



Doctoral Thesis

**Design, Synthesis, and Biological Evaluation of
Chemical Proteomic Probes for MoA Studies in *C.
neoformans*, and of PROTAC Compounds for the
Treatment of Metastatic Breast Cancer**

Thesis submitted in accordance with the requirements of the
University of Liverpool for the degree of Doctor in Philosophy

By

Rachel Crick

January 2021

Supervisors: Dr G Nixon and Prof. P O'Neill

Declaration

I declare that this thesis has been composed by myself. The results are my own unless explicitly stated otherwise. This work has not been submitted, in whole or in part, for any other degree or professional qualification.

This research was carried out at the University of Liverpool, Department of Chemistry.

Biological testing for the *C. neoformans* project conducted by Professor William Hopes group at the Institute of Translational Biology, University of Liverpool. *In vitro* and *in vivo* work performed by Dr Suzy Gore and Ms Christine Natal. DMPK data obtained by AstraZeneca.

Biological testing for the metastatic breast cancer project conducted by Professor Philip Rudlands group at the Institute of Integrative Biology, University of Liverpool. *In vitro* and *in vivo* work performed by Dr Min Du, Dr Thamir Ismail, and Dr Guozheng Wang. DMPK data obtained by Cyprotex.

Acknowledgments

Firstly, I would like to thank my supervisors Dr Gemma Nixon and Professor Paul O'Neill for giving me the opportunity to do my PhD, and for all the help and advice over the years concerning my research and the writing of this thesis. I would also like to thank my coassessors Professor Neil Berry and Dr Andrew Carnell for overseeing my research during the beginning of my PhD. A big thank you goes to Professor William Hope and his biology team over at the Institute of Translational Biology, especially Dr Suzy Gore for performing both *in vitro* and *in vivo* experiments during the *C. neoformans* project. A huge thank you also goes to Professor Philip Rudland for allowing me to work on the triple negative breast cancer project. I would also like to express my gratitude to the rest of his team over at the Institute of Integrative Biology, particularly Dr Min Du, Dr Thamir Ismail, and Dr Guozheng Wang for gathering both *in vitro* and *in vivo* data and explaining the biological concepts to me.

A big thank you goes to all the past and present members of the fourth-floor lab, for providing advice, making me laugh, continuous AJ trips, and allowing me to pinch the occasional reagent and solvent bottle. I can honestly say that I am grateful for meeting and getting to know all of you.

Finally, I would like to thank the most important people in my life, my family. My biggest supporters are my Mum and Dad, who have loved and supported me throughout my life, never stopped believing in me, and taught me to be strong and independent. Without a doubt, they are the most incredible and lovely people I have ever met, and I am grateful that they raised me. I would also like to give a massive thank you to my best friend and boyfriend Andy who has encouraged me over the years, loved me, and helped me communicate and build my confidence. I would like to give a huge thank you to my big brother John, sister-in-law Sophie, and my two beautiful nephews Ethan and Nathan for supporting me and being such a huge positive impact in my life. Last but not least, I would like to express my love for my cat Molly, my pet fish, Two-toned Tim, Orange fish, and of course Barney. Due to how amazing my family are, I would like to dedicate this thesis to them. I could not have done it without any of them. *Hoyvin-Glavin!*

Abstract

Proteomics has revolutionised the way that proteins are studied and has contributed to major leaps in drug discovery. Identification and understanding of therapeutic protein targets have led to the development of novel therapeutic agents. There are many approaches to studying proteins, however in this thesis, we explore two distinct techniques for the use in drug discovery; chemical probes for target elucidation within *Cryptococcus neoformans*, and the exploitation of cellular protein processes for the design of novel drugs to treat triple negative breast cancer (TNBC).

Cryptococcus neoformans is an opportunistic fungal pathogen that can cause cryptococcal meningoencephalitis, and to a lesser extent pneumonia in immunocompromised individuals. The number of cases of *C. neoformans* related illnesses has risen over the years due to an increase in the occurrence of HIV/AIDS; as a result, the fungal infection has become particularly serious in sub-Saharan Africa, causing approximately 720,000 cases and 504,000 deaths per annum in this region alone. The gold standard for treating *C. neoformans* is the initial combination of amphotericin B and flucytosine, followed by a course of fluconazole as maintenance therapy. Demand to develop a new therapeutic agent to treat *C. neoformans* has risen as current treatments are associated with resistance and toxicity.

Flubendazole among other benzimidazole based anthelmintic agents have been reported to possess antifungal activity against *C. neoformans*. Flubendazole possesses an MIC value of 0.125 mg/L and a CFU log drop of 3.7 when dosed at 150 mg/kg in an *in vivo* study. Unfortunately, flubendazole has poor aqueous solubility at 0.8 μM , which limits its clinical use. Extensive structural activity relationship (SAR) information was gathered during this project to further improve the antifungal activity, as well as improving the DMPK profile of the drug. A morpholine ether derivative of flubendazole (**1h**) became the lead compound of the project due to increasing the aqueous solubility to 10 μM , while possessing an MIC value of 0.25 mg/L and a log drop of 3 when dosed at 150 mg/kg during an *in vivo* study. Continued SAR exploration was conducted to further improve the solubility profile, where the morpholine functionality was altered with closely related heterocycles and bioisosteres.

The disruption of mitotic spindle by binding to the β -tubulin subunit of microtubules is believed to be the primary mode of action for benzimidazoles. Both fungal and eukaryotic

cells possess mitotic spindle, increasing the possibility of non-selective binding and associated toxicity. Identification of distinctive binding between the human and the fungal β -tubulin will aid in the design and synthesis of selective benzimidazole analogues. Photoaffinity labelling (PAL) is a proteomic screening tool which can be utilised to identify the target/s of benzimidazoles as well as ascertain whether the binding in human and *C. neoformans* β -tubulin differs. PAL studies can also confirm whether any alternative targets such as enzymes bind to benzimidazoles, which may be therapeutically exploited. Diazirine and benzophenone-based PAL probes have been synthesised, however issues attributed to poor cell penetration has led to unsuccessful identification of binding targets.

Triple negative breast cancer (TNBC) accounts for around 15% of breast cancer cases and has a poor survival rate due to the high metastatic spread. TNBC is a virtually incurable disease and current treatments are limited to chemotherapies which have poor success rates and many adverse drug reactions, such as hair loss, cardiotoxicity, neutropenia, and myelotoxicity. With these limitations in mind, new targeted therapies are required to improve the survival rate of people with TNBC.

The S100A4 protein is a promising target for new selective therapies. The expression and upregulation of this protein is linked to the initiation of metastasis of a number of tumours including breast cancer. As such, the inhibition of S100A4 expression or function could be manipulated therapeutically to reduce the invasiveness of S100A4 expressing cancers.

The PROTAC approach to drug design is a potential way to inhibit or degrade the metastasis inducing protein (MIP) S100A4. PROTAC compounds are heterobifunctional molecules that consist of two small chemical “warheads” joined by a linker with the aim of selective protein degradation. One warhead binds to a component of the ubiquitin proteasome system (UPS) which initiates selective protein degradation, and the other warhead will bind to the target S100A4 protein. Thalidomide tethers have been shown to successfully bind to the UPS and initiate selective degradation of a number of proteins. The second warhead will be an inhibitor known as **US10113**. **US10113** has been shown to reduce tumour metastasis *via* the inhibition of S100A4. The **US10113** derived warhead can be connected to the thalidomide warhead *via* a linker to yield a PROTAC compound which can selectively degrade the S100A4 protein.

The lead PROTAC (**RGC**) can inhibit tumour metastasis and migration at nM concentrations, with complete substrate specificity and no observable toxicity up to 100 μ M, a 10,000-fold increase in activity from the initial **US10113**.

Additional inhibitors and PROTAC agents have been synthesised to improve the pharmacokinetic and pharmacodynamic profile, along with improving our understanding of the structure activity relationships with the compound. A number of these analogues have surpassed the original **US10113** in terms of metastatic inhibition and metabolic profile. As a result, several new PROTAC compounds have been synthesised with potentially promising results.

Publications

In Progress

Synthesis and Biological Assessment of the Benzisothiazolinone Class for Treatment of the Fungal Infection *Cryptococcus neoformans* - Draft written - to be submitted to Med Chem Comm – August 2020.

G. Washbourn, R. Crick, L. McEntee, A. Johnson, N. Farrington, S. Whalley, J. Livermore, C. Natal, S. Gore, J. Bibby, N. Berry, I. Charles, W. Hope, G. L. Nixon.

Design, Synthesis and Biological Evaluation of Benzimidazoles for Treatment of the Fungal Infection *Cryptococcus neoformans* - Draft in progress - to be submitted to J Med Chem – September 2020.

G. Washbourn, R. Crick, L. McEntee, A. Johnson, N. Farrington, S. Whalley, J. Livermore, C. Natal, S. Gore, J. Bibby, N. Berry, I. Charles, W. Hope, G. L. Nixon.

Identification and Profiling of Benzimidazole Ether Derivatives; Towards a Treatment for the Fungal Infection *Cryptococcus neoformans* - Draft in progress - to be submitted to J Med Chem – September 2020.

R. Crick, G. Washbourn, L. McEntee, A. Johnson, N. Farrington, S. Whalley, J. Livermore, C. Natal, S. Gore, J. Bibby, N. Berry, I. Charles, W. Hope, G. L. Nixon.

Proteolysis-targeting chimera (PROTAC) compounds to degrade S100A4 and inhibit breast cancer metastasis- Draft in progress.

R. Crick, S. Pate, T. Gibson, H. Zhu, R. Lloyd-Hughes, M. Du, G. Wang, T. M. Ismail, R. Barraclough, P. S. Rudland, G. L. Nixon.

Contents

Declaration.....	2
Acknowledgments.....	3
Abstract.....	4
Publications.....	7
Contents.....	8
Abbreviations.....	15
CHAPTER 1	22
1.1 <i>Cryptococcus neoformans</i>	23
1.1.1 What is <i>Cryptococcus neoformans</i> ?	23
1.1.2 Aetiology of <i>C. neoformans</i>	23
1.1.3 Epidemiology of <i>C. neoformans</i>	24
1.1.4 Pathogenesis of <i>C. neoformans</i>	26
1.1.5 Clinical Manifestations.....	27
1.1.5.1 Pulmonary Infection.....	27
1.1.5.2 CNS Infection.....	27
1.1.5.3 Skin Infection	28
1.1.6 Current Treatment for <i>C. neoformans</i>	28
1.1.6.1 Amphotericin B	29
1.1.6.1.1 Mode of Action of Amphotericin B.....	29
1.1.6.1.2 Limitation of Amphotericin B.....	31
1.1.6.2 Flucytosine	32
1.1.6.2.1 Mode of Action of Flucytosine	32
1.1.6.2.2 Limitations of Flucytosine	33
1.1.6.3 Fluconazole	34
1.1.6.3.1 Mode of action of fluconazole	35
1.1.6.3.2 Limitations of Fluconazole	35
1.1.7 Development of the ideal anti-cryptococcal drug	36
1.2 Photoaffinity Labelling Probes (PAL).....	37
1.2.1 Binding site determination	37
1.2.2 PAL probe design	37
1.2.2.1 Photo reactive group	38

1.2.2.1.1 Phenyl azides.....	39
1.2.2.1.2 Benzophenones.....	40
1.2.2.1.3 Diazirines.....	40
1.2.2.2 Binding group/ linker	42
1.2.2.3 Reporter tag.....	42
1.2.2.3.1 Fluorophore and biotin tags	42
1.2.2.3.2 Click chemistry probes	43
1.2.3 Analytical platforms for ABPP	45
1.2.3.1 Gel electrophoresis	45
1.2.3.2 Liquid chromatography-mass spectrometry (LC-MS)	45
1.3 Benzimidazoles.....	47
1.3.1 Chemistry of Benzimidazoles	47
1.3.1.1 Tautomerism and Isomerism	47
1.3.2 Biological Activity of Benzimidazoles.....	48
1.3.2.1 Tubulin and Microtubules.....	48
1.3.2.1.1 Role of Microtubules.....	49
1.3.2.1.2 Inhibition of Microtubules	49
1.3.2.2 Current Uses of Benzimidazoles	50
1.3.2.3 Anti-cryptococcal Activity of Benzimidazoles	52
1.3.2.3.1 Benzimidazole Cryptococcal Data.....	53
1.3.3 Identification of Benzimidazole Targets	56
1.3.3.1 Previous Established β -Tubulin Probes.....	56
1.3.3.1.1 Colchicine	56
1.3.3.1.2 Phenylahistin.....	57
1.3.3.1.3 Paclitaxel.....	57
1.3.3.1.4 Vinblastine	58
1.3.3.1.5 Discodermolide	59
1.3.3.2 Alternative Benzimidazole Targets	60
1.4 References	62
CHAPTER 2	70
2.1 Benzimidazoles as Antifungals	71
2.1.1 Metabolism of Benzimidazoles	72
2.1.2 Aqueous Solubility of Benzimidazoles	75
2.1.3 Previous SAR Work.....	77

2.1.3.1 MIC and DMPK Data of Benzimidazole Analogues	77
2.1.3.2 <i>In Vivo</i> Data of Benzimidazole Analogues.....	80
2.2 Aims.....	85
2.3 Discussion.....	85
2.3.1 Synthesis	85
2.3.2 Results.....	93
2.3.2.1 MIC Data	93
2.3.2.2 DMPK Data.....	95
2.4 Benzimidazole Ketone Derivative	96
2.5 Discussion.....	96
2.5.1 Synthesis	96
2.5.2.1 MIC Data	106
2.5.2.2 DMPK Data.....	106
2.6 Salt Formulations	107
2.6.1 Introduction to <i>In Vivo</i> Data.....	107
2.6.2 Introduction to Salt Formulation	109
2.7 Aims.....	111
2.8 Discussion.....	112
2.8.1 Synthesis	112
2.8.1.1 PTSA Salt Formation.....	112
2.8.1.2 HCl Salt Formation	113
2.8.3 Solubility Tests	113
2.8.4 Results.....	114
2.8.4.1 <i>In Vivo</i> Data	114
2.8.4.2 DMPK Data.....	114
2.9 Metabolites.....	114
2.9.1 Concept.....	114
2.9.1.1 Morpholine	115
2.9.1.1.1 Metabolism and Bioactivation	115
2.10 Aims.....	117
2.11 Discussion.....	117
2.11.1 Synthesis	117
2.11.1.1 Carbamate Hydrolysis Metabolite	117
2.11.1.2 <i>N</i> -Oxide Metabolite	118

2.11.2 Results.....	118
2.11.2.1 MIC Data	118
2.11.2.2 DMPK Data.....	119
2.12 Conclusion.....	120
2.13 Future Work.....	121
2.14 Experimental.....	Error! Bookmark not defined.
2.14.1 General Experimental Details	124
2.14.1.1 Chemicals	124
2.14.1.2 Thin Layer Chromatography	124
2.14.1.3 Column Chromatography.....	124
2.14.1.4 Compound Characterisation	124
2.14.2 General Procedures	125
2.14.2.1 SAR Morpholine Derivatives	127
2.14.2.2 Benzimidazole Ketone Synthesis	143
2.14.2.3 Salt Synthesis	147
2.14.2.4 Benzimidazole Metabolite Synthesis	148
2.15 References	150
CHAPTER 3	154
3.1 Elucidation of Target Binding.....	154
3.1.1 Benzimidazole SAR Analysis.....	156
3.1.2 Design of Benzimidazole PAL Probe.....	158
3.2 Aims.....	160
3.3 Discussion.....	161
3.3.1 Synthesis of Benzimidazole Carbamate Alkyne Probes	161
3.3.1.1 Synthesis of Prop-2-yn-1-yl (5-phenoxy/thio-1 <i>H</i> -benzo[<i>d</i>]imidazol-2-yl)carbamate Analogues.....	161
3.3.1.6 Synthesis of 3-Trifluoro-3-Phenyldiazirine Probe	173
3.3.1.7 Results of 3-Trifluoro-3-Phenyldiazirine Probe.....	179
3.3.1.8 Synthesis of Minimalist Photo-Crosslinker	179
3.3.1.8.1 Introduction of the Minimalist Photo-Crosslinker	179
3.3.1.8.2 Synthesis of Methyl (5-(2-(3-(but-3-yn-1-yl)-3 <i>H</i> -diazirin-3-yl)ethoxy)-1 <i>H</i> -benzo[<i>d</i>]imidazol-2-yl)carbamate.....	182
3.3.1.9 Synthesis of Benzophenone Probe	186
3.3.1.9.1 Introduction to benzophenones	186

3.3.1.9.2 Synthesis of Prop-2-yn-1-yl (5-(4-fluorobenzoyl)-1 <i>H</i> -benzo[<i>d</i>]imidazol-2-yl)carbamate	188
3.3.1.9.3 Results for Benzophenone Probe	189
3.4 Conclusions	190
3.5 Future Work	191
3.5.1 Synthesis of Phenyl Azide PAL Probe	191
3.5.2 Photolytic Studies	191
3.5.3 Alternative Scaffold Exploration	191
3.6 Experimental	193
3.6.1 General Experimental Details	193
3.6.1.1 Diazirine Synthesis	193
3.6.1.2 General procedures	193
3.6.2 Benzimidazole Carbamate Alkyne Probes	195
3.6.2.1 Prop-2-yn-1-yl (5-phenoxy/ thio-1 <i>H</i> -benzo[<i>d</i>]imidazol-2-yl)carbamate Analogues	195
3.6.2.2 5-((4-(Prop-2-yn-1-yloxy)phenoxy)-1 <i>H</i> -benzo[<i>d</i>]imidazol-2-yl)carbamate Analogues	214
3.6.2.3 5-((4-(Prop-2-yn-1-yloxy)phenyl)thio)-1 <i>H</i> -benzo[<i>d</i>]imidazol-2-yl)carbamate Analogues	220
3.6.2.4 Synthesis of 3-Trifluoro-3-Phenyldiazirine Probe	223
3.6.2.5 Synthesis of Minimalist Photo-Crosslinker	229
3.6.2.5.1 Carbendazim based aliphatic PAL probe	229
3.6.2.5.2 Bisaryl ether based aliphatic PAL probe	233
3.6.2.6 Synthesis of Benzophenone Probe	236
3.7 References	239
CHAPTER 4	244
4.1 Introduction to Breast Cancer	245
4.1.1 Risk Factors	245
4.1.2 Classification	248
4.2.1 Oestrogen Receptor Positive and Progesterone Receptor Positive Cancer	248
4.2.1.1 Treatment of Oestrogen Receptor Positive and Progesterone Receptor Positive Cancer	248
4.2.2 HER2 Positive Cancer	250
4.2.2.1 Treatment of HER2 Positive Cancer	251
4.2.3 Triple Negative Breast Cancer	253
4.2.3.1 Closer Examination into TNBC	253
4.2.3.2 Current Treatment of TNBC	254

4.2.3.3 Treatment Regimens.....	257
4.2.3.4 Issues with the Current Treatment of TNBC.....	258
4.2.3.5 New Targeted Therapy for TNBC.....	262
4.2 S100 Protein.....	263
4.2.1 S100A4 Protein.....	263
4.2.1.1 Targeting the S100A4 Protein.....	265
4.3 Targeting Proteins.....	268
4.3.1 Limitations of Small Molecule Inhibitors.....	268
4.3.2 Alternative Approaches.....	269
4.4 Introduction to Protein Degradation.....	270
4.4.1 The Ubiquitin Proteasome System.....	270
4.4.1.1 E3 Ligases in Detail.....	272
4.4.2 Proteolysis Targeting Chimera (PROTAC).....	273
4.4.2.1 The Development of PROTAC agents.....	274
4.5 Conclusions.....	283
4.6 References.....	284
CHAPTER 5.....	296
5.1 Discovery of US10113 Inhibitor.....	297
5.2 Aims.....	300
5.3 Development of a PROTAC agent.....	301
5.3.1 Discussion.....	301
5.3.1.1 Synthesis.....	301
5.3.1.1.1 Synthesis of Inhibitor.....	302
5.3.1.1.2 Synthesis of Amide Tether and PROTAC Agent.....	303
5.3.1.2 Biology.....	307
5.3.1.2.1 Cell Line Classification.....	307
5.3.1.2.2 Drosophila Data.....	308
5.3.1.2.3 Rat Cell Line Data.....	311
5.3.1.2.4 Human Cell Line Data.....	316
5.3.1.2.5 Preliminary <i>In Vivo</i> Data.....	320
5.3.1.2.6 DMPK Data.....	324
5.4 US10113 SAR Exploration.....	326
5.4.1 Metabolism Issues.....	326
5.4.2 Discussion.....	328

5.4.2.1 Synthesis of unsaturated US analogues.....	328
5.4.2.2 Synthesis of saturated US analogues.....	333
5.5 Newly synthesised PROTAC agents.....	337
5.5.1 Biology.....	338
5.6 Conclusion.....	339
5.7 Future Work.....	340
5.7.1 Inhibitor.....	340
5.7.2 Tether.....	341
5.7.3 E3 Ligase Ligand.....	342
5.7.4 Development of Chemical Probes.....	343
5.8 Experimental.....	345
5.8.1 General Experimental Details.....	345
5.8.1.1 HPLC.....	345
5.8.2 General Procedures.....	345
5.8.3 US Derivatives.....	346
5.8.4 Synthesis of Thalidomide Warhead and Amide Linker.....	360
5.8.5 US Carboxylic Acids.....	362
5.8.6 Perfluorophenyl US analogues.....	368
5.8.7 PROTAC agents.....	373
5.9 References.....	381

Abbreviations

1D	One-dimensional
2D	Two-dimensional
5-FC	5-Fluorocytosine
5-FdUMP	5-Fluorodeoxyuridine monophosphate
5-FU	5-Fluorouracil
5-FUMP	5-Fluorouridine monophosphate
AAPC	Average annual percentage change
ABC	ATP binding cassette
ABPP	Activity based protein profiling
ADR	Adverse drug reaction
AIDS	Acquired Immune Deficiency Syndrome
Ala	Alanine
AmB	Amphotericin B
AR	Androgen receptor
Arg	Arginine
ASIR	Age standardised incidence rate
ASMR	Age standardised mortality rate
ASON	Albendazole sulfone
ASOX	Albendazole sulfoxide
Asp	Aspartic Acid
ASPP	Active site protein profiling
ATP	Adenosine triphosphate
BBB	Blood brain barrier
Boc	<i>Tert</i> -Butyloxycarbonyl
BRCA	Breast cancer gene
<i>C. Gatti</i>	<i>Cryptococcus gatti</i>
<i>C. neoformans</i>	<i>Cryptococcus neoformans</i>

cc	Click chemistry
CEM	Cryptococcal meningoencephalitis
CFU	Colony-forming unit
CI	Chemical ionisation
CLint	Intrinsic Clearance
CNS	Central nervous system
CRBN	Cereblon
CRISPR-Cas9	Clustered regularly interspaced short palindromic repeats/ associated protein 9 nuclease
CRL	Cullin-RING ligase
CRUK	Cancer Research UK
CSF	Cerebrospinal fluid
CuAAC	Copper catalysed azide-alkyne cycloaddition
CYP450	Cytochrome P450
DKP	Diketopiperazine
DMPK	Drug metabolism pharmacokinetics
DNA	Deoxyribonucleic acid
DOPA	3,4-Dihydroxyphenylalanine
E_a	Activation energy
EDC.HCl	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
EDG	Electron donating group
EGFR	Epidermal growth factor receptor
EH	Epoxide hydrolase
Equiv.	Equivalents
ER	Oestrogen receptor
ERDs	Oestrogen-receptor down regulators
ES	Electrospray
Et	Ethyl
EUCAST	European Committee on Antimicrobial Susceptibility Testing

EWG	Electron withdrawing group
FDA	Food and Drug Association
FL	Fluorescent labelling
FLBZ	Flubendazole
FLU-NH₂	Reduced flubendazole metabolite
FLU-OH	Hydrolysed flubendazole metabolite
FMO	Flavin monooxygenase
GFP	Green fluorescent protein
GI	Gastrointestinal
Glu	Glutamine
GSK	GlaxoSmithKline
GST	Glutathione S-transferase
GTP	Guanosine triphosphate
GXM	Glucuronoxylomannan
HAART	Highly Active Antiretroviral Therapies
HATU	(1-[Bis(dimethylamino)methylene]-1 <i>H</i> -1,2,3-triazolo[4,5- <i>b</i>]pyridinium 3-oxide hexafluorophosphate
HBA	Hydrogen bond acceptor
HBD	Hydrogen bond donor
HECT	Homologous to the E6-AP carboxyl terminus
HER2	Human epidermal growth factor receptor 2
His	Histidine
HIV	Human Immunodeficiency Virus
HPLC	High performance liquid chromatography
HR	Hormone receptors
HR	Homologous recombination
Hr	Hour
HRMS	High resolution mass spectrometry
HTS	High throughput screening

I.V	Intravenous
IC	Inhibitory concentrations
ICP	Intracerebral pressure
ig	Immunoglobulin
IHC	Immunohistochemistry
iPr	Isopropyl
IR	Infrared
Kd	Dissociation constant
LC-MS	Liquid chromatography mass spectrometry
LDA	Lithium diisopropylamide
LH	Luteinizing hormone
LHRHs	Luteinizing hormone-releasing hormone agents
LRMS	Low resolution mass spectrometry
Lys	Lysine
<i>m</i>	<i>Meta</i>
MAO	Monoamine oxidase
MBC	Metabolite carbendazim
<i>m</i>-CPBA	<i>meta</i> -Chloroperoxybenzoic acid
MDR	Multi-drug resistant
Me	Methyl
MIC₅₀	Minimum inhibitory concentration required to inhibit the growth of 50% of organisms
Mics	Microsomes
MMDX	2(<i>S</i>)-Methoxy-4-morpholinyl doxorubicin
MoA	Mechanism of action
MP	Melting point
MRP	Multidrug resistant protein
MS	Mass spectrometry
MTOCs	Microtubule organising centres

MudPIT	Multidimensional protein identification
MW	Microwave
NHEJ	Non-homologous end joining
NMMHC	Non-muscle myosin heavy chain
NMM-IIA	Non-muscle myosin IIA
NMR	Nuclear magnetic resonance
NTP	National Toxicology Programme
<i>o</i>	<i>Ortho</i>
<i>p</i>	<i>Para</i>
PAL	Photoaffinity labelling
PARP	Poly ADP ribose polymerase
PARPi	Poly ADP ribose polymerase Inhibitor
PBS	Phosphate-buffered saline
PCC	Pyridinium chlorochromate
pCR	Pathological complete response
P-gp	P-glycoprotein
Phe	Phenylalanine
PK	Pharmacokinetics
PKC	Protein kinase C
PLH	Phenylahistin
PO	Per Os (by Mouth)
PPG	Propargyl
PPIs	Protein-protein interactions
PR	Progesterone receptor
PROTAC	Proteolysis targeting chimera
PRR	Phagocytic recognition receptors
RING	Really interesting new gene
RNA	Ribonucleic acid
RNAi	RNA interference

RO5	Rule of five
ROS	Reactive oxygen species
Rt	Room temperature
RTK	Receptor tyrosine kinase
SAR	Structure activity relationship
SC	Subcutaneous
SCX-2	Strong cation exchange-2
SD	Standard deviation
SDI	Sociodemographic Index
SDS-PAGE	Sodium dodecyl sulfate- polyacrylamide gel electrophoresis
SERMs	Selective oestrogen-receptor response modulators
SET	Single electron transfer
SIBX	Stabilized 2-Iodoxybenzoic acid
S_NAr	Nucleophilic aromatic substitution
SPR	Surface plasmon resonance
SSA	sub-Saharan Africa
STAT3	Signal transducer and activator of transcription 3
TBAF	<i>Tetra-n</i> -butylammonium fluoride
TFA	Trifluoroacetic acid
TFP	Trifluoperazine
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TNBC	Triple negative breast cancer
TOP1	Topoisomerase I
TPS	Terpene synthase
TUB1	Tubulin alpha-1 chain
TUB2	Tubulin alpha-2 chain
Ub	Ubiquitin
UK	United Kingdom

UPRTase	Uracil phosphoribosyl transferase
UPS	Ubiquitin proteasome system
US	United States
UV	Ultraviolet
VHL	Von Hippel Lindau
VNC	Ventral nerve cord
WIPE	Water/ isopropyl alcohol/ ethyl acetate

CHAPTER 1

Introduction to *Cryptococcus neoformans*,
Benzimidazoles and Photoaffinity Labelling.

1.1 *Cryptococcus neoformans*

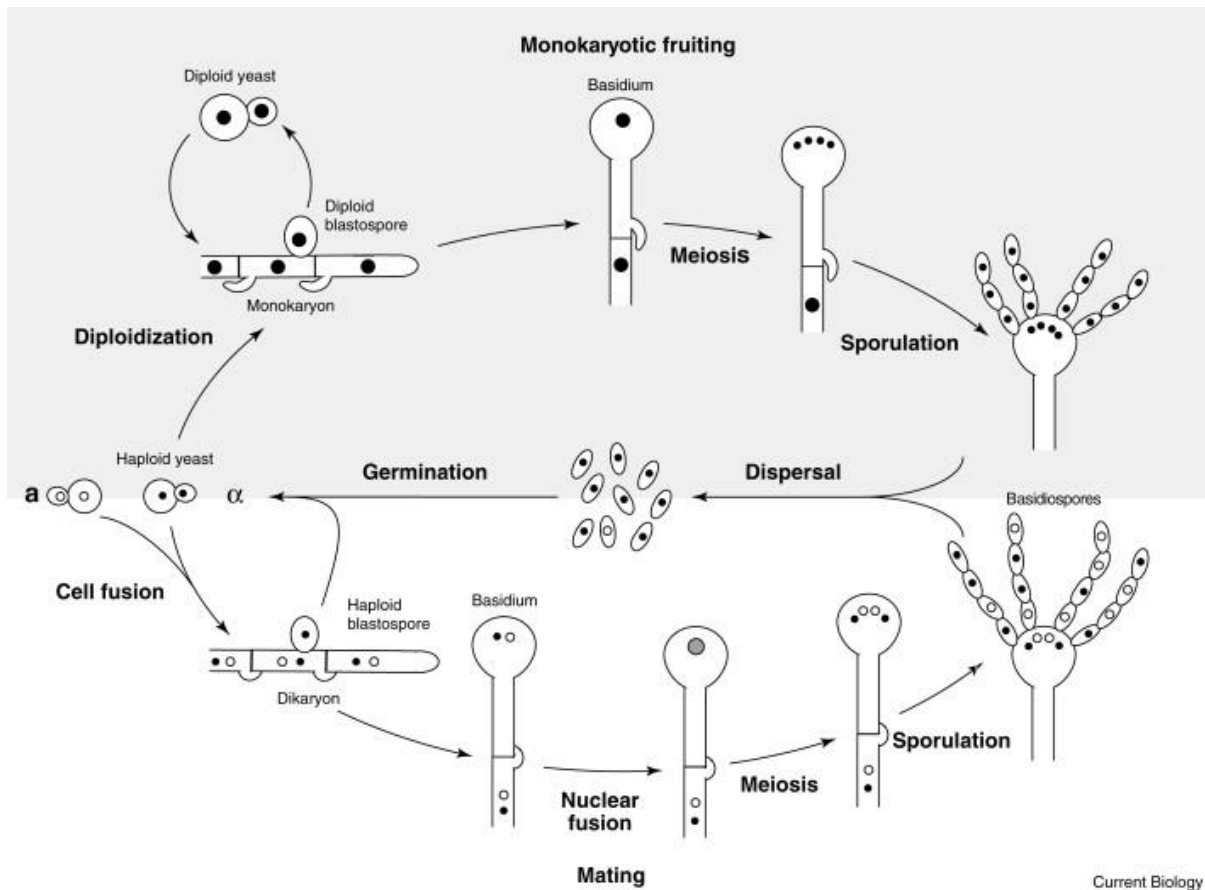
1.1.1 What is *Cryptococcus neoformans*?

Cryptococcus neoformans is an encapsulated yeast-like basidiomycete fungus, which has evolved over the past 40 million years into three distinct varieties; *C. neoformans* var. *neoformans*, *C. neoformans* var. *grubii* and *C. neoformans* var. *gattii*.^{1, 2}

1.1.2 Aetiology of *C. neoformans*

There are approximately thirty known species within the genus *Cryptococcus*. The major pathogen which causes cryptococcosis is *C. neoformans*, of which there are three varieties and five serotypes: *C. neoformans* var. *neoformans* (serotypes D and the hybrid strain AD), *C. neoformans* var. *gattii* (serotypes B and C), and *C. neoformans* var. *grubii* (serotype A).^{1, 3, 4}

The life cycle of *C. neoformans* involves sexual and asexual forms. *C. neoformans* can reproduce spores sexually *via* the two mating types MAT_a and MAT_α.⁵ The two mating types of *C. neoformans* conjugate to form a teleomorph, which consists of dikaryotic hyphae. The hyphae can develop into specialised structures known as basidia. Meiosis occurs at the terminal portion of the basidia, which results in the formation of basidiospores. Initially, these spores are unencapsulated, but develop a capsule once they are released from the basidia. Budding can occur after encapsulation to form yeast cells, which completes the sexual life cycle.^{6, 7} The asexual cycle is the encapsulated yeast form which reproduces *via* budding (monokaryotic fruiting) (see **Figure 1.1**).^{8, 9}



Current Biology

Figure 1.1: Life cycle of *C. neoformans*. The upper segment portrays asexual monokaryotic fruiting. The lower segment illustrates the traditional sexual cycle involving the two mating types (MAT α and MAT a). Reproduced from: *Sexual Reproduction and the Evolution of Microbial Pathogens*, J. Heitman.¹⁰

1.1.3 Epidemiology of *C. neoformans*

The first case of cryptococcosis was reported in 1894 by two German physicians, Otto Busse and Abraham Busk, when they isolated circular yeast like culture from a lesion in an infected woman's tibia.⁴ The fungus is an opportunistic pathogen, mainly affecting immunocompromised individuals. Early cases of cryptococcal infection (prior to 1980) were primarily associated with cancer, autoimmune diseases, and organ transplantation.^{11, 12, 13} Since the 1980's, there has been a surge in the number of cases of cryptococcosis due to the emergence of the Human Immunodeficiency Virus (HIV) pandemic. There was a 1500% increase in cryptococcal meningitis cases in USA between 1981-1990 compared to 1965-1977, due to HIV rapidly spreading worldwide.^{3, 14} Untreated HIV infection decreases an individual's CD4+ T cell count, resulting in the development of Acquired Immune Deficiency Syndrome

(AIDS).¹⁵ This immunocompromised state leaves patients at severe risk of opportunistic pathogens such as *C. neoformans* and other secondary infections. Cryptococcus is the most common cause of systemic mycosis in patients with AIDS and primarily occurs in sub-Saharan Africa (SSA). It is estimated that there were 957,900 cases of cryptococcal meningitis worldwide in 2006, with sub-Saharan Africa accounting for 720,000 of these cases, representing 75% of all instances (see **Figure 1.2**). The same study reported an average of 504,000 deaths in the same region, giving a mortality rate of 70%.¹⁶

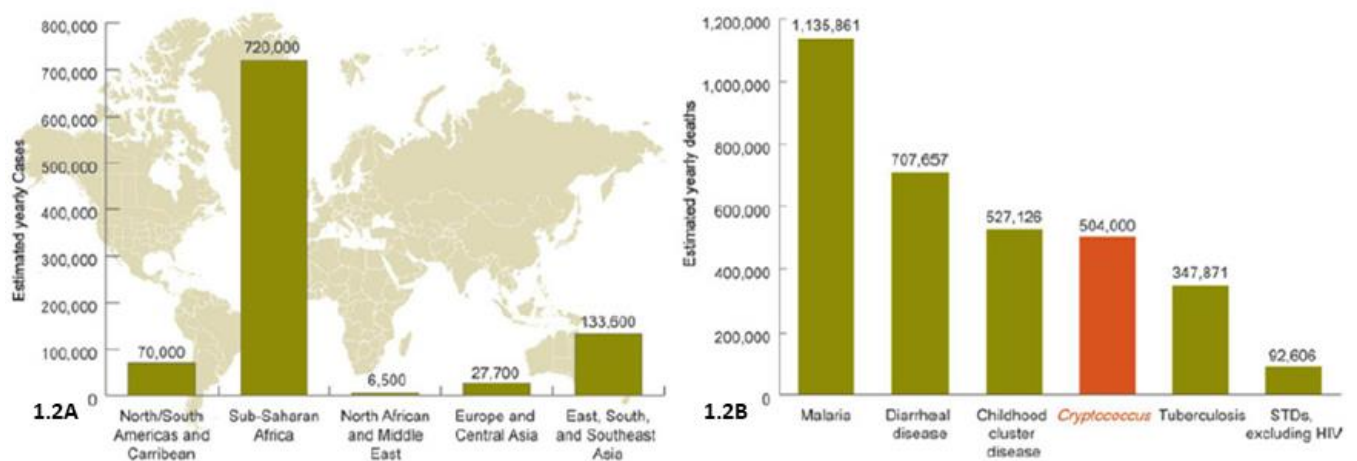


Figure 1.2: **1.2A** Global prevalence of *C. neoformans*, demonstrating that sub-Saharan Africa has the highest rate of cryptococcosis. **1.2B:** Causes of death in sub-Saharan Africa, excluding HIV/AIDS. Reproduced from: *Estimation of the current global burden of cryptococcal meningitis among person living with HIV/AIDS*, B. J. Park et al.¹⁶

The surge in cryptococcosis has decreased in recent years due to the use of Highly Active Antiretroviral Therapies (HAART) for the treatment of HIV. However, this downtrend is slower in developing countries as HAART has been introduced at a later stage due to expense issues and inadequate medical infrastructure.^{3, 17, 18, 19, 20, 21}

C. neoformans is distributed worldwide and strongly associated with soil, trees (*Eucalyptus camaldulensis*), and avian sources. Birds such as pigeons are thought to be carriers of the fungus as they consume feed infested with the fungus.²² However, it is rare for pigeons to be infected, as their body temperature exceeds the optimal and maximum growth temperature of *C. neoformans* (pigeons 42°C vs. *C. neoformans* 32-40°C).^{7, 23} Most cryptococcus cases are presumed to come from environmental sources rather than zoonotic transmission. However, there are some reports of zoonotic transmission from birds kept as pets.^{24, 25}

C. neoformans var. *grubii* (serotype A strains) cause approximately 95% of all cryptococcal infections, whereas *C. neoformans* var. *neoformans* (serotypes D and AD) and *C. neoformans* var. *gattii* (serotypes B and C) cause 4-5% of fungal infections. Serotype A is distributed worldwide. Serotypes B and C are strongly associated with tropical and subtropical regions such as southern California, Australia, Southeast Asia, and Central Africa. Serotype D is predominantly found in European countries such as Denmark, Germany, Italy, and France.³

1.1.4 Pathogenesis of *C. neoformans*

Cryptococcosis occurs *via* inhalation of airborne dehydrated yeast or basidiospores into the alveoli of the lungs. The small size of these spores (approximately 2.5-10 μm) allows them to lodge themselves deep within the tissue of the lung in the alveolar spaces and encounter alveolar macrophages. Cytokines and chemokines are released to signal the recruitment of other inflammatory cells, which results in a primary lung infection. The yeast can remain dormant in an individual for years, until the host's immune system is weakened. Once the individual is immunocompromised, the fungus can disseminate to other sites of the body causing further decline to the individual's health.^{3, 26}

The pathogenesis of cryptococcosis is determined by three main virulence factors: capsule formation, melanin synthesis, and growth at 37°C. The *C. neoformans* capsule is composed of high molecular weight polysaccharides, primarily glucuronoxylomannan (GXM). The capsule possesses antiphagocytic properties resulting in detrimental effects to the host's immune system. This impairs macrophages and phagocytes, preventing the clearance of the fungal spores. Before phagocytosis can occur, macrophages need to detect the presence of the pathogen. Macrophages have a number of phagocytic recognition receptors (PRR) on their cell surface to detect foreign cells. The capsule can shroud around the macrophage blocking the ability for the receptors to bind to pathogen and therefore blocking phagocytosis.^{27, 28} The second virulence factor is caused by *C. neoformans* ability to express the enzyme phenoloxidase, which converts dihydroxyphenols into melanin. Melanin allows the fungus to be resistant to free radical oxygen and nitrosative species produced from host immune response cells such as macrophages and neutrophils.²⁹ Melanin is also believed to interfere with the engulfing mechanism of phagocytic cells due to it being negatively charged. The negative charge may repel inherently negative cells such as phagocytes, inhibiting

clearance.³⁰ Finally, for *C. neoformans* to cause infection in a human host, the fungus must be stable at the human internal body temperature (37°C) in order to survive and proliferate.^{3, 17, 31} A number of genes have been identified which allow *C. neoformans* to grow at high temperatures without experiencing denaturation. Cryptococcal growth at high temperatures induces the upregulation of terpene synthase (TPS and TPS2) genes, which promote higher concentrations of the polysaccharide trehalose which can protect cellular proteins from denaturation.³²

1.1.5 Clinical Manifestations

C. neoformans tends to have a preference to present itself in the lungs and the central nervous system (CNS) of the host. However, due to the ability of the fungus to disseminate, cryptococcosis can present itself in any and multiple organs of the body.³

1.1.5.1 Pulmonary Infection

The fungal spores of *C. neoformans* enter *via* the respiratory tract and into the lungs of the host. This portal of entry allows the development of pulmonary cryptococcosis, which can present itself as asymptomatic in one third of cases. More commonly, pulmonary cryptococcosis is symptomatic and presents itself with a cough (54%), weight loss (26%), fever (26%), respiratory distress, and chest pain. Cryptococcal pneumonia can also develop with exposure to fungal spores.^{1, 3}

1.1.5.2 CNS Infection

Between 70-90% of patients with cryptococcosis present an infection in the CNS as subacute meningitis or cryptococcal meningoencephalitis (CEM). Patients may show symptoms of headaches, fever, lethargy, memory loss, and changes in personality. Despite *C. neoformans* capability to disseminate to any organ, the fungus shows a strong preference for the CNS. A major virulence factor of the fungus is melanin production, which requires high concentrations of dihydroxyphenols for the synthesis. Brain tissue has high concentrations of 3,4-dihydroxyphenylalanine (DOPA) which is a dihydroxyphenol. Therefore, once the fungus crosses the blood brain barrier (BBB), the fungus thrives and proliferates in the brain due the availability of DOPA, causing cryptococcal meningitis.^{1, 3, 33}

1.1.5.3 Skin Infection

Skin infections are the third most common clinical manifestation of cryptococcosis. *C. neoformans* presents itself as a wide range of skin lesions, although the most common is a papule rash with a central ulceration. The appearance of a cryptococcal skin lesion strongly signifies a disseminated cryptococcal infection. The skin lesions cannot be distinguished from other fungal skin infections; therefore, cultures are needed to diagnose the correct fungal infection for treatment.^{1, 3}

1.1.6 Current Treatment for *C. neoformans*

The current gold standard treatment for CEM is a course of amphotericin B (AmB) combined with flucytosine for two weeks, followed by maintenance treatment with fluconazole for a minimum of 10 weeks.³⁴ Non-CNS cryptococcosis can be treated with fluconazole alone ranging from six months to two years depending on the presence of an HIV infection.^{3, 35}

Most deaths from CEM occur before or during the induction phase of treatment in a medical facility. However, high mortality rates are reported in SSA during the period after discharge when a patient transitions to oral fluconazole. In the 6 months after discharge, 20-30% of patients will die, typically in 4–5 weeks after discharge. These rates far exceed anything seen in North America and Europe.^{36, 37, 38}

An explanation for this could be the inadequate induction therapy, as 5-FC which is an essential component of the combination therapy with AmB is not always available in many parts of SSA due to licensing issues and high costs.³⁹ The poor medical infrastructure also plays a part as procedures such as therapeutic lumbar puncture, which is critical in managing elevated intracerebral pressure (ICP) and reducing fungal burden, is seldom used.^{38, 40} One third to one half of patients in SSA still have viable *Cryptococcus neoformans* present in their cerebrospinal fluid (CSF).^{38, 41, 40, 42}

After induction therapy, fluconazole dosed at 400 mg/day has been the most widely used dose strategy in SSA for consolidation antifungal therapy. Unfortunately, this recommendation was established in clinical trials among patients who received a more optimal induction regimen consisting of AmB and 5-FC. It has been suggested that increasing the therapeutic dose of fluconazole may reduce the mortality rates associated with *C. neoformans*, as indicated in animal models involving rabbits with CEM. A rabbit model in

which low (100 mg), intermediate (400 mg), and high-dose (1600, 2000 mg) fluconazole were administered, showed a dose–response relationship. The higher fluconazole dose was associated with a greater reduction of *C. neoformans*.^{43, 38}

1.1.6.1 Amphotericin B

AmB is a member of the polyene macrolide class of antibiotics, which is characterised by the possession of a macrocyclic ring of carbon atoms closed by lactonization (see **Figure 1.3**). The polyene antifungal is derived from the fermentation process of *Streptomyces nodosus*. AmB is amphoteric as it forms soluble salts in both acidic and basic environments. However, it is not sufficiently soluble in water, therefore requires the addition of sodium deoxycholate to form a colloidal dispersion.^{44, 45}

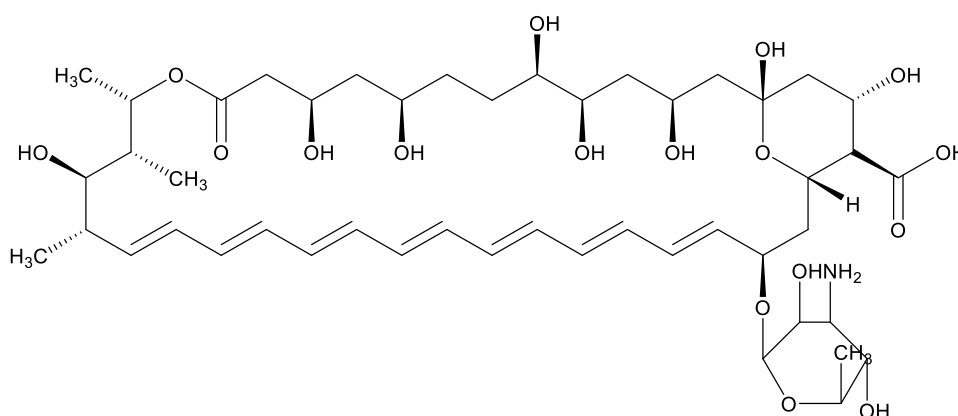


Figure 1.3: Structure of Amphotericin B

1.1.6.1.1 Mode of Action of Amphotericin B

The mechanism of action of AmB involves the binding of the polyene to ergosterol which is present in the cell membrane of the fungus. Ergosterol is essential for fungal survival as it plays a role in endocytosis, vacuole fusion, and the stabilisation of proteins in the cell membrane. The binding results in the formation of pores, which disrupts the stability of the cell membrane, resulting in the leakage of the cells contents. Unfortunately, AmB is not selective to fungal sterols, as it has been shown to bind to mammalian sterols such as cholesterol. AmB binds to cholesterol at a lower affinity than ergosterol, due to the difference in sterol structure. Ergosterol has a more cylindrical three-dimensional structure compared to mammalian cholesterol, which allows greater affinity binding (see **Figure 1.4**). Studies show that the formation of pores in the fungal cell membrane do not correlate to the death of fungal cells, which suggests that AmB exerts a second antifungal mechanism.^{3, 44, 45, 46}

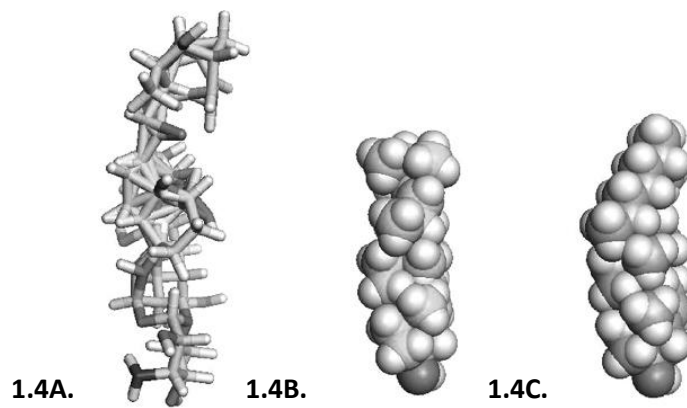


Figure 1.4: 3D structures of sterols. **1.4A:** AmB, **1.4B:** Ergosterol, **1.4C:** Cholesterol. Reproduced from: *Diagnosis and Treatment of Human Mycoses*, M. Duane et al.³

AmB is described as a pro-oxidant which may induce an oxidative burst in fungal cells, by allowing reactive oxygen species (ROS) to accumulate in the mitochondria of fungal cells. This oxidative burst leads to extensive damage to the fungal cells components such as proteins, which ultimately leads to *C. neoformans* death (see **Figure 1.5**).^{3, 44, 45, 46}

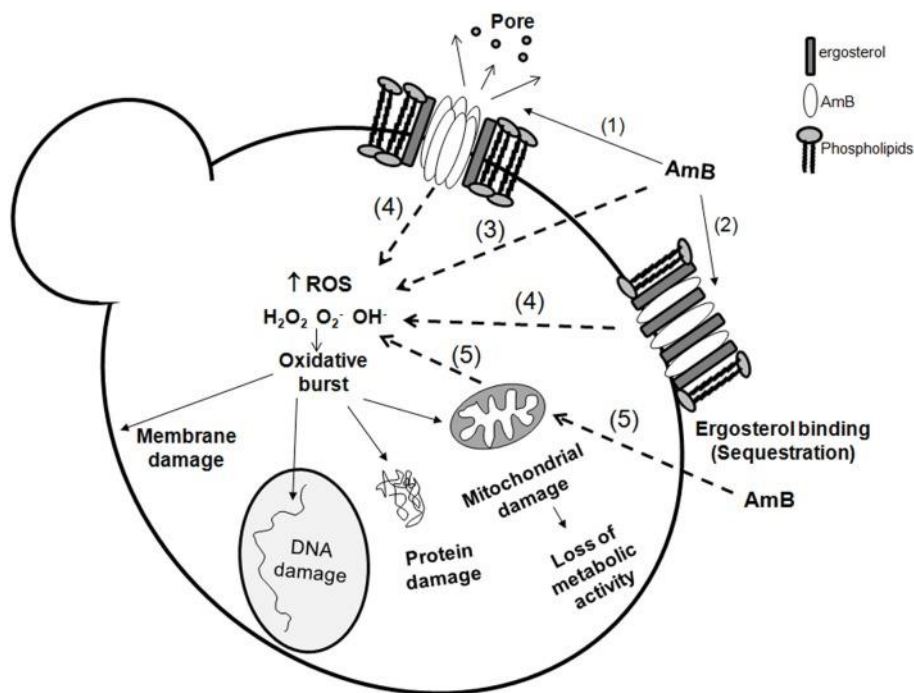


Figure 1.5: Different AmB mechanisms of action. (1) At the membrane, AmB can bind to ergosterol and form pores. (2) Ergosterol binding may also result in the induction of ergosterol sequestration resulting in membrane instability. Within the cell, AmB can induce an oxidative burst *via* several proposed mechanisms. (3) AmB acting as a prooxidant to cause the accumulation of ROS. (4) The binding of AmB to ergosterol can cause intracellular effects in the form of generating ROS. (5) AmB may influence fungal mitochondria, generating a higher output of the natural ROS from the respiratory chain. The oxidative burst can damage fungal mitochondria, proteins, DNA, and the stability of the membrane. Reproduced from: *It only takes one to do many jobs: Amphotericin B as antifungal and immunomodulatory drug*, A. C. Mesa-Arango *et al.*⁴⁶

1.1.6.1.2 Limitation of Amphotericin B

Toxicity associated with AmB limits the use and effectiveness in severely ill patients. The most common acute toxicity of AmB is infusion-related reactions which presents as a fever, chills, nausea, vomiting, headache, and hypotension/ hypertension. This adverse effect is believed to a result of AmB being a microbial product which stimulates the production of proinflammatory cytokines by the hosts' immune system. *Goodwin et al* conducted a study involving 400 patients and showed that more than half suffered from at least one infusion-related adverse effect.⁴⁷ Chronic toxicities of AmB involve nephrotoxicity, anaemia, neutropenia, and neurotoxicity. The main chronic adverse effect is nephrotoxicity, as it causes renal failure, urinary potassium wasting, hypokalaemia, and many other conditions.

As a result, nephrotoxicity is the dose limiting factor in most patients being treated with AmB.^{3, 48}

AmB requires intravenous (I.V) administration, as it is a large polyene that requires sodium deoxycholate to form a colloidal dispersion. This is a significant drawback, as this administration requires medical infrastructure which is not readily available in the developing world, such as SSA, where *C. neoformans* infections are most prevalent.⁴⁸

Another limitation to the use of AmB is fungal resistance, although acquired AmB resistance is low despite its widespread use. Resistance operates through several mechanisms, one of which is the decrease in ergosterol content in the fungal cell membrane, along with alterations in the ergosterol synthesis pathways. Fungal cells such as *C. neoformans* may also avert antifungal activity by displaying resistance to oxidative stress. Genome wide expression analysis has revealed that there is an increase in the expression of ERG genes, along with the induction of stress genes, as they can up regulate the production of catalase, and other mitochondrial enzymes. These enzymes are able to convert oxidative species into neutral species, and hence prevent the AmB induced oxidative burst.^{46, 48}

1.1.6.2 Flucytosine

Flucytosine (5-FC) is a synthetic fluorinated analogue of cytosine which displays antimycotic properties (see **Figure 1.6**). It is used as an antifungal agent which was originally developed in 1957 as an antimetabolite.⁴⁹

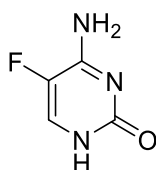


Figure 1.6: Structure of Flucytosine

1.1.6.2.1 Mode of Action of Flucytosine

Flucytosine (5-FC) acts as an antifungal agent *via* conversion to 5-Fluorouracil (5-FU). Fungal cells that possess cytosine permease can internalise flucytosine and allow the antifungal agent to be converted to 5-FU *via* the cytosine deaminase enzyme. 5-FU can be converted into a substrate for nucleic acid synthesis *via* an enzyme known as uracil phosphoribosyl

transferase. This is then incorporated into the fungal ribonucleic acid (RNA), which causes premature chain termination, and hence inhibits deoxyribonucleic acid (DNA) synthesis through the effects on thymidylate synthase. Only fungal cells that possess these enzymes can be targeted by the flucytosine antifungal agent, therefore *C. neoformans* can be targeted, and 5-FC can exert its effect.^{50, 51}

1.1.6.2.2 Limitations of Flucytosine

Flucytosine is associated with a number of side effects. One of the most common, but least harmful is gastrointestinal disturbances, such as nausea, diarrhoea, and sometimes vomiting. Gastrointestinal disturbances occur in approximately 6% of patients who are treated with 5-FC.⁵² Another adverse drug reaction (ADR) of 5-FC is hepatotoxicity. The incidence of hepatotoxicity is not clear, but it has been reported between 0-25%.^{53, 54} The most serious associated toxicity is bone marrow depression. In a study by *Kauffman and Frame*, 4 out of 15 patients treated with 5-FC developed bone marrow toxicity.⁵⁵ It has been reported that symptomatic HIV infected patients have a higher incidence of this toxicity. The mechanisms of toxicity are not fully understood. Although it is believed that toxicities such as bone marrow depression and hepatotoxicity are dose dependent.⁵⁶

Acquired resistance in *C. neoformans* to 5-FC can occur *via* two basic mechanisms. The first mechanism of resistance has occurred due to mutations occurring in the DNA, leading to the synthesis of deficient enzymes such as cytosine permease, cytosine deaminase, uracil permease and uracil phosphoribosyl transferase which are necessary for 5-FC cellular transport, uptake, and metabolism (see **Figure 1.7**). The second mechanism of resistance is the increase in the synthesis of pyrimidine nucleic bases. This increase allows the nucleic bases to out compete the fluorinated antimetabolites of 5-FC for the enzymes active site, and therefore diminish the agents antimycotic activity. Studies have shown that a defective uridine monophosphate pyrophosphorylase is the most frequently occurring type of acquired flucytosine resistance in fungal cells.⁵⁶ Resistance to flucytosine has occurred due to the drug being introduced as a monotherapy in the early 1970's; as a result the original potent activity was lost rapidly. This is the primary reason why AmB and flucytosine are now combined.^{35, 57}

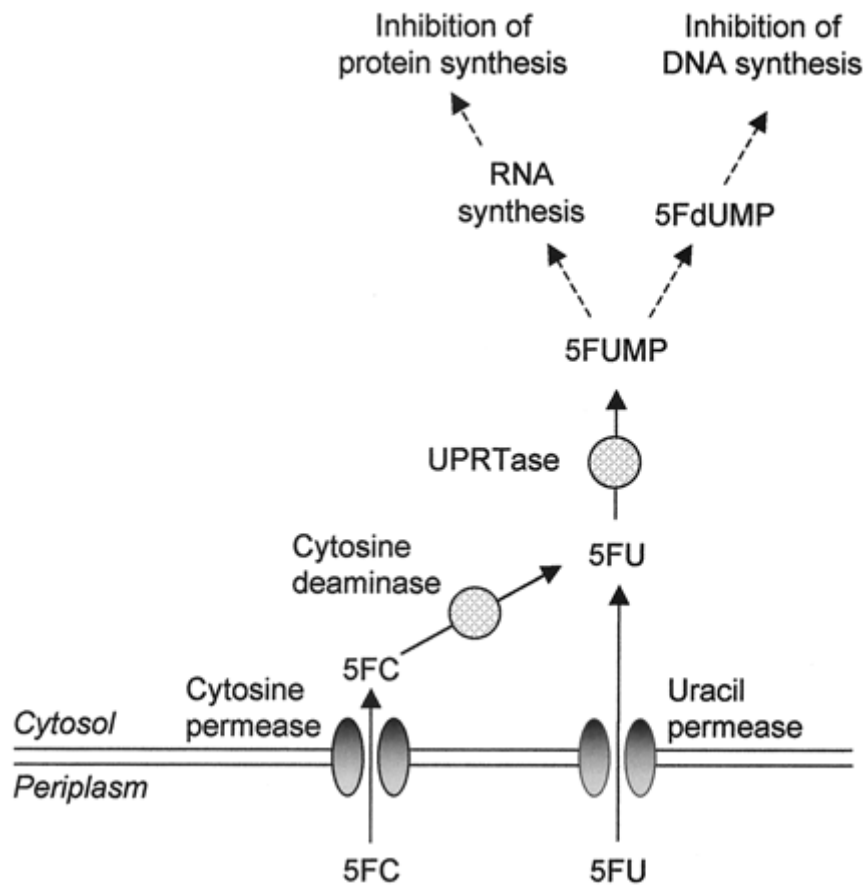


Figure 1.7: Main enzymatic pathways involved in the uptake, conversion, and mechanism of action of 5-FC. Abbrev. 5FC: 5-Fluorocytosine, 5FU: 5-Fluorouracil, 5FUMP: 5-Fluorouridine monophosphate, 5FdUMP: 5-Fluorodeoxyuridine monophosphate, UPRTase: Uracil phosphoribosyl transferase. Reproduced from: *Flucytosine-Fluconazole Cross-Resistance in Purine-Cytosine Permease-Deficient Candida lusitanae Clinical Isolates: Indirect Evidence of a Fluconazole Uptake Transporter*, Noël et al (2003).⁵⁸

5-FC has been recommended as a first line treatment, yet this therapeutic agent is not widely available in areas where cryptococcosis is most prevalent (Africa and Asia) due to licensing disputes. 5-FC is a very expensive drug (\$120/day in Uganda), therefore not readily available to the developing world where financial support is limited.^{34, 59, 60}

1.1.6.3 Fluconazole

Fluconazole is an antifungal agent which was introduced in the 1980's (see **Figure 1.8**). Fluconazole is well absorbed and highly orally bioavailable,⁶¹ as it has a 90% bioavailability after oral administration.⁶² Fluconazole is not as potent as AmB and displays lower efficacy due to fluconazole being fungistatic, while AmB and flucytosine are fungicidal. Fluconazole is

typically used as maintenance therapy after an initial course of AmB and 5-FC is completed.^{35,}

63, 64

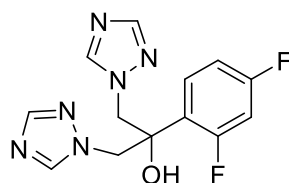


Figure 1.8: Structure of fluconazole

1.1.6.3.1 Mode of action of fluconazole

Fluconazole is an inhibitor for cytochrome P450 (CYP) 14- α -demethylase (CYP51) as the *N*-4 triazole component interacts with the heme iron atom of the CYP51. CYP51 plays an important role in the biosynthesis of lipids and steroids, particularly the conversion of lanosterol to ergosterol. Ergosterol is an important component of fungal cells as it is involved in a number of essential functions, such as stabilising the proteins in the fungal plasma membrane. The inhibition of CYP51 results in the decrease of ergosterol concentrations, which in turn disrupts the stability of the cell membrane. This disruption can make the pathogen susceptible to further damage as it interferes with the stability of the plasma membrane. It can also alter the activity of several enzymes that are bound to the plasma membrane which are involved in several processes such as nutrient transport and chitin synthesis. Ergosterol has hormone-like functions, which affects cell growth and proliferation, which can be suspended during severe ergosterol depletion and hence leads to fungal cell death.^{65, 66}

1.1.6.3.2 Limitations of Fluconazole

The fungistatic property of fluconazole can lead a patient to clinical relapse and drug resistance through a prolonged high fungal burden. This has led to a higher mortality rate compared to the AmB and flucytosine combinational therapy.^{34, 59}

The use of fluconazole has also been limited due to the emergence of resistance. There are a number of underlying mechanisms which contribute to fluconazole resistance. The first mechanism of resistance is a result of alterations to the target enzyme known as cytochrome P450 lanosterol 14- α -demethylase (Erg11p). Alterations of the enzyme could be caused by several factors. One of which may be the over expression of the enzyme, therefore higher

intracellular concentrations of fluconazole would be required to bind and complex with all the target enzymes that are present in the cell. To obtain high enough concentrations, the dose would have to be increased, which could result in toxicity making the treatment unviable. The rise of point mutations in the gene that encodes the enzyme (ERG11) is another factor. One or more amino acid substitutions can alter the topology of the active site leading to inadequate or no binding of fluconazole to the enzyme, which would lead to a reduction in antifungal activity.^{67, 68} This mechanism of resistance leads to fluconazole unable to accumulate inside the fungal cell at adequate concentrations for a desirable therapeutic effect due to enhanced drug efflux mechanisms.^{69, 70}

1.1.7 Development of the ideal anti-cryptococcal drug

Due to the limitations associated with the current treatment of *C. neoformans*, a new anti-cryptococcal agent needs to be developed. The new class of therapeutic agents must possess a number of essential properties for it to be commercialised. These properties include a good drug metabolism pharmacokinetics (DMPK) profile, low toxicity, improved efficacy compared to the current treatments on the market, and high selectivity against the target, with little to no interaction with the host.⁷¹

There are additional requirements needed for the development of an anti-cryptococcal agent. Due to the high prevalence of *C. neoformans* in SSA, the development and production of the target compound needs to be cheap, as most developing countries spend less than \$20 per year and per head on all health programmes. In SSA, this value is reduced to \$6, which includes drug expenditures.⁶⁶ Furthermore the lack of resources and medical infrastructure in SSA requires the anti-cryptococcal drug to be orally bioavailable, as administration *via* this method requires fewer resources and has a higher patient compliance.⁷² Although the development of an agent that could also be administered both orally and intravenously is ideal as some patients who are seriously ill may not be capable of taking oral medication. Due to a number of studies reporting the emergence of resistance from the current treatments of *C. neoformans*, the ideal anti-cryptococcal drug would have fungicidal activity rather than fungistatic. The former can kill the pathogen and cure the infection, whereas the latter will only suppress it, which can give rise to resistance and re-emergence of the infection.

An essential part of drug development involves the understanding of the mechanism of action (MoA) of the potential drug. Unfortunately, many researchers fail to focus on this aspect of drug discovery and continue with clinical trials. An explanation for this absence of information is that the Food and Drug Association (FDA) does not always request evidence for the MoA of a drug.⁷³ Understanding the MoA allows researchers to identify drug targets and their pathways. This knowledge can allow chemists to develop therapeutic agents with an improved efficacy and toxicity profile. Drugs can interact with multiple binding sites, increasing the risk of drug related toxicity. Identifying the target responsible for the therapeutic effect allows the synthesis of a more selective drug, which in turn may reduce toxicity.⁷⁴ Understanding MoA also allows researchers to predict resistance mechanisms. Combinational therapies can be developed to circumvent the potential resistance pathways.⁷⁵ An additional advantage of understanding the MoA is preventing the risk of late-stage failures during clinical trials. Currently, there is a 8-14% success rate for all drugs that have entered clinical trials.^{73, 76, 77} Failure during late stages causes a huge economic burden on the pharmaceutical industry, as well as wasting time and effort on an unsuccessful candidate. This misused time, and resources could have been better applied to potentially superior candidates.^{73, 78}

1.2 Photoaffinity Labelling Probes (PAL)

1.2.1 Binding site determination

To understand a potential therapeutic agent's mechanism of action, we must examine the binding between the drug and the target protein. There are a number of established methods that aid in the understanding of protein binding and function, such as X-ray crystallography⁷⁹, Nuclear Magnetic Resonance (NMR) spectroscopy⁸⁰, *in silico* docking studies,⁸¹ hydrogen/deuterium exchange⁸² and photoaffinity labelling (PAL).⁸³ PAL has experienced a surge in interest in recent years *via* the use of chemical probes. PAL probes are able to provide a vast amount of proteomic information while not being limited by post translational mechanisms.^{83, 84}

1.2.2 PAL probe design

PAL depends on the use of chemical probes that can covalently bind to an active site of a protein. There are three main features of the PAL probe: a photo reactive group, a binding

group/ linker, and a reporter tag (see **Figure 1.9**).⁸⁴ When designing the chemical probe, careful addition of substituents onto the parent compound must be made to maintain activity. PAL probes usually require structure activity relationship (SAR) information from a set of parent compound analogues to understand functional group tolerance for successful design and synthesis.^{85, 86}

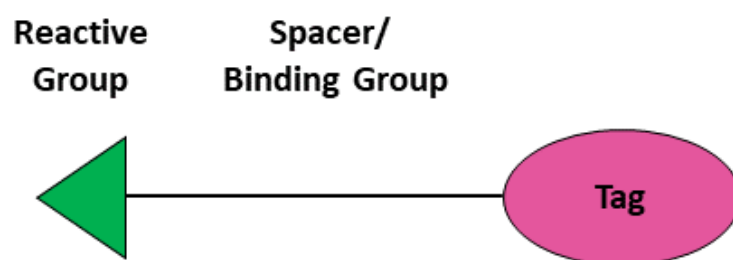


Figure 1.9: General structure of a PAL probe containing a photo reactive group (**green**), a binding group/ linker (**black**) and a reporter tag (**pink**).⁸⁴

1.2.2.1 Photo reactive group

Photo reactive groups can be employed as the “war head” of the PAL probe. The photo reactive moiety can covalently bind to amino acid residues within the target of the proteome upon irradiation by light. PAL requires an irreversible covalent bond to form between the chemical probe and target protein, for successful isolation and detection. Some proteins are only able to provide reversible ligand-protein binding. PAL can address this issue by introducing a photo reactive group, which is able to permanently attach to the protein. Upon irradiation with a specific wavelength of light, the photo reactive group forms a reactive intermediate, which rapidly binds to the target protein.^{87, 88, 89}

The photoaffinity probe needs a range of ideal properties for it to be successful in proteomic experiments. Stability is an important factor, as the reactive intermediate must not form before it has been deployed in the biological sample. The probe must be stable in the absence of light, a wide range of pH and temperature, and ideally resistant to a variety of oxidising and reducing agents. Minimal steric hindrance is another factor; the photo reactive group must not hinder any essential element of the activity-based protein profiling (ABPP) probe to ensure maximum potential for ligand-protein interaction. Another vital factor concerns the specific wavelength of light to activate the reactive intermediate; the frequency of the

wavelength must not disrupt or denature the biological molecules of interest. The photo affinity groups should be activated by wavelengths of light that exceed the absorption wavelength of proteins. Proteins tend to absorb ultraviolet (UV) light in the range of 280 and 200 nm. Aromatic amino acids such as tryptophan, tyrosine, and phenylalanine are responsible for the absorbance at 280 nm, whereas peptide bonds are responsible for absorbance at 200 nm. The resultant reactive intermediate must also be reactive enough to covalently bind with the target and form a strong enough interaction to withstand isolation and detection procedures.^{87, 88, 89}

There are many examples of photo affinity groups; the most common are phenyl azides, benzophenones, and diazirines, as they form ideal reactive intermediates for the study of proteins.^{87, 88, 89}

1.2.2.1.1 Phenyl azides

Phenyl azides can form a nitrene reactive intermediate following light activation at approximately 308 nm. Phenyl azides are commonly used for PAL as they are easy to synthesise and are readily commercially available. However, activation of the nitrene reactive intermediate occurs at shorter wavelengths of light, which have the potential to damage biological molecules. Nitrenes are less reactive than other reactive intermediates (carbenes), which results in lower photo affinity yields during ABPP procedures. Nitrenes can also undergo rearrangement reactions to form undesirable side products such as benzazines and 1,2-azacycloheptatetraenes. This induces non-specific labelling as they disperse and interact with distant nucleophiles, decreasing the photolabeling yields (see **Figure 1.10**).^{87, 89, 90} To improve the stability of the photoreactive intermediate, substituted phenyl azides can be synthesised such as difluorophenyl azide and tetrafluorophenyl azide.^{87, 90} The ring expansion rearrangement of the fluorinated phenyl azide reaction is slower than the non-fluorinated phenyl azide, due to the larger activation energy barrier (fluorinated E_a 7-8 Kcal/ mol vs non-fluorinated E_a 2-3 Kcal/mol).⁹¹

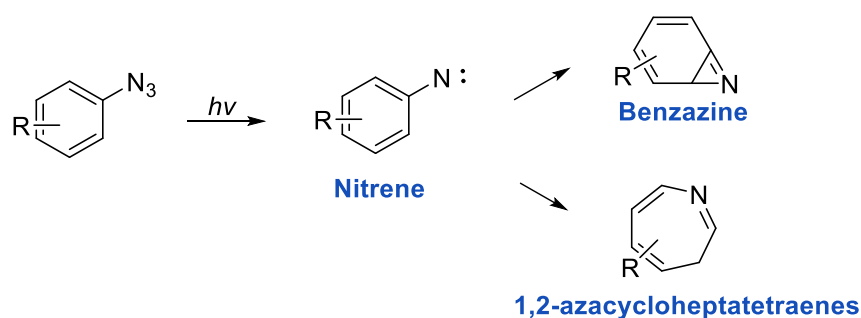


Figure 1.10: Phenyl azide reaction scheme for light-activated photochemical side reactions.^{89, 90}

1.2.2.1.2 Benzophenones

Benzophenones upon irradiation will form a reactive triplet diradical intermediate (see **Figure 1.11**). Activation of the reactive intermediate requires a longer wavelength of light, which reduces the risk of damage to biological molecules. Benzophenones are synthetically easy to prepare and are readily commercially available. However, benzophenones require a longer irradiation period to form the reactive intermediate, which increases the risk of nonspecific labelling. The benzophenone is bulky, and therefore can cause steric hindrance by blocking essential elements of the ABPP probe, and prevents entry of the probe into the target site of the protein.^{87, 92}

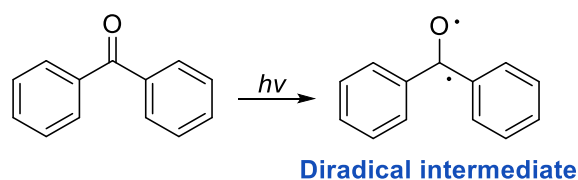


Figure 1.11: Reaction scheme of benzophenone UV activation to form a diradical intermediate.⁹³

1.2.2.1.3 Diazirines

The most used photo affinity group is the phenyldiazirine, which forms a carbene reactive intermediate (see **Figure 1.12**). Phenyldiazirines are stable to several factors such as temperature, nucleophiles, and oxidising and reducing agents. The trifluoromethyl derivative developed by *Brunner et al* in the 1980s is favoured among chemical biologists due to its exceptional chemical stability.⁹⁴ During photo irradiation, diazirines generate carbenes which can readily insert into N-H, O-H and C-H bonds of amino acid residues resulting in successful photolabelling (see **Figure 1.13**).^{95, 87, 96} Unfortunately, diazirines can isomerize into linear diazo compounds. This can present a problem as carbenes are highly reactive species with short lifetimes in the range of nanoseconds. Diazo compounds in contrast have much longer

lifetimes, are less reactive and more sensitive to protonation, and can generate carbenes or carbocations that can label nucleophilic residues of a protein. Due to the longer lifetime, the diazo compounds can diffuse from the initial site of generation to other sites of the protein, leading to nonspecific binding and obscuring results.⁹⁷ The trifluoromethyl group attached to an aromatic diazine allows it to become a more suitable photolabeling agent as it can stabilise the carbene, preventing the tendency for rearrangements. Also due to its electron withdrawing nature, the photogenerated diazo isomer is so strongly stabilised that it is considered unreactive.^{89, 92, 94}

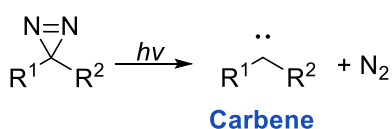


Figure 1.12: Reaction scheme for UV activation of a diazine.⁹²

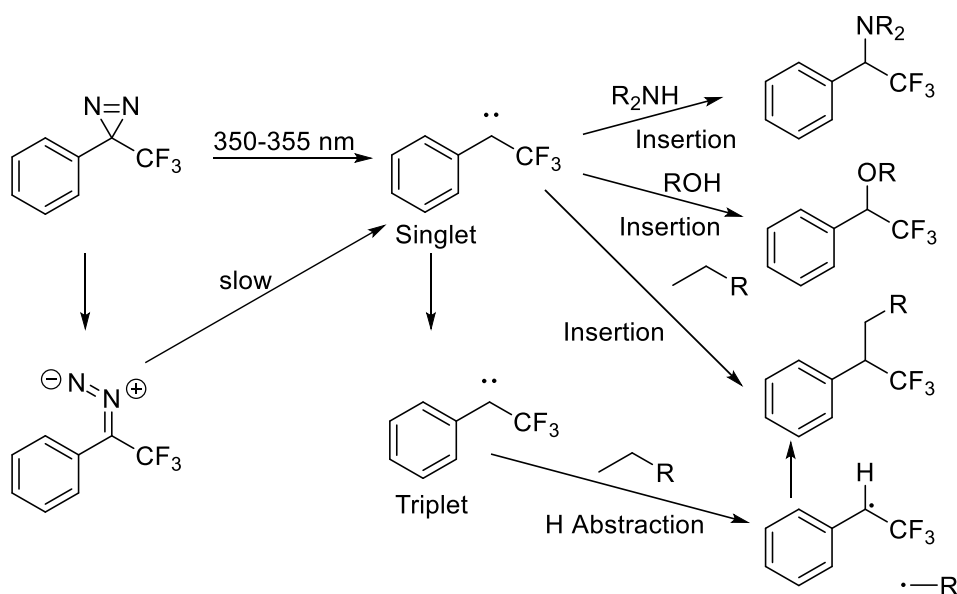


Figure 1.13: Photolytic products of an aromatic diazine. Adapted from: *Current advances of carbene-mediated photoaffinity labelling in medicinal chemistry*, S. Ge et al; *Photolytic Labelling and Its Applications in Protein Drug Discovery and Development*, Y. Chen et al; *Amino Acid Insertion Frequencies Arising from Photoproducts Generated Using Aliphatic Diazirines*, D. Ziemianowicz et al.^{86, 98, 99}

If the phenyl ring is substituted with an electron donating group (EDG), the carbene is likely to be in the singlet state with both electrons spin paired. If the phenyl ring is substituted with an electron withdrawing group (EWG), then the carbene will be in a triplet state and be

paramagnetic due to their unpaired electrons. Both states are stable; however, the triplet state has extra stability due to the laws of Hund's rule of maximum multiplicity, as the electrons in the p orbital are not experiencing electron repulsion (see **Figure 1.14**).^{100, 101}



Figure 1.14: **1.14A:** Singlet carbene with paired electrons, **1.14B:** Triplet carbene with unpaired electrons.⁹²

The carbene is activated at a high wavelength between 350-365 nm, which reduces the potential to damage biological molecules. The carbene is very reactive and has a short half-life, which minimises the risk of nonspecific binding. However, carbenes are quickly quenched with water, which can lead to low photo affinity yields, due to the aqueous environments present in biological systems.⁸⁷

1.2.2.2 Binding group/ linker

Some ABPP probes include a binding group to improve target selectivity, as they direct the reactive group towards the specific class of proteins under investigation. The chemical structure of the binding group varies considerably, due to the specificity of the target protein. The binding group is also able to function as a linker between the reactive group and the reporter tag, which allows them to be kept at a suitable distance. If the linker is too short, then the probe has the potential to cross-link with itself. If the linker is too long, the reactive group is too far away to bind to the target protein efficiently.^{83, 84, 87, 102}

1.2.2.3 Reporter tag

The reporter tag facilitates target characterisation and purification of probe-modified proteins. There are a number of examples of tags such as fluorophore, biotin, and click chemistry probes.^{89, 103}

1.2.2.3.1 Fluorophore and biotin tags

The most frequently used tags are biotin and fluorophores (see **Figure 1.15**). Biotin reporter tags can be detected *via* a simple Western blot test and avidin enrichment, prior to mass spectrometry (MS) based identification techniques. However, Western blotting tends to be time consuming, and can experience protease degradation resulting in unusual or unexpected

bands during detection.¹⁰⁴ Fluorescent tags can be detected by direct scanning of the gel with a fluorescent scanner, which tends to be faster and simpler than the Western blotting methodology. Despite the simple detection of fluorophore tags, biotin is more frequently used, due to the ability to provide both gel-based detection and purification methods.^{84, 102}

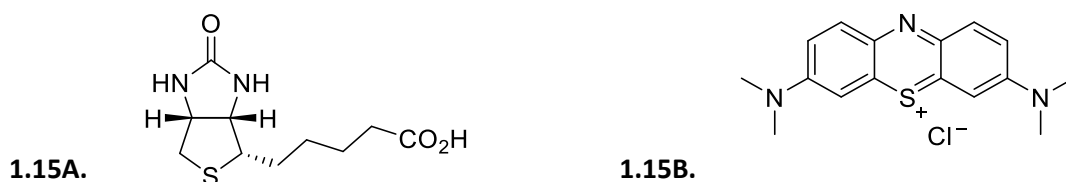


Figure 1.15: **1.15A:** Structure of biotin, **1.15B:** Structure of a fluorophore methylene blue.

Fluorophore and biotin are bulky reporter tags, which introduces a major disadvantage when the probes are used *in vivo*. The size of the tags reduces cell permeability of the ABPP probe, along with sterically hindering the probe from interacting with the binding site of the target protein.⁸⁵ The poor cell permeability has resulted in the proteomic source being homogenised to allow interaction between the chemical probe and the protein, which results in the disruption of the native cellular environment.^{105, 106}

1.2.2.3.2 Click chemistry probes

Due to the limitations of bulkier reporter tags, advancements in biorthogonal chemistry have been developed with the incorporation of clickable handles in the ABPP probe. These handles utilise click chemistry (cc). The most common example of click chemistry is the Cu (I) catalysed Huisgen 1, 3-dipolar cycloaddition of azides and alkynes, to form stable 1, 2, 3-triazoles (see **Figure 1.16**). The cc-ABPP probe consists of two parts; (i) the first structure contains the reactive group, a binding group/linker, and an alkyne handle (ii) the second structure contains a terminal azide on a binding region/ linker chain. The azide and alkyne can be on either of the two cc-ABPP probes. The first half of the cc-ABPP probe can enter cells and covalently attach to the target protein within the cell. Following cell lysis, the second half of the probe is introduced along with a copper catalyst and a reducing agent. This initiates the click chemistry reaction, which links the two chemical probes to form the stable 1, 2, 3-triazole structure, which can later be isolated and detected.^{87, 105, 106, 107, 108}

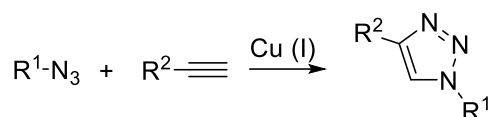


Figure 1.16: Huisgen 1, 3-dipolar cycloaddition of azides and alkynes forms a stable 1, 2, 3-triazoles.

These chemical handles are less bulky, allowing small molecule chemical probes to be utilised. The smaller size reduces the risk of poor cellular permeability and can enter cells with minimal perturbation. The small size also increases the chances of the probe entering the binding site of the target protein.^{87, 105, 106, 107, 108}

Unfortunately, there are limitations to the Huisgen's 1,3-dipolar cycloaddition reactions, as they typically use a Cu (I) metal catalyst. The use of a Cu (I) catalyst in a biorthogonal click reaction can cause mild to severe cytotoxicity within the living system being studied. Therefore, Cu-free click reactions have been developed to address this issue.¹⁰⁹

Jewett et al demonstrated that the rate of the reaction can be increased in the absence of the metal catalyst by exploiting the ring strain experienced by cyclooctyne. Typically, alkynes have bond angles of 180° as the carbon atoms are sp hybridised. Cyclooctyne has a bond angle of 160° due to the alkyne being incorporated into an eight-membered ring (see **Figure 1.17**).^{109, 110}

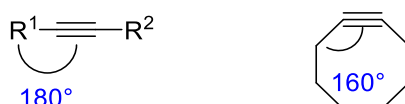


Figure 1.17: Comparison of bond angles between a linear alkyne and cyclooctyne.^{109, 110}

The large ring strain causes the cyclooctyne alkyne to react readily with phenyl azide to experience less ring strain as the system becomes sp² hybridised with a bond angle of 120° (see **Figure 1.18**).^{109, 110}

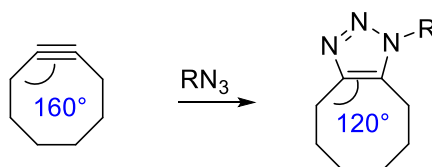


Figure 1.18: A demonstration of a Cu-free click chemistry reaction between an azide and cyclooctyne.^{109, 110}

1.2.3 Analytical platforms for ABPP

There are a number of platforms to analyse the proteomes. These platforms vary depending on the structure of the ABPP probe, due to the presence of different reporter tags.⁸⁴

1.2.3.1 Gel electrophoresis

One of the most common analytical platforms for ABPP involves gel-based methods. Probe treated proteomes undergo either one-dimensional (1D) or two-dimensional (2D) polyacrylamide gel electrophoresis, which separates the proteins based on their isoelectric point (1D) and molecular mass (2D) (see **Figure 1.19**). The protein spots on the gel can then be detected and identified by either staining (avidin blotting for biotinylated probes) or in-gel fluorescence scanning (for fluorescent probes) and MS techniques. Unfortunately, these methods tend to lack good resolution. New generations of dyes and fluorescent techniques have been developed to improve the sensitivity of 2D electrophoresis; however, this method still lacks good resolution. Despite this, gel electrophoresis is still a common method which is still used today.^{84, 111, 112, 113}

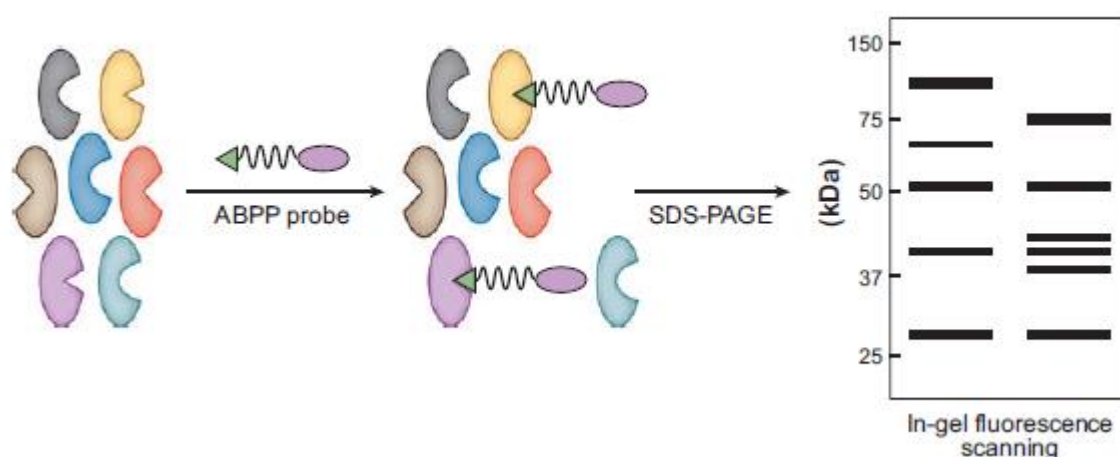


Figure 1.19: Gel based ABPP for fluorophore tagged chemical probes, which are visualised by in-gel fluorescence scanning. Reproduced from: *Activity-based protein profiling: from enzyme chemistry to proteomic chemistry*, B. F. Cravatt et al.⁸⁴

1.2.3.2 Liquid chromatography-mass spectrometry (LC-MS)

“Gel free” methods such as LC-MS have been developed to overcome the limitations of the gel electrophoresis method. Quantitative proteomics have emerged from MS methodology by analysing single intensities of proteins from mass spectral counts. There are multiple LC-

MS strategies for analysing probe treated proteomes, which can be divided into two main categories (i) analysis of the protein targets of the probes (ii) analysis of probe labelled peptides from these targets. The former category tends to be utilised for biotinylated probes. The labelled proteomes are incubated with (strept)avidin beads to enrich the probe labelled proteins. These enriched proteins are then digested on beads with trypsin and analysed by multidimensional LC-MS/MS. This methodology combines both ABPP and multidimensional protein identification (ABPP-MudPIT) (see **Figure 1.20**). This strategy enables researchers to identify the activities of individual proteomes.^{84, 111, 112}

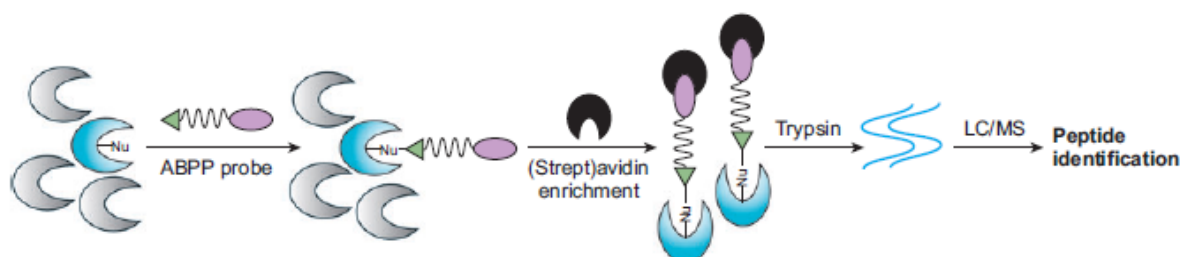


Figure 1.20: LC-MS based platform combining both ABPP and multidimensional protein identification (ABPP-MudPIT). Reproduced from: *Activity-based protein profiling: from enzyme chemistry to proteomic chemistry*, B. F. Cravatt et al.⁸⁴

LC-MS has its drawbacks as ABPP-MudPIT requires specialised equipment and is not a reliable method for comparative ABPP studies due to poor reproducibility. ABPP-MudPIT is also unable to provide topological and modification information about the proteome and is also unable to provide a simple way to identify probe labelled proteins. Other LC-MS platforms such as active site protein profiling (ASPP) has to be incorporated to provide this information (see **Figure 1.21**).^{84, 111, 114}

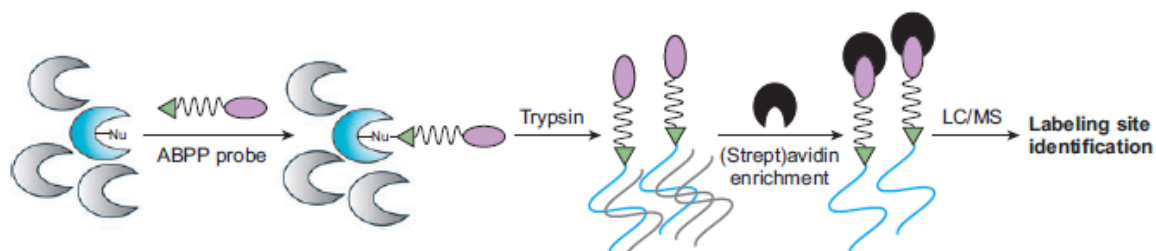


Figure 1.21: LC-MS based platform incorporating active site protein profiling. Reproduced from: *Activity-based protein profiling: from enzyme chemistry to proteomic chemistry*, B. F. Cravatt et al.⁸⁴

1.3 Benzimidazoles

1.3.1 Chemistry of Benzimidazoles

Benzimidazoles are a class of compounds that are composed of a bicyclic ring system in which benzene has been fused to the 4- and 5-position of an imidazole. The numbering system for the benzimidazole ring system is shown in **Figure 1.22**.^{115, 116}

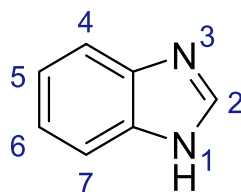


Figure 1.22: Benzimidazole structure numbering system.¹¹⁵

The two nitrogen atoms within the benzimidazole core possess both acidic and basic properties. The nitrogen in the 3 position can act as a base due to having an accessible lone pair of electrons which can accept a proton (see **Figure 1.23**). The nitrogen in the 1 position possesses a weakly acidic proton (pKa of 12.17 in H₂O), which can be deprotonated by a stronger base (see **Figure 1.24**).^{117, 118}

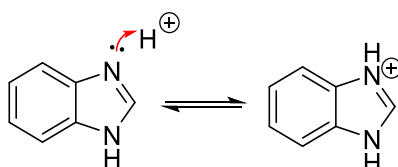


Figure 1.23: Benzimidazole exhibits basic properties when the nitrogen at the 3-position attacks a proton.

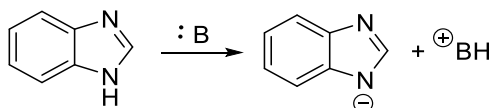


Figure 1.24: Benzimidazoles exhibit acidic behaviour when the nitrogen in the 1 position can be deprotonated *via* the addition of base.^{117, 118}

1.3.1.1 Tautomerism and Isomerism

There is a rapid exchange between the –NH– and =N– nitrogen atoms, and two tautomers can be drawn for the benzimidazole molecule (see **Figure 1.25**). Tautomerism can occur by either

an intermolecular process involving two or more benzimidazole molecules, or through interactions with a protic solvent. This results in the 5- and 6-positions and any group in that position in the ring system chemically equivalent.^{115, 119}

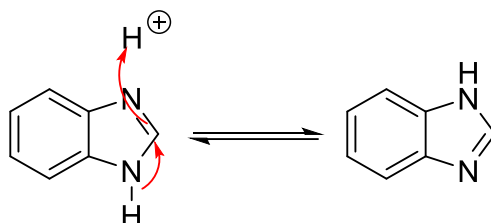


Figure 1.25: Rapid benzimidazole tautomerism between the -NH- and =N- nitrogen atoms.¹¹⁵

1.3.2 Biological Activity of Benzimidazoles

1.3.2.1 Tubulin and Microtubules

In every nucleated cell, two groups of spherical globular proteins known as alpha (α) and beta (β) tubulin are present. These two tubulin proteins can form to a α - β -tubulin heterodimer. Two molecules of energy-rich guanosine triphosphate (GTP) are bound to the heterodimers. The first GTP molecule is tightly bound to the heterodimer and will not willingly dissociate from the dimer. Whereas the second molecule freely exchanges with unbound GTP. Heterodimers can arrange themselves to form a protofilament, which is a long protein fibre composed of alternating α and β -tubulin. Typically, 12-13 protofilaments can associate to form a C-shaped protein sheet, which curls around to give a hollow cylindrical structure known as a microtubule (see **Figure 1.26**).^{120, 121, 122}

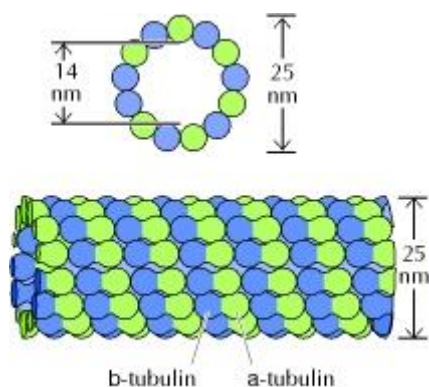


Figure 1.26: Cylindrical structure of a microtubule comprised of α - β -tubulin heterodimers.

Reproduced from: *The cell, A Molecular Approach*, G. M. Cooper.¹²¹

Microtubule organising centres (MTOCs) are situated on the end of a microtubule and are the site of microtubule growth. MTOCs aggregate to a third type of tubulin known as G-tubulin. Aggregation allows one end of the microtubule to be the site of nucleation for incoming α and β -tubulin heterodimers. There is a dynamic equilibrium between polymerisation and depolymerisation of the microtubule and the tubulin, as the dimers add to one end of the microtubule (known as the “plus” end) and leave at the other (known as the “minus” end).^{120,}

123

1.3.2.1.1 Role of Microtubules

Microtubules are essential for cell survival as they are required for cell division. During mitosis microtubules extend outward from the MTOC to form mitotic spindle, which is responsible for the separation and distribution of chromosomes in daughter cells (see **Figure 1.27**).^{120, 121,}

124

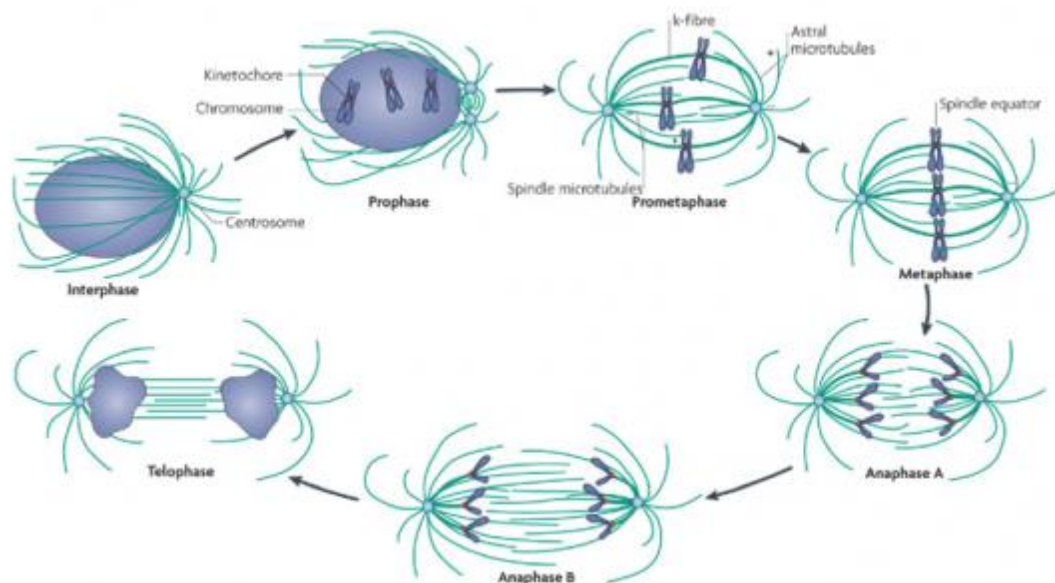


Figure 1.27: During mitosis, the microtubules form mitotic spindle by extending outwards from the MTOC to separate chromosomes for the formation of the daughter cells. Reproduced from:

*Mechanisms of chromosome behaviour during mitosis, C. E. Walczak et al.*¹²⁵

1.3.2.1.2 Inhibition of Microtubules

The dynamic equilibrium between the microtubules and tubulin can be altered by exogenous substances known as microtubule inhibitors. These substances typically inhibit microtubules *via* binding to the tubulin subunits which “caps” the microtubule as the polymerisation end (known as the “plus” end), while it continues to dissociate at the other end (known as the

“minus” end). This results in a net loss of the microtubule length, and therefore disrupts microtubule function, resulting in the destabilisation of the cell (see **Figure 1.28**).^{122, 124}

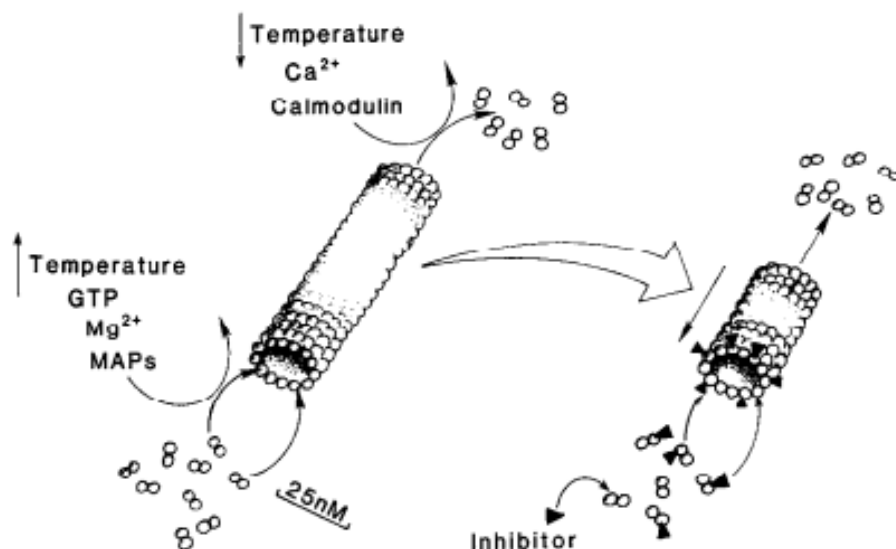


Figure 1.28: Microtubules are in dynamic equilibrium with tubulin and have several co-factors and conditions (such as temperature) to control both polymerisation and depolymerisation. The presence of a tubulin inhibitor “caps” the microtubule and allows depolymerisation to predominate over polymerisation leading to its destruction. Reproduced from: *Mode of Action of Benzimidazoles*, E. Lacey.¹²⁴

1.3.2.2 Current Uses of Benzimidazoles

Benzimidazoles have a number of indications as they have been shown to possess antimicrobial,^{126, 127} antitumor,^{128, 129} antihypertensive,^{130, 131} antiulcer,^{132, 133} antihistaminic,¹³⁴ anti-inflammatory,¹³⁵ antiviral,¹³⁶ and anthelmintic properties.¹³⁷ There are several positions in which the benzimidazole nucleus can be substituted, however the most biologically active benzimidazole based compounds tend to bear functional groups at the 2, 3 and or 5 position (see **Figure 1.29**). These compounds can be mono-, di-, or tri-substituted derivatives.¹³⁸

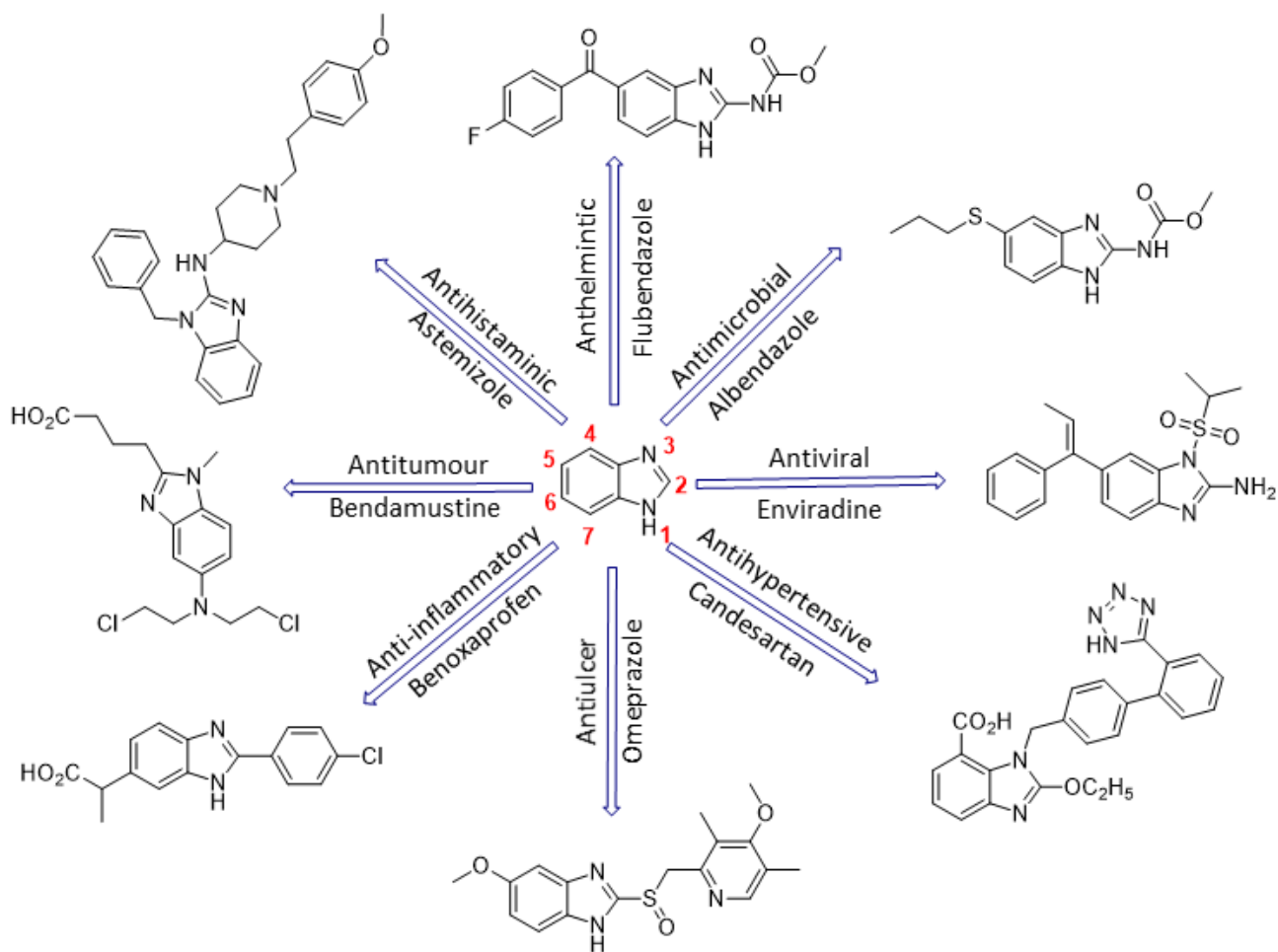


Figure 1.29: Benzimidazoles, a multifunctional nucleus displaying anthelmintic, antimicrobial, antihypertensive, antiulcer, anti-inflammatory, antitumor, and antihistaminic properties.¹³⁸

Currently the most common use of benzimidazole carbamates such as albendazole is to treat helminthic infections. The first commercially available benzimidazole was thiabendazole which was introduced to treat humans and animals such as livestock. A frequent helminthic disease is intestinal ascariasis, which is an infection of the small intestine caused by the common roundworm *Ascaris lumbricoides*. *Conterno et al* found that benzimidazoles such as mebendazole (100 mg twice a day for 3 days or 500 mg once) (see **Figure 1.30A**) and albendazole (400 mg in a single dose) (see **Figure 1.30B**) are effective in treating intestinal ascariasis, with cure rates greater than 90%.¹³⁹ Benzimidazoles have also been used as a protective barrier on crops such as fruit to prevent fungal growth.^{140, 141} Ambiol® (2-methyl-4-[dimethylaminomethyl]-5-hydroxybenzimidazole) (see **Figure 1.30C**) is often used by farmers to precondition seeds to improve germination, growth, and drought tolerance in many different species. *Rajasekaran et al* found that treating carrot seeds with 0.1 mg/L Ambiol increased germination rates by nearly fourfold.¹⁴² This pattern was also observed with

tomatoes as treatment for Ambiol increase germination rates and percentage of tomato seeds.¹⁴³

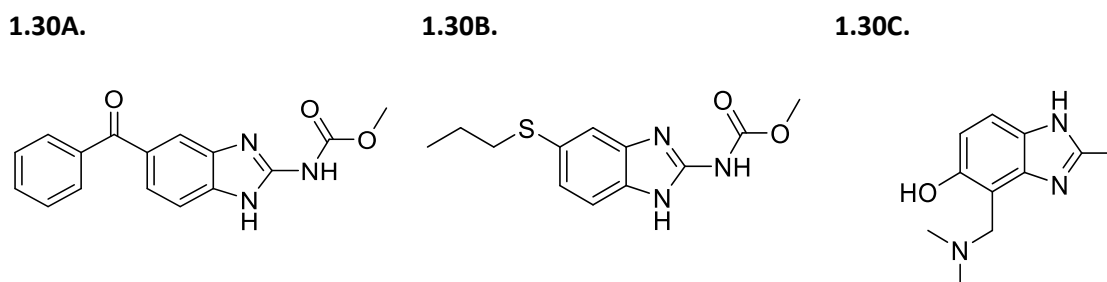


Figure 1.30: Structures of Benzimidazoles: **(1.30A)** Mebendazole and **(1.30B)** Albendazole used to treat intestinal ascariasis. **(1.30C)** Ambiol[®] used to increase crop germination and growth.

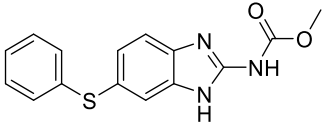
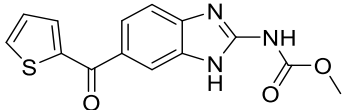
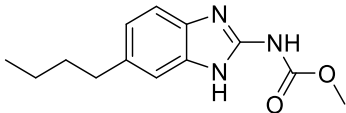
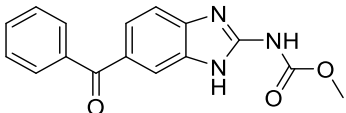
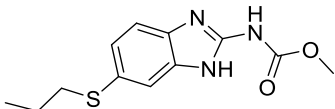
1.3.2.3 Anti-cryptococcal Activity of Benzimidazoles

Benzimidazoles such as flubendazole and albendazole exert their antifungal activity by binding to the tubulin of microtubules in fungal cells. The activity of benzimidazoles as microtubule inhibitors was first indicated in studies on fungi with the prodrug benomyl, and its active metabolite carbenazim (MBC). Cytological studies with MBC, benomyl, and other MBC prodrugs in mammalian and rat cell lines confirmed the antimitotic and microtubule inhibitory actions of benzimidazoles.¹⁴⁴ Work by Cruz *et al* found that *C. neoformans* displays *in vitro* sensitivity to benzimidazoles.¹⁴⁵ Further work within the Cruz group involved the sequencing of tubulin genes to determine selectivity.¹⁴⁶

There are two genes responsible for the formation of β -tubulin: *TUB1* and *TUB2*. Analysis of the *TUB1* and *TUB2* gene sequences show that *TUB1* has a large similarity to β -tubulin with other fungi and animals, whereas *TUB2* is comparatively more divergent. The origin of *C. neoformans* sensitivity to benzimidazoles was explored with mutational studies using a range of different fungi. These studies showed that there are six β -tubulin residues responsible for benzimidazole activity; His6, Ala165, Phe167, Glu198, Phe200, and Arg241. *C. neoformans* *TUB1* and *TUB2* sequences were compared to β -tubulin sequences from benzimidazole resistant and sensitive organisms. Residues Phe200 and Glu198 which are predictive of benzimidazole sensitivity in a range of species (helminths, fungi and protozoa) are only present in *TUB1*. This suggests that the *TUB1* sequence is responsible for benzimidazole sensitivity.¹⁴⁶

1.3.2.3.1 Benzimidazole Cryptococcal Data

The antifungal activity of eleven benzimidazole derivatives were compared with AmB against various isolates of *C. neoformans* to determine the inhibitory concentrations (IC) at both 50% and 90% inhibition (see **Table 1.1**). IC is the amount of benzimidazole carbamate needed to inhibit the growth of the *C. neoformans* isolates by 50% and 90%.

Drug	IC($\mu\text{g/mL}$) for indicated isolate(s)					
	IU4		IU8698		ATCC 14116	
	50%	90%	50%	90%	50%	90%
Fenbendazole 	0.019	0.028	0.011	0.014	<0.016	<0.016
Nocodazole 	0.13	0.22	0.09	0.12	0.20	0.29
Parbendazole 	0.16	0.23	0.016	0.23	0.16	0.22
Mebendazole 	0.23	0.43	0.18	0.32	0.19	0.40
Albendazole 	0.16	0.25	0.31	0.45	0.30	0.45

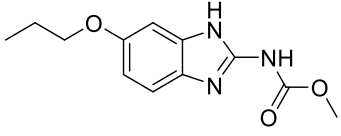
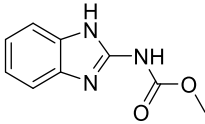
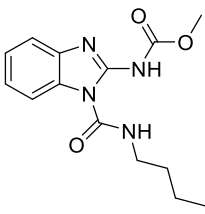
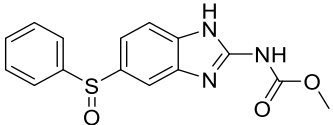
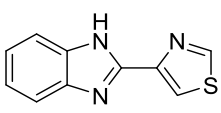
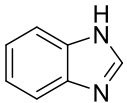
Oxibendazole 	2.3	3.2	1.0	2.1	1.4	2.1
Carbendazim 	2.0	3.6	1.5	2.0	1.6	2.3
Benomyl 	2.3	3.7	2.8	3.8	2.0	3.5
Oxfendazole 	2.5	4.0	3.2	4.0	2.8	4.0
Thiabendazole 	>4.0	>4.0	>4.0	>4.0	>4.0	>4.0
Benzimidazole 	>4.0	>4.0	>4.0	>4.0	>4.0	>4.0
Amphotericin B	0.035	0.064	0.024	0.065	0.011	0.045

Table 1.1: Inhibitory concentrations for a variety of benzimidazole compounds tested against three clinical isolates IU4, IU8698 and ATCC 14116. IC₅₀ and IC₉₀ values were estimated from plots of cell number (percent control) versus drug concentration. Reproduced from: *In vitro susceptibility of the opportunistic fungus Cryptococcus neoformans to anthelmintic benzimidazoles*, M. Cruz et al.¹⁴⁵

C. neoformans is shown to be relatively resistant to both thiabendazole and benzimidazole, even at its highest concentration. However, both thiabendazole and benzimidazole do not possess a carbamate group in the C2 position, making them structurally distinct from many of the other benzimidazole derivatives listed in the table. This suggests that the presence of a carbamate is essential for its activity against *C. neoformans*. Oxibendazole and oxfendazole showed mild activity, whereas albendazole, parabendazole, mebendazole and nocodazole showed moderate activity against *C. neoformans*. The albendazole analogue, fenbendazole was highly active and was two-fold more active than AmB. Oxfendazole, which is the sulfoxide analogue of fenbendazole displays poor antifungal activity. Sulfur containing benzimidazoles such as fenbendazole and albendazole can undergo metabolic mediated sulfur oxidation, to form inactive antifungal agents.^{147, 148} This has been confirmed by in-house testing of synthesised albendazole sulfoxide and albendazole sulfone (see **Figure 1.31**). Albendazole sulfoxide displayed weak antifungal activity, whereas albendazole sulfone had no activity against *C. neoformans*. Benomyl is structurally distinct from the other selected benzimidazoles, as it has a carbamate group attached at the C1 position. This attachment does not aid antifungal activity against *C. neoformans*, suggesting that the nitrogen in the C1 position may interact with amino acid residues within the binding site of *C. neoformans*. The data provides valuable but limited information about the SAR involved in the benzimidazole derivatives.

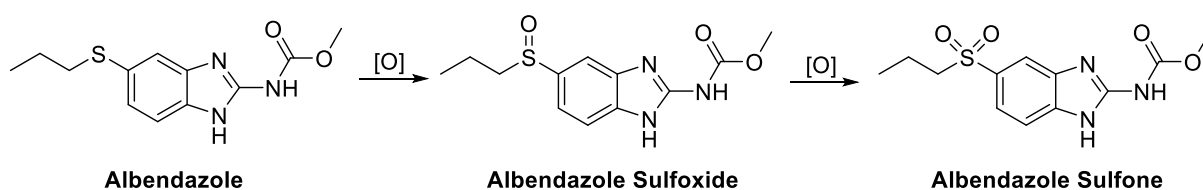


Figure 1.31: Albendazole sulfur oxidation to albendazole sulfoxide and albendazole sulfone. [O]: S-oxidation.¹⁴⁸

In-house testing to determine the minimum inhibitory concentration required to inhibit the growth of 50% of organisms (MIC_{50}) was measured on currently available agents against a *C. neoformans* isolate (ATCC 208821). Two benzimidazole derivatives; albendazole and flubendazole were shown to have better *in vitro* potency against *C. neoformans* compared to the current treatment (AmB) for cryptococcosis (see **Table 1.2**). Flubendazole displays the

greatest level of inhibition and was considerably more potent than the current treatments against CEM.

Drug	<i>C. neoformans</i> MIC ₅₀ (µg/mL)
Amphotericin B	1
5-Flucytosine	4
Fluconazole	8
Flubendazole	0.125
Albendazole	0.5

Table 1.2: *C. neoformans* MIC₅₀ of current drug agents (AmB) compared with benzimidazole derivatives. MIC₅₀ tested against H99 (ATCC 208821) isolate. Endpoint for MIC determination using EUCAST was 50%, n=3.

1.3.3 Identification of Benzimidazole Targets

The design and synthesis of photoaffinity labelled cc-ABPP probes can reveal benzimidazole targets within *C. neoformans*. The development of photoaffinity labelled probes to investigate tubulin inhibitors has already been established among different research groups across the globe to identify various binding sites. However, ABPP may reveal other targets such as enzymes.

1.3.3.1 Previous Established β-Tubulin Probes

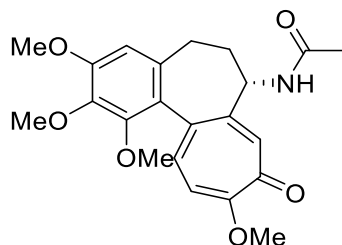
There are many different classes of compounds which exhibit microtubule depolymerisation mechanisms, such as colchicine and phenylalanine (PLH). These compounds have undergone ABPP experiments to identify their binding sites.

1.3.3.1.1 Colchicine

Colchicine (see **Figure 1.32A**) which is primarily used to treat gout possesses anti-microtubule activity. Researchers performed ABPP experiments by synthesising a range of photoaffinity labelled colchicine probes. Different photoaffinity labels have been utilised, such as the aryl azido moiety (see **Figure 1.32B**). Reports demonstrated that colchicine could bind to both the

α and β subunits of tubulin. However, photo labelled colchicine analogues displayed a preference for the β -tubulin subunit.^{149, 150}

1.32A.



1.32B.

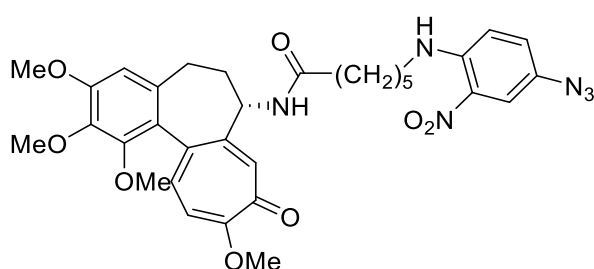


Figure 1.32: Chemical structures of colchicine and its photoaffinity derivative, **1.32A:** Colchicine, **1.32B:** 6'-(4'-Azido-2'-nitrophenylamino)hexanoyldeacetylcolchicine.¹⁵¹

1.3.3.1.2 Phenylahistin

PLH, which is a natural diketopiperazine (DKP) (see **Figure 1.33**) has been suggested to induce a microtubule depolymerisation mechanism. Until recently the precise binding site of PLH was unknown. Photoaffinity labelling has been employed to investigate the interaction between PLH and tubulin. Biotin has been the photoaffinity moiety of choice for these experiments. It was found that PLH can specifically and covalently bind to β -tubulin. PLH also displayed greater selectivity for β -tubulin over α tubulin.^{152, 153, 154}

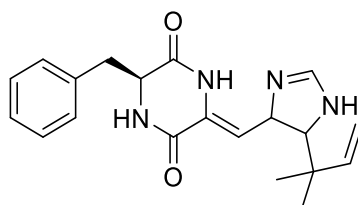


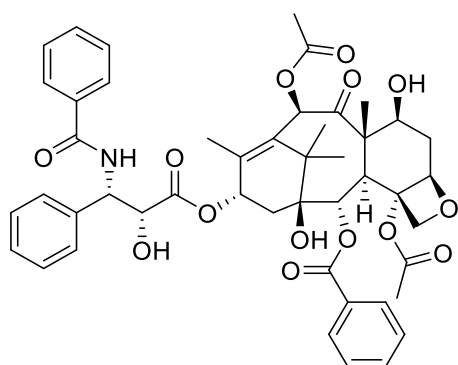
Figure 1.33: Structure of the natural diketopiperazine phenylahistin

1.3.3.1.3 Paclitaxel

Paclitaxel possesses antitumor activity by promoting microtubule assembly and stabilisation. To understand the interaction between paclitaxel and microtubules, *Dasgupta et al*

performed photolabeling studies with an azido analogue of paclitaxel (see **Figure 1.34**).¹⁵⁵ The group had previously demonstrated that the probe was a good analogue of paclitaxel, as it possessed 50% of the activity of paclitaxel.¹⁵⁶ Results from the photolabeling study showed that most of the probe was covalently incorporated in the β -tubulin subunit (80%), with 20% interacting with the α -subunit. The preference of the β -tubulin subunit was further confirmed by *Rao et al*, who found that the probe exclusively interacts with β -tubulin subunit.¹⁵⁷ The difference in binding patterns have been attributed to *Rao et al* performing photo labelled experiments in the presence of microtubule associated proteins and/ or used autoradiography to detect labelling.

1.34A



1.34B

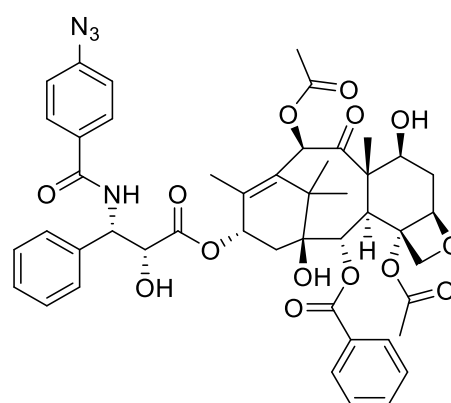


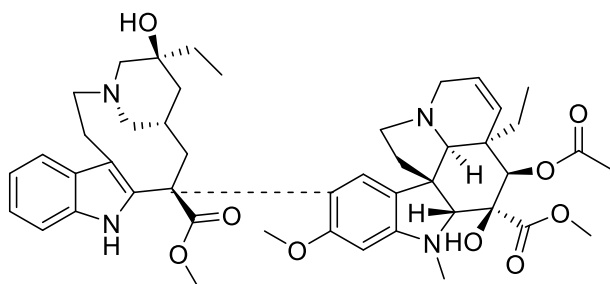
Figure 1.34: Structures of paclitaxel (**1.34A**) and its corresponding azido analogue probe (**1.34B**).¹⁵²

1.3.3.1.4 Vinblastine

Vinblastine is a Vinca alkaloid that possesses antineoplastic activity by interacting with microtubules. Vinca alkaloids have been reported to bind to tubulin dimers, which in turn disrupts microtubules in cells.¹⁵⁸ However, different biological effects occur depending on the concentration of Vinca alkaloids. At low concentrations, vinblastine can inhibit microtubule assembly. At higher concentrations, vinblastine can cause tubulin aggregation. Therefore, it was hypothesised that Vinca alkaloids have two or more binding sites on tubulin. *Safa et al* synthesised a photoactive radioactive analogue of vinblastine; *N*-(*p*-azido[3,5-³H]benzoyl)-*N'*-(β -aminoethyl)vindesine (³H]NABV) (see **Figure 1.35**).¹⁵⁹ The research group established that the *in vitro* interactions of tubulin with NABV and vinblastine were almost identical, ensuring it is suitable as a probe to study the binding sites/s of Vinca alkaloids. *Safa et al* found that vinblastine and NABV inhibited each other's binding to tubulin, although,

neither were inhibited by a different tubulin binding agent, colchicine. In the photolabeling study, [^3H]NABV was shown to covalently bind to both β and α -tubulin. When the concentration of [^3H]NABV was increased, the probe exhibited a biphasic pattern characteristic of specific and nonspecific reactions. The photolabeling study showed that there was an α : β ratio of 3:2 of binding. Thus, there are two distinct binding sites for Vinca alkaloids, one on each tubulin subunits, or a single shared by the two subunits. However, other research groups have suggested two or more binding sites.^{160, 161}

1.35A



1.35B

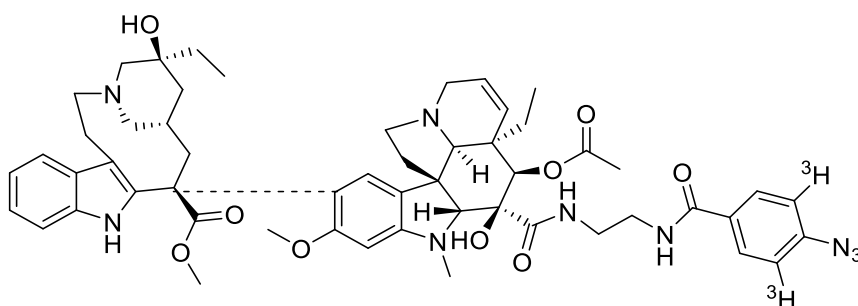


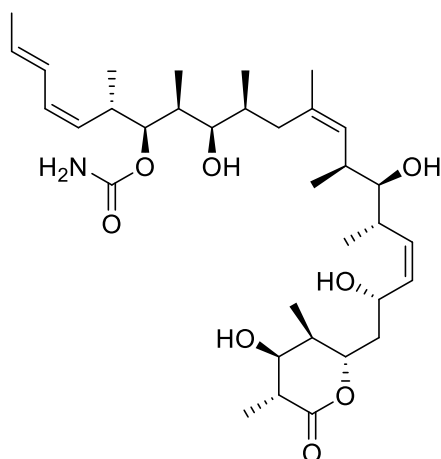
Figure 1.35: Structures of vinblastine (**1.35A**) and photoactive radioactive analogue of vinblastine (**1.35B**).¹⁵⁶

1.3.3.1.5 Discodermolide

Discodermolide is an experimental antitumor agent. Its mechanism of action is like paclitaxel in that it involves the stabilisation of microtubules, causing cell arrest and the G2/M phase of the cell cycle. *Xia et al* performed photoaffinity labelling studies to understand the interaction of the drug with microtubules. The research group synthesised C19-[4-(4- ^3H -benzoyl-phenyl)-carbamate]-discodermolide (C19-[^3H]BPCdiscodermolide) (see **Figure 1.36**). Binding studies showed that C19-BPC-discodermolide can effectively displace [^3H]paclitaxel from microtubules, just like the parent compound discodermolide. However, paclitaxel cannot

effectively displace C19-³H]BPC-discodermolide binding. They found that discodermolide analogues specifically target tubulin, as a radioactive band is localised between the α and β -tubulin subunit, like the binding pattern of paclitaxel. The group also found that when C19-BPC-discodermolide is hydrolysed by formic acid and Asp-N and Arg-C enzymes, it can bind to the β -tubulin amino acid residues near the paclitaxel binding site.^{162, 163}

1.36A



1.36B

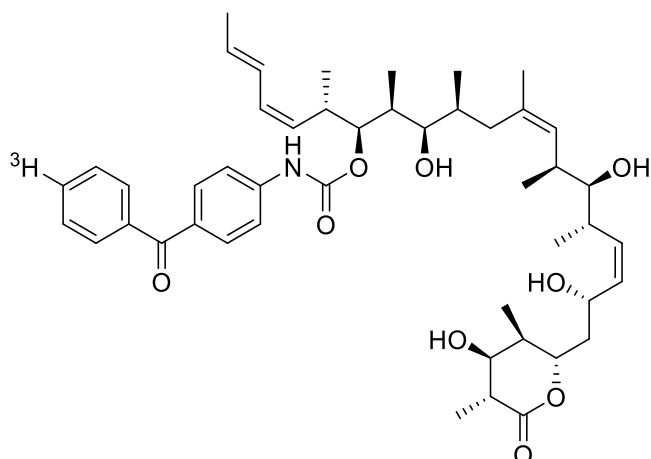


Figure 1.36: Structures of discodermolide (**1.36A**) and C19-³H]BPCdiscodermolide (**1.36B**).¹⁶⁰

1.3.3.2 Alternative Benzimidazole Targets

Over time, researchers have reported a diverse range of targets in which benzimidazoles can interact with. They can exert antiviral, antimicrobial, anti-inflammatory, and anti-cancer properties, by interacting with a various receptors, pathways, and enzymes. Benzimidazole derivatives behaving as enzyme inhibitors are frequently mentioned in the literature.^{164, 165}

Benzimidazole derivatives have been shown to possess antimicrobial properties, as they can inhibit a class of enzymes known as DNA gyrase. These enzymes belong to the Type II topoisomerases, which can disrupt the bacterial replication cycle, by halting the reunion of two strands of DNA after replication and relieving DNA supercoiling. Benzimidazoles are a promising solution to address the growing concern over bacterial resistance.^{166, 167}

Benzimidazole derivatives also display anti-tumour properties as they exhibit multi-target receptor tyrosine kinase (RTK) inhibition. RTKs are involved in the proliferation and survival of tumour cells. Therefore, inhibition of this class of enzyme results in the destruction of cancer cells.^{165, 168}

More specific to *C. neoformans* is the enzyme topoisomerase I (TOP1). A report shows that the cloning of the TOP1 gene which codes for the topoisomerase I in *C. neoformans* reveals that the fungus contains a fungal insert found in topoisomerase I from *Candida albicans* and *Saccharomyces cerevisiae* that is not present in the mammalian enzyme. This distinction between fungal and mammalian TOP1 leads to selectivity and therefore could represent a novel target for antifungal agents.¹⁶⁹

Previous and current research has reported alternative targets for benzimidazoles. Therefore, exposure of benzimidazole derivatives to *C. neoformans* may reveal several different pathways within the fungus, as opposed to just microtubule depolymerisation. PAL studies will help reveal whether other targets are interacting with our benzimidazole derivatives causing the desired anti-fungal activity.

