

CONFIDENTIAL – Property of GSK – Do Not Copy

University of Strathclyde
Department of Pure and Applied Chemistry

**Use of fragment-based drug discovery to identify
novel drug molecules for challenging biological
targets, including protein-protein interactions.**

Thesis submitted to the University of Strathclyde in fulfilment of the requirements
for the degree of Doctor of Philosophy

Gemma Victoria White

May 2016



Declaration of Copyright

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 the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

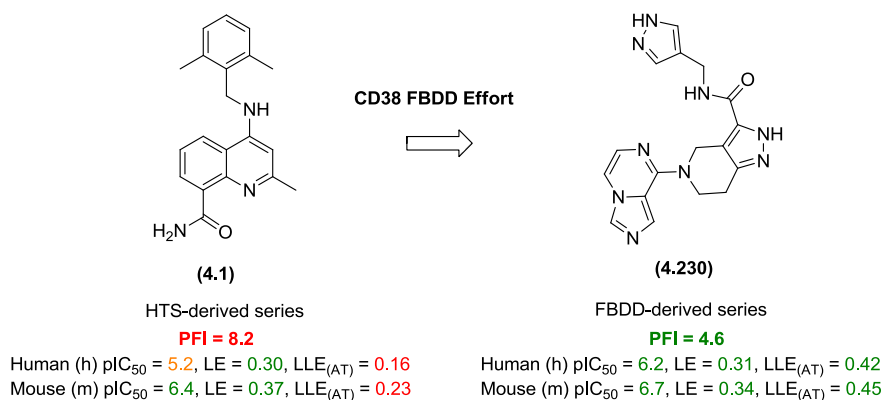
Signed:

Date:

Abstract

This research programme describes two of the first fragment-based drug discovery (FBDD) programmes pursued within our laboratories. In recent years, increased emphasis has been placed on careful control of physicochemical properties in the search for safe and efficacious medicines. FBDD has been proposed as a complementary technique to high throughput screening (HTS) resulting in the identification of small, polar, highly efficient fragments that bind to the target of interest.

The targets pursued within this programme of work each represent a distinct biological challenge: a protein-protein interaction (HIV Integrase-LEDGF) with large surface contacts to disrupt, and a hydrolase enzyme (CD38) with a large open active site. The challenge of optimising the CD38 fragments was heightened due to the inability to obtain X-ray crystal structures of the series of interest. The rigorous FBDD screening, design, and optimisation process is described, resulting in identification of lead molecules with reduced liabilities, and improved physicochemical properties compared to HTS-derived templates. These molecules provide good tools to aid in mode of action studies and future drug development.



Insight from these studies has resulted in recommendations for the most effective approach towards target tractability assessment, screening, and fragment

CONFIDENTIAL – Property of GSK – Do Not Copy

optimisation strategies. Incorporation of the outputs from this programme will have an impact on the ethos and practice of FBDD methodology and how this strategy is implemented into future ‘hit-to-lead’ efforts. It is anticipated that this will, in turn, reduce associated timelines and, therefore, make a valuable addition to such ‘hit-to-lead’ programmes.

Acknowledgements

I would firstly like to thank my academic supervisor Professor William Kerr from the University of Strathclyde for his continuous support and commitment throughout my entire PhD. His passion and enthusiasm for science has really inspired me and I have been really impressed by the meticulous analysis, and caring guidance that was provided for each of his PhD students. I would also like to acknowledge my internal examiner Dr Craig Jamieson for his efforts around reviewing this work and all the helpful comments he provided.

It is important that I thank Dr Ian Baldwin, my GSK supervisor and line manager for a large proportion of my PhD. The support Ian gave me on each of the programmes, the challenging questions that he posed, and the guidance in construction of the report was critical in helping me reach this position. Ian helped me through challenging times, and encouraged me to keep going with a smile on my face.

I would also like to thank Dr Alan Nadin for stepping in at short notice for my 30 month *viva voce*. Without Alan reading my report, understanding my work and being prepared to travel to Strathclyde for my *viva voce*, this couldn't have gone ahead.

Thank you to Dr Vipulkumar Patel for volunteering to review my thesis. I am grateful for his valuable suggestions, interest in my work and his continuous endeavour to help further enhance and broaden my knowledge.

I would also like to acknowledge all the members of the 'Fragment Chemistry Group', both past and present, for their helpful input throughout my PhD. I would particularly like to thank Dr Ian Churcher for giving me the opportunity to lead the HIV IN-LEDGF FBDD programme; Dr Rob Young for his support on the CD38 programme; Dr Lisa Sloan for proof reading sections of my thesis; and Dr Krista Goodman and Tony Dean for their continued support.

CONFIDENTIAL – Property of GSK – Do Not Copy

There are numerous people across GSK (past and present) that I need to thank. These include; Steve Richards, Sean Lynn, Curt Haffner, Dave Becherer, Eugene Stewart, Ian Wall, Denise Polhaus, Lisa Shewchuk-Chapman, Chun-wa Chung, Andrew Leach, Colin Edge, James Woolven, Emile Velthuisen, Andrew Spaltenstein, Jacques Briand, Sean Williams, Heather Barnett, Tony Cooper and Alan Hill; apologies to anyone I have missed!

I would like to say thank you to both Professor Harry Kelly and Professor William Kerr for initiating this collaborative PhD programme between GSK and the University of Strathclyde. This opportunity has proved immensely valuable to me and I believe has been instrumental in shaping me into the scientist that I am today. I would also like to say thank you to Dave Allen, Head of the Respiratory Therapy Area Unit and Chief Chemist and Dr Patrick Vallence, President of GSK Pharmaceutical R&D for their continued support for the programme and paying a special interest in the research that is completed as part of the programme. Thank you also to Andrea Malley for making everything associated with the programme run smoothly.

Finally, I need to say a huge thank you to my family. Mum, Dad and Kerry, thank you for all your support over the past five years. I promise that I will now have more time to attend all the family occasions! I want to say a special thank you to my partner Rob, thank you for being extremely understanding and very, very patient. I realise that we have had to make some sacrifices for me to complete my PhD and I am truly really grateful that you have supported me throughout. I look forward to celebrating with you all!

Abbreviations and Definitions

3P	3' Processing
Ac	Acetate
ACC	Acetyl-CoA carboxylase
ADMET	Absorption, distribution, metabolism, elimination, and toxicology
ADP	Adenosine diphosphate
ADPR	Adenosine diphosphate ribose
AIDS	Acquired immune deficiency syndrome
AMPK	5' Adenosine monophosphate-activated protein kinase
Asp	Aspartic acid
Autosolve™	Automated X-ray diffraction data analysis
Astex	External fragment based drug discovery biotech company
Boc	<i>tert</i> -Butyl carbamate protecting group
°C	Degrees Celsius
cADPR	Cyclic adenosine diphosphate ribose
CCR5	C-C chemokine receptor type 5
CD38	Cluster of differentiation 38
CD4	Cluster of differentiation 4
CHI	Chromatographic hydrophobicity index
Clint	Intrinsic clearance
CLND	Chemi-luminescent nitrogen detection
cLog P	Calculated logarithm of the partition co-efficient P
COPD	Chronic obstructive pulmonary disease
cPFI	Calculated property forecast index
CSD	Cambridge structural database
CXCR4	CXC chemokine receptor type 4
ΔH°	Standard enthalpy of formation
ΔG°	Standard Gibbs free energy
ΔS°	Standard entropy of formation
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene

CONFIDENTIAL – Property of GSK – Do Not Copy

DCM	Dichloromethane
DIPEA	<i>N,N</i> -Diisopropylethylamine
DMEDA	<i>N,N'</i> -Dimethylethylenediamine
DMF	<i>N,N</i> -Dimethylformamide
DMSO	Dimethylsulfoxide
dsDNA	Double stranded DNA
ELT	Encoded library technology
Et	Ethyl
EtOH	Ethanol
EtOAc	Ethyl acetate
F	Fluorescein
FBDD	Fragment-based drug discovery
FDA	Food and Drug Administration
FLINT	Förster/fluorescence intensity
FRET	Förster/fluorescence resonance energy transfer
g	Gram
GDB	A chemical universe database from ‘The Chemical Space Project’
Gln	Glutamine
Glu	Glutamic acid
GPCR	G-protein coupled receptor
GSK	GlaxoSmithKline
h	Hour
H	Hydrogen
HA	Heavy atoms
HAART	Highly active anti-retroviral therapy
HAC	Heavy atom count
HBA	Hydrogen bond acceptor
HBD	Hydrogen bond donor
hERG	Human ether-à-go-go - related gene
His	Histidine
HIV-1	Human immunodeficiency virus type-1

HMBC	Heteronuclear multiple quantum coherence
HRMS	High resolution mass spectrometry
HSP90	Heat shock protein 90
HTS	High throughput screening
IAM	Immobilised artificial membranes
IBD	Integrase binding domain
IMS	Industrial methylated spirit
IN	Integrase
IN _{CCD}	Integrase catalytic core domain
IN _{CTD}	Integrase C-terminal domain
IN _{NTD}	Integrase N-terminal domain
ITC	Isothermal titration calorimetry
K _B	Equilibrium binding constant
K _d	Dissociation equilibrium binding constant
k _{on}	Association rate constant
k _{off}	Dissociation rate constant
LE	Ligand efficiency
LEDGF	Lens epithelium-derived growth factor
LLE	Lipophilic ligand efficiency
LLE _(AT)	Astex lipophilic ligand efficiency
LCMS	Liquid chromatography-mass spectrometry
LKB1	Liver kinase B1
Log D	Logarithm of the partition co-efficient D
Log P	Logarithm of the partition co-efficient P
LOGSY	Ligand observed <i>via</i> gradient spectroscopy
M	Molarity
MDAP	Mass-directed auto preparative chromatography
Me	Methyl
MeCN	Acetonitrile
MeI	Methyl iodide
MeOH	Methanol

CONFIDENTIAL – Property of GSK – Do Not Copy

mg	Milligram
min	Minute
mL	Millilitre
MOA	Mechanism of action
MOE	Molecular Operating Environment
mPFI	Measured property forecast index
MS	Mass spectrometry
MS ESI	Mass spectrometry electrospray ionisation
MW	Molecular weight
NAADP	Nicotinic acid adenine dinucleotide phosphate
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NaHMDS	Sodium hexamethyldisilazide
nM	Nanomolar
NMR	Nuclear magnetic resonance
NNRTI	Non-nucleoside reverse transcriptase inhibitor
NOE	Nuclear Overhauser effect
NRTI	Nucleoside reverse transcriptase inhibitor
PBMC	Peripheral blood mononuclear cells
PDB	Protein data bank
PEI	Percentage efficiency index
PFI	Property forecast index
Ph	Phenyl
PI	Protease inhibitor
Phe	Phenylalanine
PhMe	Toluene
PIC	Pre-integration complex
PMB	<i>para</i> -Methoxybenzyl
PMBCl	<i>para</i> -Methoxybenzyl chloride
PPh ₃	Triphenylphosphine
rb	Rotatable bond

CONFIDENTIAL – Property of GSK – Do Not Copy

R _t	Retention time
RT	Room temperature
RTP	Research Triangle Park
SARs	Structure activity relationships
SD	Standard deviation
sEH	Soluble epoxide hydrolase
SFI	Solubility forecast index
SIRT	Sirtuin enzymes
SOBAX	Commercial compound catalogue
ssRNA	Single stranded RNA
SSS	Substructure search
SPR	Surface plasmon resonance
ST	Strand transfer
STD	Saturation transfer difference
THF	Tetrahydrofuran
T _m	Thermal melt
TMS	Trimethylsilyl
TMSCN	Trimethylsilyl cyanide
TR	Time-resolved
Trp	Tryptophan
Tyr	Tyrosine
μM	Micromolar
μWave	Microwave
UP	Upper Providence
UV	Ultraviolet
Val	Valine
VdW	Van der Waals

Contents

1	<i>Introduction</i>	
1.1	Introduction to the Research Area.....	1
1.2	Introduction to Lipophilicity and Efficiency Parameters.....	1
	1.2.1 Measure of Lipophilicity: Log P and Log D.....	2
	1.2.2 Property Forecast Index (PFI).....	3
	1.2.3 Drug Efficiency Parameters.....	3
1.3	Small-Molecule Hit Finding Approaches.....	7
	1.3.1 Hit Identification from High Throughput Screening (HTS)...	7
	1.3.2 Hit Identification from Encoded Library Technology (ELT)..	8
	1.3.3 Hit Identification from Fragment-based Drug Discovery (FBDD).....	9
1.4	Fragment Library Construction.....	10
1.5	Fragment-based Screening Techniques.....	12
	1.5.1 Biochemical Screening Techniques for Fragments.....	13
	1.5.1.1 Time-resolved – Förster/Fluorescence Resonance Energy Transfer (TR-FRET).....	13
	1.5.2 Biophysical Screening Techniques.....	15
	1.5.2.1 Thermal Melt (T_m).....	16
	1.5.2.2 Surface Plasmon Resonance (SPR).....	17
	1.5.2.3 Isothermal Titration Calorimetry (ITC).....	19
	1.5.2.4 NMR Techniques (Line Broadening, STD).....	20
	1.5.2.5 Pooling of Compounds for NMR Screening.....	23
	1.5.3 X-ray Crystallography as a Primary Screening Technique...	24
1.6	Fragment Library Screening.....	25
1.7	Optimisation of the Fragment and Approaches to Monitor Fragment Evolution.....	25

1.8	Small Molecule–Protein Binding Interactions.....	26
1.8.1	Thermodynamic Binding Contributions.....	26
1.8.2	Intermolecular Binding Interactions.....	30
1.8.2.1	Hydrogen Bonding (H-bonding) Interactions.....	31
1.8.2.2	Halogen Bonds.....	32
1.8.2.3	Aryl-Aryl and Aryl-Alkyl Interactions.....	32
1.9	Maintenance of Physical Properties During Drug Optimisation.....	34
1.10	FBDD-derived Compounds in the Clinic.....	35
2	<i>Inhibition of the Protein–Protein Interaction Between HIV Integrase and Lens Epithelium Derived Growth Factor (LEDGF) for the Treatment of HIV-1</i>	
2.1	HIV Disease.....	38
2.2	HIV Infection.....	38
2.3	Marketed HIV Inhibitors and the Requirement for New Therapies.....	40
2.4	HIV Integrase-LEDGF Interaction.....	41
2.5	Small Molecule Inhibitors of the HIV Integrase-LEDGF Interaction....	44
2.5.1	Tetraphenylphosphonium and Tetraphenylarsonium Inhibitors.....	45
2.5.2	Synthetic Peptides, including Cyclic Peptide Inhibitor (MZ4- 1).....	46
2.5.3	Thiazolidin-5-ylidene Analogues (D77)	47
2.5.4	Quinoline Derivatives.....	49
2.5.4.1	Boehringer Ingelheim Quinoline Analogues and Related Chemotypes.....	51
2.6	Summary.....	54
2.7	Background and Aims for Inhibition of the HIV IN-LEDGF Protein- Protein Interaction: Identifying Novel Inhibitors in Improved Physicochemical Property Space.....	55
2.8	Project Aims.....	58

2.9	Library Construction, Fragment Screening and Data Analysis.....	59
2.9.1	Library Construction.....	59
2.9.2	Identification of Appropriate Screening Methodologies.....	61
2.9.3	Screening of the ¹⁹ F and Fragment Libraries.....	63
2.9.4	Prioritisation of the Screening Output for Synthetic Preparation and Optimisation.....	69
2.10	Fragment Optimisation of Cluster 2 Analogue (2.36).....	73
2.10.1	Binding Mode of Quinoline (2.35).....	73
2.10.2	Binding Mode and Areas of Focus for Optimisation of Cluster 2 Analogue (2.36).....	74
2.10.3	SSS Around Piperidine Fragment (2.36).....	75
2.10.4	Amide Substitution to Interact with Glu170 and His171 of HIV IN _{CCD}	78
2.10.5	Identification and Synthesis of Piperidine Replacements for Fragment(2.36).....	85
2.10.6	Synthesis of 4-Piperidine and <i>t</i> -Butyl Replacements.....	88
2.10.7	Attempts to Determine an Alternative Linker Between the Core and Piperidine Moiety.....	91
2.10.8	Development of a Synthetic Route to Cyclohexyl Analogue (2.65).....	92
2.10.9	Development of a Synthetic Route to 3-Piperidine Analogue (2.63).....	95
2.11	Development of Synthetic Route to the 3-Pyrrolidine Analogue (2.64).....	100
2.12	Biochemical and Crystallography Results for 3-Piperidine and 3- Pyrrolidine Analogues.....	103
2.13	Amide Substitution to Interact with the Lipophilic Shelf.....	107
2.14	Synthesis of Quaternary Centre Analogues to Stabilise the Conformation of the Groups Occupying the Lipophilic Shelf.....	115

2.15	Copper-mediated Cyclisation Chemistry to Deliver Quaternary Carbon Centre Analogues.....	120
2.16	Benzyne Chemistry to Synthesise Benzocyclobutane Fragment (2.101).....	127
2.17	Biochemical and X-ray Crystallography Data Analysis for Quaternary Centre Analogues.....	129
2.18	Assessing the Benefits of the FBDD Approach for Identification of Inhibitors with Improved Physicochemical Properties.....	134
3	<i>Retrospective Analysis of the HIV IN_{CCD} – LEDGF FBDD Programme: Identification of the Most Efficient and Successful Method of Fragment Screening and Hit Optimisation</i>	
3.1	Assessing Target Tractability for FBDD.....	138
	3.1.1 Surveying the Literature.....	138
	3.1.2 Features of the Binding Pocket and Known Inhibitors.....	139
3.2	Determination of the Optimal Screening Triage.....	140
	3.2.1 Biophysical Screening.....	141
	3.2.2 Biochemical Screening.....	141
	3.2.3 Triage to X-ray Crystallography.....	144
	3.2.4 Identification of Biochemically Inactive Efficient Hits.....	145
	3.2.5 X-ray Crystallography.....	146
3.3	Fragment Optimisation Approach.....	146
	3.3.1 Fragmentation of Known Inhibitors.....	146
	3.3.2 Selection of Clusters for Synthetic Follow-up.....	148
	3.3.3 Mobility of Fragments.....	149
	3.3.4 Mapping of the Protein Active Site Using GRID.....	150
3.4	Summary of Successful FBDD Approaches to be used in Future Programmes.....	153

4	<i>Design of Fragment-derived Inhibitors of Cluster of Differentiation 38 (CD38) in the Absence of a Liganded X-ray Crystal Structure</i>	
4.1	CD38 Receptor Function.....	155
4.2	CD38 Structure.....	157
4.3	Therapeutic Applications.....	160
4.4	Reported CD38 Inhibitors.....	161
4.5	Relevant Background to the CD38 Programme.....	165
	4.5.1 Fragment Effort Aims.....	166
	4.5.2 CD38 Screening Cascade.....	166
	4.5.3 Challenges with the CD38 X-ray Crystallographic System.....	172
	4.5.4 The CD38 Screening Cascade.....	173
	4.5.5 CD38 Thesis Aims.....	175
4.6	Results and Discussion.....	177
	4.6.1 Analysis of the CD38 Binding Pocket.....	177
4.7	Fragment Optimisation and Growth of 7-Oxo-5-propyl-4,7-dihydropyrazolo[1,5-<i>a</i>]pyrimidine-3-carbonitrile (4.10).....	184
	4.7.1 Proposed Binding Mode for Pyrazolopyrimidinone (4.10).....	185
	4.7.2 Fragment Optimisation – Single Point Changes.....	189
	4.7.3 Exploration of 5-Position Substitution.....	197
	4.7.4 Simultaneous 5- and 6-Position Substitution.....	204
	4.7.5 Exploration of the 6-Position of Pyrazolopyrimidinone (4.10).....	206
	4.7.6 Exploration of the 5-Position of Pyrazolopyrimidinone (4.10).....	208
	4.7.6.1 Attempts to Improve the Physicochemical Properties.....	224
	4.7.6.2 Summary of Efforts to Identify Potent, Efficient CD38 Binders from the Pyrazolopyrimidinone (4.10) HTS Hit.....	234

4.8	Optimisation of 1<i>H</i>-Imidazo[4,5-<i>c</i>]pyridin-4-amine (4.5) for CD38 Binding Leading to Identification of the Imidazopyrazine Series.....	238
4.9	Optimisation of the Amide Substituent of the Pyrazolopiperidine Series Through Array Design.....	273
4.10	Profiling Representative Analogues from the Imidazopyrazine Series in hERG and P450 Screens.....	284
4.11	Summary of Efforts to Identify CD38 Binders without Crystallography of Compounds from the Series.....	287
5	<i>Conclusions</i>	
5.1	Successful FBDD Optimisation Strategies.....	295
5.2	Benefits this Research Brings to the Wider Scientific Community.....	301
6	<i>Experimental Section</i>.....	303
6.1	Experimental details – HIV Integrase-LEDGF.....	312
6.2	Experimental Details – CD38.....	339
7	<i>References</i>.....	388
8	<i>Appendices</i>.....	407

1 Introduction

1.1 Introduction to the Research Area

Small molecule drug discovery remains critically important to find safe and efficacious medicines for patients with unmet medical need. These small molecules bind to the biological target of interest modulating a biological pathway that is critical for the treatment of disease.

The physical properties of these compounds, particularly lipophilicity (the affinity of a molecule for a ‘lipid loving’ environment), has been reported in the literature to be one of the factors that affect the attrition of small molecules during development.^{1,2} In order to identify compounds in the optimal physicochemical property space, fragment-based drug discovery (FBDD) is used within medicinal chemistry groups to complement alternative hit finding methods, e.g. high throughput screening (HTS) and encoded library technology (ELT). The FBDD process identifies low molecular weight hits, which are carefully optimised using structure-guided drug design into efficient drug-like molecules.³ During hit optimisation, emphasis is placed on maintaining the molecules’ properties in the desired chemical space, whilst maximising the binding efficiency to the target.

This thesis will focus on the use of FBDD methodology as a method of identifying and optimising fragment hits in good physicochemical property space for two drug discovery projects. These projects include the challenging area of inhibiting protein-protein interactions and optimisation of fragments in the absence of X-ray crystallography.

1.2 Introduction to Lipophilicity and Efficiency Parameters

Optimisation of a screening hit to a drug molecule usually results in an increase in lipophilicity and molecular size. Increases in these parameters occur because the most beneficial binding interactions, which increase the affinity of a ligand to a protein, are often lipophilic interactions.⁴ It has been reported by Leeson and Springthorpe¹ that highly lipophilic compounds are able to bind productively but

unselectively to a number of diverse proteins and may, therefore, be referred to as “promiscuous”.⁵ Accordingly, developing safe drug molecules for lipophilic protein targets often proves challenging due to a higher chance of promiscuity, resulting in off-target activity and toxicity.^{1,4} To mitigate this issue, careful consideration of lipophilicity and efficiency of binding to the biological target need to be measured and monitored throughout the drug optimisation process and, indeed, are critical for FBDD. Therefore, before discussing the FBDD technique in detail, important lipophilicity and efficiency parameters will be introduced.

1.2.1 Measure of Lipophilicity: Log P and Log D

Log P is described as the concentration of the neutral species in octanol, divided by the concentration of the neutral species in water (**Equation 1 (a)**).⁶ Log D is pH dependant and must have the pH specified (**Equation 1 (b)**). For a neutral species Log P and Log D are always equal. Partition between octanol and water is the most widely used method for measuring Log P and Log D. More recently, higher throughput methods, such as immobilised artificial membranes (IAM)⁷ are being used within an industrial setting to gain a measure of lipophilicity. The preferred method within our laboratories is the HPLC chromatographic hydrophobicity index (CHI) IAM method referred to as chrom Log D throughout this research project.⁸

$$(a) \quad P = \frac{[\text{neutral species}]_{\text{octanol}}}{[\text{neutral species}]_{\text{un-ionised water}}}$$

$$(b) \quad D = \frac{[\text{species}]_{\text{octanol}}}{[\text{species}]_{\text{ionised water}} + [\text{species}]_{\text{water}}}$$

Equation 1 – (a) Calculation for Partition co-efficient, P; (b) Calculation for Distribution co-efficient, D.

Monitoring compound lipophilicity is critical for the development of leads into candidate drug molecules. Suitable levels of aqueous solubility can often be

achieved by maintaining the log of the partition co-efficient (Log P) or distribution co-efficient (Log D) between 1 and 3.⁹ Maintaining Log D in this range is important as highly lipophilic compounds are often poorly water soluble, resulting in low oral absorption.¹⁰ Lipophilic compounds are also more likely to be highly plasma protein bound, reducing the concentration of free drug able to reach the desired receptor and to effect a biological response.⁴ Conversely, very hydrophilic molecules suffer from poor intestinal absorption,¹¹ helping contextualise the requirement to maintain Log P/Log D between 1 and 3.

1.2.2 Property Forecast Index (PFI)¹²

During drug discovery it is common that the molecules increase in size and become increasingly lipophilic and aromatic.¹ This is often attributed to non-specific hydrophobic interactions giving increased *in vitro* potency.¹ Young *et al.* have developed the property forecast index (PFI), which is the summation of measured Chrom Log D_{pH7.4} and number of aromatic rings (**Equation 2**).¹² Analysis of PFI with respect to developability criteria (P450 inhibition, solubility, intrinsic clearance (Clint), hERG, and promiscuity) demonstrates that compounds with PFI = 4-5 have the preferred profile for progression into clinical assessment and deviation from this range gives an increased chance of attrition throughout the drug discovery process. Therefore, it is recommended that PFI is ideally maintained below 6 and does not exceed 7.¹² Development of a Chrom Log D calculator allows calculated PFI (cPFI) values to guide the synthetic strategy within the drug discovery process.

- (a) $\text{PFI} = \text{mChrom Log D}_{\text{pH7.4}} + \# \text{ aromatic rings}$
- (b) $\text{cPFI} = \text{cChrom Log D}_{\text{pH7.4}} + \# \text{ aromatic rings}$

Equation 2 – (a) measured PFI; (b) calculated PFI definitions. cChrom Log D_{pH7.4} is generated from an unpublished calculator developed within our laboratories which correlates well with measured Chrom Log D_{pH7.4} values.

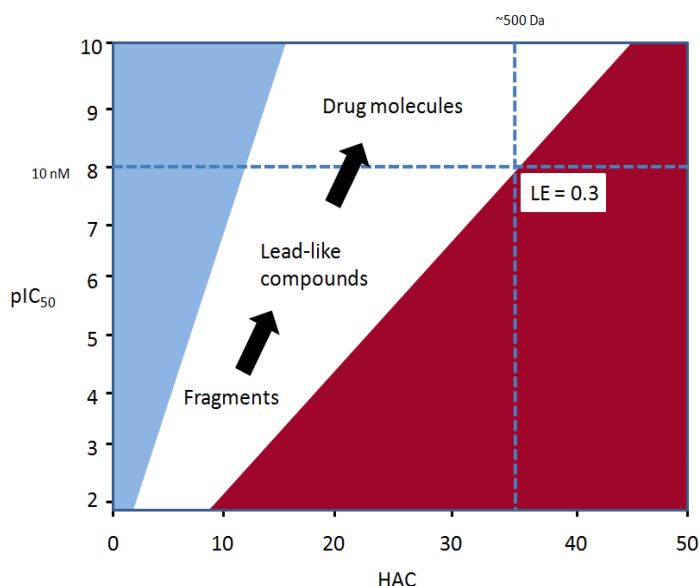
1.2.3 Drug Efficiency Parameters

Ligand efficiency (LE),^{1,3,13} lipophilic ligand efficiency (LLE),^{1,14} and Astex modified lipophilic ligand efficiency (LLE_(AT)) are measures of the binding

efficiency of the small-molecule to the biological target (**Equation 3** and **Graph 1**).^{15,16}

- (a) $LE = 1.37 (pIC_{50} / HAC)$
 - (b) $LLE = pIC_{50} - cLog P$
 - (c) $LLE_{(AT)} = ((pIC_{50} - cLog P) / (0.74 \times HAC)) + 0.11$
- HAC = Heavy atom count

Equation 3 - Definitions for (a) LE, (b) LLE and (c) $LLE_{(AT)}$.

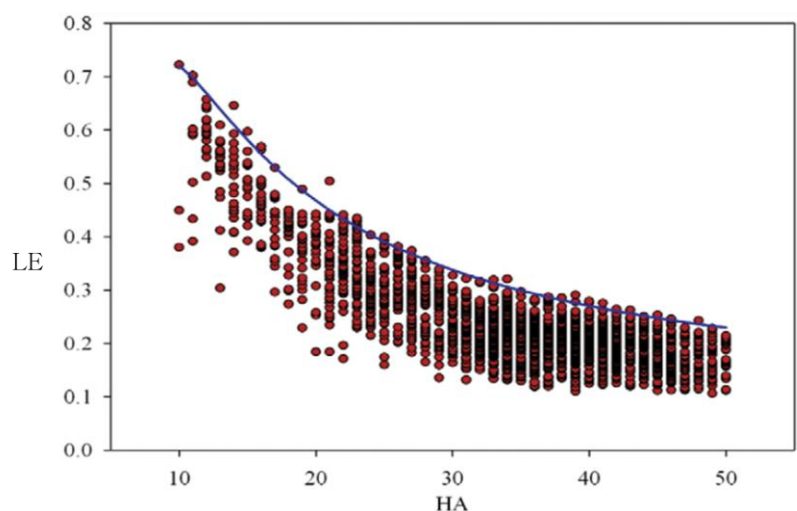


Graph 1 - Ligand efficiency with respect to pIC_{50} and molecule size.¹

LE calculates the efficiency of a binding interaction with respect to the magnitude of the ligand's physical properties, most importantly mass.¹ Achieving an $LE \geq 0.3$ has been deemed optimal.¹⁷ **Graph 1** allows approximation of the potency required to achieve $LE \geq 0.3$ for compounds with varying heavy atom count (HAC). The border between the red and white sections represents a $LE = 0.3$, with the red area showing sub-optimal LE space ($LE < 0.3$). For fragment compounds ($HAC < 20$) only weak potencies ($pIC_{50} \leq 4.5$) are required for identification of ligand efficient hits. In contrast, for a clinical candidate (~ 500 Da; $HAC \sim 35$) potency of ≤ 10 nM ($pIC_{50} \geq$

8) is required to remain in good efficiency space, as represented by the dotted lines on **Graph 1**. Therefore, the aim of FBDD is to identify fragments with a ligand efficiency of 0.3 or greater and monitor the LE through the optimisation process, aiming to increase this function or keep it constant. If a substantial reduction in LE is observed (> 10%), through a modification of the fragment, it is deemed that the change is not optimal. However, care needs to be taken not to overinterpret the LE change and an understanding of the assay variability is critical.¹³ Also, more pronounced changes are observed when growing smaller molecules, as LE does not change linearly with respect to the number of heavy atoms, as illustrated by **Graph 2**. LE levels out at ~ 45 heavy atoms, which is proposed to occur due to two phenomena:¹⁷

- 1) The accessible surface area per atom that can form interactions with the protein pocket decreases with increasing ligand size. This is due to the greater burial of smaller molecules in protein pockets and the increased surface area of larger molecules exposed to solvent.
- 2) Larger ligands have to overcome structural rearrangement to bind in the protein pocket. This provides an energy penalty to be overcome on binding.



Graph 2 – LE (pKi/HA) plotted against number of heavy atoms (HA) for a representative data set against multiple protein targets.¹³ Reprinted from Bembenek, S. D.; Tounge, B. A.; Reynolds, C. H. *Drug Discovery Today* 2009, 14, 278-283, Copyright 2009, with permission from Elsevier.

Monitoring ligand efficiency and lipophilicity is important throughout the optimisation of a screening hit regardless of the method or approach being taken. However, careful assessment and control of these factors are critical for the success of a FBDD screening and optimisation approach.

1.3 Small-Molecule Hit Finding Approaches

It is well documented that small molecule drug endeavours have experienced challenges in identifying good starting points and developing such identified units into safe, efficacious drugs.^{3,19,20} To address such issues, a number of novel hit identification methodologies have been developed, including ELT and FBDD. Multiple hit finding techniques are often used in parallel within pharmaceutical laboratories, as the chemical nature of the hits from the different screening techniques can differ significantly. The remainder of this section introduces the main hit finding techniques in turn, and concludes by comparing and contrasting the differences between the chemical outputs.

1.3.1 Hit Identification from High Throughput Screening (HTS)

The HTS approach has been extensively used since the 1990's within pharmaceutical laboratories.³ The belief was that screening large numbers of compounds (~2 million), using advanced automated technology, would increase the chance of identifying a lead molecule.²¹

Similar to views communicated in recent publications,^{1,19} within our laboratories the HTS technique has not always been found to be effective at identifying good small molecule starting points, as little importance was originally placed on the properties of compounds during library construction. As a result, a high percentage of the compounds in our library have higher molecular weight and lipophilicity than desired, which has resulted in identification of hits with poor physicochemical properties. To address this concern, evolution of our HTS library has begun, with greater consideration being placed on inherent compound properties. Reduction of the mean molecular weight and lipophilicity of compounds in our laboratory's library is expected to provide identification of compounds with improved drug-like

properties compared to historical HTS hits. It is envisaged that this will reduce the overall levels of drug candidate attrition.

In relation to the above, high molecular weight HTS compounds may have redundant functionality due to their level of complexity. Therefore, the HTS hits could have poor shape and electrostatic complementarity with the binding pocket of interest, resulting in poor binding efficiencies (**Equation 3**). Accordingly, two approaches are commonly used to transform a hit molecule into a drug candidate for progression into the clinic:

- 1) Further growth from the HTS hit to improve the binding potency. This approach can lead to high molecular weight lipophilic analogues;
- 2) Fragmentation of the HTS hit to identify the most ligand efficient starting point before growth. This approach is favoured as it maximises the efficiency of the ligand-target binding interaction and minimises redundant functionality.

1.3.2 Hit Identification from Encoded Library Technology (ELT)²²

Encoded Library Technology (ELT) complements HTS in the search for hit molecules.²³ The ELT technology provides the capability to assess over 4 billion compounds, in a single screen, from a series of combinatorial DNA encoded libraries each containing 10^6 - 10^{10} molecules.²⁴ Affinity-based screening is used to identify molecules which bind to the biological target of interest. The DNA tags of bound molecules are amplified and sequenced providing the identity of the molecule. To confirm the binding of the compound to the target, the molecule is re-synthesised off-DNA and the binding affinity is reassessed. The DNA sequence codes for the building blocks used in the combinatorial library construction and allows key binding features of the molecules to be identified. Identification of key features in the hit molecules helps determine the most efficient small molecule binder from which to begin the optimisation process.

The requirement for the compounds to be DNA tagged limits the chemistry that can be used to construct and diversify the library *via* this ELT approach. Despite this,

ELT has proved a successful hit finding technique for a variety of therapeutic targets including kinases,²² metallo proteases,²⁵ sirtuins,²⁶ and the challenging area of protein-protein interactions.²⁴ An excellent example is the identification of soluble epoxide hydrolase (sEH) inhibitors using an ELT screening campaign, providing tool molecules which showed encouraging *in vivo* target engagement (**Figure 2**).²⁷ Although there are such examples of success from using HTS²⁸ and ELT²⁴ in targeting protein-protein interactions, alternative screening strategies are still required to maximise the chance of finding efficient small molecules to develop into effective drugs.

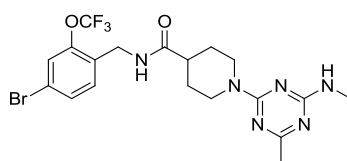


Figure 2 – sEH inhibitor identified from an ELT screening campaign. Human sEH HEK IC₅₀ = 3 nM; Rat Clp iv 1.0 mg/kg = 12 (mL/min/kg); Liability target P450 CYP2C9 IC₅₀ > 33,000 nM.²⁷

1.3.3 Hit Identification from Fragment-based Drug Discovery (FBDD)

As a relatively novel and complementary technique to HTS and ELT, FBDD is focused on identifying low molecular weight, efficient binders.²⁹ The technique is distinct from HTS and ELT as it assesses a small diverse fragment library of between 1000-3000 fragments, in both biochemical and biophysical assays in parallel (see **Section 1.5**). The screening output typically consists of low potency (mM), highly efficient fragments which bind to the chosen biological target. As the fragments are small, there is particular interest in using FBDD to identify hit molecules for inhibition of protein-protein interactions. Inhibiting protein-protein interactions with small molecules is considered very challenging as the interaction surface is large (~1000 Å²) and flat.

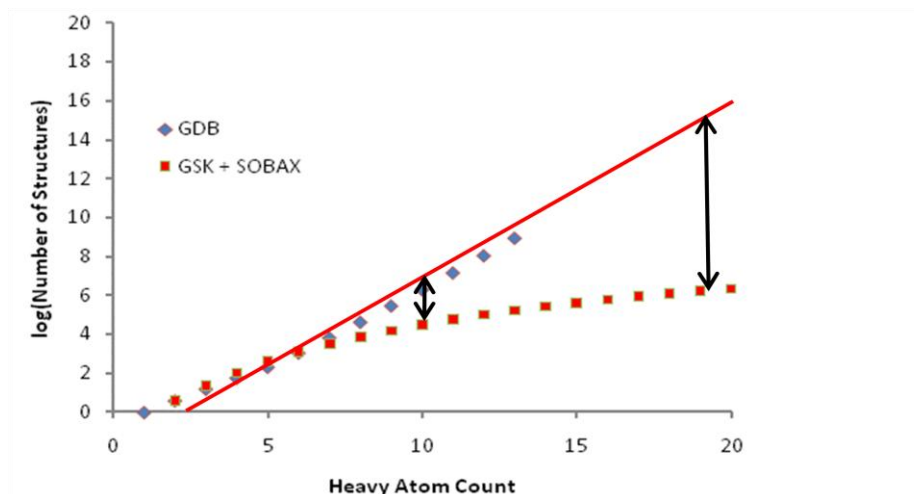
Even if fragment hits are identified from the FBDD screen, in a general sense the approach is only considered effective if careful tracking and control of molecular

properties and maintenance of LE is carried out during fragment optimisation. Providing the library is well designed, and hit optimisation is pursued with rigour, it is envisaged that compounds with lower molecular weight, higher binding efficiency, and improved physicochemical properties can be identified with respect to hits optimised from HTS and ELT screening campaigns.

FBDD is often used in parallel with HTS and/or ELT, as high lipophilicity has been linked with drug molecule attrition.¹ The FBDD methodology is of particular interest over HTS and ELT, as the hits identified are usually less lipophilic ($c\text{Log } P < 3$). Optimisation of a screening hit often results in increased lipophilicity of the molecule, therefore, starting from a more polar fragment is of appeal. Based on all of the above, to ensure that good quality hits are identified from FBDD, significant importance is placed on the design of the fragment library.

1.4 Fragment Library Construction

Three features need careful consideration when a FBDD library is compiled: the size of the library; the properties of the compounds populating it; and ensuring that diversity is maximised. The number of compounds in a FBDD screening library is typically between 1000 and 3000, with each member being carefully chosen for physicochemical properties, functional groups, and positions whereby the molecule can be grown. Despite the limit in library size, a carefully designed fragment library of between 1000-3000 diverse fragments ($\text{HAC} \leq 10$), covers proportionately more chemical space than an entire HTS library of ~ 2 million compounds ($\text{HAC} \leq 19$). This is illustrated by consideration of **Graph 3**, where the increase in chemical space covered with increasing HA count from a computational chemical universe library (GDB) (blue diamond's), is compared to the combined molecules within our laboratories and the externally available SOBAX library (red squares). As the number of HA's in a molecule increases, the difference between the theoretical chemical space and the chemical space covered by available molecules becomes significantly larger (**Graph 2**, black arrows), giving a negative effect to increasing the library size.



Graph 3 - Graph³⁰ showing the increase in chemical space, represented by GDB-13³¹ (a chemical universe database), with heavy atom count, compared to the molecules within the GSK and SOBAX libraries.

To help control the physicochemical properties of the compounds in fragment library design, the ‘rule of three’, reported by Congreve *et al.*,³² is often applied. The rule of three was derived by surveying the output of fragment screens from a range of biological targets. The results indicated that, on average, the hits possessed: MW < 300; hydrogen bond donors ≤ 3 ; hydrogen bond acceptors ≤ 3 ; and cLog P < 3. The combination of these properties is known as the ‘rule of three’.

It has been shown that using appropriate screening techniques, FBDD can identify small fragments of typically 100-300 Da.³³ The fragments bind optimally to the target of interest (**Figure 3** – (a)) at recognition sites known as hotspots. The ultimate aim is to develop these fragment hits into drug molecules, through optimal growth of the compounds to fill and interact with the remainder of the protein pocket (**Figure 3** – (b) and (c)). For this to be successful, appropriate screening methodology must be available to monitor relative affinity and efficiency of the fragments for the biological target.

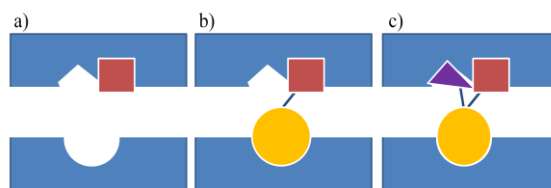


Figure 3 – (a) Diagram showing a hit molecule from an FBDD screening approach, (b) and (c) demonstrate efficient iterative growth of the fragment hit.³⁴

1.5 Fragment-based Screening Techniques

The small molecule inhibitors detected from HTS and ELT campaigns are usually in the μM - nM range. Fragment hits bind more efficiently to the target than HTS and ELT hits, but with lower affinity (mM).²⁹ The weak affinity of fragments is a consequence of their low MW and small area of interaction with the protein. This low affinity results in fragments residing outside of the detection limits for most biochemical assays, which tend to be configured for lead optimisation purposes. Re-configuration of these biochemical assays can allow binding of weaker fragments to be measured without suffering from a high number of false positives (*vide infra*).

To complement biochemical screening, advances have been made in biophysical detection methods to measure the relative affinity of weak fragments for the target.^{4,35,36} Biophysical techniques can measure the direct interaction of a ligand with a protein, and have greater sensitivity than biochemical methods, allowing detection of weak (mM) binders. Biophysical methods of note within this domain include surface plasmon resonance (SPR),³⁵⁻³⁷ isothermal titration calorimetry (ITC),^{36,38,39} thermal melt (T_m),⁴⁰ and mass-spectrometry coupled with electrospray ionisation (MS-ESI).^{36,41} For each biological target, the most appropriate biophysical and biochemical screening methodologies are identified using a small set of fragments randomly selected from a fragment library. Once an appropriate screening cascade has been constructed, the entire fragment library is assayed to identify hits, which are then progressed to X-ray crystallography. X-ray crystallography provides information relating to the location and orientation of the molecule, which is used to

guide subsequent fragment growth through synthesis. In addition to X-ray crystallography, it has been demonstrated that NMR-based technologies such as 2D ¹⁵N-HMBC can be used as an alternative.⁴²

1.5.1 Biochemical Screening Techniques for Fragments

An advantage of biochemical screening over most biophysical methods is that the assays are higher throughput and can rapidly rank compounds by binding affinity to the protein, or by extent of modulation of a biological function. As outlined above, the anticipated affinity of FBDD hits is in the mM range. For biochemical screening, the concentration of ligand required to detect the low affinity fragments could cause problems with ligand solubility and aggregation. This limit in ligand concentration often prevents the use of biochemical screening for detection of mM fragment hits. Therefore, biochemical assays are not normally used as the sole screening method, as weak but interesting fragment hits would be missed. Due to increased interest in fragment screening, biochemical assays are now being re-configured to enable measurement of weaker binders.⁴³ In this regard, the remainder of this section provides detail relating to the commonly used time-resolved Förster/fluorescence resonance energy transfer (TR-FRET) biochemical assay format, which was used in the screening cascade for identification of inhibitors of the protein-protein interaction described in **Chapter 2**.

1.5.1.1 Time-resolved – Förster/Fluorescence Resonance Energy Transfer (TR-FRET)

Time resolved – Förster/fluorescence resonance energy transfer (TR-FRET) assays are widely used in drug discovery. This arises from this general method being easily configured to different targets and, further, the developed techniques are suitable to run in a high throughput manner.⁴⁴ This assay approach relies on a signal generated through fluorescence resonance energy transfer (FRET), between a donor and an acceptor molecule, that are in close proximity to each other (**Figure 4; left image**).^{45,46} If small molecules bind to the protein active site, this inhibits the donor and acceptor from being in close proximity, preventing, or reducing TR-FRET

emission depending on the residence time of the inhibitor at the protein site (**Figure 4**; right image). The greater reduction in TR-FRET emission, the more potent the small molecule. The time-resolved (TR) fluorescence provides a delay between excitation of the donor and the fluorescence measurement (~ 100-150 μ s). This delay avoids the recording of non-specific short-lived emissions, which can reduce the accuracy of the data. This results in a sensitive and robust assay that can be run using various plate formats.

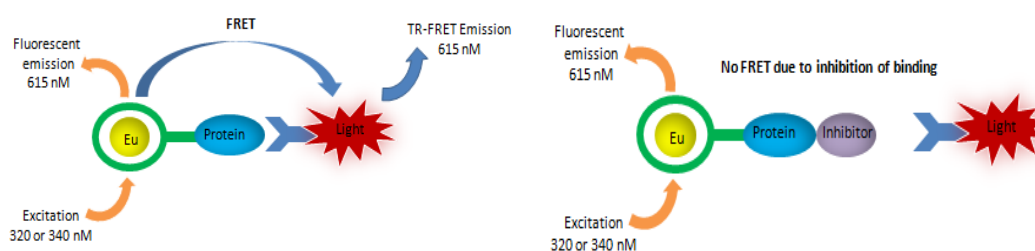


Figure 4 - TR-FRET schematic for demonstrating TR-FRET emission when donor and acceptor come into close proximity.⁴⁶

The TR-FRET technology uses the unique properties of the lanthanides: long emission half-life, stability, and ability to be linked to bio-organic molecules. Producing a fluorescence signal by direct excitation of lanthanides is difficult unless they are complexed with ‘light harvesting’ organic molecules (e.g. a cryptate). Europium cryptate was used as the fluorescence donor when the FRET technique was first developed.⁴⁴ Over the past decade, research into the technique has continued and has identified an alternative donor, terbium cryptate.^{45,47} Terbium cryptate is reported to demonstrate enhanced screening performance in comparison to europium cryptate.

Advantages of using terbium cryptate over europium cryptate include:

- increased quantum yield (improved fluorescence efficiency);
- higher molar extinction co-efficient (how strongly a chemical species absorbs light at a given wavelength); and

- multiple emission peaks at different wavelengths, which reduces the chance of fluorescence interference from small molecules, buffers, and proteins.

Terbium and europium have also been reported to simultaneously measure activity of molecules against two different protein targets in the same well using FRET technology. This is only feasible due to the different emission wavelengths between europium (665 nM) and terbium (545 nM).⁴⁸ This advance in the technology reduces time and resource constraints by measuring data against multiple targets.

Biochemical assays, including the TR-FRET format, are susceptible to interference through a number of means. Mechanisms of interference include: fluorescence from small molecules at the detection wavelength; and binding of small molecules to the tags which attach the donor and acceptor to the biological systems, disrupting the biological complex through alternative means other than binding to the pocket of interest. These interference mechanisms are only experienced in such biochemical screening formats. In contrast, parallel screening with biophysical methods provides less risk of identifying false positives.

1.5.2 Biophysical Screening Techniques

Biophysical screening techniques are essential for fragment-based screening campaigns. They are more sensitive than biochemical methods and allow for the detection of weak affinity molecules (mM). Accordingly, biophysical assays are more applicable to fragment screening than biochemical methods, as compounds can be screened at a higher concentration with less chance of interference.

There are a range of biophysical techniques available, which provide different levels of information. The information ranges from simple identification of binding a small molecule to the protein, through to more detailed information such as measurement of the dissociation constant (K_d) at a specific active site. Not all biophysical techniques are suitable for all targets, as the size or stability of the protein or enzyme play an important role. The applicability of the biophysical techniques for the target

of interest and, subsequently, identifying the most robust biophysical method prior to the full fragment screen is of utmost importance. The main biophysical techniques are briefly described below, highlighting the advantages and disadvantages of each.

1.5.2.1 Thermal Melt (T_m)

Thermal melt measures the degree of stabilisation of a protein's folded state upon binding of small molecules, by measuring the change in the melting point of the protein.^{49,50} Fluorescent dyes, such as Sypro Orange™ that show sensitivity to its environment, are added to the assay. When the protein is properly folded the dye will remain in the aqueous environment and the fluorescence signal will be quenched. As the temperature is increased and the protein begins to unfold, the dye binds to the hydrophobic region of the protein resulting in an increase in fluorescence (**Figure 5**). The increased fluorescence indicates the change in state of the protein i.e. its melting point (T_m).

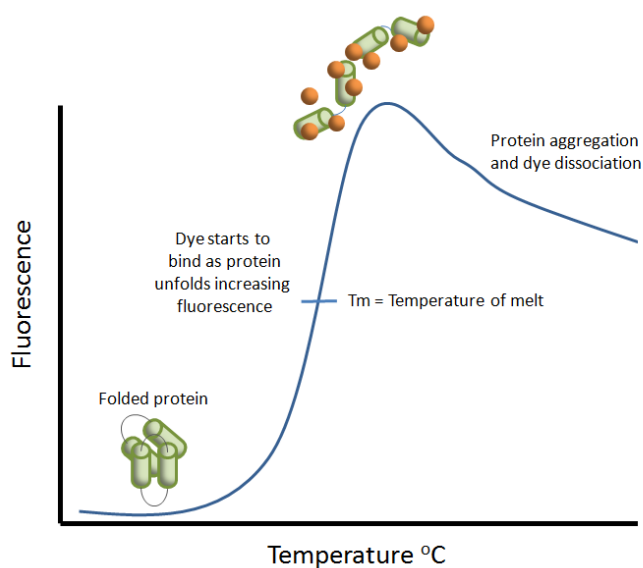


Figure 5 - Thermal melt graph demonstrating increase in fluorescence upon protein unfolding.^{49,50} Green cylinders represent the protein and the orange spheres represent the Sypro Orange™ dye.

When a small molecule binds to the target this results in increased stability of the protein which translates to an increase in its melting point. A disadvantage of the technique is that the small molecule could be binding to the protein in a biologically

less relevant manner, e.g. to the protein surface or an allosteric site. No indication of the binding location makes it difficult to build directed hypotheses from which to form synthetic plans. Structural activity relationships cannot be derived from T_m experiments as the thermal melt temperature and the binding affinity do not correlate, therefore, this technique cannot be used to fully drive medicinal chemistry.⁵⁰

An advantage of thermal melt is its high throughput nature, whilst requiring low protein consumption. However, within our laboratories T_m is rarely used as a technique for fragment screening as there is often only a small T_m upon fragment binding that can be difficult to record, or a low signal/noise ratio results.

1.5.2.2 Surface Plasmon Resonance (SPR)

Substantial effort is being invested to improve the biophysical techniques available for FBDD screening. Surface plasmon resonance (SPR) is a relatively new technique with many advantages over other biophysical methodologies.⁵¹ The technique uses a gold coated glass sensor chip, 50-100 nm thick, to which the protein is immobilised using immobilisation media such as streptavidin or carboxymethyl-dextran (carbohydrate).³⁷ The protein is attached to an exposed chemical handle on the medium through a covalent linker. The protein attachment point has to be determined with care to avoid the linker occluding the protein binding site. A solution of the ligand is then passed over the sensor so that it can bind to the active site of the protein.

To detect binding of a fragment to the protein, light is passed through a prism that is positioned on top of the sensor (**Figure 6**).⁵² The light beam is configured to strike the prism at an angle so that the beam is totally reflected (total internal reflection). The evanescent wave (the wave formed when light is reflected off an interface at an angle greater than the critical angle) passes into the gold layer, exciting the electrons.⁵³ At a particular angle and wavelength, excitation of the gold electrons produces plasmon resonance (oscillation of free electron density). The plasmon resonance results in a decrease in intensity and change in angle of the reflected wave

which is measured by a detector. The SPR produced is sensitive to changes to the binding surface e.g. the binding of small molecules. The change in resonance allows the detection of association and dissociation of small molecules to the protein.

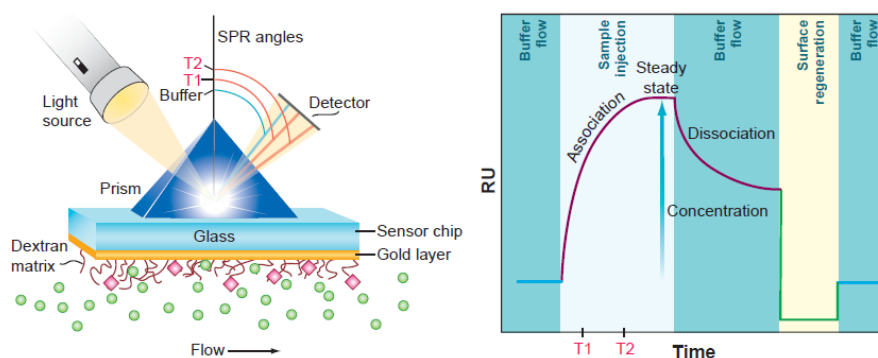


Figure 6 - SPR experiment. The ligand (green spheres) is passed over the SPR chip with protein (pink diamonds) linked to the gold layer by dextran. As ligands bind to the protein the resonance angle changes, represented by T_1 , T_2 and buffer angles.⁵²

From Wilson, D. Science, 2002, 295, 2103. Reprinted with permission from American Association for the Advancement of Science.

SPR is commonly the biophysical method of choice for fragment screening, due to the wealth of information that can be gained from the SPR binding response.^{37,50} SPR can be used to determine: binding stoichiometry; reversibility of the protein-ligand interaction; association rate constant (k_{on}); dissociation rate constant (k_{off}); dissociation equilibrium binding constant (K_d); and thermodynamic parameters, including free energy (ΔG), enthalpy (ΔH), and entropy (ΔS) (See **Section 1.8.1** for thermodynamic binding parameters). The breadth of information available gives a clear advantage over other biophysical techniques. Competition experiments may also be used to confirm that the interaction occurs at a particular binding site. SPR experiments are also able to be automated (including data analysis) and, therefore, the technique can be run in a high throughput manner for an entire fragment library.

The major disadvantage of SPR is the inability to immobilise all targets whilst maintaining protein stability. For many targets, immobilisation of the protein to the carbohydrate linkers on the gold surface has not been possible due to protein

denaturation. This limitation of the technique requires additional research into alternative biophysical methodology, or further advances in currently available techniques.

1.5.2.3 Isothermal Titration Calorimetry (ITC)

Isothermal titration calorimetry (ITC), like SPR, is capable of measuring thermodynamic binding events in solution for protein-protein or protein-ligand interactions.³⁸ For protein-ligand interactions, aliquots of ligand are injected into a glass cell containing protein solution (20 – 100 μ M). ITC measures the energy of the binding event. If energy is released during the binding event (exothermic) the ITC cell is cooled, and if energy is absorbed (endothermic) the ITC cell is heated to maintain a constant temperature (**Figure 7** - top plot). From the integration of the peaks, K_d , enthalpy, and binding stoichiometry, can all be measured from a single ITC experiment (**Figure 7** – bottom plot). Entropy and free energy of binding can then be calculated (see **Section 1.8**). Using ITC to measure binding is appealing as no pre-immobilisation or labelling of the protein is required. Disadvantages of the technique include high protein consumption, which often limits the use of ITC for routine screening. ITC experiments must be run singularly, and this considerably limits the throughput of the technique.

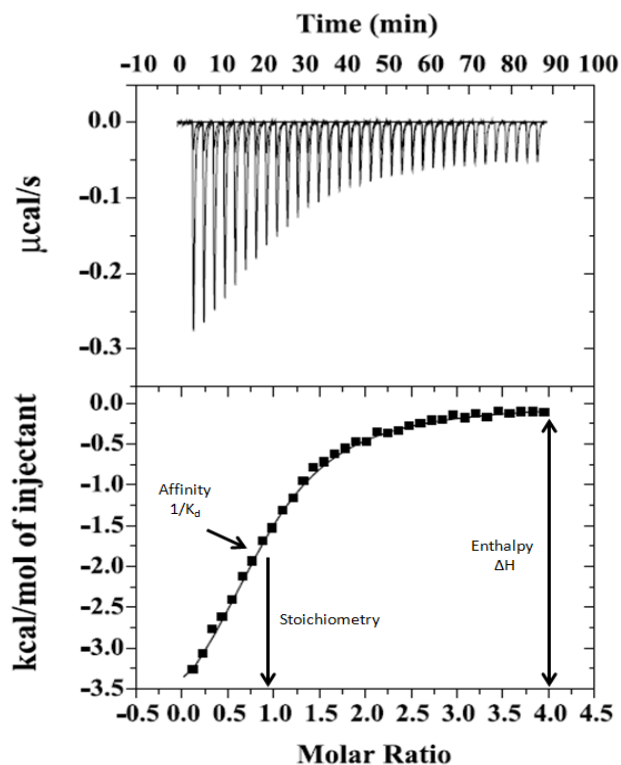


Figure 7 - ITC curve demonstrating determination of K_d , enthalpy and stoichiometry of the binding event.⁵⁴ Reprinted from Lo, M.-C.; Aulabaugh, A.; Jin, G.; Cowling, R.; Bard, J.; Malamas, M.; Ellestad, G. *Anal. Biochem.* 2004, 332, 153-159, Copyright 2004, with permission from Elsevier. Additional labels added to the graph.

1.5.2.4 NMR Techniques (Line Broadening, STD)

NMR techniques for the screening of fragment libraries, may be characterised into two distinct NMR types: protein^{42,55} and ligand.⁵⁶ Protein NMR is limited by the size of the protein which should be between 30-50 kD. Larger proteins have slower tumbling rates and shorter NMR relaxation times which reduces the sensitivity of the technique. The larger the protein, the more complex the spectrum causing challenges in data interpretation. Protein NMR also requires isotopic labelling of the protein, which is expensive and time consuming to obtain. For these reasons, protein NMR is less utilised for FBDD and the remainder of this section focuses on ligand-based NMR techniques.

Line Broadening NMR

Ligand-based NMR methods have wider application than protein-based methods, and time consuming and expensive labelling (e.g. ^{15}N) of the protein or ligand is not required. Ligand based NMR studies are information rich providing valuable structural information about the location and binding orientation of fragments to proteins.⁵⁷ In line broadening NMR, relaxation properties of the molecule are an important parameter to consider. Relaxation properties depend on tumbling rates, which depend on molecular size. Unbound ligands are small and tumble fast. They exhibit fast relaxation rates and give sharp signals. In comparison, when the ligand is bound to the protein, a slower tumbling rate is observed; this leads to a slower transverse relaxation rate. The slower relaxation rate produces broadening of ligand signals and loss of resolution (flattening) of ligand spectra. An example of line broadening spectra is shown in **Figure 8**. The top NMR in **Figure 8** is a reference spectrum. The bottom spectrum shows broadening of compound B's signals (dotted arrows), indicating that it is binding to the protein.

Advantages to using line broadening NMR include: the ability to screen mixtures of molecules, increasing the total number of compounds that can be assessed; and ranking of binding affinity to the protein. The binding affinity can be estimated enabling ranking by comparison of the extent of broadening of the compounds' NMR signals. Furthermore, the perceived disadvantage of not knowing the location of binding of any ligand molecule can be overcome by carrying out a displacement experiment using a known binder.⁵⁷

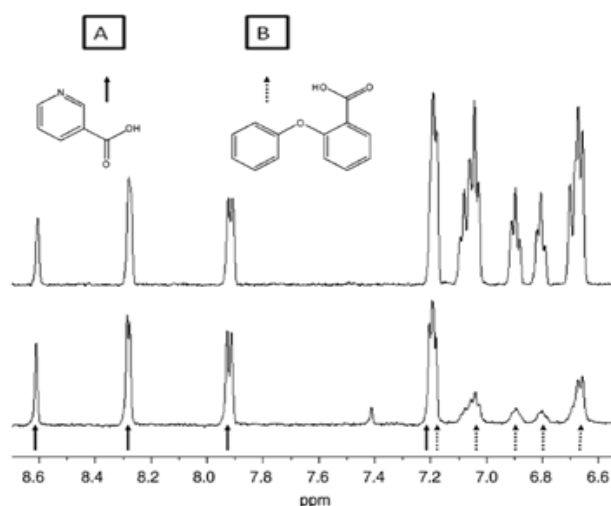


Figure 8 – Example of line broadening spectra.⁵⁸ Reprinted by permission from Professor Peng; Peng, J. W.; Lepre, C. A.; Fejzo, J.; Abdul-Manan, N.; Moore, J. M. *Methods Enzymol.* 2001, 338, 202-23, structures added to Figure for clarity.

STD NMR

An alternative NMR technique for fragment screening is a magnetisation transfer experiment known as saturation transfer difference (STD).⁵⁶ In this experimental design, the protein proton signals are selectively irradiated. Protein methyl groups are selected for irradiation as they have low shifts when the protein is folded (just above 0 ppm),⁵⁹ which avoids irradiation of the free ligand.

A reference spectrum of the free ligand is recorded which displays sharp ligand signals from fast relaxation rates. The most up-field methyl groups on the protein are irradiated which causes transfer of irradiation to any bound or partially bound ligand. The radiation transfer to the ligand results in a reduction in NMR signal height for the bound ligand, due to slower tumbling and decrease in relaxation rates. Signal intensity for unbound ligands remains unchanged when compared to the reference spectrum. The STD spectrum is obtained by subtracting the irradiated spectrum from the reference spectrum, leaving only the signals for protein bound ligand (**Figure 9**).

The success of STD NMR depends on the stability of the protein fold in solution. Furthermore, the throughput of the technique has improved by the introduction of

automation and pooling of structurally distinct fragment compounds for analysis. As described with line broadening NMR, an STD response only indicates binding to the protein and is unable to distinguish between binding sites unless additional competition experiments are performed. Additionally, from the STD signal it is difficult to rank the relative binding affinity of fragments to the protein.

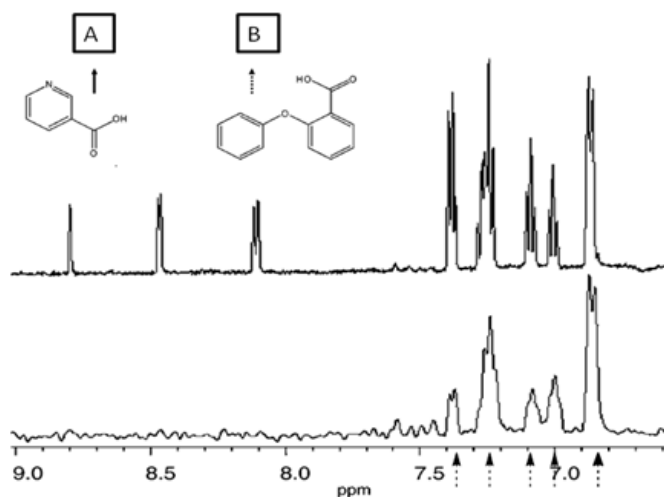


Figure 9 – STD example spectra.⁵⁸ The same compounds are used as in **Figure 8**.

The top NMR is the reference spectrum of a mixture of fragments A and B. The bottom spectra is the STD spectra which shows binding of molecule B to the protein.

Reprinted by permission from Professor Peng; Peng, J. W.; Lepre, C. A.; Fejzo, J.; Abdul-Manan, N.; Moore, J. M. Methods Enzymol. 2001, 338, 202-230, structures added to Figure for clarity.

1.5.2.5 Pooling of Compounds for NMR Screening

NMR screening of a fragment library is a time consuming and resource intensive process. To reduce the timelines, compounds are often grouped into sets of four or five, known as ‘cocktails’. Minimal structural overlap between the compounds in a pool is important. This will reduce the likelihood of multiple ligands binding from the same group and attempt to ensure that the NMR signals for each compound in a cocktail are distinct. If binding of multiple fragments from the same cocktail occurs, interpretation of the data is difficult and the cocktail requires to be deconvoluted to

confirm which fragment is binding to the protein. Cocktails can be used for any of the NMR techniques described above.

1.5.3 X-ray Crystallography as a Primary Screening Technique

X-ray crystallography in drug discovery has historically been limited in its use due to the time required to grow protein apo-crystals, or protein-ligand co-crystals.^{3,60} Diffraction data collection and analysis also require significant time investment.^{3,61,62} Improvements in both methodology and technology have reduced these time constraints. Advances in technology have increased throughput by:

1. development of sample handling equipment, allowing diffraction data to be collected 24 hours a day; and
2. development of automated software (e.g. Autosolve™) to fit ligands to the electron density map generated from the X-ray crystallography.

Improvements in the technology means fragment drug discovery programmes have been able to use X-ray crystallography as a primary screening technique.⁶³ Prior to an X-ray crystallography screen, fragments are often cocktailed, ensuring maximal shape and size distinction between the compounds.⁶⁴ This approach reduces the likelihood of multiple ligands binding, which would make fitting the electron density challenging.

Throughout the fragment optimisation process, it is often found that the key binding interactions and ligand orientation are maintained compared to that of the original fragment hit (**Figure 10**).⁶³ Using X-ray crystallography provides chemists with the binding orientation for each hit, but no way of ranking the affinity of the ligands for the target. Furthermore, in theory, X-ray crystallography can result in false negatives, but it is unable to produce false positives. X-ray crystallography also confirms that the fragments are occupying the binding pocket of interest and not an allosteric site. The disadvantages of running an X-ray crystallography screening campaign are that protein consumption is high, and data collection is expensive. The

X-ray crystallography process is still labour intensive despite recent technological advances.

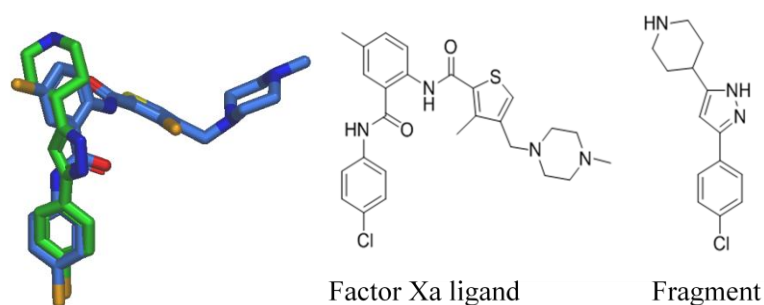


Figure 10 - Overlay of potent Factor Xa ligand (PDB code: 1MQ5; blue) with chlorophenyl fragment hit (PDB Code: 1WAY; green).⁶³ Minimal change to the fragment binding mode is observed.

1.6 Fragment Library Screening

As reported previously, biochemical screening at high concentration is often unreliable, giving a high percentage of false positives, and presumably, negatives. This can be due to poor solubility of the fragments in aqueous buffer at high concentrations, and/or assay interference through fluorescent compounds, or compounds that act as fluorescence quenchers. The information from a biophysical assay is, therefore, essential to establish the validity of a hit and aid fragment prioritisation. Those fragments found to be active post biophysical screening, or in some cases entire fragment libraries irrespective of each fragments binding affinity, are progressed to X-ray crystallography studies. X-ray crystal structures are obtained before embarking on structure-based lead optimisation, thus, completing the FBDD screening cascade.

1.7 Optimisation of the Fragment and Approaches to Monitor Fragment Evolution

Completion of fragment screening is followed by the clustering of hit compounds into distinct series. Strong anchor points (compounds exhibiting good hydrogen bonds or lipophilic contacts) and highly ligand efficient molecules are required before linking, growing, or merging fragments can begin. Growing and merging

have proved to be the most successful fragment optimisation strategies.⁶⁵ Linking has been found challenging because of the difficulty in finding linkers with the appropriate vectors whilst maintaining the optimal binding orientation of the initial fragments.⁶⁵ Whichever optimisation process is used, rigorous discipline to the fragment process is required throughout. The fragment MW and Log P/Log D need to be maintained within the desired range, and compounds need to be grown in an optimal fashion targeting specific interactions. Without discipline, FBDD has no advantage over HTS or ELT methodology, i.e. identification of the fragment hit is only the start of the FBDD process. Care needs to be taken throughout fragment evolution to ensure that, whilst the molecules are grown, the binding efficiency is maintained or improved. Maintenance of the efficiency is pursued by structure guided growth to introduce high quality protein-ligand interactions.

1.8 Small Molecule–Protein Binding Interactions

Protein-ligand interactions are mediated by complex thermodynamic processes. These processes result from the combined interactions made by bulk solvent, ligand and protein. This section describes the thermodynamic factors affecting binding affinity, and how interpretation of these can be used to improve drug molecule development. This is followed by an overview of specific interactions commonly observed between the ligand and protein.

1.8.1 Thermodynamic Binding Contributions

Fragment-based drug discovery routinely necessitates the consideration of thermodynamic contributions to binding. The thermodynamics of protein–ligand interactions is multifactorial, and therefore is not completely understood. Currently, it is difficult to rationalise binding in terms of entropic and enthalpic factors, and drug design would benefit greatly from advances in understanding of this area. The biophysical method, isothermal titration chromatography (ITC) (**Section 1.5.2.3**) is one approach that provides the ability to potentially measure thermodynamic contributions.

The equilibrium binding constant (K_B) describes the ratio of protein-ligand complex [PL] at equilibrium for a reversible reaction between free protein [P] and free ligand [L] (**Equation 4**).⁶⁶

$$K_B = \frac{[PL]}{[P][L]}$$

Equation 4 - Equilibrium binding constant (K_B) equation.

At equilibrium under constant pressure, K_B is linked to the standard Gibbs free energy change, ΔG° (**Equation 5**). The more negative the value of ΔG° , the more favourable the binding interaction. Calculation of ΔG° , and measurement of ΔH° (the heat change on making and breaking of binding interactions), by ITC, allows entropy ΔS° (energetic consequences from changes to the degree of order of the system) to be determined (**Equation 5**).

$$\Delta G^\circ = -RT \ln K_B = \Delta H^\circ - T\Delta S^\circ$$

R = the gas constant (8.314472 J mol⁻¹ K⁻¹), T = Temperature in Kelvin

Equation 5 - Gibbs free energy (ΔG°) calculations.

Literature studies have examined the correlations between ΔG° , ΔH° , and ΔS° for ligand-protein interactions.⁶⁶ No correlation has been made between ΔG° and ΔH° . When ΔH° is plotted against ΔS° , a straight line is obtained due to enthalpy-entropy compensation.⁶⁶ On consideration of protein-ligand binding events, greater ranges are observed for ΔH° and ΔS° , compared to ΔG° .

Changes in entropy and enthalpy are complex to predict as they account for many factors. Accordingly, the inability to predict these values prevents thermodynamics from driving *in silico* design. Enthalpic contributions arise from Van der Waals (VdW), hydrogen-bonding, and electrostatic interactions. Entropic contributions arise from solvation of groups and changes in protein or ligand conformation. On

consideration of the thermodynamics of a ligand-protein binding event, the ligand and protein are both fully solvated. Accordingly, various effects need to be considered such as solvation, ligand/protein flexibility, formation of specific interactions, and structural water re-organisation.

- **Solvation effects⁶⁶⁻⁶⁸**

For a protein-ligand interaction to occur, the ligand and protein must undergo desolvation. The solvent molecules on the protein surface are released into bulk solvent, resulting in an increase in entropy. New lipophilic, H-bonding, and electrostatic interactions are formed between the ligand and protein. The quality, and therefore magnitude, of these enthalpic interactions determines whether the overall process is enthalpic or entropic.

- **Flexibility⁶⁸**

If two ligands make the same pattern of H-bonds and lipophilic interactions, the more rigid molecule will have a greater binding affinity.⁶⁹ This results from the rigid ligand having less entropy, and therefore a reduced penalty to overcome on binding. Protein flexibility is an important factor. Release of protein bound water from the active site, on binding of the ligand, can result in a conformational change of the protein.⁶⁸ The conformational re-organisation gives a change in entropy (from the order/disorder occurring) and enthalpy (from change in the strength of H-bond interactions by improving/disrupting bonding angles and distances).

- **Formation of specific interactions**

The contribution of protein-ligand interactions to binding affinity can differ significantly depending on the interaction type. In general, binding affinity of ligands can be correlated with increasing lipophilicity of the active site.⁶⁶ Therefore, introduction of lipophilic interactions often result in a significant increase in binding affinity, through the release of ordered water from the active site into the bulk solvent. Displacement of water from lipophilic sites, followed

by formation of favourable lipophilic interactions between the ligand and the protein, can often result in high affinity interactions. This is exemplified by the ‘magic methyl effect’ where 100-fold increase in binding affinity is reported to have been gained in some instances.⁷⁰

H-bonding interactions less regularly provide large increases in binding affinity compared to lipophilic interactions.⁶⁸ H-bonding enthalpic effects can be completely cancelled out by the energy required to desolvate the ligand. Therefore, successful formation of H-bonds between the ligand and the protein does not always increase the binding affinity over compounds without these interactions. Burial of polar groups in the active site is disfavoured thermodynamically as protein active sites are usually lipophilic.⁷¹ The polar groups on the ligand interact preferentially with the bulk solvent, rather than with the protein. Therefore, it is more challenging to develop high affinity polar ligands. Analysis of high resolution X-ray crystal structures shows that less than 2% of polar atoms are buried without formation of good quality H-bond interactions.⁶⁸ Since H-bonding interactions are sensitive to the position of H-bond partners (donors and acceptors), considerable care is required to optimise the angle and distance of the interaction. The most highly contributing interactions are achieved when ligand-protein shape and electrostatic charge are matched.

- **Consideration of structural water^{68,72}**

Analysis of water molecules in the binding site can provide valuable information on how to optimise the binding affinity of the ligand to the protein. Replacement of tightly bound water molecules, even with multiple ligand-protein interactions, is disfavoured due to the very specific location and direction of the water-protein hydrogen bond interactions which is challenging to mimic. Consequently, water loss and ligand binding is unlikely to result in improved ligand affinity due to the loss of enthalpy. Replacement of loosely bound water, improves potency through both entropic factors (displacement of water into the

bulk solvent) and enthalpic contributions (H-bond interactions with other water molecules). This is a favourable outcome and often increases affinity, especially if the water molecules originally occupied a lipophilic environment in the protein pocket.

The majority of ligand-protein interactions are entropically driven, however, it has been shown that for ‘best-in-class’ drugs, binding is enthalpically driven.^{73,74} Development of an increased understanding of how to design enthalpically driven drug molecules is at the forefront of current research, but is still widely debated.⁷⁵ The enthalpy of high affinity compounds decreases as ligand size increases.⁶⁶ The entropic effects make greater contributions to binding in larger molecules, and this observation had historically encouraged medicinal chemists to add lipophilicity to the ligand to achieve increased affinity.⁴ In contrast, to maximise the chance of finding enthalpically driven molecules, which give better quality leads, molecular size and physical properties must be carefully controlled.^{74,76} Therefore, FBDD and its emphasis on maintaining low Log P and molecular weight, is likely to have an increased probability of finding enthalpically driven molecules when compared to traditional HTS methodology.

1.8.2 Intermolecular Binding Interactions

Protein-ligand interactions are possible due to formation of a series of intermolecular binding interactions. X-ray crystallography allows the medicinal chemists to infer the nature of the interactions from the electron density and attempt to improve the strength of the ligand-protein complex. An increase in ligand binding affinity is achieved by:

1. optimising existing intermolecular interactions;
2. modifying the ligands structure to exploit new intermolecular interactions;
and
3. minimising repulsive interactions (e.g. two carbonyl oxygen coming into close proximity).

A variety of different intermolecular interactions exist and each must be evaluated in order to understand how to improve the strength and contribution of interactions relating to ligand–protein binding affinity. This section describes the interactions in turn.

1.8.2.1 Hydrogen Bonding (H-bonding) Interactions

H-bond interactions are formed between a donor (hydrogen atom bonded to an electronegative atom) and an acceptor (lone pair on an electronegative atom). There are three parameters critical for ensuring the most productive interaction is formed:

1. The distance between a donor and acceptor atom is critical. The required distance is shorter than the combined Van der Waals (VdW) radii of the donor and acceptor heavy atoms.⁷² The optimal distance is difficult to predict, but recent analysis of the Cambridge structural database (CSD) has provided median distances (2.6-3.2 Å) for different donor and acceptor pairs.^{72,77}
2. The donor-hydrogen-acceptor angle is an important parameter. This varies between donor and acceptor pairs, but for most pairs the desired angle is in the direction of the acceptors lone pairs i.e. for a carbonyl oxygen the desired angle is 120 °. It has been observed that an ether oxygen (weak acceptor) may form H-bonds over a large variation of angles and could be referred to as an accommodation in the active site.^{72,77}
3. The optimal deviation from the plane of the acceptor moiety has been calculated to be between 25-30 °.⁷²

Acceptor strengths vary between different acceptor-donor pairs. A number of measurements have shown that:⁷²

- cyclic amides and ethers are stronger acceptors than acyclic analogues due to increased rigidity;
- electron-donating substituents on heterocycles increase acceptor strength; and
- electron-withdrawing substituents on heterocycles decrease acceptor strength.

Introduction of polar groups to form hydrogen bonding interactions often provide minimal increase in binding affinity due to the desolvation penalty to be overcome (*vide supra*). More lipophilic sulfonyl moieties are capable of forming H-bond interactions but have a reduced desolvation penalty. Consequently, the presence of a sulfonyl group H-bond in a lipophilic part of the protein active site can provide a significant increase in overall binding affinity.

1.8.2.2 Halogen Bonds

Intermolecular halogen bonds are formed by a halogen accepting electron density from an electronegative atom. All four halogens (fluorine, chlorine, bromine, and iodine) are capable of forming halogen bonds. Iodine makes the strongest interaction and fluorine the weakest interaction. The interaction is formed due to a characteristic halo of positive charge being generated on the halogen atom, known as the sigma hole (σ hole) (**Figure 11**). The σ hole is formed from the three unshared pairs of electrons forming a ring of negative charge around the central part of the halogen (for chlorine, bromine, and iodine). This leaves the outer end of the halogen positive. The σ hole is not formed when the halogen is fluorine, as its high electronegativity and sp character causes cancellation of the positive charge.

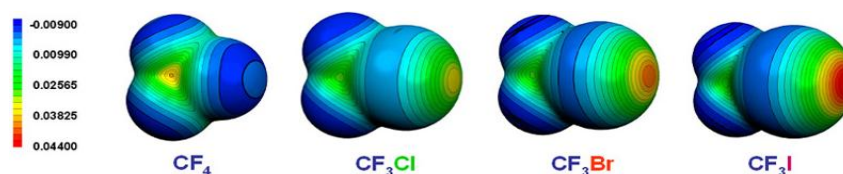


Figure 11 - Electrostatic map showing sigma (σ) hole for all four halogens.⁷⁸ Figure reproduced from Clark, T.; Hennemann, M.; Murray, J. S.; Politzer, P. J. *Halogen bonding: the σ -hole*, *Mol. Model.* 2007, 13, 291-296, with permission of Springer. Figures 1 – 4 were combined into 1 figure.

1.8.2.3 Aryl-Aryl and Aryl-Alkyl Interactions

Weak hydrogen bonds have relatively recently sparked interest as they minimise repulsive interactions between the protein and the ligand.⁷⁹ Aromatic rings have the capability to act as acceptors when donors are positioned perpendicular to the centre

of the aromatic ring, at a distance of ~ 3.5 Å away. The aromatic Tyr, Phe, and Trp amino acid side chains can interact with polarised C-H bonds (donors).⁸⁰ These weak π -interactions have wider tolerance to distance and angle in comparison to the stronger H-bonds. Aryl rings can also interact with each other. Intermolecular interactions between aryl rings can occur through four different ring geometries (**Figure 12**). In the example of benzene, three of the four interaction geometries are of similar stability (parallel-displaced, edge-to-face, and T-shaped), where as the fourth geometry, sandwich, is less stable. The sandwich conformation is less stable due to repulsion of the electron density clouds increasing the distance between the aromatic rings, making the interaction weaker. The aryl-aryl sandwich interaction can be strengthened by modification of the electronics of the ligand aryl ring.⁸¹ Introduction of electron-withdrawing substituents reduces the electron density in the π -cloud, decreasing the electrostatic repulsion between the interacting aryl ring, and hence, strengthening the interaction. Conversely, electron-donating substituents reduce the strength of the interaction by increasing π -cloud repulsion between the rings. Aryl amino acids such as His, Phe, Trp, and Tyr are often involved in aryl-aryl interactions in ligand-protein complexes, with edge-to-face being the most commonly observed alignment.

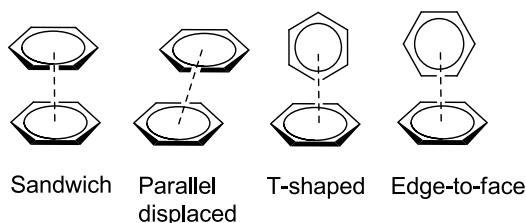


Figure 12 - Diagram of aryl-aryl interactions for benzene.⁸²

Aryl-alkyl interactions, although less frequent, are also observed in ligand-protein interactions, but are often weak in comparison to aryl-aryl interactions.⁸¹ Aryl-alkyl interactions can be strengthened by acidification of the alkyl CH. Cation- π interactions are also possible and can be comparable in strength to H-bonds.

1.9 Maintenance of Physical Properties During Drug Optimisation

To avoid off-target interactions that are caused by non-specific lipophilic interactions, careful monitoring of the physicochemical properties of the molecules is essential. Rigorous control of compound properties is also required to provide a molecule with good ADMET (absorption, distribution, metabolism, elimination, and toxicology) properties. Through consideration of the effect of Log P on ADMET properties, Waring and Leeson have suggested that the optimal Log P range for oral drug molecules is between 1 and 3.^{1,4} As illustrated by **Figure 13**, the median cLog P for patented molecules from pharmaceutical companies is above this threshold. Recent publications in the literature have highlighted and reinforced the importance of such properties and their subsequent effect in compound attrition in the drug discovery process.¹⁵ Therefore, it is important that careful control of properties to reduce lipophilicity and size of drug molecules is practiced. FBDD is an important aspect of this, promoting constant consideration of molecular weight and properties.

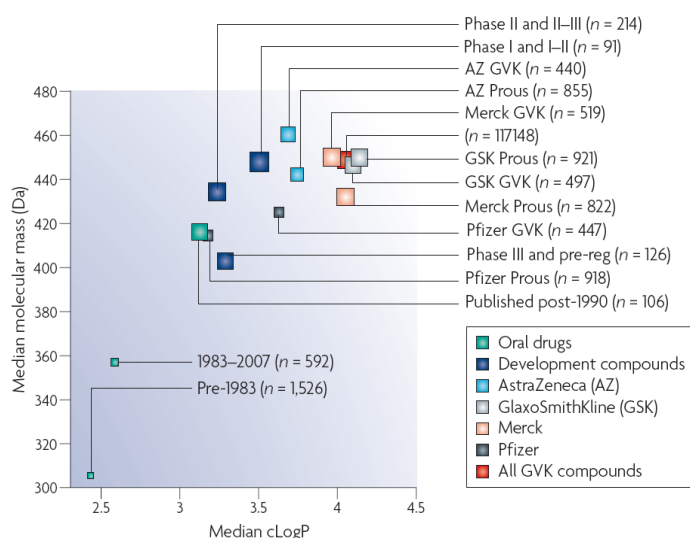


Figure 13 - Comparison of median cLog P and median molecular mass for patented drug molecules from major Pharma.¹ Reprinted by permission from Macmillan Publishers Ltd Leeson, P. D.; Springthorpe, B. *Nat. Rev. Drug Discovery* 2007, 6, 881-890, Copyright 2007.

Many of the fragment-derived compounds published in the literature are targeting kinases.^{3,20,34,83} Traditionally, kinases are one of the more tractable targets due to a well-defined binding pocket and conserved interactions. X-ray crystallography systems can be set-up for a majority of kinases providing structural information to guide the fragment optimisation. Because of this, kinases have been used to develop the FBDD process and demonstrate its advantages over, and complementarity with, other screening approaches.

1.10 FBDD-derived Compounds in the Clinic

The benefits of fragment screening and structure-guided lead optimisation are exemplified in the process used towards an Astex HSP90 inhibitor. By careful optimisation, a 100000 fold increase in activity was achieved by the introduction of only six heavy atoms (compound **(1.1)** to compound **(1.4)**, **Figure 14**). The subsequently accessed compound, AT13387 (**(1.5)**), is currently in phase two clinical assessment for oncology indications. The clinical candidate (**(1.5)**) required the addition of an *N*-methylpiperazine moiety, which was found essential to reduce hERG activity.^{84, 86}

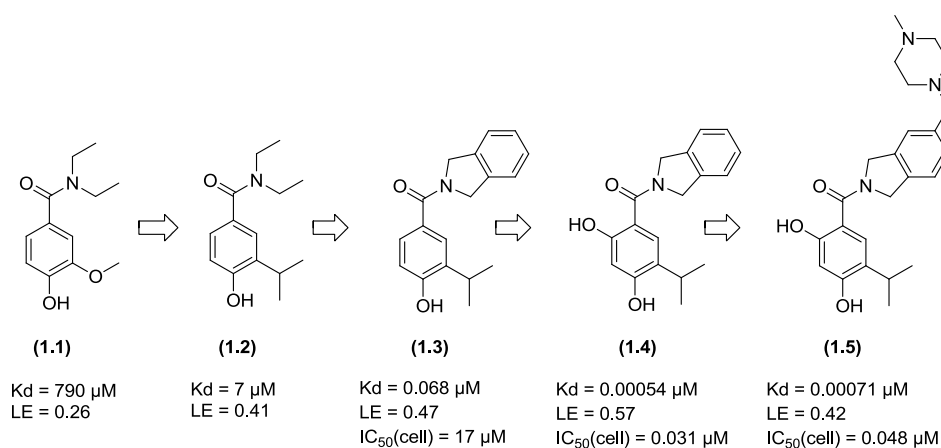


Figure 14 – Astex HSP90 inhibitor AT13387 (**(1.5)**) – Fragment to the clinic.^{84,85}

The successful application of fragment technologies has resulted in fragment derived compounds being tested in all three stages of clinical trials.⁸⁶ The B-Raf V600E inhibitor (PLX-4032) from Plexxikon reached phase three clinical assessment in

2010.⁸⁷ On August 17th 2011, PLX-4032 (branded as zelnoraf) was approved for the treatment of metastatic melanoma. Zelnoraf made history by being the first FBDD-derived molecule to reach the market and gave confidence that FBDD was a suitable and efficient approach to produce safe and efficacious drugs.

The entire development process for zelnoraf (**1.9**) was rapid and progressed from discovery to market in only six years (**Figure 15**). It is important to note that this was an example of an unselective fragment which was evolved into a highly selective drug molecule.^{88,89} Co-crystallisation of the 7-azaindole (**1.6**) fragment hit in Pim-1 (a robust kinase surrogate X-ray crystallisation system), provided a liganded X-ray crystal structure with the fragment bound in multiple conformations in the ATP binding site. Synthesis and biological screening of mono-substituted 7-azaindole fragments, including phenyl (**1.7**), resulted in improved binding affinity and a single constant binding mode as determined by X-ray crystallography in Pim-1. Design of molecules, from analysis of productive interactions made in related kinases, led to 3,5-substituted azaindole (**1.8**). Final rounds of SAR, based on X-ray crystallography in engineered forms of B-Raf, identified chloro-phenyl analogue (**1.9**) which is now marketed as zelnoraf.

