
Investigations of the Synthesis and Properties of Bromodomain and Extra-Terminal Domain Proteolysis Targeting Chimeras

John C. Priestner

February 2020



Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to GSK in accordance with the author's contract of engagement with GSK under the terms of the United Kingdom Copyright Acts. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed:

Date:

Acknowledgements

Firstly, I would like to thank my two supervisors, Dr. Darren Poole from GlaxoSmithKline and Prof. John Murphy from the University of Strathclyde. I am incredibly grateful for the guidance and support they have provided me with throughout the duration of my PhD.

I would also like to thank Prof. Harry Kelly and Prof. William Kerr for the organisation of the GSK/Strathclyde collaborative PhD program. This unique program has allowed me to gain invaluable industrial experience. Their support has enabled my development and afforded me the opportunity to travel and present my work at numerous conferences.

I am grateful to my colleagues, both past and present, within the Flexible Discovery Unit at GSK, where I spent my entire PhD. I have been fortunate to work with many talented individuals from whom I have benefited greatly, this experience would not have been the same without their friendship.

A special thanks goes to Stephen Richards for his support with NMR experiments and assignments. Without his expertise and enthusiasm this work would not have been possible. I would like to thank Andrew Baxter, Chris Luscombe and Dr. Carol Mulrooney for their input and guidance in the interpretation of complex data sets. From the Protein Degradation Unit, thanks go to Dr. Afjal Miah for his thought-provoking conversations, and Dr. Peter Stacey and Diana Klimaszewska for their help with running the HiBiT degradation assay. I would also like to thank Shenaz Bunally and Iain Reid for their help in generating the assay data disclosed in this thesis.

I must also thank my mother, father and sister Beverley for their unending support. Their encouragement and understanding have always allowed me to confidently pursue my dreams. Finally, I would like to thank my darling Chloë. Your love, patience and brilliance have kept me going through the PhD and provided meaning to my endeavours. I am sure that this would not have been the same experience without you.

I dedicate this thesis to my late mother and Chloë

Per aspera ad astra

Abstract

Hijacking a cell's ubiquitin proteasome system using PROTAC technology has the potential to be a powerful new therapeutic strategy. Some of the main advantages of this technology include the potential for PROTACs to be catalytic in activity and elicit long-lasting therapeutic effects, offering the possibility of low clinical doses and infrequent dosing regimens. However, PROTACs are large molecules and generally exhibit low cell permeability, low aqueous solubility and ultimately low bioavailability. Due to the specific defined structures of the target protein and E3 ligase binding moieties, the linker provides a facile handle to modulate physicochemical properties.

A linker library containing 66 diverse linkers that provide broad coverage of desirable physicochemical space was developed. From this library, a series of 63 BET PROTACs was rapidly synthesised in a plate-based format using the high-throughput, one-pot protocols developed for acid-ester and diacid linkers.

The linker functionality had a profound effect on the degradation profile of the PROTACs, furnishing a 4 log unit range of pDC_{50} values, with a maximum of 8.99 achieved by the PROTAC containing the 2,6-disubstituted pyridyl linker. Changes in the linker functionality of the PROTAC series furnished an approximate 3 log unit range of ChromlogD (2.95-5.55) and PAC (-0.66-2.08) values, as well as a range of CAD (1-516 $\mu\text{g/mL}$) and FaSSIF (2-1000 $\mu\text{g/mL}$) solubilities, that cover the spectrum from limited to high permeability, and negligible to high solubility.

PCA/PLS analysis of the series identified effective lipophilicity as the key property driving the cell permeability and Brd4 degradation potency of the PROTACs, with CAD solubility exhibiting a negative correlation. The three-dimensional molecular shape of the PROTACs that manifest from the linker was found to be pertinent in determining the physicochemical properties and degradation profile of the PROTACs. The most potent PROTACs occupy elongated discoid shapes, common to the 1,3-disubstituted 5- and 6-membered aromatic linkers that exhibit the optimum linker length of four bonds. These shapes are also common to PROTACs exhibiting high

permeability and low CAD solubility. Additionally, the position of functional groups in the linker was found to have a significant effect on these properties.

This work highlights that the guidelines of generic descriptors (TPSA, RBC, Fsp³, ARc, HBA and HBD count) need to be used with caution when analysing the physicochemical properties of PROTACs, due to the increased complexity of the relationships between them as a result of the effects of molecular shape and conformational folding of these large molecules. It also emphasises the need to screen linkers with a range of geometric properties, and not to discount linkers that will furnish PROTACs with higher ChromlogD values than would be desirable for a classical small molecule drug.

Table of Contents

Declaration	i
Acknowledgements	ii
Abstract	iii
Table of Contents	v
Abbreviations	ix
1: Introduction	1
1.1 The Classical Drug Discovery Paradigm	2
1.2 The Ubiquitin Proteasome System (UPS)	4
1.3 Drug Discovery Targeting the Ubiquitin Proteasome System	8
1.3.1 26 S Proteasome Inhibition	9
1.3.2 E1 Enzyme Inhibition	11
1.3.3 E2 Enzyme Inhibition	12
1.3.4 E3 Ligase Inhibition	13
1.3.4.1 von Hippel–Lindau (VHL) inhibition	13
1.3.4.2 Cereblon (CRBN) Inhibition	15
1.4 Proteolysis Targeting Chimeras (PROTACs)	18
1.5 Bromodomain and Extra-Terminal Domain (BET) Proteins	22
1.5.1 Deoxyribonucleic Acid (DNA) Structure and Function	22
1.5.2 Epigenetics	24
1.5.3 Bromodomain-Containing Proteins (BCPs).....	26
1.5.4 Small Molecule BET Inhibitors	28
1.6 Small Molecule BET PROTACs	31
2: Project Aims	36
3: Results and Discussion – <i>Investigating the Synthesis of BET PROTACs</i>	38
3.1 Introduction	39
3.1.1 Common Synthetic Methods	39
3.1.2 Decarboxylative Cross-Coupling	41
3.2 Synthetic Strategy	45
3.3 Decarboxylative Cross-Coupling Methodologies	48

3.4 Synthesis of Boronic Acid Protein Binder	52
3.5 Synthesis of PROTACs by Decarboxylative Cross-Coupling	59
3.5.1 Decarboxylative Cross-Coupling with a Boronic Acid Protein Binder	59
3.5.2 Decarboxylative Cross-Coupling with a Bromide Protein Binder	73
3.6 Synthesis of PROTACs by Amide Coupling	77
3.7 High-Throughput Synthesis of PROTACs by Amide Coupling	83
3.7.1 Acid-Ester Linkers	83
3.7.2 Diacid Linkers	94
3.8 Conclusions	98
4: Results and Discussion – Investigating the Effects of Linker Functionality on the Physicochemical Properties and Degradation Profiles of BET PROTACs	101
4.1 Introduction	102
4.1.1 Physicochemical Property Descriptors	102
4.1.2 Physicochemical Property Guidelines	107
4.2 PROTAC Linker Library	109
4.3 PROTAC Series	120
4.3.1 Brd4 Binding Affinity	121
4.3.2 Brd4 Degradation	126
4.4 Physicochemical Profile of the PROTAC Series	137
4.5 Binding Affinity and Degradation Profile of the PROTAC Series	149
4.6 PCA/PLS Analysis of the PROTAC Series	166
4.7 Geometric Properties of the PROTAC Series.....	176
4.7.1 Principal Moments of Inertia (PMI)	180
4.8 Conclusions	191
4.9 Future Work	194
5: Experimental	196
5.1 General Information	197
5.2 Analytical Techniques and Equipment	197
5.2.1 Column Chromatography	197
5.2.2 Glovebox	197
5.2.3 High-Resolution Mass Spectroscopy	197

5.2.4 Infra-red (IR) Spectroscopy	198
5.2.5 Liquid Chromatography Mass Spectrometry (LCMS)	198
5.2.5.1 Method A - Formic	198
5.2.5.1 Method B - HpH	199
5.2.6 Mass Directed Auto Purification (MDAP)	199
5.2.7 Melting Point (m.p.)	199
5.2.8 Microwave Reactor	199
5.2.9 Nuclear Magnetic Resonance Spectroscopy (NMR)	200
5.2.10 Specific Optical Rotation	200
5.2.11 Thin Layer Chromatography (TLC)	200
5.3 General Procedures	201
5.4 Experimental Details	204
5.5 Biological assays and Physicochemical Measurements	314
5.5.1 Compound Preparation	314
5.5.2 Artificial Membrane Permeability (AMP)	314
5.5.3 Brd4 BD1/BD2 Time-Resolved Förster Resonance Energy Transfer (FRET)	314
5.5.4 Brd4 HiBiT Degradation	316
5.5.5 Brd4 Human Whole Blood MCP-1	317
5.5.6 Cell Concentration Permeability (PAC)	319
5.5.7 Charged Aerosol Detector (CAD) Solubility	321
5.5.8 ChromlogD	321
5.5.9 Fasted-State Simulated Intestinal Fluid (FaSSIF) Solubility	321
5.5.10 Human Serum Albumin (HSA) Binding	322
5.5.11 Immobilised Artificial Membrane (IAM) Binding	322
5.6 Computational Methods	323
5.6.1 Docking Models	323
5.6.2 PCA/PLS analysis	323
5.6.3 Principal Moments of Inertia (PMI) Plots	324
6: References	326

7: Appendices	340
7.1 NOE Experiments	341
7.1.1 5-(6-Bromo-1-(1,3-dimethoxypropan-2-yl)-1 <i>H</i> -benzo[<i>d</i>]imidazol-2-yl)-6-isopropyl-1,3-dimethyl-3,6-dihydropyridin-2(1 <i>H</i>)-one 121	341
7.1.2 6-Butyl-5-(1-(1,3-dimethoxypropan-2-yl)-1 <i>H</i> -benzo[<i>d</i>]imidazol-2-yl)-1,3-dimethyl-3,6-dihydropyridin-2(1 <i>H</i>)-one 122	343
7.2 X-ray Cocystal Structures	345
7.2.1 (<i>R</i>)- <i>N</i> 4-(1-(1,3-dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1 <i>H</i> -benzo[<i>d</i>]imidazol-6-yl)- <i>N</i> 1-((<i>S</i>)-1-((2 <i>S</i> ,4 <i>R</i>)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-2-methylsuccinamide 168 in human Brd4 BD1 protein	345
7.3 Compound Numbering	347

Abbreviations

AB-MPS	AbbVie Multiparameter Score
AGP	α -1-Acid-Glycoprotein
AMP	Artificial Membrane Permeability
aq.	Aqueous
ARc	Aromatic Ring Count
ASO	Antisense oligonucleotide
ATP	Adenosine triphosphate
BCP	Bromodomain-Containing Protein
BD1/2	Bromodomain 1/2
BET	Bromodomain and Extra-Terminal domain
BippyPhos	5-(Di- <i>tert</i> -butylphosphino)-1', 3', 5'-triphenyl-1' <i>H</i> -[1,4']bipyrazole
Boc	<i>tert</i> -Butyloxycarbonyl
br	Broad
Brd4	Bromodomain-containing Protein 4
bRo5	Beyond the rule of five
Pin	Pinacolato
c	Calculated
CAD Solubility	Charged Aerosol Distribution Solubility (kinetic)
cChromlogD _{HAC}	Calculated Heavy Atom Count ChromlogD
CDK	Cyclin Dependent Kinase
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate
CHI	Chromatographic Hydrophobicity Index
ChromlogD	Chromatographically measured distribution coefficient between acetonitrile and water at pH 7.4
CLND	Chemiluminescent Nitrogen Detection
CMR	Calculated Molar Refractivity
cm ⁻¹	Wavenumbers
cPFI	Calculated Property Forecast Index
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRBN	Cereblon
CV	Column Volume
Da	Dalton
DCC	<i>N,N</i> -Dicyclohexylcarbodiimide
DCM	Dichloromethane
DEL	DNA-Encoded Library
DHTC	Discovery High-Throughput Chemistry
DIBAL-H	Diisobutylaluminium hydride
DIC	<i>N,N'</i> -Diisopropylcarbodiimide
DIPEA	<i>N,N</i> -Diisopropylethylamine
DMA	Dimethylacetamide
DMAP	4-Dimethylaminopyridine
DMEM	Dulbecco's Modified Eagle's Medium

DMF	Dimethylformamide
DMPU	<i>N,N'</i> -Dimethylpropyleneurea
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dtbbpy	4,4'- <i>Di-tert</i> -butyl-2,2'-bipyridine
DTT	1,4-Dithiothreitol
DUB	Deubiquitinase
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EG	Ethylene Glycol
ELT	Encoded Library Technology
equiv	Equivalent(s)
ESI	Electrospray ionisation
Et	Ethyl
ET	Extra-Terminal domain
Eu	Anti-6His europium chelate
e-state	Electrotopological-state index
FaSSIF Solubility	Fasted-State Simulated Intestinal Fluid Solubility (thermodynamic)
FBS	Fetal Bovine Serum
FDA	Food and Drug Administration
FP	Full PROTAC
FRET	Förster Resonance Energy Transfer
Fsp ³	Fraction sp ³
GBBR	Groebke–Blackburn–Bienaymé reaction
h	Hour(s)
HAC	Heavy (non-hydrogen) Atom Count
HAT	Histone Acetyltransferase
HATU	1-[Bis(dimethylamino)methylene]-1 <i>H</i> -1,2,3-triazolo[4,5- <i>b</i>]pyridinium 3-oxid hexafluorophosphate
HBA	Hydrogen Bond Acceptor
HBD	Hydrogen Bond Donor
HDAC	Histone Deacetylase
HECT	Homologous to the E6-AP Carboxyl Terminus
HEK	Human Embryonic Kidney
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HIF-1 α	Hypoxia Inducible Factor 1 α
His	Histidine
HpH	pH > 7
HPLC	High-Performance Liquid Chromatography
HRMS	High-Resolution Mass Spectroscopy
HSA	Human Serum Albumin
HWB	Human Whole Blood
Hyp	4-Hydroxyproline
IAM	Immobilised Artificial Membrane
IEDDA	Inverse electron demand Diels–Alder
IMiD	Immunomodulatory drugs
IPA	Isopropyl alcohol

<i>i</i> Pr	<i>iso</i> -Propyl
IR	Infra-red
IκB	Kappa-light-chain-enhancer of activated B cells
<i>J</i>	Coupling constant
JosiPhos	(<i>R</i>)-1-{(S _F)-2-[Bis[3,5-bis(trifluoromethyl)phenyl]phosphino]ferrocenyl}ethyldicyclohexylphosphine
<i>K</i> _d	Equilibrium dissociation constant
<i>k</i> _{inact}	Rate constant of inactivation
<i>k</i> _{off}	Rate constant of dissociation
<i>k</i> _{on}	Rate constant of binding
Lit	Literature value
logP	Partition coefficient between 1-octanol and water
LPS	Lipopolysaccharide
LT1/2	Linker Truncate 1/2
Lys	Lysine
M	Molar
Max % deg	Maximum percentage degradation
MCP-1	Monocyte Chemoattractant Protein-1
MDAP	Mass Directed Auto Purification
MDCK	Madin-Darby Canine Kidney
Me	Methyl
MeOH	Methanol
min	Minute(s)
MOA	Mode of action
mol%	Molar percent
MPNN	Message Passing Neural Networks
mRNA	Messenger ribonucleic acid
<i>M</i> _w	Molecular Weight
MW	Microwave
m.p.	Melting Point
<i>m/z</i>	Mass-to-charge ratio
NA	Not available
NF-κB	Nuclear Factor κB
NMC	NUT midline carcinoma
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser effect
npr	Normalised Principal Moments of Inertia Ratio
NUT	Nuclear protein in testis
OAc	Acetate
PBMC	Peripheral Blood Mononuclear Cells
PC1/2	Principle Component 1/2
PCA	Principal Component Analysis
pDC ₅₀	-log (Concentration at which 50% of the protein is degraded)
PEG	Polyethylene glycol
Pet. Ether	Petroleum ether
Ph	Phenyl

Phe	Phenylalanine
PhMe	Toluene
pIC ₅₀	-log (Concentration at which 50% of the protein is inhibited)
pK _a	-log (Acid dissociation constant)
PLS	Projections to latent structures by means of partial least squares
PMI	Principal Moments of Inertia
PPI	Protein–Protein Interaction
ppm	Parts per million
Pro	Proline
PROTAC	Proteolysis Targeting Chimera
PSA	Polar Surface Area
PTM	Post-Translation Modification
ΔC	Cell Concentration Permeability
QSAR	Quantitative Structure-Activity Relationship
RAE	Redox-Active Ester
RBc	Rotatable Bond Count
rDC ₅₀	-log(Concentration at which 50% of the maximum protein degradation is achieved)
RING	Really Interesting New Gene
RMS	Root Mean Square
RMSD	Root Mean Square Deviation
RNA	Ribonucleic acid
rpm	Revolutions per minute
Rpn	Regulatory particle of non-ATPase
RT	Ambient Temperature
s	Second(s)
SAR	Structure-Activity Relationship
sat.	Saturated
siRNA	Small interfering RNA
TBME	<i>tert</i> -Butyl methyl ether
^t Bu	<i>tert</i> -Butyl
THF	Tetrahydrofuran
Thr	Threonine
TLC	Thin Layer Chromatography
TPSA	Topological Polar Surface Area
tRNA	Transfer ribonucleic acid
Trp	Tryptophan
Ub	Ubiquitin
UPS	Ubiquitin Proteasome System
UV	Ultra-violet
VHL	von Hippel–Lindau
VIP	Variable Influence on Projection
vs	Versus
WDI	World Drug Index
δ	Chemical shift
[α] _D	Specific optical rotation

1: Introduction

1.1 The Classical Drug Discovery Paradigm

Historically, drug discovery has aimed to elicit a beneficial disease response by the development of therapeutics that inhibit a specific biological process. Predominantly, this is achieved through inhibition of intracellular targets with well-defined active sites to which synthetic small molecules can selectively bind, competitively inhibiting their disease-inducing functions. However, this occupancy-driven mode of action (MOA) exhibits some inherent limitations.¹

Firstly, high levels of occupancy (> 90%) are required to achieve inhibitory efficacy.² As the drug is metabolised and cleared, occupancy of the active site is reduced, biological function is restored and the beneficial disease response diminishes. This requires frequent, high dosing regimens, often leading to undesired side-effects due to off-target binding associated with higher drug concentrations.³

These high levels of target occupancy are commonly achieved using either reversible or irreversible inhibitors. Reversible inhibitors extend target occupancy through high complementarity with the target active site, inducing desirable binding interactions, shifting the equilibrium towards adduct formation, **Figure 1**. Judiciously designing small molecules that can selectively achieve this is challenging and requires consideration of properties that affect both k_{on} and k_{off} . Irreversible inhibitors form a permanent adduct with the target through the formation of a covalent bond, **Figure 1**. This extends the effects of target occupancy, as target function is only recovered after *de novo* cellular resynthesis of the target. This can extend the perceived occupancy of the inhibitor, resulting in lower doses.^{4,5} However, these inhibitors can suffer from idiosyncratic toxicity, where the covalent target-inhibitor adduct is recognised as ‘foreign’ and induces an immune response.^{6,7} Additionally, high target selectivity is required to reduce the toxicity issues associated with non-specific covalent binding.

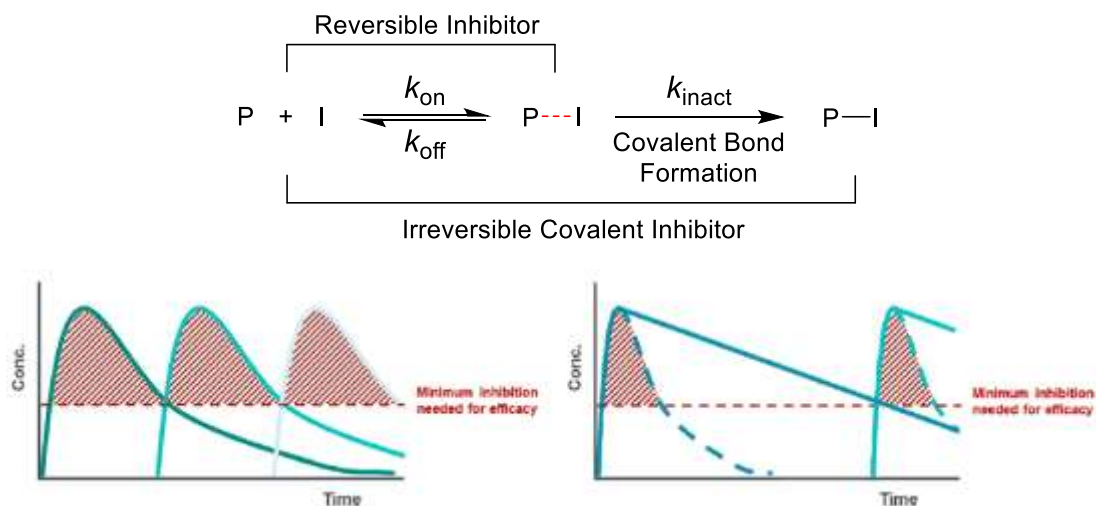


Figure 1: (Top) Comparison of the binding of reversible and irreversible inhibitors (I) with the target protein (P). **(Bottom)** Graph of cellular drug concentration vs time for reversible inhibitors **(Left)** and irreversible covalent inhibitor **(Right)**.⁸ Reprinted (adapted) with permission from Barf and co-workers.⁸ Copyright 2012 American Chemical Society.

Secondly, the occupancy-driven MOA requires targets with suitable binding sites for natural substrates and co-factors that can be imitated by synthetic small molecules, referred to as ‘druggable’ targets. Therefore, targets with poorly tractable active sites, deemed ‘undruggable’, as well as targets with secondary scaffolding functions mediated through distal sites are not amenable to this approach. This limits the targets of the occupancy-driven MOA to only a small portion of the human disease-modifying proteome.⁹

Finally, the development of resistance to inhibition/occupancy-driven therapeutics occurs in many disease indications such as cancer and bacterial infections.^{10,11} This occurs through evolutionary changes to the protein-encoding genes, changing the structure of the protein and ultimately the binding affinity of inhibitors.¹²

New approaches are being developed that overcome the inherent limitations of the occupancy-driven MOA, such as complete removal of the target protein from the cell. Through this approach, recovery of target function is dependent on *de novo* cellular resynthesis, secondary scaffolding functions are inhibited, and inhibition resistance negated.

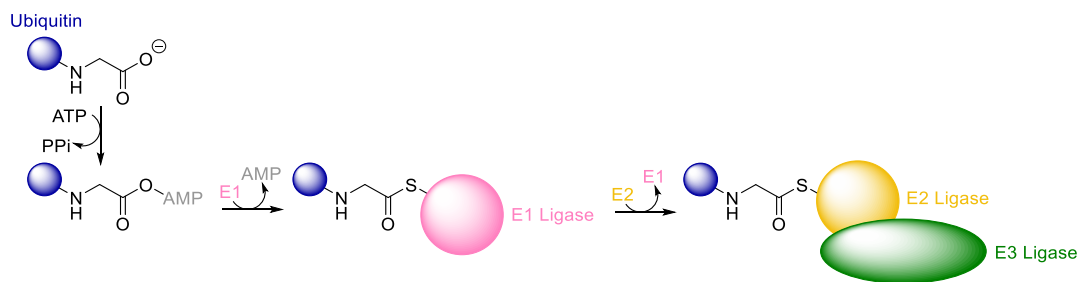
Modulation of intracellular target protein levels can be achieved by targeting *de novo* synthesis of the target protein, through genome engineering, using CRISPR-Cas9 technology, as well as gene silencing, using antisense oligonucleotides (ASOs) and small interfering RNA (siRNA).¹³⁻¹⁵ Although highly effective preclinically, these technologies most notably suffer from issues with cellular delivery, stability, biodistribution and selectivity.^{16,17}

An exciting new approach uses small molecules to hijack the ubiquitin proteasome system (UPS), the primary endogenous pathway used by eukaryotic cells to regulate intracellular protein levels.^{18,19} This approach has the potential to benefit from the administrative and pharmacological profile (physicochemical and pharmacokinetic) of small molecules, whilst avoiding the limitations of an occupancy-driven MOA. This approach will be detailed and further investigated in this thesis.

1.2 The Ubiquitin Proteasome System (UPS)

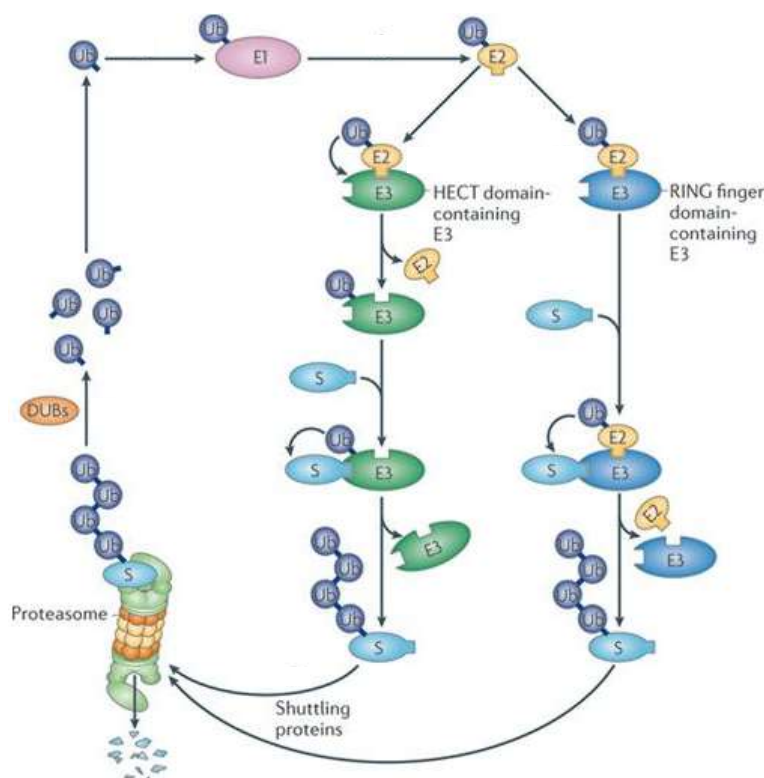
As part of cellular homeostasis, intracellular proteins are continually degraded to their constituent amino acids and new proteins resynthesised. The main pathway for degradation of cytosolic and nuclear proteins is *via* the UPS, however, some protein aggregates and damaged organelles are degraded by the autophagy-lysosomal pathway.²⁰⁻²²

Proteins are marked for degradation *via* the UPS by post-translation modification (PTM) with ubiquitin, a 8.5 kDa protein consisting of 76 amino acids, **Scheme 1**.^{23,24} First, the *C*-terminus glycine of ubiquitin is activated by the formation of a thioester linkage with an active site cysteine residue of the E1 ubiquitin-activating enzyme, in an adenosine triphosphate (ATP)-dependent reaction.²⁵ Ubiquitin is then *trans*-thiolated to an active site cysteine residue of the E2 ubiquitin-conjugating enzyme. The E3 ubiquitin ligase forms a ternary complex with the E2 enzyme and target protein, bringing them into sufficiently close proximity to achieve ubiquitin transfer onto a surface-exposed lysine residue of the target protein.²⁶



Scheme 1: Activation of ubiquitin for the post-translation modification of proteins for degradation.

The mechanisms of ubiquitin transfer can be either direct or indirect, depending on the type of E3 ligase that forms the ternary complex, **Scheme 2**.²⁷ Direct ubiquitin transfer from E2 enzymes to the target protein is achieved by the Really Interesting New Gene (RING) finger domain type E3 ligases. Indirect ubiquitin transfer occurs by *trans*-thiolation of ubiquitin to the Homologous to the E6-AP Carboxyl Terminus (HECT) domain type E3 ligases. Ubiquitination of the target protein is then achieved from the HECT E3 ligase.²⁸



Scheme 2: RING and HECT E3 ligase ubiquitination mechanisms. Ub = ubiquitin, S = substrate and DUBs = deubiquitinases.²⁷ Reprinted (adapted) with permission from Weissman and co-workers.²⁷ Copyright 2011 Springer Nature.

Further PTMs of the target protein with ubiquitin can be achieved by the described mechanism, resulting in monoubiquitination of multiple surface-exposed lysine residues of the target protein or polyubiquitination of one of the seven available lysine residues (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48 and Lys63) of ubiquitin bound to the target protein, **Figure 2**. Ubiquitination is a reversible process and cleavage of ubiquitin is performed by a series of proteases known as deubiquitinases (DUBs). This results in a myriad of permutations of ubiquitin PTMs that have been implicated in a number of other cellular processes, such as: signal transduction, enzyme activation and endocytosis.²⁹⁻³¹

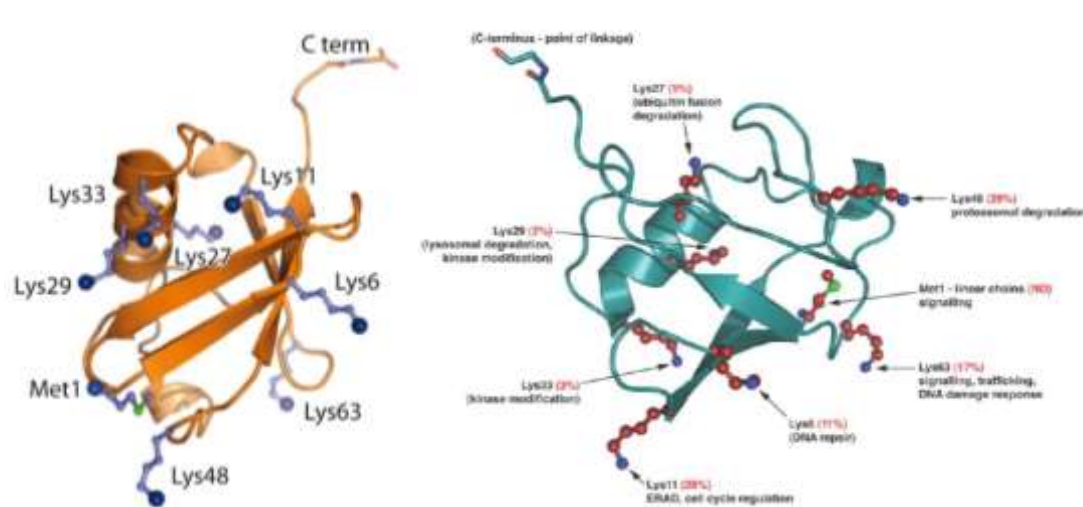


Figure 2: (Left) Structure of ubiquitin showing the seven lysine residues. Blue spheres indicate amino groups used in ubiquitin chain formation.³² Reprinted (adapted) with permission from Komander and co-workers.³² Copyright 2012 Annual Reviews. (Right) Labelling of the lysine residues with their cellular functions.³³ Reprinted (adapted) with permission from Komander and co-workers.³³ Copyright 2009 Portland Press.

Originally, it was believed that at least four ubiquitin units linked by the Lys48 residues were required for the target protein to be recognised by the proteasome for degradation.^{34,35} However, more recently, a series of shorter ubiquitin chains have been shown to be sufficient.^{36,37}

The 26 S proteasome is a 2.5 MDa complex responsible for the degradation of > 80% of proteins in mammalian cells.²⁰ It consists of two subcomplexes, the 19 S regulatory particle, responsible for recognition, unfolding and translocation of the

polyubiquitinated protein to the 20 S core catalytic particle, responsible for proteolysis, **Figure 3.**³⁸

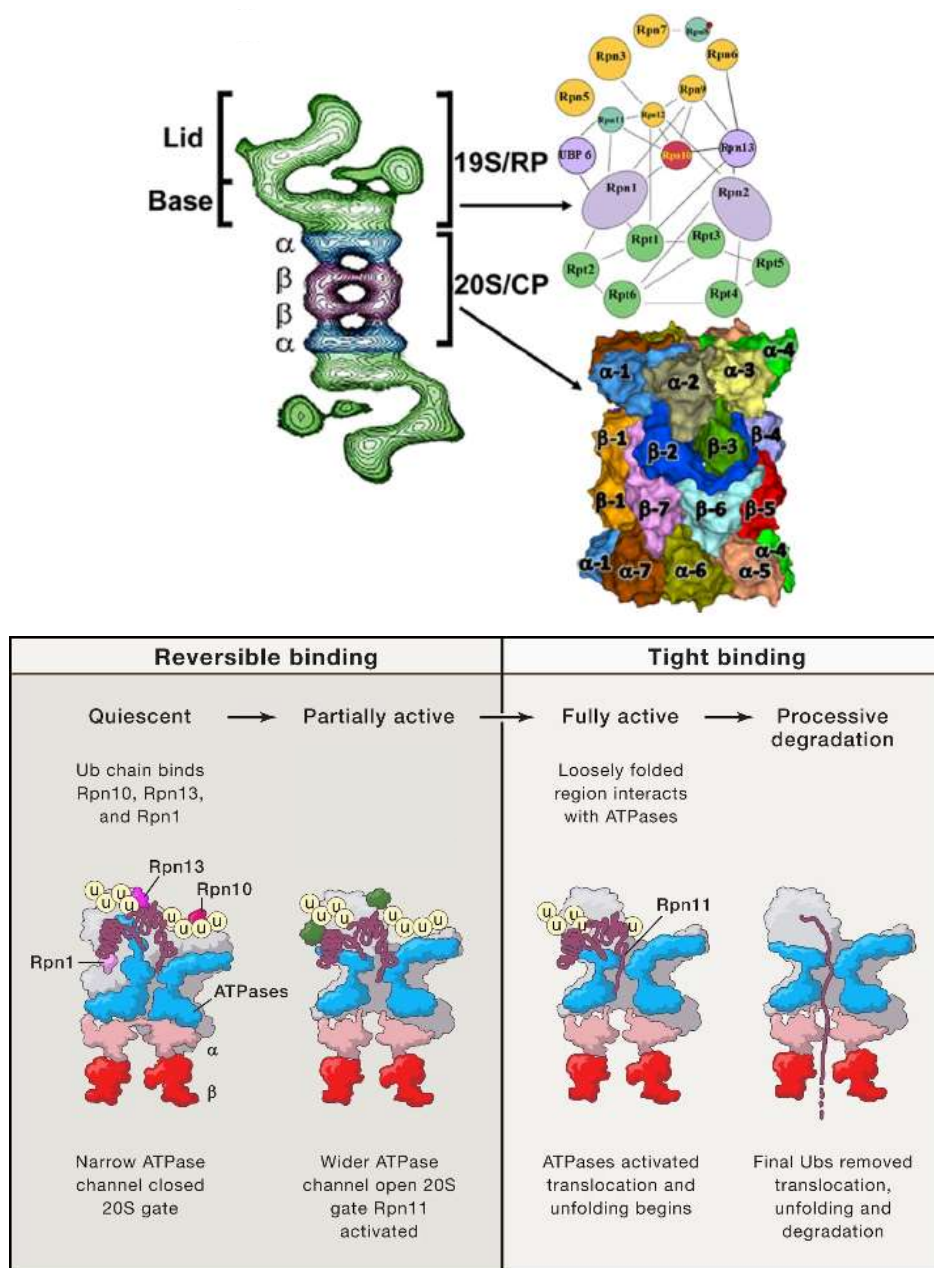
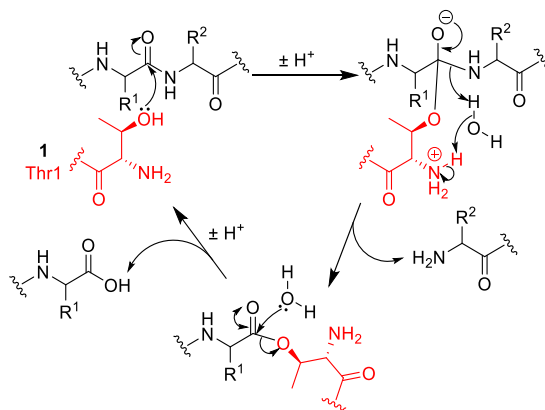


Figure 3: (Top) Schematic representation of the 26 S proteasome.¹⁹ Reprinted (adapted) with permission from Navon and co-workers.¹⁹ Copyright 2009 American Chemical Society. **(Bottom)** Mechanism of protein degradation by the 26 S proteasome.²⁰ Reprinted (adapted) with permission from Collins and co-workers.²⁰ Copyright 2017 Elsevier.

The 19 S regulatory particle is a 1 MDa complex consisting of two heteromeric complexes termed the lid, consisting of 9 subunits and the base, consisting of

10 subunits, **Figure 3**. The ubiquitin chain of the target protein is first recognised by regulatory particle of non-ATPase (Rpn)1, Rpn10 and Rpn13 of the base, resulting in unfolding and translocation of the target protein to the 20 S core catalytic particle by ATPases. At the same time, the lid DUB Rpn11 catalyses the deubiquitination of the target protein.³⁹ The 20 S core catalytic particle is a 670 kDa barrel-shaped complex comprised of 28 protein subunits that are arranged in four heptameric rings.¹⁹ The two external rings are composed of seven different α -subunits (α 1-7) that bind to the 19 S regulatory particle and selectively regulate which proteins enter the proteolytic chamber. The two internal rings, each composed of seven β -subunits, make up the proteolytic chamber. Its proteolytic activity is associated with the *N*-terminal threonine residues of the β 1, β 2 and β 5 subunits. The Thr1 residues act as nucleophiles, cleaving the protein into shorter peptide fragments *via* the mechanism described in **Scheme 3**.⁴⁰ The β 1 subunits are caspase-like, preferentially cleaving peptide bonds after acidic residues. The β 2 subunits are trypsin-like, preferentially cleaving after basic residues and the β 5 subunits are chymotrypsin-like, preferentially cleaving after hydrophobic residues.^{41,42} The cleaved peptide fragments exit the proteasome and can be used in the synthesis of new proteins.



Scheme 3: Cleavage of protein peptidic bonds by Thr1 residues to form shorter peptide fragments.^{40,43}

1.3 Drug Discovery Targeting the Ubiquitin Proteasome System

As the UPS plays an integral role in cellular homeostasis by regulating intracellular protein levels, it has become a key target for inhibition by small molecules in drug discovery. These molecules can intercept the UPS at different stages to achieve

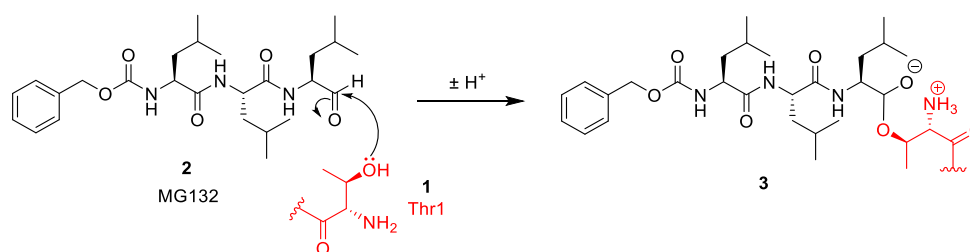
varying levels of inhibition of cellular protein homeostasis, ultimately eliciting a beneficial disease response.

1.3.1 26 S Proteasome Inhibition

The ability of 26 S proteasome inhibitors to inhibit cell proliferation and selectively induce apoptosis in proliferating cells highlighted their utility as chemotherapeutics.⁴⁴ One of the early hypotheses was that 26 S proteasome inhibition blocked the degradation of inhibitor of Kappa-light-chain-enhancer of activated B cells (I κ B), an inhibitor of Nuclear Factor κ B (NF- κ B), a protein that controls transcription of DNA, cytokine production and cell survival.^{45,46} However, it is now believed that its action is far more complicated and may involve the dysregulation of a number of cell cycle proteins, transcriptional factors, tumour inhibitors and their subsequent downstream effects.⁴⁷

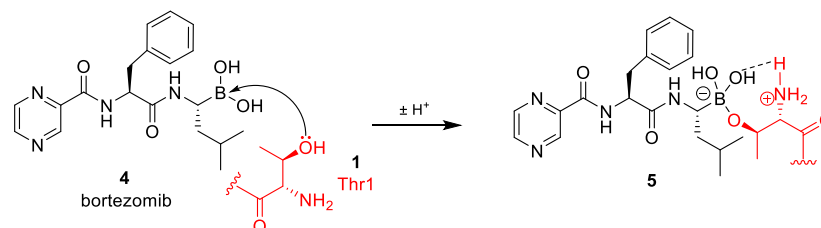
Small molecule inhibitors of the 26 S proteasome generally target the 20 S core catalytic particle by forming covalent adducts with the Thr1 residues of the β 1, β 3 and β 5 subunits of the proteolytic chamber responsible for peptide bond cleavage.

The aldehyde peptide MG132 **2** was one of the first proteasome inhibitors to be developed.^{21,48} It predominantly inhibits the chymotrypsin-like β 5 subunits, which have been shown to have the largest effect on proteasome inhibition of all β subunits.^{49,50} This inhibition is achieved through reversible hemiacetal formation (**Scheme 4**), however, MG132 **2** has a fast dissociation rate and is rapidly oxidised to the inactive carboxylic acid in cells, limiting its inhibitory effects and therapeutic utility.



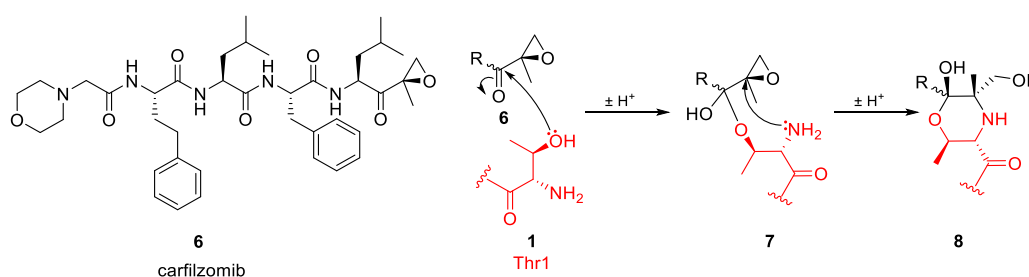
Scheme 4: Mechanism of MG132 **2** inhibition of the 26 S proteasome.⁴³

In 2005, the boronic acid peptide bortezomib **4** was granted FDA approval as a chemotherapeutic for multiple myeloma and mantle-cell lymphoma. Like MG132 **2** it predominantly targets the $\beta 5$ subunits, but it exhibits a slower dissociation rate due to the reversible formation of the boronate complex **5** and subsequent intramolecular hydrogen bond, **Scheme 5**.^{51,52} However, bortezomib **4** suffers from severe side-effects believed to be due to the off-target inhibition of serine proteases through boronate complex formation.⁵³



Scheme 5: Mechanism of bortezomib **4** inhibition of the 26 S proteasome.⁴³

In 2015, the tetrapeptide epoxyketone carfilzomib **6** was granted FDA approval as a chemotherapeutic for multiple myeloma.⁵⁴ Carfilzomib **6** is an irreversible inhibitor that selectively forms multiple covalent bonds with the Thr1 residue of the $\beta 5$ subunits, resulting in the formation of a morpholine ring, **Scheme 6**.⁵⁵ The high specificity of its covalent inhibition is believed to contribute to its high selectivity for the proteasome over off-target serine proteases, resulting in milder side-effects.^{56,57}



Scheme 6: Mechanism of carfilzomib **6** inhibition of the 26 S proteasome.⁴³

Inhibition of the 26 S proteasome has proven to be a successful way of intercepting the UPS to elicit a beneficial disease response. However, as the proteasome plays such a key role in regulating the levels of many proteins, the results of inhibition have a drastic effect on the global cellular environment, ultimately leading to cell apoptosis.

Therefore, they have found a niche as chemotherapeutics treating refractory oncology diseases, due to the severe side-effects that are associated with their effective doses.⁵⁸

1.3.2 E1 Enzyme Inhibition

The human genome codes for two ubiquitin-activating E1 enzymes, UBA1 and UBA6, which control ubiquitination of all downstream targets.⁵⁹ Inhibition of E1 enzymes prevents global ubiquitination, inhibiting degradation of all proteins that are targeted for degradation by the UPS, analogous to proteasome inhibition. However, unlike proteasome inhibition, it does not result in an accumulation of ubiquitinated proteins, instead it inhibits all pathways in which ubiquitination plays a regulatory, non-proteolytic role.⁶⁰

In 2007, the first cell-permeable E1 enzyme inhibitor PYR-41 **9** was reported.⁶¹ PYR-41 **9** is an irreversible inhibitor that covalently binds to the active site Cys632 of the E1 enzyme, abrogating its catalytic activity.⁶² It is believed that this is achieved through the nucleophilic addition of the Cys632 thiol group to either the exocyclic double-bond (arrow A) or the phenyl ring *ipso* to the pyrazolone moiety (arrow B), **Figure 4**.⁶³

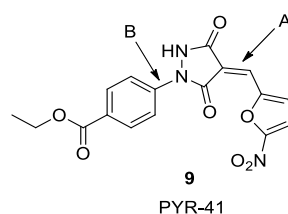


Figure 4: Sites of E1 inhibition by PYR-41 **9**.

PYR-41 universally inhibits ubiquitination and proteasomal degradation, causing an accumulation of I κ B and subsequent decrease in NF- κ B that results in cell apoptosis.⁶¹ Additionally, it was reported that an activation of p53 led to some specificity towards tumour cells. However, the overall lack of specificity of PYR-41 common to all reported E1 inhibitors that interact with numerous off-target proteins and pathways, led to PYR-41 not being progressed to clinical trials. Currently no E1 enzyme inhibitors have entered clinical trials.⁶²

1.3.3 E2 Enzyme Inhibition

To date, 38 ubiquitin-conjugating E2 enzymes have been identified in humans, making it a class of targets with potentially more specificity than E1 enzymes.⁶⁴ However, the E2 enzymes have a highly conserved active site containing the catalytic Cys93 residue. This greatly reduces the potential of selectively inhibiting a specific E2 enzyme. In 2011, CC0651 **10** was reported as the first non-competitive inhibitor of the E2 enzyme Cdc34A, binding at an allosteric site distal from the catalytic residue, **Figure 5**.⁶⁵

CC0651 **10** does not inhibit ubiquitin transfer to the E2 enzyme, instead it binds at a composite binding pocket at the periphery of the Cdc34A-ubiquitin interface, inducing conformational changes in Cdc34A to form a contiguous surface with ubiquitin, **Figure 5**.⁶⁶ This enhances complementary binding interactions between Cdc34A and ubiquitin, ultimately inhibiting transfer of ubiquitin to an E3 ligase or a surface-exposed lysine residue of a protein. Additionally, the **10**-Cdc34A-ubiquitin complex has been shown to engage the E3 ligase with slightly higher affinity than endogenous Cdc34A-ubiquitin. This allows it to also competitively inhibit the binding of ubiquitin-charged E2 enzymes, further abrogating degradation.⁶⁶

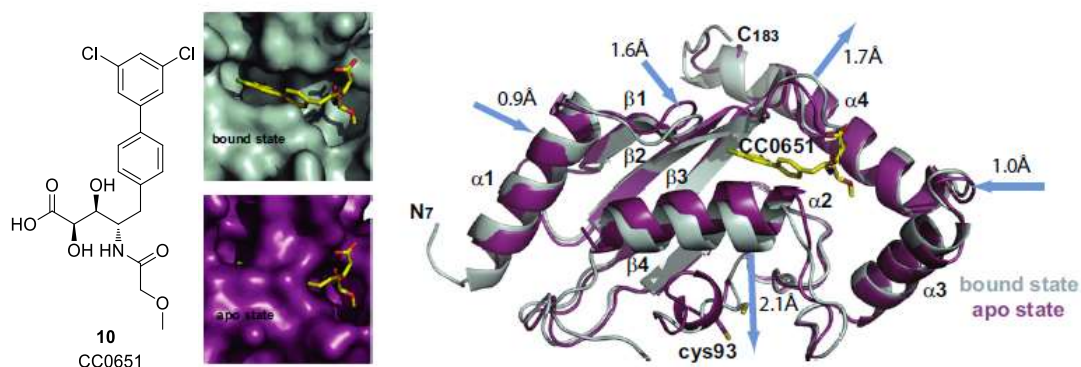


Figure 5: (Left) Structural representation of CC0651 **10**. (Middle) Surface representation of Cdc34 in apo and **10** bound conformations. (Right) Ribbons representation of apo-Cdc34 (purple) and **10**-Cdc34 (cyan). Arrows indicate structural changes induced by **10** binding.^{65,66} Reprinted (adapted) with permission from Ceccarelli and co-workers.⁶⁵ Copyright 2011 Elsevier.

The therapeutic value of this compound emanates from the inhibited degradation of the cyclin dependent kinase (CDK) inhibitor p27KIP1. This negatively regulates cell cycle progression and outlines the potential of **10** as an indication for oncology

targets.⁶⁷ However, although these initial results looked very promising, very little effort has been put into pursuing the inhibition of E2 enzymes. Instead, E3 ubiquitin ligases have been the subject of intensive research efforts in recent years.

1.3.4 E3 Ligase Inhibition

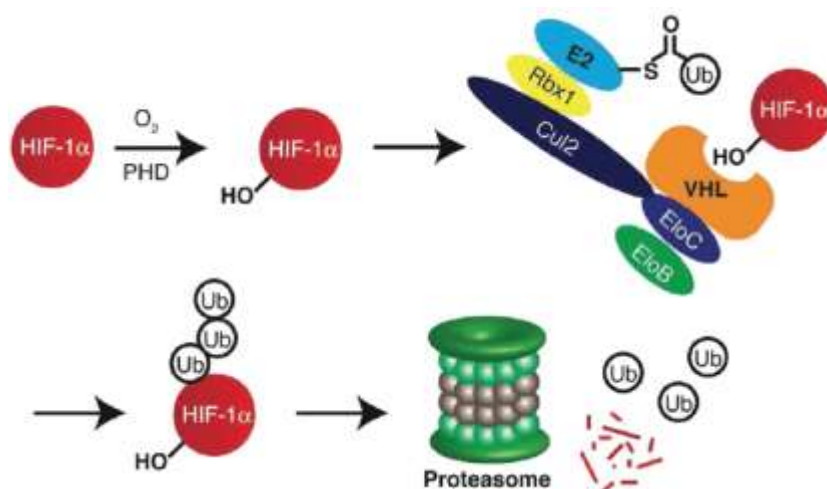
There have been > 600 E3 ubiquitin ligases identified in humans, responsible for recognition and ubiquitination of specific target proteins through the formation of a ternary complex with the ubiquitin-bearing E2 enzyme and the target protein.⁶⁸ Therefore, selective inhibition of a specific E3 ligase will have less of an effect on off-target cellular activity and the global function of the UPS, limiting the magnitude of undesirable side-effects.⁶²

The majority of E3 ligases are RING finger domain type ligases that do not possess a classic active site, but instead mediate direct ubiquitination of the target protein through protein–protein interactions (PPIs) in the ternary complex.⁶⁷ These interactions are inherently hard to inhibit due to the flat topology and large surface area of the PPI interface. Most research has focused on developing inhibitors targeting the ligase-substrate PPIs, starting from endogenous substrates and using extensive structure-activity relationship (SAR) optimisation to develop small molecule inhibitors. This has been successfully achieved with several E3 ligase targets, most notably, von Hippel–Lindau (VHL) and cereblon (CRBN), which will be outlined in this section.

1.3.4.1 von Hippel–Lindau (VHL) inhibition

The primary substrate of the E3 ligase VHL is the 850 amino acid protein, hypoxia inducible factor 1 α (HIF-1 α). This transcription factor regulates over 2% of human genes, particularly those related to oxygen sensing and the hypoxic response.^{69,70} While HIF-1 α is constitutively expressed, under normoxic conditions its levels are suppressed by hydroxylation of the Pro564 residue, subsequent VHL-mediated ubiquitination and proteasomal degradation, **Scheme 7**. Inhibition of the PPIs between HIF-1 α and VHL would mimic the physiological response to hypoxia, upregulating

the expression of related genes, potentially providing therapeutic use as a treatment for chronic anaemia associated with chronic kidney disease and chemotherapy.⁷¹



Scheme 7: Degradation of HIF-1 α under normoxic conditions.⁷² Reprinted (adapted) with permission from Buckley and co-workers.⁷² Copyright 2012 American Chemical Society.

In 2012, Crews and co-workers reported the first small molecule inhibitor of VHL **11**, with a K_d value of 5.4 μM .⁷² Starting from a 4-hydroxyproline (Hyp) core that mimics the key binding interaction between the Hyp564 residue of hydroxylated HIF-1 α and VHL, they used *in silico* design in conjunction with structure-guided medicinal chemistry to develop **11**, a competitive inhibitor of VHL that closely mimics the binding of HIF-1 α , **Figure 6**.⁷³

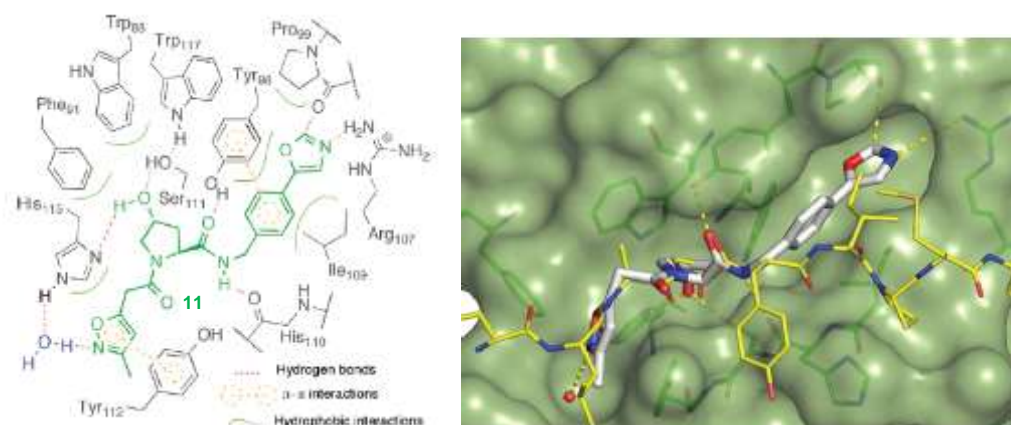


Figure 6: (Left) Graphical representation showing the key interactions of **11** (green) with VHL (grey). (Right) Crystal structure of the binding conformation of **11** (white) and HIF-1 α (yellow) with VHL.⁷² Reprinted (adapted) with permission from Buckley and co-workers.⁷² Copyright 2012 American Chemical Society.

After extensive SAR campaigns and further optimisation, Ciulli and co-workers reported a 500-fold decrease in K_d with **12**, **Figure 7**.^{74,75} This increase in binding affinity was achieved by substituting the oxazole for a 4-methylthiazole, negating the introduction of any strain in the dihedral angle when binding. Additionally, the isoxazole was replaced with the bioisosteric acetamide. The ^tBu group makes favourable hydrophobic contacts with Phe91 and Trp88, directing the carbonyl of the acetamide to make a hydrogen bond interaction with a structural water molecule within the binding pocket, **Figure 7**. Further elaboration of the acetamide methyl group revealed that a range of functionality is tolerated in this region without any change in the binding potency.

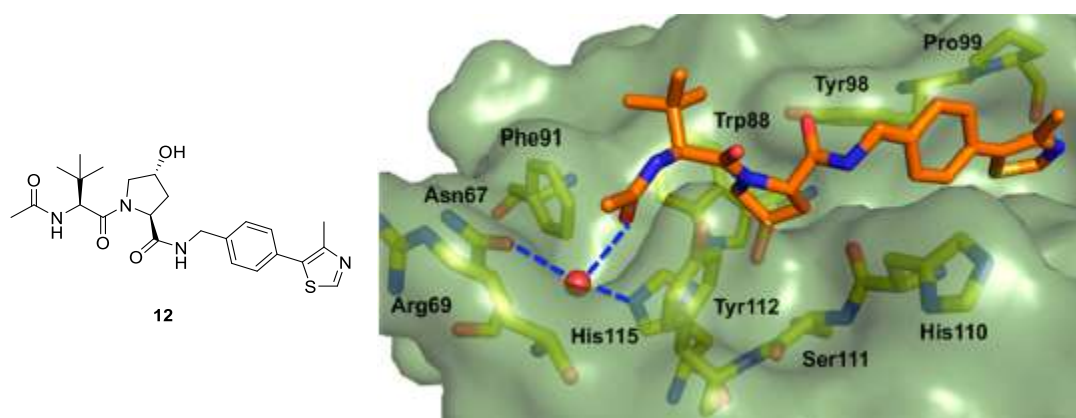


Figure 7: Crystal structure of **12** (orange) with VHL (green surface and yellow residues).⁷⁴ Reprinted (adapted) with permission from Galdeano and co-workers.⁷⁴ Copyright 2014 American Chemical Society.

1.3.4.2 Cereblon (CRBN) Inhibition

Despite the severe teratogenic effects observed in the use of thalidomide **13** for the treatment of morning sickness, the immunomodulatory drugs (IMiDs), **13**, lenalidomide **14** and pomalidomide **15** (**Figure 8**) are currently used as the standard of care for the treatment of refractory multiple myeloma.^{76,77} Although their ability to inhibit the proliferation of multiple myeloma cells and modulate the immune system is well documented, their mechanism of action was not elucidated until 2010.^{78,79} Ito and co-workers identified that thalidomide inhibits CRBN, a ubiquitously expressed

protein that is part of the E3 cullin-4-RING ligase complex known as CRL4, **Figure 8.**⁸⁰

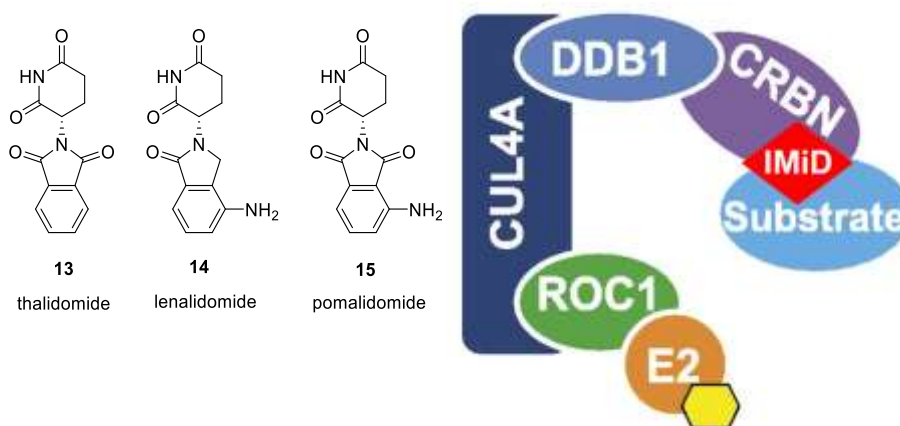


Figure 8: (Left) Chemical structures of thalidomide **13**, lenalidomide **14** and pomalidomide **15**. (Right) Schematic of the E3 cullin-4-RING ligase complex CRL4 with IMiD and substrate bound.⁸¹ Reprinted (adapted) with permission from Fink and co-workers.⁸¹ Copyright 2015 Ash Publications.

The IMiDs **13**, **14** and **15** are characterised by a common glutarimide ring and have similar affinities for CRBN, with K_d values of ~ 250 nM, ~ 178 nM and ~ 157 nM, respectively.⁸² The glutarimide ring sits in a hydrophobic pocket formed by Trp380, Trp386 and Trp400, with Phe402 at the base, **Figure 9.**⁸³ The glutarimide carbonyl (C6) and the adjacent amide (C2 and N1) form hydrogen-bonding interactions with Trp380 and His378 respectively.⁸⁴ The phthaloyl ring has a negligible contribution to the binding and is positioned outside of the binding pocket.

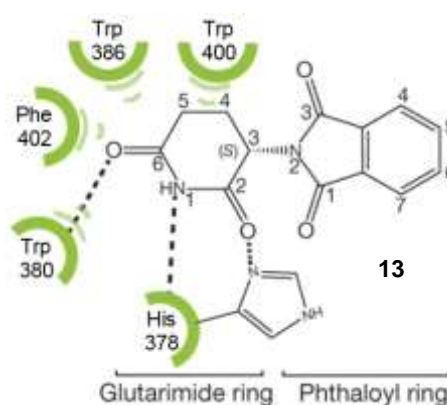
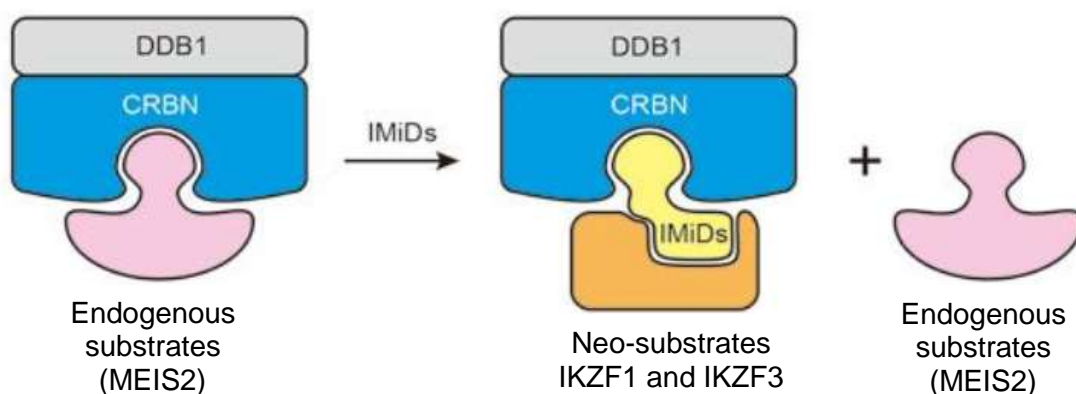


Figure 9: Key binding interactions between **13** and CRBN.⁸¹ Reprinted (adapted) with permission from Fink and co-workers.⁸¹ Copyright 2015 Ash Publications.

The inhibitory effects of the IMiDs are twofold. Primarily they inhibit the binding, ubiquitination and subsequent proteasomal degradation of endogenous substrates of CRBN, such as the transcription factor MEIS2 (myeloid ecotropic viral integration site 1 homologue 2), **Scheme 8**. Additionally, a portion of the bound IMiD is exposed on the surface of the CRBN–IMiD complex, presenting a neomorphic interface.⁸⁵ The transcription factors IKZF1 (also known as IKAROS) and IKZF3 (also known as AIOLOS) bind as neo-substrates to the CRBN-IMiD complex and undergo ubiquitination and subsequent proteasomal degradation. The downstream effects of modulating the levels of these three transcription factors has been implicated in the therapeutic effects of the IMiDs.^{86,87}

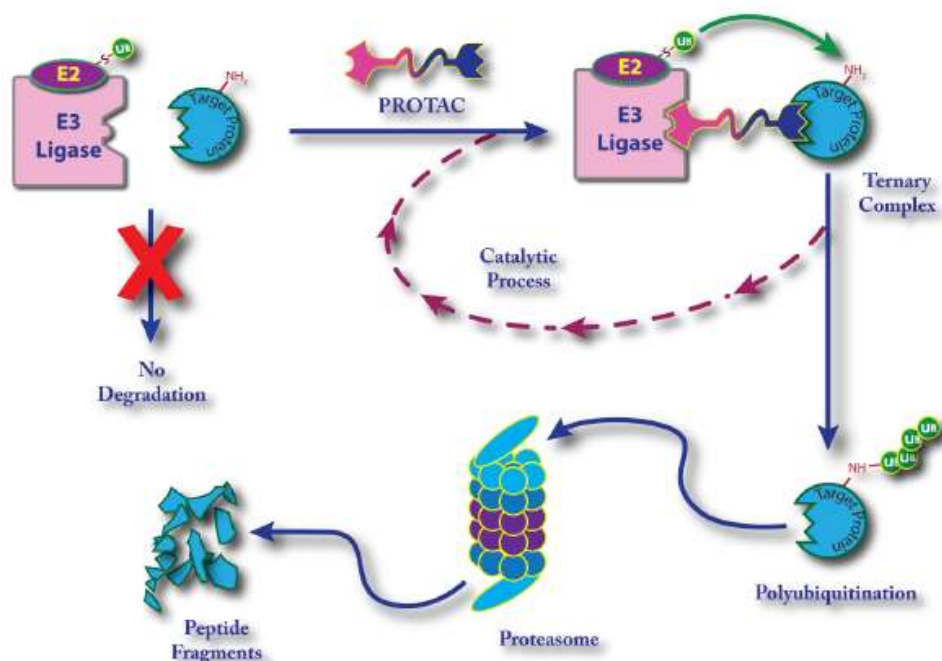


Scheme 8: Schematic representation of IMiDs (yellow) inhibiting the binding of endogenous substrates (pink) to CRBN and recruiting neo-substrates (orange).⁶² Reprinted (adapted) with permission from Huang and co-workers.⁶² Copyright 2016 Springer Nature.

Inhibitors have also been developed for the E3 ligases IAP and Mdm2, however the vast majority of the > 600 E3 ligases remain undruggable. More recently, the focus of research efforts has shifted from inhibiting the UPS to hijacking it to selectively degrade target proteins and elicit a beneficial disease response. This new technology employs a novel class of molecules termed proteolysis targeting chimeras (PROTACs) and will be outlined below.

1.4 Proteolysis Targeting Chimeras (PROTACs)

PROTACs are heterobifunctional molecules that bind to both a target protein and an E3 ligase, bringing them into close proximity to form a ternary complex, **Scheme 9**. This results in the transfer of ubiquitin from the E3 ligase to a surface-exposed lysine residue on the target protein. Upon polyubiquitination, the target protein is recognised by the 26 S proteasome and degraded.



Scheme 9: Schematic representation of how PROTACs hijack a cell's UPS to degrade a target protein.¹ Reprinted (adapted) with permission from Tinworth and co-workers.¹ Copyright 2016 Royal Society of Chemistry.

The PROTAC approach offers several advantages over the classical inhibitor-based approach. Firstly, target protein degradation not only reduces the number of active proteins that need to be inhibited but also counteracts compensatory protein overexpression, often observed upon loss of protein function by inhibition.⁸⁸ Furthermore, degradation of a target protein inhibits its secondary scaffolding functions, which are not targeted by a classical inhibitor.⁸⁹

Additionally, ternary complex formation is reversible, giving PROTACs the potential to be catalytic in activity. This means that one PROTAC molecule can label many

target proteins for degradation. Additionally, PROTACs can have long-lasting therapeutic effects, as the return of protein function is dependent on the rate of protein resynthesis. This gives PROTACs the potential for low clinical doses and infrequent dosing regimens.¹⁷

A unique advantage of PROTACs is that they operate through an event-driven MOA, where only a transient binding event is required for ubiquitin transfer to the target protein. This contrasts the occupancy-driven MOA of classical inhibitors and offers the possibility of degrading target proteins that have a low tractability for inhibitors due to non-specific active sites with flat topology. PROTACs which target what is currently referred to as the undruggable proteome may elicit novel pharmacology that allows the development of new therapeutics.⁹⁰

PROTAC-mediated target protein degradation is dependent on ternary complex formation, which is governed by a combination of binary equilibrium constants, **Figure 10**. Therefore, PROTACs exhibit a bell-shaped dose-dependent response, termed the ‘hook effect’.⁸⁹ Initially, as the concentration of PROTAC increases, the concentration of productive ternary complexes increases until a saturation point, **Figure 11**. Beyond this point, as the concentration of PROTAC increases, the concentration of productive ternary complexes decreases as the concentration of unproductive binary complexes increases.

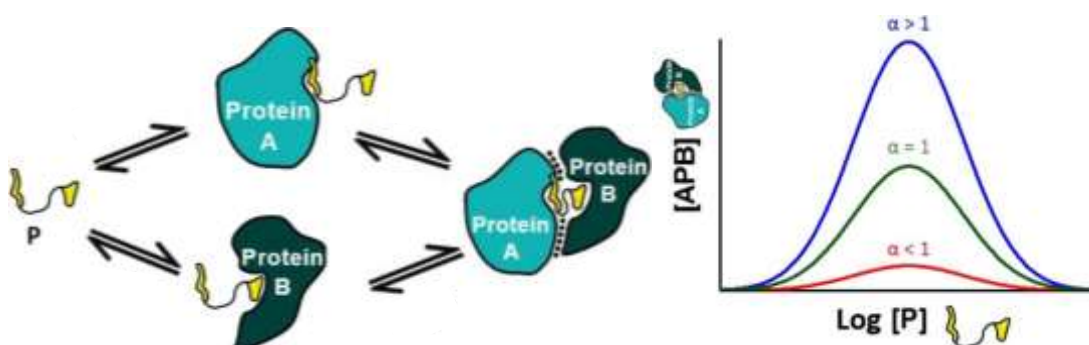


Figure 10: (Left) Ternary complex equilibria for protein A and B (green) and PROTAC (yellow). (Right) Graph showing the effect of cooperativity (α) on the relationship between PROTAC concentration and ternary complex formation.⁹¹ Reprinted (adapted) with permission from Hughes and co-workers.⁹¹ Copyright 2017 Portland Press.

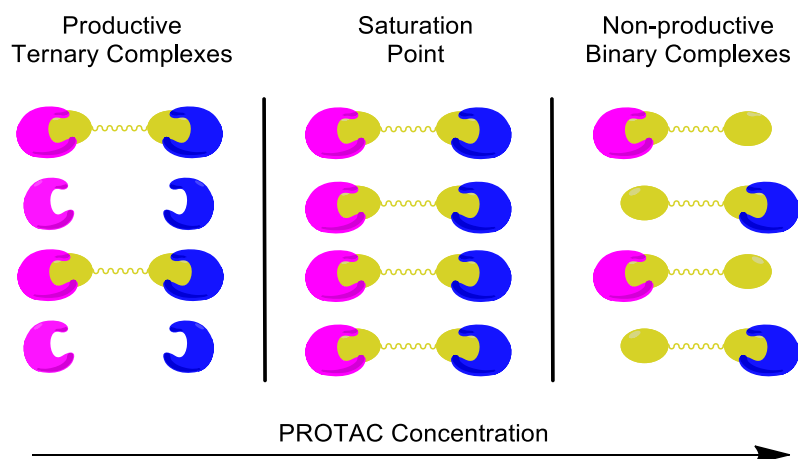


Figure 11: Schematic representation of the relationship between PROTAC concentration and ternary complex formation.

The extent of this hook effect is determined by the ternary complex cooperativity (α), defined as $K_d^{\text{binary}}/K_d^{\text{ternary}}$. Positive cooperativity ($\alpha > 1$) is observed for ternary complexes that exhibit favourable stabilising PPIs between the target protein, PROTAC and E3 ligase.⁹² A negative cooperativity ($\alpha < 1$) is observed when PPIs abrogate ternary complex formation. The majority of the ternary complex PPIs are between the large surfaces of the target protein and E3 ligase, however, these can be supplemented by intermediary interactions with the PROTAC.

The first PROTAC was reported by Crews and Deshaies in 2001, **Figure 12**.⁹³ PROTAC **16** targeted the methionine aminopeptidase MetAP2, which is an oncology target that has been implicated in the growth of new blood vessels in tumours.⁹⁴ It utilised the known MetAP2 small molecule inhibitor ovalicin, linked to a phosphopeptide E3 ligase binder, to successfully degrade MetAP2 *in vitro*. This landmark paper highlighted the feasibility of the PROTAC approach, but due to its reliance on a highly charged peptidic E3 ligase binder, it did not exhibit cellular efficacy.

Another interesting point to highlight is that ovalicin irreversibly inhibits MetAP2 by forming a covalent bond between a methionine residue and the epoxide ring on its cyclohexyl ring, thus negating the unique catalytic activity of PROTACs.

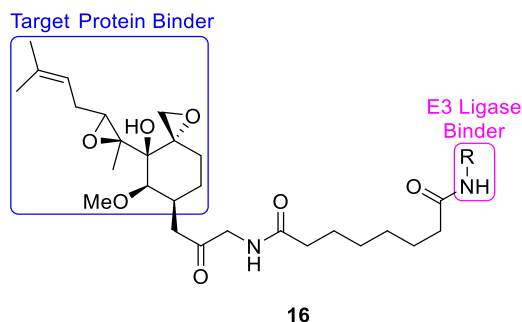


Figure 12: First published peptidic PROTAC **16**. $R = \text{GGGGGGRAEDS*GNES*EGE-COOH}$ or $\text{GGGGGGDRIIDS*GLDS*M-COOH}$. * = Phosphorylated residues.⁹³

The first non-peptidic PROTAC was reported in 2008, when a nutlin, recently described to bind to the E3 ligase Mdm2, was combined with an androgen receptor inhibitor, **Figure 13**.⁹⁵

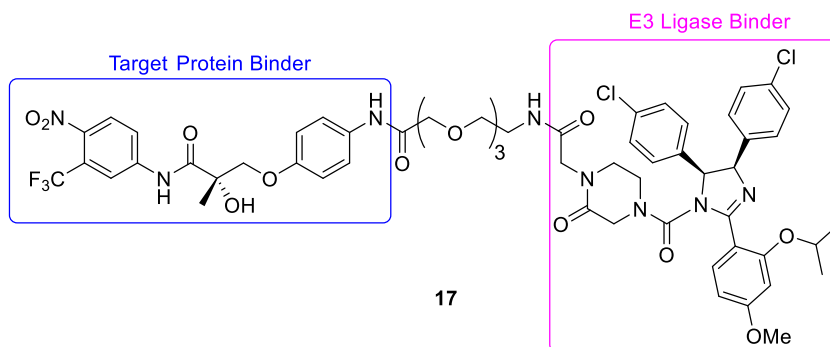


Figure 13: First published non-peptidic PROTAC **17**.⁹⁵

PROTAC **17** exhibited partial degradation of its target protein in a cellular environment. However, its modest level of cellular efficacy was attributed to its large and complex chemical structure that confers undesirable physicochemical properties.

Since the identification of small molecule E3 ligase inhibitors in 2010, **Section 1.3.4.1** and **1.3.4.2**, there have been a myriad of publications containing small molecule PROTACs. These PROTACs generally utilise CRBN or VHL E3 ligase inhibitors as the E3 ligase-binding moiety, with a plethora of linkers connecting them to the target protein-binding moiety. A range of target protein-binding moieties have been exemplified by repurposing previously developed inhibitors of the desired target

protein. This is achieved by linking the inhibitor through a vector that does not affect its binding to the target protein.

Inhibition of bromodomain and extra-terminal domain (BET) proteins has been of interest to GSK, amongst other pharmaceutical companies, over recent years, as dysregulation of BET protein activity has been strongly linked to a range of diseases.^{96,97} The importance of this family of proteins and the development of inhibitors for them will be outlined below.

1.5 Bromodomain and Extra-Terminal Domain (BET) Proteins

1.5.1 Deoxyribonucleic Acid (DNA) Structure and Function

Deoxyribonucleic acid (DNA) is the carrier of genetic information and codes for the synthesis of proteins from their constituent amino acids. Seminal work by Wilkins and Franklin was further developed by Watson and Crick to elucidate the complex structure of DNA.⁹⁸⁻¹⁰⁰

DNA is formed of two complementary strands that are antiparallel and assembled to form a double helical structure, **Figure 14**. Each chain is composed of monomeric units called nucleotides, which are made up of a nucleobase, deoxyribose sugar and phosphate group.

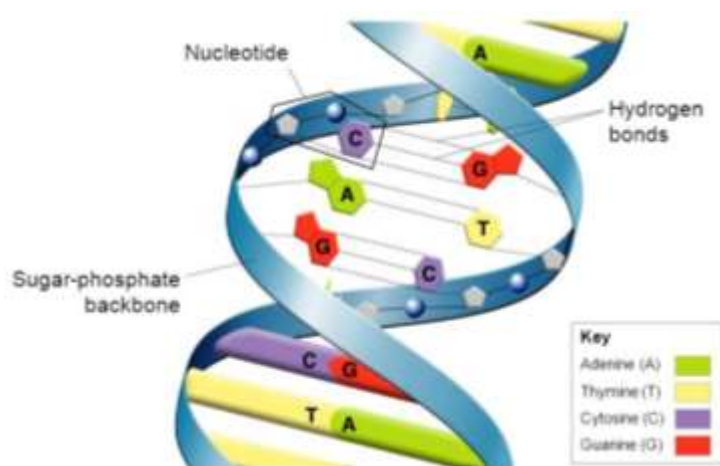


Figure 14: Schematic representation of the structure of DNA.¹⁰¹

The sugar-phosphate backbone maintains the structural integrity of each strand, whilst hydrogen bonds between complementary nucleobases tether the two strands together.¹⁰² Additionally the specific sequence of nucleobases provides the genetic code. DNA is made up of the following four nucleobases; adenine (A), cytosine (C), guanine (G) and thymine (T).¹⁰³ In Watson-Crick base pairing, A forms a complementary base pair with T, and C with G, **Figure 15**.

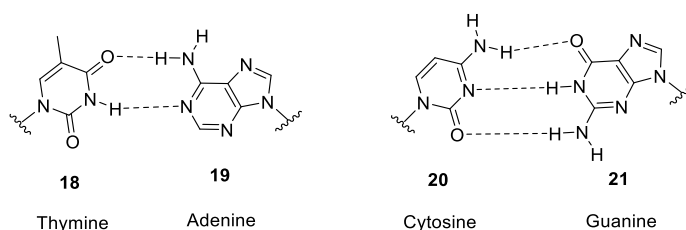


Figure 15: DNA nucleobase complementary base pairs.

DNA replication is semi-conservative and occurs with the separation of its two strands. Nucleobases present within the cell form complementary base pairs with the exposed nucleobases of each single strand to form two identical double-stranded DNA fragments. This process is mediated by various proteins as part of the cell cycle of division and growth.¹⁰³

The sections of DNA that code for the synthesis of specific proteins are termed genes. During protein synthesis, genes are transcribed to form messenger ribonucleic acid (mRNA), a complementary single stranded nucleic acid formed from the transcription process. The newly formed mRNA strand leaves the nuclear envelope and is translated by ribosomes.¹⁰³ A sequence of three nucleobases in mRNA makes up a codon, which interacts with a complementary anti-codon of transfer RNA (tRNA) through complementary base pairing in a process known as translation. Each tRNA complex carries a specific amino acid defined by its anti-codon and thus determines the amino acid sequence of the growing polypeptide chain, **Figure 16**. Once the stop codon of mRNA is reached, the polypeptide chain dissociates from the ribosome and through the formation of intramolecular bonds and folding, a functional protein is furnished. This protein can be modified for a range of different functions through PTMs. PTMs

are generally mediated by enzymes and commonly include phosphorylation, glycosylation and ubiquitination.¹⁰⁴

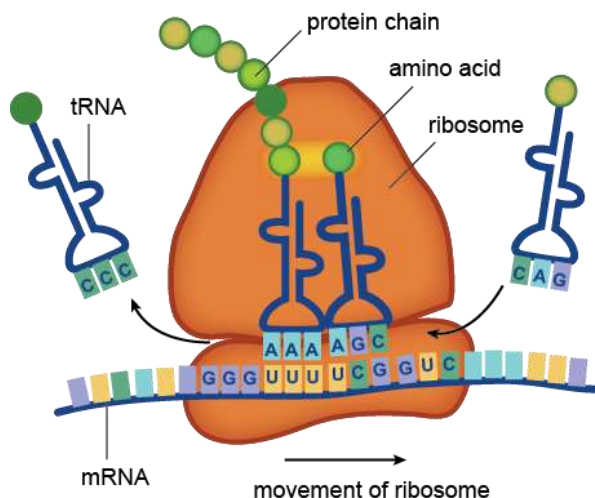


Figure 16: Ribosomal gene translation.¹⁰⁵

1.5.2 Epigenetics

The decoding of the human genome in 2003 elucidated the specific sequence of nucleobases present in human DNA. This enabled the identification of approximately 30,000 protein-encoding genes. However, in comparison to the many phenotypes in human development, there are far fewer protein-encoding genes. One explanation for this is that it is a result of epigenetics, which is defined as heritable changes in gene expression through mechanisms that are not facilitated by changes to the underlying DNA sequence.^{106,107}

The nucleus of each cell contains 46 chromosomes, each made up of a single double-stranded DNA helix that carries the genetic information required for gene expression. A single DNA helix is approximately 6 cm in length when fully extended and thus must be compressed to fit inside the cell nucleus, which is 5-10 μM in diameter. This is achieved by wrapping the DNA around small protein modules known as histones (H2A, H2B H3 and H4), that assemble into an octameric core termed a nucleosome.¹⁰⁸ Each nucleosome contains and compresses 147 nucleobase pairs of DNA and is the monomeric unit that makes up a chromatin fibre, **Figure 17**.

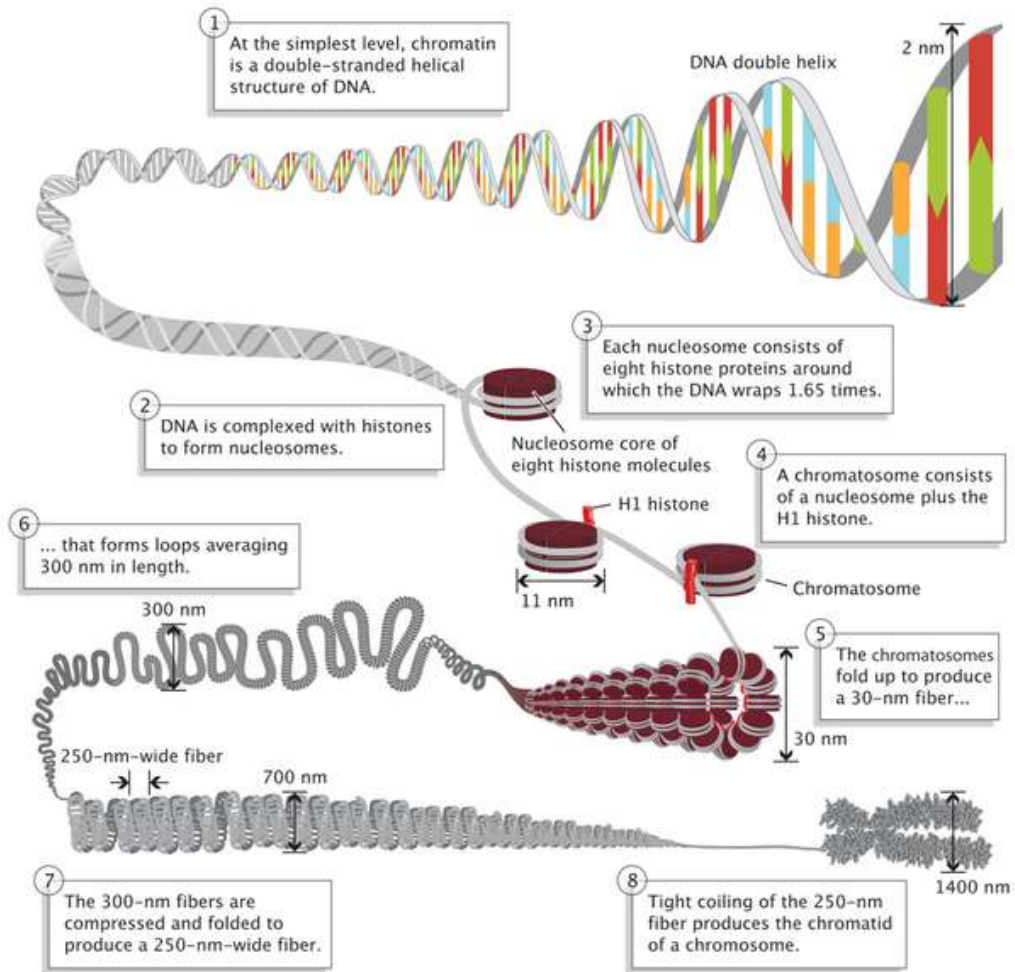


Figure 17: *The structure of chromatin within chromosomes.*¹⁰⁹

The compression of DNA into these highly organised structures is dynamic, with epigenetic proteins regulating the wrapping of DNA around histone proteins. This is achieved through epigenetic PTM of the terminal histone tails that are exposed at the nucleosome surface. These site-specific modifications include acetylation, methylation and phosphorylation.¹¹⁰ The specific sequence of PTMs constitutes the ‘histone code’ that modulates the structure of chromatin and thus gene expression.¹¹¹

Of the PTMs, lysine acetylation is one of the most abundant and has been the focus of extensive research efforts. Approximately 25% of histone amino acid residues are lysine and arginine. At physiological pH these residues are positively charged and form electrostatic binding interactions with the negatively charged phosphate

backbones of DNA, tightly wrapping it around the histone protein. This compressed state is termed heterochromatin and the DNA is deactivated to transcription. Selective acetylation of the ϵ -amino group of lysine is mediated by histone acetyltransferases (HATs) and results in charge neutralisation. This reduces the binding interaction between DNA and the histone protein to furnish a more open form of DNA termed euchromatin, which is more accessible and activated to transcription.¹¹² Selective deacetylation of the ϵ -amino group of lysine is mediated by histone deacetylases (HDACs); this restores its positive charge and reforms heterochromatin. However, it is believed that HATs and HDACs do not mediate transcription activation independently, instead they recruit numerous chromatin remodeling complexes.¹¹³

The epigenetic PTM of lysine acetylation is written into the histone code by HATs and erased by HDACs. These epigenetic marks are read by bromodomains to facilitate gene expression, **Figure 18**. The bromodomain-containing proteins (BCPs) recruit and localise the cellular transcriptional machinery to these specific epigenetic marks.

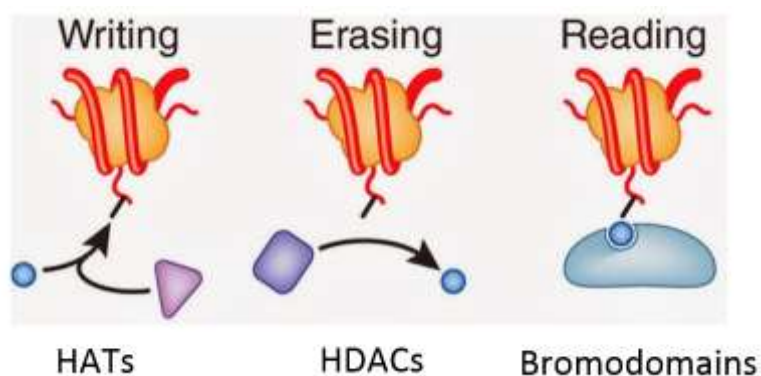


Figure 18: Epigenetic PTMs of lysine acetylation are written into the histone code by HATs, erased by HDACs and read by bromodomains, to facilitate gene expression.¹¹⁴

1.5.3 Bromodomain-Containing Proteins (BCPs)

Bromodomains are currently the only known protein modules that selectively bind to the acetylated lysine residues of DNA.¹¹⁵ Currently, there are 56 known bromodomains that are present in the 42 known BCPs, which are classified by their sequence homology into eight distinct subgroups, **Figure 19**.¹¹⁶

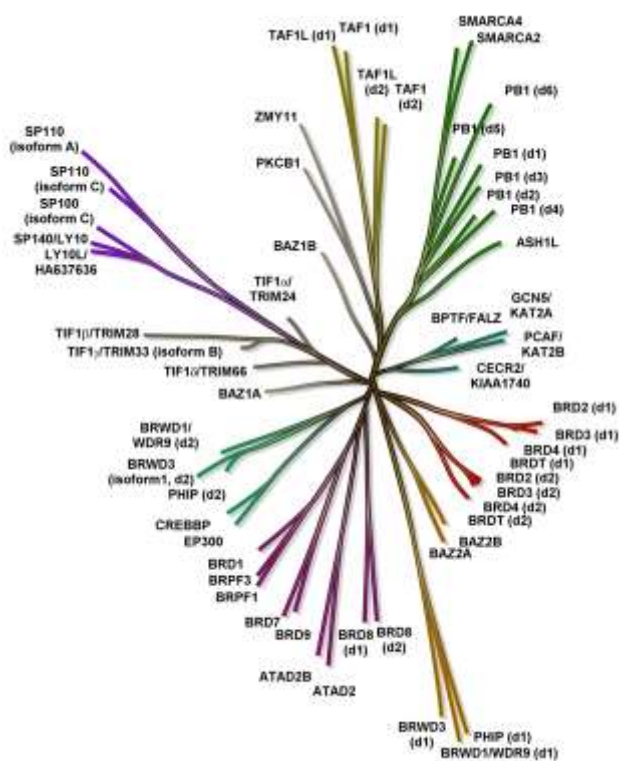


Figure 19: Human bromodomain phylogenetic tree, classifying the 56 known bromodomains into eight subgroups based on sequence homology.¹¹⁷ Reprinted (adapted) with permission from Chung and co-workers.¹¹⁷ Copyright 2012 American Chemical Society.

Bromodomain protein modules contain approximately 110 amino acids and have a highly conserved structure, consisting of four anti-parallel α -helices (α_A , α_B , α_C and α_Z) that are linked by the flexible hydrophobic loop regions ZA, AB and BC, **Figure 20.**¹¹⁸

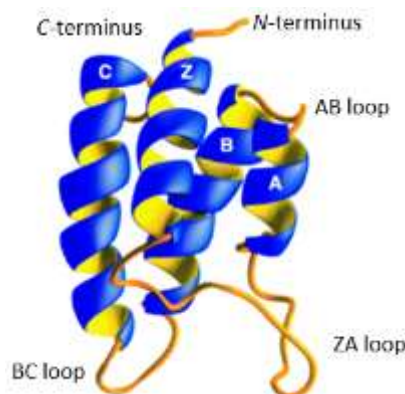


Figure 20: Bromodomain protein crystal structure.¹¹⁸ Reprinted (adapted) with permission from Dhalluin and co-workers.¹¹⁸ Copyright 1999 Springer Nature.

A hydrophobic binding pocket capable of binding an acetylated lysine residue is present between the ZA and BC loops. All bromodomains contain a highly conserved asparagine residue deep inside the pocket that is capable of forming a direct hydrogen bond between its side chain amide nitrogen and the acetyl carbonyl group of acetylated lysine, **Figure 21**.¹¹⁹ This binding is further supplemented by a water-mediated hydrogen bond interaction between a tyrosine residue deep inside the pocket and one of four structural water molecules that are conserved in the majority of bromodomains.

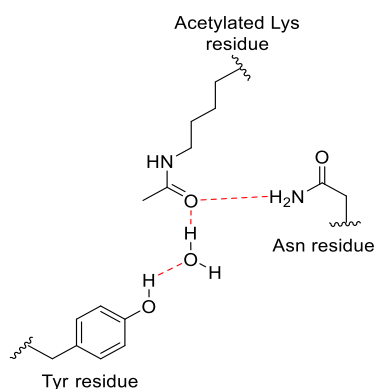


Figure 21: Key binding interactions between acetylated lysine residues and bromodomains.

Although bromodomains contain highly conserved structures, subtle variations in the amino acid sequence, predominantly present in the ZA and BC loops, enable bromodomains to selectively bind site-specific acetylated lysine residues.¹²⁰

The dysregulation of BCP activity has been strongly linked to cancer, autoimmune and inflammatory diseases. This is a result of the overexpression of specific genes and their proteins, and the subsequent downstream effects.^{96,121–123} Of the BCPs, the BET proteins have been the focus of intensive research efforts aiming to develop therapeutics for these disease classes.

1.5.4 Small Molecule BET Inhibitors

The BET family of proteins consists of Brd2, Brd3, Brd4 and BrdT. They contain highly conserved structures characterised by the presence of two bromodomains, BD1 and BD2, and an extra-terminal (ET) domain, **Figure 22**.

The BET proteins are transcriptional coactivators, which relay signals from master regulatory transcription factors, such as MYC in cancer and NF- κ B in inflammation, to RNA polymerase II which facilitates transcriptional elongation.^{124–126} In addition to chromatin-dependent transcriptional signaling, BET proteins are also important mediators of cell cycle progression and spermiogenesis.^{127,128}

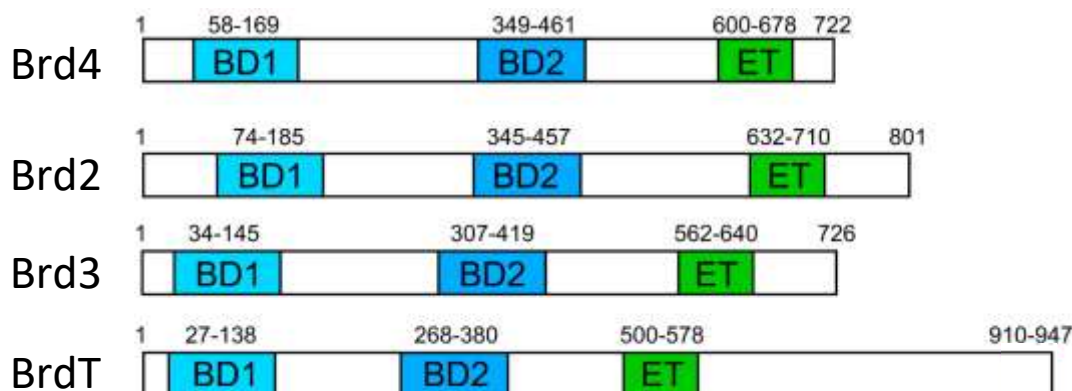


Figure 22: Amino acid domain structure of human BET proteins.¹²⁹ Reprinted (adapted) with permission from Shi and co-workers.¹²⁹ Copyright 2014 Elsevier.

The oncogenic potential of BET proteins was first identified in nuclear protein in testis (NUT) midline carcinoma (NMC). NMC is a malignant solid tumour in which oncogenesis is promoted by the fusion of Brd4 with NUT protein. In 2010, Bradner and co-workers reported JQ1 **22**, a cell-permeable small molecule that competitively inhibits BET bromodomains with high potency and specificity over non-BET bromodomains, **Figure 23**.¹³⁰ The crystal structure was solved for Brd4, indicating that **22** occupies a similar binding conformation to acetylated lysine residues, forming a hydrogen bond between its triazole ring, which acts as an acetylated lysine mimetic, and the conserved Asn140 residue deep inside the pocket.

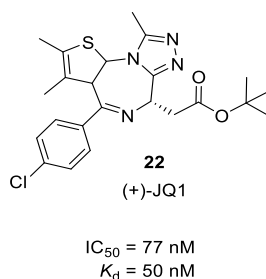


Figure 23: BET inhibitor JQ1 **22**.^{131,132}

JQ1 **22** binding displaces the Brd4-NUT fusion oncoprotein from chromatin, inducing cell-cycle arrest, subsequent cell apoptosis and thus a decrease in tumour development in mouse xenografts.¹³³ However, at efficacious doses, a decrease in lymphoid and immune cell compartments was also observed, resulting in significant body weight loss and early euthanasia.¹³⁴ Therefore, **22** was not progressed through clinical trials.

Of the BET family of proteins, Brd4 provides the most therapeutic potential, as most toxicological side-effects are believed to be associated predominantly with BrdT (infertility) and Brd3 (thrombocytopenia).^{128,135} However, due to the highly conserved structures of the BET bromodomains, obtaining selectivity between the different isoforms is very challenging.

Since the seminal work of Bradner and co-workers, a plethora of BET inhibitors targeting a variety of diseases have been published in the literature. These compounds use a range of acetylated lysine mimetics, including: isoxazoles, pyridones and tetrahydroquinolines, amongst others, **Figure 24**.¹³⁶

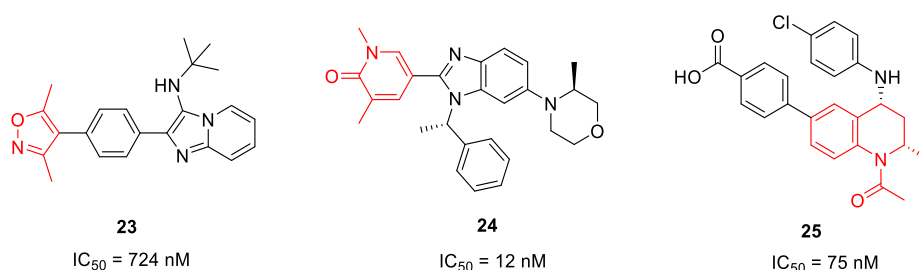


Figure 24: BET inhibitors using different functional groups highlighted in red as acetylated lysine mimetics.^{137–139}

There are currently more than 10 BET inhibitors in clinical trials that exhibit selectivity for the BET proteins over non-BET proteins, as well as isoform selectivity within the BET family.¹³⁶ However, these inhibitors operate through an occupancy-driven MOA, requiring high occupancy levels (> 90%) to achieve inhibitory efficacy.² This can be increasingly difficult to obtain as inhibition of BET proteins is often counteracted by compensatory cellular overexpression of the proteins. Shimamura and co-workers reported this observation for Brd4 within 24 h, when treating lung cancer cell lines with JQ1 **22**.¹⁴⁰ This may be due to an increase in the thermal stability of the protein

upon binding of the inhibitor, competitive inhibition of cellular degradation machinery, or interruption of a negative feedback loop that regulates Brd4 protein levels. Ultimately, this increases the concentration of inhibitor required for efficacy and often leads to undesired side-effects due to off-target binding associated with higher drug concentrations.³

The PROTAC approach operates through an event-driven MOA where only a transient binding event is required for ubiquitin transfer from the E3 ligase to the target protein. This circumvents the requirement for high levels of occupancy. Additionally, as the protein is degraded, it is less susceptible to the effects of cellular overexpression of the target protein. Therefore, using previously developed BET inhibitors as the target protein-binding moiety in PROTACs has become increasingly popular. The development and performance of a selection of BET PROTACs will be outlined and discussed below.

1.6 Small Molecule BET PROTACs

In June 2015, both Crews and co-workers, and Ciulli and co-workers published BET PROTACs using the pan-BET inhibitor JQ1 **22** with a CRBN and VHL E3 ligase binder respectively, **Figure 25**.^{141,142}

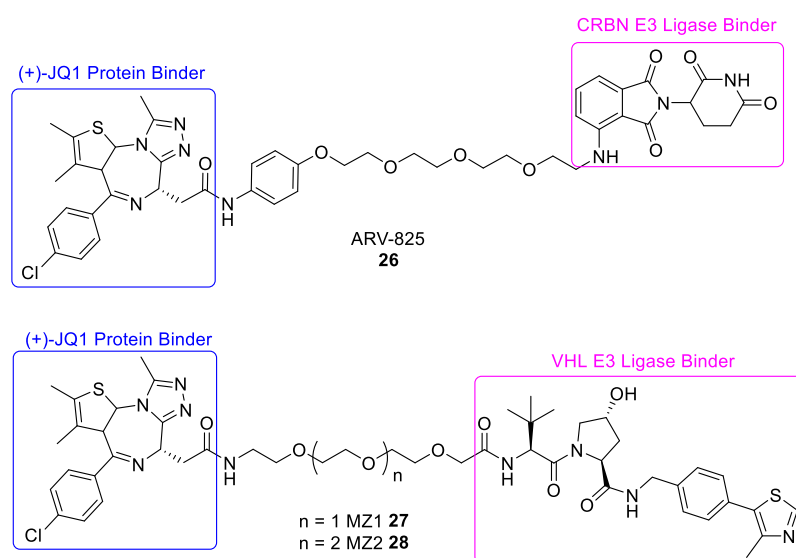


Figure 25: Published BET PROTACs using the BET inhibitor JQ1 **22**.

Crews and co-workers judiciously chose the PEG linker of ARV-825 **26** due to its high conformational flexibility, which they hypothesised would maximise the chance of productive ternary complex formation. Linking the JQ1 inhibitor **22** to form the PROTAC ARV-825 **26** led to a small decrease in binding affinity for Brd4 BD1/BD2 from 12/10 nM to 90/28 nM. ARV-825 **26** induced potent degradation of Brd4, with an observed pDC₅₀ (the concentration at which 50% of the protein is degraded) of < 1 nM.¹⁴¹ Furthermore, co-treatment with the proteasome inhibitor MG132 **2** completely blocked degradation, confirming degradation is *via* the UPS. ARV-825 **26** achieved a more pronounced and sustained downregulation of the c-MYC oncogene relative to JQ1 **22**, maintaining suppression for up to 24 h after the PROTAC was washed out. Additionally, ARV-825 **26** was shown to more efficiently suppress cell proliferation and induce cell apoptosis than JQ1 **22**. This may be a result of not only inhibiting the primary function of Brd4, but also removing any secondary scaffolding functions through degradation of the protein.

Ciulli and co-workers observed complete degradation of Brd4 upon treatment with MZ1 **27**, with no protein observable by western blotting after 24 h, **Figure 26**.¹⁴² The stereochemistry of the hydroxyl group of the central hydroxy-proline moiety in MZ1 **27** is crucial for VHL engagement. Treatment of cells with MZ1 with the inverse stereochemistry at this stereocentre resulted in no observable degradation, confirming the degradation was VHL-mediated and *via* the UPS.

Interestingly, increasing the linker length to four PEG units in MZ2 **28**, resulted in a decrease in Brd4 degradation relative to MZ1 **27**. It was previously hypothesised that longer, more flexible linkers, that are better able to accommodate a range of conformations, have an increased probability of forming a productive ternary complex. However, this result suggests that in fact there is an optimal length to achieve productive ternary complex formation and ultimately degradation. Additionally, MZ1 **27** showed a pronounced selectivity for the degradation of Brd4 over Brd2 and Brd3 at concentrations below 0.5 μ M. This was unexpected, given that JQ1 **22** is a pan-BET inhibitor with similar binding affinities to all BET isoforms. This suggests that additional selectivity is achieved in the ternary complex.

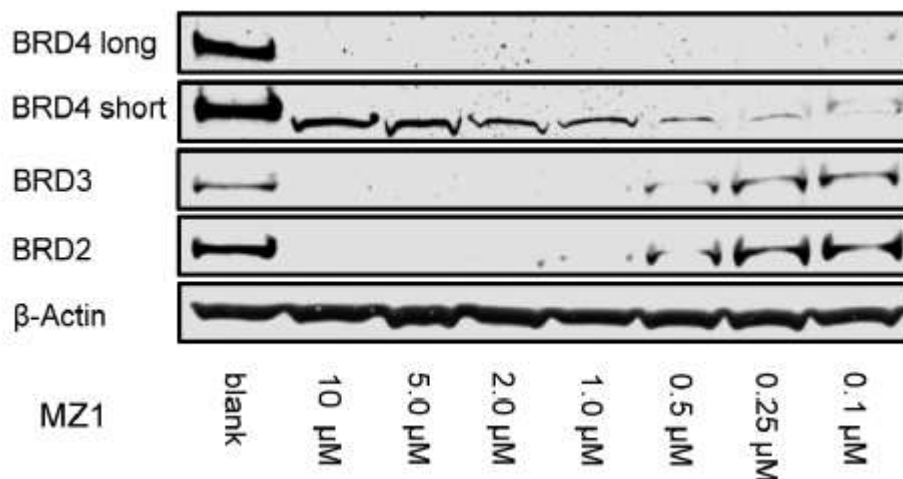


Figure 26: Western blot of Brd4, Brd3 and Brd2 degradation 24 h after dosing with MZ1 27.¹⁴² Reprinted (adapted) with permission from Zengerle and co-workers.¹⁴² Copyright 2015 American Chemical Society.

In 2017, Ciulli and co-workers solved the crystal structure for the ternary complex of MZ1 27 with Brd4 BD2 and human VHL, **Figure 27**.¹⁴³

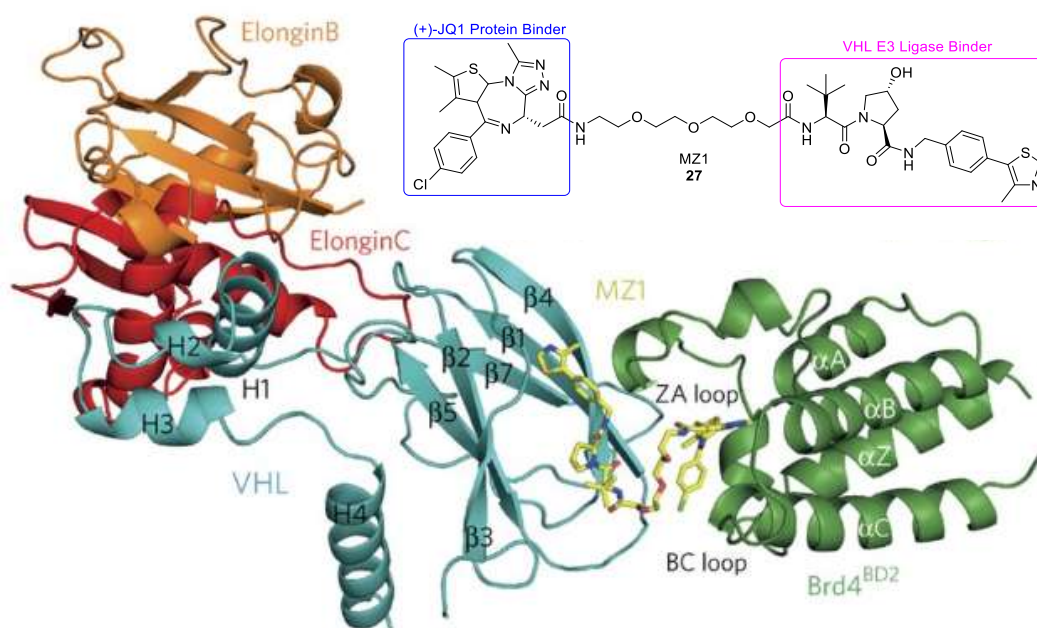


Figure 27: Crystal structure for the ternary complex of MZ1 27 with Brd4 BD2 (green) and human VHL (turquoise).¹⁴³ Reprinted (adapted) with permission from Gadd and co-workers.¹⁴³ Copyright 2017 Springer Nature.

Extensive PPIs between Brd4 BD2 and VHL form a bowl-shaped interface in which MZ1 **27** is bound. The bowl has a hydrophobic base formed through key interactions between the WPF shelf of Brd4 BD2, made up of a Trp, Pro and Phe, and complementary residues on VHL, as well as the second helical turn of the ZA loop of Brd4 BD2 and the hydrophobic side chains of VHL β 4. Two electrostatic interactions exist between the charged residues of the Brd4 BD2 ZA loop and BC loop with complementary VHL residues, **Figure 28**.

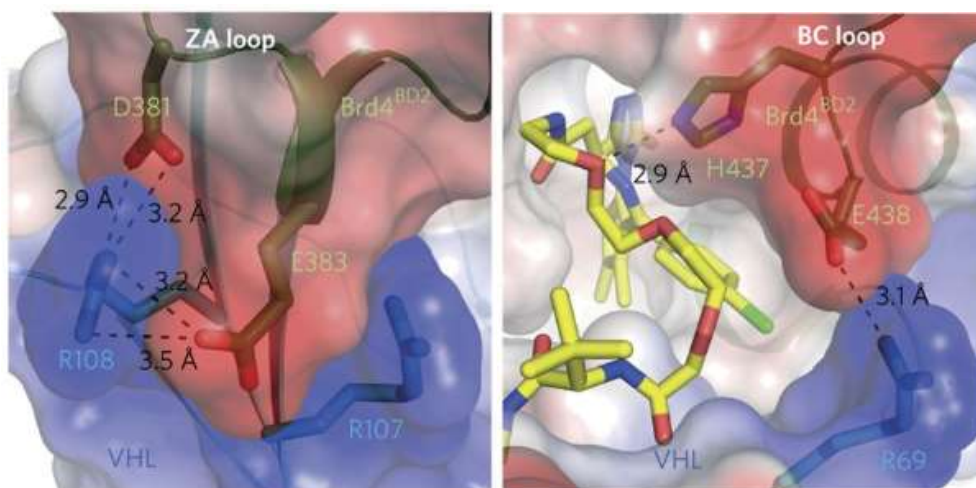


Figure 28: Electrostatic potential map highlighting the interactions between charged residues of the Brd4 BD2 ZA loop (red) (**Left**) and BC loop (red) (**Right**) with complementary VHL residues (blue). Dashed lines indicate hydrogen bonds.¹⁴³ Reprinted (adapted) with permission from Gadd and co-workers.¹⁴³ Copyright 2017 Springer Nature.

MZ1 **27** can fold in such a way that the PEG linker is packed between the *t*Bu group of the VHL E3 ligase binder and the *p*-chlorophenyl ring of JQ1. This positions the PEG linker to make a van der Waals interaction with the BC loop of Brd4 BD2, as well as a hydrogen bond between the ether oxygen closest to the JQ1 amide linkage and the His437 residue of Brd4 BD2, **Figure 28 (Right)**.

This interesting observation confirms that the linker not only plays a critical role in determining the proximity of the target protein and E3 ligase in the ternary complex, but that it is also capable of forming PPIs with both the target protein and E3 ligase, increasing positive cooperativity and driving the equilibrium towards productive ternary complex formation.

Due to small but significant differences in the amino acid sequence of the BET isoforms, these PPIs are less extensive for Brd2 and Brd3, reducing the positive cooperativity and the population of productive ternary complexes relative to Brd4. This offers the possibility of developing therapeutics that selectively target specific BET isoforms using a pan-BET inhibitor.

Since these first publications by Crews and Ciulli, there has been a plethora of examples of PROTACs targeting BET proteins and kinases, as well as a range of other targets.¹⁴⁴⁻¹⁵⁰

However, although this new modality is showing great promise, PROTACs are large molecules with molecular weights of 800-1200 Da. This results in them generally exhibiting poor physicochemical properties, including low cell permeability and low aqueous solubility.⁹⁰ Due to the specific defined structures of the target protein and E3 ligase-binding moieties, the linker provides a facile handle to modulate physicochemical properties. Therefore, the development and application of new linker strategies and synthetic methodologies that may allow PROTACs to access more favourable physicochemical space are of high priority.

2: Project Aims

2 Project Aims

It is evident from the literature review in **Section 1** that hijacking a cell's ubiquitin proteasome system using PROTAC technology has the potential to be a powerful new therapeutic strategy. Some of the main advantages of this technology include the potential for PROTACs to be catalytic in activity and elicit long-lasting therapeutic effects, offering the possibility of low clinical doses and infrequent dosing regimens.¹⁷ However, PROTACs are large molecules with molecular weights of 800-1200 Da, and this results in them generally exhibiting low cell permeability, low aqueous solubility and ultimately low bioavailability.⁹⁰ Due to the specific defined structures of the target protein and E3 ligase-binding moieties, the linker provides a facile handle to modulate physicochemical properties. However, PROTACs are commonly synthesised *via* bespoke synthetic sequences that are not amenable to iterative changes in the linker. Modern decarboxylative cross-coupling methodologies may facilitate the synthesis of PROTACs in an iterative manner, allowing for the high-throughput screening of a plethora of linkers.

The aims of this project were to:

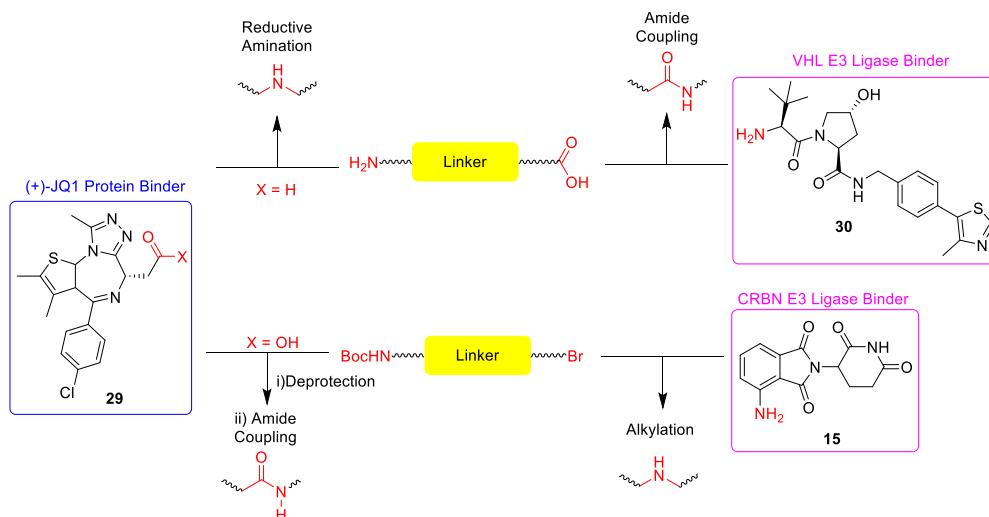
- Assess the amenability of modern decarboxylative cross-coupling methodologies to the synthesis of PROTACs.
- Develop a synthetic strategy and protocol for the synthesis of PROTACs in an iterative, high-throughput manner, that will allow the synthesis of PROTACs with a plethora of linkers in a plate-based format.
- Design a linker library that provides broad coverage of desirable physicochemical space, including a range of linker lengths and functional groups.
- Investigate the effects of linker functionality on the physicochemical properties and degradation profiles of PROTACs by synthesising a series of PROTACs from the linker library using the developed synthetic strategy and protocol.
- Identify trends between linker functionality and both physicochemical properties and degradation profiles which can be used to inform future linker choices and develop PROTACs with more desirable physicochemical properties.

3: Results and Discussion – *Investigating the Synthesis of BET PROTACs*

3.1 Introduction

3.1.1 Common Synthetic Methods

Due to the modular structure of PROTACs, a wide variety of chemical reactions can be used to synthesise them in a step-wise manner. This usually relies on the orthogonal reactivity of the terminal functional groups of the linker, allowing the sequential coupling of target protein and E3 ligase binder. Commonly used reactions include amide couplings, S_N2 alkylations and reductive aminations, **Scheme 10**.¹⁵¹ The choice of chemical reaction depends on the functionality of the linking vector of the target protein and E3 ligase binders. The VHL and CRBN E3 ligase binders commonly have terminal amines, whilst the functionality of the protein binder can be more varied.

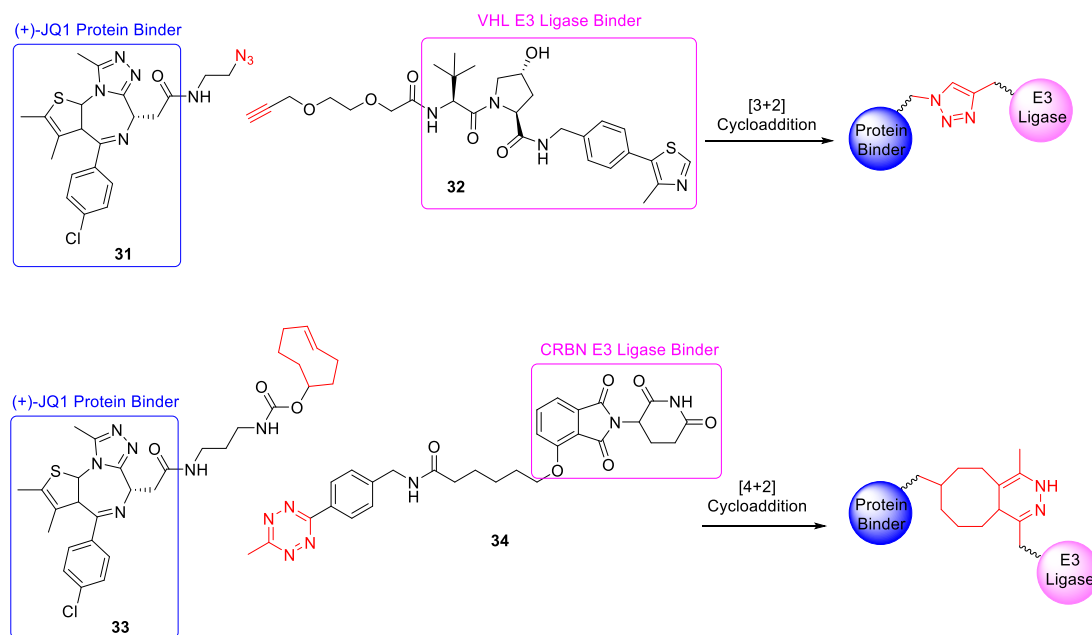


Scheme 10: Commonly used reactions to synthesise PROTACs.

However, these reactions leave legacy amide or amine functional groups in the PROTACs. This increases the number of hydrogen bond donors (HBD) and acceptors (HBA), limiting the ability to control the effects of these properties on the physicochemical profile of the PROTACs.

Click chemistry is also amenable to the synthesis of PROTACs and requires the independent functionalisation of both the target protein and E3 ligase-binding moieties to install the required functionality to perform the click reaction between the two linking vectors. This can be achieved using alkynes and azides in a [3+2] cycloaddition

that forms a 1,2,3-triazole, or through a [4+2] inverse electron demand Diels–Alder (IEDDA) cycloaddition between tetrazine and *trans*-cyclooctene, **Scheme 11**.^{152,153}



Scheme 11: Synthesis of PROTACs using cycloaddition reactions.

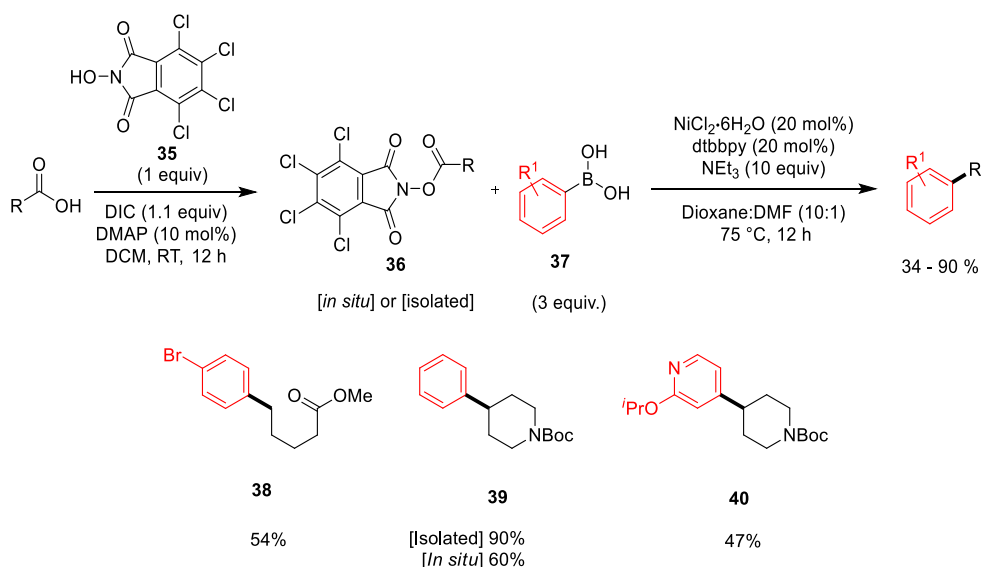
These click reactions can be performed in a cellular environment and Heightman and co-workers showed that the precursors **33** and **34** could penetrate a cell, couple to form the active PROTAC and successfully degrade Brd4 protein.¹⁵³ This was further validated as no degradation was observed with treatment of the pre-formed PROTAC, presumably due to its poor permeability. This approach to PROTAC synthesis takes advantage of the enhanced physicochemical properties and permeability of the PROTAC precursors to maximise the amount of PROTAC that enters the cell and nucleus. However, it also limits the functionality of the linker to the product of the cycloadditions. As the linker plays a key role in determining the orientation of the target protein and E3 ligase in the ternary complex through PPIs, this inherent limitation may affect the success of target protein ubiquitination and subsequent degradation.

The application of modern synthetic methodologies to the synthesis of PROTACs may expand the chemical space tolerated and covered by the linkers and allow for the synthesis of PROTACs in a high-throughput manner.

3.1.2 Decarboxylative Cross-Coupling

In recent years, transition metal-catalysed decarboxylative cross-coupling has emerged as a powerful synthetic strategy for forming carbon-carbon bonds, from carboxylic acids that are available at low cost in great structural diversity.¹⁵⁴

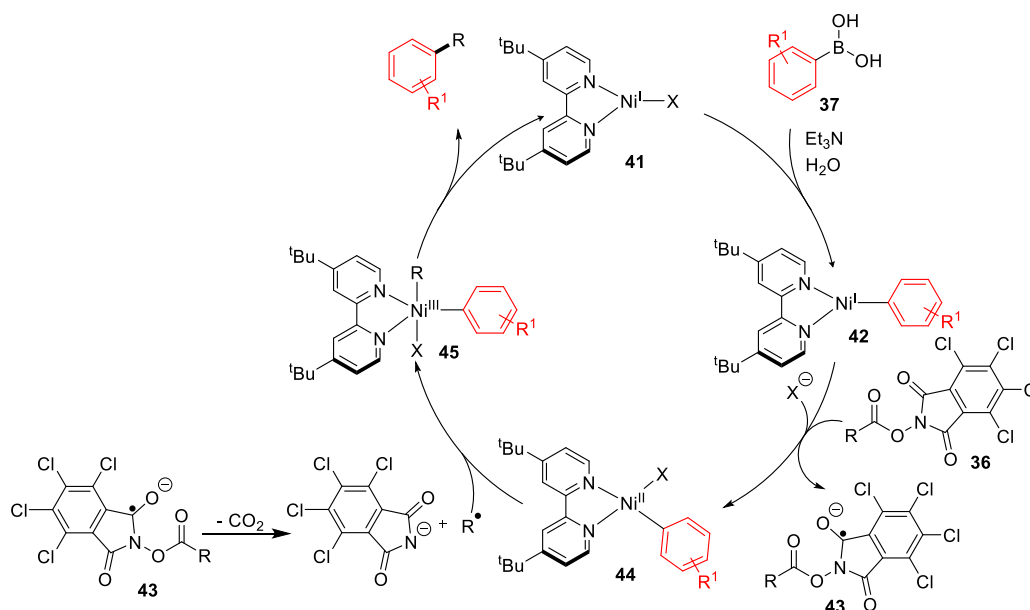
In 2016, Baran and co-workers reported the nickel-catalysed decarboxylative cross-coupling of redox-active esters (RAEs) formed from carboxylic acids and *N*-hydroxytetrachlorophthalimide with aromatic and heteroaromatic boronic acids, **Scheme 12**.¹⁵⁵ The RAE can be formed *in situ* or isolated, and the methodology tolerates both primary and secondary carboxylic acids with good functional group tolerance.



Scheme 12: Decarboxylative cross-coupling of RAEs with boronic acids.¹⁵⁵

The proposed mechanism of the reaction provides an interesting insight into the bespoke reaction conditions, **Scheme 13**. The nickel(I) complex **41** is initially formed from nickel(II) through a single-electron transfer, potentially from triethylamine. The first step of the catalytic cycle is a transmetalation with the boronic acid **37** to form the intermediate **42**. The requirement of 10 equivalents of triethylamine and the hexahydrate nickel catalyst suggests that either the boronate complex or the hydroxonickel complex is being formed prior to transmetalation. Additionally, the requirement for 3 equivalents of the boronic acid starting material suggests there is

some level of protodeborylation under the reaction conditions and an excess is required to drive the formation of the desired product.