1.4 References

- 1 T. G. Mitchell and J. R. Perfect, *Clin. Microbiol. Rev.*, 1995, **8**, 515–548.
- 2 K. Nielsen, G. M. Cox, A. P. Litvintseva, E. Mylonakis, S. D. Malliaris, D. K. Benjamin, S. S. Giles, T. G. Mitchell, A. Casadevall, J. R. Perfect and J. Heitman, *Infect. Immun.*, 2005, **73**, 4922–4933.
- 3 M. G. R. Duane R. Hospenthal, in *Diagnosis and Treatment of Human Mycoses*, Humana Press, Totowa, NJ, 2008, pp. 255–276.
- 4 K. J. Kwon-Chung, J. A. Fraser, T. Á. L. Doering, Z. A. Wang, G. Janbon, A. Idnurm and Y. S. Bahn, *Cold Spring Harb. Perspect. Med.*, 2014, **4**, 1–27.
- 5 K. J. Kwon-Chung, *Mycologia*, 1976, **68**, 942–946.
- 6 M. Gary, M Cox, J. R Perfect, Microbiology and epidemiology of *Cryptococcus neoformans* infection, <https://www.uptodate.com/contents/microbiology-and-epidemiology-of-cryptococcus-neoformans-infection>, (Accessed February 2020).
- 7 T. C. Sorrell and D. H. Ellis, *Rev Iberoam Micol*, 1997, **14**, 42–43.
- 8 B. L. Wickes, M. E. Mayorga, U. Edman and J. C. Edman, *Proc. Natl. Acad. Sci.*, 1996, **93**, 7327–7331.
- 9 J.-M. Vreulink, T. Boekhout, H. Vismer and A. Botha, *Fungal Ecol.*, 2020, **47**, 100943–100945.
- 10 J. Heitman, *Curr. Biol.*, 2006, **16**, 711–725.
- 11 S. A. Schmalzle, U. K. Buchwald, B. L. Gilliam and D. J. Riedel, *Mycoses*, 2016, **59**, 542–552.
- 12 W. Elsegeiny, K. A. Marr and P. R. Williamson, *Front. Immunol.*, 2018, **9**, 1–9.
- 13 N. Singh, B. D. Alexander, O. Lortholary, F. Dromer, K. L. Gupta, G. T. John, R. del Busto, G. B. Klintmalm, J. Somani, G. M. Lyon, K. Pursell, V. Stosor, P. Muñoz, A. P. Limaye, A. C. Kalil, T. L. Pruett, J. Garcia-Diaz, A. Humar, S. Houston, A. A. House, D. Wray, S. Orloff, L. A. Dowdy, R. A. Fisher, J. Heitman, M. M. Wagener and S. Husain, *J. Infect. Dis.*, 2007, **195**, 756–764.
- 14 S. Antinori, *ISRN AIDS*, 2013, **2013**, 1–22.
- 15 A. A. Okoye and L. J. Picker, *Immunol. Rev.*, 2013, **254**, 54–64.
- 16 B. J. Park, K. A. Wannemuehler, B. J. Marston, N. Govender, P. G. Pappas and T. M. Chiller, *Aids*, 2009, **23**, 525–530.
- 17 S. Rohatgi and L. Pirofski, *Future Microbiol.*, 2015, **10**, 565–581.
- 18 F. Dromer, S. Mathoulin-Pélissier, A. Fontanet, O. Ronin, B. Dupont and O. Lortholary, *AIDS*, 2004, **18**, 555–562.
- 19 J. E. Kaplan, D. Hanson, M. S. Dworkin, T. Frederick, J. Bertolli, M. Lou Lindegren, S. Holmberg and J. L. Jones, *Clin. Infect. Dis.*, 2000, **30**, S5–S14.

- 20 S. A. Mirza, M. Phelan, D. Rimland, E. Graviss, R. Hamill, M. E. Brandt, T. Gardner, M. Sattah, G. Ponce de Leon, W. Baughman and R. A. Hajjeh, *Clin. Infect. Dis.*, 2003, **36**, 789–794.
- 21 J. E. Vidal and D. R. Boulware, *Rev. Inst. Med. Trop. Sao Paulo*, 2015, **57**, 38–45.
- 22 A. Kamari, A. Sepahvand and R. Mohammadi, *Curr. Med. Mycol.*, 2017, **3**, 20–25.
- 23 J. R. Perfect, *FEMS Yeast Res.*, 2006, **6**, 463–468.
- 24 J. D. Nosanchuk, S. Shoham, B. C. Fries, D. S. Shapiro, S. M. Levitz and A. Casadevall, *Ann. Intern. Med.*, 2000, **132**, 205–208.
- 25 K. Lagrou, J. V Eldere, S. Keuleers, F. Hagen, R. Merckx, J. Verhaegen, W. E. Peetermans and T. Boekhout, *J. Intern. Med.*, 2005, **257**, 385–388.
- 26 N. M. Walsh, M. R. Botts, A. J. McDermott, S. C. Ortiz, M. Wüthrich, B. Klein and C. M. Hull, *PLOS Pathog.*, 2019, **15**, 1–30.
- 27 S. Mukherjee, M. Feldmesser and A. Casadevall, *J. Infect. Dis.*, 1996, **173**, 1222–1231.
- 28 O. Zaragoza, M. L. Rodrigues, M. De Jesus, S. Frases, E. Dadachova and A. Casadevall, *Adv Appl Microbiol.*, 2009, **68**, 133–216.
- 29 Y. Wang, P. Aisen and A. Casadevall, *Infect. Immun.*, 1995, **63**, 3131–3136.
- 30 A. Casadevall, C. Coelho and A. Alanio, *Front. Immunol.*, 2018, **9**, 1–8.
- 31 K. L. Buchanan and J. W. Murphy, *Emerg. Infect. Dis.*, 1998, **4**, 71–83.
- 32 E. W. Petzold, U. Himmelreich, E. Mylonakis, T. Rude, D. Toffaletti, G. M. Cox, J. L. Miller and J. R. Perfect, *Infect. Immun.*, 2006, **74**, 5877–5887.
- 33 H. C. Eisenman, S.-K. Chow, K. K. Tsé, E. McClelland and A. Casadevall, *Virulence*, 2011, **2**, 329–336.
- 34 K. Hartland, J. Pu, M. Palmer, S. Dandapani, P. N. Moquist, B. Munoz, L. Didone, S. L. Schreiber and D. J. Krysan, *ACS Infect. Dis.*, 2016, **2**, 93–102.
- 35 M. S. Saag, R. J. Graybill, R. A. Larsen, P. G. Pappas, J. R. Perfect, W. G. Powderly, J. D. Sobel and W. E. Dismukes, *Clin. Infect. Dis.*, 2000, **30**, 710–718.
- 36 R. A. Murphy, H. Sunpath, B. Taha, S. Kappagoda, K. T. M. Maphasa, D. R. Kuritzkes and L. Smeaton, *Int. J. Tuberc. Lung Dis.*, 2010, **14**, 903–908.
- 37 D. R. Boulware, D. B. Meya, C. Muzoora, M. A. Rolfes, K. Huppler Hullsiek, A. Musubire, K. Taseera, H. W. Nabeta, C. Schutz, D. A. Williams, R. Rajasingham, J. Rhein, F. Thienemann, M. W. Lo, K. Nielsen, T. L. Bergemann, A. Kambugu, Y. C. Manabe, E. N. Janoff, P. R. Bohjanen and G. Meintjes, *N. Engl. J. Med.*, 2014, **370**, 2487–2498.
- 38 R. A. Murphy, T. J. Hatlen and M. Y. S. Moosa, *AIDS Res. Hum. Retroviruses*, 2018, **34**, 399–403.
- 39 T. Chen, L. Mwenge, S. Lakhi, D. Chanda, P. Mwaba, S. F. Molloy, A. Gheorghe, U. K. Griffiths, R. S. Heyderman, C. Kanyama, C. Kouanfack, S. Mfinanga, A. K. Chan, E.

- Temfack, S. Kivuyo, M. C. Hosseinipour, O. Lortholary, A. Loyse, S. Jaffar, T. S. Harrison and L. W. Niessen, *Clin. Infect. Dis.*, 2019, **69**, 588–595.
- 40 B. O. Adeyemi and A. Ross, *African J. Prim. Heal. Care Fam. Med.*, 2014, **6**, 672–680.
- 41 M. A. Rolfes, K. H. Hullsiek, J. Rhein, H. W. Nabeta, K. Taseera, C. Schutz, A. Musubire, R. Rajasingham, D. A. Williams, F. Thienemann, C. Muzoora, G. Meintjes, D. B. Meya and D. R. Boulware, *Clin. Infect. Dis.*, 2014, **59**, 1607–1614.
- 42 M. A. Rolfes, J. Rhein, C. Schutz, K. Taseera, H. W. Nabeta, K. Huppler Hullsiek, A. Akampura, R. Rajasingham, A. Musubire, D. A. Williams, F. Thienemann, P. R. Bohjanen, C. Muzoora, G. Meintjes, D. B. Meya and D. R. Boulware, *Open Forum Infect. Dis.*, 2015, **2**, ofv157-165.
- 43 M. Kartalija, M., Kaye, K., Tureen, J., Liu, Q., Täuber, M., Elliott, B., & Sande, *J. Infect. Dis.*, 1996, **173**, 1216–1221.
- 44 H. A. Gallis, R. H. Drew and W. W. Pickard, *Clin. Infect. Dis.*, 1990, **12**, 308–329.
- 45 J. M. T. Hamilton Miller, *Bacteriol. Rev.*, 1973, **37**, 166–196.
- 46 A. C. Mesa-Arango, L. Scorzoni and O. Zaragoza, *Front. Microbiol.*, 2012, **3**, 1–10.
- 47 S. D. Goodwin, J. D. Cleary, C. A. Walawander, J. W. Taylor and T. H. Grasela, *Clin. Infect. Dis.*, 1995, **20**, 755–761.
- 48 R. Laniado-Laborín and M. N. Cabrales-Vargas, *Rev. Iberoam. Micol.*, 2009, **26**, 223–227.
- 49 A. Loyse, F. Dromer, J. Day, O. Lortholary and T. S. Harrison, *J. Antimicrob. Chemother.*, 2013, **68**, 2435–2444.
- 50 F. C. Odds, A. J. P. Brown and N. A. R. Gow, *Trends Microbiol.*, 2003, **11**, 272–279.
- 51 W. D. Ashe and D. E. Van Reken, *Clin. Pediatr. (Phila.)*, 1977, **16**, 364–366.
- 52 J.M Benson et al, *Clin Pharm*, 1988, **7**, 424–438.
- 53 J. E. Bennett, *Ann. Intern. Med.*, 1977, **86**, 319.
- 54 A. M. Stamm, R. B. Diasio, W. E. Dismukes, S. Shadomy, G. A. Cloud, C. A. Bowles, G. H. Karam and A. Espinel-Ingroff, *Am. J. Med.*, 1987, **83**, 236–242.
- 55 C. A. Kauffman and P. T. Frame, *Antimicrob. Agents Chemother.*, 1977, **11**, 244–247.
- 56 A. Vermes, H. J. Guchelaar and J. Dankert, *J. Antimicrob. Chemother.*, 2000, **46**, 171–179.
- 57 K. Vu, G. R. Thompson, C. Roe, J. Sykes, E. Dreibe, S. Lockhart, W. Meyer, D. M. Engelthaler and A. Gelli, *Open Forum Infect. Dis.*, 2017, **4**, S115–S116.
- 58 T. Noe, F. Franc, P. Paumard, C. Chastin, D. Bre and J. Villard, *Society*, 2003, **47**, 1275–1284.
- 59 D. J. Sloan, M. J. Dedicoat and D. G. Lalloo, *Curr. Opin. Infect. Dis.*, 2014, **22**, 455–463.
- 60 B. Myers, I. S. Africa and S. Africa, *Lancet Oncol.*, 2005, **5**, 530–531.

- 61 V. S. Pore, S. G. Agalave, P. Singh, P. K. Shukla, V. Kumar and M. I. Siddiqi, *Org. Biomol. Chem.*, 2015, **13**, 6551–6561.
- 62 J. H. Rex, M. G. Rinaldi and M. A. Pfaller, *Antimicrob. Agents Chemother.*, 1995, **39**, 1–8.
- 63 M. E. Brandt, M. A. Pfaller, R. A. Hajjeh, R. J. Hamill, P. G. Pappas, A. L. Reingold, D. Rimland and D. W. Warnock, *Antimicrob. Agents Chemother.*, 2001, **45**, 3065–3069.
- 64 K. G. Davey, E. M. Johnson, A. D. Holmes, A. Szekely and D. W. Warnock, *J. Antimicrob. Chemother.*, 1998, **42**, 217–220.
- 65 N. H. Georgopapadakou, *Curr. Opin. Microbiol.*, 1998, **1**, 547–557.
- 66 S. Yu, X. Chai, H. Hu, Y. Yan, Z. Guan, Y. Zou, Q. Sun and Q. Wu, *Eur. J. Med. Chem.*, 2010, **45**, 4435–4445.
- 67 W. Hope, N. R. H. Stone, A. Johnson, L. McEntee, N. Farrington, A. Santoro-Castelazo, X. Liu, A. Lucaci, M. Hughes, J. D. Oliver, C. Giamberardino, S. Mfinanga, T. S. Harrison, J. R. Perfect and T. Bicanic, *MBio*, 2019, **10**, e02575-19.
- 68 K. Venkateswarlu, M. Taylor, N. J. Manning, M. G. Rinaldi and S. L. Kelly, *Antimicrob. Agents Chemother.*, 1997, **41**, 748–51.
- 69 S. Perea, J. L. López-Ribot, W. R. Kirkpatrick, R. K. McAtee, R. A. Santillán, M. Martínez, D. Calabrese, D. Sanglard and T. F. Patterson, *Antimicrob. Agents Chemother.*, 2001, **45**, 2676–2684.
- 70 M. Chang, E. Sionov, A. Khanal Lamichhane, K. J. Kwon-Chung and Y. C. Chang, *Antimicrob. Agents Chemother.*, 2018, **62**, e01751-17.
- 71 L. Di, E. Kerns and G. Carter, *Curr. Pharm. Des.*, 2009, **15**, 2184–2194.
- 72 G. Nichols-English and S. Poirier, *J. Am. Pharm. Assoc.*, 2000, **40**, 475–85.
- 73 Editorial, *Nat. Med.*, 2010, **16**, 347.
- 74 Cerisier, Petitjean, Regad, Bayard, Réau, Badel and Camproux, *Molecules*, 2019, **24**, 2529.
- 75 P. K. Mukherjee, D. J. Sheehan, C. A. Hitchcock and M. A. Ghannoum, *Clin. Microbiol. Rev.*, 2005, **18**, 163–194.
- 76 H. D. and J. Munro, *Nat. Rev. Drug Discov.*, 2019, **18**, 495–496.
- 77 C. Hale, New MIT Study Puts Clinical Research Success Rate at 14 Percent, <https://www.centerwatch.com/articles/12702-new-mit-study-puts-clinical-research-success-rate-at-14-percent>, (Accessed February 2020).
- 78 E. Gregori-Puigjane, V. Setola, J. Hert, B. A. Crews, J. J. Irwin, E. Lounkine, L. Marnett, B. L. Roth and B. K. Shoichet, *Proc. Natl. Acad. Sci.*, 2012, **109**, 11178–11183.
- 79 I. Schlichting, in *Protein-Ligand Interactions*, Humana Press, New Jersey, 2005, vol. 305, pp. 155–166.
- 80 S. Maity, R. K. Gundampati and T. K. S. Kumar, *Nat. Prod. Commun.*, 2019, **14**, 1–17.

- 81 K. Rawal, T. Khurana, H. Sharma, S. Verma, S. Gupta, C. Kubba, U. Strych, P. J. Hotez and M. E. Bottazzi, *PeerJ Prepr.*, 2019, 1–177.
- 82 M. J. Chalmers, S. A. Busby, B. D. Pascal, G. M. West and P. R. Griffin, *Expert Rev. Proteomics*, 2011, **8**, 43–59.
- 83 A. E. Speers and B. F. Cravatt, *ChemBioChem*, 2004, **5**, 41–47.
- 84 B. F. Cravatt, A. T. Wright and J. W. Kozarich, *Annu. Rev. Biochem.*, 2008, **77**, 383–414.
- 85 A. E. Speers, G. C. Adam and B. F. Cravatt, *J. Am. Chem. Soc.*, 2003, **125**, 4686–4687.
- 86 S.-S. Ge, B. Chen, Y.-Y. Wu, Q.-S. Long, Y.-L. Zhao, P.-Y. Wang and S. Yang, *RSC Adv.*, 2018, **8**, 29428–29454.
- 87 E. Smith and I. Collins, *Future Med. Chem.*, 2015, **7**, 159–183.
- 88 A. McFedries, A. Schwaid and A. Saghatelian, *Chem. Biol.*, 2013, **20**, 667–673.
- 89 J. Sumranjit and S. Chung, *Molecules*, 2013, **18**, 10425–10451.
- 90 E. Leyva, D. de Loera and S. Leyva, *Tetrahedron Lett.*, 2008, **49**, 6759–6761.
- 91 J. Michalak, H. Bin Zhai and M. S. Platz, *J. Phys. Chem.*, 1996, **100**, 14028–14036.
- 92 M. M. Hassan and O. O. Olaoye, *Molecules*, 2020, **25**, 2285.
- 93 J. Huang, A. Russell, J. Tsarevsky, N. V. Matyjaszewski, US8349410, 2008, 233–238.
- 94 J. Brunner, H. Senn and F. M. Richards, *J. Biol. Chem.*, 1980, **255**, 3313–8.
- 95 J. R. Hill and A. A. B. Robertson, *J. Med. Chem.*, 2018, **61**, 6945–6963.
- 96 L. Dubinsky, B. P. Krom and M. M. Meijler, *Bioorg. Med. Chem.*, 2012, **20**, 554–570.
- 97 N. Yamamoto, F. Bernardi, A. Bottoni, M. Olivucci, M. A. Robb and S. Wilsey, *J. Am. Chem. Soc.*, 1994, **116**, 2064–2074.
- 98 Y. Chen and E. M. Topp, *J. Pharm. Sci.*, 2019, **108**, 791–797.
- 99 D. S. Ziemianowicz, R. Bomgarden, C. Etienne and D. C. Schriemer, *J. Am. Soc. Mass Spectrom.*, 2017, **28**, 2011–2021.
- 100 J. Wang, J. Kubicki, H. Peng and M. S. Platz, *J. Am. Chem. Soc.*, 2008, **130**, 6604–6609.
- 101 B. Noller, P. Hemberger, I. Fischer, C. Alcaraz, G. A. Garcia and H. Soldi-Lose, *Phys. Chem. Chem. Phys.*, 2009, **11**, 5384–5391.
- 102 K. T. Barglow and B. F. Cravatt, *Nat. Methods*, 2007, **4**, 822–827.
- 103 M. H. Wright and S. A. Sieber, *Nat. Prod. Rep.*, 2016, **33**, 681–708.
- 104 G. C. Cook, *Parasitol. Today*, 1990, **6**, 133–136.
- 105 A. E. Speers and B. F. Cravatt, *Curr. Protoc. Chem. Biol.*, 2009, **1**, 29–41.
- 106 J. Martell and E. Weerapana, *Molecules*, 2014, **19**, 1378–1393.
- 107 H. C. Kolb and K. B. Sharpless, *Res. Focus Rev.*, 2003, **8**, 401–422.

- 108 J. E. Moses and A. D. Moorhouse, *Chem. Soc. Rev.*, 2007, **36**, 1249–1262.
- 109 J. C. Jewett and C. R. Bertozzi, *Chem. Soc. Rev.*, 2010, **39**, 1272–1279.
- 110 J. Dommerholt, F. P. J. T. Rutjes and F. L. van Delft, *Top. Curr. Chem.*, 2016, **374**, 16.
- 111 S. A. Sieber and B. F. Cravatt, *Chem. Commun.*, 2006, **22**, 2311.
- 112 A. Galmozzi, E. Dominguez, B. F. Cravatt and E. Saez, in *Methods in Enzymology*, 2014, pp. 151–169.
- 113 S. H. L. Verhelst and M. Bogyo, *Biotechniques*, 2005, **38**, 175–177.
- 114 *J. Proteome Res.*, 2005, **4**, 1897–1899.
- 115 J. B. Wright, *Chem. Rev.*, 1951, **48**, 397–541.
- 116 Salahuddin, M. Shaharyar and A. Mazumder, *Arab. J. Chem.*, 2017, **10**, S157–S173.
- 117 T. J. Barbarich, P. F. Driscoll, S. Izquierdo, L. N. Zakharov, C. D. Incarvito and A. L. Rheingold, *Inorg. Chem.*, 2004, **43**, 7764–7773.
- 118 H. Sari and A. K. Covington, *J. Chem. Eng. Data*, 2005, **50**, 1425–1429.
- 119 F. Su, Z. Sun, W. Su and X. Liang, *J. Mol. Struct.*, 2018, **1173**, 690–696.
- 120 A. Jordan, J. A. Hadfield, N. J. Lawrence and A. T. McGown, *Med. Res. Rev.*, 1998, **18**, 259–296.
- 121 G. M. Cooper, in *The Cell: A Molecular Approach. 2nd edition*, Sinauer Associates, Sunderland (MA), 2000, p. Chapter 11.
- 122 J. J. Correia and R. C. Williams, *Annu. Rev. Biophys. Bioeng.*, 1983, **12**, 211–235.
- 123 A. D. Sanchez and J. L. Feldman, *Curr. Opin. Cell Biol.*, 2017, **44**, 93–101.
- 124 E. Lacey, *Parasitol. Today*, 1990, **6**, 112–115.
- 125 C. E. Walczak, S. Cai and A. Khodjakov, *Nat. Rev. Mol. Cell Biol.*, 2010, **11**, 91–102.
- 126 Y.-N. Wang, R. R. Y. Bheemanaboina, G.-X. Cai and C.-H. Zhou, *Bioorg. Med. Chem. Lett.*, 2018, **28**, 1621–1628.
- 127 Y. Özkay, Y. Tunalı, H. Karaca and İ. Işıkdag, *Arch. Pharm. Res.*, 2011, **34**, 1427–1435.
- 128 M. A. Omar, Y. M. Shaker, S. A. Galal, M. M. Ali, S. M. Kerwin, J. Li, H. Tokuda, R. A. Ramadan and H. I. El Diwani, *Bioorg. Med. Chem.*, 2012, **20**, 6989–7001.
- 129 M. Rashid, A. Husain, M. Shaharyar and M. Sarafroz, *Anticancer. Agents Med. Chem.*, 2014, **14**, 1003–1018.
- 130 M. T. Khan, M. T. Razi, S. U. Jan, M. Mukhtiar, R. Gul, - IzharUllah, A. Hussain, A. M. Hashmi, M. T. Ahmad, N. A. Shahwani and I. Rabbani, *Pak. J. Pharm. Sci.*, 2018, **31**, 1067–1074.
- 131 G. Rusu, G. Dănilă and M. Nechifor, *Rev. Med. Chir. Soc. Med. Nat. Iasi*, **97**, 269–71.
- 132 A. Noor, N. G. Qazi, H. Nadeem, A. Khan, R. Z. Paracha, F. Ali and A. Saeed, *Chem.*

- Cent. J.*, 2017, **11**, 85.
- 133 E. Cereda, M. Turconi, A. Ezhaya, E. Bellora, A. Brambilla, F. Pagani and A. Donetti, *Eur. J. Med. Chem.*, 1987, **22**, 527–537.
- 134 R. Iemura, T. Kawashima, T. Fukuda, K. Ito and G. Tsukamoto, *J. Heterocycl. Chem.*, 1987, **24**, 31–37.
- 135 M. Gaba, S. Singh and C. Mohan, *Eur. J. Med. Chem.*, 2014, **76**, 494–505.
- 136 M. Tonelli, G. Paglietti, V. Boido, F. Sparatore, F. Marongiu, E. Marongiu, P. La Colla and R. Loddo, *Chem. Biodivers.*, 2008, **5**, 2386–2401.
- 137 Q. A. McKeller and E. W. Scott, *J. Vet. Pharmacol. Ther.*, 1990, **13**, 223–247.
- 138 Y. Bansal and O. Silakari, *Bioorganic Med. Chem.*, 2012, **20**, 6208–6236.
- 139 L. O. Conterno, M. D. Turchi, I. Corrêa and R. A. Monteiro de Barros Almeida, *Cochrane Database Syst. Rev.*, 2020, **2020**, 1–99.
- 140 D. A. Denham, *Parasitol. Today*, 1990, **6**, 136.
- 141 W. C. Campbell, *Parasitol. Today*, 1990, **6**, 130–133.
- 142 R. Lada, A. Stiles, M. A. Surette, C. Caldwell, J. Nowak, A. V. Sturz and T. J. Blake, *Acta Hortic.*, 2004, **631**, 105–116.
- 143 M. T. MacDonald, R. R. Lada, A. R. Robinson and J. Hoyle, *J. Plant Growth Regul.*, 2010, **29**, 357–365.
- 144 E. Lacey, *Int. J. Parasitol.*, 1988, **18**, 885–936.
- 145 M. C. Cruz, M. S. Bartlett and T. D. Edlind, *Antimicrob. Agents Chemother.*, 1994, **38**, 378–380.
- 146 M. C. Cruz and T. Edlind, *Microbiology*, 1997, **143**, 2003–2008.
- 147 A. S. Surur, L. Schulig and A. Link, *Arch. Pharm. (Weinheim)*, 2018, **352**, 1–11.
- 148 M. H. la P. Olivia Soria-Arteche, Rafael Castillo, Alicia Hernández-Campos and K. G.-F. Gabriel Navarrete-Vázquez, José Luis Medina-Franco, *J. Mex. Chem. Soc.*, 2005, **49**, 353–358.
- 149 L. J. Floyd, L. D. Barnes and R. F. Williams, *Biochemistry*, 1989, **28**, 8515–8525.
- 150 S. Rao, S. B. Horwitz and I. Ringel, *J. Natl. Cancer Inst.*, 1992, **84**, 2820–2824.
- 151 R. F. Williams, C. L. Mumford, G. A. Williams, L. J. Floyd, M. J. Aivaliotis, R. A. Martinez, A. K. Robinson and L. D. Barnes, *J. Biol. Chem.*, 1985, **260**, 13794–1380.
- 152 L. G. Lis, M. A. Smart, A. Luchniak, M. L. Gupta and V. J. Gurvich, *ACS Med. Chem. Lett.*, 2012, **3**, 745–748.
- 153 Y. Yamazaki, K. Kohno, H. Yasui, Y. Kiso, M. Akamatsu, B. Nicholson, G. Deyanat-Yazdi, S. Neuteboom, B. Potts, G. K. Lloyd and Y. Hayashi, *ChemBioChem*, 2008, **9**, 3074–3081.

- 154 Y. Yamazaki, Y. Kido, K. Hidaka, H. Yasui, Y. Kiso, F. Yakushiji and Y. Hayashi, *Bioorganic Med. Chem.*, 2011, **19**, 595–602.
- 155 D. Dasgupta, H. Park, G. C. B. Harriman, G. I. Georg and R. H. Himes, *J. Med. Chem.*, 1994, **37**, 2976–2980.
- 156 G. I. Georg, Z. S. Cheruvallath, R. H. Himes, M. R. Mejillano and C. T. Burke, *J. Med. Chem.*, 1992, **35**, 4230–4237.
- 157 S. Rao, N. E. Krauss, J. M. Heerding, C. S. Swindell, I. Ringel, G. A. Orr and S. B. Horwitz, *J. Biol. Chem.*, 1994, **269**, 3132–3134.
- 158 M. N. George, P., Journey, L. J., & Goldstein, *J. Natl. Cancer Inst.*, 1965, **35**, 355–375.
- 159 A. R. Safa, E. Hamel and R. L. Felsted, *Biochemistry*, 1987, **26**, 97–102.
- 160 S. N. Lee, J. C., Harrison, D., & Timasheff, *J. Biol. Chem.*, 1975, **250**, 9276–9282.
- 161 L. Wilson, *Biochemistry*, 1975, **14**, 5586–5592.
- 162 J. H. Nettles, *Science (80-.)*, 2004, **305**, 866–869.
- 163 S. Xia, C. S. Kenesky, P. V. Rucker, A. B. Smith, G. A. Orr and S. B. Horwitz, *Biochemistry*, 2006, **45**, 11762–11775.
- 164 G. Kaur, M. Kaur and O. Silakari, *Mini-Reviews Med. Chem.*, 2014, **14**, 747–767.
- 165 X. Luan, C. Gao, N. Zhang, Y. Chen, Q. Sun, C. Tan, H. Liu, Y. Jin and Y. Jiang, *Bioorganic Med. Chem.*, 2011, **19**, 3312–3319.
- 166 R. Janupally, V. U. Jeankumar, K. A. Bobesh, V. Soni, P. B. Devi, V. K. Pulla, P. Suryadevara, K. S. Chennubhotla, P. Kulkarni, P. Yogeewari and D. Sriram, *Bioorganic Med. Chem.*, 2014, **22**, 5970–5987.
- 167 P. S. Charifson, A.-L. Grillot, T. H. Grossman, J. D. Parsons, M. Badia, S. Bellon, D. D. Deininger, J. E. Drumm, C. H. Gross, A. LeTiran, Y. Liao, N. Mani, D. P. Nicolau, E. Perola, S. Ronkin, D. Shannon, L. L. Swenson, Q. Tang, P. R. Tessier, S.-K. Tian, M. Trudeau, T. Wang, Y. Wei, H. Zhang and D. Stamos, *J. Med. Chem.*, 2008, **51**, 5243–5263.
- 168 H. T. Abdel-Mohsen, M. A. Abdullaziz, A. M. El Kerdawy, F. A. F. Ragab, K. J. Flanagan, A. E. E. Mahmoud, M. M. Ali, H. I. El Diwani and M. O. Senge, *Molecules*, 2020, **25**, 770.
- 169 M. Del Poeta, D. L. Toffaletti, T. H. Rude, C. C. Dykstra, J. Heitman and J. R. Perfect, *Genetics*, 1999, **152**, 167–178.

CHAPTER 2

Structural Activity Relationship Exploration of
Flubendazole Derivatives.

2.1 Benzimidazoles as Antifungals

The two benzimidazoles flubendazole and albendazole display promising antifungal activity against *C. neoformans*, with *in vitro* MIC values of 0.125 and 0.5 mg/L.

An *in vivo* mouse study to measure how fungal density altered with different doses of flubendazole was performed (see **Figure 2.1**). The control shows that with time the fungal density within the brain increases. As the mouse is treated with various doses of flubendazole, the fungal density falls, with the highest dose of 400 mg/kg demonstrating the greatest antifungal effect. There is very little difference between the 200 mg and 400 mg dose, demonstrating that solubility limited absorption is likely to be occurring.

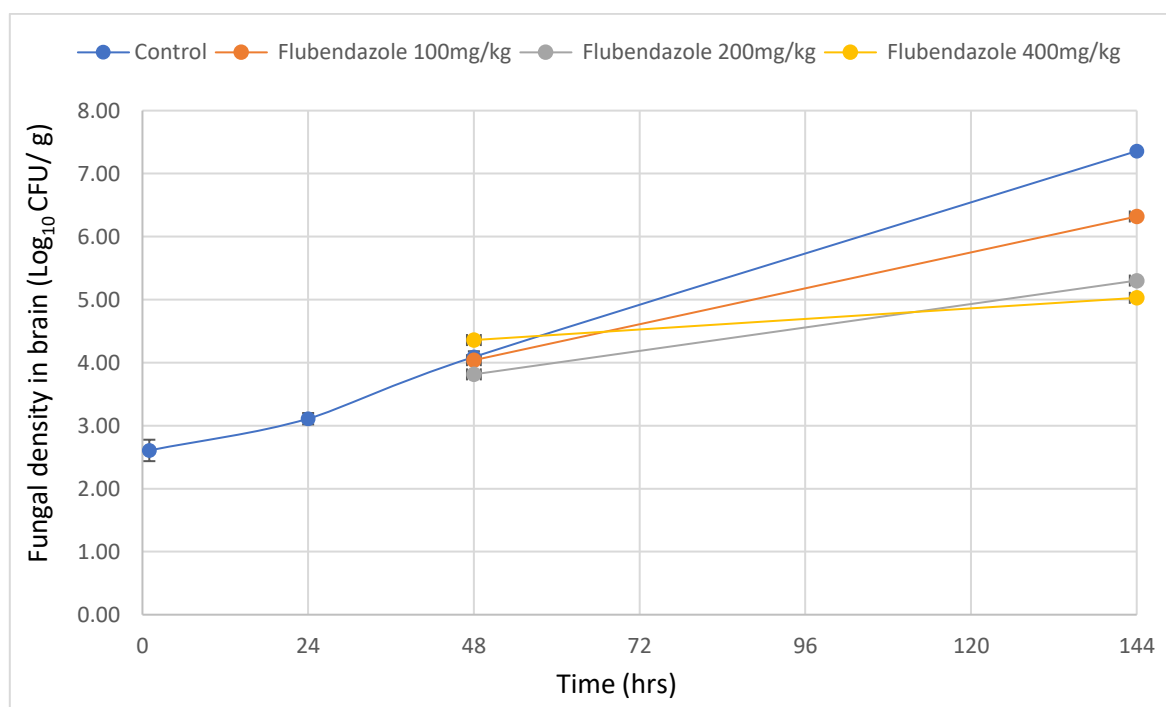


Figure 2.1: Graph illustrating how fungal density in the brain of a mouse changes when treated with various doses of flubendazole (FLBZ) SC once a day over 6 days against *C. neoformans* H99. Control shown in blue, Flubendazole dosed at 100 mg/ kg shown in orange, 200 mg/kg shown in grey, and 400 mg/kg shown in yellow. Data shown as average Log₁₀ CFU/ g, \pm SD n=3.

The two benzimidazole analogues are associated with various DMPK issues, such as metabolism and aqueous solubility of which are discussed throughout this chapter.

2.1.1 Metabolism of Benzimidazoles

Flubendazole and albendazole are known to experience several metabolic transformations in the liver. Flubendazole is rapidly metabolised by phase I and II enzymes. Cytochrome P450 (CYP450) enzymes reduce a flubendazole's carbonyl to a hydroxyl group, to generate a phase I metabolite (FLU-OH) (see **Figure 2.2A**).¹ Hydrolase enzymes can cleave the carbamate group to reveal a primary amine to produce the reduced form (FLU-NH₂) (see **Figure 2.2B**).² A study into the metabolism of benzimidazoles in the nematode *Haemonchus contortus* by *Stuchlíková et al* have shown that FLU-OH is the main phase I metabolite.¹

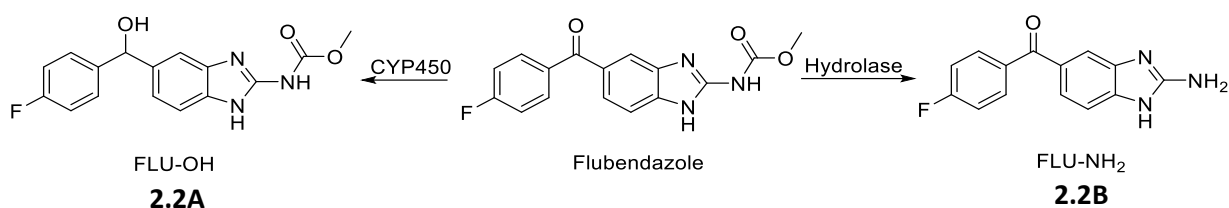


Figure 2.2: Structure of flubendazole and its phase I metabolites. **2.2A:** Hydrolysed flubendazole (FLU-OH) typically formed *via* CYP450 enzymes, **2.2B:** Reduced flubendazole (FLU-NH₂) typically formed from hydrolase.

Stuchlíková et al also state that *N*-/*O*-glycosidation, methylation and *O*-acetylation are common phase II biotransformations of flubendazole, in which various groups attach to the FLU-OH and FLU-NH₂ metabolites (see **Figure 2.3**).¹

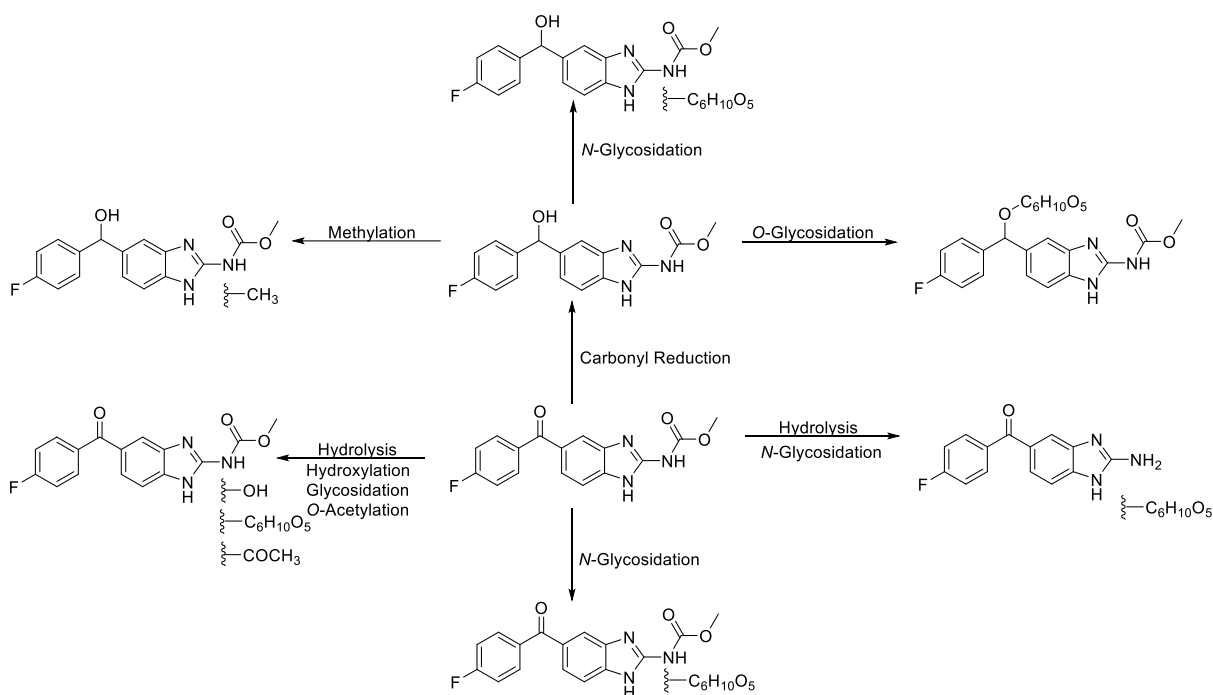


Figure 2.3: Structure of flubendazole and its phase II metabolites *via* N/O-glycosidation, O-acetylation, and methylation.¹

Albendazole is typically metabolised by flavin monooxygenase (FMO) and CYP450 enzymes in the liver to give chiral metabolites albendazole sulfoxide [(+)-ASOX; (-)ASOX], and non-chiral metabolites albendazole sulfone (ASON) (see **Figure 2.4**). The FMO enzymes favour the formation of the (+)-ASOX, whereas the CYP2C6 and/or CYP2A1 favour (-)-ASOX. CYP3A4, which is the dominant enzyme involved in the metabolism of albendazole, forms equal ratio of the (+) and (-) enantiomers.³

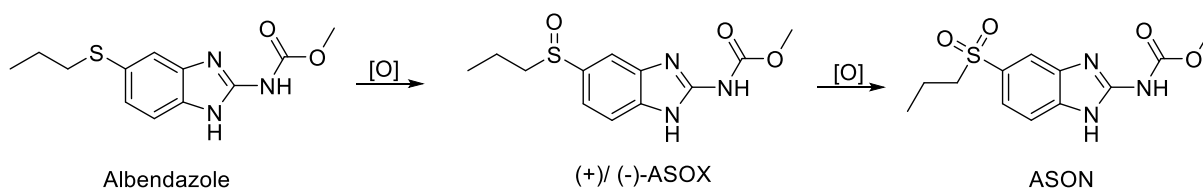


Figure 2.4: Structure of albendazole and its sulfoxide (ASOX) and sulfone (ASON) phase I metabolites.⁴

The (+)/(-)-ASOX metabolites display weaker antifungal activity against *C. neoformans* compared to its corresponding sulfide, similar to the oxfendazole, a sulfoxide analogue of fenbendazole (see **Figure 2.5**) (see **Chapter 1, Table 1.1**). In-house testing of albendazole

sulfone confirmed that sulfone metabolites do not possess any antifungal activity against *C. neoformans*.

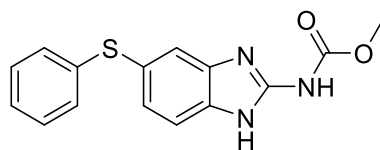


Figure 2.5: Structure of fenbendazole

Albendazole undergoes hydroxylation at the carbamate to form a carbamic acid, and at the end of the aliphatic chain *via* the CYP2J2 enzyme (see **Figure 2.6**).⁵

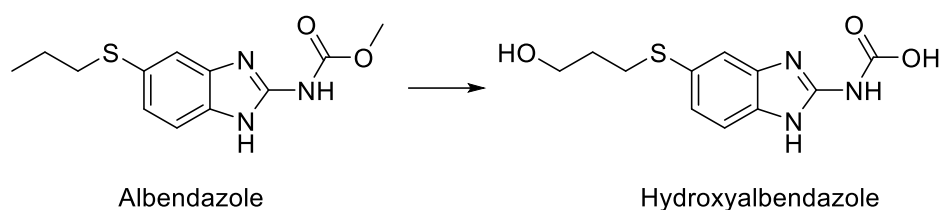


Figure 2.6: Albendazole and its hydroxylation metabolite.⁵

Albendazole also undergoes extensive phase II biotransformations, predominantly glycosidation at the carbamate nitrogen (see **Figure 2.7**).¹

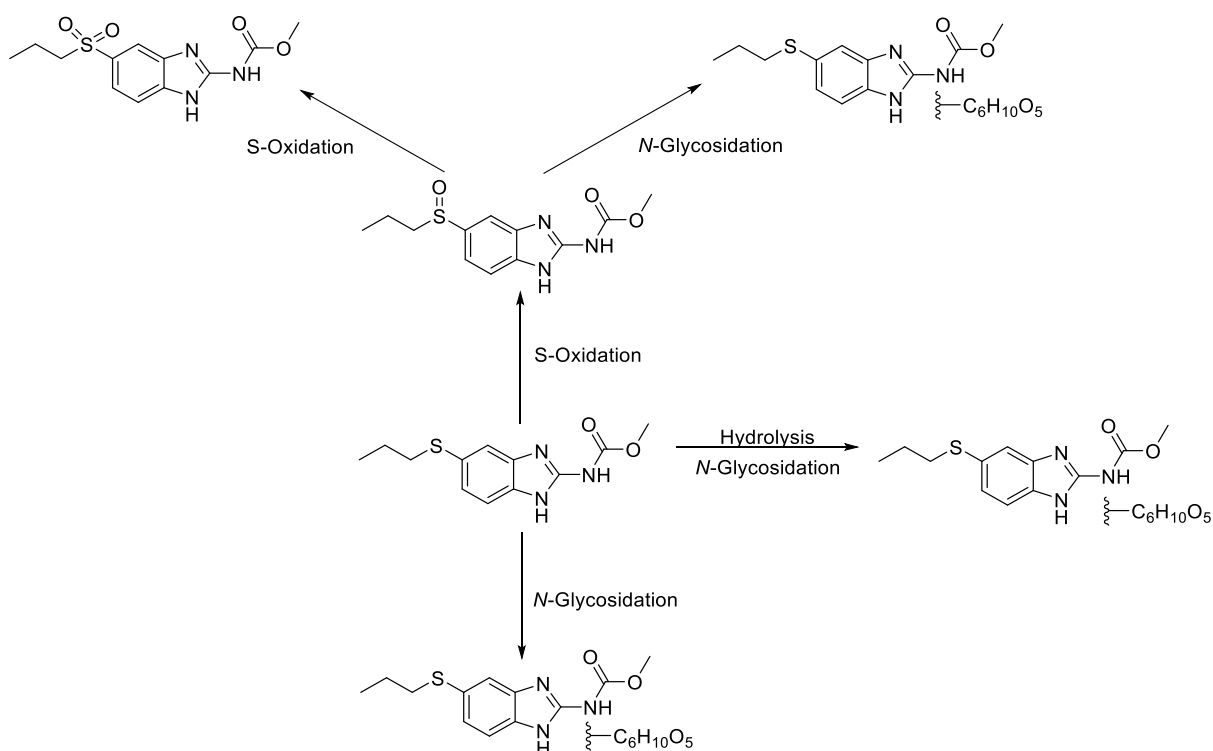


Figure 2.7: Albendazole and its phase II metabolites *via* N-glycosidation.¹

To reduce hepatic clearance of flubendazole and albendazole, medicinal chemistry strategies can be applied. Attaching large groups near the sulfur may block S-oxidation reactions, as steric hindrance can block the binding of CYP450 and FMO enzymes. The replacement of the flubendazole carbonyl group for more metabolically stable functionalities may increase the compounds metabolic half-life.^{6, 7}

2.1.2 Aqueous Solubility of Benzimidazoles

Solubility issues are strongly associated with benzimidazole carbamates. The poor aqueous solubility of albendazole leads to problems with gastrointestinal (GI) absorption and results in a low oral bioavailability of 5%.⁸

Flubendazole is also associated with poor aqueous solubility due to the presence of the carbamate group.⁹ Work by *Ceballos et al* have reported the poor aqueous solubility of different benzimidazoles, including flubendazole, albendazole, mebendazole, their metabolites albendazole sulfoxide and oxfendazole (see **Table 2.1**).¹⁰

Compound	Aqueous Solubility (pH 7.4) ($\mu\text{g}/\text{mL}$)
Albendazole	4.6 ± 1.3
Albendazole sulfoxide	83.6 ± 7.6
Mebendazole	4.6 ± 1.3
Oxafendazole	5.5 ± 1.1
Flubendazole	5.8 ± 7.0

Table 2.1: Evaluation of the aqueous solubility ($\mu\text{g}/\text{mL}$, \pm SD n= 6) of different benzimidazole compounds.¹⁰

The DMPK of flubendazole and albendazole were measured at AstraZeneca to assess the $\text{LogD}_{7.4}$, aqueous solubility, human microsomal clearance, and rat hepatic clearance values of the compounds (see **Table 2.2**). The compounds possess low aqueous solubility and moderate metabolic clearance, confirming previous reports in the literature.^{8, 10, 11}

Compound	MIC (mg/L)	$\text{LogD}_{7.4}$	Aqueous solubility (μM)	Human microsomes Cl_{int} ($\mu\text{L}/\text{min}/\text{mg}$)	Rat hepatic clearance Cl_{int} ($\mu\text{L}/\text{min}/\text{mg}$)
Flubendazole	0.125	2.9	0.8	44	39
Albendazole	0.5	3.2	4	48	30

Table 2.2: Measured DMPK data for benzimidazoles flubendazole and albendazole. Results coded with a traffic light system; MIC: green= good (0.015-0.25 mg/L), amber= acceptable (0.5-4 mg/L) and red= bad (>4 mg/L); DMPK: green= good, amber= acceptable and red=bad. ND: No Data available.

To address the inherent solubility issues of benzimidazole carbamates, water soluble groups such as morpholine can be introduced.^{12, 13}

2.1.3 Previous SAR Work

The Nixon group at the University of Liverpool have investigated the structure activity relationships (SAR) of flubendazole and albendazole to improve efficacy, metabolism, and aqueous solubility (see **Figure 2.8**).

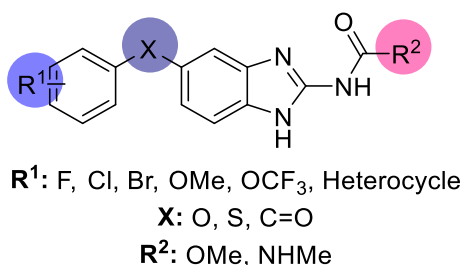


Figure 2.8: SAR exploration into the benzimidazole carbamate compound.

Altering the substituents at the R¹ and R² position to assess the effects on activity and DMPK properties have been employed. Halides around the phenyl ring and bulky groups close to the labile sulfide group have been added to impart steric hindrance. A morpholine has also been attached at the R¹ position to improve aqueous solubility.

The flubendazole carbonyl group at X has been replaced with alternative connections such as sulfides, ethers, amides and CH₂ linkers to address the metabolic issues and improve potency.

All commercially available benzimidazoles tested against various strains of *C. neoformans* possessed a methyl carbamate (see **Chapter 1, Table 1.1**). Modifying the methyl group with other substituents such as ethyl, propargyl and benzyl provided essential SAR information. Ethyl carbamates regularly maintained antifungal activity, typically at a lower inhibitory value compared to the methyl counterpart. Propargyl carbamates displayed promising activity (**Chapter 3: Table 3.1**). Benzyl carbamates resulted in a loss of activity, which was attributed to their bulky nature.

2.1.3.1 MIC and DMPK Data of Benzimidazole Analogues

The Nixon group has generated a large library of benzimidazole analogues for the purpose of gathering SAR information and finding a potential therapeutic agent for the treatment of *C. neoformans*. Various analogues containing sulfide, ether, and ketone functionalities are shown with their associated *in vitro* MIC data (see **Table 2.3**).

Compound	Substituents			DMPK				
	R ¹	X	R ²	MIC (mg/L)	LogD _{7.4}	Aqueous Solubility (μ M)	Human microsomes Cl _{int} (μ L/min/mg)	Rat hepatic clearance Cl _{int} (μ L/min/mg)
Flubendazole	4-F	C=O	OMe	0.125	3.9	0.3	20	33
1a	4-F	S	OMe	0.03	4.2	0.5	10	32
1b	4-OMe	S	OMe	0.03	4	0.3	12	17
1c	2-Br	S	OMe	0.03	4.8	0.1	22	9
1d	4-F	O	OMe	0.06	3.6	0.3	8	26
1e	3-F	O	OMe	0.125	3.7	0.5	10	13
1f	3-Cl, 4-F	O	OMe	0.25	4.1	0.7	11	32
1g	4-F	O	O <i>i</i> Pr	0.25	3.7	1	8.14	71.4

Table Continued...

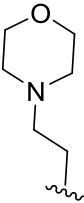
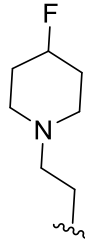
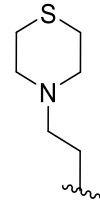
Compound	Substituents				MIC (mg/L)	DMPK			
	R ¹	X	R ²	LogD _{7.4}		Aqueous Solubility (μM)	Human microsomes Cl _{int} (μL/min/mg)	Rat hepatic clearance Cl _{int} (μL/min/mg)	
1h		O	OMe	2.9	0.25 - 4	10	30	48	
1i		O	OMe	3	>4	12	23	10	
1j		O	OMe	3.7	1-2	0.6	300	127	
1k	2-F	C=O	OMe	2.8	1	3	10	47	
1l	4-Cl	C=O	OMe	3.6	0.25	2	169	71	
1m	4-OMe	C=O	OMe	3	0.125	0.3	15	19	
1n	4-OCF ₃	C=O	OMe	3.8	>4	0.4	7	9	
1o	4-OMe	S	NHMe	ND	1	ND	ND	ND	
1p	3-OMe	S	NHMe	ND	>4	ND	ND	ND	

Table 2.3: MIC and DMPK results of various benzimidazole analogues. Results coded with a traffic light system; MIC: green= good (0.015-0.25 mg/L), amber= acceptable (0.5-4 mg/L) and red= bad (>4 mg/L); DMPK: green= good, amber= acceptable and red=bad. ND: No Data available.

Many of the benzimidazole analogues synthesised did not address the aqueous solubility issues. The only compound that had a vast improvement in solubility was compound **1n**. Unfortunately, this compound could not be optimised further as the MIC values were so poor. Another compound that showed promise was **1h**, which has a morpholine ring attached at the R¹ position (see **Figure 2.8**). Compound **1h** gave a moderate solubility of 10 µg/L and initially a favourable MIC of 0.25 mg/L. The MIC value for compound **1h** has varied over repeated MIC testing. Initially the MIC value was 0.25 mg/L, however over time this value dropped to 4 mg/L. The initial MIC value of **1h** allowed the compound to be put forward for *in vivo* testing (see **Section 2.1.3.2**). Despite the recent 4 mg/L value, **1h** performs well *in vivo*, which has attributed to the improved DMPK profile.

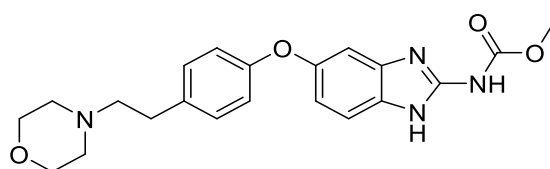


Figure 2.9: Structure of lead compound **1h**

2.1.3.2 *In Vivo* Data of Benzimidazole Analogues

An *in vivo* study demonstrated the effects on fungal density in mice with compound **1h**, with different vehicles and route of administration (see **Figure 2.10**). Initially compound **1h** was tested subcutaneously (**grey** line). The fungal density in mice did not reduce significantly, putting this method of administration into question. **1h** was also tested orally *via* two vehicles. The first vehicle was oil based in which the compound was dissolved in 10% DMSO, 10% ethanol and 80% castor oil (**yellow** line). No reduction in fungal density was observed and was comparable to the control. Furthermore, the mice responded very badly to the vehicle as dissection revealed that **1h** had precipitated within the body of the mice and caused severe adverse drug reactions (ADR) such as diarrhoea, which reduced the amount of compound in their system and led to their death. The other vehicle was an aqueous formulation which consisted of 10% DMSO, 10% tween, and 80% PBS (**orange** line). This vehicle was able to reduce the fungal density without causing ADRs. The aqueous solution in the vehicle protonated the morpholine ring, allowing it to pass through the GI tract with greater ease by improving absorption.

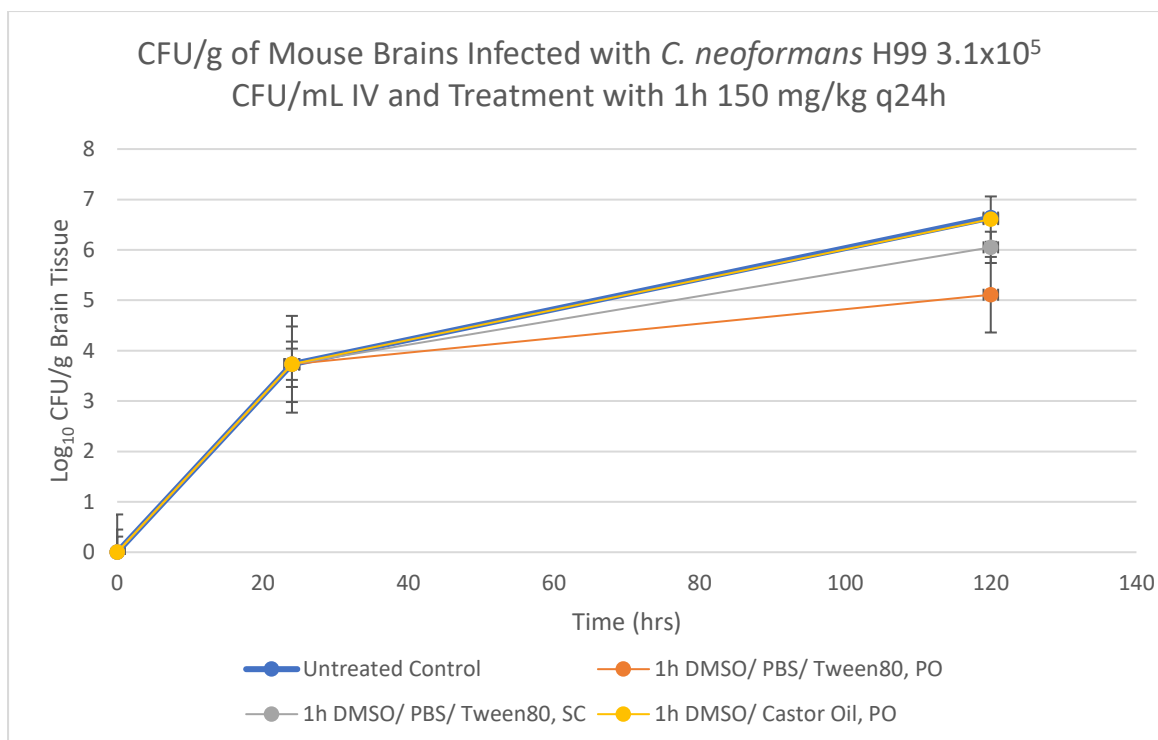


Figure 2.10: *In vivo* data of the effect of compound **1h** on fungal density in mice. Tested by different routes of administration and vehicles at a dose of 150 mg/ kg every 24 hours over 5 days. Untreated control shown in **blue**, oral administration of DMSO (10%)/ Tween80 (10%)/ PBS(80%) vehicle shown in **orange**, subcutaneous administration of DMSO(10%)/ Tween80 (10%)/ PBS(80%) vehicle shown in **grey**, oral administration of DMSO(20%)/ castor oil (80%) shown in **yellow**. Data shown as average $\text{Log}_{10} \text{CFU/g}$, $\pm \text{SD}$ n=3.

It was proposed that the morpholine ring on compound **1h** had improved the DMPK properties such that sufficient drug exposures could be attained to achieve this promising *in vivo* data. To enhance efficacy and DMPK further, new analogues related to **1h** were to be synthesised.

Flubendazole itself was submitted for an *in vivo* study in a mouse model to assess activity, choice of vehicle, administration route, and dosing range.

Flubendazole performs very well *in vivo* as there is a significant drop in fungal density in the mouse brain compared to the control. Flubendazole was tested subcutaneously and orally at 150mg/kg using the aqueous vehicle of 10% DMSO, 10% tween and 80% PBS (see **Figure 2.11**).

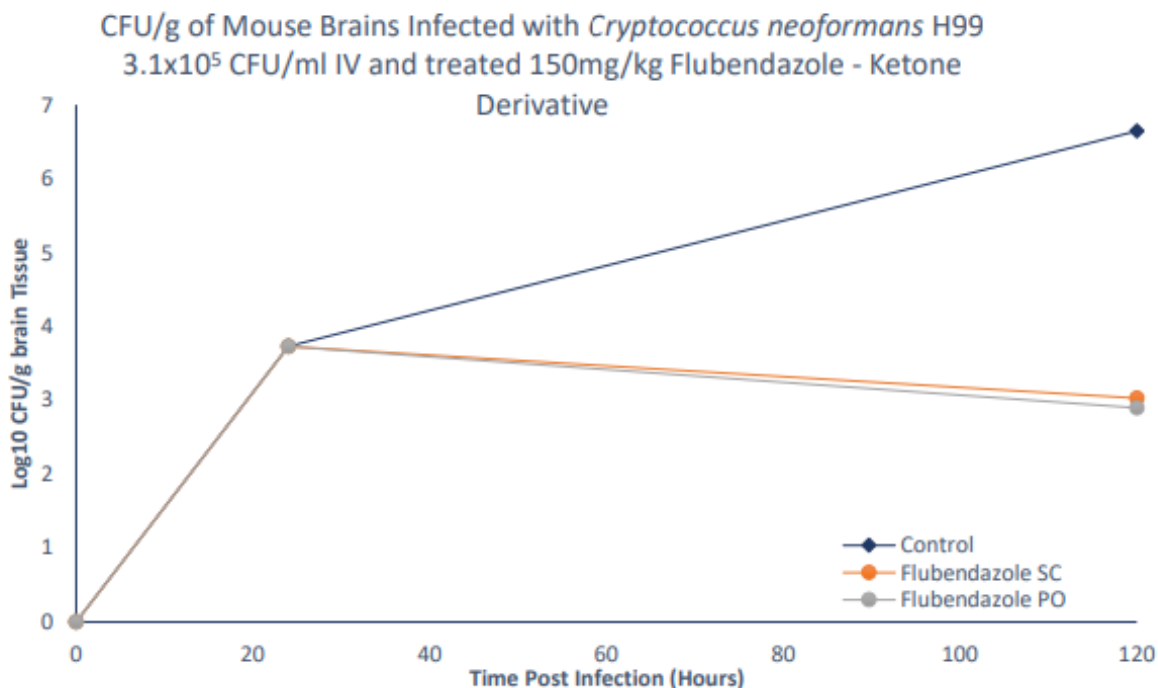


Figure 2.11: *In vivo* data of the effect of flubendazole on fungal density in mice. Dosed at 150 mg/kg every 24 hours for 2 days using the DMSO (10%)/ Tween80 (10%)/ PBS (80%) as the vehicle, administered both orally and subcutaneously. Untreated control shown in **blue**, subcutaneous administration of flubendazole shown in **orange**, and oral administration shown in **grey**. Data shown as average Log₁₀ CFU/ g, ± SD n=3.

Different dosages of flubendazole were explored from 100 mg/kg to the maximum 150 mg/kg (see **Figure 2.12**). For the subcutaneous administration of the drug given once a day at four doses, there was no significant decrease in efficacy. Even at the lowest dose (**light blue** line), the reduction in fungal density was significant. One run involved flubendazole being administered once, and as expected caused the smallest drop in fungal density (**orange** line). However, it still exhibited a significant drop compared the control. This suggests that flubendazole can be administered in a single dose subcutaneously. This is ideal in regions with poor medical infrastructure such as sub-Saharan Africa where *Cryptococcus* infections are most prominent.

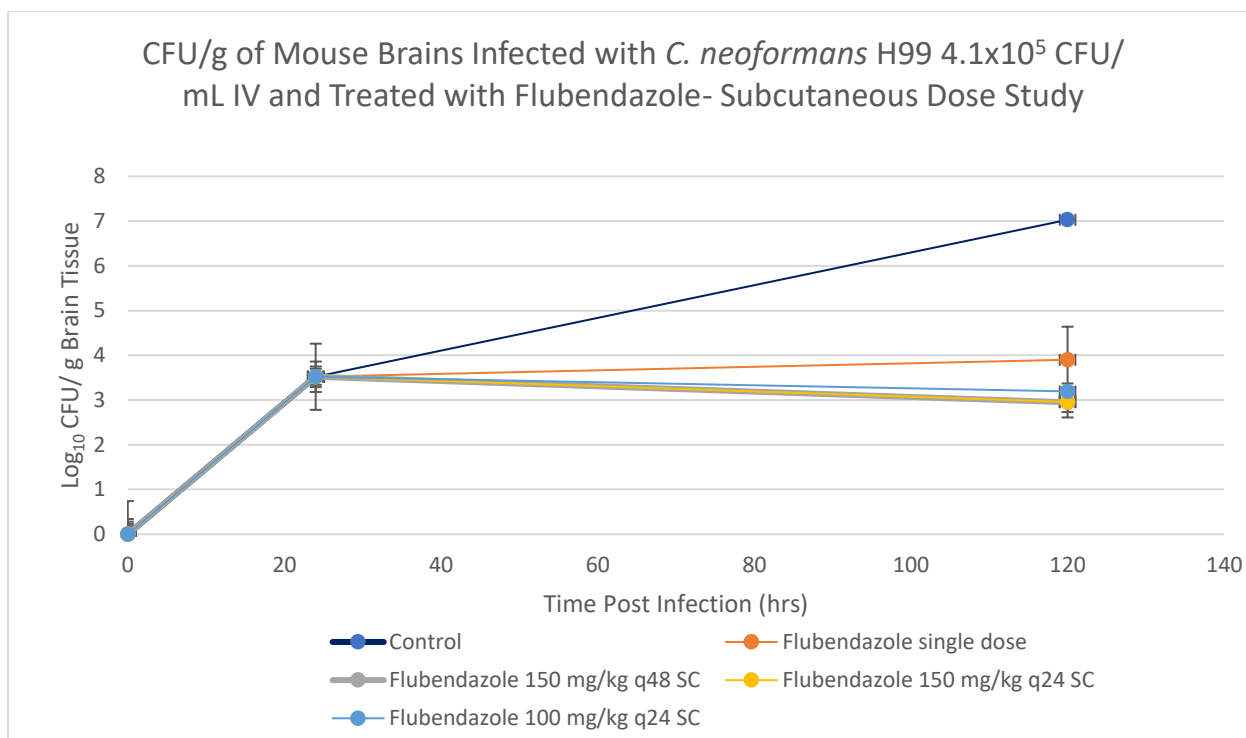


Figure 2.12: *In vivo* data of flubendazole administered subcutaneously at different doses over a course of 6 days in mice using the DMSO (10%)/ Tween80 (10%)/ PBS (80%) vehicle. Untreated control shown in **dark blue**, single dose of flubendazole shown in **orange**, flubendazole dosed at 150 mg/ kg administered over 48 hours shown in **grey**, flubendazole dosed at 150 mg/ kg administered over 24 hours shown in **yellow**, flubendazole dosed at 100 mg/kg dosed every 24 hours shown in **light blue**. Data shown as average Log₁₀ CFU/ g, ± SD n=3.

The effects of efficacy when administered orally was also investigated using the same aqueous vehicle of 10% DMSO, 10% tween and 80% PBS (see **Figure 2.13**). Oral dosing showed similar behaviour to that of subcutaneous in which no significant difference in efficacy when dosage was decreased from 150 mg/kg to 25 mg/kg once daily, with 25 mg/kg offering a substantial drop in fungal density.

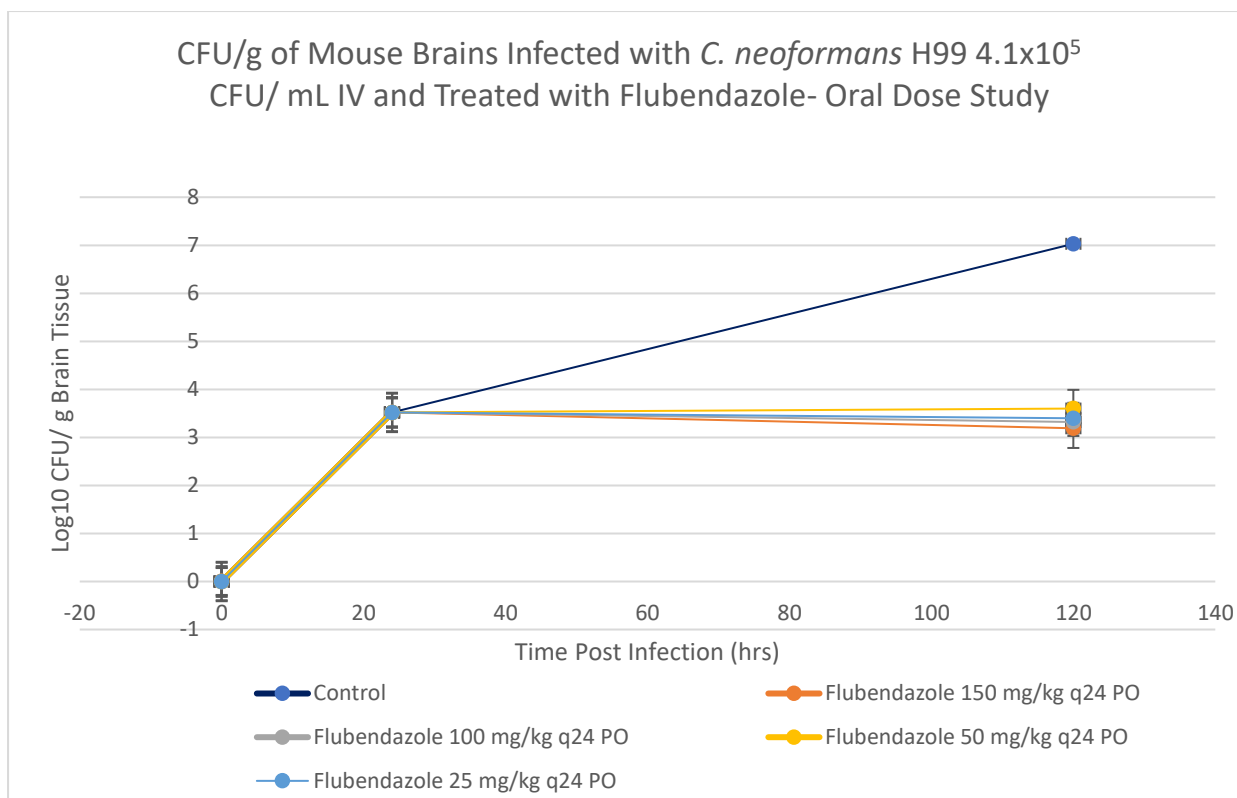


Figure 2.13: *In vivo* data of flubendazole administered orally at different doses administered over 24 hours over a course of 6 days in mice using the DMSO (10%)/ Tween80 (10%)/ PBS (80%) vehicle. Untreated control shown in **dark blue**, flubendazole dosed at 150 mg/ kg shown in **orange**, 100 mg/kg shown in **grey**, 50 mg/kg shown in **yellow**, and 25 mg/kg shown in **light blue**. Data shown as average Log₁₀ CFU/ g, ± SD n=3.

Flubendazole is an excellent drug displaying promising antifungal effects *in vitro* and *in vivo*. However, the DMPK profile, particularly the solubility hinders its effects. **Figures 2.12 and 2.13** show that flubendazole can significantly drop the fungal load in mice. Both the subcutaneous and oral dosing studies display similar reductions in fungal density. The subcutaneous study reveals that as we increase the dose of flubendazole, small reductions in fungal density occur. Although, the oral study does not conform to the same pattern of magnitude. The oral dose demonstrates that a higher dose of flubendazole is not necessary, which may minimise the toxic effects to the patient. However, the lack of continued CFU reduction at higher doses may be due to the inherent insolubility of benzimidazoles, as they are unable to be fully absorbed through the GI tract. The poor solubility potentially prevents flubendazole achieving adequate bioavailability to cause an enhanced therapeutic effect.¹⁴ In addition, a significant drop in CFU is experienced upon administration of flubendazole, yet

the fungus is not eliminated, limiting the effectiveness of the treatment. Therefore, ways to enhance the solubility of flubendazole are necessary, to assess whether solubility is the cause of the stagnant antifungal effect occurring in the oral dose study.

The morpholine compound **1h** improved upon the inherent solubility issues, therefore flubendazole with a morpholine ring is a potential candidate, provided it can maintain similar antifungal activity.

2.2 Aims

The aims for this project were to gather further SAR information by synthesising a range of new benzimidazole compounds.

Requirements for the project include...

1. Synthesise new morpholine derivatives of benzimidazoles based on the structure of compound **1h** to further improve the DMPK profile.
2. Synthesise a ketone variant of the morpholine compound **1h** to maintain flubendazole activity while improving the DMPK profile.

2.3 Discussion

2.3.1 Synthesis

Modifications to the morpholine were conducted on the ethylene linked phenyl ether analogues (see **Figure 2.14**). The targets were selected as they are structurally similar to the hit **1h** and may possess equivalent or greater activity against *C. neoformans*. A morpholine bioisostere in the form of the spirocyclic oxetane (**3a**), has been selected for its ability to increase the aqueous solubility of a compound, while also preventing metabolic attack and cleavage, of which morpholine is often susceptible to.¹⁵ Pyrrolidine (**3b**) has also been selected to determine whether a smaller heterocyclic ring is able to improve its efficacy. Morpholine derivatives **3c-3i** will also be introduced, to assess whether aliphatic substituents and bridges will enhance the potency compared to the parent compound.

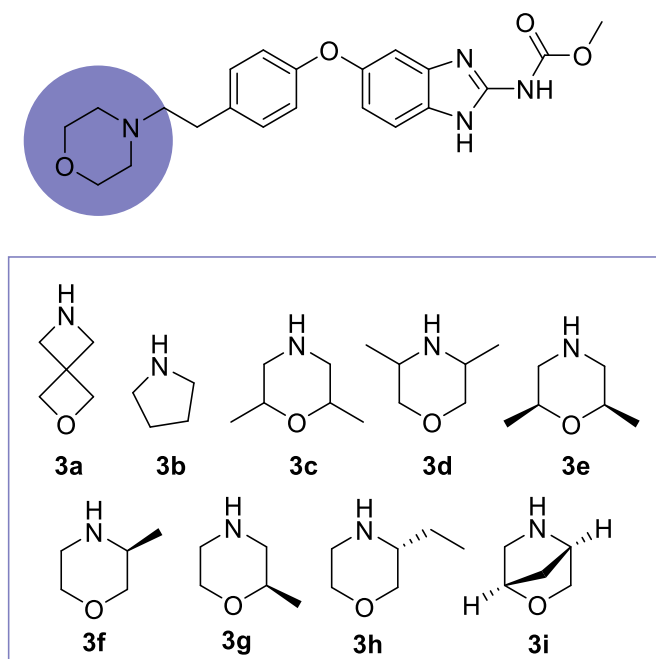
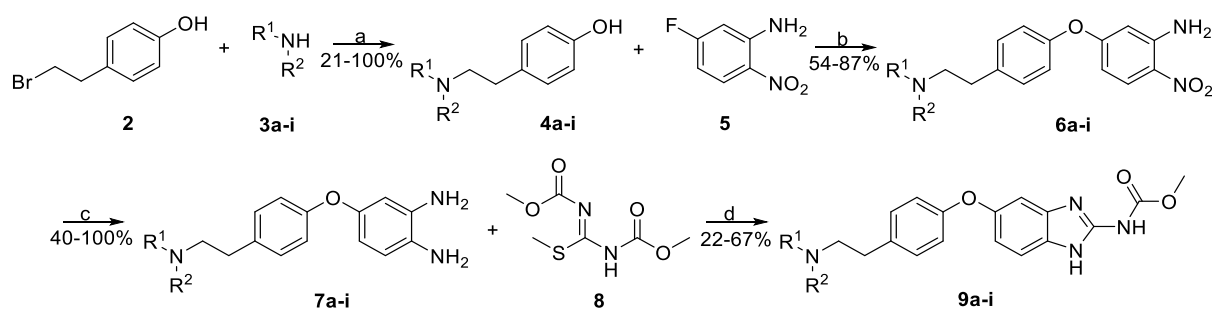


Figure 2.14: SAR exploration into the morpholine ring. Morpholine bioisosteric spirocyclic oxetane (**3a**), pyrrolidine (**3b**), and morpholine derivatives (**3c-3i**).

The synthesis of the **1h** derivatives involves a 4-step procedure in which the compound was derivatised at the first reaction (see **Scheme 2.1**). Previous attempts to incorporate the amine functionality at a later stage of the synthesis was unsuccessful due to the presence of the insoluble carbamate group. The carbamate group is soluble in DMSO and displays limited solubility in methanol. This hinders the reaction going to near or full conversion, as a result the amine is integrated at the earliest step.



Scheme 2.1: Reagent and conditions: (a) DIPEA, MeCN, reflux, 16 hr; (b) K_2CO_3 , DMF, reflux, 16 hr; (c) $SnCl_2 \cdot H_2O$, EtOH, reflux, 16 hr; (d) AcOH, MeOH, reflux, 16 hr.

For the first reaction an amine (**3a-i**) is coupled to 4-(2-bromoethyl)phenol (**2**) via a nucleophilic substitution reaction (see **Figure 2.15**).¹⁶

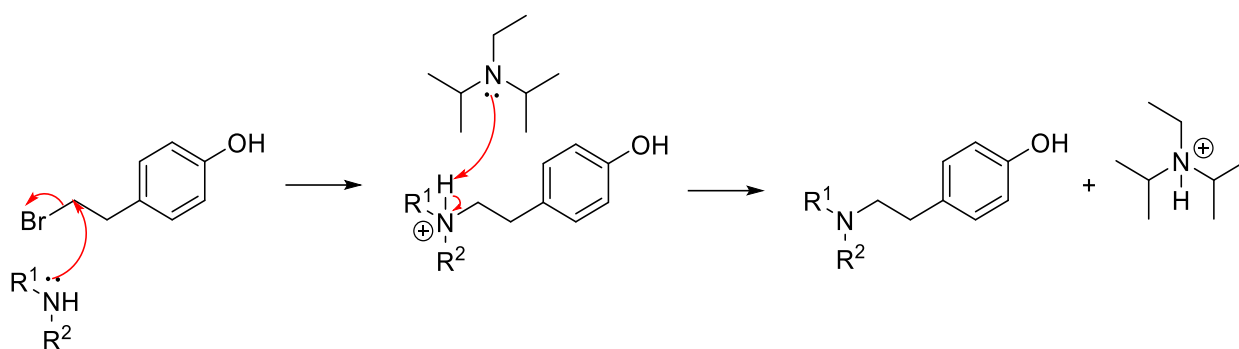


Figure 2.15: Nucleophilic substitution mechanism between 4-(2-bromoethyl)phenol and an amine.

Previously DMF and K_2CO_3 were used within the Nixon group. Unfortunately, this choice of solvent and base gave rise to a number of side products such as a self-coupled phenol. DMF increases the pKa of the phenol derivative to 15.40, allowing it to compete with the amine for deprotonation. This leads to the formation of self-coupled by-products (see **Table 2.4**). Acetonitrile increases the phenols pKa to 26.60 resulting in a less acidic hydroxyl proton compared to the selected amines which averages around 16.61 (see **Table 2.4**). This allows the reaction to be more selective by only deprotonating the amine.

Solvent	Phenol	Morpholine
H ₂ O	9.89 ¹⁷	8.49 ¹⁸
DMF	15.40 ¹⁹	-
MeCN	26.60 ²⁰	16.61 ²¹

Table 2.4: The pKa values of phenol and morpholine in different solvents.

The work up procedure described by *Grice et al* involved concentrating acetonitrile, filtering the solid and washing with diethyl ether, as the desired product is only slightly soluble in these solvents.²² However, not all analogues of compound **4** produced a solid, and instead formed a crude oil. Examination of crude ¹H NMR's showed that the amine starting material remained. This was also observed on the TLC plates from the reactions. An aqueous work up to remove the amine was typically avoided, as the enhanced polar nature of the compound led to the product being trapped in the water layer. Purification *via* column chromatography was required to remove the amine starting material. Due to the polar nature of the compounds and the close rf values, co-elution of the product and amine often occurred. Therefore, the product isolated was not always pure, but did allow some of the amine to be removed. The compounds were characterised by ¹H, ¹³C NMR and mass spectrometry. The

¹H NMR displayed the expected aromatic protons along with the aliphatic chain and the associated amine peaks. Due to the variability of purity, the NMR spectrums often showed smaller peaks from the associated amine that could not be fully removed.

The relative stereochemistry of amines **3c** (2,6-dimethylmorpholine) and **3d** (3,5-dimethylmorpholine) are unknown as they are not specified by the supplier Fluorochem. Amines **3c** and **3d** are presumed to be a mixture of *cis* and *trans* isomers. Amine **3e** (*cis*-2,6-dimethylmorpholine) was used to address the issue of not knowing the stereochemistry of **3c**. Variation in the biological evaluation of compounds **9c** and **9e** can be attributed to differences in the composition of the final product.

The yields for the reactions vary depending on the amine as they range from 21-100%. Amine **3d** (3,5-dimethylmorpholine) contributed to the lowest yield. Amine **3a** (2-oxa-6-azaspiro[3.3]heptane) performed best with a quantitative yield. The low yields were attributed to the polar nature of the compound and associated difficulties of purification due to the enhanced solubility.

For the second reaction compound **4a-i** couples to 5-fluoro-2-nitroaniline (**5**) in a S_NAr reaction.²³ Compound **4a-i** was suspended in a mixture of K₂CO₃ and DMF to allow deprotonation of the hydroxyl proton. 5-Fluoro-2-nitroaniline (**5**) was later added and allowed to stir at reflux overnight. 5-Fluoro-2-nitroaniline (**5**) was selected as the *para* nitro group enables the S_NAr reaction. The reaction is comprised of two steps; the first is the addition step, in which a nucleophile attacks the C-F bond resulting in the loss of aromaticity and forming a Meisenheimer intermediate. The second step is elimination, in which the halogen bond is broken, restoring aromaticity, yielding the desired product. The nitro group's resonance stabilises the intermediate, allowing the reaction to be favourable (see **Figure 2.16**). The fluorine also contributed to the reaction. The fluoro substituent has a high negative inductive effect, which can stabilise the negative charge in the Meisenheimer complex. The addition step is rate determining, as a high energy barrier must be overcome due to the loss of aromaticity. The negative inductive effect of the fluorine reduces the energy barrier (due to the stability of the reactive intermediate), accelerating the ordinarily slow addition reaction.

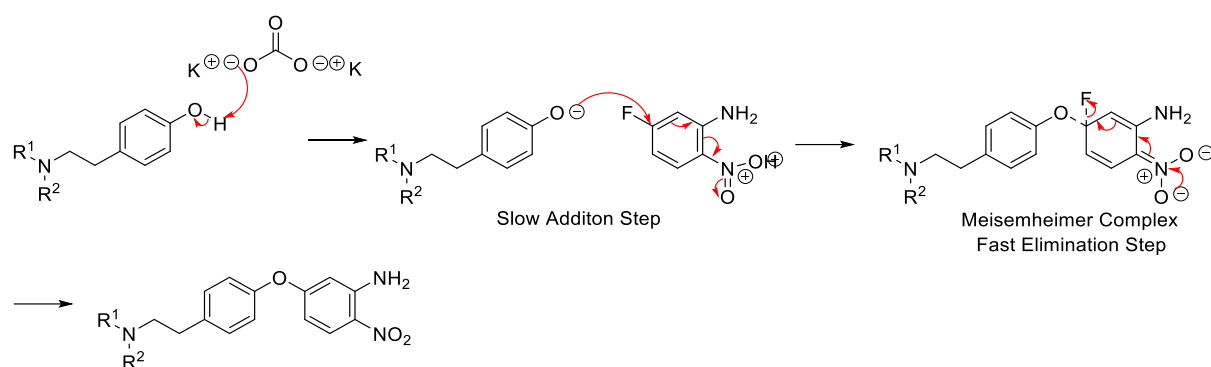


Figure 2.16: S_NAr mechanism between compound **4a-i** and 5-fluoro-2-nitroaniline (**5**).

Yields varied among the different analogues from 54-87%. The yields were typically low due to the high polarity of compound **4** series. Traditionally this reaction involved an aqueous work up to remove the water-soluble by-products. Yet again, the work up was avoided as the products may become trapped in the aqueous layer. The crude mixture was purified *via* column chromatography. The 5-fluoro-2-nitroaniline (**5**) starting material was removed easily due to having a much higher *rf* value compared to the product. However, some analogues such as **6d** (3,5-dimethylmorpholine analogue) and **6i** ((1*S*,4*S*)-2-oxa-5-azabicyclo[2.2.1]heptane analogue) carried through the same amine impurity from the previous reaction, which could be observed in the 1H NMR spectrum.

All compounds were characterised by 1H , ^{13}C NMR, and mass spectrometry. 1H NMR showed the 5-fluoro-2-nitroaniline peaks in the aromatic region. The NH_2 peak could not be observed due to hydrogen/ deuterium exchange as deuterated methanol (CD_3OD-d_4) was chosen as the NMR solvent.

The nitro group was reduced to an amine to give a diamine product using tin (II) chloride dihydrate.²³ Intermediate **6a-i** was added to a suspension of ethanol and tin (II) chloride dihydrate and refluxed overnight. The nitro groups are believed to be reduced *via* a single electron transfer reaction (SET), in which the nitro group is converted to a nitroso group, hydroxylamine, and then finally an amine (see **Figure 2.17**).

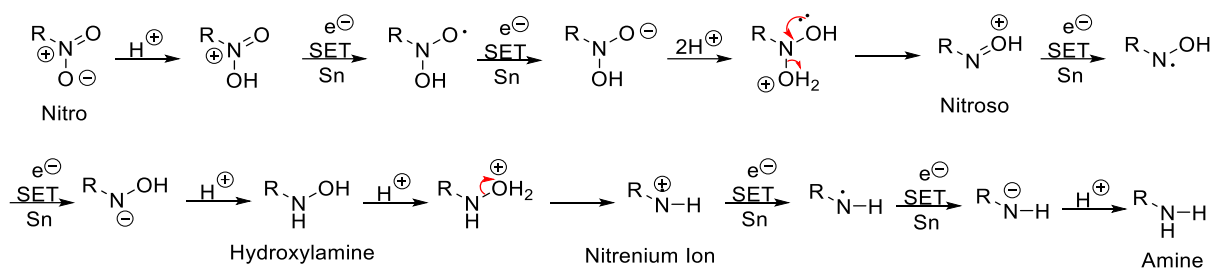
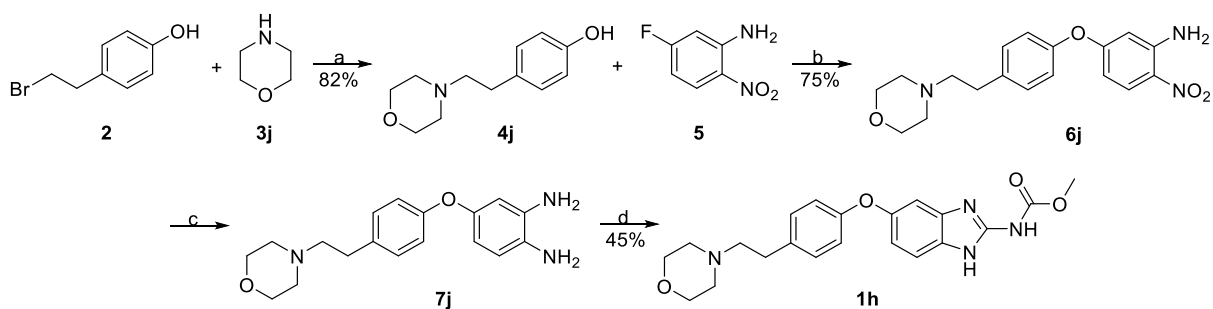


Figure 2.17: Single electron transfer (SET) mechanism for the reduction of a nitro group.²⁴

Yields for these reactions varied amongst the different amines (40->100%). The yields range from moderate to poor due to solubility issues. One yield was over 100% (**7c**) due to the presence of impurities that could not be removed *via* purification. Previous reactions avoided aqueous work ups due to the polar nature of the product. However, for this reaction, it could not be avoided as the tin waste needed to be removed. The reaction mixture was basified to pH 9-10 and the tin waste was collected *via* Buchner filtration. Unfortunately, residual tin remained in the filtrate, making the aqueous work up necessary. It was at this stage that the yield significantly dropped, as some product was lost in the aqueous layer. Multiple extractions of the aqueous layer and “salting out” with NaCl was implemented to reclaim the organic material. However, there was no notable increase in yield. Purification *via* column chromatography was required to separate the starting material **6a-i** from the product, in which some product was lost due to the highly polar nature of compound **7a-i**.

Alternative procedures to reduce the nitro group were carried out within the Nixon group such as hydrogenation²³ and the use of iron.²⁵ Reduction using iron also gave poorer yields compared to tin. Tin (II) chloride dihydrate is often cited for the reduction of nitroaniline derivatives, as workup tends to be simple and does to give rise to many side products.²³ This method has performed well in previous reactions of the benzimidazole series and has only become an issue when dealing with polar materials such as the amines. When hydrogenation was attempted, various side products were found in the resultant crude material, leading to complicated column purifications and poor yields.

Interestingly, large quantities of **1h** had to be synthesised for extensive *in vivo* testing. This was beyond the laboratories capabilities as the large-scale synthesis of **1h** had low yields due complications with work ups and purifications at the nitro reduction step. Therefore, the synthesis of **1h** was outsourced to WuXi (see **Scheme 2.2**).²⁶



Scheme 2.2: Reagents and Conditions of the WuXi scale up: (a) DIPEA, MeCN, reflux, 16 hr; (b) K_2CO_3 , DMF, reflux, 16 hr; (c) H_2 , Pd/C (10%), DMF/ethanol, rt, 16 hr; (d) 1,3-Bis(methoxycarbonyl)-2-methyl-2-thiopseudourea, acetic acid, methanol, screw fix sealed tube, 65 °C, 16 hr.²⁶

The synthetic route by WuXi was similar to the route developed by the Dr Nixon group. Although, the tin (II) chloride dihydrate reaction was replaced with a hydrogenation procedure involving H_2 , Pd/C (10%). The yield was not determined as the TLC of the reaction mixture indicated that the reaction had gone to completion. The reaction mixture was filtered to remove the catalyst and carried through to the final step of the synthesis to obtain **1h** in 98.5% purity. WuXi did not encounter the same hydrogenation issues, which is promising as this procedure can replace tin (II) chloride dihydrate in the future. The presence of residual tin in a pharmaceutical is worrying due to its toxic nature, so this route could be an advantage in future synthesis.²⁷

To discuss the final step of the procedure, we must first consider the formation of the intermediate 1,3-bis(methoxycarbonyl)-2-methylisothiurea (**8**). S-Methylisothiurea hemisulphate (**10**) and methyl chloroformate (**11**) were stirred in water at 0°C. The reaction was then warmed to room temperature and the product precipitated out as a white solid upon basification to pH 10 (see **Figure 2.18**).²⁸

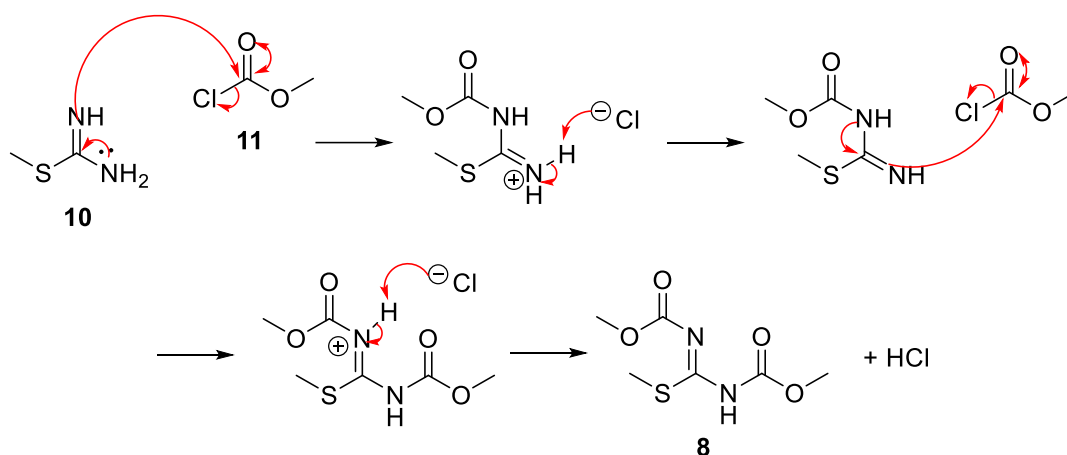


Figure 2.18: Proposed thiourea synthesis mechanism.

The compound was filtered and once isolated, transferred to a screw fix sealed tube with compound **7a-j**, MeOH and AcOH and left to stir at reflux overnight.²³ It is believed that one amine of compound **7a-i** nucleophilically attacks the imine centre of **8**, resulting in the loss of methanethiol. The second amine of **7a-i** attacks the same imine centre of **8** causing the removal of methyl carbamate. Tautomerisation between the imidazole and carbamate nitrogens afford the final benzimidazole structure (see **Figure 2.19**).²⁹ Previous analogues synthesised in the Nixon group were very insoluble, causing a solid to precipitate in the screw fix sealed tube. However, due to the addition of the amine (**3a-i**), solubility has increased, resulting in the loss of the characteristic precipitate. The product was obtained by cooling and reducing the solvent. The residue was suspended in diethyl ether and filtered through a sinter funnel. The starting material and any side products are removed during filtration and washing with diethyl ether. This method was effective with **9a** (2-oxa-6-azaspiro[3.3]heptane analogue) and **9e** ((2*S*,6*R*)-2,6-dimethylmorpholine analogue). However, the remaining analogues required alterations to the work up procedure. The addition of a small amount of 1M NaOH was effective in encouraging the product to precipitate if the reaction mixture had become too acidic. For the synthesis of **8**, HCl is produced *in situ* (see **Figure 2.18**). NaOH is added to encourage the formation of **8** as a free base, allowing the compound to precipitate out of solution. This reaction was temperamental and inconsistent amounts of NaOH were added, meaning some batches were more acidic than others. Therefore, the addition of NaOH was required to obtain **9** as a free base if excess acid remained in the reaction mixture. Often it was noticed that a sticky oil was amongst the otherwise pure product. To address

this issue, small amounts of acetone or ethyl acetate were added to the crude mixture and triturated to remove the oil, leaving the solid behind.

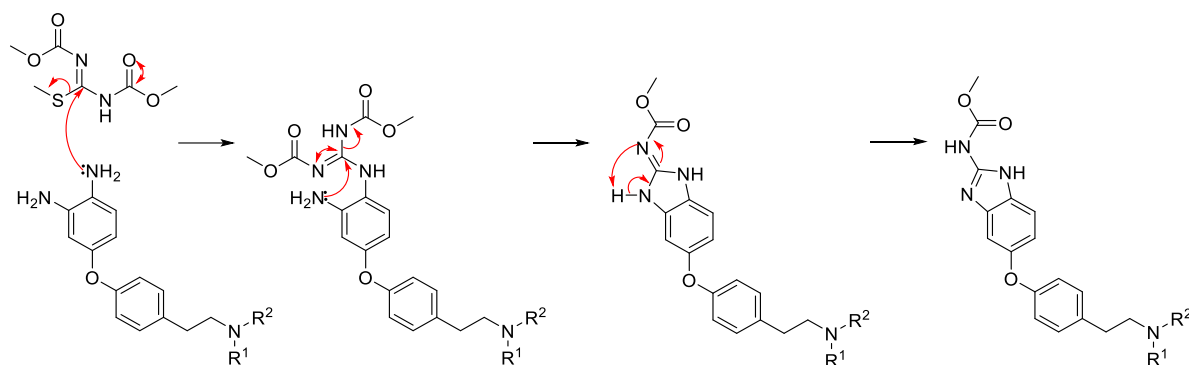


Figure 2.19: Mechanism of the ring closure and formation of the benzimidazole carbamate

The yields varied greatly as there is a range between 27-53%. An explanation for the low and variant yields has been mentioned previously. The heightened solubility of the final product brought difficulties to the workup and purification, as some of the product may have passed into the filtrate along with the by-products. Some product was lost during the trituration stage, as the solid was readily disturbed during removal of solvent, resulting in a lower yield. Another issue is the intermediate 1,3-bis(methoxycarbonyl)-2-methylisothiourea (**8**). It was observed that over time, the compound degraded, particularly if it has been made on a larger scale. The compound once formed is suspended in water, and upon filtration some residual water remains. The water can hydrolyse **8** by nucleophilically attacking the imine centre in the same manner shown in **Figure 2.18**. This means less **8** is within the reaction vessel, lowering the yields.

These compounds were characterised by ¹H, ¹³C NMR, and mass spectrometry. The characteristic methyl carbamate peak was shown at around 3.18 ppm with an integration of three. With all the other peaks associated with the compound present in the spectrum. Purity data was obtained *via* HPLC.

2.3.2 Results

2.3.2.1 MIC Data

All new compounds were tested for biological activity against whole cell *C. neoformans* using the European Committee on Antimicrobial Susceptibility Testing (EUCAST) method, which is a standardised method to make results more comparable from different research projects

(see **Table 2.5**).³⁰ Fungal cells were grown on agar plates and the media used for Minimum Inhibitory Concentration (MIC) testing was a dextrose solution and DMSO. The plates, once inoculated, were incubated for 48 hours, and the MIC value given at the lowest concentration where there was no observable growth.^{31, 32}

Compound	Functionality	MIC (mg/L)
Flubendazole	4-F	0.125
1h	Morpholine (3j)	0.25- (4)
9a	2-Oxa-6-azaspiro[3.3]heptane (3a)	Repeat
9b	Pyrrolidine (3b)	>32
9c	2,6-Dimethylmorpholine (3c)	ND- Covid-19
9d	3,5-Dimethylmorpholine (3d)	ND- Covid-19
9e	(2 <i>S</i> ,6 <i>R</i>)-2,6-Dimethylmorpholine (3e)	>32
9f	(<i>S</i>)-3-Methylmorpholine (3f)	ND- Covid-19
9g	(<i>R</i>)-2-Methylmorpholine (3g)	ND- Covid-19
9h	(<i>R</i>)-3-Ethylmorpholine (3h)	ND- Covid-19
9i	(1 <i>S</i> ,4 <i>S</i>)-2-Oxa-5-azabicyclo[2.2.1]heptane (3i)	ND- Covid-19

Table 2.5: MIC data of **1h** derivatives. Values indicated by traffic light system. Green – good - 0.015 – 0.25 mg/L, amber – acceptable - 0.5 – 4 mg/L, red – poor - >4 mg/L. ND- No data. Some data has yet to be obtained due to closure of the laboratory amid the Covid-19 crisis.

The MIC testing results are limited; however, we can determine that **9b** (pyrrolidine analogue) and **9g** ((*R*)-2-methylmorpholine analogue) have poorer antifungal activity compared to our lead compound **1h**, and the parent compound flubendazole. Compound **9a** (2-oxa-6-azaspiro[3.3]heptane analogue) required additional testing due to inconclusive MIC results. Results for the other morpholine derivatives are currently on hold due to Covid-19 disruption. Therefore, results and evaluations will be considered once the data is available.

2.3.2.2 DMPK Data

Compound	LogD _{7.4}	Aqueous Solubility (μM)	Rat Heps CLint (μL/ min/ mg)	Human Heps CLint (μL/ min/ mg)
Flubendazole	2.9	0.8	39	44
1h	2.9	10	30	48
9a	ND- Failed	ND- Failed	ND- Failed	ND- Failed
9b	1.9	876	5.94	<3
9c	ND- Covid-19	ND- Covid-19	ND- Covid-19	ND- Covid-19
9d	ND- Covid-19	ND- Covid-19	ND- Covid-19	ND- Covid-19
9e	ND- Covid-19	ND- Covid-19	ND- Covid-19	ND- Covid-19
9f	ND- Covid-19	ND- Covid-19	ND- Covid-19	ND- Covid-19
9g	ND- Failed	ND- Failed	ND- Failed	ND- Failed
9h	ND- Covid-19	ND- Covid-19	ND- Covid-19	ND- Covid-19
9i	ND- Covid-19	ND- Covid-19	ND- Covid-19	ND- Covid-19

Table 2.6: Measured DMPK data for **1h** derivatives. Values indicated as the traffic light system. Green – good, amber – acceptable and red – poor. ND: No data. Some data has yet to be obtained due to closure of the laboratory amid the Covid-19 crisis.

DMPK testing for the **1h** derivatives were inconclusive as compounds **9a** (2-oxa-6-azaspiro[3.3]heptane analogue) and **9g** ((*R*)-2-methylmorpholine analogue) failed to gather any data in the standard analytical method provided by AstraZeneca (see **Table 2.6**). Therefore, DMPK testing at an alternative company is required to develop and optimise an appropriate method. Fortunately, compound **9b** (pyrrolidine analogue) provided DMPK data. The LogD data was only 1.9 which is lower than **1h** and flubendazole. However, the aqueous solubility was far greater than **1h** at 876 μM. This is a near 88-fold improvement. The rat and

human metabolic clearances were also improved compared to the lead compound. Unfortunately, **9b** (pyrrolidine analogue) has no antifungal activity, therefore optimisation is required to balance antifungal activity with an improved DMPK profile. The DMPK data for rest of the morpholine derivatives are on hold due to the Covid-19 crisis. Once normal lab work resumes DMPK data can be obtained and the SAR can be further understood.

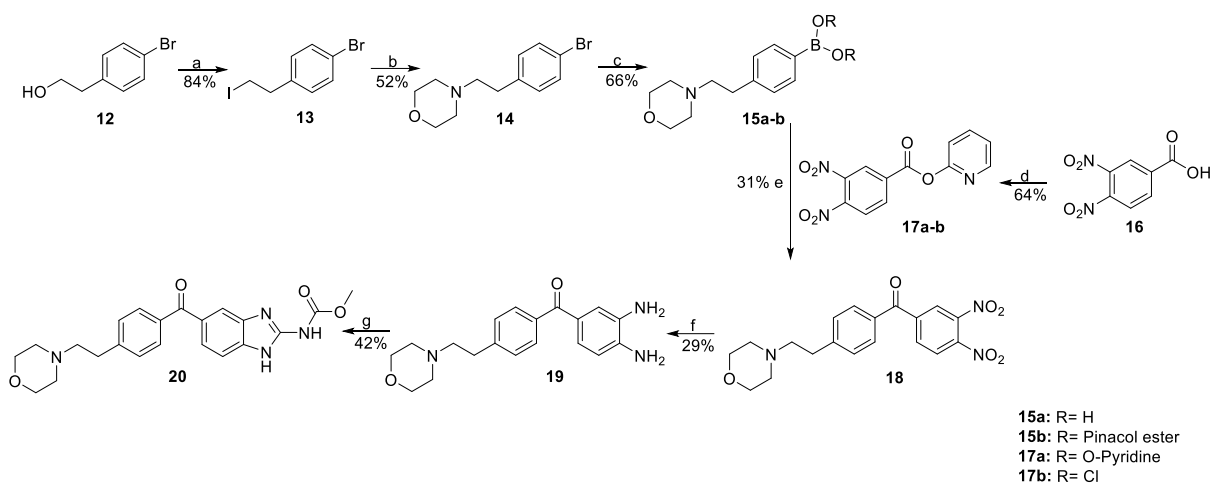
2.4 Benzimidazole Ketone Derivative

As mentioned in Section **2.1.2**, the ketone analogue flubendazole has good antifungal activity, but poor aqueous solubility. The incorporation of the morpholine ring in compound **1h** improved the DMPK properties of the benzimidazole carbamate, allowing it to behave well in the *in vivo* study (see **Figure 2.10**). The attachment of a morpholine ring on a ketone has been proposed to generate a potent antifungal that may enhance the activity of flubendazole. The increase in aqueous solubility may lead to more of the ketone analogue crossing the GI tract, reducing the fungal load further.

2.5 Discussion

2.5.1 Synthesis

The synthesis of the novel ketone analogue involves a 7-step method, in which the ketone is synthesised *via* Suzuki-Miyaura cross-coupling methodology (see **Scheme 2.3**).



Scheme 2.3: Reagent and Conditions: (a) PPh_3 , I_2 , imidazole, DCM, rt, 16 hr; (b) Morpholine, DIPEA, MeCN, reflux, 16 hr; (c) $n\text{-BuLi}$, $\text{B}(\text{O}i\text{Pr})_3$, THF -78°C , 1 h, then 2M HCl, rt, 1 hr; (d) Di(pyridine-2-yl) carbonate, DMAP, DCM, rt, 2 days; (e) $\text{Pd}(\text{OAc})_2$, PPh_3 , dioxane, MW, 100°C , 30 mins; (f) $\text{SnCl}_2 \cdot \text{H}_2\text{O}$, EtOH, reflux, 16 hr; (g) 1,3-Bis(methoxycarbonyl)-2-methylisothiourea (**8**), AcOH, MeOH, reflux, 16 hr.

The first reaction involved a nucleophilic substitution reaction to convert the hydroxyl group of 2-(4-bromophenyl)ethanol (**12**) to an alkyl halide (**13**) *via* iodination. The synthesis involved a simple $\text{S}_{\text{N}}2$ reaction between the hydroxyl group of **12** and iodine, catalysed by triphenylphosphine (PPh_3). This reaction is a modification of the Appel reaction which traditionally converts an alcohol into an alkyl chloride using PPh_3 and carbon tetrachloride (see **Figure 2.20**).³³ The catalyst PPh_3 reacts with iodine to form a phosphonium salt which is able to equilibrate. Imidazole plays the role of a general base and an acid scavenger for HI. Phenylhalogenphosphoranes in solution can exist in an equilibrating mixture of the phosphorene and the phosphonium salt and the imidazole and iodide can be displaced by the alcohol. The iodide is able to attack the adjacent carbon to the phosphonium species in an $\text{S}_{\text{N}}2$ reaction to remove the activated oxygen and form the final alkyl iodide compound with triphenylphosphine oxide as the by-product.^{34, 35, 36}

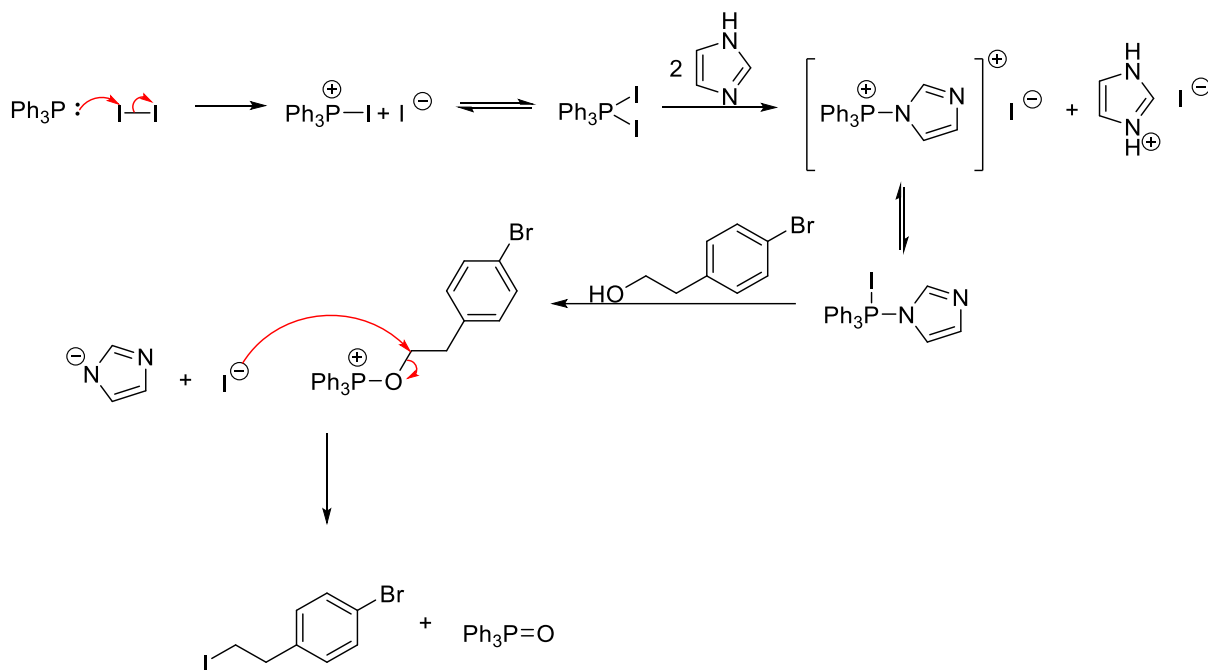


Figure 2.20: Mechanism of the Appel reaction for the conversion of a hydroxyl group to an alkyl iodide.

The preparation of compound **13** was sourced from *Lanni et al.*³⁷ Initially the reaction achieved a yield of 44%, this is low as *Lanni et al* suggests a yield of 75-90% (depending on substrate) should be obtained. Regrettably, the reaction may not have been completely anhydrous, resulting in quenching *via* the residual water. Fresh anhydrous DCM was added to a flame dried flask which had been purged with argon. By improving the anhydrous conditions, the yield increased to 79%, comparable with the literature. Starting material remained in the crude mixture, leading to purification *via* column chromatography. Characterisation of the compound *via* ^1H , ^{13}C NMR and mass spectroscopy indicated that the desired product had been synthesised. The ^1H NMR spectrum showed the CH_2 peaks from the alkyl chain moved slightly up field compared to the NMR of the starting material. The biggest indicator for the formation of the compound came from the ^{13}C NMR spectrum, in which the carbon alpha to the iodine shifted significantly up field to 6 ppm from 60 ppm, due to the shielding from the iodine.

The second step involved the addition of the morpholine ring by displacing the iodide from **13**. The synthetic procedure for this reaction is the same as previously mentioned in **Section 2.3.1**. The reaction mixture underwent an aqueous work up with no complications, as TLC analysis of the organic and aqueous layers indicated that the product remained in the organic

phase. Residual morpholine and **13** remained in the crude mixture, therefore purification *via* column chromatography was required and gave a yield of 96% of isolated material. No residual morpholine was present compared to the previously mentioned reactions in **Section 2.3.1**, as **14** possess a higher *rf* value compared to **4a-i** in **Section 2.3.1**, resulting in no co-elution with morpholine.

There are several ways to synthesise a ketone. One popular method is cross-coupling reactions. These reactions typically involve the coupling between a boronic species and an activated carbonyl compound.

Two common organoboron compounds used in Suzuki-Miyaura reactions are boronic acids and boronic esters. Boronic esters are easier to handle compared to boronic acids as they are simpler to purify and characterise. There are a number of boronic esters to select from. The pinacol ester, generated from bis(pinacolato)diboron is frequently cited. The reaction typically requires a catalyst such as Pd(PPh₃)₄, Pd(OAc)₂, Pd₂(dba)₃ or Pd(dppf)Cl₂.³⁸ The mechanism (see **Figure 2.21**) involves the catalyst first being converted to an active Pd(0) catalyst. Oxidative addition occurs in which an aryl halide and a Pd(0) coordinate to form a Pd(II) species. The halide is replaced by an acetate ligand *via* a ligand exchange reaction to form an acetoxypalladium intermediate. This highly reactive intermediate then undergoes a transmetalation step with bis(pinacolato)diboron, followed by an elimination, in which the arylboronic ester is formed with regenerated Pd(0) catalyst.

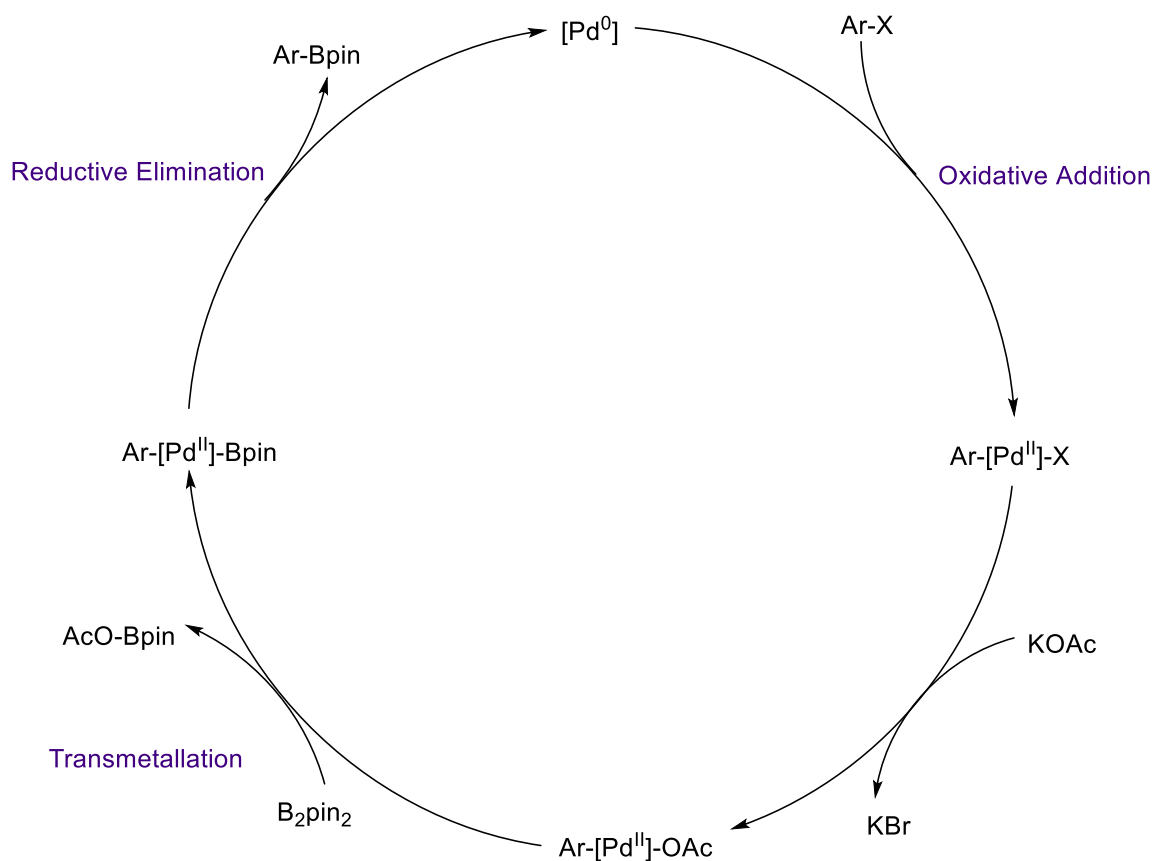


Figure 2.21: Miyaura borylation mechanism.³⁹

For the synthesis of the desired boronic ester **15b**, several reactions were trialed. The palladium catalyst used for the Miyaura borylation reaction was Pd(dppf)Cl₂, as other methods involving the popular Pd(PPh₃)₄ failed (see **Table 2.7**). The reaction involved 0.05 equivalents of the catalyst, 3 equivalents of the base K₂CO₃, 1.3 equivalents of bis(pinacolato)diboron and 1 equivalent of **14** in dioxane. The reaction appears to be scale dependent as an initial reaction on 35 mg of **14** gave a 23% yield. However, a larger scale of 0.34 g gave an acceptable yield of 64%.

Reaction	Catalyst	Diboron Equiv.	KOAc equiv.	Result (Y/N)
1	Pd(PPh ₃) ₄	1.5	3	N
2	Pd(PPh ₃) ₄	3	3	N
3	Pd(PPh ₃) ₄	3	6	N
4	Pd(dppf)Cl ₂	1.3	3	Y

Table 2.7: Trial reaction conditions for Miyaura borylation involving different catalysts, diboron and base equivalents.³⁸

For the Suzuki-Miyaura reaction, the carboxylic acid of **16** needs to be activated.³⁸ Therefore, the acid was converted to an acid chloride, a typical substrate in a cross-coupling reaction involving boronic esters. The acid chloride (**17b**) was synthesised by refluxing **16** with neat thionyl chloride to afford a yellow solid at 81% yield.⁴⁰

The boronic ester **15b** and acid chloride **17b** underwent various trial reactions in the presence of assorted palladium catalysts, bases, and solvents (see **Figure 2.22** and **Table 2.8**).^{38, 41, 42} Unfortunately, none of these reactions yielded a ketone.

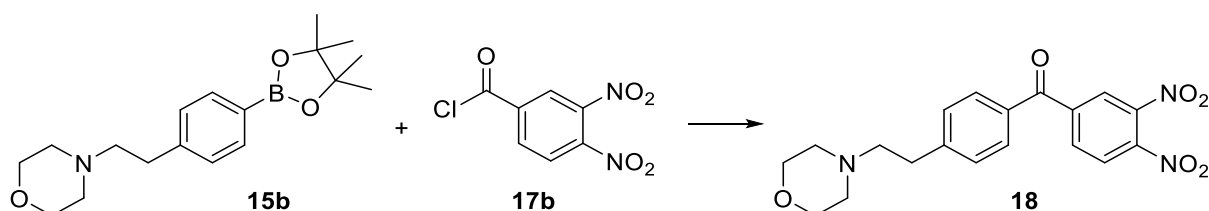


Figure 2.22: Suzuki-Miyaura cross-coupling reaction between a boronic ester (**15b**) and an acid chloride (**17b**) to give the ketone (**18**).

Reaction	Catalyst	Base	Solvent	Conditions	Result (Y/ N)
1	Pd(PPh ₃) ₄ (0.045 equiv.)	K ₂ CO ₃ (1.5 equiv.)	THF and H ₂ O	65°C, 16 h	N
2	Pd(dppf)Cl ₂ (0.05 equiv.)	K ₃ PO ₄ .H ₂ O (3 equiv.)	DMF and H ₂ O (4:1)	Rt, 16 h	N
3	Pd(dppf)Cl ₂ (0.1 equiv.)	K ₃ PO ₄ .H ₂ O (3 equiv.)	DMF and H ₂ O (4:1)	100°C, 16 h	N
4	Pd(PPh ₃) ₂ Cl ₂ (0.1 equiv.)	K ₃ PO ₄ .H ₂ O (1.5 equiv.)	Toluene	100°C, 16 h	N

Table 2.8: Trial conditions for the Suzuki-Miyaura reaction.^{38, 42}

It was proposed that the biphasic solvent system used in most of the test reactions caused the lack of conversion. Water in the presence of an acid chloride could be converted to a carboxylic acid. The carbonyl group would therefore not be sufficiently activated to allow a successful cross-coupling reaction. *Tatamiani et al* developed a Suzuki-Miyaura reaction involving a 2-pyridyl activated ester instead of an acid chloride (see **Figure 2.23**).⁴³ However, this procedure requires a boronic acid to react with **14** rather than a boronic ester.

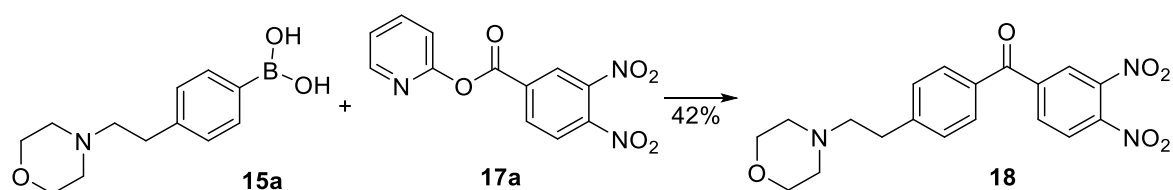


Figure 2.23: Suzuki-Miyaura cross-coupling reaction between a boronic acid (**15a**) and a 2-pyridyl activated carbonyl species (**17a**) to give the ketone (**18**).

The most common way to synthesise a boronic acid is to first convert a bromobenzene compound to either a organolithium species, using *n*-BuLi or a Grignard reagent, from magnesium.^{44, 45, 46} The intermediate can react with trialkyl borates such as B(OMe)₃, B(OEt)₃ and B(O*i*Pr)₃, followed by hydrolysis to give the final boronic acid (see **Figure 2.24**).

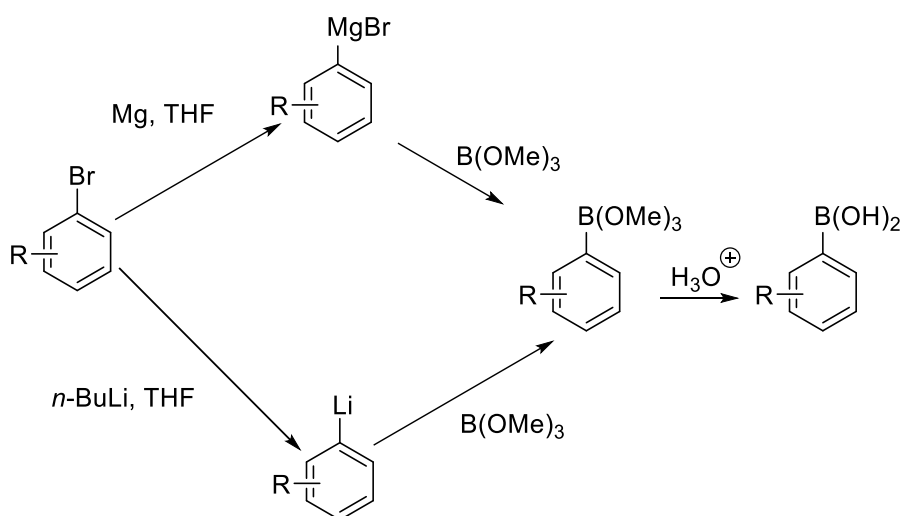


Figure 2.24: Formation of a boronic acid. Adapted from: *Base free aryl coupling of diazonium compounds and boronic esters: self-activation allowing an overall highly practical process*, H. Bonin *et al*; *The synthesis, photophysical and electrochemical properties of a series of novel 3,8,13-substituted triindole derivatives*, T. Zhu *et al.*^{44, 46}

For the synthesis of the boronic acid, this methodology was used utilising $B(OiPr)_3$ as the alkyl borate and *n*-BuLi as the base. This reaction was successful and yielded a dark brown solid. TLC analysis of the reaction mixture showed small amounts of starting materials were present. However, purification was not attempted due to the presence of the highly polar boronic acid and morpholine. Therefore, a crude mixture with a yield of 66% was carried through to the next step.

Tatamidani et al developed a synthesis using 2-pyridyl esters for Suzuki couplings due to the ester's high reactivity. They have been reported to be very capable in cross-coupling reaction as they are highly reactive towards nucleophiles, due to possessing a very good leaving group.

The synthesis of the ester involves reacting di(pyridin-2-yl) carbonate with **16**, which is catalysed by DMAP. The procedure involves leaving the reaction overnight, giving a yield of 56%. However, the yield can be slightly increased to 64% by leaving the reaction over two days. The yield is low as during aqueous workup, some of the ester (**17a**) is converted back to the acid (**16**) due to the high electrophilicity of the activated ester. Purification was attempted *via* column chromatography, sadly co-elution occurred due to the extremely close

rf values between the acid and ester. ^1H NMR suggested that crude material could be carried through to the next step as only a small amount of starting material was present.

For the formation of the ketone, a Suzuki-Miyaura cross-coupling reaction was employed. *Tatamidani et al* described the procedure in which the 2-pyridyl activated ester (**17a**) reacts with the boronic acid, **15a** in the presence of a palladium catalyst. The paper describes the use of the catalyst $\text{Pd}(\text{OAc})_2$ for the formation of the ketone (see **Figure 2.25**). The reaction proceeds by the palladium catalyst coordinating to the nitrogen of the 2-pyridyl ester. Once in close proximity, the carbonyl of the ester undergoes nucleophilic attack from the palladium to form a tetrahedral intermediate. Afterwards, C-O bond cleavage occurs with the reformation of the carbonyl and the generation of an acylpalladium. The acylpalladium undergoes transmetalation with the boronic acid, followed by reductive elimination to give the ketone and regenerate the active $\text{Pd}(0)$ catalyst.

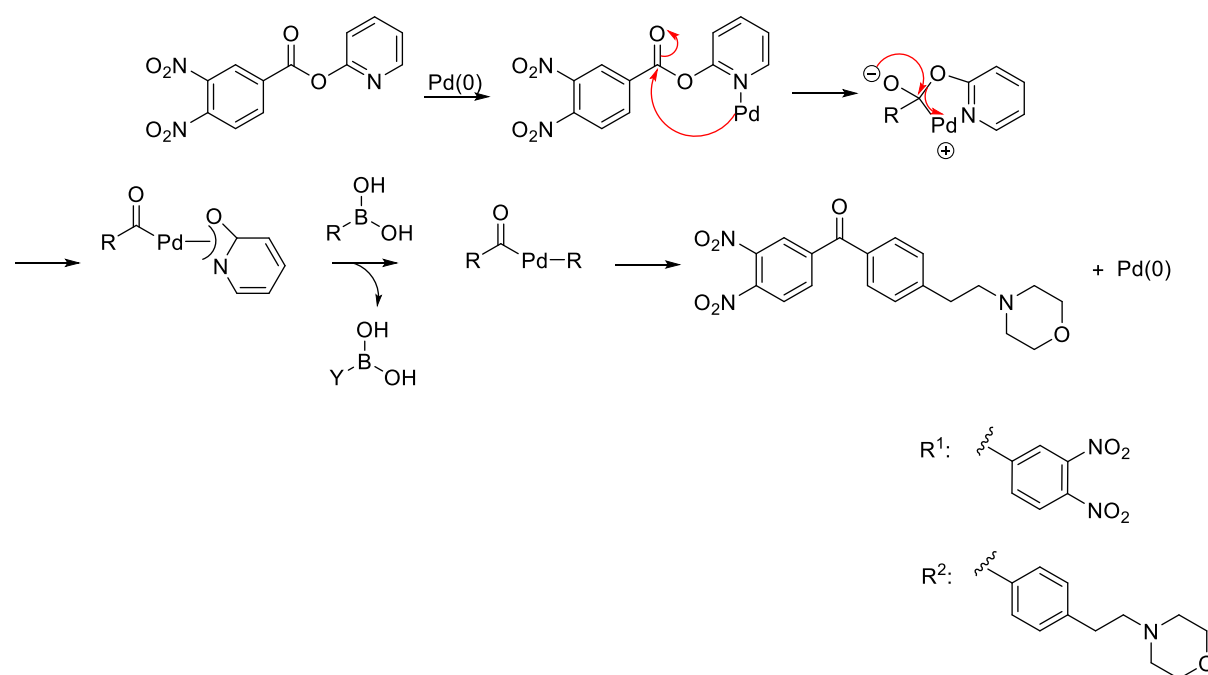


Figure 2.25: Proposed mechanism for palladium-catalysed cross-coupling reaction between 2-pyridyl ester (**17a**) and an organoboron (**15a**) compound.⁴³

Tatamidani et al report high yields between 97-63%. However, the substrates in the report differed from **18**, leading to a poorer yield of 22%. Originally the reaction was carried out in a screw fix sealed tube and left for 2 days. To improve the yield, the reaction was heated in a microwave at 110°C for 30 minutes. This improved the yield slightly to 31%. However, a

lower yield should be expected as not only is the substrate different, but both starting materials **15a** and **17a** carried trace amounts of impurity due to complicated purifications. This lowers the amount of desired starting materials in the reaction and the impurities have the potential to interfere with the reaction. TLC analysis indicated that several compounds were in the reaction mixture. The spots originated from the starting material impurities and self-coupled boronic by-products. The ketone obtained *via* column chromatography was almost pure as both TLC and NMR analysis showed only trace amount of impurity. The biggest indicator for the synthesis of the ketone came from the ^{13}C NMR spectrum as a ketone carbon peak was observed at 192.3 ppm.

Reduction of **18's** nitro groups was performed by the same procedure stated in **Section 2.3.1**. However, due to the compound having two nitro groups, the number of equivalents of $\text{SnCl}_2 \cdot \text{H}_2\text{O}$ was doubled to 10. The yield for this reaction was low at 29%. The low yield for this reaction is caused by the same issues addressed in **Section 2.3.1**. Tin waste must be removed from the reaction mixture, which unfortunately involves an aqueous work up. Some of **19** was lost to the aqueous layer due to its highly polar functional groups and is not easily retrieved. Further loss comes from purification *via* column chromatography as **19** can become stuck on the silica. The ^1H NMR confirms that the diamine had been formed, as the aromatic peaks shift upfield to 6.63 and 6.44-6.41 ppm from 8.86, 8.55, 8.49 ppm, due to the loss of the deshielding nitro group. Some impurity was present as flushing of the column was required to obtain **19** in a high enough yield. The crude material was carried forward to the final step of the synthesis.

The final step involving a ring closure of **19** to form the benzimidazole core with the carbamate attachment has been described previously in **Section 2.3.1**. This procedure performs best on a >100 mg scale. However, due to the number of steps and poor yields in the synthesis, the reaction was performed on a 21 mg scale. Fresh 1,3-bis(methoxycarbonyl)-2-methylisothiourea (**8**) was synthesised to help improve the yield of the reaction. Due to the low yield and the polar nature of the compound, precipitation had to be encouraged with a few drops of 1M NaOH. This led to a small amount of solid to precipitate, which could be filtered and washed. The same issue described in **Section 2.3.1** with the presence of a sticky oil amongst the solid was experienced. Trituration with EtOAc led to the isolation of **20** as a

beige solid. The yield for this reaction was moderate at 42%, which was a welcoming surprise from the low scale and polarity of the starting materials.

2.5.2.1 MIC Data

Compound	MIC (mg/ L)
Flubendazole	0.125
1h	0.25
20	ND- Covid-19

Table 2.9: MIC data for **20** compared to flubendazole and **1h**. Values indicated by traffic light system. Green – good - 0.015 – 0.25 mg/L, amber – acceptable - 0.5 – 4 mg/L, red – poor - >4 mg/L. ND- No data. Data has yet to be obtained due to closure of the laboratory amid the Covid-19 crisis.

Compound **20** is essentially a hybrid of **1h** and flubendazole to improve the antifungal activity and DMPK properties associated with each compound. The MIC data for **20** still needs to be obtained. The data is not yet available due to the Covid-19 crisis.

