

Strategies Towards the Development of Guanidinophenyl Piperazine Derivatives as Potential Anti-Tubercular Agents



A thesis presented to Trinity College Dublin, University of Dublin for the
degree of Master in Science

By

Matas Misevicius, B.A. (Mod.)

Under the supervision of Professor Isabel Rozas

3rd July 2026

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I confirm that the work has been completed in full compliance with relevant university policies, including the Academic Integrity policy, the Trinity Policy on Good Research Practice, and the guidance on AI and GenAI in Teaching, Learning, Assessment and Research.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

A handwritten signature in dark ink, reading "Matas Misevicius". The signature is written in a cursive style with a long horizontal stroke at the end.

Matas Misevicius

Acknowledgements

First and foremost, I wish to thank my supervisor Prof. Isabel Rozas. Words cannot describe what an amazing person she is, but I will try anyways. She has been a constant source of support, inspiration and guidance. Truly, there was no one else I would rather have as a PI. She encouraged me to push myself and be the best scientist I can be even when I could not do it myself. From the bottom of my heart, I will be eternally grateful. Thank you for taking a chance on me.

Dr. Marco Minecci, I could not have done these past two years without you. You always lent your ear to me when I was stumped on organic synthesis and working alongside you always made me want to push myself further. But most of all, thank you for your friendship. I think the three of us made an amazing team.

I would like to thank our collaborators overseas who handled the biological aspects of this project; Prof. Sylvie Garneau-Tsodikova and Dr. Nishad Thamban Chandrika in the University of Kentucky, USA and to Prof. Anna Hirsch and Dr. Andraes Kany in the University of Saarland, Germany. Additionally, I would like to thank the TCD School of Chemistry staff, especially Dr. John O'Brien for his timely NMR assistance, Dr. Gary Hessman for sometimes running 20 samples at a time and to Dr. Manuel Ruether for when the NMR machine breaks.

I would like to also thank all my friends who supported me but especially James and Caroline for keeping me sane. You guys listened to my machinations and supported me throughout. From writing (and winning) grants together with James to science communication and 'educating the masses' with Caroline, you have both made this masters that much more formative and memorable. I know you will be amazing chemists in the future!

Finally, I would like to thank my family in Ireland, my mother Marija and her partner Brian for always telling me to strive to do more and supporting me when I attempted to do so. And to my uncles, Daz and Ciaran who were a source of levity throughout and for letting me crash in Dublin on those late nights in the office.

Thank you all!

Abstract

This thesis presents the development and evaluation of guanidinophenyl piperazine derivatives as novel anti-tubercular agents targeting enoyl-acyl carrier reductase (InhA), a key enzyme in mycolic acid synthesis of *Mycobacterium tuberculosis*. In 2023, according to the World Health Organisation (WHO) Global Tuberculosis Report, an estimated 10.8 million people fell ill with TB while a further 1.25 million people died as a result of the disease. With complicated treatment regimes and anti-biotic resistance on the rise there is an urgent need for new tuberculosis therapies. Hence, this project deals with the further development of **1**, a trifluoromethyl substituted guanidinophenyl piperazine which has shown to be active in *Mycobacterium smegmatis*.

Beginning with computational studies, bioisosteres were identified for the trifluoromethyl functional group due to recent concerns over its negative environmental effects. Potential bioisosteres were considered through the systematic comparison of size, electronic properties, and lipophilicity. Molecular docking investigations, including a flexible side-chain approach, revealed novel binding modes for these guanidino piperazine derivatives in the InhA active site, incorporating hydrogen bond and π -cation interactions with the NADH cofactor and InhA which could be explain the activity of these compounds.

New derivatives were synthesised and thusly evaluated as potential anti-tubercular agents. Biological evaluation against *M. tuberculosis* strain mc2 6230 identified derivative **7** (-SF₅) and derivative **21** (alkylguanidine) as the most promising candidates with MIC values of 4-8 $\mu\text{g mL}^{-1}$ and 8 $\mu\text{g/mL}^{-1}$, respectively. However, preliminary ADMET studies revealed hepatotoxicity concerns for these compounds, with selectivity indices below acceptable thresholds. Thusly it is proposed parent compound **1** (-CF₃) should be the focus of further development as it demonstrated a balance between moderate anti-tubercular activity (32 $\mu\text{g mL}^{-1}$) and an excellent safety profile. Thusly, given the molecular modelling results, new avenues of further development have been identified.

Abbreviations

Boc:	<i>tert</i> -Butyloxycarbonyl
DCM:	Dichloromethane
DFT:	Density Functional Theory
DMSO:	Dimethylsulfoxide
DNA:	Deoxyribonucleic acid
Et ₂ O:	Diethyl Ether
EtOH:	Ethanol
EtOAc:	Ethyl Acetate
ESI :	Electrospray Ionisation
FAME:	Fatty Acid Methyl Ester
FDA:	Food and Drug Administration
HB:	Hydrogen Bond
HCl:	Hydrochloric acid
HgCl ₂ :	Mercury (II) chloride
HPLC:	High-Pressure Liquid Chromatography
HRMS:	High-Resolution Mass Spectrometry
HSQC:	Heteronuclear Single Quantum Coherence Spectroscopy
IC ₅₀ :	Half Maximal Inhibitory Concentration
InhA:	Enoyl-acyl carrier reductase
J:	Coupling Constant
MAME:	Mycolic Acid Methyl Ester
MDR:	Multi-Drug Resistance
MeCN:	Acetonitrile

MeOH:	Methanol
MIC:	Minimum Inhibitory Concentration
mmol:	Millimole
NMR:	Nuclear Magnetic Resonance
RT:	Room Temperature
SI:	Selectivity Index
THF:	Tetrahydrofuran
TFA:	Trifluoroacetic acid
TLC:	Thin Layer Chromatography
WHO:	World Health Organisation
XDR:	Extensively-Drug Resistance

Contents

Chapter 1 – Introduction	1
1.1 Overview of Mycobacterium Tuberculosis	1
1.1.1 Epidemiology and Burden	1
1.1.2 Pathogenesis	2
1.1.3 Diagnosis and Current Treatment Regimens	5
1.1.4 Current Drug Development Efforts and Targets	8
1.2 Previous Work within the Rozas Group	14
1.2.1 Guanidine-based Anti-Mycobacterial Agents	14
Chapter 2 – Objectives	17
2.1 Objective Overview	17
Chapter 3 – Guanidine-based Derivatives as Antitubercular Agents	18
3.1 Searching for Greener Bioisosteres of Trifluoromethyl	18
3.1.1 Comparison of Size	21
3.1.2 Comparison of Electronics	22
3.1.3 Comparison of Lipophilicity	24
3.1.4 Conclusion – Bioisosteres of -CF ₃	28
3.1.5 A Note on ‘Greenness’	30
3.2 Molecular Modelling and Theoretical ADMET Properties	30
3.2.1 InhA Enzyme and Considerations of its Active Site	31
3.2.2 Binding Pose Refinement – Building on Previous Work	33
3.2.3 Docking Investigations of Novel Derivatives	44
3.2.4 Theoretical Pharmacokinetic Profile Assessment	58
3.2.5 Evaluation of ‘Lipinski-Ghose-Veber’ Profile of Proposed Derivatives	59
3.2.6 Evaluation of the Theoretical Pharmacokinetics of Proposed Derivatives	61
3.3 Chemical Synthesis	65
3.3.1 Introduction and Retrosynthetic Analysis	65
3.3.2 Synthesis of Lipophilic Pocket Derivatives	67
3.3.2.1 Synthesis of 1-Phenylpiperazine Derivatives	67
3.3.2.2 Synthesis of 1-(4-nitrophenyl)-4-Phenylpiperazine Derivatives	76
3.3.2.3 Synthesis of 4-(4-Phenylpiperazin-1-yl)Aniline Derivatives	78
3.3.2.4 Guanidylation of 4-(4-Phenylpiperazin-1-yl)Aniline Derivatives	82
3.3.3 Synthesis of Alkylguanidine Derivatives	85
3.3.3.1 Synthesis of 4-(4-(4-(trifluoromethyl)phenyl)piperazin-1-yl)Aniline 70	85
3.3.3.2 Synthesis of 1-Alkyl-3-(4-(4-(4-(trifluoromethyl)phenyl)piperazin-1-yl)phenyl)guanidines	85

3.3.4 Synthesis of Triazole Derivatives	89
3.3.4.1 Synthesis of 1-Phenyl-4-(prop-2-yn-1-yl)piperazine Derivatives	89
3.3.4.2 Synthesis of tert-Butyl (2-(4-((4-Phenylpiperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)carbamate Derivatives	90
3.3.4.3 Synthesis of 1-(2-(4-((4-Phenylpiperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl) guanidine Derivatives	92
3.4 Biological Testing	95
3.4.1 Anti-Tubercular Biological Evaluation	95
3.4.2 Experimental Determination of ADMET Properties	97
3.4.3 Commentary on Theoretical vs. Experimental Determination of ADMET Properties	100
Chapter 4 – Conclusions and Future Work.....	102
4.1 Conclusions.....	102
4.1.1 Computational Studies	102
4.1.2 Chemical Synthesis	103
4.1.3 Biological Testing.....	103
4.2 Future Work	104
4.2.1 Expanded Computational Studies	104
4.2.2 Design and Chemical Synthesis	105
4.2.3 Biological Development	106
Chapter 5 – Experimental Methods and Characterisation	107
5.1 Computational Methods	107
5.1.1 DFT Methods.....	107
5.1.2 Docking Methodology	107
5.2 Synthetic Chemistry.....	108
5.2.1 Materials and Methods	108
5.2.2 Synthesis and Characterisation.....	111
5.2.3 HPLC Purity Assessment	151
5.3 Biological and Biochemical Methods.....	152
5.3.1 MIC Determination Against <i>M. tuberculosis</i> H37Rv mc ² 6230.	152
5.3.2 Lipophilicity Determination	152
5.3.3 Metabolic Stability in Mouse Liver Microsomes	153
5.3.4 Stability in Mouse Plasma.....	153
5.3.5 Cytotoxicity.....	153
References	155
Appendix	164

Chapter 1 - Introduction

1.1 Overview of Mycobacterium Tuberculosis

1.1.1 Epidemiology and Burden

Tuberculosis (TB) is an airborne disease caused by the bacillus *Mycobacterium tuberculosis* (*M. tb.*) and has plagued human societies since antiquity.¹ In 2023, according to the World Health Organisation (WHO) Global Tuberculosis Report, an estimated 10.8 million people fell ill with TB while a further 1.25 million people died as a result of the disease.² TB was marked as the 10th leading cause of death worldwide in 2021 and is the leading cause of death by a single infectious agent year on year post the COVID-19 pandemic.³ While TB is a global disease, it is distributed heterogeneously and especially pervasive in the South-East Asian and African WHO regions with eight countries accounting for two-thirds of new global cases (*Fig. 1*). Overall, a shocking 2 billion people worldwide are infected with TB (*i.e.* roughly a quarter of the world's population), however in a latent, inactive state (LTB).²

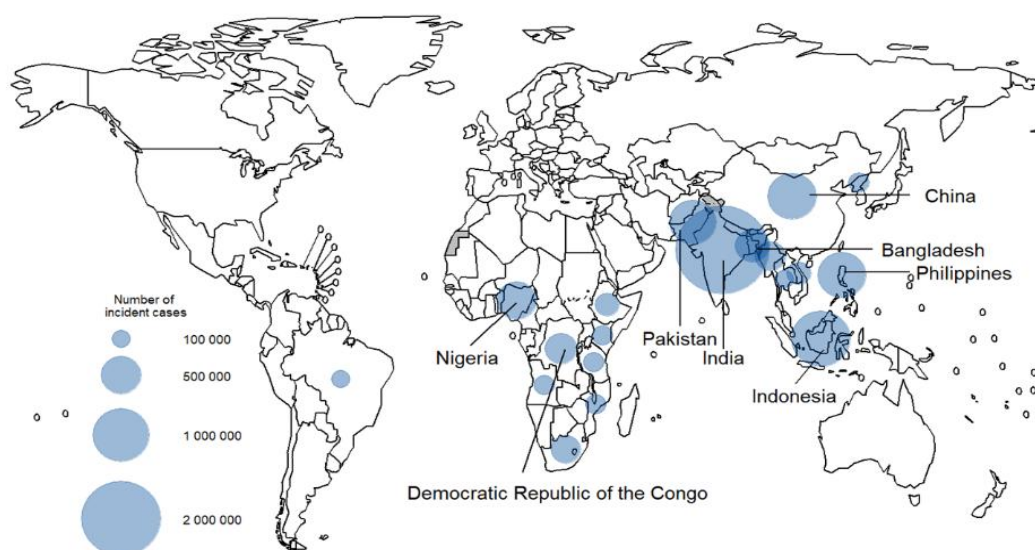


Figure 1 - Estimated number of TB cases in 2023, for countries with at least 100,000 incident cases. Reproduced from Ref. 2.

People with LTB are asymptomatic and the disease is non-transmissible; however, there is always an underlying risk that this state can progress into an active and transmissible one, particularly in the context of immunocompromised individuals, *i.e.*

co-infection with HIV.⁴ Moreover, there is a burgeoning public health challenge with the advent of multi-drug resistant or rifampicin-resistant TB (MDR/RR-TB) and extensively drug resistant TB (XDR-TB) worldwide, particularly in Eastern Europe (Russia and Ukraine), Asia and sub-Saharan Africa.⁵ In 2023, an estimated 400,000 people developed MDR/RR-TB;² notwithstanding the complications and difficulties in treating the disease for patients by themselves (discussed later), it is estimated that DR-TB (Drug-resistant TB, all profiles) will cost the global economy US\$16.7 trillion between 2015 and 2050 which will account for ~20-25% of total estimated global cost due to anti-microbial resistance overall.⁶ Regardless of the good progress in the fight against TB, it can be seen that it is far from being eradicated and hence there is an urgent need for new therapies to address this global health concern. Therefore, the focus of this work is the design, synthesis and biological evaluation of potential drug candidates that may be used as novel anti-tubercular agents.

1.1.2 Pathogenesis

The etiologic agent of TB, *Mycobacterium tuberculosis* was first successfully isolated by Robert Koch in 1882.⁷ *M. tb.* is a gram-positive, slow growing bacterium with a highly complex lipid cell wall unique to mycobacteria, which plays a major role in their pathogenesis and survival within the host organism.^{8, 9} While *M. tb.* is a human infectious agent, the study of related mycobacteria such as *M. bovis* BCG and *M. smegmatis* (both animal infectious agents) as model systems has helped greatly in understanding the mechanisms underlying the disease.¹⁰

TB is primarily a pulmonary disease and patients with pulmonary TB present with symptoms such as persistent cough and haemoptysis (coughing up blood – in advanced disease). However, it can affect the entire body (extrapulmonary) and thus

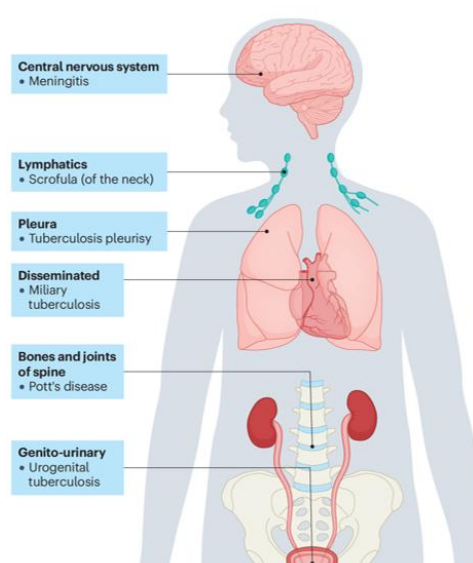


Figure 2 - Pulmonary and extrapulmonary infection sites of *M. tb.* Reproduced from Ref 11.¹

patients with active TB often present with symptoms such as fever, lack of appetite and weight loss which can make diagnosis challenging (Fig. 2).¹¹

The life cycle of *M. tb.* can be seen in Fig. 3.⁵ For initial infection a human must deeply inhale airborne *M. tb.* (1) which would be passed on by active TB patients (6), airborne *M. tb.* is capable of remaining as such for prolonged periods of time, increasing risk of infection.^{12, 13} A person may be able to eliminate the infection through innate or adaptive immune mechanisms, firstly alveolar macrophages that engulf the bacteria and translocate it across the epithelium and into interstitial tissue, from here it can be disseminated to other organ systems *via* the lymphatic system.¹⁴ From there, numerous immune responses take place and in particular macrophages, CD4+ T lymphocytes and granuloma formation creates the pillars of defence against *M. tb.*¹⁵ However, about two-thirds of patients may rapidly progress from TB infection (3) to active disease (6) within ~2-18 months.¹⁶ TB disease can present as a dynamic spectrum hence the stages in between can pragmatically categorised as LTB as patients are asymptomatic [stages (4) and (6)].⁴

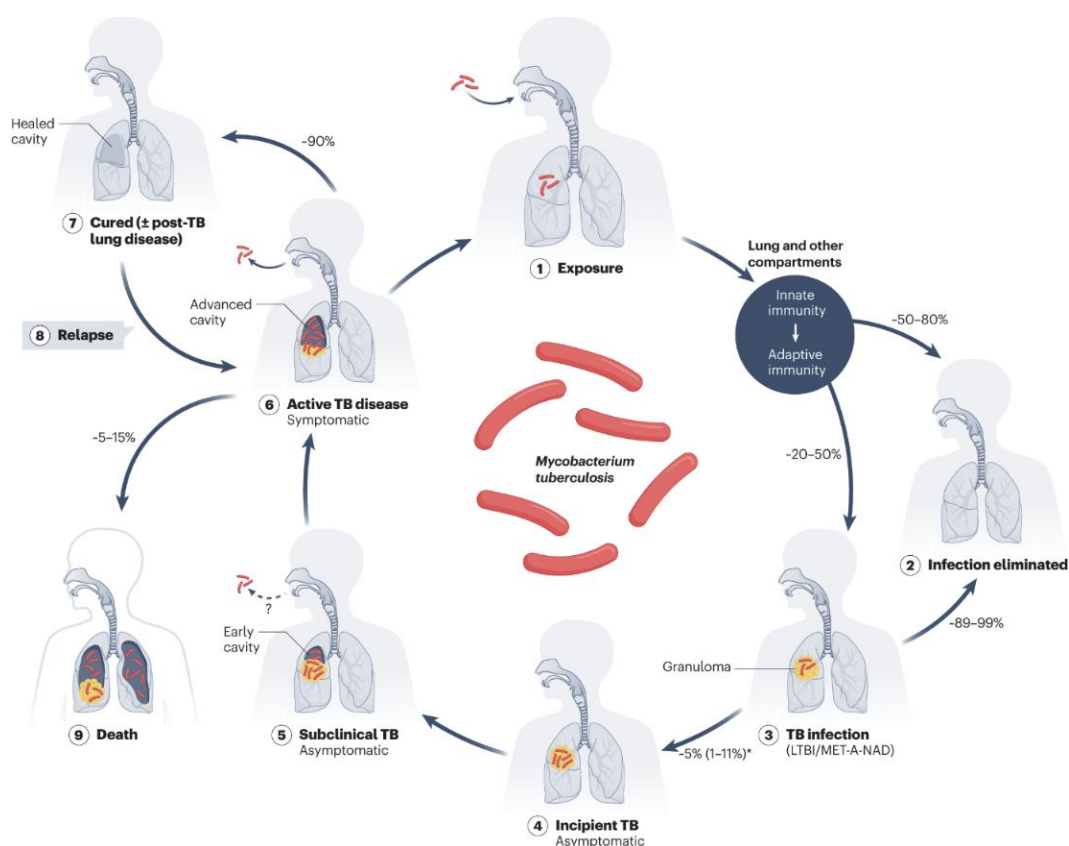


Figure 3 - The life cycle of the *M. tb.* Bacterium. Reproduced from Ref. 5.

M. tb. often establishes this asymptomatic latent transmission, in fact only about 5-10% of patients with LTB will progress to an active TB state.¹⁷ The reason for this is the aforementioned granuloma, a particular feature of *M. tb.* infection. During intracellular survival as the bacteria resists the immune response, engulfment by macrophages triggers innate immune signalling pathways, which leads to the recruitment of other immune cells (*i.e.* more macrophages, monocytes, lymphocytes etc.) whereby they organise in a spherical shape, surrounded by collagen and fibrinogen to seal the *M. tb.* and prevent progression to active disease (3 and 4a, Fig. 4).^{18, 19} The success of this for the host can be seen in calcified sterile granulomas in both human patients and animal models.²⁰ However, although it intended to restrain *M. tb.*, it has been shown that some bacilli can survive inside these structures albeit in a dormant state and while originally thought of as a purely host driven response, it has been shown that orderly formation of granulomas require the concerted action of both the host and bacilli.^{21, 22}

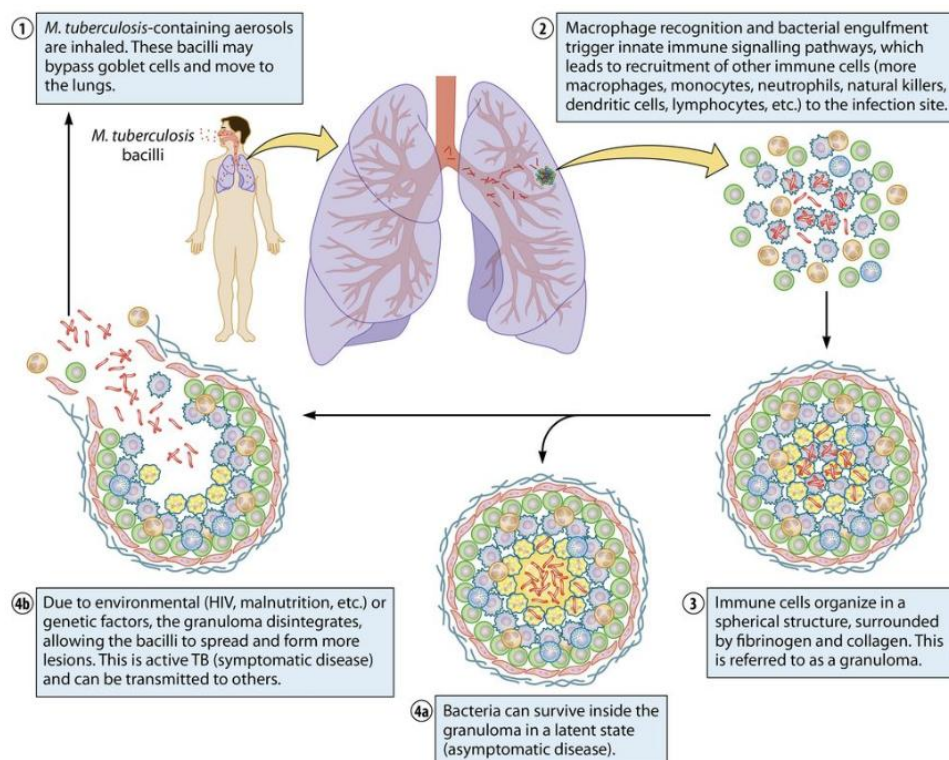


Figure 4 - The pathophysiology of *M. tuberculosis* infection: (1) initial infection, (2) immune response, (3) granuloma formation, (4a) LTB state of granuloma and (4b) granuloma disintegration active TB. Reproduced from Ref. 23.

The sequestered bacilli can be released due to the necrosis in the centre of the granuloma which leads to the formation of a late-stage caseous granuloma.²⁴ This causes the loss of its rigidity and subsequent rupturing releasing the bacilli while at the same time forming cavities in the lung wall thereby causing damage to lung tissues and the reactivation of the disease into active TB.²⁵ HIV co-infection is usually considered the primary reason of TB reactivation but other environmental reasons include malnutrition, chemotherapy, sepsis, addiction and more.²⁶⁻²⁸

1.1.3 Diagnosis and Current Treatment Regimens

A major challenge in the fight against TB has been the timely diagnosis of disease which has historically been difficult.²⁹ Nowadays, there are a wide range of diagnostics available including but not limited to: sputum smear microscopy, chest radiography, culture-based diagnostics, relatively new rapid molecular diagnostics such as nucleic acid amplification tests which detect mycobacterial DNA directly. For those with LTB specifically, immunological approaches have been developed such as interferon-gamma release assays or the tuberculin skin test which measure the host's cellular immune response to mycobacterial antigens..³⁰ Of course, some diagnostics are more expensive and resource intensive than others and hence are not as widely available in all countries.^{31, 32}

The treatment of TB depends directly on the strain of *M. tb.* and hence determines the antibiotics used which categorised hierarchically as first, second and sometimes “third line” drugs (sometimes grouped together with second line drugs), some of which can be seen in *Fig. 5*. Current treatments require the use of multiple antibiotics given to patients (*i.e.* drug cocktail) and while effective, must be administered for several months.³³

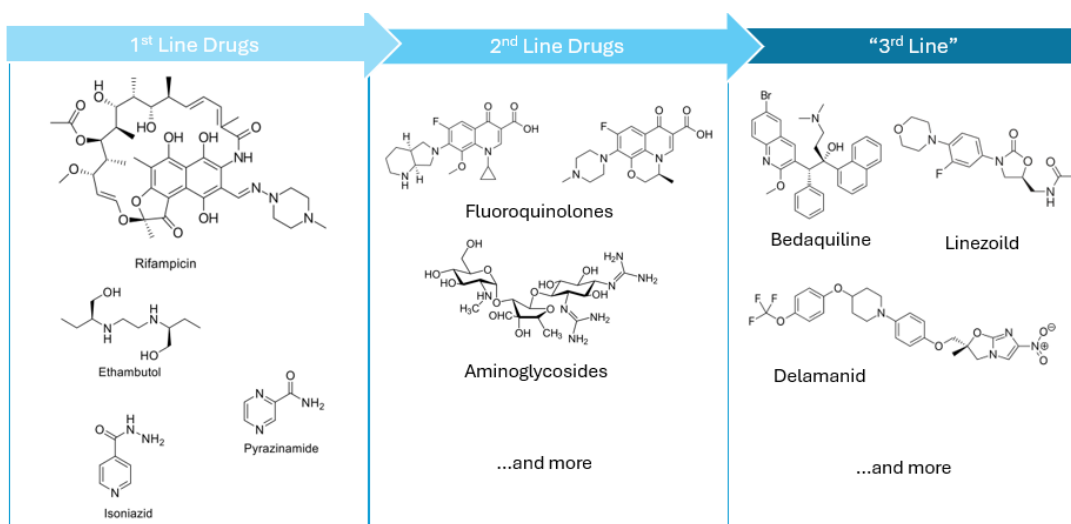


Figure 5 - 1st, 2nd and "3rd" Line anti-tubercular agents.

The treatments given to deal with the different strains of *M. tb.* as recommended by the WHO are outlined in Table 1.³⁴ The standard treatment for DS-TB for adults involves a 2 month intensive phase of isoniazid (H or INH), rifampicin (R or RIF), pyrazinamide (Z) and ethambutol (E) followed by a 4 months of rifampicin and isoniazid (2HRZE/4HR).

Table 1 - Recommended treatment regimens by the WHO for Tuberculosis.³⁴

Regimen	<i>M. tb.</i> Strain	Drug Cocktail Composition	Notes
2HRZE/4HR	DS-TB	Isoniazid, rifampicin, pyrazinamide and ethambutol for 2 months followed by isoniazid and rifampicin for 4 months (can extend up to 9 months).	HIV patients' treatment regimens need to be adjusted.
2HPMZ/2HPM	DS-TB	Isoniazid, rifapentine, moxifloxacin and pyrazinamide for 2 months followed by isoniazid, rifapentine and moxifloxacin for 2 months.	
BPaLM/BPaL	MDR/RR-TB or pre-XDR-TB	Bedaquiline, pretomanid, linezolid (and moxifloxacin) for 6 months.	
BDLLfxC/BDLLfx/BDLC	MDR/RR-TB or pre-XDR-TB	Bedaquiline, delamanid, linezolid, levofloxacin and clofazimine for 6 months.	Children/Breastfeeding Women
BLMZ, BLLfxCZ & BDLLfxZ	MDR/RR-TB with no fluoroquinolone resistance	Comprise bedaquiline in various combinations with levofloxacin/ moxifloxacin, linezolid, clofazimine, delamanid and pyrazinamide for 9 months.	

Individualised	XDR-TB	Last resort regimens for those who failed previous regimens. Mix of first, second and third line drugs.
6(H)RZE-Lfx	Isoniazid resistant but rifampicin susceptible.	rifampicin, ethambutol, pyrazinamide and levofloxacin for 6 months

However, the treatment for DS-TB can extend even up to 9 months and, hence, the medication adherence is a challenge which heightens the risk of treatment failure and in particular encourages the development of resistance.³⁵ Further complicating matters in this lengthy treatment regimens are the adverse side effects, which while may not present in every patient, can make treatment more difficult for some; side effects include: hepatitis, cutaneous reactions, gastrointestinal intolerance, haematological reactions and renal failure.^{36, 37} Therefore, the standard practice is to carry out directly observed therapy where patient compliance and monitoring of adverse reactions can be surveyed over the course of a treatment regimen.³⁸ Another complication involves the various drug-drug interactions *i.e.* one medication affects another, reducing efficacy of both in this case.

While there are many recorded interactions of the rifampicin and isoniazid with various central nervous system and cardiovascular drugs, particular care is taken in patients with a HIV co-infection.³⁹ TB is the leading cause of death and hospitalisation of people with HIV globally. In 2023, HIV-associated TB patients had a mortality rate of 24% compared to 11% of those without HIV.² Even when HIV affected individuals are on antiretroviral therapy (ART), the risk of developing active TB can increase anywhere between 2-20 times depending on the progression of HIV, likewise TB has been shown to increase levels of HIV-1 replication, propagation and genetic diversity; both diseases synergistically contribute to each other's progression.⁴⁰ For example, rifampicin has been shown to reduce the blood plasma concentration of concomitantly administered non-nucleoside reverse transcriptase inhibitors (NNRTIs) and protease inhibitors (PIs) resulting in inadequate ART plasma concentrations and inferior ART outcomes, most likely through its effects on cytochrome P450 isoenzymes and drug efflux pump P-glycoprotein (P-gp).⁴¹

Therefore, for successful treatments, TB and anti-retroviral therapies need to be well integrated with each other.⁴²

Resistance to the standard treatment regimen (*i.e.* the development of MDR-TB) has been clinically shown to begin with resistance to isoniazid (INH).⁴³ The primary reason thought for this is patient non-compliance.⁴⁴ However, it has been shown that even in patients with >60% non-compliance, only 1% develop INH-resistance, while perfect adherence still leads to isoniazid resistance in 1% of those patients due to the pharmacokinetic variability of the drug.⁴⁵ Hence, it can be assumed that this resistance results from a mix of factors. Rifampicin (RIF) resistance can subsequently emerge if a patient has INH-resistant TB. However since INH-resistance is not routinely tested for, many patients go undiagnosed and hence develop RIF-resistance also.^{5, 46} The recommended treatment for MDR-TB, until recently, was at least 18 months with a combination of 4 active drugs, often being injectables (levofloxacin, kanamycin, moxifloxacin); this regimen did not achieve success in more than 60% of the cases,⁴⁷ and was also the regime with the highest incidence of drug adverse effects.⁴⁸ The current recommendation, after a highly successful trial, is a combination of bedaquiline, pretomanid, and linezolid (BPaL) for 6 months with a high success in patients with highly drug resistant TB.⁴⁹ In a way, a revolutionising development for patient outcomes.

As previously mentioned, about 400,000 people developed MDR/RR-TB and, while it only represents about ~5% of global TB burden, evidence shows that resistances to all first line and second line drugs is on the rise.^{2, 50} Hence, there is an ever pressing need for new therapies to fight against TB and overcome resistance.

1.1.4 Current Drug Development Efforts and Targets

Historically the pipeline for novel anti-tubercular agents remained stagnant for nearly four decades until the approval of bedaquiline in 2012 by the FDA. This marked the first novel anti-tubercular approved in 40 years(*Fig. 6*).⁵¹

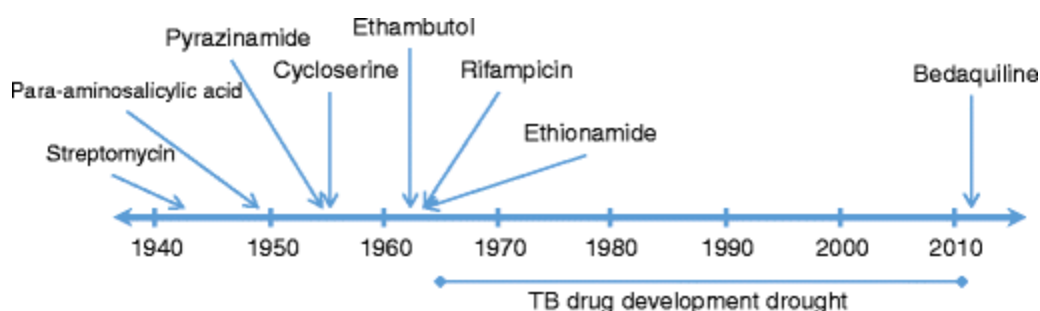


Figure 6 - Timeline for new TB drug development 1940 to 2012. Reproduced from Ref. 52.

It is widely accepted that the treatment of TB requires multiple compounds working in concert targeting various proteins and cellular processes within *M. tb.*; hence, the goal of TB drug discovery is to ultimately create shorter and less toxic treatment regimens.⁵³ The need for this combination therapy is due to several reasons, outlined in Table 2: heterogeneity of bacterial populations, its thick waxy cell wall and slow growth.

Table 2 - Overview of challenges in anti-TB drug discovery

Challenge	Explanation
Heterogeneity of bacterial populations ^{54, 55}	<i>M. tb.</i> has a remarkable ability to survive host immune attack and drug treatment. <i>M. tb.</i> subpopulations can vary wildly in their lesion structure, metabolic activity and drug penetrance. They exhibit asymmetric growth on a rapid time scale and phenotypic heterogeneity between generations. Environmental heterogeneity in the form of granuloma composition and acid microenvironment pH influence growth rate and metabolism.
Consequence	All together these factors influence drug susceptibility, therefore a cocktail of drugs is needed to effectively kill the bacteria.
Cell Wall	The cell wall of <i>M. tb.</i> accounts for around 60% of its mass, composed of complex lipids which give it a high degree of hydrophobicity (Fig. 7). ⁵⁶ The plasma membrane is surrounded by a layer of peptidoglycan (highly cross linked) covalently bonded to arabinogalactan polysaccharide which then ends with long mycolic acid chains (fatty acids). ⁵⁷
Consequence	This makes the bacterium inherently resistance to many drugs (particularly large, hydrophilic drugs) negatively affecting drug penetration and therefore effectiveness. ⁵⁸
Slow growth of bacterium and safety restrictions	<i>M. tb.</i> is a slow growing bacterium requiring biosafety level three in laboratories handling it. ⁵⁹
Consequence	Confer drug tolerance and impediment of fast drug discovery efforts i.e. time consuming laboratory studies.

Current treatment regimens target less than 0.5% of bacterial proteins. Over the years many molecular targets that are essential to the growth and survival of *M. tb.* have been considered in anti-TB drug discovery efforts and there are several ‘battlefronts’ that are seen in the literature: bioenergetics & metabolism, cell wall biosynthesis, DNA/RNA and protein synthesis, redox disruption and, the relatively new, host-directed therapies.⁶⁰⁻⁶² Inhibition of cell wall biosynthesis will be discussed in detail as it represents the highest number of candidates in the drug development pipeline and is the strategy of interest to this project.

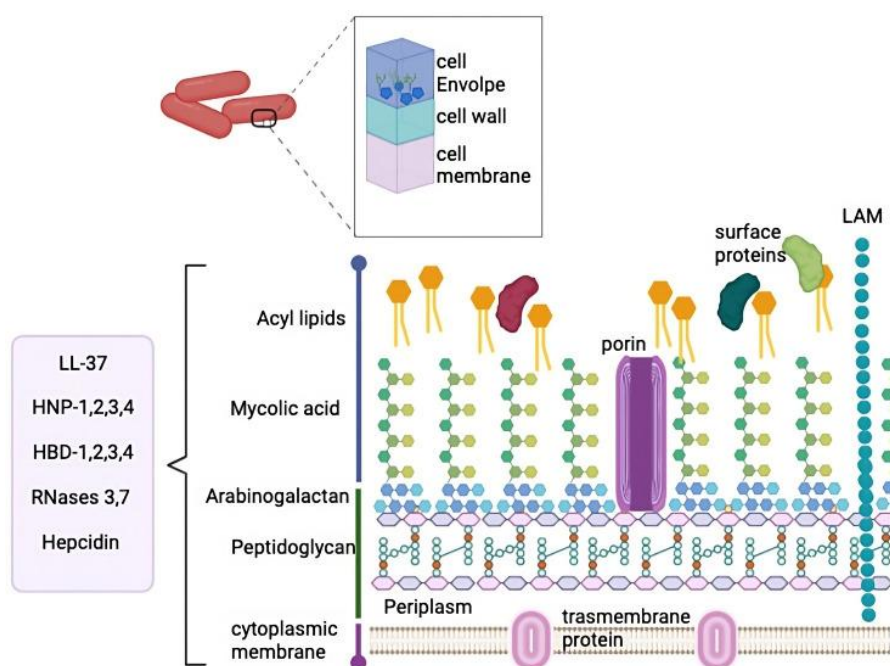


Figure 7 - *M. tuberculosis* cell wall composed of several complex layers. Reproduced from Ref. 57.

Currently approved drugs targeting cell wall biosynthesis can be seen in Table 3 along with their respective targets.⁶³ As can be seen, there has been a dearth of approved drugs targeting cell wall biosynthesis with the most recent one being delamanid in 2014. Notably, there has been a consistent effort to develop orally available drugs.

Table 3 - Approved anti-tubercular drugs targeting cell wall biosynthesis of *M. tb*.

Drug Name	Therapeutic Target	Route of Administration	Year of Approval
Isoniazid	Enoyl-acyl carrier (ACP) reductase (InhA)	Oral / Intramuscular / Intravenous	1953
Thioamides (Ethionamide)	InhA	Oral / Intramuscular / Intravenous	1968
Ethambutol	Arabinosyl transferase enzyme B and C.	Oral	1968
Cycloserine	D-alanine racemase (Alr) and D-alanine-D-alanine ligase (Ddl).	Oral	1968
Nitroimidazoles (Delamanid)	DprE2 enzyme	Oral	2014

Current drugs inhibiting cell wall synthesis that are in phase II or pre-phase III clinical trials (although there are many more preclinical candidates) are shown in *Fig. 8* grouped by their respective targets. Alpibectir, which is a spiroisoxazoline first discovered in 2017 as part of a family Small Molecules Aborting Resistance or SMART, works by inducing an alternative bioactivation in *M. tb*. of the drug ethionamide which targets enoyl-acyl carrier reductase (InhA), an enzyme involved in mycolic acid synthesis. In this way, alpibectir has been shown to completely reverse resistance to ethionamide hence the name SMART.⁶⁴ In March of 2025, a trial investigating the effectiveness of alpibectir-ethionamide in combination with other 1st line drugs in pulmonary TB patients entered Phase II clinical trials run by BioVersys and GSK.⁶⁵

Quabodepistat, BTZ-043, PBT169 and TBA-7371 represent a class of cell wall biosynthesis inhibitors that target decaprenylphosphoryl- β -D-ribose-2'-epimerase (DprE1). DprE1 is a flavoenzyme that catalyses the two-step epimerisation of decaprenyl-phospho-ribose (DPR) to decaprenyl-phospho-arabinose (DPA), the precursor for arabinogalactan (an essential polysaccharide in the composition of the mycobacterial cell wall) working in concert with a colocalised enzyme, DprE2.⁶⁶ In the mid-2010s, this target was seen as promising and indeed it has been considering the four Phase II drug candidates aiming this enzyme.⁶⁷ Quabodepistat is a 3,4-dihydrocarbostyryl derivative first presented at the American Society for Microbiology

Microbe Conference in 2018 and published in 2020. It showed potent anti-mycobacterial by itself and significant synergistic effects together with delamanid, bedaquiline, or levofloxacin in mouse models.⁶⁸ In 2022, it entered a Phase II trial combined with delamanid and bedaquiline as therapy for DS-TB expected to be completed later this year.⁶⁹ In the case of BTZ-043 and PBT169, which belong to a class of nitro-benzothiazinones, they were first discovered in 2007 to have antibacterial and antifungal activity. In 2009, BTZ-043 was found to have good anti-TB activity in mouse models and was seen as a promising candidate to be included in combination therapies.⁷⁰ In 2023, it entered Phase II in a combination therapy trial with bedaquiline, delamanid, moxifloxacin.⁷¹ TBA-7371 is an azaindole inhibitor of DpreE1 first discovered in 2013 as a non-covalent inhibitor with 'drug-like' properties and shown to be efficacious in mice models.⁷² It completed its early Phase II trial in a 14 day treatment regimen in patients with DS-TB.⁷³

many years, its mechanism of action was not known until 2012 finding that it targets mycobacterial membrane protein large 3 (MmpL3), which is responsible for the export of mycolic acids in the form of trehalose monomycolate to the periplasmic space and therefore formation of the mycobacterial cell wall.^{78, 79}

Overall, it can be seen that targeting cell wall biosynthesis in *M. tb.* is a promising avenue for therapeutic intervention, and accordingly, this is the strategy employed in this work: targeting of InhA and subsequent inhibition of mycolic acid synthesis.

1.2 Previous Work within the Rozas Group

1.2.1 Guanidine-based Anti-Mycobacterial Agents

In 2012, Forge *et al.* published a series of 1,4-diarylpiperazine derivatives which showed anti-tubercular activity (*Fig. 9*).⁸⁰

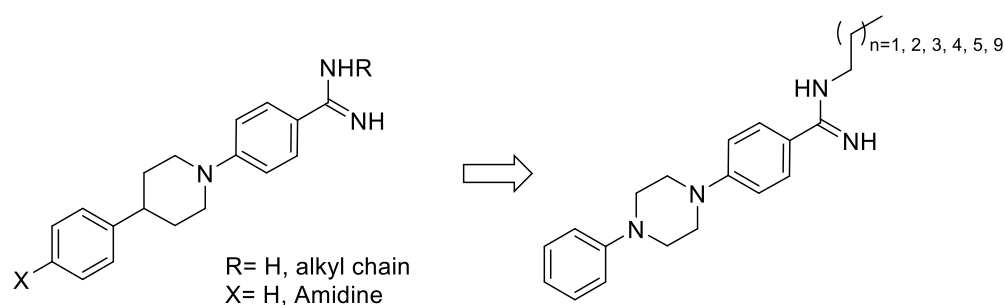


*Figure 9 – Structures 1,4-diarylpiperazine derivatives synthesised by Forge et al. which showed good anti-mycobacterial activity.*⁸⁰

These derivatives showed good anti-mycobacterial activity against drug-sensitive ($MIC_{50} = 0.052 - 0.264 \mu M$) and drug-resistant ($MIC_{50} = 0.95 - 3.85 \mu M$) strains of *M. tb.* with the alkyl-substituted amidines exhibiting the best activity. They also showed good selectivity between *M. tb.* and C3A hepatocytes which would indicate these compounds may avoid one of most common pitfalls of developing drugs which is hepatotoxicity. Furthermore, they were shown to kill *M. tb.* residing in mouse J774A.1 macrophages; indicating their potential as granuloma penetrating drugs.⁸⁰

At the time of this publication, a target had not yet been elucidated, however, their structural similarity to Rozas' in-house library of guanidine-like derivatives previously synthesised prompted investigation of the potential of Families A, B and C as anti-tubercular agents. This led to the discovery of FR20 having very good antimycobacterial activity against *M. smegmatis* (a non-pathogenic bacterium often used as a surrogate model for *M. tb*) as determined by resazurin microtiter assay (REMA) mycobacterial cell viability assays conducted in-house conducted in collaboration with Prof. Stephen Gordon (Conway Institute, UCD, Ireland) (Fig. 10).⁸¹

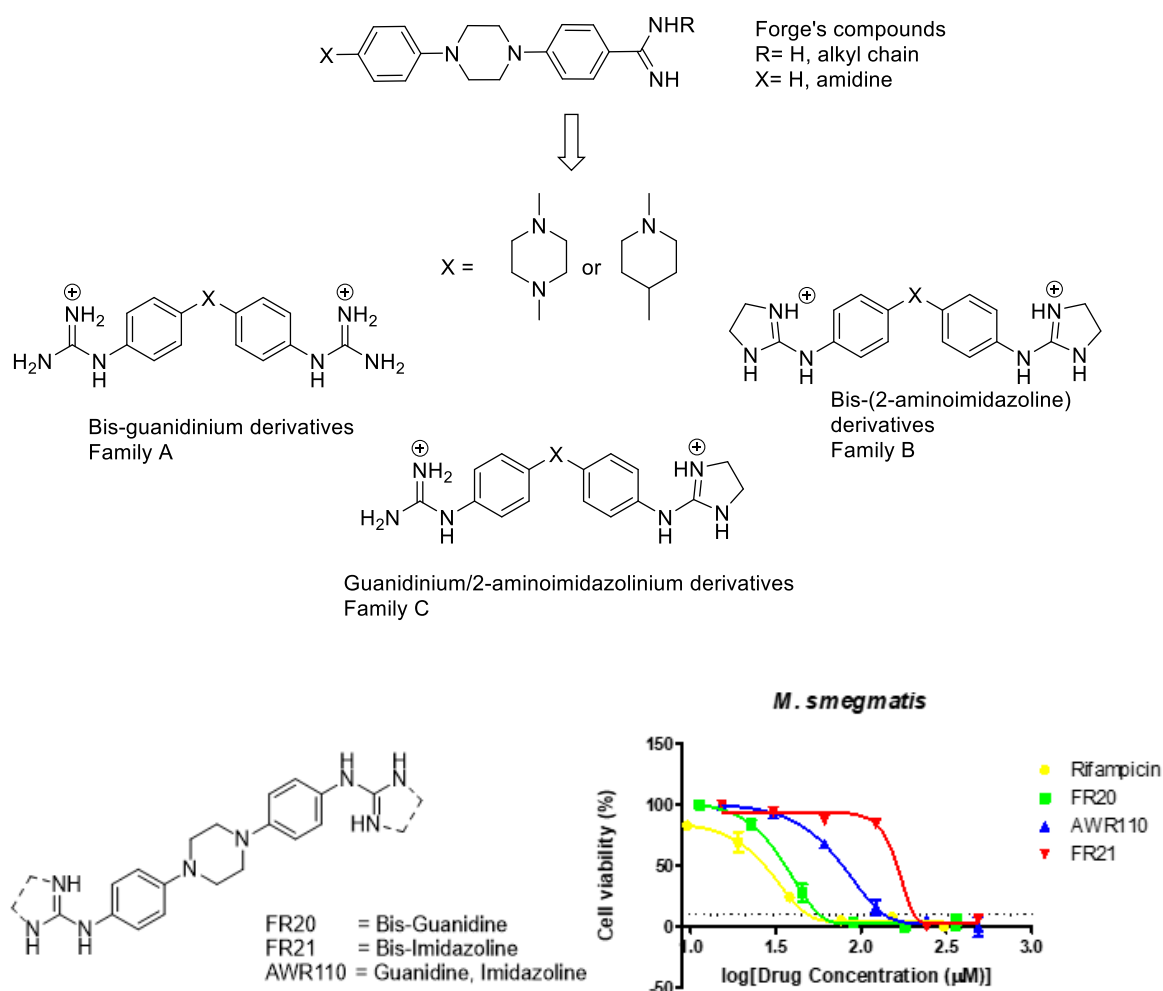


Figure 10 - Structures of in-house Rozas minor groove binders (top), the structures of FR20, FR21 and AWR110 and cell viability curves generated from REMA analysis of our in-house MGBs against *M. smegmatis* (the dotted line represents 10% cell viability mark which corresponds to the MIC).

The positive biological results obtained prompted investigations into the targeted protein. Several proteins were considered as potential targets including the previously mentioned DrpE1, but no inhibition of it was shown. Thus, a screening of pharmacophoric elements pointed the mechanism of action towards inhibition of mycolic acid synthesis and most likely InhA.

Accordingly, in collaboration with Dr. Apoorva Bhatt (University of Birmingham, UK), fatty acyl methyl esters (FAMES) and mycolic acid methyl esters (MAMES) TLC assays were performed to investigate the potential inhibitory effect of compound FR20 on mycolic acid synthesis through InhA inhibition. The inhibition of InhA is characterised by membrane lysis, which results in a reduction in MAMES and an accumulation of FAMES.⁸² Within a short timeframe of exposure to FR20 at its MIC, a reduction in mycolates (MAMES) and accumulation of FAMES was observed, confirming InhA inhibition by FR20.

From here molecular modelling investigations helped to identify the potential interactions of compound FR20 with the active site of InhA finding that an amphiphilic system might be more beneficial to binding. And indeed, this was the case, with compound **1** (72a in Fig. 11), which after being synthesised showed near identical activity to rifampicin in *M. smegmatis*, making it a prime candidate for further development which is the focus of this work.

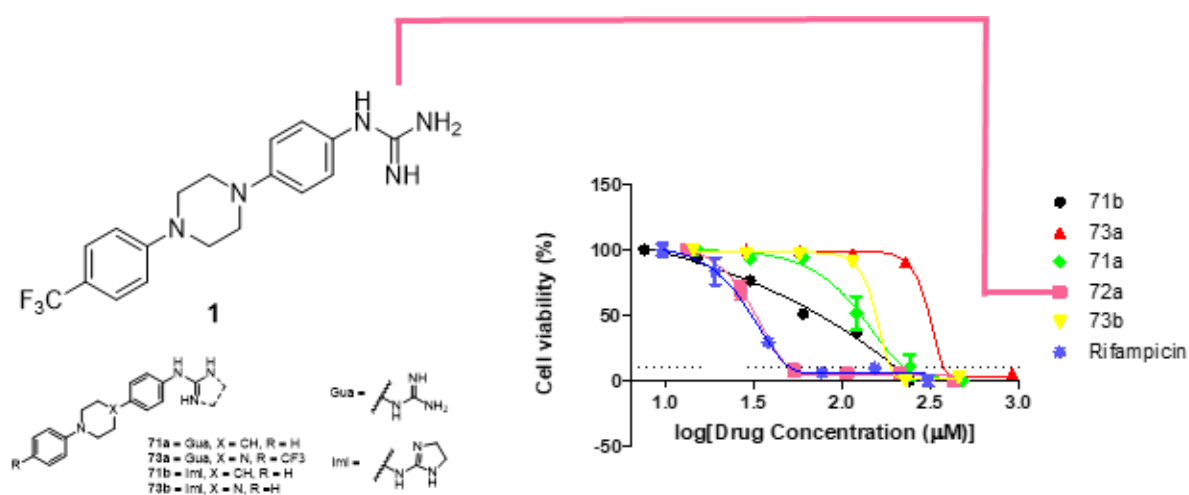


Figure 11 – Structure of ‘lead’ compound **1** and cell viability curves generated from REMA analysis of compounds several amphiphilic derivatives including **1** and rifampicin (the dotted line represents 10% cell viability mark which corresponds to the MIC).

Chapter 2 – Objectives

2.1 Objective Overview

The thesis is focused on the further development of compound **1** as anti-tubercular agent given the promising results obtained against *M. smegmatis*:

- (a) To begin, given the presence of the trifluoromethyl ($-\text{CF}_3$) group in the structure of **1** and given the recent concerns regarding the use of per- and poly-fluorinated alkyl and aromatic substances, greener bioisosteres for $-\text{CF}_3$ will be searched for and in this way inform the new structures of the derivatives of **1**.
- (b) Next, building on the previous computational studies carried out within Rozas' group, several newly proposed derivatives (including novel alkylguanidine and triazole derivatives) will be theoretically investigated for their binding to InhA. *In-silico* docking studies will be used to determine active site affinity of the newly proposed derivatives, and their theoretical ADME properties will also be explored.
- (c) Then, the most promising derivatives will be synthesised following organic transformations utilised in the group and new synthetic pathways will be developed as needed.
- (d) Finally, their effectiveness as anti-tubercular agents will be assessed directly against *M. tb.* *in vitro* in collaboration with Prof. Sylvie Garneau-Tsodikova (University of Kentucky, USA) and their ADMET properties will be experimentally assessed in collaboration with Prof. Anna Hirsh (University of Saarland, Germany).

Chapter 3 – Guanidine-based Derivatives as Antitubercular Agents

3.1 Searching for Greener Bioisosteres of Trifluoromethyl

The first reported use of the trifluoromethyl ($-\text{CF}_3$) functional group in a pharmacological context was back in 1928 where Lehman investigated the effect of nitro and amine substituted benzotrifluorides on the central nervous system of frogs.⁸³ Since then the use of $-\text{CF}_3$ and more broadly fluorine in medicinal chemistry exploded, with about 20% of all FDA drugs from 1991 to 2022 containing a fluorinated structural motif. In fact, in 2018 and 2019 they represented 45% and 41% of FDA approved drugs, respectively;⁸⁴ and the global fluorinated drug market was worth US\$ 182.64 billion in 2024.⁸⁵ The reason for the broad use of fluorine is many fold and detailed in *Table 4* which overall affects the ADMET properties of drugs.

*Table 4 - Properties of fluorine and their effect on drugs in a medicinal chemistry context.*⁸⁶

Reason	Explanation
Size	Fluorine is the second smallest atom after hydrogen, thus replacement of H (1.20 Å vdW radius) with F (1.47 Å vdW radius) atoms does not drastically change the overall steric aspects of the structure but can modulate other properties of the compound.
Stability of C-F bond	The C-F bond is the strongest bond that carbon can possible form (<i>i.e.</i> bond dissociation energy can be up to 130 kcal mol ⁻¹); therefore, fluoro-organic compounds show a level of metabolic stability not exhibited by other groups.
Electronegativity	As the most electronegative element, it induces bond polarisation therefore affecting the lipophilicity of compounds.
Non-covalent bonding	It is a weak HB acceptor.
Overall, F can affect the ADMET properties of a drug with its implementation.	

Historically, trifluoromethyl has been used as a bioisosteric replacement for several functional groups. The most obvious ‘classical’ example would be the replacement of a methyl ($-\text{CH}_3$) group for $-\text{CF}_3$ which have often been used to modulate the lipophilicity of molecules.⁸⁷ It has also found novel applications as a bioisostere of $-\text{CH}_2\text{NO}_2$

(described as a non-drug like structure) in the form of $-\text{CH}_2\text{CF}_3$ in CB_1 positive allosteric modulators – increasing potency and metabolic stability.⁸⁸ Additionally, it has been shown to be a bioisostere for the ethyl group.⁸⁹

However, in early 2023, five European countries (*i.e.* Germany, Denmark, Norway, the Netherlands and Sweden) proposed the total restriction of per- and polyfluoroalkyl substances (PFAS) under the European Chemicals Agency (ECHA). This proposal defines PFAS as "any substance that contains at least one fully fluorinated methyl ($-\text{CF}_3$) or methylene ($-\text{CF}_2$) group without any H, Cl, Br, or I attached to it". It also explicitly proposed the banning of (trifluoromethyl)benzene containing substances.⁹⁰ The scientific committees are still deliberating on the proposal as of September 2025. The reason for this sudden interest in the regulation of PFAS is because these compounds and their degradation products can persist in the environment for a long time and their removal from water (ground and surface), soil and biota is very difficult (near impossible) and very expensive. Therefore there are calls for all non-essential uses to be phased out.⁹¹ In the medicinal chemistry context it has recently been noted by O'Hagan and Young that the metabolism of fluoro-pharmaceuticals resulting in fluoride or fluorinated by-products is not very well understood and it can take up to 30 years for microbiota to evolve in response to this environmental problem.⁹² Additionally, it has been shown that aryl- CF_3 containing drugs such as the antidepressant drug fluoxetine (*i.e.* prozac), can actively release trifluoroacetic acid into our environment and mechanisms of their degradation to release fluorinated products are only now being investigated.^{93, 94}

Therefore, while it took many years to establish $-\text{CF}_3$ as a mainstay of pharmaceutical structures, perhaps the time has come to reconsider its use and search for alternative functionalities with similar properties but with an overall 'greener' endpoint. As previously mentioned, one of the objectives of this MSc consists of the computational assessment of the similarities/dissimilarities between $-\text{CF}_3$ and a wide range of functional groups in terms of size (volume and surface), electronics and lipophilicity. The chosen potential bioisosteres for $-\text{CF}_3$ and the reason for their inclusion can be seen in *Fig. 12* and *Table 5*.

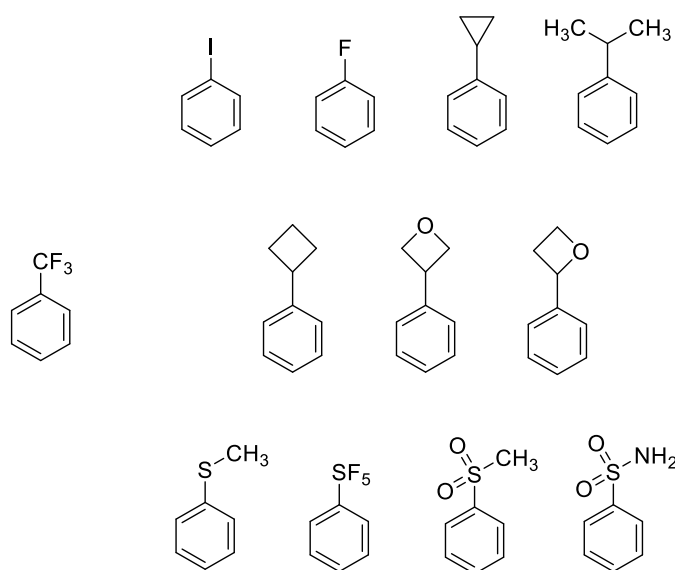


Figure 12 - Structures of analogues of (Trifluoromethyl)benzene under consideration as bioisosteres.

Table 5 - Selected functionalities as potential bioisosteres of $-CF_3$.

$R-C_6H_5$	Reason for Inclusion
-F and -I	Classic halogen bioisosteres.
-cycloPr, isoPr and cycloBu	- Though potentially larger than $-CF_3$, branched and linear alkanes may exhibit similar lipophilic properties.
-3-oxetyl and -2-oxetyl	Oxetanes are becoming attractive structures in medicinal chemistry due to their low molecular weight and physicochemical properties hence they have also been included. ⁹⁵
$-SF_5$	A promising $-CF_3$ bioisostere where it has been discovered that its use instead of $-CF_3$ leads to either similar biological activity as in the case of antimalarial mefloquine (which contains two $-CF_3$ groups), ⁹⁶ or in the case of anti-depressants fluoxetine and fenfluramide they show similar or better affinity for their respective 5-HT receptors. ⁹⁷
$-SCH_3$, $-SO_2CH_3$ and $-SO_2NH_2$	Following the idea of $-S-$ containing functionalities, we have also included methylsulfide, methylsulfonyl and sulfonamide.

The potential similarities of these functionalities to $-CF_3$ will be assessed based on their theoretical volume and surface area, as well as their electronics, which will be explored using molecular electrostatic potential (MEP) maps, calculated molecular polarity index (MPI) and polar surface area (PSA). Furthermore, given the broad use of $-CF_3$ as a lipophilic substituent in medicinal chemistry and given our own interest in

bioisosteres that are equally if not more lipophilic than -CF₃, the cLogP has been calculated for the proposed derivatives in addition to their experimental logP *via* shake-flask tests using UV spectrometry, thereby assessing cLogP accuracy also.

3.1.1 Comparison of Size

The size of the different functionalities proposed as -CF₃ bioisosteres was assessed using MEPcalculated volumes (V_{MEP}) at the MP2/cc-pVTZ or MP2/cc-pVTZ-PP for -I derivative computational levels. The volume was calculated as that of the corresponding phenyl derivative minus the volume of a benzene molecule. Therefore, an error is introduced corresponding to the volume of a H atom; however, this value will be the same for all calculated volumes and hence not affect the overall trend. The results of these calculations can be seen in *Table 6*.

Table 6 - Calculated volume of substituents from MEP calculations minus the volume of a benzene molecule calculated at the MP2/cc-pVTZ (and MP2/cc-pVTZ-PP for -I derivative) computational level.

Substituent	V_{MEP} (Å ³)
-CF ₃	37.12
-I	42.12
-F	5.42
-cyclo-Pr	56.72
-iso-Pr	68.62
-cyclo-Bu	78.32
-3-Oxetyl	66.32
-2-Oxetyl	66.42
-SF ₅	61.62
-SCH ₃	52.32
-SO ₂ CH ₃	66.92
-SO ₂ NH ₂	60.02

As expected, in comparison to -CF₃, -F is the smallest substituent by far with a volume of 5.42 Å³ versus the next one, -CF₃, which shows a volume of 37.12 Å³. Given a ±10 Å³ range, most substituents have a much larger volume than -CF₃; however, -I with a 42.12 Å³ volume falls within this range. Therefore, from a volume perspective based on MEP calculations, iodine could be considered a bioisostere of -CF₃. Moreover, if an

active site can accommodate an increased volume with an upper limit of $+20 \text{ \AA}^3$, -cyclo-Pr and -SCH₃ may also be considered potential bioisosteres.

3.1.2 Comparison of Electronics

To assess the potential similarities between the functionalities studied in electrostatic terms, the MEP on the 0.001 a.u. electron density isosurface was calculated at the MP2/cc-pVTZ (MP2/cc-pVTZ-PP[27] for Ph-I) level for the phenyl derivatives in vacuum. A representative figure of the MEPs generated for (a) Ph-CF₃, (b) Ph-I, (c) Ph-SF₅ and (d) Ph-cycBu (*Fig. 13*). The $V_{\min}$ value (green coloured in the maps in *Fig. 15*) represents the most negative electrostatic potential on the molecular surface indicating a region of high electron density whereas, conversely, $V_{\max}$ (black dots in the maps in *Fig. 13*) represents a region of depleted electron density.