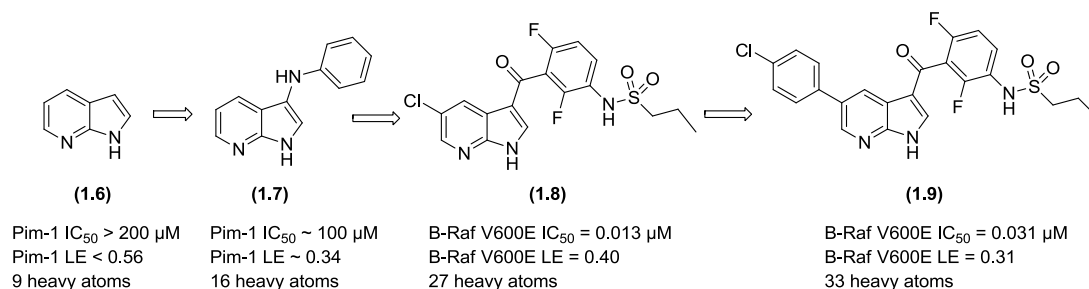


Figure 15 – Development of zelnoraf (**1.9**) from 7-azaindole fragment (**1.6**).⁸⁷

Based on all of the above, there is a growing need to establish how FBDD can impact on human disease. Evolution of the FBDD process has resulted in a number of laboratories progressing molecules through to clinical trials. Having stated this, zelnoraf (**1.9**) remains the only fragment-derived compound to be granted Food and

CONFIDENTIAL – Property of GSK – Do Not Copy

Drug Administration (FDA) approval. With this, the FBDD methodology is now being extended to more challenging protein targets, including protein-protein interactions and membrane bound proteins.^{85,90} FBDD research will undoubtedly continue to evolve with key aspects and advances being identified and reported. Use of FBDD as a screening and hit optimisation technique, alongside traditional lead optimisation strategies, provides a powerful method to identify drug molecules with good physicochemical properties, and with the overall aim of reducing attrition through the drug discovery process.

Using FBDD with disciplined structure led iterative design, the work described in this thesis endeavours to impact on and deliver lead-like molecules for two targets from distinct therapeutic areas. This research also aims to contribute to future advances for this rapidly advancing hit finding methodology.

2 Inhibition of the Protein–Protein Interaction Between HIV Integrase and Lens Epithelium Derived Growth Factor (LEDGF) for the Treatment of HIV-1

2.1 HIV Disease

Infection of individuals by the human immunodeficiency virus type 1 (HIV–1) leads to the contraction of acquired immune deficiency syndrome (AIDS). The number of cases of HIV continues to rise, and currently over thirty-five million individuals worldwide suffer from the virus.⁹¹ In an attempt to combat this disease, the Food and Drug Administration (FDA) has approved thirty-seven drugs (August 2014) over six inhibitor classes. Regardless of this, there is still a need for new therapies.⁹² The unmet medical need arises from undesirable drug-drug interactions; drug-food interactions; non-adherence to therapy; and drug resistance due to mutation of the enzyme target.⁹² To circumvent the issue of resistance caused by protein mutation, design of HIV inhibitors to inhibit a novel target, the protein-protein interaction between HIV integrase and lens epithelium derived growth factor (LEDGF, also known as p75), have been reported.⁹¹⁻⁹³ Targeting this antiviral mechanism is of interest to academic and industrial groups internationally and is the focus of this research chapter.

2.2 HIV Infection

The improved understanding of the HIV lifecycle has identified new viral infection processes to target with small molecules. The viral replication cycle⁹² is initiated through recognition of the viral envelope by the cell surface receptor, cluster of differentiation 4 (CD4) (**Figure 16**). Co-receptors [C-C chemokine receptor type 5 (CCR5) or CXC chemokine receptor type 4 (CXCR4)] engage with the CD4 receptor and promote receptor internalisation within the cell and release of the viral core. The viral core is uncoated giving single stranded RNA (ssRNA) which is converted into double stranded DNA (dsDNA) by reverse transcriptase. HIV integrase (IN) transports the viral DNA into the nucleus by a process referred to as strand transfer (ST), before integrating it into the host DNA. The viral DNA is then transcribed to full length, and the resulting spliced messenger RNA is translated by the cellular

machinery in the cytoplasm to give viral proteins and RNA which assemble into virions. The virions bud (grow from the cell) and undergo proteolytic maturation from HIV protease causing the viral particles to become infectious.

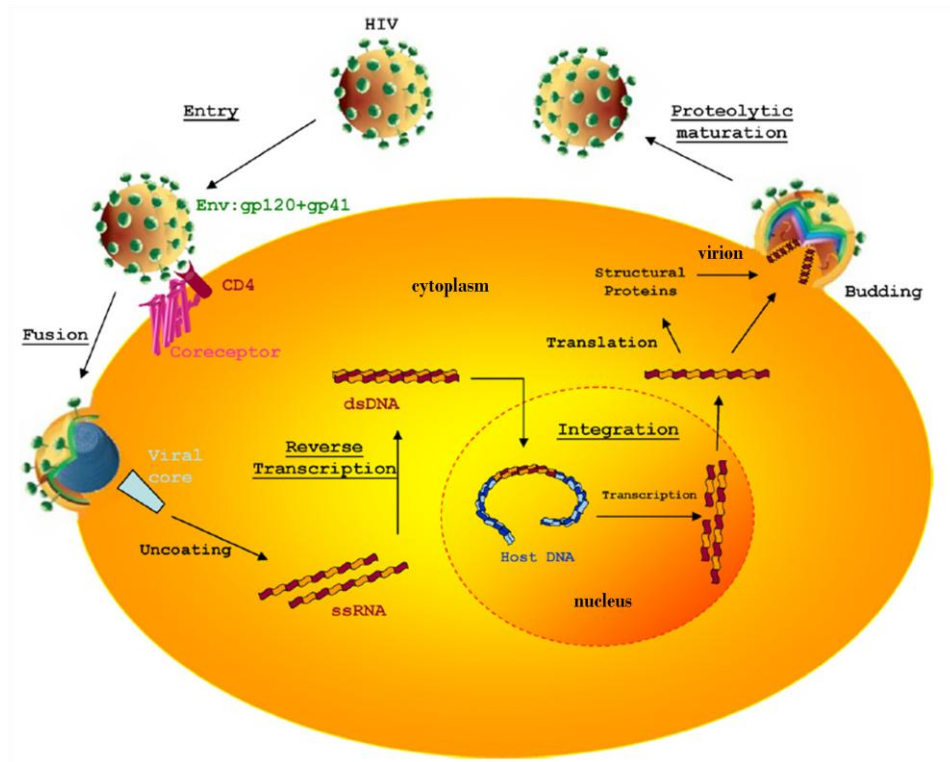


Figure 16 – HIV infection cycle.⁹² Reprinted from *Novel targets for HIV therapy*, Greene, W. C.; Debyser, Z.; Ikeda, Y.; Freed, E. O.; Stephens, E.; Yonemoto, W.; Buckheit, R. W.; Esté, J. A.; Cihlar, T. *Antiviral Res.* 2008, 80, 251-265, Copyright 2008, with permission from Elsevier.

The HIV lifecycle highlights five distinct targets for small molecule intervention: viral entry, fusion, reverse transcription, integration (including ST) and proteolytic maturation.⁹² All of these viral processes have been targeted from a therapeutic standpoint to prevent HIV replication, and approval of multiple drug classes, as well as multi-class combination products, has been successful. However, research continues as the virus mutation rate is fast, giving resistance to the developed therapies.

2.3 Marketed HIV Inhibitors and the Requirement for New Therapies

In 1987 GlaxoSmithKline marketed the first small molecule reverse transcriptase inhibitor (AZT or zidovudine[®] (2.1)) for treatment of HIV (**Figure 17**). Since then, thirty-seven drugs, or drug combinations, spanning the six HIV target classes have been approved (**Appendix i** and **ii**). Continuous research into alternative therapies is fuelled by constant mutation of the virus, leading to enhanced resistance. The challenge of resistance, combined with the undesirable side effects resulting from many of the therapies, gives rise to the unmet medical need which is the driver for research in laboratories globally.

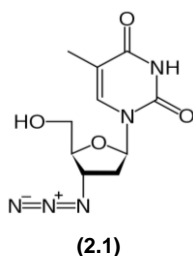


Figure 17 – The first approved HIV small molecule, zidovudine[®] (2.1).

The current regime for HIV is ‘highly active anti-retroviral therapy’ (HAART). The HAART treatment consists of a combination of drugs, from different inhibitor classes, targeting different stages in the HIV lifecycle.⁹³ Unfortunately, HAART often causes side effects ranging from mild complaints, including nausea, vomiting, dizziness, and diarrhoea; but often include more severe complaints including liver failure and neutropenia (low number of neutrophils increasing susceptibility to

infection).⁹⁴ Clinical experience suggests that a harsher treatment regime is required as the disease progresses.⁹³ As the HIV levels in the host increase, the reduction in viral load becomes difficult to sustain through treatment without the use of combination therapies.⁹³

Given the clinical pressures associated with this disease area, efforts continue to identify proteins critical in the viral lifecycle. It is hoped that identification of a new viral target could provide a method of developing anti-HIV therapies, which may not suffer from problems of resistance or side effects. In turn, finding alternative drugs to replace the combination therapies would be a monumental advancement for anti-HIV treatment.

In a more specific sense, targeting the integration step in the HIV viral lifecycle has resulted in approval of raltegravir (**2.2**) by Merck (**Figure 18**).⁹⁵ The small molecule raltegravir (**2.2**) is a well-tolerated, highly potent, oral anti-HIV drug, which binds to the catalytic site of HIV integrase and inhibits ST (see **Section 2.4** for information about the catalytic site). Despite the improved side effect profile when comparing raltegravir to other therapies, it has been difficult to increase drug blood concentrations in patients who show resistance.⁹³ As an alternative research approach to preventing the integration step and overcome resistance, investigations have subsequently included targeting inhibition of the protein-protein interaction between HIV integrase and LEDGF.⁹⁶

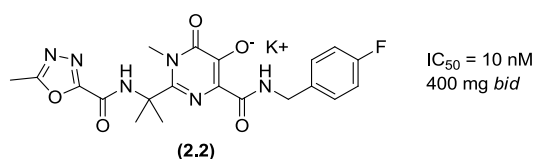


Figure 18 - HIV integrase inhibitor, raltegravir (2.2).

2.4 HIV Integrase-LEDGF Interaction

HIV integrase is a 32 kDa polypeptide, containing 288 amino acids. The enzyme consists of three functional domains.⁹⁷

CONFIDENTIAL – Property of GSK – Do Not Copy

1. the integrase *N*-terminal domain (IN_{NTD}), known as the ‘zinc binding site’, consisting of amino acids 1–50;
2. the integrase core domain (IN_{CCD}), which binds two metals (Zn²⁺ and Cd²⁺), consisting of amino acids 50–212; and
3. the integrase *C*-terminal domain (IN_{CTD}), important for non-specific binding of viral and host DNA, consisting of amino acids 213–288.

Although crystal structures of the individual domains and pairs of domains have been reported, a crystal structure of the full-length protein remains elusive.⁹⁸

The IN_{CCD} is the domain of interest for this programme of work as it contains the LEDGF binding site (**Figure 19**). Isolation of the pre-integration complex (PIC) showed that dimeric IN_{CCD} binds to the integrase binding domain (IBD) of LEDGF (**Figure 20**). The HIV IN-LEDGF protein–protein interaction is essential for complexation of HIV integrase and chromatin during the viral DNA integration process. During this binding event between LEDGF IBD and HIV IN_{CCD}, the structure of either protein is not significantly altered.^{96,99}

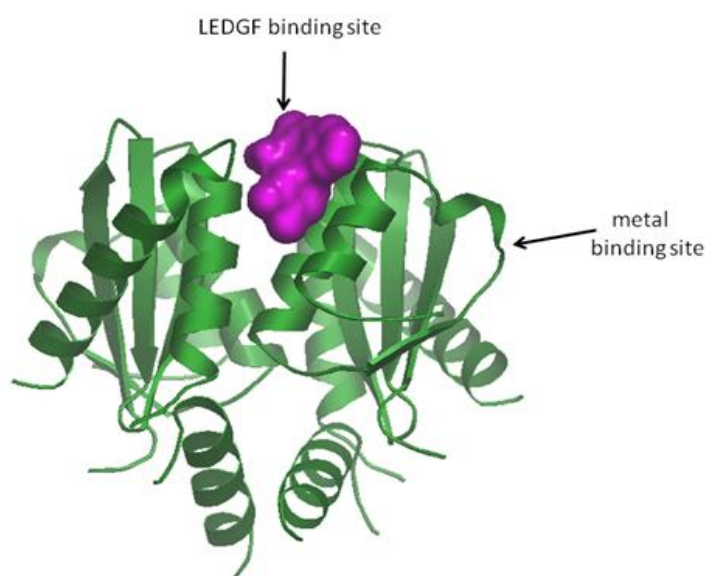


Figure 19 – IN_{CCD} dimer (green) showing the metal binding site and the LEDGF binding site with an unpublished inhibitor bound (magenta).¹⁰⁰

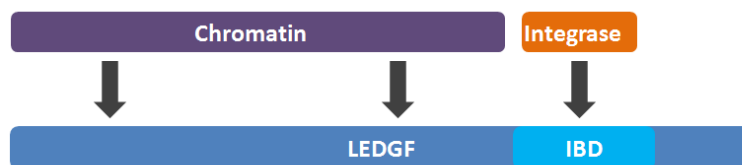


Figure 20 – Schematic of interaction between the LEDGF IBD, HIV IN_{CCD} and chromatin.⁹⁶

LEDGF IBD interacts with HIV IN_{CCD} through recognition of a loop in the IBD. The loop projects into a small hydrophobic pocket, created by the HIV IN_{CCD} dimer (**Figure 21**). Interactions identified as critical for formation of the protein-protein interaction between the LEDGF IBD loop and HIV IN_{CCD} include:¹⁰¹

1. LEDGF Asp366 (D336) forms a bidentate hydrogen bond interaction with the main chain amides of Glu170 (E170) and His171 (H171) on IN_{CCD} (chain A) (**Figure 21**);
2. The LEDGF Ile365 (I365) side chain sits in a lipophilic region at the base of the protein pocket (**Figure 21**) and the backbone amide interacts with the main chain backbone carbonyl of IN_{CCD} Gln168 (chain A);
3. LEDGF Lys364 side chain forms a salt bridge with IN_{CCD} Glu170 side chain (chain A); and
4. LEDGF Leu368 (L368) forms lipophilic interactions with IN_{CCD} Ala128 (A128) (Chain B).

Small molecules to inhibit this protein–protein interaction could bind to either protein, HIV IN_{CCD} or LEDGF. However, HIV IN_{CCD} has been the focus of research as LEDGF is essential for other cellular processes, which if unintentionally inhibited, would result in toxic effects.⁹² No human homolog of HIV IN exists, giving minimal chance of interference with other processes within the body.¹⁰² Another reason for targeting HIV IN_{CCD}, in preference to LEDGF, is that the dimerisation of IN_{CCD} creates a distinct pocket suitable for targeting with small molecules, compared to the protruding loop structure of the LEDGF IBD (**Figure 21**).

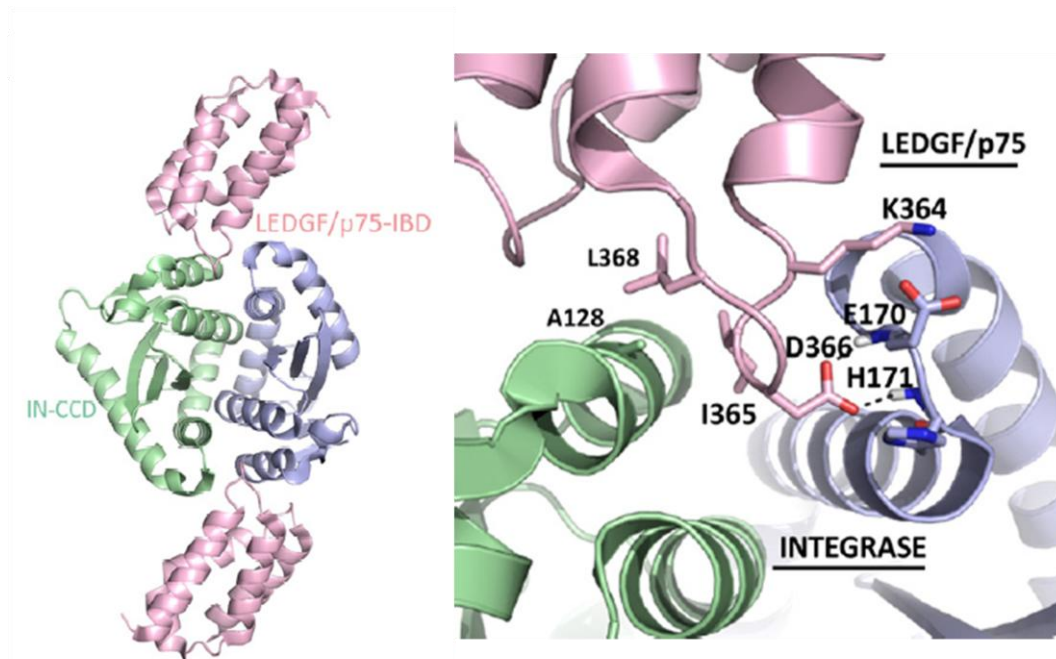


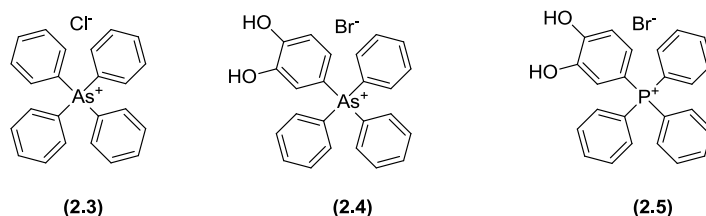
Figure 21 – LEDGF IBD interacting with HIV IN_{CCD} (PDB code: 2B4J). LEDGF IBD is coloured in pink, and the HIV IN dimer is coloured in green and blue. The numbered residues have been shown to be critical in the formation of the protein-protein interaction.¹⁰¹ Reprinted from *The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy*, Christ, F.; Debyser, Z. *Virology* 2013, 435, 102-109, Copyright 2013, with permission from Elsevier.

2.5 Small Molecule Inhibitors of the HIV Integrase-LEDGF Interaction

Protein-protein interactions are considered challenging targets for inhibition with small molecules. The interactions occur over a large area of protein surface (~ 1000 Å²), where a combination of many weak binding interactions combine to give the protein-protein complex.¹⁰³ Large molecules, often peptides, are commonly used to inhibit protein-protein interaction targets.¹⁰⁴⁻¹⁰⁷ Currently, no small molecule inhibitors of HIV IN-LEDGF have reached the market, but it is an area of active research in both industry and academia.¹⁰⁸⁻¹¹³ The exploration within this area has provided a range of structural chemotypes, the most significant of which are described in the following sections.

2.5.1 Tetraphenylphosphonium and Tetraphenylarsonium Inhibitors

In 2001 Molteni and co-workers reported the discovery of tetraphenylarsonium (**2.3**) and (**2.4**), and tetraphenylphosphonium (**2.5**) (**Figure 22**) compounds as HIV integrase binders which were confirmed by crystallography (**Figure 23**).¹¹⁴ The compounds targeted a small, negatively charged, 5 Å deep cleft. The positively charged arsenic or phosphonium centre formed a dipole interaction with the backbone carbonyl of HIV IN_{CCD} Gln168 (chain A). Two phenyl rings formed π -stacking interactions with HIV IN_{CCD} Trp131 and Trp132. Although the small molecules made interactions with HIV IN_{CCD}, the lipophilic quaternary species was resting on the protein surface with minimal protrusion into the IBD pocket. The most potent compound reported in this class is tetraphenylphosphonium (**2.5**) (IC₅₀ of 13.5 μ M) (**Figure 22**).⁹⁹ No advancements in this series have been published which is possibly due to the compounds not being suitable for drug molecule development. The positively charged lipophilic compounds are likely to be highly protein bound, limiting the free fraction available to target the protein-protein interaction. Hence, focus has moved onto exploration of less lipophilic, more drug-like chemotypes.



Inhibition of HIV-1 IC₅₀ = 150 μ M Inhibition of HIV-1 IC₅₀ = 13.5 μ M

Figure 22 – Tetraphenylarsonium (**2.3**) and (**2.4**) and tetraphenylphosphonium (**2.5**) HIV IN-LEDGF inhibitors.

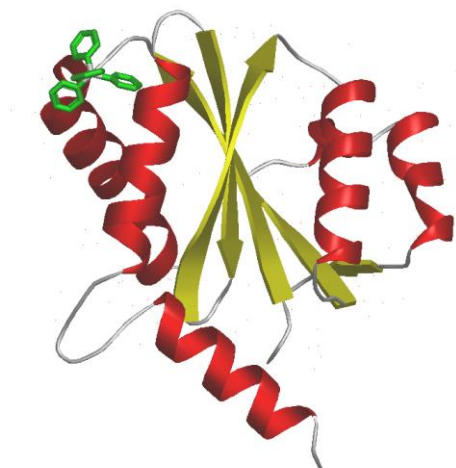


Figure 23 – X-ray crystallography of tetraphenylarsonium derivative (**2.3**) binding to HIV integrase (PDB code: 1HYV).^{114,115}

2.5.2 Synthetic Peptides, including Cyclic Peptide Inhibitor (MZ4-1)

As protein-protein interactions are challenging targets for drug discovery, the strategy of mimicking one of the proteins with a synthetic peptide is a common approach.¹¹⁶ In this regard, the HIV IN_{CCD}–LEDGF interaction is no exception. Short peptides, which bind to HIV IN_{CCD} mimicking LEDGF, have been synthesised. The measured binding affinity of the LEDGF₃₆₁₋₃₇₀ (**2.6**) peptide to HIV IN_{CCD} (5.1 $\mu\text{M} \pm 0.1$) was encouraging (**Figure 24**),^{99,117} but further truncation of the peptide resulted in a loss of binding affinity, demonstrating that each residue plays an important role in the binding of the peptide to HIV IN_{CCD}.¹¹⁷ This finding raises concern that binding to HIV IN_{CCD} with non-peptidic small molecules could be difficult, as interactions over a large surface area appear necessary and may only be obtainable with large molecules (MW ~ 500).

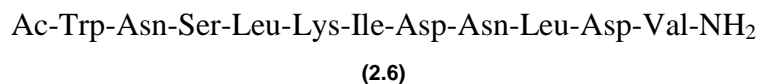
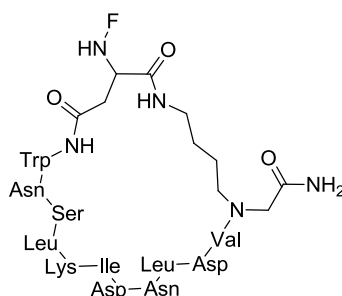


Figure 24 – Structure of LEDGF₃₆₁₋₃₇₀ peptide (**2.6**).

Linear peptides such as LEDGF₃₆₁₋₃₇₀ (**2.6**) are unsuitable as drug molecules as they rapidly undergo proteolytic degradation and have limited cellular absorption profiles.¹¹⁸ In an attempt to overcome this, cyclisation of the peptide was performed. Cyclisation was aimed at improving the peptide's stability, activity, bioavailability and selectivity. A set of cyclised forms of LEDGF₃₆₁₋₃₇₀ (**2.6**) were synthesised in attempts to mimic the bioactive conformation of LEDGF. From this strategy, cyclic peptide c(MZ4-1) (**2.7**) (**Figure 25**) was identified and which displayed comparable activity to linear of LEDGF₃₆₁₋₃₇₀ (**2.6**).¹⁰⁸ Subsequent biological testing indicated that cyclic peptide c(MZ4-1) (**2.7**) had low micromolar activity, with improved stability compared to linear LEDGF₃₆₁₋₃₇₀ (**2.6**), and exhibited antiviral effects eight days post infection.



(2.7)

Figure 25 – Structure of fluorescein (F) labelled cyclic peptide c(MZ4-1) (**2.7**).

Despite encouraging antiviral results from cyclic peptide c(MZ4-1) (**2.7**) and the knowledge that many interactions along the HIV IN_{CCD} surface are critical for potency, identification of small molecule inhibitors has still been pursued. Examples of these are described in the following sections.

2.5.3 Thiazolidin-5-ylidene Analogues (D77)

Thiazolidin-5-ylidene analogue (D77) (**2.8**) was the first non-peptidic small molecule reported to inhibit the HIV IN_{CCD} – LEDGF interaction (**Figure 26**).¹¹⁹ Molecular docking into a crystal structure of HIV IN_{CCD} bound to LEDGF (PDB code: 2B4J), showed that D77 (**2.8**) bound to HIV IN_{CCD} in the LEDGF IBD recognition site

(Figure 27) and did not disrupt the HIV IN_{CCD} dimerisation. Antiviral activity of D77 (2.8) is reported as EC₅₀ = 23.8 µg/mL in MT-4 cells.

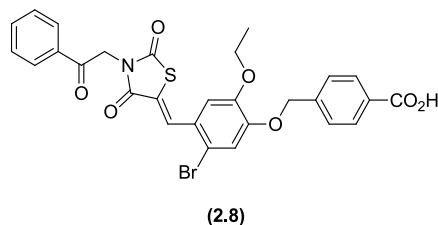


Figure 26 – Structure of thiazolidin-5-ylidene analogue (D77) (2.8).

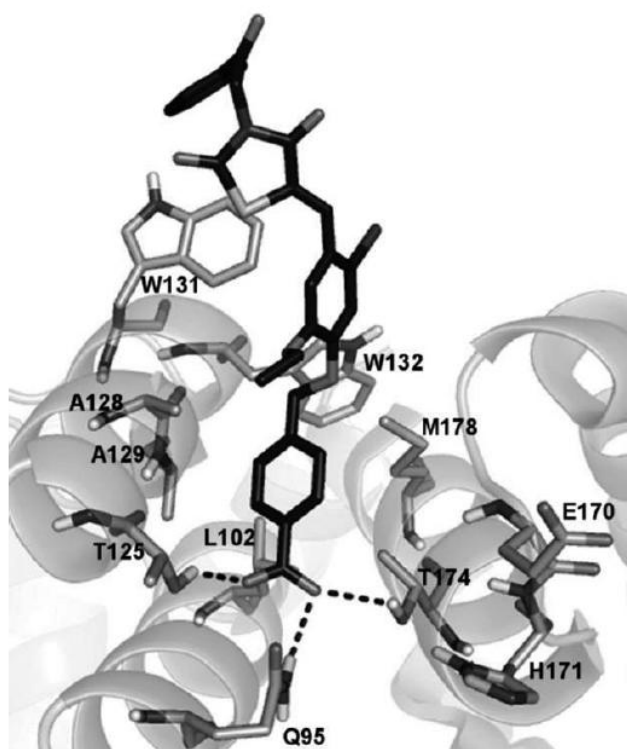


Figure 27 – Docking model of thiazolidin-5-ylidene analogue (D77) (2.8) in HIV IN_{CCD} dimer (PDB code: 2B4J).⁹⁹ Reprinted from *Inhibitors of the Interactions between HIV-1 IN and the Cofactor LEDGF/p75*, De Luca, L.; Ferro, S.; Morreale, F.; De Grazia, S.; Chimirri, A. *Chem. Med. Chem.* 2011, 6, 1184-1191, Copyright © 2011 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, with permission from Wiley.

The docking model predicted that D77 (2.8) is bound in the hydrophobic pocket created by Leu102, Ala129, Ala128, Trp131 and Trp132 from HIV IN_{CCD} chain A

and Met178 from chain B. Specific interactions were predicted between the carboxylic acid group of D77 (**2.8**) with the side chain of Thr174, Gln95 and Thr125. Hydrophobic interactions were observed between Trp131 and the two aryl rings either side of the thiazolidine ring. Despite the encouraging antiviral activity, further tests revealed that the thiazolidine template had cytotoxic effects at 128.8 μM in MT-4 cells, preventing the progression of D77 (**2.8**) towards market.

2.5.4 Quinoline Derivatives

Debyser and co-workers have reported the use of virtual screening to identify novel small molecule inhibitors of the HIV IN_{CCD}-LEDGF interaction.⁹¹ A set of 200,000 commercially available compounds were filtered using average chemical properties of known small molecule protein-protein inhibitors. The remaining 160,000 compounds were screened against a pharmacophore model created from analysis of the X-ray crystal structure of HIV IN_{CCD} bound to LEDGF IBD. The 25 highest scoring compounds, based on docking interactions, were selected for biochemical evaluation, identifying quinoline analogue (**2.9**) (36% inhibition of LEDGF at 100 μM in a bead-based Alpha Screen PPI assay) (**Figure 28**). Subsequent, analogue searching and consideration of the tetrazole bioisosteres led to carboxylic acid analogue (**2.11**) which had an enzymatic IC₅₀ of $1.37 \pm 0.36 \mu\text{M}$ and evidence of antiviral activity in an MTT/MT4 cell viability assay (EC₅₀ = $2.35 \pm 0.28 \mu\text{M}$) at concentrations where no cytotoxic effects were observed (CC₅₀ = $59.8 \pm 0.50 \mu\text{M}$).¹²⁰

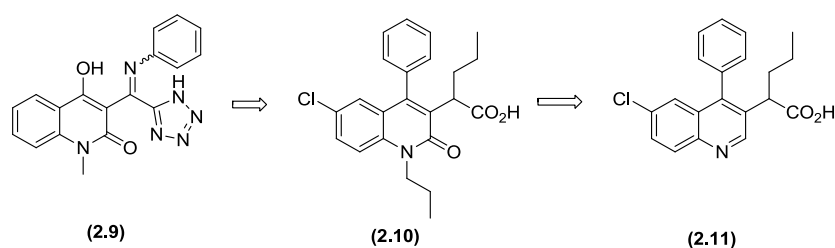


Figure 28 – Quinoline analogues identified from a pharmacophore screening.⁹¹

X-ray crystallography of carboxylic acid (**2.11**) (**Figure 29**) demonstrated excellent overlap with the docked pose previously established by molecular modelling.

Hydrogen bonds from the carboxylic acid group of analogue (**2.11**) to Glu170, His171 and Thr173 were observed. In addition, the phenyl ring was buried in a hydrophobic region, and the quinoline fused phenyl rests on the helix of chain B (Ala128), which is often referred to as a lipophilic shelf. It was noted that the small molecule occupies the space filled by the protruding loop of LEDGF (**Figure 29, Diagram b**).

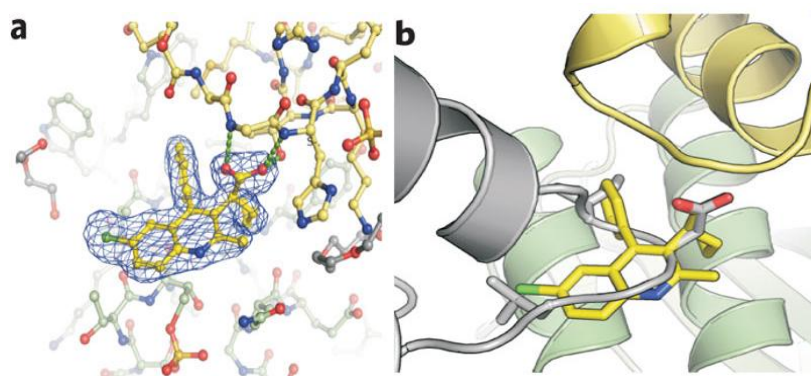


Figure 29 – Crystal structure of quinoline analogue (**2.11**) in yellow.⁹¹ **Diagram a** – shows electron density of quinoline (**2.11**); **Diagram b** – shows HIV IN_{CCD} dimer in yellow and green with LEDGF superimposed in grey (PDB code: 2B4J). Reprinted by permission from Macmillan Publishers Ltd: Christ, F.; Voet, A.; Marchand, A.; Nicolet, S.; Desimie, B. A.; Marchand, D.; Bardiot, D.; Van der Veken, N. J.; Van Remoortel, B.; Strelkov, S. V.; De Maeyer, M.; Chaltin, P.; Debyser, Z. *Nat. Chem. Biol.* 2010, 6, 442-448, Copyright 2010.

Further modifications to the quinoline core identified thiophene analogue (**2.12**) (**Figure 30**) showing enzymatic IC₅₀ of 0.58 ± 0.30 μM, with excellent antiviral activity (EC₅₀ = 0.76 ± 0.08 μM and CC₅₀ = 72.2 ± 5.15 μM).⁹⁶ More recently, Debyser and co-workers reported the discovery of *t*-butyl ether analogue (**2.13**), which was the first small molecule in this series to display antiviral activity in the low nano-molar range (EC₅₀ = 0.069 ± 0.003 μM and CC₅₀ = 96.0 ± 16.0 μM).¹²⁰ Importantly, the *t*-butyl ether (**2.13**) inhibits two distinct steps in the integration process: the HIV IN_{CCD}-LEDGF protein-protein interaction; and ST (IC₅₀ = 146 nM). Encouragingly, antiviral activity has been shown to be unaffected by known HIV IN

resistant mutants, suggesting that inhibition of this viral target could give benefits over HAART.

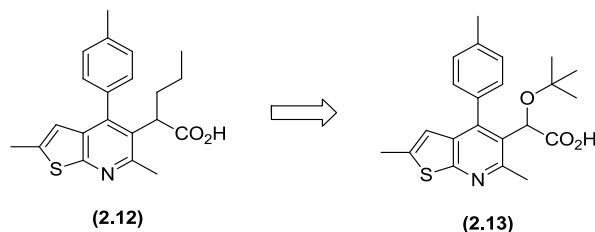


Figure 30 – Thiophene *t*-butyl ether analogue (**2.13**).¹²⁰

2.5.4.1 Boehringer Ingelheim Quinoline Analogues and Related Chemotypes

Researchers at Boehringer Ingelheim have identified related quinoline analogues through a HTS campaign that inhibit the HIV integration process.¹²¹⁻¹²³ As previously described, HIV Integrase catalyses the insertion of viral cDNA into infected cells. This process requires HIV Integrase to trim the 3'-ends of the viral DNA, resulting in reactive intermediates necessary for strand transfer (ST). HTS hit quinoline (**2.14**) (**Figure 31**) was found to inhibit the 3' processing step (3P) where a guanine-thymine (GT) dinucleotide is removed from both 3' DNA ends ($IC_{50}^{3P} = 9000$ nM).¹²² Medicinal chemistry efforts resulted in optimisation to ether analogue (**2.15**), which was confirmed by X-ray crystallography to bind in the HIV IN_{CCD} IBD pocket and inhibit both the 3P mechanism ($IC_{50}^{3P} = 28$ nM) and the binding of HIV IN_{CCD} to LEDGF ($IC_{50} = 1500$ nM).¹²² Significant optimisation work resulted in identification of clinical candidate BI224436 (**2.16**) with biochemical potency $IC_{50}^{3P} = 15 \pm 4$ nM and inhibition of the HIV IN_{CCD}-LEDGF interaction with $IC_{50} = 11 \pm 1$ nM.¹²³ *In vitro* cellular activity for (**2.16**) was encouraging with $EC_{50} < 15$ nM in peripheral blood mononuclear cells (PBMC's) across a series of integrase containing viruses.¹²³ Despite the pre-clinical profiling and early human dose escalation study in males looking promising, the clinical trial was put on hold and the programme was licensed to Gilead Sciences from where additional patent applications have since been filed.^{124,125} It is clear that the quinoline template is the most promising series of HIV IN_{CCD}-LEDGF inhibitors to date. Numerous analogues of the quinoline template have been published or patented by laboratories internationally, which also

highlights the current interest and importance of identifying inhibitors for this target
(**Figure 32**).¹²²

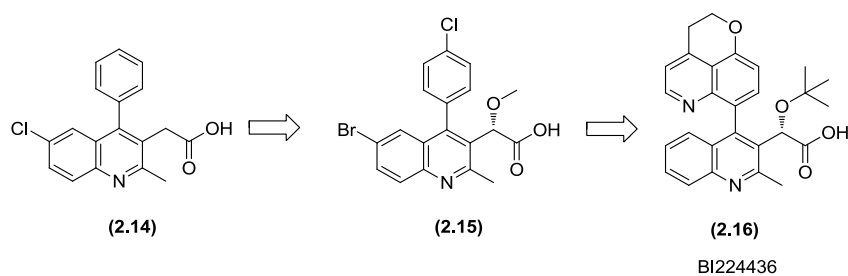


Figure 31 – Optimisation of HTS hit quinoline (**2.14**) resulted in identification of clinical candidate BI224436 (**2.16**).

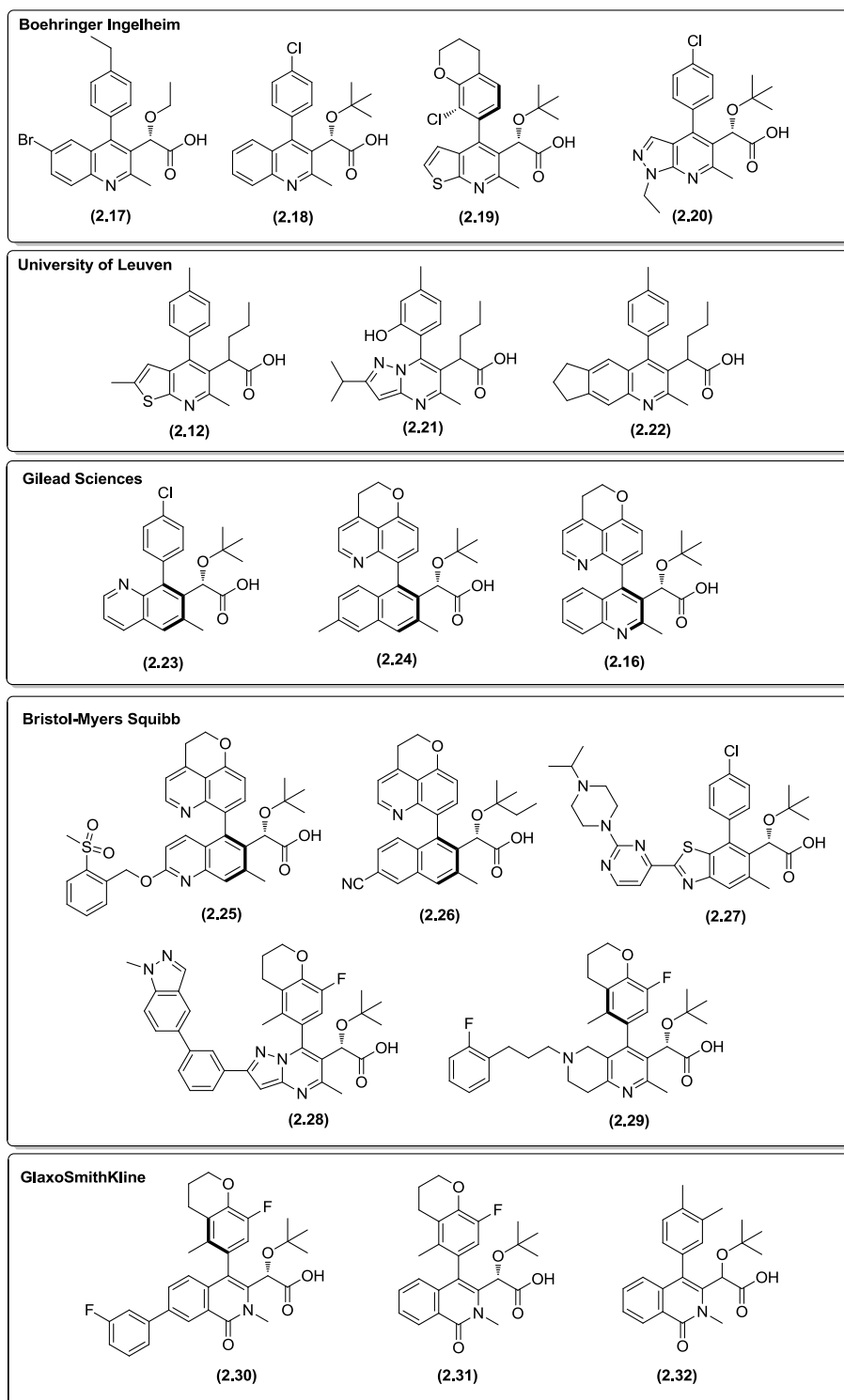


Figure 32 – Examples of patented molecules, with similar pharmacophoric features to the quinoline template, from laboratories globally. Patent examples have been taken from a review published by Debyser *et al.*¹²²

2.6 Summary

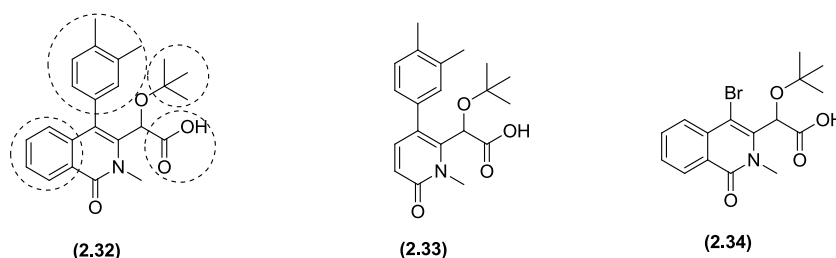
To summarise, despite recent advances there is still a significant unmet medical need for the treatment of HIV-1. This arises from the severe side effects experienced by a high percentage of patients, and increasing resistance to current therapies.

By January 2012, twenty-five drugs had been approved by the FDA, increasing to thirty-seven by August 2014. The approved drugs target five distinct processes in the HIV-1 lifecycle. To complement these approved drugs, new processes in the HIV-1 lifecycle are being targeted, including the HIV IN_{CCD}–LEDGF protein-protein interaction. It is proposed that targeting this protein-protein interaction may overcome the problem of resistance. No human homolog of HIV IN exists, giving minimal chance of interference with other cellular processes, thus making this the partner of focus in the interaction.

Multiple classes of HIV IN_{CCD}–LEDGF inhibitors have been reported in the literature, including peptides, small molecules, and, more recently, dual mechanism inhibitors. The recent explosion in patent literature around the quinoline pharmacophore demonstrates that inhibition of the HIV IN–LEDGF is feasible and is of significant interest within the wider scientific community. Targeting the HIV IN_{CCD}–LEDGF protein-protein interaction to find novel inhibitors is the focus of this research programme, the outcomes of which are described in the remainder of **Chapter 2**.

2.7 Background and Aims for Inhibition of the HIV IN-LEDGF Protein-Protein Interaction: Identifying Novel Inhibitors in Improved Physicochemical Property Space

There is clear evidence that the HIV IN_{CCD}-LEDGF target is chemically tractable.^{13,92,93,117} However, many of the reported inhibitors have a common pharmacophore which provides a very rigid structure, filling four remote regions of the LEDGF binding site on HIV IN_{CCD}. Fragmentation of quinolinone (**2.32**) demonstrates the importance of the two peripheral aromatic rings (**Figure 33**). All molecules within this pharmacophoric class have the carboxylic acid moiety which forms the only specific H-bond interactions between the ligand and HIV IN_{CCD}; all other interactions are lipophilic. The majority of the patented molecules contain the *t*-butyl ether moiety (**Figure 32**), which provides a significant potency improvement when introduced into the thiophene template (**2.13**) (**Figure 30**).



pIC₅₀ 6.6, LE = 0.31, LLE_(AT) = 0.23 pIC₅₀ = 5.0, LE = 0.27, LLE_(AT) = 0.23 pIC₅₀ = 3.9,^a LE = 0.24, LLE_(AT) = 0.19

Figure 33 – Fragmentation of quinolinone (**2.32**) demonstrates a significant reduction in potency by removal of each of the aromatic rings. Four remote groups (circled on quinolinone (**2.32**)) appear essential to achieve the required level of HIV IN_{CCD}-LEDGF inhibition. ^a This compound has only been reported active in 2 out of 3 screening occasions – 3rd screening occasion measured < 4.3.

Within the medicinal chemistry community, PFI is used as a ‘guide’ to aid discovery of clinical candidates, leading to reduced chance of attrition in the clinic (Section 1.2.2). Calculation of the cPFI for the patented HIV IN_{CCD}-LEDGF inhibitors shown in **Figure 32**, shows that the cPFI is greater than 6 in all cases, with indazole analogue (**2.28**) having a cPFI exceeding 10 (**Table 1**). Maintaining PFI below 6 is

CONFIDENTIAL – Property of GSK – Do Not Copy

recommended, therefore, identification of inhibitors with improved PFI would be of significant interest.

Cmpd No.	Patent No.	cPFI	MW
2.28	WO2012033735 ¹²⁶ WO2013025584 ¹²⁵ WO2013134142 ¹²⁷	10.7	634
2.25	WO2013103724 ¹²⁸	9.4	627
2.30	WO2012102985 ¹²⁹	8.7	548
2.27	WO2013159064 ¹³⁰	8.5	594
2.24	WO2012003497 ¹³¹	8.5	456
2.26	WO2013103738 ¹³²	8.1	481
2.29	WO2013123148 ¹³³	7.8	579
2.22	WO2012065963 ¹³⁴	7.7	373
2.12	WO2010130842 ¹³⁵	7.7	353
2.17	WO2007131350 ¹³⁶	6.8	428
2.20	WO2009062308 ¹³⁷	6.7	402
2.23	WO2012003498 ¹³¹	6.6	384
2.19	WO2009062288 ¹³⁸	6.4	446
2.16	WO2012138669 ¹³⁹ WO2012138670 ¹²⁴	6.4	442
2.18	WO2009062285 ¹⁴⁰	6.3	384
2.21	WO2011076765 ¹⁴¹	6.3	381
2.32	WO2012102985 ¹²⁹ WO2013002357 ¹⁴²	6.3	393
2.31	WO2012102985 ¹²⁹ WO2013002357 ¹⁴³	6.2	454

Table 1 – MW and cPFI for patented HIV IN_{CCD}-LEDGF Inhibitors shown in **Figure 32**.

Novel inhibitors of the HIV IN_{CCD}-LEDGF protein-protein interaction have been discovered within our laboratories through complementary screening techniques

CONFIDENTIAL – Property of GSK – Do Not Copy

including, HTS, ELT, and computational-based approaches. Hits generally lacked polar functionality ($cLog P > 3$), correlating with the lipophilic nature of the HIV IN_{CCD}-LEDGF binding pocket. Subsequent sub-structure searches of our laboratories' compound collection, to discover new chemical frameworks with improved physicochemical properties, failed to identify hits with clear SAR. Additionally, the hits could not be structurally validated as HIV IN_{CCD} binders. The combination of a lack of structural information, with unclear SAR, cast doubt over the validity of the hits obtained. As a result, further optimisation efforts on HTS and ELT hits were abandoned.

Failure to identify novel HIV IN_{CCD} binders with improved physicochemical properties led to consideration of an FBDD screening approach. It was envisaged that FBDD would aid in identification of smaller, more polar, HIV IN_{CCD}-LEDGF inhibitors with an improved PFI, in comparison to hits identified through HTS and ELT screening. However, the lack of novel chemical structures reported in the literature; the large rigid lipophilic structure of known inhibitors; and the remote nature of the areas of the HIV IN_{CCD} pocket that appear to be required to be filled to gain inhibition, caused significant concern regarding the suitability of FBDD for this target.

Despite an FBDD approach being viewed as high risk, if successful it would be of extremely high value to the worldwide HIV research and patient population. Furthermore and importantly, this study will also inform the FBDD community of the potential for using such a strategy to identify novel inhibitors in improved property space for challenging targets such as HIV IN_{CCD}-LEDGF.

2.8 Project Aims

The FBDD campaign to identify novel inhibitors of the HIV IN_{CCD}-LEDGF protein-protein interaction with improved PFI over published analogues was pursued within our laboratories.

The aims of this specific research programme are:

- identification of an appropriate screening cascade through collaboration with assay scientists;
- biochemical and biophysical data analysis to triage fragment hits through the screening cascade;
- clustering and ranking of fragment screening hits by consideration of physicochemical properties, and biochemical, biophysical, and X-ray data;
- formulation of plans for optimisation of fragment hits through sub-structure search (SSS) and synthesis;
- evaluation and progress tracking of the fragment hit optimisation;
- continuous analysis of biochemical, biophysical, and X-ray crystallography data for each fragment hit; and
- identification of a novel chemical series for inhibition of HIV IN_{CCD}-LEDGF with improved properties (PFI < 6) compared to literature examples.

This was amongst the first attempts to utilise FBDD screening in our laboratories. Accordingly, this heightens the importance of recording of successful and unsuccessful strategies for improving the success of future FBDD efforts both within our laboratories and globally.

The candidate has performed all preparative chemistry, analysis of data, and review of X-ray crystal structures to form design ideas, unless referenced otherwise.

2.9 Library Construction, Fragment Screening and Data Analysis

2.9.1 Library Construction

Identification of fragment hits suitable for development into lead molecules with good physicochemical properties, is critical for success of the project. The FBDD process requires scientific rigour throughout and must begin with careful selection of compounds from which to form the fragment library. Two distinct fragment screening sets are available for FBDD within our laboratories: a fragment library and a ¹⁹F library. As a consequence of increased knowledge regarding the properties of fragment hits, the fragment library has continuously evolved since its inception.^{30,144,145} The fragment library screened against HIV IN_{CCD}-LEDGF was constructed by consideration of commercially and internally available compounds. Compounds with physical properties outside the desired range were removed using a set of filters derived from unpublished knowledge from within our laboratories (**Figure 34**).¹⁴⁶

$$\begin{aligned}9 &\leq \text{heavy atom count} \leq 17 \\-2 &\leq \text{cLog P} \leq 2 \\ \text{rotatable bonds} &< 4 \\1 &\leq \text{number of rings} \leq 3 \\ \text{aryl rings} &\leq 2 \\ \text{HBA} &< 8 \\ \text{HBD} &< 4 \\ \text{HBA} + \text{HBD} &< 8 \\ \text{basic and acidic groups} &< 3 \text{ in total}\end{aligned}$$

Figure 34 - Filters for selection of compounds with appropriate properties for inclusion in the fragment library.¹⁴⁶

CONFIDENTIAL – Property of GSK – Do Not Copy

The next step was to further refine the set of molecules by computationally removing compounds with unwanted functionality, reactive groups and toxicophores, resulting in a set of ~ 11000 compounds.¹⁴⁶ This set was reviewed within our laboratory to prioritise structures that deemed suitable for medicinal chemistry follow-up, reducing the set to ~ 6000 fragments. The reduction in number of compounds through this process, resulted in the set being reduced to a suitable number of fragments on which to perform quality control experiments. Quality control (purity > 98%) experiments included: LCMS; NMR; light scattering (to test solubility); and SPR against four distinct targets to assess promiscuity. These quality filters reduced the compound set to 3500. This original fragment screening set (3500 compounds) was later reduced to a 'core set' of 1056 compounds ensuring maximum structural diversity. This 'core set' was first used in this programme and is now the core set for all FBDD efforts within our laboratories.

An alternative proprietary library, the ¹⁹F library, was historically designed¹⁴⁷ to demonstrate the applicability of ¹⁹F NMR screening to the drug discovery process. The ¹⁹F library consists of 1465 compounds, each containing fluorine. In contrast to the construction of the fragment library, strict filters were not applied as this set was not specifically designed for identification of fragment starting points. The set broadly encompasses Lipinski's 'rule of 5'¹⁴⁸ (**Figure 35**) and has an average molecular weight of 252.

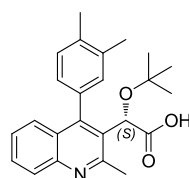
hydrogen bond donors ≤ 5
hydrogen bond acceptors ≤ 10
MW ≤ 500 Daltons
Log P ≤ 5
number of rotatable bonds < 10

Figure 35 – Lipinski's 'rule of 5'.¹⁴⁸

2.9.2 Identification of Appropriate Screening Methodologies

A carefully designed fragment library only provides the desired benefits if it is screened appropriately. A well constructed fragment screening cascade is critical for the success of any fragment effort. For the HIV IN-LEDGF project, consideration of both biochemical and biophysical screening techniques was pursued. These techniques were tested by relevant scientists for applicability and sensitivity to aid in cascade design.^{149,150}

NMR screening was found to be the biophysical method of choice, as the spectral shifts and line-widths were indicative of a well-folded protein, and the signal is stable over a 12 hour period.¹⁴⁹ In addition, line broadening NMR and STD, could be used in competition format with quinoline (**2.35**) (**Figure 36**), to confirm the binding location of the fragment. Thermal melt (T_m) and SPR methods were deemed unsuitable. Despite T_m ¹⁵¹ providing large T_m shifts of ~ 20 °C for potent ($pIC_{50} = 6 - 7$) inhibitors, only small shifts in the thermal melt curve ($2 - 3$ °C) were observed for weaker compounds ($pIC_{50} < 5$). These small shifts were difficult to separate from experimental error. Preliminary investigation into SPR highlighted problems in identifying an appropriate linker to immobilise the protein to the SPR plate.¹⁵²



(2.35)

$pIC_{50} = 6.6$ LE = 0.32 LLE_(AT) = 0.17
cLog P = 5.3
cPFI = 6.43

Figure 36 – Representative quinoline molecule (**2.35**), the *S*-enantiomer of (**2.32**).

For routine biochemical screening a TR-FRET biochemical assay¹⁵³ was used. The assay used His₆-HIV IN (50-212, F185K), GST-LEDGF IBD (349-427), anti-His₆-APC, and anti-GST-Europium (**Figure 37**). The standard top concentration of 100 μ M ($pIC_{50} \geq 4$) was not sensitive enough for detection of weak binders. Therefore,

investigation into improvement of the sensitivity of the TR-FRET assay was required. Currently our laboratory's main compound collection is stored as 10 mM DMSO stock solutions. It was known that the TR-FRET assay tolerated 1% DMSO concentration, and on further investigation was shown to tolerate up to 5% DMSO, with no reduction in accuracy or sensitivity. The optimised assay¹⁵³ enables a top concentration of 500 μ M ($pIC_{50} \geq 3.3$) which was deemed sufficient for identification of binding of weak fragments. A counter-screen was run in parallel using His₆-GST in place of His₆-HIV IN (50-212, F185K) and GST-LEDGF IBD (349-427) for detection of assay interferers (**Figure 37**). Assay interference could result from fluorescence quenchers, fluorescence emitters, and binding to the GST-His₆ tags. Once compounds achieve biochemical potency of $pIC_{50} \geq 6.0$, with no activity in the counter-screen, they were progressed to a panel of cellular assays.

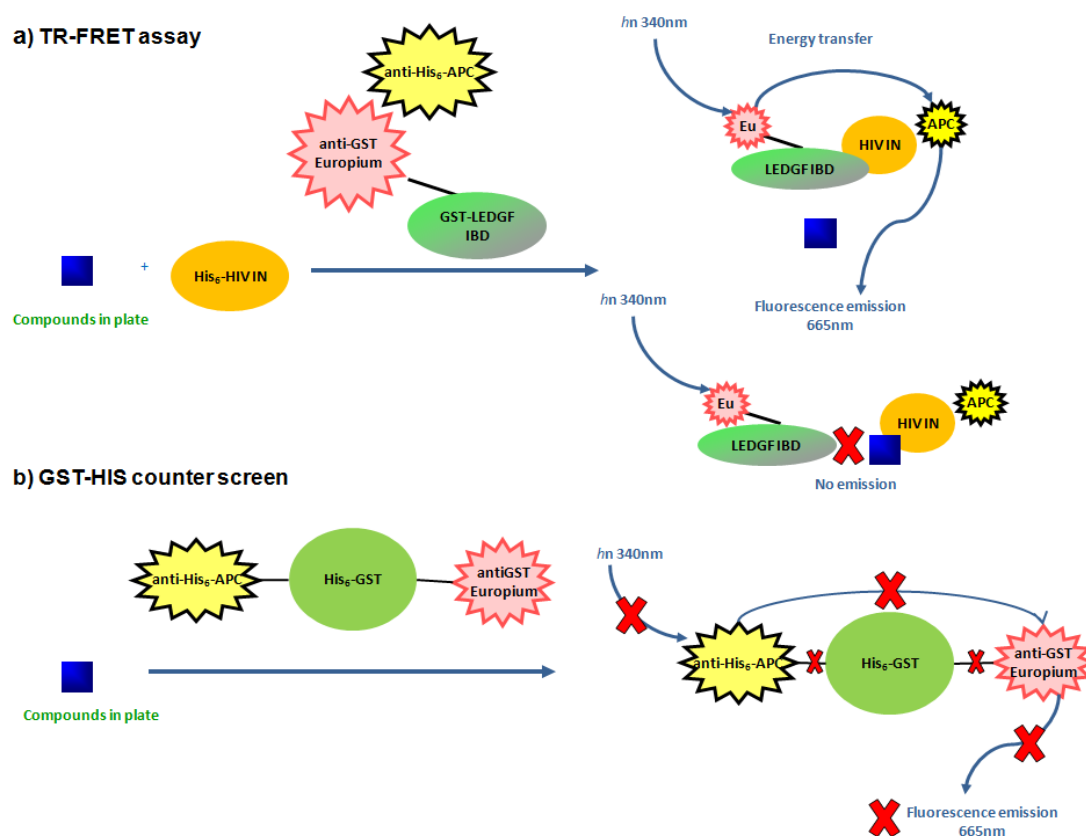


Figure 37 - a) Schematic of TR-FRET assay; **b)** Schematic of GST-His₆ counterscreen.¹⁵³

2.9.3 Screening of the ¹⁹F and Fragment Libraries

Based on analysis of the applicability of available screening techniques, the strategy adopted for this programme was to treat the two libraries (¹⁹F and fragment library) independently. The smaller library, the ¹⁹F library (1465 fluorine containing compounds), was screened first to assess the potential hit rate from the biochemical (TR-FRET)¹⁵³ and biophysical (NMR)¹⁴⁷ assays. As the fragments were expected to bind with weak affinity to HIV IN_{CCD}, biophysical and biochemical techniques were used in parallel to increase the chances of detection of these weak μM - mM hits. It was anticipated that the most efficient, but lowest molecular weight compounds, would not be detected by biochemical screening alone, as they would be below the limit of the assay (mM binders). In response to this concern, the ¹⁹F library was screened against HIV IN_{CCD} using NMR ¹⁹F line broadening.¹⁴⁷ It is important to realise that observation of line broadening only shows that the ligand is interacting with the protein, and does not confirm binding at the HIV IN_{CCD} IBD pocket. Therefore, compounds were further triaged by attempted reversal of the line broadening, by addition of the quinoline analogue (**2.35**), a known binder at the HIV IN_{CCD} IBD pocket (**Figure 38**). The line broadening experiment was performed by recording an NMR of the ligand e.g. piperidine (**2.36**) (top NMR in **Figure 38**) prior to addition of 20 μM HIV IN_{CCD}. Binding of piperidine (**2.36**) to HIV IN_{CCD} was observed by line broadening (see peaks with red arrows in **Figure 38**). To demonstrate that the binding of the ligand was to the desired site, the HIV IN_{CCD} IBD, quinoline (**2.35**) was added to compete with piperidine (**2.36**) and reverse the line broadening (bottom spectrum in **Figure 38**).

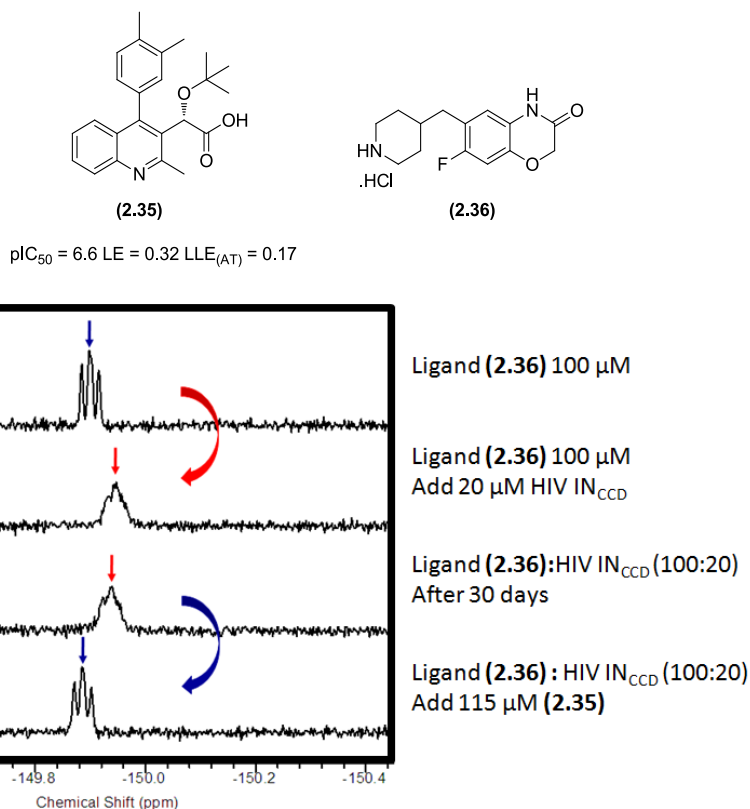


Figure 38 - NMR line broadening studies to identify ^{19}F fragments that bind in the HIV IN_{CCD} IBD pocket.¹⁴⁷ Piperidine (2.36) bound to HIV IN_{CCD} and which was reversed by known HIV IN_{CCD} IBD binder quinoline (2.35).

The ^{19}F library was screened through the high concentration TR-FRET biochemical assay (full curve format, with a top concentration of 500 μM) in parallel with the NMR line broadening experiment. In collaboration with colleagues from our laboratories,¹⁵⁴ compounds were selected for X-ray crystallography using the following criteria:

- 1) demonstration of reversible ^{19}F differential line broadening to HIV IN_{CCD} – 29 compounds; and
- 2) demonstration of irreversible ^{19}F line broadening (after addition of quinoline inhibitor (2.35)), combined with measurable biochemical activity in the high concentration TR-FRET – an additional 22 compounds. This provided an opportunity to identify allosteric binders and potential alternative mechanisms for HIV IN_{CCD}-LEDGF inhibition, e.g. protein movement.¹⁵⁴

In total, these 51 fragments from the ^{19}F library were selected for X-ray crystallography, giving a 3.5% hit rate after the biochemical and biophysical triage (**Figure 39**). To facilitate follow-up of the large hit rate, a high throughput soaking crystallisation system was developed¹⁵⁰ and was used in an attempt to provide crystal structures of HIV IN_{CCD} complexed with each of the 51 fragments. Typically, the liganded crystals diffracted to $\sim 2.2 \text{ \AA}$ using our laboratories X-ray source, or $\sim 1.7 \text{ \AA}$ using an external X-ray source (synchrotron). Protein was soaked in buffer with 2.5 - 50 mM compound and all crystal soaks contained 5% DMSO. Of the 51 fragments sent for X-ray crystallography, only 5 translated to refined structures giving a 10% success rate. The overall hit rate of the complete ^{19}F screen was 0.34% (5 crystal structures from 1465 compounds in ^{19}F library). This is a lower hit rate than observed for more tractable targets such as kinases. For example, a recent kinase FBDD screen provided a hit rate of 1.9% (20 crystal structures / 1065 compounds screened).¹⁵⁵ Having stated this, the method of triage of these kinase hits to crystallography was different to the triage of hits to target the HIV IN-LEDGF interaction. Accordingly, direct comparison of hit rates for different target classes is less readily achieved. Every FBDD project is approached using different screening techniques and triage criteria, as the triage is tailored to the given target. Despite the screening against HIV IN_{CCD} giving a low hit rate, it was encouraging that small fragments can be identified as HIV IN_{CCD} binders.

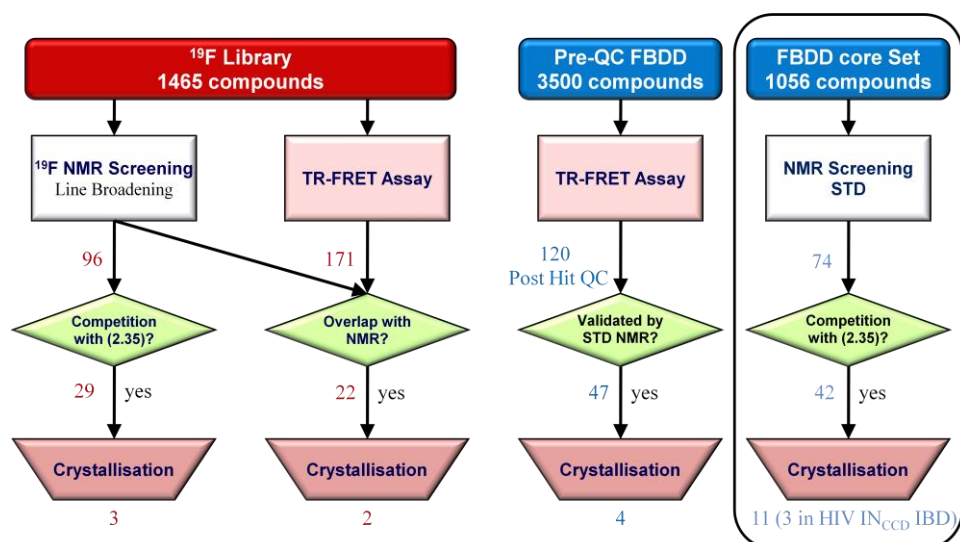


Figure 39 - Screening cascade and the number of hits progressed to each assay for the ^{19}F and fragment sets.

Under ideal circumstances, the entire fragment library (3500 compounds) would have been screened through biophysical and biochemical methods in parallel (as performed for the ^{19}F library). The size of the FBDD library, however, dictated the requirement for triage by biochemical assay to the biophysical technique. Running the NMR biophysical screen by cocktailing the 3500 compounds was a potential solution to this, but due to time constraints this approach was not a viable option. As a compromise, the 3500 compounds were screened through the TR-FRET biochemical assay,¹⁵³ in duplicate, at a single concentration (1 mM). Compounds with percentage inhibition (% I) of > 40 (254 compounds), were re-screened in full dose response curve format to establish the half maximal inhibitory concentration (IC_{50}) which is converted into pIC_{50} ($\text{pIC}_{50} = -\log\text{IC}_{50}$). The concentrations for the dose response were chosen with a top concentration of 1 mM. The %I cut-off of 40 was selected to identify weak affinity binders which could potentially exhibit high ligand efficiency. A disadvantage of using this low %I cut-off is that there is a high chance of progressing false positives as a consequence of the high concentration (1 mM) at which the ligand is screened. Of the 254 compounds which were re-screened in dose response format, all compounds with measurable pIC_{50} values, and without activity in the GST-His₆ counter screen, were progressed to NMR STD studies. To

identify the compounds for progression and avoid losing weak binders through the triage process, all TR-FRET dose response curves of compounds with $\text{pIC}_{50} < 3.0$ were inspected, and analogues with a maximum inhibition of 30 - 50% were also progressed to NMR STD studies along with compounds exhibiting pIC_{50} values of ≥ 3.0 . The triage process resulted in a set of 120 compounds being progressed to NMR, with ligand efficiency ranging between 0.2 - 0.6, with $\text{LLE}_{(\text{AT})}$ between 0.23 – 0.82.

NMR studies were used as an alternative to the fluorine line broadening experiment used for the ^{19}F set. Of the 120 compounds triaged for NMR STD studies, 17 strongly binding fragments (signal to noise (S/N) > 11) and 30 weakly binding (S/N < 11) were identified. NMR displacement studies with quinoline compound (**2.35**) were not performed, due to the interest in finding allosteric binders. As a result, all 47 fragments were progressed to X-ray crystallisation studies. Soaks were carried out at 50 mM over 14 - 28 days to give the best chance of crystallographic success, resulting in 4 validated crystallographic hits (**Figure 39**).

The decision to progress only the biochemical actives to the NMR assay, rather than screen the entire library, was made with the knowledge that the final fragment set would have a smaller core-set selected (1056 compounds). When the new set became available, this was immediately screened in cocktails against HIV IN_{CCD} by NMR STD in competition format with the known inhibitor, quinoline fragment (**2.35**). The aim of this exercise was to identify compounds below the activity threshold of the biochemical assay, which could provide alternative fragment starting points. Screening the core-set provided 42 fragments which were triaged for X-ray crystallography. From these 42 fragments, 3 novel fragment crystal structures in the HIV IN_{CCD} IBD pocket were obtained (**Figure 39**) and 11 structures were observed to occupy a small allosteric site. All 14 compounds were below the limit of the biochemical assay confirming that weak but efficient fragments were likely to have been missed in the biochemical screen of the original FBDD set (3500 compounds).

Further detail around these 14 structures will not be discussed as this extends outside this programme of work.

The next step in the FBDD process was to perform sub-structure searching (SSS) and similarity searching to provide SAR for each of the 98 compounds (query molecules) sent for X-ray crystallography.¹⁵⁶ The SSS process involves taking portions of the query molecules and searching for compounds which contain this motif with additional functionality appended. In comparison, similarity searching involves identification of compounds with similar shape but with variation in the positioning of atoms. Combination of both these processes aids in identification of the features important for binding, and growth vectors from which additional binding affinity can be achieved. Both the GSK compound collection and commercial sources were considered for this SAR study. The available compounds considered were filtered on properties using the ‘rule of 3’ (cLog P < 3, HBA ≤ 3, HBD ≤ 3, MW < 300, rotatable bonds ≤ 3) (**Section 1.4**).³² The array of compounds that resulted from this analysis were visually inspected to select a set of 10-30 compounds for each query molecule, exploring potential growth vectors and small single point changes to the core. The resulting SSS compound set was run in the full-curve dose response TR-FRET assay (top concentration 300 µM) with the aim of providing:

- 1) a small SAR package for each fragment template progressed to X-ray crystallography (subsequent extensive SSS was carried out around compounds yielding crystal structures); and
- 2) compounds with improved potency that are structurally related to fragment hits that yielded no, or partial, electron density in X-ray crystallographic studies.

Collecting this biochemical data in parallel to X-ray crystallography attempts, results in a body of data being available providing information about key binding interactions and potential growth vectors to supplement structural data for each query molecule if it is achieved. This data package gives sufficient data from which to design the first iteration of fragments for synthesis.

2.9.4 Prioritisation of the Screening Output for Synthetic Preparation and Optimisation

The initial fragment and ^{19}F library screening cascades yielded a total of nine crystal structures worthy of further consideration. The nine crystal structures were clustered by the possible growth vectors (**Figure 40**), rather than the more traditional structural chemotype clustering used during HTS follow-up. When considering fragments it is possible that two compounds of similar structure could be binding in distinctly different binding modes in the protein pocket, resulting in unique SAR. Clustering of compounds by overlay of X-ray crystal structures, and considering growth vectors, and ability to access various areas of the protein pocket, can result in compounds which are structurally distinct being grouped together. This is advantageous for fragment chemistry plans, allowing access to different parts of the pocket from multiple templates. This is of further assistance, as synthesis of analogues grown in a particular vector is often more tractable from one template class than another. In addition to this, clustering fragments by growth vector also ensures that fragment development is approached in two distinct ways:

- 1) fragment growth explores the protein pocket from *different vectors* into the same region, as different entry vectors may give varying selectivity and / or activity profiles; and
- 2) fragments can be grown into a number of *distinct areas* of the pocket, thus preventing studies from concentrating on the same hypothesis, which may ultimately fail to produce active analogues.

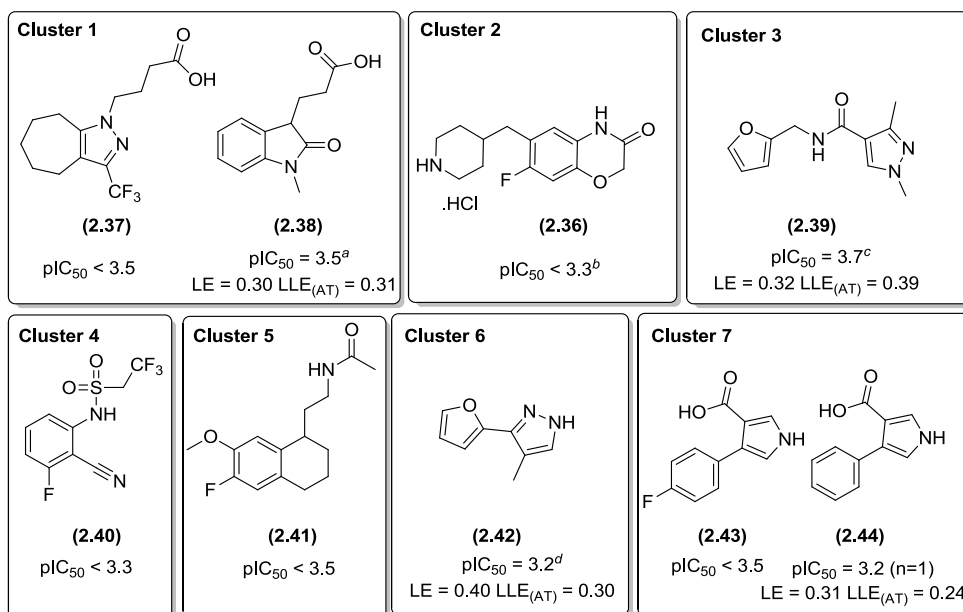


Figure 40 - Nine fragments clustered into seven fragment clusters.

^aThis compound has only been reported active in 3 of the 6 screening occasions. ^bThis compound measured < 3.52 on 3 screening occasions and < 3.3 on 1 screening occasion. ^cOnly Batch 1 is represented in the mean pIC_{50} —Batch 1 was active on 3 out of 5 screening occasions and variability was observed ($pIC_{50} = 3.35, 4.29, < 4.3, 3.33, < 3.52$). ^dThis compound has only been reported active in 2 of the 3 screening occasions ($pIC_{50} = < 4.3, 3.39, 3.06$).

Seven clusters were formed, with clusters 1 - 5 containing fragments which bind in the HIV IN_{CCD} LEDGF IBD pocket. Cluster 6 contains a fragment which binds at a new site along the dimer interface of the HIV IN_{CCD} (2nd site) (**Figure 41**). This pocket is very small and contains a mutation (F185K) which is required for solubilisation of the protein for X-ray crystallography and biochemical screening purposes.¹⁵³ In the native form of the protein, the size of the pocket would be smaller, leaving insufficient room for the compound to bind. In conjunction with this, there is no precedence for compounds binding at the dimeric interface pocket being able to inhibit the HIV IN-LEDGF interaction. Therefore, it was deemed appropriate to terminate studies on cluster 6.

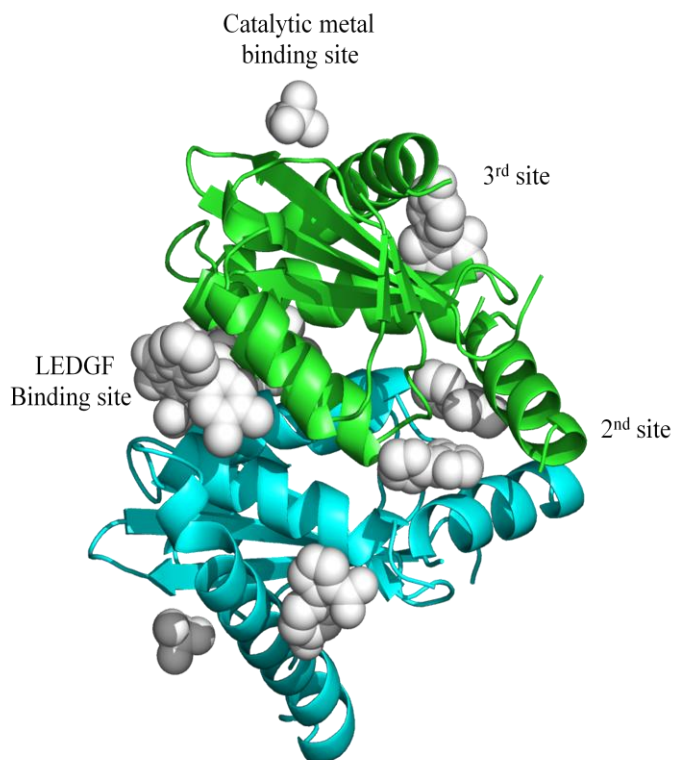


Figure 41 – Cartoon of the HIV IN_{CCD} dimer showing the distinct pockets occupied by fragment hits.¹⁰⁰

Cluster 7 contains two compounds which bind at a crystal contact site (3rd site) (**Figure 41**). Without the crystal contact, which is an artifact of the X-ray crystallography system, no pocket exists and the molecule sits on the surface of the protein. It is unlikely that this interaction could be used to inhibit the HIV IN-LEDGF interaction; therefore, studies around this cluster were also terminated.

The remaining clusters (clusters 1-5), were prioritised for optimisation by consideration of: structural confidence; ligand interactions; physicochemical properties; data from related analogues; potential synthetic direction; and synthetic tractability. A summary of the properties of the compounds in clusters 1 – 5, compared with quinoline (**2.35**), is shown in **Table 2**; with green denoting a good property, orange borderline, and red a potential issue.

Cmpd no.	Cluster	Structural confidence	pIC ₅₀	LE	LLE (AT)	cLog P	cPFI
(2.35)	Quinoline	high	6.6	0.32	0.17	5.3	6.4
(2.37)	1	high	< 3.5 weak ^d			3.0	3.4
(2.38) ^a	1	high	3.5	0.30	0.31	1.1	1.1
(2.36)	2	high	< 3.3 inactive ^e			1.9	1.9
(2.39)	3	high	3.7	0.32	0.39	0.5	3.5
(2.40)	4	high	< 3.3 weak ^d			1.8	3.4
(2.41)	5	low – multiple conformations	< 3.5 inactive ^e			2.5	4.9

Table 2 - Summary of information used to prioritise clusters for wet chemistry compared to literature analogue quinoline (2.35). ^aThis compound has only been reported active in 3 of the 6 screening occasions. ^bThis compound measured < 3.52 on 3 screening occasions and < 3.3 on 1 screening occasion. ^cOnly Batch 1 is represented in the mean pIC₅₀–Batch 1 was active on 3 out of 5 screening occasions and variability was observed (pIC₅₀ = 3.35, 4.29, < 4.3, 3.33, < 3.52). ^dCompounds are classed as ‘weak’ binders when 20-50% response is observed at the maximum screening concentration. ^eCompounds are classed as ‘inactive’ when the response at the maximum screening concentration is <20%.

Pleasingly, all fragments identified had improved cPFI when compared to quinoline (2.35). The lowest cPFI values were for cluster 1 indolinone analogue (2.38) and cluster 2 piperidine (2.36). It was anticipated that the cPFI will increase during the fragment optimisation process, and therefore, starting with a compound with lower cPFI would give the best chance of keeping to the desired range (PFI < 6).

Tetrahydronaphthalene (2.41) (cluster 5) had the highest cPFI value and demonstrated multiple binding modes in X-ray crystallography. From the electron density map there are potentially up to four binding modes, with two appearing dominant. Due to the inconclusive binding mode of the tetrahydronaphthalene (2.41), and no measurable activity in the biochemical assay, no chemistry was carried

out on this cluster. Instead, a further SSS of our laboratories compound collection around this fragment was performed. The aim of the SSS was to find structurally related fragments with improved binding affinity, and subsequently identify a preferred binding orientation. From the biochemical data on the SSS set and follow-up crystallography attempts, determination of a preferred binding mode for tetrahydronaphthalene (**2.41**) was unsuccessful, therefore, this cluster was of no further interest for this study.

Despite variation in the cPFI values, the remaining clusters (1-4) were all considered as good fragment starting points, as they had robust structural data combined with good to moderate lipophilicity and cPFI values. All 4 clusters were explored by synthesis within our laboratories, with fluoro analogue (**2.36**) from cluster 2 being the focus of this research project.

2.10 Fragment Optimisation of Cluster 2 Analogue (2.36)

For structurally enabled targets, utilisation of the X-ray crystallographic data is considered to improve the chance of success for an FBDD approach. Therefore, before optimising cluster 2 analogue (**2.36**), consideration was turned to understanding the binding mode and key interactions for the known binder quinoline (**2.35**).

2.10.1 Binding Mode of Quinoline (2.35)

Quinoline (**2.35**) binds to HIV IN_{CCD} in the LEDGF binding domain pocket (**Figure 42**). Hydrogen bond interactions are made from the carboxylic acid to the backbone NH of both Glu170 (Gln170 NH-O, 2.66 Å) and His171 (Gln170 NH-O, 2.83 Å). The remainder of the molecule has a series of lipophilic interactions in the HIV IN_{CCD} pocket, with the *t*-butyl and dimethyl phenyl moieties being deeply buried in a pocket. The quinoline reaches out to solvent whilst resting on a lipophilic shelf, and the location of this group is considered important due to the clash with the LEDGF protein preventing formation of the HIV IN-LEDGF complex. The structural knowledge from quinoline (**2.35**) was compared with that of piperidine (**2.36**) and was used to help guide the optimisation process.

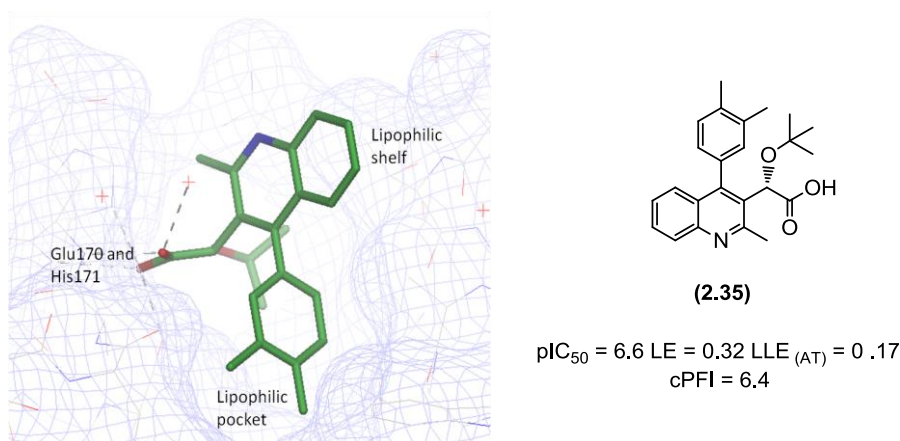


Figure 42 – X-ray crystal structure of quinoline (**2.35**) bound in the HIV IN-LEDGF binding site (PDB code: 1ODVR).

2.10.2 Binding Mode and Areas of Focus for Optimisation of Cluster 2

Analogue (**2.36**)

The location and binding orientation of piperidine (**2.36**) was confirmed by X-ray crystallography. The X-ray data was complemented by the compound exhibiting ^{19}F line broadening in NMR, which is reversed on addition of quinoline (**2.35**). Piperidine (**2.36**) resides deep in the HIV IN_{CCD} pocket in comparison to quinoline (**2.35**) and forms two H-bond interactions (**Figure 43**). Firstly, the piperidine nitrogen H-bonds to the backbone carbonyl of Gln168 (Gln168 O-NH, 2.92 Å, 168 °), and secondly the amide carbonyl from the benzoxazinone core H-bonds to the amide side chain of Gln95 (Gln95 NH-O, 2.71 Å, 142 °) (**Figure 43**).

The H-bond interaction with the backbone carbonyl of Gln168 was novel to the best of our knowledge, which increased interest in this cluster.

Four areas of work were identified for cluster 2:

1. SSS of the GSK compound collection and externally available compounds, to establish SAR around the template;
2. amide substitution to mimic the H-bond interaction with Glu170 and His171, as demonstrated by quinoline analogue (**2.35**) (**Figure 42**);

CONFIDENTIAL – Property of GSK – Do Not Copy

- exploration of piperidine replacements, including improvement of the H-bond angle to the carbonyl of Gln168 (**Figure 43**, top right image);
- growing onto the lipophilic shelf occupied by the quinoline core of compound (**2.35**) (**Figure 42**).

Each of these four areas will be discussed in turn.

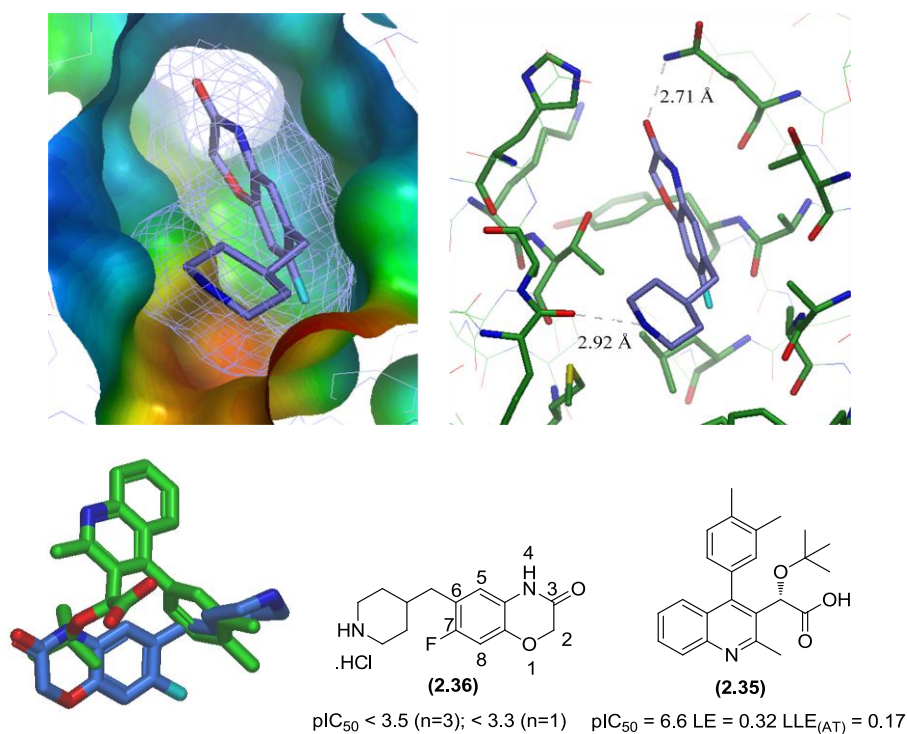


Figure 43 – Top left image - X-ray crystal structure of piperidine (2.36**) bound in HIV_{CCD} (PDB code: 1LIQU) with density map; Top right image - X-ray crystal structure of piperidine (**2.36**) showing hydrogen bond interactions to the protein; Bottom image - overlay of piperidine (**2.36**) with quinoline (**2.35**).**

2.10.3 SSS Around Piperidine Fragment (**2.36**)

The heavy atom count of piperidine (**2.36**) was greater than that of the molecules that populate the fragment library (19 compared to ≤ 17 , respectively). The increased HA count resulted from formation of the fluorine library without rigorous property

filters. Prior to synthetic evaluation, SSS of piperidine (**2.36**) was used to assess the effect of fragmentation of the molecule, determine if the core is optimal, and explore potential growth vectors.