2.5.2.2 DMPK Data

Compound	LogD _{7.4}	Aqueous Solubility (μM)	Rat Heps CLint (μL/ min/ mg)	Human Heps CLint (μL/ min/ mg)
Flubendazole	2.9	0.8	39	44
1h	2.9	10	30	48
20	ND- Covid-19	ND- Covid-19	ND- Covid-19	ND- Covid-19

Table 2.10: Measured DMPK data for compound **20** compared to flubendazole and **1h**. Values indicated as the traffic light system. Green – good, amber – acceptable and red – poor. ND: No data. Data has yet to be obtained due to closure of the laboratory amid the Covid-19 crisis.

To determine whether the incorporation of the morpholine ring into the flubendazole ketone scaffold has improved the DMPK properties, DMPK data was to be obtained. The DMPK data is not yet available due to the Covid-19 crisis.

2.6 Salt Formulations

2.6.1 Introduction to *In Vivo* Data

When **1h**, was administered to mice in an *in vivo* study, some difficulties were encountered (see **Section 2.1.3.2**). When the compound was administered using different vehicles, both orally and subcutaneously, major shifts in the reduction of fungal density were not observed. The greatest antifungal effect came from an oral aqueous formulation (see **Figure 2.26**).

The improved inhibitory effect of **1h** when administered orally with the aqueous vehicle led to the investigation of fungal density at different doses (see **Figure 2.27**). There are apparent dose dependent effects, with the antifungal activity enhancing as we increase the dose from 25 to 100 mg/kg. However, even at the highest dose of 100 mg/kg, the fungal load is still in the same range as the aqueous oral formulation (150 mg/kg) seen in **Figure 2.27**. From this, we hypothesised that **1h** could be experiencing solubility issues as it only has an aqueous solubility of 10 μM (see **Table 2.3**). The low solubility, despite the presence of the morpholine ring could be preventing the compound crossing the blood brain barrier fully, leading to weak antifungal effects.

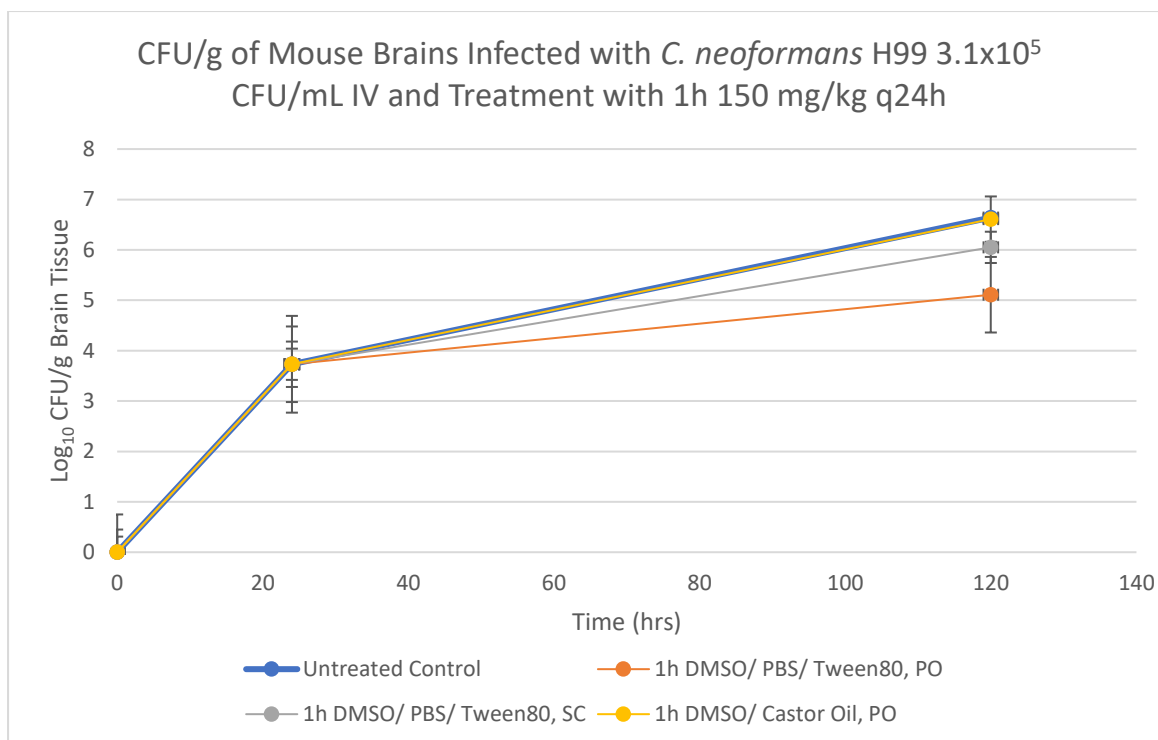


Figure 2.26: *In vitro* data of **1h** administered both orally and subcutaneously at 150 mg/kg every 24 hours for 6 days using oil and aqueous formations in mice. Untreated control shown in **black**, oral administration using an aqueous formulation of DMSO (10%)/ Tween80 (10%)/ PBS (80%) shown in **red**, subcutaneous administration using aqueous formulation shown in **green**, and oral administration using an oil formulation of DMSO (20%)/ Castor oil (80%) shown in **blue**. Data shown as average Log₁₀ CFU/ g, ± SD n=3.

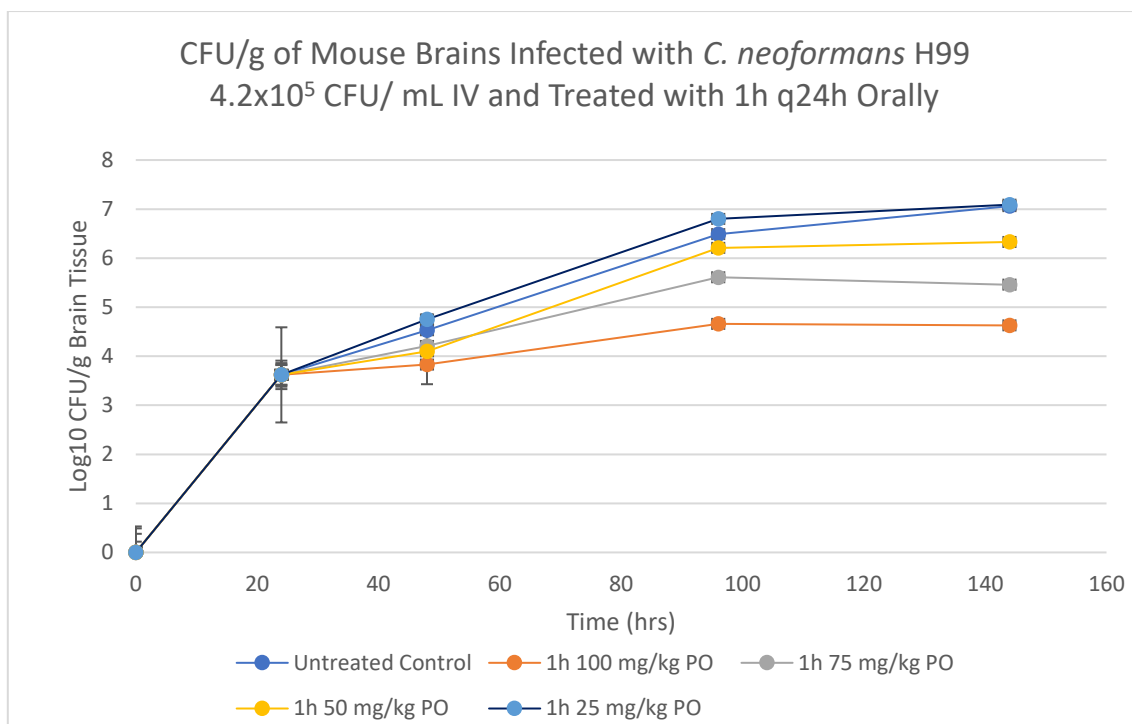


Figure 2.27: *In vitro* data of **1h** administered orally at different doses every 24 hours for 6 days in mice using the DMSO (10%)/ Tween80 (10%)/ PBS (80%) vehicle. Untreated control shown in **dark blue**, **1h** doses at 100 mg/kg shown in **orange**, 75 mg/kg shown in **grey**, 50 mg/kg shown in **yellow**, and 25 mg/kg shown in **light blue**. Data shown as average Log₁₀ CFU/ g, ± SD n=3.

There are a number of ways to improve the aqueous solubility of a drug such as complexation with cyclodextrins. However, salt formulation remains a popular choice in the pharmaceutical industry to change the rate of dissolution.^{47, 48, 49, 50}

2.6.2 Introduction to Salt Formulation

The term pharmaceutical salt refers to an ionisable drug that has been combined with a counter-ion to form a neutral complex. The pharmaceutical industry employs the use of salts for a number of reasons, one of which is to address poor aqueous solubility. Other physiochemical properties such as stability, toxicity, poor absorption and manufacturing issues can be addressed.⁵¹ In many cases salts display preferential properties compared with the parent molecule.⁵² The use of salts is wide spread, as the Food and Drug Administration (FDA) approvals of pharmaceuticals in the salt form is approximately 50%.⁴⁷ As well as this, 54% of the top 200 prescription drugs in the USA consist of a pharmaceutical salt.⁵³

There are a number of different salt formulations available to improve the physiochemical properties of a parent drug, and as well as facilitating the manufacturing process. The choice

is split into two categories: anionic or cationic counter ions. For anionic counter ions the most popular choices over the years are chloride, besylate, citrate, fumarate, maleate, tartrate, and succinate. The development of pharmaceutical salts has increased over time, with the pharmaceutical industry showing growing interest for the aforementioned anions over the years (see **Table 2.11**).⁴⁷

Anion	Overall (%)	Pre 1982 (%)	1982-1986 (%)	1987-1991 (%)	1992-2019 (%)
Acetate	3.3	1.5	8.0	12.7	
Besylate	0.8	0.4	2.0		3.3
Chloride	53.4	52.9	52.0	63.8	63.3
Citrate	2.7	2.6	2.0		3.3
Fumarate	1.7	0.4		2.1	3.3
Maleate	4.2	5.5	2.0		3.3
Succinate	1.2	0.7			3.3
Tartrate	3.8	3.7		2.1	6.7

Table 2.11: Distribution of selected anions used in pharmaceutical salts.⁴⁷

Chloride ion salts are the most favoured anionic salt and therefore would be a good starting point for the formation of **1h**. However, there are a number of different anions which gained interest between 1992-2019 such as tartrate, succinate, and maleate. Acetate ions were popular, particularly between the years 1987-1991, however, this ion appears to have been abandoned in recent years.

For cationic counter ions, the most popular choices are calcium, potassium, sodium, and tromethamine. The trends for the use of cationic ions are shown in **Table 2.12**.

Cation	Overall (%)	Pre 1982 (%)	1982-1986 (%)	1987-1991 (%)	1992-2019 (%)
Calcium	6.9	7.3			9.5
Meglumine	2.9	5.2			
Potassium	6.3	6.3			14.3
Sodium	75.3	72.9	91.7	92.3	66.7
Tromethamine	1.7			7.7	9.5
Zinc	1.2	1.0	8.3		

Table 2.12: Distribution of selected cations used in pharmaceutical salts.⁴⁷

The sodium ion dominates the use of cations in pharmaceutical salts, thus would be a good starting point. Calcium, potassium, and tromethamine are also promising ions which have gained a lot of attention in recent years. Meglumine and zinc, which were once popular, have been abandoned for over two decades and therefore will not be considered in this project.

A number of albendazole salts have been formulated to improve the aqueous solubility of benzimidazole carbamates. Work by *Bolla et al* have proposed the use of benzene sulfonic acid (BSA), *p*-toluene sulfonic acid (PTSA), and carboxylic acids such as oxalic and maleic acid.⁵⁴

2.7 Aims

The aims for this project were to further improve the aqueous solubility of **1h**.

Requirements for the project include...

1. Development of a salt from the parent compound **1h** to improve the DMPK properties.
2. Conduct solubility tests to compare the aqueous solubility to **1h**.
3. Submit **1h** salt to an *in vivo* study to assess whether the heightened aqueous solubility improved the antifungal effects.

2.8 Discussion

For this project two ions were selected for the salt formation of the parent compound **1h**: HCl, and PTSA. **1h** synthesised by WuXi was used for the formation of the salts.

2.8.1 Synthesis

2.8.1.1 PTSA Salt Formation

A variety of different methods were used to produce the PTSA salt (**21a**). Initially the procedure proposed by *Bolla et al* involved combining PTSA, **1h** and methanol (MeOH), before grinding them together in a pestle and mortar. The paste was transferred to a glass vial with a combination of MeOH and MeNO₂ as the crystallising solvent. The mixture was left for 3 days until the crystals could be harvested. This method presented a number of issues. The grinding of the materials in the pestle and mortar resulted in poor transfer from the mortar to the vial, resulting in a lower yield. Upon evaporation of the crystallising solvent, a white solid remained. The solid was transferred to a sinter funnel and washed with MeOH to separate the salt from any residual starting material. The ¹H NMR spectrum of the filtrate indicated that only a small amount of the salt formed, and its corresponding proton peaks were dwarfed by the PTSA aromatic peaks. Due to the low yield and unsuitability of the procedure, it was abandoned.⁵⁴

The procedure was modified in which **1h**, PTSA and MeOH were added to a glass vial and the ground together with a spatula. Some compound remained on the spatula, however less material was lost compared to the previous method. Unfortunately, the same issues arose from trying to harvest the crystals and similar disappointing ¹H NMR results were obtained.

An alternative procedure to synthesise the PTSA salt (**21a**) was required. Combining a 1:1 ratio of 1M PTSA in MeOH and **1h** had more success. **1h** was dissolved in the solution *via* gentle heating. The crystals were left to form at room temperature for 2 days. The crystals were harvested by washing the solid with cold MeOH, to separate the salt from any of the remaining parent compound. The ¹H NMR analysis of the filtrate showed the presence of the salt, although the initial ¹H NMR displayed a 4:1 ratio between PTSA and the salt. The solid was washed with DCM to remove excess PTSA, resulting in a salt PTSA ratio of 1:1.

2.8.1.2 HCl Salt Formation

Synthesis of the HCl salt (**21b**) involved combining a 1:1 ratio of 1M HCl in MeOH and **1h**. The suspension was gently heated until all the compound had dissolved. The solution was left at room temperature until a precipitate had formed and the crystals could be harvested. The solid was washed with cold MeOH and the filtrate was concentrated under reduced pressure. The yields for this procedure ranged from 71-92%. Increasing the scale improved the yield.

2.8.3 Solubility Tests

Solubility tests were conducted to assess whether the aqueous solubility had improved, and to gather information for the dosing regimen in the future *in vivo* work. The solubility of the salts in both water and methanol were investigated. The minimum solubility of the salts in distilled water were also determined (see **Table 2.13**).

Salt	Solvent Solubility		Minimum Solubility Dose in H ₂ O (mL)					
	H ₂ O	MeOH	0.025	0.03	0.04	0.05	0.075	0.1
1h	X	X	X	X	X	X	X	X
PTSA (21a)	Y	Y	X	X	X	X	X	Y
HCl (21b)	Y	Y	X	X	Y	Y		

Table 2.13: Solubility tests of **1h** salts. Solvent solubility 2.5 mg in 0.25 mL of solvent. Minimum solubility 1 mg in 0.025-0.1 mL of water. n=3.

The HCl salt (**21b**) was soluble in both water and methanol and would be dissolved in a minimum of 0.04 mL, indicating very promising solubility. PTSA salt (**21a**) was also able to be dissolved in both water and MeOH, however has a much higher minimum solubility of 0.1 mL. Compound **1h** was tested as a comparison to the salts. **1h** was not soluble in water, and only soluble in MeOH upon heating. Additionally, **1h** did not dissolve in any of the minimal solubility tests volumes, not even with 1 mg/mL with prolonged heating. From this preliminary test, the HCl salt (**21b**) was better performing, particularly compared to the parent compound **1h**. Therefore, the HCl salt (**21b**) was submitted as the pharmaceutical salt for an *in vivo* study.

2.8.4 Results

2.8.4.1 *In Vivo* Data

The HCl salt (**21b**) was submitted for an *in vivo* study for comparison to the parent compound (**1h**). It was believed that the salt formulation can improve the inherent solubility issues associated with benzimidazole carbamates. The poor solubility of **1h** and other benzimidazole analogues were believed to contribute to reduced antifungal activity in *in vivo* studies (see **Section 2.1.3.2**). The salt formulation should increase the rate of dissolution leading to higher drug concentrations in the cells to exhibit a therapeutic effect. The increased rate of dissolution may prevent **1h** precipitating in the mice which has been observed in previous *in vivo* studies with the parent compound. A lower fungal burden compared to the free base drug (**1h**) seen in **Figures 2.26** and **2.27** should be observed. The data collected from the *in vivo* study has not yet been made available, once released assessments can be made.

2.8.4.2 DMPK Data

Compound	LogD _{7.4}	Aqueous Solubility (μM)	Rat Heps CLint (μL/ min/ mg)	Human Heps CLint (μL/ min/ mg)
Flubendazole	2.9	0.8	39	44
1h	2.9	10	30	48
21b	3.1	2	13.9	4.19

Table 2.14: Measured DMPK data for **21b** compared to flubendazole and **1h**. Values indicated as the traffic light system. Green – good, amber – acceptable and red – poor.

The DMPK properties have not shown a vast improvement with the HCl salt formulation (**21b**). From the data analysed at AstraZeneca, the aqueous solubility has worsened. This is highly unexpected as promising results were achieved in the solubility tests (see **Table 2.14**). The rat and human metabolic clearance showed a slight improvement compared to **1h** and flubendazole, however the LogD slightly increased. Repetition of the DMPK testing is required to assess whether the data is accurate.

2.9 Metabolites

2.9.1 Concept

A number of benzimidazole derivatives have been synthesised within the group with their activity tested by MIC. A number of these compounds have shown promising results, and

hence been submitted for an *in vivo* study. Interestingly, the *in vitro* and *in vivo* data from many of these compounds were not comparable, with some demonstrating significant differences in activity. The use of different vehicles and routes of administration did not alter this trend. Therefore, it is surprising that **1h** performs so well *in vivo*, when it only has an MIC value of 4 mg/L.

Upon metabolism, a parent compound can be biotransformed into an active metabolite with enhanced biological activity. It is questioned whether the activity of **1h** stems from the improved DMPK profile or from another reason, such as forming an active metabolite.

To assess the activity of **1h** further, metabolites were synthesised and tested to check whether it possessed enhanced activity against *C. neoformans* compared to the parent **1h**.

2.9.1.1 Morpholine

2.9.1.1.1 Metabolism and Bioactivation

The main site of metabolism of benzimidazole carbamates have been discussed in **Section 2.1.1** in which hydrolysis of the carbamate and biotransformations of a thiol or ketone have been discussed. The morpholine group of **1h** allows further metabolism, of which is discussed below.

Aliphatic nitrogen heterocycles such as morpholine can be rapidly metabolised by a number of enzymes such as CYP450 and monoamine oxidase (MAO). The most common biotransformations are *N*-oxidation, *N*-conjugation, oxidative *N*-dealkylation, ring oxidation, and ring opening. These biotransformations occur at the morpholine due to the electron rich nitrogen and/ or α carbon. Hydroxylation at the α carbon yields unstable carbinolamine and iminium ion intermediates which have been implicated in bioactivation of aliphatic nitrogen heterocycles.⁵⁵ Interestingly, the morpholine ring undergoes limited metabolism compared to other aliphatic nitrogen heterocycles. Furthermore, bioactivation is rarely reported with morpholine rings, with doxorubicin being transformed into 2(*S*)-methoxy-4-morpholinyl doxorubicin (MMDX) being one of only a few examples. Doxorubicin is an anthracycline used as an anti-cancer agent. When MMDX was incubated with an ovarian cancer cell line in the presence of NADPH and liver microsomes the efficacy significantly increased. This suggests that that MMDX has become bioactivated.⁵⁶ Enzymatic α -hydroxylation of the morpholine

ring in compound **22** yields carbinolamine (**23**) which forms an electrophilic iminium ion (**24**), a species that could effectively cross-link DNA (nucleophile) *in vivo* (see **Figure 2.28**).^{55, 57, 58}

Glucuronidation is a common phase II metabolite pathway in many heterocyclic compounds. However, it is very rare in morpholine containing drugs.⁵⁵ Only morinidazole is reported to undergo glucuronidation at its tertiary nitrogen yielding the major circulating human metabolite.⁵⁹

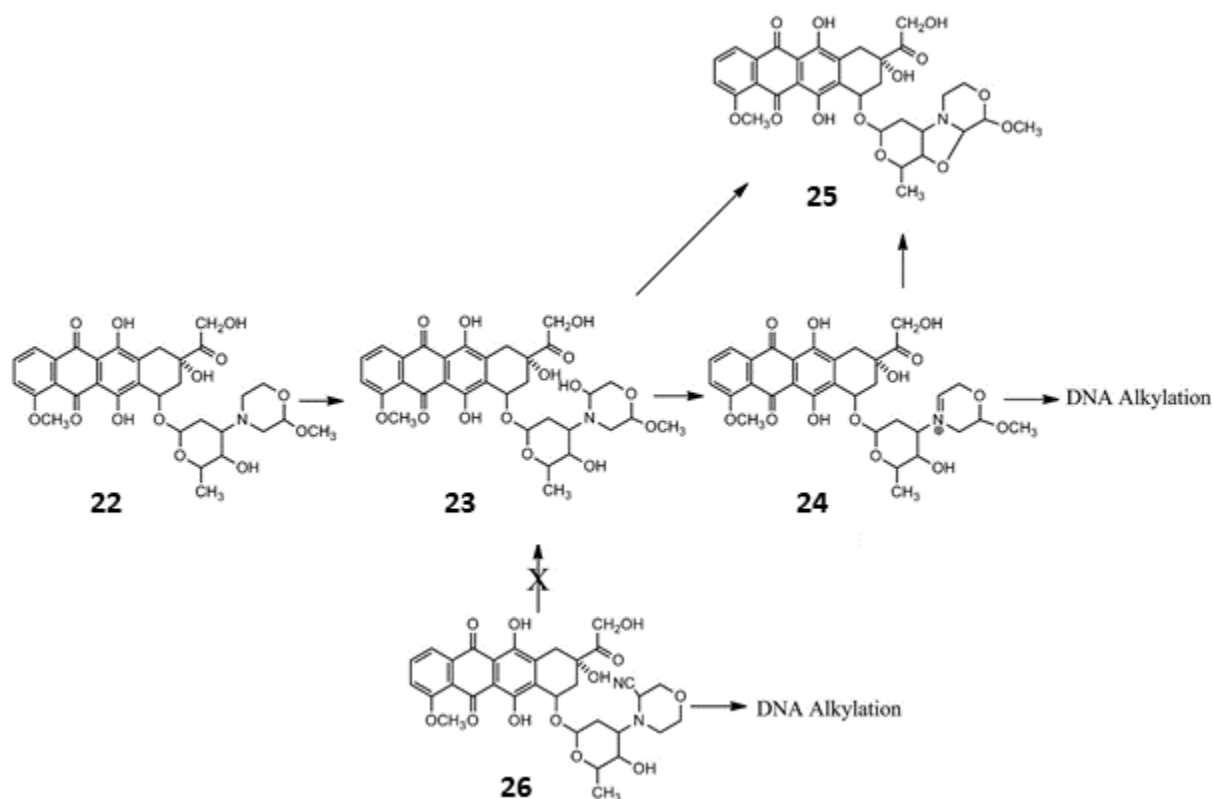


Figure 2.28: Bioactivation of methoxymorpholinyl doxorubicin (MMDX). Reproduced from: *Biotransformation and bioactivation reactions of alicyclic amines in drug molecules*, J. Bolleddula et al.⁵⁵

2.10 Aims

The aims for this project were to investigate the metabolism of **1h**.

Requirements for the project include...

1. Synthesis of potential **1h** metabolites.
2. Submit **1h** metabolites for MIC testing to assess whether they have improved antifungal activity.
3. If the theory that **1h** can be bioactivated *in vivo* is supported, the chosen metabolite will be submitted for an *in vivo* study to assess to further improve antifungal activity.

2.11 Discussion

2.11.1 Synthesis

2.11.1.1 Carbamate Hydrolysis Metabolite

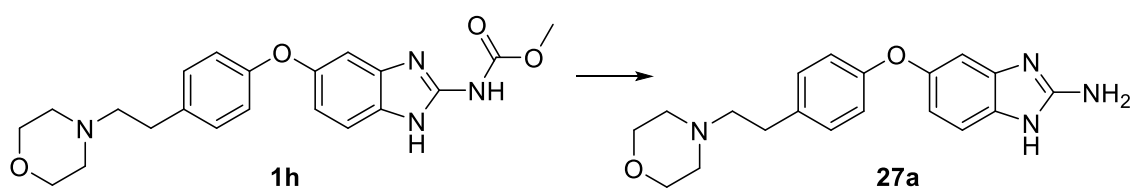


Figure 2.29: Conversion of compound **1h** to the **1h-NH₂** derivative (**27a**)

For the synthesis of the FLU-NH₂ derivative (**27a**), a procedure by *Kaipnazarov et al* was employed.⁶⁰ A suspension of **1h**, water and NaOH was heated at reflux to cleave the carbamate group (see **Figure 2.29**). The resultant solid from the reaction was characterised *via* ¹H NMR and showed that the desired product had not been synthesised as the carbamate methyl peak at 3.81 ppm was still present. The reaction failed as **1h** did not fully dissolve in the reaction solvent, and instead was suspended in water.

An alternative procedure by *Glunz et al* used a combination of water and methanol.⁶¹ Upon reflux, **1h** dissolved in the solvent mixture and the reaction was monitored by TLC analysis for 3 days, by which time the starting material was completely converted to the desired metabolite, with a yield of 48%. The product was characterised *via* ¹H NMR and the characteristic methyl carbamate peak was no longer present. ¹³C NMR, mass spectrometry, IR analysis, and HPLC data was gathered for the metabolite, confirming that the compound had been synthesised with an 89% purity.

2.11.1.2 N-Oxide Metabolite

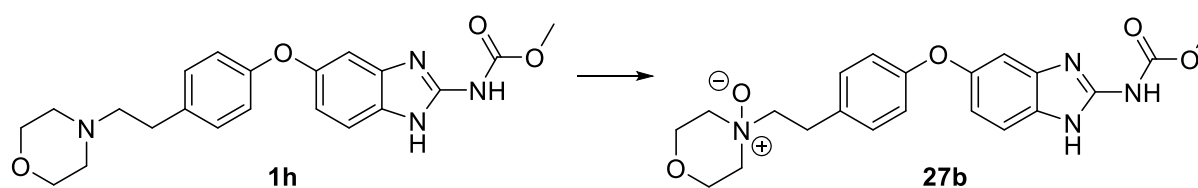


Figure 2.30: Conversion of compound **1h** to the **1h N-oxide** derivative (**27b**)

A procedure from *Khuthier et al* was employed to synthesise the *N*-oxide metabolite (**27b**) (see **Figure 2.30**).⁶² The procedure involved mixing **1h**, *m*-CPBA, and chloroform at room temperature. The reaction was unable to go into completion, therefore purification *via* column chromatography was required to yield a white solid at 69%. The ¹H NMR of the metabolite and the parent compound are indistinguishable and therefore gives limited information about the formation of the metabolite. The ¹³C NMR indicated that the product had been synthesised as the carbon peaks have shifted up field due to the presence of the *N*-oxide. This result has been further confirmed by mass spectroscopy, IR, and HPLC analysis, of which the compound was obtained with an 97% purity.

2.11.2 Results

2.11.2.1 MIC Data

The metabolites of **1h** were submitted for MIC testing to assess whether the metabolites exhibited any enhanced activity against *C. neoformans* (see **Table 2.15**). The **1h-NH₂** metabolite (**27a**) showed no activity. This is expected as the carbamate was thought to be an essential part of the structure. Commercially available benzimidazoles without the carbamate such as thiabendazole and benzimidazole possessed no activity against various strains of *C. neoformans* (see **Chapter 1, Table 1.1**). The *N*-oxide metabolite of **1h** (**27b**) displayed poor antifungal activity and was therefore not considered to be bioactivated.

Compound	EUCAST (mg/L)	Upper Limit of Assay
Flubendazole	0.125	16
1h	0.25	16
27a	>16	16
27b	4	16

Table 2.15: MIC data for metabolites **27a** and **27b** compared to flubendazole and **1h**. Values indicated as the traffic light system. Green – good - 0.015 – 0.25 mg/L, amber – acceptable - 0.5 – 4 mg/L, red – poor - >4 mg/L.

2.11.2.2 DMPK Data

Compound	LogD _{7.4}	Aqueous Solubility (μM)	Rat Heps CLint (μL/ min/ mg)	Human Heps CLint (μL/ min/ mg)
Flubendazole	2.9	0.8	39	44
1h	2.9	10	30	48
27a	2.4	437	17.4	<3
27b	NA	NA	NA	NA

Table 2.16: Measured DMPK data for metabolites **27a** and **27b** compared to flubendazole and **1h**. Values indicated as the traffic light system. Green – good, amber – acceptable and red – poor. NA: not available.

Metabolite **27a** possesses an impressive DMPK profile as the aqueous solubility was considerably higher than flubendazole and the lead **1h** compound. The carbamate group has been attributed to flubendazole and its derivatives lack of solubility. The loss of this group has led to a considerable improvement in solubility compared to flubendazole. Although, the carbamate group is essential for the activity against *C. neoformans*, and its necessity was identified in the initial SAR assessment of benzimidazoles (see **Chapter 1, Table 1.1**). This is further confirmed by the lack of activity of metabolite **27a** (see **Table 2.15**). Possible mimics of the carbamate group could be investigated to assess whether the vital interactions of the carbamate group can be replicated in a more water-soluble group. The DMPK data for **27b** was not obtained due to standard analytical methods at AstraZeneca not responding to the *N*-oxide metabolite. An adapted DMPK method will be developed to gain DMPK data in the future.

The MIC results indicate that bioactivation may not be the cause for the positive results of **1h** in the *in vivo* study, and instead be due to the slightly improved DMPK properties. Only two metabolites have been investigated, and other metabolites can be synthesised and assessed to evaluate whether they have contributed to **1h**'s promising *in vivo* results.

2.12 Conclusion

In conclusion, several flubendazole analogues have been synthesised for SAR exploration. Previously, compound **1h** displayed promising results in terms of the DMPK profile. The aqueous solubility was 10 μM , which is a vast improvement from flubendazole's 0.8 μM . The use of a polar nitrogen containing heterocycle provided this improvement. To test whether the aqueous solubility can be further improved, various analogues of **1h** were synthesised. The morpholine of the **1h** ether series was replaced with various heterocyclic groups such as 2-oxa-6-azaspiro[3.3]heptane (**3a**) and pyrrolidine (**3b**). There are delays in most of the MIC and DMPK data for these compounds due to the Covid-19 crisis. However, compound **9b** displayed an impressive improvement in the aqueous solubility of 876 μM . Unfortunately, this analogue did not possess any antifungal activity as the MIC value was >32 mg/L. This suggests that further optimisation is required to find a new lead compound.

The MIC value of **1h** was lower than flubendazole. This may be due to favourable interactions between the ketone's oxygen and the amino acid residues in the tubulin binding site. A hybrid (**20**) was synthesised to combine the lower MIC value of flubendazole with the increased aqueous solubility of **1h**. The MIC and DMPK data for **20** has yet to be obtained due to the laboratory lockdown amid the Covid-19 crisis. Once data has been received, the compounds SAR can be assessed, which may result in a new lead compound. However, the alternative heterocyclic analogues (**9a-i**) of **1h**, may perform better than the morpholine and therefore be incorporated into the ketone scaffold for further improvement in the MIC and DMPK profile.

Salt formulations were tried and tested to further improve the solubility and handling of **1h** in *in vivo* testing. The HCl salt provided enhanced solubility compared to **1h** during solubility tests. Although, the DMPK data obtained from AstraZeneca showed a lower aqueous solubility compared to the parent compound **1h**. Further testing is required to see whether the DMPK data is reliable.

2.13 Future Work

In terms of future work, MIC and DMPK data for the morpholine derivatives (**9a-i**) and the ketone analogue (**20**) will be obtained once normal laboratory work returns after the Covid-19 crisis. This information is essential to assess whether the project is headed in the right direction to improve the aqueous solubility of benzimidazoles, while also maintaining antifungal activity.

Analysis of the DMPK data for the **1h** metabolite (**27a**) has shown that the aqueous solubility of the compound can vastly improve with the absence of the carbamate group. Unfortunately, a loss of this moiety causes a lack of antifungal activity against *C. neoformans*. The carbamate groups could be replaced with a bioisostere to improve the solubility of the compound, while maintaining the antifungal activity. Minimal work has been conducted, as only two compounds containing a urea moiety have been synthesised so far (**1o** and **1p** (see **Table 2.3**), as they are known to improve solubility.⁶³ Unfortunately, both **1o** and **1p** possessed poor antifungal activity (see **Table 2.3**), which suggests that even small changes to the carbamate structure may lead to a loss in activity. Further exploration in the SAR may be beneficial, as other carbamate bioisosteres such as a sulphonamide or amide groups can be synthesised (see **Figure 2.31**).^{64, 65} The choice of carbamate isosteres are limited as carbamates are often used as bioisosteres to enhance a compounds drug like properties. Therefore, only limited work can be attempted in this area.

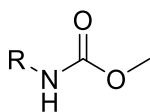
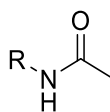
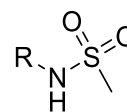
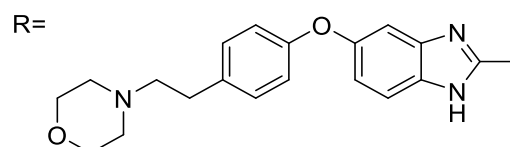
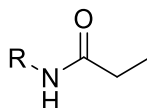
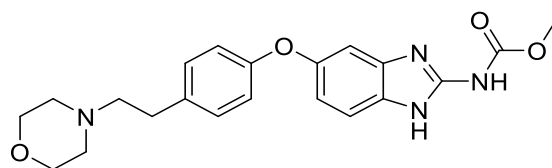
2.31A**2.31B****2.31D****2.31C**

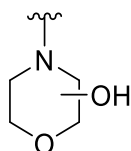
Figure 2.31: Possible carbamate bioisosteres. **2.31A:** Original carbamate moiety. **2.31B:** Methyl amide moiety. **2.31C:** Ethyl amide moiety. **2.31D:** Sulfonamide moiety.

The metabolite investigation is limited as only two possible metabolites have been synthesised. They were required to assess whether bioactivation could be the cause of **1h**'s enhanced antifungal activity *in vivo*, compared to the previous *in vitro* findings. Only the *N*-oxide (**27b**) and the cleaved carbamate (**27a**) metabolites have been investigated and showed no antifungal activity against *C. neoformans*. Therefore, further investigation is required to fully determine whether bioactivation plays a role in the improved activity. Other metabolites to be synthesised include oxidation, hydroxylation, and degraded morpholine products (see **Figure 2.32**).^{59, 66, 28}

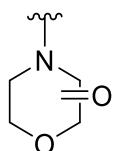
1h



2.32A



2.32B



2.32C

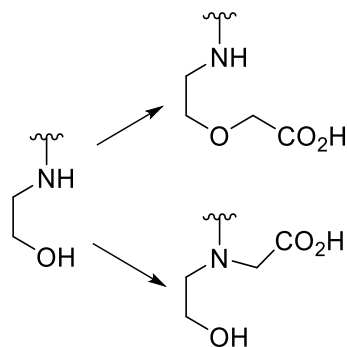


Figure 2.32: Possible metabolites of 1h. **2.32A:** Hydroxylation product.^{59, 66} **2.32B:** Oxidation product.^{59, 66} **2.32C:** Morphine link degradation leading to the formation of carboxylic acids.^{59, 28}

2.14 Experimental

2.14.1 General Experimental Details

2.14.1.1 Chemicals

All chemicals and solvents were purchased from Sigma Aldrich, Fluorochem, Alfa Aesar, Fischer, and Acros Organics with no further purification required before use. Anhydrous DMF and DCM was obtained from an MBRAUN MBSPS5 solvent purification system.

2.14.1.2 Thin Layer Chromatography

TLC was carried out on Merck silica gel 60 F-254 aluminium backed plates. Compounds were observed *via* exposure to UV light or treatment with appropriate staining solutions such as potassium permanganate or ninhydrin, and then developed *via* a heat gun.

2.14.1.3 Column Chromatography

Column chromatography was performed using silica (40-63 μM) supplied by Sigma Aldrich. All columns used had the silica made into a slurry with the appropriate solvent system to provide the column mobile phase and layered onto sand. The product was absorbed onto silica and loaded onto the top of the mobile phase. Column was then eluted using the desired eluent system and the desired fractions were collected.

2.14.1.4 Compound Characterisation

Compounds were characterised by ^1H (400 MHz) and ^{13}C (101MHz) NMR spectra were recorded on a Bruker AMX400 spectrometer (^1H 400 MHz; ^{13}C 101 MHz) in the required deuterated solvent. Deuterated solvents were obtained from Sigma Aldrich with an internal standard tetramethyl silane (TMS). Chemical shifts are reported as ppm (parts per million) and reported from downfield to upfield. Coupling constant (J) are reported in Hz. Splitting reported as singlet – s, broad singlet – br s, multiplet – m, d – doublet, t – triplet, q – quartet, s – septet, dd – doublet of doublets, ddd – doublet of doublet of doublets.

Low Resolution Mass spectrometry (LRMS) and High-Resolution Mass Spectrometry (HRMS) were recorded using the analytical service within the Chemistry Department at the University of Liverpool. LRMS and HRMS were conducted on a VG analytical 7070E machine, Frisons TRIO mass spectrometers or Agilent QTOF 7200 using chemical ionisation (CI) or electrospray (ES).

Infra-red spectra were recorded in the range of 4000-500 cm^{-1} using a JASCO FT/IR 4200 spectrometer, or Bruker ALPHA FT-IR platinum ATR spectrometer.

Melting points (MP) were carried out using a Gallenkamp melting point machine and values were recorded in degrees Celsius with three concordant results.

HPLC was carried out using Agilent 1200 HPLC equipped with a ZORBAX Eclipse Plus C18 column (4.6mm x 10 mm, 3.5 μm) at 25 $^{\circ}\text{C}$. Flow rate 1 mL/min for 15 minutes using MeCN (+0.05% Formic Acid (FA))/Water (+0.05% Formic Acid) with compounds dissolved in methanol. UV detector recorded signals at 254 nm. Method: min, gradient: 2% MeCN (+0.05% FA) hold to 1 min, 2-98% MeCN (+0.05% FA) in 11 min, then hold at 98% MeCN (+0.05% FA) to 15 min.

2.14.2 General Procedures

General Procedure A: Nucleophilic Addition of Alicyclic Nitrogen Heterocycles.¹⁹¹

4-(2-Bromoethyl)phenol (**2**) (1.0 equiv.), the selected amine (**3a-i**) (1.1 equiv.), DIPEA (1.5 equiv.) and anhydrous MeCN (0.25 M) were allowed to stir at reflux for 16 hours under a nitrogen atmosphere. After this time, the reaction was cooled and concentrated under reduced pressure. The crude solid was filtered and washed with DCM. The filtrate was concentrated *in vacuo*. Purification *via* column chromatography eluting with 0-10% MeOH in EtOAc afforded the desired product (**4a-i**).

General Procedure B: $\text{S}_{\text{N}}\text{Ar}$ Reaction¹⁹¹

Compound **4a-i** (1.5 equiv.) and potassium carbonate (2.0 equiv.) were suspended in anhydrous DMF (0.25 M) and allowed to stir at room temperature for 15 minutes under a nitrogen atmosphere. 5-Fluoro-2-nitroaniline (**5**) (1.0 equiv.) was added to the flask and allowed to stir at reflux for 16 hours. After this time, the reaction was concentrated under reduced pressure. The residue was diluted in EtOAc and washed with H_2O (x1), saturated aq. NaHCO_3 solution (x1), and brine (x1). The organic phase was dried over MgSO_4 , filtered, and concentrated *in vacuo* to yield an orange/ yellow oil. Purification *via* column chromatography eluting with 0-20% MeOH in EtOAc afforded the desired product (**6a-i**).

General Procedure C: Single Nitro Reduction¹⁹²

Compound **6a-i** (1 equiv.), tin (II) chloride dihydrate (5.0 equiv.) and ethanol (0.15 M) were added to a flask and allowed to stir at reflux for 16 hours. After this time, the reaction was concentrated under reduced pressure. The residue was diluted with EtOAc, the pH was adjusted to 9-10 with 25% NaOH solution, filtered, and washed with EtOAc. The filtrate was washed with H₂O (x3) and brine (x1). The organic phase was dried over MgSO₄, filtered, and concentrated *in vacuo* to yield a crude brown oil. Purification *via* column chromatography eluting with 0-20% MeOH in EtOAc afforded the desired product (**7a-i**).

General Procedure D: Synthesis of 1,3-Bis(Methoxycarbonyl)-2-methylisothiourea¹⁹⁷

2-Methyl-2-thiopseudourea hemisulphate salt (**10**) (1 equiv.) was suspended in distilled water (1 M) and cooled to 0°C. Methyl chloroformate (**11**) (4.3 equiv.) was added and the reaction was vigorously stirred at 0°C for 20 minutes. After this time, the reaction was allowed to warm to room temperature and basified to pH 9-10 using 25% NaOH aqueous solution. The white precipitate was filtered and washed with cold water to yield a white solid (**8**). This intermediate was not characterised and carried through to the subsequent reaction.

General Procedure E: Ring Closure Reaction to Form Benzimidazole Carbamate¹⁹²

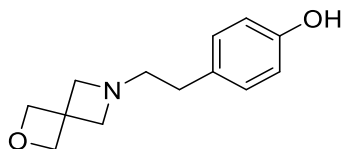
Compound **7a-i** (1.0 equiv.), 1,3-bis(methoxycarbonyl)-2-methylisothiourea (**8**) (1.5 equiv.), acetic acid (2 M) and methanol (0.4 M) were added to a screw fix sealed tube and allowed to stir at reflux for 16 hours. After this time, the reaction was cooled and concentrated under air. The residue was suspended in diethyl ether, filtered, and washed with diethyl ether to afford the desired product (**9a-i**).

If precipitation does not occur, formation of a solid can be encouraged *via* the addition of 1M NaOH, followed by sinter filtration and washing with diethyl ether. Any impurities were removed *via* trituration with small amounts of acetone or ethyl acetate.

2.14.2.1 SAR Morpholine Derivatives

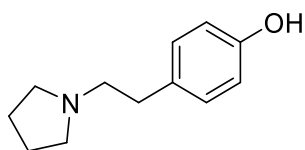
2.14.2.1.1 4-(Amino)phenol Derivatives

4a: 4-(2-(2-Oxa-6-azaspiro[3.3]heptan-6-yl)ethyl)phenol



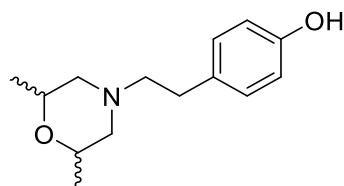
Reaction repeated as general procedure A using 2-oxa-6-azaspiro[3.3]heptane (**3a**) gave the title compound (**4a**) as a brown solid (0.11 g, 100%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.02 (d, 2H, *J* = 8.0 Hz, Ar-H), 6.72 (d, 2H, *J* = 8.0 Hz, Ar-H), 4.72 (s, 4H, CH₂), 3.45 (s, 4H, CH₂), 2.68 (t, 2H, *J* = 8.0, 16.0 Hz, CH₂), 2.56 (t, 2H, *J* = 8.0, 16.0 Hz, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 155.6, 129.7, 129.2, 114.9, 80.7, 63.0, 34.0, 32.5. HRMS (ES+) *m/z* calculated for C₁₃H₁₈NO₂: 220.1337. Found [M+H]⁺: 220.1335 (Diff -0.91 ppm).

4b: 4-(2-(Pyrrolidin-1-yl)ethyl)phenol²³⁶



Reaction repeated as general procedure A using pyrrolidine (**3b**) gave the title compound (**4b**) as a brown oil (0.15 g, >100%). This intermediate was not characterised and carried through to the subsequent reaction.

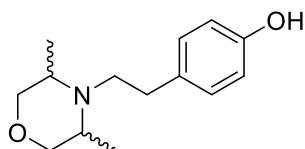
4c: 4-(2-(2,6-Dimethylmorpholino)ethyl)phenol (relative stereochemistry unknown)



Reaction repeated as general procedure A using 2,6-dimethylmorpholine (**3c**) gave the title compound (**4c**) as a white solid (0.57 g, 49%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.04 (d, 2H, *J* = 8.0 Hz, Ar-H), 6.72 (d, 2H, *J* = 8.0 Hz, Ar-H), 3.75-3.67 (m, 2H, C-H), 2.91-2.88 (m, 2H, CH₂), 2.75-2.71 (m, 2H, CH₂), 2.56-2.52 (m, 2H, CH₂), 1.79 (t, 2H, *J* = 12.0, 24.0 Hz, CH₂), 1.16 (d, 6H, *J* = 8.0 Hz, (CH₃)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 155.1, 130.4, 129.1, 114.8, 71.1, 60.4, 59.0, 31.1,

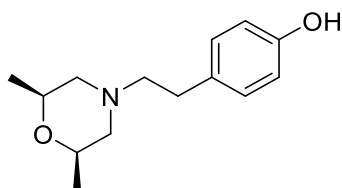
17.9. HRMS (ES+) m/z calculated for $C_{14}H_{22}NO_2$: 236.1652. Found $[M+H]^+$: 236.1648 (Diff – 1.69 ppm).

4d: 4-(2-(3,5-Dimethylmorpholino)ethyl)phenol (relative stereochemistry unknown)



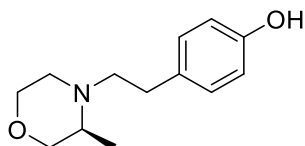
Reaction repeated as general procedure A using 3,5-dimethylmorpholine (**3d**) gave the title compound (**4d**) as a white solid (78 mg, 21%). The material was carried through to the next step without any further purification. HRMS (ES+) m/z calculated for $C_{14}H_{22}NO_2$: 236.1652. Found $[M+H]^+$: 236.1649 (Diff – 1.27 ppm).

4e: 4-(2-(*cis*-2,6-Dimethylmorpholino)ethyl)phenol



Reaction repeated as general procedure A using *cis*-2,6-dimethylmorpholine (**3e**) gave the title compound (**4e**) as a white solid (0.17 g, 49%). 1H NMR (400 MHz, CD_3OD-d_4) δ 6.91 (d, 2H, $J = 8.0$ Hz, Ar-H), 6.60 (d, 2H, $J = 8.0$ Hz, Ar-H), 3.62-3.55 (m, 2H, C-H), 2.77 (m, 2H, CH_2), 2.63-2.59 (m, 2H, CH_2), 2.44-2.39 (m, 2H, CH_2), 1.66 (t, 2H, $J = 12.0, 20.0$ Hz, CH_2), 1.04 (d, 6H, $J = 8.0$ Hz, $(CH_3)_2$). ^{13}C NMR (101 MHz, CD_3OD-d_4) δ 155.4, 130.4, 129.2, 114.9, 71.3, 60.5, 59.0, 31.5. HRMS (ES+) m/z calculated for $C_{14}H_{22}NO_2$: 236.1652. Found $[M+H]^+$: 236.1644 (Diff -3.39 ppm).

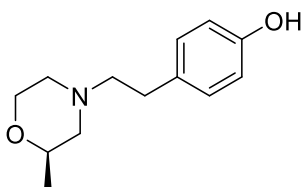
4f: (S)-4-(2-(3-Methylmorpholino)ethyl)phenol



Reaction repeated as general procedure A using (*S*)-3-methylmorpholine (**3f**) gave the title compound (**4f**) as a white solid (0.44 g, 80%). 1H NMR (400 MHz, CD_3OD-d_4) δ 7.04 (d, 2H, $J = 8.0$ Hz, Ar-H), 6.72 (d, 2H, $J = 8.0$ Hz, Ar-H), 3.84-3.81 (m, 2H, CH_2), 3.71-3.63 (m, 2H, CH_2), 3.26

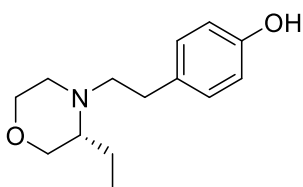
(t, 2H, $J = 8.0, 20.0$ Hz, CH₂), 2.95-2.80 (m, 2H, CH₂), 2.61-2.56 (m, 3H, CH₂, C-H), 1.03 (s, 3H, CH₃). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 155.4, 130.7, 129.2, 114.9, 72.1, 66.6, 55.8, 54.9, 50.9, 30.4, 12.6. HRMS (ES+) m/z calculated for C₁₃H₂₀NO₂: 222.1496. Found [M+H]⁺: 222.1490 (Diff – 2.70 ppm).

4g: (*R*)-4-(2-(2-Methylmorpholino)ethyl)phenol



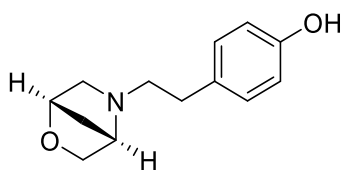
Reaction repeated as general procedure A using (*R*)-2-methylmorpholine (**3g**) gave the title compound (**4g**) as a yellow solid (0.39 g, 78%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.04 (d, 2H, $J = 8.0$ Hz, Ar-H), 6.72 (d, 2H, $J = 8.0$ Hz, Ar-H), 3.89-3.86 (m, 1H, C-H), 3.71-3.66 (m, 2H, CH₂), 2.95-2.87 (m, 2H, CH₂), 2.75 (t, 2H, $J = 8.0, 16.0$ Hz, CH₂), 2.59 (t, 2H, $J = 8.0, 16.0$ Hz, CH₂), 2.26-2.17 (m, 1H, CH₂), 1.91 (t, 1H, $J = 12.0, 20.0$ Hz, CH₂), 1.16 (s, 3H, CH₃). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 155.4, 130.2, 129.2, 114.9, 71.4, 65.8, 60.6, 59.6, 52.5, 31.4, 17.9. HRMS (ES+) m/z calculated for C₁₃H₂₀NO₂: 222.1496. Found [M+H]⁺: 222.1490 (Diff – 2.70 ppm).

4h: (*R*)-4-(2-(3-Ethylmorpholino)ethyl)phenol



Reaction repeated as general procedure A using (*R*)-3-ethylmorpholine (**3h**) gave the title compound (**4h**) as a brown oil (37.2 mg, 31%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 6.91 (d, 2H, $J = 8.0$ Hz, Ar-H), 6.60 (d, 2H, $J = 8.0$ Hz, Ar-H), 3.70-3.60 (m, 1H, CH₂), 3.56-3.50 (m, 1H, CH₂), 3.47-3.36 (m, 1H, CH₂), 3.29-3.24 (m, 1H, CH₂), 2.81-2.74 (m, 2H, CH₂), 2.65-2.58 (m, 1H, CH₂), 2.57-2.50 (m, 1H, CH₂), 2.48-2.42 (m, 2H, CH₂), 2.29-2.23 (m, 1H, C-H), 1.63-1.54 (m, 2H, CH₂), 0.79 (t, 3H, $J = 4.0, 12.0$ Hz, CH₃). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 151.9, 136.4, 131.0, 117.8, 72.2, 70.0, 68.0, 58.3, 56.4, 29.6, 19.7, 10.9. HRMS (ES+) m/z calculated for C₁₄H₂₂NO₂: 236.1652. Found [M+H]⁺: 236.1650 (Diff – 0.85 ppm).

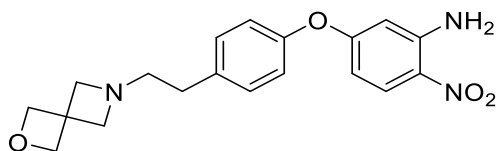
4i: 4-(2-((1S,4S)-2-Oxa-5-azabicyclo[2.2.1]heptan-5-yl)ethyl)phenol



Reaction repeated as general procedure A using (1S,4S)-2-oxa-5-azabicyclo[2.2.1]heptane (**3i**) gave the title compound (**4i**) as a yellow solid (27.6 mg, 28%). The material was carried through to the next step without any further purification. HRMS (ES+) m/z calculated for $C_{13}H_{18}NO_2$: 220.1339. Found $[M+H]^+$: 220.1331 (Diff – 3.63 ppm).

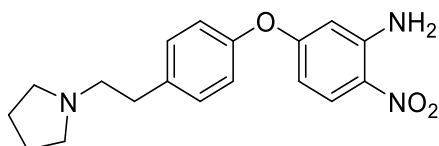
2.14.2.1.2 2-Nitro-5-((amino)phenoxy)aniline Derivatives

6a: 5-(4-(2-(2-Oxa-6-azaspiro[3.3]heptan-6-yl)ethyl)phenoxy)-2-nitroaniline



Reaction repeated as general procedure B using 4-(2-(2-oxa-6-azaspiro[3.3]heptan-6-yl)ethyl)phenol (**4a**) gave the title compound (**6a**) as a yellow solid (0.12 g, 68%). 1H NMR (400 MHz, CD_3OD-d_4) δ 8.07 (d, 1H, $J = 12.0$ Hz, Ar-H), 7.31 (d, 2H, $J = 8.0$ Hz, Ar-H), 7.05 (d, 2H, $J = 8.0$ Hz, Ar-H), 6.32 (d, 1H, $J = 4.0$ Hz, Ar-H), 6.28 (dd, 1H, $J = 4.0, 12.0$ Hz, Ar-H), 4.76 (s, 4H, CH_2), 3.51 (s, 4H, CH_2), 2.76 (t, 2H, $J = 8.0, 16.0$ Hz, CH_2), 2.69 (t, 2H, $J = 8.0, 12.0$ Hz, CH_2). ^{13}C NMR (101 MHz, CD_3OD-d_4) δ 163.8, 153.7, 143.0, 133.1, 130.1, 127.7, 125.2, 120.6, 106.2, 100.0, 80.7, 63.0, 60.0, 39.0, 32.5. HRMS (ES+) m/z calculated for $C_{19}H_{22}N_3O_4$: 336.1610. Found $[M+H]^+$: 336.1609 (Diff -0.30 ppm).

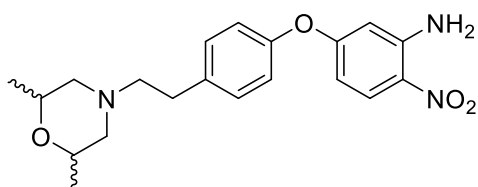
6b: 2-Nitro-5-(4-(2-(pyrrolidin-1-yl)ethyl)phenoxy)aniline



Reaction repeated as general procedure B using 4-(2-(pyrrolidin-1-yl)ethyl)phenol (**4b**) gave the title compound (**6b**) as a yellow oil (0.15 g, 58%). 1H NMR (400 MHz, $CDCl_3$) δ 8.09 (d, 1H, $J = 8.0$ Hz, Ar-H), 7.25 (d, 2H, $J = 8.0$ Hz, Ar-H), 7.0 (d, 2H, $J = 8.0$ Hz, Ar-H), 6.32 (dd, 1H, $J = 4.0, 8.0$ Hz, Ar-H), 6.20 (br s, 2H, NH_2), 6.17 (d, 1H, $J = 4.0$ Hz, Ar-H), 2.93-2.80 (m, 4H, CH_2), 2.71

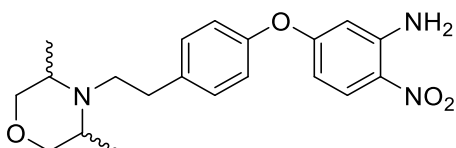
(m, 4H, CH₂), 1.87 (m, 4H, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 164.4, 152.8, 146.9, 137.0, 130.3, 128.7, 127.7, 120.9, 107.7, 103.3, 57.9, 54.1, 64.5, 23.5. HRMS (ES+) *m/z* calculated for C₁₈H₂₂N₃O₃: 328.1661. Found [M+H]⁺: 328.1662 (Diff 0.30 ppm).

6c: 5-(4-(2-(2,6-Dimethylmorpholino)ethyl)phenoxy)-2-nitroaniline (relative stereochemistry unknown)



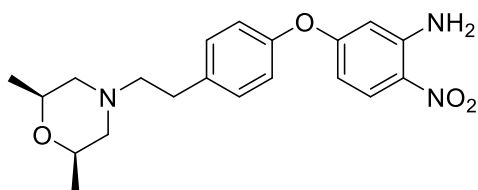
Reaction repeated as general procedure B using 4-(2-(2,6-dimethylmorpholino)ethyl)phenol (**4c**) gave the title compound (**6c**) as a yellow solid (0.48 g, 54%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.07 (d, 1H, *J*= 12.0 Hz, Ar-H), 7.33 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.05 (d, 2H, *J*= 8.0 Hz, Ar-H), 6.30-6.27 (m, 2H, Ar-H), 3.76-3.68 (m, 2H, C-H), 2.94-2.91 (m, 2H, CH₂), 2.89-2.85 (m, 2H, CH₂), 2.64-2.60 (m, 2H, CH₂), 1.82 (t, 2H, *J*= 12.0, 24.0 Hz, CH₂), 1.18 (d, 6H, *J*= 4.0 Hz, (CH₃)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 154.2, 132.1, 130.0, 114.2, 73.5, 61.1, 59.1, 32.3. HRMS (ES+) *m/z* calculated for C₂₀H₂₆N₃O₄: 372.1925. Found [M+H]⁺: 372.1918 (Diff – 1.88 ppm)

6d: 5-(4-(2-(3,5-Dimethylmorpholino)ethyl)phenoxy)-2-nitroaniline (relative stereochemistry unknown)



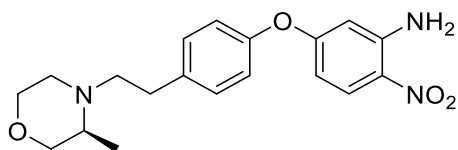
Reaction repeated as general procedure B using 4-(2-(3,5-dimethylmorpholino)ethyl)phenol (**4d**) gave the title compound (**6d**) as a yellow oil (86 mg, 72%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.32 (d, 1H, *J*= 12.0 Hz, Ar-H), 7.04 (d, 2H, *J*= 8.0 Hz, Ar-H), 6.73 (d, 2H, *J*= 8.0 Hz, Ar-H), 6.30-6.25 (m, 2H, Ar-H), 3.72-3.69 (m, 2H, C-H), 3.45-3.41 (m, 2H, CH₂), 3.26-3.19 (m, 2H, CH₂), 2.99-2.93 (m, 2H, CH₂), 2.80-2.75 (m, 2H, CH₂), 1.03 (t, 6H, *J*= 4.0, 12.0 Hz, (CH₃)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 165.3, 156.6, 145.9, 135.5, 127.8, 126.2, 123.1, 122.6, 106.4, 104.1, 74.6, 60.6, 53.0, 33.4, 12.3, 12.2. HRMS (ES+) *m/z* calculated for C₂₀H₂₆N₃O₄: 372.1925. Found [M+H]⁺: 372.1918 (Diff -1.88 ppm).

6e: 5-(4-(2-(*cis*-2,6-Dimethylmorpholino)ethyl)phenoxy)-2-nitroaniline



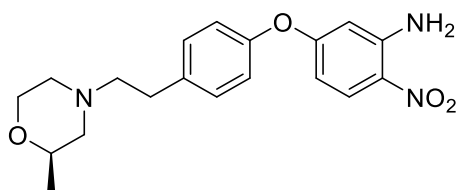
Reaction repeated as general procedure B using 4-(2-(*cis*-2,6-dimethylmorpholino)ethyl)phenol (**4e**) gave the title compound (**6e**) as a yellow solid (0.22 g, 82%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.07 (d, 1H, *J*= 12.0 Hz, Ar-H), 7.22 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.05 (d, 2H, *J*= 8.0 Hz, Ar-H), 6.31-6.27 (m, 2H, Ar-H), 3.76-3.69 (m, 2H, C-H), 2.94-2.91 (m, 2H, CH₂), 2.89-2.85 (m, 2H, CH₂), 2.64-2.60 (m, 2H, CH₂), 1.82 (t, 2H, *J*= 12.0, 24.0 Hz, CH₂), 1.18 (d, 6H, *J*= 8.0 Hz, CH₃). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 164.4, 157.0, 144.1, 137.3, 130.0, 129.3, 127.9, 120.6, 106.3, 102.8, 71.3, 60.2, 59.0, 31.4, 18.0. HRMS (ES+) *m/z* calculated for C₂₀H₂₆N₃O₄: 372.1923. Found [M+H]⁺: 372.1926 (Diff 0.27 ppm).

6f: (S)-5-(4-(2-(3-Methylmorpholino)ethyl)phenoxy)-2-nitroaniline



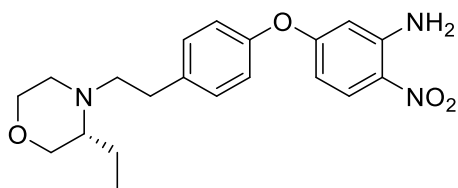
Reaction repeated as general procedure B using (*S*)-4-(2-(3-methylmorpholino)ethyl)phenol (**4f**) gave the title compound (**6f**) as a yellow oil (0.51 g, 72%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.60 (d, 1H, *J*= 12.0 Hz, Ar-H), 7.32 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.05 (d, 2H, *J*= 8.0 Hz, Ar-H), 6.31-6.26 (m, 2H, Ar-H), 3.86-3.82 (m, 1H, CH₂), 3.73-3.63 (m, 2H, CH₂), 3.29-3.24 (m, 1H, CH₂), 2.99-2.92 (m, 2H, CH₂), 2.81-2.74 (m, 2H, CH₂), 2.63-2.48 (m, 3H, CH₂, C-H), 1.05 (s, 3H, CH₃). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 164.0, 156.2, 141.2, 137.1, 126.2, 125.0, 122.9, 120.9, 106.3, 104.1, 72.3, 66.8, 62.7, 55.2, 54.7, 31.4, 12.8. HRMS (ES+) *m/z* calculated for C₁₉H₂₄N₃O₄: 358.1767. Found [M+H]⁺: 358.1770 (Diff 0.84 ppm).

6g: (R)-5-(4-(2-(2-Methylmorpholino)ethyl)phenoxy)-2-nitroaniline



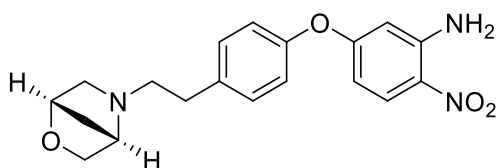
Reaction repeated as general procedure B using (*R*)-4-(2-(2-methylmorpholino)ethyl)phenol (**4g**) gave the title compound (**6g**) as a yellow oil (0.55 g, 87%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.05 (d, 1H, *J*= 12.0 Hz, Ar-H), 7.32 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.04 (d, 2H, *J*= 8.0 Hz, Ar-H), 6.31-6.25 (m, 2H, Ar-H), 3.89-3.83 (m, 1H, C-H), 3.72-3.62 (m, 2H, CH₂), 2.93-2.90 (m, 2H, CH₂), 2.85-2.83 (m, 2H, CH₂), 2.63-2.59 (m, 2H, CH₂), 2.23-2.16 (m, 1H, CH₂), 1.88 (t, 1H, *J*= 12.0, 20.0 Hz, CH₂), 1.15 (s, 3H, CH₃). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 167.2, 155.9, 143.2, 138.7, 129.1, 126.4, 125.5, 120.7, 107.4, 104.9, 78.8, 67.1, 63.3, 62.9, 57.0, 35.4, 19.0. HRMS (ES+) *m/z* calculated for C₁₉H₂₄N₃O₄: 358.1767. Found [M+H]⁺: 358.1771 (Diff 1.12 ppm).

6h: (R)-5-(4-(2-(3-Ethylmorpholino)ethyl)phenoxy)-2-nitroaniline



Reaction repeated as general procedure B using (*R*)-4-(2-(3-ethylmorpholino)ethyl)phenol (**4h**) gave the title compound (**6h**) as a yellow solid (34.1 mg, 67%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.81 (m, 1H, Ar-H), 7.45 (d, 2H, *J*=8.0 Hz, Ar-H), 7.12 (d, 2H, *J*= 8.0 Hz, Ar-H), 6.88 (d, 1H, *J*= 12.0 Hz, Ar-H), 6.60 (m, 1H, Ar-H), 3.74-3.69 (m, 1H, CH₂), 3.68-3.63 (m, 1H, CH₂), 3.52-3.49 (m, 1H, CH₂), 3.45-3.40 (m, 1H, CH₂), 2.79-2.71 (m, 2H, CH₂), 2.68-2.62 (m, 1H, CH₂), 2.57-2.53 (m, 1H, CH₂), 2.44-2.41 (m, 2H, CH₂), 2.33-2.30 (m, 1H, C-H), 1.55-1.51 (m, 2H, CH₂), 0.81 (t, 3H, *J*= 4.0, 12.0 Hz, CH₃). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 166.4, 156.4, 148.6, 138.1, 130.7, 127.0, 125.5, 120.4, 110.3, 105.7, 70.8, 70.5, 65.0, 58.9, 53.5, 37.6, 22.2, 12.1. HRMS (ES+) *m/z* calculated for C₂₀H₂₆N₃O₄: 372.1925. Found [M+H]⁺: 372.1928 (Diff 0.81 ppm).

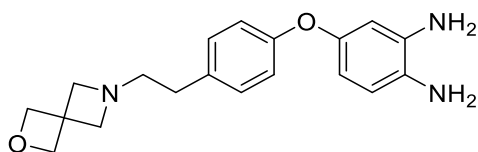
6i: 5-(4-(2-((1S,4S)-2-Oxa-5-azabicyclo[2.2.1]heptan-5-yl)ethyl)phenoxy)-2-nitroaniline



Reaction repeated as general procedure B using 4-(2-((1S,4S)-2-oxa-5-azabicyclo[2.2.1]heptan-5-yl)ethyl)phenol (**4i**) gave the title compound (**6i**) as a yellow solid. ^1H NMR (400 MHz, $\text{CD}_3\text{OD}-d_4$) δ 7.46 (d, 1H, $J = 12.0$ Hz, Ar-H), 7.21 (d, 2H, $J = 8.0$ Hz, Ar-H), 6.93 (d, 2H, $J = 8.0$ Hz, Ar-H), 6.32-6.28 (m, 2H, Ar-H), 3.94-3.91 (m, 1H, C-H), 3.72-3.69 (m, 2H, CH_2), 3.22-3.12 (m, 2H, CH_2), 2.82-2.78 (m, 2H, CH_2), 2.66-2.57 (m, 2H, CH_2), 2.55-2.51 (m, 1H, C-H), 1.98-1.88 (m, 2H, CH_2).

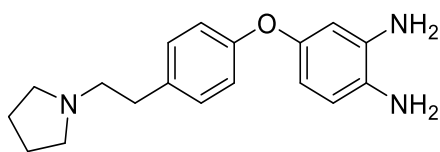
2.14.2.1.3 4-(4-(2-(Amino)phenoxy)benzene-1,2-diamine Derivatives

7a: 4-(4-(2-(2-Oxa-6-azaspiro[3.3]heptan-6-yl)ethyl)phenoxy)benzene-1,2-diamine



Reaction repeated as general procedure C using 5-(4-(2-(2-oxa-6-azaspiro[3.3]heptan-6-yl)ethyl)phenoxy)-2-nitroaniline (**6a**) gave the title compound (**7a**) as a brown oil (71.5 mg, 65%). ^1H NMR (400 MHz, $\text{CD}_3\text{OD}-d_4$) δ 7.14 (d, 2H, $J = 8.0$ Hz, Ar-H), 6.84 (d, 2H, $J = 8.0$ Hz, Ar-H), 6.70 (d, 1H, $J = 8.0$ Hz, Ar-H), 6.42 (d, 1H, $J = 4.0$ Hz, Ar-H), 6.23 (dd, 1H, $J = 4.0, 8.0$ Hz, Ar-H), 3.84 (s, 4H, CH_2), 3.38 (s, 4H, CH_2), 2.87 (t, 2H, $J = 4.0, 12.0$ Hz, CH_2), 2.66 (t, 2H, $J = 8.0, 12.0$ Hz, CH_2). ^{13}C NMR (101 MHz, $\text{CD}_3\text{OD}-d_4$) δ 157.7, 150.1, 136.7, 132.0, 130.2, 139.6, 129.4, 117.7, 109.6, 107.6, 80.2, 62.6, 58.1, 41.0, 32.0. HRMS (ES+) m/z calculated for $\text{C}_{19}\text{H}_{24}\text{N}_3\text{O}_2$: 325.1790. Found $[\text{M}+\text{H}]^+$: 325.1793 (Diff 0.92 ppm).

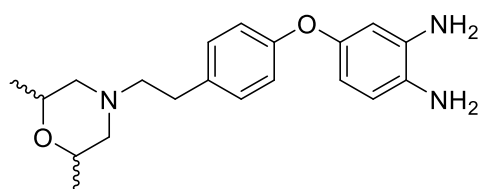
7b: 4-(4-(2-(Pyrrolidin-1-yl)ethyl)phenoxy)benzene-1,2-diamine



Reaction repeated as general procedure C using 2-nitro-5-(4-(2-(pyrrolidin-1-yl)ethyl)phenoxy)aniline (**6b**) gave the title compound (**7b**) as a crude brown oil. This intermediate was not characterised and carried through to the subsequent reaction.

7c: 4-(4-(2-(2,6-Dimethylmorpholino)ethyl)phenoxy)benzene-1,2-diamine

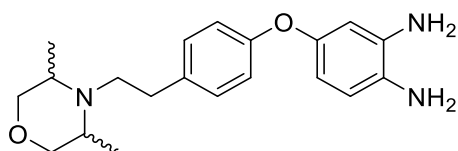
(relative stereochemistry unknown)



Reaction repeated as general procedure C using 5-(4-(2-(2,6-dimethylmorpholino)ethyl)phenoxy)-2-nitroaniline (**6c**) gave the title compound (**7c**) as a brown oil (0.67 g, >100%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.43 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.37 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.06 (d, 2H, *J* = 8.0 Hz, Ar-H), 6.54-6.50 (m, 2H, Ar-H), 4.02-3.95 (m, 2H, C-H), 3.67-3.64 (m, 2H, CH₂), 3.47-3.41 (m, 2H, CH₂), 3.19-3.15 (m, 2H, CH₂), 2.83 (t, 2H, *J* = 12.0, 24.0 Hz, CH₂), 1.28 (d, 6H, *J* = 8.0 Hz, (CH₃)₂).

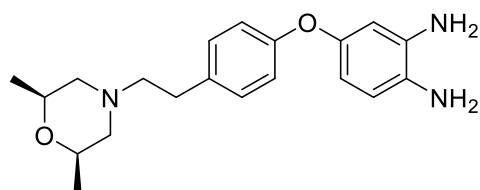
7d: 4-(4-(2-(3,5-Dimethylmorpholino)ethyl)phenoxy)benzene-1,2-diamine

(relative stereochemistry unknown)



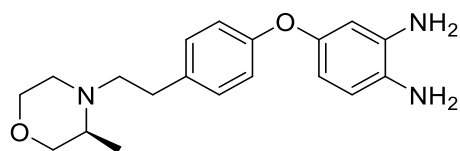
Reaction repeated as general procedure C using 5-(4-(2-(3,5-dimethylmorpholino)ethyl)phenoxy)-2-nitroaniline (**6d**) gave the title compound (**7d**) as a brown oil (53.6 mg, 68 %). This intermediate was not characterised and carried through to the subsequent reaction.

7e: 4-(4-(2-(*cis*-2,6-Dimethylmorpholino)ethyl)phenoxy)benzene-1,2-diamine



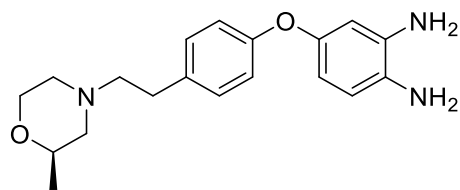
Reaction repeated as general procedure C using 5-(4-(2-(*cis*-2,6-dimethylmorpholino)ethyl)phenoxy)-2-nitroaniline (**6e**) gave the title compound (**7e**) as a brown oil (0.13 g, 65%). This intermediate was not characterised and carried through to the subsequent reaction.

7f: (*S*)-4-(4-(2-(3-Methylmorpholino)ethyl)phenoxy)benzene-1,2-diamine



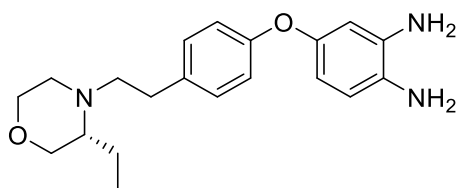
Reaction repeated as general procedure C using (*S*)-5-(4-(2-(3-methylmorpholino)ethyl)phenoxy)-2-nitroaniline (**6f**) gave the title compound (**7f**) as a dark brown/ red oil (0.19 g, 40%). This intermediate was not characterised and carried through to the subsequent reaction.

7g: (*R*)-4-(4-(2-(2-Methylmorpholino)ethyl)phenoxy)benzene-1,2-diamine



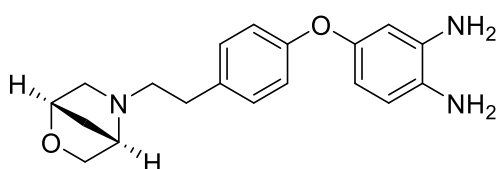
Reaction repeated as general procedure C using (*R*)-5-(4-(2-(2-methylmorpholino)ethyl)phenoxy)-2-nitroaniline (**6g**) gave the title compound (**7g**) as a brown oil (0.24 g, 48%). This intermediate was not characterised and carried through to the subsequent reaction.

7h: (R)-4-(4-(2-(3-Ethylmorpholino)ethyl)phenoxy)benzene-1,2-diamine



Reaction repeated as general procedure C using (R)-5-(4-(2-(3-ethylmorpholino)ethyl)phenoxy)-2-nitroaniline (**6h**) gave the title compound (**7h**) as a crude brown oil (17.3 mg, 55%). This intermediate was not characterised and carried through to the subsequent reaction.

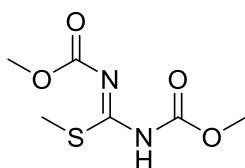
7i: 4-(4-(2-((1S,4S)-2-Oxa-5-azabicyclo[2.2.1]heptan-5-yl)ethyl)phenoxy)benzene-1,2-diamine



Reaction repeated as general procedure C using 5-(4-(2-((1S,4S)-2-oxa-5-azabicyclo[2.2.1]heptan-5-yl)ethyl)phenoxy)-2-nitroaniline (**6i**) gave the title compound (**7i**) as a crude brown oil (11.3 mg, 45%). This intermediate was not characterised and carried through to the subsequent reaction.

2.14.2.1.4 Methylisothiurea Derivatives

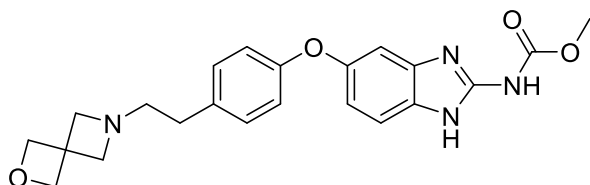
8: 1,3-Bis(Methoxycarbonyl)-2-methylisothiurea²³⁷



Reaction repeated as general procedure D using 2-methyl-2-thiopseudourea hemisulfate salt gave the title compound (**8**) as a white solid (1.18 g, 80%). This intermediate was not characterised and carried through to the subsequent reaction.

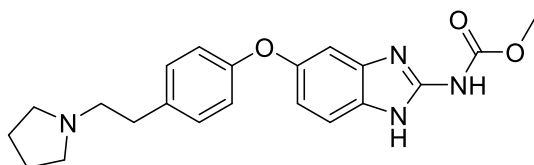
2.14.2.1.5 Methyl (5-(4-(2-(2-(Amino)-1H-benzo[d]imidazol-2-yl)carbamate Derivatives

9a: Methyl (5-(4-(2-(2-oxa-6-azaspiro[3.3]heptan-6-yl)ethyl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



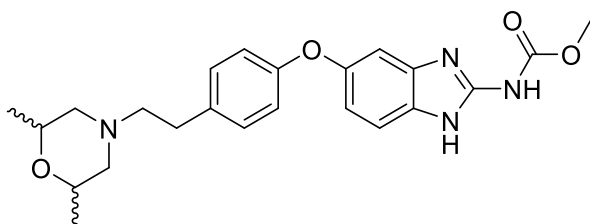
Reaction repeated as general procedure E using 4-(4-(2-(2-oxa-6-azaspiro[3.3]heptan-6-yl)ethyl)phenoxy)benzene-1,2-diamine (**7a**) gave the title compound (**9a**) as a pale-yellow solid (47.8 mg, 53%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.18 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.13 (d, 2H, *J* = 8.0 Hz, Ar-H), 6.87 (d, 1H, *J* = 4.0 Hz, Ar-H), 6.81 (d, 2H, *J* = 8.0 Hz, Ar-H), 6.58 (dd, 1H, *J* = 4.0, 8.0 Hz, Ar-H), 3.82 (s, 4H, CH₂), 3.37 (s, 4H, CH₂), 3.58 (s, 3H, CH₃), 2.88 (t, 2H, *J* = 4.0, 12.0 Hz, CH₂), 2.57 (t, 2H, *J* = 8.0, 12.0 Hz, CH₂). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 155.6, 152.4, 149.8, 147.0, 136.2, 134.1, 128.7, 128.4, 122.3, 116.2, 106.9, 104.1, 80.3, 64.1, 62.0, 54.4, 32.0. HRMS (ES⁺) *m/z* calculated for C₂₂H₂₅N₄O₄: 409.1798. Found [M+H]⁺: 409.1800 (Diff 0.49 ppm). IR *v*_{max}/cm⁻¹: (solid) 3330 (s), 2951 (m), 1626 (s), 1473 (m), 1230 (s). MP: 202-204°C. Purity HPLC 94.5%, *R*_t = 5.7 min.

9b: Methyl (5-(4-(2-(pyrrolidin-1-yl)ethyl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



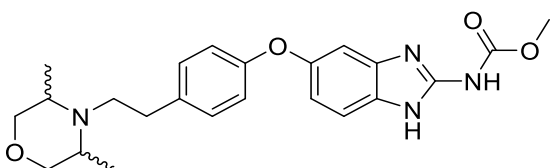
Reaction repeated as general procedure E using 4-(4-(2-(pyrrolidin-1-yl)ethyl)phenoxy)benzene-1,2-diamine (**7b**) gave the title compound (**9b**) as a beige solid (60 mg, 67%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.38 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.18 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.05 (d, 1H, *J* = 4.0 Hz, Ar-H), 6.88 (d, 2H, *J* = 8.0 Hz, Ar-H), 6.84 (dd, 1H, *J* = 4.0, 8.0 Hz, Ar-H), 3.84 (s, 3H, CH₃), 2.84-2.72 (m, 4H, CH₂), 2.68 (m, 4H, CH₂), 1.85 (m, 4H, CH₂). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 157.2, 154.5, 153.2, 146.8, 144.0, 137.6, 134.0, 129.4, 117.6, 113.7, 106.5, 103.2, 58.0, 53.6, 52.0, 33.7, 22.8. HRMS (ES⁺) *m/z* calculated for C₂₁H₂₅N₄O₃: 381.1926. Found [M+H]⁺: 381.1930 (Diff 1.05 ppm). IR *v*_{max}/cm⁻¹: (solid) 3328 (s), 2948 (m), 1626 (s), 1472 (m), 1242 (s). MP: 198-201°C decomposed. Purity HPLC 94.3%, *R*_t = 5.5 min.

9c: Methyl (5-(4-(2-(2,6-dimethylmorpholino)ethyl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate (relative stereochemistry unknown)



Reaction repeated as general procedure E using 4-(4-(2-(2,6-dimethylmorpholino)ethyl)phenoxy)benzene-1,2-diamine (**7c**) gave the title compound (**9c**) as a white solid (0.26 g, 31%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.23 (d, 2H, *J* = 12.0 Hz, Ar-H), 7.01 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.95 (d, 2H, *J* = 8.0 Hz, Ar-H), 6.56 (d, 1H, *J* = 4.0 Hz, Ar-H), 6.53 (dd, 1H, *J* = 4.0, 8.0 Hz, Ar-H), 3.80 (s, 3H, CH₃), 3.74-3.69 (m, 2H, C-H), 2.90-2.89 (m, 2H, CH₂), 2.84-2.80 (m, 2H, CH₂), 2.61-2.57 (m, 2H, CH₂), 1.80 (t, 2H, *J* = 12.0, 24.0 Hz, CH₂), 1.17 (d, 6H, *J* = 8.0 Hz, (CH₃)₂). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 155.2, 152.5, 150.7, 146.5, 139.2, 134.7, 130.6, 129.0, 123.3, 114.6, 110.4, 104.3, 71.6, 60.2, 59.8, 54.6, 31.3, 18.0. HRMS (ES+) *m/z* calculated for C₂₃H₂₉N₄O₄: 425.2189. Found [M+H]⁺: 425.2186 (Diff – 0.71 ppm). IR ν_{max}/cm⁻¹: (solid) 3336 (s), 2971 (m), 2922 (w), 1634 (s), 1488 (m), 1427 (m), 1258 (s). MP: 206-209°C. Purity HPLC 92.6%, R_t = 5.5 min.

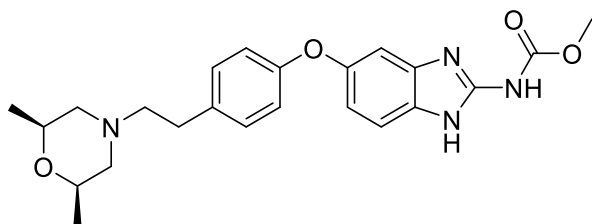
9d: Methyl (5-(4-(2-(3,5-dimethylmorpholino)ethyl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate (relative stereochemistry unknown)



Reaction repeated as general procedure E using 4-(4-(2-(3,5-dimethylmorpholino)ethyl)phenoxy)benzene-1,2-diamine (**7d**) gave the title compound (**9d**) as a light brown solid (14.7 mg, 22%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.52 (m, 1H, Ar-H), 7.42 (m, 1H, Ar-H), 7.35 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.16 (m, 1H, Ar-H), 7.03 (d, 2H, *J* = 8.0 Hz, Ar-H), 3.76-3.53 (m, 7H, CH₃, (CH₂)₂), 2.68-2.66 (m, 2H, CH₂), 2.64-2.61 (m, 2H, CH₂), 2.49-2.46 (m, 2H, C-H), 1.10 (d, 6H, *J* = 8.0 Hz, (CH₃)₂). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 155.4, 154.8, 152.6, 148.3, 140.4, 135.2, 130.2, 128.1, 120.3, 120.2, 115.4, 107.8, 104.8, 74.5, 60.4, 55.7, 51.8,

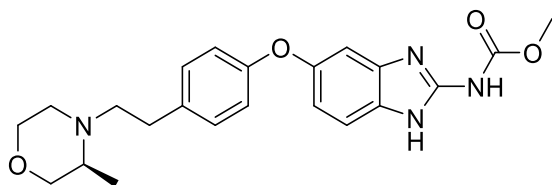
33.3, 14.9. HRMS (ES+) m/z calculated for $C_{23}H_{29}N_4O_4$: 425.2189. Found $[M+H]^+$: 425.2188 (Diff -0.24 ppm). IR ν_{max}/cm^{-1} : (solid) 3339 (s), 2983 (m), 2894 (w), 1641 (s), 1471 (m), 1412 (m), 1263 (s). MP: 203-205°C. Purity HPLC 92.9%, R_t = 5.8 min.

9e: Methyl (5-(4-(2-(*cis*-2,6-dimethylmorpholino)ethyl)phenoxy)-1*H*-benzo[*d*]imidazol-2-yl)carbamate



Reaction repeated as general procedure E using 4-(4-(2-(*cis*-2,6-dimethylmorpholino)ethyl)phenoxy)benzene-1,2-diamine (**7e**) gave the title compound (**9e**) as a yellow solid (42.5 mg, 27%). 1H NMR (400 MHz, DMSO-*d*6) δ 7.35 (d, 1H, J = 8.0 Hz, Ar-H), 7.18 (d, 2H, J = 8.0 Hz, Ar-H), 7.0 (d, 1H, J = 4.0 Hz, Ar-H), 6.85 (d, 2H, J = 8.0 Hz, Ar-H), 6.75 (dd, 1H, J = 4.0, 8.0 Hz, Ar-H), 3.70 (s, 3H, CH_3), 3.66-3.61 (m, 2H, C-H), 2.82-2.79 (m, 2H, CH_2), 2.73-2.68 (m, 2H, CH_2), 2.47-2.40 (m, 2H, CH_2), 1.67-1.61 (m, 2H, CH_2), 1.05 (d, 6H, J = 8.0 Hz, $(CH_3)_2$). ^{13}C NMR (101 MHz, DMSO-*d*6) δ 155.4, 154.7, 151.3, 147.0, 138.7, 135.5, 131.6, 128.6, 122.2, 116.1, 107.8, 105.1, 75.9, 63.4, 60.8, 55.5, 41.6, 22.6. HRMS (ES+) m/z calculated for $C_{23}H_{29}N_4O_4$: 425.2189. Found $[M+H]^+$: 425.2193 (Diff 0.94 ppm). IR ν_{max}/cm^{-1} : (solid) 3345 (s), 2981 (m), 2926 (w), 1629 (s), 1467 (m), 1411 (m), 1260 (s). MP: 204-207°C. Purity HPLC 92.4%, R_t = 5.7 min.

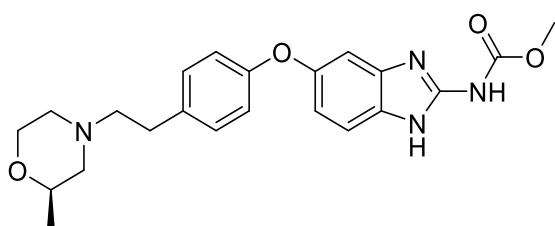
9f: Methyl (*S*)-(5-(4-(2-(3-methylmorpholino)ethyl)phenoxy)-1*H*-benzo[*d*]imidazol-2-yl)carbamate



Reaction repeated as general procedure E using (*S*)-4-(4-(2-(3-methylmorpholino)ethyl)phenoxy)benzene-1,2-diamine (**7f**) gave the title compound (**9f**) as a beige solid (0.09 g, 38%). 1H NMR (400 MHz, DMSO-*d*6) δ 7.69 (m, 1H, Ar-H), 7.45 (m, 1H, Ar-H), 7.38 (d, 2H, J = 8.0 Hz, Ar-H), 7.22 (m, 1H, Ar-H), 7.18 (d, 2H, J = 8.0 Hz, Ar-H), 3.88-3.85

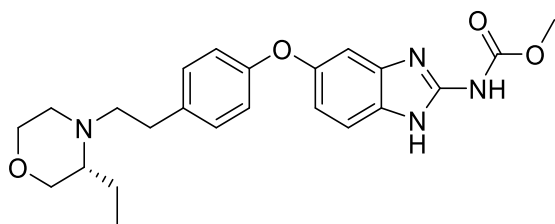
(m, 2H, CH₂), 3.78-3.76 (m, 2H, CH₂), 3.71 (s, 3H, CH₃), 2.97-2.91 (m, 2H, CH₂), 2.81-2.75 (m, 2H, CH₂), 2.66-2.43 (m, 3H, CH₂, C-H), 1.09 (s, 3H, CH₃). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 156.0, 154.3, 150.6, 148.2, 138.7, 130.1, 130.0, 128.4, 120.9, 116.6, 106.4, 105.8, 73.4, 69.2, 56.1, 55.5, 50.8, 34.7, 14.5. HRMS (ES+) *m/z* calculated for C₂₂H₂₇N₄O₄: 411.2032. Found [M+H]⁺: 411.2030 (Diff -0.49 ppm). IR ν_{max}/cm⁻¹: (solid) 3341 (s), 2988 (m), 2935 (w), 1640 (s), 1474 (m), 1423 (m), 1263 (s). MP: 200-204°C decomposed. Purity HPLC 80.19%, R_t= 5.7 min.

9g: Methyl (R)-(5-(4-(2-(2-methylmorpholino)ethyl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



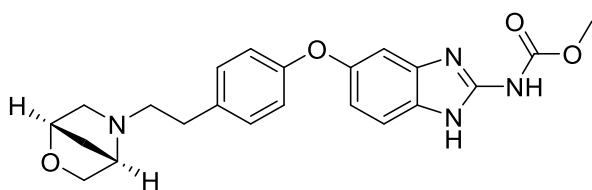
Reaction repeated as general procedure E using (R)-4-(4-(2-(2-methylmorpholino)ethyl)phenoxy)benzene-1,2-diamine (**7g**) gave the title compound (**9g**) as a pale-yellow solid (0.13 g, 43%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.68 (d, 1H, *J*= 12.0 Hz, Ar-H), 7.34 (m, 1H, Ar-H), 7.28 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.06 (m, 1H, Ar-H), 6.91-6.83 (m, 2H, Ar-H), 3.90-3.86 (m, 1H, C-H), 3.72-3.64 (m, 5H, CH₃, CH₂), 2.93-2.91 (m, 2H, CH₂), 2.87-2.84 (m, 2H, CH₂), 2.66-2.62 (m, 2H, CH₂), 2.18-2.15 (m, 1H, CH₂), 1.79 (t, 1H, *J*= 12.0, 24.0 Hz, CH₂), 1.13 (s, 3H, CH₃). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 155.6, 155.1, 148.2, 136.7, 132.8, 125.7, 125.4, 123.5, 108.2, 104.2, 77.8, 68.5, 63.9, 62.2, 59.1, 35.8, 19.2. HRMS (ES+) *m/z* calculated for C₂₂H₂₇N₄O₄: 411.2032. Found [M+H]⁺: 411.2031 (Diff - 0.24 ppm). IR ν_{max}/cm⁻¹: (solid) 3346 (s), 2976 (m), 2920 (w), 1636 (s), 1473 (m), 1465 (m), 1260 (s). MP: 202-206°C decomposed. Purity HPLC 86.1%, R_t= 5.8 min.

9h: Methyl (R)-(5-(4-(2-(3-ethylmorpholino)ethyl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



Reaction repeated as general procedure E using (R)-4-(4-(2-(3-ethylmorpholino)ethyl)phenoxy)benzene-1,2-diamine (**7h**) gave the title compound (**9h**) as a light-yellow solid (8.4 mg, 39%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.67 (m, 1H, Ar-H), 7.41 (m, 1H, Ar-H), 7.31 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.12-7.09 (m, 3H, Ar-H), 3.72 (s, 3H, CH₃), 3.68-3.62 (m, 1H, CH₂), 3.61-3.58 (m, 1H, CH₂), 3.54-3.50 (m, 1H, CH₂), 3.48-3.42 (m, 1H, CH₂), 2.75-2.70 (m, 2H, CH₂), 2.68-2.63 (m, 1H, CH₂), 2.61-2.57 (m, 1H, CH₂), 2.54-2.50 (m, 2H, CH₂), 2.29-2.26 (m, 1H, C-H), 1.55-1.52 (m, 2H, CH₂), 0.86 (t, 3H, *J* = 4.0, 12.0 Hz, CH₃). HRMS (ES+) *m/z* calculated for C₂₃H₂₉N₄O₄: 425.2191. Found [M+H]⁺: 425.2195 (Diff -0.94 ppm). IR ν_{max}/cm⁻¹: (solid) 3338 (s), 2982 (m), 2918 (w), 1644 (s), 1470 (m), 1461 (m), 1251(s). MP: 208-211°C decomposed. Purity HPLC 89.3%, R_t = 5.8 min.

9i: Methyl (5-(4-(2-((1S,4S)-2-oxa-5-azabicyclo[2.2.1]heptan-5-yl)ethyl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate

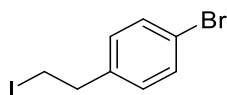


Reaction repeated as general procedure E using 4-(4-(2-((1S,4S)-2-oxa-5-azabicyclo[2.2.1]heptan-5-yl)ethyl)phenoxy)benzene-1,2-diamine (**7i**) gave the title compound (**9i**) as a light-yellow solid (4 mg, 29%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.56 (m, 1H, Ar-H), 7.38 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.32 (m, 1H, Ar-H), 7.16 (m, 1H, Ar-H), 7.06 (d, 2H, *J* = 8.0 Hz, Ar-H), 3.94-3.68 (m, 6H, CH₃, CH₂, C-H), 2.65-2.37 (m, 6H, (CH₂)₃), 2.21-2.18 (m, 1H, C-H), 1.86-1.83 (m, 2H, CH₂). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 155.3, 152.4, 149.9, 147.5, 138.4, 136.2, 130.1, 129.3, 120.7, 115.3, 107.6, 105.1, 83.0, 77.7, 68.4, 68.3, 55.9, 51.6, 42.1, 32.9, 32.8. HRMS (ES+) *m/z* calculated for C₂₂H₂₅N₄O₄: 409.1876. Found [M+H]⁺: 409.1872 (Diff -

0.98 ppm). IR $\nu_{\max}/\text{cm}^{-1}$: (solid) 3331 (s), 2983 (m), 1622 (s), 1486 (m), 1446 (m), 1261 (s). MP: 213-217°C decomposed. Purity HPLC 93.9%, R_t = 5.7 min.

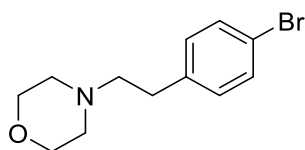
2.14.2.2 Benzimidazole Ketone Synthesis

13: 1-Bromo-4-(2-iodoethyl)benzene²⁰²

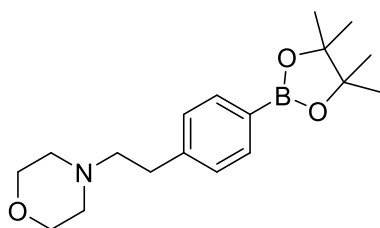


PPh_3 (0.39 g, 1.49 mmol, 1.5 equiv.), I_2 (0.38 g, 1.49 mmol, 1.5 equiv.), DMAP (49 mg, 0.40 mmol, 0.4 equiv.), and anhydrous DCM (10 mL) were stirred at room temperature for 5 minutes in a flame dried flask under an argon atmosphere. 2-(4-Bromophenyl)ethan-1-ol (**12**) (0.2 g, 0.99 mmol, 1 equiv.) was added and left to stir overnight at room temperature. After this time, saturated $\text{Na}_2\text{S}_2\text{O}_3$ aqueous solution was added to quench excess iodine and extracted with DCM (x3). The combined organic phase was washed with brine, dried over MgSO_4 , filtered, and concentrated under reduced pressure to yield a crude off white solid. Purification *via* column chromatography eluting with 100% *n*-hexane gave the title compound (**13**) as a white solid (0.26 g, 84%). ^1H NMR (400 MHz, CDCl_3) δ 7.44 (d, 2H, J = 8.0 Hz, Ar-H), 7.07 (d, 2H, J = 8.0 Hz, Ar-H), 3.23 (t, 2H, J = 8.0, 12.0 Hz, CH_2), 3.13 (t, 2H, J = 8.0, 12.0 Hz, CH_2). ^{13}C NMR (101 MHz, CDCl_3) δ 139.4, 131.8, 130.1, 120.8, 39.5, 5.0. HRMS (ES+) m/z calculated for $\text{C}_8\text{H}_9^{81}\text{Br}^{127}\text{I}$: 310.8932. Found $[\text{M}+\text{H}]^+$: 310.8936 (Diff 1.29 ppm).

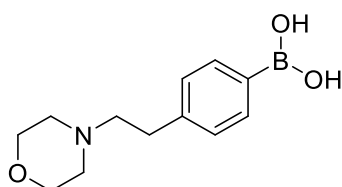
14: 4-(4-Bromophenethyl)morpholine²³⁷



Reaction repeated as general procedure A using 1-bromo-4-(2-iodoethyl)benzene (**13**) gave the title compound (**14**) as a brown oil (0.12 g, 52%). ^1H NMR (400 MHz, $\text{CD}_3\text{OD}-d_4$) δ 7.43 (d, 2H, J = 8.0 Hz, Ar-H), 7.22 (d, 2H, J = 8.0 Hz, Ar-H), 3.59-3.49 (m, 4H, $(\text{CH}_2)_2$), 3.30-3.26 (m, 2H, CH_2), 3.10-3.06 (m, 2H, CH_2), 1.34-1.30 (m, 4H, $(\text{CH}_2)_2$). ^{13}C NMR (101 MHz, $\text{CD}_3\text{OD}-d_4$) δ 139.5, 131.0, 130.4, 119.5, 69.4, 60.2, 55.4, 29.3. HRMS (ES+) m/z calculated for $\text{C}_{12}\text{H}_{17}^{81}\text{BrNO}$: 270.0493. Found $[\text{M}+\text{H}]^+$: 270.0490 (Diff -1.11 ppm).

15b: 4-(4-(4,4,5,5-Tetramethyl-1,3,2-dioxaborolan-2-yl)phenethyl)morpholine²³⁸

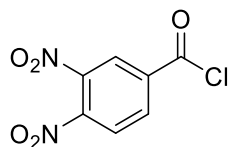
To a flame dried flask purged with nitrogen, 4-(4-bromophenethyl)morpholine (**14**) (0.34 g, 1.26 mmol, 1 equiv.), bis(pinacolato)diboron (0.42 g, 1.64 mmol, 1.3 equiv.), KOAc (0.37 g, 3.18 mmol, 3 equiv.), Pd(dppf)Cl₂ (46 mg, 0.06 mmol, 0.05 equiv.) and dioxane (15 mL) were added and allowed to stir overnight at reflux. After this time, the reaction was cooled and concentrated to yield a brown oil. Purification *via* column chromatography eluting with 40-70% EtOAc in *n*-hexane gave the title compound (**15b**) as a brown solid (0.26 g, 64%). ¹H NMR (400 MHz, CD₃OD-*d*₄) δ 7.69 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.32 (d, 2H, *J* = 8.0 Hz, Ar-H), 3.61-3.53 (m, 4H, (CH₂)₂), 3.30-3.24 (m, 2H, CH₂), 3.13-3.10 (m, 2H, CH₂), 1.33-1.30 (m, 4H, (CH₂)₂), 1.26 (s, 12H, (CH₃)₄). ¹³C NMR (101 MHz, CD₃OD-*d*₄) δ 139.6, 134.2, 130.1, 127.6, 89.4, 68.9, 60.1, 55.7, 35.7, 25.0. HRMS (ES+) *m/z* calculated for C₁₈H₂₉⁸¹BrNO₃: 318.2240. Found [M+H]⁺: 318.2246 (Diff 1.89 ppm).

15a: (4-(2-Morpholinoethyl)phenyl)boronic acid

To a flame dried flask purged with nitrogen 4-(4-bromophenethyl)morpholine (**14**) (0.26 g, 0.96 mmol, 1 equiv.) and anhydrous THF (3.2 mL, 0.3 M) were added and cooled to -78°C. *n*-BuLi (1.4 mL, 2.31 mmol, 2.4 equiv.) was added dropwise over 30 minutes and allowed to stir for 1 hour. B(O*i*Pr)₃ (0.27 mL, 1.44 mmol, 1.5 equiv.) in THF (0.8 mL, 1.2 M) was added. The reaction was allowed to stir at room temperature for 1 hour. To the flask, 2M HCl aqueous solution (3.8 mL, 0.25 M) was added and allowed to stir at room temperature for a further 1 hour. After this time, the reaction mixture was extracted with DCM (x3). The combined organic phase was dried over MgSO₄, filtered, and concentrated under reduced pressure to

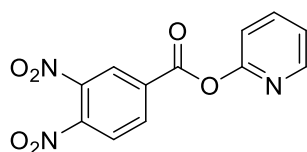
give the title compound (**15a**) as a crude brown oil (0.15 g, 66%). This intermediate was not characterised and carried through to the subsequent reaction.

17b: 3,4-Dinitrobenzoyl chloride²⁰⁹

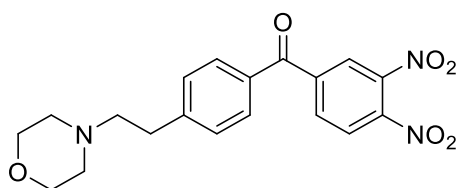


To a flask 3,4-dinitrobenzoic acid (**16**) (1 g, 4.71 mmol, 1 equiv.) and thionyl chloride (9.4 mL, 0.5 M) were added and allowed to stir at reflux for 6 hours. After this time, the reaction was cooled and concentrated under reduced pressure. The residue was co-evaporated with toluene (x3) to give the title compound (**17b**) as a yellow solid (0.87 g, 81%). This intermediate was not characterised and carried through to the subsequent reaction.

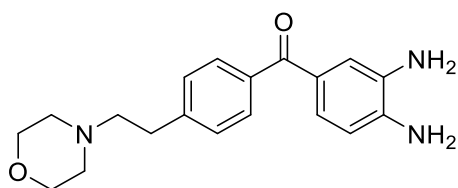
17a: Pyridin-2-yl 3,4-dinitrobenzoate²¹²



To a flask 3,4-dinitrobenzoic acid (**16**) (1 g, 4.71 mmol, 1 equiv.), DMAP (58 mg, 0.47 mmol, 0.1 equiv.) and DCM (23.6 mL, 0.2 M) were added and allowed to stir at room temperature for 15 minutes. Di(pyridin-2-yl) carbonate (1 g, 4.71 mmol, 1 equiv.) was added and left to stir at room temperature under a nitrogen atmosphere for 2 days. After this time, the reaction mixture was washed with, saturated aqueous NaHCO₃ solution, water, and brine. The organic phase was dried over MgSO₄, filtered, and concentrated under reduced pressure to give the title compound (**17a**) as a crude white solid (0.87 g, 64%). ¹H NMR (400 MHz, CDCl₃) δ 8.79 (s, 1H, Ar-H), 8.60 (d, 1H, *J*= 8.0 Hz, Ar-H), 8.49 (d, 1H, *J*= 8.0 Hz, Ar-H), 8.05 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.98 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.92 (t, 1H, *J*= 8.0, 16.0 Hz, Ar-H), 7.37 (t, 1H, *J*= 4.0, 8.0 Hz, Ar-H). ¹³C NMR (101 MHz, CDCl₃) δ 161.1, 157.3, 148.9, 140.1, 135.2, 134.5, 127.0, 126.4, 125.5, 125.3, 123.1, 116.1. HRMS: (Cl⁺) *m/z* calculated for C₁₄H₁₁N₂O₆: 303.0612. Found [M+H]⁺: 303.0623 (Diff 3.63 ppm).

18: (3,4-Dinitrophenyl)(4-(2-morpholinoethyl)phenyl)methanone²¹²

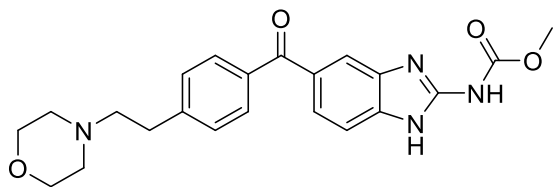
To a screw fix sealed tube pyridin-2-yl 3,4-dinitrobenzoate (**17a**) (0.2 g, 0.69 mmol, 1 equiv.), (4-(2-morpholinoethyl)phenyl)boronic acid (**15a**) (0.33 g, 1.38 mmol, 2 equiv.), Pd(OAc)₂ (4.6 mg, 0.02 mmol, 0.03 equiv.) PPh₃ (16 mg, 0.06 mmol, 0.09 equiv.) and dioxane (1.4 mL, 0.5 M) were added and allowed to stir under microwave conditions at 110°C for 30 minutes. After this time, the reaction mixture was diluted in EtOAc, washed with saturated aqueous NaHCO₃ solution, water, and brine. The organic phase was dried over MgSO₄, filtered, and concentrated under reduced pressure to yield a crude brown oil. Purification *via* column chromatography eluting with 80% EtOAc in *n*-hexane gave the title compound (**18**) as a brown oil (84 mg, 31%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.79 (m, 1H, Ar-H), 8.60 (d, 1H, *J*= 4.0 Hz, Ar-H), 8.51 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.25(d, 2H, *J*= 8.0 Hz, Ar-H), 7.72 (d, 2H, *J*= 8.0 Hz, Ar-H), 3.73 (t, 4H, *J*= 4.0, 8.0 Hz, 2(CH₂)), 2.88-2.84 (m, 2H, CH₂), 2.76-2.72 (m, 2H, CH₂), 2.58-2.55 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 191.7, 148.3, 146.4, 141.3, 138.4, 132.2, 131.7, 131.6, 128.6, 128.5, 127.9, 62.1, 60.3, 59.1, 35.8. HRMS: (ES+) *m/z* calculated for C₁₉H₂₀N₃O₆: 386.1348. Found [M+H]⁺: 386.1356 (Diff 2.07 ppm).

19: (3,4-Diaminophenyl)(4-(2-morpholinoethyl)phenyl)methanone¹⁹²

Reaction repeated as general procedure C using (3,4-dinitrophenyl)(4-(2-morpholinoethyl)phenyl)methanone (**18**) and double equivalents of SnCl₂·H₂O gave the title compound (**19**) as a brown oil (21.1 mg, 29 %). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.56 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.47 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.12 (d, 1H, *J*= 8.0 Hz, Ar-H), 6.93 (m, 1H, Ar-H), 6.60 (d, 1H, *J*= 8.0 Hz, Ar-H), 3.82 (t, 4H, *J*= 4.0, 12.0 Hz, 2(CH₂)), 2.75-2.71 (m, 1H, CH₂), 2.65-2.61 (m, 2H, CH₂), 2.56-2.53 (m, 1H, CH₂), 2.51-2.49 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 201.7, 147.3, 145.5, 138.3, 135.3, 132.9, 130.4, 129.0, 121.6, 118.5, 115.3, 66.1,

60.8, 53.4, 31.4. HRMS: (ES+) m/z calculated for $C_{19}H_{24}N_3O_2$: 326.1865. Found $[M+H]^+$: 326.1861 (Diff -1.23 ppm).

20: Methyl (5-(4-(2-morpholinoethyl)benzoyl)-1H-benzo[d]imidazol-2-yl)carbamate¹⁹²

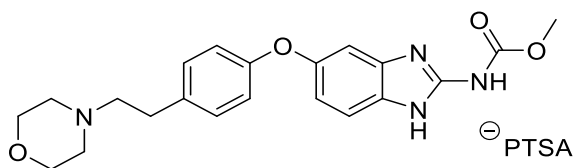


Reaction repeated as general procedure E using (3,4-diaminophenyl)(4-(2-morpholinoethyl)phenyl)methanone (**19**) gave the title compound (**20**) as a white solid (12 mg, 42%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.83 (s, 1H, Ar-H), 7.64 (d, 1H, J = 8.0 Hz, Ar-H), 7.52 (d, 2H, J = 8.0 Hz, Ar-H), 7.50 (d, 1H, J = 8.0 Hz, Ar-H), 7.43 (d, 2H, J = 8.0 Hz, Ar-H), 3.86 (t, 4H, J = 4.0, (CH₂)₂), 3.76 (s, 3H, CH₃), 2.65-2.61 (m, 2H, CH₂), 2.58-2.54 (m, 2H, CH₂), 2.52-2.48 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 199.8, 156.8, 147.2, 142.1, 139.1, 136.8, 133.2, 130.1, 128.7, 122.0, 120.1, 115.7, 66.3, 60.4, 54.4, 51.8, 31.1. HRMS: (ES+) m/z calculated for $C_{22}H_{25}N_4O_4$: 409.1876. Found $[M+H]^+$: 409.1881 (Diff 1.22 ppm). IR ν_{max}/cm^{-1} : (solid) 3334 (s), 2975 (m), 1732 (s), 1622 (s), 1467 (m), 1260 (s). MP: 274-278°C decomposed. Purity HPLC 100%, R_t = 6.0 min.

2.14.2.3 Salt Synthesis

21a: Methyl (5-(4-(2-morpholinoethyl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate

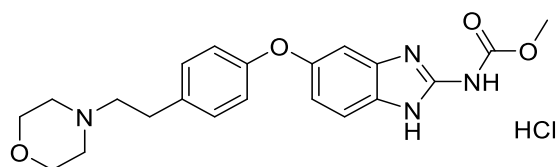
PTSA Salt



To a beaker methyl (5-(4-(2-morpholinoethyl)benzyl)-1H-benzo[d]imidazol-2-yl)carbamate (**1h**) (10 mg, 0.03 mmol, 1 equiv.) and 1M PTSA in MeOH were added, heated until the solid dissolved, and then left at room temperature for 2 days. The crystals formed were harvested by filtering with a sinter funnel and washed with cold MeOH. The filtrate was concentrated under reduced pressure to yield a crude white solid. Excess PTSA was removed *via* trituration with DCM to give the title compound (**21a**) as a white solid (10 mg, 67%). ¹H NMR (400 MHz, CD₃OD-*d*₄) δ 7.72 (d, 2H, J = 12.0 Hz, Ar-H), 7.62 (d, 1H, J = 12.0 Hz, Ar-H), 7.33 (d, 2H, J = 8.0

Hz, Ar-H), 7.26-7.23 (m, 3H, Ar-H), 7.14 (dd, 1H, $J = 4.0, 8.0$ Hz, Ar-H), 7.02 (d, 2H, $J = 12.0$ Hz, Ar-H), 4.10-4.06 (m, 2H, CH₂), 3.97 (s, 3H, CH₃), 3.85-3.78 (m, 2H, CH₂), 3.61-3.57 (d, 2H, $J = 16.0$ Hz, CH₂), 3.42-3.37 (m, 2H, CH₂), 3.25-3.18 (m, 2H, CH₂), 3.12-3.08 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*₄) δ 155.7, 154.2, 150.3, 148.2, 143.6, 138.9, 138.3, 136.9, 130.8, 130.7, 127.6, 123.2, 116.8, 108.4, 106.2, 67.7, 61.1, 54.8, 53.1, 26.8, 33.1. IR $\nu_{\text{max}}/\text{cm}^{-1}$: (solid) 3297 (s), 2975 (m), 1622 (s), 1502 (m), 1477 (m), 1377 (s) 1235 (s). Starting material (**1h**) IR $\nu_{\text{max}}/\text{cm}^{-1}$: (solid) 3328 (s), 2947 (m), 1626 (s), 1472 (m), 1225 (s). MP: 216-220°C decomposed. Purity HPLC 92.7%, $R_t = 5.4$ min.

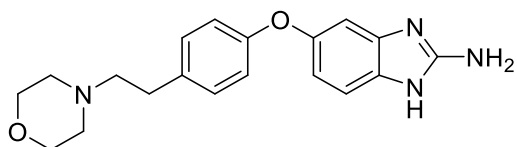
21b: Methyl (5-(4-(2-morpholinoethyl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate HCl Salt



To a beaker methyl (5-(4-(2-morpholinoethyl)benzyl)-1H-benzo[d]imidazol-2-yl)carbamate (**1h**) (70 mg, 0.18 mmol, 1 equiv.) and 1M HCl solution in MeOH (0.18 mL, 1 M) were added, heated gently until dissolved and left at room temperature overnight. The resultant crystals were filtered and washed with cold MeOH. The filtrate was concentrated under reduced pressure to yield a crude pale pink solid. The solid was purified by trituration with DCM to give the title compound (**21b**) as a pale pink solid (71 mg, 93%). ¹H NMR (400 MHz, CD₃OD-*d*₄) δ 7.40 (d, 1H, $J = 8.0$ Hz, Ar-H), 7.20 (d, 2H, $J = 8.0$ Hz, Ar-H), 7.07 (d, 1H, $J = 4.0$ Hz, Ar-H), 6.90 (d, 2H, $J = 8.0$ Hz, Ar-H), 6.86 (dd, 1H, $J = 4.0, 12.0$ Hz, Ar-H), 3.86 (s, 3H, CH₃), 3.75-3.73 (m, 2H, CH₂), 3.51-3.49 (m, 2H, CH₂), 3.16-3.14 (m, 2H, CH₂), 2.83-2.79 (m, 2H, CH₂), 2.63-2.60 (m, 2H, CH₂), 2.59-2.56 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*₄) δ 156.4, 155.7, 151.4, 150.3, 138.2, 136.5, 130.2, 123.6, 115.8, 106.2, 105.6, 68.9, 62.2, 56.3, 55.3, 34.7. IR $\nu_{\text{max}}/\text{cm}^{-1}$: (solid) 3186 (s), 2972 (m), 2631 (w), 1633 (s), 1466 (m), 1226 (s). Starting material (**1h**) IR $\nu_{\text{max}}/\text{cm}^{-1}$: (solid) 3328 (s), 2947 (m), 1626 (s), 1472 (m), 1225 (s). MP: 222-225°C decomposed. Purity HPLC 97.6%, $R_t = 5.3$ min.

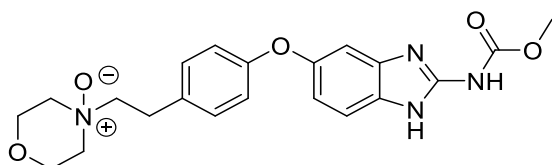
2.14.2.4 Benzimidazole Metabolite Synthesis

27a: 5-(4-(2-Morpholinoethyl)phenoxy)-1H-benzo[d]imidazol-2-amine²³⁰



Methyl (5-(4-(2-morpholinoethyl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate (**1h**) (50 mg, 0.126 mmol, 1 equiv.) and KOH (9.2 mg, 0.164 mmol, 1.3 equiv.) were dissolved in the mixture of MeOH (0.5 mL, 0.26 M) and distilled water (0.14 mL, 0.9 M) and allowed to stir at reflux for 3 days. After this time, the reaction mixture was cooled and concentrated under reduced pressure. The residue was diluted in distilled water and extracted with EtOAc (x3). The combined organic layers were washed with brine, dried over MgSO₄, filtered, and concentrated *in vacuo* to give the title compound (**27a**) as a pink solid (19 mg, 48 %). ¹H NMR (400 MHz, CD₃OD-*d*₄) δ 7.15-7.11 (m, 3H, Ar-H), 6.85-6.82 (m, 3H, Ar-H), 6.67-6.64 (m, 1H, Ar-H), 3.72-3.70 (m, 4H, (CH₂)₂), 2.79-2.75 (m, 2H, CH₂), 2.59-2.54 (m, 6H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*₄) δ 157.6, 151.3, 150.7, 136.2, 133.5, 131.2, 129.4, 117.2, 115.5, 110.9, 103.5, 66.2, 60.7, 53.3, 31.6. HRMS: (ES⁺) *m/z* calculated for C₁₉H₂₃N₄O₂: 339.1821. Found [M+H]⁺: 339.1828 (Diff 2.06 ppm). IR *v*_{max}/cm⁻¹: (solid) 3469 (m), 3339 (s), 2953 (m), 1634 (s), 1481 (m), 1271 (s). MP: 143-145°C. Purity HPLC 88.8%, R_t= 4.8 min.

27b: 4-(4-((2-((Methoxycarbonyl)amino)-1H-benzo[d]imidazol-5-yl)oxy)phenethyl)morpholine 4-oxide²³¹



Methyl (5-(4-(2-morpholinoethyl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate (**1h**) (0.1 g, 0.25 mmol, 1 equiv.) was suspended in chloroform (0.15 mL, 1M) and cooled at 0°C. *m*-CPBA (43.5 mg, 0.25 mmol, 1 equiv.) was added to the flask and allowed to stir at room temperature for 12 hours. After this time, the solvent was removed under reduced pressure to yield a crude white solid. Purification *via* column chromatography eluting with 0-20% MeOH in DCM gave the title compound (**27b**) as a white solid (69 mg, 69 %). ¹H NMR (400 MHz, CD₃OD-*d*₄) δ 7.42 (d, 1H, *J*= 12.0 Hz, Ar-H), 7.28 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.09 (d, 1H, *J*= 4.0 Hz, Ar-H), 6.93 (d, 2H, *J*= 8.0 Hz, Ar-H), 6.87 (dd, 1H, *J*= 4.0, 8.0 Hz, Ar-H), 4.25 (t, 2H, *J*= 16.0, 24.0 Hz, CH₂), 3.86 (s, 3H, CH₃), 3.85-3.82 (m, 2H, CH₂), 3.53-3.49 (m, 4H, (CH₂)₂), 3.22-3.14 (m, 4H, (CH₂)₂).

^{13}C NMR (101 MHz, $\text{CD}_3\text{OD}-d_4$) δ 157.7, 155.1, 152.2, 144.3, 136.2, 131.1, 130.0, 129.9, 117.7, 115.2, 108.2, 106.1, 71.8, 66.5, 63.6, 55.1, 52.0, 26.8. HRMS: (ES+) m/z calculated for $\text{C}_{21}\text{H}_{25}\text{N}_4\text{O}_5$: 413.1825. Found $[\text{M}+\text{H}]^+$: 413.1835 (Diff 2.42 ppm). IR $\nu_{\text{max}}/\text{cm}^{-1}$: (solid) 3329 (s), 2984 (m), 1631 (s), 1498 (s), 1477 (m), 1259 (s), 969 (s). MP: 215-217°C decomposed. Purity HPLC 96.9%, R_t = 5.5 min.

2.15 References

- 1 L. R. Stuchlíková, P. Matoušková, I. Vokřál, J. Lamka, B. Szotáková, A. Sečkařová, D. Dimunová, L. T. Nguyen, M. Várady and L. Skálová, *Int. J. Parasitol. Drugs Drug Resist.*, 2018, **8**, 50–58.

- 2 L. Urbizu, A. Confalonieri, S. Sánchez Bruni, C. Lanusse and L. I. Alvarez, *Chemotherapy*, 2012, **58**, 295–298.
- 3 M. P. Marques, O. M. Takayanagui and V. L. Lanchote, *Brazilian J. Med. Biol. Res.*, 2002, **35**, 261–269.
- 4 P. Giri, L. Gupta, S. Naidu, V. Joshi, N. Patel, S. Giri and N. R. Srinivas, *Drug Metab. Lett.*, 2018, **12**, 101–116.
- 5 Z. Wu, D. Lee, J. Joo, J. H. Shin, W. Kang, S. Oh, D. Y. Lee, S. J. Lee, S. S. Yea, H. S. Lee, T. Lee and K. H. Liu, *Antimicrob. Agents Chemother.*, 2013, **57**, 5448–5456.
- 6 Z. Zhang and W. Tang, *Acta Pharm. Sin. B*, 2018, **8**, 721–732.
- 7 S. Lepri, M. Ceccarelli, N. Milani, S. Tortorella, A. Cucco, A. Valeri, L. Goracci, A. Brink and G. Cruciani, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, E3178–E3187.
- 8 S. Ghanbarzadeh, A. Khalili, A. Jouyban, S. Emami, Y. Javadzadeh, M. Solhi and H. Hamishehkar, *Res. Pharm. Sci.*, 2016, **11**, 435.
- 9 J. Ma, J. Liang, J. Han, M. Zheng and H. Zhao, *J. Chem. Eng. Data*, 2019, **64**, 1237–1243.
- 10 L. Ceballos, L. Moreno, J. J. Torrado, C. Lanusse and L. Alvarez, *BMC Vet. Res.*, 2012, **8**, 1–10.
- 11 M. A. Ibrahim and F. K. Al-Anazi, *Saudi Pharm. J.*, 2013, **21**, 215–223.
- 12 J. A. Cisneros, M. J. Robertson, B. Q. Mercado and W. L. Jorgensen, *ACS Med. Chem. Lett.*, 2017, **8**, 124–127.
- 13 M. A. Walker, *Expert Opin. Drug Discov.*, 2014, **9**, 1421–1433.
- 14 K. T. Savjani, A. K. Gajjar and J. K. Savjani, *ISRN Pharm.*, 2012, **2012**, 1–10.
- 15 M. Rogers-Evans, H. Knust, J. M. Plancher, E. M. Carreira, G. Wuitschik, J. Burkhard, D. B. Li and C. Guérot, *Chimia (Aarau)*, 2014, **68**, 492–499.
- 16 W. Wu, C. You, C. Yin, Y. Liu, X. Q. Dong and X. Zhang, *Org. Lett.*, 2017, **19**, 2548–2551.
- 17 A. Zhai, *Determination of Phenols in Drinking Water with Agilent Bond Elut Plexa SPE and HPLC*, Agilent Technologies, 2012.
- 18 M. R. Crampton and I. A. Robotham, *J. Chem. Res.*, 1997, 22–23.
- 19 F. Maran, D. Celadon, M. G. Severin and E. Vianello, *J. Am. Chem. Soc.*, 1991, **113**, 9320–9329.
- 20 K. Roy and P. L. A. Popelier, *J. Phys. Org. Chem.*, 2009, **22**, 186–196.
- 21 J. D. Soper, E. Saganic, D. Weinberg, D. A. Hrovat, J. B. Benedict, W. Kaminsky and J. M. Mayer, *Inorg. Chem.*, 2004, **43**, 5804–5815.
- 22 C. A. Grice, K. L. Tays, B. M. Savall, J. Wei, C. R. Butler, F. U. Axe, S. D. Bembenek, A. M. Fourie, P. J. Dunford, K. Lundeen, F. Coles, X. Xue, J. P. Riley, K. N. Williams, L. Karlsson and J. P. Edwards, *J. Med. Chem.*, 2008, **51**, 4150–4169.
- 23 W. Wang, D. Kong, H. Cheng, L. Tan, Z. Zhang, X. Zhuang, H. Long, Y. Zhou, Y. Xu, X. Yang and K. Ding, *Bioorganic Med. Chem. Lett.*, 2014, **24**, 4250–4253.
- 24 J. March and M. B. Smith, in *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2001, pp.

- 1703–1869.
- 25 J. Lu, I. Maezawa, S. Weerasekara, R. Erenler, T. D. T. Nguyen, J. Nguyen, L. Z. Swisher, J. Li, L.-W. Jin, A. Ranjan, S. K. Srivastava and D. H. Hua, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 3392–3397.
- 26 W. Apptec, *ULVP-20190314*, 2019.
- 27 V. Kumar, S. B. Bharate, R. A. Vishwakarma and S. S. Bharate, *ACS Omega*, 2018, **3**, 8365–8377.
- 28 B. Combourieu, P. Besse, M. Sancelme, J. P. Godin, A. Monteil, H. Veschambre and A. M. Delort, *Appl. Environ. Microbiol.*, 2000, **66**, 3187–3193.
- 29 PCT Int. Appl., 2009087305, 2009.
- 30 EUCAST, European Committee on Antimicrobial Susceptibility Testing (EUCAST), [Http://www.eucast.org/documents/sops/](http://www.eucast.org/documents/sops/), (Accessed March 2020).
- 31 H. Nishikawa, Y. Fukuda, J. Mitsuyama, M. Tashiro, A. Tanaka, T. Takazono, T. Saijo, K. Yamamoto, S. Nakamura, Y. Imamura, T. Miyazaki, H. Takeya, Y. Yamamoto, K. Yanagihara, H. Mukae, S. Kohno and K. Izumikawa, *J. Antimicrob. Chemother.*, 2017, **72**, 1709–1713.
- 32 T. Shibata, T. Takahashi, E. Yamada, A. Kimura, H. Nishikawa, H. Hayakawa, N. Nomura and J. Mitsuyama, *Antimicrob. Agents Chemother.*, 2012, **56**, 5892–5897.
- 33 D. Das, J. M. H. Anal and L. Rokhum, *J. Chem. Sci.*, 2016, **128**, 1695–1701.
- 34 R. S. P. J Garegg, T. Regberg, J. Stawinski, *J. Chem. Soc. Perkin Trans. II*, 1987, 271–274.
- 35 R. L. Mackman, D. Siegel, J. E. Hartrum, PCT WO2016/018697A1, 2016, 79–81.
- 36 F. A. Cotton and P. A. Kibala, *J. Am. Chem. Soc.*, 1987, **109**, 3308–3312.
- 37 T. B. Lanni, K. L. Greene, C. N. Kolz, K. S. Para, M. Visnick, J. L. Mobley, D. T. Dudley, T. J. Baginski and M. B. Limmatta, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 756–760.
- 38 J. Buchspies and M. Szostak, *Catalysts*, 2019, **9**, 53.
- 39 A. J. J. Lennox and G. C. Lloyd-Jones, *Chem. Soc. Rev.*, 2014, **43**, 412–443.
- 40 D. S. Carter, R. T. Jacobs, Y. R. Freund, P. W. Berry, T. Akama, E. E. Easom, C. S. Lunde, F. Rock, R. Stefanakis, J. McKerrow, C. Fischer, C. A. Bulman, K. C. Lim, B. M. Suzuki, N. Tricoche, J. A. Sakanari, S. Lustigman and J. J. Plattner, *ACS Infect. Dis.*, 2020, **6**, 180–185.
- 41 J. Buchspies and M. Szostak, *Catalysts*, 2019, **9**, 53.
- 42 M. Haddach and J. R. McCarthy, *Tetrahedron Lett.*, 1999, **40**, 3109–3112.
- 43 H. Tatamidani, F. Kakiuchi and N. Chatani, *Org. Lett.*, 2004, **6**, 3597–3599.
- 44 S. S. Moleele, J. P. Michael and C. B. De Koning, *Tetrahedron*, 2006, **62**, 2831–2844.
- 45 J. Thongpaen, T. E. Schmid, L. Toupet, V. Dorcet, M. Mauduit and O. Baslé, *Chem. Commun.*, 2018, **54**, 8202–8205.
- 46 T. Zhu, G. He, J. Chang, D. Zhao, X. Zhu and H. Zhu, *Dye. Pigment.*, 2012, **95**, 679–688.
- 47 G. S. Paulekuhn, J. B. Dressman and C. Saal, *J. Med. Chem.*, 2007, **50**, 6665–6672.