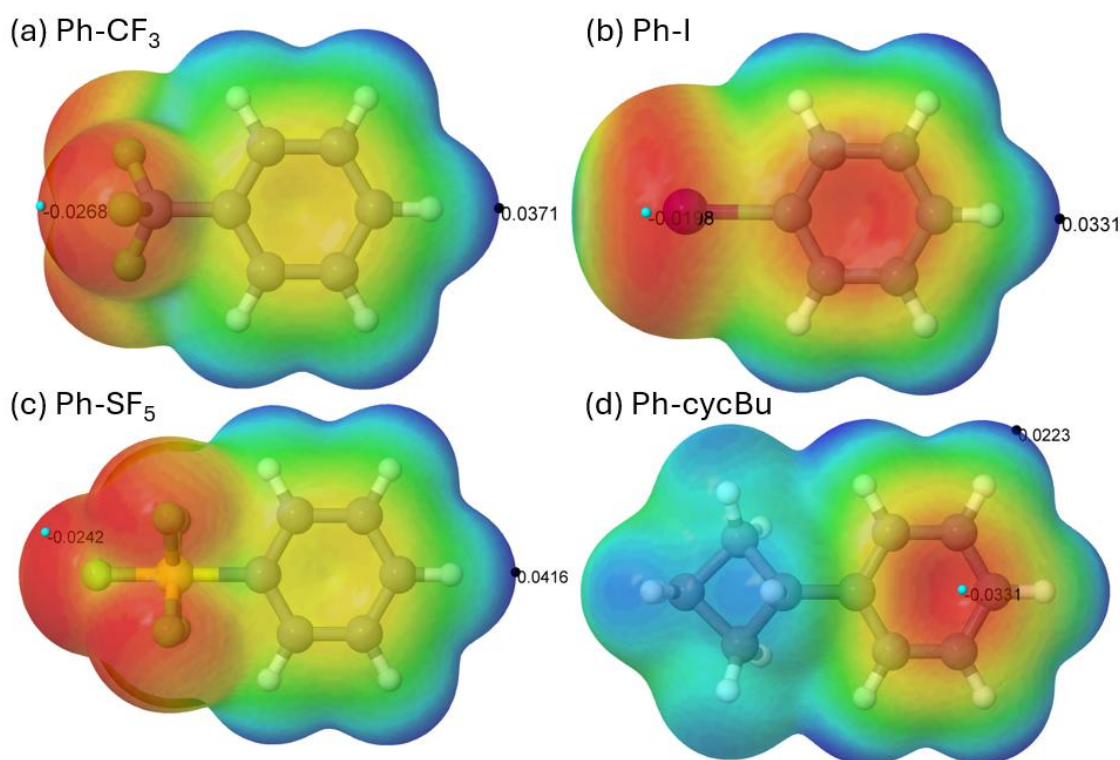


Figure 13 - Molecular electrostatic potential (MEP) on the 0.001 a.u. electron density isosurface for (a) Ph-CF₃, (b) Ph-I, (c) Ph-SF₅ and (d) Ph-cycBu calculated at MP2/cc-pVTZ level. $V(\max, I)$ and $V(\min, I)$ values (hartrees) are indicated. Colour scheme ranges from red (-0.1 hartrees) to blue (+0.2 hartrees).

The results of MEP analysis, $V_{\max,l}$ and $V_{\min,l}$ as well as the calculated molecular polarity index (MPI, kcal mol⁻¹) and polar surface area (Å²) can be found in Table 7. On all derivatives studied, $V_{\max,l}$ was found to be on the phenyl ring itself and, depending on the substituent, in the *para*- (-CF₃, -I, -SF₅, -SO₂CH₃ and -SO₂NH₂), *meta*- (-F and -cyclo-Bu) or *ortho*- (-cyclo-Pr, -iso-Pr, -3-oxetyl and -2-oxetyl) positions. However, in the case of Ph-SO₂CH₃ and Ph-SO₂NH₂, the real $V_{\max,l}$ was found to be located on the -CH₃ and -NH₂ moieties specifically with the reported $V_{\max,l}$ value in Table 7 being a partial $V_{\max,l}$. Given a tolerance of ±10% for $V_{\max,l}$ and taking into account its location, only -I can be considered similar to -CF₃ with -SF₅ being a close second, just in the limit. Regarding $V_{\min,l}$, its location was found to be mostly over the heteroatoms (*i.e.* I, F, O and S); however, in the case of the aliphatic derivatives (*e.g.* (d) Ph-cyc-Bu in Fig. 17), the $V_{\min,l}$ was found on the phenyl ring (though partial minima were found on the phenyl ring for the non-aliphatic substituents). Using the same boundary as previously but for $V_{\min,l}$, only -SF₅ can be considered similar to -CF₃.

Table 7 - Calculated MEP $V_{\max,l}$ and $V_{\min,l}$ (a.u., provided by Multiwfn), Molecular Polarity Index (kcal mol⁻¹), and Polar Surface Area (Å², provided by Multiwfn) at MP2/cc-pVTZ and MP2/cc-pVTZ-PP (for Ph-I) computational levels [10% range from the -CF₃ values considered; $V_{(\max,l)} \pm 0.0056$ and $V_{(\min,l)} \pm 0.0040$]. Values in bold are within the 10% range considered.

Derivative	MEP $V_{(\max,l)}$ (Hartree)	MEP $V_{(\min,l)}$ (Hartree)	Molecular Polarity Index (kcal mol ⁻¹)	Polar Surface Area (Å ²)
Ph-CF ₃	0.0371 (para)	-0.0268	10.33	82.08
Ph-I	0.0331 (para)	-0.0198	8.88	64.29
Ph-F	0.0320 (meta)	-0.0311	9.85	64.61
Ph-cyclo-Pr	0.0267 (ortho)	-0.0333	8.23	64.00
Ph-iso-Pr	0.0262 (ortho)	-0.0325	6.93	47.47
Ph-cyclo-Bu	0.0223 (meta)	-0.0331	6.64	41.05
Ph-3-Oxetyl	0.0320 (ortho)	-0.0613	10.51	90.60
Ph-2-Oxetyl	0.0283 (ortho)	-0.0610	10.40	89.26
Ph-SF ₅	0.0416 (para)	-0.0242	11.12	100.34
Ph-SCH ₃	0.0291 (para)	-0.0360	9.14	73.64
Ph-SO ₂ CH ₃	0.0427 (para)	-0.0681	16.99	126.38
Ph-SO ₂ NH ₂	0.0693 (para)	-0.0617	15.75	107.35

The MPI calculated for the phenyl derivatives is a measure of the polarity of the whole molecule wherein the higher the polarity the larger the MPI value. It is calculated by Eqn. (1) where it is obtained from the van der Waals surface and electrostatic potential integrating over the whole molecular surface.⁹⁸

$$\text{MPI} = \left(\frac{1}{A} \right) \iint_S V(r) \, dS \quad (1)$$

This index characterises the polarity of the whole molecule contrasting with dipole moment which only measures local polarity caused by an uneven distribution of the electrostatic potential. The MPI values obtained show that the phenyl derivatives with the -F, -2-oxetyl, -3-oxetyl and -SF₅ substituents exhibit similar MPI to that of Ph-CF₃ with Ph-I and Ph-CF₃ being close.

Finally, the polar surface area (PSA) of these phenyl derivatives were calculated from the electrostatic potential (*Table 7*). PSA is associated with the ability of a compound to form HBs and represents the area over all heteroatoms (*i.e.* N, O, S and F) including their corresponding attached hydrogens. Considering a boundary of $\pm 10 \text{ \AA}^2$ from the PSA value obtained for Ph-CF₃, only the oxetyl and -SCH₃ derivatives fall within this range.

Overall, taking into account MEPs, MPI and PSA calculated from the electrostatic potential of the phenyl derivatives it can be said that: -SF₅, -I, -F, -2- and -3-oxetyl and -SCH₃ could be considered bio-equivalent to -CF₃ in terms of electronics.

3.1.3 Comparison of Lipophilicity

To assess the lipophilicity of molecules in the form of the theoretical logP coefficient, several web-based freeware methods were employed, and the results obtained are outlined in *Table 8*.

Additionally, as previously mentioned, an in-house quantum chemical logP coefficient (n-octanol/water partition, Mat-LogP) calculation was also carried out based on work reported by Truhlar and Cramer *et al.* according to Eqn. (2):⁹⁹

$$\log P = -\frac{\Delta G_{o/w}^{\circ}}{2.303RT} = -\frac{G_o^{\circ} - G_w^{\circ}}{2.303RT} \quad (2)$$

Table 8 - Outline of freeware web-based logP calculators used to theoretically assess the lipophilicity of phenyl derivatives studied.

On SwissADME:¹⁰⁰
- XLOGP3 – atomistic method including corrective factors and knowledge-based library.¹⁰¹ - WLOGP – SwissADME’s own implementation of a purely atomistic method based on fragmental system of Wildman and Crippen.¹⁰² - MLOGP – archetype of topological method relying on a linear relationship with 13 molecular descriptors.¹⁰³ - SILICOS-IT – hybrid method using 27 fragments and 7 topological descriptors. - iLOGP – SwissADME’s in-house physics based method relying on free energies of solvation in n-octanol and water calculated by the generalised Born and solvent accessible surface area model.¹⁰⁴
On Marvin:
LogP atomistic method based on a library of 893 compounds. ¹⁰⁵
On Molinspiration:
miLogP (from https://www.molinspiration.com/) based on group contributions from a library of approximately 12,000 compounds from Molinspiration’s suite of chemoinformatics software.

The SMD (SMD18 for Ph-I) free energies (G_o° in n-octanol and G_w° in water) were calculated at the MP2/cc-pVTZ (MP2/cc-pVTZ-PP for Ph-I) computational level at 298.15 K.

Upon analysis of the calculated logP coefficients, two web-based calculators seemed to differ largely from the average value (consensus logP), iLogP and WLOGP and hence they were removed from the general discussion of results. The remaining values can be seen in *Fig. 14* where the upper and lower limits of the values generated for Ph-CF₃ are indicated by black lines.

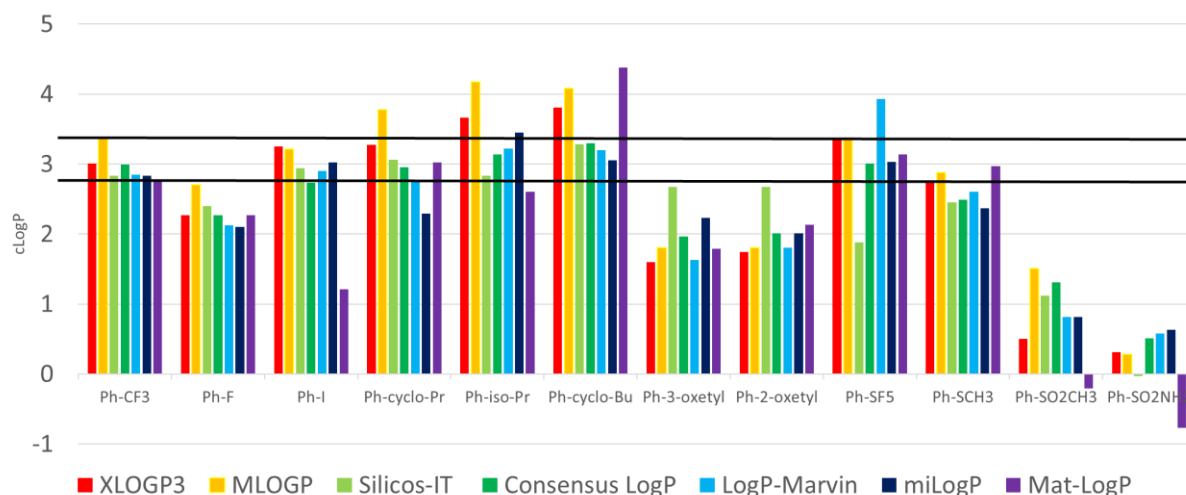


Figure 14 - Plots of selected LogP values as calculated from different web operators from SwissADME (XLOGP3, MLOGP, Silicos-IT, Consensus LogP), Marvin (LogP-Marvin) and Molinspiration (miLogP) as well as the in-house LogP (Mat-LogP) calculated at MP2/cc-pVTZ with SMD (MP2/cc-pVTZ-PP with SMD18 for Ph-I) level. The black lines indicate the maximum and minimum LogP values obtained for Ph-CF₃.

Considering the upper and lower reference limits, all seven methods predict a lower logP value for derivatives with -F, -2- and -3-oxetyl, -SO₂CH₃ and -SO₂NH₂, indicating a lipophilicity very different to -CF₃ (*i.e.* much lower in some cases) and, hence, they are not good lipophilic equivalents of -CF₃. Additionally, for the -SCH₃ derivative only three methods, XLOGP3, MLOGP and Mat-LogP, predict a similar logP to -CF₃ and, therefore, it can be considered to be not a very adequate equivalent to -CF₃ in terms of lipophilicity. Furthermore, for the -cyclo-Pr and -cyclo-Bu derivatives four methods predict a similar logP to -CF₃ (*i.e.* Silicos-IT, Consensus-LogP, LogP-Marvin and miLogP) and while MLOGP and XLOGP overestimate it for these compounds, Mat-LogP overestimates it for -cyclo-Bu and underestimates it for -iso-Pr. Five methods predict a similar lipophilicity for the -cyclo-Pr derivative (XLOGP, Consensus-LogP, Silicos-IT, LogP-Marvin and Mat-LogP), and similarly, five methods predict a similar lipophilicity to -CF₃ for -SF₅ (XLOGP3, MLOGP, Consensus-LogP, miLogP and Mat-LogP). Finally, all but one method (Mat-LogP) predicts that -I has a similar lipophilicity to -CF₃. Given these results, it can be concluded that -I followed by -SF₅ followed by -cyclo-Pr are potentially recommended equivalents to -CF₃ in terms of lipophilicity.

To validate the in-house Mat-LogP calculation, the experimental logP of the compounds was determined for the phenyl derivatives *via* the shake-flask method at

room temperature through UV-vis absorption spectroscopy analysis.¹⁰⁶ All compounds were commercially available except for the -2- and -3-oxetyl derivatives which were synthesised (see Chapter 5). Unfortunately, Ph-cyclo-Bu was excluded due to synthetic and economic limitations, however, this would not represent a problem given the predicted lack of equivalence to -CF₃ in the previously discussed theoretical parameters. The experimentally determined partition coefficient (LogP_{exp}), the QM calculated partition coefficient (Mat-LogP) and available literature experimentally determined partition coefficients for some of the compounds studied (LogP_{exp, ref}) can be found in *Table 9*.

Table 9 - LogP values experimentally measured by us (LogP_{exp}), theoretically calculated by us (Mat-LogP) and experimental values found in reference¹³⁷ (LogP_{exp, ref}).

Substituent	LogP _{exp}	Mat-LogP	LogP _{exp, ref} ¹⁰⁷
-CF ₃	2.49	2.75	-
-I	2.95	1.21	3.25
-F	2.1	2.27	2.27
-cyclo-Pr	2.9	3.02	-
-iso-Pr	3.3	2.6	3.66
-3-Oxetyl	1.13	1.79	-
-2-Oxetyl	1.42	2.13	-
-SF ₅	3.09	3.14	-
-SCH ₃	2.82	2.97	2.74
-SO ₂ CH ₃	0.7	-0.21	0.5
-SO ₂ NH ₂	-1.37	-0.77	-

Firstly, to validate our own LogP_{exp} values, we compared them to those in the literature (*i.e.* Ph-I, Ph-F, Ph-iso-Pr, Ph-SCH₃ and Ph-SO₂CH₃) and found excellent correlation (LogP_{exp, ref} = 1.18 x LogP_{exp} – 0.31; R² = 0.98, F = 178.5; s = 0.18; n = 5). Comparing our theoretically (Mat-LogP) and experimentally (LogP_{exp}) determined values it was found that -I stands out as a major outlier. In order to correct this deviation, it was thought first to use refined SMD parameters *i.e.* SMD18 as it is recommended for any calculations involving iodides.¹⁰⁸ However, this did not make any difference to the overall Mat-logP values. Hence, removing the value for Ph-I, a good correlation with LogP_{exp} was found (Mat-LogP = 0.88 x LogP_{exp} + 0.33; R² = 0.86, F = ??; s = ??; n = 10).

Therefore, it can be concluded that this methodology may be a good guide in the assessment of lipophilicity except for compounds containing iodine.

Finally, considering the results for LogP_{exp} as a reference, it can be proposed that -F and -SCH₃ (within ±15% range) as well as -I and -cyclo-Pr (within a ±20% range) may be good bioequivalents of -CF₃ in terms of lipophilicity.

3.1.4 Conclusion – Bioisosteres of -CF₃

The summary of all theoretical and experimental findings can be found in *Table 10* identifying bioisosteres of -CF₃ based on different size, electronic and lipophilic criteria, a shaded box indicates it fulfils the selection criteria. In terms of size (*i.e.* MEP volumes) it was found that -I is a good equivalent to -CF₃ and with an expanded limit range, potentially -cyclo-Pr and -SCH₃ may also be useful alternatives. In terms of electronic properties (*i.e.* MEP, MPI and PSA), it was found that -I, -F, -3- and -2-oxetyl, -SF₅ and -SCH₃ were reasonable equivalents to -CF₃. And finally, in terms of theoretical lipophilicity it was found that -I, -cyclo-Pr and -SF₅ could be potential equivalents to -CF₃. Experimental lipophilicity determined that -I, -F, -cyclo-Pr and -SCH₃ were good equivalents of -CF₃ in terms of lipophilicity. The calculations and measurements seem to indicate that -I is the best option to replace -CF₃ since it is equivalent in all aspects studied *i.e.* size, electronics and lipophilicity.

Table 10 - Equivalences found for -CF₃ when comparing size, electronic and lipophilic parameters for all compounds studied. Shaded box indicates substituent satisfies the selection criteria for the parameter.

Substituent	Theoretical			Experimental
	Size	Electronic	Lipophilic	Lipophilic
-I				
-F				
-cyclo-Pr				
-iso-Pr				
-cyclo-Bu				
-3-Oxetyl				
-2-Oxetyl				
-SF ₅				
-SCH ₃				

-SO₂CH₃

-SO₂NH₂

As a proof of concept to test these outcomes, we selected a known -CF₃ containing drug, Efavirenz (non-nucleoside reverse transcriptase inhibitor), the crystal structure of which, in complex with its target (reverse transcriptase, RT), is known (PDB:7KJW),¹⁰⁹ and the -CF₃ was replaced by -I in a docking exercise. Both structures were optimised at M06-2X/6-311+G(d,p) and M06-2X/[6-311+G(d,p) (light atoms) plus LANL2DZdp (iodine)] computational levels for efavirenz and I-efavirenz, respectively. These optimised structures were then docked in RT using as template the crystal structure after removal of the ligand and using a ligand flexible-rigid target approach with AutoDock Vina 1.2.0. In Figure 15 it is shown how both structures (efavirenz in pink and I-analogue in blue) fit equally well in the binding site of RT and the G-scores obtained (-9.56 and -9.81 kcal mol⁻¹, respectively) were very similar, being that of the I-efavirenz slightly better. Induced fit docking was also performed resulting in a similar trend (-9.56 kcal mol⁻¹ for I-Efavirenz and -9.50 kcal mol⁻¹ for Efavirenz) (Fig. 17).

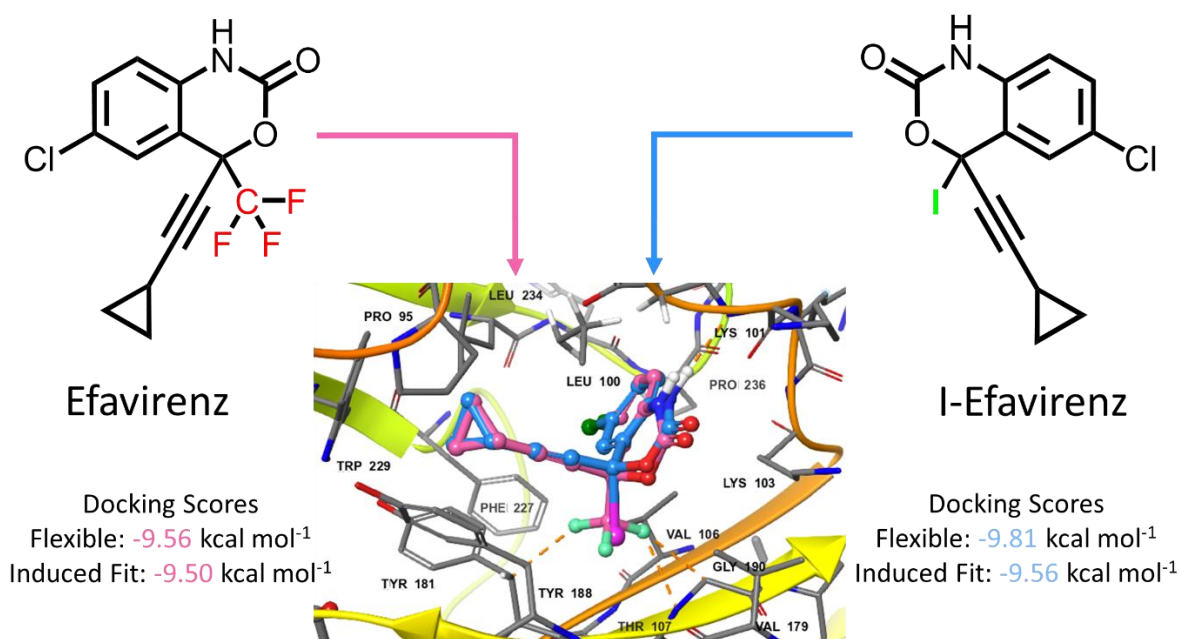


Figure 15 - Superposition of the optimal poses of efavirenz (pink) and its I-analogue (blue) docked in the binding site of RT (7KJW) using AutoDock Vina 1.2.0. Ligands were optimised

using M06-2X/6-311+G(d,p) and M06-2X/[6-311+G(d,p) (light atoms) plus LANL2DZdp (iodine)] for efavirenz and I-analogue, respectively.

More details and data supporting these outcomes is available in the supporting information material of the published version of this research.¹¹⁰

3.1.5 A Note on ‘Greenness’

The greenness of a compound of course does not only depend on its end-impact on the environment, but also on its preparation. For all functionalities chosen there are numerous examples of their green synthesis: halogens¹¹¹, alkanes¹¹², oxetanes¹¹³, sulfones¹¹⁴ and pentafluorosulfanyls.¹¹⁵

3.2 Molecular Modelling and Theoretical ADMET Properties

Given the results presented in Section 3.1 for bioequivalents of -CF₃, and our interest in alkylguanidines and triazole derivatives for potential InhA inhibitors, the compounds shown in Fig. 17 were considered for molecular modelling investigations. Several aspects were explored; firstly, the characteristics of InhA as a protein and its active site, along with mechanism of action on its substrate; secondly, the refinement of the theoretical binding pose based on previous work within the Rozas group; and thirdly, the evaluation of novel derivatives in their binding to InhA as well as a theoretical ADMET assessment for the newly proposed derivatives. Alkyl chain derivatives have been proposed as Forge *et al.* has shown alkyl chain length can positively influence anti-mycobacterial activity.⁸⁰ The motivations for the triazole derivatives will be discussed later within this section.

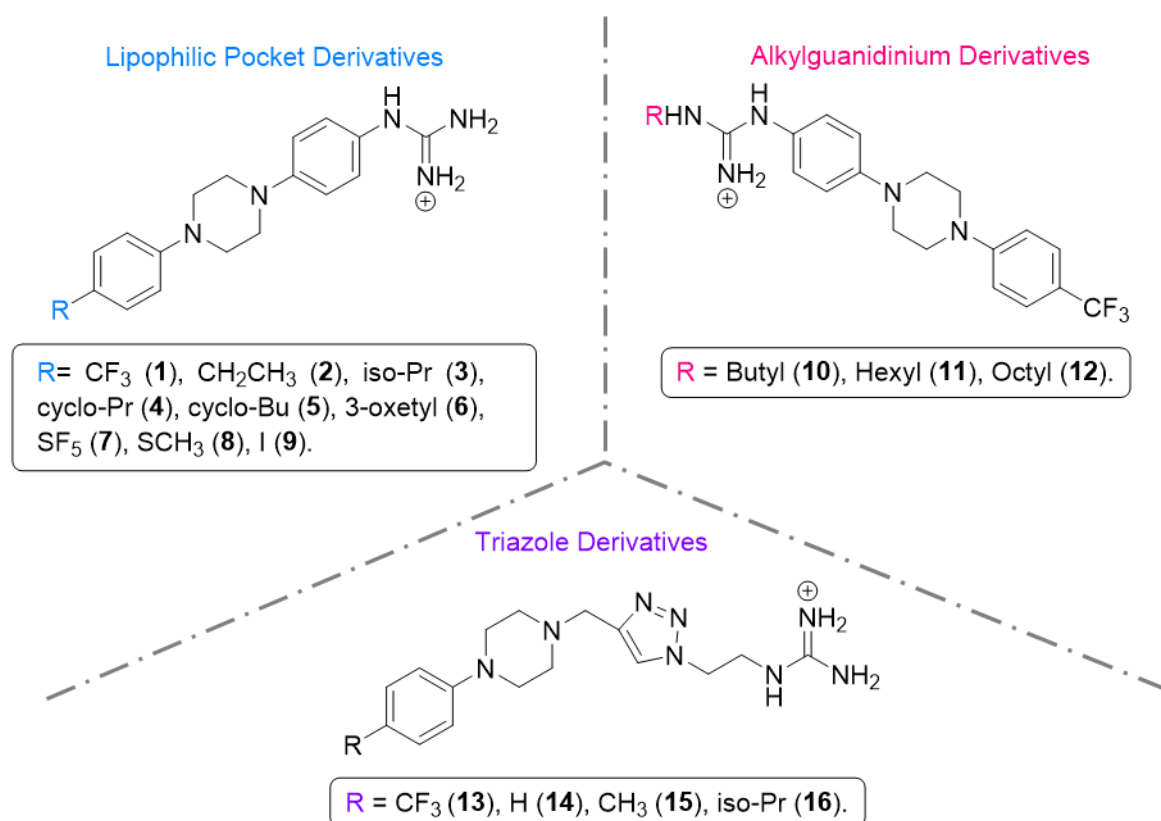


Figure 16 - Structures of compounds to be investigated using molecular modelling to characterise their interaction with InhA (lipophilic pocket, alkyl chain and triazole derivatives).

3.2.1 InhA Enzyme and Considerations of its Active Site

As previously mentioned enoyl acyl carrier protein reductase (InhA) is involved in mycolic acid synthesis. Specifically, it is involved in the NADH-dependent reduction of enoyl-ACP at the later stages of the FAS-II mycolic acid synthesis pathway and is also involved in fatty acid synthesis.¹¹⁶ It has been proposed to work through a pericyclic 'ene-shaped' transition state which forms an observed C2-ene adduct from NADH and enoyl-CoA (localised within the active site of InhA). An alternative pathway for formation is *via* an enolate formed from direct hydride transfer through a Michael-addition reaction. The C2-ene adduct can then be resolved through an elimination reaction, and thus the enolate abstracts a proton to form the reduced acyl-CoA product (Fig. 17).¹¹⁷

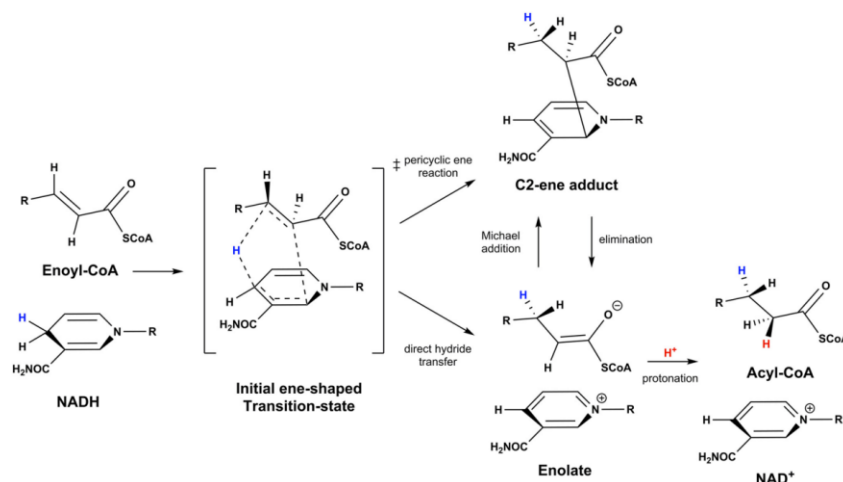


Figure 17 - Proposed mechanism of InhA catalysed NADH-mediated reduction of Enoyl-CoA substrate. Taken from reference 117.

The InhA protein itself is a homotetramer composed of a repeating unit of a single domain with a Rossmann fold (tertiary folds in proteins' bind cofactors) in the core providing space for the NADH binding site.¹¹⁸ Its natural substrates are characterised by the long aliphatic chains starting at C16 and going as high as C56 (R in Fig. 18). Examining the protein with a C16 substrate analogue and NAD⁺ (PDB ID: 1BVR) a better picture of the active site can be achieved (see Fig 18.A).¹¹⁹ The catalytic site is located in a deep hydrophobic pocket wherein Phe149-Tyr158-Lys165 play an important role in cofactor binding, stabilisation of 'ene' transition state and proton transfer. The entrance to the active site, termed the 'major portal', is solvent exposed. A 'minor portal' exists in the hydrophobic pocket; however, its opening is restricted and highly dependent on substrate binding.^{119, 120} From one InhA structure to another, Phe149 and Tyr158 can exist in an 'in' or 'out' conformation depending on the bound substrate.¹²¹ The 2D ligand-protein interaction diagrams can be seen for NAD⁺ and the C16 substrate in Fig. 18 B and C showing the stabilisation of NAD⁺ and the lipophilic and solvent exposed portions of the active site respectively.

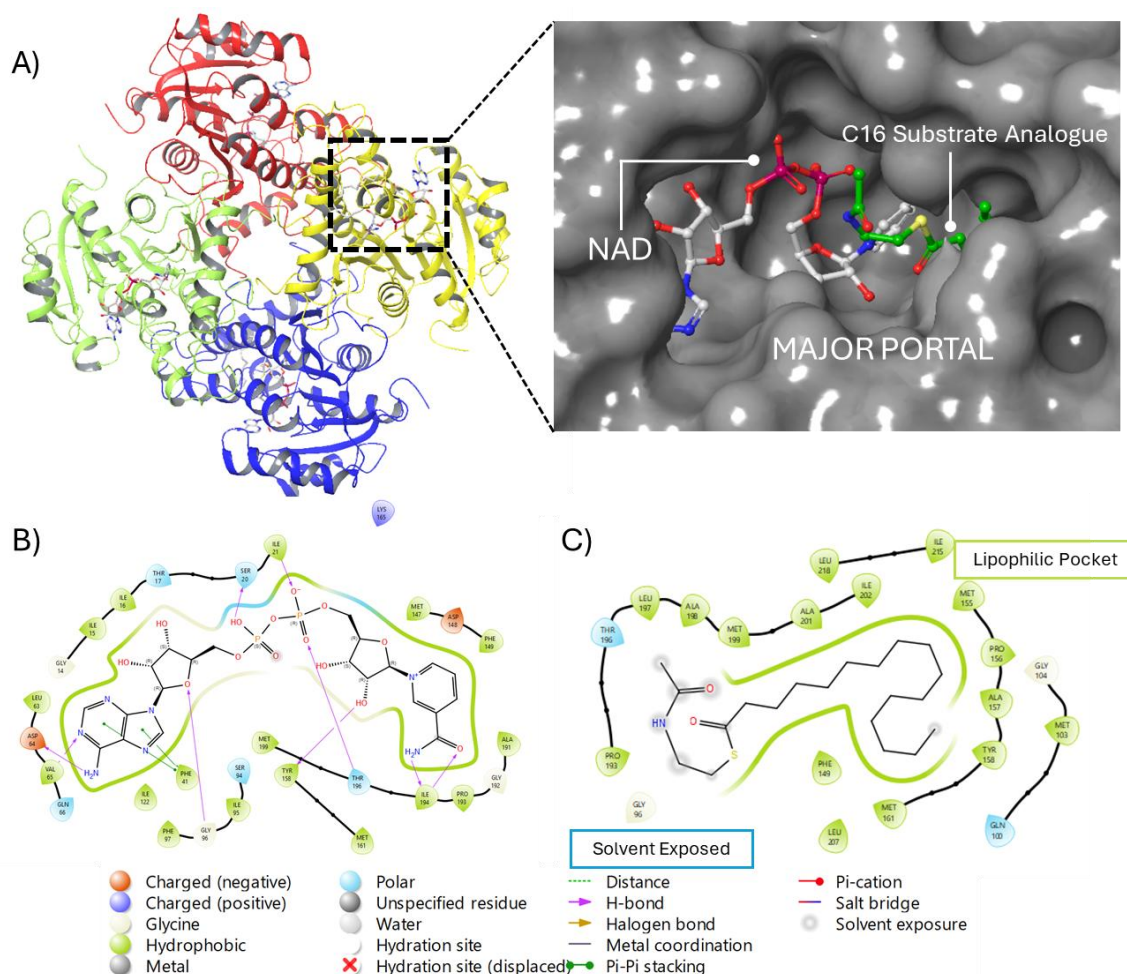


Figure 18 – A) *InhA* protein tetramer (PDB ID: IBVR) and its active site showing the major portal opening and localisation of NAD⁺ and C16 substrate analogue in the active site. B) 2D Ligand-protein interaction diagram of NAD⁺ and C) 2D Ligand-protein interaction diagram of the C16 substrate analogue.

3.2.2 Binding Pose Refinement – Building on Previous Work

In previous work in the Rozas group (Adeyemi Rahman, PhD work, 2023) after target identification, molecular modelling studies were carried out using as template the crystal structure data of *InhA* co-crystallised with a methyl thiazole (PDB ID: 4BQR), which was removed previous to the modelling.^{81, 122} Geometry optimisation of the bis-guanidine derivatives studied by Dr Rahman returned an L-shaped conformer similar to the methyl thiazole (Fig. 19 A). Docking of the amphiphilic derivative in the form of compound **1** yielded the following observations of an L-shaped binding motif: (i) the polar guanidine ‘warhead’ interacts with the ribose portion of NADH by means of a HB (2.11 Å) within the lipophilic pocket; (ii) this interaction left the lipophilic portion of the

molecule (*i.e.* the 4-CF₃-phenyl group) to interact with a ‘secondary’ lipophilic pocket (Fig. 19 B).⁸¹ This templated approach is quite common in research and indeed in the first steps in many molecular modelling campaigns when ligand-protein co-crystal structure data is unavailable. However, this can lead to biased binding modes as it is highly system dependent.¹²³

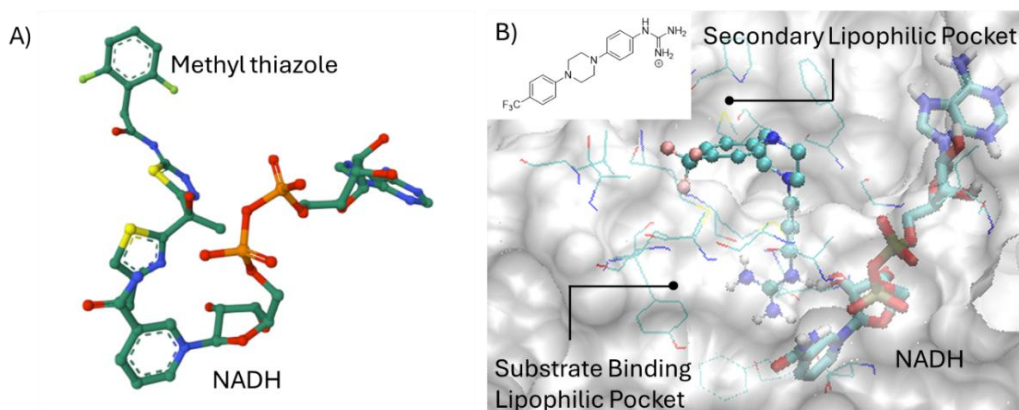


Figure 19 - A) Binding pose of methyl thiazole derivative in InhA (PDB ID: 4BQR). B) Molecular docking performed in the Rozas group showing an L-shaped binding motif of compound **1** with the guanidine facing towards the substrate binding pocket while the lipophilic phenyl ring interacts with a ‘secondary’ lipophilic pocket.

Further complications arise due to the ability of piperazine to exhibit multiple conformations (Fig. 20). It is generally accepted that the most stable conformer is the equatorial-equatorial (eq-eq) chair conformation. However, there are three more possible conformers, with axial-eq and axial-axial being within 2 kcal mol⁻¹ from the most stable at room temperature (though at the B3LYP/6-311-G(d,p) level of theory which conformer predominates is unknown).¹²⁴ Indeed, the energy difference between L-shaped and linear conformation of **1** at the M06-2X/6-311+G(d,p) computational level is approximately 1 kcal mol⁻¹, and given the dynamic nature of protein-ligand interactions, without a co-crystal structure or complex molecular dynamics simulations (outside of the scope of this project), it is unknown which conformer is preferred in the interactions with InhA. Nevertheless, in order to simplify the forthcoming molecular docking studies an effort was made to conformationally sample binding poses and eliminate biologically irrelevant binding modes.

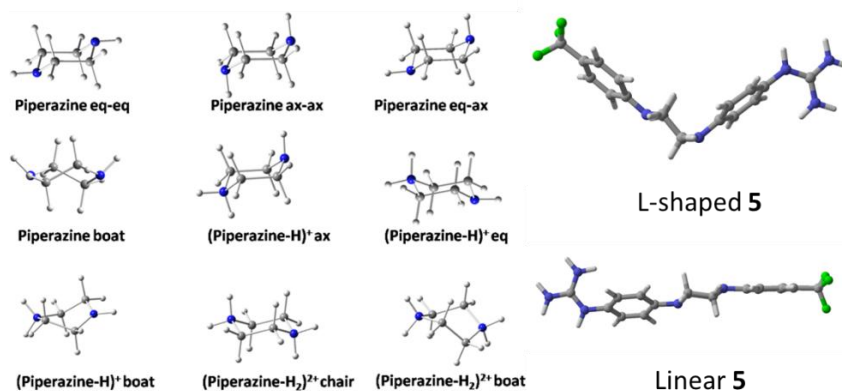


Figure 20 – Left - Theoreticaly optimised structures of the low-lying conformations of piperazine and its protonated cationic forms. The geometries of the major conformers were optimised at B3LYP/6-311-G(d,p) level of theory. Reproduced from Ref. 124. Right – Optimised structures of 5 (L-shaped and linear) at the M06-2X/6-311+G(d,p) level.

To perform this sampling, the first tool that was considered was DiffDock which is a novel state-of-the-art method for prediction of binding poses. First unveiled in 2023 by MIT researchers at the 11th International Conference on Learning Representations, it is a diffusion-based machine learning method for docking small molecules into proteins. It is different to the “sampling and scoring” approach used by traditional docking approaches (detailed later) and considers binding as a regression problem, where randomly sampled initial poses are denoised *via* diffusion over the protein surface over translational, rotational and torsional degrees of freedom. They report that the model performs better than Autodock Vina and Glide which are traditional docking software where the RMSD of 20% of predicted poses falls under 2 Å from the crystal structure (1.4% and 19.6% for Autodock Vina and Glide respectively).¹²⁵ In this way multiple conformers of the piperazine could be sampled, however, its major drawback is that currently it is only implemented for single molecule to protein docking thereby being unable to also accommodate NADH which in the present study is essential.

However, there is another docking method, Attracting Cavities 2.0, which in principle permits higher degree of ligand flexibility, similar to DiffDock, while also allowing for the presence of cofactors. Developed by the same team behind SwissADME, this docking method replaces the rough energy landscape of the macromolecule by a

smooth attractive energy landscape and the simple rotation, translation and geometry optimisations of the ligand on this smooth landscape results in an efficient sampling algorithm for docking. The scoring function *i.e.* how ‘good’ the ligands fits into the protein’s binding site is composed of the CHARMM forcefield terms and the fast analytical continuum treatment of solvation (FACTS) model.¹²⁶ Therefore, lead compound **1** was docked using Attracting Cavities 2.0 into InhA (PDB ID: 5OIR). The top three binding modes and their respective AC and SwissParam Scores can be seen in *Table 11*.

The ΔG_{AC} and ΔG_{SP} scores provided by Attractive Cavities and SwissParam respectively, are physics based estimations of binding free energies of the ligand. Thus, ΔG_{AC} is a scoring function that uses molecular mechanics generalised Born surface area continuum solvation (MM-GBSA) method and may be written as:

$$\Delta G_{AC} = E_{intra,lig} + E_{intra,prot} + E_{vdW} + E_{elec} + \Delta G_{solv,elec} + \Delta G_{solv,np}$$

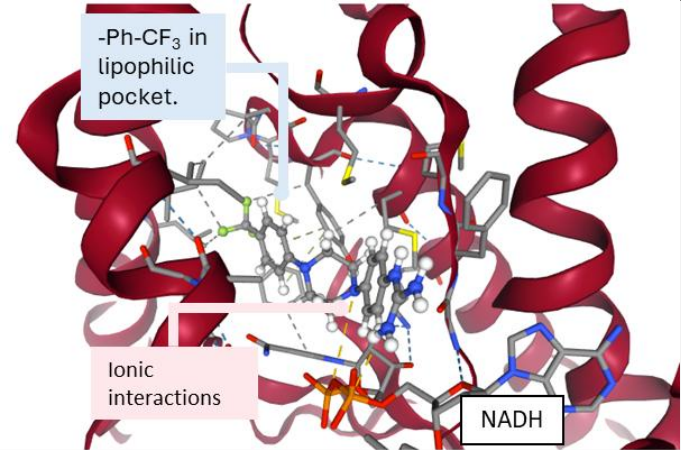
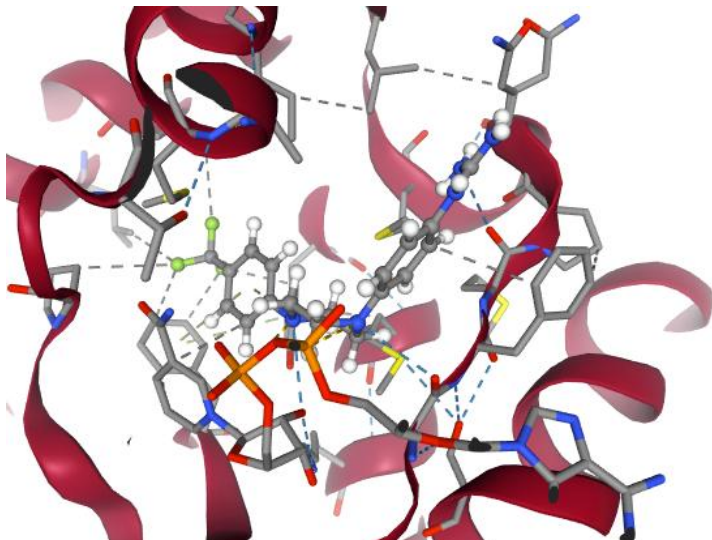
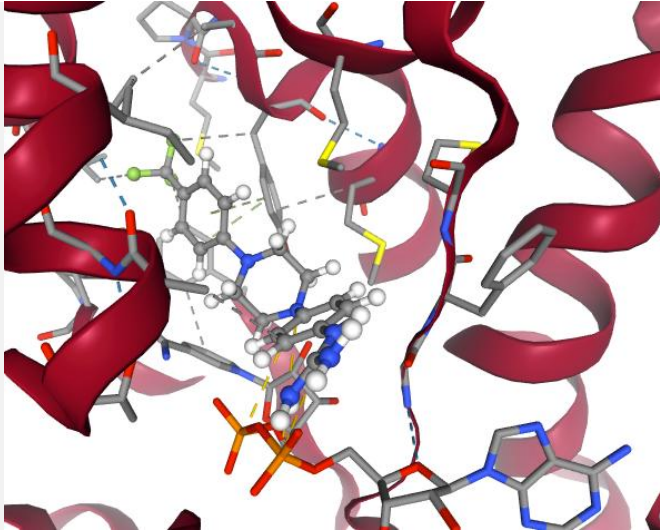
Where $E_{intra,lig}$ and $E_{intra,prot}$ are the internal energies of the ligand and the protein, respectively; E_{vdW} and E_{elec} are the van der Waals and Coulomb electrostatic energies of interaction between the ligand and the protein, respectively; $\Delta G_{solv,elec}$ is the electrostatic solvation energy of the complex calculated with FACTS; and $\Delta G_{solv,np}$ is the nonpolar solvation energy that is estimated as being proportional to the solvent accessible surface area of the complex.^{127, 128}

Meanwhile, SwissParam is a weighted sum of polar and nonpolar terms fitted to reproduce the experimental binding free energies of 214 ligand-protein complexes in the Ligand Protein Database (LPDB) meaning it is a hybrid scoring function using both force-field (physics-based) and empirical parameters. Thus, ΔG_{SP} is defined as:

$$\Delta G_{SP} = -3.99 + 0.0946 \times (E_{vdW} + \Delta G_{desolv,np}) + 0.0612 \times (E_{elec} + \Delta G_{desolv,elec})$$

with the van der Waals (E_{vdW}) and Coulomb electrostatic (E_{elec}) interaction energies between the ligand and the protein, and the nonpolar ($\Delta G_{desolv,np}$) and polar ($\Delta G_{desolv,elec}$) desolvation energies of the ligand and protein upon complexation calculated with the FACTS approach.¹²⁹ However, it is important to note that ΔG_{AC} and ΔG_{SP} can be used to rank poses of the same ligand to the same target, they are only approximations and hold no physical significance.

Table 11 – Top 3 binding poses generated by Attractive Cavities 2.0 along with their respective AC and SwissParam Scores.

	Binding Pose	AC Score	SwissParam Score
1	 -Ph-CF₃ in lipophilic pocket. Ionic interactions NADH	-113.98	-8.25
2		-113.41	-8.28
3		-113.37	-8.06

Examining the resulting top three binding modes (*Tab. 11*), it is notable that their calculated scores are nearly identical across both scoring functions despite differences in binding poses. In all binding modes the phenyl-CF₃ moiety sits in the lipophilic pocket, usually occupied by the substrate. Then the guanidinium moiety in Pose 1 and 3 interacts with NADH's phosphates in what Attracting Cavities 2.0 defines as 'ionic interactions' however no explicit hydrogen bond between the molecule and cofactor seems to be predicted. In Pose 2, we see a hydrogen bond form between the guanidinium and Met98. Binding Pose 2 exhibits the L-shaped binding mode (most likely due to the hydrogen bond formed with Met98) as well as Binding Pose 3. Binding Pose 1 where the piperazine is in a boat conformation exhibits the most linear binding pose. The lack of differentiation between these poses suggests that the ligand exhibits conformational flexibility in the active site (as expected) and that the scoring function struggles to distinguish between these different conformations which may thermodynamically accessible within the binding site.

The binding mode, where phenyl-CF₃ is buried in the lipophilic pocket and the guanidinium points outward towards the solvent accessible major portal is different to the previous binding mode identified in earlier templated molecular modelling studies where the guanidinium moiety points inwards towards the lipophilic substrate binding pocket (*Fig. 19 B*). The structural amphiphilic features of the lead compound appear to be optimally suited to both penetrate the lipophilic substrate binding pocket (formed by hydrophobic residues such as Phe149, Met155, Tyr158, Met199, Ile215 and Leu218) while the polar guanidinium headgroup points out toward the solvent accessible area and interacts with NADH. This would be the expected binding mode considering the natures of the compound and the putative binding site. Two factors support this newly identified position of the phenyl-CF₃ moiety; literature strongly supports that hydrophobic elements of drugs are more likely to penetrate hydrophobic pockets with hydrophobic interactions being the most common interactions in protein-ligand complexes.¹³⁰ It has also been shown that the burial of a polar head group in a lipophilic pocket can have a free energy penalty of up to 21 kJ mol⁻¹ thereby making it more unlikely.¹³⁰

Finally, the last tool considered to decide between the linear and L-shaped binding modes was to use AutoDock Vina 1.2 to dock both conformers into InhA. The use of AutoDock Vina 1.2 as the docking software of choice for the study of derivatives in *Fig. 16* is one of practicality since it allows for side chain flexibility, seen as important to obtain accurate binding modes, and it is an open source free docking software. Unlike the previous two scoring functions (ΔG_{AC} and ΔG_{SP}), AutoDock Vina does not employ an explicitly physics based scoring function and the creators see the scoring function as a ‘machine learning’ exercise which exclusively uses empirical functions. These functions describe van der Waals intermolecular forces, coulombic interactions between oppositely charged particles and HBs between the ligand and protein. In this way, docking software employing empirical scoring functions are much faster than force-field based scoring functions such as those used in Attracting Cavities 2.0.¹³¹,

132

Both conformers were optimised at the M06-2X/6-311+G(d,p) computational level and docked into InhA (PDB ID: 5OIR) in a flexible ligand-rigid protein approach (*Fig. 21*). The linear and L-shaped dockings had a G-score of -8.8 kcal mol⁻¹ and -9.1 kcal mol⁻¹ respectively. This indicates that the L-shaped conformation is slightly more ‘stable’ and both conformers’ lipophilic moiety (phenyl-CF₃) sits in the lipophilic substrate binding pocket (L-shaped to a lesser extent). However the linear conformer is able to form a hydrogen bond with the phosphate group of NADH at bond length of 2.20 Å (echoing Attracting Cavities 2.0) while in the L-shaped conformer the guanidine head group points up and out of the active site and a bad contact *i.e.* clash occurs between it and a polar Gln100 residue implying in a dynamic case, it would shift away from this positions. Several attempts were made to try replicate the binding mode found with Attracting Cavities 2.0 whereby even a slightly curved conformer was able to form ionic interactions with NADH or hydrogen bonding with Met98. The box size and position *i.e.* search space of docking was changed several times to no difference (Vina only allows for ligand torsional flexibility). This result is probably due to the extreme L-shaped conformation whereas in Binding Pose 2 and 3 (*Table 11*) it was curved to a lesser extent. In line with these findings, only the linear conformation will be considered to optimise time and computational costs however it is important to

be aware it is not the only conformer possible. This should not pose a major issue as we are primarily concerned with the effect of substitution of the phenyl ring to target the lipophilic substrate binding pocket and the G-score seems to be insensitive to curvature *i.e.* -8.8 kcal mol⁻¹ (linear conformer) versus -9.1 kcal mol⁻¹ (L-shaped conformer).

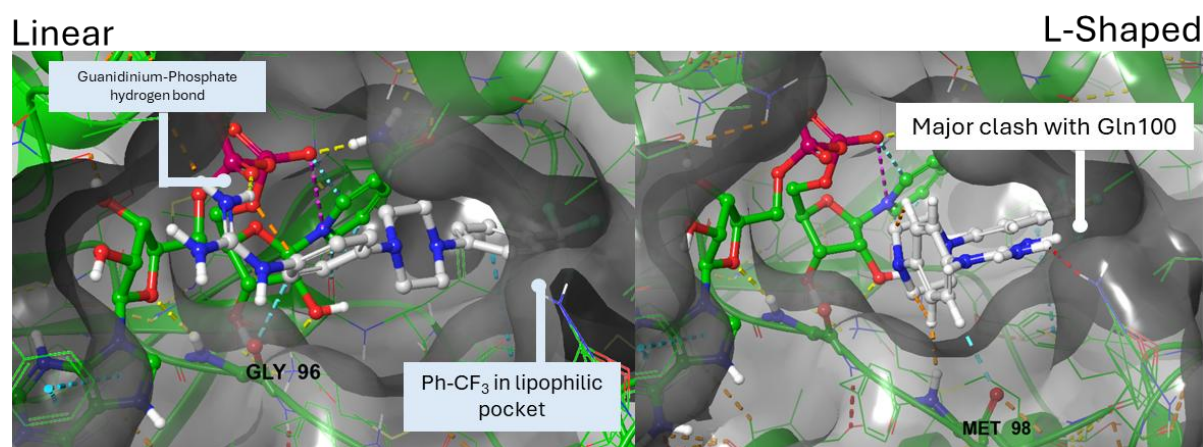


Figure 21 – Left) Linear conformer binding pose in InhA. Right) L-shaped conformer's binding pose in InhA.

Finally, the last issue to consider in docking experiments is receptor flexibility. Typical ligand-protein interaction studies rely on a single rigid receptor (as already shown); nonetheless, this approach only has a success rate between 50 and 75% of generating a biologically relevant binding mode (depending on the docking software). However, it is obvious that upon binding, both the ligand and protein undergo conformational changes.¹³³ Generally, the use of a crystal structure in docking for new drug discovery efforts has found that cross-docking (*i.e.* docking of a non-native ligand into a crystal structure solved in the presence of a different ligand) biases the active site towards the native ligand.¹³⁴ To overcome this issue, allowing for side chain flexibility or whole system dynamics to capture this conformational change is standard practice within academia and industry (though rigid proteins are still used due to the vast number of degrees of freedom in protein flexibility).^{135, 136} For the purposes of the present work, whole system flexibility enabled by molecular dynamics (MD) simulations was not feasible for time and resources constraints. Additionally, induced fit approaches are popular and allow receptor conformational

changes upon ligand binding, however again, they were not possible in the present work due to time limitations.¹³⁷ Thus, a flexible side chain approach was undertaken instead given AutoDock Vina's ability to incorporate selective side chain flexibility. While AutoDock Vina can technically accommodate any number of flexible side chains, computational cost increases exponentially and raises the chances of false positives.¹³⁸ Because of this, most successful flexible side chain dockings only select two to five residues.¹³⁹ However, it has been shown convincingly that only several key residues undergo conformational changes upon ligand binding. For example, across a non-redundant dataset of 188 proteins, in both bound and unbound forms, 90% of cases show at most 5 rotamer changes accounting for conformational movements in response to ligand binding.¹⁴⁰ Also, it has been found that keeping the backbone of a protein rigid while allowing for some side chain flexibility is consistent with ligand-protein binding in nature across a set of 305 proteins with 2369 ligand-bound and 1679 unbound crystal structures.¹⁴¹

To the best of our knowledge, there has not been an established flexible ligand-flexible side chain docking approach to InhA in the literature and therefore one was developed to suit the present case. The first set of flexible side chains chosen were the catalytic triad members Tyr158 and Phe149. As previously mentioned, crystallographic analysis has shown Try158 and Phe149 can exist in an 'in' and 'out' conformation when bound with a substrate and no substrate respectively, therefore binding affects their conformations and binding pose. Lys158 has been shown to fluctuate very little as it is primarily interacting with NADH.^{121, 142} Therefore, Tyr158 and Phe149 were made flexible in docking studies with lead compound **1**. Upon introducing side-chain flexibility, the G-score obtained did not change significantly *i.e.* from -8.8 kcal mol⁻¹ to -8.9 kcal mol⁻¹ (from grey to green conformations – *Fig. 22*) though there was a slight reorientation of the piperazine ring. There was also no major change in the HB formed between the guanidinium and phosphate groups at 2.20 Å, probably because Tyr158 and Phe149 are already in their out conformation given the InhA crystal structure was resolved with a ligand bound. Nevertheless, the substitution on the phenyl ring with different ligands may result in the disruption of π -stacking interactions formed by

Phe149 and Tyr158 and, hence, it was decided to keep these residues flexible in future docking experiments.

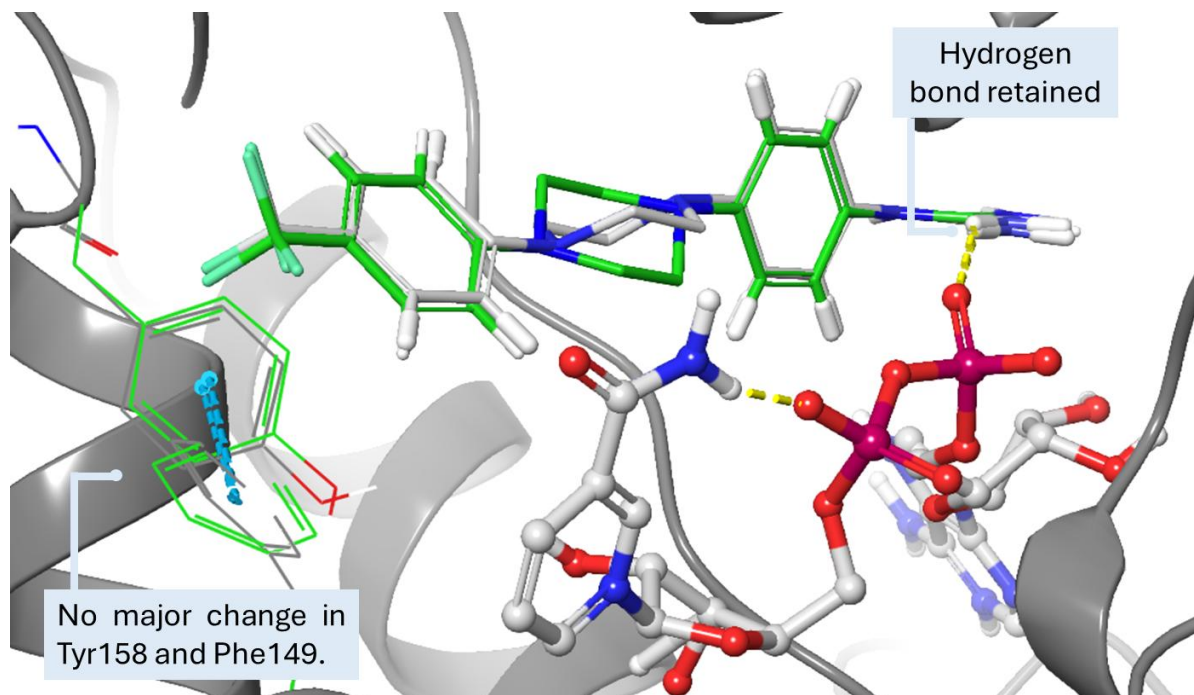


Figure 22 – Binding pose of flexible ligand-rigid protein (grey) in comparison to the flexible ligand-flexible protein (Tyr158 and Phe149) binding pose of compound **1** in InhA.

Next, different residues in the hydrophobic pocket were considered, particularly those found on the loop regions of InhA as these residues are most likely to exhibit conformational changes upon ligand binding because loop regions lack of stable secondary structures such as α -helices or β -sheets.¹⁴³ When introducing flexibility in Met103 (Fig. 23), the G-score again did not change significantly *i.e.* $-8.90 \text{ kcal mol}^{-1}$ to $-8.92 \text{ kcal mol}^{-1}$ with the residue shifting towards the phenyl ring of **1** but with no change in the overall binding mode. Thus, it was determined that Met103 was already in its optimum rotamer form. Making Ile215 flexible resulted in a small increase in G-score from $-8.9 \text{ kcal mol}^{-1}$ to $-8.78 \text{ kcal mol}^{-1}$; however, it resulted in the loss of the HB between guanidinium and phosphate groups, pushing **1** slightly out of the hydrophobic pocket and away from Tyr158 and Phe149. This left the guanidinium with little interactions with the active site, sitting near the entrance of the major portal; hence, it was determined that this binding mode was not biologically relevant (Fig. 23).

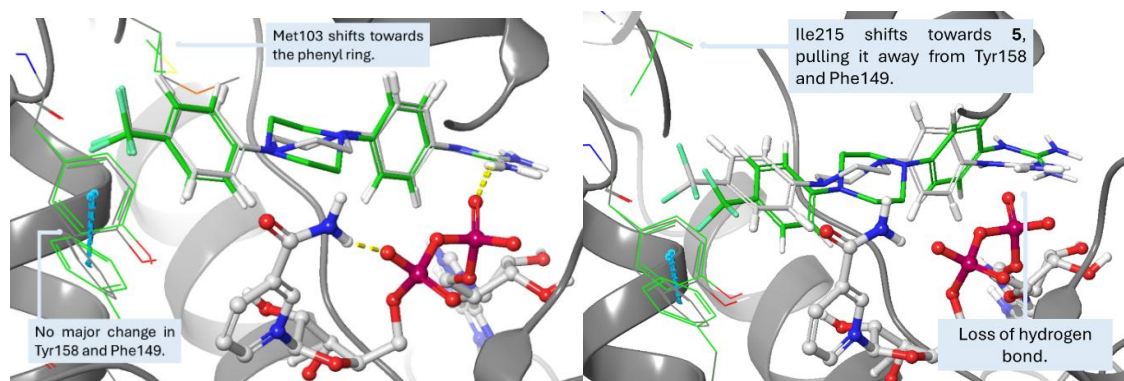


Figure 23 – Left) Flexible docking (green) of **1** in InhA with Phe149, Tyr158 and Met103 flexible (grey = rigid side chains). Right) Flexible docking (green) of **1** in InhA with Phe149, Tyr158 and Ile215 being flexible.

After several iterations introducing flexibility in Met199, Met155 and Met98 with little to no changes in the overall binding mode or G-score, it was found that by allowing Phe97 to be flexible, a bridged HB- π -cation intermolecular bonding network was formed between the phosphate group of NADH, the guanidinium moiety and Phe97; a well-established phenomenon in biological systems involving a guanidinium moiety (Fig. 24).¹⁴⁴ The resultant G-score improves from -8.9 to -9.1 kcal mol⁻¹ and represents a biologically relevant binding mode which provides a plausible explanation for the similar activity of **1** and rifampicin in *M. smegmatis*. Allowing the nearest residue to Phe97, Met98, to be flexible did not disrupt this intermolecular bonding network. Therefore, for subsequent docking experiments, this docking protocol will be employed.

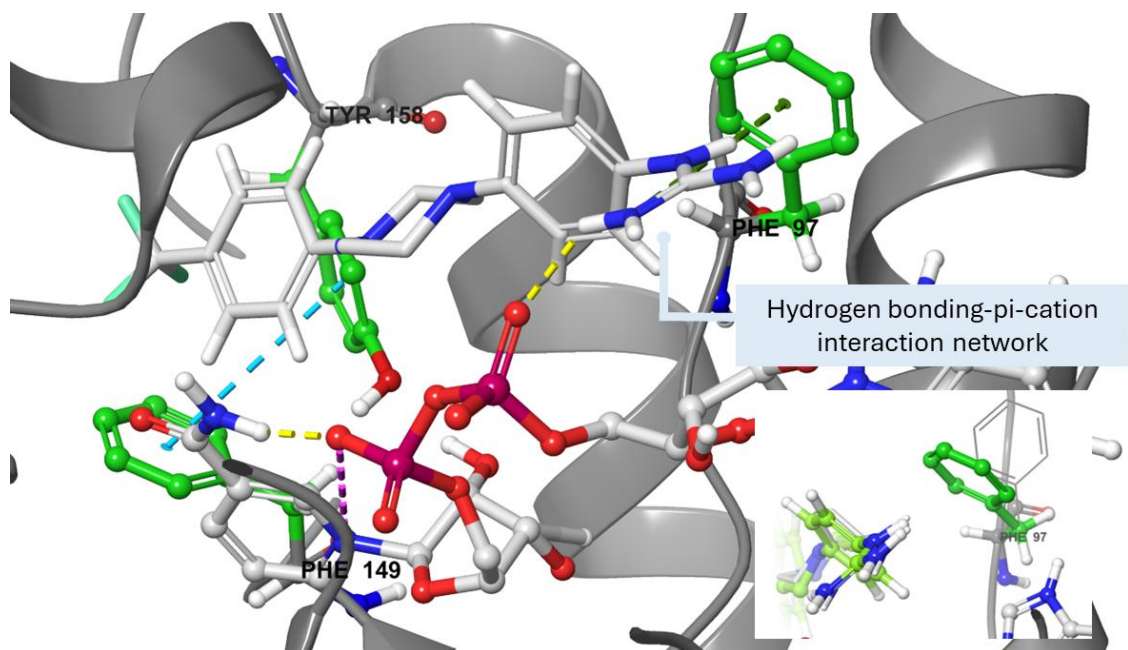


Figure 24 - Flexible docking of **1** in InhA with Phe97, Phe149 and Tyr158 being flexible. Hydrogen bonding-pi-cation interaction network can be seen between phosphate, guanidinium and Phe97. The rotation of Phe97 from its rigid rotamer can be seen in the bottom left.