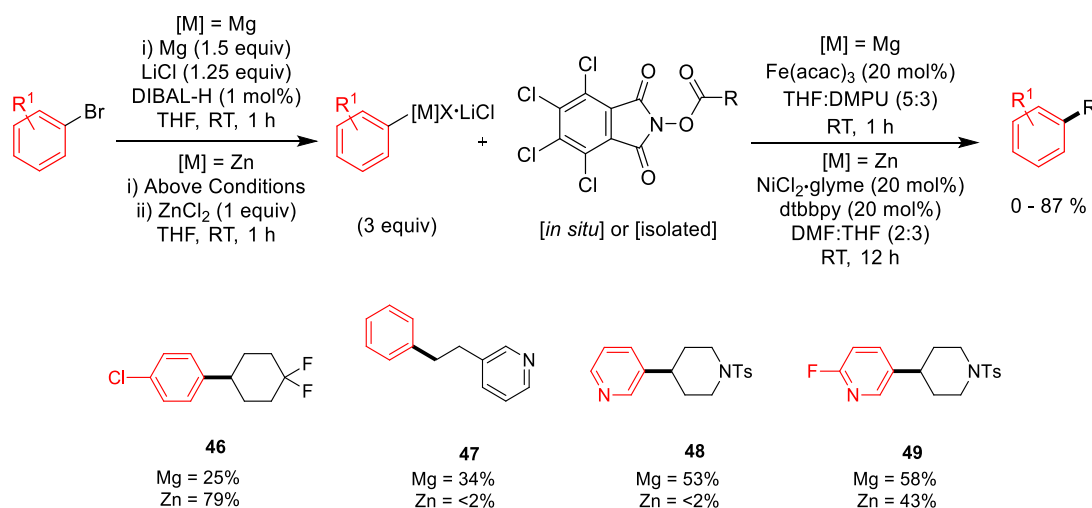
Scheme 13: Proposed mechanism for the decarboxylative cross-coupling of RAEs with boronic acids.¹⁵⁵

A single-electron transfer to **36** forms the nickel(II) complex **44** and the resulting radical anion **43** fragments, with loss of CO₂ to form the tetrachlorophthalimide anion and the R radical. The R radical combines with the nickel(II) complex **44**, forming the nickel(III) complex **45**, which undergoes reductive elimination to furnish the product and complete the catalytic cycle.

One advantage of using this decarboxylative cross-coupling methodology to form PROTACs is that it results in the formation of a carbon-carbon bond. Therefore, no residual amide or amine functional groups are formed, as with the previously described methodologies. This reduces the number of legacy HBDs and HBAs, enhancing the ability to control the effects of these properties on the physicochemical profile of the PROTACs.

A further publication by Baran and co-workers in 2017 extended the decarboxylative cross-coupling of RAEs to Grignard and organozinc reagents, using iron and nickel catalysis respectively, **Scheme 14**.¹⁵⁶ A wide range of yields was reported for this

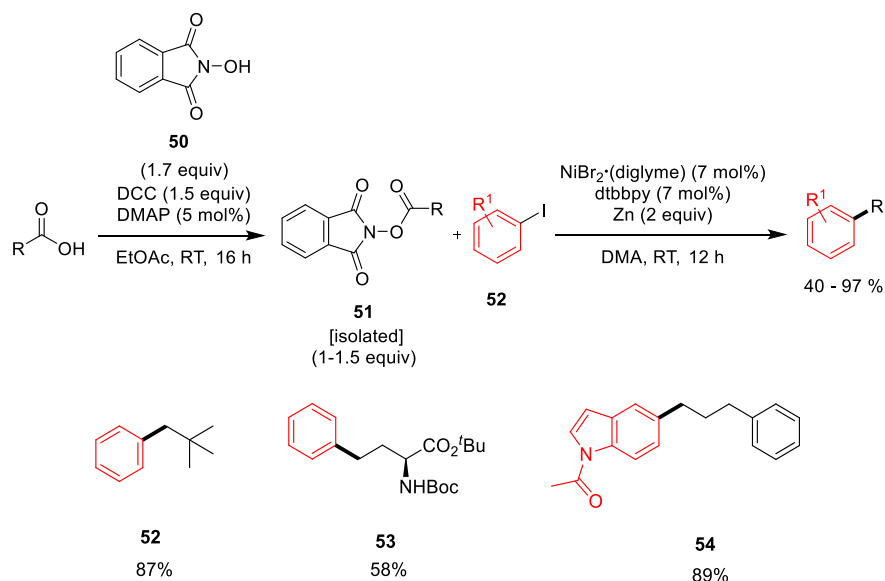
methodology, with few high yielding examples of nitrogen-containing heterocycles. However, this methodology uses aromatic bromides that are abundantly commercially available and tolerates tertiary carboxylic acids. A similar reaction mechanism is proposed as for the decarboxylative cross-coupling of boronic acids, with the transmetalation occurring from either the Grignard or organozinc reagent to the active nickel or iron complex respectively.



Scheme 14: Decarboxylative cross-coupling of RAEs with Grignard and organozinc reagents.¹⁵⁶

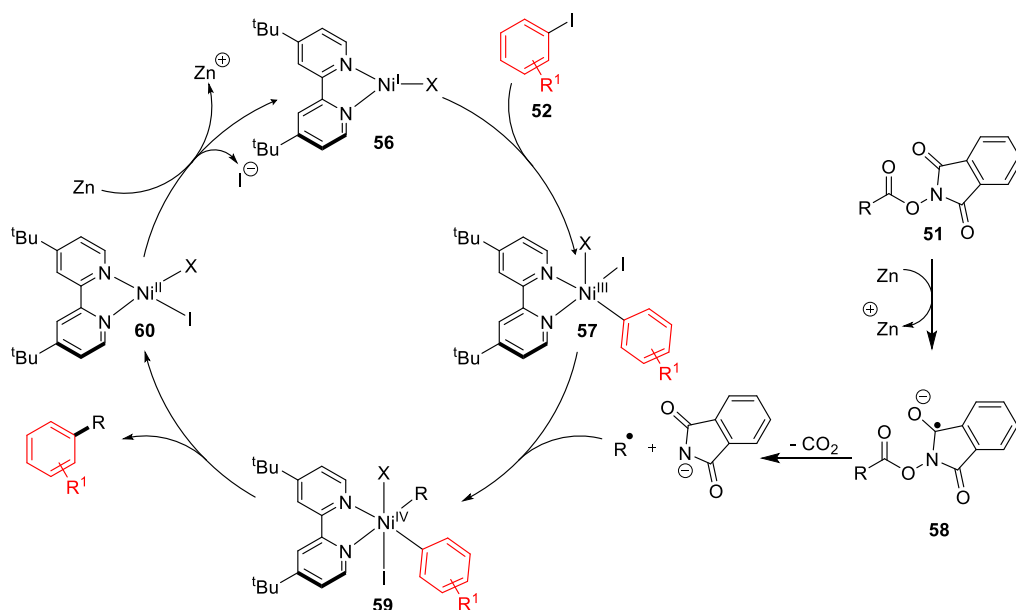
At a similar time, Weix and co-workers published the nickel-catalysed decarboxylative cross-coupling of RAEs formed from carboxylic acids and *N*-hydroxyphthalimide with aryl iodides, **Scheme 15**.¹⁵⁷

The methodology tolerates both primary and secondary carboxylic acids, as well as esters, protected amines and pinacol boronic esters. However, only aryl iodides are tolerated, with no examples of aryl bromides or chlorides. Additionally, there is only one example of coupling with a heteroaromatic iodide.



Scheme 15: Decarboxylative cross-coupling of RAEs with aryl iodides.¹⁵⁷

One drawback to this method is that the reactions are set up in a glovebox, due to the requirement for inert conditions. Additionally, there is a requirement of 2 equivalents of zinc metal as, instead of forming an organozinc reagent through insertion of the zinc between the aromatic-iodine bond, it is acting as a single-electron reducing agent. Although a definitive mechanism has not been elucidated, based on the preliminary mechanistic studies performed by Weix and co-workers, the catalytic cycle in **Scheme 16** is proposed. The nickel(I) complex **56** is initially formed from nickel(II) through a single-electron transfer, potentially from either Zn^0 or **51**. The first step of the catalytic cycle is an oxidative addition of the aryl iodide **52** to form the nickel(III) complex **57**. As this methodology only tolerates aryl iodides, it suggests that the oxidative addition is the rate-limiting step, as substrates with stronger aryl-halide bonds are not tolerated. A single-electron transfer from Zn^0 to **51** forms the radical anion **58**, which fragments with loss of CO_2 to form the phthalimide anion and the R radical. The R radical combines with the nickel(III) complex **57**, forming the nickel(IV) complex **59**, which undergoes reductive elimination to furnish the product and the nickel(II) complex **56**. A further single-electron transfer from Zn^0 completes the catalytic cycle.



Scheme 16: Proposed mechanism for the decarboxylative cross-coupling of RAEs with aryl iodides.

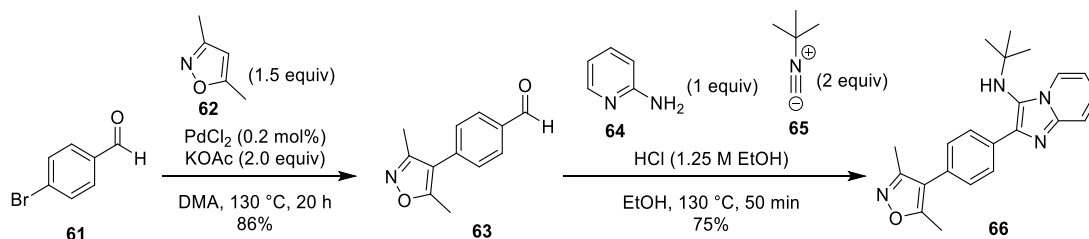
Application of these modern synthetic methodologies to the synthesis of PROTACs may expand the chemical space tolerated and covered by the linkers. As well as allowing the synthesis of PROTACs in more favourable physicochemical space in a high-throughput manner.

3.2 Synthetic Strategy

To investigate the effects of linker functionality on the physicochemical properties of PROTACs, the identity of the target protein and E3 ligase-binding moieties needs to remain constant. This will allow any changes to the physicochemical properties to be independently attributed to changes in the linker functionality. The BET proteins have been of interest to GSK, amongst other pharmaceutical companies, over recent years, as dysregulation of their activity has been strongly linked to cancer, autoimmune and inflammatory diseases. Pan-BET inhibitors exhibit similar binding affinities for all BET isoforms. However, potency and degradation data will be reported for Brd4 only, as this has been highlighted as the isoform with the most therapeutic potential.

The pan-BET inhibitor **66** was chosen from the literature as the model target protein binder for this project as it could be easily accessed following the synthetic route

outlined in **Scheme 17**.¹³⁷ It was envisaged that the requisite functional groups for the decarboxylative cross-coupling could be easily installed using commercially available derivatives of pyridine **64**. The aldehyde **63** was formed *via* a palladium-catalysed C-H activation of isoxazole **62** using 0.2 mol% catalyst loading. Aldehyde **63** was then used in an acid-catalysed multicomponent Groebke–Blackburn–Bienaymé reaction (GBBR) developed in the Flexible Discovery Unit at GSK Stevenage to synthesise isoxazole **66** in 75% isolated yield.¹⁵⁸



Scheme 17: Synthesis of isoxazole **66** via the GBBR.

Isoxazole **66** is a potent BET inhibitor with a Brd4 BD1/BD2 FRET pIC₅₀ of 6.6/6.3 and a moderate physicochemical profile in terms of solubility and permeability, **Figure 29**.

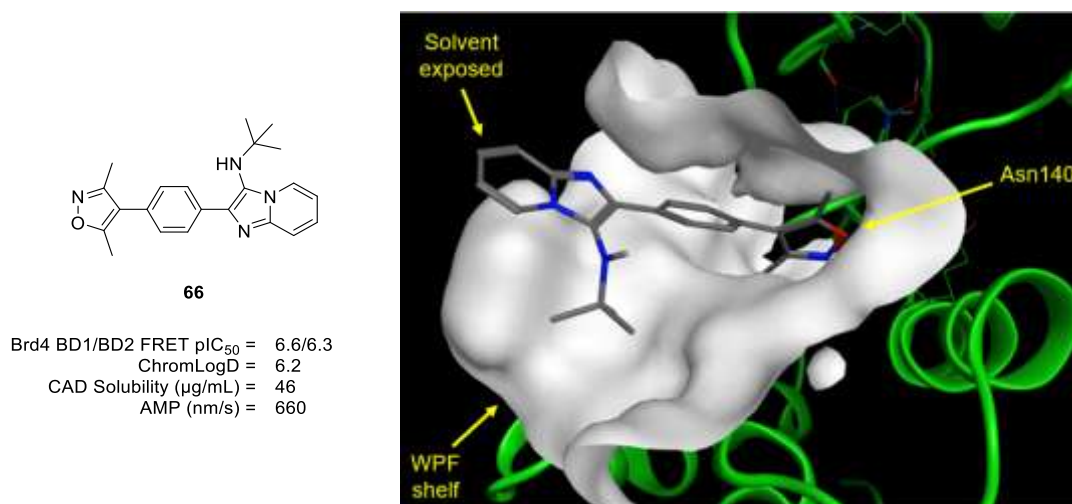
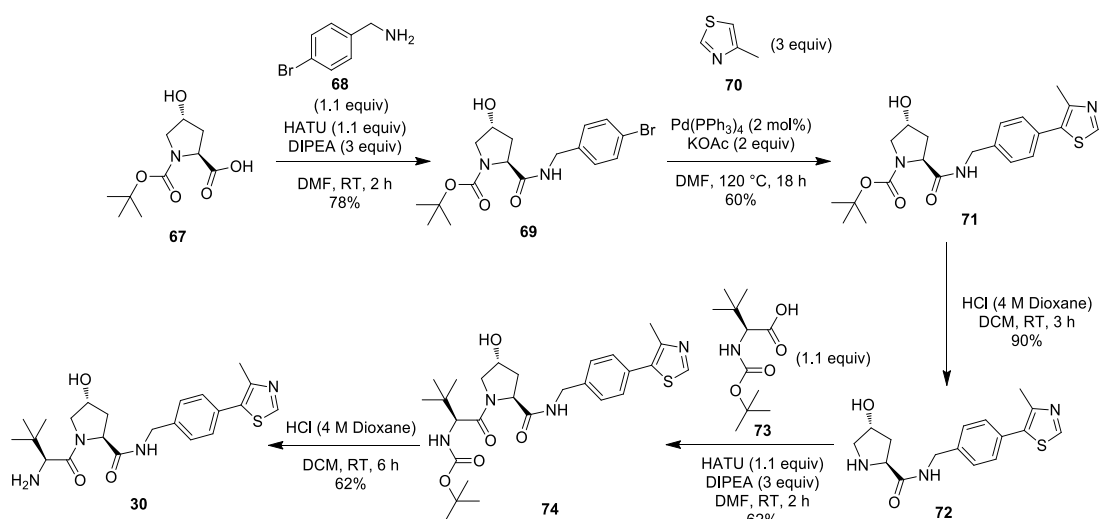


Figure 29: (Left) Brd4 potency and physicochemical data for isoxazole **66**. (Right) Docking model of isoxazole **66** in Brd4 BD1.

The crystal structure for isoxazole **66** has been solved for Brd4 BD1. The key binding interactions include a hydrogen bond from Asn140 deep inside the pocket to the

isoxazole, mimicking the binding of acetylated lysine residues. There are van der Waals interactions between the ^tBu substituent and the WPF shelf, as well as an edge-to-face π -stacking interaction between the imidazopyridine core and Trp81.

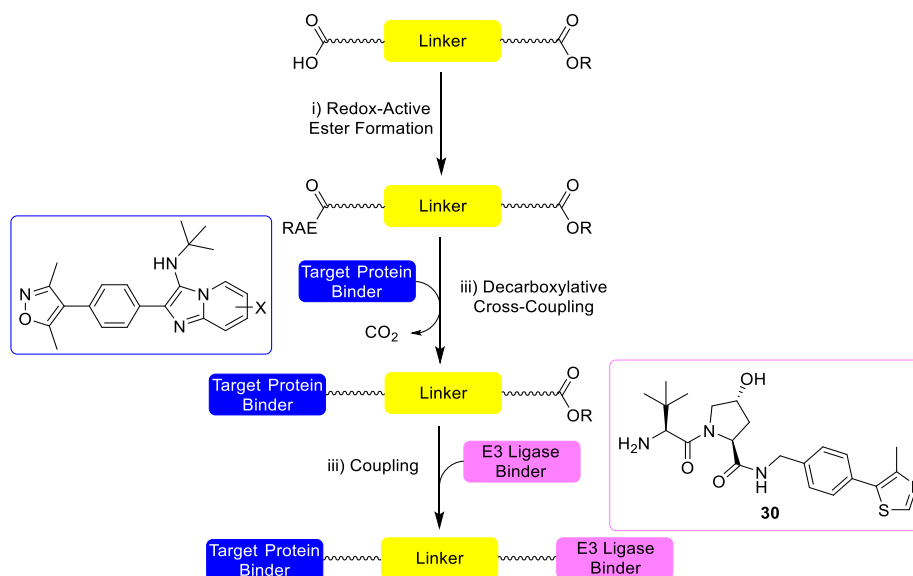
Importantly, another key feature of this crystal structure is that the 6- and 7-carbon of the imidazopyridine ring are solvent-exposed and can provide a vector for linking to an E3 ligase binder. Due to the poor tractability of E3 ligase-binding sites, very few small molecule E3 ligase binders have been developed. The VHL E3 ligase binder **30** is the most commonly used in PROTACs and has the most therapeutic potential. VHL E3 Ligase binder **30** has a terminal amine that has been successfully linked to a range of protein binders to form PROTACs. However, these PROTACs generally exhibit inherently poor physicochemical properties, thus amine **30** was chosen as the E3 ligase binder for further investigation. VHL E3 ligase binder **30** was synthesised following the route outlined in **Scheme 18**.



Scheme 18: Synthesis of VHL E3 ligase binder **30** from hydroxyproline **67**.

This project offered the exciting opportunity to assess the applicability of modern decarboxylative cross-coupling methodologies to the synthesis of PROTACs, ultimately enabling the investigation of the effects of linker functionality on the physicochemical properties of PROTACs, as proposed in the project aims (**Section 2**). It was envisaged that this could be achieved with the linker strategy in **Scheme 19**, using bifunctional linkers with terminal carboxylic acid and ester functionality. These

functional groups exhibit orthogonal reactivity under decarboxylative cross-coupling reaction conditions, facilitating the synthesis of PROTACs in an iterative manner, allowing for the high-throughput screening of a plethora of linkers.

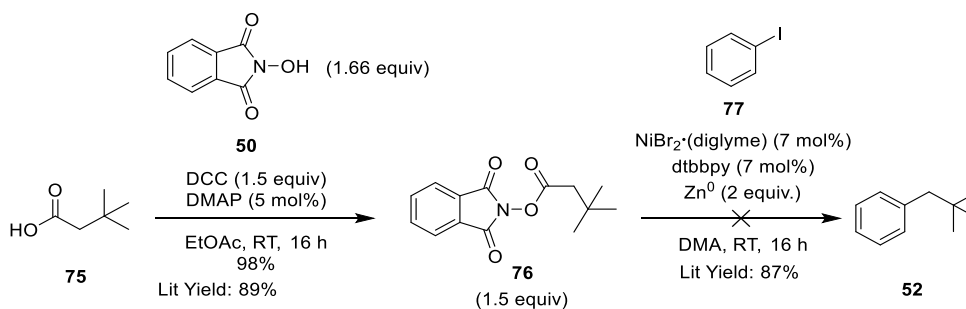


Scheme 19: Linker strategy for the investigation of the effects of linker functionality on the physicochemical properties of PROTACs.

Prior to assessing the applicability of modern decarboxylative cross-coupling methodologies to the synthesis of PROTACs, a selection of examples from the literature was repeated to validate the chemistry.

3.3 Decarboxylative Cross-Coupling Methodologies

The first methodology assessed was the decarboxylative cross-coupling of aryl iodides with primary RAEs published by Weix and co-workers, due to the ability to directly employ the aryl halide, **Scheme 20**.¹⁵⁷ The coupling of carboxylic acid **75** with *N*-hydroxyphthalimide **50** following the described literature procedure furnished the desired RAE **76** in a comparable isolated yield of 98%.



Scheme 20: Decarboxylative cross-coupling of aryl iodides with primary RAEs using the methodology developed by Weix and co-workers.¹⁵⁷

RAE **76** is stable to column chromatography and, once isolated, can be stored at ambient temperature for over four weeks without any observable degradation. The reaction of carboxylic acid **75** and *N*-hydroxyphthalimide **50** was repeated using the carbodiimide coupling reagent EDC. The byproduct of this coupling reagent is water soluble and it was envisaged that an aqueous workup would furnish RAE **76** in an equivalent purity without column chromatography. However, EDC gave a lower isolated yield of 35%.

The decarboxylative cross-coupling of RAE **76** with aryl iodide **77** was performed following the described literature procedure in a glovebox using oven-dried glassware. The active catalyst is pre-formed by stirring NiBr₂·(diglyme) with dtbbpy in DMA, before being added to a solution of RAE **76**, aryl iodide **77** and zinc in DMA, which is warmed at 28 °C for 16 h. The reaction was monitored at regular intervals and, after 16 h, 2% desired product was observed by LCMS. This is significantly lower than the published yield of 87%.

The reaction was repeated using new bottles of catalyst and ligand, as well as at 50 °C, with no increase in yield. All solvents and the reaction mixture were degassed by three freeze-pump-thaw cycles prior to use in oven-dried glassware in the glovebox to exclude water and oxygen from the reaction. However, Weix and co-workers reported an 81% isolated yield of product **52** when the reaction was setup and performed outside of the glovebox on a 6 mmol scale, suggesting that this was not the main factor limiting the yield.

The literature procedure specifically uses 325 mesh zinc flakes purchased from Alfa Aesar, which have a particle size of $< 40 \mu\text{m}$ and had been discontinued prior to the start of this project. Therefore, the initial optimisation reactions were performed using 40-60 nm zinc powder from Sigma Aldrich, as it was envisaged that an increased surface area would increase the rate of single-electron transfer from the zinc. A range of zinc sources and particle sizes was screened, using 20 mol% pre-formed $\text{NiBr}_2 \cdot (\text{dtbbpy})$, **Table 1**.

Table 1: Zinc sources screened for the decarboxylative cross-coupling of RAE **76** and aryl iodide **77**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Zinc Source	Particle Size	52 / %
1	Sigma Aldrich	40-60 nm	2
2	Sigma Aldrich	$< 10 \mu\text{m}$	3
3	Acros Organics	Dust/Not Specified	2
4	Sigma Aldrich	Powder/Not Specified	0

The results in **Table 1** show that the zinc dust with a particle size $< 10 \mu\text{m}$ gave the highest LCMS peak area of 3%. However, this negligible increase is within the error of the LCMS integration. A selection of nickel catalysts was screened using the improved conditions with zinc dust ($< 10 \mu\text{m}$), **Table 2**.

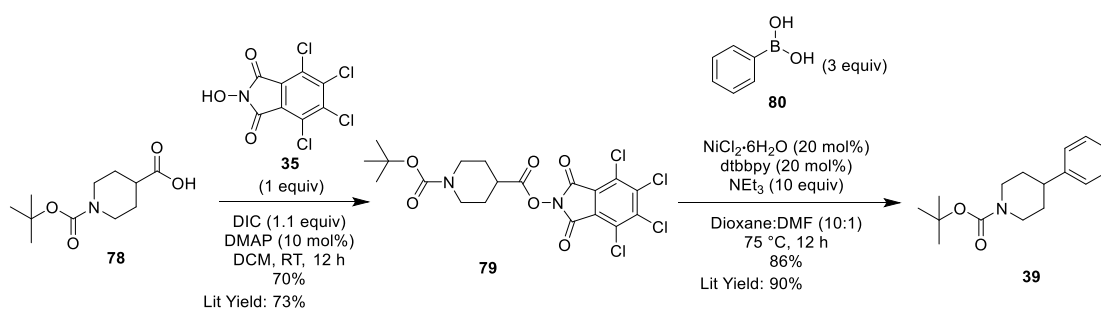
Table 2: Nickel sources screened for the decarboxylative cross-coupling of RAE **76** and aryl iodide **77**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Catalyst	52 / %
1	$\text{NiBr}_2 \cdot (\text{diglyme})$	3
2	NiBr_2	0
3	$\text{NiBr}_2 \cdot x\text{H}_2\text{O}$ ($x=1-2$)	0
4	$\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$	0

Formation of the active catalyst results in a distinct colour change to give a blue-green reaction mixture, which was observed for all the catalysts screened. Surprisingly using NiBr_2 (**Table 2**, entry 2) resulted in a decrease in conversion to the desired product, with no product being observed for **Table 2** entries 2, 3 and 4.

Due to the poor yields obtained when repeating literature examples, this methodology was not selected to assess the applicability of modern decarboxylative cross-coupling methodologies to the synthesis of PROTACs.

The next methodology assessed was the decarboxylative cross-coupling of aromatic boronic acids with primary and secondary RAEs published by Baran and co-workers, **Scheme 21**.¹⁵⁵ The coupling of carboxylic acid **78** with *N*-hydroxyphthalimide **35** following the described literature procedure furnished the desired RAE **79** in a comparable isolated yield of 70%.



Scheme 21: Decarboxylative cross-coupling of boronic acid **80** with RAE **79**.

The decarboxylative cross-coupling of RAE **79** with phenylboronic acid **80** was performed following the described literature procedure. The active catalyst is pre-formed by stirring $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$ and dtbbpy in DMF for 3 h at ambient temperature and can be stored for several days without appreciable deterioration.¹⁵⁵ The RAE was added to the reaction vessel and vacuum degassed three times prior to the addition of the remaining reagents and finally the active catalyst solution. The reaction vessel was heated at 75 °C for 12 h and afforded the desired product **39** in 86% isolated yield following purification by column chromatography. The observed yield is comparable to the published yield of 90% and was repeatable.

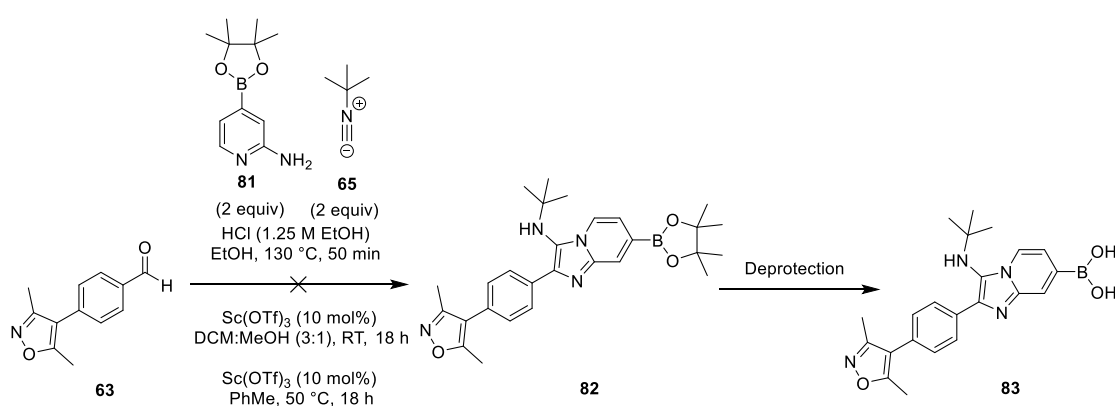
Interestingly, substitution of the catalyst for anhydrous NiCl_2 resulted in a decrease in yield, with no desired product observed by LCMS. This would suggest that 10 equivalents of triethylamine and the hexahydrate nickel catalyst is required for productive transmetalation of the boronic acid *via* the formation of either the boronate or hydroxonickel complex. Additionally, none of the excess boronic acid starting

material was observed by LCMS after 12 h, indicative of a competing protodeborylation reaction and the requirement for 3 equivalents of boronic acid.

After successfully repeating the literature in comparable yields, the applicability of the decarboxylative cross-coupling methodology developed by Baran and co-workers to the synthesis of PROTACs was investigated.

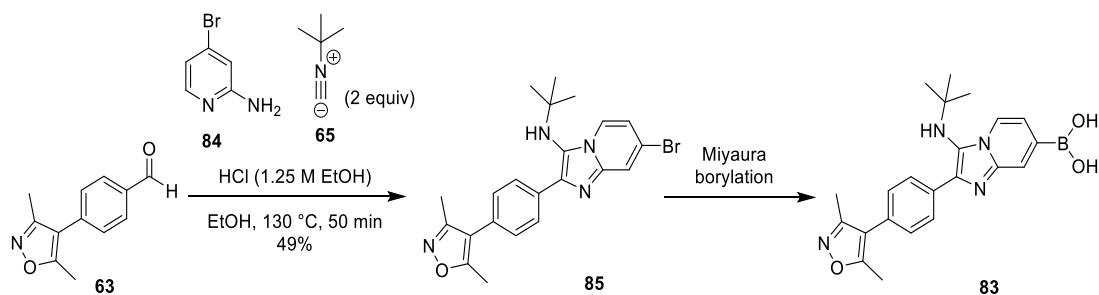
3.4 Synthesis of Boronic Acid Protein Binder

It was envisaged that the boronic acid derivative of the protein binder **83** could be synthesised from the commercially available boron pinacol ester **81** via a GBBR and subsequent deprotection, **Scheme 22**.¹⁵⁸ However, no desired product was formed using the same acid-catalysed conditions used to synthesise the unsubstituted protein binder **66**. Only aldehyde **63** could be observed by LCMS, which suggests that the pinacol ester **82** undergoes protodeborylation under the reaction conditions.



Scheme 22: Synthesis of boron pinacol ester **82** via the GBBR with aldehyde **63**.

Milder Lewis acid-catalysed conditions using Sc(OTf)₃ at both ambient temperature and 50 °C, in 3:1 DCM:MeOH and PhMe respectively, furnished no desired product.¹⁵⁹ It was then proposed that boronic acid **83** could be accessed from the bromide **85** via a Miyaura borylation, **Scheme 23**. The bromide **85** was synthesised using the previously described acid-catalysed GBBR conditions in 49% isolated yield following purification by column chromatography.



Scheme 23: Synthesis of bromide **85** via the GBBR with aldehyde **63**.

However, although the product appeared 99% pure by LCMS, there were a number of uncharacterisable impurities observed in the NMR. A selection of 10 solvents was screened for the trituration of these impurities using the method described in **Section 5.4**. The crude bromide **85** was weighed before and after trituration to obtain an isolated yield, with 0% reported for samples that were impure by NMR, **Table 3**. Both THF and TBME yielded 0% of the clean desired product, with THF completely dissolving the product and impurities and TBME also partially dissolving both. The alcoholic solvents (**Table 3**, entries 3, 4, and 5) gave moderate yields of the clean desired product, with a maximum of 48% with IPA. The ethereal solvents Et₂O and ⁱPr₂O showed improved yields of 75% and 91% respectively. The highest yields were achieved by hydrocarbon solvents (**Table 3**, entries 8 and 9), with a maximum isolated yield of 95% after trituration with cyclohexane. However, repeating the trituration on a larger scale resulted in a lower final isolated yield for the synthesis and purification of bromide **85** of 36%.

Table 3: Solvents screened for the trituration of bromide **85**.

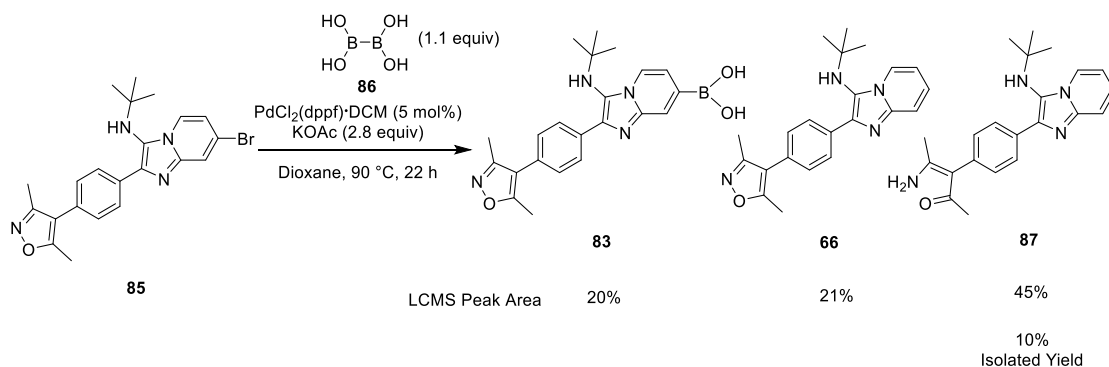
Entry	Solvent	85 Isolated Yield / %
1	THF	0
2	TBME	0
3	MeOH	15
4	EtOH	25
5	IPA	48
6	Et ₂ O	75
7	ⁱ Pr ₂ O	91
8	Pet. Ether	92
9	Cyclohexane	95

Further optimisation of the reaction conditions for the synthesis of bromide **85** was performed, **Table 4**. Increasing the duration of the reaction from 50 min to 120 min resulted in an increased isolated yield of 48%, (**Table 4** entry 2). Further increases in the duration of the reaction did not increase the yield. Increasing the number of equivalents of the amine or aldehyde to 2 equivalents to drive the initial imine formation furnished improved yields of 64% and 66% respectively, (**Table 4** entries 3 and 4).

Table 4: Optimisation of GBBR conditions for the synthesis of bromide **85** at 130 °C with 2 equiv *t*BuNC.

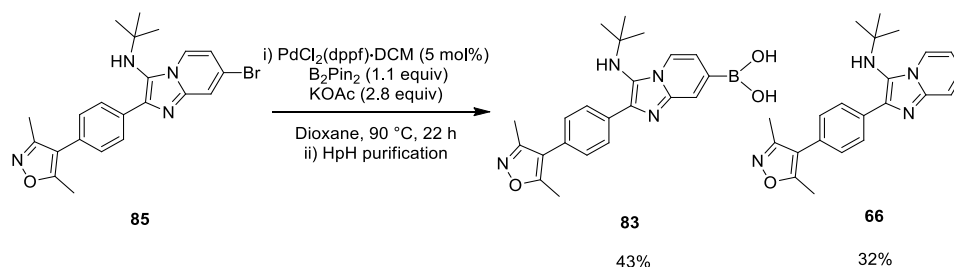
Entry	Duration / min	Amine / equiv	Aldehyde / equiv	85 Isolated Yield / %
1	50	1	1	36
2	120	1	1	48
3	120	1	2	64
4	120	2	1	66

Synthesis of boronic acid **83** directly from bromide **85** by a Miyaura borylation using hypodiboric acid **86** and the conditions in **Scheme 24** was attempted. LCMS analysis of the reaction mixture observed the formation of the desired boronic acid product **83** in 20% by peak area. However, protodeborylated product **66** was formed in 21% by peak area, as well as the isoxazole ring-opened product **87** in 45% by peak area. Purification by column chromatography isolated **87** in 10% yield. Boronic acid **83** and **66** appeared to coelute, however this may be a consequence of further protodeborylation of boronic acid **83** during purification.



Scheme 24: Miyaura borylation of bromide **85**.

Synthesis of boronic acid **83** from bromide **85** by a Miyaura borylation using B_2Pin_2 to access the boron pinacol ester, which could be deprotected to furnish boronic acid **83** was attempted, **Scheme 25**. The pinacol esters are less susceptible to hydrolysis and have a greater potential of being separated from the protodeborylated product.



Scheme 25: Miyaura borylation of bromide **85**.

The pinacol ester was hydrolysed to the boronic acid during purification by HpH reverse phase column chromatography to furnish boronic acid **83** in a 43% isolated yield. Boronic acid **83** was characterised by LCMS and 1H NMR in d_6 -DMSO. However, after 12 h, extensive protodeborylation of boronic acid **83** in d_6 -DMSO was observed. The instability of the boronic acid **83** not only limits its isolation but also its potential use in a subsequent nickel-catalysed decarboxylative cross-coupling, with reaction conditions that are likely to enhance the rate of protodeborylation.

Analysis of the ^{13}C NMR shifts of isoxazole **66** gives an indication of the electron density at each position, with higher chemical shift emanating from a more deshielded environment, indicative of an electron-deficient carbon, **Figure 30**.

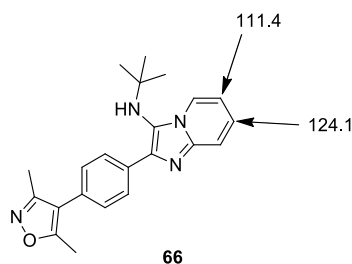
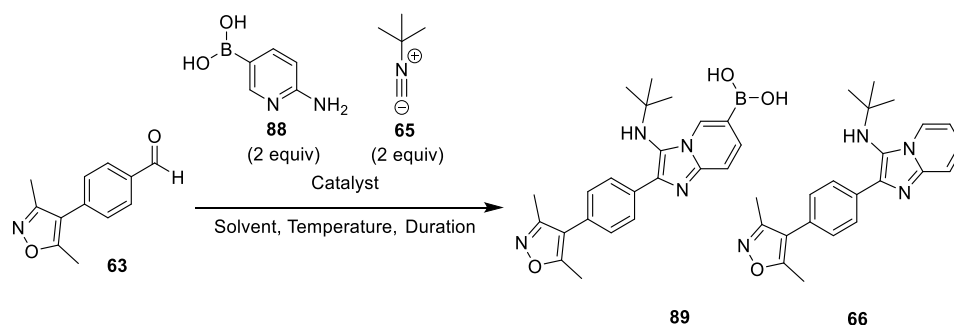


Figure 30: ^{13}C NMR shifts of isoxazole **66** in $CDCl_3$.

The carbon at the 7-position has a ^{13}C NMR shift of 124.1 ppm in $CDCl_3$, indicative of an electron-deficient aromatic carbon. This has the effect of further decreasing the

electron density at the already electron-deficient boronic acid in **83**, potentially enhancing its rate of base-catalysed protodeborylation *via* the formation of the boronate complex.¹⁶⁰ The carbon at the 6-position has a ¹³C NMR shift of 111.4 ppm in CDCl₃, indicative of an aromatic carbon with a higher electron density. It was postulated that a boronic acid at this position would have a higher electron density and thus a decreased propensity for boronate formation and subsequent base-catalysed protodeborylation. As this carbon is also solvent-exposed, it could provide a vector for linking to an E3 ligase binder under nickel-catalysed decarboxylative cross-coupling reaction conditions.

A range of GBBR conditions using both Bronsted acid and Lewis acid catalysis was screened for the formation of boronic acid **89** from commercially available boronic acid **88**, **Scheme 26**, and the outcomes are visualised in **Table 5**.



Scheme 26: Synthesis of boronic acid **89** from boronic acid **88** via the GBBR.

Table 5: Conditions screened for the synthesis of boronic acid **89** from **Scheme 26**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Duration / min	Temperature / °C	Catalyst	Solvent	89 / %	66 / %	63 / %
1	50	RT	Sc(OTf) ₃	DCM:MeOH (3:1)	0	0	100
2	120	50	Sc(OTf) ₃	DCM:MeOH (3:1)	55	18	27
3	120	130	HCl	EtOH	87	13	0
4*	120	130	HCl	EtOH	87	13	0
5	90	130	HCl	EtOH	87	13	0
6	120	120	HCl	EtOH	90	10	0
7	120	110	HCl	EtOH	91	9	0
8	120	110	HCl	Dioxane	0	8	92

* = Performed under a N₂ atmosphere.

The aldehyde **63** is the limiting reagent in this reaction and was used to assess the relative formation of the desired product **89** from the LCMS absorption trace. Forcing conditions with 2 equivalents of boronic acid **88** and isocyanide **65** were used to drive the formation of the desired product. The highest conversions to the cyclised product were observed with acid catalysis (**Table 5**, entries 3-7). No reaction was observed with the Lewis acid $\text{Sc}(\text{OTf})_3$ at ambient temperature. However, increasing the temperature to 50 °C furnished an improved peak area of boronic acid **89** of 55%, with 18% protodeborylated product **66**. Acid-catalysed GBBR conditions (**Table 5**, entry 3) gave an improved formation of boronic acid **89** of 87%, with only 13% protodeborylated product **66**. Performing the reaction both under N_2 (**Table 5**, entry 4) and for a shorter duration of 90 min (**Table 5**, entry 5) resulted in no change in formation of boronic acid **89**. Decreasing the temperature from 130 °C (**Table 5**, entry 3) to 120 °C (**Table 5**, entry 6) and 110 °C (**Table 5**, entry 7) furnished the highest yields of boronic acid **89** of 90% and 91% respectively by peak area. A concurrent decrease in protodeborylation was observed, with complete consumption of the starting material aldehyde **63**. However, further decreases in temperature resulted in incomplete consumption of aldehyde **63**. A solvent swap from EtOH (**Table 5**, entry 7) to dioxane (**Table 5**, entry 8) resulted in no observable formation of the desired boronic acid **89** and 8% formation of protodeborylated product **66**.

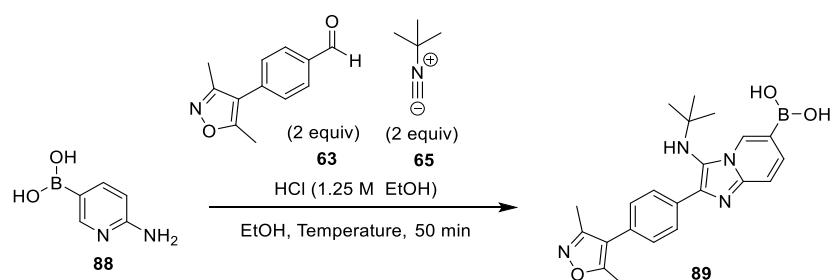
Table 5 entry 7 provided the optimum conditions with full consumption of the starting material aldehyde **63** and the maximum formation of boronic acid **89** of 91% by LCMS peak area. TLC analysis of the reaction mixture to generate conditions for purification by normal phase column chromatography was performed. A range of solvent combinations were screened, with the most forcing conditions of 30% MeOH in EtOAc and 30% MeOH in DCM in conjunction with acetic acid and triethylamine additives. Analysis of the TLC plates with UV light, KMnO_4 staining and TLC-MS, observed no movement of the desired boronic acid **89** from the baseline. Purification by HpH reverse phase column chromatography was performed and the boronic acid **89** eluted between 70-80% MeCN in 10 mM ammonium bicarbonate solution (pH 10), furnishing **89** in 71% isolated yield. Protodeborylated product **66** was present as a minor impurity as it appeared to coelute closely after **89**. However, the lower isolated

yield of boronic acid **89** in comparison to its peak area by LCMS may be due to base-catalysed protodeborylation under the aqueous column conditions.

Recrystallisation of the product from the reaction mixture by the addition of Et₂O as an antisolvent and cooling to -78 °C was unsuccessful. The solvent was removed under reduced pressure and the crude product was dissolved in water, acidified to pH 1 and washed with EtOAc and then cyclohexane. It was envisaged that the boronic acid product would be protonated at this pH and thus show a higher affinity for the aqueous layer, whilst impurities would be extracted in the washings.¹⁶¹ The aqueous layer was basified to pH 5 and boronic acid **89** precipitated as a white solid and was isolated by filtration in 23% yield. This yield is significantly lower than the observed peak area by LCMS due to the protodeborylation of boronic acid **89** under the aqueous extraction conditions.

Trapping of the boronic acid product using an aminopropyl SPE cartridge was successful and furnished **89** in the highest isolated yield at 81%, with no protodeborylated product **66** observed. However, the excess boronic acid starting material **88** was also trapped by the aminopropyl SPE cartridge and present as a minor impurity, which could compete with boronic acid **89** in the subsequent decarboxylative cross-coupling reactions.

It was envisaged that making boronic acid **88** the limiting reagent and using an excess of aldehyde **63** would force complete consumption of boronic acid **88** and improve the purification of boronic acid **89**, **Scheme 27** and **Table 6**.



Scheme 27: Synthesis of boronic acid **89** from boronic acid **88** via the GBBR.

Using the previously employed GBBR conditions (**Table 6**, entry 1) with 2 equivalents of aldehyde **63** and isocyanide **65**, furnished boronic acid **89** in an isolated yield of 90%. As observed when using an excess of boronic acid **88**, an improved formation of boronic acid **89** was obtained at 110 °C (**Table 6**, entry 2), with a maximum isolated yield of 92%. Once isolated, boronic acid **89** is a stable solid that was stored at ambient temperature under noninert conditions without any observable degradation.

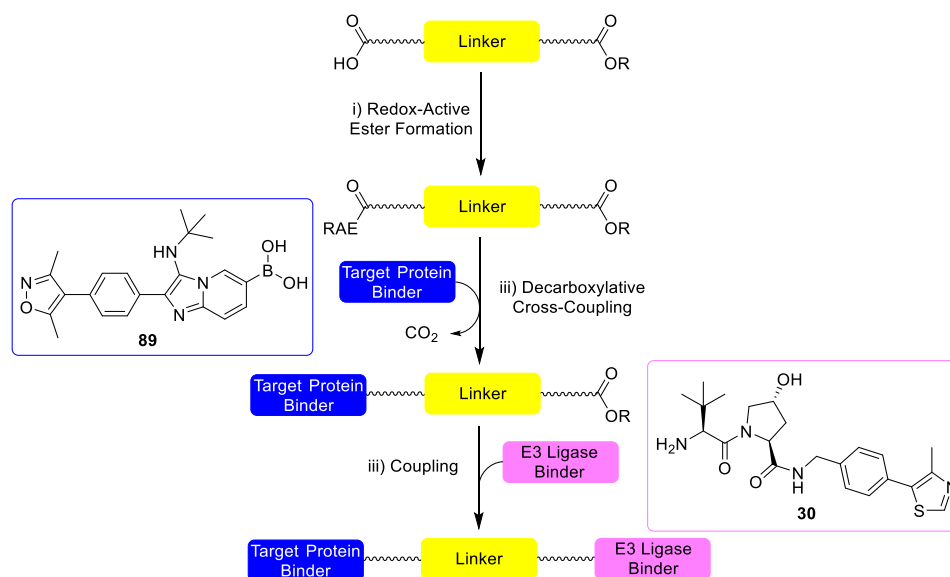
Table 6: Reaction conditions screened for the synthesis of boronic acid **89** from boronic acid **88** via the GBBR.

Entry	Duration / min	Temperature / ° C	Catalyst	Solvent	89 Isolated Yield / %
1	120	130	HCl	EtOH	90
2	120	110	HCl	EtOH	92

3.5 Synthesis of PROTACs by Decarboxylative Cross-Coupling

3.5.1 Decarboxylative Cross-Coupling with a Boronic Acid Protein Binder

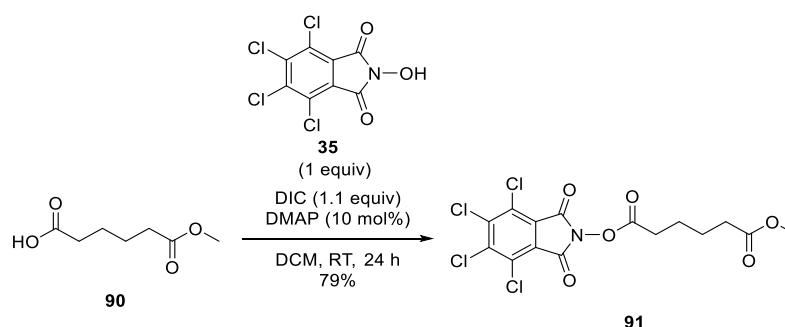
With the boronic acid protein binder **89** in hand, the applicability of the decarboxylative cross-coupling methodology developed by Baran and co-workers to the synthesis of PROTACs could be assessed using the linker strategy in **Scheme 28**.



Scheme 28: Synthetic strategy used to assess the applicability of the methodology developed by Baran and co-workers to the synthesis of PROTACs.

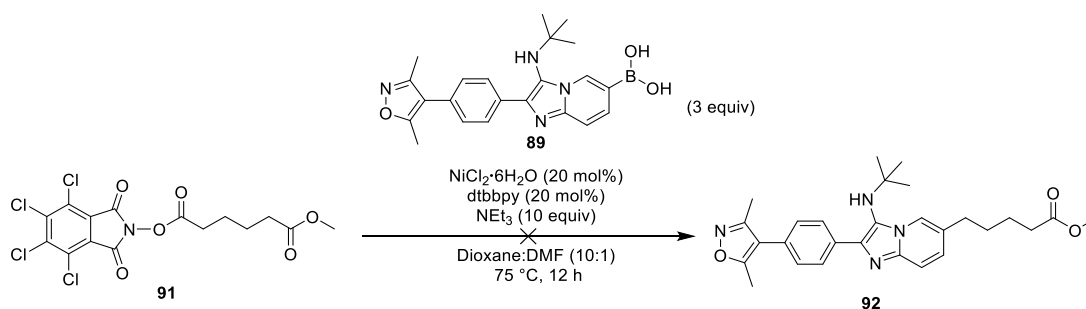
The precursor to the 4-carbon linker **90** was chosen as the model linker for initial investigations as it contains only carbon functionality and was exemplified by Baran and co-workers in published decarboxylative cross-coupling reactions.¹⁵⁵

The carbodiimide-mediated coupling of the model linker **90** with *N*-hydroxyphthalimide **35** furnished the RAE **91** in 79% isolated yield, **Scheme 29**. RAE **91** was stable to purification by column chromatography and was stored at ambient temperature under noninert conditions without any observable degradation.



Scheme 29: Synthesis of RAE **91** from carboxylic acid **90**.

The decarboxylative cross-coupling of RAE **91** with boronic acid **89** was attempted using the standard literature conditions reported by Baran and co-workers, **Scheme 30**.¹⁵⁵ The active catalyst was pre-formed by stirring the nickel catalyst and dtbbpy ligand for 3 h in DMF. The reaction was performed under an inert atmosphere of N₂ and the active catalyst was added to the reaction mixture containing all other reagents in dioxane and immediately heated at 75 °C for 12 h.



Scheme 30: Decarboxylative cross-coupling of RAE **91** and boronic acid **89**.

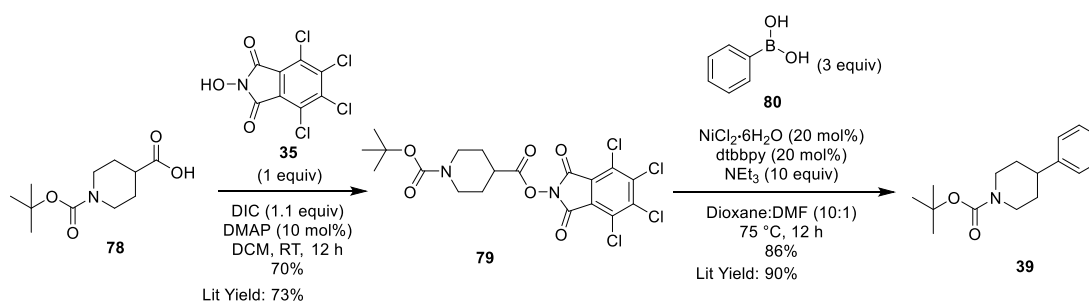
After 12 h and an aqueous work-up, the crude product was analysed by LCMS with no desired product **92** observed. The main product observed was protodeborylated

starting material **66**. Further attempts to force the formation of the desired product **92** by increasing the temperature (**Table 7**, entry 2) and active catalyst loading were unsuccessful and subsequently increased the amount of protodeborylated starting material **66**, **Table 7**.

Table 7: Conditions screened for the decarboxylative cross-coupling of RAE **91** and boronic acid **89** from **Scheme 30**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

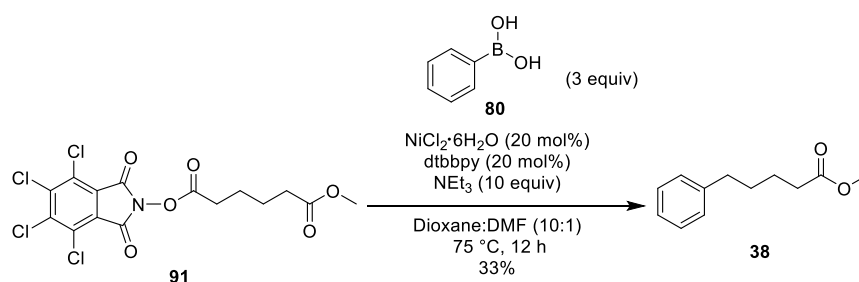
Entry	Temperature / ° C	Catalyst / mol%	Ligand / mol%	66 / %	89 / %	92 / %
1	75	20	20	43	46	0
2	100	20	20	45	26	0
3	75	100	100	80	0	0

A series of experiments was designed to determine why the desired product **92** was not formed in the decarboxylative cross-coupling reaction. First, one of the literature examples by Baran and co-workers was repeated in comparable yield, **Scheme 31**. This confirmed the formation of the active catalyst and that the reagents and inert reaction conditions are not inhibiting the formation of the desired product **92**.



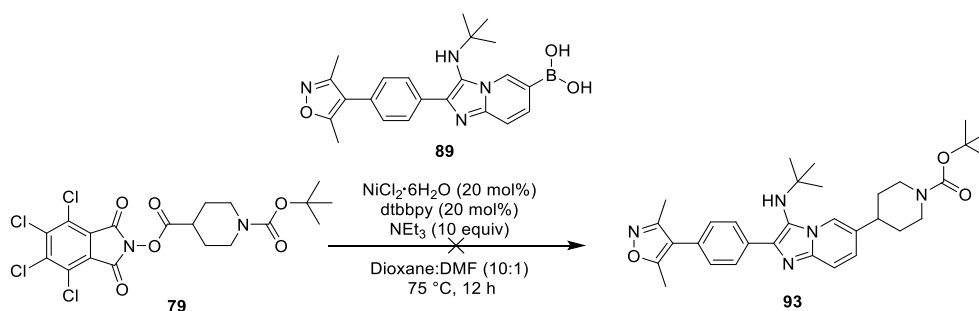
Scheme 31: Repeat of the decarboxylative cross-coupling of RAE **79** with boronic acid **80** reported by Baran and co-workers.

The decarboxylative cross-coupling of the boronic acid **80**, which has been exemplified in decarboxylative cross-coupling reactions by Baran and co-workers, with the model linker RAE **91** furnished the desired product **38** in 33% isolated yield, **Scheme 32**. This confirmed that the RAE **91** is compatible with the decarboxylative cross-coupling reaction conditions.



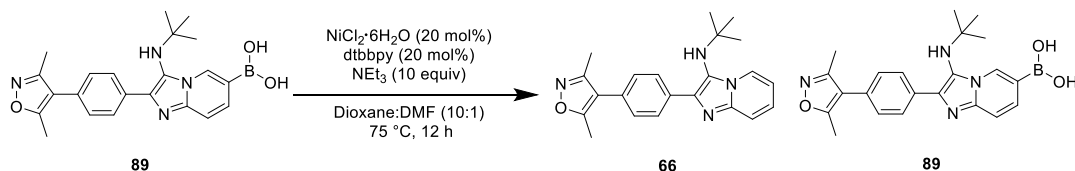
Scheme 32: Decarboxylative cross-coupling of RAE **91** with boronic acid **80**.

The decarboxylative cross-coupling of the RAE **79**, which has been exemplified in decarboxylative cross-coupling reactions by Baran and co-workers, with the boronic acid **89** furnished a 1:1 mixture of boronic acid starting material and protodeborylated starting material **66**, **Scheme 33**. After isolating all possible variables independently, this suggests that the boronic acid **89** is not compatible with the decarboxylative cross-coupling reaction conditions.



Scheme 33: Decarboxylative cross-coupling of RAE **79** and boronic acid **89**.

Stability studies with the boronic acid **89** under the reaction conditions were performed, **Scheme 34** and **Table 8**. The reaction mixtures were analysed by LCMS and reagents systematically and independently excluded to understand their effect on the stability of the boronic acid **89**.



Scheme 34: Stability studies of the boronic acid **89** under decarboxylative cross-coupling reaction conditions.

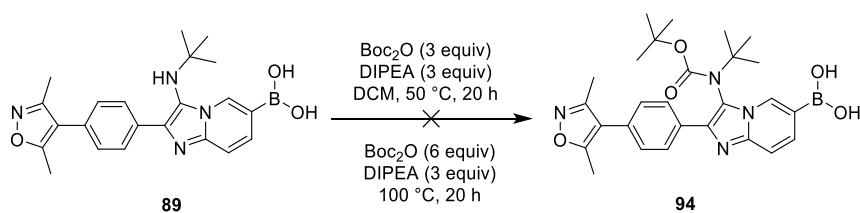
Table 8: Stability studies of the boronic acid **89** under decarboxylative cross-coupling reaction conditions. 10 mg of **89** was stirred under the conditions described at 75 °C for 12 h. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Catalyst	Ligand	Base	66 / %	89 / %
1	✓	✓	✓	95	5
2	✓	✓	✗	85	15
3	✓	✗	✓	52	48
4	✗	✓	✓	0	100

Stirring the boronic acid **89** under the reaction conditions in the absence of a RAE (**Table 8**, entry 1) yielded the highest amount of protodeborylated starting material **66** at 95%. Excluding the triethylamine base (**Table 8**, entry 2) decreased the amount of protodeborylation to 85%. A significant decrease in the amount of protodeborylation was observed in the absence of ligand (**Table 8**, entry 3), with the formation of only 52% protodeborylated starting material **66** by LCMS peak area. No protodeborylation was observed in the absence of the NiCl₂·6H₂O catalyst (**Table 8**, entry 4). These results suggest that the active catalyst complex is catalysing the protodeborylation of boronic acid **89**.

An in-depth analysis of decarboxylative cross-coupling literature examples published by Baran and co-workers discovered that no substrates with a free NH (primary or secondary amine) have been exemplified for both the boronic acid and RAE. All potential substrates containing a free NH have been Boc protected as amides and thus their nitrogen lone pair is delocalised into the amide. It was hypothesised that the free NH of the secondary amine adjacent to the ^tBu group in boronic acid **89** may coordinate to the catalyst and inhibit the desired coupling.

It was envisaged that protecting this free NH would block this pathway and facilitate the desired coupling, with subsequent deprotection furnishing the desired product. Boc-protection of the free NH with Boc₂O at 50 °C was unsuccessful, with no protected product observed, **Scheme 35**. Performing the reaction neat, using Boc₂O as the solvent and heating at 100 °C yielded no improvement. It was believed that this may be due to the large size of the Boc-protecting group and the steric encumbrance of the free NH.



Scheme 35: Boc protection of the free NH in boronic acid **89**.

The ^tPr derivative of the protein binder **95** has similar potency and physicochemical properties to those of the protein binder **89**, **Figure 31**.¹³⁷ A-values give a measure of steric size and are derived from the energy difference between the sterically preferred cyclohexane conformation with the substituent equatorial and the axial conformation.¹⁶² The A-value of ^tBu is more than two times larger than that of the ⁱPr substituent and it was envisaged that the free NH adjacent to the ⁱPr in protein binder **95** would be more amenable to Boc-protection due to the lower steric encumbrance.

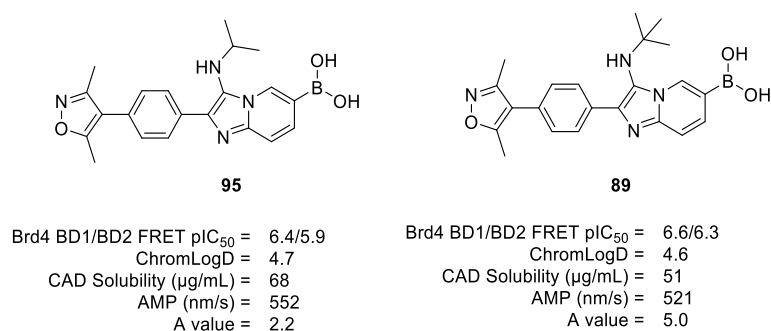
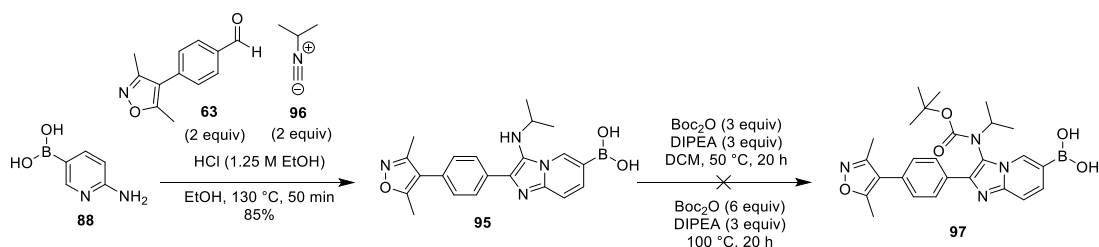


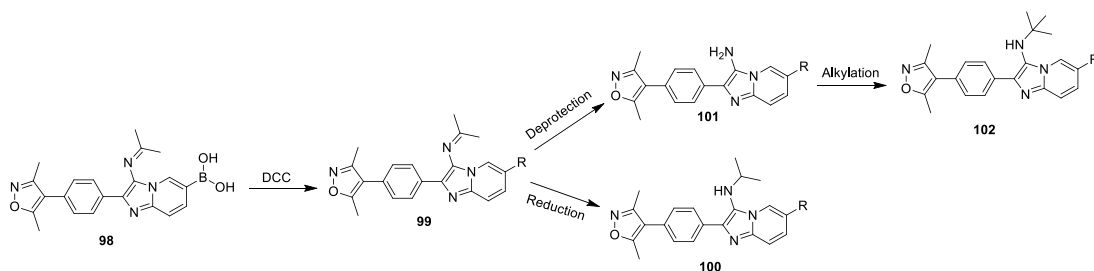
Figure 31: Comparison of the Brd4 potency and physicochemical properties of protein binders **95** and **89**.

Protein binder **95** was synthesised using the acid-catalysed GBBR conditions previously employed, in 85% isolated yield, **Scheme 36**. However, Boc-protection of the free NH in **95** with Boc₂O and forcing conditions was unsuccessful.



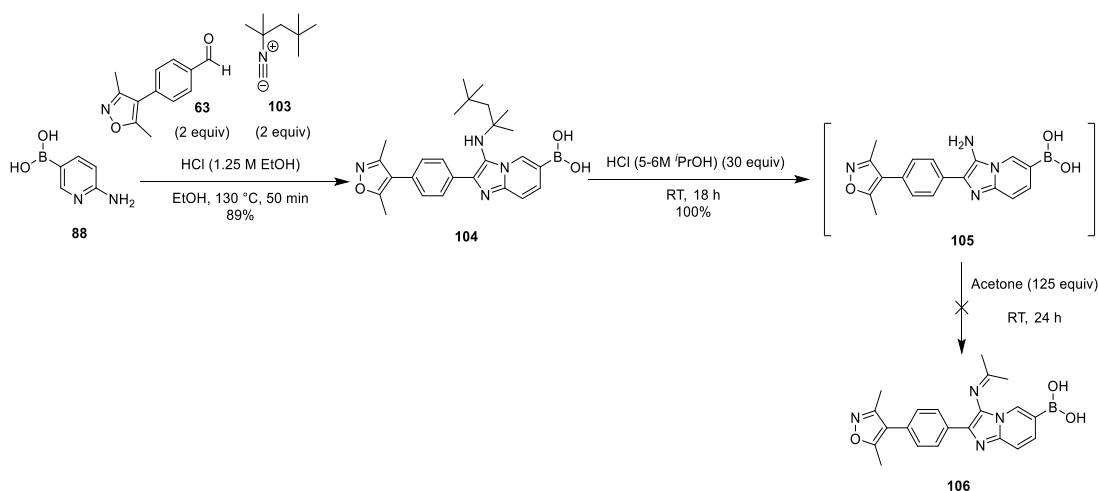
Scheme 36: Synthesis of boronic acid **95** from boronic acid **88** via the GBBR.

It was proposed that protein binder **98**, with the free NH protected as an imine, would be a suitable substrate for the desired decarboxylative cross-coupling reaction. The desired decarboxylative cross-coupling product could then be deprotected to furnish the free amine, which could be alkylated to furnish the *t*Bu protein binder **102**, or reduction of the imine would furnish the *i*Pr protein binder **100**, **Scheme 37**.



Scheme 37: Synthetic strategies to mask the free NH of the protein binder.

Protein binder **104** was synthesised using the acid-catalysed GBBR conditions previously employed, in 89% isolated yield, **Scheme 38**. The *t*-octyl substituent was removed under acidic conditions to furnish protein binder **105** in quantitative yield. However, the desired product **106** was not formed upon addition of acetone to the reaction mixture, *via* acid-catalysed imine formation.

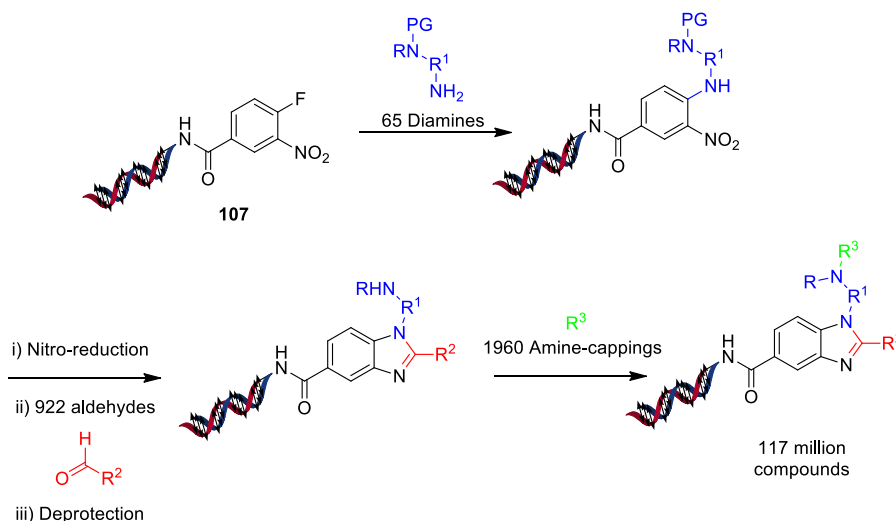


Scheme 38: Synthesis of protein binder **104** from boronic acid **88** via the GBBR.

After extensive investigation and optimisation of both the decarboxylative cross-coupling reaction conditions and the protein binder, it was judged that the series of isoxazole imidazopyridine protein binders that contain a free NH is not amenable

to the modern decarboxylative cross-coupling methodology developed by Baran and co-workers.

In view of this, our aims were directed towards a series of pyridone benzimidazole pan-BET inhibitors developed in a previous BET campaign in the Flexible Discovery Unit at GSK Stevenage. The generic scaffold was identified through an encoded library technology (ELT) screen.¹⁶³ This entailed screening a 6His-Brd4 (1–477) tandem bromodomain construct against a small molecule DNA-encoded library (DEL), using affinity-based selection technology to identify hits. The small molecule DEL was constructed using a split-and-pool strategy with three sets of building blocks to furnish 117 million compounds, **Scheme 39**.¹⁶⁴ Starting from DNA-appended **107**, S_NAr with 65 mono-protected diamine building blocks furnished a pool of benzenediamines. *In situ* nitro reduction and condensation with 922 aldehyde building blocks constructed the benzimidazole ring. Finally, amine deprotection and reaction with 1960 amine-capping groups afforded the complete library. Further DNA tags were added to the appended DNA after each reaction, creating a unique DNA sequence for each compound that can be used to identify it.



Scheme 39: Synthesis of small molecule DEL from DNA-appended **107**.

Analysis of the affinity data identified benzimidazole **108** as a hit, **Figure 32**. No additional enrichment was provided by specific amine-capping groups (R³),

suggesting minimal contribution to the binding. This resulted in the identification of pan-BET inhibitor **109**.

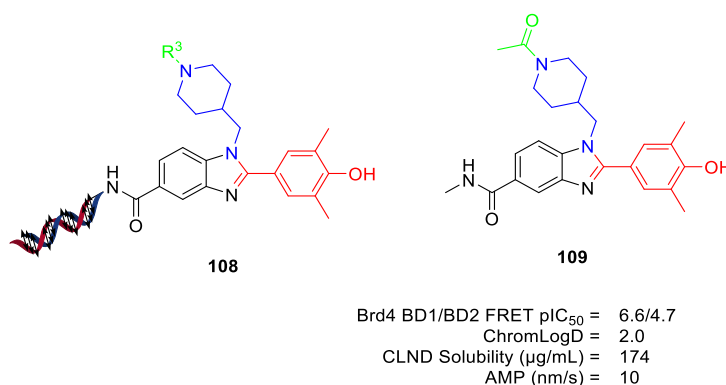


Figure 32: Benzimidazole **108** identified from the ELT screen using the described small molecule DEL and the BET inhibitor **109** that was developed from it.

The X-ray crystal structure of BET inhibitor **109** was solved for Brd4 BD1, **Figure 33**. The 2,6-dimethylphenol group mimics the binding of acetylated lysine residues, with the hydroxyl group accepting a hydrogen bond from Asn140 and donating one *via* a structural water to Tyr97, whilst the methyl substituent *ortho* to the hydroxyl makes van der Waals interactions with the lipophilic pocket adjacent to Phe83. The phenyl ring of benzimidazole makes an edge-to-face π -stacking interaction with Trp81, whilst the aromatic ring nitrogen accepts a hydrogen bond from a structural water that is in contact with Pro82.

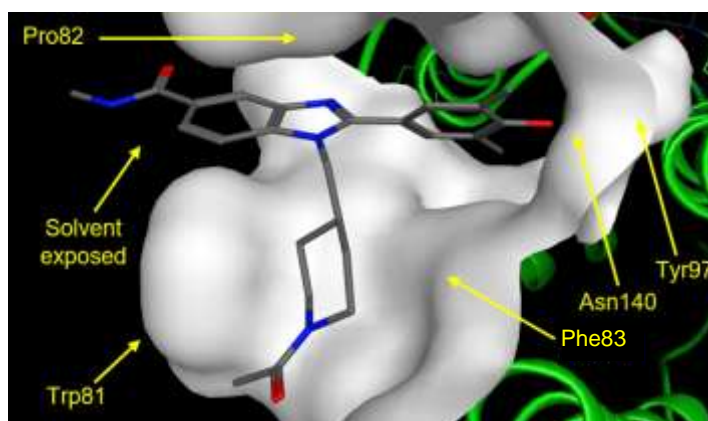


Figure 33: X-ray crystal structure of BET inhibitor **109** in Brd4 BD1.

Due to a propensity of phenols to be rapidly metabolised by phase II enzymes, BET inhibitor **109** exhibited a high clearance and short half-life in *in vivo* pharmacokinetic experiments. Substitution of the 2,6-dimethylphenol group for the analogous acetylated lysine mimetic, 1,3-dimethylpyridin-2(1H)-one, maintained binding affinity whilst decreasing clearance and increasing half-life, **Figure 34**.

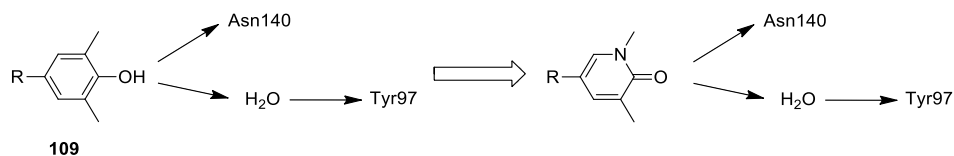


Figure 34: Scaffold swap of acetylated lysine mimetic in BET inhibitor **109** from the 2,6-dimethylphenol group to 1,3-dimethylpyridin-2(1H)-one.

The methylamide substituent of the benzimidazole in BET inhibitor **109** that provided a vector for linking to DNA in the ELT screen is directed out towards solvent. This provides no additional binding affinity and was removed with no observable potency penalty. The piperidine acetamide is pointing out towards solvent as well, consistent with no additional enrichment in the ELT screen with specific amine-capping groups (R³). An extensive SAR campaign around the piperidine diamine portion that makes van der Waals interactions with the WPF shelf identified the dimethoxypropyl substituent. The culmination of these observations and SAR campaigns was the identification of BET inhibitor **110**, a potent pan-BET inhibitor with a favourable balance of physicochemical properties, **Figure 35**.

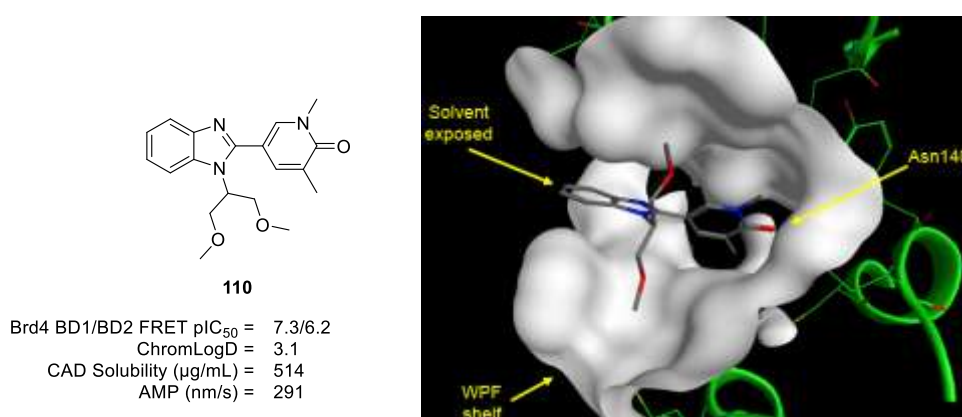
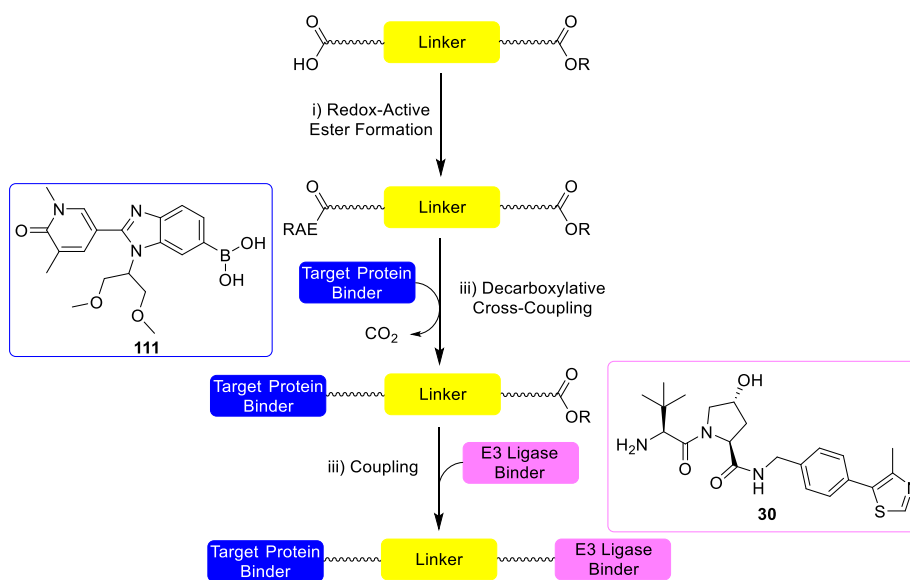


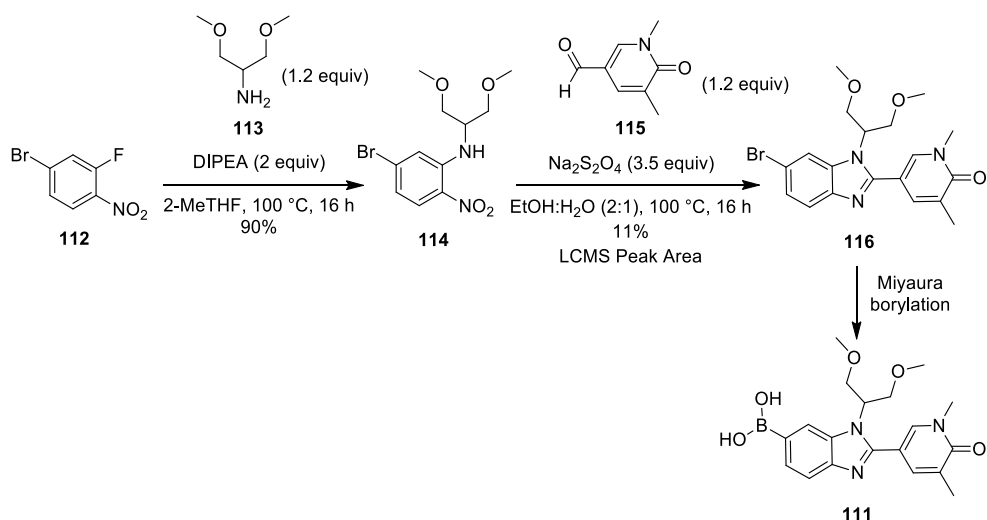
Figure 35: (Left) *Brd4* potency and physicochemical data for BET inhibitor **110**. (Right) Docking model of BET inhibitor **110** in Brd4 BD1.

BET inhibitor **110** contains no free NH groups and the solvent-exposed 5- and 6-carbon of the benzimidazole ring have been previously used to link to large fragments of DNA. It was thus envisaged that BET inhibitor **110** could be used as the target protein-binding moiety to assess the applicability of modern decarboxylative cross-coupling methodologies to the synthesis of PROTACs, by the linker strategy outlined in **Scheme 40**.



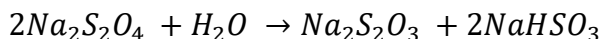
Scheme 40: Synthetic strategy used to assess the applicability of the methodology developed by Baran and co-workers to the synthesis of PROTACs.

It was proposed that boronic acid **111** could be accessed from the bromide **116** via a Miyaura borylation, **Scheme 41**. The S_NAr reaction of aryl fluoride **112** with amine **113** furnished bromide **114** in 91% isolated yield. The *in situ* nitro reduction of **114** with Na₂S₂O₄ and subsequent condensation with aldehyde **115** formed bromide **116** in 11% by LCMS peak area.¹⁶⁵ No other byproducts were formed and both starting materials were observed.



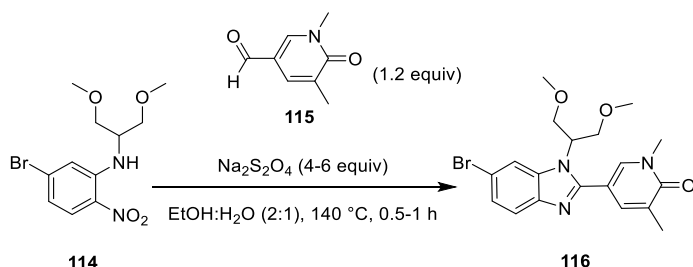
Scheme 41: Synthesis of bromide **116** from aryl fluoride **112** over two steps.

$\text{Na}_2\text{S}_2\text{O}_4$ decomposes under aqueous conditions at elevated temperatures and under acidic conditions, *via* the reaction in **Scheme 42**.¹⁶⁶ This reaction is autocatalytic as NaHSO_3 is weakly acidic ($\text{p}K_a = 6.97$) and therefore promotes further degradation. It was postulated that the $\text{Na}_2\text{S}_2\text{O}_4$ is degrading as the reaction mixture heats up to a temperature that can facilitate the reduction of the nitro group. Therefore, the reaction was repeated in a microwave reactor as this would allow the solution to reach the desired reaction temperature within seconds.



Scheme 42: Decomposition of $\text{Na}_2\text{S}_2\text{O}_4$ under acidic aqueous conditions.¹⁶⁶

All reactions were performed in a Biotage[®] Initiator+ microwave (MW) reactor at 140 °C as this was the maximum temperature recommended for the EtOH:H₂O solvent mixture in the microwave guide, **Scheme 43**.¹⁶⁷



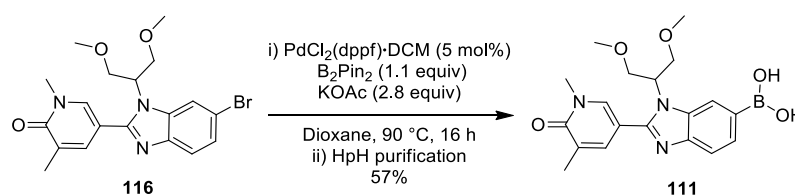
Scheme 43: Synthesis of bromide **116** from bromide **114** in a MW reactor.

Performing the reaction with 6 equivalents of Na₂S₂O₄ and heating at 140 °C for 1 h gave 63% formation of bromide **116** by LCMS peak area, with no starting material **114** observed, (**Table 9**, entry 1). Decreasing the equivalents of Na₂S₂O₄ to 4 (**Table 9**, entry 2), resulted in a significant decrease in formation of bromide **116**, confirming the requirement for a large excess of Na₂S₂O₄. Decreasing the reaction duration to 0.5 h gave the highest formation of bromide **116** of 84% (**Table 9**, entry 3), furnishing the desired product in 77% isolated yield following purification by column chromatography. This suggests that bromide **116** is unstable under the reaction conditions and that a shorter duration is preferable.

Table 9: Reaction conditions screened for the synthesis of bromide **116** from bromide **114** in a MW reactor. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Na ₂ S ₂ O ₄ / equiv	Duration / h	116 Peak Area / %	116 Isolated Yield / %
1	6	1	63	×
2	4	1	59	×
3	6	0.5	84	77

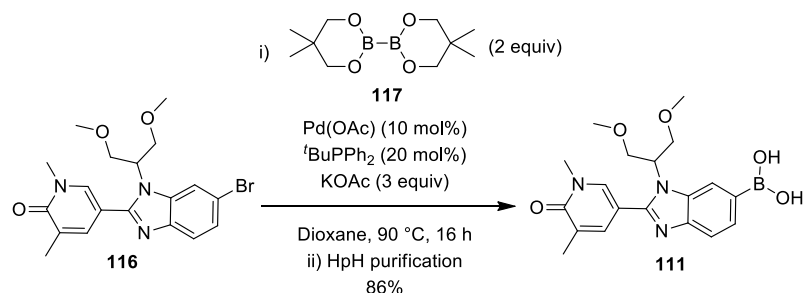
The Miyaura borylation of bromide **116** with B₂Pin₂ was performed using the reaction conditions outlined in **Scheme 44**. A 2:1 mixture of pinacol boronic ester and boronic acid was observed by LCMS after 16 h, with no starting material bromide **114** observed. The pinacol boronic ester was hydrolysed to the desired boronic acid during purification by HpH reverse phase column chromatography to furnish boronic acid **111** in 57% isolated yield.



Scheme 44: Synthesis of boronic acid **111** from bromide **116** via a Miyaura borylation.

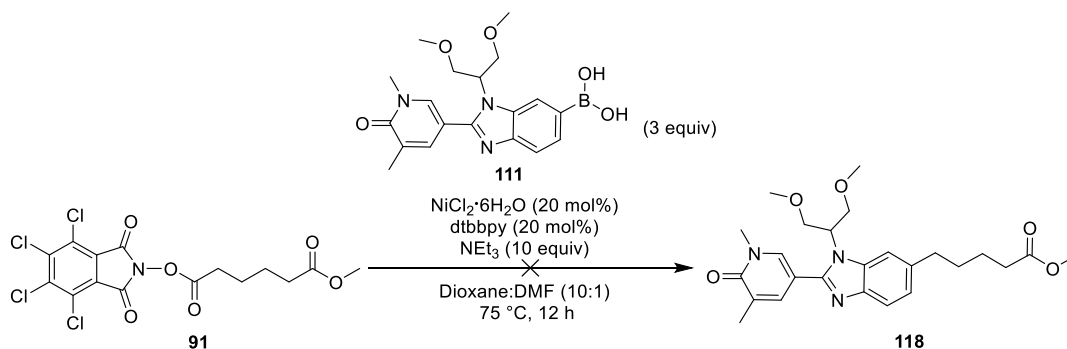
The isolated yield of boronic acid **111** was lower than expected and it was proposed that this may be due to the incomplete deprotection of the pinacol boronic ester under the purification conditions. Repeating the reaction with hypodiboric acid **86**, in an

attempt to avoid the hydrolysis of the pinacol boronic ester, was unsuccessful with no conversion of bromide **116** to boronic acid **111**. Performing the Miyaura borylation of bromide **116** using the bis-(neopentylglycolato)diboron **117**, which is more susceptible to hydrolysis, furnished boronic acid **111** in 86% isolated yield after purification by HpH reverse phase column chromatography, **Scheme 45**.¹⁶⁸



Scheme 45: Synthesis of boronic acid **111** from bromide **116** via a Miyaura borylation.

The decarboxylative cross-coupling of boronic acid **111** with RAE **91** was attempted using the standard literature conditions reported by Baran and co-workers, **Scheme 46**.



Scheme 46: Decarboxylative cross-coupling of boronic acid **111** with RAE **91**.

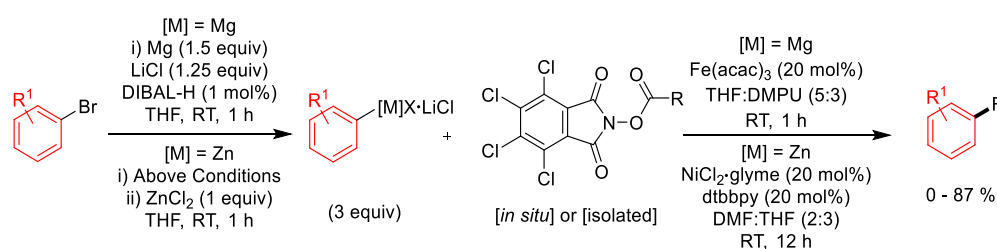
After 12 h and an aqueous work-up, the crude product was analysed by LCMS with no desired product **118** observed. The main peak observed was starting material boronic acid **111**, with protodeborylated starting material **110** present in a 1:6 ratio relative to boronic acid **111**. This confirms that the boronic acid **111** has an increased stability under the decarboxylative cross-coupling reaction conditions in comparison to boronic acid **89**. However, the decarboxylative cross-coupling methodology developed by Baran and co-workers is not amenable to the desired coupling of RAE **91** and boronic acid **111**.

The conditions published by Baran and co-workers have been extensively optimised by screening a plethora of permutations of 8 catalysts, 6 ligands, 12 bases, 23 solvents, 15 additives and a range of temperatures and times to produce the bespoke reaction conditions. Therefore, further optimisation attempts of the reaction conditions were futile.

However, as the bromide **116** had already been synthesised to access boronic acid **111**, the applicability of a complementary decarboxylative cross-coupling methodology developed by Baran and co-workers was investigated.

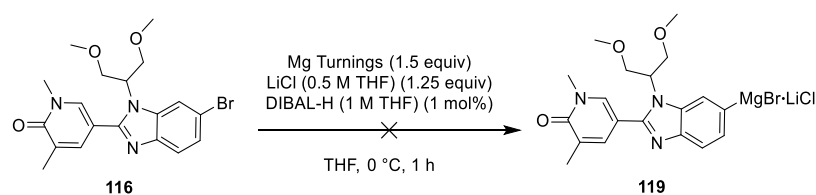
3.5.2 Decarboxylative Cross-Coupling with a Bromide Protein Binder

Baran and co-workers reported the decarboxylative cross-coupling of RAEs with Grignard and organozinc reagents synthesised from the corresponding aromatic bromide, **Scheme 47**.¹⁵⁶



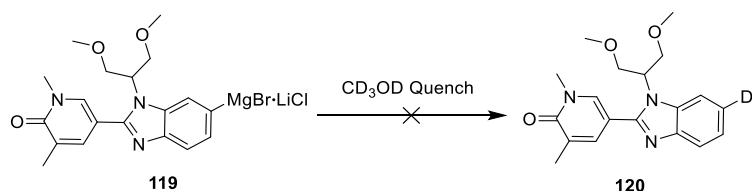
Scheme 47: Decarboxylative cross-coupling of RAEs with Grignard and organozinc reagents synthesised from the corresponding aromatic bromide.

The bromide **116** was synthesised as previously described, **Scheme 42**. The formation of the Grignard reagent **119** was attempted using the conditions reported by Baran and co-workers, **Scheme 48**, in oven-dried glassware under an inert N₂ atmosphere. The magnesium turnings were activated by the addition of DIBAL-H, and LiCl was used to help solubilise the generated organometallic species and thus remove clusters accumulating on the surface of the metal, resulting in more vacant sites for subsequent reactions.¹⁶⁹



Scheme 48: Formation of the Grignard **119** from bromide **116**.

After 1 h at 0 °C, an aliquot of the reaction mixture was titrated with a mixture of iodine and LiCl (0.5 M in THF) solution to determine the molarity of Grignard **119**. However, no colour change was observed upon the addition of a large excess of the reaction mixture, suggesting insufficient formation of the desired Grignard **119**. A further aliquot of the reaction mixture was quenched with deuterated methanol and analysed by LCMS. Only starting material bromide **116** was observed, with no deuterated product that would be indicative of Grignard formation, **Scheme 49**.



Scheme 49: Quenching of Grignard **119** with deuterated methanol.

The reaction was repeated with new bottles of reagents and solvents, with no observable formation of Grignard **119** when titrated with iodine or quenched with deuterated methanol. A range of reaction conditions outlined in **Table 10** were screened for the formation of Grignard **119**. The reaction in **Scheme 48** was repeated with magnesium turnings at ambient temperature with no observable formation of **119** when titrated with iodine or quenched with deuterated methanol (**Table 10**, entry 2). Repeating the reaction with magnesium powder (**Table 10**, entry 3 and 4), which has a larger surface area for the reaction, with the activating agents DIBAL-H and iodine at ambient temperature furnished no observable formation of **119**. Sonicating the magnesium prior to the addition of the activating agent was attempted to remove any magnesium oxide on the metal surface. The reaction was then repeated with DIBAL-H, iodine and 1,2-dibromoethane activating agents at reflux. 1,2-Dibromoethane forms ethene upon activation, which was observed by bubbling in reaction mixture.

However, no formation of Grignard **119** was observable when the reaction mixture was titrated with iodine or quenched with deuterated methanol (**Table 10**, entry 5, 6 and 7). Performing the reaction in the absence of LiCl (0.5 M THF) and grinding the magnesium with a pestle and mortar to remove any magnesium oxide on the metal surface prior to the reaction gave no improvement in Grignard **119** formation, (**Table 10**, entry 8 and 9). Repeating the reaction, using this same magnesium with ultra-dry solid LiCl, enabling an increase in the concentration of the reaction, resulted in no observable formation of Grignard **119**, (**Table 10**, entry 10 and 11).

Table 10: Conditions screened for formation of the Grignard **119** from bromide **116**.

Entry	Magnesium Source	Activating Agent	Temperature
1	Turnings	DIBAL-H	0 °C
2	Turnings	DIBAL-H	RT
3	Powder	DIBAL-H	RT
4	Powder	Iodine	RT
5 ^a	Turnings	DIBAL-H	Reflux
6 ^a	Turnings	Iodine	Reflux
7 ^a	Turnings	1,2-Dibromoethane	Reflux
8 ^{b, c}	Turnings	DIBAL-H	Reflux
9 ^{b, c}	Turnings	Iodine	Reflux
10 ^{b, d}	Turnings	DIBAL-H	Reflux
11 ^{b, d}	Turnings	Iodine	Reflux

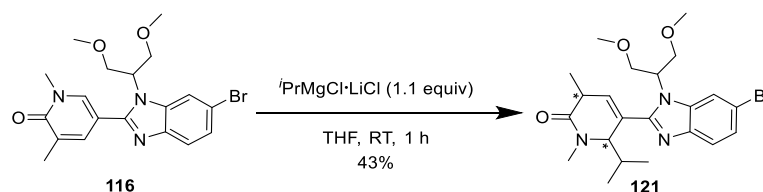
a = Sonication of the reaction mixture for 5 min

b = Magnesium ground using a pestle and mortar

c = No LiCl (0.5 M THF) solution

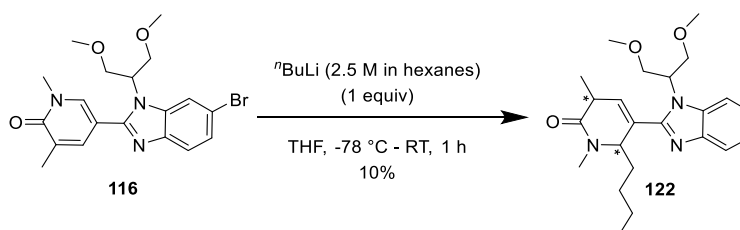
d = Ultra-dry solid LiCl

The formation of Grignard **119** through metal-halogen exchange using a Grignard reagent was attempted, **Scheme 50**.¹⁷⁰ However, Turbo Grignard added at the γ -position of the $\alpha,\beta,\delta,\gamma$ -unsaturated carbonyl of the pyridone core, to furnish bromide **121** as an approximate 1:1.4 diastereomeric mixture in 43% isolated yield.



Scheme 50: Addition of Turbo Grignard to bromide **116**.

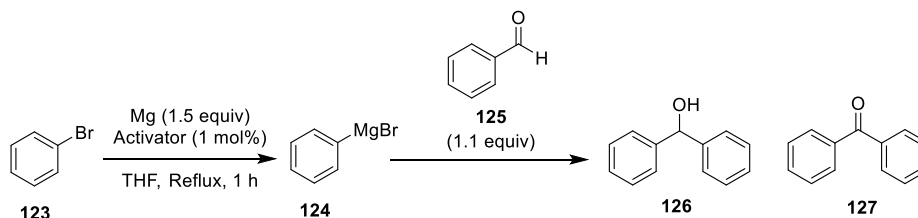
The formation of Grignard **119** through metal-halogen exchange was repeated using organolithium reagents, which are harder nucleophiles than Grignard reagents and therefore have a lower propensity for conjugate addition. However, *n*-butyllithium also added at the γ -position of the $\alpha,\beta,\delta,\gamma$ -unsaturated carbonyl of the pyridone core, to furnish bromide **122** as an approximate 1:1.5 diastereomeric mixture in 10% isolated yield, **Scheme 51**.



Scheme 51: Addition of *n*-butyllithium to bromide **116**.

Repeating the reaction with *tert*-butyllithium requires 2 equivalents, as the *tert*-butyl bromide product formed from lithium-halogen exchange can undergo E2 elimination with a second equivalent of *tert*-butyllithium acting as a base, to ultimately form isobutylene.¹⁷¹ Treating bromide **116** with *tert*-butyllithium resulted in no observable lithium-halogen exchange when the reaction mixture was quenched with deuterated methanol.

Formation of a Grignard reagent was performed with the unfunctionalised bromide **123**. As with the previous reactions, oven-dried glassware and an inert N_2 atmosphere were employed. Grignard formation was successful with both DIBAL-H and iodine activation, and the Grignard was trapped with benzaldehyde **125** to confirm its formation, **Scheme 52**.



Scheme 52: Formation and trapping of Grignard reagent **124**.

Both activating agents gave combined isolated yields of the products **126** and **127** of ~80%, confirming significant formation of the Grignard reagent **124**, **Table 11**.

Table 11: Activating agents screened for the formation of Grignard reagent **124**.

Entry	Activating Agent	126	127
		Isolated Yield / %	Isolated Yield / %
1	DIBAL-H	48	30
2	Iodine	50	33

This suggests that the bromide **116** is not compatible with Grignard formation and therefore is not compatible with the decarboxylative cross-coupling methodology reported by Baran and co-workers.

The applicability of modern decarboxylative cross-coupling methodologies to the synthesis of PROTACs has been diligently and methodically investigated. From the experimental data obtained, the methodologies developed by Baran and co-workers are not amenable to the coupling of protein binders containing complex heterocycles, with RAEs formed from simple carbon linkers.

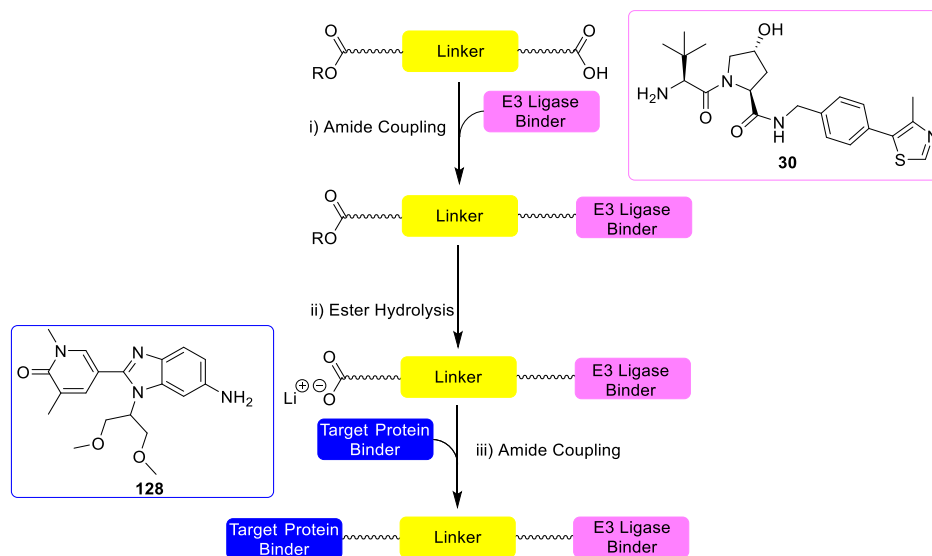
A more robust and reliable methodology for synthesising PROTACs is required to allow the investigation of the effects of linker functionality on the physicochemical properties of PROTACs.

3.6 Synthesis of PROTACs by Amide Coupling

The venerable amide coupling reaction is the most commonly used reaction in Medicinal Chemistry. In 2014, Boström and Brown analysed a representative data set of 125 publications in the *Journal of Medicinal Chemistry* and found that amide couplings appeared in > 50% of the publications.¹⁷² Amide coupling is most commonly achieved through the condensation of a carboxylic acid and amine, using coupling reagents that activate the carboxylic acid.¹⁷³

Carboxylic acids and esters exhibit orthogonal reactivity under amide coupling reaction conditions. Therefore, it was envisaged that the original linker strategy using bifunctional linkers with terminal carboxylic acid and ester functionality could be used

to synthesise PROTACs in an iterative manner. This would enable the screening of a plethora of linkers and the investigation of the effects of linker functionality on the physicochemical properties of PROTACs, **Scheme 53**.

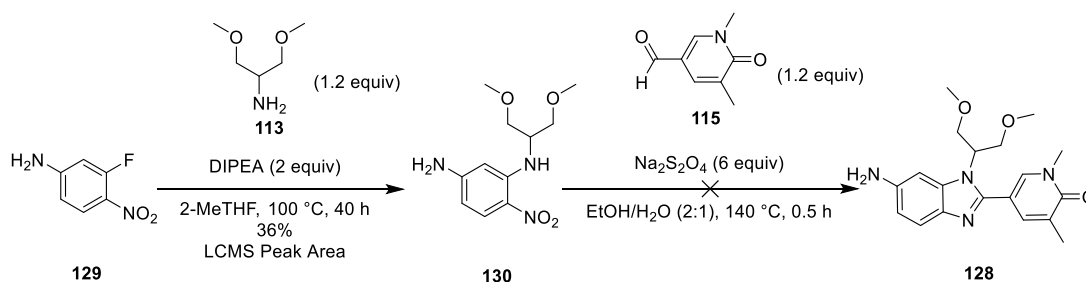


Scheme 53: Synthetic strategy used to synthesise PROTACs with a plethora of linker to assess the effects of linker functionality on their physicochemical properties.

Due to the robustness and wide functional group tolerance of amide coupling reactions, it was envisaged that this linker strategy could be developed into a plate-based one-pot methodology that would allow the efficient and flexible high-throughput synthesis of a variety of PROTACs. To accommodate this, the synthetic sequence was reordered so that the first coupling occurs with the primary amine of **30**, prior to ester hydrolysis and the second coupling with the aniline of **128**. Anilines are less reactive towards amide couplings than primary amines, due to delocalisation of the nitrogen lone pair that is in conjugation with the aromatic ring.¹⁷⁴ Thus, the amide coupling with VHL E3 ligase binder **30** is expected to be higher yielding, with less residual uncoupled amine that could interfere with future coupling reactions.

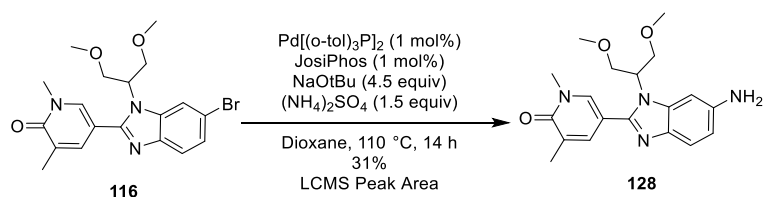
However, it should be noted that the PROTACs will now contain two legacy amide bonds from the couplings. This increases the number of HBDS and HBAs, limiting the ability to control the effects of these properties on the physicochemical profile of the PROTACs.

It was envisaged that amine **128** could be accessed from aryl fluoride **129** following the synthetic sequence in **Scheme 54**. Bis-aniline **130** was formed in 36% by LCMS peak area after 40 h. This may be due to the electron-donating effects of the aniline lone pair, increasing the electron density in the ring and thus decreasing its propensity to undergo the desired S_NAr reaction. Large quantities of bis-aniline **130** are required and the starting material aryl fluoride **129** is expensive and only available in small quantities. Additionally, it was proposed that there may be some selectivity issues between the addition of either the primary or secondary amine in **130** to the aldehyde **115**, that may limit the yield of the desired amine product **128**. This would require protection of the free amine in **129**, adding additional steps to the synthetic sequence as well as introducing potential issues with protecting group stability under the reaction conditions.



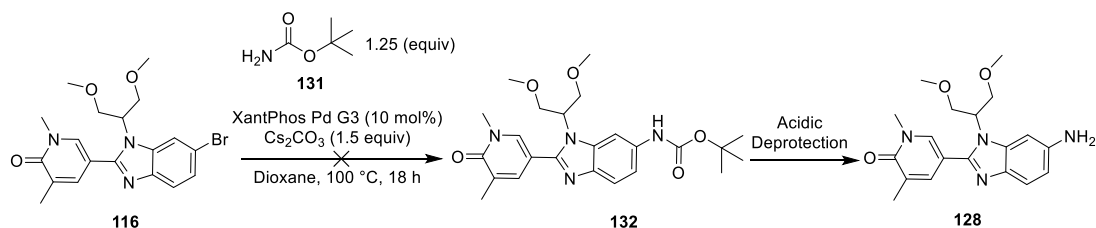
Scheme 54: Synthesis of amine **128** from aryl fluoride **129** over two steps.

Therefore, it was proposed that amine **128** could be accessed from bromide **116**, which had already been synthesised in large quantities. This was attempted using a Buchwald-Hartwig type palladium-catalysed cross-coupling, with $(\text{NH}_4)_2\text{SO}_4$ as the nitrogen source. The reaction was performed following the described literature procedure in a glovebox under an inert N_2 atmosphere using oven-dried glassware, **Scheme 55**.¹⁷⁵ LCMS analysis of the reaction mixture after 14 h observed 31% amine **128** by peak area. The major byproduct of the reaction was the debrominated starting material **110**, formed in 28% by peak area.



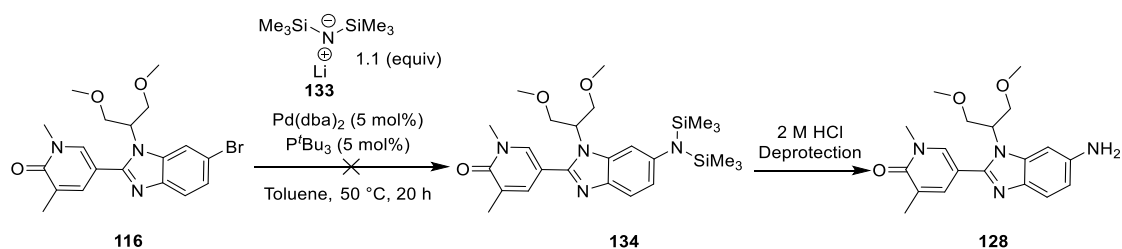
Scheme 55: Synthesis of amine **128** via the Buchwald-Hartwig type coupling of bromide **116** and (NH₄)₂SO₄. The reaction mixture was analysed by LCMS and the peak areas reported as a percentage.

The Buchwald-Hartwig type palladium-catalysed cross-coupling, using the carbamate **131** to access the Boc-protected amine **132**, which could be subsequently deprotected under acidic conditions to furnish amine **128**, was attempted. The reaction was performed under an inert N₂ atmosphere following the described literature procedure, **Scheme 56**.¹⁷⁶ However, there was no observable formation of the Boc-protected amine **132** or amine **128** by LCMS analysis after 18 h.



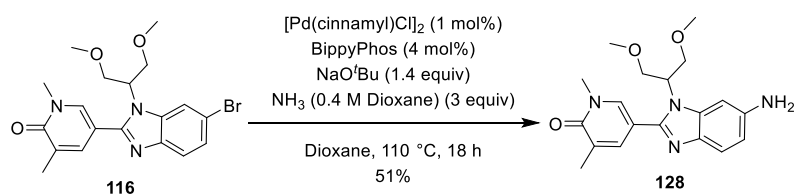
Scheme 56: Synthesis of amine **128** via the deprotection of **132** formed from the Buchwald-Hartwig type coupling of bromide **116** and carbamate **131**.

The palladium-catalysed cross-coupling of bromide **116** with LiHMDS **133** to access bis(trimethylsilyl)amine **134**, which could be subsequently deprotected under acidic conditions to furnish amine **128**, was attempted. The reaction was performed following the described literature procedure in a glovebox under an inert N₂ atmosphere using oven-dried glassware, **Scheme 57**.¹⁷⁷ However, there was no observable formation of bis(trimethylsilyl)amine **134** or amine **128** by LCMS analysis after 20 h or the addition of 2 M HCl.



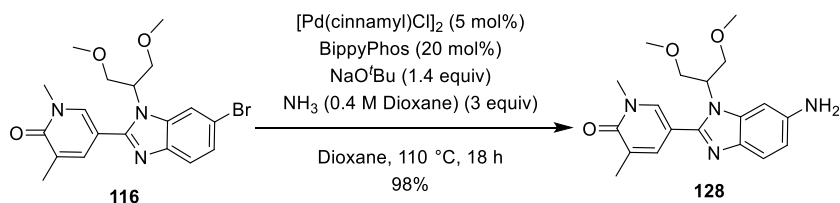
Scheme 57: Synthesis of amine **128** via the deprotection of bis(trimethylsilyl)amine **134** formed from bromide **116**.

The Buchwald-Hartwig type palladium-catalysed cross-coupling using ammonia as the nitrogen source to directly access amine **128** was attempted. The reaction was performed following the described literature procedure in a glovebox under an inert N₂ atmosphere using oven-dried glassware, **Scheme 58**.¹⁷⁸ The desired amine **128** was formed in 51% isolated yield following purification by column chromatography.



Scheme 58: Synthesis of amine **128** via the Buchwald-Hartwig type coupling of bromide **116** with NH₃.

The reaction in **Scheme 58** requires inert conditions and the omission of water. As NaO^tBu is a hygroscopic solid, the moisture it absorbs from the atmosphere may be limiting the reaction. A new bottle of NaO^tBu was dried in a vacuum oven at 40 °C for 7 days before being stored in a glovebox. Repeating the reaction with the newly dried base and increasing the catalyst and ligand loading 5-fold furnished an improved isolated yield of 98%, **Scheme 59**.

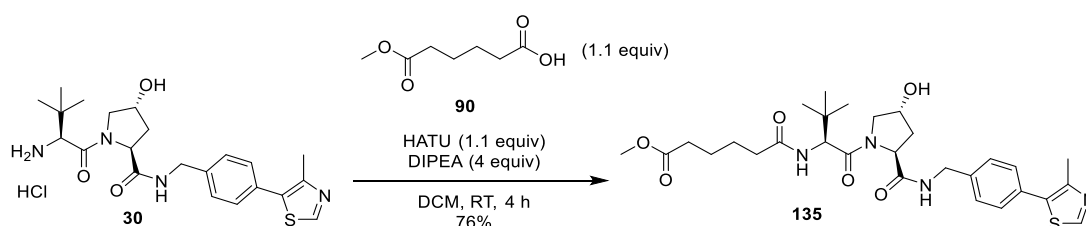


Scheme 59: Synthesis of amine **128** via the Buchwald-Hartwig type coupling of bromide **116** with NH₃.

With the protein binder **128** in hand, the synthesis of a PROTAC with the model 4-carbon linker **90** was performed, following the synthetic strategy outlined in **Scheme 53**.

The Compounds Arrays Team at GSK Stevenage is a bespoke group that regularly uses an array format to couple many different amines and carboxylic acids in a plate-based format to allow for rapid and extensive SAR studies of target compounds. After consultation with its members, standard amide coupling conditions using the carboxylic acid activating agent HATU and the base DIPEA in either DMF or DCM were decided upon. These robust conditions have been extensively used by the department and show the greatest functional group tolerance.¹⁷⁹

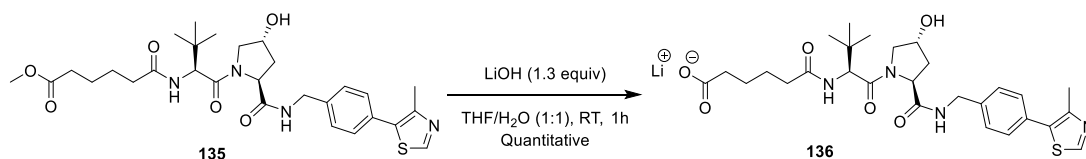
Standard amide coupling conditions employ 1.1 equivalents of HATU and 3 equivalents of DIPEA. However, as amine **30** is the HCl salt, an extra equivalent of DIPEA was used to provide the free base. Carboxylic acid **90**, HATU and DIPEA were sonicated in the reaction solvent for 30 s to pre-form the activated ester prior to the addition of amine **30**. The amide coupling of amine **30** with carboxylic acid **90** furnished the product **135** in 86% and 84% by LCMS peak area after 4 h in DMF and DCM respectively. Multiple washes of the reaction mixture with LiCl solution was required to sufficiently remove enough DMF for the crude product to be purified by column chromatography. Therefore, using DCM as a solvent was preferable and gave an isolated yield of 76% following purification by column chromatography, **Scheme 60**.



Scheme 60: Synthesis of **135** via the amide coupling of amine **30** and carboxylic acid **90**.

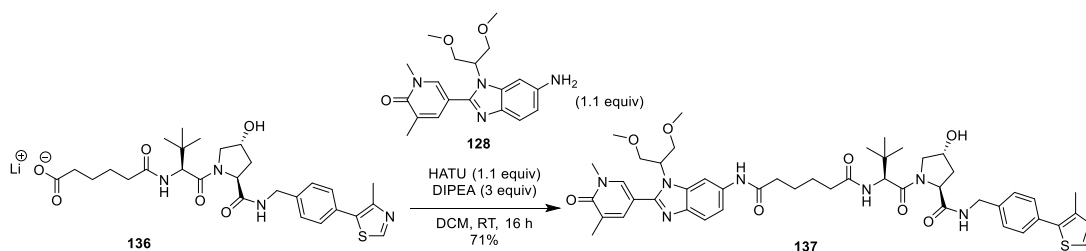
Hydrolysis of ester **135** was performed by the dropwise addition of a solution of ester **135** in THF to a stirred solution of LiOH in water, **Scheme 61**. A small excess of LiOH was required to achieve quantitative conversion to the lithium carboxylate salt **136**.

The reaction solvent was removed under reduced pressure and the crude product was telescoped through to the next step, without the need for purification.



Scheme 61: Hydrolysis of ester **135** to form the lithium carboxylate salt **136**.

The amide coupling of the lithium carboxylate salt **136** with amine **128** required a longer duration of 16 h, due to the lower reactivity of the aniline. PROTAC **137** was isolated in 71% yield, following an aqueous workup and purification by MDAP, **Scheme 62**. This equates to an isolated yield of 54% over three steps, with an average yield per step of 82%.



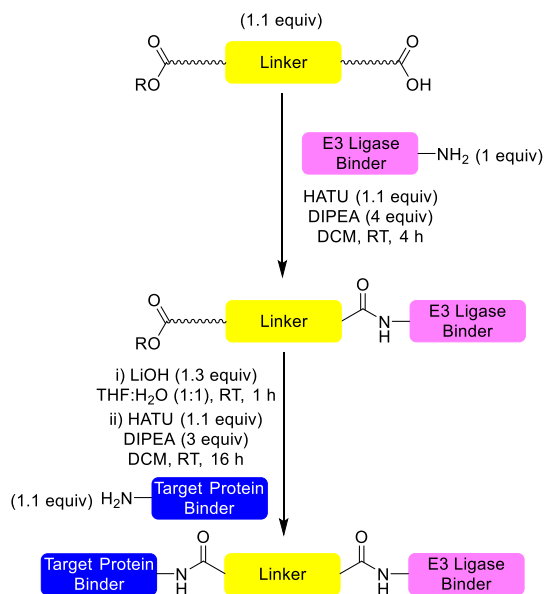
Scheme 62: Synthesis of PROTAC **137** via the amide coupling of lithium carboxylate salt **136** and amine **128**.

A high yielding methodology and linker strategy has been successfully developed, which could be applied to efficiently synthesise PROTACs with a plethora of linkers, enabling the investigation of the effects of linker functionality on the physicochemical properties of PROTACs.

3.7 High-Throughput Synthesis of PROTACs by Amide Coupling

3.7.1 Acid-Ester Linkers

A high yielding methodology and linker strategy was successfully developed in **Section 3.6** and is described in **Scheme 63**.



Scheme 63: Methodology and linker strategy developed and described in Section 3.6.

A Reaxys search of all molecules that contain both a carboxylic acid and ester functional group was refined by the additional functionality, price and availability from suppliers to furnish a library of 66 linkers, **Figure 36**.

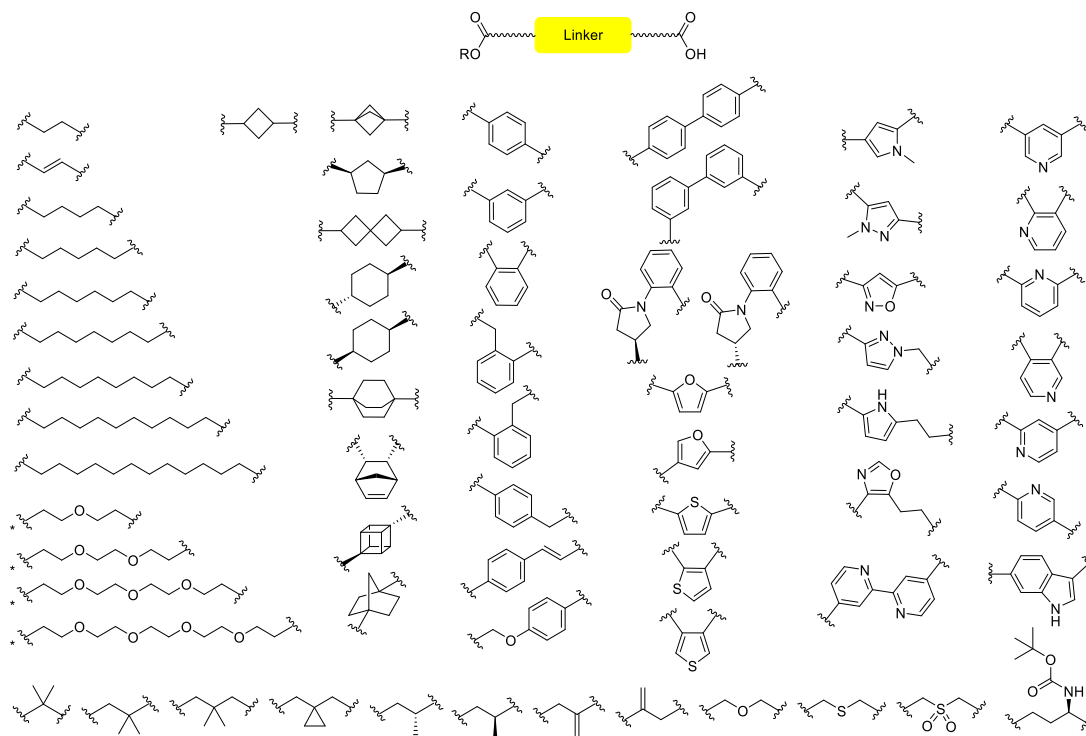


Figure 36: Terminal carboxylic acid-ester linker library. * = Diacid linker.

The protocol in **Scheme 63** was used to synthesise the 23 PROTACs shown in **Figure 37** from the linker library described in **Figure 36**. However, when synthesising PROTAC **153b**, the lithium carboxylate salt exhibited poor solubility in DCM, resulting in only a 7% isolated yield of PROTAC **153b**. Therefore, all future reactions were performed in DMF, as the lithium salts exhibited a visibly greater solubility in this solvent. Additionally, the homogenous reaction mixtures could be directly purified by MDAP in DMF, without requiring an aqueous workup.

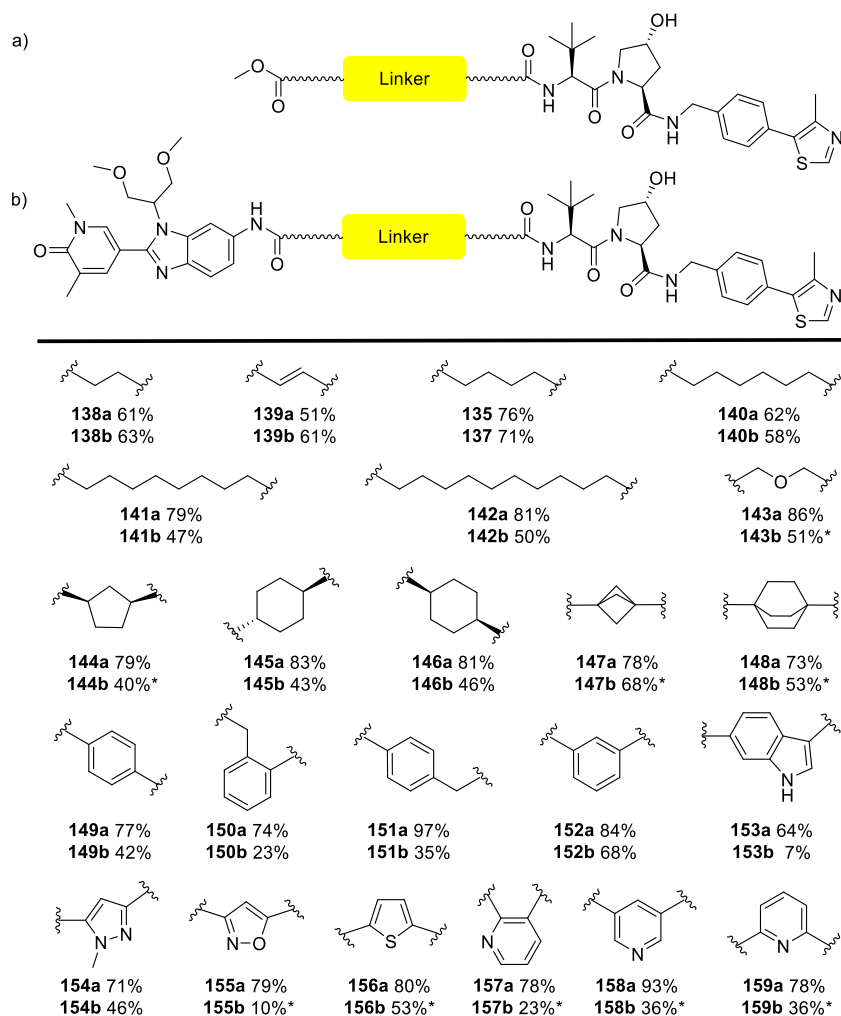
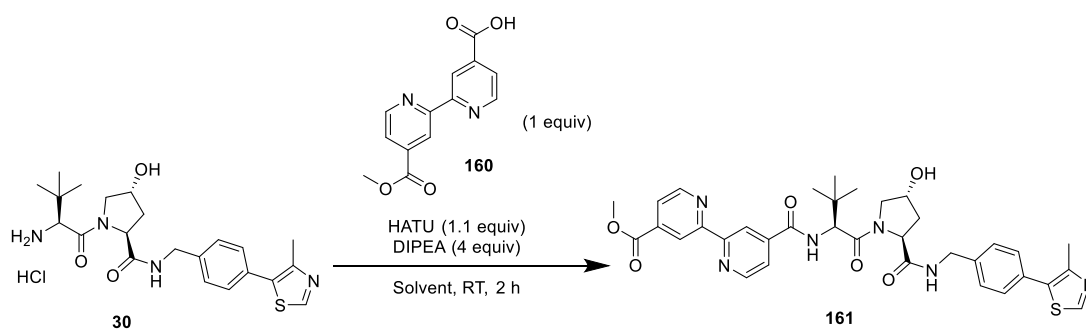


Figure 37: Twenty-three PROTACs synthesised using the methodology and linker strategy outlined in **Scheme 63**. Yields are reported for the both the first (a) and second (b) couplings. *second coupling performed in DMF.

It was envisaged that the protocol described in **Scheme 63** could be developed into a one-pot methodology that would allow the high-throughput synthesis of PROTACs in

a plate-based format. This required several current limitations to be addressed, such as the purification and isolation of the product of the first coupling reaction, as well as the dilute reaction concentration of 0.05 M.

The 2,2'-bipyridyl linker **160** was chosen as the model linker for the one-pot optimisation as it was predicted to exhibit inherently low solubility due to its rigid aromatic structure. Therefore, it provides a good starting point to develop robust conditions that will show good functional group tolerance and facilitate the synthesis of the remaining PROTACs in the series. First, the amide coupling of the 2,2'-bipyridyl linker **160** and VHL E3 ligase binder **30** was screened in DMF and DCM, with varying concentrations, **Scheme 64**. LCMS analysis of the reaction mixtures at regular intervals confirmed completion of the reaction after 2 h, with no further increase in formation of product **161** observed after 64 h.



Scheme 64: Synthesis of product **161** from the amide coupling of VHL E3 ligase binder **30** and 2,2'-bipyridyl linker **160**.