The weak binding affinity of piperidine (**2.36**) was a concern for progression of the cluster, as routine biochemical screening was to be carried out solely through the biochemical assay. The SSS of our laboratories' compound collection aimed to identify related analogues with robust measurable biochemical potency. Identification of biochemically active analogues was critical to allow potency changes from single point modifications of the fragment to be quantified, and enable meaningful SAR to be developed. Two biochemically active analogues, closely related to piperidine (**2.36**) were identified. These 2 fragments, bromide (**2.45**) and benzoxazinone (**2.46**) (**Figure 44**), both came from SSS for fragmented analogues of piperidine (**2.36**).

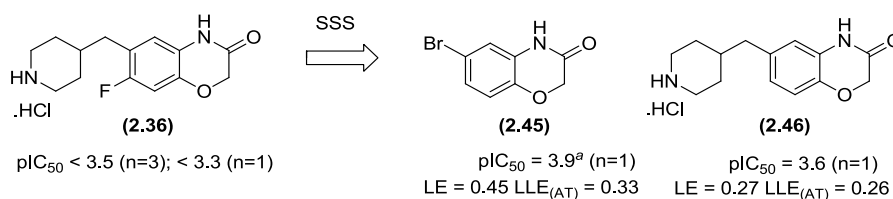
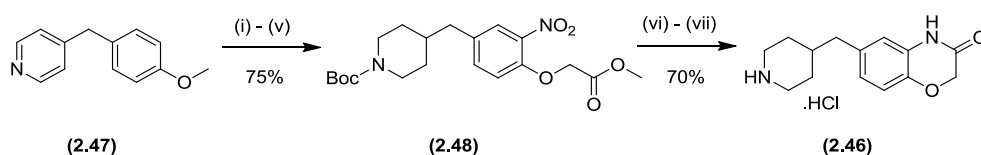


Figure 44 - Compounds identified from SSS and fragmentation of piperidine (**2.36**). Biochemical pIC_{50} are n = 1 for bromide (**2.45**) and piperidine (**2.46**) due to insufficient sample being available. ^aBiochemical pIC_{50} for Batch 1.

Benzoxazinone (**2.46**) ($pIC_{50} = 3.6$, $LE = 0.27$, $LLE_{(AT)} = 0.26$) demonstrated that removal of the fluorine from piperidine (**2.36**) gave weak biochemical activity, with moderate ligand efficiencies. The reason for improvement in activity by removal of the fluorine was unclear. However, removal of the fluorine moiety simplifies the synthetic route to cluster 2 analogues allowing the chemistry related to the published route in **Scheme 1** to be utilised for synthesis.¹⁵⁷ Synthesis of fluorinated analogues has been reported,¹⁵⁸ proceeding *via* intermediates closely related to piperidine

(**2.48**), but due to commercial starting reagent availability the synthesis is two steps longer compared to the corresponding *des*-fluoro analogues.



Scheme 1 - Published synthetic route to benzoxazinone fragment (**2.46**).¹⁵⁷ Reagents and conditions: (i) 48% aq HBr, reflux, 8 h; (ii) H₂, PtO₂, H₂SO₄, MeOH, 20 °C, 18 h; (iii) HNO₃, AcOH, 20 °C, 2 h; (iv) (Boc)₂O, NEt₃, H₂O, THF, 20 °C, 18 h; (v) BrCH₂CO₂Me, K₂CO₃, acetone, reflux, 18 h; (vi) H₂ Pd-C, 3 bar, 50 °C, 18 h; (vii) HCl, Et₂O, 2-propanol, reflux, 18 h.

Bromide (**2.45**) (**Figure 44**) (pIC₅₀ = 3.9, LE = 0.45, LLE_(AT) = 0.33) gave increased biochemical potency and ligand efficiencies compared to piperidine (**2.46**). In our experience, efficiencies of this magnitude are difficult to achieve for protein–protein interactions from fragments, due to the widespread contact area and dispersed interaction points required. NMR displacement studies of bromide (**2.45**) and benzoxazinone (**2.46**) with quinoline (**2.35**), demonstrated that both compounds occupied the binding site of interest. Despite biophysical confirmation that bromide (**2.45**) was binding in the HIV IN_{C_{CD}} LEDGF IBD pocket, X-ray crystallography was unsuccessful.

Although the bromide (**2.45**) and *des*-fluoro compound (**2.46**) were identified from SSS, no biochemically active analogues were found that targeted changes to the core or explored fragment growth. This is largely due to the desired analogues not being available in our compound collection, and therefore, design and synthesis of specific analogues was required.

2.10.4 Amide Substitution to Interact with Glu170 and His171 of HIV IN_{CCD}

Compounds for synthesis were designed from consideration of the X-ray crystallography data for piperidine (**2.36**). Modification of the compound properties in an attempt to gain additional X-ray crystal structures for the template was desirable. Substitution from the amide nitrogen of benzoxazinone (**2.36**) or (**2.46**), or bromide (**2.45**) would prevent the formation of intermolecular H-bond interactions through the amide of the core to another copy of the molecule. From previous unreported experience, carboxamides can be detrimental to solubility which is considered to be due to the energy penalty required to break these intermolecular H-bond interactions. In the case of (**2.36**) or (**2.46**), intermolecular hydrogen bond interactions through the amide functionality of two copies of the molecule, would result in flat planar structures, with aromatic character, which are likely to pack in a lattice resulting in an high melting point and poor solubility. It was hypothesised that improving the solubility of the fragment, by disrupting formation of this lattice, would enhance the possibility of crystallographic success, aiding in construction of fragment optimisation plans. Additionally, amide substitution could target H-bond interactions with the Glu170 and His171 backbone nitrogen as observed for quinoline (**2.35**) and LEDGF when the HIV IN_{CCD}-LEDGF complex is formed (**Figure 45**).

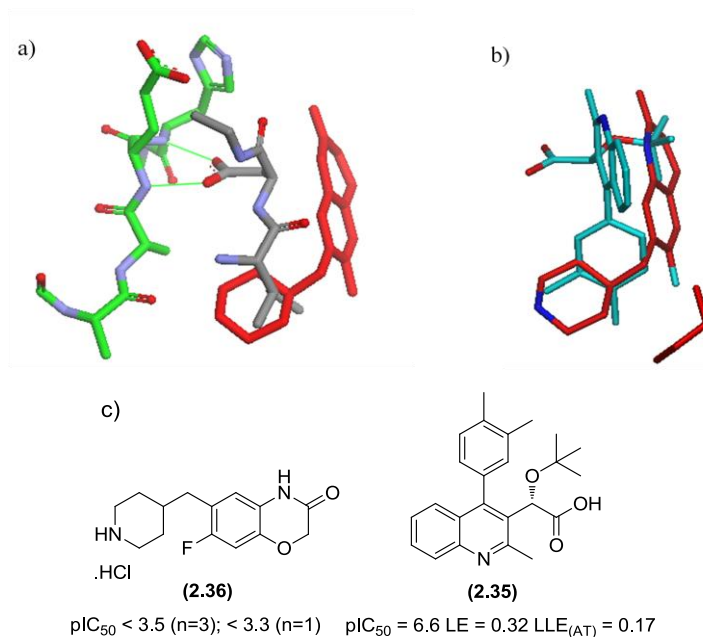


Figure 45 – a) Section of HIV IN_{CCD} interacting with a section of LEDGF (PDB code: 2B4J), overlaid with crystal structure of piperidine (**2.36**) (PDB code:1LIQU); b) piperidine (**2.36**) (PDB code:1LIQU) overlaid with quinoline (**2.35**) (PDB code:1ODVR); and c) Structures of piperidine (**2.36**) and quinoline (**2.35**).

LEDGF forms two H-bond interactions with Glu170 and His171 of HIV IN_{CCD} (**Figure 45**, diagram a). Alanine scanning of linear peptide LEDGF₃₆₁₋₃₇₀ has been reported in the literature, showing that these hydrogen bond interactions only make a small contribution to the binding of LEDGF to HIV IN_{CCD}.^{13,117} It is clear from the alanine scanning that HIV IN_{CCD}-LEDGF interactions are made up of many hydrophobic contacts around the surface of the HIV IN_{CCD} dimer, each providing only a small binding contribution.

Despite this observation, docking of benzoxazinone (**2.36**) with one (**2.49**), two (**2.50**) and three (**2.51**) carbon linked carboxylic acid chains (**Figure 46**), verified the feasibility of forming H-bond interactions with Glu170 and His171. The docking of the one carbon analogue (**2.49**) demonstrated minimal movement of the benzoxazine core when compared to the X-ray crystal structure of piperidine (**2.36**), whilst

making H-bond interactions with Gln168 and Gln95 (**Figure 47**). Greater movement of the core was predicted for fragments **(2.50)** and **(2.51)** compared to the docking of one carbon analogue **(2.49)**.

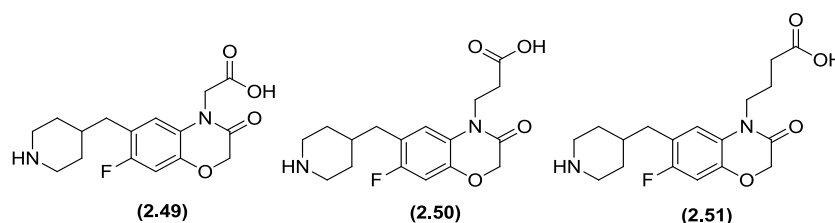


Figure 46 - Molecules docked using LigX minimisation in Molecular Operating Environment (MOE).

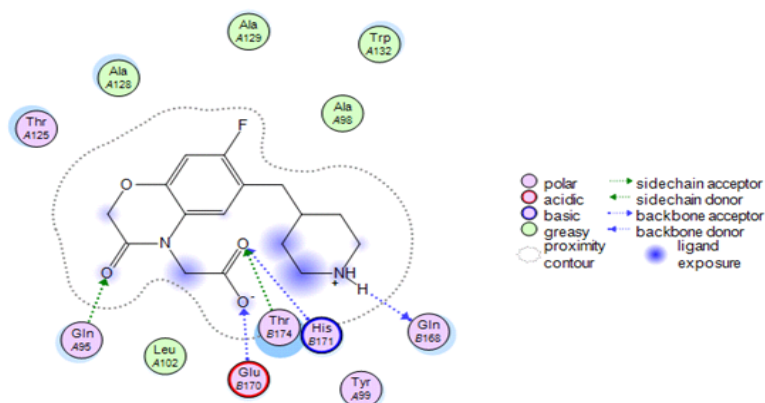


Figure 47 - LigX minimisation in MOE for methylene carboxylic analogue fragment **(2.49)**.

Synthesis of bromides **(2.52)**, **(2.53)** and **(2.54)** (**Figure 48**) was performed to test the hypothesis around optimal linker lengths. The bromide analogues were synthesised in preference to the more complex piperidine analogues (**Figure 46**) as the effect of growth from the more potent bromide fragment starting point would be quantifiable. This decision assumes bromide **(2.45)** (for which X-ray crystallography was unsuccessful) is binding with the benzoxazinone in the same position as the core of piperidine **(2.36)**. It was envisaged that addition of the carboxylic acid group to

bromide (**2.45**), would give enhanced aqueous solubility increasing the chance of crystallographic success.

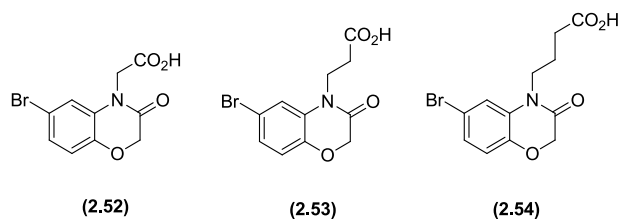
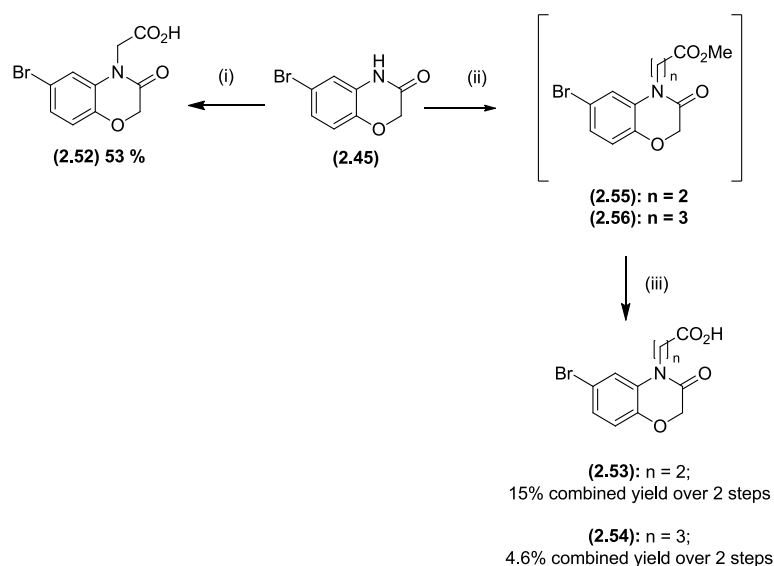


Figure 48 - Bromide analogues synthesised.



Scheme 2 - Synthesis of carboxylic acid analogues (**2.52**), (**2.53**) and (**2.54**) with different linker lengths. Reagents and conditions: (i) methyl chloroacetate, K_2CO_3 , DMF, 20 °C, 23 h; then 2.0 N NaOH, 20 °C, 1 h, 53%; (ii) alkylating agent (for (**2.55**) = ethyl 3-bromopropanoate; for (**2.56**) = ethyl 4-bromobutanoate), K_2CO_3 , DMF, 20 °C, 21.5 h; (iii) 2.0 N NaOH, MeOH, THF, 20 °C, 1 h, (**2.53**) 15%, (**2.54**) 4.5% (over steps (ii) and (iii)).

The carboxylic acid analogues (**2.52**), (**2.53**) and (**2.54**) were synthesised *via* alkylation of bromide (**2.45**) with the relevant alkylating agent to give the corresponding ester intermediates (**Scheme 2**). The alkylation reactions were performed at room temperature giving incomplete consumption of starting material

CONFIDENTIAL – Property of GSK – Do Not Copy

and low yields. This could have been circumvented by heating the reaction in an attempt to drive it to completion. The ester hydrolysis was performed by addition of 2N NaOH, either to the alkylation reaction directly (as for synthesis of (2.52) from (2.45)), or to crude ester intermediate in MeOH and THF (as for synthesis of (2.53) and (2.54)).

All three carboxylic acid analogues (2.52), (2.53) and (2.54) were inactive in the biochemical assay ($pIC_{50} < 3.3$ ($n=1$)). Also, a newly synthesised batch of bromide (2.45) was inactive. To understand if, and how the compounds were binding, the four compounds were submitted for X-ray crystallography. Refined crystal structures were obtained for the one (2.52) and two (2.53) carbon linked carboxylic acid analogues. An overlay of the crystal structures of (2.52) and (2.53) with (2.36) is shown in **Figure 49**, diagram a.

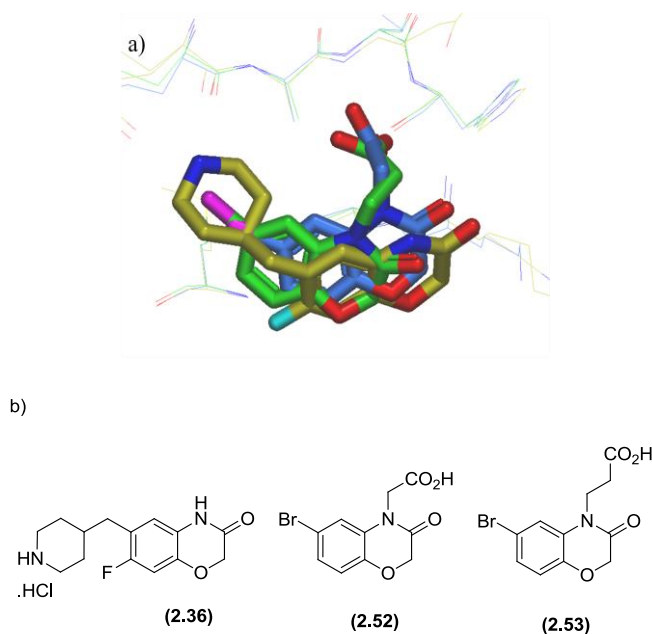


Figure 49 – (a) Overlay of crystal structures of piperidine (2.36) (PDB:1LIQU), one carbon linked (2.52) (PDB:1VLLP), and two carbon linked (2.53) (PDB:1VORG) analogues to compare shifts of the core; (b) structures of piperidine (2.36), one carbon linked (2.52), and two carbon linked (2.53) analogues.

Gratifyingly, X-ray crystallography showed both carboxylic acid analogues, (**2.52**) and (**2.53**), formed H-bonds with Glu170 and His171 backbone nitrogens through the carboxylic acid as designed. Minimal shift of the core was seen for methylene linked acid fragment (**2.52**), with respect to (**2.36**), as predicted by modelling. A more pronounced shift was observed for the two carbon linked analogue (**2.53**), which causes a disruption of the hydrogen bond interaction between the carbonyl oxygen and the side chain of Gln95.

On inspection of the biochemical assay dose response curves, the one carbon linked analogue (**2.52**) shows the beginning of a curve (13% maximum response at the top concentration in the assay) (for example see **Figure 50 - graph b**). The start of a curve suggests the affinity of the one carbon linked analogue (**2.52**) is just below the limit of the assay. In contrast, the dose response of the two carbon linked analogue (**2.53**) is completely flat (for example see **Figure 50 – graph c**). The benzoxazinone core of the two carbon linked analogue (**2.53**) contains no H-bond substituents to anchor the molecule in position as demonstrated by piperidine (**2.36**) which has a flat dose response. Therefore, it was hypothesised that the piperidine group of fragment (**2.36**), may be orientating the molecule across the pocket towards Gln95. Removal of the methylene-linked piperidine from fragment (**2.36**) to give bromide (**2.45**), could have, in turn, allowed the core to take up a more preferred binding position in the HIV IN_{CCD} IBD pocket. This hypothesis has not been proved conclusively, since X-ray crystal structures of small fragments such as bromide (**2.45**), without the carboxylic acid chain on the amide, have failed to provide liganded structures.

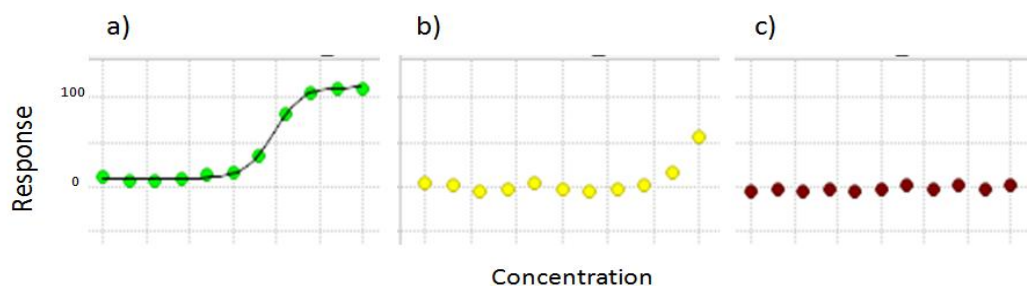


Figure 50 – Biochemical dose response curves. Full-curve showing maximum response of 100% shown in green. The beginning of a curve (which is typical for weak fragments) is shown in yellow. An inactive compound showing a flat line is shown in brown. Curve a) is representative of compounds with a measurable pIC_{50} ; curve b) is representative of the dose response curve for weak compounds i.e. carbon linked analogue **(2.52)** – showed a maximum response of 13%; curve c) is representative of an inactive compound with a maximum response of $\sim 0\%$ i.e. two carbon linked analogue **(2.53)**.

It should be noted that addition of the carboxylic acid group did improve the success rate in X-ray crystallography as hypothesised. The success rate in X-ray crystallography had previously been poor ($\sim 10\%$), despite soaking at high concentration (50 mM) for long soak times (up to 28 days). For compounds with the carboxylic acid, soak times were reduced to, in some cases, 1 day. This improvement in crystallographic success was attributed to the marginally improved solubility of the fragments (**Figure 51**), combined with the enhanced binding affinity. The improvement in crystallographic cycle time helped advance the cluster more quickly in comparison to the other fragment templates.

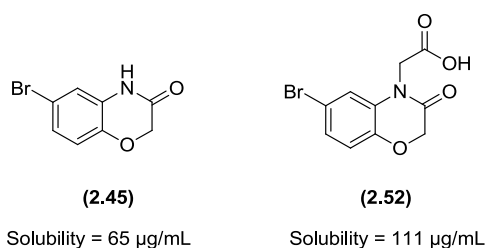
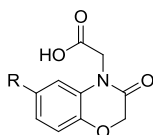


Figure 51 – Solubility comparison for bromides **(2.45)** and **(2.52)**.

2.10.5 Identification and Synthesis of Piperidine Replacements for Fragment (2.36)

Due to the increased crystallographic success with the carboxylic acid present, an additional SSS of our laboratories' compound collection was carried out. The methylene linked carboxylic acid functionality was maintained, and replacements for the bromide in (2.52) were sought (Table 3). Compounds mimicking the size and shape of the carbon linked piperidine in fluoro-analogue (2.36) were screened including sulfone (2.58) and amide (2.59). To ensure that steric bulk did not hinder binding, compounds with smaller substituents including dimethylamide (2.60), methyl (2.61) and the unsubstituted compound (2.62) were also screened. All the compounds were inactive apart from methyl (2.61) which did not give a measurable pIC₅₀, but gave weak binding (between 20 – 50% response at the maximum screening concentration), and *t*-butyl (2.57) with a measurable pIC₅₀ of 4.4.



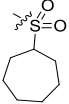
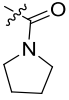

Compound	R	pIC ₅₀	LE	LLE _(AT)	cPFI
(2.57)	<i>t</i> -Bu	4.4	0.32	0.23	2.8
(2.52)	Br	< 3.3 inactive ^{a,b}	< 0.28	< 0.23	1.9
(2.58)		< 3.8 inactive ^a	0.21	0.20	2.0
(2.59)		< 3.3 (n=1) inactive ^a	< 0.21	< 0.3	0.7
(2.60)		< 3.3 (n=1) inactive ^a	< 0.23	< 0.36	0.1
(2.61)	Me	< 3.3 weak ^c	< 0.28	< 0.27	1.4
(2.62)	H	< 3.3 (n=1) inactive ^a	< 0.30	< 0.33	1.4

Table 3 – Selection of results from the SSS to explore the replacement of bromide in (2.52). ^aCompounds are classed as ‘inactive’ when the response at the maximum screening concentration is <20%. ^b13% maximum response observed at the top concentration - the start of a dose response curve. ^cCompounds are classed as ‘weak’ binders when 20-50% response is observed at the maximum screening concentration.

t-Butyl containing (2.57) (pIC₅₀ = 4.4, LE = 0.32, LLE_(AT) = 0.23) demonstrated biochemical potency with good LE despite having an increase in cPFI compared to bromide (2.52). Pleasingly, the crystal structure of *t*-butyl analogue (2.57) was obtained, and overlaid exactly with the crystal structure of bromide (2.52) (Figure 52, diagram a). The SAR around the 6-position of the benzoxazinone core, suggested that large lipophilic groups are required to demonstrate biochemical potency (Table 3). When the crystal structures of bromide (2.52), *t*-butyl (2.57) and

LEDGF (in grey in **Figure 52**, diagram (b)) are overlaid, a larger overlap of *t*-butyl (**2.57**) with LEDGF is observed in comparison to bromide (**2.52**). It is apparent that the *t*-butyl moiety can mimic the hydrophobic interactions made by the LEDGF Ile residue.

Identification of *t*-butyl (**2.57**) was important as it allowed the effect on potency of small changes to the fragment to be quantified. In addition, *t*-butyl (**2.57**) simplified the synthetic route to cluster 2 analogues in comparison to piperidine (**2.36**). Simplification of the synthetic route would facilitate the more rapid exploration of the cluster's growth vectors. Despite the increased potency and tractability, the *t*-butyl moiety in fragment (**2.57**) ($pIC_{50} = 4.4$, $LE = 0.32$, $LLE_{(AT)} = 0.23$) was not optimal as the $LLE_{(AT)}$ was less than 0.3. It was envisaged that physical properties of the compound could be modified in this region once other vectors had been explored. To confirm the feasibility of this property modification, the focus turned to identification of suitable 6-position replacements of the *t*-butyl group, aiming for a combination of measurable potency and improved physicochemical properties.

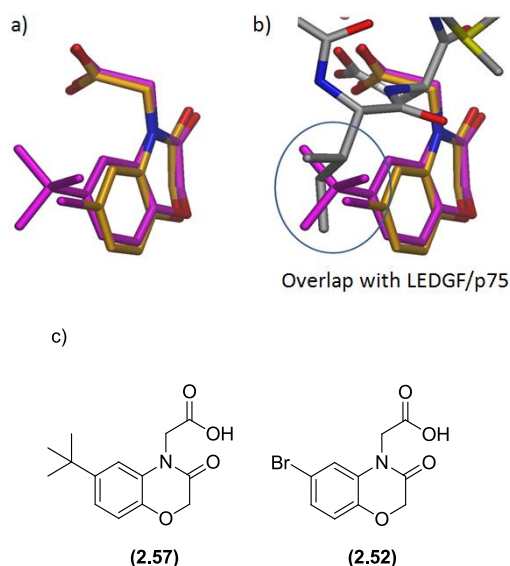


Figure 52 – (a) Overlay of crystal structures of *t*-butyl (**2.57**) (PDB code:1WQHJ) and bromide (**2.52**) (PDB code:1VLLP); (b) overlay of crystal structures of *t*-butyl (**2.57**) and bromide (**2.52**) with a portion of LEDGF (PDB code:2B4J); and (c) structures of *t*-butyl (**2.57**) and bromide (**2.52**).

2.10.6 Synthesis of 4-Piperidine and *t*-Butyl Replacements

The *t*-butyl analogue (**2.57**) ($pIC_{50} = 4.4$, $LE = 0.32$, $LLE_{(AT)} = 0.23$) was the most potent cluster 2 analogue identified. Pleasingly *t*-butyl analogue (**2.57**) had clear electron density for the entire compound in the X-ray crystal structure, which is not always achievable for weakly binding fragments. When obtaining X-ray crystal structures of fragments, often partial density is observed for parts of the fragments which are most ordered in the protein – i.e. where hydrogen bond interactions are made, or excellent shape complementarity and electrostatic interactions are achieved. For flexible chains, it is usual for poor electron density to be observed. Despite the good electron density for *t*-butyl analogue (**2.57**), the sub-optimal $LLE_{(AT)}$ was a concern and we chose to re-evaluate piperidine analogue (**2.46**), which had reduced lipophilicity, before focussing solely on the *t*-butyl moiety. It was assumed that piperidine (**2.46**) bound in the same fashion as piperidine (**2.36**) and, therefore, has the advantage of being anchored by two H-bonds, at either end of the HIV IN_{CCD} IBD pocket, with Gln95 and Gln168. Further analysis of the X-ray crystal structure showed that the H-bond angle from the piperidine of fragment (**2.36**) to Gln168 is poor (**Figure 54**) with the C=O–HN angle deviating from the optimal 120 °. Consideration turned to approaches to improve this hydrogen bond interaction between HIV IN_{CCD} and piperidine (**2.46**).

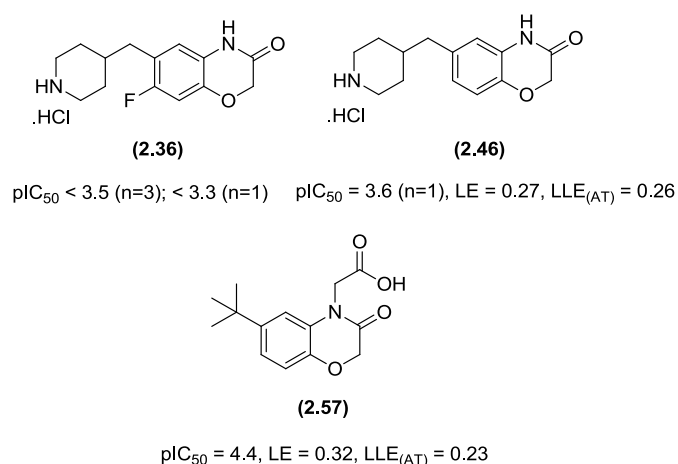


Figure 53 – Structures of piperidine (**2.36**), (**2.46**) and bromide (**2.57**).

The piperidine ring nitrogen has an estimated pK_{aH} of ~ 11 in **(2.46)** and, therefore, would be expected to be protonated at physiological pH.¹⁵⁹ The equatorial NH is predicted to make an H-bond interaction with the Gln168 backbone amide. Although the piperidine NH donor is in the plane of the Gln168 amide carbonyl as desired, the hydrogen is not appropriately juxtapositioned as depicted by arrow a (**Figure 54**, diagram a). Moving the piperidine nitrogen around the ring one position, would provide a NH pointing in the direction of arrow c, compound **(2.63)** (**Figure 54**, diagram a). It was hypothesised that movement of the rest of the ligand could occur to improve the geometry of the H-bond. Alternatively, a pyrrolidine analogue **(2.64)** could provide a preferred H-bond interaction, pointing at the Gln168 backbone carbonyl oxygen lone pair, as depicted by arrow b, if the core of the ligand remains fixed in the protein pocket (**Figure 54**, diagram b). Without movement of the core, the pyrrolidine analogue **(2.64)** was predicted to give an improved angle for donation of the hydrogen towards the lone pair of the carbonyl oxygen compared to either of the piperidine analogues, however, movement of the core could not be ruled out.

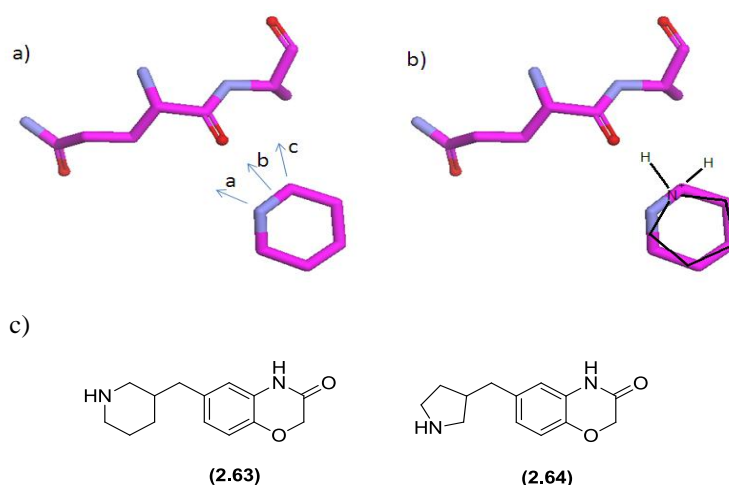


Figure 54 - a) Diagram demonstrating angles of H-bonds with backbone carbonyl; **b)** proposed pyrrolidine analogue overlaid with piperidine; and **c)** proposed piperidine **(2.63)** and pyrrolidine analogues **(2.64)**.

Along with synthesis of the alternative piperidine and pyrrolidine analogues, it was important to examine the binding contribution of the H-bond from 4-piperidine

fragment (**2.46**). As the H-bond angle was poor, the interaction was likely to be inefficient and would have an increased desolvation penalty to overcome in comparison to a direct analogue without the basic centre.

Three compounds were targeted to help determine the importance of the 4-piperidine moiety and identify the optimal position of the nitrogen (**Figure 55**):

1. 3-piperidine analogue (**2.63**),
2. 3-pyrrolidine analogue (**2.64**), and
3. cyclohexyl analogue (**2.65**).

The desired compounds (**2.63**), (**2.64**), and (**2.65**) were prepared in the more potent *des*-fluoro series. The piperidine molecules (**2.36**) and (**2.46**) (**Figure 55**) were closely related in structure, and from docking it was assumed that the two compounds would be binding in an identical fashion. The designed molecules, the 3-piperidine (**2.63**) and the 3-pyrrolidine (**2.64**) had good cPFI values. Cyclohexyl (**2.65**) had an increased cPFI which had moved into a less desirable region (> 6), but this constituted an important tool compound for synthesis to determine the importance of the hydrogen bond to Gln168.

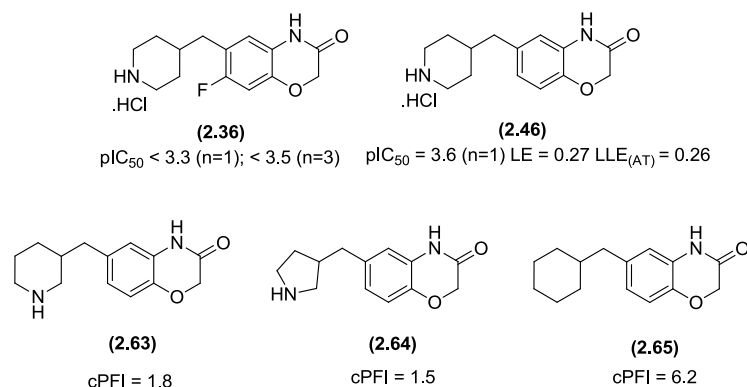


Figure 55 - Structures and data for fragments (2.36) and (2.46). Piperidine (2.63), pyrrolidine (2.64) and cyclohexyl (2.65) molecules designed for synthesis.

2.10.7 Attempts to Determine an Alternative Linker Between the Core and Piperidine Moiety

Before committing to separate complex multi-step syntheses of each analogue, replacing the methylene linker with a heteroatom, or removing the linker altogether, was explored. If alternative heteroatom linkers are appropriate, this would allow facile late stage introduction of a range of pendant groups. However, analogues from our compound collection with alternative linkers were not biochemically active. Therefore, it was hypothesised that the carbon linker was required in order to angle the piperidine out of the plane of the benzoxazinone core (**Figure 56**).

Closer analysis of the X-ray crystal structure of cluster 2 hit (2.36) shows a dihedral angle of -74.0° between the core and the piperidine branch point. This is comparable to the dihedral angle obtained when the piperidine (2.36) occupies the lowest energy conformation (dihedral angle = -76.7°) (**Figure 56**).

The keto analogue (2.66) (dihedral angle = -21.9°) and ether analogue (2.67) (dihedral angle = 170.2°) displayed dihedral angles placing the linker close to the plane of the benzoxazinone core. Positioning the piperidine out of the plane as required, as in compound (2.36) would give a high energy conformation for the carbonyl (2.66) and ether (2.67) analogues. From this analysis, it was decided that

chemistry to the carbon linked analogues needed to be pursued to obtain the correct geometry to interact with Gln168.

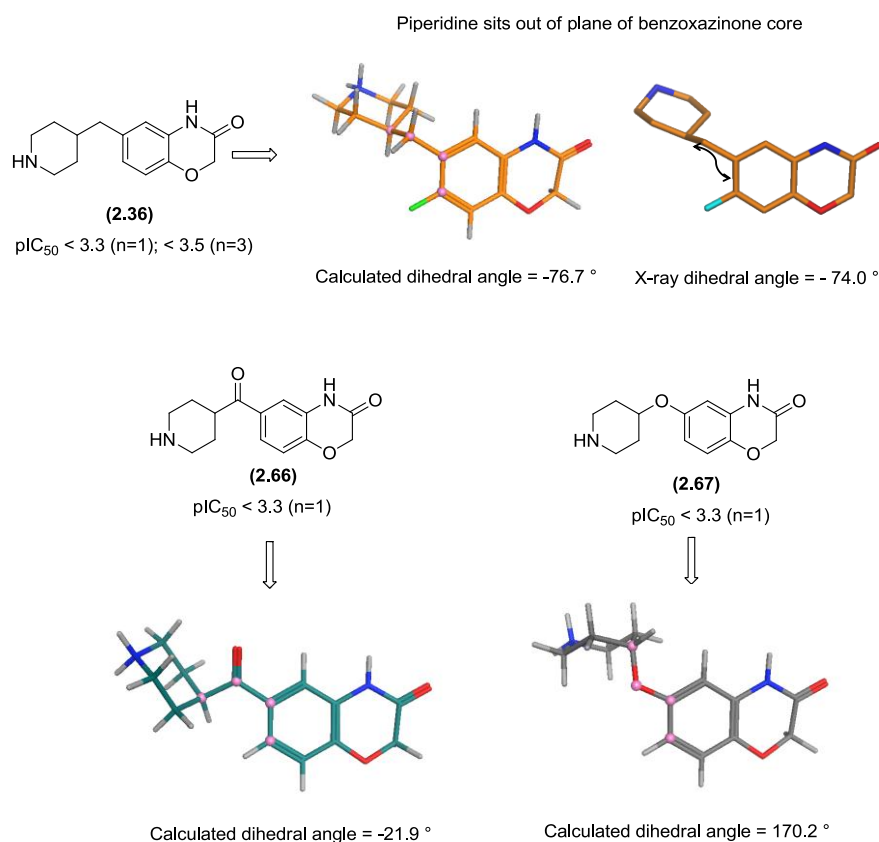
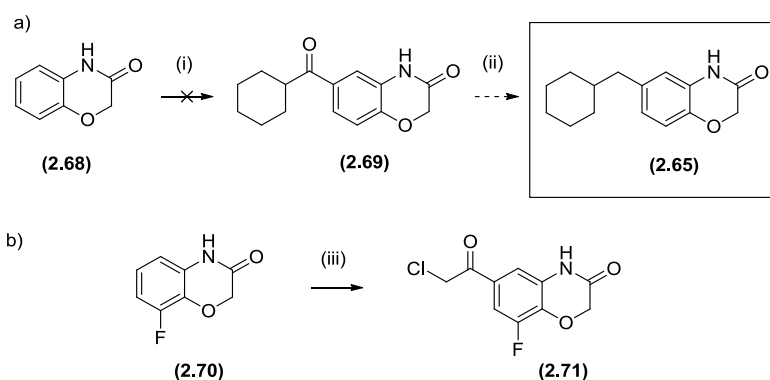


Figure 56 - Representative compounds from SSS of alternative linkers. Dihedral angles of the calculated low energy conformations (from a stochastic search in MOE) for the piperidine linker, with respect to the benzoxazinone core, compared to the dihedral angle for X-ray crystallographic conformation of piperidine (**2.36**) (PDB code: 1LIQU). The atoms the dihedral angles are measured between are shown in pink.

2.10.8 Development of a Synthetic Route to Cyclohexyl Analogue (**2.65**)

Syntheses of carbon linked analogues (**2.63**), (**2.64**), and (**2.65**), proved to be challenging. Use of Friedel-Crafts acylation¹⁶⁰ as a key synthetic step was attempted, and a representative example is shown in **Scheme 3**.

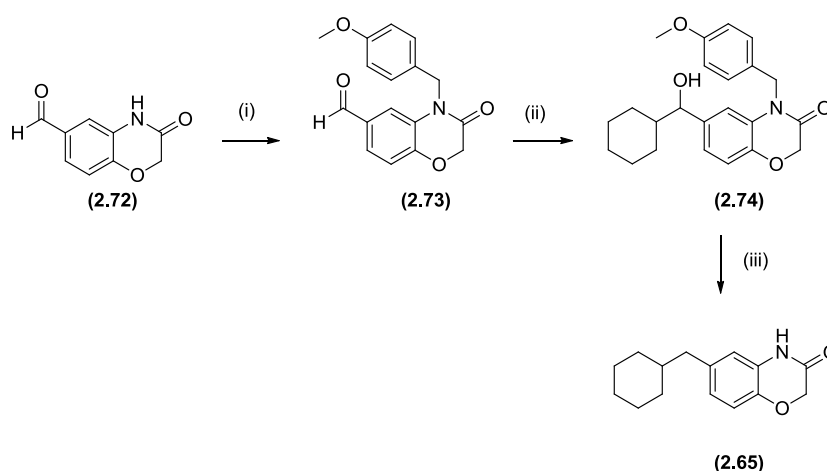
Preparation of the keto analogue (**2.69**) by portionwise addition of AlCl_3 (5 eq) and cyclohexyl acid chloride to a solution of benzoxazinone (**2.68**) in anhydrous DMF, followed by heating to $80\text{ }^\circ\text{C}$ was unsuccessful. No conversion of starting material was observed. This is unsurprising since Friedel-Crafts acylation of similarly substituted aryl compounds is not well preceded. This suggests that this synthetic sequence may not be feasible due to steric hindrance or reduction in electrophilicity of secondary acyl chlorides over primary analogues.



Scheme 3 - a) Attempted synthesis of methylene linked cyclohexyl analogue at the 6-position on benzoxazine core; and **b)** Friedel Crafts literature example.¹⁶⁰

Reagents and conditions: **a)** (i) CyCOCl , AlCl_3 , DMF, RT to $80\text{ }^\circ\text{C}$; (ii) TFA, Et_3SiH ; **b)** Literature example - (iii) ClCH_2COCl , AlCl_3 , $80\text{ }^\circ\text{C}$, 5 h, (50-65%).¹⁶⁰

As an alternative strategy to access the carbon linked analogues, investigation of adding Grignard reagents to the commercially available aldehyde (**2.72**) was initiated (**Scheme 4**). Protection of the amide functionality in aldehyde (**2.72**) was chosen, to prevent the proton on the amide nitrogen ($\text{pK}_a \sim 20$) from being deprotonated by the Grignard reagent. The *p*-methoxybenzyl (PMB) protecting group was used as this proved to be useful in previous related chemistry undertaken. Facile deprotection of the PMB group occurs by heating in toluene with tosic acid in the microwave,¹⁶¹ alternatively the PMB group could be removed at the same stage as the planned deoxygenation of the benzyl alcohol moiety with TFA and Et_3SiH (*vide infra*). Amide protection with PMBBr and potassium carbonate in DMF was successful to give compound (**2.73**) in 81% yield.



Scheme 4 – Synthetic route to access methylene-linked cyclohexyl (**2.65**). Reagents and conditions: (i) PMBBr, K_2CO_3 , DMF, 20 °C, 81%; (ii) CyMgCl (2.0 M in diethyl ether), THF, -78 °C to RT, 38%; (iii) TFA, Et_3SiH , 50-60 °C, 58%.

The key step for the synthetic route was the nucleophilic addition of cyclohexyl Grignard into the aldehyde of (**2.73**) (**Scheme 4**, step ii). The Grignard reagent (2.0 M in diethyl ether) was added dropwise to a solution of the aldehyde core (**2.73**) in THF at -78 °C. Reaction quenching, work-up and reverse phase purification afforded the cyclohexyl substituted alcohol (**2.74**) in 38% yield. Low consumption of aldehyde (**2.73**) was observed and reaction optimisation was required (discussed later), however, isolated material was taken on in an attempt to reach the final methylene linked cyclohexyl product (**2.65**). Dehydroxylation and deprotection of the amide were performed simultaneously by heating in TFA at 50 - 60 °C with triethylsilane. The crude was purified to afford cyclohexyl analogue (**2.65**) in 58% yield.

The cyclohexyl analogue (**2.65**) ($\text{pIC}_{50} = 3.6$, $\text{LE} = 0.27$, $\text{LLE}_{(\text{AT})} = 0.08$) demonstrated weak biochemical activity and moderate LE, and indeed comparable to that of piperidine (**2.46**) ($\text{pIC}_{50} = 3.6$, $\text{LE} = 0.27$, $\text{LLE}_{(\text{AT})} = 0.26$). This confirmed the hypothesis that the piperidine nitrogen H-bond to Gln168 in piperidine (**2.46**) was

not contributing significantly to its potency. The $LLE_{(AT)}$ for cyclohexyl (**2.65**) is dramatically reduced compared to piperidine (**2.46**), demonstrating that the lipophilicity is not playing a significant role in binding, and that the cyclohexyl group in analogue (**2.65**) is less optimal than the 4-piperidine in fragment (**2.46**) (**Figure 55**). The piperidine basic centre in (**2.46**) has a greater desolvation penalty to overcome in comparison to cyclohexyl (**2.65**) and this penalty cancels out the increased binding obtained by the inefficient H-bond. These observations give increased confidence that improvement of the H-bond angle for piperidine (**2.46**) to Gln168 could lead to an increase of potency with maintenance of LLE_{AT} . An increase in H-bond efficiency could help overcome the desolvation penalty associated with removal of the piperidine from the bulk solvent.

2.10.9 Development of a Synthetic Route to 3-Piperidine Analogue (2.63)

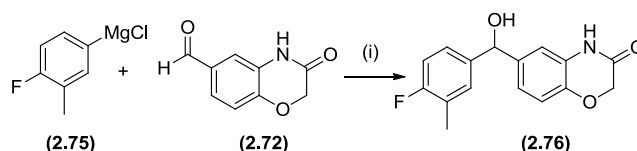
The focus of the synthetic chemistry now turned to the development of a synthetic route to form the 3-piperidine analogue (**2.63**) (**Figure 55**). In the formation of cyclohexyl alcohol (**2.74**), the Grignard reagent addition to aldehyde (**2.73**) (**Scheme 4**) proceeded in low yield. The low yield and poor consumption of starting material (**2.73**) led to the assumption that the Grignard reagent could be acting as a base. If the Grignard reagent had acted as a base prior to nucleophilic attack of the aldehyde (**2.73**), no product would be formed. The Grignard reagent with a $pK_a \sim 45$, could be deprotonating:

1. the methylene protons between the amide carbonyl and the ether oxygen of the benzoxazinone core of aldehyde (**2.73**) ($pK_a = 25 - 30$); or
2. the methylene protons of the *p*-methoxy benzyl group of aldehyde (**2.73**) ($pK_a = 25 - 30$).

To explore optimisation of the Grignard reaction conditions, addition of the organometallic reagent to unprotected amide (**2.72**) was attempted. This removed the potential for interference of the benzylic protons with the chemistry. For this study, aryl Grignard (**2.75**) (**Scheme 5**) was selected as an alternative Grignard

reagent to cyclohexylmagnesium chloride to expand the SAR for the template. Aryl Grignard (**2.75**) was selected for the following reasons:

- the product from reacting aldehyde (**2.72**) and aryl Grignard reagent (**2.75**) was interesting from a modelling perspective, as the fluoro-tolyl group was predicted to occupy a large lipophilic area in the pocket;
- Grignard reagent (**2.75**) does not contain a basic centre or protecting group which may interfere with the chemistry; both a basic centre and protecting group would be present if the desired 3-piperidine (**2.63**) or 3-pyrrolidine (**2.64**) analogues were targeted, **Figure 55**;
- Grignard (**2.75**) had a UV chromophore which will aid in following the reaction profile.



Scheme 5 – Optimisation of the Grignard reagent addition to aldehyde (**2.72**).

Reagents and conditions: (i) THF, 0 – RT, 7.2% (+ impure batch 55%).

Aryl Grignard (**2.75**) (2 eq) was reacted with aldehyde (**2.72**), **Scheme 5**. It was assumed that the first equivalent of the Grignard reagent was sacrificed as a base. The yield of the alcohol (**2.76**) was approximately 62%, consisting of a pure batch (7.2% yield) and an impure batch (55% yield containing aliphatic impurities by NMR considered to be grease). This indicates that this one step protocol is potentially superior to using PMB as an amide protecting group in terms of yield and atom efficiency. The improved Grignard reagent conditions provided a potential pathway to future analogues, but further synthetic route development was still required. Disappointingly, fluoro-tolyl analogue (**2.76**) was inactive in the biochemical assay ($pIC_{50} < 3.3$).

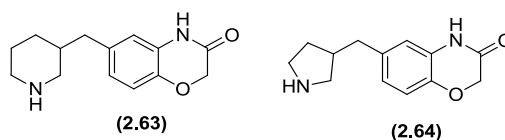
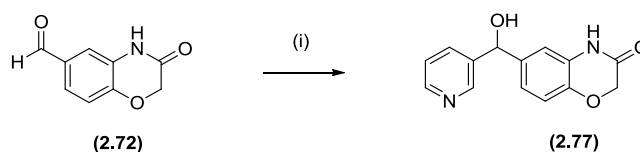


Figure 57 - 3-Piperidine (**2.63**) and 3-pyrrolidine (**2.64**).

An alternative synthetic approach was taken for formation of both the 3-piperidine (**2.63**) and 3-pyrrolidine (**2.64**) analogues. To access 3-piperidine (**2.63**) the approach involved nucleophilic addition of pyridin-3-ylmagnesium bromide reagent into aldehyde (**2.72**). The requisite Grignard reagent was prepared *in situ* by dropwise addition of *i*PrMgCl.LiCl solution in THF to commercially available 3-bromopyridine at room temperature. Literature examples report that *i*PrMgCl.LiCl has enhanced reactivity in Grignard reagent formation compared to *i*PrMgCl. The improved reactivity is due to the breakdown of polymeric aggregates formed by isopropylmagnesium chloride.¹⁶² Due to the first equivalent of Grignard being sacrificed as a base, two equivalents of the 3-pyridine Grignard reagent were prepared with respect to aldehyde (**2.72**) which was added dropwise to the solution of 3-pyridine Grignard (**Scheme 6**). The addition was reversed compared to the Grignard reaction in **Scheme 5** to ensure that the Grignard reagent remained in excess. The conversion to product by LCMS was approximately 20%, however, after work-up and purification the alcohol (**2.77**) was isolated in 2.9%. In an attempt to enhance the yield, the addition was reversed, and the Grignard reagent was added to the aldehyde (**2.72**), but no improvement in conversion was observed.

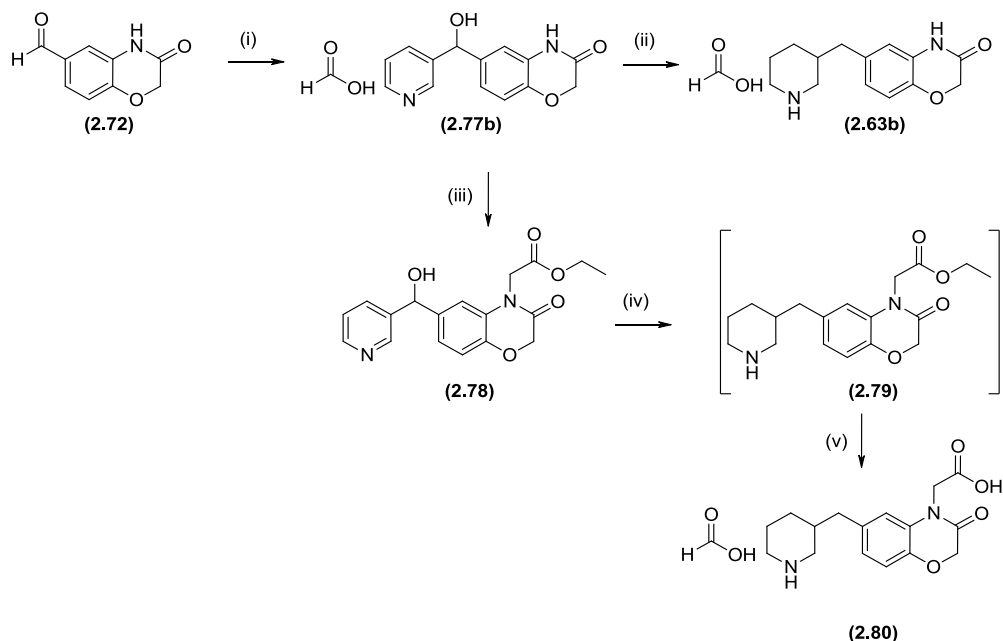


Scheme 6 – Formation of alcohol (**2.77**) from aldehyde (**2.72**). Reagents and conditions: (i) 3-Bromopyridine, *i*PrMgCl.LiCl, THF, 4 h, RT, 2.9%.

Literature examples demonstrated that addition of anhydrous cerium chloride can suppress side reactions that occur in Grignard reactions, particularly those caused by

enolisable groups.¹⁶³ The effect is explained by the oxophilicity demonstrated by cerium chloride, causing activation of carbonyl groups. Cerium chloride also reduces the basicity of the Grignard reagent through co-ordination to the magnesium.¹⁶³ With this knowledge, the reaction was repeated in the presence of cerium chloride. The 3-pyridine Grignard reagent was formed *in situ* by addition of *i*PrMgCl.LiCl to 3-bromopyridine in THF at 20 °C. On completion of the Grignard formation (2 eq) the solution was added dropwise, by syringe, to a solution of the aldehyde (**2.72**) and CeCl₃ (2 eq) (**Scheme 7**, reaction (i)) giving 63% conversion by LCMS after 2 h. After work-up and purification, the desired alcohol (**2.77b**) was obtained in 24% yield as the formic acid salt.

Sufficient material for synthesis of the desired analogues, 3-piperidines (**2.63b**) and (**2.80**), had been isolated, and therefore, the reaction was not repeated. However, increasing the equivalents of CeCl₃ to 3 eq may have increased the yield of the Grignard addition to aldehyde (**2.72**), by reducing the basicity of the Grignard by co-ordination to the magnesium. In disagreement with this, literature was found suggesting that only catalytic amounts of CeCl₃ may be required.¹⁶⁴ If this chemistry was repeated, preformation of the cerium reagent, prior to addition of aldehyde (**2.72**), may also have a positive effect on the reaction yield.



Scheme 7 - Synthetic route to 3-piperidine analogues (**2.63b**) and (**2.80**). Reagents and conditions: (i) 3-bromopyridine, *i*PrMgCl.LiCl, CeCl₃, THF, 20 °C, 2 h, 24%; (ii) H₂, AcOH, 10% Pd/C, 80 °C, 30 bar, 45%; (iii) K₂CO₃, DMF, ClCH₂CO₂Et, 20 °C, 90 h, 92%; (iv) H₂, AcOH, 10 % Pd/C, 80 °C, 30 bar; (v) THF, EtOH, 2.0 N NaOH, 6 days, (7.7% over steps iv + v).

Simultaneous dehydroxylation and pyridine reduction of alcohol (**2.77b**) under acid mediated (acetic acid 0.05 M) hydrogenation conditions were performed using a H-cube hydrogenation system. High pressure (30 bar) and temperature (80 °C) were required to effect the transformations. Despite cycling the reaction mixture through the palladium on carbon cartridge four times, the reaction did not proceed to completion. LCMS analysis revealed mass ions for the desired product (**2.63**) in ~ 65% by UV, with the remaining starting material being converted to 2 close eluting peaks for the diastereomers formed from pyridyl reduction only, to form (**2.81**) and (**2.82**) (**Figure 58**). Piperidine (**2.63b**) was isolated by reverse phase purification in 45% yield as the formic acid salt. The by-products of the reaction (**2.81**) and (**2.82**) were not isolated.

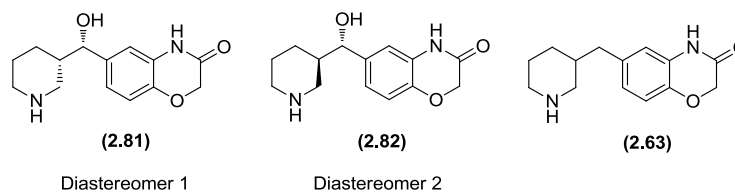


Figure 58 - Products from H-cube reduction (**Scheme 7**, reaction (ii)).

Disappointingly, 3-piperidine (**2.63b**) was inactive in the biochemical assay ($pIC_{50} < 3.3$). Based on this, efforts turned to anchoring the 3-piperidine analogue in the HIV IN_{CCD} by addition of the methylene carboxylic acid to give carboxylic acid (**2.80**) (**Scheme 7**).

Ester analogue (**2.78**) was synthesised by amide alkylation of intermediate (**2.77b**) using ethyl 2-chloroacetate and K_2CO_3 in DMF at room temperature (**Scheme 7**). Hydrogenation of the alcohol in (**2.78**) was carried out simultaneously with pyridine reduction using analogous conditions to those described previously. The ester of crude intermediate (**2.79**) was hydrolysed to the carboxylic acid using 2N NaOH in THF and EtOH to afford carboxylic acid (**2.80**). The low yield (7.7%) for this reduction and ester hydrolysis is explained by the zwitterionic nature of the final compound, which gave poor recovery from work-up and purification due to poor solubility in aqueous and organic solvents. Disappointingly, 3-piperidine (**2.80**) showed no biochemical activity and further analogues in this sub-series were not pursued.

2.11 Development of Synthetic Route to the 3-Pyrrolidine Analogue (**2.64**)

The final piperidine replacement for synthesis was the 3-pyrrolidine analogue (**2.64**) (**Figure 59**). As previously discussed, the 3-pyrrolidine analogue was designed to improve the angle of the H-bond donor to the Gln168 backbone carbonyl lone pair (**Figure 60**, diagram **b**) in comparison to the donor in 4-piperidine (**2.36**) (**Figure 60**, diagram **a**, arrow **a**); and also minimise the shift of the core away from the hypothesised optimal position of *t*-butyl fragment (**2.57**) (**Figure 60**, diagram **c**).

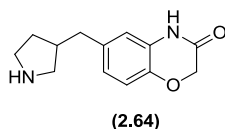


Figure 59 – 3-Pyrrolidine (2.64).

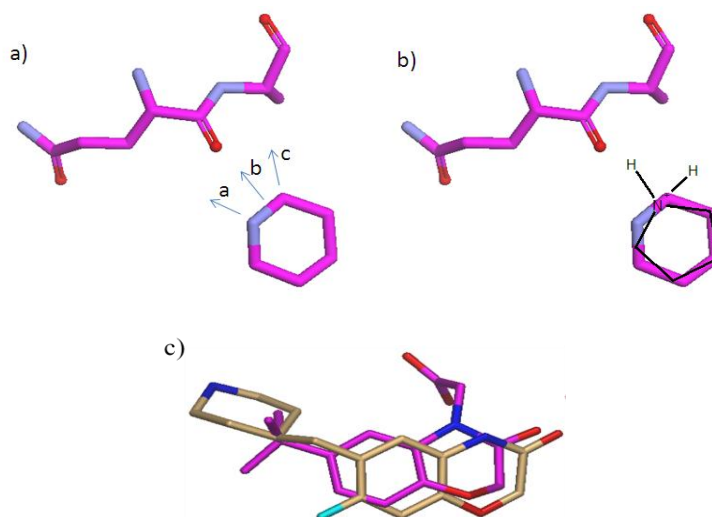
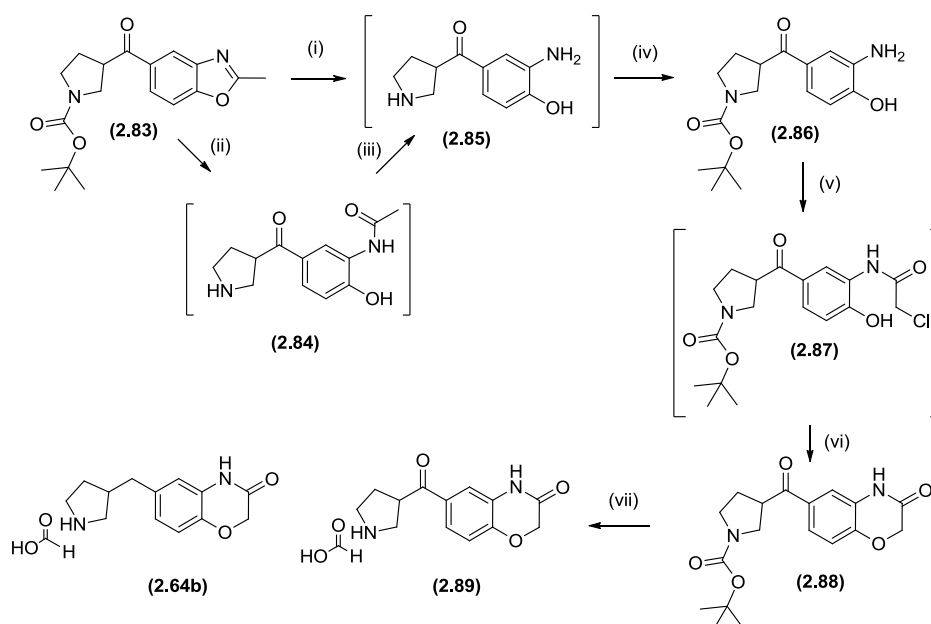


Figure 60 - a) Diagram demonstrating angles of H-bonds of Gln168 backbone carbonyl with fragment (2.36) piperidine; b) proposed pyrrolidine analogue overlaid with piperidine from the X-ray crystal structure of fragment (2.36); c) X-ray crystal structures of *t*-butyl (2.57) overlaid with piperidine (2.36) (PDB code: 1LIQU). It is hypothesised that *t*-butyl (2.57) (PDB code: 1WQHZ) adopts the preferred position for the core.

An alternative route to synthesis of pyrrolidine (2.64) was explored based on the learnings from the pyridine synthesis. The route involved the introduction of the pyrrolidine ring early in the synthesis and formation of the benzoxazine ring from 1,3-benzoxazole. The benzoxazole intermediate (2.83), available within our laboratories, was ring opened in ethanol with 5 M HCl at 70 °C in the microwave to afford *N*-acetyl compound (2.84). Further heating of the reaction at 100 °C afforded the desired aniline (2.85). As one would expect, the two steps could be performed simultaneously by heating (2.83) at 100 °C. Selective Boc protection of the more

nucleophilic pyrrolidine nitrogen was achieved giving a 70% yield over steps i and iv in **Scheme 8**.



Scheme 8 - Synthetic route to 3-pyrrolidine analogue (**2.64b**). Reagents and conditions: (i) 5.0 N HCl, EtOH, μ Wave, 100 °C, not isolated; (ii) 5.0 N HCl, EtOH, μ Wave, 70 °C, not isolated; (iii) 5.0 N HCl, EtOH, μ Wave, 100 °C, not isolated; (iv) Et_3N , Boc_2O , DCM, 20 °C, (70% over steps i and iv); (v) ClCOCH_2Cl , MeCN, K_3PO_4 , 20 °C, 21 h, then, (vi) 5.0 M NaOH(aq), 20 °C, 1 h, (83%, over steps v and vi); (vii) TFA, Et_3SiH , 20 °C, 24 h, 19% (**2.64b**), 16% (**2.89**).

Amide (**2.87**) was formed from addition of chloroacetyl chloride to a solution of K_3PO_4 and aniline (**2.86**), in anhydrous MeCN at 0 °C followed by warming to room temperature. Use of the weak base potassium K_3PO_4 , allowed selective acylation of the aniline (**2.86**) without concomitant acylation of the phenol. Increasing the pH to ~ 14, by addition of 5 M aqueous NaOH, resulted in deprotonation of the phenol, enabling cyclisation, *via* the favoured 6-*exo*-tet pathway to afford benzoxazinone (**2.88**) in 83% yield (over 2 steps, **Scheme 8**). The ketone was reduced with simultaneous Boc deprotection with TFA and Et_3SiH to give the 3-pyrrolidine (**2.64b**) in 19% yield. Additionally the ketone intermediate (**2.89**) was isolated in 16% yield.

2.12 Biochemical and Crystallography Results for 3-Piperidine and 3-Pyrrolidine Analogues

Both the 3-piperidine and 3-pyrrolidine analogues, designed to improve the H-bond interaction with Gln168, were below the limit of the biochemical assay (**Figure 61**). The 3-pyrrolidine (**2.64b**) was active on the first screening occasion ($pIC_{50} = 4.3$, $LE = 0.35$, $LLE_{(AT)} = 0.37$), but on retest was confirmed to be below the limit of the biochemical assay ($pIC_{50} < 4.3$, $n = 1$; and $pIC_{50} < 3.3$, $n = 2$).

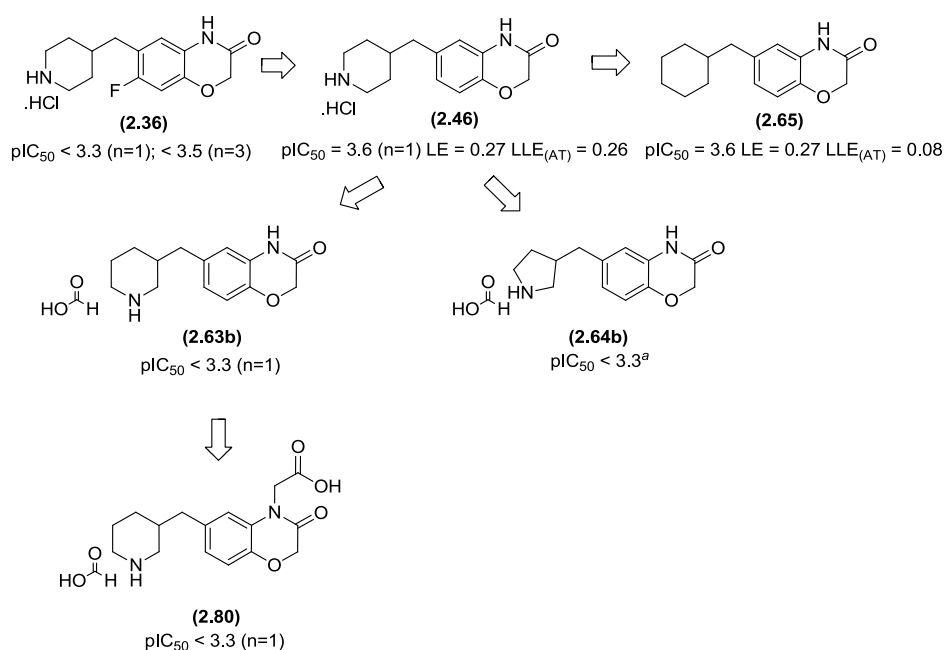


Figure 61 - Biochemical data for 4-piperidine replacements. ^aThis compound was screened on 4 screening occasions ($pIC_{50} = 4.27$, < 4.3 , < 4.3 , < 3.3) – the active result is classed as a false positive.

Although the piperidine and pyrrolidine analogues did not have an increase in biochemical potency, it was important to determine whether the modelling hypothesis was correct by obtaining X-ray crystal structures. The 3-piperidine (**2.80**) did indeed form a H-bond interaction with the Gln168 backbone carbonyl (Gln168 O – NH, 2.89 Å) (**Figure 62**), with an optimal angle in the plane of the Gln168 backbone carbonyl. Additionally, the H-bond angle of 3-piperidine (**2.80**) to Gln168 (Gln168 O – NH, 2.89 Å, C-O-HN = 164 °) is marginally improved in comparison to

the 4-piperidine analogue (**2.36**) (Gln168 O – NH, 2.92 Å, C-O-HN = 168 °) (**Figure 62**). This angle could be further improved to the optimal angle of 120 °C.

The piperidine (**2.63b**) was predicted to have an improved H-bond angle to Gln168, compared to 3-piperidine (**2.80**), as the core would be more amenable to movement due to the lack of the carboxylic acid tether. However, crystallography failed to provide a crystal structure of 3-piperidine (**2.63b**), and this hypothesis remains unvalidated. The methylene linked carboxylic acid group of (**2.80**) makes two H-bonds with backbone nitrogen of Glu170 (Glu170 NH-O, 3.00 Å) and His171 (His171 NH-O, 3.13 Å). There is little movement of the benzoxazinone core along the pocket compared to (**2.36**), but a twist of the core to improve the distances of the H-bonds to Glu170 and His171 is observed, **Figure 62**. It is hypothesised that this twist does not affect the interactions with the base of the protein pocket, as it is the solvent exposed region of the molecule which has the greatest movement compared to 4-piperidine (**2.36**).

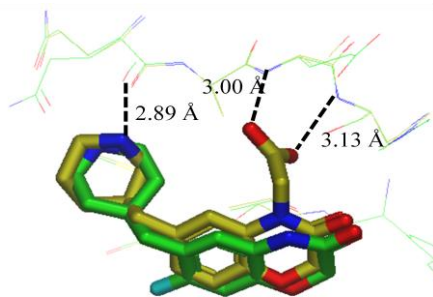


Figure 62 - Overlay of X-ray crystal structures of 4-piperidine (**2.36**) (PDB code: 1LIQU) with 3-piperidine (**2.80**) (PDB code: 5HJH7) in HIV IN_{CCD}.

X-ray crystallography of the 3-pyrrolidine analogue (**2.64b**) (**Figure 63**) also demonstrated an improved H-bond angle with the Gln168 backbone carbonyl (Gln168 O – HN, 3.08 Å, C-O-HN = 154 °) compared to 4-piperidine (**2.36**) (Gln168 O – NH, 2.92 Å, C-O-HN = 168 °). Despite the improved angle, it was observed that the H-bond distance of 3-pyrrolidine (**2.64b**) (3.08 Å) to Gln168 backbone carbonyl was longer than the H-bond to Gln168 from the 4-piperidine analogue (**2.36**) (2.92 Å). It can be seen that the benzoxazinone core of pyrrolidine (**2.64b**) shifts along the

pocket towards Gln168 in comparison to the original cluster 2 hit (**2.36**), but remains in the same plane (**Figure 63**). The combination of an increased H-bond distance to Gln168 and the movement along the protein pocket were expected to be detrimental to binding potency.

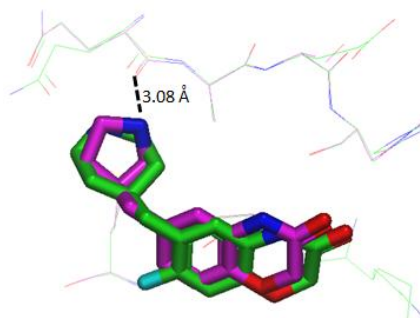


Figure 63 - Overlay of X-ray crystal structures of 4-piperidine (**2.36**) (PDB code: 1LIQU) with 3-pyrrolidine (**2.64b**) (PDB code: 66GKH) in HIV IN_{CCD}.

Overlays of the crystal structures of the original cluster 2 hit (**2.36**), the 3-pyrrolidine analogue (**2.64b**) and *t*-butyl analogue analogue (**2.57**) (**Figure 64**) shows that the movement of the benzoxazinone core of pyrrolidine (**2.64b**), has positioned the core closer to the position of that of *t*-butyl (**2.57**). The *t*-butyl species (**2.57**), *vide supra*, was thought to be in a more optimal position compared to the original cluster 2 hit (**2.36**). The *t*-butyl analogue core (**2.57**) twisted towards Glu170 and His171 to pick up the two H-bond interactions with the backbone in comparison to 3-pyrrolidine (**2.64b**).

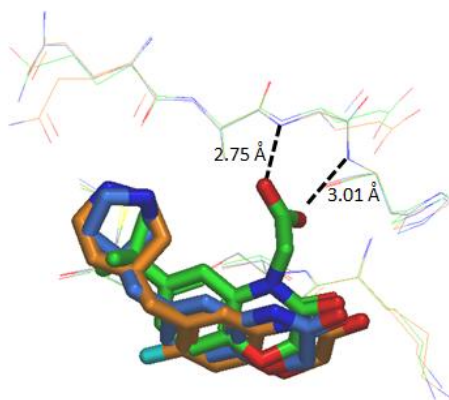


Figure 64 - Overlay of X-ray crystal structures of 4-piperidine (**2.36**) (PDB code: 1LIQU), 3-pyrrolidine (**2.64b**) (PDB code: 66GKH), and *t*-butyl (**2.57**) (PDB code: 1WQHZ) in HIV IN_{CCD}.

3-Pyrrolidine analogue (**2.64b**) and 3-piperidine compound (**2.80**) both had activity below the limit of the biochemical assay. Despite the lack of biochemical activity, crystallography confirmed that there were good quality H-bonds, demonstrating that forming new polar interactions does not always lead to an increase in potency. It was hypothesised that this was due to the greater desolvation (entropic) penalty associated with removing the fragment from the bulk solvent on binding to the protein (see Thermodynamic Binding contributions – **Section 1.8.1**).

In addition to this thermodynamic penalty, there were concerns over the ability of the TR-FRET biochemical assay to measure compounds binding deep in the HIV IN_{CCD} pocket, such as 4-piperidine (**2.36**) and 3-piperidine (**2.63b**) (**Figure 65**). For measurement in the biochemical assay, the HIV IN_{CCD}-LEDGF protein-protein interaction needs to be inhibited. The HIV IN_{CCD}-LEDGF interaction is made up from lipophilic contacts around the outer surface of the HIV IN_{CCD} pocket. LEDGF does not protrude deep into the HIV IN_{CCD} binding pocket. Growing out of the pocket from a well-anchored fragment, towards the location where LEDGF binds, was hypothesised to improve the chance of inhibiting the protein-protein interaction allowing robust SAR to be generated. Ideally ITC experiments would be used to identify the most strongly anchored fragment from comparison of K_d values.

Unfortunately, due to resource constraints, ITC experiments could not be carried out, so consideration turned to identifying the most appropriate fragment to grow out of the pocket to maximise disruption of the protein-protein interaction.

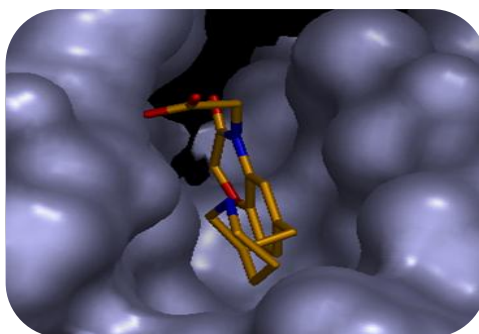


Figure 65 - 3-Piperidine analogue (**2.80**) ($pIC_{50} < 3.3$ ($n=1$)) sits deep in the HIV IN_{CCD} pocket (PDB code:5HJH7).

2.13 Amide Substitution to Interact with the Lipophilic Shelf

Growth of fragments out of the binding pocket towards solvent was an uncommon fragment optimisation strategy as it was unlikely to lead to an increase in binding energy, compared to maximising interactions deep in the protein pocket. Consideration of the binding mode of quinoline (**2.35**) confirmed that a phenyl ring (B-ring) sat on a lipophilic shelf in an area where LEDGF binds out towards solvent (**Figure 66**). For the related quinolinone template (**2.32**) which bound in an analogous fashion, removal of the B-ring to give pyridone (**2.33**), resulted in a significant reduction in activity, suggesting the lipophilic shelf as an important area for gaining biochemical potency.

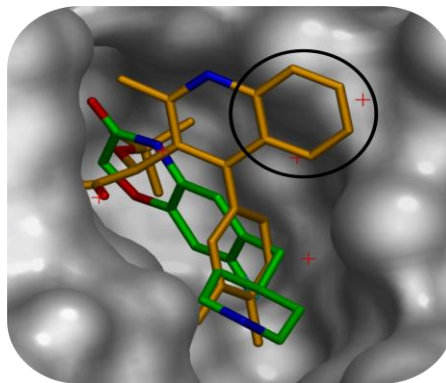
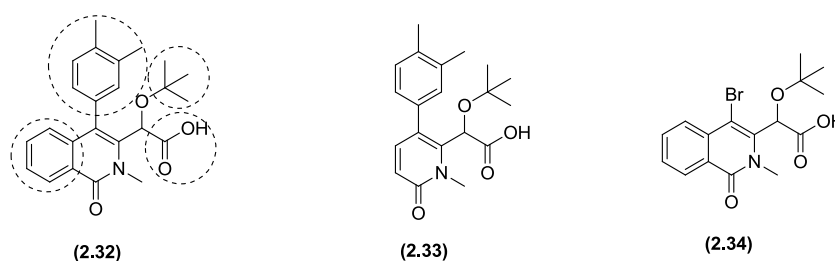


Figure 66 - Overlay of quinoline (**2.35**) (PDB code: 1OVDR) with piperidine (**2.36**) (PDB code: 1LIQU) in HIV IN_{CCD}. The B-ring occupying the lipophilic shelf is circled.



pIC_{50} 6.6, LE = 0.31, $LLE_{(AT)}$ = 0.23 pIC_{50} = 5.0, LE = 0.27, $LLE_{(AT)}$ = 0.23 pIC_{50} = 3.9, LE = 0.24, $LLE_{(AT)}$ = 0.19

Figure 67 – Fragmentation of quinolinone (**2.32**) demonstrates a significant reduction in potency by removal each of the aromatic rings. Four remote groups (circled on quinolinone (**2.32**)) appear essential to achieve the required level of HIV IN_{CCD}-LEDGF inhibition in this series. *“This compound has only been reported active in 2 out of 3 screening occasions – 3rd screening occasion measured < 4.3.*

The potency provided by the B-ring of quinolinone (**2.32**), is rationalised by overlap of the B-ring with multiple residues of LEDGF inhibiting formation of the HIV IN_{CCD}-LEDGF complex. All cluster 2 fragments bound deep in the HIV_{CCD} protein pocket and have minimal overlap with LEDGF which corresponds with the lack of measurable biochemical activity. To date the *t*-butyl compound (**2.57**) (**Figure 68**, diagram c) (pIC_{50} = 4.4, LE = 0.32, $LLE_{(AT)}$ = 0.23) demonstrated one of the highest potencies with good ligand efficiency. The biochemical assay reproducibility of

(**2.57**), accompanied by successful X-ray crystallography and binding in NMR STD studies, gave high confidence in the validity of the hit.

Based on all of this, efforts turned to growing onto the lipophilic shelf to maximise steric clash with LEDGF. Docking of benzyl fragment (**2.90**) into the HIV IN_{CCD} binding pocket showed that the benzyl group extruding from the methylene linker would interact with the lipophilic shelf, whilst picking up H-bonds with Glu170 and His171 from the carboxylic acid moiety (**Figure 68**, diagram **a**).¹⁶⁵ Despite increasing, the cPFI for benzyl (**2.90**) still remains in the desired range at cPFI = 5.6.

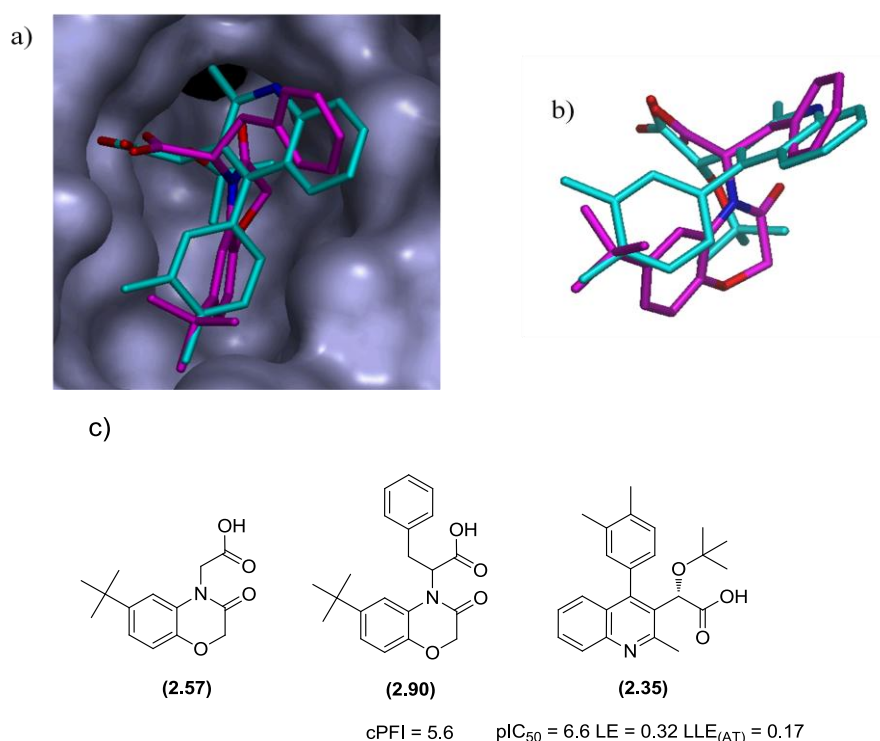


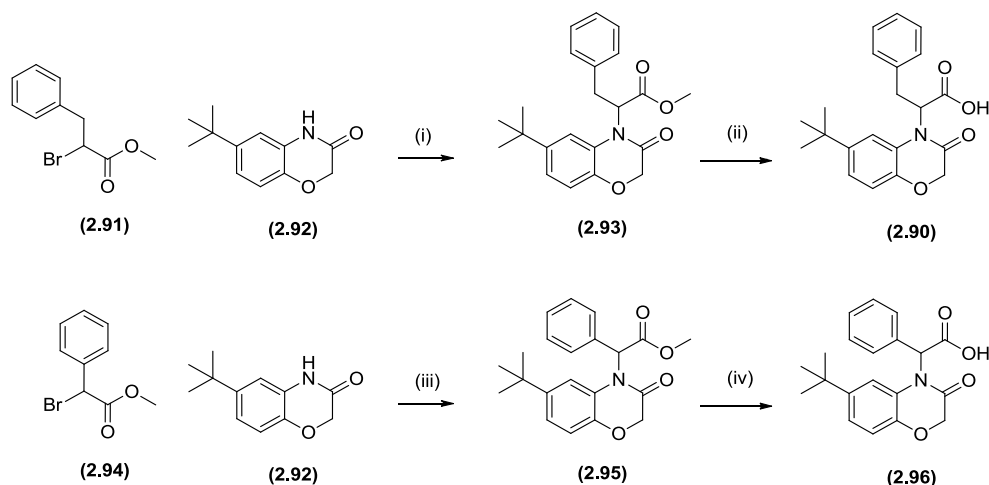
Figure 68 – (a) Overlay of X-ray crystal structures of compound (**2.35**) (PDB code: 1ODVR) with docking of fragment (**2.90**) in HIV IN_{CCD}; (b) Overlay of X-ray crystal structure of compound (**2.35**) with docking of fragment (**2.90**) – not showing HIV IN_{CCD}; and c) structures of relevant molecules.

Synthesis of benzyl analogue (**2.90**)¹⁶⁶ initially proved challenging. Alkylation on the amide nitrogen of commercially available benzoxazinone (**2.92**) with alkyl

bromide (**2.91**) (**Scheme 9**) did not proceed as expected using the standard methodology of using K_2CO_3 as base and DMF as the solvent. Forming the corresponding iodide *in situ*, by addition of tetrabutylammonium iodide, was also unsuccessful. Use of a stronger base, LHMDS, to formally deprotonate the amide nitrogen of (**2.92**), followed by addition of the alkylating agent, gave the required product albeit in low yield (12%) (**Scheme 9**). The low yield for the formation of methyl ester (**2.93**) (**Scheme 9**) could be due to:

1. the hindered nature of the centre being formed;
2. LHMDS deprotonating the methylene protons between the amide carbonyl and the ether oxygen on the benzoxazinone core of (**2.92**), ($pK_a = 25 - 30$);
3. formation of the α,β -unsaturated byproduct from bromide (**2.91**) – tentative assignment by LCMS.

Further optimisation of the reaction was deemed unnecessary as ester hydrolysis of (**2.93**) to give the acid (**2.90**) provided sufficient material for the required assays. The phenyl analogue (**2.96**) was synthesised using the K_2CO_3 alkylation procedure on (**2.92**), as for previous analogues, in preference to the LHMDS conditions due to concerns of deprotonation occurring at the methylene on the cluster 2 core.



Scheme 9 - Synthetic route to the methylene acid analogues substituted with benzyl (**2.90**) or phenyl moieties (**2.96**).¹⁶⁶ Reagents and conditions: (i) LHMDS, DMF, RT, 12%; (ii) LiOH, THF, MeOH, 38%; (iii) methyl 2-bromo-2-phenylacetate, K₂CO₃, DMF, 20 °C, 27 h, 55%; (iv) LiOH, MeOH, THF, 20 °C, 45%.¹⁶⁶

Despite both analogues demonstrating biochemical activity, phenyl (**2.96**) (pIC₅₀ = 4.2, LE = 0.23, LLE_(AT) = 0.11) and benzyl (**2.90**) (pIC₅₀ = 3.8, LE = 0.20, LLE_(AT) = 0.08) gave a reduction in ligand efficiencies compared to *t*-butyl (**2.57**) (pIC₅₀ = 4.4, LE = 0.32, LLE_(AT) = 0.23). Due to the ligand efficiency decrease, the benzyl of the molecule is either not occupying the lipophilic shelf as expected from docking, or the interaction on the lipophilic shelf is not as beneficial as anticipated. X-ray crystallography was unsuccessful and, therefore, a firm conclusion cannot be made.

The lack of potency increase seen by growing onto the lipophilic shelf, led to consideration of whether *t*-butyl compound (**2.57**) was well anchored in the protein pocket as first hypothesised. Fragment (**2.57**) and structurally related analogues (**2.97**),¹⁶⁶ (**2.98**)¹⁶⁶ and (**2.52**) do appear well anchored in the HIV IN_{CCD} pocket, as little movement is seen from the different analogues in X-ray crystallography (**Figure 69**). The carboxylic acid of the cluster 2 fragments shown in **Figure 69** forms two H-bond interactions with backbone nitrogen of Glu170 and His171. In addition to these H-bond interactions, the fragments are held in position by the good shape complementarity of the fragments' core with the lipophilic pocket.

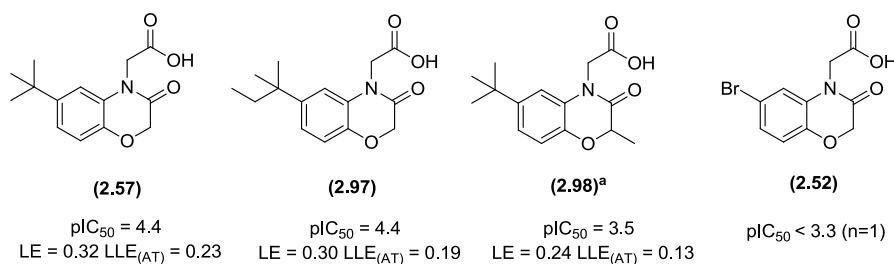
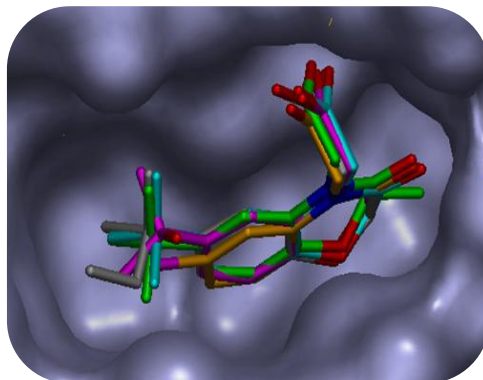


Figure 69 - Little movement of the fragments in the HIV IN_{CCD} pocket is observed when fragment (2.57) (PDB code: 1WQHZ) is overlaid with structurally related analogues (2.97) (PDB code: 52SDT), (2.98) (PDB code: 1GWMZ), and (2.52) (PDB code: 1VLPL). ^aThis compound has only been reported active in 2 out of 4 screening occasions – ($pIC_{50} = 3.31, <3.3, <4.3, 3.6$).

Initial exploration to grow onto the lipophilic shelf from fragment (2.57) (Figure 70) has resulted in a reduction in biochemical potency and ligand efficiencies. Addition of a benzyl group which picks up good lipophilic contacts is expected to give ~ 1000 fold increase in potency to maintain the $LLE_{(AT)}$.¹⁵ Crystal structures of methyl (2.99)¹⁶⁶ and phenyl (2.96)¹⁶⁶ analogues show that there is minimal movement in the position of the core compared with the core of (2.57) (Figure 70). In methyl analogue (2.99) ($pIC_{50} = 4.1$, $LE = 0.28$, $LLE_{(AT)} = 0.18$), the methyl group points directly towards the lipophilic shelf.