3.2.3 Docking Investigations of Novel Derivatives

The geometries of the proposed novel derivatives (Fig. 16), categorised as lipophilic pocket and alkylguanidine derivatives, were first optimised at the M06-2X/6-311+G(d,p) (M06-2X/6-311+G(d,p) [light atoms] + LanL2DZ [Iodine] for -I derivative) computational level, with solvent effects included by means of a continuum model. Next, they were docked into InhA (PDB ID: 5OIR) in rigid receptor and flexible receptor protocols as outlined in Section 3.2.2. The results of these docking experiments can be seen in Table 12. In terms of rigid receptor docking, the G-scores for both lipophilic pocket derivatives and alkylguanidine derivatives are near an overall average value of 8.9 kcal mol⁻¹ indicating similar affinity for InhA as **1**. To avoid redundant analysis, the 2D interaction diagram of **7**, the -SF₅ derivative, can be seen in Fig. 25 as representative of this group of compounds. In general, the lipophilic portion of the molecule sits in the substrate binding pocket interacting with lipophilic residues (Met103, Phe149, Tyr158, Ala198, Met199, Ile202, and Ile215 and Leu218) while the guanidinium moiety interacts with the phosphate group of NADH (exception being **8**

[-SCH₃] where the highest G-score pose has the opposite orientation with the guanidinium sitting in the lipophilic substrate binding pocket).

*Table 12 – Docking experiment G-scores (kcal mol⁻¹) of lipophilic pocket and alkyl chain derivatives of **5** obtained using AutoDock Vina 1.2. Ligands optimised at the M06-2X/6-311+G(d,p) (M06-2X/6-311+G(d,p) [light atoms] + LanL2DZ [Iodine] for -I derivative) computational level, with solvent effects included by means of a continuum model.*

	R-Substituent	G-Score (kcal mol ⁻¹)	
		Rigid	Flexible
1	-CF ₃	-8.8	-9.1
Lipophilic Pocket Derivatives			
2	-CH ₂ CH ₃	-9.1	-8.5
3	-isoPr	-8.7	-9.2
4	-cycloPr	-9.0	-9.2
5	-cycloBu	-9.0	-9.6
6	-3-oxetyl	-8.5	-8.9
7	-SF ₅	-9.0	-9.8
8	-SCH ₃	-8.5	-7.1
9	-I	-8.2	-8.2
Alkylguanidine Derivatives			
10	-Butyl	-9.1	-9.2
11	-Hexyl	-9.1	-9.4
12	-Octyl	-9.3	-9.5

The lowest scoring derivative, compound **9** (-I), showed to be the weakest binding ligand of InhA compared to **1** (G-score = -8.2 kcal mol⁻¹) despite being predicted to exhibit similar lipophilicity to -CF₃. This may stem from the loss of the HB between the guanidium and phosphate; however, a π - π stacking interaction is predicted to occur between the phenyl-iodide moiety and Phe149. A more likely reason though may be the poor parametrization of halogen atoms and halogen bonds (possible between the phenyl iodide moiety and Tyr158) that the empirical docking algorithm of AutoDock Vina employs. There has been a modified AutoDock Vina code which changes the van der Waal radii of halogen as well as halogen bonding parameters to refine the empirical scoring function termed XBSF (Halogen Bonding Scoring Function).¹⁴⁵ However, unfortunately due to time restraints this could not be explored.

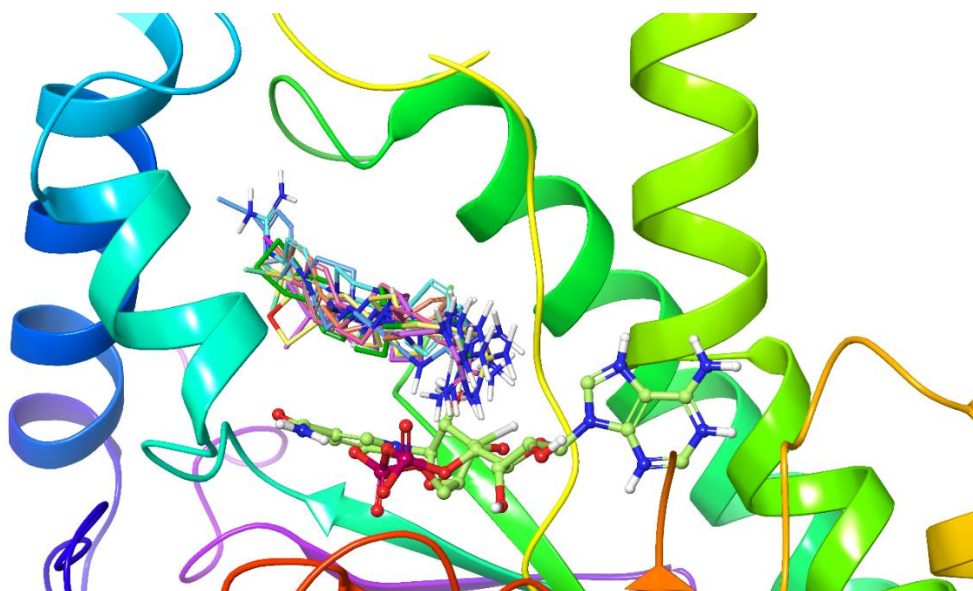
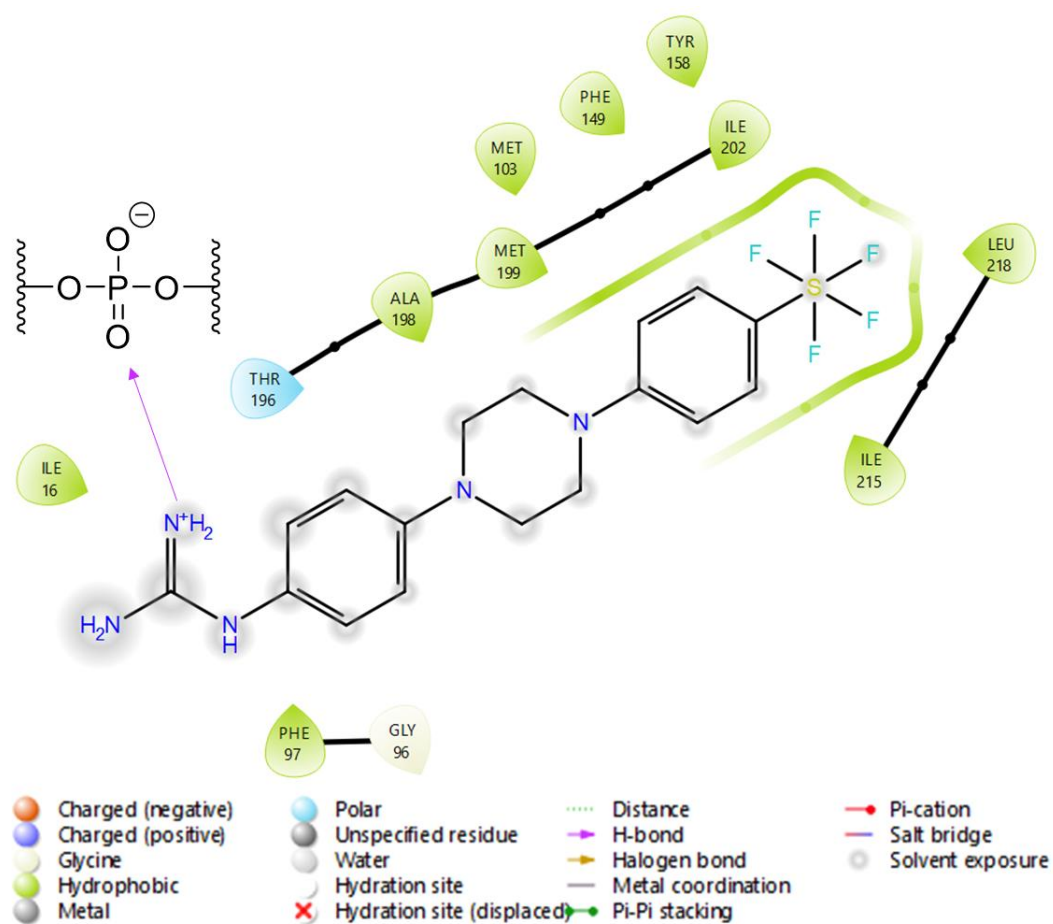


Figure 25 – Top) 2D Interaction diagram of **7** (-SF₅) with InhA. Bottom) Overlay of the top G-score binding poses generated by AutoDockVina for lipophilic pocket derivatives with InhA.

Analysing the alkylguanidine derivatives, only **10** (-Butyl) follows the general binding mode established previously, but, in the case of **11** (-Hexyl) and **12** (-Octyl), the highest G-score poses predict the alkyl guanidinium moiety to sit in the substrate binding pocket while the lipophilic portion of the molecule points towards the major portal (the 2D interaction diagram of **12** can be seen in Fig. 26).

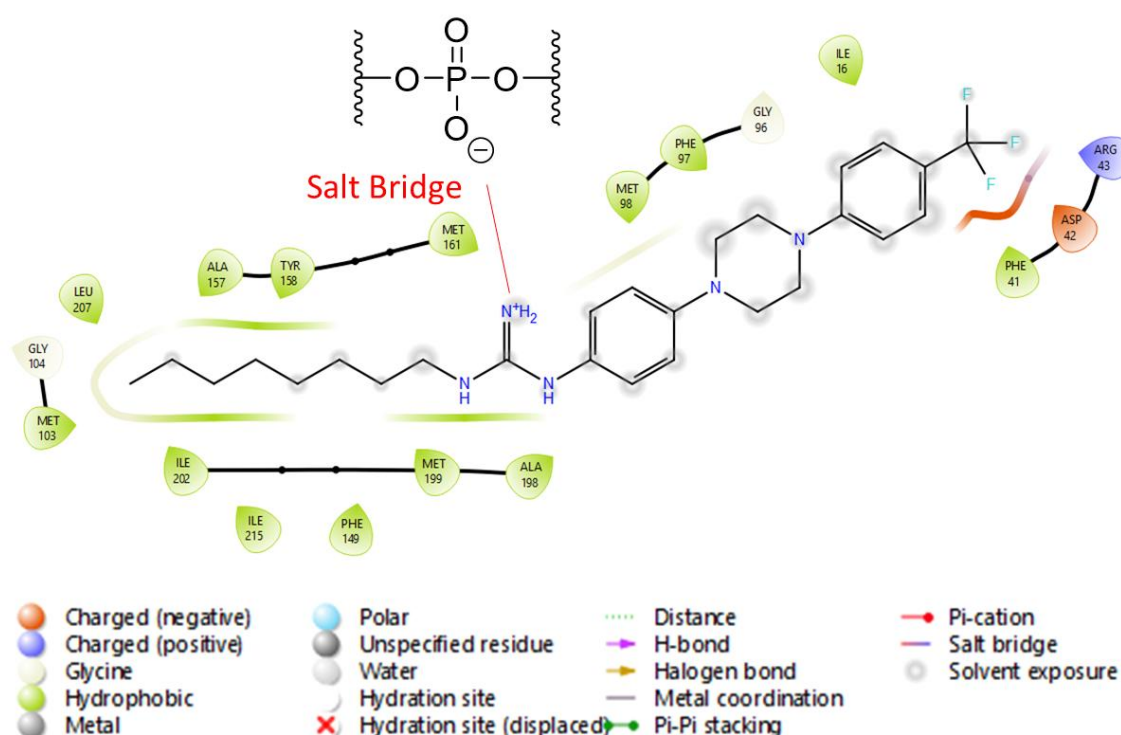


Figure 26 - 2D Interaction diagram of **12** (-Octyl) with InhA.

Within these binding modes the formation of a salt bridge between the guanidinium moiety and phosphate group of NADH (the strongest single non-covalent interaction) can be observed, while the alkyl chain curls itself into the substrate binding pocket. It should be noted, however, that the ‘original’ binding mode of the lipophilic pocket derivatives is manifested in the alkylguanidine in lower scoring poses. The G-score values of the ‘original’ binding mode for **11** and **12** is $-8.9 \text{ kcal mol}^{-1}$ (compared to $-9.1 \text{ kcal mol}^{-1}$ and $-9.3 \text{ kcal mol}^{-1}$ for the best poses, respectively); therefore, at least using

AutoDock Vina's empirical scoring function, it is not possible to tell which binding pose predominates for this set of derivatives.

Turning to the flexible derivatives, the newly developed flexible receptor model has increased G-scores for most derivatives. In the case of **8** (-SCH₃) and **9** (-I), they have maintained identical values. Notably, in the case of **2** (-CH₂CH₃), allowing for side chain flexibility has decreased the G-score from -9.1 kcal mol⁻¹ to -8.5 kcal mol⁻¹.

Examining the derivatives **2**, **3**, **4** and **5**, distinct binding modes are observed. AutoDock Vina predicts that derivatives **3** (-isoPr) and **4** (-cycloPr) exhibit similar binding affinities as **1**, both ligands possessing a G-score of -9.2 kcal mol⁻¹ similar to that of **5** (-9.1 kcal mol⁻¹). Derivative **2** (-CH₂CH₃) is predicted to have a worse binding affinity (-8.5 kcal mol⁻¹) than **1**. This can be accounted for by its reduced capacity for protein-ligand interactions *i.e.* inability to form any HB or π - π stacking interactions as well as reduced number of hydrophobic contacts with the active site. On the other hand, **5** (-cycloBu) is predicted to have a higher binding affinity (-9.6 kcal mol⁻¹) than **1**, and this is predicted to result from a number of favourable interactions with InhA such as the formation of a HB between the guanidinium and a phosphate group, the formation of a π -cation interaction with Phe97 and the optimal positioning of the cyclobutyl substituent within the hydrophobic binding pocket, maximising van der Waal interactions (2D interaction diagram can be seen in *Fig. 27*). Compound **4** (-cycloPropyl) interacts with InhA in a similar manner, except that instead of forming a π -cation interaction with Phe97 *via* the guanidinium, it is predicted to form a π - π stacking interaction with Phe149.

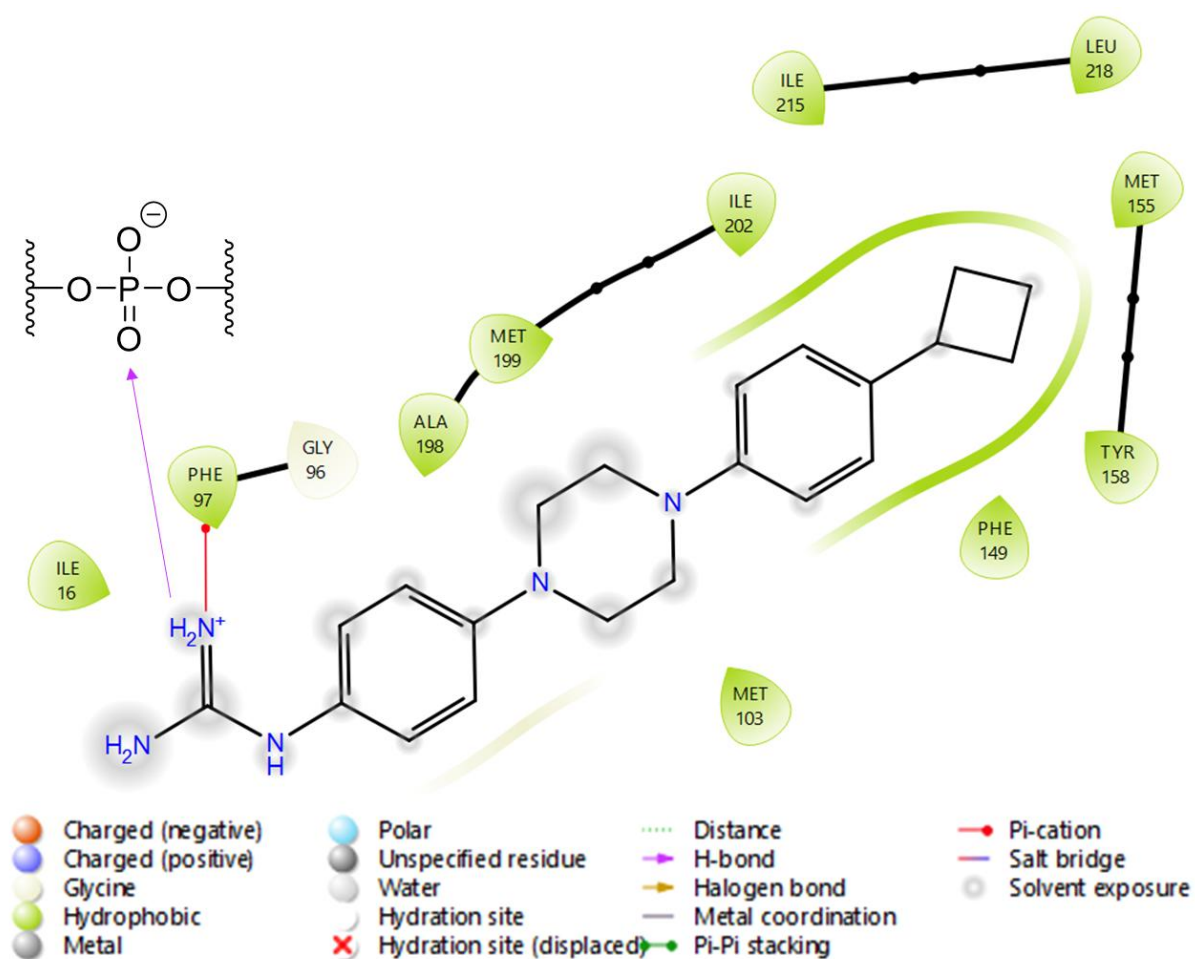


Figure 27 – 2D Interaction diagram of **5** (-cycloBu) with InhA (flexible).

In terms of the sulphur-based derivatives, **7** and **8** have very different G-scores and binding modes. Derivative **15** (-SF₅) is predicted to have a higher binding affinity for InhA compared to **1** (-9.8 kcal mol⁻¹ vs -9.1 kcal mol⁻¹) while **8** (-SCH₃) is predicted to have a much worse binding affinity (-7.1 kcal mol⁻¹) than **1**. Examining the 2D interaction diagram of **7** with InhA (Fig. 28), it is identical to **5** (-cycloBu) with a HB π -cation network established by the guanidinium while the -SF₅ substituent maximises hydrophobic contacts. In the case of **8** (-SCH₃), no HB or π -cation interaction with the protein is found, with the ligand being pushed slightly out of the active site. Regarding **6** (-3-oxetyl), it is predicted to have a similar binding to **1** (-8.9 kcal mol⁻¹ vs. -9.1 kcal mol⁻¹) and while hydrophobic contacts with the lipophilic binding pocket are reduced due to the polar nature of the oxetyl derivative, it is predicted to engage in a π - π and π -cation non-covalent interactions with Phe97. Finally, the iodine-substituted

derivative **9** exhibits no significant change between the rigid receptor and flexible receptor models.

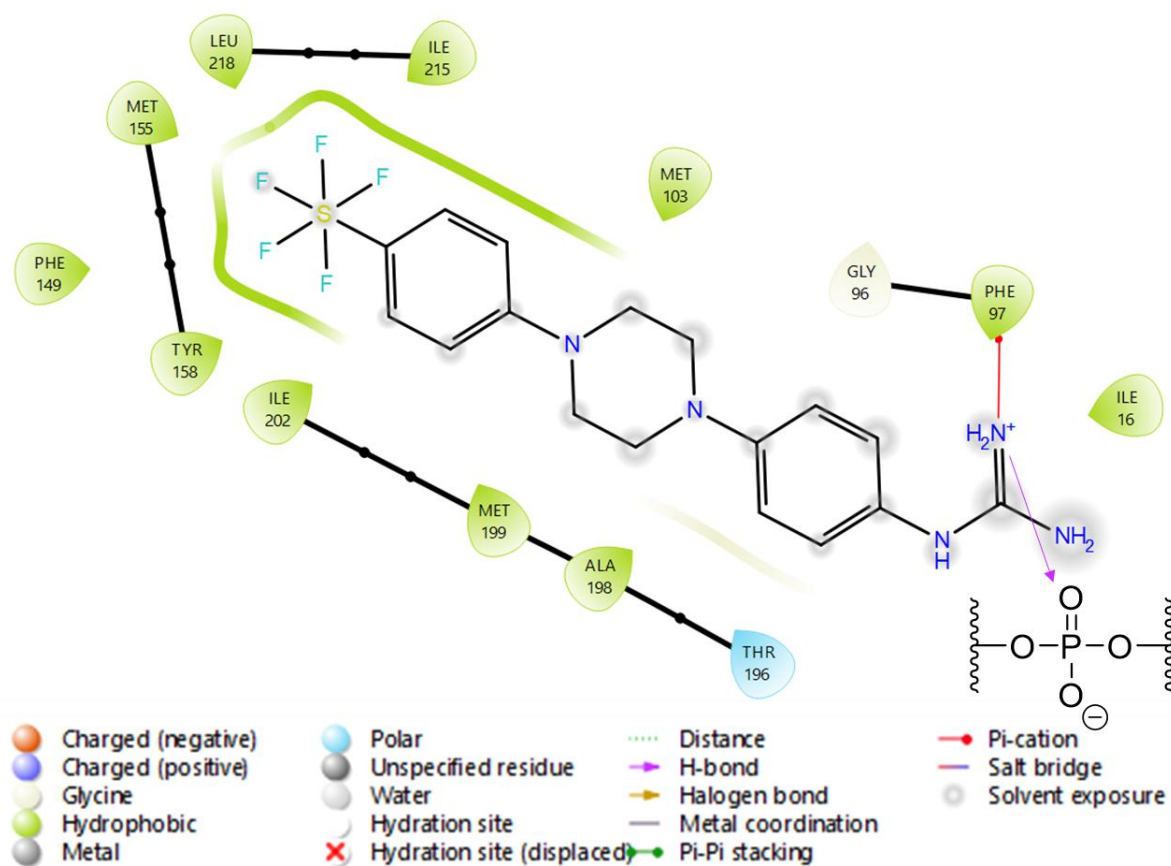


Figure 28 - 2D Interaction diagram of **7** (-SF₅) with InhA (flexible).

Examining the alkyl chain derivatives **10**, **11** and **12**, all ligands exhibit two types of binding modes with similar G-scores: alkyl guanidinium inside the lipophilic substrate binding pocket (In) and alkyl guanidinium facing the major portal (Out). The respective G-scores for each derivative can be seen below in Table 13.

Table 13 - Docking experiment G-scores (kcal mol⁻¹) for alkyl chain derivatives of **5** in their 'In' and 'Out' binding modes obtained using AutoDock Vina 1.2.

	R-Substituent	G-Score (kcal mol ⁻¹)	
		Rigid	Flexible
1	-CF ₃	-8.8	-9.1
Alkyl Chain Derivatives Flexible Docking Results			
		In	Out
10	-Butyl	-9.2	-9.3
11	-Hexyl	-9.4	-9.0
12	-Octyl	-9.5	-9.0

The 2D interaction diagram between **12** (-Octyl) and InhA can be seen in Fig. 29 as representative of the alkyl chain derivative group. In the case of **11** (-Hexyl) and **12** (-Octyl), the 'In' binding modes interact with the receptor more favourably than their 'Out' binding modes as per the G-scores, -9.4 and -9.5 kcal mol⁻¹, respectively, versus -9.0 kcal mol⁻¹ (for both 'Out' binding modes). The likely reason for this higher G-score is the burial of the hexyl and octyl alkyl chains in the substrate binding pocket which then allows for a salt bridge interaction between the guanidinium and phosphate groups which is a stronger interaction than the HB formed by the guanidinium and phosphate in the reverse 'Out' binding mode. For derivative **10** (-Butyl), the alkyl chain is mostly likely not long enough to maximise hydrophobic contacts in the predicted 'In' binding mode as it is in **11** and **12**.

'In' Conformation

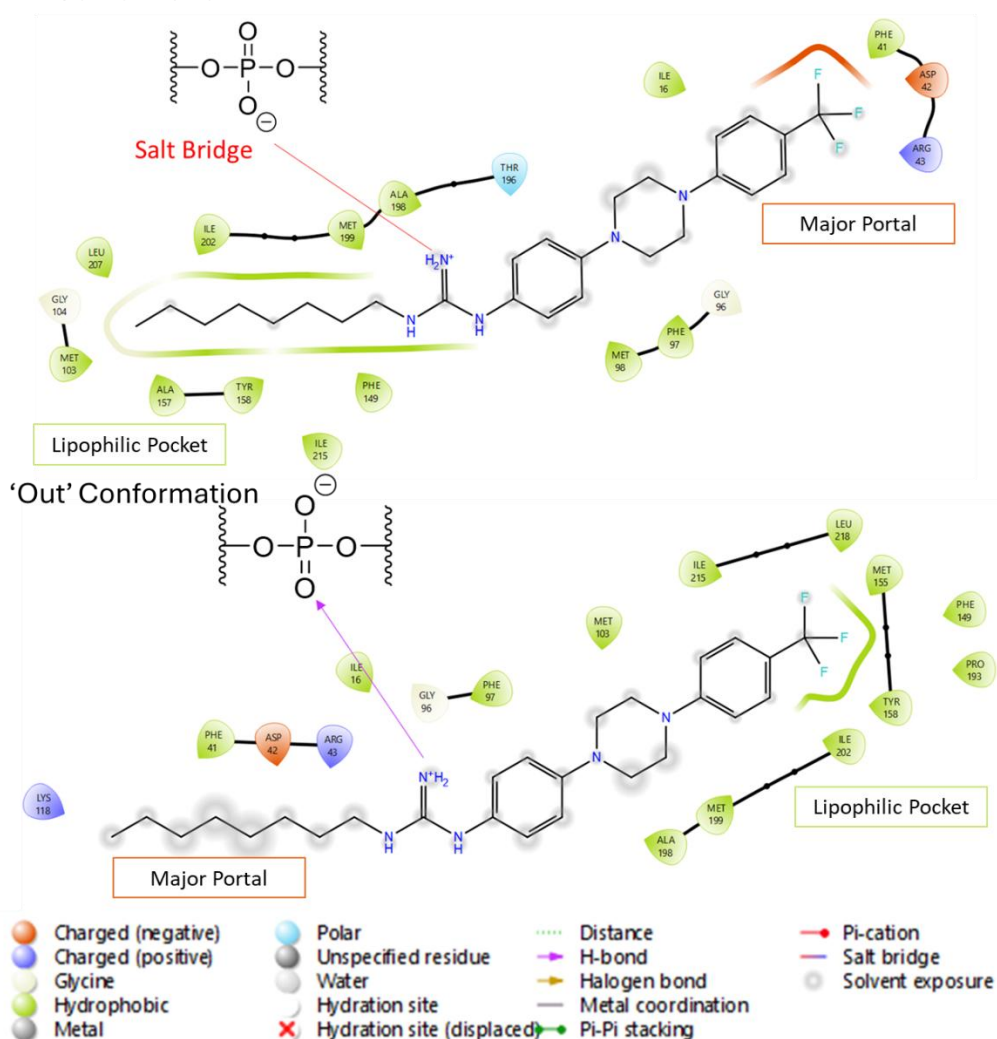
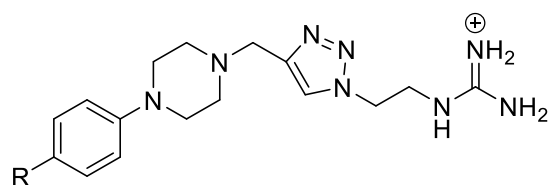


Figure 29 - 2D Interaction diagram for **23** (-Octyl) in its 'In' and 'Out' binding modes with InhA.

The final group of compounds to consider are the following triazole derivatives as can be seen in Fig. 30. Derivative **13** (-CF₃) was proposed to test the effect of replacing the aromatic phenyl group of parent compound **1** while also exploring the flexibility of the guanidinium group. Triazole derivatives **14** (-CH₃), **15** (-isoPr) and **16** (-H) have been proposed, and while they would not optimally fill the lipophilic substrate pocket, they are relatively synthetically accessible in a short space of time and, therefore, they have been included in the search for novel InhA inhibitor structures.

Triazole Derivatives



R = CF₃ (**13**), CH₃ (**14**), isoPr (**15**), H (**16**).

Figure 30 - Structures of proposed triazole derivatives.

The primary motivation behind the replacement of the aromatic group was to explore how a different aromatic group may affect the anti-tubercular activity. Triazoles have many applications in different fields ranging from material sciences to medicinal chemistry, becoming an essential chemical building block.^{146, 147} In anti-tubercular drug discovery, this heterocycle has found applications as a pharmacophoric element in many preclinical candidates (*Fig. 31*). For example, **17** is a derivative of delamanid that shows promising activity *via* DrpE1 inhibition,¹⁴⁸ and compounds **18**, **19** and **20** were reported to be non-covalent InhA inhibitors.¹⁴⁹⁻¹⁵¹ All this reported data, has inspired the replacement of the phenyl ring with a triazole.

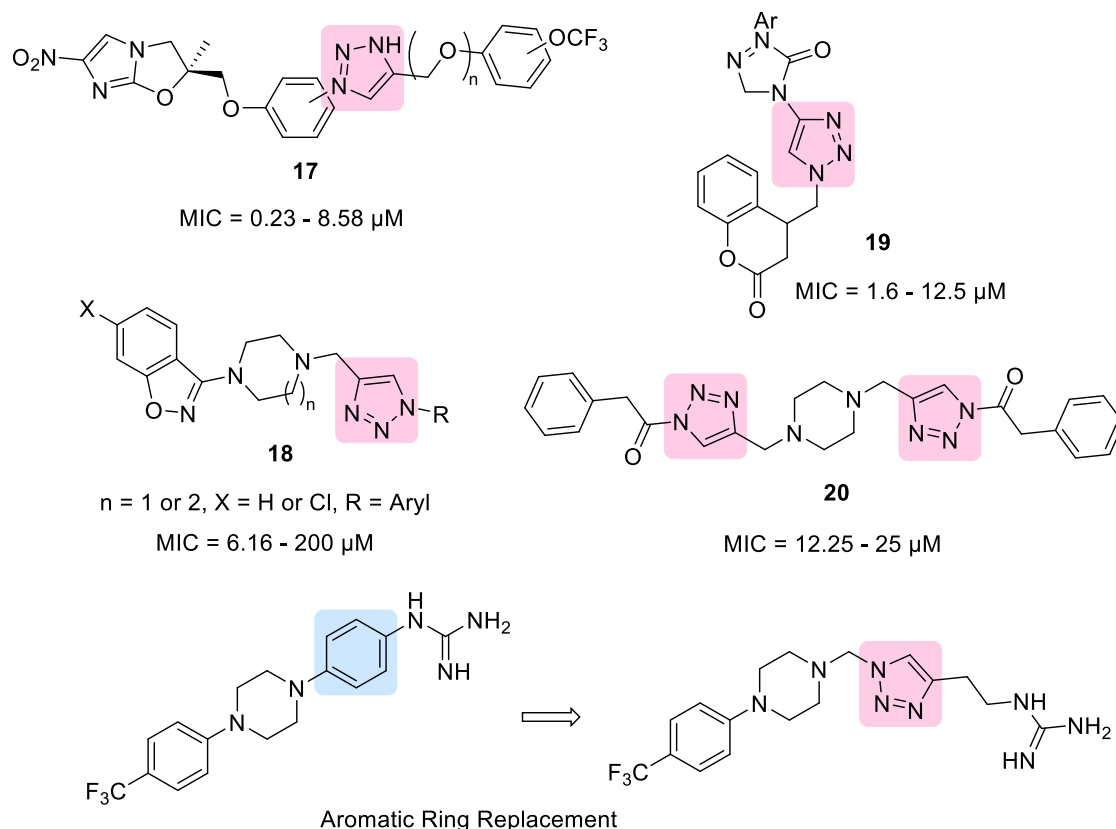


Figure 31 - Structures of triazole containing anti-tubercular pre-clinical candidates and the proposed bioisosteric replacement of the phenyl ring.

Thus, the proposed novel triazole derivatives have been docked into InhA (PDB ID: 5OIR) using the rigid receptor protocol (and flexible receptor for **13** [-CF₃]) as outlined in Section 3.2.2. Geometry optimisation was carried out at the M06-2X/6-311+G(d,p) computational level, with solvent effects included by means of a continuum model. There are two distinct binding modes close in G-score that these derivatives exhibit when docked into InhA and the results of these docking experiments can be seen in Table 14. Derivative **13** (-CF₃) and **15** (-isoPr) is predicted to have similar binding affinity for InhA as **1** while derivatives **14** (-CH₃) and **16** (-H) are predicted to have a worse binding affinity (which is expected given the results of the previous section).

Table 14 – Docking experiment G-scores (kcal mol^{-1}) for triazole derivatives of **5** in their binding mode 1 and mode 2 with InhA obtained using AutoDock Vina 1.2.

	R-Substituent	G-Score (kcal mol^{-1})	
		Rigid	Flexible
1	-CF ₃	-8.8	-9.1
Triazole Derivatives Docking Results			
		Mode 1 Rigid (Flexible)	Mode 2 Rigid (Flexible)
13	-CF ₃	-9.3 (-9.2)	-9.3 (-9.2)
14	-CH ₃	-8.7	-8.5
15	-isoPr	-8.9	-8.7
16	-H	-8.3	-8.3

Binding modes 1 and 2 are depicted in Fig. 32, using compound **13** as representative of the triazole derivatives. In both binding modes, the lipophilic portion of the molecule occupies the lipophilic binding pocket as previously observed with the phenyl derivatives. However, the additional flexibility existing in these derivatives allows for the triazole-alkyl-guanidine moiety to either swing left or right of the major portal. In binding mode 1, the alkyl guanidinium moiety occupies a polar pocket composed of Thr17, Ser19, Ser20 and Thr196, whereas in binding mode 2, the triazole forms a HB with Met97 as the guanidinium is located above it, being in contact with the Gln100 residue, an unlikely position for this cation.

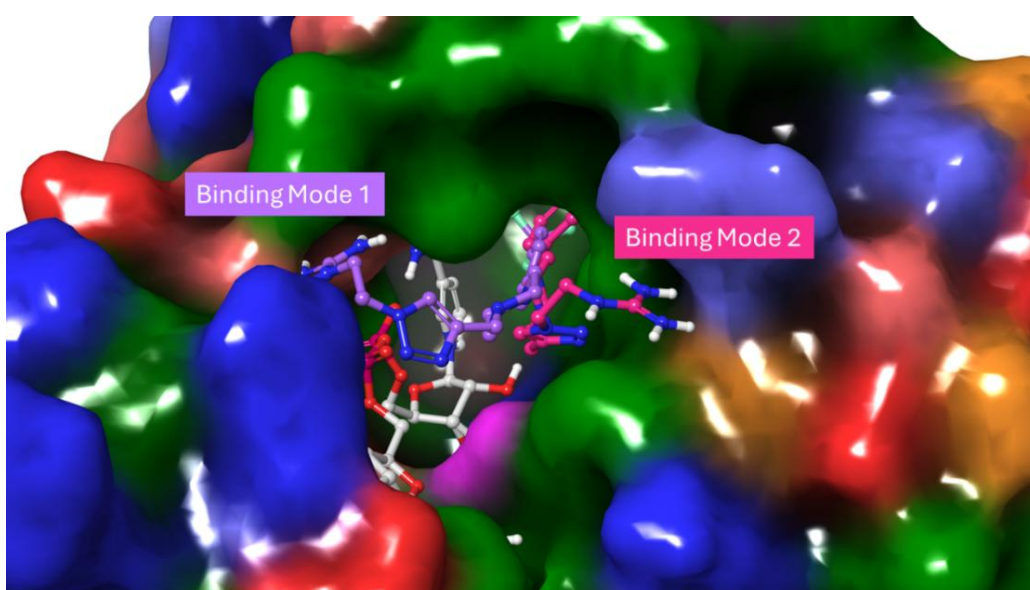


Figure 32 – Binding mode 1 (purple) and 2 (pink) of triazole derivatives of **9** in InhA (derivative **13**) shown as representative of this group of compounds).

However, the ability of the guanidinium group to form strong interactions with water molecules *via* HB networks on the surface of proteins is well established;¹⁵² therefore, it is logical to expect these structures to participate in such interactions given their position at the mouth of the major portal of InhA.

The final exercise of molecular modelling in this project involved alkylguanidinium derivative of **1** (compound **21**) that was synthesised serendipitously (see Section 3.3) in the form of a dialkylated guanidinium (*Fig. 33*).

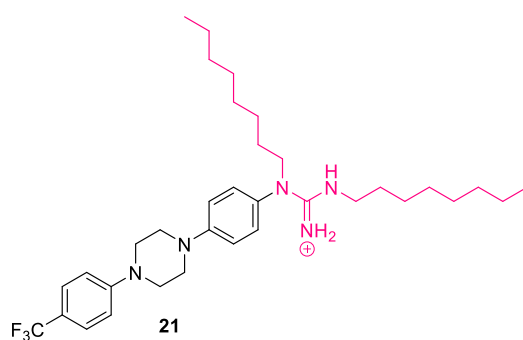


*Figure 33 – Structure of dialkylated guanidine derivative **21**.*

The structure of derivative **21** was optimised, as the other ligands, at the M06-2X/6-311+G(d,p) computational level, with aqueous solvent effects included by means of a continuum model and docked into InhA (PDB ID: 5OIR) in a rigid receptor approach using AutoDock Vina 1.2. The top 8 poses generated were all within a 1 kcal mol⁻¹ range, from -8.8 to -8.7 kcal mol⁻¹, and their overlayed structures can be seen in *Fig. 34*.

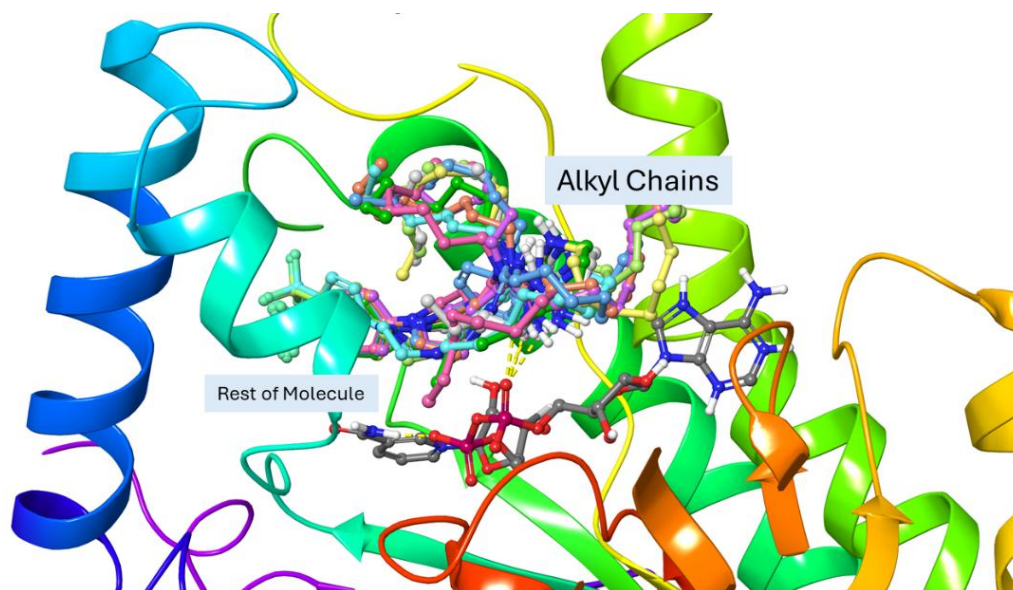


Figure 34 - Overlay of rigid poses of alkyl chain derivative **21** in InhA.

The binding mode of **21** follows the general binding mode established by the lipophilic pocket derivatives and, in some poses, HBs with the phosphate group of NADH is predicted to occur. The poses all differ mostly in the conformation of alkyl chains which is understandable given the high degree of freedom associated with long alkyl chains. It is known that AutoDock Vina performs well for ligands with up to nine rotatable bonds; however, it struggles with more flexible ligands.¹²⁶ Given this fact, AutoDock Vina with its empirical scoring function is not the best suited docking programme to study such a flexible ligand that presents 18 active torsions. Unfortunately, due to resource and time constraints, a more suitable commercial induced fit physics-based docking software could not be used to further probe the binding mode of **21**.

Summarising, over the course of the molecular modelling studies undertaken, several objectives were achieved:

- Establishment of a refined methodological framework for investigating the theoretical binding affinity of derivatives of compound **1** for InhA; in particular, a flexible side chain docking protocol was developed by incorporating Phe98, Phe149 and Tyr158 residues, thus enabling the discovery of biologically relevant binding modes of those derivatives, which explain their activity.

- The evaluation of lipophilic pocket derivatives resulted in the identification of several promising candidates with enhanced predicted binding affinity to InhA compared to the parent compound **1**. In general, derivatives **5** (-cycloBu) and **7** (-SF₅) demonstrated superior G-scores (-9.6 and -9.8 kcal mol⁻¹, respectively) than **1** (-9.2 kcal mol⁻¹). This improvement may be attributed to maximisation of hydrophobic contacts within the binding pocket.
- Evaluation of alkylguanidine derivatives **11** (-Hexyl) and **12** (-Octyl) revealed the possibility of dual binding modes due to the lipophilic alkyl chain. In terms of the triazole derivatives, compound **13** (-CF₃) is predicted to have comparable binding affinity as **1**.
- The highly flexible dialkylated derivative **21** highlighted the limitations of Vina's empirical scoring function for systems with a large number of torsions.

3.2.4 Theoretical Pharmacokinetic Profile Assessment

Drug design and development is an expensive and lengthy process with a high attrition rate. One study in the US market reported that the mean cost of developing a drug between 2000 to 2018 was US\$ 221 million; inclusion of drug development failure, rises the mean cost to US\$ 662 million (figures adjusted to 2025 dollar value).¹⁵³ Also, it has been estimated that only 1 in 10 drug candidates succeed clinical trials.¹⁵⁴ While one major reason for this small success is poor efficacy, which can relate to the validation of the biological target, another reason is linked to the chemical structure of the drug itself which can lead to bad clinical outcomes in terms of absorption, distribution, metabolism and excretion (ADME) as well as safety.¹⁵⁵ That means, how the structure of drug influences its pharmacokinetics within the body which directly influences its *viability* as a drug candidate. Pharmacokinetics describes the movement of the drug throughout the body *i.e.* its ADME profile. This has led to the development of some guidelines to adhere to in the search for 'drug-like' molecules with good oral availability *i.e.* searching for molecules with potential to succeed.¹⁵⁶ In tuberculosis drug discovery, specifically, unfavourable physicochemical properties lead to a high number of clinical failures;¹⁵⁷ for this reason, *in silico* assessment of ADME properties is now wide-spread and usually

occurs at the hit-to-lead optimisation stage, filtering out structures with potential for adverse pharmacokinetic properties.¹⁵⁸

Hence, what exactly are these ADME principles? Several formulations of these principles have been iterated throughout the years but the first comprehensive attempt was carried out by Lipinski in his seminal 1996 review, wherein it is predicted poor oral activity of drugs when their chemical structure has: (i) more than 5 hydrogen bond donors (HBD), (ii) more than 10 hydrogen bond acceptors (HBA), (iii) a molecular weight (MW) greater than 500 g mol⁻¹ and (iv) a calculated partition coefficient (cLogP) greater than 5.¹⁵⁹ Further developments of this rule were made by Ghose, who proposed that drug-like molecules should have a cLogP of -0.4 to 5.6, a MW between 150 and 480, a molar refractivity (MR – total polarizability of a mole of a substance) between 40 and 130 and the total number of heavy atoms (#HA) to be between 20 and 70.¹⁶⁰ Additionally, Veber included the requisite of fewer than 10 rotatable bonds (#RB) and a polar surface area less than 140 Å² for compounds with potential as oral drugs.¹⁶¹ However, it is important to note that the rules established by Lipinski, Ghose and Veber for good oral bioavailability should be regarded as guidelines rather than strict rules, with the ‘beyond rule of five’ (bRo5) space attracting more and more interest in drug discovery and many compounds which violate these guidelines have become good drug candidates.¹⁶² These bRo5 compounds include large synthetic macrocycles, peptides and natural products.^{163, 164} In any case, given the difficulties in treatment regimens of *M. tb.* (Section 1.1.3) and the fact that the main objective of this work is to find small molecule inhibitors of InhA, an assessment of drug-likeness was undertaken for the proposed derivatives of **1**. Physicochemical and pharmacokinetic properties for these derivatives were generated using the SwissADME computational suite (<http://www.swissadme.ch/>).¹⁰⁰

3.2.5 Evaluation of ‘Lipinski-Ghose-Veber’ Profile of Proposed Derivatives

The calculated physicochemical properties *i.e.* those proposed by Lipinski, Ghose and Veber as important for oral bioavailability, were calculated for all derivatives proposed and the results can be seen in *Table 15*; violations are presented in red. All

lipophilic pocket derivatives, including **1**, adhere to the Lipinski, Ghose and Veber guidelines indicating that they may have good oral bioavailability. This does not extend to all alkylguanidine derivatives where **11** (-Hexyl) and **12** (-Octyl) have 11 and 13 #RB respectively and a MR higher than 130, breaking both of Veber's guidelines. Dioctyl derivative **21** has a MW of 588.8 and a ClogP larger than 7.65 thus breaking Lipinski and Ghose's guidelines; also, it has 20 #RB and a MR of 181.32 also breaking Veber's rules. This implies that **21** may have poor oral bioavailability.

Table 15 - Physicochemical properties of all derivatives proposed as calculated by SwissADME [ClogP – Consensus partition coefficient as an average of all logP calculators as implemented in SwissADME]. Violations of the rules are indicated in red.

	R-substituent	MW (g mol ⁻¹)	HBA	HBD	CLogP	#RB	TPSA	MR	#HA
1	0	364.4	3	3	2.42	5	70.12	104.21	26
Lipophilic Pocket Derivatives									
2	-CH ₂ CH ₃	324.4	0	3	2.01	5	70.12	108.99	24
3	-isoPr	338.5	0	3	2.34	5	70.12	113.79	25
4	-cycloPr	336.5	0	3	2.16	5	70.12	111.68	25
5	-cycloBu	350.5	0	3	2.49	5	70.12	116.49	26
6	-3-oxetyl	352.5	1	3	1.50	5	79.35	112.76	26
7	-SF ₅	422.4	5	3	2.79	5	95.42	110.61	28
8	-SCH ₃	342.5	0	3	1.90	5	95.42	110.93	24
9	-I	422.3	0	3	2.10	4	70.12	111.93	23
Alkyl Chain Derivatives									
10	-Butyl	420.5	3	3	3.85	9	56.13	123.54	30
11	-Hexyl	448.6	3	3	4.52	11	56.13	133.15	32
12	-Octyl	476.6	3	3	5.27	13	56.13	142.76	34
21	-diOctyl	588.8	3	2	7.65	20	47.34	181.32	42
Triazole Derivatives									
13	-CF ₃	397.4	6	3	0.92	8	100.83	106.42	28
14	-CH ₃	343.5	3	3	0.22	7	100.83	106.38	25
15	-isoPr	371.5	3	3	0.79	8	100.83	115.99	27
16	-H	329.4	3	3	-0.12	7	100.83	101.41	24

The high cLogP value (7.65) of compound **21** is of concern because it indicates the potential for rapid metabolic turnover, low solubility and poor absorption, moreover considering what is suggested in the literature about a logP value of 1 to 3 being ideal

for good ADMET profiles.¹⁶⁵ In the context of anti-tubercular drug discovery, it has been shown that lipophilic compounds are better at killing *M. tb.* than their hydrophilic counterparts and high lipophilicity is associated with granuloma penetration and bactericidal activity against latent/dormant *M. tb.*¹⁶⁶ Therefore, the lipophilicity of anti-tubercular agents is one of the most important physicochemical properties to consider as it not only affects the ADMET properties of a drug, but also it directly influences its ability to act as an anti-infective against *M.tb.* All triazole derivatives obey the established guidelines except in the case of **13** (-CF₃) which has more than five HBAs breaking Lipinski's guidelines.

Fraction sp³ (*i.e.* number of sp³ hybridised carbons as a fraction of total carbon count) is a relatively novel physicochemical property used as a measure of the complexity of molecular structure, and it was first proposed by Lovering.¹⁶⁷ The calculated Fsp³ can be seen in *Table 16* and is recommended by SwissADME to be at least 0.25 for optimal saturation of drug candidates. All derivatives meet this criterion except for **9** (-I) which has a Fsp³ of 0.24.

3.2.6 Evaluation of the Theoretical Pharmacokinetics of Proposed Derivatives

SwissADME also provides an implementation to predict and evaluate the individual ADME profile of any molecule under investigation. The predictions for passive human gastrointestinal absorption (HIA) and blood-brain barrier (BBB) permeation are presented in the BOILED-Egg model. This model relates the predicted logP of a compound and its total Polar Surface Area (PSA) to its permeation in a statistically robust manner using 880 known compounds and their experimentally determined ability to permeate gastrointestinal and brain tissue and, therefore, it establishes boundaries between these two properties where brain or gastrointestinal permeation is more likely.¹⁶⁸ The results of this analysis can be seen in *Table 16* in addition to other properties such as the fraction of sp³ character (Fsp³), whether the derivative is a substrate of the permeability glycoprotein (P-gp sub.), the solubility as well as the guanidine (piperazine) pK_{aH} values as calculated using the ChemAxon Marvin package (discussed later in the section). Accordingly, it is predicted that all derivatives have a

high probability of being absorbed into the bloodstream *via* the gastrointestinal (GI) tract except for **14** (triazole, -CH₃), **16** (triazole, -H) and **21** (-diOctyl) which are predicted to have low GI absorption. Furthermore, it is predicted that **1** (-CF₃) and all mono alkylated guanidine derivatives **10** (-Butyl), **11** (-Hexyl), **12** (-Octyl) would be able to cross the BBB while the rest of the derivatives cannot. This is worrisome given that in principle, penetration of the BBB is undesirable as it can lead to neurotoxicity in the central nervous system and reduce therapeutic efficacy overall against *M. tb*. However, this is not always a bad thing when it comes to anti-tubercular agents, as isoniazid (first line anti-TB drug that targets InhA) penetrates the BBB and is used to treat tubercular meningitis.¹⁶⁹

Table 16 – Theoretical ADME properties and Fsp³ character of proposed derivatives as calculated by SwissADME. pK_{aH} values were calculated by ChemAxon Marvin package [Solubility key: VS – Very Soluble, S – Soluble, MS – Moderately Soluble, PS – Poorly Soluble].

	R-substituent	Fsp3	HIA	BBB	P-gp sub.	Solubility	pK _{aH}
1	-CF ₃	0.28	High	Yes	Yes	MS	12.88
Lipophilic Pocket Derivatives							
2	-CH ₂ CH ₃	0.32	High	No	Yes	S	12.88
3	-isoPr	0.35	High	No	No	MS	12.88
4	-cycloPr	0.35	High	No	Yes	S	12.88
5	-cycloBu	0.38	High	No	Yes	MS	12.88
6	-3-oxetyl	0.35	High	No	Yes	S	12.88
7	-SF ₅	0.24	High	No	No	MS	12.88
8	-SCH ₃	0.28	High	No	Yes	S	12.87
9	-I	0.24	High	No	No	MS	12.88
Alkylguanidine Derivatives							
10	-Butyl	0.41	High	Yes	No	MS	13.22
11	-Hexyl	0.46	High	Yes	No	MS	13.2
12	-Octyl	0.5	High	Yes	No	PS	13.15
21	-diOctyl	0.62	Low	No	Yes	PS	13.28
Triazole Derivatives							
13	-CF ₃	0.47	High	No	Yes	S	12.39 (6.23)
14	-CH ₃	0.47	Low	No	Yes	S	12.39 (6.26)
15	-isoPr	0.53	High	No	Yes	S	12.39 (6.25)
16	-H	0.44	Low	No	Yes	VS	12.39 (6.24)

Another factor affecting the pharmacokinetics of drugs is whether they are substrates of the permeability protein 1 (P-gp) which is an ATP-driven drug efflux pump that affects drug distribution throughout the body.¹⁷⁰ This transporter is conserved throughout the animal kingdom and is expressed in a wide variety of organs in their barrier tissues such as in the liver, kidneys, GI and also the brain.¹⁷¹ In SwissADME, this prediction is made *via* a machine learning model based on a training set of 1033 substrates/non-substrates validated against a test set of 415 substrates/non-substrates with approximately equal amounts of each in each set.¹⁰⁰ Consequently, while **9** is predicted to cross the BBB based on its physicochemical properties, it is also predicted to be a P-gp substrate which may result in active efflux from the brain. Other compounds predicted to act as substrates for P-gp are the lipophilic pocket derivatives **2** (-CH₂CH₃), **4** (-cycloPr), **5** (-cycloBu), **7** (-SF₅) and **8** (-SCH₃), derivative **21** (-diOctyl) and all triazole derivatives.

Two more factors considered in this study affecting the pharmacokinetics of a drug are its water solubility and ionizability. Solubility in water is an important property which affects not only the absorption within the body *via* oral intake, but also the ease of handling and future formulation of the drug.¹⁷² Water solubility in SwissADME is given as a qualitative assessment based on a scale of solubility which is calculated for a compound based on a system of 16 fragmental contributions modulated by the squared root of MW.¹⁰⁰ All derivatives studied are predicted to be moderately soluble to soluble except for **12** (-Octyl) and **21** (-diOctyl), which are predicted to be poorly water soluble, while **16** (triazole, -H) is predicted to be very soluble. In terms of ionizability, all structures in this work contain an ionisable guanidine moiety, which as previously mentioned in Section 1.2., is exclusively in its protonated guanidinium form *in vivo*. The mean pK_{aH} of the lipophilic pocket derivatives is 12.88 while the alkylguanidine derivatives are predicted to have a higher pK_{aH} ranging from 13.15 to 13.28, reflecting the inductive stabilisation of the cation by the alkyl substituents. Triazole pK_{aH} values can be seen in brackets ranging from 6.23 to 6.29 and, thus, at blood pH (7.4) will be unionised. This has implications for the distribution of guanidine-based drugs since it is obvious that charged particles struggle to cross

apolar phospholipid cell membranes, even more so in the case of *M. tb.* with its fatty cell wall.¹⁷³ Ionised drugs do in fact cross cell membranes but, rather than through passive diffusion, they utilise active transport mechanisms such as the previously mentioned P-gp.¹⁷⁴ An interesting transport mechanism of cationic amphiphilic drugs is 'degradative transport', in which the protonated moiety catalyses phospholipid ester hydrolysis, generating single-chain lipids that fragment and disperse, carrying drug molecules across the bilayer with them.¹⁷⁵

The final factor affecting the pharmacokinetics of drugs is their interaction with cytochromes P450 (CYP) metabolic enzymes, which are a large group of isoenzymes which are reportedly involved in 90% of all enzymatic reactions within the body. Furthermore, three fourths of all enzyme-related drug metabolism is caused by CYP 1A2, 2C19, 2C9, 2D6 and 3A4 (CYP3A enzyme representing the largest fraction of this, led by 3A4).¹⁷⁶

Therefore, to assess CYP-related metabolism of the proposed derivatives, a SwissADME implemented algorithm was used. This algorithm has been trained on 35,539 compounds that are inhibitors and non-inhibitors of the aforementioned CYP enzymes with roughly equal amounts of each.¹⁰⁰ The results of this analysis (*Table 17*) indicate that none of the proposed derivatives are inhibitors of CYP2C9 and inhibition of CYP1A2 is not expected except for derivative **8** (-SCH₃). Similarly, only derivative **7** (-SF₅) is predicted to act as a CYP3A4 inhibitor, and regarding CYP2C19, derivative **5** (-cycloBu) together with the alkylguanidine derivatives (**10**, **11**, **12** and **21**) are predicted to exert inhibitory effects.

Finally, inhibition of CYP2D6 is predicted for all lipophilic pocket derivatives, including **1**, as well as all mono-alkylated guanidine derivatives. CYP2D6 metabolism involves the addition or removal of certain functional groups, specifically the hydroxylation of phenyl rings as well as demethylation and dealkylation.¹⁷⁷ Inhibition of this enzyme is likely to be the most significant factor compromising the pharmacokinetic profile of our derivatives in the human body as it can reduce the available level of compound within blood plasma and, therefore, not reach its intended destination.

Table 17 - Theoretical profile of Cytochrome P450 enzymes of the proposed derivatives as determined by SwissADME.

	R-substituent	CYP1A2	CYP2C19	CYP2C9	CYP2D6	CYP3A4
9	-CF ₃	No	No	No	Yes	No
Lipophilic Pocket Derivatives						
2	-CH ₂ CH ₃	No	No	No	Yes	No
3	-isoPr	No	No	No	Yes	No
4	-cycloPr	No	No	No	Yes	No
5	-cycloBu	No	Yes	No	Yes	No
6	-3-oxetyl	No	No	No	Yes	No
7	-SF ₅	Yes	No	No	Yes	Yes
8	-SCH ₃	No	No	No	Yes	No
9	-I	No	No	No	Yes	No
Alkyl Chain Derivatives						
10	-Butyl	No	Yes	No	Yes	No
11	-Hexyl	No	Yes	No	Yes	No
12	-Octyl	No	Yes	No	Yes	No
21	-diOctyl	No	Yes	No	No	No
Triazole Derivatives						
13	-CF ₃	No	No	No	No	No
14	-CH ₃	No	No	No	No	No
15	-isoPr	No	No	No	No	No
16	-H	No	No	No	No	No

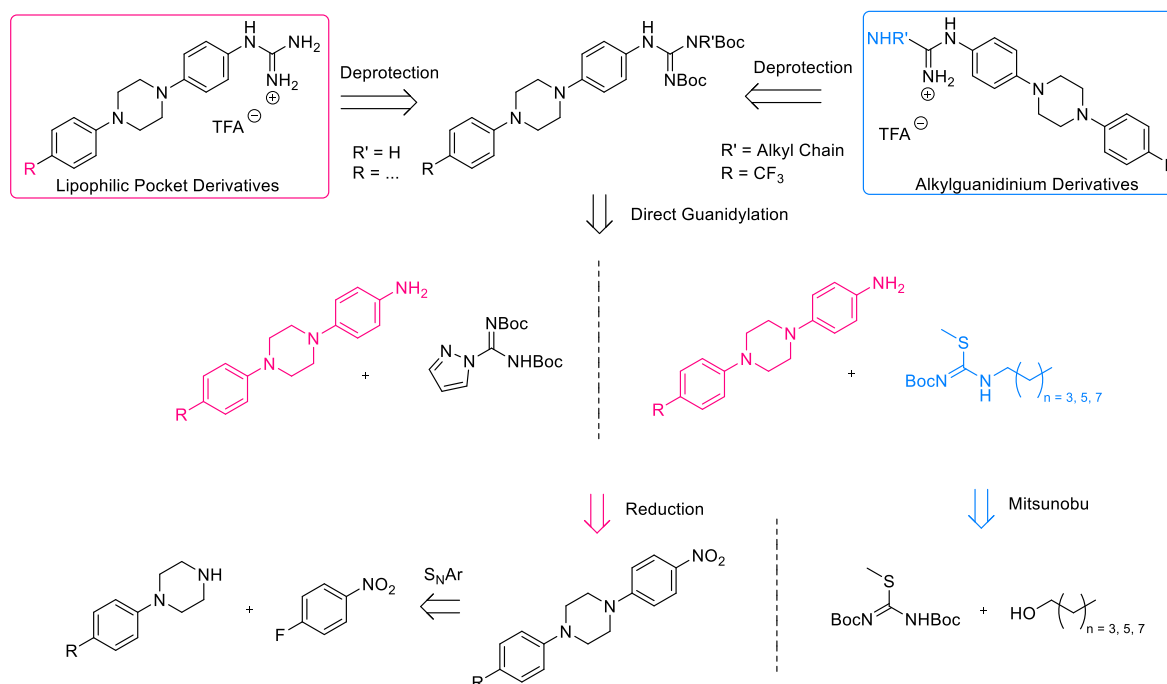
3.3 Chemical Synthesis

3.3.1 Introduction and Retrosynthetic Analysis

Based on the computational studies performed, the new derivatives were synthesised using already existing methodologies developed by the Rozas group. The general philosophy of the retrosynthetic analysis (*Scheme 1*) is as follows: the substituted 1,4-diarylpiperazine guanidinium and alkylguanidinium trifluoroacetate salts are obtained from their respective protected guanidines. The protected guanidines are themselves formed from the reaction of substituted 4(4-phenylpiperazin-1-yl)anilines (pink) and the respective guanidylating reagents. The aniline derivatives are the reduced forms of the preceding 1-(4-nitrophenyl)-4-

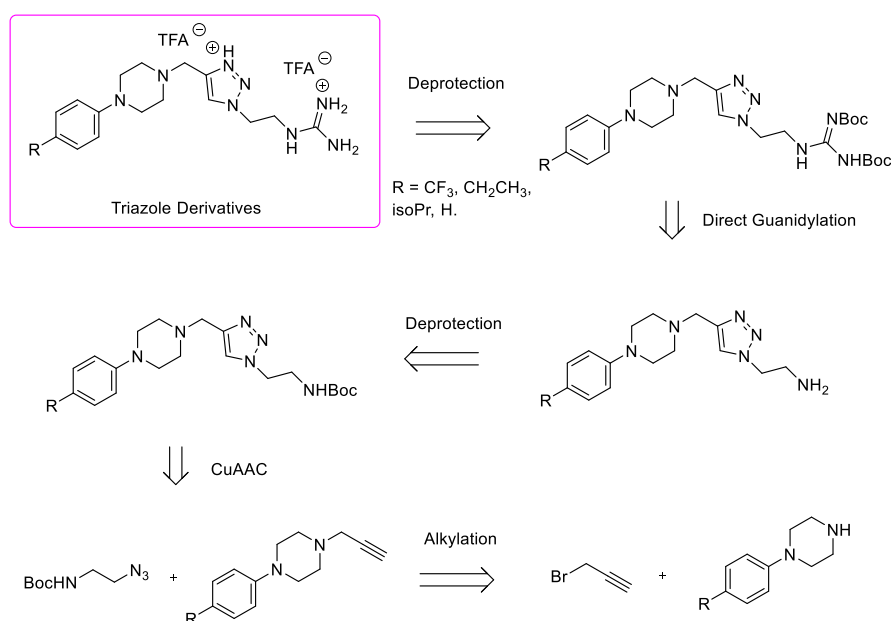
phenylpiperazines which are formed by the nucleophilic aromatic substitution reaction (S_NAr) between the respective substituted phenyl piperazine and 4-fluoro-nitrobenzene.