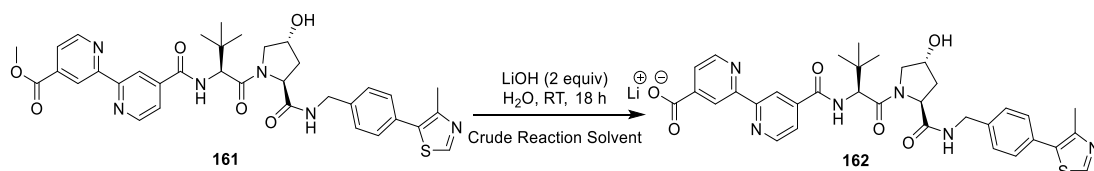
The results for the solvent and concentration screen were obtained by LCMS analysis of the reaction mixture and are visualised in **Table 12**. A maximum formation of product **161** of 86% by LCMS peak area was achieved in DMF at 0.067 M, **Table 12**, entry 1. It appears that as the concentration of the reaction mixture increases from 0.067 M to 0.10 M and beyond, the reactions performed in DCM exhibit a higher formation of product **161** than the reactions performed in DMF at the same concentration, **Table 12**, entries 2, 3, 5 and 6. A maximum formation of product **161** in DCM of 83% by LCMS peak area was achieved at 0.20 M, **Table 12**, entry 6, with a small decrease in formation to 79% when doubling the concentration to 0.40 M, **Table 12**, entry 7. Although DMF gives the highest formation of product **161**, a more

concentrated reaction mixture is desirable to allow the reactions to be performed in 4 mL vials in a plate format. Therefore, the most favourable conditions for this reaction are in DCM at 0.20 M and 0.40 M, **Table 12**, entry 6 and 7.

Table 12: Conditions screened for the synthesis of product **161** from the amide coupling of VHL E3 ligase binder **30** and 2,2'-bipyridyl linker **160**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Solvent	Concentration / M	161 / %	30 / %
1	DMF	0.067	86	5
2	DMF	0.10	74	6
3	DMF	0.20	81	6
4	DCM	0.067	80	5
5	DCM	0.10	79	5
6	DCM	0.20	83	5
7	DCM	0.40	79	5

Addition of the reaction mixtures to an aqueous solution of lithium hydroxide resulted in minimal formation of the desired lithium carboxylate salt **162**, **Scheme 65** and **Table 13**.



Scheme 65: Hydrolysis of crude ester **161** to lithium carboxylate salt **162**.

Table 13: Hydrolysis of crude ester **161** to lithium carboxylate salt **162**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Crude Reaction Solvent	162 / %	161 / %
1	DMF	5	88
2	DCM	8	76

Therefore, the crude reaction mixtures were blown down under a stream of N₂ and placed in a vacuum oven at 40 °C for 18 h to remove any residual reaction solvent. The hydrolysis of ester **161** was repeated, screening a range of cosolvents, temperatures and times, **Table 14**. Hydrolysis of ester **161**, which had been synthesised

in DCM using MeOH as the cosolvent at ambient temperature, gave 22% formation of the desired lithium carboxylate salt **162** by LCMS peak area after 18 h, **Table 14**, entry 1. Repeating the reaction with ester **161**, which had been synthesised in DMF, gave a lower formation of lithium carboxylate salt **162** of 10%, **Table 14**, entry 2. This may be due to the incomplete removal of DMF, which has a much higher boiling point. This further increases the favourability of synthesising ester **161** in DCM. Additionally, repeating the reaction using THF as the cosolvent gave a lower formation of 7%, **Table 14**, entry 3. These three reactions were repeated at 60 °C. A maximum formation of lithium carboxylate salt **162** of 84% by LCMS peak area was observed when using ester **161** that had been synthesised in DCM, with MeOH as the cosolvent after 18 h.

Table 14: Reaction conditions screened for the hydrolysis of ester **161**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	LiOH / equiv	Hydrolysis Cosolvent	Hydrolysis Concentration / M	Temperature / °C	Duration / h	162 / %	161 / %
1	2	MeOH	0.1	RT	18	22	73
2*	2	MeOH	0.1	RT	18	10	79
3	2	THF	0.1	RT	18	7	81
4	2	MeOH	0.1	60	18	84	11
5*	2	MeOH	0.1	60	18	28	62
6	2	THF	0.1	60	18	8	89

* = Ester **161** synthesised using DMF as the reaction solvent. All other entries used ester **161** synthesised in DCM.

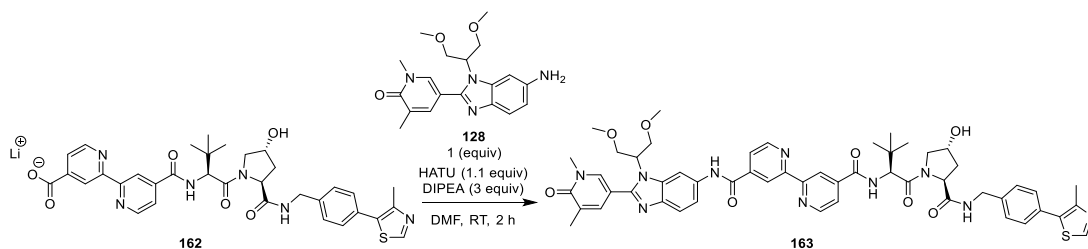
The hydrolysis of ester **161** synthesised in DCM using MeOH as the cosolvent was repeated, screening a range of concentrations and equivalents of LiOH, **Table 15**. LCMS analysis of the reaction mixtures at regular time intervals confirmed that the reaction was not complete after 6 h. Therefore, all reactions were run for 18 h to ensure that the duration was not a limiting factor. Changing the equivalents of LiOH from 2 to 4 gave comparable formation of lithium carboxylate salt **162** by LCMS peak area at 0.1 M, with any discrepancies being within the error of the LCMS analysis, **Table 15**, entries 1, 2 and 3. Doubling the concentration of the reaction mixture resulted in a lower formation of lithium carboxylate salt **162** of 77% with 2 equivalents LiOH,

Table 15, entries 1 and 4. However, this decrease in formation was not observed with 3 equivalents of LiOH, **Table 15**, entries 2 and 5.

Table 15: Reaction conditions screened for the hydrolysis of ester **161** with MeOH as the cosolvent. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	LiOH / equiv	Hydrolysis Concentration / M	Temperature / °C	Duration / h	162 / %	161 / %
1	2	0.1	60	18	84	11
2	3	0.1	60	18	84	11
3	4	0.1	60	18	83	10
4	2	0.2	60	18	77	18
5	3	0.2	60	18	83	12

Finally, the effect of LiOH concentration on the second amide coupling was screened. All reactions were performed in DMF as the lithium salts exhibited a visibly greater solubility in this solvent. Additionally, the reaction mixtures could be directly purified by MDAP, further increasing the potential for automation of the methodology. All reactions were performed at ambient temperature at 0.4 M concentration using the conditions in **Scheme 66**.



Scheme 66: Synthesis of PROTAC **163** from the amide coupling of lithium carboxylate salt **162** and amine **128**.

LCMS analysis of the reaction mixtures at regular time intervals up to 16 h confirmed that the reactions were complete after 2 h. A maximum formation of PROTAC **163** of 41% was observed with lithium carboxylate salt **162** that had been synthesised with 4 equivalents of LiOH, **Table 16**, entry 3. A lower formation of PROTAC **163** is observed with lithium carboxylate salt **162** that had been synthesised with 3 equivalents of LiOH, **Table 16**, entry 1 and 2. This may be due to a higher amount of unhydrolysed ester product observed in the LCMS, suggesting incomplete

hydrolysis of ester **161** with 3 equivalents of LiOH. The LCMS traces with both 4 and 10 equivalents of LiOH have the same amount of ester. However, a lower formation of PROTAC **163** of 30% by LCMS peak area was observed with 10 equivalents of LiOH. This may be due to the excess hydroxide-promoted hydrolysis of HATU or the activated ester formed between lithium carboxylate salt **162** and HATU, prior to coupling with amine **128**. Therefore, the optimum ester hydrolysis conditions with 4 equivalents of LiOH at 0.2 M were selected.

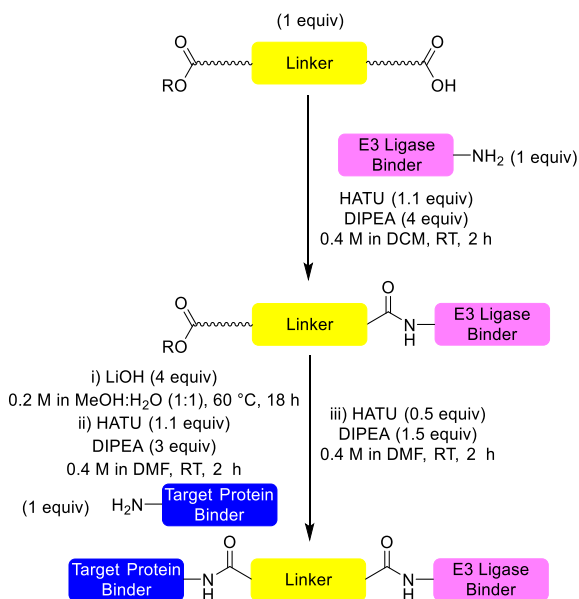
Table 16: Reaction conditions screened for the synthesis of PROTAC **163** from the amide coupling of lithium carboxylate salt **162** and amine **128**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	LiOH / equiv	HATU / equiv	163 / %	162 / %	161 / %
1	3	1.1	22	5	18
2	3	2	24	5	11
3	4	1.1	41	8	7
4	10	1.1	30	10	7
5*	3	1.1	56	0	7

* = Second addition of HATU (0.5 equiv) and DIPEA (1.5 equiv) after 2 h, with stirring for a further 2 h.

Increasing the equivalents of HATU to 2 equivalents, to counteract the effect of the excess hydroxide, resulted in a minimal increase in formation of PROTAC **163**, **Table 16**, entries 1 and 2. However, addition of a further 0.5 equivalents of HATU and 1.5 equivalents of DIPEA after 2 h resulted in complete consumption of lithium carboxylate salt **162** and an additional increase in formation of PROTAC **163** after a further 2 h, **Table 16**, entry 5. The reaction mixture was purified directly by MDAP to furnish PROTAC **163** in 36% isolated yield, with an average yield per step of 71%.

A high-throughput, one-pot methodology was successfully developed for the coupling of acid-ester linkers and is describe in **Scheme 67**. This protocol was used to synthesise the 39 PROTACs shown in **Figure 38** in a plate-based format. The solid reagents: VHL E3 ligase binder **30**, amine **128** and HATU were weighed out into separate vials using a Quantos QX96 automated weighing machine. The liquid reagents: DIPEA, LiOH stock solution and solvents were dispensed using a multi-channel pipette. Additionally, the final reactions mixtures were purified directly by MDAP.



Scheme 67: One-pot protocol for the high-throughput synthesis of PROTACs in a plate-based format from acid-ester linkers.

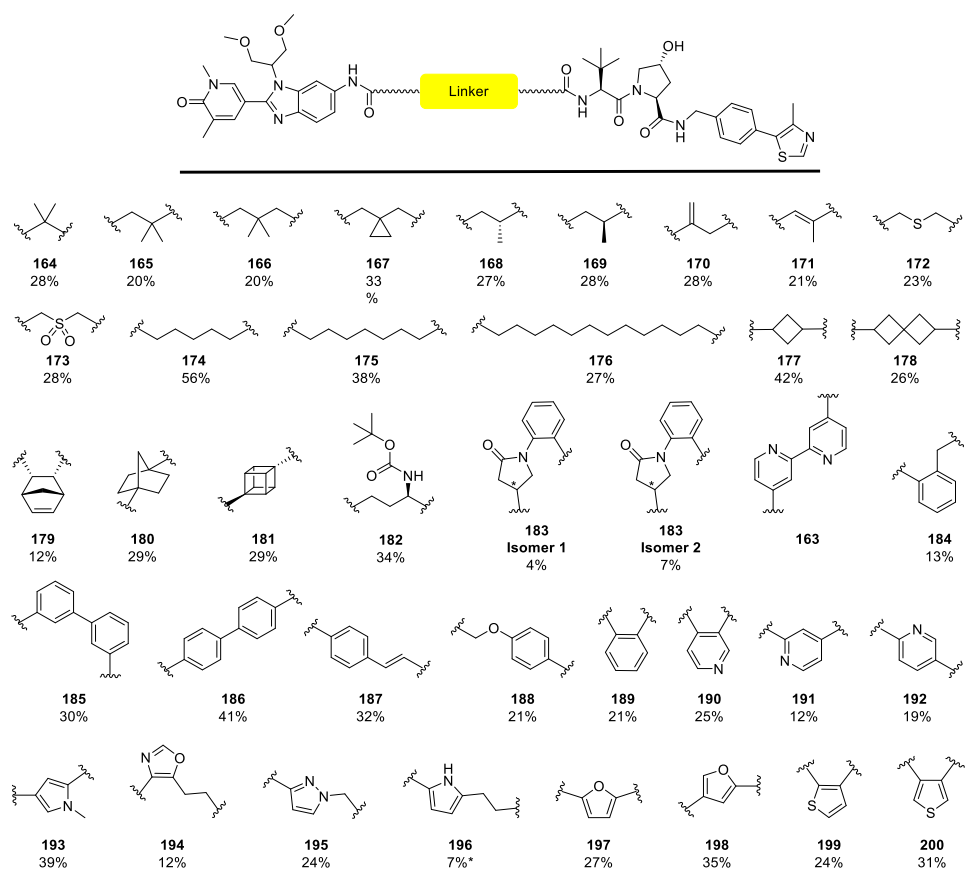


Figure 38: Thirty-nine PROTACs synthesised using the one-pot protocol outlined in Scheme 67.

* = Impure product following purification by MDAP.

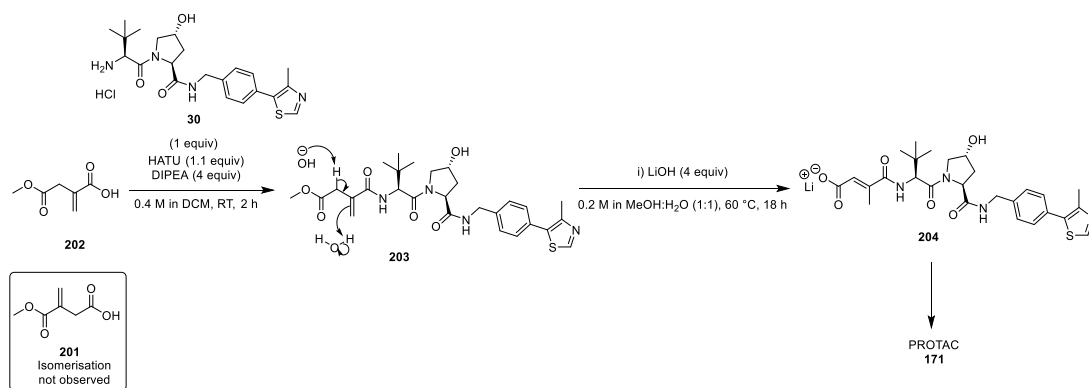
The high level of automation of this methodology enabled the synthesis of 24 PROTACs in one week. This productivity was only limited by the size of the plate used and the availability of the purification equipment. Therefore, it is reasonable to propose that this methodology would be amenable to much larger plate dimensions.

The protocol described in **Scheme 67** was used to synthesise the 39 PROTACs shown in **Figure 38** with a range of yields from 4-56%, equating to an average yield of 24%, with an average yield per step of 62%.

This methodology showed good functional group tolerance for alkyl, spirocyclic, aromatic and heteroaromatic linkers, including the sulfide **172** and sulfone **173**. Additionally, the Boc-protecting group in PROTAC **182** was not affected by the mild reaction conditions. Acidic deprotection of the Boc-protecting group in PROTAC **182** would furnish a primary amine that could be used to modify the physicochemical profile of the PROTAC further, by adding solubilising groups or a moiety that makes favourable interactions with the protein surface in the ternary complex.

The pyrrole linker PROTAC **196** was isolated in only 7% yield following purification by MDAP. NMR analysis of the isolated product confirmed that it was in fact an approximately 1:1 mixture of the desired PROTAC **196** and the VHL E3 ligase binder pyrrole carboxylic acid, presumably formed through base-mediated hydrolysis of the pyrrole amide bond under the HpH purification conditions. This highlights the reactivity of this particular pyrrole linker and its unsuitability for incorporation into a PROTAC.

An interesting observation was made for PROTACs **170** and **171**. Both PROTACs were synthesised from acid-esters linkers containing a gem-disubstituted alkene, **201** and **202** respectively. However, the more substituted and stable double-bond PROTAC **171** was formed as the major product *via* a base-catalysed isomerisation of the double-bond under the reaction conditions, **Scheme 68**. Interestingly only a negligible amount of the double-bond isomerisation product was observed with linker **201**.



Scheme 68: Base-catalysed double-bond isomerisation of ester **203** to form PROTAC **171**.

It was proposed that the selectivity of double-bond isomerisation was driven by the pK_a of the α -proton of the double-bond. Comparing the pK_a 's of comparable substrates gives an insight to the observed selectivity. The α -proton of the ester ethyl 2-phenylacetate has a pK_a of 22.7 in DMSO.¹⁸⁰ In comparison, the α -proton of the amide *N,N*-dimethyl-2-phenylacetamide has a pK_a of 26.6 in DMSO.¹⁸⁰ This suggests that the α -proton of the double-bond and ester in linker **202** is also more acidic. Therefore, linker **202** will be more readily deprotonated, forming the more substituted and stable double-bond that was observed for PROTAC **171**. It is likely that this isomerisation occurs prior to ester hydrolysis, as the deprotonation of the α -proton to δ,γ -unsaturated carboxylic acids requires much stronger bases, such as *n*-BuLi, which has a $pK_{aH} > 50$ in DMSO. Whereas LiOH will not deprotonate the α -proton, following formation of the carboxylate salt, as the pK_{aH} of the hydroxide anion is only 31.4 in DMSO.¹⁸¹

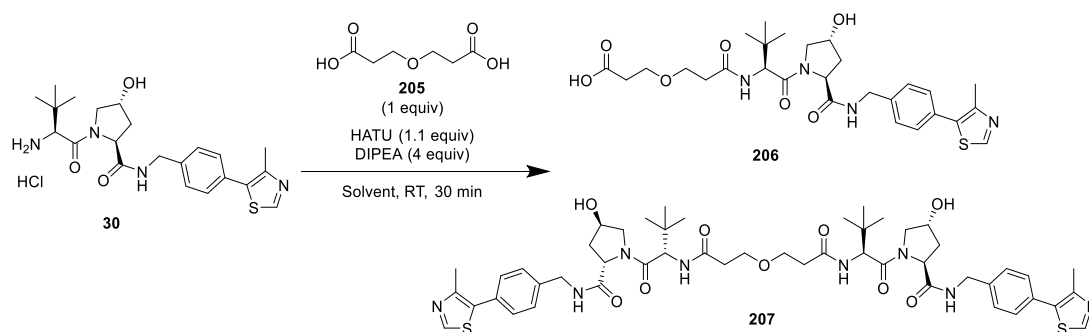
Separation of racemic acid-ester linkers with an unresolved chiral centre was also achieved by this methodology. Coupling of the racemic acid-ester to the VHL E3 ligase binder **30**, which has three known chiral centres, forms a mixture of diastereoisomers. As diastereoisomers have different physical properties, they can commonly be separated by column chromatography and MDAP. This process was used to synthesise the diastereoisomers **168**, **169** and **183** Isomers **1** and **2**.

The absolute stereochemistry of PROTAC **168** was confirmed by X-ray crystallographic analysis by obtaining a cocrystal structure of PROTAC **168** with

human Brd4 BD1, **Appendix 7.2**. A cocrystal structure of PROTAC **169** could not be obtained with sufficient resolution for elucidation of the absolute stereochemistry. However, as both PROTACs **168** and **169** were synthesised from the same racemic starting material and separated by purification, the absolute stereochemistry of the linker methyl group of PROTAC **169** was assigned relative to PROTAC **168**. A cocrystal structure of PROTACs **183 Isomer 1** and **183 Isomer 2** with human Brd4 BD1 could not be obtained, and thus the absolute stereochemistry at the unknown chiral centre was not assigned.

3.7.2 Diacid Linkers

Due to the prohibitive price of acid-ester PEG linkers, the diacid PEG1, 2, 3, and 4 linkers were utilised. The diacid PEG1 linker **205** was used as the model linker for the optimisation of a one-pot methodology for the synthesis of PROTACs from diacid linkers. First, the amide coupling of diacid linker **205** and amine **30** was screened in DMF and DCM, using the reaction conditions outlined in **Scheme 69**. LCMS analysis of the reaction mixtures at regular intervals confirmed completion of the reaction after 30 min, with no further increase in formation of carboxylic acid **206** observed after 3 h.



Scheme 69: Synthesis of carboxylic acid **206** from the amide coupling of amine **30** and diacid linker **205**.

A maximum formation of carboxylic acid **206** of 40% by LCMS peak area was achieved in DCM at 0.06 M, **Table 17**, entry 2. Performing the reaction in DMF, **Table 17**, entry 1, resulted in a lower formation of carboxylic acid **206** of 23%, with the dicoupled product **207** being formed in 49% by LCMS peak area.

Table 17: Conditions screened for the synthesis of carboxylic acid **206** from the amide coupling of amine **30** and diacid linker **205**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Solvent	Concentration / M	30 / %	206 / %	207 / %
1	DMF	0.06	14	23	49
2	DCM	0.06	15	40	32

A range of concentrations were screened for the coupling of diacid linker **205** and amine **30** in DCM and are visualised in **Table 18**.

Table 18: Conditions screened for the synthesis of carboxylic acid **206** from the amide coupling of amine **30** and diacid linker **205**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Concentration / M	30 / %	206 / %	207 / %
1	0.02	15	30	40
2	0.05	14	41	30
3	0.06	15	40	32
4	0.1	14	38	34
5	0.4	15	39	35

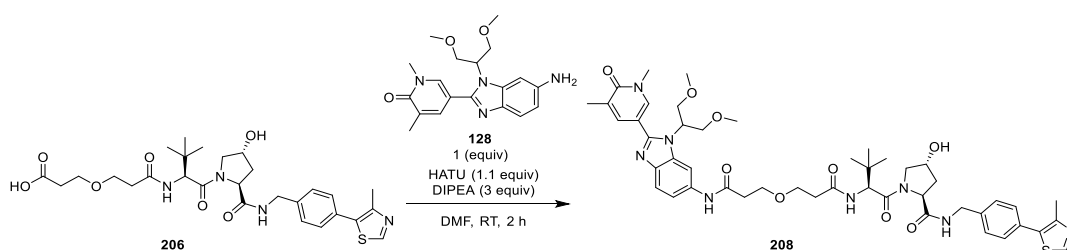
A maximum formation of carboxylic acid **206** of 41% by LCMS peak area was achieved by decreasing the concentration to 0.05 M, **Table 18**, entry 2. However, further decreasing the concentration to 0.02 M furnished a lower formation of carboxylic acid **206** of 30%, with an increase in formation of the dicoupled product **207**, **Table 18**, entry 1. Increasing the concentration to 0.1 M and 0.4 M furnished an increase in the formation of the dicoupled product **207** and concurrent decrease in the formation of carboxylic acid **206**.

A range of equivalents of the diacid PEG1 linker **205** was screened, **Table 19**. Increasing the equivalents of diacid linker **205** to 1.2 furnished a maximum formation of carboxylic acid **206** of 46%, **Table 19**, entry 2. Interestingly, further increasing the equivalents of diacid linker **205** to 1.5 resulted in a decrease in formation of carboxylic acid **206** and a concurrent increase in the formation of dicoupled product **207** **Table 19**, entry 3.

Table 19: Conditions screened for the synthesis of carboxylic acid **206** from the amide coupling of amine **30** and diacid linker **205**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	205 / equiv	30 / %	206 / %	207 / %
1	1.0	14	41	30
2	1.2	14	46	32
3	1.5	15	30	40

The reaction mixtures were blown down under a stream of N₂ and placed in a vacuum oven at 40 °C for 18 h. The coupling of carboxylic acid **206** and amine **128** was performed using the methodology previously optimised for acid-ester linkers, **Scheme 70**.



Scheme 70: Synthesis of PROTAC **208** from the amide coupling of carboxylic acid **206** and amine **128**.

LCMS analysis of the reaction mixture at regular time intervals confirmed its completion after 2 h. The reaction mixture was directly purified by MDAP and the resulting isolated yields are visualised in **Table 20**.

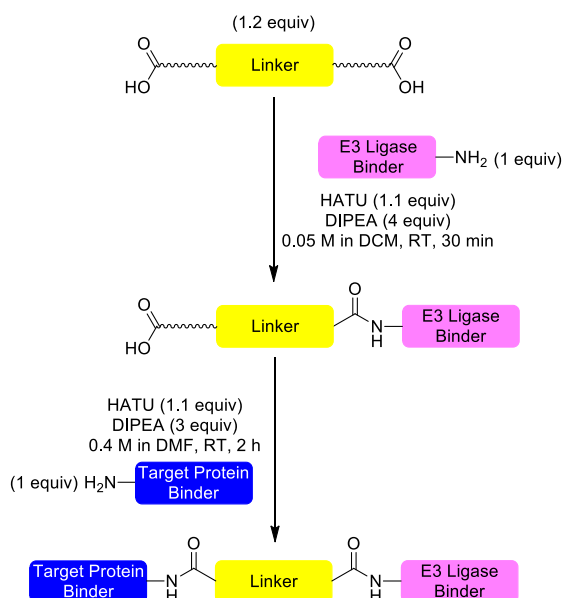
Table 20: Conditions screened for the synthesis of PROTAC **208** from the amide coupling of carboxylic acid **206** and amine **128**.

Entry	205 / equiv	208 Isolated Yield / %
1	1.0	26
2	1.2	28
3	1.5	20

The previously observed trend in the formation of carboxylic acid **206** with a range of equivalents of diacid linker **205**, **Table 20**, was observed for the formation of

PROTAC 208. A maximum isolated yield of **PROTAC 208** of 28% was achieved with 1.2 equivalents of diacid linker **205**.

A high-throughput, one-pot methodology for the coupling of diacid linkers was successfully developed and is described in **Scheme 71**. This protocol was used to synthesise the four **PROTACs** shown in **Figure 39** in a plate-based format.



Scheme 71: One-pot protocol for the high-throughput synthesis of **PROTACs** in a plate-based format from diacid linkers.

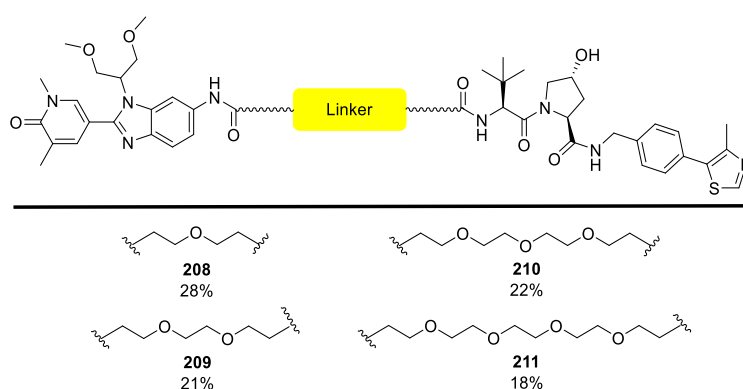


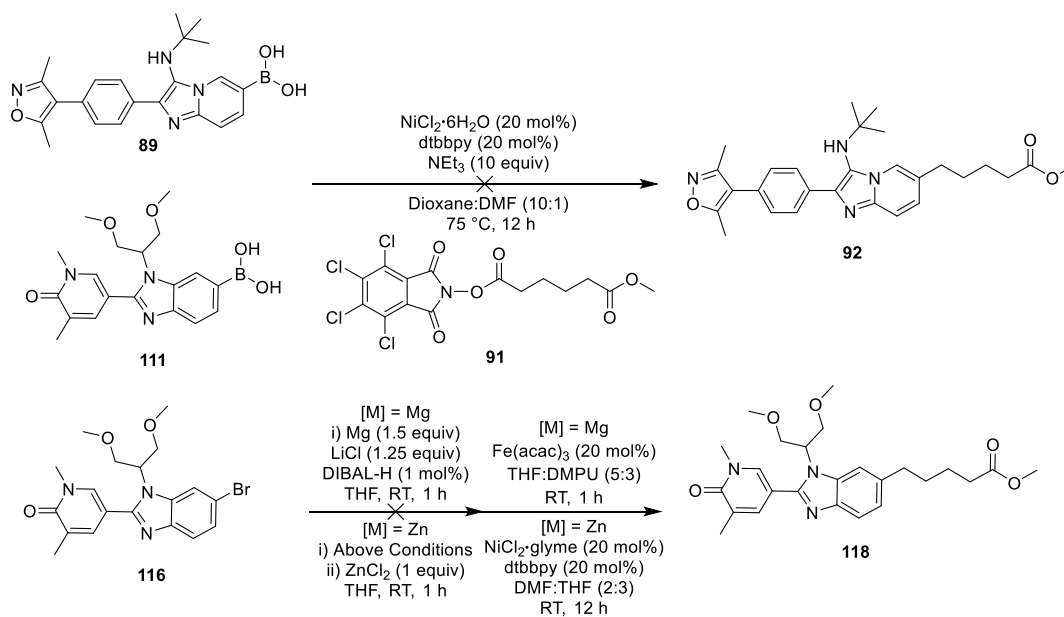
Figure 39: Four **PROTACs** synthesised using the one-pot protocol outlined in **Scheme 71** with diacid linkers.

The protocol in **Scheme 71** was highly automated, solid reagents were weighed out using a Quantos QX96 automated weighing machine, liquid reagents were dispensed

using a multichannel pipette and the final reactions mixtures were purified directly by MDAP. This enabled the synthesis of the four PROTACs shown in **Figure 39** with a range of yields from 18-28%, equating to an average yield of 22%, with an average yield per step of 47%.

3.8 Conclusions

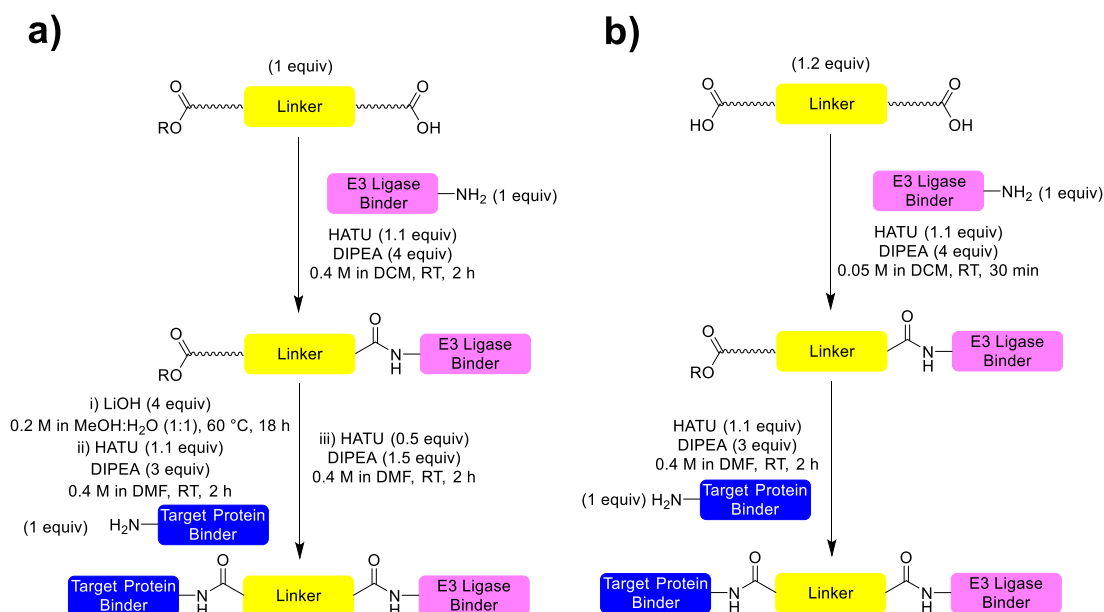
The applicability of modern decarboxylative cross-coupling methodologies to the synthesis of PROTACs was assessed. The methodologies developed by Baran and co-workers for the coupling of aromatic boronic acids, Grignard and organozinc reagents with alkyl RAE were investigated using the pan-BET inhibitors **89**, **111** and **116** and the model 4-carbon linker precursor **91**, **Scheme 72**. However, no desired coupling was observed. It was postulated that these methodologies are not amenable to substrates with free NHs and complex electron-deficient nitrogen-containing heterocycles.



Scheme 72: Decarboxylative cross-coupling methodologies investigated for the synthesis of PROTACs.

A high yielding methodology and linker strategy that uses bifunctional linkers with terminal carboxylic acids and ester functionality has been developed and used to synthesise 23 PROTACs with a range of linkers. This methodology was successfully

developed into a one-pot protocol that facilitated the high-throughput synthesis of 39 PROTACs in a plate-based format, **Scheme 73a**. An analogous one-pot protocol was developed for diacid linkers that facilitated the high-throughput synthesis of four PROTACs in a plate-based format, **Scheme 73b**.



Scheme 73: One-pot protocol for the high-throughput synthesis of PROTACs in a plate-based format from **a)** acid-ester linkers **b)** diacid linkers.

The series of 65 PROTACs shown in **Figure 40** will be used to investigate the effect of linker functionality on the physicochemical properties of PROTACs, **Section 4**.

Additionally, a collaborative project with the GSK Stevenage Discovery High-Throughput Chemistry (DHTC) department has been initiated, using these one-pot protocols to synthesise PROTACs with different protein binders, E3 ligase binders and linkers in a 1536-well plate. The aim of this project is to provide an initial high-throughput screening platform to identify favourable protein binder, linker and E3 ligase binder combinations for new target proteins that can provide the foundations for further optimisation.

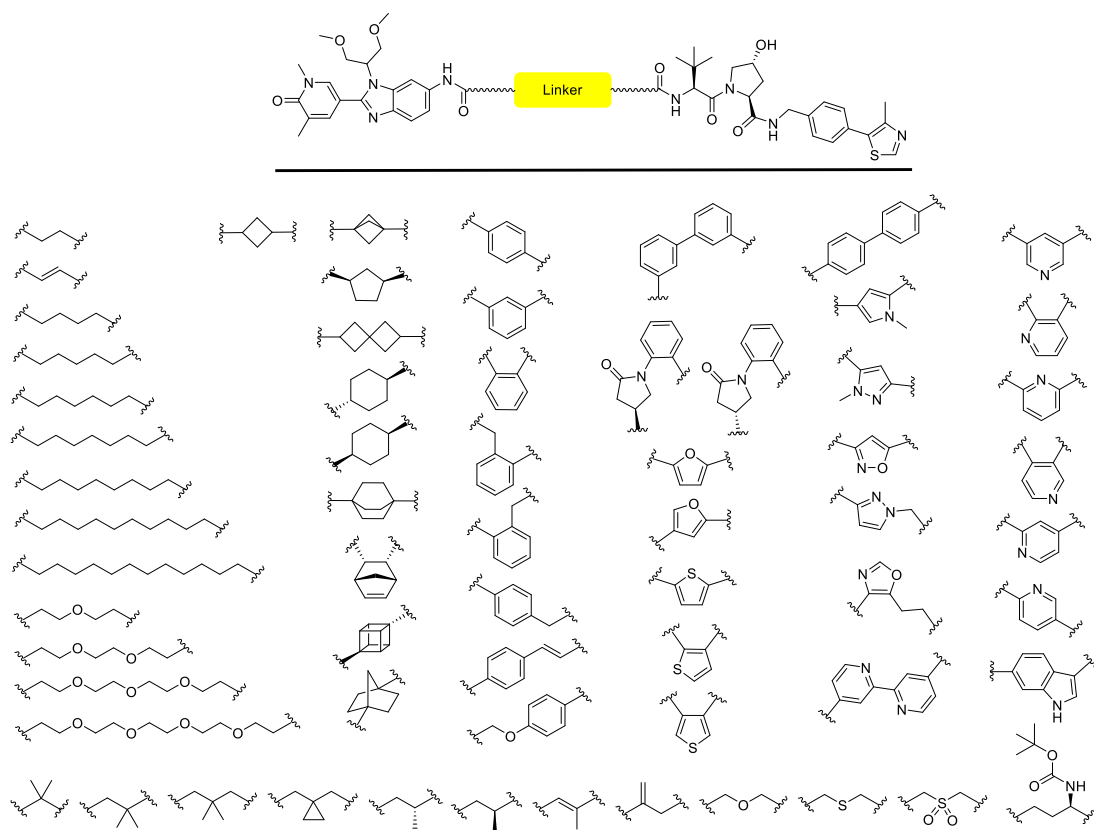


Figure 40: Sixty-five PROTACs synthesised from the linker library in **Figure 36**.

4: Results and Discussion –
Investigating the Effects of Linker
Functionality on the Physicochemical
Properties and Degradation Profiles
of BET PROTACs

4.1 Introduction

Hijacking a cell's ubiquitin proteasome system using PROTAC technology has the potential to be a powerful new therapeutic strategy. Some of the main advantages of this technology include the potential for PROTACs to be catalytic in activity and elicit long-lasting therapeutic effects, offering the possibility of low clinical doses and infrequent dosing regimens.

However, although this new modality is showing great promise, PROTACs are large molecules with molecular weights of 800-1200 Da. This results in them generally exhibiting poor physicochemical properties, including low cell permeability and low aqueous solubility. This results in a reduced *in vivo* efficacy relative to *in vitro* models and is currently limiting the therapeutic potential of PROTACs.

Due to the specific defined structures of the target protein and E3 ligase-binding moieties, the linker provides a facile handle to modulate physicochemical properties.

4.1.1 Physicochemical Property Descriptors

A number of properties are used to define the physicochemical profile of molecules. Solubility, lipophilicity and permeability are all properties that affect the absorption of molecules *in vivo*.

Solubility is a measure of the propensity of a molecule to act as a solute in a specific solvent to form a homogenous solution wherein the solute molecules are dispersed and individually surrounded by solvent molecules.¹⁸² Solubility is commonly reported as a concentration.¹⁸³ At GSK, kinetic solubility is determined by the precipitation of a compound from solution and is measured using charged aerosol detection (CAD). Thermodynamic solubility is determined by the dissolution of a compound into solution and is measured using fasted-state simulated intestinal fluid (FaSSIF). High solubility is required for orally dosed drugs as only the soluble fraction is available for absorption in the gut, whilst parenteral drugs must be solubilised prior to dosing as a solution.

Lipophilicity is a measure of a molecule's ability to dissolve in fats, oils, lipids, and other non-polar solvents. Thus, *in vivo*, it reflects molecular desolvation and transfer from an aqueous phase to cell membranes and protein-binding sites, which are generally hydrophobic in nature.¹⁸⁴ Lipophilicity is experimentally measured as a partition coefficient between an aqueous buffer and the hydrophobic solvent *n*-octan-1-ol as either logP or logD.¹⁸⁵ logP is a measure of the intrinsic lipophilicity of a molecule and represents the partition of its unionised form. However, as most drug molecules contain ionisable moieties, pH can profoundly influence the charge of the molecules and thus their distribution between the phases. logD is a measure of the effective lipophilicity of a molecule and represents its partition at a defined pH. Unless stated otherwise, this is most commonly quoted at the pH of blood plasma, 7.4. ChromlogD is a chromatographically measured logD using a C18 reverse phase column at pH 7.4. It has a positive offset of approximately 2 log units from logD, which is retained to highlight the different origins of the data.¹⁸⁶

Permeability is a measure of a molecule's ability to pass through a lipid membrane and is a function of its lipophilicity and solubility. A major point of contention is whether this process is mediated through the lipid bilayer of the cell membrane or facilitated by transmembrane proteins.¹⁸⁷ These proteins are liberally dispersed throughout cell membranes and many have been shown to enable both passive permeation and active transport of specific molecules through the membrane.¹⁸⁸ However, the definitive mechanism of permeation is yet to be elucidated and is likely molecule-specific.

A number of assays can be used to measure permeability, each with their own limitations. Artificial membrane permeability (AMP) gives a measure of the rate of passive permeability through a lipid bilayer, negating the contribution of transmembrane proteins. Cellular assays using MDCK and HeLa cells give a more holistic measurement of permeability, such as cell concentration permeability (PAC).¹⁸⁹ However, in the absence of reliable assays, functional endpoints such as whole blood potency data for intracellular targets can be used as a surrogate for permeability.

The crux of successful drug development is finding the most favourable balance of physicochemical properties and efficacy. The difficulty of this is highlighted by comparing the properties of GSK oral candidates between 2001-2009 with the top 100 oral prescription drugs in 2009, **Figure 41**.¹⁹⁰

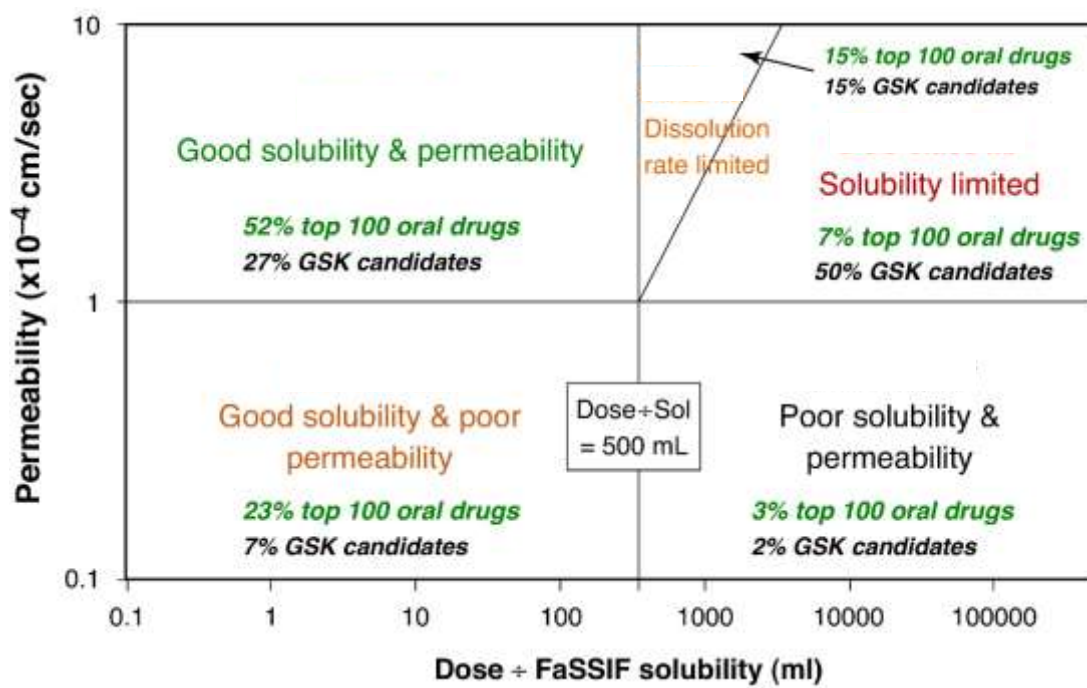


Figure 41: Comparison of GSK oral candidates between 2001-2009 with the top 100 oral prescription drugs in 2009.¹⁹⁰ Reprinted (adapted) with permission from Bayliss and co-workers.¹⁹⁰ Copyright 2016 Elsevier.

Only 27% of the GSK oral candidates and 52% of the top 100 oral drugs are in the most desirable top left quadrant. This is due to the complex interdependent relationship between different physicochemical properties and efficacy. Focusing on permeability as an example, this property technically shows a positive correlation with both solubility and lipophilicity independently. However, as these two properties show a negative correlation to each other, a balance must be achieved to maximise permeability.¹⁸⁶ Additionally, other properties need to be considered, as increasing lipophilicity has been correlated with high clearance and promiscuity for off-target tissues.¹⁹⁰

A range of other physicochemical descriptors that show correlation with the above properties can be used to obtain a comprehensive physicochemical profile. The size of

molecules is described by molecular weight (M_w) and calculated molar refractivity (CMR). Both properties show a positive correlation with size and are respectively derived from the molecule's molecular formula and its refractive index in aqueous solution.¹⁹¹ M_w shows a positive correlation with lipophilicity and thus permeability. However, a recent study highlighted a decrease in passive permeability and increase in transporter-mediated efflux for molecules with $M_w > 1000$ Da.¹⁹²

HBDs and HBAs form key binding interactions with the target protein that drive binding affinity and potency. The numbers of HBDs and HBAs are also reported as a physicochemical descriptor. Increasing the number of HBDs and HBAs that can form hydrogen bonds with water decreases the solvation energy, increasing the molecule's solubility. However, this results in a subsequent decrease in permeability as this solvation network must be shed prior to permeation through a membrane.¹⁸²

Topological polar surface area (TPSA) is a measure of the total surface area of all polar moieties in a molecule in \AA^2 . TPSA is a 2D calculation that only considers the atoms' connectivity and not their positioning in space. This property exhibits a positive correlation with solubility and number of HBDs and HBAs.¹⁹³ A negative correlation with permeability is also observed, due to unfavourable interactions between polar moieties and the lipophilic membrane.

Aromatic ring count (ARc) is also reported as a physicochemical descriptor, as aromatic rings enhance the planarity of molecules and promote π - π interactions. This results in an increase in the crystal lattice energy, thus reducing the aqueous solubility of the molecules.¹⁹⁴ A concurrent increase in lipophilicity is also observed, resulting in the associated trends with an increase in this property. However, ARc is a crude measurement, as all aromatic and heteroaromatic rings have the same contribution, irrespective of their structural composition and properties.¹⁹⁵

Fraction sp^3 (F_{sp^3}) is another measurement used to quantify saturation and molecular complexity. It is derived from the number of sp^3 hybridised carbons relative to the total carbon count. A higher F_{sp^3} value correlates with increased molecular complexity and other trends associated with a lower ARc.¹⁹⁶

The rotatable bond count (RBC) also gives an insight into the size and saturation of the molecule, exhibiting a positive correlation with lipophilicity. Recently, the flexibility added by increasing the RBC has been implicated in conferring permeability through dynamically exposed polarity in high M_w molecules.¹⁹⁷

Protein-binding measurements are also reported as physicochemical descriptors, as they can affect a molecule's pharmacokinetic and pharmacodynamic profile. They are routinely measured using high-throughput HPLC methods with biomimetic columns.

Human serum albumin (HSA) is the most abundant protein in human blood plasma and is responsible for binding endogenous compounds such as fatty acids, hormones and toxic metabolites and delivering them to their targeting tissues and organs.¹⁹⁸ HSA is also capable of binding drug molecules, with a preference for molecules containing acidic moieties, due to exposed basic residues.¹⁸² HSA binding gives a reliable indication of the free fraction of the molecule in plasma. Only unbound molecules are capable of entering the target tissue and eliciting the desired therapeutic effect. HSA binding exhibits a positive correlation with lipophilicity and ARc.¹⁸⁶

Immobilised artificial membrane (IAM) binding is measured using phosphatidylcholine, the most abundant phospholipid in mammalian cell membranes, as a surrogate.¹⁹⁹ IAM binding (CHI_{IAM}) gives an indication of a molecule's promiscuity and propensity to be sequestered into non-specific phospholipid binding sites, potentially incurring unwanted toxicity and side-effects.²⁰⁰ The net negative charge of the phosphate head groups in phosphatidylcholine results in a preference for binding molecules containing basic moieties. CHI_{IAM} exhibits a positive correlation with lipophilicity.²⁰¹

Due to the complexity of finding the most favourable balance of physicochemical properties and efficacy, medicinal chemists have developed general guidelines by analysing large data sets of molecules.

4.1.2 Physicochemical Property Guidelines

In 1997, Lipinski and co-workers published a seminal paper in which they analysed the physicochemical properties of 2245 molecules registered in phase II clinical trials from the world drug index (WDI) database. From the trends they identified, they defined five guidelines for molecules with good oral absorption properties, which exhibited oral bioavailability in humans.²⁰²

These guidelines are commonly referred to as the ‘Lipinski rule of five’ and define an upper limit for the physicochemical properties outlined in **Table 21**. Molecules can break one of the guidelines, but if two of the guidelines are broken then the molecule is predicted to have poor oral absorption and not exhibit oral bioavailability. In 2002, Veber and co-workers further developed these rules to include TPSA.²⁰³

Table 21: Guidelines developed by Lipinski & Veber for oral absorption and bioavailability of small molecules.^{202,203}

Properties	Guidelines
M _w / Da	≤ 500
HBA	≤ 10
HBD	≤ 5
logP	≤ 5
TPSA / Å ²	≤ 140
RBc	≤ 10

These arbitrary guidelines have been simplified to rounded numbers for ease of use, at the expense of a loss of detail. Additionally, the guidelines are most applicable to small molecules as < 11% of the data set had a M_w > 500 Da.

As PROTACs are large molecules, with molecular weights of 800-1200 Da, they sit in what is defined as ‘beyond the rule of five’ (bRo5) space. In 2018, DeGoey and co-workers from AbbVie developed guidelines for molecules in this space by analysing 1116 molecules from their compound collection with M_w > 500 Da which exhibited oral bioavailability. These rules define a revised upper limit for the physicochemical properties outlined in **Table 22**.²⁰⁴ Additionally, these guidelines corroborate with those developed by Kihlberg and co-workers in 2014, who analysed

485 clinical candidates and marketed drugs with molecular weights > 500 Da and < 3000 Da.²⁰⁵

Table 22: Guidelines developed by DeGoey and co-workers for oral bioavailability of molecules in bRo5 chemical space.²⁰⁴

Properties	Guidelines
M _w / Da	≤ 1132
HBA	≤ 14
HBD	≤ 6
logP	≤ 13.3
TPSA / Å ²	≤ 229
RBc	≤ 19

DeGoey and co-workers went on to define a simple multiparametric scoring function (AB-MPS) that correlates oral and parenteral bioavailability with clogD, RBc and ARc, **Equation 1**.²⁰⁴ Molecules with an AB-MPS score < 15 are predicted to exhibit oral bioavailability. Whilst molecules with an AB-MPS score < 28 are predicted to exhibit parenteral bioavailability.

$$\text{AB-MPS} = \text{Abs}(\text{clogD} - 3) + \text{RBc} + \text{ARc}$$

Equation 1: Equation used to calculate AB-MPS from clogD, RBc and ARc.

In 2019, Edmondson and co-workers from AstraZeneca further developed the rules outlined by DeGoey and co-workers to establish guidelines for PROTACs more specifically, **Table 23**.²⁰⁶ This was achieved by analysing 38 PROTACs with a range of different E3 ligase and protein-binding moieties.

Table 23: Guidelines developed by Edmondson and co-workers for oral bioavailability of PROTACs.²⁰⁶

Properties	Guidelines
M _w / Da	≤ 1000
HBA	≤ 17
HBD	≤ 5
logP	≤ 9
TPSA / Å ²	≤ 250
RBc	≤ 20

Synthesis of a series of PROTACs with different linkers whilst the target protein and E3 ligase-binding moieties remain constant will allow the effects of the linker functionality on the physicochemical properties of the PROTACs to be investigated.

The physicochemical property guidelines outlined above can be used to analyse the physicochemical space covered by this series, highlighting linkers that are occupying favourable physicochemical space. This may enable the identification of physicochemical trends that can be used to direct future linker choices and the development of PROTACs with more desirable physicochemical properties.

4.2 PROTAC Linker Library

A Reaxys search of all molecules that contain both a carboxylic acid and ester functional group was refined by the additional functionality, price and availability from suppliers, to furnish a library of 66 linkers, **Figure 42**.

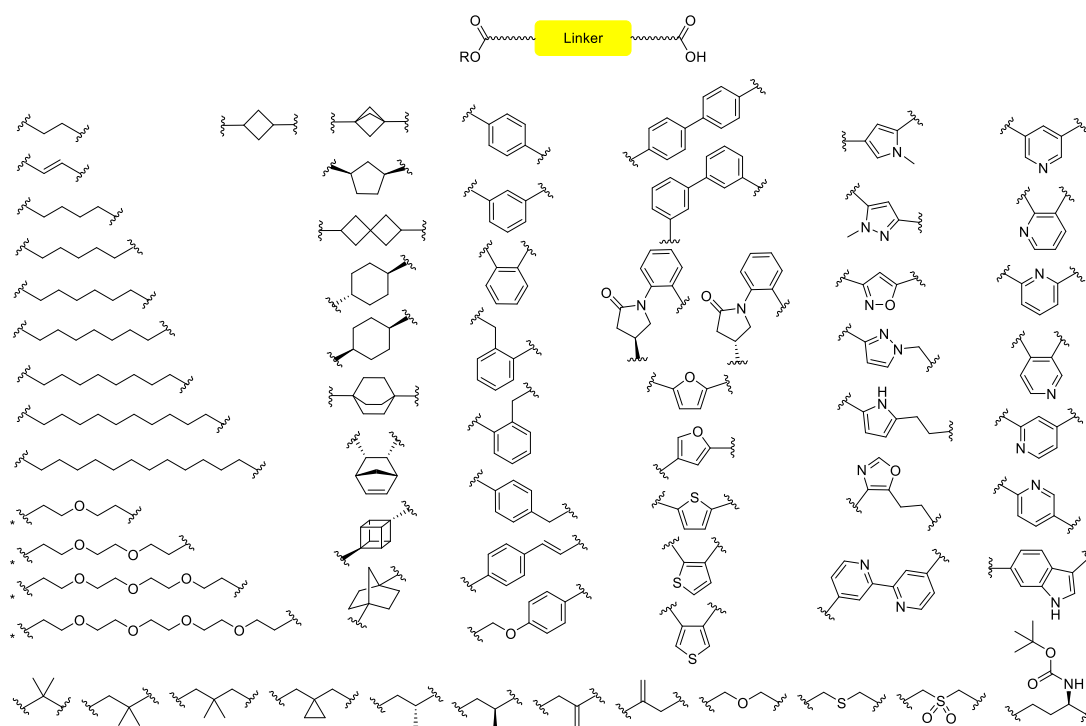


Figure 42: Terminal carboxylic acid-ester linker library. * = Diacid linker.

This library of linkers covers a range of lengths, from 2-12 carbons and 1-4 PEG units. This enabled the optimal linker length for degradation to be determined, whilst

providing broad coverage of physicochemical space. This library also covers a range of functionalities such as: simple alkyl, PEG, spirocyclic, aromatic, heteroaromatic, and some more substituted alkyl linkers.

The measurable and calculated properties of the PROTACs, listed in **Table 24**, were investigated to obtain a comprehensive overview of their physicochemical profiles. Relating these properties to both potency and degradation data allowed the physicochemical space to be evaluated and favourable properties identified.

Table 24: *Measurable and calculated properties used to assess the physicochemical profile of PROTACs.*

Measured Properties	Calculated Properties
ChromlogD	M _w
CAD Solubility	HBA
FaSSIF Solubility	HBD
AMP	logP
PΔC	TPSA
HSA Binding	RBC
IAM Binding	F _{sp} ³
Binding affinity (FRET and HWB pIC ₅₀)	ARc
Degradation (pDC ₅₀ , max % deg)	

Prior to synthesising the entire series of PROTACs from the linker library, some of the measured physicochemical properties were modelled to determine whether the series is providing sufficient coverage of the physicochemical space. These calculated values can then be validated against the measure properties of six PROTACs from the series that had been synthesised, **Figure 43**.

The graph in **Figure 44** compares the measured ChromlogD, on the y-axis, with the calculated heavy atom count ChromlogD (cChromlogD_{HAC}), on the x-axis. cChromlogD_{HAC} has been trained on a data set of compounds with > 70 heavy (non-hydrogen) atoms and generally shows good correlation with the measured ChromlogD for high M_w compounds.

The black line signifies direct correlation between these values. The red dashed line represents the trendline formed from the data points of the six PROTACs. As these

two lines are parallel, it confirms that the $c\text{ChromlogD}_{\text{HAC}}$ model is a good predictor of ChromlogD for the PROTACs. However, the red trendline does not exactly overlay the black line, suggesting that $c\text{ChromlogD}_{\text{HAC}}$ overestimates the measured ChromlogD, and an arbitrary shift factor is required to achieve a more accurate prediction from the model.

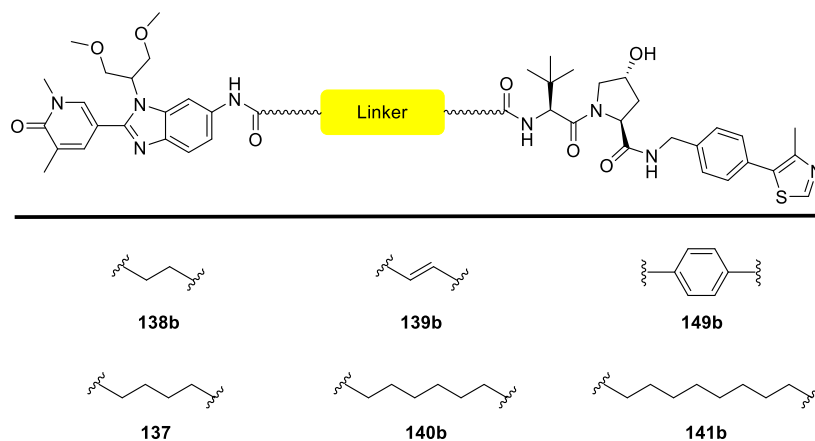


Figure 43: Six BET PROTACs synthesised from the linker library in **Figure 42**.

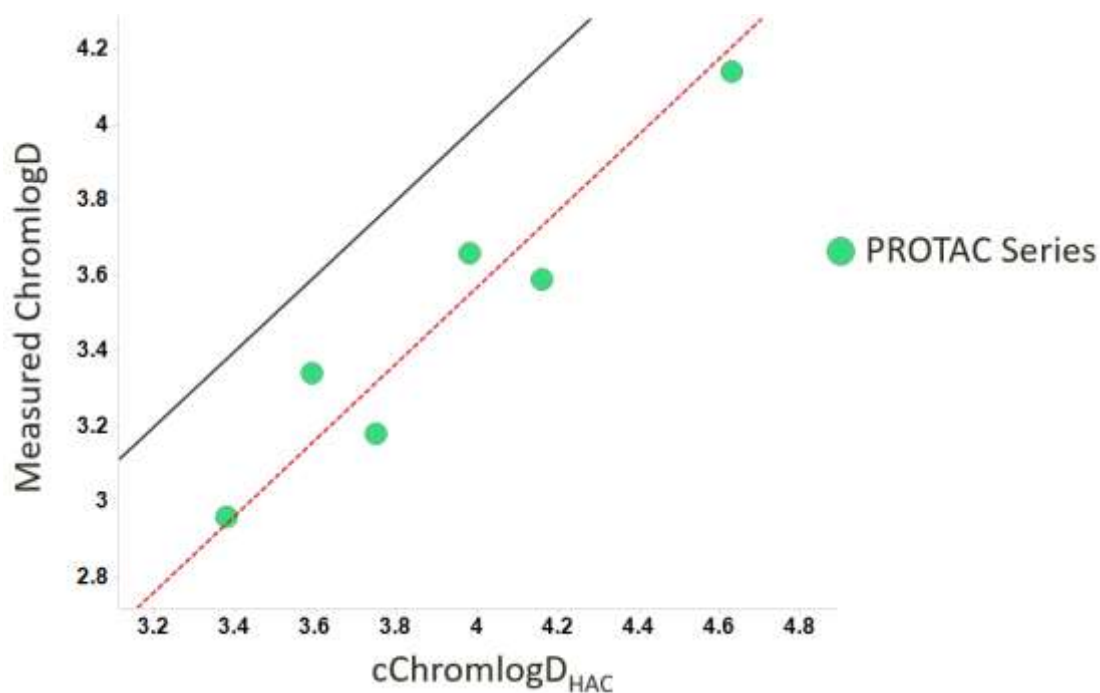


Figure 44: Comparison of the measured ChromlogD and $c\text{ChromlogD}_{\text{HAC}}$ values for the six PROTACs shown in **Figure 43**.

The AMP, which is a measure of passive permeability, can also be modelled by plotting lipophilicity vs size. For small molecules, AMP is modelled by plotting ChromlogD vs CMR. This model has been trained on a data set of marketed drugs and has historically shown good correlation for small molecules with a CMR < 14. However, the calculated CMR values for this PROTAC series are > 24.

It is possible to more accurately model the passive permeability of larger molecules by plotting calculated property forecast index (cPFI) vs aromatic ring adjusted size, **Figure 45**. cPFI is calculated from cChromlogD plus the number of aromatic rings, whilst aromatic ring adjusted size accounts for the intramolecular folding of larger molecules, thus reducing their effective size.

Molecules above the black dashed line in **Figure 45** are predicted to be permeable. Each point on the graph represents a VHL PROTAC registered at GSK with AMP data and is colour-coded to green if its measured AMP suggests it is permeable or red if not. The white points represent the calculated properties of the PROTAC series.

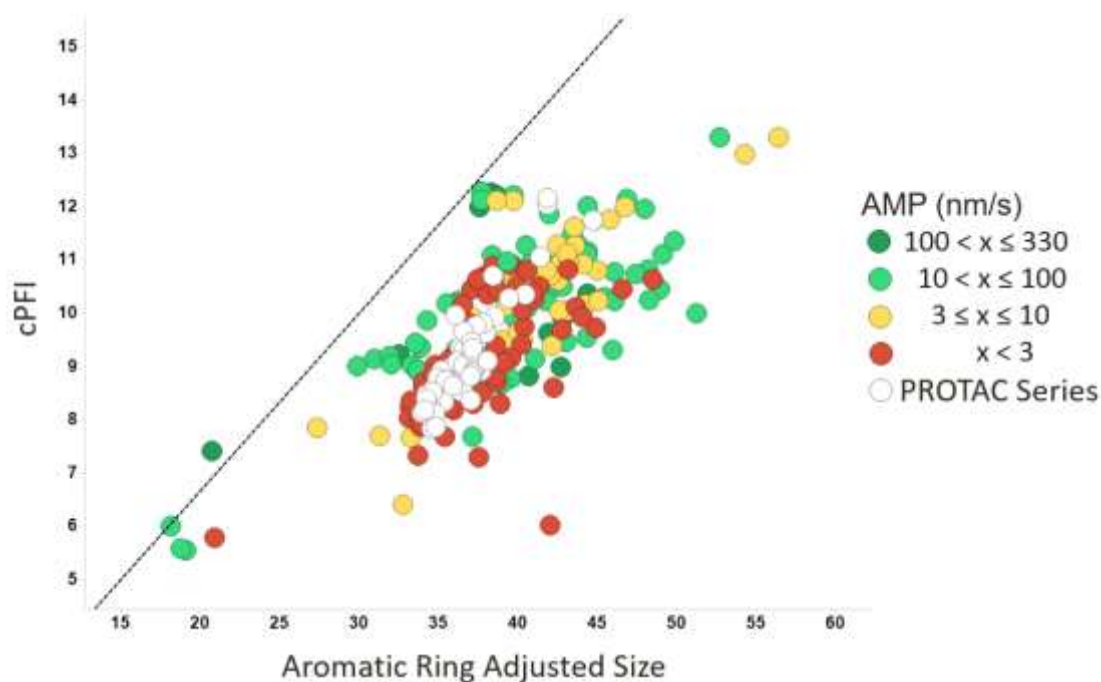


Figure 45: Plot of cPFI vs aromatic ring adjusted size. Each point on the graph represents a VHL PROTAC registered at GSK with AMP data. The points are coloured relative to their AMP values using the key in the figure. The white points represent the calculated properties of the PROTAC series. Molecules above the black dashed line are predicted to be permeable.

The model in **Figure 45** shows a lack of a clear trend in predicting permeability for the GSK VHL PROTACs that show some AMP permeability. Additionally, it predicts none of the PROTACs to be synthesised from the linker library to be permeable.

However, it is possible to compare the Brd4 human whole blood (HWB) and FRET potencies (pIC_{50}) for the six PROTACs from the series that have already been synthesised, **Figure 46**. These potencies are measured *in vitro* in cell-based and non-cell-based assays respectively and thus can be used to arbitrarily measure permeability.

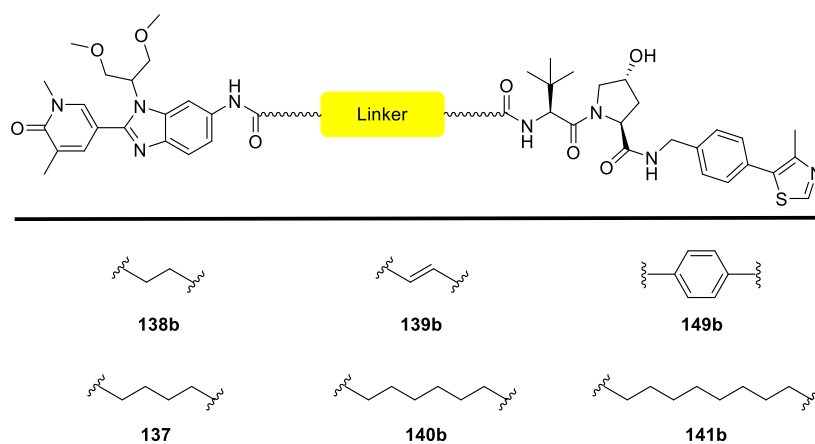


Figure 46: Six BET PROTACs synthesised from the linker library in **Figure 42**.

The six PROTACs from the series are represented by green points in **Figure 47**, with all BET inhibitors with the pyridone benzimidazole scaffold registered at GSK represented by the blue points. The black line signifies equivalence between the Brd4 HWB and FRET pIC_{50} , with the red lines showing a variance of ± 1 log unit.

The six PROTACs exhibited a small decrease in HWB pIC_{50} , within 1 log unit of their FRET pIC_{50} values. This is minimal in comparison to the larger drop off exhibited by the GSK registered BET inhibitors that exceeds 1 log unit. As BET proteins are intracellular targets, this suggests that the PROTACs are exhibiting permeability and thus the cPFI vs aromatic ring adjusted size model does not accurately predict permeability for this series of PROTACs.

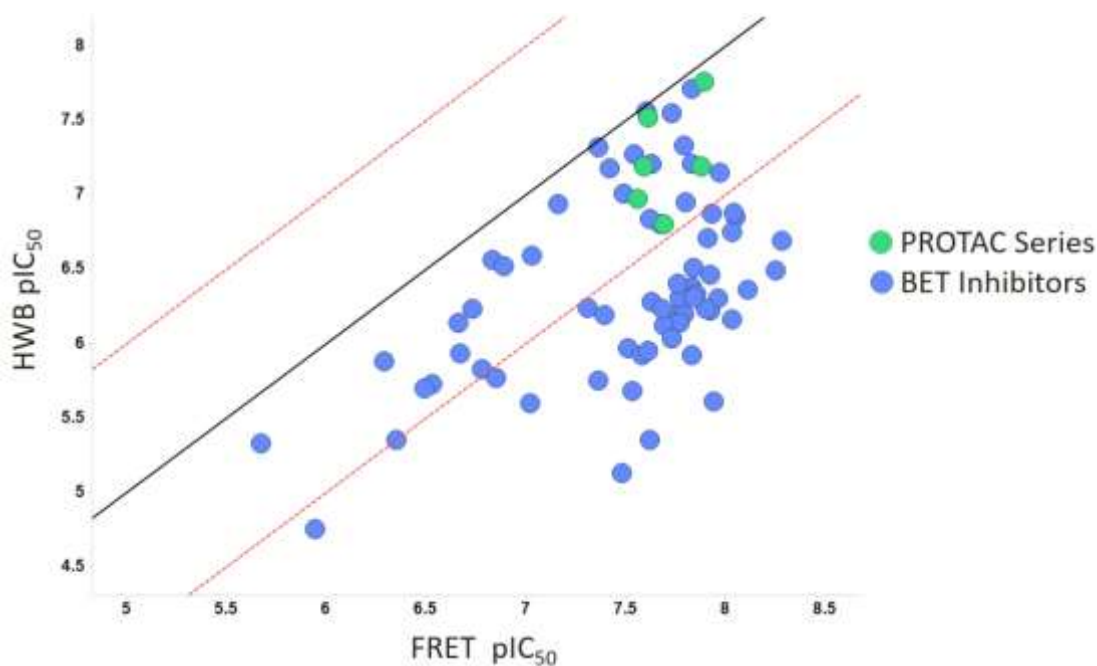


Figure 47: Plot of Brd4 HWB pIC_{50} vs Brd4 BD1 FRET pIC_{50} for the six PROTACs shown in Figure 46.

This highlights the complexity of accurately modelling measured physicochemical properties, such as solubility and permeability, due to the inherent limitations imposed by the data sets that are used to train the models. The calculated physicochemical properties, which are more empirical and can be calculated more reliably (cChromlogD/clogP), will be used to assess the physicochemical space covered by the PROTAC series synthesised from the linker library. These values can be used in conjunction with published guidelines to predict the bioavailability, which encompasses both solubility and permeability.

The guidelines developed by Edmondson and co-workers were applied to the calculated properties of the PROTAC series to be synthesised from the linker library.²⁰⁶ It is important to note that the primary aim of this analysis is not to assess the feasibility of developing an orally bioavailable PROTAC. Instead, these guidelines are used to analyse the physicochemical space covered by the PROTAC series, using properties indicative of oral bioavailability as guidelines for favourable physicochemical space.

It should also be noted that all the previously described guidelines (**Table 21, 22 and 23**) are arbitrary, and they do not represent a firm barrier between properties that do and do not facilitate oral bioavailability. Instead, they should be viewed as a continuum, and considered in conjunction with all properties to predict the probability of achieving oral bioavailability.

The colour-coding of properties for the guidelines developed by Edmondson and co-workers is presented in **Table 25**.²⁰⁶ A green colouring signifies that the value of the calculated property is significantly below that of the guidelines. A yellow colouring signifies that the value of the calculated property is below that of the guidelines, but within a significant tolerance. A red colouring signifies that the value of the calculated property exceeds that of the guidelines. The AB-MPS was also included and is colour coded using the guidelines developed by DeGoey and co-workers for parenteral bioavailability.²⁰⁴

Table 25: Colour-coding specified by Edmondson and co-workers to evaluate the physicochemical properties of PROTACs.^{204,206}

Properties	Green	Yellow	Red
M _w / Da	< 700	700-1000	> 1000
HBA	≤ 12	13-17	> 17
HBD	≤ 3	4	≥ 5
logP	< 6	6-9	> 9
TPSA / Å ²	< 200	200-250	> 250
RBC	< 15	15-20	> 20
AB-MPS	< 27.4	27.4-34.5	> 34.5

The guidelines in **Table 25** were applied to the calculated properties of the PROTAC series to be synthesised from the linker library, **Figure 48**. Each horizontal row represents a PROTAC with a different linker and the colours represent where the calculated physicochemical properties of the PROTAC sit relative to the guidelines for each property.

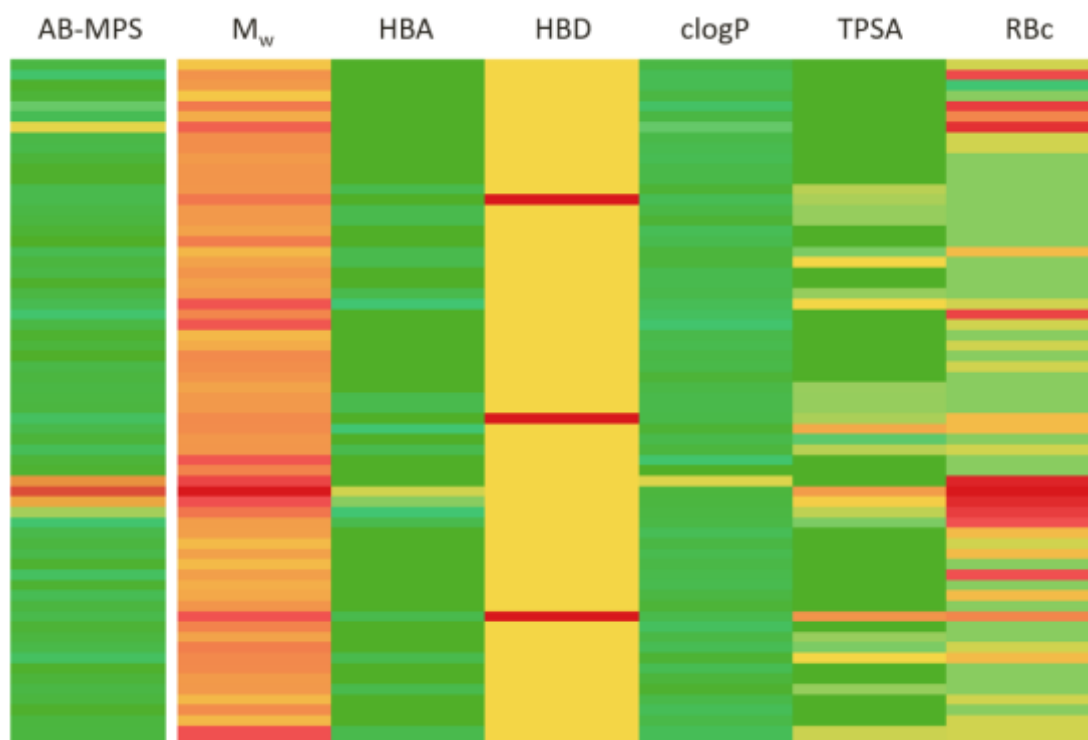
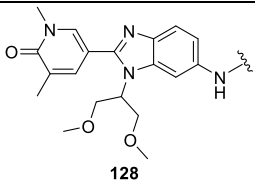
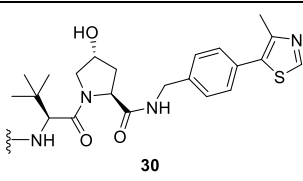


Figure 48: Visual representation of the calculated physicochemical properties for the PROTAC series to be synthesised from the linker library relative to the guidelines in **Table 25**.

Figure 48 shows that the PROTAC series is providing broad coverage of chemical space, within and outside of the guidelines. The only columns that are not predominantly green are M_w and number of HBDs. However, this is difficult to avoid, as these properties are inherited from the E3 ligase and protein binder, **Table 26**.

Table 26: Physicochemical properties of BET inhibitor **128** and VHL E3 ligase binder **30**.

Properties	 128	 30	Total
M_w / Da	355.3	429.5	784.8
HBA	4	4	8
HBD	1	3	4
clogP	1.9	1.8	3.7
TPSA / Å ²	70.3	94.6	164.9
RBC	7	7	14

Additionally, the AB-MPS predicts the majority of the compounds to be parenterally bioavailable, suggesting a favourable balance of solubility and permeability. However, none of the PROTACs are predicted to exhibit oral bioavailability. The application of these guidelines to the 13 VHL PROTACs published by Edmondson and co-workers is reproduced in **Figure 49**.²⁰⁶ These PROTACs utilised a range of target protein-binding moieties for a selection of targets, including six BET targeting PROTACs. These PROTACs cover a narrow range of properties, with the majority of them exhibiting high M_w , number of HBAs and RBc. **Figure 49** highlights that the PROTAC series to be synthesised from the linker library generally occupies more favourable physicochemical space than the VHL PROTACs published by Edmondson and co-workers.

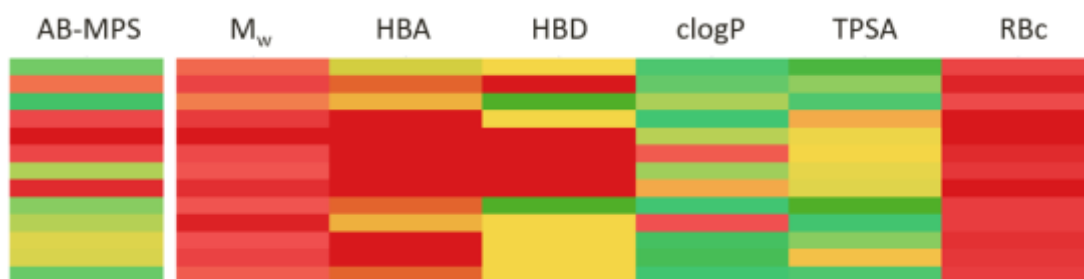


Figure 49: Visual representation of the calculated physicochemical properties for the VHL PROTAC series published by Edmondson and co-workers relative to their guidelines in **Table 25**.²⁰⁶

The PROTAC series to be synthesised from the linker library can also be compared to an in-house GSK series. The FAK-VHL series is characteristic of a generic linker optimisation campaign at GSK. The target protein and E3 ligase-binding moieties are kept constant whilst the linker is periodically changed, **Figure 50**. The guidelines developed by Edmondson and co-workers were applied to this series of 51 PROTACs, **Figure 51**.

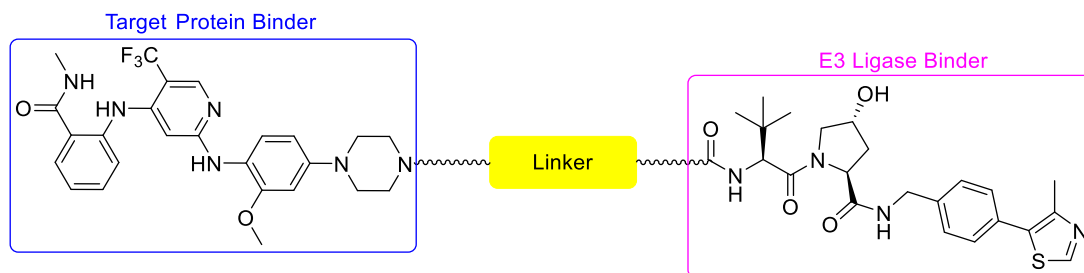


Figure 50: Target protein and E3 ligase-binding moieties of the FAK-VHL linker optimisation series.



Figure 51: Visual representation of the calculated physicochemical properties for the GSK FAK-VHL PROTAC series relative to the guidelines in **Table 25**.²⁰⁶

The PROTACs from the FAK-VHL series generally exhibit high M_w , number of HBDs and RBc. This is unavoidable for this series, as the target protein and E3 ligase-binding moieties have a combined M_w of 930 Da with 6 HBDs.

This highlights that the PROTAC series to be synthesised from the linker library is providing broad coverage of physicochemical space, analogous to that of an in-house GSK series. However, the calculated properties of the PROTAC series are in much more favourable physicochemical space as defined by Edmondson and co-workers. This should enable the effect of the linker functionality on the physicochemical properties of PROTACs to be analysed over a broad and favourable range.

This range is visualised in the radar plot in **Figure 52**. The guidelines for the upper limits of the six physicochemical properties developed by Lipinski & Veber, DeGoey and Edmondson are represented by green, red and blue lines respectively. The calculated properties for the PROTAC series are represented by the yellow lines, with the outer line representing the top of the range and the inner line representing the bottom. The broad coverage of favourable physicochemical space within and beyond the guidelines is achieved with a linker library that covers a diverse range of linker lengths and functionalities.

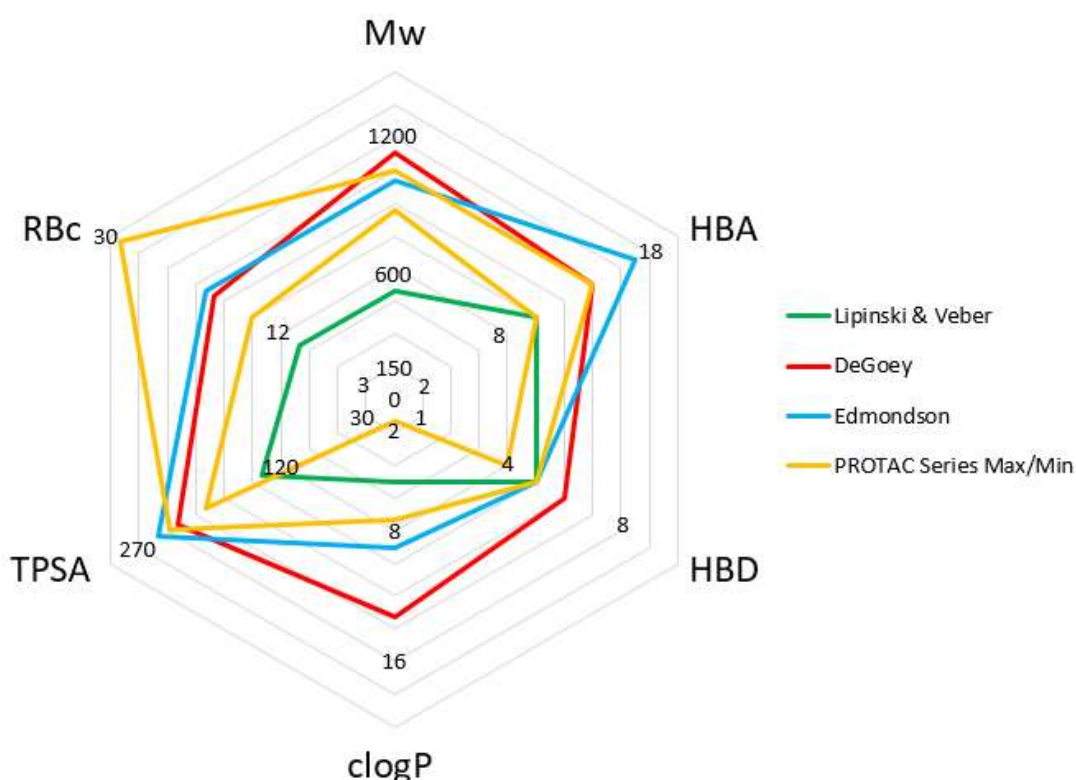


Figure 52: Radar plot of the chemical space covered by the PROTAC series to be synthesised from the linker library, as well as the guidelines published by Lipinski & Veber, DeGoey and Edmondson.

The distribution of properties within the physicochemical space defined by the PROTAC series was analysed, **Figure 53**. Each horizontal row represents a PROTAC with a different linker and the colours represent where the calculated physicochemical properties of the PROTAC sit relative to the range defined by the data set. Red signifies the top of the range and blue the bottom. **Figure 53** shows that the calculated properties

for the PROTAC series and linker library are evenly distributed within the range. The only properties that show limited variance are the number of HBAs and HBAs. However, these properties have a tight range as they are inherited from the target protein and E3 ligase-binding moieties and are already close to the limit of the chemical space chosen to sample. **Figure 53** also shows that PROTACs with linkers exploring the edges of the defined physicochemical space are not all clustered in subsets. This should allow the effects of different properties to be analysed independently. Additionally, the PROTAC series covers a range of Fsp³ (0.3-0.6) and ARc (5-7) values, highlighting the broad chemical diversity of the linker library.

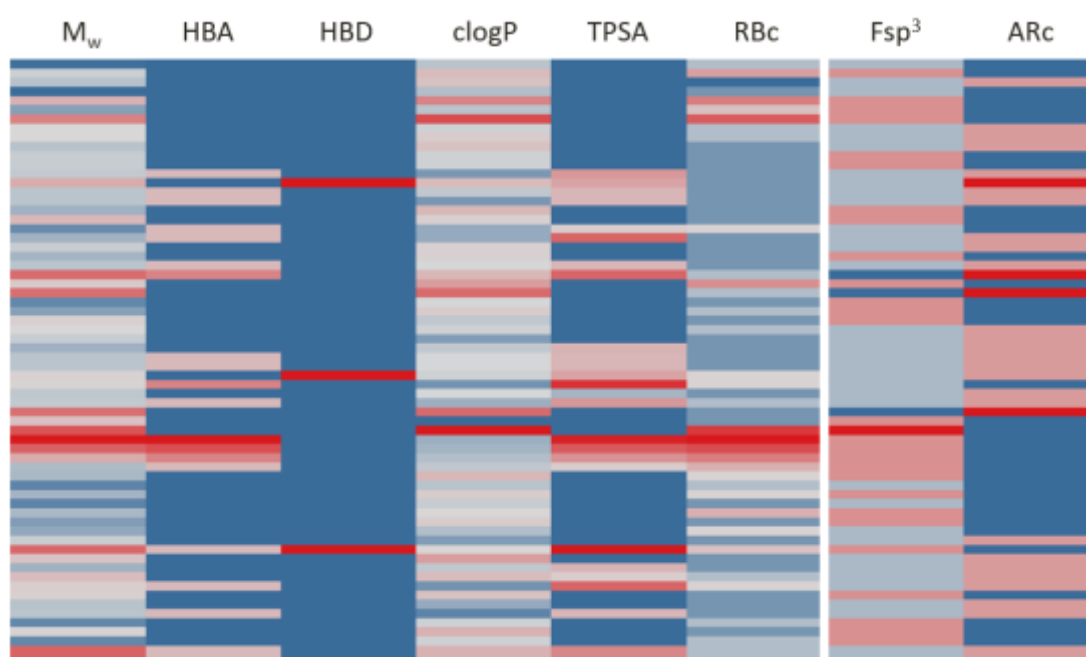


Figure 53: Visual representation of the distribution of the calculated physicochemical properties for the PROTAC series to be synthesised from the linker library.

4.3 PROTAC Series

The PROTAC series in **Figure 54** was synthesised from the linker library in **Figure 42** following the high-throughput, one-pot methodologies that were developed for the coupling of acid-ester and diacid linkers in a plate-based format, **Scheme 73**. This furnished 65 PROTACs in 4-56% isolated yield. However, as the absolute stereochemistry of PROTACs **183 Isomer 1** and **Isomer 2** could not be elucidated,

they were omitted from further analysis, furnishing a series of 63 PROTACs. The compound numbering of the PROTAC series can be viewed using the fold-out page in **Section 7.3**.

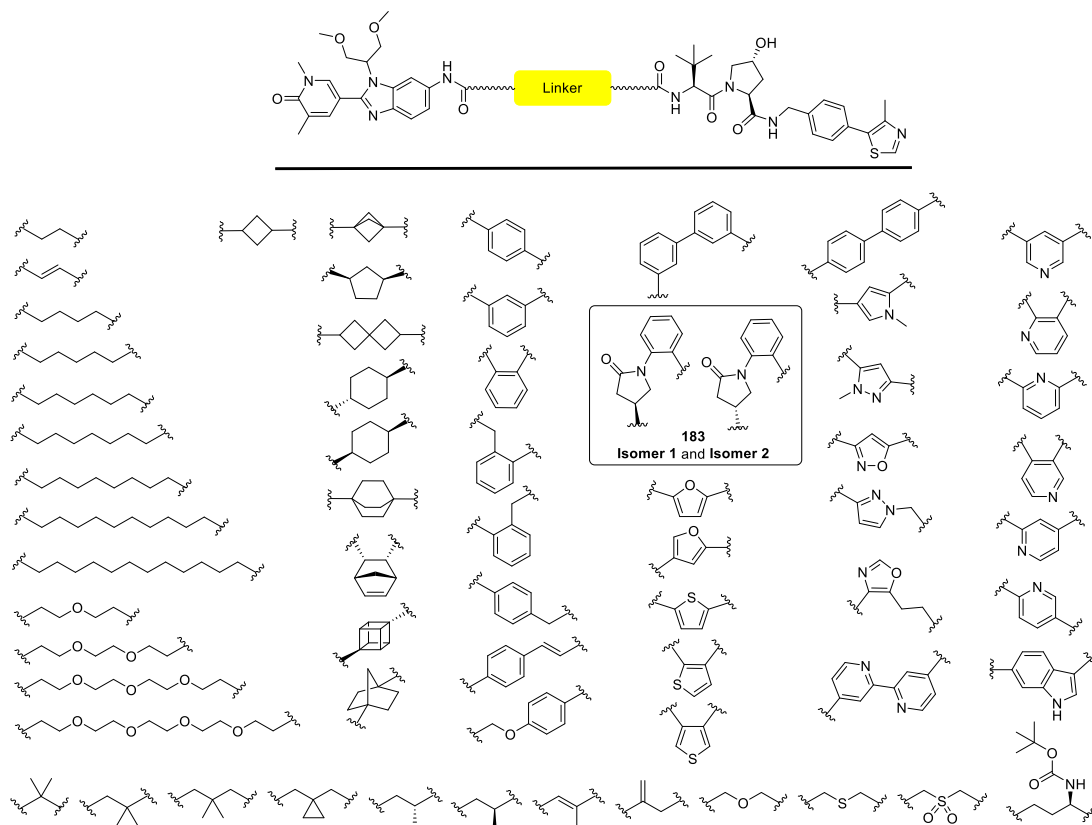


Figure 54: Sixty-five PROTACs synthesised from the linker library in **Figure 42**, with PROTACs **183 Isomer 1 and Isomer 2** omitted from further analysis highlighted.

All PROTACs were individually made up as 10 mM solutions in DMSO and stored at 4 °C prior to testing in the in-house biological assays outlined in **Table 24** and **Section 5.5**. Binding affinity and degradation data are reported as the average of duplicate and triplicate repeats, giving a value of $n = 2$ and $n = 3$ respectively. The complete data set for the degradation results are included in the digital appendices, highlighting results with an n value < 3 .

4.3.1 Brd4 Binding Affinity

The Brd4 binding affinity of the PROTAC series was measured using the *in vitro* FRET assay described in **Section 5.5** to generate pIC_{50} values. The Brd4 BD1 FRET

pIC₅₀ values for each PROTAC from the series are visualised by green points in **Figure 55**. The blue points represent the FRET pIC₅₀ values of the protein binders **110** and **128**. The unfunctionalised protein binder **110** has a FRET pIC₅₀ of 7.33, represented by the dashed line. All of the PROTACs from the series except PROTAC **169** exhibit a higher binding affinity to Brd4 BD1 than the protein binders **110** and **128**. Although PROTAC **169** exhibits a lower FRET pIC₅₀ of 7.32, this difference is within the error of the assay. This confirms that linking protein binder **110** to the E3 ligase binder **30** through the 6-position on the benzimidazole ring has no deleterious effects on its Brd4 binding affinity.

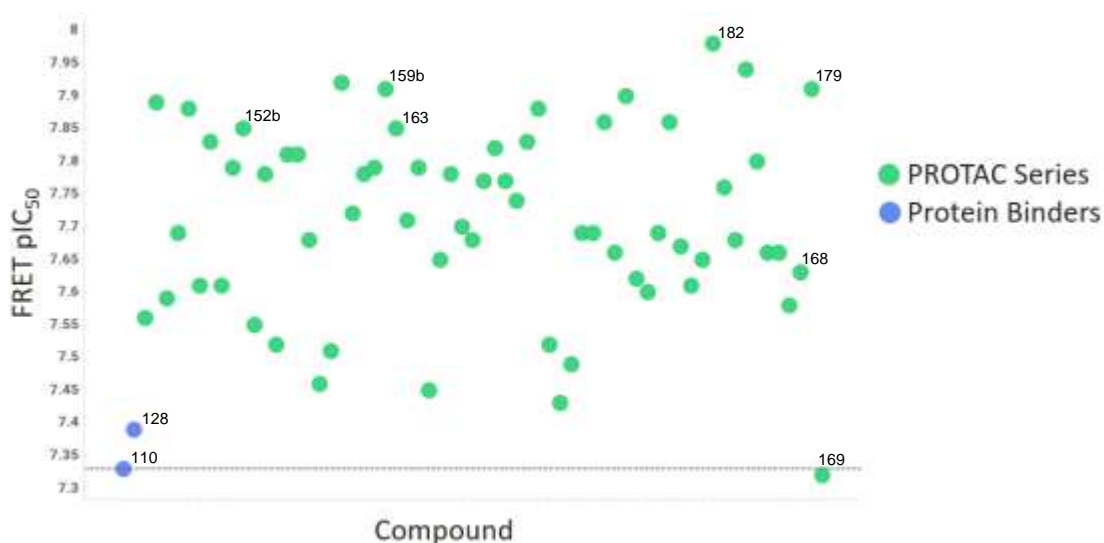


Figure 55: Plot of Brd4 BD1 FRET pIC₅₀ values for the PROTAC series shown in green, with protein binders **110** and **128** represented by blue points.

Interestingly, PROTAC **168** with the (*R*)-Me linker, the diastereoisomer of PROTAC **169**, has a higher FRET pIC₅₀ of 7.63. These PROTACs have identical calculated physicochemical properties (M_w, HBA, HBD, clogP, TPSA, RBC, ARc and Fsp³). This suggests that additional binding affinity is inferred by the geometric conformation of the PROTAC. The X-ray crystal structure of PROTAC **168** was obtained by cocrystallisation with human Brd4 BD1, **Figure 56**. The cocrystal structure highlights additional binding interactions between the VHL E3 ligase-binding moieties of PROTAC **168** and the Brd4 protein, in particular, between Lys91 and the carbonyl adjacent to the hydroxy-proline ring of the VHL E3 ligase-binding moiety. This

additional binding interaction may be less accessible for PROTAC **169** due to steric and torsional constraints, resulting in no additional binding affinity and increase in FRET pIC_{50} relative to protein binder **110**. This suggests that the PROTACs from the series that exhibit a FRET pIC_{50} above the dashed line in **Figure 55** are obtaining some additional binding affinity from supplementary binding interactions between the linker or VHL E3 ligase-binding moiety and Brd4 BD1. PROTAC **182**, containing the 3-carbon linker with a Boc-protected primary amine substituent, exhibited the highest FRET pIC_{50} of 7.98. This may be due to additional hydrogen-bonding interactions between the Boc-protected primary amine that can act as both a HBA and HBD with complementary Brd4 BD1 residues.

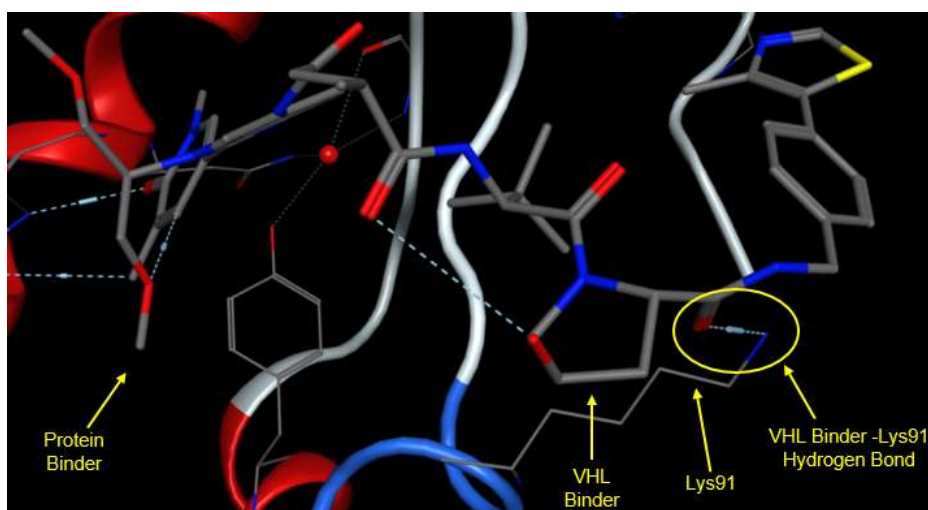


Figure 56: X-ray cocrystal structure of PROTAC **168** with human Brd4 BD1 protein, highlighting the hydrogen-bonding interaction between the VHL E3 ligase-binding moiety and Lys91.

The cell-based Brd4 binding affinity of the PROTAC series was measured using the HWB assay described in **Section 5.5** to generate pIC_{50} values. As this is a cellular assay, these measurements will give an insight into the permeability of the molecules. The Brd4 HWB pIC_{50} values for each PROTAC from the series are visualised by green points in **Figure 57**. The blue points represent the HWB pIC_{50} values of the protein binders **110** and **128**. The unfunctionalised protein binder **110** has a HWB pIC_{50} of 7.08, represented by the dashed line. Both protein binders exhibited a decrease in HWB pIC_{50} relative to their FRET pIC_{50} of 0.25 and 0.23 respectively. A wide range of HWB pIC_{50} values is observed for the PROTAC series, ± 1.5 log units of the protein binder

110. This suggests that PROTACs exhibiting a HWB pIC_{50} below the dashed line either observe a lower permeability or higher % protein-binding than the protein binders **110** and **128**. Therefore, there is less available free PROTAC in the cell that can engage the Brd4 protein, furnishing a lower HWB pIC_{50} . PROTAC **199** exhibits the lowest HWB pIC_{50} of 5.53, which is significantly lower than its FRET pIC_{50} of 7.65. This decrease in over 2 log units may be due to the poor cellular permeability of the compound, which exhibits an AMP of < 3 nm/s, in comparison to protein binder **110**, which exhibits an AMP of 291 nm/s.

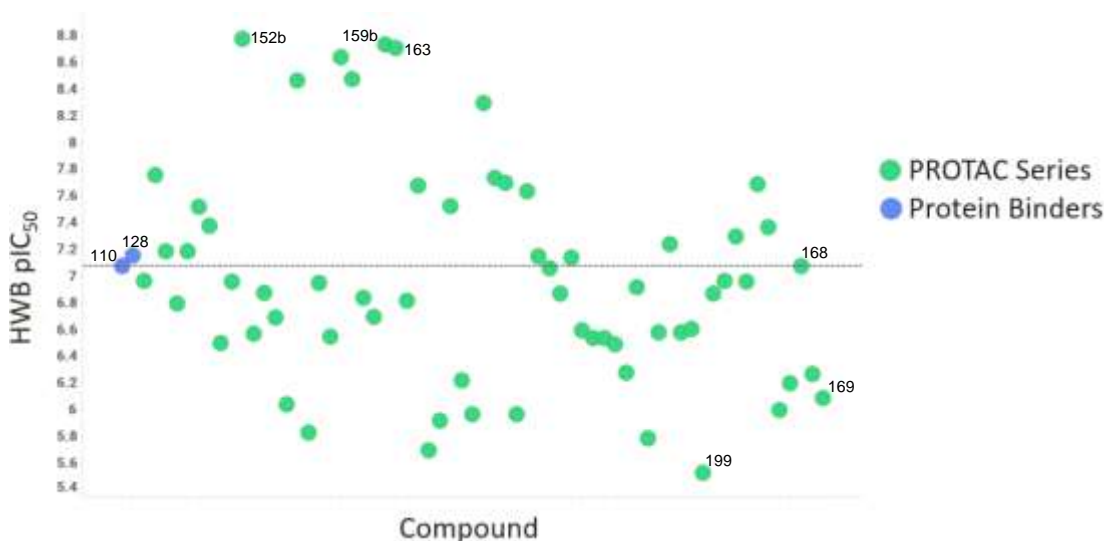


Figure 57: Plot of Brd4 HWB pIC_{50} values for the PROTAC series shown in green, with protein binders **110** and **128** represented by blue points.

The differences between the FRET pIC_{50} and HWB pIC_{50} are visualised in **Figure 58**. The solid black line represents equivalence between the two pIC_{50} measurements, with the red dashed lines representing ± 1 log unit. PROTACs above the solid black line exhibit a HWB pIC_{50} higher than their FRET pIC_{50} , whereas PROTACs below the line exhibit a HWB pIC_{50} lower than their FRET pIC_{50} .

No correlation between FRET pIC_{50} and HWB pIC_{50} can be observed for the PROTAC series. PROTAC **152b** exhibits the highest HWB pIC_{50} of 8.78, **Figure 58**, which is significantly higher than its FRET pIC_{50} of 7.85. As this is a cell-based measurement, if PROTAC **152b** is capable of productive ternary complex formation, this may be due to the degradation of Brd4 protein *via* the cell's UPS. This would reduce the overall

amount of Brd4 protein in the cell and thus increase the observed HWB pIC_{50} . Therefore, PROTAC **152b** may also be exhibiting a lower cell permeability than protein binder **110**, however, due to the catalytic nature of PROTACs this effect may be being masked by the increase in HWB pIC_{50} .

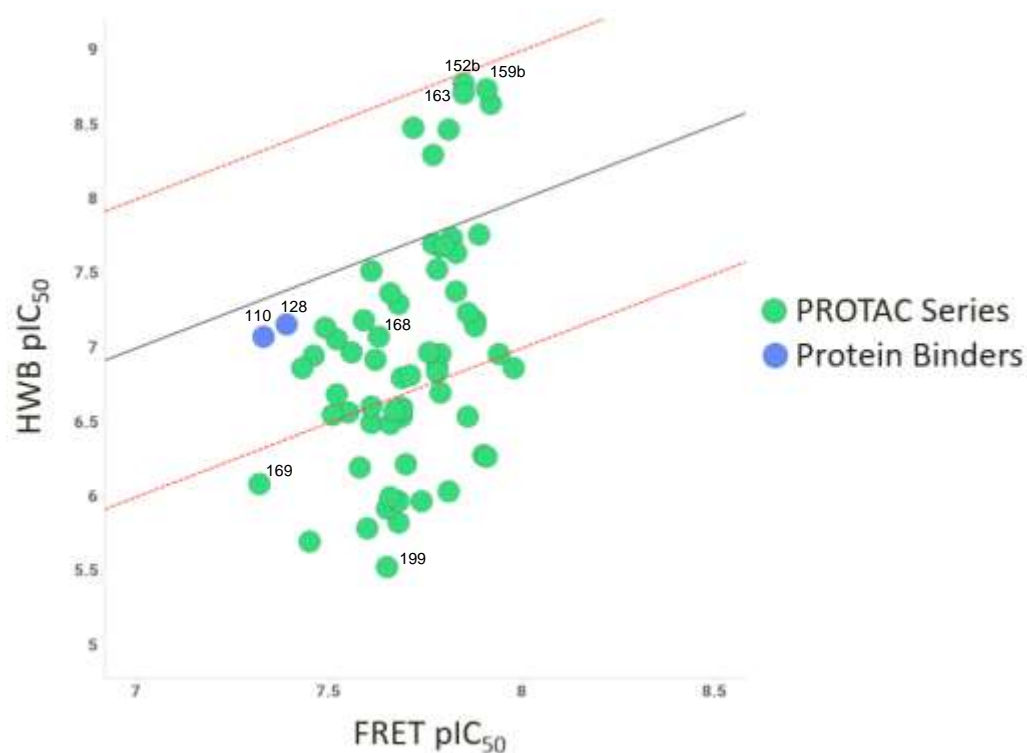


Figure 58: Plot of Brd4 HWB pIC_{50} vs FRET pIC_{50} for the PROTAC series shown in green, with protein binders **110** and **128** represented by blue points. The solid black line represents equivalence between the two pIC_{50} measurements, with the red dashed lines representing ± 1 log unit.

The differences between the FRET pIC_{50} and HWB pIC_{50} are more explicitly quantified in **Figure 59**. The dashed black line represents the decrease in HWB pIC_{50} relative to FRET pIC_{50} observed for protein binder **110**. Due to the large M_w of the PROTACs, their negligible passive permeability and similar ChromlogD measurements to protein binder **110**, it is not surprising that 48 of the PROTACs in the series exhibit a larger decrease in HWB pIC_{50} relative to their FRET pIC_{50} that may be attributed to a lower permeability.

The complex trends between physicochemical properties and binding affinity will be outlined and discussed further below.

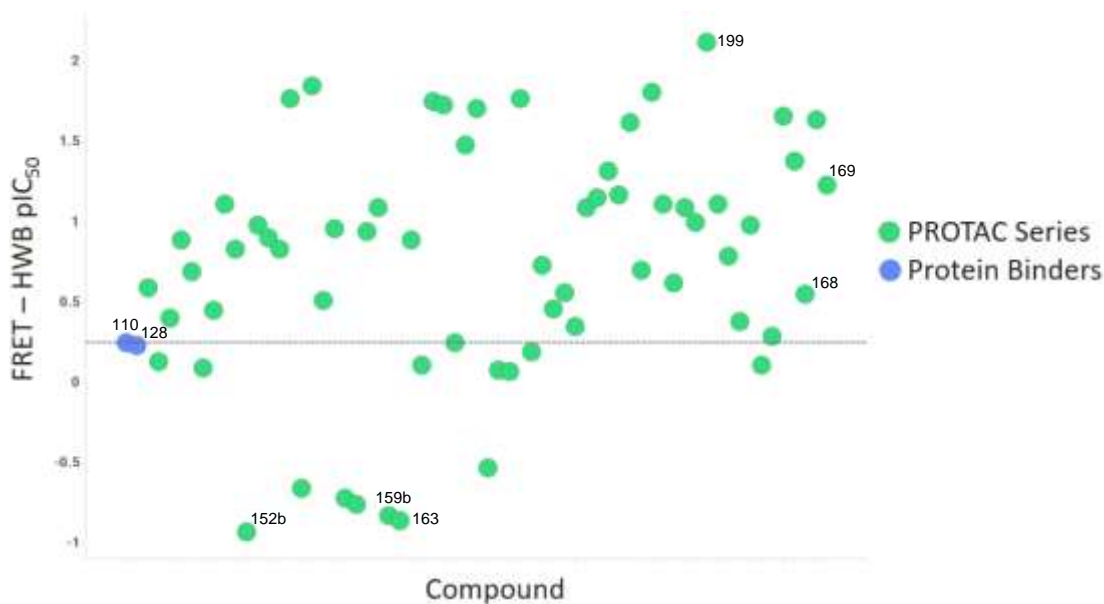


Figure 59: Plot of Brd4 FRET - HWB pIC₅₀ for the PROTAC series shown in green, with protein binders 110 and 128 represented by blue points.

4.3.2 Brd4 Degradation

Brd4 degradation data was obtained using the in-house HiBiT assay developed by Promega, **Figure 60** and **Section 5.5**.²⁰⁷ Gene editing CRISPR-Cas9 technology was used to encode the 11 amino acid HiBiT tag just before the Brd4 gene in DNA, resulting in a natural level of cellular expression of HiBiT-tagged Brd4. These cells are incubated with varying concentrations of PROTAC for 18 h. The cells are then lysed and incubated with the 18 kDa LgBiT tag. The HiBiT and LgBiT tags have picomolar affinity and bind to form the NanoBiT, which is a luciferase protein that luminesces. The intensity of luminescence is analysed to quantify the amount of Brd4 protein remaining.

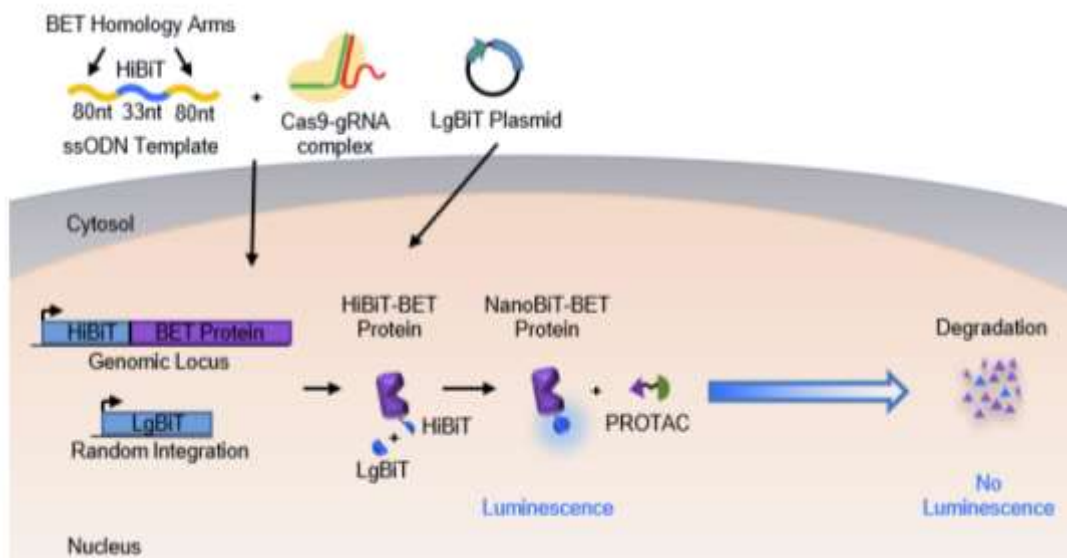


Figure 60: Schematic representation of the HiBiT degradation assay used to quantify *Brd4* degradation.²⁰⁷ Reprinted (adapted) with permission from Riching and co-workers.²⁰⁷ Copyright 2018 American Chemical Society.

The degradation profile of each PROTAC was analysed by plotting % *Brd4* remaining vs log concentration of the PROTAC, **Figure 61**. The three key parameters reported from the *Brd4* degradation assay are maximum % degradation (max % deg), pDC₅₀ and relative DC₅₀ (rDC₅₀).

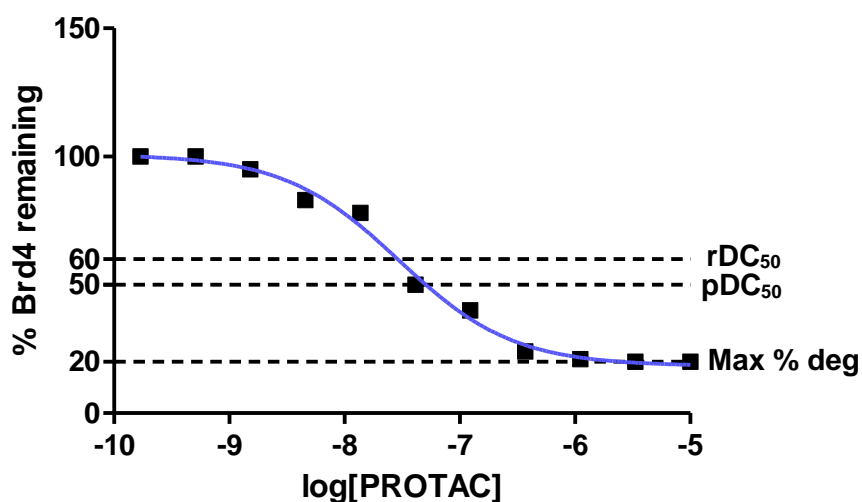


Figure 61: Degradation profile for a model PROTAC with key degradation parameters labelled.

The max % deg is reported from 100% minus the minimum value of % Brd4 remaining, giving a max % deg of 80% in **Figure 61**. The rDC₅₀ is reported at 50% of the max % deg, 60% in **Figure 61**. The pDC₅₀ is always reported at 50% Brd4 remaining, with a pDC₅₀ value of < 5 reported for PROTACs that did not achieve 50% Brd4 degradation.

The Brd4 max % deg values for each PROTAC from the series are visualised by green points in **Figure 62**. The blue points represent the max % deg values of the protein binders **110** and **128**, with the pink point representing the VHL E3 ligase binder **30**. As expected, negligible (< 10%) Brd4 degradation is observed with the unfunctionalised protein binder **110** and VHL E3 ligase binder **30**, as they are unable to simultaneously bind both the Brd4 protein and E3 ligase to form a ternary complex that facilitates ubiquitination and subsequent degradation of Brd4. However, a higher max % deg of 27% is observed for the protein binder **128**, this may be due to the **128** bound Brd4 protein being recognised as foreign by the cell and targeted for degradation.

A wide range of max % deg values are observed for the PROTAC series. PROTAC **163** exhibited the highest Brd4 max % deg of 98%, with PROTACs **152b**, **159b** and **180** within the assay error of the highest value. PROTACs **152b** and **159b** both contain a linker with a 2,6-substituted aryl ring, these trends will be outlined and investigated further below. PROTAC **147b**, with the short rigid bicyclo[1.1.1]pentane linker, exhibited the lowest max % deg of 12%.

PROTAC **169** exhibited the second lowest Brd4 max % deg of 16%, whilst its diastereoisomer with the (*R*)-Me linker **168** exhibited a significantly higher max % deg of 35%. This may be due to the higher binding affinity, FRET and HWB pIC₅₀ values, observed for PROTAC **168**, indicating that it is better able to accommodate a favourable binding conformation, increasing the duration of Brd4 engagement and thus the likelihood of ubiquitination of Brd4. However, it could also be due to the (*R*)-Me linker of PROTAC **168** being better able to form a productive ternary complex conformation due to potential clashes between the (*S*)-Me linker of PROTAC **169** and the Brd4 protein or VHL E3 ligase. Additionally, PROTAC **169** exhibited a larger

decrease in HWB pIC_{50} relative to its FRET pIC_{50} , which may suggest that it has a lower permeability than PROTAC **168**.

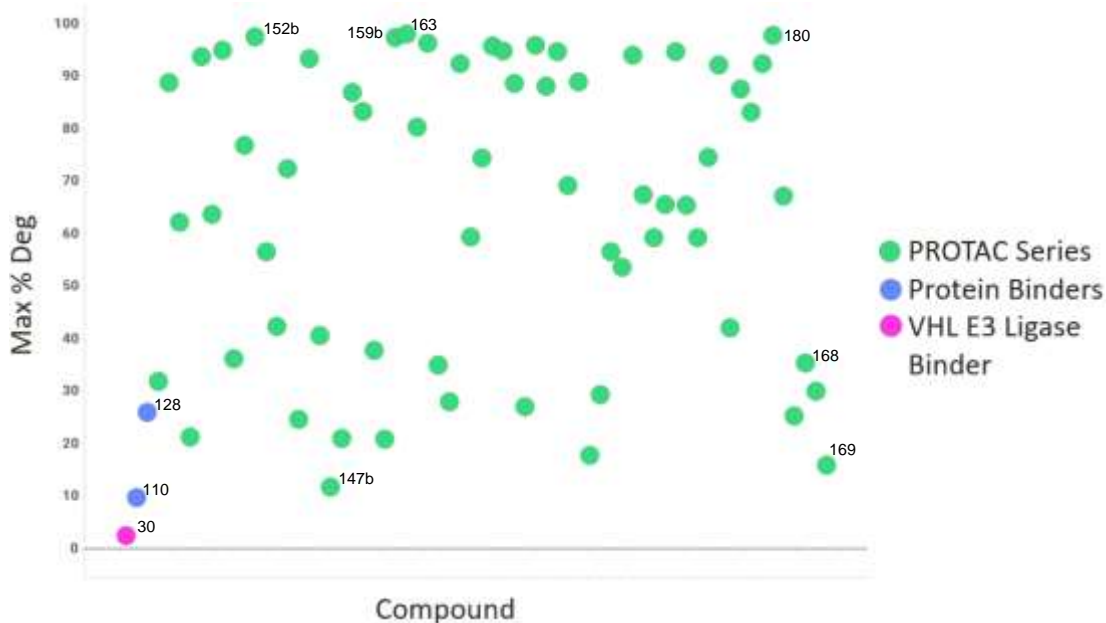


Figure 62: Plot of Brd4 max % deg for the PROTAC series shown in green, with protein binders **110** and **128** represented by blue points and VHL E3 ligase binder **30** by the pink point.

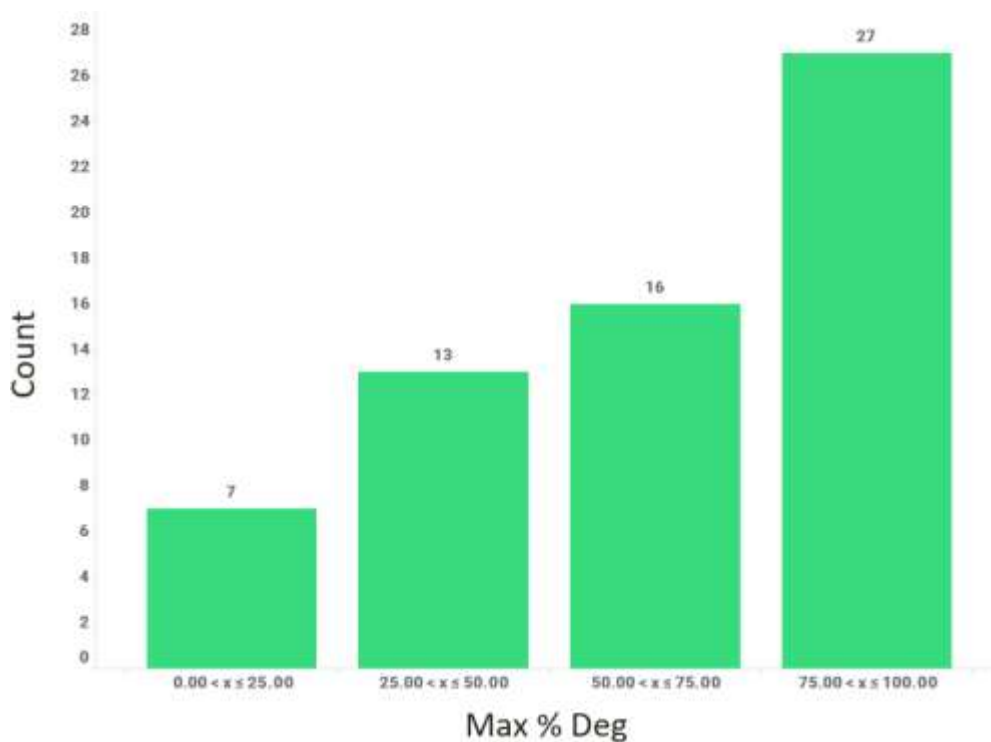


Figure 63: Histogram plot of the measured Brd4 max % deg values of the PROTAC series.

The distribution of max % deg values for the PROTAC series can be visualised in the histogram in **Figure 63**. The PROTAC series shows broad coverage of max % deg values, with more than 40% of the PROTAC series exhibiting a max % deg > 75%.

The Brd4 pDC₅₀ values for each PROTAC from the series are visualised by green points in **Figure 64**. The blue points represent the pDC₅₀ values of the protein binders **110** and **128**, with the pink point representing the VHL E3 ligase binder **30**. The lowest pDC₅₀ value measurable in the Brd4 degradation assay is 5, any PROTACs that do not achieve 50% degradation are assigned a pDC₅₀ of < 5, represented by the black dashed line in **Figure 64**. As expected, a pDC₅₀ of < 5 is observed with the protein binders **110** and **128**, and the VHL E3 ligase binder **30**, as they are unable to simultaneously bind both the Brd4 protein and E3 ligase to form a ternary complex that facilitates ubiquitination and subsequent degradation of Brd4 protein. Twenty PROTACs (**138b**, **139b**, **144b**, **146b**, **147b**, **148b**, **150b**, **153b**, **156b**, **157b**, **164**, **165**, **168**, **169**, **173**, **179**, **187**, **190**, **210** and **211**) from the series exhibited a pDC₅₀ value of < 5 as they did not achieve 50% degradation of the Brd4 protein. The details to distinguish the degradation profiles of these compounds is lost with the generic assignment of < 5, therefore, other measurements such as max % deg will be relied upon.

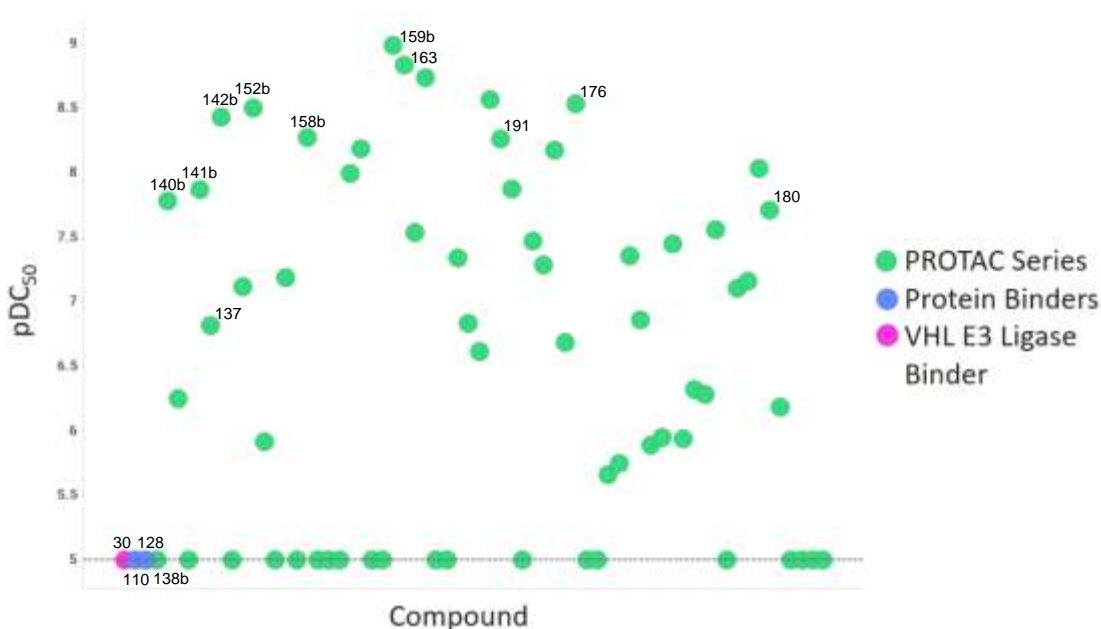


Figure 64: Plot of Brd4 pDC₅₀ for the PROTAC series shown in green, with protein binders **110** and **128** represented by blue points and VHL E3 ligase binder **30** by the pink point.

PROTAC **159b** exhibited the highest pDC₅₀ of 8.99, which equates to nanomolar potency and is considered highly potent. PROTACs **152b**, **158b** and **191** containing a similar 2,6-substituted aryl ring linker exhibited pDC₅₀ values > 8; these trends will be outlined and investigated further below.

Focusing on the PROTACs containing unfunctionalised carbon only linkers: 2-carbon **138b**, 4-carbon **137**, 6-carbon **140b**, 8-carbon **141b**, 10-carbon **142b** and 12-carbon **176**, an increase in pDC₅₀ is observed as the length of the linker increases, with PROTAC **176** exhibiting the highest pDC₅₀ of 8.54. This suggests that the longer, more flexible linkers are better able to accommodate a ternary complex conformation that facilitates ubiquitination and subsequent degradation of the Brd4 protein. However, PROTAC **176** and PROTAC **142b** exhibit similar pDC₅₀ values that are within the error of the assay. This suggests that beyond an optimum linker length, further increases in conformational flexibility furnish no further increases in pDC₅₀. It is expected that even further increases in linker length and conformational flexibility would result in a decrease in pDC₅₀, due to an increase in the number of possible conformations and therefore a decrease in the population of the productive ternary complex conformation required for degradation.

PROTAC **159b** and **163** achieved the highest pDC₅₀ values of 8.99 and 8.84 respectively. These PROTACs contain more rigid linkers and therefore have a lower amount of conformational flexibility. This suggests that preorganisation of the PROTACs into a conformation that can facilitate productive ternary complex formation furnishes an increase in Brd4 degradation.

A clear positive correlation is observed between pDC₅₀ and max % deg for the PROTAC series, **Figure 65**. Interestingly, although both PROTACs **159b** and **180** achieved a max % deg of 98%, PROTAC **180** exhibited a significantly lower pDC₅₀ value of 7.71. This indicates that a higher dose concentration of PROTAC **180** is required to elicit 50% of the maximum Brd4 degradation, **Figure 66**.