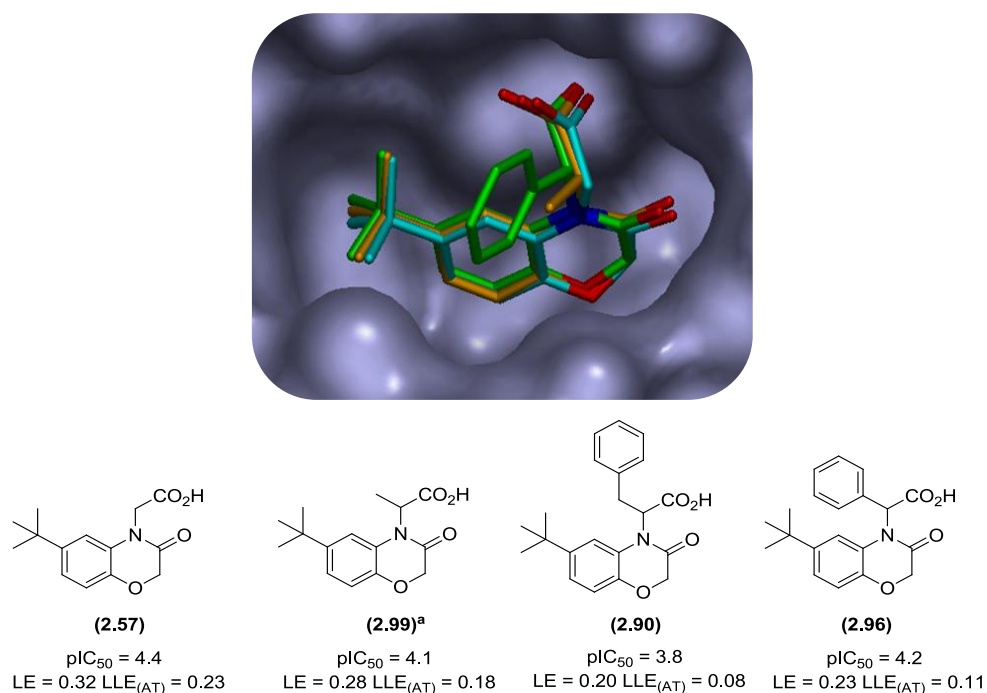


Figure 70 – Initial compounds synthesised to reach onto the lipophilic shelf.

Overlay of X-ray crystal structures of **(2.57)** (PDB code: 1WQHZ), **(2.99)** (PDB code: 249WD), and **(2.96)** (PDB code: 5EV7R). *“This compound has only been reported active in 2 out of 3 screening occasions – (pIC₅₀ = 3.98, 4.21, <4.3).*

Phenyl analogue **(2.96)** (pIC₅₀ = 4.2, LE = 0.23, LLE_(AT) = 0.11) gave equivalent biochemical potency to methyl analogue **(2.99)**, resulting in lower ligand efficiency due to the additional heavy atoms in the phenyl-based system. X-ray crystallography of phenyl analogue **(2.96)** showed that the phenyl ring sits orthogonal to the desired orientation to lay flat on the lipophilic shelf. Quantum mechanical (QM) calculations for phenyl analogue **(2.96)**¹⁰⁰ demonstrate that the lowest energy orientation for the phenyl ring would be 77 ° away from the desired plane. The lowest energy conformation for the phenyl ring is similar to that observed in the X-ray crystal structure of phenyl analogue **(2.96)** (**Figure 70**). The desired orientation for the phenyl ring in fragment **(2.96)** is out of the desired plane due to the clash of the *ortho* phenyl protons with the proton at the stereogenic centre and the proton at the 5-position of the core (**Figure 71**). The phenethyl analogue **(2.90)** showed the lowest biochemical potency and ligand efficiency of all the compounds designed to target

the lipophilic shelf, despite it having extra flexibility to theoretically orientate itself into the correct plane. However, the increased flexibility results in an entropic penalty, which could be detrimental to the binding affinity for HIV IN_{CCD}.

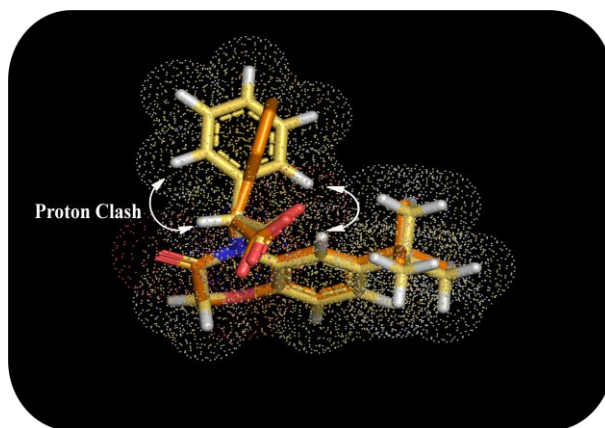


Figure 71 - Fragment **(2.96)** lowest energy conformation overlaid with the desired orientation of fragment **(2.96)**.

The phenyl ring of **(2.96)** occupied the correct position in the pocket to interact with the lipophilic shelf; however the ring is twisted out of the required plane for maximum lipophilic interactions. Priority was then given to stabilising the desired phenyl orientation; by increasing the stability of the high energy conformation and making it lie flat on the lipophilic shelf. By tying the phenyl or benzyl group round to form a quaternary carbon centre (as in the fragments shown in **Figure 72**), the energy barrier to binding could be reduced. For the phenyl to sit in the correct orientation on the lipophilic shelf, QM calculations¹⁰⁰ suggest that a 4.2 kcal energy penalty needs to be overcome, which equates to approximately 1000-fold reduction in potency.¹⁶⁷ Restricting the movement of the phenyl by synthesis of the quaternary carbon centre analogues (**Figure 72**) could, therefore, theoretically give up to a 1000-fold increase in potency compared to phenyl **(2.96)**.

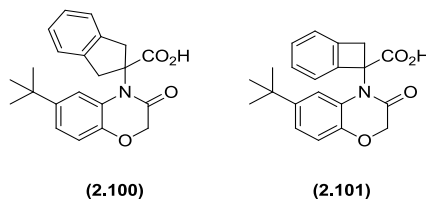


Figure 72 – Constraining the orientation of the benzyl or phenyl ring to maximise contacts with the lipophilic shelf.

2.14 Synthesis of Quaternary Centre Analogues to Stabilise the Conformation of the Groups Occupying the Lipophilic Shelf

Consideration turned to the design of a small set of quaternary carbon molecules, to test the effect of stabilisation of the desired phenyl conformation to interact more efficiently with the lipophilic shelf (**Figure 73**). Although the molecules of interest (**2.100-2.103**) are highly lipophilic (cLog P = 5.0 – 5.7) it was rationalised that the lipophilicity could be decreased in subsequent iterations, for example by addition of a nitrogen giving pyridine analogues such as (**2.104**). Replacing one carbon atom for one nitrogen atom decreases the cLogP by approximately 1.4 units, improving the LLE_(AT), thus enhancing the overall developability properties of these constrained templates. The pyridine nitrogen was expected to be tolerated as it would point out towards solvent.

It was decided that synthesis of all four quaternary carbon analogues should be attempted, as it was still unclear if the quinoline ring from (**2.35**) was in the optimal position. It is possible that a small deviation from the quinoline position may be beneficial for binding efficiency and biochemical potency. To understand the use of and validate the docking methodology, the quaternary centre compounds were ranked in an attempt to predict which would bind more efficiently. Docking of the quaternary centre molecules (**Figure 73**),¹⁰⁰ were each overlaid with the X-ray crystal structure of quinoline (**2.35**) and as reference benzoxazinone (**2.57**). From the overlay it was determined how well the designed analogues were predicted to occupy the area filled by the quinoline of (**2.35**). The docking positioned the 5,6 asymmetrical fragment (*R*-enantiomer) (**2.103**) with the best overlap with the

quinoline core of **(2.35)** (**Figure 73**, diagram **c**). The 4,6 system (*R*-enantiomer) **(2.101)** lies in a similar position but does not reach as far out over the lipophilic shelf in comparison to **(2.35)** (**Figure 73**, diagram **a**) or 5,6 analogue **(2.100)**. The 6,6 fragment (*R*-enantiomer) **(2.102)** does not overlay as well with the quinoline from **(2.35)** (**Figure 73**, diagram **d**), but does reach further out in comparison to the 4,6 **(2.101)** and 5,6 asymmetrical **(2.103)** fragments. From consideration of the docking study, the 5,6 symmetrical system **(2.103)** was deemed the least likely of the 4 compounds to occupy the area of the lipophilic shelf filled by the quinoline of **(2.35)** (**Figure 73**, diagram **b**).

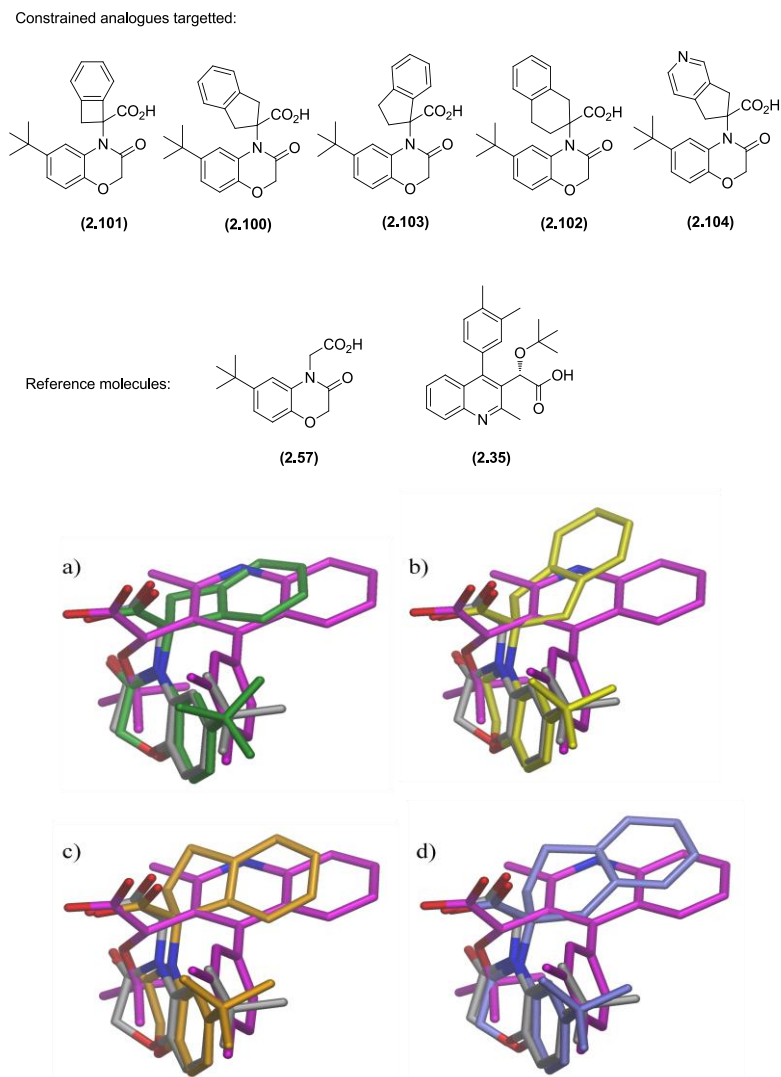
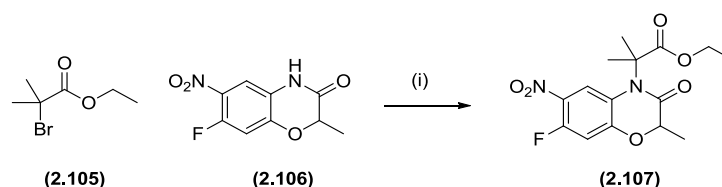


Figure 73 – a) Docking of **(2.101)** overlaid with X-ray crystal structures of **(2.57)** and **(2.35)**; b) Docking of **(2.100)** overlaid with X-ray crystal structures of **(2.57)** and **(2.35)**; c) Docking of **(2.103)** overlaid with X-ray crystal structures of **(2.57)** and **(2.35)**; and d) Docking of **(2.102)** overlaid with X-ray crystal structures of **(2.57)** and **(2.35)**.

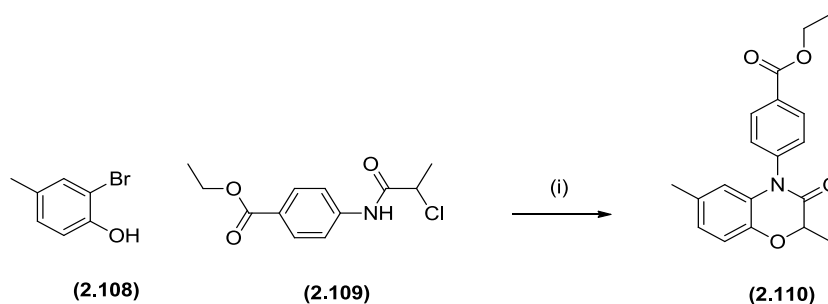
Based on literature precedence, the following methodologies were considered for the preparation of the required analogues. Alkylation to form *N*-amide substituted analogues such as **(2.107)** had been reported by Huang and co-workers in good yield (62 – 83% yield) (**Scheme 10**).¹⁶⁸ Synthesis within our laboratory using alkylation to form phenyl substituted carboxylic acid **(2.91)**, had demonstrated only a moderate

yield for this transformation (**Scheme 9**).¹⁶⁶ Formation of a quaternary centre analogue by *N*-amide alkylation was likely to be even more challenging due to greater steric hindrance around the bond forming centre. The desired analogues would be more sterically hindered than the literature example, or examples previously attempted in our laboratory, and therefore this chemistry was not attempted.



Scheme 10 - *N*-Amide alkylation on benzoxazinone core.¹⁶⁸ Reagents and conditions: (i) K_2CO_3 , MeCN, reflux 3 h, reported yields for all alkylations is 62-83% (yield not specified for individual reactions).

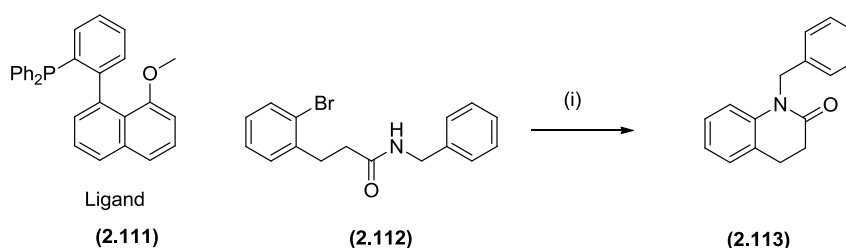
Benzoxazinone core formation *via* intramolecular copper-mediated cyclisation of the amide nitrogen onto aryl bromides had been reported by Feng and co-workers (**Scheme 11**).¹⁶⁹ In this reaction alkylation of the phenol occurs before cyclisation of the amide nitrogen onto the aryl bromide. As the final step is intramolecular rather than intermolecular, this assists formation of the desired compound.



Scheme 11 - Intramolecular amide and aryl-bromide copper-mediated cyclisation to form the benzoxazinone core.¹⁶⁹ Reagents and conditions: (i) CuI (20 mol%), Cs_2CO_3 , DBU, DMSO, 3 h, 130 °C microwave irradiation, 59%.

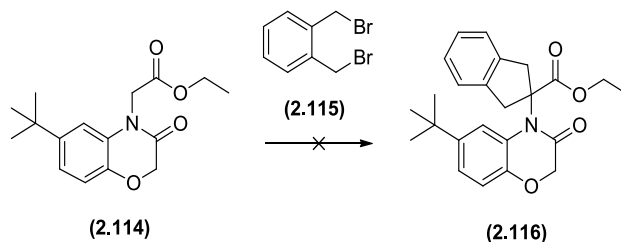
An alternative method considered was intramolecular palladium-catalysed cyclisation. Furuta and co-workers reported the development of phenylnaphthyl

phosphines for the application of Pd-catalysed intramolecular amidations (**Scheme 12**).¹⁷⁰ Lactam (**2.113**) was formed in 91% yield by heating benzyl amide (**2.112**) with Pd(OAc)₂, Cs₂CO₃ and ligand (**2.111**) in 1,4-dioxane at 100 °C for 3.5 h. No analogues with substituents on the methylene attached to the amide nitrogen were reported in this paper. The effect of methylene substitution on the reaction was unclear, it may hinder the reaction due to sterics reducing the yield, or alternatively, provide a restricted rotamer effect enhancing the yield.



Scheme 12 - Intramolecular palladium-mediated cyclisation.¹⁷⁰ Reagents and conditions: (i) Pd(OAc)₂ (6.0 mol%), Cs₂CO₃, ligand (**2.111**) (5.0 mol%), 1,4-dioxane, 100 °C, 3.5 h, 91%.

An expedient means of forming the symmetrical 5,6 quaternary centre analogue (**2.100**) was *via* a *di*-alkylation reaction of the ester α centre of (**2.114**) (**Scheme 13**). Disappointingly, attempts to perform this reaction by deprotonating with excess LHMDS failed to effect any reaction.¹⁷¹ It was anticipated that deprotonation and subsequent alkylation could also occur at the methylene between the amide carbonyl and ether oxygen on the benzoxazinone core, giving mixtures of products. LCMS analysis showed that no alkylation at either methylene centre had occurred. Only starting material remained in the reaction mixture and no side products were observed. It is presumed that LHMDS is basic enough to affect the desired deprotonation, but the anion was not sufficiently reactive to effect the alkylation. Increasing the anion nucleophilicity through solvent variation was not explored in favour of alternative precedented methodology.



Scheme 13 – Di-alkylation of ester α centre. Reagents and conditions: (i) LHMDS (2.0 M in THF) (5 eq), 20 °C.

The final strategy considered to form the quaternary carbon analogues was through the use of quaternary carbon amino acid and amino ester intermediates prepared within our laboratories.¹⁷² The requisite amino acids could be synthesised by Strecker chemistry from the analogous ketones, thus enabling a convergent synthetic strategy using common intermediates to investigate the orientation of groups on the lipophilic shelf.

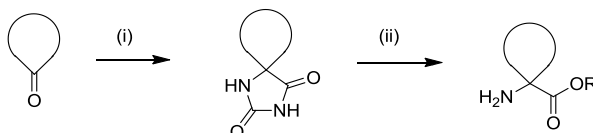
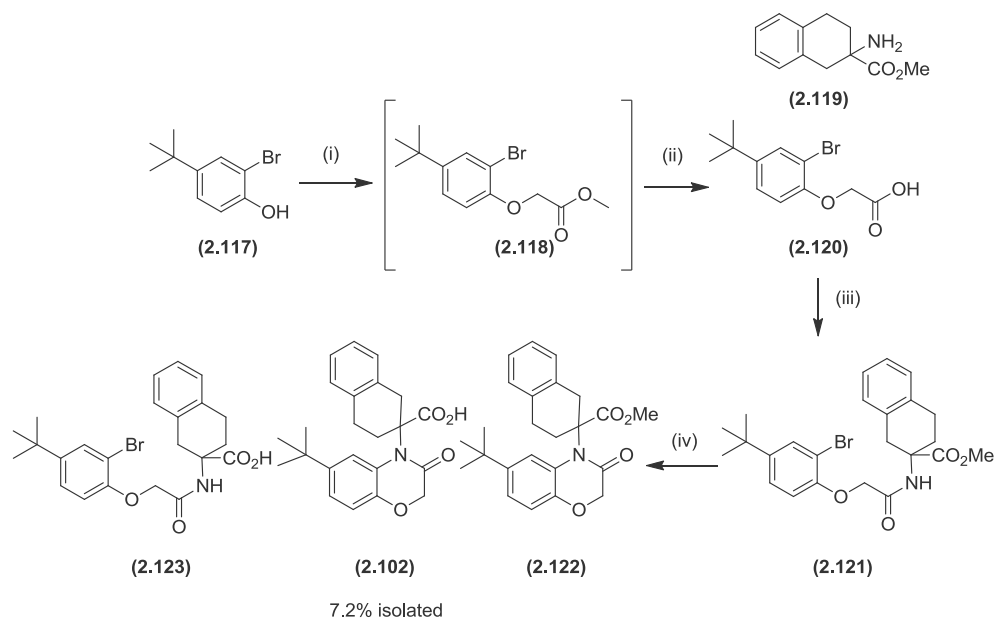


Figure 74 – Synthetic route to the required α,α -disubstituted amino esters from commercially available ketone analogues. Reagents and conditions: (i) KCN, $(\text{NH}_4)_2\text{CO}_3$; (ii) 2M NaOH, μWave , 160 °C.¹⁷²

2.15 Copper-mediated Cyclisation Chemistry to Deliver Quaternary Carbon Centre Analogues

It was envisaged that forming the carbon quaternary centre early before introduction onto the benzoxazinone template, may improve the chance of obtaining a synthetic strategy to access the quaternary carbon analogues of interest (**Figure 73**). The commercially available amino ester (**2.119**) was used to investigate the suitability of the route to the quaternary carbon analogues (**Scheme 14**). Intermediate (**2.118**) was formed *in situ* by alkylation of the phenol (**2.117**) with methyl chloroacetate. The ester was then hydrolysed to afford carboxylic acid (**2.120**) in 27% yield over two

steps. The commercially available amino ester (**2.119**) was added to a solution of the acid chloride formed from (**2.120**) in THF, to afford amide (**2.121**) in a 35% yield.

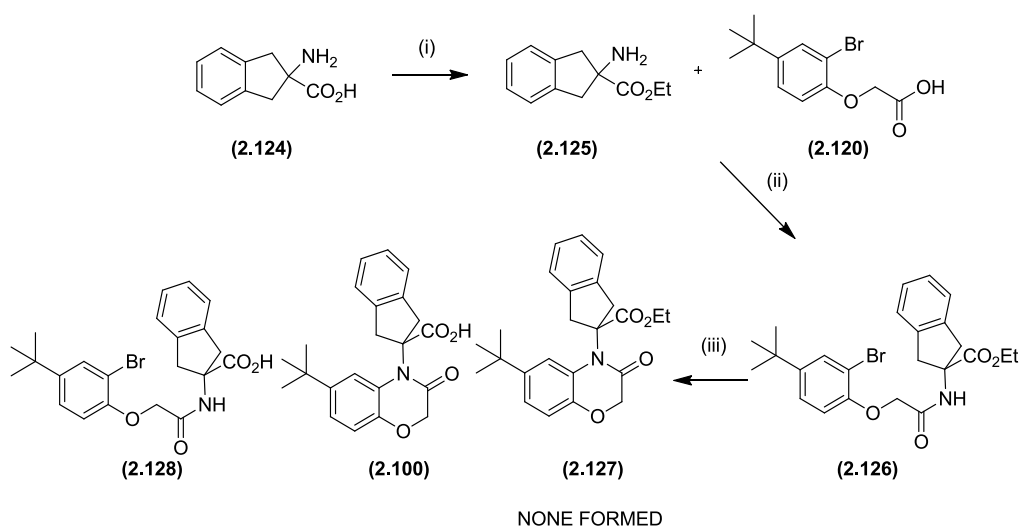


Scheme 14 – Copper-mediated cyclisation route to quaternary carbon centre analogues using α,α -disubstituted amino methyl ester (**2.119**). Reagents and conditions: (i) K_2CO_3 , methyl chloroacetate, DMF, RT, overnight, not isolated; (ii) 2.0 N NaOH, THF, MeOH, 20 °C, 44 h, 27 % over 2 steps; (iii) (1-Chloro-2-methyl-1-propen-1-yl)dimethylamine, DIPEA, THF, 20 °C, 30 min, 35%; (iv) CuI, CS_2CO_3 , DBU, DMSO, μ Wave, 130 °C, 1 h 40 min, 7.2% (**2.102**).

The key step in this reaction sequence was the intramolecular cyclisation¹⁶⁹ of the amide nitrogen in (**2.121**) onto the aryl bromide forming benzoxazinone (**2.122**). This reaction was expected to be problematic due to the formation of the hindered centre. The reaction profile by LCMS showed formation of small amounts of the desired carboxylic acid product (**2.102**) and hydrolysed starting material (**2.123**). There was no evidence of ester (**2.122**) and the majority of material observed by LCMS was starting material. Addition of further base and CuI with further heating did not appear to push the reaction further. Pleasingly, after purification, the benzoxazinone (**2.102**) was isolated in 7.2% yield. It was initially thought that a combination of the CuI and / or moisture in the DMSO with heating, may be the

cause of ester cleavage, and the resultant carboxylic acid could chelate to the copper preventing the intramolecular cyclisation reaction from proceeding. This hypothesis warranted further attention during synthesis of the quaternary carbon analogues.

A sufficient quantity of 6,6-analogue (**2.102**) had been isolated for the appropriate testing, so optimisation of reaction conditions based on these observations was performed during the synthesis of the 5,6-symmetrical analogue (**2.100**) (Scheme 15). In an effort to reduce the production of the carboxylic acid, which we believed contributed to the low yield for the copper-mediated cyclisation to form the 6,6-product (**2.102**), we substituted the methyl ester for the ethyl ester as this was predicted to be less prone to hydrolysis under the reaction conditions.



Scheme 15 - Copper-mediated cyclisation route to quaternary carbon centre analogues using α,α -disubstituted amino ethyl ester (**2.125**). Reagents and conditions – (i) c.H₂SO₄, EtOH, 50 - 20 °C, 116 h, 71%; (ii) (1-Chloro-2-methyl-1-propen-1-yl)dimethylamine, DIPEA, THF, 20 h, 42%; (iii) CuI (20 mol%), Cs₂CO₃, DBU, microwave, DMSO, 130 °C, 4.5 h, 16% (**2.100**).

The α,α -disubstituted amino ethyl ester (**2.125**) was formed by heating commercially available amino acid (**2.124**) in ethanol under acidic conditions (pH = 1). Reaction

of the α,α -disubstituted amino ethyl ester (**2.125**) with carboxylic acid (**2.120**) afforded the amide (**2.126**) which was subjected to the copper-catalysed intramolecular cyclisation conditions. Similar reaction profiles were seen (after 30 min of heating in the microwave at 130 °C) for formation of 6,6-carboxylic acid (**2.123**) and (**2.128**) from methyl ester (**2.121**) and ethyl ester (**2.126**), respectively (**Scheme 14** and **Scheme 15**). The reaction profile for the cyclisation of (**2.126**) by LCMS showed there was formation of a small amount of the carboxylic acid product (**2.100**) (peak (d) - **Figure 75**). The side product by LCMS had mass ions indicative of hydrolysed starting material (**2.128**) (peak (e) - **Figure 75**). Despite expecting the ethyl ester to be more stable than the methyl ester under the copper-mediated cyclisation conditions, no benzoxazinone with the ester intact was observed. The majority of the material in the reaction was starting material (**2.126**) (peak (f) - **Figure 75**), as previously observed when attempting the reaction with methyl ester (**2.121**). However, for the ethyl ester analogue, upon prolonged heating of 4.5 h at 130 °C in the microwave, 70% conversion to the desired carboxylic acid was observed (peak (d) - **Figure 75**). Carboxylic acid (**2.100**) was isolated in 16% yield which was a lower yield than expected from the LCMS conversion in the reaction. Concerns were raised over the stability of the 5,6-analogue (**2.100**) and whether decomposition was occurring during purification, as the LCMS profile of the crude did not represent the percentage yield obtained (**Figure 75**). Despite the low yield from the copper-mediated reaction, it provided a suitable route to other quaternary centre analogues.

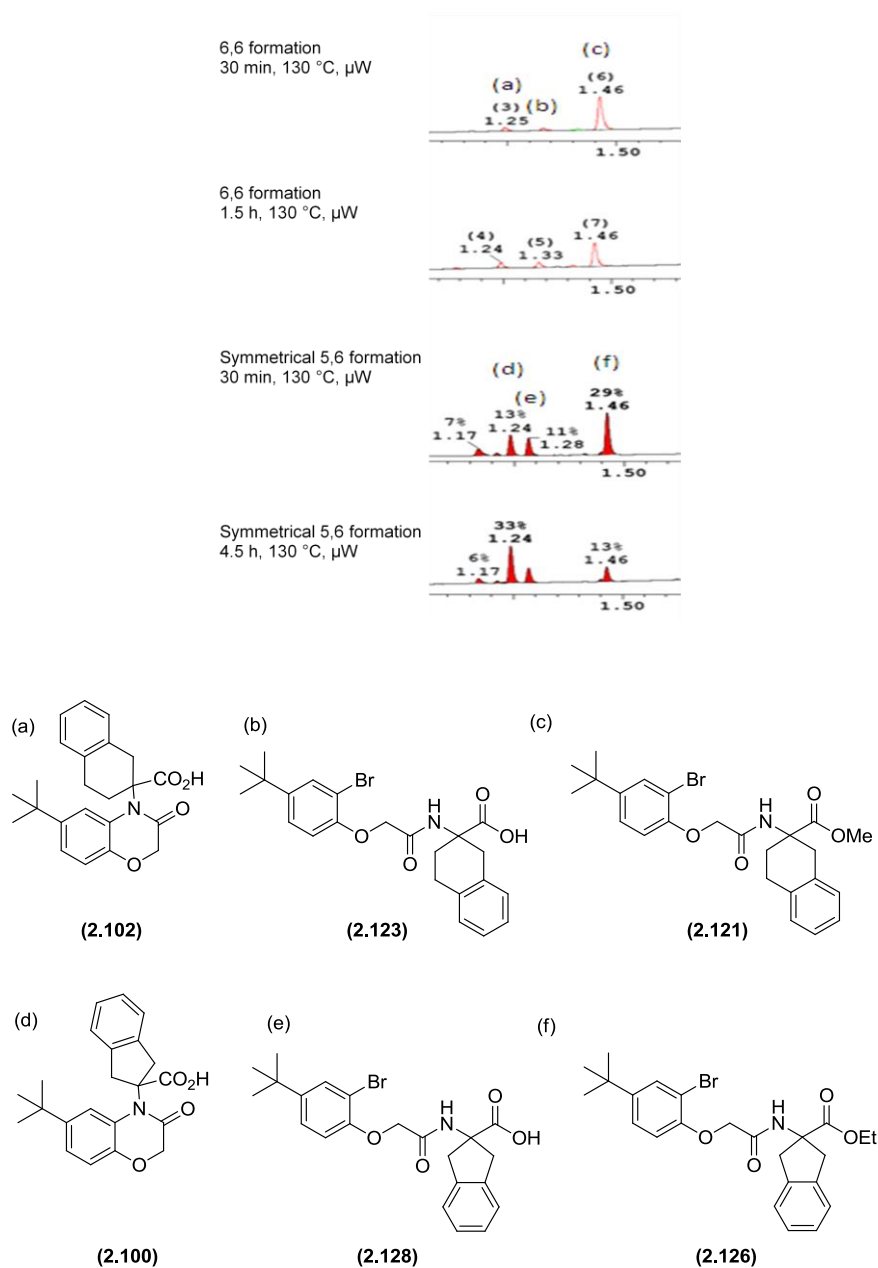
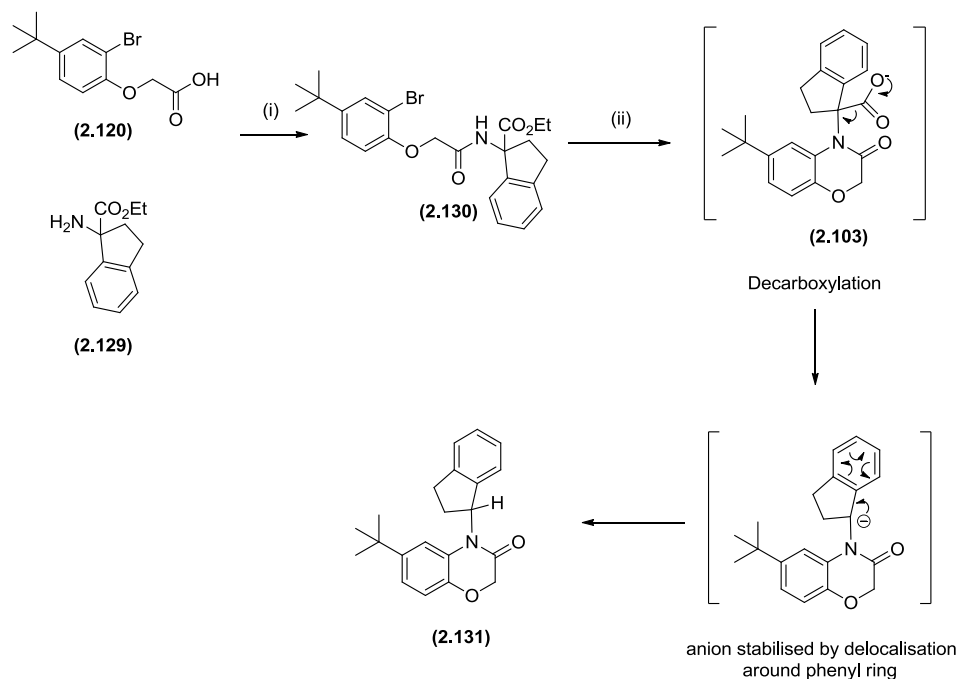


Figure 75 - LCMS profiles for copper-mediated cyclisation reactions to form 6,6-**(2.102)** and 5,6- **(2.100)** systems.

The LCMS reaction profiles for the methyl **(2.121)** and ethyl ester **(2.126)** copper-mediated cyclisations showed that the percentage of hydrolysed starting material, **(2.123)** or **(2.128)** did not increase significantly as the reaction proceeds (**Figure 75** – peak (b) and peak (e)). This suggests that the carboxylic acid intermediate was able

to undergo the copper cyclisation, contrary to initial belief. Investigation of the copper-mediated reaction for intermolecular introduction of amino acid derivatives to closely related aryl bromides gave moderate yields without protection of the carboxylic acid moiety.¹⁷³ Therefore, contrary to our original hypothesis, the chelation of the carboxylic acid to copper may be beneficial in driving the reaction forward and is an area worthy of further investigation.

The copper-mediated cyclisation was also used in an attempt to form the asymmetric 5,6-system (**2.103**). Amide (**2.130**) was formed *via in situ* acid chloride formation from carboxylic acid (**2.120**), before addition of amino ester (**2.129**) as previously described. The copper-mediated cyclisation to form 5,6-system (**2.103**) appeared to proceed, as consumption of (**2.130**) was observed. Purification and characterisation highlighted that the loss of the carboxylic acid moiety had occurred to afford (**2.131**). A proposed mechanism for loss of the carboxylic acid is shown in **Scheme 16**. The intramolecular cyclisation proceeded as expected to give the carboxylic acid (**2.103**), the molecular ion of which was observed as a minor product in the crude reaction mixture by LCMS. The decarboxylation of the acid is driven by the stabilisation of the resultant anion under basic conditions, through delocalisation of the negative charge to the phenyl group. As this issue could not be overcome, efforts were discontinued to synthesise the asymmetric 5,6-system (**2.103**).



Scheme 16 - Copper-mediated intramolecular cyclisation for the attempted formation of (2.103). Reagents and conditions: (i) (1-chloro-2-methyl-1-propen-1-yl)dimethylamine, DIPEA, THF, 62 h at 20 °C followed by 5 h at 45 °C, 23%; (ii) Cs₂CO₃, DBU, CuI (20 mol %), DMSO, 130 °C, 8 h, 23%.

Awareness of the instability of (2.103) raised concerns over the stability of the benzocyclobutane analogue (2.101), as the carboxylic acid moiety is α to the phenyl group (as in (2.103)). From docking, benzocyclobutane (2.101) demonstrated best one of the best overlaps with the B-ring of quinoline (2.35). This made the attempted synthesis a high priority and, therefore, synthesis and the resultant stability of the product were investigated.

The route for the synthesis of the benzocyclobutane (2.101) was redesigned as the required benzocyclobutane amino acid (2.133) or amino ester (2.132) intermediates (Figure 76) were not known. There was no literature precedence for Strecker chemistry on the corresponding ketone (2.134) and attempts to form amino acid monomer (2.133) using Strecker chemistry *via* the hydantoin were unsuccessful.¹⁷²

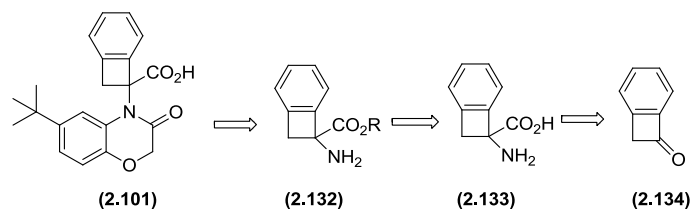
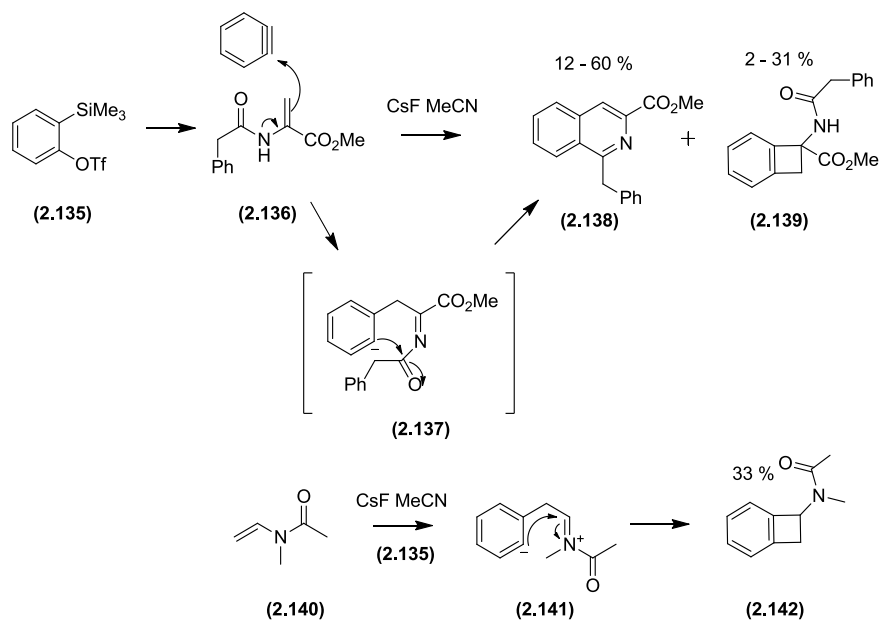


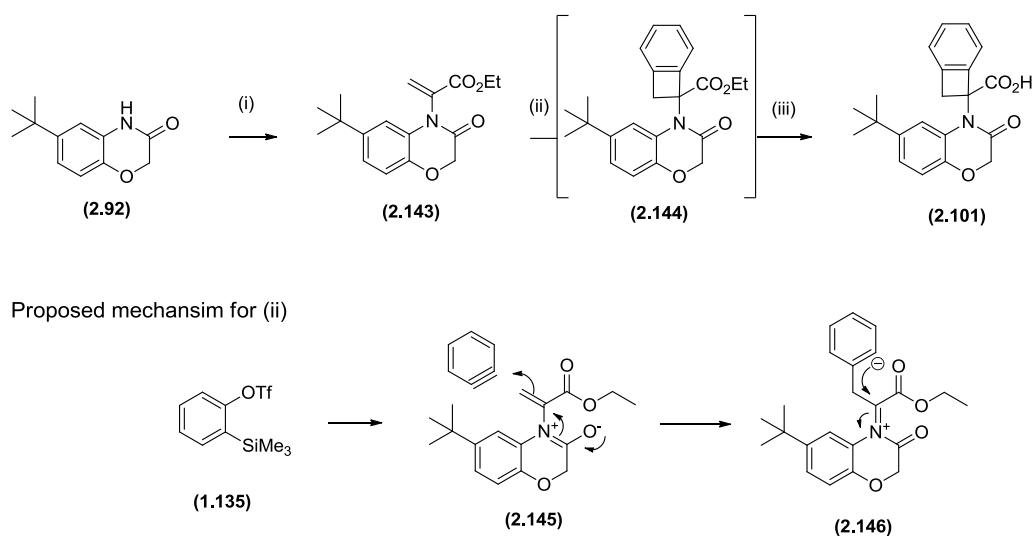
Figure 76 – Retrosynthesis of **(2.101)** showing monomers required for synthesis *via* copper catalysed intramolecular cyclisation.

2.16 Benzyne Chemistry to Synthesise Benzocyclobutane Fragment (2.101)

It has been reported by Blackburn and Ramtohl that benzocyclobutane **(2.139)** is formed as a by-product (2-31%) in the synthesis of isoquinoline-3-carboxylates *via* reaction of 2-amidoacrylate esters with arynes (**Scheme 17**).¹⁷⁴ Deprotonation of the amide proton of **(2.136)** occurs in the presence of CsF. Intermediate **(2.137)**, formed from attack of the deprotonated intermediate on benzyne, undergoes intramolecular attack of the amide carbonyl by the resultant negative charge on the phenyl ring. A hemiacetal intermediate is formed before dehydration occurs giving isoquinoline **(2.138)** as the major product. The quaternary carbon **(2.139)** is formed as a by-product by attack of the anion in intermediate **(2.137)** on the iminium carbon. Methylation of the amide nitrogen (as in compound **(2.140)**) results in nucleophilic attack onto the iminium ion of **(2.141)** being favoured, rather than at the amide carbonyl, to form the benzocyclobutane by-product. Application of this reaction was used to investigate synthesis of the desired benzocyclobutane quaternary carbon analogue **(2.101)** (**Scheme 18**).



Scheme 17 – The synthesis of isoquinoline-3-carboxylates *via* reaction of 2-amidoacrylate esters with aryne.¹⁷⁴



Scheme 18 - Formation of benzocyclobutane (2.101) and proposed mechanism.

Reagents and conditions – (i) ethyl propiolate, PPh₃, DCM, 0 °C - 20 °C, 65%; (ii) 2-(trimethylsilyl)phenyl trifluoromethanesulfonate, cesium fluoride, MeCN, 20 °C, 68 h; (iii) THF, EtOH, 2.0 N NaOH, 20 °C, 94 h, 3% over steps (ii) and (iii).

Alkene intermediate (**2.143**) was formed in good yield (65%) from reaction of the benzoxazinone core and ethyl 2-propynoate with PPh₃.^{174,175} PPh₃ mediates the reaction by addition to the acetylene (ethyl 2-propynoate) prior to the attack of the anion of the amide of (**2.92**) on this species. Subsequent loss of PPh₃ affords the desired product (**2.143**).

The benzocyclobutane system was formed by reaction of alkene (**2.143**) with commercially available 2-(trimethylsilyl)phenyl trifluoromethanesulfonate (**2.135**) and CsF in MeCN (**Scheme 18**). It is proposed that benzyne is formed *in situ*, which undergoes subsequent nucleophilic attack by the alkene. Following this, the negative charge from the carbonyl oxygen delocalises through the conjugated system to form the iminium ion intermediate (see proposed mechanism in **Scheme 18**). The resultant anion on the phenyl ring of intermediate (**2.145**) then participates in intermolecular nucleophilic attack on the iminium carbon. Hydrolysis of the ester resulted in the final product (**2.101**), affording 6.5 mg of material for screening, a 3.4% yield over two steps for this complex process.

2.17 Biochemical and X-ray Crystallography Data Analysis for Quaternary Centre Analogues

A small increase in potency of 6,6-analogue (**2.102**) in comparison to benzyl compound (**2.90**) (pIC₅₀ = 3.8, LE = 0.20, LLE_(AT) = 0.08) was observed, with increases in both the LE and the LLE_(AT) (**Figure 77**). The 6,6-compound (**2.102**) demonstrated a small increase in potency compared to methylene linked carboxylic acid (**2.57**) (pIC₅₀ = 4.4, LE = 0.32, LLE_(AT) = 0.23) but a reduction in the ligand efficiencies was observed suggesting no profitable improvement in the lipophilic shelf interaction. The symmetric 5,6-compound (**2.100**) showed comparable potency and ligand efficiencies to 6,6-species (**2.102**). Docking had highlighted this 5,6-molecule (**2.100**) as having least overlap with the lipophilic shelf where the B-ring from quinoline (**2.35**) is positioned (**Figure 78**). However, X-ray crystallography of 5,6-analogue (**2.100**) showed the shift of the 5,6-group away from the lipophilic shelf, was more pronounced than predicted by docking. Consequently, it is

unsurprising that an increase in potency was not observed, as the lipophilic 5,6-moiety is more directed out into solvent than predicted (**Figure 78**).

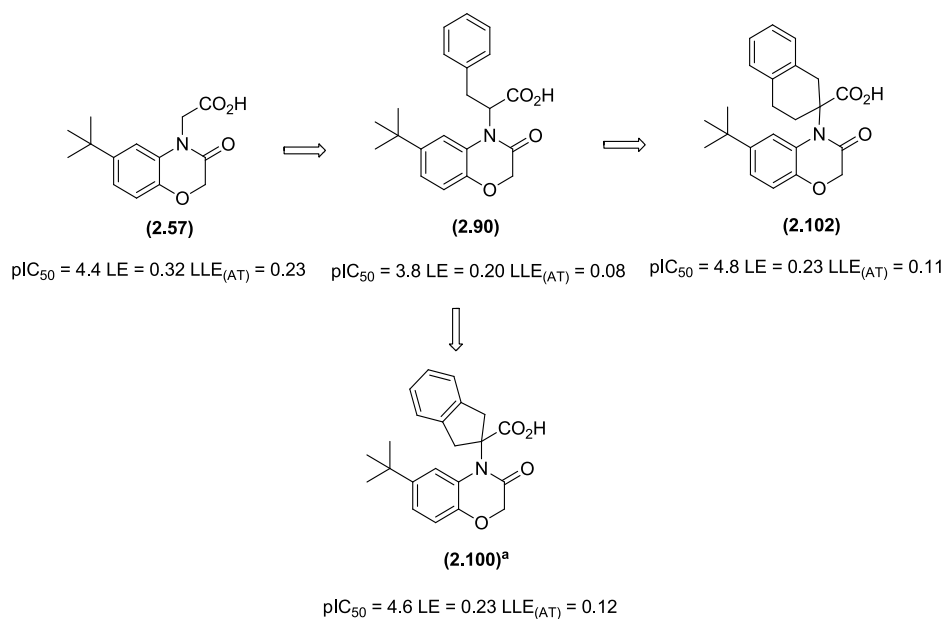


Figure 77 – Data for analogues made to rigidify benzyl group in (2.90) to stabilise the correct orientation to occupy the lipophilic shelf. ^aThis compound has only been reported active in 3 out of 4 screening occasions – ($pIC_{50} = < 3.3, 4.57, 4.58, 4.77$).

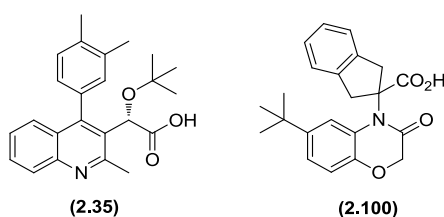
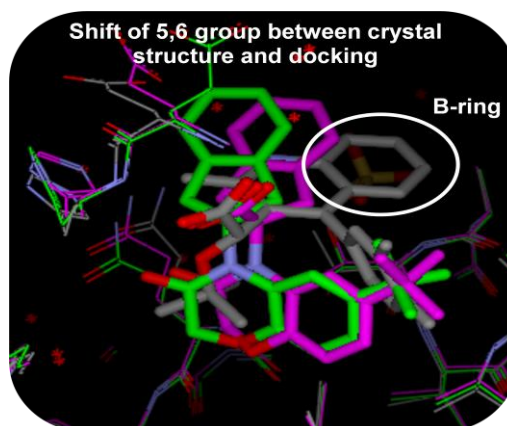


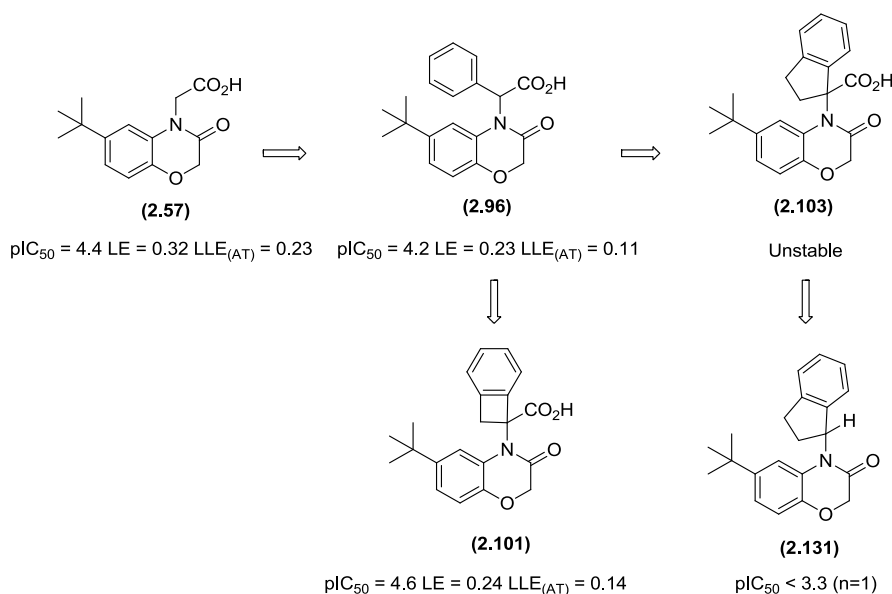
Figure 78 - Docking of 5,6-compound **(2.100)** overlaid with crystal structures of 5,6-compound **(2.100)** (PDB code: 8JXSR) and quinoline **(2.35)** (PDB code: 1OVDR).

To help rationalise the measured biochemical potency, quantum mechanics calculations were run for each of the quaternary carbon analogues **(2.100)**, **(2.101)**, **(2.102)** and **(2.103)**. The calculations showed that the 6,6-analogue **(2.102)** was not the preferred quaternary carbon molecule,¹⁰⁰ and that the 4,6 **(2.101)** and 5,6-asymmetric **(2.103)** analogues have a lower energy penalty (~ 2 kcal) to occupy the desired binding orientation compared to the 6,6-species **(2.102)**. A 1.4 kcal reduction in strain energy should theoretically result in approximately ten times difference in binding,¹⁷⁶ providing the molecule has good shape and electrostatic complementarity with the pocket.

Reducing the strain energy of the desired phenyl group orientation in **(2.96)**, by formation of the asymmetric 5,6-compound **(2.103)**, was unsuccessful due to instability of the compound (**Scheme 16**). The desired asymmetric 5,6-compound **(2.103)** suffered from decarboxylation to afford 5,6-species **(2.131)**, which was unsurprisingly inactive in the biochemical assay ($pIC_{50} < 3.3$), due to the absence of

the anchoring H-bond interactions from the carboxylic acid with Glu170 and His171, as predicted by prior knowledge from crystallographic data (**Figure 79**).

Benzocyclobutane (**2.101**) ($pIC_{50} = 4.6$, $LE = 0.24$, $LLE_{(AT)} = 0.14$) (**Figure 79**) was comparable in activity to phenyl (**2.96**) and *t*-butyl analogue (**2.57**). The benzocyclobutane analogue (**2.101**) was predicted to have good overlap with the B-ring from docking and, therefore, have less internal strain energy than the 6,6-species (**2.102**) ($pIC_{50} = 4.8$, $LE = 0.23$, $LLE_{(AT)} = 0.11$) (**Figure 78**). Despite this, the potencies and efficiencies of the two molecules are comparable. Unfortunately, X-ray crystallography for both the benzocyclobutane (**2.101**) and 6,6 (**2.102**) analogues was unsuccessful, meaning the docking poses could not be verified.



Compounds submitted for cellular assays

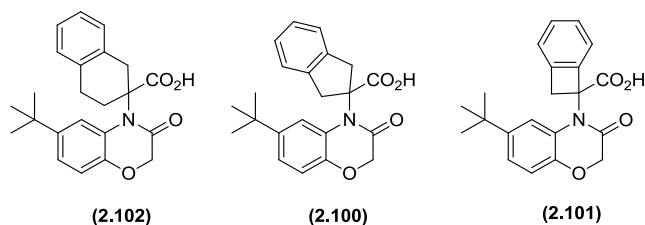


Figure 79 – Top: Data for analogues made to rigidify the phenyl group in (**2.96**) to fix in the correct plane to occupy the lipophilic shelf. **Bottom:** Compounds submitted for cellular assays.

CONFIDENTIAL – Property of GSK – Do Not Copy

Disappointingly, the 5,6 asymmetric and benzocyclobutane systems displayed most promise from docking studies, but were found to be chemically unstable due to decarboxylation. Despite further investigation into altering the electronics of the system, decarboxylation still occurred. For the sake of completeness, all stable quaternary carbon analogues with sufficient material were submitted to the appropriate cellular assays. As expected from previous experience within the programme, suggesting that biochemical pIC_{50} of above 5.0 was required to observe cellular potency, all three analogues were inactive. Without increased biochemical potency and demonstrated cellular activity, these fragments were of no further interest. Therefore, despite this research identifying novel biochemically validated inhibitors of the HIV IN-LEDGF interaction, changes in programme direction and resource constraints prevented further investigation of this template.

2.18 Assessing the Benefits of the FBDD Approach for Identification of Inhibitors with Improved Physicochemical Properties

The aim of this research project was to identify novel inhibitors of the HIV IN_{CCD}-LEDGF protein-protein interaction in improved property space compared to reported inhibitors (**Table 4**). It was hypothesised that a disciplined FBDD approach could be used to identify ligand efficient fragment starting points which could be optimised into drug candidates.

Cmpd No.	Patent No.	cPFI	MW
2.28	WO2012033735 ¹²⁶ WO2013025584 ¹²⁵ WO2013134142 ¹²⁷	10.7	634
2.25	WO2013103724 ¹²⁸	9.4	627
2.30	WO2012102985 ¹²⁹	8.7	548
2.27	WO2013159064 ¹³⁰	8.5	594
2.24	WO2012003497 ¹³¹	8.5	456
2.26	WO2013103738 ¹³²	8.1	481
2.29	WO2013123148 ¹³³	7.8	579
2.22	WO2012065963 ¹³⁴	7.7	373
2.12	WO2010130842 ¹³⁵	7.7	353
2.17	WO2007131350 ¹³⁶	6.8	428
2.20	WO2009062308 ¹³⁷	6.7	402
2.23	WO2012003498 ¹³¹	6.6	384
2.19	WO2009062288 ¹³⁸	6.4	446
2.16	WO2012138669 ¹³⁹ WO2012138670 ¹²⁴	6.4	442
2.18	WO2009062285 ¹⁴⁰	6.3	384
2.21	WO2011076765 ¹⁴¹	6.3	381
2.32	WO2012102985 ¹²⁹ WO2013002357 ¹⁴²	6.3	393
2.31	WO2012102985 ¹²⁹ WO2013002357 ¹⁴³	6.2	454

Table 4 – MW and cPFI for patented HIV IN_{CCD}-LEDGF Inhibitors shown in **Figure 32** (Section 2.7).

At the beginning of this research programme, to the best of our knowledge, no fragment based programmes had been published specifically targeting the HIV IN_{CCD}-LEDGF protein-protein interaction. At the conclusion of this research project in 2013, Debyser and co-workers published details around the use of the FBDD technique to identify and optimise fragment binders for inhibition of the HIV IN_{CCD}-LEDGF interaction.¹¹¹ The ‘privileged’ 8-hydroxyquinoline fragment (QA) was docked and optimised using a pharmacophore virtual screening approach, resulting in compounds with measurable antiviral activity *in vitro* being reported (Q2-8) (**Figure 80**).

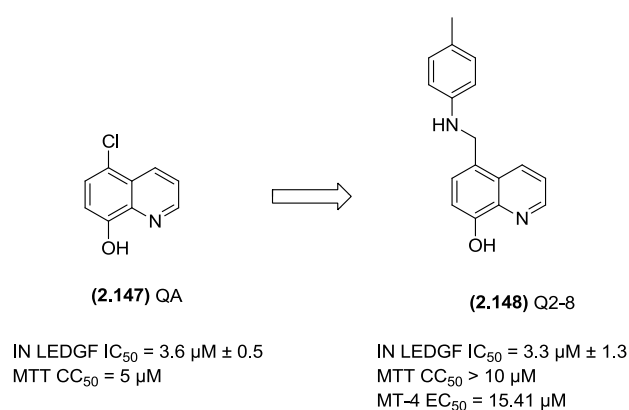
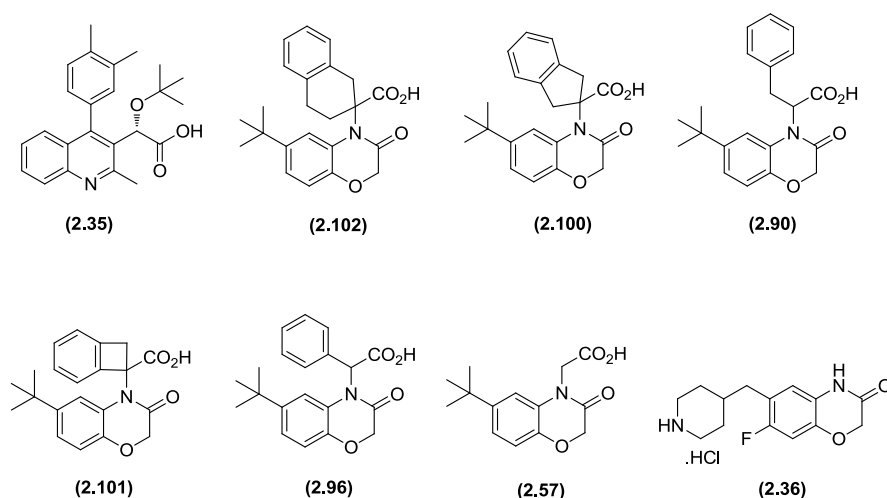


Figure 80 – 8-Hydroxyquinoline fragment (QA) developed into Q2-8 which shows measurable anti-viral activity *in vitro*.

The studies from Debyser and co-workers, combined with the fragments identified as part of this research project, gives confidence that FBDD can be used to identify efficient small molecule binders, even for challenging protein-protein interaction targets. Further optimisation of these hits could provide valuable tools to help validate the mechanism of action as a therapeutically relevant approach, provided that careful control of the physicochemical properties is practiced, which can be predicted using cPFI. Comparison of the cPFI for patented inhibitors (**Table 4**) and the FBDD derived analogues from Debyser and this research project (**Table 5**), clearly demonstrates that not all FBDD derived molecules have an improved physicochemical profile. Debyser’s 8-hydroxyquinoline fragment (QA) displayed improved cPFI (5.1) (**Table 5**) compared to the patented analogues (**Table 4**). Despite this, once optimised into Q2-8, the physicochemical advantage was lost (Q2-

CONFIDENTIAL – Property of GSK – Do Not Copy

8 cPFI = 8.2) (**Table 5**). This demonstrates the importance of careful control of physicochemical properties during lead optimisation.



Cmpd No.	cChrom LogD	cPFI	MW
2.148 (Q2-8)	5.2	8.2	264
2.35	3.4	6.4	377
2.147 (QA)	3.1	5.1	179
2.102	3.1	5.1	379
2.100	2.7	4.7	365
2.90	3.6	5.6	353
2.101	2.4	4.4	351
2.96	3.1	5.1	339
2.57	1.8	2.8	263
2.36 (FBDD hit)	0.9	1.9	264

Table 5 – MW and cPFI for the FBDD hit and the molecules prepared during the exploration for HIV IN_{CCD}–LEDGF Inhibitors **Figure 32**.

Encouragingly, the benzoxazinone template identified as part of this research project, had a significant chrom Log D and cPFI advantage over both the FBDD derived analogues from Debyser, and the patented inhibitors. Publication of the benzoxazinone template, when intellectual property allows, will provide a good

quality starting point from which tool molecules or drug candidates with desirable property profiles (cPFI < 7) could be identified to help fight the relentless battle against the global disease AIDS.

Identifying molecules in optimal physicochemical property space is not only a challenge for this target, but applies across the entire medicinal chemistry community. Encouragingly, this work has demonstrated that a disciplined FBDD approach aids in this endeavour, and can add value over other hit finding methodologies (e.g. HTS and ELT) providing the techniques is given the appropriate level of time and resource. To attempt to determine the most efficient FBDD approach, a retrospective analysis of the FBDD technique for HIV IN_{CCD}-LEDGF was carried out (**Chapter 3**).

3 *Retrospective Analysis of the HIV IN_{CCD} – LEDGF FBDD Programme: Identification of the Most Efficient and Successful Method of Fragment Screening and Hit Optimisation*

The HIV IN_{CCD}-LEDGF interaction was the one of the first programmes to be explored using FBDD methodology in our laboratories. This heightens the importance of assessing the approach and identifying successful strategies to use in future FBDD projects. Therefore, as part of this research, a thorough retrospective review of the programme was undertaken.

Throughout the HIV IN_{CCD}-LEDGF programme, lessons have been learnt and insight gained into the most efficient FBDD strategies. The retrospective review of the HIV IN_{CCD}-LEDGF FBDD approach examined three areas:

1. how to assess target tractability of FBDD targets early in the programme;
2. determination of the optimal screening triage; and
3. how to select and optimise the fragments efficiently.

These three areas are discussed in turn and the authors recommendations for subsequent FBDD efforts are highlighted. This insight will now be implemented within future fragment programmes, globally across our laboratories, publicising the potential to identify high quality lead molecules using FBDD. When the timing is appropriate, from an intellectual property perspective, this work will be presented at external conferences and published, to inform and influence the wider fragment community.

3.1 Assessing Target Tractability for FBDD

3.1.1 Surveying the Literature

HIV IN_{CCD}-LEDGF was considered tractable due to literature evidence of inhibition of the protein-protein interaction by small molecules with clear SAR.^{91,93,96,99} At the

outset of the programme, literature described inhibition of the interaction by a number of structurally distinct chemotypes. However, as the FBDD effort progressed new literature became limited to the majority of potent inhibitors having a common pharmacophore related to quinoline (**2.35**) (**Figure 81**). If at the outset of the FBDD effort all reported analogues were of a single chemotype, this would have raised concerns over target tractability. From the literature it was also apparent that a number of academic groups and pharmaceutical companies had already independently explored inhibition of HIV IN_{CCD}-LEDGF, without reporting identification of novel templates.^{93,102,108,119,177} This highlighted the significant challenge of inhibiting this protein-protein interaction.

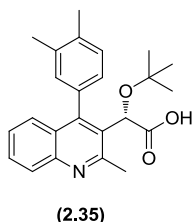


Figure 81 - Quinoline (2.35).

3.1.2 Features of the Binding Pocket and Known Inhibitors

X-ray crystallography confirmed that quinoline (**2.35**) bound to HIV IN_{CCD}. All quinoline analogues that inhibited the HIV IN_{CCD}-LEDGF interaction were lipophilic compounds with pendant carboxylic acids, which carried the potential metabolic problem of glucuronidation. Apart from the H-bond interaction between HIV IN_{CCD} protein backbone and the carboxylic acid moiety, all other interactions in the HIV IN_{CCD} pocket reported were lipophilic. The lipophilic nature of the pocket and requirement for the carboxylic acid, reduced the likelihood of identifying small polar fragments from the fragment screen. Identification of fragments with low lipophilicity and efficient binding interactions to the protein is critical for the success of an FBDD project.

For future fragment programmes the literature should be surveyed for screening attempts against the target of interest. Known inhibitors should be identified and the

reported SAR considered. Evidence of a low hit rate and the necessity for undesired molecular characteristics for binding (i.e. lipophilicity), should result in questioning the appropriateness of the target for FBDD. However, a fragment screen to assess target tractability may be appropriate for less explored targets.

From the outset of the FBDD effort it was apparent that targeting the inhibition of the HIV IN_{CCD}-LEDGF interaction would not proceed *via* a classical fragment approach. HIV IN_{CCD} did not have a well-defined binding pocket with conserved interactions as seen in more tractable targets such as kinases. Despite inhibition of the protein-protein interaction being considered less tractable, the FBDD approach was still pursued in order to identify a novel binding motif and new interactions within the HIV IN_{CCD} pocket. Novel interactions identified would not only be used to develop a structurally distinct new series, but could also be incorporated into the quinoline series (2.35).

As with HIV IN_{CCD}-LEDGF interaction, targets considered less tractable are often of highest value to our and other global laboratories. Due to this, even if the ‘hit rate’ was predicted to be low, the expense of screening the entire FBDD library with the risk of not identifying any developable hits, is outweighed by the high value of identifying an novel fragment hit. The caveat is that the high concentration biochemical, biophysical and X-ray crystallography systems must be high throughput and robust, as optimisation of the fragment hits will likely begin from weak mM hits, especially for protein-protein interactions.

3.2 Determination of the Optimal Screening Triage

Without robust biochemical and biophysical assays in place, as was the case for HIV IN_{CCD}-LEDGF, hits from an FBDD screen cannot be followed up in a data-driven manner. This was a significant challenge for follow up of the fragment hits identified from the HIV IN_{CCD}-LEDGF screen. The majority of fragments binding to HIV IN_{CCD} were below the limit of the biochemical assay and development of a routine high throughput biophysical assay had not been successful. For all future

fragment targets, especially less tractable targets, high importance must be placed on optimisation of the assay techniques before the FBDD screening begins. If robust biochemical, biophysical and X-ray crystallography systems are not developable, the target is unlikely to be suitable for FBDD due to problems in confirming and following up the weak hits from the FBDD screen.

3.2.1 Biophysical Screening

For the HIV IN_{CCD}-LEDGF target, development of an SPR assay failed due to problems attaching the HIV IN_{CCD} protein to the SPR plate. SPR was the preferred biophysical assay because it could have provided measures of K_d , on and off rates and binding stoichiometry. SPR is also a high throughput method which can be used routinely during fragment optimisation, as well as for initial screening of the FBDD library. Despite the wealth of binding information provided by ITC, this assay was not suitable for screening the entire FBDD1 library due to the high protein consumption required. As an alternative approach, competitive NMR STD experiments were developed with quinoline (**2.35**). The NMR STD experiments were used to triage the fragment hits to crystallography. Unfortunately, the NMR assay could not be resourced routinely to help guide the fragment optimisation process. Therefore, this fragment programme was run without a routine biophysical assay. Through the HIV IN_{CCD}-LEDGF effort it has been shown that the lack of biophysical information makes it challenging to rank weak biochemical hits. A routine biophysical assay is highly recommended for all future fragment efforts, which is a conclusion that parallels with literature consensus.^{3,20}

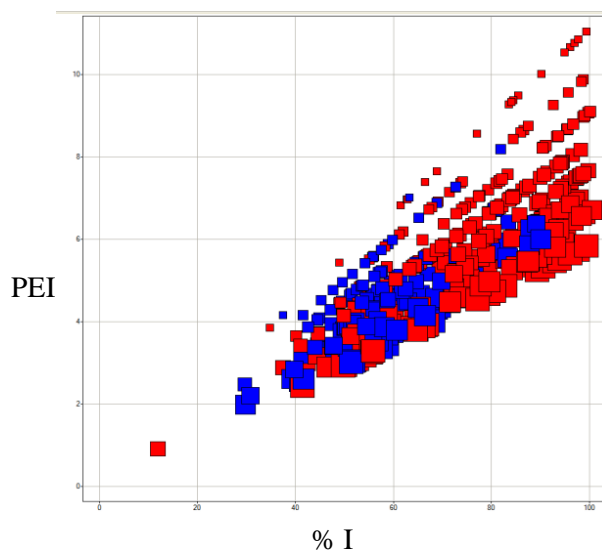
3.2.2 Biochemical Screening

Prior to screening FBDD1 for inhibition of the HIV IN_{CCD} – LEDGF interaction, the biochemical assay was re-configured to allow high concentration screening. After optimisation, a top concentration of 500 μ M was used from 10 mM DMSO stock solutions to triage compounds to NMR.

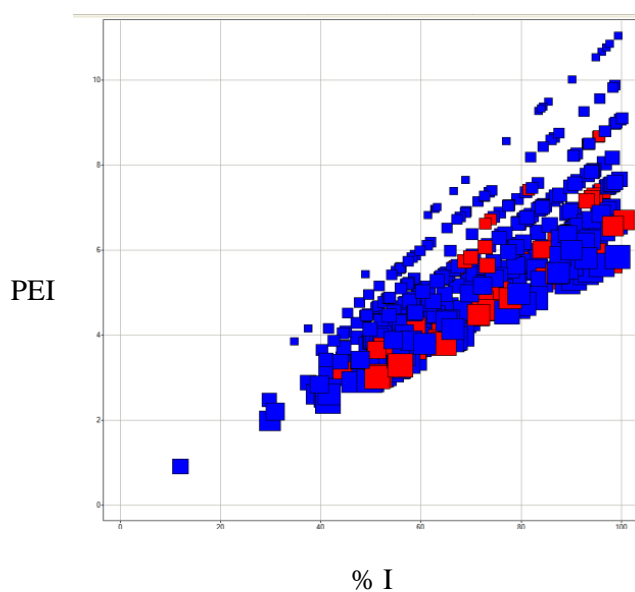
CONFIDENTIAL – Property of GSK – Do Not Copy

For HIV IN_{CCD}-LEDGF, single concentration screening was carried out in duplicate. Compounds with at least one replicate giving %I \geq 40 were progressed to full-curve analysis (all compounds in **Graph 4**). Compounds with activity in full curve analysis are displayed in **Graph 4** in red. The %I \geq 40 proved appropriate as compounds across the entire %I range confirmed in full curve analysis were hits in NMR analysis (red compounds in **Graph 5**).

However, the three fragment hits confirmed in X-ray crystallography from FBDD1 all had %I > 60. Despite this, future fragment efforts are advised to take the %I cut-off down to at least %I > 40 to avoid missing weak but efficient binders.



Graph 4 – All compounds selected for full curve from single shot analysis. Response (%I) vs. PEI (%I / HAC). Size of points increases with HAC. Active compounds in full curve are shown in red.



Graph 5 – All compounds selected for full curve analysis from single shot analysis. Response (%I) vs. PEI (%I / HAC). Size of points increases with HAC. Compounds selected for X-ray from NMR STD biophysical study are shown in red.

3.2.3 Triage to X-ray Crystallography

Analysis of the fragment triage for inhibition of the HIV IN_{CCD}-LEDGF interaction, showed the optimal triage to crystallography would have been to use NMR competition studies with quinoline (**2.35**). The compounds showing greatest percentage displacement by quinoline (**2.35**) gave the highest translation to liganded X-ray crystal structures (**Table 6**). All compounds which were screened through the NMR displacement assay and gave crystal structures in HIV IN_{CCD} showed > 75% displacement in NMR by quinoline (**2.35**). This knowledge, when selecting compounds for crystallography, would have reduced the significant resource and cost requirements as fewer compounds would have been progressed. Understanding if this filter is applicable to other targets could help reduce the resource and time implications for future projects when progressing fragments to X-ray which are unlikely to bind. Currently there is insufficient data to support using this as a general filter across other targets. However, this could be considered as more FBDD projects are embarked upon.

Some fragments displaced by quinoline (**2.35**) by > 75% gave crystal structures in an allosteric pocket (second site). It is suspected that the compound must be able to bind in both pockets despite only crystallising in the allosteric site. Chemistry was not performed to optimise the allosteric site binders as it was difficult to assess the biological relevance of interactions within this pocket.

Displacement by (2.35)	No. of compounds	X-ray HIV IN _{CCD} -LEDGF Pocket	X-ray pocket 2	Ambiguous
Yes (> 75%)	16-17	3	4	1
Partial (< 50-75%)	12	0	3	0
Maybe (ambiguous)	12-13	0	2	1

Table 6 - Percentage displacement of fragments by quinoline (2.35) in NMR studies correlated with success in X-ray crystallography at the HIV IN_{CCD} – LEDGF and the second site.

3.2.4 Identification of Biochemically Inactive Efficient Hits

The pre-quality controlled fragment library of 3500 compounds was triaged using the biochemical screen, prior to the biophysical NMR screen. This was necessary as the NMR-based biophysical assay did not have sufficient throughput to screen the entire set of 3500 compounds. Weak HIV IN_{CCD} binders, with activity below the limit of the biochemical assay, were not identified.

After completion of the fragment screen a sub-set of the FBDD1 library, consisting of 1056 compounds, was cocktailed for NMR study. These cocktails of compounds were screened through the NMR assay aiming to identify weak binders which were eliminated due to the biochemical cut-off of 40% inhibition in the original FBDD1 screen. Forty-two fragments were progressed to crystallography after NMR displacement studies with quinoline (2.35). Of the forty-two fragments, eleven compounds were crystallised with HIV IN_{CCD}, with three compounds occupying the desired pocket. Screening of the FBDD1 sub-set was carried out too late to impact on cluster 2 optimisation. However, the NMR screen of the FBDD1 sub-set has emphasised the importance of orthogonal biochemical and biophysical screening. The FBDD research on HIV IN_{CCD}-LEDGF has shown that, wherever possible, orthogonal biochemical and biophysical screening should be carried out for future fragment programmes. This maximises the chance of identification of small, weak, highly efficient binders which could give crystal structures and minimises the risk of

missing weak hits. These validated hits provide valuable fragment starting points and / or information around novel interactions in the pocket.

3.2.5 X-ray Crystallography

The low success rate in crystallography for HIV IN_{CCD} resulted in huge cost and resource implications which would not be sustainable for future fragment efforts. As previously described, the number of compounds submitted to crystallography from the fragment screen could have been reduced through improved selection criteria. In this regard, only compounds which demonstrated > 75% displacement by quinoline (2.35) in NMR, gave crystal structures in HIV IN_{CCD}. Given that NMR is a more rapid and less expensive technique in comparison to X-ray crystallography, establishing a %I filter early for a biochemical target, on a small set of compounds, has been highlighted as a suitable method of triaging large sets of compounds to crystallography for future fragment efforts. The triage reduces the number of fragments attempted in crystallography and reduces the resource and cost burden.

3.3 Fragment Optimisation Approach

X-ray crystallography provides valuable structural information about the binding location, orientation, and specific interactions for a fragment. However, identification of a fragment from a well-constructed screening cascade is just the beginning of the FBDD approach. Without careful growth and optimisation, an FBDD approach provides no advantage over alternative hit finding methodologies. Therefore, three fragment optimisation exercises are discussed, with conclusions and recommendations for future fragment efforts.

3.3.1 Fragmentation of Known Inhibitors

Fragmentation of the quinoline racemate (2.32) (Figure 82) identified that the entire molecule is required to achieve sub-micromolar biochemical potency. The smaller fragments were no more potent and less efficient than the *t*-butyl benzoxazinone (2.57), endorsing (2.57) as a good start point.

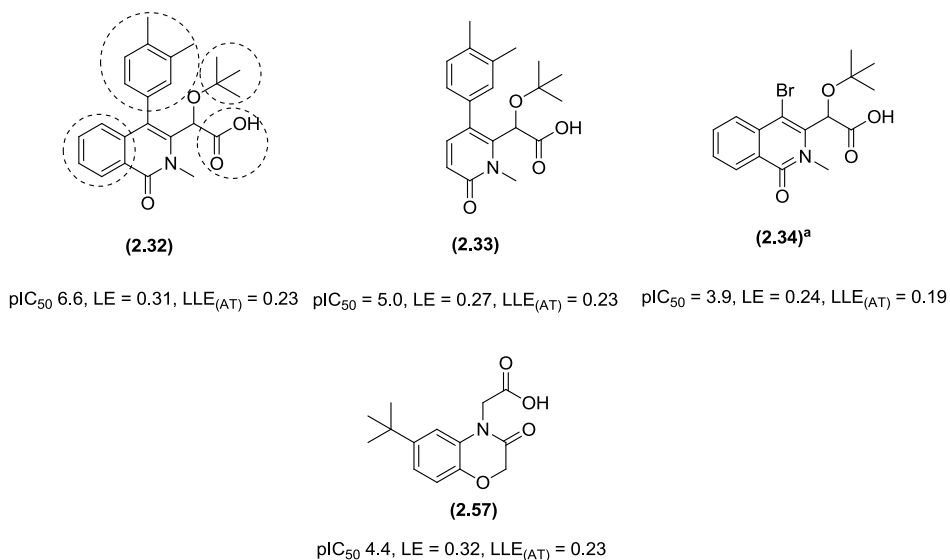


Figure 82 – Fragmented analogues of quinolinone (**2.32**) compared to benzoxazinone (**2.57**). ^aThis compound has only been reported active in 2 out of 3 screening occasions – 3rd screening occasion measured < 4.3.

Fragmentation of the quinoline analogue (**2.32**) highlighted that interactions with multiple remote areas of the pocket was required to achieve potency. This mirrors the observation that growth of benzoxazinone in a single vector/direction, failed to give the desired activity. Growth in multiple vectors contradicts the usual FBDD approach where small growth in a single vector is usually made, as was the route pursued for HIV IN_{CCD}-LEDGF. Growing in multiple directions would result in higher molecular weight compounds than usually targeted during a fragment optimisation process. However, following the information from the fragmentation exercise, and synthesising larger compounds occupying all of the key binding pockets, was predicted to produce more potent analogues. This demonstrates that not all targets are suitable for small growth in a single vector. Future fragment efforts are advised to deconstruct known inhibitors at the beginning of the project, to assess which interactions of the ligand with the protein are required for potency, and to fully understand the binding site.

3.3.2 Selection of Clusters for Synthetic Follow-up

Cluster 3 pyrazole (**2.39**) ($pIC_{50} = 3.7$, $LE = 0.32$, $LLE_{(AT)} = 0.39$) was the most potent hit identified from the fragment screen (**Figure 83**). Pyrazole (**2.39**) displayed good LE and $LLE_{(AT)}$ and had the lowest $cLog P$ (0.5) of the fragments identified. The pyrazole fragment also gave the highest resolution in crystallography and was initially considered the best fragment starting point.

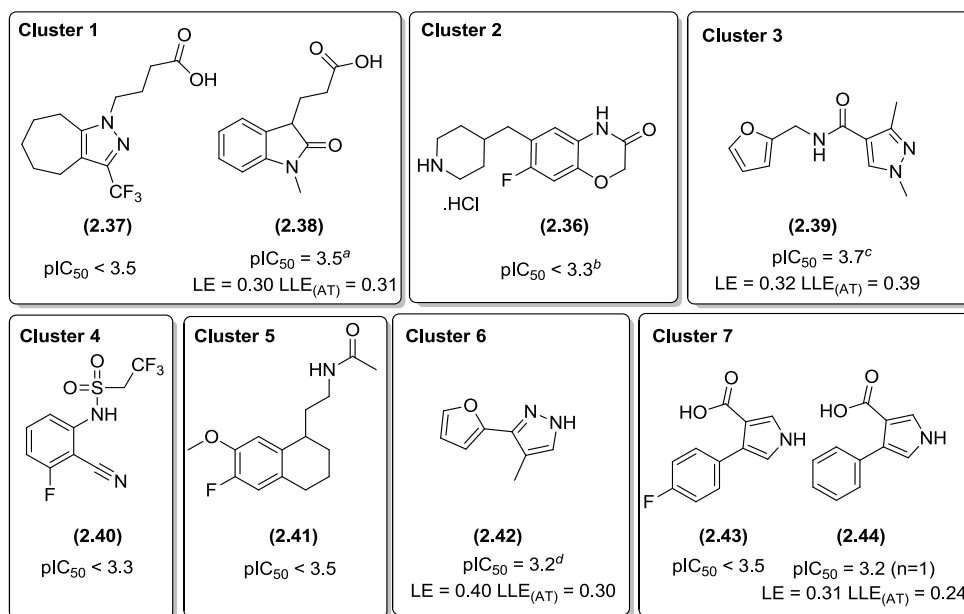


Figure 83 - Nine fragments clustered into seven fragment clusters. ^aThis compound has only been reported active in 3 of the 6 screening occasions. ^bThis compound measured < 3.52 on 3 screening occasions and < 3.3 on 1 screening occasion. ^cOnly Batch 1 is represented in the mean pIC_{50} —Batch 1 was active on 3 out of 5 screening occasions and variability was observed ($pIC_{50} = 3.35, 4.29, < 4.3, 3.33, < 3.52$). ^dThis compound has only been reported active in 2 of the 3 screening occasions ($pIC_{50} = < 4.3, 3.39, 3.06$).

Confusing SAR from SSS around cluster 3 led to resynthesis of a selection of analogues. Resynthesis of SSS compounds highlighted them as false positives (pIC_{50} 4-5), as on remake, they were all substantially less active. Structural design to improve interactions of pyrazole (**2.39**) with the HIV IN_{CCD} pocket, failed to produce active analogues.¹⁷⁸ From FBDD research on the HIV IN_{CCD} —LEDGF interaction, it has been highlighted that the most potent compounds identified from the fragment

screen may not be the best fragment starting points. To minimise false positives such as pyrazole (**2.39**), hit validation using orthogonal biochemical and biophysical screening methods is recommended as a requirement.

3.3.3 Mobility of Fragments

Optimisation of piperidine (**2.36**) in HIV IN_{CCD} resulted in significant movement of the benzoxazinone core as the molecule was grown (**Figure 84**). Multiple different binding orientations were observed, resulting in unclear SAR. This made it challenging to optimise piperidine (**2.36**) in an iterative fashion. The smaller fragments such as bromide (**2.45**) are hypothesised to be binding in the energetically preferred mode, suggesting that the new substituents added as the compounds were grown were dictating the binding mode.

A good fragment starting point is one that is strongly anchored in the pocket by interactions with the protein. The *t*-butyl analogue (**2.57**) displayed H-bond interactions with the backbone nitrogen's of Glu170 and His171 anchoring the benzoxazinone core, and was therefore selected as a fragment to grow.

For future fragment programmes careful selection and optimisation of tightly anchored fragments is recommended. It is preferable to have starting points with multiple H-bond interactions from different parts of the fragment, minimising the potential movement in the protein. If starting points with multiple hydrogen bond interactions are not identified, growth and optimisation should focus on finding additional interactions to lock the template into a conserved binding orientation. As an alternative, fragments which fit tightly in the protein pocket, with optimal shape complementarity, are also considered good starting points. The tightly anchored hit fragments would provide clear SAR significantly simplifying the optimisation process.

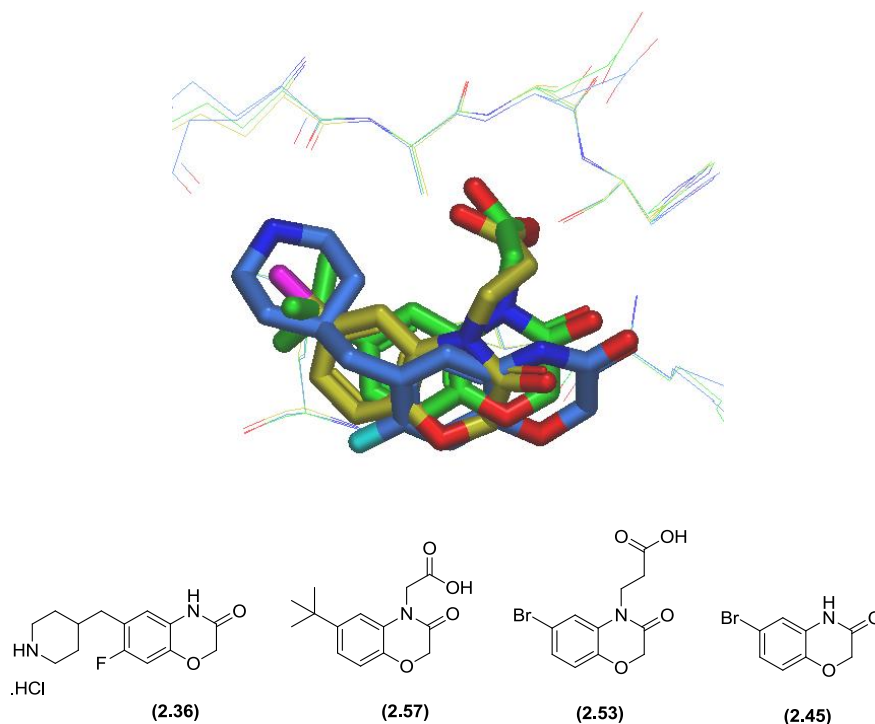


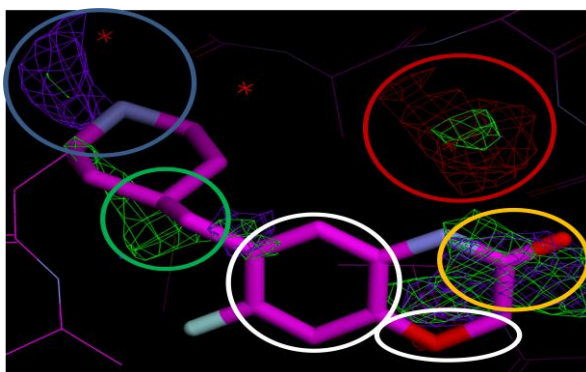
Figure 84 - Movement of benzoxazinone core of piperidine (2.36) (PDB:1LIQU), *t*-butyl compound (2.57) (PDB:1WQHZ) and bromide (2.53) (PDB:1VORG).

3.3.4 Mapping of the Protein Active Site Using GRID

The computational programme GRID was first reported by Goodford in 1985¹⁷⁹ and is used to identify energetically favoured binding sites on proteins. Despite the use of GRID being reported for lead optimisation projects,¹⁸⁰ the use of GRID to advance FBDD projects within our laboratories was only explored towards the conclusion of the HIV IN_{CCD} – LEDGF fragment effort. The information gained from GRID was not used during optimisation of piperidine (2.36) due to the time of GRID becoming available, therefore, the technique is discussed in a retrospective fashion.

The GRID programme arranges small probes, such as water, methane, and ammonia in a 3D grid in the protein pocket, to help characterise the target. The GRID software calculates the energy of the probes and moves them to low energy positions in the binding pocket of interest. The information is displayed as a GRID map which shows the optimal positions for the probes, highlighting areas where particular

interactions are favourable. This probe map can be used to help guide fragment selection and optimisation.



Key to GRID – Green = CH₄, Blue = NH₄⁺, Red = C=O

Figure 85 - GRID map for CH₄, NH₄⁺ and C=O probes overlaid with crystal structure of piperidine (**2.36**) (PDB:1LIQU) in the HIV IN_{CCD} binding pocket.

The information derived from the GRID maps for the HIV IN_{CCD} – LEDGF binding pocket is discussed below;

- The CH₄ grid map circled in green (**Figure 85**) shows an area where lipophilic groups are predicted to be energetically favourable. This is backed up by fragment binding i.e. compound (**2.57**) (**Figure 86**) where the *t*-butyl group occupies the area of the HIV IN_{CCD} pocket circled in green, which has been filled by the CH₄ probe.

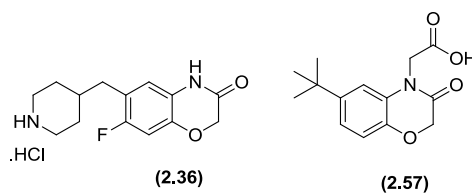


Figure 86 - *t*-Butyl (**2.36**) and piperidine (**2.57**).