Scheme 1 - Retrosynthetic analysis of the proposed lipophilic pocket and alkyl chain guanidine-based derivatives.



The retrosynthetic analysis of the triazole derivatives (*Scheme 2*) follows a similar pattern to the previous route described where the trifluoroacetate guanidinium salt is obtained from the respective protected guanidine. This is installed on an alkylamine handle obtained from the formation of the triazole moiety *via* copper-catalysed azide-alkyne cycloaddition (CuAAC) from the respective azide and phenylpiperazine alkyne. The phenylpiperazine alkyne is formed from the alkylation of the respective phenyl piperazine *via* an S_N2 reaction.

Scheme 2 - Retrosynthetic analysis of triazole derivatives.



When needed new methodologies were developed and all synthetic intermediates and final compounds were characterised by ¹H and ¹³C NMR, MS and IR. Final compounds were assessed for purity using HPLC before biological testing with 95% purity set as the cutoff.

3.3.2 Synthesis of Lipophilic Pocket Derivatives

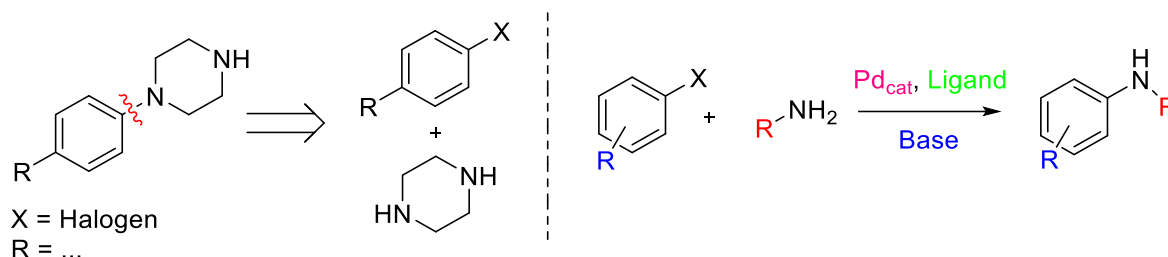
Referring to the *Scheme 1*, the synthesis of the lipophilic pocket derivatives begins with the S_NAr reaction between the substituted phenylpiperazine and 4-fluoro-nitrobenzene. However, apart from 1-(4-trifluoromethylphenyl)piperazine, *none* of the other proposed para-substituted phenylpiperazines are available commercially. Therefore, their synthesis begins the overall synthesis of the lipophilic pocket derivatives.

3.3.2.1 Synthesis of 1-Phenylpiperazine Derivatives

The most obvious disconnection in the preparation of phenylpiperazines is the C-N bond present between the ring moieties (*Scheme 3*), especially given the breadth of C-N bond formation methods present in the literature and its widespread use in small molecule synthesis.¹⁷⁸ In the 1990s, Buchwald and Hartwig independently developed

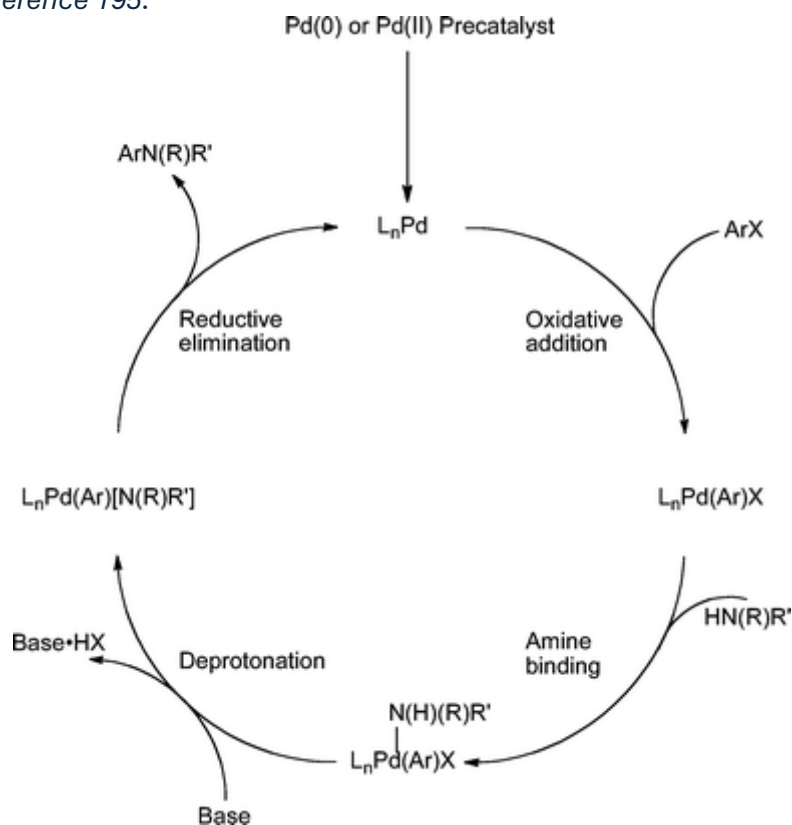
Pd and Cu catalysed N-arylation using suitable catalyst, ligand and base from the reaction of a halobenzene and suitable amine (Scheme 3).^{179, 180}

Scheme 3 - C-N disconnection in phenyl piperazine and the general conditions for the Pd-catalysed Buchwald-Hartwig Amination Reaction.



Therefore, by employing Buchwald-Hartwig amination, the necessary 1-phenylpiperazine derivatives were prepared from the piperazine and the respective commercially available substituted halobenzenes (Scheme 4).

Scheme 4 – Generalised catalytic cycle for Pd-catalysed amination with phosphine ligands. Taken from reference 195.



The reaction proceeds with the oxidative addition of a Pd(0) species into the aryl halide; the afforded complex then is coordinated by the respective amine. This

increases the N-H's acidity and therefore allows for proton abstraction by the base affording a palladium amide complex which undergoes reductive elimination to give the arylated amine product and regeneration of the Pd(0) catalyst.

After screening aryl iodides, protecting groups, Pd catalyst and ligand (Table 18) optimal conditions were determined to be:

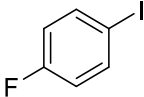
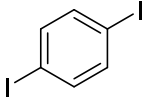
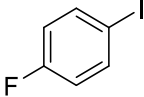
(i) The use of Pd₂(dba)₃ at 2.5% loading was superior to Pd(OAc)₂, which showed no conversion due to insufficient reduction to the Pd(0) species. While piperazine can act as a reductant at temperature >100°C, in this case it was insufficient.

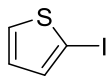
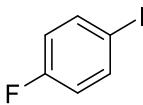
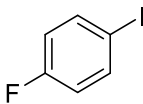
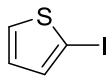
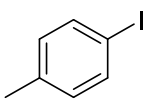
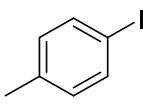
(ii) The XPhos ligand at 5% loading proved superior to Xantphos and what allowed for the reaction to occur compared to partial conversion in entries 1,2 and 5 (as adjudged by crude ¹H NMR). Most likely due to its bulky, electron rich monodentate nature facilitating the reaction steps outlined above.¹⁸¹

(iii) Toluene served as the optimal solvent, specifically chosen for its ability to suppress iodide inhibition by maintain iodide's insolubility in the reaction mixture.¹⁸² CsCO₃ was the preferred base over NaOtBu which gave zero conversion.

(iv) Temperature was increased to 100 °C and the reactions were allowed to proceed under Ar gas over 24 hours. Boc-protected piperazine (72%) performed slightly better than the benzhydryl-protected (65%) piperazine.

Table 18 – Screening of conditions of the Buchwald-Hartwig Amination reaction for subsequent synthesis.

Entry	Aryl Iodide	PG	Pd Source	Ligand	Temp (°C)	Solvent	Base	Conversion
1		Boc	Pd ₂ (dba) ₃	Xantphos	90	PhMe	CsCO ₃	Partial ^a
2		Boc	Pd ₂ (dba) ₃	Xantphos	90	PhMe	CsCO ₃	Partial ^a
3		Boc	Pd(OAc) ₂	Xantphos	90	PhMe	CsCO ₃	None ^a

4		Boc	$\text{Pd}_2(\text{dba})_3$	Xantphos	90	PhMe	CsCO_3	None ^a
5		Benzh.	$\text{Pd}_2(\text{dba})_3$	Xantphos	100	PhMe	CsCO_3	Partial ^a
6		Benzh.	$\text{Pd}_2(\text{dba})_3$	Xantphos	100	PhMe	NaOtBu	None ^a
7		Benzh.	$\text{Pd}_2(\text{dba})_3$	Xantphos	100	PhMe	CsCO_3	None ^a
8		Boc.	$\text{Pd}_2(\text{dba})_3$	XPhos	100	PhMe	CsCO_3	72% ^b
9		Benz.	$\text{Pd}_2(\text{dba})_3$	XPhos	100	PhMe	CsCO_3	65% ^b

^a Catalyst/Ligand Loading: 3/3%, 1.4 eq. of base, Ar and 24 h. As adjudged by crude NMR. ^b Catalyst/Ligand loading: 2.5/5%, 2 eq. of Base, Ar and 24 h. Isolated Yield.

As previously mentioned, only substituted iodobenzenes **22** and **23** are commercially available at a reasonable economic cost. In *Fig. 35*, the prices for an amount of the substituted iodobenzenes of interest can be seen (as per the first result for each compound on sigmaaldrich.com/ie). Compound **24** (-SF₅) is not available, while **25** (-cycloPr) and **26** (-cycloBu) are prohibitively expensive. Derivative **27** (-3-oxetyl) is only available as a bromobenzene, still at a high cost.

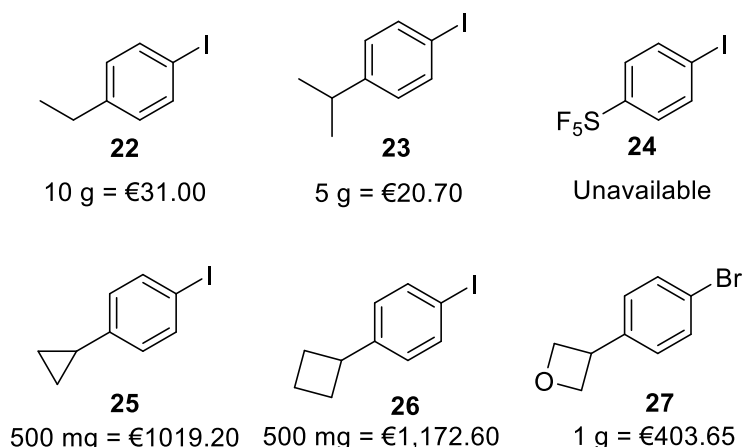
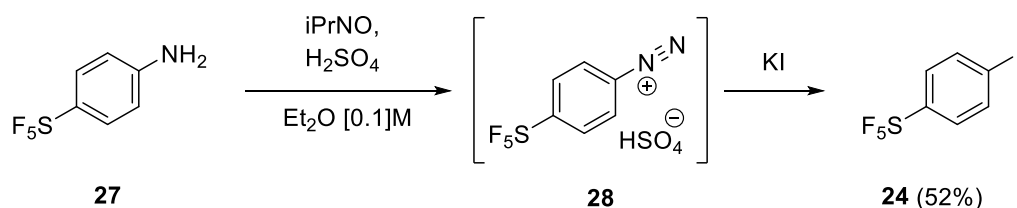


Figure 35 – Cost of an amount of substituted iodobenzenes of interest as per the Sigma Aldrich website.

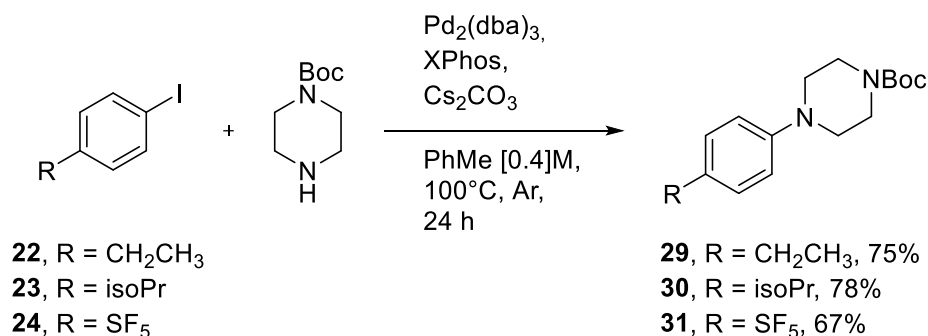
Therefore, there was a need to synthesise the substrates for the Buchwald Hartwig reactions from cheaper starting materials. Beginning with **24**, the much cheaper aniline derivative **27** was used with a diazonium salt to install the functional group of interest. Diazonium salts are one of the most versatile (and explosive) functional groups as they can participate in a plethora of functional group conversion reactions once they are made.¹⁸³ In our particular case, we were interested the Sandmeyer-like reaction with potassium iodide which can convert the diazonium salt into an iodine.¹⁸⁴ Thus, **27** was reacted with sulphuric acid at 0 °C in diethyl ether followed by the slow addition of isopropyl nitrite. When bubbling had subsided, indicating the completion of diazonium salt **28**'s formation, potassium iodide was added and allowed to react at room temperature for 10 minutes. After dilution with more diethyl ether, the organic layer was recovered and washed with sodium thiocyanate to reduce excess iodide. The crude mixture was purified using flash column chromatography (hexane:EtOAc, 1:0 to 9:1) to afford 4-pentafluorosulfonyl-iodobenzene **24** at 52% yield (*Scheme 5*) to use in subsequent Buchwald-Hartwig aminations.

Scheme 5 – Synthesis of 4-pentafluorosulfayl-iodobenzene (24) from 27 via diazonium salt and Sandmeyer-like reaction.



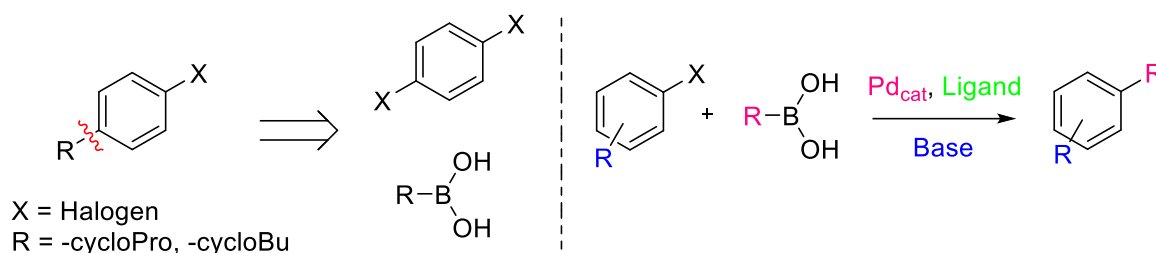
For the synthesis of the cyclopropyl, cyclobutyl and 3-oxetyl derivatives **4**, **5** and **6**, a different strategy was employed; thus, using our optimised Buchwald-Hartwig conditions, substituted phenylpiperazines **29–31** were synthesised in yields ranging from 67 to 78%, ready for subsequent deprotection (Scheme 6).

Scheme 6 – Buchwald-Hartwig amination for the synthesis of phenylpiperazine 29, 30 and 31.



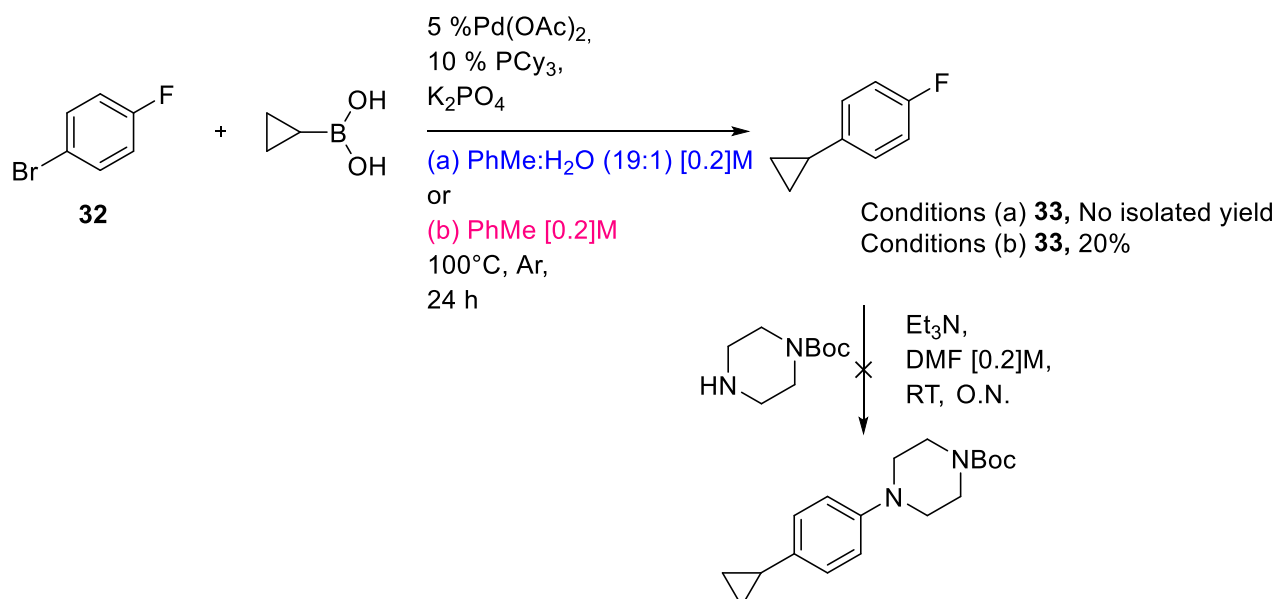
To prepare the required phenylpiperazines for the synthesis of **4** (-cycloPro) and **5** (-cycloBu) it was first required the formation of a $\text{C}(\text{sp}^3)\text{--}\text{C}(\text{sp}^2)$ bond as shown in the retrosynthetic analysis in Scheme 7. One of the most widely used methods for this is the Suzuki-Miyaura cross-coupling reaction between a suitable halogen substrate and the respective boronic acid.

Scheme 7 – Alkyl-substituted-benzene retrosynthesis and General Suzuki-Miyaura cross-coupling reaction scheme.



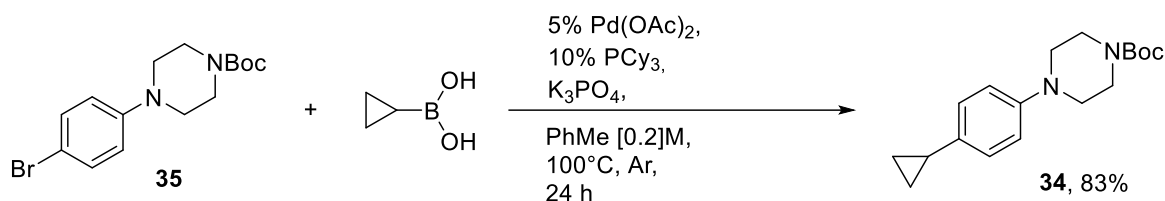
Following the procedure published by Wallace *et al.* for the coupling of cyclopropyl boronic acid with substituted bromobenzenes¹⁸⁵ and using **32** as a model substrate, the corresponding compound was prepared (Scheme 8). The initial conditions used a 19:1 PhMe:H₂O mixture as the solvent system and while crude ¹H NMR showed partial conversion, purification *via* flash column chromatography (1:0 to 9:1 hexane:EtOAc) returned mix fractions given that the R_f of **32** and **33** are identical. Therefore, in a subsequent reaction, water was removed and the purification was carried out for a longer period of time with a eluent concentration gradient of hexane:EtOAc starting at 1:0 and ending at 95:5, allowing for the recovery of some pure **33** product at 20% yield. An S_NAr reaction was attempted between **33** and 1-Boc-piperazine, however the reaction failed.

Scheme 8 – Synthesis of 1-cyclopropyl-4-fluorobenzene **41**



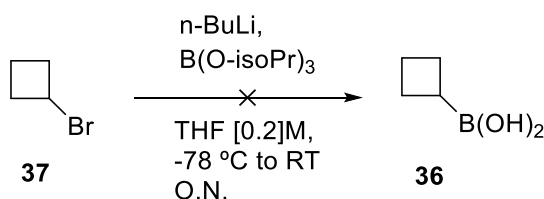
Nevertheless, to synthesise the required cyclopropyl derivative (*i.e.* **34**), the reaction was carried out again using commercially available phenylpiperazine **35** and cyclopropylboronic acid, yielding **34** at 83% (*Scheme 11*). Purification was carried out using a PhMe:EtOAc (1:0 to 95:5) solvent system which was able to discriminate between starting material **35** and the product (as they also have the same R_f), allowing for a smoother albeit longer purification.

Scheme 11 – Synthesis of cyclopropyl-substituted phenylpiperazine 34.



An attempt was made to synthesise the cyclobutyl analogue by preparing the corresponding boronic acid **36** *via* a modified method reported by Sweetman and Guiry for the preparation of boronic acids from bromocyclobutane **37**.¹⁸⁶ Unfortunately, the synthesis was not successful, and this pathway had to be abandoned due to time constraints.

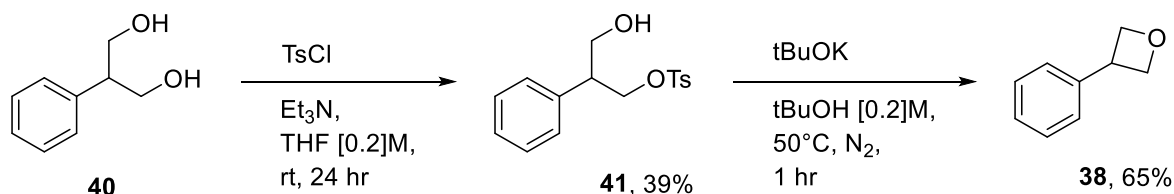
Scheme 12 – Attempted synthesis of cyclobutyl-boronic acid 36.



The last phenylpiperazine whose synthesis was attempted, was that of the oxetyl derivatives. 3-Phenyloxetane **38** and 2-phenyloxetane **39** had been previously synthesised for use in experimental logP determination (Section 3.1.3).¹¹⁰ Compound **38** was synthesised *via* classical intramolecular Williamson ether synthesis, which begins with the mono-tosylation of commercially available 2-aryl-1,3-diol **40** followed

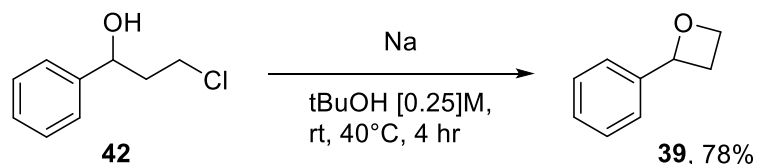
by base promoted cyclisation of the mono-tosylated product **41** (Scheme 13). The crude mixture was diluted with EtOAc, washed with water and brine and purified *via* distillation at 110 °C yielding **38** at 65% yield (Scheme 13).

Scheme 13 - Synthesis of 3-phenyloxetane **38**.



2-Phenyloxetane **39** was similarly synthesised *via* classical Williamson ether synthesis using sodium metal to promote cyclisation from commercially available 3-chloro-1-phenylpropan-1-ol **42** at 78% yield (Scheme 14).

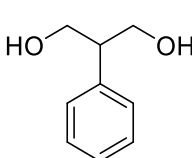
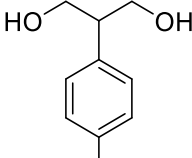
Scheme 14 – Synthesis of 2-phenyloxetane **39**.



To synthesise the corresponding phenylpiperazine required for the preparation of **6** (-3-oxetyl), an iodine has to be installed in the para-position of the phenyl ring. This was attempted on the 2-aryl-1,3-diol **40** using a wide variety of methods found in literature to synthesise 2-(4-iodophenyl)propane-1,3-diol **43** (Table 19). Generally, these methods rely on creating or using an electrophilic iodine species which then undergoes nucleophilic attack by the arene. In the case of entry 1, the use of *N*-iodosuccinamide as an iodination reagent using mild conditions is well described; however, it relies on an activated arene.¹⁸⁷ Unfortunately in our case, **40** is not activated enough to act as a nucleophile. Therefore, in order to make the iodine more electrophilic, the use of mercuric (II) chloride, which has been described as a stoichiometric activating agent,¹⁸⁸ was considered; however, the I₂-HgCl₂ species was

also not electrophilic enough and the reaction failed. Another method reported the use of *N*-iodosuccinimide and trifluoroacetic acid to generate *in-situ* an electrophilic iodine trifluoroacetate species has had success with deactivated substrates.¹⁸⁹ Thus, **40** was reacted with *N*-iodosuccinimide and trifluoroacetic acid, neat, resulting in the formation of **43**. The crude mixture was poured in cold water and extracted with DCM, the combined extracts were washed with a saturated sodium dithionite solution and dried over sodium sulphate and purified using flash column chromatography (hexane:EtOAc 1:0 to 7:3) giving **43** at a 23% yield. Unfortunately, due to time constraints this synthesis could not be brought forward using the strategy used for the synthesis of **38**.

Table 19 - Reaction conditions for the formation of iodo-substituted aryl diol **43**.

 40 Conditions →  43		
Entry	Conditions	Results
1	NIS, MeCN, RT, O.N. ¹⁸⁷	Fail
2	I ₂ , HgCl ₂ , DCM [0.25]M, RT, O.N. ¹⁸⁸	Fail
3	NIS, TFA [0.33]M, 1 hr, 1. ¹⁸⁹	Success, 23%

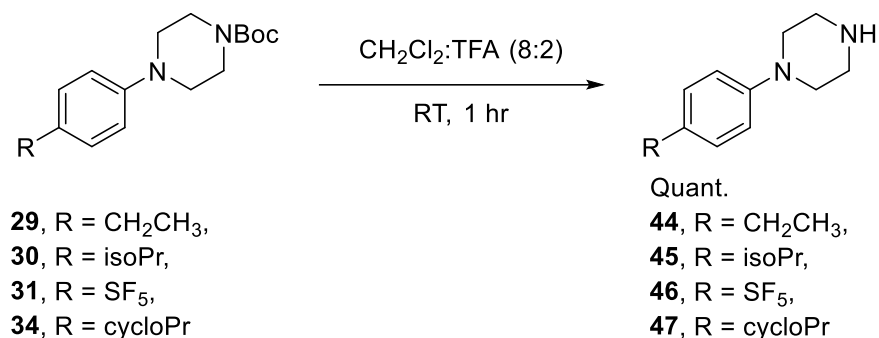
The attempted synthesis of the final lipophilic pocket derivative **9** (-I) is described in Section 3.3.2.2.

3.3.2.2 Synthesis of 1-(4-nitrophenyl)-4-Phenylpiperazine Derivatives

Before the synthesis of the substituted 1-(4-nitrophenyl)-4-phenylpiperazine derivatives can proceed, the preceding substituted 1-phenylpiperazines must be deprotected. Thus, 1-phenylpiperazines **29**, **30**, **31** and **34** were deprotected using

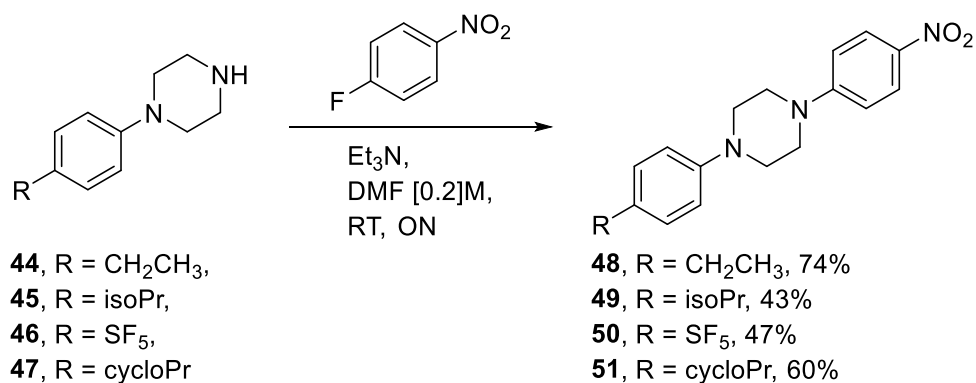
standard conditions in DCM:TFA (8:2) at room temperature for an hour yielding **44-47** quantitatively (Scheme 15).

Scheme 15 - Deprotection of Boc-protected 1-phenylpiperazines **29-31, 34**.



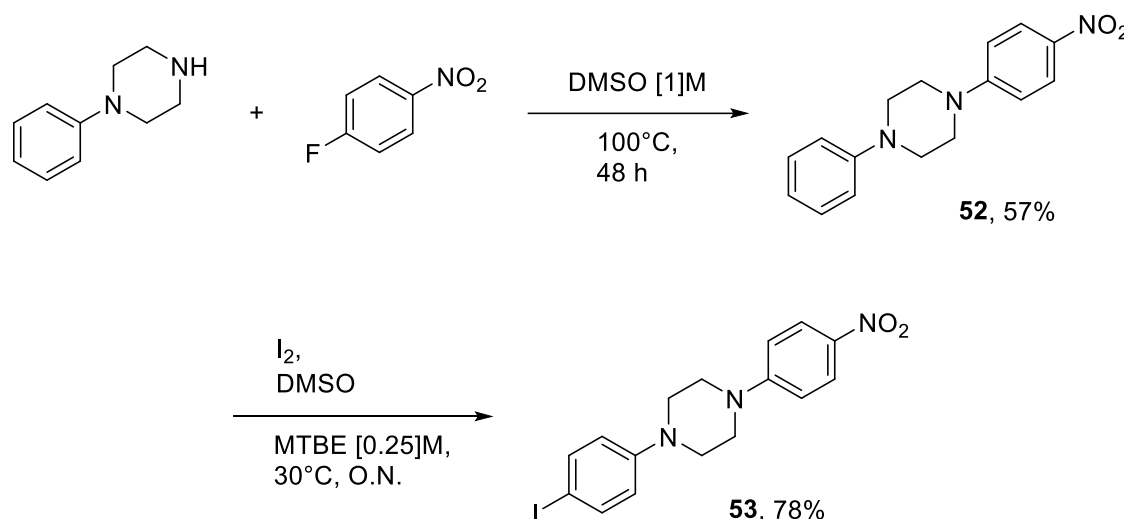
The resultant deprotected 1-phenylpiperazines **44-47** were reacted with 4-fluoronitrobenzene in DMF at room temperature with triethylamine acting as the base in an S_NAr reaction resulting in substituted 1-(4-nitrophenyl)-4-phenylpiperazines **48-51** with yields ranging from 43-74% (Scheme 16). Compounds **48** and **49** were considered pure enough for the following steps while compounds **50** and **51**, which were determined as not sufficiently pure *via* ¹H NMR and were purified using flash column chromatography with hexane:EtOAc (1:0 to 7:3).

Scheme 16 – Synthesis of substituted 1-(4-nitrophenyl)-4-phenylpiperazines **56-59**.



For the synthesis of **9** (-I), the arene 1-(4-nitrophenyl)-4-phenylpiperazine **52** was firstly synthesised *via* an S_NAr with 1-phenyl piperazine and 4-fluornitrobenzene (Scheme 17) at 57% yield. Subsequently, **52** was iodinated according to a modified method reported by Bovonsombat *et al.* wherein DMSO activates the molecular iodine electrophile *via* halogen bonding (milder than the conditions used for **43**).¹⁹⁰ Thus, **52** was reacted with molecular iodine and DMSO in methyl-*tert*-butyl ether yielding **53** at 78% (Scheme 17).

Scheme 17 - Synthesis of 1-(4-nitrophenyl)-4-phenylpiperazine **52** and 1-(4-nitrophenyl)-4-phenylpiperazine **53**.

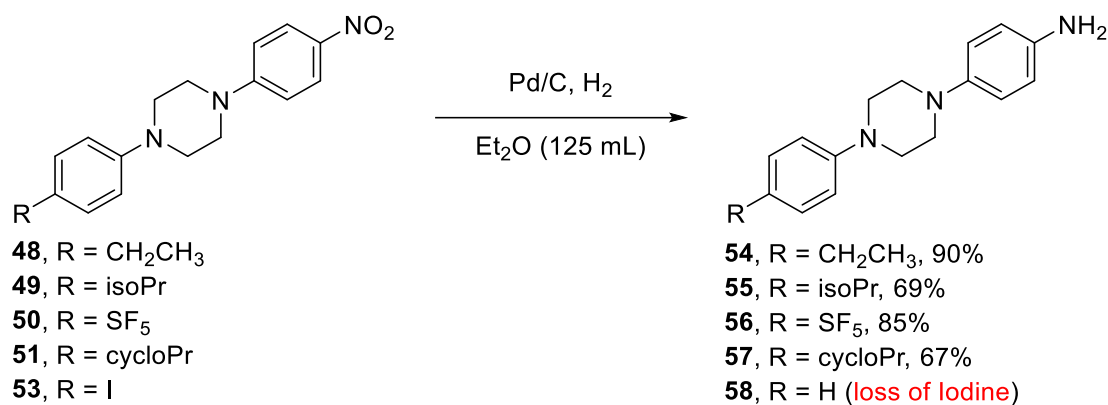


3.3.2.3 Synthesis of 4-(4-Phenylpiperazin-1-yl)Aniline Derivatives

The next step in the preparation of the lipophilic pocket derivatives is the reduction of the previously synthesised 1-(4-nitrophenyl)-4-phenylpiperazine derivatives. Thus, using a standard method *via* Pd/C catalysis in EtOAc, derivatives **48-51** and **53** were reduced to amino derivatives **62-66** in yields ranging from 67 to 90% (Scheme 18). However, the iodo-derivative **54** yielded **58** where the carbon-iodine bond was cleaved by action of the palladium catalyst. This was an unexpected consequence of the nitro reduction, nevertheless, this result can be rationalised by the ability of palladium to insert itself oxidatively into the carbon-iodine bond (such as in the

Buchwald-Hartwig reaction). This meant there was a need to explore palladium-free nitro reduction methods.

*Scheme 18 – Synthesis of 4-(4-phenylpiperazin-1-yl)aniline derivatives **54-58**.*



To that effect, an investigation of different literature methods tolerant of iodine for nitro group reduction of **52** to form aniline **59** was undertaken. The results of the different methods are shown in *Table 20* and the respective crude aromatic signals in the ^1H NMR region for each entry are presented in *Fig. 36* highlighting the ortho-H peak due to the presence of the para- NO_2 group on the phenyl ring at approximately 8.10 ppm and the two triplets which result from para- and meta-Hs of the unsubstituted phenyl ring at approximately 7.75 and 6.85 ppm.

*Table 20 – Reaction conditions for the reduction of the nitro group of **52**.*

Reaction of 52 (4-(4-iodophenyl)piperazine-1-yl)nitrobenzene to 59 (4-(4-iodophenyl)piperazine-1-yl)aniline under various conditions.		
Entry	Conditions	Results
1	LiAlH_4 , THF [0.1]M, RT, O.N.	Reduction and Cleavage
2	$\text{Na}_2\text{S}_2\text{O}_4$, EtOH:H ₂ O (3:1) [0.2]M, RT, O.N.	Reduction Failed and Cleavage
3	$\text{SnCl}_2 \cdot 2\text{H}_2\text{O}$, EtOH [0.2]M, 70°C, Reflux, 2 hr.	Reduction and Cleavage
4	Sulphur, NaHCO_3 , DMF [0.3]M, 130°C, 18 hr	Reduction Failed and Cleavage
5	Fe dust, CaCl_2 , EtOH:H ₂ O (21:1) [0.95], RT, O.N.	Reduction Failed and Cleavage

6	Zn (6 eq.), AcOH:EtOH:H ₂ O (3:2:1) [0.1]M, RT, O.N.	Reduction and Cleavage
7	D-glucose, KOH, H ₂ O:DMSO (1:1) [0.25]M, 100°C, 24 hr.	Reduction and Cleavage

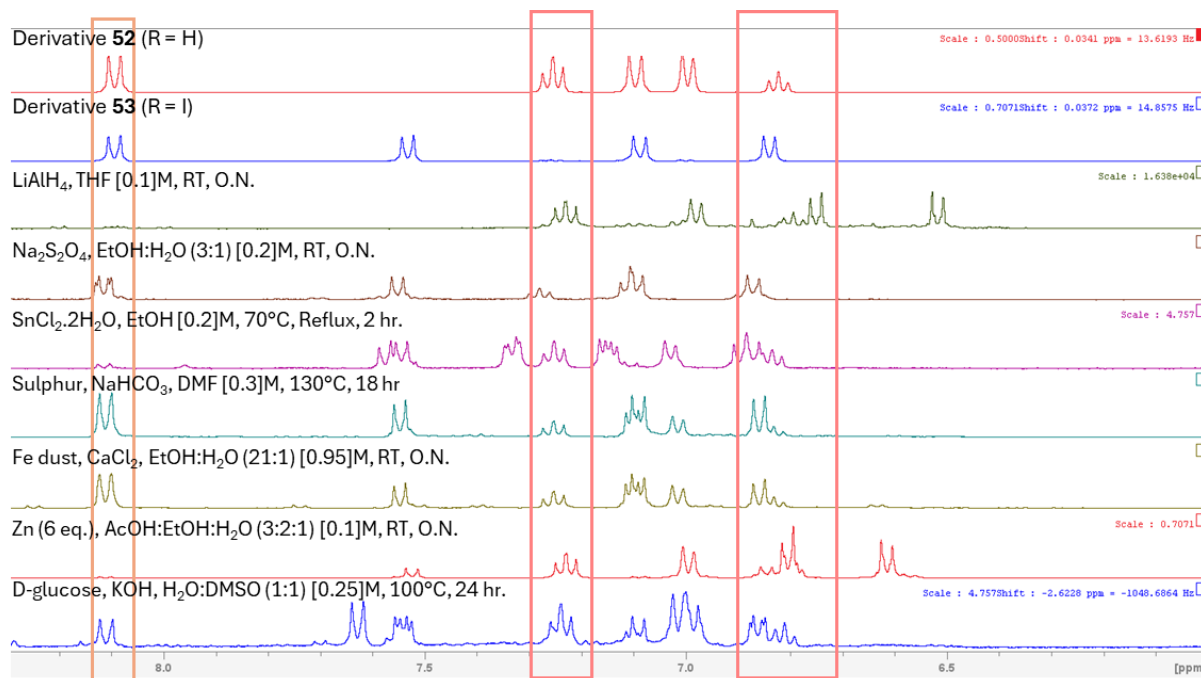


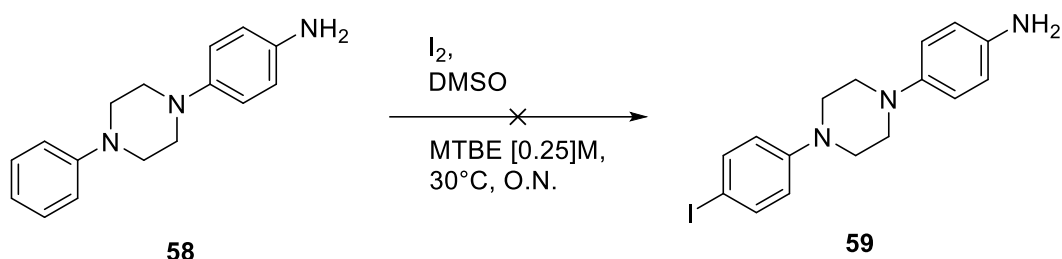
Figure 36 – ¹H NMR Spectra of derivatives **52** and **53** along with the crude spectra for entries in Table 20.

Lithium aluminium hydride was used for aromatic nitro group reduction as far as back as the 1960s (entry 1);¹⁹¹ however, this reagent caused both the reduction of the nitro group and cleavage of the C-I bond. The reduction of nitroarenes under mild conditions using sodium dithionite as previously reported was attempted;¹⁹² unfortunately, reduction was not achieved and this was coupled with carbon-iodine cleavage. Metal-based reduction methods employing Sn, Zn and Fe were attempted (entries 3, 5, and 6).¹⁹³⁻¹⁹⁵ Stannous chloride was successful in reducing the nitro group however also facilitated the cleavage of the C-I bond. Iron dust and calcium chloride at room temperature overnight did not reduce the nitro group but participated in cleavage. The two attempts using Zn metals showed different results, while the Zn in the presence of acid (Entry 4) was successful in reduction, it also cleaved the C-I bond. A method using elemental sulphur and sodium hydrogen bicarbonate was also explored;¹⁹⁶ however, only C-I bond cleavage was observed.

Finally, a method using D-glucose and potassium hydroxide to generate hydrogen gas *in-situ* was reported to be able to reduce nitro groups in the presence of iodine;¹⁹⁷ however, the nitro group remained while partial cleavage occurred.

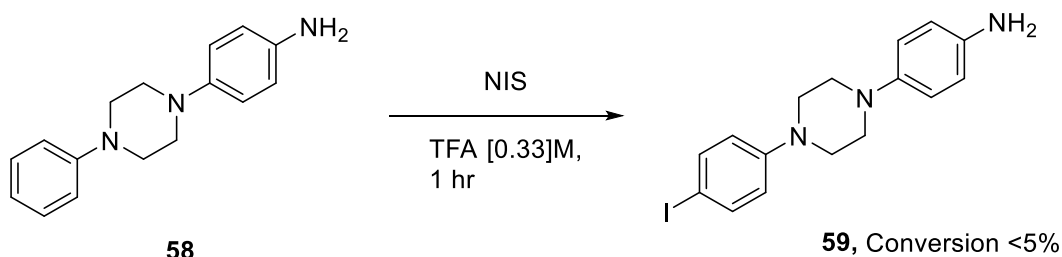
Concurrently, the direct iodination of **58** (unsubstituted 4-(4-phenylpiperazin-1-yl)aniline) was also explored though this pathway was less preferred given the possibility of product mixtures due to incomplete conversions. Using the same conditions as for **52**, *i.e.* iodine, DMSO and MTBE, did not return the aimed compound.

Scheme 19 – Attempted synthesis of substituted 4-(4-phenylpiperazin-1-yl)aniline 59.



In the synthesis of **59** an attempt was made to use *N*-iodosuccinamide and trifluoroacetic acid as for the iodinated diol **43**; however, crude ^1H NMR showed very little conversion and isolation was difficult due to the presence of excess trifluoroacetic acid which protonates both the aniline and piperazine rings.

Scheme 20 – Synthesis of aniline 4-(4-phenylpiperazin-1-yl)aniline 59.

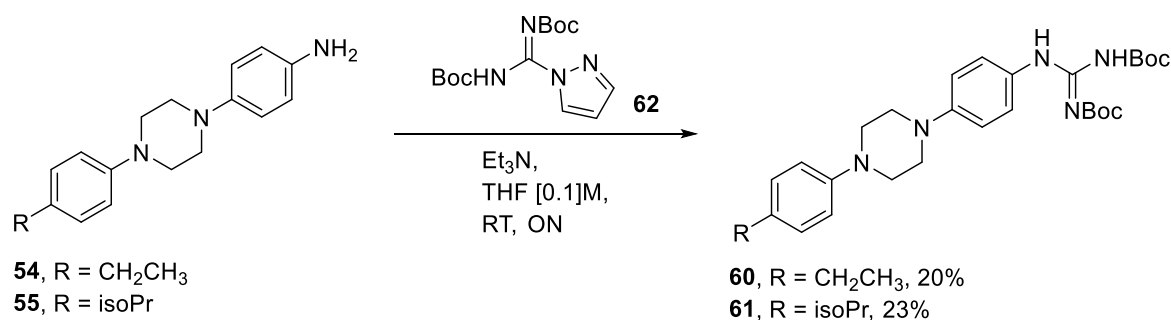


Unfortunately, due to time constraints and difficulty in its synthesis it was decided to abandon the synthesis of **9** (-I).

3.3.2.4 Guanidylation of 4-(4-Phenylpiperazin-1-yl)Aniline Derivatives

Once the synthesis of the 4-(4-phenylpiperazin-1-yl)aniline derivatives was completed, the corresponding Boc-protected mono-guanidine derivatives **60** and **61** were prepared by reacting **54** and **55** with commercially available *N,N'*-di-Boc-1H-pyrazole-1-carboxamide **62**. The rationale behind this reagent is that the pyrazole moiety should be a good enough leaving group to encourage the nucleophilic attack of the aniline on the carboxamide hence removing the need to use *N,N*-bis-Boc-S-methylisothiourea in conjunction with mercuric (II) chloride (standard guanidylation reagents used within the Rozas group).¹⁹⁸ Thus, **53** and **55** were treated with **62** and Et₃N in THF at room temperature overnight yielding **60** and **61** at 20% and 23% respectively (*Scheme 21*). Full consumption of **53** and **54** was not achieved and purification was carried out *via* flash column chromatography with hexane:EtOAc (1:0 to 8:2), though mixed fractions were prevalent given the closeness in R_f of guanidylating reagent **62** (R_f = 0.65, hexane:EtOAc 7:3), **60** (R_f = 70) and **61** (R_f = 70). Low yield was expected given the poor nucleophilicity of anilines and this was worrying for the guanidylation of anilines **56** (-SF₅) and **57** (-cycloPr) given the small amounts available.

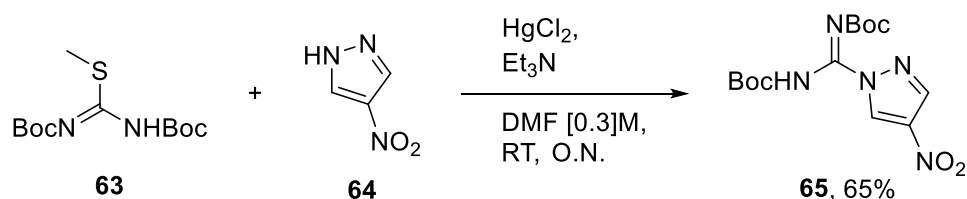
Scheme 21 - Synthesis of mono-guanidine derivatives 60 and 61.



Examining the literature for more reactive guanidylating reagents, it was found that the nitro 3-substitution of the pyrazole ring (**62**) results in more efficient guanidine synthesis. However, this reagent is not commercially available and, thus, had to be

prepared. In a modified method to that of the reported by Yong *et al.*, *N,N*-bis-Boc-S-methylisothiurea **63** was reacted with nitropyrazole **64** in DMF with triethylamine acting as the base and mercury (II) chloride as the desulfurisation agent affording **65** at 65% yield (Scheme 22)

Scheme 22 - Synthesis of guanidylation reagent **65**.



As a test of this reagent's synthetic efficiency compared to *N,N*-bis-Boc-S-methylisothiurea **63**, all three were used as reagents in the guanidylation of aniline **66** affording **67** (Table 21). **65** performs the best at 82% yield while **64** and **62** perform similarly to each other at 73 and 75% yield respectively. Important to note however is these high yields are a function of the more nucleophilic aniline **66**, given that unsubstituted anilines has a reported yield of 11% with reagent **62**.¹⁹⁹

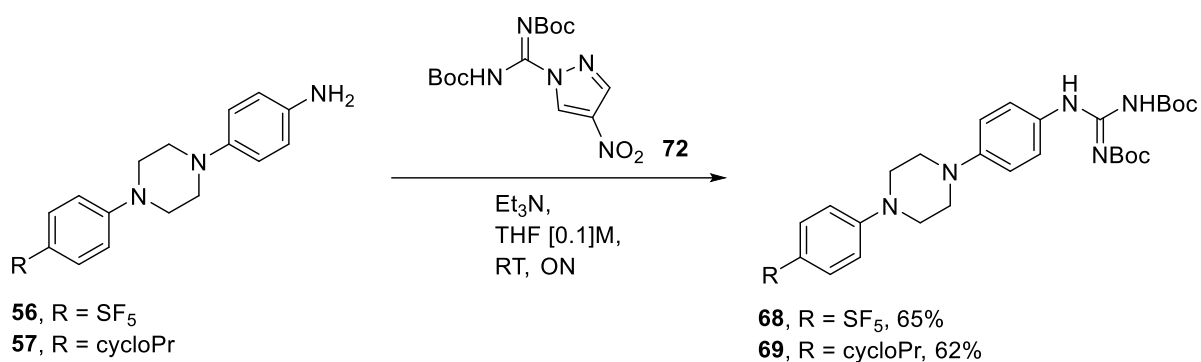
Table 21 - Conditions for the guanidylation of aniline **67** and their respective yields.

Entry	Conditions	Yield
1	<i>N,N</i> -bis-Boc-S-methylisothiurea 63 , HgCl_2 , Et_3N , THF[0.2]M, RT, O.N.	73%
2	Pyrazole 62 , Et_3N , THF[0.2]M, RT, O.N.	75%
3	Pyrazole 65 , Et_3N , THF[0.2]M, RT, O.N.	82%

Therefore, for the guanidylation of aniline derivatives **56** (-SF₅) and **57** (-cycloPr), guanidylation reagent **65** was used using the same standard conditions in THF with Et_3N base (Scheme 23). This afforded boc-protected guanidines **68** and **69** at 65 and 62% yield respectively (Scheme 23). Notably, the preceding aniline had been fully

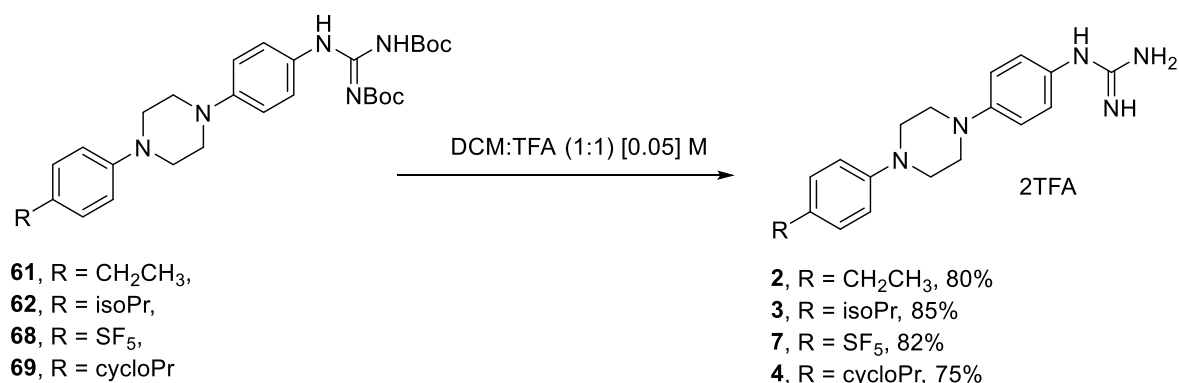
consumed in both cases (in contrast to the anilines **54** [-CH₂CH₃] and **55** [-isoPr] which were not fully consumed) however again the closeness in R_f of the products and guanidylation reagent made mixed fractions appear in both cases, even when using larger columns with silica gel pre-treated with 1%v/v Et₃N in hexane and a smoother gradient change in the mobile phase Hexane:EtOAc (1:0 to 85:15) *i.e.* 1:0 to 95:5 to 90:10 *etc.*

*Scheme 23 – Synthesis of Boc-protected guanidines **76** and **77**.*



Finally, to afford the guanidinium trifluoroacetate salts for biological testing, Boc-protected guanidines **61**, **62**, **68** and **69** were deprotected using standard Boc deprotection methods for guanidines employed in the Rozas group (*Scheme 24*). This afforded the trifluoroacetate salts in yields ranging from 75 to 85%.

*Scheme 24 - Deprotection of Boc-protected guanidines **61**, **62**, **68** and **69**.*



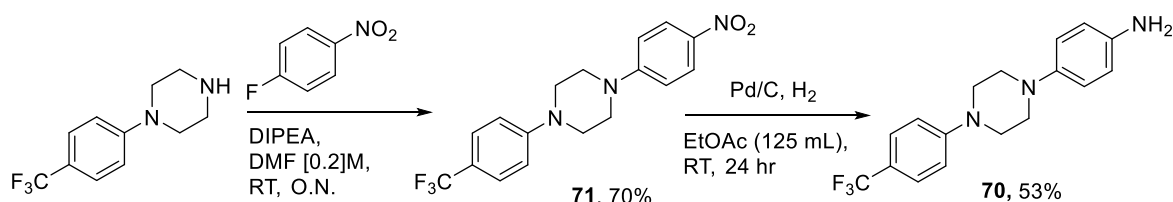
3.3.3 Synthesis of Alkylguanidine Derivatives

Referring to the retrosynthetic analysis in *Scheme 5*, the synthesis of the alkylguanidine derivatives was planned to be carried out *via* the alkylation of guanidylating reagents and using this in the guanidylation of the CF₃-substituted aniline handle. Thus, our synthesis begins with the synthesis of the required aniline.

3.3.3.1 Synthesis of 4-(4-(4-(trifluoromethyl)phenyl)piperazin-1-yl)Aniline **70**

For the synthesis of 4-(4-(4-(trifluoromethyl)phenyl)piperazin-1-yl)aniline **70**, a similar synthetic route was followed as with the other aniline derivatives, as developed by Dr. Rahman in Rozas' group.⁸¹ Beginning with the the S_NAr reaction between commercially available CF₃-substituted 1-phenylpiperazine and 4-fluoronitrobenzene affording **71** at 70% yield. This was then reduced using standard conditions (previously discussed) affording aniline **70** at 53% yield (*Scheme 25*).

Scheme 25 – Synthesis of aniline 70.

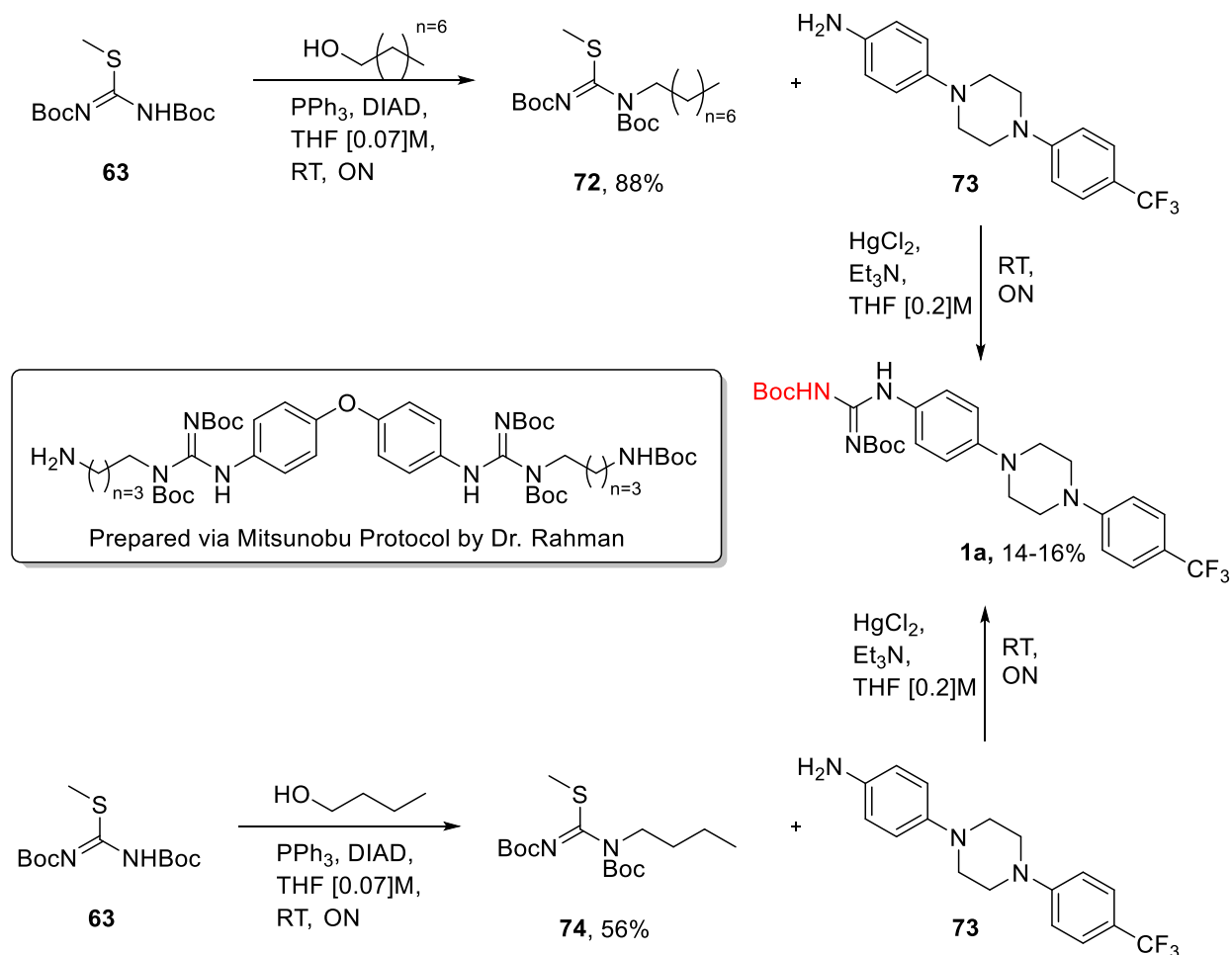


3.3.3.2 Synthesis of 1-Alkyl-3-(4-(4-(4-(trifluoromethyl)phenyl)piperazin-1 yl)phenyl)guanidines

The next step was the preparation of an alkylated guanidylating reagent using a Mitsunobu-like protocol as reported by Kim *et al.* for the synthesis of di-substituted guanidines.²⁰⁰ Thus, compound **63** was treated with n-octanol in the presence of triphenylphosphine and diisopropyl azodicarboxylate (DIAD) in THF overnight affording **72** at 88% yield (*Scheme 26*). This was then reacted with **73** using standard guanidylation conditions; however, post purification compound **1a** was afforded with the octyl chain removed, was obtained unexpectedly at 14% yield. This method had been successfully employed by Dr. Rahman for the generation of amino-alkylguanidine derivative **74** (structure in *Scheme 26*), which was stable on silica gel,

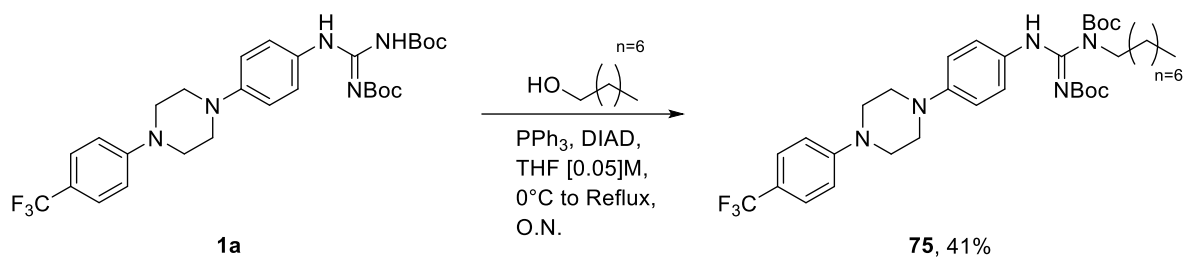
yet was unsuccessful in this case.⁸¹ To investigate whether alkyl chain length is responsible of this result, **73** was guanidylated using the butyl chain guanidylation derivative **74**. Unfortunately, yet again, **1a** was afforded at 16% yield.

*Scheme 26 - Synthesis of guanidylation reagents **72** and **74**, boc-protected guanidine derivative **1a** and structure of biaryl derivative synthesised via a mitsunobu protocol by Dr. Rahman.*



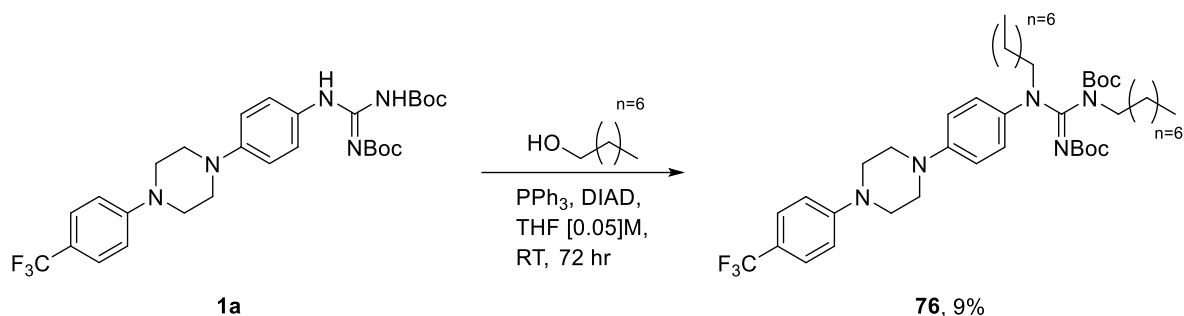
However, it was realised that the Boc-protected guanidine **1a** contains acidic protons itself and therefore could potentially act as a nucleophile in the Mitsunobu reaction. Thus, 2 equivalents of **1a** were treated to the modified protocol employed in the synthesis of **73** and **74** (involving 1 equivalent of n-octanol), affording **75** at a 41% yield (Scheme 27). This method of synthesising disubstituted guanidines was previously reported by Feichtinger *et al.* using the synthesis of alkylated arginine derivatives in the 1990s.²⁰¹

*Scheme 27 - Synthesis of octyl chain Boc-protected guanidine **75**.*



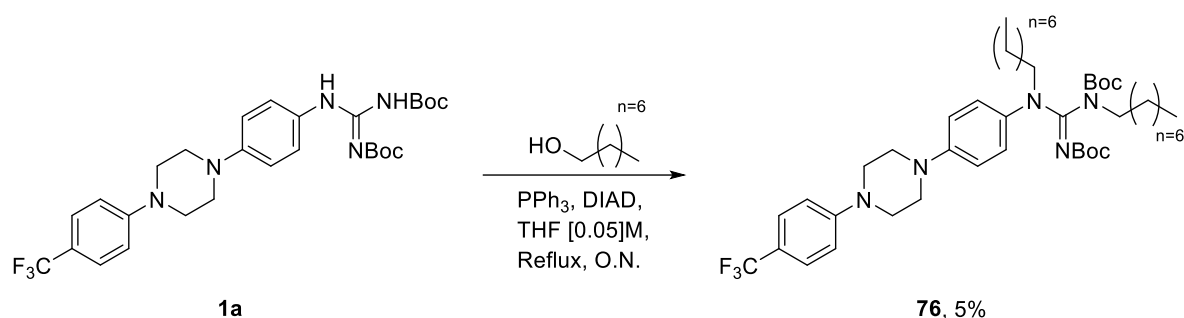
Additionally, in a parallel effort, 1 equivalent of **1a** was treated with 1.5 equivalents of *n*-octanol and 1.5 equivalents each of DIAD and PPh_3 , unexpectedly affording **76** at 9% after 3 days at room temperature (Scheme 28).