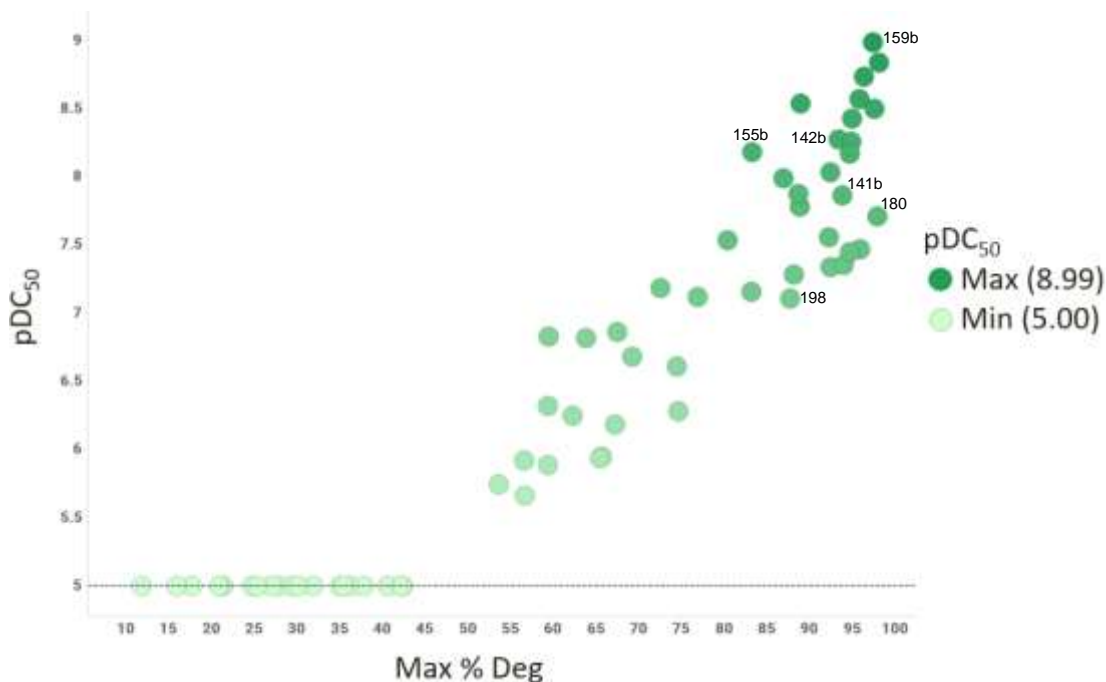


Figure 65: Plot of pDC₅₀ vs max % deg for the PROTAC series shown in green. The shading of the points corresponds to their pDC₅₀ values, with dark green representing a pDC₅₀ of 8.99 and light green representing a pDC₅₀ of < 5.00.

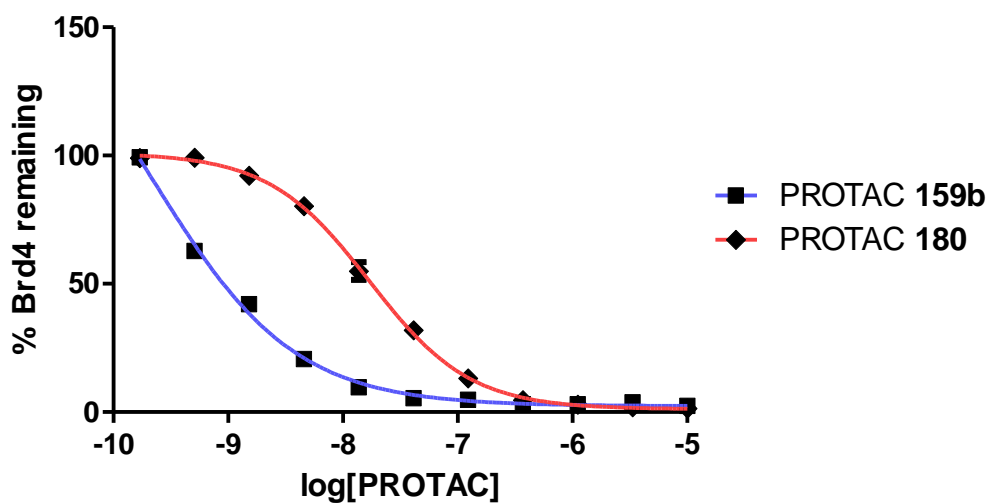


Figure 66: Plot of % Brd4 remaining vs log[PROTAC] for PROTACs 159b (blue) and 180 (red).

This may be because PROTAC 180 is less able to occupy a conformation that facilitates ubiquitination and subsequent degradation of Brd4, due to the incurrence of steric and torsional strain. However, it may also suggest that PROTAC 180 exhibits a

lower cell permeability; therefore, a high dosing concentration of PROTAC **180** is required to obtain the same level of intracellular concentration as PROTAC **159b**. A similar conclusion can be postulated for PROTACs **142b** and **141b**, containing an unfunctionalised 10- and 8-carbon linker, and PROTACs **155b** and **198** containing a substituted isoxazole and furan ring respectively.

An interesting phenomenon termed the hook effect (**Section 1.4, Figure 10 and 11**) is observed for the PROTAC with the 4-carbon linker **137, Figure 67**. As the concentration of PROTAC increases the concentration of productive ternary complexes and subsequent Brd4 degradation increases to a saturation point. As the concentration of PROTAC increases beyond this point, non-productive binary complexes are favoured resulting in a decrease in degradation. The appearance of this hook effect may be due to negative cooperativity in the ternary complex, due to unfavourable protein-protein and protein-PROTAC interactions.

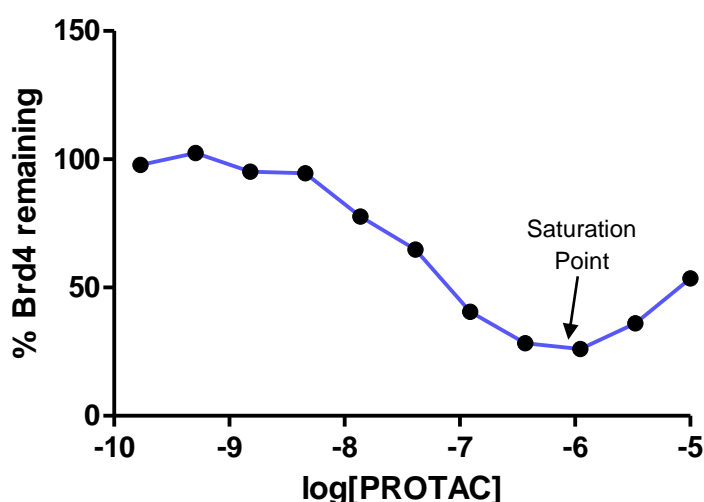


Figure 67: Plot of % Brd4 remaining vs log[PROTAC] for PROTAC **137**.

The distribution of pDC₅₀ values for the PROTAC series can be visualised in the histogram in **Figure 68**. The PROTAC series shows broad coverage of over 4 log units of potency, with approximately 60% of the PROTACs exhibiting a pDC₅₀ value > 6, equating to micromolar potency, which is a common benchmark of desirable potency (pIC₅₀) for classical inhibitors.

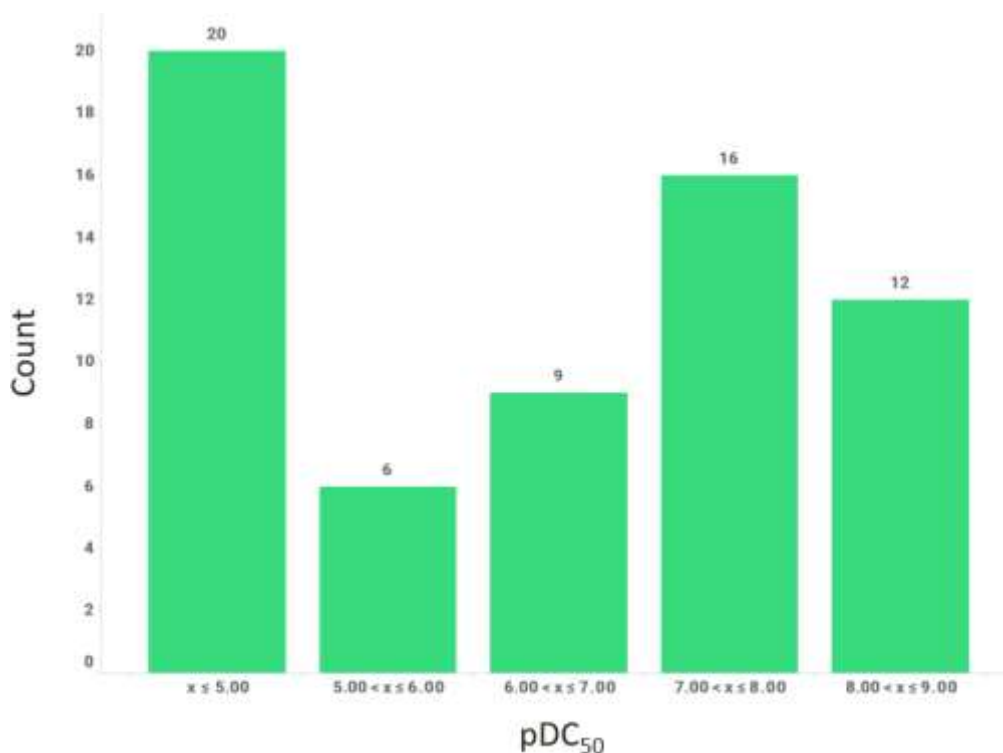
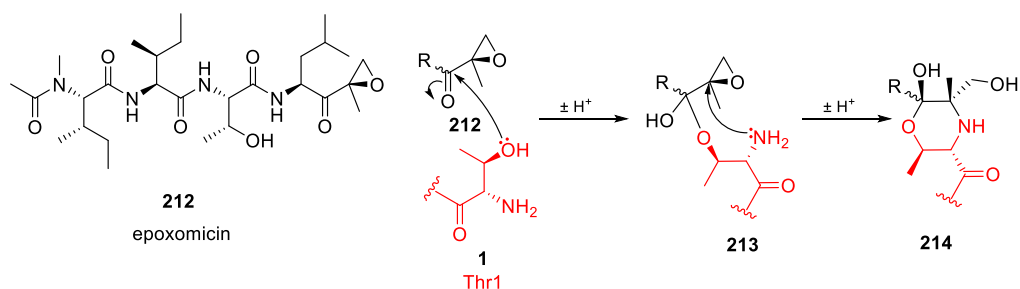


Figure 68: Histogram plot of the measured Brd4 pDC₅₀ values of the PROTAC series.

As Brd4 protein is an intracellular protein, the degradation data in **Figure 62** and **64** confirms that the PROTAC series using the pyridone benzimidazole protein binder **110** and VHL E3 ligase binder **30** is capable of eliciting potent degradation of Brd4 protein in a cellular setting. This reinforces that the PROTACs are exhibiting some level of cell permeability and aqueous solubility.

To confirm that the observed degradation was occurring *via* the UPS and the 26 S proteasome, a series of control experiments was performed using epoxomicin **212**. Epoxomicin **212** is an irreversible inhibitor of the 26 S proteasome that selectively forms multiple covalent bonds with the Thr1 residue of the $\beta 5$ subunits, resulting in the formation of a morpholine ring, **Scheme 74**.



Scheme 74: Mechanism of epoxomicin **212** inhibition of the 26 S proteasome.

The Brd4 HiBiT degradation assay was performed using the procedure described in **Section 5.5.4.1**, incubating the cells with and without epoxomicin for a shorter duration of 4 h. Protein binder **110** exhibited negligible degradation of Brd4 protein with and without epoxomicin, **Figure 69**. The error bars represent the standard deviation of the data points, which show a good correlation for protein binder **110** with and without epoxomicin. This suggests that any degradation does not occur *via* the UPS and 26 S proteasome.

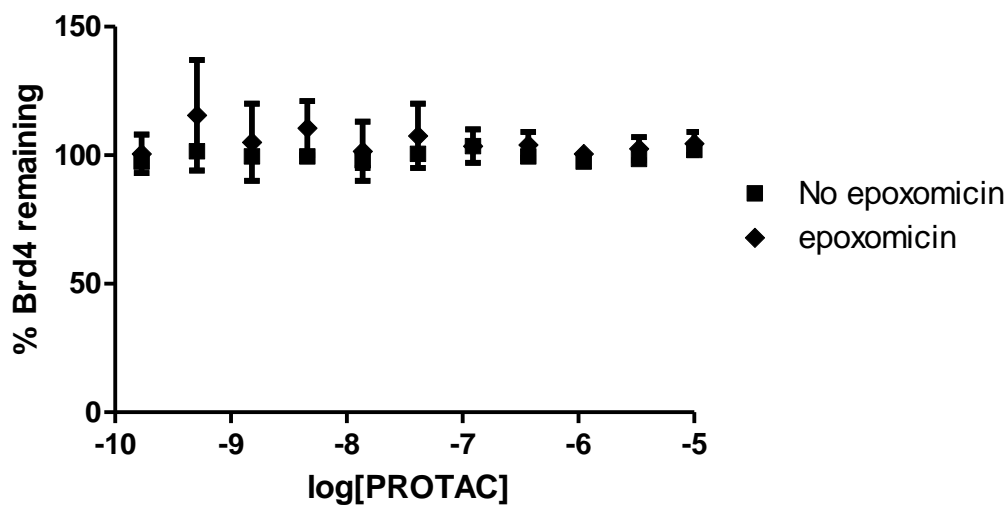


Figure 69: Plot of % Brd4 remaining vs log[PROTAC] for protein binder **110**, without epoxomicin (square markers) and with epoxomicin (diamond markers). Error bars represent the standard deviation of the data points.

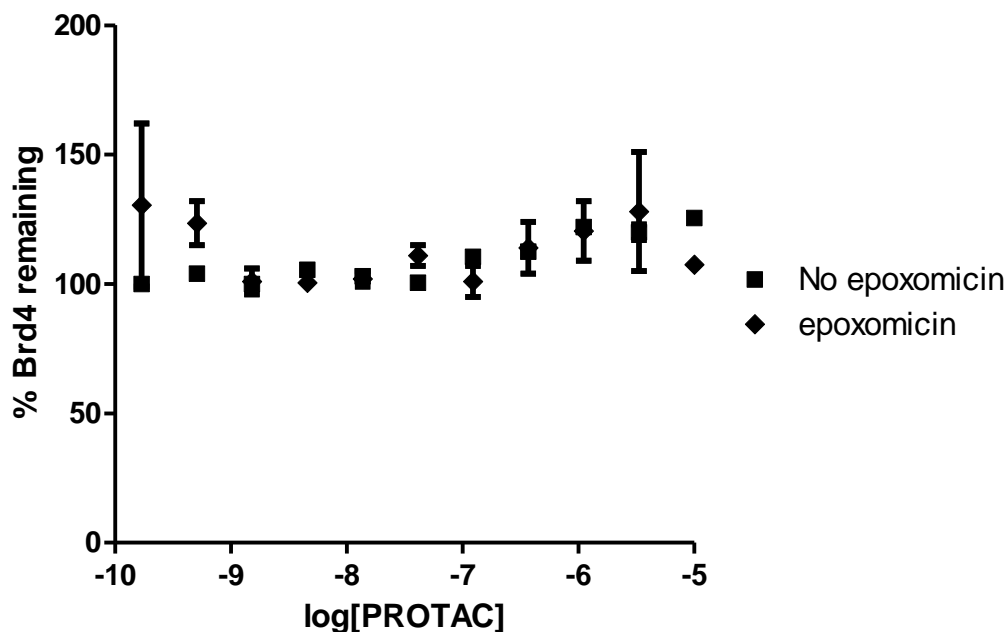


Figure 70: Plot of % Brd4 remaining vs log[PROTAC] for VHL E3 ligase binder **30**, without epoxomicin (square markers) and with epoxomicin (diamond markers). Error bars represent the standard deviation of the data points.

VHL E3 ligase binder **30** exhibited negligible degradation of Brd4 protein with and without epoxomicin, **Figure 70**. The error bars represent the standard deviation of the data points, which show a good correlation for VHL E3 ligase binder **30** with and without epoxomicin. This suggests that any degradation does not occur *via* the UPS and 26 S proteasome.

PROTAC **159b** exhibited potent degradation of Brd4 protein without epoxomicin, **Figure 71**. Achieving a max % deg of 94% and a pDC₅₀ of 8.95. However, negligible degradation was observed for PROTAC **159b** with the 26 S proteasome inhibitor epoxomicin **212**. This confirms that the Brd4 degradation elicited by PROTAC **159b** is occurring *via* the UPS and 26 S proteasome. Interestingly, this potent degradation is observed after 4 h, with minimal increase in max % deg and pDC₅₀ for the longer incubation time of 18 h. This may suggest that PROTAC **159b** is exhibiting high clearance from the cell, reducing its residence time and thus its potency.

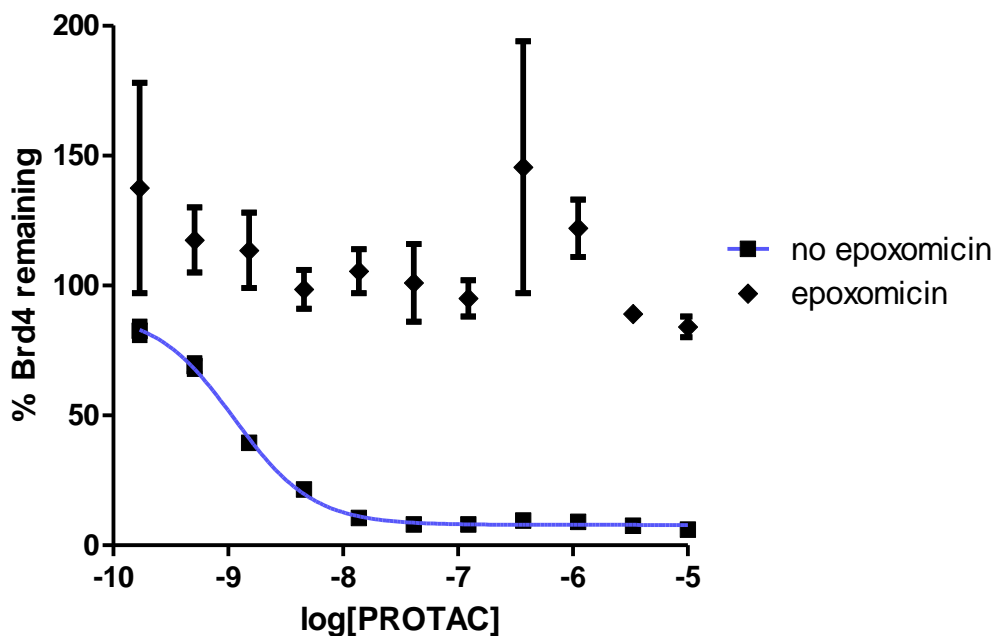


Figure 71: Plot of % Brd4 remaining vs log[PROTAC] for PROTAC 159b, without epoxomicin (square markers) and with epoxomicin (diamond markers). Error bars represent the standard deviation of the data points.

The effects of linker functionality on the physicochemical properties of the PROTAC series will be investigated further below. Brd4 binding affinity and degradation data will be used as functional endpoints to aid in the elucidation of these trends.

4.4 Physicochemical Profile of the PROTAC Series

The physicochemical properties of the PROTAC series were measured using the in-house biological assays described in **Section 5.5**.

The AMP, which is a measure of the passive permeability of molecules through a lipid bilayer, is visualised for each PROTAC of the series in **Figure 72**. The PROTACs are represented by green points and are shaded relative to their measured ChromlogD values, with dark green representing a ChromlogD of 5.55 and light green representing a ChromlogD of 2.95. The dashed black line represents a measured AMP of < 3 nm/s, which is assigned to PROTACs with an AMP too small to be measured. Compounds exhibiting an AMP > 200 nm/s are considered to be highly permeable, 200-10 nm/s

are considered moderately permeable and < 10 nm/s are considered poorly permeable.¹⁸⁶

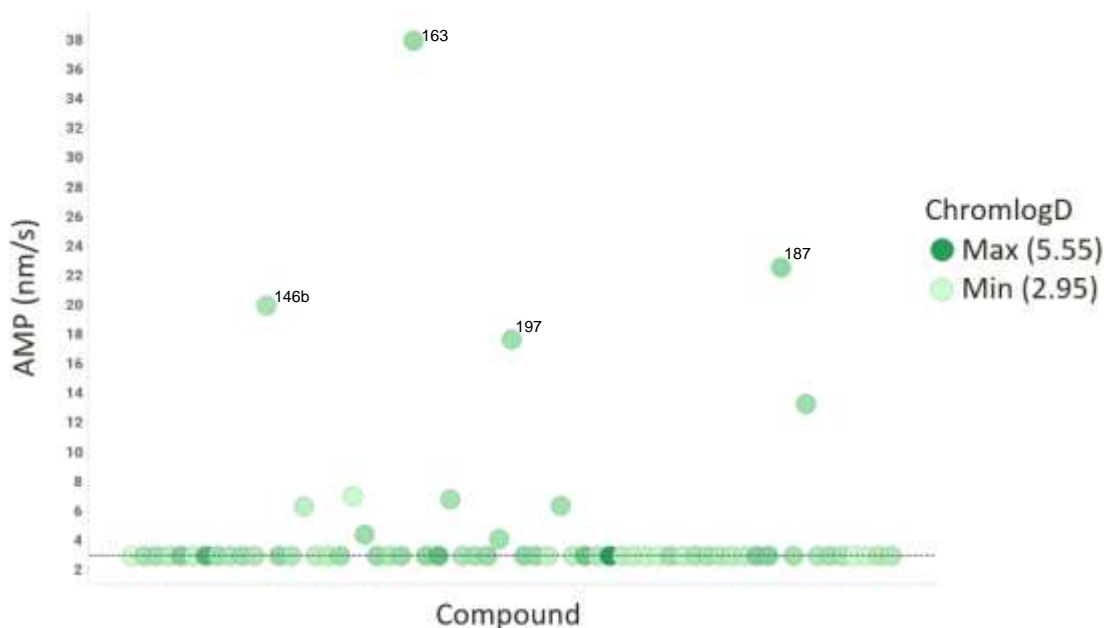


Figure 72: Plot of AMP for the PROTAC series shown in green, with shading corresponding to their measured ChromlogD values, with dark green representing a ChromlogD of 5.55 and light green representing a ChromlogD of 2.95.

The PROTACs visualised in **Figure 72** are shaded relative to their measured ChromlogD values. The expected positive correlation between lipophilicity (ChromlogD) and permeability is not observed for AMP. PROTAC **163** exhibited the highest AMP of 38 nm/s, suggesting it is moderately permeable. The majority of the PROTACs in the series exhibit an AMP of < 3 nm/s, suggesting that a negligible amount of PROTAC is passively permeating through the artificial lipid bilayer. However, more than 40% of the PROTAC series exhibited a Brd4 max % deg $> 75\%$, **Figure 63**. As Brd4 protein is an intracellular target, this suggests that the PROTACs are in fact exhibiting sufficient levels of permeability to achieve potent degradation of Brd4 protein. Furthermore, this suggests that the PROTACs are permeating into the cells *via* a mechanism other than passive permeability, which may involve transmembrane proteins. Therefore, a more accurate assay for determining the permeability and cellular concentration of the PROTAC series is required.

The cell concentration permeability (P Δ C) of the PROTAC series was measured in HeLa cells using the detergent rupture method described in **Section 5.5**. P Δ C was calculated by taking the log of the concentration of the test compound in cells divided by the control test compound concentration with no cells. A P Δ C value of < -1 is indicative of poor permeability, -1 to 0 limited permeability, 0 to 1 good permeability and > 1 is indicative of significant accumulation of the test compound in cells. The P Δ C values of the PROTAC series are visualised in **Figure 73**. A P Δ C value was not obtained for PROTAC **170** due to an insufficient amount of solid sample required for testing. A 3 log unit range of P Δ C values is observed, which suggests that the linker functionality has a significant effect on the permeability of the PROTACs. Approximately 75% of the PROTAC series exhibits a P Δ C \geq 0, which suggests they are exhibiting good permeability. There is a clear positive correlation between ChromlogD, which is a measure of lipophilicity, and P Δ C, with PROTAC **176** exhibiting the highest ChromlogD value of 5.55 and the highest P Δ C of 2.08. This suggests that lipophilicity is a key factor driving the cell permeability of PROTACs. Furthermore, there is no correlation between AMP and P Δ C. PROTAC **163** exhibits the highest AMP of 38 nm/s and one of the highest P Δ C values of 1.70. However, PROTAC **192** exhibits a higher P Δ C of 1.87 in spite of being assigned an AMP of < 3 nm/s. Additionally, both PROTACs exhibit a ChromlogD of 3.83 and 3.84 respectively. This further suggests that passive permeability is not the main contributor to PROTAC cell permeability. It should also be noted that although PROTACs that exhibit higher lipophilicity appear to be more permeable, it is likely that they will exhibit a higher affinity to other proteins in the cell. As only the unbound fraction of PROTAC is capable of facilitating Brd4 engagement, ubiquitination and degradation, this may result in a lower degradation potency. However, there is no observable correlation between higher P Δ C values and higher levels of HSA or IAM binding.

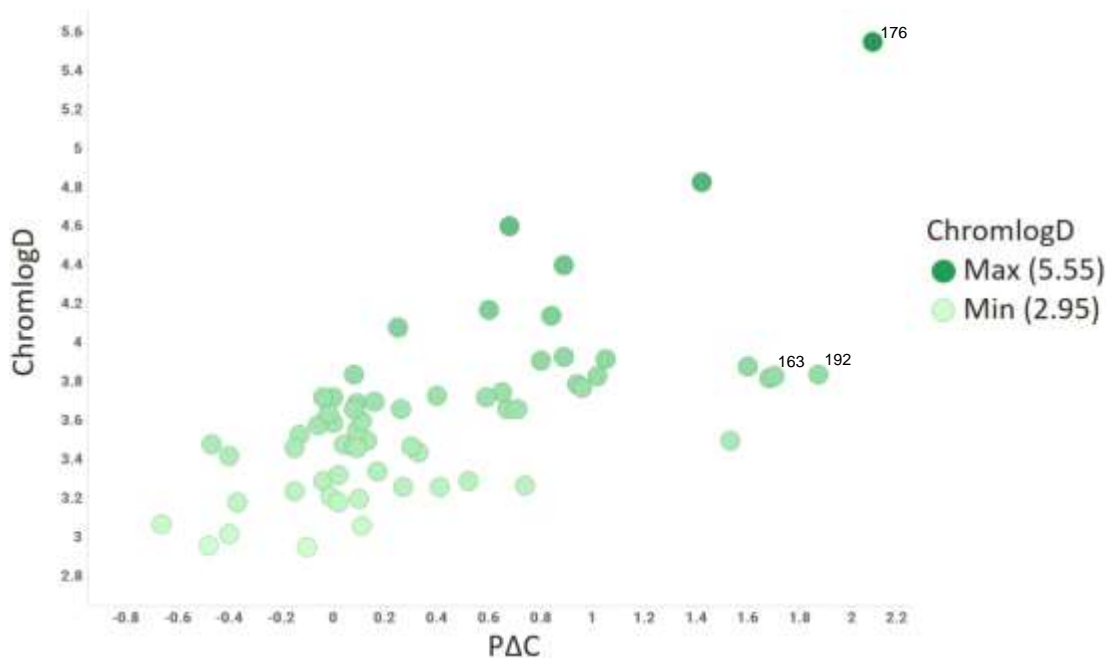


Figure 73: Plot of ChromlogD vs PAC for the PROTAC series shown in green, with shading corresponding to their measured ChromlogD values, with dark green representing a ChromlogD of 5.55 and light green representing a ChromlogD of 2.95.

The kinetic (CAD) and thermodynamic (FaSSIF) solubilities of the PROTACs in the series are visualised in **Figure 74**. FaSSIF solubility data were not obtained for PROTACs **152b**, **153b**, **168**, **169** and **184** due to an insufficient amount of free-flowing solid sample required for testing. The PROTACs are represented by green points and are shaded relative to their measured ChromlogD values. The solid black line represents equivalence between FaSSIF and CAD solubility measurements and the dashed black line represents the maximum measurable FaSSIF solubility of 1000 $\mu\text{g/mL}$. Compounds exhibiting a solubility measurement $> 200 \mu\text{g/mL}$ are considered to exhibit high solubility, 200-30 $\mu\text{g/mL}$ exhibit intermediate solubility and $< 30 \mu\text{g/mL}$ exhibit poor solubility. The majority of the PROTAC series exhibits a good correlation between FaSSIF and CAD solubility measurements within the error of the assays. Forty-four PROTACs from the series exhibit intermediate to high solubility for both kinetic and thermodynamic solubility measurements.

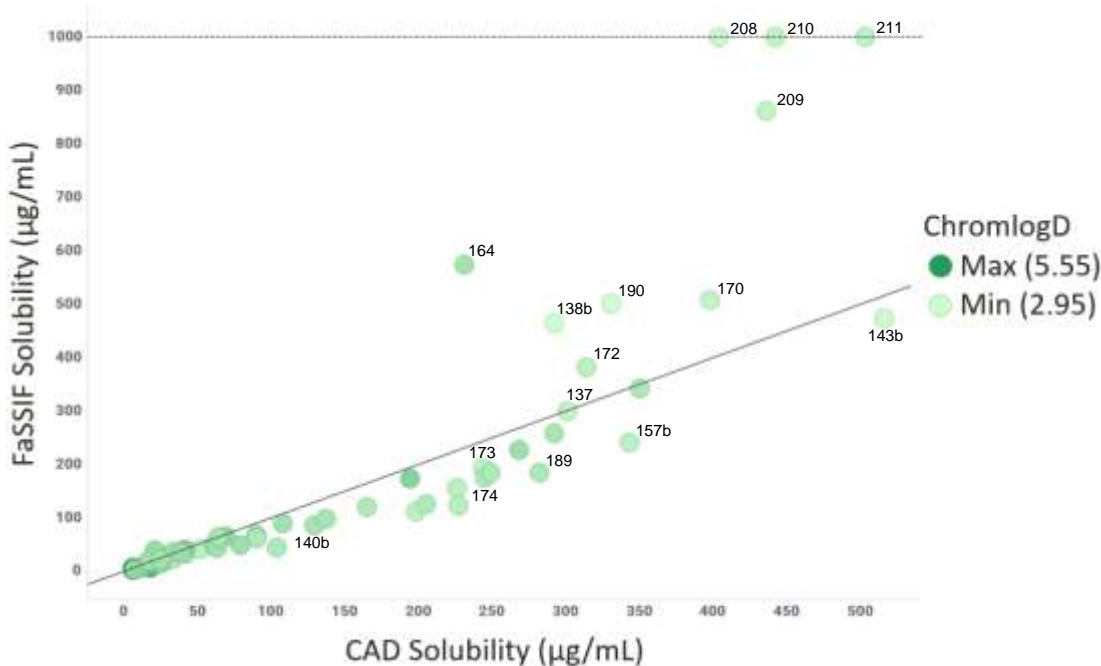


Figure 74: Plot of FaSSIF solubility vs CAD solubility for the PROTAC series shown in green, with shading corresponding to their measured ChromlogD values, with dark green representing a ChromlogD of 5.55 and light green representing a ChromlogD of 2.95.

The PROTACs **208**, **209**, **210** and **211** containing PEG linkers exhibited the highest FaSSIF solubility, with PROTACs **208**, **210** and **211** achieving the maximum measurable solubility of 1000 µg/mL. PEG moieties are commonly used to improve the solubility of compounds due to their hydrophilic properties and ability to form multiple hydrogen bonds with water.²⁰⁸ As the PROTACs were not isolated by recrystallisation, FaSSIF solubility measurements were performed on amorphous solid samples. As the most stable crystalline form of the compounds were not isolated, the FaSSIF solubility measurements may overestimate the compounds' thermodynamic solubility, accounting for the large discrepancy between FaSSIF and CAD solubility measurements for these PROTACs and PROTAC **164**.

PROTAC **143b** exhibited the highest CAD solubility of 516 µg/mL. This PROTAC also contains an ether linker and benefits from its aforementioned hydrophilic properties. Substituting the oxygen for a sulfur in PROTAC **172** resulted in a significant decrease in both CAD and FaSSIF solubility measurements. This may be due to the more favourable formation of hydrogen bonds between the harder and more

electronegative ether oxygen than the larger and softer sulfide sulfur.²⁰⁹ This would decrease the molecule's solvation energy, increasing its solubility. A further decrease in solubility for PROTAC **173** containing a sulfone linker, which is an even weaker hydrogen bond acceptor was observed.²¹⁰ These differences are not described by the HBA count, which is only contributed to by nitrogen and oxygen atoms that behave as HBA, resulting in a count of 11, 10 and 12 for the PROTACs respectively.

Additionally, substitution of a carbon atom of the phenyl ring in PROTAC **189** for a nitrogen in PROTACs **157b** and **190** furnishes an increase in solubility. This is a result of an increase in the TPSA and HBA count, conferring favourable hydrogen-bonding interactions between the solvent and the pyridyl nitrogen than can act as a HBA. The position of the pyridyl ring nitrogen also has a subtle effect on the solubility of the PROTACs.

PROTAC **138b**, with an unfunctionalised 2-carbon linker, exhibited a FaSSIF solubility of 464 µg/mL and a CAD solubility of 292 µg/mL. A higher solubility was observed for PROTAC **170**, containing an additional gem-disubstituted alkene on the 2-carbon linker, although a concurrent increase in ChromlogD from 2.92 to 3.2 was observed. This may be due to a reduced ability of the molecules to pack and form favourable intra- and intermolecular bonding interactions in the solid form, reducing its lattice energy and thus increasing its solubility.

Increasing the length of the all-carbon linker from 2-carbons in PROTAC **138b** to 4-carbons PROTAC **137**, 5-carbons PROTAC **174** and 6-carbons PROTAC **140b** is accompanied by a subsequent increase in ChromlogD and decrease in both kinetic and thermodynamic solubility. This general trend is observed for the entire PROTAC series and is highlighted by the shading of the green points in **Figure 74**.

An exponential increase in solubility is observed for PROTACs with decreasing ChromlogD measurements, **Figure 75**. A similar trend is also observed for the correlation between PAC and solubility.

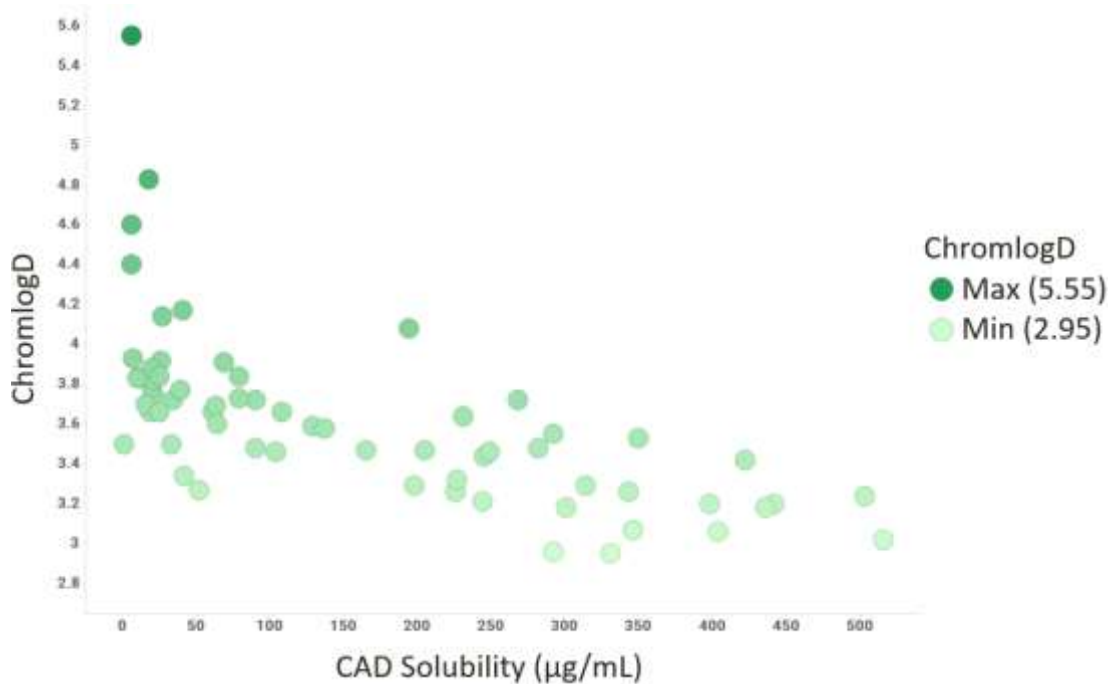


Figure 75: Plot of ChromlogD vs CAD solubility for the PROTAC series shown in green, with shading corresponding to their measured ChromlogD values, with dark green representing a ChromlogD of 5.55 and light green representing a ChromlogD of 2.95.

No clear trend is observed between TPSA and either CAD or FaSSIF solubility, highlighted by the plot of TPSA vs CAD solubility in **Figure 76**. Forty PROTACs from the series have a TPSA of 199 \AA^2 . However, these compounds contain a range of carbon lengths, substitutions and configurations that have no effect on TPSA, but a significant effect on the PROTACs' lattice and solvation energy and thus their solubility. Therefore, a range of CAD solubility values of 6-422 $\mu\text{g/mL}$ is observed for the PROTACs with a TPSA of 199 \AA^2 . However, substitution of a carbon atom of the phenyl ring in PROTAC **189** for a nitrogen in PROTACs **157b** and **190** furnishes an increase in both TPSA and CAD solubility. A similar trend is observed for PROTACs **208**, **209**, **210** and **211** when increasing the number of EG units in the linker. This suggests that the expected increase in TPSA and CAD solubility is observed to some extent, however, other factors are having a significant effect on the solubility. Therefore, it may be favourable to use a 3D PSA descriptor that may account for some of these factors.¹⁹⁷ However, this is beyond the scope of this investigation.

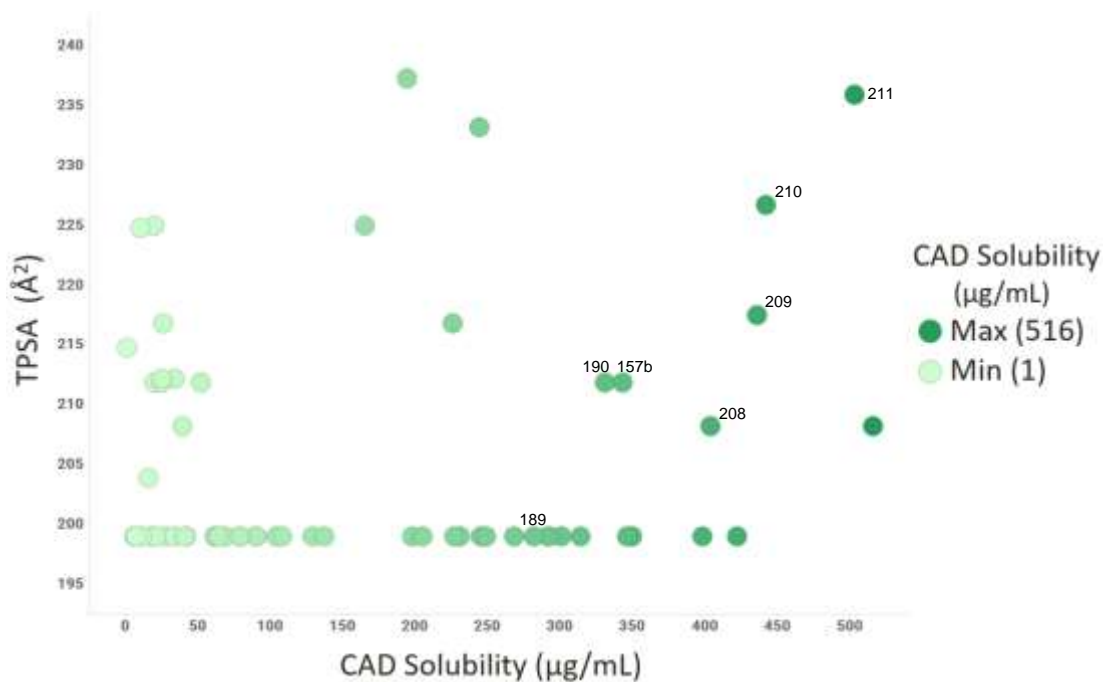


Figure 76: Plot of TPSA vs CAD solubility for the PROTAC series shown in green, with shading corresponding to their measured CAD solubility values, with dark green representing a CAD solubility of 516 µg/mL and light green representing a CAD solubility of 1 µg/mL.

A weak negative correlation between M_w and CAD solubility is observed for the PROTAC series, **Figure 77**. Only PROTACs **209**, **210** and **211**, with PEG linkers and PROTAC **182**, with a Boc-protected primary amine on the linker, show significant deviation from this trend due to the presence of HBAs and HBDs that furnish an improved solubility in spite of the higher M_w .

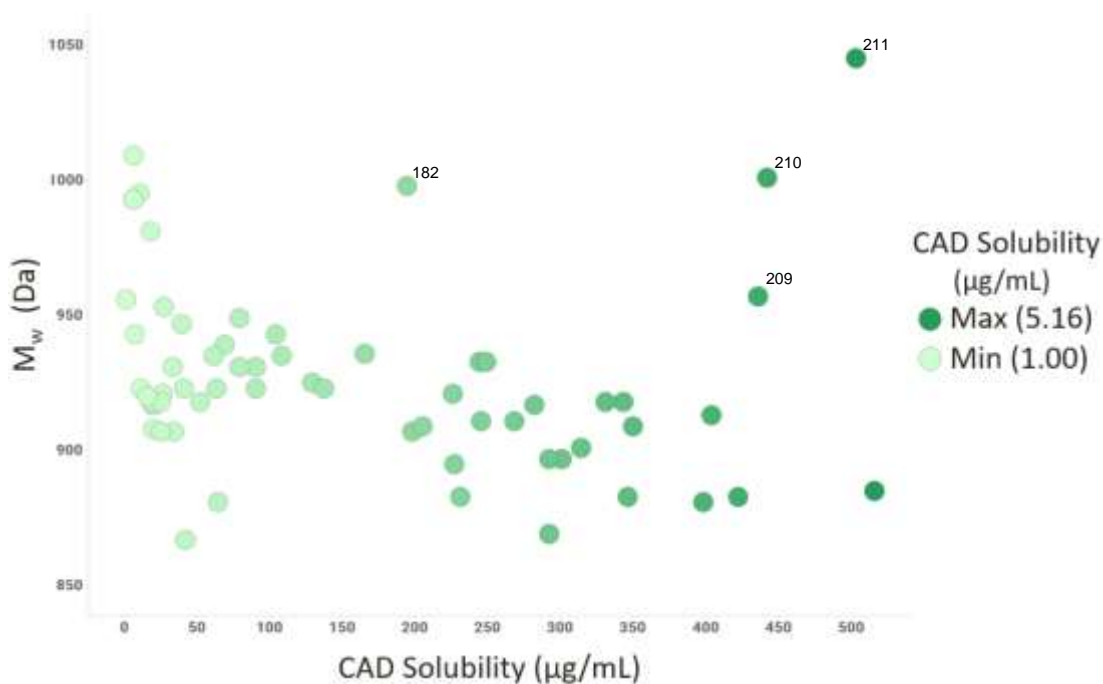


Figure 77: Plot of M_w vs CAD solubility for the PROTAC series shown in green, with shading corresponding to their measured CAD solubility values, with dark green representing a CAD solubility of 5.16 $\mu\text{g/mL}$ and light green representing a CAD solubility of 1 $\mu\text{g/mL}$.

A weak positive correlation between ChromlogD and M_w is observed for the PROTAC series, **Figure 78**. Only PROTACs **209**, **210** and **211**, with PEG linkers, show significant deviation from this trend due to the hydrophilic properties of the EG units that results in a constant ChromlogD in spite of the higher M_w .

Focusing on a series of unfunctionalised all-carbon linkers, PROTACs **137**, **174**, **140b**, **175**, **141b**, **142b** and **176**. A positive linear correlation between ChromlogD and M_w is observed as the length of the carbon chain increases from 4 to 12 carbons. Substitution of a carbon for an oxygen atom in PROTAC **174** furnishes PROTAC **208**, resulting in a minimal increase in M_w and a concurrent decrease in ChromlogD from 3.44 to 3.06.

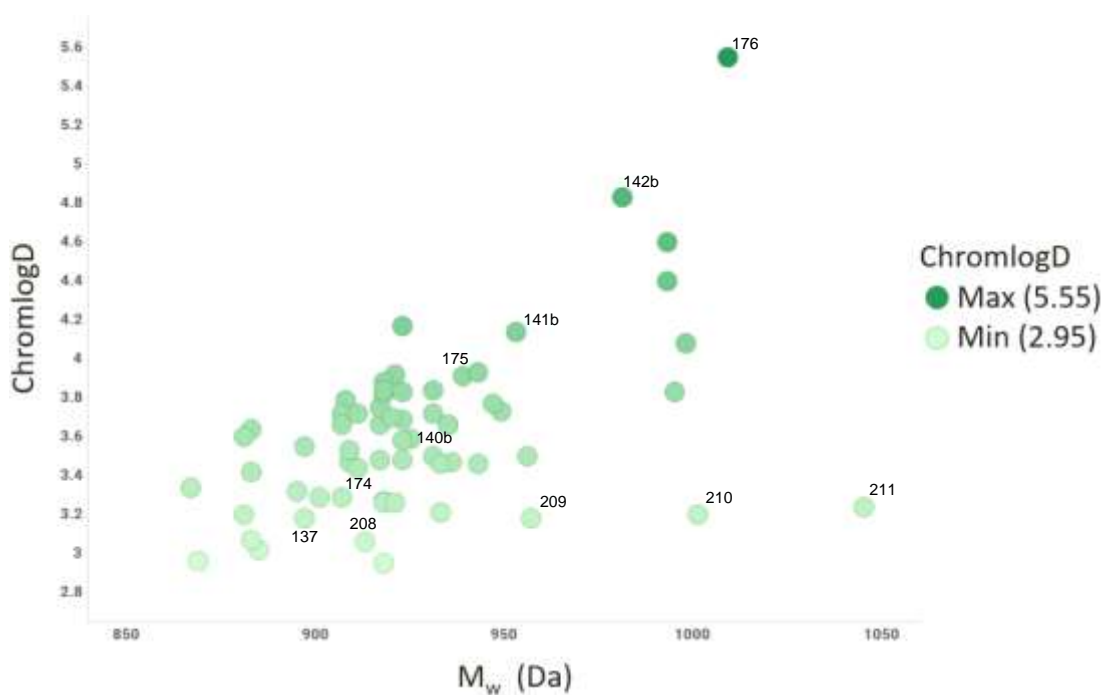


Figure 78: Plot of ChromlogD vs M_w for the PROTAC series shown in green, with shading corresponding to their measured ChromlogD values, with dark green representing a ChromlogD of 5.55 and light green representing a ChromlogD of 2.95.

No clear correlation between ChromlogD and RBC is observed when analysing the entire PROTAC series, **Figure 79**. However, as previously stated, PROTACs **209**, **210** and **211**, with PEG linkers, exhibit similar ChromlogD values in spite of an increase in M_w and RBC due to the hydrophilic properties of the additional EG units.

Focusing on a series of unfunctionalised all-carbon linkers, PROTACs **137**, **174**, **140b**, **175**, **141b**, **142b** and **176**, a positive linear correlation between ChromlogD and RBC is observed as the length of the carbon chain increases from 4 to 12 carbons.

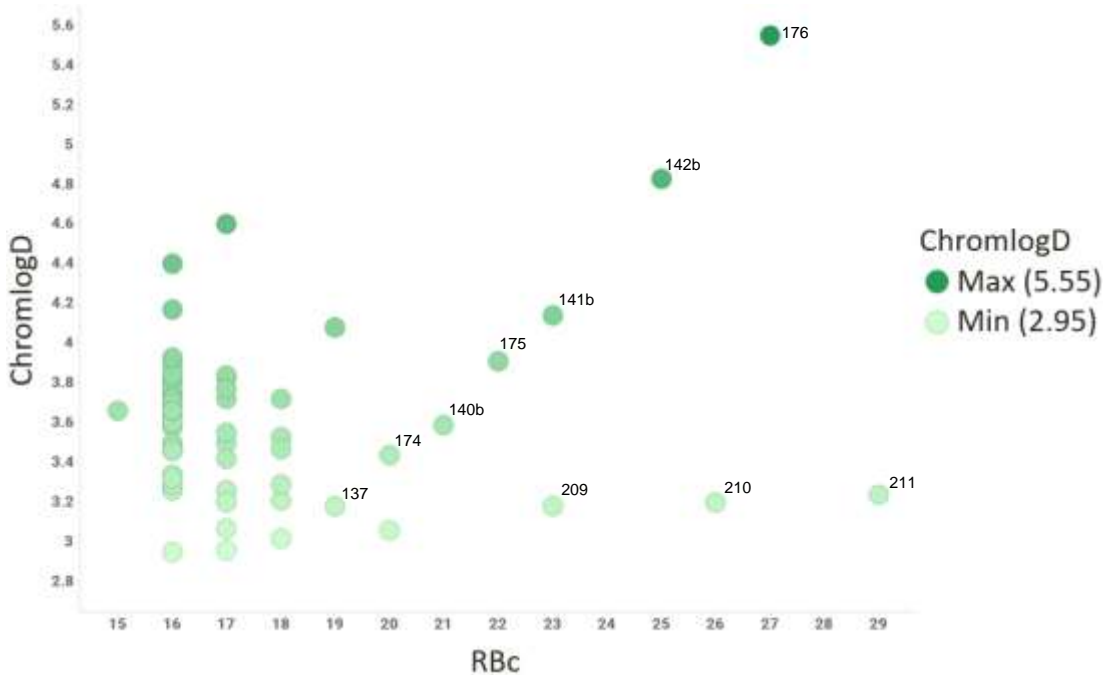


Figure 79: Plot of ChromlogD vs RBC for the PROTAC series shown in green, with shading corresponding to their measured ChromlogD values, with dark green representing a ChromlogD of 5.55 and light green representing a ChromlogD of 2.95.

A weak positive correlation between ChromlogD and HSA binding is observed for the PROTAC series, **Figure 80**. PROTAC **176** has the highest measured ChromlogD of 5.55, furnishing the highest HSA binding of 98%. PROTAC **153b**, containing an indole linker which has a pK_a of 21.0 for the ring NH in DMSO, exhibits a HSA binding of 93%, despite having a ChromlogD of only 3.50.²¹¹ This may suggest that the indole NH is capable of making additional binding interactions with the exposed basic residues of HSA, increasing the HSA binding of PROTAC **153b**. However, as HSA binding also correlates with ARc, the indole ring may be making a significant contribution.

A range of HSA binding of 73-98% is observed for the PROTAC series. Decreasing the amount of HSA binding is favourable as only unbound molecules are capable of entering the target tissue and eliciting the desired therapeutic effect. However, HSA binding enables the transportation of the molecule to the target tissue and protects the compounds from oxidation, metabolism and excretion. The optimum amount of HSA

binding achieves a desirable balance of these properties and is both target and substrate specific. Therefore, Brd4 HWB potency and degradation will need to be considered.

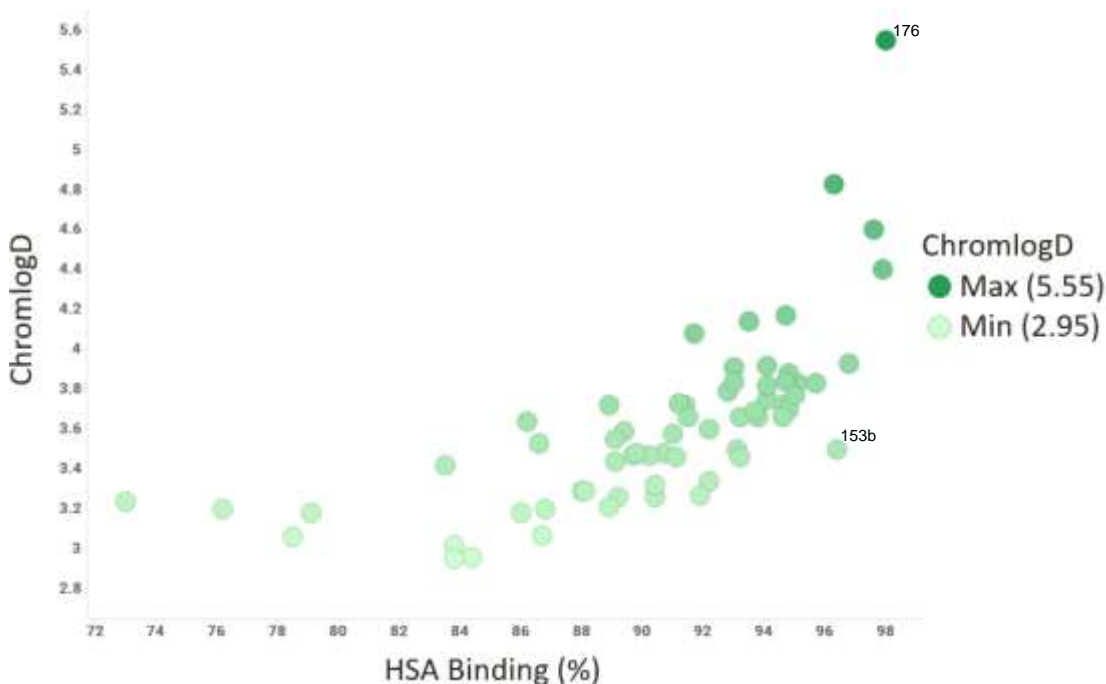


Figure 80: Plot of ChromlogD vs HSA binding for the PROTAC series shown in green. The shading of the points corresponds to their measured ChromlogD values, with dark green representing a ChromlogD of 5.55 and light green representing a ChromlogD of 2.95.

A weak positive correlation between ChromlogD and CHI_{IAM} is observed for the PROTAC series, with a range of CHI_{IAM} values between 23.0 and 42.0 for PROTACs **211** and **142b** respectively, **Figure 81**. A CHI_{IAM} value > 50 indicates strong phospholipid binding associated with high levels of promiscuity that may incur unwanted toxicity and side-effects. A CHI_{IAM} value < 10 indicates weak binding and is characteristic of molecules that exhibit poor permeability.²⁰⁰ PROTAC **185**, containing the 1,3-substituted biaryl linker, exhibited a CHI_{IAM} value of 37.0 and a ChromlogD of 4.60. Introducing heteroatoms into aromatic rings is a commonly employed strategy to reduce lipophilicity.²¹² Replacing the two aryl rings with pyridyl rings, furnished PROTAC **163**, which exhibits a lower ChromlogD of 3.83. However, PROTAC **163** also exhibits a higher CHI_{IAM} value of 37.3. The net negative charge of the phosphate head groups in phosphatidylcholine results in a preference for binding

molecules containing basic moieties, such as the pyridine ring nitrogen, accounting for the increase in CHI_{IAM} value.

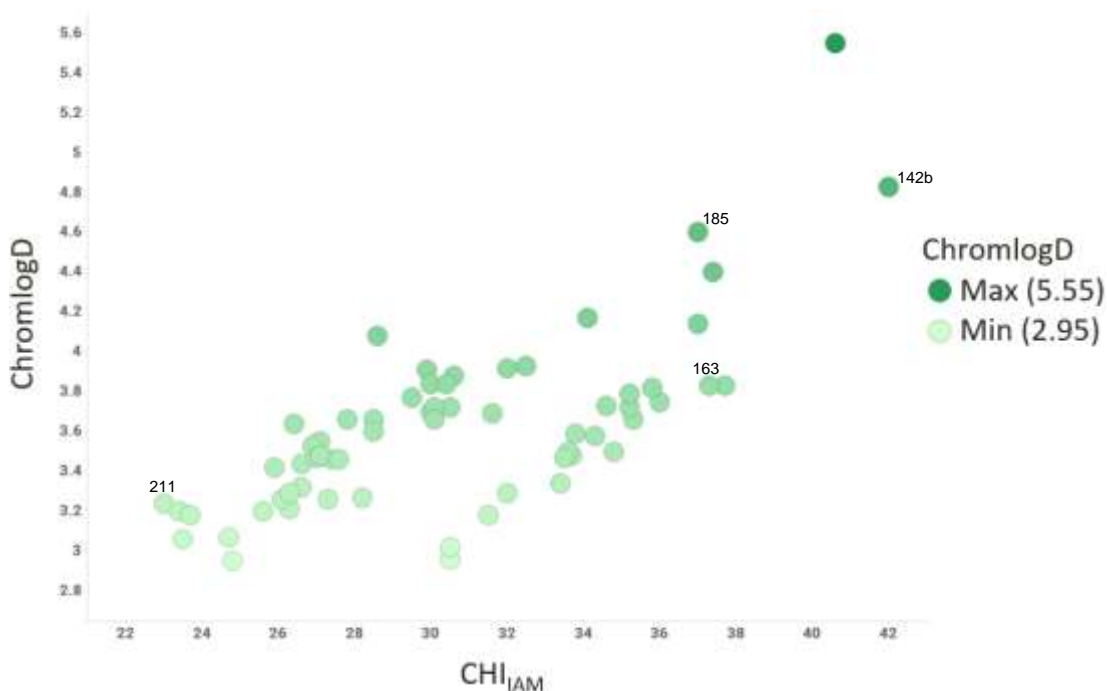


Figure 81: Plot of ChromlogD vs CHI_{IAM} for the PROTAC series shown in green. The shading of the points corresponds to their measured ChromlogD values, with dark green representing a ChromlogD of 5.55 and light green representing a ChromlogD of 2.95.

The correlation between the trends observed for the physicochemical properties above with Brd4 binding affinity and degradation was investigated further.

4.5 Binding Affinity and Degradation Profile of the PROTAC Series

No correlation between Brd4 pDC_{50} and FRET pIC_{50} is observed when considering the entire PROTAC series, **Figure 82**. The lowest pDC_{50} value measurable in the Brd4 degradation assay is 5; any PROTACs that do not achieve 50% degradation are assigned a pDC_{50} of < 5 , represented by the dashed black line. A weak positive correlation can be observed by omitting PROTACs exhibiting a $pDC_{50} < 5$. This suggests that the PROTACs' binding affinity is not driving the 4 log unit range of pDC_{50} values observed for the series.

The majority of the PROTACs Brd4 binding affinity emanates from the protein binder **110**, which was maintained for all PROTACs in the series. Any additional binding affinity was achieved through the additional interactions between the linker or VHL E3 ligase binder **30** that were previously described. Therefore, a small range of FRET pIC_{50} values between 7.32-7.98 is observed.

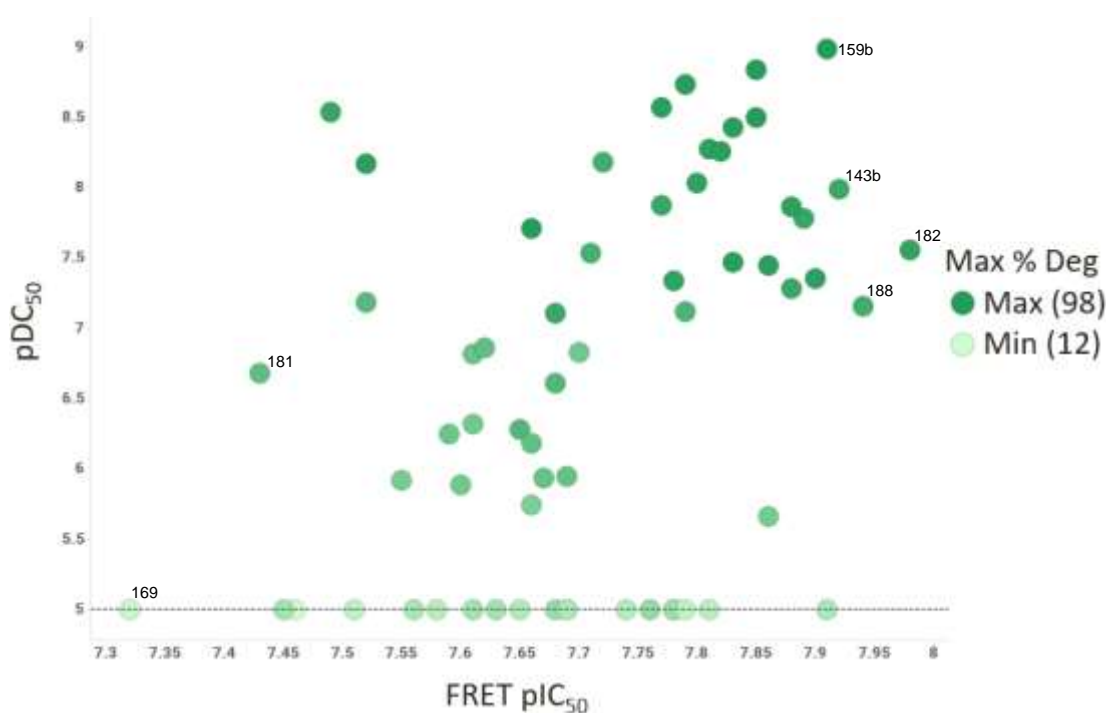


Figure 82: Plot of pDC_{50} vs FRET pIC_{50} for the PROTAC series shown in green. The shading of the points corresponds to their max % deg values, with dark green representing a max % deg of 98% and light green representing a max % deg of 12%.

Sufficient Brd4 binding affinity is required to achieve adequate Brd4 engagement to enable the transfer of ubiquitin to a surface-exposed lysine residue on the Brd4 protein. PROTAC **181** exhibits a pDC_{50} value of 6.68, a max % deg of 69% and a FRET pIC_{50} of 7.43. This confirms that PROTACs with a FRET $pIC_{50} \geq 7.43$ are capable of eliciting sufficient Brd4 engagement for adequate ubiquitin transfer to furnish degradation of the Brd4 protein *via* the UPS. However, PROTAC **169** exhibited a max % deg of only 16%. This may be due to insufficient productive ternary complex formation for ubiquitin transfer and not a lack of Brd4 engagement. Therefore, the actual threshold FRET pIC_{50} required for sufficient Brd4 engagement may be much

lower than this value. Additional data points exploring FRET pIC₅₀ values < 7.3 would be required to more accurately probe this threshold.

PROTAC **159b** exhibited the highest pDC₅₀ of 8.99 with a FRET pIC₅₀ of 7.91. The three PROTACs **143b**, **188** and **182** that exhibit higher FRET pIC₅₀ values than PROTAC **159b** all exhibit a significantly lower pDC₅₀. This may suggest that there is an upper limit of binding affinity, beyond which additional Brd4 engagement after sufficient ubiquitination actually limits the catalytic effects of the PROTAC. However, further investigations and data points at higher FRET pIC₅₀ values are required to confidently identify this limit.

The relationship between Brd4 pDC₅₀ and HWB pIC₅₀ is visualised for the PROTAC series in **Figure 83**. The solid black line represents equivalence between Brd4 pDC₅₀ and HWB pIC₅₀. PROTACs above the line exhibit a more potent degradation than inhibition in the assays described in **Section 5.5**, suggesting that they exhibit favourable characteristics for productive ternary complex formation and Brd4 degradation. PROTACs below the line exhibit more potent inhibition than degradation.

A positive correlation between Brd4 pDC₅₀ and HWB pIC₅₀ can be observed by omitting PROTACs exhibiting a pDC₅₀ < 5. This correlation is somewhat expected as both assays use similar concentrations of test compound with 18 h incubation. Additionally, as both measurements are taken from *in vitro* cell-based assays they are affected by similar factors such as solubility and permeability. Furthermore, all eukaryotic cells have a UPS and therefore PROTAC-mediated Brd4 degradation may contribute to the measured HWB pIC₅₀.²¹³

The points in **Figure 83** are shaded relative to their max % deg values, with dark green representing a max % deg of 98% and light green representing a max % deg of 12%. PROTAC **168** exhibited a HWB pIC₅₀ of 7.08 and was assigned a pDC₅₀ < 5 after only achieving a max % deg of 36%. PROTAC **186** exhibited a similar HWB pIC₅₀ of 7.06, but a significantly higher pDC₅₀ of 8.17 and max % deg of 95%. Additionally, both PROTACs **168** and **186** exhibited similar FRET pIC₅₀ values of 7.63 and 7.52 respectively. Similar trends can be identified for PROTACs **199** and **198** that exhibit

almost a 2 log unit difference in HWB pIC_{50} , with similar degradation profiles and FRET pIC_{50} values of 7.65 and 7.68 respectively. This suggests that if Brd4 degradation is contributing to the measured HWB pIC_{50} , it is not the determining factor. Therefore, HWB pIC_{50} may be a function of FRET pIC_{50} and permeability.

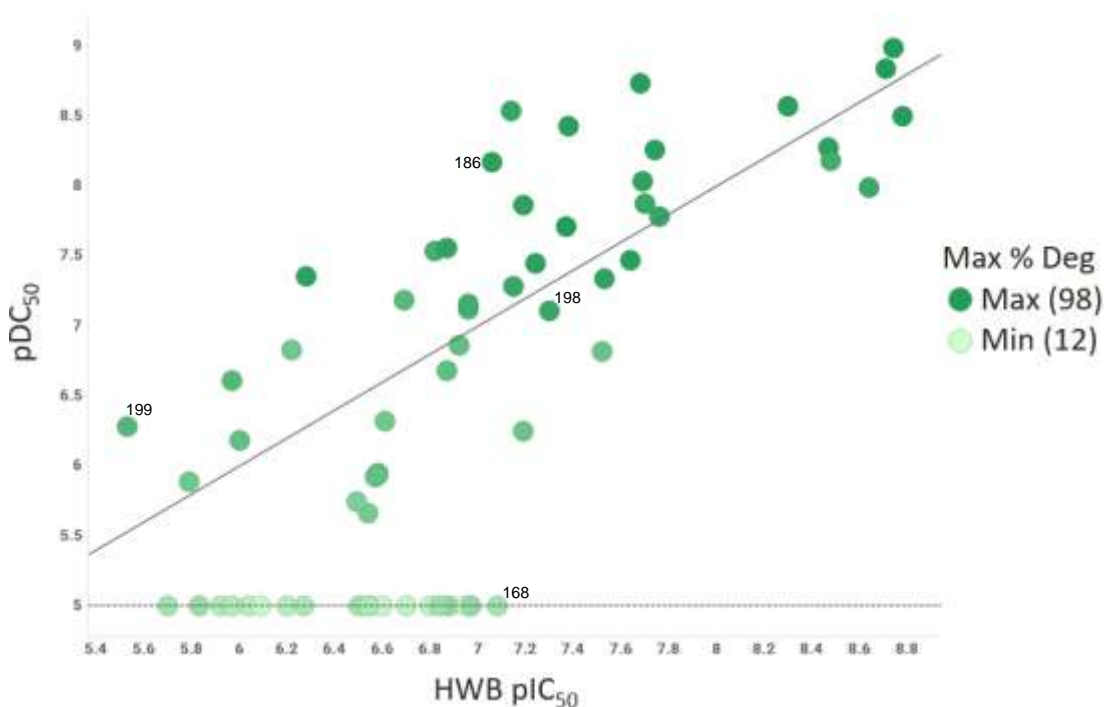


Figure 83: Plot of pDC_{50} vs HWB pIC_{50} for the PROTAC series shown in green. The shading of the points corresponds to their max % deg values, with dark green representing a max % deg of 98% and light green representing a max % deg of 12%.

No correlation is observed between AMP and HWB pIC_{50} for the PROTAC series, **Figure 84**. This is highlighted by comparing PROTAC **163**, which exhibited a HWB pIC_{50} of 8.74 and the highest AMP of 38 nm/s, with PROTAC **159b**, which exhibited a similar HWB pIC_{50} of 8.71 and an AMP < 3 nm/s. Furthermore, PROTAC **164** exhibited the second lowest HWB pIC_{50} of 5.70, as well as exhibiting an AMP of 7 nm/s. The points in **Figure 84** are shaded relative to the PROTACs' measured pDC_{50} values, with dark green representing a pDC_{50} of 8.99 and light green representing a pDC_{50} of < 5. PROTACs **159b** and **163** exhibit the highest pDC_{50} values of 8.99 and 8.84 respectively. They also exhibit an AMP of < 3 nm/s and 38 nm/s respectively. Therefore, there is no correlation observed between AMP and pDC_{50} .

As Brd4 protein is an intracellular target, this suggests that the PROTACs are in fact exhibiting sufficient levels of permeability to achieve potent degradation of Brd4 protein. This suggests that the PROTACs are permeating into the cells *via* a mechanism other than passive permeability, which may involve transmembrane proteins.

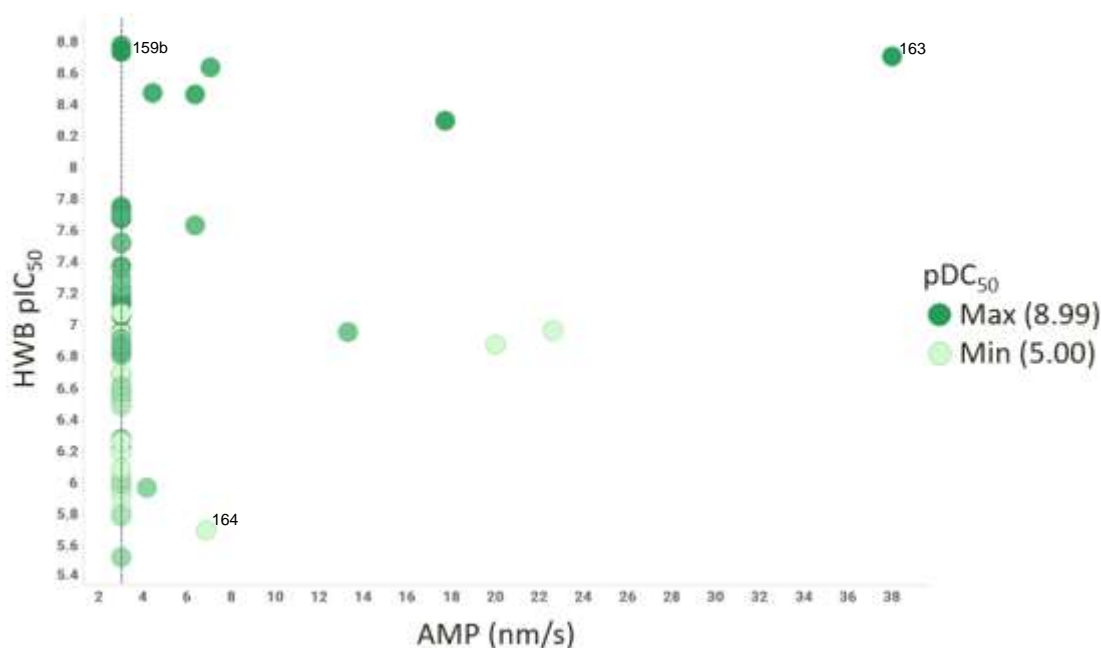


Figure 84: Plot of HWB pIC₅₀ vs AMP for the PROTAC series shown in green, with shading corresponding to their measured pDC₅₀ values, with dark green representing a pDC₅₀ of 8.99 and light green representing a pDC₅₀ of < 5.00.

The relationship between max % deg and HWB pIC₅₀ is visualised for the PROTAC series in **Figure 85**. All PROTACs exhibiting a HWB pIC₅₀ ≥ 7.52 also exhibit a max % deg ≥ 80%. If HWB pIC₅₀ is a function of permeability, this may suggest that PROTACs exhibiting a HWB pIC₅₀ ≥ 7.52 are sufficiently permeable to elicit potent degradation of Brd4 protein. However, the ability of the linker to adopt a productive ternary complex for ubiquitination and degradation of the Brd4 protein also needs to be considered. This is highlighted by PROTACs **150b** and **137**, which exhibit similar max % deg of 59% and 64%, with significantly different HWB pIC₅₀ values of 6.22 and 7.52 respectively. This may suggest that PROTAC **150b** is considerably less permeable but better able to adopt a productive ternary complex that facilitates a similar Brd4 degradation profile in spite of the lower cellular concentration. This is

reflected by the similar pDC_{50} values of 6.83 and 6.82 for PROTACs **150b** and **137** respectively.

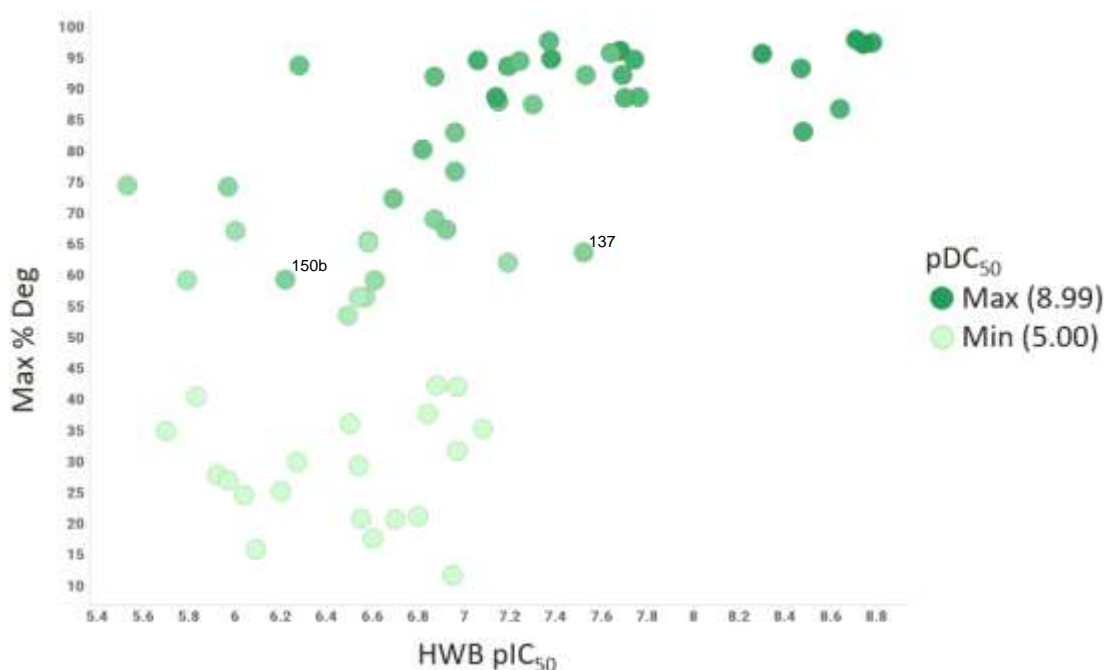


Figure 85: Plot of max % deg vs HWB pIC_{50} for the PROTAC series shown in green. The shading of the points corresponds to their pDC_{50} values, with dark green representing a pDC_{50} of 8.99 and light green representing a pDC_{50} of < 5.00.

It may therefore be possible to use HWB pIC_{50} as a surrogate for permeability by comparing it to the FRET pIC_{50} . The correlation between these measurements is visualised for the PROTAC series in **Figure 86**. The solid black line represents equivalence between HWB and FRET pIC_{50} measurements, with the thin red dashed lines representing ± 0.5 log units and the thick red dashed lines representing ± 1 log unit. PROTACs that exhibit a smaller decrease in HWB pIC_{50} relative to their FRET pIC_{50} are likely to be more permeable.

Although there is no clear trend between HWB and FRET pIC_{50} measurements when considering the entire PROTAC series, PROTACs exhibiting the highest pDC_{50} values, indicated by the shading of the green points, generally exhibit a higher HWB and FRET pIC_{50} . Seven PROTACs from the series exhibit a HWB pIC_{50} significantly higher than their FRET pIC_{50} . These PROTACs also exhibit potent pDC_{50} values > 8 . This suggests that the HWB pIC_{50} for these PROTACs is not just a function of FRET

pIC₅₀ and permeability and that degradation may be making a significant contribution. This caveat must be considered when using the difference between FRET and HWB pIC₅₀ as an approximation for permeability.

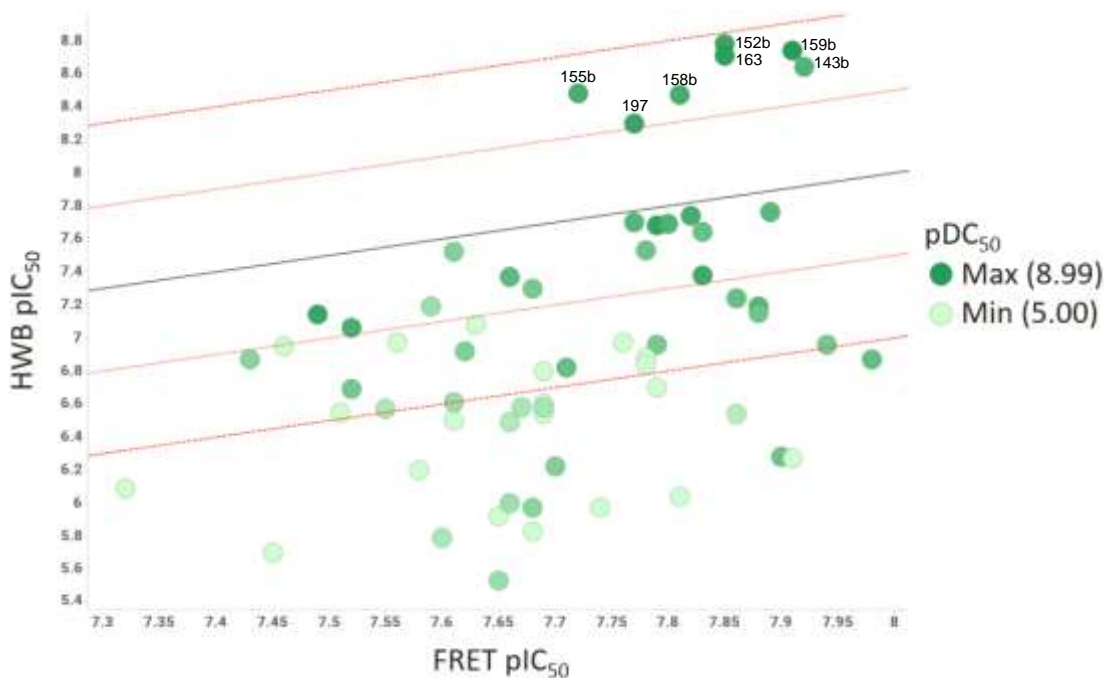


Figure 86: Plot of HWB pIC₅₀ vs FRET pIC₅₀ for the PROTAC series shown in green. The shading of the points corresponds to their pDC₅₀ values, with dark green representing a pDC₅₀ of 8.99 and light green representing a pDC₅₀ of < 5.00.

A negative correlation between pDC₅₀ and FRET - HWB pIC₅₀ is observed for the PROTAC series when omitting PROTACs exhibiting a pDC₅₀ < 5, **Figure 87**. This suggests that permeability and thus intracellular concentration is a key factor in determining the pDC₅₀ of the PROTAC. However, this trend is only clear when the PROTACs exhibiting a pDC₅₀ < 5 are omitted from the analysis. This suggests that the ability of the PROTAC to occupy a conformation that facilitates productive ternary complex formation, enabling favourable bonding interactions with and between both the Brd4 protein and VHL E3 ligase is a key factor in determining the observed pDC₅₀. The PROTACs exhibiting a pDC₅₀ < 5 generally contain a short linker, such as PROTACs **138**, **139**, **165**, **168** and **169** with only two atoms between the amide carbonyls.

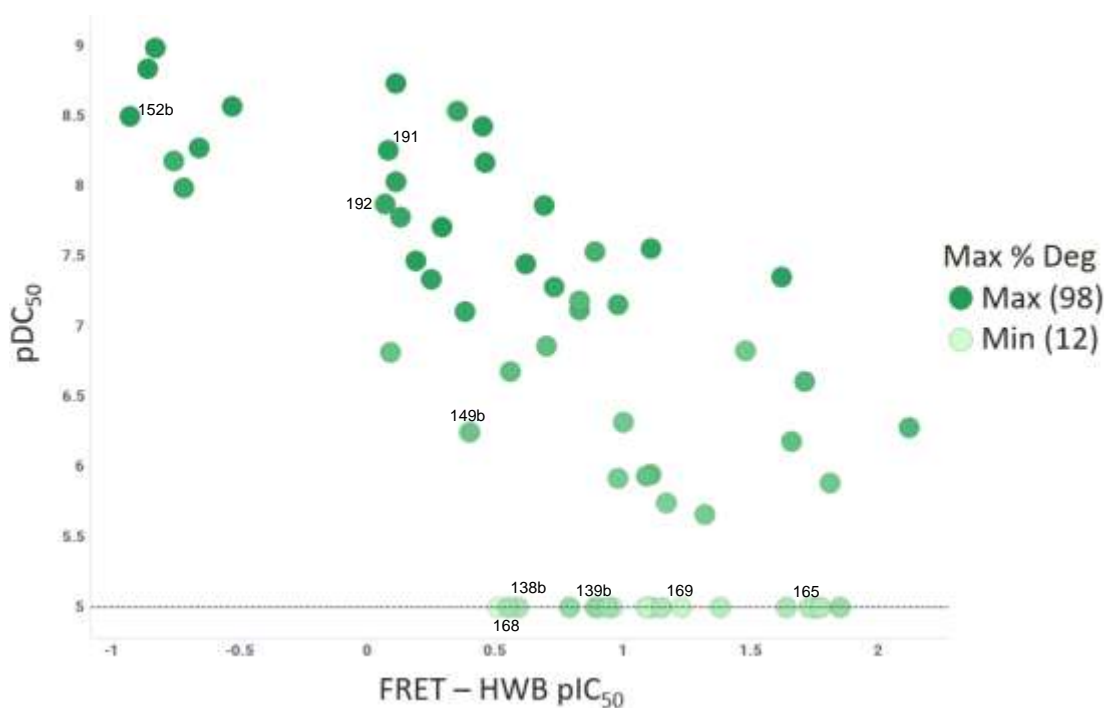


Figure 87: Plot of pDC_{50} vs $FRET - HWB pIC_{50}$ for the PROTAC series shown in green. The shading of the points corresponds to their max % deg values, with dark green representing a max % deg of 98% and light green representing a max % deg of 12%.

Comparing compounds with the same $FRET - HWB pIC_{50}$ should allow for the productivity of ternary complex formation to be assessed independently. Both PROTACs **192** and **191**, containing a pyridyl linker that is linked to the protein binder **128** through the 2-position, exhibited a $FRET - HWB pIC_{50}$ of 0.07 and 0.08 respectively. However, PROTAC **191** exhibits a pDC_{50} of 8.26 that is significantly higher than PROTAC **192**, which exhibits a pDC_{50} of 7.88. This suggests that the *meta*-substituted pyridyl linker in PROTAC **191** is better able to accommodate productive ternary complex formation than the *para*-substituted pyridyl linker of PROTAC **192**. This trend is more pronounced for PROTACs **152b** and **149b** with *meta*- and *para*-substituted phenyl linkers respectively. These PROTACs also exhibit a large difference in their $FRET - HWB pIC_{50}$ values. This may suggest that the more folded structure of the *meta*-substituted PROTAC **152b** is also exhibiting a higher permeability as it is better able to mask its polar moieties when passing through the hydrophobic membrane. However, Brd4 degradation may be contributing to the observed $FRET - HWB pIC_{50}$ for the more potent PROTAC **152b**.

A weak negative correlation between FRET - HWB pIC_{50} and $P\Delta C$ is observed for the PROTAC series, **Figure 88**. The points in **Figure 88** are shaded relative to their pDC_{50} values, with dark green representing a pDC_{50} of 8.99 and light green representing a pDC_{50} of < 5 . The PROTACs that exhibit the largest decrease in HWB pIC_{50} appear to occupy the top left-hand quadrant of the graph and exhibit pDC_{50} values < 7 . Whilst the PROTACs that exhibit the highest permeability and $P\Delta C$ values > 1 occupy the bottom right-hand quadrant of the graph and exhibit the highest pDC_{50} values. This suggests that FRET - HWB pIC_{50} may be a function of permeability, however, Brd4 degradation and other factors may be influencing the observed HWB pIC_{50} .

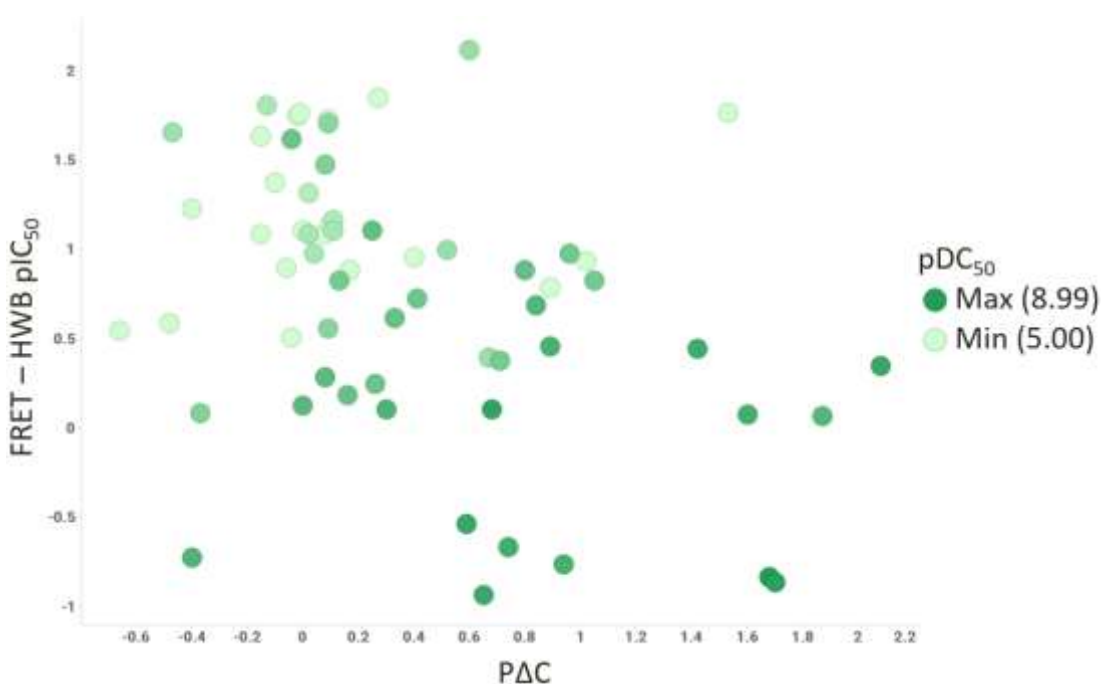


Figure 88: Plot of FRET - HWB pIC_{50} vs $P\Delta C$ for the PROTAC series shown in green. The shading of the points corresponds to their pDC_{50} values, with dark green representing a pDC_{50} of 8.99 and light green representing a pDC_{50} of < 5.00 .

A positive correlation between pDC_{50} and $P\Delta C$ is observed for the PROTAC series by omitting PROTACs exhibiting a $pDC_{50} < 5$ from the analysis, **Figure 89**. This suggests that cell permeability and thus intracellular concentration makes a significant contribution to the observed Brd4 degradation. $P\Delta C$ also exhibits a positive correlation with ChromlogD. This suggests that increases in the lipophilicity of the PROTACs to drive permeability are favourable for Brd4 degradation, in spite of the potential

negative side-effects associated with higher lipophilicity compounds, such as lower aqueous solubility and higher levels of undesirable protein-binding. The caveat to this is that these compounds would likely exhibit a lower oral bioavailability, thus limiting their potential dosing methods. However, the primary aim of this analysis is not to assess the feasibility of developing an orally bioavailable PROTAC.

It is also apparent that cell permeability is not the only factor in determining the Brd4 degradation potency, as PROTACs **197** and **176** exhibit similar pDC₅₀ values of 8.57 and 8.54 respectively, with significantly different PAC values of 0.59 and 2.08. This suggests that a lower intracellular concentration of PROTAC **197** is required to elicit a similar degradation profile to PROTAC **176**. This suggests that PROTAC **197** is better able to adopt a conformation for productive ternary complex formation due to its rigid 2,5-disubstituted furan linker, whilst PROTAC **176**, with the 12-carbon linker, has a high level of conformational flexibility, which may enhance its permeability through conformational folding to mask polar groups as it passes through the cell's lipid bilayer. However, this also increases the number of accessible conformations and therefore reduces the population of the conformation that facilitates productive ternary complex formation and Brd4 degradation. The position of the ring nitrogen of the pyridyl-linked PROTACs **158b**, **159b** and **191** appears to have an effect on both the pDC₅₀ and PAC values. PROTAC **159b** exhibits the highest pDC₅₀ of the entire PROTAC series of 8.99. PROTAC **158b** and **191** both exhibit lower pDC₅₀ values of 8.27 and 8.26 respectively. However, PROTAC **158b** exhibits a significantly lower PAC of 0.74. This may suggest that the *meta*-substituted pyridyl linkers occupy a folded conformation when passing through the cell's lipid bilayer to mask the polar groups of the protein and E3 ligase-binding moieties. For PROTAC **158b**, this results in the exposure of the polar pyridyl nitrogen HBA and thus a decrease in the PAC, whereas, the pyridyl nitrogen is better masked in this folded conformation for PROTACs **159b** and **191**. Due to the similar pDC₅₀ values of PROTACs **158b** and **191**, this also suggests that PROTAC **158b** is better able to adopt a conformation that maximises ternary complex formation. This may be due to the position of the HBA pyridyl nitrogen maximising favourable bonding interactions between the linker and the Brd4 protein and VHL E3 ligase. However, PROTAC **158b** also exhibits a

significantly lower ChromlogD, which correlates with a lower promiscuity of protein-binding. Therefore, both PROTACs might have a similar protein unbound intracellular concentration. As only the unbound fraction of PROTAC is capable of facilitating Brd4 engagement, ubiquitination and degradation, this may account for the similar degradation profiles.

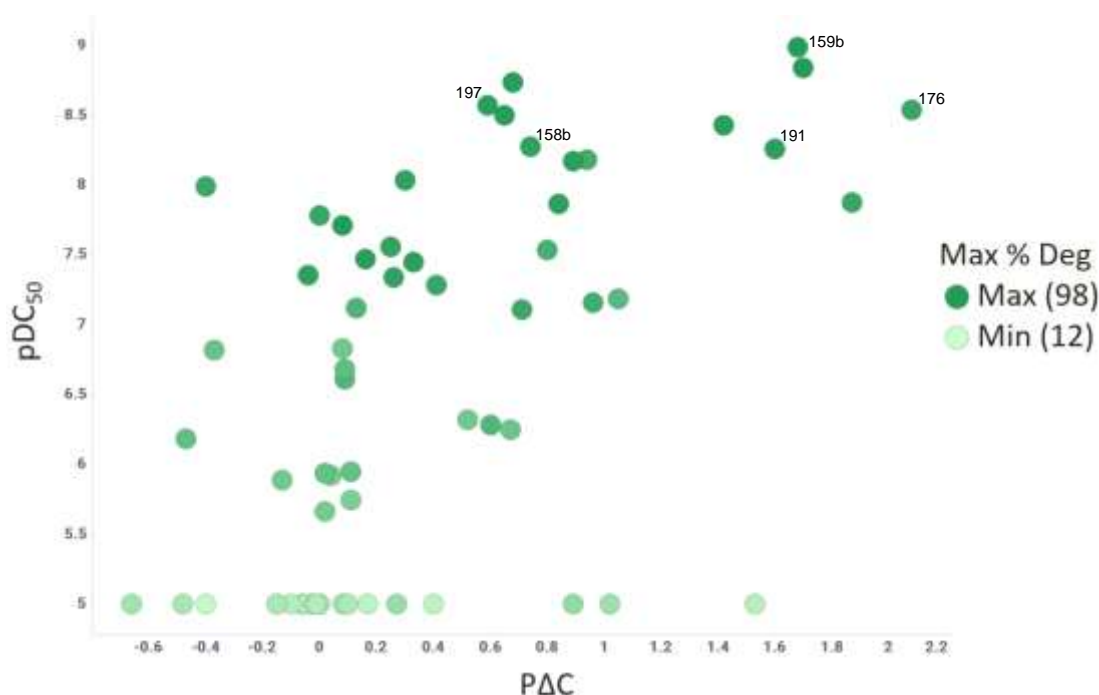


Figure 89: Plot of pDC₅₀ vs PΔC for the PROTAC series shown in green. The shading of the points corresponds to their max % deg values, with dark green representing a max % deg of 98% and light green representing a max % deg of 12%.

FRET pIC₅₀, HWB pIC₅₀ and pDC₅₀ show no correlation with solubility for the PROTAC series. This is visualised by the graph of pDC₅₀ vs CAD solubility, **Figure 90**. However, all PROTACs exhibiting a pDC₅₀ > 8.1 also exhibit a CAD solubility of ≤ 52 μg/mL. As ChromlogD and CAD solubility exhibit a negative correlation, this may suggest that the more lipophilic compounds exhibit a higher permeability, as well as more favourable hydrophobic binding interactions with the Brd4 protein and VHL E3 ligase in the ternary complex, furnishing a higher pDC₅₀.

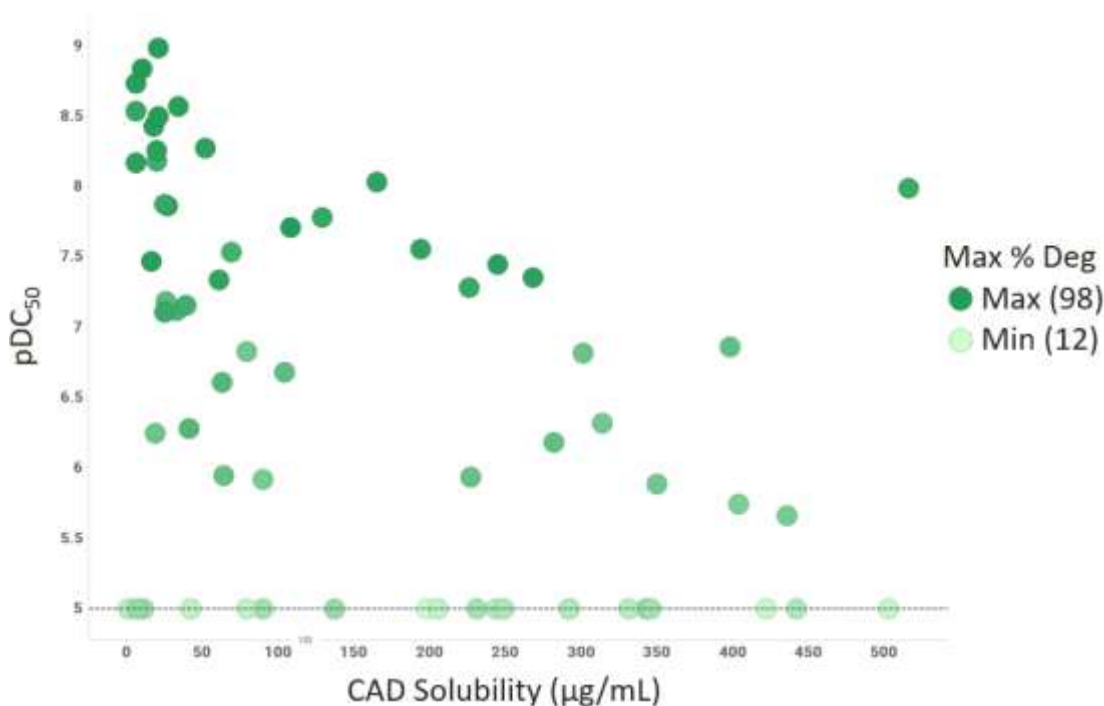


Figure 90: Plot of pDC_{50} vs CAD solubility for the PROTAC series shown in green. The shading of the points corresponds to their max % deg values, with dark green representing a max % deg of 98% and light green representing a max % deg of 12%.

The correlation between pDC_{50} and ChromlogD for the PROTAC series is visualised in **Figure 91**. ChromlogD exhibits a stronger correlation with both pDC_{50} and PAC than $clogP$. This suggests that the effective lipophilicity is a more accurate and useful measurement of lipophilicity than the intrinsic lipophilicity. This is likely due to presence of ionisable moieties in the PROTACs.

As ChromlogD increases, so does the maximum pDC_{50} achieved, up to the two most potent degraders, PROTACs **159b** and **163**, which exhibit pDC_{50} values of 8.99 and 8.84 respectively. Further increases in ChromlogD beyond 3.83 furnish a decrease in the maximum pDC_{50} achieved. Classically, in a small molecule drug discovery program, assuming an ARc of 2-3, a ChromlogD of 3-4 would be desirable.¹⁸⁶ For the PROTAC series generated, the majority of PROTACs exhibiting a ChromlogD > 4 also exhibit potent Brd4 degradation, with PROTAC **176** achieving a pDC_{50} of 8.54 with a ChromlogD of 5.55. This suggests that for this PROTAC series, higher values of ChromlogD may be tolerated. However, further investigations in animal models would be required to understand the effects of the increased lipophilicity on metabolic

clearance, promiscuity and bioavailability *in vivo*. A similar trend is observed between HWB pIC₅₀ and ChromlogD, whilst FRET pIC₅₀ exhibits no correlation with ChromlogD. This suggests that PROTAC **159b**, with a ChromlogD of 3.82, exhibits the most desirable balance of physicochemical properties to elicit the most potent Brd4 degradation.

PROTAC **185**, containing the 1,3-substituted biaryl linker, exhibits a ChromlogD of 4.60 and a pDC₅₀ of 8.74. Substitution of the biaryl linker for the bipyridyl linker in PROTAC **163** furnishes a lower ChromlogD of 3.83 and a higher pDC₅₀ of 8.84. This substitution also results in an increase in the HBA count, TPSA and CAD solubility. The observed increase in pDC₅₀ may be due to more desirable physicochemical properties and thus an increase in cellular concentration, highlighted by the increase in HWB pIC₅₀ from 7.68 to 8.71. However, the bipyridyl linker may also be better able to occupy a ternary complex conformation that facilitates ubiquitination of the Brd4 protein.

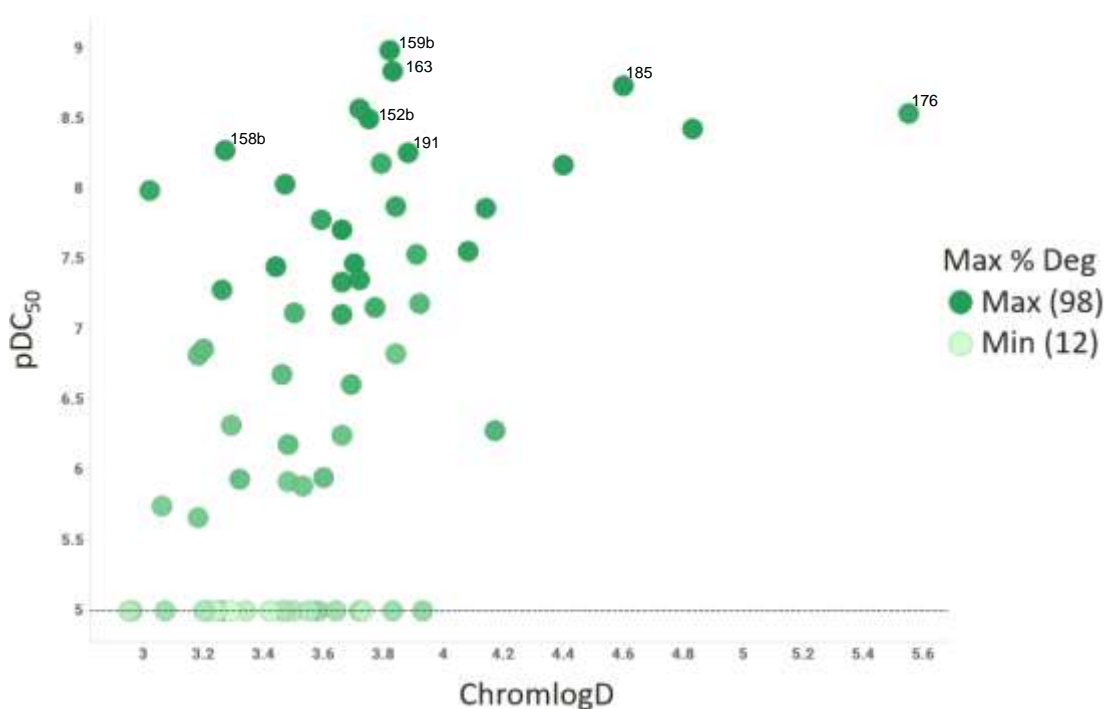


Figure 91: Plot of pDC₅₀ vs ChromlogD for the PROTAC series shown in green. The shading of the points corresponds to their max % deg values, with dark green representing a max % deg of 98% and light green representing a max % deg of 12%.

However, an increase in ChromlogD is observed for the substitution of a carbon atom in the 1,3-substituted aryl linker of PROTAC **152b** for a nitrogen atom, this furnishes PROTAC **159b** and an increase in ChromlogD from 3.75 to 3.82. This increase is negligible and within the error of the assay, however, it is accompanied by a significant increase in pDC₅₀ from 8.50 to 8.99. Additionally, although PROTAC **159b** has a higher calculated TPSA, no increase in CAD solubility from 21 µg/mL is observed. This suggests that PROTAC **159b** may be occupying a conformation that masks the additional polarity of the pyridyl nitrogen in solution. However, the HBA pyridyl nitrogen may be able to make more favourable bonding interactions with the Brd4 protein and VHL E3 ligase, increasing the cooperativity and resulting in an increase in degradation.

The position of the pyridyl nitrogen also appears to have an effect on the observed pDC₅₀ values for the PROTACs. PROTAC **191**, containing the 2,4-substituted pyridyl linker exhibited a higher ChromlogD of 3.88, whilst PROTAC **158b**, containing the 3,5-substituted pyridyl linker exhibited a higher CAD solubility of 52 µg/mL. Both PROTACs exhibited significantly lower pDC₅₀ values than PROTAC **159b** of 8.26 and 8.27 respectively. This may further suggest that PROTAC **159b** exhibits the most desirable balance of physicochemical properties for potent Brd4 degradation. However, it may also suggest that the nitrogen of the 2,6-substituted pyridyl linker in PROTAC **159b** is in the optimum position to make favourable bonding interactions with the Brd4 protein and VHL E3 ligase, facilitating favourable PPIs in the ternary complex.

No correlation is observed between pDC₅₀ and RBC when considering the entire PROTAC series, **Figure 92**. However, when considering PROTACs **208**, **209**, **210** and **211**, with PEG linkers, a decrease in pDC₅₀ is observed as the RBC increases. As previously stated, these PROTACs exhibit similar ChromlogD values in spite of an increase in M_w due to the hydrophilic properties of the additional EG units. Additionally, all PEG PROTACs exhibit a HWB pIC₅₀ around 6.5. This suggests that beyond an optimum length, further increases in conformational flexibility results in an increase in the number of possible conformations and therefore a decrease in the population of the productive ternary complex conformation required for degradation.

However, focusing on the series of unfunctionalised all-carbon linkers, PROTACs **137**, **174**, **140b**, **141b**, **142b** and **176**, a general positive correlation between pDC_{50} and RBC is observed, with the exception of the 7-carbon linker PROTAC **175**. This trend conflicts with that observed for the previously described PEG linker series. This may be due to the fact that for the unfunctionalised all-carbon linker series an increase in RBC correlates with an increase in ChromlogD, which may result in an increase in permeability and intracellular concentration and thus Brd4 degradation.

This highlights the complexity of the relationship between degradation and physicochemical properties, which are both governed by multiple interdependent variables.

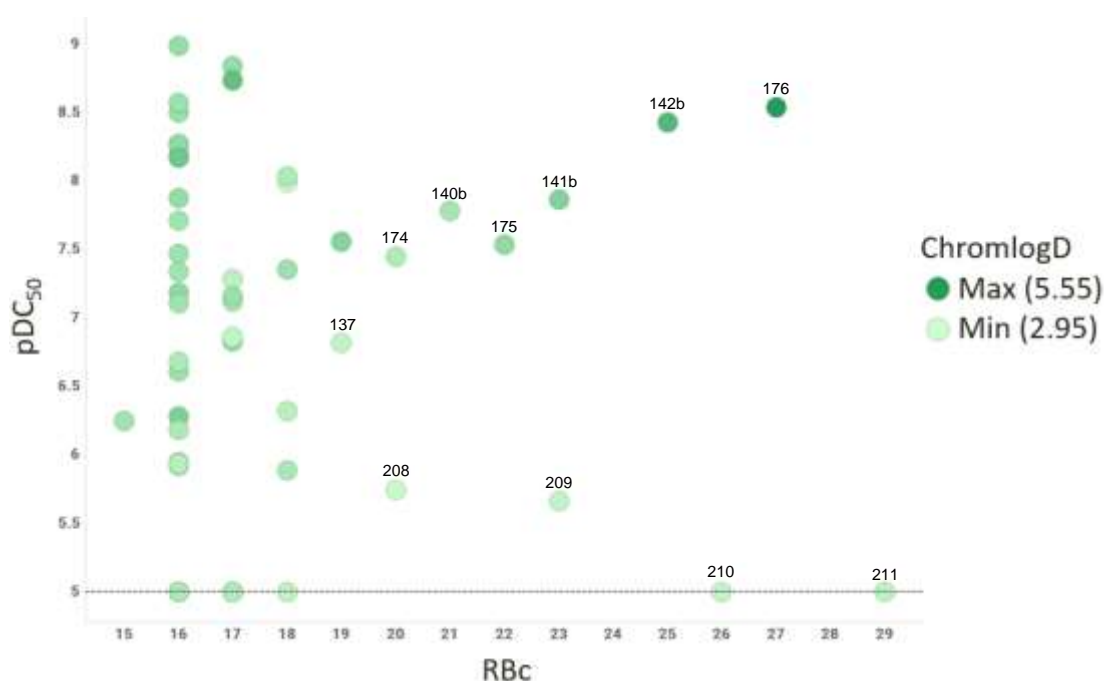


Figure 92: Plot of pDC_{50} vs RBC for the PROTAC series shown in green. The shading of the points corresponds to their measured ChromlogD values, with dark green representing a ChromlogD of 5.55 and light green representing a ChromlogD of 2.95.

There is no observable correlation between pDC_{50} and M_w , ARc or F_{sp}^3 . The correlation between pDC_{50} and HSA binding for the PROTAC series is visualised in **Figure 93**. A positive correlation between pDC_{50} and HSA binding is observed by omitting PROTACs exhibiting a $pDC_{50} < 5$ from the analysis. PROTACs exhibiting a $pDC_{50} > 8$ also exhibit a HSA binding $\geq 90\%$. However, HSA binding also observes a

positive correlation with ChromlogD, which may also be influencing this trend. A similar trend is observed between HWB pIC_{50} and HSA binding. Dosing in an animal model would be required to determine the HSA binding required to achieve the optimum balance of unbound molecules that are capable of entering the target tissue and eliciting the desired therapeutic effect, with sufficient HSA binding that enables the transportation of the molecule to the target tissue and protects the compounds from oxidation, metabolism and excretion.

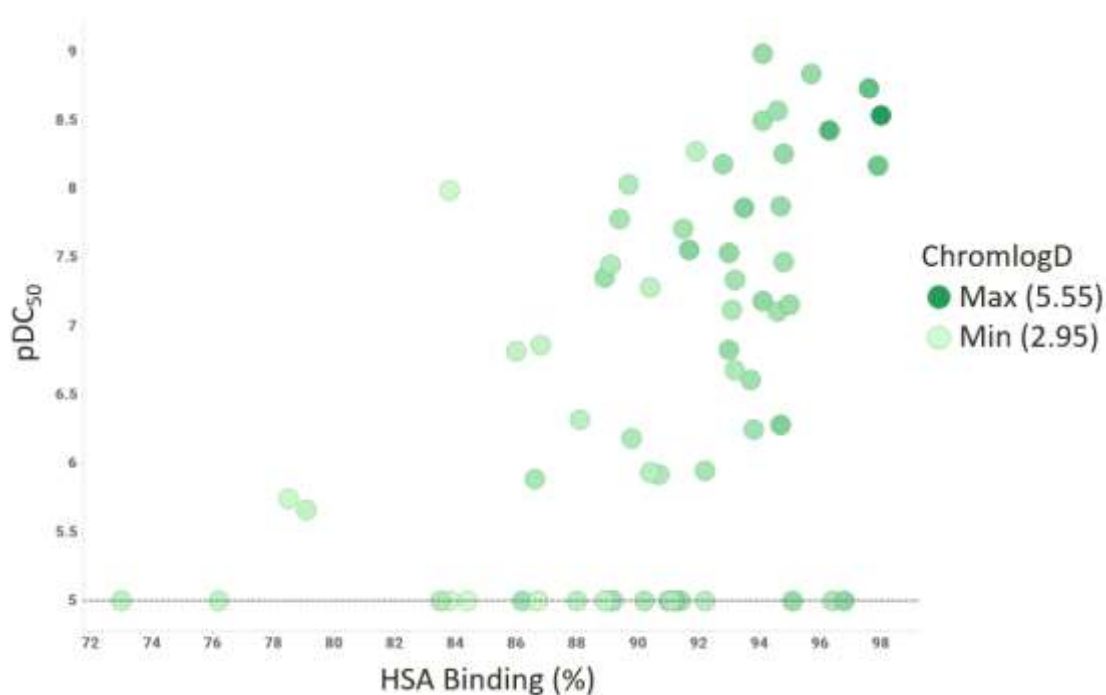


Figure 93: Plot of pDC_{50} vs HSA binding for the PROTAC series shown in green. The shading of the points corresponds to their measured ChromlogD values, with dark green representing a ChromlogD of 5.55 and light green representing a ChromlogD of 2.95.

There is no observable correlation between pDC_{50} and CHI_{IAM} when analysing the entire PROTAC series, **Figure 94**. However, a weak positive correlation can be observed by omitting PROTACs exhibiting a $pDC_{50} < 5$ from the analysis. However, as CHI_{IAM} also exhibits a positive correlation with ChromlogD, this property may be responsible for the observed trend with pDC_{50} . There is no observable correlation between HWB pIC_{50} and CHI_{IAM} .

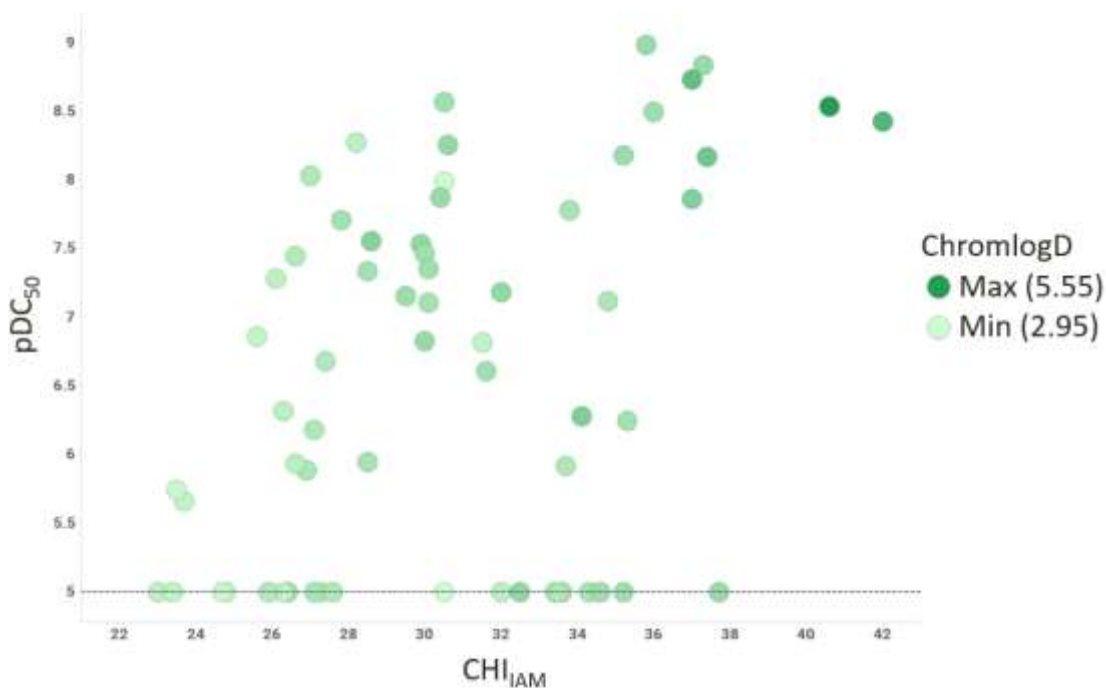


Figure 94: Plot of pDC_{50} vs CHI_{IAM} for the PROTAC series shown in green. The shading of the points corresponds to their measured ChromlogD values, with dark green representing a ChromlogD of 5.55 and light green representing a ChromlogD of 2.95.

There is no observable correlation between AMP, PAC, CAD or FaSSIF solubility and AB-MPS. However, > 90% of the PROTAC series exhibits an AB-MPS < 28, which suggests that they will be parenterally bioavailable, indicating a favourable balance of solubility and permeability. Additionally, no correlation between AB-MPS and HWB pIC_{50} or FRET - HWB pIC_{50} is observed. This may suggest that it is not possible to independently analyse the complex relationship of these interdependent variables with bioavailability, or that approximating bioavailability using AB-MPS, which uses crude molecular descriptors such as RBC and ARc, does not provide an accurate measure of bioavailability. Bioavailability studies in animal models would be required to acquire the relevant data to elucidate any trends with physicochemical properties and AB-MPS.

In this section, the effects of linker functionality on the physicochemical properties, Brd4 binding affinity and degradation profile of the PROTAC series, were investigated and observable trends were identified and rationalised. However, quantifying these

complex trends between multiple interdependent properties is challenging. Therefore, a more in-depth statistical analysis of the data is required.

4.6 PCA/PLS Analysis of the PROTAC Series

Principal component analysis (PCA) forms the basis for multivariate data analysis. It enables the representation of multivariate data sets in a two-dimensional plane, such that an overview of the data that is highly interpretable is obtained.

A range of physicochemical descriptors was used as the input variables to describe each PROTAC of the series. These descriptors were calculated using the GSK in-house models and included: cChromlogD (standard, HAC and MPNN),²¹⁴ logP, protein-binding (HSA, IAM, AGP),²¹⁵ pK_a acid and base classes, PCv6, and Abraham descriptors. The definitions of these descriptors are outlined in **Section 5.6**. The measured ChromlogD and CAD solubility were also included. However, the in-house models are unable to calculate all of the descriptors for compounds that contain > 70 heavy atoms, PROTACs **163**, **176**, **182**, **185**, **186**, **210** and **211**. Therefore, each PROTAC of the series was truncated at positions A and B in **Figure 95** to furnish linker truncates 1 (LT1) and 2 (LT2).

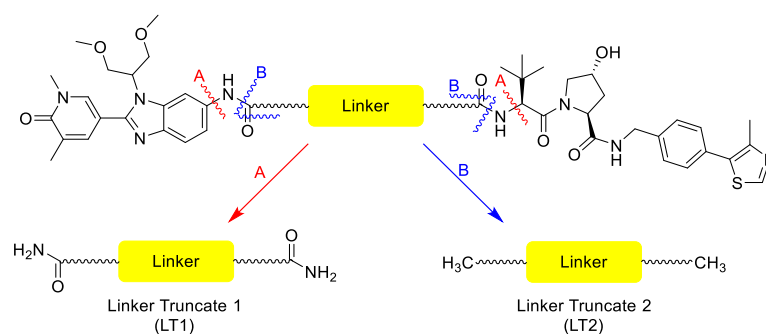


Figure 95: Truncation of the full PROTAC to generate LT1 and LT2.

The electrotopological-state index (e-state) descriptors, which encode information about the topological environment of the subject group and the electronic interactions due to all other atoms in the molecule, were calculated for LT1 and LT2, in addition to the previously described descriptors. The resulting data are included in the digital appendices. Including LT1 and LT2 in the analysis will also enable the effects of the

linker to be analysed independently from the full PROTAC (FP). The PCA analysis was performed using the method and software described in **Section 5.6**. A variable space was constructed with as many dimensions as there are descriptors, with each variable representing one co-ordinate axis. Each PROTAC was plotted in the variable space to create a swarm of points in the multi-dimensional space. The first principal component (PC1) is computed by least squares analysis to maximise the variance of the co-ordinates on the line and minimise the residual variance. The second principal component (PC2) is computed to generate a line that is orthogonal to PC1 and defines a two-dimensional plane. The co-ordinate values of the descriptors on this plane are called scores (t_1 and t_2). These co-ordinates were used to generate a scores scatter plot for the PROTAC series, **Figure 96**.

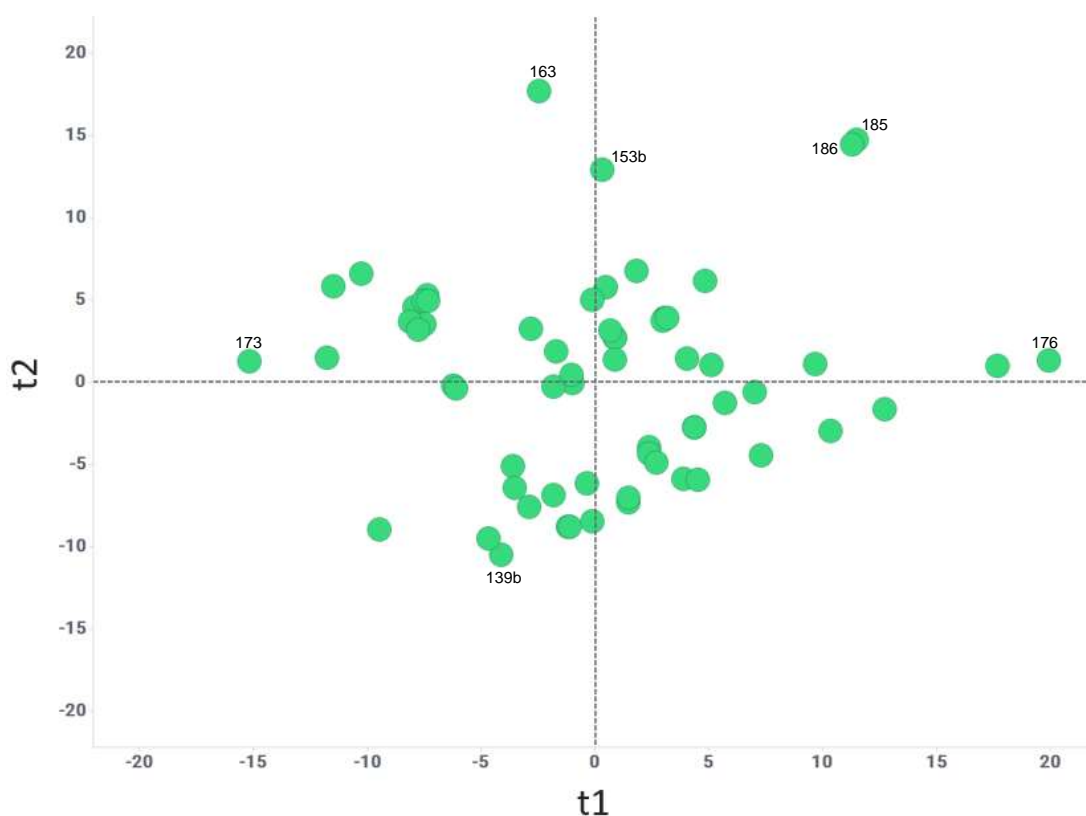


Figure 96: Scores scatter plot for the PROTAC series.

The scatter plot highlights that the PROTACs exhibit broad coverage of the descriptor space and are not all clustered into subsets. PROTACs that exhibit similar descriptor profiles, such as PROTACs **185** and **186**, occupy similar space on the scatter plot. PROTAC **176** exhibits the highest t_1 co-ordinate, whilst PROTAC **173** exhibits the

lowest. As these PROTACs represent the highest and one of the lowest measured ChromlogD values, this may suggest that lipophilicity is making a key contribution to the fitting of PC1. Additionally, PROTACs containing biaryl or fused aromatic linkers such as PROTACs **153b**, **163**, **185** and **186** are exhibiting the highest t2 co-ordinate, whilst non-aromatic linkers such as **139b** exhibit the lowest. This may suggest that aromaticity is making a significant contribution to the fitting of PC2.

Projections to latent structures by means of partial least squares (PLS) is a regression extension of PCA that can relate two data sets, A (physicochemical descriptors) and B (responses), to each other by a linear multivariate model. The responses that make up data set B include: FRET pIC₅₀, HWB pIC₅₀, pDC₅₀ and max % deg. This should allow us to identify physicochemical descriptors that are driving these responses and are therefore important properties in determining the Brd4 degradation profile of the PROTACs.

Two variable spaces were constructed for data set A and B. Each PROTAC was plotted in the variable spaces to create two point-swarms in each of the multi-dimensional spaces. The first PLS component (PLS C1) is computed in much the same way as PC1, however it also maximises the correlation between the positions of the points in the multi-dimensional spaces A and B. The second PLS component is computed to improve the approximation of, and correlation between, the position of PLS C1 in the multi-dimensional spaces A and B. The O2PLS method was used for this analysis as there are more than two responses in data set B that are interdependent.

The output of the PCA/PLS analysis is a variable influence on projection (VIP) parameter which summarises the importance of the physicochemical descriptors to explain the correlation between the descriptors in data set A and to correlate the responses in dataset B. VIP values ≥ 1 indicate important descriptors, whilst values < 1 generally indicate unimportant descriptors. Descriptors with a VIP value < 1 were excluded and the model refitted using the refined data set A.

The fit of the optimised PCA/PLS model can be assessed by plotting the observed response variables from data set B against the predicated response variables generated by the model. The graph for pDC₅₀ is shown in **Figure 97**. The regression line is

represented by the solid black line and has an R^2 value of 0.25. This suggests that the PCA/PLS model has identified some level of correlation between the physicochemical property descriptors and the pDC_{50} response. However, a large portion of the response remains unmodelled.