- The NH₄⁺ probe is circled in blue and occupies an area of binding pocket close to the basic piperidine nitrogen of (**2.36**). The close proximity of the

NH₄⁺ probe and piperidine nitrogen suggests the basic centre is well tolerated in this region of HIV IN_{CCD}.

- The red circle highlights an area of HIV IN_{CCD} where carbonyl groups are energetically favourable. The red circled area is also where the carboxylic acid group of (**2.57**) is positioned to interact with the backbone nitrogens of Glu170 and His171. The good overlay of the carboxylic acid and the carbonyl probe suggests the position of the carboxylic acid is favourable for binding.
- The yellow circle highlights an area filled by CH₄ and NH₄⁺ probes. For cluster 2 analogues, including piperidine (**2.36**), this area of HIV IN_{CCD} is filled by the amide carbonyl of the benzoxazinone core. Replacement of the amide carbonyl with a lipophilic or basic group could improve interactions with the HIV IN_{CCD} protein.
- The white circles highlight areas where no probes occupy. This suggests that the aryl ring of the benzoxazinone core and the ether oxygen are acting as a spacer between the areas filled by probes circled in green and yellow. The GRID suggests that these spacer groups, despite being required to link the groups, are likely to have minimal contribution to the binding affinity to HIV IN_{CCD}. This information suggests that replacement or removal of these groups could have been investigated.

On termination of the HIV IN_{CCD}-LEDGF fragment effort, the fragment optimisation strategy was reviewed. Exploration of ring-opened analogues resulting in removal of one of the rings from the benzoxazinone core, were highlighted as major gaps in fragment optimisation of piperidine (**2.36**).

The retrospective analysis highlighted that the use of GRID would help guide fragment optimisation giving rapid advancement of the fragment templates. Early application is essential to characterise the protein pocket, and the resulting maps could be used to grow the fragments to pick up favourable lipophilic and polar interactions. An important strategy for further investigation is to use the GRID maps

to identify areas where the fragment can be reduced in size. This is performed by removing heavy atoms from regions that the probes do not occupy, maximising the binding efficiency of the fragment.

3.4 Summary of Successful FBDD Approaches to be used in Future Programmes

The FBDD effort was carried out in our laboratories in an attempt to identify ligand efficient, HIV IN_{CCD} – LEDGF inhibitors. Despite termination of the programme, this work provided valuable insight into successful and unsuccessful FBDD strategies, which will be used for future FBDD efforts. Four recommendations have been made:

- 1) **Target Tractability:** Tractability should be assessed from literature review. If the target is deemed poorly tractable, the value of success versus the cost of screening and fragment follow-up needs careful consideration.
- 2) **Screening:** For FBDD success, care must be taken to identify an appropriate screening cascade. It is highly recommended that orthogonal biophysical and biochemical screening is carried out in parallel for the entire screening library. This avoids missing biochemically inactive, but highly efficient binders. This supports recommendations from the literature.²⁰
- 3) **Fragment Optimisation:** FBDD only has an advantage in the pursuit of ligand efficient molecules in good physicochemical property space, providing rigorous control is considered throughout the optimisation process. Fragmentation of known binders gives clear direction on which areas of the protein pocket are likely to be beneficial for potency. It is important to understand that the best fragment starting points are not always the most potent and efficient, but are those with the optimal growth vectors and limited mobility in the protein pocket.

- 4) **Computational Techniques:** Numerous computational techniques are available to help guide the fragment optimisation process. The computational programme GRID can be used to guide fragment growth, by determining the nature of groups which would be energetically favourable in specific regions of the protein pocket. It is believed that GRID can also highlight parts of the molecule acting solely as linkers, and assist with the redesign of these regions of the molecule to improve the binding efficiency.

In summary, retrospective analysis of the HIV IN_{CCD}-LEDGF effort has resulted in recommendations which will be applied to future fragment efforts. It is predicted that careful application of these recommendations will maximise the success of such FBDD projects and minimise the timelines to identify lead molecules suitable for progression towards a drug candidate.

4 Design of Fragment-derived Inhibitors of Cluster of Differentiation 38 (CD38) in the Absence of a Liganded X-ray Crystal Structure

4.1 CD38 Receptor Function

Cluster of differentiation 38 (CD38) is an evolutionarily conserved membrane-bound multi-functional protein, which was identified in 1980 by monoclonal antibody typing of lymphocytes.¹⁸¹ Clusters of differentiation are involved in many different functions including cell-signalling pathways.^{182,183} Initial studies around CD38 focused on lymphoid tissues, but research has widened to other tissues as its expression is now considered ubiquitous.¹⁸³ It has been identified that the CD38 enzyme metabolises nicotinamide adenine dinucleotide (NAD) to generate two Ca^{2+} messengers, cyclic adenosine diphosphate ribose (cADPR)¹⁸⁴ and ADP-ribose (ADPR). Formation of the Michaelis complex¹⁸⁵ by co-ordination of the side chain oxygen from Ser193 to NAD, results in release of nicotinamide (**Figure 87**). The resultant enzyme-stabilised cation intermediate can be attacked by either the adenine nitrogen (route 1 - **Figure 87**) forming cADPR, or alternatively by water (route 2 - **Figure 87**), resulting in formation of ADPR. CD38 generates approximately one molecule of cADPR for every hundred NAD molecules metabolised, converting the remaining 99% to ADPR.¹⁸⁶

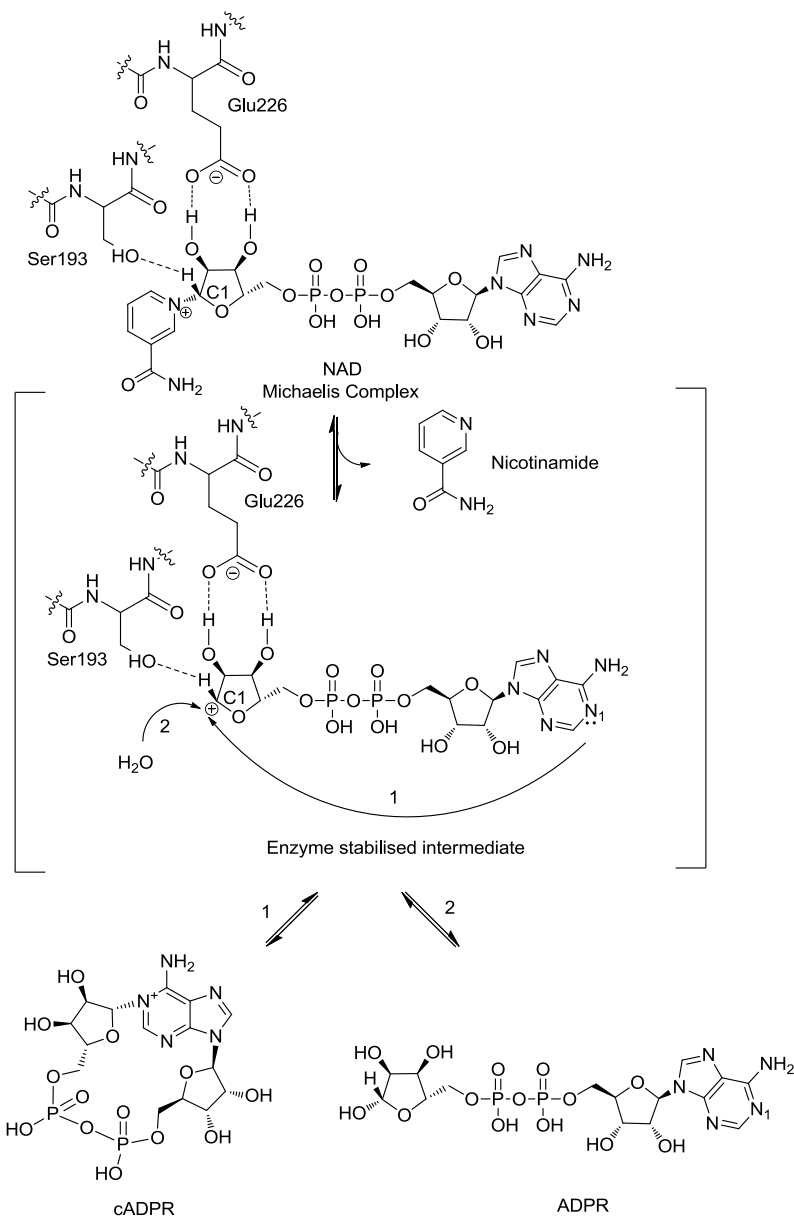


Figure 87 – CD38 promoted conversion of NAD⁺ to cADPR and ADPR.

It has been postulated that CD38 is the major regulator of intracellular NAD levels. This initially caused some confusion in the scientific community, as it is a transmembrane bound protein, with its active site in the extracellular domain.¹⁸⁷ From further studies, it has been proposed that CD38 promotes transfer of the Ca²⁺ messenger (cADPR) into cells,¹⁸⁸ or internalisation of CD38 is feasible, but no conclusive evidence has been obtained.¹⁸² The modulation of intracellular calcium levels, by formation of the second messenger cADPR, is not the only function of

CD38. In addition to this, it has been reported that CD38 may form other secondary messengers when in different pH environments (**Figure 88**). For example, NAADP was proposed to be formed through base-exchange reactions under acidic pH.¹⁸⁹⁻¹⁹¹ Despite this observation, formation and control of intracellular NAADP was not thought to be solely reliant on CD38, and therefore NAADP could not be definitively controlled by modulation of CD38 activity.¹⁸³ More recently, it has been reported that *in vivo* CD38 appears to be a NAADP degrading rather than a NAADP forming enzyme, thus contradicting the original observation.¹⁹²

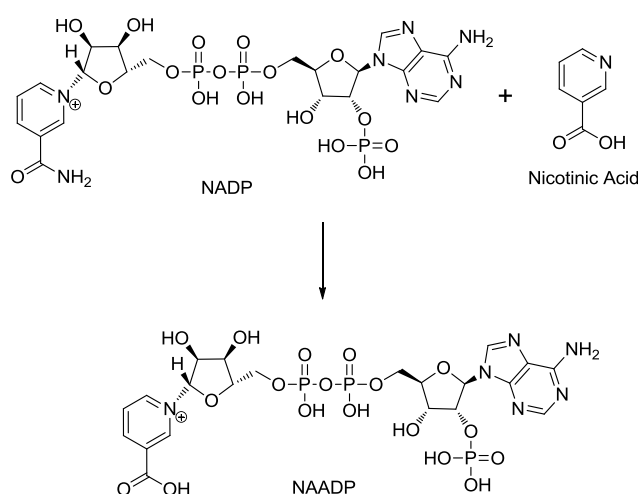


Figure 88 – It is postulated that NADP is converted to NAADP under acidic pH.¹⁸³

4.2 CD38 Structure

CD38 is a single chain 45 KDa protein.¹⁸³ Research to understand the location and function of the catalytic site has been performed by a number of research groups. Even with deletion of 15 amino acid residues from the C-terminal domain, the hydrolase activity of CD38 was maintained.¹⁹³ In contrast, deletion of 27 amino acids from the C-terminal domain, abolished the hydrolase function of CD38. This finding led to the conclusion that the active site residues were likely to be located in the C-terminal domain. Site directed mutagenesis of Glu-226 to Asp, Asn, Gln, or Gly by Lee removed enzymatic activity indicating that this was an important

catalytic residue.¹⁹⁰ A crystal structure of human CD38 with cADPR confirmed the proximity of the substrate to Glu226.¹⁹⁰

The crystal structure of the extracellular domain of CD38 was solved to 1.9 Å (**Figure 89**).¹⁹⁴ The protein is split into two domains (*C*-terminal and *N*-terminal) separated by a large hinge like cleft. From the crystal structure, it was observed that the extracellular domain is made up from a series of β -sheets in the *C*-terminal domain and of helical structures in the *N*-terminal domain (**Figure 89**). The relative conformation of the two domains is stabilised by a series of disulfide bonds between cysteine residues. Glu226 and the catalytic site exist at the hinge region of the two domains. The cavity at the hinge region between the two domains forms a solvent exposed cleft, which can be targeted by potential inhibitors of the CD38 protein.

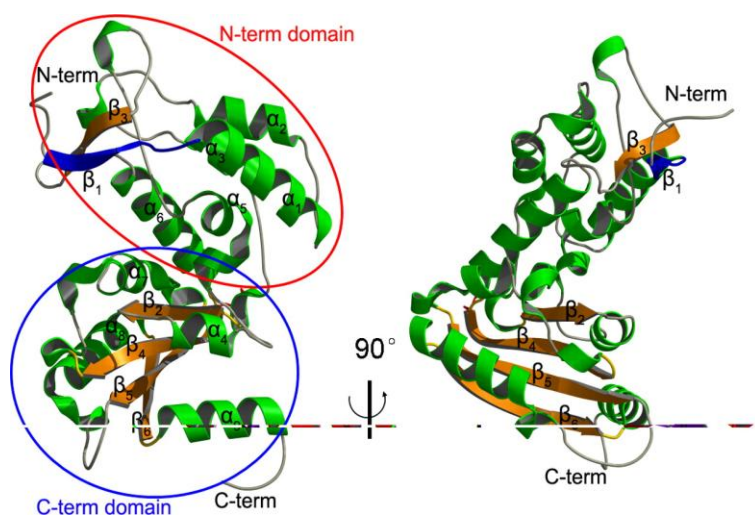


Figure 89 – Crystal structure of CD38 showing the secondary structure. Glu-226 and the catalytic site exist at the hinge region between the *C*-terminal and *N*-terminal domain.¹⁹⁴ Reprinted from Malavasi, M., Deaglio, S., Funaro, A., Ferrero, E., Horenstein, A. L., Ortolan, E., Vaisitti, T., Aydin, S. *Physiological Reviews* 2008, 88 3, 841-886, permission not required for thesis use.

Public domain X-ray crystal structures were not available for small molecule ligands bound to CD38. However, liganded CD38 crystal structures of NAD, 2'-fluoroNAD

(2'-FNAD) (**Figure 90**) and truncated analogues of 2'-FNAD were known and published.¹⁹⁵ Crystal structures of full NAD bound (with the nicotinamide intact) are only feasible when the catalytic Glu226 is mutated, preventing the release of nicotinamide. Soaking catalytically active CD38 with NAD, results in formation of the covalent intermediate from attack of Glu226 carboxylic acid side chain on the C₁ of the NAD ribose sugar, releasing nicotinamide.¹⁸⁴ The covalent acetal intermediate formed is cleaved by either nicotinamide (reversing the reaction), the adenine ring to form (cADPR) or water to give ADPR. The instability of this covalent intermediate makes obtaining a crystal structure of this intermediate complex a significant challenge.

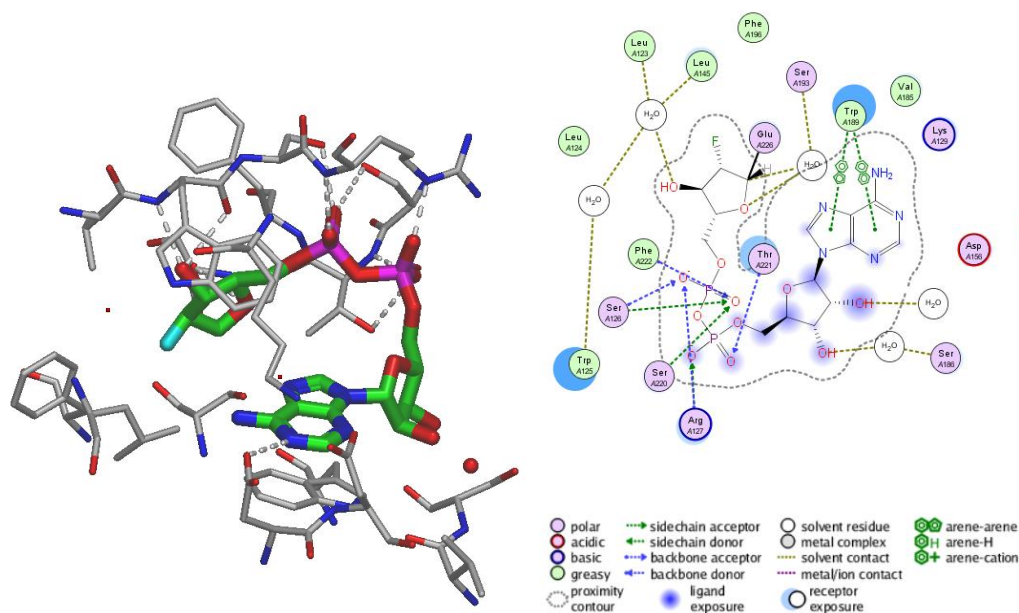


Figure 90 –2'-FNAD crystallised in CD38 (PDB: 3OFS) using dynamic conformation X-ray crystallography in a single crystal.

Many CD38 small molecule inhibitors bind in an uncompetitive manner with respect to the NAD substrate based on kinetic studies from within our laboratories.¹⁹⁶ In agreement with this result, crystallography in the absence of substrate was unsuccessful.¹⁹⁷ Due to development of an X-ray crystallography system being perceived to be critical for the success of an FBDD campaign, the opportunity to

soak fragments into CD38 covalently bound with substrate (NAD or related analogues) was explored.

Disappointingly, the covalent CD38–NAD ester intermediate was not stable enough over the time span required to soak small molecule inhibitors into the CD38–NAD crystals. Crystal structures of the CD38–NAD covalent complex have only been reported when soaking for 30 seconds at 4 °C with high levels of glycerol (30%).¹⁹⁵ Two changes were made in an attempt to increase the lifetime of the covalent CD38 – NAD intermediate, making it suitable for soaking in small molecule inhibitors.¹⁹⁷ First, Glu226 was mutated to Gln; and secondly 2'-FNAD was used instead of NAD. The covalent intermediate between the Gln primary amide side chain and the 2'-FNAD sugar takes approximately one hour to form, and subsequent nucleophilic attack of the covalent linker is electronically hindered by the fluorine atom on the 2'-FNAD. These two modifications stabilise the CD38 (Gln226 mutant) 2'-FNAD complex increasing the complex stability from 30 seconds to one to two days. The stability of the CD38 - 2'-FNAD complex allowed crystal structures of small molecule inhibitors which bind in an uncompetitive manner to be solved. Relevant X-ray crystallographic results of small molecule CD38 inhibitors bound in the CD38 (Gln226) - 2'-FNAD system will be described later in this thesis; however, few liganded X-ray crystal structures were solved through either soaking or co-crystallisation experiments.

4.3 Therapeutic Applications

Obtaining crystallographic evidence of small molecule inhibitors binding to CD38 has proved challenging. Despite this, the CD38 enzyme is currently of significant interest within the scientific community due to the number of important therapeutic applications of potential CD38 inhibitors. Literature reports have linked the function of CD38 to the potential treatment of numerous medical conditions including cachexia,¹⁹⁸ blood cancers,¹⁹⁹ myeloid and acute lymphoblastic leukaemia,²⁰⁰ type 2-diabetes,²⁰¹ and obesity.²⁰² The link of CD38 and sirtuin enzymes (SIRT) has recently increased interest in inhibition of CD38 due to the strong link of SIRT1 with

obesity and muscle metabolism.²⁰²⁻²⁰⁴ The remainder of this section will describe the hypothesised link to SIRT1 and the potential method for the treatment of obesity and metabolic syndrome.

The understanding of energy metabolism has recently advanced considerably, and it has been widely suggested that inhibition of CD38 could potentially be a method of treating obesity.¹⁸² As well as regulating NAD levels, CD38 regulates sirtuin enzymes.²⁰³ CD38 decreases the levels of intracellular NAD resulting in a reduced availability of NAD to the NAD-dependant deacetylase, SIRT1. In addition to the decrease in levels of NAD, nicotinamide levels are increased from CD38 mediated metabolism of NAD, acting as a SIRT1 inhibitor. It is proposed that activation of SIRT1 regulates glucose and fat metabolism preventing obesity and other metabolic conditions.²⁰⁵ Inhibition of CD38 increases the levels of NAD available to SIRT1 (**Figure 91**). SIRT1 uses NAD as a substrate to de-acetylate liver kinase B1 (LKB1) that phosphorylates and activates 5'-adenosine monophosphate-activated protein kinase (AMPK). Activation of AMPK inactivates acetyl-CoA carboxylase (ACC) leading to the reduction in levels of malonyl-CoA decreasing any effects from metabolic syndromes.



Figure 91 – Inhibition of CD38 increase NAD levels. SIRT1 activates AMPK which phosphorylates ACC leading to a decrease in malonyl-CoA and subsequent metabolic syndrome.

4.4 Reported CD38 Inhibitors

Although there is significant interest in the identification of CD38 inhibitors, very few small molecule compounds have been reported. A representative set of compounds reported to be CD38 inhibitors is shown in **Figure 92**. To the best of our knowledge, apart from for Luteolin, proposed binding modes or X-ray crystal

structures have not been reported for the CD38 inhibitors. The reported docking pose of Luteolin proposed a hydrogen bond interaction from a Luteolin phenol to the catalytic residue Glu226 (**Figure 93**) suggesting NAD competitive binding to CD38 expelling the substrate from the active site.²⁰⁶ Additional hydrogen bond interactions are also proposed between Luteolin and Glu146 and Asp155. A π -stacking interaction of Luteolin with Trp189 would help stabilise the binding pose. It is important to note that the majority of the CD38 inhibitors reported contain functionality such as polyphenols, quinone motifs and hydrolysable ester motifs that could cause concern around off-target effects and toxicity if attempting to develop into a marketable drug molecule.^{204,206,207} Despite the undesirable functionality, these reported CD38 inhibitors have provided evidence that small molecules can be used to control the level of conversion of NAD to cADPR and ADPR, giving optimism that small molecule treatment for one of the many diseases that NAD levels are linked to, might be found.

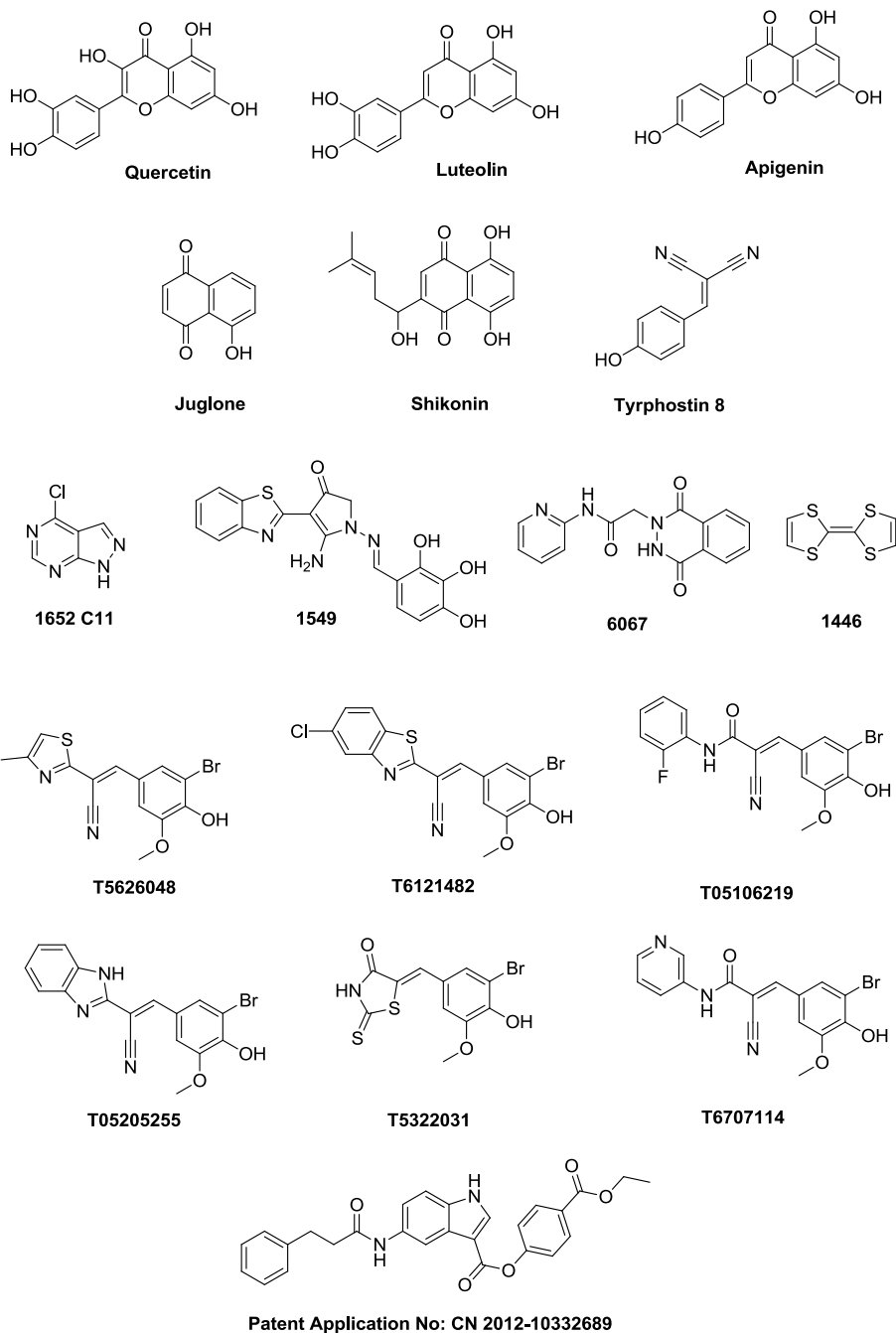


Figure 92 – Reported CD38 inhibitors.^{204,206,207}

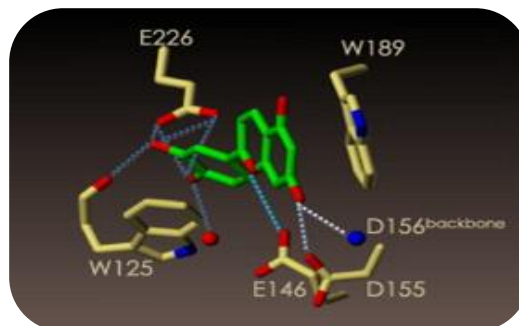
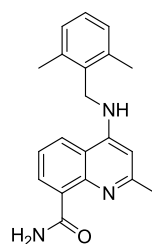


Figure 93 – Molecular modelling of Luteolin in the active site of CD38.²⁰⁶ *Reprinted from Flavonoids as inhibitors of human CD38, Kellenberger, E.; Kuhn, I.; Schuber, F.; Muller-Steffner, H. Bioorg. Med. Chem. Lett. 2011, 21, 3939-3942, Copyright 2011, with permission from Elsevier.*

4.5 Relevant Background to the CD38 Programme

The inhibition of CD38 is under investigation in our laboratories due to its effect on numerous physiological processes. Development of tool molecules to help understand and validate the proposed disease association *in vivo* is of high importance. If target validation and mechanism of action studies prove successful, these tool molecules could provide good leads to develop into drug molecules. Exploration of quinoline carboxamide analogues related to **(4.1)** (**Figure 94**) has been undertaken.²⁰⁸ Several developability concerns have been raised for this series, including low aqueous solubility (48.5 µg/mL) which correlates with the sub-optimal solubility forecast index (PFI) of 8.2. Disappointingly, P450 and hERG inhibition have been observed for a number of close analogues of **(4.1)**, which has been attributed to the basicity of the amino quinoline motif (pKa of **(4.1)** = 8.7). Although lead optimisation strategies have tried to successfully overcome these developability concerns for the quinoline carboxamide series, an alternative series was sought as a back-up to increase the chance of identifying a good tool molecule and developable CD38 inhibitors. Learnings from the HIV IN_{CCD}-LEDGF FBDD effort suggested that a fragment approach improves the chance of identifying inhibitors in good physicochemical property space. Therefore, a fragment effort was initiated to identify novel, efficient CD38 inhibitors, with improved physicochemical properties and reduced developability concerns compared to quinoline **(4.1)**.



(4.1)

Human (h) pIC₅₀ = 5.2, LE = 0.30, LLE_(AT) = 0.16
Mouse (m) pIC₅₀ = 6.4, LE = 0.37, LLE_(AT) = 0.23

P450 2D6 pIC₅₀ = 8.2; P450 2C9 pIC₅₀ = 5.0
Solubility = 48.5 µg/mL; chromLog D_{7.4} = 5.2;
hERG pIC₅₀ = 6.3; PFI = 8.2

Figure 94 – Example molecule from the quinoline carboxamide series under investigation for inhibition of CD38.

4.5.1 Fragment Effort Aims

The aims of the fragment effort (**Figure 95**) would address the concerns associated with the quinoline carboxamide series (**4.1**). The fragment programme was required to deliver a novel lead molecule in good physicochemical property space ($\text{PFI} < 6.0$), that can be used as a tool compound. The molecule was required to have $\text{pIC}_{50} \geq 6.0$ (against both mouse and human orthologues), with minimal hERG or p450 liabilities.

Human and mouse $\text{pIC}_{50} \geq 6.0$

LE and $\text{LLE}_{(\text{AT})} > 0.3$

hERG $\text{pIC}_{50} \leq 4.5$

P450 $\text{pIC}_{50} \leq 5.0$

$\text{PFI} < 6.0$

Solubility $\geq 100 \mu\text{g/mL}$

Figure 95 – Fragment effort criteria for a novel lead molecule.

To identify fragment hits with good physicochemical properties, two hit identification strategies were used:

- 1) screening of the fragment compounds through biophysical techniques^{209,210} to triage to X-ray crystallography,¹⁵³
- 2) re-analysis of historical high throughput screening (HTS) data.¹⁷²

4.5.2 CD38 Screening Cascade

Two fragment sets were screened simultaneously; our laboratories' FBDD core set (1045 compounds), and a knowledge-based set of 93 compounds constructed by fragmentation of and similarity searching around known CD38 binders.²¹¹ The biochemical assay was not sufficiently sensitive to be used as part of the fragment screening cascade, as the lower limit of detection of the assay was $\text{pIC}_{50} = 4.0$. The lack of high concentration biochemical screening detection capabilities for CD38 was not a concern, as learnings from the HIV IN_{CCD} -LEDGF project were followed resulting in multiple biophysical methods being developed.

SPR²¹² was carried out on both libraries in single concentration format at 500 μ M (**Figure 96**). Compounds that showed a change in the refractive index were re-analysed in dose response format with a top concentration of 1 mM. In total, 64 fragments showed binding. For a fragment to be classified as binding, either a complete curve was observed allowing measurement of a K_d, or linear binding observed where the top saturation of the curve was not reached. Without saturation of the curve, a K_d measurement cannot be quoted.

All 64 fragments which showed binding in SPR were progressed into saturation transfer difference (STD) NMR studies to gain increased confidence of binding of the fragment to the protein. Post-project analysis of the HIV IN_{CCD}-LEDGF effort clearly demonstrated increased success in X-ray crystallography for compounds which were displaced by a known binder. Therefore, reversal of the STD NMR signal for the CD38 fragments by competition studies, using a close analogue of **(4.1)** was pursued. In total 31 compounds were displaced with the known inhibitor, confirming binding in the active site of the CD38 protein, and all were submitted for X-ray crystallographic studies.¹⁹⁷ Of the 31 compounds, 7 compounds provided liganded CD38 crystal structures giving a total hit rate of 0.61%. The knowledge-based set gave an increased crystallographic hit rate (4.3%), compared to the FBDD core set (0.29%) (**Figure 96**), demonstrating the value of using knowledge of the binding site (and substrates) to select a set of molecules which could bind to the target.

To supplement the 7 crystallographic hits, the FBDD expansion set (1409 compounds), the kinase hinge binder set (1458 compounds), and compounds structurally related to the initial hits (71 compounds) were screened through the SPR assay.¹⁹⁷ The kinase hinge binding set was selected due to the compounds having hydrogen bond donor-acceptor-donor motifs, or truncated donor-acceptor/acceptor-donor which are common in many kinase inhibitors. The CD38 protein contains a region of three amino acids with carboxylic acid side chains (explained in further detail in **Section 4.6.1**). The close proximity of the carboxylic acid side chains

means that they are unlikely to all be deprotonated at physiological pH, as repulsion of the subsequent negative charges would occur. The exact protonation states of each carboxylic acid side chain are unclear but compounds with a pattern of closely located hydrogen bond acceptors and donors had a high chance of binding to this carboxylic acid region blocking binding of the substrate. All compounds with a measurable K_d from SPR were progressed to X-ray crystallography, NMR STD and biochemical screening in parallel. Of the 2938 compounds screened, only one compound from this second round of screening gave a liganded X-ray crystal structure with CD38.

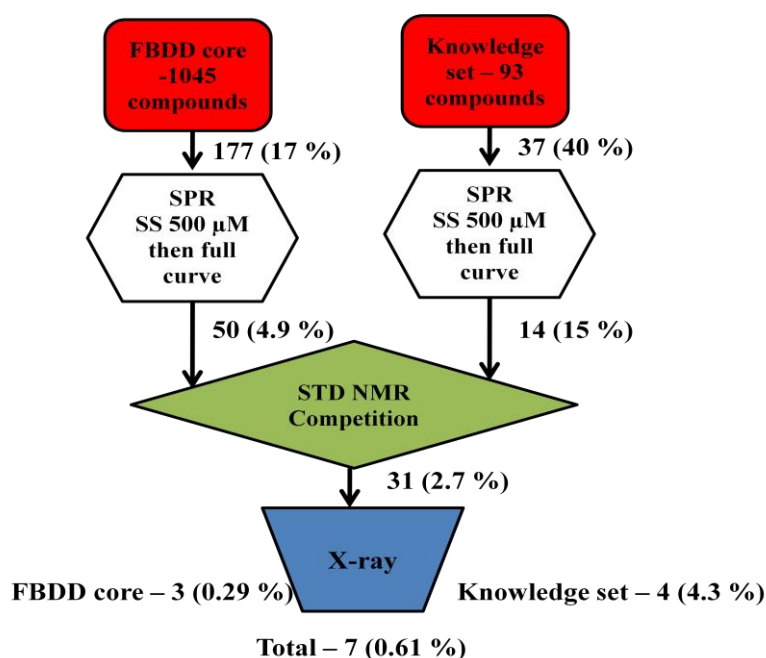


Figure 96 – CD38 fragment screening cascade with percentage-hit rates for each screening set.

Although the CD38 protein pocket exists as a large open cavity, all 8 X-ray crystallography hits occupy the same region of the protein pocket (**Figure 97**). The similarity in ligand location is hypothesised to be driven by the ligands forming π -stacking interactions with Trp189, and having one or more interactions with acid residues; Glu146, Asp155 or Asp 156 (**Figure 97**). Despite the similarity in binding

location and interactions, the eight X-ray crystal structures were grouped using vectorial clustering into five clusters (**Figure 98**).^{165,166} Each cluster provided a distinct growth vector to explore through sub-structure searching of our laboratories' compound collection and synthesis. Optimisation of the ligand-protein binding of 1*H*-imidazo[4,5-*c*]pyridin-4-amine (**4.5**) from cluster 3, to CD38, is described as part of this thesis. This compound was of particular interest as the X-ray structure showed that imidazopyridine (**4.5**) made three hydrogen bond interactions with the protein. As discussed in the learnings from HIV IN_{CCD}-LEDGF, the presence of multiple hydrogen bonds was hoped to minimise movement of the fragment in the protein pocket as the fragment is grown. Although follow-up of binding of compounds in clusters 1, 2, 4 and 5 was also carried out in our laboratory, this is not described as it remains outside the scope of this work.

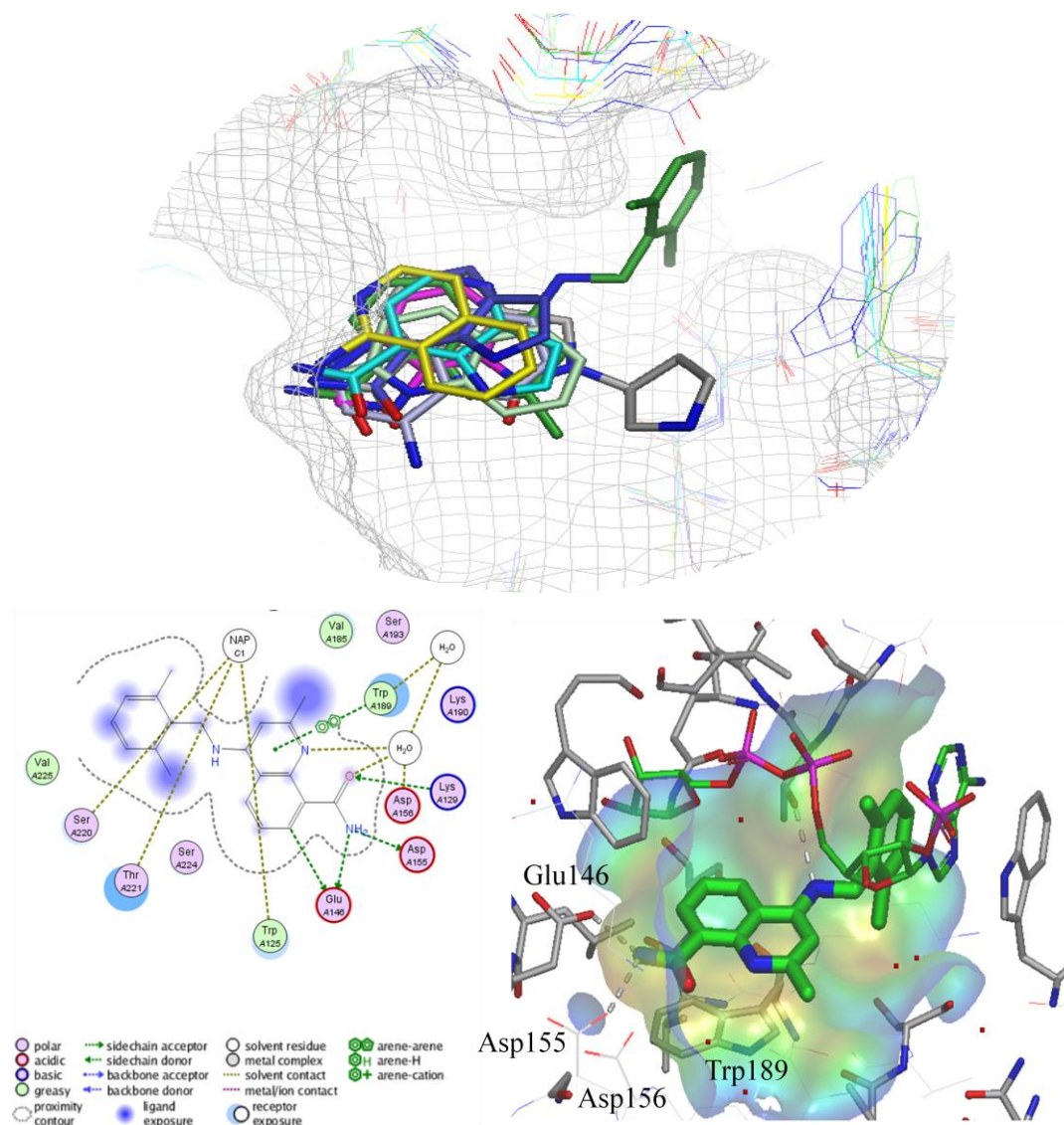
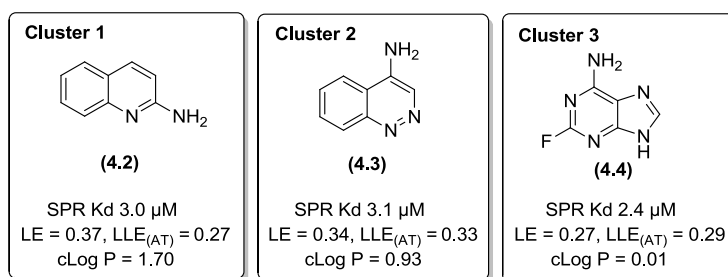


Figure 97 – Top image: X-ray structures of fragment hits in the CD38 pocket. Ligands are shown in bold bonds with the CD38 protein surface represented by a grid. **Bottom image:** X-ray crystal structure of quinoline carboxamide (**4.1**) (PDB: 1GECC).

Crystal structures derived from the FBDD core set



Crystal structures derived from knowledge based screening, kinase hinge binding set and FBDD expansion set

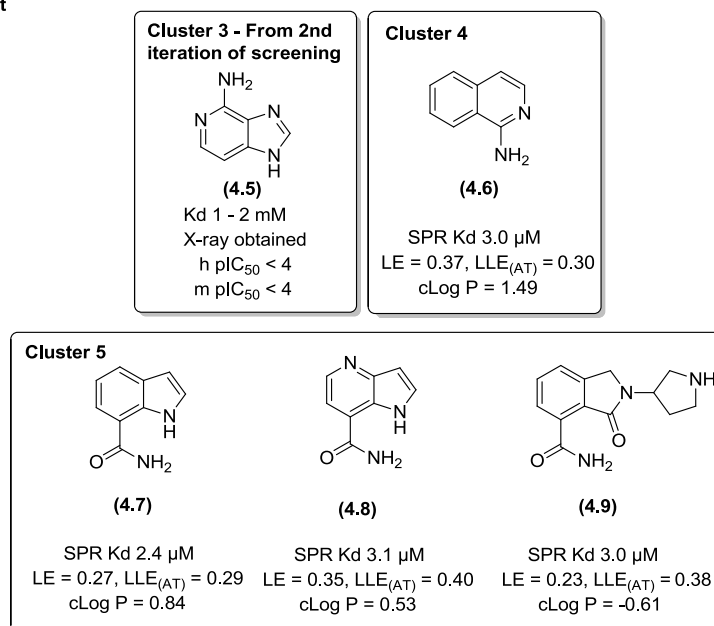
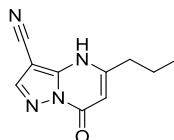


Figure 98 – Clustering of hits from the CD38 screening cascade.

As for previous fragment efforts, to complement the hits from the fragment and knowledge-based screens, a HTS top-up screen was carried out, reviewing the biochemical data with respect to efficiency (LE and LLE_(AT)) rather than potency.^{165,166,172,213} Historically, when reviewing HTS data, research groups have focused on potency and have disregarded weakly potent but efficient binders. Although more potent compounds are likely to achieve the potency requirements for the programme in a shorter length of time, these compounds often suffered from poor physicochemical properties, which may cause developability issues. This process highlighted related exemplars of **(4.5)** with good efficiency (LE and LLE_(AT) > 0.3) and a distinct template, pyrazolopyrimidinone **(4.10)** (**Figure 99**). Follow-up of this ligand efficient screening hit was carried out and is described as part of this thesis.

Pyrazolopyrimidinone (**4.10**) was of particular interest due to it having activity in both the mouse and human enzyme assays, regardless of its failure to produce a liganded CD38 X-ray crystal structure.



(4.10)

Mouse pIC_{50} = 7.1, LE = 0.65, $LLE_{(AT)}$ = 0.73
Human pIC_{50} = 5.1, LE = 0.47, $LLE_{(AT)}$ = 0.55
PFI = 2.6

Figure 99 – Pyrazolopyrimidinone hit identified from re-analysis of the HTS top-up data from a ligand efficiency perspective.

4.5.3 Challenges with the CD38 X-ray Crystallographic System

Despite liganded CD38 crystal structures being obtained for the eight fragment compounds shown in **Figure 98**, routinely securing liganded crystal structures proved challenging.¹⁹⁷ Throughout the remainder of this thesis, despite numerous attempts, no additional liganded X-ray crystal structures of the fragment series were obtained. Variables within the X-ray crystallographic conditions were explored using mouse, rat, human and dog CD38 constructs, exploring soaking into apo-crystals, back soaking out ligands, and co-crystallisation.¹⁹⁷ Variation of the length of experiment, a range of temperatures, variation in the structure of the substrate and different glycosylation levels of the protein, failed to provide a suitable system to help guide structure based optimisation and growth. It was hoped that as the fragments increased in binding affinity, the chance of obtaining liganded X-ray crystal structures would improve. Unfortunately, this was not true for the fragment series developed in this research programme. This challenge was not unique to these compounds. The quinoline carboxamide series (**4.1**), was also unable to routinely provide liganded CD38 crystal structures, despite X-ray crystallography of close analogues confirming that they bind at the CD38 catalytic site. The inability to develop an X-ray crystal system, suitable to guide structure based design, differentiates this programme of work from a majority of other FBDD programmes embarked upon within our laboratories, including HIV IN_{CCD} -LEDGF (**Chapter 2**).

Accordingly, this provides a unique challenge to understand how successful FBDD can be without routine X-ray crystallography to guide fragment growth and optimisation.

4.5.4 The CD38 Screening Cascade

The CD38 project was also distinct from other drug discovery programmes undertaken within international laboratories, due to the lack of a cellular assay to profile the molecules and help guide the drug discovery effort. This programme of work to identify CD38 inhibitors was initiated after positive early target validation studies.^{196,208,212} A strong body of evidence has been building in the literature suggesting that CD38 inhibitors may be beneficial for a number of diseases with unmet medical need.^{190,199-201,203,205,206} Despite this, additional tool molecules with good physicochemical properties and pharmacokinetic profiles are required to further validate these observations *in vivo*.

A typical drug discovery programme can be broken down into five phases: target validation; enzymatic assays; cell-based screening and profiling; *in vivo* mode of action, PK/PD and safety studies; and, finally, clinical studies leading to market launch (**Figure 100**).²¹⁴ A standard fragment screening campaign would aim to use biophysical and biochemical enzymatic screening to identify and perform iterative design to improve the potency and physicochemical properties for the series. When sufficient enzymatic potency is obtained (~ 1 µM), evidence of cellular activity is usually sought as a bridge between enzymatic screening and *in vivo* studies. At this point, measures of other important developability criteria are assessed, e.g. hERG and P450 profiles.

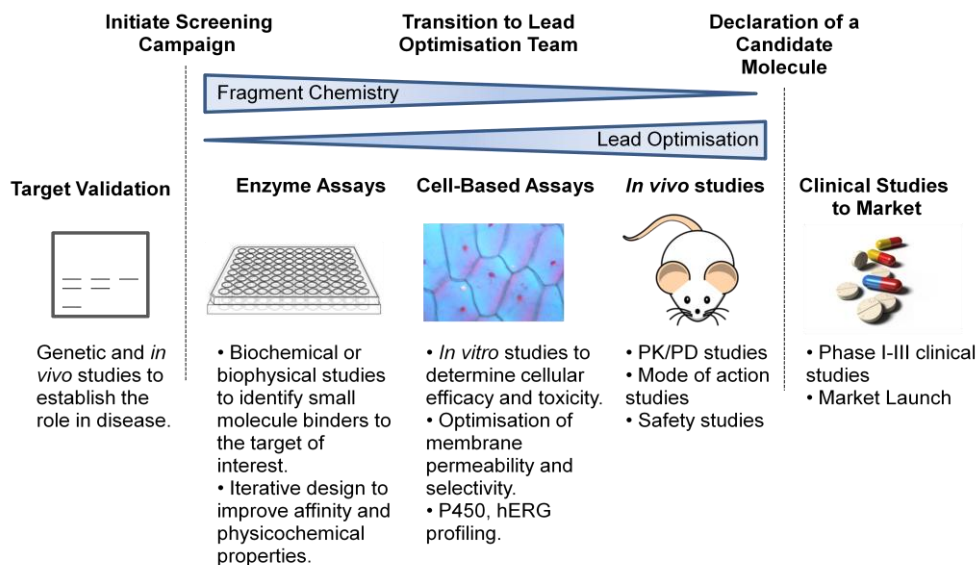


Figure 100 – Development of a typical drug molecule – from target validation to market.²¹⁴

As stated above, a cellular assay can be used as an intermediate screen before progressing a molecule into *in vivo* studies. A cell assay gives confidence that the compound has the ability to have the desired effect at the target and that the molecule is cell permeable. For CD38, the set-up of a cell-based assay was not successful and, therefore, the enzymatic assay against mouse, human and rat constructs was used in parallel to develop compounds suitable for *in vivo* studies. The mouse enzyme assay uses a fluorescence intensity (FLINT) MOA format, where remaining NAD that is not hydrolysed due to inhibition of CD38, is fluorometrically measured using alcohol dehydrogenase to convert the NAD to NADH.²¹² In contrast to this, the rat and human enzyme assays use a calorimetric readout. The calorimetric readout monitors inhibition of product formation arising from the catalytic CD38 mechanism.²¹² In the absence of a cell-based assay, measuring enzyme activity *via* two different assay formats, gives confidence that the compounds are inhibiting the function of CD38. In preparation for use of the molecules identified from this research programme as *in vivo* tools, it is important that molecules are developed with: low-moderate lipophilicity, good solubility (PFI < 7) and low molecular weight to maximise the chance of good solubility and membrane permeability. Molecules with the desired

profile for *in vivo* studies would be transferred to the relevant team for further exploration.

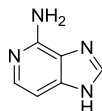
4.5.5 CD38 Research Aims

This CD38 research programme was initiated post fragment screening and clustering. The aims of this research programme described here include:

- Optimisation of fragment hit 1*H*-imidazo[4,5-*c*]pyridin-4-amine (**4.5**) (**Figure 101**) and/or HTS-derived pyrazolopyrimidinone (**4.10**) to identify novel CD38 inhibitors with the desired property and potency profile (**Figure 102**) to provide tool molecules. This will be effected by:
 - Understanding the interactions made between the ligand and the protein; and
 - Targeted growth of the compounds to pick up new interactions with the protein pocket whilst maintaining good efficiency (LE and LLE_(AT) > 0.30) and PFI < 6.0;

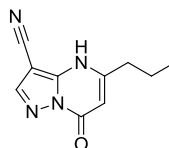
These aims were carried out by:

- Formulation of plans for optimisation of fragment hits through SSS and synthesis;
- Design and docking of key structures using LigX in the Molecular Operating Environment (MOE); and
- Continuous analysis of biochemical, biophysical, and X-ray crystallography, where available.



(4.5)

m pIC₅₀ < 4
h pIC₅₀ < 4
X-ray obtained
K_d = 1-2 mM



(4.10)

m pIC₅₀ = 7.1, LE = 0.65, LLE_(AT) = 0.73
h pIC₅₀ = 5.1, LE = 0.47, LLE_(AT) = 0.55
No X-ray, no binding by SPR

Figure 101 - Fragment hit 1*H*-imidazo[4,5-*c*]pyridin-4-amine (**4.5**) and HTS-derived pyrazolopyrimidinone (**4.10**).

Human and mouse pIC₅₀ ≥ 6.0

LE and LLE_(AT) > 0.3

hERG pIC₅₀ ≤ 4.5

P450 pIC₅₀ ≤ 5.0

PFI < 6.0

Solubility ≥ 100 µg/mL

Figure 102 – Fragment effort criteria.

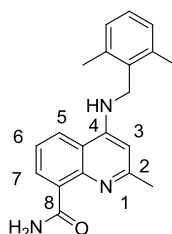
The candidate has performed all preparative chemistry, analysis of data, and review of X-ray crystal structures to form design and synthesis strategies, unless referenced otherwise.

4.6 Results and Discussion

4.6.1 Analysis of the CD38 Binding Pocket

Prior to fragment optimisation and growth of imidazopyridine (**4.5**) (**Figure 101**) and HTS-derived pyrazolopyrimidinone (**4.10**), analysis of the CD38 binding site was performed. It was envisaged that the increased knowledge of the CD38 binding site would allow careful optimisation of fragment compounds and improve the CD38 protein-ligand interactions. Improvement of the protein–ligand interaction can be carried out by maximising the shape complementarity and/or pairing of electrostatics between the fragment and protein. An X-ray crystal structure of quinoline carboxamide (**4.1**) had been solved.¹⁹⁷ Quinoline carboxamide (**4.1**) (**Figure 103**) binds uncompetitively with the NAD substrate, which in the X-ray crystal structure is covalently bound to Gln226 (**Figure 104**). The adenine ring forms a triple π -stacking interaction between the dimethyl phenyl ring of the ligand, adenine ring of the substrate, and Trp176. The substrate is flexible, but this triple π -stack demonstrates how binding of the ligand, protein, and substrate simultaneously could help stabilise the ligand-protein complex. The quinoline ligand (**4.1**) has multiple hydrogen bonds to the protein amino acid side chains (**Figure 104**). At pH 6.5 in the crystal soaking system, the quinoline nitrogen is protonated and forms an intramolecular hydrogen bond to the carboxamide carbonyl oxygen. The amide NH₂ of (**4.1**) forms two hydrogen bond interactions, one to the side chain of Glu146 (NH – O 3.03 Å) and one to the side chain of Asp155 (NH – O 2.84 Å). The 4-NH picks up a hydrogen bond interaction to the side chain of Thr221 (NH – O 2.93 Å). In addition to these hydrogen bond interactions, the quinoline core of (**4.1**) forms a π -stacking interaction with Trp189 (quinoline – Trp189 distance ranges between 3.2 – 3.5 Å). The 2,5-dimethylphenyl ring sits in close proximity to Trp176, but without substrate present is insufficiently close to pick up a π -stacking interaction (closest point between 2,5-dimethylphenyl and Trp176 is ~ 5 Å). Instead a triple π -stacking

interaction is formed sandwiching the adenine ring of the substrate between the 2,5-dimethylphenyl and Trp176 (**Figure 104**). Consideration of ways to improve binding of the molecules to CD38 by forming additional interactions with the NAD substrate may be helpful during this project.



(4.1)

h pIC₅₀ = 5.2, LE = 0.30, LLE_(AT) = 0.16
m pIC₅₀ = 6.4, LE = 0.37, LLE_(AT) = 0.23

Figure 103 - Structure and data for CD38 binder, quinoline carboxamide (**4.1**) which is assumed to be protonated at physiological pH. The biochemical enzyme data is presented as human pIC₅₀ (h pIC₅₀) and murine pIC₅₀ (m pIC₅₀).

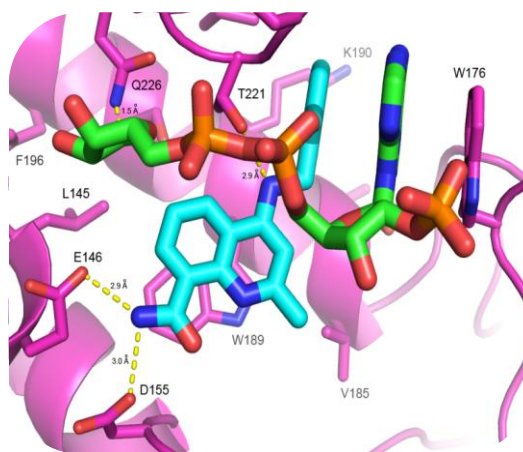


Figure 104 - X-ray crystal structure of quinoline carboxamide (**4.1**) (PDB:1GECC – with substrate present) in human CD38 showing that uncompetitive binding with respect to the substrate occurs. A triple π -stack is formed between the dimethyl phenyl ring of the ligand, adenine ring of the substrate, and Trp176.¹⁹⁷ A Glu226 to Gln mutation is present.

Two amino acid residues are involved in the CD38-mediated conversion of NAD to ADPR or cADPR, Glu226 and Ser193 (highlighted by the orange circle in **Figure 105**). No direct hydrogen bond interactions from (**4.1**) are made to either of these catalytic residues. It is postulated that quinoline carboxamide (**4.1**) inhibits the function of CD38 by preventing NAD from occupying the binding pocket and achieving the required close proximity to the catalytic residues. Quinoline carboxamide (**4.1**) competes with NAD by binding in the region where nicotinamide from NAD would reside. To demonstrate this clash, the X-ray crystal structures of quinoline carboxamide (**4.1**) and NAD are superimposed (**Figure 106**). The adenine ring also clashes with the ligand, but this is unlikely to contribute to the inhibition due to the flexibility of NAD and the additional room in the CD38 binding pocket.

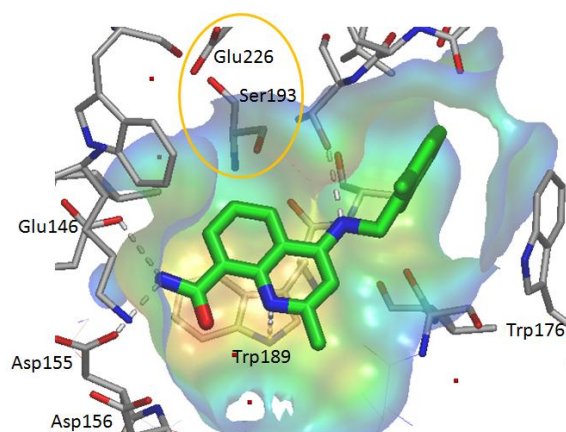


Figure 105 – X-ray crystal structure of quinoline carboxamide (**4.1**) soaked into human CD38 apo crystals at 5 mM for 24 h (PDB: 1OCRI). Minimal change to the binding mode is observed compared to PDB: 1GECC (with substrate).

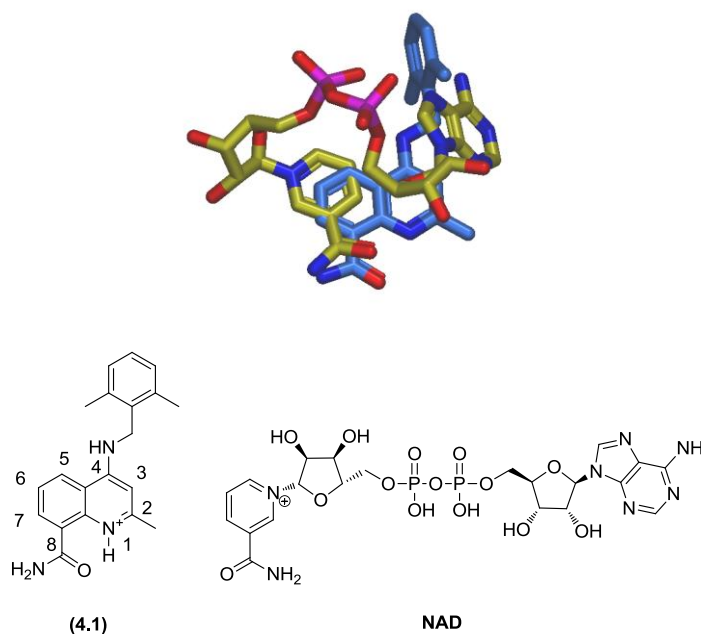


Figure 106 - X-ray crystal structure of quinoline carboxamide (**4.1**) in human CD38 (PDB: 1OCRI – wild type human CD38) overlaid with the crystal structure of NAD bound to CD38 with Glu226 mutated to Gln (PDB: 2I65).

From consideration of the available X-ray crystallography data, the CD38 protein pocket has been sub-classified into 5 areas (**Figure 107**). The 5 areas are:

1. *Catalytic pocket*. A polar region where the catalytic residues Glu226 and Ser193 reside. This is where the tetrahydrofuran-3,4-diol moiety attached to the nicotinamide portion of NAD binds for conversion to ADPR or cADPR.
2. *Carboxylic acid region*. Glu146, Asp155, and Asp156 form an acidic wall in the protein pocket. X-ray crystallography has shown that the nicotinamide amide of NAD forms hydrogen bond interactions with this part of the protein. The nicotinamide interactions are mimicked by quinoline carboxamide (**4.1**). All the hits from the fragment screen form an interaction with this region of the protein (**Figure 107**).
3. *Trp189 region*. Both nicotinamide and quinoline carboxamide (**4.1**), form favourable π -stacking interactions with Trp189. In addition to this, all the fragment hits bind over Trp189 (**Figure 107**).

4. *Trp176 shelf*. This is a protein channel leading out to solvent. Quinoline carboxamide (**4.1**) reaches into this channel. Trp176 is available for forming a π -stacking interaction the adenine of the substrate. However, flexibility of this binding loop has been observed and marked differences are seen for this region between mouse and human orthologues (**Figure 108**). This results in the design of analogues to form specific interactions to this region of the protein being challenging.
5. *Human/mouse selectivity region*. In the active site of the protein an amino acid change is observed between human and mouse protein constructs: Thr158 to Arg158, respectively. When growing into this region, differences in the human and mouse binding potency could be seen due to this amino acid change. Although the long term aim was to identify a medicine to treat humans, but activity against the mouse or rat CD38 protein was essential to carry out the necessary animal studies to advance the molecule towards human trials.

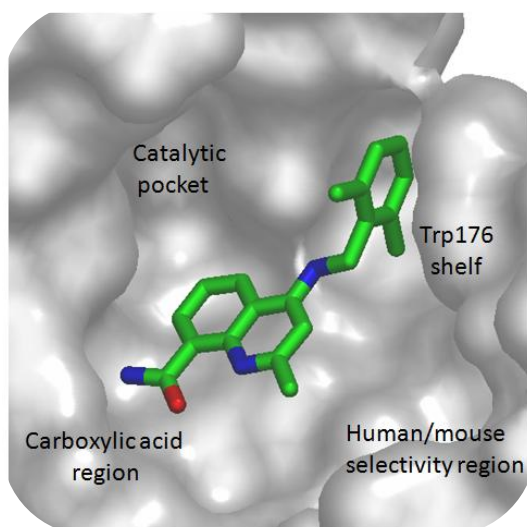


Figure 107 – CD38 binding pocket sub-classified into five areas for binding (PDB: 1OCRI – wild type human CD38).

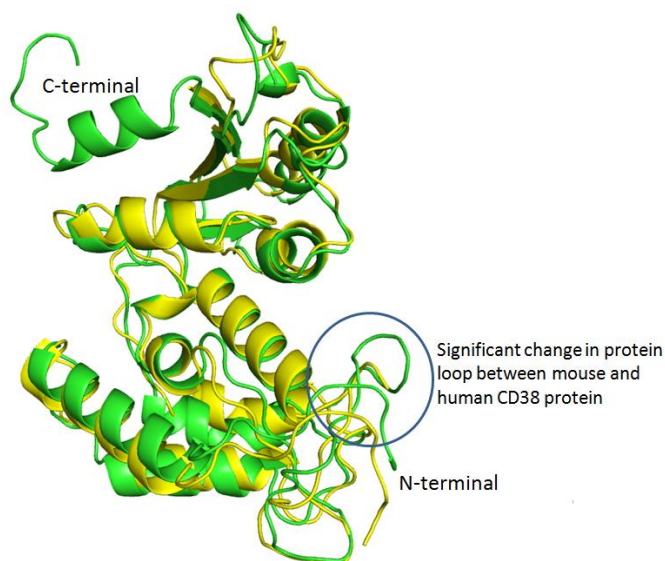


Figure 108 – Difference in Trp176 binding loop between mouse and human protein APO structures.¹⁹⁷ Green and yellow ribbon represent human and mouse constructs, respectively.

In response to the recommendations from the HIV IN_{CCD} project, before beginning fragment optimisation for CD38, the computational programme GRID was run to help characterise the binding pocket and understand areas where profitable interactions could be made. The grid probes for CD38 were strong and well defined (**Figure 109**) and showed that interactions along the acid region were predicted to be beneficial, along with reaching into the area occupied by the dimethyl phenyl of quinoline carboxamide (**4.1**). In addition to this, the GRID analysis suggested that methyl (represented by green GRID) or amino groups (represented by blue GRID) placed in the selectivity region or on the Trp176 shelf may improve binding. GRID probes also fill the catalytic pocket, however, this region was predicted to be challenging to target due to the very polar, narrow nature of the pocket.

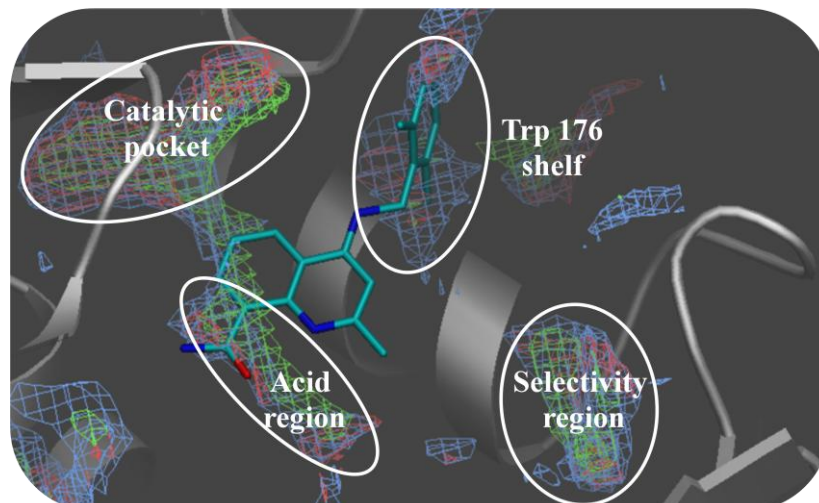


Figure 109 – GRID probes: Key to GRID – Green = CH₄, Blue = NH₄⁺, Red = C=O.

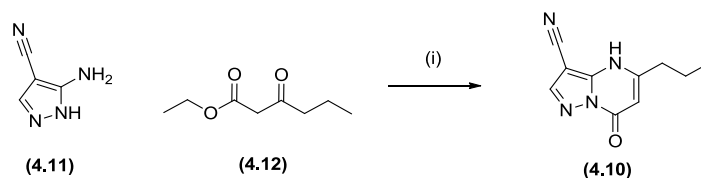
In summary, analysis of the CD38 binding pocket through consideration of X-ray crystal structures of historical compounds and the required binding interactions of NAD, combined with the use of the computational technique GRID, helped build a clear hypothesis of regions of the CD38 pocket that would be profitable to occupy with small molecule inhibitors. It was hypothesised that the key binding interactions are likely to be made to the acid region of the protein through H-bond interactions (**Figure 109**). Additional binding affinity could potentially be achieved by growing into the Trp176 shelf region, or into the selectivity region taking care to retain activity in both mouse and human constructs. Consideration of the possible growth vectors, and assessing different approaches to access these key regions of the CD38 active site from the fragment starting points, was deemed critical to develop the fragment molecules into lead templates.

4.7 Fragment Optimisation and Growth of 7-Oxo-5-propyl-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.10)

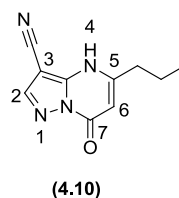
Pyrazolopyrimidinone (**4.10**) was identified through re-analysis of the HTS top-up data from a ligand efficiency perspective (**Figure 110**). Pyrazolopyrimidinone (**4.10**) (h pIC₅₀ = 5.1, LE = 0.47, LLE_(AT) = 0.55; m pIC₅₀ = 7.1, LE = 0.65, LLE_(AT) = 0.73) was chosen for follow-up through substructure searching (SSS) and synthesis. The fragment-like compound (**4.10**) had demonstrated affinity in both the mouse and human enzyme assays with good LE and LLE_(AT), both above 0.3 for the mouse and human orthologues. For previous fragment efforts, compounds that showed good affinity and efficiency have been found to be false positives (compounds that exhibit enzyme activity through other means than from binding to the protein). The false positive nature of compound samples is postulated to be due to minor impurities causing assay interference, often through compound fluorescence, resulting in problems with reproducibility of the assay result. Historically, isolation and characterisation of such compound contaminants in our laboratories has proved challenging and unsuccessful. Although the enzyme potency for pyrazolopyrimidinone (**4.10**) had been confirmed on retest from the 10 mM DMSO stock solution, the compound was re-synthesised to rule out any false positive component to this measurement.

Pyrazolopyrimidinone (**4.10**) was synthesised by condensation of amino pyrazole (**4.11**) and β-keto ester (**4.12**) under acidic reflux (**Scheme 19**). Regioselectivity in this reaction was influenced by the more nucleophilic *exo*-amino group, preferentially reacting with the ketone, prior to cyclisation and dehydration to form the 5,6 ring system. Pyrazolopyrimidinone (**4.10**) was obtained in a 40% yield. Disappointingly, enzymatic screening of the re-synthesised batch of (**4.10**), resulted in low inhibition values against both mouse and human compared to that of the original sample (h pIC₅₀ < 4.0; m pIC₅₀ = 5.1, LE = 0.47, LLE_(AT) = 0.55) (**Figure 110**). Human enzyme potency had dropped below the detection limit of the assay. Inspection of LCMS and NMR data for both batches did not provide insight into the

potency discrepancy, as the two batches appeared identical. Despite the reduction in potency, pyrazolopyrimidinone (**4.10**) still had good efficiency against the mouse protein, comparable to or greater than all other hits from the HTS and fragment screens. For this reason, it was decided that this template still warranted further investigation.



Scheme 19 – Synthesis of pyrazolopyrimidinone (**4.10**). Reagents and conditions: (i) acetic acid:water 2:1, 100 °C, 17 h, 40%.



'Stock' sample: h pIC₅₀ = 5.1, LE = 0.47, LLE_(AT) = 0.55
 m pIC₅₀ = 7.1, LE = 0.65, LLE_(AT) = 0.73

Resynthesised sample: h pIC₅₀ < 4.0
 m pIC₅₀ = 5.1, LE = 0.47, LLE_(AT) = 0.55

Figure 110 – CD38 enzyme data for two batches of pyrazolopyrimidinone (**4.10**).

4.7.1 Proposed Binding Mode for Pyrazolopyrimidinone (**4.10**)

Despite attempts to obtain a liganded X-ray structure of pyrazolopyrimidinone (**4.10**) in mouse or human CD38 protein, an X-ray crystal structure remained elusive.¹⁹⁷ Historical fragment programmes explored in our laboratories have demonstrated that even small modifications to a fragment can result in X-ray crystallographic success. With the hope that future pyrazolopyrimidinone fragment analogues may provide a liganded CD38 crystal structure for the series, SSS and single point changes through chemical synthesis around fragment (**4.10**) were initiated. In the absence of a crystal structure, it was challenging to design an appropriate set of compounds, which answered specific questions relating to the binding location and orientation of the

fragment. To aid in design of compounds for synthesis, computational docking using MOE was performed in collaboration with our laboratories' computational modeller.¹⁶⁵ Compound **(4.10)** was docked from outside the CD38 protein, with no bias towards the binding location or orientation. Multiple binding modes of pyrazolopyrimidinone **(4.10)** in the CD38 binding pocket were obtained. Two binding modes were of particular interest because both showed π -stacking over Trp189 which lines the base of the active site, and interactions to the carboxylic acid region of the protein (**Figure 111** and **Figure 112**). On consideration of the proposed binding modes compared with the GRID map (**Figure 109**), both binding modes satisfy the interactions in the acid region of the protein pocket and the propyl provides vectors to explore towards the selectivity region or the Trp176 shelf.

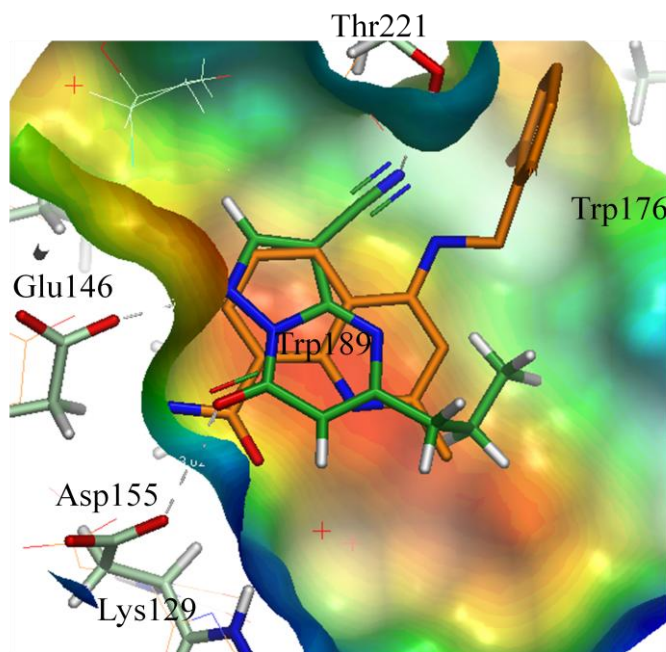


Figure 111 – Overlay of quinoline carboxamide **(4.1)** in human CD38 (PDB: 1OCRI) with docking pose 1 of pyrazolopyrimidinone **(4.10)**.

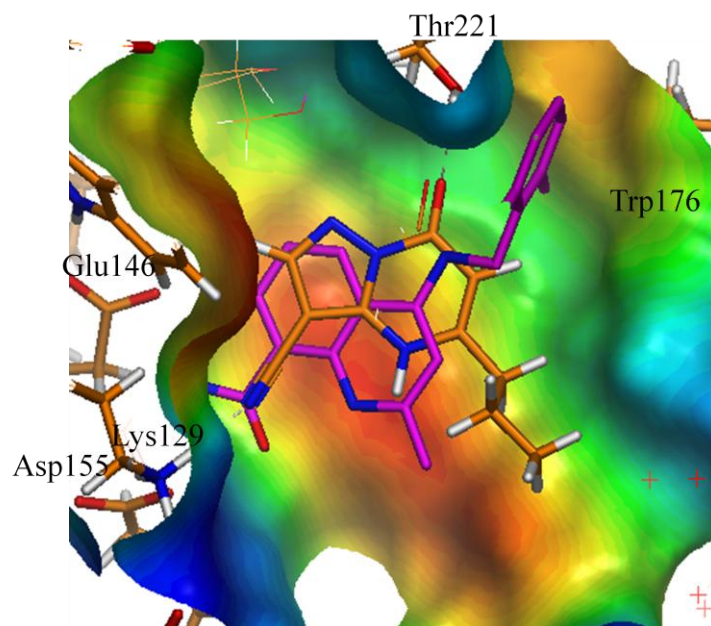


Figure 112 - Overlay of quinoline carboxamide (**4.1**) in human CD38 (PDB: 1OCRI) with docking pose 2 of pyrazolopyrimidinone (**4.10**).

Docking pose 1 of pyrazolopyrimidinone (**4.10**) (**Figure 111**) demonstrated a hydrogen bond through the carbonyl oxygen at the 7-position to Asp155 (O – HO 3.02 Å) and to the flexible Lys129 (O – HN 3.01 Å). In addition to this, a third hydrogen bond interaction was observed to the carboxylic acid region, through the 1-*H* tautomeric hydrogen to Glu146 (NH – O 2.70 Å). The protonation states of the three carboxylic acids in this region (Glu146, Asp155 and Asp156) remain uncertain. What is clear is that it is not likely that all three carboxylic acid side chains would be deprotonated, due to repulsion of the subsequent negative charges. In addition to this, the tautomeric form of (**4.10**) bound in the CD38 protein pocket remains inconclusive. Calculations of the relative stability of each tautomer showed that in water, pyrazolopyrimidinone (**4.10**) would exist exclusively as the 4-*H* tautomer (**Figure 113**).²¹⁵ The calculations were performed on the 5-methyl analogue of (**4.1**) to simplify the experiment and reduce the computational time required. When bound in the CD38 binding site in binding pose 1, the carbonyl and pyrazole 1-*N* of (**4.10**) are unlikely to both act as acceptors, as this would require both Glu146 and Asp155 to be protonated. Therefore, in the CD38 active site, the 4-*H* tautomer is

unlikely to exist. The computationally calculated energy difference between the 7-*H* and 1-*H* tautomers shows that the 1-*H* tautomer is the least stable and, therefore, it is hypothesised that the 7-*H* tautomer is most likely to occur in the CD38 active site. However, conclusive determination of the exact tautomeric form required for binding would be difficult to obtain, as the tautomeric form could be affected by the protonation states of the amino acid residues in the protein pocket. A fourth hydrogen bonding interaction was observed from the nitrile nitrogen to Thr221 (N – HO 3.02 Å). The combination of the four hydrogen bonds gave confidence that the fragment could be anchored strongly reducing the probability of the molecule moving in the protein pocket as the fragment was grown. This gave encouragement that the template was suitable for optimisation, as starting from a fixed fragment was a key learning from the HIV IN_{CCD}-LEDGF effort.

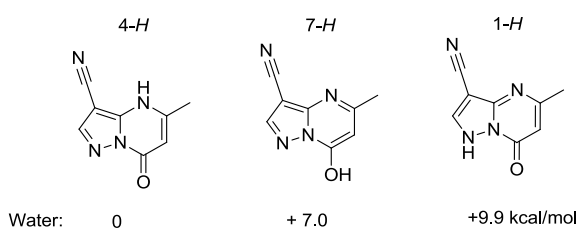


Figure 113 – Relative tautomer stability of pyrazolopyrimidinone (**4.10**) calculated using LMP2/ccPVTZ(-F)/water//B3LYP/6-31+G**/water.²¹⁵

Docking pose 2 of pyrazolopyrimidinone (**4.10**) (**Figure 112**) also showed multiple hydrogen bonds in the CD38 protein pocket. The compound docked in the lower energy 4-*H* tautomeric form, with hydrogen bond interactions from the nitrile nitrogen to Lys129 (N – HN 2.99 Å). A second hydrogen bond interaction was made from the carbonyl oxygen to Thr221 (O – HO 2.82 Å).

Overlay of docking poses 1 and 2 of pyrazolopyrimidinone (**4.10**), emphasises the difference in binding location and orientation (**Figure 114**). Both docking poses showed that substitution from the 2-position would allow growth into the substrate catalytic pocket, but at slightly different angles. Substitution from the 2-position looked more feasible from docking pose 1. For docking pose 2 a clash with the

protein would occur, unless a movement of the ligand or protein is possible. For both docking poses, growth from the 5 and 6-positions looked feasible as the protein pocket was very open in the direction of these vectors. In addition to this, docking pose 2 gave the opportunity to grow from the 4-*N* nitrogen. For docking pose 1, growth from 4-*N* would disrupt the hydrogen bond interaction from the 1-*N* hydrogen to Glu146 (**Figure 111**). Before exploration of growth in a particular vector, fragment optimisation was required. This was carried out through a combination of SSS and synthesis.

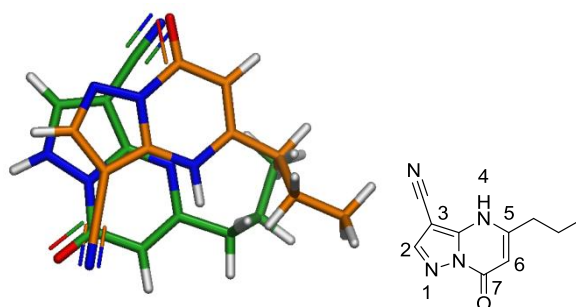


Figure 114 – Overlay of **docking pose 1** and **docking pose 2** of pyrazolopyrimidinone (**4.10**).

4.7.2 Fragment Optimisation – Single Point Changes

Optimisation of pyrazolopyrimidinone fragment (**4.10**) was performed prior to fragment growth. The fragment optimisation process was designed to better understand and improve the interactions between the ligand and protein. Single point changes were made to the fragment core to understand the importance of each atom to the protein-ligand interaction. This process was used to validate the proposed docking poses. Docking pose 1 of (**4.10**) made hydrogen bond interactions from the carbonyl oxygen and 1-*N* of the pyrazolopyrimidinone core. If this proposed docking mode was correct, moving the 1-*N* nitrogen to give (**4.13**), removal of the carbonyl to give analogue (**4.14**), or replacement to give analogue (**4.15**), was expected to be detrimental to the binding of the fragment. Also, conversion of the carbonyl oxygen to the NH₂ (**4.16**) would help determine whether the ligand needed to make these interactions through donors or acceptors (**Figure 115**). If binding pose 1 is correct, NMe analogue (**4.17**) would have been unlikely to bind even if a

hydrogen bond donor could interact with the carboxylic acid region. The additional methyl group in **(4.17)** was predicted to clash with the carboxylic acid region of the protein. If the binding mode is as predicted for docking pose 2, the single point changes in compounds **(4.13)** – **(4.16)** would not affect the ability of the nitrile moiety to hydrogen bond to the carboxylic acid region of the protein, therefore binding to CD38 would be expected. *Des*-carbonyl **(4.14)** and methyl analogue **(4.15)**, may lose some binding affinity as no hydrogen bond interaction can be made with Thr221. However, the binding affinity of **(4.14)** and **(4.15)** could be comparable to **(4.10)** as the energy gained from the hydrogen bond interaction may not compensate for the desolvation penalty that the additional polarity in **(4.10)** provides compared to **(4.14)** and **(4.15)**. If docking pose 2 is correct, *exo*-amino analogues **(4.16)** and **(4.17)** would have the ability to satisfy the hydrogen bond interactions with the carboxylic acid region of the CD38 protein through the nitrile moiety, and with Thr221 through the aniline nitrogen. It is hypothesised there would also be sufficient room for the methyl group of NMe **(4.17)** to be tolerated.

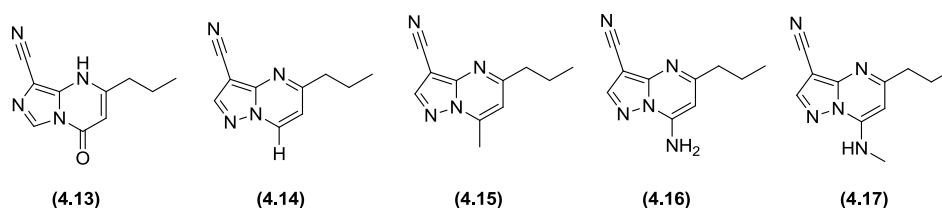
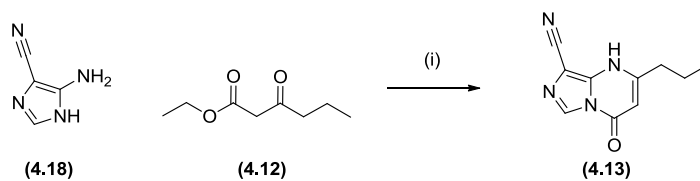


Figure 115 – Single point core changes to explore importance of 1-*N* nitrogen and carbonyl group.

Imidazopyrimidinone **(4.13)** was synthesised using analogous chemistry to pyrazolopyrimidinone **(4.10)**, through condensation with ethyl 3-oxohexanoate **(3.2)** under reflux in acetic acid and water. Compounds **(4.14)** – **(4.17)** were prepared via an external chemistry partner.

CONFIDENTIAL – Property of GSK – Do Not Copy



Scheme 20 – Synthesis of imidazopyrimidinone (**4.13**). Reagents and conditions: (i) acetic acid:water 2:1, 100 °C, 9.3%.

Compound No.	Structure	Human pIC ₅₀	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.10)		< 4.0	5.1	0.47	0.55
(4.13)		< 4.0	< 5.9		
(4.14)		< 4.0	< 4.0		
(4.15)		< 4.0	< 4.0		
(4.16)		< 4.0	< 4.0		
(4.17)		< 4.0	4.0 ^a	0.34	0.28

Table 7 – Mouse and human enzyme data for single point changes to pyrazolopyrimidinone (**4.10**). ^aThis compound has only been reported active in 2 out of 3 screening occasions – 3rd screening occasion measured < 4.48.

From the data presented in **Table 7**, compounds **(4.13)** – **(4.16)** were all inactive in both the mouse and human enzyme assays showing that they were less efficient starting points compared to pyrazolopyrimidinone **(4.10)**. Moving the 1-*N* nitrogen of pyrazolopyrimidinone, **(4.10)** to give imidazopyrimidinone **(4.13)** abolished the original mouse activity. Removal of the carbonyl group to give **(4.14)**, or replacement of the carbonyl with methyl **(4.15)** or amino substituent **(4.16)**, also removed all activity. *Des*-carbonyl analogue **(4.14)** and methyl analogue **(4.15)**, lacked the 1-*N* or 4-*N* proton of the 5,6 ring system. The proton on the 5,6-ring system is believed, from docking pose 1, to be important for binding. Amino analogues **(4.16)** and **(4.17)** also removed the proton from the 5,6-core, and changed the carbonyl hydrogen bond acceptor for a hydrogen bond donor. Switching of the donor and acceptor atoms in the pyrazolopyrimidinone core proved detrimental to binding of the fragments to the CD38 protein, despite NMe **(4.17)** providing borderline CD38 mouse biochemical activity. Although the biochemical data does not prove the binding mode, for these single point changes the data correlates closer with docking pose 1.

Both docking poses 1 and 2, showed the nitrile moiety acted as a hydrogen bond acceptor to the CD38 protein. To help further validate the docking poses, the nitrile group from HTS hit **(4.10)** was removed and replaced with a variety of small functional groups (**Figure 116**).

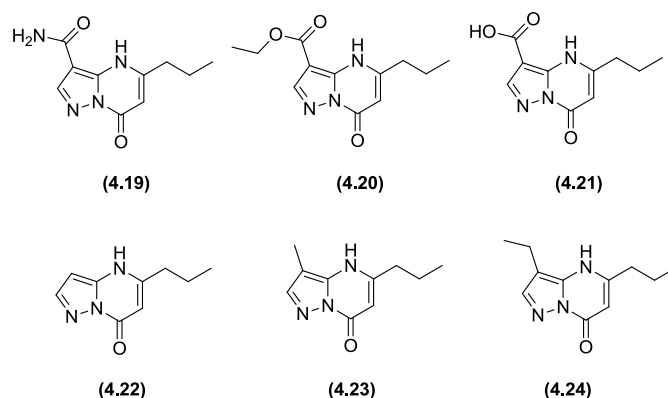


Figure 116 – Compounds designed to explore substitution at the 3-position of pyrazolopyrimidinone **(4.10)**.

Primary carboxamide (**4.19**) was a high priority compound. Overlay of docking pose 2 of nitrile (**4.10**), with quinoline carboxamide (**4.1**), showed that the nitrile could potentially be replaced with a primary carboxamide to form (**4.19**) (**Figure 117**). Docking of carboxamide (**4.19**) using MOE, confirmed that the carboxamide moiety of (**4.19**) could form hydrogen bond interactions to the carboxylic acid region of the CD38 protein (**Figure 118**), namely between the carboxamide NH₂ to Glu146 (NH – O 2.50 Å) and Asp155 (NH – O 2.92 Å). Two additional hydrogen bond interactions were made from (**4.19**) to the CD38 protein, one from the carboxamide carbonyl oxygen to Lys129 (O – HN 2.88 Å) and the other from the pyrazolopyrimidinone carbonyl oxygen to Thr221 (O – HO 3.29 Å). It is important to note that the tautomeric form of the pyrazolopyrimidinone core was important for the orientation of the carboxamide group. The docking was performed assuming the molecule existed in the 4-*H* tautomer, giving an intramolecular H-bond between the 4-*H* of the core and the amide carbonyl oxygen. The carboxylic acid (**4.21**) could potentially form similar interactions to the carboxylic acid region of the protein as the primary carboxamide, but ester (**4.20**) was expected to be too bulky to occupy the CD38 binding pocket in the docking pose 2 orientation. Replacement of the nitrile group with a proton (**4.22**) or an alkyl group (methyl (**4.23**) or ethyl (**4.24**)) would prevent the molecules from docking in the docking pose 2 orientation.