- 48 K. Iwata, M. Karashima and Y. Ikeda, *Cryst. Growth Des.*, 2016, **16**, 4599–4606.
- 49 B. Sekula and A. Bujacz, *J. Med. Chem.*, 2016, **59**, 82–89.
- 50 Z. Yang, W. Wu, J. Wang, L. Liu, L. Li, J. Yang, G. Wang, D. Cao, R. Zhang, M. Tang, J. Wen, J. Zhu, W. Xiang, F. Wang, L. Ma, M. Xiang, J. You and L. Chen, *J. Med. Chem.*, 2014, **57**, 7977–7989.
- 51 D. Gupta, D. Bhatia, V. Dave, V. Sutariya and S. V. Gupta, *Molecules*, 2018, **23**, 1–15.
- 52 N. P. Aateka Patel, Stuart A Jones, Albert Ferro, *Br. J. Cardiol.*, 2009, **16**, 281–286.
- 53 D. L. Prohotsky and F. Zhao, *J. Pharm. Sci.*, 2012, **101**, 1–6.
- 54 G. Bolla and A. Nangia, *CrystEngComm*, 2018, **20**, 6394–6405.
- 55 J. Bolleddula, K. Dement, J. P. Driscoll, P. Worboys, P. J. Brassil and D. L. Bourdet, *Drug Metab. Rev.*, 2014, **46**, 379–419.
- 56 D. H. M. Lau, A. D. Lewis and B. I. Sikic, *JNCI J. Natl. Cancer Inst.*, 1989, **81**, 1034–1038.
- 57 D. J. St. Jean and C. Fotsch, *J. Med. Chem.*, 2012, **55**, 6002–6020.
- 58 M. Ulgen and N. Sevinc, *Curr. Drug Metab.*, 2017, **18**, 291–305.
- 59 R. Gao, L. Li, C. Xie, X. Diao, D. Zhong and X. Chen, *Drug Metab. Dispos.*, 2012, **40**, 556–567.
- 60 T. N. Kaipnazarov, K. B. Abdireimov, N. S. Mukhamedov, R. Y. Okmanov, B. Tashkhodjaev, G. E. Berdimbetova and K. M. Shakhidoyatov, *Russ. J. Org. Chem.*, 2013, **49**, 108–111.
- 61 PCT Int. Appl. 2004072101, 2004, 23–26.
- 62 A.-H. Khuthier and M. A. Sheat, *J. fur Prakt. Chemie*, 1989, **331**, 187–194.
- 63 L. A. Pinck and M. A. Kelly, *J. Am. Chem. Soc.*, 1925, **47**, 2170–2172.
- 64 A. Blaser, B. D. Palmer, H. S. Sutherland, I. Kmentova, S. G. Franzblau, B. Wan, Y. Wang, Z. Ma, A. M. Thompson and W. A. Denny, *J. Med. Chem.*, 2012, **55**, 312–326.
- 65 Swissbioisostere, [Http://www.swissbioisostere.ch/](http://www.swissbioisostere.ch/), (Accessed March 2020).
- 66 H. Alrabiah, A. A. Kadi, M. W. Attwa, A. S. Abdelhameed and G. A. E. Mostafa, *RSC Adv.*, 2019, **9**, 6409–6418.
- 67 A. Presser, G. Lainer, N. Kretschmer, W. Schuehly, R. Saf, M. Kaiser and M.-M. Kalt, *Molecules*, 2018, **23**, 2902.
- 68 C. Huanming, WO2017173999A1, 2017, 64.
- 69 Z. Yujun, WO2018188660A1, 2018, 171.

CHAPTER 3

Synthesis of Benzimidazole based Photoaffinity Labelled Probes.

3.1 Elucidation of Target Binding

Photoaffinity labelling (PAL) is a powerful tool for studying protein-protein interactions (PPIs). The approach is attractive as it enables us to visualise and monitor specific proteins without interfering with their biological function, allowing us to observe protein functions naturally. Traditionally proteins were studied *via* fluorescent labelling (FL) using tags such as ethidium bromide, fluorescein, and green fluorescent protein (GFP) (see **Figure 3.1**). The fluorophore

can selectively attach itself to the protein to aid visualisation. However, this method has its limitations as tags such as GFP are bulky with a weight of 300 kDa, which can disrupt the natural function of the protein.^{1, 2, 3} Fluorescent proteins also tend to form low affinity oligomers, limiting binding to the protein of interest.⁴ Although smaller tags have been developed for fluorescent labelling, PAL has become more commonly used over FL. Small PAL probes covalently bind to the protein providing valuable information about PPI without perturbing the proteins.¹

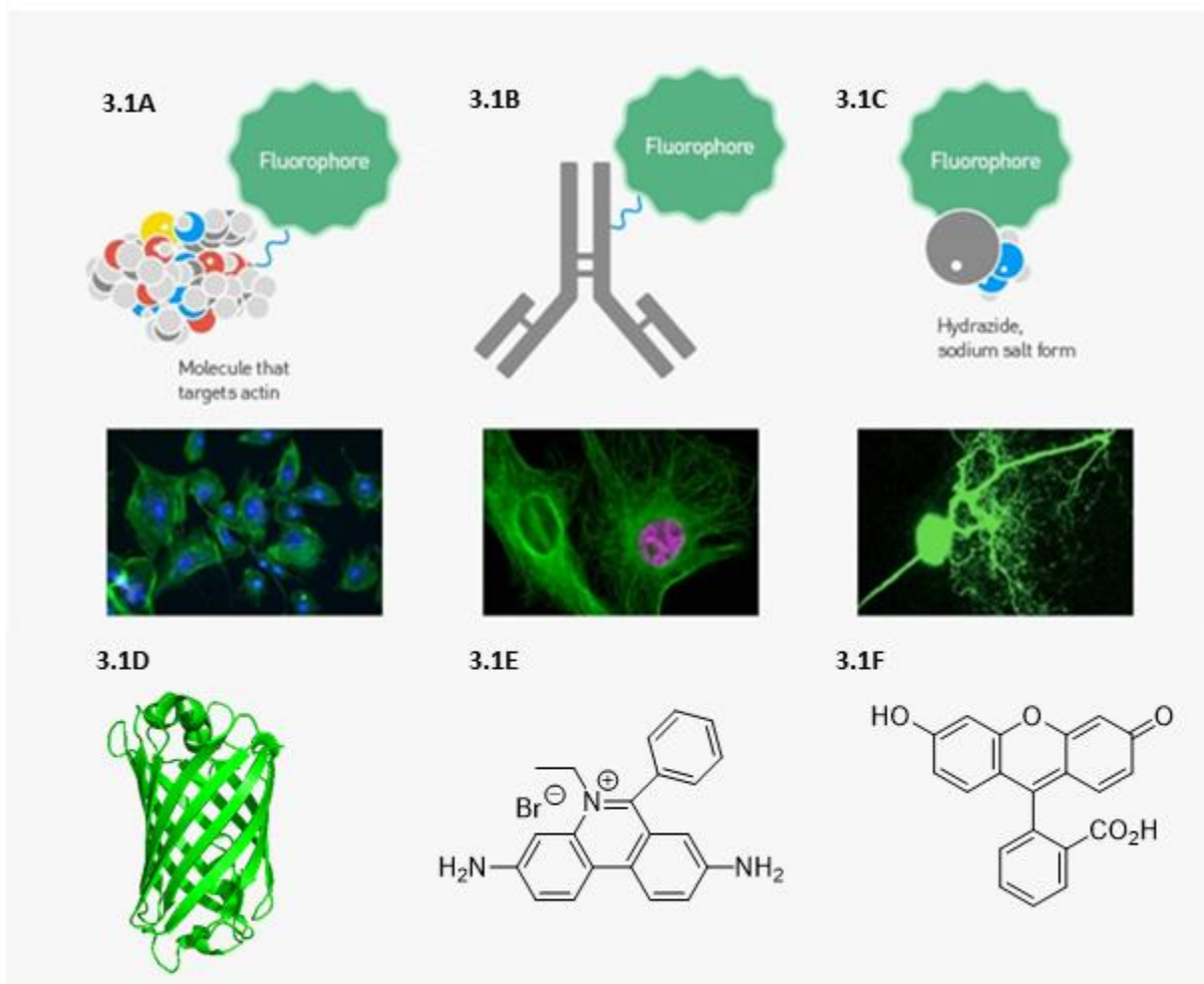


Figure 3.1: A fluorophore can perform several labelling tasks. **(3.1A):** Labelling proteins such as actin, **(3.1B):** Antibodies, **(3.1C):** Salt forms for whole cell staining.⁵ Common fluorophores include **(3.1D):** Green fluorescent protein (GFP),⁶ **(3.1E):** Ethidium bromide, **(3.1F):** Fluorescein. Adapted From: *Different Ways to Add Fluorescent Labels, Thermofischer, molecular probes school of fluorescence.*⁵

Extensive work has been carried out by the Nixon group to develop a novel antifungal agent to treat *C. neoformans* using benzimidazole as the basic template. Benzimidazoles such as flubendazole and albendazole display good activity against the fungus and are believed to target the β -tubulin monomer of microtubules. However, other targets such as the fungal topoisomerase I have been reported to be targets for benzimidazoles.⁷ Furthermore, benzimidazole carbamates have been reported to inhibit the enzyme fumarate reductase, an enzyme found in the mitochondria of many helminths which converts fumarate to succinate. This provides vital energy during anaerobic glucose metabolism.⁸ Albendazole can inhibit the enzyme up to 50% in *Ascaris suum*, a nematode that affects pigs.⁹ Thiabendazole inhibits the enzyme in *Haemonchus contortus*, which affects sheep and goats.¹⁰ Genomic studies have found that *C. neoformans* possesses fumarate reductase and is upregulated during stress. The presence of this enzyme may allow the benzimidazole to engage in promiscuous binding, resulting in an unselective treatment.^{11, 12}

Providing a selective treatment is vital as promiscuous binding can lead to off target effects resulting in adverse drug reactions. Analysing the binding profile of the drug will provide valuable knowledge to predict issues with toxicity and metabolic difficulties. It may be possible to modify the structure of the hit compound, improving the suitability, and making it an attractive candidate for further preclinical trials.

A PAL probe binding to the protein of interest can also reveal drug-protein interactions to improve the potency of the therapeutic agent, creating a more effective drug candidate.

3.1.1 Benzimidazole SAR Analysis

For this project, the probe will have the basic structure of the current lead candidate, with a photo affinity moiety and reporter tag attached (see **Figure 3.2**).

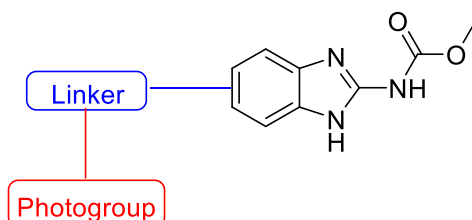


Figure 3.2: Basic scaffold of a benzimidazole PAL probe.

To develop a potent chemical probe, the current lead compounds for the project need to be studied. Extensive SAR work has been carried out by the Nixon group and gathered copious amounts of important information concerning the benzimidazole carbamate (see **Figure 3.3**).

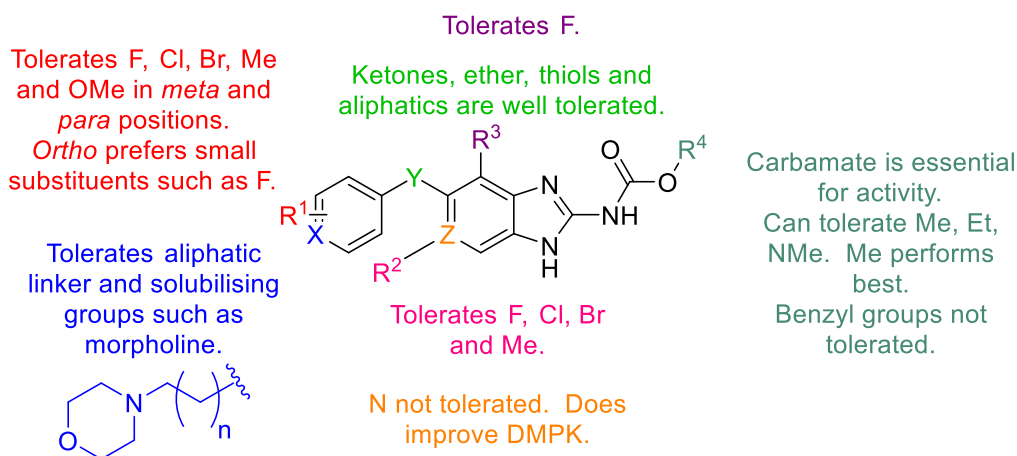


Figure 3.3: SAR information of benzimidazoles against *C. neoformans* gathered by the Nixon group.

Several trends have been established in **Figure 3.3**. Numerous substituents varying in size and electronegativity have been attached to the R¹ position. Bulkier groups such as bromine are better suited in the *meta* and *para* positions, as activity is lost when placed in the *ortho* position. The *ortho* position tends to prefer smaller substituents such as fluorine.

The structure can be expanded beyond the phenyl group at position X, as aliphatic chains can be added varying in length with solubilising groups such as morpholine being tolerated (see **Chapter 2, Table 2.3**). Nitrogen containing aliphatic heterocycles may improve the DMPK properties, however, not all are active such as piperidine (see **Chapter 2, Table 2.3**).

Ketones, ethers, thiols, and aliphatic linkers are well tolerated at position Y, with some displaying higher potency compared to the initial hit flubendazole and the current gold standard treatment (see **Chapter 2, Table 2.3**). Unfortunately, amide linkages provide no activity against *C. neoformans*.

Substituents on the phenyl ring of the benzimidazole core in position R² can tolerate fluorine, chlorine, bromine, and methyl groups. Only small substituents such as fluorine are tolerated at the R³ position. The potency tends to be lower and is therefore considered to be an unhelpful modification.

Pyridine rings on the benzimidazole core (position Z) improve the DMPK properties of the compound, however, diminish antifungal activity.

The carbamate is essential for antifungal activity. Additionally, the carbamate can be altered. Ethoxy and urea derivatives maintain antifungal activity, however methoxy groups performs best out of all the other substituents. Bulkier groups such as benzyls were not tolerated.

3.1.2 Design of Benzimidazole PAL Probe

Applying the SAR information, a PAL probe can be developed. A promising hit at the beginning of the project was the biaryl ether derivative **1** (see **Figure 3.4**).

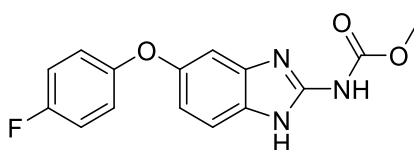
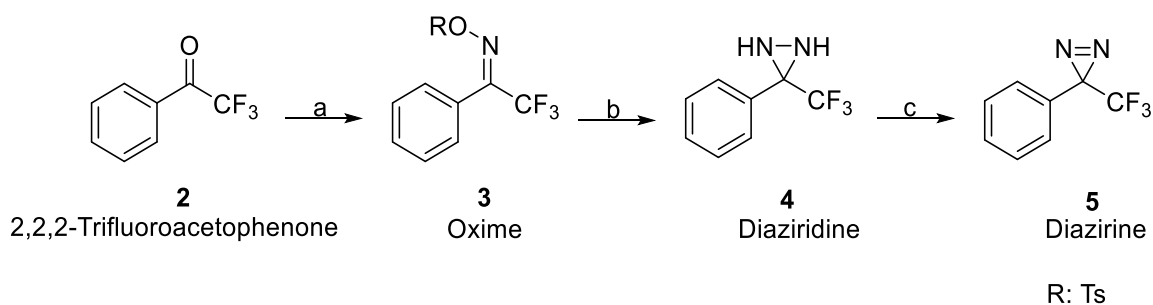


Figure 3.4: A biaryl ether analogue (**1**) was the Nixon groups lead compound in 2016.

The probe was designed around the structure of **1**. The type of photoreactive group and reporter tag that would be incorporated into structure **1** had to be decided.

There are several photoreactive groups available, such as phenyl azide, benzophenones, and diazirines. All these groups have their advantages and disadvantages (see **Section 1.2.2**). However, diazirines remain a popular choice in chemical biology due to their thermal stability, generation of the highly reactive carbene intermediate, and the ability to undergo a number of bond insertions such as C-N, C-O, C-H, and C-C.^{13, 14}

There are many different variants of the phenyl diazirine, however the 3-trifluoro-3-phenyl diazirine is the most effective photoreactive group in PAL due to its relatively small size, thermal and chemical stability, low rate of rearrangements, and high reactivity of the intermediate.^{15, 16, 17, 18, 19, 20} The most common way to synthesise a 3-trifluoro-3-phenyl diazirine is by converting a 2,2,2-trifluoroacetophenone (**2**) to an oxime (**3**) *via* oximation with hydroxylamine. The oxime undergoes tosylation with tosyl chloride.²¹ The tosyl oxime is converted to a diaziridine (**4**) *via* the addition of ammonia,²² followed by oxidation to the desired diazirine (**5**) (see **Scheme 3.1**).^{23, 24}



Scheme 3.1: Synthesis of a diazirine from the 2,2,2-trifluoroacetophenone starting material.

Reagent and Conditions: (a) $\text{NH}_2\text{OH}\cdot\text{HCl}$, pyridine, EtOH, 60-85°C, then *p*-TsCl, NEt_3 , rt; (b) NH_3 , 80°C; (c) Ag_2O or I_2 , NEt_3 , rt.²⁰

As mentioned in **Chapter 1.2.2.3**, reporter tags are required for characterisation of the protein-ligand complex. Traditionally tags such as biotin were used, however poor cellular penetration was experienced due to the large size of the tag. Copper catalysed azide-alkyne cycloaddition (CuAAC) was able to solve this issue with the attachment of an alkyne or azide handle on the probe. The alkyne is usually preferred as they are smaller and tend to perform better in CuAAC experiments, as azides can provide lower background signals from the proteomic workflow techniques such as SDS-PAGE.^{25, 26, 27} Although, it should be noted that azide handles have their own advantages as they possess high chemical stability under physiological conditions,^{28, 29} and can undergo a number of organic transformations that alkynes struggle with, such as Cu(I) catalysed and strain promoted [3 + 2] cycloaddition and Staudinger ligation.^{30, 31, 32}

The PAL probe needs to have the benzimidazole core and a carbamate for antifungal activity. Due to the favourable properties and high reactivity of the photo-induced intermediate, the 3-trifluoro-3-phenyldiazirine was selected as the photoreactive group. An alkyne handle would be incorporated to assist in the CuAAC reaction. From careful consideration of the SAR data, two initial probes (**50a** and **50b**) were designed based on the structure of the then hit compound (see **Figure 3.5**).

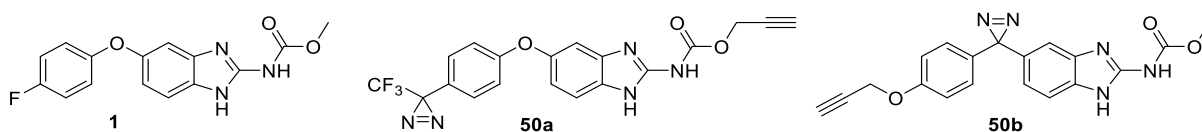


Figure 3.5: Structure of hit compound **1** and the potential PAL probes **50a**, **50b**

Compound **50a** keeps the ether linkage while having the 3-trifluoro-3-phenyldiazirine functionality. The alkyne handle is located on the carbamate handle. Large groups such as benzyl have not performed well in *in vitro* studies. Therefore, it is not known whether the 2-propyn-1-oxy group will be tolerated.

The other proposed structure (**50b**) has a number of structural modifications. The ether linkage has been replaced by a diazine functionality, which should bind in a similar way to a ketone, such as flubendazole. The 2-propyn-1-oxy group is attached to the phenyl ring, leaving the methyl carbamate intact. Based on the initial SAR data gathered, this should be better tolerated. The alkyne handle is shown as an ether linkage, however an aliphatic chain as well as other functionalities can be used.

Due to the amount of literature, chemical stability, and the ease of synthesis of the 3-trifluoro-3-phenyl diazine compound **50a** was initially synthesised.

3.2 Aims

The aim of this project is to design and synthesise a PAL probe based on the benzimidazole structure.

The individual aims for this project include...

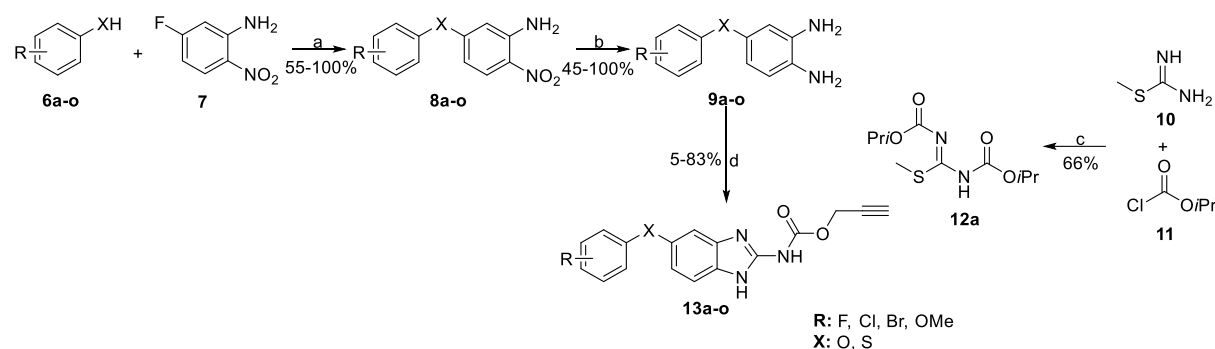
1. Synthesise a probe and gain *in vitro* MIC data to determine whether it possesses antifungal activity.
2. Submit active PAL probe for protein-ligand investigation to gain binding information.
3. Further proteolysis studies can be carried out to identify the exact residues the probe has labelled. Followed by homology modelling to characterise the structural features of the binding site on the β -tubulin.
4. Design and synthesise an inactive probe for comparison studies between the two probes, analysing their binding and covalent attachment to amino acid residues.

3.3 Discussion

3.3.1 Synthesis of Benzimidazole Carbamate Alkyne Probes

3.3.1.1 Synthesis of Prop-2-yn-1-yl (5-phenoxy/thio-1*H*-benzo[*d*]imidazol-2-yl)carbamate Analogues

Tolerance of a 2-propyn-1-oxy group at the carbamate was not known as groups larger than ethoxy possessed no antifungal activity. Initial synthesis involved the synthesis of 2-propyn-1-oxy carbamates to determine tolerance and to gain additional SAR data (see **Scheme 3.2**).



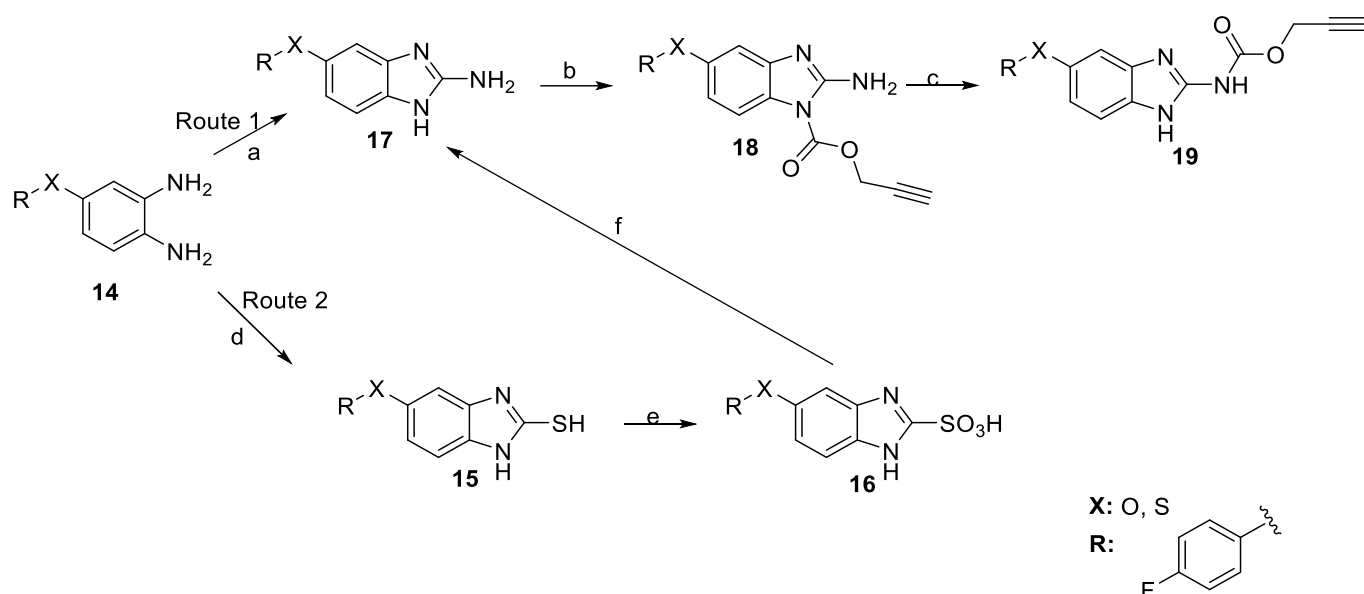
Scheme 3.2: Reagent and Conditions: (a) K_2CO_3 , DMF, reflux, 16 hr (b) $SnCl_2 \cdot H_2O$, EtOH, reflux, 16 hr; (c) H_2O $0^\circ C$, 20 min, then NaOH, rt; (d) AcOH, MeOH, reflux, 16 hr.

The ether and thiol synthesis involves the same S_NAr conditions described previously (see **Chapter 2, Section 2.3.1**).³³ Although the initial synthesis of compound **8a** differed, as the phenol (**6a**) was deprotonated with sodium hydride rather than potassium carbonate. Later, the synthesis changed by using potassium carbonate as the base. This resulted in a much higher yield of 88-100% compared to 56% with NaH. The various ether intermediates were characterised by 1H and ^{13}C NMR as well as mass spectrometry. The 1H NMR shows the presence of additional aromatic peaks as well as the amine peak from the nitroaniline between 8.0-6.0 ppm depending on the intermediate.

The reduction of the nitro group to form a diamine has been mentioned previously (see **Chapter 2, Section 2.3.1**).³⁴ The yields for these reactions vary greatly between 45-100%, although most analogues have yields in the 60% region. The best performing diamine was **9e** (*ortho*-F, ether), and the poorest was **9i** (*ortho*-OMe, thiol). The varying yields could be the result of a difference in solubility, reducing the amount of material able to react and be consumed. The compounds were characterised by 1H and ^{13}C NMR, and mass spectrometry. In the 1H NMR, the aromatic peaks from the diamine shift upfield to 6.0-5.0 ppm from 8.0-6.0

ppm. The diamine peaks can also be observed as a broad singlet to 4.9-3.4 ppm from 6.0-7.5 ppm with an integration of 4.

Before the synthesis of the 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea intermediate (**12a**) was discovered, the ring closure reaction to form the benzimidazole core proved challenging (see **Scheme 3.3**).



Scheme 3.3: Reagent and conditions: (a) BrCN, MeOH, H₂O, 50°C, 1 hr, (b) propargyl chloroformate, pyridine, 0°C to rt, 3 hr, (c) pyridine, 60°C, 4 hr (d) EtOH, H₂O, CS₂, reflux, 3 hr, (e) KMnO₄, NaOH, H₂O, reflux, 30 min, (f) NH₄OH, 145°C, 15 hr.

Scheme 3.3 branches into two routes to synthesise the key 1H-benzo[d]imidazol-2-amine (**17**) intermediate. Both routes merge upon the formation of compound **17** for the final two steps, involving the incorporation of the propargyl carbamate.

Starting from the diamine, route 1 ring closes **14** via cyanogen bromide following a procedure by *Frei et al.*³⁵ This synthesis afforded low yields of 35% compared to the literatures 91%. Difficulty was found when purifying the crude material via column chromatography due to the polar nature brought about by the additional primary amine on **17**. Route 2 provides an alternative synthesis as **14** is treated with carbonyl disulfide to form the benzo[d]imidazole-2-thiol (**15**). Literature states that the thiol can be oxidised to sulfonic acid by potassium permanganate (KMnO₄), followed by conversion to a primary amine by ammonium chloride to give **17**.³⁶ This reaction was initially trialled on 1H-benzo[d]imidazole-2-thiol, however the

procedure was unable to convert the thiol to the acid, as only starting material and degradation products were recovered. Therefore route 2 was abandoned.

In step B the alkyl chloroformate should be able to be incorporated into **17** to form **18**. Reports have claimed that the propargyl chloroformate is introduced to the amine in position 1, due to higher reactivity compared to the primary amine in position 2 (see **Figure 3.6**). The reports suggest that the propargyl carbamate can migrate to the primary amine (position 2) when **18** is refluxed in pyridine.³⁷

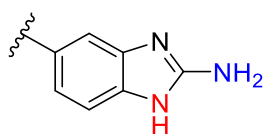


Figure 3.6: Positioning of amines on 1*H*-benzo[*d*]imidazol-2-amine. Position 1 shown in red.
Position 2 shown in blue.

It has been suggested that the amino group in position 1 is more reactive than the amino group in the position 2, due to delocalisation of the nitrogen lone pairs in the imidazole ring (see **Figure 3.7**). This causes the position 1 amine to be more readily available for nucleophilic attack of propargyl chloroformate. Additionally, the proton in position 1 is more acidic, and therefore more susceptible to deprotonation by pyridine.^{38, 39}

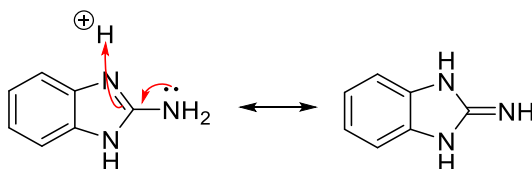


Figure 3.7: The lone pair of electrons from 1*H*-benzo[*d*]imidazol-2-amine's primary amine can delocalise.

Unfortunately, only starting material was present in the reaction mixture after **18** was refluxed with pyridine. Therefore, pyridine was unable to migrate the carbamate group to position 2. It was proposed that the amine in position 1 could be protected, prior to addition of the propargyl chloroformate. Unfortunately, attempts to synthesise a protected amine in the position 1 failed and only yielded starting materials.

Due to the unsuccessful attempts to attach the propargyl carbamate to the benzimidazole, the route was abandoned. Employing route 1 to synthesise **17**, would still result in the final

benzimidazole alkyne probe (**19**) not being synthesised, due to the failure of the migration step.

Another attempt to synthesise alkyne probes, involved an alternative key intermediate, termed 1,3-bis(2-propyn-1-oxycarbonyl)-2-methylisothiurea (**12a**) (see **Figure 3.8**).

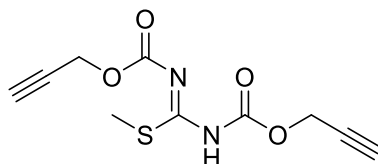
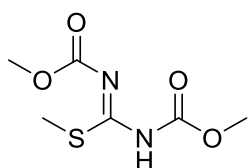


Figure 3.8: Structure of 1,3-bis(2-propyn-1-oxycarbonyl)-2-methylisothiurea (**12a**).

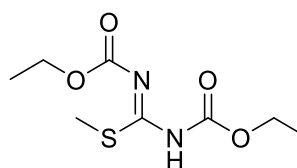
There is limited literature precedence for the synthesis of compound **12a**, however experimental procedures for the preparation of 1,3-bis(methoxycarbonyl)-2-methylisothiurea (**12b**) and 1,3-bis(ethoxycarbonyl)-2-methylisothiurea (**12c**) (see **Figure 3.9**) are available and have been mentioned previously in **Chapter 2, Section 2.3.1**.^{40, 41}

3.9A



12b

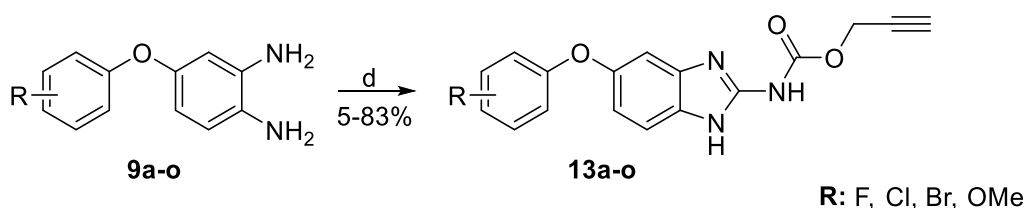
3.9B



12c

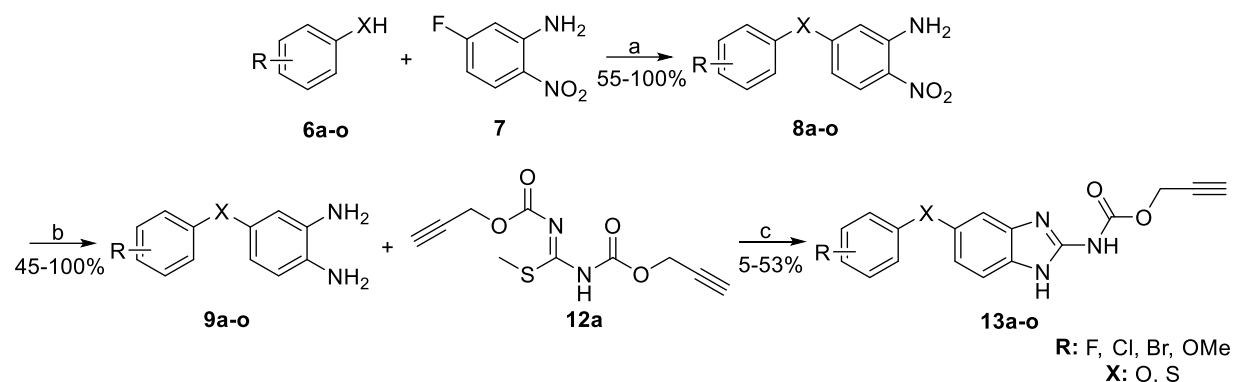
Figure 3.9: Structures of 1,3-bis(alkoxycarbonyl)-2-methylisothiurea derivatives. **3.9A:** 1,3-bis(methoxycarbonyl)-2-methylisothiurea (**12b**). **3.9B:** 1,3-bis(ethoxycarbonyl)-2-methylisothiurea (**12c**)

Two sources of literature have provided an alternative way to utilise this intermediate for the formation of the final benzimidazole carbamate structure (see **Scheme 3.4**).^{42, 43}



Scheme 3.4: Reagent and conditions: (a) **12a**, AcOH, MeOH, reflux, 16 hr.

Applying a similar method for the synthesis of methoxy and ethoxy thioureas, the preparation of **12a** was successful, illustrating that the reaction conditions were transferable.⁴¹ The intermediate was obtained as a pungent smelling white solid with modest yield of 50%. Once **12a** was successfully synthesised, **Scheme 3.5** was proposed.



Scheme 3.5: Reagent and conditions: (a) K_2CO_3 , DMF, reflux, 16 hr, (b) $SnCl_2 \cdot H_2O$, EtOH, reflux, 16 hr, (c) AcOH, MeOH, 65°C, 16 hr.

Steps A and B of **Scheme 3.5** have already been discussed previously. Step C involves **9a-o** and 1,3-bis(2-propyn-1-oxycarbonyl)-2-methylisothiourea (**12a**) being stirred in a combination of AcOH and MeOH in a screw fix sealed tube to form **9a-o**. Modest yields ranging between 38-83% were obtained depending on the substituted phenol or thiophenol used.³⁴ The yields vary considerably due to the scale of the reaction. It has been observed that the more **9a-o** used (particularly over 100 mg), the greater the yield of **13a-o** will be.

3.3.1.2 Results of Prop-2-yn-1-yl (5-phenoxy/thio-1H-benzo[d]imidazol-2-yl)carbamate Analogues

3.3.1.2.1 MIC Results

Several alkyne probes have been synthesised and submitted for MIC testing (see **Table 3.1**). A number of these compounds were synthesised, not only to test whether the 2-propyn-1-oxo group could be tolerated at the carbamate, but also to gather more SAR data for the project.

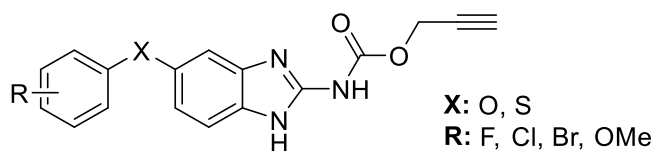


Figure 3.10 Basic structure of the alkyne probe

Compound	R	X	EUCAST (mg/L)
Flubendazole	<i>Para</i> -F	C=O	0.25
13a	<i>Para</i> -F	O	0.25
13b	<i>Para</i> -F	S	0.06
13c	<i>Meta</i> -F	O	0.125
13d	<i>Meta</i> -F	S	0.06
13e	<i>Ortho</i> -F	O	0.06
13f	<i>Ortho</i> -F	S	0.06
13g	<i>Ortho</i> -Br	S	0.125
13h	<i>Meta</i> -OMe	S	0.06
13i	<i>Ortho</i> -OMe	S	0.125
13j	<i>Para</i> -OMe	S	0.06
13k	<i>Para</i> -Cl	O	0.25
13l	<i>Para</i> -Cl	S	0.125
13m	<i>Meta</i> -Cl	S	0.06
13n	<i>Ortho</i> -Cl	O	0.125
13o	<i>Ortho</i> -Cl	S	0.06

Table 3.1: MIC results for the Propargyl benzimidazole carbamates (**13a-o**). Results coded with a traffic light system; green= good (0.015-0.25 mg/L), amber=acceptable (0.5-4 mg/L) and red=bad (>4 mg/L).

All compounds synthesised displayed antifungal activity, confirming that the 2-propyn-1-oxy group is well tolerated at the carbamate. Most of the propargyl benzimidazole carbamate analogues performed better than the original flubendazole compound. Typically, but not always the thioether derivatives perform better than the ether equivalents e.g., **13a** vs **13b**, **13c** vs **13d**, **13k** vs **13l** and **13n** vs **13o**. This is a very promising result, allowing the viable synthesis of **50a** (see **Figure 3.5**).

3.3.1.3 Synthesis of (5-((4-(Prop-2-yn-1-yloxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate Analogues

The tolerance of the alkyne was also tested on the proposed structure **26** and **32** (see **Figure 3.11**).

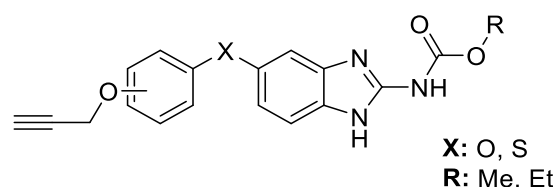
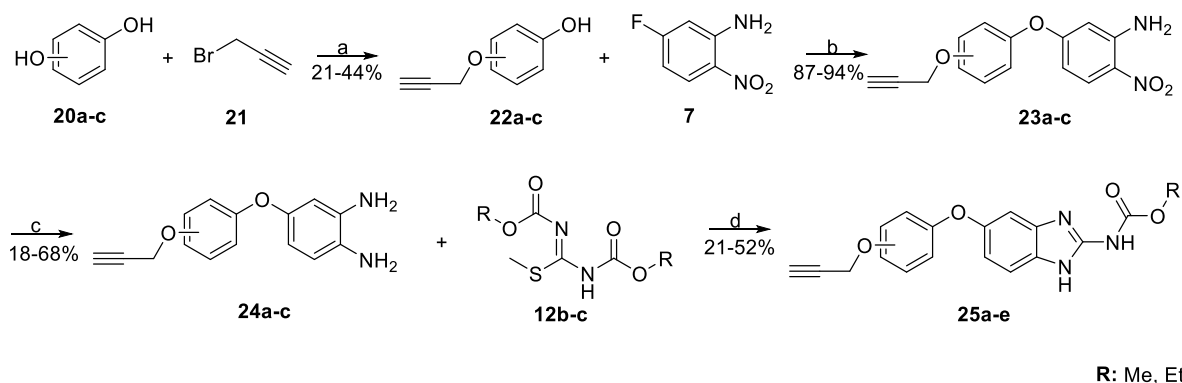


Figure 3.11: Basic structure of the compound **25** and **31** series.

Thiol and ether equivalents of the compound were investigated. The initial procedures differ slightly depending on which analogue was synthesised. Therefore, the ether synthesis which involves a 4-step process will be discussed first (see **Scheme 3.6**).



Scheme 3.6: Reagent and Conditions: (a) NaH, DMF, reflux, 16 hr; (b) K₂CO₃, DMF, reflux, 16 hr; (c) SnCl₂.H₂O, EtOH, reflux, 16 hr; (d) AcOH, MeOH, reflux, 16 hr.

The first step involves the reaction of one hydroxyl group of hydroquinone (**20a-c**) with propargyl bromide (**21**) to create the alkyne handle on **22a-c**.⁴⁴ This synthesis had low yields between 21-44%, with the *meta* hydroxyl derivative (**22b**) performing best and the *ortho* (**22c**)

being worst. The yield is low due to the formation of the disubstituted species. Propargyl bromide was slowly added to the reaction mixture *via* a dropping funnel in DMF. A mixture of mono and double addition products was present in the crude mixture. Purification *via* column chromatography was required to separate the single addition product from the starting materials and the double addition by-product.

The yields for the S_NAr products (**23a-c**) were between 87-94%, with again the *meta* derivative (**23b**) possessing a higher yield compared to the *ortho* (**23c**) and *para* (**23a**) derivatives. The ¹H NMR showed the additional aromatic peaks between 8.1-6.3 ppm and the amine peak as a broad singlet between 6.3-6.1 ppm from the nitroaniline.

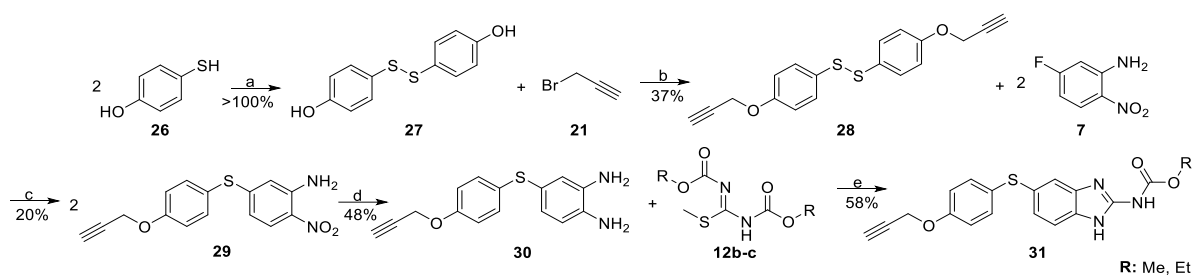
The yields for nitro reduction (**24a-c**) varied considerably from 18-68%. The 18% yield from the *para* derivative (**24a**) is due to human error and has the potential to be considerably higher if the reaction was repeated. The ¹H NMR indicated the presence of the diamine peaks at 5.0-4.9 ppm as a broad singlet.

The synthesis of the methoxy equivalent (**12b**) has been described previously (see **Chapter 2, Section 2.3.1**). However, due to the alkyne handle being attached to the phenyl ring, the carbamate group can be varied to gather further SAR data and more information about the activity of the potential probe. The ethoxy equivalent (**12c**) was synthesised in the same manner. However, when the mixture was basified to a consistent pH of 9-10, no precipitate formed. Therefore, the reaction mixture was extracted with ethyl acetate to afford a light-yellow solid in an 83% yield. These intermediates were not characterised and instead carried through to the next and final step of the synthesis.

For the ring closure reaction (**25a-e**) yields between 21-52% were obtained with the *para* derivative (**25a-b**) performing worst and *ortho* (**25d-e**) affording the highest yield. The ethoxy derivatives of the intermediate gave yields between 30 and 31%. These yields were low despite being performed on an acceptable scale (>100 mg). The yield could have been increased by the slight addition of sodium hydroxide to encourage precipitation (see **Chapter 2, Section 2.14.2.1, General Procedure E**) however, this methodology had not yet been developed.

3.3.1.4 Synthesis of (5-((4-(Prop-2-yn-1-yloxy)phenyl)thio)-1H-benzo[d]imidazol-2-yl)carbamate Analogues

For the synthesis of the thioether derivative the beginning stages are altered due to the heightened nucleophilicity of the sulfur atom (see **Scheme 3.7**).



Scheme 3.7: Reagent and Conditions: (a) DMSO, 60°C, 1 hr; (b) NaH, DMF, reflux, 16 hr; (c) K₂CO₃, DMF, reflux, 16 hr; (d) SnCl₂.H₂O, EtOH, reflux, 16 hr; (e) AcOH, MeOH, reflux, 16 hr.

Due to the high nucleophilicity of the sulfur compared to the oxygen of the hydroxyl group, the sulfur of **26** needs to be protected to prevent the formation of a prop-2-yn-1-ylthiol intermediate. There are a number of thiol masking agents available such as acetyls and the use of methoxymethyls (MOM).^{45, 46, 47} Despite these protection methods being available, disulfide formation was selected. Two equivalents of mercaptophenol (**26**) were dissolved in DMSO and heated at 60°C for 1 hour. A white solid formed upon the addition of iced water and was isolated *via* filtration. The disulfide (**27**) is formed when the thiol and DMSO reacting together to form a sulfoxide-thiol adduct. The resulting adduct reacts further with the remaining thiol to produce a disulfide, sulfide, and water (see **Figure 3.12**).⁴⁸

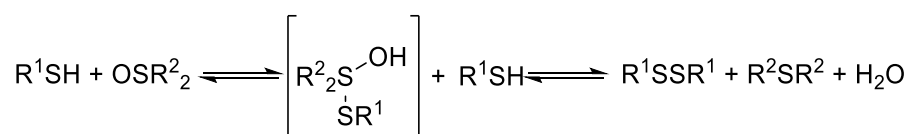


Figure 3.12: Formation of a disulfide bond *via* DMSO.⁴⁸

The yield for this reaction was over 100% due to excess DMSO being present. The literature procedure by *Vibert et al* stated that the crude material is carried through to the next step.⁴⁹ The ¹H NMR shows two aromatic peaks and the OH peak corresponding to the material, and a large amount of DMSO and water. Due to the symmetry of the compound, conclusive evidence that the product had been formed could not be obtained by 1D or 2D NMR. Mass spectrometry indicated the presence of only the starting material and was believed that

disulfide cleavage occurred during analysis. To gain structural data, a melting point range was obtained. The starting material has a melting point of 30-31°C.⁵⁰ Whereas the melting point of the disulfide is 149-151°C.⁵⁰ The melting point obtained for our product was 94.5-97.6°C. The large range and lower melting point are due to the presence of impurities and solvents in the sample. Despite this, the significant increase in the melting point suggested that the S-S bond had formed. Further confirmation will occur from the characterisation of later products in this synthetic route.

For the addition of the alkyne handle, a similar reaction shown in **Scheme 3.6** was used. **28** was formed using two equivalents of propargyl bromide (**21**) for both hydroxyl groups of **27**. The slow addition of propargyl bromide was not required due to the protection of the thiols. The yield of this reaction was 37%, which is lower than anticipated as we did not encounter the same double addition issues seen with compounds **22a-c**. The presence of water in the crude starting material and the unknown yield of **27** may have contributed to the low yield. The reaction procedure by *Vibert et al* reported a yield of 82%, however their reagents differed as they used allylbromide instead of propargyl bromide. Issues with compound characterisation using NMR and mass spectrometry occurred again due to the symmetry of the compound. The aromatic peaks from the phenyl ring and alkyl peaks from the alkyne handle were present, but the integration gave no clarity. Mass spectrometry of the compound showed the presence of a 4-(prop-2-yn-1-yloxy)benzenethiol, suggesting that fragmentation occurred during analysis. The TLC analysis of the reaction material and pure material after column chromatography shows that the R_f value increases significantly compared to the starting material due to the two hydroxyl groups being masked by the propargyl groups. Melting point analysis could not be obtained as the product was a yellow oil.

The procedure by *Vibert et al* suggest that the disulfide bond can be broken with activated zinc powder in acetic acid. However, this step was abandoned as disulfide bond cleavage readily occurred during mass spectrometry. We believed, that under reflux over enough time, **28** would fragment exposing the thiol to form **29** upon reacting with 5-fluoro-2-nitroaniline (**7**). This proved to be correct as it yielded **29** in a 20% yield. Further encouragement to break the disulfide bond to expose the thiol would be recommended if the reaction were to be repeated, to ensure a higher conversion of **28** to **29**. **29** was reduced to **30** using tin (II)

chloride dihydrate with a yield of 48%, which is typical for this type of reaction. **31a** and **31b** were synthesised *via* the ring closure procedure. Both the methoxy (**31a**) and ethoxy carbamate (**31b**) were synthesised to compare the MIC values and to gain more SAR information. **31a** had a yield of 58%, while **31b** had a very poor yield of 8%. Typically, the ethoxy equivalents have lower yields compared to the methoxy. Although it should be noted that the bis(ethoxycarbonyl)-2-methylisothiourea (**12c**) was not freshly synthesised and instead was stored in a freezer for many months. Hydrolysis of the imine centre may have occurred, lowering yields. The bis(methoxycarbonyl)-2-methylisothiourea (**12b**) was freshly made and this probably contributed to the higher yields. Again, the methodology for the encouragement of precipitation by the small addition of sodium hydroxide (see **Chapter 2, Section 2.14.2.1, General Procedure E**) had not been developed, and therefore the yields for both **31a** and **31b** could have been increased if this methodology had been implemented.

3.3.1.5 Results of Benzimidazole Carbamate Alkyne Probes

3.3.1.5.1 MIC Results

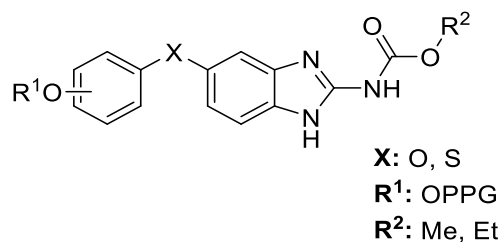


Figure 3.13: Basic structure of the 25 and 31 series

Compound	R ¹	X	R ²	EUCAST (mg/L)
Flubendazole	<i>Para</i> -F	C=O	Me	0.25
25a	<i>Para</i> -OPPG	O	Me	0.25
25b	<i>Para</i> -OPPG	O	Et	0.5
25c	<i>Meta</i> -OPPG	O	Me	0.25
25d	<i>Ortho</i> -OPPG	O	Me	0.25
25e	<i>Ortho</i> -OPPG	O	Et	0.5
31a	<i>Para</i> -OPPG	S	Me	0.06
31b	<i>Para</i> -OPPG	S	Et	0.06

Table 3.2: MIC results for the alkyne handle benzimidazole carbamates derivatives (25a-e and 31a-b). Results coded with a traffic light system; green= good (0.015-0.25 mg/L), amber=acceptable (0.5-4 mg/L) and red=bad (>4 mg/L). OPPG= Propargyl ether.

The MIC values show that the alkyne handle on the phenyl ring is well tolerated. General trends can be taken from the results. In the ether series the methoxy equivalents behave better than the ethoxy. However, with the thioethers, both the methoxy and ethoxy possess very similar activity. Additional thioether derivatives can be synthesised to gain further understanding about the reactivity trends of this series.

3.3.1.6 Synthesis of 3-Trifluoro-3-Phenyldiazirine Probe

The tolerance of the alkyne handle at the carbamate allows use to synthesise **49a** (initial proposed compound). To create the ether function, similar reactions discussed in **Chapter 2.3.1** can be employed. The diazirine can be synthesised from a 2,2,2-trifluoroacetophenone derivative, which is commercially available.

Several synthetic routes were carried out to synthesise a chemical probe, however, constant adjustments had to be performed due to incompatibilities of certain intermediates with various reagents.

Phenyl diazirines are surprisingly chemically stable to several reagents and conditions. However not to all the reactions involved in the synthesis of a benzimidazole carbamate (see **Figure 3.14**).

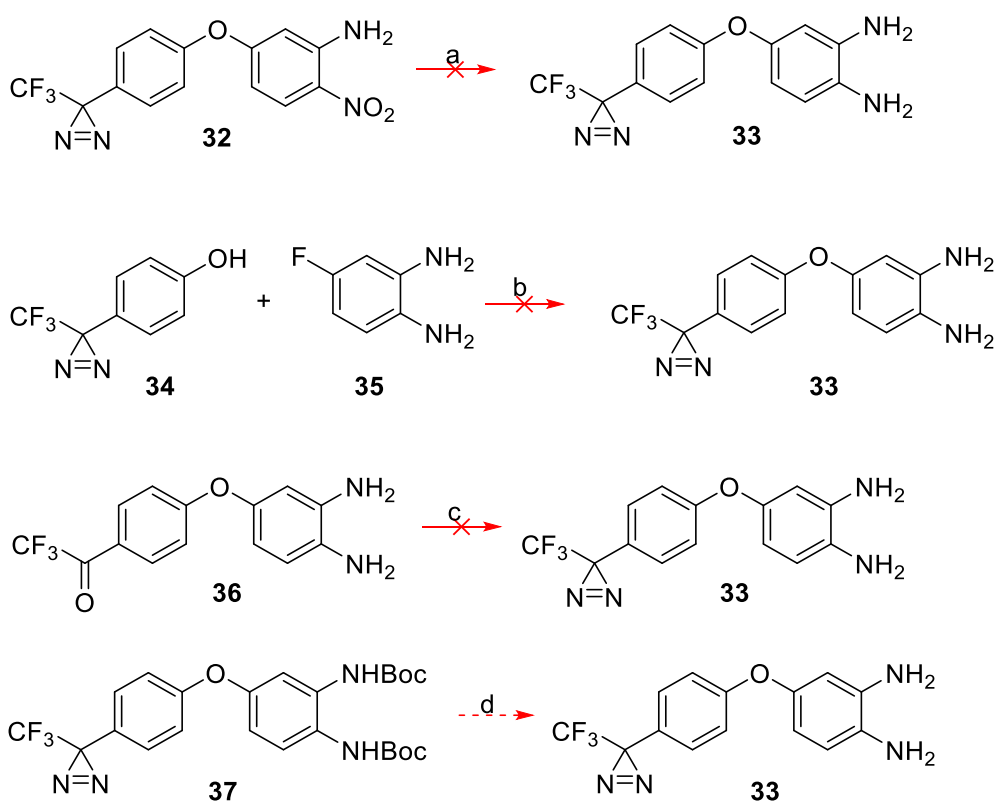


Figure 3.14: Triled and proposed reactions that are not compatible with the synthetic route for the synthesis of a phenyl diazirine probe. *Reactions and conditions:* (a) $\text{SnCl}_2 \cdot \text{H}_2\text{O}$, EtOH, reflux, 16hr; (b) K_2CO_3 , DMF, reflux, 16hr; (c) (I) $\text{NH}_2\text{OH} \cdot \text{HCl}$, MeOH, 60°C, 16hr, (II) TsCl , DCM, NEt_3 , rt, 2hr, (III) NH_3 , -78°C to rt, 16hr, (IV) I_2 , NEt_3 , MeOH, rt, 16hr; (d) TFA, DCM, rt.

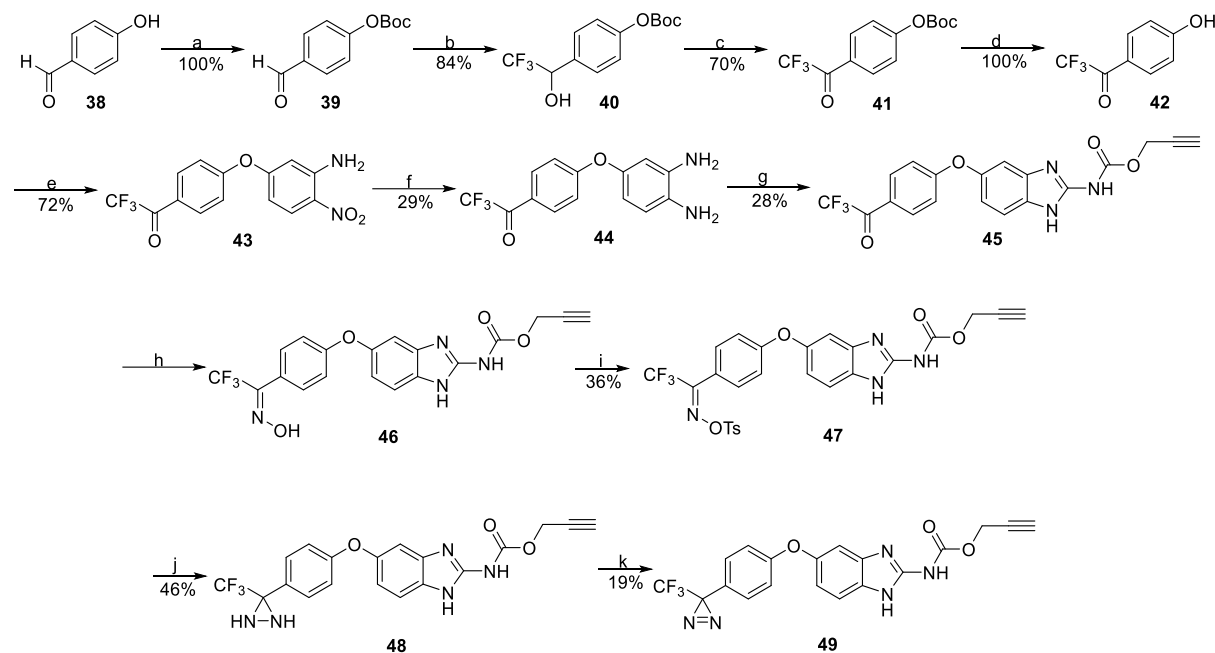
Reduction of the nitro group of **32** to give the diamine (**33**) resulted in a very messy reaction with many by-products. Tin (II) chloride can behave as a Lewis acid and diazirines are known to degrade in the presence of Lewis acids.^{51, 52} Therefore, this reaction is not suitable.

The 2,2,2-trifluoroacetophenone derivative (**34**) did not couple with the aniline (**35**). The nitro group is required as it stabilises the Meisenheimer complex. In the absence of the nitro group no ether was formed, and only leaves starting materials in the reaction mixture.

Due to the issues with the first two reactions, the diazirine was introduced at a later stage in the synthetic procedure. The 2,2,2-trifluoroacetophenone derivative (**36**) was coupled to 5-fluoro-2-nitroaniline (**7**) *via* the typical S_NAr reaction. The nitro group can then be reduced to form the diamine without any issues. The diazirine synthesis was attempted with the diamine present, however this was unsuccessful, resulting in several by-products which were not easily purified. The diamines were believed to interfere with the reaction, as similar compounds containing amines are frequently protected prior to the synthesis of the diazirine.^{53, 54, 55, 22}

The protection of the diamine to form **37** was not attempted, as previous synthetic knowledge within the group states that very harsh conditions must be used for the eventual deprotection step. Diazirines can withstand boc deprotection methods such as TFA, however 60% TFA in DCM for 2 days would be too acidic for the diazirine to survive.^{53, 56} Therefore, this reaction was not attempted.

Due to the previous issues and preconceived problems, an alternative synthetic route was designed (see **Scheme 3.8**).



Scheme 3.8: *Reagent and Conditions:* (a) Boc₂O, PPh₃, THF, 65°C, 16 hr; (b) TMS-CF₃, TBAF, THF, 0°C to rt, 1 hr then HCl (1M), THF, rt, 2 hr; (c) SIBX, EtOAc, reflux, 16 hr; (d) TFA, DCM, rt, 3 hr; (e) 5-fluoro-2-nitroaniline (**7**), K₂CO₃, DMF, reflux, 16 hr; (f) SnCl₂·H₂O, EtOH, reflux, 16 hr; (g) 1,3-Bis(prop-2-yn-1-yloxy)carbonyl-2-methylisothiourea (**12a**), AcOH, MeOH, reflux, 16 hr; (h) NH₂OH·HCl, MeOH, NaOAc, 50°C, 16 hr; (i) TsCl, K₂CO₃, DMSO, 70°C, 16 hr; (j) NH₃, -78°C to 80°C, 16 hr; (k) I₂, MeOH, NEt₃, rt, 3 hr.

Before the reaction could proceed the hydroxyl group of **38** needed to be protected. The hydroxyl group was protected with boc anhydride, which gave **39** in a quantitative yield. The adopted procedure by *Monterio et al* reported a yield of 96% on the same starting material.⁵⁷

The trifluoromethyl group was introduced by converting an aldehyde to a trifluoromethyl carbinol (**40**). This reaction used the Ruppert Surya Prakash procedure using Rupperts reagent, TMSCF₃. A high yield of 98% was obtained. *Pauff et al* obtained a yield of 94-95% depending on the substrate. The substrates slightly differed from **39**, as the hydroxyl group was protected with a methyl, methoxy, or an isopropoxy substituent.⁵⁸ The starting material was converted to **40** in a 84% yield.

The alcohol **40** was oxidised to ketone **41**. There were issues with the synthesis, as several oxidation reactions were attempted. The oxidation of trifluoromethyl alcohols to their

corresponding ketones are described as being particularly difficult, compared to its mono and difluoromethyl counterparts. It has been stated that the trifluoroalkyl group is extremely electron withdrawing, making oxidation very challenging.^{59, 60, 61, 62} Classical oxidation procedures involving reagents like MnO₂, are unsuccessful or require much harsher conditions.⁵⁹

Initially Dess-Martin methodology was used. A procedure by *Linderman et al* reported a yield of 76% on a similar substrate to **40**.⁶³ Unfortunately this procedure was not successful, as the crude ¹³C NMR showed the absence of the ketone peak and only the alcohol remained. The procedure was adapted slightly by the addition of water, which can catalyse the reaction using Dess-Martin Periodinane (DMP). The new procedure by *Meyer et al* was used to no avail.⁶⁴

A procedure by *Mori et al* was utilised involving the chromium agent pyridinium chlorochromate (PCC).⁶⁵ *Mori et al* afforded a yield of 98%. When this reaction was attempted on a small scale of 0.13 g, a very poor yield of 13% was obtained. The crude ¹H NMR of the reaction mixture indicated that both the ketone and starting material were present, with a larger ratio of starting material. The reaction was conducted under strict anhydrous conditions using dry materials and molecular sieves to prevent unwanted side reactions and lower yields. The work up procedure of the reaction was extremely difficult. The procedure called for filtration *via* celite to remove the reduced chromium by-products, followed by washes with ether. The filtration process took a long time, with a lot of sticky material remaining on the reaction flask walls despite washing and agitation. Therefore, it was believed that much of the material remained in the flask leading to a poor recovery. The reaction was modified by adding silica gel to the reaction vessel, a method discussed by *Luzzio et al*.⁶⁶ The presence of the silica gel allows the reduced chromium waste to be adsorbed onto the silica, leading to easier removal during filtration. This also leads to easier disposal of solid chromium waste, which is considered a very hazardous material.⁶⁶ The addition of the silica gel, did in fact aid the work up procedure and increased the yield to 54%.

Due to the poor yields of **41** and health and safety and environmental concerns over the use of chromium agents, a different type of oxidising agent was used. Success was found with stabilised 2-iodoxybenzoic acid (SIBX). SIBX is an attractive oxidising agent as it uses mild conditions and is selective as opposed to some chromium agents that oxidise beyond the ketone to the carboxylic acid (see **Figure 3.15**). SIBX is environmentally friendly as it does not

contain any toxic heavy metals. A procedure by *Cheng et al* was used which involved SIBX to be refluxed in a solution of **40** and EtOAc.⁶² The reduced SIBX was filtered leaving the starting material and ketone in the filtrate. The procedure can be left anywhere between 9 to 30 hours; however, this reaction was left for 24 hours. Trace amounts of the starting material appeared on the TLC plate, and the crude material was purified *via* column chromatography to afford a yellow oil in 70% yield. The yields obtained by *Cheng et al* were from 90% upwards depending on the substrate used. Our yield is a little lower compared to the literature, but much greater than the yield we obtained from the previous procedures.

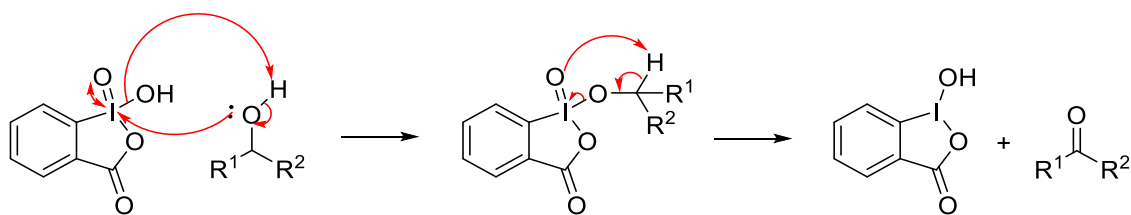


Figure 3.15: SIBX oxidation mechanism.⁶⁷

The boc group was removed to unveil the hydroxyl group on **42** at a quantitative yield, using TFA in DCM. **43** was formed under the standard S_NAr procedure with a yield of 72%. The nitro reduction *via* tin (II) chloride dihydrate to form **44** occurred at a poor 23% yield. These reactions tend to vary a lot throughout the project and can be sensitive to different functionalities attached to the nitroaniline. Finally, the benzimidazole carbamate **45** can be synthesised with a 28% yield *via* the ring closure reaction using the 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea intermediate (**12a**). The yield was low as the reaction was conducted on a small scale of 22 mg. More material was carried through for the subsequent reactions, with slight increases in the percentage yields.

For the synthesis of the diazirine, several reactions must take place. First the trifluoromethylketone on **45** must be converted into an oxime using hydroxylamine. *Savarin et al* methodology was used in which methanol was used as the solvent and heated to 64°C to help **45** dissolve.⁶⁸ Not all of **45** was able to dissolve in the solution, and instead formed a suspension. *Savarin et al* allowed their reactions to stir between 1-48 hours, this reaction was left for 16 hours. The TLC of the reaction mixture confirmed that not all the **45** fully converted to the oxime (**46**), due to the slight insolubility in methanol. Increasing the equivalents of hydroxylamine and base was not effective as the less reactive carbamate ketone begins to

react with the reagents. Therefore, the equivalents were kept the same and the lower yield was accepted.

The crude oxime (**46**) was carried through to the subsequent step to form a tosyloxime intermediate (**47**) *via* tosylation. Tosylation is necessary to create a good leaving group for the next step of the synthesis with ammonia. A tosylation reaction is typically carried out at room temperature. Due to the insolubility of benzimidazole carbamates, traditional reaction conditions of stirring tosyl chloride in a solution of oxime and DCM would not work. The only solvents the oxime will dissolve in are methanol and DMSO. Methanol cannot be used as the alcohol would also be prone to tosylation, reducing yields. Also, as a methanol-tosyl adduct forms, the volume of the solvent decreases, allowing the oxime (**46**) to precipitate out of solution, resulting in even lower yields of **47**. A procedure by *Butkevich et al* utilises DMSO as the solvent for a tosylation reaction.⁶⁹ The reaction was heated to 80°C which further aided the solubility of **46**. Purification to remove excess tosyl chloride in the reaction mixture was performed and yielded **47** in a 36% yield. Butkevich's reaction was carried out on much more soluble compounds compared to **46** and therefore achieved yields of 71%.

For the formation of the diaziridine ring, **47** was treated with ammonia in methanol. The lone pair of ammonia can attack the electrophilic imine centre of **47** to form an unstable tetrahedral intermediate (**50**). The amino group can then form a sigma bond with the other amino group, and with a proton transfer generate tosic acid (see **Figure 3.16**). A suspension formed due to the slight insolubility of **47** and was heated gently to encourage dissolution. Non-anhydrous conditions have impacted the yield, as methanol and water are also able to nucleophilically attack the imine centre of **47**. The crude material was filtered through a sinter funnel and washed with methanol before being carried through to the final oxidation step. Crude yields between 33-46% were obtained, which is considerably lower than the yields found in the literature which are typically between 87-90%.^{22, 70, 71}

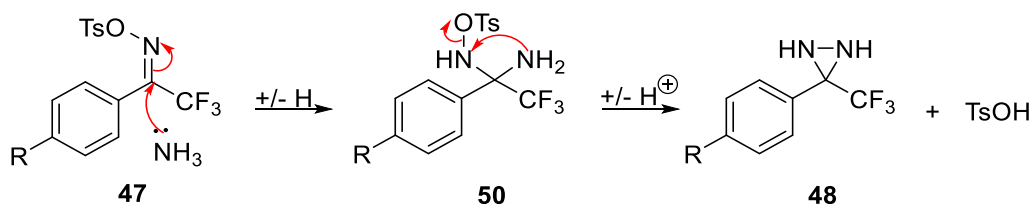


Figure 3.16: Mechanism for the formation of a diaziridine ring.⁵³

A lot of different methods are available to form the photoreactive group such as a combination of ammonia and lithium amide, potassium manganate, or silver. Success has been found with iodine and triethylamine to afford **49**. The most cited oxidation reaction of diaziridines involve iodine and triethylamine in DCM. However, this leads to very low yields as **48** is insoluble in DCM. The procedure was adapted by using methanol as the solvent, following similar methods to *Cheng et al.*⁷² *Cheng et al* were able to obtain a yield of 70%, while the oxidation of **49** only afforded a yield of 19%. The yields are considerably lower due to the probable low ratio of diaziridine in the reaction mixture of **48** and the solubility issues associated with the benzimidazole carbamate.

3.3.1.7 Results of 3-Trifluoro-3-Phenyldiazirine Probe

3.3.1.7.1 MIC Results

The final phenyl diazirine probe (**49**) was submitted for MIC testing to confirm whether the PAL probe possessed any antifungal activity, allowing it to be suitable for PAL experiments (see **Table 3.3**).

Compound	EUCAST (mg/L)
49	>4

Table 3.3: MIC results for the phenyl diazirine photoreactive probe (**49**). Results coded with a traffic light system; green= good (0.015-0.25 mg/L), amber=acceptable (0.5-4 mg/L) and red=bad (>4 mg/L).

The MIC results confirm that **49** possesses no antifungal activity. This was repeated multiple times to test whether the result was accurate. This result was surprising as the structure of **50** is very similar to other flubendazole analogues synthesised by the Nixon group. Upon reviewing the SAR data, it was found that the trifluoromethyl group is not well tolerated when attached to the phenyl ring and can cause a loss in antifungal activity. Therefore, the trifluoromethyl group stabilising the diazirine functionality may be the cause of the loss of activity. Alternative diazirine groups were explored to address this potential issue. One solution came in the form of utilising the aliphatic diazirine.

3.3.1.8 Synthesis of Minimalist Photo-Crosslinker

3.3.1.8.1 Introduction of the Minimalist Photo-Crosslinker

Due to the potential issue with the trifluoromethyl group, an alternative way to synthesise a stable diazirine structure was sought. Recently aliphatic diazirines have gained attention

within research groups. Initially researchers were reluctant to use aliphatic diazirines as they generate smaller amounts of carbene upon photo irradiation, resulting in lower yields of protein-probe adducts. As well as this, less stabilisation is experienced resulting in the greater formation of the undesired diazo intermediate. The diazo intermediate has a longer half-life, allowing it to diffuse from the initial site, leading to non-specific labelling.⁷³ Despite these issues, the aliphatic chain has gained much attention recently due to their compact sizes. *Li et al* described the aliphatic chain as a minimalist linker in which the photoreactive and reporter group are incorporated on the same aliphatic chain. The smaller size leads to minimisation of the steric effects on the natural interactions between the ligand and protein. *Li et al* described the formation of a chain with a terminal alkyne, which can be used in subsequent *ex vivo* target identification by conjugation to a suitable reporter using well established bioorthogonal click chemistry.⁷⁴

The linkers have a minimum chain length to ensure successful synthesis. *Li et al* state that the minimum number of carbons on each side of the diazirine is two (see **Figure 3.17**).⁷⁴

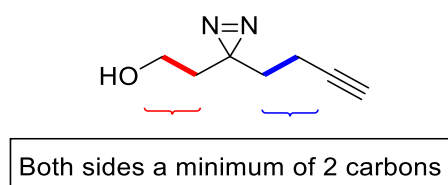


Figure 3.17: Chain length requirement for the synthesis of the minimalist linker.⁷⁴

It is essential to have a minimal length of two carbons on either side of the diazirine moiety. The chain highlighted in red must be at least two carbons in length. This is to ensure functional groups such as a hydroxyl group are synthetically accessible from the commercially available ethyl acetoacetate. The chain length also prevents steric hindrance between the diazirine and the benzimidazole core. The chain highlighted in blue must also be at least two carbons in length to prevent side reactions. Side reactions occur as Michael acceptors and acidic α protons can react during the synthesis of the diazirine. These side reactions can be prevented by increasing the bond length.^{74, 75}

The minimalist linker can be developed with several different functionalities which can be attached to the binding group of the ligand *via* numerous robust acylation and alkylation reactions (see **Figure 3.18**).⁷⁴

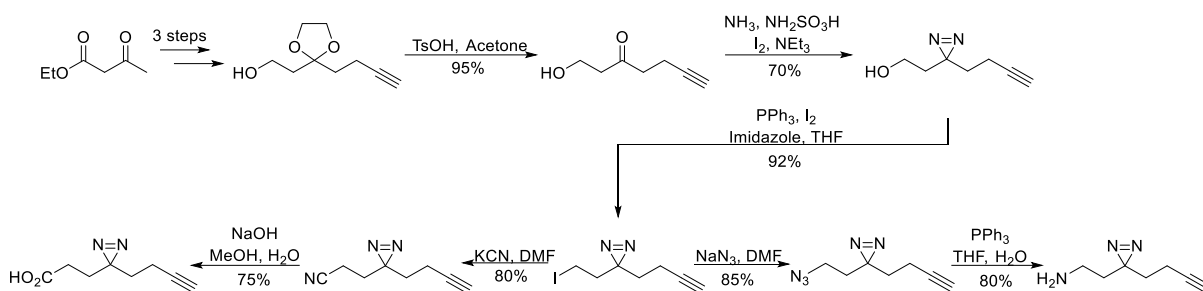


Figure 3.18: Synthesis of various minimalist linkers.⁷⁴

With the trifluoromethylphenyl diazirine, the synthesis of the diazirine is left until the end of the synthetic route. This led to poor yields, due to the solubility issues associated with the benzimidazole carbamate. Incorporation of the diazirine before the nitro reduction is the main source of the problem. *Hashimoto et al* were able to reduce a range of different functionalities such as esters, carboxylic acids, and nitriles while keeping the diazirine intact.⁷⁶ However, the conditions for the catalyst for nitro reduction would not be suitable and therefore could not be carried out. No selective reduction for nitro groups over the diazirine could be found in the literature. Therefore, it has been presumed that the selective nitro reduction would not be achievable (see **Figure 3.19**).

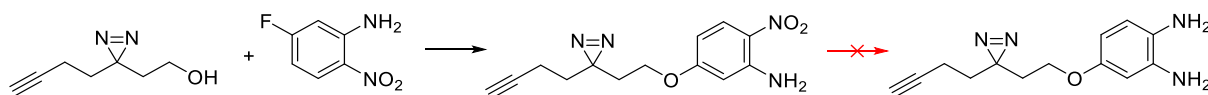


Figure 3.19: Scheme to illustrate the synthetic incompatibility of a diazirine under nitro reduction methods.

Therefore, the synthesis of the full minimalist linker which could be inserted into the binding group was abandoned. Instead, the ketone functionality of the minimalist linker will be incorporated, and the diazirine installed in the final stage. The proposed synthetic route for the benzimidazole probe with the minimalist linker involves 9 steps (see **Scheme 3.9**). Two variants of the benzimidazole carbamate with a minimalist linker were synthesised (see **Figure 3.20**). **61a** has the binding group of carbendazim, which possesses moderate antifungal activity against *C. neoformans* (see **Chapter 1, Table 1.1**). **60b** is structurally like other more potent antifungals due to the presence of the second phenyl ring. The minimalist linker can be positioned around the phenyl ring, but for now it is in the *para* position.

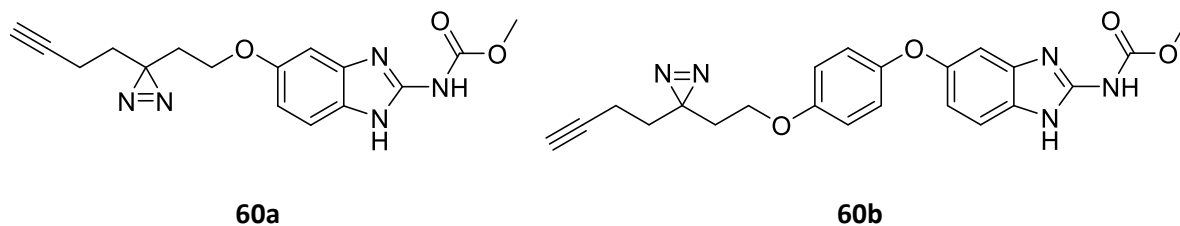
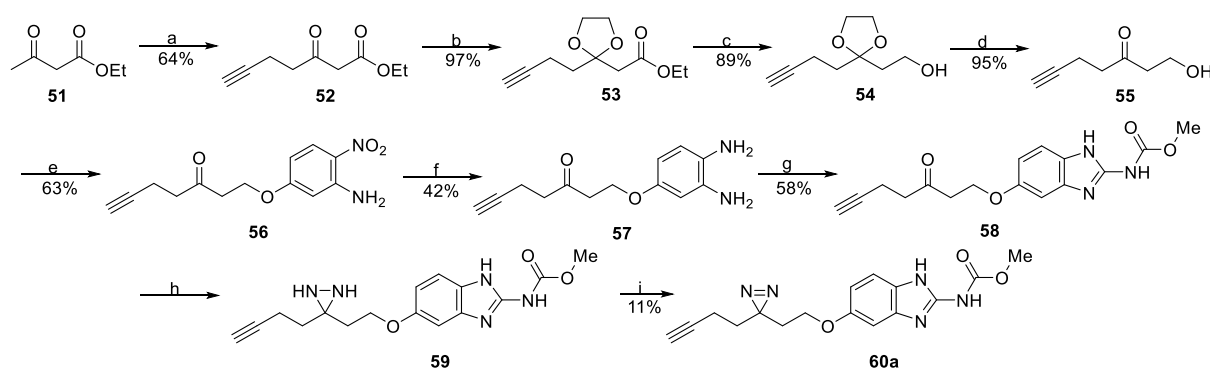


Figure 3.20: Structure of minimalist linker probes **60a** and **60b**.

3.3.1.8.2 Synthesis of Methyl (5-(2-(3-(but-3-yn-1-yl)-3*H*-diazirin-3-yl)ethoxy)-1*H*-benzo[*d*]imidazol-2-yl)carbamate

For the synthesis of **60a**, a 9-step procedure is required, starting with the synthesis of the ketone derivative of the minimalist linker (see **Scheme 3.9**).



Scheme 3.9: *Reagent and Conditions:* (a) *n*-BuLi, DIPA, THF -78°C, then propargyl bromide (**21**), rt, 16 hr; (b) ethylene glycol, TsOH, toluene, reflux, 6 hr; (c) LiAlH₄, THF, rt, 3 hr; (d) TsOH, acetone, rt, 2 hr; (e) 5-fluoro-2-nitroaniline (**7**), K₂CO₃, DMF, reflux, 16 hr; (f) SnCl₂·H₂O, EtOH, reflux, 16 hr; (g) 1,3-Bis(methoxycarbonyl)-2-methylisothiourea (**12b**), AcOH, MeOH, reflux, 16 hr; (h) NH₃, NH₂SO₃H, MeOH, reflux, 4 hr; (i) Oxalyl chloride, DMSO, DCM, NEt₃, -78°C to rt, 90 min.

The synthesis of the minimalist linker was taken from *Li et al.*⁷⁴ The first step involved the installation for the alkyne handle (**52**) on the commercially available ethyl acetoacetate (**51**). LDA made from *n*-BuLi and DIPA deprotonates the terminal proton, allowing nucleophilic attack of the propargyl bromide (**21**) *via* an S_N2 reaction. Fresh LDA was synthesised as previous attempts to use premade LDA 2.0 M in THF from Sigma Aldrich resulted in poorer yields. The yields obtained from fresh LDA ranged from 64-72%. The percentage yields described in the procedure was 85%, which is higher than the yields we obtained.

For the synthesis of **53**, the ketone was protected with a cyclic acetal group. Compound **52**, ethylene glycol, tosic acid, and toluene were refluxed under Dean Stark conditions to remove the water and drive the reaction into completion. Previously alternative protection methods

were tried such as trimethylorthoformate to form two acyclic methyl acetals.⁷⁷ This method obtained high yields between 88-91%, but was abandoned for the Dean Stark procedure due to the slightly higher yields. The Dean Stark method gave yields between 84-97%, which is slightly higher than *Li et al* reported yield of 90%.⁷³

Once the reactive ketone is protected, the ester can be reduced to the primary alcohol (**54**). The yields of these reactions varied and could be quite low depending on the number of equivalents for LiAlH₄. It was found that using an excess of the reducing agent with 5 equivalents was able to achieve a yield of 89-91%, as using 2.5 equivalents only gave a yield of 64%. This is quite a dramatic leap in yield and is more comparable with the literature which was able to obtain **54** in a 96% yield.

The ketone was deprotected to form **55** by mixing **54** with tosic acid in acetone at room temperature. The yields for this reaction were high at 95-96%. Depending on the scale, the reaction times varied between 2-16 hours. The high yields are comparable to *Li et al* with yields of 95%. At this stage, the basic backbone of the minimalist linker was provided. This functionalised alkyl alcohol could be incorporated into the benzimidazole structure for the probe.

56 was formed in a yield of 55-63% by following the same basic S_NAr procedure. The nitro group of **56** was reduced *via* the typical method with hydrated stannous chloride, to afford **57** in a 42-66% yield. The benzimidazole carbamate structure (**58**) was then formed *via* the ring closure reaction with 1,3-bis(methoxycarbonyl)-2-methylisothiourea (**12b**). The methoxy group was chosen over the ethoxy and the propan-1-yloxy group as the alkyne handle is included on the minimalist linker and the ethoxy group tends to afford lower yields and poorer MIC values *via* EUCAST testing. A low to moderate yield of 33-58% was achieved.

The ketone of **58** was converted to a diazirine. The ketone was treated with 7N ammonia in methanol, followed by the addition of sulfamic acid as a base to form the precursory diaziridine functionality (**59**). The carbonyl centre of the ketone undergoes nucleophilic attack by ammonia to form a simple imine. The sulfamic acids amine group attacks the imine centre to form an unstable tetrahedral intermediate. The negative charge on the nitrogen can then form a sigma bond to the other nitrogen while cleaving the sulfonic acid (see **Figure 3.21**). The methanol in the ammonia solution was able to dissolve most of the insoluble **59**

at reflux, resulting in a high conversion. The yield for this reaction was not calculated as the crude mixture was carried through to the next step.

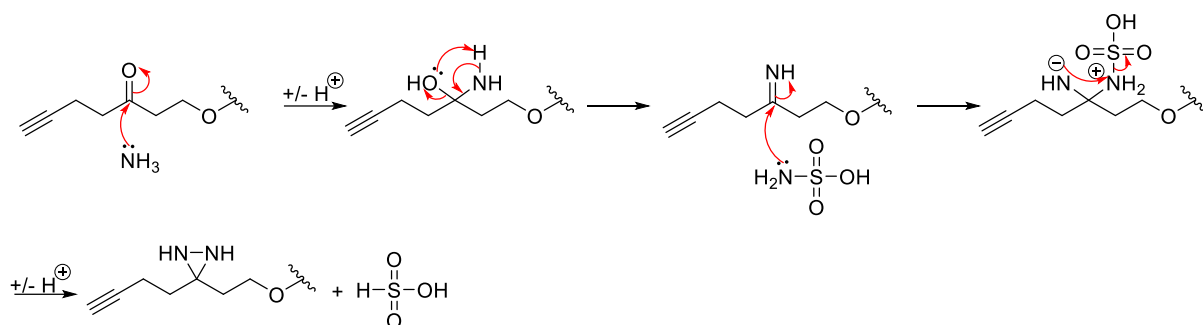
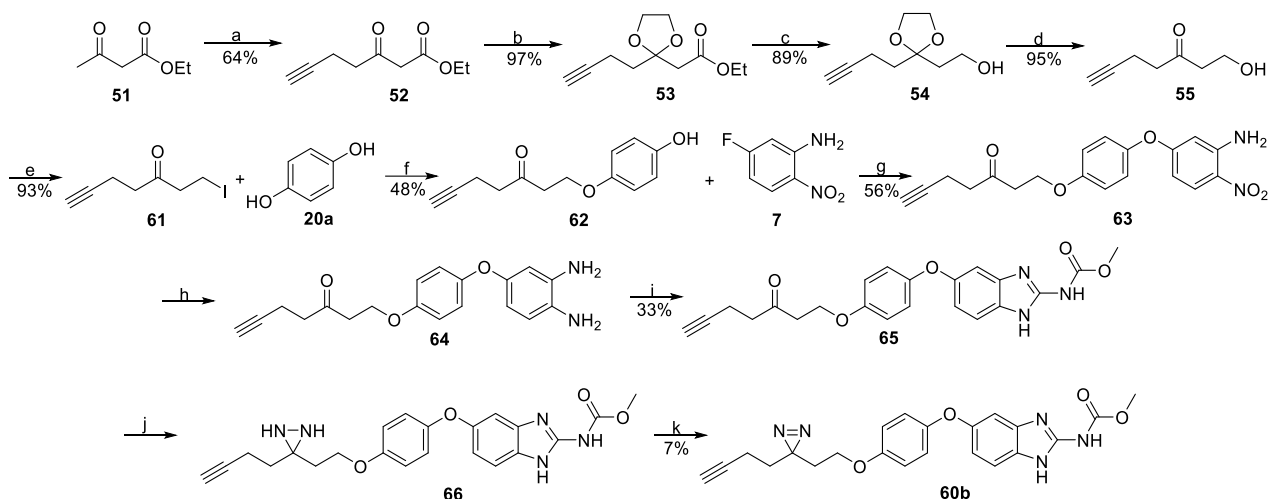


Figure 3.21: Mechanism of the formation of the diaziridine ring. Adapted from: *Base-Mediated One-Pot Synthesis of Aliphatic Diazirines for Photoaffinity Labelling*, L. Wang *et al.*⁷⁸

The diaziridine (**59**) can then be oxidised to the diazirine (**60a**) using a procedure reported by *Kumar et al* which uses a co-solvent of DMSO and DCM in the presence of oxalyl chloride.⁷⁹ *Kumar et al* performed the reaction on a thiazole derivative and yielded the diazirine at 97% as a colourless syrup. This method achieved yields of 8-11%. The yield is considerably lower than the procedure described by *Kumar et al*, as the reaction is initially performed at -78°C and then warmed to room temperature. Heat is usually required for the benzimidazole to dissolve in the DMSO, therefore, a slight suspension forms, resulting in a reduced yield. The yield was attempted to be improved slightly by using a different oxidising agent. *Wang et al* formed the diazirine *in situ* with the base KOH, under air for 2 hours. *Wang et al* used several different bases such as lithium amide, sodium ethoxide, potassium methoxide, and sodium hydroxide, but found that potassium hydroxide gave the best yield. The yield achieved for the oxidation of a simple aliphatic diaziridine was 67%. Most of these reactions were carried out with neat liquid ammonia and the presence of methanol decreased the yields. The reaction was attempted with methanol to aid solubility, however only a yield of 4% was obtained. Therefore, the previous method with oxalyl chloride was continued, with yields of 8-11%.⁷⁸

3.3.1.8.3 Synthesis of Methyl (5-(4-(2-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)ethoxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate

For the synthesis of **60b**, the procedure differs slightly and has additional steps making it an 11-step route (see **Scheme 3.10**).



Scheme 3.10: Reagent and conditions: (a) *n*-BuLi, DIPA, THF, -78°C 15 mins, then propargyl bromide (**21**), rt, 16 hr; (b) ethylene glycol, TsOH, toluene, reflux, 2 hr; (c) LiAlH₄, THF, rt, 3 hr; (d) TsOH, acetone, rt, 6 hr; (e) PPh₃, I₂, imidazole, THF, rt, 16 hr; (f) K₂CO₃, acetone, reflux, 16 hr; (g), K₂CO₃, DMF, reflux, 16 hr; (h) SnCl₂.H₂O, EtOH, reflux, 16 hr; (i) 1,3-Bis(methoxycarbonyl)-2-methylisothiourea (**12b**), AcOH, MeOH, reflux, 16 hr; (j) NH₃, MeOH, 40°C, 3 hr, then NH₂SO₃H, rt, 16 hr; (k) Oxalyl chloride, DMSO, DCM, NEt₃, -78°C to rt, 90 min.

The additional phenyl ring meant that extra steps were needed. The hydroxyl minimalist linker was synthesised in the same manner as **Scheme 3.9**. However, for the attachment of the minimalist linker to the phenyl ring, the hydroxyl group of **55** had to undergo iodination to form an alkyl halide (**61**). **55** was stirred in a suspension of iodine, triphenylphosphine, imidazole and THF at room temperature to afford **61** in a 93% yield. The 4-hydroxyphenol (**12a**) was added very slowly *via* a dropping funnel to minimise the formation of a di-substituted phenyl to afford **62**. The yields for this reaction were a little low at 48%, due to the inevitable formation of the doubly substituted phenyl ring. Once synthesised, the reaction can undergo the same procedures in **Scheme 3.9**. The S_NAr reaction afforded **63** in a 56% yield, like the formation of **56**. The nitro reduction with hydrated stannous chloride yielded **64**. **65** was formed *via* the ring closure reaction using 1,3-bis(methoxycarbonyl)-2-methylisothiourea (**12b**) in a 33% yield. The formation of the diazirine ring used the same

procedures in **Scheme 3.9**, obtaining **60b** in a 7% yield. This is very low but is expected due to the insolubility and low scale of **65** from an 11-step synthetic procedure.

3.3.1.8.4 Results of the Minimalist Photo-Crosslinkers

3.3.1.8.4.1 MIC results

Compound	EUCAST (mg/L)
60a	>4
60b	>4

Table 3.4: MIC results for the minimalist photoreactive probes (**60a** and **60b**). Results coded with a traffic light system; green= good (0.015-0.25 mg/L), amber=acceptable (0.5-4 mg/L) and red=bad (>4 mg/L).

The MIC values obtained for both **60a** and **60b** PAL probes show they do not possess antifungal activity. This was also seen with compound **49**, which was initially attributed to the presence of the trifluoromethyl group. From these results it can be assumed that the diazine ring itself is causing issues or is degrading during MIC testing. Further investigation is needed, as it is not known whether chemical degradation is occurring which could result in a lack of activity. It should be noted that issues have been encountered when benzimidazole analogues undergo MIC testing, such as inconsistent values and lack of activity from repeat testing. Occasional “crashing out” of the benzimidazole analogues in DMSO have occurred, altering the carefully prepared concentrations, thus providing incorrect MIC values. Compounds prepared in this chapter could be experiencing the same problems, resulting in a lack of antifungal activity. Therefore, consistency in MIC testing is required to rule this issue out for these compounds.

3.3.1.9 Synthesis of Benzophenone Probe

3.3.1.9.1 Introduction to benzophenones

Benzophenones have become established as a photophore which is able to covalently attach to hydrophobic regions of binding proteins. There are several reasons to use benzophenones such as being chemically stable to acids and bases, stability in ambient light, activation not being photo dissociative, and preferential labelling of hydrophobic binding regions. C-H insertion dominates leading to adducts that are stable to chemical and enzymatic polypeptide

cleavage. However, the size of the benzophenone and its hydrophobicity can perturb the cell, leading to a disruption of natural biological processes.^{80, 81, 82}

The benzophenone forms a diradical species. When activated at 330-360 nm, the benzophenone goes into a triplet state excitation. Hydrogen abstraction on the oxygen radical occurs on a nearby amino acid residue. This then leads to a radical recombination to create a covalent crosslink (see **Figure 3.22**).^{80, 81, 82}

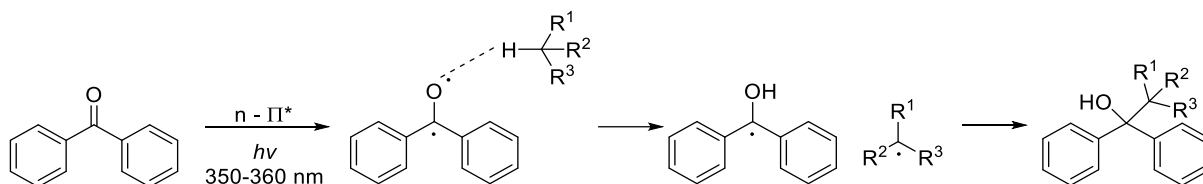
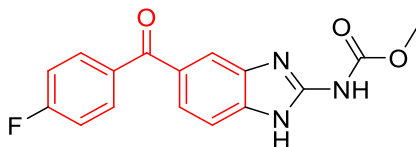


Figure 3.22: Mechanism of a benzophenone chemical probe.⁸⁰

The benzophenone structure is desirable for this project as two commercially available benzimidazole compounds that possess the benzophenone functionality display promising antifungal activity (see **Chapter 1, Table 1.1** and **Figure 3.23**).

3.23A



3.23B

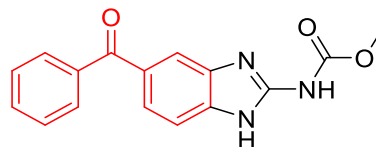
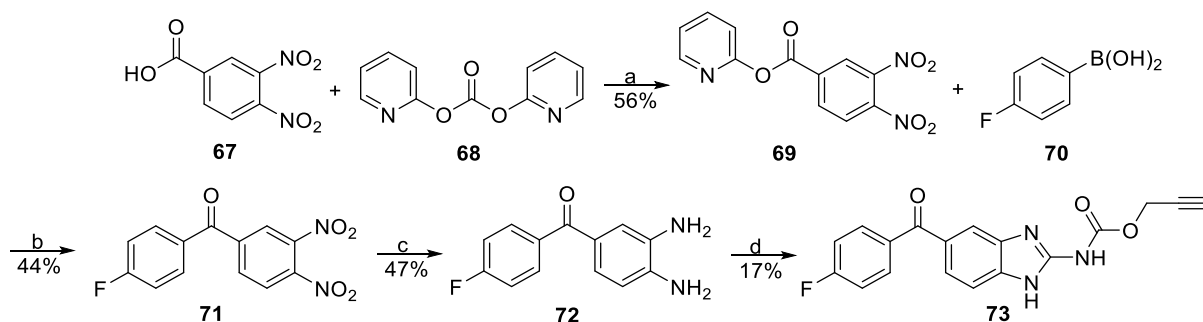


Figure 3.23: Benzimidazole structures with benzophenone functionality (shown in red). **3.23A:** Flubendazole; **3.23B:** Mebendazole.

3.3.1.9.2 Synthesis of Prop-2-yn-1-yl (5-(4-fluorobenzoyl)-1H-benzo[d]imidazol-2-yl)carbamate

A flubendazole based probe (**73**) with the already present benzophenone structure was proposed to be synthesised *via* a 4-step process (see **Scheme 3.11**).



Scheme 3.11: Reagent and Conditions: (a) DMAP, DCM, rt, 2 days; (b) Pd(PPh₃)₄, 1,4-dioxane, 50°C, 16 hr; (c) SnCl₂·H₂O, EtOH, reflux, 16 hr; (d) 1,3-Bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea (**12a**), AcOH, MeOH, reflux, 16 hr.

The formation of the ketone for the clickable flubendazole photoprobe involves a Suzuki-Miyaura reaction like the reaction as seen in **Chapter 2, Scheme 2.3**. A simple commercially available boronic acid can be used for the synthesis. An activated ester (**70**) can be synthesised by mixing **67** and **68** together for 2 days to obtain a yield of 65%.⁸³ Purification was challenging as **68** and **69** co-eluted due to similar *rf* values observed on the TLC plate of the crude reaction mixture. The crude material was carried through to the next step for the Suzuki-Miyaura reaction. The crude material was characterised, and the additional pyridine aromatic peaks can be observed in the ¹H NMR spectrum.

Previous work within the group synthesising different ketones with similar functionalities found success with the use of Pd(PPh₃)₄ in 1,4-dioxane.⁸³ This Suzuki-Miyaura methodology was used while mixing the reagents **69** and **70** in a screw fix sealed tube to afford **71** in a 44% yield. The synthesis of the product was determined by structural analysis. The ¹H NMR shows the presence of the different aromatic peaks from the two substituted phenyl rings. The ¹³C NMR confirms the synthesis of **71** as a ketone peak is observed at 193 ppm.

The two nitro groups of **71** were reduced to the diamine with double the equivalents of tin chloride dihydrate. Typically, TLC analysis for nitro reductions have three spots: the starting material (**71**), the product spot (**72**) and baseline by-products and/or crude material from a previous reaction. For this reaction 4 spots were observed. The additional spot was found

between the starting material and the desired product. This additional spot was isolated from the reaction material *via* purification with column chromatography. ¹H NMR analysis indicated that the additional spot was a monoamine product. The nitro reduction for this material was repeated to gain more **72** and improve the yield. The yield of **72** from the first reaction was 49%. However, with combination with the second nitro reduction, the yield improved to 57%. If the reaction were to be repeated, extra tin chloride dihydrate could be added to ensure the reaction is driven to completion, or the reaction can be refluxed for longer encouraging the second nitro group to reduce. The synthesis of **72** was confirmed by the presence of the diamine peak shown as a broad singlet peak at 3.30 ppm which integrated to 4.

The final step of the synthesis involves the ring closure reaction. The 1,3-bis(prop-2-yn-1-yloxy carbonyl)-2-methylisothiourea (**12a**) was selected as the reagent for this reaction as a clickable reporter tag was required. The yield of **73** of this reaction was lower compared to similar reactions involving the propargyl thiourea with a yield of 17%. The diamine (**72**) was carried through crude to the final reaction, which may have reduced yields as a lower amount of diamine was present in the reaction mixture.

3.3.1.9.3 Results for Benzophenone Probe

3.3.1.9.3.1 MIC Results

Compound	EUCAST (mg/L)
73	>4

Table 3.5: MIC results for the benzophenone chemical probe (**73**). Results coded with a traffic light system; green= good (0.015-0.25 mg/L), amber=acceptable (0.5-4 mg/L) and red=bad (>4 mg/L).

Compound **73** was shown to possess very little antifungal activity (see **Table 3.5**). The structure and purity of the compound was confirmed by numerous analytical techniques such as ¹H and ¹³C NMR, mass spectrometry, IR, melting point, and HPLC. Therefore, it is not known why compound **73** possessed poor antifungal activity. Issues with the plates for the MIC work have been a problem as they have provided unreliable and divergent results on already tested compounds. Another explanation could be due to the presence of the propargyl group on the flubendazole structure. Propargyl groups tend to be well tolerated on the carbamate (see **Tables 3.1** and **3.2**), with some possessing better antifungal activity than the original

flubendazole. However, the addition of the propargyl group on the flubendazole seems to kill the activity for no clear reason. The presence of the propargyl groups could be shifting the alignment of the ketone and *para* fluoro phenyl ring resulting in a loss of the activity. Compound **73** requires retesting alongside positive controls to confirm whether it possesses any antifungal activity and whether it can be used as a probe.

3.4 Conclusions

A number of different PAL probes were synthesised using different photo group functionalities; however, none were found to possess any antifungal activity during MIC testing. Various benzimidazole carbamate analogues displayed poor or no antifungal activity due to the presence of certain polar groups. The diazirine could be interfering with the physicochemical properties of the compound and hence block activity and cell permeability, resulting in a loss of fungal activity. Doubt can be cast of the MIC results for compound **73**, as the SAR data presented in **Chapter 2** suggests that **73** will possess at least some antifungal activity. Repeating the MIC testing for the compounds discussed within this chapter alongside positive controls can determine whether the previous results obtained are accurate, or whether the biological assays are faulty.

Due to these issues, the PAL probes could not be submitted for PAL experiments to determine the binding of benzimidazole carbamates to their targets and determine selectivity within the drug compounds.

The benzimidazole carbamates themselves have experienced multiple issues with their DMPK properties as they possess limited aqueous solubility, resulting in toxicity. The final lead compound **1h** (see **Chapter 2, Table 2.3**) had slightly improved DMPK properties, although had poorer antifungal activity compared to the original flubendazole. Due to these limitations, the benzimidazole carbamate template was abandoned, and attention has shifted to other scaffolds for the improved and novel treatment of *C. neoformans*. High throughput screening (HTS) has been carried out by the group to generate new templates for onward investigation.

3.5 Future Work

3.5.1 Synthesis of Phenyl Azide PAL Probe

Future work may include the synthesis of the phenyl azide PAL probe (see **Figure 3.24**), as phenyl azides have gained a lot of attention and have been frequently published over the years.^{84 85 86 87}

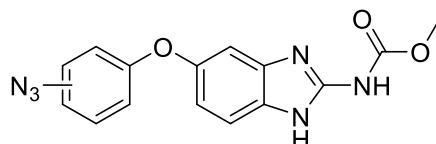


Figure 3.24: Proposed structure of a benzimidazole phenyl azide.

However, due to the benzimidazole project being closed for the lack of efficacy, this synthesis may no longer be necessary.

3.5.2 Photolytic Studies

The photolytic reactivity of **60a** and **60b** can be investigated by forming a probe-methanol adduct. A solution of **60a** and **60b** in methanol can be exposed to UV light to generate the reactive carbene intermediate to produce a methanol photo-adduct; **74a** and **74b** (see **Figure 3.25**). These adduct can be confirmed by mass spectrometry and NMR spectroscopy. These experiments will confirm whether the diazirine probe is behaving as expected, and therefore opens the possibility of the probes being used for the elucidation of benzimidazole binding sites within other diseases.

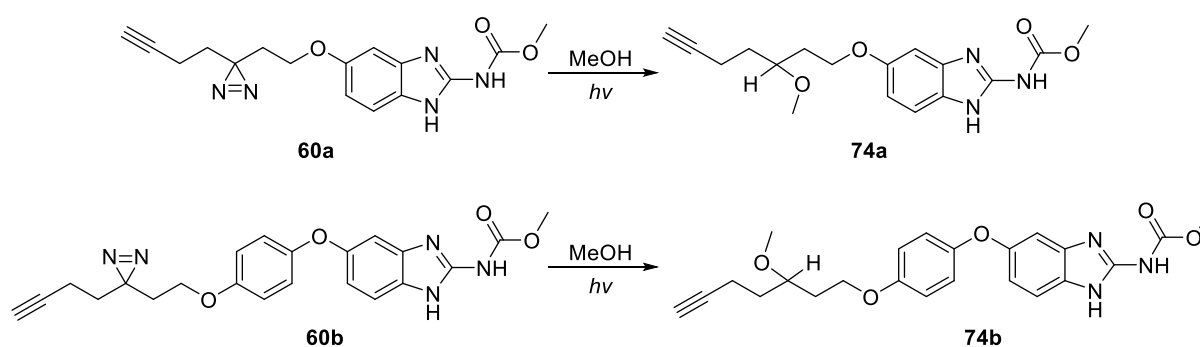


Figure 3.25: Photolytic products of **60a** and **60b** in methanol.

3.5.3 Alternative Scaffold Exploration

Recently, high throughput screening (HTS) conducted by the Nixon group for compounds with activity against *C. neoformans* has highlighted a new lead scaffold. The lead scaffold is a cyclopropyl phenyl ketone derivative, which has been reported to have a MIC value of 0.559

mg/L. This hit compound is currently being synthesised within the group following *Ajay et al* methodology to test whether the reported MIC value can be replicated.⁸⁸

SAR work is required to carry out structural modifications on the metabolically liable hydroxyl group on the structure (see **Figure 3.26**). Fluorine, a metabolically stable substituent could replace the troubling hydroxyl group. The DMPK properties of the original compound will be determined to identify various physicochemical issues such as poor aqueous solubility. Structural modifications by adding various polar groups such as morpholine could help improve the solubility profile (see **Section 2.3.1**).

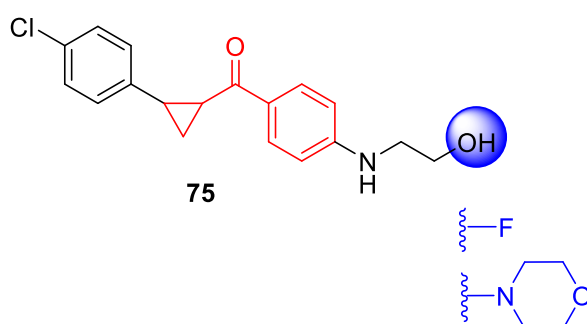


Figure 3.26: New lead antifungal (**75**). Cyclopropyl phenyl ketone shown in red. Structural modification *via* medicinal chemistry shown in blue.

If the template looks promising as a potential antifungal agent against *C. neoformans*, similar strategies for target ID approaches in the form of synthesising a PAL probe can use the same approaches and procedures discussed in this chapter. Some proposed PAL probes based on **75** have been proposed (see **Figure 3.27**).

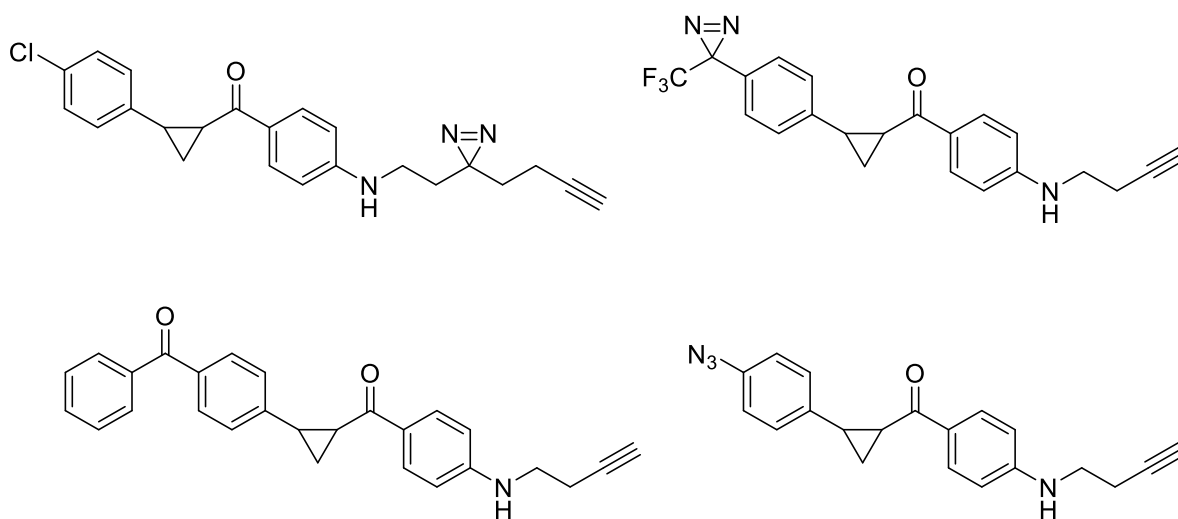


Figure 3.27: Potential PAL probes on the cyclopropyl phenyl ketone derivative template.

3.6 Experimental

3.6.1 General Experimental Details

For general experimental details please see **Section 2.14.1, Chapter 2**.

3.6.1.1 Diazirine Synthesis

For the synthesis and purification of a diazirine containing compound, glassware was covered in aluminium foil and all light sources were blocked to prevent photolysis.

^1H and ^{13}C NMR was performed in a Norell® amberised NMR tube which can protect photosensitive materials from visible and ultraviolet radiation between 650 and 300 nm.

3.6.1.2 General procedures

General Procedure A: Synthesis of 2-Nitro-5-phenoxyaniline derivatives²⁷⁰

5-Fluoro-2-nitroaniline (**7**) (1 equiv.), substituted phenol/thiophenol (**6a-o**) (1 equiv.), K_2CO_3 (2 equiv.) and anhydrous DMF (0.35 M) were added to a flask and allowed to stir at reflux for 16 hours under a nitrogen atmosphere. After this time, the reaction was diluted with H_2O and extracted with EtOAc (x3). The combined organic extracts were washed with brine (x1), dried over MgSO_4 , filtered, and concentrated under reduced pressure. Purification *via* column chromatography eluting with 20-100% EtOAc in *n*-hexane afforded the desired product.

General Procedure B: Tin chloride reduction of a nitro group²⁷¹

The nitro compound (**8a-o**) (1 equiv.), $\text{SnCl}_2 \cdot \text{H}_2\text{O}$ (5 equiv.) and ethanol (0.15 M), were allowed to stir at reflux for 16 hours. After this time, the reaction was concentrated under reduced pressure and basified with aqueous 25% NaOH to a consistent pH 10. The SnCl_2 was removed *via* Buchner filtration and washed with water and EtOAc. The organic layer was separated and washed with water (x3) and brine (x1). The organic phase was concentrated *in vacuo* to yield a crude brown solid/ oil. Purification *via* column chromatography eluting with 30% EtOAc in *n*-hexane afforded the desired product.

General Procedure C: Synthesis of 1,3-Bis(Alkoxy carbonyl)-2-methylisothiurea²⁷⁸

A suspension of *S*-methylisothiurea hemisulfate (**10**) (1 equiv.) and distilled water (1 M) were cooled to 0°C . Alkyl chloroformate (**11a-c**) (4.3 equiv.) was added and allowed to stir vigorously at room temperature for 20 minutes. After this time, the reaction was allowed to warm to room temperature and basified with 25% aqueous NaOH to a consistent pH 10. The

resultant precipitate was isolated *via* Buchner filtration and washed with water to afford the desired product.

General Procedure D: Synthesis of benzimidazole carbamate *via* ring closure²⁷¹

A diamine (**9a-o**) (1 equiv.) and 1,3-bis(alkoxycarbonyl)-2-methylisothiurea (**12a-c**) (1 equiv.) were dissolved in AcOH (2 M) and MeOH (0.4 M) and allowed to stir at reflux for 16 hours in a screw fix sealed tube. After this time, the reaction vessel was cooled to room temperature and diluted with Et₂O. The resultant precipitate was isolated *via* sinter filtration and washed with Et₂O to afford the desired product.

General Procedure E: Synthesis of (Prop-2-yn-1-yloxy)phenol²⁸¹

Benzenediol (**20a-c**) (1 equiv.) was dissolved in anhydrous DMF and cooled to 0°C. NaH (1.3 equiv.) was added portion-wise to the reaction and allowed to stir at room temperature for 10 minutes under a nitrogen atmosphere. Propargyl bromide (**21**) (1 equiv.) was added and allowed to stir at reflux for 16 hours. After this time, the reaction was cooled and concentrated under reduced pressure. The residue was dissolved in EtOAc, washed with water (x1), saturated aq. NaHCO₃ solution (x1) and brine (x1). The organic phase was dried over MgSO₄, filtered, and concentrated *in vacuo* to yield a crude oil. Purification *via* column chromatography eluting with 0-30% EtOAc in *n*-hexane afforded the desired product.

General Procedure F: Diaziridine synthesis²⁶⁰

To a flame dried flask under nitrogen the ketone (**58/ 65**) (1 equiv.) was added and cooled to -78°C. 7N NH₃ in MeOH solution (1 M) was added dropwise and allowed to stir at 40°C for 5 hours under a nitrogen atmosphere. After this time, the reaction was cooled to -78°C and hydroxylamine-*O*-sulfonic acid (1.15 equiv.) in anhydrous MeOH was added dropwise over 15 minutes and allowed to stir at 40°C for a further hour. Upon completion, the solvent was removed under reduced pressure, filtered *via* sinter filtration, washed with MeOH, and concentrated *in vacuo* to afford the desired product.

General Procedure G: Diazirine synthesis³¹⁶

WARNING: During this and following procedures the diazirine must be protected from light by covering the flask and other subsequent glassware in aluminium foil.

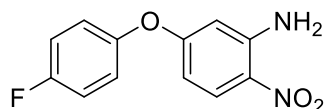
Oxalyl chloride (1.3 equiv.) was added to a stirred solution of DMSO (1.6 equiv.) in DCM (0.04 M) at -78°C under a nitrogen atmosphere. After 5 minutes, a suspension of diaziridine (**59/66**) (1 equiv.) in DCM (0.2 M) was added and allowed to stir at -78°C for 15 min. Triethylamine (5 equiv.) was added, and the mixture was allowed to stir at room temperature for 90 minutes. The reaction was quenched with water and extracted with EtOAc (x3), dried over MgSO_4 , filtered, and concentrated under reduced pressure. The crude product was purified by *via* column chromatography to afford the desired product.

3.6.2 Benzimidazole Carbamate Alkyne Probes

3.6.2.1 Prop-2-yn-1-yl (5-phenoxy/ thio-1*H*-benzo[*d*]imidazol-2-yl)carbamate Analogues

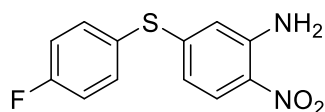
3.6.2.1.1 (Phenoxy/thiol)-2-nitroaniline Derivatives

8a: 5-(4-Fluorophenoxy)-2-nitroaniline³²⁵

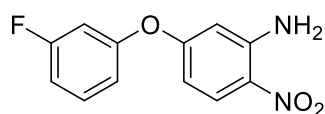


Reaction repeated as general procedure A using 4-fluorophenol (**6a**) gave the title compound (**8a**) as a light orange solid (1.0 g, 90%). ^1H NMR (400 MHz, CDCl_3) δ 8.09 (d, 1H, $J = 9.6$ Hz, Ar-H), 7.07 (m, 4H, Ar-H), 6.35 (br s, 2H, NH_2), 6.29 (dd, 1H, $J = 2.4, 7.2$ Hz, Ar-H), 6.17 (d, 1H, $J = 2.8$ Hz, Ar-H). ^{13}C NMR (101 MHz, CDCl_3) δ 166.3, 162.4, 161.2, 143.4, 128.6, 127.9, 125.7, 124.6, 122.2, 118.6, 118.1, 102.4. HRMS (CI+) m/z calculated for $\text{C}_{12}\text{H}_{10}^{19}\text{FN}_2\text{O}_3$: 249.0677. Found $[\text{M}+\text{H}]^+$: 249.0672 (Diff -2.01 ppm).

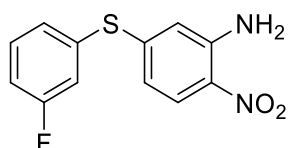
8b: 5-((4-Fluorophenyl)thio)-2-nitroaniline³²⁶



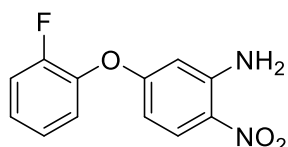
Reaction repeated as general procedure A using 4-fluorothiophenol (**6b**) gave the title compound (**8b**) as a yellow solid (1.1 g, 98%). ^1H NMR (400 MHz, CDCl_3) δ 7.93 (d, 1H, $J = 8.8$ Hz, Ar-H), 7.53 (m, 2H, Ar-H), 7.13 (m, 2H, Ar-H), 6.35 (m, 2H, Ar-H), 6.25 (br s, 2H, NH_2). ^{13}C NMR (101 MHz, CDCl_3) δ 162.4, 147.6, 140.3, 132.8, 132.4, 126.9, 125.7, 120.0, 118.1, 115.6. LRMS (CI+) m/z calculated for $\text{C}_{12}\text{H}_{10}\text{N}_2\text{O}_2^{19}\text{FS}$: 265.0. Found $[\text{M}+\text{H}]^+$: 265.0 (Diff -3.4 ppm).

8c: 5-(3-Fluorophenoxy)-2-nitroaniline²⁷¹

Reaction repeated as general procedure A using 3-fluorophenol (**6c**) gave the title compound (**8c**) as a yellow solid (1.18 g, 99%). ¹H NMR (400 MHz, CDCl₃) δ 8.13 (d, 1H, *J* = 9.2 Hz, Ar-H), 7.36 (m, 1H, Ar-H), 6.94 (m, 1H, Ar-H), 6.88 (dd, 1H, *J* = 2.4, 6.0, 10.4 Hz, Ar-H), 6.82 (dt, 1H, *J* = 2.4, 4.4, 9.6 Hz, Ar-H), 6.34 (dd, 1H, *J* = 2.4, 9.6, 12.0 Hz, Ar-H), 6.23 (d, 1H, *J* = 2.8 Hz, Ar-H), 6.14 (br s, 2H, NH₂). ¹³C NMR (101 MHz, CDCl₃) δ 166.2, 161.3, 159.4, 144.4, 126.8, 126.2, 125.7, 116.5, 110.9, 107.4, 106.9, 102.0. HRMS (CI+ *m/z* calculated for C₁₂H₁₀¹⁹FN₂O₃: 249.0675. Found [M+H]⁺: 249.0681 (Diff 2.41 ppm).

8d: 5-((3-Fluorophenyl)thio)-2-nitroaniline

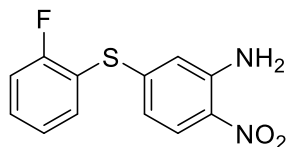
Reaction repeated as general procedure A using 3-fluorothiophenol (**6d**) gave the title compound (**8d**) as a brown oil (2.06 g, 100%). ¹H NMR (400 MHz, CDCl₃) δ 8.01 (d, 1H, *J* = 9.6 Hz, Ar-H), 7.40 (m, 2H, Ar-H), 7.30 (m, 1H, Ar-H), 7.23 (m, 1H, Ar-H), 7.13 (m, 1H, Ar-H), 6.47 (m, 1H, Ar-H), 6.06 (br s, 2H, NH₂). ¹³C NMR (101 MHz, CDCl₃) δ 164.0, 147.6, 144.8, 142.1, 131.2, 131.1, 139.9, 126.9, 121.1, 120.8, 116.0, 115.4. HRMS (CI+) *m/z* calculated for C₁₂H₁₀O₂S₂N₂F₂: 265.0447. Found [M+H]⁺: 265.0450 (Diff 1.13 ppm).

8e: 5-(2-Fluorophenoxy)-2-nitroaniline

Reaction repeated as general procedure A using 2-fluorophenol (**6e**) gave the title compound (**8e**) as a dark-yellow solid (0.95 g, 80%). ¹H NMR (400 MHz, CDCl₃) δ 8.12 (d, 1H, *J* = 9.6 Hz, Ar-H), 7.22 (m, 4H, Ar-H), 6.33 (dd, 1H, *J* = 2.4, 6.8, 9.2 Hz, Ar-H), 6.15 (d, 1H, *J* = 2.4 Hz, Ar-H), 6.12 (br s, 2H, NH₂). ¹³C NMR (101 MHz, CDCl₃) 163.4, 159.0, 153.2, 146.9, 128.6, 126.6, 125.3,

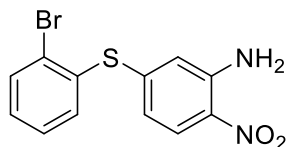
123.3, 117.2, 107.0, 102.6. HRMS (CI+) m/z calculated for $C_{12}H_{10}^{19}FN_2O_2$: 249.0675. Found $[M+H]^+$: 249.0679 (Diff 1.61 ppm).

8f: 5-((2-Fluorophenyl)thio)-2-nitroaniline



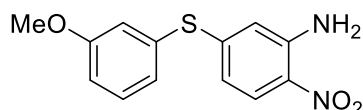
Reaction repeated as general procedure A using 2-fluorothiophenol (**6f**) gave the title compound (**8f**) as a dark yellow solid (1.05 g, 98%). 1H NMR (400 MHz, $CDCl_3$) δ 8.02 (d, 1H, $J=9.2$ Hz, Ar-H), 7.80 (d, 1H, $J=8.8$ Hz, Ar-H), 7.56 (m, 2H, Ar-H), 7.22 (d, 1H, $J=8.8$ Hz, Ar-H), 7.01 (m, 2H, Ar-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.6, 148.5, 141.0, 132.6, 132.4, 127.3, 126.1, 125.2, 121.3, 120.0, 119.8, 110.9. HRMS (CI+) m/z calculated for $C_{12}H_{10}N_2O_2^{19}FS$: 265.0447. Found $[M+H]^+$: 265.0451 (Diff 1.51 ppm).

8g: 5-((2-Bromophenyl)thio)-2-nitroaniline³²⁶



Reaction repeated as general procedure A using 2-bromothiophenol (**6g**) gave the title compound (**8g**) as a yellow solid (0.21 g, 100 %). 1H NMR (400 MHz, $CDCl_3$) δ 8.01 (d, 1H, $J=8.0$ Hz, Ar-H), 7.73 (dd, 1H, $J=4.0, 8.0$ Hz, Ar-H), 7.68 (d, 1H, $J=8.0$ Hz, Ar-H), 7.62 (d, 1H, $J=8.0$ Hz, Ar-H), 7.57 (dd, 1H, $J=4.0, 8.0$ Hz, Ar-H), 7.38-7.36 (t, 1H, $J=8.0, 16.0$ Hz, Ar-H), 7.31-7.27 (t, 1H, $J=8.0, 16.0$ Hz, Ar-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 146.7, 144.9, 136.2, 134.0, 132.4, 130.8, 130.4, 129.2, 128.5, 126.9, 125.5, 116.2. HRMS (CI+) m/z calculated for $C_{12}H_{10}N_2O_2S^{81}Br$: 325.9724. Found $[M+H]^+$: 325.9738 (Diff 4.29 ppm).

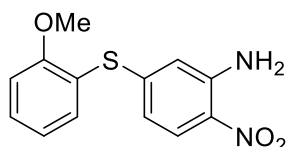
8h: 5-((3-Methoxyphenyl)thio)-2-nitroaniline³²⁶



Reaction repeated as general procedure A using 3-methoxythiophenol (**6h**) gave the title compound (**8h**) as a brown solid (1.42 g, 90%). 1H NMR (400 MHz, $CDCl_3$) δ 7.82 (d, 1H, $J=8.8$

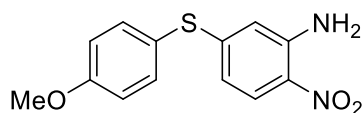
Hz, Ar-H), 7.77 (d, 1H, J = 6.2 Hz, Ar-H), 7.21, (br s, 2H, NH₂), 7.15 (m, 2H, Ar-H), 7.12 (m, 1H, Ar-H), 6.70 (d, 1H, J = 8.8 Hz, Ar-H), 6.40 (m, 1H, Ar-H), 3.82 (s, 3H, CH₃). ¹³C NMR (101 MHz, CDCl₃) δ 165.3, 149.8, 146.5, 144.9, 135.6, 132.7, 127.7, 122.8, 122.5, 120.8, 116.1, 114.1, 56.7. HRMS (CI+) m/z calculated for C₁₃H₁₃O₃N₂S: 277.0647. Found [M+H]⁺: 277.0652 (Diff 1.80 ppm).

8i: 5-((2-Methoxyphenyl)thio)-2-nitroaniline³²⁶

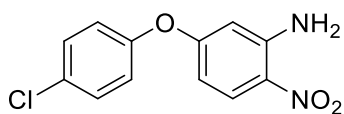


Reaction repeated as general procedure A using 2-methoxythiophenol (**6i**) gave the title compound (**8i**) as a brown solid (1.95 g, 90%). ¹H NMR (400 MHz, CDCl₃) δ 7.96 (d, 1H, J = 8.8 Hz, Ar-H), 7.49 (m, 2H, Ar-H), 7.03 (m, 2H, Ar-H), 6.39 (dd, 1H, J = 2.0, 9.2, 11.2 Hz, Ar-H), 6.33 (d, 1H, J = 1.6 Hz, Ar-H), 6.01 (br s, 2H, NH₂), 3.85 (s, 3H, CH₃). ¹³C NMR (101 MHz, CDCl₃) δ 159.8, 148.8, 144.6, 137.1, 131.8, 126.4, 121.5, 115.3, 113.8, 111.7, 56.7. HRMS (CI+) m/z calculated for C₁₃H₁₃O₃N₂S: 277.0647. Found [M+H]⁺: 277.0637 (Diff -3.61 ppm).

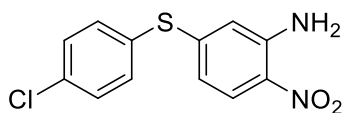
8j: 5-((4-Methoxyphenyl)thio)-2-nitroaniline³²⁶



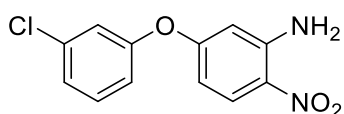
Reaction repeated as general procedure A using 4-methoxythiophenol (**6j**) gave the title compound (**8j**) as a dark yellow solid (1.18 g, 97%). ¹H NMR (400 MHz, CDCl₃) δ 7.92 (d, 1H, J = 9.2 Hz, Ar-H), 7.83 (d, 1H, J = 6.2 Hz, Ar-H), 7.22 (d, 2H, J = 8.8 Hz, Ar-H), 7.19 (d, 1H, J = 8.8 Hz, Ar-H), 7.02 (d, 2H, J = 8.8 Hz, Ar-H), 6.41 (br s, 2H, NH₂), 3.82 (s, 3H, CH₃). ¹³C NMR (101 MHz, CDCl₃) δ 161.2, 143.2, 140.5, 132.6, 131.7, 131.4, 124.2, 122.6, 120.3, 118.0, 114.0, 111.1, 53.5. HRMS (CI+) m/z calculated for C₁₃H₁₃N₂O₃S: 277.0647. Found [M+H]⁺: 277.0644 (Diff -1.08 ppm).

8k: 5-(4-Chlorophenoxy)-2-nitroaniline²⁷¹

Reaction repeated as general procedure A using 4-chlorophenol (**6k**) gave the title compound (**8k**) as a yellow solid (0.76 g, 74%). ¹H NMR (400 MHz, CDCl₃) δ 8.02 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.64 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.44 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.33 (br s, 2H, NH₂), 6.58 (d, 1H, *J* = 12.0 Hz, Ar-H), 6.41-6.39 (m, 1H, Ar-H). ¹³C NMR (101 MHz, CDCl₃) δ 165.7, 154.6, 143.9, 127.0, 126.8, 126.4, 117.1, 107.8, 104.3. HRMS (CI+) *m/z* calculated for C₁₂H₁₀N₂O₃³⁵Cl: 265.0380. Found [M+H]⁺: 265.0385 (Diff 1.89 ppm).

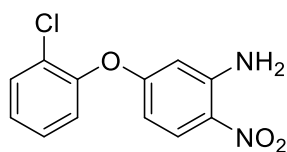
8l: 5-((4-Chlorophenyl)thio)-2-nitroaniline³²⁶

Reaction repeated as general procedure A using 4-chlorothiophenol (**6l**) gave the title compound (**8l**) as a dark orange solid (1.16 g, 94%). ¹H NMR (400 MHz, CDCl₃) δ 7.81 (m, 3H, Ar-H), 7.76 (d, 1H, *J* = 6.2 Hz, Ar-H), 7.38 (br s, 2H, NH₂), 7.29 (d, 2H, *J* = 8.8 Hz, Ar-H), 6.92 (d, 1H, *J* = 8.8 Hz, Ar-H). ¹³C NMR (101 MHz, CDCl₃) δ 152.3, 146.5, 136.5, 130.2, 130.0, 129.7, 128.4, 124.1, 119.9. HRMS (CI+) *m/z* calculated for C₁₂H₁₀N₂O₂S³⁵Cl: 281.0151. Found [M+H]⁺: 281.0159 (Diff 2.85 ppm).

8m: 5-(3-Chlorophenoxy)-2-nitroaniline²⁷¹

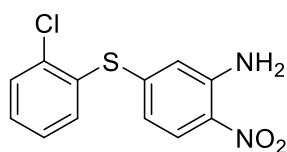
Reaction repeated as general procedure A using 3-chlorophenol (**6m**) gave the title compound (**8m**) as a yellow oil (0.45 g, 55%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.94 (d, 1H, *J* = 12.0 Hz, Ar-H), 7.45 (s, 1H, Ar-H), 7.41 (m, 1H, Ar-H), 7.33 (d, 1H, *J* = 12.0 Hz, Ar-H), 7.02 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.82 (s, 1H, Ar-H), 6.45 (m, 1H, Ar-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 165.2, 160.3, 140.4, 139.5, 132.2, 130.9, 126.9, 120.6, 119.1, 105.0, 104.7. HRMS (CI+) *m/z* calculated for C₁₂H₁₀N₂O₃³⁵Cl: 265.0380. Found [M+H]⁺: 265.0382 (Diff 0.75 ppm).