*Scheme 28 - Synthesis of dioctyl Boc-protected guanidine **76**.*



Several attempts were made to replicate this reaction on a shorter timescale. The most successful attempt was using 1 equivalents of guanidine **1a**, 2 equivalents of *n*-octanol and 2.2 equivalents each of PPh_3 and DIAD under reflux affording **76** at 5% yield (Scheme 29). Due to time constraints the reaction conditions could not be optimised.

Scheme 29 – Synthesis of dioctyl Boc-protected guanidine **76**.



To understand the unusual substitution pattern, we calculated the pK_{aH} of the N-H bond in **75** utilising the RowanSci (virtual platform) implementation of Grimme and Pracht's pK_{a} calculation where they use linear free energy relationships to calculate different macroscopic pK_{a} states of a molecule using QM-based methods.^{202,203} Thus, it was calculated that the N-H bond in **75** has a high pK_{a} of at 12.89 (Fig. 37).

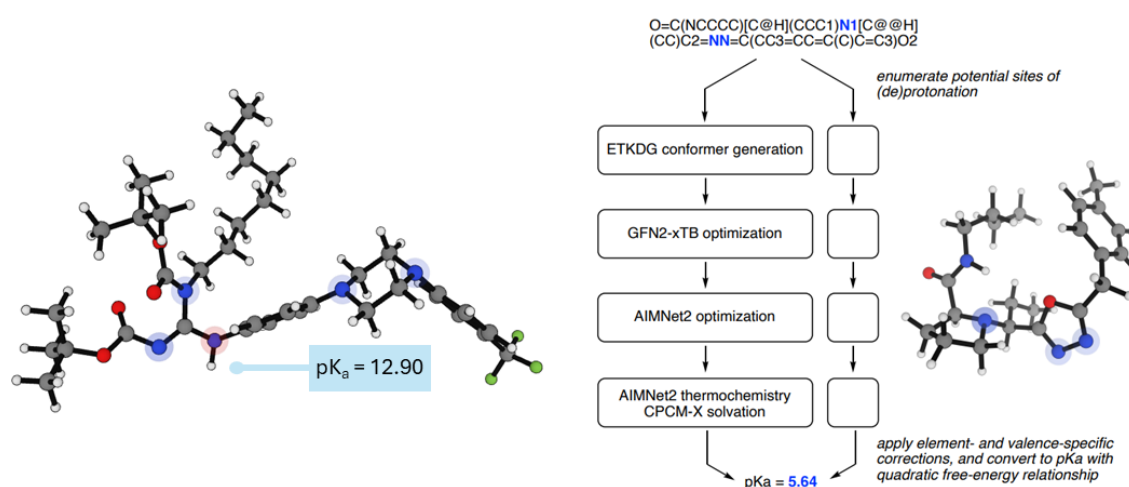


Figure 37 – Left) RowanSci workflow optimised structure of **75** and the pK_{a} of the N-H bond. Right) RowanSci workflow for accurate pK_{a} determinations (taken from reference 203).

In general, for Mitsunobu reactions to be successful, the amino pro-nucleophile must have a pK_{aH} of less than 11.²⁰⁴ In our case, the calculated pK_{aH} , while an approximation, provides an explanation for the low yield. Furthermore, the R_{f} values of **75** and **76** are 0.55 and 0.60, respectively, compared to 0.5 for **1a**; therefore

purification was very difficult. Therefore, the synthesis of **10** (-Butyl) and **11** (-Hexyl) using this method was abandoned.

Deprotection of **75** and **76** using the standard Boc-deprotection method for guanidines as before, afforded **12** and **21** in 75% and 85% yields respectively (*Scheme 30*).

Scheme 30 - Boc-deprotection of derivatives 75 and 76.



3.3.4 Synthesis of Triazole Derivatives

The synthesis of the triazole derivatives begins with the alkylation of the 1-phenylpiperazine derivatives of interest. Some are commercially available as in the case of the para substituted -CF₃, -CH₃ and -H 1-phenylpiperazines and in the case of the -isoPr substituted phenylpiperazine, which was not commercially available, methods for its preparation were established in Section 3.3.1.

3.3.4.1 Synthesis of 1-Phenyl-4-(prop-2-yn-1-yl)piperazine Derivatives

The preparation of the substituted 1-phenyl-4-(prop-2-yn-1-yl)piperazine derivatives involved the S_N2 reaction between 1-phenylpiperazines **45**, **77** and **78** with propargyl bromide affording derivatives **79**, **80** and **81** at 14-71% yields (*Scheme 31*). However, given the low isolated yield obtained for **80** (14%) a different approach had to be taken. Thus, **45** and **82** were reacted with propargyl bromide in DMF at 90 °C with K₂CO₃ (*Scheme 31*). However, post-purification of the new reaction *via* flash column chromatography, **80** was afforded at an even lower 6% yield. Therefore, it was decided

to use crude **80** (and **83**) in subsequent synthesis given little difference in ^1H NMR spectra between the crude mixture and purified forms of the compounds (example of **80** in Fig. 38). The major impurity would have been propargyl bromide in both crude mixtures.

Scheme 31 - Synthesis of 1-Phenyl-4-(prop-2-yn-1-yl)piperazine derivatives.

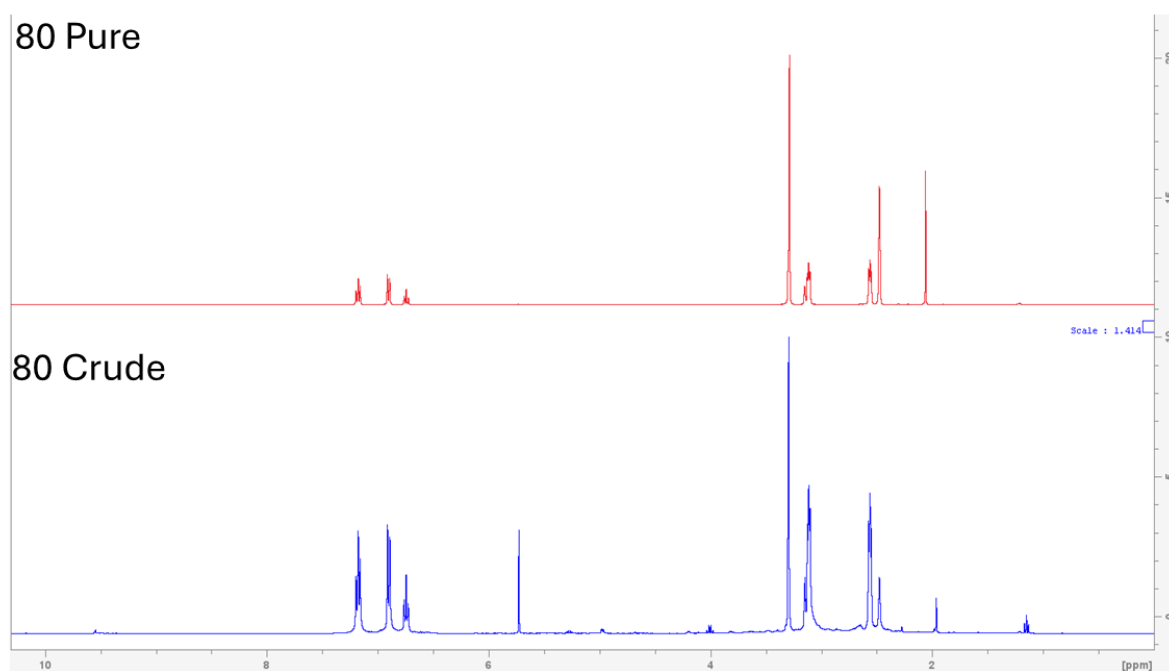
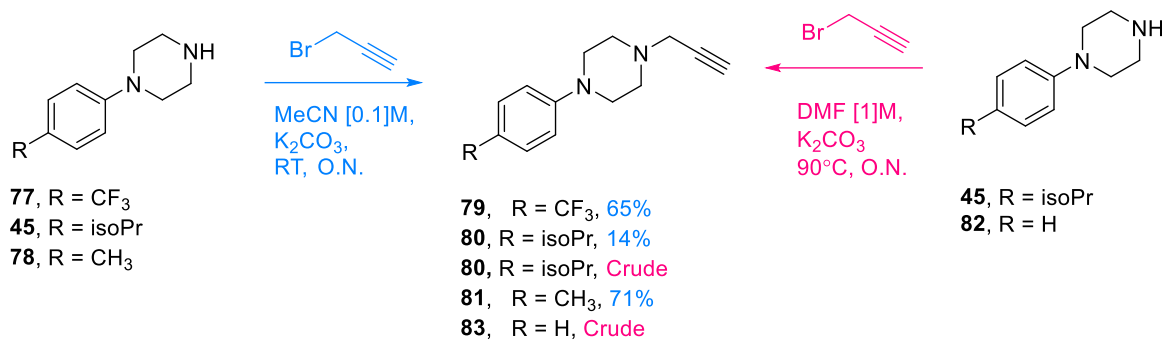


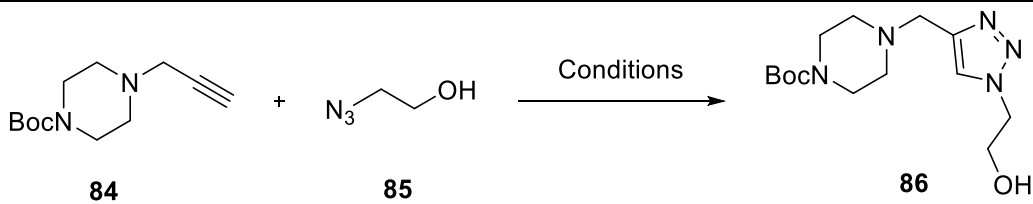
Figure 38 – Pure and crude ^1H NMR spectra of **80**.

3.3.4.2 Synthesis of tert-Butyl (2-(4-((4-Phenylpiperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)carbamate Derivatives

To install the triazole moiety and the amine handle required for guanidylation onto derivatives **79-81** and **83**, conditions for the Cu-catalysed azide-alkyne cycloaddition

Conditions were explored for the CuAAC reaction between model alkyne and azide **84** and **85**, the results of which can be seen in *Table 22*. CuSO₄·5H₂O (0.08M) and Na*Ascorbate were used as the Cu source and reducing agent in entries 1 and 2. The use of methanol in entry 1 resulted in complete alkyne consumption and purification *via* flash column chromatography (DCM:MeOH 1:0 to 9:1) affording **86** in 89% yield, while the use of MeCN resulted in no reaction most likely due to the strong association of MeCN to the Cu(I) species thereby preventing the formation of the Cu(I)-acetylide/azide aggregates previously discussed in *Fig. 39* and hampering catalysis. Cu(OAc)₂·2H₂O as the copper source in *tert*-BuOH was also successful affording **86** at 80% yield; however, it is easier to work with methanol at room temperature than *tert*-BuOH given its propensity to be a solid at room temperature.

Table 22 – Reaction conditions for the CuAAC reaction between 91 and 92.

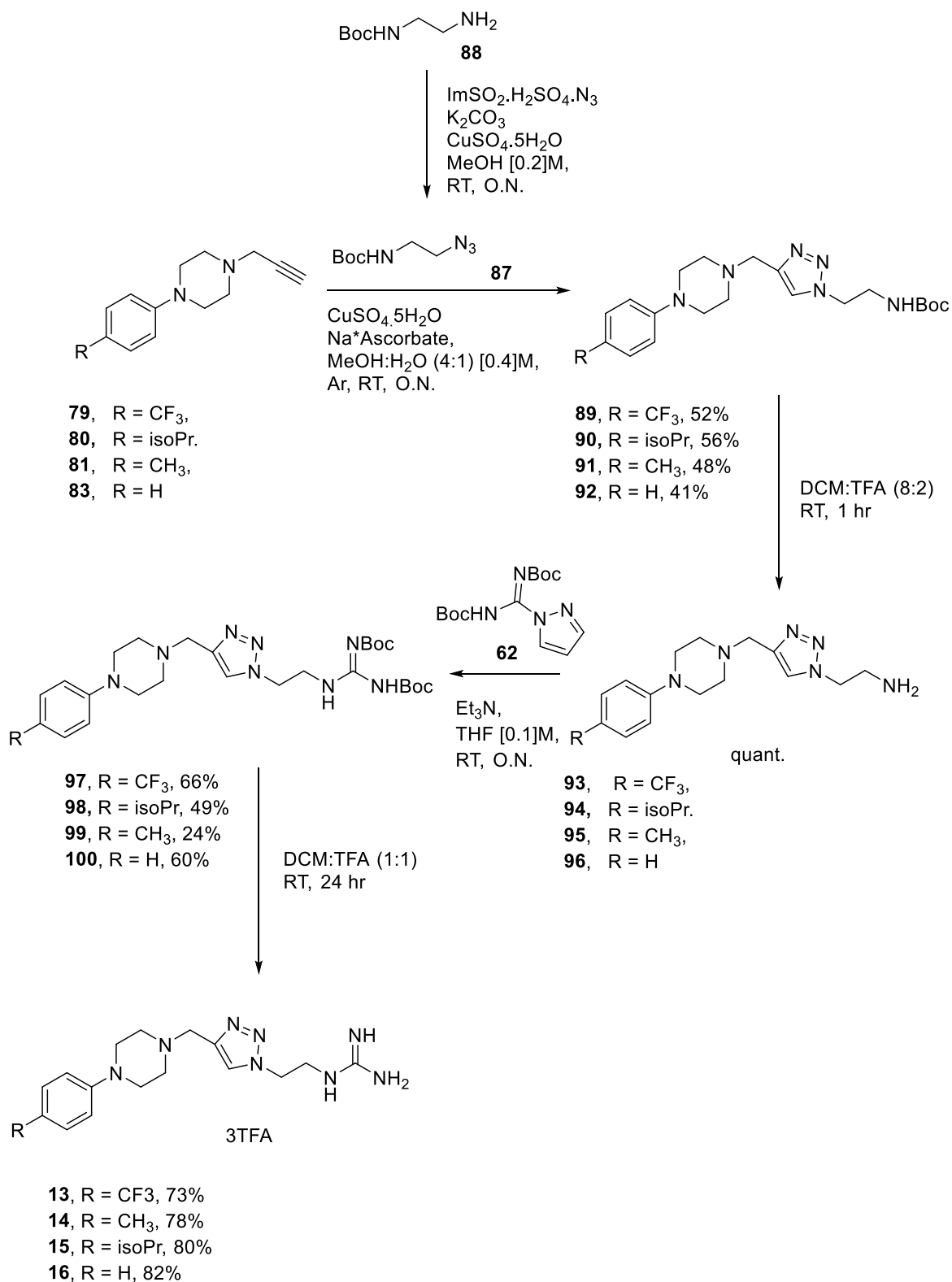
		
Entry	Conditions	Yield
1	CuSO ₄ ·5H ₂ O (0.08M), Na*Ascorbate (0.16M), Methanol:H₂O (4:1) [0.4]M, Ar, RT, O.N.	89%
2	CuSO ₄ ·5H ₂ O (0.08M), Na*Ascorbate (0.16M), MeCN:H₂O (4:1) [0.4]M, RT, Ar, O.N.	No Reaction
3	Cu(OAc)₂·2H₂O (0.08M), Na*Ascorbate (0.16M), tert-BuOH:H₂O (4:1) [0.4]M, Ar, RT, O.N.	80%

3.3.4.3 Synthesis of 1-(2-(4-((4-Phenylpiperazin-1-yl)methyl)-1*H*-1,2,3-triazol-1-yl)ethyl)guanidine Derivatives

The synthesis of the triazole derivatives of interest (*Scheme 32*) begins with the CuAAC reaction between alkyne derivatives **79-81** and **83** with azide **87** (synthesised *via* standard diazotransfer conditions from **88**). This afforded Boc-protected triazole-amino derivatives **89-92** in 41-56% yields which were subsequently Boc-deprotected

using the standard method utilised for previous derivatives, affording **93-96** in quantitative yields. These amine derivatives were then guanidylated using reagent **62**, which caused conversion issues with the aniline derivatives; however, in this case primary amines are sufficiently nucleophilic. Thus, Boc-protected guanidines **97-100** were afforded in 23-66% yields which were subsequently Boc-deprotected affording guanidines **13-16** in 73-82% yields.

Scheme 32 - Synthesis of 1-(2-(4-((4-phenylpiperazin-1-yl)methyl)-1H-1,2,3-Triazol-1-yl)ethyl)guanidine Derivatives



3.4 Biological Testing

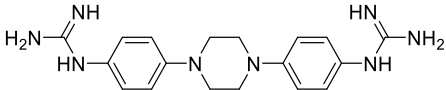
To submit the aimed compounds for biological testing, their purity was first assessed by HPLC with a requirement of $\geq 95\%$ (chromatograms can be seen in appendix). Unfortunately, lipophilic pocket derivative **10** ($-\text{CH}_2\text{CH}_3$) and triazole derivative **22** ($-\text{CH}_3$) did not meet this criteria and therefore were excluded from biological assessment.

3.4.1 Anti-Tubercular Biological Evaluation

All compounds selected were submitted to the laboratory of Professor Sylvie Garneau-Tsodikova (University of Kentucky, USA) to be evaluated against the *M. tb.* strain mc² 6230 and the resultant minimum inhibitory concentration (MIC) can be found in Table 23. Additionally, compound **1** and the original bis-guanidine derivative (which targets DNA as well as InhA) were tested, along with isoniazid (INH) and kanamycin (KAN) as controls. The MIC reported is the average found in a triplicate MIC determination at different time periods (this indicates the variability found in MIC determinations while a single value indicates the same MIC all three times).

Table 23 - Compounds were tested for their activity in the bacterial cell against *Mycobacterium tuberculosis* strain mc² 6230 Kanamycin (KAN) and Isoniazid (INH) were used as controls.

Compound	Substituent	MIC ($\mu\text{g mL}^{-1}$)
1	$-\text{CF}_3$	32
Lipophilic Pocket Derivatives		
3	-isoPr	16-32
4	-cycloPr	32
7	$-\text{SF}_5$	4-8
Alkylguanidinium Derivatives		
12	-Octyl	>64
21	-diOctyl	8
Triazole Derivatives		
13	$-\text{CF}_3$	>64
15	-isoPr	>64
16	-H	>64

Bis-Guanidine		2-8
INH		0.5
KAN		0.5

Within the lipophilic pocket derivatives, compound **7** bearing the -SF₅ substituent has emerged as the most promising candidate out of the compounds tested with MIC values ranging from 4-8 µg mL⁻¹. While this represents an 8-16 fold decreased anti-tubercular activity compared to INH and KAN, it demonstrates a significant improvement from the parent compound **1** and over other members of this class *i.e.* **3** (-isoPr) and **4** (-cycloPr), which have weaker anti-tubercular activity ranging from 16-32 µg mL⁻¹. The superior anti-tubercular activity of **7** may be attributed to better targeting of the lipophilic pocket (as described in the molecular modelling in Section 3.2.3). It may also be attributed to better cellular penetration. The similar activity of **1** (-CF₃) to **3** (-isoPr) and **4** (-cycloPr) also provides evidence for their potential as bioisosteres or bioequivalents of -CF₃.

The alkylguanidinium series shows the most dramatic change in structure-activity relationship. Hence, **12** (-Octyl), the mono-alkylated guanidinium derivative exhibits a MIC > 64 µg mL⁻¹, which is too high to translate into positive treatment outcomes *in vivo*. However, the di-alkylated derivative **21** (-diOctyl) exhibits a MIC of 8 µg mL⁻¹ indicating promising anti-tubercular activity. Given the fact that activity is expected to be linear with increased alkylation, and also according to our molecular modelling results, derivatives **12** and **21** have been re-synthesised and sent for testing once again to confirm this counterintuitive outcome. Recent results have confirmed this unexpected behaviour therefore indicating that for this series of compounds, activity does not scale with increases in alkylation and that the target for **21** may differ from InhA.

Finally, all tested triazole derivatives (**13**, **15** and **16**) demonstrated MIC values >64 µg mL⁻¹, indicating they are inactive against TB. This indicates that the introduction of the triazole moiety and increased flexibility near the guanidinium, negatively affects their

binding to InhA regardless of substituent on the phenyl ring and therefore this strategy is not suitable for the anti-tubercular guanidino piperazine class of compounds.

The bis-guanidine derivative has been included as well to evaluate the approach of the amphiphilic design philosophy. While it shows a superior MIC range (2-8 $\mu\text{g mL}^{-1}$) than the two most promising derivatives **7** (-SF₅) and **21** (-diOctyl), it is not suitable for further development as anti-tubercular agent because it is also a DNA minor groove binder.²⁰⁹ The amphiphilic derivatives, including **7** and **21**, were designed to avoid potential DNA binding and to specifically target InhA thus preventing any cytotoxicity and increasing safety profile.

3.4.2 Experimental Determination of ADMET Properties

Preliminary experimental ADMET properties for the compounds tested (*i.e.* half-time of elimination - $t_{1/2}$ - and clearance - Cl_{int} -) were assessed in the laboratory of Professor Anna Hirsch (Helmholtz Institute for Pharmaceutical Research Saarland, Germany) and the results obtained can be found in *Table 24*. In this early phase of drug discovery, industry-standard assays to assess mouse liver microsome metabolic stability, HepG2 hepatocyte viability and lipophilicity ($\text{LogD}_{7.4}$) were carried out.

Table 24 - Preliminary experimental ADMET properties for tested compounds.

	Substituent	$\text{LogD}_{7.4}$	Mouse Liver Microsomes $t_{1/2}$ [min] / Cl_{int} [$\mu\text{L}/\text{mL}/\text{min}$]	HepG2 MTT Viability @ 100 μM [%]	HepG2 CC_{50} μM
1	-CF ₃	2.83	>120 / <11.6	97.0 ± 1.1	-
Lipophilic Pocket Derivatives					
3	-isoPr	-	>120 / <11.6	16.6 ± 2.6	31.3 ± 9.0 μM
4	-cycloPr	3.86	>120 / <11.6	17.0 ± 3.9	26.9 ± 4.5 μM
7	-SF ₅	3.36	>120 / <11.6	17.3 ± 1.3	15.8 ± 3.1 μM
Alkylguanidine Derivatives					
12	-Octyl	-	>120 / <11.6	94.5 ± 7.3	-
21	-diOctyl	-	66.3 ± 4.2	18.7 ± 3.1	8.6 ± 0.9 μM
Triazole Derivatives					
13	-CF ₃	-	>120 / <11.6	94.1 ± 6.8	-
15	-isoPr	-	>120 / <11.6	94.5 ± 7.3	
16	-H	-	>120 / <11.6	99 ± 14	

The LogD_{7.4} evaluation is on-going and has so far only been carried out for derivatives **1**, **4** and **7**. As previously mentioned in Section 3.2.5, the ideal logP range is between 1 and 3 therefore derivatives **4** (-cycloPr) and **7** (-SF₅) with 3.86 and 3.36 respectively are slightly out of this range being moderately lipophilic. However, parent compound **1** has a logD_{7.4} within this range at 2.83, which would bode well for its pharmacokinetics.

In terms of pharmacokinetic parameters, the mouse liver microsome half-life ($t_{1/2}$) values were determined. This parameter is widely used as an initial cell-based model to correlate with human liver microsome stability.²¹⁰ All compounds but **21** (-diOctyl) exhibit excellent metabolic stability with mouse liver microsomes $t_{1/2}$ values >120 minutes, indicating a low intrinsic clearance (Cl_{int}) of <11.6 $\mu\text{l ml}^{-1} \text{ min}^{-1}$ (meaning hepatic blood flow increase has minimal effect on the clearance of the drug). However, compound **21**(-diOctyl) shows reduced metabolic stability compared to parent compounds **12** and **1**, which could indicate rapid clearance that would require higher or more frequent dosing. For the rest of the compounds, the results obtained are promising in terms of lacking drug-drug interactions caused by CYP450 binding (inhibition or induction).²¹¹

The HepG2 MTT *viability* assay at 100 μM serves as a preliminary cytotoxicity screen used widely in drug discovery.²¹² While **1**, **12** (-Octyl) and all triazole derivatives exhibit excellent high HepG2 *viability*, lipophilic pocket derivatives **3** (-isoPr), **4** (-cycloPr) and **7** (-SF₅) and dialkylatedguanidinium **21** exhibit severe hepatotoxicity (16.6 – 18.7% *viability*) which would generally preclude them from further development. In *Fig. 40*, the experimentally determined HepG2 *viability* assays for first line InhA inhibitor drugs and their combinations can be seen.²¹³ Parent compound **1**, although it is not as effective as an anti-tubercular agent as **7** and **21**, it does show a better HepG2 *viability* and lower hepatotoxicity at 100 μM than isoniazid, pyrazinamide and rifampicin in combination therapy at low doses and higher still at 200 mM and 100 μM of the 1st line drugs.

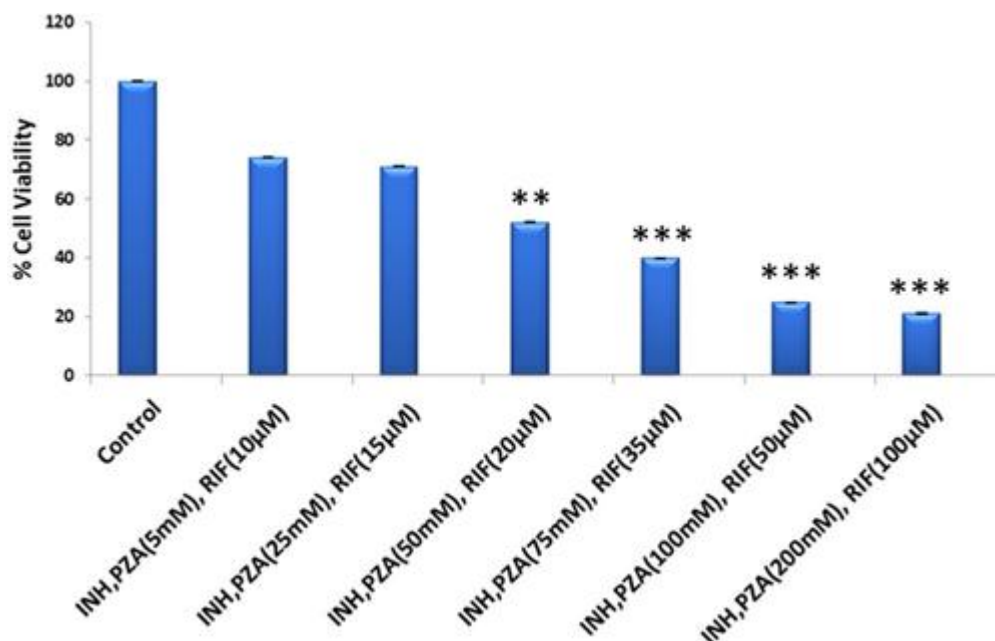


Figure 40 - Effect of Isoniazid, Rifampicin, and Pyrazinamide on HepG2 cell viability. Values are mean $\pm$ SD of three independent experiments. ** $p < 0.01$, *** $p < 0.001$ as compared to control.

To further investigate the cytotoxicity of **3**, **4**, **7** and **21** their 50% cytotoxic concentration (CC_{50}) was determined in the Hep2G cell line. From here their selectivity index (SI) was determined which can be defined as the ratio of a toxic concentration of a sample and its effective bioactive concentration (3):

$$SI = \frac{CC_{50}}{MIC} \quad (3)$$

Thus, using this formula, the SI for **3**, **4**, **7** and **21** are respectively 2.6, 1.1, 1.3 and 0.8. For further development, the SI should be at least higher than 10 and then these low SI values indicate that the compounds with good therapeutic potential may also be hepatotoxic.²¹⁴ It can also be noted that while the lipophilic pocket derivatives are metabolically stable, their hepatotoxicity could be related to their slightly high lipophilicity and therefore tissue accumulation-related toxicity.

From all the results obtained, it has been shown that **1** ($-CF_3$) seems to be the most viable compound for further development. While its MIC is modest ($32 \mu g mL^{-1}$), its excellent safety profile and metabolic stability provides a strong structure for optimisation. The substitution of $-CF_3$ for $-SF_5$ from **1** to **7** has improved activity, however, has opened the door to cytotoxicity, potentially due to increased

lipophilicity, while substituents such as -cycloPr and -isoPr have resulted in cytotoxicity without a coupled increase in activity. Similarly, the di-octyl substitution pattern in **21**, while increasing activity, negatively impacts the safety profile of the compound. Whereas mono-alkylation in of **1** to form **12** does not impact its metabolic stability and hepatotoxicity, it does however result in very poor inhibition of *M. tb.* growth (pending currently ongoing repeated tests). While it is disappointing that the triazole derivatives have no significant anti-tubercular activity, their safety profile suggests they could be good drug candidates for different targets.

3.4.3 Commentary on Theoretical vs. Experimental Determination of ADMET Properties

Overall, the in silico ADME predictions were consistent with the experimental data for identifying “good” versus “problematic” compounds. However, it is important to note that the theoretical ADME prediction primarily assessed fitness to guidelines-based physiochemical windows for good oral bioavailability (Lipinski, Ghose and Veber) while theoretical PK dealt with prediction of CYP interaction profiles and potential to be a substrate of P-gp. As such, reliably capturing subtle differences in metabolic stability i.e. in mouse liver microsome and toxicity in the form of hepatotoxicity in HepG2 cells is difficult.

Parent compound **1** behaves very much as the theoretical ADME predictions suggested i.e. it has an acceptable lipophilicity and adheres fully to guidelines for good oral-bioavailability, and this is borne out in its metabolic stability and good HepG2 cell viability. In contrast, di-octyl derivative **21** was flagged as being highly lipophilic (cLogP value of 7) with multiple Ro5 violations thereby being a poor drug-like molecule and this was confirmed experimentally by reduced microsomal stability and marked hepatotoxicity.

The absolute ranking of lipophilicity from the cLogP predictions did not perfectly match the experimental ranking for compounds. However, retrospectively it may allow us to rationalise the severe HepG2 toxicity of **3** (-isoPr), **4** (-cycloPr), **7** (-SF5) and **21** (-diOctyl), which are among the most lipophilic members compared to the rest of the series and, therefore, it may be prone to accumulation driven toxicity as they sit at the upper edge of “acceptable” lipophilicity (experimentally and theoretically). In

contrast, more moderately polar triazole derivatives with lower cLogP exhibit excellent HepG2 cell viability and therefore consistent with a more benign safety profile. However, compound 12 (-Octyl) appears to have a good safety profile regardless of cLogP or Ro5 violations.

Hence, one can conclude that the theoretical ADME tools provided a qualitatively useful framework: they successfully identified 1 as a pharmacokinetically robust scaffold and 21 as a high risk, high lipophilicity outlier. However, their ability to predict hepatotoxicity and to accurately correlate metabolic risk within a mid lipophilicity window was limited, underscoring the need to treat in silico ADME as a coarse filter that must be complemented by early microsomal and cell based experimental assays in drug development.

Chapter 4 – Conclusions and Future Work

4.1 Conclusions

Among the different objectives of this work, the first aim was the identification of bioisosteres for the trifluoromethyl ($-\text{CF}_3$) functional group given the environmental concerns regarding its use. Upon the identification of several potential bioequivalents ($-\text{I}$, $-\text{SF}_5$, $-\text{cycloPr}$, $-\text{isoPr}$, or $-\text{cycloBu}$), computer-aided drug design studies were undertaken, designing novel derivatives of the lead compound **1**, assessing their binding potential to InhA and analysing the binding modes of this guanidino piperazine class of compounds to rationalise their biological activity. Based on these studies, the most promising derivatives were prepared following synthetic methodologies known in the Rozas' group and novel ones as needed. Following the successful synthesis of several derivatives, their MIC as anti-tubercular agents and their experimental ADMET properties were assessed by collaborators.

4.1.1 Computational Studies

Following the identification of potential bioisosteres of $-\text{CF}_3$ in terms of size, electronics and lipophilicity (*i.e.* $-\text{I}$, $-\text{SF}_5$, $-\text{cycloPr}$, $-\text{SCH}_3$) new derivatives of **1** were proposed. Docking was then used to probe the ligand-protein binding site. Beginning with conformational sampling of **1**, and utilising the software Attracting Cavities 2.0 in order to eliminate template-based bias in the determination of binding modes, a different binding mode was identified for these guanidino piperazine derivatives. Following this a flexible-ligand/flexible-side-chain docking approach was developed incorporating key side chains: Phe149, Tyr158 and Phe97. This allowed for the identification of a possible binding mode which is biologically relevant, incorporating a bridged HB/ π -cation interactions between the phosphate group of the NADH cofactor, the guanidinium moiety and Phe97. Dual binding modes were found for the alkyl guanidine and triazole derivatives. Compounds **5** ($-\text{cycloBu}$) and **7** ($-\text{SF}_5$) were identified as having a superior theoretical binding to InhA than **1**. A theoretical ADMET assessment was also undertaken, showing a good 'Lipinski-Ghose-Veber' profile indicating good drug-likeness and oral bioavailability. Their theoretical

pharmacokinetics were evaluated using SwissADME indicating some had potential for blood-brain barrier penetration and good gastrointestinal absorption.

4.1.2 Chemical Synthesis

The synthetic protocols previously established by our group were used in the synthesis of the lipophilic pocket derivatives; however, the lack of economically *viable* starting materials was an issue which had to be overcome. Thus, in the case of the -SF₅ and -cycloPr derivatives, methods for their introduction onto the parent scaffold were developed. Buchwald-Hartwig conditions were optimised for the introduction of the -CH₂CH₃, -isoPr and -SF₅ onto the scaffold achieving good yields (67-75%). Due to time constraints and synthetic inaccessibility, preparation of several derivatives had to be abandoned. Subsequently, by following the established methods for the synthesis of the guanidino piperazine within the Rozas' group, the TFA salts were afforded in very good yields (75-82%). The use of a more active guanidylating reagent was beneficial since it allowed for higher yields of the -SF₅ and -cycloPr derivatives, considering that little amounts of their aniline precursors could be produced. For the alkylguanidine derivatives, the -Octyl and -diOctyl derivatives were successfully synthesised implementing Mitsunobu-like conditions, affording the TFA salts at 75 and 85% yields, respectively. Finally, the incorporation of a 1,2,3-triazole ring was successful by employing the CuAAC reaction. The corresponding TFA salts were afforded in very good yields (73-82%).

4.1.3 Biological Testing

Biological testing carried out by our collaborators (Professor Sylvie Garneau-Tsodikova [University of Kentucky, USA] and Professor Anna Hirsch [Helmholtz Institute for Pharmaceutical Research Saarland, Germany]) has yielded interesting results and consequences for our guanidino piperazine compounds in terms of anti-tubercular activity and experimental ADME properties.

In terms of the anti-tubercular results, the novel derivatives which show the best activity are **7** (-SF₅) and **21** (-diOctyl) at MIC values of 4-8 and 8 µg mL⁻¹. However, their potential hepatotoxicity and low selectivity index precludes them from further development. Notably, the similar activity of **3** (-isoPr) and **4** (-cycloPr) to lead compound **1** shows their ability of these functionalities to act as bioequivalents of -CF₃, assuming the binding mode is as computationally predicted. Unfortunately, the triazole derivatives show no significant anti-tubercular activity and therefore their development as therapies for tuberculosis cannot go further; however, their excellent safety profile and drug-likeness means that they could potentially be repurposed for different targets. Looking at the biological results as a whole, it can be deduced that compound **1** is the most promising derivative for further development given its moderate anti-tubercular activity (32 µg mL⁻¹) and excellent mouse liver microsome metabolic stability and safety profile.

4.2 Future Work

There are three fundamental areas of future development given the conclusions of this project are presented next.

4.2.1 Expanded Computational Studies

In order to confirm that the binding modes and interactions characterised in Section 3.2 are correct, higher order computational studies should be undertaken. Firstly, to confirm the binding mode and its stability, molecular dynamics simulations should be performed to further probe the binding interaction of these guanidino piperazine derivatives within the InhA active site. Upon confirmation of the novel binding mode or discovery of a new one, subsequent docking experiments in the design of novel guanidino piperazine derivatives as anti-tubercular agents should employ induced fit docking software (*i.e.* as implemented in Gold or Glide)²¹⁵ in order to allow for the full range of flexibility within the active site given InhA's active site is largely dominated by loops which are highly flexible.

4.2.2 Design and Chemical Synthesis

The substitution of the para- CF_3 functional group on the lipophilic portion of the molecule for $-\text{SF}_5$ resulted in increased anti-tubercular activity and similar activity in the case of -isoPr and -cycloPr; however, in all cases it resulted in a worse safety profile with those changes causing a potential increase in hepatotoxicity. Thus, it shows that from a pharmacokinetic perspective, the $-\text{CF}_3$ moiety is the best para-substituent in terms of safety. Furthermore, dialkylation of the guanidine moiety, while it increased anti-tubercular activity, resulted in potential hepatotoxicity. This, therefore, leaves only the piperazine and phenyl rings as suitable for modification in the design of more effective anti-tubercular activity. Therefore, the proposed structural changes can be seen below in *Fig. 41*. Replacement of guanidiniums for isouronium species has been shown to be a *viable* strategy for biological activity.²¹⁶ Substituting a phenyl meta-C near the guanidine with a N would result in the formation of intramolecular HBs and has been well described.²¹⁷ This could potentially drive the formation of the HB/ π -cation interaction system by keeping the guanidine in its pre-organised flat position. Furthermore, the isosteric replacement of piperazine rings in the cores of molecules with different N-cycles has had varying success in many biologically active compounds.²¹⁸ Therefore, it would be interesting to see the effects that this has on our compounds. Finally, it would be interesting to see how the substitution pattern on the lipophilic portion of the molecule affects overall activity.

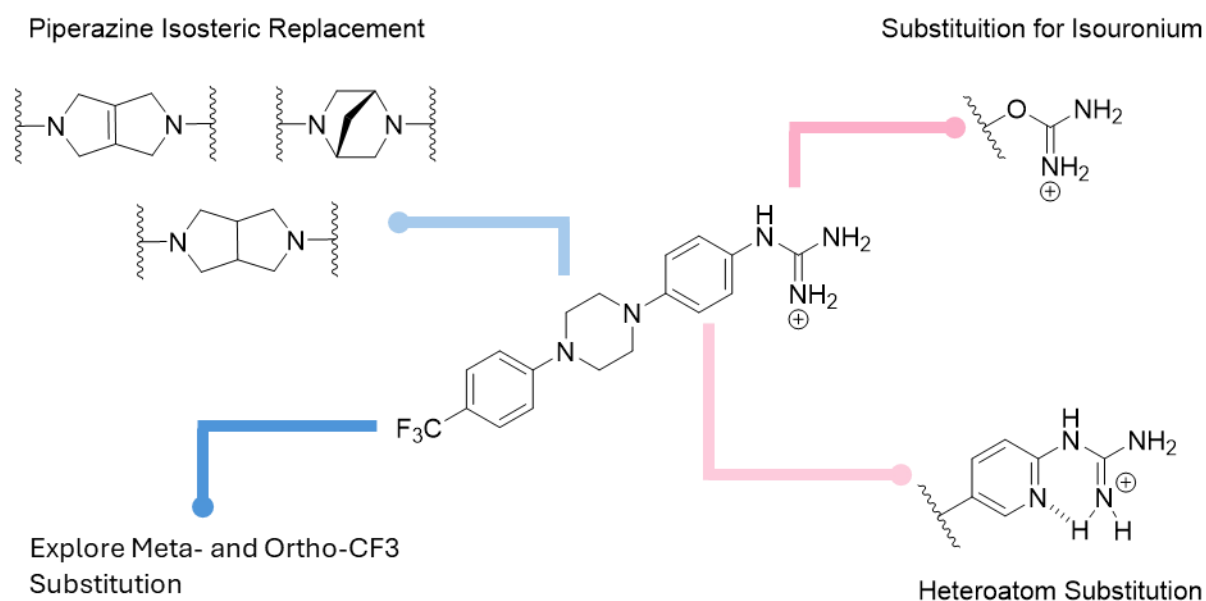


Figure 41 - Proposed structural modifications for novel guanidino piperazine derivatives to act as anti-tubercular agents.

4.2.3 Biological Development

Further biological testing to evaluate **1** should be undertaken, such as testing in drug resistant strains of *M. tb*. Additionally, as it is an intracellular microbe, drug susceptibility assays against *M. tb* infected macrophages should be carried out in order to determine the ability of this class of compounds to cross the macrophage membrane and therefore deal with both active and latent TB.

Chapter 5 – Experimental Methods and Characterisation

5.1 Computational Methods

5.1.1 DFT Methods

Density functional theory (DFT) was used to optimise the structures of all compounds at the M06-2X/6-311+g(d,p) computational level using the Gaussian16 computational package.^{219, 220} Harmonic vibrational frequencies were computed at the same level to confirm the stationary points are local minima. Solvent effects (*i.e.* water) were included in the optimisation by means of the Solvation Model based on Density (SMD) approach.²²¹

5.1.2 Docking Methodology

The 1.97 Å electron microscopy structure of Inha was retrieved from the Protein Data Bank (PDB ID: 5OIR). The molecular docking experiments were performed with AutoDock Vina 1.2.0.¹³² Protein preparation (solvent removal) and ligand preparation were carried out using AutoDockTools 1.5.7 using Maestro (Schrödinger Release 2024-2: Maestro, Schrödinger, LLC, New York, NY, 2024) as visualisation software. DFT optimised structures of ligands were used as inputs for semi-flexible docking, *i.e.* flexible ligand and rigid protein, using the AutoDock Vina's built-in scoring function. The appropriate grid box was generated based on the positioning of co-crystallised ligand within the original electron microscopy structure. The flexible docking included Phe97, Phe149 and Tyr158 as flexible side chains. Exhaustiveness was set at 50 for all docking experiments.

5.2 Synthetic Chemistry

5.2.1 Materials and Methods

All commercial chemicals were obtained from either Sigma-Aldrich, Fischer Scientific or Fluorochem and used without further purification. Deuterated solvents for NMR analysis were purchased from Fischer Scientific. Column chromatography was performed using Sigma-Aldrich silica gel 100-200 mesh. Solvents for synthesis purposes were used at HPLC grade. Analytical TLC was performed using Merck Kieselgel 60 F-254 silica gel plates. Visualisation was completed using a UV light (254 nm). NMR spectra were recorded on a Bruker DPX-400 Advance spectrometer at 400.13 and 100.6 MHz for ^1H and ^{13}C respectively. Additional spectra recorded on a Bruker AV-600 spectrometer were carried out at 600.1 MHz and 150.9 MHz for ^1H and ^{13}C NMR respectively. Shifts are reported as referenced to the internal solvent signal. NMR data was processed using TopSpin, Bruker's NMR data analysis software. HRMS spectra were measured on a Micromass LCT electrospray TOF instrument with a WATERS 2690 autosampler and methanol/acetonitrile as carrier solvent.

Method A – Buchwald-Hartwig Cross-Coupling

To a solution of iodobenzene (1 eq.), 1-Boc-piperazine (1 eq.) and Cs_2CO_3 (2 eq.) in anhydrous PhMe [0.4] M (under an Ar atmosphere) was added $\text{Pd}_2(\text{dba})_3$ (0.025 eq.) and XPhos (0.05 eq.). The reaction was allowed to proceed at 100°C for 24 hours. Upon completion the mixture was filtered *via* syringe using a PTFE syringe filter (0.45 μm pore size). It was diluted with EtOAc, washed with water and brine. The organic phase was dried over Na_2SO_4 , concentrated *in vacuo* and purified by silica column chromatography with a hexane:EtOAc solvent system.

Method B – Suzuki-Miyaura Cross-Coupling

To a solution of bromobenzene (1 eq.), boronic acid (2 eq.) and K_2CO_3 (2 eq.) in anhydrous PhMe [0.2] M (under an Ar atmosphere) was added $\text{Pd}(\text{OAc})_2$ (0.1 eq.), PCy_3 (0.2 eq.). The reaction was allowed to proceed at 100°C for 24 hours. Upon completion the mixture was filtered *via* syringe using a PTFE syringe filter (0.45 μm pore size). It was diluted with EtOAc, washed with water and brine. The organic phase

was dried over Na_2SO_4 , concentrated *in vacuo* and purified by silica column chromatography with a hexane:PhMe solvent system.

Method C – Boc deprotection of Amine Derivatives

Boc-protected amine (1 eq.) was dissolved in an 8:2 DCM:TFA mixture [0.1] M. The reaction was allowed to proceed for 1 hour then concentrated *in vacuo*.

Method D – Nucleophilic Aromatic Substitution Reactions

To a solution of 1-phenylpiperazine (1 eq.) and 1-fluoro-4-nitrobenzene (1.1 eq.) in anhydrous DMF [0.2] M, Et_3N (1.1 eq.) was added and the reaction was allowed to proceed overnight at room temperature. Upon completion, the mixture was diluted with EtOAc, washed with water and brine. The organic phase was dried over $\text{Na}_2\text{S}_2\text{O}_4$, concentrated *in vacuo*.

Method E – Pd/C Reduction of Nitro Derivatives

To a solution of nitro derivative in EtOAc (125 mL) in a 250 mL roundbottom flask, was added 10% Pd/C under an inert atmosphere. The reaction was then allowed to proceed under a balloon of H_2 for 24 hours. The reaction mixture was filtered through celite and concentrated *in vacuo*.

Method F – Guanidylation of Aniline Derivatives

To a solution of amine (1 eq.) and guanidylation reagent (1.1 eq.) in anhydrous THF [0.1]M was added Et_3N (1.1 eq.). The reaction was allowed to proceed at room temperature overnight. The reaction mixture was concentrated *in vacuo* and purified by silica column chromatography with a hexane:EtOAc solvent system.

Method G – Deprotection of Guanidine Derivatives

The corresponding Boc-protected guanidine (1 eq.) was dissolved in a 1:1 DCM:TFA mixture [0.05] M. The reaction was allowed to proceed for 24 hours then concentrated *in vacuo*.

Method H – Mitsunobu Reaction

To a solution of pseudothiourea (1 eq.), triphenylphosphine (1.5 eq.) in anhydrous THF [0.05] M (under an Ar atmosphere) was added alcohol (1 eq.) and diisopropyl azodicarboxylate (1.5 eq.). The reaction was allowed to proceed overnight at room

temperature. The reaction mixture was concentrated *in vacuo* and purified by flash column chromatography with a hexane:EtOAc solvent system.

Method I – Alkylation of 1-Phenylpiperazines

To a solution of 1-phenylpiperazine (1 eq.) and K_2CO_3 (1.5 eq.) in anhydrous MeCN [0.1] M was added propargyl bromide (1.2 eq.). The reaction mixture was allowed to proceed at room temperature overnight. The reaction mixture was concentrated *in vacuo*, dissolved in EtOAc and washed with water and brine. The organic phase was dried over $Na_2S_2O_4$ and concentrated *in vacuo*, then purified by silica column chromatography with a hexane:EtOAc solvent system.

Method J – Alkylation of 1-Phenylpiperazines

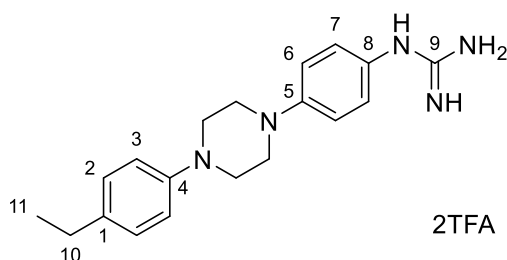
To a solution of 1-phenylpiperazine (1 eq.) and K_2CO_3 (3.5 eq.) in anhydrous DMF [0.1] M was added propargyl bromide (1.6 eq.). The reaction mixture was allowed to proceed at 90°C overnight. The reaction mixture was concentrated *in vacuo*, dissolved in EtOAc and washed with water and brine. The organic phase was dried over $Na_2S_2O_4$ and concentrated *in vacuo*. The crude mixture was use without further purification.

Method Q – Copper-catalysed Azide-Alkyne Cycloaddition Reactions

To a solution of alkyne (1 eq.) and azide (1 eq.) in anhydrous MeOH [0.05] M (additionally degassed with Ar), was added $CuSO_4 \cdot 5H_2O$ in water (0.08 M, 0.2 eq.) and $Na^+ Ascorbate$ (0.16 M, 0.4 eq.) under an Ar atmosphere. The reaction was allowed to proceed overnight at room temperature. Upon completion, the reaction mixture was diluted with EtOAc and washed with saturated NH_4Cl solution and brine. The organic phase was dried over $Na_2S_2O_4$, concentrated *in vacuo* and purified by silica flash column chromatography using a DCM:MeOH solvent system.

5.2.2 Synthesis and Characterisation

TFA salt of 1-(4-(4-(4-Ethylphenyl)piperazin-1-yl)phenyl)guanidine (**2**)



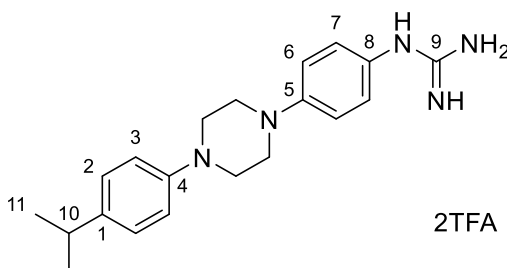
Following method G from **61** (30 mg, 0.05 mmol). **2** was obtained as a pale yellow powder (14 mg, 80%, R_f = 0.0 [Hexane:EtOAc 7:3]).

δ_H (600 MHz, DMSO- d_6): 9.32 (NH), 7.23 (NH), 7.21 (d, 2H, H_2 , J = 9.2 Hz), 7.16 (d, 2H, H_6 , J = 8.8 Hz), 7.10 (d, 2H, H_3 , J = 9.2 Hz), 7.02 (d, 2H, H_7 , J = 8.8 Hz), 3.34 (t, 4H, 2 x CH_2), 3.21 (t, 4H, 2 x CH_2), 2.49 (d, 3H, H_{11} , J = 7.8 Hz), 1.20 (d, 6H, H_{10} , J = 7.8 Hz).

δ_C (100 MHz, Acetone- d_6): 158.72 (C_9), 156.90 (C_4), 149.81 (C_5), 131.22 (C_8), 128.88 (C_1), 126.97 (C_1), 126.71 (C_6), 116.26 (C_7), 116.12 (C_3), 49.59 ($\underline{CH}_2$), 49.34 ($\underline{CH}_2$), 28.04 (C_{10}), 16.51 (C_{11}).

HRMS-ESI (+): m/z calculated for $C_{19}H_{26}N_5$: 324.2183. Found $[M+H] = 325.2401$.

TFA salt of 1-(4-(4-(4-Isopropylphenyl)piperazin-1-yl)phenyl)guanidine (**3**)



Following method **G** from **62** (25 mg, 0.046 mmol). **3** was obtained as a pale yellow powder (13.3 mg, 85%, R_f = 0.0 [Hexane:EtOAc 7:3], HPLC Purity 99% [R_t = 36.91 min]).

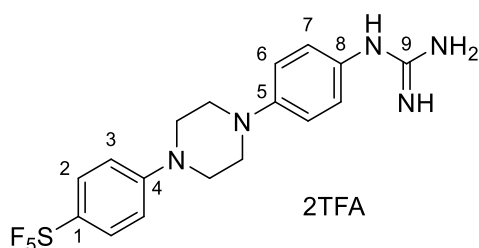
δ_H (600 MHz, DMSO- d_6): 9.36 (NH), 7.18 (NH), 7.12 (m, 4H, H_2 and H_6), 7.07 (d, 2H, H_3 , J = 8.8 Hz), 6.95 (d, 2H, H_7 , J = 8.8 Hz), 3.32 (t, 4H, 2 x CH_2), 3.24 (t, 4H, 2 x CH_2), 2.81 (m, 1H, H_{10}), 1.17 (d, 6H, H_{11} , J = 6.96 Hz).

δ_c (150 MHz, Acetone- d_6): 158.74 (C₉), 156.67 (C₅), 149.25 (C₄), 140.36 (C₈), 130.10 (C₁), 127.24 (C₂), 127.06 (C₆), 116.72 (C₃), 116.60 (C₇), 49.58 ($\underline{C}H_2$), 49.42 ($\underline{C}H_2$), 33.01 (C₁₀), 24.58 (C₁₀).

HRMS-ESI (+): m/z calculated for C₂₀H₂₇N₅: 337.2266. Found [M+H]= 338.2340.

ν_{max} (ATR)/cm⁻¹: 3490, 3160 (br, N-H, guanidinium salt), 2965-2850 (C-H aliphatic), 1683 (s, C=O TFA, C=N guanidinium), 1196-1118 (s, TFA C-F).

TFA salt of 1-(4-(4-(4-(Pentafluoro- λ 6-sulfaneyl)phenyl)piperazin-1-yl)phenyl)guanidine (**7**)



Following method G from **68** (50 mg, 0.08 mmol). **2** was obtained as a pale yellow powder (27.8 mg, 82%, R_f = 0.0 [Hexane:EtOAc 7:3], HPLC Purity 97% [R_t = 37.23 min]).

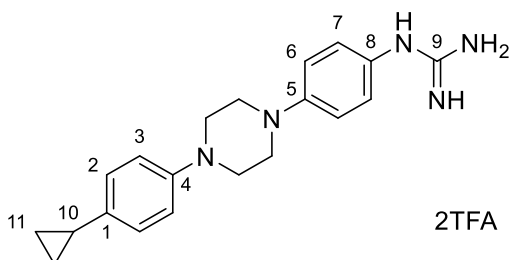
δ_H (600 MHz, DMSO- d_6): 9.31 (NH), 7.69 (d, 2H, H₂, J = 9.3 Hz), 7.16 (NH), 7.12 (d, 2H, H₆, J = 8.8 Hz), 7.09 (d, 2H, H₃, J = 9.3 Hz), 7.06 (d, 2H, H₇, J = 8.8 Hz), 3.47 (t, 4H, 2 x CH₂), 3.33 (t, 4H, 2 x CH₂).

δ_c (150 MHz, DMSO- d_6): 152.90 (C₉), 150.17 (C₄), 148.65 (C₅), 130.12 (C₈), 127.41 (C₂), 127.06 (C₆), 126.38 (C₁), 116.62 (C₃), 114.05 (C₇), 48.19 ($\underline{C}H_2$), 47.10 ($\underline{C}H_2$),

HRMS-ESI (+): m/z calculated for C₁₇H₂₀F₅N₅S: 421.1360. Found [M+H] = 422.1436.

ν_{max} (ATR)/cm⁻¹: 3160 (br, N-H, guanidinium), 2926-2839 (C-H, aliphatic), 1669 (s, C=O TFA + C=N guanidinium 1617), 1597 (C=N, guanidinium), 1232-1200 (s, C-F, TFA), 1134-1106 (S-F).

TFA salt of 1-(4-(4-(4-Cyclopropylphenyl)piperazin-1-yl)phenyl)guanidine (**4**)



Following method G from **69** (30 mg, 0.06 mmol). **2** was obtained as a pale yellow powder (14 mg, 75%, R_f = 0.0 [Hexane:EtOAc 7:3], HPLC Purity 98% [R_t = 36.95 min]).

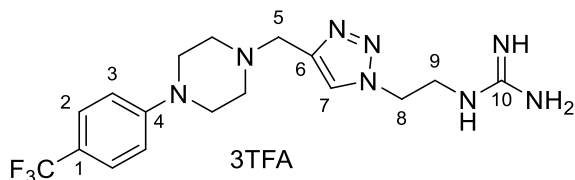
δ_H (600 MHz, DMSO- d_6): 9.35 (NH), 7.19 (NH), 7.12 (d, 2H, H_2 , J = 9.1 Hz), 7.06 (d, 2H, H_3 , J = 9.1 Hz), 6.98 (d, 2H, H_6 , J = 8.8 Hz), 6.91 (d, 2H, H_7 , J = 8.8 Hz), 3.31 (t, 4H, 2 x CH_2), 3.21 (t, 4H, 2 x CH_2), 1.84 (m, 1H, H_{10}), 0.87 (m, 2H, H_{11A}), 0.56 (m, 2H, H_{11B}).

δ_C (100 MHz, Acetone- d_6): 158.52 (C_9), 158.31 (C_5), 150.26 (C_4), 149.24 (C_8), 130.10 (C_1), 127.05 (C_2), 126.56 (C_6), 116.66 (C_3), 116.52 (C_7), 49.33 (CH_2), 48.57 (CH_2), 31.17 (C_{10}), 14.75 (C_{10}).

HRMS-ESI (+): m/z calculated for $C_{30}H_{41}N_5O_4$: 335.2110. Found $[M+H] = 336.2183$.

ν_{max} (ATR)/ cm^{-1} : 3349, 3150 (br, N-H, guanidinium salt), 2961(C-H aliphatic), 1668 (s, C=O TFA, C=N guanidinium), 1250-1176 (s, TFA C-F).

TFA salt of 1-(2-(4-((4-(4-(Trifluoromethyl)phenyl)piperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)guanidine (**13**)



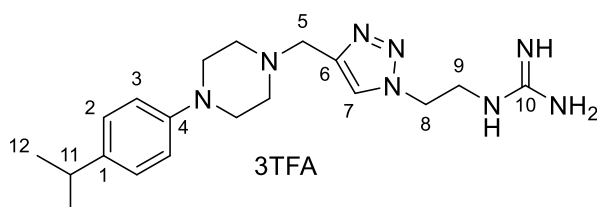
Following method G from **97** (25 mg, 0.04 mmol). **13** was obtained as a yellow gum (16 mg, 73%, R_f = 0.0 [Hexane:EtOAc 7:3], HPLC Purity 97% [R_t = 29.59 min]).

δ_{H} (600 MHz, DMSO- d_6): 10.38 (NH), 8.27 (s, 1H, H₇), 7.61 (NH), 7.58 (d, 2H, H₂, J = 8.4 Hz), 7.13 (d, 2H, H₃, J = 8.4 Hz), 4.56 (t, 2H, H₈, J = 5.8 Hz), 3.67 (q, 2H, H₉, J = 5.8 Hz), [3.77, 3.52, 3.20, 2.91 (s, br, 4 x CH₂)].

δ_{C} (150 MHz, DMSO- d_6): 158.40 (C₁₀), 157.24 (C₄), 130.12 (C₇), 126.83 (C₂) 126.62 (C₁), 126.16 (C₆), 115.41 (C₃), 65.40 (CH₂), 49.18 (CH₂), 41.07 (C₈).

HRMS-ESI (+): m/z calculated for C₁₇H₂₃F₃N₈ : 396.1998. Found [M+H] = 397.2074.

TFA salt of 1-(2-(4-((4-(4-Isopropylphenyl)piperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)guanidine (**14**)



Following method G from **97** (22 mg, 0.041 mmol). **14** was obtained as a yellow gum (15 mg, 80%, R_f = 0.0 [Hexane:EtOAc 7:3], HPLC Purity 99% [R_t = 31.61 min]).

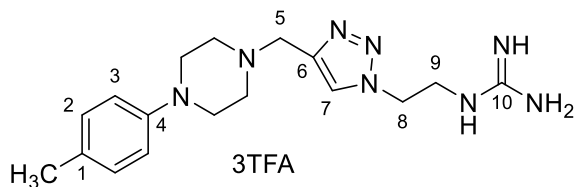
δ_{H} (600 MHz, DMSO- d_6): 10.24 (NH), 8.27 (s, 1H, H₇), 7.64 (NH), 7.14 (d, 2H, H₂, J = 8.4 Hz), 6.9 (d, 2H, H₃, J = 8.4 Hz), 4.57 (t, 2H, H₈, J = 5.0 Hz), 3.66 (br, m, 2H, H₉, J = 5.8 Hz), [3.77, 3.50, 3.20, 2.9 (s, br, 4 x CH₂)], 2.81 (m, 1H, H₁₁, J = 7.0 Hz), 1.16 (d, 6H, H₁₂, J = 7.0 Hz).

δ_{C} (150 MHz, DMSO- d_6): 158.81 (C₁₀), 156.78 (C₄), 147.58 (C₇), 129.62 (C₁), 126.82 (C₂), 126.16 (C₆), 116.22 (C₃), 50.42 (CH₂), 48.18 (CH₂), 41.55 (C₈), 32.48 (C₁₁), 24.02 (C₁₂).

HRMS-ESI (+): m/z calculated for C₁₉H₃₀N₈ : 370.2593. Found [M+H] = 371.2665.

ν_{max} (ATR)/cm⁻¹: 3349, 3162 (br, N–H, guanidinium), 2962 (C–H, aliphatic), 1666 (s, C=O TFA + C=N guanidinium), 1517 (Ar–C=C, triazole C=N/C=C), 1456, 1364, 1176–1128 (s, C–F TFA, C–N), 964–935 (triazole ring, C–N).

TFA salt of 1-(2-(4-((4-(p-Tolyl)piperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)guanidine (**15**)



Following method G from **100** (18 mg, 0.034 mmol). **15** was obtained as a yellow gum (11.2 mg, 83%, R_f = 0.0 [Hexane:EtOAc 7:3]).

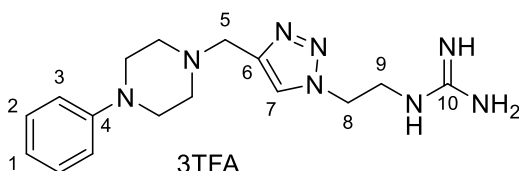
δ_H (600 MHz, DMSO- d_6): 10.24 (NH), 8.27 (s, 1H, H_7), 7.52 (NH), 7.12 (d, 2H, H_2 , J_{2-3} = 8.4 Hz), 6.88 (d, 2H, H_3 , J_{3-2} = 8.4 Hz), 4.59 (t, 2H, H_8 , J_{8-9} = 5.8 Hz), 3.65 (br, m, 2H, H_9 , J_{9-8} = 5.8 Hz), [3.78, 3.51, 3.19, 2.90 (br s, 4 $\times$ CH₂), 2.19 (s, CH₃)

δ_C (150 MHz, DMSO- d_6): 158.82 (C₁₀), 156.76 (C₄), 146.58 (C₇), 128.55 (C₁), 125.33 (C₂) 126.16 (C₆), 114.33 (C₃), 50.41 (CH₂), 48.16 (CH₂), 41.55 (C₈), 32.48 (C₁₁), 24.02 (C₁₂).