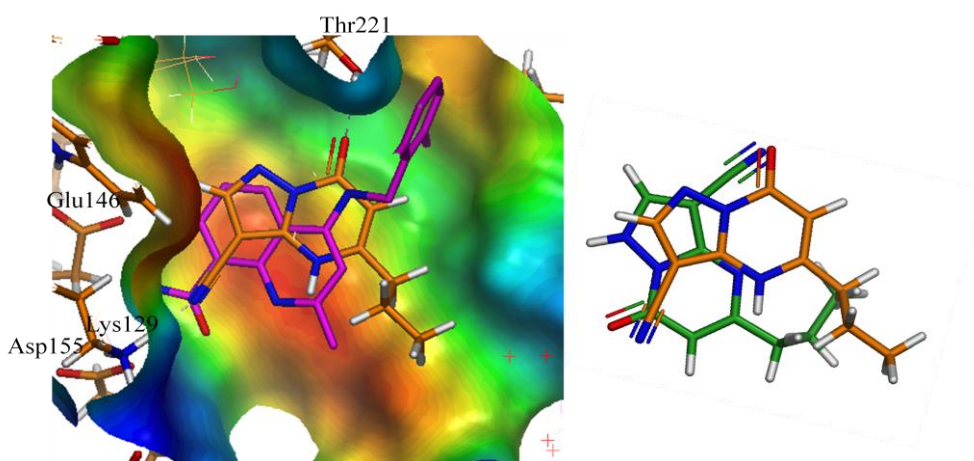


Figure 117 – Right: Overlay of quinoline carboxamide (**4.1**) in human CD38 (PDB: 1OCRI) with docking pose 2 of pyrazolopyrimidinone (**4.10**); **Left:** Overlay of docking pose 1 and docking pose 2 of pyrazolopyrimidinone (**4.10**).

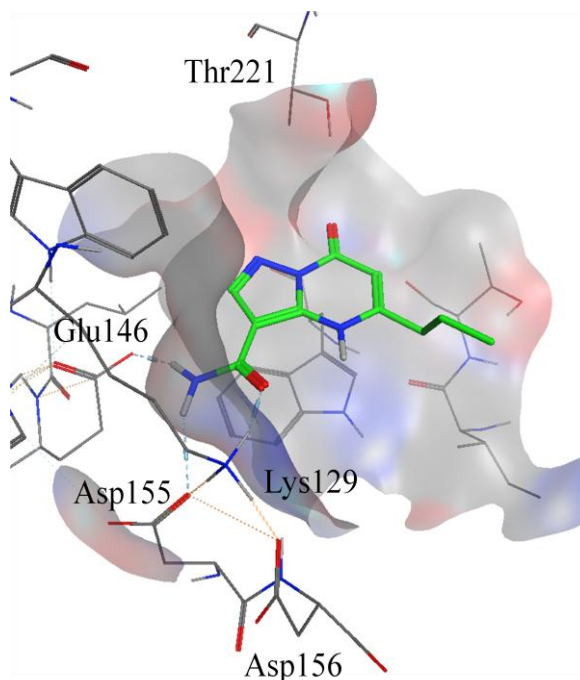
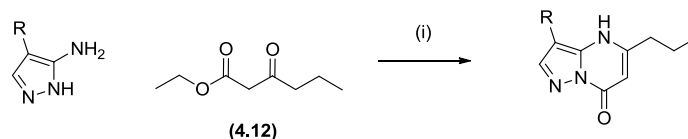


Figure 118 – Docking of amide analogue (**4.19**).

The nitrile replacement analogues shown in **Figure 116** all contain the pyrazolopyrimidinone carbonyl group and hydrogen on the nitrogen on the 5,6-heterocyclic core to interact with the carboxylic acid region of the protein (as in docked pose 1 (**4.10**)). The primary carboxamide (**4.19**), ester (**4.20**) and carboxylic acid (**4.21**) groups could potentially interact with Thr221 as in docking pose 2. Although sufficient space is observed in this region of the protein from docking, the removal of the nitrile group to afford (**4.22**) would show the importance of the nitrile moiety, but would not confirm that the nitrile was definitely participating in a hydrogen bond interaction to the protein. The nitrile group is electron-withdrawing, and removal of the group, or replacement with alkyl groups (Me (**4.23**), Et (**4.24**)) will result in the 5,6-pyrazolopyrimidinone ring system being more electron-rich. If the electronics of the 5,6-ring system play an important role in binding to the protein and π -stacking with Trp189, all activity may be abolished.

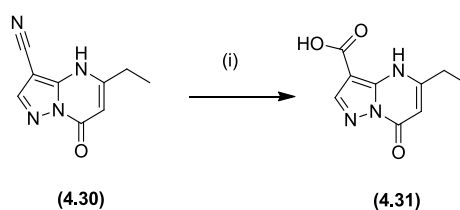
Exploration of nitrile replacements for pyrazolopyrimidinone (**4.10**) was carried out by condensation of the appropriate amino-pyrazoles with ethyl 3-oxohexanoate (**4.12**) in acetic acid and water. Yields ranged between 22 – 63% (**Scheme 21**).



R	Amino pyrazole No.	Product No.	Yield %
CONH ₂	(4.25)	(4.19)	62
CO ₂ Et	(4.26)	(4.20)	63
H	(4.27)	(4.22)	30
Me	(4.28)	(4.23)	43
Et	(4.29)	(4.24)	22

Scheme 21 – Synthesis of pyrazolopyrimidinone (**4.10**) analogues with various 3-substituents. Reagents and conditions: (i) acetic acid:water 2:1, 100 °C.

Prior to this programme of work, carboxylic acid (**4.31**) had been synthesised for a different programme in 75% yield from basic hydrolysis of nitrile (**4.30**) (**Scheme 22**). Although the ethyl-substituted carboxylic acid analogue (**4.31**) had a shorter alkyl chain compared to propyl (**4.10**), it was felt that (**4.31**) would establish whether replacement of the nitrile group with a carboxylic acid was feasible. If binding of (**4.31**) was observed, the corresponding propyl analogue would be synthesised.



Scheme 22 – Synthesis of carboxylic acid (**4.31**). Reagents and conditions: (i) water, 10 M NaOH, reflux, 75%.

Disappointingly, all nitrile replacements were inactive in both the mouse and human biochemical assays (**Table 8**). This suggests that the nitrile group must either be forming a hydrogen bonding interaction to the CD38 protein (both docking pose 1 and 2 support this), and/or be essential for the electronics of the pyrazolopyrimidinone core. Even conversion of the nitrile group to give amide (**4.19**), which appeared to be promising from a docking perspective (**Figure 118**), was inactive in both enzyme assays. For future chemistry, it appeared essential that the pyrazolopyrimidinone core and nitrile group remain fixed. As with all hydrogen bond interactions, the directionality is important in establishing the efficiency of the interaction; this is particularly important for the nitrile moiety. In relation to this, Turner has analysed the hydrogen bond distances and directionality from compounds within the Cambridge Structural Database (CSD).²¹⁶ This analysis showed that the mean hydrogen bond distance for a nitrile group is 2.211 Å (NH-NC), with the most common θ values ranging between 140 ° - 170 °.²¹⁶ Despite the hydrogen bond distances corresponding with the quoted value, the nitrile hydrogen bond angles show that docking pose 1 has a smaller angle than commonly observed, but docking pose 2 is in the expected range (docking pose 1 θ = 127.3 ° and docking pose 2 θ = 161.6 °). Although the hydrogen bond angle is not ideal for docking pose 1, protein backbone flexibility, which is frozen out in the docking procedure, could improve this situation.

To conclude, small single point changes to the core and pendant substituents have failed to identify more potent and efficient starting points. Although it is currently inconclusive, initial data suggests that the pyrazolopyrimidinone core is more likely to bind as proposed by docking pose 1. Further data is required to improve confidence in this hypothesis.

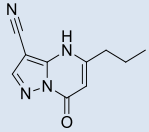
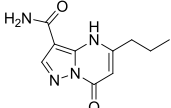
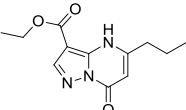
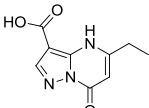
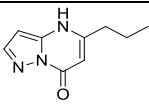
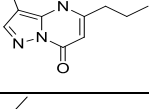
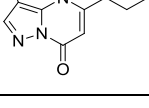
Compound No.	Structure	Human pIC ₅₀	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.10)		< 4.0	5.1	0.47	0.55
(4.19)		< 4.0	< 4.0		
(4.20)		< 4.0	< 4.0		
(4.21)		< 4.0 (n=1)	< 4.0 (n=1)		
(4.22)		< 4.0	< 4.0		
(4.23)		< 4.0	< 4.0		
(4.24)		< 4.0	< 4.0		

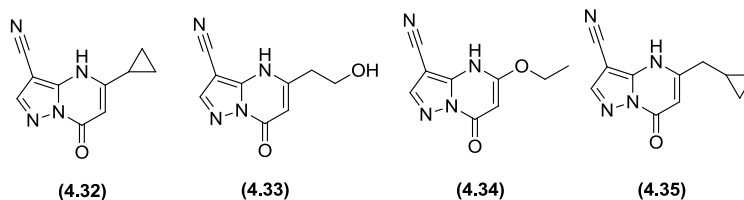
Table 8 - Mouse and human enzyme data for nitrile replacements of pyrazolopyrimidinone (**4.10**).

4.7.3 Exploration of 5-Position Substitution

The propyl group at the 5-position of pyrazolopyrimidinone (**4.10**) had been kept constant to understand the consequence of making changes to the pyrazolopyrimidinone core. It had been confirmed that the pyrazolopyrimidinone core and 3-nitrile were both critical for binding to the CD38 protein. The next set of compounds designed (**Figure 119**) focused on small changes around the 5-position substituent of (**4.10**), keeping the remainder of the molecule fixed. Some compounds were identified through SSS of our laboratories' compound collection, and others required synthesis. The set of compounds was designed to understand the

consequence of truncation of the propyl group, adding polarity to it, or growing it by a single heavy atom (**Figure 119**). The 5-substituent would occupy a similar position in docking poses 1 and 2 (**Figure 117**) and growth along this vector was expected to be beneficial in both binding orientations.

From SSS of our laboratories compound collection



Require synthesis

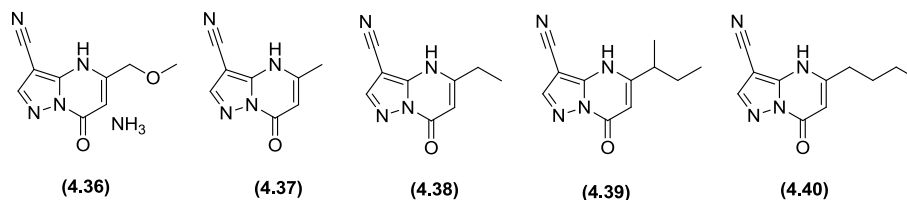
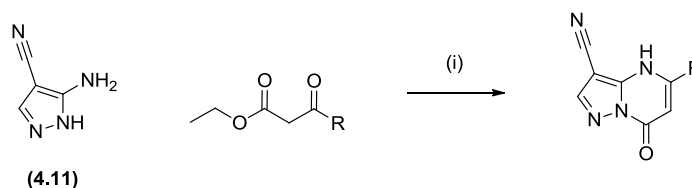


Figure 119 – Compounds designed to explore small changes to the 5-substituent on pyrazolopyrimidinone (**4.10**).

The compounds were prepared by condensation of commercially available pyrazolopyrimidinone (**4.11**) with the corresponding commercially available β -keto esters (**Scheme 23**). Encouragingly, biochemical screening against mouse and human constructs provided some enzyme activity (**Table 9**).



R	Beta-keto ester	Product No.	Yield %
CH ₂ OMe	(4.41)	(4.36b)	57
Me	(4.42)	(4.37)	12
Et	(4.43)	(4.38b)	20
sec-Bu	(4.44)	(4.39)	16
n-Bu	(4.45)	(4.40)	6.8

4.36b and 4.38b refers to the ammonium salt of 4.36 and 4.38, respectively

Scheme 23 – Synthesis of pyrazolopyrimidinone (**4.10**) analogues with various 5-substituents. Reagents and conditions: (i) acetic acid:water 2:1, 100 °C.

Truncating the propyl group of (**4.10**) to ethyl (**4.38b**) (h pIC₅₀ < 4.0; m pIC₅₀ = 4.2, LE = 0.41, LLE_(AT) = 0.55) and methyl (**4.37**) (h pIC₅₀ < 4.0; m pIC₅₀ = 4.1, LE = 0.43, LLE_(AT) = 0.62) gave a drop in mouse potency, with a slightly decreased LE suggesting that the propyl group is preferable in the 5-position (**Table 9**). Constraining the three carbon atoms into a cyclopropane ring (**4.32**) (h pIC₅₀ = 4.5, LE = 0.41, LLE_(AT) = 0.55; m pIC₅₀ = 6.8, LE = 0.62, LLE_(AT) = 0.76) gave an increase in mouse potency and efficiency, compared to (**4.10**), along with human activity. The cyclopropane ring has different properties in comparison to the propyl chain. Cyclopropane rings display partial π -character, as described by Walsh's orbital description,^{217,218} and is rigidified in comparison to the flexible propyl chain. From docking it is unclear why the π -character would increase the binding efficiency of the fragment to the protein, but does suggest that aromaticity may be favoured in this region (*vide infra*).

CONFIDENTIAL – Property of GSK – Do Not Copy

Compound No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.10)		< 4.0 (n=2)			5.1 (n=2)	0.47	0.55
(4.38b)		< 4.0			4.2	0.41	0.55
(4.37)		< 4.0			4.1	0.43	0.62
(4.32)		4.5	0.41	0.55	6.8	0.62	0.76
(4.33)		< 4.0			4.1	0.37	0.67
(4.36b)		< 4.0			< 4.5		
(4.34)		< 4.0			< 4.0		
(4.39)		< 4.0 (n=2)			5.5 (n=2)	0.47	0.52
(4.40)		< 4.0 (n=3)			4.5 (n=3)	0.39	0.43
(4.35)		< 4.0			< 4.0		

Table 9 - Mouse and human enzyme data for propyl replacements of pyrazolopyrimidinone (4.10). Data in this table is n=1 unless stated otherwise.

Adding polarity to the propyl chain of **(4.10)** is detrimental to potency as observed for oxygen analogues **(4.33)**, **(4.34)** and **(4.36b)**. Addition of an extra carbon to the propyl chain to form *sec*-butyl **(4.39)** (h pIC₅₀ < 4.0; m pIC₅₀ = 5.5 LE = 0.47, LLE_(AT) = 0.52) gave an increase in potency with comparable LE and LLE_(AT) to propyl **(4.10)**. LLE_(AT) was unchanged, but the increased fraction of sp³ centres (fsp³) close to the 5,6-heterocyclic could be favourable as this could lead to an increase in aqueous solubility by disrupting π -stacking interactions between individual molecules in the crystal lattice, increasing the aqueous solubility of the compound compared to propyl **(4.10)**. The *n*-butyl analogue **(4.40)**, was inactive in the human assay, and gave a drop in the mouse potency and efficiency compared to propyl **(4.10)**. The drop in potency and efficiency for *n*-butyl **(4.40)** is likely to involve entropic factors; the butyl chain is very flexible with four rotatable bonds giving a large number of possible molecular conformations. Extending the cyclopropane ring further out by one carbon **(4.33)**, removed activity against both mouse and human. The methylene linker between the cyclopropane and the pyrazolopyrimidinone core, in the lowest energy conformation, would place the cyclopropane ring out of the plane of the core (**Section 2.10.7**). The dihedral angle between the core and the cyclopropane is predicted to be approximately 90 ° from computational analysis of the lowest energy conformation. The CD38 binding pocket is a large flat shelf which is hypothesised to prefer planar molecules, maximising the interactions with the protein. This hypothesis fits with the biochemical SAR observed. Although the CD38 protein appears to have a preference for flatter molecules, care needs to be taken to maintain good aqueous solubility profile of at least 100 μ g/mL (CLND). Despite some promising efficacy developments from this series, data for compounds with small aliphatic groups in the 5-position (**Table 9**) could not help validate either docking pose as the groups will be occupying a similar area of space in a large open region of the CD38 pocket for both poses.

The effect of 2 and 6-position substitution of pyrazolopyrimidinone **(4.10)** was explored before deciding which vector, or vectors, to grow from to gain additional

binding potency. Docking of propyl (**4.10**) suggests that 6-Me would be tolerated in both docking poses (**Figure 120**). Substitution in the 2-position would not be tolerated for docking pose 2 as the substituent would clash with the protein surface. Substitution from the 2-position in docking pose 1 looked feasible, but the pocket in this region is very tight, increasing the risk of the substituent failing to fit. The molecules were synthesised from commercially available intermediates using the condensation chemistry used to prepare previous analogues (**Scheme 24**).

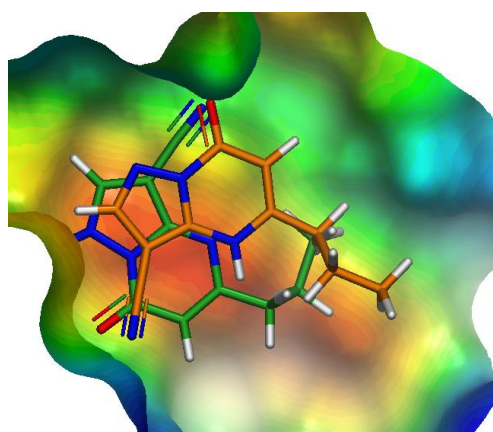
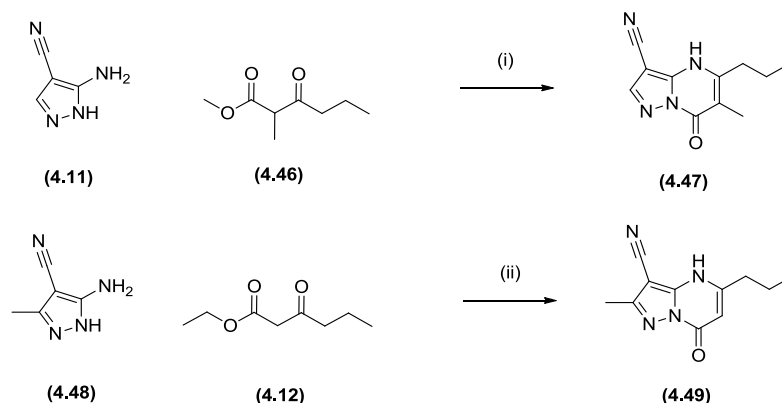


Figure 120 - Overlay of **docking pose 1** and **docking pose 2** of pyrazolopyrimidinone (**4.10**).



Scheme 24 – Synthesis of pyrazolopyrimidinone (**4.10**) analogues to explore 2 and 6 position substitution. Reagents and conditions: (i) acetic acid:water 2:1, 100 °C, 33%; (ii) acetic acid:water 2:1, 100 °C, 51%.

6-Methyl (**4.47**), 2-methyl (**4.49**), and 2-amino (from our laboratory's compound collection) (**4.50**), were all inactive in the human assay (**Table 10**). 2-Substitution was detrimental to mouse activity for both the 2-methyl (**4.49**) and the 2-amino (**4.50**) compounds. This was disappointing, as docking pose 1 (**Figure 120**) for (**4.10**) would place the 2-methyl substituent in the substrate pocket, which was hoped to give improved binding of the fragments to CD38. A reduction in efficiencies for 6-methyl (**4.47**) compared to the original HTS hit (**4.10**) was observed. Despite this, the LE and LLE_(AT) still remained greater than 0.3 suggesting that disubstitution from the pyrazolopyrimidinone core may be tolerated. Based on this observation it was decided to investigate this hypothesis further.

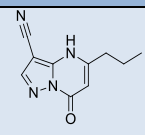
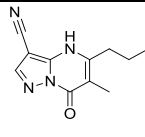
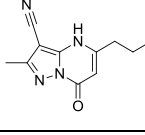
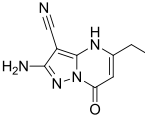
Cmpd No.	Structure	Human pIC ₅₀	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.10)		< 4.0	5.1	0.47	0.55
6-position					
(4.47)		< 4.0	5.0	0.43	0.48
2-position					
(4.49)		< 4.0	< 4.0		
(4.50)		< 4.0 (n=1)	< 4.0 (n=1)		

Table 10 - Mouse and human enzyme data for 2 and 6 substituted analogues of pyrazolopyrimidinone (**4.10**).

Unfortunately, despite continued efforts, no X-ray crystal structure of a pyrazolopyrimidinone analogues of (**4.10**) could be obtained. To further investigate

the hypothesis around growing from the 6-position of the pyrazolopyrimidinone template, consideration once again turned to which of the two docking poses was most representative of the binding pose of **(4.10)**. Although the SAR was not conclusive, docking pose 1 could rationalise most of the SAR (particularly for the single point changes carried out at the beginning of this research programme) (**Table 7** and **Table 8**). Careful design of the next iterations of compounds would help to further validate the preferred docking pose. Exploration of the template has shown that the 5- and 6-position growth of the pyrazolopyrimidinone core was feasible for binding to CD38. These two vectors will be explored in the next section of this research programme.

4.7.4 Simultaneous 5- and 6-Position Substitution

Exploration of 5,6-substituted pyrazolopyrimidinone analogues was carried out through SSS of our laboratory's compound collection. Compounds with small groups at both vectors were available for screening **(4.51) – (4.58)**. Simultaneous 5,6-substitution was predicted to be tolerated in both docking poses 1 and 2.

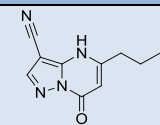
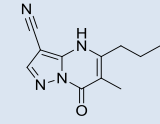
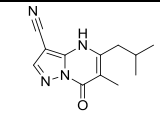
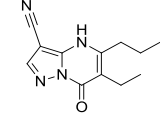
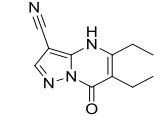
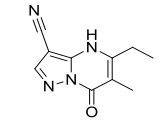
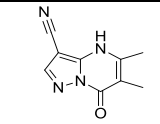
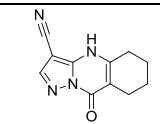
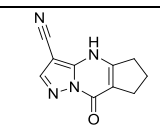
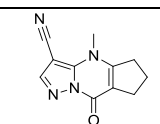
Cmpd No.	Structure	Human pIC₅₀	Mouse pIC₅₀	Mouse LE	Mouse LLE_(AT)
(4.10)		< 4.0	5.1	0.47	0.55
(4.47)		< 4.0	5.0	0.43	0.48
(4.51)		< 4.0	< 4.0		
(4.52)		< 4.0	4.6	0.37	0.38
(4.53)		< 4.0	< 5.0		
(4.54)		< 4.0	5.4	0.49	0.58
(4.55)		< 4.0	5.4	0.53	0.67
(4.56)		< 4.0	4.5	0.39	0.47
(4.57)		< 4.0	4.3	0.39	0.53
(4.58)		< 4.0	< 4.0		

Table 11 - Mouse and human enzyme data for 5,6-substituted analogues of pyrazolopyrimidinone (**4.10**). The data in this table is n=1.

No human activity was observed for compounds (4.51) – (4.58), and the mouse enzyme activity was very similar across the set. The most potent and efficient CD38 binders had a methyl substituent at the 6-position and ethyl (4.54) or methyl (4.55) at the 5-position. Disubstitution with small 5,6-substituents was well tolerated, but as larger groups were introduced (5-*i*Bu, 6-Me (4.51); or 5-Et, 6-Et (4.53)) the mouse enzyme activity was reduced. Having groups in both the 5- and 6-position simultaneously would affect the conformation of the groups due to steric repulsion. Tricyclic analogues (4.56) and (4.57) were inactive against human and equipotent within experimental error against mouse. The tricyclic cores (4.56) and (4.57) were both less active and efficient compared to propyl (4.10) and were likely to be less soluble due to the more planar structure giving rise to π -stacking between copies of the molecule. 4-*N* Alkylation of cyclic analogue (4.57) to give *N*-methyl (4.58) was not tolerated and the compound was inactive against both orthologues. This observation fits in with docking pose 1 (Figure 111), where the hydrogen on the pyrazolopyrimidinone core formed a hydrogen bond interaction with Glu146 from the carboxylic acid region. If binding of the pyrazolopyrimidinone series was as docking pose 2, *N*-methyl (4.58) would be expected to be tolerated in the protein pocket.

4.7.5 Exploration of the 6-Position of Pyrazolopyrimidinone (4.10)

Screening following SSS of our laboratories' compound collection concluded that simultaneous growth in the 5- and 6- positions of the pyrazolopyrimidinone core was unlikely to be beneficial for CD38 binding. Efficient CD38 binding could be obtained by growth from the 5-position, but a substituent in the 6-position was likely to affect the binding orientation of the 5-substituent. Before deciding to focus on growth in the 5-position, exploration of growth from the 6-position in the absence of a 5-substituent was required. Review of our laboratories' compound collection established that groups with small 6-substituents were readily available for screening (4.59) – (4.64). Small substituents, methyl (4.59), ethyl (4.60) and methoxy (4.64), showed activity in the mouse enzyme assay, but were less potent than the original HTS hit (4.10).

It had been demonstrated that growth from the 6-position was tolerated, but gave less efficient analogues compared to propyl (**4.10**). Small 6-substituents were tolerated but as the substituent was increased in size (larger than ethyl), mouse biochemical activity was lost. This observation fits with both docking pose 1 and 2 as a clash with the protein wall would be predicted for larger substituents. Focus returned to optimising the 5-substituent, which had given some promise when explored with small substituents (**Table 9**). Although small substituents (methyl and ethyl) were tolerated at the 6-position with a 5-substituent present, it was decided to optimise the 5-position in the absence of such a group, and then introduce the 6-substituent later, if this was deemed to be necessary. This strategy was adopted because the presence of a 6-substituent may prevent the 5-substituents from occupying the most preferred orientation.

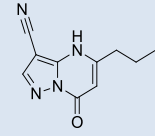
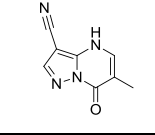
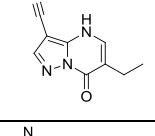
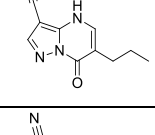
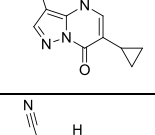
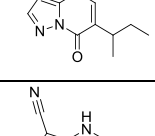
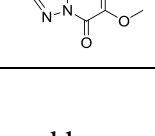
Cmpd No.	Structure	Human pIC ₅₀	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.10)		< 4.0 (n=2)	5.1 (n=2)	0.47	0.55
(4.59)		< 4.0	4.6	0.48	0.68
(4.60)		< 4.0	4.5	0.44	0.57
(4.61)		< 4.0	< 4.0		
(4.62)		< 4.0	< 4.0		
(4.63)		< 4.0 (n=2)	< 4.0 (n=2)		
(4.64)		< 4.0	4.7	0.46	0.70

Table 12 - Mouse and human enzyme data for 6-substituted analogues of pyrazolopyrimidinone (**4.10**). Data is n=1 unless stated otherwise.

4.7.6 Exploration of the 5-Position of Pyrazolopyrimidinone (**4.10**)

Docking poses 1 and 2 suggest that substitution from the 5-position would point into a large open area of the protein pocket, giving room for a larger substituent than the small alkyl groups which have previously been explored. Investigation of small changes to the propyl of HTS hit (**4.10**) had identified cyclopropane analogue (**4.32**) (h pIC₅₀ = 4.5, LE = 0.41, LLE_(AT) = 0.55; m pIC₅₀ = 6.8, LE = 0.62, LLE_(AT) = 0.76) which demonstrated both mouse and human enzyme activity. The cyclopropane ring has partial π -character associated with it, suggesting that in this region, aromaticity

may be beneficial for potency. From docking, it is unclear why π -character would be favoured in this region of the protein, but there is a possibility that a CH- π interaction is being formed between the cyclopropane ring and Ile185 which sits below the molecule. To explore this hypothesis, a second SSS iteration of our laboratories compound collection was performed. The SSS focused on compounds with aromatic groups in the 5-position. These compounds were excluded from the initial SSS due to concerns that the compounds were made up from three aromatic rings directly linked. This arrangement of rings could make the compounds planar, giving rise to π -stacking between copies of the molecule, thereby limiting solubility. Despite these concerns, exploration of 5-position aromaticity had been identified as the most promising approach to improve binding affinity of the fragments to CD38.

The SSS and resultant screening yielded data for compounds **(4.65) – (4.80)** (**Table 13**). A molecular weight cut-off of 300 Da was applied, as well as limiting the number of aromatic rings to three. Mouse and human enzyme screening provided a range of results. As hypothesised, aromaticity at the 5-position was tolerated with phenyl **(4.65)** (h pIC₅₀ = 4.2, LE = 0.32, LLE_(AT) = 0.37; m pIC₅₀ = 5.7, LE = 0.43, LLE_(AT) = 0.48) having activity against both the mouse and human enzymes. Phenyl **(4.65)** was less potent than cyclopropane **(4.32)** and therefore less efficient. Introduction of polarity into the phenyl ring gave 2- and 3-pyridyl analogues, respectively, **(4.66)** and **(4.67)**. 3-Pyridyl **(4.67)** gave mouse activity (m pIC₅₀ = 5.1, LE = 0.39, LLE_(AT) = 0.55) but no human activity. 2-Pyridyl **(4.66)** was inactive against both orthologues.

Thiophene rings are often used in medicinal chemistry as a phenyl isostere.²¹⁹ The lone pair of electrons from the sulfur are delocalised through the system making it aromatic. The long C-S bonds of 1.7 Å, compared to the C-C bonds in benzene (1.4 Å), combined with the electron density cloud on sulfur make thiophene and benzene close in relative size. On enzymatic screening, thiophene **(4.68)** proved to be the most potent and efficient aryl-analogue (h pIC₅₀ = 5.5, LE = 0.44, LLE_(AT) = 0.52; m pIC₅₀ = 6.9, LE = 0.56, LLE_(AT) = 0.63). Suitable thiophene analogues were not

CONFIDENTIAL – Property of GSK – Do Not Copy

available but analogous phenyl analogues were screened in the enzyme assays. Phenyl substitution in the *ortho*-position was not tolerated, (4.69) and (4.70), but *m*- (4.71) and *p*-phenyl substitution ((4.73) and (4.74)) gave a possible vector for growth. As part of the SSS, benzylic and phenethyl 5-substitution were explored (compounds (4.75) – (4.80)). All five analogues screened were inactive against the human enzyme, with four of the five showing low levels of mouse activity.

CONFIDENTIAL – Property of GSK – Do Not Copy

Cmpd No.	Structure	Human pIC₅₀	Human LE	Human LLE_(AT)	Mouse pIC₅₀	Mouse LE	Mouse LLE_(AT)
(4.10)		< 4.0 (n=2)			5.1 (n=2)	0.47	0.55
(4.32)		4.5	0.41	0.55	6.8	0.62	0.76
(4.65)		4.2	0.32	0.37	5.7	0.43	0.48
(4.66)		< 4.0 (n=2)			< 4.5 (n=2)		
(4.67)		< 4.0			5.1	0.39	0.55
(4.68)		5.5	0.44	0.52	6.9	0.56	0.63
(4.69)		< 4.0			< 4.5		
(4.70)		< 4.0			4.5	0.32	0.34
(4.71)		5.0	0.36	0.38	6.4	0.46	0.48
(4.72)		< 4.0			5.1	0.37	0.37
(4.73)		4.4	0.32	0.33	5.8	0.42	0.43
(4.74)		4.7	0.34	0.34	5.7	0.41	0.41
(4.75)		< 4.0			5.5	0.40	0.45

Table continued on next page

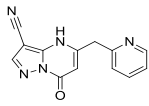
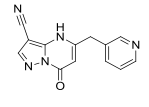
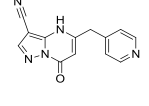
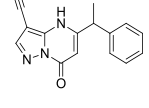
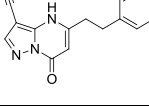
(4.76)		< 4.0			4.3	0.31	0.47
(4.77)		< 4.0			< 4.0		
(4.78)		< 4.0			4.6	0.33	0.49
(4.79)		< 4.0			5.7	0.39	0.42
(4.80)		< 4.0			4.7	0.32	0.35

Table 13 - Mouse and human enzyme data for 5-substituted aryl analogues of pyrazolopyrimidinone (**4.10**). Data is n=1 unless stated otherwise.

Analysis of the 5-aryl analogues of (**4.10**) (**Table 13**) highlighted that directly attached aryl groups to the core, with *m*- or *p*-substituents could give improved ligand-CD38 binding. Docking of the *m*-toluene analogue (**4.71**) suggested that the most likely binding orientation is docking pose 1 for HTS hit (**4.10**), picking up hydrogen bond interactions with Thr221 and Glu146 (**Figure 122**). Toluene (**4.71**) was unable to bind in docking pose 2 without significant movement of the compound to avoid a clash with the protein surface (Thr186) (**Figure 122**). Docking pose 1 showed that *m*-toluene analogue (**4.71**) filled the CD38 pocket well, but had space in the *p*-position which is near to Arg158 (mouse) or Thr158 (human). Compounds were designed to identify alternative *p*-substituents that could potentially interact with the Arg or Thr residues ((**4.81**) and (**4.82**)), and to assess the effect of simultaneous *m*- and *p*-substitution (**4.83**) (**Figure 121**). Disubstituted phenyl compound (**4.83**), was available in our laboratory's compound collection, but (**4.81**) and (**4.82**) required synthesis from commercially available starting materials (**Scheme 25**). The same synthetic procedure was used as for previous

pyrazolopyrimidinone analogues, but the isolated yields were low. This reduced reaction efficiency was believed to be contributed to by the poor solubility observed when handling these compounds.

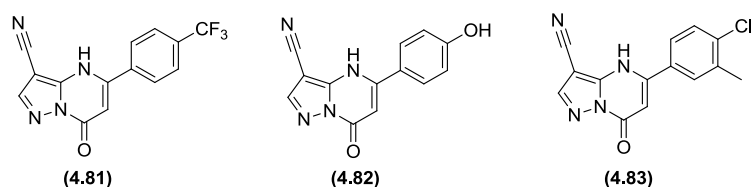


Figure 121 – Compounds designed to explore *p*-substitution of 5-phenyl analogues.

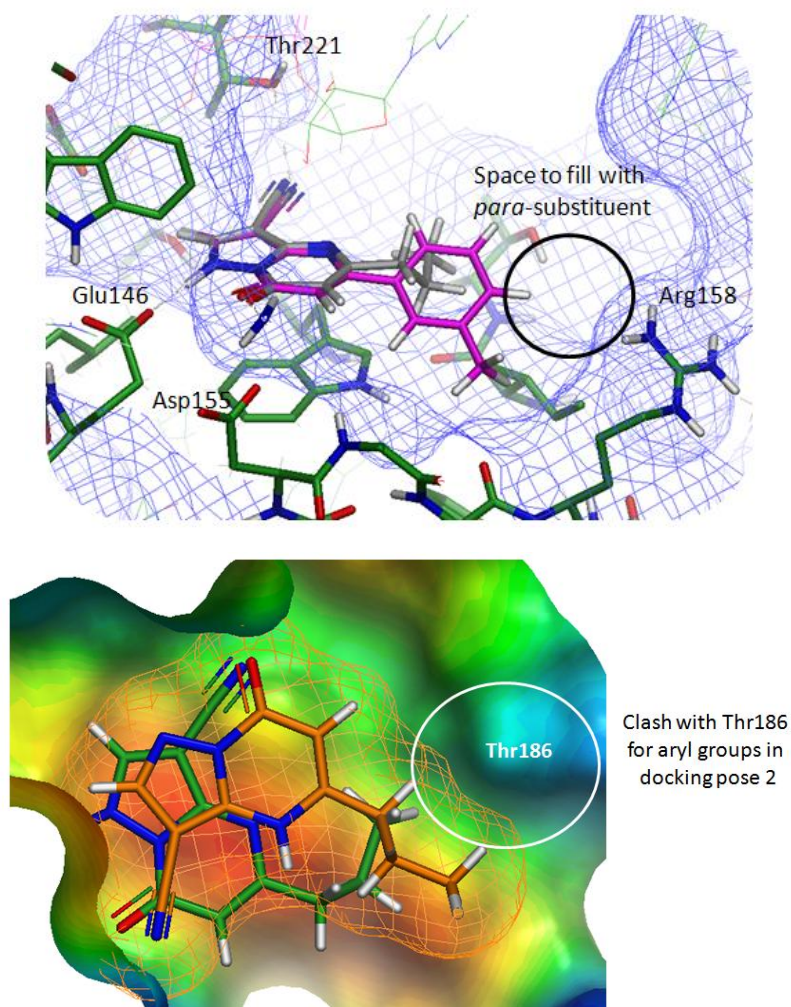
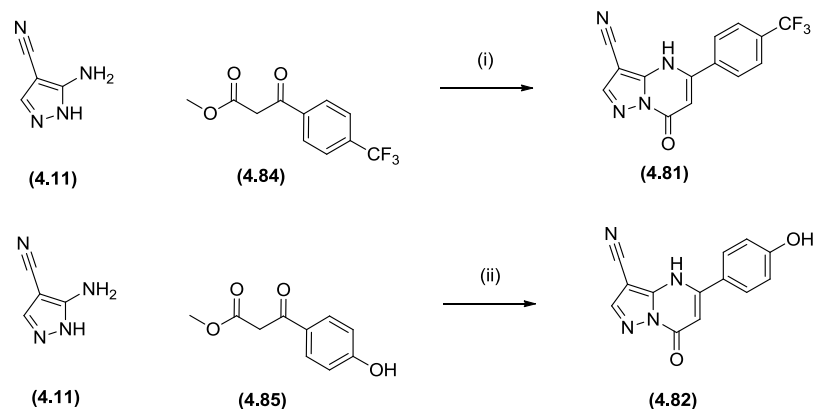


Figure 122 – **Top image:** Docking of *m*-toluene analogue (4.71) in CD38 (mouse) overlaid with docking pose 1 of HTS hit (4.10); **Bottom image:** Overlay of docking

pose 1 and docking pose 2 of pyrazolopyrimidinone (**4.10**). The surface of (**4.10**) docking pose 2 shows good shape complementarity with the area of protein by Thr186.



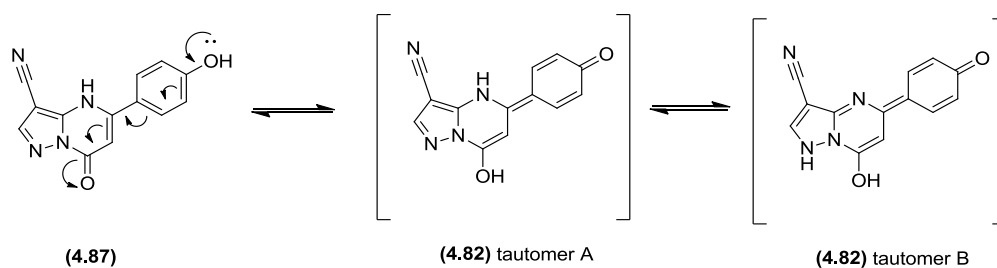
Scheme 25 – Synthesis of *p*-substituted 5-aryl analogues of HTS hit (**4.10**).

Reagents and conditions: (i) acetic acid:water 2:1, 100 °C, 7%; (ii) acetic acid:water 2:1, 100 °C, 9%.

Disappointingly, trifluoromethyl (**4.81**) was inactive in the mouse assay and displayed no human potency benefit compared to methyl (**4.73**) (**Table 14**). The drop in mouse potency could be rationalised by the additional electron density of the trifluoromethyl group clashing with Arg158. Phenol (**4.82**) and di-substituted analogue (**4.83**), were inactive against both mouse and human. It was reasoned that the phenol group in (**4.82**), being mildly acidic, would have an attractive interaction with Arg158 giving improved binding (assuming docking pose 1 is correct). A possible hypothesis for the reduction in CD38 binding for phenol (**4.82**) could be due to the electron-donating nature of the phenol influencing the tautomeric form of the core (**Scheme 26**). The increased phenolic nature of the carbonyl caused by donation of electrons through the π -system, could disrupt the hydrogen bond acceptor ability of the pyrazolopyrimidinone series. From docking it was unclear why the disubstituted analogue (**4.83**) would not bind in the CD38 pocket, which gave concern over the validity of the proposed docking pose.

Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.10)		< 4.0			5.1	0.47	0.55
(4.74)		4.7 (n=1)	0.34	0.34	5.7 (n=1)	0.41	0.41
(4.73)		4.4 (n=1)	0.32	0.33	5.8 (n=1)	0.42	0.43
(4.81)		4.3	0.27	0.27	< 5.0		
(4.82)		< 4.0			< 4.5		
(4.83)		< 4.0			< 4.5		

Table 14 – Biochemical data for *para*-substituted phenyl analogues.



Scheme 26 – The electron-donating nature the phenol in (4.82) could potentially influence the tautomeric form of the pyrazolopyrimidinone core.

From the limited number of *p*-substituted analogues synthesised, it became apparent that substitution in this area was not well tolerated. Despite the raised concerns over

the validity of the docking pose for *m*-toluene (**4.71**) (See below; **Figure 123**), *m*-substituted phenyl analogues were explored. Exploration of the *m*-substituted analogues would provide further insight into the validity of the proposed binding mode. The *m*-methyl group fills the CD38 pocket well in the proposed docking mode, but places the hydrophobic group in close proximity to the backbone carbonyl of Leu157 (CH₂-H - O 2.98 Å) (**Figure 123**). Accordingly, it was proposed that replacement of the methyl group with a hydrogen bond donor could improve binding efficiency of the pyrazolopyrimidinones to CD38. From the docking pose (**Figure 123**) it appeared feasible to substitute simultaneously in both *m*-positions. Four compounds, (**4.91**) – (**4.94**), were synthesised to explore these proposals.

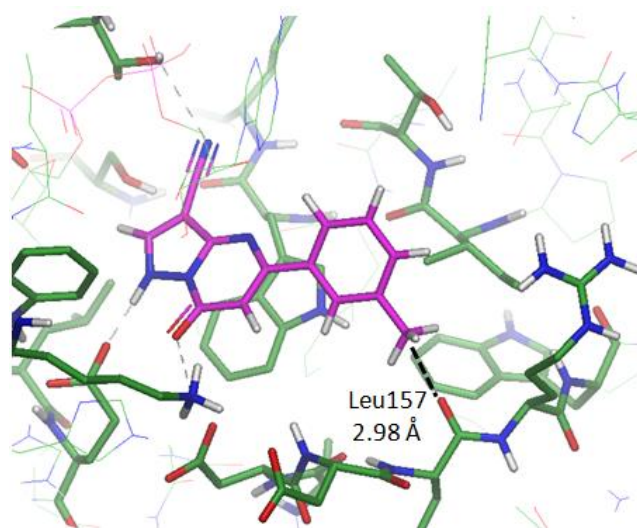
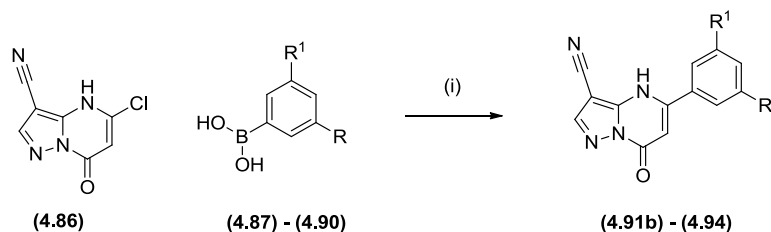


Figure 123 – Docking of *m*-toluene analogue (**4.71**).

Aryl chloride (**4.86**) was available in our laboratories' compound collection. Suzuki chemistry was used to couple the appropriate arylboronic acids to provide analogues (**4.91**) – (**4.94**) (**Scheme 27**). Previous attempts on using the pyrazolopyrimidinone template in metal-mediated reactions had proved challenging due to the metal chelating ability of the carbonyl and 1-*N* nitrogen.²²⁰ From previous experience, the use of K₂CO₃ as the base, SPhos as the ligand, and Pd(OAc)₂ as the catalyst, gave the best results when heating at 100 °C under microwave irradiation. Due to the electron-rich nature of the aryl chloride (**4.86**), SPhos was the ligand of choice as

Buchwald has reported that it is remarkably efficient for cross coupling of unactivated aryl chlorides.²²¹ The same paper stated that the ligand exhibited excellent substrate scope, improving the chance of success for couplings of aryl chloride (**4.86**).



Boronic acid	mol% SPhos	mol% Pd(OAc) ₂	R	R ¹	Product No.	Yield %
(4.87)	20	16	Me	Me	(4.91b) ^a	NA*
(4.88)	20	16	F	Me	(4.92)	23
(4.89)	20	16	F	Cl	(4.93)	8
(4.90)	20	24	OH	H	(4.94)	8

^a Isolated as ammonium salt

* Only part of the crude product was purified

Scheme 27 – Suzuki reactions to afford substituted phenyl analogues (**4.91b**) – (**4.94**). Reagents and conditions: (i) K₂CO₃, SPhos, Pd(OAc)₂, 100 °C, microwave.

Synthesis of dimethylphenyl (**4.91b**) resulted in sufficient quantity of compound for screening, but was low yielding. The low yield was attributed to potential chelation of the metal between the carbonyl oxygen and pyrazole nitrogen of the pyrazolopyrimidinone core, along with poor solubility of the crude mixture, which hampered extraction and purification. Disubstituted analogues (**4.92**) and (**4.93**) were prepared *via* the same procedure as for (**4.91b**), where the reagents were combined, the reaction mixture was purged with nitrogen, and the vial was sealed prior to heating in the microwave at 100 °C. Higher yields were achieved for disubstituted analogues (**4.92**) and (**4.93**), compared to dimethylphenyl (**4.91b**), which could be explained by the change in purification conditions to acidic rather than basic modifiers in the mobile phase of the preparative HPLC system. Although the methodology afforded the desired analogues, the low yields were concerning, particularly for synthesis of phenol (**4.94**). It was anticipated that the use of phenol

boronic acid (**4.90**) may reduce the yield further due to chelating effects of the phenol to the palladium. To minimise any effects from chelation, the catalyst was pre-activated with the ligand, before addition of the aryl chloride (**4.86**), boronic acid (**4.90**) and base. Purification using silica chromatography afforded phenol (**4.94**) in 8% yield.

Biochemical data for substituted phenyl analogues (**4.91b**) – (**4.94**) was disappointing (**Table 15**). The compounds were all inactive against both mouse and human enzymes, apart from the dimethylphenyl analogue (**4.91b**), which showed weak mouse activity ($m\text{ pIC}_{50} = 5.1$, $LE = 0.35$, $LLE_{(AT)} = 0.34$). This additional data brings into question the validity of the docking pose of *m*-toluene analogue (**4.71**) (**Figure 123**), as the docking pose predicted that disubstitution in both *m*-positions should be tolerated.

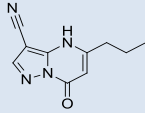
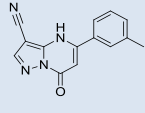
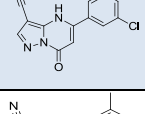
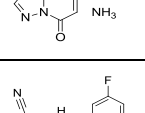
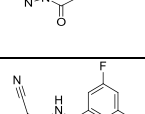
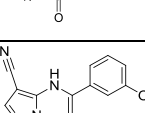
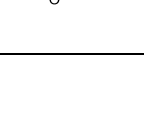
Cmpd No.	Structure	Human pIC₅₀	Human LE	Human LLE_(AT)	Mouse pIC₅₀	Mouse LE	Mouse LLE_(AT)
(4.10)		< 4.0			5.1	0.47	0.55
(4.71)		5.0 (n=1)	0.36	0.38	6.4 (n=1)	0.46	0.48
(4.72)		< 4.0 (n=1)			5.1 (n=1)	0.37	0.37
(4.91b)		< 4.5			5.1 ^a	0.34	0.35
(4.92)		< 4.0			< 5.0		
(4.93)		< 4.0			< 4.5		
(4.94)		< 4.5			< 5.0		

Table 15 – Biochemical data for *m*-substituted phenyl analogues (4.91b) – (4.94).

^aThis compound has only been reported active in 1 of the 4 screening occasions.

In light of these results for the disubstituted phenyl analogues, synthesis of monosubstituted *m*-phenyl analogues of (4.71) was executed. Fragment-based drug design usually focuses on making small incremental changes to the fragment in an attempt to pick up additional ligand efficient interactions, and this process relies on a crystal structure or robust docking model. For pyrazolopyrimidinone (4.71), no X-ray crystal structure could be obtained and the proposed docking pose was in question as the emerging biochemical SAR profile did not fully align with that being proposed. The lack of a ligand-protein X-ray crystal structure, or validated docking

pose, made it difficult to design specific analogues. As a result, the *m*-position was explored by identifying:

- 1) additional compounds from our laboratories' compound collection with larger groups in the *meta*-position (**Figure 124**); and
- 2) identifying commercially available β -keto esters to condense with aminopyrazole (**4.11**), and boronic acids or boronate esters from which pyrazolopyrimidinone analogues of (**4.71**) could be synthesised *via* Suzuki methodology (**Scheme 28** and **Scheme 29**).

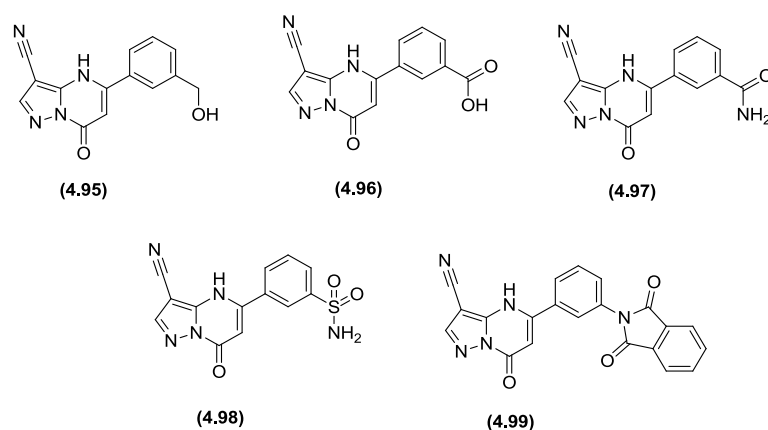
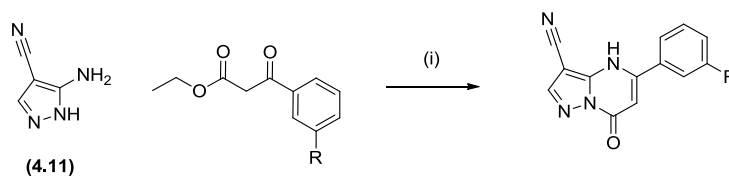


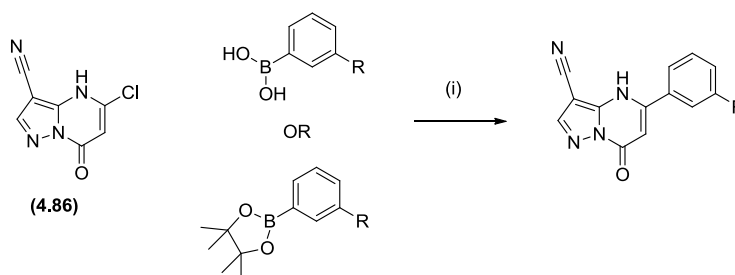
Figure 124 – Compounds identified from our laboratories' compound collection to explore the effect of larger substituents at the *m*-position.

The requisite β -Keto esters and boronic acids/boronate esters are readily available. Having stated this, a focused set of monomers was chosen to explore how varying the electronic properties in the *m*-position of (**4.71**) affected the binding affinity, alongside varying the steric bulk in this region.



R	Beta-keto ester	Product No.	Yield %
Br	(4.100)	(4.103)	15
OMe	(4.101)	(4.104)	5.1
CF ₃	(4.102)	(4.105)	31

Scheme 28 – Synthesis of pyrazolopyrimidinone (**4.71**) analogues to explore *m*-substitution. Reagents and conditions: (i) acetic acid:water 2:1, 100 °C.



Boron species	R	Product No.	Yield %
(4.106) boronic acid	<i>i</i> -Pr	(4.111)	32
(4.107) boronate ester	CO ₂ Et	(4.112) ^a	N/A ^a
(4.108) boronate ester	CH ₂ NHBoc	(4.113b) ^b	32
(4.109) boronic acid	C(O)NH <i>i</i> -Bu	(4.114)	47
(4.110) boronate ester	C(O)N(morpholine)	(4.115b) ^{a,b}	N/A ^a

^a not all sample purified - only 100 mg of crude product was purified

^b Isolated as ammonium salt

Scheme 29 - Suzuki reactions to afford *m*-substituted phenyl analogues (**4.111**) – (**4.115b**). Reagents and conditions: (i) K₂CO₃, SPhos (19 mol%), Pd(OAc)₂ (16 mol%), water, MeCN, 100 °C.

Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.10)		< 4.0			5.1	0.47	0.55
(4.71)		5.0 (n=1)	0.36	0.38	6.4 (n=1)	0.46	0.48
(4.103)		4.4	0.32	0.31	6.2	0.45	0.44
(4.104)		4.1 ^a	0.28	0.34	5.8	0.40	0.46
(4.105)		4.4	0.27	0.28	6.3	0.39	0.40
(4.95)		4.1	0.28	0.41	5.5 ^b	0.38	0.50
(4.111)		< 4.0			< 5.0		
(4.96)		4.1 (n=1)	0.27	0.34	5.6 (n=1)	0.37	0.44
(4.112)		< 4.0			4.2 ^c	0.25	0.28
(4.97)		< 4.0 (n=1)			5.5 (n=1)	0.36	0.51
(4.98)		< 4.0 (n=1)			< 5.4 (n=1)		
(4.113b)		< 4.5			< 5.0		
(4.114)		< 4.0			< 4.0		
(4.115b)		< 4.5			< 5.0		
(4.99)		5.8 (n=1)	0.27	0.33	6.5 (n=1) *r=6.1 (n=1)	0.31 *r=0.29	0.37 *r=0.35

Table 16 – Biochemical data for the exploration of the *m*-position of pyrazolopyrimidinone (4.71). * r = rat biochemical pIC₅₀. Rat biochemical data was for the ammonium salt of (4.71). ^aThis compound has only been reported active in 1 of the 4

screening occasions. ^bThis compound has only been reported active in 1 of the 3 screening occasions.

^aThis compound has only been reported active in 1 of the 4 screening occasions.

Human and mouse biochemical data for the additional *m*-substituted analogues failed to provide compounds which showed an increase in potency, whilst maintaining the ligand efficiency, compared to *m*-toluene analogue (**4.71**). A variety of small substituents (all slightly larger than methyl) with varying electronics (Br (**4.103**); OMe (**4.104**); CF₃ (**4.105**)) appeared to be tolerated giving similar potencies to analogue (**4.71**). Growth of the *m*-substituent resulting in additional steric bulk compared to methyl analogue (**4.71**) was detrimental to both human and mouse activity. It is clear that directly attached amides, esters and sulfonamides (with the carbonyl/sulfone group attached to the phenyl ring) diminished activity and efficiency, especially with larger substituents pendant from these units. An interesting result was achieved for phthalimide (**4.99**) (h pIC₅₀ = 5.8, LE = 0.27, LLE_(AT) = 0.33; m pIC₅₀ = 6.5, LE = 0.31, LLE_(AT) = 0.37). Both human and mouse potency had been increased, but a drop in LE and LLE_(AT) was observed due to the number of additional heavy atoms that had been added and the lipophilic nature of the phthalimide group. Despite the ligand efficiencies being reduced, the phthalimide group did demonstrate that there was a substantial amount of room in the protein pocket to grow, assuming the binding mode had not changed. Building from the phthalimide ring was not a valid option as phthalimide is particularly prone to hydrolysis giving subsequent ring opening, providing developability concerns.²²² As confidence was still being built in relation to docking pose 1, it was difficult to determine if the aromaticity of the phthalimide ring played a role in providing the observed potency. Even if placing aromaticity out in this region by alternative connectivity was successful, the high aromatic nature of the compound (4 aryl rings) was a concern. Phthalimide (**4.99**) has a measured PFI of 6.68 (Chrom LogD = 2.68; number of aromatic rings = 4). The measured chromatographic LogD is in an acceptable range, but the number of aromatic rings takes the PFI outside the desired range of 5 – 6. The measured aqueous solubility (CLND = 64 µg/mL) confirms that phthalimide (**4.99**) has lower solubility than desired. Without reduction of the

number of aromatic rings, the likelihood of the series providing a robust lead without suffering from developability hurdles (due to toxicity or poor pharmacokinetics) was reduced. Although rat biochemical data was not routinely measured for this series, several analogues were submitted for screening against the rat construct. Rat biochemical activity was of interest because mode of action studies prior to clinical testing in man were planned in rat. For the quinoline carboxamide series it had proved challenging to achieve rat biochemical activity (quinoline carboxamide **4.1** biochemical activity - h pIC₅₀ = 5.2, LE = 0.30, LLE_(AT) = 0.16; m pIC₅₀ = 6.4, LE = 0.37, LLE_(AT) = 0.22; r pIC₅₀ = 4.5, LE = 0.26, LLE_(AT) = 0.12). It was encouraging to see that for the analogues screened in the rat biochemical assay, pyrazolopyrimidinones (**4.71**) and (**4.99**) had comparable rat biochemical activity when compared to mouse. This rat activity differentiated this series from the quinoline carboxamide (**4.1**), and gave encouragement that with further work to improve the physicochemical properties, this series could provide a valuable tool molecule to help establish the physiological effects of inhibiting CD38.

4.7.6.1 Attempts to Improve the Physicochemical Properties

In an attempt to improve the properties of the series and explore alternative vectors to increase potency, synthesis of 5-membered heterocyclic analogues in the 5-position was initiated. Replacement of the phenyl ring of analogue (**4.65**), with thiophene to form (**4.68**), had given an increase in human potency and binding efficiencies. To follow-up on this result, compounds (**4.116**) – (**4.118**) were designed to test the feasibility of growing from a 5-membered heterocycle (**Figure 125**). In parallel, pyrazole (**4.119**) and furan (**4.120**) analogues were designed to test the impact of the positions of heteroatoms in the 5-membered ring. Docking of pyrazole (**4.119**) and furan (**4.120**) suggest that the heteroatoms are likely to be too far away from the protein to pick up specific hydrogen bonding interactions, but may interact with the protein through a water network. As mentioned previously, there were no obvious π -systems from the protein side chains that the heterocycles could π -stack with.

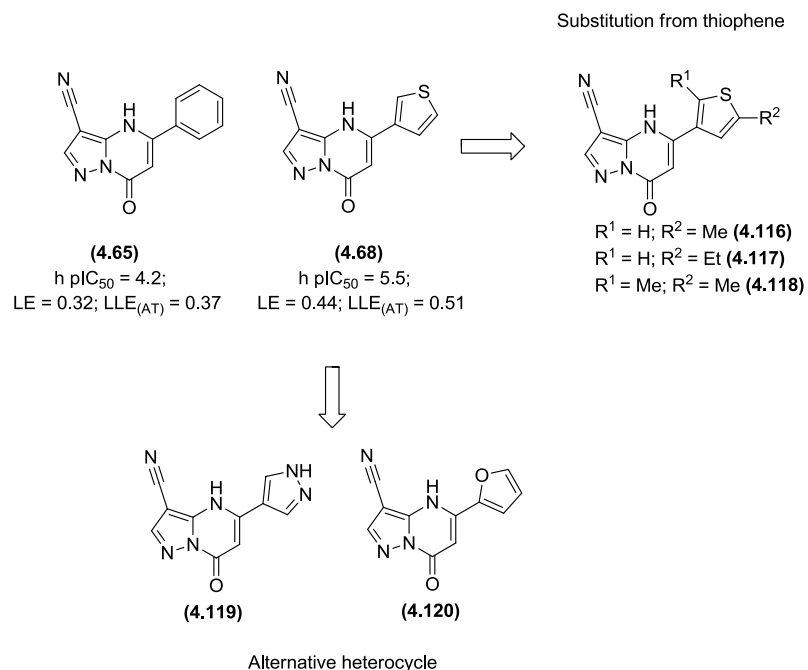
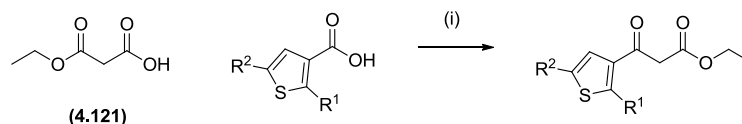


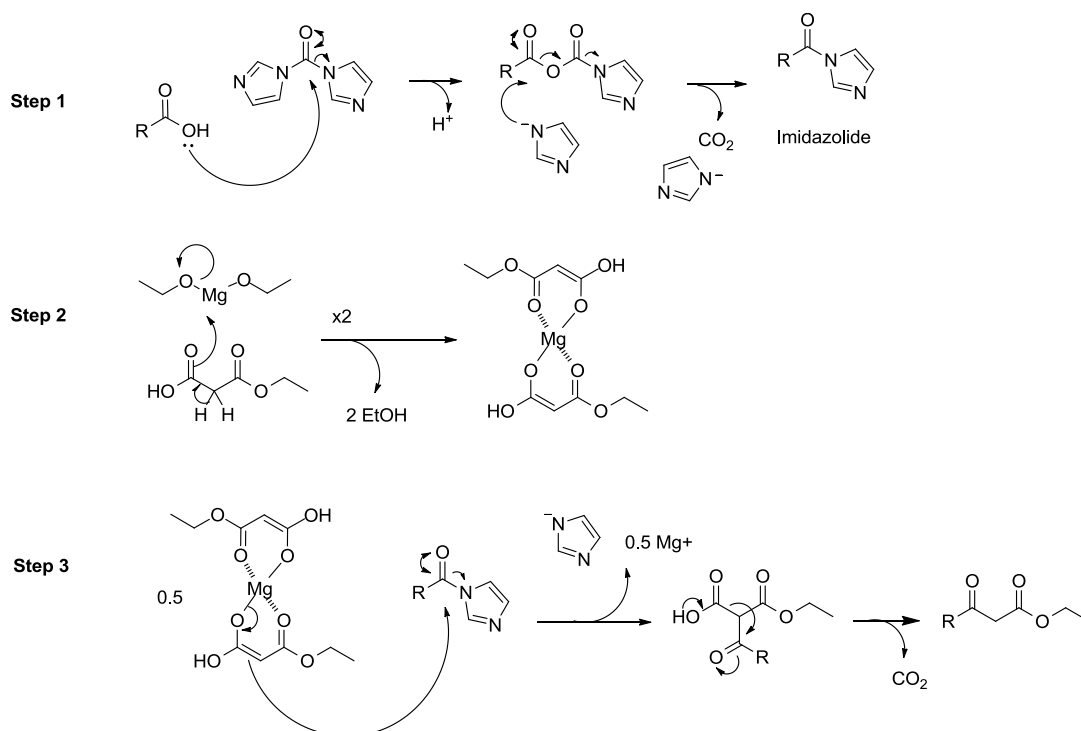
Figure 125 – Design ideas to test the potential of having a 5-membered heterocycle at the 5-position of the pyrazolopyrimidinone core.

Substituted thiophene analogues (**4.116**) – (**4.118**) were synthesised through formation of the corresponding β -keto esters (**4.125**) – (**4.127**) (**Scheme 30**) using C-acylation chemistry reported by Brooks *et al.*²²³ Although not critical for the synthesis of β -keto ester analogues (**4.125**) – (**4.127**), it is important to note that the methodology is suitable for compounds containing acid- or base-sensitive functionality. The procedure involves formation of the imidazolide by reaction of the corresponding carboxylic acid with CDI (**Scheme 31 - Step 1**). The activated carboxylate is subsequently treated with a solution of the neutral magnesium chelate of 3-ethoxy-3-oxopropanoic acid (**Scheme 31 – Step 3**). The β -keto esters (**4.125**) – (**4.127**) were taken on crude into the cyclisation step but it is important to note that by LCMS essentially clean conversion of carboxylic acids to the corresponding β -keto esters was achieved.



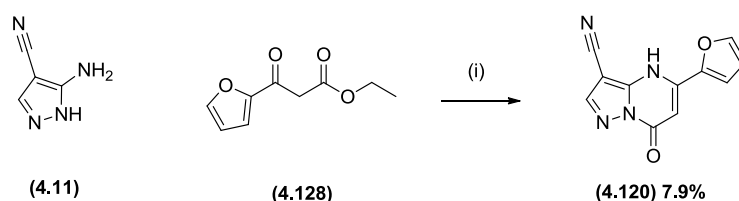
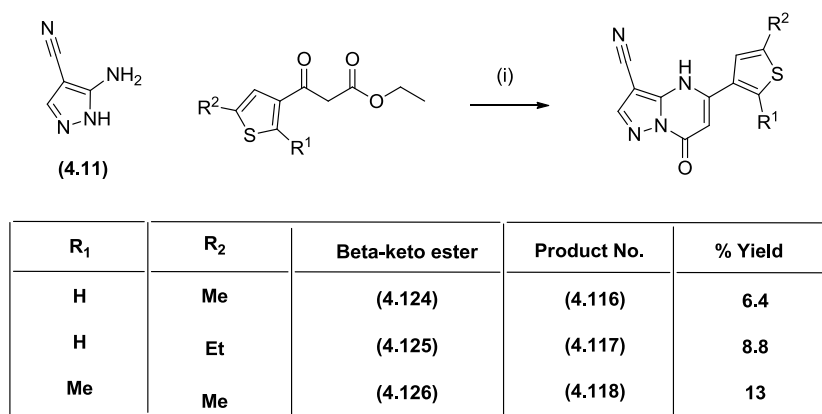
R ¹	R ²	Carboxylic acid	Beta-keto ester
H	Me	(4.122)	(4.125)
H	Et	(4.123)	(4.126)
Me	Me	(4.124)	(4.127)

Scheme 30 - β -Keto ester formation. Reagents and conditions: (i) CDI, Mg(OEt)₂, THF, RT, taken on crude into cyclisation step.²²³



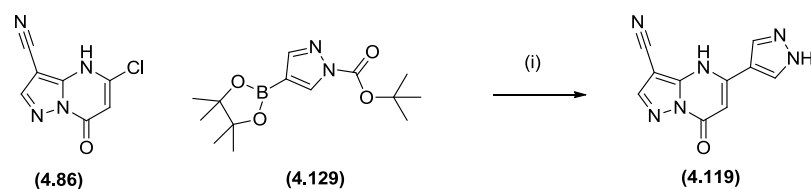
Scheme 31 – Mechanism of β -keto ester formation.

Substituted thiophene pyrazolopyrimidinones (**4.116**) – (**4.118**) and furan analogue (**4.120**) were synthesised by heating aminopyrazole (**4.11**) with the corresponding β -keto esters as previously described (**Scheme 32**). Furan β -keto ester (**4.128**) was commercially available.



Scheme 32 – Synthesis of thiophene and furan β -keto esters. Reagents and conditions: (i) acetic acid:water 2:1, 100 °C, 8%.

To complete the designed set of compounds (**Figure 125**), pyrazole (**4.119**) was prepared *via* Suzuki–Miyaura methodology. *N*-Boc-protected pyrazoleboronate (**4.129**) was reacted with the aryl chloride (**4.86**) as it was anticipated that with no protection the reaction would be hindered by coordination of the free pyrazole NH to the palladium. However, the Boc group was labile under the reaction conditions, which was likely to have contributed to the low yield achieved for the transformation (1%) along with the poor solubility. The reaction was not repeated or optimised as sufficient material was isolated for screening purposes, but alternative protecting groups are predicted to improve the yield.



Scheme 33 - Suzuki reactions to afford pyrazole analogue (**4.120**). Reagents and conditions: (i) K₂CO₃, SPhos (20 mol%), Pd(OAc)₂ (16 mol%), 100 °C, 1%.

CONFIDENTIAL – Property of GSK – Do Not Copy

Human and mouse biochemical data for the 5-membered heterocycles was disappointing, with all new analogues showing a reduction in both human and mouse potency. Furan (**4.120**) was active against both human and mouse enzymes, but due to having the same heavy atom count as thiophene (**4.68**) was less efficient. However, the data suggests that a hydrogen bond acceptor was tolerated in the 2-position of the 5-membered heterocycle, which gave insight into design of the next iteration of compounds.

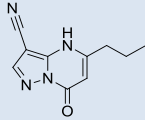
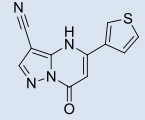
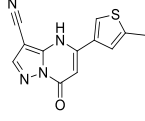
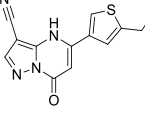
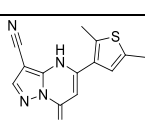
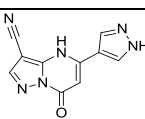
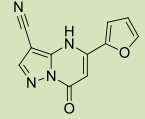
Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.10)		< 4.0			5.1	0.47	0.55
(4.68)		5.5 (n=1)	0.44	0.52	6.9 (n=1)	0.56	0.63
(4.116)		< 4.0			< 4.5		
(4.117)		< 4.0			5.3	0.38	0.38
(4.118)		< 4.0			5.7	0.41	0.44
(4.119)		< 4.0			< 5.0		
(4.120)		4.4	0.36	0.45	6.2	0.50	0.59

Table 17 - Biochemical data for heterocyclic analogues of thiophene-pyrazolopyrimidinone (**4.68**).

Figure 126 delineates the rationale around design of additional 5-membered heterocyclic analogues to follow-up on the biochemical results shown in **Table 17**. Growth of thiophene (**4.68**) to give (**4.116**) resulted in a reduction in human and murine potency. To explore this further and ensure that this effect was not specific to the thiophene heterocycle, an alternative heterocyclic analogue, methyl pyrazole (**4.130**), was designed. Methyl pyrazole (**4.130**) retained a hydrogen bond acceptor atom in the same position as the sulfur and explored the opportunity for growth with the methyl group. Furan (**4.120**) encouragingly showed biochemical activity against both mouse and human, but with a drop-off compared to thiophene (**4.68**). As discussed earlier, it is thought that the pyrazolopyrimidinone core binds to the protein in the tautomeric form shown in **Figure 126**. It was reasoned that the oxygen atom of furan (**4.120**) sits away from the pyrimidinone nitrogen atom to avoid repulsion of the lone pairs of electrons from the two hydrogen bond acceptors. To establish whether the binding from the furan oxygen and thiophene sulfur are additive, combination of the two donor atoms was considered. 5-Membered heterocycles containing both sulfur and oxygen do not exist, and it was hypothesised that a thiazole ring with the two hydrogen bond acceptor atoms could test the hypothesis as the thiazole nitrogen could act as a surrogate for the oxygen. However, there were concerns over the potential reactivity of the methine between the sulfur and nitrogen atoms, and unsubstituted thiazole ring (**4.131**) would not demonstrate the ability to grow to pick up additional potency. It was thus deemed logical that the methylthiazole (**3.132**) was chosen for synthesis.

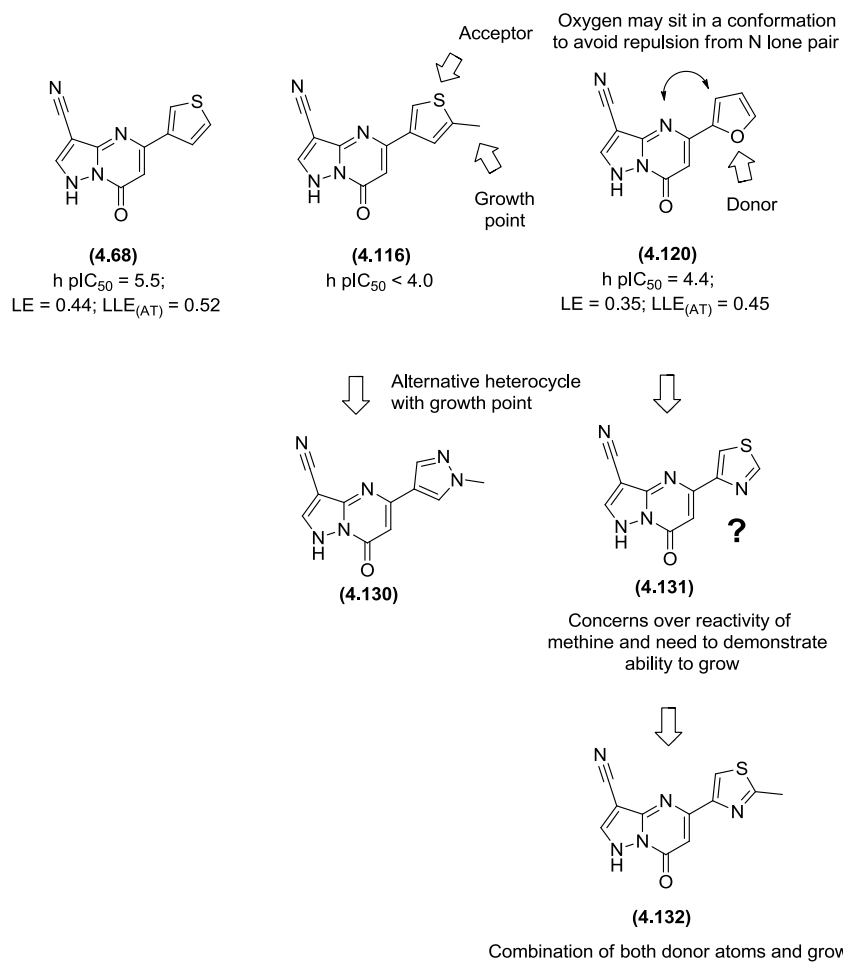
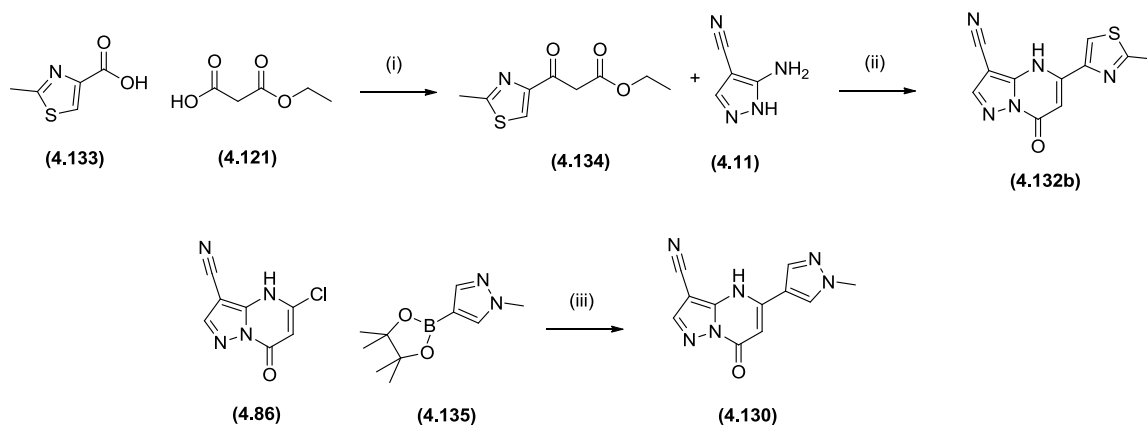


Figure 126 – Rationale around design of additional 5-membered heterocyclic analogues in response to the biochemical data in **Table 17**.



Scheme 34 – Synthesis of methyl thiazole (**4.132b**) and methyl pyrazole (**4.130**).

Reagents and conditions: (i) CDI, Mg(OEt)₂, THF, RT, taken on crude into cyclisation step;²²³ (ii) acetic acid:water 2:1, 100 °C, 3%; (iii) K₂CO₃, SPhos, Pd(OAc)₂, 100 °C, 7%. Note that **4.132b** is the ammonium salt of **4.132**.

Synthesis of methylthiazole (**4.132b**) and methylpyrazole (**4.130**) was performed using the standard synthetic procedures (**Scheme 34**). Disappointingly, little improvement in yield was seen for the methyl-pyrazole analogue (**4.130**) compared to unsubstituted pyrazole (**4.119**) as had been hypothesised. Despite the low yield, sufficient compound was isolated for screening and subsequent biochemical data for both analogues was encouraging (**Table 18**). Methylthiazole (**4.132b**) was equipotent to thiophene (**4.68**) against both CD38 constructs and had comparable efficiencies. The additional heteroatom in the thiazole, compared to the thiophene, helps maintain the LLE_(AT) whilst tolerating growth. Methylpyrazole (**4.130**) was less potent against both isoforms, but helped to demonstrate the importance of the heteroatom positions in the 5-membered heterocycle. The position of the heteroatoms could affect binding through changes to the conformation of the molecule or electrostatic interactions. A molecular dynamics conformation minimisation was carried out in MOE to examine for the lowest energy conformations of methyl thiazole (**4.132**) and methyl pyrazole (**4.130**) in the gas phase. No difference was observed between the dihedral angle of heterocycles relative to the core, confirming that this did not account for the difference in binding

affinity (**Figure 127**). The effect of the position of the heteroatom on the predicted key hydrogen bond interaction was considered. The predicted binding mode for the pyrazolopyrimidinone analogues indicated a hydrogen bond interaction from the pyrazolopyrimidinone carbonyl oxygen to Asp155 and Lys129 (**Figure 111**). The position of the heteroatom in the 5-membered heterocycle may have been affecting the pyrazolopyrimidinone core, causing the carbonyl to tautomerise to a more phenolic form (**Figure 128**). From the few compounds that have been synthesised, there appears to be a correlation between the biochemical potency and the ability to push electron through the π -system onto the carbonyl oxygen (**Figure 128**). This trend correlates with the difference in strength of hydrogen bond acceptor ability between the pyrazoles with respect to the furan. Laurence *et al.* quote that pyrazoles are stronger hydrogen bond acceptors compared to furans due to their hydrogen bond basicity.²²⁴ In this case, the donating ability of the pyrazoles may be increasing the proportion of the phenolic tautomeric form, hence reducing the ability of the oxygen to act as a hydrogen bond acceptor. Further analogue synthesis and molecular dynamics data would have been necessary to confirm this hypothesis, however, the biochemical data demonstrates how positioning the heteroatoms within each ring system can make a substantial difference in binding affinity to the target protein.

Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.10)		< 4.0			5.1	0.47	0.55
(4.68)		5.5 (n=1)	0.44	0.52	6.9 (n=1)	0.56	0.63
(4.132b)		5.5	0.42	0.54	6.8	0.52	0.64
(4.130)		4.1 ^a	0.31	0.50	<4.0 ^b	0.40	0.59

Table 18 – Second iteration of heterocyclic analogues of thiophene pyrazolopyrimidinone (**4.68**). ^aThis compound has only been reported active in 2 of the 6 screening occasions. ^bThis compound has only been reported active in 1 of the 6 screening occasions and therefore the active value is assumed a false result.

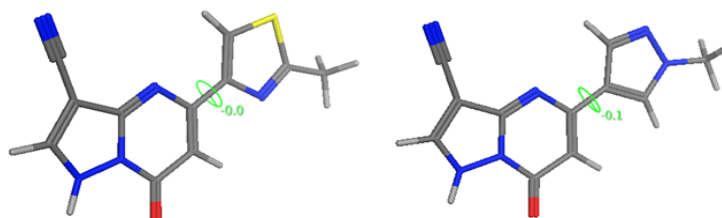


Figure 127 – Molecular dynamics conformational stochastic search using MOE for methyl-thiazole (**4.132**) and methyl-pyrazole (**4.130**). No difference in dihedral angle between the core and the 5-membered heterocycle was observed.

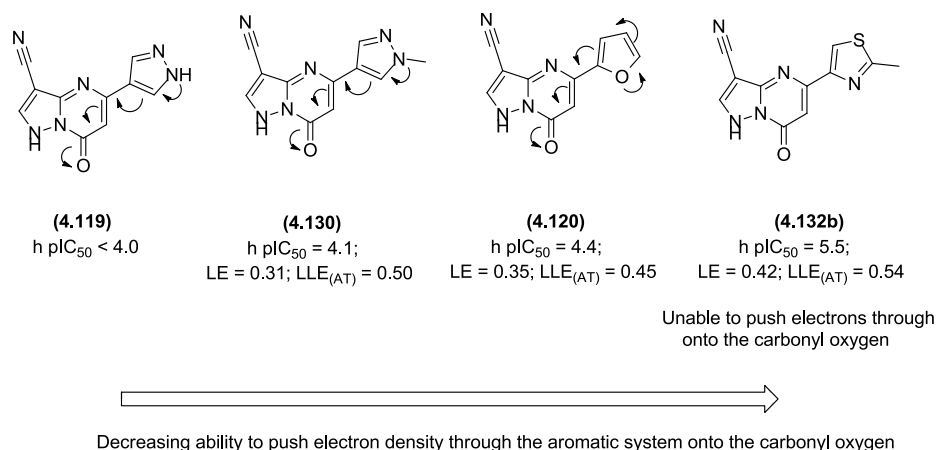


Figure 128 – Ability to delocalise charge through the π -system onto the carbonyl oxygen.

4.7.6.2 Summary of Efforts to Identify Potent, Efficient CD38 Binders from the Pyrazolopyrimidinone (4.10) HTS Hit

The initial aim of this project was optimisation of the HTS derived pyrazolopyrimidinone (4.10) using FBDD concepts to improve binding to the CD38 protein ($h\text{ pIC}_{50} > 6.0$), whilst maintaining good binding efficiency and physicochemical properties ($PFI < 6.0$; LE and $LLE_{(AT)} > 0.3$). The growth vectors of pyrazolopyrimidinone (4.10) were each assessed, demonstrating that potency, with maintained or improved binding efficiency, could only be obtained from 5-substitution from the pyrazolopyrimidinone core. SAR suggested that aromaticity was preferred in the 5-position, and that heterocycles were tolerated.

Pleasingly, this research programme has successfully increased both the mouse and human binding potency to CD38 (Table 19). Thiophene (4.68) and methyl thiazole (4.132b) are both more potent (against mouse and human) and more efficient than propyl (4.10), with good PFI values. CLND solubility confirmed that the molecules are around the lower end of the desired value, 100 $\mu\text{g/mL}$ (Table 19). Despite these positive observations around the series, there were relatively few indications that the series could be further developed to pick up the additional potency required to turn the lead molecule ($\text{pIC}_{50} > 6.0$) into a candidate molecule ($\text{pIC}_{50} > 7.0$). All

CONFIDENTIAL – Property of GSK – Do Not Copy

attempts to grow from the 5/6-membered heterocycles had failed to increase the biochemical potency and resulted in a drop in binding efficiencies. Phthalimide (**4.99**) was the most potent analogue against human enzyme, however, due to its increased lipophilicity and heavy atom count had lower binding efficiencies. Despite this, phthalimide (**4.99**) demonstrated that additional potency can be obtained by substantial growth, but CLND solubility for phthalimide (**4.99**) (64 $\mu\text{g/mL}$) combined with an PFI of 6.7 gave concern that the low solubility of the compound would be a barrier to progression. Despite the series having anchored a reasonable level of potency and efficiency, without a crystal structure of the pyrazolopyrimidinone series, the next iteration of compounds was not clear. In response to this, the potency and $\text{LLE}_{(\text{AT})}$ obtainable for the template was compared to other fragment-derived series identified from the FBDD screen and subsequent SSS. The aim of this analysis was to determine which series had the best profile for obtaining the desired potency, efficiency, and physicochemical package.

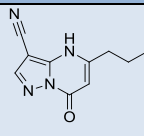
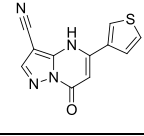
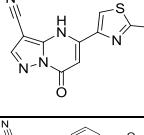
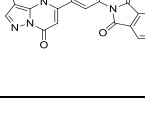
Cmpd No.	Structure	h pIC_{50}	h LE	h $\text{LLE}_{(\text{AT})}$	m pIC_{50}	m LE	m $\text{LLE}_{(\text{AT})}$	PFI	CLND $\mu\text{g/mL}$
(4.10)		< 4.0			5.1	0.47	0.55	2.6	101
(4.68)		5.5 (n=1)	0.44	0.52	6.9 (n=1)	0.56	0.63	4.3	≥ 134
(4.132b)		5.5	0.42	0.54	6.8	0.52	0.64	3.9	≥ 88
(4.99)		5.8 (n=1)	0.27	0.33	6.5 (n=1)	0.31	0.37	6.7	64

Table 19 – Most advanced analogues of pyrazolopyrimidinone (**4.10**).

The relative potency and $LLE_{(AT)}$ were compared for the pyrazolopyrimidinone series e.g. thiophene (**4.68**), and the imidazopyrazine series (**Figure 129**), which had been identified from SSS around the imidazopyridine fragment hit (**4.5**) (*vida infra*). The correlation of the efficiency ($LLE_{(AT)}$) with human activity from the pyrazolopyrimidinone and imidazopyrazine series, was encouraging (**Graph 6**). As human pIC_{50} (x-axis) increases, the molecules had increasing $LLE_{(AT)}$ (y-axis). An X-ray crystal structure was obtained for pyrazolopyridine (**4.5**), which was used to guide the design of molecules for the imidazopyrazine series. In order to assess whether the potency could be increased for imidazopyrazine analogues, whilst maintaining binding efficiency, the strategic decision was made to put optimisation of pyrazolopyrimidinone (**4.10**) on hold. The imidazopyrazine series offered additional opportunities for growth, providing greater potential to increase binding affinity to CD38, which did not appear feasible in the pyrazolopyrimidinone series.

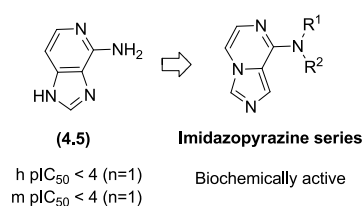
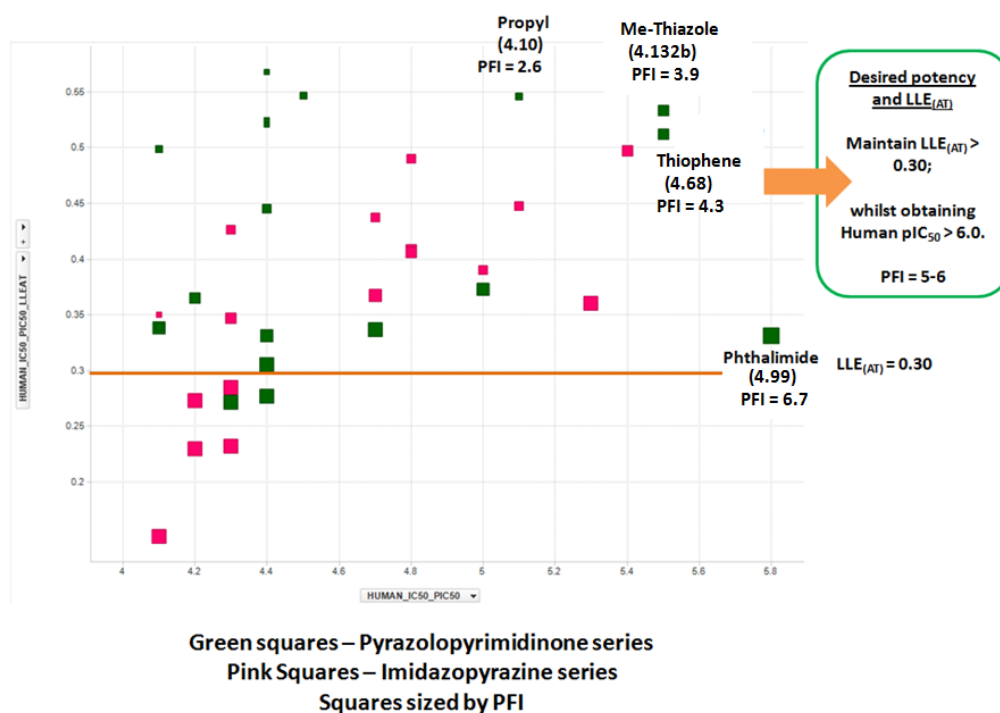


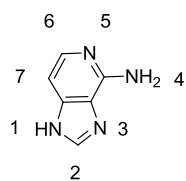
Figure 129 – Imidazopyrazine series derived from pyrazolopyridine (**4.5**).