8n: 5-(2-Chlorophenoxy)-2-nitroaniline²⁷¹



Reaction repeated as general procedure A using 2-chlorophenol (**6n**) gave the title compound (**8n**) as a yellow solid (0.75 g, 73%). ¹H NMR (400 MHz, CDCl₃) δ 7.88 (d, 1H, *J*= 12.0 Hz, Ar-H), 7.52 (m, 2H, Ar-H), 7.46 (br s, 2H, NH₂), 7.31 (m, 1H, Ar-H), 7.16 (m, 1H, Ar-H), 6.94 (s, 1H, Ar-H), 6.45 (d, 1H, *J*= 8.0 Hz, Ar-H). ¹³C NMR (101 MHz, CDCl₃) δ 164.3, 150.2, 148.9, 144.2, 130.7, 128.5, 128.1, 127.8, 122.2, 119.6, 105.0, 104.4. HRMS (Cl⁺) *m/z* calculated for C₁₂H₁₀N₂O₃³⁵Cl: 265.0380. Found [M+H]⁺: 265.0388 (Diff 3.02 ppm).

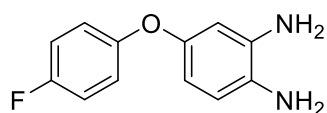
8o: 5-((2-Chlorophenyl)thio)-2-nitroaniline



Reaction repeated as general procedure A using 2-chlorothiophenol (**6o**) gave the title compound (**8o**) as a yellow solid (0.89 g, 92%). ¹H NMR (400 MHz, CDCl₃) δ 7.99 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.75 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.54 (d, 1H, *J*= 6.4 Hz, Ar-H), 7.50 (br s, 2H, NH₂), 7.22 (d, 1H, *J*= 8.8 Hz, Ar-H), 7.01 (m, 2H, Ar-H), 6.88 (d, 1H, *J*= 8.0 Hz, Ar-H). ¹³C NMR (101 MHz, CDCl₃) δ 149.6, 140.0, 136.2, 135.8, 135.7, 134.9, 133.3, 133.0, 130.6, 124.2, 12.09, 119.6. HRMS (Cl⁺) *m/z* calculated for C₁₂H₁₀N₂O₂S³⁵Cl: 281.0151. Found [M+H]⁺: 281.0158 (Diff 2.49 ppm).

3.6.2.1.2 (Phenoxy/thiol)benzene-1,2-diamine Derivatives

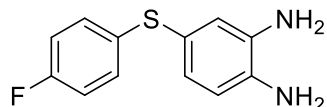
9a: 4-(4-Fluorophenoxy)benzene-1,2-diamine³¹²



Reaction repeated as general procedure B using 5-(4-fluorophenoxy)-2-nitroaniline (**8a**) gave the title compound (**9a**) as a brown solid (0.85 g, 95%). ¹H NMR (400 MHz, CDCl₃) δ 6.94 (m, 4H, Ar-H), 6.66 (d, 1H, *J*= 8.4 Hz, Ar-H), 6.40 (d, 1H, *J*= 2.8 Hz, Ar-H), 6.35 (dd, 1H, *J*= 2.8, 8.4 Hz, Ar-H), 3.39 (br s, 4H, (NH₂)₂). ¹³C NMR (101 MHz, CDCl₃) δ 155.3, 152.7, 142.6, 136.0,

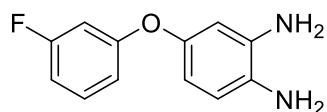
128.0, 119.2, 118.5, 117.6, 116.1, 115.7, 107.6, 103.6. LRMS (Cl⁺) m/z calculated for C₁₂H₁₂¹⁹FN₂O: 219.09. Found [M+H]⁺: 219.09 (Diff -3.66 ppm).

9b: 4-((4-Fluorophenyl)thio)benzene-1,2-diamine³²⁶



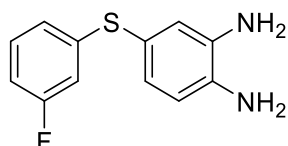
Reaction repeated as general procedure B using 5-((4-fluorophenyl)thio)-2-nitroaniline (**8b**) gave the title compound (**9b**) as a dark brown oil (0.97 g, 99%). ¹H NMR (400 MHz, CDCl₃) δ 7.15 (m, 2H, Ar-H), 6.90 (m, 2H, Ar-H), 6.81 (dd, 1H, *J* = 1.6, 6, 7.6 Hz, Ar-H), 6.76 (d, 1H, *J* = 2.0 Hz, Ar-H), 6.62 (d, 1H, *J* = 8.0 Hz, Ar-H), 3.43 (bs, 4H, (NH₂)₂). ¹³C NMR (101 MHz, CDCl₃) δ 162.6, 137.2, 135.4, 133.8, 130.3, 123.5, 121.6, 117.0, 116.1, 115.8. HRMS (Cl⁺) m/z calculated for C₁₂H₁₂N₂¹⁹FS: 235.0705. Found [M+H]⁺: 235.0706 (Diff 0.43 ppm).

9c: 4-(3-Fluorophenoxy)benzene-1,2-diamine²⁷¹



Reaction repeated as general procedure B using 5-(3-fluorophenoxy)-2-nitroaniline (**8c**) gave the title compound (**9c**) as a crude brown oil (0.77 g, 69 %). ¹H NMR (400 MHz, CDCl₃) δ 7.93 (d, 1H, *J* = 8.8 Hz, Ar-H), 7.35 (m, 1H, Ar-H), 6.92 (m, 1H, Ar-H), 6.85 (m, 1H, Ar-H), 6.79 (m, 1H, Ar-H), 6.65 (dd, 1H, *J* = 2.4, 8.8 Hz, Ar-H), 6.36 (d, 1H, *J* = 2.4 Hz, Ar-H). ¹³C NMR (101 MHz, CDCl₃) δ 159.9, 156.0, 141.4, 131.0, 130.9, 129.3, 115.7, 112.0, 111.6, 111.0, 108.4, 108.0. HRMS (Cl⁺) m/z calculated for C₁₂H₁₂¹⁹FN₂O: 219.0933. Found [M+H]⁺: 219.0926 (Diff -3.19 ppm).

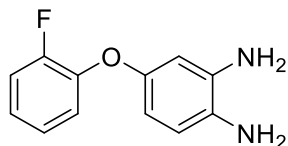
9d: 4-((3-Fluorophenyl)thio)benzene-1,2-diamine



Reaction repeated as general procedure B using 5-((3-fluorophenyl)thio)-2-nitroaniline (**8d**) gave the title compound (**9d**) as a brown oil (1.16 g, 63 %). ¹H NMR (400 MHz, CDCl₃) δ 7.22 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.02 (m, 2H, Ar-H), 6.73 (m, 2H, Ar-H), 6.67 (d, 1H, *J* = 7.6 Hz, Ar-H),

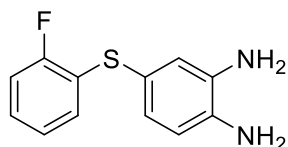
6.50 (d, 1H, $J = 8.0$ Hz, Ar-H), 3.41 (br s, 4H, $(\text{NH}_2)_2$). ^{13}C NMR (101 MHz, CDCl_3) δ 167.2, 140.2, 138.5, 130.2, 129.6, 129.0, 127.4, 120.3, 119.7, 118.1, 115.2, 114.6. LRMS (Cl⁺) m/z calculated for $\text{C}_{12}\text{H}_{12}\text{N}_2^{19}\text{FS}$: 235.1. Found $[\text{M}+\text{H}]^+$: 235.1.

9e: 4-(2-Fluorophenoxy)benzene-1,2-diamine



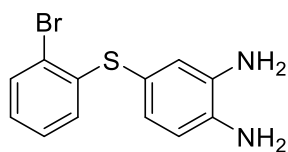
Reaction repeated as general procedure B using 5-(2-fluorophenoxy)-2-nitroaniline (**8e**) gave the title compound (**9e**) as a brown oil (0.18 g, 100%). ^1H NMR (400 MHz, CDCl_3) δ 6.75 (d, 1H, $J = 7.2$ Hz, Ar-H), 6.64 (d, 1H, $J = 8.4$ Hz, Ar-H), 6.56 (d, 1H, $J = 8.4$ Hz, Ar-H), 6.47 (d, 1H, $J = 8.4$ Hz, Ar-H), 6.40 (d, 1H, $J = 7.2$ Hz, Ar-H), 6.36 (m, 2H, Ar-H), 3.50 (br s, 4H, $(\text{NH}_2)_2$). ^{13}C NMR (101 MHz, CDCl_3) δ 139.5, 136.7, 135.5, 133.6, 130.2, 128.7, 124.6, 124.4, 122.8, 117.8, 116.7. LRMS (Cl⁺) m/z calculated for $\text{C}_{12}\text{H}_{12}\text{F}^{19}\text{N}_2\text{O}$: 219.1. Found $[\text{M}+\text{H}]^+$: 219.1.

9f: 4-((2-Fluorophenyl)thio)benzene-1,2-diamine



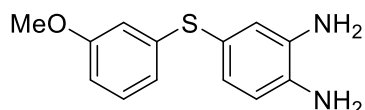
Reaction repeated as general procedure B using 5-((2-fluorophenyl)thio)-2-nitroaniline (**8f**) gave the title compound (**9f**) as a brown oil (0.89 g, 96%). ^1H NMR (400 MHz, CDCl_3) δ 7.12 (m, 2H, Ar-H), 7.06 (d, 1H, $J = 6.2$ Hz, Ar-H), 6.99 (m, 2H, Ar-H), 6.40 (d, 1H, $J = 8.0$ Hz, Ar-H), 6.33 (d, 1H, $J = 8.0$ Hz, Ar-H), 4.76 (br s, 4H, $(\text{NH}_2)_2$). ^{13}C NMR (101 MHz, CDCl_3) δ 166.4, 142.2, 135.1, 134.0, 130.6, 129.9, 125.0, 120.3, 119.4, 118.9, 115.8, 113.1. LRMS (Cl⁺) m/z calculated for $\text{C}_{12}\text{H}_{12}^{19}\text{FN}_2\text{S}$: 235.1. Found $[\text{M}+\text{H}]^+$: 235.1.

9g: 4-((2-Bromophenyl)thio)benzene-1,2-diamine³²⁶



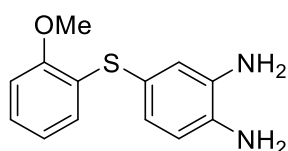
Reaction repeated as general procedure B using 5-((2-bromophenyl)thio)-2-nitroaniline (**8g**) gave the title compound (**9g**) as a brown oil (0.17g, 91%). ¹H NMR (400 MHz, CDCl₃) δ 7.33 (s, 1H, Ar-H), 7.15 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.11 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.06 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.98 (t, 1H, *J* = 8.0, 16.0 Hz, Ar-H), 6.93 (t, 1H, *J* = 8.0, 12.0 Hz, Ar-H), 6.83 (t, 1H, *J* = 8.0, 16.0 Hz, Ar-H). ¹³C NMR (101 MHz, CDCl₃) δ 140.0, 139.4, 134.7, 132.6, 128.8, 128.3, 127.6, 127.0, 126.2, 122.5, 121.9, 121.0. LRMS (CI+) *m/z* calculated for C₁₂H₁₂N₂S⁸¹Br: 294.99. Found [M+H]⁺: 294.99 (Diff -5.2 ppm).

9h: 4-((3-Methoxyphenyl)thio)benzene-1,2-diamine³²⁶



Reaction repeated as general procedure B using 5-((4-methoxyphenyl)thio)-2-nitroaniline (**8h**) gave the title compound (**9h**) as a brown oil (0.8 g, 63%). ¹H NMR (400 MHz, CDCl₃) δ 7.17 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.99 (m, 2H, Ar-H), 6.86 (d, 1H, *J* = 8.8 Hz, Ar-H), 6.30 (m, 2H, Ar-H), 6.21 (d, 1H, *J* = 8.0 Hz, Ar-H), 4.74 (br s, 4H, (NH₂)₂), 3.82 (s, 3H, CH₃). ¹³C NMR (101 MHz, CDCl₃) δ 160.9, 138.2, 136.0, 135.7, 134.1, 130.8, 125.4, 120.7, 120.4, 118.3, 116.9, 111.3, 58.1. HRMS (CI+) *m/z* calculated for C₁₃H₁₅N₂OS: 247.0905. Found [M+H]⁺: 247.0903 (Diff -0.81 ppm).

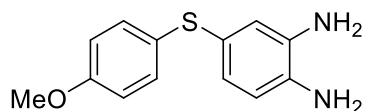
9i: 4-((2-Methoxyphenyl)thio)benzene-1,2-diamine³²⁶



Reaction repeated as general procedure B using 5-((2-methoxyphenyl)thio)-2-nitroaniline (**8i**) gave the title compound (**9i**) as a brown oil (0.79 g, 45%). ¹H NMR (400 MHz, CDCl₃) δ 7.35 (t, 1H, *J* = 2.4, 7.6, 12.4 Hz, Ar-H), 7.23 (m, 2H, Ar-H), 7.11 (d, 2H, *J* = 8.0 Hz, Ar-H), 6.51 (m, 2H, Ar-

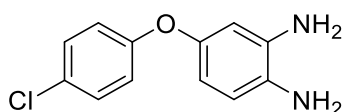
H), 4.76 (br s, 4H, (NH₂)₂), 3.83 (s, 3H, CH₃). ¹³C NMR (101 MHz, CDCl₃) δ 154.2, 139.1, 136.5, 129.6, 129.4, 125.7, 123.3, 120.6, 117.9, 117.4, 113.0, 55.2. HRMS (CI+) m/z calculated for C₁₃H₁₅N₂OS: 247.0905. Found [M+H]⁺: 247.0912 (Diff 2.83 ppm).

9j: 4-((4-Methoxyphenyl)thio)benzene-1,2-diamine³²⁶



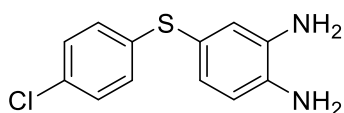
Reaction repeated as general procedure B using 5-((4-methoxyphenyl)thio)-2-nitroaniline (**8j**) gave the title compound (**9j**) as a brown oil (0.65 g, 62%). ¹H NMR (400 MHz, CDCl₃) δ 7.44 (d, 2H, *J* = 8.8 Hz, Ar-H), 7.23 (d, 1H, *J* = 7.2 Hz, Ar-H), 7.02 (d, 2H, *J* = 8.8 Hz, Ar-H), 6.44 (m, 2H, Ar-H), 4.94 (br s, 4H, (NH₂)₂), 3.82 (s, 3H, CH₃). ¹³C NMR (101 MHz, CDCl₃) δ 162.0, 140.1, 134.4, 129.6, 128.5, 122.3, 117.8, 117.4, 116.6, 112.1, 54.3. HRMS (CI+) m/z calculated for C₁₃H₁₅N₂OS: 247.0905. Found [M+H]⁺: 247.0911 (Diff 2.43 ppm).

9k: 4-(4-Chlorophenoxy)benzene-1,2-diamine²⁷¹



Reaction repeated as general procedure B using 5-(4-chlorophenoxy)-2-nitroaniline (**8k**) gave the title compound (**9k**) as a brown oil (0.38 g, 57%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.65 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.48 (d, 2H, *J* = 8.0 Hz, Ar-H), 6.74 (d, 1H, *J* = 12.0 Hz, Ar-H), 6.41 (m, 1H, Ar-H), 6.22 (m, 1H, Ar-H). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 154.2, 148.6, 130.5, 130.1, 129.8, 126.6, 117.0, 116.6, 105.8, 104.1. HRMS (CI+) m/z calculated for C₁₂H₁₂N₂O³⁵Cl: 235.0638. Found [M+H]⁺: 235.0644 (Diff 2.55 ppm).

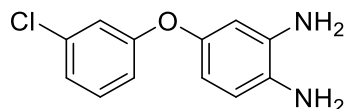
9l: 4-((4-Chlorophenyl)thio)benzene-1,2-diamine³²⁶



Reaction repeated as general procedure B using 5-((4-chlorophenyl)thio)-2-nitroaniline (**8l**) gave the title compound (**9l**) as a brown oil (0.69 g, 66%). ¹H NMR (400 MHz, CDCl₃) δ 7.81 (d, 2H *J* = 8.8 Hz, Ar-H), 7.13 (m, 3H, Ar-H), 6.40 (m, 2H, Ar-H), 4.71 (br s, 4H, (NH₂)₂). ¹³C NMR

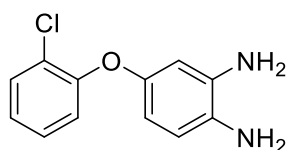
(101 MHz, CDCl₃) δ 140.2, 136.8, 130.1, 129.8, 128.4, 127.7, 124.3, 117.2, 117.0, 115.1. HRMS (Cl⁺) *m/z* calculated for C₁₂H₁₂N₂S³⁵Cl: 251.0409. Found [M+H]⁺: 251.0402 (Diff -2.79 ppm).

9m: 4-(3-Chlorophenoxy)benzene-1,2-diamine²⁷¹



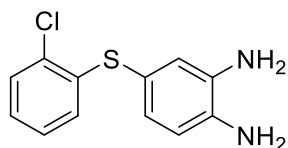
Reaction repeated as general procedure B using 5-(3-chlorophenoxy)-2-nitroaniline (**6m**) gave the title compound (**9m**) as a brown oil (0.25 g, 63%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.44 (s, 1H, Ar-H), 7.40 (m, 1H, Ar-H), 7.08 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.89 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.84 (m, 1H, Ar-H), 6.66 (m, 1H, Ar-H), 6.21 (s, 1H, Ar-H). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 160.2, 150.1, 143.6, 134.8, 130.0, 139.7, 120.3, 119.9, 118.2, 115.4, 115.2, 107.5, 104.7. LRMS (Cl⁺) *m/z* calculated for C₁₂H₁₂N₂O³⁵Cl: 235.06. Found [M+H]⁺: 235.06.

9n: 4-(2-Chlorophenoxy)benzene-1,2-diamine²⁷¹



Reaction repeated as general procedure B using 5-(2-chlorophenoxy)-2-nitroaniline (**8n**) gave the title compound (**9n**) as a brown oil (0.35 g, 52%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.45-7.39 (m, 2H, Ar-H), 7.02 (m, 1H, Ar-H), 6.67 (m, 1H, Ar-H), 6.49 (m, 1H, Ar-H), 6.31 (s, 1H, Ar-H), 6.21 (d, 1H, *J* = 8.0 Hz, Ar-H). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 150.4, 150.2, 144.5, 128.4, 127.4, 127.0, 125.9, 122.3, 121.7, 117.8, 113.5, 104.0. HRMS (Cl⁺) *m/z* calculated for C₁₂H₁₂N₂O³⁵Cl: 235.0638. Found [M+H]⁺: 235.0644 (Diff 2.55 ppm).

9o: 4-((2-Chlorophenyl)thio)benzene-1,2-diamine

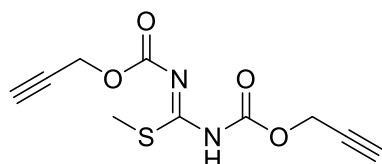


Reaction repeated as general procedure B using 5-((2-chlorophenyl)thio)-2-nitroaniline (**8o**) gave the title compound (**9o**) as a brown oil (0.57 g, 72%). ¹H NMR (400 MHz, CDCl₃) δ 7.88 (d, 1H, *J* = 8.8 Hz, Ar-H), 7.61 (d, 1H, *J* = 6.4 Hz, Ar-H), 7.55 (d, 1H, *J* = 8.8 Hz, Ar-H), 6.86 (m, 2H,

Ar-H), 6.44 (d, 1H, $J = 7.2$ Hz, Ar-H), 6.23 (m, 1H, Ar-H), 4.65 (br s, 4H, $(\text{NH}_2)_2$). ^{13}C NMR (101 MHz, CDCl_3) δ 143.7, 138.6, 135.2, 134.7, 130.6, 130.4, 124.7, 124.0, 122.2, 116.9, 116.8, 115.6. HRMS (CI+) m/z calculated for $\text{C}_{12}\text{H}_{12}\text{N}_2\text{S}^{35}\text{Cl}$: 251.0409. Found $[\text{M}+\text{H}]^+$: 251.0402 (Diff -2.79 ppm).

3.6.2.1.3 Methylisothiurea Derivative.

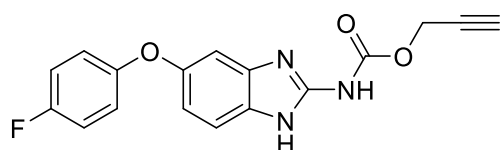
12a: 1,3-Bis(Prop-2-yn-1-yloxycarbonyl)-2-methylisothiurea



Reaction repeated as general procedure C using propargyl chloroformate (**11a**) gave the title compound (**12a**) as a white solid (1 g, 66%). This intermediate was not characterised and carried through to the subsequent reaction.

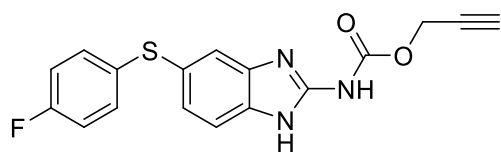
3.6.2.1.4 Prop-2-yn-1-yl(phenoxy/thio)-1H-benzo[d]imidazol-2-yl)carbamate Derivatives

13a: Prop-2-yn-1-yl (5-(4-fluorophenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



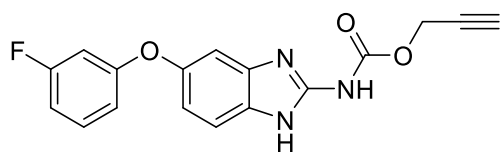
Reaction repeated as general procedure D using 4-(4-fluorophenoxy)benzene-1,2-diamine (**9a**) and 1,3-bis(prop-2-yn-1-yloxycarbonyl)-2-methylisothiurea (**12a**) gave the title compound (**13a**) as an off-white solid (0.27 g, 79%). ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 7.89 (d, 1H, $J = 8.4$ Hz, Ar-H), 7.19 (m, 2H, Ar-H), 7.01 (m, 3H, Ar-H), 6.81 (dd, 1H, $J = 2.4, 6, 8.4$ Hz, Ar-H), 4.82 (m, 2H, CH_2), 3.60 (m, 1H, C-H). ^{13}C NMR (101 MHz, $\text{DMSO}-d_6$) δ 157.0, 154.8, 151.9, 148.9, 139.3, 127.1, 119.7, 117.0, 116.6, 113.4, 105.5, 104.2, 79.2, 77.9, 53.1. HRMS (ES+) m/z calculated for $\text{C}_{17}\text{H}_{13}\text{N}_3\text{O}_3^{19}\text{F}$: 326.0941. Found $[\text{M}+\text{H}]^+$: 326.0938 (Diff -0.92 ppm). IR $\nu_{\text{max}}/\text{cm}^{-1}$ (solid) 3333 (s), 1588 (s), 1262 (s), 1203 (m), 1080 (m). MP: 236 – 238 °C decomposed. Purity HPLC 97.5%, $R_t = 8.6$ min.

13b: Prop-2-yn-1-yl (5-((4-fluorophenyl)thio)-1H-benzo[d]imidazol-2-yl)carbamate



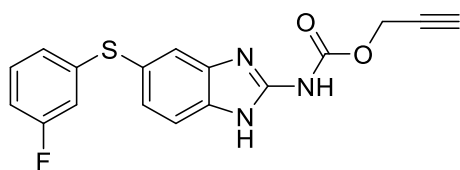
Reaction repeated as general procedure D using 4-((4-fluorophenyl)thio)benzene-1,2-diamine (**9b**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea (**12a**) gave the title compound (**13b**) as an off-white solid (0.48 g, 83%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.83 (br s, 2H, N-H), 7.47 (m, 1H, Ar-H), 7.45 (d, 1H, *J* = 8.4 Hz, Ar-H), 7.21 (m, 5H, Ar-H), 4.84 (s, 2H CH₂), 3.61 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 159.9, 151.1, 149.3, 140.8, 133.4, 131.1, 126.4, 126.0, 124.9, 117.4, 116.6, 116.5, 115.0, 79.0, 78.5, 53.4. HRMS (ES+) *m/z* calculated for C₁₇H₁₃N₃O₂¹⁹F: 342.0712. Found [M+H]⁺: 342.0716 (Diff 1.17 ppm). IR ν_{max}/cm⁻¹ (solid) 3337 (s), 1582 (s), 1270 (s), 1162 (m), 1106 (m), 1022 (s), 613 (m). MP: 246-247°C decomposed. Purity HPLC 96.1%, R_t = 9.6 min.

13c: Prop-2-yn-1-yl (5-(3-fluorophenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



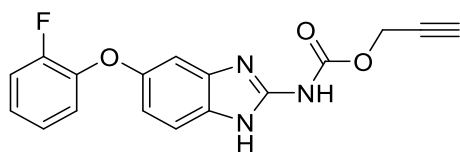
Reaction repeated as general procedure D using 4-(3-fluorophenoxy)benzene-1,2-diamine (**9c**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea (**12a**) gave the title compound (**13c**) as an off-white solid (7.2 mg, 5%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.79 (m, 2H, Ar-H), 7.68 (d, 1H, *J* = 8.8 Hz, Ar-H), 7.33 (m, 1H, Ar-H), 6.88 (m, 3H, Ar-H), 4.81 (m, 2H, CH₂), 3.42 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 167.8, 160.2, 156.2, 152.0, 150.3, 139.7, 133.5, 130.6, 116.3, 116.1, 114.7, 110.0, 109.7, 103.5, 77.8, 75.8, 53.2. HRMS (ES+) *m/z* calculated for C₁₇H₁₃N₃O₃¹⁹F: 326.0941. Found [M+H]⁺: 326.0944 (Diff 0.92 ppm). IR ν_{max}/cm⁻¹ (solid) 3336 (s), 1560 (s), 1255 (s), 1198 (s), 1097 (m). MP: 226-229 °C decomposed. Purity HPLC 90.4%, R_t = 8.9 min.

13d: Prop-2-yn-1-yl (5-((3-fluorophenyl)thio)-1H-benzo[d]imidazol-2-yl)carbamate



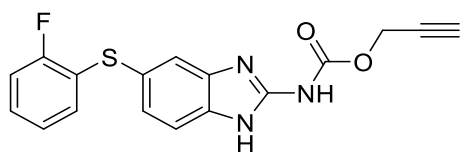
Reaction repeated as general procedure D using 4-((3-fluorophenyl)thio)benzene-1,2-diamine (**9d**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea (**12a**) gave the title compound (**13d**) as an off-white solid (1.26 g, 79%). ¹H NMR (400 MHz, DMSO-*d*6) δ 7.74 (d, 1H, *J*= 8.4 Hz, Ar-H), 7.41 (d, 1H, *J*= 7.2 Hz, Ar-H), 7.06 (m, 3H, Ar-H), 6.74 (m, 2H, Ar-H), 4.83 (m, 2H, CH₂), 3.43 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*6) δ 160.2, 152.4, 149.5, 140.7, 140.1, 135.6, 131.8, 125.3, 125.0, 124.0, 118.7, 115.7, 115.3, 113.6, 78.4, 77.6, 53.6. LRMS (ES+) *m/z* calculated for C₁₇H₁₃N₃O₂¹⁹F: 342.07. Found [M+H]⁺: 342.07 (Diff 0.2 ppm). IR $\nu_{\max}/\text{cm}^{-1}$ (solid) 3335 (s), 1591 (s), 1277 (s), 1159 (s), 1112 (m), 1028 (s), 615 (m). MP: 249-251 °C decomposed. Purity HPLC 85.3%, *R*_t= 9.6 min.

13e: Prop-2-yn-1-yl (5-(2-fluorophenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



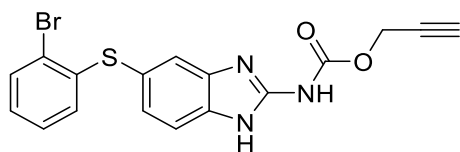
Reaction repeated as general procedure D using 4-(2-fluorophenoxy)benzene-1,2-diamine (**9e**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea (**12a**) gave the title compound (**13e**) as an off-white solid (0.16 g, 59 %). ¹H NMR (400 MHz, DMSO-*d*6) δ 11.76 (br s, 2H, N-H), 7.39 (d, 2H, *J*= 8.4 Hz, Ar-H), 7.17 (m, 2H, Ar-H), 7.05 (m, 1H, Ar-H), 7.01 (d, 1H, *J*= 2.4 Hz, Ar-H), 6.83 (dd, 1H, *J*= 2.4, 8.8 Hz, Ar-H), 4.82 (d, 2H, *J*= 2.4 Hz, CH₂), 3.59 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*6) δ 154.9, 152.3, 151.7, 144.9, 144.3, 125.6, 124.8, 121.1, 117.6, 117.4, 114.2, 112.6, 79.3, 78.3, 53.1. HRMS (ES+) *m/z* calculated for C₁₇H₁₃N₃O₃¹⁹F: 326.0941. Found [M+H]⁺: 326.0940 (Diff -0.31 ppm). IR $\nu_{\max}/\text{cm}^{-1}$ (solid) 3329 (s), 1577 (s), 1258 (s), 1186 (s), 1086 (m). MP: 239 – 242 °C decomposed. Purity HPLC 97.5%, *R*_t= 8.6 min.

13f: Prop-2-yn-1-yl (5-((2-fluorophenyl)thio)-1H-benzo[d]imidazol-2-yl)carbamate



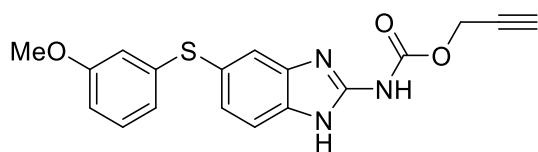
Reaction repeated as general procedure D using 4-((2-fluorophenyl)thio)benzene-1,2-diamine (**9f**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiurea (**12a**) gave the title compound (**13f**) as an off-white solid (1.05 g, 81%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.92 (br s, 2H, N-H), 7.72 (d, 1H, *J* = 8.4 Hz, Ar-H), 7.43 (m, 3H, Ar-H), 7.31 (d, 1H, *J* = 8.4 Hz, Ar-H), 6.98 (m, 2H, Ar-H), 4.81 (m, 2H, CH₂), 3.40 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 164.4, 155.1, 148.3, 140.0, 135.8, 131.2, 129.6, 129.4, 127.9, 124.2, 122.3, 116.8, 116.2, 115.7, 78.5, 75.8, 54.6. HRMS (ES⁺) *m/z* calculated for C₁₇H₁₃N₃O₂¹⁹F: 342.0712. Found [M+H]⁺: 342.0706 (Diff -1.75 ppm). IR *v*_{max}/cm⁻¹ (solid) 3328 (s), 1595 (s), 1262 (s), 1157 (s), 1118 (m), 1031 (s), 613 (m). MP: 252-254 °C decomposed. Purity HPLC 98.5%, R_t = 9.6 min.

13g: Prop-2-yn-1-yl (5-((2-bromophenyl)thio)-1H-benzo[d]imidazol-2-yl)carbamate



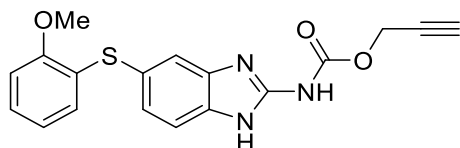
Reaction repeated as general procedure D using 4-((2-bromophenyl)thio)benzene-1,2-diamine (**9g**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiurea (**12a**) gave the title compound (**13g**) as an off-white solid (0.11 g, 48%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.61 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.57 (m, 1H, Ar-H), 7.52 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.25 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.19 (t, 1H, *J* = 8.0, 12.0 Hz, Ar-H), 7.05 (t, 1H, *J* = 8.0, 12.0 Hz, Ar-H), 6.89 (d, 1H, *J* = 8.0 Hz, Ar-H), 4.84 (s, 2H, CH₂), 3.61 (s, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 153.9, 149.3, 139.8, 137.9, 135.4, 133.4, 132.4, 132.0, 128.1, 126.4, 125.0, 124.2, 120.7, 116.9, 78.4, 72.8, 58.0. LRMS (ES⁺) *m/z* calculated for C₁₇H₁₃N₃SO₂⁸¹Br: 401.99. Found [M+H]⁺: 401.99. IR *v*_{max}/cm⁻¹ 3330 (s), 1583 (s), 1269 (s), 1211 (m), 1084 (m), 615 (m). MP: 231-236 °C decomposed. Purity HPLC 84.4%, R_t = 9.7 min.

13h: Prop-2-yn-1-yl (5-((3-methoxyphenyl)thio)-1H-benzo[d]imidazol-2-yl)carbamate



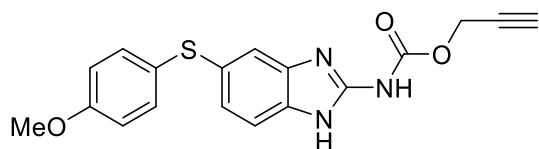
Reaction repeated as general procedure D using 4-((3-methoxyphenyl)thio)benzene-1,2-diamine (**9h**) and 1,3-bis(prop-2-yn-1-yloxy carbonyl)-2-methylisothiurea (**12a**) gave the title compound (**13h**) as an off-white solid (0.71 g, 62%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.94 (br s, 2H, N-H), 7.77 (d, 1H, *J* = 8.4 Hz, Ar-H), 7.40 (d, 1H, *J* = 8.4 Hz, Ar-H), 7.13 (m, 3H, Ar-H), 7.01 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.95 (d, 1H, *J* = 8.0 Hz, Ar-H), 4.82 (m, 2H, CH₂), 3.81 (s, 3H, CH₃), 3.38 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 162.3, 155.6, 149.9, 139.5, 139.3, 136.4, 128.0, 127.4, 126.2, 126.1, 119.7, 116.3, 115.9, 111.6, 78.9, 75.8, 55.3, 52.8. HRMS (ES+) *m/z* calculated for C₁₈H₁₆N₃O₃S: 354.0912. Found [M+H]⁺: 354.0910 (Diff -0.56 ppm). IR ν_{max}/cm⁻¹ (solid) 3333 (s), 1596 (s), 1273 (s), 1155 (s), 1127 (m), 617 (m). MP: 243-247°C. decomposed. Purity HPLC 96.1%, R_t = 8.6 min.

13i: Prop-2-yn-1-yl (5-((2-methoxyphenyl)thio)-1H-benzo[d]imidazol-2-yl)carbamate



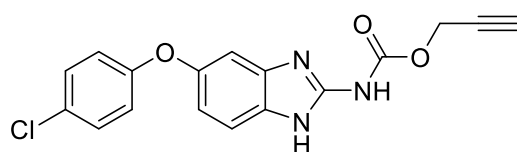
Reaction repeated as general procedure D using 4-((2-methoxyphenyl)thio)benzene-1,2-diamine (**9i**) and 1,3-bis(prop-2-yn-1-yloxy carbonyl)-2-methylisothiurea (**12a**) gave the title compound (**13i**) as an off-white solid (0.92 g, 68%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.98 (br s, 2H, N-H), 7.78 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.41 (m, 2H, Ar-H), 7.26 (m, 3H, Ar-H), 6.98 (m, 1H, Ar-H), 4.92 (m, 2H, CH₂), 3.81 (s, 3H, CH₃), 3.39 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 159.4, 155.6, 149.2, 138.2, 138.1, 131.0, 129.7, 126.5, 126.3, 122.8, 121.6, 119.1, 118.4, 113.2, 78.9, 75.7, 55.3, 53.1. HRMS (ES+) *m/z* calculated for C₁₈H₁₆N₃O₃S: 354.0912. Found [M+H]⁺: 354.0908 (Diff -1.13 ppm). IR ν_{max}/cm⁻¹ (solid) 3342 (s), 1610 (s), 1281 (s), 1158 (s), 1174 (m), 621 (m). MP: 253-256°C decomposed. Purity HPLC 95.9%, R_t = 8.4 min.

13j: Prop-2-yn-1-yl (5-((4-methoxyphenyl)thio)-1H-benzo[d]imidazol-2-yl)carbamate



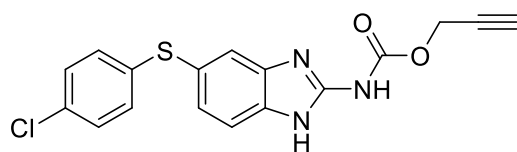
Reaction repeated as general procedure D using 4-((4-methoxyphenyl)thio)benzene-1,2-diamine (**9j**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea (**12a**) gave the title compound (**13j**) as an off-white solid (0.77 g, 83%). ¹H NMR (400MHz, DMSO-*d*₆) δ 12.02 (br s, 2H, N-H), 7.66 (s, 1H, Ar-H), 7.41 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.39 (d, 1H, *J*= 8.4 Hz, Ar-H), 7.18 (m, 3H, Ar-H), 4.83 (m, 2H, CH₂), 3.83 (s, 3H, CH₃), 3.31 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 160.1, 155.6, 147.8, 138.2, 134.6, 129.6, 127.1, 124.5, 123.9, 118.1, 116.0, 113.8, 78.1, 76.5, 55.2, 54.8. HRMS (ES+) *m/z* calculated for C₁₈H₁₆N₃O₃S: 354.0912. Found [M+H]⁺: 354.0911 (Diff -0.28 ppm). IR *v*_{max}/cm⁻¹ (solid) 3334 (s), 1591 (s), 1272 (s), 1160 (s), 1123 (m), 618 (m). MP: 247-251 °C decomposed. Purity HPLC 93.7%, R_t= 8.8 min.

13k: Prop-2-yn-1-yl (5-(4-chlorophenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



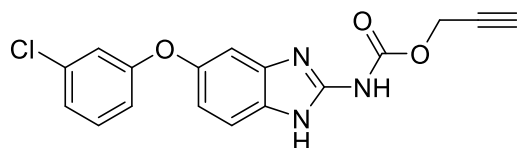
Reaction repeated as general procedure D using 4-(4-chlorophenoxy)benzene-1,2-diamine (**9k**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea (**12a**) gave the title compound (**13k**) as a pale-yellow solid (0.39 g, 71%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.64 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.50 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.68 (m, 1H, Ar-H), 7.22 (m, 1H, Ar-H), 6.98 (m, 1H, Ar-H), 4.71 (m, 2H, CH₂), 3.43 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 156.1, 155.7, 153.2, 150.2, 140.3, 130.6, 128.7, 126.5, 117.0, 116.3, 116.0, 105.8, 79.6, 75.3, 54.1. LRMS (ES+) *m/z* calculated for C₁₇H₁₃N₃O₃³⁵Cl: 342.07. Found [M+H]⁺: 342.07 (Diff -0.88 ppm). IR *v*_{max}/cm⁻¹ (solid) 3329 (s), 1596 (s), 1277 (s), 1154 (m), 1033 (s) 1021 (m), 811 (m). MP: Decomposed at 232-236 °C decomposed. Purity HPLC 82.1%, R_t= 9.7 min.

13l: Prop-2-yn-1-yl (5-((4-chlorophenyl)thio)-1H-benzo[d]imidazol-2-yl)carbamate



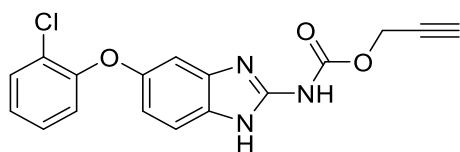
Reaction repeated as general procedure D using 4-((4-chlorophenyl)thio)benzene-1,2-diamine (**9l**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea (**12a**) gave the title compound (**13l**) as an off-white solid (0.49 g, 50%). ¹H NMR (400 MHz, DMSO-*d*6) δ 11.96 (br s, 2H, N-H), 7.81 (d, 2H, *J* = 8.4 Hz, Ar-H), 7.61 (s, 1H, Ar-H), 7.41 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.24 (d, 2H, *J* = 8.4 Hz, Ar-H), 7.18 (d, 1H, *J* = 8.4 Hz, Ar-H), 4.85 (m, 2H, CH₂), 3.73 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*6) δ 155.9, 145.8, 137.2, 136.6, 134.2, 133.7, 133.5, 133.4, 127.3, 123.0, 122.4, 120.2, 117.6, 78.5, 76.4, 53.1. HRMS (ES+) *m/z* calculated for C₁₇H₁₃N₃O₂S³⁵Cl: 258.0417. Found [M+H]⁺: 258.0414 (Diff -1.16 ppm). IR ν_{max}/cm⁻¹ (solid) 3438 (s), 1623 (s), 1295 (s), 1163 (s), 1114 (s), 795 (m), 616 (m). MP: 236-238°C decomposed. Purity HPLC 84.9%, R_t = 10.3 min.

13m: Prop-2-yn-1-yl (5-(3-chlorophenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



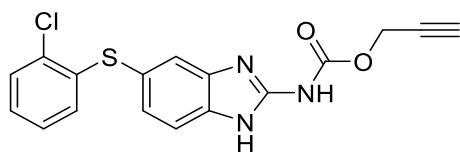
Reaction repeated as general procedure D using 4-(3-chlorophenoxy)benzene-1,2-diamine (**9m**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea (**12a**) gave the title compound (**13m**) as an orange solid (0.13 g, 36%). ¹H NMR (400 MHz, DMSO-*d*6) δ 7.65 (d, 1H, *J* = 12.0 Hz, Ar-H), 7.46 (s, 1H, Ar-H), 7.38 (m, 1H, Ar-H), 7.22 (s, 1H, Ar-H), 7.12 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.92 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.81 (m, 1H, Ar-H), 4.84 (m, 2H, CH₂), 3.32 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*6) δ 162.3, 150.4, 150.1, 146.5, 138.8, 123.1, 130.2, 130.0, 129.6, 120.3, 119.7, 115.6, 115.0, 107.1, 79.7, 75.9, 54.3. HRMS (ES+) *m/z* calculated for C₁₇H₁₃N₃O₃³⁵Cl: 342.0645. Found [M+H]⁺: 342.0639 (Diff -1.75 ppm). IR ν_{max}/cm⁻¹ (solid) 3337 (s), 1591 (s), 1269 (s), 1166 (m), 1021 (m), 813 (m). MP: 237-240 °C decomposed. Purity HPLC 92.4%, R_t = 9.3 min.

13n: Prop-2-yn-1-yl (5-(2-chlorophenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



Reaction repeated as general procedure D using 4-(2-chlorophenoxy)benzene-1,2-diamine (**9n**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea (**12a**) gave the title compound (**13n**) as a light brown solid (0.16 g, 31%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.66 (d, 1H, *J* = 12.0 Hz, Ar-H), 7.43-7.38 (m, 2H, Ar-H), 7.25 (s, 1H, Ar-H), 7.00 (m, 1H, Ar-H), 6.89 (m, 1H, Ar-H), 6.43 (d, 1H, *J* = 8.0 Hz, Ar-H), 4.88 (m, 2H, CH₂), 3.32 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 155.5, 152.8, 151.2, 150.3, 142.1, 133.9, 130.7, 125.4, 125.0, 124.9, 116.5, 110.6, 106.5, 80.8, 74.9, 55.4. LRMS (ES+) *m/z* calculated for C₁₇H₁₃N₃O₃³⁵Cl: 342.07. Found [M+H]⁺: 342.07 (Diff -1.59 ppm). IR ν_{\max} /cm⁻¹ (solid) 3328 (s), 1595 (s), 1263 (s), 1157 (m), 1116 (m), 807 (m). MP: 234-237 °C decomposed. Purity HPLC 83.3%, R_t = 9.4 min.

13o: Prop-2-yn-1-yl (5-((2-chlorophenyl)thio)-1H-benzo[d]imidazol-2-yl)carbamate

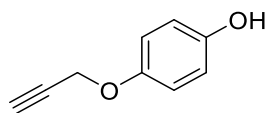


Reaction repeated as general procedure D using 4-((2-chlorophenyl)thio)benzene-1,2-diamine (**9o**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea (**12a**) gave the title compound (**13o**) as an off-white solid (0.31 g, 52%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.89 (br s, 2H, N-H), 7.73 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.59 (m, 1H, Ar-H), 7.40 (m, 2H, Ar-H), 7.18 (m, 3H, Ar-H), 4.84 (m, 2H, CH₂), 3.33 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 154.2, 149.6, 136.5, 136.2, 132.8, 132.7, 132.3, 129.4, 126.3, 126.0, 125.7, 118.5, 116.1, 78.4, 76.2, 55.3. HRMS (ES+) *m/z* calculated for C₁₇H₁₃N₃O₂S³⁵Cl: 258.0417. Found [M+H]⁺: 258.0411 (Diff -2.33 ppm). IR ν_{\max} /cm⁻¹ (solid) 3411 (s), 1628 (s), 1321 (s), 1162 (s), 1103 (s), 811 (m), 611 (m). MP: 257-260°C decomposed. Purity HPLC 90.8%, R_t = 10.2 min.

3.6.2.2 (5-((4-(Prop-2-yn-1-yloxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate Analogues

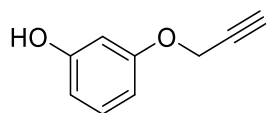
3.6.2.2.1(Prop-2-yn-1-yloxy)phenol Derivatives

22a: 4-(Prop-2-yn-1-yloxy)phenol³²⁷



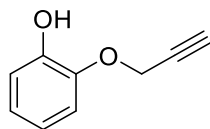
Reaction repeated as general procedure E using hydroquinone (**20a**) gave the title compound (**22a**) as a crude yellow oil (0.42 g, 31 %). ¹H NMR (400 MHz, CDCl₃) δ 6.83 (d, 2H, *J* = 9.2 Hz, Ar-H), 6.76 (d, 2H, *J* = 9.2 Hz, Ar-H), 6.24 (br s, 1H, O-H), 4.59 (m, 2H, CH₂), 2.50 (m, 1H, C-H). ¹³C NMR (101 MHz, CDCl₃) δ 151.6, 150.3, 116.5, 16.2, 78.9, 75.9, 56.9. HRMS (CI+) *m/z* calculated for C₉H₉O₂: 149.0604. Found [M+H]⁺: 149.0603 (Diff -0.67 ppm).

22b: 3-(Prop-2-yn-1-yloxy)phenol³²⁸



Reaction repeated as general procedure E using resorcinol (**20b**) gave the title compound (**22b**) as a crude yellow oil (0.59 g, 44 %). ¹H NMR (400 MHz, CDCl₃) δ 7.93 (t, 1H, *J* = 2.2, 7.6, 12.2 Hz, Ar-H), 6.67 (d, 1H, *J* = 7.2 Hz, Ar-H), 6.52 (d, 1H, *J* = 7.2 Hz, Ar-H), 6.43 (s, 1H, Ar-H), 6.23 (br s, 1H, O-H), 4.53 (m, 2H, CH₂), 2.51 (m, 1H, C-H). ¹³C NMR (101 MHz, CDCl₃) δ 159.3, 152.4, 132.5, 110.0, 109.8, 103.4, 78.7, 76.1, 56.6. HRMS (CI+) *m/z* calculated for C₉H₉O₂: 149.0604. Found [M+H]⁺: 149.0606 (Diff 1.34 ppm).

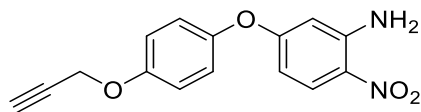
22c: 2-(Prop-2-yn-1-yloxy)phenol³²⁹



Reaction repeated as general procedure E using catechol (**20c**) gave the title compound (**22c**) as a crude yellow oil (0.28 g, 21 %). ¹H NMR (400 MHz, CDCl₃) δ 7.11 (d, 1H, *J* = 7.2 Hz, Ar-H), 6.88 (t, 1H, *J* = 2.2, 7.2, 12.4 Hz, Ar-H), 6.43 (m, 2H, Ar-H), 6.21 (br s, 1H, O-H), 4.55 (m, 2H, CH₂), 2.49 (m, 1H, C-H). ¹³C NMR (101 MHz, CDCl₃) δ 150.2, 147.5, 122.3, 121.8, 118.6, 117.4, 78.9, 75.8, 56.8. HRMS (CI+) *m/z* calculated for C₉H₉O₂: 149.0604. Found [M+H]⁺: 149.0611 (Diff 4.70 ppm).

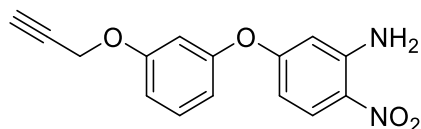
3.6.2.2.2 (Prop-2-yn-1-yloxy)phenoxy)benzene-1,2-diamine Derivatives

23a: 2-Nitro-5-(4-(prop-2-yn-1-yloxy)phenoxy)aniline



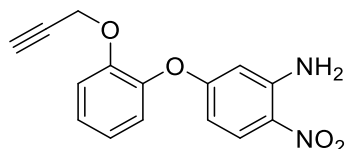
Reaction repeated as general procedure A using 4-(prop-2-yn-1-yloxy)phenol (**22a**) gave the title compound (**23a**) as a yellow solid (0.75 g, 93%). ¹H NMR (400 MHz, CDCl₃) δ 8.09 (d, 1H, *J*= 9.6 Hz, Ar-H), 7.02 (m, 4H, Ar-H), 6.31 (dd, 1H, *J*= 2.4, 9.6, 12.0 Hz, Ar-H), 6.17 (br s, 2H, NH₂), 6.12 (d, 1H, *J*= 2.8 Hz, Ar-H), 4.71 (m, 2H, CH₂), 2.56 (m, 1H, C-H). ¹³C NMR (101 MHz, CDCl₃) δ 164.7, 154.9, 148.4, 146.9, 128.7, 127.6, 122.0, 116.3, 107.5, 102.9, 78.3, 75.8, 56.3. HRMS (CI⁺) *m/z* calculated for C₁₅H₁₃N₂O₄: 285.2790. Found [M+H]⁺: 285.2779 (Diff -3.86 ppm).

23b: 2-Nitro-5-(3-(prop-2-yn-1-yloxy)phenoxy)aniline



Reaction repeated as general procedure A using 3-(prop-2-yn-1-yloxy)phenol (**22b**) gave the title compound (**23b**) as a yellow solid (1.06 g, 94%). ¹H NMR (400 MHz, CDCl₃) δ 7.92 (d, 1H, *J*= 9.2 Hz, Ar-H), 7.42 (t, 1H, *J*= 2.4, 9.2, 14.4 Hz, Ar-H), 6.72 (m, 2H, Ar-H), 6.50 (m, 3H, Ar-H), 6.23 (br s, 2H, NH₂), 4.73 (m, 2H, CH₂), 2.53 (m, 1H, C-H). ¹³C NMR (101 MHz, CDCl₃) δ 166.1, 160.2, 159.6, 143.4, 129.5, 125.4, 125.1, 110.3, 108.9, 108.7, 105.2, 78.2, 75.9, 56.1. HRMS (CI⁺) *m/z* calculated for C₁₅H₁₃N₂O₄: 285.2790. Found [M+H]⁺: 285.2799 (Diff 3.15 ppm).

23c: 2-Nitro-5-(2-(prop-2-yn-1-yloxy)phenoxy)aniline

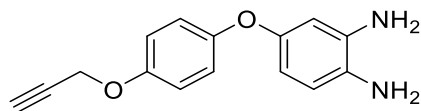


Reaction repeated as general procedure A using 2-(prop-2-yn-1-yloxy)phenol (**22c**) gave the title compound (**23c**) as a yellow solid (0.47 g, 87%). ¹H NMR (400 MHz, CDCl₃) δ 7.82 (d, 1H, *J*= 9.2 Hz, Ar-H), 6.86 (m, 4H, Ar-H), 6.75 (m, 2H, Ar-H), 6.28 (br s, 2H, NH₂), 4.76 (m, 2H, CH₂), 2.52 (m, 1H, C-H). ¹³C NMR (101 MHz, CDCl₃) δ 164.3, 153.2, 145.1, 144.8, 127.8, 126.9, 123.6,

120.1, 118.5, 118.2, 108.7, 103.3, 78.2, 76.3, 56.8. HRMS (CI+) m/z calculated for $C_{15}H_{13}N_2O_4$: 285.2790. Found $[M+H]^+$: 285.2791 (Diff 1.35 ppm).

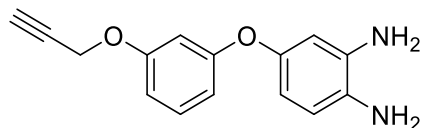
3.6.2.2.3 (Prop-2-yn-1-yloxy)phenoxy)benzene-1,2-diamine Derivatives

24a: 4-(4-(Prop-2-yn-1-yloxy)phenoxy)benzene-1,2-diamine



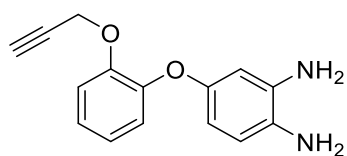
Reaction repeated as general procedure B using 2-nitro-5-(4-(prop-2-yn-1-yloxy)phenoxy)aniline (**23a**) gave the title compound (**24a**) as a brown oil (0.12 g, 18%). 1H NMR (400 MHz, $CDCl_3$) δ 7.53 (d, 2H, J = 8.0 Hz, Ar-H), 6.88 (d, 2H, J = 8.0 Hz, Ar-H), 6.62 (d, 1H, J = 7.2 Hz, Ar-H), 6.28 (m, 2H, Ar-H), 4.86 (br s, 4H, $(NH_2)_2$), 4.67 (m, 2H, CH_2), 2.53 (m, 1H, C-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 152.3, 150.4, 149.6, 138.7, 126.7, 120.4, 115.7, 114.3, 108.5, 102.0, 77.9, 75.8, 56.2. HRMS (CI+) m/z calculated for $C_{15}H_{15}N_2O_2$: 256.3233. Found $[M+H]^+$: 256.3229 (Diff -1.56 ppm).

24b: 4-(3-(Prop-2-yn-1-yloxy)phenoxy)benzene-1,2-diamine



Reaction repeated as general procedure B using 2-nitro-5-(3-(prop-2-yn-1-yloxy)phenoxy)aniline (**23b**) gave the title compound (**24b**) as a brown solid (0.50 g, 68%). 1H NMR (400 MHz, $CDCl_3$) δ 7.41 (t, 1H, J = 2.2, 7.6, 12.4 Hz, Ar-H), 6.43 (m, 4H, Ar-H), 6.33 (d, 1H, J = 7.2 Hz, Ar-H), 6.11 (d, 1H, J = 7.6 Hz, Ar-H), 4.88 (br s, 4H, $(NH_2)_2$), 4.86 (m, 2H, CH_2), 2.51 (m, 1H, C-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 160.2, 159.8, 149.1, 138.2, 130.3, 130.0, 118.6, 112.4, 107.4, 106.5, 104.2, 103.8, 78.2, 75.9, 55.1. HRMS (CI+) m/z calculated for $C_{15}H_{15}N_2O_2$: 256.3233. Found $[M+H]^+$: 256.3226 (Diff -2.73 ppm).

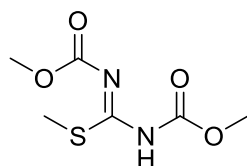
24c: 4-(2-(Prop-2-yn-1-yloxy)phenoxy)benzene-1,2-diamine



Reaction repeated as general procedure B using 2-nitro-5-(3-(prop-2-yn-1-yloxy)phenoxy)aniline (**23c**) gave the title compound (**24c**) as a brown solid (0.22 g, 53%). ¹H NMR (400 MHz, CDCl₃) δ 6.95 (m, 4H, Ar-H), 6.63 (d, 1H, *J* = 9.2 Hz, Ar-H), 6.43 (s, 1H, Ar-H), 6.18 (d, 1H, *J* = 9.2 Hz, Ar-H), 4.96 (br s, 4H, (NH₂)₂), 4.88 (m, 2H, CH₂), 2.52 (m, 1H, C-H). ¹³C NMR (101 MHz, CDCl₃) δ 154.2, 149.6, 145.3, 136.6, 130.8, 122.4, 119.3, 119.0, 117.1, 115.2, 112.6, 106.9, 79.0, 76.1, 55.4. HRMS (CI⁺) *m/z* calculated for C₁₅H₁₅N₂O₂: 256.3233. Found [M+H]⁺: 256.3236 (Diff 1.17 ppm).

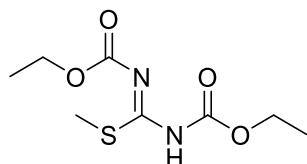
3.6.2.2.4 Methyisothioureia Derivatives

12b: 1,3-Bis(methoxyoxycarbonyl)-2-methylisothioureia³³⁰



Reaction repeated as general procedure C using methyl chloroformate (**11b**) gave the title compound (**12b**) as a white solid (1.92 g, 65%). This intermediate was not characterised and carried through to the subsequent reaction.

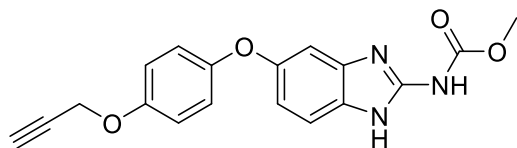
12c: 1,3-Bis(ethoxyoxycarbonyl)-2-methylisothioureia³³¹



Reaction repeated as general procedure C using ethyl chloroformate (**11c**). After basification, the reaction mixture was extracted with EtOAc (x3). The organic extracts were combined and dried over MgSO₄, filtered, and concentrated under reduced pressure to give the title compound (**12c**) as pale-yellow oil (0.33 g, 83%). This intermediate was not characterised and carried through to the subsequent reaction.

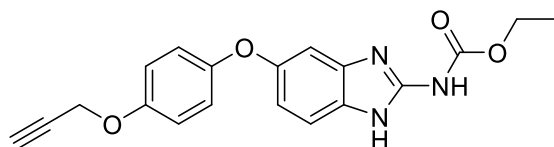
3.6.2.2.5 (Prop-2-yn-1-yloxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate Derivatives

25a: Methyl (5-(4-(prop-2-yn-1-yloxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



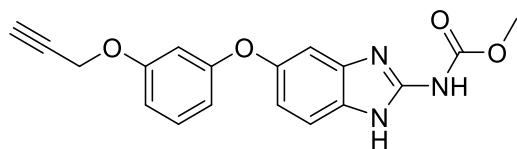
Reaction repeated as general procedure D using 4-(4-(prop-2-yn-1-yloxy)phenoxy)benzene-1,2-diamine (**24a**) and 1,3-bis(methoxycarbonyl)-2-methylisothiurea (**12b**) gave the title compound (**25a**) as an off-white solid (0.17 g, 21%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 12.03 (br s, 2H, N-H), 7.66 (d, 1H, *J* = 7.6 Hz, Ar-H), 7.21 (d, 1H, *J* = 7.2 Hz, Ar-H), 7.11 (m, 1H, Ar-H), 6.87 (m, 4H, Ar-H), 4.87 (m, 2H, CH₂), 3.72 (s, 3H, CH₃), 2.52 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 155.2, 151.4, 150.8, 148.7, 148.0, 135.2, 126.7, 120.4, 115.3, 114.6, 106.8, 78.7, 76.2, 55.2, 51.8. HRMS (ES+) *m/z* calculated for C₁₈H₁₆N₃O₄: 338.1141. Found [M+H]⁺: 338.1140 (Diff -0.30 ppm). IR ν_{\max} /cm⁻¹ (solid) 3352 (s), 1683 (s), 1203 (s), 1193 (s), 1183 (s). MP: 262-264°C decomposed. Purity HPLC 100%, R_t = 7.1 min.

25b: Ethyl (5-(4-(prop-2-yn-1-yloxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



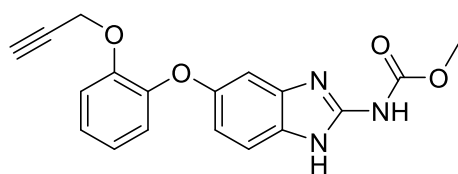
Reaction repeated as general procedure D using 4-(4-(prop-2-yn-1-yloxy)phenoxy)benzene-1,2-diamine (**24a**) and 1,3-bis(ethoxycarbonyl)-2-methylisothiurea (**12c**) gave the title compound (**25b**) as a beige solid (0.25 g, 30%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.98 (br s, 2H, N-H), 7.67 (d, 1H, *J* = 7.6 Hz, Ar-H), 7.23 (d, 1H, *J* = 7.2 Hz, Ar-H), 7.08 (m, 1H, Ar-H), 6.41 (m, 4H, Ar-H), 4.88 (m, 2H, CH₂), 3.72 (q, 2H, *J* = 2.2, 7.4 Hz, CH₂), 2.51 (m, 1H, C-H), 1.52 (t, 3H, *J* = 2.2, 7.4, 9.6 Hz, CH₃). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 154.3, 150.8, 150.6, 149.2, 147.6, 135.4, 134.2, 126.4, 119.5, 115.5, 114.7, 108.9, 78.6, 75.9, 64.1, 55.1, 15.2. HRMS (ES+) *m/z* calculated for C₁₉H₁₈N₃O₄: 352.1292. Found [M+H]⁺: 352.1293 (Diff 0.28 ppm). IR ν_{\max} /cm⁻¹ (solid) 3353 (s), 1681 (s), 1209 (s), 1197 (s), 1179 (s). MP: 258-262°C decomposed. Purity HPLC 91.3%, R_t = 7.1 min.

25c: Methyl (5-(3-(prop-2-yn-1-yloxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



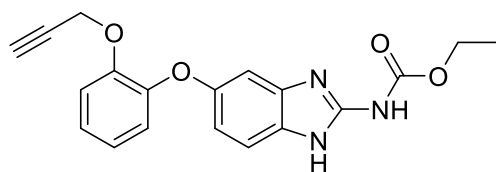
Reaction repeated as general procedure D using 4-(3-(prop-2-yn-1-yloxy)phenoxy)benzene-1,2-diamine (**24b**) and 1,3-bis(methoxycarbonyl)-2-methylisothiurea (**12a**) gave the title compound (**25c**) as an off-white solid (0.28 g, 42%). ^1H NMR (400 MHz, DMSO-*d*₆) δ 11.96 (br s, 2H, N-H), 7.63 (d, 1H, *J* = 9.2 Hz, Ar-H), 7.33 (t, 1H, *J* = 2.4, 9.2, 14.4 Hz, Ar-H), 7.12 (m, 1H, Ar-H), 7.01 (m, 4H, Ar-H), 4.86 (m, 2H, CH₂), 3.73 (s, 3H, CH₃), 2.52 (m, 1H, C-H). ^{13}C NMR (101 MHz, DMSO-*d*₆) δ 160.6, 159.3, 154.2, 149.9, 148.7, 130.6, 130.3, 115.8, 113.7, 110.2, 106.4, 103.7, 78.2, 76.2, 56.3, 55.7. HRMS (ES+) *m/z* calculated for C₁₈H₁₆N₃O₄: 338.1141. Found [M+H]⁺: 338.1139 (Diff -0.59 ppm). IR ν_{max} /cm⁻¹ (solid) 3348 (s), 1686 (s), 1202 (s), 1213 (s), 1180 (s). MP: 264-265°C decomposed. Purity HPLC 95.3%, R_t = 7.9 min.

25d: Methyl (5-(2-(prop-2-yn-1-yloxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



Reaction repeated as general procedure D using 4-(2-(prop-2-yn-1-yloxy)phenoxy)benzene-1,2-diamine (**24c**) and 1,3-bis(methoxycarbonyl)-2-methylisothiurea (**12b**) gave the title compound (**25d**) as an off-white solid (78 mg, 52%). ^1H NMR (400 MHz, DMSO-*d*₆) δ 11.98 (br s, 2H, N-H), 7.53 (d, 1H, *J* = 9.2 Hz, Ar-H), 7.02 (s, 1H, Ar-H), 6.85 (m, 4H, Ar-H), 4.85 (m, 2H, CH₂), 3.71 (s, 3H, CH₃), 2.53 (m, 1H, C-H). ^{13}C NMR (101 MHz, DMSO-*d*₆) δ 155.8, 154.2, 151.6, 149.2, 145.1, 134.8, 130.4, 123.9, 121.7, 120.2, 115.2, 114.8, 112.2, 78.4, 76.3, 56.8, 52.5. HRMS (ES+) *m/z* calculated for C₁₈H₁₆N₃O₄: 338.1141. Found [M+H]⁺: 338.1137 (Diff -1.18 ppm). IR ν_{max} /cm⁻¹ (solid) 3341 (s), 1683 (s), 1212 (s), 1226 (s), 1174 (s). MP: 266-270°C decomposed. Purity HPLC 99.7%, R_t = 7.7 min.

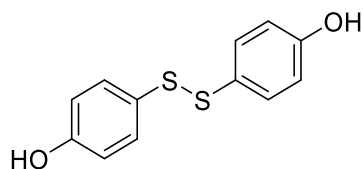
25e: Ethyl (5-(2-(prop-2-yn-1-yloxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



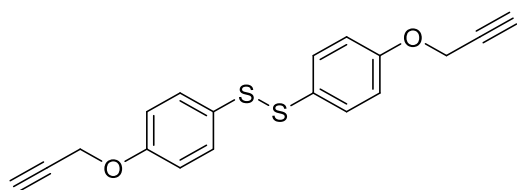
Reaction repeated as general procedure D using 4-(2-(prop-2-yn-1-yloxy)phenoxy)benzene-1,2-diamine (**24c**) and 1,3-bis(ethoxycarbonyl)-2-methylisothiourea (**12c**) gave the title compound (**25e**) as an off-white solid (50 mg, 31%). ^1H NMR (400 MHz, DMSO-*d*6) δ 11.92 (br s, 2H, N-H), 7.68 (d, 1H, J = 9.2 Hz, Ar-H), 7.06 (s, 1H, Ar-H), 6.92 (m, 2H, Ar-H), 6.84 (m, 3H, Ar-H), 4.83 (m, 2H, CH₂), 4.13 (q, 2H, J = 2.2, 7.6 Hz, CH₂), 2.51 (m, 1H, C-H), 1.21 (t, 3H, J = 2.2, 7.4, 9.6 Hz, CH₃). ^{13}C NMR (101 MHz, DMSO-*d*6) δ 155.5, 153.4, 150.8, 148.8, 145.2, 138.5, 130.0, 122.2, 119.8, 119.4, 119.0, 114.3, 113.5, 112.7, 106.2, 78.5, 75.8, 61.3, 57.4, 12.8. HRMS (ES+) m/z calculated for C₁₉H₁₈N₃O₄: 352.1292. Found [M+H]⁺: 352.1298 (Diff 1.70 ppm). IR $\nu_{\text{max}}/\text{cm}^{-1}$ (solid) 3342 (s), 1676 (s), 1209 (s), 1228 (s), 1168 (s). MP: 261-264°C decomposed. Purity HPLC 87.6%, R_t = 7.9 min.

3.6.2.3 (5-((4-(Prop-2-yn-1-yloxy)phenyl)thio)-1H-benzo[d]imidazol-2-yl)carbamate Analogues

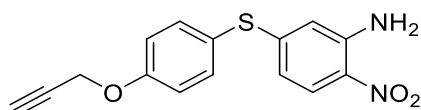
27: 4,4'-Disulfanedioldiphenol²⁸⁶



4-Sulfonyl phenol (**26**) (1 g, 7.93 mmol, 1 equiv.) was dissolved in DMSO (4 mL, 2 M) and allowed to stir vigorously at 60°C for 1 hour. After this time, the reaction was diluted with iced H₂O and filtered to give the title compound (**27**) as a pale-yellow solid (1.75 g, 176 %). ^1H NMR (400 MHz, CDCl₃) δ 7.33 (d, 4H, J = 8.8 Hz, Ar-H), 6.77 (d, 4H, J = 8.8 Hz, Ar-H), 6.08 (br s, 2H, O-H). HRMS (CI+) m/z calculated for C₁₂H₁₀O₂S₂: 251.0202. Found [M+H - C₆H₅OS]⁺: 127.0217. MP: 94.5-97.6°C, starting material 29.5°C.

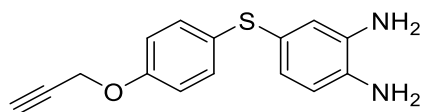
28: 1,2-Bis(4-(Prop-2-yn-1-yloxy)phenyl)disulfane²⁸⁶

4,4'-Disulfanediyldiphenol (**27**) (1.25 g, 4.99 mmol, 1 equiv.), propargyl bromide (**21**) (1.7 mL, 14.98 mmol, 3 equiv.) and K_2CO_3 (4.1 g, 29.96 mmol, 6 equiv.) were dissolved in acetone (6.2 mL, 0.8 M) and allowed to stir at reflux for 16 hours under a nitrogen atmosphere. After this time, the reaction was cooled and concentrated under reduced pressure. The residue was diluted with water and extracted with Et_2O (x3). The combined organic layers were washed with water (x1) and brine (x1), dried over $MgSO_4$, filtered, and concentrated *in vacuo* to yield a crude yellow oil. Purification *via* column chromatography eluting with 1.5-2% EtOAc in *n*-hexane gave the title compound (**28**) as a yellow oil (0.6 g, 37 %). 1H NMR (400 MHz, $CDCl_3$) δ 7.41 (d, 4H, J = 8.8 Hz, Ar-H), 6.90 (d, 4H, J = 8.8 Hz, Ar-H), 4.66 (m, 4H, CH_2), 2.52 (m, 2H, C-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 157.7, 132.0, 129.5, 115.6, 78.2, 75.9, 56.9. LRMS (ES+) m/z calculated for C_9H_8OS : 165.0. Found $[M+H-C_9H_7OS]^+$: 165.0.

29: 2-Nitro-5-((4-(prop-2-yn-1-yloxy)phenyl)thio)aniline

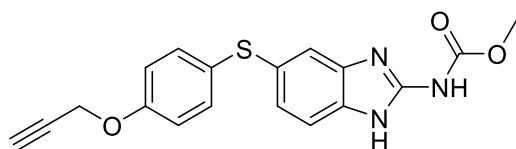
Reaction repeated as general procedure A using 1,2-bis(4-(prop-2-yn-1-yloxy)phenyl)disulfane (**28**) gave the title compound (**29**) as an orange solid (0.96 g, 20%). 1H NMR (400 MHz, $CDCl_3$) δ 8.00 (d, 1H, J = 12.0 Hz, Ar-H), 7.50 (d, 1H, J = 12.0 Hz, Ar-H), 7.06 (d, 2H, J = 8.0 Hz, Ar-H), 6.38 (d, 1H, J = 8.0 Hz, Ar-H), 6.27 (d, 1H, J = 4.0 Hz, Ar-H), 6.01 (br s, 2H, NH_2), 4.76 (m, 2H, CH_2), 2.52 (m, 1H, C-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 150.1, 140.0, 139.7, 137.1, 135.3, 133.1, 132.4, 128.9, 116.5, 113.0, 77.9, 75.1, 56.2. HRMS (ES+) m/z calculated for $C_{15}H_{13}N_2O_3S$: 277.025. Found $[M+H]^+$: 277.026 (Diff 3.61 ppm).

30: 4-((4-(Prop-2-yn-1-yloxy)phenyl)thio)benzene-1,2-diamine



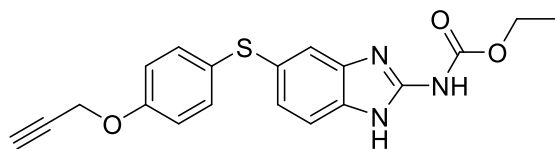
Reaction repeated as general procedure B using 2-nitro-5-((4-(prop-2-yn-1-yloxy)phenyl)thio)aniline (**29**) gave the title compound (**30**) as a brown oil (57.1 mg, 48 %). ¹H NMR (400 MHz, CDCl₃) δ 7.43 (d, 2H, *J* = 8.4 Hz, Ar-H), 7.25 (m, 1H, Ar-H), 7.13 (d, 2H, *J* = 8.4 Hz, Ar-H), 6.48 (m, 2H, Ar-H), 4.91 (br s, 4H, (NH₂)₂), 4.69 (m, 2H, CH₂), 2.53 (m, 1H, C-H). ¹³C NMR (101 MHz, CDCl₃) δ 157.2, 140.2, 136.8, 124.5, 122.3, 121.7, 119.6, 119.1, 115.6, 78.2, 76.4, 55.9. LRMS (ES+) *m/z* calculated for C₁₅H₁₅N₂OS: 271.1. Found [M+H]⁺: 271.1 (Diff -0.55 ppm).

31a: Methyl (5-((4-(prop-2-yn-1-yloxy)phenyl)thio)-1*H*-benzo[*d*]imidazol-2-yl)carbamate



Reaction repeated as general procedure D using 4-((4-(prop-2-yn-1-yloxy)phenyl)thio)benzene-1,2-diamine (**30**) and 1,3-bis(methoxycarbonyl)-2-methylisothiourea (**12b**) gave the title compound (**31a**) as a brown solid (0.54 g, 58%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.66 (br s, 2H, N-H), 7.39 (m, 2H, Ar-H), 7.26 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.10 (dd, 1H, *J* = 8.0 Hz, Ar-H), 6.98 (d, 2H, *J* = 12.0 Hz, Ar-H), 4.79 (m, 2H, CH₂), 3.75 (s, 3H, CH₃), 3.57 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 157.0, 155.0, 148.4, 140.8, 132.4, 128.7, 125.3, 116.4, 112.2, 79.5, 78.9, 56.0, 53.0. HRMS (ES+) *m/z* calculated for C₁₈H₁₅N₃O₃SNa: 376.0732. Found [M+Na]⁺: 376.0725 (Diff -1.86 ppm). IR *v*_{max}/cm⁻¹ (solid) 3348 (s), 1662 (s), 1174 (s), 1286 (s), 1142 (s). MP: 268-271°C decomposed. Purity HPLC 98.4%, *R*_t = 8.6 min.

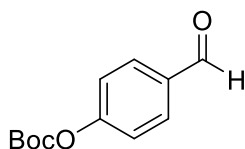
31b: Ethyl 5-((4-(prop-2-yn-1-yloxy)phenyl)thio)-1H-benzo[d]imidazol-2-yl)carbamate



Reaction repeated as general procedure D using 4-((4-(prop-2-yn-1-yloxy)phenyl)thio)benzene-1,2-diamine (**30**) and 1,3-bis(ethoxycarbonyl)-2-methylisothiourea (**12c**) gave the title compound (**31b**) as a brown solid (9.7 mg, 8%). ^1H NMR (400 MHz, DMSO-*d*₆) δ 11.68 (br s, 2H, N-H), 7.58 (m, 1H, Ar-H), 7.44 (d, 2H, *J* = 8.4 Hz, Ar-H), 7.38 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.14 (m, 3H, Ar-H), 4.88 (m, 2H, CH₂), 4.23 (q, 2H, *J* = 2.2, 7.4 Hz, CH₂), 2.52 (m, 1H, C-H), 1.14 (t, 3H, *J* = 2.2, 7.6, 9.6 Hz, CH₃). ^{13}C NMR (101 MHz, DMSO-*d*₆) δ 158.9, 155.4, 147.3, 140.5, 132.8, 128.7, 125.3, 125.2, 119.6, 116.2, 114.0, 78.5, 76.1, 60.8, 59.2, 12.9. HRMS (ES⁺) *m/z* calculated for C₁₉H₁₇N₃O₃SNa: 390.0971. Found [M+Na]⁺: 390.0968 (Diff -0.77 ppm). IR ν_{max} /cm⁻¹ (solid) 3342 (s), 1668 (s), 1171 (s), 1292 (s), 1151 (s). MP: 266-268°C decomposed. Purity HPLC 96.5%, *R*_t = 8.6 min.

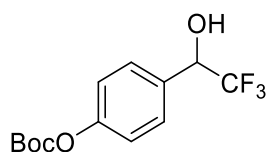
3.6.2.4 Synthesis of 3-Trifluoro-3-Phenyldiazirine Probe

39: *Tert*-butyl (4-formylphenyl) carbonate²⁹⁴



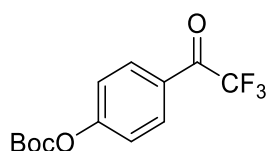
4-Hydroxy benzaldehyde (**38**) (1 g, 8.19 mmol, 1 equiv.), boc anhydride (3.60 g, 16.38 mmol, 2 equiv.) and PPh₃ (0.22 g, 0.82 mmol, 10% mol) were dissolved in anhydrous THF (27 mL, 0.3 M) and allowed to stir at 65°C for 16 hours under a nitrogen atmosphere. After this time, the solvent was removed under reduced pressure. The residue was diluted in water and extracted with EtOAc (x3). The combined organic layers were washed with brine (x1), dried over MgSO₄, filtered, and concentrated *in vacuo* to yield a crude clear oil. Purification *via* column chromatography eluting with 0-3% EtOAc in *n*-hexane gave the title compound (**39**) as a clear oil (1.82 g, 100%). ^1H NMR (400 MHz, CDCl₃) δ 7.91 (d, 2H, *J* = 8.8 Hz, Ar-H), 7.36 (d, 2H, *J* = 8.4 Hz, Ar-H), 1.57 (s, 9H, (CH₃)₃). ^{13}C NMR (101 MHz, CDCl₃) δ 190.7, 155.6, 151.0, 133.6, 131.3, 121.9, 83.8, 27.5. HRMS (ES⁺) *m/z* calculated for C₁₂H₁₅O₄ : 223.0972. Found [M+H]⁺: 223.097 (Diff -0.90 ppm).

40: Tert-butyl (4-(2,2,2-trifluoro-1-hydroxyethyl)phenyl) carbonate²⁹⁵



Tert-butyl (4-formylphenyl) carbonate (**39**) (2.22 g, 9.99 mmol, 1 equiv.) and trimethyl(trifluoromethyl)silane (2.1 mL, 13.99 mmol, 1.4 equiv.) were dissolved in anhydrous THF (20 mL, 0.5 M) and cooled to 0°C. TBAF (2 drops (cat)) was added to the reaction and allowed to stir at room temperature for 2 hours under a nitrogen atmosphere. After this time 1 M HCl (20 mL, 0.5 M) was added and the reaction was allowed to stir at room temperature for a further 2 hours. After completion, the reaction was diluted with water and extracted with EtOAc (x3). The combined organic layers were washed with 0.1 M HCl aq. solution (x1), water (x1) and brine (x1), dried over MgSO₄, filtered, and concentrated under reduced pressure to give the title compound (**40**) as a white solid (2.45 g, 84%). ¹H NMR (400 MHz, CDCl₃) δ 7.49 (d, 2H, *J* = 8.4 Hz, Ar-H), 7.22 (d, 2H, *J* = 9.2 Hz, Ar-H), 5.01 (q, 1H, *J* = 6.8, 13.2, 20.0 Hz, C-H), 1.56 (s, 9H, (CH₃)₃). ¹⁹F NMR (376 MHz, CDCl₃) δ -78.49 (s, 3F). ¹³C NMR (101 MHz, CDCl₃) δ 151.9, 150.2, 131.8, 131.3, 121.5, 121.2, 83.8, 72.1, 27.9. LRMS (CI+) *m/z* calculated for C₈H₈¹⁹F₃O₂: 193.0. Found [M+H-Boc]⁺: 193.0.

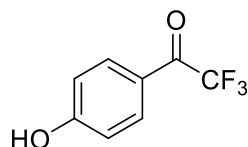
41: Tert-butyl (4-(2,2,2-trifluoroacetyl)phenyl) carbonate²⁹⁹



Tert-butyl (4-(2,2,2-trifluoro-1-hydroxyethyl)phenyl) carbonate (**40**) (0.10 g, 0.34 mmol, 1 equiv.) and SIBX (0.42 g, 0.68 mmol, 2 equiv.) were dissolved in EtOAc (5 mL) and allowed to stir at reflux for 16 hours under a nitrogen atmosphere. After this time, the precipitate was removed *via* Buchner filtration. The filtrate was concentrated under reduced pressure to yield a crude yellow oil. Purification *via* column chromatography eluting with 100 % *n*-hexane gave the title compound (**41**) as a white solid (70.3 mg, 70%). ¹H NMR (400 MHz, CDCl₃) δ 8.12 (d, 2H, *J* = 8.8 Hz, Ar-H), 7.38 (d, 2H, *J* = 8.8 Hz, Ar-H), 1.58 (s, 9H, (CH₃)₃). ¹⁹F NMR (376 MHz, CDCl₃) δ -71.5 (s, 3F). ¹³C NMR (101 MHz, CDCl₃) δ 179.4, 156.2, 150.6, 132.0, 130.2, 121.8,

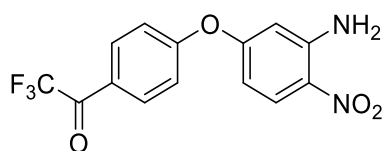
118.0, 84.7, 27.6. LRMS (CI+) m/z calculated for $C_8H_5^{19}F_3O_2$: $C_8H_6^{19}F_3O_2$. Found $[M+H-Boc]^+$: 191.0.

42: 2,2,2-Trifluoro-1-(4-hydroxyphenyl)ethan-1-one³³²



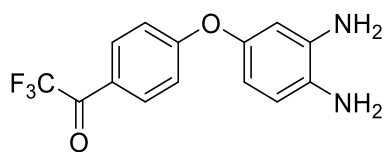
Tert-butyl (4-(2,2,2-trifluoroacetyl)phenyl) carbonate (**41**) (0.46 g, 1.59 mmol, 1 equiv.) was dissolved in TFA (2 mL) and DCM (8 mL) and allowed to stir at room temperature for 3 hours. After this time, the reaction was diluted with DCM, cooled to 0°C, neutralised with saturated aq. $NaHCO_3$ solution, warmed to room temperature, and separated. The organic layer was washed with water (x1) and brine (x1). The organic phase was dried over $MgSO_4$, filtered, and concentrated *in vacuo* to give the title compound (**42**) as a pale-yellow oil (0.30 g, 100%). 1H NMR (400 MHz, $CDCl_3$) δ 8.02 (d, 2H, J = 8.0 Hz, Ar-H), 7.48 (br s, 1H, O-H), 6.99 (d, 2H, J = 8.0 Hz, Ar-H). ^{19}F NMR (376 MHz, $CDCl_3$) δ -70.96 (s, 3F). ^{13}C NMR (101 MHz, $CDCl_3$) δ 179.4, 162.8, 133.2, 122.3, 118.2, 116.2. HRMS (CI+) m/z calculated for $C_8H_6^{19}F_3O_2$: 191.0322. Found $[M+H]^+$: 191.0321 (Diff -0.52 ppm).

43: 1-(4-(3-Amino-4-nitrophenoxy)phenyl)-2,2,2-trifluoroethan-1-one



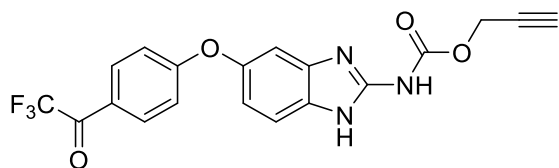
Reaction repeated as general procedure A using 2,2,2-trifluoro-1-(4-hydroxyphenyl)ethan-1-one (**42**) gave the title compound (**43**) as a yellow oil (0.90 g, 72%). 1H NMR (400 MHz, $CDCl_3$) δ 8.17 (m, 1H, Ar-H), 8.01 (d, 2H, J = 8.0 Hz, Ar-H), 6.98 (d, 2H, J = 12.0 Hz, Ar-H), 6.49 (dd, 1H, J = 8.0, 12.0 Hz, Ar-H), 6.43 (m, 1H, Ar-H), 6.25 (br s, 2H, NH_2). ^{19}F NMR (376 MHz, $CDCl_3$) δ -70.94 (s, 3F). ^{13}C NMR (101 MHz, $CDCl_3$) δ 168.2, 165.7, 162.6, 146.8, 133.1, 133.1, 129.4, 129.3, 129.2, 119.3, 129.2, 119.5, 118.4, 116.1, 105.7, 103.8. HRMS (CI+) m/z calculated for $C_{14}H_{10}^{19}F_3N_2O_4$: 327.0594. Found $[M+H]^+$: 327.0576 (Diff -5.50 ppm).

44: 1-(4-(3,4-Diaminophenoxy)phenyl)-2,2,2-trifluoroethan-1-one



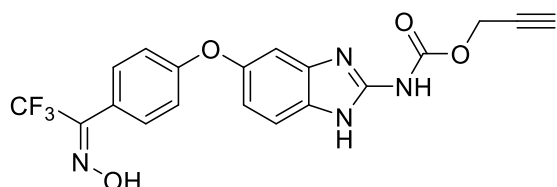
Reaction repeated as general procedure B using 1-(4-(3-amino-4-nitrophenoxy)phenyl)-2,2,2-trifluoroethan-1-one (**43**) gave the title compound (**44**) as a brown oil (0.19 g, 29%). ¹H NMR (400 MHz, CDCl₃) δ 7.40 (d, 1H, *J* = 8.4 Hz, Ar-H), 6.98 (m, 4H, Ar-H), 6.54 (dd, 1H, *J* = 2.4, 8.4, 10.8 Hz, Ar-H), 6.40 (d, 1H, *J* = 2.4 Hz, Ar-H). ¹³C NMR (101 MHz, CDCl₃) δ 172.4, 158.5, 153.9, 139.3, 129.0, 128.9, 128.6, 118.1, 117.9, 117.0, 112.4, 111.8. LRMS (ES+) *m/z* calculated for C₁₄H₁₂N₂O₂¹⁹F₃: 296.1. Found [M+H]⁺: 296.1.

45: Prop-2-yn-1-yl (5-(4-(2,2,2-trifluoroacetyl)phenoxy)-1*H*-benzo[*d*]imidazol-2-yl)carbamate



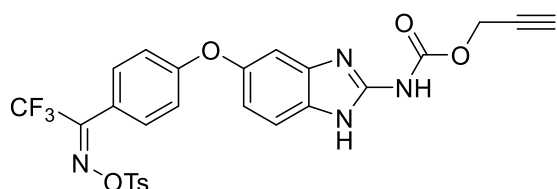
Reaction repeated as general procedure D using 1-(4-(3,4-diaminophenoxy)phenyl)-2,2,2-trifluoroethan-1-one (**44**) and 1,3-bis(prop-2-yn-1-yloxy-carbonyl)-2-methylisothiourea (**12a**) gave the title compound (**45**) as a brown solid (8.76 mg, 28%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.40 (d, 1H, *J* = 8.4 Hz, Ar-H), 6.98 (m, 4H, Ar-H),), 6.54 (dd, 1H, *J* = 2.4, 8.4, 10.8 Hz, Ar-H), 6.40 (d, 1H, *J* = 2.4 Hz, Ar-H), 4.82 (m, 2H, CH₂), 3.86 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 172.4, 158.5, 153.9, 139.3, 129.0, 128.9, 128.6, 126.5, 121.3, 118.1, 117.9, 117.0, 112.4, 111.8, 78.6, 76.2, 54.1. HRMS (ES+) *m/z* calculated for C₁₉H₁₃N₃O₄¹⁹F₃: 403.0858. Found [M+H]⁺: 403.0855 (Diff -0.74 ppm).

46: Prop-2-yn-1-yl (*E*)-(5-(4-(2,2,2-trifluoro-1-(hydroxyimino)ethyl)phenoxy)-1*H*-benzo[*d*]imidazol-2-yl)carbamate³⁰⁵



To a solution of prop-2-yn-1-yl (5-(4-(2,2,2-trifluoroacetyl)phenoxy)-1*H*-benzo[*d*]imidazol-2-yl)carbamate (**45**) (22.2 mg, 0.056 mmol, 1.0 equiv.) in methanol (10 mL) was added hydroxylamine hydrochloride (7.2 mg, 0.110 mmol, 2.0 equiv.) and sodium acetate (11.5 mg, 0.140 mmol, 2.5 equiv.) and allowed to stir at 50 °C for 16 hours under a nitrogen atmosphere. After this time, the reaction was cooled and concentrated under reduced pressure to give the title compound (**46**) as a crude brown suspension. This intermediate was not characterised and carried through to the subsequent reaction.

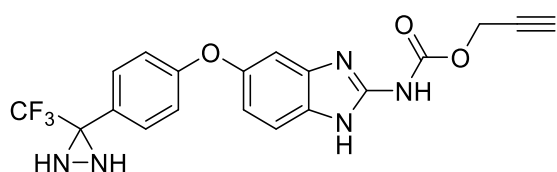
47: Prop-2-yn-1-yl (*E*)-(5-(4-(2,2,2-trifluoro-1-((tosyloxy)imino)ethyl)phenoxy)-1*H*-benzo[*d*]imidazol-2-yl)carbamate³⁰⁶



Prop-2-yn-1-yl (*E*)-(5-(4-(2,2,2-trifluoro-1-(hydroxyimino)ethyl)phenoxy)-1*H*-benzo[*d*]imidazol-2-yl)carbamate (**46**) (26.4 mg, 0.063 mmol, 1 equiv.), K₂CO₃ (16 mg, 0.11 mmol, 1.8 equiv.) and TsCl (13 mg, 0.069 mmol, 1.1 equiv.) were suspended in DMSO (10 mL) and allowed to stir at 70°C for 16 hours under a nitrogen atmosphere. After this time, the reaction was cooled and partially concentrated. The resultant residue was suspended in EtOAc and the solid was isolated *via* Buchner filtration and washed with 5% aq. NaOH solution (x3) and saturated aq. NaHCO₃ solution (x2). The crude brown solid/ oil was purified *via* column chromatography eluting with 0-10% MeOH in EtOAc to give the title compound (**47**) as a brown solid (13 mg, 36%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 8.44 (d, 2H, *J* = 8.0 Hz, Ar-H), 8.12 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.44 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.41 (m, 1H, Ar-H), 7.38 (d, 2H, *J* = 8.4 Hz, Ar-H), 7.25 (m, 1H, Ar-H), 6.93 (m, 1H, Ar-H), 4.86 (m, 2H, CH₂), 3.82 (m, 1H, C-H), 2.44 (s, 3H, CH₃). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 166.2, 160.2, 154.0, 153.4, 148.7, 140.9, 140.3,

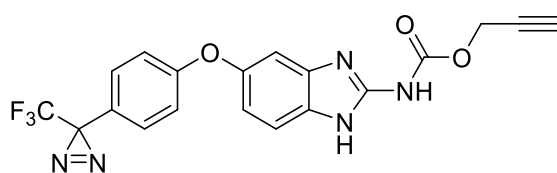
139.6, 131.4, 130.2, 130.0, 128.5, 128.3, 125.9, 121.8, 118.5, 115.7, 115.4, 105.6, 80.6, 76.8, 55.8, 20.3. LRMS (ES+) m/z calculated for $C_{26}H_{20}N_4O_6^{19}F_3S$: 573.11. Found $[M+H]^+$: 573.11 (Diff -1.35 ppm).

48: Prop-2-yn-1-yl (5-(4-(3-(trifluoromethyl)diaziridin-3-yl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



Prop-2-yn-1-yl (E)-(5-(4-(2,2,2-trifluoro-1-((tosyloxy)imino)ethyl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate (**47**) (13 mg, 0.02 mmol, 1equiv.) was added to a screw fix sealed tube and cooled to $-78^{\circ}C$. $4N NH_3$ in MeOH (10 mL) was added and allowed to stir for 15 minutes. The reaction was gently heated to encourage dissolution, then allowed to stir at room temperature for 16 hours. After this time, the reaction was concentrated under air flow, filtered *via* sinter filtration, and washed with methanol to give the title compound (**48**) as a brown solid (9.5 mg, 46%). This intermediate was not characterised and carried through to the subsequent reaction.

49: Prop-2-yn-1-yl (5-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



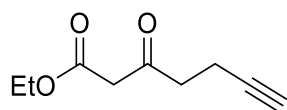
Prop-2-yn-1-yl (5-(4-(3-(trifluoromethyl)diaziridin-3-yl)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate (**48**) (22 mg, 0.05 mmol, 1 equiv.) was dissolved in anhydrous MeOH (5 mL) and gently heated to encourage dissolution. NEt_3 (2 drops) was added to the reaction mixture. A solution of I_2 (20 mg, 0.08 mmol, 1.5 equiv.) in MeOH (2 mL) was added portion-wise until a persistent brown coloured solution was observed. Excess I_2 was quenched *via* the addition of 5% aqueous $Na_2S_2O_3$ solution, followed by filtering the material with cold water and methanol to give a crude solid. Purification *via* column chromatography eluting with 0-15% MeOH in EtOAc gave the title compound (**49**) as a brown solid (4.2 mg, 19%). 1H NMR (400

MHz, DMSO-*d*₆) δ 12.12 (br s, 2H, N-H), 7.62 (d, 1H, *J* = 7.6 Hz, Ar-H), 7.31 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.03 (m, 4H, Ar-H), 4.89 (m, 2H, CH₂), 2.53 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 160.3, 155.8, 151.2, 148.0, 136.8, 132.3 132.3, 130.7, 126.5, 120.8, 114.2, 107.9, 103.3, 85.8, 78.6, 75.7, 54.1. LRMS (ES+) *m/z* calculated for C₁₉H₁₃N₅O₃¹⁹F₃: 416.10. Found [M+H]⁺: 416.10 (Diff -0.6 ppm).

3.6.2.5 Synthesis of Minimalist Photo-Crosslinker

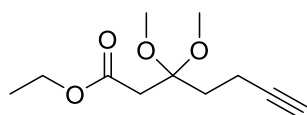
3.6.2.5.1 Carbendazim based aliphatic PAL probe

52: Ethyl 3-oxohept-6-ynoate³¹¹



Diisopropylamine (4.74 mL, 33.8 mmol, 2.2 equiv.) was dissolved in anhydrous THF (15.4 mL, 1 M) and allowed to stir at -78°C under an argon atmosphere. *n*-BuLi (13.5 mL, 33.8 mmol, 2.2 equiv.) was added and continued to stir at -78°C for 15 minutes. Ethyl acetoacetate (**51**) (1.94 mL, 15.4 mmol, 1 equiv.) was added and allowed to stir at 0°C for 15 minutes. Propargyl bromide (**21**) (0.72 mL, 15.4 mmol, 1 equiv.) was added and allowed to stir at room temperature for 16 hours. After this time, the reaction was poured into saturated aq. NH₄Cl and extracted with EtOAc (x3). The combined organic extracts were washed with brine (x1), dried over MgSO₄, filtered, and concentrated under reduced pressure to yield a crude dark brown oil. Purification *via* column chromatography eluting with 100% *n*-hexane gave the title compound (**52**) as a yellow oil (1.66 g, 64%). ¹H NMR (400 MHz, CDCl₃) δ 4.06 (q, 2H, *J* = 7.2, 14.4 Hz, CH₂), 3.36 (s, 2H, CH₂), 2.69 (t, 2H, *J* = 7.2, 14.4 Hz, CH₂), 2.33 (m, 2H, CH₂), 1.88 (s, 1H, C-H), 1.15 (t, 3H, *J* = 7.2, 14.4 Hz, CH₃). ¹³C NMR (101 MHz, CDCl₃) δ 200.6, 166.8, 82.4, 69.0, 61.3, 49.0, 41.4, 13.8, 12.5. HRMS (ES+) *m/z* calculated for C₉H₁₂O₃: 191.0685. Found [M+Na]⁺: 191.0678 (Diff -3.66 ppm).

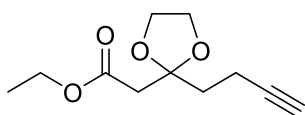
53b: Ethyl 3,3-dimethoxyhept-6-ynoate³¹⁴



Ethyl 3-oxohept-6-ynoate (**52**) (0.2 g, 1.19 mmol, 1 equiv.) trimethyl orthoformate (0.14 mL, 1.31 mmol, 1.1 equiv.), tosic acid monohydrate (0.11 g, 0.96 mmol, 0.5 equiv.) and 4Å

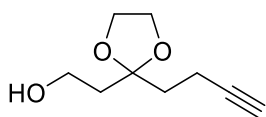
activated molecular sieves (0.3 g) were added to anhydrous methanol (1.2 mL, 1 M) and allowed to stir at room temperature for 16 hours under a nitrogen atmosphere. After this time, the reaction mixture was neutralised with NaOMe and concentrated under reduced pressure. The residue was dissolved in Et₂O and washed with saturated aq. NH₄Cl solution (x1), 1 M aq. NaOH solution (x1), and brine (x1). The organic layer was dried over MgSO₄, filtered, and concentrated *in vacuo* to yield a crude brown oil. Purification *via* column chromatography eluting with 5% EtOAc in *n*-hexane gave the title compound (**53b**) as a brown oil (0.22 g, 88%). ¹H NMR (400 MHz, CDCl₃) δ 4.20 (q, 2H, *J*= 7.2, 14.4 Hz, CH₂), 3.74 (s, 6H, (CH₃)₂), 2.82 (t, 2H, *J*= 7.2, 14.4 Hz, CH₂), 2.46 (m, 4H, (CH₂)₂), 1.96 (m, 1H, C-H), 1.29 (t, 3H, *J*= 7.2, 14.4 Hz, CH₃). ¹³C NMR (101 MHz, CDCl₃) δ 166.9, 115.6, 82.4, 68.9, 61.7, 49.1, 21.5, 14.1, 12.5. LRMS (ES+) *m/z* calculated for C₁₁H₁₉O₄: 215.13. Found [M+H]⁺: 215.13 (Diff -3.22 ppm).

53a: Ethyl 2-(2-(but-3-yn-1-yl)-1,3-dioxolan-2-yl)acetate³¹¹



Ethyl 3-oxohept-6-ynoate (**52**) (0.53 g, 3.15 mmol, 1 equiv.), tosic acid monohydrate (0.02 g, 0.13 mmol, 0.04 equiv.), ethylene glycol (1.58 mL, 2 M) and toluene (10 mL) were added to a flask and fixed to insulated Dean Stark head apparatus and allowed to stir at 140°C for 1 hour under a nitrogen atmosphere. After this time, the water was drained, and the toluene solution was concentrated under reduced pressure to give the title compound (**53a**) as a yellow oil (0.65 g, 97%). ¹H NMR (400 MHz, CDCl₃) δ 4.16 (q, 2H, *J*= 8.0, 12.0, 20.0 Hz, CH₂), 4.02-3.97 (m, 4H, (CH₂)₂), 2.65 (s, 2H, CH₂), 2.33-2.28 (m, 2H, CH₂), 2.12 (t, 2H, *J*= 8.0, 16.0 Hz, CH₂), 1.95-1.93 (m, 1H, C-H), 1.27 (t, 3H, *J*= 8.0, 16.0 Hz, CH₃). ¹³C NMR (101 MHz, CDCl₃) δ 167.2, 115.1, 83.4, 68.7, 62.6, 60.9, 40.1, 38.4, 14.2, 11.6. HRMS (ES+) *m/z* calculated for C₁₁H₁₇O₄: 213.1127. Found [M+H]⁺: 213.1122 (Diff -2.35 ppm).

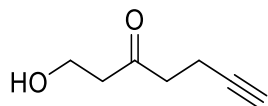
54: 2-(2-(But-3-yn-1-yl)-1,3-dioxolan-2-yl)ethan-1-ol³¹¹



To a flask LiAlH₄ (1.62 mL, 3.25 mmol, 1.5 equiv.) was suspended in anhydrous THF (1.8 mL, 1.2 M) and cooled to 0°C. Ethyl 2-(2-(but-3-yn-1-yl)-1,3-dioxolan-2-yl)acetate (**53a**) (0.46 g,

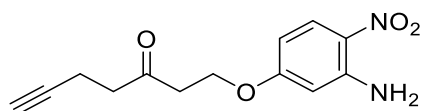
2.17 mmol, 1 equiv.) was added dropwise and allowed to stir at reflux for 2 hours under a nitrogen atmosphere. After this time, the reaction was cooled to 0°C and water (5 mL) was slowly added with vigorous stirring. 1 M aq. KOH solution (4 mL) was added. The biphasic solution was filtered over a pad of silica and washed with Et₂O. The organic phase of the filtrate was separated, washed with brine, dried over MgSO₄, filtered, and concentrated under reduced pressure to give the title compound (**54**) as a pale-yellow oil (0.33 g, 89%). ¹H NMR (400 MHz, CDCl₃) δ 4.60 (br s, 1H, O-H), 4.01-3.96 (m, 4H, (CH₂)₂), 2.41-2.38 (m, 2H, CH₂), 2.88-2.26 (m, 1H, C-H), 1.82 (m, 2H, CH₂), 1.77 (t, 2H, *J*=8.0, 16.0 Hz, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 115.8, 83.9, 68.4, 61.8, 60.2, 36.5, 19.6, 10.2. HRMS (ES+) *m/z* calculated for C₉H₁₅O₃: 171.1021. Found [M+H]⁺: 171.1026 (Diff -2.92 ppm).

55: 1-Hydroxyhept-6-yn-3-one³¹¹



To a flask 2-(2-(but-3-yn-1-yl)-1,3-dioxolan-2-yl)ethan-1-ol (**54**) (0.14 g, 0.82 mmol, 1 equiv.) and TsOH monohydrate (40 mg, 0.21 mmol, 0.25 equiv.) was dissolved in acetone (2.7 mL, 0.3 M) and allowed to stir at reflux for 6 hours. After this time, the reaction mixture was concentrated under reduced pressure. The residue was dissolved in EtOAc, washed with saturated aq. NaHCO₃ solution (x1), water (x1) and brine (x1), dried over MgSO₄, filtered, and concentrated *in vacuo* to yield a crude oil. Purification *via* column chromatography eluting with 50% EtOAc in *n*-hexane gave the title compound (**55**) as a yellow oil (95 mg, 95%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 3.97 (t, 2H, *J*= 8.0, 16.0 Hz, CH₂), 3.16 (t, 2H, *J*= 8.0, 16.0 Hz, CH₂), 2.77 (t, 2H, *J*= 8.0, 12.0 Hz, CH₂), 2.96-2.94 (m, 1H, C-H), 2.22-2.19 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 213.6, 84.5, 70.1, 55.2, 45.3, 39.8. LRMS (ES+) *m/z* calculated for C₇H₁₁O₂: 127.08. Found [M+H]⁺: 127.08 (Diff -1.2 ppm).

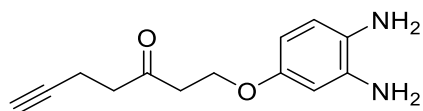
56: 1-(3-Amino-4-nitrophenoxy)hept-6-yn-3-one



Reaction repeated as general procedure A using 1-hydroxyhept-6-yn-3-one (**55**) gave the title compound (**56**) as a yellow oil (0.30 g, 63%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.88 (d, 1H, *J*=

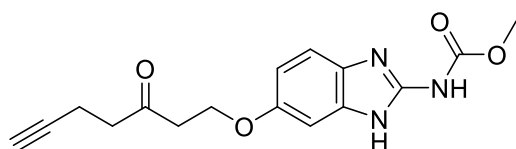
8.0 Hz, Ar-H), 6.33 (d, 1H, $J = 4.0$ Hz, Ar-H), 6.28 (dd, 1H, $J = 4.0, 12.0$ Hz, Ar-H), 4.32 (t, 2H, $J = 8.0, 16.0$ Hz, CH₂), 3.88 (t, 2H, $J = 8.0, 16.0$ Hz, CH₂), 3.41 (t, 2H, $J = 8.0, 12.0$ Hz, CH₂), 2.93-2.91 (m, 1H, C-H), 2.26-2.22 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 212.1, 168.2, 140.3, 126.4, 125.7, 104.4, 101.6, 84.9, 70.2, 60.2, 40.3, 39.5, 12.8. HRMS (ES+) m/z calculated for C₁₃H₁₅N₂O₄: 263.1034. Found [M+H]⁺: 263.1041 (Diff 2.66 ppm).

57: 1-(3,4-Diaminophenoxy)hept-6-yn-3-one



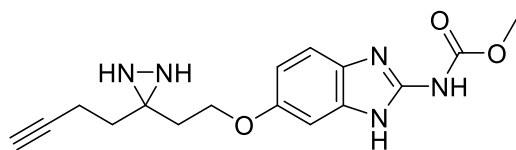
Reaction repeated as general procedure B using 1-(3-amino-4-nitrophenoxy)hept-6-yn-3-one (**56**) gave the title compound (**57**) as a brown oil (84 mg, 42%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 6.71 (d, 1H, $J = 12.0$ Hz, Ar-H), 6.66 (d, $J = 8.0$ Hz, Ar-H), 6.12 (dd, 1H, $J = 4.0, 12.0$ Hz, Ar-H), 4.31 (t, 2H, $J = 8.0, 16.0$ Hz, CH₂), 3.86 (t, 2H, $J = 8.0, 16.0$ Hz, CH₂), 2.95-2.93 (m, 1H, C-H), 2.64-2.61 (m, 2H, CH₂), 2.21-2.18 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 211.3, 167.2, 142.1, 128.6, 114.2, 106.3, 101.8, 86.9, 70.1, 59.7, 40.1, 38.9, 12.6. LRMS (ES+) m/z calculated for C₁₃H₁₇N₂O₂: 233.13. Found [M+H]⁺: 233.13.

58: Methyl (6-((3-oxohept-6-yn-1-yl)oxy)-1H-benzo[d]imidazol-2-yl)carbamate



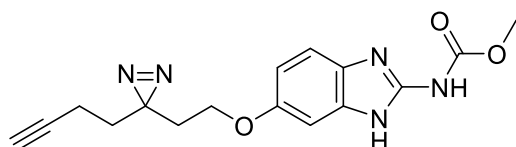
Reaction repeated as general procedure D using 1-(3,4-diaminophenoxy)hept-6-yn-3-one (**57**) and 1,3-bis(methoxycarbonyl)-2-methylisothiourea (**12b**) gave the title compound (**58**) as a beige solid (0.11 g, 58%). ¹H NMR (400 MHz, DMSO-*d*6) δ 7.78 (d, 1H, $J = 12.0$ Hz, Ar-H), 7.22 (d, 1H, $J = 8.0$ Hz, Ar-H), 7.14 (dd, 1H, $J = 4.0, 12.0$ Hz, Ar-H), 4.32 (t, 2H, $J = 8.0, 16.0$ Hz, CH₂), 3.83 (t, 2H, $J = 8.0, 16.0$ Hz, CH₂), 3.72 (s, 3H, CH₃), 2.93-2.90 (m, 1H, C-H), 2.64-2.60 (m, 2H, CH₂), 2.26-2.21 (m, 2H, CH₂). ¹³C NMR (101 MHz, DMSO-*d*6) δ 212.2, 154.3, 153.2, 148.9, 138.4, 128.1, 114.2, 107.2, 102.1, 85.4, 70.4, 58.6, 52.1, 42.6, 40.1, 12.1. HRMS (ES+) m/z calculated for C₁₆H₁₈N₃O₄: 316.1297. Found [M+H]⁺: 316.1289 (Diff -2.53 ppm).

59: Methyl (6-(2-(3-(but-3-yn-1-yl)diaziridin-3-yl)ethoxy)-1H-benzo[d]imidazol-2-yl)carbamate



Reaction repeated as general procedure F using methyl (6-((3-oxohept-6-yn-1-yl)oxy)-1H-benzo[d]imidazol-2-yl)carbamate (**58**) gave the title compound (**59**) as a crude brown solid. This intermediate was not characterised and carried through to the subsequent reaction.

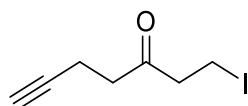
60a: Methyl (6-(2-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)ethoxy)-1H-benzo[d]imidazol-2-yl)carbamate



Reaction repeated as general procedure G using methyl (6-(2-(3-(but-3-yn-1-yl)diaziridin-3-yl)ethoxy)-1H-benzo[d]imidazol-2-yl)carbamate (**59**) gave the title compound (**60a**) as a brown solid (8.8 mg, 11%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.68 (d, 1H, *J* = 12.0 Hz, Ar-H), 7.21 (dd, 1H, *J* = 4.0, 12.0 Hz, Ar-H), 7.02 (d, 1H, *J* = 8.0 Hz, Ar-H), 4.41 (t, 2H, *J* = 8.0, 16.0 Hz, CH₂), 3.73 (s, 3H, CH₃), 2.88-2.86 (m, 1H, C-H), 2.31-2.29 (m, 2H, CH₂), 1.69-1.66 (m, 2H, CH₂), 1.11-1.07 (m, 2H, CH₂). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 155.6, 151.3, 148.6, 138.2, 126.7, 114.7, 112.5, 103.2, 84.9, 70.2, 68.4, 60.3, 54.4, 44.4, 43.2, 12.8. HRMS (ES⁺) *m/z* calculated for C₁₆H₁₈N₅O₃: 328.1409. Found [M+H]⁺: 328.1412 (Diff 0.91 ppm).

3.6.2.5.2 Bisaryl ether based aliphatic PAL probe

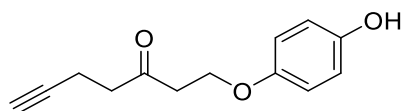
61: 1-Iodohept-6-yn-3-one³¹¹



PPh₃ (30 mg, 0.11 mmol, 1 equiv.) and imidazole (7.7 mg, 0.11 mmol, 1 equiv.) were added to a flask of 1-hydroxyhept-6-yn-3-one (**55**) (14.2 mg, 0.11 mmol, 1 equiv.) in anhydrous THF (0.1 mL, 1.1 M) and allowed to stir at room temperature for 5 minutes. Iodine (37 mg, 0.15 mmol, 1.3 equiv.) was added and allowed to stir for 4 hours at room temperature under a nitrogen

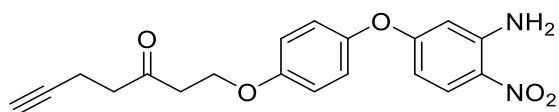
atmosphere. After this time, saturated aq. Na₂S₂O₃ solution was added to destroy excess iodine and extracted with DCM (x3). The combined organic phases were washed with brine, dried over MgSO₄, filtered, and concentrated under reduced pressure to yield a crude brown oil. Purification *via* column chromatography eluting with 30-40% EtOAc in *n*-hexane gave the title compound (**61**) as an orange solid (24.8 mg, 93%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 3.45-3.41 (m, 2H, CH₂), 2.99 (t, 2H, *J*= 8.0, 12.0 Hz, CH₂), 2.85-2.82 (m, 2H, CH₂), 2.46 (m, 2H, *J*= 8.0, 16.0 Hz, CH₂), 2.21-2.18 (m, 1H, C-H). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 213.3, 85.6, 70.2, 44.1, 40.3, 12.1, 1.2. LRMS (CI+) *m/z* calculated for C₇H₁₀O¹²⁷I: 236.98. Found [M+H]⁺: 236.97 (Diff -1.44 ppm).

62: 1-(4-Hydroxyphenoxy)hept-6-yn-3-one



Reaction repeated as general procedure E using 1-iodohept-6-yn-3-one (**61**) and hydroquinone (**20a**) gave the title compound (**62**) as an orange oil (0.71 g, 48%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.32 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.06 (d, 2H, *J*= 8.0 Hz, Ar-H), 4.33 (t, 2H, *J*= 8.0, 12.0 Hz, CH₂), 3.57 (t, 2H, *J*= 8.0, 16.0 Hz, CH₂), 2.88-2.86 (m, 1H, C-H), 2.69 (t, 2H, *J*= 8.0, 12.0 Hz, CH₂), 2.21-2.18 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 213.6, 153.2, 151.1, 115.8, 113.4, 85.6, 70.6, 60.3, 45.7, 39.8, 12.4. HRMS (ES+) *m/z* calculated for C₁₃H₁₅O₃: 219.1021. Found [M+H]⁺: 219.1015 (Diff -2.74 ppm).

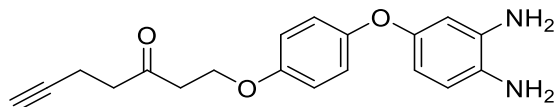
63: 1-(4-(3-Amino-4-nitrophenoxy)phenoxy)hept-6-yn-3-one



Reaction repeated as general procedure A using 1-(4-hydroxyphenoxy)hept-6-yn-3-one (**62**) gave the title compound (**63**) as a yellow oil (0.64 g, 56 %). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.92 (d, 1H, *J*= 12.0 Hz, Ar-H), 7.61 (d, 2H, *J*= 8.0 Hz, Ar-H), 6.88 (d, 2H, *J*= 8.0 Hz, Ar-H), 6.64 (d, 1H, *J*= 8.0 Hz, Ar-H), 6.32 (dd, 1H, *J*= 4.0, 8.0 Hz, Ar-H), 4.54 (t, 2H, *J*= 8.0, 16.0 Hz, CH₂), 3.32 (t, 2H, *J*= 8.0, 12.0 Hz, CH₂), 2.96-2.94 (m, 1H C-H), 2.55-2.50 (m, 2H, CH₂), 2.19-2.16 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 213.1, 168.7, 155.4, 151.2, 146.9, 128.4, 120.3,

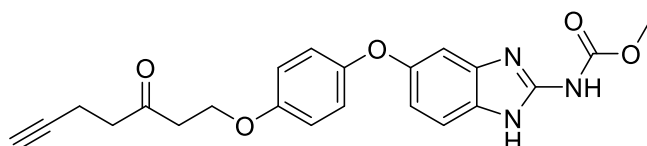
118.6, 116.9, 106.5, 102.8, 86.3, 70.4, 60.1, 43.0, 39.6, 12.1. HRMS (ES+) m/z calculated for $C_{19}H_{19}N_2O_5$: 355.1296. Found $[M+H]^+$: 355.1302 (Diff 1.69 ppm).

64: 1-(4-(3,4-Diaminophenoxy)phenoxy)hept-6-yn-3-one



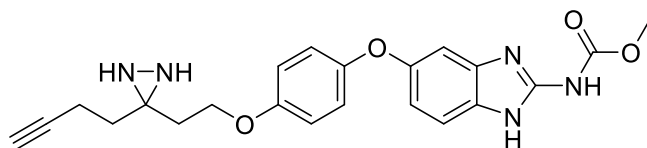
Reaction repeated using general procedure B using 1-(4-(3-amino-4-nitrophenoxy)phenoxy)hept-6-yn-3-one (**63**) gave the title compound (**64**) as a brown oil. This intermediate was not characterised and carried through to the subsequent reaction.

65: Methyl (5-(4-((3-oxohept-6-yn-1-yl)oxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



Reaction repeated as general procedure D using 1-(4-(3,4-diaminophenoxy)phenoxy)hept-6-yn-3-one (**64**) and 1,3-bis(methoxycarbonyl)-2-methylisothiourea (**12b**) gave the title compound (**65**) as a brown solid (0.18 g, 33 %). 1H NMR (400 MHz, DMSO- d_6) δ 7.72 (d, 1H, J = 8.0 Hz, Ar-H), 7.49 (d, 2H, J = 8.0 Hz, Ar-H), 6.74 (d, 2H, J = 8.0 Hz, Ar-H), 7.44 (d, 1H, J = 8.0 Hz, Ar-H), 6.22 (dd, 1H, J = 4.0, 8.0 Hz, Ar-H), 4.51 (t, 2H, J = 8.0, 12.0 Hz, CH_2), 3.72 (s, 3H, CH_3), 3.34 (t, 2H, J = 8.0, 12.0 Hz, CH_2), 2.98-2.94 (m, 1H, C-H), 2.49-2.46 (m, 2H, CH_2), 2.17-2.13 (m, 2H, CH_2). ^{13}C NMR (101 MHz, DMSO- d_6) δ 212.7, 157.4, 154.0, 151.8, 148.3, 144.9, 135.5, 130.1, 120.6, 117.2, 116.6, 106.0, 84.4, 70.2, 60.6, 51.7, 43.7, 39.4, 12.7. HRMS (ES+) m/z calculated for $C_{22}H_{22}N_3O_5$: 408.1559. Found $[M+H]^+$: 408.1544 (Diff -3.68 ppm).

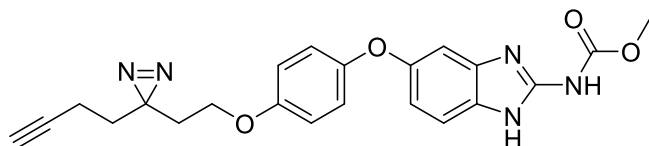
66: Methyl (5-(4-(2-(3-(but-3-yn-1-yl)diaziridin-3-yl)ethoxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



Reaction repeated as general procedure F using methyl (5-(4-((3-oxohept-6-yn-1-yl)oxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate (**65**) gave the title compound (**66**) as a

crude brown solid. This intermediate was not characterised and carried through to the subsequent reaction.

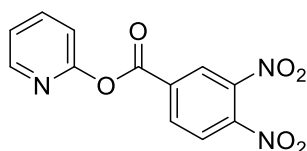
60b: Methyl (5-(4-(2-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)ethoxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate



Reaction repeated as general procedure G using methyl (5-(4-(2-(3-(but-3-yn-1-yl)diaziridin-3-yl)ethoxy)phenoxy)-1H-benzo[d]imidazol-2-yl)carbamate (**66**) gave the title compound (**60b**) as a brown solid (0.13 mg, 7%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.60 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.24 (d, 2H, *J* = 8.0 Hz, Ar-H), 7.11 (d, 2H, *J* = 8.0 Hz, Ar-H), 6.95 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.41 (dd, 1H, *J* = 4.0, 8.0 Hz, Ar-H), 4.31 (t, 2H, *J* = 8.0, 12.0 Hz, CH₂), 3.71 (s, 3H, CH₃), 2.87-2.84 (m, 1H, C-H), 2.51-2.48 (m, 2H, CH₂), 2.17-2.13 (m, 2H, CH₂); 1.74-1.70 (m, 2H, CH₂). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 155.8, 155.6, 153.3, 152.0, 149.0, 138.4, 130.2, 119.5, 116.2, 116.6, 111.2, 101.7, 105.8, 84.7, 69.8, 69.6, 55.5, 51.4, 50.2, 41.7, 13.0. HRMS (ES⁺) *m/z* calculated for C₂₂H₂₂N₅O₄: 420.1672. Found [M+H]⁺: 420.1678 (Diff 1.90 ppm).

3.6.2.6 Synthesis of Benzophenone Probe

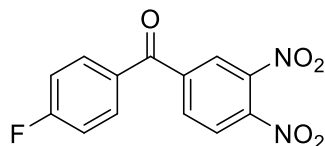
69: Pyridin-2-yl 3,4-dinitrobenzoate³¹⁹



3,4-Diaminobenzoic acid (**67**) (2 g, 9.43 mmol, 1 equiv.), di(pyridin-2-yl) carbonate (**68**) (2.04 g, 9.43 mmol, 1 equiv.) and DMAP (0.12 g, 0.94 mmol, 0.1 equiv.) were dissolved in anhydrous DCM (47.1 mL, 0.2 M) and allowed to stir at reflux for 2 days under a nitrogen atmosphere. After this time, the reaction mixture was washed with saturated aq. NaHCO₃ solution (x1), water (x1), and brine (x1), dried over MgSO₄, filtered, and concentrated under reduced pressure to yield a crude yellow oil. Purification *via* column chromatography eluting with 30-100% EtOAc in *n*-hexane gave the title compound (**69**) as a white solid (1.46 g, 56%). ¹H NMR (400 MHz, CDCl₃) δ 8.79 (m, 1H, Ar-H), 8.49 (d, 1H, *J* = 8.0 Hz, Ar-H), 8.42 (d, 1H, *J* = 8.0 Hz, Ar-H), 8.05 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.98 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.94-7.90 (m, 1H, Ar-H), 7.38-

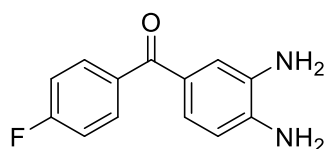
7.35 (m, 1H, Ar-H). ^{13}C NMR (101 MHz, CDCl_3) δ 161.0, 157.2, 148.9, 140.1, 135.2, 134.1, 127.0, 126.4, 125.5, 125.3, 123.1, 116.1. LRMS (CI+) m/z calculated for $\text{C}_{12}\text{H}_8\text{N}_3\text{O}_6$: 290.04. Found $[\text{M}+\text{H}]^+$: 290.04.

71: (3,4-Dinitrophenyl)(4-fluorophenyl)methanone³¹⁹



Pyridin-2-yl 3,4-dinitrobenzoate (**69**) (1.46 g, 5.05 mmol, 1 equiv.), 4-fluorophenyl boronic acid (**70**) (1.41 g, 10.10 mmol, 2 equiv.), and $\text{Pd}(\text{PPh}_3)_4$ (0.58 g, 0.51 mmol, 0.1 equiv.) were suspended in anhydrous 1,4-dioxane (10.1 mL, 0.14 M) in a screw fix sealed tube and allowed to stir at 50°C for 16 hours. After this time, the reaction was cooled and concentrated under reduced pressure. The residue was suspended in EtOAc and washed with saturated aq. NaHCO_3 solution (x1), water (x1), and brine (x1). The organic phase was dried over MgSO_4 , filtered, and concentrated *in vacuo* to yield a crude brown oil. Purification *via* column chromatography eluting with 0-40% EtOAc in *n*-hexane gave the title compound (**71**) as an orange oil (0.64 g, 44 %). ^1H NMR (400 MHz, CDCl_3) δ 8.52 (s, 1H, Ar-H), 8.35 (d, 1H, $J = 8.0$ Hz, Ar-H), 8.09 (d, 1H, $J = 8.0$ Hz, Ar-H), 7.92 (t, 1H, $J = 8.0, 16.0$ Hz, Ar-H), 7.65-7.58 (m, 2H, Ar-H), 7.52-7.50 (m, 1H, Ar-H). ^{13}C NMR (101 MHz, CDCl_3) δ 200.1, 165.3, 150.1, 143.6, 143.2, 136.2, 135.7, 130.8, 122.4, 124.1, 116.9. HRMS (CI+) m/z calculated for $\text{C}_{13}\text{H}_8\text{N}_2\text{O}_5^{19}\text{F}$: 291.0417. Found $[\text{M}+\text{H}]^+$: 291.0423 (Diff 2.06 ppm).

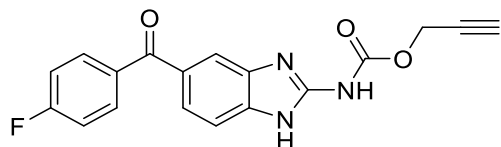
72: (3,4-Diaminophenyl)(4-fluorophenyl)methanone³³³



Reaction repeated as general procedure B using (3,4-dinitrophenyl)(4-fluorophenyl)methanone (**71**) (0.64 g) and double equivalents of $\text{SnCl}_2 \cdot \text{H}_2\text{O}$ (4.2 g, 10 equiv.) gave the title compound (**72**) as a brown oil (0.24 g, 47%). ^1H NMR (400 MHz, CDCl_3) δ 7.40-7.36 (m, 2H, Ar-H), 7.24 (s, 1H, Ar-H), 7.16-7.13 (m, 2H, Ar-H), 7.08-7.04 (m, 1H, Ar-H), 6.60-6.57 (m, 1H, Ar-H), 3.30 (br s, 4H, $(\text{NH}_2)_2$). ^{13}C NMR (101 MHz, CDCl_3) δ 199.9, 167.4, 139.7,

135.2, 133.8, 133.6, 133.1, 128.4, 116.5, 115.9, 115.4. LRMS (CI+, CH4) m/z calculated for C₁₃H₁₂N₂O¹⁹F: 231.09. Found [M+H]⁺: 231.09.

73: Prop-2-yn-1-yl (5-(4-fluorobenzoyl)-1H-benzo[d]imidazol-2-yl)carbamate



Reaction repeated as general procedure D using (3,4-diaminophenyl)(4-fluorophenyl)methanone (**72**) and 1,3-bis(prop-2-yn-1-yloxy carbonyl)-2-methylisothiourea (**12a**) gave the title compound (**73**) as a brown solid (60 mg, 17%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.92 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.91 (s, 1H, Ar-H), 7.44-7.40 (m, 2H, Ar-H), 7.25 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.18-7.14 (m, 2H, Ar-H), 4.75-4.71 (m, 2H, CH₂), 3.38-3.36 (m, 1H, C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 200.2, 166.5, 154.2, 139.6, 139.2, 137.4, 135.9, 130.5, 130.3, 126.9, 120.7, 115.0, 114.3, 80.1, 78.8, 55.0. HRMS (ES+) m/z calculated for C₁₈H₁₃N₃O₃¹⁹F: 338.0941. Found [M+H]⁺: 338.0935 (Diff -1.77 ppm). IR ν_{max}/cm⁻¹ (solid) 3321 (s), 2983 (m), 1733 (s), 1645 (s), 1462 (m), 1310 (s), 1288 (s). MP: 288-291 °C decomposed. Purity HPLC 95.6%, R_t = 9.9 min.

3.7 References

- 1 D. P. Murale, S. C. Hong, M. M. Haque and J. S. Lee, *Proteome Sci.*, 2017, **15**, 1–34.
- 2 R. Y. Tsien, *Annu. Rev. Biochem.*, 1998, **67**, 509–544.
- 3 J. Zhang, R. E. Campbell, A. Y. Ting and R. Y. Tsien, *Nat. Rev. Mol. Cell Biol.*, 2002, **3**, 906–918.
- 4 E. L. S. Costantini, Lindsey M., Matteo Fossati, Maura Francolini, *Traffic*, 2012, **13**, 643–649.
- 5 Thermofisher, <https://www.thermofisher.com/uk/en/home/life-science/cell-analysis/cell-analysis-learning-center/molecular-probes-school-of-fluorescence/imaging-basics/labeling-your-samples/different-ways-to-add-fluorescent-labels.html>, (Accessed April 2020).
- 6 Chromotek, <https://resources.chromotek.com/blog/green-fluorescent-proteins-tools>, (Accessed April 2020).
- 7 S. Bansal, S. Sur and V. Tandon, *Biochemistry*, 2019, **58**, 809–817.
- 8 S. Sharma and N. Anand, in *Approaches to Design and Synthesis of Antiparasitic Drugs*, Elsevier Science, Amsterdam, The Netherlands, 1997, pp. 195–238.
- 9 M. M. Barrowman, S. E. Marriner and J. A. Bogan, *Biosci. Rep.*, 1984, **4**, 879–883.
- 10 R. K. Prichard, *Int. J. Parasitol.*, 1973, **3**, 409–417.
- 11 A. R. Florio, S. Ferrari, E. De Carolis, R. Torelli, G. Fadda, M. Sanguinetti, D. Sanglard and B. Posteraro, *BMC Microbiol.*, 2011, **11**, 1–13.
- 12 L. Da, S. Derengowski, H. C. Paes, P. Albuquerque, A. H. F. P. Tavares, L. Fernandes, I. Silva-Pereira and A. Casadevall, *Eukaryot. Cell*, 2013, **12**, 761–774.
- 13 J. Brunner, *Annu. Rev. Biochem.*, 1993, **62**, 483–514.
- 14 A. L. MacKinnon and J. Taunton, *Curr. Protoc. Chem. Biol.*, 2009, **1**, 55–73.
- 15 Y. Hatanaka, E. Yoshida, H. Nakayama and Y. Kanaoka, *Bioorg. Chem.*, 1989, **17**, 482–485.
- 16 T. Hiramatsu, Y. Guo and T. Hosoya, *Org. Biomol. Chem.*, 2007, **5**, 2916–2919.
- 17 Z. Qiu, L. Lu, X. Jian and C. He, *J. Am. Chem. Soc.*, 2008, **130**, 14398–14399.
- 18 J. Das, *Chem. Rev.*, 2011, **111**, 4405–4417.
- 19 R. A. G. Smith and J. R. Knowles, *J. Am. Chem. Soc.*, 1973, **95**, 5072–5073.
- 20 Y. H. Hashimoto, in *Photoaffinity Labeling for Structural Probing Within Protein*, Springer Japan, Tokyo, 2017, pp. 45–78.
- 21 Z. Song and Q. Zhang, *Org. Lett.*, 2009, **11**, 4882–4885.
- 22 L. Wang, Y. Murai, T. Yoshida, A. Ishida, K. Masuda, Y. Sakihama, Y. Hashidoko, Y. Hatanaka and M. Hashimoto, *Org. Lett.*, 2015, **17**, 616–619.