HRMS-ESI (+): m/z calculated for C₁₇H₂₆N₈ : 342.2280. Found [M+H] = 343.2317.

ν_{max} (ATR)/cm⁻¹: 3488, 3349, 3159 (br, N-H, guanidinium), 2965-2926 (C-H, aliphatic), 1675 (s, C=O TFA + C=N guanidinium), 1517 (Ar-C=C, triazole C=N/C=C), 1397, 1330, 1255-1195 (s, C-F TFA, C-N), 1117-1016 (triazole ring, C-N).

TFA salt of 1-(2-(4-((4-Phenylpiperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)guanidine (**16**)



Following method G from **101** (18 mg, 0.034 mmol). **15** was obtained as a yellow gum (11.2 mg, 83%, R_f = 0.0 [Hexane:EtOAc 7:3], HPLC Purity 95% [R_t = 21.09 min]).

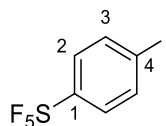
δ_{H} (600 MHz, DMSO- d_6): 11.32 (NH), 8.26 (s, 1H, H₇), 7.64 (NH), 7.27 (t, 2H, H₂, J = 8.0 Hz), 6.99 (d, 2H, H₃, J = 8.0 Hz), 6.88 (t, 1H, H₁, J = 8.0 Hz), 4.57 (t, 2H, H₈, J = 5.9 Hz), 3.67 (q, 2H, H₉, J = 5.9 Hz), [3.85, 3.51, 3.20, 2.96 (s, br, 4 x CH₂)].

δ_{C} (150 MHz, DMSO- d_6): 158.87 (C₁₀), 157.35 (C₄), 149.96 (C₇), 129.63 (C₂), 127.74 (C₁), 120.65 (C₆), 116.50 (C₃), 50.86 (CH₂), 49.23 (CH₂), 46.04 (C₈), 41.05 (C₉).

HRMS-APCI (+): m/z calculated for C₁₆H₂₄N₈: 328.2124. Found [M-H] = 329.2195.

ν_{max} (ATR)/cm⁻¹: 3349, 3150 (br, N-H, guanidinium), 2965-2926 (C-H, aliphatic), 1666 (s, C=O TFA + C=N guanidinium), 1601, 1496 (Ar-C=C, triazole C=N/C=C), 1397, 1330, 1179-1127 (s, C-F TFA, C-N), 1029 (triazole ring, C-N).

4-Pentafluorosulfonyl-iodobenzene (**24**)



To a solution of aniline (1 g, 4.56 mmol, 1 eq.) in Et₂O [0.1] M at 0°C was added H₂SO₄ (491.56 mg, 268.81 μ L, 1.1 eq.) dropwise, ensuring that the temperature was maintained below 0°C. This was followed by the addition of isopropyl nitrite (18.24 mL, [0.25] M). An aqueous solution of KI (1.51 g, 9.12 mmol, 2 eq., 10 mL water) was added. When bubbling subsided, NaCNS was added. The reaction mixture was diluted with Et₂O. The organic phase was washed with a saturated NaCNS solution, dried over Na₂S₂O₄ and concentrated by silica flash column chromatography (Hexane 100%), affording an orange crystalline powder (782.9 mg, 52%, R_f = 0.95 [Hexane:EtOAc 9:1]). The structural data are consistent with previously reported characterisation for **24**.²²²

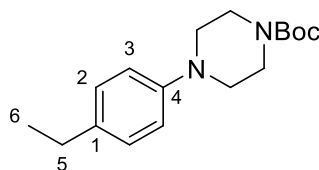
δ_{H} (400 MHz, Acetone- d_6): 8.05 (d, 2H, H₃, J = 9.1 Hz), 7.70 (d, 2H, H₂, J = 9.1 Hz).

δ_{C} (150 MHz, Acetone- d_6): 132.24 (C₁), 129.29 (C₃), 127.67 (C₂), 98.55 (C₄)

HRMS-ESI (+): m/z calculated for C₆H₄F₅IS: 329.8999. Not found.

HRMS-APCI: m/z calculated for C₆H₄F₅IS: 329.8999. Not found.

1-(4-Ethylphenyl)piperazine -Butyl 4-(4-ethylphenyl)piperazine-1-carboxylate (**29**)



Following method A, from **22** (698.18 mg, 3 mmol). **29** was obtained as a white powder (653.42 mg, 75%, R_f = 0.33 [Hexane:EtOAc 9:1]).

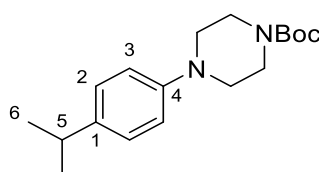
δ_H (400 MHz, Acetone- d_6): 7.10 (d, 2H, H_3 , J = 8.8 Hz), 6.91 (d, 2H, H_2 , J = 8.8 Hz), 3.54 (t, 4H, 2 x CH_2), 3.08 (t, 4H, 2 x CH_2), 2.56 (q, 2H, H_5 , J = 7.6 Hz), 1.81 (t, 3H, H_6 , J = 7.6 Hz), 1.47 (s, 9H, Boc- H).

δ_C (150 MHz, Acetone- d_6): 154.72 (Boc- $\underline{C}O$), 150.31 (C_4), 136.14 (C_1), 128.83 (C_2), 117.31 (C_3), 79.38 (Boc- $\underline{C}(CH_3)_3$), 50.19 ($\underline{C}H_2$), 31.97 (C_5), 28.24 (Boc- $\underline{C}(CH_3)_3$), 15.97 (C_6).

HRMS-ESI (+): m/z calculated for $C_{17}H_{26}N_2O_2$: 290.1994. Found $[M+H] = 291.2062$.

ν_{max} (ATR)/ cm^{-1} : 3005 (w, Ar-H), 2975-2858 (C-H, aliphatic), 1685 (C=O).

tert-Butyl 4-(4-isopropylphenyl)piperazine-1-carboxylate (**30**)



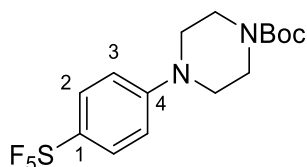
Following method A, from **23** (738.27 mg, 3 mmol). **30** was obtained as a white powder (711.86 mg, 78%, R_f = 0.33 [Hexane:EtOAc 9:1]).

δ_H (400 MHz, DMSO- d_6): 7.14 (d, 2H, H_2 , J = 8.8 Hz), 6.92 (d, 2H, H_3 , J = 8.8 Hz), 3.54 (t, 4H, 2 x CH_2), 3.08 (t, 4H, 2 x CH_2), 2.82 (m, 1H, H_5), 1.47 (s, 9H, Boc- H), 1.21 (d, 6H, H_6 , J = 6.9 Hz).

δ_c (150 MHz, DMSO- d_6): 154.19 (Boc-C=O), 149.81 (C₄), 140.23 (C₁), 126.69 (C₂), 116.68 (C₃), 78.82 (Boc-C(CH₃)₃), 49.56 (CH₂), 33.07 (C₅), 27.64 (Boc-C(CH₃)₃), 23.55 (C₆).

HRMS-ESI (+): m/z calculated for C₁₇H₂₆N₂O₂: 304.2151. Found [M+H] = 305.2223.

tert-Butyl 4-(4-(pentafluoro- λ^6 -sulfaneyl)phenyl)piperazine-1-carboxylate (**31**)



Following method A, from **24** (495.08 mg, 1.5 mmol). **31** was obtained as an orange powder (390.3 mg, 67%, R_f = 0.25 [Hexane:EtOAc 9:1]).

δ_H (600 MHz, Acetone- d_6): 7.69 (d, 2H, H₃, J = 9.3 Hz), 7.07 (d, 2H, H₂, J = 9.3 Hz), 3.58 (t, br, 4H, 2 x CH₂), 3.37 (t, 4H, 2 x CH₂), 1.47 (s, 9H, Boc-H),

δ_c (150 MHz, Acetone- d_6): 154.14 (Boc-C=O), 152.93 (C₁), 126.97 (C₂), 113.80 (C₃), 78.82 (Boc-C(CH₃)₃), 47.18 (CH₂), 27.63 (Boc-C(CH₃)₃).

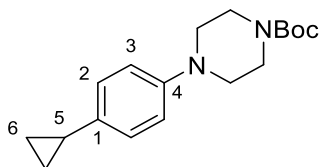
δ_F (400 MHz, Acetone- d_6): 87.64 (quintet, J = 148.50 Hz), 64.05 (d, J = 149.08 Hz).

HRMS-ESI (+): m/z calculated for C₁₅H₂₁F₅N₂O₂S: 388.1244. Not found.

HRMS-APCI: m/z calculated for C₁₅H₂₁F₅N₂O₂S: 388.1244. Not found.

ν_{max} (ATR)/cm⁻¹: 2932(C-H, aliphatic), 1671 (s, C=O), 1596, 1512 (Ar-C=C), 1367 (*tert*-butyl), 1289-1172 (C-O, C-N), 1127-1082 (S-F).

tert-Butyl 4-(4-cyclopropylphenyl)piperazine-1-carboxylate (**43**)



Following method B, from **35** (341.25 mg, 1 mmol). **31** was obtained as a white powder (251.00 mg, 83%, $R_f=0.40$ [PhMe:Hexane 95:5]).

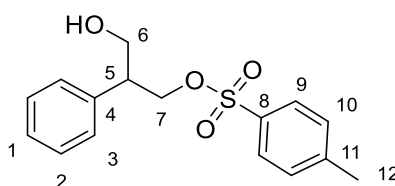
δ_H (600 MHz, Acetone- d_6): 6.99 (d, 2H, H_2 , $J = 8.8$ Hz), 6.89 (d, 2H, H_3 , $J = 8.8$ Hz), 3.53 (t, br, 4H, 2 x CH_2), 3.07 (t, 4H, 2 x CH_2), 1.84 (m, 1H, H_5), 1H, H 1.47 (s, 9H, Boc-H), 0.87 (m, 2H, H_{6A}), 0.58 (m, 2H, H_{6B}).

δ_C (150 MHz, Acetone- d_6): 154.58 (Boc-C=O), 150.08 (C_4), 135.62 (C_1), 126.68 (C_2), 117.17 (C_3), 79.36 (Boc-C(CH_3)₃), 50.17 (CH_2), 28.17 (Boc-C(CH_3)₃), 14.79 (C_5), 8.61 (C_6).

HRMS-ESI (+): m/z calculated for $C_{18}H_{26}N_2O_2$: 302.1994. Found $[M+H] = 303.2067$.

ν_{max} (ATR)/ cm^{-1} : 3002 (w, Ar-H), 2977-2819 (C-H, aliphatic), 1681 (s, C=O), 1612, 1519 (Ar-C=C), 1493-1454 (CH_2), 1389-1363 (*tert*-butyl), 1281-1157 (C-N, Ar-C-N), 919, 819.

3-Hydroxy-2-phenylpropyl 4-methylbenzenesulfonate (**49**)



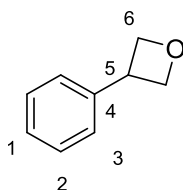
2-Phenyl-1,3-diol (1.6 g, 10.51 mmol, 2.1 eq.) was dissolved in dry THF (12.5 mL, 0.4 M), followed by the addition of TEA (0.85 mL, 6 mmol, 1.2 eq.) under N_2 . TsCl (953.2 mg, 5 mmol, 1 eq.) was added to the mixture and was allowed to proceed for 24 hours at room temperature ($R_f = 0.65$, [DCM:EtOAc 7:3]). Upon completion, reaction mixture was diluted with dichloromethane (100 mL) and was washed with 1 M HCl (1 x 80 mL), 10%_{w/v} $NaHCO_3$ (2 x 80 mL) and brine (1 x 50 mL). The organic phase was dried over Na_2SO_4 , filtered and concentrated at low pressure. The crude mixture was purified *via* flash column chromatography (Hexane:EtOAc 1:0 to 7:3) yielding a white powder (618 mg, 39%).

δ_H (600 MHz, DMSO- d_6) = 7.65 (d, 2H, H_9 , $J = 7.9$ Hz), 7.43 (d, 2H, H_8 , $J = 7.9$ Hz), 7.29 – 7.18 (m, 3H, H_1 , H_2), 7.14 (d, 2H, H_3 , $J = 8.6$ Hz), 4.85 (t, 1H, OH, $J_{OH-6} = 5.0$ Hz), 4.37 –

4.20 (m, 2H, H₆), 3.52 (dd, 2H, H_{7a}, H_{7b} J_{7a-5} = 5.5 Hz, J_{7b-5} = 6.6 Hz), 3.07 – 2.97 (m, 1H, H₅), 2.41 (s, 3H, H₁₂).

δ_c (100 MHz, DMSO-*d*₆) = 145.3 (C₈), 139.4 (C₄), 132.6 (C₁₁), 130.5 (C₃), 128.7 (C₂), 128.4 (C₁₀), 127.9 (C₉), 127.7 (C₁), 72.0 (C₆), 62.4 (C₇), 47.5 (C₅), 21.5 (C₁₂).

3-Phenyloxetane (**46**)



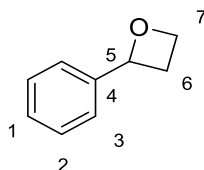
Monotosylated 2-phenyl-1,3-diol **X** (618 mg, 2.01 mmol, 1 eq.) was dissolved in tBuOH (8.15 mL, 0.2 M) at 50°C under Ar. tBuOK (338.31 mg, 3.02 mmol, 1.5 eq.) was added and the reaction was allowed to proceed for 1 hour (*R_f* = 0.85 [DCM:EtOAc 7:3]). Upon completion, reaction mixture was diluted with EtOAc (ca. 200 mL) and washed with H₂O (2 x 150 mL) and brine (1x 100 mL). The organic phase was dried over Na₂SO₄ and concentrated *in vacuo*. The crude mixture was purified *via* distillation at 110°C yielding a colourless oil (400 mg, 65%). The structural data are consistent with previously reported characterisation for **46**.²²³

δ_H NMR (400 MHz, DMSO-*d*₆) = 7.40 – 7.30 (m, 4H, H₂, H₃), 7.24 (t, 1H, H₁, J = 6.9 Hz), 4.92 (dd, 1H, H_{6a}, J_a = 8.1, J_b = 6.0 Hz), 4.59 (t, 1H, H_{6b}, J_b = 8.1 Hz), 4.22 (q, 1H, H₅).

δ_c NMR (100 MHz, DMSO-*d*₆) = 142.1 (C₄), 129.0 (C₂), 127.2 (C₁, C₃), 77.9 (C₆), 39.7 (C₅).

HRMS-APCI = *m/z* calculated for C₉H₁₀O: 134.0732. Found [M+H] = 135.0807.

2-Phenyloxetane (**47**)



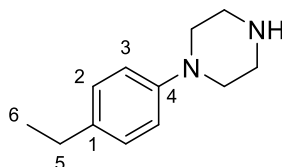
Sodium metal (344.7 mg, 15 mmol, 1.5 eq.) was added to 40 mL of tBuOH under Ar atmosphere. The reaction was allowed to proceed at 40°C for two hours until all the sodium has reacted. 3-chloro-1-phenylpropan-1-ol (1.7 g, 10 mmol, 1 eq.) was then added to the reaction mixture and the reaction was allowed to proceed at 40°C for 1 hour ($R_f = 0.80$, [DCM:EtOAc 7:3]). The crude mixture was diluted with H₂O (250 mL) and the product was extracted with EtOAc (2 x 100 mL). The combined organic phases were dried over Na₂SO₄ and concentrated at low pressure. The crude product was then purified *via* flash column chromatography (Hexane with a gradient of EtOAc of 0% to 30%) yielding a colourless oil (1046.6 mg, 78% yield). The structural data are consistent with previously reported characterisation for **47**.²²⁴

δ_H NMR (400 MHz, DMSO-*d*₆) = 7.47 – 7.32 (m, 4H, H₂, H₃), 7.28 (t, 1H, H₁, $J = 6.9$ Hz), 5.70 (t, 1H, H₅, $J_a = J_7 = 7.6$ Hz), 4.70 – 4.60 (m, 1H, H_{7a}), 4.57 – 4.46 (m, 1H, H_{7b}), 3.01 – 2.89 (m, 1H, H_{6a}), 2.59 – 2.44 (m, 1H, H_{6a}).

δ_C NMR (400 MHz, DMSO-*d*₆) = 144.1 (C₄), 128.8 (C₂), 128.1 (C₁), 125.7 (C₃), 82.1 (C₅), 67.6 (C₇), 30.6 (C₆).

HRMS-APCI = m/z calculated for C₉H₁₀O: 134.0732. Found [M+H] = 135.0814

1-(4-Ethylphenyl)piperazine (**44**)



Following method C, from **29** (435.9 mg, 2.29 mmol). **44** was obtained as a pale yellow powder (quant., $R_f = 0.1$ [Hexane:EtOAc 7:3]). The structural data are consistent with previously reported characterisation for **44**.²²⁵

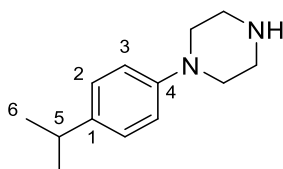
δ_H (400 MHz, Acetone- d_6): 7.10 (d, 2H, H_3 , $J = 8.8$ Hz), 6.91 (d, 2H, H_2 , $J = 8.8$ Hz), 3.09 (t, 4H, 2 x CH_2), 2.97 (t, 4H, 2 x CH_2), 2.54 (q, 2H, H_5 , $J = 7.6$ Hz), 1.18 (t, 3H, H_6 , $J = 7.6$ Hz).

δ_C (100 MHz, Acetone- d_6): 150.93 (C_4), 135.30 (C_1), 128.79 (C_2), 116.66 (C_3), 50.71 (CH_2), 46.32 (CH_2), 28.24 (C_5), 15.97 (C_6).

HRMS-ESI (+): m/z calculated for $C_{12}H_{18}N_2$: 190.1470. Found $[M+H] = 191.1546$.

ν_{max} (ATR)/ cm^{-1} : 3284 (br, N-H), 2961-2818 (C-H, aliphatic), 1671, 1514 (Ar-C=C, N-H bend), 1323-1053 (C-N, Ar-C-N), 913, 821, 783.

1-(4-Isopropylphenyl)piperazine (**45**)



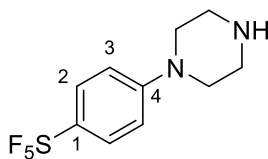
Following method C, from **30** (366.3 mg, 1.79 mmol). **45** was obtained as a pale yellow powder (quant. $R_f = 0.1$ [Hexane:EtOAc 7:3])..

δ_H (400 MHz, DMSO- d_6): 7.10 (d, 2H, H_2 , $J = 8.2$ Hz), 6.92 (d, 2H, H_3 , $J = 8.2$ Hz), 3.12 (t, 4H, 2 x CH_2), 2.99 (t, 4H, 2 x CH_2), 2.80 (m, 1H, H_5), 1.18 (d, 6H, H_6 , $J = 7.0$ Hz).

δ_C (100 MHz, DMSO- d_6): 150.27 (C_4), 139.66 (C_1), 126.68 (C_2), 116.10 (C_3), 49.56 (CH_2), 45.54 (CH_2), 33.03 (C_5), 23.56 (C_6).

HRMS-ESI (+): m/z calculated for $C_{13}H_{20}N_2$: 204.1626. Found $[M+H] = 205.1704$.

1-(4-(Pentafluoro- λ^6 -sulfaneyl)phenyl)Piperazine (**46**)



Following method C, from **31** (380 mg, 0.98 mmol). **46** was obtained as a red-orange powder (quant., $R_f = 0.1$ [Hexane:EtOAc 7:3]).

δ_H (400 MHz, Acetone- d_6): 7.75 (d, 2H, H_2 , $J = 9.3$ Hz), 7.2 (d, 2H, H_3 , $J = 9.3$ Hz), 3.81 (t, br, 4H, 2 x CH_2), 3.65 (t, 4H, 2 x CH_2).

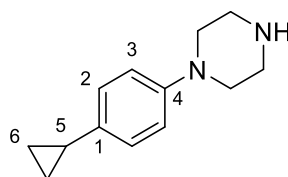
δ_C (100 MHz, Acetone- d_6): 158.41 (C_1), 152.15 (C_2), 127.19 (C_3), 114.68 (C_4), 44.90 (CH_2), 43.33 (CH_2).

δ_F (376 MHz, Acetone- d_6): 87.97 (quintet, $J = 148.50$ Hz), 63.79 (d, $J = 148.50$ Hz).

HRMS-ESI (+): m/z calculated for $C_{10}H_{13}F_5N_2S$: 288.0795. Found $[M+H] = 289.0798$.

ν_{max} (ATR)/ cm^{-1} : 3200 (br, N-H), 2968-2830 (C-H, aliphatic), 1596, 1513 (Ar-C=C), 1329-1206 (C-N, Ar-C-N), 1158-1088 (S-F).

1-(4-Cyclopropylphenyl)piperazine (**47**)



Following method C, from **34** (140 mg, 0.462 mmol). **47** was obtained as a pale yellow powder (quant., $R_f = 0.1$ [Hexane:EtOAc 7:3]).

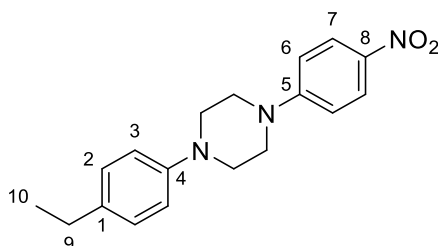
δ_H (400 MHz, Acetone- d_6): 7.04 (d, 2H, H_2 , $J = 8.8$ Hz), 6.97 (d, 2H, H_3 , $J = 8.8$ Hz), 3.62 (t, br, 4H, 2 x CH_2), 3.52 (t, 4H, 2 x CH_2), 1.87 (m, 1H, H_5), 0.89 (m, 2H, H_{6A}), 0.60 (m, 2H, H_{6B}).

δ_C (100 MHz, Acetone- d_6): 148.28 (C_4), 136.53 (C_1), 126.43 (C_2), 117.08 (C_3), 46.88 (CH_2), 43.84 (CH_2), 14.20 (C_5), 8.10 (C_6).

HRMS-ESI (+): m/z calculated for $C_{13}H_{18}N_2$: 202.1470. Found $[M+H] = 203.1546$.

ν_{max} (ATR)/ cm^{-1} : 3030 (w, Ar-H, cyclopropyl C-H), 2837 (C-H, aliphatic), 1664 (N-H), 1508 (Ar-C=C), 1464-1421 (CH_2 , Ar-C-C) 1031-956 (cyclopropyl ring).

1-(4-Ethylphenyl)-4-(4-nitrophenyl)piperazine (**48**)



Following method D, from **44** (130 mg, 0.68 mmol). **48** was obtained as a pale yellow powder (156.92, 74%, R_f = 0.1 [Hexane:EtOAc 7:3]).

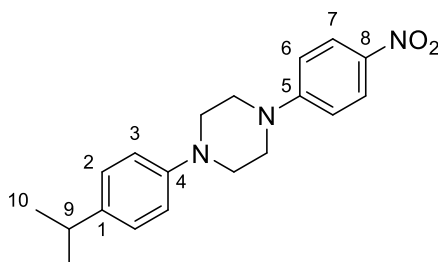
δ_H (400 MHz, Acetone- d_6): 8.13 (d, 2H, H_7 , J = 9.5 Hz), 7.13 (d, 4H, Overlap of H_2 & H_2 , J = J = 9.5 Hz), 6.96 (d, 2H, H_6 , J = 8.8 Hz), 3.69 (t, 4H, 2 x CH_2), 3.33 (t, 4H, 2 x CH_2), 2.56 (q, 2H, H_9 , J_{9-10} = 8.2 Hz), 1.19 (t, 3H, H_{10} , J_{10-9} = 8.2 Hz).

δ_C (100 MHz, Acetone- d_6): 155.07 (C_5), 149.35 (C_4), 138.13 (C_8), 135.62 (C_1), 128.53 (C_2), 125.53 (C_7), 116.44 (C_6), 112.87 (C_3), 49.08 ($\underline{C}H_2$), 46.85 ($\underline{C}H_2$), 27.67 (C_5), 15.44 (C_6).

HRMS-APCI: m/z calculated for $C_{18}H_{21}N_3O_2$: 311.1634. Found $[M+H] = 312.1702$.

ν_{max} (ATR)/ cm^{-1} : 3030 (w, Ar-H), 2967-2839 (C-H, aliphatic), 1588 (NO_2), 1514 (Ar-C=C), 1389-1317 (NO_2), 1276-1105 (C-N, Ar-C-N), 973, 820-784 (NO_2).

1-(4-Isopropylphenyl)-4-(4-nitrophenyl)piperazine (**49**)



Following method D, from **45** (203.8 mg, 0.99 mmol). **49** was obtained as a pale yellow powder (139.9 mg, 43%, R_f = 0.8 [Hexane:EtOAc 7:3]).

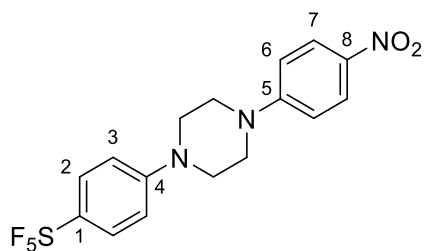
δ_{H} (600 MHz, Acetone- d_6): 8.13 (d, 2H, H_7 , $J = 9.5$ Hz), 7.16 (d, 2H, H_3 , $J = 9.5$ Hz), 7.12 (d, 2H, H_2 , $J = 9.5$ Hz), 6.97 (d, 2H, H_6 , $J = 9.5$ Hz), 3.69 (t, 4H, 2 x CH_2), 3.34 (t, 4H, 2 x CH_2), 2.85 (m, 1H, H_9), 1.22 (d, 6H, H_{10} , $J_{10-9} = 6.9$ Hz).

δ_{C} (150 MHz, Acetone- d_6): 155.10 (C_5), 149.38 (C_4), 140.18 (C_8), 138.16 (C_1), 126.81 (C_2), 125.53 (C_7), 116.35 (C_6), 112.83 (C_3), 49.08 (CH_2), 46.99 (CH_2), 32.98 (C_5), 23.54 (C_6).

HRMS-APCI: m/z calculated for $\text{C}_{19}\text{H}_{23}\text{N}_3\text{O}_2$: 325.1790. Not found.

HRMS-APCI: m/z calculated for $\text{C}_6\text{H}_4\text{F}_5\text{IS}$: 409.0883. Not found.

1-(4-Nitrophenyl)-4-(4-(pentafluoro-l6-sulfaneyl)phenyl)Piperazine (**58**)



Following method D, from **46** (590 mg, 0.98 mmol). **50** was obtained as a yellow powder (190 mg, 47%, $R_f = 0.5$ [Hexane:EtOAc 7:3]) post purification by silica flash column chromatography.

δ_{H} (600 MHz, Acetone- d_6): 8.14 (d, 2H, H_7 , $J = 9.5$ Hz), 7.72 (d, 2H, H_3 , $J = 9.5$), 7.10 (d, 2H, H_2 , $J = 9.5$), 7.09 (d, 2H, H_6 , $J = 9.5$ Hz), 3.77 (t, 4H, 2 x CH_2), 3.65 (t, 4H, 2 x CH_2).

δ_{C} (150 MHz, Acetone- d_6): 154.75 (C_5), 152.40.35 (C_1), 138.17 (C_8), 143.6 (C_4), 127.07 (C_2), 125.60 (C_7), 113.11 (C_6), 112.56 (C_3), 46.37 (CH_2), 46.11 (CH_2).

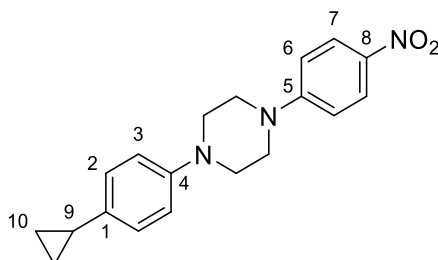
δ_{F} (376 MHz, Acetone- d_6): 87.94 (quintet, $J = 147.50$ Hz), 64.14 (d, $J = 148.05$ Hz).

HRMS-ESI (+): m/z calculated for $\text{C}_{16}\text{H}_{16}\text{F}_5\text{N}_3\text{O}_2\text{S}$: 409.0883. Not found.

HRMS-APCI: m/z calculated for $\text{C}_6\text{H}_4\text{F}_5\text{IS}$: 409.0883. Not found.

ν_{max} (ATR)/ cm^{-1} : 2968-2862 (C-H, aliphatic), 1588 (NO_2), 1491 (Ar-C=C), 1394-1316 (NO_2), 1276-1105 (C-N, Ar-C-N, S-F), 1113-1092 (S-F).

1-(4-Cyclopropylphenyl)-4-(4-nitrophenyl)Piperazine (**51**)



Following method D, from **47** (286.2 mg, 0.46 mmol). **51** was obtained as a yellow powder (190 mg, 47%, R_f = 0.5 [Hexane:EtOAc 7:3]) post purification by silica flash column chromatography.

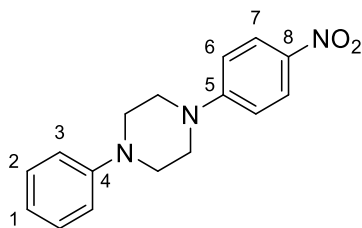
δ_H (600 MHz, Acetone- d_6): 8.13 (d, 2H, H_7 , J = 9.2 Hz), 7.12 (d, 2H, H_3 , J = 9.5 Hz), 7.02 (d, 2H, H_2 , J = 9.5 Hz), 6.94 (d, 2H, H_6 , J = 9.2 Hz), 3.68 (t, 4H, 2 x CH_2), 3.32 (t, 4H, 2 x CH_2), 1.86 (m, 1H, H_9), 0.88 (m, 2H, H_{6A}), 0.59 (m, 2H, H_{6B}).

δ_C (150 MHz, Acetone- d_6): 155.08 (C_5), 149.30 (C_4), 138.15 (C_8), 135.13 (C_1), 126.31 (C_2), 125.55 (C_7), 116.41 (C_6), 112.84 (C_3), 49.17 ($\underline{CH}_2$), 46.90 ($\underline{CH}_2$), 14.23 (C_9), 8.06 (C_{10}).

HRMS-ESI (+): m/z calculated for $C_{19}H_{21}F_5N_3O_2$: 423.1634. Not found.

HRMS-APCI: m/z calculated for $C_{19}H_{21}F_5N_3O_2$: 423.1634. Not found.

1-(4-Nitrophenyl)-4-Phenylpiperazine (**52**)



To a solution of 1-phenylpiperazine (516.6 mg, 5 mmol, 1 eq.) in DMSO (5 mL, 1 M) was added 4-fluoronitrobenzene (811 mg, 5 mmol, 1 eq.). The reaction was allowed to proceed at 100°C for 48 hours. Upon completion, the reaction mixture was allowed to cool followed by washing with diethyl ether under vacuum, affording **52** (736.67 g,

57%). The structural data are consistent with previously reported characterisation for **52**.²²⁶

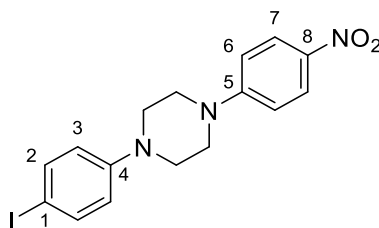
δ_{H} (600 MHz, DMSO- d_6): 8.11 (d, 2H, H₇, J = 9.2 Hz), 7.25 (d, 2H, H₂, J₂₋₁ = J = 7.77 Hz), 7.10 (d, 2H, H₃, J = 7.75 Hz), 7.01 (d, 2H, H₆, J = 9.2 Hz), 3.69 (t, 4H, 2 x CH₂), 3.38 (t, 4H, 2 x CH₂).

δ_{C} (150 MHz, DMSO- d_6): 155.10 (C₅), 151.02 (C₄), 137.47 (C₈), 129.51 (C₂), 126.31 (C₇), 116.03 (C₆), 113.16 (C₃), 48.24 (CH₂), 46.63 (CH₂).

HRMS-ESI (+): m/z calculated for C₁₉H₂₁F₅N₃O₂: 423.1634. Not found.

HRMS-APCI: m/z calculated for C₁₉H₂₁F₅N₃O₂: 423.1634. Not found.

1-(4-Iodophenyl)-4-(4-nitrophenyl)Piperazine (**53**)



To a solution of **52** (432.5 mg, 1.5 mmol, 1 eq.) in methyl *tert*-butyl ether (6 mL, 0.25 M) was added DMSO (5.25 mmol, 372.89 μ L, 3.5 eq.) and iodine (2.7 mmol, 410.18 mg, 1.5 eq.). The reaction was allowed to proceed at 30°C overnight. Upon completion, the crude mixture was diluted by EtOAc (30 mL) and washed with a saturated NaCNS (3 x 30 mL) followed by water (3 x 30 mL) and dried over Na₂S₂O₄ and concentrated *in vacuo* affording **53** as a brown powder (592 mg, 78%, R_f = 0.1 [Hexane:EtOAc 7:3]).

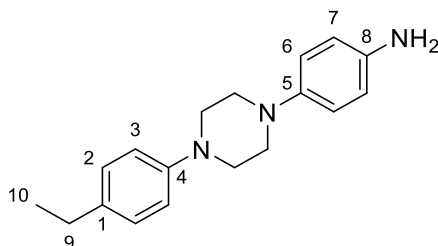
δ_{H} (600 MHz, Acetone- d_6): 8.13 (d, 2H, H₇, J = 9.2 Hz), 7.57 (d, 2H, H₂, J₂₋₁ = J = 9.1 Hz), 7.11 (d, 2H, H₃, J = 9.1 Hz), 6.88 (d, 2H, H₆, J = 9.2 Hz), 3.72 (t, 4H, 2 x CH₂), 3.43 (t, 4H, 2 x CH₂).

δ_{C} (150 MHz, DMSO- d_6): 154.96 (C₅), 150.81 (C₄), 137.84 (C₂), 129.01 (C₈), 125.53 (C₇), 129.51 (C₂), 126.31 (C₇), 118.08 (C₃), 116.03 (C₆), **80.41 (C₁)**, 47.81 (CH₂), 46.50 (CH₂).

HRMS-ESI (+): m/z calculated for C₁₉H₂₁F₅N₃O₂: 409.0297. Not found.

HRMS-APCI: m/z calculated for C₁₉H₂₁F₅N₃O₂: 409.0287. Not found.

4-(4-(4-Ethylphenyl)piperazin-1-yl)Aniline (**54**)



Following method E, from **49** (156 mg, 0.50 mmol). **54** was obtained as a white powder (128.6 mg, 90%, R_f = 0.05 [Hexane:EtOAc 7:3]).

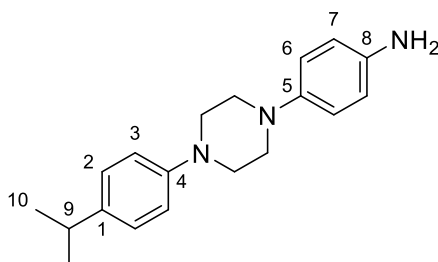
δ_H (400 MHz, Acetone-*d*₆): 7.11 (d, 2H, H₂, J = 8.4 Hz), 6.98 (d, 2H, H₃, J = 8.4 Hz), 6.96 (d, 2H, H₆, J = 8.4 Hz), 6.61 (d, 2H, H₇, J = 8.4 Hz), 3.28 (t, 4H, 2 x CH₂), 3.33 (t, 4H, 2 x CH₂), 2.56 (q, 2H, H₉, J₉₋₁₀ = 7.7 Hz), 1.19 (t, 3H, H₁₀, J₁₀₋₉ = 7.7 Hz).

δ_C (100 MHz, Acetone-*d*₆): 149.76 (C₄), 147.61 (C₅), 144.83 (C₈), 135.25 (C₁), 128.53 (C₂), 120.31 (C₇), 117.02 (C₃), 116.37 (C₆), 50.04 (CH₂), 49.68 (CH₂), 27.66 (C₉), 15.44 (C₁₀).

HRMS-ESI (+): m/z calculated for C₁₈H₂₃N₃: 281.1892. Found [M+H] = 282.1962.

ν_{max}(ATR)/cm⁻¹: 3358, 3207 (br, N-H), 2956-2832 (C-H, aliphatic), 1611, 1514 (Ar-C=C, N-H bend) 1319-1101 (C-N, Ar-C-N), 859, 821, 794, 736, 724, 704.

4-(4-(4-Isopropylphenyl)piperazin-1-yl)aniline (**55**)



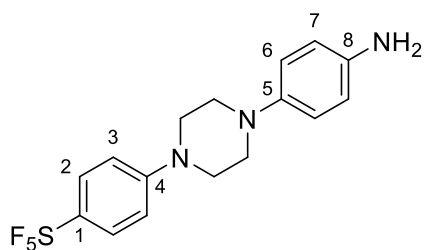
Following method E, from **49** (139.9 mg, 0.43 mmol). **55** was obtained as a white yellow powder (85.5 mg, 69%, R_f = 0.05 [Hexane:EtOAc 7:3]).

δ_{H} (400 MHz, Acetone- d_6): 7.12 (d, 2H, H_2 , $J = 8.7$ Hz), 6.93 (d, 2H, H_3 , $J = 8.7$ Hz), 6.80 (d, 2H, H_6 , $J = 8.9$ Hz), 6.61 (d, 2H, H_7 , $J = 8.9$ Hz), 3.24 (t, 4H, 2 x CH_2), 3.10 (t, 4H, 2 x CH_2), 2.8 (m, 1H, H_9), 1.19 (d, 6H, H_{10} , $J_{10-9} = 7.3$ Hz).

δ_{C} (150 MHz, Acetone- d_6): 155.10 (C_5), 149.38 (C_4), 140.18 (C_8), 138.08 (C_1), 126.81 (C_2), 125.53 (C_7), 116.35 (C_6), 112.83 (C_3), 49.03 (CH_2), 46.86 (CH_2), 33.08 (C_9), 23.62 (C_{10}).

HRMS-ESI (+): m/z calculated for $\text{C}_{19}\text{H}_{25}\text{N}_3$: 295.2048. Found $[\text{M}+\text{H}] = 296.2122$.

4-(4-(4-(Pentafluoro- λ 6-sulfaneyl)phenyl)piperazin-1-yl)aniline (**56**)



Following method E, from **50** (190 mg, 0.5 mmol). **55** was obtained as a pale yellow powder (149 mg, 85%, $R_f = 0.10$ [Hexane:EtOAc 7:3]).

δ_{H} (600 MHz, Acetone- d_6): 7.70 (d, 2H, H_2 , $J = 9.2$ Hz), 7.12 (d, 2H, H_6 , $J = 9.0$ Hz), 6.99 (d, 2H, H_3 , $J = 9.2$ Hz), 6.62 (d, 2H, H_7 , $J = 9.0$ Hz), 3.54 (t, 4H, 2 x CH_2), 3.29 (t, 4H, 2 x CH_2).

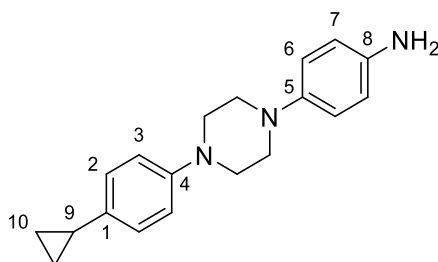
δ_{C} (150 MHz, Acetone- d_6): 153.05 (C_4), 147.35 (C_8), 144.91 (C_5), 127.10 (C_2), 120.33 (C_7), 118.98 (C_1), 117.28 (C_3), 113.68 (C_6), 49.08 (CH_2), 46.85 (CH_2).

δ_{F} (376 MHz, Acetone- d_6): 87.95 (quintet, $J = 147.14$ Hz), 64.09 (d, $J = 148.05$ Hz).

HRMS-ESI (+): m/z calculated for $\text{C}_{16}\text{H}_{18}\text{F}_5\text{N}_3\text{S}$: 379.1142. Found $[\text{M}+\text{H}] = 380.1218$.

ν_{max} (ATR)/ cm^{-1} : 3453, 3373 (br, N-H), 2837 (C-H, aliphatic), 1593, 1512 (Ar-C=C, N-H bend) 1328-1162 (C-N, Ar-C-N, S-F), 1104-1091 (S-F).

4-(4-(4-Cyclopropylphenyl)piperazin-1-yl)aniline (**57**)



Following method E, from **51** (68.86 mg, 0.21 mmol). **55** was obtained as a pale yellow powder (53.3 mg, 85%, R_f = 0.10 [Hexane:EtOAc 7:3]).

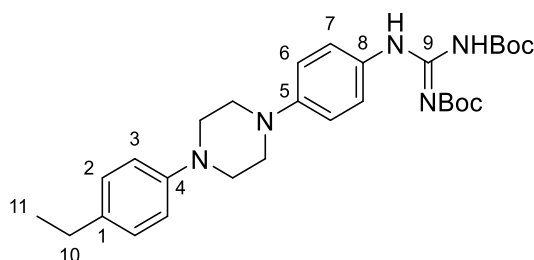
δ_H (600 MHz, Acetone- d_6): 7.12 (d, 2H, H_6 = , J = 9.0 Hz), 6.99 (d, 2H, H_3 , J = 9.2 Hz), 6.98 (d, 2H, H_2 , J = 9.2 Hz), 6.61 (d, 2H, H_7 , J = 9.0 Hz), 3.54 (t, 4H, 2 x CH_2), 3.29 (t, 4H, 2 x CH_2).

δ_C (150 MHz, Acetone- d_6): 149.6 (C_4), 147.56 (C_5), 128.93 (C_1), 126.34 (C_2), 120.23 (C_7), 119.29 (C_8), 116.96 (C_3), 116.28 (C_6), 50.00 (CH_2), 49.68 (CH_2), 14.21 (C_9), 7.87 (C_{10}).

HRMS-ESI (+): m/z calculated for $C_{19}H_{23}N_3$: 293.1892. Found $[M+H] = 294.1969$.

ν_{max} (ATR)/ cm^{-1} : 2965-2836 (C-H aliphatic, cyclopropyl C-H), 1590, 1516 (Ar-C=C, N-H), 1392-1090 (C-N, C-H, Ar-C-N), 1051-995 (C-N, cyclopropyl ring).

Boc-protected 1-(4-(4-(4-Ethylphenyl)piperazin-1-yl)phenyl)guanidine (**60**)



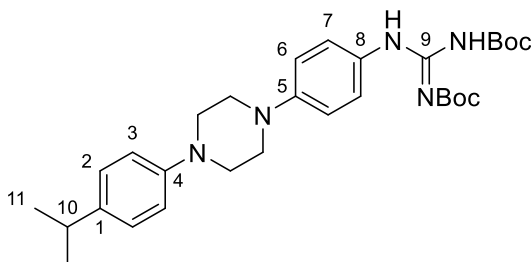
Following method F from **54** (120 mg, 0.43 mmol). **60** was obtained as a white powder (45 mg, 20%, R_f = 0.8 [Hexane:EtOAc 7:3]).

δ_H (400 MHz, Acetone- d_6): 11.75 (NH), 10.09 (NH), 7.53 (d, 2H, H_2 , J = 9.2 Hz), 7.10 (d, 2H, H_6 , J = 8.4 Hz), 7.01 (d, 2H, H_3 , J = 9.2 Hz), 6.94 (d, 2H, H_7 , J = 8.4 Hz), 3.30 (m, br,

8H, 4 x CH₂), 2.55 (d, 3H, H₁₁, J = 7.7 Hz), 1.56 (s, 9H, Boc-(CH₃)₃), 1.44 (s, 9H, Boc-(CH₃)₃), 1.17 (d, 6H, H₁₁, J = 7.7 Hz).

HRMS-ESI (+): m/z calculated for C₂₉H₄₁N₅O₄: 523.3159. Found [M+H] = 538.3385.

Boc-protected 1-(4-(4-(4-Isopropylphenyl)piperazin-1-yl)phenyl)guanidine (**61**)



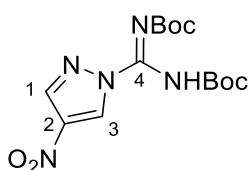
Following method F from **55** (80 mg, 0.27 mmol). **59** was obtained as a white powder (33.3 mg, 23%, R_f = 0.8 [Hexane:EtOAc 7:3]).

δ_H (400 MHz, Acetone-*d*₆): 11.77 (NH), 10.11 (NH), 7.55 (d, 2H, H₂, J = 9.0 Hz), 7.15 (d, 2H, H₆, J = 8.7 Hz), 7.03 (d, 2H, H₃, J = 9.0 Hz), 6.97 (d, 2H, H₇, J = 8.7 Hz), 3.32 (m, br, 8H, 4 x CH₂), H₁₀ obscured by deuterated solvent signal at 2.05 ppm, 1.59 (s, 9H, Boc-(CH₃)₃), 1.46 (s, 9H, Boc-(CH₃)₃), 1.22 (d, 6H, H₁₁, J = 6.8 Hz).

δ_C (100 MHz, Acetone-*d*₆): 163.72 (C₉), 153.77 (Boc-C=O), 153.31 (Boc-C=O), 149.90 (C₄), 149.07 (C₅), 139.95 (C₁), 129.08 (C₈), 126.76 (C₆), 123.44 (C₂), 116.36 (C₇), 116.04 (C₃), 83.42 (Boc-C(CH₃)₃), 78.42 (Boc-C(CH₃)₃), 49.53 (CH₂), 49.27 (CH₂), 33.04 (C₁₀), 27.59 (Boc-C(CH₃)₃), 27.27 (Boc-C(CH₃)₃), 23.56 (C₁₀).

HRMS-ESI (+): m/z calculated for C₃₀H₄₃N₅O₄: 537.3315. Found [M+H] = 538.3385.

tert-Butyl (E)-(((*tert*-butoxycarbonyl)amino)(4-nitro-1H-pyrazol-1-yl)methylene)carbamate (**65**)



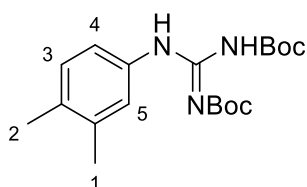
To a solution of N,N-bis-Boc-S-methylisothiurea (7.6 g, 26.4 mg, 1.2 eq.) in anhydrous DMF (88 mL, 0.3 M) was added HgCl₂ (6.57 g, 24.2 mmol, 1.1 eq.), Et₃N (10.1 mL, 72.6 mmol, 3.3 eq.) and nitropyrazole (2.48 g, 22 mmol, 1 eq.). The reaction was allowed to proceed at room temperature overnight. The reaction mixture was filtered through celite and concentrated *in vacuo* followed by dilution with EtOAc (100 mL), washed with water (2 x 100 mL), brine (1 x 100 mL) and dried over Na₂S₂O₄, concentrated *in vacuo*. **65** was afforded as a white powder (5.08 g, 65%, R_f = 0.7 (Hexane:EtOAc 8:2)) post purification by silica gel column chromatography. The structural data are consistent with previously reported characterisation for **65**.²²⁷

δ_H (600 MHz, Acetone-*d*₆): 9.72 (s, br, NH), 9.17 (s, 1H, H₃), 8.43 (s, 1H, H₁), 1.53 (s, 9H, Boc-(CH₃)₃).

δ_C (150 MHz, Acetone-*d*₆): 170.00 (C₄), 149.95 (Boc-C=O), 138.88 (C₂), 137.32 (C₃), 128.35 (C₁) 83.76 (Boc-C(CH₃)₃), 80.11 (Boc-C(CH₃)₃), 27.32 (Boc-C(CH₃)₃)

HRMS-APCI: m/z calculated for C₁₄H₂₁N₅O₆: 355.1492. Found [M-H] = 354.1418.

Boc-protected 1-(3,4-Dimethylphenyl)guanidine (**67**)

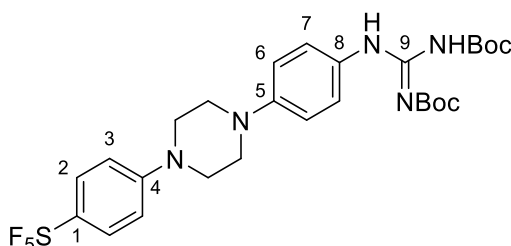


Following method F from 3,4-dimethyl-aniline (60.5 mg, 0.5 mmol). **67** was obtained as a pale yellow powder (149 mg, 84%, R_f = 0.8 [Hexane:EtOAc 7:3]). The structural data are consistent with previously reported characterisation for **67**.²²⁸

δ_H (400 MHz, Acetone-*d*₆): 11.74 (NH), 10.13 (NH), 7.50 (d, 1H, H₃, J = 7.9 Hz), 7.30 (s, 1H, H₅), 7.10 (d, 1H, H₄, J = 7.9 Hz), 2.24-2.21 (d, 2H, H₁ and H₂).

HRMS-ESI (+): m/z calculated for C₁₉H₂₉N₃O₄: 363.2158. Found [M+H] = 364.2235.

Boc-protected 1-(4-(4-(4-(Pentafluoro- λ 6-sulfaneyl)phenyl)piperazin-1-yl)phenyl)guanidine (**68**)



Following method F from **56** (149 mg, 0.39 mmol). **68** was obtained as a pale yellow powder (157.5 mg, 65%, R_f = 0.8 [Hexane:EtOAc 7:3]).

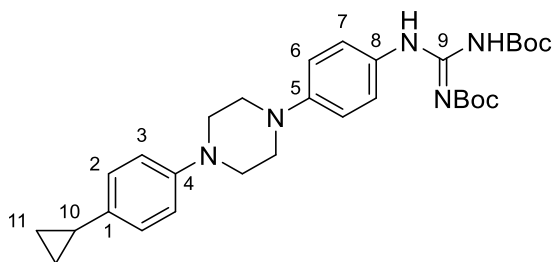
δ_H (400 MHz, Acetone- d_6): 11.78 (NH), 10.13 (NH), 7.71 (d, 2H, H_2 , J = 9.3 Hz), 7.56 (d, 2H, H_6 , J = 8.9 Hz), 7.13 (d, 2H, H_3 , J = 9.3 Hz), 7.04 (d, 2H, H_7 , J = 8.9 Hz), 3.56 (t, 4H, 2 x CH_2), 3.38 (t, 4H, 2 x CH_2), 1.59 (s, 9H, Boc-(CH_3)₃), 1.46 (s, 9H, Boc-(CH_3)₃).

δ_C (100 MHz, Acetone- d_6): 163.74 (C_9), 153.66 (Boc- $C=O$), 153.68 (C_4), 152.86 (Boc- $C=O$), 148.65 (C_5), 129.38 (C_8), 127.00 (C_1), 127.12 (C_2), 123.56 (C_6), 116.19 (C_7), 113.55 (C_3), 83.44 (Boc- $C(CH_3)_3$), 78.47 (Boc- $C(CH_3)_3$), 49.91 (CH_2), 49.34 (CH_2), 27.59 (Boc- $C(CH_3)_3$), 27.30 (Boc- $C(CH_3)_3$).

HRMS-ESI (+): m/z calculated for $C_{17}H_{20}F_5N_5S$: 421.1360. Found $[M+H]^+$ = 622.2476.

ν_{max} (ATR)/ cm^{-1} : 3282 (br, N-H, guanidine), 2891 (C-H, aliphatic), 1718 (C=O), 1554 (C=N, guanidine), 1149-1101 (S-F).

Boc-protected 1-(4-(4-(4-Cyclopropylphenyl)piperazin-1-yl)phenyl)guanidine (**69**)



Following method F from **57** (53 mg, 0.18 mmol). **69** was obtained as a white powder (59.78 mg, 65%, R_f = 0.8 [Hexane:EtOAc 7:3]).

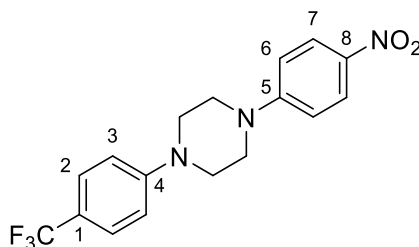
δ_H (400 MHz, Acetone- d_6): 11.78 (NH), 10.11 (NH), 7.55 (d, 2H, H_2 , J = 9.0 Hz), 7.03 (d, 2H, H_3 , J = 9.0 Hz), 7.01 (d, 2H, H_6 , J = 8.8 Hz), 6.94 (d, 2H, H_7 , J = 8.8 Hz), 3.33 (t, 4H, 2 x CH_2), 3.29 (t, 4H, 2 x CH_2), 1.85 (m, 1H, H_{10}), 1.58 (s, 9H, Boc- $(CH_3)_3$), 1.46 (s, 9H, Boc- $(CH_3)_3$), 0.87 (m, 2H, H_{11A}), 0.60 (m, 2H, H_{11B}).

δ_C (100 MHz, Acetone- d_6): 163.75 (C_9), 153.76 (Boc- $C=O$), 153.25 (Boc- $C=O$), 149.68 (C_4), 149.00 (C_5), 134.91 (C_1), 129.22 (C_8), 126.34 (C_6), 123.61 (C_2), 116.49 (C_7), 116.06 (C_3), 83.46 (Boc- $C(CH_3)_3$), 78.45 (Boc- $C(CH_3)_3$), 49.58 (CH_2), 49.24 (CH_2), 27.59 (Boc- $C(CH_3)_3$), 27.27 (Boc- $C(CH_3)_3$), 14.26 (C_{10}), 7.98 (C_{11}).

HRMS-ESI (+): m/z calculated for $C_{30}H_{41}N_5O_4$: 535.3159. Found $[M+H] = 536.3195$.

ν_{max} (ATR)/ cm^{-1} : 3293 (br, N-H, guanidine), 2980-2828 (C-H aliphatic, cyclopropyl C-H), 1717 (C=O), 1560 (C=N, guanidine), 1059-1029 (C-N, cyclopropyl ring).

1-(4-Nitrophenyl)-4-(4-(trifluoromethyl)phenyl)piperazine (**70**)



Following method D, from 1-(4-(trifluoromethyl)phenyl)piperazine (690.3 mg, 3 mmol). **70** was obtained as a yellow powder (737.79 mg, 70%, R_f = 0.75 [Hexane:EtOAc 7:3]).

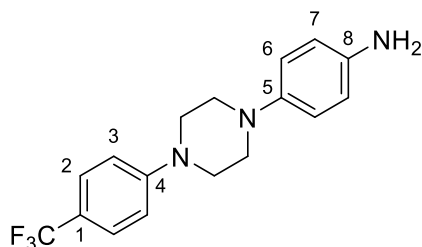
δ_{H} (600 MHz, Acetone- d_6): 8.09 (d, 2H, H_7 , $J = 9.7$ Hz), 7.54 (d, 2H, H_3 , $J = 8.8$ Hz), 7.09 (d, 2H, H_2 , $J = 8.8$), 7.07 (d, 2H, H_6 , $J = 9.7$ Hz), 3.67 (t, 4H, 2 x CH_2), 3.49 (t, 4H, 2 x CH_2).

δ_{C} (150 MHz, Acetone- d_6): 154.83 (C_8), 153.10 (C_4), 137.35 (C_5), 126.72 (C_1), 126.70 (C_3), 127.23 (C_7), 114.33 (C_6), 112.91 (C_3), 46.54 ($\underline{\text{CH}_2}$), 46.08 ($\underline{\text{CH}_2}$).

δ_{F} (376 MHz, Acetone- d_6): 59.38 (s).

HRMS-ESI (+): m/z calculated for $\text{C}_{17}\text{H}_{16}\text{F}_3\text{N}_3\text{O}_2$: 351.1195. Not found.

4-(4-(4-(Trifluoromethyl)phenyl)piperazin-1-yl)aniline (**71**)



Following method E, from **70** (700 mg, 1.99 mmol). **71** was obtained as a white yellow powder (338.93 mg, 53%, $R_f = 0.10$ [Hexane:EtOAc 7:3]).

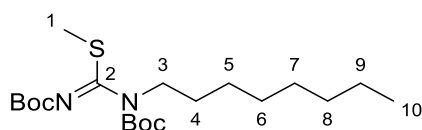
δ_{H} (600 MHz, Acetone- d_6): 7.52 (d, 2H, H_7 , $J = 9.3$ Hz), 7.12 (d, 2H, H_6 , $J = 9.3$ Hz), 6.76 (d, 2H, H_2 , $J = 9.1$), 6.52 (d, 2H, H_3 , $J = 9.1$ Hz), 3.67 (t, 4H, 2 x CH_2), 3.49 (t, 4H, 2 x CH_2).

δ_{C} (150 MHz, Acetone- d_6): 153.76 (C_8), 143.09 (C_5), 142.55 (C_4), 126.65 (C_7), 118.90 (C_2), 115.21 (C_3), 114.77 (C_8), 50.86 ($\underline{\text{CH}_2}$), 47.74 ($\underline{\text{CH}_2}$).

δ_{F} (376 MHz, Acetone- d_6): 59.47 (s).

HRMS-ESI (+): m/z calculated for $\text{C}_{17}\text{H}_{16}\text{F}_3\text{N}_3$: 321.1453. Found $[\text{M}+\text{H}] = 322.1527$.

Boc-protected Methyl octylcarbamimidothioate (**72**)



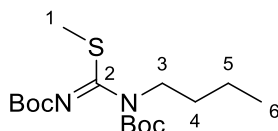
Following method H, from n-octanol (0.31 mL, 1.97 mmol). **72** was obtained as a colourless oil (711.5 mg, 88%, $R_f = 0.9$ [Hexane:EtOAc 9:1]). The structural data are consistent with previously reported characterisation for **72**.²²⁹

δ_H (400 MHz, Acetone- d_6): 3.49 (t, 2H, H_3 , $J_{3-4} = 7.8$ Hz), 2.40 (s, 3H, H_1), 1.50 (s, 9H, Boc-(CH₃)₃), 1.49 (s, 9H, Boc-(CH₃)₃), 0.90 (t, 3H, H_{10} , $J_{9-10} = 6.8$ Hz).

δ_C (100 MHz, Acetone- d_6): 162.17 (Boc-C=O), 157.49 (Boc-C=O), 151.51 (C₂), 81.33 (Boc-C(CH₃)₃), 80.94 (Boc-C(CH₃)₃), 48.61 (C₃), 31.65 (C₄), 27.32 (Boc-C(CH₃)₃), 27.29 (Boc-C(CH₃)₃), 26.54-14.26 (C₅-C₉), 13.48 (C₁₁).

HRMS-ESI (+): m/z calculated for C₂₀H₃₈N₂O₄S: 402.2552. Found [M+H] = 403.2629.

Boc-protected Methyl butylcarbamimidothioate (**74**)



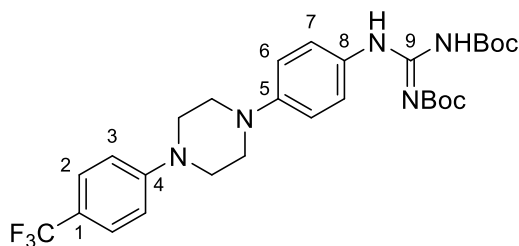
Following method H, from n-butanol (0.18 mL, 2 mmol). **72** was obtained as a colourless oil (391 mg, 56%, $R_f = 0.9$ [Hexane:EtOAc 9:1]).

δ_H (400 MHz, Acetone- d_6): 3.49 (t, 2H, H_3 , $J_{3-4} = 7.7$ Hz), 2.40 (s, 3H, H_1), 1.65 (m, 2H, H_4 , $J_4 = 7.7$ Hz), 1.50 (s, 9H, Boc-(CH₃)₃), 1.49 (s, 9H, Boc-(CH₃)₃), 1.35 (m, 2H, H_5 , $J_5 = 7.7$ Hz), 0.94 (t, 3H, H_6 , $J = 7.7$ Hz)

δ_C (100 MHz, Acetone- d_6): 162.17 (Boc-C=O), 157.49 (Boc-C=O), 151.51 (C₂), 81.33 (Boc-C(CH₃)₃), 80.94 (Boc-C(CH₃)₃), 48.61 (C₃), 31.65 (C₄), 27.32 (Boc-C(CH₃)₃), 27.29 (Boc-C(CH₃)₃), 26.54-14.26 (C₅-C₉), 13.48 (C₁₁).

HRMS-ESI (+): m/z calculated for C₂₀H₃₈N₂O₄S: 402.2552. Found [M+H] = 403.2629.

Boc-protected 1-(4-(4-(4-(Trifluoromethyl)phenyl)piperazin-1-yl)phenyl)guanidine (**1a**)

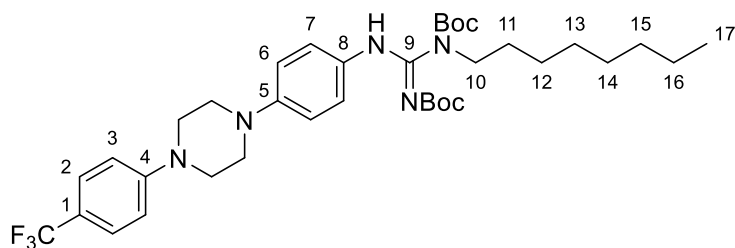


Following method F from **71** (321.34 mg, 1 mmol). **1a** was obtained as a white powder (78.85 mg, 14%, R_f = 0.8 [Hexane:EtOAc 7:3]).

δ_H (400 MHz, Acetone- d_6): 11.78 (NH), 10.11 (NH), 7.63 (d, 2H, H_2 , J = 8.8 Hz), 7.52 (d, 2H, H_1 , J = 8.9 Hz), 7.65 (d, 2H, H_3 , J = 8.8 Hz), 7.02 (d, 2H, H_7 , J = 8.9 Hz), 3.54 (t, 4H, 2 $\times$ CH₂), 3.36 (t, 4H, 2 $\times$ CH₂), 1.58 (s, 9H, Boc-(CH₃)₃), 1.46 (s, 9H, Boc-(CH₃)₃).