Graph 6 – Graph showing biochemical data for the **pyrazolopyrimidinone** and **imidazopyrazine** analogues from SSS around the imidazopyridine fragment hit (4.5); human pIC₅₀ vs human LLE_(AT).

4.8 Optimisation of 1*H*-Imidazo[4,5-*c*]pyridin-4-amine (4.5) for CD38 Binding Leading to Identification of the Imidazopyrazine Series

1*H*-Imidazo[4,5-*c*]pyridin-4-amine (**4.5**) was identified from the second iteration of fragment screening (**Section 4.5.2**). The hit was identified by SPR²¹² screening ($K_d \approx 1\text{-}2\text{ mM}$), but it was undetectable in the biochemical assay (mouse (m) and human (h) $pIC_{50} < 4.0$). In spite of the low level of binding of the fragment to the CD38 protein, an X-ray crystal structure was obtained (**Figure 130**).¹⁹⁷ The X-ray crystal structure showed imidazopyridine (**4.5**) π -stacking with Trp189, with the imidazo-nitrogen atoms forming hydrogen bond interactions with Glu146 (2.41 Å) and Asp156 through water. The distances are indicative of good hydrogen bonding interactions, however, it is unclear which of the imidazo-nitrogen atoms are acting as hydrogen bond donors or acceptors, as the protonation states of the acidic amino acid side chains in this region of the protein are unclear. The X-ray crystal structure shows that the imidazopyridine fragment (**4.5**) binds uncompetitively with the truncated substrate by-product, where nicotinamide has been cleaved from NAD (highlighted by the yellow circle in **Figure 130**). Additional copies of the fragment were seen π -stacking between the substrate purine moiety and Trp176 (circled in white). The orientation of the additional copy of imidazopyridine fragment (**4.5**) is inconclusive from the electron density, but suggested an area of the protein pocket to grow towards to gain additional binding affinity.



(4.5)

h pIC₅₀ < 4 (n=1)

m pIC₅₀ < 4 (n=1)

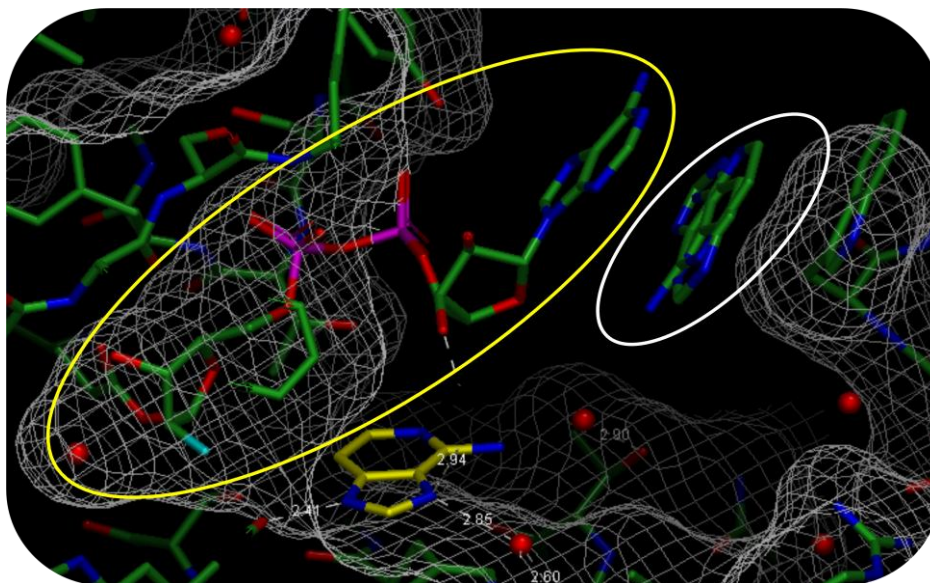
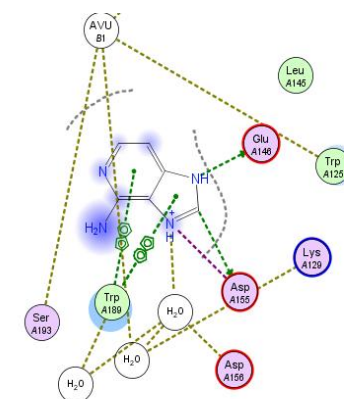
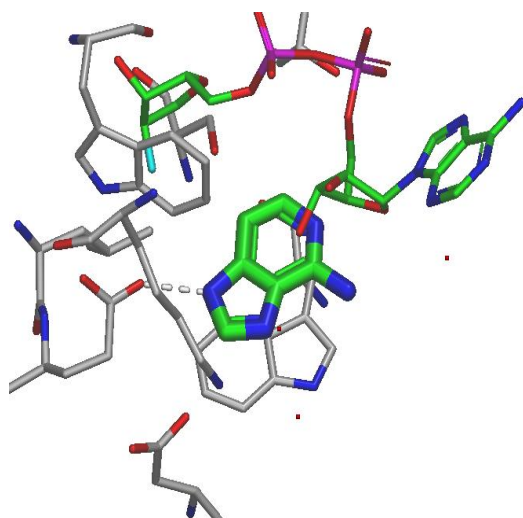


Figure 130 – X-ray crystal structure of imidazopyridine (4.5) PDB Code: 7HIZG (mouse). Top image: Hydrogen bond interactions to the CD38 protein; Bottom image: Imidazopyridine (4.5) binds un-competitively with the truncated substrate (circled in yellow). A second copy of imidazopyridine (circled in white) is seen triple π -stacking between the substrate and Trp176.

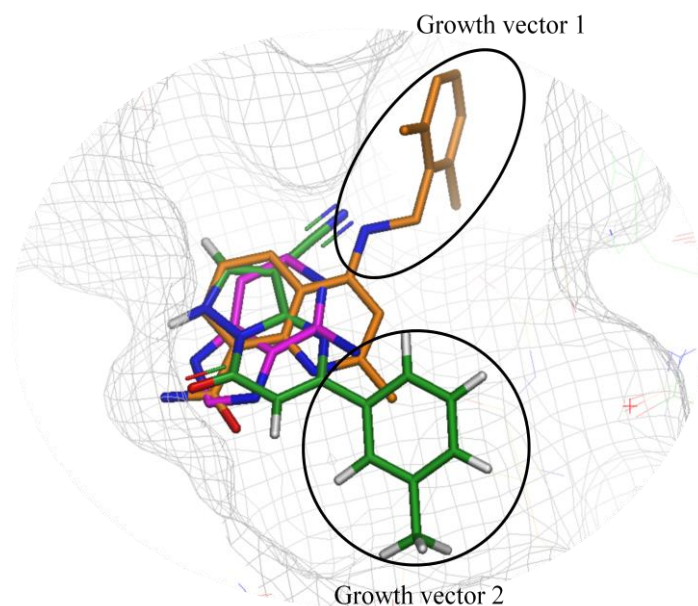
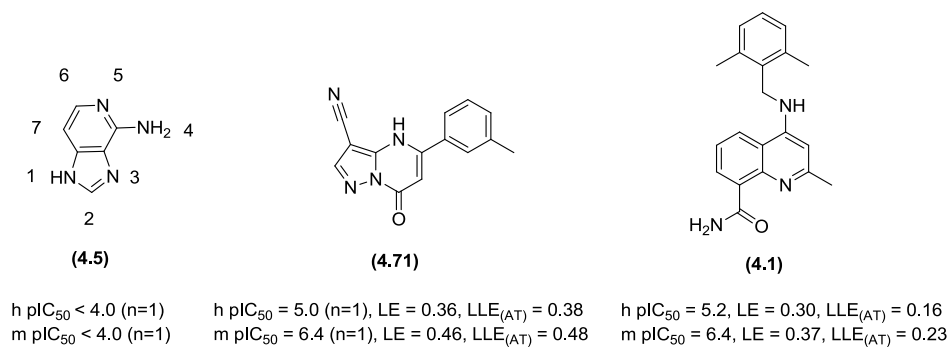


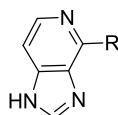
Figure 131 - X-ray crystal structure of imidazopyridine **(4.5)** (PDB Code: 7HIZG), overlaid with X-ray crystal structure of quinoline carboxamide **(4.1)** (PDB Code: 1GECC) and docking¹⁶⁵ of pyrazolopyrimidinone (docking pose 1) **(4.71)**. Potential areas to grow fragment **(4.5)** are highlighted with circles.

An overlay of the X-ray crystal structure of imidazopyridine **(4.5)** with the crystal structure of quinoline carboxamide **(4.1)** and a docked pose of pyrazolopyrimidinone **(4.71)**, highlighted two potential growth areas to explore to attempt to improve the potency of imidazopyridine **(4.5)** (**Figure 131**). Growth vector one projects into the area occupied by the 2,6-dimethylphenyl of quinoline carboxamide **(4.1)**. Growth of imidazopyridine **(4.5)** to fill this area of the pocket was not feasible unless the pyridine nitrogen (5-N) was removed or relocated to another position of the 5,6 heterocyclic system. Substitution from the 4-N atom of the imidazopyridine core

CONFIDENTIAL – Property of GSK – Do Not Copy

provided a good handle to explore growth vector two and was deemed a higher priority.

SSS of our laboratories' compound collection provided analogues enabling exploration of growth vector two (**Table 20**). All analogues screened were below the limit of the biochemical assay against both mouse and human enzymes, apart from cyclohexyl (**4.139**) which gave weak mouse activity (h pIC₅₀ < 4.0; m pIC₅₀ = 4.1, LE = 0.35, LLE_(AT) = 0.17). Cyclohexyl (**4.5**) had good LE, but poor LLE_(AT) due to the lipophilic nature of the R group. To progress the imidazopyridine series further, it was essential to show that potency could be gained against both CD38 constructs with good efficiencies. Improvement of LLE_(AT) could be achieved by addition of polarity, or by finding a more efficient location for the lipophilicity.



Cmpd No.	Structure	Human pIC ₅₀	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.5)		< 4.0 (n=1)	< 4.0 (n=1)		
(4.136)		< 4.1 (n=1)	< 4.1 (n=1)		
(4.137)		< 4.0	< 4.0		
(4.138)		< 4.0	< 4.0		
(4.139)		< 4.0	4.1 ^a	0.35	0.17
(4.140)		< 4.0	< 4.0		
(4.141)		< 4.0	< 4.0		

Table 20 – SSS data to explore growth vector two of imidazopyridine (**4.5**). ^aThis compound has only been reported active in 1 of the 4 screening occasions (pIC₅₀ = <4.98, <4.0, <4.0, 4.05).

CONFIDENTIAL – Property of GSK – Do Not Copy

Reanalysis of HTS expansion data from a LE perspective within our laboratories^{172,213} highlighted pyrazolopyrimidine analogues, cyclohexyl (**4.142**) (h pIC₅₀ = 4.9, LE = 0.42, LLE_(AT) = 0.28; m pIC₅₀ = 6.3, LE = 0.54, LLE_(AT) = 0.40) and thiophene (**4.143**) (h pIC₅₀ < 4.0; m pIC₅₀ = 4.3, LE = 0.37, LLE_(AT) = 0.31). Comparison of pyrazolopyrimidine cyclohexyl (**4.142**) and thiophene (**4.143**), with the corresponding imidazopyridine analogues ((**4.139**) and (**4.141**), respectively) showed that the pyrazolopyrimidine template was more potent and efficient. Despite repeated attempts to obtain a pyrazolopyrimidine liganded CD38 crystal structure, this has not been achieved. Inspection of the crystal structure of imidazopyridine (**4.5**) (**Figure 132**) shows that changing the imidazole ring to a pyrazole ring could still result in two hydrogen bonds being made in the carboxylic acid region of the protein (to Glu146 (2.41 Å) and Asp 156 (3.16 Å)). Clockwise rotation of the pyrazolopyrimidine core, with respect to the binding mode of imidazopyridine (**4.5**), could improve the H-bonding distances and enhance the binding efficiency.

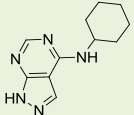
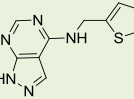
Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.142)		4.9	0.42	0.28	6.3	0.54	0.40
(4.143)		< 4.0			4.3	0.37	0.31

Table 21 – Biochemical data for pyrazolopyrimidine analogues cyclohexyl (**4.142**) and thiophene (**4.143**).

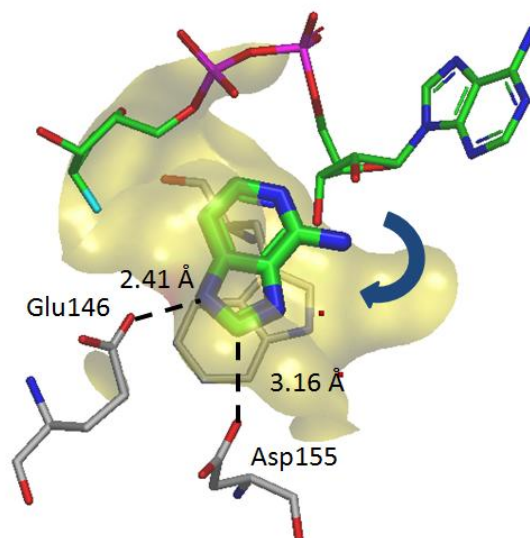
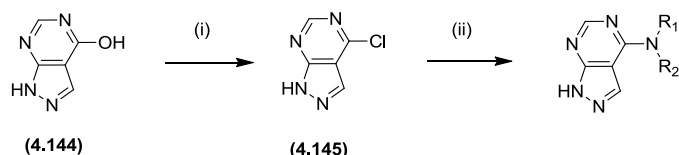


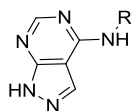
Figure 132 - X-ray crystal structure of imidazopyridine (**4.5**) (PDB Code: 7HIZG) showing potential hydrogen bonds that could be made for the pyrazolopyrimidine analogues.

Pyrazolopyrimidine analogues were prepared from commercially available 1*H*-pyrazolo[3,4-*d*]pyrimidin-4-ol (**4.144**). Conversion of pyrazolo[3,4-*d*]pyrimidin-4-ol (**4.144**) to aryl chloride (**4.145**) in POCl₃ with dimethylaniline at 110 °C gave 82% conversion to product.²²⁵ S_NAr of appropriate amines on aryl chloride (**4.145**) was carried out. The amines were selected to systematically explore the importance of the ring size, to ascertain if the cyclohexyl was optimal before starting to grow the compound. Amines were also chosen to assess if polarity or substitution from the cyclohexyl ring was tolerated. It should be noted that specific design of amines in the absence of a crystal structure of the pyrazolopyrimidine core was difficult. Also, if the proposed binding mode was correct, the amines would be occupying a large open shelf in the CD38 pocket making the orientation more challenging to predict. Although the selection of amines was part of this research project, these compounds were synthesised elsewhere in our laboratories using the route described in **Scheme 35**.²²⁵



Scheme 35 – Synthetic route to pyrazolopyrimidine analogues. Reagents and conditions: (i) POCl₃, *N,N*-dimethylaniline, 110 °C, 1.5 h, 82%; (ii) isopropanol or ethyl acetate, amine, 80 – 100 °C, μ wave.

Analysis of the biochemical data for pyrazolopyrimidine analogues (**4.146**) – (**4.154**) (**Table 22**) demonstrated that the cyclohexyl ring could be reduced in size to cyclopentyl without a loss of binding efficiency. Conversely, compounds with smaller rings, cyclopropyl (**4.146**) and cyclobutyl (**4.147**) were inactive against both human and mouse orthologues. THF (**4.149**) and THP (**4.150**) analogues showed that oxygen atoms were not tolerated. Methyl-substituted cyclohexyl analogues suggested that growth was feasible, but with an initial drop in binding potency and efficiency. Further growth could regain this lost binding affinity. It was encouraging to see both mouse and human activity for methylpiperidine (**4.153**) (h pIC₅₀ = 4.4, LE = 0.35, LLE_(AT) = 0.39; m pIC₅₀ = 5.9, LE = 0.48, LLE_(AT) = 0.51). The basic nitrogen gave improved potency, LE and particularly LLE_(AT) for methylpiperidine (**4.153**), over cyclohexyl (**4.142**). Disappointingly, growth of methyl-piperidine (**4.153**) to give ethyl piperidine (**4.154**), only gave weak mouse activity.



Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.142)		4.9	0.42	0.28	6.3	0.54	0.40
(4.146)		< 4.0			< 4.0		
(4.147)		< 4.0			< 4.0		
(4.148)		5.1 ^a	0.47	0.37	5.6 ^b	0.51	0.41
(4.149)		< 4.0			< 4.0		
(4.150)		< 4.0			< 4.0		
(4.151)		4.2	0.34	0.18	4.7	0.38	0.22
(4.152)		< 4.0			4.7	0.38	0.22
(4.153)		4.4 (n=1)	0.35	0.39	5.9 (n=1)	0.48	0.51
(4.154)		< 4.0			4.2	0.32	0.32

Table 22 – Biochemical data for pyrazolopyrimidine analogues (4.146) – (4.154).

^aThis compound has only been reported active in 3 of the 6 screening occasions (pIC₅₀ = <4.0, <4.0, <4.0, 5.43, 4.38, 5.45). ^bThis compound has only been reported active in 4 of the 6 screening occasions (pIC₅₀ = 4.45, <4.0, <4.48, 6.71, 6.79).






Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.142)		4.9 (n=2)	0.42	0.28	6.3 (n=2)	0.54	0.40
(4.155)		< 4.0			< 4.0		
(4.156)		< 4.0			5.7	0.52	0.61
(4.157)		< 4.0			4.9	0.42	0.35
(4.158)		< 4.0			< 4.0		

Table 23 – Biochemical data of exemplars from SSS of our laboratories’ compound collection for secondary amines attached to the pyrazolopyrimidine core. All screening data for compounds in this table are n=1 unless stated otherwise.

Exemplars from SSS of our laboratories’ compound collection showed that small secondary and tertiary amines attached to the pyrazolopyrimidine core (**Table 22** and **Table 23**) gave no potency or efficiency advantage over cyclohexyl (**4.142**). Growth of these small amines to pick up additional interactions with CD38 was a feasible strategy to gain additional potency, as these compounds may be too small to pick up efficient interactions in the CD38 pocket. It was suggested that the small amine was orientated out into a large open cavity meaning that a large proportion of the fragment was solvent exposed.

An additional strategy was used to hybridise the pyrazolopyrimidine analogues with compounds that had provided liganded CD38 X-ray crystal structures. An alternative CD38 fragment identified during the FBDD screening cascade was 4-amino-isoquinoline (**4.6**) (**Section 4.5.2**). SSS of our laboratories’ compound collection around cluster 4-amino-isoquinoline (**4.6**), identified indazole analogue (**4.159**).¹⁹⁶ Indazole (**4.159**) showed activity against both mouse and human orthologues, with moderate LLE_(AT), but had high aromatic ring count so was not an attractive starting point. Despite this, a liganded CD38 X-ray crystal structure of

indazole (**4.159**) was obtained (**Figure 134**) giving guidance on how to grow and improve interactions with the CD38 pocket for the pyrazolopyrimidine template.¹⁹⁷

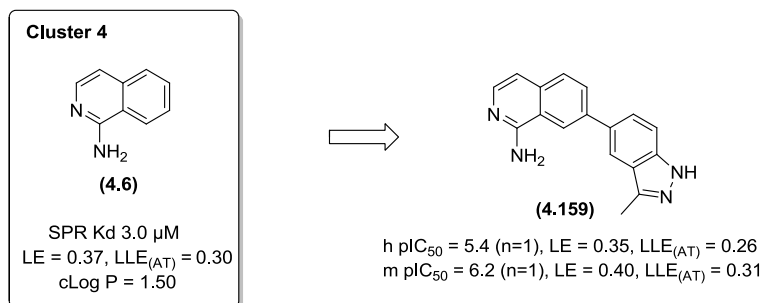


Figure 133 – Output from SSS around cluster 4 amino-isoquinoline (**4.6**).

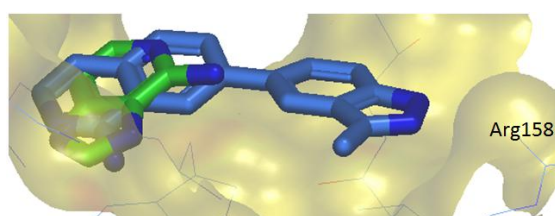


Figure 134 – X-ray crystal structure of indazole (**4.159**) (PDB Code: 8RTDI) overlaid with X-ray crystal structure of imidazopyridine (**4.5**) (PDB Code: 7HIZG).

Overlay of the X-ray crystal structures of indazole (**4.159**) and imidazopyridine (**4.5**) showed that both fragments participated in interactions with the carboxylic region of the protein (Glu146 and Asp155). The indazole motif of (**4.159**) extended into the area defined as growth vector two, picking up a hydrogen bond interaction with Arg158 (N–HN 3.02 Å) in the mouse construct. This region was predicted by GRID to provide the opportunity for profitable interactions. Assuming the pyrazolopyrimidine core bound in a similar orientation to the imidazopyridine (**4.5**), fusion of a pyrazole onto the piperidine ring of (**4.157**), to give (**4.160**), reaching towards Arg158 was suggested by docking as means of improving potency (**Figure 135**). Docking studies showed fusion onto the piperidine ring of (**4.157**) would have better overlap with (**4.160**) compared to fusion onto the cyclohexyl of (**4.142**) or the piperidine ring of (**4.153**). The exact position of the hydrogen bond donor and acceptor atoms was not critical, as arginine side chains are flexible giving an

increased chance of picking up a hydrogen bonding interaction. In unpublished exploration of an alternative drug target, nitrile analogues were shown to form profitable hydrogen bond interactions with arginine side chains, therefore, nitrile substitution from the cyclohexyl ring also warranted investigation. Although the selection of amines was part of this research project, the compounds were synthesised elsewhere in our laboratory using the S_NAr chemistry described in **Scheme 35**.²²⁵ It is important to note that pyrazole analogue (**4.160**) had greater sp^3 character than indazole (**4.159**). The fully aromatic indazole analogue of (**4.160**) was not prepared as the molecule was expected to exhibit poor solubility due to π -stacking between copies of itself due to the planarity of the molecule.

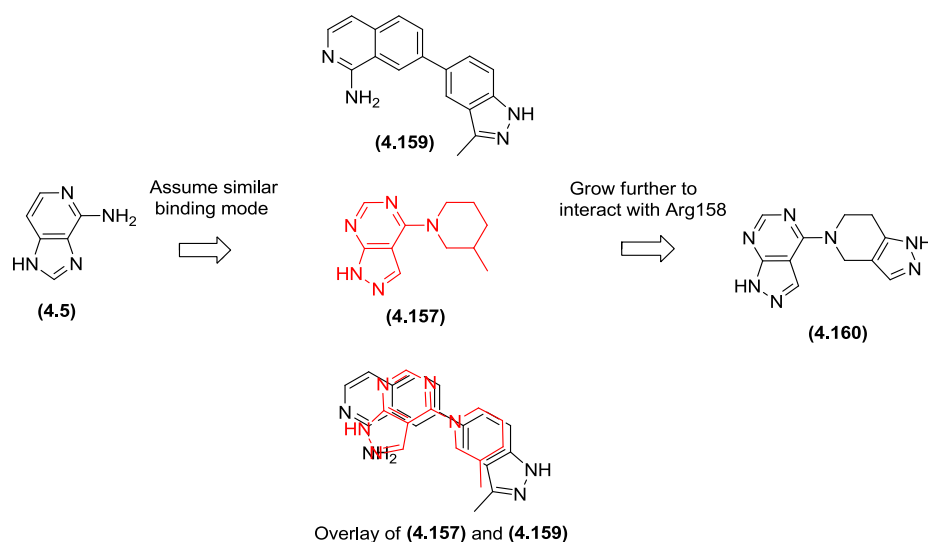
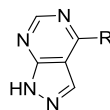


Figure 135 – Design of extended analogues of piperidine (**4.157**) to pick up additional interactions with the CD38 protein.

Analysis of the biochemical data for the extended analogues of methyl piperidine compound (**4.157**) (**Table 24**) showed no additional potency gain. Pyrazole (**4.162**) (h pIC_{50} = 4.2, LE = 0.32, $LLE_{(AT)}$ = 0.42; m pIC_{50} = 4.9, LE = 0.37, $LLE_{(AT)}$ = 0.48) was the only analogue to show human potency, albeit only weakly. Despite no increase in mouse or human potency being observed for compounds (**4.161**) – (**4.165**), it was encouraging to see that polarity could be tolerated in this region, resulting in an increase in $LLE_{(AT)}$ compared to methyl analogue (**4.157**). This

knowledge was noted and used to underpin further hybridisation into other fragment templates for CD38.



Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.157)		< 4.0 (n=1)			4.9 (n=1)	0.42	0.35
(4.161)		< 4.0			4.7	0.36	0.46
(4.162)		4.2	0.32	0.42	4.9	0.37	0.48
(4.163)		< 4.0			4.9	0.35	0.44
(4.164)		< 4.0			5.2	0.34	0.42
(4.165)		< 4.0			5.2	0.42	0.54

Table 24 – Biochemical data for analogues of piperidine (**4.157**) designed to pick up additional interactions to Arg158.

Alongside follow-up of the pyrazolopyrimidine core, exploration of alternative 5,6 heterocycles was pursued within our laboratories.¹⁷² This exercise highlighted the imidazopyrazine core as a suitable pyrazolopyrimidine replacement (**Table 25**). Initially cyclohexyl pyrazolopyrimidine (**4.142**) was of greater interest than the cyclohexyl imidazopyrazine (**4.166**), due to the pyrazolopyrimidine core showing biochemical activity against both mouse and human constructs. However, the interest in the imidazopyrazine core increased when thiophene (**4.167**) demonstrated higher biochemical potency and efficiency over its direct pyrazolopyrimidine analogue (**4.143**). This data suggests that there are subtle differences in SAR between the two bicyclic templates.

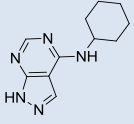
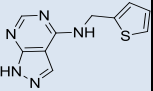
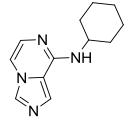
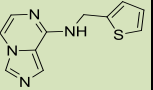
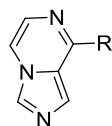
Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.142)		4.9	0.42	0.28	6.3	0.54	0.40
(4.143)		< 4.0			4.3	0.37	0.31
(4.166)		< 4.0 (n=1)			5.6 (n=1)	0.48	0.33
(4.167)		4.2	0.36	0.29	5.8	0.50	0.42

Table 25 – Pyrazolopyrimidine compared with imidazopyrazine analogues.

Compounds from SSS and synthesis by other members of our laboratories²²⁵⁻²²⁷ generated the biochemical data shown in **Table 26** around heterocyclic changes to thiophene (**4.167**). It can be seen from the biochemical data that thiophene (**4.167**) is equipotent against human CD38 to its phenyl isostere (**4.168**) and bromo-phenyl (**4.169**). In both CD38 constructs, a LE drop is observed from thiophene (**4.167**) to phenyl (**4.168**) and bromophenyl compound (**4.169**) analysis due to increased heavy atom count. The three pyridyl isomers (**4.170**) – (**4.172**) were all equipotent with thiophene (**4.167**) suggesting that no specific hydrogen bond interactions were formed with the protein. However, the chromLog D of pyridyl analogues is lower than for thiophene, giving an improvement in LLE_(AT). The improved properties of the pyridyl ring were exploited further by adding small substituents. Bromopyridine (**4.173**) (h pIC₅₀ = 6.2, LE = 0.47, LLE_(AT) = 0.43; m pIC₅₀ = 5.6, LE = 0.43, LLE_(AT) = 0.39) was the most potent imidazopyrazine analogue screened. Although the bromine atom added lipophilicity to the analogue, the pyridyl nitrogen helped maintain good lipophilic ligand efficiency. It was intriguing as to why an enrichment

in potency was observed when the *o*-pyridyl and *p*-bromine were combined in a single molecule, when compounds, *p*-bromide (**4.169**) and *o*-pyridyl (**4.172**), showed no potency advantage over thiophene (**4.167**) or phenyl (**4.168**). It was hypothesised that the position of the bromide substituent may be important and the *o*-pyridyl nitrogen was conformationally restricting the molecule in this orientation by forming an intramolecular hydrogen bond interaction to the aryl NH.



Compound No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.167)		4.2	0.36	0.29	5.8	0.50	0.42
(4.168)		4.0 (n=1)	0.32	0.23	5.5 (n=1)	0.44	0.35
(4.169)		4.2	0.32	0.18	6.0	0.46	0.31
(4.170)		4.6	0.37	0.40	6.0	0.48	0.51
(4.171)		4.2	0.34	0.37	6.2	0.50	0.53
(4.172)		4.4	0.35	0.38	5.8	0.47	0.50
(4.173)		6.2	0.47	0.43	6.6	0.50	0.46

Table 26 – Analogues to explore changes to thiophene (**4.167**).

A conformational search using MOE was carried out to identify the lowest energy conformations for bromide analogues (**4.173**) and (**4.169**) (**Figure 136**). The lowest energy conformations place the phenyl or pyridyl rings in very different orientations. As hypothesised, in the lowest energy conformation, the pyridyl ring of analogue (**4.173**) formed an intramolecular hydrogen bond from the pyridyl nitrogen to the NH of the *exo*-nitrogen of the imidazopyrazine core. The hydrogen bond keeps the pyridyl ring in the same plane as the imidazopyrazine core. In contrast, the lowest energy conformation of bromophenyl analogue (**4.169**) sits with the pendant ring

orthogonal to the core (dihedral angle = - 88.6 °). Docking of both bromide analogues into the CD38 protein was carried out and compared to the lowest energy molecular conformations (**Figure 136**). The bromo pyridyl analogue (**4.173**) docked in a very similar pose to the lowest energy conformation of the molecule (**Figure 137**). The docked pose of bromophenyl (**4.169**) predicts that the phenyl ring would bind in a flatter conformation in the CD38 protein pocket than seen in the lowest energy conformation prediction, thus, maximising the VdW interactions between the molecule and the protein. A more pronounced twist of the phenyl ring of (**4.169**) would raise the bromine away from the surface of the protein. The difference observed between the docked pose for (**4.169**), and the predicted lowest energy conformation, suggested there is a likely energy penalty to overcome on binding of the molecule to the protein. This could account for the difference in biochemical potency between bromide analogues (**4.169**) and (**4.173**), despite the fact that the docked poses for both molecules were quite similar.

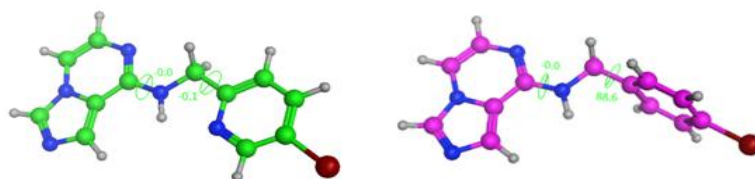


Figure 136 – Lowest energy conformations (Forcefield: MMFF94x) for bromopyridyl (**4.173**) and bromophenyl (**4.169**) showing the difference in dihedral angles.

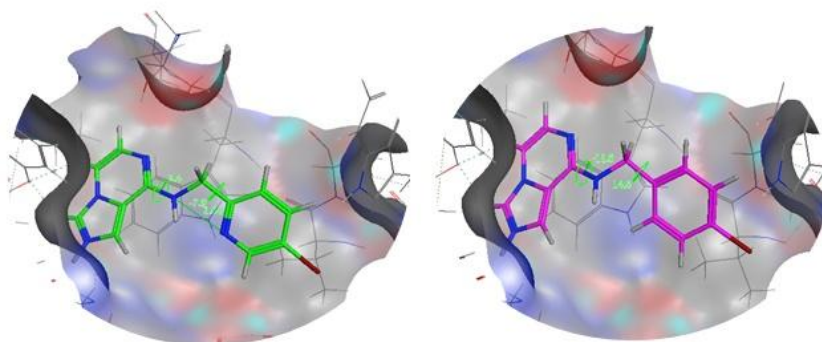


Figure 137 – MOE docking for bromopyridyl (**4.173**) and bromophenyl (**4.169**) compounds.

From the biochemical data and docking poses it was hypothesised that the intramolecular hydrogen bond interaction in **(4.173)** placed the bromine in a region of space that gave an increase in the binding affinity of the molecule to the protein. To understand this hypothesis further, molecules with enforced constraint, isoindoline **(4.174)** and tetrahydroisoquinoline **(4.175)** were designed and synthesised. By formally tying the molecule into the desired conformation for binding, this would remove the energy penalty to form the planar confirmation. Both the isoindoline and tetrahydroisoquinoline analogues were predicted to place the bromine atom in a similar location (**Figure 138**).

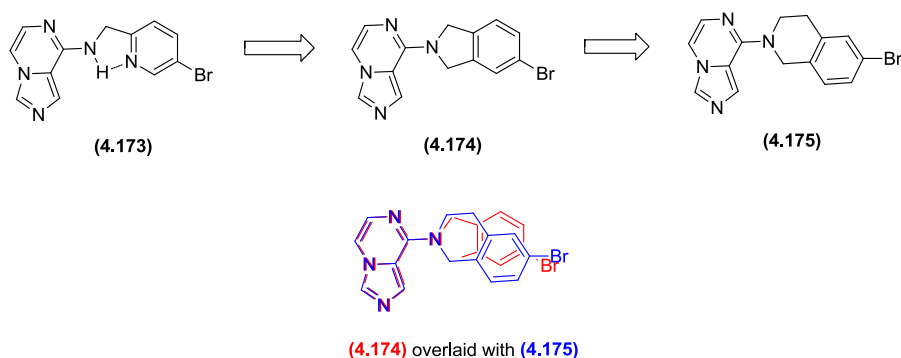
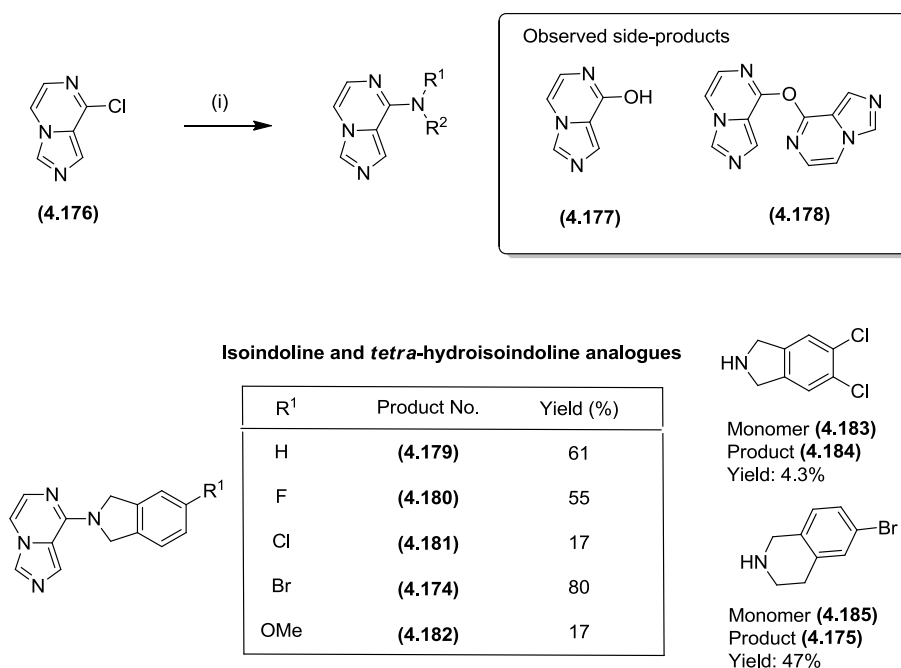


Figure 138 – Mimicking the intramolecular hydrogen bond interaction of bromopyridyl compound **(4.173)** by synthesis of isoindoline **(4.174)** tetrahydroisoquinoline **(4.175)**.

Our laboratories' compound collection was searched for 5-substituted isoindoline building blocks to react with 8-chloroimidazo[1,5-*a*]pyrazine **(4.176)**. Pleasingly, 5-bromoisoindoline and 6-bromo-1,2,3,4-tetrahydroisoquinoline **(4.185)** were available to test the hypothesis that cyclisation to replace the intramolecular hydrogen bond in bromide **(4.173)** could increase binding affinity of the molecule to CD38 (**Scheme 36**). To explore the SAR around the bromine atom, a variety of 5-substituted isoindoline molecules were synthesised (H, F, Cl, OMe) (**Scheme 36**). The substituents varied in size, and would provide information relating to the amount of room in the relevant region of the protein pocket and the nature of any hydrogen

bond interactions. From docking it was feasible that there was sufficient space for simultaneous 4- and 5-isoindoline substitution. This hypothesis could be tested by reaction of dichloroisoindoline (**4.183**) with 8-chloroimidazo[1,5-*a*]pyrazine (**4.176**) to afford dichloroisoindoline (**4.184**). Reaction of chloride (**4.176**) was carried out with K_2CO_3 and the appropriate amine in dry DMSO at 60 °C. Anhydrous conditions helped to minimise observed side products, hydroxyl (**4.177**) and ether (**4.178**). Purification methods improved over time, as the products were challenging to handle as they were poorly soluble in most solvents.



Scheme 36 – Synthesis of isoindoline and *tetra*-hydroisoindoline analogues.

Reagents and conditions: (i) amine, K_2CO_3 , anhydrous DMSO, 60 °C.

Despite the concerns over the poor solubility of the isoindoline analogues during synthesis, clear SAR and encouraging activity was observed (Table 27). By adding two extra bonds to constrain (**4.172**) to form the unsubstituted isoindoline (**4.179**) an increase in the biochemical potency and LE against both orthologues was observed. Despite the increase in potency, the $LLE_{(AT)}$ remained unchanged compared to *o*-pyridyl (**4.172**) due to isoindoline (**4.179**) having increased lipophilicity. A similar trend is observed for comparison of bromopyridyl compound (**4.173**) with bromoisoindoline (**4.174**), however, the increase in mouse biochemical potency was

more pronounced resulting in increased LE and $LLE_{(AT)}$. Increasing the size of the isoindoline substituent had a positive influence on the potency, with the methoxy analogue (**4.182**) displaying the greatest potency against both the human and mouse constructs. Disubstitution in compound (**4.184**) was also tolerated, which suggested that the monosubstituted isoindoline could place the substituent in two different orientations in the protein pocket. Although tetrahydroisoquinoline (**4.175**) was active against both orthologues, it was less potent and less efficient than the corresponding isoindoline (**4.174**), and therefore of lesser interest. The high PFI for this series of compounds ($PFI \geq 6.0$) fits with the measured modest solubility of the isoindoline and tetrahydroisoquinoline analogues (bromide isoindoline (**4.174**) CLND = 12 $\mu\text{g/mL}$ and tetrahydroisoquinoline (**4.175**) CLND = 34 $\mu\text{g/mL}$). Based on all of these outputs, despite the poor solubility, some promise had been established within this series with the consideration that solubility improvement could be gained by introduction of sp^3 character.



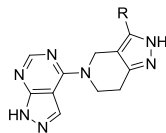
Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)	PFI
(4.172)		4.4	0.35	0.38	5.8	0.47	0.50	4.0
(4.173)		6.2	0.47	0.43	6.6	0.43	0.39	5.3
(4.179)		5.5	0.42	0.38	6.7	0.51	0.47	6.0
(4.180)		6.4	0.46	0.42	7.5	0.54	0.50	6.3
(4.181)		6.7	0.48	0.40	7.4	0.53	0.45	7.1
(4.174)		6.8	0.49	0.40	7.4	0.53	0.44	7.3
(4.182)		7.0	0.48	0.46	7.8	0.53	0.52	-
(4.184)		6.9	0.47	0.36	7.2	0.49	0.38	7.9
(4.175)		5.8	0.40	0.28	6.5	0.45	0.33	8.0

Table 27 – Biochemical data for isoindoline analogues compared to pyridyl analogues (4.172) and (4.173).

With the sub-optimal solubility in mind, attention turned to alternative semi-saturated bicyclic systems to replace the isoindoline, which could help improve the aqueous solubility of the compounds. Pyrazole analogues (4.162), (4.163) and (4.164) made in the pyrazolopyrimidine series (Table 28) showed good mouse potency and had increased 3-dimensional character and polarity compared to the isoindolines. These characteristics were expected to provide an improved solubility profile over the isoindolines. The corresponding imidazopyrazine analogues were made in our laboratories (Table 29),²²⁶ to determine if changing the 5,6 heterocyclic core from pyrazolopyrimidine to imidazopyrazine was tolerated. The initial SAR was encouraging with all analogues displaying activity against both mouse and human

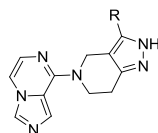
CONFIDENTIAL – Property of GSK – Do Not Copy

enzymes, exhibiting good LE, LLE_(AT) and PFI profiles. *p*-Methoxybenzyl (PMB) analogue (**4.191**) was made as an intermediate, but did demonstrate that there was a substantial amount of room for amide substitution, and that potency could be gained by occupying this region.



Compound No.	R-group	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.162)	H	4.2	0.32	0.42	4.9	0.37	0.48
(4.163)	Me	< 4.0			4.9	0.35	0.44
(4.164)	CO ₂ H	< 4.0			5.2	0.34	0.42

Table 28 – Bicyclic pyrazole analogues made in the pyrazolopyrimidine series.



Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)	PFI
(4.186)	Me	4.8	0.35	0.41	6.7	0.48	0.55	4.0
(4.187)	CO ₂ H	4.7 ^a	0.31	0.37	6.1	0.40	0.46	2.6
(4.188)		6.0	0.37	0.49	6.7	0.42	0.53	3.6
(4.189)		5.7	0.34	0.42	6.8	0.41	0.49	4.1
(4.190)		5.7	0.33	0.39	6.2	0.35	0.42	4.6
(4.191)		6.7	0.31	0.33	6.4	0.29	0.32	6.6

Table 29 – Bicyclic pyrazole SAR with the imidazopyrazine 5,6 heterocyclic core.

^aThis compound has only been reported active in 2 of the 4 screening occasions (pIC₅₀ = <4.52, <4.52, 4.87, 4.61).

At this stage, the focus of the research programme described here turned to optimisation of the amide substituent on methyl amide (**4.188**) through design and synthesis. *p*-Methoxybenzyl (**4.191**) gave confidence that there was room in the CD38 pocket for large amide substituents. Despite *p*-methoxybenzyl (**4.191**) having moderate efficiencies, the compound had four aromatic rings resulting in an increased PFI over the alkyl analogues in **Table 29**. Replacement of the PMB group with a saturated system, was predicted to improve the solubility, and increase the chance of the template progressing successfully towards a drug candidate. To help design saturated amide substituents for synthesis, the potential binding modes of methylamide (**4.188**) and PMB-amide (**4.191**) were investigated using MOE (**Figure 139**). Additionally, thought turned back to the X-ray structure of pyrazolopyridine fragment (**4.5**) (**Figure 140**). A second copy of pyrazolopyridine fragment (**4.5**) was seen π -stacking between Trp176 and the purine moiety of the substrate. Docking of PMB-amide (**4.191**) in mouse CD38 suggested that the PMB group could be π -stacking similarly to the second copy of pyrazolopyridine fragment (**4.5**). The substrate is mobile and can adapt to the position of the PMB substituent. A similar docking pose was obtained for methylamide (**4.188**).

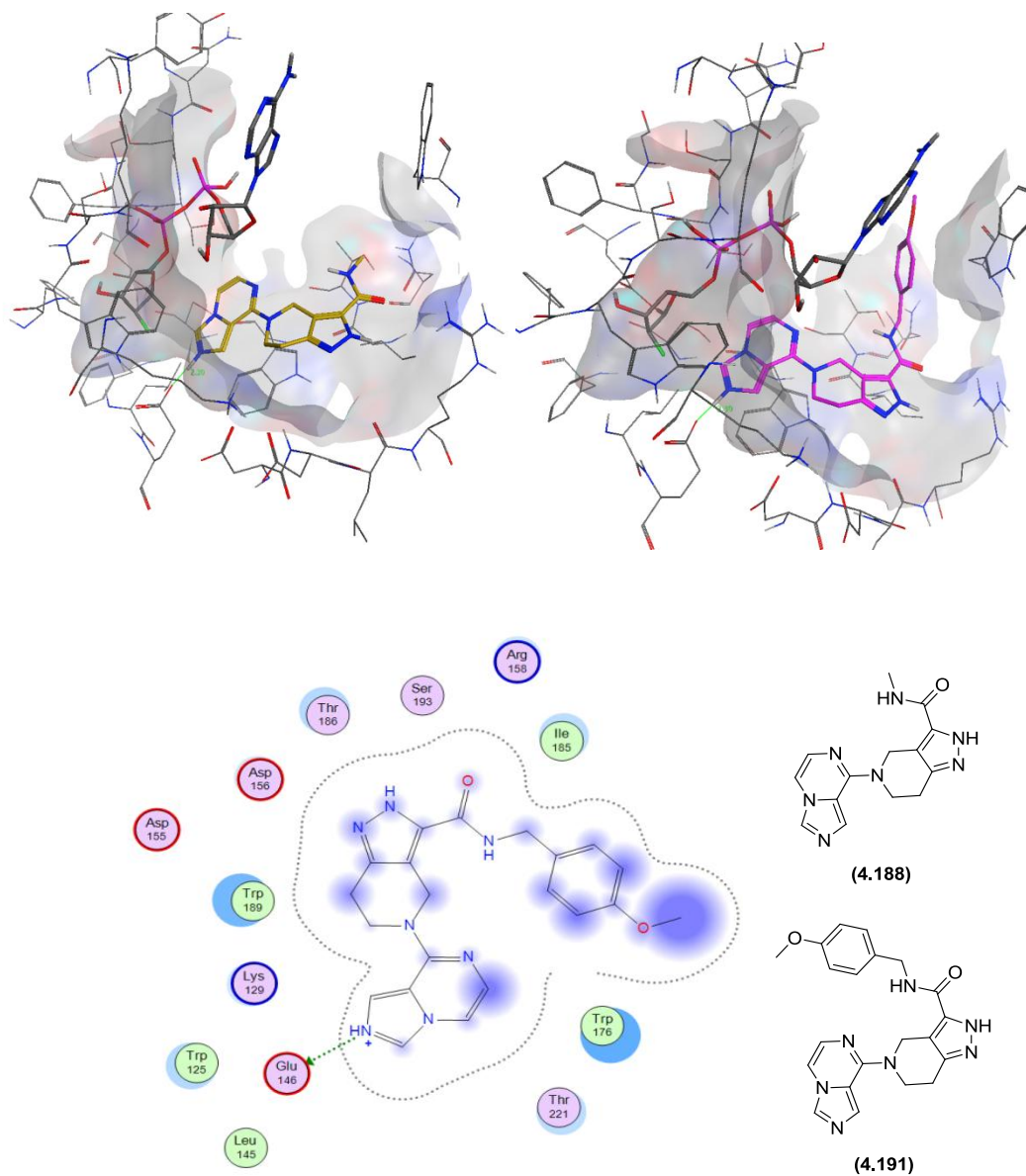


Figure 139 – Top left: Docking of methylamide (**4.188**) in mouse CD38; Top right: Docking of PMB (**4.191**) in mouse CD38; Bottom: Site view of docking of PMB (**4.191**) in mouse CD38.

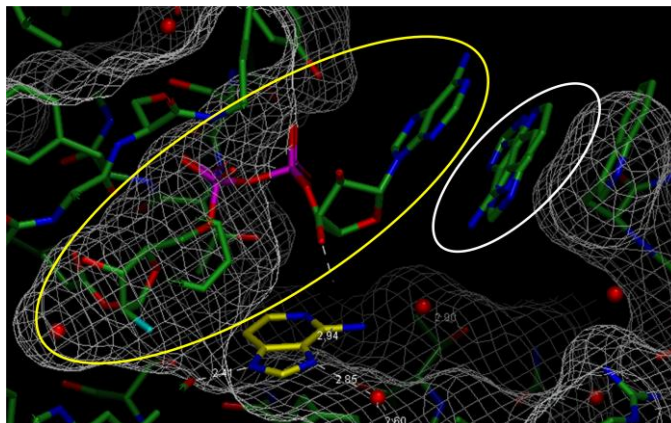


Figure 140 - X-ray crystal structure of imidazopyridine (**4.5**) PDB Code: 7HIZG (mouse). Imidazopyridine (**4.5**) binds uncompetitively with the truncated substrate and shows a second copy of (**4.5**) π -stacking between Trp176 and the purine moiety of the substrate (the orientation is undefined).

From the docking of PMB (**4.191**) and methyl amide (**4.188**) it can be seen that there is no limit on space in the region of the protein as it points out towards solvent. The main interactions to be exploited are with Trp176, which were highlighted by GRID maps at the beginning of the project (**Figure 109**). Introduction of saturated rings may pick up CH- π interactions with Trp176. Cyclopropyl analogues (**4.192**) and (**4.193**) (**Figure 141**) were of particular interest due to the partial sp^2 character associated with the alkyl ring. As the proposed rings increased in size the lipophilicity increased, so accordingly, THP (**4.195**) was designed in preference of the corresponding cyclohexyl analogue to address the issue of increasing lipophilicity. The oxygen atom was predicted to be tolerated as it would occupy a solvent exposed region and CH- π interactions could be made with Trp176 (**Figure 142**). Despite it being desirable to remove an aromatic ring, it was decided to include phenyl isostere thiophene (**4.196**) in the set. The thiophene ring was chosen for the following reasons. Firstly, the pendant amide moiety is expected to occupy a region of the protein where π -stacking interactions could be exploited to improve the binding affinity of the molecule to the protein. Secondly, the PMB-amide (**4.191**) was less efficient than the smaller alkyl substituents in **Table 29**. Docking suggested the methoxy group of PMB (**4.191**) (**Figure 139**) pointed out to solvent and was

unlikely to give any benefit to the binding affinity. Thirdly, in addition to this, the phenyl ring was predicted to be participating in π -stacking interactions, however, the size of the ring was unlikely to be important as it was not anticipated to be space filling. If thiophene (**4.196**) had comparable potency to PMB (**4.191**), the reduction in heavy atom count of thiophene (**4.196**) compared to PMB (**4.191**) would help improve the efficiency of binding. It is important to note that the designed compounds all had the methylene linker apart from cyclopropyl (**4.192**). Directly attached larger rings were predicted to clash with the protein unless significant movement of the rest of the molecule was observed.

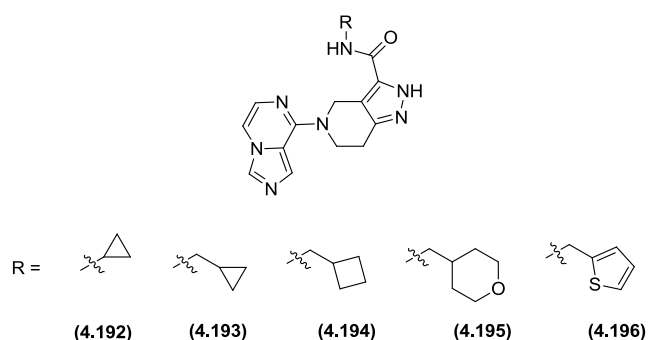


Figure 141 – Amide compounds prioritised for synthesis from docking.

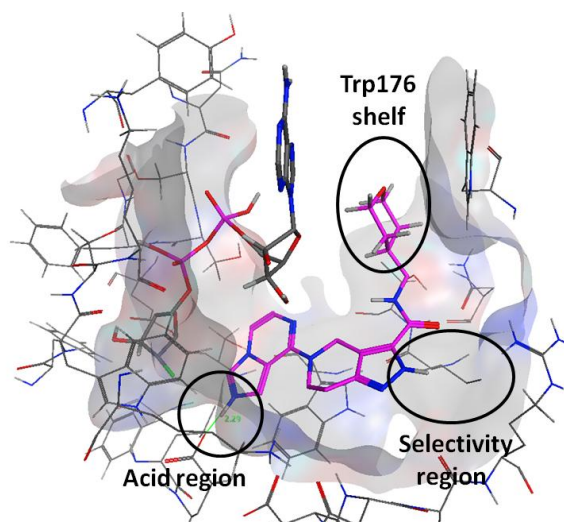
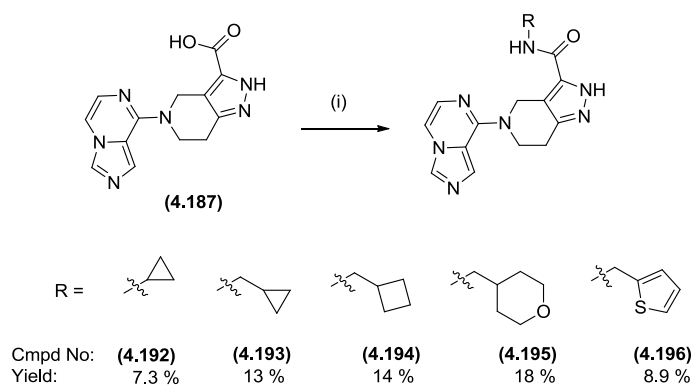
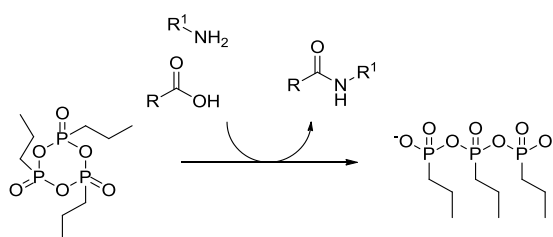


Figure 142 – Docking of THP (**4.195**) in CD38 (mouse). The areas of the pocket highlighted by GRID maps as regions to achieve profitable ligand-protein interactions are circled.

Carboxylic acid (**4.187**) was available within our laboratories for synthesis of amide analogues (**4.192**) – (**4.196**). The amides were synthesised using propane phosphonic acid anhydride (T3P[®]) coupling conditions (**Scheme 37**). T3P[®] is a mild, non-hazardous, easy to use reagent and only gives water soluble by-products (**Scheme 38**).²²⁸⁻²³⁰ Although not important for these particular synthetic examples, T3P[®] is suitable for synthesis of chiral analogues as minimal racemisation is observed.²³¹



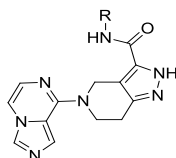
Scheme 37 – Synthesis of amide analogues (**4.192**) – (**4.196**). Reagents and conditions: (i) T3P[®] (50% in EtOAc), Et₃N, 2-MeTHF, RT – 40 °C.



Scheme 38 - T3P[®] amide coupling mechanism.

Biochemical screening against both mouse and human CD38 constructs showed that the saturated analogues (**4.192**) – (**4.195**) gave no potency or efficiency advantage over methyl amide (**4.188**) (**Table 30**). As predicted when designing the analogues, although thiophene (**4.196**) was less efficient than methyl amide (**4.188**), it was more potent and more efficient than PMB (**4.191**) (particularly against the mouse

construct). Thiophene (**4.196**) was of elevated interest as it showed good human potency and efficiencies (h pIC₅₀ = 7.1, LE = 0.36, LLE_(AT) = 0.36). The main concern with thiophene (**4.196**) was the number of aromatic rings and the associated measured PFI of 6.3 (cPFI = 6.2), which would predict sub-optimal aqueous solubility, borne out in the measured CLND solubility of 18 µg/mL.



Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)	cPFI
(4.188)	Me	6.0	0.37	0.49	6.7	0.42	0.53	3.5
(4.191)		6.7	0.31	0.33	6.4	0.29	0.32	6.2
(4.192)		5.6	0.32	0.38	6.3	0.36	0.42	4.2
(4.193)		5.6	0.31	0.35	7.1	0.39	0.43	4.6
(4.194)		5.7	0.30	0.31	6.8	0.36	0.37	5.1
(4.195)		5.7	0.28	0.36	5.7	0.28	0.36	4.0
(4.196)		7.1	0.36	0.36	7.0	0.36	0.36	6.2

Table 30 – Biochemical data for second iteration of amide analogues.

An additional observation from the biochemical data for the amide analogues (**Table 30**), was that PMB (**4.191**), THP (**4.195**) and thiophene (**4.196**) were each equipotent against both mouse and human, or had higher human potency. Throughout this programme of work, the general trend has been higher mouse biochemical potency. Overlay of the crystal structures of mouse and human protein showed significant variations in the shape of the protein pocket in the region the amides grow into. The amide analogues were docked in the mouse protein and were predicted to form interactions with Trp176. It can be seen from the overlay in **Figure 143** that the

location of Trp176 varies substantially between mouse and human enzymes. Previously, many of the fragment analogues synthesised were unlikely to reach far enough across the protein pocket, to make desirable interactions in the Trp176 region. This may explain why the larger amide analogues were equipotent against both mouse and human, or had higher human potency.

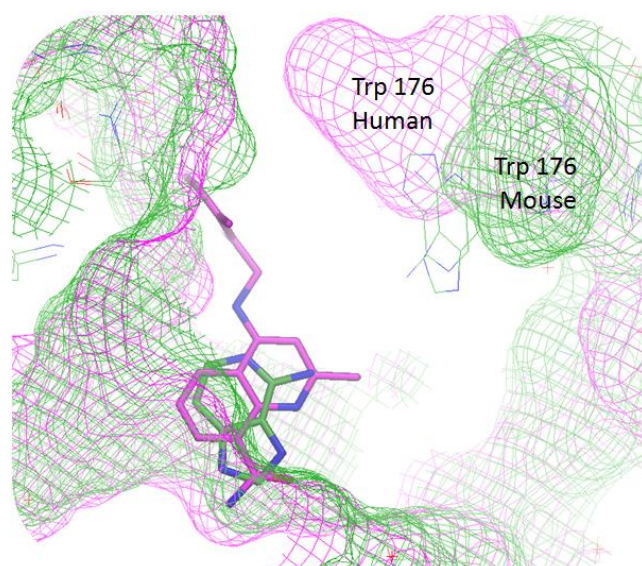


Figure 143 – Overlay of mouse (green) and human (pink) protein surfaces with imidazopyridine (**4.5**).

Despite repeated attempts to obtain X-ray crystal structures of numerous imidazopyrazine analogues, none had provided a liganded CD38 structure. The only X-ray crystal structure available to help guide the docking was of imidazopyridine (**4.5**). Thiophene (**4.196**) was docked using MOE into the human protein (**Figure 144**). Docking of thiophene (**4.196**) in human CD38 showed that the imidazopyrazine core was likely to rotate clockwise in the protein pocket maintaining the π -stacking interactions with Trp189. The protonated imidazole nitrogen was predicted to form a hydrogen bond to the side chain of Glu146. The docking suggested the pyrazolopiperidine ring system made good VdW contacts with the CD38 protein surface, and picked up a hydrogen bond between the pyrazole NH to the backbone carbonyl of Leu157. An intramolecular hydrogen bond between the

amide NH and the pyrazole nitrogen could form a pseudo 5-membered ring and would help position the thiophene ring in the area which is filled by Trp176 in the mouse CD38 construct. The thiophene ring appeared to make good contacts with the protein surface, and had the potential to form a π -H interaction with the backbone hydrogen of Asp175. The difference in binding modes of the amide analogues in mouse (**Figure 142**) and human protein (**Figure 144**) could account for the sudden change in trend observed between relative potency in the two constructs.

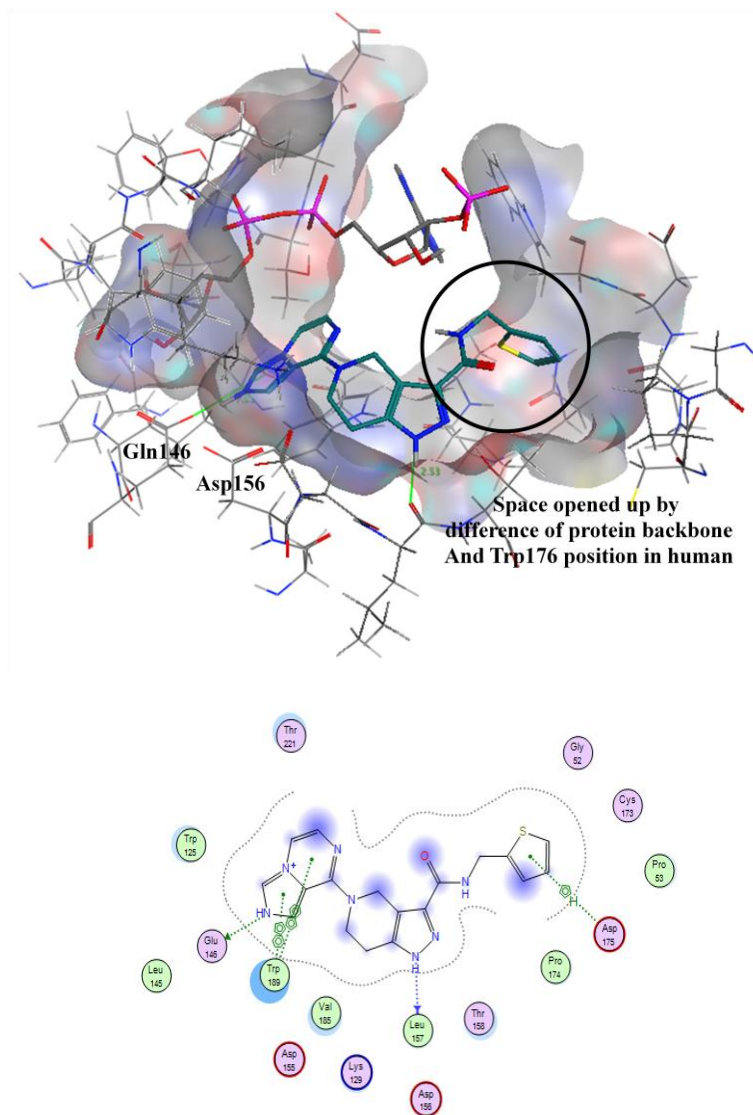


Figure 144 – Top image: Docking of thiophene (**4.196**) in human CD38 protein;
Bottom image: Site image showing ligand-protein interactions for docking of thiophene (**4.196**) in human CD38 protein.

Amide substitution from the pyrazolopiperidine ring system gave increased potency over the truncated analogues, particularly in the human CD38 protein. Before exploration of additional amide analogues of methylamide (**4.188**), comparative isoindoline amide analogues were designed to ensure that the pyrazolopiperidine ring system was preferable from a potency perspective (**Figure 145**).

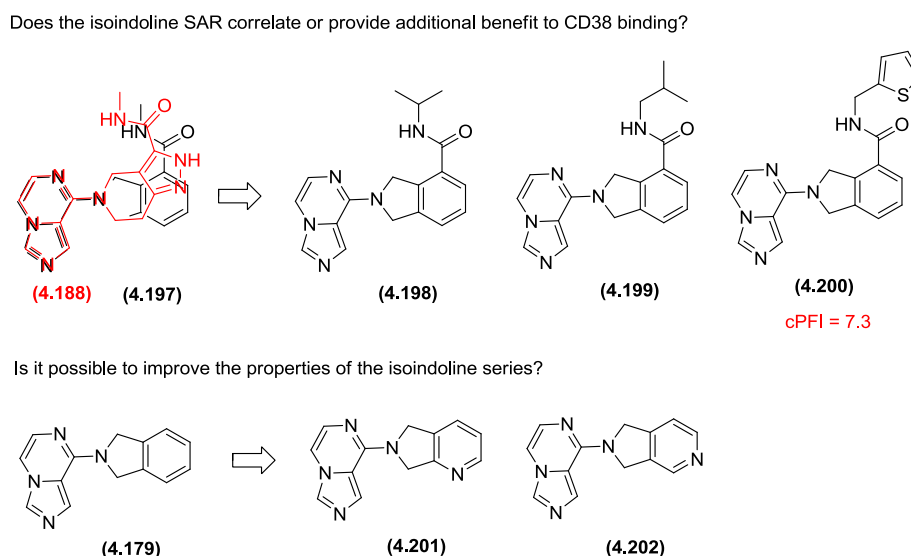
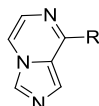


Figure 145 – Comparison of isoindoline and pyrazolopiperidine ring systems attached to the imidazopyrazine core.

The pyrazolopiperidine ring system has more 3D character than the corresponding isoindoline, suggesting better solubility for the pyrazolopiperidine series compared to the isoindoline series. This hypothesis was confirmed by the relative cPFI for the two structural templates ((**4.186**) cPFI = 4.0 compared to (**4.179**) cPFI = 6.0 (**Table 31**)). Despite this, the amide vector was similar in both the isoindoline and pyrazolopiperidine templates (**Figure 145**) and it was felt important to explore the relative potency of the two templates. The unsubstituted isoindoline (**4.179**) was more potent and had greater LE compared to the unsubstituted pyrazolopiperidine (**4.186**) (**Table 31**). Therefore, it was logical to assume that amide substitution from the isoindoline could give more potent analogues than the pyrazolopiperidine amide analogues in **Table 30**. If this improvement in potency was observed, the isoindoline

template would require additional polarity, or reduction in sp² character to improve the aqueous solubility for progression towards a drug candidate. Aza-isoindoline analogues (**4.201**) and (**4.202**) were prepared to assess if polarity could be tolerated in the phenyl ring, alongside isopropyl (**4.198**) and isobutyl amide (**4.199**). The decision was made not to synthesise amides with aromatic substituents, despite aryl analogues being most active for the pyrazolopiperidine series, as the PFI would be expected to be outside the desired range (e.g. thiophene amide (**4.200**) cPFI = 7.3).

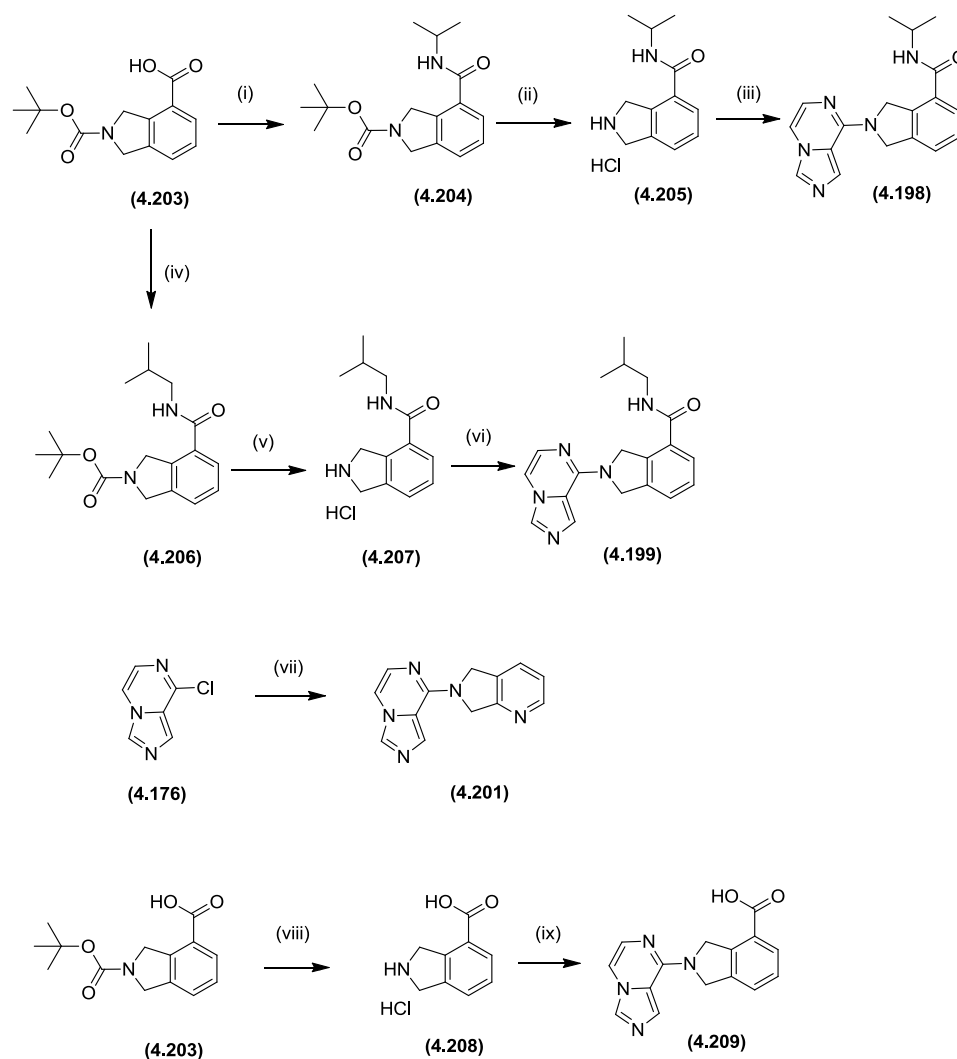


Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)	PFI
(4.186)		4.8	0.35	0.41	6.7	0.48	0.55	4.0
(4.179)		5.5	0.42	0.38	6.7	0.51	0.47	6.0

Table 31 – Comparison of potency, binding efficiency and PFI for pyrazolopiperidine analogues compared with the isoindoline analogues.

Isopropyl amide (**4.198**) and *iso*-butyl amide (**4.199**) were synthesised using a three step process (**Scheme 39**). Carboxylic acid (**4.203**) was available in our laboratories and was used in T3P[®] promoted amide coupling reactions to form Boc protected isopropyl amide (**4.204**) and isobutyl amide (**4.206**). Removal of the Boc group using 4N HCl in dioxane was performed to give isoindolines (**4.205**) and (**4.207**) as the hydrochloride. These species were used directly in S_NAr reactions with 8-chloroimidazo[1,5-*a*]pyrazine (**4.176**), using conditions previously described, to give the target compounds (isopropyl amide (**4.198**) and *iso*-butyl amide (**4.199**)). Aza-isoindoline (**4.201**) was prepared from reaction of commercially available 6,7-dihydro-5*H*-pyrrolo[3,4-*b*]pyridine with 8-chloroimidazo[1,5-*a*]pyrazine (**4.176**). As an additional comparator to the pyrazolopiperidine analogues, carboxylic acid (**4.209**) was prepared from Boc deprotection of (**4.203**), followed by S_NAr with 8-

chloroimidazo[1,5-*a*]pyrazine (**4.176**). The alternative aza-indoline analogue (**4.202**) (**Figure 145**) was prepared elsewhere within our laboratories.²³²

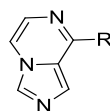


Scheme 39 – Synthetic routes to isoindoline analogues, isopropyl amide (**4.198**), isobutyl amide (**4.199**), aza-isoindoline (**4.201**) and carboxylic acid (**4.209**).

Reagents and conditions: (i) Propan-2-amine, T3P[®] (50% in EtOAc), Et₃N, 2-MeTHF, RT, 55%; (ii) 4.0 N HCl in dioxane, DCM, RT, 88%; (iii) 8-chloroimidazo[1,5-*a*]pyrazine, K₂CO₃, DMSO, 60 °C, 15%; (iv) 2-methylpropan-1-amine, T3P[®] (50% in EtOAc), Et₃N, 2-MeTHF, RT, 27%; (v) 4.0 N HCl in dioxane, DCM, RT, 92%; (vi) 8-chloroimidazo[1,5-*a*]pyrazine, K₂CO₃, DMSO, 60 °C, 33%; (vii) 6,7-dihydro-5H-pyrrolo[3,4-*b*]pyridine, K₂CO₃, DMSO, 60 °C, 16%; (viii) 4.0

N HCl in dioxane, DCM, RT, then heating at 45 °C, 93%; (ix) 8-chloroimidazo[1,5-*a*]pyrazine, K₂CO₃, DMSO, 60 °C, 55%.

Biochemical data for the isoindoline analogues showed carboxylic acid analogue (**4.209**) gave additional potency in both the human and mouse enzyme assays compared to pyrazolopiperidine carboxylic acid (**4.187**) (Table 32). Interestingly, isopropyl amide isoindoline (**4.198**) (h pIC₅₀ = 5.4, LE = 0.31, LLE_(AT) = 0.33; m pIC₅₀ = 7.0, LE = 0.40, LLE_(AT) = 0.42) gave increased mouse potency and efficiencies compared to pyrazolopiperidine isopropylamide (**4.190**), but exhibited a slightly lower human potency. Growth of the amide substituent to give *iso*-butyl amide (**4.199**) (h pIC₅₀ = 5.7, LE = 0.31, LLE_(AT) = 0.31; m pIC₅₀ = 8.0, LE = 0.44, LLE_(AT) = 0.43) gave the same trend, with increased mouse potency compared to pyrazolopiperidine isopropylamide (**4.190**), but with no change to the human potency. Pleasingly, the two aza-isoindoline analogues (**4.201**) and (**4.202**) gave comparable or improved binding affinities compared to isoindoline (**4.179**), resulting in improved LLE_(AT).



Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)	PFI
(4.187)		4.7 ^a	0.31	0.37	6.1	0.40	0.46	2.6
(4.190)		5.7	0.33	0.39	6.2	0.35	0.42	4.6
(4.179)		5.5	0.42	0.38	6.7	0.51	0.47	6.0
(4.209)		5.3	0.35	0.35	7.3	0.48	0.48	3.1
(4.198)		5.4	0.31	0.33	7.0	0.40	0.42	5.6
(4.199)		5.7	0.31	0.31	8.0	0.44	0.43	6.2
(4.201)		5.4	0.41	0.49	6.6	0.50	0.58	4.1
(4.202)		6.1	0.46	0.54	7.3	0.56	0.63	3.9

Table 32 – Biochemical data for pyrazolopiperidine amide analogues compared to isoindoline amide and aza-isoindoline analogues. ^aThis compound has only been reported active in 2 of the 4 screening occasions (pIC₅₀ = <4.52, <4.52, 4.87, 4.61).

Biochemical data for the isoindoline analogues in **Table 32** showed that the mouse activity was consistently higher than the human activity. This difference in activity against the constructs was particularly pronounced for the three substituted isoindolines, carboxylic acid (**4.209**), *i*-Pr amide (**4.198**) and *i*-Bu amide (**4.199**). In an attempt to rationalise this data, docking of the isoindoline analogues was

investigated. The carboxylic acid or amide moiety are predicted to occupy the area of the CD38 protein pocket where there is a significant change to the position of Trp176. Docking of isoindoline isopropyl amide (**4.198**), in mouse CD38, suggested a hydrogen bond to Glu146 from the protonated imidazole nitrogen of the imidazopyrazine core. The isoindoline ring was predicted to sit planar on a flat shelf, with the amide substituent sitting close to and making good CH- π interactions and VDW contacts with Trp176.

A 2D overlay of isoindoline methyl amide (**4.197**) and with pyrazolopiperidine methyl amide (**4.188**) showed that the preferred orientation of the amide moiety would differ for the two templates. With the assumption that the pyrazolopyrimidine core occupies an identical orientation for the two templates, the amide group of the isoindoline methyl amide (**4.197**) sits closer to Trp176 than the amide in pyrazolopiperidine methyl amide (**4.188**). This close proximity to Trp176 predicted for the isoindoline series could account for the pronounced mouse activity measured for these analogues. To bind into the human CD38 protein, the amide moiety would need to flip and occupy the other orientation, pointing the amide NH away from the core (**Figure 146**). From the 2D overlay of isoindoline methyl amide (**4.197**) with pyrazolopiperidine methyl amide (**4.188**), it can be seen that the amide nitrogen substituent occupies a very similar position. This could account for the similarity in human activity observed between the isoindoline and pyrazolopiperidine templates.

From the synthesis of isoindoline analogues in **Table 32** it had been demonstrated that the isoindoline template had no potency advantage for the human CD38 construct over the pyrazolopiperidine series. The isoindoline had less sp^3 character than the pyrazolopiperidine analogues which contributed to the reduced solubility of isoindoline *i*-Pr amide (**4.198**) (CLND sol = 34 $\mu\text{g/mL}$) compared to the pyrazolopiperidine *i*-Pr amide (**4.190**) (CLND sol = 143 $\mu\text{g/mL}$). The increased lipophilicity of the isoindoline analogues compared to the pyrazolopiperidine series gave a decrease in $LLE_{(AT)}$. Synthesis and screening of aza-isoindoline analogues (**4.201**) and (**4.202**) did show that the pyridine nitrogen could be tolerated which

would help improve the $LLE_{(AT)}$, however, from docking alone it was difficult to predict which location would be most favourable for the nitrogen to determine which substituted aza-isoindoline analogues to prioritise for synthesis. With these considerations in mind, the decision was made to revert back to optimisation of the pyrazolopiperidine amide substitution, as the physicochemical properties of this template would facilitate additional growth to increase the CD38 human potency, whilst maintaining good efficiencies.

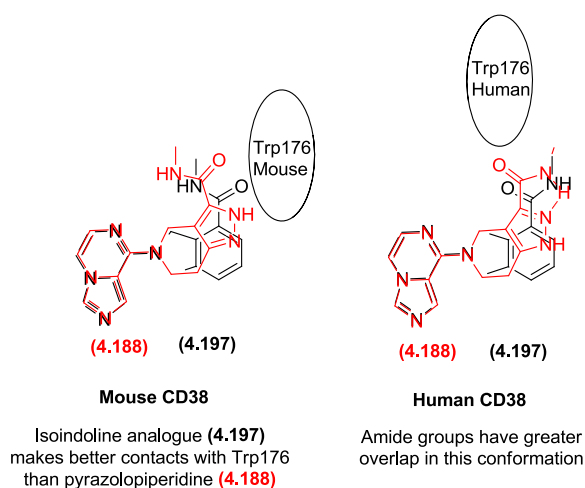
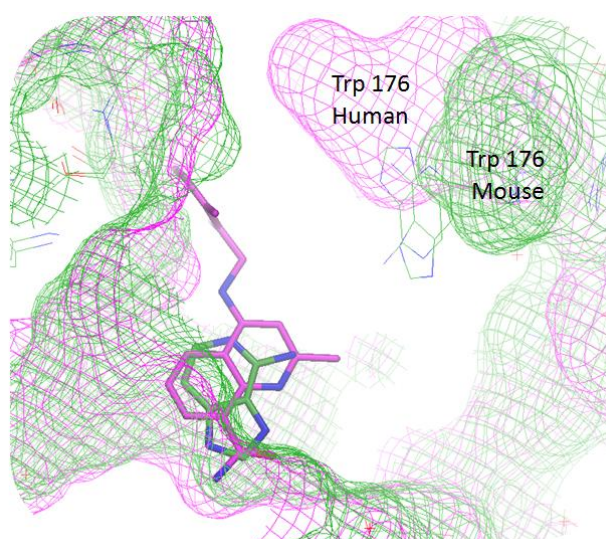
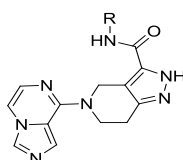


Figure 146 – Top: Overlay of human (pink) and mouse (green) protein surfaces highlighting differences in location of Trp176; Bottom: Proposed binding mode of

isoindoline methyl amide (**4.197**) overlaid with pyrazolopiperidine (**4.188**) in mouse and human CD38 protein.

4.9 Optimisation of the Amide Substituent of the Pyrazolopiperidine Series Through Array Design

The original fragment hit imidazopyridine (**4.5**) has been optimised and grown through careful design using docking, and has resulted in identification of potent CD38 inhibitors (**Table 33**). It is clear that the most potent analogues against human CD38 are thiophene (**4.196**) and PMB (**4.191**), both of which have an aryl substituent pendant from the amide nitrogen, which is detrimental to PFI. The low CLND solubility for these aryl amide analogues was a concern for progression of the template towards a drug molecule.



Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)
(4.188)	Me	6.0	0.37	0.49	6.7	0.42	0.53
(4.191)		6.7	0.31	0.33	6.4	0.29	0.32
(4.192)		5.6	0.32	0.38	6.3	0.36	0.42
(4.193)		5.6	0.31	0.35	7.1	0.39	0.43
(4.194)		5.7	0.30	0.31	6.8	0.36	0.37
(4.195)		5.7	0.28	0.36	5.7	0.28	0.36
(4.196)		7.1	0.36	0.36	7.0	0.36	0.36

Table 33 – Biochemical data for second iteration of amide analogues.

In an attempt to find an alternative amide substituent to improve the solubility profile, an array of compounds was designed by the author for synthesis within an array team elsewhere within our laboratories. The aim of the array was to try to improve the physicochemical properties, whilst maintaining or improving the human potency. It was agreed that the synthesis of 30 amide analogues would be feasible. As no X-ray crystal structure had been obtained for the imidazopyrazine series, docking was used to guide the design and selection of analogues. It was important to ensure that the array targeted a diverse range of compounds, exploring a number of potential poses.

The array was designed by computationally evaluating all the primary and secondary amine monomers with a molecular weight < 300 available from our laboratories' compound collection. The molecular weight cut-off was kept at this level to allow for molecular weight from protecting groups (such as Boc and PMB). The amine monomers were computationally coupled to the carboxylic acid (**4.187**), before removal of the protecting groups. Properties were calculated and the following filters were applied: $-1.5 \leq \text{cLogP} \leq 3$; $\text{MW} < 400$; $\# \text{ rotatable bonds (rb)} \leq 5$. The lower cLog P filter was applied to ensure that the compounds were not too polar, which could give a large desolvation penalty when binding to the CD38 protein. The upper cLog P cut-off was applied to prevent the compounds becoming too lipophilic, as high lipophilicity was likely to be associated with an increased chance of promiscuity and poor solubility for the compounds represented by high PFI values. The size of the molecules was capped by the MW cut-off as the aim was to find an efficient molecule. A rotatable bond cut-off was used to remove large flexible chains from the set, which would have a high entropic penalty to overcome on binding to the protein. The enumerated amides were visually inspected to remove analogues which had potentially unstable functionality (e.g. acetals, methyl esters). Undesirable functionality such as unsubstituted thiazoles and pyrroles were removed due to their potential reactivity/instability. The filters and visual inspection reduced the enumerated analogues down to 75 compounds.

The 75 enumerated products were docked into human CD38 (PSILO code: 1GECC) with substrate present.¹⁶⁵ Prior docking of imidazopyrazine analogues had suggested the most likely docking poses participated in a hydrogen bond interaction from the protonated imidazole nitrogen of the imidazopyrazine core to the side chain Glu146. This knowledge was incorporated into the docking of the enumerated amide analogues. The 75 enumerated products were protonated on the imidazole nitrogen of the imidazopyrazine core and docked into human CD38.

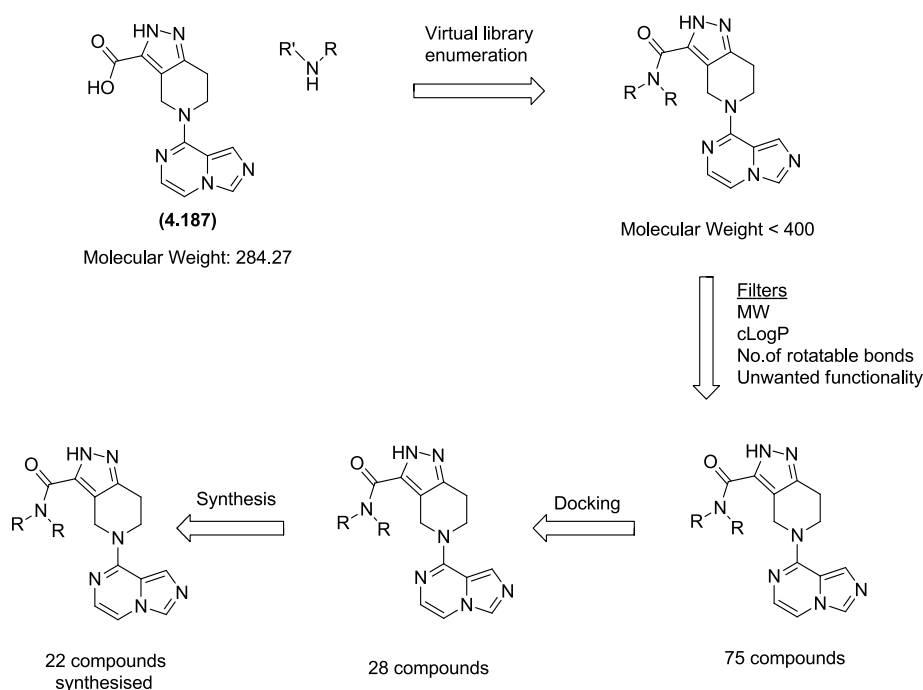


Figure 147 – Pyrazolopiperidine amide array design.

The docking poses produced were examined in collaboration with the programme modeller.¹⁶⁵ The amide analogues orientated with the protonated nitrogen of the imidazole from the imidazopyrazine core interacting with Glu146 were selected. A total of 51 docking poses satisfied this condition, of which 40 were unique compounds (duplicates came from expansion of tautomers and stereoisomers). The 40 compounds were examined by eye and close analogues removed giving a selected list of 24 amide analogues for synthesis. As X-ray crystallography of the

imidazopyrazine series had been unsuccessful, the predicted docking pose could be incorrect. To allow for this, a further 4 enumerated amide compounds were selected from those that did not dock well in the predicted binding orientation. These compounds were added to the set of compounds for synthesis to give additional diversity and act as a control for the hypothesis that the imidazole-Glu146 pose is preferred. The 28 selected amide compounds selected for synthesis are shown **Figure 148**. These could be grouped into six clusters based around structural similarity of the pendant amine.

Cluster 1: Secondary cyclic aliphatic amines – The four monomers all had halogens or π -systems (nitrile (**4.213**) or alkene (**4.211**)) which could potentially pick up interactions with the π -system of Trp176. The protons on the carbon rings could also point into the Trp176 π -system. Diversity in ring size was included.

Cluster 2: Primary aliphatic amines – This cluster explored branched, bulky amides.

Cluster 3: Primary aliphatic cyclic amines – A diversity of ring sizes was included. CH interactions were predicted with the π -system of Trp176.

Cluster 4: Methylene-linked alicycles – Docking suggested oxygen could be tolerated in the ring to help control the lipophilicity. A diversity of ring systems were explored.

Cluster 5: Methylene-linked 5-membered aryl heterocycles – Although the main focus of the array was to identify amides with a high percentage of sp^3 character, aryl analogues were included in the set to probe less lipophilic heterocycles. The methylene linked thiophene amide (**4.196**) was the most potent amide synthesised and the corresponding analogues with an additional methyl group on the methylene carbon linker or the amide nitrogen were prioritised for synthesis (from amine monomers (**4.232**) and (**4.233**)).

Cluster 6: Additional 4 compounds chosen for diversity – The first amine explored branching from the β -carbon from the amine (methoxy (**4.234**)). The second monomer explored the feasibility of adding another hydrogen bond acceptor to the heterocyclic ring to help lower the lipophilicity (oxadiazole (**4.235**)). The morpholine analogue (**4.236**) explored adding oxygen to a secondary cyclic aliphatic amine, as none of these were present in cluster 1. Finally, a six-membered aryl ring with a methylene linker to the amine nitrogen was included (pyrimidine (**4.237**)). To date, no six-membered heterocycles had been screened and none were present in cluster 5.

