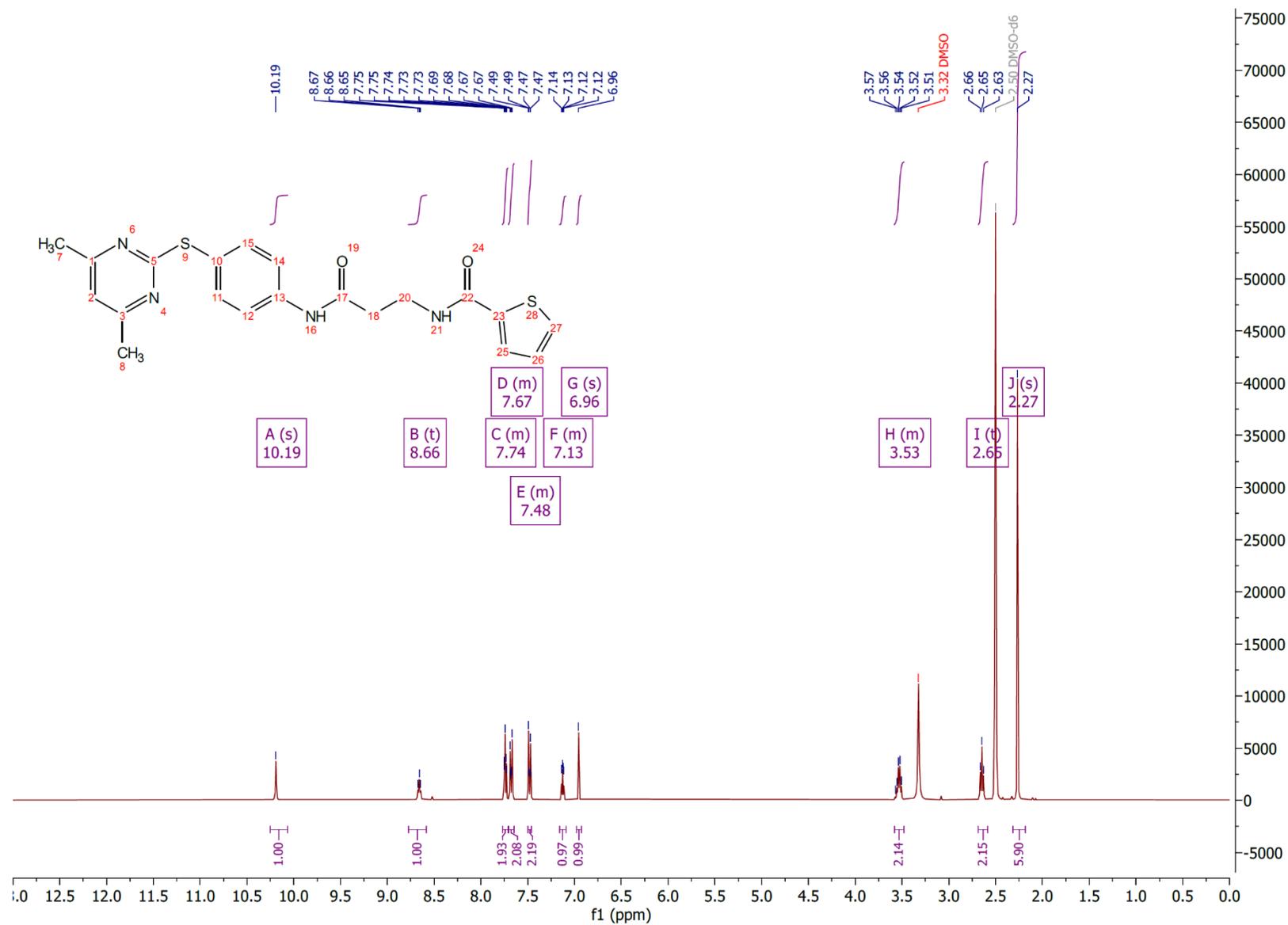
Figure 56: ^{1}H NMR spectra of compound **77** (400 MHz, dimethylsulfoxide-d6).

Figure 57: LCMS result for compound **77** with a retention time of 2.91 min ($C_{20}H_{21}N_4O_2S_2^+$: 413.1, observed: 413.1).