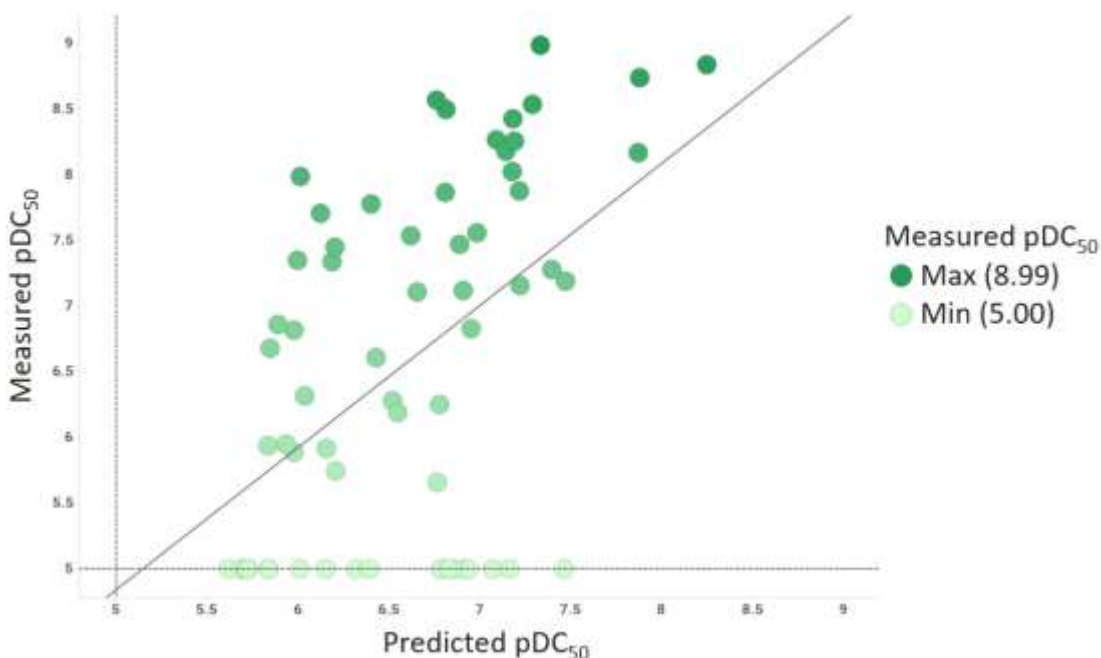


Figure 97: Plot of measured pDC_{50} vs predicted pDC_{50} for the PCA/PLS model generated for the PROTAC series shown in green. The shading of the points corresponds to their measured pDC_{50} values, with dark green representing a pDC_{50} of 8.99 and light green representing a pDC_{50} of < 5.00.

Of the four responses, the best fit is achieved for pDC_{50} . The next highest R^2 value of 0.17 was achieved for the max % deg response. This is expected as both the pDC_{50} and max % deg responses correlate. R^2 values of 0.14 and 0.12 were achieved for FRET pIC_{50} and HWB pIC_{50} respectively, which suggests a weaker correlation between the physicochemical descriptors and these responses was obtained from the PCA/PLS analysis.

The PCA/PLS analysis was repeated using only the physicochemical descriptors of LT1 and LT2. This excludes any contribution of the FP and will allow the effects of the linker functionality on the Brd4 binding affinity and degradation responses to be isolated. The R^2 values of the regression lines for the plots of the measured vs predicted responses decreased when the physicochemical descriptors of the FP were omitted. An

R^2 value of 0.2, 0.14, 0.11 and 0.13 was achieved for pDC_{50} , max % deg, FRET pIC_{50} and HWB pIC_{50} respectively. This suggests that the physicochemical properties of the FP are making some contribution to the prediction of the model. As both the protein and VHL E3 ligase-binding moieties remain constant in the PROTAC series, this may suggest that there is important information in how the properties of the linker manifest in the FP. The small drop off in the R^2 values of pDC_{50} and max % deg suggest that the physicochemical descriptors of the linker truncates are making a significant contribution to the correlation observed with the degradation responses. This suggests that the linker functionality plays a key role in determining the degradation profile of the PROTACs.

The 20 PROTACs that did not achieve a max % deg of 50% and were assigned a $pDC_{50} < 5$ appear to show the largest deviation from the regression line. These PROTACs are assigned a value of 5 in the analysis as the $<$ function cannot be computed. However, their actual pDC_{50} values could be anywhere between an infinitely small number and 4.9 recurring. Therefore, these inaccurate responses may be hindering the analysis.

The PCA/PLS analysis was repeated using a refined data set containing the physicochemical descriptors for the FP, LT1 and LT2 of the 43 PROTACs that achieved a $pDC_{50} > 5$. The optimised PCA/PLS model provided an improved fit for pDC_{50} , achieving an R^2 value of 0.56 for the regression line in **Figure 98**. The regression line also exhibits a slope of approximately one, which suggests that there is no statistical bias and that linear changes in the measured response would furnish a proportional change in the predicted response. The regression line also appears to provide a better fit for the PROTACs with the highest pDC_{50} values. Extending the range by obtaining more data points for PROTACs exhibiting a pDC_{50} between 5 and 6 would improve the accuracy of the model. PROTACs **143b**, **208** and **209**, containing ether linkages, appear to show the largest deviation from the regression line. This suggests that the complex effect of the ether linkages on the physicochemical profile of these PROTACs and their measured responses is not well predicted by the model. Although a significant R^2 value of 0.56 was achieved for the fit of the refined PCA/PLS model for pDC_{50} , all of the information in the model is not explained. This model is

not completely predictive and an R^2 value closer to one would be preferable. Therefore, the limitations of this model must be acknowledged and caution exercised when using it for predictive purposes. The accuracy of the model could be validated and developed further by virtually designing PROTACs, predicting their pDC_{50} responses and then synthesising and measuring their actual pDC_{50} , before feeding this information back into the model.

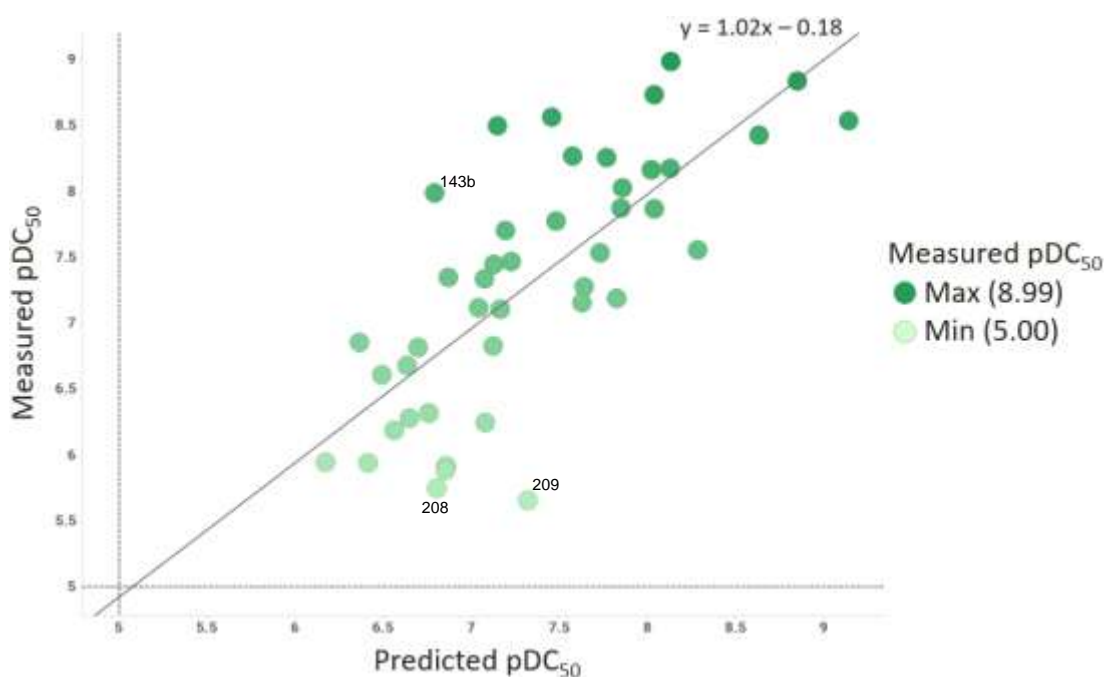


Figure 98: Plot of measured pDC_{50} vs predicted pDC_{50} for the PCA/PLS model generated for the refined PROTAC series shown in green. The shading of the points corresponds to their measured pDC_{50} values, with dark green representing a pDC_{50} of 8.99 and light green representing a pDC_{50} of < 5.00.

Both max % deg and HWB pIC_{50} also exhibited improved R^2 values of 0.36 and 0.22 respectively. However, the model obtained a lower R^2 value of 0.08 for FRET pIC_{50} . This suggests that the PROTACs that exhibited a $pDC_{50} < 5$ provided useful FRET pIC_{50} measurements, enabling a stronger correlation from the larger data set. However, the majority of the PROTACs' Brd4 binding affinity emanates from the protein-binding moiety, which was maintained for all PROTACs in the series, resulting in a small range of FRET pIC_{50} values that shows little correlation with the physicochemical descriptors. Significantly higher R^2 values were achieved for the pDC_{50} , max % deg and HWB pIC_{50} responses. These responses are all measured from

in vitro cell-based assays and are therefore influenced by similar factors such as permeability and solubility.

The relationship between the physicochemical descriptors and the responses for the predictive component of the model can be visualised by the loadings scatter plot, **Figure 99**. The loadings p1 and p2 are calculated from the sum of the vectors of the descriptors in both multi-dimensional spaces A and B from the PLS C1 and C2 respectively. Therefore, p1 describes the correlation between the physicochemical descriptors and responses defined by the model, whilst p2 reflects structured information within the descriptor space that is not correlated with the responses. The physicochemical descriptors are represented by green points, whilst the responses are represented by blue points.

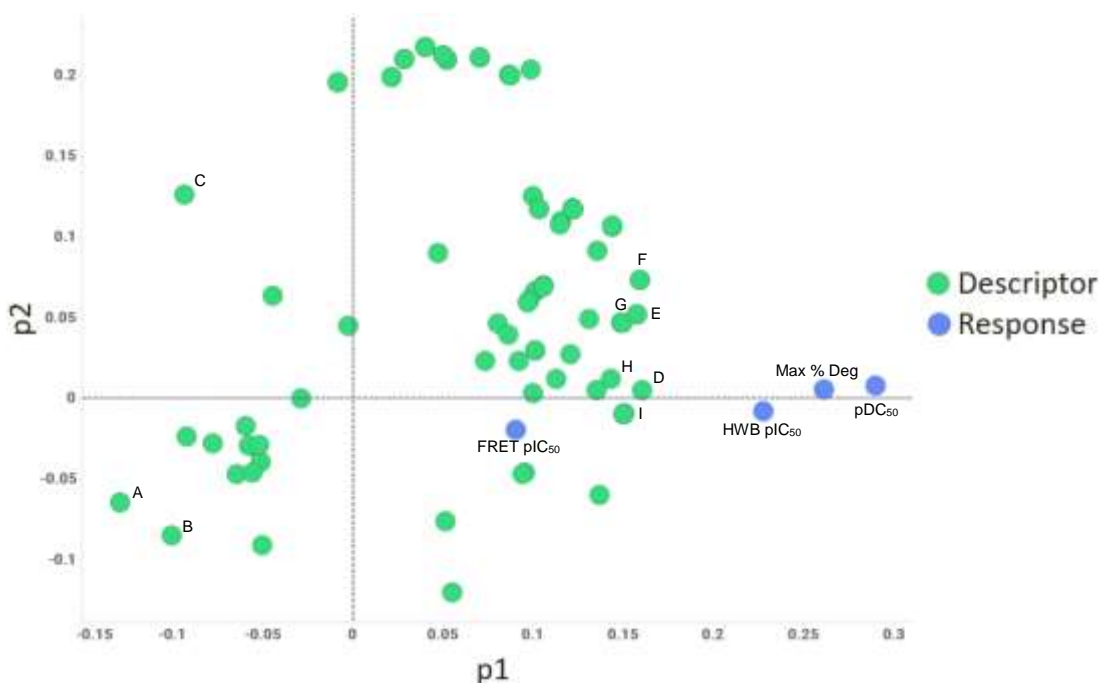


Figure 99: Loadings scatter plot for the PCA/PLS model generated for the PROTAC series. Physicochemical descriptors are represented by green points and responses are represented by blue points. A = LT1 PCv6.fraction HBD, B = Measured CAD solubility, C = FP CAD solubility, D = FP cChromlogD MPNN, E = LT1/2 PCv6.M_w, F = LT1/2 PCv6.HAC, G = LT1/LT2 PCv6.CMR, H = Measured ChromlogD and I = LT1/LT2 PCv6.vx and LT1/LT2 Abraham.vx.

Responses pDC₅₀, max % deg and HWB pIC₅₀ occupy similar space on the p1 axis, which suggests that there is a high degree of correlation between these responses. The

FRET pIC_{50} response has a smaller magnitude of $p1$, which suggests that the model is less predictive for this response. The descriptor $p1$ co-ordinates indicate the effect they have on the responses; the larger the magnitude of $p1$ the more the descriptor influences the responses. Positive $p1$ co-ordinates have an enhancing effect on the responses, whilst negative $p1$ co-ordinates have a deleterious effect on the responses.

The descriptors labelled A, B and C in Figure 99 correspond to the fraction of HBD in LT1, and the measured and calculated CAD solubility of the FPs respectively. These descriptors have a negative $p1$ co-ordinate and therefore have a deleterious effect on the responses. The points labelled D to I are all descriptors of ChromlogD, number of heavy atoms and size and have an enhancing effect on the responses. These descriptors all correlate with lipophilicity, which suggests that decreasing CAD solubility and increasing lipophilicity drives pDC_{50} , max % deg and HWB pIC_{50} . These two properties are inversely proportional and the described changes are generally used to improve the permeability of compounds, which may suggest that permeability is a key factor in determining the magnitude of the responses. It should be noted that descriptors of the FP, LT1 and LT2 are contributing to the formation of the model and the correlation with the responses. In some cases, the descriptors of LT1 and LT2 are plotted at the same co-ordinates, such as points E, F, G and I, which suggests that they make the same contribution. However, point A only represents the descriptor of LT1. This suggests that there is valuable information being computed by the different representations of the linker truncates. The points labelled E, F, G and I all correspond to descriptors for the size of the linker, which suggests that it makes a significant contribution to the observed responses. As the size of the linker will have an effect on the productivity of ternary complex formation, this may suggest that the geometric properties of the linker are having some effect on the responses.

The PCA/PLS model can be re-expressed as a regression model, considering the correlation between the physicochemical descriptors and only one of the response variables. This simplifies the interpretation of the model but information regarding the correlation between the responses is lost. This was performed for pDC_{50} and the regression coefficients for the most important descriptors are shown in the coefficients plot in **Figure 100**.

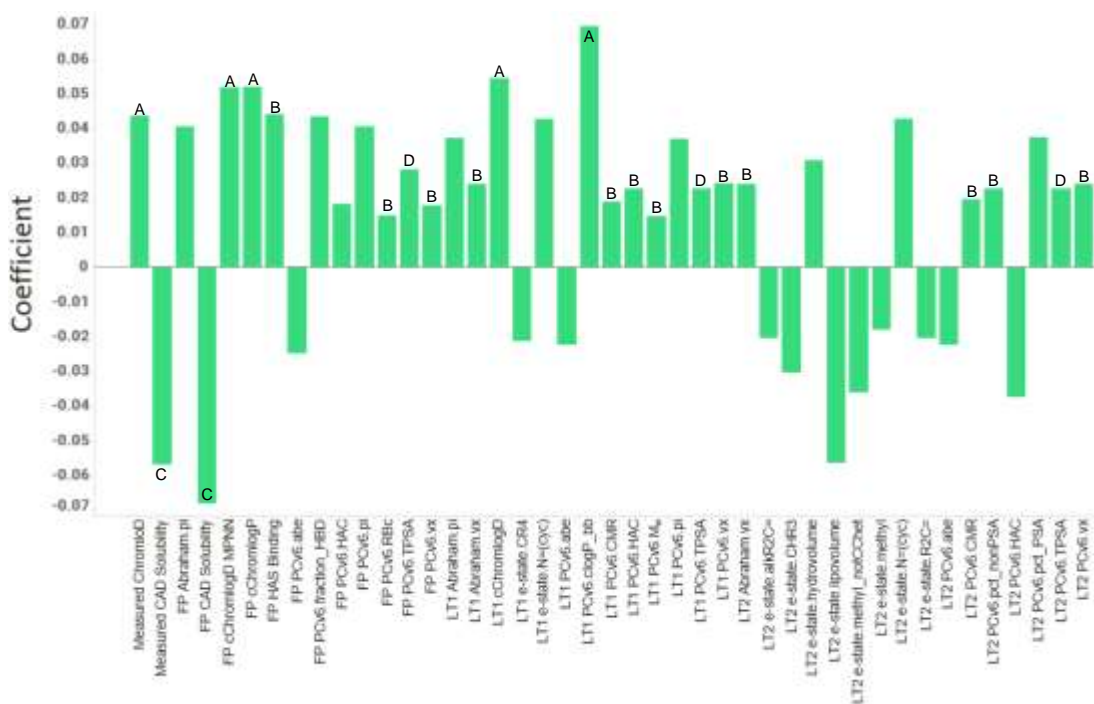


Figure 100: Coefficients plot for the PCA/PLS model generated for the PROTAC series.

The size and sign of the coefficients indicate the influence of each physicochemical descriptor on the pDC₅₀ response. There is approximately an equal number of descriptors for the FP, LT1 and LT2 represented in the coefficients plot, which suggests that the physicochemical properties of both the FP and linker have a significant effect on the observed pDC₅₀ of the PROTACs. However, it should be noted that some of the descriptors are repeated for all three structural representations of the PROTAC, such as PCv6.abe, .vx, .TPSA and .HAC. Therefore, the average influence of these coefficients needs to be considered.

The descriptors for ChromlogD and clogP exhibit the highest positive coefficients for pDC₅₀ and are labelled A in **Figure 100**. This suggests that increasing the lipophilicity of the PROTACs results in an increase in Brd4 degradation. This may be due to the positive correlation exhibited between lipophilicity and permeability. Additionally, descriptors that exhibit a positive correlation with lipophilicity, such as the HAC, RBC, M_w, CMR, .vx and HSA binding, labelled B in **Figure 100**, also exhibit positive coefficients for pDC₅₀. As increasing the amount of HSA binding also reduces the free concentration of PROTAC in the cell that is capable of engaging the Brd4 protein and

VHL E3 ligase, the observed positive correlation may be due to lipophilicity. However, HAC, RBC, M_w , CMR, and .vx are all descriptors that also correlate with molecular size.

Measured and calculated CAD solubility descriptors for the FP, labelled C in **Figure 100**, exhibit the largest negative coefficients for pDC_{50} . This suggests that lowering the CAD solubility of the PROTACs actually results in an increase in Brd4 degradation. However, only the fraction of solubilised PROTAC is capable of permeating into cells and eliciting degradation. Therefore, this may suggest that the solubility of the PROTACs is not limiting the degradation. As CAD solubility exhibits a negative correlation with both lipophilicity and permeability, the observed negative coefficient for CAD solubility may be a manifestation of its relationship with these two physicochemical properties.

Interestingly, the TPSA descriptors, labelled D in **Figure 100**, exhibit a positive coefficient for all three structural representations of the PROTAC. This suggests that increasing TPSA whilst decreasing the CAD solubility is favourable for Brd4 degradation. One way of achieving this is through the introduction of a heteroatom that is dynamically masked in solution and exposed in the hydrophobic binding pocket.

The lipophilicity of the PROTACs appears to be a key physicochemical property in determining Brd4 degradation potency. However, a number of descriptors that correlate with the size and shape of the linker also appear to make a significant contribution. The size and geometric properties of the linker will have an effect on the productivity of ternary complex formation. Therefore, these properties need to be considered when assessing the effects of the linker on the physicochemical properties of PROTACs and their resulting degradation profiles.

The physicochemical descriptors and responses were remodelled in DS Biovia quantitative structure-activity relationship (QSAR) workbench to investigate whether an improved fit for the data could be achieved by treating the relationship between the descriptors and responses as non-linear.²¹⁶ However, no improvement was achieved.

4.7 Geometric Properties of the PROTAC Series

The physicochemical properties of the PROTACs that achieved the most and least potent Brd4 degradation are visualised in **Figure 101** and **Table 27**. PROTAC **159b**, containing the 2,6-disubstituted pyridyl linker, exhibited the most potent degradation, with a pDC₅₀ of 8.99 and a max % deg of 97%. PROTAC **147b**, containing the bicyclo[1.1.1]pentane linker, exhibited the least potent degradation, with a pDC₅₀ of < 5 and a max % deg of 12%. The radar plot in **Figure 101** shows that both PROTACs exhibit desirable physicochemical properties that are within the guidelines outlined by DeGoey and Edmondson in **Section 4.1.2**. Additionally, it highlights how similar the physicochemical properties are for both PROTACs, which are reported in **Table 27** with the Brd4 binding affinity and degradation data.

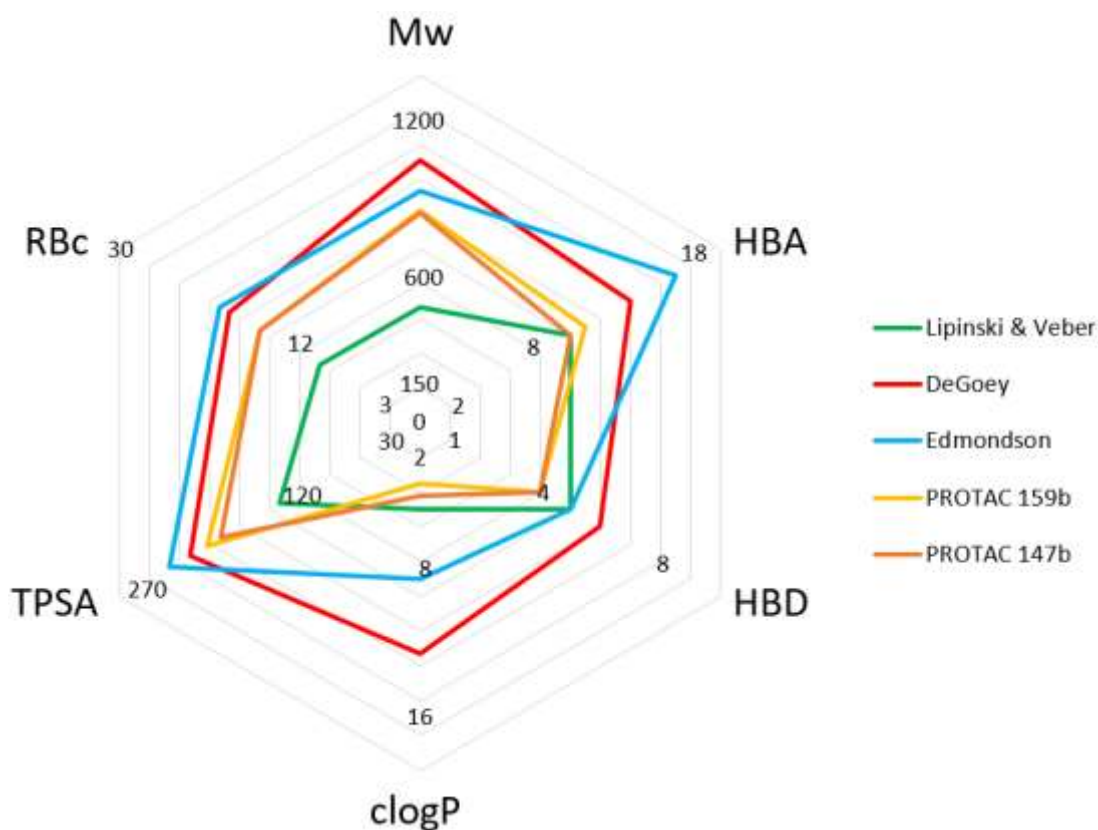


Figure 101: Radar plot of the physicochemical properties of PROTACs **159b** and **147b**, as well as the guidelines published by Lipinski & Veber, DeGoey and Edmondson.

Table 27: *Brd4* binding affinity, degradation and physicochemical properties for PROTACs **159b** and **147b**.

Brd4 Binding / Degradation	159b	147b
FRET pIC ₅₀	7.91	7.46
HWB pIC ₅₀	8.74	6.95
pDC ₅₀	8.99	< 5
Max % Deg (%)	97	12
Physicochemical Properties		
M _w (Da)	918	907
HBA	11	10
HBD	4	4
clogP	3.5	4.2
TPSA (Å ²)	212	199
RBc	16	16
ARc	6	5
Fsp ³	0.4	0.5
AB-MPS	23.9	22.9
CAD Solubility (µg/mL)	21	198
FaSSIF Solubility (µg/mL)	39	112
AMP (nm/s)	< 3	< 3
PΔC	1.68	-0.04
ChromlogD	3.82	3.29
HSA Binding (%)	94	88
CHI _{IAM}	35.8	32.0

Both PROTACs exhibit FRET pIC₅₀ values > 7, which suggests potent inhibition of the Brd4 protein *in vitro*. Although PROTAC **159b** exhibits a higher TPSA, PROTAC **147b** exhibits a higher CAD and FaSSIF solubility, as well as a lower ChromlogD. The more lipophilic PROTAC **159b** exhibits a significantly higher PΔC value of 1.68, suggesting that it is highly permeable, whilst PROTAC **147b** exhibits a PΔC value of -0.04, which suggests that it exhibits limited permeability. This suggests that lipophilicity and permeability are key driving forces in determining degradation.

However, PROTAC **153b**, containing the indole linker, exhibits a PΔC value of 1.53 and achieves a max % deg of only 25%. This suggests that whilst the binding affinity and physicochemical properties of the PROTACs play a key role in determining the intracellular concentration and Brd4 engagement, the ability of the PROTAC to facilitate the formation of a productive ternary complex that enables ubiquitination of a Brd4 surface-exposed lysine residue is paramount for potent degradation.

A crude measurement of the PROTACs' linker length was taken by counting the fewest number of bonds between the amide carbonyl groups. The correlation between pDC₅₀ and crude linker length is visualised in **Figure 102**. PROTAC **159b**, containing a 2,6-substituted pyridyl linker, exhibited the most potent degradation of Brd4 protein of the PROTAC series, with a pDC₅₀ of 8.99. This PROTAC has a linker length of four bonds between the amide carbonyls, which suggests that this is the optimum length for ternary complex formation. The PROTACs with a linker length of four that exhibit a pDC₅₀ > 8 all contain a 1,3-disubstituted 5- or 6-membered aromatic ring as the linker.

Substituting the pyridyl linker in PROTAC **159b** for the phenyl linker in PROTAC **152b** results in a lower pDC₅₀ of 8.50. This suggests that the pyridyl nitrogen of PROTAC **159b** is capable of making additional bonding interactions that favour productive ternary complex formation. However, PROTAC **158b**, containing a 3,5-disubstituted pyridyl linker, and PROTAC **191**, containing a 2,4-disubstituted pyridyl linker, exhibit lower pDC₅₀'s of 8.27 and 8.26 respectively. This may suggest that the pyridyl nitrogen of PROTAC **159b** is optimally positioned to facilitate these bonding interactions. Whilst the pyridyl nitrogen of PROTACs **158b** and **191** actually form more unfavourable bonding interactions than the phenyl-linked PROTAC **152b**, which exhibits a higher pDC₅₀, in spite of exhibiting a lower PΔC value. However, the differences in pDC₅₀ may also be due to the ability of the linker to occupy a conformation that facilitates productive ternary complex formation and subsequent ubiquitination of the Brd4 protein.

PROTAC **197**, which contains a 2,5-substituted furan linker, exhibited a pDC₅₀ of 8.57. The furan oxygen is a weaker HBA than the pyridyl nitrogen and thus may make weaker additional bonding interactions, resulting in a lower pDC₅₀ than PROTAC **159b**, but a higher pDC₅₀ than PROTAC **152b**.²¹⁷ As with the pyridyl linkers, the position of the furan oxygen is pertinent, as PROTAC **198**, containing a 2,4-substituted furan linker, exhibited a lower pDC₅₀ of 7.11. A smaller decrease in pDC₅₀ than expected was observed between PROTAC **197** and **152b** for the removal of an optimally positioned HBA. This may suggest that the more acute angle between the 1,3-disubstituted linker of a 6-membered aromatic ring, PROTAC **159b**, is better able

to accommodate the required geometry for ternary complex formation than that of a 5-membered ring.

An increase in pDC_{50} is observed when comparing the ether-linked PROTAC **143b**, which exhibits a pDC_{50} of 7.99, with PROTAC **197**. This may be due to an increase in the rigidity of the linker, locking it into a favourable conformation for ternary complex formation. This reduces the number of possible conformations, thus increasing the population of the productive ternary complex conformation for Brd4 degradation. PROTAC **144b** containing a 1,3-substituted cyclopentyl linker exhibits a pDC_{50} of < 5 . This may suggest that the aromatic linker is making favourable π -bonding interactions that facilitate ternary complex formation. It may also suggest that a preferred geometry is obtained by the planar aromatic linker.

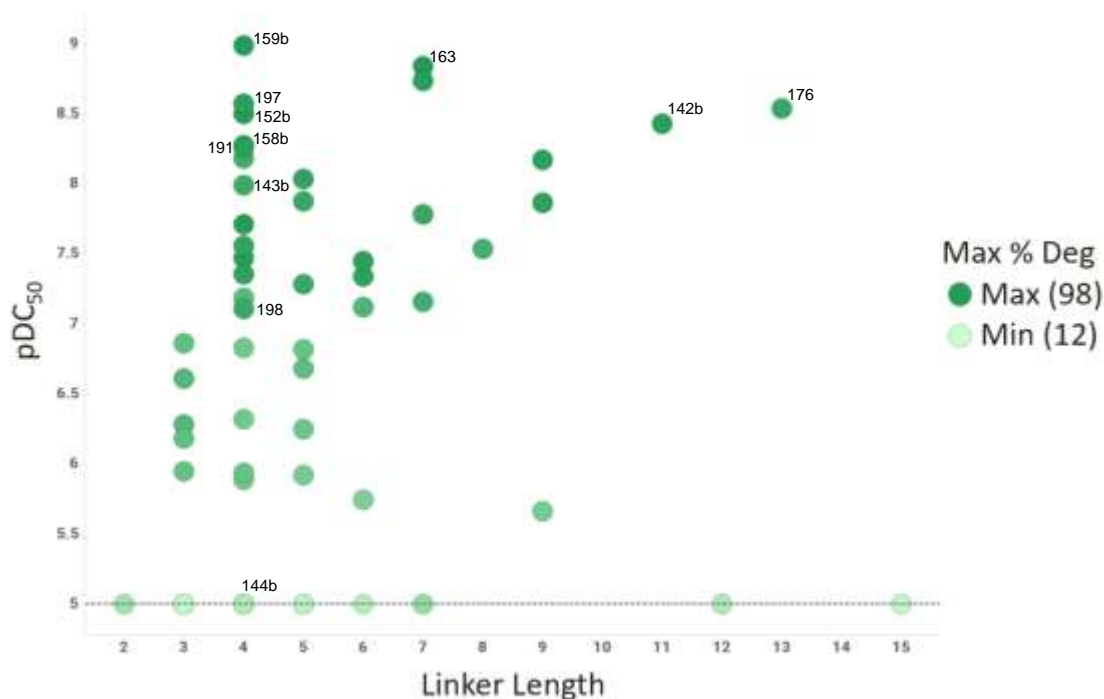


Figure 102: Plot of pDC_{50} vs linker length for the PROTAC series shown in green. The shading of the points corresponds to their max % deg values, with dark green representing a max % deg of 98% and light green representing a max % deg of 12%.

This crude measurement of linker length does not consider the lengths of different bonds or the geometry of the bonds, which may be restricted in spirocycles or aromatic rings. Additionally, PROTACs with a range of linker lengths between 4-13 bonds

exhibit potent Brd4 degradation, with pDC_{50} values > 8.1 and max % deg $> 85\%$. This may suggest that the longer, more flexible linkers are capable of folding back on themselves to facilitate productive ternary complex formation. The propensity of these PROTACs to fold in such a way is exemplified by the lowest energy conformation of PROTAC **176**, containing a 12-carbon linker, calculated using the method described in **Section 5.6.3, Figure 103**. Therefore, a more accurate descriptor for molecular shape is required.

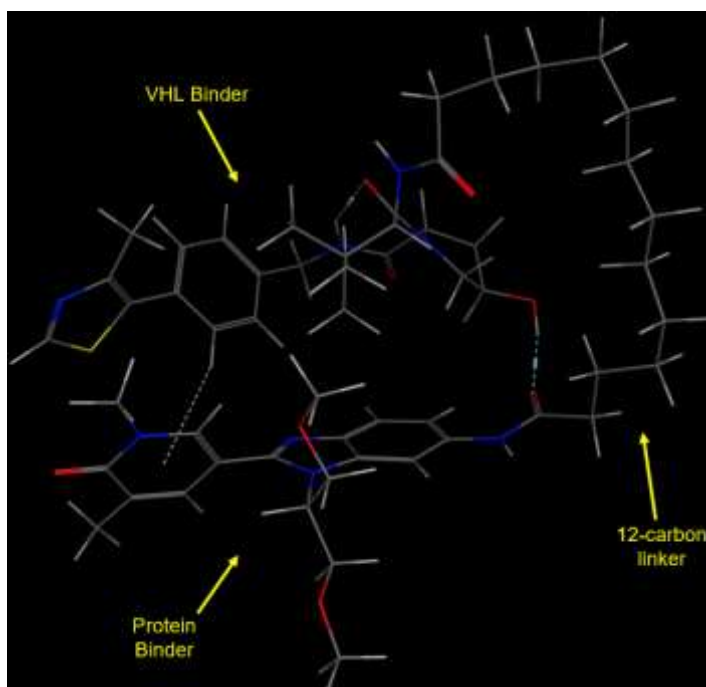


Figure 103: *Lowest energy conformation of PROTAC 176.*

4.7.1 Principal Moments of Inertia (PMI)

In 2003, Saur and co-workers reported the use of principal moments of inertia (PMI) plots to describe three-dimensional molecular shapes in a two-dimensional triangular graph.²¹⁸ This is achieved by calculating two normalised principal moments of inertia ratios (npr), npr1 and npr2. As these are normalised ratios, any effects due to the size of the molecules are eliminated. The corners of the triangular PMI plot represent three geometric objects: rod [0,1], disc [0.5,0.5] and sphere [1,1].²¹⁹ These objects can be represented by acetylene, benzene and adamantane respectively.

The lowest energy conformation of each of the 63 PROTACs in the series was calculated using the method described in **Section 5.6.3**. These conformations are shown in the PMI plot in **Figure 104** and are coloured relative to their pDC_{50} values. The PROTAC series achieves broad coverage of rod to disc-like molecular shape, with few PROTACs probing the sphere-like region in the top right-hand corner of the PMI plot.

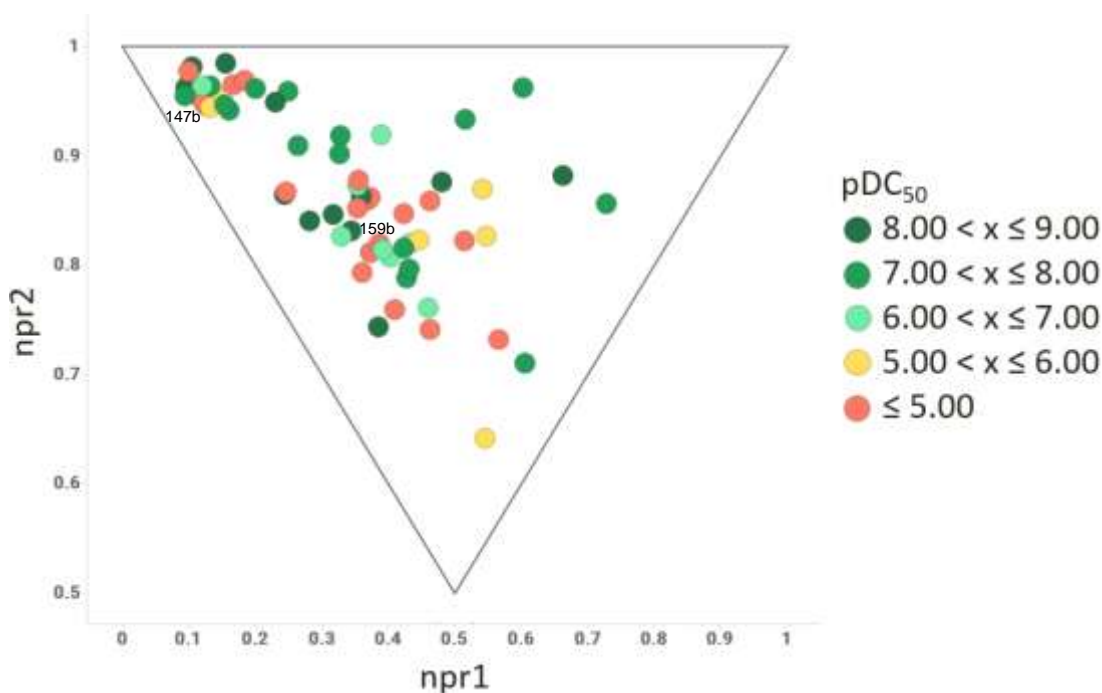


Figure 104: PMI plot of the lowest energy conformation of each PROTAC in the PROTAC series. The points are coloured relative to their pDC_{50} values using the key in the figure.

PROTAC **159b**, with the 2,6-disubstituted pyridyl linker, exhibits the highest pDC_{50} of 8.99 and occupies an elongated discoid shape in its lowest energy conformation. Whilst PROTAC **147b**, with the short rigid bicyclo[1.1.1]pentane linker, exhibits the lowest max % deg of 12% and occupies a rod-like shape in its lowest energy conformation.

There is no observable trend between three-dimensional molecular shape and pDC_{50} when evaluating only the lowest energy conformations of the PROTAC series. However, the lowest energy conformations generally contain favourable bonding interactions between the protein and VHL E3 ligase-binding moieties. During ternary

complex formation, the majority of the protein-binding moiety will be masked inside the Brd4 binding pocket. Therefore, it is likely that the molecular shape of the PROTAC in the productive ternary complex is not the lowest energy conformation.

All accessible conformations within 7 kcal/mol of the lowest energy conformation of each PROTAC were calculated using the method described in **Section 5.6.3**, using a root mean square deviation (RMSD) ≥ 0.25 Å to differentiate distinct conformations. Conformations with an energy > 7 kcal/mol were omitted from the analysis as they represent a relative population $< 10^{-5}$ at 298 K.²²⁰ Approximately 9500 distinct conformations were generated for the entire PROTAC series, **Figure 105**.

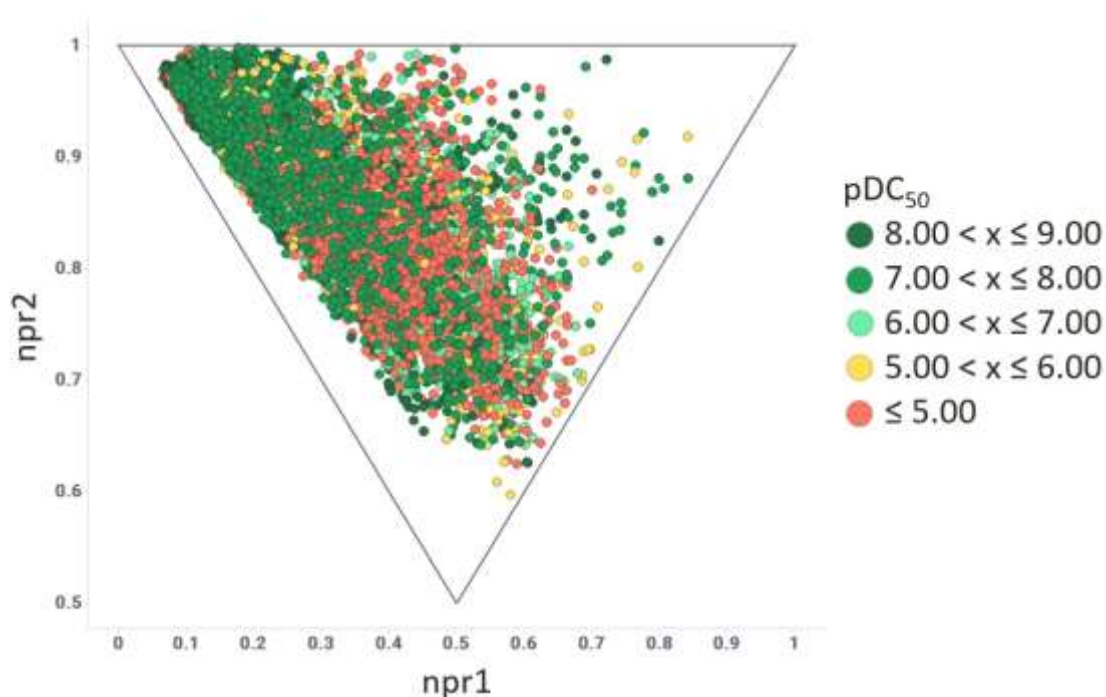


Figure 105: PMI plot of all accessible conformations of each PROTAC in the PROTAC series that are within 7 kcal/mol of its lowest energy conformation. The points are coloured relative to their pDC_{50} values using the key in the figure.

There is a clear preference for conformations with more rod-like molecular shapes. This is consistent with the rigid structure of the pyridone benzimidazole protein binder and the proposed near linear binding conformation of the VHL E3 ligase binder. The coverage of the rod to disc-like axis is predominantly due to conformational folding at the linker. As with the lowest energy conformations, there is little coverage of the

sphere-like molecular shape, which is unlikely to represent a conformation that the PROTACs would occupy during ternary complex formation.

All accessible conformations of PROTACs exhibiting a $pDC_{50} > 8$ are highlighted in **Figure 106**. There is a clear clustering of conformations in the quadrant between $npr1$ 0.2-0.4 and $npr2$ 0.8-0.9, which is occupied by all PROTACs exhibiting a $pDC_{50} > 8$. Additionally, this quadrant becomes more heavily populated as the pDC_{50} of the PROTACs increase.

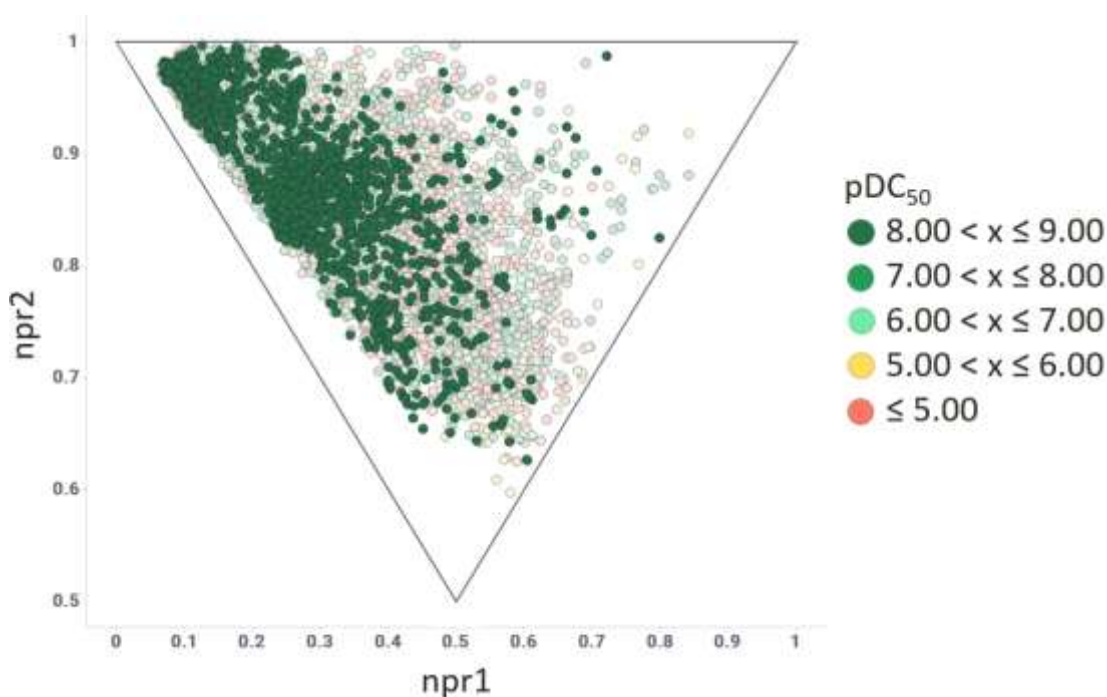


Figure 106: PMI plot of all accessible conformations of each PROTAC in the PROTAC series that are within 7 kcal/mol of its lowest energy conformation. The points are coloured relative to their pDC_{50} values using the key in the figure. The conformations of PROTACs exhibiting a $pDC_{50} > 8.00$ are highlighted.

All accessible conformations of PROTACs exhibiting a max % deg < 25% are highlighted in **Figure 107**. These PROTACs provide more sporadic coverage of the triangular graph, with little coverage of the quadrant between $npr1$ 0.2-0.4 and $npr2$ 0.8-0.9. Additionally, conformations that are in this quadrant are only accessible at energies > 5 kcal/mol.

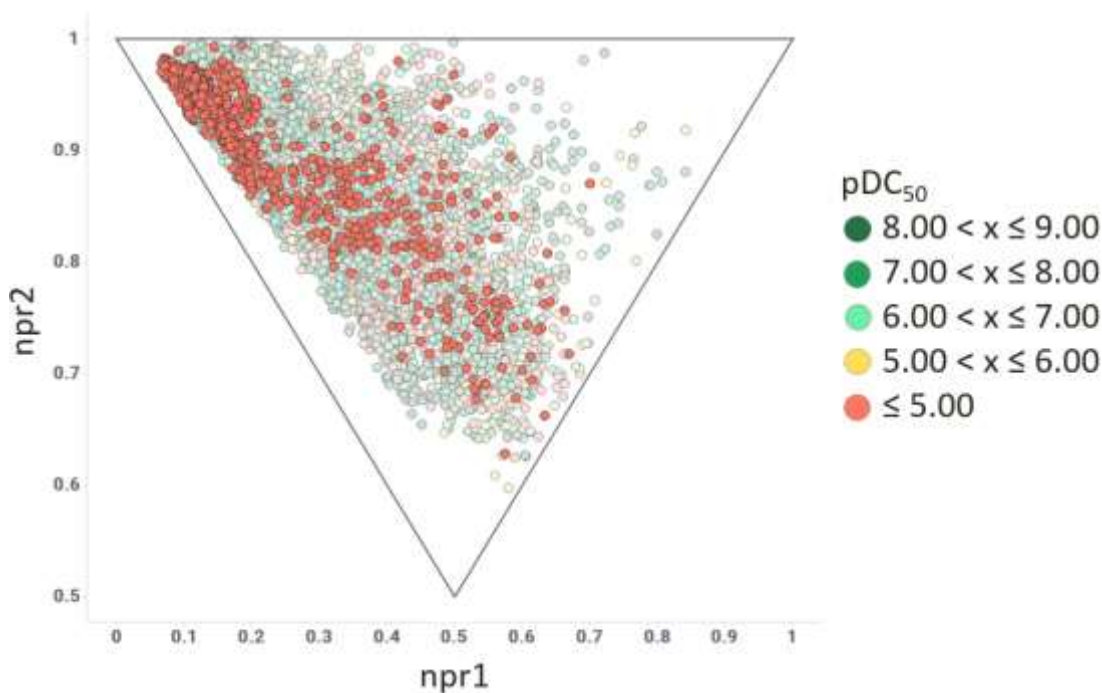


Figure 107: PMI plot of all accessible conformations of each PROTAC in the PROTAC series that are within 7 kcal/mol of its lowest energy conformation. The points are coloured relative to their pDC_{50} values using the key in the figure. The conformations of PROTACs exhibiting a max % deg < 25% are highlighted.

The lowest energy conformations of each PROTAC were truncated at positions A and B in **Figure 108** to furnish LT1 and LT2. The nprs were recalculated for the linker truncate PROTAC series and are visualised in **Figure 109**.

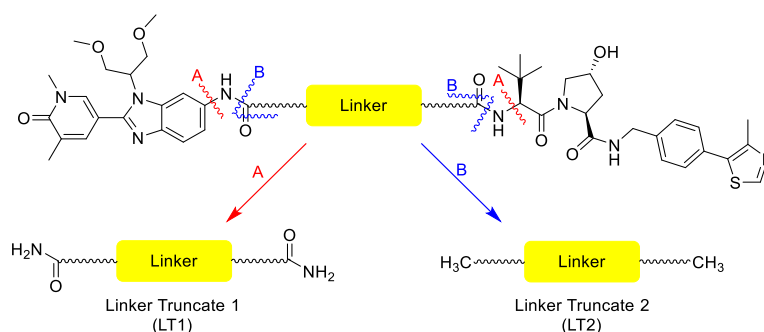


Figure 108: Truncation of the full PROTAC to generate LT1 and LT2.

Truncating the linker eliminates the contribution of the protein and VHL E3 ligase-binding moieties, which are common to all PROTACs, to the molecular shape,

thus focusing on the shape of the linker. PROTACs exhibiting a $pDC_{50} > 8$ are highlighted in **Figure 109**.

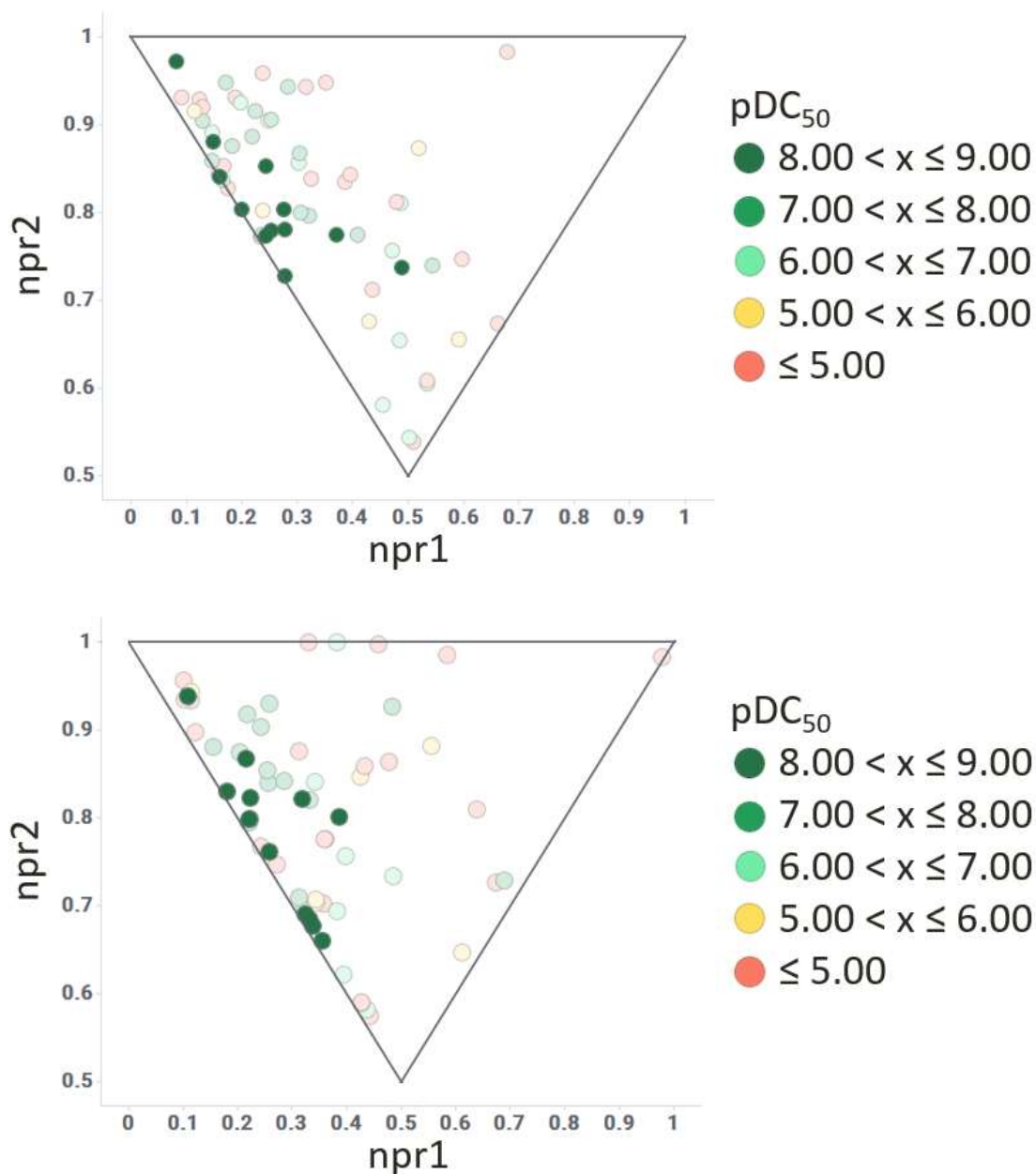


Figure 109: PMI plot of the lowest energy conformation of each PROTAC in the PROTAC series for LT1 (Top) and LT2 (Bottom). The points are coloured relative to their pDC_{50} values using the key in the figure. PROTACs exhibiting a $pDC_{50} > 8.00$ are highlighted.

There is a clear preference for linkers that sit along the rod to disc shape axis with a npr2 of 0.6-1.0 for both LT1 and LT2. These linkers occupy an elongated discoid shape, common to the 1,3-disubstituted 5- and 6-membered aromatic linkers. As these

conformations represent the lowest energy conformations only, this suggests that preorganisation of the linker and FP into these favourable conformations is desirable for potent Brd4 degradation.

The PMI plot of all of the accessible conformations of each PROTAC in the PROTAC series that are within 7 kcal/mol of its lowest energy conformation is reproduced in **Figure 110**. The points are coloured relative to their measured HWB pIC_{50} values using the key in the figure.

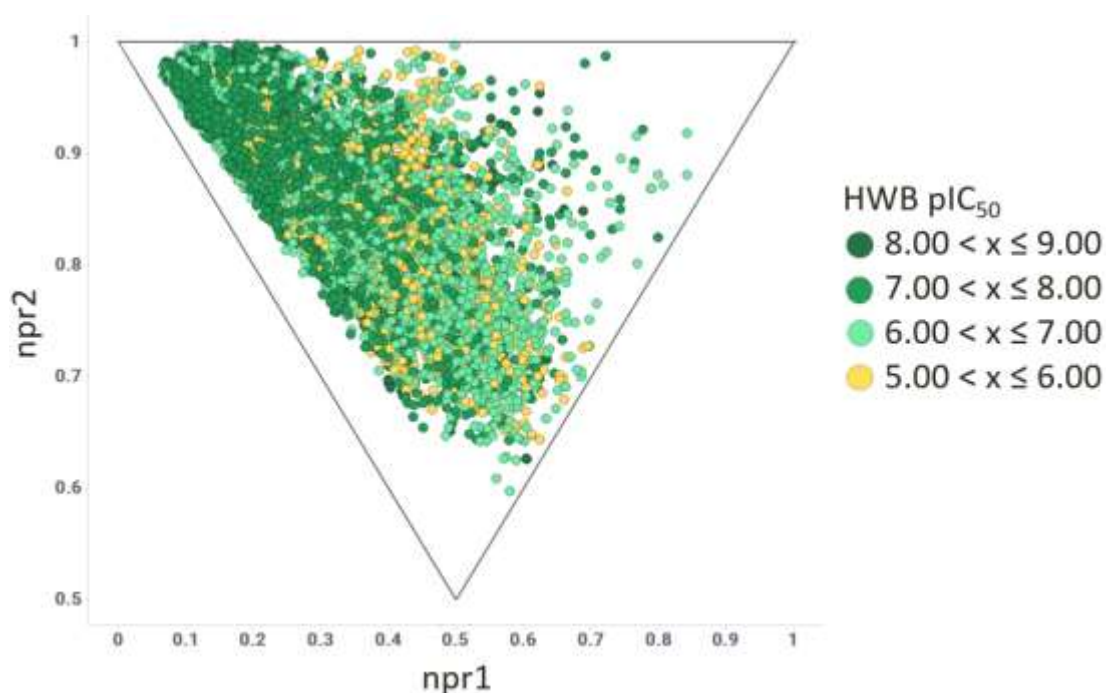


Figure 110: PMI plot of all accessible conformations of each PROTAC in the PROTAC series that are within 7 kcal/mol of its lowest energy conformation. The points are coloured relative to their HWB pIC_{50} values using the key in the figure.

The HWB pIC_{50} shows a similar correlation with molecular shape as pDC_{50} . This correlation is somewhat expected as both assays use similar concentrations of test compound with 18 h incubation. Additionally, as both measurements are taken from *in vitro* cell-based assays, they are affected by similar factors such as solubility and permeability. However, the PROTACs exhibiting a HWB $pIC_{50} < 6$ show minimal coverage of the top left-hand corner of the triangular plot, which is well populated by PROTACs exhibiting higher HWB pIC_{50} values, **Figure 111**. This may suggest that

PROTACs that can occupy a rod-like conformation have a higher cell permeability and thus exhibit a higher HWB potency.

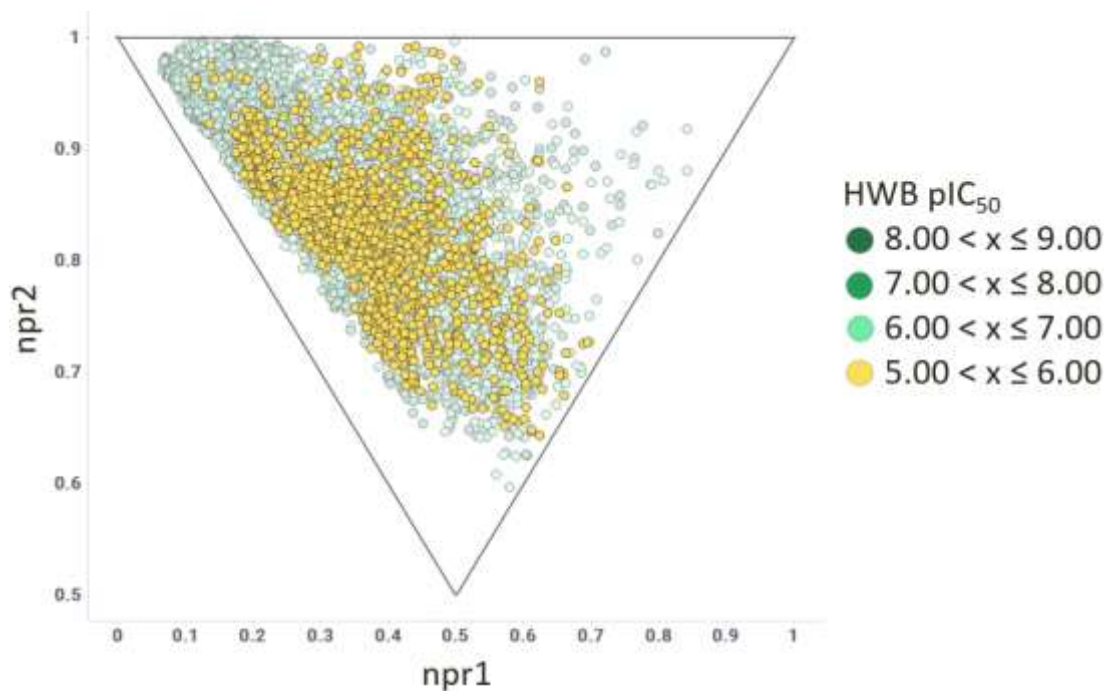


Figure 111: PMI plot of all accessible conformations of each PROTAC in the PROTAC series that are within 7 kcal/mol of its lowest energy conformation. The points are coloured relative to their HWB pIC_{50} values using the key in the figure. PROTACs exhibiting a HWB $pIC_{50} < 6$ are highlighted.

The molecular shape of the PROTAC is not the only factor that contributes to the Brd4 degradation potency. This is highlighted when considering the PMI plots of all accessible conformations of the PROTACs **142b**, **176**, **210** and **211**, **Figure 112**. These PROTACs are of similar length, containing the 10-carbon, 12-carbon, PEG3 and PEG4 linkers respectively. They also occupy similar molecular space in the centre of the PMI plot. However, PROTACs **142b** and **176** exhibit pDC_{50} values > 8 , whilst PROTACs **210** and **211** exhibit pDC_{50} values < 5 . This suggests that the unfunctionalised all-carbon linkers make more favourable binding interactions with both the Brd4 protein and VHL E3 ligase, favouring productive ternary complex formation and potent Brd4 degradation, whilst the ether moieties of the PEG linkers may be making unfavourable interactions with the Brd4 protein and VHL E3 ligase. Therefore, the composition and characteristics of the linker may also modulate the Brd4 degradation potency. However, the difference in pDC_{50} values may also be due to the

physicochemical properties of the PROTACs that govern intracellular concentration. PROTACs **142b** and **176** exhibit ChromlogD values of 4.83 and 5.55, which are significantly higher than PROTACs **210** and **211**, which exhibit ChromlogD values of 3.2 and 3.24. Additionally, the more lipophilic PROTACs **142b** and **176** exhibit P Δ C values of 1.42 and 2.08 respectively, which is indicative of high permeability. The less lipophilic PROTACs **210** and **211** exhibit much lower P Δ C values of 0.1 and -0.15, which suggests that they exhibit limited permeability. Therefore, the higher pDC₅₀ values observed for PROTACs **142b** and **176** may be due to a higher intracellular concentration of the PROTACs. However, it should be noted that PROTACs **142b** and **176** also exhibit CHI_{IAM} values > 40, in comparison to PROTACs **210** and **211** that exhibit values < 24. This will affect the free concentration of PROTAC in the cell that is capable of engaging the Brd4 protein and VHL E3 ligase.

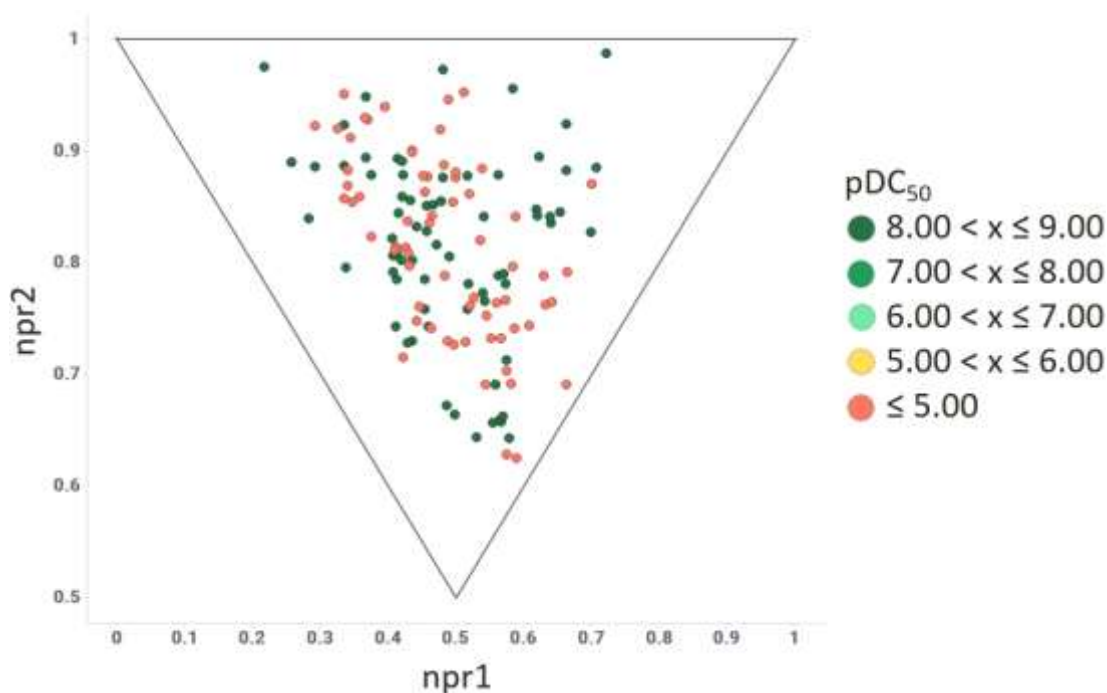


Figure 112: PMI plot of all accessible conformations of PROTACS **142b**, **176**, **210** and **211** that are within 7 kcal/mol of its lowest energy conformation. The points are coloured relative to their pDC₅₀ values using the key in the figure.

There is no observable correlation between ChromlogD and molecular shape, however, there appears to be a correlation between CAD solubility and molecular shape, **Figure 113**. Low solubility PROTACs appear to occupy more rod-like shape,

whilst PROTACs that exhibit a higher solubility tend towards the centre of the PMI plot. This may be due to less efficient packing and intermolecular interactions between PROTACs with more three-dimensional shape, resulting in a lower lattice energy and higher solubility.

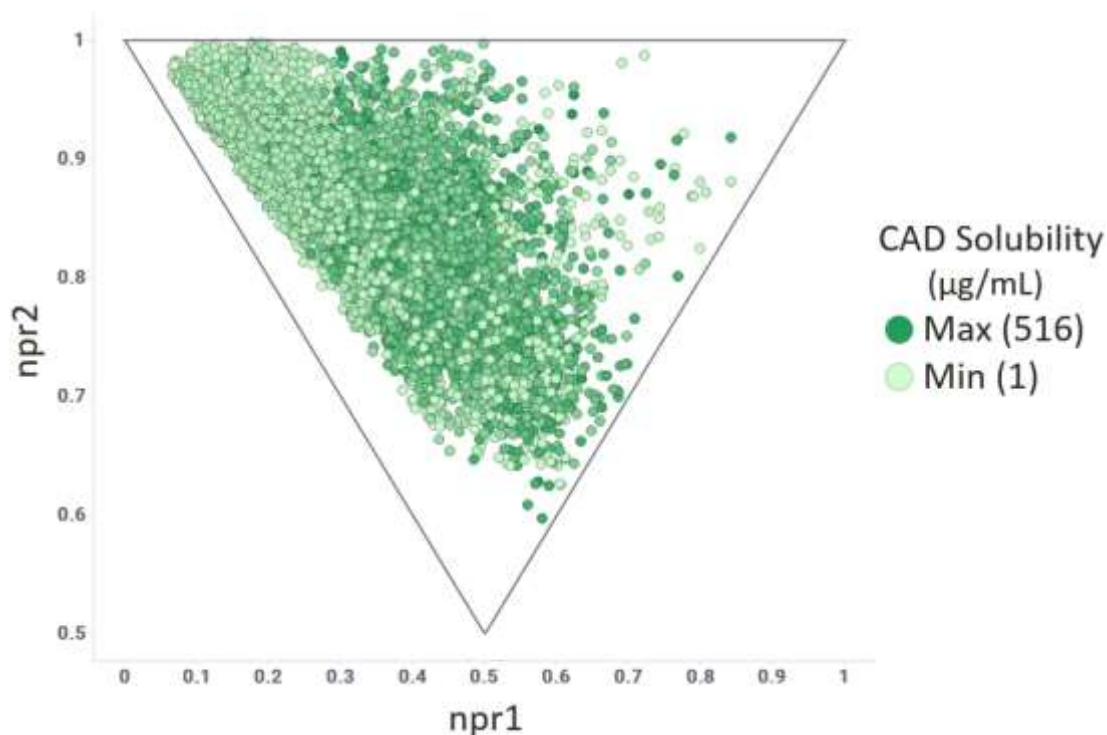


Figure 113: PMI plot of all accessible conformations of each PROTAC in the PROTAC series that are within 7 kcal/mol of its lowest energy conformation. The points are coloured relative to their CAD solubility, with dark green representing a CAD solubility of 516 µg/mL and light green representing a CAD solubility of 1 µg/mL.

The opposite trend appears to be observed for PΔC, with higher permeability PROTACs occupying more rod and elongated discoid-like shapes, **Figure 114**, whilst lower permeability compounds appear to occupy more three-dimensional shape, towards the centre of the PMI plot. This may suggest that a more linear or folded conformation is preferable for permeating through the lipid bilayer, which may involve the use of transmembrane proteins. This may be due to a smaller perceived size of the PROTACs or their ability to mask polar moieties in these conformations.

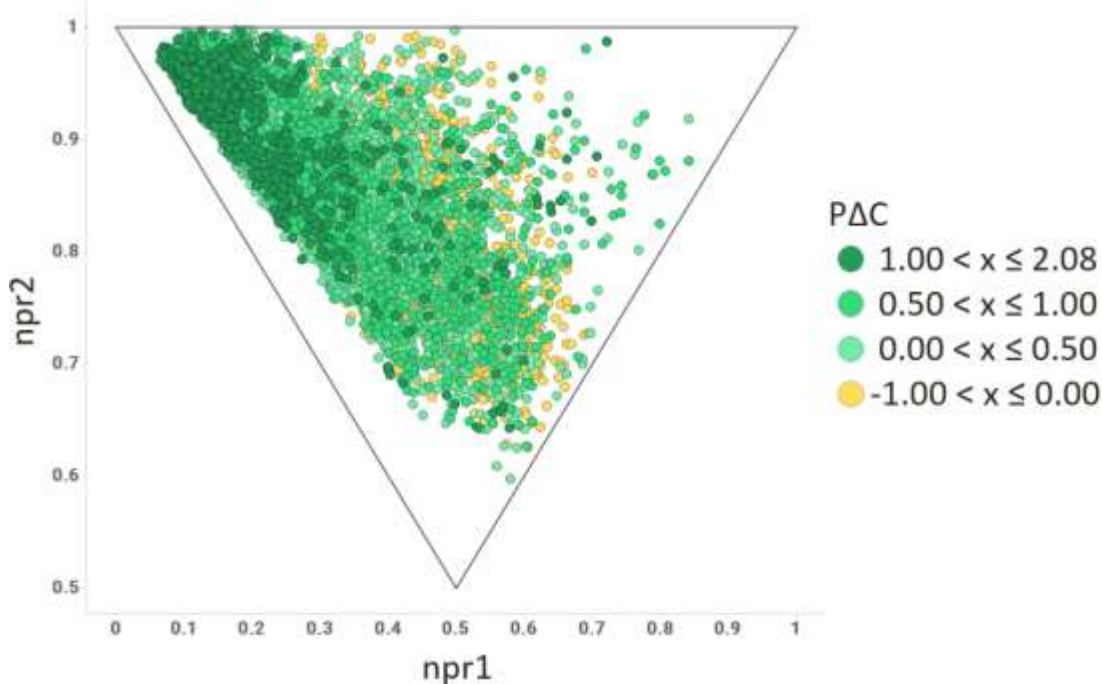


Figure 114: PMI plot of all accessible conformations of each PROTAC in the PROTAC series that are within 7 kcal/mol of its lowest energy conformation. The points are coloured relative to their $P\Delta C$ values using the key in the figure.

The calculated conformations of the PROTACs that are within 7 kcal/mol of the lowest energy conformation achieve broad coverage of rod to disc-like molecular shape, with the most potent degraders occupying elongated discoid shapes in the quadrant between $npr1$ 0.2-0.4 and $npr2$ 0.8-0.9. By analysing the lowest energy conformations of LT1 and LT2, there is a clear trend of linkers that occupy an elongated discoid shape, common to the 1,3-disubstituted 5- and 6-membered aromatic linkers, achieving the highest pDC_{50} values. The molecular shape of the PROTAC is not the only factor that contributes to the Brd4 degradation potency. PROTACs need to be sufficiently permeable and make favourable binding interactions in the ternary complex. Physicochemical properties that have previously shown correlation with the Brd4 degradation profile of the PROTAC series also correlated with molecular shape. Low CAD solubility PROTACs appear to occupy more rod-like shape, whilst PROTACs that exhibit a higher CAD solubility tend towards the centre of the PMI plot. The opposite trend appears to be observed for $P\Delta C$, with higher permeability PROTACs occupying more rod and elongated discoid-like shapes, whilst lower permeability

compounds appear to occupy more three-dimensional shape, towards the centre of the PMI plot. PROTACs that exhibit a low CAD solubility and high PΔC have been shown to exhibit the highest pDC₅₀ values. Therefore, PROTACs that occupy elongated discoid-like shapes are likely to exhibit the most potent Brd4 degradation.

4.8 Conclusions

A linker library containing 66 diverse linkers with terminal carboxylic acid and ester functionality was developed, **Figure 42**. This library provides broad coverage of desirable physicochemical properties defined by the guidelines developed by Edmondson and co-workers, **Figure 48**.²⁰⁶ A series of 63 PROTACs shown in **Figure 54** was rapidly synthesised from the linker library using the high-throughput, one-pot methodologies developed in **Section 3, Scheme 73**. The combination of protein binder **110** and VHL E3 ligase binder **30** has been shown to be capable of eliciting potent degradation of Brd4 protein *via* the UPS *in vitro*. The linker functionality was found to have a profound effect on the degradation profile of the PROTACs, furnishing a 4 log unit range of pDC₅₀ values, with a maximum of 8.99 achieved by PROTAC **159b**, containing the 2,6-disubstituted pyridyl linker. Additionally, the series exhibited a wide range of max % deg values from 12-98%.

The linker functionality also had a significant effect on the physicochemical properties of the PROTACs in the series. Although, only a small range of AMP values was observed, with the majority of the PROTACs exhibiting an AMP of < 3 nm/s, indicative of negligible passive permeability, more than 40% of the PROTAC series exhibited a Brd4 max % deg > 75%, **Figure 63**, with approximately 75% of the PROTAC series also exhibiting a PΔC value > 0, **Figure 89**, which is indicative of good cell permeability. As Brd4 protein is an intracellular target, this suggests that the PROTACs are in fact exhibiting sufficient levels of permeability to achieve potent degradation of Brd4 protein. Therefore, this suggests that the PROTACs are permeating into the cells *via* a mechanism other than passive permeability, which may involve transmembrane proteins. In small molecule drug discovery, AMP is commonly used as the standard measurement of permeability, but for this series of

PROTACs it does not afford a useful measurement for predicting cell permeability in the series.

Changes in the linker functionality of the PROTAC series furnished an approximate 3 log unit range of ChromlogD values from 2.95-5.55, as well as a range of CAD (1-516 $\mu\text{g/mL}$) and FaSSIF (2-1000 $\mu\text{g/mL}$) solubilities, which cover the spectrum from negligibly soluble to highly soluble compounds. Lipophilicity, permeability and solubility are all properties that affect the absorption of molecules *in vivo*. For the PROTAC series, there is a clear positive correlation between ChromlogD and $\text{P}\Delta\text{C}$, which both also exhibit a positive correlation with pDC_{50} . CAD solubility exhibits a negative correlation with ChromlogD, with all PROTACs exhibiting a $\text{pDC}_{50} > 8.1$ also exhibiting a CAD solubility of $\leq 52 \mu\text{g/mL}$. This suggests that lipophilicity is the key property driving the cell permeability, absorption and degradation potency of the PROTACs, whilst the solubility is not a limiting factor, where provided there is appreciable solubility, the PROTAC is able to enter the cell and degrade Brd4 protein.

Furthermore, although a range of TPSA, RBC, Fsp^3 and ARc values were generated for the PROTAC series, these descriptors show no clearly observable correlation with the degradation or measured physicochemical properties of the PROTAC series.

PCA/PLS analysis of the physicochemical descriptors and responses described in **Section 4.6** highlighted the importance of the linker functionality on all of these properties, as the coefficients plot contained an equal number of descriptors for the full PROTAC, linker truncate 1 and linker truncate 2. The loadings plot in **Figure 99** confirmed and quantified the previously described relationship between pDC_{50} , lipophilicity and solubility. The analysis also identified a number of descriptors that correlate with the size of the linker as pertinent in determining the physicochemical properties and degradation profiles of the PROTACs.

PROTACs with a range of linker lengths between 4-13 bonds exhibited potent Brd4 degradation, with pDC_{50} values > 8.1 and max % deg $> 85\%$. However, an optimum linker length of four bonds exhibited by linkers containing 1,3-disubstituted 5- or 6-membered aromatic rings elicited the most potent degradation. Additionally, shorter, rigid linkers with linear conformations elicited the lowest degradation. This suggests

that preorganisation of the PROTAC into a conformation that is conducive to productive ternary complex formation is favourable for driving degradation, although PROTACs with longer linkers are capable of folding to achieve productive ternary complex formation.

The three-dimensional molecular shapes of all accessible conformations of each PROTAC with a relative population $> 10^{-5}$ at 298 K were analysed using a PMI plot. The most potent PROTACs occupy elongated discoid shapes in the quadrant between npr1 0.2-0.4 and npr2 0.8-0.9, common to the 1,3-disubstituted 5- and 6-membered aromatic linkers. Physicochemical properties that have previously shown correlation with the Brd4 degradation profile of the PROTAC series also correlated with molecular shape. Low CAD solubility PROTACs appear to occupy more rod-like shape, whilst PROTACs that exhibit a higher CAD solubility tend towards the centre of the PMI plot. The opposite trend was observed for PΔC, with higher permeability PROTACs occupying more rod and elongated discoid-like shapes, whilst lower permeability compounds appear to occupy more three-dimensional shape, towards the centre of the PMI plot. This suggests that PROTACs that occupy elongated discoid-like shapes are likely to exhibit low CAD solubility, high permeability and elicit the most potent Brd4 degradation. However, the molecular shape of the PROTAC is not the only factor that contributes to the solubility, permeability and degradation potency. The position of functional groups in the linker can have a significant effect on these properties, as it influences the formation of favourable binding interactions in the ternary complex and the exposure of polar moieties. Therefore, crude descriptors such as HBA, HBD and TPSA should be used with caution.

This analysis has highlighted that the guidelines of generic descriptors need to be used with caution when analysing the physicochemical properties of PROTACs, due to the increased complexity of the relationships between them as a result of the effects of molecular shape and conformational folding of these large molecules. It also emphasises the need to screen linkers with a range of geometric properties, and not to discount linkers that will furnish PROTACs with higher ChromlogD values than would be desirable for a classical small molecule drug.

4.9 Future Work

A number of areas have been identified for future study to expand on this work. Firstly, whilst it has not been possible for this PROTAC series thus far, isolation of a ternary complex X-ray crystal structure would allow further understanding of the features of the linker that provide for productive ternary complex formation. With the caveat that this solid phase conformation may not be identical to that which elicits Brd4 degradation in cells, the information about the structure of the ternary complex could direct further SAR studies around linker structure, allowing exploration of how this can increase degradation potency, whilst also aiming to keep the overall properties of the PROTACs in the desired physicochemical space.

The generality of the trends and correlations identified between the physicochemical descriptors and degradation properties of the PROTAC series could be validated by synthesising a selection of PROTAC series from the linker library using different protein-binding and E3 ligase-binding moieties. This could be promptly achieved using the high-throughput, one-pot protocols developed in **Section 3**, for the rapid synthesis of PROTACs in a plate-based format from both acid-ester and diacid linkers. As mentioned in **Section 3.8**, this is currently being investigated by the DHTC department at GSK Stevenage.

Additionally, increasing the scope of the linker library to provide more comprehensive coverage of broader physicochemical space, specifically increasing the range of the HBA and HBD count, would allow these trends and correlations to be identified with a greater confidence. This could potentially allow the formulation of generic guidelines for these properties. Although a plethora of acid-ester and diacid linkers are commercially available, the concept of using bifunctional linkers with terminal functional groups that exhibit orthogonal reactivity under the reaction conditions could be used to expand the scope of this methodology and the linker library to include other functional groups, with varying HBA and HBD counts, along with varying other properties.

As the development of PROTAC technology is still in its infancy, there is a reliance on empiricism in PROTAC design. This efficient high-throughput approach to the rapid synthesis of a large number of PROTACs in a plate-based format could prove to be an invaluable tool for generating vast amounts of valuable data quickly, furthering our understanding of the nuances of successful PROTAC design.

Whilst trends have been identified in the impact of linker functionality on physicochemical properties, further work needs to be conducted to understand how these relate to *in vivo* properties and bioavailability in the PROTAC series, and whether trends can be identified more generally for PROTACs. The acquisition of pharmacokinetic data for the most potent PROTACs of the series by dosing in an animal model would allow the holistic effects of key trends to be assessed. This includes the effect of high lipophilicity on the metabolic clearance, promiscuity and free fraction of PROTAC in both the blood and cell, as well as the effect of low solubility on dosing and absorption. This would allow the true consequences of increasing lipophilicity and decreasing solubility to increase the *in vitro* degradation potency to be evaluated, and how these properties can be balanced to provide adequate bioavailability in oral or parenteral dosing regimens.

The first PROTACs are currently in phase I clinical trials, with the first readouts being eagerly awaited, as they will provide decisive insights into the feasibility of using PROTAC technology as a new therapeutic strategy.

5: Experimental

5.1 General Information

Solvents and reagents were purchased from commercial suppliers and used as received. Anhydrous solvents were used unless otherwise stated. Oven-dried glassware was dried in a vacuum oven at 150 °C when indicated.

5.2 Analytical Techniques and Equipment

5.2.1 Column Chromatography

Automated column chromatography was performed using a Teledyne ISCO CombiFlash[®] R_f with premade RediSep[®] silica cartridges. UV-active fractions were analysed by LCMS and product-containing fractions were combined and the solvent removed *in vacuo* to afford the desired product. Conditions are reported in the order: elution gradient, solvent system and duration (florisil implies solid loading).

5.2.2 Glovebox

An MBraun MB-200B glovebox with an inert N₂ atmosphere was used.

5.2.3 High-Resolution Mass Spectroscopy

High-resolution mass spectroscopy was performed on a micromass Q-ToF Ultima[®] time-of-flight mass spectrometer, and analytes were separated on an Agilent[®] 1100 Liquid Chromatograph equipped with a Phenomenex[®] Luna C18 (2) reverse phase column (100 x 2.1 mm, 3 μm). HRMS analysis utilised liquid chromatography conditions of 0.500 mL min⁻¹ flow rate, 35.0 °C and an injection volume of 2.00-5.00 μL.

The solvents employed were:

A: 0.1% v/v solution of formic acid in water.

B: 0.1% v/v solution of formic acid in MeCN.

Liquid chromatography was performed using the elution gradient shown in **Table 28**.

Table 28: *Elution gradient for HRMS.*

Time / min	Flow rate / mL min ⁻¹	% A	% B
0	1	95	5
6.0	1	0	100
8.5	1	0	100
12.0	1	95	5

Mass-to-charge (m/z) ratios are reported in Daltons.

5.2.4 Infra-red (IR) Spectroscopy

IR spectra were recorded using a Perkin Elmer[®] Spectrum 1 machine. Absorption maxima (ν_{\max}) are reported in wavenumbers (cm^{-1}).

5.2.5 Liquid Chromatography Mass Spectrometry (LCMS)

LCMS analysis was performed on a Waters[®] Acquity UPLC instrument equipped with a BEH column (50 x 2.1 mm, 1.7 μm) and a Waters[®] Micromass ZQ MS using alternate-scan positive and negative electrospray ionisation. Analytes were detected as a summed UV wavelength spectrum between 210-350 nm. Two methods were used:

5.2.5.1 Method A - Formic

0.5 μL injection volume, 40.0 °C, 1.00 mL min⁻¹ flow rate.

The solvents employed were:

A: 0.1% v/v solution of formic acid in water.

B: 0.1% v/v solution of formic acid in MeCN.

Liquid chromatography was performed using the elution gradient shown in **Table 29**.

Table 29: *Elution gradient for LCMS formic method.*

Time / min	Flow rate / mL min ⁻¹	% A	% B
0	1	97	3
1.5	1	5	95
1.9	1	5	95
2.0	1	97	3

5.2.5.1 Method B - HpH

0.3 μL injection volume, 40.0 $^{\circ}\text{C}$, 1.00 mL min^{-1} flow rate.

The solvents employed were:

A: 10 mM ammonium bicarbonate in water adjusted to pH 10 with ammonia solution.

B: MeCN

Liquid chromatography was performed using the elution gradient shown in **Table 30**.

Table 30: *Elution gradient for LCMS HpH method.*

Time / min	Flow rate / mL min^{-1}	% A	% B
0.00	1	97	3
0.05	1	97	3
1.50	1	5	95
1.90	1	5	95
2.00	1	97	3

5.2.6 Mass Directed Auto Purification (MDAP)

MDAP was performed on a Waters FractionLynx system comprising of a Waters 600 pump with extended pump heads, Waters 2700 autosampler, Waters 996 diode array and Gilson 202 fraction collector. HPLC separation was conducted on a XBridge C18 column (100 x 30 mm, 5 μm) at ambient temperature, eluting with 10 mM ammonium bicarbonate in water (pH 10) with MeCN using an appropriate elution gradient determined by LCMS analysis. Mass spectra were recorded on a Waters ZQ mass spectrometer using alternate-scan positive and negative electrospray ionisation with a 150-1500 amu scan range, 0.5 s scan time with an 0.25 s inter-scan delay. The software used was MassLynx 3.5 with FractionLynx 4.1.

5.2.7 Melting Point (m.p.)

Melting points were measured on a Buchi M-565 melting point apparatus.

5.2.8 Microwave Reactor

Reactions performed in a microwave (MW) reactor were performed in a Biotage[®] Initiator⁺ with a maximum power of 300 W and 2.45 GHz.

5.2.9 Nuclear Magnetic Resonance Spectroscopy (NMR)

NMR spectroscopy was performed using a Bruker AV-400 spectrometer ($^1\text{H} = 400 \text{ MHz}$, $^{13}\text{C} = 101 \text{ MHz}$) at $30 \text{ }^\circ\text{C}$ unless otherwise stated.

Chemical shifts (δ) are reported in parts per million (ppm). Peak assignments were chosen based on chemical shifts, integrations and coupling constants, considering 2D analyses where necessary, or the following solvent peaks: CDCl_3 ($^1\text{H} = 7.27 \text{ ppm}$, $^{13}\text{C} = 77.0 \text{ ppm}$), $d_6\text{-DMSO}$ ($^1\text{H} = 2.50 \text{ ppm}$, $^{13}\text{C} = 39.5 \text{ ppm}$) or CD_3OD ($^1\text{H} = 4.87$ and 3.31 ppm , $^{13}\text{C} = 49.1 \text{ ppm}$). Coupling constants (J) are quoted to the nearest 0.1 Hz and refer to 3J coupling unless otherwise stated. Splitting patterns are described as singlet (s), doublet (d), triplet (t), quartet (q), quintet (quin), sextet (sxt), septet (sept), multiplet (m), or broad (br).

5.2.10 Specific Optical Rotation

Specific optical rotation was measured using a JASCO P-1030 polarimeter with a sodium lamp (589 nm) at $20 \text{ }^\circ\text{C}$ and are reported as $[\alpha]_D^{20} = \pm^\circ$ (c concentration, solvent), where a concentration of 1.00 equates to 10 mg/mL .

5.2.11 Thin Layer Chromatography (TLC)

TLC was performed using polyester-backed precoated silica plates (0.200 mm particle size). Spots were visualised under ultraviolet (UV) light of $\lambda_{\text{max}} = 254 \text{ nm}$ or 365 nm . The plates were developed using KMnO_4 or vanillin solutions.

5.3 General Procedures

5.3.1 General Procedure A: Synthesis of imidazopyridine 89 via the GBBR

4-(3,5-Dimethylisoxazol-4-yl)benzaldehyde **63** (100 mg, 0.500 mmol), (6-aminopyridin-3-yl)boronic acid **88** (140 mg, 1.00 mmol) and either EtOH (5.0 mL) or 3:1 DCM:MeOH (3.6/1.2 mL) were added to a microwave vial and sealed. 2-Isocyno-2-methylpropane **65** (0.11 mL, 1.0 mmol) and either HCl (1.25 M in EtOH) (0.040 mL, 0.050 mmol) or Sc(OTf)₃ (24 mg, 0.050 mmol) were added to the reaction vessel and heated at the specified temperature for the specified duration.

5.3.2 General Procedure B: Synthesis of imidazopyridine derivatives via the GBBR

4-(3,5-Dimethylisoxazol-4-yl)benzaldehyde **63** (450 mg, 2.24 mmol), (6-aminopyridin-3-yl)boronic acid **88** (154 mg, 1.12 mmol) and EtOH (11 mL) were added to a microwave vial and sealed. The requisite isocyanide (2.24 mmol) and HCl (1.25 M in EtOH) (0.089 mL, 0.11 mmol) were added to the reaction vessel and heated at 110 °C for 50 min in a MW reactor.

5.3.3 General Procedure C: Amide coupling of VHL E3 ligase binder 30 with acid-ester linkers

Linker (1.1 equiv), HATU (1.1 equiv), DIPEA (4 equiv) and DCM (20 mL/mmol) were added to a round-bottomed flask and sonicated for 30 s. (2*S*,4*R*)-1-((*S*)-2-amino-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide·HCl **30** (1 equiv) was added and the reaction mixture stirred at ambient temperature for 4 h. The solvent was removed *in vacuo* and the crude product was purified by flash column chromatography, eluting with 3:1 EtOAc:EtOH in cyclohexane. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product.

5.3.4 General Procedure D: Amide coupling of VHL-linkers with protein binder 128

LiOH (1.3 equiv) and water (2.6 mL/mmol) were added to a round-bottomed flask and stirred for 5 min. A solution of the VHL-linker in THF (2.6 mL/mmol) was added and the reaction mixture stirred at ambient temperature for 1 h. The solvent was removed *in vacuo* and HATU (1.1 equiv), DIPEA (3 equiv) and DCM or DMF (20 mL/mmol) were added to the round-bottomed flask and sonicated for 30 s. 5-(6-Amino-1-(1,3-dimethoxypropan-2-yl)-1*H*-benzo[*d*]imidazol-2-yl)-1,3-dimethylpyridin-2(1*H*)-one **128** (1 equiv) was added and the reaction mixture stirred at ambient temperature for 16 h. The reaction mixture was diluted with water and extracted with three portions of DCM. The combined organic phases were dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by either MDAP or reverse phase column chromatography, eluting with MeCN in 10 mM ammonium bicarbonate solution (pH 10). UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product.

5.3.5 General Procedure E: One-pot coupling of VHL E3 ligase binder 30, acid-ester linkers and protein binder 128

Linker (1 equiv), HATU (1.1 equiv), DIPEA (4 equiv) and DCM (2.5 mL/mmol) were added to a 4 mL vial and sonicated for 30 s. (2*S*,4*R*)-1-((*S*)-2-amino-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide·HCl **30** (1 equiv) was added and the reaction mixture stirred at ambient temperature for 2 h. The crude reaction mixture was blown down under a stream of N₂ and placed in a vacuum oven at 40 °C for 18 h. The crude product was dissolved in MeOH (2.5 mL/mmol) and added to a stirred solution of LiOH (4 equiv) in water (2.5 mL/mmol). The reaction mixture was heated at 60 °C for 18 h. The crude reaction mixture was blown down under a stream of N₂ and placed in a vacuum oven at 40 °C for 18 h. HATU (1.1 equiv), DIPEA (3 equiv) and DMF (2.5 mL/mmol) were added to the vial and the reaction mixture was sonicated for 30 s. 5-(6-amino-1-(1,3-dimethoxypropan-2-yl)-1*H*-benzo[*d*]imidazol-2-yl)-1,3-dimethylpyridin-2(1*H*)-one

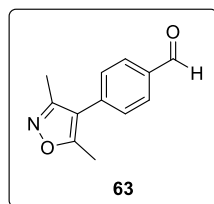
128 (1 equiv) was added and the reaction mixture stirred at ambient temperature for 2 h. HATU (0.5 equiv) and DIPEA (1.5 equiv) were added and the reaction mixture stirred at ambient temperature for a further 2 h. The reaction mixture was diluted with DMSO and purified by MDAP eluting with MeCN in 10 mM ammonium bicarbonate solution (pH 10). UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product.

5.5.6 General Procedure F: One-pot coupling of VHL E3 ligase binder 30, diacid linkers and protein binder 128

Linker (1.2 equiv), HATU (1.1 equiv), DIPEA (4 equiv) and DCM (20 mL/mmol) were added to a 10 mL vial and sonicated for 30 s. (2*S*,4*R*)-1-((*S*)-2-amino-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide·HCl **30** (1 equiv) was added and the reaction mixture stirred at ambient temperature for 30 min. The crude reaction mixture was blown down under a stream of N₂ and placed in a vacuum oven at 40 °C for 18 h. HATU (1.1 equiv), DIPEA (3 equiv) and DMF (2.5 mL/mmol) were added to the vial and the reaction mixture was sonicated for 30 s. 5-(6-amino-1-(1,3-dimethoxypropan-2-yl)-1*H*-benzo[*d*]imidazol-2-yl)-1,3-dimethylpyridin-2(1*H*)-one **128** (1 equiv) was added and the reaction mixture stirred at ambient temperature for 2 h. The reaction mixture was diluted with DMSO and purified by MDAP eluting with MeCN in 10 mM ammonium bicarbonate solution (pH 10). UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product.

5.4 Experimental Details

4-(3,5-Dimethylisoxazol-4-yl)benzaldehyde **63**

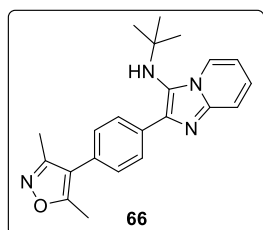


Bromobenzaldehyde **61** (4.6 g, 25 mmol), KOAc (4.9 g, 50 mmol) and PdCl₂ (8.9 mg, 0.050 mmol) were added to a round-bottomed flask. 3,5-Dimethylisoxazole **62** (3.7 mL, 38 mmol) and DMA (125 mL) were added to the reaction vessel and heated at 130 °C for 20 h under an inert N₂ atmosphere. The reaction mixture was diluted with water and extracted with three portions of DCM. The combined organic phases were washed with LiCl solution, dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-30% EtOAc in cyclohexane over 30 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **63** (4.5 g, 22 mmol, 86% yield) as a white solid.

m.p. 73-75 °C [Lit: NA]; **δ_{H}** (400 MHz, CDCl₃) δ ppm 10.06 (s, 1 H), 7.98-7.95 (m, 2 H), 7.46-7.43 (m, 2 H), 2.43 (s, 3 H), 2.31 (s, 3 H); **δ_{C}** (101 MHz, CDCl₃) δ ppm 191.5, 166.0, 158.3, 137.0, 135.4, 130.2, 129.5, 115.9, 11.8, 10.9; **LCMS** (Method B): t_{R} = 0.94, (m/z) [M+H]⁺ not observed (97% purity); **HRMS** (ESI): calculated for C₁₂H₁₂NO₂ (m/z) [M+H]⁺ = 202.0868, found 202.0873; **IR** (ν_{max} /cm⁻¹) 2847, 2751, 1696, 1606, 1566, 1212, 1131, 836.

Spectroscopic data are consistent with the literature.²²¹

N-(*tert*-Butyl)-2-(4-(3,5-dimethylisoxazol-4-yl)phenyl)imidazo[1,2-*a*]pyridin-3-amine **66**



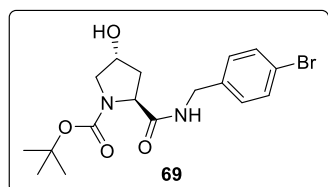
4-(3,5-Dimethylisoxazol-4-yl)benzaldehyde **63** (400 mg, 2.00 mmol), pyridin-2-amine **64** (190 mg, 2.00 mmol) and EtOH (20 mL) were added to a microwave vial and sealed. 2-Isocyano-2-methylpropane **65** (0.45 mL, 4.0 mmol) and HCl (1.25 M in EtOH) (0.16 mL, 0.20 mmol) were added to the reaction vessel and heated at 130 °C for 50 min in a MW reactor. The reaction mixture

was transferred to a round-bottomed flask and the solvent removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-60% EtOAc in cyclohexane over 20 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **66** (540 mg, 1.50 mmol, 75% yield) as a white solid.

m.p. 64-77 °C [Lit: NA]; δ_{H} (400 MHz, CDCl₃) δ ppm 8.24 (d, $J=6.6$ Hz, 1 H), 8.05 (d, $J=8.1$ Hz, 2 H), 7.55 (d, $J=9.0$ Hz, 1 H), 7.33 (d, $J=8.0$ Hz, 2 H), 7.17-7.14 (m, 1 H), 6.81-6.78 (m, 1 H), 3.10 (s (br), 1 H), 2.39 (s, 3 H), 2.25 (s, 3 H), 1.03 (s, 9 H); δ_{C} (101 MHz, CDCl₃) δ ppm 165.1, 158.7, 142.1, 138.9, 134.6, 129.3, 128.8, 128.4, 124.1, 123.5, 123.4, 117.3, 116.5, 111.4, 56.4, 30.3, 11.6, 10.8; **LCMS** (Method B): $t_{\text{R}} = 1.20$, (m/z) [M+H]⁺ 361 (99% purity); **HRMS** (ESI): calculated for C₂₂H₂₅N₄O (m/z) [M+H]⁺ = 361.2028, found 361.2032; **IR** (ν_{max} /cm⁻¹) 3302, 2968, 2926, 1630, 1504, 1363, 1239, 755.

Spectroscopic data are consistent with the literature.¹³⁷

tert*-Butyl (2*S*,4*R*)-2-((4-bromobenzyl)carbamoyl)-4-hydroxypyrrolidine-1-carboxylate **69*

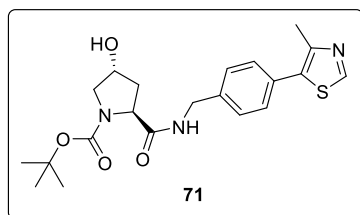


(2*S*,4*R*)-1-(*tert*-Butoxycarbonyl)-4-hydroxypyrrolidine-2-carboxylic acid **67** (400 mg, 1.73 mmol) and DMF (2 mL) were added to a microwave vial and cooled to 0 °C. DIPEA (0.90 mL, 5.2 mmol) and HATU (723 mg, 1.90 mmol) were added and the reaction mixture stirred for 15 min. (4-Bromophenyl)methanamine **68** (0.24 mL, 1.9 mmol) was added and the reaction mixture stirred at ambient temperature for 2 h. The reaction mixture was diluted with water and extracted with three portions of EtOAc. The combined organic phases were washed with LiCl solution, dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-100% EtOAc in cyclohexane over 25 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was

removed *in vacuo* to afford the product **69** (542 mg, 1.36 mmol, 78% yield) as a white solid.

m.p. 50-62 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 8.45-8.41 (m, 1 H), 7.50-7.46 (m, 2 H), 7.23-7.22 (m, 2 H), 4.99-4.97 (m, 1 H), 4.30-4.14 (m, 4 H), 3.45-3.37 (m, 1 H), 3.29-3.26 (m, 1 H), 2.08-1.99 (m, 1 H), 1.85-1.81 (m, 1 H), 1.26 (s, 9 H) (Major rotamer reported); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 172.4, 153.6, 139.0, 130.9, 129.5, 119.7, 78.5, 68.0, 58.8, 54.8, 41.4, 38.7, 27.8.; **LCMS** (Method B): $t_{\text{R}} = 0.95$, (m/z) $[\text{M}+\text{H}]^+$ 300 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{17}\text{H}_{24}\text{BrN}_2\text{O}_4$ (m/z) $[\text{M}+\text{H}]^+ = 399.0919$, found 399.0919; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3412, 3302, 3080, 2982, 2941, 1660, 1405, 1159; $[\alpha]_{\text{D}}^{20} = -33$ (c 0.77, MeOH).

tert*-Butyl (2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carboxylate **71*



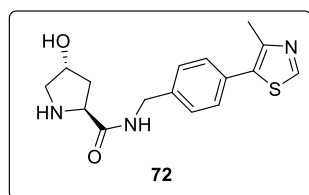
tert-Butyl (2*S*,4*R*)-2-((4-bromobenzyl)carbamoyl)-4-hydroxypyrrolidine-1-carboxylate **69** (2.1 g, 5.2 mmol) and DMF (14 mL) were added to a microwave vial. 4-Methylthiazole **70** (1.4 mL, 16 mmol), KOAc (1.0 g, 10 mmol) and $\text{Pd}(\text{PPh}_3)_4$ (120 mg, 0.104 mmol) were

added and the vial sealed. The reaction mixture was heated at 120 °C for 18 h under an inert N_2 atmosphere. The reaction mixture was diluted with water and extracted with three portions of EtOAc. The combined organic phases were washed with LiCl solution, dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by reverse phase column chromatography eluting with 0-40% MeCN in 10 mM ammonium bicarbonate solution (pH 10) over 25 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **71** (1.32 g, 3.11 mmol, 60% yield) as a white solid.

m.p. 74-91 °C; δ_{H} (400 MHz, CD_3OD) δ ppm 8.87 (s, 1 H), 7.44-7.43 (m, 4 H), 4.52-4.30 (m, 4 H), 3.61-3.49 (m, 2 H), 2.47 (s, 3 H), 2.28-2.22 (m, 1 H), 2.05-2.01 (m, 1 H), 1.33 (s, 9 H) (Major rotamer reported); δ_{C} (101 MHz, CD_3OD) δ ppm 175.7,

156.4, 153.1, 149.6, 140.5, 133.4, 132.1, 130.7, 130.7, 129.8, 129.8, 81.8, 70.2, 61.0, 56.2, 44.0, 41.0, 28.7, 28.7, 28.7, 15.9; **LCMS** (Method B): $t_R = 0.87$, (m/z) $[M+H]^+$ 418 (98% purity); **HRMS** (ESI): calculated for $C_{21}H_{28}N_3O_4S$ (m/z) $[M+H]^+ = 418.1801$, found 418.1794; **IR** (ν_{max}/cm^{-1}) 3298, 3080, 2977, 2923, 1661, 1536, 1403, 1158; $[\alpha]_D^{20} = -29$ (c 0.78, MeOH).

(2S,4R)-4-Hydroxy-N-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide 72

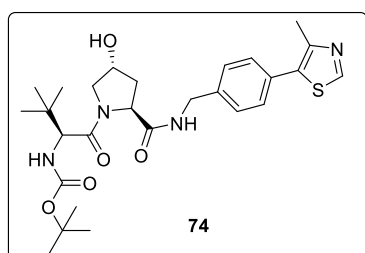


tert-Butyl (2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carboxylate **71** (1.2 g, 2.9 mmol) and DCM (14 mL) were added to a microwave vial. HCl (4 M in dioxane) (3.6 mL, 14 mmol) was added

and the reaction mixture stirred at ambient temperature for 3 h. The solvent was removed *in vacuo*. The crude product was purified by reverse phase column chromatography eluting with 0-30% MeCN in 10 mM ammonium bicarbonate solution (pH 10) over 20 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **72** (0.830 g, 2.56 mmol, 90% yield) as a white solid.

m.p. 112-119 °C; δ_H (400 MHz, CD_3OD) δ ppm 8.87 (s, 1 H), 7.45-7.44 (m, 2 H), 7.39-7.38 (m, 2 H), 4.44 (s, 2 H), 4.37-4.35 (m, 1 H), 3.93 (t, $J=8.3$ Hz, 1 H), 3.03-3.00 (m, 1 H), 2.91-2.88 (m, 1 H), 2.48 (s, 3 H), 2.19-2.15 (m, 1 H), 1.89-1.84 (m, 1 H); δ_C (101 MHz, CD_3OD) δ ppm 177.4, 153.0, 149.3, 140.5, 133.5, 131.9, 130.7, 129.3, 73.7, 61.0, 56.2, 43.6, 41.2, 15.9; **LCMS** (Method B): $t_R = 0.69$, (m/z) $[M+H]^+$ 318 (98% purity); **HRMS** (ESI): calculated for $C_{16}H_{20}N_3O_2S$ (m/z) $[M+H]^+ = 318.1276$, found 318.1269; **IR** (ν_{max}/cm^{-1}) 3273, 3069, 2929, 2870, 2440, 1642, 1481, 1405; $[\alpha]_D^{20} = -16$ (c 0.72, MeOH).

tert*-Butyl((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamate **74*

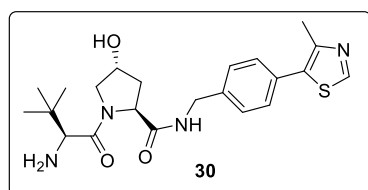


(*S*)-2-((*tert*-Butoxycarbonyl)amino)-3,3-dimethylbutanoic acid **73** (846 mg, 3.66 mmol) and DMF (3.5 mL) were added to a round-bottomed flask and cooled to 0 °C. DIPEA (1.7 mL, 10 mmol) and HATU (1.39 g, 3.66 mmol) were added and the reaction

mixture stirred at ambient temperature for 15 min. (2*S*,4*R*)-4-Hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide **72** (1.06 g, 3.32 mmol) was added and the reaction mixture stirred at ambient temperature for 2 h. The reaction mixture was diluted with water and extracted with three portions of EtOAc. The combined organic phases were washed with LiCl solution, dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by reverse phase column chromatography eluting with 0-60% MeCN in 10 mM ammonium bicarbonate solution (pH 10) over 20 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **74** (1.09 g, 2.05 mmol, 62% yield) as a brown oil.

δ_{H} (400 MHz, CD₃OD) δ ppm 8.86 (m, 1 H), 7.46-7.44 (m, 2 H), 7.42-7.40 (m, 2 H), 4.6 (t, $J=8.2$ Hz, 1 H), 4.54-4.51 (m, 2 H), 4.37-4.34 (m, 1 H), 4.30 (s, 1 H), 3.89-3.87 (m, 1 H), 3.81-3.79 (m, 1 H), 2.47 (s, 3 H), 2.24-2.20 (m, 1 H), 2.12.-2.07 (m, 1 H), 1.44 (s, 9 H), 1.01 (s, 9 H); δ_{C} (101 MHz, CD₃OD) δ ppm 174.5, 173.1, 157.9, 152.9, 149.2, 140.4, 133.5, 131.7, 130.5, 129.1, 80.8, 71.2, 60.9, 60.5, 58.1, 43.6, 39.0, 36.9, 28.9, 27.1, 16.0; **LCMS** (Method B): t_{R} = 1.01, (m/z) [M+H]⁺ 531 (100% purity); **HRMS** (ESI): calculated for C₂₇H₃₉N₄O₅S (m/z) [M+H]⁺ = 531.2641, found 531.2640; **IR** (ν_{max} /cm⁻¹) 3430, 3307, 3080, 2970, 1674, 1629, 1165, 752; [α]_D²⁰ = -28 (*c* 1.2, MeOH).

(2*S*,4*R*)-1-((*S*)-2-Amino-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide **30**

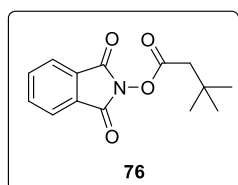


tert-Butyl ((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamate **74**
(1.07 g, 2.02 mmol) and DCM (10 mL) were added to

a round-bottomed flask and stirred at ambient temperature. HCl (4 M in dioxane) (2.52 mL, 10.1 mmol) was added and the reaction mixture stirred at ambient temperature for 6 h. The solvent was removed *in vacuo* and the crude product was purified by reverse phase column chromatography, eluting with 0-40% MeCN in 10 mM ammonium bicarbonate solution (pH 10) over 30 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **30** (533 mg, 1.24 mmol, 62% yield) as a white solid.

m.p. 78-83 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 8.97 (s, 1 H), 8.49 (t, $J=6.1$ Hz, 1 H), 7.44-7.35 (m, 4 H), 5.05 (d, $J=3.7$ Hz, 1 H), 4.44 (t, $J=8.1$ Hz, 1 H), 4.39-4.33 (m, 2 H), 4.26-4.22 (m, 1 H), 3.59-3.56 (m, 2 H), 3.18, (s, 1 H), 2.44 (s, 3 H), 2.04-2.01 (m, 1 H), 1.92-1.88 (m, 1 H), 0.89 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 173.4, 172.0, 151.3, 147.7, 139.5, 131.4, 129.6, 128.6, 127.4, 68.9, 59.2, 58.5, 55.9, 41.6, 37.7, 35.3, 26.1, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.77$, (m/z) $[\text{M}+\text{H}]^+$ 431 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{22}\text{H}_{31}\text{N}_4\text{O}_3\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 431.2115$, found 431.2112; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3297, 3086, 2958, 2877, 1666, 1629, 1556, 1364; $[\alpha]_{\text{D}}^{20} = -31$ (c 0.85, MeOH).

1,3-Dioxoisindolin-2-yl 3,3-dimethylbutanoate **76**



3,3-Dimethylbutanoic acid **75** (1.50 mL, 11.8 mmol) and EtOAc (115 mL) were added to a round-bottomed flask. *N,N'*-Dicyclohexylcarbodiimide (DCC) (3.65 g, 17.7 mmol), DMAP (72 mg, 0.59 mmol) and 2-hydroxyisindoline-1,3-dione

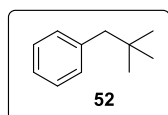
50 (3.19 g, 19.6 mmol) were added to the reaction vessel and stirred at ambient

temperature for 16 h. The reaction mixture was filtered through Celite with EtOAc washings and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-20% EtOAc in cyclohexane over 40 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **76** (3.03 g, 11.8 mmol, 98% yield) as a white solid.

m.p. 90-95 °C [Lit: 92-93 °C]¹⁵⁷; δ_{H} (400 MHz, *d*₆-DMSO) δ ppm 7.96 (m, 4 H), 2.61 (s, 2 H), 1.13 (s, 9 H); δ_{C} (101 MHz, *d*₆-DMSO) δ ppm 167.9, 161.7, 135.3, 127.9, 123.7, 43.2, 30.7, 28.8; **LCMS** (Method B): $t_{\text{R}} = 1.25$, (*m/z*) [M+H]⁺ not observed (75% purity due to sample degradation in LCMS, 1 spot by TLC).

Spectroscopic and physical data are consistent with the literature.¹⁵⁷

Neopentylbenzene **52**



A range of zinc sources was screened for the synthesis of **52** using the described procedure (*vide infra*) and the conditions visualised in **Table 31**.

The reactions were set up using oven-dried glassware in a glovebox under an inert N₂ atmosphere. NiBr₂·diglyme (18 mg, 0.050 mmol), dtbbpy (13 mg, 0.050 mmol) and DMA (0.31 mL) were added to a microwave vial and stirred vigorously for 15 min to furnish solution A. Iodobenzene **77** (0.028 mL, 0.30 mmol), 1,3-dioxoisindolin-2-yl 3,3-dimethylbutanoate **76** (98 mg, 0.38 mmol) and zinc (33 mg, 0.50 mmol) were added to a separate microwave vial. Solution A (1 mL) was added to the vial, which was sealed and removed from the glovebox. The reaction mixture was stirred at 28 °C for 16 h and monitored by LCMS.

Table 31: Zinc sources screened for the decarboxylative cross-coupling of RAE **76** with aryl iodide **77**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Zinc Source	Particle Size	52 / %
1	Sigma Aldrich	40-60 nm	2
2	Sigma Aldrich	< 10 μ m	3
3	Acros Organics	Dust/Not Specified	2
4	Sigma Aldrich	Powder/Not Specified	0

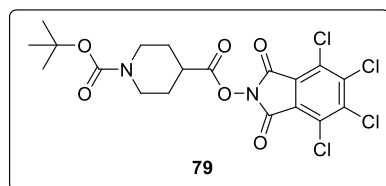
A range of nickel catalysts was screened for the synthesis of **52** using the described procedure (*vide infra*) and the conditions visualised in **Table 32**.

The reactions were set up using oven-dried glassware in a glovebox under an inert N₂ atmosphere. Nickel catalyst (0.050 mmol), dtbbpy (13 mg, 0.050 mmol) and DMA (0.31 mL) were added to a microwave vial and stirred vigorously for 15 min to furnish solution A. Iodobenzene **77** (0.028 mL, 0.25 mmol), 1,3-dioxoisindolin-2-yl 3,3-dimethylbutanoate **76** (98 mg, 0.38 mmol) and zinc source (33 mg, 0.50 mmol) were added to a separate microwave vial. Solution A (1 mL) was added to the vial, which was sealed and removed from the glovebox. The reaction mixture was stirred at 28 °C for 16 h and monitored by LCMS.

Table 32: Nickel sources screened for the decarboxylative cross-coupling of RAE **76** with aryl iodide **77**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Nickel Catalyst	Nickel Catalyst / mg	52 / %
1	NiBr ₂ ·(diglyme)	18	3
2	NiBr ₂	11	0
3	NiBr ₂ ·xH ₂ O (x=1-2)	13	0
4	NiCl ₂ ·6H ₂ O	12	0

1-(*tert*-Butyl) 4-(4,5,6,7-tetrachloro-1,3-dioxoisindolin-2-yl) piperidine-1,4-dicarboxylate **79**



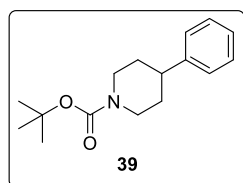
1-(*tert*-Butoxycarbonyl)piperidine-4-carboxylic acid **78** (495 mg, 2.16 mmol) and DCM (21 mL) were added to a round-bottomed flask. DIC (0.370 mL, 2.38 mmol), DMAP (26 mg, 0.22 mmol), and

4,5,6,7-tetrachloro-2-hydroxyisoindoline-1,3-dione **35** (650 mg, 2.16 mmol), were added to the reaction vessel and stirred at ambient temperature for 12 h. The reaction mixture was filtered through Celite with DCM washings and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-25% EtOAc in cyclohexane over 30 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **79** (807 mg, 1.50 mmol, 70% yield) as a white solid.

m.p. 182-190 °C [Lit: 180-183 °C]¹⁵⁵; δ_{H} (400 MHz, CDCl₃) δ ppm 4.07-3.96 (m, 2 H), 3.04-3.01 (m, 2 H), 2.94-2.89 (m, 1 H), 2.06-2.03 (m, 2 H), 1.88-1.82 (m, 2 H), 1.47 (s, 9 H); δ_{C} (101 MHz, CDCl₃) δ ppm 170.2, 157.5, 154.5, 141.1, 130.5, 124.7, 79.9, 42.4, 38.5, 28.4, 27.7; **LCMS** (Method B): $t_{\text{R}} = 0.84$, (m/z) [M+H]⁺ not observed (64% purity due to sample degradation in LCMS, 1 spot by TLC).

Spectroscopic and physical data are consistent with the literature.¹⁵⁵

tert*-Butyl 4-phenylpiperidine-1-carboxylate **39*



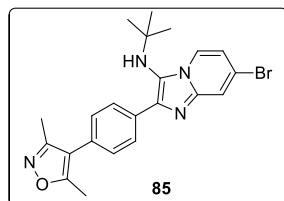
NiCl₂·6H₂O (9.5 mg, 0.040 mmol) and dtbbpy (11 mg, 0.040 mmol) were added to a microwave vial and sealed. The vial was evacuated and backfilled with N₂ three times. DMF (0.8 mL) was added and the reaction mixture stirred at ambient temperature for 3 h to furnish solution A. 1-(*tert*-Butyl 4-(4,5,6,7-tetrachloro-1,3-dioxoisoindolin-2-yl) piperidine-1,4-dicarboxylate **79** (0.10 g, 0.20 mmol) and phenylboronic acid **80** (73 mg, 0.60 mmol) were added to a microwave vial and sealed. The vial was evacuated and backfilled with N₂ three times. Dioxane (8 mL) was added and the reaction mixture stirred at ambient temperature for 1 min. Triethylamine (0.28 mL, 2.0 mmol) was added and the reaction mixture stirred for a further 5 min. Solution A was added and the reaction mixture immediately placed in a preheated oil bath at 75 °C for 12 h. The reaction solution was diluted with water and extracted with three portions of EtOAc. The combined organic phases were washed with LiCl solution, dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-10%

EtOAc in cyclohexane over 30 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **39** (45 mg, 0.17 mmol, 86% yield) as a colourless oil.

δ_{H} (400 MHz, CDCl_3) δ ppm 7.38-7.25 (m, 5 H), 4.37-4.26 (m, 2 H), 2.92-2.81 (m, 2 H), 2.74-2.66 (m, 1 H), 1.90-1.87 (m, 2 H), 1.74-1.66 (m, 2 H), 1.55 (s, 9 H); δ_{C} (101 MHz, CDCl_3) δ ppm 154.9, 145.8, 128.5, 126.8, 126.3, 79.4, 44.5, 42.7, 33.2, 28.5; **LCMS** (Method B): $t_{\text{R}} = 1.37$, (m/z) $[\text{M}+\text{H}]^+$ not observed (100% purity).

Spectroscopic data are consistent with the literature.¹⁵⁵

7-Bromo-*N*-(*tert*-butyl)-2-(4-(3,5-dimethylisoxazol-4-yl)phenyl)imidazo[1,2-*a*]pyridin-3-amine **85**



4-(3,5-Dimethylisoxazol-4-yl)benzaldehyde **63** (400 mg, 2.00 mmol), 4-bromopyridin-2-amine **84** (172 mg, 1.00 mmol) and EtOH (10 mL) were added to a microwave vial and sealed. 2-Isocyno-2-methylpropane **65** (0.225 mL, 2.0 mmol) and HCl (1.25 M in EtOH) (0.080 mL, 0.010 mmol) were added to the reaction vessel and heated at 130 °C for 50 min in a MW reactor. The reaction mixture was transferred to a round-bottomed flask and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-20% EtOAc in cyclohexane over 30 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* and the crude product triturated with cyclohexane (0.05 mL/mg) and filtered through a Buchner funnel. The solid product was dissolved in chloroform and the solvent removed *in vacuo* to afford the product **85** (289 mg, 0.651 mmol, 66% yield) as a white solid.

m.p. 197-200 °C; δ_{H} (400 MHz, CDCl_3) δ ppm 8.11-8.09 (m, 1 H), 8.00-7.97 (m, 2 H), 7.72-7.71 (m, 1 H), 7.34-7.31 (m, 2 H), 6.89 (dd, $J=6.9$ and 1.5 Hz, 1 H), 2.44 (s, 3 H), 2.30 (s, 3 H), 1.07 (s, 9 H); δ_{C} (101 MHz, CDCl_3) δ ppm 165.2, 158.7, 142.1, 139.6, 134.1, 129.7, 129.0, 128.4, 123.9, 123.8, 119.5, 118.0, 116.5, 115.4, 56.6, 30.4, 11.7, 10.9; **LCMS** (Method B): $t_{\text{R}} = 1.35$, (m/z) $[\text{M}+\text{H}]^+$ 439 (94% purity); **HRMS** (ESI):

calculated for $C_{22}H_{24}BrN_4O$ (m/z) $[M+H]^+ = 439.1133$, found 439.1135; IR (ν_{max}/cm^{-1}) 3342, 3068, 2973, 2926, 2868, 1617, 1193, 805.

A range of solvents, visualised in **Table 33**, was screened for the trituration of **85** using the described procedure (*vide infra*).

To approximately 10 mg of **85** synthesised using the described procedure (*vide supra*), was added 0.5 mL of solvent. The mixture was sonicated for 5 s and filtered through cotton wool in a pipette. The filtrate was isolated and the remaining solid dissolved in chloroform and isolated. All samples were blown down under a stream of N_2 for 24 h. Both samples were analysed by LCMS and the isolated yields of pure product reported in **Table 33**.

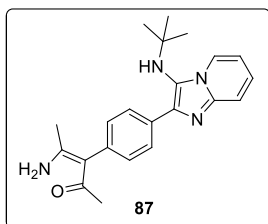
Table 33: Solvents screened for the trituration of bromide **85**.

Entry	Solvent	85 Isolated Yield / %
1	THF	0
2	TBME	0
3	MeOH	15
4	EtOH	25
5	IPA	48
6	Et ₂ O	75
7	<i>i</i> Pr ₂ O	91
8	Pet. Ether	92
9	Cyclohexane	95

A range of reaction durations and equivalents of **63** and **84**, visualised in **Table 34**, were screened for the synthesis of **85** using the described procedure (*vide supra*).

Table 34: Reaction duration and equivalents of **63** and **84** screened for the synthesis of bromide **85**.

Entry	Duration / min	63 / mmol	84 / mmol	85 Isolated Yield / %
1	50	1.0	1.0	36
2	120	1.0	1.0	48
3	120	2.0	1.0	64
4	120	1.0	2.0	66

(Z)-4-amino-3-(4-(3-(*tert*-butylamino)imidazo [1,2-a]pyridin-2-yl)phenyl)pent-3-en-2-one 87

7-Bromo-*N*-(*tert*-butyl)-2-(4-(3,5-dimethylisoxazol-4-yl)phenyl)imidazo[1,2-a]pyridin-3-amine **85** (100 mg, 0.230 mmol), hypodiboric acid **86** (22 mg, 0.25 mmol), PdCl₂(dppf)·DCM (9.3 mg, 0.011 mmol), KOAc (63 mg, 0.64 mmol) and dioxane (1.8 mL) were added to a microwave

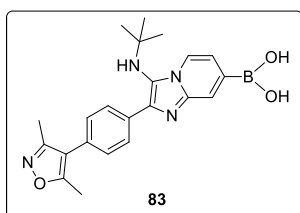
vial and sealed. The reaction mixture was heated at 90 °C for 22 h under an inert N₂ atmosphere. The reaction mixture was diluted with water and extracted with three portions of EtOAc. The combined organic phases were dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-100% EtOAc in cyclohexane over 20 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **87** (8.3 mg, 0.023 mmol, 10% yield) as a colourless oil.

87

δ_H (400 MHz, *d*₆-DMSO) δ ppm 10.39 (s (br), 1 H), 8.39 (d, *J*=7.0 Hz, 1 H), 8.07 (d, *J*=8.1 Hz, 2 H), 7.56 (s (br), 1 H), 7.45 (d, *J*=8.8 Hz, 1 H), 7.18-7.16 (m, 3 H), 6.87 (t, *J*=6.8 Hz, 1 H), 4.58 (s, 1 H), 1.72 (s, 3 H), 1.66 (s, 3 H), 0.98 (s, 9 H); **δ_C** (101 MHz, *d*₆-DMSO) δ ppm 193.4, 160.6, 141.0, 139.3, 138.2, 133.4, 131.4, 131.4, 127.8, 124.0, 123.7, 116.5, 110.9, 108.2, 55.7, 29.9, 28.7, 20.8; **LCMS** (Method B): *t_R* = 0.63, (*m/z*) [M+H]⁺ 363 (99% purity); **HRMS** (ESI): calculated for C₂₂H₂₇N₄O (*m/z*) [M+H]⁺ = 363.2185, found 363.2184; **IR** (ν_{max}/cm⁻¹) 3302, 2964, 1602, 1473, 1534, 1217, 1025, 748.

Note: **δ_N** (60.8 MHz, *d*₆-DMSO) δ ppm 104 (360 predicted for isoxazole).