- 23 M. Al-Omari, K. Banert and M. Hagedorn, *Angew. Chemie - Int. Ed.*, 2005, **45**, 309–311.
- 24 N. Burkard, T. Bender, J. Westmeier, C. Nardmann, M. Huss, H. Wieczorek, S. Grond and P. Von Zezschwitz, *European J. Org. Chem.*, 2010, **11**, 2176–2181.
- 25 M. H. Wright and S. A. Sieber, *Nat. Prod. Rep.*, 2016, **33**, 731–733.
- 26 A. E. Speers, G. C. Adam and B. F. Cravatt, *J. Am. Chem. Soc.*, 2003, **125**, 4686–4687.
- 27 G. Charron, M. M. Zhang, J. S. Yount, J. Wilson, A. S. Raghavan, E. Shamir and H. C. Hang, *J. Am. Chem. Soc.*, 2009, **131**, 4967–4975.
- 28 N. J. Agard, J. M. Baskin, J. A. Prescher, A. Lo and C. R. Bertozzi, *ACS Chem. Biol.*, 2006, **1**, 644–648.
- 29 D. Vellucci, A. Kao, R. M. Kaake, S. D. Rychnovsky and L. Huang, *J. Am. Soc. Mass Spectrom.*, 2010, **21**, 1432–1445.
- 30 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chemie Int. Ed.*, 2002, **41**, 2596–2599.
- 31 N. J. Agard, J. A. Prescher and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2004, **126**, 15046–15047.
- 32 M. Köhn and R. Breinbauer, *Angew. Chemie Int. Ed.*, 2004, **43**, 3106–3116.
- 33 C. A. Grice, K. L. Tays, B. M. Savall, J. Wei, C. R. Butler, F. U. Axe, S. D. Bembenek, A. M. Fourie, P. J. Dunford, K. Lundeen, F. Coles, X. Xue, J. P. Riley, K. N. Williams, L. Karlsson and J. P. Edwards, *J. Med. Chem.*, 2008, **51**, 4150–4169.
- 34 W. Wang, D. Kong, H. Cheng, L. Tan, Z. Zhang, X. Zhuang, H. Long, Y. Zhou, Y. Xu, X. Yang and K. Ding, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 4250–4253.
- 35 A. B. H. Blackwell, R. Frei, *Angew. Chemie - Int. Ed.*, 2012, **51**, 5226–5229.
- 36 S. A. Mavrova ATs, Yancheva D, Anastassova N, Anichina K, Zvezdanovic J, Djordjevic A, Markovic D, *Bioorg. Med. Chem.*, 2015, **23**, 6317–6326.
- 37 C. L. V. Malarthi Raghunath, *Int J Pharm Pharm Sci*, 2014, **6**, 372–375.
- 38 J. B. Wright, *Chem. Rev.*, 1951, **48**, 397–541.
- 39 N. K. M. Alamgir, D. Black, in *Bioactive Heterocycles III*, Springer-Verlag, Heidelberg Berlin, 2007, pp. 87–118.
- 40 K. Weinhardt, C. C. Beard, C. Dvorak, M. Marx, J. Patterson, A. Roszkowski, M. Schuler, S. H. Unger, P. J. Wagner, *J. Med. Chem.*, 1984, **27**, 616–627.
- 41 B. Combourieu, P. Besse, M. Sancelme, J.-P. Godin, A. Monteil, H. Veschambre and A.-M. Delort, *Appl. Environ. Microbiol.*, 2000, **66**, 3187–3193.
- 42 S. Ram, D. S. Wise, L. L. Wotring, J. W. McCall, *J. Med. Chem.*, 1992, **35**, 539–547.
- 43 J. W. Jin Zhang, Lei-lei Fu, Mao Tian, Hao-qiu Liu, Jing-jing Li, Yan Li, Jun He, Jian Huang, Liang Ouyang, Hui-yuan Gao, *Bioorg. Med. Chem.*, 2015, **23**, 976–984.

- 44 M. Srinivasan, S. Sankararaman, H. Hopf, I. Dix and P. G. Jones, *J. Org. Chem.*, 2001, **66**, 4299–4303.
- 45 B. Zeysing, C. Gosch and A. Terfort, *Org. Lett.*, 2000, **2**, 1843–1845.
- 46 I. Jain, R. Sharma and P. Malik, *Synth. Commun.*, 2019, **42**, 2952–2960.
- 47 H. J. Jeong, E. Y. Yoon, M. K. Kim, J. H. Lee, Y. J. Yoon and S. G. Lee, *Bull. Korean Chem. Soc.*, 2003, **24**, 1689–1691.
- 48 T. J. Wallace and J. J. Mahon, *J. Am. Chem. Soc.*, 1964, **86**, 4099–4103.
- 49 S. R. A. François Vibert, Marque, E. Bloch and E. Besson, *J. Phys. Chem.*, 2015, **119**, 5434–5439.
- 50 E. B. Hotelling, R. J. Windgassen, E. P. Previc and M. B. Neuwirth, *J. Org. Chem.*, 1959, **24**, 1598–1600.
- 51 R. A. Moss, J. M. Fede and S. Yan, *J. Am. Chem. Soc.*, 2000, **122**, 9878–9879.
- 52 M. Hashimoto, Y. Murai, G. D. and Y. Hatanak, in *Hydrogenation*, InTech, 2012, vol. i, pp. 121–136.
- 53 J. R. Hill and A. A. B. Robertson, *J. Med. Chem.*, 2018, **61**, 6945–6963.
- 54 F. M. Müskens, R. J. Ward, D. Herkt, H. van de Langemheen, A. B. Tobin, R. M. J. Liskamp and G. Milligan, *Mol. Pharmacol.*, 2019, **95**, 196–209.
- 55 B. Zhu, H. Zhang, S. Pan, C. Wang, J. Ge, J. S. Lee and S. Q. Yao, *Chem. - A Eur. J.*, 2016, **22**, 7824–7836.
- 56 F. Beiroth, T. Koudelka, T. Overath, S. D. Knight, A. Tholey and T. K. Lindhorst, *Beilstein J. Org. Chem.*, 2018, **14**, 1890–1900.
- 57 L. S. Monteiro, F. Paiva-Martins, S. Oliveira, I. Machado and M. Costa, *Bioorg. Chem.*, 2019, **89**, 102983.
- 58 S. M. Pauff and S. C. Miller, *J. Org. Chem.*, 2013, **78**, 711–716.
- 59 B. Imperiali and R. H. Abeles, *Tetrahedron Lett.*, 1986, **27**, 135–138.
- 60 B. N. M. Kolb, J. Barth, *Tetrahedron Lett.*, 1986, **27**, 1579–1582.
- 61 Y. Tanaka, T. Ishihara and T. Konno, *J. Fluor. Chem.*, 2012, **137**, 99–104.
- 62 Y. W. Cheng, Huicheng, Yu Pei, Faqiang Leng, Jingya Li, Apeng Liang, Dapeng Zou, Yangjie Wu, *Tetrahedron Lett.*, 2013, **54**, 4483–4486.
- 63 R. J. Linderman and D. M. Graves, *J. Org. Chem.*, 1989, **54**, 661–668.
- 64 S. D. Meyer and S. L. Schreiber, *J. Org. Chem.*, 1994, **59**, 7549–7552.
- 65 K. Mori, N. Umehara and T. Akiyama, *Adv. Synth. Catal.*, 2015, **357**, 901–906.
- 66 F. A. Luzzio, R. W. Fitch, W. J. Moore and K. J. Mudd, *J. Chem. Educ.*, 1999, **76**, 974–975.
- 67 H. Jiang, T.-Y. Sun, X. Wang, Y. Xie, X. Zhang, Y.-D. Wu and H. F. Schaefer, *Org. Lett.*,

- 2017, **19**, 6502–6505.
- 68 C. G. Savarin, C. Grisé, J. A. Murry, R. A. Reamer and D. L. Hughes, *Org. Lett.*, 2007, **9**, 981–983.
- 69 A. N. Butkevich, A. Corbu, L. Meerpoel, I. Stansfield, P. Angibaud, P. Bonnet and J. Cossy, *Org. Lett.*, 2012, **14**, 4998–5001.
- 70 L. Wang, T. Yoshida, Y. Muto, Y. Murai, Z. P. Tachrim, A. Ishida, S. Nakagawa, Y. Sakihama, Y. Hashidoko, K. Masuda, Y. Hatanaka and M. Hashimoto, *European J. Org. Chem.*, 2015, **2015**, 3129–3134.
- 71 Y. A. Tsegaw, P. E. Kadam, N. Tötsch, E. Sanchez-Garcia and W. Sander, *J. Am. Chem. Soc.*, 2017, **139**, 12310–12316.
- 72 W. W. L. Cheng, Z.-W. Chen, J. R. Bracamontes, M. M. Budelier, K. Krishnan, D. J. Shin, C. Wang, X. Jiang, D. F. Covey, G. Akk and A. S. Evers, *J. Biol. Chem.*, 2018, **293**, 3013–3027.
- 73 S. S. Ge, B. Chen, Y. Y. Wu, Q. S. Long, Y. L. Zhao, P. Y. Wang and S. Yang, *RSC Adv.*, 2018, **8**, 29428–29454.
- 74 Z. Li, P. Hao, L. Li, C. Y. J. Tan, X. Cheng, G. Y. J. Chen, S. K. Sze, H. M. Shen and S. Q. Yao, *Angew. Chemie - Int. Ed.*, 2013, **52**, 8551–8556.
- 75 M. A. Ibrahim, H. W. B. Johnson, J. W. Jeong, G. L. Lewis, X. Shi, R. T. Noguchi, M. Williams, J. W. Leahy, J. M. Nuss, J. Woolfrey, M. Banica, F. Bentzien, Y. Chou, A. Gibson, N. Heald, P. Lamb, L. Mattheakis, D. Matthews, A. Shipway, X. Wu, W. Zhang, S. Zhou and G. Shankar, *J. Med.*, 2012, **55**, 1368–1381.
- 76 M. Hashimoto, Y. H. Kato and Y. Hatanaka, *Chem. Pharm. Bull.*, 2007, **55**, 1540–1543.
- 77 R. V. Stevens and A. W. M. Lee, *J. Am. Chem. Soc.*, 1979, **101**, 7032–7035.
- 78 L. Wang, Z. P. Tachrim, N. Kurokawa, F. Ohashi, Y. Sakihama, Y. Hashidoko and M. Hashimoto, *Molecules*, 2017, **22**, 2–9.
- 79 N. S. Kumar and R. N. Young, *Bioorganic Med. Chem.*, 2009, **17**, 5388–5395.
- 80 G. D. Prestwich, G. Dormán, J. T. Elliott, D. M. Marecak and A. Chaudhary, *Photochem. Photobiol.*, 1997, **65**, 222–234.
- 81 M. M. Hassan and O. O. Olaoye, *Molecules*, 2020, **25**, 2285.
- 82 D. J. Lapinsky, *Bioorg. Med. Chem.*, 2012, **20**, 6237–6247.
- 83 H. Tatamidani, F. Kakiuchi and N. Chatani, *Org. Lett.*, 2004, **6**, 3597–3599.
- 84 K. Kawada, E. K. Dolence, H. Morita, L. E. Gerschenson, T. Kometani, D. S. Watt, D. J. Orlicky, T. A. Fitz and A. Balapure, *J. Med. Chem.*, 1989, **32**, 256–264.
- 85 C. Feng, D. Chan, J. Joseph, M. Muuronen, W. H. Coldren, N. Dai, I. R. Corrêa, F. Furche, C. M. Hadad and R. C. Spitale, *Nat. Chem. Biol.*, 2018, **14**, 276–283.
- 86 D. Szymanski, M. Papanastasiou, L. Pandarinathan, N. Zvonok, D. R. Janero, S. Pavlopoulos, P. Vouros and A. Makriyannis, *J. Med. Chem.*, 2018, **61**, 11199–11208.

- 87 J. Liang, L. Zhang, X. L. Tan, Y. K. Qi, S. Feng, H. Deng, Y. Yan, J. S. Zheng, L. Liu and C. L. Tian, *Angew. Chemie - Int. Ed.*, 2017, **56**, 2744–2748.
- 88 A. Ajay, V. Singh, S. Singh, S. Pandey, S. Gunjan, D. Dubey, S. Kumar, B. N. Singh, V. Chaturvedi, R. Tripathi, R. Ramchandran and R. P. Tripathi, *Bioorg. Med. Chem.*, 2010, **18**, 8289–8301.
- 89 M. A. Ibrahim, H. W. B. Johnson, J. W. Jeong, G. L. Lewis, X. Shi, R. T. Noguchi, M. Williams, J. W. Leahy, J. M. Nuss, J. Woolfrey, M. Banica, F. Bentzien, Y.-C. Chou, A. Gibson, N. Heald, P. Lamb, L. Mattheakis, D. Matthews, A. Shipway, X. Wu, W. Zhang, S. Zhou and G. Shankar, *J. Med. Chem.*, 2012, **55**, 1368–1381.
- 90 E. C. Cortés, R. S. Mendoza, M. S. Gutiérrez and O. G.-M. De Cortés, *J. Heterocycl. Chem.*, 2004, **41**, 273–276.
- 91 C. C. Huang, C. H. Lin and S. A. Dai, *RSC Adv.*, 2015, **5**, 74874–74880.
- 92 V. M. Lau, W. C. Pfalzgraff, T. E. Markland and M. W. Kanan, *J. Am. Chem. Soc.*, 2017, **139**, 4035–4041.
- 93 W. Zhang, Z. Li, M. Zhou, F. Wu, X. Hou, H. Luo, H. Liu, X. Han, G. Yan, Z. Ding and R. Li, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 799–807.
- 94 C. Muanming, WO2017173999, 2017, 64.
- 95 C. J. Peto, WO2016187620, 2016, 114.
- 96 S. M. Burns, WO2018175324, 2018, 189.
- 97 L. Pu-rui, L. Min, L. Hong and N. Bin-ke, (*Xi'an Mod. Chem. Res. Institute, Xi'an 710065, China*).

CHAPTER 4

Introduction to Triple Negative Breast Cancer,
the S100A4 Protein and Proteolysis Targeting
Chimera

4.1 Introduction to Breast Cancer

Breast cancer is the most common malignancy in women worldwide, in 2018 2.1 million women were diagnosed.¹ It has been reported that one in four women are diagnosed with breast cancer at some point in their life.² Approximately 627,000 women in the UK died of breast cancer in 2018, accounting for 15% of all cancer deaths among women.¹ Breast cancer can occur in men; however, it is rare, effecting less than 1% of all male cancer cases.³ Breast cancer can be a highly curable disease when detected early. Devastatingly, detection of late grade cancer can often lead to death. Access to high-quality medical care leading to early diagnosis and proper surgical and medical treatment can mean the difference between life and death.⁴

4.1.1 Risk Factors

There are many risk factors for the development of breast cancer. The main risk factor is age, as 80% of all breast cancer cases occur in women between 50-69 years of age.⁵ Low rates of breastfeeding can also lead to an increased risk in breast cancer. This may contribute to the elevated incidence rates in high-income countries, as artificial breast milk is more available and favoured in these regions.⁶ Other risk factors include alcohol and tobacco consumption,⁶ ⁷ prolonged use of the oral contraceptive pill,⁸ ⁹ exposure to menopausal hormone therapy,¹⁰ menstrual and reproductive history,¹¹ obesity,¹² and race/ ethnicity.¹³ Family history of breast cancer is also a major contributing factor, which can be linked to predisposing genes such as BRCA1 and BRCA2. These genes can make women susceptible to the development of breast cancer,¹⁴ as they are responsible for 90% of all hereditary breast cancer cases.^{15, 16}

The chances of a woman developing breast cancer and their survival can be determined by their country of residence and socioeconomic status. Breast cancer occurs globally in both developing and developed countries, however there is a higher incidence in developed European countries. In 2018 the top three countries with the highest incidence of breast cancer were Belgium (113.2 women per 100,000), Luxembourg (109.3 women per 100,000) and Netherlands (105.9 women per 100,000). The UK has a rate of 93.6 per 100,000 and the US has a rate of 84.9 per 100,000.¹⁷ Breast cancer is more likely to be diagnosed in developed countries, as the public are more aware of the risk factors. Women in high-income countries have access to advanced medical care, which can lead to earlier diagnosis and treatment for breast cancer. For example, women 55 years of age and over have regular mammogram

screenings to detect abnormalities in the breast tissue. These favourable factors have led to a fall in both incidence and mortality rates since the year 2000. Even though incidence rates are lower in the developing countries, the mortality rate can remain high due to lack of medical resources. The incidence and subsequent mortality rates are increasing in low income countries due to changes in reproductive patterns in women, and delays in both diagnosis and treatment.^{18, 19}

A comprehensive study by *Hu et al* has explored global trends of incidence and mortality rates to show the inequalities in diagnosis and survival. The institute for health metrics and evaluation introduced the Sociodemographic Index (SDI) in the Global Burden of Disease study 2015 (GBD 2015). The SDI is a quantitative measure of development in a country or region. The paper describes current patterns and trends in breast cancer incidence and mortality in countries across the globe with data collected between 1990-2016 (see **Figure 4.1**). The SDI evaluated 195 countries using the global health data exchange database to classify countries according to their development as high, high-middle, middle, middle-low, and low. The study found that the age standardised incidence rate (ASIR) in high and high-middle countries plateaued and decreased after the rapid increase in the early 1990's, with the high-income category decreasing 0.1% per year since 2000. The ASIR increased in the middle, middle-low-, and low-income countries during the whole study. The average annual percentage change (AAPC) in ASIR in the middle-income category was an alarming 2.1%. The age standardised mortality rate (ASMR) decreased in high-income SDI category from 24.2% in 1990 to 17.6% in 2016. The high-middle SDI also experienced a decline in 1994, with an accelerated decrease between 2004-2016. The middle SDI also declined, but less so between 2002-2016 with an average decrease of 0.5% per year. Unfortunately, the decrease in ASMR stopped when we reach the low-middle and low-income SDIs. In the low-middle category ASMR rose by 0.7% between 2002-2016 and low SDI increased by 0.8% between 2009-2016. These trends indicate that incidence and mortality rates are declining in the developed world. Unfortunately, the rates continue to rise in the developing world. This demonstrates that a huge effort to improve health care in middle to low-income countries is necessary. Development of new therapeutic agents and identification of new biomarkers for early detection are urgently required.²⁰

	1990			2016			Trend 1			Trend 2			Trend 3		
	Case	ASR	95% UI	Case	ASR	95% UI	Period	APC (%)	Period	APC (%)	Period	APC (%)	Period	APC (%)	AAPC (%)
Global	815.4	41.0	39.8-43.1	1681.9	45.6	43.6-48.2	1990-1995	1.5*	1995-2000	0.2*	2000-2016	0.1*	2000-2016	0.1*	0.4*
High SDI	463.0	83.6	82.0-85.1	719.4	88.9	86.5-93.0	1990-1995	1.3*	1995-2000	0.2*	2000-2016	-0.1*	2000-2016	-0.1*	0.2*
High-middle SDI	153.1	37.1	36.0-38.6	329.0	46.5	42.9-50.5	1990-1995	3.2*	1995-2016	0.4*					0.9*
Middle SDI	116.0	19.3	18.1-22.1	408.8	33.2	30.4-36.0	1990-2000	2.6*	2000-2009	2.0*	2009-2016	1.5*	2009-2016	1.5*	2.1*
Low-middle SDI	69.7	17.7	14.9-21.9	187.0	23.1	21.1-27.8	1990-1999	1.3*	1999-2010	0.4*	2010-2016	1.8*	2010-2016	1.8*	1*
Low SDI	15.6	16.3	13.6-20.9	37.0	18.8	17.1-20.8	1990-1995	0.9*	1995-2010	0.3*	2010-2016	0.8*	2010-2016	0.8*	0.5*

	1990			2016			Trend 1			Trend 2			Trend 3		
	Case	ASR	95% UI	Case	ASR	95% UI	Period	APC (%)	Period	APC (%)	Period	APC (%)	Period	APC (%)	AAPC (%)
Global	336.9	17.2	16.4-18.8	535.3	14.6	13.8-15.6	1990-1994	0.8*	1994-2002	-0.6*	2002-2016	-1.1*	2002-2016	-1.1*	-0.7*
High SDI	141.1	24.2	23.8-24.7	162.1	17.6	16.9-18.3	1990-1995	-0.6*	1995-2010	-1.6*	2010-2016	-1*	2010-2016	-1*	-1.3*
High-middle SDI	65.5	15.9	15.2-17.0	97.9	13.7	12.2-15.7	1990-1994	2.9*	1994-2004	-0.4*	2004-2016	-1.9*	2004-2016	-1.9*	-0.6*
Middle SDI	65.5	11.6	10.7-13.7	138.5	11.8	10.8-12.7	1990-2002	0.8*	2002-2016	-0.5*					0.1*
Low-middle SDI	50.7	13.6	11.4-17.3	104.5	13.8	12.1-17.2	1990-1999	0.7*	1999-2012	-0.6*	2012-2016	0.7*	2012-2016	0.7*	0
Low SDI	13.9	15.7	13.0-20.2	31.9	17.6	15.7-19.9	1990-1996	0.9*	1995-2009	0	2009-2016	0.8*	2009-2016	0.8*	0.5*

Figure 4.1: Breast cancer age-standardised rates. **4.1A:** Breast cancer standardised incidence rates in 1990, 2016 and joint point trend analysis between 1990 and 2016 y SDI settings. **4.1B:** Breast cancer standardised mortality rates in 1990, 2016 and joint point trend analysis between 1990 and 2016 y SDI settings. Reproduced from: *Global patterns and trends in the breast cancer incidence and mortality according to sociodemographic indices: an observational study based on the global burden of diseases*, K. Hu et al.²⁰

*P<0.06

AAPC: Average annual percent change; APC: Annual percent change; ASR: Age-standardised rate; SDI: Sociodemographic index; 95% UI, 95% Uncertainty interval

4.1.2 Classification

There are three broad subtypes in the biological classification of breast cancer: oestrogen and progesterone receptor positive (ER/ PR), human epidermal growth factor receptor 2 (HER2) positive, and triple negative breast cancer (TNBC). The prognosis, treatment regimen, and survival outcome depend on the type of tumour/s the patient has.²¹

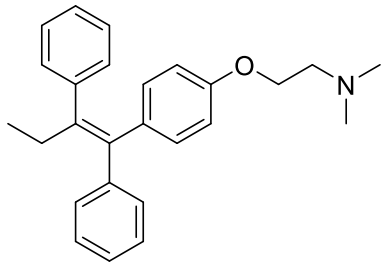
4.2.1 Oestrogen Receptor Positive and Progesterone Receptor Positive Cancer

Oestrogen receptors (ER) and progesterone receptors (PR) can be found inside breast cancer cells that rely on oestrogen and progesterone to grow. Breast cancer cells which express oestrogen receptors are referred to as ER-positive tumours. Similarly, if the breast cancer cells have progesterone receptors, they are called PR-positive tumours. Approximately two thirds of all breast cancers cases are described as ER and/or PR positive.²²

4.2.1.1 Treatment of Oestrogen Receptor Positive and Progesterone Receptor Positive Cancer

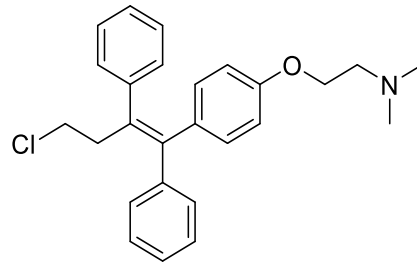
Tumours expressing oestrogen and/or progesterone receptors can be treated with hormonal therapy, which is able to block the binding of oestrogen or progesterone onto the respective receptors. The inhibition of the hormone receptors (HR) slows or completely hinders tumour growth. The prevention of hormone binding can be achieved in several ways.²³ Selective oestrogen-receptor response modulators (SERMS) such as tamoxifen (see **Figure 4.2A**) and toremifene (see **Figure 4.2B**) can bind to oestrogen receptors, rather than oestrogen, which slows the growth of breast cancer cells.^{24, 25, 26, 27} Aromatase inhibitors such as formestane (see **Figure 4.2C**) and anastrozole (see **Figure 4.2D**) treat postmenopausal women by inhibiting the synthesis of oestrogen. Aromatase inhibitors bind to the crucial aromatase enzyme, blocking the conversion of androgens to oestrogen.^{27, 28} Oestrogen-receptor down regulators (ERDs) such as fulvestrant (see **Figure 4.2E**) treat advanced stage HR positive metastatic breast cancer in postmenopausal women, when alternative endocrine therapy such as tamoxifen are no longer effective. ERDs are antagonists which bind to the oestrogen receptor to block hormone effects on cancer cells. Additionally, they exert selective oestrogen receptor degradation *via* intracellular proteolysis by causing receptor misfolding.^{27, 29, 30} Lastly, luteinizing hormone-releasing hormone agents (LHRHs) such as goserelin (see **Figure 4.2F**) suppresses the production of oestrogen by downregulating the production of luteinizing hormone (LH) in the pituitary gland. This limits the growth and proliferation of cancer cells.²⁷ Structures of the common endocrine therapies are shown in **Figure 4.2**.

4.2A



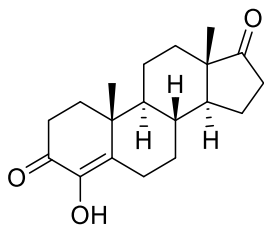
Tamoxifen

4.2B



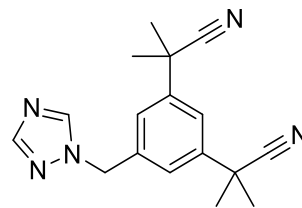
Toremifene

4.2C



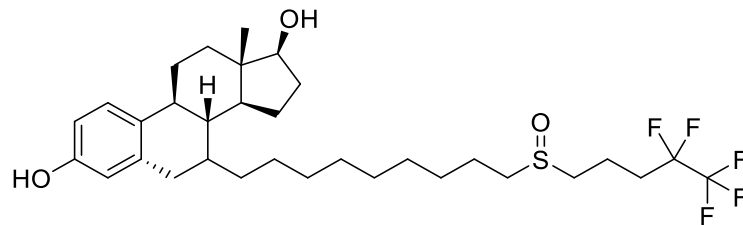
Formestane

4.2D



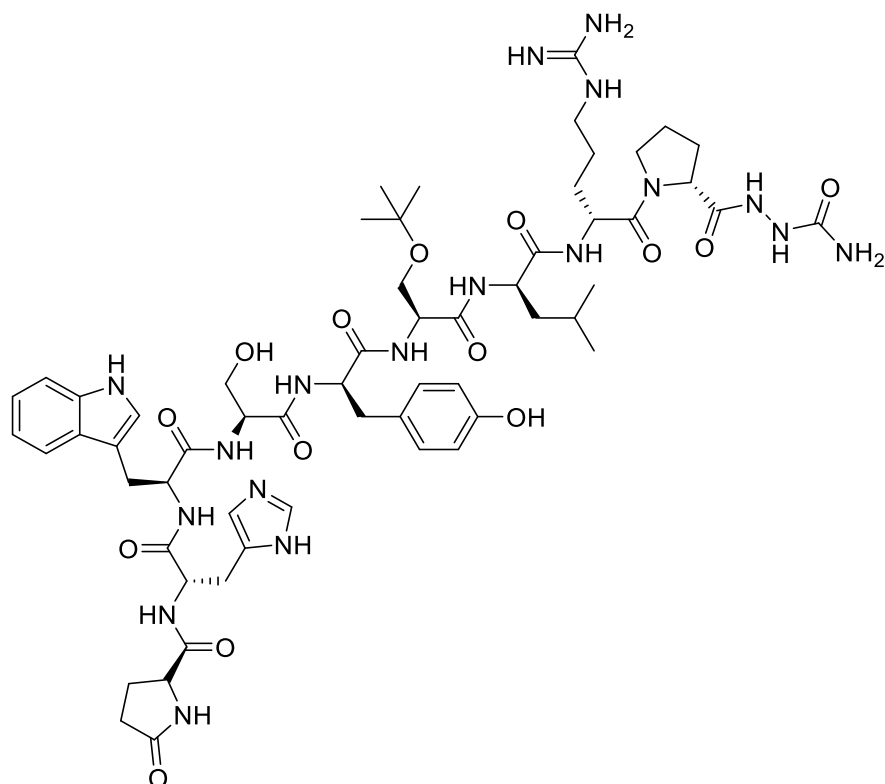
Anastrozole

4.2E



Fulvestrant

4.2F



Goserelin

Figure 4.2: Structures of common endocrine therapy treatments for HR positive breast cancer. **4.2A:** Tamoxifen, **4.2B:** Toremifene, **4.2C:** Formestane, **4.2D:** Anastrozole, **4.2E:** Fulvestrant and **4.2F:** Goserelin.

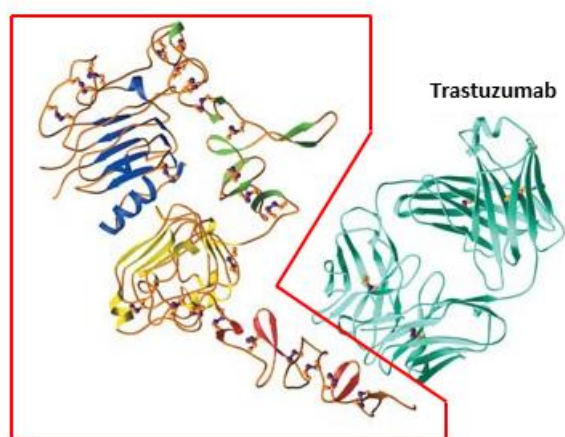
4.2.2 HER2 Positive Cancer

HER2 proteins are transmembrane growth factor receptors that activate intracellular signalling pathways in response to extracellular signals. In normal breast cancer cells HER2 helps healthy tissue grow, divide, and repair itself. HER2 positive breast cancer involves cells that over express HER2 and account for 1 in 5 of all breast cancers. Up-regulation of HER2 contributes to tumour progression by increasing proliferation and survival of the primary and secondary tumours caused by metastasis. Targeted therapy for HER2 positive cancer is available, which many patients respond well to.^{31, 32, 33}

4.2.2.1 Treatment of HER2 Positive Cancer

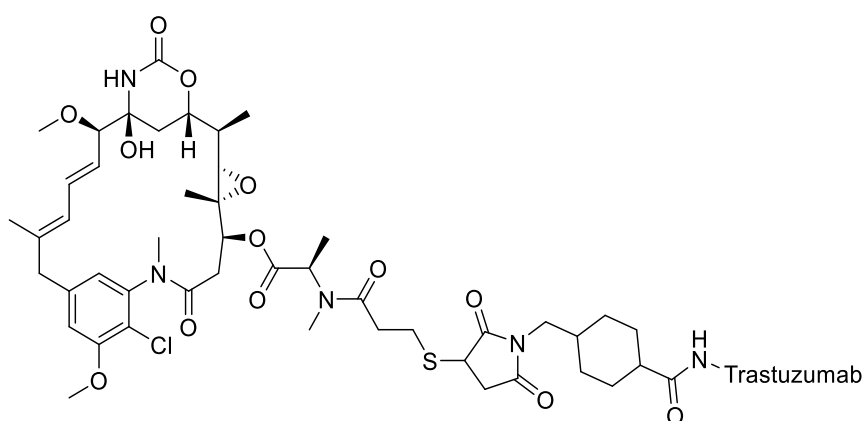
There are several targeted treatments available for the treatment of HER2 positive breast cancer: trastuzumab (see **Figure 4.3A**), trastuzumab emtansine (see **Figure 4.3B**) and neratinib (see **Figure 4.3C**). Trastuzumab otherwise known as Herceptin is an effective treatment which operates *via* three main mechanisms of action: **(1)** HER2 degradation in which binding to the receptor triggers selective intracellular proteolysis.^{34, 35} **(2)** Antibody dependent cellular cytotoxicity, which attracts immune cells to the HER2 positive tumours allowing them to lyse and destroy breast cancer cells.³⁴ **(3)** Interfering with the essential MAPK and P13K/Akt pathway, which is involved in promoting division and development of cells.^{34 36} Trastuzumab emtansine (kadcyla) is an antibody-drug conjugate. Trastuzumab is covalently attached to the cytotoxic agent DM1. These drugs are combined to allow trastuzumab to carry DM1 successfully to HER2 positive cancer cells. Kadcyla enters cells *via* receptor mediated internalisation and then digested by lysosomes. The metabolic products of DM1 are released and able to cause mitotic arrest by binding to the microtubule subunit tubulin, preventing cellular division.^{37, 38} Lastly, neratinib is an able to inhibit both HER2 and epidermal growth factor receptor (EGFR) kinases by covalently binding to the cysteine side chain in the proteins. This inhibits the cancer cells ability to receive growth signals, preventing proliferation.^{37 39 40} Structures of the common HER2 positive therapies are shown in **Figure 4.3**.

4.3A



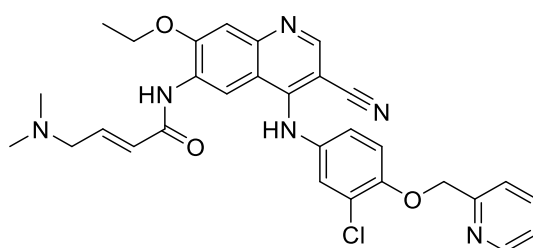
Ribbon Diagram of HER2 (highlighted in red box) and Trastuzumab (cyan) complex. Adapted from: *Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab*, H. Ch et al.⁴¹

4.3B



Trastuzumab emtansine

4.3C



Neratinib

Figure 4.3: Structures of common HER2 positive targeted therapies. **4.3A:** Trastuzumab, **4.3B:** Trastuzumab emtansine and **4.3C:** Neratinib.

4.2.3 Triple Negative Breast Cancer

Triple-negative breast cancer are cells that do not possess oestrogen and progesterone receptors nor have an excess of the HER2 protein. Approximately 10-20% of all breast cancer cases are described as TNBC. The lack of targets means that hormonal therapy and selective HER2 therapies will not be effective to obstruct the growth and proliferation of these cancer cells. Patients suffering with TNBC must rely on surgery, radiotherapy, and chemotherapy, involving aggressive cytotoxic agents to survive.^{42, 43}

4.2.3.1 Closer Examination into TNBC

Patients suffering from TNBC have the poorest prognosis due to the cancers being higher grade, basal like and more invasive. TNBC tends to be highly metastatic, which leads to poor patient survival as patients do not die from the primary tumour, as this can be surgically removed, but from its metastasis to other organs.^{43, 44}

Most deaths occur in the first five years following initial diagnosis. The five-year survival rate of TNBC is 77%, which is significantly lower compared to HR positive and HER2 positive cancers, which has a 5-year survival rate of 93%.⁴⁵ As well as this, TNBC is associated with higher recurrence rates compared to other breast cancers,⁴⁶ as the early peak of distance recurrences is at 3 years after diagnosis.⁴⁵ The survival rate of TNBC is lower than other subtypes as they tend to be discovered at a later stage and higher grade. TNBC is more common in younger women who do not qualify for routine mammograms; therefore, the cancer can metastasise before it is discovered. The lack of selective treatment to treat TNBC plays a significant role.⁴⁷

There are a few risk factors which can lead to individuals developing TNBC. Most breast cancer cases affect 50-69-year-old women, however, TNBC is common in younger women below the age of 50. Race also plays an important part as African American and Hispanic women tend to be diagnosed more frequently than Asian and non-Hispanic white women. TNBC is also often associated with hereditary conditions compared to other breast cancer subtypes. 70% of cancer patients which have the inherited BRCA mutation (particularly BRCA1) are triple negative.^{42, 46, 48}

4.2.3.2 Current Treatment of TNBC

Due to the lack of targets on the breast cancer cells of TNBC, selective treatments are not yet available. Treatment involves a combination of surgery, radiation therapy, and chemotherapy.^{49, 50}

There are two broad types of chemotherapy to treat cancer, which are neoadjuvant and adjuvant therapy. Neoadjuvant therapy involves treating the patient with chemotherapy to shrink the tumour/s before surgery to make the operation easier and more successful. Neoadjuvant therapy is also used to conserve the breast tissue to prevent the need for a mastectomy. Adjuvant therapy is administered after surgery to suppress secondary tumour formation, reduce the risk of recurrence and to destroy any cancer cells that might remain from the primary tumour after surgery. Both neoadjuvant and adjuvant therapy rely on chemotherapy and use a variety of cytotoxic agents such as platinum based agents, anthracyclines, taxanes, PARP inhibitors, and immunotherapy.^{49, 50}

4.2.3.2.1 Platinum Agents

Platinum agents such as cisplatin (see **Figure 4.4A**) and carboplatin (see **Figure 4.4B**) have a four-step mechanism of action: (i) cellular uptake, (ii) aquation/ activation, (iii) DNA platinumation and (iv) cellular processing of Pt-DNA. Platinum agents enter the cells *via* active transporters such as the copper transporters CTR1 and CTR2, which have been identified as the major access points.⁵¹ Once the platinum agent has entered the cell, it can be activated. Aquation of the platinum agents with water molecules is necessary to form an activated complex. The aquation rates for carboplatin are slower than cisplatin, which results in carboplatin being less potent than cisplatin.⁵² Once aquated/ activated, the complexes can react with nucleophilic centres on purine bases of DNA such as guanosine and adenosine. The two labile coordination sites on the platinum centre allow crosslinking of adjacent guanine bases.⁵³ Crosslinking distorts the DNA double helix, causing DNA lesions which can be recognised by cellular machinery. Platinum agents can initiate apoptosis by inhibiting transcription. Transcription of DNA by RNA polymerase can halt at the platinum cross-link. Recognition of the DNA lesion results in the recruitment of the transcription-coupled repair machinery. If the machinery is unable to repair the lesion, apoptosis is initiated.⁵⁴

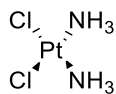
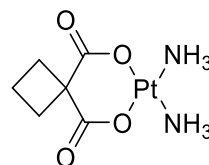
4.4A**4.4B**

Figure 4.4: Structures of common platinum agents. **4.4A:** Cisplatin, **4.4B:** Carboplatin.

4.2.3.2.2 Anthracyclines

A popular example of an anthracycline drug is doxorubicin (see **Figure 4.5**). Doxorubicin has two main mechanisms of action. The first is the intercalation into DNA and disrupting topoisomerase-II-mediated DNA repair. Topoisomerase II is an enzyme which relaxes supercoiling in DNA during transcription. Disruption of this process leads to irreparable DNA damage and results in cell death. The second mechanism is the generation of free radicals causing cytotoxicity and damaging cellular membranes, DNA, and the cells proteins.^{42, 49, 55, 56}

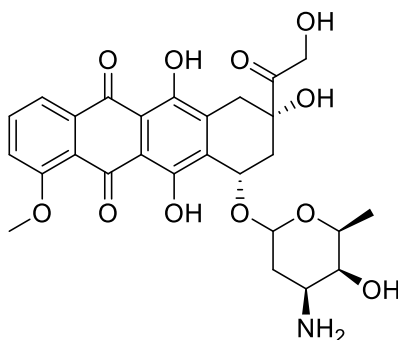


Figure 4.5: Structure of the anthracycline doxorubicin.

4.2.3.2.3 Taxanes

A common example of a taxane drug is paclitaxel (see **Figure 4.6**). Paclitaxel binds to tubulin subunits, causing microtubule stabilisation by promoting polymerisation while stopping depolymerisation. This stabilisation leads to the arrest of the cell cycle during its G1 or M phase of mitosis.⁵⁷ As a result, proliferation is halted and cell death occurs.^{49, 58}

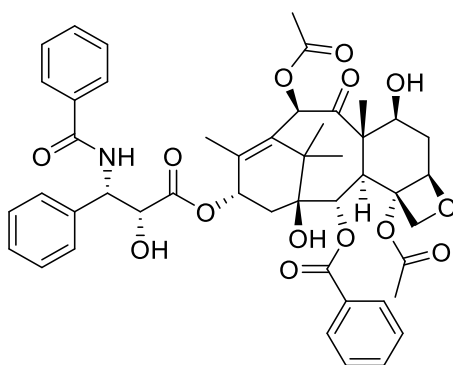


Figure 4.6: Structure of the taxane paclitaxel.

4.2.3.2.4 PARP Inhibitors

An example of a poly ADP-ribose polymerase (PARP) inhibitor is olaparib (see **Figure 4.7**). Olaparib is an FDA approved drug to treat advanced stage breast cancer. It is effective in metastatic patients possessing the BRCA mutation with TNBC. Olaparib inhibits the PARP enzyme which is essential in DNA repair, genomic stability, and apoptosis. Studies have indicated that drugs which interfere or inhibit PARP make it harder for cancer cells with the BRCA mutations to repair DNA damage. Therefore, cancer cells are less likely to survive DNA lesions inflicted by cytotoxic agents.^{42, 49, 59}

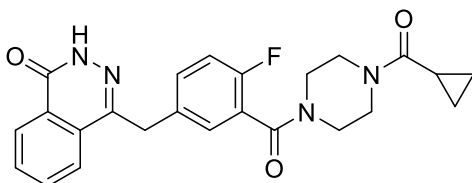


Figure 4.7: Structure of the PARP inhibitor olaparib.

4.2.3.2.5 Immunotherapy

Immunotherapy is a treatment which aids the patient's immune system to detect cancer for immune attack. A common agent is the monoclonal antibody atezolizumab, which displays good clinical activity in TNBC.⁶⁰ It blocks the interaction between PD-L1 and programme cell death protein 1 (PD-1). PD-L1 can allow cells to evade the hosts immune response, particularly the cytotoxic T cells. Some tumours can over express PD-L1 suppressing the immune response further, so the cancer cells can survive and proliferate. Inhibition of PD-L1

stops immune suppression, allowing the immune system to detect and destroy cancer cells.^{42, 49, 61}

4.2.3.3 Treatment Regimens

At present, there is no gold standard combination of cytotoxic agents for the treatment of TNBC, however, anthracycline based therapy is highly effective. A combination of adjuvant anthracycline and taxane seem to be beneficial to TNBC patients, as they can reduce the mortality rate by a third.^{62, 63, 64, 65} A retrospective study by *Wang et al* analysed the use of neoadjuvant anthracycline-taxane based therapy in patients with TNBC. TNBC patients had significantly higher pCR rate of 38% compared to non-TNBC patients which had a pCR rate of 12%.^{66, 67} Anthracyclines combined with antimetabolites such as 5-Fluorouracil (5-FU) (**4.8A**) (see **Figure 4.8**) have also proven to be effective against TNBC. *Tourneau et al* reported that response rates in patients with TNBC can be increased by combining anthracyclines with 5-FU in an intensive neoadjuvant chemotherapy with a dose of 700 mg m⁻². The pCR rate increased from 13% to 47% from the regimen.⁶⁸ Taxanes are highly beneficial in the treatment of TNBC as response rates improve when taxanes are included in a chemotherapy combination.^{69, 70, 71, 72} The addition of platinum-based agents to chemotherapy regimens have been shown to be beneficial in a neoadjuvant setting in BRCA mutated TNBC cases. A phase II study involving only TNBC patients with the BRCA1 mutation demonstrated that the use of cisplatin monotherapy was highly beneficial as pCR rates increased from 61 to 90%. This result confirms that patients with the BRCA mutation respond better with platinum-based regimens.^{73, 74, 75} Iniparib (**4.8B**), a PARPi has shown to improve response rates when combined with gemcitabine (**4.8C**) and carboplatin (control). The clinical benefit rate improved to 62% compared with 21% from the gemcitabine and carboplatin combination (control). The overall response rate also improved to 48% compared to 16% from the control, although this effect was not observed in larger registrational trials.^{67, 76}

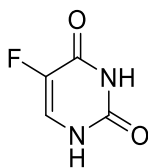
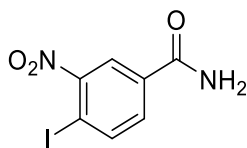
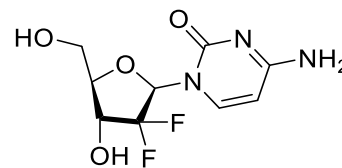
4.8A**4.8B****4.8C**

Figure 4.8: Structures of drugs used in regimens to treat TNBC. **4.8A:** Antimetabolite 5-FU, **4.8B:** PARP inhibitor iniparib, and **4.8C:** Antimetabolite gemcitabine.

4.2.3.4 Issues with the Current Treatment of TNBC

HR and HER2 positive breast cancers tend to have more successful treatment outcomes due to the selective treatments available. The reliance on cytotoxic drugs for the treatment of TNBC often leads to patients presenting adverse drug reactions (ADRs). Severe ADRs can halt the treatment regimen, limiting the choice of therapy and weakening the patient. The emergence of resistance mechanisms also causes issues, as treatments become less effective, resulting in poor patient survival.

4.2.3.4.1 Anthracyclines

Anthracyclines such as doxorubicin have varied adverse drug reactions with the most severe being cardiotoxicity. Doxorubicin causes an 8-fold increase in the risk of developing potentially fatal cardiotoxicity.⁷⁷ There are several proposed theories concerning the molecular mechanisms for the toxicity. The most prominent hypothesis is the formation of reactive oxygen species (ROS) causing oxidative stress and cell damage.⁷⁸ However, this is disputed as the use of antioxidants such as dexrazoxane (see **Figure 4.9**) make little difference to the development of heart failure.⁷⁹ Therefore, a combination of destructive mechanisms may cause cardiotoxicity. Dysfunction of cardiomyocyte calcium ion signalling pathways may hinder the heart's cardiac muscles ability to contract, causing the loss of cardiac function.^{80, 81} Doxorubicin's cardiotoxicity is not the only ADR to worry about, as it can commonly cause alopecia, anaemia, anxiety, bone marrow suppression, dizziness, depression, diarrhoea, muscle weakness, and blurred vision. Doxorubicin can also cause embolism, thrombosis, secondary oral neoplasms, and severe cutaneous adverse reactions; however, these events are rare.⁸²

It is not just adverse reactions that limit the use of anthracyclines, as the emergence of resistance mechanisms have also proved to be an issue. Multi-drug resistant (MDR) genes have led to resistance by several different mechanisms, such as reducing the accumulation of cytotoxic agents by decreasing drug influx and increasing drug efflux. They can also alter expression and/ or the activity of certain target cellular proteins. Studies have found that certain proteins can be overexpressed to allow resistance mechanisms. A common protein is the MDR1 product P-glycoprotein (P-gp). P-gp plays a major role in modulating cellular permeability to drugs by acting as an energy-dependent drug efflux pump and therefore lowers cytotoxic agent concentrations within cancer cells.⁸³ ATP binding cassette (ABC) transporters are strongly associated with chemoresistance in breast cancer to anthracyclines.⁸⁴ The multidrug resistant protein-1 (MRP1/ ABCC1), breast cancer resistance protein (BCRP/ABCG2), and the multidrug resistance protein-8 (MRP-8/ABCC11) are highly expressed in TNBC compared to other breast cancer subtypes. They are responsible for the efflux of cytotoxic agents from the breast cancer cell.^{56, 85, 86}

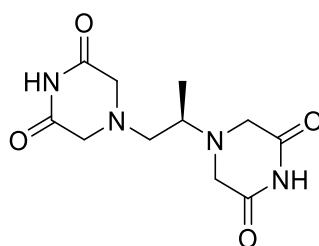


Figure 4.9: Structure of the cardioprotective agent dexrazoxane.

4.2.3.4.2 Taxanes

Taxanes such as paclitaxel and docetaxel (see **Figure 4.10**) must have their dose limited due to the emergence of neutropenia, particularly febrile neutropenia.⁸⁷ Reports show that administering docetaxel at 100 mg m⁻² over one hour every three weeks is associated with severe neutropenia in 75.4% of patients.⁸⁸ Paclitaxel dosed at 135-300 mg m⁻² over 24 hours causes neutropenia in 52% of patients.⁸⁹ Taxanes are also associated with neuropathy, which predominantly affects the sensory system. This side effects appears to be dose-related as it is usually observed 24-72 hours after a single dose exceeding 250 mg m⁻², or after multiple doses of 135-200 mg m⁻². This neuropathy usually manifests as a burning sensation in the hands and feet.^{90,91} Fluid retention is also a common complaint among patients being treated with docetaxel, in the form of oedema, ascites, and pleural effusions. If paclitaxel treatment

is continued the adverse effects can progress to motor weakness.⁹² Taxanes can commonly cause several other ADRs such as alopecia, arthralgia, and myalgia.⁹²

Resistance mechanisms of taxane are like the resistance mechanisms shown with anthracyclines, which involves the overexpression of the MDR-1 gene product P-gp. Taxanes experience an increase in efflux and decrease in influx by various transporters on the breast cancer cells.⁹³ Another resistance mechanism involves mutations that alter the β -tubulins active site, which hinders the binding of the taxane. Binding impairment ceases microtubule stabilisation and inhibits taxane induced G2-M arrest.^{94, 95, 96}

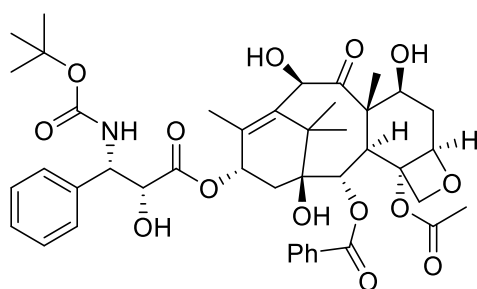


Figure 4.10: Structure of the taxane Docetaxel.

4.2.3.4.3 Platinum Agents

Platinum agents such as cisplatin have a range of dose-limiting side effects, with the most severe being nephrotoxicity. It can present in several ways, with the most common being acute kidney injury which occurs in 20-30% of patients. Other renal damage includes tubular degeneration, necrosis, and mineralisation of tubular epithelial cells.⁹⁷ Another adverse reaction is ototoxicity, which again is dose-dependent and cumulative. This side effect is prominent in children receiving the treatment. This type of toxicity is irreversible and causes loss of balance and hearing difficulties.⁹⁸ Due to cytotoxic agents' ability to inhibit DNA synthesis, cells with high rates of proliferation are adversely affected. This will fortunately damage cancer cells, but normal tissues such as bone marrow elements, gastrointestinal epithelial cells, and hair follicles will also be damaged. This can cause bone marrow suppression, alopecia, severe nausea, and vomiting.⁹⁹

Chemoresistance to platinum-based agents can occur due to the expression of DNA mismatch repair enzymes which restore DNA that are damaged by cytotoxic agents.¹⁰⁰ Fedier *et al* reported that the loss of DNA-mismatch repair proteins MSH1 and MSH2 resulted in lowered sensitivity to alkylating agents, resulting in chemosensitivity.¹⁰¹ Furthermore, an increase in

glutathione and glutathione-S-transferase activity results in an increase in enzyme degradation of alkylating agents such as cisplatin.^{67, 101, 102}

4.2.3.4.4 PARP Inhibitors

PARP inhibitors have been associated with several side effects such as agranulocytosis, anaemia, loss of appetite, cough, asthenia, and gastrointestinal discomfort.¹⁰³ However, haematological toxicities have been described as a frequent adverse reaction. A study by *Zhou et al* compared the toxic effects of various PARPi in various cancers such as ovarian, non-small and small lung cancer, breast cancer, melanoma, and gastric cancer. Four trials evaluated olaparib, seven evaluated veliparib (**4.11A**), and one evaluated niraparib (**4.11B**) (see **Figure 4.11**). Of the 12 trials, 10 reported anaemia with an incidence of 9.1%, 9 reported neutropenia with an incidence of 32.9%, and 9 reported thrombocytopenia with an incidence of 15.9%.¹⁰⁴

PARPi resistance in cancer cells have been observed and operate *via* five main mechanisms. The first mechanism involves restoration of the homologous recombination (HR). HR involves the exchange of DNA strands of similar or identical nucleotide sequences, which is used to repair double strand DNA breaks.¹⁰⁵ Despite PARPi efforts the BRCA1 and BRCA2 protein functionality and in turn HR can be restored *via* mutations.¹⁰⁶ The second mechanism involves alteration to the target PARP1 protein. PARP1 is not essential for cell survival. Therefore, the cell can remove the protein without disastrous effects to the cancer cell, and aid in the resistance to PARPi.^{107, 108} The third mechanism is the suppression of non-homologous end joining (NHEJ), which promotes the favoured error-free homologous recombination pathway instead, allowing better DNA repair and chemoresistance.¹⁰⁷ The fourth mechanism is the stabilisation of the replication fork, to induce alternate mechanisms of DNA repair resulting in PARPi resistance.^{106, 107} Finally, the last mechanism involves interference to drug concentration. This has been mentioned previously with anthracyclines. The MDR proteins are involved in the efflux of PARPi by upregulation of the ATP-dependent efflux pump ABCB1 (MDR1) gene. This gene encodes P-gp which results in resistance to PARPi by enhanced efflux movement.^{109, 110, 111}

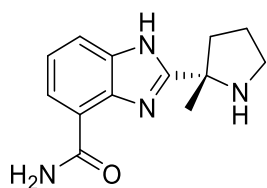
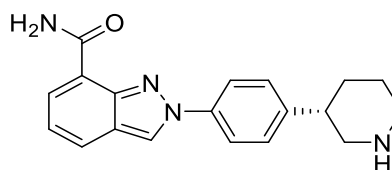
4.11A**4.11B**

Figure 4.11: Common structures of PARP inhibitors. **4.11A:** Veliparib, **4.11B:** Niraparib.

4.2.3.4.5 Immunotherapy

Immunotherapy agents such as atezolizumab have been associated with several adverse events such as fatigue, nausea, anorexia, diarrhoea, fever, dyspnea, and rashes. Other immune mediated adverse events include hypothyroidism, colitis, pneumonitis, and hepatitis. Liver toxicity is a major issue as reports found that 2-6% of participants during clinical trials have elevated levels of serum aminotransferase and were above 5 times the ULN in 1-2% of patients. However, severe liver damage such as jaundice was not reported, as detection of any abnormalities in routine liver tests would have resulted in the participant being withdrawn from the study. If the participants were continued on the treatment regimen, more extensive liver damage may be observed.^{112, 113}

4.2.3.5 New Targeted Therapy for TNBC

The current treatments for TNBC are associated with resistance mechanisms and adverse drug reactions, which can be extremely severe. These concerns have highlighted the need for a new novel treatment which specifically targets TNBC. HR and HER2 positive breast cancers have higher survival rates for several reasons, one of which being access to targeted therapies that can successfully treat the patient, often with lessened toxic effects that are associated with traditional cytotoxic agents.^{114, 115} Immunotherapy has made a huge difference to the treatment of TNBC as it is possible to selectively target cancer cells due to the precise nature of the immune system. Immunotherapy can also circumvent chemoresistance mechanisms as the immune system is inherently flexible and adaptative. Unfortunately, patient response can be varied, calling for identification of additional biomarkers to predict efficacy. Treatments that only target singular molecular mutation can result in resistance and affect patient survival. The efficacy of immunotherapy can be improved by combination therapy. Unfortunately, combination with harsh cytotoxic agents is the only current option. Therefore, efforts to develop novel selective treatments are urgently required. A selective treatment

can be combined with safer and effective treatments like immunotherapy to further improve patient survival.^{116, 117, 118} The search for druggable targets to treat TNBC are being conducted. One target that is promising is the S100A4 protein, which is currently being investigated at University of Liverpool.

4.2 S100 Protein

S100 proteins are a family that consist of 25 members and are categorized into three main subgroups: only exerting intracellular regulatory effects, exerting both intracellular and extracellular functions, and only exerting extracellular regulatory effects.¹¹⁹ These small dimeric Ca²⁺ binding proteins regulate intracellular Ca²⁺ levels and are involved in several Ca²⁺ signalling pathways.^{120, 121} These proteins are involved in several essential cell processes such as regulation of proliferation, differentiation, apoptosis, energy metabolism, inflammation, and migration/ invasion *via* interactions with various proteins such as enzymes, cytoskeleton subunits, receptors, transcription factors, and nucleic acids.^{119, 122, 123} The S100 proteins have gained interest as they have the potential to become therapeutic targets for various human disorders such as arthritis, cancer, and Alzheimer's disease.^{122, 124, 125, 126}

The S100 proteins are associated with various types of cancer due to involvement with proliferation, differentiation, and apoptosis. The S100A2 protein is considered a tumour-suppressing protein as it can be downregulated in many cancers, and the loss in nuclear expression is associated with a poor prognosis.^{127, 128} However, S100A2 can be upregulated, which is also believed to contribute to the progression of cancers, although the functions remain unclear.¹²⁷ The S100A7 protein can promote aggressive features in breast cancer by binding to c-Jun activation domain binding protein 1 (early response transcription factor), which results in the stimulation of Akt and NF- κ B. This upregulation plays a role in the development and progression of cancer *via* escalation of transcription.¹²⁹ The S100A4 protein is strongly associated with cancer, particularly TNBC due to its role in stimulating cell survival, motility, and invasion.^{119, 130, 131}

4.2.1 S100A4 Protein

As the S100A4 and other S100 proteins are strongly associated with the development and spread of many cancers they can be referred to as metastatic inducing proteins (MIPs).¹³² The S100A4 protein is a candidate as a biomarker for metastatic potential. *Rudland et al* reported

a significant difference in the 19-year survival rate for patients with invasive but inoperable stage I and stage II breast cancer depending on the S100A4 status of their tumours. 80% of patients of S100A4 negative survived compared to 11% who were S100A4 positive.¹³³ The presence of the S100A4 protein can be used as a potential predictor for metastasis and mortality in early stage breast cancer, as patients with S100A4 positive early tumours had a significantly worse 10-year survival than those with S100A4 negative tumours.^{132, 134}

The S100A4 proteins are homodimers, which are stabilised by non-covalent interactions between two helices from each subunit. Each subunit has two EF-hand Ca^{2+} - binding motifs: A C-terminal canonical EF-hand and an N-terminal pseudo EF-hand that are brought into close proximity by a small two-stranded antiparallel β sheet.^{135, 136, 137} When calcium binds to S100A4, a calcium-dependent conformational change occurs which exposes the hydrophobic binding pocket in each monomer (see **Figure 4.12**). This allows binding to extracellular or intracellular proteins to interact with the protein target and generates a biological effect.^{130, 138, 139, 140, 141}

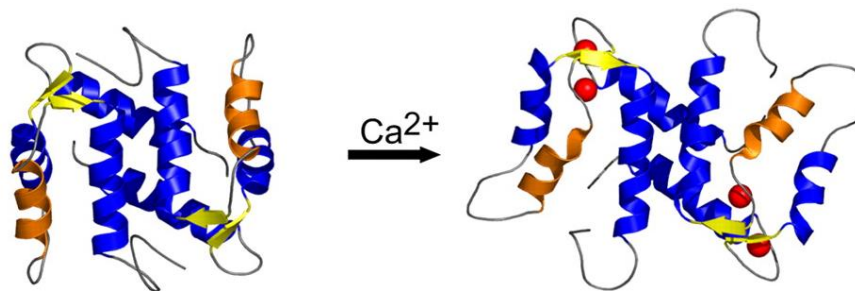


Figure 4.12: Ribbon diagrams of the NMR solution structure of apo-S100A4 (left) and a homology model of calcium-bound S100A4 (right). S100A4 undergoes a calcium-dependent conformational change which exposes the hydrophobic binding pocket allowing targets such as myosin IIA to bind.

Red sphere: Calcium ions. Reproduced from: *S100A4, a mediator of metastasis*, S. C. Garrett et al.¹³⁰

The MIP S100A4 can cause metastasis and invasiveness due to its role in proliferation, motility, and invasion. The mechanism of action involves S100A4 interacting with cytoskeleton proteins such as non-muscle myosin heavy chain (NMMHC) IIA, tropomyosin, and actin, all which aid in cell migration. Intracellular S100A4 is found to increase the stability of lamellipodia, which are cytoskeletal protein actin projections found on the leading edge of a cell, allowing the cell to migrate.¹⁴² Lamellipodia is stabilised *via* the interaction with non-

muscle myosin IIA (NMM-IIA) and therefore strongly associated with metastasis in tumour cells.^{143, 144}

The S100A4 as well as other members of the S100 family regulate the p53 tumour suppressor protein. S100A4 interacts with the C terminus of p53 and inhibits protein kinase C (PKC) phosphorylation of the tumour suppressor *in vitro*, which stops p53 binding to its DNA-binding sequence.^{145, 146, 147, 148} Therefore, it is expected to be an inhibitor of the p53 protein. However, further examination of the p53 regulated genes in S100A4 expressing cells show several genes are upregulated, while some are initially downregulated only to be upregulated later. These genes include both pro-survival and pro-apoptotic.¹⁴⁵ The effect of expression of these seemingly opposite p53 regulated genes could result from differences in the cell lines in the *in vitro* tests. Several reports have shown that the upregulation of S100A4 does increase invasive behaviours. However, this is believed to be a result of interacting with cytoskeleton proteins rather than the p53 proteins.¹³⁰ Further proof that the upregulation of the S100A4 proteins promotes the metastasis of tumours comes from a series of genomic knockout studies presented by various research groups. The removal of the S100A4 has resulted in a loss of metastatic behaviour in TNBC cancer cells.^{149, 150, 151, 152}

4.2.1.1 Targeting the S100A4 Protein

The S100 family have been identified as potential targets for cancer treatment due to their role in metastasis. The S100A4 is not the only S100 protein gaining attention as a potential therapeutic target. Inhibitors of the S100B and S100A1 proteins may halt other diseases such as Alzheimer's.^{153, 154}

S100A4 modulators are already clinically available such as calcimycin (see **Figure 4.13**), which is a calcium ionophore. Calcimycin can inhibit transcription of S100A4 in colon cancer by inhibiting the Wnt/ β -catenin pathway which reduces S100A4 transcription.¹⁵⁵ The Wnt pathway is involved in several essential cellular processes such as cell fate determination, cell migration, cell polarity, and organogenesis.¹⁵⁶ The overexpression of certain genes can over promote the Wnt signalling pathway and lead to uncontrolled cell growth and survival, and ultimately oncogenesis.¹⁵⁷

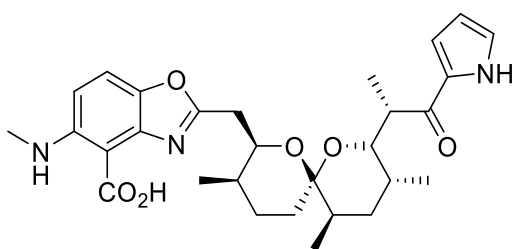


Figure 4.13: Structure of S100A4 inhibitor calcimycin.

Niclosamide (see **Figure 4.14**), which traditionally is an antihelminth has been shown to possess antitumor effects by preventing metastasis. It inhibits transcription activity of the S100A4 protein by targeting multiple signalling pathways, one of which is the previously mentioned Wnt/ β -catenin pathway.¹⁵⁸ The antitumour activity of niclosamide has led the drug to be investigated in a number of clinical trial studies. These studies are assessing the effectiveness in treating patients with colon cancer and recurrent or metastatic castration-resistant prostate cancer.^{159, 160}

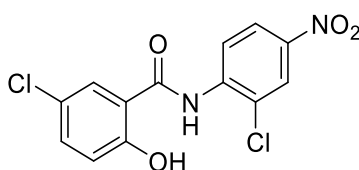


Figure 4.14: Structure of S100A4 inhibitor niclosamide

Ochiya et al have inhibited the S100A4-methionine aminopeptidase 2 interaction *via* a synthetic peptide known as NBD. The S100A4-binding domain of methionine aminopeptidase 2 is an effector protein which plays a role in endothelial cell growth. NBD was found to bind to the S100A4 and inhibit capillary formation *in vitro* and new blood vessel formation *in vivo*, halting angiogenesis.¹⁶¹

Trifluoperazine (TFP) (see **Figure 4.15**) is commonly used as an antipsychotic drug.¹⁶² *Malashkavich et al* have studied TFP interactions with the S100A4 protein. They identified TFP as an inhibitor as it interferes with the interaction between S100A4 and NMM-IIA. They determined this by studying a crystal structure of the TFP-S100A4 complex. They found that five Ca^{2+} -S100A4/ TFP dimers form a tightly packed pentameric ring. Therefore, TFP was able to inhibit S100A4 activity by inducing S100A4 oligomerisation.¹⁶³

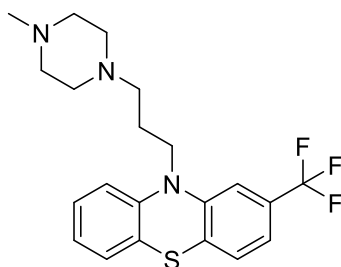


Figure 4.15: Structure of antipsychotic drug trifluoperazine (TFP) which is known to interact with S100A4.

Sorafenib (see **Figure 4.16**), which is a multikinase inhibitor may have indications for osteosarcoma, which is the most common type and aggressive form of primary bone cancer in humans. Sorafenib was shown to downregulate S100A4 expression two fold in sorafenib treated cells, as well as other tumour progression factors such as CXCR4 and the oncogene Fos.¹⁶⁴ A phase II clinical trial with sorafenib in relapsed and unresected high-grade osteosarcoma demonstrated partial and minor responses to sorafenib, with a median progression-free survival of 4 months and an overall survival time of 7 months.^{164, 165}

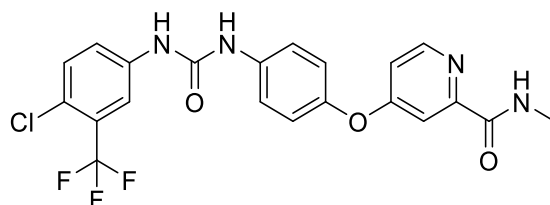


Figure 4.16: Structure of multikinase inhibitor sorafenib which can downregulate S100A4 expression.

Efficacy of these drugs are limited, as the S100 protein have long half-lives, preventing the achievement of sufficiently low steady-state protein levels to elicit a therapeutic effect. It should also be noted that transcription modulators can also possess toxic, off-target effects as they are able to regulate the expression of numerous proteins under the control of the same or related transcriptional regulatory assemblies.¹⁶⁶ Fortunately, inhibitors are not the only option to interfere with S100 expression. Gene therapy may be able to modulate the expression of the proteins in cancer patients, as preclinical animal models for the treatment of heart disease have found that it is beneficial to upregulate S100A1.^{135, 166,}

Many inhibitors have been identified to target the S100A4 protein and exert their various biological effects. The development of a novel inhibitor to selectively target the protein is a promising way to stop the invasive and aggressive nature of TNBC. However, even though

this project focuses on TNBC, S100A4 has been found to be present in several different cancers and other diseases such as Alzheimer's, and therefore the treatment can have other promising indications.

4.3 Targeting Proteins

4.3.1 Limitations of Small Molecule Inhibitors

Small molecule inhibitors have been heavily relied upon as a major source of targeted therapy. These inhibitors compete with other natural targets and are occupancy driven.¹⁶⁷ Small molecule inhibitors are designed to enter cells and block the function of certain intracellular proteins to elicit a therapeutic effect. These inhibitors are usually constrained by Lipinski's rule of five (RO5), which was developed at Pfizer to allow the inhibitors to be orally bioavailable. The size of the compound, number of hydrogen bond donors (HBDs) and hydrogen bond acceptors (HBAs) as well as the LogP values are constricted to increase the probability of developing a successful drug.^{168, 169} Small molecule inhibitors are currently the most popular choice for the inhibition of intracellular proteins.

Alternative therapies are increasingly investigated to address the limitations associated with small molecule inhibitors (see **Figure 4.17**). These limitations are, (i) Receptors and enzymes are the typical targets for small molecule inhibitors. Unfortunately, around 80% of these targets lack "druggable" active sites such as transcription factors and non-enzymatic proteins. Targets such as protein kinases have also been difficult to inhibit due to their often "featureless" and "undruggable" active sites.^{170, 171} Therefore, conventional drugs are unusable on many of these proteins.¹⁷² (ii) Adverse drug reactions are common with small molecule inhibitors and vary on severity depending on the type of drug treatment. Inhibitors need to reach the therapeutic window for effective treatment; however, this often requires high concentrations and regular dosing to reach the suitable intracellular levels for inhibition of the target. These dosing regimens can cause off-target inhibition leading to adverse drug reactions, due to lack of selectivity and the competitive nature of the drug.¹⁷³ (iii) Inhibition of multidomain compounds like scaffolding proteins can be difficult as small molecules typically only inhibit the activity of one domain. Therefore, the target protein can adapt by interacting with other intracellular or extracellular ligands to continue normal function, leading to an ineffectual inhibitor.¹⁷⁴ (iv) Resistance mechanisms can arise with most drug

treatments. Protein overexpression and accumulation can occur to compensate for the inhibitory effects of the drug. The inhibitor may be unable to cause a therapeutic effect and higher doses may be required.^{175 176} However, it is not just protein overexpression that can cause issues, as alterations to the shape of the target proteins active site can occur due to point mutations occurring in the DNA. The alteration of the proteins binding site prevents the drug interacting with the protein, ceasing the formation of the drug-target complex. This is particularly common amongst cancer treatments due to the high replication rates of tumours. Epidermal growth factor receptor (EGFR) and androgen receptor (AR) are found to be highly mutated.^{177, 178} Therefore, resistance mechanisms are often a leading cause of poor prognosis and eventual death of the patient.¹⁷³

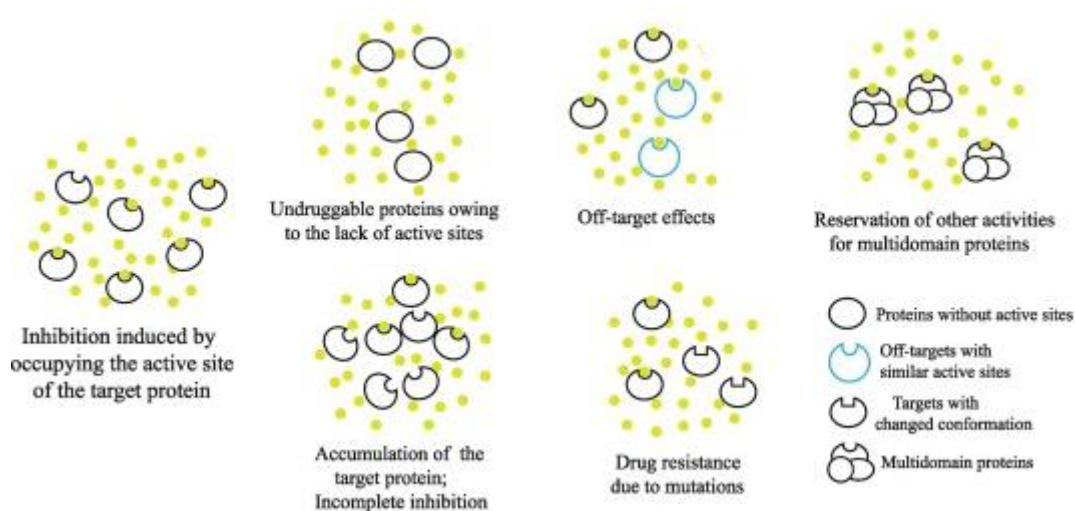


Figure 4.17: Limitations of small molecule inhibitors. Reproduced from: *Small-molecule PROTACs: An emerging and promising approach for the development of targeted therapy drugs*, S. An et al.¹⁷³

4.3.2 Alternative Approaches

Exploration into alternative targeted therapies that may be better suited for the treatment of certain diseases is required to lessen the reliance on small molecule inhibitors. There are many options available, for instance nucleic acid-based strategies such as oligonucleotides and RNA interference (RNAi) can knockdown and decrease the expression of certain disease-causing proteins. This means that an appropriate binding site and suitable ligand does not have to be found for the function of the protein to be inhibited. Unfortunately, genetic knockdown techniques are often limited as metabolic distribution, high clearance rates, and off-target binding are associated with this strategy, and therefore not suitable for the clinical use.^{179, 180, 181}

Clustered regularly interspaced short palindromic repeats/ associated protein 9 nuclease (CRISPR-Cas9) is another technique that can reduce the expression of disease inducing proteins.¹⁸² Again, this technology is limited by several issues. Off-target effects are a common concern as non-selective editing can occur leading to the potential elimination of essential proteins.^{183, 184} CRISPR-Cas9 is also associated with low efficiency, which has been observed in both *in vitro* and *in vivo* models. This efficiency can vary among different targets.^{185, 186, 187}

A relatively new approach that has gained a lot of interest in recent years is proteolysis targeting chimeras (PROTACs). The traditional methods previously mentioned that interfere with the expression of proteins have mostly been biomacromolecules.¹⁸⁸ However, small molecule PROTACs can be developed to complete similar tasks. PROTACs can induce protein degradation *via* hijacking of the hosts intracellular proteolysis pathway known as the ubiquitin proteasome system (UPS).¹⁷³

4.4 Introduction to Protein Degradation

Protein degradation is an essential process to ensure the correct functioning of the cells components as proteins play a critical role in maintaining life of an organism.^{189, 190} Proteins may need to be degraded due to their capability of aggregating or misfolding. The misfolding and misregulation of proteins can contribute to a whole range of diseases depending on the proteins involved, such as Alzheimer's, Parkinson's disease, and type II diabetes.^{191, 192} Therefore, the correct formation of proteins is monitored and when misfolding is detected refolding occurs. If the protein is unable to be corrected, it can be degraded *via* two main intracellular processes; the ubiquitin proteasome system (UPS) or autophagy.^{193, 194} The UPS pathway is responsible for the degradation of short-lived proteins and soluble misfolded proteins.¹⁹⁴ Autophagy is responsible for long-lived proteins, insoluble aggregates, whole organelles (mitochondria and peroxisomes), macromolecular compounds, and intracellular parasites such as certain bacteria.^{188, 195, 196} The S100 protein family are small proteins that are degraded *via* the UPS pathway as opposed to autophagy.^{197, 198}

4.4.1 The Ubiquitin Proteasome System

Protein degradation can occur *via* the UPS pathway. This process involves attachment of a poly ubiquitin chain to a target protein to mark it for degradation. Ubiquitin (Ub) is a small

protein comprised of 76 amino acids.¹⁹⁹ This protein is essential for maintaining homeostasis of the body by controlling several cellular functions, such as aiding in the degradation of damaged or misfolded proteins.²⁰⁰

The UPS involves three major proteins which are involved in three critical steps: (i) the E1 enzyme or the ubiquitin activating enzyme, (ii) E2 or the ubiquitin conjugating enzyme and (iii) E3 ligase or the ubiquitin ligase enzyme.²⁰¹ For the UPS to degrade a protein, a ubiquitin molecule needs to be activated by the E1 enzyme. The E1 enzyme adenylates the C-terminus of Ub in an ATP-dependent manner and then forms a thioester bond between the Ub C-terminus and the catalytic E1 cysteine residue.²⁰² Once activated, Ub can be transferred to the E2 enzyme *via* a transesterification reaction between the E1 and E2 enzymes. The Ub molecule forms a new thioester bond with a cysteine residue located in the active site of the E2 conjugating enzyme. The Ub can be transferred from the E2 enzyme to the E3 ligase *via* the formation of an isopeptide bond between the ϵ -amino group of a substrate lysine and a C-terminal carboxylate of the Ub. Ub can be transferred either to a lysine (Lys), or (less frequently) to the N-terminus, of either a substrate, or to another Ub molecule to generate mono-, multi-, or poly-ubiquitin chains.²⁰³ The generation of different Ub structures provides diverse functions of the pathway to allow different fates for the substrate protein. Monoubiquitination can regulate DNA repair and gene expression. Monoubiquitinating on multiple sites of the substrate has been reported to be important for receptor endocytosis. Polyubiquitination with a chain of at least 4 Ubs typically leads to the protein undergoing proteasomal degradation, as the substrate is shuttled to the proteasome where it can be proteolytically broken down (see **Figure 4.18**).^{204, 205, 206}

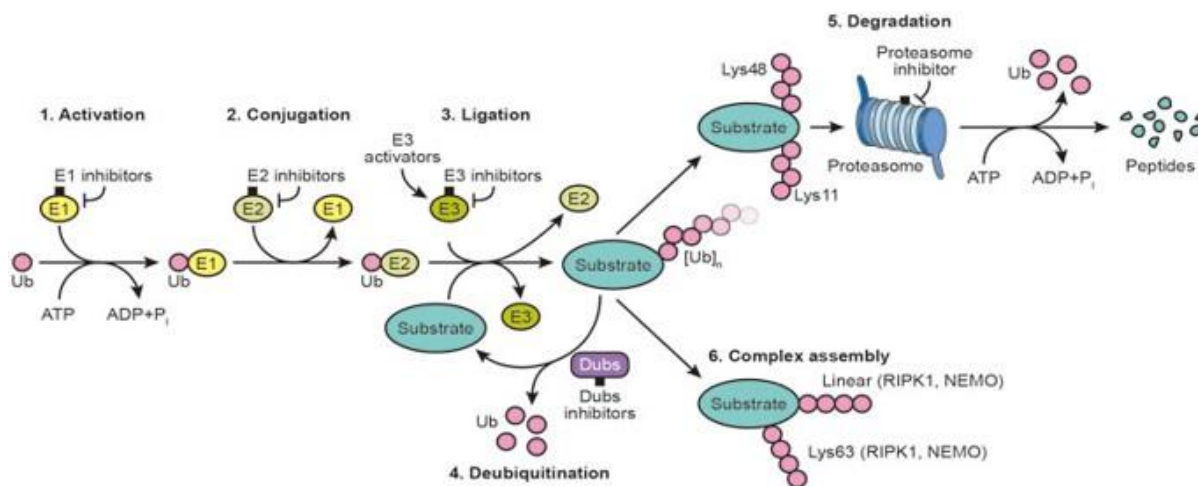


Figure 4.18: Illustration of the ubiquitin system. Ubiquitination is an ATP-dependent process involving three classes of enzymes: E1 activating enzyme, E2 conjugating enzyme, and the E3 ligase. Reproduced from: *Drugging the undruggables: exploring the ubiquitin system for drug development*, X. Huang et al.²⁰⁷

4.4.1.1 E3 Ligases in Detail

The E3 ligases can provide specificity as they recognise and bind to specific substrate sequences and degrons. This allows Ub modifications to occur on specific proteins. There is only one type of E1 enzyme in eukaryotic cells (Uba1) and 10-30 E2 enzymes. Each E2 enzyme may interact with several E3 enzymes. The number of E3 ligases is currently unknown, but it is believed to be significantly higher than the number of E2 enzymes.²⁰⁸

There are different classes of E3 ligase enzymes, one of which is known as the homologous to the E6-AP carboxyl terminus (HECT) domain type E3 ligase. This E3 ligase transfers Ub to the target protein *via* a two-step reaction. Ub is initially transferred to a catalytic cysteine on the E3 and then from the E3 to the substrate. However, the most abundant E3 ligase is characterised by a binding domain consisting of two zinc ions in a crossed-braces arrangement. This is referred to as the really interesting new gene (RING) domain. A U-box domain which adopts the same RING fold but does not contain zinc is also a common type of E3 ligase. RING E3 ligases mediate a direct transfer of Ub to the substrate, as the RING scaffold positions the Ub-charged E2 with respect to the substrate proteins. RING E3s can function as monomers, homodimers, or heterodimers. Similarly, U-box domains can work as monomers and homodimers. Some RING E3s are composed of multiple subunits, such as the cullin-RING ligases (CRLs).^{203, 209} The cullin proteins in CRLs provide a rigid scaffold for assembly. All

cullins associate with a RING protein *via* their C-terminus and can recruit a wide variety of adaptor proteins, such as substrate receptors that provide substrate specificity as the N-terminus.^{203, 210}

4.4.2 Proteolysis Targeting Chimera (PROTAC)

PROTACs are heterobifunctional molecules that consists of a substrate and a E3 ligase ligand. These ligands are connected by a linker/ tether to promote protein degradation *via* formation of a ternary complex with the recruited E3 ligase (see **Figure 4.19**). The substrate protein undergoes polyubiquitination, which can be recognised and subsequently degraded by the 26S proteasome. Therefore, PROTACs utilise and exploit the UPS pathway to regulate protein levels and instigate protein knockdown to selectively remove disease promoting proteins.¹⁶⁷

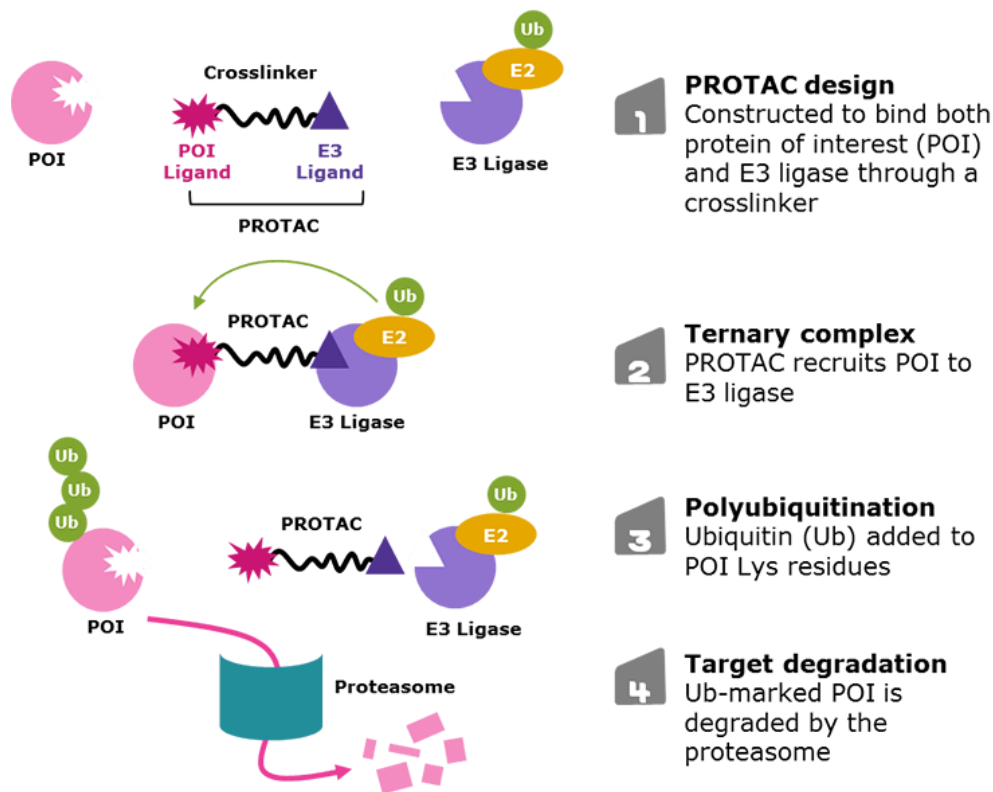


Figure 4.19: PROTAC structure and function. A PROTAC is a heterobifunctional molecule that recruits the substrate to the E3 ligase for polyubiquitination to occur. The Ub marked substrate is proteolytically broken down *via* the proteasome. Reproduced from: *Degrader Building Blocks for Targeted Protein Degradation, Sigma Aldrich*.²¹¹

The degradation induced by the PROTAC is catalytic as they can dissociate after promoting polyubiquitination of the substrate. Therefore, PROTACS can be administered at very low doses, lowering the potential for adverse drug reactions.¹⁶⁷

4.4.2.1 The Development of PROTAC agents

Gosink et al in 1995 reported artificially induced protein degradation *via* modifying E2 conjugating enzymes by fusing them to immunoglobulin (ig) binding motifs. This allowed Ub to be directly transferred from the E2 enzyme to specific substrates. This research discovered that most of the ubiquitinated substrates underwent modest proteasomal processing in a cell lysate. This research for the first time showed that artificially induced ubiquitination could result in proteasomal degradation. The implications of this work were huge as it held the potential to be applied therapeutically, as ubiquitination could be induced by artificial adaptors.²¹²

Kenten et al published a patent in 1999 from Proteinix, where a series of bifunctional inducers of protein degradation were designed. Protein degradation was able to rely on much smaller “ubiquitination recognition elements”, with smaller molecular sizes of around 1 KDa. This indicated that manipulation of protein degradation could rely on non-protein-based agents. However, no specific data was enclosed, so the exact efficiency of these molecule is not known. Unfortunately, the lack of specific data hindered its scientific impact at the time.²¹³

A paper published in 2001 from Deshaies and Crews laboratories began to stir interest in the field of PROTACs. They reported cellular degradation of the aminopeptidase METAP2 protein using a hybrid of ovalicin (a small molecule MetAP2 inhibitor) linked to I κ B α phosphopeptide epitope, which was known to bind to the Ub E3 ligase SCF β TrCP. This paper first described and introduced us to the term PROTAC. They demonstrated that the PROTAC agents were able to form a ternary complex between the degradation target (MetAP2) and the E3 ligase SCF β TrCP, resulting in a rapid (30 min) proteasomal dependent MetAP2 degradation in *Xenopus* extracts at high PROTAC concentrations. This discovery was ground-breaking, however, due to the reliance on a highly charged peptide moiety for ligase recruitment, it would be unlikely to be useful in a clinical setting. Therefore, further development was required before it could be used in drug discovery programmes.²¹⁴ The same groups published another paper using peptide based PROTACS to target and degrade oestrogen (ER) and androgen (AR) receptors, to halt the progression of breast and prostate cancer. However,

the reliance on peptide ligands required microinjection to demonstrate degradation within intact cells. Therefore, further development was required for PROTACs to be used as a mainstream treatment.^{215, 216}

Peptide based drugs are highly beneficial as they can provide high specificity and low toxicity. Although, their therapeutic applications are limited.²¹⁷ Peptide based drugs possess a short plasma half-life due to various excretory mechanisms that inactivate and clear peptides. Another drawback is their poor oral bioavailability. Many digestive enzymes are designed to break down amide bonds of ingested proteins, thus these enzymes are effective at cleaving the same bonds in peptide-based drugs. Additionally, the high polarity and molecular weight of the peptides severely limit the intestinal permeability. Administration *via* injection is often relied upon for peptide-based drugs, which results in lower patient compliance compared to the preferred administrative method of oral delivery.^{217, 218} As a result, small molecules are preferred due to lower molecular weight and polarity, leading to an improved half-life and oral bioavailability.

The design of PROTACs evolved to have therapeutic applications. The Crews' research group in 2008 developed a nonpeptidic PROTAC that could partially degrade AR by utilising the E3 ligase MDM2 in HeLa cells at a concentration of 10 μ M. The PROTAC consisted of a non-steroidal androgen receptor ligand (SARM), and a MDM2 ligand known as nutilin, which were connected by a PEG-based linker (see **Figure 4.20**). The MDM2 protein can function as a E3 ubiquitin ligase and degrade AR.^{210, 216}

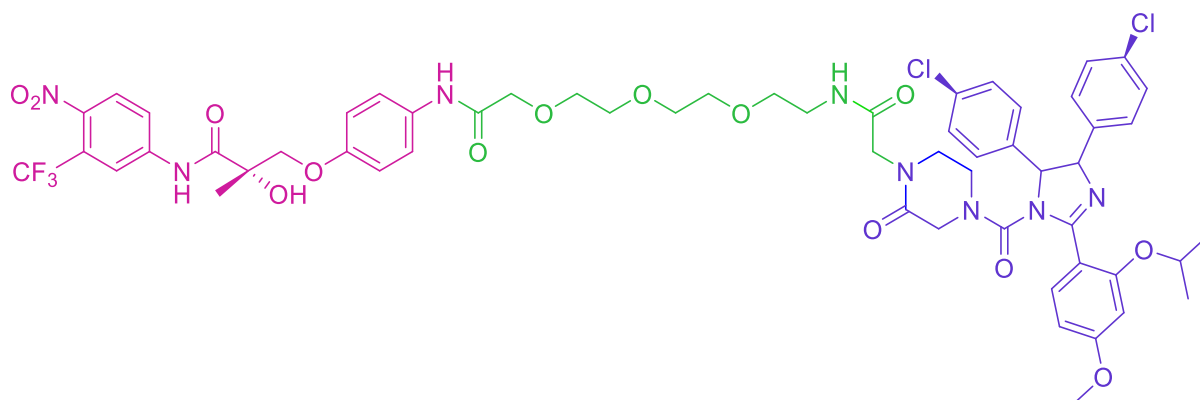
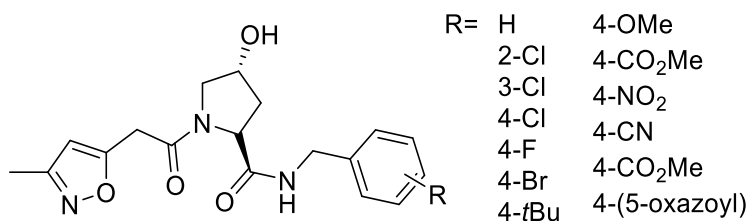


Figure 4.20: Structure of nonpeptidic PROTAC able to partially degrade AR, developed by Crews *et al.* SARM ligand which binds to AR shown in pink, PEG linker shown in green, and nutlin structure which binds to MDM2 shown in blue. Adapted from: *Targeted intracellular protein degradation induced by a small molecule: En route to chemical proteomics*, A. R. Schneekloth *et al.*²¹⁰

In 2004, the same group reported the first cell permeable PROTAC which was able to degrade the FKBP12 protein *via* a ligand they described as AP21998. The ligand to Von Hippel Lindau (VHL) E3 ligase was a seven amino acid sequence ALAPYIP derived from HIF-1 α , which is a substrate of VHL.²¹⁹ This sequence has been shown to be the minimum recognition domain for the von Hippel-Lindau tumour suppressor protein (VHL), part of the VBC-Cul2 E3 ubiquitin ligase complex. The membrane permeability of this peptide was accomplished by adding a poly-D-arginine tag. The tag also provided resistance to nonspecific proteolysis. They found that the use of this tag resulted in the degradation of androgen receptor in cultured cells. They reported the first *in vivo* examples of direct small molecule induced recruitment of target proteins to the proteasome for degradation upon addition to cultured cells.^{220, 221}

Significant advancements in PROTAC technology for the use in drug discovery occurred between 2010-2012 due to collaborations between Craig Crews (Yale University), Alessio Ciulli (University of Cambridge) and small groups of drug discovery scientists at GSK (Stevenage, U.K.).^{222, 223} They collectively identified more drug-like E3 ligase binders (see **Figure 4.21**).^{224, 225} They worked together to identify nonpeptidic ligands to mimic the natural peptidic VHL substrate.^{226, 227} Their efforts allowed the identification, optimisation, and incorporation of several binders for the E3 VHL, which in turn has provided invaluable information about the SAR and resulted in PROTACS showing highly potent cellular effects.²¹⁶ The development of non-peptidic, smaller and not excessively polar moieties allowed them to be more accessible for drug discovery, as they possessed better drug-like properties.²¹⁶

4.21A



4.21B

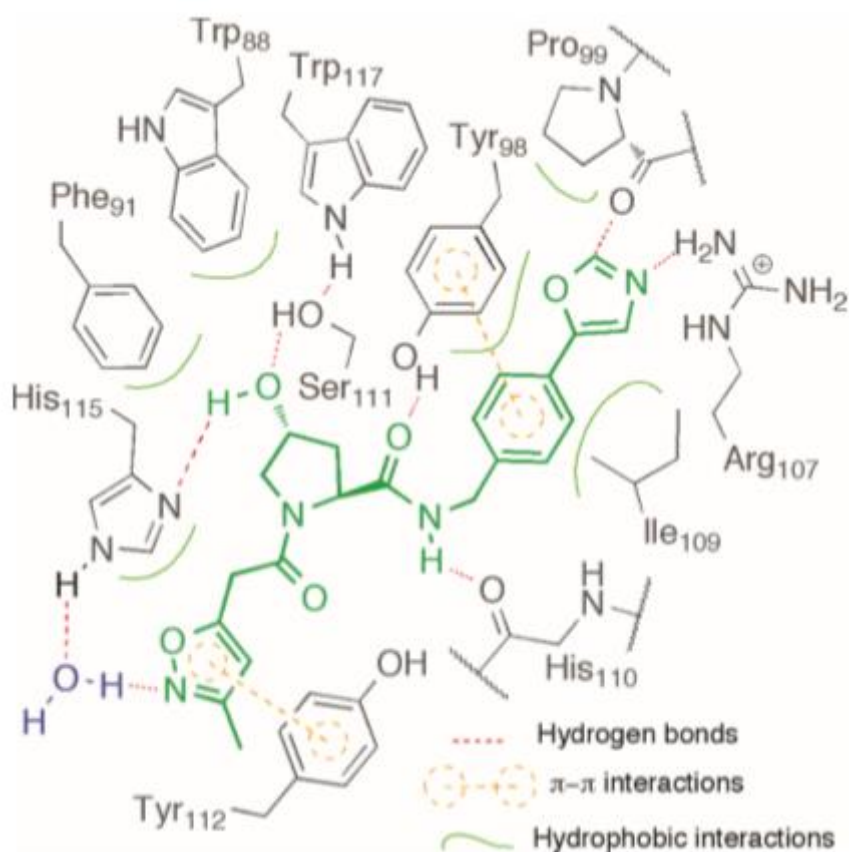


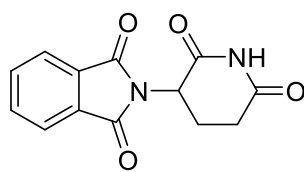
Figure 4.21: Discovery of VHL E3 ligase ligands. **4.21A:** SAR exploration of VHL ligands. **4.21B:** Graphical representation displaying the key interactions between VHL and VHL ligand with 4-(5-oxazolyl) as the R group. Reproduced from: *Targeting the von Hippel-Lindau E3 ubiquitin ligase using small molecules to disrupt the VHL/HIF-1 α interaction*, D. L. Buckley et al.²²⁵

Bondeson et al, part of Crews' research group collaborated with GSK in 2013. They reported two PROTACs that were capable of specifically reducing the protein levels of serine-threonine kinase (RIPK2) by >90% at nanomolar concentrations by exploiting VHL degradation. The mechanism of action of the PROTAC was confirmed by demonstration of the formation of the ternary complex. The PROTAC agent was also shown to be catalytic *via* experiments with a

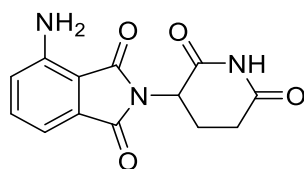
radiolabelled RIPK2. They demonstrated that a single PROTAC molecule mediates the degradation of multiple RIPK2 substrate molecules. The event was also shown to be highly selective *via* the use of proteomic techniques.^{216, 228}

Further work by several research groups reported degradation of the BET family of epigenetic bromodomain containing proteins using cereblon (CRBN) and VHL ligases.^{229, 230, 231} The incorporation of the E3 ligase CRBN ligand into PROTAC molecules opens the use of the phthalimide class of compounds which have been known since the 1960s (see **Figure 4.22**).²³² *Lu et al* were able to use pomalidomide as a ligand for CRBN to form a PROTAC that was able to degrade the BRD4 protein using the ARV-825 ligand.²³⁰ New PROTACS are reported by many research groups, which has recognised the potential for the agents in disease treatment.²¹⁶

4.22A



4.22B



4.22C

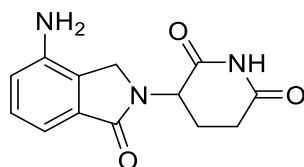


Figure 4.22: Structures of phthalimides. **4.22A:** Thalidomide, **4.22B:** Pomalidomide, **4.22C:** Lenalidomide.

The VHL and CRBN proteins are substrate recognition subunits of two important cullin RING E3 ubiquitin ligase complexes. They represent the two most common E3 ligases to be recruited by bifunctional PROTACs. Cuillis research group hypothesised that the E3 ligases

themselves could be hijacked against one another using the PROTAC strategy.²³³ They developed the first small molecule dimeriser of an E3 ligase to induce its own degradation, which they termed a “homo-PROTAC”. Homo-PROTACs can trigger a suicide-type chemical knockdown inside cells. They developed a highly active compound termed CM11 (see **Figure 4.23**), to induce a potent and preferential selective degradation of VHL, resulting in the complete knockdown of the pVHL30 protein.²³⁴

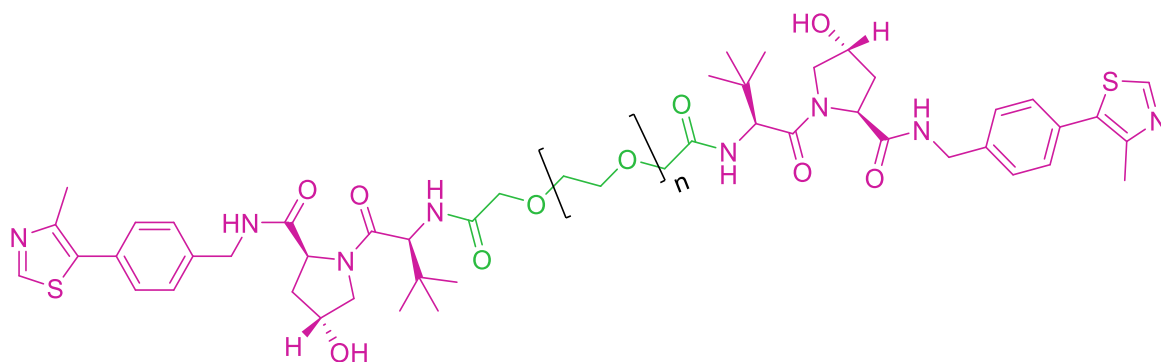


Figure 4.23: Structure of Homo-PROTAC CM11. N= 5. VHL ligands shown in pink and PEG linker shown in green. Adapted from: *Homo-PROTACs: bivalent small-molecule dimerizers of the VHL E3 ubiquitin ligase to induce self-degradation*, C. Maniaci et al.²³⁴

CRBN is a highly conserved protein that forms a complex with Cula4 and damaged DNA binding protein 1 (DDB1) to form a CRL-4-type E3 Ub ligase complex known as CRL4CRBN. This complex can ubiquitinate several substrate proteins marking them for degradation.²³⁵ Phthalimides such as thalidomide have been shown to bind to the CRBN and hence hijack the E3 ligase for selective protein degradation.^{235, 236}

New PROTAC agents showing highly promising results have emerged from the literature, particularly in the field of cancer therapy. The notable Crews’ research group has reported the development of their new PROTAC, which is the first to degrade endogenous KRAS^{G12C} protein. The KRAS^{G12C} is one of the most frequently mutated oncogenes in cancer, involved in cell proliferation, cell differentiation, and cell death. The PROTAC consists of a LC-2 ligand which can covalently bind to KRAS^{G12C} and a MRTX849 warhead which recruits the E3 ligase VHL (see **Figure 4.24**). The PROTAC induces rapid and sustained KRAS^{G12C} degradation, leading to suppression of MAPK signalling in both homogenous and heterozygous KRAS^{G12C} cell lines.²³⁷ They achieved KRAS ubiquitination and degradation with DC₅₀ values ranging

from 0.25 to 0.76 μM . They observed rapid engagement and sustained KRAS^{G12C} degradation for up to 72 hours in several KRAS^{G12C} mutant cell lines. *Zeng et al* also tried to develop a PROTAC (XY-4-88) for the degradation of the endogenous KRAS^{G12C} at 20 μM over 24 hours, unfortunately they were not successful.²³⁸ They used thalidomide to recruit the CRBN E3 ligase, whereas *Crews et al* used a MRTX based warhead to recruit the alternative VHL E3 ligase for degradation.²³⁷

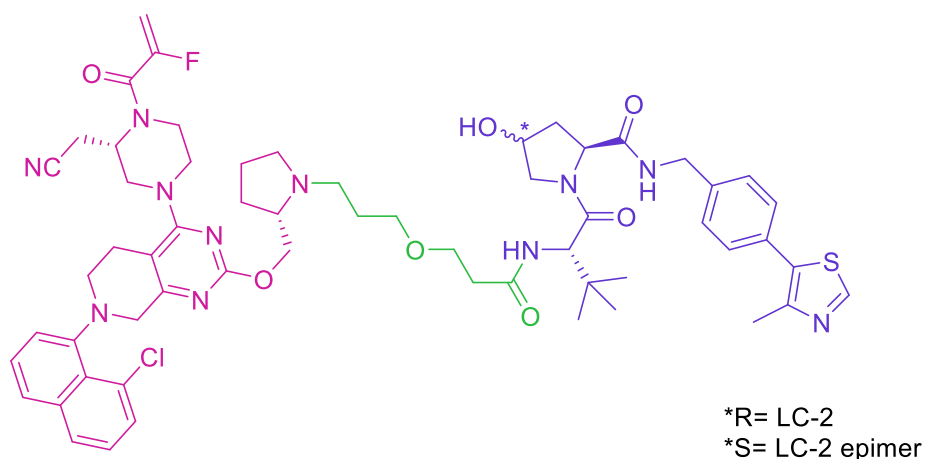


Figure 4.24: Structure of PROTAC able to degrade the KRAS^{G12C} protein. PROTAC is comprised of a LC-2 ligand that can covalently bind to the KRAS^{G12C} protein shown in **blue**, and a MRTX849 ligand able to bind to the VHL E3 ligase shown in **pink**, linked together with an amide shown in **green**. Adapted from: *Targeted Degradation of Oncogenic KRAS^{G12C} by VHL-recruiting PROTACs*, M. J. Bond *et al.*²³⁷

PROTACs are frequently referenced to target the BRD4 protein accumulation as an anticancer therapy. Pre-existing BRD4 inhibitors can develop resistance and cause the accumulation of the protein, exacerbating the cancer. *Zhang et al* developed PROTACs that were based on a potent dihydroquinazolinone BRD4 inhibitor and the CRBN E3 ligase warheads lenalidomide and pomalidomide. One of their lead compounds (see **Figure 4.25**) achieved a submicromolar IC₅₀ value of 0.81 μM , inhibiting the growth of the cancer and displaying higher potency (4 times more potent) than the original inhibitor.²³⁹

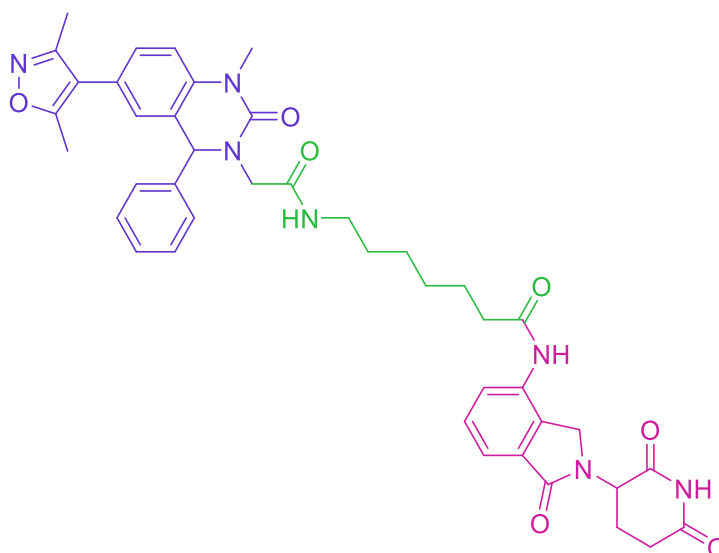


Figure 4.25: An example PROTAC developed by *Zhang et al* consisting of a BRD4 inhibitor shown in **blue**, a CRBN ligand (lenalidomide) shown in **pink**, and an amide linker shown in **green**. Adapted from: *Discovery of a new class of PROTAC BRD4 degraders based on a dihydroquinazolinone derivative and lenalidomide/pomalidomide*, F. Zhang *et al.*²³⁹

Bai et al reported targeting a signal transducer and activator of transcription 3 (STAT3) as a potential anticancer target. They designed and synthesised the PROTAC SD-36, which is a small molecule degrader of STAT3. It consists of a STAT3 SH2 inhibitor and a CRBN E3 ligase warhead lenalidomide (see **Figure 4.26**). SD-36 potently induced degradation of STAT3 protein *in vitro* and *in vivo* and demonstrated high selectivity over other STAT proteins. STAT3 degradation led to inhibition of its transcription network in leukaemia and lymphoma cells and caused cell-cycle arrest and/ or apoptosis. SD-36 achieved complete and long-lasting tumour regression in multiple xenograft mouse models, and therefore, is a promising cancer therapeutic strategy.²⁴⁰

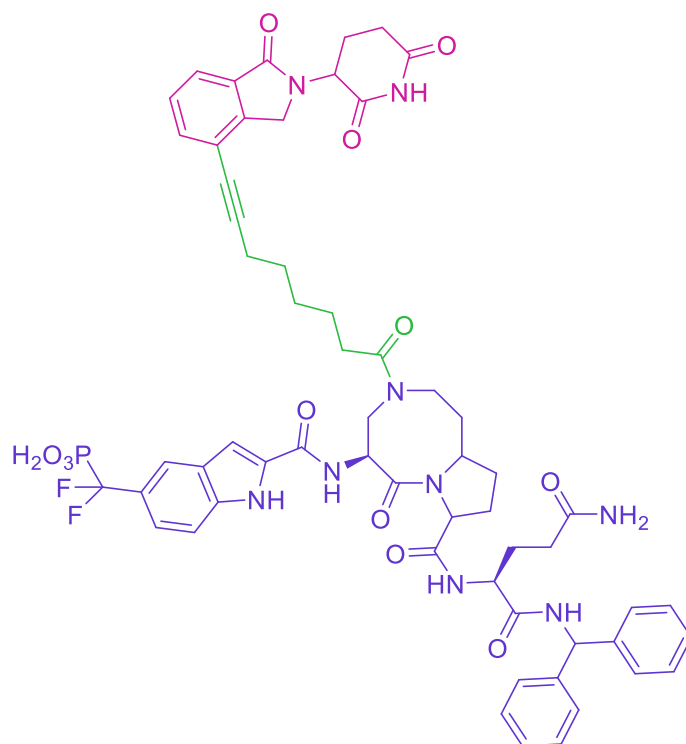


Figure 4.26: Structure of SD-36. Lenalidomide (CRBN ligand) shown in pink, connected to a STAT3 SH2 ligand shown in blue by an amide linker shown in green. Adapted from: *A Potent and Selective Small-Molecule Degrador of STAT3 Achieves Complete Tumor Regression In Vivo*, L. Bai et al.²⁴⁰

A biopharmaceutical company called Arvinas was founded by Dr Crews of Yale University and focuses on the development of PROTAC agents for the treatment of cancer and neurological diseases. They have several PROTAC agents in the pipeline, however two PROTAC agents have reached clinical trials.^{241, 242, 243}

ARV-110 is a protein degrader that targets the androgen receptor (AR) for the treatment of metastatic castration resistant prostate cancer (mCRPC). The current treatments for mCRPC are enzalutamide and abiraterone, which can be useful for patients. AR point mutations, increased levels of androgen, and/or increased AR gene are common resistance mechanisms associated with the current treatments, leading to poor prognosis for those suffering from the condition. The drug is currently in a phase I clinical study and has been reported to be well tolerated amongst the participants. Prostate-specific antigen (PSA) reductions >50% have been observed at doses greater than 280 mg.^{241, 242}

ARV-471 is an oral oestrogen receptor targeting PROTAC for the treatment of locally advanced or metastatic ER positive/ HER2 negative breast cancer. The PROTAC is currently at the phase I stage of clinical trials. ARV-471 has been well tolerated and ER degradation up to 90% (average of 62%) has been reported. A Phase 2 dose expansion of ARV-471 is expected to begin in the first half of 2021.^{241, 243}

4.5 Conclusions

PROTAC agents have undergone major advancements since 1995 so they can be utilised in drug discovery with clinical applications. The design and synthesis of PROTAC agents that can selectively degrade a disease inducing protein can be a game changer in the fight against cancer, as well as other diseases. A PROTAC to chemically degrade the metastasis inducing S100A4 for the treatment of TNBC could address the absence of targeted therapies for this class of cancer. A PROTAC agent that incorporates a S100A4 inhibitor can be tethered to a E3 ligase recruiting warhead to induce polyubiquitination of the MIP and reduce the levels or eliminate the protein. This in turn can lower the rate of metastasis and improve patient survival.

4.6 References

- 1 WHO, <https://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/>, (Accessed April 2020).
- 2 L. A. Torre, F. Bray, R. L. Siegel and J. Ferlay, *CA Cancer J Clin*, 2015, **65**, 87–108.
- 3 <https://www.wcrf.org/dietandcancer/cancer-trends/breast-cancer-statistics>, WCRF.
- 4 Editorial, *The Breast*, 2012, **21**, 423–425.
- 5 M. Kamińska, T. Ciszewski, K. Łopacka-szatan, P. Miotła and E. Starosławska, *Prz Menopauzalny*, 2015, **14**, 196–202.
- 6 P. Murawa, D. Murawa, B. Adamczyk and K. Połom, *Reports Pract. Oncol. Radiother.*, 2014, **19**, 165–172.
- 7 N. E. Allen, V. Beral, D. Casabonne, S. W. Kan, G. K. Reeves and A. Brown, *JNCI*, 2009, **101**, 296–305.
- 8 C. G. on H. F. in B. Cancer1, *Lancet*, 1996, **347**, 1713–1727.
- 9 S. S. J. Ibrar Anjum, Amna Khalid, Farwa Ejaz, *EC Gynaecol.*, 2019, **8**, 911–915.
- 10 C. Group, *Lancet*, 2019, **394**, 1159–1168.
- 11 S. Rinaldi, M. Khalis, B. Charbotel, C. Biessy, L. Dossus, I. Huybrechts, E. Fort, N. Mellas, S. Elfakir, H. Charaka, C. Nejjari, I. Romieu and K. El Rhazi, *PLoS One*, 2018, **13**, 1–12.
- 12 K. Lee, L. Kruper, C. M. Dieli-conwright, J. E. Mortimer and J. E. Mortimer, *Curr. Oncol. Rep.*, 2019, **21**, 1–6.
- 13 B. Daly, O. I. Olopade and G. Health, *JAMA*, 2015, **313**, 141–142.
- 14 M.-W. Seong, S. Cho, D.-Y. Noh, W. Han, S.-W. Kim, C.-M. Park, H.-W. Park, S. Kim, J. Kim and S. Park, *Clin. Genet.*, 2009, **76**, 152–160.
- 15 A. Mehrgou and M. Akouchekian, *Med. J. Islam. Repub. Iran*, 2016, **30**, 1–12.
- 16 R. Nanda, L. P. Schumm, S. Cummings, J. D. Fackenthal, L. Sveen, F. Ademuyiwa, M. Cobleigh, L. Esserman, N. M. Lindor, S. L. Neuhausen and O. I. Olopade, *JAMA*, 2005, **294**, 1925–1933.
- 17 WCRF, <https://www.wcrf.org/dietandcancer/breast-cancer>, (Accessed April 2020).
- 18 C. Vieira, G. Biller, I. G. Uemura, I. I. Carlos, A. Ruiz, M. Paula and C. Iv, *Clinics*, 2017, **4**, 244–253.
- 19 S. Agarwal, J. Ying, K. M. Boucher and J. P. Agarwal, *PLoS One*, 2017, **12**, 1–10.
- 20 K. Hu, P. Ding, Y. Wu, W. Tian, T. Pan and S. Zhang, *BMJ Open*, 2019, **9**, 1–8.
- 21 R. Haque, S. A. Ahmed, G. Inzhakova, J. Shi, C. Avila, J. Polikoff, L. Bernstein, S. M. Enger and M. F. Press, *Cancer Epidemiol. Biomarkers Prev.*, 2012, **21**, 1848–1855.
- 22 T. J. Ballinger, J. B. Meier and V. M. Jansen, *Front. Oncol.*, 2018, **8**, 1–11.

- 23 Cancer.net, [Cancer.net/research-and-advocacy/asco-care-and-treatment-recommendations-patients/estrogen-and-progesterone-receptor-testing-breast-cancer](https://www.cancer.net/research-and-advocacy/asco-care-and-treatment-recommendations-patients/estrogen-and-progesterone-receptor-testing-breast-cancer), (Accessed April 2020).
- 24 K. An, *Asian Spine J.*, 2016, **10**, 787–791.
- 25 M. Sporn, in *Holland-Frei Cancer Medicine. 6th edition.*, BC Decker, Hamilton (ON), 6th edn., 2003, pp. 30–44.
- 26 A. J. Trop, A. J. Epidemiol and J. Campo, *Lancet*, 1992, **340**, 1143–1145.
- 27 Breastcancer.org, https://www.breastcancer.org/symptoms/diagnosis/hormone_status/treatment_hrp_os, (Accessed April 2020).
- 28 W. Miller, *Semin. Oncol.*, 2003, **30**, 3–11.
- 29 A. C. Lai and C. M. Crews, *Nat. Rev. Drug Discov.*, 2017, **16**, 101–114.
- 30 C. K. Osborne, A. Wakeling and R. I. Nicholson, *Br. J. Cancer*, 2004, **90**, 2–6.
- 31 J. A. Freudenberg and M. I. Greene, *Exp Mol Pathol.*, 2009, **87**, 1–11.
- 32 M. M. Moasser, *Oncogene*, 2007, **26**, 6469–6487.
- 33 Cancer.org, <https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancer-diagnosis/breast-cancer-her2-status.html>, (Accessed April 2020).
- 34 T. Vu and F. X. Claret, *Front. Oncol.*, 2012, **2**, 1–6.
- 35 C. B. F. Thien and W. Y. Langdon, *Biochem J*, 2005, **166**, 153–166.
- 36 C. J. Dinsmore and P. Soriano, *Dev Biol.*, 2018, **444**, 79–97.
- 37 Breastcancer.org, <https://www.breastcancer.org/symptoms/diagnosis/her2>, (Accessed April 2020).
- 38 P. M. Lorusso, D. Weiss, E. Guardino, S. Girish and M. X. Sliwkowski, *Clin. Cancer Res.*, 2011, **17**, 6437–6448.
- 39 J. Singh, R. C. Petter, T. A. Baillie and A. Whitty, *Nat. Rev. Drug Discovery*, 2011, **10**, 307–317.
- 40 M. Segovia-mendoza, M. E. González-gonzález, D. Barrera and L. Díaz, *Am J Cancer Res*, 2015, **5**, 2531–2561.
- 41 H.-S. Cho, K. Mason, K. X. Ramyar, A. M. Stanley, S. B. Gabelli, D. W. D. Jr and D. J. Leahy, *Nature*, 2003, **421**, 756–760.
- 42 Breastcancer.org, https://www.breastcancer.org/symptoms/diagnosis/trip_neg, (Accessed April 2020).
- 43 R. Ismail-Khan and M. M. Bui, *Cancer Control*, 2010, **17**, 173–176.
- 44 G. P. Gupta and J. Massagué, *Cell*, 2006, **127**, 679–695.
- 45 A. D. Hanna, A. Lam, S. Tham, A. F. Dulhunty and N. A. Beard, *Mol. Pharmacol.*, 2014,

- 2, 438–449.
- 46 R. Dent, M. Trudeau, K. I. Pritchard, W. M. Hanna, H. K. Kahn, C. A. Sawka, L. A. Lickley, E. Rawlinson, P. Sun and S. A. Narod, *Clin. Cancer Res.*, 2007, **13**, 4429–4435.
- 47 Breastcancer.org,
https://www.breastcancer.org/symptoms/diagnosis/trip_neg?what, (Accessed April 2020).
- 48 J. M. Dolle, J. R. Daling, E. White, L. A. Brinton, D. R. Doody, P. L. Porter and K. E. Malone, *Cancer Epidemiol Biomarkers Prev.*, 2009, **18**, 1157–1166.
- 49 R. Eid, M. Lambertini and H. R. Kourie, *Int. Journal Women's Heal.*, 2019, **11**, 431–437.
- 50 H. A. Wahba and H. A. El-Hadaad, *Cancer Biol. Med.*, 2015, **12**, 106–16.
- 51 S. B. Howell, R. Safaei, C. A. Larson and M. J. Sailor, *Mol. Pharmacol.*, 2010, **77**, 887–894.
- 52 A. Pasqua, J. Goodisman, D. J. Kerwood, B. B. Toms, J. C. Dabrowiak, A. J. Di Pasqua, J. Goodisman, D. J. Kerwood, B. B. Toms, R. L. Dubowy and J. C. Dabrowiak, *Chem. Res. Toxicol.*, 2006, **19**, 139–149.
- 53 P. M. Takahara, A. C. Rosenzweig, C. A. Frederickt and S. J. Lippard, *Nature*, 1995, **377**, 649–652.
- 54 S. L. Timothy Johnstone, Ga Young, *Anticancer Res*, 2014, **34**, 471–476.
- 55 A. Shafei, W. El-Bakly, A. Sobhy, O. Wagdy, A. Reda, O. Aboelenin, A. Marzouk, K. El Habak, R. Mostafa, M. A. Ali and M. Ellithy, *Biomed. Pharmacother.*, 2017, **95**, 1209–1218.
- 56 R. A. C. Thorn, C. Oshiro, S. Marsh, T. Hernandez-Boussard, H. McLeod, T. Klein, *Pharmacogenet Genomics.*, 2011, **21**, 440–446.
- 57 I. Arnal and R. H. Wade, *Curr. Biol.*, 1995, **5**, 900–908.
- 58 B. A. Weaver, *Perspect. Cell Biol. Hum. Heal.*, 2014, **25**, 2677–2681.
- 59 K. N. Dziadkowiec, E. Gąsiorowska, E. Nowak-Markwitz and A. Jankowska, *Menopausal Rev.*, 2016, **4**, 215–219.
- 60 L. A. Emens, C. Cruz, J. P. Eder, F. Braiteh, C. Chung, S. M. Tolaney, I. Kuter, R. Nanda, P. A. Cassier, J.-P. Delord, M. S. Gordon, E. ElGabry, C.-W. Chang, I. Sarkar, W. Grossman, C. O'Hear, M. Fassò, L. Molinero and P. Schmid, *JAMA Oncol.*, 2019, **5**, 74–82.
- 61 N. L. Syn, M. W. L. Teng, T. S. K. Mok and R. A. Soo, *Lancet Oncol.*, 2017, **18**, e731–e741.
- 62 D. F. Hayes, A. D. Thor, L. G. Dressler, D. Weaver, S. Edgerton, D. Cowan, G. Broadwater, L. J. Goldstein, S. Martino, J. N. Ingle, I. C. Henderson, L. Norton, E. P. Winer, C. A. Hudis, M. J. Ellis and D. Ph, *N Engl J Med*, 2007, **357**, 1496–1506.
- 63 P. Ellis, P. Barrett-Lee, L. Johnson, D. Cameron, A. Wardley, S. O'Reilly, M. Verrill, I.

- Smith, J. Yarnold, R. Coleman, H. Earl, P. Canney, C. Twelves, C. Poole, D. Bloomfield, P. Hopwood, S. Johnston, M. Dowsett, J. M. Bartlett, I. Ellis, C. Peckitt, E. Hall and J. M. Bliss, *Lancet*, 2009, **373**, 1681–1692.
- 64 L. A. Carey, E. C. Dees, L. Sawyer, L. Gatti, D. T. Moore, F. Collichio, D. W. Ollila, C. I. Sartor, M. L. Graham and C. M. Perou, *Clin Cancer Res*, 2007, **13**, 2329–2335.
- 65 F. Cardoso, E. Guidelines and W. Group, *Ann. Oncol.*, 2013, **24**, vi7–vi23.
- 66 J. G. Shu Wang, Houpu Yang, Fuzhong Tong, Jiaqing Zhang, Deqi Yang, Hongjun Liu, Yingming Cao, Peng Liu, Po Zhou, Lin Cheng, Miao Liu, *Gan To Kagaku Ryoho*, 2009, **36**, 255–258.
- 67 H. A. Wahba and H. A. El-hadaad, *Cancer Biol Med*, 2015, **12**, 106–116.
- 68 J. P. C. Le Tourneau, S. Dettwiler, V. Laurence, S. Alran, P. Beuzeboc, *Breast Cancer Res Treat*, 2007, **106**, 4010.
- 69 J. Hugh, J. Hanson, M. C. U. Cheang, T. O. Nielsen, C. M. Perou, C. Dumontet, J. Reed, M. Krajewska, I. Treilleux, M. Rupin, E. Magherini, J. Mackey, M. Martin and C. Vogel, *J. Clin. Oncol.*, 2009, **27**, 1168–1176.
- 70 D. A. Berry, D. Ph, C. Cirrincione, I. C. Henderson, M. L. Citron, D. R. Budman, L. J. Goldstein, S. Martino, E. A. Perez, H. B. Muss, L. Norton, C. Hudis and E. P. Winer, *J.A.M.A*, 2006, **295**, 1658–1667.
- 71 D. P. H. Colin, R. James, J. E. Quinn, P. B. Mullan, P. G. Johnston, *Oncologist*, 2007, **12**, 142–150.
- 72 K. C. Lakshmaiah, A. Anand, K. G. Babu and L. Dasappa, *World J Oncol*, 2017, **8**, 110–116.
- 73 D. P. Silver, A. L. Richardson, A. C. Eklund, Z. C. Wang, Z. Szallasi, Q. Li, P. D. Ryan, N. M. Tung, A. De Nicolo, S. Ganesan, A. Miron and C. Colin, *J. Clin. Oncol. Orig.*, 2010, **28**, 1145–1153.
- 74 T. Byrski, T. Huzarski, R. Dent, E. Marczyk, M. Jasiowka, J. Gronwald, J. Jakubowicz, C. Cybulski, R. Wisniowski, D. Godlewski, J. Lubinski and S. A. Narod, *Breast Cancer Res. Treat.*, 2014, **147**, 401–405.
- 75 J. H. Park, J. Ahn and S. Kim, *ESMO Open*, 2018, **3**, 1–16.
- 76 J. O’Shaughnessy, C. Osborne, J. Pippen, M. Yoffe, D. Patt, G. Monaghan, C. Rocha, V. Ossovskaya, B. Sherman and C. Bradley, *J. Clin. Oncol.*, 2009, **27**, 3–3.
- 77 B. J. Cusack, S. P. Young and R. D. Olson, *Cancer Chemother. Pharmacol.*, 1995, **35**, 213–218.
- 78 T. Šimůnek, M. Štěrba, O. Popelová, M. Adamcová, R. Hrdina and V. Geršl, *Pharmacol. Reports*, 2009, **61**, 154–171.
- 79 E. Van Dalen, H. Van der Pal, H. Caron and L. Kremer, *Cochrane Collab.*, 2015, **4**, 1–53.
- 80 F. Zorzato, T. Facchinettis, C. Erba and G. Xiii, *J. Biol. Chem.*, 1985, **260**, 7349–7355.

- 81 A. D. Hanna, A. Lam, S. Tham, A. F. Dulhunty and N. A. Beard, *Mol. Pharmacol.*, 2014, **86**, 438–449.
- 82 NICE, <https://bnf.nice.org.uk/drug/doxorubicin-hydrochloride.html>, (Accessed April 2020).
- 83 U. A. Germann, *Eur. J. Cancer*, 1996, **32**, 927–944.
- 84 F. J. Sharom, *Pharmacogenomics*, 2008, **9**, 105–127.
- 85 A. Yamada, T. Ishikawa, I. Ota, M. Kimura, D. Shimizu, M. Tanabe, T. Chishima, T. Sasaki, Y. Ichikawa, S. Morita, K. Yoshiura, K. Takabe and I. Endo, *Breast Cancer Res Treat.*, 2013, **137**, 773–782.
- 86 L. Xu, Z. Zhao, K. Wang, H. Zhou and C. Xing, *Biomed. Res.*, 2017, **28**, 5078–5083.
- 87 L. B. Michaud, V. Valero and G. Hortobagyi, *Drug Saf.*, 2000, **23**, 401–428.
- 88 Aventis, FDA, *Cent. DRUG Eval. Res.*, 2002, **NDA 20-449**, 1–379.
- 89 M. Johnson, *Mead Johnson Oncol. Prod.*, 2011, 1–43.
- 90 R. Pazdur, A. P. Kudelka, J. J. Kavanagh, P. R. Cohen and M. N. Raber, *Cancer Treat. Rev.*, 1993, **19**, 351–386.
- 91 V. Dieras and J. Guastalla, *Br. J. Cancer*, 2003, **89**, 16–22.
- 92 M. Markman, *Expert Opin. Drug Saf.*, 2003, **2**, 141–146.
- 93 S. Murray, E. Briasoulis, H. Linardou, D. Bafaloukos and C. Papadimitriou, *Cancer Treat. Rev.*, 2012, **38**, 890–903.
- 94 S. B. H. George A Orr, Pascal Verdier-Pinard, Hayley McDaid, *Oncogene*, 2003, **22**, 7280–7295.
- 95 S. Wang, Z. Wang, L. Boise, P. Dent and S. Grant, *Biochem. Biophys. Res. Commun.*, 1999, **259**, 67–72.
- 96 T. Kosciuk, M. A. Miller, M. J. Oudin, L. Barbier, C. Sch, S. Han, O. Jonas, D. A. Lauffenburger and F. B. Gertler, *Mol Cancer Ther*, 2017, **16**, 143–156.
- 97 R. P. Miller, R. K. Tadagavadi, G. Ramesh and W. B. Reeves, *Toxins (Basel)*, 2010, **2**, 2490–2518.
- 98 L. P. Rybak, D. Mukherjea, S. Jajoo and V. Ramkumar, *Tohoku J. Exp. Med.*, 2009, **219**, 177–86.
- 99 Chemical Information Technology Imperial College London, http://www.ch.ic.ac.uk/local/projects/s_liu/Html/Toxicity.html, (Accessed April 2020).
- 100 H. P. G Giaccone, *Oncologist*, 1996, **1**, 82–87.
- 101 F. Edier, V. A. S. Chwarz, H. W. Alt, R. D. C. Arpini, U. H. Aller and D. F. Ink, *Int. J. Cancer*, 2001, **93**, 571–576.
- 102 E. A. O. Reilly, L. Gubbins, S. Sharma, R. Tully, M. Ho, Z. Guang, K. Weiner-gorzel, J.