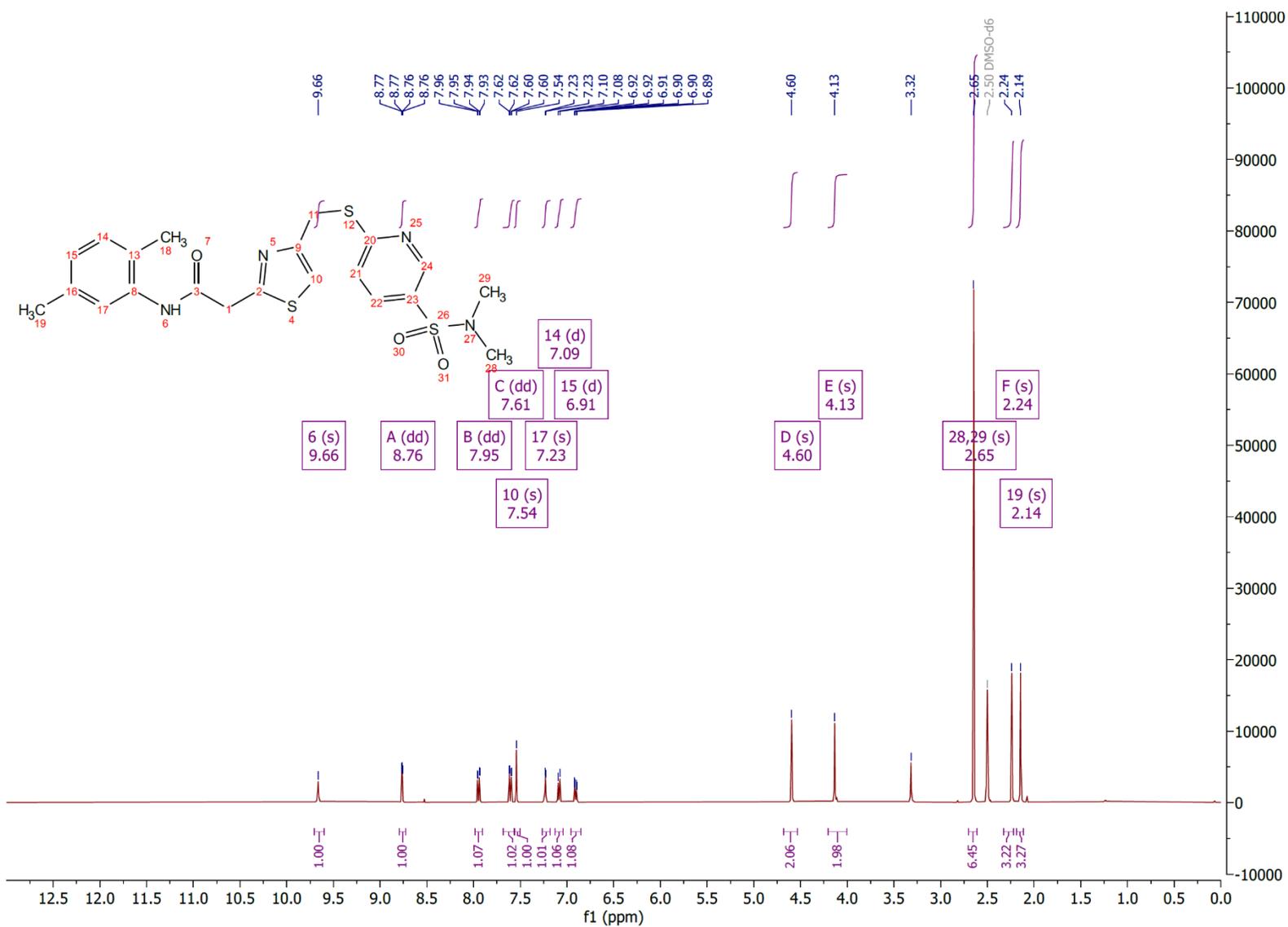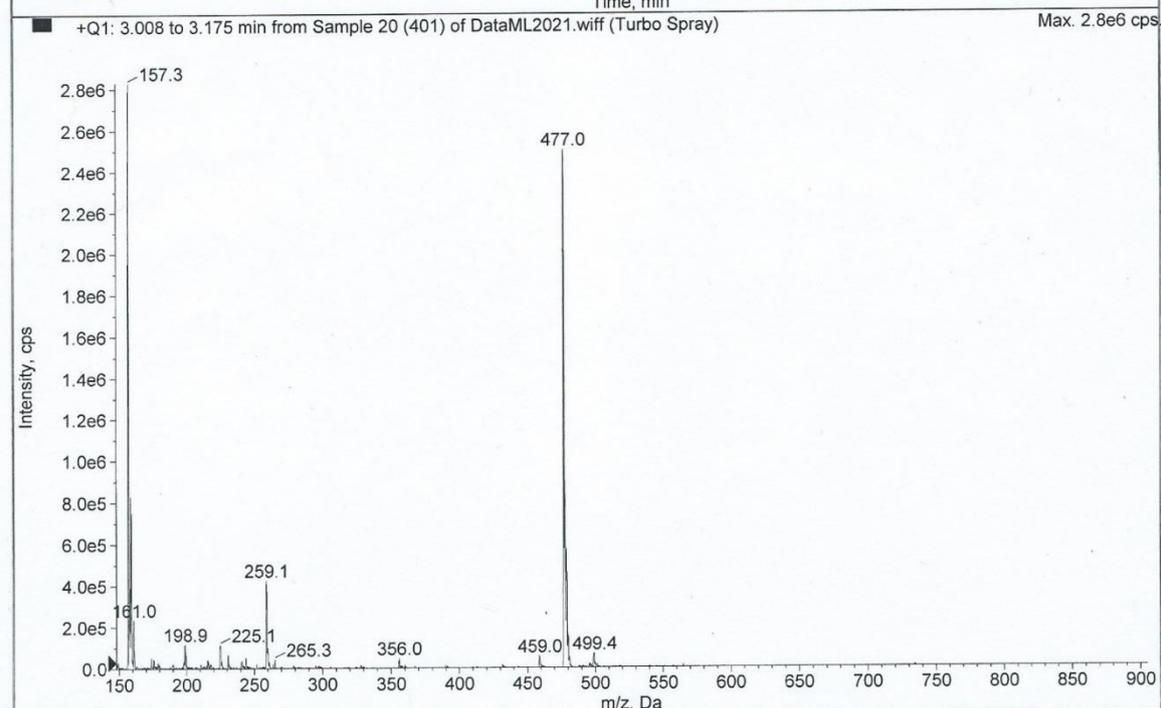
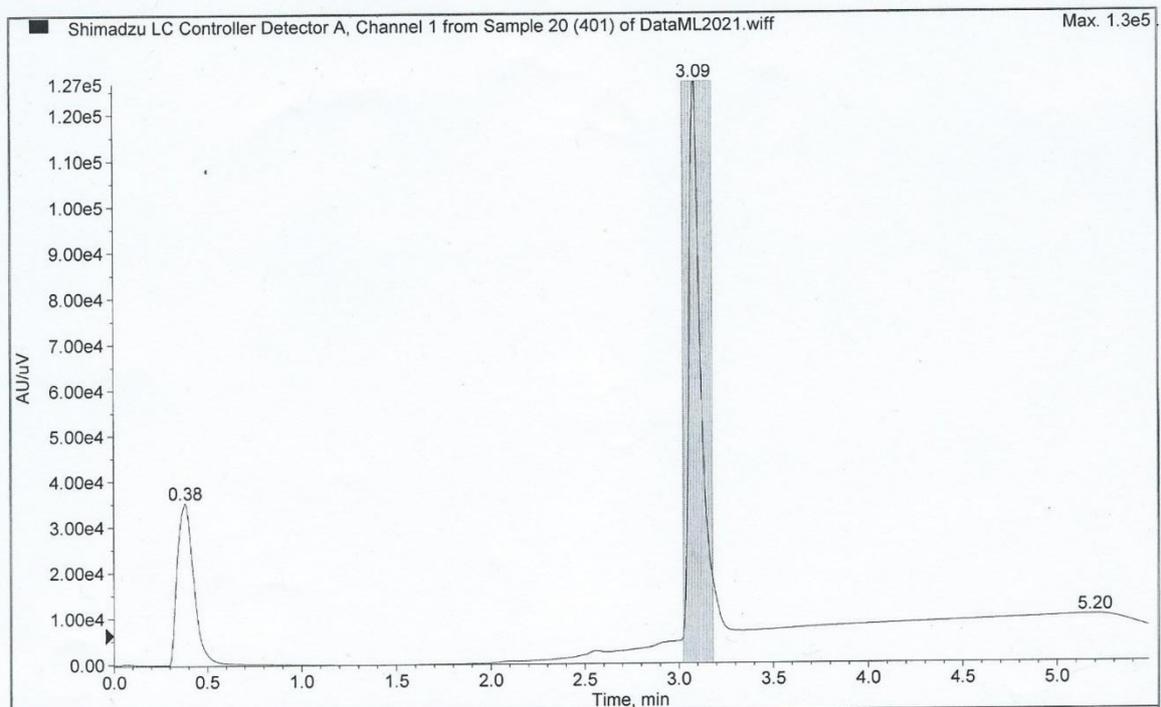**Figure 58:** ¹H NMR spectra of compound **78** (400 MHz, dimethylsulfoxide-d6).

Figure 59: LCMS result for compound **78** with a retention time of 3.09 min ($C_{21}H_{25}N_4O_3S_3^+$: 477.1, observed: 477.0).