(3-(*tert*-Butylamino)-2-(4-(3,5-dimethylisoxazol-4-yl)phenyl)imidazo[1,2-a]pyridin-7-yl)boronic acid **83**



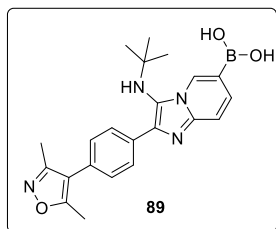
7-Bromo-*N*-(*tert*-butyl)-2-(4-(3,5-dimethylisoxazol-4-yl)phenyl)imidazo[1,2-a]pyridin-3-amine **85** (100 mg, 0.230 mmol), B_2Pin_2 (63 mg, 0.25 mmol), $PdCl_2(dppf) \cdot DCM$ (9.3 mg, 0.011 mmol), KOAc (63 mg, 0.64 mmol) and dioxane (1.8 mL) were added to a microwave vial and sealed. The reaction mixture was heated at 90 °C for 22 h under an inert N_2 atmosphere. The reaction mixture was diluted with water and extracted with three portions of EtOAc. The combined organic phases were dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by reverse phase column chromatography eluting with 0-50% MeCN in 10 mM ammonium bicarbonate solution (pH 10) over 20 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **83** (41 mg, 0.10 mmol, 43% yield) as a yellow oil.

83

δ_H (400 MHz, d_6 -DMSO) δ ppm 8.53 (d, $J=5.9$ Hz, 1 H), 8.09-8.07 (m, 2 H), 7.82 (s, 1 H), 7.54-7.50 (m, 3 H), 2.46 (s, 3 H), 2.30 (s, 3 H), 1.09 (s, 9 H); **LCMS** (Method A): $t_R = 0.73$, (m/z) $[M+H]^+$ 405 (97% purity).

Note: Sample degraded before full characterisation could be obtained.

(3-(*tert*-Butylamino)-2-(4-(3,5-dimethylisoxazol-4-yl)phenyl)imidazo[1,2-a]pyridin-6-yl)boronic acid **89**



A range of conditions using both Bronsted acid and Lewis acid catalysis was screened for the synthesis of **89** using **general procedure A** and the conditions described in **Table 35**. The reaction mixtures were analysed by LCMS.

Table 35: Conditions screened for the synthesis of boronic acid **89**, **Scheme 26**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

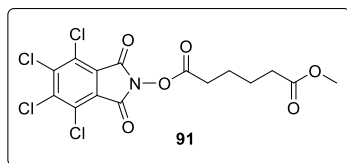
Entry	Duration / min	Temperature / °C	Catalyst	Solvent	89 / %	66 / %	63 / %
1	50	RT	Sc(OTf) ₃	DCM:MeOH (3:1)	0	0	100
2	120	50	Sc(OTf) ₃	DCM:MeOH (3:1)	55	18	27
3	120	130	HCl	EtOH	87	13	0
4*	120	130	HCl	EtOH	87	13	0
5	90	130	HCl	EtOH	87	13	0
6	120	120	HCl	EtOH	90	10	0
7	120	110	HCl	EtOH	91	9	0
8	120	110	HCl	Dioxane	0	8	92

* = Performed under an inert N₂ atmosphere.

89 was synthesised following the **general procedure B** with 2-isocyano-2-methylpropane **65** (0.253 mL, 2.24 mmol). The reaction solvent was removed *in vacuo* and the resulting residue was dissolved in MeOH and loaded on to an aminopropyl SPE cartridge, washed with 3 CV of MeOH and eluted with NH₃ (2 M in MeOH). The filtrate was isolated and the solvent removed *in vacuo* to afford the product **89** (418 mg, 1.03 mmol, 92% yield) as an orange solid.

m.p. 213-238 °C; **δ_H** (400 MHz, CD₃OD) δ ppm 8.62 (s (br), 1 H), 8.09 (d, *J*=8.4 Hz, 2 H) 7.91 (s (br), 1 H), 7.57-7.52 (m, 3 H), 2.46 (s, 3 H), 2.30 (s, 3 H), 1.09 (s, 9 H); **δ_C** (101 MHz, CD₃OD) δ ppm 167.4, 160.1, 138.3, 132.6, 132.6, 130.8, 130.2, 130.0, 129.1, 117.7, 117.5, 57.2, 30.8, 11.7, 10.9; **LCMS** (Method B): *t_R* = 0.99, (*m/z*) [M+H]⁺ 405 (100% purity); **HRMS** (ESI): calculated for C₂₂H₂₆N₄O₃ (*m/z*) [M+H]⁺ = 405.2098, found 405.2098; **IR** (ν_{max}/cm⁻¹) 3337, 3284, 2968, 1621, 1521, 1405, 1205, 844.

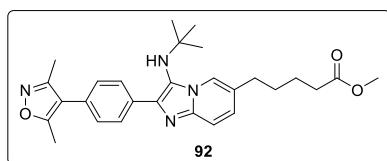
Note: Some ¹³C peaks are not visible due to partial ^tBuNH-C double-bond tautomer giving rise to peak broadening.

Methyl (4,5,6,7-tetrachloro-1,3-dioxoisindolin-2-yl) adipate 91

6-Methoxy-6-oxohexanoic acid **90** (0.593 mL, 4.00 mmol) and DCM (40 mL) were added to a round-bottomed flask. DIC (0.686 mL, 4.40 mmol), DMAP (49 mg, 0.40 mmol), and 4,5,6,7-tetrachloro-2-hydroxyisindoline-1,3-dione **35** (1.20 g, 4.00 mmol) were added to the reaction vessel and stirred at ambient temperature for 24 h. The reaction mixture was filtered through Celite with DCM washings and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-50% EtOAc in cyclohexane over 20 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **91** (1.40 g, 3.15 mmol, 79% yield) as a white solid.

m.p. 113-116 °C [Lit: 118-120 °C]¹⁵⁵; δ_{H} (400 MHz, CDCl₃) δ ppm 3.69 (s, 3 H), 2.70 (t, $J=7.0$ Hz, 2 H), 2.39 (t, $J=7.0$ Hz, 2 H), 1.84-1.79 (m, 4 H); δ_{C} (101 MHz, CDCl₃) δ ppm 173.4, 168.8, 157.5, 141.0, 130.5, 124.7, 51.6, 33.4, 30.6, 24.1, 24.0; **LCMS** (Method B): $t_{\text{R}} = 0.71$, (m/z) [M+H]⁺ not observed (50% purity due to sample degradation in LCMS, 1 spot by TLC).

Spectroscopic and physical data are consistent with the literature.¹⁵⁵

Methyl 5-(3-(*tert*-butylamino)-2-(4-(3,5-dimethylisoxazol-4-yl)phenyl)imidazo[1,2-*a*]pyridin-6-yl)pentanoate 92

NiCl₂·6H₂O and dtbbpy were added to a microwave vial and sealed. The vial was evacuated and backfilled with N₂ three times. DMF (0.4 mL) was added and the reaction mixture stirred at ambient temperature for 3 h to furnish solution A. Methyl (4,5,6,7-tetrachloro-1,3-dioxoisindolin-2-yl) adipate **91** (44 mg, 0.10 mmol) and (3-(*tert*-butylamino)-2-(4-(3,5-dimethylisoxazol-4-yl)phenyl)imidazo[1,2-*a*]pyridin-6-yl)boronic acid **89** (120 mg, 0.300 mmol) were added to a microwave vial and sealed. The vial was evacuated and backfilled with N₂ three times. Dioxane (4 mL) was added and the reaction mixture stirred at ambient

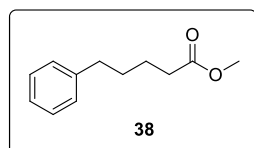
temperature for 1 min. Triethylamine (0.14 mL, 1.0 mmol) was added and the reaction mixture stirred for a further 5 min. Solution A was added and the reaction mixture immediately placed in a preheated oil bath at the specified temperature for 12 h.

A range of conditions was screened for the synthesis of **92** from **91** and **89** using the described procedure (*vide supra*) and the conditions in **Table 36**.

Table 36: Conditions screened for the decarboxylative cross-coupling of RAE **91** and boronic acid **89**, **Scheme 30**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Temperature / ° C	Catalyst / mol%	Ligand / mol%	66 / %	89 / %	92 / %
1	75	20	20	43	46	0
2	100	20	20	45	26	0
3	75	100	100	80	0	0

Methyl 5-phenylpentanoate **38**



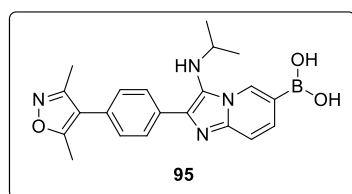
$\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$ (10 mg, 0.040 mmol) and dtbbpy (11 mg, 0.040 mmol) were added to a microwave vial and sealed. The vial was evacuated and backfilled with N_2 three times. DMF (0.8 mL) was added and the reaction mixture stirred at ambient temperature for 3 h to furnish solution A. Methyl (4,5,6,7-tetrachloro-1,3-dioxoisindolin-2-yl)adipate **91** (89 mg, 0.20 mmol) and phenylboronic acid **80** (73 mg, 0.60 mmol) were added to a microwave vial and sealed. The vial was evacuated and backfilled with N_2 three times. Dioxane (8 mL) was added and the reaction mixture stirred at ambient temperature for 1 min. Triethylamine (0.28 mL, 2.0 mmol) was added and the reaction mixture stirred for a further 5 min. Solution A was added and the reaction mixture immediately placed in a preheated oil bath at 75 °C for 12 h. The reaction mixture was diluted with water and extracted with three portions of EtOAc. The combined organic phases were washed with LiCl solution, dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-50% EtOAc in cyclohexane over 30 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was

removed *in vacuo* to afford the product **38** (17 mg, 0.070 mmol, 33% yield) as a colourless oil.

δ_{H} (400 MHz, CD₃OD) δ ppm 7.25-7.14 (m, 5 H), 3.64 (s, 3 H), 2.63-2.60 (m, 2 H), 2.35-2.33 (m, 2 H), 1.64-1.62 (m, 4 H); δ_{C} (101 MHz, CD₃OD) δ ppm 176.1, 143.7, 129.7, 129.6, 127.0, 52.3, 36.8, 34.9, 32.3, 25.9; **LCMS** (Method B): $t_{\text{R}} = 1.53$, (m/z) [M+H]⁺ not observed (75% purity).

Spectroscopic data are consistent with the literature.²²²

(2-(4-(3,5-Dimethylisoxazol-4-yl)phenyl)-3-(isopropylamino)imidazo[1,2-a]pyridin-6-yl)boronic acid (95)



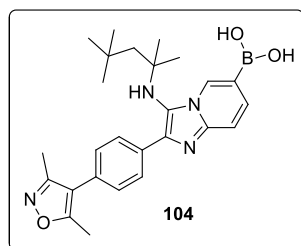
95 was synthesised following the **general procedure B** with 2-isocyanopropane **96** (0.211 mL, 2.24 mmol). The reaction solvent was removed *in vacuo* and the resulting residue was dissolved in MeOH and loaded on to an

aminopropyl SPE cartridge, washed with 3 CV of MeOH and eluted with NH₃ (2 M in MeOH). The filtrate was isolated and the solvent removed *in vacuo* to afford the product **95** (370 mg, 0.948 mmol, 85% yield) as an orange solid.

m.p. 212-224 °C; δ_{H} (400 MHz, CD₃OD) δ ppm 8.51 (s (br), 1 H), 8.14 (d, $J=8.1$ Hz, 2 H), 7.89 (s (br), 1 H), 7.57-7.52 (m, 3 H), 3.43 (sept, $J=6.2$ Hz, 1 H), 2.47 (s, 3 H), 2.31 (s, 3 H), 1.13 (d, $J=6.2$ Hz, 6 H); δ_{C} (101 MHz, CD₃OD) δ ppm 167.4, 160.1, 130.9, 128.9, 128.4, 127.8, 117.5, 111.7, 50.1, 23.6, 11.7, 10.9; **LCMS** (Method B): $t_{\text{R}} = 1.00$, (m/z) [M+H]⁺ 391 (100% purity); **HRMS** (ESI): calculated for C₂₁H₂₄BN₄O₃ (m/z) [M+H]⁺ = 391.1942, found 391.1943; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 2969, 2921, 2809, 1656, 1619, 1444, 1103, 812.

Note: Some ¹³C peaks are not visible due to partial ⁱPrNH-C double-bond tautomer giving rise to peak broadening.

(2-(4-(3,5-Dimethylisoxazol-4-yl)phenyl)-3-((2,4,4-trimethylpentan-2-yl)amino)imidazo[1,2-a]pyridin-6-yl)boronic acid **104**

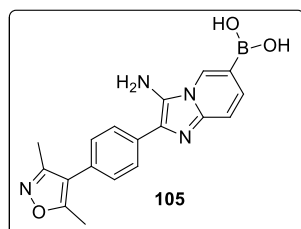


104 was synthesised following the **general procedure B** with 2-isocyano-2,4,4-trimethylpentane **103** (0.392 mL, 2.24 mmol). The reaction solvent was removed *in vacuo* and the resulting residue was dissolved in MeOH and loaded on to an aminopropyl SPE cartridge, washed with 3 CV of MeOH and eluted with NH₃ (2 M in MeOH). The filtrate was isolated and the solvent removed *in vacuo* to afford the product **104** (460 mg, 1.00 mmol, 89% yield) as an orange solid.

m.p. 211-213 °C; **δ_H** (400 MHz, CD₃OD) δ ppm 8.63 (s (br), 1 H), 8.02 (d, *J*=8.4 Hz, 2 H), 7.91 (s (br), 1 H), 7.56-7.52 (m, 3 H), 2.46 (s, 3 H), 2.30 (s, 3 H), 1.60 (s, 2 H), 1.07 (s, 6 H), 0.99 (s, 9 H); **δ_C** (101 MHz, CD₃OD) δ ppm 167.4, 160.1, 132.6, 130.8, 130.5, 129.2, 126.1, 117.6, 111.5, 61.6, 57.9, 32.5, 32.4, 29.7, 11.6, 10.9; **LCMS** (Method B): *t_R* = 1.26, (*m/z*) [M+H]⁺ 461 (100% purity); **HRMS** (ESI): calculated for C₂₆H₃₄BN₄O₃ (*m/z*) [M+H]⁺ = 461.2724, found 461.2728; **IR** (ν_{max}/cm⁻¹) 3144, 3059, 2951, 2873, 1620, 1365, 1240, 846.

Note: One ¹³C peak is not visible due to partial ¹³C OctNH-C double-bond tautomer giving rise to peak broadening. The carbon adjacent to the boronic acid is also not visible due to peak broadening.

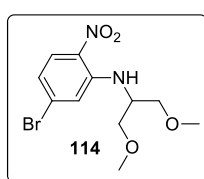
(3-Amino-2-(4-(3,5-dimethylisoxazol-4-yl)phenyl)imidazo[1,2-a]pyridin-6-yl)boronic acid **105**



(2-(4-(3,5-Dimethylisoxazol-4-yl)phenyl)-3-((2,4,4-trimethylpentan-2-yl)amino)imidazo[1,2-a]pyridin-6-yl)boronic acid **104** (392 mg, 0.851 mmol) and HCl (5-6 M in *i*PrOH) (4.26 mL, 25.5 mmol) were added to a round-bottomed flask and stirred at ambient temperature for 18 h. The solvent was removed *in vacuo* to afford the product **105** (327 mg, 0.864 mmol, 100% yield) as an orange solid.

m.p. 248-251 °C; δ_{H} (400 MHz, CD₃OD) δ ppm 8.81 (s (br), 1 H), 8.08 (d, $J=8.8$ Hz, 1 H), 7.99-7.97 (m, 2 H), 7.78 (d, $J=8.8$ Hz, 1 H), 7.60-7.56 (m, 2 H), 2.47 (s, 3 H), 2.31 (s, 3 H); δ_{C} (101 MHz, CD₃OD) δ ppm 167.6, 160.0, 137.6, 132.6, 131.3, 130.5, 129.3, 128.6, 128.6, 128.2, 119.0, 119.0, 117.4, 111.6, 11.7, 10.9; **LCMS** (Method B): $t_{\text{R}} = 0.76$, (m/z) [M+H]⁺ 349 (92% purity); **HRMS** (ESI): calculated for C₁₈H₁₈BN₄O₃ (m/z) [M+H]⁺ = 349.1472, found 349.1476; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3166, 3059, 2820, 2725, 1655, 1619, 1320, 833.

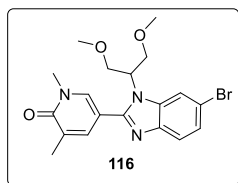
5-Bromo-*N*-(1,3-dimethoxypropan-2-yl)-2-nitroaniline **114**



4-Bromo-2-fluoro-1-nitrobenzene **112** (6.00 g, 27.3 mmol), DIPEA (9.5 mL, 55 mmol), 1,3-dimethoxypropan-2-amine **113** (4.03 mL, 32.7 mmol) and 2-MeTHF (140 mL) were added to a round-bottomed flask and heated to reflux (100 °C) for 16 h under an inert N₂ atmosphere. The reaction mixture was diluted with water and extracted with three portions of EtOAc. The combined organic phases were dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-7.5% EtOAc in cyclohexane over 20 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **114** (7.96 g, 24.5 mmol, 90% yield) as an orange solid.

m.p. 78-81 °C; δ_{H} (400 MHz, CD₃OD) δ ppm 8.01 (d, $J=9.4$ Hz, 1 H), 7.31 (s, 1 H), 6.80 (d, $J=9.4$ Hz, 1 H), 3.99 (quin, $J=5.5$ Hz, 1 H), 3.59-3.57 (m, 4 H), 3.39 (s, 6 H); δ_{C} (101 MHz, CD₃OD) δ ppm 145.3, 130.9, 127.7, 118.3, 118.3, 116.9, 71.4, 58.2, 52.2; **LCMS** (Method B): $t_{\text{R}} = 1.26$, (m/z) [M+H]⁺ 230 (98% purity); **HRMS** (ESI): calculated for C₁₁H₁₆BrN₂O₄ (m/z) [M+H]⁺ = 319.0293, found 319.0289; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3364, 3107, 2990, 2891, 2815, 1601, 1561, 1103.

5-(6-Bromo-1-(1,3-dimethoxypropan-2-yl)-1H-benzo[d]imidazol-2-yl)-1,3-dimethylpyridin-2(1H)-one **116**



5-Bromo-*N*-(1,3-dimethoxypropan-2-yl)-2-nitroaniline **114**

(2.94 g, 9.21 mmol), 1,5-dimethyl-6-oxo-1,6-dihydropyridine-3-carbaldehyde **115** (1.53 g, 10.1 mmol) and EtOH (29 mL) were added to a microwave vial. Na₂S₂O₄ (9.62 g, 55.3 mmol) and water (14.5 mL) were added, the reaction vessel was sealed and immediately heated at 140 °C for 30 min in a MW reactor. The reaction mixture was diluted with water and extracted with three portions of EtOAc. The combined organic phases were dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-50% 3:1 EtOAc:EtOH in cyclohexane over 20 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **116** (3.00 g, 7.14 mmol, 77% yield) as a white solid.

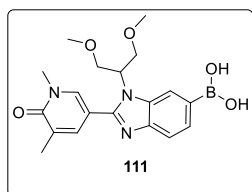
m.p. 179-181 °C; **δ_H** (400 MHz, CDCl₃) δ ppm 7.88-7.87 (m, 1 H), 7.71-7.70 (m, 1 H), 7.68-7.67 (m, 1 H), 7.63 (d, *J*=8.4 Hz, 1 H), 7.40-7.37 (m, 1 H), 4.86-4.79 (m, 1 H), 3.95-3.91 (m, 2 H), 3.84-3.80 (m, 2 H), 3.62 (s, 3 H), 3.30 (s, 6 H), 2.25 (s, 3 H); **δ_C** (101 MHz, CDCl₃) δ ppm 186.2, 138.6, 137.4, 134.9, 130.0, 125.8, 121.3, 115.7, 114.8, 108.6, 92.5, 92.2, 70.8, 59.3, 57.9, 38.2, 17.4; **LCMS** (Method B): *t_R* = 1.00, (*m/z*) [M+H]⁺ 420 (100% purity); **HRMS** (ESI): calculated for C₁₉H₂₃BrN₃O₃ (*m/z*) [M+H]⁺ = 420.0923, found 420.0919; **IR** (ν_{max}/cm⁻¹) 2980, 2927, 2889, 2810, 1656, 1618, 1442, 1102.

A range of equivalents of Na₂S₂O₄ and reaction durations was screened using the described procedure (*vide supra*) and the conditions visualised in **Table 37**.

Table 37: Reaction conditions screened for the synthesis of bromide **116** from bromide **114** in a MW reactor. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Na ₂ S ₂ O ₄ / equiv	Duration / h	116 Peak Area / %	116 Isolated Yield / %
1	6	1	63	✘
2	4	1	59	✘
3	6	0.5	84	77

(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)boronic acid **111**

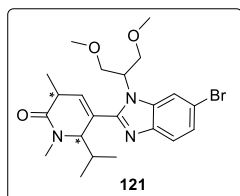


5-(6-Bromo-1-(1,3-dimethoxypropan-2-yl)-1H-benzo[d]imidazol-2-yl)-1,3-dimethylpyridin-2(1H)-one **116** (127 mg, 0.302 mmol), bis-(neopentylglycolato)diboron **117** (136 mg, 0.604 mmol), Pd(OAc)₂ (7.0 mg, 0.030 mmol), KOAc (89 mg, 0.91 mmol), ^tBuPPh₂ (15 mg, 0.060 mmol) and dioxane (15 mL) were added to a microwave vial and sealed. The reaction mixture was heated to 90 °C for 16 h under an inert atmosphere of N₂. The reaction mixture was filtered through Celite with EtOAc washings. The solvent was removed *in vacuo* and the crude product was purified by reverse phase column chromatography eluting with 0-30% MeCN in 10 mM ammonium bicarbonate solution (pH 10) over 40 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **111** (100 mg, 0.260 mmol, 86% yield) as a white solid.

m.p. 172-178 °C; **δ_{H}** (400 MHz, CD₃OD) δ ppm 8.05-8.03 (m, 1 H), 8.01-7.98 (m, 1 H), 7.80-7.76 (m, 1 H), 7.69-7.65 (m, 1 H), 7.59-7.55 (m, 1 H), 4.93-4.88 (m, 1 H), 4.12-4.09 (m, 2 H), 3.85-3.83 (m, 2 H), 3.66 (s, 3 H), 3.25 (s, 6 H), 2.20 (s, 3 H); **δ_{C}** (101 MHz, CD₃OD) δ ppm 164.8, 154.0, 144.9, 140.7, 139.9, 134.7, 130.0, 129.0, 119.4, 119.2, 110.9, 71.6, 59.6, 59.5, 38.9, 17.4; **LCMS** (Method A): t_{R} = 0.47, (m/z) [M+H]⁺ 386 (100% purity); **HRMS** (ESI): calculated for C₁₉H₂₅BN₃O₅ (m/z) [M+H]⁺ = 386.1887, found 386.1890; **IR** (ν_{max} /cm⁻¹) 3304, 3075, 2927, 2889, 2815, 1654, 1608, 1334.

Note: The carbon adjacent to the boronic acid is not visible due to peak broadening.

5-(6-Bromo-1-(1,3-dimethoxypropan-2-yl)-1H-benzo[d]imidazol-2-yl)-6-isopropyl-1,3-dimethyl-3,6-dihydropyridin-2(1H)-one **121**



5-(6-Bromo-1-(1,3-dimethoxypropan-2-yl)-1H-benzo[d]imidazol-2-yl)-1,3-dimethylpyridin-2(1H)-one **116** (25 mg, 0.060 mmol) was added to an oven-dried microwave vial and sealed. The vial was evacuated and backfilled with N₂

three times. THF (0.5 mL) was added and the solution stirred for 10 min. ⁱPrMgCl·LiCl (1.3 M in THF) (0.050 mL, 0.065 mmol) was added and the reaction mixture stirred at ambient temperature for 1 h. The reaction mixture was quenched with sat. aq. NH₄Cl and the solvent was removed *in vacuo*. The crude product was purified by MDAP. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the products **121a** (5.0 mg, 0.011 mmol, 18% yield) as a white solid and **121b** (7.0 mg, 0.015 mmol, 25% yield) as a white solid.

121a

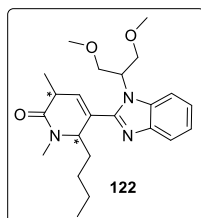
δ_{H} (400 MHz, CD₃OD) δ ppm 8.00-7.98 (m, 1 H), 7.57-7.55 (m, 1 H), 7.43-7.41 (m, 1 H), 6.51-6.50 (m, 1 H), 5.04-4.99 (m, 1 H), 4.16-4.12 (m, 1 H), 3.99-3.13 (m, 3 H), 3.79-3.76 (m, 1 H), 3.37 (s, 3 H), 3.24-3.20 (m, 1 H), 3.19 (s, 3 H), 3.09 (s, 3 H), 2.14-2.09 (m, 1 H), 1.48 (s, 3 H), 0.96 (d, $J=7.0$ Hz, 3 H), 0.68 (d, $J=7.0$ Hz, 3 H); **LCMS** (Method B): $t_{\text{R}} = 1.15$, (m/z) [M+H]⁺ 456 (80% purity).

121b

δ_{H} (400 MHz, CD₃OD) δ ppm 8.00-7.98 (m, 1 H), 7.57-7.55 (m, 1 H), 7.43-7.41 (m, 1 H), 6.47-6.46 (m, 1 H), 5.12-4.08 (m, 1 H), 4.77-4.76 (m, 1 H), 4.17-4.17 (m, 1 H), 3.99-3.96 (m, 1 H), 3.94-3.91 (m, 1 H), 3.77-3.74 (m, 1 H), 3.36 (s, 3 H), 3.30-3.26 (m, 1 H), 3.18 (s, 3 H), 3.12 (s, 3 H), 2.26-2.21 (m, 1 H), 1.36 (s, 3 H), 0.96 (d, $J=7.0$ Hz, 3 H), 0.74 (d, $J=7.0$ Hz, 3 H); **LCMS** (Method B): $t_{\text{R}} = 1.18$, (m/z) [M+H]⁺ 456 (93% purity).

Confirmation of regiochemistry was obtained *via* a series of NOE experiments reported in **Appendix 7.1.1**.

6-Butyl-5-(1-(1,3-dimethoxypropan-2-yl)-1*H*-benzo[*d*]imidazol-2-yl)-1,3-dimethyl-3,6-dihydropyridin-2(1*H*)-one **122**



5-(6-Bromo-1-(1,3-dimethoxypropan-2-yl)-1*H*-benzo[*d*]imidazol-2-yl)-1,3-dimethylpyridin-2(1*H*)-one **116** (50 mg, 0.12 mmol) was added to an oven-dried microwave vial and sealed. The vial was evacuated and backfilled with N₂ three times. THF (1.2 mL) was added and the reaction mixture cooled to -78 °C. *n*BuLi (2.5 M in hexanes) (0.05 mL, 0.12 mmol) was added and the reaction mixture stirred at ambient temperature for 1 h. The reaction mixture was quenched with sat. aq. NH₄Cl and extracted with three portions of EtOAc. The combined organic phases were dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by MDAP. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the products **122a** (4.0 mg, 0.0097 mmol, 6% yield) as a white solid and **122b** (3.0 mg, 0.0073 mmol, 4% yield) as a white solid.

122a

δ_{H} (400 MHz, CD₃OD) δ ppm 7.77-7.74 (m, 1 H), 7.67-7.65 (m, 1 H), 7.31-7.28 (m, 2 H), 6.46 (d, $J=4.4$ Hz, 1 H), 5.03-4.97 (m, 1 H), 4.85-4.83 (m, 1 H), 4.20-4.15 (m, 1 H), 3.97-3.93 (m, 2 H), 3.83-3.79 (m, 1 H), 3.35 (s, 3 H), 3.23-3.22 (m, 1 H), 3.16 (s, 3 H), 3.06 (s, 3 H), 1.84-1.81 (m, 1 H), 1.47 (d, $J=7.6$ Hz, 3 H), 1.32-1.24 (m, 4 H), 1.06-1.02 (m, 1 H), 0.86 (t, $J=7.1$ Hz, 3 H); **LCMS** (Method B): $t_{\text{R}} = 1.09$, (m/z) [M+H]⁺ 401 (48% purity).

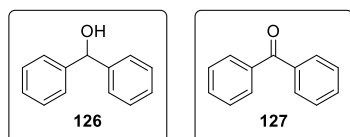
122b

7.77-7.75 (m, 1 H), 7.67-7.65 (m, 1 H), 7.31-7.27 (m, 2 H), 6.37 (d, $J=2.5$ Hz, 1 H), 5.09-5.02 (m, 1 H), 4.85-4.83 (m, 1 H), 4.21-4.16 (m, 1 H), 3.99-3.91 (m, 2 H), 3.81-3.77 (m, 1 H), 3.34 (s, 3 H), 3.15 (s, 3 H), 3.07 (s, 3 H), 1.90-1.86 (m, 1 H), 1.59-1.50

(m, 1 H), 1.38 (d, $J=7.6$ Hz, 3 H), 1.33-1.24 (m, 4 H), 1.10-1.07 (m, 1 H), 0.86 (t, $J=7.1$ Hz, 3 H); **LCMS** (Method B): $t_R = 1.12$, (m/z) $[M+H]^+$ 401 (41% purity).

Confirmation of regiochemistry was obtained *via* a series of NOE experiments reported in **Appendix 7.1.2**.

Diphenylmethanol **126** and benzophenone **127**



Magnesium turnings (32.5 mg, 1.34 mmol) were ground with a pestle and mortar and added to an oven-dried microwave vial and sealed. The vial was evacuated and

backfilled with N_2 three times. Iodine (2.0 mg, 6.4 μ mol) in THF (0.5 mL) was added and the reaction mixture refluxed (75 °C) for 30 min. An aliquot (0.05 mL) of bromobenzene (0.134 mL, 1.27 mmol) in THF (0.25 mL) was added and the reaction mixture refluxed (75 °C) for 30 min. The rest of the bromobenzene in THF solution was added and the reaction mixture refluxed (75 °C) for 30 min. The reaction was cooled to 0 °C in an ice bath and benzaldehyde (0.142 mL, 1.40 mmol) in THF (0.14 mL) was added and the reaction mixture warmed to ambient temperature and stirred for 16 h. The reaction mixture was quenched with sat. aq. NH_4Cl and extracted with three portions of EtOAc. The combined organic phases were dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by flash column chromatography eluting with 0-30% EtOAc in cyclohexane over 20 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the products **126** (118 mg, 0.640 mmol, 50% yield) as a white solid and **127** (79 mg, 0.42 mmol, 33% yield) as a colourless oil.

126

m.p. 64-66 °C [Lit: 66 °C]²²³; δ_H (400 MHz, $CDCl_3$) δ ppm 7.41-7.23 (m, 10 H) 5.85 (d, $J=3.5$ Hz, 1 H) 2.19 (d, $J=3.5$ Hz, 1 H); δ_C (101 MHz, $CDCl_3$) δ ppm 143.8, 128.5, 127.6, 126.5, 76.3; **LCMS** (Method B): $t_R = 1.01$, (m/z) $[M+H]^+$ 183 (100% purity).

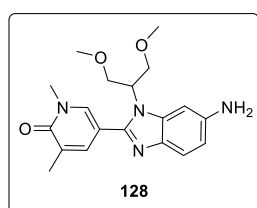
Spectroscopic and physical data are consistent with the literature.²²³

127

δ_{H} (400 MHz, CDCl_3) δ ppm 7.82-7.79 (m, 4 H), 7.60-7.56 (m, 2 H), 7.50-7.46 (m, 4 H); δ_{C} (101 MHz, CDCl_3) δ ppm 196.7, 137.7, 132.4, 130.1, 128.3; **LCMS** (Method B): $t_{\text{R}} = 1.15$, (m/z) $[\text{M}+\text{H}]^+$ 183 (96% purity).

Spectroscopic data are consistent with the literature.²²⁴

5-(6-Amino-1-(1,3-dimethoxypropan-2-yl)-1H-benzo[d]imidazol-2-yl)-1,3-dimethylpyridin-2(1H)-one 128

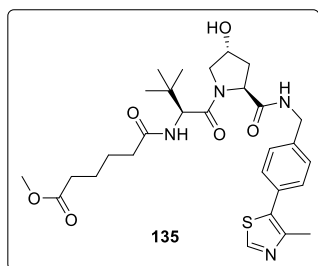


In a glovebox, $[\text{Pd}(\text{cinnamyl})\text{Cl}]_2$ (37 mg, 0.071 mmol), BippyPhos (145 mg, 0.286 mmol) and dioxane (7.1 mL) were added to an oven-dried microwave vial and stirred at ambient temperature for 2 min under an inert N_2 atmosphere. Sodium *tert*-butoxide (192 mg, 2.00 mmol) was added and the reaction mixture stirred for 2 min. 5-(6-Bromo-1-(1,3-dimethoxypropan-2-yl)-1H-benzo[d]imidazol-2-yl)-1,3-dimethylpyridin-2(1H)-one **116** (600 mg, 1.43 mmol) was added to the reaction mixture and the vial was sealed. Ammonia (0.4 M dioxane) (11 mL, 4.3 mmol) was added to the reaction mixture. The vial was removed from the glovebox and heated at 110 °C for 16 h. The reaction mixture was filtered through Celite with DCM washings. The solvent was removed *in vacuo* and the crude product was purified by flash column chromatography eluting with 0-100% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **128** (520 mg, 1.46 mmol, 98% yield) as a white solid.

m.p. 230-234 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 7.94-7.92 (m, 1 H), 7.62-7.60 (m, 1 H), 7.26 (d, $J=8.3$ Hz, 1 H), 6.82-6.80 (m, 1 H), 6.55-6.52 (m, 1 H), 4.89 (s, 2 H), 4.72-4.65 (m, 1 H), 3.95-3.90 (m, 2 H), 3.74-3.70 (m, 2 H), 3.51 (s, 3 H), 3.17 (s, 6 H), 2.06 (s, 3 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 162.2, 149.3, 145.1, 139.0, 138.1, 135.8, 135.3, 127.9, 119.9, 112.0, 109.1, 96.2, 70.1, 59.0, 57.0, 38.0, 17.6; **LCMS** (Method B): $t_{\text{R}} = 0.66$, (m/z) $[\text{M}+\text{H}]^+$ 357 (100% purity); **HRMS** (ESI):

calculated for $C_{19}H_{25}N_4O_3$ (m/z) $[M+H]^+ = 357.1927$, found 357.1925; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3415, 3320, 3219, 2887, 2804, 1655, 1602, 1117.

Methyl 6-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-6-oxohexanoate **135**

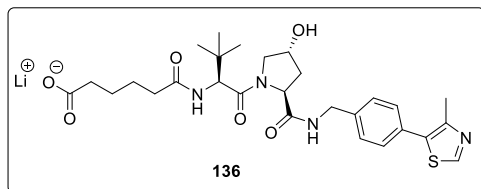


6-Methoxy-6-oxohexanoic acid **90** (0.245 mL, 1.65 mmol), HATU (630 mg, 1.65 mmol), DIPEA (1.1 mL, 6.0 mmol) and DCM (30 mL) were added to a round-bottomed flask and sonicated for 30 s. (2*S*,4*R*)-1-((*S*)-2-amino-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-

carboxamide·HCl **30** (702 mg, 1.50 mmol) was added and the reaction mixture stirred at ambient temperature for 4 h. The solvent was removed *in vacuo* and the crude product was purified by reverse phase column chromatography eluting with 0-50% MeCN in 10 mM ammonium bicarbonate solution (pH 10) over 20 CV. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **135** (686 mg, 1.15 mmol, 76% yield) as a beige solid.

m.p. 48-58 °C; **δ_H** (400 MHz, d_6 -DMSO) δ ppm 8.70 (s, 1 H), 7.38-7.32 (m, 5 H), 6.24 (d, $J=8.8$ Hz, 1 H), 4.72 (t, $J=8.1$ Hz, 1 H), 4.58-4.51 (m, 3 H), 4.36-4.33 (m, 1 H), 4.08 (d, $J=11.4$ Hz, 1 H), 3.65 (s, 3 H), 3.62-3.60 (m, 1 H), 2.51 (s, 1 H), 2.32-2.11 (m, 5 H), 1.64-1.62 (m, 4 H), 0.93 (s, 9 H); **δ_C** (101 MHz, d_6 -DMSO) δ ppm 173.7, 173.0, 171.7, 170.5, 150.1, 148.3, 137.9, 131.3, 130.8, 129.3, 127.9, 120.8, 69.8, 58.3, 57.3, 56.5, 51.4, 43.0, 35.7, 35.7, 34.7, 26.2, 24.4, 24.4, 15.8; **LCMS** (Method B): $t_R = 0.91$, (m/z) $[M+H]^+ 573$ (96% purity); **HRMS** (ESI): calculated for $C_{29}H_{41}N_4O_6S$ (m/z) $[M+H]^+ = 573.2747$, found 573.2744; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3305, 3078, 2952, 2873, 1730, 1623, 1531, 1435; **$[\alpha]_D^{20} = -37$** (c 0.42, MeOH).

Lithium 6-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-6-oxohexanoate 136



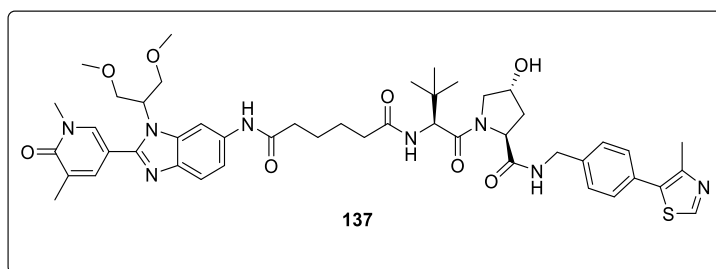
Lithium hydroxide (4.0 mg, 0.15 mmol) and water (0.2 mL) were added to a round-bottomed flask and stirred for 5 min. A solution of methyl 6-(((S)-1-((2S,4R)-4-

hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-6-oxohexanoate **135** (64 mg, 0.11 mmol) in THF (0.2 mL) was added and the reaction mixture stirred at ambient temperature for 1 h. The solvent was removed *in vacuo* to afford the crude product **136** (63 mg, 0.12 mmol, 100% yield) as a white solid.

LCMS (Method B): $t_R = 0.65$, $[M+H]^+$ 558 (100% purity).

The crude product was used without further purification.

N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N6-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)adipamide 137

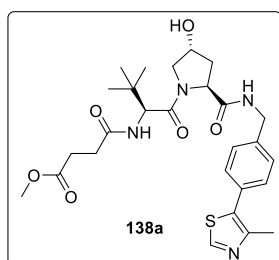


Lithium 6-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-6-oxohexanoate **136** (59 mg, 0.10 mmol), HATU (43 mg, 0.11 mmol), DIPEA (0.054 mL, 0.31 mmol) and DCM (2.1 mL) were added to a round-bottomed flask and sonicated for 30 s. 5-(6-amino-1-(1,3-dimethoxypropan-2-yl)-1H-benzo[d]imidazol-2-yl)-1,3-dimethylpyridin-2(1H)-

one **128** (37 mg, 0.10 mmol) was added and the reaction mixture stirred at ambient temperature for 16 h. The reaction mixture was diluted with water and extracted with three portions of DCM. The combined organic phases were dried through a hydrophobic frit and the solvent was removed *in vacuo*. The crude product was purified by MDAP. UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **137** (66 mg, 0.070 mmol, 71% yield) as a white solid.

m.p. 134-148 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.89 (s, 1 H), 8.97 (s, 1 H), 8.53 (t, $J=6.1$ Hz, 1 H), 8.21-8.19 (m, 1 H), 8.03-8.01 (m, 1 H), 7.85 (d, $J=9.2$ Hz, 1 H), 7.66-7.65 (m, 1 H), 7.52 (d, $J=8.8$ Hz, 1 H), 7.43-7.37 (m, 4 H), 7.30-7.27 (m, 1 H), 5.11-5.09 (m, 1 H), 4.83-4.75 (m, 1 H), 4.55 (d, $J=9.2$ Hz, 1 H), 4.46-4.40 (m, 2 H), 4.38-4.34 (m, 1 H), 4.24-4.19 (m, 1 H), 3.98-3.94 (m, 2 H), 3.77-3.74 (m, 2 H), 3.69-3.64 (m, 2 H), 3.53 (s, 3 H), 3.17 (s, 6 H), 2.45 (s, 3 H), 2.36-2.29 (m, 3 H), 2.21-2.15 (m, 1 H), 2.09 (s, 3 H), 2.05-2.01 (m, 1 H), 1.93-1.89 (m, 1 H), 1.64-1.51 (m, 4 H), 0.94 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.9, 171.9, 170.9, 169.7, 161.6, 151.5, 151.4, 147.7, 139.4, 139.1, 138.9, 137.3, 134.2, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 118.9, 114.7, 107.9, 102.7, 69.5, 68.8, 58.6, 58.4, 56.7, 56.3, 56.3, 41.6, 37.9, 37.4, 36.3, 35.2, 34.7, 26.3, 25.2, 24.9, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.90$, (m/z) $[M+H]^+$ 897 (100% purity); **HRMS** (ESI): calculated for $C_{47}H_{61}N_8O_8S$ (m/z) $[M+2H]^{2+} = 449.2208$, found 449.2207; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3305, 3078, 2952, 2873, 1730, 1623, 1531, 1435; $[\alpha]_{\text{D}}^{20} = -11$ (c 1.00, MeOH).

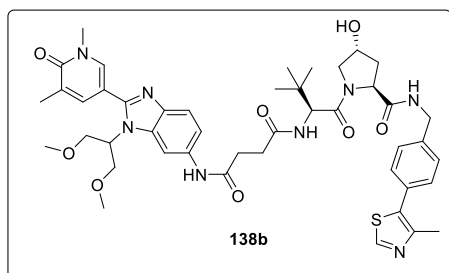
Methyl 4-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-4-oxobutanoate 138a



138a was synthesised following the **general procedure C** with 4-methoxy-4-oxobutanoic acid (109 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-40% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **138a** (250 mg, 0.454 mmol, 66% yield) was isolated as a white solid.

m.p. 82-94 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 8.98 (s, 1 H), 8.53 (t, $J=6.1$ Hz, 1 H), 7.95 (d, $J=9.5$ Hz, 1 H), 7.43-7.38 (m, 4 H), 5.10 (d, $J=3.7$ Hz, 1 H), 4.53 (d, $J=9.5$ Hz, 1 H), 4.45-4.41 (m, 2 H), 4.37-4.33 (m, 1 H), 4.24-4.20 (m, 1 H), 4.68-4.65 (m, 1 H), 4.61-4.60 (m, 1 H), 3.56 (s, 3 H), 2.58-2.38 (m, 7 H), 2.05-2.01 (m, 1 H), 1.93-1.88 (m, 1 H), 0.93 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 172.7, 171.8, 170.5, 169.5, 151.4, 147.7, 139.4, 131.1, 129.6, 128.6, 127.4, 68.8, 58.6, 56.4, 56.3, 51.2, 41.6, 37.9, 35.3, 29.5, 28.9, 26.3, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.85$, (m/z) $[\text{M}+\text{H}]^+$ 545 (99% purity); **HRMS** (ESI): calculated for $\text{C}_{27}\text{H}_{37}\text{N}_4\text{O}_6\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 545.2434$, found 545.2437; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3299, 3075, 2954, 2868, 1731, 1622, 1532, 1435; $[\alpha]_{\text{D}}^{20} = -36$ (c 1.00, MeOH).

N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N4-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)succinamide 138b



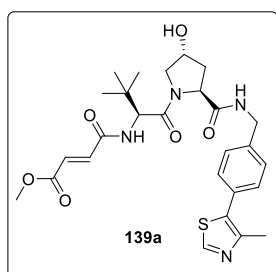
138b was synthesised following the **general procedure D** with the VHL-linker methyl 4-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-4-oxobutanoate **138a** (237 mg,

0.435 mmol) in DCM. The crude product was purified by reverse phase column chromatography eluting with 0-50% MeCN in 10 mM ammonium bicarbonate solution (pH 10) over 30 CV. **138b** (240 mg, 0.273 mmol, 63% yield) was isolated as a white solid.

m.p. 142-156 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.89 (s, 1 H), 8.97 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 8.21-8.19 (m, 1 H), 8.03-8.01 (m, 1 H), 7.94 (d, $J=9.2$ Hz, 1 H), 7.66-7.65 (m, 1 H), 7.52 (d, $J=8.4$ Hz, 1 H), 7.43-7.37 (m, 4 H), 7.28-7.26 (m, 1 H), 5.11-5.09 (m, 1 H), 4.83-4.75 (m, 1 H), 4.56 (d, $J=9.2$ Hz, 1 H), 4.46-4.40 (m, 2 H), 4.38-4.34 (m, 1 H), 4.24-4.21 (m, 1 H), 3.97-3.93 (m, 2 H), 3.77-3.74 (m, 2 H), 3.69-3.62 (m, 2 H), 3.53 (s, 3 H), 3.17 (s, 6 H), 2.68-2.35 (m, 4 H), 2.45 (s, 3 H), 2.09 (s, 3 H),

2.06-2.02 (m, 1 H), 1.93-1.89 (m, 1 H), 0.95 (s, 9 H); δ_C (101 MHz, d_6 -DMSO) δ ppm 171.8, 171.2, 170.3, 169.5, 161.6, 151.4, 151.4, 147.7, 139.4, 139.0, 138.9, 137.2, 134.3, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 118.9, 114.6, 107.9, 102.6, 69.4, 68.8, 58.7, 58.4, 56.7, 56.4, 56.3, 41.6, 37.9, 37.4, 35.3, 31.9, 30.1, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_R = 0.88$, (m/z) $[M+H]^+$ 869 (99% purity); **HRMS** (ESI): calculated for $C_{45}H_{57}N_8O_8S$ (m/z) $[M+H]^+ = 869.4020$, found 869.4014; **IR** (ν_{max}/cm^{-1}) 3299, 3075, 2963, 2927, 1656, 1614, 1539, 1485; $[\alpha]_D^{20} = -32$ (c 1.00, MeOH).

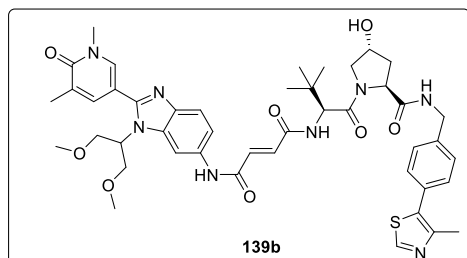
Methyl (*E*)-4-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-4-oxobut-2-enoate **139a**



139a was synthesised following the **general procedure C** with (*E*)-4-methoxy-4-oxobut-2-enoic acid (108 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-50% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **139a** (227 mg, 0.381 mmol, 51% yield) was isolated as a white solid.

m.p. 116-128 °C; δ_H (400 MHz, d_6 -DMSO) δ ppm 8.98 (s, 1 H), 8.67 (d, $J=9.2$ Hz, 1 H), 8.57 (t, $J=6.1$ Hz, 1 H), 7.45-7.37 (m, 5 H), 6.60 (d, $J=15.4$ Hz, 1 H), 5.12 (d, $J=3.7$ Hz, 1 H), 4.62 (d, $J=9.2$ Hz, 1 H), 4.45-4.42 (m, 2 H), 4.38-4.35 (m, 1 H), 4.24-4.21 (m, 1 H), 3.73 (s, 3 H), 4.71-4.65 (m, 2 H), 2.45 (s, 3 H), 2.06-2.02 (m, 1 H), 1.93-1.89 (m, 1 H), 0.97 (s, 9 H); δ_C (101 MHz, d_6 -DMSO) δ ppm 171.8, 168.9, 165.5, 162.6, 151.4, 147.7, 139.4, 137.6, 131.1, 129.6, 128.6, 128.5, 127.4, 68.8, 58.7, 57.1, 56.4, 51.9, 41.6, 37.9, 35.3, 26.3, 15.9; **LCMS** (Method B): $t_R = 0.89$, (m/z) $[M+H]^+$ 543 (91% purity); **HRMS** (ESI): calculated for $C_{27}H_{37}N_4O_6S$ (m/z) $[M+H]^+ = 543.2277$, found 543.2281; **IR** (ν_{max}/cm^{-1}) 3300, 3059, 2948, 2868, 1662, 1621, 1524, 1435; $[\alpha]_D^{20} = -77$ (c 1.00, MeOH).

N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*4-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)fumaramide 139b

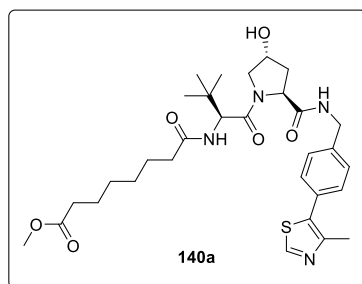


139b was synthesised following the **general procedure D** with the VHL-linker methyl (*E*)-4-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-

yl)amino)-4-oxobut-2-enoate **139a** (217 mg, 0.400 mmol) in DCM. The crude product was purified by MDAP. **139b** (210 mg, 0.242 mmol, 61% yield) was isolated as a white solid.

m.p. 143-151 °C; **δ_{H}** (400 MHz, *d*₆-DMSO) δ ppm 10.52 (s, 1 H), 9.89 (s, 1 H), 8.63-8.57 (m, 2 H), 8.36 (s, 1 H), 8.06-8.05 (m, 1 H), 7.68-7.66 (m, 1 H), 7.59 (d, *J*=8.8 Hz, 1 H), 7.46-7.39 (m, 5 H), 7.31 (d, *J*=15.4 Hz, 1 H), 7.10 (d, *J*=15.4 Hz, 1 H), 5.14 (s (br), 1 H), 4.88-4.83 (m, 1 H), 4.69 (d, *J*=9.5 Hz, 1 H), 4.48-4.43 (m, 2 H), 4.40-4.36 (m, 1 H), 4.25-4.22 (m, 1 H), 4.01-3.96 (m, 2 H), 3.80-3.78 (m, 2 H), 3.74-3.68 (m, 2 H), 3.54 (s, 3 H), 3.19 (s, 6 H), 2.45 (s, 3 H), 2.09 (s, 3 H), 2.07-2.04 (m, 1 H), 1.95-1.91 (m, 1 H), 0.99 (s, 9 H); **δ_{C}** (101 MHz, *d*₆-DMSO) δ ppm 171.8, 169.1, 163.5, 162.2, 161.7, 151.6, 151.4, 147.7, 139.4, 139.2, 138.6, 137.1, 133.9, 133.5, 133.5, 133.3, 131.1, 129.6, 128.6, 127.5, 127.4, 118.0, 115.3, 107.5, 103.4, 69.4, 68.9, 58.7, 58.4, 56.9, 56.9, 56.4, 41.6, 37.9, 37.5, 35.4, 26.3, 16.9, 15.9; **LCMS** (Method B): *t*_R = 0.92, (*m/z*) [M+H]⁺ 867 (100% purity); **HRMS** (ESI): calculated for C₄₅H₅₅N₈O₈S (*m/z*) [M+H]⁺ = 867.3864, found 867.3844; **IR** (ν_{max} /cm⁻¹) 3288, 3081, 2974, 2927, 2873, 1652, 1605, 1101; [α]_D²⁰ = -29 (*c* 0.38, MeOH).

Methyl 8-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-8-oxooctanoate **140a**

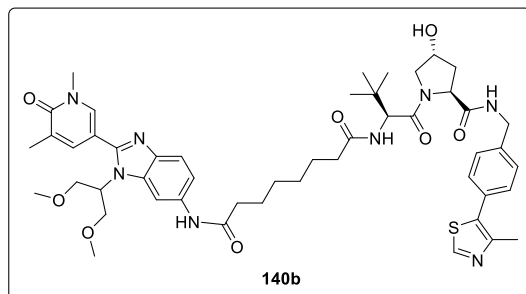


140a was synthesised following the **general procedure C** with 8-methoxy-8-oxooctanoic acid (0.149 mL, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-50% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **140a** (280 mg, 0.466 mmol, 62% yield) was isolated as a

white solid.

m.p. 56-64 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 8.98 (s, 1 H), 8.52 (t, $J=6.1$ Hz, 1 H), 7.81 (d, $J=9.2$ Hz, 1 H), 7.42-7.38 (m, 4 H), 5.09 (d, $J=3.3$ Hz, 1 H), 4.54 (d, $J=9.2$ Hz, 1 H), 4.45-4.41 (m, 2 H), 4.36-4.34 (m, 1 H), 4.24-4.20 (m, 1 H), 3.68-3.62 (m, 2 H), 3.58 (s, 3 H), 2.44 (s, 3 H), 2.29-2.22 (m, 3 H), 2.13-2.08 (m, 1 H), 2.04-2.01 (m, 1 H), 1.93-1.88 (m, 1 H), 1.53-1.42 (m, 4 H), 1.28-1.20 (m, 4 H), 0.93 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 173.3, 172.0, 171.9, 169.7, 151.4, 147.7, 139.4, 131.1, 129.6, 128.6, 127.4, 68.8, 58.6, 56.3, 56.2, 51.1, 41.6, 37.9, 35.1, 34.7, 33.2, 28.2, 28.1, 26.3, 25.2, 24.3, 15.9; **LCMS** (Method B): $t_{\text{R}} = 1.00$, (m/z) $[\text{M}+\text{H}]^+$ 601 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{31}\text{H}_{47}\text{N}_4\text{O}_6\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 601.3060$, found 601.3065; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3311, 3070, 2949, 2863, 1736, 1629, 1534, 1436; $[\alpha]_{\text{D}}^{20} = -35$ (c 1.00, MeOH).

N*1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*8-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)octanediamide **140b*

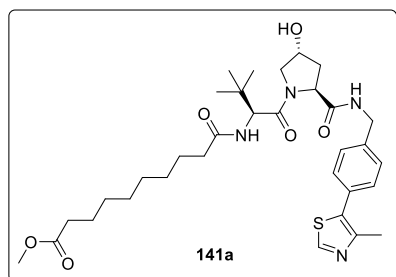


140b was synthesised following the **general procedure D** with the VHL-linker methyl 8-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-8-

oxooctanoate **140a** (265 mg, 0.441 mmol) in DCM. The crude product was purified by MDAP. **140b** (238 mg, 0.257 mmol, 58% yield) was isolated as a white solid.

m.p. 135-142 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.88 (s, 1 H), 8.97 (s, 1 H), 8.53 (t, $J=6.1$ Hz, 1 H), 8.21-8.19 (m, 1 H), 8.03-8.01 (m, 1 H), 7.81 (d, $J=9.5$ Hz, 1 H), 7.67-7.65 (m, 1 H), 7.52 (d, $J=8.8$ Hz, 1 H), 7.42-7.38 (m, 4 H), 7.30-7.28 (m, 1 H), 5.10 (d, $J=3.3$ Hz, 1 H), 4.83-4.79 (m, 1 H), 4.55 (d, $J=9.5$ Hz, 1 H), 4.44-4.40 (m, 2 H), 4.37-4.34 (m, 1 H), 4.23-4.20 (m, 1 H), 3.98-3.94 (m, 2 H), 3.77-3.74 (m, 2 H), 3.69-3.63 (m, 2 H), 3.53 (s, 3 H), 3.17 (s, 6 H), 2.44 (s, 3 H), 2.33-2.31 (m, 2 H), 2.29-2.24 (m, 1 H), 2.16-2.11 (m, 1 H), 2.08 (s, 3 H), 2.05-2.01 (m, 1 H), 1.93-1.89 (m, 1 H), 1.63-1.58 (m, 2 H), 1.56-1.45 (m, 2 H), 1.35-1.27 (m, 4 H), 0.94 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 172.0, 171.0, 171.9, 169.7, 161.6, 151.5, 151.4, 147.7, 139.4, 139.1, 138.9, 137.3, 134.3, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 107.9, 118.9, 114.7, 102.7, 69.5, 68.8, 58.6, 58.4, 56.7, 56.3, 56.2, 41.6, 37.9, 37.4, 36.4, 35.1, 34.8, 28.4, 28.4, 26.3, 25.0, 25.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.95$, (m/z) $[\text{M}+\text{H}]^+$ 925 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{49}\text{H}_{65}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 925.4646$, found 925.4637; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3415, 3304, 3075, 2927, 2826, 1652, 1609, 1539; $[\alpha]_{\text{D}}^{20} = -23$ (c 0.27, MeOH).

Methyl 10-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-10-oxodecanoate 141a

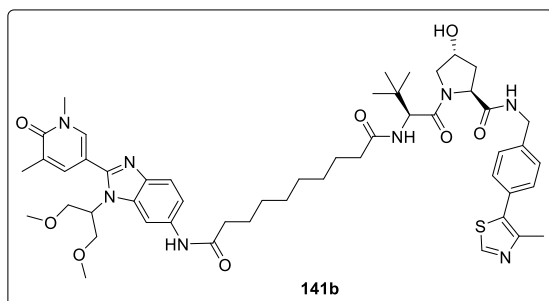


141a was synthesised following the **general procedure C** with 10-methoxy-10-oxodecanoic acid (179 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-40% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **141a** (373 mg, 0.593 mmol, 79% yield) was

isolated as a white solid.

m.p. 54-60 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 8.98 (s, 1 H), 8.52 (t, $J=6.1$ Hz, 1 H), 7.81 (d, $J=9.5$ Hz, 1 H), 7.42-7.38 (m, 4 H), 5.09 (d, $J=3.7$ Hz, 1 H), 4.54 (d, $J=9.5$ Hz, 1 H), 4.45-4.41 (m, 2 H), 4.36-4.34 (m, 1 H), 4.24-4.20 (m, 1 H), 3.68-3.62 (m, 2 H), 3.58 (s, 3 H), 2.44 (s, 3 H), 2.29-2.22 (m, 3 H), 2.12-2.08 (m, 1 H), 2.04-2.01 (m, 1 H), 1.93-1.88 (m, 1 H), 1.53-1.42 (m, 4 H), 1.26-1.20 (m, 8 H), 0.93 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 173.3, 172.0, 171.9, 169.7, 151.4, 147.7, 139.5, 131.1, 129.6, 128.6, 127.4, 68.8, 58.6, 56.3, 56.2, 51.1, 41.6, 37.9, 35.1, 34.8, 33.2, 28.5, 28.5, 28.5, 28.4, 26.3, 25.3, 24.4, 15.9; **LCMS** (Method B): $t_{\text{R}} = 1.11$, (m/z) $[\text{M}+\text{H}]^+ 629$ (100% purity); **HRMS** (ESI): calculated for $\text{C}_{33}\text{H}_{49}\text{N}_4\text{O}_6\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 629.3373$, found 629.3374; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3311, 3070, 2929, 2857, 1736, 1628, 1531, 1436; $[\alpha]_{\text{D}}^{20} = -13$ (c 1.00, MeOH).

N*1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*10-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)decanediamide **141b*

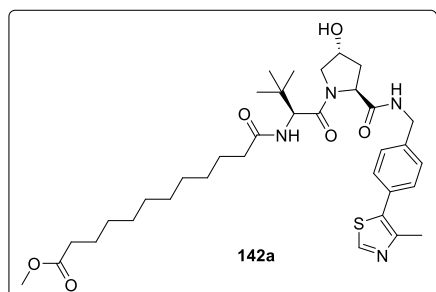


141b was synthesised following the **general procedure D** with the VHL-linker methyl 10-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-

10-oxodecanoate **141a** (362 mg, 0.576 mmol) in DCM. The crude product was purified by MDAP. **141b** (260 mg, 0.273 mmol, 47% yield) was isolated as a white solid.

m.p. 115-131 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 9.92 (s, 1 H), 8.97 (s, 1 H), 8.53 (t, $J=6.1$ Hz, 1 H), 8.23 (s, 1 H), 8.03-8.01 (m, 1 H), 7.81 (d, $J=9.5$ Hz, 1 H), 7.67-7.65 (m, 1 H), 7.54 (d, $J=8.4$ Hz, 1 H), 7.42-7.37 (m, 4 H), 7.32-7.30 (m, 1 H), 5.10 (s (br), 1 H), 4.85-4.81 (m, 1 H), 4.55 (d, $J=9.2$ Hz, 1 H), 4.44-4.41 (m, 2 H), 4.36-4.34 (m, 1 H), 4.24-4.20 (m, 1 H), 3.98-3.95 (m, 2 H), 3.77-3.75 (m, 2 H), 3.68-3.63 (m, 2 H), 3.53 (s, 3 H), 3.17 (s, 6 H), 2.44 (s, 3 H), 2.34-2.31 (m, 2 H), 2.28-2.23 (m, 1 H), 2.14-2.10 (m, 1 H), 2.08 (s, 3 H), 2.05-2.01 (m, 1 H), 1.93-1.89 (m, 1 H), 1.63-1.59 (m, 2 H), 1.54-1.44 (m, 2 H), 1.34-1.22 (m, 8 H), 0.93 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 172.5, 171.9, 171.0, 169.7, 161.7, 151.4, 151.3, 147.7, 139.4, 139.4, 139.1, 137.1, 134.5, 133.4, 131.1, 129.6, 128.6, 127.5, 127.4, 118.6, 115.0, 107.9, 102.8, 69.5, 68.6, 58.6, 58.4, 56.9, 56.3, 56.2, 41.6, 37.9, 37.5, 36.4, 35.1, 34.8, 28.7, 28.7, 28.6, 26.4, 26.3, 25.4, 25.0, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 1.03$, (m/z) $[\text{M}+\text{H}]^+$ 953 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{51}\text{H}_{69}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 953.4959$, found 953.4945; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3415, 3300, 3065, 2925, 2857, 1652, 1603, 1534; $[\alpha]_{\text{D}}^{20} = -20$ (c 1.00, MeOH).

Methyl 12-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-12-oxododecanoate 142a

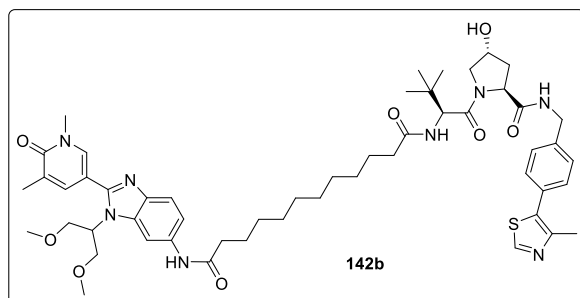


142a was synthesised following the **general procedure C** with 12-methoxy-12-oxododecanoic acid (202 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-40% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **142a**

(398 mg, 0.606 mmol, 81% yield) was isolated as a white solid.

m.p. 44-50 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 8.98 (s, 1 H), 8.52 (t, $J=6.1$ Hz, 1 H), 7.80 (d, $J=9.2$ Hz, 1 H), 7.43-7.37 (m, 4 H), 5.09 (d, $J=3.7$ Hz, 1 H), 4.54 (d, $J=9.2$ Hz, 1 H), 4.45-4.41 (m, 2 H), 4.37-4.33 (m, 1 H), 4.24-4.19 (m, 1 H), 3.70-3.62 (m, 2 H), 3.57 (s, 3 H), 2.45 (s, 3 H), 2.29-2.23 (m, 3 H), 2.12-2.08 (m, 1 H), 2.04-2.01 (m, 1 H), 1.93-1.88 (m, 1 H), 1.53-1.40 (m, 4 H), 1.27-1.19 (m, 12 H), 0.94 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 173.3, 172.0, 171.9, 167.7, 151.4, 147.7, 139.4, 131.1, 129.6, 128.6, 127.4, 68.8, 58.6, 56.2, 56.2, 51.1, 41.6, 37.9, 35.1, 34.8, 33.2, 28.8, 28.7, 28.6, 28.6, 28.6, 28.4, 26.3, 25.3, 24.4, 15.9; **LCMS** (Method B): $t_{\text{R}} = 1.22$, (m/z) $[\text{M}+\text{H}]^+$ 657 (98% purity); **HRMS** (ESI): calculated for $\text{C}_{35}\text{H}_{53}\text{N}_4\text{O}_6\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 657.3686$, found 657.3683; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3312, 3075, 2927, 2852, 1736, 1629, 1530, 1436; $[\alpha]_{\text{D}}^{20} = -29$ (c 1.00, MeOH).

N*1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*12-(((*S*)-1-(((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)dodecanediamide **142b*

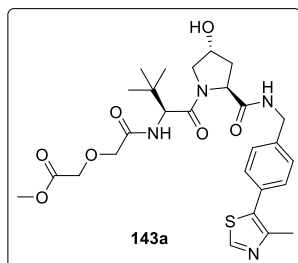


142b was synthesised following the **general procedure D** with the VHL-linker methyl 12-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-

dimethyl-1-oxobutan-2-yl)amino)-12-oxododecanoate **142a** (374 mg, 0.570 mmol) in DCM. The crude product was purified by MDAP. **142b** (282 mg, 0.287 mmol, 50% yield) was isolated as a white solid.

m.p. 122-128 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 9.87 (s, 1 H), 8.97 (s, 1 H), 8.52 (t, $J=6.2$ Hz, 1 H), 8.20 (s, 1 H), 8.02 (s, 1 H), 7.80 (d, $J=9.2$ Hz, 1 H), 7.65 (s, 1 H), 7.52 (d, $J=8.8$ Hz, 1 H), 7.44-7.37 (m, 4 H), 7.29 (d, $J=8.6$ Hz, 1 H), 5.10 (d, $J=3.7$ Hz, 1 H), 4.83-4.79 (m, 1 H), 4.54 (d, $J=9.5$ Hz, 1 H), 4.44-4.41 (m, 2 H), 4.37-4.34 (m, 1 H), 4.24-4.20 (m, 1 H), 3.97-3.94 (m, 2 H), 3.79-3.74 (m, 2 H), 3.68-3.63 (m, 2 H), 3.53 (s, 3 H), 3.17 (s, 6 H), 2.44 (s, 3 H), 2.34-2.30 (m, 2 H), 2.28-2.23 (m, 1 H), 2.12-2.07 (m, 4 H), 2.05-2.01 (m, 1 H), 1.93-1.88 (m, 1 H), 1.63-1.59 (m, 2 H), 1.54-1.40 (m, 2 H), 1.36-1.20 (m, 12 H), 0.94 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 172.0, 171.9, 171.0, 169.6, 161.7, 151.5, 151.3, 147.7, 139.4, 139.1, 138.9, 137.3, 134.3, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 118.9, 114.7, 108.0, 102.7, 69.5, 68.8, 58.6, 58.4, 56.7, 56.3, 56.2, 41.6, 37.9, 37.4, 36.4, 35.1, 34.8, 28.9, 28.8, 28.7, 28.7, 28.7, 28.6, 26.4, 25.4, 25.1, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 1.11$, (m/z) $[\text{M}+\text{H}]^+ 981$ (100% purity); **HRMS** (ESI): calculated for $\text{C}_{53}\text{H}_{73}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 981.5272$, found 981.5269; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3291, 3074, 2926, 2857, 1651, 1598, 1540, 1439; $[\alpha]_{\text{D}}^{20} = -25$ (c 1.00, MeOH).

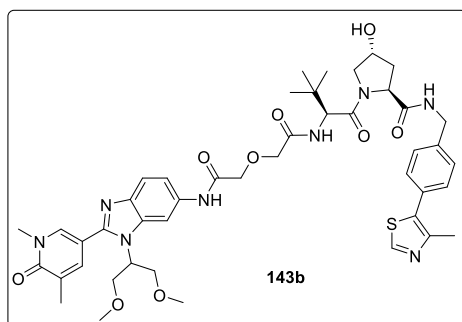
Methyl 2-(2-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-2-oxoethoxy)acetate **143a**



143a was synthesised following the **general procedure C** with 2-(2-methoxy-2-oxoethoxy)acetic acid (122 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-70% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **143a** (361 mg, 0.644 mmol, 86% yield) was isolated as a white solid.

m.p. 108-115 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 8.99 (s, 1 H), 8.57 (t, $J=6.1$ Hz, 1 H), 7.54 (d, $J=9.5$ Hz, 1 H), 7.42-7.39 (m, 4 H), 4.57 (d, $J=9.5$ Hz, 1 H), 4.46-4.38 (m, 2 H), 4.37-4.34 (m, 1 H), 4.27-4.23 (m, 3 H), 4.06-4.04 (m, 2 H), 3.68-3.65 (m, 4 H), 3.63-3.60 (m, 2 H), 2.45 (s, 3 H), 2.08-2.04 (m, 1 H), 1.93-1.89 (m, 1 H), 0.95 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.7, 170.2, 169.1, 167.9, 151.4, 147.6, 139.4, 131.1, 129.6, 128.6, 127.4, 69.7, 68.8, 67.7, 58.7, 56.5, 55.7, 51.5, 41.6, 37.9, 35.7, 26.1, 15.8; **LCMS** (Method B): $t_{\text{R}} = 0.86$, (m/z) $[\text{M}+\text{H}]^+$ 561 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{27}\text{H}_{37}\text{N}_4\text{O}_7\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 561.2383$, found 561.2383; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3413, 3323, 3069, 2960, 2878, 1751, 1672, 1640; $[\alpha]_{\text{D}}^{20} = -26$ (c 1.00, MeOH).

(2S,4R)-1-(((S)-2-(2-(2-(((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)amino)-2-oxoethoxy)acetamido)-3,3-dimethylbutanoyl)-4-hydroxy-N-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide **143b**

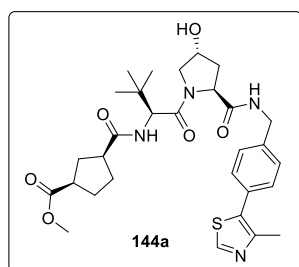


143b was synthesised following the **general procedure D** with the VHL-linker methyl 2-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-2-oxoethoxy)acetate **143a** (168 mg,

0.300 mmol) in DMF. The crude product was purified by MDAP. **143b** (134 mg, 0.151 mmol, 51% yield) was isolated as a white solid.

m.p. 148-153 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.24 (s, 1 H), 8.96 (s, 1 H), 8.57 (t, $J=6.1$ Hz, 1 H), 8.25 (s, 1 H), 8.05-8.00 (m, 2 H), 7.67 (m, 1 H), 7.56 (d, $J=8.8$ Hz, 1 H), 7.42-7.37 (m, 5 H), 5.17-5.13 (m, 1 H), 4.85-4.80 (m, 1 H), 4.62 (d, $J=9.2$ Hz, 1 H), 4.48-4.45 (m, 1 H), 4.43-4.37 (m, 2 H), 4.27-4.20 (m, 5 H), 3.99-3.96 (m, 2 H), 3.79-3.75 (m, 2 H), 3.73-3.70 (m, 1 H), 3.68-3.65 (m, 1 H), 3.54 (s, 3 H), 3.18-3.15 (m, 6 H), 2.44 (s, 3 H), 2.11 (s, 3 H), 2.08-2.04 (m, 1 H), 1.95-1.90 (m, 1 H), 0.99 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.7, 169.1, 168.9, 167.8, 161.7, 151.7, 151.4, 147.7, 139.4, 139.0, 137.2, 133.4, 133.3, 131.1, 129.6, 128.6, 127.9, 127.5, 127.4, 118.9, 115.1, 107.7, 103.3, 70.9, 70.5, 69.4, 68.8, 58.7, 58.4, 56.8, 56.5, 56.2, 41.6, 37.9, 37.5, 35.5, 26.2, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.85$, (m/z) $[\text{M}+\text{H}]^+$ 885 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{45}\text{H}_{57}\text{N}_8\text{O}_9\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 885.3969$, found 885.3969; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3286, 3074, 2926, 2868, 1656, 1603, 1540, 1429; $[\alpha]_{\text{D}}^{20} = -14$ (c 1.00, MeOH).

Methyl (1R,3S)-3-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)cyclopentane-1-carboxylate 144a



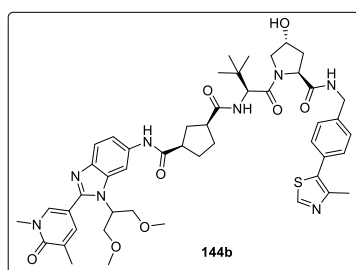
144a was synthesised following the **general procedure C** with (1S,3R)-3-(methoxycarbonyl)cyclopentane-1-carboxylic acid (142 g, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-70% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **144a** (349 mg, 0.597 mmol, 79% yield) was isolated as a white

solid.

m.p. 78-83 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 8.98 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 7.81 (d, $J=9.2$ Hz, 1 H), 7.42-7.38 (m, 4 H), 4.54 (d, $J=9.2$ Hz, 1 H), 4.45-4.41 (m, 2 H), 4.37-4.34 (m, 1 H), 4.24-4.20 (m, 1 H), 3.68-3.62 (m, 3 H), 3.59 (s, 3 H), 2.90-2.84 (m, 1 H), 2.80-2.74 (m, 1 H), 2.44 (s, 3 H), 2.05-1.97 (m, 2 H), 1.93-1.77 (m,

5 H), 1.68-1.62 (m, 1 H), 0.93 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 175.3, 173.9, 171.9, 169.6, 151.4, 147.6, 138.5, 131.1, 129.6, 128.6, 127.4, 68.8, 58.7, 56.3, 56.3, 51.3, 43.6, 43.1, 41.6, 37.9, 35.3, 32.8, 30.2, 28.9, 26.3, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.94$, (m/z) $[\text{M}+\text{H}]^+$ 585 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{30}\text{H}_{41}\text{N}_4\text{O}_6\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 585.2747$, found 585.2742; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3413, 3296, 3079, 2958, 2873, 1718, 1625, 1535; $[\alpha]_{\text{D}}^{20} = -48$ (c 1.00, MeOH).

(1R,3S)-N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N3-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)cyclopentane-1,3-dicarboxamide 144b



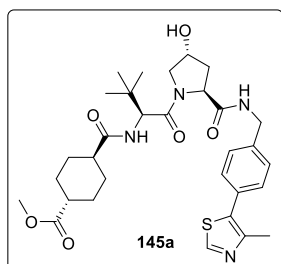
144b was synthesised following the **general procedure D** with the VHL-linker methyl **(1R,3S)-3-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)cyclopentane-1-carboxylate 144a** (175 mg, 0.300 mmol) in DMF. The crude product

was purified by MDAP. **144b** (109 mg, 0.120 mmol, 40% yield) was isolated as a yellow solid.

m.p. 165-169 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.99 (s, 1 H), 8.97 (s, 1 H), 8.53 (t, $J=6.1$ Hz, 1 H), 8.26 (s, 1 H), 8.04 (s, 1 H), 7.86 (d, $J=9.2$ Hz, 1 H), 7.66 (s, 1 H), 7.54 (d, $J=8.4$ Hz, 1 H), 7.42-7.38 (m, 4 H), 7.32 (d, $J=8.6$ Hz, 1 H), 5.14-5.10 (m, 1 H), 4.85-4.80 (m, 1 H), 4.57 (d, $J=9.2$ Hz, 1 H), 4.46-4.40 (m, 2 H), 4.38-4.34 (m, 1 H), 4.25-4.21 (m, 1 H), 3.99-3.94 (m, 2 H), 3.78-3.75 (m, 2 H), 3.70-3.65 (m, 2 H), 3.53 (s, 3 H), 3.18-3.16 (m, 6 H), 2.96-2.91 (m, 1 H), 2.85-2.81 (m, 1 H), 2.44 (s, 3 H), 2.11 (s, 3 H), 2.06-1.99 (m, 3 H), 1.93-1.85 (m, 4 H), 1.78-1.73 (m, 1 H), 0.95 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 174.4, 173.4, 171.9, 169.6, 161.7, 151.4, 151.3, 147.7, 139.4, 139.1, 138.4, 137.2, 134.5, 133.2, 131.1, 129.6, 128.6, 127.5, 127.4, 118.7, 115.0, 107.5, 102.8, 69.4, 68.8, 58.7, 58.4, 56.9, 56.4, 56.3, 45.8, 44.0, 41.6, 37.9, 37.5, 35.4, 33.6, 30.4, 29.7, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.91$, (m/z) $[\text{M}+\text{H}]^+$ 909 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{48}\text{H}_{61}\text{N}_8\text{O}_8\text{S}$ (m/z)

$[M+2H]^{2+} = 909.4333$, found 909.4319; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3286, 3079, 2937, 2878, 1656, 1603, 1540, 1487; $[\alpha]_D^{20} = -71$ (c 1.00, MeOH).

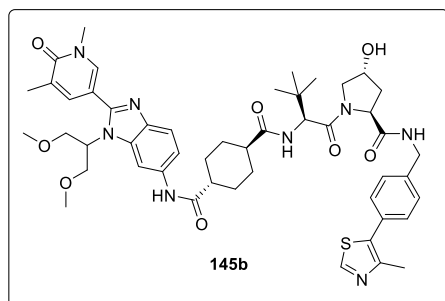
Methyl (1*S*,4*R*)-4-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)cyclohexane-1-carboxylate **145a**



145a was synthesised following the **general procedure C** with (1*R*,4*R*)-4-(methoxycarbonyl)cyclohexane-1-carboxylic acid (154 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-40% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **145a** (373 mg, 0.623 mmol, 83% yield) was isolated as a white solid.

m.p. 110-118 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 9.01 (s, 1 H), 8.53 (t, $J=6.1$ Hz, 1 H), 7.71 (d, $J=9.2$ Hz, 1 H), 7.45-7.38 (m, 4 H), 5.10 (d, $J=3.7$ Hz, 1 H), 4.51 (d, $J=9.5$ Hz, 1 H), 4.44-4.41 (m, 2 H), 4.37-4.33 (m, 1 H), 4.24-4.20 (m, 1 H), 3.68-3.61 (m, 2 H), 3.58 (s, 3 H), 2.45 (s, 3 H), 2.38-2.32 (m, 1 H), 2.29-2.23 (m, 1 H), 2.05-2.01 (m, 1 H), 1.93-1.87 (m, 3 H), 1.82-1.77 (m, 1 H), 1.70-1.66 (m, 1 H), 1.42-1.26 (m, 4 H), 0.93 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 175.3, 174.6, 171.9, 169.6, 151.4, 147.7, 139.4, 131.1, 129.6, 128.6, 127.4, 68.8, 58.7, 56.3, 56.1, 51.2, 42.3, 41.7, 41.7, 37.9, 35.3, 28.9, 28.0, 27.7, 27.6, 26.3, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.97$, (m/z) $[M+H]^+ 599$ (100% purity); **HRMS** (ESI): calculated for $\text{C}_{31}\text{H}_{43}\text{N}_4\text{O}_6\text{S}$ (m/z) $[M+H]^+ = 599.2903$, found 599.2905; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3300, 3081, 2951, 2868, 1728, 1626, 1524, 1435; $[\alpha]_D^{20} = -24$ (c 1.00, MeOH).

(1*S*,4*R*)-N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-N4-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)cyclohexane-1,4-dicarboxamide 145b

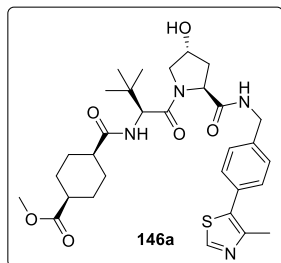


145b was synthesised following the **general procedure D** with the VHL-linker methyl (1*S*,4*R*)-4-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)cyclohexane-1-carboxylate **145a**

(180 mg, 0.300 mmol) in DCM. The crude product was purified by MDAP. **145b** (118 mg, 0.128 mmol, 43% yield) was isolated as a white solid.

m.p. 170-175 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 9.89 (s, 1 H), 8.99 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 8.22 (s, 1 H), 8.03 (s, 1 H), 7.79 (d, $J=9.5$ Hz, 1 H), 7.66 (s, 1 H), 7.53 (d, $J=8.8$ Hz, 1 H), 7.43-7.38 (m, 4 H), 7.32 (d, $J=8.8$ Hz, 1 H), 5.15-5.12 (m, 1 H), 4.84-4.80 (m, 1 H), 4.53 (d, $J=9.5$ Hz, 1 H), 4.47-4.42 (m, 2 H), 4.39-4.35 (m, 1 H), 4.27-4.21 (m, 1 H), 3.98-3.95 (m, 2 H), 3.79-3.75 (m, 2 H), 3.70-3.62 (m, 2 H), 3.53 (s, 3 H), 3.18 (s, 6 H), 2.47-2.40 (m, 4 H), 2.36-2.30 (m, 1 H), 2.09 (s, 3 H), 2.07-2.02 (m, 1 H), 1.94-1.83 (m, 4 H), 1.78-1.72 (m, 1 H), 1.53-1.36 (m, 4 H), 0.95 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 174.8, 173.9, 171.9, 169.6, 161.7, 151.4, 151.4, 147.7, 139.5, 139.0, 138.7, 137.2, 134.5, 133.3, 131.1, 129.6, 128.6, 127.4, 127.4, 118.7, 114.8, 107.6, 102.6, 69.4, 68.8, 58.7, 58.4, 56.8, 56.3, 56.1, 44.3, 42.4, 41.6, 37.9, 37.5, 35.3, 29.2, 28.4, 28.3, 27.9, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.94$, (m/z) $[\text{M}+\text{H}]^+$ 923 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{49}\text{H}_{63}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 923.4490$, found 923.4478; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3286, 3064, 2937, 2868, 1651, 1603, 1519, 1482; $[\alpha]_{\text{D}}^{20} = -25$ (c 1.00, MeOH).

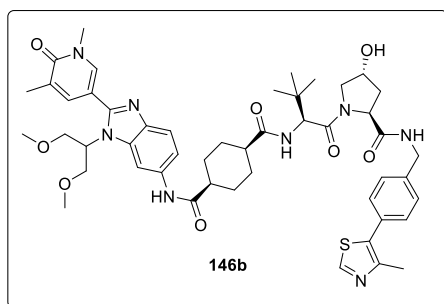
Methyl (1*R*,4*S*)-4-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)cyclohexane-1-carboxylate **146a**



146a was synthesised following the **general procedure C** with (1*S*,4*S*)-4-(methoxycarbonyl)cyclohexane-1-carboxylic acid (154 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-40% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **146a** (363 mg, 0.606 mmol, 81% yield) was isolated as a white solid.

m.p. 110-115 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 8.97 (s, 1 H), 8.52 (t, $J=6.1$ Hz, 1 H), 7.66 (d, $J=9.5$ Hz, 1 H), 7.43-7.37 (m, 4 H), 5.10 (d, $J=3.7$ Hz, 1 H), 4.50 (d, $J=9.2$ Hz, 1 H), 4.44-4.40 (m, 2 H), 4.36-4.32 (m, 1 H), 4.25-4.20 (m, 1 H), 3.67-3.62 (m, 2 H), 3.60 (s, 3 H), 2.59-2.55 (m, 1 H), 2.45-2.42 (m, 4 H), 2.05-2.00 (m, 2 H), 1.93-1.88 (m, 2 H), 1.61-1.45 (m, 6 H), 0.92 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 174.5, 174.6, 171.9, 169.6, 151.4, 147.7, 139.5, 131.1, 129.6, 128.6, 127.4, 68.8, 58.7, 56.3, 56.0, 51.3, 41.6, 41.1, 38.9, 37.9, 35.3, 26.9, 26.3, 25.8, 25.7, 25.4, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.98$, (m/z) $[\text{M}+\text{H}]^+$ 599 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{31}\text{H}_{43}\text{N}_4\text{O}_6\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 599.2903$, found 599.2903; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3299, 3070, 2951, 2873, 1727, 1626, 1524, 1435; $[\alpha]_{\text{D}}^{20} = -34$ (c 1.00, MeOH).

(1*R*,4*S*)-*N*1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*4-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)cyclohexane-1,4-dicarboxamide **146b**

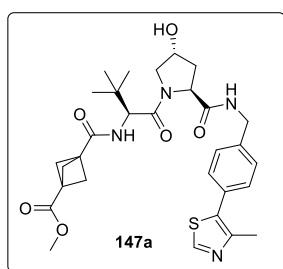


146b was synthesised following the **general procedure D** with the VHL-linker methyl (1*R*,4*S*)-4-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)cyclohexane-1-carboxylate **146a**

(180 mg, 0.300 mmol) in DCM. The crude product was purified by MDAP. **146b** (126 mg, 0.136 mmol, 46% yield) was isolated as a white solid.

m.p. 160-168 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.80 (s, 1 H), 8.98 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 8.24 (s, 1 H), 8.05 (s, 1 H), 7.67 (s, 1 H), 7.63 (d, $J=9.2$ Hz, 1 H), 7.54 (d, $J=8.8$ Hz, 1 H), 7.43-7.38 (m, 4 H), 7.35 (d, $J=8.6$ Hz, 1 H), 5.16-5.12 (m, 1 H), 4.86-4.82 (m, 1 H), 4.56 (d, $J=9.5$ Hz, 1 H), 4.45-4.41 (m, 2 H), 4.38-4.34 (m, 1 H), 4.26-4.20 (m, 1 H), 4.00-3.95 (m, 2 H), 3.79-3.74 (m, 2 H), 3.70-3.63 (m, 2 H), 3.54 (s, 3 H), 3.18 (s, 6 H), 2.58-2.53 (m, 1 H), 2.53-2.51 (m, 1 H), 2.45 (s, 3 H), 2.09 (s, 3 H), 2.06-2.01 (m, 2 H), 1.95-1.83 (m, 4 H), 1.64-1.56 (m, 3 H), 1.51-1.43 (m, 1 H), 0.95 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 174.6, 173.9, 171.9, 169.7, 161.7, 151.4, 151.2, 147.7, 139.5, 139.2, 137.9, 137.1, 134.8, 133.1, 131.1, 129.6, 128.6, 127.5, 127.4, 118.4, 115.2, 107.1, 102.8, 69.4, 68.8, 58.7, 58.5, 56.9, 56.3, 56.0, 41.8, 41.6, 39.5, 37.9, 37.5, 35.4, 27.4, 26.4, 26.1, 26.0, 25.7, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.96$, (m/z) $[M+H]^+$ 923 (98% purity); **HRMS** (ESI): calculated for $C_{49}H_{63}N_8O_8S$ (m/z) $[M+H]^+ = 923.4490$, found 923.4456; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3291, 3064, 2921, 2867, 1651, 1603, 1535, 1482; $[\alpha]_D^{20} = -51$ (c 1.00, MeOH).

Methyl 3-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)bicyclo[1.1.1]pentane-1-carboxylate 147a

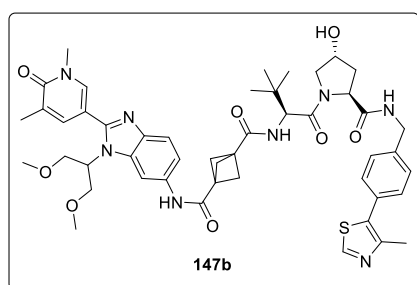


147a was synthesised following the **general procedure C** with 3-(methoxycarbonyl)bicyclo[1.1.1]pentane-1-carboxylic acid (141 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-70% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **147a** (341 mg, 0.585 mmol, 78% yield) was isolated as a white solid.

m.p. 123-130 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.09 (s, 1 H), 8.56 (t, $J=6.1$ Hz, 1 H), 7.49 (d, $J=9.2$ Hz, 1 H), 7.42-7.39 (m, 4 H), 4.58 (d, $J=9.5$ Hz, 1 H), 4.46-4.40 (m, 2 H), 4.36-4.34 (m, 1 H), 4.26-4.22 (m, 1 H), 3.78-3.74 (m, 1 H), 3.68-3.59 (m, 5 H), 2.45 (s, 3 H), 2.24-2.21 (s, 6 H), 2.07-2.03 (m, 1 H), 1.93-1.89 (m, 1 H), 0.94 (s,

9 H); δ_C (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.4, 169.1, 168.0, 151.5, 147.5, 139.5, 131.2, 129.5, 128.6, 127.4, 68.8, 58.7, 56.4, 56.2, 51.7, 51.4, 41.7, 38.7, 37.9, 36.3, 35.6, 26.3, 15.8; **LCMS** (Method B): $t_R = 0.91$, (m/z) $[M+H]^+$ 583 (100% purity); **HRMS** (ESI): calculated for $C_{30}H_{39}N_4O_6S$ (m/z) $[M+H]^+ = 583.2590$, found 583.2593; **IR** (ν_{max}/cm^{-1}) 3296, 3074, 2963, 2878, 1725, 1620, 1535, 1439; $[\alpha]_D^{20} = -32$ (c 1.00, MeOH).

N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N3-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)bicyclo[1.1.1]pentane-1,3-dicarboxamide 147b



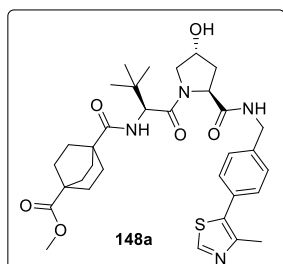
147b was synthesised following the **general procedure D** with the VHL-linker methyl 3-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)bicyclo [1.1.1]pentane-1-carboxylate **147a** (175 mg,

0.300 mmol) in DMF. The crude product was purified by MDAP. **147b** (184 mg, 0.203 mmol, 68% yield) was isolated as a white solid.

m.p. 181-187 °C; δ_H (400 MHz, d_6 -DMSO) δ ppm 9.63 (s, 1 H), 8.99 (s, 1 H), 8.56 (t, $J=6.1$ Hz, 1 H), 8.21 (s, 1 H), 8.05 (s, 1 H), 7.67 (s, 1 H), 7.56 (d, $J=8.8$ Hz, 1 H), 7.47-7.39 (m, 6 H), 5.18-5.12 (m, 1 H), 4.86-4.81 (m, 1 H), 4.60 (d, $J=9.5$ Hz, 1 H), 4.48-4.39 (m, 2 H), 4.38-4.35 (m, 1 H), 4.27-4.23 (m, 1 H), 4.00-3.95 (m, 2 H), 3.78-3.74 (m, 2 H), 3.70-3.66 (m, 1 H), 3.64-3.60 (m, 1 H), 3.53 (s, 3 H), 3.17 (s, 6 H), 2.45 (s, 3 H), 2.30 (s, 6 H), 2.09 (s, 3 H), 2.08-2.04 (m, 1 H), 1.94-1.90 (m, 1 H), 0.96 (s, 9 H); δ_C (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.1, 168.5, 167.7, 161.7, 151.6, 151.4, 147.7, 139.4, 139.1, 138.8, 137.2, 133.7, 133.1, 131.0, 129.6, 128.6, 127.5, 127.4, 118.5, 115.6, 107.5, 103.7, 69.4, 68.8, 58.7, 58.4, 56.9, 56.4, 56.2, 51.6, 51.6, 41.7, 38.9, 37.9, 37.9, 37.5, 35.7, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_R = 0.89$, (m/z) $[M+H]^+$ 907 (100% purity); **HRMS** (ESI): calculated for $C_{48}H_{59}N_8O_8S$ (m/z)

$[M+2H]^{2+} = 454.2127$, found 454.2133; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3275, 3074, 2963, 2926, 2878, 1651, 1603, 1519; $[\alpha]_D^{20} = -16$ (*c* 1.00, MeOH).

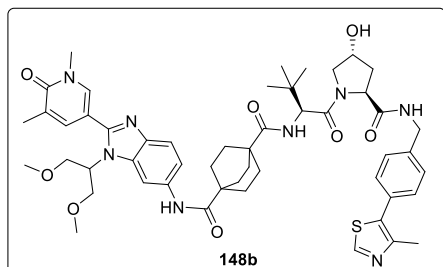
Methyl 4-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)bicyclo[2.2.2]octane-1-carboxylate **148a**



148a was synthesised following the **general procedure C** with 4-(methoxycarbonyl)bicyclo[2.2.2]octane-1-carboxylic acid (175 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-50% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **148a** (340 mg, 0.544 mmol, 72% yield) was isolated as a white solid.

m.p. 112-118 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 8.98 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 7.42-7.38 (m, 4 H), 6.72 (d, $J=9.5$ Hz, 1 H), 4.54 (d, $J=9.2$ Hz, 1 H), 4.45-4.32 (m, 3 H), 4.34-4.23 (m, 1 H), 3.65-3.62 (m, 1 H), 3.59-3.55 (m, 5 H), 2.45 (s, 3 H), 2.06-2.01 (m, 1 H), 1.92-1.88 (m, 1 H), 1.73-1.68 (m, 12 H), 0.92 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 177.1, 175.6, 171.8, 169.5, 151.4, 147.6, 139.4, 131.1, 129.6, 128.6, 127.4, 68.8, 58.7, 56.3, 56.0, 51.4, 41.7, 38.2, 38.2, 37.8, 35.8, 27.6, 27.4, 26.3, 15.8; **LCMS** (Method B): $t_{\text{R}} = 1.02$, (m/z) $[M+H]^+ 625$ (100% purity); **HRMS** (ESI): calculated for $\text{C}_{33}\text{H}_{45}\text{N}_4\text{O}_6\text{S}$ (m/z) $[M+H]^+ = 625.3060$, found 625.3058; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3307, 3079, 2952, 2868, 1725, 1619, 1513, 1435; $[\alpha]_D^{20} = -41$ (*c* 1.00, MeOH).

N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N4*-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)bicyclo[2.2.2]octane-1,4-dicarboxamide 148b

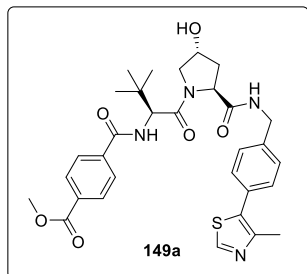


148b was synthesised following the **general procedure D** with the VHL-linker methyl 4-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)bicyclo[2.2.2]octane-1-

carboxylate **148a** (187 mg, 0.300 mmol) in DMF. The crude product was purified by MDAP. **148b** (150 mg, 0.158 mmol, 53% yield) was isolated as a white solid.

m.p. 177-182 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.18 (s, 1 H), 8.98 (s, 1 H), 8.56 (t, $J=6.1$ Hz, 1 H), 8.13 (s, 1 H), 8.05 (s, 1 H), 7.67 (s, 1 H), 7.53-7.46 (m, 2 H), 7.42-7.39 (m, 4 H), 6.75 (d, $J=9.2$ Hz, 1 H), 5.17-5.12 (m, 1 H), 4.85-4.80 (m, 1 H), 4.57 (d, $J=9.2$ Hz, 1 H), 4.47-4.44 (m, 1 H), 4.41-4.35 (m, 2 H), 4.29-4.25 (m, 1 H), 4.00-3.95 (m, 2 H), 3.79-3.75 (m, 2 H), 3.68-3.65 (m, 1 H), 3.61-3.58 (m, 1 H), 3.55 (s, 3 H), 3.18 (s, 6 H), 2.45 (s, 3 H), 2.09 (s, 3 H), 2.07-2.04 (m, 1 H), 1.94-1.90 (m, 1 H), 1.89-1.84 (m, 6 H), 1.80-1.76 (m, 6 H), 0.95 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 175.8, 175.5, 171.8, 169.5, 161.7, 151.4, 151.4, 147.7, 139.4, 139.0, 138.9, 137.2, 134.3, 131.1, 133.1, 129.7, 128.7, 127.5, 127.5, 118.4, 115.9, 107.7, 103.7, 69.4, 68.8, 58.8, 58.4, 56.8, 56.4, 56.0, 41.7, 39.2, 38.4, 37.8, 37.5, 35.9, 27.9, 27.7, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.95$, (m/z) $[\text{M}+\text{H}]^+$ 949 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{51}\text{H}_{65}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+2\text{H}]^{2+} = 475.2362$, found 475.2372; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3296, 3074, 2921, 2873, 1651, 1619, 1513, 1476; $[\alpha]_{\text{D}}^{20} = -25$ (c 1.00, MeOH).

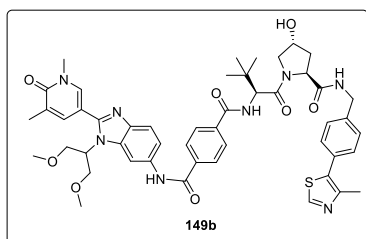
Methyl 4-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)benzoate
149a



149a was synthesised following the **general procedure C** with 4-(methoxycarbonyl)benzoic acid (149 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-50% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **149a** (343 mg, 0.579 mmol, 77% yield) was isolated as a white solid.

m.p. 152-158 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 8.98 (s, 1 H), 8.56 (t, $J=6.1$ Hz, 1 H), 8.26 (d, $J=8.8$ Hz, 1 H), 8.03-8.01 (m, 2 H), 7.98-7.97 (m, 2 H), 7.43-7.39 (m, 4 H), 5.13 (d, $J=3.3$ Hz, 1 H), 4.78 (d, $J=9.2$ Hz, 1 H), 4.48-4.36 (m, 3 H), 4.26-4.22 (m, 1 H), 3.88 (s, 3 H), 3.75-3.72 (m, 2 H), 2.45 (s, 3 H), 2.07-2.04 (m, 1 H), 1.95-1.91 (m, 1 H), 1.04 (m, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.2, 165.8, 165.7, 151.4, 147.7, 139.4, 138.2, 131.7, 131.1, 129.6, 128.9, 128.6, 128.1, 127.4, 68.9, 58.8, 57.5, 56.4, 52.3, 41.6, 37.9, 35.5, 26.5, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.99$, (m/z) $[\text{M}+\text{H}]^+$ 593 (91% purity); **HRMS** (ESI): calculated for $\text{C}_{31}\text{H}_{37}\text{N}_4\text{O}_6\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 594.2434$, found 593.2444; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3314, 3075, 2954, 2879, 1721, 1626, 1521, 1277; $[\alpha]_{\text{D}}^{20} = -48$ (c 1.00, MeOH).

N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-N4-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)terephthalamide **149b**

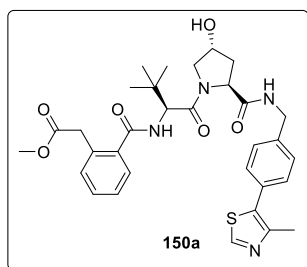


149b was synthesised following the **general procedure D** with the VHL-linker methyl 4-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)benzoate **149a** (327 mg,

0.551 mmol) in DCM. The crude product was purified by MDAP. **149b** (211 mg, 0.230 mmol, 42% yield) was isolated as a white solid.

m.p. 155-165 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.42 (s, 1 H), 9.89 (s, 1 H), 8.57 (t, $J=6.1$ Hz, 1 H), 8.34 (s, 1 H), 8.22 (d, $J=8.8$ Hz, 1 H), 8.09-8.03 (m, 5 H), 7.69-7.68 (m, 1 H), 7.61-7.59 (m, 2 H), 7.44-7.40 (m, 4 H), 5.16 (s (br), 1 H), 4.89-4.85 (m, 1 H), 4.82 (d, $J=8.8$ Hz, 1 H), 4.50-4.39 (m, 3 H), 4.30-4.24 (m, 1 H), 4.05-4.01 (m, 2 H), 3.81-3.79 (m, 2 H), 3.77-3.74 (m, 2 H), 3.55 (s, 3 H), 3.20 (s, 6 H), 2.45 (s, 3 H), 2.10 (s, 3 H), 2.08-2.05 (m, 1 H), 1.97-1.92 (m, 1 H), 1.07 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.3, 165.9, 164.6, 161.7, 151.6, 151.4, 147.7, 139.4, 139.3, 139.3, 137.2, 137.1, 136.6, 134.1, 133.1, 131.1, 129.6, 128.7, 127.7, 127.5, 127.4, 127.4, 118.4, 116.5, 107.9, 104.5, 69.4, 68.9, 58.8, 58.4, 57.4, 57.0, 56.4, 41.7, 37.9, 37.5, 35.6, 26.5, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.97$, (m/z) $[\text{M}+\text{H}]^+$ 917 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{49}\text{H}_{57}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 917.4020$, found 917.4009; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3415, 3288, 3059, 2927, 1652, 1609, 1540, 1483; $[\alpha]_{\text{D}}^{20} = -23$ (c 1.00, MeOH).

Methyl 2-(2-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)phenyl)acetate **150a**

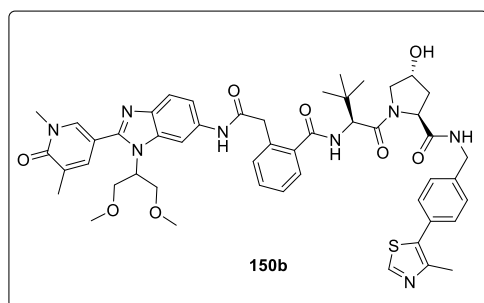


150a was synthesised following the **general procedure C** with 2-(2-methoxy-2-oxoethyl)benzoic acid (161 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-50% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **150a** (443 mg, 0.730 mmol, 97% yield) was isolated as a white solid.

m.p. 102-108 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 8.97 (s, 1 H), 8.55 (t, $J=6.1$ Hz, 1 H), 7.93 (d, $J=8.8$ Hz, 1 H), 7.46-7.29 (m, 8 H), 5.11 (d, $J=3.7$ Hz, 1 H), 4.69 (d, $J=9.2$ Hz, 1 H), 4.48 (t, $J=8.1$ Hz, 1 H), 4.43-4.37 (m, 2 H), 4.27-4.24 (m, 1 H), 3.38-3.31 (m, 2 H), 3.74-3.68 (m, 2 H), 3.57 (s, 3 H), 2.44 (s, 3 H), 2.08-2.03 (m, 1 H), 1.95-1.90 (m, 1 H), 1.01 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.8, 171.4, 169.4,

168.2, 151.4, 147.7, 139.4, 132.6, 131.2, 131.1, 131.1, 129.7, 129.6, 128.6, 127.9, 127.4, 126.8, 68.8, 58.8, 57.1, 56.3, 51.5, 41.6, 38.0, 37.9, 35.4, 26.4, 15.9; **LCMS** (Method B): $t_R = 0.99$, (m/z) $[M+H]^+$ 607 (100% purity); **HRMS** (ESI): calculated for $C_{32}H_{39}N_4O_6S$ (m/z) $[M+H]^+ = 607.2590$, found 607.2592; **IR** (ν_{max}/cm^{-1}) 3304, 3070, 2955, 2873, 1732, 1626, 1520, 1434; $[\alpha]_D^{20} = -34$ (c 1.00, MeOH).

(2*S*,4*R*)-1-((*S*)-2-(2-(2-((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)amino)-2-oxoethyl)benzamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide 150b



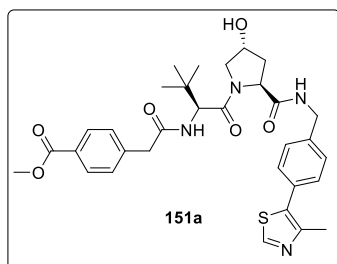
150b was synthesised following the **general procedure D** with the VHL-linker methyl 2-(2-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)phenyl)acetate **150a** (180 mg,

0.300 mmol) in DCM. The crude product was purified by MDAP. **150b** (100 mg, 0.104 mmol, 35% yield) was isolated as a yellow solid.

m.p. 157-165 °C; δ_H (400 MHz, d_6 -DMSO) δ ppm 10.56 (s, 1 H), 8.91 (s, 2 H), 8.53 (t, $J=6.1$ Hz, 1 H), 8.31 (s, 1 H), 8.03 (s, 1 H), 7.65 (s, 1 H), 7.56 (d, $J=8.4$ Hz, 1 H), 7.50-7.34 (m, 8 H), 7.26 (d, $J=8.3$ Hz, 1 H), 5.15-5.12 (m, 1 H), 4.83-4.78 (m, 1 H), 4.65 (d, $J=8.4$ Hz, 1 H), 4.51 (t, $J=8.1$ Hz, 1 H), 4.43-4.38 (m, 1 H), 4.37-4.32 (m, 1 H), 4.31-4.25 (m, 1 H), 3.98-3.86 (m, 4 H), 3.85-3.81 (m, 1 H), 3.79-3.69 (m, 3 H), 3.53 (s, 3 H), 3.13 (s, 3 H), 3.11 (s, 3 H), 2.38 (s, 3 H), 2.09 (s, 3 H), 2.07-2.03 (m, 1 H), 1.97-1.91 (m, 1 H), 1.05 (s, 9 H); δ_C (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.8, 168.4, 168.4, 161.7, 151.6, 151.3, 147.7, 139.4, 139.2, 138.7, 137.1, 136.8, 133.9, 133.2, 133.1, 131.0, 130.3, 129.9, 129.7, 128.7, 128.6, 127.5, 127.4, 126.8, 118.8, 115.1, 107.2, 103.2, 69.4, 68.9, 58.8, 58.4, 58.4, 56.9, 56.3, 41.6, 40.6, 37.9, 37.5, 34.7, 26.5, 16.9, 15.8; **LCMS** (Method B): $t_R = 0.97$, (m/z) $[M+H]^+$ 931 (97% purity); **HRMS** (ESI): calculated for $C_{50}H_{59}N_8O_8S$ (m/z) $[M+H]^+ = 931.4177$, found 931.4150;

IR ($\nu_{\max}/\text{cm}^{-1}$) 3259, 3064, 2926, 2868, 1651, 1603, 1540, 1487; $[\alpha]_D^{20} = -35$ (*c* 1.00, MeOH).

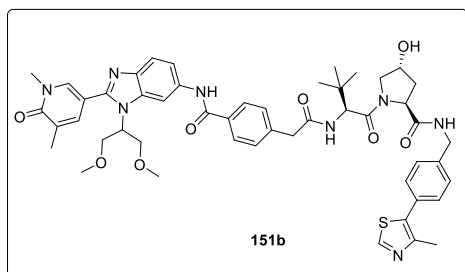
Methyl 4-(2-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl) carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-2-oxoethyl) benzoate **151a**



151a was synthesised following the **general procedure C** with 2-(4-(methoxycarbonyl)phenyl)acetic acid (161 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-55% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **151a** (354 mg, 0.554 mmol, 74% yield) was isolated as a white solid.

m.p. 106-110 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 8.98 (s, 1 H), 8.55 (t, $J=6.1$ Hz, 1 H), 8.20 (d, $J=9.5$ Hz, 1 H), 7.90-7.87 (m, 2 H), 7.48-7.38 (m, 6 H), 5.10 (d, $J=3.3$ Hz, 1 H), 4.52 (d, $J=9.5$ Hz, 1 H), 4.45-4.41 (m, 2 H), 4.36-4.32 (m, 1 H), 4.24-4.20 (m, 1 H), 3.83 (s, 3 H), 3.76-3.74 (m, 1 H), 3.67-3.60 (m, 2 H), 3.57-3.55 (m, 1 H), 2.45 (s, 3 H), 2.05-2.02 (m, 1 H), 1.92-1.88 (m, 1 H), 0.91 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.4, 169.2, 166.1, 151.4, 147.7, 142.3, 139.5, 131.1, 129.6, 129.4, 129.0, 128.9, 128.6, 127.4, 68.4, 58.7, 56.5, 56.3, 52.0, 41.7, 41.6, 37.9, 35.4, 26.3, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.98$, (m/z) $[\text{M}+\text{H}]^+$ 607 (96% purity); **HRMS** (ESI): calculated for $\text{C}_{32}\text{H}_{39}\text{N}_4\text{O}_6\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 607.2590$, found 607.2586; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3300, 3070, 2958, 2879, 1715, 1627, 1531, 1435; $[\alpha]_D^{20} = -41$ (*c* 1.00, MeOH).

(2*S*,4*R*)-1-(((*S*)-2-(2-(4-((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)carbamoyl)phenyl)acetamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide **151b**

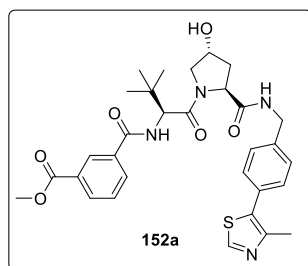


151b was synthesised following the **general procedure D** with the VHL-linker methyl 4-(2-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-2-oxoethyl)benzoate **151a** (180 mg,

0.300 mmol) in DCM. The crude product was purified by MDAP. **151b** (63 mg, 0.068 mmol, 23% yield) was isolated as a white solid.

m.p. 169-175 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.24 (s, 1 H), 8.98 (s, 1 H), 8.56 (t, $J=6.1$ Hz, 1 H), 8.32 (s, 1 H), 8.20 (d, $J=9.1$ Hz, 1 H), 8.07 (s, 1 H), 7.95-7.90 (m, 2 H), 7.69 (s, 1 H), 7.63-7.59 (m, 2 H), 7.49-7.35 (m, 6 H), 5.15-5.12 (m, 1 H), 4.88-4.84 (m, 1 H), 4.55 (d, $J=9.5$ Hz, 1 H), 4.46-4.42 (m, 2 H), 4.37-4.34 (m, 1 H), 4.25-4.21 (m, 1 H), 4.04-4.01 (m, 2 H), 3.81-3.76 (m, 3 H), 3.69-3.63 (m, 2 H), 3.58-3.53 (m, 4 H), 3.19 (s, 6 H), 2.45 (s, 3 H), 2.11 (s, 3 H), 2.06-2.02 (m, 1 H), 1.94-1.89 (m, 1 H), 0.94 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.9, 169.5, 169.5, 165.3, 161.7, 151.6, 151.4, 147.7, 140.4, 139.5, 139.2, 138.8, 137.2, 134.2, 133.2, 133.1, 131.1, 129.6, 129.0, 128.6, 127.5, 127.5, 127.4, 118.5, 116.2, 107.5, 104.2, 69.4, 68.8, 58.7, 58.7, 58.7, 57.0, 56.5, 41.6, 41.6, 37.9, 37.5, 35.5, 26.3, 17.0, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.95$, (m/z) $[\text{M}+\text{H}]^+$ 931 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{50}\text{H}_{59}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 931.4177$, found 931.4181; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3302, 3064, 2931, 2869, 1646, 1609, 1540, 1482; $[\alpha]_{\text{D}}^{20} = -33$ (c 1.00, MeOH).

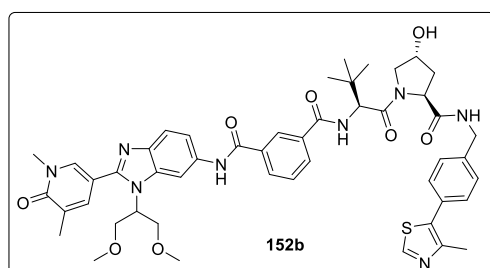
Methyl 3-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)benzoate
152a



152a was synthesised following the **general procedure C** with 3-(methoxycarbonyl)benzoic acid (149 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-55% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **152a** (374 mg, 0.631 mmol, 84% yield) was isolated as a white solid.

m.p. 107-111 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 8.98 (s, 1 H), 8.56 (t, $J=6.1$ Hz, 1 H), 8.39 (s, 1 H), 8.37 (d, $J=9.2$ Hz, 1 H), 8.14-8.09 (m, 2 H), 7.61 (t, $J=7.7$ Hz, 1 H), 7.43-7.37 (m, 4 H), 5.13 (d, $J=3.7$ Hz, 1 H), 4.79 (d, $J=9.2$ Hz, 1 H), 4.48-4.36 (m, 3 H), 4.26-4.23 (m, 1 H), 3.90 (s, 3 H), 3.76-3.72 (m, 2 H), 2.45 (s, 3 H), 2.08-2.03 (m, 1 H), 1.95-1.91 (m, 1 H), 1.04 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.9, 169.3, 165.9, 165.8, 151.4, 147.7, 139.4, 134.7, 132.3, 131.6, 131.1, 129.6, 129.6, 128.7, 128.6, 128.5, 127.4, 68.9, 58.8, 57.5, 56.4, 52.3, 41.6, 37.9, 35.5, 26.5, 15.9; **LCMS** (Method B): $t_{\text{R}} = 1.00$, (m/z) $[\text{M}+\text{H}]^+$ 593 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{31}\text{H}_{37}\text{N}_4\text{O}_6\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 593.2434$, found 593.2438; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3304, 3081, 2953, 2873, 1720, 1626, 1520, 1435; $[\alpha]_{\text{D}}^{20} = -38$ (c 1.00, MeOH).

N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-N3-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)isophthalamide **152b**

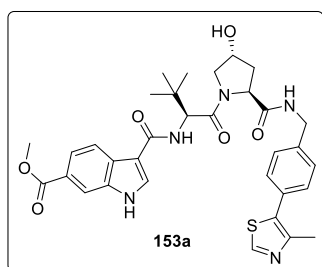


152b was synthesised following the **general procedure D** with the VHL-linker methyl 3-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)benzoate **152a** (180 mg,

0.300 mmol) in DCM. The crude product was purified by MDAP. **152b** (188 mg, 0.205 mmol, 68% yield) was isolated as a white solid.

m.p. 176-183 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.44 (s, 1 H), 8.97 (s, 1 H), 8.57 (t, $J=6.1$ Hz, 1 H), 8.44 (s, 1 H), 8.34 (s, 1 H), 8.24 (d, $J=9.2$ Hz, 1 H), 8.12 (d, $J=7.7$ Hz, 1 H), 8.08-8.04 (m, 2 H), 7.69 (s, 1 H), 7.65-7.57 (m, 3 H), 7.43-7.39 (m, 4 H), 5.17 (d, $J=3.7$ Hz, 1 H), 4.87-4.82 (m, 2 H), 4.50-4.38 (m, 3 H), 4.27-4.23 (m, 1 H), 4.04-3.99 (m, 2 H), 3.81-3.71 (m, 4 H), 3.55 (s, 3 H), 3.19 (s, 6 H), 2.44 (s, 3 H), 2.12 (s, 3 H), 2.09-2.05 (m, 1 H), 1.97-1.93 (m, 1 H), 1.07 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.9, 169.4, 166.2, 165.0, 161.7, 151.9, 151.4, 147.7, 139.7, 139.4, 139.0, 137.3, 135.1, 134.3, 133.8, 131.1, 130.6, 130.4, 129.6, 128.7, 128.4, 127.5, 127.4, 126.9, 118.9, 116.0, 107.9, 104.3, 69.5, 68.9, 58.8, 58.4, 57.5, 56.9, 56.5, 41.7, 37.9, 37.5, 35.6, 26.5, 17.0, 15.9, 13.4; **LCMS** (Method B): $t_{\text{R}} = 0.98$, (m/z) $[\text{M}+\text{H}]^+$ 917 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{49}\text{H}_{57}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 917.4020$, found 917.4009; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3286, 3069, 2952, 2873, 1656, 1609, 1519, 1476; $[\alpha]_{\text{D}}^{20} = -25$ (c 1.00, MeOH).

Methyl 3-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)-1H-indole-6-carboxylate 153a

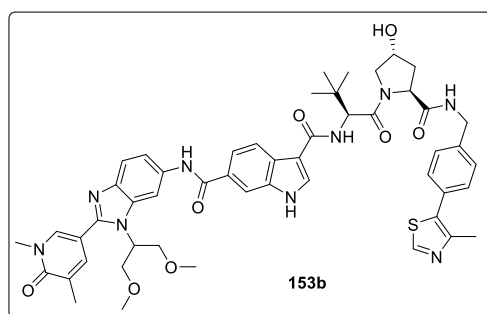


153a was synthesised following the **general procedure C** with 6-(methoxycarbonyl)-1H-indole-3-carboxylic acid (181 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-50% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **153a** (318 mg, 0.478 mmol, 64% yield) was isolated as a white solid.

m.p. 170-179 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 12.01 (s, 1 H), 8.99 (s, 1 H), 8.59-8.54 (m, 2 H), 8.20 (d, $J=8.4$ Hz, 1 H), 8.09 (s, 1 H), 7.75-7.73 (m, 1 H), 7.59 (d, $J=9.2$ Hz, 1 H), 7.44-7.39 (m, 4 H), 5.15 (d, $J=3.7$ Hz, 1 H), 4.87 (d, $J=9.5$ Hz, 1 H), 4.49-4.38 (m, 3 H), 4.28-4.24 (m, 1 H), 3.87 (s, 3 H), 3.80-3.75 (m, 2 H), 2.46 (s, 3 H), 2.08-2.05 (m, 1 H), 1.97-1.92 (m, 1 H), 1.05 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm

171.9, 169.9, 166.9, 163.7, 151.4, 147.7, 139.5, 135.5, 132.0, 131.1, 129.8, 129.6, 128.6, 127.4, 123.0, 121.0, 120.7, 113.7, 110.2, 68.9, 58.7, 56.4, 56.1, 51.8, 41.6, 37.9, 35.5, 26.5, 15.9; **LCMS** (Method B): $t_R = 0.98$, (m/z) $[M+H]^+$ 632 (95% purity); **HRMS** (ESI): calculated for $C_{33}H_{38}N_5O_6S$ (m/z) $[M+H]^+ = 632.2543$, found 632.2541; **IR** (ν_{max}/cm^{-1}) 3431, 3267, 3091, 2955, 2873, 1704, 1622, 1537; $[\alpha]_D^{20} = -30$ (c 1.00, MeOH).

N6-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N3-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-1H-indole-3,6-dicarboxamide 153b



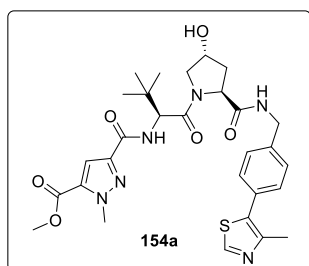
153b was synthesised following the **general procedure D** with the VHL-linker methyl 3-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)-1H-indole-6-carboxylate **153a**

(190 mg, 0.300 mmol) in DCM. The crude product was purified by MDAP. **153b** (19 mg, 0.020 mmol, 7% yield) was isolated as a white solid.

m.p. 182-190 °C; δ_H (400 MHz, d_6 -DMSO) δ ppm 12.12 (s, 1 H), 10.59 (s, 1 H), 9.00 (s, 1 H), 8.69 (s, 1 H), 8.59-8.56 (m, 2 H), 8.84 (s, 1 H), 8.23 (d, $J=8.4$ Hz, 1 H), 8.17 (s, 1 H), 7.98 (d, $J=9.0$ Hz, 1 H), 7.85-7.80 (m, 2 H), 7.74 (s, 1 H), 7.57 (d, $J=9.5$ Hz, 1 H), 7.45-7.40 (m, 4 H), 5.10-5.05 (m, 1 H), 4.88 (d, $J=9.2$ Hz, 1 H), 4.49-4.38 (m, 4 H), 4.28-4.24 (m, 1 H), 4.12-4.09 (m, 2 H), 3.89-3.86 (m, 2 H), 3.79-3.75 (m, 2 H), 3.58 (s, 3 H), 3.23 (s, 6 H), 2.46 (s, 3 H), 2.13 (s, 3 H), 2.09-2.05 (m, 1 H), 1.97-1.92 (m, 1 H), 1.06 (s, 9 H); δ_C (101 MHz, d_6 -DMSO) δ ppm 171.9, 169.9, 166.3, 163.8, 161.7, 151.4, 149.4, 147.6, 141.8, 139.5, 137.1, 135.9, 135.6, 131.5, 131.2, 131.1, 129.6, 128.9, 128.8, 128.6, 128.2, 128.0, 127.4, 119.9, 112.3, 120.4, 119.4, 115.1, 110.1, 105.3, 101.3, 68.9, 68.7, 58.7, 58.6, 58.4, 56.4, 56.1, 41.7, 37.9, 37.9, 35.5, 26.6, 16.9, 15.9; **LCMS** (Method B): $t_R = 0.96$, (m/z) $[M+H]^+$ 956 (100% purity); **HRMS** (ESI): calculated for $C_{51}H_{58}N_9O_8S$ (m/z) $[M+H]^+ = 956.4129$, found 956.4138;

IR ($\nu_{\max}/\text{cm}^{-1}$) 3265, 3058, 2963, 1656, 1608, 1535, 1439, 1371; $[\alpha]_D^{20} = -37$ (*c* 1.00, MeOH).

Methyl 3-(((*S*)-1-(((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)-1-methyl-1*H*-pyrazole-5-carboxylate **154a**

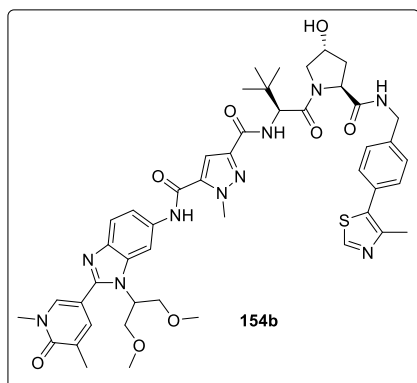


154a was synthesised following the **general procedure C** with 5-(methoxycarbonyl)-1-methyl-1*H*-pyrazole-3-carboxylic acid (152 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-50% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **154a** (318 mg, 0.533 mmol, 71% yield) was isolated as a white

solid.

m.p. 113-117 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 8.98 (s, 1 H), 8.59 (t, $J=5.9$ Hz, 1 H), 7.52 (d, $J=9.5$ Hz, 1 H), 7.43-7.38 (m, 4 H), 7.21 (s, 1 H), 5.14 (d, $J=3.7$ Hz, 1 H), 4.70 (d, $J=9.5$ Hz, 1 H), 4.46 (t, $J=8.3$ Hz, 1 H), 4.42-4.35 (m, 2 H), 4.33-4.25 (m, 1 H), 4.16 (s, 3 H), 3.86 (s, 3 H), 3.71-3.64 (m, 2 H), 2.45 (s, 3 H), 2.09-2.06 (m, 1 H), 1.94-1.89 (m, 1 H), 0.99 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.6, 169.1, 159.4, 159.1, 151.4, 147.7, 144.0, 139.4, 133.9, 131.1, 129.7, 128.7, 127.4, 110.7, 68.8, 58.7, 56.6, 56.1, 52.3, 41.7, 39.9, 37.9, 36.0, 26.2, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.99$, (m/z) $[\text{M}+\text{H}]^+$ 597 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{29}\text{H}_{37}\text{N}_6\text{O}_6\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 597.2495$, found 597.2493; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3415, 3299, 3065, 2958, 2873, 1728, 1628, 1538; $[\alpha]_D^{20} = -37$ (*c* 1.00, MeOH).