δ_C (100 MHz, Acetone- d_6): 163.72 (C₉), 153.65 (Boc-C=O), 153.20 (C₄), 152.84 (Boc-C=O), 148.45 (C₅), 129.15 (C₈), 126.85 (C₂), 123.28 (C₆), 119.23 (C₁), 116.12 (C₇), 83.42 (Boc-C(CH₃)₃), 78.44 (Boc-C(CH₃)₃), 49.85 (CH₂), 49.28 (CH₂), 27.57 (Boc-C(CH₃)₃), 27.28 (Boc-C(CH₃)₃).

Boc-protected 1-Octyl-3-(4-(4-(4-(trifluoromethyl)phenyl)piperazin-1-yl)phenyl)guanidine (**75**)



Following method H, from **1a** (150 mg, 0.27 mmol). **75** was obtained as a colourless oil (37.5 mg, 42%, R_f = 0.55 [Hexane:EtOAc 8:2]).

δ_H (600 MHz, Acetone- d_6): 11.76 and 10.17 (NH, conformational tautomers), 7.58 (d, 2H, H_2 , J = 9.2 Hz), 7.55 (d, 2H, H_6 , J = 8.5 Hz), 7.16 (d, 2H, H_7 , J = 8.5), 7.03 (d, 2H, H_3 , J = 9.2 Hz), 4.06 (t, 2H, H_{10} , J = 6.8 Hz), 3.51 (t, 4H, 2 $\times$ CH₂), 3.36 (t, 4H, 2 $\times$ CH₂), 1.58 (s, 18H, 2 $\times$ Boc-(C(CH₃)₃), 1.32 (m, 12H, H_{11} - H_{16}), 0.89 (t, 3H, H_{17} , J_{17-16} = 7.5 Hz).

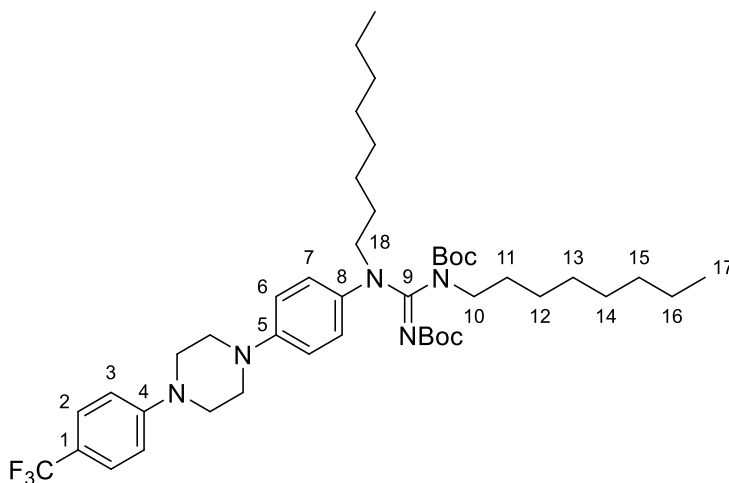
δ_c (150 MHz, Acetone- d_6): 164.07 (C₉), 153.92 (Boc-C=O), 153.68 (Boc-C=O), 148.71 (C₅), 129.19 (C₈), 126.47 (C₆), 123.70 (C₂), 119.21 (C₁), 116.22 (C₃), 114.76 (C₇), 83.51 (Boc-C(CH₃)₃), 48.91(CH₂), 47.64(CH₂), 31.64 (C₁₀), 27.3, 22.49 (C₁₁-C₁₆), 13.48 (C₁₇)

δ_f (376 MHz, Acetone- d_6): -61.6 (s).

HRMS-ESI (+): m/z calculated for C₃₆H₅₂F₃N₅O₄: 675.3971. Found [M+H] = 676.4035 and [M+Na] = 698.3836.

$\nu_{\max}$ (ATR)/cm⁻¹: ~3400 (br, N-H), 2927, 2855 (s, C-H, aliphatic octyl chain), 1719 (C=O), 1561 (C=N, guanidine), 1370 (*tert*-butyl), 1154-1107 (C-F).

Boc-protected 1,3-Dioctyl-1-(4-(4-(4-(trifluoromethyl)phenyl)piperazin-1-yl)phenyl)guanidine (**77**)



Following method H, from **1a** (200 mg, 0.35 mmol). **75** was obtained as a colourless oil (24.8 mg, 9%, R_f = 0.6 [Hexane:EtOAc 8:2]).

δ_H (400 MHz, Acetone- d_6): 7.55 (d, 2H, H₂, J = 8.9 Hz), 7.16 (d, 2H, H₃, J = 8.9 Hz), 7.02 (d, 2H, H₆, J = 8.8), 6.85 (d, 2H, H₇, J = 8.8 Hz), 3.76 (t, 2H, H₁₀, J = 8.2 Hz), 3.51 (t, 4H, 2 x CH₂), 3.36 (t, 4H, 2 x CH₂), 3.12 (t, 2H, H₁₈, J₁₈₋₁₉ = 7.7 Hz), 1.52 (s, 9H, Boc-(C(CH₃)₃),

1.38 (s, 9H, Boc-(C(CH₃)₃), 1.32 (m, 12H, H₁₁-H₁₆), 1.31 (m, H₁₀-H₁₆), 0.88 (m, 6H, H₁₇ and H_{17'}).

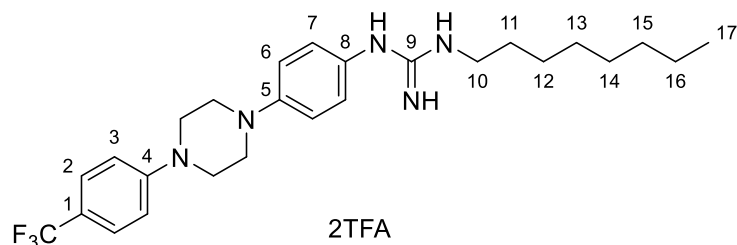
δ_c (100 MHz, Acetone-*d*₆): 153.70 (C₄), 152.66 (Boc-C=O), 152.51 (Boc-C=O), 148.07 (C₅), 140.56 (C₈), 126.22 (C₂), 123.83 (C₁), 121.76 (C₆), 116.97 (C₇), 114.59 (C₃), 81.05 (Boc-C(CH₃)₃), 80.35 (Boc-C(CH₃)₃), 49.58 (CH₂), 47.64 (CH₂), 31.67 (C₁₀), 31.66 (C₁₈), 27.55-26.84 (C₁₁-C₁₆), 13.51 (C_{17'}), 13.49 (C₁₇).

δ_F (376 MHz, Acetone-*d*₆): -61.59 (s).

HRMS-ESI (+): m/z calculated for C₄₄H₆₈F₃N₅O₄: 787.5223. Found [M+H]⁺ = 788.5289 and [M+Na]⁺ = 810.5114.

ν_{max} (ATR)/cm⁻¹: 2926, 2855 (s, C-H, aliphatic octyl chains), 1716 (C=O), 1615 (C=N, guanidine), 1366 (*tert*-butyl), 1200-1112 (C-F).

TFA salt of 1-Octyl-3-(4-(4-(4-(trifluoromethyl)phenyl)piperazin-1-yl)phenyl)guanidine (**12**)



Following method G from **75** (25 mg, 0.04 mmol). **2** was obtained as a yellow gum (14.29 mg, 75%, R_f = 0.0 [Hexane:EtOAc 7:3]).

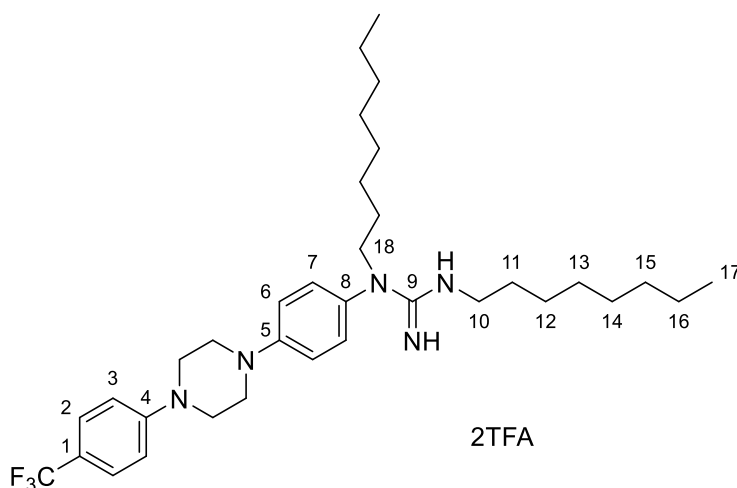
δ_H (600 MHz, DMSO-*d*₆): 11.28 and 10.00 (NH, conformational tautomers), 8.34 (NH), 7.55 (d, 2H, H₂, J = 9.0 Hz), 7.19 (d, 2H, H₆, J = 9.1 Hz), 7.14 (d, 2H, H₃, J = 9.0 Hz), 7.10 (d, 2H, H₇, J = 9.1 Hz), 4.20 (t, 2H, H₁₀, J = 6.4 Hz), 3.39 (m, 8H, 4 x CH₂), 1.29 (m, 12H, H₁₁-H₁₆), 0.87 (t, 3H, H₁₇, J₁₇₋₁₆ = 7.0 Hz).

δ_c (150 MHz, DMSO-*d*₆): 153.22 (C₄), 150.23 (C₅), 129.96 (C₈), 127.37 (C₆), 126.64 (C₂), 118.21 (C₁), 116.51 (C₇), 114.80 (C₃), 47.53 (CH₂), 46.83 (CH₂), 31.19 (C₁₀), 30.71-22.07 (C₁₁-C₁₆), 13.96 (C₁₇)

HRMS-ESI (+): *m/z* calculated for C₂₆H₃₆F₃N₅: 475.2923. Found [M+CO₂+H]⁺ = 520.2905.

$\nu_{\max}$ (ATR)/cm⁻¹: 2918, 2850 (s, C-H, aliphatic octyl chain), 1675 (s, C=O TFA and C=N guanidiniumC=O), 1539 (C=N, guanidine), 1142-1114 (C-F).

TFA salt of 1,3-dioctyl-1-(4-(4-(4-(trifluoromethyl)phenyl)piperazin-1-yl)phenyl)guanidine (**21**)



Following method G from **77** (20 mg, 0.03 mmol). **21** was obtained as a yellow gum (15 mg, 85%, *R_f* = 0.0 [Hexane:EtOAc 7:3]).

δ_H (600 MHz, Acetone-*d*₆): 9.07 (NH), 9.06 (NH), 7.55 (d, 2H, H₂, *J* = 9.2 Hz), 7.14 (d, 2H, H₃, *J* = 9.2 Hz), 7.08 (d, 2H, H₇, *J* = 9.2 Hz), 7.06 (d, 2H, H₆, *J* = 9.2 Hz), 3.45 (t, 4H, 2 x CH₂), 3.34 (t, 4H, 2 x CH₂), 3.20 (br, 4H, H₁₀ & H₁₈), 1.27 (m, 12H, H₁₁-H₁₆, H_{11'}-H_{16'}), 0.87 (t, 6H, H₁₇, *J*₁₇ = 7.0 Hz).

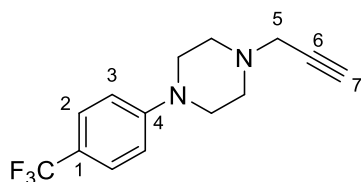
δ_c (150 MHz, Acetone-*d*₆): 154.30 (C₉), 153.60 (C₄), 149.94 (C₅), 127.41 (C₆), 126.73 (C₂), 126.68 (C₈), 118.5 (C₁), 116.48 (C₇), 114.87 (C₃), 49.21 (CH₂), 47.38 (CH₂), 48.17 & 47.19 (C₁₀ & C₁₈), 31.68-22.57 (C₁₁-C₁₆), 14.43 (C₁₇).

δ_F (376 MHz, Acetone-*d*₆): -59.49 (s), -74.53 (TFA).

HRMS-ESI (+): m/z calculated for C₄₄H₆₈F₃N₅O₄: 787.5223. Found [M+H] = 788.5289 and [M+Na] = 810.5114.

v_{max}(ATR)/cm⁻¹: 3283 (br, N-H, guanidinium), 2931, 2858 (s, C-H, aliphatic octyl chains), 1616 (C=O TFA, C=N guanidinium), 1142-1116 (C-F).

1-(Prop-2-yn-1-yl)-4-(4-(trifluoromethyl)phenyl)piperazine (**79**)



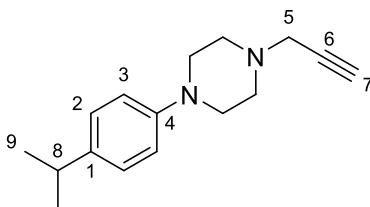
Following method I, from **77** (460 mg, 2 mmol). **79** was obtained as a pale yellow powder (347.8 mg, 65%, R_f = 0.3 [Hexane:EtOAc 7:3]).

δ_H (600 MHz, DMSO-*d*₆): 7.50 (d, 2H, H₂, J = 8.6 Hz), 7.07 (d, 2H, H₃, J = 8.6 Hz), 3.34 (d, 1H, H₇, J₇₋₅ = 2.3 Hz), 4.39 (t, 4H, 2 x CH₂), 3.19 (t, 1H, H₇, J₇₋₅ = 2.3 Hz), 2.58 (t, 4H, 2 x CH₂).

δ_C (150 MHz, DMSO-*d*₆): 153.68 (C₄), 126.59 (C₂), 126.36 (C₁), 114.73 (C₃), 79.58 (C₆), 76.39 (C₇), 51.26 (CH₂), 47.33 (CH₂), 46.43 (C₅).

HRMS-ESI (+): m/z calculated for C₁₄H₁₅F₃N₂: 268.1187. Found [M+H]⁺ = 269.1264.

1-(4-Isopropylphenyl)-4-(prop-2-yn-1-yl)piperazine (**80**)



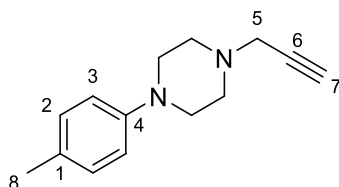
Following method J, from **45** (1.64 mg, 10 mmol). Crude **80** was obtained as a brown oil (538 mg).

δ_{H} (600 MHz, DMSO- d_6): 7.07 (d, 2H, H₂, J = 8.6 Hz), 6.86 (d, 2H, H₃, J = 8.6 Hz), 3.32 (d, 1H, H₇, J₇₋₅ = 2.3 Hz), 3.32 (s, 2H, H₅), 3.18 (t, 1H, H₇, J₇₋₅ = 2.3 Hz), 3.09 (t, 4H, 2 x CH₂), 2.79 (sept, 1H, H₈, J = 6.9 Hz), 2.59 (t, 4H, 2 x CH₂), 1.16 (d, H₆, H₉, J = 6.9 Hz).

δ_{C} (150 MHz, DMSO- d_6): 149.27 (C₄), 139.04 (C₁), 126.72 (C₂), 115.74 (C₃), 79.37 (C₆), 75.96 (C₇), 51.33 (CH₂), 48.59 (CH₂), 46.13 (C₅), 32.58 (C₈), 24.15 (C₉).

HRMS-ESI (+): m/z calculated for C₁₆H₂₂N₂: 242.1783. Found [M+H]⁺ = 243.1862.

1-(Prop-2-yn-1-yl)-4-(p-tolyl)piperazine (**81**)



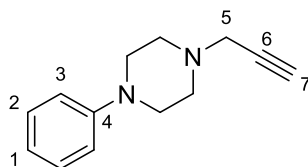
Following method I, from **78** (460 mg, 2 mmol). **79** was obtained as a pale yellow powder (347.8 mg, 65%, R_f = 0.3 [Hexane:EtOAc 7:3]).

δ_{H} (600 MHz, DMSO- d_6): 7.02 (d, 2H, H₂, J = 8.4 Hz), 6.83 (d, 2H, H₃, J = 8.4 Hz), 3.32 (H₇, overlap with H₂O peak), 3.30 (s, 2H, H₅), 3.18 (t, 1H, H₇, J₇₋₅ = 2.9 Hz), 3.08 (t, 4H, 2 x CH₂), 2.58 (t, 4H, 2 x CH₂), 2.20 (s, 2H, H₈).

δ_{C} (150 MHz, DMSO- d_6): 149.39 (C₄), 129.81 (C₂), 128.10 (C₁), 116.20 (C₃), 79.76 (C₆), 76.27 (C₇), 51.60 (CH₂), 49.04 (CH₂), 20.51 (C₈).

HRMS-ESI (+): m/z calculated for C₁₄H₁₈N₂: 214.1470. Found [M+H]⁺ = 215.1548.

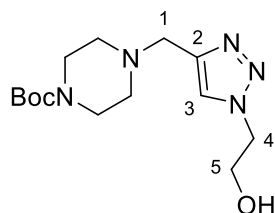
1-Phenyl-4-(prop-2-yn-1-yl)piperazine (**82**)



Following method J, from **45** (1.64 mg, 10 mmol). Crude **82** was obtained as a brown oil.

Crude δ_{H} (400 MHz, DMSO- d_6): 7.18 (t, 2H, H₂, J = J₂₋₁ = 8.3 Hz) 6.90 (d, 2H, H₁, J = 8.3 Hz), 6.74 (t, 2H, H₃, J = 8.2 Hz), 3.30 (H₇, overlap with H₂O peak), 3.30 (s, 2H, H₅), 3.15 (br t, 1H, H₇), 3.11 (t, 4H, 2 x CH₂), 2.56 (t, 4H, 2 x CH₂), 2.20 (s, 2H, H₈).

Tert-Butyl 4-((1-(2-hydroxyethyl)-1H-1,2,3-triazol-4-yl)methyl)piperazine-1-carboxylate (**86**)



Following method Q, from **84** (112.5 mg, 0.5 mmol). **86** was afforded a white powder (138.57 mg, 89%, R_f = 0.25 [Hexane:EtOAc 7:3]).

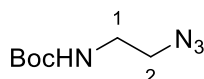
δ_{H} (400 MHz, Acetone- d_6): 7.90 (s, 1H, H₃), 4.49 (t, 2H, H₄, J₄₋₅ = 5.3 Hz), 3.96 (t, 2H, H₅, J₅₋₄ = 5.3 Hz), 3.72 (s, 2H, H₁), 3.42 (t, 4H, 2 x CH₂), 2.51 (t, 4H, 2 x CH₂), 1.44 (s, s, 9H, Boc-(C(CH₃)₃)).

δ_{C} (100 MHz, Acetone- d_6): 154.25 (Boc-C=O), 124.4 (C₁), 123.26 (C₂), 78.96 (Boc-C(CH₃)₃), 60.63 (C₅), 52.86 (C₁), 52.44 (CH₂), 52.27 (CH₂), 27.63 (Boc-C(CH₃)₃).

HRMS-ESI (+): m/z calculated for C₁₄H₂₅N₅O₃: 311.1957. Not found.

HRMS-APCI: m/z calculated for C₇H₁₄N₄O₂: 311.1957 Not found.

tert-Butyl (2-azidoethyl)carbamate (**87**)



To a solution of amine **88** (801 mg, 5 mmol, 1 eq.) and ImSO₂.H₂SO₄.N₃ (1.62 g, 6 mmol, 1 eq.) in anhydrous MeOH (12.5 mL, 0.2 M) was added an aqueous solution of K₂CO₃ (26 mL, 20 mmol, 0.77 M) and an aqueous solution of CuSO₄.5H₂O (0.5 mL, 0.05 mmol, 0.1 M). The reaction was allowed to proceed at room temperature

overnight. The reaction mixture was concentrated *in vacuo*, diluted with water and extracted with EtOAc (4 x 75 mL). The combined organic phases were washed with brine (1 x 100 mL) and dried over Na₂S₂O₄. Concentration *in vacuo* afforded **87** as an orange crystalline substance (1.32 g) and used without further purification. The structural data are consistent with previously reported characterisation for **72**.²³⁰

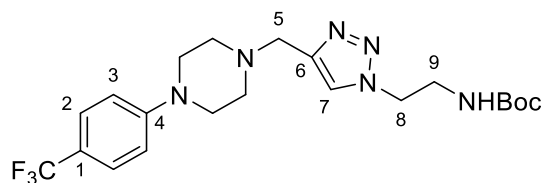
δ_H (600 MHz, DMSO-*d*₆): 7.04 (NH), 3.29 (t, 2H, H₂, J = 5.8 Hz), 3.12 (q, 2H, H₁, J = 5.8 Hz), 1.39 (s, 9H, Boc-(C(CH₃)₃)).

δ_C (150 MHz, DMSO-*d*₆): 155.55 (Boc-C=O), 77.84 (Boc-C(CH₃)₃), 49.99 (C₂), 40.03 (C₁), 28.18 (Boc-C(CH₃)₃).

HRMS-ESI (+): m/z calculated for C₇H₁₄N₄O₂: 186.1117. Not found.

HRMS-APCI: m/z calculated for C₇H₁₄N₄O₂: 186.1117 Not found.

tert-Butyl (2-(4-((4-(4-(trifluoromethyl)phenyl)piperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)carbamate (**89**)



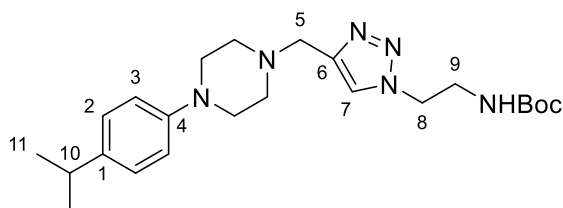
Following method Q, from **79** (134.1 mg, 0.5 mmol). **89** was afforded a brown powder (119.1 mg, 52%, R_f = 0.45 [DCM:MeOH 9:1]).

δ_H (600 MHz, Acetone-*d*₆): 7.85 (s, 1H, H₇), 7.51 (d, 2H, H₂, J = 8.8 Hz), 8.8 (d, 2H, H₃, J = 8.8 Hz), 6.20 (NH), 4.51 (t, 2H, H₈, J = 6.0 Hz), 3.68 (s, H₅), 3.57 (q, 2H, H₉, J = J_{9-NH} = 6.0 Hz), 3.32 (t, 4H, 2 x CH₂), 2.64 (t, 4H, 2 x CH₂), 1.40 (s, s, 9H, Boc-(C(CH₃)₃)).

δ_C (150 MHz, Acetone-*d*₆): 155.74 (Boc-C=O), 153.77 (C₄), 143.60 (C₆), 127.91 (C₁), 123.5 (C₇), 126.16 (C₂), 114.30 (C₃), 78.18 (Boc-C(CH₃)₃), 54.32 (CH₂), 53.01 (C₅), 52.37 (CH₂), 49.39 (C₈), 40.51 (C₉), 27.69 (Boc-C(CH₃)₃).

HRMS-ESI (+): m/z calculated for C₂₁H₂₉F₃N₆O₂: 454.4982. Found [M+H] = 455.2387.

tert-Butyl (2-(4-((4-(4-isopropylphenyl)piperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)carbamate (**90**)



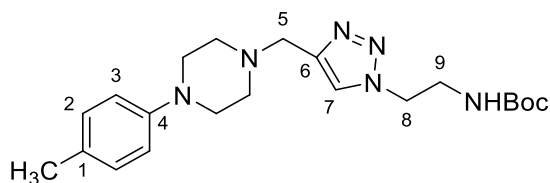
Following method Q, from **80** (538 mg, 2.22 mmol). **90** was afforded a brown oil (533.4 mg, 56%, $R_f = 0.6$ [DCM:MeOH 9:1]).

δ_H (600 MHz, DMSO- d_6): 7.94 (s, 1H, H₇), 7.07 (d, 2H, H₂, $J = 8.6$ Hz), 6.99 (NH), 6.83 (d, 2H, H₃, $J = 8.6$ Hz), 4.37 (t, 2H, H₈, $J = 6.0$ Hz), 3.58 (s, H₅), 3.37 (q, 2H, H₉, $J = J_{9-NH} = 6.0$ Hz), 3.05 (t, 4H, 2 x CH₂), 2.78 (sept, 1H, H₁₀, $J = 7.0$ Hz), 2.54 (t, 4H, 2 x CH₂), 1.35 (s, s, 9H, Boc-(C(CH₃)₃), 1.15 (d, 6H, H₁₀, $J = 7.0$ Hz).

δ_C (150 MHz, DMSO- d_6): 156.03 (Boc-C=O), 149.69 (C₄), 143.59 (C₆), 139.37 (C₁), 130.09 (C₇), 127.04 (C₂), 116.30 (C₃), 78.49 (Boc-C(CH₃)₃), 53.23 (CH₂), 52.71 (C₅), 49.48 (CH₂), 49.12 (C₈), 40.55 (C₉), 32.94 (C₁₀), 27.69 (Boc-C(CH₃)₃), 24.59 (C₁₁).

HRMS-ESI (+): m/z calculated for C₂₃H₃₆N₆O₂: 428.2900. Found [M+H] = 429.2978.

tert-Butyl (2-(4-((4-(*p*-tolyl)piperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)carbamate (**91**)



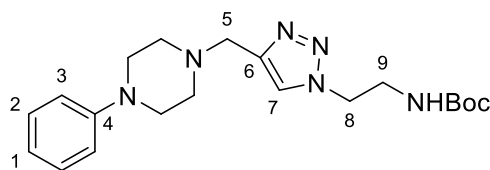
Following method Q, from **81** (107.2 mg, 0.5 mmol). **91** was afforded a white powder (95.7 mg, 48%, R_f = 0.45 [DCM:MeOH 9:1]).

δ_H (600 MHz, Acetone- d_6): 7.82 (s, 1H, H₇), 7.03 (d, 2H, H₂, J = 8.6 Hz), 6.83 (d, 2H, H₃, J = 8.6 Hz), 6.19 (NH), 4.51 (t, 2H, H₈, J = 6.0 Hz), 3.66 (s, H₅), 3.58 (q, 2H, H₉, J = J_{9-NH} = 6.0 Hz), 3.11 (t, 4H, 2 x CH₂), 2.61 (t, 4H, 2 x CH₂), 2.27 (s, 3H, CH₃) 1.40 (s, s, 9H, Boc-(C(CH₃)₃)).

δ_C (150 MHz, Acetone- d_6): 155.40 (Boc-C=O), 149.98 (C₄), 143.95 (C₆), 129.42 (C₂), 128.35 (C₇), 123.39 (C₁), 116.12 (C₃), 53.16 (C₅), 49.38 (CH₂), 40.5 (C₈), 40.51 (C₉), 27.68 (Boc-C(CH₃)₃), 19.51 (CH₃).

HRMS-ESI (+): m/z calculated for C₂₁H₃₂N₆O₂: 400.2587. Found [M+H] = 401.2669.

tert-Butyl (2-(4-((4-phenylpiperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)carbamate (**92**)



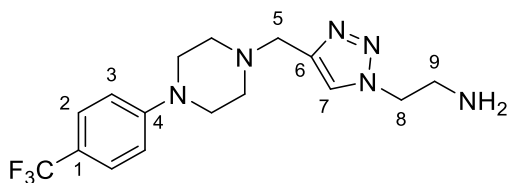
Following method Q, from **83** (107.2 mg, 0.5 mmol). **92** was afforded as a brown oil (319.1 mg, 41%, R_f = 0.45 [DCM:MeOH 9:1]).

δ_H (600 MHz, DMSO- d_6): 7.94 (s, 1H, H₇), 7.20 (t, 2H, H₂, J = J_{2-1} = 7.5 Hz), 7.04 (NH), 6.91 (d, 2H, H₂, J = 7.5 Hz), 6.76 (d, 1H, H₁, J = 7.5 Hz), 4.37 (t, 2H, H₈, J = 5.7 Hz), 3.59 (s, H₅), 3.30 (q, 2H, H₉, J = J_{9-NH} = 5.7 Hz, obscured by solvent peak), 3.10 (t, br, 4H, 2 x CH₂), 2.55 (t, br, 4H, 2 x CH₂), 1.39 (s, s, 9H, Boc-(C(CH₃)₃)).

δ_C (150 MHz, DMSO- d_6): 155.54 (Boc-C=O), 151.13 (C₄), 143.01 (C₆), 128.85 (C₂), 118.85 (C₁), 115.42 (C₃), 77.91 (Boc-C(CH₃)₃), 52.61 (C₅), 52.30 (CH₂), 50.04 (CH₂), 48.96 (C₈), 48.24 (C₉), 32.94 (C₁₀), 27.69 (Boc-C(CH₃)₃), 28.22 (C₁₁).

HRMS-ESI (+): m/z calculated for C₂₃H₃₆N₆O₂: 386.2430. Found [M+H] = 387.2512.

2-(4-((4-(4-(Trifluoromethyl)phenyl)piperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethan-1-amine (**93**)



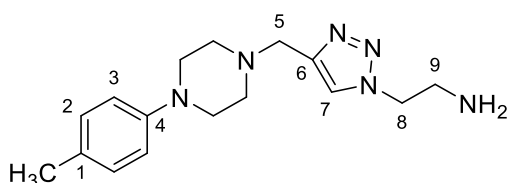
Following method C, from **89** (119.1 mg, 0.262 mmol). **93** was obtained as a brown oil (quant., $R_f = 0.1$ [DCM:EtOAc 9:1]).

δ_H (600 MHz, DMSO- d_6): 8.33 (s, 1H, H₇), 8.04 (br, NH₂), 7.58 (d, 2H, H₂, $J = 8.8$ Hz), 7.15 (d, 2H, H₃, $J = 8.8$ Hz), 4.68 (t, 2H, H₈, $J = 6.1$ Hz), 4.56 (s, H₅), 3.60 (t, 3.38 (q, 2H, H₉, $J = J_{9-NH} = 6.0$ Hz), [4.05, 3.53, 3.21, 3.10 (br, 4 x CH₂)].

δ_C (150 MHz, DMSO- d_6): 153.77 (C₄), 136.80 (C₆), 128.01 (C₆), 126.89 (C₇), 126.79 (C₂), 119.80 (C₁), 115.53 (C₃), 50.55 (CH₂), 49.80 (C₅), 47.72 (CH₂), 44.85 (C₈), 40.54 (C₉).

HRMS-ESI (+): m/z calculated for C₁₆H₂₁F₃N₆: 354.1780. Found [M+H] = 355.1852.

2-(4-((4-(p-Tolyl)piperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethan-1-amine (**95**)



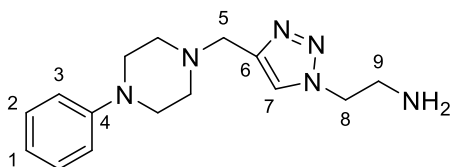
Following method C, from **90** (95.7 mg, 0.239 mmol). **95** was obtained as a brown oil (quant., $R_f = 0.1$ [DCM:EtOAc 9:1]).

δ_H (600 MHz, DMSO- d_6): 8.33 (s, 1H, H₇), 8.04 (br, NH₂), 7.07 (d, 2H, H₂, $J = 8.4$ Hz), 6.89 (d, 2H, H₃, $J = 8.4$ Hz), 4.68 (t, 2H, H₈, $J = 6.0$ Hz), 4.56 (s, H₅), 3.60 (t, 2H, H₈, $J_{8-7} = 6.8$ Hz), 3.38 (q, 2H, H₉, $J = J_{9-NH} = 6.0$ Hz), 2.2 (s, 3H, CH₃), [3.77, 3.52, 3.20, 2.91 (s, br, 4 x CH₂)].

δ_c (150 MHz, DMSO- d_6): 147.83 (C₄), 136.77 (C₆), 130.06 (C₂), 129.59 (C₆), 128.06 (C₇), 119.80 (C₁), 116.77 (C₂), 50.79 (CH₂), 49.71 (C₅), 47.69 (C₈), 46.43 (CH₂), 40.54 (C₉), 20.51 (CH₃).

HRMS-ESI (+): m/z calculated for C₁₆H₂₄N₆: 300.2062. Found [M+H] = 301.2135.

2-(4-((4-Phenylpiperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethan-1-amine (**96**)



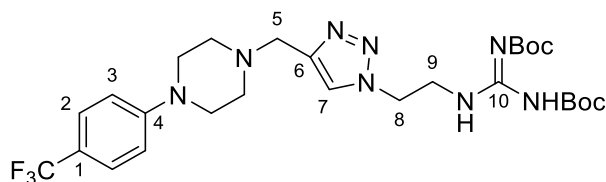
Following method C, from **91** (319.1 mg, 0.82 mmol). **96** was obtained as a brown oil (quant., R_f = 0.1 [DCM:EtOAc 9:1]).

δ_H (600 MHz, DMSO- d_6): 8.33 (s, 1H, H₇), 8.06 (br, NH₂), 7.26 (t, 2H, H₂, J = 8.0 Hz), 6.99 (d, 2H, H₃, J = 8.0 Hz), 6.88 (t, 1H, H₁, J = 8.0 Hz), 4.68 (t, 2H, H₈, J = 6.3 Hz), 4.56 (s, H₅), 3.38 (q, 2H, H₉, J = J_{9-NH} = 6.3 Hz), 2.2 (s, 3H, CH₃) [3.85, 3.53, 3.21, 2.96 (s, br, 4 x CH₂)].

δ_c (150 MHz, DMSO- d_6): 149.95 (C₄), 136.79 (C₆), 129.60 (C₂), 128.04 (C₆), 128.06 (C₇), 120.69 (C₁), 116.49 (C₂), 50.77 (CH₂), 49.00 (C₅), 47.10 (C₈), 45.98 (CH₂), 40.52 (C₉).

HRMS-ESI (+): m/z calculated for C₁₆H₂₄N₆: 286.1906. Found [M+H] = 287.1982.

Boc-protected 1-(2-(4-((4-(4-(Trifluoromethyl)phenyl)piperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)guanidine (**93**)



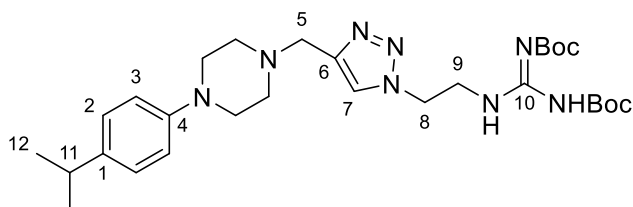
Following method F from **93** (92.8 mg, 0.26 mmol). **93** was obtained as a white powder (102.8 mg, 66%, $R_f = 0.65$ [DCM:MeOH 9:1]).

δ_H (600 MHz, DMSO- d_6): 11.45 (NH), 8.39 (NH), 8.01 (s, 1H, H_7), 7.58 (d, 2H, H_2 , $J = 9.4$ Hz), 7.03 (d, 2H, H_3 , $J = 9.4$ Hz), 4.53 (t, 2H, H_8 , $J = 6.2$ Hz), 3.77 (q, 2H, H_9 , $J = 6.2$ Hz), 3.61 (s, H_5), 3.25 (t, 4H, 2 x CH_2), 2.53 (t, 4H, 2 x CH_2), 1.45 (s, 9H, Boc- $C(CH_3)_3$), 1.40 (s, 9H, Boc- $C(CH_3)_3$).

δ_C (150 MHz, DMSO- d_6): 163.5 (C_{10}), 155.97 (Boc- $\underline{C}=\underline{O}$), 153.74 (Boc- $\underline{C}=\underline{O}$), 152.24 (C_4), 143.61 (C_7), 126.62 (C_2), 126.60 (C_1), 124.58 (C_6), 114.52 (C_3), 83.45 (Boc- $\underline{C}(CH_3)_3$), 78.83 (Boc- $\underline{C}(CH_3)_3$), 52.96 (C_5), 52.37 ($\underline{CH}_2$), 49.80 (C_8), 47.43 ($\underline{CH}_2$), 28.45 (Boc- $\underline{C}(CH_3)_3$), 28.03 (Boc- $\underline{C}(CH_3)_3$).

HRMS-ESI (+): m/z calculated for $C_{27}H_{39}F_3N_8O_4$: 596.3046. Found $[M+H] = 597.3134$.

Boc-protected 1-(2-(4-((4-(4-Isopropylphenyl)piperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)guanidine (**98**)



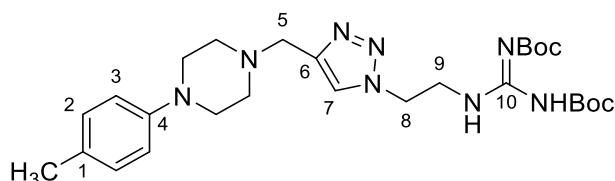
Following method F from **94** (407.3 mg, 1.24 mmol). **98** was obtained as a pale yellow crystalline powder (345.1 mg, 49%, $R_f = 0.65$ [DCM:MeOH 9:1]).

δ_H (600 MHz, DMSO- d_6): 11.46 (NH), 8.45 (NH), 8.31 (s, 1H, H_7), 7.13 (d, 2H, H_2 , $J = 8.7$ Hz), 6.90 (d, 2H, H_3 , $J = 8.7$ Hz), 4.62 (t, 2H, H_8 , $J = 6.4$ Hz), 4.52 (s, 2H, H_5), 3.79 (q, 2H, H_9 , $J = 6.4$ Hz), 3.25 (t, 4H, 2 x CH_2), 2.53 (t, 4H, 2 x CH_2), 2.81 (sept, 1H, H_{11} , $J = 6.8$ Hz), 1.46 (s, 9H, Boc- $C(CH_3)_3$), 1.41 (s, 9H, Boc- $C(CH_3)_3$), 1.16 (d, 6H, H_{12} , $J = 6.8$ Hz), [3.74, 3.48, 3.19, 2.97 (s, br, 4 x CH_2)].

δ_c (150 MHz, DMSO- d_6): 162.78 (C_{10}), 155.42 (Boc-C=O), (Boc-C=O), 151.69 (C_4), 147.59 (C_7), 127.4 (C_6) 127.28 (C_1), 126.82 (C_2), 116.19 (C_3), 82.99 (Boc-C(CH₃)₃), 78.49 (Boc-C(CH₃)₃), 54.88 (C_5), 50.30 (CH₂), 49.35 (C_8) 48.51 (CH₂), 27.94 (Boc-C(CH₃)₃), 27.55 (Boc-C(CH₃)₃), 32.48 (C_{11}), 24.00 (C_{12}).

HRMS-ESI (+): m/z calculated for C₂₇H₃₉F₃N₈O₄ : 570.3642. Found [M+H] = 571.3714.

Boc-protected 1-(2-(4-((4-(p-Tolyl)piperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)guanidine (**99**)



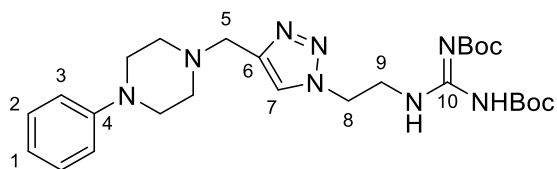
Following method F from **95** (71.8 mg, 0.24 mmol). **98** was obtained as a pale yellow crystalline powder (31.7 mg, 24%, R_f = 0.65 [DCM:MeOH 9:1]).

δ_H (600 MHz, DMSO- d_6): 11.45 (NH), 8.39 (NH), 7.99 (s, 1H, H_7), 7.01 (d, 2H, H_2 , J = 8.6 Hz), 6.80 (d, 2H, H_3 , J = 8.6 Hz), 4.53 (t, 2H, H_8 , J = 5.8 Hz), 3.77 (q, 2H, H_9 , J = 5.8 Hz), 3.59 (s, 2H, H_5) 3.03 (t, 4H, 2 x CH₂), 2.53 (t, 4H, 2 x CH₂), 2.19 (s, CH₃), 1.45 (s, 9H, Boc-C(CH₃)₃), 1.40 (s, 9H, Boc-C(CH₃)₃).

δ_c (150 MHz, DMSO- d_6): 163.47 (C_{10}), 156.06 (Boc-C=O), 152.27 (Boc-C=O), 149.48 (C_4), 143.76 (C_7), 129.80 (C_2), 128.07 (C_1), 127.40 (C_6) 116.07 (C_3), 83.55 (Boc-C(CH₃)₃), 78.82 (Boc-C(CH₃)₃), 53.10 (C_5), 52.75 (CH₂), 49.15 (C_8) 48.73 (CH₂), 28.45 (Boc-C(CH₃)₃), 28.03 (Boc-C(CH₃)₃), 20.50 (CH₃).

HRMS-ESI (+): m/z calculated for C₂₇H₄₂N₈O₄ : 542.3329. Found [M+H] = 543.3417.

Boc-protected 1-(2-(4-((4-Phenylpiperazin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethyl)guanidine (**100**)



Following method F from **96** (234.8 mg, 0.82 mmol). **100** was obtained as a pale yellow crystalline powder (260.4 mg, 60%, $R_f = 0.70$ [DCM:MeOH 9:1]).

δ_H (600 MHz, DMSO- d_6): 11.46 (NH), 8.39 (NH), 7.99 (s, 1H, H_7), , 7.20 (t, 2H, H_2 , $J = 8.0$ Hz), 6.90 (d, 2H, H_3 , $J = 8.0$ Hz), 6.77 (t, 1H, H_1 , $J = 8.0$ Hz), 4.53 (t, 2H, H_8 , $J = 6.0$ Hz), 3.77 (q, 2H, H_9 , $J = 6.0$ Hz), 3.60 (s, 2H, H_5) 3.10 (t, 4H, 2 x CH_2), 2.53 (t, 4H, 2 x CH_2), 1.45 (s, 9H, Boc- $(C(CH_3)_3)$), 1.40 (s, 9H, Boc- $(C(CH_3)_3)$).

δ_C (150 MHz, DMSO- d_6): 163.46 (C_{10}), 155.97 (Boc- $\underline{C}=\underline{O}$), 152.27 (Boc- $\underline{C}=\underline{O}$), 151.51 (C_4), 143.76 (C_7), 129.36 (C_2), 124.53 (C_1), 124.40 (C_6) 116.07 (C_3), 83.55 (Boc- $\underline{C}(CH_3)_3$), 78.82 (Boc- $\underline{C}(CH_3)_3$), 53.10 (C_5), 52.75 ($\underline{CH}_2$), 49.15 (C_8) 48.73 ($\underline{CH}_2$), 28.45 (Boc- $\underline{C}(CH_3)_3$), 28.03 (Boc- $\underline{C}(CH_3)_3$), 20.50 (CH_3).

HRMS-APCI (+): m/z calculated for $C_{26}H_{40}N_8O_4$: 528.3173. Found $[M-H] = 527.3095$.

5.2.3 HPLC Purity Assessment

HPLC purity analysis was conducted using a Varian ProStar system equipped with a Varian Prostar 335 diode array detector and a manual injector (20 μ L). For purity assessment, UV detection at 245 nm was performed and peak purity was confirmed using a purity channel. The stationary phase consisted of an ACE 5 C18-AR column (150 $\times$ 4.6 mm), and the mobile phase used the following gradient system, eluting at 1 mL/min: aqueous ammonium formate buffer (30mM, pH 3.0) for 10 min, linear ramp to 85% methanol buffered with the same system over 25 min, hold at 85% buffered methanol for 10 min. Minimum requirement for purity was set at 95%.

5.3 Biological and Biochemical Methods

5.3.1 MIC Determination Against *M. tuberculosis* H37Rv mc² 6230.

Antibacterial activity was determined using the standard double-dilution method. *M. tuberculosis* mc² 6230 was provided by Prof. William R. Jacobs, Jr. (Albert Einstein College of Medicine, NY, USA). The *M. tuberculosis* mc² 6230 was grown in 7H9 broth supplemented with tyloxapol (0.05%), glycerol (0.5%), OADC (10%), casaminoacids (0.2%), and pantothenate (24 µg/mL). The mycobacteria were grown on agar plates corresponding to the growth medium. Cells were transferred from the plates to the respective liquid medium until the OD₆₀₀, measured on a DEN-1 McFarland densitometer, was equal to that of a 0.5 McFarland standard (VWR catalog # 75785-176). The culture was then diluted to 1:100 and added to the compounds dissolved in the appropriate medium. Plates were incubated at 37 °C until growth was observed in the wells containing no compound (2 to 3 weeks). To positively assess the growth inhibition, 5 µL of a 1 mg/mL solution of resazurin was added to each well and incubated overnight at 37 °C.

5.3.2 Lipophilicity Determination

LogD_{7.4} was analyzed using an HPLC-based method. The UV retention time of reference compounds with known LogD_{7.4} was determined and plotted toward their LogD_{7.4}. Linear regression was used to determine the LogD_{7.4} of unknown compounds. Analysis was performed using a Vanquish Flex HPLC system with variable wavelength detector (Thermo Fisher, Dreieich, Germany) with the following conditions: EC150/2 NUCLEODUR C18 Pyramid column, 5 µm (Macherey Nagel, Düren, Germany); eluent A: 50 mM NH₄OAc pH 7.4, eluent B: acetonitrile, and flow: 0.6 mL/min. The gradient was set to 0–100% B from 0 to 2.5 min, 100% B from 2.5 to 3.0 min, 100–0% B from 3.0 to 3.2 min, and 0% B from 3.2–5.0.

5.3.3 Metabolic Stability in Mouse Liver Microsomes

For the evaluation of phase I metabolic stability, the compound (1 μ M) was incubated with 0.5 mg/mL pooled mouse liver microsomes (Xenotech, Kansas City, USA), 2 mM NADPH, 10 mM MgCl_2 at 37 °C for 120 min on a microplate shaker (Eppendorf, Hamburg, Germany). The metabolic stability of testosterone, verapamil and ketoconazole was determined in parallel to confirm the enzymatic activity of mouse liver microsomes. The incubation was stopped after defined time points by precipitation of aliquots of enzymes with 2 volumes of cold internal standard solution (15 nM diphenhydramine in 10% methanol/acetonitrile). Samples were stored on ice until the end of the incubation and precipitated protein was removed by centrifugation (15 min, 4 °C, 4,000 g). The remaining test compound at the different time points was analyzed by HPLC-MS/MS (Vanquish Flex coupled to a TSQ Altis Plus, Thermo Fisher, Dreieich, Germany) and used to determine half-life ($t_{1/2}$).

5.3.4 Stability in Mouse Plasma

To determine stability in plasma, the compound (1 μ M) was incubated with pooled CD-1 mouse plasma (Neo Biotech, Nanterre, France). Samples were taken at defined time points by mixing aliquots with 4 volumes of ice-cold internal standard solution (12.5 nM diphenhydramine in 10% methanol/acetonitrile). Samples were stored on ice until the end of the incubation and precipitated protein was removed by centrifugation (15 min, 4 °C, 4,000 g, 2 centrifugation steps). The remaining test compound at the different time points was analyzed by HPLC-MS/MS (Vanquish Flex coupled to a TSQ Altis Plus, Thermo Fisher, Dreieich, Germany). The plasma stability of procain, propantheline and diltiazem were determined in parallel to confirm the enzymatic activity.

5.3.5 Cytotoxicity

To obtain information regarding the toxicity of our compounds, their impact on the viability of human cells was investigated. HepG2 cells (2×10^4 cells per well) were seeded in 96-well, flat-bottomed culture plates in 100 μ L culture medium (DMEM containing 10% fetal calve serum, 1% penicillin-streptomycin). Twenty-four hours after seeding the cells, medium was removed and replaced by medium containing

test compounds in a final DMSO concentration of 1%. Compounds were tested in duplicates at a single concentration or, for CC₅₀ determination, at 8 concentrations that were prepared *via* 2-fold serial dilutions in 1% DMSO/medium. Epirubicin and doxorubicin were used as positive controls in serial dilutions starting from 10 µM, and rifampicin was used as a negative control (at 100 µM). The living cell mass was determined 48 h after treatment with compounds by adding 0.1 volumes of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) solution (5 mg/mL sterile PBS) (Sigma, St. Louis, MO) to the wells. After incubating the cells for 30 min at 37 °C (atmosphere containing 5% CO₂), medium was removed and MTT crystals were dissolved in 75 µL of a solution containing 10% SDS and 0.5% acetic acid in DMSO. The optical density (OD) of the samples was determined photometrically at 570 nm in a PHERAstar Omega plate reader (BMG labtech, Ortenberg, Germany). To obtain percent *viability* for each sample, their ODs were related to those of DMSO controls. At least two independent measurements were performed for each compound. The calculation of CC₅₀ was performed using the nonlinear regression function of GraphPad Prism 10 (GraphPad Software, San Diego, CA, USA).

References

1. S. M. A. Mousavi-Sagharchi, A. Ghorbani, M. Meskini and S. D. Siadat, *Journal of Infection and Public Health*, 2025, **18**, 102649.
2. WHO, *Global tuberculosis report 2024*.
3. T. Berida and C. W. Lindsley, *Journal of Medicinal Chemistry*, 2024, **67**, 21633-21640.
4. C. E. Barry, H. I. Boshoff, V. Dartois, T. Dick, S. Ehrt, J. Flynn, D. Schnappinger, R. J. Wilkinson and D. Young, *Nature Reviews Microbiology*, 2009, **7**, 845-855.
5. K. Dheda, F. Mirzayev, D. M. Cirillo, Z. Udwadia, K. E. Dooley, K.-C. Chang, S. V. Omar, A. Reuter, T. Perumal, C. R. Horsburgh, M. Murray and C. Lange, *Nature Reviews Disease Primers*, 2024, **10**, 22.
6. J. O'Neill, *Tackling drug-resistant infections globally: final report and recommendations*, 2016, 84 pp.
7. E. Cambau and M. Drancourt, *Clinical Microbiology and Infection*, 2014, **20**, 196-201.
8. S. V. Gordon and T. Parish, *Microbiology*, 2018, **164**, 437-439.
9. L. J. Alderwick, J. Harrison, G. S. Lloyd and H. L. Birch, *Cold Spring Harb Perspect Med*, 2015, **5**, a021113.
10. M. U. Shiloh and P. A. Champion, *Curr Opin Microbiol*, 2010, **13**, 86-92.
11. D. Natali, G. Cloatre, C. Brosset, P. Verdalle, A. Fauvy, J.-P. Massart, Q. Vo Van, N. Gerard, C. C. Dobler and P. Hovette, *Breathe*, **16**, 200216.
12. A. N. Leung, *Radiology*, 1999, **210**, 307-322.
13. V. Nduba, L. N. Njagi, W. Murithi, Z. Mwongera, J. Byers, G. Logioia, G. Peterson, R. M. Segnitz, K. Fennelly, T. R. Hawn and D. J. Horne, *Nat Commun*, 2024, **15**, 7604.
14. S. B. Cohen, B. H. Gern, J. L. Delahaye, K. N. Adams, C. R. Plumlee, J. K. Winkler, D. R. Sherman, M. Y. Gerner and K. B. Urdahl, *Cell Host & Microbe*, 2018, **24**, 439-446.e434.
15. M. de Martino, L. Lodi, L. Galli and E. Chiappini, *Front Pediatr*, 2019, **7**, 350.
16. M. A. Behr, P. H. Edelstein and L. Ramakrishnan, *Bmj*, 2018, **362**, k2738.
17. D. B. Young, H. P. Gideon and R. J. Wilkinson, *Trends Microbiol*, 2009, **17**, 183-188.
18. M. R. Cronan, *Frontiers in Immunology*, 2022, **Volume 13 - 2022**.
19. S. Jaiswal, S. Fatima, E. Velarde de la Cruz and S. Kumar, *Tuberculosis*, 2025, **152**, 102615.
20. P. L. Lin, C. B. Ford, M. T. Coleman, A. J. Myers, R. Gawande, T. Ioerger, J. Sacchettini, S. M. Fortune and J. L. Flynn, *Nature Medicine*, 2014, **20**, 75-79.
21. M. Silva Miranda, A. Breiman, S. Allain, F. Deknuydt and F. Altare, *Clin Dev Immunol*, 2012, **2012**, 139127.
22. M. R. Cronan, *Front Immunol*, 2022, **13**, 820134.
23. L. Luies and I. d. Preez, *Clinical Microbiology Reviews*, 2020, **33**, 10.1128/cmr.00036-00020.
24. A. Fayyazi, B. Eichmeyer, A. Soruri, S. Schweyer, J. Herms, P. Schwarz and H. J. Radzun, *The Journal of Pathology*, 2000, **191**, 417-425.
25. D. G. Russell, P.-J. Cardona, M.-J. Kim, S. Allain and F. Altare, *Nature Immunology*, 2009, **10**, 943-948.
26. C. Herbert, L. Luies, D. T. Loots and A. A. Williams, *BMC Infectious Diseases*, 2023, **23**, 536.
27. A. VanValkenburg, V. Kaipilyawar, S. Sarkar, S. Lakshminarayanan, C. Cintron, S. Prakash Babu, S. Knudsen, N. M. Joseph, C. R. Horsburgh, P. Sinha, J. J. Ellner, P. B. Narasimhan, W. E. Johnson, N. S. Hochberg and P. Salgame, *Front Immunol*, 2022, **13**, 1011166.
28. S. Ehlers and U. E. Schaible, *Front Immunol*, 2012, **3**, 411.
29. S. Datta and C. A. Evans, *The Lancet Infectious Diseases*, 2020, **20**, 1002-1004.
30. G. Bartolomeu-Gonçalves, J. M. d. Souza, B. T. Fernandes, L. F. A. Spoladori, G. F. Correia, I. M. d. Castro, P. H. G. Borges, G. Silva-Rodrigues, E. R. Tavares, L. M.

- Yamauchi, M. Pelisson, M. R. E. Perugini and S. F. Yamada-Ogatta, *Diseases*, 2024, **12**, 202.
31. P. Yoopetch, O. Wu, J. Jittikoon, M. Thavorncharoensap, S. Youngkong, N. Praditsitthikorn, S. Mahasirimongkol, T. Anothaisintawee, W. Udomsinprasert and U. Chaikledkaew, *Scientific Reports*, 2024, **14**, 17693.
 32. I. Devoid, A. K. Sillah, J. Sutherland, O. Owolabi, O. Ivanova, C. Govathson, K. Hirasen, M. Davies, K. Lönnroth, I. Loum, A. Touray, S. Charlabous, D. Evans and M. Quaife, *Int J Tuberc Lung Dis*, 2022, **26**, 1162-1169.
 33. G. Sotgiu, R. Centis, L. D'Ambrosio and G. B. Migliori, *Cold Spring Harb Perspect Med*, 2015, **5**, a017822.
 34. WHO, *WHO operational handbook on tuberculosis. Module 4: treatment and care*.
 35. A. Mabhula and V. Singh, *MedChemComm*, 2019, **10**, 1342-1360.
 36. E. J. Forget and D. Menzies, *Expert Opin Drug Saf*, 2006, **5**, 231-249.
 37. Y. Fang, H. Xiao, S. Tang, L. Liang, W. Sha and Y. Fang, *Clin Respir J*, 2016, **10**, 449-454.
 38. T. R. Frieden and J. A. Sbarbaro, *Bull World Health Organ*, 2007, **85**, 407-409.
 39. N. Riccardi, D. Canetti, P. Rodari, G. Besozzi, L. Sadari, M. Dettori, L. R. Codecasa and G. Sotgiu, *Curr Res Pharmacol Drug Discov*, 2021, **2**, 100007.
 40. L. C. K. Bell and M. Noursadeghi, *Nat Rev Microbiol*, 2018, **16**, 80-90.
 41. A. Kwara, G. Ramachandran and S. Swaminathan, *Expert Opin Drug Metab Toxicol*, 2010, **6**, 55-68.
 42. K. Cohen and G. Meintjes, *Curr Opin HIV AIDS*, 2010, **5**, 61-69.
 43. M. O'Donnell, *Ann Am Thorac Soc*, 2018, **15**, 306-307.
 44. J. G. Pasipanodya and T. Gumbo, *Clinical Infectious Diseases*, 2013, **57**, 21-31.
 45. S. Srivastava, J. G. Pasipanodya, C. Meek, R. Leff and T. Gumbo, *J Infect Dis*, 2011, **204**, 1951-1959.
 46. F. Fregonese, S. D. Ahuja, O. W. Akkerman, D. Arakaki-Sanchez, I. Ayakaka, P. Baghaei, D. Bang, M. Bastos, A. Benedetti, M. Bonnet, A. Cattamanchi, P. Cegielski, J.-Y. Chien, H. Cox, M. Dedicoat, C. Erkens, P. Escalante, D. Falzon, A. J. Garcia-Prats, M. Gegia, S. H. Gillespie, J. R. Glynn, S. Goldberg, D. Griffith, K. R. Jacobson, J. C. Johnston, E. C. Jones-López, A. Khan, W.-J. Koh, A. Kritski, Z. Y. Lan, J. H. Lee, P. Z. Li, E. L. Maciel, R. M. Galliez, C. S. C. Merle, M. Munang, G. Narendran, V. N. Nguyen, A. Nunn, A. Ohkado, J. S. Park, P. P. J. Phillips, C. Ponnuraja, R. Reves, K. Romanowski, K. Seung, H. S. Schaaf, A. Skrahina, D. v. Soolingen, P. Tabarsi, A. Trajman, L. Trieu, V. V. Banurekha, P. Viikklepp, J.-Y. Wang, T. Yoshiyama and D. Menzies, *The Lancet Respiratory Medicine*, 2018, **6**, 265-275.
 47. World Health Organization, *Global Tuberculosis Report 2022*, 2022.
 48. Z. Lan, N. Ahmad, P. Baghaei, L. Barkane, A. Benedetti, S. K. Brode, J. C. M. Brust, J. R. Campbell, V. W. L. Chang, D. Falzon, L. Guglielmetti, P. Isaakidis, R. R. Kempker, M. Kipiani, L. Kuksa, C. Lange, R. Laniado-Laborín, P. Nahid, D. Rodrigues, R. Singla, Z. F. Udwadia, D. Menzies, N. Ahmad, P. Baghaei, L. Barkane, A. Benedetti, S. K. Brode, J. C. M. Brust, J. R. Campbell, V. W. L. Chang, D. Falzon, L. Guglielmetti, P. Isaakidis, R. R. Kempker, M. Kipiani, L. Kuksa, Z. Lan, C. Lange, R. Laniado-Laborín, P. Nahid, D. Rodrigues, R. Singla, Z. F. Udwadia and D. Menzies, *The Lancet Respiratory Medicine*, 2020, **8**, 383-394.
 49. F. Conradie, A. H. Diacon, N. Ngubane, P. Howell, D. Everitt, A. M. Crook, C. M. Mendel, E. Egizi, J. Moreira, J. Timm, T. D. McHugh, G. H. Wills, A. Bateson, R. Hunt, C. V. Niekerk, M. Li, M. Olugbosi and M. Spigelman, *New England Journal of Medicine*, 2020, **382**, 893-902.
 50. M. Serajian, C. Testagrose, M. Prosperi and C. Boucher, *Scientific Reports*, 2025, **15**, 5104.
 51. R. Mahajan, *Int J Appl Basic Med Res*, 2013, **3**, 1-2.