- Mccaffrey, M. Harrison, F. Furlong, M. Kell and A. Mccann, *BBA Clin.* **3**, 2015, **3**, 257–275.
- 103 NICE, <https://bnf.nice.org.uk/drug/olaparib.html>, (Accessed April 2020).
- 104 X. Z. Jian xin Zhou, Li jin Feng, *Drug Des. Devel. Ther.*, 2017, **11**, 3009–3017.
- 105 X. L. and W.-D. Heyer, *Cell Res*, 2008, **18**, 99–113.
- 106 A. D. D’Andrea, *DNA Repair (Amst.)*, 2018, **71**, 172–176.
- 107 K. E. McCann, *Curr. Opin. Obstet. Gynecol.*, 2019, **31**, 12–17.
- 108 M. Makvandi, A. Pantel, L. Schwartz, E. Schubert, K. Xu, C. Hsieh, C. Hou, H. Kim, C. Weng, H. Winters, R. Doot, M. D. Farwell, D. A. Pryma, R. A. Greenberg, D. A. Mankoff, F. Simpkins, R. H. Mach and L. L. Lin, *J. Clin. Invest.*, 2018, **128**, 2116–2126.
- 109 S. Durmus, R. W. Sparidans, A. van Esch, E. Wagenaar, J. H. Beijnen and A. H. Schinkel, *Pharm. Res.*, 2015, **32**, 37–46.
- 110 A. Mweempwa and M. K. Wilson, *Cancer Drug Resist*, 2019, **2**, 608–617.
- 111 P. Francica and S. Rottenberg, *Genome Med.*, 2018, **10**, 1–3.
- 112 Mi. L. Yan Tie, Hui Yang, Rui Zhao, Heng Zheng, Doake Yang, Jingyi Zhao, *Drug Des. Devel. Ther.*, 2019, **13**, 523–538.
- 113 Bethesda, in *LiverTox: Clinical and Research Information on Drug-Induced Liver Injury*, <https://www.ncbi.nlm.nih.gov/books/NBK547852/>, 2010.
- 114 L. A. Hoeflerlin, C. E Chalfant and M. A. Park, *J. Surg. Sci.*, 2013, **1**, 3–7.
- 115 G. Jerusalem, J. Collignon, H. Schroeder and L. Lousberg, *Breast Cancer Targets Ther.*, 2016, **8**, 93–107.
- 116 C. L. Ventola, *P T*, 2017, **42**, 514–521.
- 117 K. R. Spencer, J. Wang, A. W. Silk, S. Ganesan, H. L. Kaufman and J. M. Mehnert, *Am. Soc. Clin. Oncol. Educ. B.*, 2016, **36**, e493–e503.
- 118 M. Suneel Kamath, Inconvenient Truths about Immunotherapy for Advanced Cancer, <https://thedoctorweighsin.com/inconvenient-truths-cancer-immunotherapy/>, (Accessed May 2020).
- 119 C. L. G. Donato, R., B.R. Cannon, G. Sorci, F. Riuzzi, K. Hsu, D.J. Weber, *Curr Mol Med.* **2**, 2013, **13**, 24–57.
- 120 E. Carafoli and J. Krebs, *J. Biol. Chem. VOL.*, 2016, **291**, 20849–20857.
- 121 E. Permyakov and R. Kretsinger, *J. Inorg. Biochem.*, 2009, **103**, 77–86.
- 122 J. Zhong and T. A. C. Rao, Xiaoquan, Xia Chang, Braunstein Zachary, *Front. Immunol.*, 2018, **8**, 1–11.
- 123 F. Fei, J. Qu, C. Li, X. Wang, Y. Li and S. Zhang, *Cell Biosci.*, 2017, **7**, 1–10.
- 124 L. A. Cerezo, M. Kuklová, M. Grigorian, M. Neidhart, S. Gay, J. Vencovský and L.

- Šenolt, *Ann Rheum Dis*, 2011, **70**, A85–A94.
- 125 M. Rema, M. Tomc, S. Gay, M. Neidhart, E. Lukanidin, K. Pavelka and M. Grigorian, *Rheumatology*, 2014, **53**, 1520–1526.
- 126 L. A. P. D R Marshak, *Prog Clin Biol Res*, 1992, **379**, 289–307.
- 127 S. Wolf, C. Haase-Kohn and J. Pietzsch, *Amino Acids*, 2011, **41**, 849–861.
- 128 J. van Dieck, D. P. Teufel, A. M. Jaulent, M. R. Fernandez-Fernandez, T. J. Rutherford, A. Wyslouch-Cieszynska and A. R. Fersht, *J. Mol. Biol.*, 2009, **394**, 922–930.
- 129 E. D. Emberley, Y. Niu, L. Curtis, S. Troup, S. K. Mandal, J. N. Myers, S. B. Gibson, L. C. Murphy and P. H. Watson, *Cancer Res*, 2005, **65**, 5696–5703.
- 130 S. C. Garrett, K. M. Varney, D. J. Weber and A. R. Bresnick, *J. Biol. Chem. VOL.*, 2006, **281**, 677–680.
- 131 A. S. Actor and P. Many, *Am. J. Pathol.*, 2010, **176**, 528–535.
- 132 D. M. Helfman, E. J. Kim, E. Lukanidin and M. Grigorian, *Br. J. Cancer*, 2005, **92**, 1955–1958.
- 133 P. S. Rudland, A. Platt-higgins, C. Renshaw, C. R. West, J. H. R. Winstanley, L. Robertson and R. Barraclough, *Cancer Res.*, 2000, **60**, 1595–1603.
- 134 W.-Y. Lee, W.-C. Su, P.-W. Lin, H.-R. Guo, T.-W. Chang and H. H. W. Chen, *Oncology*, 2004, **66**, 429–438.
- 135 A. R. Bresnick, D. J. Weber and D. B. Zimmer, *Nat Rev Cancer.*, 2015, **15**, 96–109.
- 136 N. G. Dulyaninova, P. D. Ruiz, M. J. Gamble, J. M. Backer and P. Forscher, *Mol. Biol. Cell*, 2018, **29**, 632–642.
- 137 K. M. Vallely, R. R. Rustandi, K. C. Ellis, O. Varlamova, A. R. Bresnick and D. J. Weber, *Biochemistry*, 2002, **41**, 12670–12680.
- 138 W. J. C. L Mäler, B C Potts, *J Biomol NMR*, 1999, **13**, 233–47.
- 139 S. Bhattacharya, E. Large, C. W. Heizmann, B. A. Hemmings and W. J. Chazin, *Biochemistry*, 2003, **43**, 14416–14426.
- 140 R. R. Rustandi, D. M. Baldisseri and D. J. Weber, *Nat. Struct. Biol.*, 2000, **7**, 570–574.
- 141 P. Pathuri, L. Vogeley and H. Luecke, *J Mol Biol.*, 2008, **383**, 62–77.
- 142 C. Ballestrem, B. Wehrle-haller, B. Hinz and B. A. Imhof, *Mol. Biol. Cell*, 2000, **11**, 2999–3012.
- 143 Z.-H. Li and A. R. Bresnick, *Cancer Res.*, 2006, **66**, 5173–5180.
- 144 F. Fei, J. Qu, M. Zhang, Y. Li and S. Zhang, *Oncotarget*, 2017, **8**, 73219–73239.
- 145 M. Grigorian, S. Andresen, E. Tulchinsky, M. Kriajevska, C. Carlberg, C. Kruse, M. Cohn, N. Ambartsumian, A. Christensen, G. Selivanova and E. Lukanidin, *J. Biol. Chem.*, 2001, **276**, 22699–22708.

- 146 J. Baudier, C. Delphin, D. Grunwald and S. Khochbin, *Proc Natl Acad Sci USA*, 1992, **89**, 11627–11631.
- 147 C. Delphin, M. Ronjat, J. C. Deloulme, L. Debussche, Y. Higashimoto, K. Sakaguchi and J. Baudier, *J. Biol. Chem.*, 1999, **274**, 10539–10544.
- 148 M. Doi, J. Lin, Q. Yang, Z. Yan, J. Markowitz, P. T. Wilder, F. Carrier, D. J. Weber and S. B. Lin, *J. Biol. Chem.*, 2004, **279**, 34071–34077.
- 149 M. Fujiwara, T. G. Kashima, A. Kunita, I. Kii, D. Komura, A. E. Grigoriadis, A. Kudo, H. Aburatani and M. Fukayama, *Tumor Biol.*, 2011, **32**, 611–622.
- 150 G. Zhang, M. Li, J. Jin, Y. Bai and C. Yang, *Asian Pacific J Cancer Prev*, 2011, **12**, 2075–2080.
- 151 T. Ochiya and E. H. Takenaga, Keizo, *Angiogenesis*, 2014, **17**, 17–26.
- 152 M. Dahlmann, U. Sack, P. Herrmann, M. Lemm, I. Fichtner, P. M. Schlag and U. Stein, *Oncotarget*, 2012, **3**, 783–797.
- 153 E. Roltsch, L. Holcomb, K. A. Young, A. Marks and D. B. Zimmer, *J. Neuroinflammation*, 2010, **7**, 1–12.
- 154 L. Afanador, E. A. Roltsch, L. Holcomb, K. S. Campbell, D. A. Keeling, Y. Zhang and D. B. Zimmer, *Cell Calcium*, 2014, **56**, 68–80.
- 155 U. Sack, W. Walther, D. Scudiero, M. Selby, J. Aumann, C. Lemos and J. C. Adams, *Mol. Biol. Cell*, 2011, **22**, 3344–3354.
- 156 Y. Komiya and R. Habas, *Organog.* 42, 2008, **4**, 68–75.
- 157 R. Morgan, R. Ankrah, S. El-Tanani, P. M. Loadman, L. Pattterson, P. S. Rudland and M. El-Tanani, in *Introduction to Cancer Metastasis*, Elsevier, 2017, pp. 375–394.
- 158 J.-X. Pan, K. Ding and C.-Y. Wang, *Chin. J. Cancer*, 2012, **31**, 178–184.
- 159 NIH, <https://www.cancer.gov/about-cancer/treatment/clinical-trials/search/v?id=NCI-2016-01245&r=1>, (Accessed December 2020).
- 160 NIH, <https://www.cancer.gov/about-cancer/treatment/clinical-trials/search/v?id=NCI-2017-00294&r=1>, (Accessed December 2020).
- 161 T. Ochiya, K. Takenaga, M. Asagiri, K. Nakano, H. Satoh, T. Watanabe and S. Imajoh-ohmi, *Mol. Ther. — Methods Clin. Dev.*, 2015, **2**, 1–11.
- 162 M. F. Jia Qin, Aleksey V. Zima, Maura Porta, Lothar A. Blatter, *Pflugers Arch.*, 2009, **458**, 643–651.
- 163 V. N. Malashkevich, N. G. Dulyaninova, U. A. Ramagopal, M. A. Liriano and K. M. Varney, *PNAS*, 2010, **107**, 8605–8610.
- 164 I. Walter, B. Wolfesberger, I. Miller, G. Mair, S. Burger, B. Gallè and R. Steinborn, *Oncol. REPORTS* 31, 2014, **31**, 1147–1156.
- 165 G. Grignani, E. Palmerini, P. Dileo, S. D. Asaftei, L. D. Ambrosio and Y. Pignochino, *Ann. Oncol.*, 2012, **23**, 508–516.

- 166 C. Weber, I. Neacsu, B. Krautz, P. Schlegel, S. Sauer, P. Raake, J. Ritterhoff, A. Jungmann, A. Remppis, M. Stangassinger, W. Koch, H. Katus, O. Müller, P. Most and S. Pleger, *Gene Ther.*, 2014, **21**, 131–138.
- 167 X. Sun, H. Gao, Y. Yang, M. He, Y. Wu, Y. Song, Y. Tong and Y. Rao, *Signal Transduct. Target. Ther.*, 2019, **4**, 1–33.
- 168 P. Wu, T. E. Nielsen and M. H. Clausen, *Drug Discov. Today*, 2016, **21**, 5–10.
- 169 L. Z. Benet, C. M. Hosey, O. Ursu and T. I. Oprea, *Adv Drug Deliv Rev.*, 2016, **101**, 189–98.
- 170 M. Egbert, A. Whitty, G. M. Keserü and S. Vajda, *J. Med. Chem.*, 2019, **62**, 10005–10025.
- 171 D. A. Degoey, H. J. Chen, P. B. Cox and M. D. Wendt, *J. Med. Chem.*, 2018, **61**, 2636–2651.
- 172 C. M. Crews, *Chem. Biol.*, 2010, **17**, 551–555.
- 173 S. An and L. Fu, *EBioMedicine*, 2018, **36**, 553–562.
- 174 L. M. Graves, J. S. Duncan, M. C. Whittle and G. L. Johnson, *Biochem. J.*, 2013, **450**, 1–8.
- 175 B. P. Rubin and A. Duensing, *Lab. Investig.*, 2006, **86**, 981–986.
- 176 R. Krishnamurty and D. J. Maly, *ACS Chem. Biol.*, 2010, **5**, 121–138.
- 177 S. Dogan, R. Shen, D. C. Ang, M. L. Johnson, S. P. D’Angelo, P. K. Paik, E. B. Brzostowski, G. J. Riely, M. G. Kris, M. F. Zakowski and M. Ladanyi, *Clin. Cancer Res.*, 2012, **18**, 6169–6177.
- 178 K. Eisermann, D. Wang, Y. Jing, L. E. Pascal and Z. Wang, *Transl. Androl. Urol.*, 2013, **2**, 137–147.
- 179 M. Toure and C. M. Crews, *Angew. Chemie Int. Ed.*, 2016, **55**, 1966–1973.
- 180 M. Mansoor and A. J. Melendez, *Gene Regul. Syst. Bio.*, 2008, **2**, 275–295.
- 181 C. A. Sledz and B. R. G. Williams, *Blood*, 2005, **106**, 787–794.
- 182 M. Adli, *Nat. Commun.*, 2018, **9**, 1–13.
- 183 Y. Fu, J. A. Foden, C. Khayter, M. L. Maeder, D. Reyon, J. K. Joung and J. D. Sander, *Nat. Biotechnol.*, 2013, **31**, 822–826.
- 184 Y. Lin, T. J. Cradick, M. T. Brown, H. Deshmukh, P. Ranjan, N. Sarode, B. M. Wile, P. M. Vertino, F. J. Stewart and G. Bao, *Nucleic Acids Res.*, 2014, **42**, 7473–7485.
- 185 M. Kosicki, S. S. Rajan, F. C. Lorenzetti, H. H. Wandall, Y. Narimatsu, E. Metzakopian and E. P. Bennett, in *Progress in Molecular Biology and Translational Science*, Elsevier, 2017, pp. 49–67.
- 186 I. Simoff, M. Karlgren, M. Backlund, A.-C. Lindström, F. Z. Gaugaz, P. Matsson and P. Artursson, *J. Pharm. Sci.*, 2016, **105**, 1017–1021.

- 187 M. Karlgren, I. Simoff, M. Keiser, S. Oswald and P. Artursson, *Drug Metab. Dispos.*, 2018, **46**, 1776–1786.
- 188 Y. Wang, X. Jiang, F. Feng, W. Liu and H. Sun, *Acta Pharm. Sin. B*, 2020, **10**, 207–238.
- 189 M. H. Seo and P. M. Kim, *Curr. Opin. Struct. Biol.*, 2018, **50**, 162–170.
- 190 J. M. Berg, J. L. Tymoczko and L. Stryer, in *Biochemistry*, W H Freeman, New York, 2002.
- 191 M. Stefani, *Biochim. Biophys. Acta - Mol. Basis Dis.*, 2004, **1739**, 5–25.
- 192 C. Soto and S. Pritzkow, *Nat. Neurosci.*, 2018, **21**, 1332–1340.
- 193 J. Lee, Y. Xu, T. Zhang, L. Cui, L. Saidi and Y. Ye, *J. Biol. Chem.*, 2018, **293**, 14359–14370.
- 194 A. Ciechanover and Y. T. a. Kwon, *Exp. Mol. Med.*, 2015, **47**, e147–e163.
- 195 N. M. Kocaturk and D. Gozuacik, *Front. Cell Dev. Biol.*, 2018, **6**, 1–27.
- 196 F. Navid, G. Layh-Schmitt, K. A. Sikora, A. Cougnoux and R. A. Colbert, *Arthritis Rheumatol.*, 2018, **70**, 746–755.
- 197 R. Mukherjee, A. Das, S. Chakrabarti and O. Chakrabarti, *Biochim. Biophys. Acta - Mol. Cell Res.*, 2017, **1864**, 1227–1235.
- 198 A. T. Melvin, G. S. Woss, J. H. Park, M. L. Waters and N. L. Allbritton, *Cell Biochem Biophys*, 2013, **67**, 75–89.
- 199 S. H. Lecker, A. L. Goldberg and W. E. Mitch, *J. Am. Soc. Nephrol.*, 2006, **17**, 1807–1819.
- 200 N. Nakamura, *Int. J. Mol. Sci.*, 2018, **19**, 1–5.
- 201 Y. Tu, C. Chen, J. Pan, J. Xu, Z. G. Zhou and C. Y. Wang, *Int. J. Clin. Exp. Pathol.*, 2012, **5**, 726–738.
- 202 J. W. H. Brenda A. Schulman, *Nat Rev Mol Cell Biol.*, 2009, **10**, 319–331.
- 203 A. M. W. Meredith B. Metzger, Jonathan N. Pruneda, Rachel E. Klevit, *Biochim Biophys Acta*, 2014, **1843**, 47–60.
- 204 K. Tanaka, *Proc. Japan Acad. Ser. B Phys. Biol. Sci.*, 2009, **85**, 12–36.
- 205 W. Li and Y. Ye, *Cell. Mol. Life Sci.*, 2008, **65**, 2397–2406.
- 206 M. Sadowski, R. Suryadinata, A. R. Tan, S. N. A. Roesley and B. Sarcevic, *IUBMB Life*, 2012, **64**, 136–142.
- 207 X. Huang and V. M. Dixit, *Cell Res.*, 2016, **26**, 484–498.
- 208 L. A. Passmore and D. Barford, *Biochem. J.*, 2004, **379**, 513–525.
- 209 F. E. Morreale and H. Walden, *Cell*, 2016, **165**, 248–248.
- 210 A. R. Schneekloth, M. Pucheault, H. S. Tae and C. M. Crews, *Bioorg Med Chem Lett.*, 2008, **18**, 5904–5908.

- 211 Sigma Aldrich, <https://www.sigmaaldrich.com/technical-documents/articles/technology-spotlights/protac-building-blocks.html>, (Accessed May 2020).
- 212 M. M. Gosink and R. D. Vierstra, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 9117–9121.
- 213 J. H. Kenten, S.F. Roberts, US6306663B1, 2001.
- 214 K. M. Sakamoto, K. B. Kim, A. Kumagai, F. Mercurio, C. M. Crews and R. J. Deshaies, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 8554–8559.
- 215 K. M. Sakamoto, K. B. Kim, R. Verma, A. Ransick, B. Stein, C. M. Crews and R. J. Deshaies, *Mol. Cell. Proteomics*, 2003, **2**, 1350–1358.
- 216 I. Churcher, *J. Med. Chem.*, 2018, **61**, 444–452.
- 217 L. Otvos and J. D. Wade, *Front. Chem.*, 2014, **2**, 1–4.
- 218 J. L. Lau and M. K. Dunn, *Bioorganic Med. Chem.*, 2018, **26**, 2700–2707.
- 219 W. C. Hon, M. I. Wilson, K. Harlos, T. D. W. Claridge, C. J. Schofield, C. W. Pugh, P. H. Maxwell, P. J. Ratcliffe, D. I. Stuart and E. Y. Jones, *Nature*, 2002, **417**, 975–978.
- 220 M. Ohh, C. W. Park, M. Ivan, M. A. Hoffman, T. Y. Kim, L. E. Huang, N. Pavletich, V. Chau and W. G. Kaelin, *Nat. Cell Biol.*, 2000, **2**, 423–427.
- 221 J. S. Schneekloth, F. N. Fonseca, M. Koldobskiy, A. Mandal, R. Deshaies, K. Sakamoto and C. M. Crews, *J. Am. Chem. Soc.*, 2004, **126**, 3748–3754.
- 222 C. M. C. Dennis L Buckley, Kanak Raina, Nicole Darricarrere, John Hines, Jeffrey L Gustafson, Ian E. Smith, Afjal H Miah, John D Harling, *ACS Chem Biol.*, 2015, **10**, 1831–1837.
- 223 Campos, G. Campos, C. M. Crew, A. Ciulli, D. Buckley et al, WO2013/106646A3, 2013.
- 224 P. C. M. C. Dennis L. Buckley, Dr. Jeffrey L. Gustafson, Dr. Inge Van Molle, Dr. Anke G. Roth, Dr. Alessio Ciulli, *Angew Chem Int Ed Engl.*, 2012, **51**, 11463–11467.
- 225 D. L. Buckley, I. Van Molle, P. C. Gareiss, H. S. Tae, J. Michel, D. J. Noblin, W. L. Jorgensen, A. Ciulli and C. M. Crews, *J. Am. Chem. Soc.*, 2012, **134**, 4465–4468.
- 226 I. Van Molle, A. Thomann, D. L. Buckley, E. C. So, S. Lang, C. M. Crews and A. Ciulli, *Chem Biol.*, 2012, **19**, 1310–1312.
- 227 C. Galdeano, M. S. Gadd, P. Soares, S. Scaffidi, I. Van Molle, I. Birced, S. Hewitt, D. M. Dias and A. Ciulli, *J. Med. Chem.*, 2014, **57**, 8657–8663.
- 228 D. P. Bondeson, A. Mares, I. E. D. Smith, E. Ko, S. Campos, A. H. Miah, K. E. Mulholland, N. Routly, D. L. Buckley, L. Jeffrey, N. Zinn, P. Grandi, S. Shimamura, G. Bergamini, M. Bantscheff, C. Cox, D. A. Gordon, R. R. Willard, J. J. Flanagan, L. N. Casillas, B. J. Votta, W. Den Besten, L. Kruidenier, P. S. Carter, J. D. Harling, I. Churcher and M. Craig, *Nat Chem Biol.*, 2015, **11**, 611–617.
- 229 B. J. E. Winter Georg E., Buckley Dennis L., Paulk Joshiawa, Roberts Justin M., Souza Amanda, Dhe-Paganon Sirano, *Science (80-.)*, 2015, **348**, 1376–1381.

- 230 J. Lu, Y. Qian, M. Altieri, H. Dong, J. Wang, K. Raina, J. D. Winkler, A. P. Crew, K. Coleman and C. M. Crews, *Chem Biol.*, 2015, **22**, 755–763.
- 231 M. Zengerle, K. H. Chan and A. Ciulli, *ACS Chem. Biol.*, 2015, **10**, 1770–1777.
- 232 R. O. Nulsen, *Am. J. Obstet. Gynecol.*, 1961, **81**, 1245–1248.
- 233 M. Girardini, C. Maniaci, S. J. Hughes, A. Testa and A. Ciulli, *Bioorganic Med. Chem.*, 2019, **27**, 2466–2479.
- 234 C. Maniaci, S. J. Hughes, A. Testa, W. Chen, D. J. Lamont, S. Rocha, D. R. Alessi, R. Romeo and A. Ciulli, *Nat. Commun.*, 2017, **8**, 1–14.
- 235 T. Mori, T. Ito, S. Liu, H. Ando, S. Sakamoto, Y. Yamaguchi, E. Tokunaga, N. Shibata, H. Handa and T. Hakoshima, *Sci. Rep.*, 2018, **8**, 1–14.
- 236 Y. Liu, X. Huang, X. He, Y. Zhou, X. Jiang, S. Chen-Kiang, S. R. Jaffrey and G. Xu, *FASEB J.*, 2015, **29**, 4829–4839.
- 237 M. J. Bond, L. Chu, D. A. Nalawansa, K. Li and C. Crews, *ChemRxiv. Prepr.*, , DOI:<https://doi.org/10.26434/chemrxiv.12091176.v1>.
- 238 M. Zeng, Y. Xiong, N. Safaee, R. P. Nowak, K. A. Donovan, C. J. Yuan, B. Nabet, T. W. Gero, F. Feru, L. Li, S. Gondi, L. J. Ombelets, C. Quan, P. A. Jänne, M. Kostic, D. A. Scott, K. D. Westover, E. S. Fischer and N. S. Gray, *Cell Chem. Biol.*, 2020, **27**, 19-31.e6.
- 239 F. Zhang, Z. Wu, P. Chen, J. Zhang, T. Wang, J. Zhou and H. Zhang, *Bioorganic Med. Chem.*, 2020, **28**, 115228.
- 240 L. Bai, H. Zhou, R. Xu, Y. Zhao, K. Chinnaswamy, D. McEachern, J. Chen, C. Y. Yang, Z. Liu, M. Wang, L. Liu, H. Jiang, B. Wen, P. Kumar, J. L. Meagher, D. Sun, J. A. Stuckey and S. Wang, *Cancer Cell*, 2019, **36**, 498–511.
- 241 GlobeNewsWire, <https://www.globenewswire.com/news-release/2020/12/14/2144367/0/en/Arvinas-Releases-Interim-Clinical-Data-Further-Demonstrating-the-Powerful-Potential-of-PROTAC-Protein-Degraders-ARV-471-and-ARV-110.html>, (Accessed December 2020).
- 242 Arvinas, <https://www.arvinas.com/pipeline-programs/androgen-receptor>, (Accessed December 2020).
- 243 Arvinas, <https://www.arvinas.com/pipeline-programs/estrogen-receptor>, (Accessed December 2020).

CHAPTER 5

Structural Activity Relationship Exploration of
S100A4 Inhibitors, and Design and Synthesis
of Proteolysis Targeting Chimera Agents

5.1 Discovery of US10113 Inhibitor

To develop a targeted treatment for TNBC, a therapeutic agent that selectively binds to the S100A4 protein may be helpful. There are several ways to target S100 proteins, which have been discussed previously (see **Chapter 4, Section 4.2.1.1**). Pentamidine is the most common agent used experimentally to investigate small molecular targeting of S100A4 for the inhibition of cell migration. This drug has several indications involving antiprotozoal, antimicrobial, and antitumor activity. The mechanism of action for their activity is not fully known but believed to be involved in the disruption of DNA, RNA, and protein synthesis. The antitumor properties are believed to come from acting as a small molecule antagonist that disrupts interaction between S100P and the RAGE receptor and/or the interaction between S100P and p53.^{1,2} However, pentamidine has shown to be active against S100A4, but only at high, nonspecific concentrations.^{3,4}

In the hunt for a new therapeutic agent, Cancer Research UK (CRUK) conducted a screen of over 2,500 compounds for potential activity against S100A4 proteins. Out of all these compounds, one scaffold termed **CT070909** was found to be a hit. The hit displayed a 90% inhibition at 30-50 μM , using robotic Surface Plasmon Resonance (SPR).⁵ In-house synthesis of “hit like” compounds were generated by the Dr Carnell research group (University of Liverpool) to provide three lead inhibitors; **US10113 (1a)**, **US10253-H (1b)**, and **US1041-MOF (1c)** (see **Figure 5.1**).

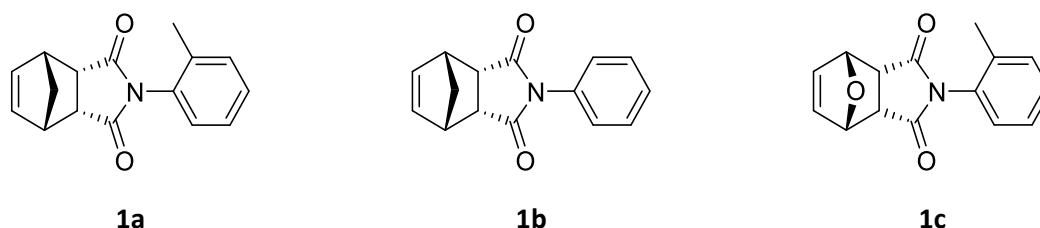


Figure 5.1: Structures of original inhibitors. **5.1A:** US10113 (**1a**), **5.1B:** US10253-H/ CT070909 (**1b**) and **5.1C:** US1041-MOF (**1c**).

The original inhibitors were assayed for inhibition of S100A4-dependent metastasis, in a *Drosophila* model involving transgenic fruit flies overexpressing the mutant Ras^{Val12} and S100A4 (see **Section 5.1.1.1**). **US10113 (1a)** inhibited the binding between S100A4 and NNMIIA in the transgenic fly larvae at 84% for 33 μM and 98% for 200 μM . This inhibition was greater than the other inhibitors as **US10253-H (1b)** and **US1041-MOF (1c)**, inhibited at only 63% to 67% at 33 μM and 200 μM (see **Figure 5.2A**).

No significant toxicity was observed for **US10113 (1a)** at 200 μM as the number of live pupae at the end of the experiment remained the same. The other two **US** compounds displayed significant toxicity (see **Figure 5.2C**). An associated 50% reduction in overall GFP-fluorescence reflecting and a fall in tumour burden was observed for **US10113 (1a)** (see **Figure 5.2B**).⁶ Ras^{Val12} / S100A4 over Ras^{Val12} larvae demonstrates clearly that S100A4 promotes extensive dissemination to the ventral nerve cord (VNC), as well as elsewhere in the larvae, as GFP is spreading from the eye lobes to other tissue (see **Figure 5.3**).⁶

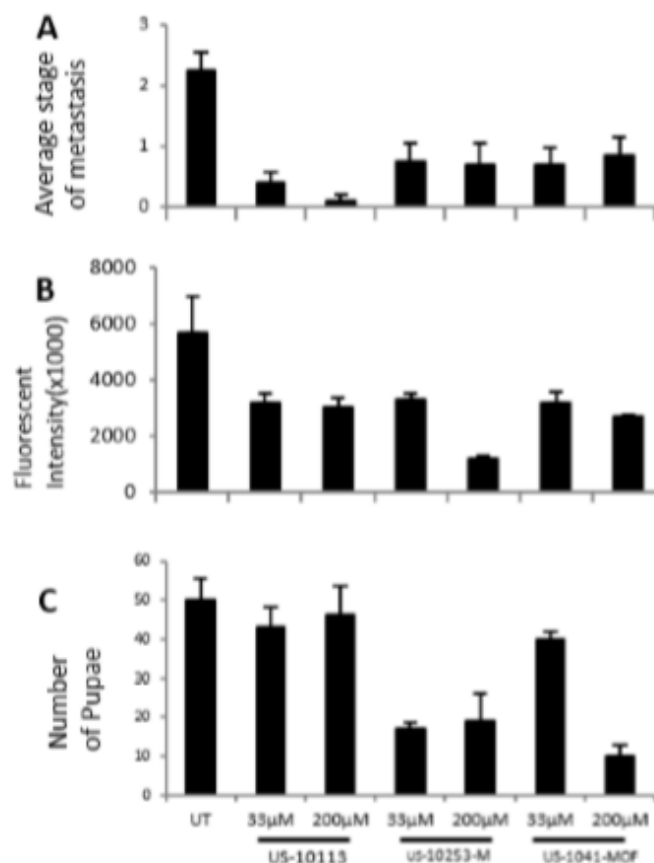


Figure 5.2: Effects of original inhibitors on metastatic spread in Ras^{Val12}. S100A4 recombinant fruit fly larvae for: untreated (UT), **US10113 (1a)**, **US10253-H (1b)** and **US1041-MOF (1c)** at 33 and 200 μM . **5.2A:** Average stage score of metastasis, **5.2B:** Fluorescent intensity from dissected CNS including primary tumours and **5.2C:** Toxicity assay showing number of pupae surviving. Results show mean \pm SE for 20 larvae in **5.2A**, **5.2B** and 50 in **5.2C**, all repeated thrice.

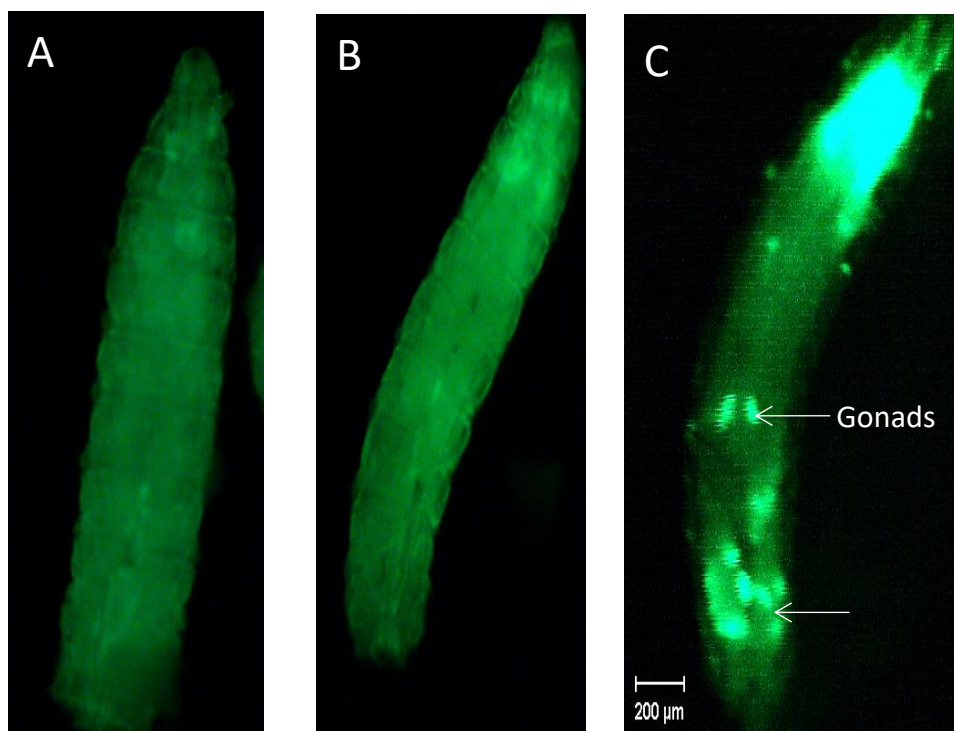
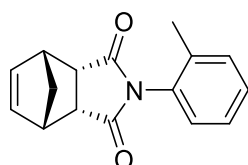


Figure 5.3: S100A4 combines with Ras^{Val12} to induce extensive metastases in *Drosophila* larvae. GFP fluorescence was recorded for larvae with the following genotype: **5.3A:** *S100A4* wt alone, **5.3B:** Ras^{Val12} alone and **5.3C:** $Ras^{Val12}/S100A4$. Only the double transgenic flies produced extensive metastases. Bar=200 μ m.

Due to the enhanced inhibitory effect and low toxicity in the *Drosophila* model, **US10113 (1a)** was selected as the lead inhibitor. **US10113 (1a)** also showed a similar Kd to that of pentamidine at approximately 10^{-6} M.

Compound **US10113** can exist as two stereoisomers, possessing the *exo* and the *endo* configurations respectively (see **Figure 5.4**). The *endo* has been found to be biologically active, whereas the *exo* is believed to be inactive. Synthesis of both the *exo* and *endo* isomers of **US10113** was required as the *exo* isomer could act as a control group for comparison.

5.4A



5.4B

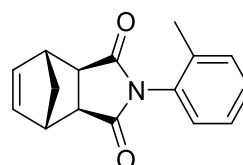


Figure 5.4: Different configurations of **US10113**. **5.4A:** *Endo* isomer= active (**1a**). **5.4B:** *Exo* isomer= inactive (**1d**).

The **US10113 (1a)** inhibitor has also been shown to display inhibitory effects at the micromolar range in Rama 37 cell lines (see **Figure 5.5** and **Section 5.1.1.2**), showing the potential to stop the metastatic spread that is strongly associated with TNBC. The migration is not fully inhibited even at the highest concentration of 100 μM , suggesting the inhibitor alone may not fully treat TNBC.

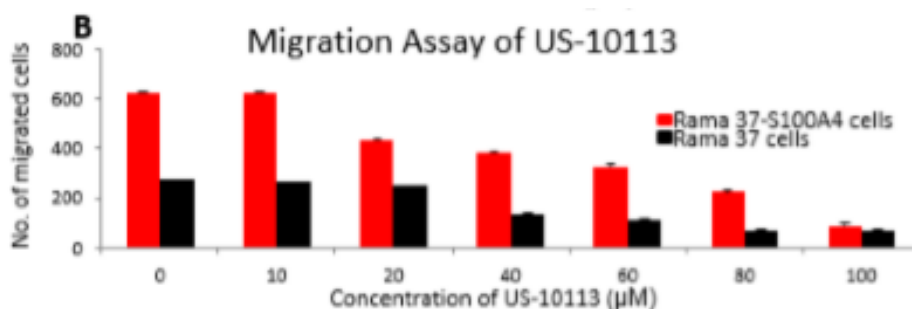


Figure 5.5: Migration assay for **US10113 (1a)** at μM dose on cultured rat mammary cells. **Black:** Benign parental Rama 37 cells. **Red:** Malignant, metastatic S100A4-overexpressing Rama 37-S100A4 cells.

The reliance on small molecule inhibitors is limited (see **Chapter 4, Section 4.3.1**), hence the development of new chemical strategies is increasingly investigated to enhance cancer treatments, as well as other diseases. Incorporating the lead inhibitor into a PROTAC compound may significantly reduce metastasis further than the original inhibitor and produce a selective targeted therapy for the treatment of cancer.

5.2 Aims

The design and synthesis of a PROTAC agent able to degrade the oncogenic S100A4 protein and treat TNBC is the broad aim of the project. By degrading the upregulated S100A4 protein, metastasis can be slowed down or potentially stopped all together. This allows the primary tumour to stay within the breast tissue, which could drastically improve the survival outcome of patients, as the primary tumour is able to be removed *via* surgery.

A PROTAC agent that incorporates the **US10113** inhibitor (**1a**) tethered to a E3 ligase recruiting warhead such as a phthalimide or VHL ligand 1 can be synthesised to polyubiquitinate the S100A4, marking it for proteolytic degradation *via* the proteasome (see **Figure 5.6**).

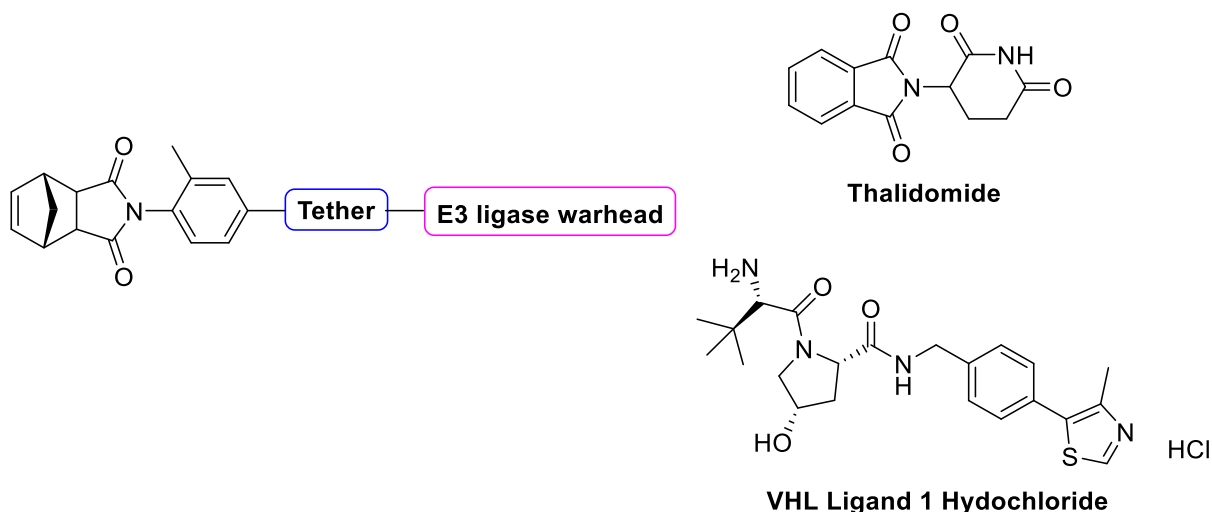


Figure 5.6: Basic structure of the proposed PROTAC compound.

The aim of this project can be broken down into the following:

- Design and synthesise a PROTAC compound incorporating the original hit inhibitor **US10113 (1a)** and the inactive equivalent (**1d**) as a control.
- Assess the antimetastatic effect and DMPK profile of the PROTAC compound.
- Perform extensive SAR exploration into the lead **US10113 (1a)** inhibitor, assessing whether metabolic stability and potency can be improved.
- Applying SAR knowledge to develop new PROTAC compounds with potentially enhanced activity and stability.

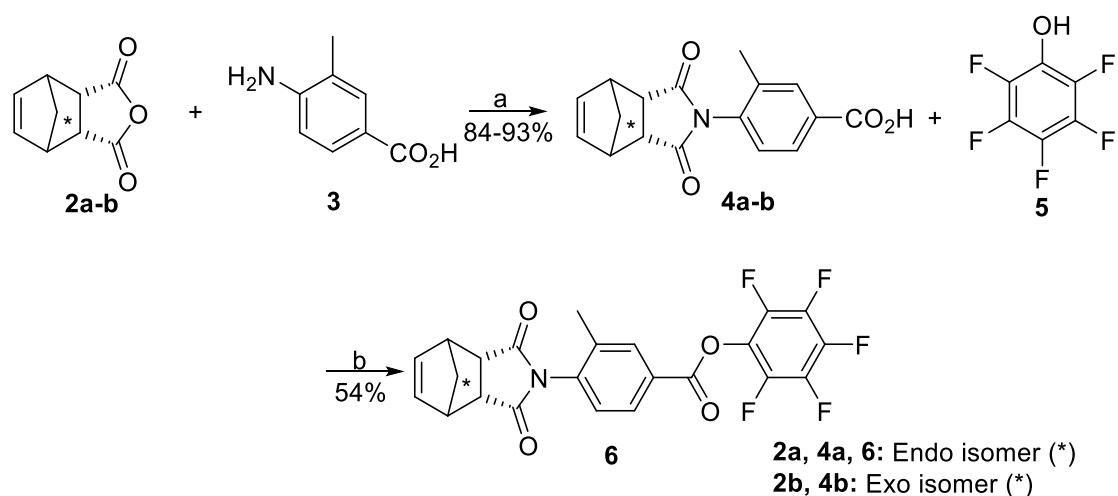
5.3 Development of a PROTAC agent

5.3.1 Discussion

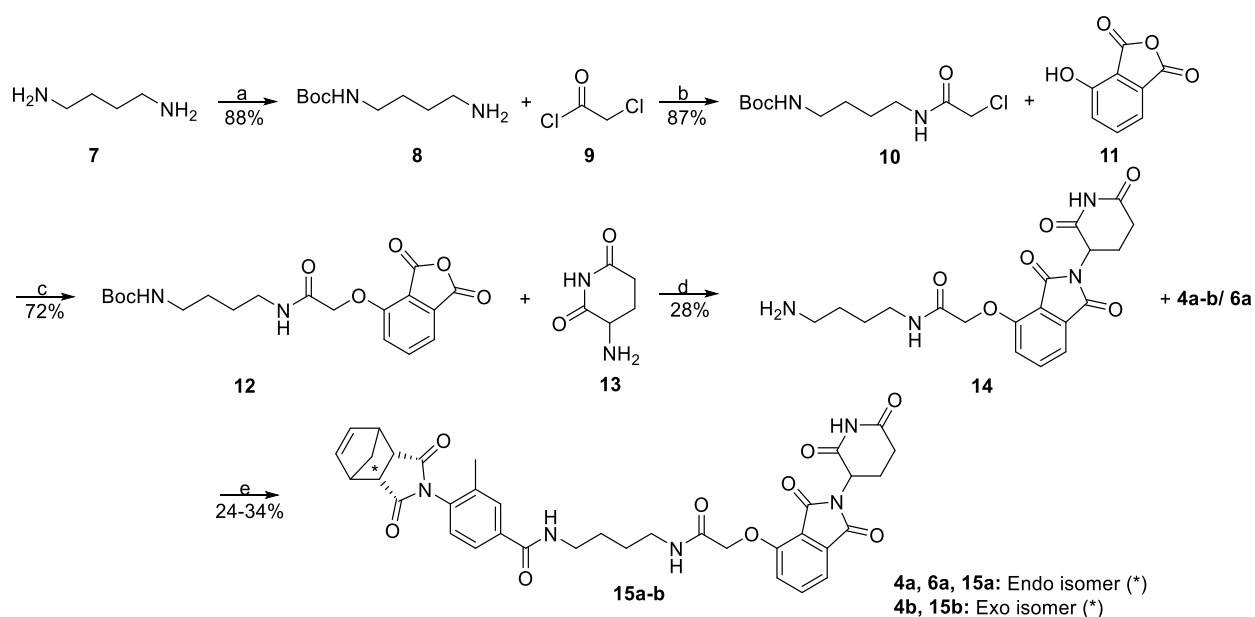
5.3.1.1 Synthesis

For the **US10113 (1a)** inhibitor to be tethered to the E3 ligase warhead, the **US10113** parent compound needs a point of attachment. The original active (*endo*) **US10113 (1a)** and inactive (*exo*) **iUS10113 (1d)** inhibitors were synthesised to assess the differences in activity. An amide tether reported by *Winter et al* was chosen to connect the inhibitor to the E3 warhead.⁷ The carboxyl functionality was added to the **US10113 (1a)** in the *para* position to allow maximum flexibility and lower repulsion. A thalidomide ligand was selected as the E3 ligase warhead due to the perceived ease of synthesis.

The synthesis of the PROTAC involves 7 steps over 2 separate routes that are joined in step 7 (see **Scheme 5.1** and **5.2**).



Scheme 5.1: Reagent and Conditions: (a) AcOH, reflux, 16 hr; (b) EDC.HCl, DMAP, DMF, rt, 16 hr.



Scheme 5.2: Reagent and Conditions: (a) Boc_2O , DCM, rt, 16 hr; (b) DIPEA, THF, 0°C to rt, 16 hr; (c) NaH, THF, 0°C to reflux, 16 hr; (d) KOAc, AcOH, MW 150°C, 12 mins; (e) DMF, rt, 16 hr.

5.3.1.1.1 Synthesis of Inhibitor

Synthesis of the **US10113** carboxylic acid (**3a** and **3b**) (see **Scheme 5.1**) involved the *endo* or *exo* isomer of the anhydride (**2a-b**) reacting with a 3-methyl, 4-amino benzoic acid (**3**) in acid within a screw fix sealed tube overnight. The formation of a pink precipitate for both compounds was observed 1-2 hours into the reaction, however, the reaction was continued overnight in accordance with *Li et al* and to ensure a higher yield.⁸ The reaction TLC showed

faint traces of starting material, therefore purification *via* column chromatography was performed. The active *endo* carboxylic acid (**4a**) had a yield of 84%, whereas the *exo* (**4b**) had a yield of 93%. The intermediates **4a** and **4b** have been directly used in **Scheme 5.2** to synthesise **15a-b** *via* a HATU mediated amide coupling reaction. **4a** has also been further modified with pentafluorophenol (**5**) to activate the carboxylic acid, allowing simpler coupling reactions between **14** and **6** in **Scheme 5.2**. Experimental details are discussed further in **Section 5.3.1.1.2**.

5.3.1.1.2 Synthesis of Amide Tether and PROTAC Agent

Scheme 5.2 initially involves the synthesis of the amide tether and the CRBN warhead. Mono boc protection of 1,4-diaminobutane (**7**) was performed with 0.1 equivalents of boc anhydride. Throughout the project, the yield of this reaction has varied. Low yields of 15-20% have occurred when boc anhydride is rapidly added to a mixture of **7** and solvent. A high volume of solvent is used to ensure a highly dilute reaction to prevent double addition. Despite the high volume of solvent used, the double addition side product was formed which led to poor yields. Purification of this material was incredibly difficult. Good separation between the starting material, desired mono addition (**8**) and the double addition side product was shown on the crude TLC plate. However, this separation did not translate during purification *via* column chromatography. Coelution was a common problem despite numerous eluent systems used and variations in silica length. Small amounts of **8** were able to be obtained before coelution occurred. Further purification of the coeluted material was carried out, which delayed the synthesis, and further lowered overall yields. To improve purification, the work up procedure was manipulated. Originally the work up involved washing the crude mixture with distilled water and brine. A double extraction method was attempted to separate the multiple intermediates by exploiting the differences in polarity of different work up solvents (see **Figure 5.7**).

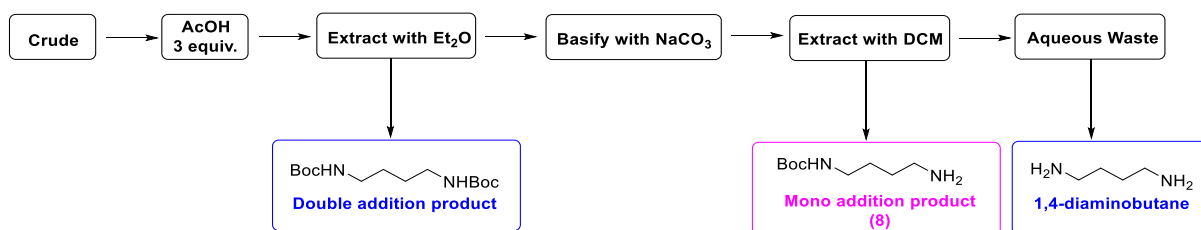


Figure 5.7: Flow chart diagram for the double extraction work up procedure.⁹

The crude mixture was acidified with acetic acid and then the double addition product was extracted with diethyl ether, leaving both the starting material and the mono addition product (**8**) in the water layer. To separate the mono boc product (**8**) from the starting material (**7**), the aqueous layer was basified with NaCO₃, and extracted with DCM to give the desired mono boc protected product (**8**). The starting material remained in the aqueous layer and was suitably disposed of. This double extraction was proposed by Dr Findlay on his personal chemistry blog.⁹ Findlay proposed a 25-45% yield of **8**, with a 90% recovery using this extraction method. However, when this was performed, a 70% recovery and a 21% yield was obtained. This is still a very poor yield, especially at the first step of a long synthesis. Therefore, the reaction was adapted further.

Success was found with a very slow addition of the boc anhydride to a solution of **7** and DCM. A solution of Boc anhydride in DCM was added slowly *via* a dropping funnel at ~1 drop per 5 seconds over 5 hours. This resulted in a significant increase in yield at an impressive 88%, which is similar to yields reported in the literature of 96-100%.^{10, 11, 12} ¹H NMR also confirmed that pure **8** was synthesised, and did not require purification *via* column chromatography.

It should also be noted that this compound is commercially available, and upon occasion **8** was bought and used in **Scheme 5.2**. Although, the synthesis has been adapted so **8** can be synthesised in bulk for the synthetic route.

The second step involves reacting **8** with 2-chloroacetylchloride (**9**) to give **10**. This reaction was relatively straight forward with no known issues. This reaction was adapted from *Fanning et al* to give a yield of 87%, similar to the papers reported yield of 74-83%.^{13, 14} Purification was accomplished *via* column chromatography with no complications. The intermediate is chemically stable, as ¹H NMR has indicated that after several months the compound remains intact when stored in the fridge at 4°C.

The next step in the synthesis involves the assembly of the thalidomide warhead. The synthesis of **12** involves the deprotonation of **11**'s hydroxyl group *via* NaH before an S_N2 reaction with **10**.¹⁵ Initially the synthesis worked well and yields of 64-72% were obtained. However, as the project went on, the yield of the reaction lowered. The ability of NaH to effectively deprotonate **11** was hindered. This has been attributed to the possible contamination of NaH. The protective mineral oil coating NaH was stripped *via* multiple

hexane extractions to enhance the reactivity of NaH. The stripped NaH was able to regain the high yields. Purification *via* column chromatography was required as the reaction did not go into completion, however, no issues such as coelution occurred.

The thalidomide warhead was formed when **12** and **13** were reacted together under acidic conditions. Originally, traditional reflux condensation reactions were performed, which resulted in longer reaction times and lower yields. Microwave conditions involving heating the reaction vessel to 150°C for 12 minutes improved the yields and therefore became the preferred method. Purification was tricky as the intermediate was highly polar. The boc group masking the amine functionality was removed during the reaction due to the presence of acetic acid. TLC analysis using normal phase TLC plates suggested that the crude mixture was far too polar for purification *via* standard purification methods. *Fischer et al* recommended the use of preparative HPLC.¹⁵ Due to complications with the equipment, a manual method was used instead. A column packed with reverse phase silica gel was used to purify the crude mixture. The yield *via* this purification method was low at 12-16%, as coelution occurred during the purification. Alternative techniques such as preparative TLC were attempted with success, however this method was very time consuming due to the small scale of each purification and contributed to low yields of 14%. The use of normal phase silica gel was found to be successful with the correct eluent system. A solvent system referred to as WIPE (water/ isopropanol/ ethyl acetate) 1:2:7 was modified to water/methanol/ ethyl acetate 1:2:7 and used for the purification of **14**. The presence of water and methanol did break down some of the stationary silica phase, causing it to be present in the recovered material. The silica was separated from **14** *via* trituration with acetone. This method was able to obtain **14** in a yield of 28%. This is still a low yield, but higher than other previously attempted purification methods.

Initially, the final synthesis involved a standard amide coupling reaction between **4a/4b** and **14** using the coupling agent HATU. Unfortunately, this resulted in low yields and very difficult purifications due to the high polarity of **RGC (15a)** and **iRGC (15b)** among other highly polar starting materials and by-products. An aqueous work up was not attempted for this procedure, as the polar product would be trapped in the water layer. Therefore, the aqueous work up was skipped, and purification was immediately attempted. Several different purification methods were tried. Purification *via* column chromatography using reverse

phase silica gel was initially attempted with some success, however, small amounts of HATU and 3-aminopiperidine-2,6-dione (**13**) were found in the ^1H NMR spectrum of the recovered material. **iRGC (15b)** was further purified *via* ion exchange cartridges. Ion exchange cartridges are available for the purification of polar compounds formed from amide coupling reactions. These cartridges were successfully used for the purification of **iRGC (15b)**, providing a recovery yield of 24%. The yields are still considered low which could be due to the challenging purification, but also due to the reaction itself. The polar compound **14** was insoluble to several solvents, and therefore full solubility in DMF was not achieved, which contributed to a lower yield.

RGC (15a) was initially purified *via* reverse phase silica gel in the same manner as **iRGC (15b)**, but later switched to normal phase silica gel using the modified WIPE-127 solvent system discussed previously in the synthesis of **14**. Some success was found, although coelution often occurred and some silica gel was present in the product. Silica was removed *via* trituration with acetone.

Modifications to the synthesis of **RGC (15a)** were attempted to lessen the number of polar compounds present in the crude material. Initially the carboxylic acid of **4a** was converted to an acid chloride, a proven method in amide coupling reactions.¹⁶ Unfortunately, telescoping the reaction mixture of **4a's** acid chloride derivative to react with amine (**14**) resulted in an odd ^1H NMR spectrum. Degradation of the amine and odd integration patterns were observed. Excess thionyl chloride may have disrupted the reaction and caused degradation. Lowering the equivalents of thionyl chloride still caused degradation, and as a result the reaction was abandoned.

A milder synthesis to activate the carboxylic acid on **4a** was undertaken. **4a** was coupled to pentafluorophenol, aided by the coupling reagent 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), to give **6**.¹⁷ The attachment of pentafluorophenol allows the carboxylic ester to be highly activated. Purification *via* an aqueous work up and column chromatography afforded **6** in a 54% yield. No complications during the purification methods were experienced.

The pure activated **6** was then coupled directly to the amine (**14**) and allowed to stir at 50°C in DMF. Excess **14** and other by-products were observed on the reaction TLC plate, resulting

in the low yield. The material was purified *via* column chromatography with normal phase silica using the modified WIPE127 solvent system. Silica was separated from the product *via* trituration with acetone. If trace amounts of starting materials or by-products remained, ion exchange cartridges were used to afford pure product. This method improved the yield of **RGC (15a)** to 34%, a 27% difference compared to the original HATU coupling method. The final PROTAC compound was analysed *via* ^1H and ^{13}C NMR, mass spectroscopy, IR spectroscopy, and purity data were gathered using analytical HPLC. The purity of the compound was found to be 95%.

It should be noted that the final PROTAC agent is susceptible to chemical instability. **RGC (15a)** was initially stored at 4°C. Over the course of 3-4 months, the PROTAC was reanalysed *via* ^1H NMR. It was found that over time the PROTAC degraded, as the integration patterns between the inhibitor and the amide tether altered. It was concluded that the amide bond of the tether was hydrolysed by trace amounts of water. The hydrolysis caused the separation between the two segments of the PROTAC agent. The ^1H NMR spectrum suggests that the thalidomide warhead is still attached to the tether and therefore there is favourable degradation at the connection between the tether and the inhibitor (**US10113**). The PROTAC agents are now stored in a freezer at -18°C.

5.3.1.2 Biology

The PROTAC's **RGC (15a)** and to a lesser extent **iRGC (15b)** underwent various *in vitro* testing in a range of different models such as the Drosophila model, the rat cell line (Rama 37), and the TNBC human cell line MDA-MB-231. The effectiveness of the compounds was analysed and compared to the original inhibitor **US10113 (1a)**.

5.3.1.2.1 Cell Line Classification

5.3.1.2.1.1 Rama Cell line

The rat mammary cell line is frequently used in the project. The cell line described as Rama 37 are benign, non-metastatic cells that have no detectable levels of the S100A4 protein. This cell line is used as a control. The Rama 37-S100A4 cell line are engineered cells that over express S100A4 and are used to monitor the inhibition of metastasis.^{18, 19}

5.3.1.2.1.2 Human Cell Line

The human cell line known as MDA-MB-231 is commonly used in the project. MDA-MB-231 is a highly aggressive and invasive TNBC cell line, which lacks hormone and HER2 receptors.

Therefore, this cell line is highly suitable for the study between cancer migration and invasive upon the addition of S100A4 inhibitors and related PROTAC agents.^{20, 21}

5.3.1.2.1.3 *Drosophila Model*

The *Drosophila* model involving transgenic fruit flies are used to help define the S100A4 functions that cause metastatic dissemination. In flies, the overexpression of mutant Ras^{Val12} and S100A4 increases the activation of the stress kinase JNK and the production of the matrix metalloproteinase MMP1, which leads to metastatic spread. Genetic or chemical blockades of JNK and MMP1 suppresses the metastatic activity associated with high S100A4 concentrations. To investigate the protein, human S100A4 can be expressed in the fruit fly *Drosophila melanogaster*. The expression is targeted in the developing eye lobes and is absent in the brain or the rest of the CNS.^{22, 23, 24} *Drosophila* have conserved signal transduction pathways for cell cycle, growth control, and cell-to-cell communication.²⁵ Additionally, 70% of human cancer genes are conserved in the *Drosophila* model, but importantly none of the S100 family proteins are present.²⁶ Overexpression of oncogenic Ras (Ras^{Val12}) causes the formation of tumours in the epithelial tissues of *Drosophila*.²⁷ These tumours can be transformed into a malignant phenotype by disruption of suppressor genes such as scribble (*scrib*) and lethal 2.²⁸ They show, after multiple crosses, that the S100A4 gene is required to disseminate Ras^{Val12} tumour cells from the optic lobes to the ventral nerve cord (VNC) and further afield in fly larvae.²⁷ Therefore, these transgenic flies are a fantastic model to monitor inhibitors on the effect of S100A4 induced cell migration.⁶

5.3.1.2.2 *Drosophila Data*

The PROTAC **15a**, also referred to as **RGC** underwent *in vitro* testing in the *Drosophila* model Ras^{Val12} S100A4 transgenic fly. A series of data was gathered using different techniques.

Western blot analysis and quantification of S100A4 in flies with **RGC (15a)** were tested and the bands were normalised to actin over three experiments for standard error (see **Figure 5.8**). **RGC (15a)** at 10 nM begin to rapidly decrease the S100A4 concentration and is no longer observed by Western blot analysis at 50 nM. This indicates that **RGC (15a)** can successfully degrade the protein at low concentrations.

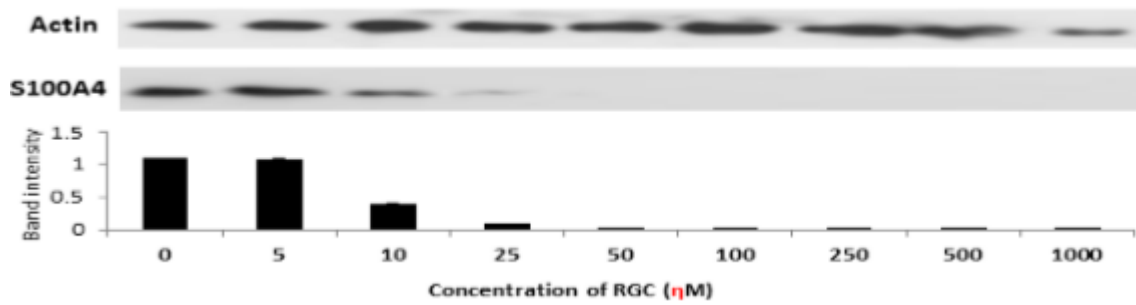


Figure 5.8: Western blot and quantification of S100A4 in flies with **RGC (15a)**, mean of 25 larvae \pm SE n=3.

To assess whether the degradation and disappearance of the S100A4 influences the metastatic properties of the cancer cells, the inhibition of the metastatic spread was quantified (see **Figure 5.8**). The concentration is plotted against average stage of metastasis. The average stage of metastasis describes the size of a cancer, how fast it has grown, and whether metastatic spread has occurred. Stage 1 refers to cancer that is relatively small and contained within the organ it originated in. Stage 2 usually means that the tumour is larger than in stage 1, but the cancer has not started to spread into the surrounding tissues. However, it can sometimes mean the cancer cells have spread into the lymph nodes close to the tumour (depends on the type of cancer). Stage 3 refers to cancer that is larger than stage 2 and may have spread into surrounding tissues, and cancer cells are present in the lymph nodes. Stage 4 (not shown on the graph) means the cancer has spread from where it started to another body organ, this is typically called secondary or metastatic cancer.²⁹

Figure 5.9 shows that the cancer in the flies start off at stage 2 of metastasis and at 10 nM the staging halves bringing average stage of metastasis to stage 1. Upon 50 nM metastatic spread of the cancer cells stopped as the average stage of metastasis was close to stage 0. This supports the hypothesis that as the concentration of the S100A4 decreases by degradation, the metastasis capabilities on the cancer cells also decrease.

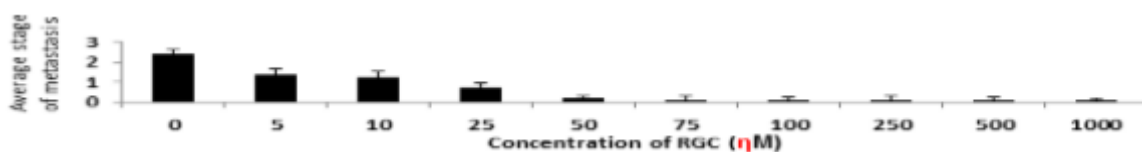


Figure 5.9: Inhibition of metastatic spread by **RGC (15a)**; mean \pm SE of 50 larvae.

To further illustrate the effectiveness of the PROTAC **RGC (15a)**, the metastasis spread was compared to the original inhibitor and the CRBN E3 ligase warhead; thalidomide alone (see **Figure 5.10 and 5.11**).

The **US10113 (1a)** inhibitor has some effect on the metastatic spread (see **Figure 5.10**). The average stage of metastasis starts off at 3, which is higher than with the PROTAC agent. At 70 μM the stage drops to 1, and then below stage one at 100 μM . The **RGC (15a)** is considerably more potent than the inhibitor (50 nM (**RGC (15a)**) versus 100 μM (**US10113 (1a)**)). The high potency of **RGC (15a)** suggests that the drug is considerably more effective than the inhibitor alone. Lower doses at a nanomolar scale may be administered, lowering the potential of adverse drug reactions, resulting in a more attractive therapeutic agent for the treatment of TNBC.

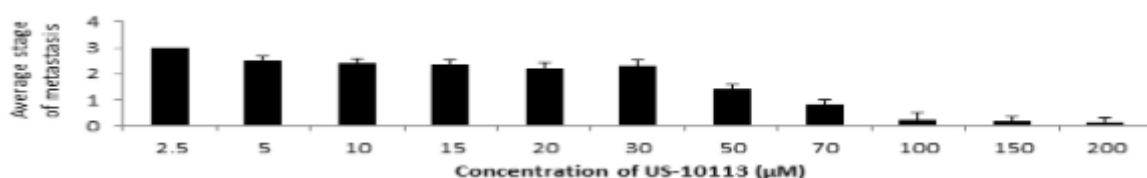


Figure 5.10: Inhibition of metastatic spread by **US10113 (1a)**; mean \pm SE of 50 larvae.

If we consider the CRBN warhead thalidomide (see **Figure 5.11**), we can see that the warhead alone has no effect on the inhibition of metastasis spread, as the average stage of metastasis remains at 3. The S100A4 is not being brought into close proximity to the E3 ligase, and therefore attachment of the polyubiquitin chain to mark it for degradation is not occurring. This confirms that a heterobifunctional molecule containing both the inhibitor and the CRBN ligand are necessary for the functioning of a PROTAC.

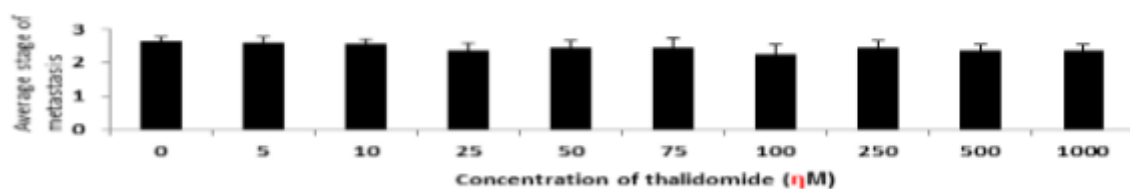


Figure 5.11: Inhibition of metastatic spread by thalidomide; mean \pm SE of 50 larvae.

As mentioned previously, the highly potent nature of the PROTAC agent would suggest that lower doses of the anticancer agent could be administered, and therefore reduce the

incidence of adverse drug reactions. This was further supported by the toxicity analysis obtained for **RGC (15a)** (see **Figure 5.12, 5.21, 5.31**).

The toxicity was derived from the number of pupae that survived the increasing doses of **RGC (15a)**. **Figure 5.12** shows that most of the pupae survived suggesting that the PROTAC agent has a low toxicity profile. The number of pupae begin to lower at 2000 nM. This is not considered a huge problem, as inhibition of metastatic spread was observed at 50 nM (see **Figure 5.8**), suggesting that the dosing regimen would be lower than 200 nM.

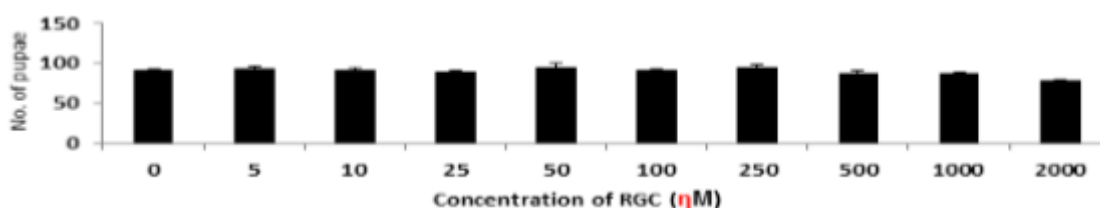


Figure 5.12: Toxicity of **RGC (15a)** in flies. Mean \pm SE n=3.

The toxicity of **RGC (15a)** was compared to the original inhibitor (see **Figure 5.13**). **US10113 (1a)** was dosed at the micromolar range to suit its inhibitory concentration. The inhibitor was shown to possess toxicity as the number of pupae lowered at the first administered dose of 33 μ M. Only half of the pupae population survived at 200 μ M. The same toxicity may be seen if **RGC (15a)** was administered at the same concentration. However, the PROTAC does not need to be dosed this high as it displays highly potent behaviour at the nanomolar range. From comparing the results, the PROTAC agent is considerably less toxic than the original inhibitor.

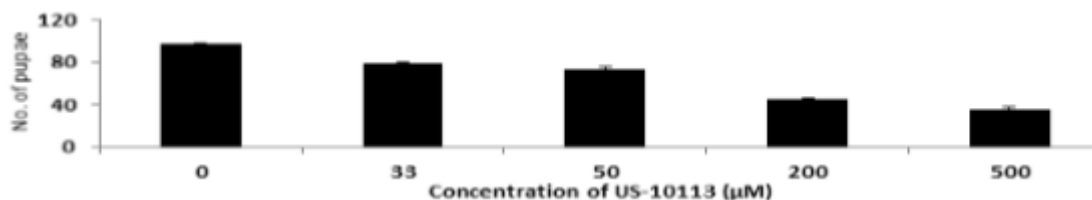


Figure 5.13: Toxicity of **US10113 (1a)** in flies. Mean \pm SE n=3.

5.3.1.2.3 Rat Cell Line Data

RGC (15a) was also applied to the TNBC rat cell line Rama 37 to assess the effects and compare the results to the *Drosophila* model.

Assessment of **RGC (15a)** capability to degrade S100A4 was studied *via* Western blot analysis (see **Figure 5.14**). Similarly, to the Drosophila model, the band was normalised to actin. Slight degradation of S100A4 was observed at 50 nM and had completely disappeared between 100-250 nM. The dosing for degradation is slightly higher in the Rama37 cell line compared to the Drosophila model. Although, degradation is still occurring at the nanomolar scale.

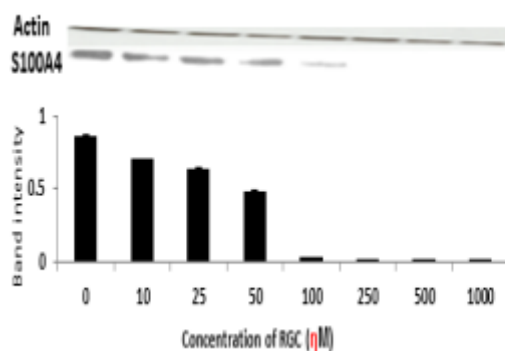


Figure 5.14: Western blot and quantification of S100A4 in Rama 37/ S100A4 clone cells with **RGC (15a)**. Mean \pm SE n=3.

The low toxicity we see in the toxicity models of all three cell lines tested (Drosophila, Rama 37 and MDA-MB-231) have been partly attributed to the low concentrations required to cause degradation. Low toxicity may also be ascribed to the high selectivity of the PROTAC agent towards the S100A4 protein. Western blot analysis of the S100P protein, which is closely related to the S100A4 protein was performed as a direct comparison (see **Figure 5.15**). The band intensity of the S100P remained the same as the concentration of **RGC (15a)** went from 10-1000 nM. This suggests high selectivity for the S100A4 protein, which may also contribute to its low toxicity profile.

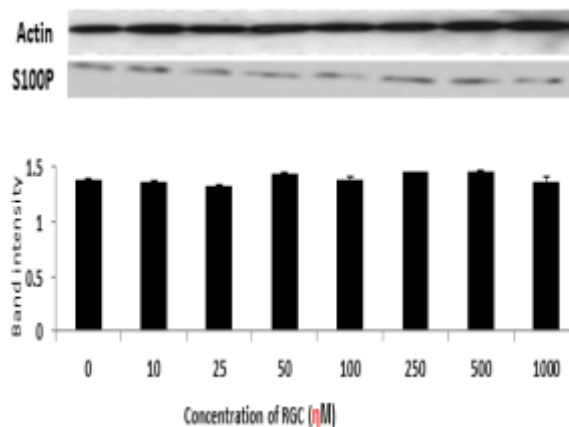


Figure 5.15: Western blot quantification of S100P in Rama 37/ S100P clone cells with **RGC (15a)**.
Mean \pm SE n=3.

The migration assay of **RGC (15a)** was performed on the Rama 37 cell line (see **Figure 5.16**) and the percentage relative migration dropped considerably at 10 nM to approximately 10% and continued at this level as the concentration of the PROTAC increased. The PROTAC agent in the rat cell line performs similarly to the Drosophila model.

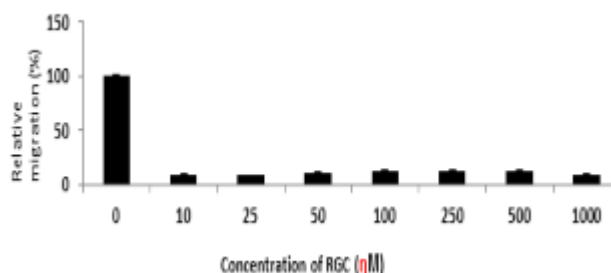


Figure 5.16: Migration assay of Rama37 S100A4 clone with **RGC (15a)**. Mean \pm SE n=3.

An invasion assay of the active PROTAC was also performed (see **Figure 5.17**). The migration and invasion are common features for a cell motility; however, they differ. The migration assay refers to normal cell movement, whereas invasion describes cells actively invading surrounding tissues due to chemoattractant agents.^{30, 31} The invasion assay shows that **RGC (15a)** can reduce relative invasion by more than 50% at 20 nM, and is continually reduced as the concentration increases up to 1000 nM. This further supports that **RGC (15a)** is highly potent and potentially able to treat invasive TNBC.

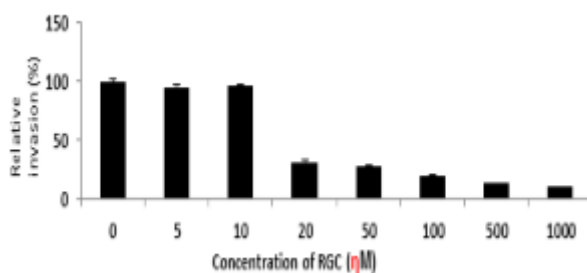


Figure 5.17: Invasion assay of Rama 37/ S100A4 clone cells with **RGC (15a)**. Mean \pm SE n=3.

The migration assay of **RGC (15a)** was compared against the original inhibitor **US10113 (1a)** and the CRBN ligand thalidomide (see **Figure 5.18** and **5.19**).

The **US10113 (1a)** was able to reduce the relative migration by 50% at 50 μ M and migration had almost stopped at 100 μ M (see **Figure 5.18**). **US10113 (1a)** is dosed at the micromolar range compared to the PROTACs nanomolar range. Therefore, **US10113 (1a)** can inhibit the S100A4 protein, but is noticeably less potent.

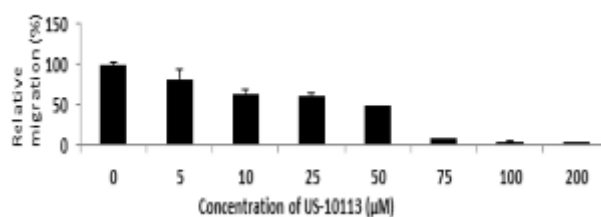


Figure 5.18: Migration assay of Rama37/ S100A4 clone cells with **US10113 (1a)**. Mean \pm SE n=3.

The CRBN binding warhead, thalidomide displayed no effect on the relative migration of the cancer cells, like the results in the Drosophila model (see **Figure 5.19**). This supports the concept that thalidomide does not possess any inhibitory or degradation effects on the protein itself and must be tethered to the inhibitor to degrade S100A4.

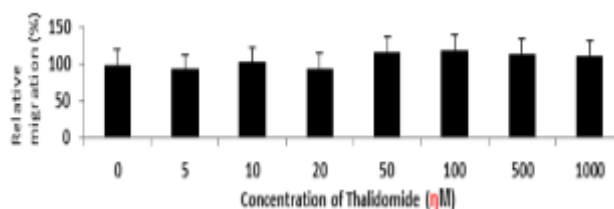


Figure 5.19: Migration assay of Rama 37/ S100A4 clone cells with thalidomide. Mean \pm SE n=3.

A migration assay of the effects of **RGC (15a)** on the S100P protein was also investigated (see **Figure 5.20**) to further support the evidence that the PROTAC can selectively target the S100A4 protein. Between 5-1000 nM the PROTAC has no substantial effect against the relative migration of the cancer cells, suggesting selectivity against S100A4 compared to the closely related S100P protein. However, at 2000 nM the relative migration decreases below 50%. There could be several possibilities for this decline, one of which could be that the high concentration of the compound causes damage to the cells and therefore limits their migration pattern. Also, at a high enough dose the PROTAC can have a degradation effect on the S100P protein. This issue with selectivity may not be an issue as the PROTAC agent is unlikely to be dosed at such a high concentration, as it typically displays a potent effect at 50 nM.

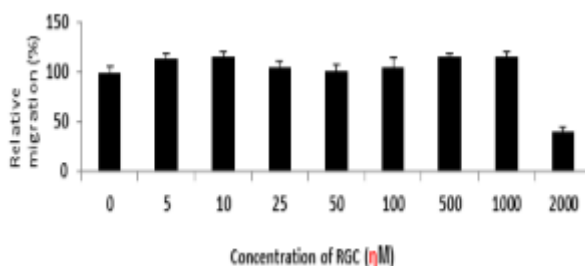


Figure 5.20: Migration assay of Rama 37/ S100P clone cells with **RGC (15a)**. Mean \pm SE n=3

The toxicity of the **RGC (15a)** against the rat cell line was also assessed (see **Figure 5.21**). As the dose is increased from 0.05-100 μ M, toxicity is negligible. In the Drosophila model, the toxicity assay was performed on a nanomolar scale, and it was assumed that if the PROTAC was applied on a micromolar scale, toxicity would be observed. In this assay, even at the micromolar range, toxicity remains very low, increasing its suitability to be administered to patients with TNBC.

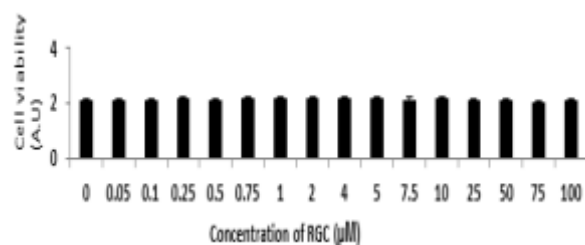


Figure 5.21: Toxicity of **RGC (15a)** on Rama 37 cells. Mean \pm SE n=3.

The toxicity profile of **US10113 (1a)** was also assessed in the rat cell line (see **Figure 2.22**), with results comparable to those found in the Drosophila model (see **Figure 5.12**). The toxicity model was performed on the micromolar scale for a direct comparison with **RGC (15a)**. Toxicity is low, however, slightly more toxic than **RGC (15a)** as the number of cells starts to decrease at 25 μM and continues up to the maximum concentration of 500 μM .

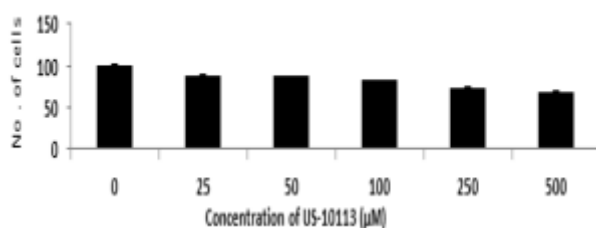


Figure 2.22: Toxicity of **US10113 (1a)** on Rama 37 cells. Mean \pm SE n=3.

5.3.1.2.4 Human Cell Line Data

The TNBC human cell line MDA-MB-231 was also used for the testing of **RGC (15a)**. Similar results were observed for this assay compared to the rat and Drosophila model.

The results of the Western blot quantification translate to the results shown for the Rama 37 and the Drosophila model. The expression of S100A4 protein is significantly reduced as **RGC (15a)** is administered. In as little as 10 nM, the band intensity of S100A4 reduces to almost 0, proving to be an effective degradatory agent for the protein (see **Figure 5.23**).

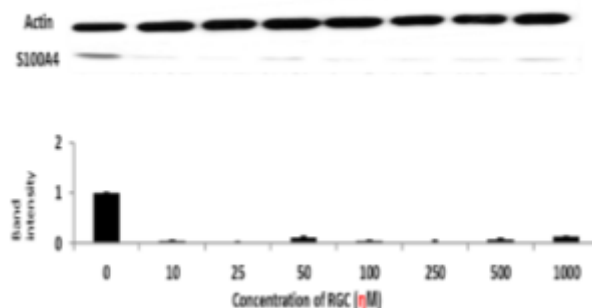


Figure 5.23: Western blot and quantification. S100A4 expression in MDA-MB-231 cell with **RGC (15a)**. Mean \pm SE n=3.

Selectivity of the PROTAC agent was again tested against the S100P protein (see **Figure 5.24**). **RGC (15a)** had little to no effect on the band intensity of the protein, with only a slight decrease at 50 nM as the band intensity went from 1.5 to 1.25. At 1000 nM the band intensity went to 0.8. A reduction is observed, but not on the same scale as with the S100A4 protein.

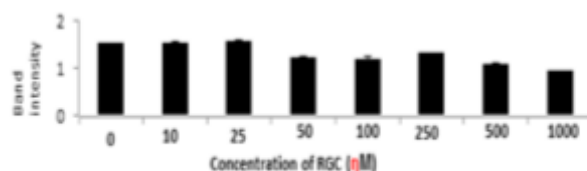


Figure 5.24: Quantification of Western blot of S100P in MDA-MB-231 cells with **RGC (15a)**. Mean \pm SE n=3.

This Western blot analysis was compared to the inhibitor, and applied to both the S100A4 and S100P to determine selectivity towards S100A4 (see **Figure 5.25**). The inhibitor in the human cell line is able to degrade the S100A4 protein, but at a much higher concentration between 100-200 μ M. The inhibitor was shown to display little to no effect on the S100P protein. Therefore, selectivity of the inhibitor is maintained and the antimetastatic effect is achieved at a much higher concentration than the highly potent **RGC (15a)** PROTAC.

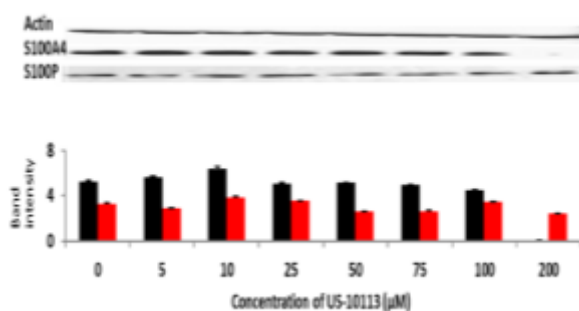


Figure 5.25: Western blot quantification. S100A4 (**black**) and S100P (**red**) expression in MDA-MB-231 with **US10113 (1a)**. Mean \pm SE n=3.

Compound **iRGC (15b)**, which is the *exo* inactive isomer of the PROTAC was assessed *via* Western blot quantification against the S100A4 protein (see **Figure 5.26**). **iRGC (15b)** up to 0.5 μ M was unable to substantially degrade the S100A4 protein. The *exo* isomer is not able to degrade the protein until a higher dose is reached (1 μ M), supporting the SAR evidence and allowing it to be used as a control for future testing.

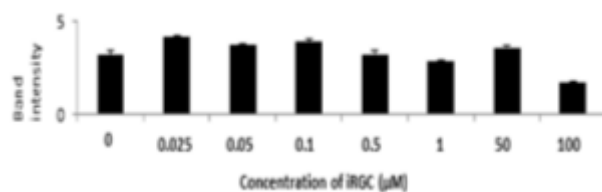


Figure 5.26: Quantification of Western blot S100A4 expression in MDA-MB-231 cells with iRGC (15b). Mean ± SE n=3.

The migration assay of RGC (15a) was performed on the human cell line and similar results have been found compared to the rat and fly cell lines. RGC (15a) significantly reduces relative migration at 10 nM and keeps migration at a constant low value as the concentration is increased from 10 -1000 nM (see Figure 5.27).

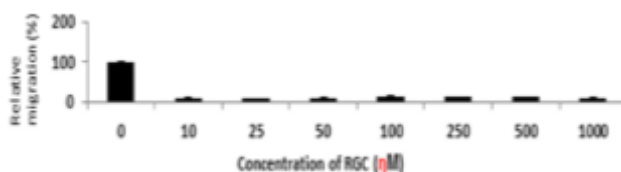


Figure 5.27: Migration assay of MDA-MB-231 cell with RGC (15a). Mean ± SE n=3.

The migration assay of US10113 (1a) on the human cell line was also assessed and shows a reduction of the relative migration at a higher concentration than the RGC (15a) compound (see Figure 5.28). At 50 µM relative migration reduces to approximately 50% and almost no migration is observed between 75-200 µM.

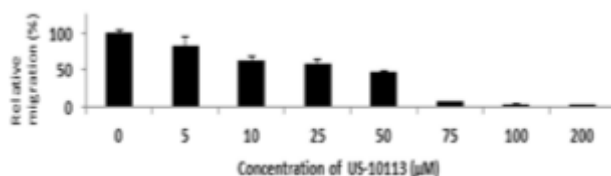


Figure 5.28: Migration assay of MDA-MB-231 cells with US10113 (1a). Mean ± SE n=3.

The thalidomide warhead for the CRBN was tested and showed the comparable results with the other two assays. Thalidomide displayed no effects on the relative migration of the cancer cells by itself (see Figure 5.29).

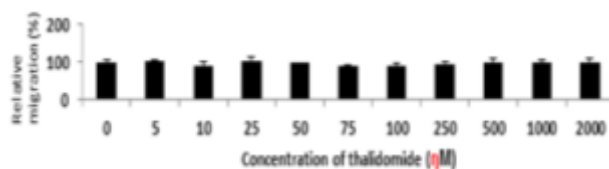


Figure 5.29: Migration assay of MDA-MB-231 cells with thalidomide. Mean \pm SE n=3.

The migration assay was also applied to the **iRGC (15b)** compound (see **Figure 5.30**). Again, if we compare the data from the active *endo* **RGC (15a)**, drastic differences in the migratory patterns are witnessed. **iRGC (15b)** has no inhibitory effects on migration and therefore supports the claim that the *exo* isomer possesses no antimetastatic effects.

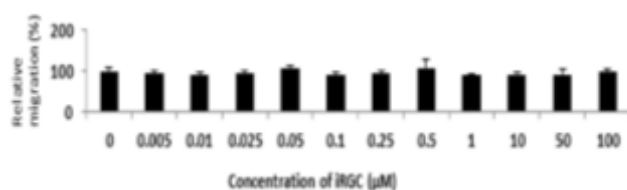


Figure 5.30: Migration assay of MDA-MB-231 cells with **iRGC (15b)**. Mean \pm SE n=3.

No toxicity of **RGC (15a)** was shown in the human cell line as we range from 10-1000 nM (see **Figure 5.31**). The three different assays have all shown a very low toxicity profile, signifying high promise as a future treatment.

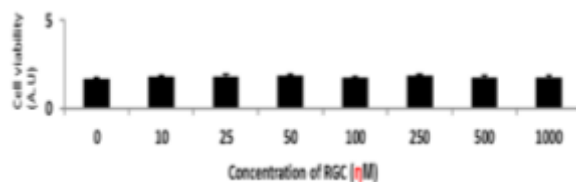


Figure 5.31: Toxicity of **RGC (15a)** on MDA-MB-231 cells; mean \pm SE of 8 wells.

The **US10113 (1a)** inhibitor displays very low toxicity on the human cell line, with only a slight reduction in cell viability at 100 μ M (see **Figure 5.32**). Both the PROTAC and inhibitor are well tolerated in the human cell line. However, the PROTAC is a more suitable treatment as it can be dosed at the nanomolar range and has greater effects on relative migration.

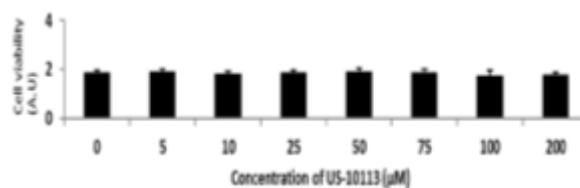


Figure 5.32: Toxicity of **US10113 (1a)** on MDA-MB-231 cells; mean \pm SE of 8 wells.

5.3.1.2.5 Preliminary *In Vivo* Data

Preliminary *in vivo* data was gathered for the current lead PROTAC and inhibitor. Toxic effects were determined by blood tests and immunohistochemical stained sections by autopsy for levels of S100A4. Immunohistochemistry (IHC) applies antibodies to determine the tissue distribution of an antigen of interest in health and disease.³² The expression and distribution of the S100A4 protein was analysed *via* the IHC technique when treated with various doses of **RGC (15a)**, **US10113 (1a)** and thalidomide (see **Table 5.1**).

As **RGC (15a)** is applied intravenously, the amount of S100A4 proteins present in the spleen and colon is reduced as we increase the dosage of the PROTAC. This suggests that the PROTAC can block metastasis to a point and the protein is remaining in the tumour. This then has the potential to stop the formation of secondary tumours. **US10113 (1a)** and thalidomide give similar results from the staining, suggesting that the inhibitor alone is unable to prevent metastasis to the same level as **RGC (15a)**. The thalidomide through various *in vitro* evaluations (see **Figure 5.11**, **5.19** and **5.29**) has shown to possess no biological effects on the degradation of the protein. Therefore, metastasis will occur and the S100A4 will be present in the spleen and colon tissue as the cancer cells migrate.

In the subcutaneous route of administration, **RGC (15a)** can significantly reduce the spread of the protein to the parenchymal tissues further and lower the protein levels in the tumour itself. This suggests that only a low dose is required, and a subcutaneous route into a mammary fat pat is preferable over iv.

The protein is still shown in the spleen and colon tissues, suggesting that some migration is still occurring. Therefore, further development of the PROTAC agent is required to make it a more effective treatment against cancer metastasis.

Route of Administration	Inhibitor (mg/ kg)	Tissue Staining ^a		
		Spleen	Colon	Tumour
		Percent \pm SD (n) (P) ^b	Percent \pm SD (n) (P) ^b	Percent \pm SD (n) (P) ^b
Intravenous cells (iv)	Untreated	32.5 \pm 8.7 (4)	37.5 \pm 8.7 (4)	40 \pm 18 (5)
	RGC (15a) (0.07)	17.0 \pm 4.5 (5) (P= 0.01)	12.5 \pm 6.1 (5) (P= 0.0014)	17.5 \pm 10.6 (2) (P= 0.17)
	RGC (15a) (0.7)	10.5 \pm 2.7 (5) (P= 0.001)	10.0 \pm 0.0 (5) (P= 0.008)	50 \pm 0 (1) (P= 0.64)
	RGC (15a) (3.5)	13.7 \pm 7.5 (4) (P= 0.017)	17.5 \pm 10.4 (4) (P= 0.026)	45 \pm 0 (1) (P= 0.81)
	US10113 (1a) (13.5)	26.2 \pm 8.5 (4) (P= 0.34)	40.0 \pm 10.8 (4) (P= 0.73)	42 \pm 15 (4) (P= 0.83)
	Thalidomide (1.35)	26.5 \pm 13.2 (4) (P= 0.46)	37.0 \pm 4.5 (5) (P= 0.91)	35 \pm 21 (3) (P= 0.76)
Cells s/ c into mammary fat pad	Untreated	23.5 \pm 3.7 (7)	39.3 \pm 10.2 (7)	63.6 \pm 18.2 (7)
	RGC (15a) (0.07)	9.9 \pm 2.6 (7) (P< 0.0001)	17.1 \pm 9.1 (8) (P= 0.001)	13.5 \pm 18.6 (7) (P= 0.0003)
	RGC (15a) (3.5)	12.5 \pm 5.0 (8) (P= 0.0004)	16.9 \pm 10.7 (8) (P= 0.0012)	29.6 \pm 30.2 (7) (P= 0.026)

Table 5.1: Immunohistochemical- staining for S100A4 in inhibitor- treated mice

^a: Immunohistochemical staining of parenchymal (spleen, colon) or carcinoma (tumour) cells as a percentage \pm SD (n= number of tissues or tumours analysed).

^b: Probability of difference from untreated mice calculated using students t-test (2 sided), P<0.05 considered significant.

Further preliminary *in vivo* data was collected using two different routes of administration (see **Table 5.2**). Different doses of **RGC (15a)**, **US10113 (1a)** and thalidomide were utilised to assess their effects on the invasive and metastatic nature of TNBC cells. **RGC (15a)** was able to lower the incidence of metastasis *via* the iv route and was most significant at the 0.07 and 0.7 mg/ kg dose. The incidence of metastasis elevated slightly as the dose of **RGC (15a)** increased. **US10113 (1a)** was not particularly effective in inhibiting metastasis, as its incidence was 4 out of 5. For the iv route, no invasion data was collected. The orthotopic injection into mammary fat pad was able to obtain invasion and metastasis data. **RGC (15a)** at both 0.07 and 3.50 mg/ kg were able to reduce the incidence of invasiveness, however, the lower dose of 0.07 mg/ kg performed better than the higher dose. This was also seen for metastasis, as 0.07 stopped all metastasis in 7 mice, whereas the higher dose had an incidence of 4 out of 8.

As the dose of **RGC (15a)** is increased, the degradative behaviour and effects on invasion and metastasis are hindered. This could be due to the hook effect. The hook effect is an immunological phenomenon whereby the effectiveness of antibodies to form immune complexes is sometimes impaired when concentrations of an antigen are very high. This phenomenon has often been associated with PROTAC agents, in which their high concentrations prevent substrate degradation due to saturation of one of the bound target proteins. This prevents efficient interaction between the target protein and the E3 ligase.³³

³⁴ The presence of the hook effect could contribute to two adverse consequences: (i) An increase in off-target degradation. Hence, further investigations into the selectivity for S100A4 beyond S100P are needed to ensure no adverse hook effect consequences occurs in patients. (ii) Potential for the PROTAC-S100A4 binary complex to cause different biological effects other than the desired target degradation. This can complicate toxicological profile of the drug. Extensive exploration into large dose ranges may be required to fully evaluate the toxicity and the different pharmacological responses that can occur.³⁵ However, the hook effect might not be the reason for the PROTACs *in vivo* behaviour as this is not observed in the previous *in vitro* experiments. Repeated *in vivo* investigations are required to determine whether the PROTACs lack of degradative behaviour at high doses are accurate.

Route of ^a administration	Inhibitor	Dose (mg/ kg)	Invasion ^b incidence	P ^c	Metastasis ^d incidence	P ^c
Intravenous cells together with drug (iv)	None	-	n.a.		4 of 5	-
	Thalidomide	1.35	n.a.		3 of 5	0.22
	US10113 (1a)	13.5	n.a.		4 of 5	0.50
	RGC (15a)	0.07	n.a.		1 of 5	0.024
	RGC (15a)	0.70	n.a.		1 of 5	0.025
	RGC (15a)	3.50	n.a.		2 of 5	0.083
Orthotopic cells into mammary fat pad (s/ c) together with drug (iv)	None	-	8 of 8	-	8 of 8	-
	RGC (15a)	0.07	1 of 7	0.0014	0 of 7	0.0002
	RGC (15a)	3.50	2 of 8	0.0035	4 of 8	0.038

Table 5.2: Effect of the PROTAC, **RGC (15a)** on invasion and metastasis in TNBC mouse cells

^a: Mouse TNBC cells injected intravenously (iv) or orthotopically into mammary fat pad (subcutaneously) and the drug injected every 2 days (iv). Mice sacrificed after total of 4 weeks.

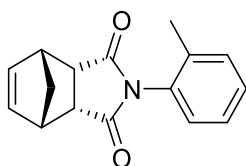
^b: Invasion incidence: number with muscle invasion. Number of mice with tumours, n.a. = not applicable.

^c: Probability (P) of the difference from no inhibitor calculated using Fisher's Exact test, 2-sided, P<0.05 considered significant.

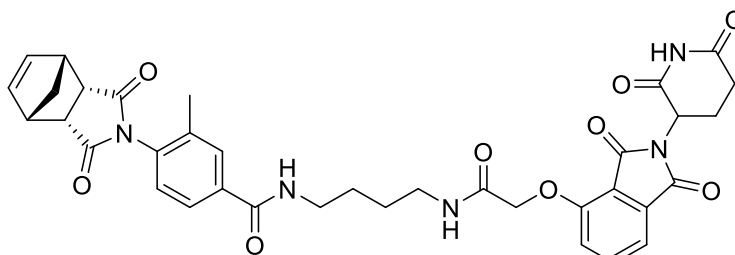
^d: Metastasis incidence: number with experimental lung metastasis/ number of mice for iv route or number with lung metastasis/ number of mice with tumour for s/ c route.

5.3.1.2.6 DMPK Data

DMPK data was outsourced to Cyprotex to assess various properties of **RGC (15a)**, which include physicochemical properties and the metabolic profile. The data was compared to the DMPK data of the **US10113 (1a)** (see **Table 5.3**). Both compounds are highly soluble in water. The $\text{LogD}_{7.4}$ is lower for the PROTAC than the inhibitor, making it less lipophilic and therefore may struggle to cross cell membranes to cause a pharmacological effect. The *in vitro* DMPK assays show striking differences between the two compounds. Mouse and human microsomal clearance were significantly higher in the PROTAC agent, this is not surprising due to the large number of metabolic sites that the PROTAC agent has. The hepatocyte clearance is also high but cannot be compared to the inhibitor as this data was not obtained. The hepatocyte clearance in the human model was lower than the mouse, which could be promising. The metabolic stability is an issue with this compound due to the presence of the amide tether and the norbornene double bond, which is further discussed later in this chapter (see **Section 5.4.1**).



US10113 (1a)



RGC (15a)

	S100A4 Inhibitor	PROTAC
Lead Structures	US10113 (1a)	RGC (15a)
Physiochemical Properties		
Molecular Weight	253.11	681.70
LogD _{7.4}	2.11	1.33
cLogP	1.285	0.197
tPSA	37.38	188.36
Aqueous Solubility (μM)	>100	>100
Heavy Atom Count (HAC)	19	50
Number of Hydrogen Bond Donors and Acceptors (HBD/A)	HBD: 0 HBA: 2	HBD: 3 HBA: 8
<i>In vitro</i> DMPK Assays		
Human microsomal clearance CL _{int} (μL/ min/ mg)	22.2	391
Mouse microsomal clearance CL _{int} (μL/ min/ mg)	156	628
Human hepatocyte clearance CL _{int} (μL/ min/ 10 ⁶)	ND	83.4
Mouse hepatocyte clearance CL _{int} (μL/ min/ 10 ⁶)	ND	293
Human plasma protein binding (% bound)	59	IP
Plasma stability (% remaining at 60 minutes- human)	82.5	38.3
Plasma Stability (% remaining at 60 minutes- mouse)	65.6	5.1

Table 5.3: Comparison of DMPK data between **RGC (15a)** and **US10113 (1a)**.

5.4 US10113 SAR Exploration

5.4.1 Metabolism Issues

The SAR work was conducted on the **US** series as only 3 compounds (**1a-c**) were initially synthesised and assessed (see **Figure 5.1**). Further SAR exploration of the inhibitors was required to assess whether the inhibitory effects could be enhanced *via* the addition of various substituents. Different substituents were attached to the *ortho* and *meta* position of the **US10113** phenyl ring. The *para* position has been left alone as this is the site of tether attachment. As investigation into **US10113** SAR expands, alterations to tether attachments may be explored, allowing various substituents to be attached to the *para* position.

US10113 (1a) possesses a metabolically susceptible double bond due to the presence of the norbornene ring. The double bond of the norbornene leads to the formation of metabolites, lowering its therapeutic effect. The double bond can be oxidised to an epoxide *via* the CYP450 peroxygenase enzymes. Epoxides have shown to behave as alkylating agents as they are highly electrophilic. They can react with the nucleophilic sites of both DNA and proteins, causing significant changes in the genetic information and the functioning of proteins. In addition to this, many epoxides display carcinogenicity.³⁶ A study conducted by National Toxicology Programme (NTP) have established a database of epoxides that display carcinogenicity in mice and rats. This work has shown that many of these chemicals induce tumours at multiple organ sites in rats and mice.³⁷ The destructive nature of epoxides can be resolved by hydrolysis to form a diol *via* epoxide hydrolase enzymes (EHs), or by forming conjugates with glutathione *via* glutathione S-transferase (GST) (see **Figure 5.33**).^{38, 39} SAR analysis to generate a metabolically stable inhibitor will be conducted to lower its problematic high clearance rate (see **Table 5.3**). To stop the rapid metabolism, the double bond can be removed or masked by other functionalities.

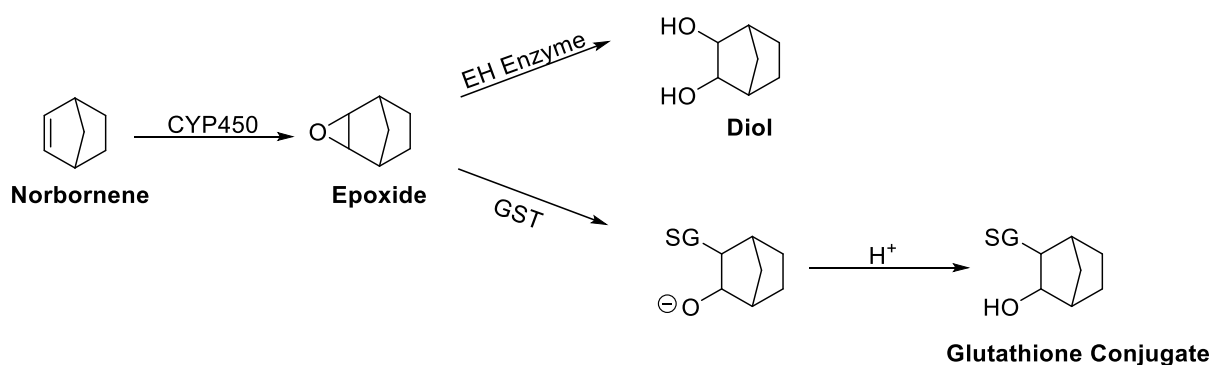


Figure 5.33: Metabolic pathways of the norbornene ring. Adapted from: *Structure and function of the cytochrome P450 peroxygenase enzymes*, A. W. Munro et al; *Purification of Glutathione S-Transferases from Human Liver by Glutathione-Affinity Chromatography*, P. C. Simons et al.^{38 39}

For this section of the project, various **US10113** analogues were synthesised to gain understanding of the SAR and develop a more potent inhibitor, while maintaining target specificity and low toxicity. Investigation into whether the double bond is required for the inhibitor's activity is also essential to develop a more metabolically stable inhibitor (see **Figure 5.34**). The methylene bridge of norbornene highlighted in pink will remain unchanged. The oxygen on **US1041-MOF (1c)** has been associated with higher toxicity compared to the lead inhibitor **US10113 (1a)** (see **Figure 5.2**), as a result, no heteroatoms will be installed at this position.

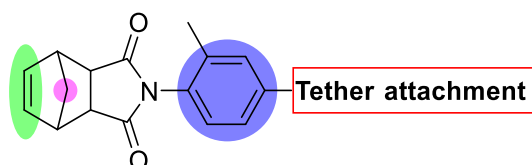
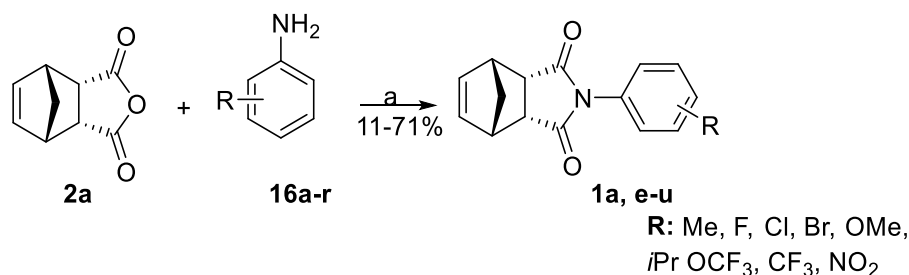


Figure 5.34: SAR exploration of the **US10113** inhibitor.

5.4.2 Discussion

5.4.2.1 Synthesis of unsaturated US analogues

For the initial examination of the relationship between the substituents on the phenyl ring and its inhibition capabilities, a single step reaction was used (see **Scheme 5.3**).



Scheme 5.3: Reagent and Conditions: (a) AcOH, reflux, 16 hr.

A series of US derivatives with varied R groups in the *ortho* and *meta* position were synthesised and tested for their ability to reduce cell migration in Rama 37 cells.

The unsaturated **US** derivatives were synthesised *via* the reaction of *cis*-5-norbornene-*endo*-2,3-dicarboxylic anhydride (**2a**) with a variety of functionalised aniline derivatives (**16a-r**) in the presence of acetic acid.

Initially, the synthesis of **US10113** analogues involved a conventional condenser reflux. It was later discovered that refluxing the reaction in a screw fix sealed tube improved the yield of **1a** from 32% to 54%. The yields for these **US** derivatives vary greatly, the R group is believed to be responsible due to various stereo and electronic effects (see **Table 5.4**). For example, the methoxy group is an inductively electron withdrawing group, which is deactivated towards the *ortho* and *para* position. However, the methoxy group is mesomerically electron donating, making it activating towards the *ortho* group. The yield for **1h** with the methoxy group in the *ortho* position is higher at 42% compared to **1m** in the *meta* position which has a yield of 27%. The mesomeric effect is stronger than the inductive effect, and therefore the *ortho* substituent had a better yield compared to the substituent in the *meta* position. The halides fluorine and chlorine are inductively electron withdrawing, which deactivates the *ortho* and *para* position. **1j** and **1k** which have the halides in the *meta* position have higher yields than **1e** and **1f** which have halides in the deactivated *ortho* position. Sterics also play an important role, as compounds **1g** and **1n**, which have an isopropyl substituent obtain yields of 18%. The steric bulk and unfavourable and repulsive interactions between the isopropyl and the

anhydride functionality cause the yield to be much lower than most of the other **US** derivatives in the series.

Another explanation for the lower yield compared to the procedure by *Li et al* could be due to the work up procedure.⁸ Literature states that the crude reaction mixture should be washed sparingly with cold methanol and acetonitrile. The recovered yields after washing lowered considerably. The work up procedure was later abandoned, and the reaction mixture was reduced and purified *via* column chromatography.

These compounds also display hygroscopic behaviour as a prominent water peak is usually observed in the ¹H NMR spectrum and has previously interfered with purity analysis (CHN). Therefore, azeotropic distillation with toluene was used to remove the excess water.

Compound	R	Yield (%)
US10113 (1a)	<i>Ortho</i> -Me	54
1e	<i>Ortho</i> -F	38
1f	<i>Ortho</i> -Cl	44
1g	<i>Ortho</i> -iPr	18
1h	<i>Ortho</i> -OMe	42
1i	<i>Ortho</i> -CF ₃	42/ 54
1j	<i>Meta</i> -F	40
1k	<i>Meta</i> -Cl	54
1l	<i>Meta</i> -Me	56
1m	<i>Meta</i> -OMe	27
1n	<i>Meta</i> -iPr	18
1o	<i>Meta</i> -CF ₃	41
1p	<i>Meta</i> -NO ₂	38
1q	<i>Meta</i> -Br	65
1r	<i>Meta</i> -OCF ₃	44/ 71
1s	<i>Ortho</i> -OMe, <i>meta</i> -F	39
1t	Di <i>ortho</i> -Cl	34
1u	Di <i>meta</i> -CF ₃	11

Table 5.4: US derivatives synthesised and their corresponding yields.

5.4.2.1.1 Biology

The **US10113** derivatives **1a, e-u**, were tested for their ability to reduce cell migration at 50 μM in Rama 37 cells that overexpress the S100A4 protein, and in wild type Rama 37 cells. Biological testing was carried out by Rudlands group in Life Sciences at the University of Liverpool (see **Table 5.5**).

Compound	R	Relative Migration (%)
DMSO	-	100
US10113 (1a)	<i>Ortho</i> -Me	78
1e	<i>Ortho</i> -F	85
1f	<i>Ortho</i> -Cl	87
1g	<i>Ortho</i> -iPr	97
1h	<i>Ortho</i> -OMe	86
1i	<i>Ortho</i> -CF ₃	97
1j	<i>Meta</i> -F	76
1k	<i>Meta</i> -Cl	54
1l	<i>Meta</i> -Me	59
1m	<i>Meta</i> -OMe	55
1n	<i>Meta</i> -iPr	100
1o	<i>Meta</i> -CF ₃	140
1p	<i>Meta</i> -NO ₂	88
1q	<i>Meta</i> -Br	102
1r	<i>Meta</i> -OCF ₃	68
1s	<i>Ortho</i> -OMe, <i>meta</i> -F	IP
1t	Di <i>ortho</i> -Cl	IP
1u	Di <i>meta</i> -CF ₃	IP

Table 5.5: US 1a, e-u series and their corresponding relative migration in Rama 37 cells, n=3.

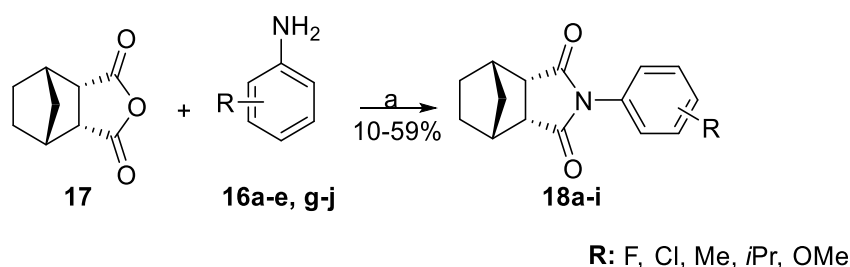
5.4.2.1.2 SAR Analysis

In Rama 37 cells that overexpress S100A4, the *ortho*-substituted compounds **1e-i** displayed worse activity compared to the original **US10113 (1a)**. Compound **1g** with the isopropyl substituent in the *ortho* position exacerbated relative migration the most, and therefore has no therapeutic benefit. These striking differences in relative migration brought about by different substituents confirms the need for further investigation. Compounds **1j-r** which are part of the *meta*-substituted compounds were found to significantly increase the activity against S100A4 induced metastasis with all compounds reducing cell migration more so than the original **US10113 (1a)**. **1k** which has a chlorine in the *meta* position was the most active compound, reducing cell migration to 54%. Compounds **1l** and **1m** which have either methyl or a methoxy group in the *meta* position also displayed highly promising results at 59% and 55%. However, not all the *meta* substituted series had improved activity as **1n** with an isopropyl group displayed the same relative migration value as the control (DMSO). **1o** and **1q** which have either a trifluoromethyl or bromine in the *meta* position drastically increased relative migration, with **1n** (*Meta-iPr*) promoting metastatic spread the worst. The encouragement and discouragement of relative migration due to different substituents may need to be investigated further *via* mechanistic studies to understand differing ligands pharmacological effects on the S100A4 protein.

Some of the best inhibitory effects from the **1a-u** series came from the initial *meta* substituted compounds (**1j-m**). Therefore, *meta* substituted **US** analogues could be incorporated into the PROTAC scaffold, to develop a more potent compound compared to the original **RGC (15a)**.

5.4.2.2 Synthesis of saturated US analogues

The **US 1a-u** series showed promising and interesting results, although the metabolically liable double bond remained in the compounds structure. It was unknown whether the double bond was essential for the **US10113 (1a)** activity. Therefore, early investigations involved removing the double bond and only having a single bond in the compound. The synthesis in **Scheme 5.4** was simple and followed a similar procedure as in **Scheme 5.3**.



Scheme 5.4: Reagent and Conditions: (a) AcOH, reflux, 16 hr.

Like the **1a-u** series, **US** compounds with *ortho* and *meta* substituents were synthesised as a direct comparison between the saturated and unsaturated compounds. The reaction follows the same procedure as **Scheme 5.3**, but instead using the saturated starting material, *cis*-3,6-*endomethylene* tetrahydro phthalic anhydride (**17**). The yield for these reactions were moderate just like in the **1a-u** series (see **Table 5.6**). **18b** (*meta*-Cl) was able to obtain the highest yield of 59% and **18e** (*ortho*-*i*Pr) gave the lowest yields of 10%. For this series, no washing with cold solvent was performed, and instead pure compounds were obtained *via* column chromatography. The yields for this series are comparably lower than the unsaturated series as the reaction was performed on a much lower scale of typically 50 mg compared to 100 mg.

Compound	R	Yield (%)
18a	<i>Ortho</i> -F	25
18b	<i>Ortho</i> -Cl	20
18c	<i>Ortho</i> -Me	38
18d	<i>Ortho</i> -OMe	50
18e	<i>Ortho</i> - <i>i</i> Pr	10
18f	<i>Meta</i> -F	53
18g	<i>Meta</i> -Cl	59
18h	<i>Meta</i> -Me	55
18i	<i>Meta</i> -OMe	50

Table 5.6: **US** derivatives saturated series (**18a-i**) and their corresponding yields.

5.4.2.2.1 Biology

We are currently awaiting the biology results for the saturated compounds from the Rudland group at the Life Sciences department at the University of Liverpool. Once data is obtained, direct comparisons between the saturated and unsaturated compounds can be made to gain

an understanding of whether the double bond is essential for the inhibition of the S100A4 protein.

Compound	R	Relative Migration (%)
18a	<i>Ortho-F</i>	IP
18b	<i>Ortho-Cl</i>	IP
18c	<i>Ortho-Me</i>	IP
18d	<i>Ortho-OMe</i>	IP
18e	<i>Ortho-iPr</i>	IP
18f	<i>Meta-F</i>	IP
18g	<i>Meta-Cl</i>	IP
18h	<i>Meta-Me</i>	IP
18i	<i>Meta-OMe</i>	IP

Table 5.7: The relative migration of the saturated **US** series (**18a-i**) is currently in progress.

The metabolic stability of the saturated **18c** was determined and compared against the original inhibitor **US10113 (1a)** (see **Table 5.8**). The DMPK testing was outsourced to Cyprotex in which they calculated the percentage of compound remaining in a human and mouse model after 45 minutes. Results showed that **18c** had an improved metabolic profile compared to **US10113 (1a)**. In the mouse model both compounds experience a high metabolic turnover, but the amount of **18c** was almost 4 times higher compared to **US10113 (1a)**. In the human model, both compounds fared better, but as expected less **US10113 (1a)** remained after 45 minutes. Only a 15.8% loss of **18c** occurred in human model, which is a significant improvement compared to the original inhibitor. Although, the metabolic turnover is still quite high for **18c**, therefore continued modifications to the structure of the inhibitor may be needed to slow the clearance rate of the compound further.

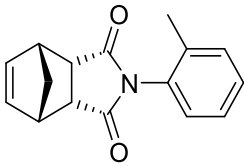
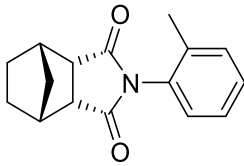
Metabolic Stability % Compound remaining after 45 mins	US10113 (1a) 	18c 
Mouse	8.01 ± 0.16*	31.9 ± 0.93 [†]
Human	60.7 ± 0.79 [‡]	84.2 ± 0.89 [‡]

Table 5.8: Metabolic stability in mouse and human models for compounds **US10113 (1a)** and **18c**. ±

SE CL_{int} *n=3, [†]n=4, [‡]n=5.

5.4.22.2 SAR Analysis

Many of the synthesised **US** derivatives are still awaiting testing to determine their activity in comparison to **US10113 (1a)**. The SAR analysis that has been obtained has provided some important information. So far, derivatives with enhanced activity compared to **US10113 (1a)** have been identified, with trends and explanations theorised.

By far the most active compounds were **1k** (*meta*-Cl), **1l** (*meta*-Me), and **1m** (*meta*-OMe). This suggests that the *meta* compounds occupy a binding site within the S100A4 protein that the *ortho* moieties are unable to, or that the *meta* moieties have a much greater interaction with certain residues in this position. The enhanced inhibitory effects of **1k** (*meta*-Cl) suggests that favourable halogen bonding between the ligand and the protein is occurring. However, this halogen bonding is not as strongly experienced with **1j** (*meta*-F) as relative migration is 76%, only 2% lower than the original **US10113 (1a)** (78%). This could be due to the difference in atomic radius, as chlorine is much larger than fluorine, and therefore able to occupy more space within the binding site of S100A4 and strengthen the interaction between the ligand and protein in the complex. This can be further confirmed when we investigate the activity of **1m** (*meta*-OMe) and **1l** (*meta*-Me) which also possess highly promising inhibitory effects. The methyl and methoxy substituents are large in comparison to the fluorine and therefore may cause favourable interactions, resulting in stronger inhibition of S100A4. Although, as we increase the size of the substituent, such as **1n** (*meta*-iPr), the relative migration increases, suggesting unfavourable interactions despite its bulky nature. There will be a size limit within the binding pocket, therefore bulky substituents such as isopropyl groups may not have favourable interactions with the protein. Additionally, there could be an issue with the

presence of fluorine atoms in the moiety at the *meta* position. **1o** (*meta*-CF₃) and **1r** (*meta*-OCF₃) are only slightly smaller than the methoxy and methyl substituents, and yet have poorer activity compared to **1l** (*meta*-Me) and **1m** (*meta*-OMe). These varied results could be due to differences in halogen bonding. Chlorine is a greater halogen bond acceptor compared to fluorine; therefore, chlorine can interact with a residue capable of displaying halogen bond donor activity greater than the fluorine. This can be further confirmed as the second most active compound **1m** (*meta*-OMe) can act as a hydrogen bond acceptor, suggesting favourable interactions. The activity of **1q** (*meta*-Br) does not follow the expected trend. Bromine is a stronger halogen bond acceptor than chlorine, and yet encourages metastatic spread. Bromine has a larger atomic radius compared to chlorine; therefore, there may be a fine balance between size and halogen bonding potential.

Hydrophobic properties of the substituents should also be considered. The original inhibitor **US10113** (**1a**) has a small hydrophobic methyl group attached to the *ortho* position of the phenyl ring. The two most active **US** derivatives, **1k** (*meta*-Cl) and **1l** (*meta*-Me) also possess the hydrophobic substituents in the *meta* position. Therefore, hydrophobicity may also play a role in activity for the inhibitors. Therefore, a combination of size, halogen/ hydrogen bond acceptance and hydrophobicity may play a role in the activity of inhibitors.

The disubstituted **US** derivatives may give an important insight into the SAR. Different substituents placed in the *ortho* and *meta* positions of the phenyl ring may strengthen the interactions between the inhibitor and the binding pocket of S100A4.

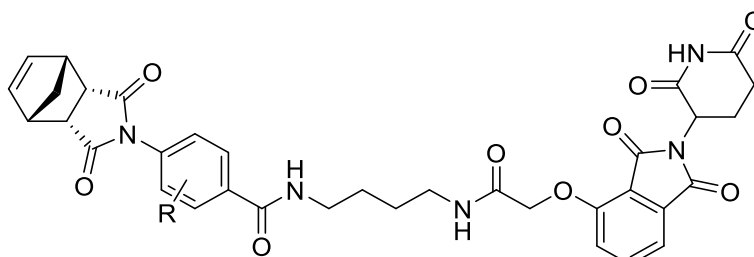
The results for the saturated series of **US** derivatives have yet to be obtained. These results will provide vital information on whether the metabolically liable double bond of the norbornene ring is essential.

5.5 Newly synthesised PROTAC agents

The results so far for the **US** SAR work have indicated that more active inhibitors compared to **US10113** (**1a**) exist and could be incorporated into the PROTAC structure to enhance activity. New PROTAC agents have been synthesised using the same **RGC** (**15a**) methodology following **Schemes 5.1** and **5.2**. Activation of the **1** and **18** series carboxylic acids with pentafluorophenol (**5**) was applied for a cleaner synthesis and higher yields. Purification *via* column chromatography using normal phase silica gel eluting with the modified WIPE127

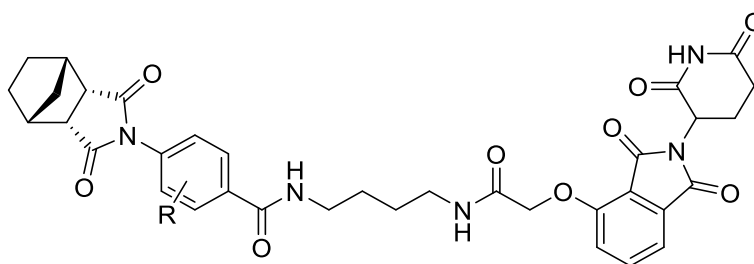
solvent system, followed by trituration, and if required a final cartridge purification yielded PROTAC compounds with purity data ranging between 87.7-99.6%. For the unsaturated series four new PROTAC compounds were synthesised (see **Figure 5.35A**). The **US** compound **15c**, with the chlorine atom in the typically unfavourable *ortho* position was tethered to the PROTAC linker to further examine the relationship between **US** compounds and a full PROTAC compound. Compounds **15d-g** had substituents in the *meta* position which corresponded to the most active US compounds during SAR investigation (see **Table 5.5**). For the saturated series, 4 compounds **19a-d** were synthesised as a direct comparison in activity to the unsaturated series (see **Figure 5.35B**). The results from the **US** saturated series (**18a-f**) have yet to be gathered, however PROTAC agents were synthesised in the event of positive results.

5.35A



R= *ortho*-Cl (**15c**), *meta*-F (**15d**), *meta*-Cl (**15e**), *meta*-OMe (**15f**), *meta*-Me (**15g**)

5.35B



R= *meta*-F (**19a**), *meta*-Cl (**19b**), *meta*-OMe (**19c**), *meta*-Me (**19d**)

Figure 5.35: Structures of newly synthesised PROTAC agents: **5.35A:** Unsaturated series involving *ortho* and *meta* chloro, *meta* fluoro, methoxy, and methyl substituents. **5.35B:** Saturated series involving *meta* fluoro, chloro, methoxy, and methyl substituents.

5.5.1 Biology

Preliminary relative migration data for the newly synthesised PROTACs have shown some promising results and provided further SAR information. **15c** (unsaturated series, *o*-Cl)

which has a substituent in the *ortho* position performed the worse in the relative migration study, as increasing the concentration from 5 to 100 nM did not impede migration. **15c** appeared to encourage migration, reaching higher levels compared to the untreated control. This further confirms that substituents in the *meta* position are favoured over the *ortho* position. Compounds **19c** (saturated series, *m*-OMe) and **19d** (saturated series, *m*-Me) appear to have promising data and outperform the original RGC (**15a**). **19d** is the best performing PROTAC so far, as **19d** at 5 nM was able to prevent migration further than RGC (**15a**) at 100 nM. This verifies that the double bond present on the US warhead is not required for binding to the S100A4 protein. This is fantastic news as the removal of the double bond may address the metabolic stability issues associated with RGC (**15a**). DMPK data between the saturated and unsaturated series (**15a-g** versus **19a-d**) have yet to be gathered but will be compared to see if the metabolic issues are being resolved. Continued analysis to confirm the preliminary migration results are currently being undertaken.

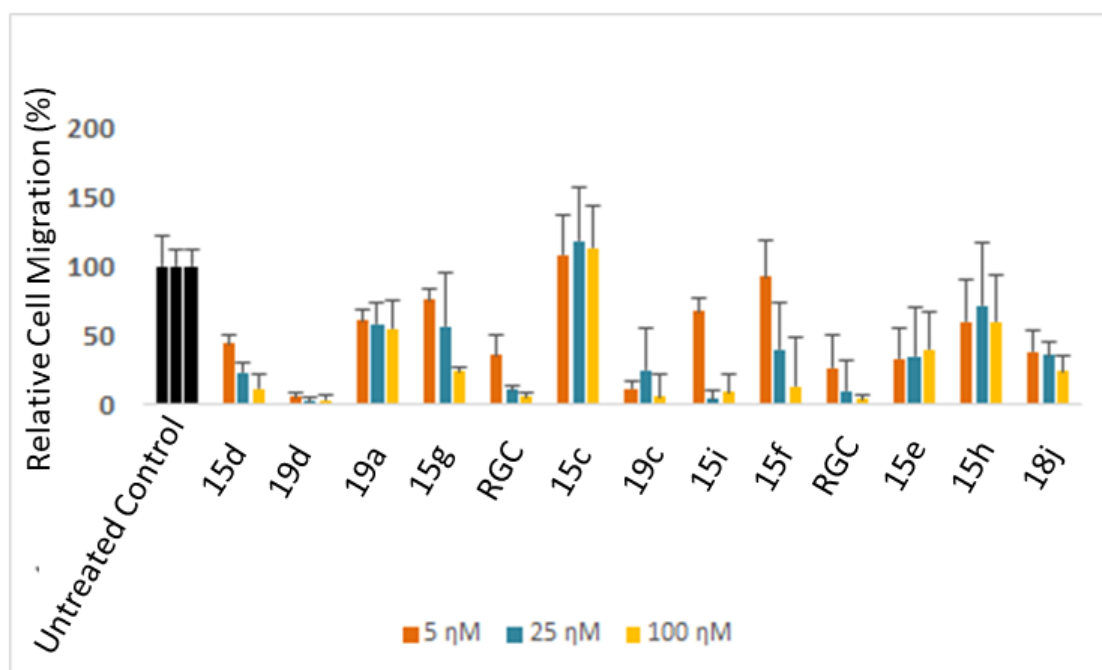


Figure 5.36: Relative cell migration (%) of newly synthesised PROTAC compounds at 5, 25, 100 nM concentrations on Rama37-S100A4 cells. \pm SD n=3

5.6 Conclusion

TNBC is associated with high mortality rates and aggressive chemotherapy due to metastatic spread and the lack of available targeted therapies. The abundant S100A4 protein has been

linked to the cells enhanced metastatic spread. The **US10113 (1a)** inhibitor has demonstrated modest inhibition of the protein on a micromolar scale. Interesting, a new therapeutic technique referred to as PROTACs can degrade the protein rather than rely on inhibition by utilising the proteasomal ubiquitin pathway.

The PROTAC **RGC (15a)** is a combination of the S100A4 inhibitor (**US10113 (1a)**) and the CRBN binding warhead (thalidomide). **RGC (15a)** underwent *in vitro* and *in vivo* testing on three established cell lines analysing toxicity, selectivity, and effect on cancer cell migration and degradation. The data was compared to a series of controls (**US10113 (1a)**), thalidomide and the S100P protein) amongst the various cell lines. **RGC (15a)** has shown to improve the inhibitory effects up to 10,000-fold over the inhibitor alone at the nanomolar range with minimal toxicity at micromolar levels.

Many **US** derivatives have been synthesised for SAR exploration with some data suggesting that a more potent inhibitor could be employed with an improved DMPK profile. Additional PROTAC agents have been synthesised and are awaiting testing to ascertain whether a more promising PROTAC is available.

5.7 Future Work

5.7.1 Inhibitor

Further SAR exploration is needed to gain more information concerning the interactions between the S100A4 protein and the inhibitor. This can also be aided by molecular modelling to observe the positioning of amino acid residues within the binding pocket of the S100A4, and to see how the inhibitor is able to fit within the protein.

Metabolism needs to be improved upon as the double bond has been identified as being metabolically liable. Synthesis of a cyclopropane analogue of **US10113 (1a)** could also address the metabolism issues associated with the norbornene double bond (see **Figure 5.37**).⁴⁰ This along with the saturated series will reveal whether the double bond is necessary for activity. Depending on the results, further manipulation of the **US** structure may be needed to make it more metabolically stable.

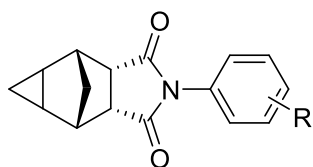


Figure 5.37: Structure of cyclopropane analogue of **US10113 (1a)**.

5.7.2 Tether

Alterations to the amide tether of the PROTAC have been proposed. **RGC (15a)** has a linker length of 10 but adjusting the length will allow us to investigate the flexibility of the molecule. If the linker is too short, the flexibility will be limited. This will hinder the binding of the protein of interest to the E3 ligase, preventing the formation of the ternary complex. If the linker is too long, the protein of interest and the E3 ligase will not be in close proximity, resulting in no degradation.⁴¹ Synthesis of shorter and longer linker lengths will determine the optimum length to ensure maximum degradation.

The amide bond is also susceptible to hydrolysis from several catabolic pathways involving amidases.⁴²⁻⁴³ The amide tether is also susceptible to chemical instability as mentioned previously (see **Section 5.3.1.1**). Fortunately, many tethers are available for PROTAC agents such as ethers and pegylated chains (see **Figure 5.38**).^{44, 45} These linkers frequently incorporate ring systems such as triazole to stabilise linker orientation and add solubility.^{45, 46} Utilising different tethers may improve the metabolic and chemical stability, enhancing the DMPK profile of **RGC (15a)**.