N*5-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*3-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-1-methyl-1*H*-pyrazole-3,5-dicarboxamide **154b*

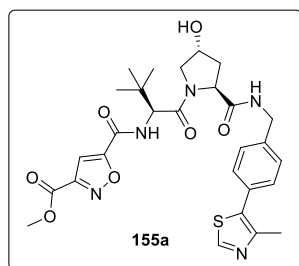


154b was synthesised following the **general procedure D** with the VHL-linker methyl 3-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)-1-methyl-1*H*-pyrazole-5-carboxylate **154a** (179 mg, 0.300 mmol) in DCM. The crude product was purified by MDAP. **154b** (127 mg, 0.138 mmol,

46% yield) was isolated as a white solid.

m.p. 206-213 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.75 (s, 1 H), 9.02 (s, 1 H), 8.67-8.61 (m, 2 H), 8.36 (s, 1 H), 7.96 (d, $J=8.8$ Hz, 1 H), 7.83 (d, $J=8.8$ Hz, 1 H), 7.73 (s, 1 H), 7.68 (s, 1 H), 7.50 (d, $J=9.5$ Hz, 1 H), 7.44-7.41 (m, 4 H), 5.11-5.06 (m, 1 H), 4.74 (d, $J=9.5$ Hz, 1 H), 4.48 (t, $J=8.1$ Hz, 1 H), 4.43-4.37 (m, 3 H), 4.31-4.26 (m, 1 H), 4.21 (s, 3 H), 4.12-4.07 (m, 2 H), 3.89-3.85 (m, 2 H), 3.73-3.65 (m, 2 H), 3.57 (s, 3 H), 3.23 (s, 6 H), 2.46 (s, 3 H), 2.12 (s, 3 H), 2.11-2.07 (m, 1 H), 1.95-1.90 (m, 1 H), 1.01 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.6, 169.1, 161.7, 159.7, 157.6, 151.5, 149.6, 147.5, 143.7, 142.0, 139.6, 137.0, 135.9, 135.9, 131.2, 131.0, 129.6, 129.0, 128.7, 128.2, 127.5, 119.5, 115.2, 108.8, 105.8, 101.1, 68.8, 68.8, 58.7, 58.6, 58.5, 56.7, 56.0, 41.7, 39.8, 37.9, 37.9, 36.2, 26.2, 16.9, 15.8; **LCMS** (Method B): $t_{\text{R}} = 1.01$, (m/z) $[\text{M}+\text{H}]^+$ 921 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{47}\text{H}_{57}\text{N}_{10}\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 921.4082$, found 921.4081; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3254, 3079, 2942, 2825, 1656, 1625, 1545, 1429; $[\alpha]_{\text{D}}^{20} = -47$ (c 1.00, MeOH).

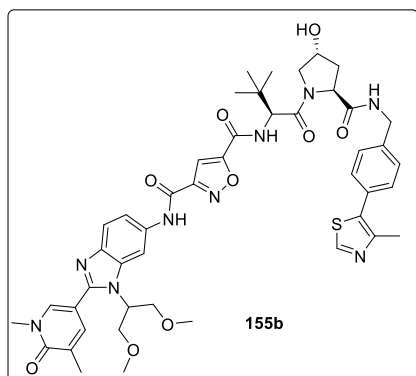
Methyl 5-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)isoxazole-3-carboxylate 155a



155a was synthesised following the **general procedure C** with 3-(methoxycarbonyl)isoxazole-5-carboxylic acid (96.0 mg, 0.561 mmol). The crude product was purified by flash column chromatography eluting with 0-60% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **155a** (241 mg, 0.401 mmol, 79% yield) was isolated as a white solid.

m.p. 127-131 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 9.00 (s, 1 H), 8.76 (d, $J=9.2$ Hz, 1 H), 8.59 (t, $J=6.1$ Hz, 1 H), 7.86 (s, 1 H), 7.74-7.78 (m, 4 H), 4.75 (d, $J=9.2$ Hz, 1 H), 4.48-4.42 (m, 2 H), 4.39-4.36 (m, 1 H), 4.26-4.22 (m, 1 H), 3.92 (s, 3 H), 3.74-3.67 (m, 3 H), 2.45 (s, 3 H), 2.08-2.05 (m, 1 H), 1.94-1.90 (m, 1 H), 1.02 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.7, 168.4, 164.2, 159.2, 156.3, 155.0, 151.5, 151.2, 147.6, 139.5, 129.6, 128.6, 127.4, 106.7, 68.8, 58.8, 57.2, 56.5, 53.0, 41.6, 37.9, 35.6, 26.3, 15.8; **LCMS** (Method B): $t_{\text{R}} = 0.94$, (m/z) $[\text{M}+\text{H}]^+$ 584 (95% purity); **HRMS** (ESI): calculated for $\text{C}_{28}\text{H}_{34}\text{N}_5\text{O}_7\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 584.2179$, found 584.2178; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3302, 3079, 2947, 2873, 1736, 1625, 1524, 1429; $[\alpha]_{\text{D}}^{20} = -43$ (c 1.00, MeOH).

N3-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N5-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)isoxazole-3,5-dicarboxamide **155b**

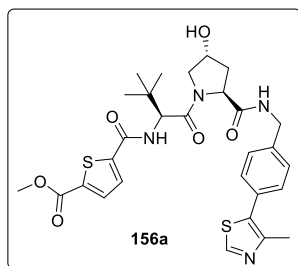


155b was synthesised following the **general procedure D** with the VHL-linker methyl 5-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)isoxazole-3-carboxylate **155a** (175 mg, 0.300 mmol) in DMF. The crude product was purified by MDAP. **155b** (28 mg, 0.031 mmol, 10% yield) was isolated as a

yellow solid.

m.p. 170-175 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.86 (s, 1 H), 8.99 (s, 1 H), 8.75 (d, $J=9.2$ Hz, 1 H), 8.63-8.58 (m, 1 H), 8.30 (s, 1 H), 8.08 (s, 1 H), 7.87 (s, 1 H), 7.71-7.62 (m, 3 H), 7.46-7.40 (m, 4 H), 5.20-5.12 (m, 1 H), 4.89-4.85 (m, 1 H), 4.78 (d, $J=9.2$ Hz, 1 H), 4.51-4.40 (m, 2 H), 4.39-4.37 (m, 1 H), 4.28-4.23 (m, 1 H), 4.03-4.00 (m, 2 H), 3.81-3.78 (m, 2 H), 3.75-3.69 (m, 2 H), 3.55 (s, 3 H), 3.19 (s, 6 H), 2.45 (s, 3 H), 2.10 (s, 3 H), 2.09-2.06 (m, 1 H), 1.96-1.91 (m, 1 H), 1.06-1.01 (m, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.7, 168.5, 161.7, 159.6, 158.9, 156.4, 153.4, 152.0, 151.4, 147.7, 139.4, 139.4, 139.3, 137.1, 133.1, 132.8, 131.1, 129.6, 128.6, 127.5, 127.4, 118.6, 116.3, 107.2, 106.1, 104.9, 69.4, 68.9, 58.8, 58.4, 57.2, 57.0, 56.5, 41.6, 37.9, 37.5, 35.6, 26.4, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.95$, (m/z) $[\text{M}+\text{H}]^+$ 908 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{46}\text{H}_{54}\text{N}_9\text{O}_9\text{S}$ (m/z) $[\text{M}+2\text{H}]^{2+} = 454.6922$, found 454.6926; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3275, 3069, 2947, 2867, 1656, 1603, 1524, 1429; $[\alpha]_{\text{D}}^{20} = -42$ (c 1.00, MeOH).

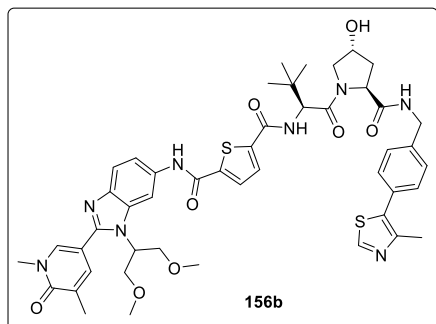
Methyl 5-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)thiophene-2-carboxylate 156a



156a was synthesised following the **general procedure C** with 5-(methoxycarbonyl)thiophene-2-carboxylic acid (0.154 g, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-60% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **156a** (372 mg, 0.603 mmol, 80% yield) was isolated as a white solid.

m.p. 120-128 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 8.98 (s, 1 H), 8.57 (t, $J=6.1$ Hz, 1 H), 8.48 (d, $J=9.2$ Hz, 1 H), 8.16 (d, $J=4.4$ Hz, 1 H), 7.79 (d, $J=4.0$ Hz, 1 H), 7.43-7.39 (m, 4 H), 4.75 (d, $J=9.2$ Hz, 1 H), 4.48-4.42 (m, 2 H), 4.38-4.36 (m, 1 H), 4.26-4.22 (m, 1 H), 4.02-3.96 (m, 1 H), 3.84 (s, 3 H), 3.74-3.69 (m, 2 H), 2.45 (s, 3 H), 2.08-2.04 (m, 1 H), 1.95-1.90 (m, 1 H), 1.04 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.8, 168.9, 161.6, 160.4, 151.1, 147.6, 145.2, 139.5, 135.9, 133.8, 131.2, 129.6, 129.4, 128.6, 127.4, 68.9, 58.8, 57.5, 56.4, 52.5, 41.6, 37.9, 35.5, 26.5, 15.8; **LCMS** (Method B): $t_{\text{R}} = 1.00$, (m/z) $[\text{M}+\text{H}]^+$ 599 (97% purity); **HRMS** (ESI): calculated for $\text{C}_{29}\text{H}_{35}\text{N}_4\text{O}_6\text{S}_2$ (m/z) $[\text{M}+\text{H}]^+ = 599.1998$, found 599.1993; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3310, 3079, 2958, 2873, 1709, 1619, 1535, 1439; $[\alpha]_{\text{D}}^{20} = -75$ (c 1.00, MeOH).

N*2-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*5-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)thiophene-2,5-dicarboxamide **156b*

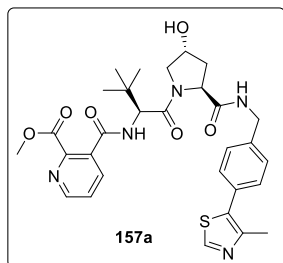


156b was synthesised following the **general procedure D** with the VHL-linker methyl 5-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)thiophene-2-carboxylate **156a** (180 mg, 0.300 mmol) in DMF. The crude product was purified by MDAP.

156b (145 mg, 0.157 mmol, 52% yield) was isolated as a yellow solid.

m.p. 183-189 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.44 (s, 1 H), 8.99 (s, 1 H), 8.58 (t, $J=5.9$ Hz, 1 H), 8.35-8.30 (m, 2 H), 8.17 (d, $J=4.0$ Hz, 1 H), 8.08 (s, 1 H), 8.02 (d, $J=4.0$ Hz, 1 H), 7.69 (s, 1 H), 7.64 (d, $J=8.8$ Hz, 1 H), 7.57 (d, $J=8.4$ Hz, 1 H), 7.45-7.40 (m, 4 H), 5.20-12 (m, 1 H), 4.90-4.86 (m, 1 H), 4.77 (d, $J=9.2$ Hz, 1 H), 4.49-4.43 (m, 2 H), 4.40-4.37 (m, 1 H), 4.28-4.22 (m, 1 H), 4.05-4.00 (m, 2 H), 3.81-3.78 (m, 2 H), 3.75-3.71 (m, 2 H), 3.55 (s, 3 H), 3.19 (s, 6 H), 2.46 (s, 3 H), 2.12 (s, 3 H), 2.08-2.05 (m, 1 H), 1.96-1.91 (m, 1 H), 1.05 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.0, 161.7, 160.7, 159.3, 151.7, 151.4, 147.7, 143.6, 143.1, 139.4, 139.3, 138.5, 137.1, 133.5, 133.0, 131.1, 129.6, 129.3, 129.1, 128.6, 127.6, 127.4, 118.5, 116.4, 107.0, 104.5, 69.4, 68.9, 58.8, 58.4, 57.4, 57.1, 56.4, 41.7, 37.9, 37.5, 35.5, 26.5, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.96$, (m/z) $[\text{M}+\text{H}]^+$ 923 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{47}\text{H}_{55}\text{N}_8\text{O}_8\text{S}_2$ (m/z) $[\text{M}+2\text{H}]^{2+} = 462.1831$, found 462.1837; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3284, 3074, 2931, 2857, 1652, 1604, 1537, 1432; $[\alpha]_{\text{D}}^{20} = -75$ (c 1.00, MeOH).

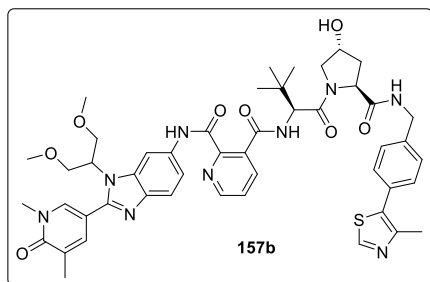
Methyl 3-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)picolinate
157a



157a was synthesised following the **general procedure C** with 2-(methoxycarbonyl)nicotinic acid (150 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-80% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **157a** (350 mg, 0.584 mmol, 78% yield) was isolated as a white solid.

m.p. 116-120 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 8.97 (s, 1 H), 8.69-8.67 (m, 2 H), 8.57 (t, $J=6.1$ Hz, 1 H), 7.93 (d, $J=7.8$ Hz, 1 H), 7.61-7.59 (m, 1 H), 7.44-7.36 (m, 4 H), 5.14 (d, $J=3.5$ Hz, 1 H), 4.71 (d, $J=9.4$ Hz, 1 H), 4.50-4.45 (m, 2 H), 4.41-4.38 (m, 1 H), 4.25-4.22 (m, 1 H), 3.76-3.71 (m, 5 H), 2.44 (s, 3 H), 2.09-2.05 (m, 1 H), 1.96-1.93 (m, 1 H), 1.03 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.9, 169.1, 166.3, 165.9, 151.4, 149.8, 148.1, 147.7, 139.4, 136.5, 132.4, 131.1, 129.6, 128.6, 127.4, 125.1, 68.8, 58.8, 57.2, 56.4, 52.2, 41.6, 37.9, 35.3, 26.4, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.85$, (m/z) $[\text{M}+\text{H}]^+$ 594 (99% purity); **HRMS** (ESI): calculated for $\text{C}_{30}\text{H}_{36}\text{N}_5\text{O}_6\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 594.2386$, found 594.2382; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3280, 3064, 2963, 2873, 1736, 1630, 1535, 1423; $[\alpha]_{\text{D}}^{20} = -44$ (c 1.00, MeOH).

N2-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-N3-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)pyridine-2,3-dicarboxamide **157b**

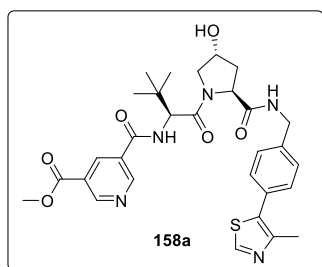


157b was synthesised following the **general procedure D** with the VHL-linker methyl 3-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)picolinate **157a** (178 mg, 0.300 mmol) in DMF. The crude

product was purified by MDAP. **157b** (64 mg, 0.070 mmol, 23% yield) was isolated as a beige solid.

m.p. 187-195 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.96 (s, 1 H), 9.00 (s, 1 H), 8.75 (d, $J=4.8$ Hz, 1 H), 8.70 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 8.49 (d, $J=9.2$ Hz, 1 H), 8.37 (s, 1 H), 8.07 (d, $J=8.8$ Hz, 1 H), 7.91 (d, $J=7.7$ Hz, 1 H), 7.83 (d, $J=8.8$ Hz, 1 H), 7.774 (s, 1 H), 7.70-7.67 (m, 1 H), 7.43-7.35 (m, 4 H), 5.10-5.06 (m, 1 H), 4.72 (d, $J=9.2$ Hz, 1 H), 4.49 (d, $J=7.9$ Hz, 1 H), 4.43-4.38 (m, 3 H), 4.28-4.22 (m, 1 H), 4.11-4.06 (m, 2 H), 3.88-3.84 (m, 2 H), 3.79-3.71 (m, 2 H), 3.58 (s, 3 H), 3.21 (s, 6 H), 2.43 (s, 3 H), 2.12 (s, 3 H), 2.10-2.06 (m, 1 H), 1.98-1.93 (m, 1 H), 1.04 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.3, 166.8, 163.7, 161.7, 151.5, 149.3, 148.8, 148.6, 142.0, 147.5, 139.5, 137.5, 136.6, 135.8, 132.9, 131.9, 130.9, 129.6, 128.7, 128.6, 128.2, 127.4, 125.6, 119.1, 115.1, 104.8, 100.8, 68.8, 68.6, 58.8, 58.6, 58.6, 57.3, 56.1, 41.6, 37.9, 37.9, 35.6, 26.3, 16.9, 15.8; **LCMS** (Method B): $t_{\text{R}} = 0.92$, (m/z) $[\text{M}+\text{H}]^+$ 918 (95% purity); **HRMS** (ESI): calculated for $\text{C}_{48}\text{H}_{56}\text{N}_9\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 918.3973$, found 918.3978; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3261, 3064, 2942, 2884, 1656, 1610, 1529, 1418; $[\alpha]_{\text{D}}^{20} = -22$ (c 1.00, MeOH).

Methyl 5-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)nicotinate
158a

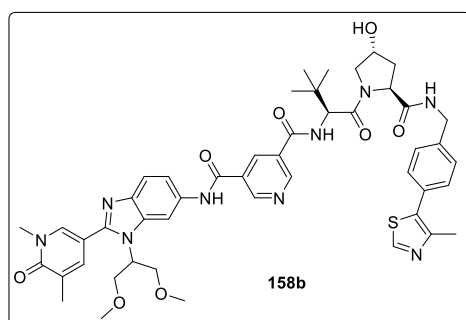


158a was synthesised following the **general procedure C** with 5-(methoxycarbonyl)nicotinic acid (150 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-70% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **158a** (413 mg, 0.696 mmol, 93% yield) was isolated as a white solid.

m.p. 88-92 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.20-9.18 (m, 2 H), 8.98 (s, 1 H), 8.80 (d, $J=8.8$ Hz, 1 H), 8.67 (s, 1 H), 8.57 (t, $J=6.1$ Hz, 1 H), 7.43-7.37 (m, 4 H), 5.14 (d, $J=3.7$ Hz, 1 H), 4.18 (d, $J=9.2$ Hz, 1 H), 4.49-4.42 (m, 2 H), 4.40-4.37 (m, 1 H), 4.27-4.23 (m, 1 H), 3.93 (s, 3 H), 3.77-3.72 (m, 2 H), 2.45 (s, 3 H), 2.09-2.04 (m, 1 H), 1.96-

1.92 (m, 1 H), 1.05 (s, 9 H); δ_C (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.0, 164.8, 164.5, 152.7, 151.8, 151.4, 147.7, 139.4, 136.1, 131.1, 129.8, 129.6, 128.6, 127.4, 125.0, 68.8, 58.8, 57.6, 56.4, 52.6, 41.7, 37.9, 35.5, 26.5, 15.9; **LCMS** (Method B): t_R = 0.91, (m/z) $[M+H]^+$ 594 (100% purity); **HRMS** (ESI): calculated for $C_{30}H_{36}N_5O_6S$ (m/z) $[M+H]^+$ = 594.2386, found 594.2385; **IR** (ν_{max}/cm^{-1}) 3298, 3075, 2958, 2879, 1715, 1662, 1622, 1508; $[\alpha]_D^{20} = -33$ (c 1.00, MeOH).

N3-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N5-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)pyridine-3,5-dicarboxamide 158b



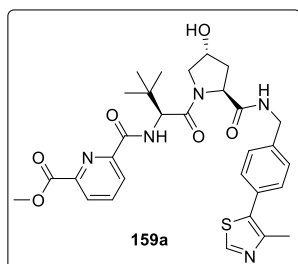
158b was synthesised following the **general procedure D** with the VHL-linker methyl 5-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)nicotinate **158a** (178 mg, 0.300 mmol) in DMF. The crude product was

purified by MDAP. **158b** (87 mg, 0.095 mmol, 32% yield) was isolated as a white solid.

m.p. 207-215 °C; δ_H (400 MHz, d_6 -DMSO) δ ppm 11.15 (s, 1 H), 9.32 (s, 1 H), 9.18 (s, 1 H), 9.02 (s, 1 H), 8.87 (s, 1 H), 8.77 (s, 1 H), 8.70 (d, $J=9.2$ Hz, 1 H), 8.59 (t, $J=6.1$ Hz, 1 H), 8.37 (s, 1 H), 7.99 (d, $J=9.0$ Hz, 1 H), 7.86 (d, $J=8.8$ Hz, 1 H), 7.74 (s, 1 H), 7.44-7.39 (m, 4 H), 5.12-5.08 (m, 1 H), 4.83 (d, $J=8.8$ Hz, 1 H), 4.50-4.39 (m, 4 H), 4.27-4.23 (m, 1 H), 4.12-4.09 (m, 2 H), 3.90-3.86 (m, 2 H), 3.78-3.75 (m, 2 H), 3.58 (s, 3 H), 3.25 (s, 6 H), 2.45 (s, 3 H), 2.12 (s, 3 H), 2.10-2.05 (m, 1 H), 1.97-1.91 (m, 1 H), 1.09 (s, 9 H); δ_C (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.1, 164.7, 163.7, 161.7, 151.5, 151.0, 150.6, 149.5, 147.5, 142.1, 139.5, 136.5, 135.8, 135.2, 131.2, 131.0, 129.7, 129.5, 129.5, 128.6, 128.6, 128.2, 127.4, 119.6, 115.0, 105.8, 100.8, 68.9, 68.8, 58.8, 58.6, 58.6, 57.7, 56.4, 41.7, 37.9, 37.9, 35.5, 26.5, 16.9, 15.8; **LCMS** (Method B): t_R = 0.92, (m/z) $[M+H]^+$ 918 (100% purity); **HRMS** (ESI):

calculated for $C_{48}H_{56}N_9O_8S$ (m/z) $[M+H]^+ = 918.3973$, found 918.3980; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3259, 3069, 2952, 2825, 1646, 1619, 1535, 1434; $[\alpha]_D^{20} = -28$ (c 1.00, MeOH).

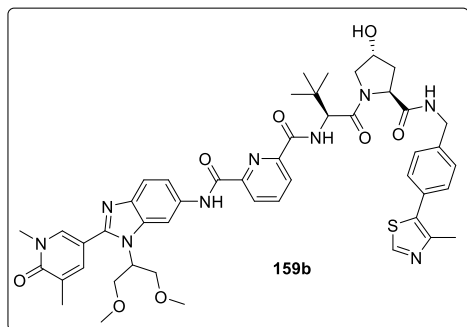
Methyl 6-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)picolinate
159a



159a was synthesised following the **general procedure C** with 6-(methoxycarbonyl)picolinic acid (150 mg, 0.827 mmol). The crude product was purified by flash column chromatography eluting with 0-70% 3:1 EtOAc:EtOH in cyclohexane over 30 CV. **159a** (355 mg, 0.586 mmol, 78% yield) was isolated as a white solid.

m.p. 122-126 °C; **δ_H** (400 MHz, d_6 -DMSO) δ ppm 8.99 (s, 1 H), 8.61 (t, $J=6.1$ Hz, 1 H), 8.52 (d, $J=9.5$ Hz, 1 H), 8.30-8.20 (m, 3 H), 7.44-7.40 (m, 4 H), 4.75 (d, $J=9.5$ Hz, 1 H), 4.47 (t, $J=8.3$ Hz, 1 H), 4.40-4.36 (m, 2 H), 4.32-4.28 (m, 1 H), 3.93 (s, 3 H), 3.75-3.68 (m, 3 H), 2.46 (s, 3 H), 2.10-2.06 (m, 1 H), 1.95-1.91 (m, 1 H), 1.03 (s, 9 H); **δ_C** (101 MHz, d_6 -DMSO) δ ppm 171.6, 168.9, 164.3, 162.1, 151.4, 149.2, 147.6, 146.7, 139.7, 139.4, 131.1, 129.7, 128.7, 127.5, 127.5, 125.2, 68.9, 58.7, 56.7, 56.5, 52.8, 41.7, 37.9, 36.1, 26.1, 15.8; **LCMS** (Method B): $t_R = 0.97$, (m/z) $[M+H]^+ = 594$ (98% purity); **HRMS** (ESI): calculated for $C_{30}H_{36}N_5O_6S$ (m/z) $[M+H]^+ = 594.2386$, found 594.2386; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3290, 3085, 2958, 2873, 1730, 1630, 1524, 1434; $[\alpha]_D^{20} = -19$ (c 1.00, MeOH).

N*2-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*6-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)pyridine-2,6-dicarboxamide **159b*

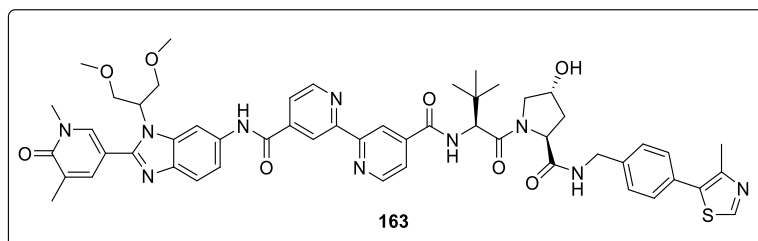


159b was synthesised following the **general procedure D** with the VHL-linker methyl 6-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)picolinate **159a** (178 mg, 0.300 mmol) in DMF. The crude product was

purified by MDAP. **159b** (100 mg, 0.109 mmol, 36% yield) was isolated as a beige solid.

m.p. 162-169 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 11.04 (s, 1 H), 9.10 (d, $J=9.5$ Hz, 1 H), 8.95 (s, 1 H), 8.56 (t, $J=6.1$ Hz, 1 H), 8.37 (s, 1 H), 8.35-8.32 (m, 1 H), 8.28-8.25 (m, 2 H), 8.06 (s, 1 H), 7.71-7.65 (m, 2 H), 7.62-7.59 (m, 1 H), 7.43-7.38 (m, 4 H), 5.17-5.15 (m, 1 H), 4.88-4.82 (m, 2 H), 4.50-4.47 (m, 1 H), 4.43-4.40 (m, 2 H), 4.32-4.25 (m, 1 H), 4.06-4.00 (m, 2 H), 3.82-3.76 (m, 4 H), 3.55 (s, 3 H), 3.18-3.16 (m, 6 H), 2.42 (s, 3 H), 2.10 (s, 3 H), 2.09-2.05 (m, 1 H), 1.96-1.93 (m, 1 H), 1.13 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.1, 163.4, 161.9, 161.7, 152.0, 151.4, 149.5, 148.8, 147.7, 139.7, 139.7, 139.4, 139.1, 137.2, 133.5, 133.1, 131.1, 129.6, 128.7, 127.5, 127.4, 125.1, 125.0, 119.0, 115.8, 107.7, 104.2, 69.4, 68.9, 58.9, 58.4, 56.9, 57.5, 56.4, 41.7, 37.9, 37.5, 35.6, 26.5, 17.0, 15.8; **LCMS** (Method B): $t_{\text{R}} = 0.98$, (m/z) $[\text{M}+\text{H}]^+$ 918 (99% purity); **HRMS** (ESI): calculated for $\text{C}_{48}\text{H}_{56}\text{N}_9\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 918.3973$, found 918.3961; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3291, 3070, 2931, 2873, 1651, 1609, 1524, 1439; $[\alpha]_{\text{D}}^{20} = +28$ (c 1.00, MeOH).

N*4-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*4'-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-[2,2'-bipyridine]-4,4'-dicarboxamide **163*

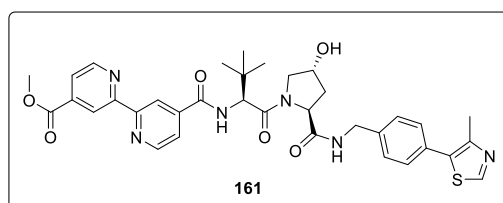


4'-(Methoxycarbonyl)-[2,2'-bipyridine]-4-carboxylic acid **160** (25.8 mg, 0.100 mmol), HATU (41.8 mg, 0.110 mmol), DIPEA (0.070 mL, 0.40 mmol) and DCM (0.25 mL) were added to a 4 mL vial and sonicated for 30 s. (2*S*,4*R*)-1-((*S*)-2-amino-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide·HCl **30** (46.7 mg, 0.100 mmol) was added and the reaction mixture stirred at ambient temperature for 2 h. The crude reaction mixture was blown down under a stream of N₂ and placed in a vacuum oven at 40 °C for 18 h. The crude product was dissolved in MeOH (0.2 mL) and added to a stirred solution of LiOH (9.6 mg, 0.40 mmol) in water (0.2 mL). The reaction mixture was heated at 60 °C for 18 h. The crude reaction mixture was blown down under a stream of N₂ and placed in a vacuum oven at 40 °C for 18 h. HATU (41.8 mg, 0.110 mmol), DIPEA (0.050 mL, 0.30 mmol) and DMF (0.25 mL) were added to the vial and the reaction mixture was sonicated for 30 s. 5-(6-amino-1-(1,3-dimethoxypropan-2-yl)-1*H*-benzo[*d*]imidazol-2-yl)-1,3-dimethylpyridin-2(1*H*)-one **128** (35.6 mg, 0.100 mmol) was added and the reaction mixture stirred at ambient temperature for 2 h. HATU (20.9 mg, 0.0550 mmol) and DIPEA (0.030 mL, 0.15 mmol) were added and the reaction mixture stirred at ambient temperature for a further 2 h. The reaction mixture was purified directly by MDAP to afford the product **163** (36 mg, 0.036 mmol, 36% yield) as a beige solid.

m.p. 180-190 °C; δ_{H} (400 MHz, *d*₆-DMSO) δ ppm 10.74 (s, 1 H), 8.98 (s, 1 H), 8.95 (d, *J*=5.0 Hz, 1 H), 8.93 (s, 1 H), 8.88 (d, *J*=5.0 Hz, 1 H), 8.77 (s, 1 H), 8.70 (d, *J*=9.2 Hz, 1 H), 8.58 (t, *J*=6.1 Hz, 1 H), 8.33 (s, 1 H), 8.06 (s, 1 H), 8.01 (d, *J*=5.0 Hz,

1 H), 7.88 (d, $J=5.0$ Hz, 1 H), 7.69 (s, 1 H), 7.64-7.61 (m, 2 H), 7.44-7.40 (m, 4 H), 5.15 (d, $J=3.5$ Hz, 1 H), 4.89-4.83 (m, 2 H), 4.50 (t, $J=8.1$ Hz, 1 H), 4.46-4.38 (m, 2 H), 4.28-4.24 (m, 1 H), 4.04 (t, $J=9.6$ Hz, 2 H), 3.82-3.74 (m, 4 H), 3.55 (s, 3 H), 3.20 (s, 6 H), 2.46 (s, 3 H), 2.11 (s, 3 H), 2.09-2.06 (m, 1 H), 1.97-1.93 (m, 1 H), 1.07 (s, 9 H); δ_C (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.0, 165.5, 163.7, 161.7, 155.5, 155.1, 152.1, 151.4, 150.1, 149.8, 147.7, 143.5, 142.9, 140.0, 139.4, 139.0, 137.2, 133.4, 133.3, 131.1, 129.6, 128.6, 127.5, 127.4, 122.4, 122.2, 118.9, 118.9, 118.9, 118.4, 116.1, 107.8, 104.7, 69.5, 68.9, 58.8, 58.4, 57.6, 56.9, 56.5, 41.7, 37.9, 37.5, 35.6, 26.5, 16.9, 15.9; **LCMS** (Method B): $t_R = 0.99$, (m/z) $[M+H]^+$ 995 (100% purity); **HRMS** (ESI): calculated for $C_{53}H_{59}N_{10}O_8S$ (m/z) $[M+H]^+ = 995.4238$, found 995.4235; **IR** (ν_{max}/cm^{-1}) 3296, 3080, 2960, 2877, 1622, 1538, 1437, 1218; $[\alpha]_D^{20} = -18$ (c 0.29, MeOH).

Methyl 4'-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)-[2,2'-bipyridine]-4-carboxylate **161**



4'-(Methoxycarbonyl)-[2,2'-bipyridine]-4-carboxylic acid **160** (25.8 mg, 0.100 mmol), HATU (41.8 mg, 0.110 mmol), DIPEA (0.070 mL, 0.40 mmol) and either DMF or

DCM were added to a 4 mL vial and sonicated for 30 s. (2S,4R)-1-((S)-2-amino-3,3-dimethylbutanoyl)-4-hydroxy-N-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide·HCl **30** (46.7 mg, 0.100 mmol) was added and the reaction mixture stirred at ambient temperature for 2 h.

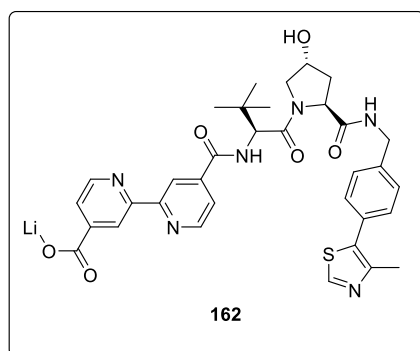
A selection of solvents and concentrations were screened using the described procedure (*vide supra*) and the conditions visualised in **Table 38**. The reaction mixtures were analysed by LCMS.

Table 38: Conditions screened for the synthesis of ester **161** from the amide coupling of VHL E3 ligase binder **30** and 2,2'-bipyridyl linker **160**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Solvent	Solvent / mL	Concentration / M	161 / %	30 / %
1	DMF	1.5	0.067	86	5
2	DMF	1.0	0.10	74	6
3	DMF	0.50	0.20	81	6
4	DCM	1.5	0.067	80	5
5	DCM	1.0	0.10	79	5
6	DCM	0.50	0.20	83	5
7	DCM	0.25	0.40	79	5

The crude reaction mixture was blown down under a stream of N₂ and placed in a vacuum oven at 40 °C for 18 h to furnish **161**. The crude product **161** was used without further purification.

Lithium 4'-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)-[2,2'-bipyridine]-4-carboxylate **162**



Methyl 4'-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)-[2,2'-bipyridine]-4-carboxylate **161** (67 mg, 0.10 mmol) synthesised *via* the described procedure (*vide supra*) was dissolved in either MeOH or THF and added to a stirred solution of

LiOH in water (to create a 1:1 solvent mixture) and stirred at ambient temperature for 18 h.

A selection of solvents, concentrations and temperatures were screened using the described procedure (*vide supra*) and the conditions visualised in **Table 39**. The reaction mixtures were analysed by LCMS.

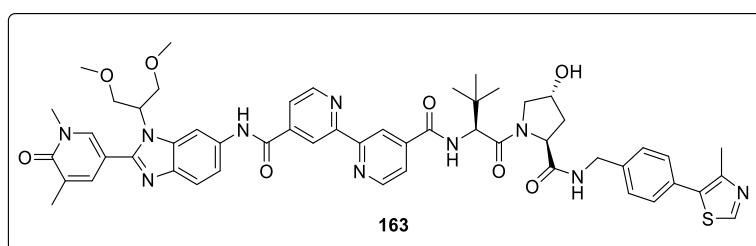
Table 39: Reaction conditions screened for the hydrolysis of ester **161** to lithium carboxylate salt **162**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	LiOH / mg	LiOH / mmol	Cosolvent	Cosolvent / mL	Temperature / °C	162 / %	161 / %
1	5.0	0.20	MeOH	0.4	RT	22	73
2*	5.0	0.20	MeOH	0.4	RT	10	79
3	5.0	0.20	THF	0.4	RT	7	81
4	5.0	0.20	MeOH	0.4	60	84	11
5*	5.0	0.20	MeOH	0.4	60	28	62
6	5.0	0.20	THF	0.4	60	8	89
7	5.0	0.20	MeOH	0.4	60	84	11
8	7.5	0.30	MeOH	0.4	60	84	11
9	10	0.40	MeOH	0.4	60	83	10
10	5.0	0.20	MeOH	0.2	60	77	18
11	7.5	0.30	MeOH	0.2	60	83	12

* = Crude reaction mixture performed in DMF. All others performed in DCM.

The crude reaction mixture was blown down under a stream of N₂ and placed in a vacuum oven at 40 °C for 18 h to furnish **162**. The crude product **162** was used without further purification.

N4-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N4'-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-[2,2'-bipyridine]-4,4'-dicarboxamide 163



HATU, DIPEA (0.05 mL, 0.30 mmol) and DMF (0.25 mL) were added to a vial containing lithium 4'-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)-[2,2'-bipyridine]-4-carboxylate **162** (66.9 mg, 0.100 mmol) synthesised *via* the described procedure (*vide supra*) and sonicated for 30 s. 5-(6-amino-1-(1,3-dimethoxypropan-2-

yl)-1*H*-benzo[*d*]imidazol-2-yl)-1,3-dimethylpyridin-2(1*H*)-one **128** (35.6 mg, 0.100 mmol) was added and the reaction mixture stirred at ambient temperature for 2 h.

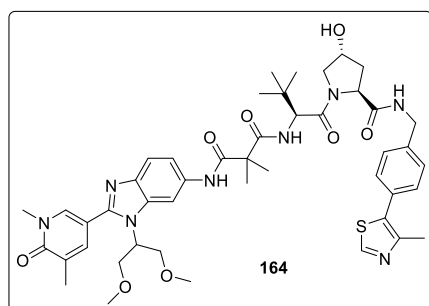
Lithium 4'-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamoyl)-[2,2'-bipyridine]-4-carboxylate **162** synthesised using varying amounts of LiOH was screened with different amounts of HATU using the described procedure (*vide supra*) and the conditions visualised in **Table 40**. The reaction mixtures were analysed by LCMS.

Table 40: Reaction conditions screened for the synthesis of **163** from the amide coupling of lithium carboxylate salt **162** and amine **128**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	LiOH / equiv	HATU / mg	HATU / mmol	163 / %	162 / %	161 / %
1	3	42	0.11	22	5	18
2	3	76	0.20	24	5	11
3	4	42	0.11	41	8	7
4	10	42	0.11	30	10	7
5*	4	42	0.11	56	0	7

* = Second addition of HATU (0.055 mmol) and DIPEA (0.15 mmol) after 2 h, with stirring for a further 2 h.

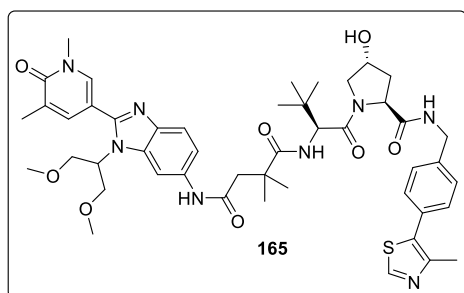
N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-N3-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-2,2-dimethylmalonamide **164**



164 was synthesised following the **general procedure E** with the linker 3-methoxy-2,2-dimethyl-3-oxopropanoic acid (44.0 mg, 0.300 mmol). **164** (74 mg, 0.084 mmol, 28% yield) was isolated as a white solid.

m.p. 136-144 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.47 (s, 1 H), 8.97 (s, 1 H), 8.57-8.53 (m, 1 H), 8.08 (s, 1 H), 8.02 (s, 1 H), 7.66 (s, 1 H), 7.52 (d, $J=8.8$ Hz, 1 H), 7.45 (d, $J=8.6$ Hz, 1 H), 7.40-7.37 (m, 4 H), 7.27 (d, $J=8.8$ Hz, 1 H), 5.14 (s (br), 1 H), 4.84-4.80 (m, 1 H), 4.53 (d, $J=8.8$ Hz, 1 H), 4.48-4.45 (m, 1 H), 4.38-4.34 (m, 2 H), 4.29-4.26 (m, 1 H), 3.99-3.95 (m, 2 H), 3.76-3.73 (m, 2 H), 3.69-3.66 (m, 1 H), 3.63-3.61 (m, 1 H), 3.53 (s, 3 H), 3.16 (s, 6 H), 2.42 (s, 3 H), 2.09 (s, 3 H), 2.07-2.03 (m, 1 H), 1.94-1.88 (m, 1 H), 1.53 (s, 3 H), 1.46 (s, 3 H), 0.93 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 172.2, 171.7, 171.5, 169.3, 161.7, 151.9, 151.4, 147.7, 139.8, 139.3, 139.0, 137.2, 133.3, 133.2, 131.1, 129.7, 128.7, 127.4, 127.4, 118.6, 116.6, 107.9, 104.9, 69.4, 68.8, 58.8, 58.4, 56.8, 56.8, 56.5, 56.4, 41.7, 37.9, 37.5, 36.0, 26.1, 23.8, 23.2, 16.9, 15.8; **LCMS** (Method B): $t_{\text{R}} = 0.94$, (m/z) $[\text{M}+\text{H}]^+$ 883 (98% purity); **HRMS** (ESI): calculated for $\text{C}_{46}\text{H}_{59}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 883.4177$, found 883.4176; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3307, 3079, 2937, 2878, 1651, 1619, 1519, 1482; $[\alpha]_{\text{D}}^{20} = -30$ (c 0.16, MeOH).

N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N4-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-2,2-dimethylsuccinamide 165

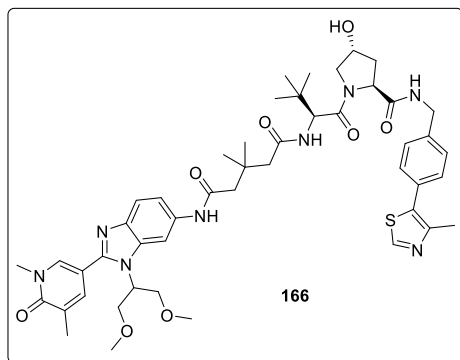


165 was synthesised following the **general procedure E** with the linker 4-methoxy-2,2-dimethyl-4-oxobutanoic acid (48.0 mg, 0.300 mmol). **165** (55 mg, 0.061 mmol, 20% yield) was isolated as a white solid.

m.p. 141-148 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.06 (s, 1 H), 8.97 (s, 1 H), 8.55-8.51 (m, 1 H), 8.25 (s, 1 H), 8.05 (s, 1 H), 7.66 (s, 1 H), 7.56 (d, $J=8.4$ Hz, 1 H), 7.51 (d, $J=8.8$ Hz, 1 H), 7.42-7.38 (m, 4 H), 7.29 (d, $J=8.4$ Hz, 1 H), 5.10 (s (br), 1 H), 4.86-4.81 (m, 1 H), 4.54 (d, $J=9.2$ Hz, 1 H), 4.47-4.44 (m, 1 H), 4.38-4.34 (m, 2 H), 4.29-4.26 (m, 1 H), 3.99-3.94 (m, 2 H), 3.77-3.75 (m, 2 H), 3.69-3.62 (m, 2 H), 3.54 (s, 3 H), 3.16 (s, 6 H), 2.68-2.65 (m, 1 H), 2.62-2.60 (m, 1 H), 2.44 (s, 3 H), 2.09 (s, 3 H), 2.06-2.02 (m, 1 H), 1.94-1.90 (m, 1 H), 1.30 (s, 3 H), 1.25 (s, 3 H), 0.97 (s, 9 H);

δ_{C} (101 MHz, d_6 -DMSO) δ ppm 175.6, 171.8, 169.7, 169.5, 161.7, 151.4, 151.4, 147.7, 139.5, 139.4, 139.2, 137.1, 134.2, 133.1, 131.1, 129.7, 128.6, 127.5, 127.5, 118.5, 115.3, 107.8, 103.2, 69.4, 68.8, 58.7, 58.4, 56.9, 56.6, 56.2, 45.4, 41.7, 40.8, 37.8, 37.5, 35.7, 26.3, 25.7, 25.6, 16.9, 15.8; **LCMS** (Method B): $t_{\text{R}} = 0.93$, (m/z) $[\text{M}+\text{H}]^+$ 897 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{47}\text{H}_{61}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 897.4333$, found 897.4332; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3280, 3069, 2931, 2868, 1651, 1609, 1540, 1487; $[\alpha]_{\text{D}}^{20} = -25$ (c 0.13, MeOH).

***N*1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*5-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-3,3-dimethylpentanediamide 166**

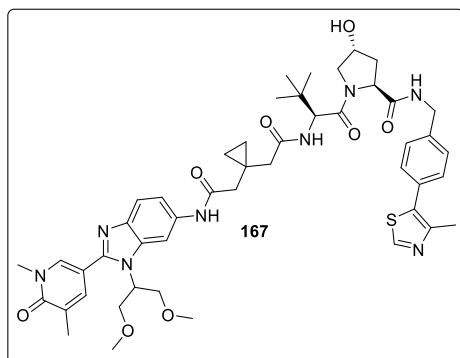


166 was synthesised following the **general procedure E** with the linker 5-methoxy-3,3-dimethyl-5-oxopentanoic acid (52.0 mg, 0.300 mmol). **166** (58 mg, 0.064 mmol, 20% yield) was isolated as a white solid.

m.p. 167-174 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.18 (s, 1 H), 8.96 (s, 1 H), 8.54-8.52 (m, 1 H), 8.21 (s, 1 H), 8.19-8.17 (m, 1 H), 8.04 (s, 1 H), 7.65 (s, 1 H), 7.54 (d, $J=8.8$ Hz, 1 H), 7.41-7.36 (m, 4 H), 7.29 (d, $J=8.8$ Hz, 1 H), 5.10 (s (br), 1 H), 4.85-4.80 (m, 1 H), 4.54 (d, $J=9.2$ Hz, 1 H), 4.47-4.44 (m, 1 H), 4.42-4.37 (m, 2 H), 4.26-4.23 (m, 1 H), 3.98-3.95 (m, 2 H), 3.77-3.69 (m, 4 H), 3.54 (s, 3 H), 3.16 (s, 6 H), 2.48-2.45 (m, 1 H), 2.43 (s, 3 H), 2.39-2.33 (m, 3 H), 2.09 (s, 3 H), 2.06-2.01 (m, 1 H), 1.95-1.91 (m, 1 H), 1.07 (s, 6 H), 0.97 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.8, 170.9, 170.1, 169.8, 161.7, 151.5, 151.3, 147.6, 139.4, 139.1, 138.8, 137.2, 134.0, 133.3, 131.1, 129.7, 128.6, 127.5, 127.4, 118.7, 115.2, 107.6, 103.2, 69.5, 68.8, 58.6, 58.4, 56.8, 56.7, 56.2, 47.4, 46.0, 41.6, 38.0, 37.5, 34.9, 33.6, 28.3, 28.2, 26.4, 16.9, 15.8; **LCMS** (Method B): $t_{\text{R}} = 0.95$, (m/z) $[\text{M}+\text{H}]^+$ 911 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{48}\text{H}_{63}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 911.4490$, found 911.4486;

IR ($\nu_{\max}/\text{cm}^{-1}$) 3270, 3069, 2952, 2878, 1656, 1603, 1545, 1489; $[\alpha]_D^{20} = +13$ (*c* 0.22, MeOH).

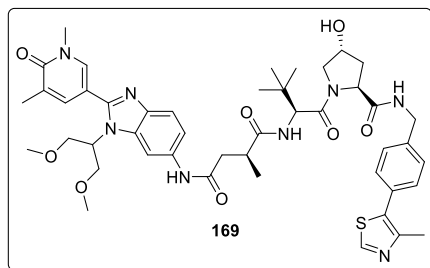
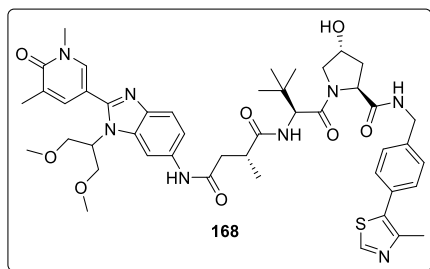
(2*S*,4*R*)-1-((*S*)-2-(2-(1-(2-((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)amino)-2-oxoethyl)cyclopropyl)acetamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide 167



167 was synthesised following the **general procedure E** with the linker 2-(1-(2-methoxy-2-oxoethyl)cyclopropyl)acetic acid (52.0 mg, 0.300 mmol). **167** (90 mg, 0.099 mmol, 33% yield) was isolated as a white solid.

m.p. 151-155 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 9.94 (s, 1 H), 8.96 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 8.23 (s, 1 H), 8.04 (s, 1 H), 7.93 (d, $J=9.2$ Hz, 1 H), 7.66 (s, 1 H), 7.53 (d, $J=8.8$ Hz, 1 H), 7.42-7.37 (m, 4 H), 7.29 (d, $J=8.6$ Hz, 1 H), 5.13 (s (br), 1 H), 4.84-4.79 (m, 1 H), 4.57 (d, $J=9.5$ Hz, 1 H), 4.47-4.37 (m, 3 H), 4.25-4.22 (m, 1 H), 3.98-3.95 (m, 2 H), 3.77-3.75 (m, 2 H), 3.70-3.68 (m, 2 H), 3.53 (s, 3 H), 3.17 (s, 6 H), 2.62-2.60 (m, 1 H), 2.45-2.42 (m, 4 H), 2.26-2.19 (m, 2 H), 2.10 (s, 3 H), 2.06-2.02 (m, 1 H), 1.94-1.90 (m, 1 H), 0.93 (s, 9 H), 0.59-0.53 (m, 2 H), 0.49-0.45 (m, 2 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.8, 170.1, 169.9, 169.7, 161.7, 151.5, 151.3, 147.7, 139.4, 139.2, 138.9, 137.3, 134.1, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 118.9, 114.8, 108.0, 102.8, 69.5, 68.8, 58.7, 58.4, 56.7, 56.5, 56.3, 43.2, 41.7, 41.6, 37.9, 37.4, 35.1, 26.3, 16.9, 16.1, 15.9, 11.8, 11.1; **LCMS** (Method B): $t_{\text{R}} = 0.93$, (m/z) $[\text{M}+\text{H}]^+$ 909 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{48}\text{H}_{61}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 909.4363$, found 909.4348; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3286, 3074, 2958, 2884, 1651, 1603, 1529, 1487; $[\alpha]_D^{20} = -8$ (*c* 0.26, MeOH).

(R)-N4-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N1-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-2-methylsuccinamide 168 and **(S)-N4-(1-(1,3-dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N1-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-2-methylsuccinamide 169**



168 and **169** were synthesised following the **general procedure E** with the linker 4-methoxy-2-methyl-4-oxobutanoic acid (44.0 mg, 0.300 mmol). **168** (72 mg, 0.081 mmol, 27% yield) and **169** (75 mg, 0.085 mmol, 28% yield) were both isolated as white solids.

168

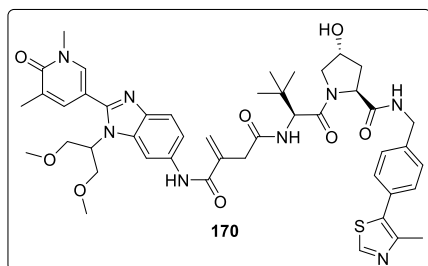
m.p. 163-169 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.96 (s, 1 H), 8.98 (s, 1 H), 8.53 (t, $J=6.1$ Hz, 1 H), 8.25 (s, 1 H), 8.01 (s, 1 H), 7.75 (d, $J=9.5$ Hz, 1 H), 7.65 (s, 1 H), 7.52 (d, $J=8.8$ Hz, 1 H), 7.42-7.38 (m, 4 H), 7.24 (d, $J=8.8$ Hz, 1 H), 5.12 (s (br), 1 H), 4.83-4.78 (m, 1 H), 4.53 (d, $J=9.5$ Hz, 1 H), 4.46-4.39 (m, 2 H), 4.37-4.35 (m, 1 H), 4.25-4.22 (m, 1 H), 3.97-3.92 (m, 2 H), 3.76-3.73 (m, 2 H), 3.68-3.66 (m, 1 H), 3.62-3.60 (m, 1 H), 3.53 (s, 3 H), 3.16 (s, 6 H), 3.08-3.05 (m, 1 H), 2.67-2.63 (m, 1 H), 2.44 (s, 3 H), 2.40-2.36 (m, 1 H), 2.09 (s, 3 H), 2.06-2.02 (m, 1 H), 1.93-1.89 (m, 1 H), 1.06 (s, 3 H), 0.93 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 174.5, 171.9, 169.6, 169.5, 161.7, 151.5, 151.4, 147.7, 139.4, 139.1, 138.9, 137.3, 134.2, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 118.9, 114.6, 107.9, 102.7, 69.4, 68.8, 58.7, 58.4, 56.7, 56.3, 56.2, 41.6, 40.8, 37.9, 37.4, 35.6, 35.4, 26.3, 17.5, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.87$, (m/z) $[M+H]^+$ 883 (99% purity); **HRMS** (ESI): calculated for $C_{46}H_{59}N_8O_8S$ (m/z)

$[M+H]^+ = 883.4177$, found 883.4158; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3296, 3069, 2952, 2878, 1651, 1603, 1540, 1487; $[\alpha]_D^{20} = -57$ (c 0.38, MeOH).

169

m.p. 157-161 °C; **δ_H** (400 MHz, d_6 -DMSO) δ ppm 10.00 (s, 1 H), 8.97 (s, 1 H), 8.51 (t, $J=6.1$ Hz, 1 H), 8.17 (s, 1 H), 8.02 (s, 1 H), 7.97 (d, $J=9.2$ Hz, 1 H), 7.65 (s, 1 H), 7.53 (d, $J=8.8$ Hz, 1 H), 7.42-7.38 (m, 4 H), 7.29 (d, $J=8.6$ Hz, 1 H), 5.07 (s (br), 1 H), 4.83-4.79 (m, 1 H), 4.54 (d, $J=9.5$ Hz, 1 H), 4.46-4.39 (m, 2 H), 4.36-4.33 (m, 1 H), 4.25-4.22 (m, 1 H), 3.98-3.94 (m, 2 H), 3.77-3.74 (m, 2 H), 3.70-3.68 (m, 1 H), 3.64-3.62 (m, 1 H), 3.53 (s, 3 H), 3.17 (s, 6 H), 3.09-3.05 (m, 1 H), 2.64-2.61 (m, 1 H), 2.44 (s, 3 H), 2.39-2.34 (m, 1 H), 2.08 (s, 3 H), 2.06-2.01 (m, 1 H), 1.92-1.88 (m, 1 H), 1.13 (s, 3 H), 0.97 (s, 9 H); **δ_C** (101 MHz, d_6 -DMSO) δ ppm 174.7, 171.9, 169.9, 169.5, 161.6, 151.5, 151.4, 147.7, 139.4, 139.2, 138.9, 137.3, 134.1, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 118.9, 114.8, 108.0, 102.7, 69.5, 68.8, 58.7, 58.4, 56.8, 56.3, 56.1, 41.6, 40.0, 37.8, 37.4, 35.7, 35.5, 26.3, 18.4, 16.9, 15.9; **LCMS** (Method B): $t_R = 0.91$, (m/z) $[M+H]^+ 883$ (100% purity); **HRMS** (ESI): calculated for $C_{46}H_{59}N_8O_8S$ (m/z) $[M+H]^+ = 883.4177$, found 883.4174; **IR** ($\nu_{\max}/\text{cm}^{-1}$) 3286, 3069, 2958, 2884, 1656, 1598, 1529, 1482; $[\alpha]_D^{20} = +0.04$ (c 0.28, MeOH).

N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N4-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-2-methylenesuccinamide 170

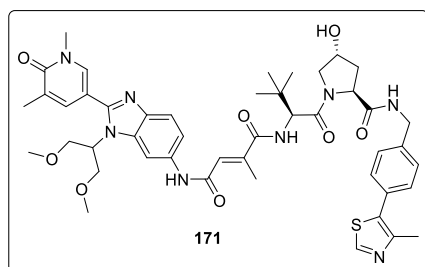


170 was synthesised following the **general procedure E** with the linker 3-(methoxycarbonyl)but-3-enoic acid **201** (43.0 mg, 0.300 mmol). **170** (38 mg, 0.043 mmol, 28% yield) was isolated as a white solid.

m.p. 153-159 °C; **δ_H** (400 MHz, d_6 -DMSO) δ ppm 10.12 (s, 1 H), 8.98 (s, 1 H), 8.55 (t, $J=6.1$ Hz, 1 H), 8.26 (s, 1 H), 8.04-8.00 (m, 2 H), 7.67 (m, 1 H), 7.55 (d, $J=8.4$ Hz,

1 H), 7.44-7.38 (m, 5 H), 5.98 (s, 1 H), 5.63 (s, 1 H), 5.11 (s (br), 1 H), 4.84-4.79 (m, 1 H), 4.54 (d, $J=9.5$ Hz, 1 H), 4.46-4.43 (m, 2 H), 4.36-4.33 (m, 1 H), 4.26-4.20 (m, 1 H), 3.99-3.94 (m, 2 H), 3.78-3.75 (m, 2 H), 3.68-3.61 (m, 2 H), 3.54 (s, 3 H), 3.52-3.47 (m, 1 H), 3.28-3.26 (m, 1 H), 3.17 (s, 6 H), 2.44 (s, 3 H), 2.09 (s, 3 H), 2.05-2.01 (m, 1 H), 1.92-1.88 (m, 1 H), 0.93 (s, 9 H); δ_c (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.4, 169.4, 166.0, 161.7, 151.7, 151.4, 147.7, 139.9, 139.4, 139.4, 138.9, 137.2, 133.9, 133.3, 131.1, 129.6, 128.6, 127.4, 127.4, 121.8, 118.8, 115.5, 107.9, 103.5, 69.4, 68.8, 58.7, 58.4, 56.8, 56.4, 56.3, 41.6, 39.0, 37.9, 37.4, 35.5, 26.2, 16.9, 15.9; **LCMS** (Method B): $t_R = 0.89$, (m/z) $[M+H]^+$ 881 (98% purity); **HRMS** (ESI): calculated for $C_{46}H_{57}N_8O_8S$ (m/z) $[M+H]^+ = 881.4042$, found 881.4031; **IR** (ν_{max}/cm^{-1}) 3286, 3074, 2952, 2868, 1651, 1603, 1535, 1482; $[\alpha]_D^{20} = -29$ (c 0.12, MeOH).

N4-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N1-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-2-methylene succinamide 171

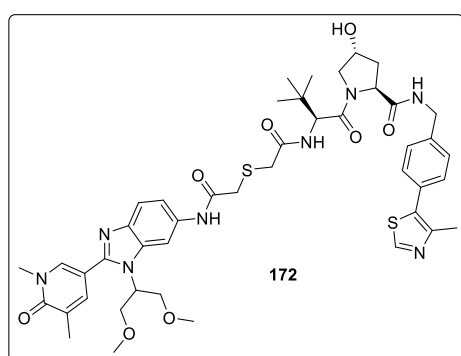


171 was synthesised following the **general procedure E** with the linker 4-methoxy-2-methylene-4-oxobutanoic acid **202** (43.0 mg, 0.300 mmol). **171** (55 mg, 0.062 mmol, 21% yield) was isolated as a white solid.

m.p. 164-170 °C; δ_H (400 MHz, d_6 -DMSO) δ ppm 10.30 (s, 1 H), 8.98 (s, 1 H), 8.57 (t, $J=6.1$ Hz, 1 H), 8.29 (s, 1 H), 8.03 (s, 1 H), 7.68-7.65 (m, 2 H), 7.56 (d, $J=8.4$ Hz, 1 H), 7.43-7.35 (m, 5 H), 6.60 (s, 1 H), 5.15 (s (br), 1 H), 4.86-4.81 (m, 1 H), 4.65 (d, $J=9.2$ Hz, 1 H), 4.49-4.37 (m, 3 H), 4.27-4.24 (m, 1 H), 4.00-3.97 (m, 2 H), 3.79-3.76 (m, 2 H), 3.73-3.67 (m, 2 H), 3.54 (s, 3 H), 3.18 (s, 6 H), 2.45 (s, 3 H), 2.34 (s, 3 H), 2.09 (s, 3 H), 2.07-2.06 (m, 1 H), 1.95-1.90 (m, 1 H), 1.03 (s, 9 H); δ_c (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.2, 168.5, 163.6, 161.7, 151.7, 151.4, 147.7, 144.2, 139.5, 139.4, 138.9, 137.2, 133.9, 133.5, 131.1, 129.7, 128.6, 127.4, 127.4, 125.1, 119.0, 114.8, 107.9, 102.9, 69.5, 68.9, 58.8, 58.4, 56.9, 56.8, 56.5, 41.7, 37.9, 37.5, 35.7, 26.4,

16.9, 15.9, 14.4; **LCMS** (Method B): $t_R = 0.94$, (m/z) $[M+H]^+$ 881 (100% purity); **HRMS** (ESI): calculated for $C_{46}H_{57}N_8O_8S$ (m/z) $[M+H]^+ = 881.4020$, found 881.4017; **IR** (ν_{max}/cm^{-1}) 3286, 3074, 2947, 2873, 1656, 1609, 1513, 1482; $[\alpha]_D^{20} = -28$ (c 0.19, MeOH).

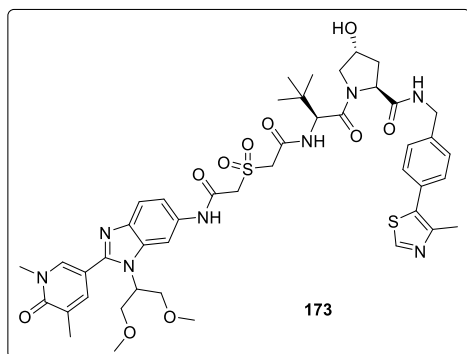
(2*S*,4*R*)-1-((*S*)-2-(2-((2-((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)amino)-2-oxoethyl)thio)acetamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide **172**



172 was synthesised following the **general procedure E** with the linker 2-((2-methoxy-2-oxoethyl)thio)acetic acid (49.0 mg, 0.300 mmol). **172** (69 mg, 0.076 mmol, 23% yield) was isolated as a white solid.

m.p. 115-123 °C; **δ_H** (400 MHz, d_6 -DMSO) δ ppm 10.23 (s, 1 H), 8.97 (s, 1 H), 8.57-8.55 (m, 1 H), 8.20 (s, 1 H), 8.03 (s, 1 H), 7.66 (m, 1 H), 7.55 (d, $J=8.4$ Hz, 1 H), 7.42-7.37 (m, 5 H), 7.32 (d, $J=8.8$ Hz, 1 H), 5.16 (s (br), 1 H), 4.84-4.80 (m, 1 H), 4.55 (d, $J=9.5$ Hz, 1 H), 4.47-4.43 (m, 2 H), 4.39-4.34 (m, 1 H), 4.26-4.22 (m, 1 H), 3.99-3.95 (m, 2 H), 3.77-3.75 (m, 2 H), 3.69-3.64 (m, 2 H), 3.55-3.47 (m, 7 H), 3.17 (s, 6 H), 2.44 (s, 3 H), 2.09 (s, 3 H), 2.06-2.02 (m, 1 H), 1.94-1.89 (m, 1 H), 0.97 (s, 9 H); **δ_C** (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.3, 168.7, 167.4, 161.7, 151.7, 151.4, 147.7, 139.4, 139.4, 138.9, 137.3, 133.8, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 119.0, 114.8, 107.9, 103.0, 69.5, 68.8, 58.7, 58.4, 56.8, 56.7, 56.4, 41.6, 37.9, 37.5, 36.1, 35.5, 34.5, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_R = 0.90$, (m/z) $[M+H]^+$ 901 (92% purity); **HRMS** (ESI): calculated for $C_{45}H_{57}N_8O_8S_2$ (m/z) $[M+H]^+ = 901.3741$, found 901.3736; **IR** (ν_{max}/cm^{-1}) 3275, 3079, 2952, 2878, 1651, 1609, 1540, 1487; $[\alpha]_D^{20} = -7$ (c 0.20, MeOH).

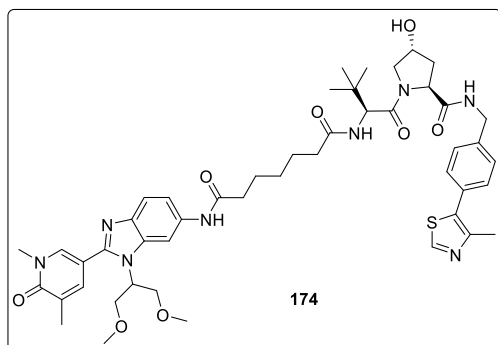
(2*S*,4*R*)-1-((*S*)-2-(2-((2-((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)amino)-2-oxoethyl)sulfonyl)acetamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide 173



173 was synthesised following the **general procedure E** with the linker 2-((2-methoxy-2-oxoethyl)sulfonyl)acetic acid (59.0 mg, 0.300 mmol). **173** (78 mg, 0.084 mmol, 28% yield) was isolated as a white solid.

m.p. 176-184 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.52 (s, 1 H), 8.97 (s, 1 H), 8.60-8.37 (m, 2 H), 8.22 (s, 1 H), 8.05 (s, 1 H), 7.66 (s, 1 H), 7.60 (d, $J=8.8$ Hz, 1 H), 7.43-7.38 (m, 4 H), 7.31 (d, $J=8.8$ Hz, 1 H), 5.15 (s (br), 1 H), 4.86-4.82 (m, 1 H), 4.60-4.37 (m, 2 H), 4.51-4.37 (m, 6), 4.24-4.21 (m, 1 H), 3.98-3.95 (m, 2 H), 3.78-3.76 (m, 2 H), 3.72-3.70 (m, 1 H), 3.63-3.61 (m, 1 H), 3.53 (s, 3 H), 3.18 (s, 6 H), 2.44 (s, 3 H), 2.09 (s, 3 H), 2.06-2.03 (m, 1 H), 1.94-1.90 (m, 1 H), 0.98 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.8, 168.7, 161.7, 161.5, 160.2, 151.9, 151.4, 147.7, 139.4, 139.2, 139.2, 137.1, 133.3, 133.3, 131.1, 129.6, 128.6, 127.5, 127.4, 119.0, 115.0, 107.4, 103.3, 69.5, 68.8, 59.7, 58.7, 57.7, 56.9, 56.8, 56.5, 56.4, 41.6, 37.9, 37.5, 35.7, 26.2, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.89$, (m/z) $[\text{M}+\text{H}]^+$ 933 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{45}\text{H}_{57}\text{N}_8\text{O}_{10}\text{S}_2$ (m/z) $[\text{M}+\text{H}]^+ = 933.3639$, found 933.3633; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3280, 3069, 2921, 2873, 1646, 1603, 1540, 1489; $[\alpha]_{\text{D}}^{20} = -10$ (c 0.26, MeOH).

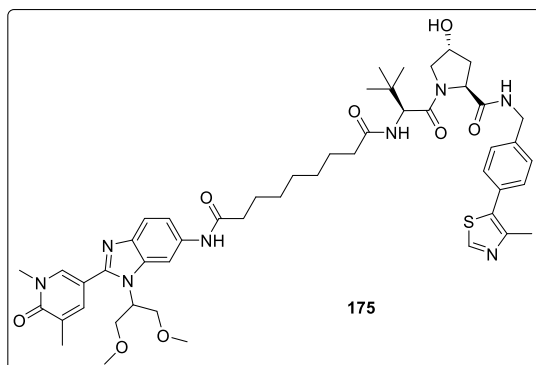
***N*1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*7-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)heptanediamide 174**



174 was synthesised following the **general procedure E** with the linker 7-methoxy-7-oxoheptanoic acid (52.0 mg, 0.300 mmol). **174** (154 mg, 0.170 mmol, 56% yield) was isolated as a white solid.

m.p. 148-156 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 9.89 (s, 1 H), 8.97 (s, 1 H), 8.53 (t, $J=6.1$ Hz, 1 H), 8.21 (s, 1 H), 8.02 (s, 1 H), 7.83 (d, $J=9.5$ Hz, 1 H), 7.66 (s, 1 H), 7.52 (d, $J=8.4$ Hz, 1 H), 7.42-7.37 (m, 4 H), 7.29 (d, $J=8.6$ Hz, 1 H), 5.12 (s, 1 H), 4.83-4.79 (m, 1 H), 4.55 (d, $J=9.5$ Hz, 1 H), 4.45-4.41 (m, 2 H), 4.37-4.34 (m, 1 H), 4.24-4.20 (m, 1 H), 3.98-3.95 (m, 2 H), 3.77-3.74 (m, 2 H), 3.69-3.64 (m, 2 H), 3.53 (s, 3 H), 3.17 (s, 6 H), 2.44 (s, 3 H), 2.33-2.26 (m, 3 H), 2.18-2.13 (m, 1 H), 2.09 (s, 3 H), 2.05-2.02 (m, 1 H), 1.93-1.89 (m, 1 H), 1.65-1.60 (m, 2 H), 1.58-1.49 (m, 2 H), 1.34-1.29 (m, 2 H), 0.93 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 172.0, 171.9, 171.0, 169.7, 161.7, 151.5, 151.4, 147.7, 139.4, 139.1, 138.9, 137.3, 134.3, 133.4, 131.0, 129.6, 128.6, 127.4, 127.4, 118.9, 114.7, 108.0, 102.7, 69.5, 68.8, 58.7, 58.4, 56.7, 56.3, 56.3, 41.6, 37.9, 37.4, 36.3, 35.2, 34.8, 28.4, 26.3, 25.2, 24.8, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.92$, (m/z) $[\text{M}+\text{H}]^+$ 911 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{48}\text{H}_{63}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 911.4490$, found 911.4493; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3280, 3074, 2937, 2863, 1656, 1598, 1540, 1482; $[\alpha]_{\text{D}}^{20} = -24$ (c 0.23, MeOH).

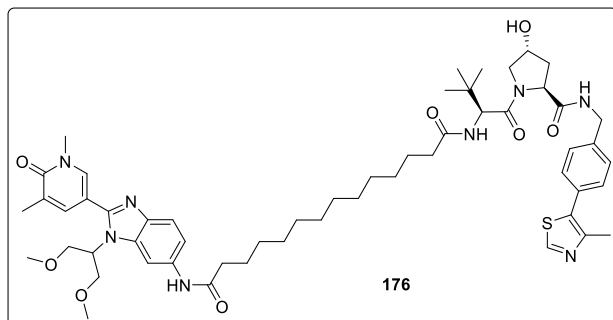
N*1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*9-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)nonanediamide **175*



175 was synthesised following the **general procedure E** with the linker 9-methoxy-9-oxononanoic acid (61.0 mg, 0.300 mmol). **175** (110 mg, 0.120 mmol, 38% yield) was isolated as a white solid.

m.p. 127-135 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.93 (s, 1 H), 8.97 (s, 1 H), 8.53 (t, $J=6.1$ Hz, 1 H), 8.25 (s, 1 H), 8.06 (s, 1 H), 7.81 (d, $J=9.5$ Hz, 1 H), 7.66 (s, 1 H), 7.55 (d, $J=8.8$ Hz, 1 H), 7.42-7.37 (m, 4 H), 7.33 (d, $J=8.8$ Hz, 1 H), 5.11 (s (br), 1 H), 4.86-4.82 (m, 1 H), 4.54 (d, $J=9.5$ Hz, 1 H), 4.44-4.41 (m, 2 H), 4.37-4.34 (m, 1 H), 4.23-4.20 (m, 1 H), 3.98-3.95 (m, 2 H), 3.78-3.75 (m, 2 H), 3.68-3.63 (m, 2 H), 3.53 (s, 3 H), 3.17 (s, 6 H), 2.44 (s, 3 H), 2.34-2.31 (m, 2 H), 2.29-2.24 (m, 1 H), 2.14-2.11 (m, 1 H), 2.08 (s, 3 H), 2.05-2.01 (m, 1 H), 1.93-1.88 (m, 1 H), 1.63-1.59 (m, 2 H), 1.55-1.44 (m, 2 H), 1.34-1.23 (m, 6 H), 0.91 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 172.0, 171.9, 171.1, 169.7, 161.7, 151.4, 151.2, 147.7, 139.4, 139.3, 137.6, 137.1, 134.7, 133.1, 131.1, 129.6, 128.6, 127.5, 127.4, 118.4, 115.2, 107.0, 102.9, 69.4, 68.8, 58.6, 58.4, 57.0, 56.3, 56.2, 41.6, 37.9, 37.5, 36.4, 35.1, 34.8, 28.6, 28.5, 28.5, 26.3, 25.4, 25.0, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.98$, (m/z) $[\text{M}+\text{H}]^+ 939$ (100% purity); **HRMS** (ESI): calculated for $\text{C}_{50}\text{H}_{67}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 939.4803$, found 939.4808; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3280, 3074, 2937, 2931, 2863, 1656, 1609, 1535; $[\alpha]_{\text{D}}^{20} = -26$ (c 0.26, MeOH).

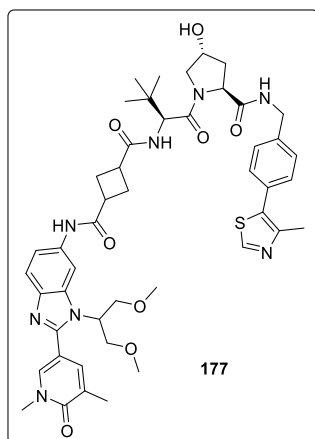
N*1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*14-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)tetradecanediamide **176*



176 was synthesised following the **general procedure E** with the linker 14-methoxy-14-oxotetradecanoic acid (82.0 mg, 0.300 mmol). **176** (83 mg, 0.082 mmol, 27% yield) was isolated as a white solid.

m.p. 119-128 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.88 (s, 1 H), 8.97 (s, 1 H), 8.52 (t, $J=6.1$ Hz, 1 H), 8.20 (s, 1 H), 8.01 (s, 1 H), 7.79 (d, $J=9.2$ Hz, 1 H), 7.65 (s, 1 H), 7.52 (d, $J=8.8$ Hz, 1 H), 7.42-7.37 (m, 4 H), 7.28 (d, $J=8.6$ Hz, 1 H), 5.09 (s (br), 1 H), 4.83-4.78 (m, 1 H), 4.53 (d, $J=9.5$ Hz, 1 H), 4.44-4.41 (m, 2 H), 4.36-4.33 (m, 1 H), 4.23-4.18 (m, 1 H), 3.97-3.94 (m, 2 H), 3.77-3.74 (m, 2 H), 3.67-3.64 (m, 2 H), 3.52 (s, 3 H), 3.16 (s, 6 H), 2.44 (s, 3 H), 2.33-2.30 (m, 2 H), 2.26-2.22 (m, 1 H), 2.12-2.09 (m, 1 H), 2.09 (s, 3 H), 2.03-2.01 (m, 1 H), 1.92-1.88 (m, 1 H), 1.62-1.58 (m, 2 H), 1.52-1.41 (m, 2 H), 1.31-1.20 (m, 16 H), 0.93 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 172.0, 171.9, 171.0, 169.7, 161.6, 151.4, 151.3, 147.7, 139.5, 139.1, 138.9, 137.2, 134.3, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 118.9, 114.7, 107.9, 102.7, 69.5, 68.8, 58.6, 58.4, 56.7, 56.2, 56.2, 41.6, 37.9, 37.4, 36.5, 35.1, 34.8, 28.9, 28.9, 28.9, 28.9, 28.9, 28.8, 28.7, 28.6, 26.3, 25.4, 25.1, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 1.19$, (m/z) $[\text{M}+2\text{H}]^{2+}$ 505 (95% purity); **HRMS** (ESI): calculated for $\text{C}_{55}\text{H}_{77}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+$ = 1009.5585, found 1009.5587; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3280, 3069, 2921, 2852, 1651, 1609, 1545, 1482; $[\alpha]_{\text{D}}^{20} = -25$ (c 0.14, MeOH).

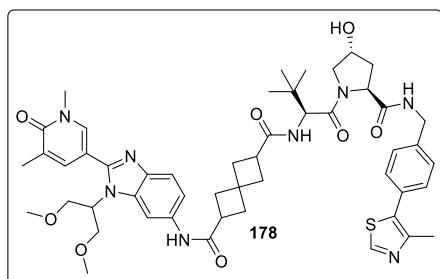
N1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*3-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)cyclobutane-1,3-dicarboxamide 177



177 was synthesised following the **general procedure E** with the linker 3-(methoxycarbonyl)cyclobutane-1-carboxylic acid (47.0 mg, 0.300 mmol). **177** (112 mg, 0.130 mmol, 42% yield) was isolated as a white solid.

m.p. 171-177 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.80 (s, 1 H), 9.00 (s, 1 H), 8.56-8.53 (m, 1 H), 8.26 (s, 1 H), 8.04 (s, 1 H), 7.76 (d, $J=9.2$ Hz, 1 H), 7.66 (s, 1 H), 7.54 (d, $J=8.8$ Hz, 1 H), 7.43-7.38 (m, 4 H), 7.32-7.29 (m, 1 H), 5.13 (s (br), 1 H), 4.85-4.79 (m, 1 H), 4.59-4.55 (m, 1 H), 4.46-4.41 (m, 2 H), 4.38-4.35 (m, 1 H), 4.24-4.21 (m, 1 H), 3.99-3.96 (m, 2 H), 3.78-3.75 (m, 2 H), 3.70-3.65 (m, 2 H), 3.53 (s, 3 H), 3.18 (s, 6 H), 3.17-3.08 (m, 1 H), 2.45 (s, 3 H), 2.43-2.15 (m, 5 H), 2.09 (s, 3 H), 2.06-2.02 (m, 1 H), 1.95-1.89 (m, 1 H), 0.95 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 172.9, 172.0, 171.9, 169.6, 161.7, 151.4, 151.4, 147.7, 139.5, 139.0, 138.6, 137.2, 134.3, 133.3, 131.1, 129.6, 128.6, 127.5, 127.4, 118.7, 114.9, 107.6, 102.8, 69.4, 68.9, 58.7, 58.4, 56.9, 56.4, 56.4, 41.6, 37.9, 37.5, 35.5, 35.3, 33.7, 28.2, 27.2, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.90$, (m/z) $[\text{M}+\text{H}]^+ 895$ (100% purity); **HRMS** (ESI): calculated for $\text{C}_{47}\text{H}_{59}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 895.4177$, found 895.4169; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3280, 3064, 2952, 2873, 1651, 1603, 1524, 1482; $[\alpha]_{\text{D}}^{20} = -46$ (c 0.25, MeOH).

N2-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N6-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)spiro[3.3]heptane-2,6-dicarboxamide 178

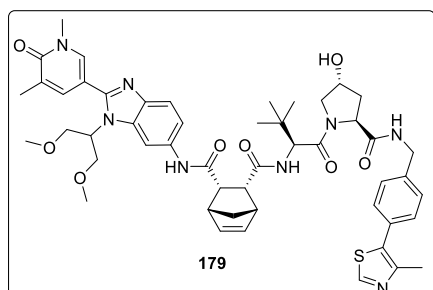


178 was synthesised following the **general procedure E** with the linker 6-(methoxycarbonyl)spiro[3.3]heptane-2-carboxylic acid (59.0 mg, 0.300 mmol). **178** (73 mg, 0.078 mmol, 26% yield) was isolated as a white solid.

m.p. 171-177 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.78 (s, 1 H), 8.97 (s, 1 H), 8.55-8.52 (m, 1 H), 8.22 (s, 1 H), 8.02 (s, 1 H), 7.67-7.63 (m, 2 H), 7.52 (d, $J=8.4$ Hz, 1 H), 7.42-7.38 (m, 4 H), 7.29-7.26 (m, 1 H), 5.12 (s (br), 1 H), 4.83-4.79 (m, 1 H), 4.53 (d, $J=9.2$ Hz, 1 H), 4.44-4.41 (m, 2 H), 4.37-4.35 (m, 1 H), 4.24-4.20 (m, 1 H), 3.98-3.94 (m, 2 H), 3.77-3.74 (m, 2 H), 3.69-3.63 (m, 2 H), 3.52 (s, 3 H), 3.19 (s, 6 H), 3.15-3.07 (m, 2 H), 2.44 (s, 3 H), 2.30-2.10 (m, 6 H), 2.09-2.07 (s, 3 H), 2.06-2.00 (m, 3 H), 1.94-1.88 (m, 1 H), 0.92 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 173.5, 172.5, 171.9, 169.6, 161.7, 151.5, 151.4, 147.7, 139.4, 139.1, 138.9, 137.3, 134.3, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 118.9, 114.7, 108.0, 102.7, 69.5, 68.8, 58.7, 58.4, 56.8, 56.4, 56.3, 41.6, 38.1, 37.9, 37.6, 37.4, 36.9, 35.3, 34.4, 32.7, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.95$, (m/z) $[\text{M}+\text{H}]^+$ 935 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{50}\text{H}_{63}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 935.4490$, found 935.4490; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3286, 3069, 2921, 2873, 1651, 1603, 1529, 1482; $[\alpha]_{\text{D}}^{20} = -32$ (c 0.16, MeOH).

178 was isolated as a diastereomeric mixture.

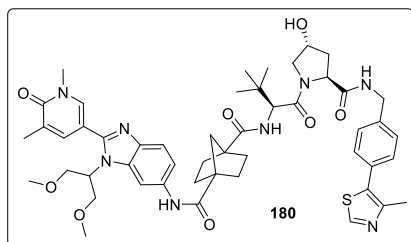
(1*R*,2*S*,3*R*,4*S*)-*N*2-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*3-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)bicyclo[2.2.1]hept-5-ene-2,3-dicarboxamide 179



179 was synthesised following the **general procedure E** with the linker (1*S*,2*R*,3*S*,4*R*)-3-(methoxycarbonyl)bicyclo[2.2.1]hept-5-ene-2-carboxylic acid 59.0 mg, 0.300 mmol). **179** (33 mg, 0.036 mmol, 12% yield) was isolated as a white solid.

m.p. 165-171 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 9.64 (s, 1 H), 8.99 (s, 1 H), 8.48 (t, $J=6.1$ Hz, 1 H), 8.14 (s, 1 H), 8.03 (s, 1 H), 7.66 (s, 1 H), 7.57 (d, $J=9.2$ Hz, 1 H), 7.47 (d, $J=8.8$ Hz, 1 H), 7.40-7.36 (m, 4 H), 7.20 (d, $J=8.4$ Hz, 1 H), 6.18-6.13 (m, 2 H), 5.12 (s (br), 1 H), 4.82-4.78 (m, 1 H), 4.43-4.32 (m, 4 H), 4.22-4.19 (m, 1 H), 3.95-3.91 (m, 2 H), 3.75-3.73 (m, 2 H), 3.60-3.56 (m, 2 H), 3.53 (s, 3 H), 2.42-2.36 (m, 2 H), 3.16 (s, 6 H), 3.06-3.04 (m, 1 H), 3.00-2.98 (m, 1 H), 2.42 (s, 3 H), 2.08 (s, 3 H), 2.03-1.99 (m, 1 H), 1.90-1.88 (m, 1 H), 1.29-1.23 (m, 2 H), 0.76 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.8, 170.9, 170.1, 169.9, 161.6, 151.4, 151.0, 147.7, 139.4, 139.1, 139.1, 137.2, 135.0, 134.5, 134.2, 133.0, 131.1, 129.6, 128.6, 127.4, 127.4, 118.4, 114.8, 107.6, 102.1, 69.3, 68.8, 58.6, 58.4, 56.8, 56.6, 56.1, 50.9, 50.1, 48.3, 46.7, 46.6, 41.6, 37.9, 37.5, 34.9, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.92$, (m/z) $[\text{M}+\text{H}]^+$ 933 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{50}\text{H}_{61}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 933.4333$, found 933.4350; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3280, 3069, 2968, 2873, 1656, 1603, 1535, 1482, 1439; **$[\alpha]_{\text{D}}^{20} = +78$** (c 0.13, MeOH).

N*1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*4-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)bicyclo[2.2.1]heptane-1,4-dicarboxamide **180*

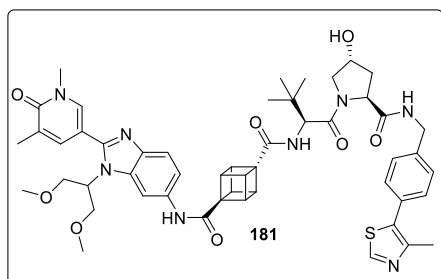


180 was synthesised following the **general procedure E** with the linker 4-(methoxycarbonyl)bicyclo[2.2.1]heptane-1-carboxylic acid (59.0 mg, 0.300 mmol). **180** (81 mg, 0.087 mmol, 29% yield) was isolated as a

white solid.

m.p. 174-178 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.29 (s, 1 H), 8.99 (s, 1 H), 8.56 (t, $J=6.1$ Hz, 1 H), 8.17 (s, 1 H), 8.03 (s, 1 H), 7.68 (s, 1 H), 7.53 (d, $J=8.4$ Hz, 1 H), 7.47 (d, $J=8.8$ Hz, 1 H), 7.42-7.40 (m, 4 H), 6.88 (d, $J=9.2$ Hz, 1 H), 5.16 (s (br), 1 H), 4.84-4.80 (m, 1 H), 4.63 (d, $J=9.5$ Hz, 1 H), 4.48-4.37 (m, 3 H), 4.30-4.24 (m, 1 H), 4.00-3.96 (m, 2 H), 3.78-3.75 (m, 2 H), 3.70-3.66 (m, 1 H), 3.64-3.61 (m, 1 H), 3.55 (s, 3 H), 3.19 (s, 6 H), 2.45 (s, 3 H), 2.09 (s, 3 H), 2.07-1.99 (m, 5 H), 1.95-1.90 (m, 3 H), 1.80-1.73 (m, 4 H), 0.97 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 173.2, 172.8, 171.8, 169.4, 161.7, 151.5, 151.4, 147.7, 139.4, 139.2, 138.9, 137.3, 134.1, 133.3, 131.1, 129.7, 128.6, 127.4, 127.4, 118.7, 115.5, 108.0, 103.5, 69.4, 68.8, 58.8, 58.4, 56.8, 56.4, 56.0, 54.4, 53.5, 43.6, 41.7, 37.9, 37.4, 35.8, 33.0, 32.8, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.95$, (m/z) $[\text{M}+\text{H}]^+$ 935 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{50}\text{H}_{63}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 935.4490$, found 935.4499; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3302, 3074, 2958, 2873, 1651, 1614, 1482, 1434, 1371; $[\alpha]_{\text{D}}^{20} = -24$ (c 0.28, MeOH).

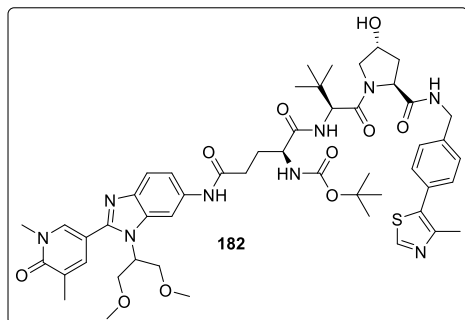
N*2-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*6-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)cubane-1,4-dicarboxamide **181*



181 was synthesised following the **general procedure E** with the linker 4-methoxycarbonylcubane-carboxylic acid (62.0 mg, 0.300 mmol). **181** (82 mg, 0.087 mmol, 29% yield) was isolated as a white solid.

m.p. 183-189 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.66 (s, 1 H), 8.98 (s, 1 H), 8.55 (t, $J=6.2$ Hz, 1 H), 8.21 (s, 1 H), 8.03 (s, 1 H), 7.70-7.65 (m, 2 H), 7.54 (d, $J=8.8$ Hz, 1 H), 7.45-7.39 (m, 5 H), 5.15 (s (br), 1 H), 4.84-4.80 (m, 1 H), 4.63 (d, $J=9.5$ Hz, 1 H), 4.46-4.44 (m, 2 H), 4.38-4.34 (m, 1 H), 4.25-4.21 (m, 4 H), 4.15-4.12 (m, 3 H), 4.00-3.96 (m, 2 H), 3.78-3.75 (m, 2 H), 3.70-3.62 (m, 2 H), 3.53 (s, 3 H), 3.17 (s, 6 H), 2.45 (s, 3 H), 2.09 (s, 3 H), 2.08-2.04 (m, 1 H), 1.94-1.89 (m, 1 H), 0.97 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.8, 170.4, 169.7, 169.5, 161.7, 151.6, 151.4, 147.7, 139.4, 139.3, 138.9, 137.3, 133.9, 133.3, 131.1, 129.6, 128.6, 127.4, 127.4, 118.7, 115.4, 107.9, 103.5, 69.4, 68.8, 58.7, 58.4, 57.6, 56.8, 56.8, 56.4, 56.1, 46.4, 46.2, 41.6, 37.9, 37.5, 35.5, 26.4, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.92$, (m/z) $[\text{M}+\text{H}]^+$ 943 (99% purity); **HRMS** (ESI): calculated for $\text{C}_{51}\text{H}_{59}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+2\text{H}]^{2+} = 472.2127$, found 472.2127; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3296, 3069, 2963, 2868, 1646, 1609, 1519, 1487; $[\alpha]_{\text{D}}^{20} = -30$ (c 0.25, MeOH).

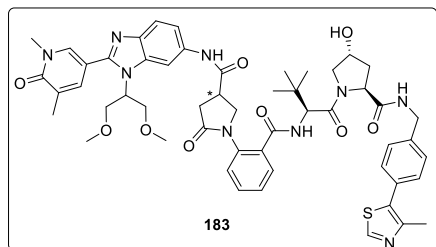
tert*-Butyl ((*R*)-5-((1-(1,3-dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)amino)-1-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-1,5-dioxopentan-2-yl)carbamate **182*



182 was synthesised following the **general procedure E** with the linker (*R*)-2-((*tert*-butoxycarbonyl)amino)-5-methoxy-5-oxopentanoic acid (78.0 mg, 0.300 mmol). **182** (102 mg, 0.100 mmol, 34% yield) was isolated as a white solid.

m.p. 158-164 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.00 (s, 1 H), 8.97 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 8.15 (s, 1 H), 8.03 (s, 1 H), 7.91 (d, $J=9.2$ Hz, 1 H), 7.66 (s, 1 H), 7.55 (d, $J=8.8$ Hz, 1 H), 7.42-7.36 (m, 5 H), 6.94-6.93 (m, 1 H), 5.11 (s (br), 1 H), 4.84-4.80 (m, 1 H), 4.55 (d, $J=9.5$ Hz, 1 H), 4.45-4.41 (m, 2 H), 4.37-4.34 (m, 1 H), 4.24-4.20 (m, 1 H), 4.10-4.05 (m, 1 H), 3.99-3.95 (m, 2 H), 3.77-3.75 (m, 2 H), 3.69-3.62 (m, 2 H), 3.53 (s, 3 H), 3.18 (s, 6 H), 2.44 (s, 3 H), 2.42-2.40 (m, 1 H), 2.24-2.19 (m, 1 H), 2.09 (s, 3 H), 2.06-2.02 (m, 1 H), 1.95-1.83 (m, 3 H), 1.40 (s, 9 H), 0.94 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.5, 171.9, 170.6, 169.6, 161.7, 155.3, 151.6, 151.4, 147.7, 139.4, 139.3, 139.0, 137.3, 133.4, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 119.0, 114.9, 107.9, 103.0, 78.2, 69.5, 68.8, 58.7, 58.4, 56.8, 56.4, 56.3, 55.2, 41.6, 37.9, 37.4, 35.3, 32.2, 28.4, 28.2, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 1.00$, (m/z) $[\text{M}+\text{H}]^+$ 998 (99% purity); **HRMS** (ESI): calculated for $\text{C}_{51}\text{H}_{68}\text{N}_9\text{O}_{10}\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 998.4810$, found 998.4799; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3296, 3069, 2974, 2868, 1656, 1614, 1519, 1487, 1434; $[\alpha]_{\text{D}}^{20} = -31$ (c 0.26, MeOH).

(2*S*,4*R*)-1-((2*S*)-2-(2-(4-((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)carbamoyl)-2-oxopyrrolidin-1-yl)benzamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide 183 Isomer 1 and Isomer 2



183 Isomer 1 and Isomer 2 were synthesised following the **general procedure E** with the linker 2-(4-(methoxycarbonyl)-2-oxopyrrolidin-1-yl)benzoic acid (79.0 mg, 0.300 mmol). **183 Isomer 1** (13 mg, 0.013 mmol, 4% yield) and **183**

Isomer 2 (23 mg, 0.023 mmol, 7% yield) were both isolated as white solids.

183 Isomer 1

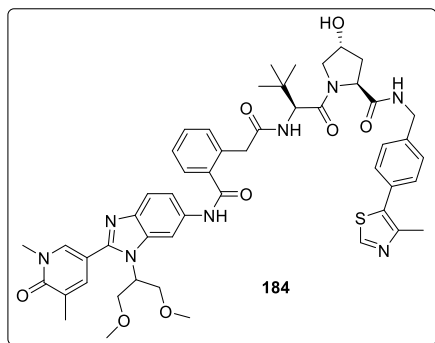
m.p. 179-183 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.27 (s, 1 H), 8.98 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 8.24-8.22 (m, 2 H), 8.06 (s, 1 H), 7.68-7.65 (m, 2 H), 7.58-7.55 (m, 2 H), 7.51-7.50 (m, 1 H), 7.44-7.38 (m, 6 H), 5.12 (s (br), 1 H), 4.88-4.83 (m, 1 H), 4.58 (d, $J=9.2$ Hz, 1 H), 4.45-4.43 (m, 2 H), 4.37-4.35 (m, 1 H), 4.25-4.19 (m, 1 H), 4.03-3.93 (m, 4 H), 3.80-3.77 (m, 2 H), 3.70-3.64 (m, 2 H), 3.54 (s, 3 H), 3.52-3.49 (m, 1 H), 3.19 (s, 6 H), 2.62-2.55 (m, 1 H), 2.47-2.45 (m, 1 H), 2.44 (s, 3 H), 2.10 (s, 3 H), 2.06-2.03 (m, 1 H), 1.94-1.89 (m, 1 H), 0.95 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.9, 171.8, 171.7, 169.3, 165.6, 161.6, 151.5, 151.4, 147.7, 139.4, 139.0, 139.0, 137.2, 136.5, 134.3, 133.3, 131.1, 130.8, 129.6, 128.7, 128.6, 127.5, 127.5, 127.4, 126.6, 126.4, 118.5, 115.8, 107.6, 103.9, 69.5, 68.8, 58.7, 58.4, 56.7, 56.4, 52.6, 41.6, 37.9, 37.9, 37.5, 36.5, 35.4, 35.3, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.89$, (m/z) $[\text{M}+\text{H}]^-$ 998 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{53}\text{H}_{62}\text{N}_9\text{O}_9\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 1000.4391$, found 1000.4393; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3275, 3074, 2952, 2873, 1651, 1609, 1535, 1487; $[\alpha]_{\text{D}}^{20} = -66$ (c 0.03, MeOH).

183 Isomer 2

m.p. 171-177 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.33 (s, 1 H), 8.98 (s, 1 H), 8.58 (t, $J=6.1$ Hz, 1 H), 8.27-8.22 (m, 2 H), 8.09 (s, 1 H), 7.68-7.66 (m, 2 H), 7.61-7.55 (m,

3 H), 7.45-7.36 (m, 6 H), 5.12 (s (br), 1 H), 4.90-4.86 (m, 1 H), 4.59 (d, $J=9.5$ Hz, 1 H), 4.46-4.42 (m, 2 H), 4.37-4.33 (m, 1 H), 4.23-4.20 (m, 1 H), 4.10-4.00 (m, 3 H), 3.87-3.84 (m, 1 H), 3.81-3.77 (m, 2 H), 3.70-3.63 (m, 3 H) 3.55 (s, 3 H), 3.20 (s, 6 H), 2.64-2.59 (m, 1 H), 2.44 (s, 3 H), 2.43-2.40 (m, 1 H), 2.10 (s, 3 H), 2.06-2.03 (m, 1 H), 1.93-1.89 (m, 1 H), 0.95 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 172.4, 171.8, 171.4, 169.3, 165.7, 161.7, 151.4, 151.3, 147.7, 139.4, 139.4, 139.4, 137.0, 136.4, 134.3, 133.0, 131.1, 130.8, 129.6, 128.8, 128.6, 128.0, 128.0, 127.6, 127.4, 126.7, 118.1, 116.0, 109.3, 104.0, 69.4, 68.8, 58.7, 58.4, 57.0, 56.6, 53.6, 41.6, 37.9, 37.9, 37.5, 36.6, 35.4, 33.9, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.95$, (m/z) $[\text{M}+\text{H}]^+$ 998 (98% purity); **HRMS** (ESI): calculated for $\text{C}_{53}\text{H}_{62}\text{N}_9\text{O}_9\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 1000.4391$, found 1000.4401; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3296, 3064, 2963, 2889, 1651, 1539, 1482, 1450; $[\alpha]_{\text{D}}^{20} = -30$ (c 0.08, MeOH).

(2S,4R)-1-((S)-2-(2-(2-((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)carbamoyl)phenyl)acetamido)-3,3-dimethylbutanoyl)-4-hydroxy-N-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide 184

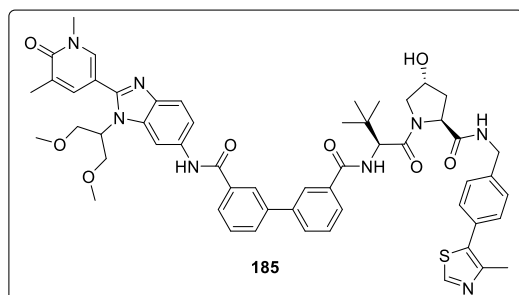


184 was synthesised following the **general procedure E** with the linker 2-(2-(methoxycarbonyl)phenyl)acetic acid (58.0 mg, 0.300 mmol). **184** (35 mg, 0.038 mmol, 13% yield) was isolated as a white solid.

m.p. 161-167 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 11.02 (s, 1 H), 8.97 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 8.49-8.46 (m, 2 H), 8.09 (s, 1 H), 7.69-7.60 (m, 3 H), 7.52-7.45 (m, 2 H), 7.42-7.38 (m, 6 H), 5.12 (s (br), 1 H), 4.90-4.85 (m, 1 H), 4.53 (d, $J=9.2$ Hz, 1 H), 4.46-4.93 (m, 2 H), 4.36-4.33 (m, 1 H), 4.26-4.21 (m, 1 H), 4.04-4.00 (m, 2 H), 3.96-3.94 (m, 1 H), 3.80-3.76 (m, 3 H), 3.68-3.62 (m, 2 H), 3.55 (s, 3 H), 3.20 (s, 6 H), 2.44 (s, 3 H), 2.10 (s, 3 H), 2.05-2.00 (m, 1 H), 1.93-1.89 (m, 1 H), 0.92 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.8, 170.5, 169.2, 167.1, 161.7, 151.4, 151.4,

147.7, 139.4, 139.4, 137.0, 136.9, 136.7, 134.6, 134.1, 133.0, 131.1, 130.4, 130.1, 129.6, 128.6, 128.5, 127.6, 127.4, 126.7, 118.4, 115.9, 106.9, 103.6, 69.4, 68.8, 58.7, 58.4, 57.1, 56.7, 56.3, 41.6, 39.3, 37.9, 37.5, 35.5, 26.2, 16.9, 15.9; **LCMS** (Method B): $t_R = 0.97$, (m/z) $[M+H]^+$ 931 (100% purity); **HRMS** (ESI): calculated for $C_{50}H_{59}N_8O_8S$ (m/z) $[M+H]^+ = 931.4177$, found 931.4174; **IR** (ν_{max}/cm^{-1}) 3270, 3069, 2958, 2878, 1651, 1603, 1535, 1487; $[\alpha]_D^{20} = -33$ (c 0.05, MeOH).

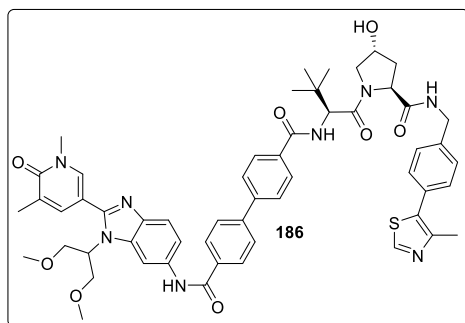
N*3-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*3'-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-[1,1'-biphenyl]-3,3'-dicarboxamide **185*



185 was synthesised following the **general procedure E** with the linker 3'-(methoxycarbonyl)-[1,1'-biphenyl]-3-carboxylic acid (77.0 mg, 0.300 mmol). **185** (90 mg, 0.091 mmol, 30% yield) was isolated as a white solid.

m.p. 191-197 °C; δ_H (400 MHz, d_6 -DMSO) δ ppm 10.44 (s, 1 H), 8.99 (s, 1 H), 8.56 (t, $J=6.1$ Hz, 1 H), 8.36-8.23 (m, 3 H), 8.24 (s, 1 H), 8.06-7.90 (m, 5 H), 7.70-7.66 (m, 2 H), 7.63-7.57 (m, 3 H), 7.43-7.37 (m, 4 H), 5.15 (d, $J=3.7$ Hz, 1 H), 4.88-4.83 (m, 2 H), 4.49-4.46 (m, 1 H), 4.44-4.38 (m, 2 H), 4.29-4.23 (m, 1 H), 4.05-4.01 (m, 2 H), 3.81-3.75 (m, 4 H), 3.56 (s, 3 H), 3.19 (s, 6 H), 2.44 (s, 3 H), 2.10 (s, 3 H), 2.08-2.04 (m, 1 H), 1.96-1.92 (m, 1 H), 1.07 (s, 9 H); δ_C (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.4, 166.5, 165.2, 161.7, 151.8, 151.3, 147.7, 139.8, 139.7, 139.4, 139.4, 139.0, 137.3, 135.8, 134.9, 133.8, 133.4, 131.1, 130.0, 129.1, 128.9, 128.6, 127.5, 127.4, 127.4, 127.2, 127.0, 127.0, 126.1, 126.0, 118.8, 116.1, 107.9, 104.4, 69.5, 68.9, 58.8, 58.4, 57.4, 56.8, 56.4, 41.6, 37.9, 37.8, 35.6, 26.5, 16.9, 15.9; **LCMS** (Method B): $t_R = 1.07$, (m/z) $[M+H]^+$ 993 (100% purity); **HRMS** (ESI): calculated for $C_{55}H_{61}N_8O_8S$ (m/z) $[M+H]^+ = 993.4333$, found 993.4340; **IR** (ν_{max}/cm^{-1}) 3291, 3074, 2926, 2873, 1656, 1603, 1519, 1482; $[\alpha]_D^{20} = -43$ (c 0.18, MeOH).

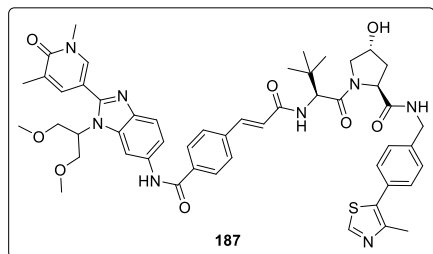
N*4-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*4'-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-[1,1'-biphenyl]-4,4'-dicarboxamide **186*



186 was synthesised following the **general procedure E** with the linker 4'-(methoxycarbonyl)-[1,1'-biphenyl]-4-carboxylic acid (77.0 mg, 0.300 mmol). **186** (123 mg, 0.120 mmol, 41% yield) was isolated as a white solid.

m.p. 192-200 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.36 (s, 1 H), 8.99 (s, 1 H), 8.57 (t, $J=5.9$ Hz, 1 H), 8.84 (s, 1 H), 8.14 (d, $J=8.4$ Hz, 2 H), 8.09 (d, $J=8.8$ Hz, 1 H), 8.06 (s, 1 H), 8.02 (d, $J=8.4$ Hz, 2 H), 7.93 (d, $J=8.4$ Hz, 2 H), 7.88 (d, $J=8.4$ Hz, 2 H), 7.69 (s, 1 H), 7.63-7.60 (m, 2 H), 7.44-7.40 (m, 4 H), 5.16 (d, $J=3.7$ Hz, 1 H), 4.88-4.84 (m, 1 H), 4.82 (d, $J=9.2$ Hz, 1 H), 4.49 (t, $J=7.9$ Hz, 1 H), 4.45-4.39 (m, 2 H), 4.28-4.24 (m, 1 H), 4.05-4.02 (m, 2 H), 3.82-3.79 (m, 2 H), 3.77-3.75 (m, 2 H), 3.55 (s, 3 H), 3.20 (s, 6 H), 2.46 (s, 3 H), 2.10 (s, 3 H), 2.09-2.05 (m, 1 H), 1.96-1.92 (m, 1 H), 1.08 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.4, 166.0, 164.8, 161.7, 151.8, 151.4, 147.7, 141.9, 141.7, 139.7, 139.4, 139.0, 137.3, 134.3, 133.9, 133.5, 133.4, 131.1, 129.6, 128.6, 128.4, 128.3, 127.5, 127.4, 126.7, 126.6, 118.0, 116.0, 107.9, 104.2, 69.5, 68.9, 58.8, 58.4, 57.3, 56.8, 56.4, 41.7, 37.9, 37.5, 35.6, 26.5, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 1.05$, (m/z) $[\text{M}+\text{H}]^+$ 993 (99% purity); **HRMS** (ESI): calculated for $\text{C}_{55}\text{H}_{61}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 993.4333$, found 993.4340; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3275, 3064, 2952, 2873, 1656, 1609, 1535, 1487; $[\alpha]_{\text{D}}^{20} = -25$ (c 0.23, MeOH).

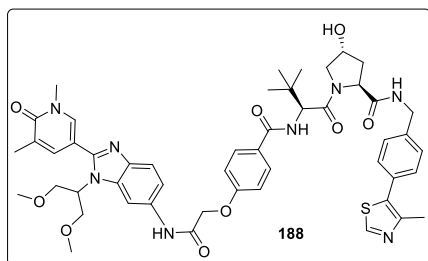
(2*S*,4*R*)-1-((*S*)-2-((*E*)-3-(4((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)amino)-3-oxoprop-1-en-1-yl)benzamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide **187**



187 was synthesised following the **general procedure E** with the linker (*E*)-3-(4-(methoxycarbonyl)phenyl)acrylic acid (62.0 mg, 0.300 mmol). **187** (91 mg, 0.096 mmol, 32% yield) was isolated as a yellow solid.

m.p. 199-203 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.33 (s, 1 H), 8.98 (s, 1 H), 8.58 (t, $J=6.1$ Hz, 1 H), 8.32 (s, 1 H), 8.22 (d, $J=9.5$ Hz, 1 H), 8.07-8.04 (m, 3 H), 7.73-7.69 (m, 3 H), 7.62 (s, 2 H), 7.51 (d, $J=15.8$ Hz, 1 H), 7.45-7.39 (m, 4 H), 7.16 (d, $J=15.8$ Hz, 1 H), 5.14 (s, 1 H), 4.89-4.84 (m, 1 H), 4.70 (d, $J=9.2$ Hz, 1 H), 4.48-4.44 (m, 2 H), 4.39-4.37 (m, 1 H), 4.25-4.22 (m, 1 H), 4.04-4.01 (m, 2 H), 3.81-3.79 (m, 2 H), 3.74-3.70 (m, 2 H), 3.55 (s, 3 H), 3.20 (s, 6 H), 2.46 (s, 3 H), 2.10 (s, 3 H), 2.09-2.05 (m, 1 H), 1.96-1.91 (m, 1 H), 1.00 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.4, 164.7, 164.4, 161.7, 151.7, 151.4, 147.7, 139.4, 139.1, 138.0, 138.0, 137.2, 137.2, 135.4, 134.0, 133.2, 131.1, 129.6, 128.6, 128.2, 127.5, 127.4, 127.4, 124.1, 118.6, 116.1, 107.6, 104.3, 69.4, 68.9, 58.7, 58.4, 56.9, 56.7, 56.4, 41.6, 38.0, 37.5, 35.4, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.99$, (m/z) $[\text{M}+\text{H}]^+$ 943 (98% purity); **HRMS** (ESI): calculated for $\text{C}_{51}\text{H}_{59}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 943.4177$, found 943.4186; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3286, 3058, 2963, 2873, 1651, 1603, 1529, 1487, 1439; $[\alpha]_{\text{D}}^{20} = -34$ (c 0.17, MeOH).

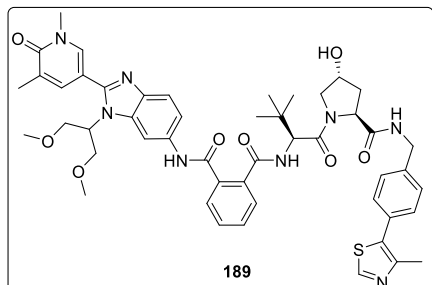
(2*S*,4*R*)-1-((*S*)-2-(4-(2-((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)carbamoyl)phenoxy)acetamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide 188



188 was synthesised following the **general procedure E** with the linker 4-(2-methoxy-2-oxoethoxy)benzoic acid (63.0 mg, 0.300 mmol). **188** (64 mg, 0.067 mmol, 21% yield) was isolated as a white solid.

m.p. 162-168 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.18 (s, 1 H), 8.98 (s, 1 H), 8.56 (t, $J=6.1$ Hz, 1 H), 8.20 (s, 1 H), 8.03 (s, 1 H), 7.91 (d, $J=8.8$ Hz, 2 H), 7.84-7.81 (m, 1 H), 7.67 (s, 1 H), 7.58 (d, $J=8.4$ Hz, 1 H), 7.43-7.40 (m, 5 H), 7.10 (d, $J=8.8$ Hz, 2 H), 5.14 (s, 1 H), 4.85-4.81 (m, 3 H), 4.77 (d, $J=9.2$ Hz, 1 H), 4.48-4.38 (m, 3 H), 4.27-4.23 (m, 1 H), 3.99-3.96 (m, 2 H), 3.78-3.73 (m, 4 H), 3.54 (s, 3 H), 3.17 (s, 6 H), 2.45 (s, 3 H), 2.09 (s, 3 H), 2.08-2.04 (m, 1 H), 1.95-1.91 (m, 1 H), 1.03 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.5, 165.9, 165.7, 161.7, 160.2, 151.8, 151.4, 147.7, 139.6, 139.6, 139.4, 137.2, 133.4, 133.1, 131.1, 129.6, 129.5, 128.6, 127.5, 127.4, 126.9, 119.0, 115.3, 114.2, 107.9, 103.6, 69.5, 68.9, 67.2, 58.8, 58.4, 57.2, 56.8, 56.3, 41.6, 37.9, 37.5, 35.5, 26.5, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.96$, (m/z) $[\text{M}+\text{H}]^+$ 947 (94% purity); **HRMS** (ESI): calculated for $\text{C}_{50}\text{H}_{59}\text{N}_8\text{O}_9\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 947.4126$, found 947.4133; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3275, 3074, 2942, 2868, 1656, 1603, 1535, 1492, 1434; $[\alpha]_{\text{D}}^{20} = -30$ (c 0.18, MeOH).

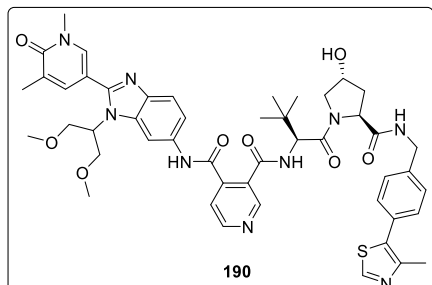
N*1-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*2-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)phthalimide **189*



189 was synthesised following the **general procedure E** with the linker 2-(methoxycarbonyl)benzoic acid (54.0 mg, 0.300 mmol). **189** (58 mg, 0.063 mmol, 21% yield) was isolated as a white solid.

m.p. 165-171 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.49 (s, 1 H), 8.99 (s, 1 H), 8.53 (t, $J=5.9$ Hz, 1 H), 8.32 (s, 1 H), 8.06-8.03 (m, 2 H), 7.68-7.55 (m, 6 H), 7.45 (d, $J=8.4$ Hz, 1 H), 7.41-7.38 (m, 4 H), 5.12 (s, 1 H), 4.85-4.81 (m, 1 H), 4.67 (d, $J=9.2$ Hz, 1 H), 4.47 (t, $J=7.9$ Hz, 1 H), 4.40-4.36 (m, 2 H), 4.28-4.23 (m, 1 H), 4.00-3.96 (m, 2 H), 3.78-3.74 (m, 2 H), 3.71-3.67 (m, 2 H), 3.56 (s, 3 H), 3.17 (s, 6 H), 2.44 (s, 3 H), 2.12 (s, 3 H), 2.07-2.04 (m, 1 H), 1.94-1.89 (m, 1 H), 0.95 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.3, 167.2, 167.2, 161.7, 151.7, 151.4, 147.7, 139.5, 139.4, 139.0, 137.3, 136.7, 134.7, 134.2, 133.3, 131.1, 129.9, 129.6, 129.4, 128.6, 128.5, 127.6, 127.6, 127.4, 118.8, 115.2, 107.9, 103.2, 69.5, 68.8, 58.8, 58.4, 57.4, 56.8, 56.2, 41.6, 37.8, 37.4, 35.6, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.92$, (m/z) $[\text{M}+\text{H}]^+$ 917 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{49}\text{H}_{57}\text{N}_8\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 917.4020$, found 917.4026; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3275, 3069, 2947, 2873, 1651, 1598, 1529, 1482; $[\alpha]_{\text{D}}^{20} = -29$ (c 0.19, MeOH).

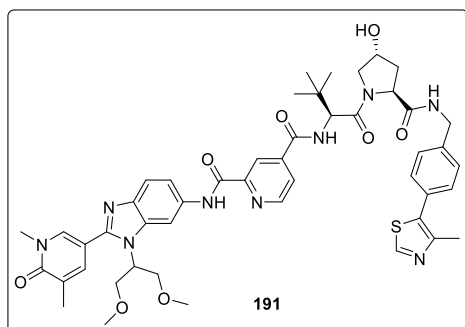
N*4-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*3-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)pyridine-3,4-dicarboxamide **190*



190 was synthesised following the **general procedure E** with the linker 4-(methoxycarbonyl)nicotinic acid (54.0 mg, 0.300 mmol). **190** (70 mg, 0.076 mmol, 25% yield) was isolated as a white solid.

m.p. 176-182 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.65 (s, 1 H), 8.99 (s, 1 H), 8.80-8.78 (m, 2 H), 8.54 (t, $J=5.9$ Hz, 1 H), 8.46 (d, $J=9.2$ Hz, 1 H), 8.29 (s, 1 H), 8.05 (s, 1 H), 7.68-7.64 (m, 2 H), 7.58 (d, $J=8.4$ Hz, 1 H), 7.44-7.38 (m, 5 H), 5.12 (s, 1 H), 4.86-4.82 (m, 1 H), 4.70 (d, $J=9.2$ Hz, 1 H), 4.48 (t, $J=8.1$ Hz, 1 H), 4.41-4.36 (m, 2 H), 4.27-4.23 (m, 1 H), 3.99-3.96 (m, 2 H), 3.78-3.75 (m, 2 H), 3.71-3.65 (m, 2 H), 3.54 (s, 3 H), 3.18 (s, 6 H), 2.44 (s, 3 H), 2.09 (s, 3 H), 2.07-2.04 (m, 1 H), 1.94-1.90 (m, 1 H), 0.97 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.0, 165.4, 165.0, 161.7, 151.9, 151.4, 151.2, 149.1, 147.7, 143.2, 139.7, 139.4, 139.0, 137.2, 133.6, 133.3, 131.1, 129.6, 129.3, 128.6, 127.5, 127.4, 121.6, 119.0, 115.2, 107.9, 103.5, 69.5, 68.8, 58.8, 58.4, 57.4, 56.8, 56.2, 41.6, 37.9, 37.5, 35.6, 26.4, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.86$, (m/z) $[\text{M}+\text{H}]^+$ 918 (98% purity); **HRMS** (ESI): calculated for $\text{C}_{48}\text{H}_{56}\text{N}_9\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 918.3973$, found 918.3981; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3265, 3048, 2947, 2863, 1651, 1603, 1545, 1476; $[\alpha]_{\text{D}}^{20} = -33$ (c 0.35, MeOH).

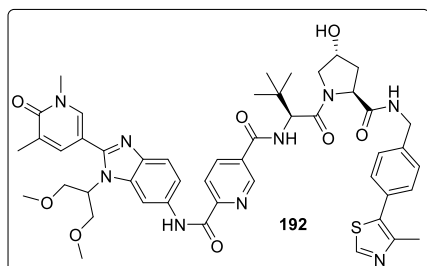
N*2-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*4-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)pyridine-2,4-dicarboxamide **191*



191 was synthesised following the **general procedure E** with the linker 2-(methoxycarbonyl)isonicotinic acid (54.0 mg, 0.300 mmol). **191** (34 mg, 0.037 mmol, 12% yield) was isolated as a yellow solid.

m.p. 174-180 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.78 (s, 1 H), 8.98 (s, 1 H), 8.88 (d, $J=5.5$ Hz, 1 H), 8.84 (d, $J=9.2$ Hz, 1 H), 8.59 (t, $J=6.1$ Hz, 1 H), 8.56 (s, 1 H), 8.41 (s, 1 H), 8.07 (s, 1 H), 8.04 (d, $J=5.0$ Hz, 1 H), 7.86 (d, $J=8.8$ Hz, 1 H), 7.70 (s, 1 H), 7.62 (d, $J=8.4$ Hz, 1 H), 7.44-7.40 (m, 4 H), 5.17 (s, 1 H), 4.88-4.83 (m, 2 H), 4.50 (t, $J=8.1$ Hz, 1 H), 4.47-4.43 (m, 1 H), 4.41-4.39 (m, 1 H), 4.27-4.24 (m, 1 H), 4.07-4.04 (m, 2 H), 3.82-3.79 (m, 2 H), 3.76-3.73 (m, 2 H), 3.54 (s, 3 H), 3.20 (s, 6 H), 2.45 (s, 3 H), 2.10 (s, 3 H), 2.08-2.05 (m, 1 H), 1.97-1.94 (m, 1 H), 1.07 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.0, 165.0, 161.8, 161.7, 152.0, 151.4, 150.5, 148.9, 147.7, 143.2, 139.9, 139.4, 139.0, 137.2, 133.4, 133.1, 131.1, 129.6, 128.6, 127.5, 127.4, 124.8, 120.5, 118.8, 115.7, 107.9, 104.2, 69.4, 68.9, 58.8, 58.4, 57.7, 56.9, 56.5, 41.7, 37.9, 37.5, 35.6, 26.5, 17.0, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.98$, (m/z) $[\text{M}+\text{H}]^+$ 918 (100% purity); **HRMS** (ESD): calculated for $\text{C}_{48}\text{H}_{56}\text{N}_9\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 918.3973$, found 918.3973; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3302, 3074, 2931, 2873, 1651, 1603, 1524, 1482; $[\alpha]_{\text{D}}^{20} = -35$ (c 0.10, MeOH).

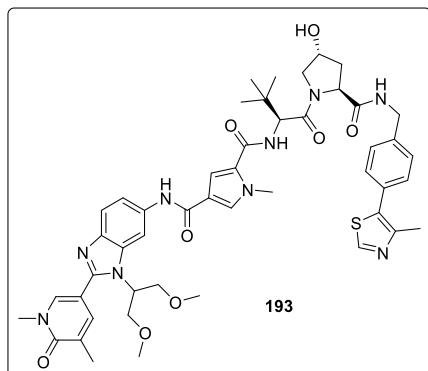
N*2-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*5-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)pyridine-2,5-dicarboxamide **192*



192 was synthesised following the **general procedure E** with the linker 6-(methoxycarbonyl)nicotinic acid (54.0 mg, 0.300 mmol). **192** (51 mg, 0.056 mmol, 19% yield) was isolated as a yellow solid.

m.p. 169-175 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.81 (s, 1 H), 9.09 (s, 1 H), 8.98 (s, 1 H), 8.70 (d, $J=8.8$ Hz, 1 H), 8.58 (t, $J=6.1$ Hz, 1 H), 8.50 (d, $J=8.3$ Hz, 1 H), 8.40 (s, 1 H), 8.26 (d, $J=8.1$ Hz, 1 H), 8.06 (s, 1 H), 7.85 (d, $J=8.4$ Hz, 1 H), 7.69 (s, 1 H), 7.61 (d, $J=8.8$ Hz, 1 H), 7.44-7.39 (m, 4 H), 5.17 (s, 1 H), 4.88-4.81 (m, 2 H), 4.50-4.40 (m, 3 H), 4.27-4.23 (m, 1 H), 4.07-4.04 (m, 2 H), 3.82-3.76 (m, 4 H), 3.55 (s, 3 H), 3.19 (s, 6 H), 2.45 (s, 3 H), 2.10 (s, 3 H), 2.08-2.06 (m, 1 H), 1.97-1.93 (m, 1 H), 1.08 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.1, 164.9, 161.7, 161.6, 152.0, 151.7, 151.4, 147.8, 147.7, 139.9, 139.4, 139.0, 137.4, 137.2, 133.4, 133.0, 132.1, 131.1, 129.6, 128.6, 127.5, 127.4, 121.7, 118.8, 115.6, 107.9, 104.2, 69.4, 68.9, 58.8, 58.4, 57.7, 56.9, 56.4, 41.7, 37.9, 37.5, 35.5, 26.5, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.97$, (m/z) $[\text{M}+\text{H}]^+$ 918 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{48}\text{H}_{56}\text{N}_9\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 918.3973$, found 918.3967; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3302, 3079, 2942, 2873, 1651, 1609, 1524, 1482; **$[\alpha]_{\text{D}}^{20} = -33$** (c 0.17, MeOH).

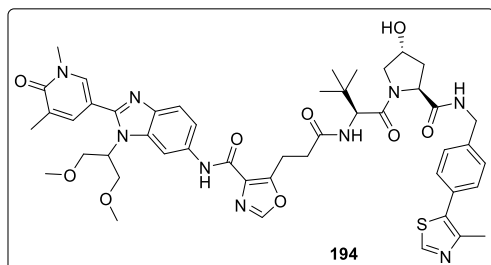
N4-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N2-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-1-methyl-1H-pyrrole-2,4-dicarboxamide **193**



193 was synthesised following the **general procedure E** with the linker 4-(methoxycarbonyl)-1-methyl-1H-pyrrole-2-carboxylic acid (55.0 mg, 0.300 mmol). **193** (108 mg, 0.120 mmol, 39% yield) was isolated as a white solid.

m.p. 195-204 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.01 (s, 1 H), 8.98 (s, 1 H), 8.55 (t, $J=6.1$ Hz, 1 H), 8.26 (s, 1 H), 8.09 (s, 1 H), 7.80 (s, 1 H), 7.69 (s, 1 H), 7.64-7.59 (m, 2 H), 7.53 (s, 1 H), 7.43-7.40 (m, 4 H), 7.33 (d, $J=9.2$ Hz, 1 H), 5.14 (s, 1 H), 4.90-4.86 (m, 1 H), 4.78 (d, $J=9.2$ Hz, 1 H), 4.47-4.37 (m, 3 H), 4.27-4.22 (m, 1 H), 4.03-4.00 (m, 2 H), 3.93 (s, 3 H), 3.81-3.77 (m, 2 H), 3.73-3.70 (m, 2 H), 3.55 (s, 3 H), 3.19 (s, 6 H), 2.45 (s, 3 H), 2.10 (s, 3 H), 2.08-2.05 (m, 1 H), 1.96-1.91 (m, 1 H), 1.02 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.7, 162.6, 161.7, 159.4, 151.4, 151.2, 147.7, 139.4, 139.4, 139.4, 137.0, 134.5, 133.0, 131.1, 130.3, 129.6, 128.6, 127.6, 127.4, 126.1, 118.1, 117.3, 116.3, 113.1, 106.8, 103.8, 69.3, 68.8, 58.7, 58.4, 57.0, 56.4, 56.2, 41.6, 37.9, 37.5, 36.7, 35.6, 26.4, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.96$, (m/z) $[\text{M}+\text{H}]^+$ 920 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{48}\text{H}_{58}\text{N}_9\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 920.4129$, found 920.4135; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3291, 3079, 2947, 2878, 1656, 1614, 1556, 1482; $[\alpha]_{\text{D}}^{20} = -35$ (c 0.23, MeOH).