52. X. Zheng and Y. Av-Gay, *Current Treatment Options in Infectious Diseases*, 2016, **8**, 299-310.
53. V. Singh, *Drug Discovery Today*, 2024, **29**, 103955.
54. E. S. Chung, W. C. Johnson and B. B. Aldridge, *Nature Reviews Microbiology*, 2022, **20**, 529-541.
55. N. E. Mvubu and K. Jacoby, *Heliyon*, 2023, **9**, e22611.
56. P. Singh, N. R. Rameshwaram, S. Ghosh and S. Mukhopadhyay, *Future Microbiology*, 2018, **13**, 689-710.
57. Y. M. Jacobo-Delgado, A. Rodríguez-Carlos, C. J. Serrano and B. Rivas-Santiago, *Front Immunol*, 2023, **14**, 1194923.
58. S. M. Batt, C. E. Burke, A. R. Moorey and G. S. Besra, *The Cell Surface*, 2020, **6**, 100044.
59. B. Mtafya, P. Qwaray, J. John, E. Sichone, A. Shoo, S. H. Gillespie, N. E. Ntinginya and W. Sabiiti, *Tuberculosis*, 2023, **138**, 102275.
60. P. K. Agnivesh, A. Roy, S. Sau, S. Kumar and N. P. Kalia, *Microbial Pathogenesis*, 2025, **198**, 107074.
61. R. Capela, R. Félix, M. Clariano, D. Nunes, M. d. J. Perry and F. Lopes, *International Journal of Molecular Sciences*, 2023, **24**, 10482.
62. N. Tian, H. Chu, Q. Li, H. Sun, J. Zhang, N. Chu and Z. Sun, *European Journal of Medical Research*, 2025, **30**, 267.
63. A. Diab, H. Dickerson and O. Al Musaimi, *Pharmaceuticals*, 2025, **18**, 70.
64. N. Blondiaux, M. Moune, M. Desroses, R. Frita, M. Flipo, V. Mathys, K. Soetaert, M. Kiass, V. Delorme, K. Djaout, V. Trebosc, C. Kemmer, R. Wintjens, A. Wohlkönig, R. Antoine, L. Huot, D. Hot, M. Coscolla, J. Feldmann, S. Gagneux, C. Loch, P. Brodin, M. Gitzinger, B. Déprez, N. Willand and A. R. Baulard, *Science*, 2017, **355**, 1206-1211.
65. H. Levett, *Journal*, 2025.
66. K. Mikušová, H. Huang, T. Yagi, M. Holsters, D. Vereecke, W. D'Haese, M. S. Scherman, P. J. Brennan, M. R. McNeil and D. C. Crick, *Journal of Bacteriology*, 2005, **187**, 8020-8025.
67. K. Mikusová, V. Makarov and J. Neres, *Curr Pharm Des*, 2014, **20**, 4379-4403.
68. N. Hariguchi, X. Chen, Y. Hayashi, Y. Kawano, M. Fujiwara, M. Matsuba, H. Shimizu, Y. Ohba, I. Nakamura, R. Kitamoto, T. Shinohara, Y. Uematsu, S. Ishikawa, M. Itotani, Y. Haraguchi, I. Takemura and M. Matsumoto, *Antimicrob Agents Chemother*, 2020, **64**.
69. R. Dawson, A. H. Diacon, S. Takuva, Y. Liu, B. Zheng, V. Karwe and J. Hafkin, *Trials*, 2024, **25**, 70.
70. V. Makarov, G. Manina, K. Mikusova, U. Möllmann, O. Ryabova, B. Saint-Joanis, N. Dhar, M. R. Pasca, S. Buroni, A. P. Lucarelli, A. Milano, E. De Rossi, M. Belanova, A. Bobovska, P. Dianiskova, J. Kordulakova, C. Sala, E. Fullam, P. Schneider, J. D. McKinney, P. Brodin, T. Christophe, S. Waddell, P. Butcher, J. Albrethsen, I. Rosenkrands, R. Brosch, V. Nandi, S. Bharath, S. Gaonkar, R. K. Shandil, V. Balasubramanian, T. Balganes, S. Tyagi, J. Grosset, G. Riccardi and S. T. Cole, *Science*, 2009, **324**, 801-804.
71. N. Heinrich, V. de Jager, J. Dreisbach, P. Gross-Demel, S. Schultz, S. Gerbach, F. Kloss, R. Dawson, K. Narunsky, L. Matt, L. Wildner, T. D. McHugh, U. Fuhr, B. H. Aldana, C. Mouhaddad, L. t. Brake, M. J. Boeree, R. E. Aarnoutse, E. M. Svensson, X. Gong, P. P. J. Phillips, A. H. Diacon, M. Hoelscher, M. Hoelscher, J. Dreisbach, P. Gross – Demel, L. Wagnerberger, N. Heinrich, A. Razid, W. Lutchmun, I. Noreña, L. Paramo Diaz, D. Sloan, W. Sabiiti, S. Gillespie, L. te Brake, E. Svensson, C. Mouhaddad, R. Aarnoutse, M. Boeree, R. Stemkens, S. Koele, A. Bateson, R. Hunt, T. D. McHugh, L. Muraro Wildner, P. Solanki, P. Phillips, X. Gong, B. Aldana, A. Crook, R. Dawson, K. Narunsky, S. Arnolds, A. Diacon, V. de Jager, S. Friedrich, I. Sanne, M. Rassool, G. Churchyard, M. Sebe, H. Makkan, L. Mokaba, N. Madikizela, J. Mdluli, J. Sithole, R. Wallis, T. Beattie,

- N. E. Ntinginya, C. Mangu, C. Manyama, I. Sabi, B. Mtafya, L. T. Minja, F. Mhimbira, B. Mbeya, T. Zumba, N. Chibunu, M. Sasamalo, K. Reither, L. Jugheli, N. Sam, G. Kibiki, H. Semvua, S. Mpagama, A. Liyoyo, B. R. Adegbite, A. A. Adegnika, M. P. Grobusch, B. Kirenga, C. Khosa, I. Timana, M. Nliwasa and M. Mukoka, *The Lancet Microbe*, 2025, **6**.
72. P. S. Shirude, R. Shandil, C. Sadler, M. Naik, V. Hosagrahara, S. Hameed, V. Shinde, C. Bathula, V. Humnabadkar, N. Kumar, J. Reddy, V. Panduga, S. Sharma, A. Ambady, N. Hegde, J. Whiteaker, R. E. McLaughlin, H. Gardner, P. Madhavapeddi, V. Ramachandran, P. Kaur, A. Narayan, S. Guptha, D. Awasthy, C. Narayan, J. Mahadevaswamy, K. G. Vishwas, V. Ahuja, A. Srivastava, K. R. Prabhakar, S. Bharath, R. Kale, M. Ramaiah, N. R. Choudhury, V. K. Sambandamurthy, S. Solapure, P. S. Iyer, S. Narayanan and M. Chatterji, *Journal of Medicinal Chemistry*, 2013, **56**, 9701-9708.
 73. M. S. Butler, W. Vollmer, E. C. A. Goodall, R. J. Capon, I. R. Henderson and M. A. T. Blaskovich, *ACS Infectious Diseases*, 2024, **10**, 3440-3474.
 74. K. V. Singh, T. M. Coque and B. E. Murray, *Antimicrobial Agents and Chemotherapy*, 1996, **40**, 2142-2146.
 75. S. Ramón-García, R. González del Río, M. P. Arenaz-Callao, H. I. Boshoff, J. Rullas, S. Anca, M. Cacho Izquierdo, E. Porras de Francisco, E. Pérez Herrán, A. Santos-Villarejo, A. Mendoza-Losana, S. Ferrer-Bazaga, C. J. Thompson, D. Barros Aguirre and R. H. Bates, *Drug Resistance Updates*, 2025, **80**, 101213.
 76. R. E. Lee, M. Protopopova, E. Crooks, R. A. Slayden, M. Terrot and C. E. Barry, *Journal of Combinatorial Chemistry*, 2003, **5**, 172-187.
 77. M. Imran, M. K. Arora, A. Chaudhary, S. A. Khan, M. Kamal, M. M. Alshammari, R. M. Alharbi, N. A. Althomali, I. M. Alzimam, A. A. Alshammari, B. H. Alharbi, A. Alshengeti, A. A. Alsaleh, S. A. Alqahtani and A. A. Rabaan, *Biomedicines*, 2022, **10**.
 78. K. A. Sacksteder, M. Protopopova, C. E. Barry, 3rd, K. Andries and C. A. Nacy, *Future Microbiol*, 2012, **7**, 823-837.
 79. T. Hu, X. Yang, F. Liu, S. Sun, Z. Xiong, J. Liang, X. Yang, H. Wang, X. Yang, L. W. Guddat, H. Yang, Z. Rao and B. Zhang, *Structure*, 2022, **30**, 1395-1402.e1394.
 80. D. Forge, D. Cappoen, J. Laurent, D. Stanicki, A. Mayence, T. L. Huang, L. Verschaeve, K. Huygen and J. J. Vanden Eynde, *Eur J Med Chem*, 2012, **49**, 95-101.
 81. A. Rahman, Trinity College Dublin, The University of Dublin, 2021.
 82. U. H. Manjunatha, S. R. SP, 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. Glynn, P. W. Smith and T. T. Diagana, *Sci Transl Med*, 2015, **7**, 269ra263.
 83. H. L. Yale, *Journal of Medicinal and Pharmaceutical Chemistry*, 1959, **1**, 121-133.
 84. M. Inoue, Y. Sumii and N. Shibata, *ACS Omega*, 2020, **5**, 10633-10640.
 85. I. Research, *Fluorinated Drugs Market Growth Analysis, Market Dynamics, Key Players and Innovations, Outlook and Forecast 2025-2032*, 2025.
 86. P. Shah and A. D. Westwell, *Journal of Enzyme Inhibition and Medicinal Chemistry*, 2007, **22**, 527-540.
 87. N. A. Meanwell, *Chem Res Toxicol*, 2016, **29**, 564-616.
 88. C.-C. Tseng, G. Baillie, G. Donvito, M. A. Mustafa, S. E. Juola, C. Zanato, C. Massarenti, S. Dall'Angelo, W. T. A. Harrison, A. H. Lichtman, R. A. Ross, M. Zanda and I. R. Greig, *Journal of Medicinal Chemistry*, 2019, **62**, 5049-5062.
 89. M. Jagodzinska, F. Huguenot, G. Candiani and M. Zanda, *ChemMedChem*, 2009, **4**, 49-51.
 90. E. C. Agency, Per- and polyfluoroalkyl substances (PFAS), <https://echa.europa.eu/hot-topics/perfluoroalkyl-chemicals-pfas>.

91. I. T. Cousins, J. C. DeWitt, J. Glüge, G. Goldenman, D. Herzke, R. Lohmann, C. A. Ng, M. Scherlinger and Z. Wang, *Environmental Science: Processes & Impacts*, 2020, **22**, 2307-2312.
92. D. O'Hagan and R. J. Young, *Medicinal Chemistry Research*, 2023, **32**, 1231-1234.
93. Z. Guo, A. A. Attar, Q. Qiqige, R. J. Lundgren and S. Joudan, *Environmental Science & Technology*, 2025, **59**, 1367-1377.
94. M. F. Khan and C. D. Murphy, *Appl Microbiol Biotechnol*, 2021, **105**, 9359-9369.
95. J. J. Rojas and J. A. Bull, *Journal of Medicinal Chemistry*, 2023, **66**, 12697-12709.
96. P. Wipf, T. Mo, S. J. Geib, D. Caridha, G. S. Dow, L. Gerena, N. Roncal and E. E. Milner, *Org Biomol Chem*, 2009, **7**, 4163-4165.
97. J. T. Welch and D. S. Lim, *Bioorg Med Chem*, 2007, **15**, 6659-6666.
98. Z. Liu, T. Lu and Q. Chen, *Carbon*, 2021, **171**, 514-523.
99. R. F. Ribeiro, A. V. Marenich, C. J. Cramer and D. G. Truhlar, *Physical Chemistry Chemical Physics*, 2011, **13**, 10908-10922.
100. A. Daina, O. Michielin and V. Zoete, *Scientific Reports*, 2017, **7**, 42717.
101. T. Cheng, Y. Zhao, X. Li, F. Lin, Y. Xu, X. Zhang, Y. Li, R. Wang and L. Lai, *Journal of Chemical Information and Modeling*, 2007, **47**, 2140-2148.
102. S. A. Wildman and G. M. Crippen, *Journal of Chemical Information and Computer Sciences*, 1999, **39**, 868-873.
103. I. Moriguchi, S. Hirono, I. Nakagome and H. Hirano, *CHEMICAL & PHARMACEUTICAL BULLETIN*, 1994, **42**, 976-978.
104. A. Daina, O. Michielin and V. Zoete, *Journal of Chemical Information and Modeling*, 2014, **54**, 3284-3301.
105. V. N. Viswanadhan, A. K. Ghose, G. R. Revankar and R. K. Robins, *Journal of Chemical Information and Computer Sciences*, 1989, **29**, 163-172.
106. Z. Congliang, W. Yan and W. Fuan, *Bulletin of the Korean Chemical Society*, 2007, **28**, 1183-1186.
107. J. Sangster, *Journal of Physical and Chemical Reference Data*, 1989, **18**, 1111-1229.
108. E. Engelage, N. Schulz, F. Heinen, S. M. Huber, D. G. Truhlar and C. J. Cramer, *Chemistry – A European Journal*, 2018, **24**, 15983-15987.
109. B. Ha, K. P. Larsen, J. Zhang, Z. Fu, E. Montabana, L. N. Jackson, D.-H. Chen and E. V. Puglisi, *Nat Commun*, 2021, **12**, 2500.
110. M. Minneci, M. Misevicius and I. Rozas, *Chemistry – A European Journal*, 2024, **30**, e202401954.
111. H. Yekkezare, H. Tajik and N. M. Mahmoodi, *Journal of Molecular Structure*, 2023, **1285**, 135454.
112. T. Wdowik and D. Gryko, *ACS Catalysis*, 2022, **12**, 6517-6531.
113. S. Paul, D. Filippini, F. Ficarra, H. Melnychenko, C. Janot and M. Silvi, *Journal of the American Chemical Society*, 2023, **145**, 15688-15694.
114. S. Liang, K. Hofman, M. Friedrich, J. Keller and G. Manolikakes, *ChemSusChem*, 2021, **14**, 4878-4902.
115. V. Debrauwer, I. Leito, M. Lõkov, S. Tshepelevitsh, M. Parmentier, N. Blanchard and V. Bizet, *ACS Organic & Inorganic Au*, 2021, **1**, 43-50.
116. A. Quemard, J. C. Sacchettini, A. Dessen, C. Vilcheze, R. Bittman, W. R. Jacobs, Jr. and J. S. Blanchard, *Biochemistry*, 1995, **34**, 8235-8241.
117. B. Vögeli, R. G. Rosenthal, G. M. M. Stoffel, T. Wagner, P. Kiefer, N. S. Cortina, S. Shima and T. J. Erb, *J Biol Chem*, 2018, **293**, 17200-17207.
118. M. V. B. Dias, I. B. Vasconcelos, A. M. X. Prado, V. Fadel, L. A. Basso, W. F. de Azevedo and D. S. Santos, *Journal of Structural Biology*, 2007, **159**, 369-380.
119. D. A. Rozwarski, C. Vilchèze, M. Sugantino, R. Bittman and J. C. Sacchettini, *Journal of Biological Chemistry*, 1999, **274**, 15582-15589.

120. A. Chollet, L. Mourey, C. Lherbet, A. Delbot, S. Julien, M. Baltas, J. Bernadou, G. Pratviel, L. Maveyraud and V. Bernardes-Génisson, *Journal of Structural Biology*, 2015, **190**, 328-337.
121. J. Rizet, L. Maveyraud, D. Rengel, V. Guillet, G. Publicola, F. Rodriguez, C. Lherbet and L. Mourey, *ChemMedChem*, 2025, **20**, e202500079.
122. 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, *Journal of Medicinal Chemistry*, 2013, **56**, 8533-8542.
123. J. P. Arcon, A. G. Turjanski, M. A. Martí and S. Forli, *Methods Mol Biol*, 2021, **2266**, 39-72.
124. A. Saha, S. Sengupta, A. Virmani and A. Kumar, *Journal of Chemical Sciences*, 2022, **134**, 98.
125. G. C. H. S. r. B. J. R. B. T. Jaakkola, *DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking*.
126. U. F. Röhrig, M. Goullieux, M. Bugnon and V. Zoete, *Journal of Chemical Information and Modeling*, 2023, **63**, 3925-3940.
127. V. Zoete, T. Schuepbach, C. Bovigny, P. Chaskar, A. Daina, U. F. Röhrig and O. Michielin, *J Comput Chem*, 2016, **37**, 437-447.
128. U. Haberthür and A. Caflisch, *Journal of Computational Chemistry*, 2008, **29**, 701-715.
129. V. Zoete, M. A. Cuendet, A. Grosdidier and O. Michielin, *Journal of Computational Chemistry*, 2011, **32**, 2359-2368.
130. R. Ferreira de Freitas and M. Schapira, *Medchemcomm*, 2017, **8**, 1970-1981.
131. O. Trott and A. J. Olson, *Journal of Computational Chemistry*, 2010, **31**, 455-461.
132. J. Eberhardt, D. Santos-Martins, A. F. Tillack and S. Forli, *Journal of Chemical Information and Modeling*, 2021, **61**, 3891-3898.
133. K. W. Lexa and H. A. Carlson, *Quarterly Reviews of Biophysics*, 2012, **45**, 301-343.
134. C. W. Murray, C. A. Baxter and A. D. Frenkel, *Journal of Computer-Aided Molecular Design*, 1999, **13**, 547-562.
135. C. Talarico, S. Gervasoni, C. Manelfi, A. Pedretti, G. Vistoli and A. R. Beccari, *International Journal of Molecular Sciences*, 2020, **21**, 2265.
136. H. Wei and J. A. McCammon, *npj Drug Discovery*, 2024, **1**, 1.
137. D. Lu, D. Luo, Y. Zhang and B. Wang, *Journal of Chemical Theory and Computation*, 2024, **20**, 6414-6423.
138. A. May and M. Zacharias, *Journal of Medicinal Chemistry*, 2008, **51**, 3499-3506.
139. P. A. Ravindranath, S. Forli, D. S. Goodsell, A. J. Olson and M. F. Sanner, *PLoS Comput Biol*, 2015, **11**, e1004586.
140. F. Gaudreault, M. Chartier and R. Najmanovich, *Bioinformatics*, 2012, **28**, i423-i430.
141. J. J. Clark, M. L. Benson, R. D. Smith and H. A. Carlson, *PLoS Comput Biol*, 2019, **15**, e1006705.
142. M. Mori, S. Villa, S. Ciceri, D. Colombo, P. Ferraboschi and F. Meneghetti, *Molecules*, 2021, **26**.
143. M. L. Danielson and M. A. Lill, *Proteins*, 2012, **80**, 246-260.
144. C. Trujillo, A. A. Rodriguez-Sanz and I. Rozas, *Molecules*, 2015, **20**, 9214-9228.
145. M. R. Koebel, G. Schmadeke, R. G. Posner and S. Sirimulla, *Journal of Cheminformatics*, 2016, **8**, 27.
146. A. Bhatnagar, R. P. Pakhariya and G. Pemawat, *Topics in Current Chemistry*, 2025, **383**, 27.
147. Q. Guan, S. Xing, L. Wang, J. Zhu, C. Guo, C. Xu, Q. Zhao, Y. Wu, Y. Chen and H. Sun, *Journal of Medicinal Chemistry*, 2024, **67**, 7788-7824.

148. G. Munagala, K. R. Yempalla, S. Singh, S. Sharma, N. P. Kalia, V. S. Rajput, S. Kumar, S. D. Sawant, I. A. Khan, R. A. Vishwakarma and P. P. Singh, *Organic & Biomolecular Chemistry*, 2015, **13**, 3610-3624.
149. S. M. Somagond, R. R. Kamble, P. K. Bayannavar, S. K. J. Shaikh, S. D. Joshi, V. M. Kumbar, A. R. Nesaragi and M. Y. Kariduraganavar, *Archiv der Pharmazie*, 2019, **352**.
150. K. M. Naidu, S. Srinivasarao, N. Agnieszka, A.-K. Ewa, M. M. K. Kumar and K. V. G. Chandra Sekhar, *Bioorganic & Medicinal Chemistry Letters*, 2016, **26**, 2245-2250.
151. T. R. Deshmukh, S. P. Khare, V. S. Krishna, D. Sriram, J. N. Sangshetti, V. M. Khedkar and B. B. Shingate, *Synthetic Communications*, 2020, **50**, 271-288.
152. S. Heiles, R. J. Cooper, M. J. DiTucci and E. R. Williams, *Chemical Science*, 2015, **6**, 3420-3429.
153. O. J. Wouters, M. McKee and J. Luyten, *JAMA*, 2020, **323**, 844-853.
154. P. H. van der Graaf, *Clinical Pharmacology & Therapeutics*, 2022, **111**, 983-985.
155. A. M. Davis and R. J. Riley, *Current Opinion in Chemical Biology*, 2004, **8**, 378-386.
156. W. Zhu, Y. Wang, Y. Niu, L. Zhang and Z. Liu, *Health Data Science*, 2023, **3**, 0098.
157. B. B. Shaik and R. Karpoomath, *Bioorganic & Medicinal Chemistry Letters*, 2024, **109**, 129846.
158. G. Biala, E. Kedzierska, M. Kruk-Slomka, J. Orzelska-Gorka, S. Hmaidan, A. Skrok, J. Kaminski, E. Havrankova, D. Nadaska and I. Malik, *Pharmaceuticals (Basel)*, 2023, **16**.
159. C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Advanced Drug Delivery Reviews*, 2001, **46**, 3-26.
160. A. K. Ghose, V. N. Viswanadhan and J. J. Wendoloski, *J Comb Chem*, 1999, **1**, 55-68.
161. D. F. Veber, S. R. Johnson, H.-Y. Cheng, B. R. Smith, K. W. Ward and K. D. Kopple, *Journal of Medicinal Chemistry*, 2002, **45**, 2615-2623.
162. M. Egbert, A. Whitty, G. M. Keserü and S. Vajda, *Journal of Medicinal Chemistry*, 2019, **62**, 10005-10025.
163. Bradley C. Doak, B. Over, F. Giordanetto and J. Kihlberg, *Chemistry & Biology*, 2014, **21**, 1115-1142.
164. D. Asano, H. Takakusa and D. Nakai, *Pharmaceutics*, 2023, **16**.
165. J. A. Arnott and S. L. Planey, *Expert Opinion on Drug Discovery*, 2012, **7**, 863-875.
166. G. Piccaro, G. Poce, M. Biava, F. Giannoni and L. Fattorini, *The Journal of Antibiotics*, 2015, **68**, 711-714.
167. F. Lovering, J. Bikker and C. Humblet, *Journal of Medicinal Chemistry*, 2009, **52**, 6752-6756.
168. A. Daina and V. Zoete, *ChemMedChem*, 2016, **11**, 1117-1121.
169. A. Davis, G. Meintjes and R. J. Wilkinson, *Current Treatment Options in Neurology*, 2018, **20**, 5.
170. Y. H. Wang, Y. Li, S. L. Yang and L. Yang, *J Chem Inf Model*, 2005, **45**, 750-757.
171. J. Ahmed, II, A. A. Abdul Hamid, K. B. Abd Halim and A. T. Che Has, *Heliyon*, 2022, **8**, e09777.
172. K. T. Savjani, A. K. Gajjar and J. K. Savjani, *ISRN Pharm*, 2012, **2012**, 195727.
173. J. Wang, L. Benier and M. Winterhalter, *ACS Omega*, 2018, **3**, 17481-17486.
174. A. M. Seddon, D. Casey, R. V. Law, A. Gee, R. H. Templer and O. Ces, *Chemical Society Reviews*, 2009, **38**, 2509-2519.
175. M. Baciú, S. C. Sebai, O. Ces, X. Mulet, J. A. Clarke, G. C. Shearman, R. V. Law, R. H. Templer, C. Plisson, C. A. Parker and A. Gee, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2006, **364**, 2597-2614.
176. S. Rendic and F. P. Guengerich, *Chem Res Toxicol*, 2015, **28**, 38-42.
177. B. Wang, L.-P. Yang, X.-Z. Zhang, S.-Q. Huang, M. Bartlam and S.-F. Zhou, *Drug Metabolism Reviews*, 2009, **41**, 573-643.
178. J. Bariwal and E. Van der Eycken, *Chemical Society Reviews*, 2013, **42**, 9283-9303.

179. F. Paul, J. Patt and J. F. Hartwig, *Journal of the American Chemical Society*, 1994, **116**, 5969-5970.
180. J. P. Wolfe, S. Wagaw and S. L. Buchwald, *Journal of the American Chemical Society*, 1996, **118**, 7215-7216.
181. D. S. Surry and S. L. Buchwald, *Angewandte Chemie International Edition*, 2008, **47**, 6338-6361.
182. D. S. Surry and S. L. Buchwald, *Chemical Science*, 2011, **2**, 27-50.
183. F. Mo, D. Qiu, L. Zhang and J. Wang, *Chemical Reviews*, 2021, **121**, 5741-5829.
184. in *Comprehensive Organic Name Reactions and Reagents*, DOI: <https://doi.org/10.1002/9780470638859.conrr558>, pp. 2471-2475.
185. D. J. Wallace and C.-y. Chen, *Tetrahedron Letters*, 2002, **43**, 6987-6990.
186. B. A. Sweetman and P. J. Guiry, *Tetrahedron*, 2018, **74**, 5567-5581.
187. M. C. Carreño, J. García Ruano, G. Sanz, M. A. Toledo and A. Urbano, *Tetrahedron Letters*, 1996, **37**, 4081-4084.
188. A. Bachki, F. Foubelo and M. Yus, *Tetrahedron*, 1994, **50**, 5139-5146.
189. M. Bergström, G. Suresh, V. R. Naidu and C. R. Unelius, *European Journal of Organic Chemistry*, 2017, **2017**, 3234-3239.
190. P. Bovonsombat, W. Lorpai boon, S. Laoboonchai, P. Sriprachaya-anunt, W. Yimkosol, N. Siriphatcharachaikul, P. Siricharoensang, T. Kangwannarakul, J. Maeda, S. Losuwanakul and M. Mahesh Abhyankar, *Tetrahedron Letters*, 2020, **61**, 152461.
191. F. A. L. Anet and J. M. Muchowski, *Canadian Journal of Chemistry*, 1960, **38**, 2526-2528.
192. J. K. Laha, P. Gupta and A. Hazra, *The Journal of Organic Chemistry*, 2024, **89**, 725-730.
193. F. D. Bellamy and K. Ou, *Tetrahedron Letters*, 1984, **25**, 839-842.
194. M. Minneci, M. Misevicius and I. Rozas, *Bioorganic & Medicinal Chemistry Letters*, 2023, **90**, 129346.
195. S. Chandrappa, K. Vinaya, T. Ramakrishnappa and K. S. Rangappa, *Synlett*, 2010, **2010**, 3019-3022.
196. M. A. McLaughlin and D. M. Barnes, *Tetrahedron Letters*, 2006, **47**, 9095-9097.
197. M. Kumar, U. Sharma, S. Sharma, V. Kumar, B. Singh and N. Kumar, *RSC Advances*, 2013, **3**, 4894-4898.
198. K. Feichtinger, C. Zapf, H. L. Sings and M. Goodman, *The Journal of Organic Chemistry*, 1998, **63**, 3804-3805.
199. Y. F. Yong, J. A. Kowalski, J. C. Thoen and M. A. Lipton, *Tetrahedron Letters*, 1999, **40**, 53-56.
200. H.-O. Kim, F. Mathew and C. Ogbu, *Synlett*, 1999, **1999**, 193-194.
201. K. Feichtinger, H. L. Sings, T. J. Baker, K. Matthews and M. Goodman, *The Journal of Organic Chemistry*, 1998, **63**, 8432-8439.
202. P. Pracht and S. Grimme, *The Journal of Physical Chemistry A*, 2021, **125**, 5681-5692.
203. C. Wagen and A. Wagen, *Efficient and Accurate pKa Prediction Enabled by Pre-Trained Machine-Learned Interatomic Potentials*, 2024.
204. W. S. Bechara, I. K. Sagamanova, L. Thai-Savard, M. Dauphinais, S. Régnier, C. Noël, S. B. D. Jarvis and A. B. Charette, *Angewandte Chemie International Edition*, 2025, **64**, e202420312.
205. M. Meldal and C. W. Tornøe, *Chemical Reviews*, 2008, **108**, 2952-3015.
206. J. Héron and D. Balcells, *ACS Catalysis*, 2022, **12**, 4744-4753.
207. M. Meldal and F. Diness, *Trends in Chemistry*, 2020, **2**, 569-584.
208. A. D. Moorhouse, J. A. Homer and J. E. Moses, *Chem*, 2023, **9**, 2063-2077.
209. F. Rodríguez, I. Rozas, M. Kaiser, R. Brun, B. Nguyen, W. D. Wilson, R. N. García and C. Dardonville, *Journal of Medicinal Chemistry*, 2008, **51**, 909-923.

210. A. L. Perryman, T. P. Stratton, S. Ekins and J. S. Freundlich, *Pharmaceutical Research*, 2016, **33**, 433-449.
211. M. Deodhar, S. B. Al Rihani, M. J. Arwood, L. Darakjian, P. Dow, J. Turgeon and V. Michaud, *Pharmaceutics*, 2020, **12**.
212. S. Kim, K. Kang, H. Kim and M. Seo, *Toxics*, 2024, **12**, 126.
213. S. Bakshi, M. Kaur, A. Verma and S. Sharma, *Journal of Biochemical and Molecular Toxicology*, 2023, **37**, e23386.
214. G. Indrayanto, G. S. Putra and F. Suhud, in *Profiles of Drug Substances, Excipients and Related Methodology*, ed. A. A. Al-Majed, Academic Press, 2021, vol. 46, pp. 273-307.
215. H. K. Nivatya, A. Singh, N. Kumar, Sonam, L. Sharma, V. Singh, R. Mishra, N. Gaur and A. K. Mishra, *Future Journal of Pharmaceutical Sciences*, 2025, **11**, 111.
216. A. Kahvedžić-Seljubic, S.-M. Nathwani, D. M. Zisterer and I. Rozas, *European Journal of Medicinal Chemistry*, 2016, **117**, 269-282.
217. B. Kelly, D. H. O'Donovan, J. O'Brien, T. McCabe, F. Blanco and I. Rozas, *The Journal of Organic Chemistry*, 2011, **76**, 9216-9227.
218. N. A. Meanwell and O. Loiseleur, *Journal of Agricultural and Food Chemistry*, 2022, **70**, 10972-11004.
219. Y. Zhao and D. G. Truhlar, *Theoretical Chemistry Accounts*, 2008, **120**, 215-241.
220. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, 2016.
221. A. V. Marenich, C. J. Cramer and D. G. Truhlar, *The Journal of Physical Chemistry B*, 2009, **113**, 6378-6396.
222. Y. Sumii, K. Sasaki, S. Tsuzuki and N. Shibata, *Journal*, 2019, **24**, 3610.
223. W. Guo, V. Laserna, J. Rintjema and A. W. Kleij, *Advanced Synthesis & Catalysis*, 2016, **358**, 1602-1607.
224. J. J. Loman, E. R. Carnaghan, T. A. Hamlin, J. M. Ovian, C. B. Kelly, M. A. Mercadante and N. E. Leadbeater, *Organic & Biomolecular Chemistry*, 2016, **14**, 3883-3888.
225. C. Zhao, X. Shi, Z. Zhang, J. Wang and G. Chen, *New Journal of Chemistry*, 2024, **48**, 14067-14070.
226. S. Kumari, C. B. Mishra and M. Tiwari, *Bioorganic & Medicinal Chemistry Letters*, 2015, **25**, 1092-1099.
227. Y. F. Yong, J. A. Kowalski, J. C. Thoen and M. A. Lipton, *Tetrahedron Letters*, 1999, **40**, 53-56.
228. F. Rodriguez, I. Rozas, J. E. Ortega, A. M. Erdozain, J. J. Meana and L. F. Callado, *Journal of Medicinal Chemistry*, 2008, **51**, 3304-3312.
229. A. Srivastava and D. Loganathan, *Glycoconjugate Journal*, 2013, **30**, 769-780.
230. M. Minneci, B. Farkaš, A. Rahman, A. Kempf, I. Nehlmeier, S. Pöhlmann and I. Rozas, *ChemMedChem*, 2026, **21**, e202501000.

Appendix

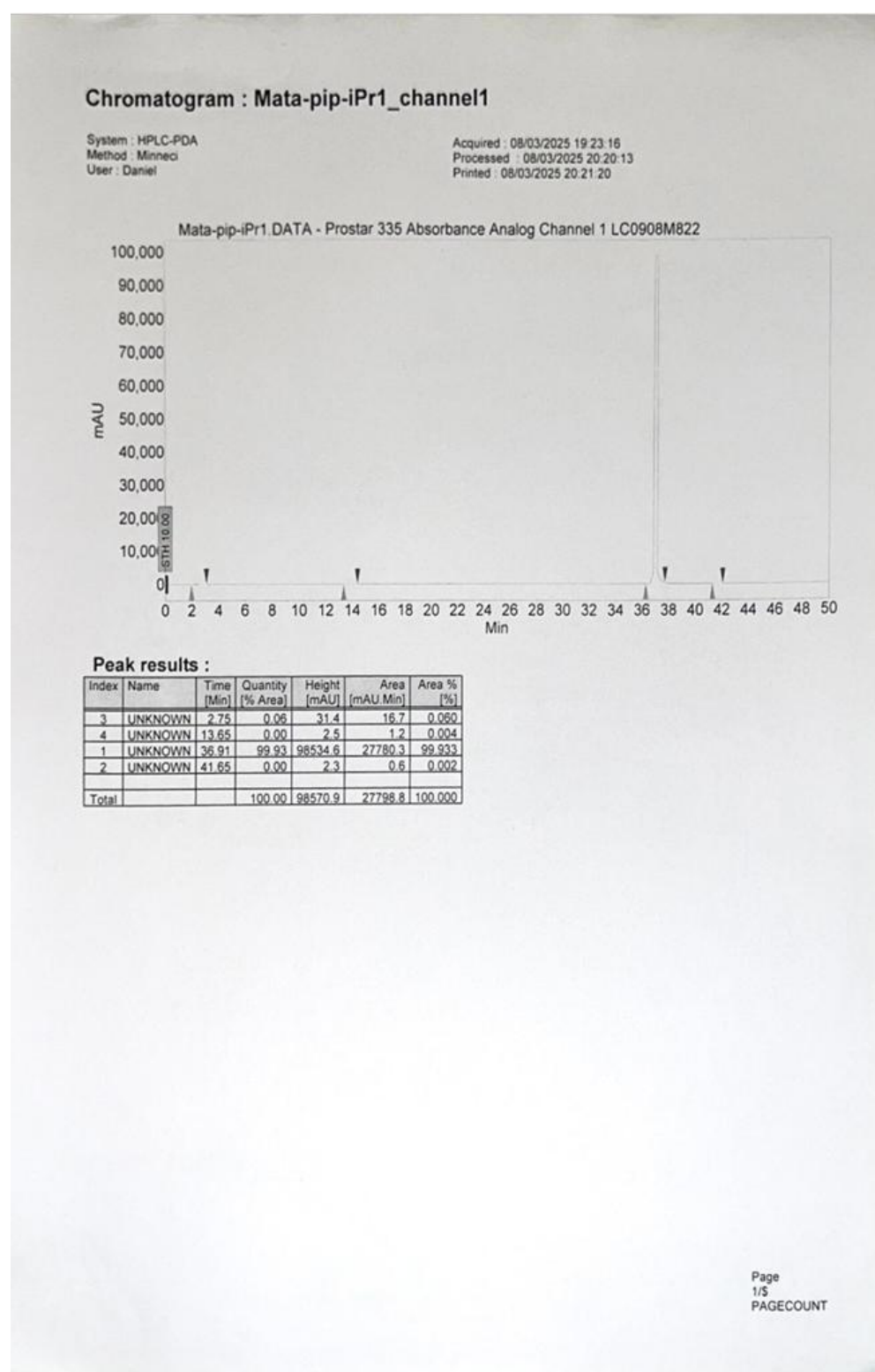
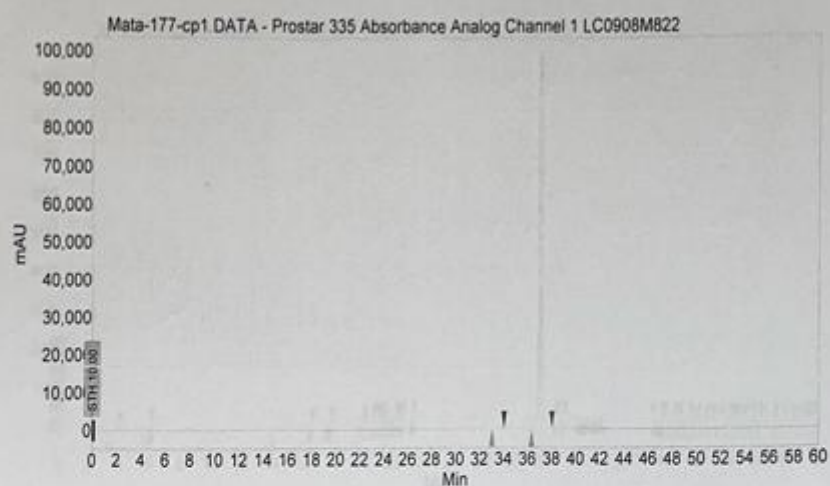


Figure A1 – Chromatogram for 3.

Chromatogram : Mata-177-cp1_channel1

System : HPLC-PDA
Method : Minnece
User : Daniel

Acquired : 07/03/2025 13:28:20
Processed : 07/03/2025 14:31:36
Printed : 07/03/2025 14:34:19



Peak results :

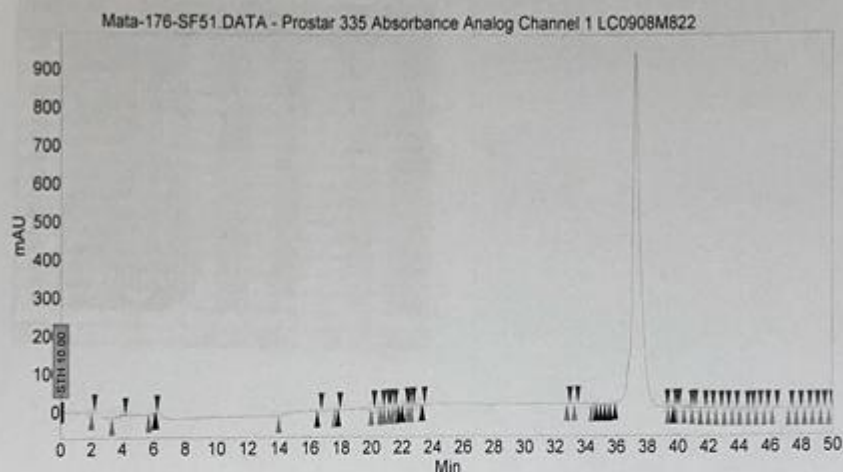
Index	Name	Time (Min)	Quantity (% Area)	Height (mAU)	Area (mAU.Min)	Area % (%)
1	UNKNOWN	33.51	1.27	682.6	321.9	1.271
2	UNKNOWN	36.95	98.73	98546.2	24997.3	98.729
Total			100.00	99328.8	25319.2	100.000

Figure A2 – Chromatogram for 4.

Chromatogram : Mata-176-SF51_channel1

System : HPLC-PDA
Method : Minnet
User : Daniel

Acquired : 06/03/2025 15:34:50
Processed : 06/03/2025 16:27:34
Printed : 06/03/2025 16:36:05



Peak results :

Index	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU.Min]	Area % [%]
1	UNKNOWN	2.04	0.02	0.8	0.1	0.016
2	UNKNOWN	3.79	0.14	1.5	0.8	0.136
3	UNKNOWN	6.01	0.11	2.9	0.6	0.109
4	UNKNOWN	6.09	0.03	2.4	0.2	0.030
5	UNKNOWN	6.13	0.01	1.4	0.1	0.013
6	UNKNOWN	15.16	1.75	5.5	9.9	1.758
7	UNKNOWN	16.45	0.01	0.5	0.1	0.015
8	UNKNOWN	17.72	0.01	0.4	0.1	0.013
9	UNKNOWN	17.80	0.00	0.3	0.0	0.005
10	UNKNOWN	20.03	0.01	0.2	0.0	0.005
11	UNKNOWN	20.63	0.01	0.5	0.1	0.011
12	UNKNOWN	20.92	0.01	0.4	0.1	0.012
13	UNKNOWN	21.17	0.00	0.2	0.0	0.002
14	UNKNOWN	21.48	0.01	0.3	0.0	0.007
15	UNKNOWN	21.71	0.01	0.4	0.0	0.008
16	UNKNOWN	21.96	0.01	0.4	0.0	0.007
17	UNKNOWN	22.19	0.01	0.4	0.0	0.008
18	UNKNOWN	22.40	0.00	0.2	0.0	0.002
19	UNKNOWN	22.65	0.00	0.3	0.0	0.004
20	UNKNOWN	23.28	0.00	0.3	0.0	0.004
21	UNKNOWN	23.31	0.00	0.3	0.0	0.004
22	UNKNOWN	32.87	0.01	0.5	0.0	0.008
23	UNKNOWN	33.39	0.01	0.6	0.1	0.012
24	UNKNOWN	34.49	0.02	0.9	0.1	0.025
25	UNKNOWN	34.80	0.03	1.0	0.2	0.027
26	UNKNOWN	35.12	0.03	1.0	0.2	0.034
27	UNKNOWN	35.48	0.08	2.2	0.4	0.078
28	UNKNOWN	35.87	0.21	5.3	1.2	0.207
29	UNKNOWN	37.23	95.87	937.4	544.3	95.873
30	UNKNOWN	39.53	0.02	0.6	0.1	0.020
31	UNKNOWN	39.77	0.02	0.8	0.1	0.016
32	UNKNOWN	39.99	0.01	0.3	0.0	0.006
33	UNKNOWN	40.57	0.07	1.2	0.4	0.066
34	UNKNOWN	41.01	0.01	0.4	0.1	0.009
35	UNKNOWN	41.60	0.04	1.1	0.2	0.041

Chromatogram : Mata-176-SF51_channel1

System : HPLC-PDA
Method : Minneo
User : Daniel

Acquired : 06/03/2025 15:34:50
Processed : 06/03/2025 16:27:34
Printed : 06/03/2025 16:36:05

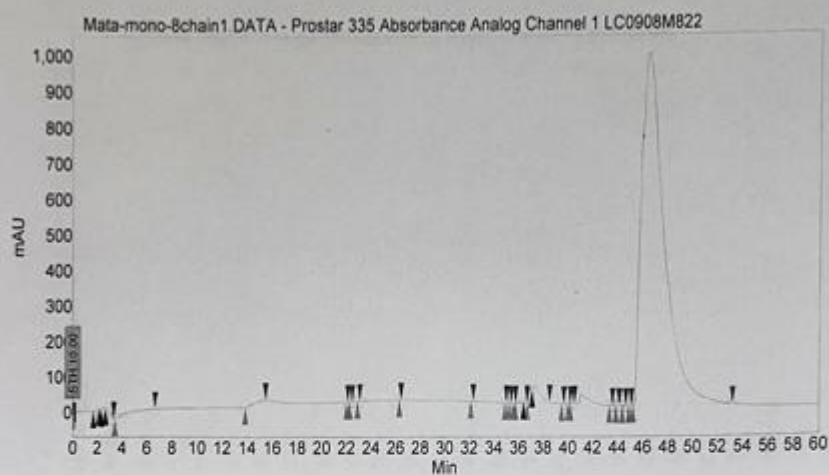
Index	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU Min]	Area % [%]
36	UNKNOWN	42.12	0.03	0.7	0.1	0.026
37	UNKNOWN	42.63	0.04	1.1	0.2	0.041
38	UNKNOWN	43.15	0.03	0.9	0.2	0.031
39	UNKNOWN	43.68	0.03	1.0	0.2	0.033
40	UNKNOWN	44.27	0.08	1.3	0.4	0.077
41	UNKNOWN	44.72	0.03	0.9	0.2	0.032
42	UNKNOWN	45.21	0.02	0.6	0.1	0.018
43	UNKNOWN	45.75	0.01	0.5	0.1	0.015
44	UNKNOWN	46.28	0.03	0.9	0.2	0.030
45	UNKNOWN	47.29	0.01	0.5	0.1	0.011
46	UNKNOWN	47.83	0.02	0.7	0.1	0.024
47	UNKNOWN	48.26	0.02	0.7	0.1	0.023
48	UNKNOWN	48.84	0.03	0.7	0.1	0.026
49	UNKNOWN	49.37	0.02	0.6	0.1	0.017
50	UNKNOWN	49.91	0.01	0.3	0.0	0.006
Total			100.00	984.4	551.9	100.000

Figure A3 – Chromatogram for 7.

Chromatogram : Mata-mono-8chain1_channel1

System : HPLC-PDA
Method : Minnet
User : Daniel

Acquired : 09/03/2025 17:10:59
Processed : 09/03/2025 18:29:32
Printed : 17/04/2025 11:35:13



Peak results :

Index	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU.Min]	Area % [%]
1	UNKNOWN	1.47	0.58	11.9	12.6	0.582
2	UNKNOWN	2.12	0.35	21.6	7.6	0.349
3	UNKNOWN	2.17	0.21	23.0	4.6	0.210
4	UNKNOWN	2.52	0.37	28.1	8.1	0.374
5	UNKNOWN	2.72	0.53	28.4	11.6	0.534
6	UNKNOWN	4.65	0.80	8.4	17.3	0.798
7	UNKNOWN	14.60	0.18	3.4	3.8	0.176
8	UNKNOWN	22.07	0.00	0.2	0.0	0.001
9	UNKNOWN	22.31	0.00	0.3	0.0	0.002
10	UNKNOWN	23.00	0.00	0.3	0.0	0.001
11	UNKNOWN	25.35	0.00	0.3	0.0	0.001
12	UNKNOWN	32.20	0.00	0.5	0.1	0.003
13	UNKNOWN	34.93	0.00	0.3	0.0	0.001
14	UNKNOWN	35.29	0.00	0.8	0.1	0.004
15	UNKNOWN	35.63	0.00	0.7	0.1	0.004
16	UNKNOWN	36.03	0.04	3.2	0.9	0.041
17	UNKNOWN	36.45	0.01	1.0	0.1	0.006
18	UNKNOWN	36.99	0.34	42.8	7.3	0.338
19	UNKNOWN	37.32	1.53	51.0	33.0	1.525
20	UNKNOWN	39.57	0.00	0.6	0.1	0.003
21	UNKNOWN	40.05	0.00	0.8	0.1	0.004
22	UNKNOWN	40.31	0.00	0.5	0.1	0.004
23	UNKNOWN	43.47	0.00	0.5	0.1	0.003
24	UNKNOWN	44.00	0.00	0.3	0.0	0.002
25	UNKNOWN	44.55	0.00	0.6	0.1	0.004
26	UNKNOWN	45.01	0.00	0.3	0.0	0.002
27	UNKNOWN	46.32	95.03	987.6	2058.3	95.029
Total			100.00	1217.5	2166.0	100.000

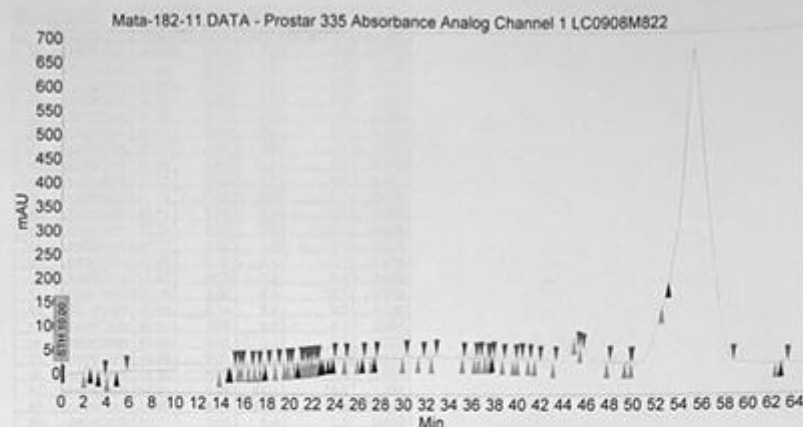
Page
1/5
PAGECOUNT

Figure A4 – Chromatogram for 12.

Chromatogram : Mata-182-11_channel1

System : HPLC-PDA
Method : Minneol
User : Daniel

Acquired : 07/03/2025 20:59:36
Processed : 07/03/2025 22:08:08
Printed : 07/03/2025 22:08:51



Peak results :

Index	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU Min]	Area % [%]
1	UNKNOWN	2.31	0.39	16.6	5.6	0.386
2	UNKNOWN	2.87	0.82	26.1	12.0	0.824
3	UNKNOWN	3.45	0.22	8.2	3.2	0.217
4	UNKNOWN	4.60	0.21	5.4	3.1	0.210
5	UNKNOWN	4.89	0.14	4.4	2.0	0.135
6	UNKNOWN	14.44	0.09	2.6	1.4	0.093
7	UNKNOWN	14.68	0.02	2.1	0.3	0.021
8	UNKNOWN	14.84	0.03	1.6	0.4	0.021
9	UNKNOWN	15.59	0.00	0.2	0.0	0.001
10	UNKNOWN	15.85	0.00	0.2	0.0	0.001
11	UNKNOWN	16.53	0.01	0.4	0.1	0.006
12	UNKNOWN	17.17	0.01	0.6	0.2	0.011
13	UNKNOWN	17.69	0.02	0.8	0.2	0.015
14	UNKNOWN	17.97	0.00	0.2	0.1	0.004
15	UNKNOWN	18.89	0.00	0.2	0.0	0.003
16	UNKNOWN	19.65	0.00	0.3	0.0	0.003
17	UNKNOWN	19.99	0.00	0.3	0.1	0.004
18	UNKNOWN	20.55	0.01	0.6	0.1	0.007
19	UNKNOWN	20.84	0.01	0.4	0.1	0.005
20	UNKNOWN	21.09	0.00	0.2	0.0	0.001
21	UNKNOWN	21.35	0.00	0.3	0.0	0.003
22	UNKNOWN	21.63	0.00	0.2	0.0	0.002
23	UNKNOWN	21.87	0.00	0.2	0.0	0.002
24	UNKNOWN	22.13	0.00	0.2	0.0	0.001
25	UNKNOWN	22.36	0.00	0.2	0.0	0.002
26	UNKNOWN	22.59	0.00	0.3	0.0	0.002
27	UNKNOWN	22.83	0.01	0.5	0.1	0.005
28	UNKNOWN	23.23	0.02	0.9	0.3	0.021
29	UNKNOWN	23.35	0.02	0.7	0.2	0.017
30	UNKNOWN	23.63	0.00	0.5	0.1	0.004
31	UNKNOWN	24.86	0.00	0.3	0.0	0.002
32	UNKNOWN	26.15	0.01	0.3	0.1	0.006
33	UNKNOWN	26.44	0.00	0.3	0.0	0.002
34	UNKNOWN	27.36	0.01	0.7	0.1	0.009
35	UNKNOWN	27.48	0.01	0.6	0.1	0.008

Page
1/3
PAGECOUNT

Chromatogram : Mata-182-11_channel1

System : HPLC-PDA
Method : Mnned
User : Daniel

Acquired : 07/03/2025 20:59:36
Processed : 07/03/2025 22:08:08
Printed : 08/03/2025 19:10:19

Index	Name	Time (Min)	Quantity (% Area)	Height (mAU)	Area (mAU Min)	Area % (%)
36	UNKNOWN	30.09	0.02	1.5	0.3	0.020
37	UNKNOWN	31.52	0.02	1.2	0.4	0.025
38	UNKNOWN	32.68	0.04	1.9	0.5	0.037
39	UNKNOWN	35.21	0.00	0.4	0.0	0.003
40	UNKNOWN	36.28	0.00	0.4	0.0	0.003
41	UNKNOWN	36.69	0.00	0.6	0.1	0.004
42	UNKNOWN	37.17	0.01	0.9	0.1	0.010
43	UNKNOWN	37.49	0.00	0.1	0.0	0.000
44	UNKNOWN	37.67	0.02	1.7	0.2	0.014
45	UNKNOWN	37.73	0.01	0.9	0.1	0.007
46	UNKNOWN	38.72	0.00	0.6	0.1	0.005
47	UNKNOWN	39.80	0.00	0.5	0.1	0.004
48	UNKNOWN	40.17	0.02	1.2	0.3	0.019
49	UNKNOWN	40.99	0.00	0.4	0.1	0.004
50	UNKNOWN	41.63	0.03	1.3	0.4	0.029
51	UNKNOWN	43.16	0.00	0.3	0.0	0.003
52	UNKNOWN	45.03	0.14	9.6	2.0	0.135
53	UNKNOWN	45.51	0.02	2.8	0.3	0.023
54	UNKNOWN	47.83	0.01	1.1	0.2	0.013
55	UNKNOWN	49.45	0.01	0.5	0.2	0.012
56	UNKNOWN	49.86	0.00	0.1	0.0	0.000
57	UNKNOWN	52.91	1.09	62.7	16.0	1.084
58	UNKNOWN	54.89	96.45	583.6	1492.3	96.446
59	UNKNOWN	62.71	0.01	0.6	0.2	0.010
60	UNKNOWN	62.95	0.02	0.7	0.3	0.019
Total			100.00	753.5	1461.2	100.000

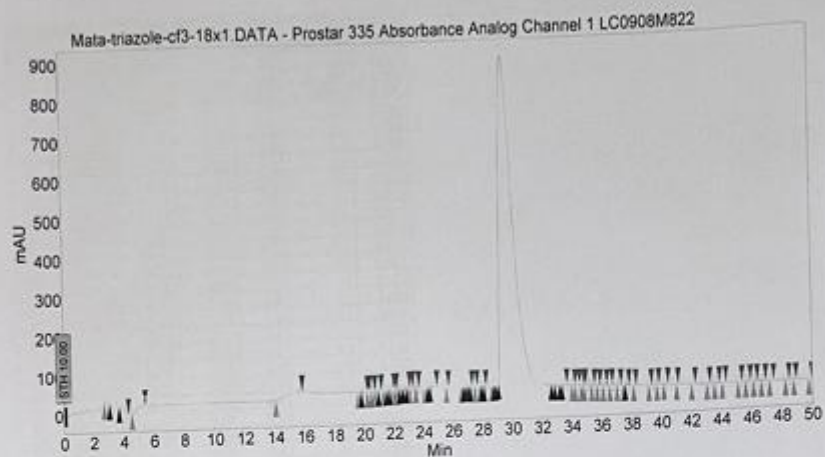
Page
2/5
PAGECOUNT

Figure A5 – Chromatogram for 21.

Chromatogram : Mata-triazole-cf3-18x1_channel1

System : HPLC-PDA
Method : Minneo
User : Daniel

Acquired : 09/03/2025 15:00:20
Processed : 09/03/2025 15:53:04
Printed : 09/03/2025 16:01:05



Peak results :

Index	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU Min]	Area % [%]
1	UNKNOWN	2.85	0.20	8.4	1.9	0.203
2	UNKNOWN	3.27	1.01	22.9	9.3	1.006
3	UNKNOWN	3.88	0.52	12.2	4.8	0.517
4	UNKNOWN	4.96	0.33	6.2	3.1	0.333
5	UNKNOWN	15.11	0.58	4.4	5.2	0.564
6	UNKNOWN	19.64	0.01	0.5	0.1	0.006
7	UNKNOWN	19.95	0.01	0.5	0.1	0.008
8	UNKNOWN	20.23	0.00	0.2	0.0	0.002
9	UNKNOWN	20.55	0.00	0.4	0.0	0.005
10	UNKNOWN	20.83	0.01	0.5	0.1	0.007
11	UNKNOWN	21.07	0.00	0.3	0.0	0.004
12	UNKNOWN	21.32	0.01	0.5	0.1	0.008
13	UNKNOWN	21.60	0.01	0.5	0.1	0.007
14	UNKNOWN	21.83	0.01	0.5	0.0	0.003
15	UNKNOWN	22.05	0.00	0.4	0.0	0.004
16	UNKNOWN	22.28	0.00	0.4	0.0	0.006
17	UNKNOWN	22.52	0.01	0.6	0.1	0.009
18	UNKNOWN	22.73	0.01	0.6	0.1	0.009
19	UNKNOWN	22.95	0.01	0.5	0.1	0.008
20	UNKNOWN	23.12	0.01	0.5	0.1	0.007
21	UNKNOWN	23.57	0.01	0.6	0.1	0.006
22	UNKNOWN	24.19	0.01	0.4	0.1	0.006
23	UNKNOWN	24.35	0.01	0.6	0.1	0.008
24	UNKNOWN	24.61	0.08	2.6	0.7	0.075
25	UNKNOWN	25.59	0.00	0.4	0.0	0.003
26	UNKNOWN	25.60	0.01	0.5	0.0	0.005
27	UNKNOWN	26.75	0.01	0.6	0.1	0.008
28	UNKNOWN	26.92	0.01	0.6	0.1	0.007
29	UNKNOWN	27.05	0.00	0.5	0.0	0.005
30	UNKNOWN	27.21	0.00	0.4	0.0	0.003
31	UNKNOWN	27.52	0.01	0.5	0.0	0.005
32	UNKNOWN	27.84	0.00	0.3	0.0	0.002
33	UNKNOWN	27.97	0.01	0.6	0.1	0.006
34	UNKNOWN	28.13	0.00	0.3	0.0	0.003
35	UNKNOWN	28.69	0.01	0.7	0.1	0.008

Page
1/5
PAGECOUNT

Chromatogram : Mata-triazole-cf3-18x1_channel1

System : HPLC-PDA
Method : Minnecl
User : Daniel

Acquired : 09/03/2025 15:00:20
Processed : 09/03/2025 15:53:04
Printed : 09/03/2025 16:01:05

Index	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU Min]	Area % [%]
36	UNKNOWN	28.83	0.01	1.2	0.1	0.013
37	UNKNOWN	29.03	0.06	3.8	0.5	0.056
38	UNKNOWN	29.59	96.77	849.3	889.9	96.769
39	UNKNOWN	32.71	0.02	1.1	0.2	0.022
40	UNKNOWN	32.96	0.02	1.2	0.2	0.024
41	UNKNOWN	33.23	0.01	0.6	0.1	0.008
42	UNKNOWN	33.49	0.02	1.3	0.2	0.024
43	UNKNOWN	34.09	0.01	0.6	0.1	0.008
44	UNKNOWN	34.39	0.01	0.8	0.1	0.010
45	UNKNOWN	34.71	0.01	0.6	0.1	0.007
46	UNKNOWN	34.87	0.00	0.2	0.0	0.001
47	UNKNOWN	35.41	0.01	0.9	0.1	0.011
48	UNKNOWN	35.77	0.01	0.9	0.1	0.011
49	UNKNOWN	36.20	0.01	1.1	0.1	0.013
50	UNKNOWN	36.63	0.01	0.8	0.1	0.009
51	UNKNOWN	37.13	0.02	1.2	0.1	0.015
52	UNKNOWN	37.52	0.01	0.6	0.0	0.005
53	UNKNOWN	37.63	0.02	1.5	0.2	0.024
54	UNKNOWN	38.16	0.00	0.4	0.0	0.003
55	UNKNOWN	39.21	0.01	0.5	0.1	0.007
56	UNKNOWN	39.76	0.00	0.4	0.0	0.002
57	UNKNOWN	40.29	0.01	0.9	0.1	0.011
58	UNKNOWN	41.04	0.00	0.2	0.0	0.001
59	UNKNOWN	42.13	0.01	0.4	0.1	0.006
60	UNKNOWN	43.16	0.00	0.2	0.0	0.003
61	UNKNOWN	43.67	0.02	0.7	0.1	0.015
62	UNKNOWN	44.16	0.01	0.6	0.1	0.008
63	UNKNOWN	45.24	0.01	0.4	0.1	0.007
64	UNKNOWN	45.73	0.01	0.5	0.1	0.009
65	UNKNOWN	46.24	0.01	0.4	0.1	0.006
66	UNKNOWN	46.75	0.01	0.5	0.1	0.009
67	UNKNOWN	47.32	0.01	0.4	0.1	0.009
68	UNKNOWN	48.39	0.00	0.3	0.0	0.005
69	UNKNOWN	48.85	0.00	0.2	0.0	0.001
70	UNKNOWN	49.88	0.01	0.4	0.0	0.006
Total			100.00	945.7	919.6	100.000

Page
2/5
PAGECOUNT

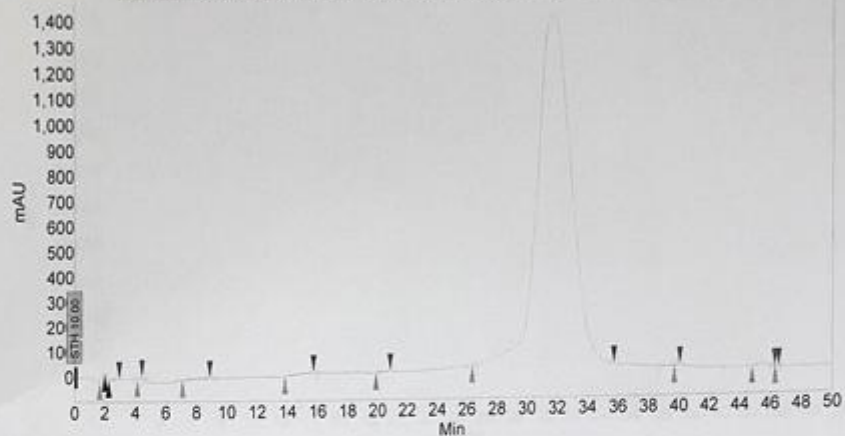
Figure A6 – Chromatogram for 13.

Chromatogram : Mata-t-iPr1_channel1

System : HPLC-PDA
Method : Minnece
User : Daniel

Acquired : 08/03/2025 18:14:30
Processed : 08/03/2025 19:07:14
Printed : 08/03/2025 19:09:50

Mata-t-iPr1.DAT - Prostar 335 Absorbance Analog Channel 1 LC0908M822



Peak results :

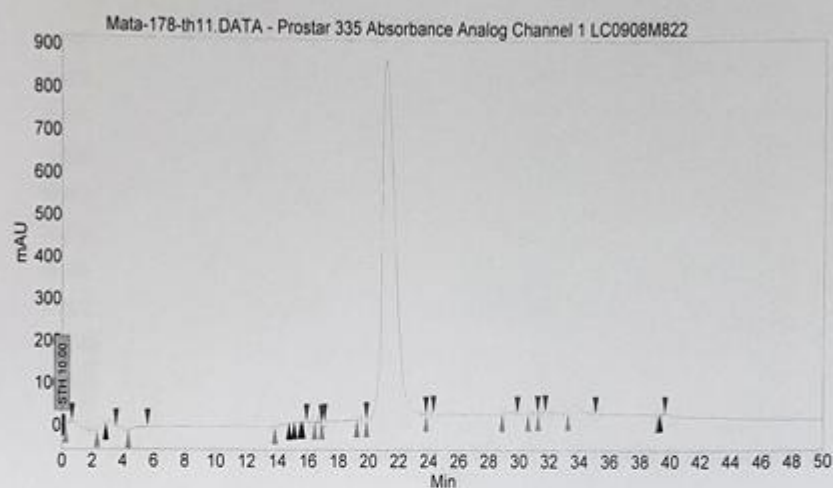
Index	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU Min]	Area % [%]
1	UNKNOWN	1.89	0.22	30.0	2.8	0.220
2	UNKNOWN	2.01	0.10	27.0	3.4	0.095
3	UNKNOWN	2.60	0.04	3.1	1.5	0.042
4	UNKNOWN	4.28	0.01	2.5	0.4	0.013
5	UNKNOWN	7.64	0.08	2.4	2.7	0.077
6	UNKNOWN	14.40	0.10	2.6	3.4	0.096
7	UNKNOWN	20.37	0.04	2.2	1.2	0.035
8	UNKNOWN	31.61	99.21	1382.1	3522.6	99.206
9	UNKNOWN	39.87	0.01	2.8	0.5	0.014
10	UNKNOWN	45.36	0.20	10.3	7.1	0.201
11	UNKNOWN	46.45	0.00	0.6	0.1	0.002
Total			100.00	1475.8	3550.8	100.000

Figure A7 – Chromatogram for 14.

Chromatogram : Mata-178-th11_channel1

System : HPLC-PDA
Method : Minnecei
User : Daniel

Acquired : 07/03/2025 17:33:02
Processed : 07/03/2025 18:25:46
Printed : 08/03/2025 19:11:10



Peak results :

Index	Name	Time [Min]	Quantity [% Area]	Height [mAU]	Area [mAU.Min]	Area % [%]
1	UNKNOWN	0.45	0.07	2.4	0.6	0.063
2	UNKNOWN	2.79	0.50	13.9	4.4	0.504
3	UNKNOWN	3.13	0.86	17.9	7.5	0.859
4	UNKNOWN	4.84	0.30	3.8	2.6	0.299
5	UNKNOWN	14.55	0.28	3.8	2.5	0.280
6	UNKNOWN	14.75	0.12	3.4	1.1	0.121
7	UNKNOWN	15.11	0.11	2.9	1.0	0.112
8	UNKNOWN	15.53	0.02	1.8	0.2	0.023
9	UNKNOWN	15.63	0.02	1.0	0.2	0.018
10	UNKNOWN	16.73	0.02	0.6	0.2	0.022
11	UNKNOWN	17.03	0.00	0.2	0.0	0.002
12	UNKNOWN	19.35	0.01	0.4	0.1	0.012
13	UNKNOWN	21.09	95.32	855.6	833.6	95.318
14	UNKNOWN	24.12	0.02	0.5	0.1	0.015
15	UNKNOWN	29.41	0.30	4.6	2.6	0.297
16	UNKNOWN	30.80	0.04	1.1	0.4	0.042
17	UNKNOWN	31.47	0.02	0.5	0.2	0.019
18	UNKNOWN	33.64	1.96	22.3	17.1	1.955
19	UNKNOWN	39.27	0.01	0.7	0.1	0.008
20	UNKNOWN	39.48	0.03	1.3	0.2	0.028
Total			100.00	938.8	874.5	100.000

Page
1/5
PAGECOUNT

Figure A7 – Chromatogram for 16.