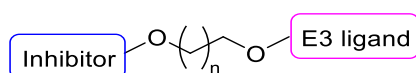
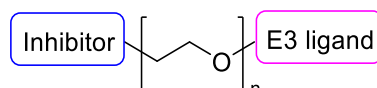
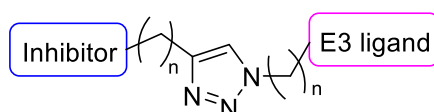
5.38A**5.38B****5.38C**

Figure 5.38: Structures of alternative linkers. **5.38A:** ether linker. **5.38B:** Pegylated linker. **5.38C:** triazole ring incorporated into aliphatic linker.

Alterations to the linker are currently undergoing, with an ether chain incorporating a tetrazole ring being synthesised first (see **Figure 5.39**). The linker is the same length as **RGC (15a)** at 10, as a direct comparison between the two PROTACs. The synthesis of this PROTAC was halted due to the Covid-19 crisis. Modifications to the linker length will also be investigated to vary flexibility once laboratory work resumes.

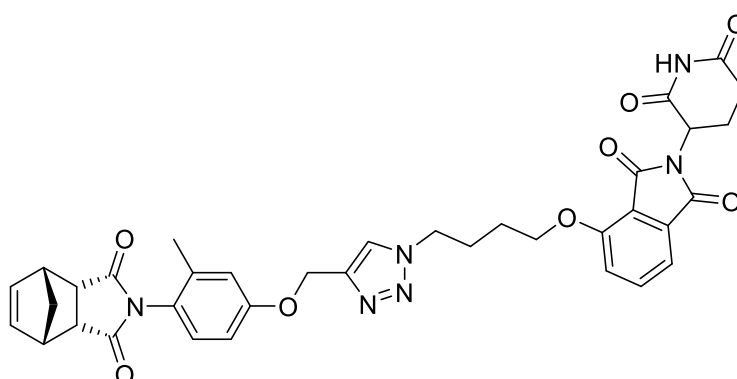


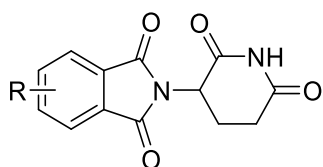
Figure 5.39: Structure of new PROTAC with an ether linker incorporating a triazole ring. Synthesis not completed.

5.7.3 E3 Ligase Ligand

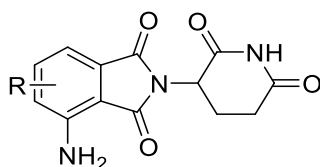
For **RGC (15a)** thalidomide was used to bind to the CRBN portion of the E3 ligase. Thalidomide has been shown to recruit the proteasome successfully, resulting in the degradation of the S100A4 protein. However, other options could be explored. Other phthalimides such as

lenalidomide and pomalidomide can be used to assess whether the other phthalimides improve binding to the E3 ligase (see **Figure 5.40**). VHL ligands are frequently mentioned in literature and utilised in PROTAC agents due to their ability to induce protein degradation.^{47, 48, 49} Possible alterations to **RGCs (15a)** warhead to include a VHL ligand could be made (see **Figure 5.41**), and the destructive behaviour of the two warheads (thalidomide versus VHL) could be compared.

5.40A



5.40B



5.40C

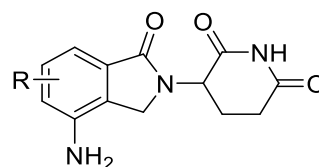


Figure 5.40: Alternative phthalimide structures for targeting E3 ligase. **5.40A:** Thalidomide, **5.40B:** Pomalidomide, **5.40C:** Lenalidomide. R: PROTAC linker and S100A4 inhibitor.

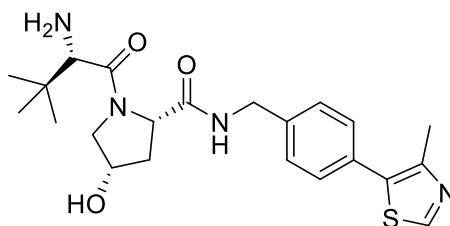


Figure 5.41: Structure of a VHL ligand. VHL ligand can replace the phthalimide structure to induce protein degradation.

5.7.4 Development of Chemical Probes

The selective binding of **RGC (15a)** needs to be investigated as small degradation of the S100P protein can occur at the micromolar range (see **Figure 5.20** and **5.24**). Further studies are needed to determine whether promiscuous binding can degrade a potentially essential protein. Instead of conducting a series of tests involving different proteins, chemical probes can be utilised. This concept has been discussed previously (see **Chapter 1, Section 1.2**). A chemical probe based on the structure of **US10113 (1a)** has been designed. A minimalist linker incorporating the photoreactive diazirine was linked to the inhibitor (see **Figure 5.42**).⁵⁰ The synthesis of the chemical probe has been halted before oxidation to the photoreactive

diazirine. Synthesis was paused to prevent diazirine degradation while waiting for shipment to our collaborator Professor Ciulli at the University of Dundee.

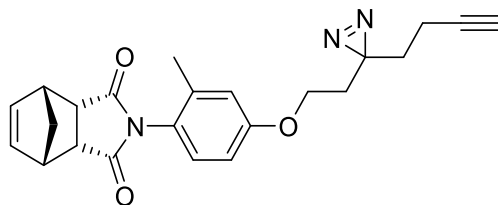


Figure 5.42: Structure of **US10113** photoreactive probe.

5.8 Experimental

5.8.1 General Experimental Details.

For general experimental details please see **Chapter 2, Section 2.14.1**.

5.8.1.1 HPLC

Flow rate 0.5 mL/ min for 15 minutes using MeOH/ Water with compounds dissolved in methanol. UV detector recorded signals at 254 nm. HPLC Method: min, gradient: 80% H₂O hold to 2 min, 80-5% H₂O in MeOH to 12 min, then hold at 80% H₂O to 15 min.

5.8.2 General Procedures

General Procedure A⁵⁸⁴

To a screw fix sealed tube, (3aR,4R,7S,7aS)-hydro-4,7-methanoisobenzofuran-1,3-dione derivative (1.6 equiv.) (**2a-b/18**) and a functionalised aniline (**3a-j/ 16a-r**) (1 equiv.) were suspended in glacial acetic acid (1.2 M) and allowed to stir overnight at reflux. After this time, the solvent was concentrated under reduced pressure. The resultant solid was dissolved in EtOAc and washed with distilled water (x2) and brine (x1). The organic phase was dried over MgSO₄, filtered, and concentrated *in vacuo* to yield a crude solid. Purification *via* column chromatography eluting with 0-50% EtOAc in *n*-hexane afforded the desired product.

General Procedure B⁵⁹³

To a flask, carboxylic acid (**4a-j**) (1 equiv.), 2,3,4,5,6-pentafluorophenol (**5**) (1 equiv.), DMAP (0.1 equiv.), and EDC.HCl (1 equiv.) were suspended in anhydrous THF (0.2 M) and allowed to stir at room temperature overnight under a nitrogen atmosphere. After this time, the crude mixture was filtered, and the filtrate was concentrated under reduced pressure. Purification *via* column chromatography eluting with 0-40% EtOAc in *n*-hexane afforded the desired product.

General Procedure C⁶²⁶

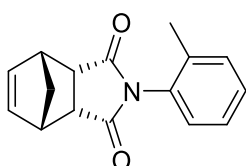
To a flask, perfluorophenyl 4-((3aR,4R,7S,7aS)-1,3-hydro-2H-4,7-methanoisindol-2-yl)benzoate derivative (**6a-j**) (1 equiv.), *N*-(4-aminobutyl)-2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamide (**14**) (1 equiv.), and DMAP (2 equiv.) were suspended in anhydrous DMF (0.2 M) and allowed to stir at 50°C for 16 hours under a nitrogen atmosphere. After this time, the solution was cooled and concentrated under reduced pressure to yield a crude brown oil. Purification *via* column chromatography using normal phase silica gel eluting

with 1% Water: 2% Methanol: 7% EtOAc), followed by trituration with acetone. If necessary, further purification *via* SCX-2 cartridge eluting with 1M NH₃ in MeOH afforded the desired product.

5.8.3 US Derivatives

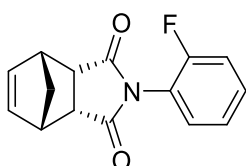
5.8.3.1 Active Unsaturated series⁵⁸⁴

US10113 (1a): (3aR,4S,7R,7aS)-2-(*o*-Tolyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione⁵⁸⁴



Reaction repeated as general procedure A using *o*-toluidine (**16a**) gave the title compound (**US10113 (1a)**) as a white solid (0.64 g, 54%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.37 (m, 2H, Ar-H), 7.30 (m, 1H, Ar-H), 7.13 (dd, 1H, *J* = 7.7, 16.2 Hz, Ar-H), 6.38 (m, 2H, C=C-H), 3.23 (m, 2H, ((CH)₂C=O)C-H), 2.95 (m, 2H, ((CH)₂CH₂)C-H), 2.09 (d, 3H, *J* = 8.0 Hz, CH₃), 1.56 – 1.47 (m, 2H, CH₂). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 177.2, 138.2, 135.7, 132.3, 131.2, 129.5, 129.1, 128.4, 127.8, 48.6, 48.0, 45.5, 44.9, 43.5, 17.6. IR ν_{\max} /cm⁻¹ (solid) 2976 (m), 1774 (s), 1376 (s), 1194 (s), 717 (s). HRMS (CI+) *m/z* calculated for C₁₆H₁₆NO₂: 254.1181. Found [M+H]⁺: 254.1183 (Diff 0.79 ppm). MP: 142–143°C. Purity HPLC 96.6%, R_t = 2.1 min.

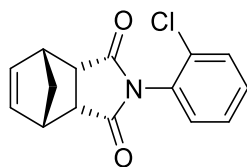
1e: (3aR,4S,7R,7aS)-2-(2-Fluorophenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione⁶²⁷



Reaction repeated as general procedure A using 2-fluoroaniline (**16b**) gave the title compound (**1e**) as a white solid (0.44 g, 38%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.53 (d, 1H, *J* = 7.6 Hz, Ar-H), 7.40 (m, 1H, Ar-H), 7.26 (m, 1H, Ar-H), 7.16 (m, 1H, Ar-H), 6.37 (s, 2H, C=C-H), 3.22 (m, 2H, ((CH)₂C=O)C-H), 2.95 (m, 2H, ((CH)₂CH₂)C-H), 1.61-1.49 (m, 2H, CH₂). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 177.4, 154.3, 136.0, 135.9, 129.8, 129.5, 128.4, 125.0, 116.1, 53.2, 46.2, 46.0, 44.5, 44.3. IR ν_{\max} /cm⁻¹ (solid) 2994 (m), 1704 (s), 1499 (s), 1461 (s), 1215 (s), 689 (s). HRMS

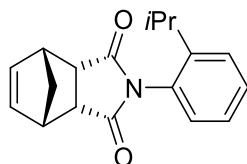
(CI+) m/z calculated for $C_{15}H_{13}^{19}FNO_2$: 258.0930. Found $[M+H]^+$: 258.0922 (Diff -3.10 ppm).
MP: 164–165°C. Purity HPLC 89.74.6%, R_t = 2.1 min.

1f: (3aR,4S,7R,7aS)-2-(2-Chlorophenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione⁶²⁷



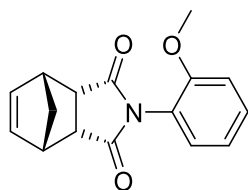
Reaction repeated as general procedure A using 2-chloroaniline (**16c**) gave the title compound (**1f**) as a light pink solid (0.46 g, 44%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.69 – 7.64 (m, 1H, Ar-H), 7.51 (m, 2H, Ar-H), 7.45 – 7.39 (m, 1H, Ar-H), 6.41 – 6.34 (m, 2H, C=C-H), 3.24 (m, 2H, ((CH)₂C=O)C-H), 2.95 (m, 2H, ((CH)₂CH₂)C-H), 1.79 (d, 1H, J = 9.6 Hz, CH₂), 1.57 – 1.47 (m, 1H, CH₂). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 176.5, 138.3, 131.8, 131.7, 131.4, 131.2, 130.8, 130.7, 130.4, 48.7, 47.9, 45.7, 45.1, 43.4. IR ν_{max}/cm^{-1} (solid) 2983 (m), 1705 (s), 1481 (s), 1176 (s), 690 (s), 646 (s). HRMS (CI+) m/z calculated for $C_{15}H_{13}^{35}ClNO_2$: 274.0635. Found $[M+H]^+$: 274.0639 (Diff 1.46 ppm). MP: 159– 160°C. Purity HPLC 95.6%, R_t = 2.1 min.

1g: (3aR,4S,7R,7aS)-2-(2-Isopropylphenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione



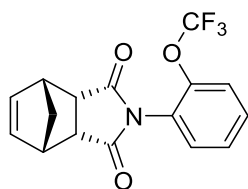
Reaction repeated as general procedure A using 2-isopropylaniline (**16d**) gave the title compound (**1g**) as a white solid (0.19 g, 18%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.51 – 7.41 (m, 2H, Ar-H), 7.33 – 7.26 (m, 1H, Ar-H), 7.08 (m, 1H, Ar-H), 6.38 (m, 2H, J = 3.6, 1.8 Hz, C=C-H), 3.28 – 3.22 (m, 2H, ((CH)₂C=O)C-H), 2.98 – 2.87 (m, 2H, ((CH)₂CH₂)C-H), 2.62 (m, 1H, (CH₃)₂C-H), 1.59 – 1.42 (m, 2H, CH₂), 1.10 (d, 6H, J = 6.8 Hz, CH₃). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 177.3, 144.8, 137.4, 137.2, 135.1, 128.7, 127.1, 126.8, 125.9, 48.6, 46.6, 46.2, 45.3, 44.9, 29.6, 22.3, 22.2. IR ν_{max}/cm^{-1} (solid) 2967 (m), 1699 (s), 1371 (s), 1187 (s), 693 (s). HRMS (CI+) m/z calculated for $C_{18}H_{20}NO_2$: 282.1494. Found $[M+H]^+$ 282.1496 (Diff 0.71 ppm). MP: 145°C. Purity HPLC 98.9%, R_t = 2.1 min.

1h: (3aR,4S,7R,7aS)-2-(2-Methoxyphenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione⁶²⁸



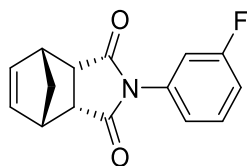
Reaction repeated as general procedure A using 2-methoxyaniline (**16e**) gave the title compound (**1h**) as an off-white solid (0.33 g, 42%). ¹H NMR (400 MHz, CDCl₃) δ 7.36 (t, 1H, *J*= 8.0, 16.0 Hz, Ar-H), 7.05 – 6.98 (m, 1H, Ar-H), 6.97 (m, 1H, Ar-H), 6.96 – 6.89 (m, 1H, Ar-H), 6.28-6.21 (m, 2H, C=C-H), 3.78 (s, 3H, CH₃), 3.48 (m, 2H, ((CH)₂C=O)C-H), 3.47 – 3.42 (m, 2H, ((CH)₂CH₂)C-H), 1.78 (t, 1H, *J*= 8.2, 10.0 Hz, CH₂), 1.60 (t, 1H, *J*= 8.5, 10.2 Hz, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 176.7, 176.6, 161.3, 134.6, 134.45, 130.6, 129.4, 129.0, 120.8, 112.2, 55.8, 55.6, 52.0, 46.6, 46.0, 45.3. HRMS (ES⁺) *m/z* calculated for C₁₆H₁₆NO₃: 270.1130. Found [M+H]⁺ 270.1114 (Diff -5.92 ppm). IR ν_{\max} /cm⁻¹ (solid) 2977 (m), 1690 (s), 1373 (s), 1189 (s), 684 (s). MP: 144-145°C. Purity HPLC 99.0%, R_t= 2.0 min.

1i: (3aR,4S,7R,7aS)-2-(2-(Trifluoromethyl)phenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione



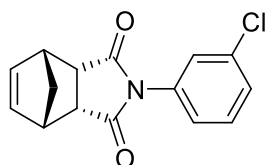
Reaction repeated as general procedure A using 2-(trifluoromethyl)aniline (**16f**) gave the title compound (**1i**) as a white solid (0.51 g, 54%). ¹H NMR (400 MHz, CDCl₃) δ 7.75 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.64-7.51 (m, 2H, *J*= 8.0, 16.0 Hz, Ar-H), 7.11 (d, 1H, *J*= 8.0 Hz, Ar-H), 6.32-6.26 (m, 2H, C=C-H), 3.50-3.45 (m, 4H, ((CH)₂C=O)C-H, ((CH)₂CH₂)C-H), 1.80-1.58 (dd, 2H, *J*= 8.0, 80.0 Hz, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 176.6, 176.4, 134.8, 134.7, 133.3, 133.1, 131.3, 130.7, 129.9, 129.8, 127.4, 52.4, 47.0, 46.1, 45.3, 45.2. ¹⁹F NMR (376 MHz, CDCl₃) δ -60.5. HRMS (ES⁺) *m/z* calculated for C₁₆H₁₃¹⁹F₃NO₂: 308.0898. Found [M+H]⁺ 308.0896 (Diff -0.65 ppm). IR ν_{\max} /cm⁻¹ (solid) 2979 (m), 1702 (s), 1497 (s), 1456 (s), 1172 (s), 690 (s). MP: 159-160°C. Purity HPLC 100.0%, R_t= 2.0 min.

1j: (3a*R*,4*S*,7*R*,7a*S*)-2-(3-Fluorophenyl)-3a,4,7,7a-tetrahydro-1*H*-4,7-methanoisindole-1,3(2*H*)-dione



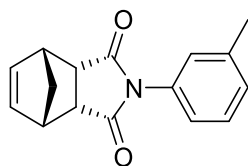
Reaction repeated as general procedure A using 3-fluoroaniline (**16g**) gave the title compound (**1j**) as a white solid (0.40 g, 40%). ¹H NMR (400 MHz, CDCl₃) δ 7.43 (t, 1H, *J* = 7.9, 16.0 Hz, Ar-H), 7.14 – 7.02 (m, 3H, Ar-H), 6.36 (s, 2H, C=C-H), 3.42 (m, 2H, ((CH)₂C=O)C-H), 2.87 (m, 2H, ((CH)₂CH₂)C-H), 1.63 (d, 1H, *J* = 10.2 Hz, CH₂), 1.46 (d, 1H, *J* = 9.9 Hz, CH₂). ¹³C NMR (101 MHz, CDCl₃) 176.6, 176.6, 165.2, 137.8, 136.5, 136.3, 129.8, 124.0, 120.5, 116.3, 52.0, 47.1, 46.7, 45.6, 45.2. IR $\nu_{\max}/\text{cm}^{-1}$ (solid) 2965 (w), 1695 (s), 1380 (m), 1268 (m), 1180 (s), 716 (m). HRMS (CI+) *m/z* calculated for C₁₅H₁₃¹⁹FNO₂: 258.0930. Found [M+H]⁺ 258.0931 (Diff -0.39 ppm). MP: 122–123°C. Purity HPLC 89.7%, *R*_t = 2.1 min.

1k: (3a*R*,4*S*,7*R*,7a*S*)-2-(3-Chlorophenyl)-3a,4,7,7a-tetrahydro-1*H*-4,7-methanoisindole-1,3(2*H*)-dione⁶²⁹



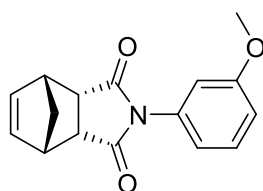
Reaction repeated as general procedure A using 3-chloroaniline (**16h**) gave the title compound (**1k**) as a white solid (0.56 g, 54%). ¹H NMR (400 MHz, CDCl₃) δ 7.43 – 7.35 (m, 2H, Ar-H), 7.32 (m, 1H, Ar-H), 7.20 (m, 1H, Ar-H), 6.36 (s, 2H, C=C-H), 3.41 (m, 2H, ((CH)₂C=O)C-H), 2.86 (m, 2H, ((CH)₂CH₂)C-H), 1.63 (d, 1H, *J* = 10.1 Hz, CH₂), 1.46 (d, 1H, *J* = 9.9 Hz, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 176.6, 176.5, 138.1, 138.0, 134.1, 130.1, 128.9, 126.7, 124.6, 123.8, 48.0, 47.9, 45.9, 43.0, 42.7. IR $\nu_{\max}/\text{cm}^{-1}$ (solid) 2983 (w), 1702 (s), 1430 (m), 1193 (s), 726 (m), 698 (s). HRMS (CI+) *m/z* calculated for C₁₅H₁₃³⁵ClNO₂: 274.0635. Found [M+H]⁺ 274.0634 (Diff -0.36 ppm). MP: 158°C. Purity HPLC 91.6%, *R*_t = 2.0 min.

1l: (3aR,4S,7R,7aS)-2-(*m*-Tolyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione⁶³⁰



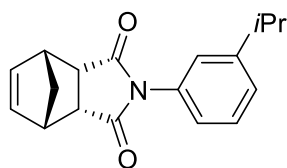
Reaction repeated as general procedure A using *m*-toluidine (**16i**) gave the title compound (**1l**) as a white solid (0.54 g, 56%). ¹H NMR (400 MHz, CDCl₃) δ 7.35 (t, 1H, *J* = 7.8 Hz, 15.9, Ar-H), 7.20 (d, 1H, *J* = 7.7 Hz, Ar-H), 7.04 (m, 2H, Ar-H), 6.35 (m, 2H, C=C-H), 3.41 (m, 2H, ((CH)₂C=O)C-H), 2.85 (m, 2H, ((CH)₂CH₂)C-H), 2.38 (s, 3H, CH₃), 1.63-1.51 (m, 2H, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 177.2, 138.0, 136.4, 136.3, 132.1, 129.6, 129.0, 127.0, 123.5, 47.9, 45.8, 45.7, 43.2, 43.0, 21.3, IR ν_{\max} /cm⁻¹ (solid) 2979 (w), 1701 (s), 1382 (s), 1197 (s), 721 (s). HRMS (Cl⁺) *m/z* calculated for C₁₆H₁₆NO₂: 254.1181. Found [M+H]⁺ 254.1174 (Diff -2.75 ppm). MP: 125°C. Purity HPLC 80.2%, R_t = 2.1 min.

1m: (3aR,4S,7R,7aS)-2-(3-Methoxyphenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione



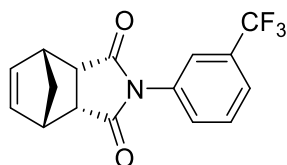
Reaction repeated as general procedure A using 3-methoxyaniline (**16j**) gave the title compound (**1m**) as a beige solid (0.14 g, 27%). ¹H NMR (400 MHz, CDCl₃) δ 7.37 (t, 1H, *J* = 8.3, 16.2 Hz, Ar-H), 6.94 (dd, 1H, *J* = 8.4, 2.6 Hz, Ar-H), 6.84 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.79 (m, 1H, Ar-H), 6.35 (m, 2H, C=C-H), 3.81 (s, 3H, CH₃), 3.41 (m, 2H, ((CH)₂C=O)C-H), 2.86 (m, 2H, ((CH)₂CH₂)C-H), 1.64 (d, 1H, *J* = 8.3 Hz, CH₂), 1.49 (d, 1H, *J* = 9.9 Hz, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 177.0, 162.5, 138.0, 135.2, 135.1, 129.9, 118.6, 114.6, 112.3, 55.5, 47.9, 45.9, 45.7, 43.1, 43.0. IR ν_{\max} /cm⁻¹ (solid) 2975 (w), 1697 (s), 1379 (s), 1187 (s), 1041 (s), 688 (s). HRMS (Cl⁺) *m/z* calculated for C₁₆H₁₆NO₃: 270.1130. Found [M+H]⁺ 270.1136 (Diff 2.22 ppm). MP: 126–127°C. Purity HPLC 84.3%, R_t = 2.0 min.

1n: (3aR,4S,7R,7aS)-2-(3-Isopropylphenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione



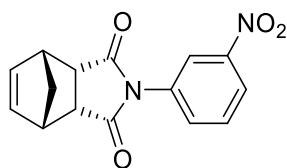
Reaction repeated as general procedure A using 3-isopropylaniline (**16k**) gave the title compound (**1n**) as a white solid (0.14 g, 18%). ¹H NMR (400 MHz, CDCl₃) δ 7.34 (t, 1H, *J* = 7.8, 16.0 Hz, Ar-H), 7.23 (m, 1H, Ar-H), 6.95 (m, 1H, Ar-H), 6.93 (m, 1H, Ar-H), 6.27 (m, 2H, C=C-H), 3.51 (m, 2H, ((CH)₂C=O)C-H), 3.42 (m, 2H, ((CH)₂CH₂)C-H), 1.79 (m, 1H, (CH₂)₂C-H), 1.61-1.56 (m, 2H, CH₂), 1.25 (s, 3H, CH₃), 1.23 (s, 3H, CH₃). ¹³C NMR (101 MHz, CDCl₃) δ 176.9, 150.0, 135.5, 134.6, 131.8, 129.0, 126.7, 124.9, 124.0, 52.8, 52.3, 47.1, 46.1, 45.8, 33.9, 23.8. IR ν_{\max} /cm⁻¹ (solid) 2957 (m), 1699 (s), 1447 (m), 1177 (s), 695 (s). HRMS (ES+) *m/z* calculated for C₁₈H₂₀NO₂: 282.1494. Found [M+H]⁺ 282.1396 (Diff 0.71 ppm). MP: 122–123°C. Purity HPLC 89.8%, R_t = 2.0 min.

1o: (3aR,4S,7R,7aS)-2-(3-(Trifluoromethyl)phenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione⁶³¹



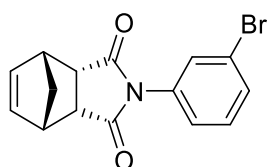
Reaction repeated as general procedure A using 3-(trifluoromethyl)aniline (**16l**) gave the title compound (**1o**) as an off-white solid (0.39 g, 41%). ¹H NMR (400 MHz, CDCl₃) δ 7.62 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.54 (t, 1H, *J* = 8.0, 16.0 Hz, Ar-H), 7.36 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.28 (m, 2H, C=C-H), 3.54-3.52 (m, 2H, ((CH)₂C=O)C-H), 3.47-3.46 (m, 2H, ((CH)₂CH₂)C-H), 1.81 (d, 1H, *J* = 8.0 Hz, CH₂), 1.63 (d, 1H, *J* = 8.0 Hz, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 176.5, 176.3, 134.9, 134.7, 136.0, 132.8, 130.2, 129.6, 125.9, 125.3, 119.6, 52.4, 45.9, 45.7, 45.6, 45.3. ¹⁹F NMR (376 MHz, CDCl₃) δ -62.7. HRMS (ES+) *m/z* calculated for C₁₆H₁₃¹⁹F₃NO₂: 308.0898. Found [M+H]⁺ 308.0894 (Diff -1.30 ppm). IR ν_{\max} /cm⁻¹ (solid) 2971 (m), 1703 (s), 1386 (s), 1190 (s), 1177 (s), 695 (s). MP: 126-128°C Purity HPLC 95.9%, R_t = 2.1 min.

1p: (3aR,4S,7R,7aS)-2-(3-Nitrophenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione⁶³¹



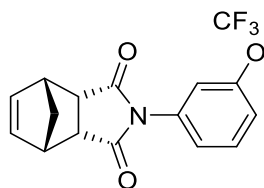
Reaction repeated as general procedure A using 3-nitroaniline (**16m**) gave the title compound (**1p**) as a light-yellow solid (0.33 g, 38%). ¹H NMR (400 MHz, CDCl₃) δ 8.23 (d, 1H, *J* = 8.0 Hz, Ar-H), 8.09-8.08 (m, 1H, Ar-H), 7.62 (t, 1H, *J* = 8.0, 16.0 Hz, Ar-H), 7.55 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.29 (m, 2H, C=C-H), 3.55-3.54 (m, 2H, ((CH)₂C=O)C-H), 3.49-3.48 (m, 2H, ((CH)₂CH₂)C-H), 1.83 (d, 1H, *J* = 8.0 Hz, CH₂), 1.65 (d, 1H, *J* = 8.0 Hz, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 176.1, 148.6, 134.8, 132.8, 132.5, 129.8, 123.2, 121.9, 52.4, 45.9, 45.7. HRMS (ES+) *m/z* calculated for C₁₅H₁₃N₂O₄: 285.0875. Found [M+H]⁺ 285.0876 (Diff 0.35 ppm). IR ν_{max}/cm⁻¹ (solid) 2961 (w), 1693 (s), 1532 (s), 1380 (s), 1253 (s), 703 (s). MP: 148-149°C. Purity HPLC 89.3%, R_t = 2.1 min.

1q: (3aR,4S,7R,7aS)-2-(3-Bromophenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione



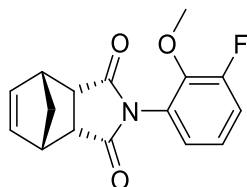
Reaction repeated as general procedure A using 3-bromoaniline (**16n**) gave the title compound (**1q**) as a beige solid (0.63 g, 65%). ¹H NMR (400 MHz, CDCl₃) δ 7.50(d, 1H, *J* = 8.0 Hz, Ar-H), 7.32-7.28 (m, 2H, Ar-H), 7.11 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.27 (m, 2H, C=C-H), 3.52-3.50 (m, 2H, ((CH)₂C=O)C-H), 3.44-3.43 (m, 2H, ((CH)₂CH₂)C-H), 1.80 (d, 1H, *J* = 8.0 Hz, CH₂), 1.62 (d, 1H, *J* = 8.0 Hz, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 176.4, 136.2, 134.7, 133.0, 131.7, 130.3, 129.7, 125.4, 122.4, 52.3, 45.8, 45.6. HRMS (ES+) *m/z* calculated for C₁₅H₁₃⁸¹BrNO₂: 318.0129. Found [M+H]⁺ 318.0127 (Diff -0.63 ppm). IR ν_{max}/cm⁻¹ (solid) 2986 (w), 1709 (s), 1422 (s), 1197 (s), 699 (s) 673 (s). MP: 141-143°C. Purity HPLC 99.7%, R_t = 2.0 min.

1r: (3aR,4S,7R,7aS)-2-(3-(Trifluoromethoxy)phenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione



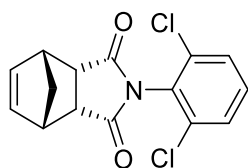
Reaction repeated as general procedure A using 3-(trifluoromethoxy)aniline (**16o**) gave the title compound (**1r**) as a white solid (0.65 g, 71%). ¹H NMR (400 MHz, CDCl₃) δ 7.42 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.02 (t, 1H, *J* = 8.0, 16.0 Hz, Ar-H), 6.95 (m, 1H, Ar-H), 6.67 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.31 (m, 2H, C=C-H), 3.56-3.54 (m, 2H, ((CH)₂C=O)C-H), 3.49-3.48 (m, 2H, ((CH)₂CH₂)C-H), 1.84 (d, 1H, *J* = 8.0 Hz, CH₂), 1.66 (d, 1H, *J* = 8.0 Hz, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 176.3, 159.8, 135.2, 134.7, 132.3, 129.6, 125.3, 123.6, 106.1, 52.4, 45.9, 45.6. ¹⁹F NMR (376 MHz, CDCl₃) δ -68.4. HRMS (ES+) *m/z* calculated for C₁₆H₁₃¹⁹F₃NO₃: 324.0847. Found [M+H]⁺ 324.0849 (Diff 0.62 ppm). IR ν_{\max} /cm⁻¹ (solid) 2969 (w), 1694 (s), 1388 (s), 1274 (m), 1176 (s), 715 (m). MP: 123-124°C. Purity HPLC 99.1%, R_t = 2.0 min.

1s: (3aR,4S,7R,7aS)-2-(3-Fluoro-2-methoxyphenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione



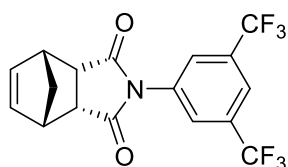
Reaction repeated as general procedure A using 3-fluoro-2-methoxyaniline (**16p**) gave the title compound (**1s**) as a plight pink solid (0.16 g, 39%). ¹H NMR (400 MHz, CDCl₃) δ 7.21-7.19 (m, 2H, Ar-H), 7.06 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.22 (m, 2H, C=C-H), 3.84 (s, 3H, CH₃), 3.52 (m, 2H, ((CH)₂C=O)C-H), 3.42 (m, 2H, ((CH)₂CH₂)C-H), 1.52 (d, 1H, *J* = 8.0 Hz, CH₂), 1.43 (d, 1H, *J* = 8.0 Hz, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 176.4, 155.6, 139.4, 137.1, 137.0, 125.1, 123.2, 120.9, 113.3, 55.7, 52.6, 46.1, 46.0, 45.6, 45.3. HRMS (ES+) *m/z* calculated for C₁₆H₁₅¹⁹FNO₃: 288.1036. Found [M+H]⁺ 288.1032 (Diff -1.39 ppm). IR ν_{\max} /cm⁻¹ (solid) 2971 (m), 1692 (s), 1375 (s), 1270 (m), 1184 (s), 711 (s). MP: 136-137°C. Purity HPLC 90.2%, R_t = 2.0 min.

1t: (3aR,4S,7R,7aS)-2-(2,6-Dichlorophenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisoindole-1,3(2H)-dione



Reaction repeated as general procedure A using 2,6-dichloroaniline (**16q**) gave the title compound (**1t**) as a white solid (78.8 mg, 34%). ¹H NMR (400 MHz, CDCl₃) δ 7.66 (d, 2H, *J*= 8.0 Hz, Ar-H), 7.22 (m, 1H, Ar-H), 6.23 (m, 2H, C=C-H), 3.56-3.4 (m, 2H, ((CH)₂C=O)C-H), 3.41-3.40 (m, 2H, ((CH)₂CH₂)C-H), 1.54 (d, 1H, *J*= 8.0 Hz, CH₂), 1.42 (d, 1H, *J*= 8.0 Hz, CH₂). ¹³C NMR (101 MHz, CDCl₃) δ 175.9, 140.3, 134.0, 132.5, 125.4, 126.6, 52.6, 45.1, 44.8. HRMS (ES+) *m/z* calculated for C₁₅H₁₂³⁵Cl₂NO₂: 308.0245. Found [M+H]⁺ 308.0246 (Diff 0.32 ppm). IR ν_{max}/cm⁻¹ (solid) 2988 (w), 1710 (s), 1426 (s), 1195 (s), 733 (m), 670 (s). MP: 139-140°C. Purity HPLC 100.0%, R_t= 2.1 min.

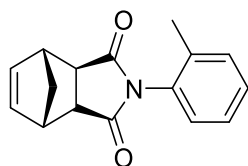
1u: (3aR,4S,7R,7aS)-2-(3,5-bis(Trifluoromethyl)phenyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisoindole-1,3(2H)-dione⁶³²



Reaction repeated as general procedure A using 3,5-bis(trifluoromethyl)aniline (**16r**) gave the title compound (**1u**) as a white solid (11.1 mg, 11%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.01-7.99 (m, 1H, Ar-H), 7.83 (s br, 2H, Ar-H), 6.31 (m, 2H, C=C-H), 3.61-3.60 (m, 2H, ((CH)₂C=O)C-H), 3.48-3.46 (m, 2H, ((CH)₂CH₂)C-H), 1.80-1.70 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 171.0, 140.6, 131.6 (q), 134.4, 124.6, 122.0, 188.9, 52.0, 45.0, 45.3. ¹⁹F NMR (376 MHz, CD₃OD-*d*4) δ -64.6. HRMS (ES-) *m/z* calculated for C₁₇H₁₀¹⁹F₆NO₂: 374.0616. Found [M-H]⁺ 374.0622 (Diff 1.60 ppm). IR ν_{max}/cm⁻¹ (solid) 2959 (w), 1692 (s), 1389 (s), 1275 (m), 1181 (s), 720 (s). MP: 149°C. Purity HPLC 93.0%, R_t= 2.1 min.

5.8.3.2 Inactive Unsaturated series⁵⁸⁴

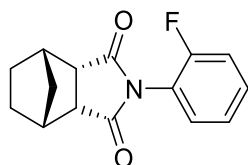
1d, (iUS10113): (3aR,4R,7S,7aR)-2-(o-Tolyl)-3a,4,7,7a-tetrahydro-1H-4,7-methanoisindole-1,3(2H)-dione⁵⁸⁴



Reaction repeated as general procedure A using *o*-toluidine (**16a**) gave the title compound (**1d**, **iUS10113**) as a white solid (0.88 g, 92%). ¹H NMR (400 MHz, CDCl₃) δ 7.45-7.41 (m, 2H, Ar-H), 7.13 (d, 2H, *J* = 8.0 Hz, Ar-H), 6.27 (m, 2H, C=C-H), 3.51 (m, 2H, ((CH)₂C=O)C-H), 3.44-3.43 (m, 2H, ((CH)₂CH₂)C-H), 1.80-1.60 (m, 2H, CH₂), 1.56 (s, 3H, CH₃). ¹³C NMR (101 MHz, CDCl₃) δ 176.9, 134.6, 133.7, 131.9, 129.1, 128.6, 126.7, 124.8, 52.2, 45.8, 45.5, 17.8. HRMS (CI+) *m/z* calculated for C₁₆H₁₆NO₂: 254.1181. Found [M+H]⁺ 254.1186 (Diff 1.97 ppm). IR *v*_{max}/cm⁻¹ (solid) 2977 (m), 1777 (s), 1378 (s), 1193 (s), 721 (s). MP: 143-144°C. Purity HPLC 90.9%, R_t = 2.1 min.

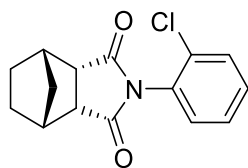
5.8.3.3 Active Saturated series⁵⁸⁴

18a: (3aR,4R,7S,7aS)-2-(2-Fluorophenyl)hexahydro-1H-4,7-methanoisindole-1,3(2H)-dione



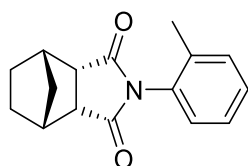
Reaction repeated as general procedure A using 2-fluoroaniline (**16b**) gave the title compound (**18a**) as a white solid (32 mg, 25%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.53 (m, 1H, Ar-H), 7.41 (m, 1H, Ar-H), 7.38 – 7.26 (m, 2H, Ar-H), 3.41-3.30 (m, 2H, ((CH)₂C=O)C-H), 2.68 (m, 2H, CH₂), 1.57 (m, 4H, (CH₂)₂), 1.34 (m, 2H, ((CH)₂CH₂)C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 177.2, 163.1, 131.6, 130.9, 125.5, 116.8, 116.6, 49.4, 41.9, 39.2, 24.7. IR *v*_{max}/cm⁻¹ (solid) 2980 (m), 1702 (s), 1503 (m), 1383 (m), 1215 (s), 649 (s). HRMS (CI+) *m/z* calculated for C₁₅H₁₅¹⁹FNO₂: 260.1087. Found [M+H]⁺ 260.1085 (Diff -0.77 ppm). MP: 146°C. Purity HPLC 90.9%, R_t = 2.0 min.

18b: (3aR,4R,7S,7aS)-2-(2-Chlorophenyl)hexahydro-1H-4,7-methanoisindole-1,3(2H)-dione



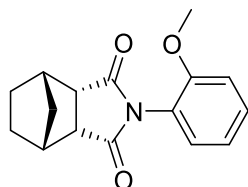
Reaction repeated as general procedure A using 2-chloroaniline (**16c**) gave the title compound (**18b**) as a white solid (27 mg, 20%). ¹H NMR (400 MHz, CDCl₃) δ 7.54 (m, 1H, Ar-H), 7.45 – 7.32 (m, 2H, Ar-H), 7.15 (m, 1H, Ar-H), 3.33-3.28 (d, 2H, *J* = 8.0 Hz, ((CH)₂C=O)C-H), 2.85 (m, 2H, CH₂), 1.81 – 1.63 (m, 5H, CH₂, ((CH)₂CH₂)C-H), 1.53 – 1.46 (m, 1H, ((CH)₂CH₂)C-H). ¹³C NMR (101 MHz, CDCl₃) δ 177.2, 130.7, 130.6, 130.5, 130.3, 129.7, 49.6, 49.1, 42.7, 42.3, 39.7, 25.0, 24.9. IR ν_{\max} /cm⁻¹ (solid) 3081 (m), 1702 (s), 1377 (m), 1216 (s), 693 (s), 643 (s). HRMS (CI+) *m/z* calculated for C₁₅H₁₅³⁵ClNO₂: 276.0791. Found [M+H]⁺ 276.0798 (Diff 2.54 ppm). MP: 148-149°C. Purity HPLC 93.5%, R_t = 2.0 min.

18c: (3aR,4R,7S,7aS)-2-(*o*-Tolyl)hexahydro-1H-4,7-methanoisindole-1,3(2H)-dione



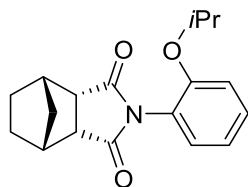
Reaction repeated as general procedure A using *o*-toluidine (**16a**) gave the title compound (**18c**) as a white solid (46 mg, 38%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.42 (m, 2H, Ar-H), 7.39 – 7.33 (m, 1H, Ar-H), 7.23 – 7.03 (dd, 1H, *J* = 8.0, 16.0 Hz, Ar-H), 3.42-3.38 (m, 2H, ((CH)₂C=O)C-H), 2.75-2.21 (m, 2H, CH₂), 2.13 (s, 3H, CH₃), 1.78 – 1.62 (m, 4H, (CH₂)₂), 1.42 (t, 2H, *J* = 8.0, 16.0 Hz, ((CH)₂CH₂)C-H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 177.8, 136.4, 135.7, 131.3, 131.1, 129.6, 128.6, 49.5, 48.9, 42.2, 42.1, 39.1, 25.1, 24.9, 17.6. IR ν_{\max} /cm⁻¹ (solid) 2977 (m), 1698 (s), 1373 (s), 1177 (s), 595(s). HRMS (CI+) *m/z* calculated for C₁₆H₁₈NO₂: 256.1337. Found [M+H]⁺ 256.1336 (Diff -0.39 ppm). MP: 133-135°C. Purity HPLC 90.7%, R_t = 2.1 min.

18d: (3aR,4R,7S,7aS)-2-(2-Methoxyphenyl)hexahydro-1H-4,7-methanoisindole-1,3(2H)-dione



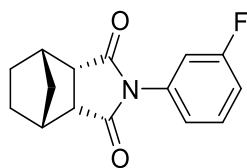
Reaction repeated as general procedure A using 2-methoxyaniline (**16e**) gave the title compound (**18d**) as a white solid (80 mg, 50%). ¹H NMR (400 MHz, CDCl₃) δ 7.38 (t, 1H, *J* = 8.0, 16.0 Hz, Ar-H), 6.96-6.93 (dd, 1H, *J* = 4.0, 8.0 Hz, Ar-H), 6.84-6.82 (dd, 1H, *J* = 4.0 Hz, Ar-H), 6.78 (t, 1H, *J* = 4.0 Hz, Ar-H), 3.82 (s, 3H, CH₃), 3.24 (m, 2H, ((CH)₂C=O)C-H), 2.87 (m, 2H, CH₂), 1.73-1.65 (m, 4H, (CH₂)₂), 1.51-1.45 (m, 2H, ((CH)₂CH₂)C-H). ¹³C NMR (101 MHz, CDCl₃) δ 177.5, 160.1, 132.9, 130.0, 119.0, 114.4, 112.7, 55.4, 48.8, 42.2, 39.8, 24.9. HRMS (CI⁺) *m/z* calculated for C₁₆H₁₈NO₃: 272.1286. Found [M+H]⁺ 272.1281 (Diff -1.84 ppm). IR ν_{max}/cm⁻¹ (solid) 2956 (m), 1704 (s), 1374 (s), 1180 (s), 766 (s). MP: 150-151°C. Purity HPLC 91.2%, R_t = 2.0 min.

18e: (3aR,4R,7S,7aS)-2-(2-Isopropylphenyl)hexahydro-1H-4,7-methanoisindole-1,3(2H)-dione



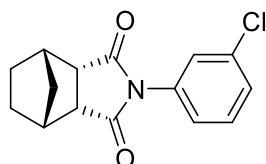
Reaction repeated as general procedure A using 2-isopropylaniline (**16d**) gave the title compound (**18e**) as a white solid (14 mg, 10%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.56 – 7.47 (m, 2H, Ar-H), 7.35 (m, 1H, Ar-H), 7.08 (dd, 1H, *J* = 1.3, 7.8 Hz, Ar-H), 3.44 (s, 2H, ((CH)₂C=O)C-H), 2.75 (m, 2H, CH₂), 1.79 – 1.63 (m, 4H, (CH₂)₂), 1.40 (m, 2H, ((CH)₂CH₂)C-H), 1.18 (dd, 6H (CH₃)₂). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 178.4, 130.1, 129.9, 129.5, 128.7, 127.2, 127.1, 49.7, 48.9, 42.3, 42.1, 39.1, 28.6, 28.1, 25.2, 24.9, 24.0. IR ν_{max}/cm⁻¹ (solid) 2933 (m), 1701 (s), 1370 (s), 1183 (s), 780 (s). HRMS (CI⁺) *m/z* calculated for C₁₈H₂₂NO₂: 284.1650. Found [M+H]⁺ 284.1648 (Diff -0.70 ppm). MP: 133.8 – 134.0°C. MP: 138-140°C. Purity HPLC 89.7%, R_t = 2.0 min.

18f: (3aR,4R,7S,7aS)-2-(3-Fluorophenyl)hexahydro-1H-4,7-methanoisindole-1,3(2H)-dione



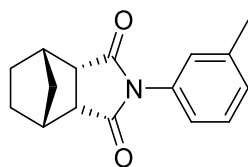
Reaction repeated as general procedure A using 3-fluoroaniline (**16g**) gave the title compound (**18f**) as a white solid (80 mg, 53%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.55-7.50 (q, 1H, *J*= 8.0, 12.0 Hz, Ar-H), 7.21 (t, 1H, *J*= 8.0, 16.0 Hz, Ar-H), 7.13 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.11-7.07 (dt, 1H, *J*= 4.0, 12.0 Hz, Ar-H), 3.50 (m, 1H, ((CH)₂C=O)C-H), 3.15 (m, 1H, ((CH)₂C=O)C-H), 2.80 (m, 2H, CH₂), 1.81-1.69 (m, 4H, (CH₂)₂), 1.44-1.42 (m, 2H, ((CH)₂CH₂)C-H). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 177.1, 170.3, 130.4, 130.3, 122.3, 115.8, 114.4, 48.8, 42.2, 39.8, 24.8. HRMS (ES-) *m/z* calculated for C₁₅H₁₃¹⁹FNO₂: 258.0931. Found [M-H]⁺ 258.0934 (Diff 1.16 ppm). IR $\nu_{\max}/\text{cm}^{-1}$ (solid) 2921 (w), 1701 (s), 1462 (s), 1367 (m), 1188 (s), 644 (s). MP: 149-150°C. Purity HPLC 90.3%, R_t= 2.1 min.

18g: (3aR,4R,7S,7aS)-2-(3-Chlorophenyl)hexahydro-1H-4,7-methanoisindole-1,3(2H)-dione⁶³¹



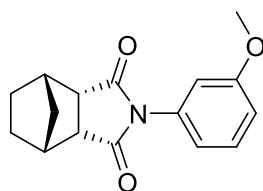
Reaction repeated as general procedure A using 3-chloroaniline (**16h**) gave the title compound (**18g**) as a white solid (0.1 mg, 59%). ¹H NMR (400 MHz, CDCl₃) δ 7.43-7.38 (m, 2H, Ar-H), 7.29 (t, 1H, *J*= 4.0 Hz, Ar-H), 7.19-7.16 (dt, 1H, *J*= 4.0, 8.0 Hz, Ar-H), 3.25 (m, 2H, ((CH)₂C=O)C-H), 2.88 (m, 2H, CH₂), 1.71-1.66 (m, 4H, (CH₂)₂), 1.47-1.43 (m, 2H, ((CH)₂CH₂)C-H). ¹³C NMR (101 MHz, CDCl₃) δ 177.1, 134.7, 132.8, 130.1, 128.9, 126.9, 124.9, 48.8, 42.2, 39.8, 24.9. HRMS (CI+) *m/z* calculated for C₁₅H₁₅³⁵ClNO₂: 276.0791. Found [M+H]⁺ 276.0790 (Diff -0.36 ppm). IR $\nu_{\max}/\text{cm}^{-1}$ (solid) 2969 (w), 1698 (s), 1355 (s), 1208 (m), 699 (s), 641 (s). MP: 148°C. Purity HPLC 100.0%, R_t= 2.1 min.

18h: (3aR,4R,7S,7aS)-2-(*m*-Tolyl)hexahydro-1*H*-4,7-methanoisindole-1,3(2*H*)-dione



Reaction repeated as general procedure A using *m*-toluidine (**16i**) gave the title compound (**18h**) as a white solid (0.12 g, 55%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.36 (t, 1H, *J*= 8.0, 16.0 Hz, Ar-H), 7.24 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.03-7.0 (m, 2H, Ar-H), 3.48-3.47 (m, 1H, ((CH)₂C=O)C-H), 3.14-3.12 (m, 1H, ((CH)₂C=O)C-H), 2.77 (m, 2H, CH₂), 1.78-1.67 (m, 4H, (CH₂)₂), 1.46-1.39 (m, 2H, ((CH)₂CH₂)C-H). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 177.6, 139.3, 131.8, 129.6, 129.1, 127.3, 123.8, 48.8, 42.2, 39.7, 24.9, 21.4. HRMS (ES+) *m/z* calculated for C₁₆H₁₈NO₂: 256.1337. Found [M+H]⁺ 256.1334 (Diff -1.17 ppm). IR ν_{max}/cm⁻¹ (solid) 2921 (w), 1701 (s), 1462 (s), 1367 (m), 1188 (s), 644 (s). IR ν_{max}/cm⁻¹ (solid) 2931 (m), 1683 (s), 1366 (s), 1164 (s), 601 (s). MP: 139-140°C. Purity HPLC 92.7%, R_t= 2.0 min.

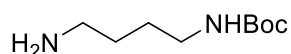
18i: (3aR,4R,7S,7aS)-2-(3-Methoxyphenyl)hexahydro-1*H*-4,7-methanoisindole-1,3(2*H*)-dione



Reaction repeated as general procedure A using 3-methoxyaniline (**16j**) gave the title compound (**18i**) as a white solid (0.80 mg, 50%). ¹H NMR (400 MHz, CDCl₃) δ 7.38 (t, 1H, *J*= 8.0, 16.0 Hz, Ar-H), 6.96-6.93 (dd, 1H, *J*= 4.0, 8.0 Hz, Ar-H), 6.84-6.82 (m, 1H, Ar-H), 6.78 (t, 1H, *J*= 4.0 Hz, Ar-H), 3.82 (s, 3H, CH₃), 3.24 (m, 2H, ((CH)₂C=O)C-H), 2.87 (m, 2H, CH₂), 1.73-1.65 (m, 4H, (CH₂)₂), 1.49-1.46 (m, 2H, ((CH)₂CH₂)C-H). ¹³C NMR (101 MHz, CDCl₃) δ 177.5, 160.1, 132.9, 130.0, 119.0, 114.4, 112.7, 55.4, 48.8, 42.2, 39.8, 24.9. HRMS (CI+) *m/z* calculated for C₁₆H₁₈NO₃: 272.1286. Found [M+H]⁺ 272.1282 (Diff -1.47 ppm). IR ν_{max}/cm⁻¹ (solid) 2957 (m), 1713 (s), 1370 (s), 1179 (s), 755 (s). MP: 141-142°C. Purity HPLC 91.1%, R_t= 2.1 min.

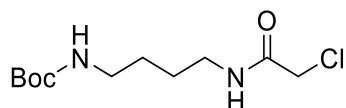
5.8.4 Synthesis of Thalidomide Warhead and Amide Linker

8: *Tert*-Butyl (4-aminobutyl)carbamate⁵⁸⁶



A solution of Boc₂O (1.24 g, 5.68 mmol, 0.1 equiv.) in CHCl₃ (30 mL, 0.9 M) was slowly added *via* a dropping funnel to a flask charged with 1,4-diaminobutane (**7**) (5 g, 26.6 mmol, 1 equiv.) and CHCl₃ (60 mL, 0.4 M), at a rate of 1 drop per 5 seconds. The solution was allowed to stir at room temperature for 16 hours. After this time, the crude mixture was washed with saturated aq. NaHCO₃ solution (x1), water (x1), and brine (x1). The organic phase was dried over MgSO₄, filtered, and concentration under reduced pressure to yield a crude brown oil. Purification *via* column chromatography eluting with 10-20% EtOAc in *n*-hexane yielded the title compound (**8**) a yellow oil (9.39 g, 88%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.89 (br s, 1H, N-H), 3.03 (t, 2H, *J*= 8.0, 12.0 Hz, CH₂), 2.62 (t, 2H, *J*= 8.0, 16.0 Hz, CH₂), 1.48-1.45 (m, 6H, (CH₂)₂, NH₂), 1.42 (s, 9H, (CH₃)₃). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 149.6, 80.1, 42.0, 35.2, 29.0, 28.7, 28.6, 28.4, 26.3. HRMS (CI+) *m/z* calculated for C₉H₂₁N₂O₂: 189.1603. Found [M+H]⁺ 189.1602 (Diff -0.53 ppm).

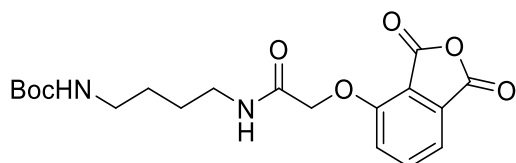
10: *Tert*-Butyl (4-(2-chloroacetamido)butyl)carbamate⁵⁹¹



Tert-butyl (4-aminobutyl) carbamate (**8**) (5 mL, 26.6 mmol, 1 equiv.) was dissolved in anhydrous THF (27 mL, 1 M) and cooled to 0°C. 2-Chloroacetyl chloride (**9**) (2.3 mL, 29.3 mmol, 1.1 equiv.) and DIPEA (4.4 mL, 26.6 mmol, 1 equiv.) were added, and allowed to stir at room temperature for 3 hours under a nitrogen atmosphere. After this time, the solution was concentrated under reduced pressure. The residue was diluted with ethyl acetate and washed with water (x3), and brine (x1). The organic phase was dried over MgSO₄, filtered, and concentrated *in vacuo* to yield the title compound (**10**) as a brown oil (6.12 g, 87%). ¹H NMR (400 MHz, CDCl₃) δ 6.66 (br s, 1H, N-H), 4.58 (br s, 1H, N-H), 4.05 (s, 2H, CH₂), 3.33 (dd, 2H, *J*= 12.9, 6.7 Hz, CH₂), 3.16 (m, 2H, CH₂), 1.57 (ddd, 4H, *J*= 9.7, 7.3, 3.9 Hz, (CH₂)₂), 1.44 (s, 9H, (CH₃)₃). ¹³C NMR (101 MHz, CDCl₃) δ 165.9, 156.0, 79.3, 42.7, 39.5, 28.4, 27.5, 26.6. HRMS

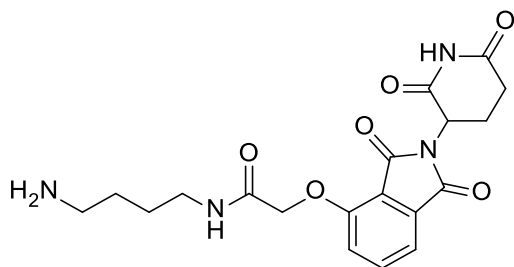
(ES+) m/z calculated for $C_{11}H_{21}^{35}ClN_2O_3Na$: 287.1141. Found $[M+Na]^+$ 287.1133 (Diff -2.79 ppm).

12: *Tert*-Butyl (4-(2-((1,3-dioxo-1,3-dihydroisobenzofuran-4-yl)oxy)acetamido)butyl)carbamate⁵⁹¹



A solution of 4-hydroxyisobenzofuran-1,3-dione (**11**) (1.0 g, 6.09 mmol, 1 equiv.) in anhydrous THF (8.7 mL, 0.7 M) was treated with crude *tert*-butyl (4-(2-chloroacetamido)butyl)carbamate (**10**) (2.42 g, 9.14 mmol, 1 equiv.) and cooled to 0°C. NaH (0.31 g, 7.92 mmol, 1.3 equiv.) was added portion wise and allowed to stir at reflux overnight under a nitrogen atmosphere. After this time, the reaction was diluted with DCM, and washed with water (x1), saturated aq. NaHCO₃ solution (x1), and brine (x1). The organic phase was dried over MgSO₄, filtered, and concentrated under reduced pressure to yield a crude orange oil. Purification *via* column chromatography eluting with 0-2% MeOH in DCM yielded the title compound (**12**) as a pale-yellow oil (2.08 g, 72%). ¹H NMR (400 MHz, CDCl₃) δ 7.64 (dd, 1H, J = 7.8, 0.6 Hz, Ar-H), 7.48 (t, 1H, J = 8.1 Hz, Ar-H), 7.14 (d, 1H, J = 8.1 Hz, Ar-H), 7.08 (br s, 1H, NH), 4.67 (br s, 1H, NH), 4.60 (s, 2H, CH₂), 3.35 – 3.29 (m, 2H, CH₂), 3.12 (d, 2H, J = 6.1 Hz, CH₂), 1.56 (m, 4H, 2(CH₂)), 1.45 (s, 9H, (CH₃)₃). ¹³C NMR (101 MHz, CDCl₃) δ 168.1, 167.5, 165.6, 156.0, 154.5, 131.1, 129.9, 124.9, 123.3, 116.9, 79.1, 68.1, 52.99, 39.5, 38.8, 28.4, 27.2, 26.6. HRMS (ES+) m/z calculated for C₁₉H₂₅N₂O₇: 393.1662. Found $[M+H]^+$ 393.1674 (Diff 3.05 ppm).

14: *N*-(4-Aminobutyl)-2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamide⁵⁹¹



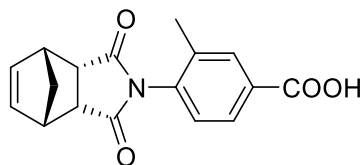
Tert-butyl (4-(2-((1,3-dioxo-1,3-dihydroisobenzofuran-4-yl)oxy)acetamido)butyl)carbamate (**12**) (0.1 g, 0.26 mmol, 1 equiv.), 3-aminopiperidine-2,6-dione (**13**) (0.04 g, 0.29 mmol, 1.1

equiv.) and KOAc (0.06 g, 0.65 mmol, 2.5 equiv.) were dissolved in AcOH (2.6 mL, 0.1 M) and heated in a microwave at 150°C for 12 minutes. After this time, the reaction mixture was cooled and concentrated under air flow to yield a crude dark blue oil. Purification *via* column chromatography with normal phase silica gel eluting with the solvent system 1% Water, 2% methanol and 7% ethyl acetate yielded the title compound (**14**) as a yellow oil (28 mg, 28%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.57-7.22 (m, 3H, Ar-H), 4.68 (m, 1H, NC-H), 3.78 (s, 2H, CH₂), 2.94 (m, 2H, CH₂), 2.54 (m, 2H, CH₂), 2.49-2.25 (m, 4H, (CH₂)₂), 1.64-1.62 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 175.4, 170.3, 169.5, 168.0, 167.8, 160.2, 146.3, 134.1, 133.7, 120.6, 115.0, 68.0, 64.0, 44.2, 38.0, 30.7, 29.4, 29.1, 23.0. LRMS (ES+) *m/z* calculated for C₁₉H₂₃N₄O₆: 403.2. Found [M+H]⁺ 403.2.

5.8.5 US Carboxylic Acids

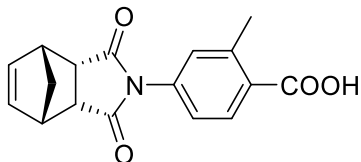
5.8.5.1 Active Unsaturated series⁵⁸⁴

4a: 4-((3a*R*,4*S*,7*R*,7a*S*)-1,3-Dioxo-1,3,3a,4,7,7a-hexahydro-2*H*-4,7-methanoisindol-2-yl)-3-methylbenzoic acid⁵⁸⁴



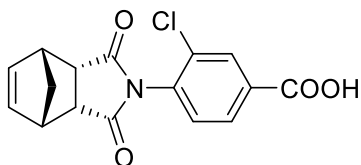
Reaction repeated as general procedure A using (3a*R*,4*S*,7*R*,7a*S*)-3a,4,7,7a-tetrahydro-4,7-methanoisobenzofuran-1,3-dione (**2a**) and 4-amino-3-methylbenzoic acid (**3a**) gave the title compound (**4a**) as a pale-pink solid (1.66 g, 84%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.96 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.89 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.21 (d, 0.5H, *J* = 8.0, Ar-H), 7.02 (d, 0.5H, *J* = 8.0 Hz, Ar-H), 6.32 (m, 2H, C=C-H), 3.60 (m, 2H, ((CH)₂C=O)C-H), 3.42 (m, 2H, ((CH)₂CH₂)C-H), 2.15 (s, 3H, CH₃), 1.72-1.69 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 179.3, 169.6, 139.8, 139.6, 138.5, 133.9, 133.7, 130.5, 129.8, 129.8, 50.5, 47.6, 47.0, 44.9, 44.5, 19.3. HRMS (ES+) *m/z* calculated for C₁₇H₁₆NO₄: 298.1079. Found [M+H]⁺ 298.1084 (Diff 1.68 ppm).

4c: 4-((3aR,4S,7R,7aS)-1,3-Dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-2-methylbenzoic acid



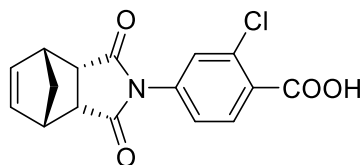
Reaction repeated as general procedure A using (3aR,4S,7R,7aS)-3a,4,7,7a-tetrahydro-4,7-methanoisobenzofuran-1,3-dione (**2a**) and 4-amino-2-methylbenzoic acid (**3b**) gave the title compound (**4c**) as a pale-pink solid (0.23 g, quant.). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.07 (s, 1H, Ar-H), 8.05 – 7.98 (m, 1H, Ar-H), 7.17 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.40 (m, 2H, C=C-H), 3.47 – 3.40 (m, 2H, ((CH)₂C=O)C-H), 2.92 (m, 2H, ((CH)₂CH₂)C-H), 2.25 (d, 3H, *J* = 6.9 Hz, CH₃), 1.68–1.56 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.6, 176.5, 170.4, 138.0, 137.9, 136.5, 136.2, 133.2, 128.7, 128.6, 127.9, 48.6, 45.9, 45.3, 43.3, 43.1, 17.9. HRMS (ES-) *m/z* calculated for C₁₇H₁₄NO₄: 296.0923. Found [M-H]⁺ 296.0928 (Diff 1.69 ppm).

4d: 3-Chloro-4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)benzoic acid



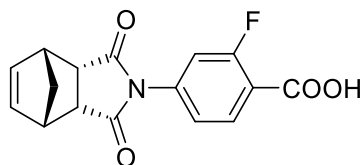
Reaction repeated as general procedure A using (3aR,4S,7R,7aS)-3a,4,7,7a-tetrahydro-4,7-methanoisobenzofuran-1,3-dione (**2a**) and 4-amino-3-chlorobenzoic acid (**3c**) gave the title compound (**4d**) as a white solid (0.76 g, 82%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.21 (s, 1H, Ar-H), 7.84 (m, 1H, Ar-H), 7.71 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.31 (m, 2H, C=C-H), 3.46–3.44 (m, 2H, ((CH)₂C=O)C-H), 3.23–3.19 (m, 2H, ((CH)₂CH₂)C-H), 1.62–1.59 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.3, 176.2, 159.8, 139.7, 136.2, 136.2, 133.1, 132.8, 130.7, 128.0, 122.4, 54.0, 45.5, 45.2, 43.2, 43.1. HRMS (ES-) *m/z* calculated for C₁₆H₁₁³⁵ClNO₄: 316.0377. Found [M-H]⁺ 316.0371 (Diff -1.90 ppm).

4e: 2-Chloro-4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)benzoic acid



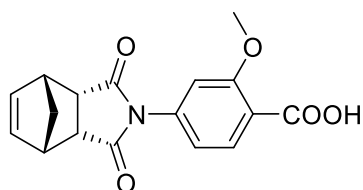
Reaction repeated as general procedure A using (3aR,4S,7R,7aS)-3a,4,7,7a-tetrahydro-4,7-methanoisobenzofuran-1,3-dione (**2a**) and 4-amino-2-chlorobenzoic acid (**3d**) gave the title compound (**4e**) as a white solid (0.43 g, 77%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.91 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.36 (m, 1H, Ar-H), 7.29 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.29 (m, 2H, C=C-H), 3.56-3.55 (m, 2H, ((CH)₂C=O)C-H), 3.45-3.44 (m, 2H, ((CH)₂CH₂)C-H), 1.79-1.69 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.9, 165.0, 143.5, 134.4, 130.4, 128.4, 127.7, 125.3, 124.8, 51.8, 48.9, 45.3. HRMS (ES-) *m/z* calculated for C₁₆H₁₁³⁵ClNO₄: 316.0377. Found [M-H]⁺ 316.0382 (Diff 1.58 ppm).

4f: 4-((3aR,4S,7R,7aS)-1,3-Dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-2-fluorobenzoic acid



Reaction repeated as general procedure A using (3aR,4S,7R,7aS)-3a,4,7,7a-tetrahydro-4,7-methanoisobenzofuran-1,3-dione (**2a**) and 4-amino-2-fluorobenzoic acid (**3e**) gave the title compound (**4f**) as a white solid (0.23 g, 40%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.01 (t, 1H, *J* = 8.0, 16.0 Hz, Ar-H), 7.16-7.11 (m, 2H, Ar-H), 6.29-6.28 (m, 2H, C=C-H), 3.56-3.55 (m, 2H, ((CH)₂C=O)C-H), 3.46-3.44 (m, 2H, ((CH)₂CH₂)C-H), 1.78-1.69 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.8, 165.8, 165.5, 153.5, 134.4, 132.1, 122.1, 115.2, 114.7, 51.7, 45.8, 45.3. ¹⁹F NMR (376 MHz, CD₃OD-*d*4) δ -109.8. HRMS (ES-) *m/z* calculated for C₁₆H₁₁¹⁹FNO₄: 300.0672. Found [M-H]⁺ 300.0676 (Diff 1.33 ppm).

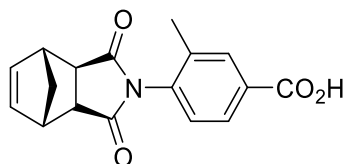
4g: 4-((3aR,4S,7R,7aS)-1,3-Dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-2-methoxybenzoic acid



Reaction repeated as general procedure A using (3aR,4S,7R,7aS)-3a,4,7,7a-tetrahydro-4,7-methanoisobenzofuran-1,3-dione (**2a**) and 4-amino-2-methoxybenzoic acid (**3f**) gave the title compound (**4g**) as a white solid (0.39 g, 70%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.20 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.63-7.58 (m, 2H, Ar-H), 6.25 (m, 2H, C=C-H), 3.95 (s, 3H, CH₃), 3.57-3.56 (m, 2H, ((CH)₂C=O)C-H), 3.42-3.41 (m, 2H, ((CH)₂CH₂)C-H), 1.64-1.60 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.4, 168.5, 163.2, 140.0, 136.7, 132.6, 120.1, 115.1, 106.3, 56.6, 51.6, 45.6, 45.1. HRMS (ES-) *m/z* calculated for C₁₇H₁₄NO₅: 312.0872. Found [M-H]⁺ 312.0877 (Diff 1.60 ppm).

5.8.5.2 Inactive Unsaturated series⁵⁸⁴

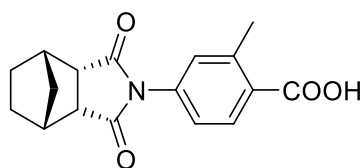
4b: 4-((3aR,4R,7S,7aR)-1,3-Dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-3-methylbenzoic acid⁵⁸⁴



Reaction repeated as general procedure A using (3aR,4R,7S,7aR)-3a,4,7,7a-tetrahydro-4,7-methanoisobenzofuran-1,3-dione (**2b**) and 4-amino-3-methylbenzoic acid (**3a**) gave the title compound (**4b**) as a pale-pink solid (1.84 g, 93%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 7.91 (d, 1H, *J*= 4.0 Hz, Ar-H), 7.84 (d, 1H, *J*= 12.0 Hz, Ar-H), 7.11 (d, 0.5H, *J*= 8.0 Hz, Ar-H), 7.06 (d, 0.5H, *J*= 8.0 Hz, Ar-H), 6.30-6.29 (m, 2H, C=C-H), 3.25-3.22 (m, 2H, ((CH)₂C=O)C-H), 2.86-2.83 (m, 2H, ((CH)₂CH₂)C-H), 2.11 (s, 3H, CH₃), 1.56-1.47 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 179.3, 179.3, 169.6, 139.8, 139.6, 133.9, 133.8, 130.5, 129.8, 129.4, 50.5, 47.7, 47.0, 19.3. HRMS (CI+) *m/z* calculated for C₁₇H₁₆NO₄: 298.1079. Found [M-H]⁺ 298.1082 (Diff 1.01 ppm).

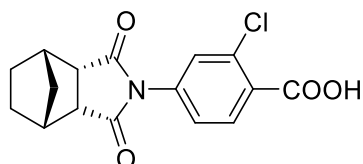
5.8.5.3 Active Saturated series ⁵⁸⁴

4h: 4-((3aR,4R,7S,7aS)-1,3-Dioxooctahydro-2H-4,7-methanoisindol-2-yl)-2-methylbenzoic acid



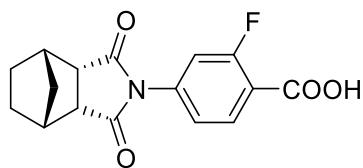
Reaction repeated as general procedure A using (3aR,4R,7S,7aS)-hexahydro-4,7-methanoisobenzofuran-1,3-dione (**17**) and 4-amino-2-methylbenzoic acid (**3b**) gave the title compound (**4h**) as a white solid (0.24 g, 46%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.03 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.85 (m, 1H, Ar-H), 7.31 (d, 1H, *J* = 8.0 Hz, Ar-H), 3.66-3.64 (m, 2H, ((CH)₂C=O)C-H), 2.56 (s, 3H, CH₃), 2.22-2.16 (m, 2H, CH₂), 1.86-1.84 (m, 2H, ((CH)₂CH₂)C-H), 1.43-1.37 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 177.2, 164.2, 139.1, 139.0, 136.4, 125.8, 124.6, 122.4, 42.5, 42.0, 37.7, 25.6, 25.3, 19.1. HRMS (ES-) *m/z* calculated for C₁₇H₁₆NO₄: 298.1080. Found [M-H]⁺ 298.1082 (Diff 0.67 ppm).

4i: 2-Chloro-4-((3aR,4R,7S,7aS)-1,3-dioxooctahydro-2H-4,7-methanoisindol-2-yl)benzoic acid



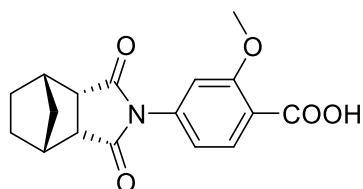
Reaction repeated as general procedure A using (3aR,4R,7S,7aS)-hexahydro-4,7-methanoisobenzofuran-1,3-dione (**17**) and 4-amino-2-chlorobenzoic acid (**3d**) gave the title compound (**4i**) as a white solid (0.15 g, 63%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.21 (d, 1H, *J* = 8.0 Hz, Ar-H), 8.11 (m, 1H, Ar-H), 7.23 (d, 1H, *J* = 8.0 Hz, Ar-H), 3.55-3.54 (m, 2H, ((CH)₂C=O)C-H), 2.31-2.27 (m, 2H, CH₂), 1.85-1.84 (m, 2H, ((CH)₂CH₂)C-H), 1.52-1.48 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.2, 164.0, 141.3, 136.5, 134.7, 123.3, 122.0, 121.7, 42.6, 41.8, 37.1, 24.7. HRMS (ES-) *m/z* calculated for C₁₆H₁₃³⁵ClNO₄: 318.0533. Found [M-H]⁺ 318.0523 (Diff -3.14 ppm).

4j: 4-((3aR,4R,7S,7aS)-1,3-Dioxooctahydro-2H-4,7-methanoisindol-2-yl)-2-fluorobenzoic acid



Reaction repeated as general procedure A using (3aR,4R,7S,7aS)-hexahydro-4,7-methanoisobenzofuran-1,3-dione (**17**) and 4-amino-2-fluorobenzoic acid (**3e**) gave the title compound (**4j**) as a white solid (0.13 g, 55%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.33 (t, 1H, *J*= 8.0, 16.0 Hz, Ar-H), 8.02 (m, 1H, Ar-H), 7.21 (m, 1H, Ar-H), 3.59-3.57 (m, 2H, ((CH)₂C=O)C-H), 2.44-2.42 (m, 2H, CH₂), 1.92-1.91 (m, 2H, ((CH)₂CH₂)C-H), 1.53-1.42 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.3, 166.6, 159.8, 145.2, 130.0, 122.3, 117.8, 111.0, 42.6, 41.7, 37.0, 26.1. HRMS (ES-) *m/z* calculated for C₁₆H₁₃¹⁹FNO₄: 302.0829. Found [M-H]⁺ 302.0833 (Diff 1.32 ppm).

4k: 4-((3aR,4R,7S,7aS)-1,3-Dioxooctahydro-2H-4,7-methanoisindol-2-yl)-2-methoxybenzoic acid

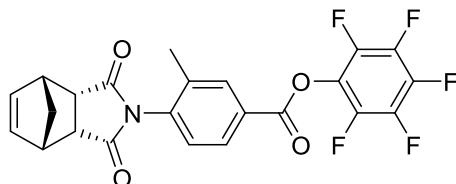


Reaction repeated as general procedure A using (3aR,4R,7S,7aS)-hexahydro-4,7-methanoisobenzofuran-1,3-dione (**17**) and 4-amino-2-methoxybenzoic acid (**3f**) gave the title compound (**4k**) as a white solid (0.19 g, 79%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.21 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.32-7.24 (m, 2H, Ar-H), 3.98 (s, 3H, CH₃), 3.45-3.43 (m, 2H, ((CH)₂C=O)C-H), 2.43-2.40 (m, 2H, CH₂), 1.74-1.73 (m, 2H, ((CH)₂CH₂)C-H), 1.52-1.34 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.2, 167.1, 162.3, 142.6, 134.7, 120.5, 115.2, 106.7, 56.2, 43.3, 42.5. HRMS (ES-) *m/z* calculated for C₁₇H₁₆NO₅: 314.1029. Found [M-H]⁺ 314.1035 (Diff 1.91 ppm).

5.8.6 Perfluorophenyl US analogues

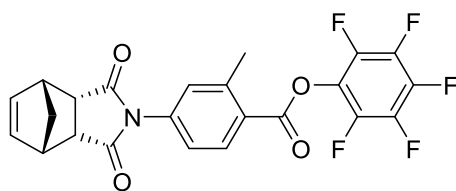
5.8.6.1 Active Unsaturated series⁶²⁶

6a: Perfluorophenyl 4-((3a*R*,4*S*,7*R*,7a*S*)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2*H*-4,7-methanoisindol-2-yl)-3-methylbenzoate



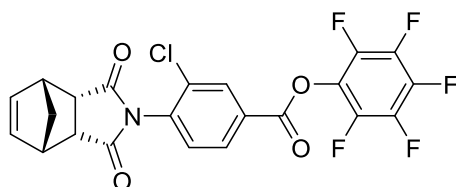
Reaction repeated as general procedure B using 4-((3a*R*,4*S*,7*R*,7a*S*)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2*H*-4,7-methanoisindol-2-yl)-3-methylbenzoic acid (**4a**) gave the title compound (**6a**) as a yellow solid (42.2 mg, 54%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.06 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.98 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.23 (d, 0.5H, *J* = 8.0 Hz, Ar-H), 7.04 (d, 0.5H, *J* = 8.0 Hz, Ar-H), 6.23 (m, 2H, C=C-H), 3.52-3.50 (m, 2H, ((CH)₂C=O)C-H), 3.22-3.20 (m, 2H, ((CH)₂CH₂)C-H), 2.13-2.09 (d, 3H, *J* = 16.0 Hz, CH₃), 1.71-1.58 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 177.2, 165.3, 146.0, 143.7, 141.1, 140.9, 139.5, 137.4, 133.6, 130.2, 126.9, 120.9, 53.4, 46.1, 45.7, 18.1. HRMS (ES⁺) *m/z* calculated for C₂₃H₁₅¹⁹F₅NO₄: 464.0921. Found [M+H]⁺ 464.0932 (Diff 2.37 ppm).

6c: Perfluorophenyl 4-((3a*R*,4*S*,7*R*,7a*S*)-1,3-Dioxo-1,3,3a,4,7,7a-hexahydro-2*H*-4,7-methanoisindol-2-yl)-2-methylbenzoate



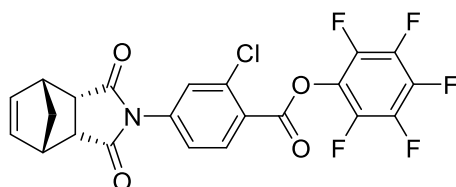
Reaction repeated as general procedure B using 4-((3a*R*,4*S*,7*R*,7a*S*)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2*H*-4,7-methanoisindol-2-yl)-2-methylbenzoic acid (**4c**) gave the title compound (**6c**) as a yellow solid (45.1 mg, 58%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.12 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.96 (m, 1H, Ar-H), 7.30 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.24 (m, 2H, C=C-H), 3.51-3.50 (m, 2H, ((CH)₂C=O)C-H), 3.24-3.23 (m, 2H, ((CH)₂CH₂)C-H), 2.45 (s, 3H, CH₃), 1.63-1.60 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.0, 165.4, 143.9, 143.2, 141.1, 141.0, 138.3, 136.6, 132.7, 125.3, 120.3, 120.1, 54.1, 45.0, 44.4, 19.0. HRMS (ES⁺) *m/z* calculated for C₂₃H₁₅¹⁹F₅NO₄: 464.0921. Found [M+H]⁺ 464.0926 (Diff 1.10 ppm).

6d: Perfluorophenyl 3-chloro-4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)benzoate



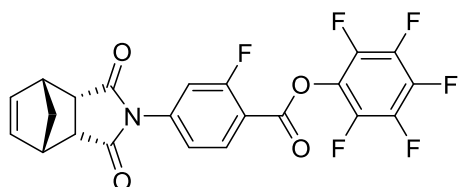
Reaction repeated as general procedure B using 3-chloro-4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)benzoic acid (**4d**) gave the title compound (**6d**) as a yellow solid (46 mg, 61%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.33 (m, 1H, Ar-H), 8.01 (m, 1H, Ar-H), 7.62 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.22 (m, 2H, C=C-H), 3.62-3.61 (m, 2H, ((CH)₂C=O)C-H), 3.46-3.44 (m, 2H, ((CH)₂CH₂)C-H), 1.66-1.58 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 177.8, 165.7, 146.9, 143.6, 140.2, 139.6, 138.4, 132.0, 130.9, 130.5, 127.7, 127.0, 53.3, 46.2, 45.8. HRMS (ES⁺) *m/z* calculated for C₂₂H₁₂³⁵Cl¹⁹F₅NO₄: 484.0375. Found [M+H]⁺ 484.0388 (Diff 2.69 ppm).

6e: Perfluorophenyl 2-chloro-4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)benzoate



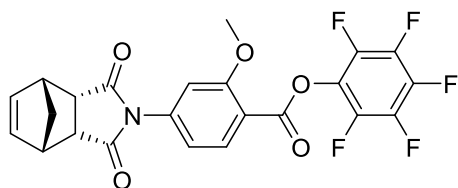
Reaction repeated as general procedure B using 2-chloro-4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)benzoic acid (**4e**) gave the title compound (**6e**) as a pale-yellow solid (31 mg, 41%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.35 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.99 (m, 1H, Ar-H), 7.46 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.33 (m, 2H, C=C-H), 3.52-3.51 (m, 2H, ((CH)₂C=O)C-H), 3.41-3.40 (m, 2H, ((CH)₂CH₂)C-H), 1.52-1.50 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.7, 163.3, 144.4, 143.0, 140.5, 140.2, 140.0, 138.4, 136.2, 130.5, 125.5, 122.0, 121.7, 52.6, 45.3, 45.1. HRMS (ES⁺) *m/z* calculated for C₂₂H₁₂³⁵Cl¹⁹F₅NO₄: 484.0375. Found [M+H]⁺ 484.0370 (Diff -1.03 ppm).

6f: Perfluorophenyl 4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-2-fluorobenzoate



Reaction repeated as general procedure B using 4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-2-fluorobenzoic acid (**4f**) gave the title compound (**6f**) as an orange solid (25 mg, 32%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.34 (m, 1H, Ar-H), 7.88 (m, 1H, Ar-H), 7.11 (m, 1H, Ar-H), 6.22 (m, 2H, C=C-H), 3.64-3.63 (m, 2H, ((CH)₂C=O)C-H), 3.21-3.20 (m, 2H, ((CH)₂CH₂)C-H), 1.74-1.71 (m, 2H, CH₂).

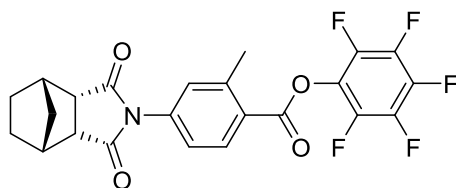
6g: Perfluorophenyl 4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-2-methoxybenzoate



Reaction repeated as general procedure B using 4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-2-methoxybenzoic acid (**4g**) gave the title compound (**6g**) as a yellow solid (38 mg, 49%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.44 (m, 1H, Ar-H), 7.56 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.24 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.21 (m, 2H, C=C-H), 3.91 (d, 3H, *J* = 12.0 Hz, CH₃), 3.62-3.60 (m, 2H, ((CH)₂C=O)C-H), 3.31-3.30 (m, 2H, ((CH)₂CH₂)C-H), 1.66-1.63 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.3, 170.1, 164.7, 143.3, 142.0, 140.6, 140.2, 140.0, 136.9, 132.5, 118.8, 107.6, 105.2, 57.2, 53.1, 45.1, 44.8. HRMS (ES⁺) *m/z* calculated for C₂₃H₁₅¹⁹F₅NO₅: 480.0870. Found [M+H]⁺ 480.0876 (Diff 1.25 ppm).

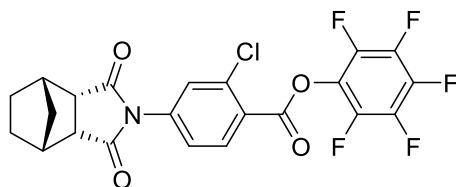
5.8.6.2 Active Saturated series

6h: Perfluorophenyl 4-((3aR,4R,7S,7aS)-1,3-dioxooctahydro-2H-4,7-methanoisindol-2-yl)-2-methylbenzoate



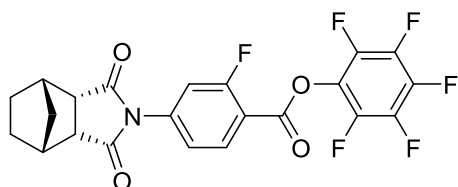
Reaction repeated as general procedure B using 4-((3aR,4R,7S,7aS)-1,3-dioxooctahydro-2H-4,7-methanoisindol-2-yl)-2-methylbenzoic acid (**4h**) gave the title compound (**6h**) as an off-white solid (86 mg, 57%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.16 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.66 (m, 1H, Ar-H), 7.31 (m, 1H, Ar-H), 3.44-3.43 (m, 2H, ((CH)₂C=O)C-H), 2.55 (s, 3H, CH₃), 2.62-2.59 (m, 2H, CH₂), 1.74-1.73 (m, 2H, ((CH)₂CH₂)C-H), 1.56-1.50 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.4, 166.3, 145.7, 143.1, 141.6, 140.2, 139.9, 131.4, 126.4, 120.8, 120.2, 42.3, 41.0, 37.8, 24.3, 19.7. HRMS (ES+) *m/z* calculated for C₂₃H₁₇¹⁹F₅NO₄: 466.1077. Found [M+H]⁺ 466.1080 (Diff 0.64 ppm).

6i: Perfluorophenyl 2-chloro-4-((3aR,4R,7S,7aS)-1,3-dioxooctahydro-2H-4,7-methanoisindol-2-yl)benzoate



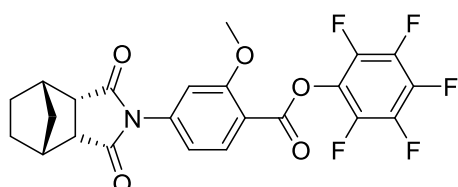
Reaction repeated as general procedure B using 2-chloro-4-((3aR,4R,7S,7aS)-1,3-dioxooctahydro-2H-4,7-methanoisindol-2-yl)benzoic acid (**4i**) gave the title compound (**6i**) as a pale-yellow solid (40 mg, 53%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.21 (m, 1H, Ar-H), 7.84 (m, 1H, Ar-H), 7.13 (m, 1H, Ar-H), 3.55 (m, 2H, ((CH)₂C=O)C-H), 2.21-2.20 (m, 2H, CH₂), 1.74-1.72 (m, 2H, ((CH)₂CH₂)C-H), 1.49-1.46 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.4, 163.3, 144.7, 144.3, 141.6, 140.5, 139.8, 136.1, 130.2, 130.0, 123.7, 121.9, 42.1, 42.0, 35.4, 26.3.

6j: Perfluorophenyl 4-((3aR,4R,7S,7aS)-1,3-dioxooctahydro-2H-4,7-methanoisindol-2-yl)-2-fluorobenzoate



Reaction repeated as general procedure B using 4-((3aR,4R,7S,7aS)-1,3-dioxooctahydro-2H-4,7-methanoisindol-2-yl)-2-fluorobenzoic acid (**4j**) gave the title compound (**6j**) as a beige solid (32 mg, 40%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.26 (m, 1H, Ar-H), 7.84 (m, 1H, Ar-H), 7.15 (d, 1H, *J* = 8.0 Hz, Ar-H), 3.53-3.52 (m, 2H, ((CH)₂C=O)C-H), 2.51-2.48 (m, 2H, CH₂), 1.77-1.76 (m, 2H, ((CH)₂CH₂)C-H), 1.52-1.47 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 177.4, 165.2, 162.8, 143.8, 142.6, 140.9, 140.2, 132.6, 120.4, 117.7, 111.7, 43.8, 43.6, 36.2, 24.8. HRMS (ES+) *m/z* calculated for C₂₂H₁₄¹⁹F₆NO₄: 470.0827. Found [M+H]⁺ 470.0836 (Diff 1.91 ppm).

6k: Perfluorophenyl 4-((3aR,4R,7S,7aS)-1,3-dioxooctahydro-2H-4,7-methanoisindol-2-yl)-2-methoxybenzoate

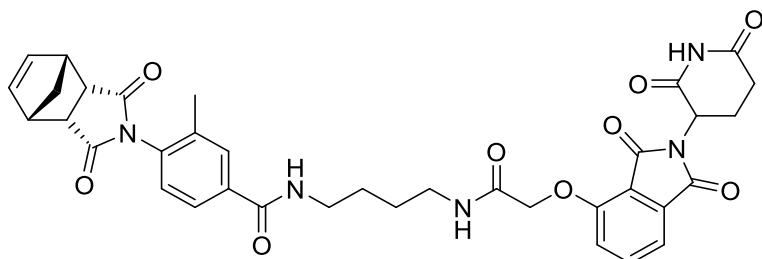


Reaction repeated as general procedure B using 4-((3aR,4R,7S,7aS)-1,3-dioxooctahydro-2H-4,7-methanoisindol-2-yl)-2-methoxybenzoic acid (**4k**) gave the title compound (**6k**) as a yellow solid (36 mg, 47%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.22 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.31-7.26 (m, 2H, Ar-H), 3.92 (s, 3H, CH₃), 3.44-3.43 (m, 2H, ((CH)₂C=O)C-H), 2.22-2.18 (m, 2H, CH₂), 1.74-1.73 (m, 2H, ((CH)₂CH₂)C-H), 1.54-1.50 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.1, 166.6, 159.9, 143.3, 142.7, 142.8, 141.5, 139.6, 130.4, 117.8, 110.2, 106.3, 55.6, 42.0, 41.8, 37.7, 25.3. HRMS (ES+) *m/z* calculated for C₂₃H₁₇¹⁹F₅NO₅: 482.1027. Found [M+H]⁺ 482.1034 (Diff -0.62 ppm).

5.8.7 PROTAC agents

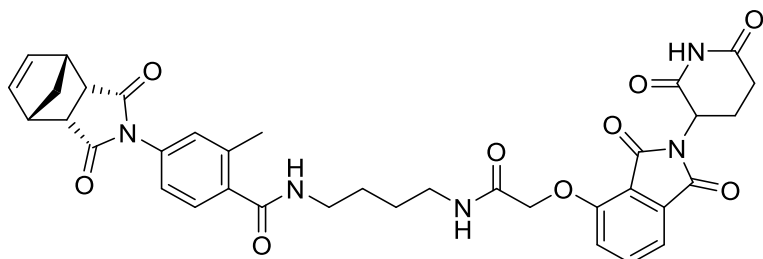
5.8.7.1 Active Unsaturated series⁵⁹¹

15a, (RGC): 4-((3aR,4S,7R,7aS)-1,3-Dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-N-(4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamido)butyl)-3-methylbenzamide



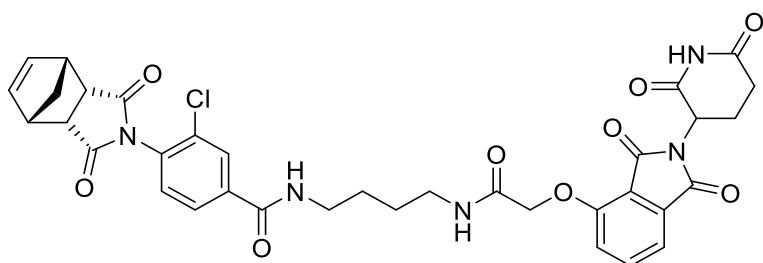
Reaction repeated as general procedure C using perfluorophenyl 4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-3-methylbenzoate (**6a**) gave the title compound (**15a**) as a yellow oil (25 mg, 34%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.51-8.47 (m, 2H, Ar-H), 8.21-8.19 (m, 2H, Ar-H), 7.88 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.32-7.25 (m, 1H, Ar-H), 6.29 (m, 2H, C=C-H), 4.71 (s, 2H, CH₂), 4.66 (m, 1H, C-H), 3.78-3.77 (m, 2H, ((CH)₂C=O)C-H), 3.62-3.60 (m, 2H, ((CH)₂CH₂)C-H), 3.37 (m, 2H, CH₂), 3.21-2.13 (m, 6H, (CH₂)₂, CH₂), 2.89 (s, 3H, CH₃), 2.05-1.98 (m, 4H, (CH₂)₂), 1.59-1.51 (m, 2H, CH₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.2, 175.1, 170.3, 169.8, 168.7, 168.4, 159.4, 142.6, 138.1, 134.0, 133.5, 133.2, 132.7, 130.7, 129.6, 128.6, 122.4, 120.0, 118.9, 69.1, 64.4, 54.5, 45.6, 45.1, 40.2, 39.4, 30.2, 27.9, 22.2, 18.3. HRMS (ES⁺) *m/z* calculated for C₃₆H₃₆N₅O₉: 682.2513. Found [M+H]⁺ 682.2522 (Diff 1.32 ppm). IR ν_{max}/cm⁻¹ (oil) 3323 (m), 3095 (m), 2979 (w), 1706 (s), 1698 (s), 1684 (s), 1385 (s), 1212 (s). Purity HPLC 94.9%, R_t = 4.9 min.

15c: 4-((3aR,4S,7R,7aS)-1,3-Dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-N-(4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamido)butyl)-2-methylbenzamide



Reaction repeated as general procedure C using perfluorophenyl 4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-2-methylbenzoate (**6c**) gave the title compound (**15c**) as a yellow oil (23 mg, 31%). ¹H NMR (400 MHz, CD₃OD-*d*₄) δ 8.42 (d, 1H, *J* = 8.0 Hz, Ar-H), 8.38-8.37 (m, 1H, Ar-H), 8.20-8.18 (m, 2H, Ar-H), 7.86 (d, 1H, Ar-H), 7.29 (m, 1H, Ar-H), 6.31 (m, 2H, C=C-H), 4.84-4.82 (m, 1H, C-H), 4.69 (s, 2H, CH₂), 3.69-3.68 (m, 2H, ((CH)₂C=O)C-H), 3.51-3.49 (m, 2H, ((CH)₂CH₂)C-H), 3.29 (m, 2H, CH₂), 3.17-3.14 (m, 2H, CH₂), 2.55 (s, 3H, CH₃), 2.22-2.13 (m, 4H, (CH₂)₂), 1.84-1.80 (m, 2H, CH₂), 1.63-1.58 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*₄) δ 176.4, 174.8, 170.1, 169.8, 168.6, 168.2, 158.9, 141.5, 138.3, 134.3, 133.1, 132.9, 132.5, 131.2, 129.4, 127.0, 121.7, 120.2, 119.0, 70.2, 65.1, 54.3, 45.7, 45.1, 40.4, 40.1, 31.6, 28.4, 21.9, 17.9. HRMS (ES+) *m/z* calculated for C₃₆H₃₆N₅O₉: 682.2513. Found [M+H]⁺ 682.2517 (Diff 0.59 ppm). Purity HPLC 87.7%, R_t = 4.2 min.

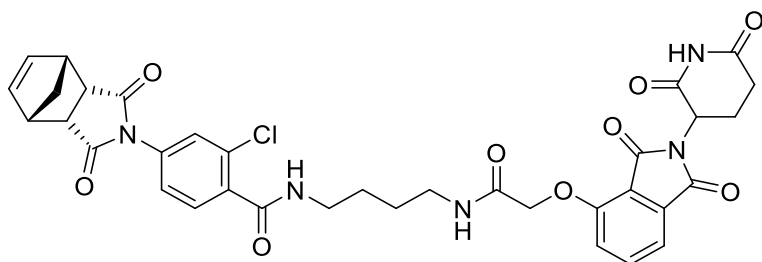
15d: 3-Chloro-4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-N-(4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamido)butyl)benzamide



Reaction repeated as general procedure C using perfluorophenyl 3-chloro-4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)benzoate (**6d**) gave the title compound (**15d**) as a yellow oil (13 mg, 18%). ¹H NMR (400 MHz, CD₃OD-*d*₄) δ 8.31 (m, 2H,

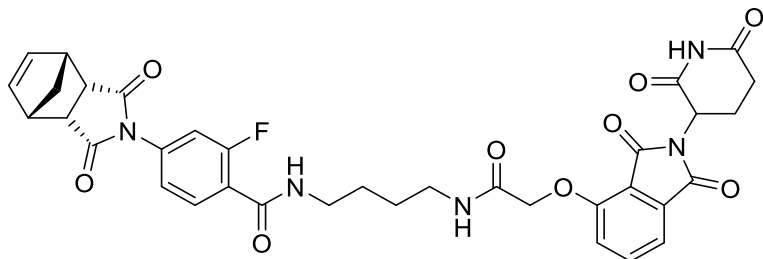
Ar-H), 8.19 (d, 1H, $J = 8.0$ Hz, Ar-H), 8.07 (d, 1H, $J = 8.0$ Hz, Ar-H), 7.65 (t, 1H, $J = 8.0, 16.0$ Hz, Ar-H), 7.26 (m, 1H, Ar-H), 6.23 (m, 2H, C=C-H), 4.82-4.80 (m, 1H, C-H), 4.66 (m, 2H, CH₂), 3.62-3.60 (m, 2H, ((CH)₂C=O)C-H), 3.54-3.53 (m, 2H, ((CH)₂CH₂)C-H), 3.26-3.23 (m, 2H, CH₂), 3.11-3.08 (m, 2H, CH₂), 2.26-2.13 (m, 4H, (CH₂)₂), 1.78-1.76 (m, 2H, CH₂), 1.58-1.53 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 177.2, 172.6, 169.2, 168.7, 168.2, 167.5, 160.3, 142.6, 140.9, 134.7, 133.8, 133.6, 133.0, 128.4, 128.1, 120.9, 118.0, 117.1, 69.4, 63.2, 55.6, 45.7, 45.2, 44.2, 37.6, 30.4, 27.3, 27.2, 21.5. HRMS (ES+) m/z calculated for C₃₅H₃₃³⁵ClN₅O₉: 702.1967. Found [M+H]⁺ 702.1973 (Diff 0.85 ppm). Purity HPLC 97.7%, R_t = 4.2 min.

15e: 2-Chloro-4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-N-(4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamido)butyl)benzamide



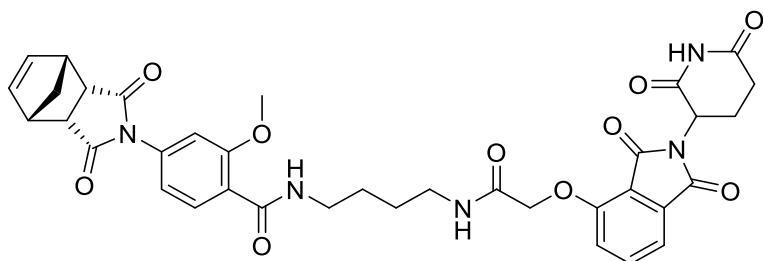
Reaction repeated as general procedure C using perfluorophenyl 2-chloro-4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)benzoate (**6e**) gave the title compound (**15e**) as a yellow oil (16 mg, 22%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.36-3.32 (m, 2H, Ar-H), 8.24-8.22 (d, 2H, Ar-H), 7.68-7.67 (m, 1H, Ar-H), 7.31 (d, 1H, $J = 8.0$ Hz, Ar-H), 6.23 (m, 2H, C=C-H), 4.86-4.84 (m, 1H, C-H), 4.62 (m, 2H, CH₂), 3.66-3.65 (m, 2H, ((CH)₂C=O)C-H), 3.49-3.47 (m, 2H, ((CH)₂CH₂)C-H), 3.30-3.29 (m, 2H, CH₂), 3.13-3.11 (m, 2H, CH₂), 2.25-2.21 (m, 4H, (CH₂)₂), 1.78-1.77 (m, 2H, CH₂), 1.56-1.50 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 177.0, 176.7, 176.2, 170.3, 170.2, 169.7, 160.6, 144.5, 140.2, 136.6, 134.6, 134.2, 131.8, 127.7, 123.1, 122.8, 119.6, 116.9, 70.2, 64.3, 54.1, 45.7, 45.6, 41.1, 40.8, 32.6, 27.3, 22.2. HRMS (ES+) m/z calculated for C₃₅H₃₃³⁵ClN₅O₉: 702.1967. Found [M+H]⁺ 702.1970 (Diff 0.43 ppm). Purity HPLC 96.7%, R_t = 4.2 min.

15f: 4-((3aR,4S,7R,7aS)-1,3-Dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-N-(4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamido)butyl)-2-fluorobenzamide



Reaction repeated as general procedure C using perfluorophenyl 4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-2-fluorobenzoate (**6f**) gave the title compound (**15f**) as a brown oil (19 mg, 26%). ¹H NMR (400 MHz, CD₃OD-*d*₄) δ 8.34-8.31 (m, 2H, Ar-H), 8.22 (m, 1H, Ar-H), 7.96 (m, 1H, Ar-H), 7.61 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.29 (d, 1H, *J* = 8.0 Hz, Ar-H), 6.22 (m, 2H, C=C-H), 4.88-4.86 (m, 1H, C-H), 4.43 (s, 2H, CH₂), 3.68-3.67 (m, 2H, ((CH)₂C=O)C-H), 3.49-3.48 (m, 2H, ((CH)₂CH₂)C-H), 3.25-3.24 (m, 2H, CH₂), 3.18-3.16 (m, 2H, CH₂), 2.26-2.23 (m, 4H, (CH₂)₂), 1.77-1.75 (m, 2H, CH₂), 1.52-1.48 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*₄) δ 175.9, 174.1, 170.8, 170.4, 169.3, 167.8, 160.4, 160.2, 140.6, 140.5, 135.2, 134.7, 134.3, 133.6, 122.8, 122.6, 118.9, 116.3, 106.3, 70.2, 69.5, 54.4, 45.5, 45.5, 40.7, 40.2, 30.0, 27.6, 27.6, 22.5. HRMS (ES⁺) *m/z* calculated for C₃₅H₃₃¹⁹FN₅O₉: 686.2262. Found [M+H]⁺ 686.2264 (Diff 0.29 ppm). Purity HPLC 94.5%, R_t = 4.2 min.

15g: 4-((3aR,4S,7R,7aS)-1,3-Dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-N-(4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamido)butyl)-2-methoxybenzamide

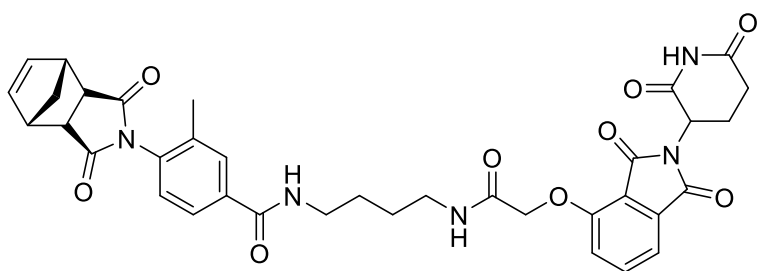


Reaction repeated as general procedure C using perfluorophenyl 4-((3aR,4S,7R,7aS)-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-methanoisindol-2-yl)-2-methoxybenzoate (**6g**) gave the title compound (**15g**) as a yellow oil (12 mg, 17%). ¹H NMR (400 MHz, CD₃OD-*d*₄) δ 8.33-

8.32 (m, 1H, Ar-H), 8.02-8.00 (m, 1H, Ar-H), 7.97-7.94 (m, 2H, Ar-H), 7.64 (m, 1H, Ar-H), 7.31 (d, 1H, $J = 8.0$ Hz, Ar-H), 6.22 (m, 2H, C=C-H), 4.76-4.76 (m, 1H, C-H), 4.42 (s, 2H, CH₂), 3.98 (s, 3H, CH₃), 3.66-3.65 (m, 2H, ((CH)₂C=O)C-H), 3.48-3.46 (m, 2H, ((CH)₂CH₂)C-H), 3.30-3.28 (m, 2H, CH₂), 3.16-3.14 (m, 2H, CH₂), 2.64-2.41 (m, 4H, (CH₂)₂), 1.76-1.74 (m, 2H, CH₂), 1.49-1.42 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*₄) δ 176.3, 174.5, 170.2, 169.9, 169.2, 166.6, 155.4, 140.2, 139.8, 136.7, 135.5, 133.0, 129.9, 128.6, 119.6, 117.7, 117.4, 105.6, 70.3, 64.1, 55.5, 53.3, 45.6, 45.1, 40.7, 39.8, 30.4, 26.8, 20.9. HRMS (ES+) m/z calculated for C₃₆H₃₆N₅O₁₀: 698.2462. Found [M+H]⁺ 698.2473 (Diff 1.58 ppm). Purity HPLC 93.5%, $R_t = 4.2$ min.

5.8.7.2 Inactive Unsaturated series⁵⁹¹

15b, (iRGC): 4-((3*aR*,4*R*,7*S*,7*aS*)-1,3-Dioxo-1,3,3*a*,4,7,7*a*-hexahydro-2*H*-4,7-methanoisindol-2-yl)-*N*-(4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamido)butyl)-3-methylbenzamide

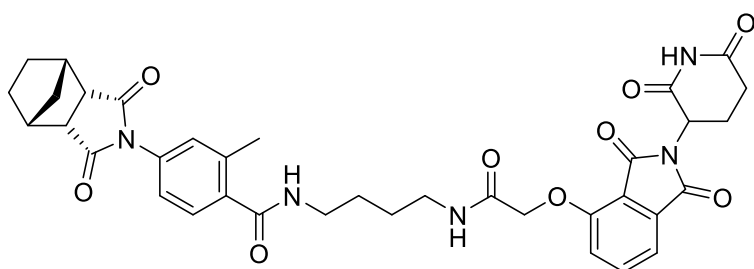


To a flask, *N*-(4-aminobutyl)-2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamide (**14**) (34.5 mg, 0.069 mmol, 1 equiv.) was added and dissolved in DMF (0.35 mL, 0.2 M). The flask was cooled to 0°C, then HATU (50 mg, 0.14 mmol, 2 equiv.) and DIPEA (0.04 mL, 0.21 mmol, 3 equiv.) was added and allowed to stir at room temperature for 10 minutes. 4-((3*aR*,4*R*,7*S*,7*aR*)-1,3-Dioxo-1,3,3*a*,4,7,7*a*-hexahydro-2*H*-4,7-methanoisindol-2-yl)-3-methylbenzoic acid (**6b**) (20 mg, 0.069 mmol, 1 equiv.) was added and continued to stir at room temperature for 16 hours. After this time, the solution was concentrated under reduced pressure to yield a crude brown oil. Purification *via* column chromatography using reverse phase silica gel eluting with 5-20% MeCN in H₂O switching to 0-30% MeOH in H₂O, followed by purification *via* a SCX-2 cartridge eluting with 1M NH₃ in MeOH afforded the title compound (**15b, iRGC**) as a yellow oil (18 mg, 24%). ¹H NMR (400 MHz, CD₃OD-*d*₄) δ 8.55 (m, 1H, Ar-H), 8.16 (t, 1H, $J = 8.0, 12.0$ Hz, Ar-H), 7.81-7.78 (m, 2H, Ar-H), 7.24-7.23 (m, 1H, Ar-H), 7.11-7.10 (m, 1H, Ar-H), 6.22 (m, 2H, C=C-H), 4.67-4.64 (m, 1H, C-H), 4.42 (s, 2H, CH₂), 3.62-3.61 (m, 2H, ((CH)₂C=O)C-H), 3.51-3.50 (m, 2H, ((CH)₂CH₂)C-H), 3.31-3.29 (m, 2H, CH₂), 3.24-

3.21 (m, 2H, CH₂), 2.76-2.71 (m, 4H, (CH₂)₂), 2.13 (s, 3H, CH₃), 1.79-1.76 (m, 2H, CH₂), 1.52-1.49 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.9, 174.6, 171.6, 171.0, 168.4, 166.3, 160.6, 141.5, 140.5, 134.6, 133.5, 133.0, 133.0, 129.5, 127.6, 125.9, 123.6, 118.0, 117.3, 69.8, 62.9, 55.9, 45.6, 45.2, 40.2, 38.6, 31.2, 28.3, 21.3, 17.9. HRMS (ES+) *m/z* calculated for C₃₆H₃₆N₅O₉: 682.2513. Found [M+H]⁺ 682.2541 (Diff 4.10 ppm). Purity HPLC 94.1%, R_t= 4.2 min.

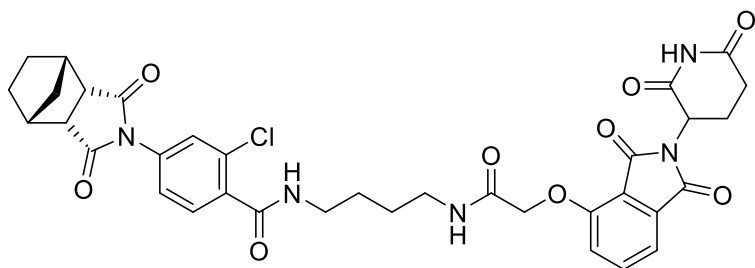
5.8.7.3 Active Saturated series⁵⁹¹

19a: 4-((3a*R*,4*R*,7*S*,7a*S*)-1,3-Dioxooctahydro-2*H*-4,7-methanoisindol-2-yl)-*N*-(4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamido)butyl)-2-methylbenzamide



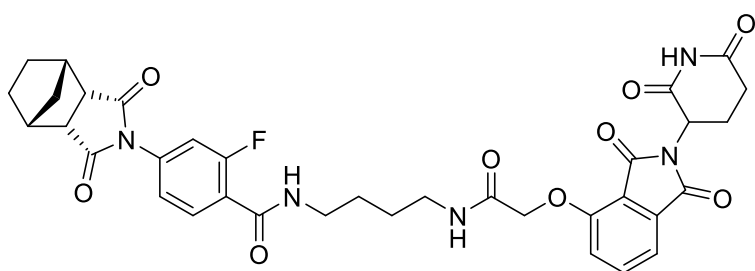
Reaction repeated as general procedure C using perfluorophenyl 4-((3a*R*,4*R*,7*S*,7a*S*)-1,3-dioxooctahydro-2*H*-4,7-methanoisindol-2-yl)-2-methylbenzoate (**6h**) gave the title compound (**19a**) as a yellow oil (12 mg, 20%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.36-8.22 (m, 3H, Ar-H), 7.84 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.54-7.51 (m, 2H, Ar-H), 4.44-4.42 (m, 1H, C-H), 4.21 (s, 2H, CH₂), 3.84-3.83 (m, 2H, ((CH)₂C=O)C-H), 3.37-3.35 (m, 2H, CH₂), 3.18-3.16 (m, 2H, CH₂), 2.67-2.37 (m, 6H, CH₃, (CH₂)₂), 2.21-2.19 (m, 2H, CH₂), 1.88-1.87 (m, 2H, ((CH)₂CH₂)C-H), 1.57-1.32 (m, 8H, (CH₂)₂, (CH₂)). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.8, 175.4, 169.6, 169.4, 168.2, 167.5, 160.4, 146.5, 140.6, 138.8, 135.1, 134.7, 134.6, 125.6, 122.5, 122.5, 119.4, 116.5, 64.3, 61.1, 41.2, 40.7, 39.4, 38.1, 37.8, 30.4, 26.8, 26.2, 21.2, 18.8. HRMS (ES+) *m/z* calculated for C₃₆H₃₈N₅O₉: 684.2669. Found [M+H]⁺ 684.2683 (Diff 2.05 ppm). Purity HPLC 94.0%, R_t= 5.0 min.

19b: 2-Chloro-4-((3aR,4R,7S,7aS)-1,3-dioxooctahydro-2H-4,7-methanoisindol-2-yl)-N-(4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamido)butyl)benzamide



Reaction repeated as general procedure C using perfluorophenyl 2-chloro-4-((3aR,4R,7S,7aS)-1,3-dioxooctahydro-2H-4,7-methanoisindol-2-yl)benzoate (**6i**) gave the title compound (**19b**) as a yellow oil (12 mg, 16%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.31-8.22 (m, 2H, Ar-H), 8.02 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.81 (d, 1H, *J* = 8.0 Hz, Ar-H), 7.35-7.30 (m, 2H, Ar-H), 4.84 (s, 2H, CH₂), 4.44-4.43 (m, 1H, C-H), 3.55-3.54 (m, 2H, ((CH)₂C=O)C-H), 3.32-3.31 (m, 2H, CH₂), 3.14-3.11 (m, 2H, CH₂), 2.68-2.41 (m, 4H, (CH₂)₂), 2.25-2.24 (m, 2H, CH₂), 1.81-1.80 (m, 2H, ((CH)₂CH₂)C-H), 1.57-1.27 (m, 8H, (CH₂)₂, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 177.6, 176.4, 174.8, 169.8, 169.6, 168.2, 158.2, 141.2, 140.7, 135.6, 134.5, 131.1, 130.5, 124.6, 123.3, 118.6, 117.5, 70.8, 65.7, 41.9, 41.5, 40.7, 40.6, 37.5, 31.9, 26.8, 24.2, 22.9. HRMS (ES+) *m/z* calculated for C₃₅H₃₅³⁵ClN₅O₉: 704.2123. Found [M+H]⁺ 704.2151 (Diff 3.98 ppm). Purity HPLC 94.5%, R_t = 4.2 min.

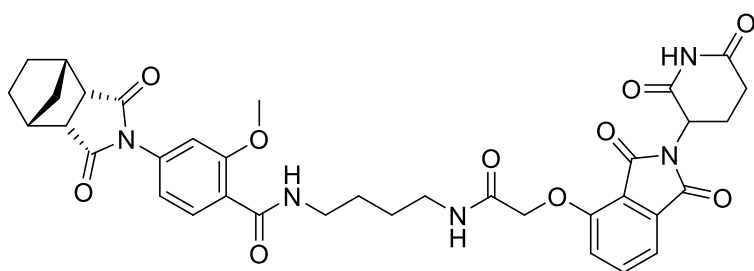
19c: 4-((3aR,4R,7S,7aS)-1,3-Dioxooctahydro-2H-4,7-methanoisindol-2-yl)-N-(4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamido)butyl)-2-fluorobenzamide



Reaction repeated as general procedure C using perfluorophenyl 4-((3aR,4R,7S,7aS)-1,3-dioxooctahydro-2H-4,7-methanoisindol-2-yl)-2-fluorobenzoate (**6j**) gave the title compound (**19c**) as a yellow oil (16 mg, 22%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.36 (m, 1H, Ar-H), 8.06-8.02 (m, 2H, Ar-H), 7.54-7.53 (m, 1H, Ar-H), 7.48 (m, 1H, Ar-H), 7.39 (d, 1H, *J* = 8.0 Hz, Ar-H), 4.46 (m, 2H, CH₂), 4.41-4.38 (m, 1H, C-H), 3.69-3.68 (m, 2H, ((CH)₂C=O)C-H), 3.55-3.52 (m, 2H,

CH₂), 3.34-3.32 (m, 2H, CH₂), 2.68-2.33 (m, 6H, (CH₂)₂, CH₂), 1.76-1.75 (m, 2H, ((CH)₂CH₂)C-H), 1.54-1.31 (m, 8H, (CH₂)₂, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.8, 172.5, 170.9, 170.5, 168.4, 167.6, 154.6, 154.5, 141.6, 140.9, 136.6, 136.3, 130.5, 121.5, 120.2, 118.3, 116.6, 109.1, 68.3, 64.6, 41.5, 41.3, 40.3, 38.3, 35.6, 30.5, 28.0, 26.2, 22.2. HRMS (ES+) *m/z* calculated for C₃₅H₃₅¹⁹FN₅O₉: 688.2419. Found [M+H]⁺ 688.2423 (Diff 0.58 ppm). Purity HPLC 99.6%, R_t= 4.2 min.

19d: 4-((3*aR*,4*R*,7*S*,7*aS*)-1,3-Dioxooctahydro-2*H*-4,7-methanoisindol-2-yl)-*N*-(4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)oxy)acetamido)butyl)-2-methoxybenzamide



Reaction repeated as general procedure C using perfluorophenyl 4-((3*aR*,4*R*,7*S*,7*aS*)-1,3-dioxooctahydro-2*H*-4,7-methanoisindol-2-yl)-2-methoxybenzoate (**6k**) gave the title compound (**19d**) as a yellow oil (13 mg, 18%). ¹H NMR (400 MHz, CD₃OD-*d*4) δ 8.51-8.50 (m, 1H, Ar-H), 8.29-8.28 (m, 1H, Ar-H), 7.92 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.78-7.76 (m, 1H, Ar-H), 7.62 (d, 1H, *J*= 8.0 Hz, Ar-H), 7.12-7.10 (m, 1H, Ar-H), 4.59-4.57 (m, 1H, C-H), 4.42-4.41 (m, 2H, CH₂), 3.98 (d, 3H, *J*= 16.0 Hz, CH₃), 3.88-3.86 (m, 2H, ((CH)₂C=O)C-H), 3.61-3.59 (m, 2H, CH₂), 3.16-3.14 (m, 2H, CH₂), 2.74-2.52 (m, 6H, CH₂, (CH₂)₂), 2.44-2.41 (m, 2H, (CH₂)₂), 1.85-1.84 (m, 2H, ((CH)₂CH₂)C-H), 1.62-1.59 (m, 4H, (CH₂)₂), 1.51-1.33 (m, 4H, (CH₂)₂). ¹³C NMR (101 MHz, CD₃OD-*d*4) δ 176.3, 175.4, 169.2, 169.1, 167.5, 166.6, 159.5, 159.0, 143.6, 143.0, 133.4, 133.3, 129.6, 124.0, 119.6, 116.7, 116.2, 106.1, 68.6, 65.1, 54.3, 42.0, 41.7, 39.1, 38.6, 38.6, 30.4, 27.7, 23.4, 21.7. HRMS (ES+) *m/z* calculated for C₃₆H₃₈N₅O₁₀: 700.2618. Found [M+H]⁺ 700.2624 (Diff 0.86 ppm). Purity HPLC 99.2%, R_t= 4.2 min.

5.9 References

- 1 E. Capoccia, C. Cirillo, A. Marchetto, S. Tiberi, Y. Sawikr, M. Pesce, A. D'alessandro, C. Scuderi, G. Sarnelli, R. Cuomo, L. Steardo and G. Esposito, *Oncol. Lett.*, 2015, **9**, 2864–2870.
- 2 S. R. Penumutchu, R. H. Chou and C. Yu, *PLoS One*, 2014, **9**, 1–16.
- 3 S. R. Jenkinson, R. Barraclough, C. R. West and P. S. Rudland, *Br. J. Cancer*, 2004, **90**, 253–262.
- 4 A. R. Bresnick, *Biophys. Rev.*, 2018, **10**, 1617–1629.
- 5 Y. Tang, X. Zeng and J. Liang, *J. Chem. Educ.*, 2010, **87**, 742–746.
- 6 T. M. Ismail, D. Bennett, A. M. Platt-Higgins, M. Al-Medhity, R. Barraclough and P. S. Rudland, *Cancer Res.*, 2017, **77**, 780–789.
- 7 G. E. Winter, D. L. Buckley, J. Paulk, J. M. Roberts, A. Souza, S. Dhe-Paganon and J. E. Bradner, *Science (80-.)*, 2015, **348**, 1376–1381.
- 8 P. Li, J. Hwang, J. M. Maier, C. Zhao, D. V. Kaborda, M. D. Smith, P. J. Pellechia and K. D. Shimizu, *Cryst. Growth Des.*, 2015, **15**, 3561–3564.
- 9 B. Findlay, The Double Extraction, <https://chemtips.wordpress.com/2012/07/31/clowns-to-the-left-of-me-jokers-to-the-right/>, (Accessed May 2020).
- 10 S. Jobin, S. Vézina-Dawod, C. Herby, A. Derson and E. Biron, *Org. Lett.*, 2015, **17**, 5626–5629.
- 11 K. Brahm, J. S. Wack, S. Eckes, V. Engemann and K. Schmitz, *Biopolymers*, 2018, **110**, 1–8.
- 12 C. Caumes, T. Hjelmgaard, O. Roy, M. Reynaud, D. Servent, C. Taillefumier and S. Faure, *Medchemcomm*, 2012, **3**, 1531–1535.
- 13 A. M. Fanning, S. E. Plush and T. Gunnlaugsson, *Chem. Commun.*, 2006, **36**, 3791–3793.
- 14 A. Prandina, S. Radix, M. Le Borgne, L. P. Jordheim, Z. Bousfiha, C. Fröhlich, H. K. S. Leiros, Ø. Samuelsen, E. Frøvdold, P. Rongved and O. A. H. Åstrand, *Tetrahedron*, 2019, **75**, 1525–1540.
- 15 E. S. Fischer, K. Böhm, J. R. Lydeard, H. Yang, M. B. Stadler, S. Cavadini, J. Nagel, F. Serluca, V. Acker, G. M. Lingaraju, R. B. Tichkule, M. Schebesta, W. C. Forrester, M. Schirle, U. Hassiepen, J. Ottl, M. Hild, R. E. J. Beckwith, J. W. Harper, J. L. Jenkins and N. H. Thomä, *Nature*, 2014, **512**, 49–53.
- 16 J. Lee et al, US201301715633 A1, 2013.
- 17 J. Chen, E. Namila, C. Bai, M. Baiyin, B. Agula and Y. S. Bao, *RSC Adv.*, 2018, **8**, 25168–25176.
- 18 E. J. Ormerod and P. S. Rudland, *Vitr. Cell. Dev. Biol.*, 1985, **21**, 143–153.

- 19 P. S. Rudland, D. J. Dunnington, B. Gusterson, P. Monaghan and C. M. Hughes, *Cancer Res.*, 1984, **44**, 2089–102.
- 20 D. T. Ross, U. Scherf, M. B. Eisen, C. M. Perou, C. Rees, P. Spellman, V. Iyer, S. S. Jeffrey, M. Van de Rijn, M. Waltham, A. Pergamenschikov, J. C. F. Lee, D. Lashkari, D. Shalon, T. G. Myers, J. N. Weinstein, D. Botstein and P. O. Brown, *Nat. Genet.*, 2000, **24**, 227–235.
- 21 R. Cailleau, M. Olivé and Q. V. J. Cruciger, *In Vitro*, 1978, **14**, 911–915.
- 22 E. Bazigou, H. Aplitz, J. Johansson, C. E. Lorén, E. M. A. Hirst, P. L. Chen, R. H. Palmer and I. Salecker, *Cell*, 2007, **128**, 961–975.
- 23 G. Halder, P. Callaerts, S. Flister, U. Walldorf, U. Kloter and W. J. Gehring, *Development*, 1998, **125**, 2181–2191.
- 24 B. Hauck, W. J. Gehring and U. Walldorf, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 564–569.
- 25 M. P. Belvin and K. V. Anderson, *Annu. Rev. Cell Dev. Biol.*, 1996, **12**, 393–416.
- 26 D. Bennett, E. Lyulcheva and N. Cobbe, *Ocul. Oncol. Pathol.*, 2015, **1**, 190–199.
- 27 M. Uhlirova and D. Bohmann, *EMBO J.*, 2006, **25**, 5294–5304.
- 28 A. M. Brumby and H. E. Richardson, *EMBO J.*, 2003, **22**, 5769–5779.
- 29 Cancer Research UK, <https://www.cancerresearchuk.org/about-cancer/what-is-cancer/stages-of-cancer>, (Accessed May 2020).
- 30 C. R. Justus, N. Leffler, M. Ruiz-Echevarria and L. V. Yang, *J. Vis. Exp.*, 2014, **88**, 1–8.
- 31 J. Pijuan, C. Barceló, D. F. Moreno, O. Maiques, P. Sisó, R. M. Marti, A. Macià and A. Panosa, *Front. Cell Dev. Biol.*, 2019, **7**, 1–16.
- 32 K. Kaliyappan, M. Palanisamy, J. Duraiyan and R. Govindarajan, *J. Pharm. Bioallied Sci.*, 2012, **4**, 307–309.
- 33 X. Huang and V. M. Dixit, *Cell Res.*, 2016, **26**, 484–498.
- 34 S. An and L. Fu, *EBioMedicine*, 2018, **36**, 553–562.
- 35 K. Moreau, M. Coen, A. X. Zhang, F. Pachi, M. P. Castaldi, G. Dahl, H. Boyd, C. Scott and P. Newham, *Br. J. Pharmacol.*, 2020, **177**, 1709–1718.
- 36 A. Tomberg, J. Pottel, Z. Liu, P. Labute and N. Moitessier, *Angew. Chemie Int. Ed.*, 2015, **54**, 13743–13747.
- 37 R. L. Melnick, *Ann. N. Y. Acad. Sci.*, 2006, **982**, 177–189.
- 38 A. W. Munro, K. J. McLean, J. L. Grant and T. M. Makris, *Biochem. Soc. Trans.*, 2018, **46**, 183–196.
- 39 T. A. Fjellstedt, R. H. Allen, B. K. Duncan and W. B. Jakoby, *J. Biol. Chem.*, 1973, **248**, 3702–3707.
- 40 C. J. Thibodeaux, W. Chang and H. Liu, *Chem. Rev.*, 2012, **112**, 1681–1709.

- 41 K. Cyrus, M. Wehenkel, E.-Y. Choi, H.-J. Han, H. Lee, H. Swanson and K.-B. Kim, *Mol. Biosyst.*, 2011, **7**, 359–364.
- 42 B. W. Weber, S. W. Kimani, A. Varsani, D. A. Cowan, R. Hunter, G. A. Venter, J. C. Gumbart and B. T. Sewell, *J. Biol. Chem.*, 2013, **288**, 28514–28523.
- 43 M. Sharma, N. N. Sharma and T. C. Bhalla, *Rev. Environ. Sci. Bio/Technology*, 2009, **8**, 343–366.
- 44 Y. Wang, X. Jiang, F. Feng, W. Liu and H. Sun, *Acta Pharm. Sin. B*, 2020, **10**, 207–238.
- 45 X. Sun, H. Gao, Y. Yang, M. He, Y. Wu, Y. Song, Y. Tong and Y. Rao, *Signal Transduct. Target. Ther.*, 2019, **4**, 1–33.
- 46 L.-W. Xia, M.-Y. Ba, W. Liu, W. Cheng, C.-P. Hu, Q. Zhao, Y.-F. Yao, M.-R. Sun and Y.-T. Duan, *Future Med. Chem.*, 2019, **11**, 2919–2973.
- 47 M. Girardini, C. Maniaci, S. J. Hughes, A. Testa and A. Ciulli, *Bioorg. Med. Chem.*, 2019, **27**, 2466–2479.
- 48 C. Steinebach, H. Kehm, S. Lindner, L. P. Vu, S. Köpff, Á. López Mármol, C. Weiler, K. G. Wagner, M. Reichenzeller, J. Krönke and M. Gütschow, *Chem. Commun.*, 2019, **55**, 1821–1824.
- 49 B. E. Smith, S. L. Wang, S. Jaime-Figueroa, A. Harbin, J. Wang, B. D. Hamman and C. M. Crews, *Nat. Commun.*, 2019, **10**, 131.
- 50 Z. Li, P. Hao, L. Li, C. Y. J. Tan, X. Cheng, G. Y. J. Chen, S. K. Sze, H.-M. Shen and S. Q. Yao, *Angew. Chemie Int. Ed.*, 2013, **52**, 8551–8556.
- 51 S. Specklin and J. Cossy, *J. Org. Chem.*, 2015, **80**, 3302–3308.
- 52 H. Sun, A. Horatscheck, V. Martos, M. Bartetzko, U. Uhrig, D. Lentz, P. Schmieder and M. Nazaré, *Angew. Chemie Int. Ed.*, 2017, **56**, 6454–6458.
- 53 E. C. Vik, P. Li, J. M. Maier, D. O. Madukwe, V. A. Rassolov, P. J. Pellechia, E. Masson and K. D. Shimizu, *Chem. Sci.*, 2020, **11**, 7487–7494.
- 54 Jolanta Obniska, R. Lesyk, D. Atamanyuk and K. Kamiński, *Acta Pol Pharm*, 2005, **62**, 213–219.
- 55 L. I. Kas'yan, O. V. Krishchik, A. O. Kas'yan and I. N. Tarabara, *Russ. J. Org. Chem.*, 2004, **40**, 1832–1834.
- 56 M. E. Salvati, WO2002067939A1, 2002, 331.
- 57 J. Vargas, A. Martínez, A. A. Santiago, M. A. Tlenkopatchev, R. Gaviño and M. Aguilar-Vega, *J. Fluor. Chem.*, 2009, **130**, 162–168.