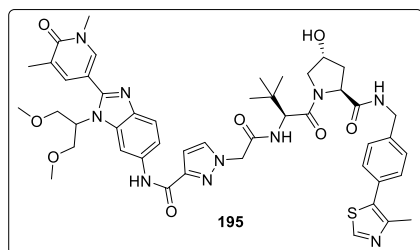
N*-1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-5-(3-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-3-oxopropyl)oxazole-4-carboxamide **194*



194 was synthesised following the **general procedure E** with the linker 3-(4-(methoxycarbonyl)oxazol-5-yl)propanoic acid (60.0 mg, 0.300 mmol). **194** (35 mg, 0.037 mmol, 12% yield) was isolated as a white solid.

m.p. 156-160 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.02 (s, 1 H), 8.97 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 8.45 (s, 1 H), 8.27 (s, 1 H), 8.06-8.03 (m, 2 H), 7.74-7.67 (m, 2 H), 7.56 (d, $J=8.8$ Hz, 1 H), 7.42-7.37 (m, 4 H), 5.14 (s (br), 1 H), 4.86-4.81 (m, 1 H), 4.56 (d, $J=9.5$ Hz, 1 H), 4.46-4.41 (m, 2 H), 4.37-4.35 (m, 1 H), 4.24-4.20 (m, 1 H), 4.03-4.00 (m, 2 H), 3.79-3.77 (m, 2 H), 3.70-3.64 (m, 2 H), 3.54 (s, 3 H), 3.32-3.26 (m, 2 H), 3.19 (s, 6 H), 2.74-2.69 (m, 1 H), 2.62-2.57 (m, 1 H), 2.44 (s, 3 H), 2.10 (s, 3 H), 2.06-2.02 (m, 1 H), 1.94-1.88 (m, 1 H), 0.93 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.9, 170.4, 169.5, 161.7, 159.4, 155.9, 151.8, 151.4, 149.7, 147.7, 139.6, 139.4, 139.0, 137.2, 133.3, 133.3, 131.1, 129.6, 128.6, 128.5, 127.4, 127.4, 118.7, 115.8, 107.9, 104.0, 69.4, 68.8, 58.7, 58.4, 56.8, 56.5, 56.3, 41.6, 37.9, 37.5, 35.2, 32.3, 26.3, 21.6, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.92$, (m/z) $[\text{M}+\text{H}]^+ 936$ (100% purity); **HRMS** (ESI): calculated for $\text{C}_{48}\text{H}_{58}\text{N}_9\text{O}_9\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 936.4078$, found 936.4075; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3291, 3058, 2974, 2894, 1656, 1614, 1513, 1487, 1434; $[\alpha]_{\text{D}}^{20} = -30$ (c 0.080, MeOH).

1-(2-((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)amino)-2-oxoethyl)-*N*-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-1*H*-pyrazole-3-carboxamide **195**

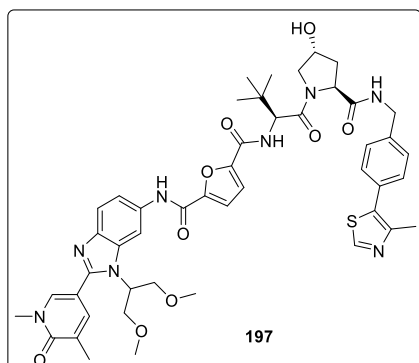


195 was synthesised following the **general procedure E** with the linker 2-(3-(methoxycarbonyl)-1*H*-pyrazol-1-yl)acetic acid (55.0 mg, 0.300 mmol). **195** (66 mg, 0.072 mmol, 24% yield) was isolated as a white

solid.

m.p. 167-173 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.04 (s, 1 H), 8.98 (s, 1 H), 8.58 (t, $J=6.1$ Hz, 1 H), 8.36 (d, $J=9.5$ Hz, 1 H), 8.27 (s, 1 H), 8.05 (s, 1 H), 7.88 (d, $J=2.6$ Hz, 1 H), 7.70-7.68 (m, 2 H), 7.55 (d, $J=8.8$ Hz, 1 H), 7.43-7.39 (m, 4 H), 6.80 (d, $J=2.2$ Hz, 1 H), 5.15-5.11 (m, 3 H), 4.85-4.80 (m, 1 H), 4.57 (d, $J=9.5$ Hz, 1 H), 4.46-4.43 (m, 2 H), 4.37-4.34 (m, 1 H), 4.25-4.21 (m, 1 H), 4.03-4.00 (m, 2 H), 3.79-3.76 (m, 2 H), 3.69-3.66 (m, 1 H), 3.62-3.60 (m, 1 H), 3.54 (s, 3 H), 3.18 (s, 6 H), 2.45 (s, 3 H), 2.09 (s, 3 H), 2.07-2.03 (m, 1 H), 1.92-1.88 (m, 1 H), 0.98 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.8, 169.1, 166.1, 161.7, 159.7, 151.7, 151.4, 147.1, 146.6, 139.5, 139.5, 138.9, 137.2, 134.0, 133.6, 133.3, 131.1, 129.6, 128.6, 127.4, 127.4, 118.6, 115.8, 107.9, 106.4, 104.1, 69.4, 68.8, 58.8, 58.4, 56.8, 56.8, 56.6, 54.0, 41.6, 37.9, 37.5, 35.5, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.89$, (m/z) $[\text{M}+\text{H}]^+$ 921 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{47}\text{H}_{57}\text{N}_{10}\text{O}_8\text{S}$ (m/z) $[\text{M}+\text{H}]^+$ = 921.4082, found 921.4081; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3286, 3074, 2947, 2873, 1651, 1598, 1535, 1487; $[\alpha]_{\text{D}}^{20} = -41$ (c 0.29, MeOH).

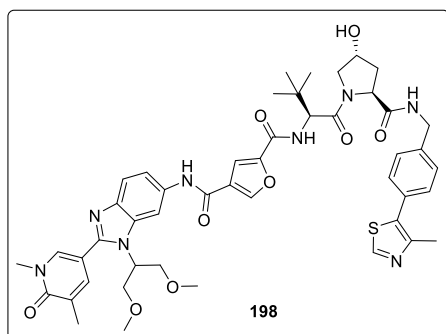
N*2-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*5-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)furan-2,5-dicarboxamide **197*



197 was synthesised following the **general procedure E** with the linker 5-(methoxycarbonyl)furan-2-carboxylic acid (51.0 mg, 0.300 mmol). **197** (75 mg, 0.083 mmol, 27% yield) was isolated as a white solid.

m.p. 172-178 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.54 (s, 1 H), 8.97 (s, 1 H), 8.58 (t, $J=5.9$ Hz, 1 H), 8.24 (s, 1 H), 8.13 (d, $J=9.5$ Hz, 1 H), 8.05 (s, 1 H), 7.68 (s, 1 H), 7.63 (d, $J=8.4$ Hz, 1 H), 7.48-7.37 (m, 7 H), 5.16 (s, 1 H), 4.87-4.81 (m, 2 H), 4.49-4.37 (m, 3 H), 4.28-4.23 (m, 1 H), 4.03-4.00 (m, 2 H), 3.79-3.70 (m, 4 H), 3.55 (s, 3 H), 3.17 (s, 6 H), 2.45 (s, 3 H), 2.10 (s, 3 H), 2.09-2.06 (m, 1 H), 1.96-1.91 (m, 1 H), 1.06 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.7, 168.9, 161.7, 156.9, 155.7, 152.2, 151.4, 148.5, 147.9, 147.7, 140.2, 139.4, 139.0, 137.2, 133.5, 132.3, 131.1, 129.6, 128.6, 127.5, 127.4, 118.9, 116.9, 115.8, 115.6, 107.8, 105.6, 69.6, 68.8, 58.8, 58.4, 56.9, 56.5, 56.5, 41.7, 38.0, 37.5, 35.6, 26.4, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.96$, (m/z) $[\text{M}+\text{H}]^+$ 907 (97% purity); **HRMS** (ESI): calculated for $\text{C}_{47}\text{H}_{55}\text{N}_8\text{O}_9\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 907.3813$, found 907.3807; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3280, 3074, 2958, 2873, 1651, 1598, 1535, 1487; $[\alpha]_{\text{D}}^{20} = -5$ (c 0.21, MeOH).

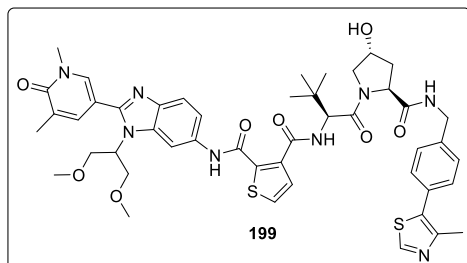
N*4-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*2-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)furan-2,4-dicarboxamide **198*



198 was synthesised following the **general procedure E** with the linker 5-(methoxycarbonyl)furan-3-carboxylic acid (51.0 mg, 0.300 mmol). **198** (96 mg, 0.11 mmol, 35% yield) was isolated as a white solid.

m.p. 179-185 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 10.15 (s, 1 H), 8.98 (s, 1 H), 8.59 (t, $J=6.1$ Hz, 1 H), 8.53 (s, 1 H), 8.22 (s, 1 H), 8.05 (s, 1 H), 7.85 (s, 1 H), 7.74 (d, $J=9.5$ Hz, 1 H), 7.76 (s, 1 H), 7.60 (d, $J=8.8$ Hz, 1 H), 7.51 (d, $J=8.6$ Hz, 1 H), 7.44-7.40 (m, 4 H), 5.15 (s, 1 H), 4.86-4.82 (m, 1 H), 4.74 (d, $J=9.5$ Hz, 1 H), 4.49-4.36 (m, 3 H), 4.27-4.24 (m, 1 H), 4.02-3.99 (m, 2 H), 3.80-3.77 (m, 2 H), 7.73-7.68 (m, 2 H), 3.53 (s, 3 H), 3.18 (s, 6 H), 2.43 (s, 3 H), 2.09 (s, 3 H), 2.08-2.06 (m, 1 H), 1.95-1.91 (m, 1 H), 1.02 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.7, 169.0, 161.7, 159.4, 156.9, 151.9, 151.4, 147.7, 147.6, 147.4, 139.7, 139.4, 139.0, 137.2, 133.4, 133.4, 131.1, 129.7, 128.7, 127.5, 127.4, 124.6, 118.9, 115.8, 112.9, 107.9, 104.2, 69.5, 68.9, 58.8, 58.4, 56.8, 56.5, 56.5, 41.7, 37.9, 37.5, 35.8, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.95$, (m/z) $[\text{M}+\text{H}]^+$ 907 (99% purity); **HRMS** (ESI): calculated for $\text{C}_{47}\text{H}_{55}\text{N}_8\text{O}_9\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 907.3813$, found 907.3820; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3291, 3074, 2958, 2883, 1646, 1598, 1524, 1482, 1434; $[\alpha]_{\text{D}}^{20} = -47$ (c 0.31, MeOH).

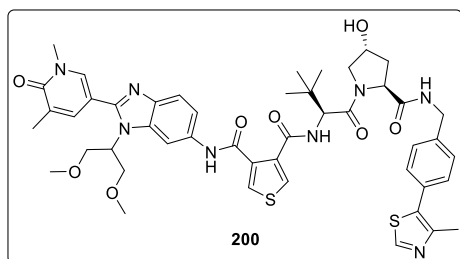
N*2-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*3-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)thiophene-2,3-dicarboxamide **199*



199 was synthesised following the **general procedure E** with the linker 2-(methoxycarbonyl)thiophene-3-carboxylic acid (56.0 mg, 0.300 mmol). **199** (69 mg, 0.075 mmol, 24% yield) was isolated as a yellow solid.

m.p. 169-173 °C; **δ_{H}** (400 MHz, d_6 -DMSO) δ ppm 12.00 (s, 1 H), 8.97 (s, 1 H), 8.94 (d, $J=8.4$ Hz, 1 H), 8.56 (t, $J=6.1$ Hz, 1 H), 8.45 (s, 1 H), 8.04 (s, 1 H), 7.86 (d, $J=5.1$ Hz, 1 H), 7.68 (s, 1 H), 7.61 (d, $J=8.4$ Hz, 1 H), 7.58 (d, $J=5.1$ Hz, 1 H), 7.41-7.38 (m, 4 H), 7.28 (d, $J=8.8$ Hz, 1 H), 5.19 (s, 1 H), 4.87-4.82 (m, 1 H), 4.73 (d, $J=8.4$ Hz, 1 H), 4.49 (t, $J=8.1$ Hz, 1 H), 4.44-4.38 (m, 2 H), 4.26-4.22 (m, 1 H), 4.02-3.99 (m, 2 H), 3.86-3.84 (m, 1 H), 3.79-3.76 (m, 3 H), 3.54 (s, 3 H), 3.17 (s, 6 H), 2.43 (s, 3 H), 2.10 (s, 3 H), 2.09-2.07 (m, 1 H), 1.98-1.94 (m, 1 H), 1.06 (s, 9 H); **δ_{C}** (101 MHz, d_6 -DMSO) δ ppm 171.8, 168.9, 164.2, 161.7, 159.4, 152.0, 151.4, 147.7, 141.2, 139.9, 139.4, 139.0, 137.2, 134.4, 133.5, 133.0, 131.1, 130.7, 129.6, 128.6, 128.2, 127.5, 127.4, 119.4, 115.6, 107.9, 103.6, 69.5, 69.0, 58.8, 58.4, 58.3, 56.9, 56.4, 41.6, 38.0, 37.5, 35.2, 26.4, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 1.02$, (m/z) $[\text{M}+\text{H}]^+$ 923 (97% purity); **HRMS** (ESI): calculated for $\text{C}_{47}\text{H}_{55}\text{N}_8\text{O}_8\text{S}_2$ (m/z) $[\text{M}+\text{H}]^+ = 923.3584$, found 923.3577; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3275, 3064, 2952, 2873, 1656, 1603, 1535, 1482; $[\alpha]_{\text{D}}^{20} = -54$ (c 0.24, MeOH).

N*3-(1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)-*N*4-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)thiophene-3,4-dicarboxamide **200*

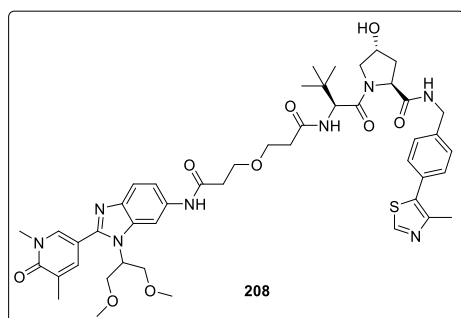


200 was synthesised following the **general procedure E** with the linker 4-(methoxycarbonyl)thiophene-3-carboxylic acid (56.0 mg, 0.300 mmol). **200** (85 mg, 0.092 mmol, 31% yield) was isolated as a white

solid.

m.p. 176-180 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 10.09 (s, 1 H), 9.28 (d, $J=8.4$ Hz, 1 H), 8.94 (s, 1 H), 8.51 (t, $J=6.1$ Hz, 1 H), 8.42 (s, 1 H), 8.28-8.25 (m, 2 H), 8.05 (s, 1 H), 7.68 (s, 1 H), 7.60 (d, $J=8.8$ Hz, 1 H), 7.43 (d, $J=8.6$ Hz, 1 H), 7.40-7.36 (m, 4 H), 5.13 (s, 1 H), 4.86-4.82 (m, 1 H), 4.64 (d, $J=8.8$ Hz, 1 H), 4.46 (t, $J=8.1$ Hz, 1 H), 4.40-4.25 (m, 3 H), 4.02-3.98 (m, 2 H), 3.81-3.73 (m, 4 H), 3.52 (s, 3 H), 3.16 (s, 6 H), 2.42 (s, 3 H), 2.10 (s, 3 H), 2.07-2.03 (m, 1 H), 1.96-1.92 (m, 1 H), 1.04 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.7, 169.3, 163.2, 162.1, 161.7, 152.0, 151.3, 147.7, 139.8, 139.4, 139.0, 137.2, 135.3, 135.3, 133.9, 133.4, 133.3, 131.8, 131.1, 129.7, 128.6, 127.5, 127.4, 119.0, 115.7, 107.8, 104.0, 69.4, 68.9, 58.7, 58.4, 57.8, 56.9, 56.2, 41.6, 37.9, 37.5, 35.0, 26.4, 16.9, 15.8; **LCMS** (Method B): $t_{\text{R}} = 0.95$, (m/z) $[\text{M}+\text{H}]^+$ 923 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{47}\text{H}_{55}\text{N}_8\text{O}_8\text{S}_2$ (m/z) $[\text{M}+\text{H}]^+ = 923.3584$, found 923.3576; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3254, 3074, 2952, 2868, 1646, 1609, 1545, 1489; $[\alpha]_{\text{D}}^{20} = +4$ (c 0.27, MeOH).

(2*S*,4*R*)-1-((*S*)-2-(3-(3-((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1*H*-benzo[*d*]imidazol-6-yl)amino)-3-oxopropoxy)propanamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide **208**



3,3'-oxydipropionic acid **205** (16.2 mg, 0.100 mmol), HATU (41.8 mg, 0.110 mmol), DIPEA (0.070 mL, 0.40 mmol) and solvent were added to a 4 mL vial and sonicated for 30 s.

(2*S*,4*R*)-1-((*S*)-2-amino-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide·HCl **30** (57.5 mg, 0.100 mmol) was added and the reaction mixture stirred at ambient temperature for 30 min.

A selection of solvents and concentrations were screened using the described procedure (*vide supra*) and the conditions visualised in **Table 41**. The reaction mixtures were analysed by LCMS.

Table 41: Conditions screened for the synthesis of carboxylic acid **206** from the amide coupling of amine **30** and diacid linker **205**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	Solvent	Solvent / mL	Concentration / M	30 / %	206 / %	207 / %
1	DMF	1.68	0.06	14	23	49
2	DCM	1.68	0.06	15	40	32
3	DCM	5.00	0.02	15	30	40
4	DCM	2.00	0.05	14	41	30
5	DCM	1.00	0.1	14	38	34
6	DCM	0.25	0.4	15	39	35

A selection of equivalents of 3,3'-oxydipropionic acid **205** were screened using the described procedure (*vide infra*) and the conditions visualised in **Table 42**. The reaction mixtures were analysed by LCMS.

3,3'-oxydipropionic acid **205**, HATU (41.8 mg, 0.110 mmol), DIPEA (0.070 mL, 0.40 mmol) and DCM (2 mL) were added to a 4 mL vial and sonicated for 30 s. (2*S*,4*R*)-1-((*S*)-2-amino-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide·HCl **30** (57.5 mg, 0.100 mmol) was added and the reaction mixture stirred at ambient temperature for 30 min.

Table 42: Conditions screened for the synthesis of carboxylic acid **206** from the amide coupling of amine **30** and diacid linker **205**. The reaction mixtures were analysed by LCMS and the peak areas reported as a percentage.

Entry	205 / mg	205 / mmol	205 / %	206 / %	207 / %
1	16	0.10	14	41	30
2	19	0.12	14	46	32
3	24	0.15	15	30	40

The crude reaction mixtures from **Table 42** were blown down under a stream of N₂ and placed in a vacuum oven at 40 °C for 18 h. HATU (41.8 mg, 0.110 mmol), DIPEA (0.052 mL, 0.30 mmol) and DMF (0.25 mL) were added to the vial and the reaction mixture was sonicated for 30 s. 5-(6-amino-1-(1,3-dimethoxypropan-2-yl)-1*H*-benzo[*d*]imidazol-2-yl)-1,3-dimethylpyridin-2(1*H*)-one **128** (35.6 mg, 0.100 mmol) was added and the reaction mixture stirred at ambient temperature for 2 h. The reaction mixture was diluted with DMSO and purified by MDAP eluting with MeCN in 10 mM ammonium bicarbonate solution (pH 10). UV-active fractions were analysed by LCMS and product-containing fractions were combined. The solvent was removed *in vacuo* to afford the product **208** in the isolated yields reported in **Table 43**.

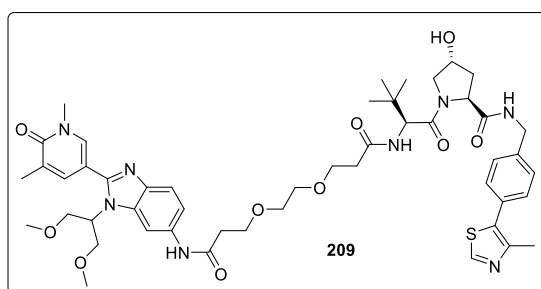
Table 43: Conditions screened for the synthesis of **208** from the amide coupling of carboxylic acid **206** and amine **128**.

Entry	205 / mg	205 / mmol	208 Isolated Yield / %
1	16	0.10	26
2	19	0.12	28
3	24	0.15	20

208 (25 mg, 0.028 mmol, 28% yield) was isolated as a white solid.

m.p. 134-138 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.96 (s, 1 H), 8.97 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 8.22 (s, 1 H), 8.02 (s, 1 H), 7.91 (d, $J=9.5$ Hz, 1 H), 7.66 (s, 1 H), 7.53 (d, $J=8.4$ Hz, 1 H), 7.43-7.38 (m, 4 H), 7.29 (d, $J=8.6$ Hz, 1 H), 5.13 (s (br), 1 H), 4.84-4.78 (m, 1 H), 4.55 (d, $J=9.2$ Hz, 1 H), 4.45-4.42 (m, 2 H), 4.37-4.34 (m, 1 H), 4.24-4.20 (m, 1 H), 3.98-3.94 (m, 2 H), 3.77-3.75 (m, 2 H), 3.72-3.60 (m, 6 H), 3.53 (s, 3 H), 3.17 (s, 6 H), 2.58-2.55 (m, 3 H), 2.44 (s, 3 H), 2.41-2.37 (m, 1 H), 2.09 (s, 3 H), 2.05-2.02 (m, 1 H), 1.93-1.89 (m, 1 H), 0.93 (s, 9 H); δ_{C} (101 MHz, d_6 -DMSO) δ ppm 171.9, 169.9, 169.5, 168.9, 161.7, 151.5, 151.4, 147.7, 139.4, 139.2, 138.4, 137.3, 134.1, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 118.9, 114.7, 107.9, 102.7, 69.5, 68.8, 66.7, 66.3, 58.7, 58.4, 56.8, 56.3, 56.3, 41.6, 37.9, 37.9, 37.1, 35.5, 35.3, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_{\text{R}} = 0.86$, (m/z) $[\text{M}+\text{H}]^+$ 913 (100% purity); **HRMS** (ESI): calculated for $\text{C}_{47}\text{H}_{61}\text{N}_8\text{O}_9\text{S}$ (m/z) $[\text{M}+\text{H}]^+ = 913.4322$, found 913.4302; **IR** ($\nu_{\text{max}}/\text{cm}^{-1}$) 3286, 3079, 2958, 2878, 1651, 1609, 1540, 1487; $[\alpha]_{\text{D}}^{20} = -22$ (c 0.23, MeOH).

(2S,4R)-1-((S)-2-(3-(2-(3-((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)amino)-3-oxopropoxy)ethoxy)propanamido)-3,3-dimethylbutanoyl)-4-hydroxy-N-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide 209

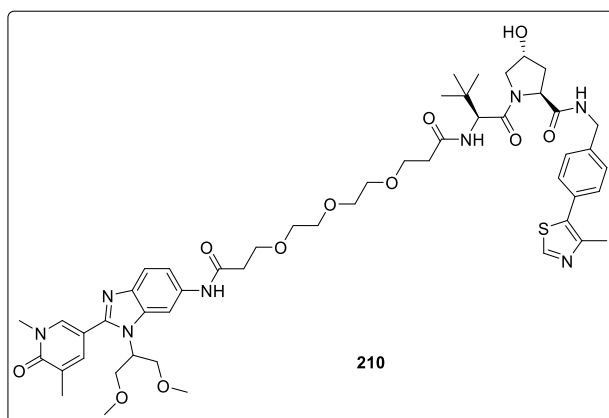


209 was synthesised following the **general procedure F** with the linker 3,3'-(ethane-1,2-diylbis(oxy))dipropionic acid (74.2 mg, 0.360 mmol). **209** (60 mg, 0.062 mmol, 21% yield) was isolated as a white solid.

m.p. 100-104 °C; δ_{H} (400 MHz, d_6 -DMSO) δ ppm 9.97 (s, 1 H), 8.97 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 8.22 (s, 1 H), 8.02 (s, 1 H), 7.89 (d, $J=9.5$ Hz, 1 H), 7.66 (s, 1 H), 7.53 (d, $J=8.4$ Hz, 1 H), 7.42-7.38 (m, 4 H), 7.29 (d, $J=8.6$ Hz, 1 H), 5.12 (s (br), 1 H), 4.83-4.79 (m, 1 H), 4.55 (d, $J=9.5$ Hz, 1 H), 4.45-4.40 (m, 2 H), 4.37-4.34 (m, 1 H), 4.24-4.20 (m, 1 H), 3.98-3.94 (m, 2 H), 3.77-3.71 (m, 4 H), 3.63-3.60 (m, 2 H), 3.53-3.47 (m, 9 H), 3.16 (s, 6 H), 2.59-2.57 (m, 2 H), 2.53-2.52 (m, 1 H), 2.44 (s, 3 H), 2.38-

2.33 (m, 1 H), 2.09 (s, 3 H), 2.05-2.02 (m, 1 H), 1.93-1.89 (m, 1 H), 0.93 (s, 9 H); δ_c (101 MHz, d_6 -DMSO) δ ppm 171.9, 169.9, 169.5, 169.0, 161.7, 151.5, 151.4, 147.7, 139.4, 139.2, 138.9, 137.3, 134.1, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 118.9, 114.7, 108.0, 102.7, 69.5, 69.5, 69.4, 68.8, 66.9, 66.2, 58.7, 58.4, 56.8, 56.3, 56.3, 41.6, 37.9, 37.4, 37.2, 35.6, 35.3, 26.3, 16.9, 15.9; **LCMS** (Method B): t_R = 0.88, (m/z) $[M+H]^+$ 957 (100% purity); **HRMS** (ESI): calculated for $C_{49}H_{65}N_8O_{10}S$ (m/z) $[M+H]^+$ = 957.4544, found 957.4562; **IR** (ν_{max}/cm^{-1}) 3291, 3074, 2958, 2878, 1651, 1609, 1540, 1482; $[\alpha]_D^{20} = -31$ (c 0.25, MeOH).

(2S,4R)-1-((S)-2-(tert-butyl)-16-((1-(1,3-Dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)amino)-4,16-dioxo-7,10,13-trioxa-3-azahexadecanoyl)-4-hydroxy-N-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide 210

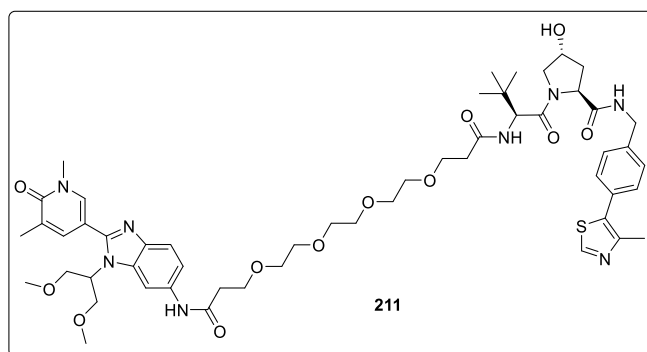


210 was synthesised following the **general procedure F** with the linker 3,3'-((oxybis(ethane-2,1-diyl))bis(oxy))dipropionic acid (90.0 mg, 0.360 mmol). **210** (67 mg, 0.067 mmol, 22% yield) was isolated as a white solid.

m.p. 84-90 °C; δ_H (400 MHz, d_6 -DMSO) δ ppm 10.01 (s, 1 H), 8.98 (s, 1 H), 8.55 (t, $J=6.1$ Hz, 1 H), 8.24 (s, 1 H), 8.03 (s, 1 H), 7.89 (d, $J=9.5$ Hz, 1 H), 7.66 (s, 1 H), 7.54 (d, $J=8.4$ Hz, 1 H), 7.43-7.38 (m, 4 H), 7.30 (d, $J=8.6$ Hz, 1 H), 5.13 (s (br), 1 H), 4.84-4.80 (m, 1 H), 4.56 (d, $J=9.5$ Hz, 1 H), 4.45-4.41 (m, 2 H), 4.37-4.35 (m, 1 H), 4.25-4.21 (m, 1 H), 3.98-3.95 (m, 2 H), 3.78-3.72 (m, 4 H), 3.64-3.62 (m, 2 H), 3.55-3.45 (m, 13 H), 3.18 (s, 6 H), 2.60-2.58 (m, 2 H), 2.56-2.53 (m, 1 H), 2.45 (s, 3 H), 2.38-2.33 (m, 1 H), 2.09 (s, 3 H), 2.06-2.04 (m, 1 H), 1.94-1.89 (m, 1 H), 0.94 (s, 9 H); δ_c (101 MHz, d_6 -DMSO) δ ppm 171.9, 169.9, 169.5, 169.0, 161.7, 151.5, 151.4, 147.7, 139.4, 139.2, 138.9, 137.3, 134.1, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 118.9,

114.7, 107.9, 102.7, 69.7, 69.6, 69.6, 69.5, 69.4, 68.8, 66.9, 66.6, 58.7, 58.4, 56.8, 56.3, 56.3, 41.6, 37.9, 37.4, 37.2, 35.6, 35.3, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_R = 0.89$, (m/z) $[M+2H]^{2+}$ 501 (100% purity); **HRMS** (ESI): calculated for $C_{51}H_{69}N_8O_{11}S$ (m/z) $[M+H]^+ = 1001.4807$, found 1001.4794; **IR** (ν_{max}/cm^{-1}) 3302, 3069, 2947, 2878, 1651, 1598, 1540, 1487; $[\alpha]_D^{20} = -27$ (c 0.23, MeOH).

N1-(1-(1,3-dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N16-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-4,7,10,13-tetraoxahexadecanediamide 211



211 was synthesised following the **general procedure F** with the linker 4,7,10,13-tetraoxahexadecanedioic acid (106 mg, 0.360 mmol). **211** (60 mg, 0.055 mmol, 18% yield) was isolated as a white solid.

m.p. 70-74 °C; δ_H (400 MHz, d_6 -DMSO) δ ppm 10.01 (s, 1 H), 8.97 (s, 1 H), 8.54 (t, $J=6.1$ Hz, 1 H), 8.24 (s, 1 H), 8.02 (s, 1 H), 7.88 (d, $J=9.5$ Hz, 1 H), 7.65 (s, 1 H), 7.53 (d, $J=8.8$ Hz, 1 H), 7.43-7.37 (m, 4 H), 7.29 (d, $J=8.6$ Hz, 1 H), 5.12 (s (br), 1 H), 4.84-4.78 (m, 1 H), 4.55 (d, $J=9.5$ Hz, 1 H), 4.44-4.41 (m, 2 H), 4.36-4.34 (m, 1 H), 4.24-4.20 (m, 1 H), 3.98-3.95 (m, 2 H), 3.77-3.44 (m, 23 H), 3.17 (s, 6 H), 2.60-2.53 (m, 3 H), 2.44 (s, 3 H), 2.37-2.32 (m, 1 H), 2.08 (s, 3 H), 2.05-2.01 (m, 1 H), 1.93-1.89 (m, 1 H), 0.93 (s, 9 H); δ_C (101 MHz, d_6 -DMSO) δ ppm 171.9, 169.9, 169.5, 169.0, 161.7, 151.5, 151.4, 147.7, 139.4, 139.2, 138.9, 137.3, 134.2, 133.4, 131.1, 129.6, 128.6, 127.4, 127.4, 118.9, 114.7, 107.9, 102.7, 69.7, 69.6, 69.6, 69.6, 69.4, 69.4, 69.4, 68.8, 66.9, 66.6, 58.7, 58.4, 56.8, 56.3, 56.3, 41.6, 37.9, 37.4, 37.2, 35.6, 35.3, 26.3, 16.9, 15.9; **LCMS** (Method B): $t_R = 0.89$, (m/z) $[M+2H]^{2+}$ 523 (96% purity); **HRMS** (ESI): calculated for $C_{53}H_{73}N_8O_{12}S$ (m/z) $[M+H]^+ = 1044.4994$, found 1044.4990; **IR** (ν_{max}/cm^{-1}) 3286, 3069, 2937, 2868, 1656, 1603, 1545, 1487; $[\alpha]_D^{20} = -21$ (c 0.14, MeOH).

5.5 Biological assays and Physicochemical Measurements

5.5.1 Compound Preparation

All test compounds were stored as 10 mM solutions in DMSO at 4 °C.

5.5.2 Artificial Membrane Permeability (AMP)

AMP was measured in a high-throughput plate-based assay that measured the rate of compound permeation through a phospholipid membrane at pH 7.4. The phospholipid membrane was prepared by dissolving egg phosphatidyl choline (1.8%) and cholesterol (1%) in *n*-decane. An aliquot (3.5 µL) was transferred to the bottom of the microfiltration filter inserts in a Transwell® plate. Phosphate buffer (50 mM Na₂HPO₄ with 0.5% (2-hydroxypropyl)-β-cyclodextrin) pH 7.4 was added to the acceptor well (100 µL) and donor well (250 µL) of the plate. The Transwell® plate was shaken at ambient temperature for 40 min (630 rpm) to allow the phospholipid to form a bilayer across the small holes in the filter. Test compound (10 mM in DMSO) (2.5 µL) was added to the donor well and the plate was shaken at ambient temperature for 3 h (630 rpm). The test compound concentration in both the acceptor and donor wells was determined by HPLC.²⁰³ The AMP is a measure of the rate of compound permeability through the phospholipid membrane and is expressed in nm/s. Reported values are an average of measurements from two independent wells.

5.5.3 Brd4 BD1/BD2 Time-Resolved Förster Resonance Energy Transfer (FRET)

A time-resolved FRET assay was used to calculate the Brd4 pIC₅₀ of test compounds. The fluorophore Alexa Fluor® 647, was tagged with the fluoro-ligand GSK2791026A (**Figure 115**), that binds to both bromodomains of Brd4 with equal affinity ($K_d \sim 50$ nM) and acts as the acceptor in the FRET pair. To determine the BD1 potency, mutant Brd4 Y390A was used. This site-directed mutation decreases the binding affinity of the Alexa Fluor® 647 ligand **212** at BD2 > 1000-fold. To determine the BD2 potency, mutant Brd4 Y97A was used, as this decreases the binding affinity of the Alexa Fluor® 647 ligand **212** at BD1 > 1000-fold. To detect the binding of the

Alexa Fluor[®] 647 ligand **212**, the 6His purification tag at the *N*-terminus of Brd4 was utilised. This acts as the epitope for anti-6His europium chelate (Eu), enabling binding of the lanthanide europium to Brd4, which acts as the donor in the FRET pair. The binding of the Alexa Fluor[®] 647 ligand **212** to Brd4 bromodomains was detected by excitation of Eu at 337 nm using a Perkin Elmer EnVision[®] laser. This leads to energy transfer from the Eu chelate to the Alexa Fluor[®] 647 ligand **212**, resulting in emission at 665 nm. Emission from Eu directly at 615 nm was used as an internal control, enabling an internal normalisation for emission through the ratio of the acceptor over the donor. Binding of the test compound at the unmutated bromodomain will inhibit the FRET and result in a decrease in the emission at 665 nm.

Serial dilution of a stock solution of test compound (10 mM in DMSO) furnished a range of test concentration points. An aliquot (50 nL) of each concentration was dispensed into a black Greiner 384-well low volume microtiter plate. All assay components were suspended in assay medium (50 mM HEPES, 150 mM NaCl, 1 mM CHAPS, 5% Glycerol and 1 mM DTT at pH 7.4 NaOH). The final concentration of Brd4 protein was at 10 nM and the Alexa Fluor[®] 647 ligand **212** was at K_d (~50 nM for Brd4). These components were premixed and an aliquot (5 μ L) of this mixture was added to each well of the plate, before incubating for 30 min at ambient temperature. Detection reagents were prepared in assay medium by diluting Eu-W1024 Anti-6xHis Antibody (AD0111 Perkin Elmer) to 1.5 nM. An aliquot (5 μ L) of this solution was added to each well of the plate. The plates were read on the Perkin Elmer EnVision[®] reader and the donor and acceptor counts were determined. From this, the ratio of acceptor/donor was calculated ($\lambda_{ex} = 337$ nm, λ_{em} donor = 615 nm, λ_{em} acceptor = 665 nm) and used for data analysis. All data was normalised to the mean of 16 high and 16 low control wells on each plate. A four parameter curve fit was used to calculate pIC_{50} using **Equation 2**.

$$y = \frac{a - d}{1 + \left(\frac{x}{c}\right)^b} + d$$

Equation 2: Equation used to calculate pIC_{50} (c) from the minimum (a), the Hill slope (b) and the maximum (d).

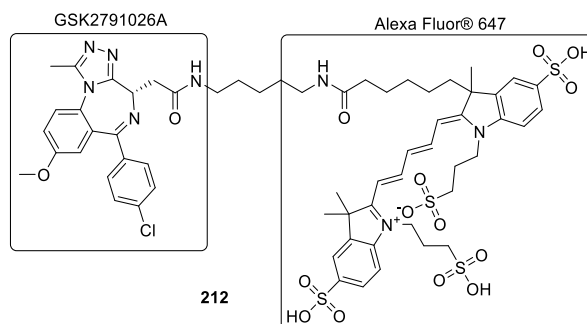


Figure 115: Alexa Fluor® 647 ligand **212**.

5.5.4 Brd4 HiBiT Degradation

The Brd4 pDC₅₀ and maximum % degradation of test compounds was measured by quantifying the reduction in luminescence from HiBiT-tagged Brd4 protein in human embryonic kidney (HEK293) cells. A 3-fold serial dilution of a stock solution of test compound (10 mM in DMSO) was performed 10 times to obtain 11 concentration points. An aliquot (25 nL) of each concentration was dispensed into a white Nunc 384-well plate. HEK293 HiBiT Brd4 cells were prepared in-house using gene editing CRISPR-Cas9 technology to encode the 11 amino acid HiBiT tag just before the Brd4 gene in DNA. This results in a natural level of cellular expression of HiBiT-tagged Brd4. A cryopreserved vial of HEK293 HiBiT Brd4 cells was rapidly thawed and suspended in assay medium (Fluorobrite™ DMEM supplemented with 10% heat inactivated FBS and pen-strep) (30 mL). The solution was centrifugated using a Heraeus™ Megafuge™ for 5 min (400 g). The supernatant was disposed, and the cell pellet was re-suspended in assay medium at a concentration of 10000 cells/25 μL. An aliquot (25 μL) of cell suspension was added to each well of the Nunc 384-well plate. One well containing only cell suspension (25 μL) was used as the 0% effect control (Z) and one well containing only assay medium (25 μL) was used as the 100% effect control (H). The plate was incubated at 37 °C with 10% CO₂ for 18 h. Nano-Glo® HiBiT Lytic Detection System (Promega prod no. N3040, containing LgBiT Protein and the Nano-Glo® HiBiT lytic substrate) (25 μL) was added to each well and the plate was shaken at ambient temperature for 10 min (500 rpm). This results in cell lysis and the HiBiT and LgBiT tags with picomolar affinity bind to form the NanoBiT, which is a luciferase protein that luminesces. The intensity of luminescence was measured

using a PHERAstar[®] microplate reader (focal height 12 mm, interval time 1 s, gain 3600) and the output data was exported and analysed using GraphPad Prism to fit a four parameter logistic regression curve.

% Brd4 remaining was calculated using the formula in **Equation 3**.

$$\% \text{ Brd4 remaining} = 100 \times \frac{(L - L_H)}{(L_Z - L_H)}$$

Equation 3: Equation used to calculate % Brd4 remaining from the PHERAstar[®] microplate reader luminescence readout. L = raw luminescence value from test compound well. L_H = raw luminescence value for 100% effect control (H). L_Z = raw luminescence value for 0% effect control (Z).

pDC₅₀ is calculated from the concentration of test compound at which 50% of L_Z is achieved. Maximum % degradation is reported for the minimum value of % Brd4 remaining, with rDC₅₀ reported at 50% of the maximum % degradation.

5.5.4.1 Epoxomicin 26 S Proteasome Inhibition

Prior to performing the Brd4 HiBiT degradation assay, **Section 5.5.4**, cells (cell suspension as above) were pre-incubated with epoxomicin at a concentration of 1 μM at 37 °C with 10% CO₂ for 1 h before dispensing into the assay plate. Subsequent assay incubation time was reduced from 18 h to 4 h to minimise potential cytotoxicity.

5.5.5 Brd4 Human Whole Blood MCP-1

This assay utilises human whole blood containing monocytes, T cells and B cells. Monocytes migrate to sites of inflammation and secrete inflammatory cytokines, thereby activating other cell types such as T cells. Stimulation of monocytes with lipopolysaccharide (LPS) leads to downstream signaling events *via* its toll-like receptor 4. This assay measures the release of one of the key cytokines, monocyte chemoattractant protein-1 (MCP-1), from peripheral blood mononuclear cells (PBMC) in whole blood incubated with test compound and then stimulated with LPS. MCP-1 levels are analysed using a bead-based ELISA method and TTP Mirrorball[®]. The assay

assesses the ability of the test compound to inhibit the production of MCP-1 from the monocytic population of whole blood.

A 3-fold serial dilution of a stock solution of test compound (10 mM in DMSO) was performed 9 times to obtain 10 concentration points. An aliquot (1 μ L) of each concentration was dispensed into a white 96-well Greiner TC flat clear bottomed plate. A negative control (100% DMSO) (1 μ L) and positive control (10 mM GSK3169442A **213, Figure 116**) (1 μ L) was also dispensed into separate wells. Human blood was collected into sodium heparin (1 ml/100 ml blood) at the GSK Stevenage blood donation unit. Whole blood (130 μ L) was added to each well of the plate. The plate was incubated at 37 °C with 5% CO₂ for 30 min. LPS in phosphate buffer saline (2.8 μ g/mL) (10 μ L) was added to each well of the plate. The plate was incubated at 37 °C with 5% CO₂ for 18 h. Phosphate buffer saline (140 μ L) was added to each well in the plate. The plate was centrifuged using a Heraeus™ Megafuge™ for 10 min (2000 rpm). Phosphate buffer saline (1% bovine serum albumin) (30 μ L) was dispensed into each well of a separate 96-well Greiner PP V-bottom plate. The plate was centrifuged using a Heraeus™ Megafuge™ for 1 min (1000 rpm). The supernatant (30 μ L) from each well of the white Greiner TC flat clear bottomed plate was transferred into a well of the Greiner PP V-bottom plate. An aliquot (20 μ L) of the resulting solution from each well was transferred to a separate well of the final assay plate. This plate was centrifuged using a Heraeus™ Megafuge™ for 1 min (1000 rpm). A 200-fold dilution of the capture beads with phosphate buffer saline (1% bovine serum albumin) was performed in a 50 mL falcon. The beads were vortexed briefly to ensure the formation of a homogenous 'MCP-1 capture bead solution'. MCP-1 capture bead solution (10 μ L) was added to each well of the assay plate. The plate was centrifuged using a Heraeus™ Megafuge™ for 1 min (1000 rpm). A plate seal was added, and the plate was covered with foil and placed on a shaker at ambient temperature for 2 h (1000 rpm). A 400-fold dilution of the BD™ CBA MCP-1 PE detection reagent with phosphate buffer saline (1% bovine serum albumin) was performed in a 50 mL falcon to form the 'MCP-1 detection antibody solution'. MCP-1 detection antibody solution (10 μ L) was added to each well of the assay plate. The plate was centrifuged using a Heraeus™ Megafuge™ for 1 min (1000 rpm). A plate

seal was added and the plate was covered with foil and placed on a shaker at ambient temperature for 2 h (1000 rpm).

FL-3 median fluorescence intensity was measured using the TTP Mirrorball[®], only events with a perimeter of 86-120 μm were recorded. Raw data was analysed using Activitybase XE. All data was normalised to the mean of 16 high (0% inhibition) and 16 low (100% inhibition) control wells on each plate. The robust algorithm was used to calculate the mean and standard deviation for each set of controls. This method down-weights the outliers during the calculation, so that their physical removal from the data set becomes unnecessary. The well result was then calculated by dividing the difference between sample measurements and the average of positive controls by the difference between positive and negative controls. This method yields response values that are a normalised percentage of the control. The dose-response curve was modelled by the four parameter logistic curve, based on **Equation 4**.

$$y = a + \frac{b - a}{1 + \left(\frac{10^c}{x}\right)^d}$$

Equation 4: Equation used to calculate pIC_{50} (c) from the known concentration (x), response (y), minimum value of y (a), maximum value of y (b) and the slope factor (d). Starting values for pre-fit set: $a = 0$, $b = 100$, $c = -6$ and $d = 1$.

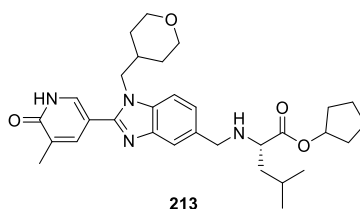


Figure 116: Positive control GSK3169442A **213**.

5.5.6 Cell Concentration Permeability (PAC)

A detergent rupture method was used to determine the concentration of test compounds in HeLa cells by Rapidfire[®] MS/MS.²²⁵ This measurement is used to calculate PAC using **Equation 5**.

$$P\Delta C = \text{Log} \left(\frac{[\text{Test Compound in cells}]}{[\text{Test Compound in no cell control}]} \right)$$

Equation 5: Equation used to calculate $P\Delta C$.

In-house cultured HeLa cells were suspended at 2×10^6 cells/mL in assay medium (MR1-4 custom media (SAFC - 81058-1000M1172) with 10% FBS). Test compound (10 mM in DMSO) (4 μL) was added to wells C1 and D1 of a Greiner 96-well V bottom plate (plate 1) and wells E1 and F1 of an analogous plate (plate 2). This process was repeated, adding test compound 2 to wells C2, D2, E2 and F2. Assay medium (2 mL) was added to rows C and D of plate 1. HeLa cell suspension (2 mL) was added to rows E and F of plate 2. Both plates were covered with a Greiner plate lid and incubated at 37 °C with 5% CO₂ in 85% humidity for 2.5 h with shaking. Plate 2 was centrifuged in a Sorvall™ ST40R at 4 °C for 5 min (1000 g). The supernatant was disposed, and phosphate buffered saline (1 mL) was added to rows E and F. The cell pellet was resuspended using an Agilent Bravo by repeat aspiration and ejection. Plate 2 was centrifuged in a Sorvall™ ST40R at 4 °C for 5 min (1000 g). The supernatant was disposed, and M-PER solution (Thermo Fischer Scientific) (180 μL) was added to rows E and F. The cell pellet was resuspended using an Agilent Bravo by repeat aspiration and ejection. Plate 2 was incubated at 37 °C with 5% CO₂ in 85% humidity for 15 min with shaking. Plate 2 was centrifuged in a Sorvall™ ST40R at 4 °C for 30 min (4100 g). The supernatant was transferred to rows E and F of plate 3. An aliquot (260 μL) of each well in rows C and D of plate 1 was transferred to rows C and D of plate 3. An aliquot (5 μL) of each well in plate 3 was transferred to the corresponding well in plate 4. RPMI medium (Sigma Aldrich) (45 μL), 50% MeCN (95 μL) and internal standard solution (containing 100 ng/mL Sulfamethazine in 5% MeCN) (5 μL) was added to all wells in rows C, D, E and F of plate 4. Plate 4 was analysed by Rapidfire® MS/MS using an AB Sciex 6500 mass spectrometer. Peak area was converted to concentration using a calibration assay with a top concentration of 300 μM of test compound and a 3-fold serial dilution to obtain eight data points.

5.5.7 Charged Aerosol Detector (CAD) Solubility

A 20-fold dilution of test compound (10 mM in DMSO) (5 µL) was performed with pH 7.4 phosphate buffered saline. The solution was equilibrated for 1 h at ambient temperature and filtered through a Millipore Multiscreen[®] HTS-PCF filter plate. The filtrate was quantified by suitably calibrated CAD.²²⁶ The dynamic range of the assay was 1-500 µg/mL when working from 10 mM DMSO stock solution.

5.5.8 ChromlogD

Chromatographic Hydrophobicity Index (CHI) was measured by HPLC using a Phenomenex Gemini[®] NX C18 reverse phase column (50 x 2 mm, 3 µm), eluting with an MeCN gradient with starting mobile phase pHs of 2, 7.4 and 10.5.²²⁷ CHI values were derived directly from the gradient retention times using a calibration line obtained from standard compounds. The CHI value was linearly transformed into ChromlogD by least-square fitting of experimental CHI values to calculated logP values for over 20,000 research compounds using **Equation 6**.¹⁸⁶

$$\text{ChromlogD} = 0.0857(\text{CHI}) - 2$$

Equation 6: Equation used to calculate ChromlogD from experimentally measured CHI values.

5.5.9 Fasted-State Simulated Intestinal Fluid (FaSSIF) Solubility

The solubility of solid compounds in FaSSIF solution was measured at pH 6.5 after 4 h equilibration at ambient temperature. FaSSIF buffer (3 mM sodium taurocholate, 0.75 mM lecithin in sodium phosphate buffer at pH 6.5) (1 mL) was added to solid test compound (1 mg) in a 2 mL vial. The resulting suspension was shaken at ambient temperature for 4 h (900 rpm) and then transferred to a Multiscreen HTS, 96-well solubility filter plate. The residual solid was removed by filtration. The supernatant solution was quantified by HPLC-UV using single-point calibration of a known concentration of the compound in DMSO. The dynamic range of the assay was 1-1000 µg/mL.^{228,229}

5.5.10 Human Serum Albumin (HSA) Binding

HSA binding was measured using a Chiral Technologies CHIRALPAK[®] HSA HPLC column (50 x 3 mm, 5 µm), eluting with 0-30% *iso*-propanol in 50 mM ammonium acetate buffer, pH 7.4. HSA binding values were derived directly from the gradient retention times using a calibration line obtained from standard compounds.^{201,230}

5.5.11 Immobilised Artificial Membrane (IAM) Binding

IAM binding was measured using a Regis Technologies PC DD2 (100 x 4.6 mm, 10 µM) immobilised phosphatidylcholine HPLC column, eluting with 0-85% MeCN in 50 mM ammonium acetate buffer, pH 7.4. CHI_{IAM} values were derived directly from the gradient retention times using a calibration line obtained from standard compounds.²³¹ The CHI_{IAM} value was linearly transformed into the logarithmic retention factor ($\log(K_{IAM})$) using **Equation 7**.²⁰¹

$$\log(K_{IAM}) = 0.046(CHI_{IAM}) + 0.42$$

Equation 7: Equation used to calculate $\log(K_{IAM})$ from experimentally measured CHI_{IAM} values.

5.6 Computational Methods

5.6.1 Docking Models

Test compound docking models were prepared by Dr. Paul Bamborough using Maestro 2018.3. The Brd4 BD1 protein structure was prepared from the published pdb4wiv crystal structure. A two-step restrained energy minimisation of the test compound was performed using the OPLS3e forcefield. A Glide grid was constructed using default parameters and the test compound was docked. The Brd4 BD1 docked test compounds were visualised in MOE 2019.10.

5.6.2 PCA/PLS analysis

PCA/PLS analysis was performed using Umetrics® SIMCA® 13.0.3 with the guidance of Chris Luscombe. The input physicochemical descriptors and responses were preprocessed by scaling and mean-centring the data sets. The PCA/PLS analysis was performed using the O2PLS method with default settings.

All calculated physicochemical descriptors were obtained using GSK in-house models. The definitions of the PCv6, Abraham and e-state descriptors are outlined in **Table 44**.

Table 44: Definitions of the PCv6, Abraham and e-state descriptors used in the PCA/PLS analysis.

Descriptor	Definition
PCv6 ²³²	
PCv6.abe	Andrews binding energy ²³³
PCv6.clogP_bb	clogP (BioByte software)
PCv6.CMR	CMR
PCv6.HAC	HAC
PCv6.M _w	M _w
PCv6.arinc	ARc
PCv6.narinc	Non-aromatic ring count
PCv6.RBc	RBc
PCv6.flex	Ratio of the number of rotatable bonds to total bonds
PCv6.alpha	Hydrogen bond acidity (HBD)
PCv6.betah	Hydrogen bond basicity (HBA)
PCv6.pi	Dipolarity/dipolarisability
PCv6.r2	Excess molar refraction
PCv6.vx	McGowan characteristic volume ²³⁴

PCv6.HBA	HBA (nitrogen, oxygen and sulfur)
PCv6.HBD	HBD (nitrogen, oxygen and sulfur)
PCv6.neg	Negatively ionisable/charged groups count
PCv6.pos	Positively ionisable/charged groups count
PCv6.total_HB	HBA + HBD
PCv6.total_charge	Sum of positive and negative charges
PCv6.fraction_HBA	HBA ÷ CMR
PCv6.fraction_HBD	HBD ÷ CMR
PCv6.TPSA	TPSA (nitrogen and oxygen only) ¹⁹³
PCv6.nonPSA	Approximation of total surface area minus PSA (Molar Volume - TPSA)
PCv6.pct_PSA	PSA scaled by size ($[TPSA \div \text{Molar Volume}] \times 100$)
<hr/>	
Abraham ²³⁵	
Abraham.alpha	Hydrogen bond acidity (HBD)
Abraham.betah	Hydrogen bond basicity (HBA)
Abraham.pi	Dipolarity/dipolarisability
Abraham.r2	Excess molar refraction
Abraham.vx	McGowan characteristic volume ²³⁴
<hr/>	
e-states ^{236,237}	
e-state.CHR3	Tertiary carbon
e-state.CR4	Quaternary carbon
e-state.methylene	Methylene
e-state.methyl	Methyl
e-state.HRC=	Monosubstituted carbon double-bond
e-state.R2C=	Disubstituted carbon double-bond
e-state.alkHRC=	Alkyl monosubstituted carbon double-bond
e-state.alkR2C=	Alkyl disubstituted carbon double-bond
e-state.NH2R	Primary amine
e-state.amide	Any amide
e-state.amideNH	Primary amide
e-state.amideOtoNH	Primary amide
e-state.anycarbonyl	Carbonyl
e-state.O=	Oxygen double-bond
e-state.hydrovolume	Measure of hydrophilicity
e-state.lipovolume	Measure of lipophilicity
e-state.N=(cyc)	Nitrogen inside cyclic ring
e-state.methyl_notCChet	Methyl not connected to a heteroatom

5.6.3 Principal Moments of Inertia (PMI) Plots

PMI plots were generated with guidance from Dr. Carol Mulrooney. The lowest energy conformation for each of the 63 PROTACs in the PROTAC series was generated in MOE 2019.01, using the MMFF94s forcefield. A conformational search was

performed on each lowest energy conformation using the LowModeMD search method, which generates conformations using a short (~1 ps) run of molecular dynamics at constant temperature, followed by an all-atom energy minimisation.²³⁸ An energy window of 7 kcal/mol from the lowest energy conformation was used, as higher energy conformers represent a relative population $< 10^{-5}$ at 298 K.²²⁰ A RMSD of ≥ 0.25 Å was required to generate distinct conformations, with an iteration limit of 30. Three PMI's, I_1 , I_2 and I_3 , were calculated for each conformation using the method described by Sauer and co-workers.²¹⁸ The two smallest PMI's were independently divided by the largest one to generate npr1 and npr2. These values are used as the [x, y] coordinates in the triangular graph of the PMI plots.

6: References

6 References

- (1) Tinworth, C. P.; Lithgow, H.; Churcher, I. *Med. Chem. Comm.* **2016**, *7*, 2206–2216.
- (2) Grimwood, S.; Hartig, P. R. *Pharmacol. Ther.* **2009**, *122*, 281–301.
- (3) Adjei, A. A. *J. Clin. Oncol.* **2006**, *24*, 4054–4055.
- (4) Cusack, K. P.; Wang, Y.; Hoemann, M. Z.; Marjanovic, J.; Heym, R. G.; Vasudevan, A. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 2019–2027.
- (5) Cambridge, S. B.; Gnad, F.; Nguyen, C.; Bermejo, J. L.; Krüger, M.; Mann, M. *J. Proteome Res.* **2011**, *10*, 5275–5284.
- (6) Uetrecht, J. *Chem. Res. Toxicol.* **2008**, *21*, 84–92.
- (7) Uetrecht, J. *Chem. Res. Toxicol.* **2009**, *22*, 24–34.
- (8) Barf, T.; Kaptein, A. *J. Med. Chem.* **2012**, *55*, 6243–6262.
- (9) Rask-Andersen, M.; Almén, M. S.; Schiöth, H. B. *Nat. Rev. Drug Discov.* **2011**, *10*, 579–590.
- (10) Holohan, C.; Van Schaeybroeck, S.; Longley, D. B.; Johnston, P. G. *Nat. Rev. Cancer* **2013**, *13*, 714–726.
- (11) Blair, J. M. A.; Webber, M. A.; Baylay, A. J.; Ogbolu, D. O.; Piddock, L. J. V. *Nat. Rev. Microbiol.* **2015**, *13*, 42–51.
- (12) Kinch, L. N.; Grishin, N. V. *Curr. Opin. Struct. Biol.* **2002**, *12*, 400–408.
- (13) Barrangou, R.; Doudna, J. A. *Nat. Biotechnol.* **2016**, *34*, 933–941.
- (14) Valeur, E.; Guéret, S. M.; Adihou, H.; Gopalakrishnan, R.; Lemurell, M.; Waldmann, H.; Grossmann, T. N.; Plowright, A. T. *Angew. Chem. Int. Ed.* **2017**, *56*, 10294–10323.
- (15) Sharma, V. K.; Watts, J. K. *Future Med. Chem.* **2015**, *7*, 2221–2242.
- (16) Waddington, S. N.; Privolizzi, R.; Karda, R.; O’Neill, H. C. *Curr. Stem Cell Rep.* **2016**, *2*, 9–20.
- (17) Churcher, I. *J. Med. Chem.* **2018**, *61*, 444–452.
- (18) Hochstrasser, M. *Curr. Opin. Cell Biol.* **1995**, *7*, 215–223.
- (19) Navon, A.; Ciechanover, A. *J. Biol. Chem.* **2009**, *284*, 33713–33718.
- (20) Collins, G. A.; Goldberg, A. L. *Cell* **2017**, *169*, 792–806.

- (21) Rock, K. L.; Gramm, C.; Rothstein, L.; Clark, K.; Stein, R.; Dick, L.; Hwang, D.; Goldberg, A. L. *Cell* **1994**, *78*, 761–771.
- (22) Baehrecke, E. H. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 505–510.
- (23) Kleiger, G.; Mayor, T. *Trends Cell Biol.* **2014**, *24*, 352–359.
- (24) Hershko, A. *Angew. Chem. Int. Ed.* **2005**, *44*, 5932–5943.
- (25) Hershko, A.; Ciechanover, A. *Annu. Rev. Biochem.* **1992**, *61*, 761–807.
- (26) Ciechanover, A. *Cell* **1994**, *79*, 13–21.
- (27) Weissman, A. M.; Shabek, N.; Ciechanover, A. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 605–620.
- (28) Cromm, P. M.; Crews, C. M. *Cell Chem. Biol.* **2017**, *24*, 1181–1190.
- (29) Zaaroor-Regev, D.; De Bie, P.; Scheffner, M.; Noy, T.; Shemer, R.; Heled, M.; Stein, I.; Pikarsky, E.; Ciechanover, A. *Proc. Natl. Acad. Sci.* **2010**, *107*, 6788–6793.
- (30) Hicke, L.; Dunn, R. *Annu. Rev. Cell Dev. Biol.* **2003**, *19*, 141–172.
- (31) Haglund, K. *Trends Biochem. Sci.* **2003**, *28*, 598–604.
- (32) Komander, D.; Rape, M. *Annu. Rev. Biochem.* **2012**, *81*, 203–229.
- (33) Komander, D. *Biochem. Soc. Trans.* **2009**, *37*, 937–953.
- (34) Chau, V.; Tobias, J.; Bachmair, A.; Marriott, D.; Ecker, D.; Gonda, D.; Varshavsky, A. *Science* **1989**, *243*, 1576–1583.
- (35) Thrower, J. S. *EMBO J.* **2000**, *19*, 94–102.
- (36) Lu, Y.; Lee, B.; King, R. W.; Finley, D.; Kirschner, M. W. *Science* **2015**, *348*, 1250834–1250834.
- (37) Braten, O.; Livneh, I.; Ziv, T.; Admon, A.; Kehat, I.; Caspi, L. H.; Gonen, H.; Bercovich, B.; Godzik, A.; Jahandideh, S. *Proc. Natl. Acad. Sci.* **2016**, *113*, 4639–4647.
- (38) Bustamante, H. A.; González, A. E.; Cerda-Troncoso, C.; Shaughnessy, R.; Otth, C.; Soza, A.; Burgos, P. V. *Front. Cell. Neurosci.* **2018**, *12*, 1–18.
- (39) Verma, R. *Science* **2002**, *298*, 611–615.
- (40) Groll, M.; Huber, R. *Biochim. Biophys. Acta Mol. Cell Res.* **2004**, *1695*, 33–44.
- (41) Nussbaum, A. K.; Dick, T. P.; Keilholz, W.; Schirle, M.; Stevanovic, S.; Dietz, K.; Heinemeyer, W.; Groll, M.; Wolf, D. H.; Huber, R. *Proc. Natl. Acad. Sci.* **1998**, *95*, 12504–12509.

- (42) Dick, T. P.; Nussbaum, A. K.; Deeg, M.; Heinemeyer, W.; Groll, M.; Schirle, M.; Keilholz, W.; Stevanović, S.; Wolf, D. H.; Huber, R. *J. Biol. Chem.* **1998**, *273*, 25637–25646.
- (43) Kisselev, A. F.; Goldberg, A. L. *Chem. Biol.* **2001**, *8*, 739–758.
- (44) Cvek, B. *Prog. Mol. Biol. Transl.* **2012**, *109*, 161–226.
- (45) Drexler, H. C. A.; Risau, W.; Konecny, M. A. *FASEB J.* **2000**, *14*, 65–77.
- (46) Manasanch, E. E.; Orlovski, R. Z. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 417–433.
- (47) Guo, N.; Peng, Z. *Asia. Pac. J. Clin. Oncol.* **2013**, *9*, 6–11.
- (48) Adams, J.; Palombella, V. J.; Sausville, E. A.; Johnson, J.; Destree, A.; Lazarus, D. D.; Maas, J.; Pien, C. S.; Prakash, S. *Cancer Res.* **1999**, *59*, 2615–2622.
- (49) Chen, P.; Hochstrasser, M. *Cell* **1996**, *86*, 961–972.
- (50) Heinemeyer, W.; Fischer, M.; Krimmer, T.; Stachon, U.; Wolf, D. H. *J. Biol. Chem.* **1997**, *272*, 25200–25209.
- (51) Chen, D.; Frezza, M.; Schmitt, S.; Kanwar, J. *Curr. Cancer Drug Targets* **2011**, *11*, 239–253.
- (52) Moreau, P.; Richardson, P. G.; Cavo, M.; Orlovski, R. Z.; San Miguel, J. F.; Palumbo, A.; Harousseau, J. *Blood* **2012**, *120*, 947–959.
- (53) Arastu-Kapur, S.; Anderl, J. L.; Kraus, M.; Parlati, F.; Shenk, K. D.; Lee, S. J.; Muchamuel, T.; Bennett, M. K.; Driessen, C.; Ball, A. J. *Clin. Cancer Res.* **2011**, *17*, 2734–2743.
- (54) Fostier, K.; De Becker, A.; Schots, R. *Onco. Targets. Ther.* **2012**, *5*, 237–244.
- (55) Groll, M.; Kim, K. B.; Kairies, N.; Huber, R.; Crews, C. M. *J. Am. Chem. Soc.* **2000**, *122*, 1237–1238.
- (56) Meng, L.; Kwok, B. H. B.; Sin, J. *Cancer Res.* **1999**, *59*, 2798–2801.
- (57) Meng, L.; Mohan, R.; Kwok, B. H. B.; Eloffson, M.; Sin, N.; Crews, C. M. *Proc. Natl. Acad. Sci.* **1999**, *96*, 10403–10408.
- (58) Kaplan, G. S.; Torcun, C. C.; Grune, T.; Ozer, N. K.; Karademir, B. *Free Radic. Biol. Med.* **2017**, *103*, 1–13.
- (59) Jin, J.; Li, X.; Gygi, S. P.; Harper, J. W. *Nature* **2007**, *447*, 1135–1138.
- (60) Landré, V.; Rotblat, B.; Melino, S.; Bernassola, F.; Melino, G. *Oncotarget* **2014**, *5*, 7988–8013.

- (61) Yang, Y.; Kitagaki, J.; Dai, R.-M.; Tsai, Y. C.; Lorick, K. L.; Ludwig, R. L.; Pierre, S. A.; Jensen, J. P.; Davydov, I. V.; Oberoi, P. *Cancer Res.* **2007**, *67*, 9472–9481.
- (62) Huang, X.; Dixit, V. M. *Cell Res.* **2016**, *26*, 484–498.
- (63) da Silva, S. R.; Paiva, S.-L.; Lukkarila, J. L.; Gunning, P. T. *J. Med. Chem.* **2013**, *56*, 2165–2177.
- (64) Ye, Y.; Rape, M. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 755–764.
- (65) Ceccarelli, D. F.; Tang, X.; Pelletier, B.; Orlicky, S.; Xie, W.; Plantevin, V.; Neculai, D.; Chou, Y.-C.; Ogunjimi, A.; Al-Hakim, A. *Cell* **2011**, *145*, 1075–1087.
- (66) Huang, H.; Ceccarelli, D. F.; Orlicky, S.; St-Cyr, D. J.; Ziemba, A.; Garg, P.; Plamondon, S.; Auer, M.; Sidhu, S.; Marinier, A. *Nat. Chem. Biol.* **2014**, *10*, 156–163.
- (67) Mattern, M. R.; Wu, J.; Nicholson, B. *Biochim. Biophys. Acta - Mol. Cell Res.* **2012**, *1823*, 2014–2021.
- (68) Cohen, P.; Tcherpakov, M. *Cell* **2010**, *143*, 686–693.
- (69) Manalo, D. J. *Blood* **2005**, *105*, 659–669.
- (70) Semenza, G. L. *Science* **2007**, *318*, 62–64.
- (71) Brahim-Horn, M. C.; Pouysségur, J. *Biochem. Pharmacol.* **2007**, *73*, 450–457.
- (72) Buckley, D. L.; Van Molle, I.; Gareiss, P. C.; Tae, H. S.; Michel, J.; Noblin, D. J.; Jorgensen, W. L.; Ciulli, A.; Crews, C. M. *J. Am. Chem. Soc.* **2012**, *134*, 4465–4468.
- (73) Buckley, D. L.; Gustafson, J. L.; Van Molle, I.; Roth, A. G.; Tae, H. S.; Gareiss, P. C.; Jorgensen, W. L.; Ciulli, A.; Crews, C. M. *Angew. Chem. Int. Ed.* **2012**, *51*, 11463–11467.
- (74) Galdeano, C.; Gadd, M. S.; Soares, P.; Scaffidi, S.; Van Molle, I.; Birced, I.; Hewitt, S.; Dias, D. M.; Ciulli, A. *J. Med. Chem.* **2014**, *57*, 8657–8663.
- (75) Van Molle, I.; Thomann, A.; Buckley, D. L.; So, E. C.; Lang, S.; Crews, C. M.; Ciulli, A. *Chem. Biol.* **2012**, *19*, 1300–1312.
- (76) Vargesson, N. *Birth Defects Res. C.* **2015**, *105*, 140–156.
- (77) Mohty, B.; El-Cheikh, J.; Yakoub-Agha, I.; Avet-Loiseau, H.; Moreau, P.; Mohty, M. *Leukemia* **2012**, *26*, 73–85.
- (78) Quach, H.; Ritchie, D.; Stewart, A. K.; Neeson, P.; Harrison, S.; Smyth, M. J.; Prince, H. M. *Leukemia* **2010**, *24*, 22–32.

- (79) Ramsay, A. G.; Clear, A. J.; Fatah, R.; Gribben, J. G. *Blood* **2012**, *120*, 1412–1421.
- (80) Ito, T.; Ando, H.; Suzuki, T.; Ogura, T.; Hotta, K.; Imamura, Y.; Yamaguchi, Y.; Handa, H. *Science* **2010**, *327*, 1345–1350.
- (81) Fink, E. C.; Ebert, B. L. *Blood* **2015**, *126*, 2366–2369.
- (82) Fischer, E. S.; Böhm, K.; Lydeard, J. R.; Yang, H.; Stadler, M. B.; Cavadini, S.; Nagel, J.; Serluca, F.; Acker, V.; Lingaraju, G. M. *Nature* **2014**, *512*, 49–53.
- (83) Chamberlain, P. P.; Lopez-Girona, A.; Miller, K.; Carmel, G.; Pagarigan, B.; Chie-Leon, B.; Rychak, E.; Corral, L. G.; Ren, Y. J.; Wang, M. *Nat. Struct. Mol. Biol.* **2014**, *21*, 803–809.
- (84) Bulatov, E.; Ciulli, A. *Biochem. J.* **2015**, *467*, 365–386.
- (85) Holstein, S. A.; Hillengass, J.; McCarthy, P. L. *J. Clin. Oncol.* **2018**, *36*, 2101–2104.
- (86) Gandhi, A. K.; Kang, J.; Havens, C. G.; Conklin, T.; Ning, Y.; Wu, L.; Ito, T.; Ando, H.; Waldman, M. F.; Thakurta, A. *Br. J. Haematol.* **2014**, *164*, 811–821.
- (87) Kronke, J.; Udeshi, N. D.; Narla, A.; Grauman, P.; Hurst, S. N.; McConkey, M.; Svinkina, T.; Heckl, D.; Comer, E.; Li, X. *Science* **2014**, *343*, 301–305.
- (88) Lai, A. C.; Crews, C. M. *Nat. Rev. Drug Discov.* **2017**, *16*, 101–114.
- (89) An, S.; Fu, L. *EBioMedicine* **2018**, *36*, 553–562.
- (90) Pei, H.; Peng, Y.; Zhao, Q.; Chen, Y. *RSC Adv.* **2019**, *9*, 16967–16976.
- (91) Hughes, S. J.; Ciulli, A. *Essays Biochem.* **2017**, *61*, 505–516.
- (92) Zhang, Y.; Loh, C.; Chen, J.; Mainolfi, N. *Drug Discov. Today Technol.* **2019**, *31*, 53–60.
- (93) Sakamoto, K. M.; Kim, K. B.; Kumagai, A.; Mercurio, F.; Crews, C. M.; Deshaies, R. J. *Proc. Natl. Acad. Sci.* **2001**, *98*, 8554–8559.
- (94) Roy Choudhury, S.; Karmakar, S.; Banik, N. L.; Ray, S. K. *J. Oncol.* **2012**, 1–15.
- (95) Schneekloth, A. R.; Pucheault, M.; Tae, H. S.; Crews, C. M. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5904–5908.
- (96) Stathis, A.; Bertoni, F. *Cancer Discov.* **2018**, *8*, 24–36.
- (97) Klein, K. *RMD Open* **2018**, *4*, 1-10.
- (98) Franklin, R. E.; Gosling, R. G. *Nature* **1953**, *171*, 740–741.
- (99) Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. *Nature* **1953**, *171*, 738–740.

- (100) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737–738.
- (101) Structure of DNA
<http://slideplayer.com/slide/6291031/> (accessed Jul 29, 2019).
- (102) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 964–967.
- (103) Alberts, B.; Johnson, A.; Lewis, J. *Molecular Biology of the Cell*, 4th ed. Garland Science: New York, 2002.
- (104) Knorre, D. G.; Kudryashova, N. V.; Godovikova, T. S. *Acta Naturae* **2009**, *1*, 29–51.
- (105) Ribosomal gene translation
http://upendrats.blogspot.com/2013_08_01_archive.html (accessed Jul 29, 2019).
- (106) Holliday, R. *Science* **1987**, *238*, 163–170.
- (107) Arrowsmith, C. H.; Bountra, C.; Fish, P. V.; Lee, K.; Schapira, M. *Nat. Rev. Drug Discov.* **2012**, *11*, 384–400.
- (108) Kornberg, R. D. *Science* **1974**, *184*, 868–871.
- (109) The structure of chromatin within chromosomes
<http://www.nature.com/scitable/topicpage/DNA-Packaging-Nucleosomes-and-Chromatin-310> (accessed Jul 29, 2019).
- (110) Sanchez, R. *Curr. Opin. Drug Discov. Devel.* **2009**, *12*, 659–665.
- (111) Gardner, K. E.; Allis, C. D.; Strahl, B. D. *J. Mol. Biol.* **2011**, *409*, 36–46.
- (112) Grunstein, M. *Nature* **1997**, *389*, 349–352.
- (113) Hewings, D. S.; Rooney, T. P. C.; Jennings, L. E.; Hay, D. A.; Schofield, C. J.; Brennan, P. E.; Knapp, S.; Conway, S. J. *J. Med. Chem.* **2012**, *55*, 9393–9413.
- (114) Writers, Readers and Erasers
<https://epiphanyasd.blogspot.com/2014/01/epigenetics-and-autism.html>
(accessed Jul 29, 2019).
- (115) Conway, S. J. *ACS Med. Chem. Lett.* **2012**, *3*, 691–694.
- (116) Sanchez, R.; Zhou, M.-M. *Curr. Opin. Drug Discov. Devel.* **2009**, *12*, 659–665.
- (117) Chung, C.; Dean, A. W.; Woolven, J. M.; Bamborough, P. *J. Med. Chem.* **2012**, *55*, 576–586.
- (118) Dhalluin, C.; Carlson, J. E.; Zeng, L.; He, C.; Aggarwal, A. K.; Zhou, M.-M.; Zhou, M.-M. *Nature* **1999**, *399*, 491–496.

- (119) Philpott, M.; Yang, J.; Tumber, T.; Fedorov, O.; Uttarkar, S.; Filippakopoulos, P.; Picaud, S.; Keates, T.; Felletar, I.; Ciulli, A. *Mol. Biosyst.* **2011**, *7*, 2899–2909.
- (120) Haynes, S. R.; Dollard, C.; Winston, F.; Beck, S.; Trowsdale, J.; Dawid, I. B. *Nucleic Acids Res.* **1992**, *20*, 2603–2603.
- (121) Wyce, A.; Degenhardt, Y.; Bai, Y.; Le, B.; Korenchuk, S.; Crouthamel, M.-C.; McHugh, C.; Vessella, R.; Creasy, C.; Tummino, P. *Oncotarget* **2013**, *4*, 2419–2429.
- (122) Jahagirdar, R.; Attwell, S.; Marusic, S.; Bendele, A.; Shenoy, N.; McLure, K. G.; Gilham, D.; Norek, K.; Hansen, H. C.; Yu, R. *Mol. Pharmacol.* **2017**, *92*, 694–706.
- (123) Nicholas, D. A.; Andrieu, G.; Strissel, K. J.; Nikolajczyk, B. S.; Denis, G. V. *Cell. Mol. Life Sci.* **2017**, *74*, 231–243.
- (124) Delmore, J. E.; Issa, G. C.; Lemieux, M. E.; Rahl, P. B.; Shi, J.; Jacobs, H. M.; Kastritis, E.; Gilpatrick, T.; Paranal, R. M.; Qi, J. *Cell* **2011**, *146*, 904–917.
- (125) Zou, Z.; Huang, B.; Wu, X.; Zhang, H.; Qi, J.; Bradner, J.; Nair, S.; Chen, L.-F. *Oncogene* **2014**, *33*, 2395–2404.
- (126) Chapuy, B.; McKeown, M. R.; Lin, C. Y.; Monti, S.; Roemer, M. G. M.; Qi, J.; Rahl, P. B.; Sun, H. H.; Yeda, K. T.; Doench, J. G. *Cancer Cell* **2013**, *24*, 777–790.
- (127) Belkina, A. C.; Denis, G. V. *Nat. Rev. Cancer* **2012**, *12*, 465–477.
- (128) Matzuk, M. M.; McKeown, M. R.; Filippakopoulos, P.; Li, Q.; Ma, L.; Agno, J. E.; Lemieux, M. E.; Picaud, S.; Yu, R. N.; Qi, J. *Cell* **2012**, *150*, 673–684.
- (129) Shi, J.; Vakoc, C. R. *Mol. Cell* **2014**, *54*, 728–736.
- (130) Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W. B.; Fedorov, O.; Morse, E. M.; Keates, T.; Kapp, S.; Bradner, J. *Nature* **2010**, *468*, 1067–1073.
- (131) Fedorov, O.; Marsden, B.; Pogacic, V.; Rellos, P.; Muller, S.; Bullock, A. N.; Schwaller, J.; Sundstrom, M.; Knapp, S. *Proc. Natl. Acad. Sci.* **2007**, *104*, 20523–20528.
- (132) Bullock, A. N.; Debreczeni, J. É.; Fedorov, O. Y.; Nelson, A.; Marsden, B. D.; Knapp, S. *J. Med. Chem.* **2005**, *48*, 7604–7614.
- (133) Wadhwa, E.; Nicolaidis, T. *Cureus* **2016**, *8*, 620–628.
- (134) Lee, D. U.; Katavolos, P.; Palanisamy, G.; Katewa, A.; Sioson, C.; Corpuz, J.; Pang, J.; DeMent, K.; Choo, E.; Ghilardi, N. *Toxicol. Appl. Pharmacol.* **2016**, *300*, 47–54.

- (135) Lamonica, J. M.; Deng, W.; Kadauke, S.; Campbell, A. E.; Gamsjaeger, R.; Wang, H.; Cheng, Y.; Billin, A. N.; Hardison, R. C.; Mackay, J. P. *Proc. Natl. Acad. Sci.* **2011**, *108*, 159–168.
- (136) Duan, Y.; Guan, Y.; Qin, W.; Zhai, X.; Yu, B.; Liu, H. *Med. Chem. Comm.* **2018**, *9*, 1779–1802.
- (137) McKeown, M. R.; Shaw, D. L.; Fu, H.; Liu, S.; Xu, X.; Marineau, J. J.; Huang, Y.; Zhang, X.; Buckley, D. L.; Kadam, A. *J. Med. Chem.* **2014**, *57*, 9019–9027.
- (138) Engelhardt, H.; Martin, L.; Smethurst, C. US2015051208A, 2015.
- (139) Gosmini, R.; Nguyen, V. L.; Toum, J.; Simon, C.; Brusq, J.-M. G.; Krysa, G.; Mirguet, O.; Riou-Eymard, A. M.; Boursier, E. V.; Trottet, L. *J. Med. Chem.* **2014**, *57*, 8111–8131.
- (140) Shimamura, T.; Chen, Z.; Soucheray, M.; Carretero, J.; Kikuchi, E.; Tchaicha, J. H.; Gao, Y.; Cheng, K. A.; Cohoon, T. J.; Qi, J. *Clin. Cancer Res.* **2013**, *19*, 6183–6192.
- (141) Lu, J.; Qian, Y.; Altieri, M.; Dong, H.; Wang, J.; Raina, K.; Hines, J.; Winkler, J. D.; Crew, A. P.; Coleman, K.; Crews, C. M. *Chem. Biol.* **2015**, *22*, 755–763.
- (142) Zengerle, M.; Chan, K.-H.; Ciulli, A. *ACS Chem. Biol.* **2015**, *10*, 1770–1777.
- (143) Gadd, M. S.; Testa, A.; Lucas, X.; Chan, K.-H.; Chen, W.; Lamont, D. J.; Zengerle, M.; Ciulli, A. *Nat. Chem. Biol.* **2017**, *13*, 514–521.
- (144) Winter, G. E.; Buckley, D. L.; Paulk, J.; Roberts, J. M.; Souza, A.; Dhe-Paganon, S.; Bradner, J. E. *Science* **2015**, *348*, 1376–1381.
- (145) Zhou, B.; Hu, J.; Xu, F.; Chen, Z.; Bai, L.; Fernandez-Salas, E.; Lin, M.; Liu, L.; Yang, C.-Y.; Zhao, Y. *J. Med. Chem.* **2018**, *61*, 462–481.
- (146) Nowak, R. P.; DeAngelo, S. L.; Buckley, D.; He, Z.; Donovan, K. A.; An, J.; Safaei, N.; Jedrychowski, M. P.; Ponthier, C. M.; Ishoey, M. *Nat. Chem. Biol.* **2018**, *14*, 706–714.
- (147) Bondeson, D. P.; Mares, A.; Smith, I. E. D.; Ko, E.; Campos, S.; Miah, A. H.; Mulholland, K. E.; Routly, N.; Buckley, D. L.; Gustafson, J. L. *Nat. Chem. Biol.* **2015**, *11*, 611–617.
- (148) Bondeson, D. P.; Smith, B. E.; Burslem, G. M.; Buhimschi, A. D.; Hines, J.; Jaime-Figueroa, S.; Wang, J.; Hamman, B. D.; Ishchenko, A.; Crews, C. M. *Cell Chem. Biol.* **2018**, *25*, 78–87.
- (149) Tinworth, C. P.; Lithgow, H.; Dittus, L.; Bassi, Z. I.; Hughes, S. E.; Muelbaier, M.; Dai, H.; Smith, I. E. D.; Kerr, W. J.; Burley, G. A. *ACS Chem. Biol.* **2019**, *14*, 342–347.

- (150) Schiedel, M.; Herp, D.; Hammelmann, S.; Swyter, S.; Lehotzky, A.; Robaa, D.; Oláh, J.; Ovádi, J.; Sippl, W.; Jung, M. *J. Med. Chem.* **2018**, *61*, 482–491.
- (151) Qiu, X.; Sun, N.; Kong, Y.; Li, Y.; Yang, X.; Jiang, B. *Org. Lett.* **2019**, *21*, 3838–3841.
- (152) Wurz, R. P.; Dellamaggiore, K.; Dou, H.; Javier, N.; Lo, M.-C.; McCarter, J. D.; Mohl, D.; Sastri, C.; Lipford, J. R.; Cee, V. J. *J. Med. Chem.* **2018**, *61*, 453–461.
- (153) Lebraud, H.; Wright, D. J.; Johnson, C. N.; Heightman, T. D. *ACS Cent. Sci.* **2016**, *2*, 927–934.
- (154) Rodríguez, N.; Goossen, L. J. *Chem. Soc. Rev.* **2011**, *40*, 5030–5048.
- (155) Wang, J.; Qin, T.; Chen, T.-G.; Wimmer, L.; Edwards, J. T.; Cornella, J.; Vokits, B.; Shaw, S. A.; Baran, P. S. *Angew. Chem. Int. Ed.* **2016**, *55*, 9676–9679.
- (156) Sandfort, F.; O'Neill, M. J.; Cornella, J.; Wimmer, L.; Baran, P. S. *Angew. Chem. Int. Ed.* **2017**, *56*, 3319–3323.
- (157) Huihui, K. M. M.; Caputo, J. A.; Melchor, Z.; Olivares, A. M.; Spiewak, A. M.; Johnson, K. A.; DiBenedetto, T. A.; Kim, S.; Ackerman, L. K. G.; Weix, D. J. *J. Am. Chem. Soc.* **2016**, *138*, 5016–5019.
- (158) Baker, B. The Application of Flow Chemistry to the Synthesis of Pharmaceutically Privileged Heterocycles, Ph.D. Thesis, University of Strathclyde, 2019.
- (159) Blackburn, C.; Guan, B.; Fleming, P.; Shiosaki, K.; Tsai, S. *Tetrahedron Lett.* **1998**, *39*, 3635–3638.
- (160) Cox, P. A.; Reid, M.; Leach, A. G.; Campbell, A. D.; King, E. J.; Lloyd-Jones, G. C. *J. Am. Chem. Soc.* **2017**, *139*, 13156–13165.
- (161) Patel, N. D.; Zhang, Y.; Gao, J.; Sidhu, K.; Lorenz, J. C.; Fandrick, K. R.; Mulder, J. A.; Herbage, M. A.; Li, Z.; Ma, S. *Org. Process Res. Dev.* **2016**, *20*, 95–99.
- (162) Eliel, E. L.; Wilen, S. H.; Mander, L. N. *Stereochemistry of Organic Compounds*, 1st ed. Wiley: New York, 1994.
- (163) Wellaway, C. R.; Amans, D.; Bamborough, P.; Barnett, H.; Bit, R. A.; Brown, J. A.; Carlson, N. R.; Chung, C.; Cooper, A. W. J.; Craggs, P. D. *J. Med. Chem.* **2020**, *63*, 714–746.
- (164) Ding, Y.; Chai, J.; Centrella, P. A.; Gondo, C.; DeLorey, J. L.; Clark, M. A. *ACS Comb. Sci.* **2018**, *20*, 251–255.
- (165) Yang, D.; Fokas, D.; Li, J.; Yu, L.; Baldino, C. *Synthesis* **2005**, *1*, 47–56.

- (166) Burlamacchi, L.; Guarini, G.; Tiezzi, E. *Trans. Faraday Soc.* **1969**, *65*, 496.
- (167) Biotage® Initiator+ microwave reactor
<https://www.biotage.com/product-page/biotage-initiator> (accessed Jan 6, 2020).
- (168) Roy, C. D.; Brown, H. C. *J. Organomet. Chem.* **2007**, *692*, 784–790.
- (169) Krasovskiy, A.; Knochel, P. *Angew. Chem. Int. Ed.* **2004**, *43*, 3333–3336.
- (170) Li-Yuan Bao, R.; Zhao, R.; Shi, L. *Chem. Commun.* **2015**, *51*, 6884–6900.
- (171) Fukuzaki, E. US9062084B2, 2015.
- (172) Brown, D. G.; Boström, J. *J. Med. Chem.* **2016**, *59*, 4443–4458.
- (173) Valeur, E.; Bradley, M. *Chem. Soc. Rev.* **2009**, *38*, 606–631.
- (174) Due-Hansen, M. E.; Pandey, S. K.; Christiansen, E.; Andersen, R.; Hansen, S. V. F.; Ulven, T. *Org. Biomol. Chem.* **2016**, *14*, 430–433.
- (175) Green, R. A.; Hartwig, J. F. *Org. Lett.* **2014**, *16*, 4388–4391.
- (176) Fujiwara, H.; Sato, K.; Mizumoto, S.; Sato, Y.; Kurihara, H.; Kubo, Y.; Nakata, H.; Baba, Y.; Tamura, T.; Kuniyoshi, H.; Hagiwara, S.; Yamamoto, M. WO2012002577, 2012.
- (177) Lee, S.; Jørgensen, M.; Hartwig, J. F. *Org. Lett.* **2001**, *3*, 2729–2732.
- (178) Crawford, S. M.; Lavery, C. B.; Stradiotto, M. *Chem. Eur. J.* **2013**, *19*, 16760–16771.
- (179) MacMillan, D. S.; Murray, J.; Sneddon, H. F.; Jamieson, C.; Watson, A. J. B. *Green Chem.* **2013**, *15*, 596.
- (180) Bordwell, F. G.; Fried, H. E. *J. Org. Chem.* **1981**, *46*, 4327–4331.
- (181) Olmstead, W. N.; Margolin, Z.; Bordwell, F. G. *J. Org. Chem.* **1980**, *45*, 3295–3299.
- (182) Young, R. J. *Tactics in Contemporary Drug Design*, 1st ed. Springer: Berlin, 2014.
- (183) Savjani, K. T.; Gajjar, A. K.; Savjani, J. K. *ISRN Pharm.* **2012**, 195727–195737.
- (184) Arnott, J. A.; Planey, S. L. *Expert Opin. Drug Discov.* **2012**, *7*, 863–875.
- (185) Fujita, T.; Iwasa, J.; Hansch, C. *J. Am. Chem. Soc.* **1964**, *86*, 5175–5180.
- (186) Young, R. J.; Green, D. V. S.; Luscombe, C. N.; Hill, A. P. *Drug Discov. Today* **2011**, *16*, 822–830.
- (187) Dobson, P. D.; Kell, D. B. *Nat. Rev. Drug Discov.* **2008**, *7*, 205–220.

- (188) Kell, D. B.; Dobson, P. D.; Oliver, S. G. *Drug Discov. Today* **2011**, *16*, 704–714.
- (189) Irvine, J. D.; Takahashi, L.; Lockhart, K.; Cheong, J.; Tolan, J. W.; Selick, H. E.; Grove, J. R. *J. Pharm. Sci.* **1999**, *88*, 28–33.
- (190) Bayliss, M. K.; Butler, J.; Feldman, P. L.; Green, D. V. S.; Leeson, P. D.; Palovich, M. R.; Taylor, A. J. *Drug Discov. Today* **2016**, *21*, 1719–1727.
- (191) Valko, K.; Butler, J.; Eddershaw, P. *Expert Opin. Drug Discov.* **2013**, *8*, 1225–1238.
- (192) Pye, C. R.; Hewitt, W. M.; Schwochert, J.; Haddad, T. D.; Townsend, C. E.; Etienne, L.; Lao, Y.; Limberakis, C.; Furukawa, A.; Mathiowetz, A. M. *J. Med. Chem.* **2017**, *60*, 1665–1672.
- (193) Ertl, P.; Rohde, B.; Selzer, P. *J. Med. Chem.* **2000**, *43*, 3714–3717.
- (194) Valkó, K. L. *J. Pharm. Biomed. Anal.* **2016**, *130*, 35–54.
- (195) Ritchie, T. J.; Macdonald, S. J. F. *Drug Discov. Today* **2009**, *14*, 1011–1020.
- (196) Lovering, F.; Bikker, J.; Humblet, C. *J. Med. Chem.* **2009**, *52*, 6752–6756.
- (197) Rossi Sebastiano, M.; Doak, B. C.; Backlund, M.; Poongavanam, V.; Over, B.; Ermondi, G.; Caron, G.; Matsson, P.; Kihlberg, J. *J. Med. Chem.* **2018**, *61*, 4189–4202.
- (198) Yang, F.; Zhang, Y.; Liang, H. *Int. J. Mol. Sci.* **2014**, *15*, 3580–3595.
- (199) van der Veen, J. N.; Kennelly, J. P.; Wan, S.; Vance, J. E.; Vance, D. E.; Jacobs, R. L. *Biochim. Biophys. Acta - Biomembr.* **2017**, *1859*, 1558–1572.
- (200) Valko, K. L.; Teague, S. P.; Pidgeon, C. *ADMET DMPK* **2017**, *5*, 14–38.
- (201) Bunally, S.; Young, R. J. *ADMET DMPK* **2018**, *6*, 74–84.
- (202) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3–25.
- (203) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. *J. Med. Chem.* **2002**, *45*, 2615–2623.
- (204) DeGoey, D. A.; Chen, H.-J.; Cox, P. B.; Wendt, M. D. *J. Med. Chem.* **2018**, *61*, 2636–2651.
- (205) Doak, B. C.; Over, B.; Giordanetto, F.; Kihlberg, J. *Chem. Biol.* **2014**, *21*, 1115–1142.
- (206) Edmondson, S. D.; Yang, B.; Fallan, C. *Bioorg. Med. Chem. Lett.* **2019**, *29*, 1555–1564.

- (207) Riching, K. M.; Mahan, S.; Corona, C. R.; McDougall, M.; Vasta, J. D.; Robers, M. B.; Urh, M.; Daniels, D. L. *ACS Chem. Biol.* **2018**, *13*, 2758–2770.
- (208) Banerjee, S. S.; Aher, N.; Patil, R.; Khandare, J. J. *Drug Deliv.* **2012**, 1–17.
- (209) Howard, D. L.; Kjaergaard, H. G. *Phys. Chem. Chem. Phys.* **2008**, *10*, 4113.
- (210) Hao, M. H. *J. Chem. Theory Comput.* **2006**, *2*, 863–872.
- (211) Bordwell, F. G.; Drucker, G. E.; Fried, H. E. *J. Org. Chem.* **1981**, *46*, 632–635.
- (212) Degorce, S. L.; Bodnarchuk, M. S.; Cumming, I. A.; Scott, J. S. *J. Med. Chem.* **2018**, *61*, 8934–8943.
- (213) Bard, J. A. M.; Goodall, E. A.; Greene, E. R.; Jonsson, E.; Dong, K. C.; Martin, A. *Annu. Rev. Biochem.* **2018**, *87*, 697–724.
- (214) cChromlogD MPNN
<https://socrates.gsk.com/search/elc/collection/prd-gsk/9329861> (accessed Jan 6, 2020).
- (215) Huang, Z.; Ung, T. *Curr. Drug Metab.* **2013**, *14*, 226–238.
- (216) DS Biovia QSAR Workbench
http://help.accelrys.com/qsarwb/2018/qsarwb.htm#qsarwb_user/introduction.htm (accessed Jan 6, 2020).
- (217) Nobeli, I.; Price, S. L.; Lommerse, J. P. M.; Taylor, R. *J. Comput. Chem.* **1997**, *18*, 2060–2074.
- (218) Sauer, W. H. B.; Schwarz, M. K. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 987–1003.
- (219) Wirth, M.; Volkamer, A.; Zoete, V.; Rippmann, F.; Michielin, O.; Rarey, M.; Sauer, W. H. B. *J. Comput. Aided. Mol. Des.* **2013**, *27*, 511–524.
- (220) Grygorenko, O. O.; Prytulyak, R.; Volochnyuk, D. M.; Kudrya, V.; Khavryuchenko, O. V.; Komarov, I. V. *Mol. Divers.* **2012**, *16*, 477–487.
- (221) Fall, Y.; Reynaud, C.; Doucet, H.; Santelli, M. *Eur. J. Org. Chem.* **2009**, 4041–4050.
- (222) Shing, K.-P.; Liu, Y.; Cao, B.; Chang, X.-Y.; You, T.; Che, C.-M. *Angew. Chem. Int. Ed.* **2018**, *57*, 11947–11951.
- (223) Santra, S.; Guin, J. *Eur. J. Org. Chem.* **2015**, 7253–7257.
- (224) Skillinghaug, B.; Sköld, C.; Rydfjord, J.; Svensson, F.; Behrends, M.; Sävmarker, J.; Sjöberg, P. J. R.; Larhed, M. *J. Org. Chem.* **2014**, *79*, 12018–12032.

- (225) Colletti, L. M.; Liu, Y.; Koev, G.; Richardson, P. L.; Chen, C.-M.; Kati, W. *Anal. Biochem.* **2008**, 383, 186–193.
- (226) Robinson, M. W.; Hill, A. P.; Readshaw, S. A.; Hollerton, J. C.; Upton, R. J.; Lynn, S. M.; Besley, S. C.; Boughtflower, B. J. *Anal. Chem.* **2017**, 89, 1772–1777.
- (227) Valkó, K.; Bevan, C.; Reynolds, D. *Anal. Chem.* **1997**, 69, 2022–2029.
- (228) Sou, T.; Bergström, C. A. S. *Drug Discov. Today Technol.* **2018**, 27, 11–19.
- (229) Wyllie, S.; Brand, S.; Thomas, M.; De Rycker, M.; Chung, C.; Pena, I.; Bingham, R. P.; Bueren-Calabuig, J. A.; Cantizani, J.; Cebrian, D. *Proc. Natl. Acad. Sci.* **2019**, 116, 9318–9323.
- (230) Valko, K.; Nunhuck, S.; Bevan, C.; Abraham, M. H.; Reynolds, D. P. *J. Pharm. Sci.* **2003**, 92, 2236–2248.
- (231) Valko, K.; Du, C. M.; Bevan, C. D.; Reynolds, D. P.; Abraham, M. H. *J. Pharm. Sci.* **2000**, 89, 1085–1096.
- (232) PCv6 Descriptors
<http://ccwiki.gsk.com/twiki/bin/view/CSC/StandardProperties#gskneg>
 (accessed Jan 6, 2020).
- (233) Andrews, P. R.; Craik, D. J.; Martin, J. L. *J. Med. Chem.* **1984**, 27, 1648–1657.
- (234) McGowan, J. C.; Sowada, R. *J. Chem. Technol. Biotechnol.* **2007**, 58, 357–361.
- (235) Platts, J. A.; Butina, D.; Abraham, M. H.; Hersey, A. *J. Chem. Inf. Comput. Sci.* **1999**, 39, 835–845.
- (236) Kier, L. B.; Hall, L. H. *Pharm. Res.* **1990**, 7, 801–807.
- (237) Kier, L. B.; Hall, L. H.; Frazer, J. W. *J. Math. Chem.* **1991**, 7, 229–241.
- (238) Labute, P. *J. Chem. Inf. Model.* **2010**, 50, 792–800.

7: Appendices

7.1 NOE Experiments

7.1.1 5-(6-Bromo-1-(1,3-dimethoxypropan-2-yl)-1H-benzo[d]imidazol-2-yl)-6-isopropyl-1,3-dimethyl-3,6-dihydropyridin-2(1H)-one **121**

Irradiation of H^a of benzimidazole **121** resulted in an NOE enhancement of CH^b₃, H^c and H^d, clearly demonstrating their proximity in space, **Figure 117**.

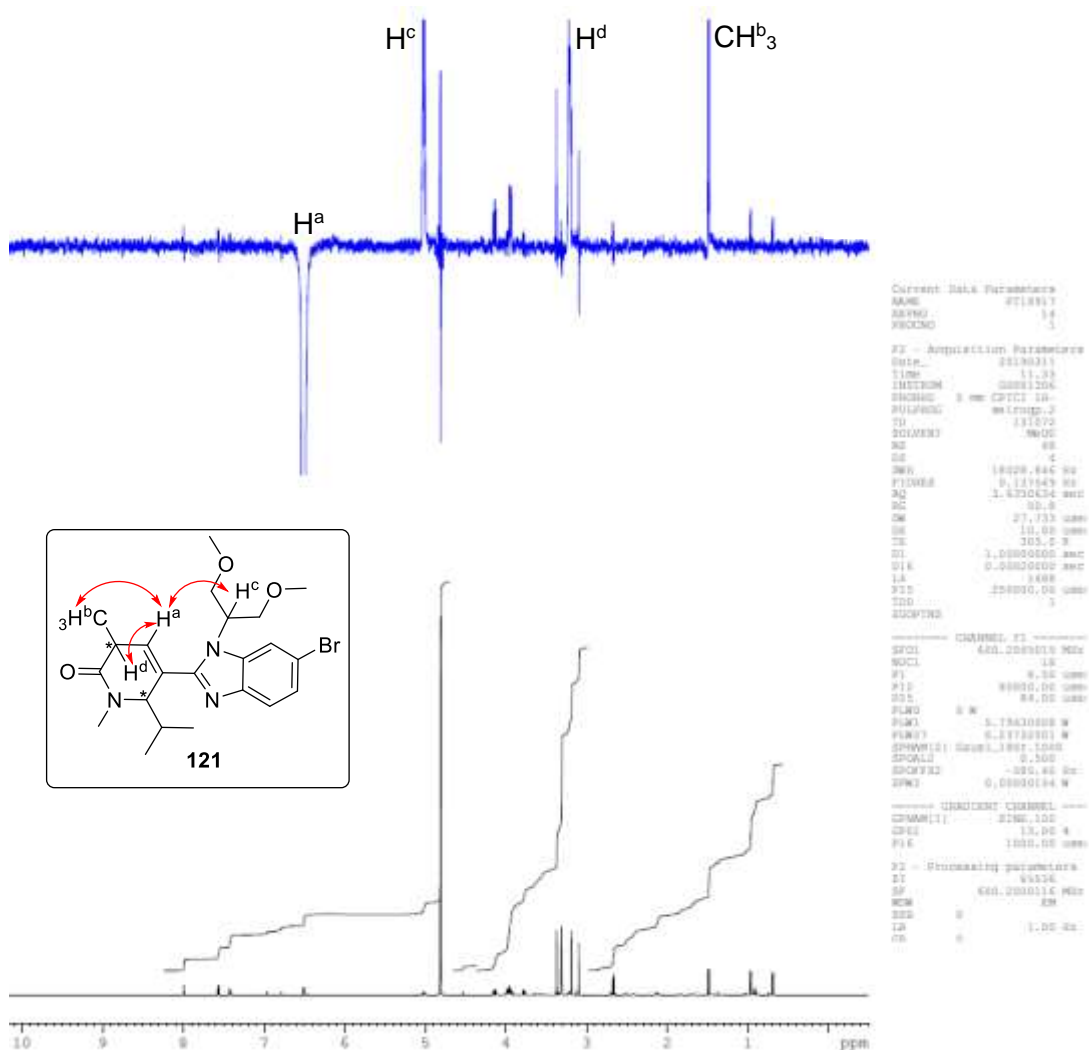


Figure 117: NOE experiment irradiating H^a of benzimidazole **121**. Observed NOE correlations are highlighted by red double-headed arrows.

Irradiation of CH^e_3 of benzimidazole **121** resulted in an NOE enhancement of CH^f_3 , H^g and H^i , clearly demonstrating their proximity in space, thus supporting the assignment of the regiochemistry of Grignard addition, **Figure 118**.

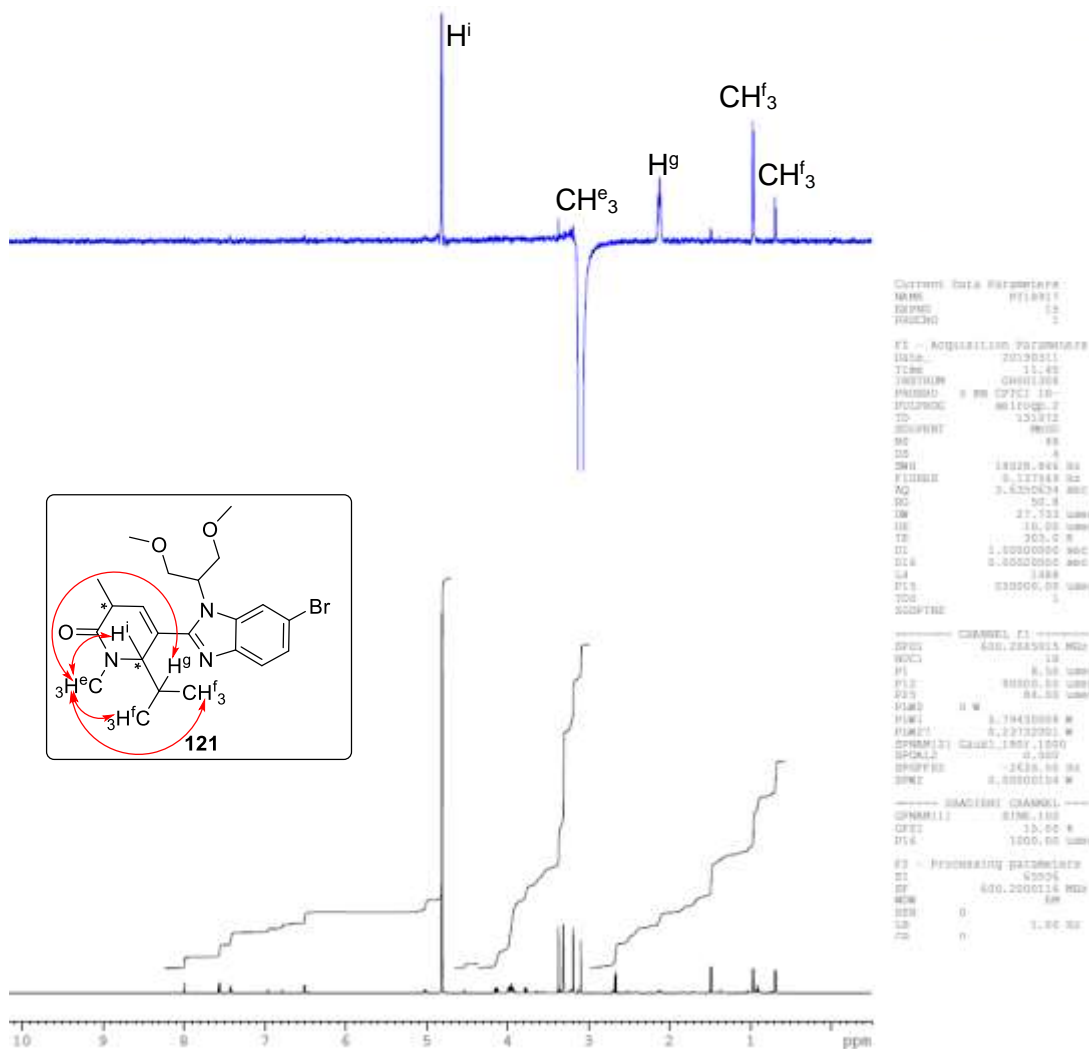


Figure 118: NOE experiment irradiating CH^e_3 for benzimidazole **121**. Observed NOE correlations are highlighted by red double-headed arrows.

7.1.2 6-Butyl-5-(1-(1,3-dimethoxypropan-2-yl)-1H-benzo[d]imidazol-2-yl)-1,3-dimethyl-3,6-dihydropyridin-2(1H)-one 122

Irradiation of H^a of benzimidazole **122** resulted in an NOE enhancement of CH^b₃ and H^c, clearly demonstrating their proximity in space, **Figure 119**.

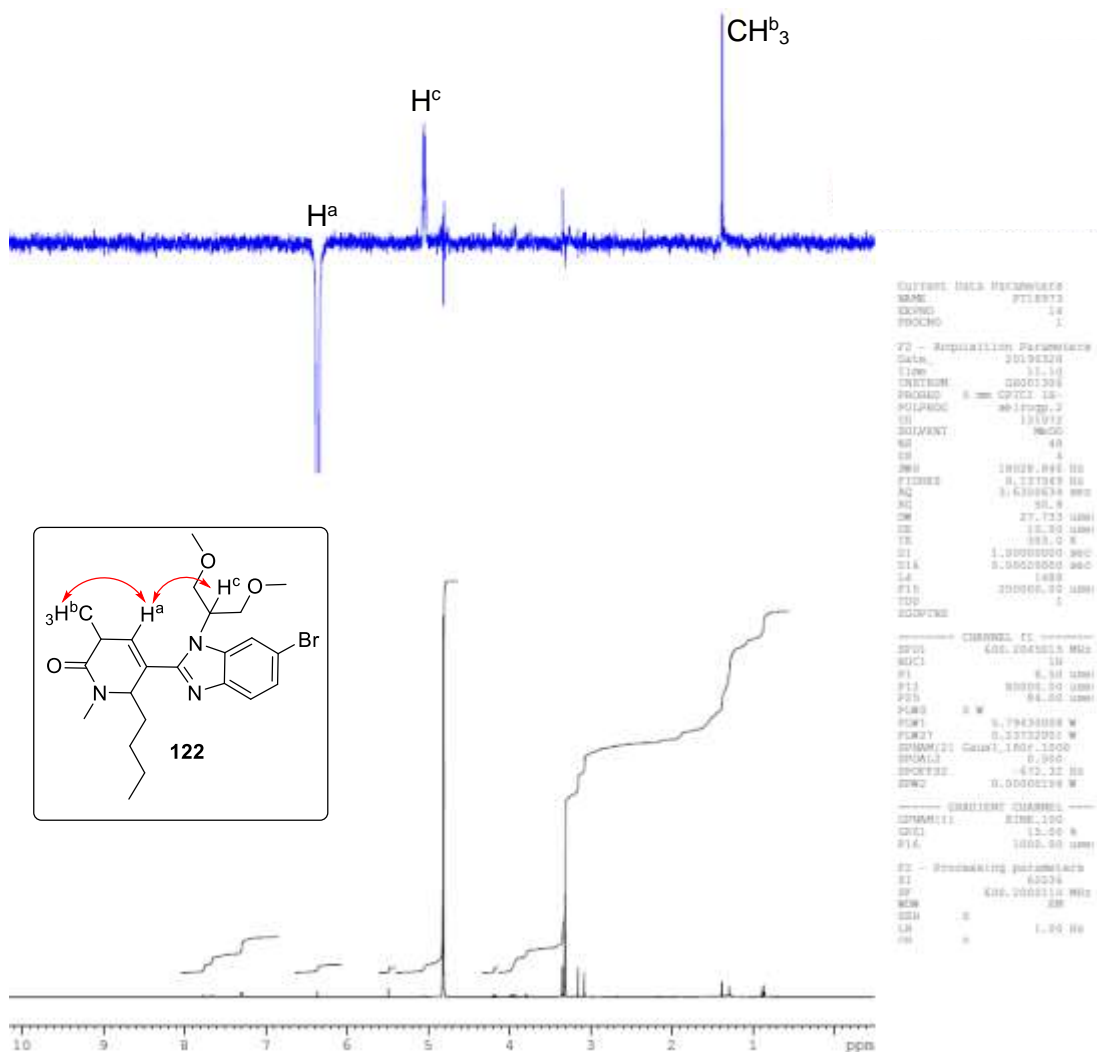


Figure 119: NOE experiment irradiating H^a of benzimidazole **122**. Observed NOE correlations are highlighted by red double-headed arrows.

Irradiation of CH^{d} of benzimidazole **122** resulted in an NOE enhancement of CH^{e} and H^{f} , clearly demonstrating their proximity in space, thus supporting the assignment of the regiochemistry of organolithium addition, **Figure 120**.

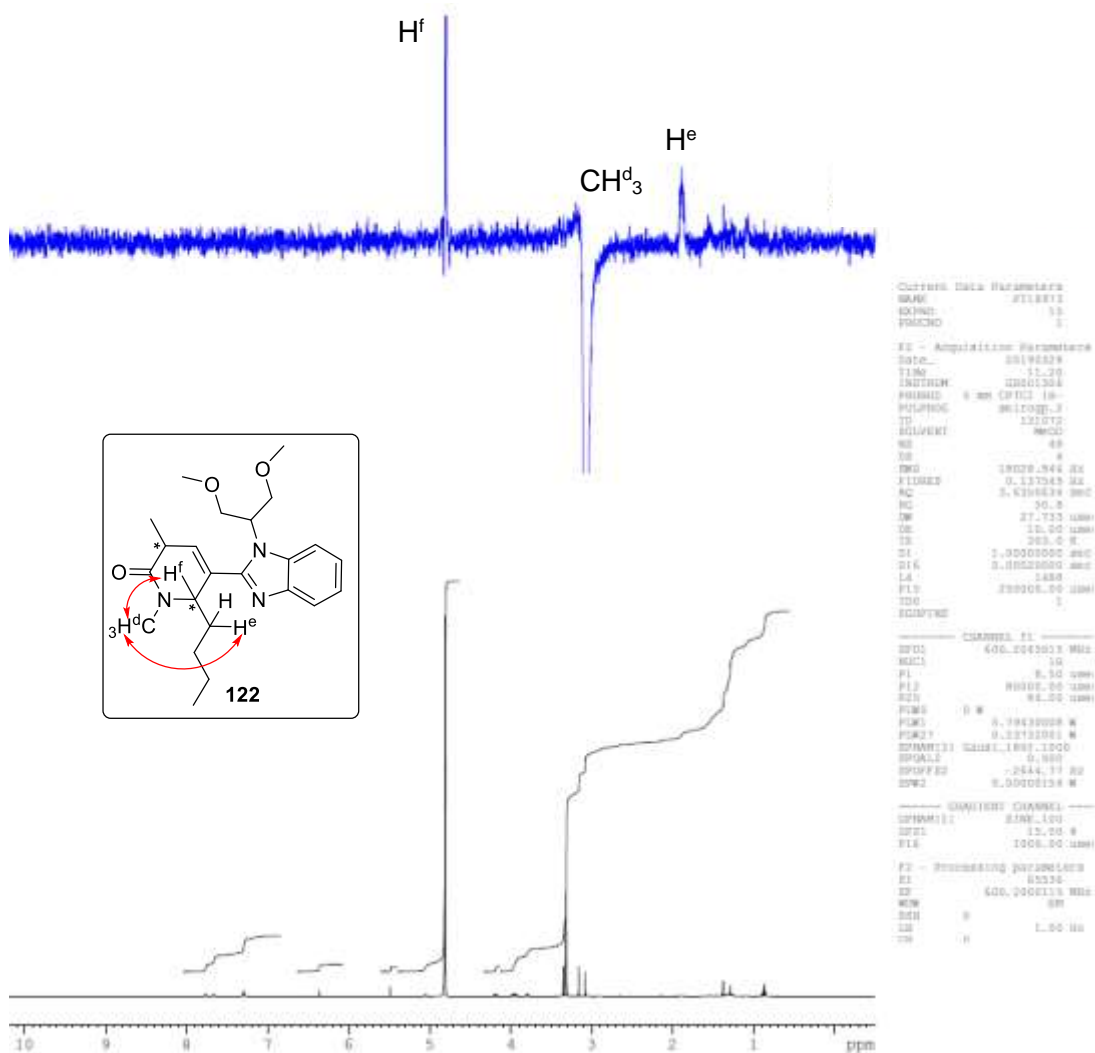


Figure 120: NOE experiment irradiating CH^{d} of benzimidazole **122**. Observed NOE correlations are highlighted by red double-headed arrows.

7.2 X-ray Cocrystal Structures

7.2.1 (R)-N4-(1-(1,3-dimethoxypropan-2-yl)-2-(1,5-dimethyl-6-oxo-1,6-dihydropyridin-3-yl)-1H-benzo[d]imidazol-6-yl)-N1-((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-2-methylsuccinamide 168 in human Brd4 BD1 protein

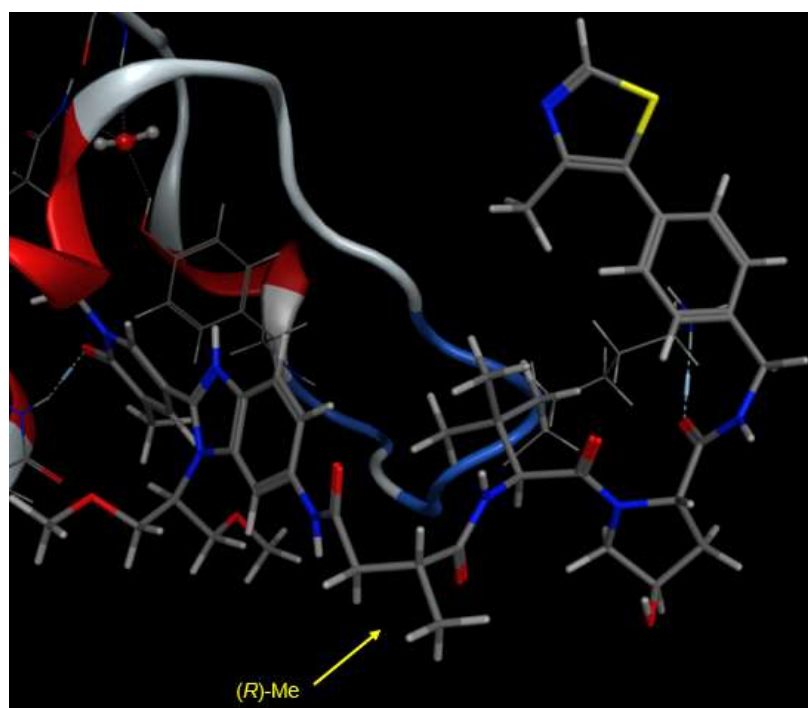


Figure 121: X-ray cocrystal structure of PROTAC 168 in human Brd4 BD1 protein, highlighting the absolute stereochemistry of the (R)-Me substituted linker.

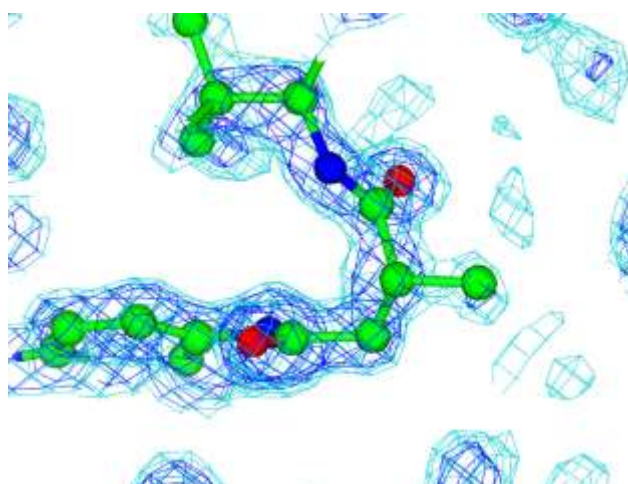


Figure 122: OMIT (2Fo-Fc) map contoured at +1.0 sigma (blue), +0.6 sigma (cyan) showing the clarity of the assignment of the (R)-Me chiral centre in PROTAC 168.

Table 45: Data collection and refinement statistics for the X-ray crystallographic analysis of the cocrystal structure of PROTAC 168 with human Brd4 BDI protein.

Data collection	
Space Group	P2 ₁ 2 ₁ 2 ₁
Cell Dimensions	
a, b, c (Å)	43.357, 52.364, 57.039
α , β , γ (°)	90.000, 90.000, 90.000
Resolution (Å)	50.00-1.24 (1.26-1.24)
R _{merge} *	0.104 (0.238)
CC(1/2)	(0.945)
Average I/ σ I	14.5 (7.7)
Completeness (%)	98.7 (81.6)
Redundancy	4.9 (4.8)
Refinement	
Resolution (Å)	50.00-1.24 (1.26-1.24)
No. Reflections	182878
No. Unique Reflections	37041
R _{work} /R _{free}	0.192/0.221
No. atoms	1404
Protein	1056
Ligand/ion	63/0
Water	285
B-factors	
Protein	16.13
Ligand/ion	17.86/0
Water	39.55
RMS deviations	
Bond lengths (Å)	0.0042
Bond angles (°)	1.171

* = Data for the highest resolution shell are given in parentheses.

7.3 Compound Numbering

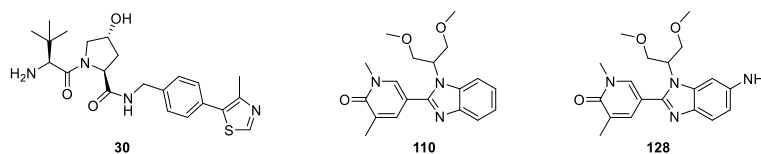


Figure 123: VHL E3 ligase binder **30** and protein binders **110** and **128** used to synthesise the PROTAC series.

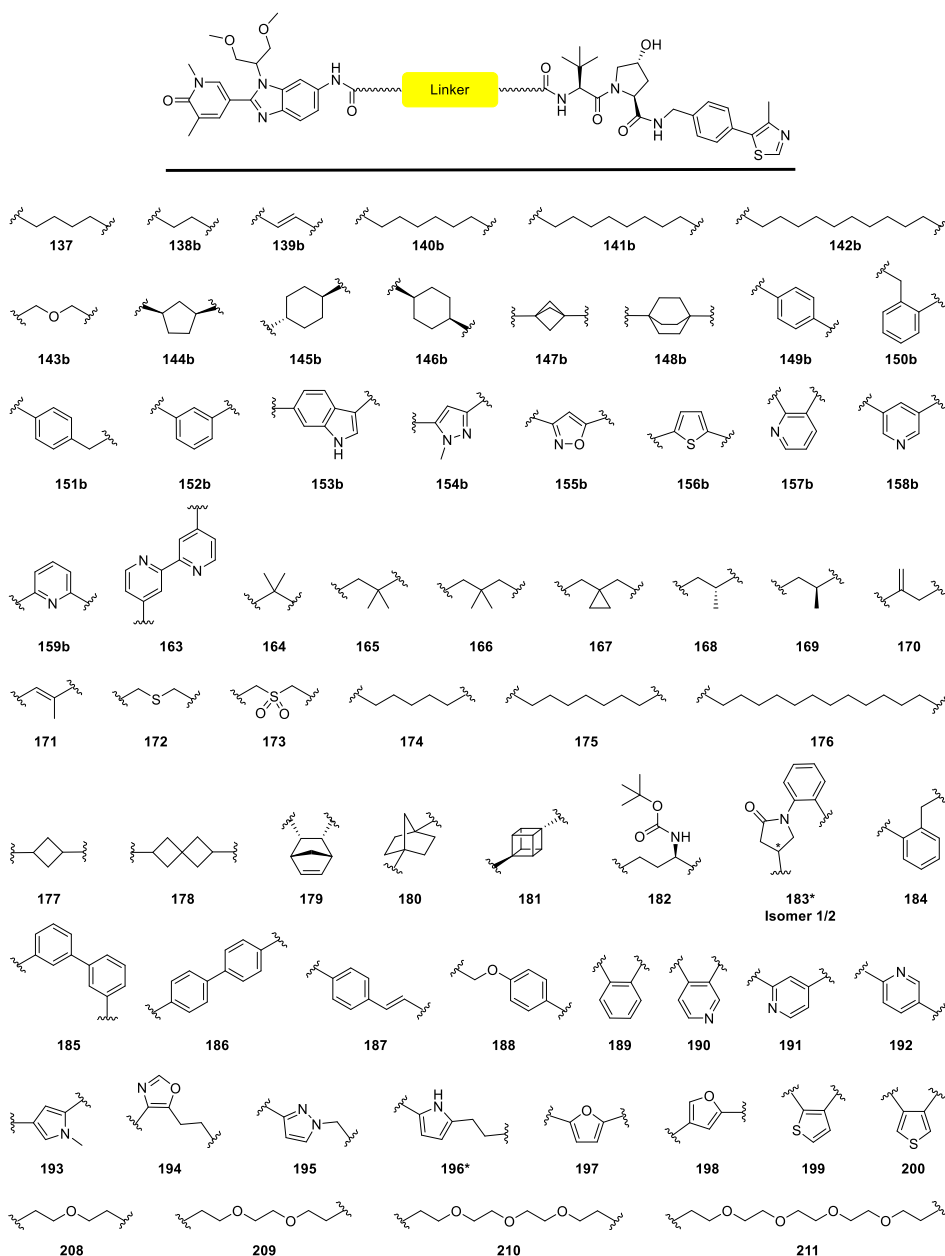


Figure 124: PROTAC series containing 66 PROTACs. * = Omitted from analysis.