Before synthesis, the cLog P profile of the 28 enumerated compounds was examined to ensure that a range of lipophilicity was included in the final set of compounds for synthesis. **Figure 149** shows that the cLog P is evenly distributed between -1.5 and 2.5.

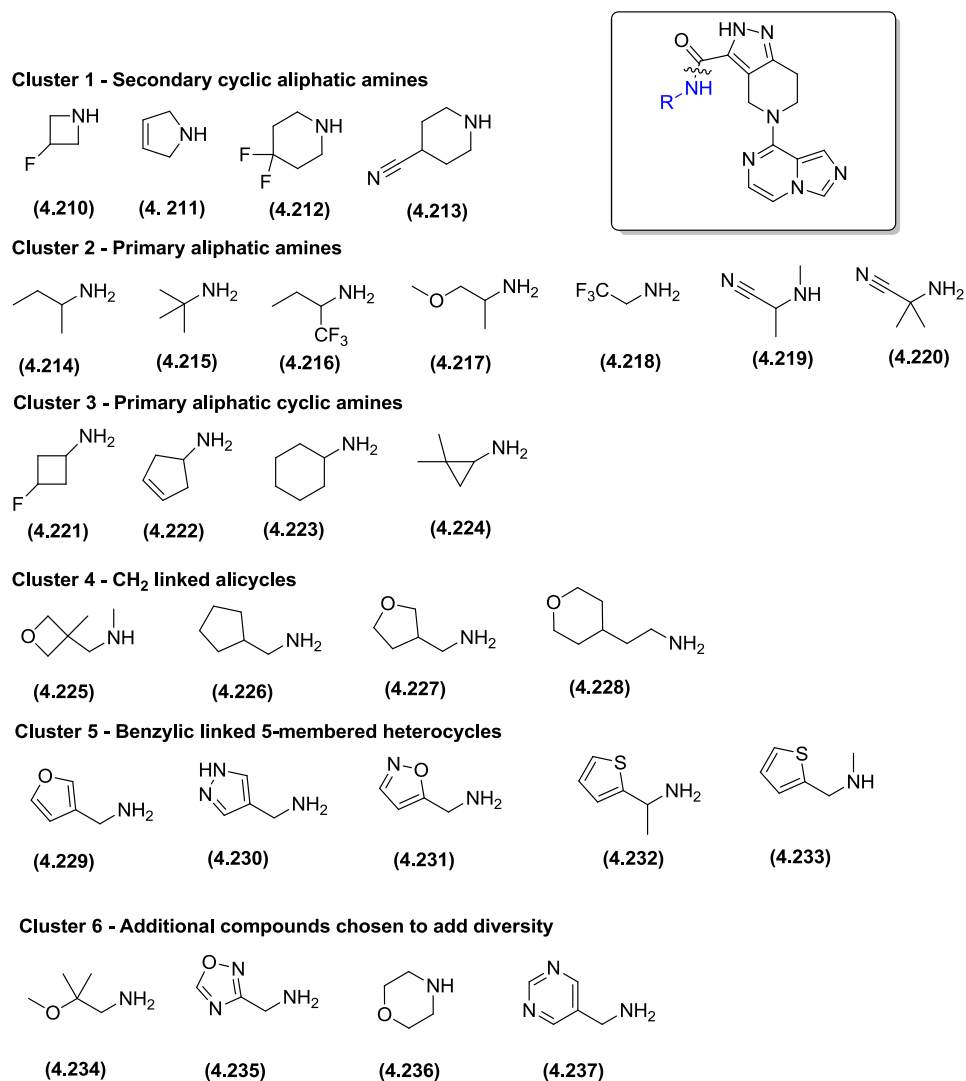


Figure 148 – 28 Amine monomers selected for the pyrazolopiperidine amide array.
Compound numbers correspond to final amides.

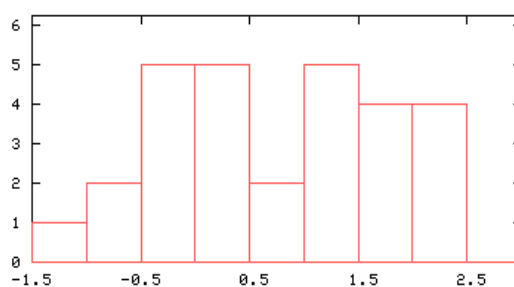
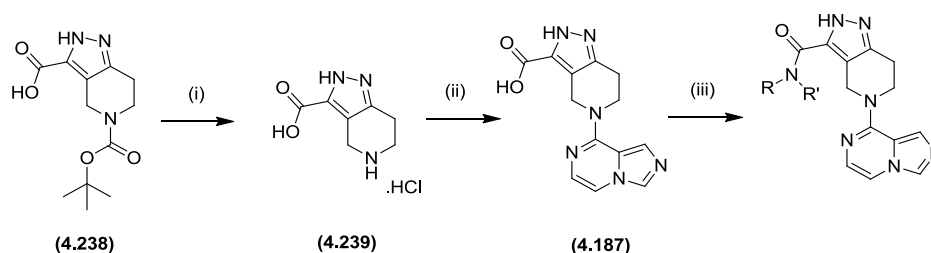


Figure 149 – cLog P range for 28 array analogues.

Scale-up of the synthesis of carboxylic acid (**4.187**) was performed to facilitate the array (**Scheme 40**). The Boc group was removed from pyrazolopiperidine (**4.238**) under acidic conditions to afford piperidine (**4.239**), which was coupled to chloroimidazo[1,5-*a*]pyrazine (**4.176**) in DMSO with sodium bicarbonate as the base. Sodium bicarbonate was used in place of the usual K_2CO_3 as it had been shown to give an improved yield for related analogues. Synthesis of the 28 amides was attempted elsewhere within our laboratories using amide coupling conditions.



Scheme 40 – Synthesis of carboxylic acid (**4.176**) for use in amide array. Reagents and conditions: (i) isopropanol, 5.0-6.0 M HCl in isopropanol, 65 °C, 18 h, 87%; (ii) 8-chloroimidazo[1,5-*a*]pyrazine, $NaHCO_3$ DMSO, 70 - 80 °C, 68 h, 96% (in two batches); (iii) amide coupling in the arrays team.

Not all the monomers were successful in the amide coupling reactions. Aliphatic nitrile analogues (**4.213**) and (**4.220**) were not attempted in this preparative campaign, due to knowledge from the array team that these amines tend to fail in amide coupling reactions and the potential reactivity to formation of glutathione adducts. The *t*-butyl (**4.215**), CF_3 (**4.216**), oxetane (**4.225**) and oxadiazole (**4.235**) building blocks failed to produce sufficient product for successful purification. The decision was made to screen the obtained compounds through the biochemical assay and use the data to decide on the importance of re-synthesising those that had failed.

Of the 22 amide compounds that were synthesised and screened in the biochemical assay, all showed human activity apart from one compound, the morpholine amide (from morpholine monomer (**4.236**) in **Figure 150**). This helped validate the

docking pose, as the morpholine analogue did not dock in the desired pose and was selected for the array for this purpose. Likewise, the amides synthesised from methoxy (4.234) and pyrimidine (4.237) monomers had human LE and $LLE_{(AT)} < 0.3$. It was interesting to see that cluster 1 analogues (secondary aliphatic amines) all fitted into this low binding efficiency category (coloured in orange in **Figure 150**). The most interesting amide analogues were those with ligand efficiencies equal to or greater than 0.3. These compounds are shown in green in **Figure 150**. Notably, the most efficient amide analogues were all synthesised from primary amine monomers.

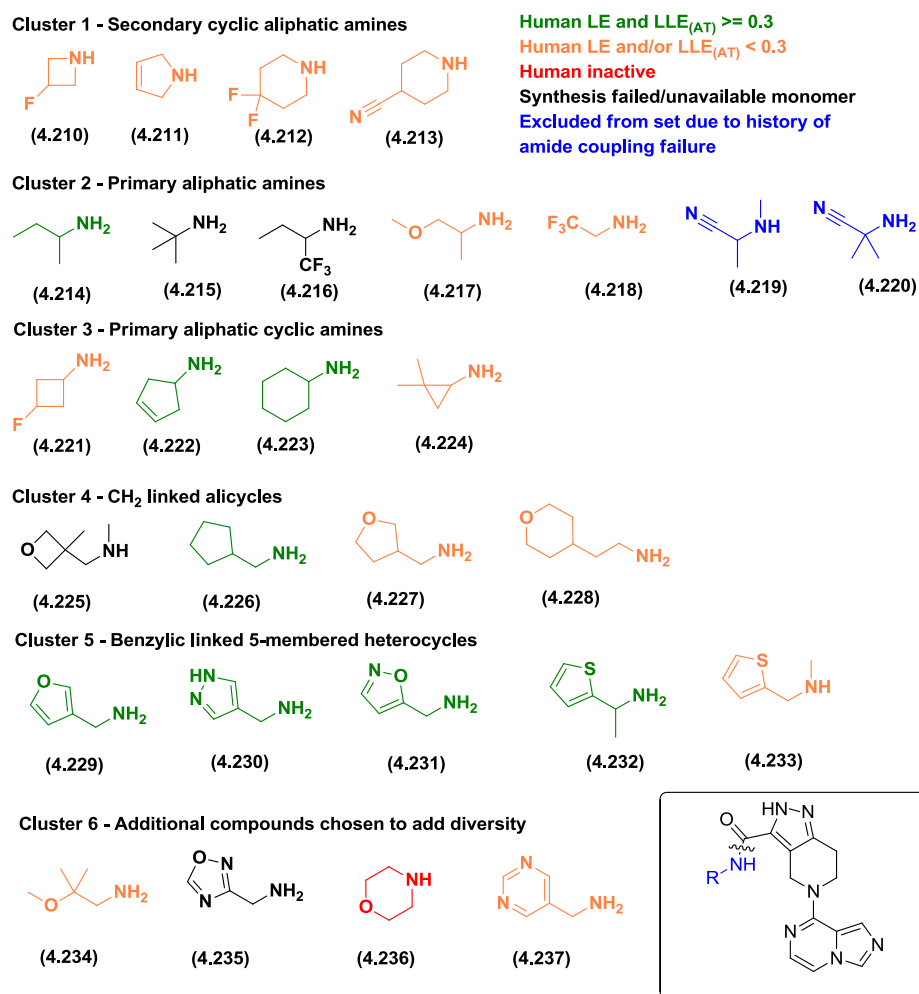
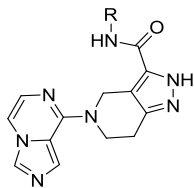


Figure 150 – Ranking of amide array compounds from human biochemical data. Compound numbers correspond to final amides.

Biochemical data for the amide analogues with both LE and $LLE_{(AT)}$ of 0.3 or greater is shown in **Table 34**. Thiophene (**4.196**) still remained the most active and ligand efficient amide analogue against the human CD38 construct. Methyl substitution from the methylene linker (methyl (**4.232**)) gave a drop in both human and mouse activity and an increase in PFI. The three compounds of most interest are oxazole (**4.231**), pyrazole (**4.230**) and cyclopentene (**4.222**) due to these species having the best balance of potency, efficiency, and aqueous solubility, which meets the criteria set out at the beginning of this CD38 fragment programme (**Figure 151**).



Cmpd No.	R Group	h pIC ₅₀	h LE	h LLE _(AT)	m pIC ₅₀	m LE	m LLE _(AT)	PFI	CLND µg/mL
(4.196)		7.1	0.36	0.36	7.0	0.36	0.36	6.3	18
(4.223)		6.1	0.31	0.32	6.1	0.31	0.32	5.9	75
(4.226)		6.3	0.32	0.33	6.6	0.33	0.34	6.0	76
(4.232)		6.7	0.33	0.34	6.3	0.31	0.32	6.9	79
(4.214)		5.7	0.31	0.35	5.9	0.32	0.36	5.1	119
(4.229)		6.7	0.34	0.39	7.0	0.36	0.41	5.9	≥ 130
(4.230)		6.2 (n=1)	0.31	0.42	6.7 (n=1)	0.34	0.45	4.6	≥ 133
(4.222)		6.2	0.33	0.39	6.8	0.36	0.42	5.1	≥ 139
(4.231)		6.6	0.33	0.45	6.9	0.35	0.46	5.1	≥ 178

Table 34 - Mouse and human biochemical data for the amide analogues with both LE and LLE_(AT) of 0.3 or greater, combined with human and mouse potency of greater than pIC₅₀ = 6.0. The increasing solubility correlates with improved LLE_(AT) particularly when considering the mouse data.

Human and mouse pIC₅₀ ≥ 6.0

LE and LLE_(AT) > 0.3

hERG pIC₅₀ ≤ 5.0 tbd

P450 pIC₅₀ ≤ 5.0 tbd

PFI < 6

Solubility ≥ 100 µg/mL

Figure 151 – Fragment effort criteria.

4.10 Profiling Representative Analogues from the Imidazopyrazine Series in hERG and P450 Screens

To determine if the programme criteria had been met, representative analogues from the isoindoline and pyrazolopyrimidine sub-series of the imidazopyrazine template were profiled in hERG and P450 screens. Rat potency data was also tabulated as a rat probe was a further aspiration from this research to help enable target validation studies. Previous attempts within our laboratories to identify potent analogues against the rat construct had proved challenging.

Data for three isoindoline analogues was collected (**Table 35**). Bromide (**4.174**) gave very good potency against all three CD38 constructs and low hERG liability. Disappointingly, activity was observed against the P450 enzyme 3A4 and it possessed low measured aqueous solubility (14 µg/mL). Fluoro-substituted isoindoline (**4.180**) also had poor solubility and moderate P450 activity, but also gave moderate hERG activity. *i*-Pr Amide (**4.198**) was lower in activity against mouse and rat, had low solubility and moderate hERG activity, but was inactive against P450 3A4. As *i*-Pr-amide (**4.198**) was inactive against P450 3A4, this gave encouragement that this was not a series specific issue and that such selectivity matters could be overcome with further optimisation

Four pyrazolopiperidine analogues with sufficient compound availability were profiled. All four analogues met the desired potency criteria ($pIC_{50} > 6.0$) across all three constructs. Thiophene (**4.196**) and cyclopentyl (**4.226**) showed moderate hERG activity, lower than desired solubility (CLND < 100 µg/mL) and moderate P450 3A4 activity. The two most encouraging compounds were cyclopentene (**4.222**) and pyrazole (**4.230**). As described previously, both analogues had good CLND solubility (CLND > 100 µg/mL). No hERG liability was present for these two analogues, but they still displayed moderate P450 3A4 activity. Despite further work being required to understand what features of the molecule were contributing to the

CONFIDENTIAL – Property of GSK – Do Not Copy

P450 activity, five of the six criteria set out at the beginning of the programme had been successfully met (**Figure 151**).

Cmpd No.	Structure	h pIC ₅₀	m pIC ₅₀	r pIC ₅₀	hERG	P450	PFI	CLND µg/mL
(4.1)		5.2	6.4	4.5 ^a	6.3	2D6 = 8.2 2C9 = 5.0 3A4VG = 5.0 (n=1)	8.2	48.5
(4.174)		6.8	7.4	6.7	4.9	3A4VR = 6.1 3A4VG = 5.6	7.3	14
(4.180)		6.4	7.5	6.6	5.1	3A4VR = 5.2 3A4VG = 5.2	6.3	11
(4.198)		5.4	7.0	5.7	4.7	No 3A4 TDI	5.6	34
(4.196)		7.1	7.0	6.1	4.6	3A4VR = 5.5 3A4VG = 5.5	6.3	18
(4.226)		6.3	6.6	6.2	5.0	3A4VR = 5.4 3A4VG = 5.1	6.0	76
(4.222)		6.2	6.8	6.4	4.4	3A4VR = 5.4 3A4VG = 4.7	5.1	≥ 139
(4.230)		6.2	6.7	6.8	< 4.3	3A4VR = 5.3 3A4VG = 4.5	4.6	≥ 133

Table 35 – Potency, hERG, P450 and solubility data for representatives from the isoindoline and pyrazolopyrimidine sub-series of the imidazopyrazine template. ^aThis compound has only been reported active in 4 out of 24 screening occasions – all inactive screening occasions were pIC₅₀ < 4.52.

4.11 Summary of Efforts to Identify CD38 Binders without Crystallography of Compounds from the series

The aim of this programme of work was to optimise imidazopyridine (**4.5**) or pyrazolopyrimidinone (**4.10**) into CD38 inhibitors with human $pIC_{50} > 6.0$, with good efficiencies (LE and $LLE_{(AT)} > 0.3$) (**Figure 152**). The optimised molecules were required to have minimal P450 or hERG liabilities, as these off-target activities had been found for the carboxamide series (**4.1**) being investigated (**Figure 153**). Throughout this programme of work, no X-ray crystal structures could be obtained for either of the series, apart from the initial imidazopyridine hit (**4.5**). The X-ray crystal structure of imidazopyridine (**4.5**) was used to help guide docking and design of analogues throughout.

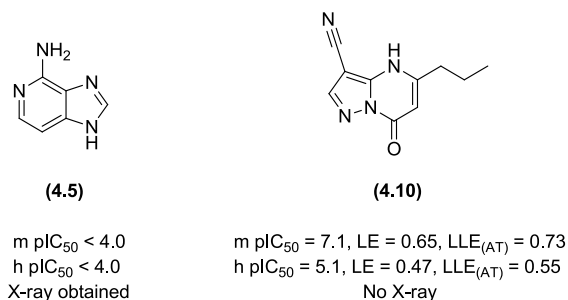


Figure 152 - Fragment hit 1*H*-imidazo[4,5-*c*]pyridin-4-amine (**4.5**) and HTS derived pyrazolopyrimidinone (**4.10**).

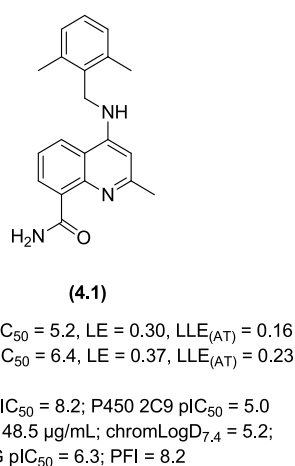


Figure 153 – Example molecule from the quinoline carboxamide (**4.1**) series under investigation for inhibition of CD38.

Chapter 4 describes how pyrazolopyrimidinone (**4.10**) was optimised, looking systematically at each growth point. Two docking poses for (**4.10**) were feasible and it was concluded that docking pose 1 represented the most likely binding orientation for the pyrazolopyrimidinone analogues. Despite building confidence in docking pose 1, the pyrazolopyrimidinone analogues reached a human potency plateau of $pIC_{50} \sim 6.0$ that could not be overcome. Thiophene (**4.68**) and thiazole (**4.132b**) were the most encouraging analogues with good measures of PFI and solubility (**Table 36**).

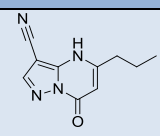
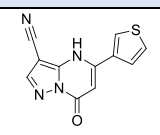
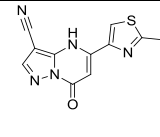
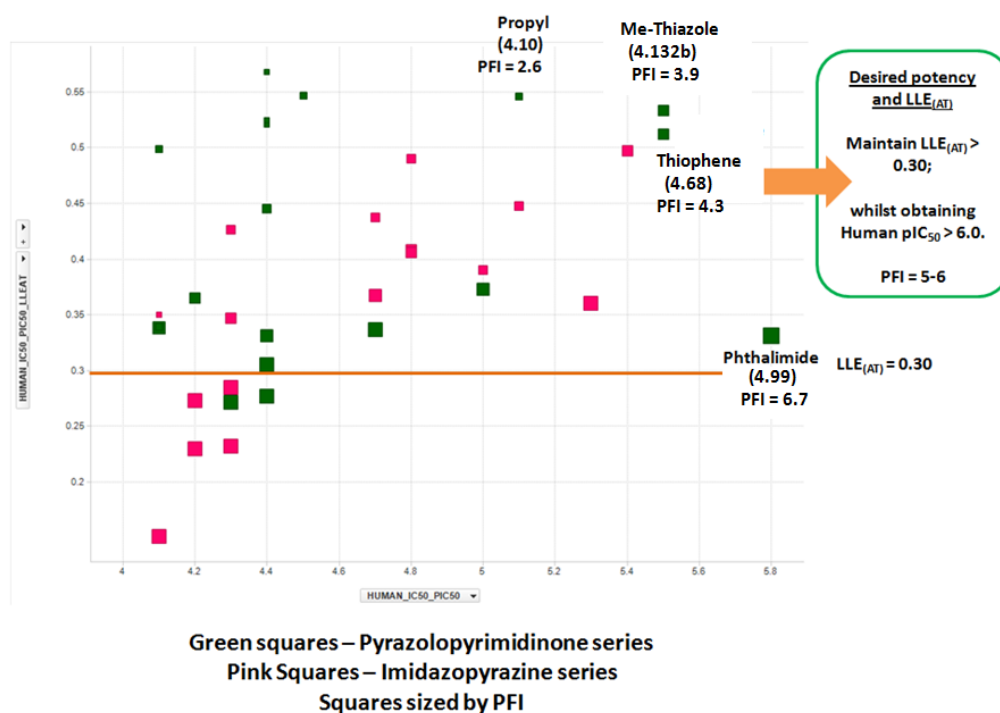
Cmpd No.	Structure	h pIC_{50}	h LE	h $LLE_{(AT)}$	m pIC_{50}	m LE	m $LLE_{(AT)}$	PFI	CLND $\mu\text{g/mL}$
(4.10)		< 4.0			5.1	0.47	0.55	2.6	≥ 101
(4.68)		5.5 (n=1)	0.44	0.52	6.9 (n=1)	0.56	0.63	4.3	≥ 134
(4.132b)		5.5	0.42	0.54	6.8	0.52	0.64	3.9	≥ 94

Table 36 – Most advanced analogues of pyrazolopyrimidinone (**4.10**).

Based on all the above, the strategic decision was made to focus on the imidazopyrazine series derived from SSS and synthesis around imidazopyridine (**4.5**). The imidazopyrazine series had demonstrated, through SSS, compounds with similar potency and efficiency to the most advanced pyrazolopyrimidinone analogues (**Graph 7**). It was reasoned that, with careful design, more potent and efficient imidazopyrazine analogues could be obtained.



Graph 7 – Graph showing pyrazolopyrimidinone and imidazopyrazine analogues; human pIC_{50} vs human $LLE_{(AT)}$.

Screening of a set of compounds obtained *via* SSS around crystallographic hit pyrazolopiperidine (**4.5**) resulted in identification of thiophene (**4.167**) with the alternative imidazopyrazine core (**Figure 154**). Due to X-ray crystal structures being unobtainable, thiophene (**4.167**) was optimised through rigorous docking and consideration of GRID maps to understand areas of the protein pocket to target for additional potency. Considering the pyridine analogue (**4.172**) conformation, the overall developing structure was constrained by ring formation to mimic the intramolecular hydrogen bond between the NH and the pyridyl nitrogen in (**4.172**), resulting in the more active isoindoline (**4.179**). Exploration of the isoindoline series demonstrated that excellent levels of potency could be obtained, but the compounds had very poor predicted physicochemical properties and measured solubility. The solubility profile was improved by changing the isoindoline moiety for the pyrazolopiperidine motif. Pyrazolopiperidine (**4.186**) gave a decrease in potency and LE compared to isoindoline (**4.179**), but a slight improvement in

LLE_(AT) was observed (**Table 37**). Design from docking resulted in amide-substituted pyrazolopiperidines, pyrazole (**4.230**) and cyclopentene (**4.222**) fulfilling the desired criteria for this research programme, apart from the modest P450 3A4 activity (**Table 38**). Confidence that the P450 3A4 activity was not a series specific issue came from the isoindoline *i*-Pr amide (**4.198**) which was inactive in the P450 3A4 assay (**Table 35**).

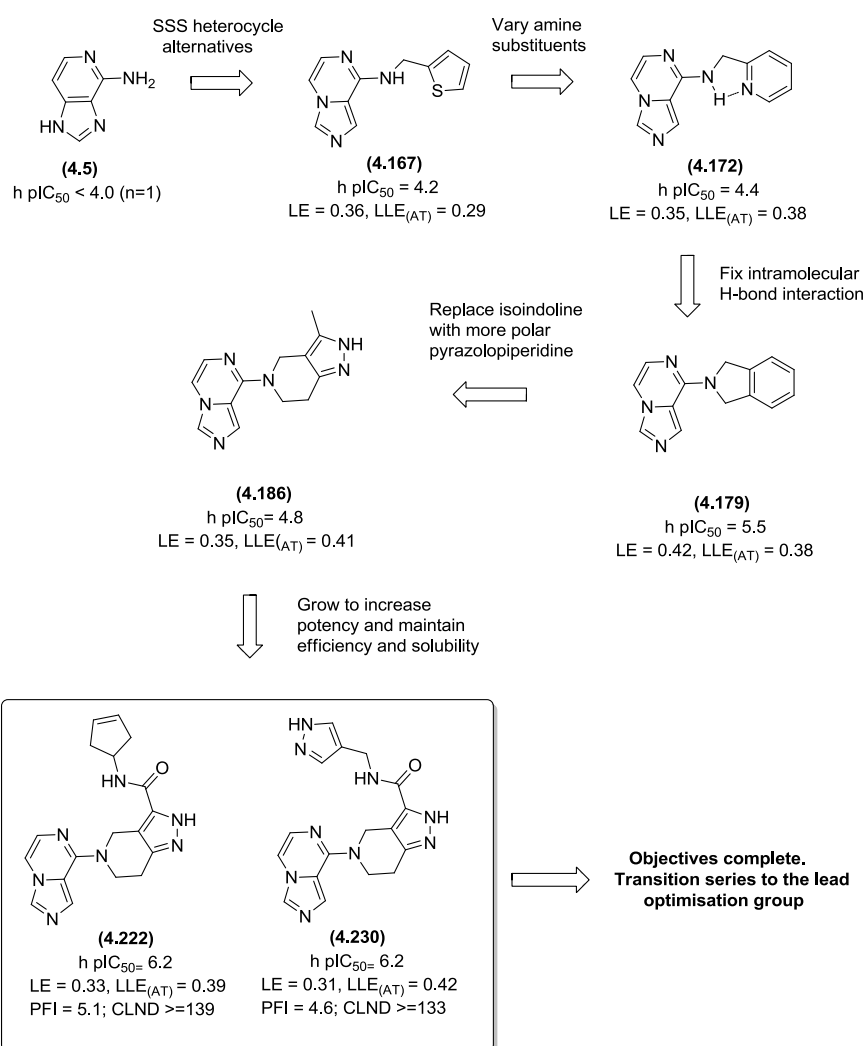
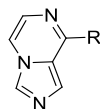


Figure 154 – Summary of the route of identification of imidazopyridine analogue (**4.222**), which meets the programme criteria, from the inactive imidazopyridine analogue (**4.5**).



Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)	PFI
(4.186)		4.8	0.35	0.41	6.7	0.48	0.55	4.0
(4.179)		5.5	0.42	0.38	6.7	0.51	0.47	6.0

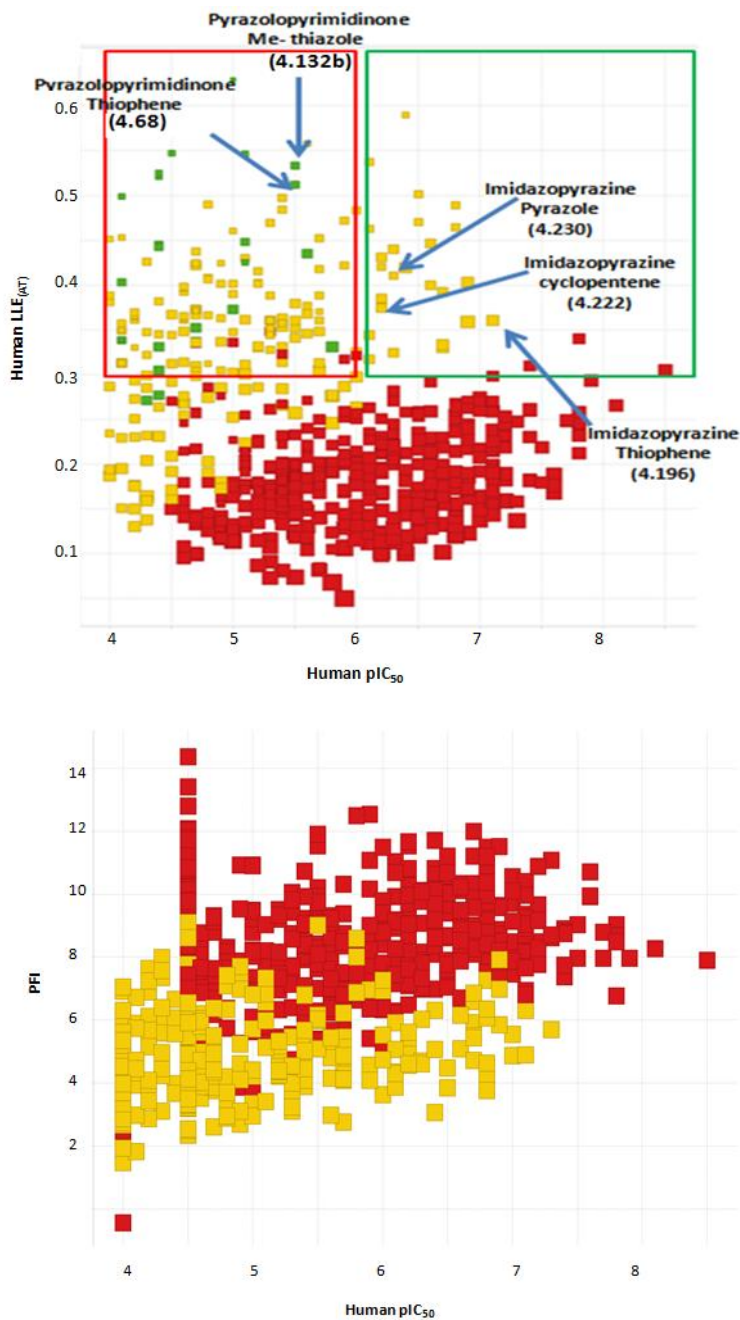
Table 37 – Comparison of potency, binding efficiency and PFI for pyrazolopiperidine analogues compared with the isoindoline analogues.

Cmpd No.	Structure	h pIC ₅₀	m pIC ₅₀	r pIC ₅₀	hERG	P450	PFI	CLND µg/mL
(4.1)		5.2	6.4	4.5 ^a	6.3	2D6 = 8.2 2C9 = 5.0 3A4VG = 5.0 (n=1)	8.2	48.5
(4.222)		6.2	6.8	6.4	4.4	3A4VR = 5.4 3A4VG = 4.7	5.1	≥ 139
(4.230)		6.2	6.7	6.8	< 4.3	3A4VR = 5.3 3A4VG = 4.5	4.6	≥ 133

Table 38 – Potency, hERG, P450 and solubility data for representatives from the isoindoline and pyrazolopyrimidine sub-series of the imidazopyrazine template. ^aThis compound has only been reported active in 4 out of 24 screening occasions – all inactive screening occasions were pIC₅₀ < 4.52.

CONFIDENTIAL – Property of GSK – Do Not Copy

At the time of making the decision to switch the synthetic focus to the imidazopyrazine from the pyrazolopyrimidinone template, all the analogues had $pIC_{50} \leq 5.8$ (in the red box or below the red box). Consideration turned to improving the CD38 potency of the imidazopyrazines whilst maintaining the LE and $LLE_{(AT)}$. The desired potency and efficiency space is represented by the green square. It can be seen that the imidazopyrazine series gave a significant number of analogues with the desired potency and $LLE_{(AT)}$ showing that the decision to change series was beneficial to the programme. It can also be seen from the graph that both the imidazopyrazine and the pyrazolopyrimidinone series had improved $LLE_{(AT)}$ compared to the quinoline carboxamide series. Comparison of the PFI of the quinoline carboxamide series (red) and imidazopyrazine series (yellow) (**Graph 8 – Bottom Graph**), clearly demonstrates that the imidazopyrazine series optimised using consideration of FBDD concepts, is in a preferred property space. This analysis validates the proposal that FBDD screening and follow-up strategy can be used to identify less lipophilic, more efficient starting points for drug discovery.



Graph 8 – Top Graph: Human pIC₅₀ vs human LLE_(AT) for three series: quinoline carboxamide (red); pyrazolopyrimidinones (green) and imidazopyrazine (yellow). The red box indicates the potency and efficiency window for the optimised pyrazolopyrimidinone series. The imidazopyrazine series was then optimised identifying analogues which fill the desired area (pIC₅₀ > 6.0 and LLE_(AT) > 0.3) green box. **Bottom Graph:** Human pIC₅₀ vs PFI for quinoline carboxamide (red) and imidazopyrazine (yellow).

CONFIDENTIAL – Property of GSK – Do Not Copy

As described in the introductory section, CD38 has been linked to many therapeutic applications. Although CD38 inhibitors have been reported in the literature, many of them have functionality that would cause concern regarding potential off-target effects and toxicity. The compounds developed during this research programme are much needed tool molecules, which can be used to help validate the proposed disease association *in vivo*. If inhibition of CD38 is confirmed to have a positive effect on one of the proposed diseases *in vivo*, these series both provide good leads to develop into medicines. Although significant work is still required to develop a drug candidate from either of these series, the improved LLE_(AT) and PFI for the FBDD derived imidazopyrazine series provides optimism that inhibitors in good physicochemical property space can be identified.

5 Conclusions

5.1 Successful FBDD Optimisation Strategies

In an attempt to find CD38 lead molecules by the most efficient process, insight from the HIV IN_{CCD}-LEDGF fragment effort was applied. Although the two series discussed for CD38 are not both derived directly from FBDD, the thought and ethics behind FBDD have been applied and will be discussed. Four key recommendations are made, and how they impacted the progression of the CD38 FBDD project will be discussed in-turn.

- 1) Target Tractability:** From the outset of the CD38 effort it was clear that the enzyme would be challenging to target with small molecule inhibitors due to: the large open binding site; the uncompetitive nature of binding with respect to the substrate; and the undesirable structural and physicochemical properties of reported inhibitors. For the HIV IN_{CCD}-LEDGF FBDD effort it was demonstrated that the FBDD approach could facilitate identification of ligands with improved binding efficiency to the desired target, in an improved property space. Due to these outcomes, and the desire to identify CD38 binders in improved property space compared to quinoline (**4.1**), an FBDD approach was initiated, despite concerns regarding the target tractability. Pleasingly, the CD38 FBDD approach identified the imidazopyrazine series (e.g. pyrazole (**4.230**)) with improved binding efficiency and PFI compared to quinoline (**4.1**) (**Figure 155**). This emphasises the value of the FBDD approach to help control physicochemical properties, as the applicability has now been demonstrated for two distinct therapeutic targets.

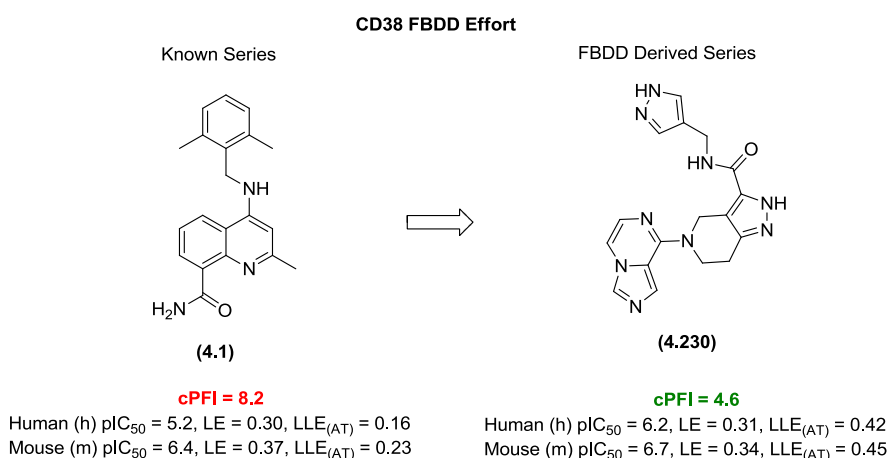
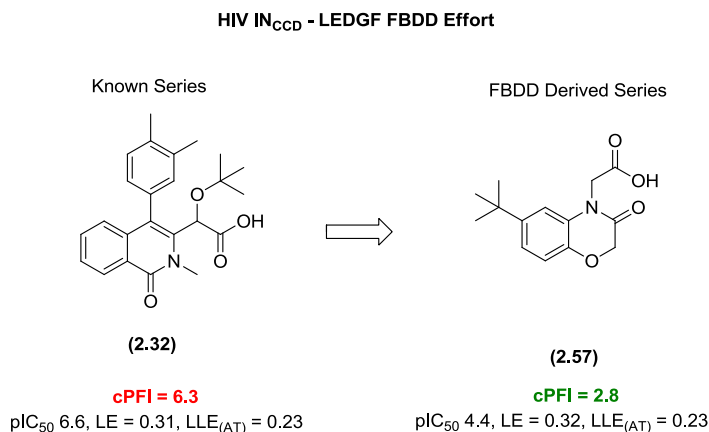


Figure 155 - Summary of templates identified as part of this research project for the HIV IN_{CCD}-LEDGF and CD38 FBDD efforts. In both examples, the FBDD -derived templates have lowered PFI.

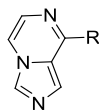
- 2) **Screening:** Both the literature and the HIV IN_{CCD}-LEDGF projects highlight the importance of running orthogonal biochemical and biophysical assays on the entire screening library to ensure identification of weak biochemically inactive, but efficient binders. The CD38 project helps validate this statement. The pyrazolopyrimidinone fragment (**4.10**) was identified from a HTS screen using a biochemical assay alone. Due to the number of compounds screened during HTS it is not feasible to run an orthogonal biophysical assay, but it is likely that

weakly efficient molecules were missed. In contrast, imidazopyridine hit fragment (**4.5**) was identified through SPR and was biochemically inactive. Without a robust biophysical assay, this fragment would not have been detectable, preventing the imidazopyrazine series from being identified on this screening occasion.

3) Fragment Optimisation:

For the pyrazolopyrimidinone series (**4.10**), the challenge from the outset was not having a liganded X-ray crystal structure to determine the binding orientation of the fragment. This project has shown that by docking combined with careful analogue design, confidence can be built in a specific binding mode allowing progression of a series without the requirement for X-ray crystallography. This work highlighted the importance of including negative controls, i.e. compounds predicted to be less active, to help validate the binding mode. Although growth of the pyrazolopyrimidinone series (**4.10**) was terminated in favour of pursuing the imidazopyrazine series, it is believed that knowledge around the increase in binding affinity by interacting with Trp176 from growth of the imidazopyrazine series, could be used to further develop the pyrazolopyrimidinone series.

During the retrospective-analysis of the HIV IN_{CCD}-LEDGF effort, it was stated that FBDD only has an advantage in the pursuit of ligand efficient molecules in good physicochemical property space, providing rigorous control is practiced throughout the optimisation process. This is exemplified by constraint of bromopyridyl (**4.173**) into bromoisindoline (**4.174**) resulting in an encouraging potency and efficiency increase against both CD38 constructs, with a detrimental impact to the PFI. Without consideration of the PFI, the isoindolines could have been developed into a potent lead series with a high chance of attrition moving towards or within the clinic due to developability issues, i.e. solubility, dose, off-target toxicity.



Cmpd No.	Structure	Human pIC ₅₀	Human LE	Human LLE _(AT)	Mouse pIC ₅₀	Mouse LE	Mouse LLE _(AT)	PFI
(4.173)		6.2	0.47	0.43	6.6	0.43	0.39	5.3
(4.174)		6.8	0.49	0.40	7.4	0.53	0.44	7.3

Table 39 – ~100-fold improvement in CD38 (mouse) potency was achieved by maximising the interactions of the small molecule with the CD38 protein by stabilising the required binding conformation. This had a detrimental effect on PFI.

Prior to growth of the hit fragment, the most efficient starting point must be identified. To complement the fragmentation of known binders, confirming the best identity (e.g. C, N, S, and O) of each atom in the new fragments should be pursued. For the CD38 FBDD effort, exploration of alternative 5,6 heterocyclic cores for imidazopyridine (**4.5**) lead to identification of pyrazolopyrimidine (**4.143**). Efforts to find alternatives to the pyrazolopyrimidine group in thiophene (**4.143**) led to the imidazopyrazine analogue (**4.167**) which displayed a significant improvement in mouse potency and efficiency (**Table 40**). Without this potency improvement, fragments such as thiophene (**4.143**) were on the limit of the biochemical assay (primary assay) which made it difficult to understand the influence of changes to the molecules. A significant amount of work was carried out exploring pyrazolopyrimidine analogues before the decision was made to switch to the imidazopyrazine core. If this decision to switch templates had been made sooner, suitable molecules for lead optimisation would have been developed in a more efficient manner reducing the cost of the programme and reducing the timelines for suitable molecules to be progressed to the clinic.


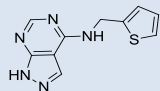
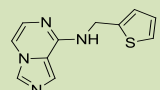
Cmpd No.	Structure	h pIC ₅₀	h LE	h LLE _(AT)	m pIC ₅₀	m LE	m LLE _(AT)
(4.5)		< 4.0 (n=1)			< 4.0 (n=1)		
(4.143)		< 4.0			4.3	0.37	0.31
(4.167)		4.2	0.36	0.29	5.8	0.50	0.42

Table 40 – Changing the arrangement of heteroatoms in the 5,6 heterocyclic core provided greater than 10-fold increase in enzyme activity (mouse).

- 4) **Computational Techniques:** As recommended in the retrospective analysis of the HIV IN_{CCD}-LEDGF FBDD effort, the computational technique GRID was used early in the CD38 FBDD effort to help characterise the enzyme pocket and understand regions of the protein pocket where profitable ligand-protein interactions may be obtained. Similarly to the HIV IN_{CCD}-LEDGF effort, GRID probes were observed in multiple remote regions of the pocket (**Figure 156**). Despite X-ray crystal structures not being obtained, docking experiments suggest that the potent imidazopyrazine analogues are making interactions in three of the regions highlighted by GRID: the selectivity region; the acid region and the Trp176 shelf. It has also been demonstrated that the linker (where no GRID occupies between the acid and selectivity regions) can be switched from the pyrazolopiperidine to the isoindoline without significant potency affects, validating this part of the molecule as a linker as the GRID maps depict. Changes to the core were critical in maintaining the physicochemical properties in a desirable range (**Figure 157**) leading to a successful outcome for the programme.

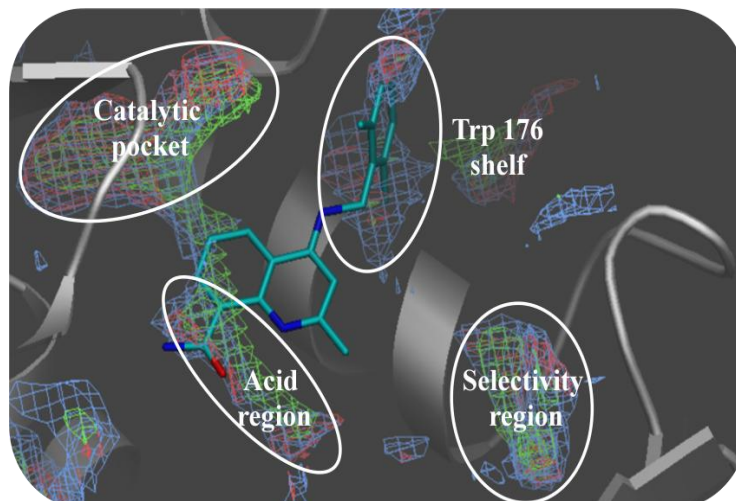


Figure 156 – GRID analysis for CD38. Key to GRID – Green = CH₄, Blue = NH₄⁺, Red = C=O.

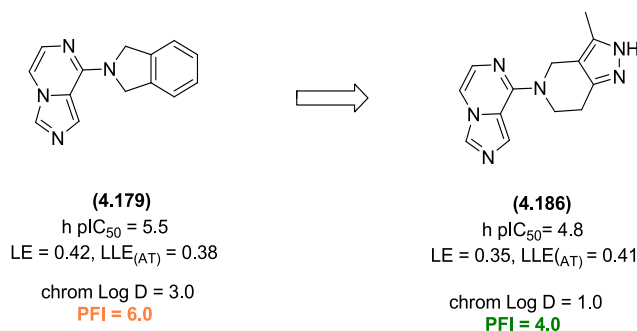


Figure 157 – Changing from the isoindoline core to the pyrazolopiperidine reduced the chrom LogD of the molecule and moved the PFI into the desired region (PFI < 6).

In summary, application of the recommendations from the retrospective analysis of the HIV IN_{CCD}-LEDGF has had great impact on the CD38 HTS/FBDD effort, resulting in identification of a valuable tool series which could be further optimised towards a drug candidate. It is predicted that careful application of these recommendations will maximise the success of future projects and minimise the timelines to identify lead molecules suitable for optimisation for progression into the clinic.

5.2 Benefits this Research Brings to the Wider Scientific Community

This research programme utilised FBDD techniques to successfully identify hits, and develop them into ligand efficient leads in good physicochemical property space for targets of high therapeutic interest. The targets pursued each represent a distinct biological challenge: a protein-protein interaction with large surface contacts to disrupt, and a hydrolase enzyme with a large open active site, combined with no structural information for the series of interest. Lead molecules with reduced structural liabilities, reducing the chance of off-target effects, have been identified for both programmes providing valuable molecules to aid in mode of action studies and drug development.

In addition to lead molecules being identified, improved understanding and advancement of FBDD concepts and optimisation tools have been discussed. Retrospective analysis of the FBDD approach pursued for each programme has highlighted key learnings which include:

- fragmentation of known binders to identify the most LE starting point and the interaction/interactions required for potency;
- the use of GRID probes to identify directions and areas to focus growth towards and determine ‘linker’ areas to modify physicochemical properties;
- continuous monitoring of physicochemical properties is critical; care must be taken not to be overly influenced by potency and efficiency alone;
- the importance of stabilising the required binding mode of the ligand (comparing with the lowest energy conformation from molecular dynamics calculations); and
- the pursuit of FBDD without X-ray crystallography (or alternative NMR techniques) is achievable through continuous rigorous docking and the use of GRID maps.

Incorporation of these learnings into future FBDD efforts will reduce the timelines associated with the ‘hit to lead’ process and, therefore, make FBDD optimisation a valuable addition to any ‘hit to lead’ effort. Reducing the timelines to identify a

CONFIDENTIAL – Property of GSK – Do Not Copy

marketable drug, and reducing attrition rates in the clinic by ensuring the molecules have good physicochemical properties, will aid in increasing the number of medicines marketed, impacting on the quality and life expectancy of millions of people worldwide.

6 Experimental Section

General procedures

All solvents and reagents, unless otherwise stated, were commercially available and were used without further purification.

Nuclear magnetic resonance (NMR)

NMR spectra were recorded using a Bruker DPX400, AV400 or AVIII600 (with cryoprobe) referenced to tetramethylsilane. Chemical shifts are reported in parts per million (ppm) and coupling constants (J) in Hz. The following abbreviations are used for multiplicities: s = singlet; br.s = broad singlet; d = doublet; t = triplet; m = multiplet; dd = doublet of doublets; ddd = double double doublet.

Liquid chromatography mass spectroscopy (LCMS)

Liquid Chromatography Mass Spectroscopy (LCMS) was conducted on either a Acquity UPLC BEH C18 column (50 mm x 2.1 mm i.d. 1.7 μ m packing diameter) at 40 °C eluting with either:

- Method A - 0.1% v/v solution of formic acid in water (solvent A) and 0.1% v/v solution of formic acid in acetonitrile (solvent B); or
- Method B - 10 mM ammonium bicarbonate in water adjusted to pH 10 with ammonia solution (solvent A) and acetonitrile (solvent B).

The UV detection is a summed signal from wavelength of 210 nm to 350 nm.

The mass spectra were recorded on a Waters ZQ spectrometer using electrospray positive and negative mode. The following elution gradients were used:

Formic acid modifier (Method A)

Time (min)	Flow rate (mL / min)	% A	% B
0	1	97	3
1.5	1	0	100
1.9	1	0	100
2.0	1	97	3

High pH modifier (Method B)

Time (min)	Flow rate (mL / min)	% A	% B
0	1	99	1
1.5	1	3	97
1.9	1	3	97
2.0	1	99	1

Mass directed auto-preparative HPLC (MDAP)

Preparative mass directed HPLC (MDAP) was conducted on a Waters *MassLynx* system comprising of a Waters 515 pump with extended pump heads, Waters 2767 autosampler, Waters 996 photodiode array detector, and Gilson 202 fraction collector on a Xbridge, Sunfire, XSelect CSH C18 column (30 mm x 150 mm i.d. 5 µm packing diameter) at ambient temperature. The mobile phase was:

- Method A - 0.1% formic acid in water (solvent A) and 0.1% formic in acetonitrile (solvent B); or
- Method B - 10 mM ammonium bicarbonate in water adjusted to pH 10 with ammonia solution (solvent A) acetonitrile (solvent B).

The UV detection is a summed signal from wavelength of 210 nm to 350 nm. Mass spectra were recorded on Waters ZQ mass spectrometer using alternate-scan positive and negative electrospray ionization. The software used was *MassLynx* 3.5 with *FractionLynx* optio or using equivalent alternative systems.

The elution gradients used were at a flow rate of 40 mL/min over 10 or 20 min:

Gradient A	5-30% B
Gradient B	15-55% B
Gradient C	30-85% B
Gradient D	50-99% B
Gradient E	80-99% B

High resolution mass spectroscopy (HRMS)

ESI (+) high resolution mass spectra (HRMS) were obtained on a Micromass Q-Tof 2 hybrid quadrupole time-of-flight mass spectrometer, equipped with a Z-spray interface, over a mass range of 100 – 1100 Da, with a scan time of 0.9 s and an interscan delay of 0.1 s. Reserpine was used as the external mass calibrant ($[M+H]^+ = 609.2812$ Da). The Q-Tof 2 mass spectrometer was operated in W reflectron mode to give a resolution (FWHM) of 16000-20000. Ionisation was achieved with a spray voltage of 3.2 kV, a cone voltage of 50 V, with cone and desolvation gas flows of 10-20 and 600 L/h, respectively. The source block and desolvation temperatures were maintained at 120 °C and 250 °C, respectively. The elemental composition was calculated using MassLynx v4.1 for the $[M+H]^+$.

An Agilent 1100 Liquid Chromatograph equipped with a model G1367A autosampler, a model G1312A binary pump and a HP1100 model G1315B diode array detector was used. The method used was generic for all experiments. All separations were achieved using a Phenomenex Luna C18 (2) reversed phase column (100 x 2.1 mm, 3 µm particle size). Gradient elution was carried out with the mobile phases as (A) water containing 0.1% (v/v) formic acid and (B) acetonitrile containing 0.1% (v/v) formic acid. The conditions for the gradient elution were initially 5% B, increasing linearly to 100% B over 6 minutes, remaining at 100% B for 2.5 min then decreasing linearly to 5% B over 1 min, followed by an equilibration period of 2.5 min prior to the next injection. The flow rate was 0.5 mL/min, temperature controlled at 35 °C with an injection volume of between 2 to 5 µL. All samples were diluted with DMSO (99.9%) prior to LCMS analysis.

Purification by column chromatography

Column chromatography was conducted on a Combiflash® Rf,²³³ automated flash chromatography system, from Teledyne Isco using disposable, normal or reverse phase, SPE Rediseq cartridges (4 g to 330 g). The CombiFlash® Rf uses RFID (Radio Frequency Identification) technology to automate setting the parameters for purification runs and fraction collection. The system is equipped with a UV variable dual-wavelength and a Foxy® fraction collector enabling automated peak cutting, collection, and tracking.

Phase separators

Isolute® phase separator cartridges²³³ are fitted with a hydrophobic Teflon frit. They were used to separate chlorinated solvent from aqueous phase under gravity.

Isolute® 103²³³

Isolute® 103 cartridges from Biotage are hydroxylated polystyrene-divinylbenzene co-polymer resin. They were used to extract organic compounds from aqueous solution using a catch and release protocol.

Microwave

Microwave chemistry was typically performed in Biotage sealed vessels, irradiating with a Biotage InitiatorTM Microwave Synthesiser.²³³

H-Cube®²³⁴

Small scale hydrogenation chemistry was performed using a ThalesNano H-cube® continuous flow hydrogenation reactor. The reaction takes place on disposable packed catalyst cartridges, CatCarts®. Hydrogen is generated *in situ* by electrolysis of water. The reaction mixture can be heated up to 100 °C and pressurised up to 100 bar.

Melting point

Melting points were measured on a Stuart automatic melting point apparatus, SMP40.²³⁵

CLND solubility

CLND solubility is a high throughput solubility determination from 10 mM DMSO stock solution of compounds, using Chemiluminescence Nitrogen Detector (CLND) end point. The solubility range covered by the method is from 1 µM to 500 µM. The signal from CLND is directly proportional to the number of moles of the molecule / number of nitrogen. This makes it possible to quantify the solubility of the compounds using a calibration based on one single compound (for example caffeine). GSK in-house kinetic solubility assay conditions: 5 mL of 10 mM DMSO stock solution diluted to 100 mL with pH 7.4 phosphate buffered saline, equilibrated for 1 h at RT, and filtered through Millipore Multiscreen HTS-PCF filter plates (MSSL BPC). The eluent is quantified by suitably calibrated flow injection CLND.

Chrom Log D

Chrom Log D has been introduced to replace the octanol / water Log D measurements and is applied together with solubility. Chrom Log D is derived from the gradient retention time of the compound in reversed phase HPLC using three different pHs for starting mobile phases with acetonitrile gradient. The compound's lipophilicity is measured in acidic, neutral, and basic environments to reveal the acid / base character of the compound. The Chrom Log D at three pHs is measured *via* fast gradient retention time using UPLC instruments. The retention time is converted to CHI (Chromatographic Hydrophobicity Index). The CHI scale is projected onto a Chrom Log D scale based on the linear regression of CHI values against cLog P values for several thousands of diverse project compounds. Both Chrom Log D and CHI Log D are derived from CHI and can be inter-converted using the following formulae:

$$\text{Chrom Log D} = 1.632 \text{ CHI Log D} + 0.396$$

$$\text{CHI Log D} = 0.613 \text{ Chrom Log D} - 0.24$$

Advantages of Chrom Log D over octanol / water Log D:

1. Compounds' impurity does not disturb the measurements;
2. Shows good correlation with calculated Log D / P;
3. Shows good correlation with CHI Log D; and
4. Shows good correlation with solubility.

GSK in-house hydrophobicity assay conditions: 10 mL of 10 mM DMSO stock solution diluted to 750 mL with octanol saturated pH_{7.4} phosphate buffer and 160 mL buffer saturated octanol in a 96 well deep well block. The block is sealed and inverted for 3 sets of 50 inversions, then centrifuged at 300 g for 20 min. Both phases are then quantified using generic gradient UV-HPLC.

Biochemical Assay Details for the pIC₅₀ Determination against the Human and Mouse CD38 Enzymes²⁰⁸

CD38 inhibitors were tested for their capacity to inhibit human CD38 enzyme activity in a colorimetric-based assay. The extracellular domain of human CD38 was expressed in *Pichia pastoris* and purified to homogeneity. The enzyme activity assay was performed in a low-volume 384-well plate in a total volume of 20 µL. A range of concentrations of test compound in 200 nL of DMSO was delivered into the assay plate wells. Columns 6 and 18 of the plate contained DMSO with no compound and served as the high signal and low signal controls (no CD38 added), respectively. All additions of assay reagents to the plate were done using a Multidrop Combi, and the plate was shaken 3–5 s after each addition. CD38 (0.8 nM) was incubated with test compound in 10 µL containing 100 mM HEPES, pH 7.4, 4 mM EDTA, and 1 mM CHAPS for 30 min prior to initiation of the reaction. The reaction was initiated by a 10 µL addition containing 5 mM sodium acetate, pH 4.5, 1 mM CHAPS, 200 µM NAD, and 500 µM GW323424X ((*E*)-2-(2-(pyridin-4-ylmethylene)hydrazinyl)pyridine). The solutions for each of the two additions were prepared fresh each day from concentrated stocks of the individual components. The final concentrations in the assay were 50 mM HEPES, 2 mM EDTA, 1 mM CHAPS, and 2.5 mM sodium

CONFIDENTIAL – Property of GSK – Do Not Copy

acetate, 100 μ M NAD, 250 μ M GW323434X, and 0.4 nM CD38. GW323434X is a 4-pyridynal compound that acts as a nucleophile that participates in the base exchange reaction with the nicotinamide on NAD to form a novel dinucleotide that absorbs at 405 nm. Catalytic formation of this novel chromophore was followed in an Envision microplate reader by reading absorbance at two time points, typically 30 min apart within the first 45 min of the reaction. These time points were established empirically to ensure the rates determined were in a linear range of product formation. Data analysis was performed in the following way using ActivityBase XE (Abase XE). The data from the 15 and 45 min reads was processed by performing a subtraction function of 45 min read value minus 15 min read value for each plate well. The resulting values for noncontrol wells were converted to % inhibition using the formula $100 \times ((U - C1)/(C2 - C1))$, where U is the value of the test well, C1 is the average of the values of the high signal (column 6) control wells, and C2 is the average of the values of the low signal (column 18) control wells. Percent inhibition (y) was plotted versus inhibitor concentration (x), and curve fitting was performed with the following four parameter equation: $y = A + ((B - A)/(1 + (10x/10C)^D))$, where A is the minimum response, B is the maximum response, C is the $\log_{10}IC_{50}$, and D is the Hill slope. The results for each compound were recorded as pIC_{50} values ($-C$ in the above equation). pIC_{50} determination against rat CD38 was performed in a similar fashion using a colorimetric based assay.

The recombinant extracellular domain of mouse CD38 was expressed in CHO CGE cells and purified to homogeneity. The pIC_{50} values for the inhibitors against mouse CD38 were generated using the enzyme in a fluorescence-based assay in which the enzyme reaction occurred in a 10 μ L volume in a low-volume 384-well assay plate. The assay quantitated CD38 catalysed NAD hydrolysis over 45 min of reaction time in which the rate was linear. A range of concentrations of test compound in 100 nL of DMSO was delivered into the assay plate wells. Columns 6 and 18 of the plate contained DMSO and served as the low signal and high signal controls, respectively. Column 18 contained a potent mouse CD38 inhibitor to define the high signal (no enzyme activity) control. Additions to the plate other than compound were done

CONFIDENTIAL – Property of GSK – Do Not Copy

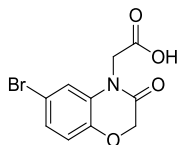
using a Multidrop Combi, and the plate was shaken 3–5 s after each addition. CD38 (0.45 nM) was incubated with test compound in 5 μ L containing 20 mM HEPES, pH 7.2, 1 mM EDTA, and 1 mM CHAPS for 30 min prior to initiation of the reaction. The reaction was initiated by a 5 μ L addition containing 20 mM HEPES, pH 7.2, 1 mM EDTA, 1 mM CHAPS, and 60 μ M NAD. The final concentrations in the assay were 20 mM HEPES, pH 7.2, 1 mM EDTA, 1 mM CHAPS, 30 μ M NAD, and 0.225 nM mouse CD38. After the reaction time, the amount of NAD remaining was quantitated by converting it to NADH using alcohol dehydrogenase (ADH). The ADH was added in 5 μ L containing 9U/mL ADH, 90 mM sodium pyrophosphate, pH 8.8, 90 mM ethanol, 1 mM EDTA, and 1 mM CHAPS. The alcohol dehydrogenase reaction was stopped by the addition of 5 μ L of 1 M HEPES, pH 7.0, 1.0 mM EDTA, and 1 mM CHAPS containing 0.8 M dithiothreitol (DTT), and the NADH fluorescence was measured in an Envision plate reader (340 nm excitation, 460 nm emission). The solutions for each of the four additions were prepared fresh each day from concentrated stocks of the individual components, except the DTT which was prepared fresh daily from solid. In this assay, an increase in enzyme activity results in a decreased measured fluorescent signal. Each compound plate was run in duplicate with (plate A) and without (plate B) ADH. Data were acquired by reading plates in pairs and subtracting the values for plate B from plate A to obtain “corrected” data (accounts for intrinsic fluorescence from test compound). Using Abase XE, “corrected” fluorescence signals for noncontrol wells were converted to percent inhibition values using the formula $100 - 100 \times ((U - C2)/(C1 - C2))$, where U is the “corrected” fluorescence signal value of the test well, C1 is the average of the “corrected” fluorescence values of the low signal (column 6; full CD38 enzyme activity) control wells, and C2 is the average of the “corrected” fluorescence values of the high signal (column 18; 100% inhibited CD38 enzyme activity) control wells. Percent inhibition data were fit using the four-parameter curve fit equation described above.

Biochemical Assay Details for the pIC₅₀ Determinations of Inhibition of the HIV IN – LEDGF Protein-Protein Interaction

His-HIV Integrase (50-212, F185K) and GST LEDGF (347-429) were expressed in E.Coli. Inhibitors were dissolved in DMSO to 10 mM and serially diluted in DMSO for assays. Reactions were performed at a 10 µl final volume in a 384-well plate format (Greiner Bio-One). The reaction buffer contained 50 mM HEPES pH 7.5, 150 mM NaCl, 20 mM MgCl₂, 0.1 mg/ml bovine serum albumin (BSA), 50 µM CHAPS, and fresh 2 mM dithiothreitol (DTT). After addition of drug in a 0.1 µl volume, hexahistidine-tagged INF185K (CCD) was added to a final concentration of 5 nM using a Multidrop Combi Reagent Dispenser (ThermoFisher Scientific). After 30 minutes of incubation at room temperature another addition was made with the dispenser to deliver GST-tagged LEDGF integrase binding domain (IBD, LEDGF 347-429) to a final concentration of 5 nM alongside the time-resolved fluorescence resonance energy transfer (TR-FRET) reagents APC-conjugated α-hexahistidine monoclonal antibody and α-GST Eulabelled monoclonal antibody (Perkin Elmer) at 5 nM final concentrations. After an additional hour of incubation at room temperature the TR-FRET signal the fluorescence at 665 nm was recorded with a ViewLux Microplate Imager (Perkin Elmer). For dose response curves, the ratio of 665 nM counts to 620 nM counts is normalised and expressed as %I.

6.1 Experimental Details – HIV Integrase-LEDGF

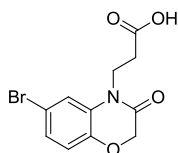
2-(6-Bromo-3-oxo-2H-benzo[b][1,4]oxazin-4(3H)-yl)acetic acid (2.52)



Methyl chloroacetate (39 μ L, 0.44 mmol) and K_2CO_3 (182 mg, 1.32 mmol) were added to a solution of 6-bromo-2H-benzo[b][1,4]oxazin-3(4H)-one (100 mg, 0.44 mmol) in DMF (5 mL). The reaction mixture was stirred at 20 $^{\circ}C$ for 23 h. NaOH (2.0 N, 1 mL) was added and the reaction mixture was stirred at 20 $^{\circ}C$ for 1 h. The reaction mixture was acidified to pH 1 with 2 N HCl and was extracted with EtOAc (x 3). The combined organic was dried by passing through a hydrophobic frit and was concentrated *in vacuo*. The residue was purified by reverse phase chromatography (MDAP method A, gradient B) to afford 2-(6-bromo-3-oxo-2H-benzo[b][1,4]oxazin-4(3H)-yl)acetic acid, as a white solid, 67 mg, 53%.

1H NMR (400 MHz, DMSO- d_6) δ 13.8 – 12.2 (br. s, 1H), 7.29 (d, J = 2.2 Hz, 1H), 7.18 (dd, J = 8.5, 2.2 Hz, 1H), 6.99 (d, J = 8.5 Hz, 1H), 4.71 (s, 2H), 4.65 (s, 2H); ^{13}C NMR (101 MHz, MeOD) δ ppm 171.1, 166.6, 145.8, 131.6, 127.9, 119.5, 119.0, 115.8, 68.4, 43.7; LCMS method A: $C_{10}H_7^{79/81}BrNO_4$ $[M-H]^-$ = 284/286 found 284/286, R_T = 0.75 min; HRMS: m/z calcd for $C_{10}H_9^{79}BrNO_4$ $[M+H]^+$ = 285.9710, found 285.9716; IR cm^{-1} 3052 (broad and complex), 2287, 1732, 1649; mp 205-208 $^{\circ}C$.

3-(6-Bromo-3-oxo-2H-benzo[b][1,4]oxazin-4(3H)-yl)propanoic acid (2.53)



Methyl 3-bromopropanoate (73 mg, 0.44 mmol) was added to a solution of 6-bromo-2H-benzo[b][1,4]oxazin-3(4H)-one (100 mg, 0.44 mmol) and K_2CO_3 (121 mg, 0.88 mmol) in DMF (2 mL). The reaction mixture was stirred at 20 $^{\circ}C$ for 21.5 h. The

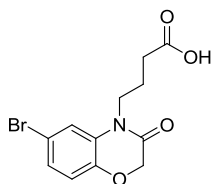
CONFIDENTIAL – Property of GSK – Do Not Copy

reaction was partitioned between EtOAc and water. The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*. The residue was purified by MDAP (method A, gradient C) to afford impure methyl 3-(6-bromo-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)propanoate (99 mg – 45 mg was taken into the ester hydrolysis step).

2.0 N NaOH (0.5 mL) was added to a solution of 2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)propanoate (45 mg, 0.14 mmol) in methanol (1 mL) and THF (1 mL) and was stirred at 20 °C for 1 h. The reaction mixture was blown down under nitrogen at 45 °C before partitioning between 2.0 N HCl (5 mL) and EtOAc (10 mL x 2). The organic was combined and dried by passing through a hydrophobic frit and concentrated *in vacuo*. The residue was purified by MDAP (method A, gradient B) to afford 3-(6-bromo-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)propanoic acid, as a white solid, 20 mg, 45% for ester hydrolysis step. Overall yield of 15% for alkylation and hydrolysis step.

¹H NMR (400 MHz, DMSO-*d*₆) δ 7.46 (d, *J* = 2.2 Hz, 1H), 7.17 (dd, *J* = 8.5, 2.2 Hz, 1H), 6.96 (d, *J* = 8.5 Hz, 1H), 4.64 (s, 2H), 4.10 (t, *J* = 7.8 Hz, 2H), 2.52 (t, *J* = 7.8 Hz, assumed 2H as signal overlaps with DMSO), carboxylic acid proton not seen as water present; ¹³C NMR (101 MHz, MeOD) δ ppm 174.6, 166.2, 146.2, 131.2, 127.9, 119.6, 119.3, 116.0, 68.5, 38.6, 32.7; LCMS method A: C₁₁H₁₁^{79/81}BrNO₄ [M+H]⁺ = 300/302, found 300/302, R_T = 0.80 min; HRMS: *m/z* calcd for C₁₁H₁₁⁷⁹BrNO₄ [M+H]⁺ = 299.9866, found 299.9871; IR cm⁻¹ 3280 - 2690 (broad and complex), 1731, 1641; mp 168-170 °C.

4-(6-Bromo-3-oxo-2H-benzo[b][1,4]oxazin-4(3H)-yl)butanoic acid (2.54)

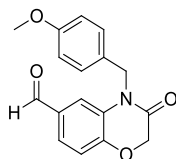


Methyl 4-bromobutanoate (79 mg, 0.44 mmol) was added to a solution of 6-bromo-2H-benzo[b][1,4]oxazin-3(4H)-one (100 mg, 0.44 mmol) and K₂CO₃ (121 mg, 0.88 mmol) in DMF (2 mL). The reaction mixture was stirred at 20 °C for 21.5 h. The reaction was partitioned between EtOAc and water. The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*. The residue was purified by reverse phase chromatography (method A, gradient C) to afford impure methyl 4-(6-bromo-3-oxo-2H-benzo[b][1,4]oxazin-4(3H)-yl)butanoate (83 mg - 50 mg was taken into the ester hydrolysis step).

2N NaOH (0.5 mL) was added to the crude (50 mg) in methanol (1 mL) and THF (1 mL) and was stirred at 20 °C for 1 h. The reaction mixture was blown down under nitrogen at 45 °C before partitioning between 2.0 N HCl (5 mL) and EtOAc (10 mL x 2). The organic was combined and dried by passing through a hydrophobic frit and concentrated *in vacuo*. The residue was purified by reverse phase chromatography (method A, gradient B) to afford 4-(6-bromo-3-oxo-2H-benzo[b][1,4]oxazin-4(3H)-yl)butanoic acid, as a white solid, 6.7 mg, 14% for ester hydrolysis step. Overall yield of 4.6% for alkylation and hydrolysis step.

¹H NMR (400 MHz, MeOD) δ 7.45 (d, *J* = 2.2 Hz, 1H), 7.14 (dd, *J* = 8.5, 2.0 Hz, 1H), 6.90 (d, *J* = 8.5 Hz, 1H), 4.59 (s, 2H), 3.95 - 4.00 (t, *J* = 7.5, 2H), 2.38 (t, *J* = 7.0 Hz, 2H), 1.95 - 1.87 (multiplet (apparent quartet with varying *J* values), 2H), carboxylic acid proton not seen; LMCS method A: C₁₂H₁₃^{79/81}BrNO₄ [M+H]⁺ = 314/316, found 314/316, R_T = 0.84 min.

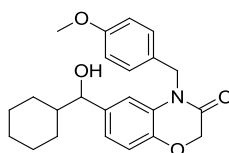
4-(4-Methoxybenzyl)-3-oxo-3,4-dihydro-2H-benzo[*b*][1,4]oxazine-6-carbaldehyde (2.73)



K_2CO_3 (468 mg, 3.39 mmol) and *p*-methoxybenzyl bromide (0.25 mL, 1.69 mmol) were added to a solution of 3-oxo-3,4-dihydro-2H-benzo[*b*][1,4]oxazine-6-carbaldehyde (200 mg, 1.13 mmol) in DMF (3 mL) and the reaction was stirred at 20 °C for 4 h. The reaction mixture was left sitting at 20 °C overnight before being blown down at 45 °C under nitrogen. The residue was partitioned between water and ethyl acetate. The organic was separated and dried by passing through a hydrophobic frit and was concentrated by blow down at 45 °C under nitrogen. The residue was purified by normal phase chromatography (cyclohexane / *t*-butyl methyl ether) to afford 4-(4-methoxybenzyl)-3-oxo-3,4-dihydro-2H-benzo[*b*][1,4]oxazine-6-carbaldehyde as a clear colourless gum, 272 mg, 81%.

^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 9.81 (s, 1H), 7.58 (dd, $J = 8.3, 1.8$ Hz, 1H), 7.54 (d, $J = 1.8$ Hz, 1H), 7.25 (d, $J = 8.6$ Hz, 2H), 7.20 (d, $J = 8.3$ Hz, 1H), 6.89 (d, $J = 8.6$ Hz, 2H), 5.15 (s, 2H), 4.93 (s, 2H), 3.71 (s, 3H); LCMS method A: $\text{C}_{17}\text{H}_{16}\text{NO}_4$ $[\text{M}+\text{H}]^+ = 298$, molecule does not ionise (material used without further characterisation), $R_T = 1.05$ min.

6-(Cyclohexyl(hydroxy)methyl)-4-(4-methoxybenzyl)-2H-benzo[*b*][1,4]oxazin-3(4H)-one (2.74)

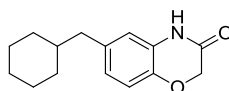


Cyclohexylmagnesium chloride (2.0 M in diethyl ether, 0.27 mL, 0.55 mmol) was added dropwise to a solution of 4-(4-methoxybenzyl)-3-oxo-3,4-dihydro-2H-benzo[*b*][1,4]oxazine-6-carbaldehyde (135 mg, 0.45 mmol) in THF (5 mL) at -78 °C under nitrogen. The reaction was slowly warmed to 20 °C over 5 h before stirring at

20 °C overnight. Further cyclohexylmagnesium chloride (2.0 M in diethyl ether, 0.27 mL, 0.55 mmol) was added under ice cooling, before stirring at 20 °C for a further 3 h. The reaction was quenched by dropwise addition of saturated aqueous ammonium chloride. The reaction mixture was extracted with EtOAc (x2). The combined organics were dried by passing through a hydrophobic frit and were blown down at 45 °C under nitrogen. The material was purified by MDAP (method A, gradient C). The appropriate fractions were combined and concentrated *in vacuo*. The residue was partitioned between saturated aqueous sodium bicarbonate and DCM. The organic was separated and dried by passing through a hydrophobic frit and was blown down at 45 °C under nitrogen to afford 6-(cyclohexyl(hydroxy)methyl)-4-(4-methoxybenzyl)-2H-benzo[*b*][1,4]oxazin-3(4H)-one, 65 mg, 38%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 7.19 (d, *J* = 8.5 Hz, 2H), 6.94 (d, *J* = 1.5 Hz, 1H), 6.90 (d, *J* = 8.1 Hz, 1H), 6.85 (d, *J* = 8.5 Hz, 2H), 6.82 (dd, *J* = 8.1, 1.5 Hz, 1H), 5.13 (d, *J* = 15.6 Hz, 1H), 5.03 (d, *J* = 15.6 Hz, 1H), 4.95 (d, *J* = 4.5, 1H), 4.73 (s, 2H), 4.09 (dd, *J* = 6.5, 4.6 Hz, 1H), 3.69 (s, 3H), 1.70 – 1.55 (m, 2H), 1.55 - 1.45 (m, *J* = 16.1 Hz, 2H), 1.30 - 1.20 (m, 1H), 1.09 (d, *J* = 13.1 Hz, 1H), 1.02 – 0.92 (m, 3H), 0.70 - 0.63 (m, 2H); LCMS method A: C₂₃H₂₈NO₄ [M+H]⁺ = 382, found 382, R_T = 1.14 min.

6-(Cyclohexylmethyl)-2H-benzo[*b*][1,4]oxazin-3(4H)-one (2.65)



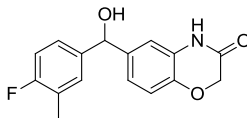
Triethylsilane (25 μL, 0.16 mmol) was added to a solution of 6-(cyclohexyl(hydroxy)methyl)-4-(4-methoxybenzyl)-2H-benzo[*b*][1,4]oxazin-3(4H)-one (20 mg, 52 μmol) in TFA (1 mL) and the reaction was heated at 50 °C for 3 h. The reaction mixture was heated at 60 °C for a further 3 h. The reaction mixture was basified by addition of saturated aqueous sodium bicarbonate solution. The aqueous was extracted with EtOAc (x 2) and DCM (x 2). The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*. The residue

CONFIDENTIAL – Property of GSK – Do Not Copy

was purified by normal phase chromatography (cyclohexane / *t*-butyl methyl ether) to afford 6-(cyclohexylmethyl)-2*H*-benzo[*b*][1,4]oxazin-3(4*H*)-one, as clear colourless crystals, 7.4 mg, 58%.

¹H NMR (400 MHz, MeOD) δ 6.82 (d, *J* = 8.3 Hz, 1H), 6.72 (dd, *J* = 8.3, 2.0 Hz, 1H), 6.67 (d, *J* = 2.0 Hz, 1H), 4.51 (s, 2H), 2.39 (d, *J* = 7.0 Hz, 2H), 1.72 - 1.61 (m, 5H), 1.52 - 1.42 (m, 1H), 1.24 - 1.16 (m, 3H), 0.99 - 0.87 (m, 2H), amide NH not observed; ¹³C NMR (101 MHz, MeOD) δ ppm 167.9, 143.3, 137.1, 127.7, 125.5, 117.6, 117.0, 68.2, 44.3, 41.2, 34.2, 27.7, 27.4; LCMS method A: C₁₅H₂₀NO₂ [M+H]⁺ = 246, found 246, R_T = 1.24 min; HRMS: *m/z* calcd for C₁₅H₂₀BrNO₂ [M+H]⁺ = 246.1489, found 246.1488; IR cm⁻¹ 2916, 2848, 2304, 2276, 2187, 1673; mp 150-151 °C.

6-((4-Fluoro-3-methylphenyl)(hydroxy)methyl)-2*H*-benzo[*b*][1,4]oxazin-3(4*H*)-one (2.76)



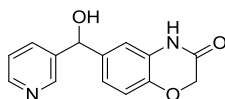
(4-Fluoro-3-methylphenyl)magnesium chloride in THF (1.13 mL, 1.13 mmol, 1.0 M) was added dropwise to an ice cooled solution of 3-oxo-3,4-dihydro-2*H*-benzo[*b*][1,4]oxazine-6-carbaldehyde (100 mg, 0.56 mmol) in THF (5 mL). The reaction mixture was stirred at 20 °C under nitrogen for 22 h. Further (4-fluoro-3-methylphenyl)magnesium chloride in THF (1.13 mL, 1.13 mmol, 1.0 M) was added dropwise under ice cooling. The reaction mixture was stirred at 20 °C under nitrogen for a further 3 h. The reaction mixture was quenched by addition of saturated aqueous ammonium chloride solution and was extracted with EtOAc (x 3). The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*. The residue was purified by normal phase chromatography (cyclohexane / *t*-butyl methyl ether) to afford 2 batches of impure 6-((4-fluoro-3-methylphenyl)(hydroxy)methyl)-2*H*-benzo[*b*][1,4]oxazin-3(4*H*)-one containing aliphatic impurities (Batch 1: clear gum, 20 mg; Batch 2: white solid, 89 mg, 55%).

CONFIDENTIAL – Property of GSK – Do Not Copy

Subsequent purification of Batch 1 using LCMS (method A, gradient B) afforded pure 6-((4-fluoro-3-methylphenyl)(hydroxy)methyl)-2*H*-benzo[*b*][1,4]oxazin-3(4*H*)-one, as a clear colourless gum, 11.7 mg, 7.2%.

¹H NMR (400 MHz, MeOD) δ 7.19 (dd, *J* = 8.0, 2.0 Hz, 1H), 7.17 - 7.12 (m, 1H), 6.97 - 6.91 (m, *J* = 8.0 Hz, 2H), 6.89 (d, *J* = 2.0 Hz, 1H), 6.87 (d, *J* = 8.0 Hz, 1H), 5.64 (s, 1H), 4.52 (s, 2H), 2.22 (d, *J* = 1.5 Hz, 3H), amide NH and alcohol OH not seen; LCMS method A: C₁₆H₁₅FNO₃ [M+H]⁺ = 288, found 288, R_T = 0.88 min.

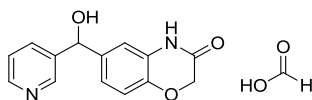
6-(Hydroxy(pyridin-3-yl)methyl)-2*H*-benzo[*b*][1,4]oxazin-3(4*H*)-one (2.77) - Scheme 6



3-Bromopyridine (0.09 mL, 0.94 mmol) was dissolved in THF (2 mL) under nitrogen. *i*PrMgCl.LiCl (1.3M in THF, 0.73 mL, 0.94 mmol) was added dropwise and the reaction mixture was stirred at RT for 1h. 3-Oxo-3,4-dihydro-2*H*-1,4-benzoxazine-6-carbaldehyde (84 mg, 0.47 mmol) in THF (2 mL) was added and the reaction mixture was stirred a RT under nitrogen for 4 h after which time the reaction conversion had halted. The reaction was quenched by addition of saturated aqueous ammonium chloride solution and was extracted with EtOAc (x 3). The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*. The residue was purified by reverse phase chromatography (method A, gradient A) to afford 6-(hydroxy(pyridin-3-yl)methyl)-2*H*-benzo[*b*][1,4]oxazin-3(4*H*)-one, 7.0 mg, 2.9%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 10.65 (s, 1H), 8.58 - 8.54 (m, 1H), 8.43 (dd, *J* = 1.4, 4.6 Hz, 1H), 7.71 - 7.65 (m, 1H), 7.34 (dd, *J* = 4.8, 7.8 Hz, 1H), 6.96 - 6.92 (m, 1H), 6.92 - 6.85 (m, 2H), (6.08 (br.s, 1H) 5.70 (s, 1H), 4.52 (s, 2H), contains ~ 10% impurity and residual formic acid; LCMS method A: C₁₄H₁₃N₂O₃ [M+H]⁺ = 257, found 257, R_T = 0.38 min, contains 11% impurity at R_T = 0.52 min.

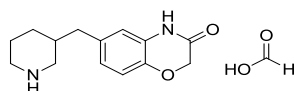
6-(Hydroxyl(pyridine-3-yl)methyl)-2H-benzo(b)(1,4)oxazin-3(4H)-one. formic acid salt (2.77b) – Scheme 7



*i*PrMgCl.LiCl (1.3 M in THF, 2.61 mL, 3.39 mmol) was added dropwise to a solution of 3-bromopyridine (0.33 mL, 3.39 mmol) in THF (3 mL) at 20 °C under nitrogen. After 1 h the reaction mixture was added drop-wise through a septum to a suspension of 3-oxo-3,4-dihydro-2H-1,4-benzoxazine-6-carbaldehyde (300 mg, 1.69 mmol) and cerium (III) chloride (835 mg, 3.39 mmol) in THF (12 mL) at 20 °C under nitrogen. After stirring for 2 h (63% conversion to product by LCMS) the reaction was quenched with saturated aqueous ammonium chloride solution before acidifying with 2.0 N HCl and extracting with EtOAc. The aqueous phase was extracted with DCM (x 3) and the combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*. The crude was purified by reverse phase chromatography (0.1% formic acid in water / MeOH), using a C18 Gold aqueous Redisep column (50 g). The appropriate fractions were combined and concentrated *in vacuo* to afford 6-(hydroxyl(pyridine-3-yl)methyl)-2H-benzo(b)(1,4)oxazin-3(4H)-one formic acid salt, as a clear colourless gum, 124 mg, 24%.

¹H NMR (400 MHz, MeOD) δ 8.55 (d, *J* = 2.0 Hz, 1H), 8.41 (dd, *J* = 5.0, 1.5 Hz, 1H), 8.09 (s, 1H), 7.82 (ddd, *J* = 7.9, 2.0, 1.5 Hz, 1H), 7.40 (dd, *J* = 7.9, 5.0 Hz, 1H), 6.97 (dd, *J* = 8.3, 2.0 Hz, 1H), 6.93 (d, *J* = 2.0 Hz, 1H), 6.91 (d, *J* = 8.3 Hz, 1H), 5.78 (s, 1H), 4.53 (s, 2H), alcohol and amine protons not observed; LCMS method A: C₁₄H₁₃N₂O₃.CH₂O₂ [M+H]⁺ for desalted fragment = 257, found 257, R_T = 0.37 min.

6-(3-Piperidinylmethyl)-2H-1,4-benzoxazin-3(4H)-one, formic acid salt (2.63b)



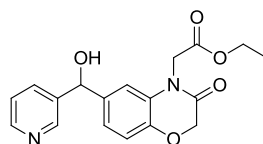
6-(Hydroxyl(pyridine-3-yl)methyl)-2H-benzo(b)(1,4)oxazin-3(4H)-one (60 mg, 0.23 mmol) in acetic acid (5 mL) was cycled (4 times) through the H-cube loaded with

CONFIDENTIAL – Property of GSK – Do Not Copy

10% Pd on carbon at 80 °C at 30 bar. The solvent was removed by blow down at 45 °C under nitrogen. The crude product was purified by reverse phase chromatography (0.1% formic acid in water / MeOH), using a C18 Gold aqueous Redisepp column (50 g). The appropriate fractions were combined and blown down at 45 °C under nitrogen to afford 6-(3-piperidinylmethyl)-2*H*-1,4-benzoxazin-3(4*H*)-one formic acid salt, as a clear colourless gum, 31 mg, 45%.

¹H NMR (400 MHz, CDCl₃) δ 10.02 (br.s, 1H), 8.52 (br.s, 2H), 6.83 (d, *J* = 8.3 Hz, 1H), 6.77 (s, 1H), 6.68 (d, *J* = 8.3 Hz, 1H), 4.55 (s, 2H), 3.31 (d, *J* = 11.5 Hz, 1H), 3.23 (d, *J* = 11.5 Hz, 1H), 2.75 (apparent t, *J* = 10.5 Hz, 1H), 2.59 - 2.39 (m, 3H), 2.13 (s, 1H), 1.91 - 1.73 (m, 3H), 1.24 - 1.13 (m, 1H); LCMS method A: C₁₄H₁₉N₂O₂.CH₂O₂ [M+H]⁺ = 247, found 247, R_T = 0.49 min.

Ethyl {6-[hydroxy(3-pyridinyl)methyl]-3-oxo-2,3-dihydro-4*H*-1,4-benzoxazin-4-yl}acetate (2.78)



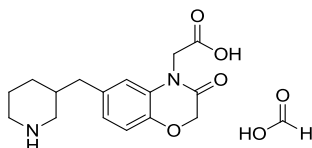
6-(Hydroxyl(pyridine-3-yl)methyl)-2*H*-benzo(*b*)(1,4)oxazin-3(4*H*)-one (57 mg, 0.22 mmol), K₂CO₃ (92 mg, 0.67 mmol) and ethyl chloroacetate (71 μL, 0.67 mmol) were stirred in DMF (4 mL) at 20 °C for 90 h. The reaction mixture was partitioned between DCM (15 mL) and water. The organic was washed with water (10 mL x 2) and was dried by passing through a hydrophobic frit and was concentrated *in vacuo* to afford ethyl {6-[hydroxy(3-pyridinyl)methyl]-3-oxo-2,3-dihydro-4*H*-1,4-benzoxazin-4-yl}acetate, as a cream solid, 70 mg, 92%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.57 (d, *J* = 2.0 Hz, 1H), 8.41 (dd, *J* = 4.8, 1.5 Hz, 1H), 7.71 (ddd, *J* = 7.7, 2.0, 1.5 Hz, 1H), 7.31 (dd, *J* = 7.8, 4.8 Hz, 1H), 7.08 (d, *J* = 1.5 Hz, 1H), 7.03 (dd, *J* = 8.1, 1.5 Hz, 1H), 6.98 (d, *J* = 8.1 Hz, 1H), 6.07 (d, *J* = 4.3 Hz, 1H), 5.73 (d, *J* = 4.3 Hz, 1H), 4.71 (apparent d, *J* = 1.8 Hz, 2H), 4.67 (s, 2H),

CONFIDENTIAL – Property of GSK – Do Not Copy

4.12 (q, $J = 7.2$ Hz, 2H), 1.17 (t, $J = 7.2$ Hz, 3H), contains residual DMF; LCMS method A: $C_{18}H_{19}N_2O_5$ $[M+H]^+ = 343$, found 343, $R_T = 0.52$ min.

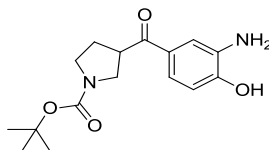
[3-Oxo-6-(3-piperidinylmethyl)-2,3-dihydro-4H-1,4-benzoxazin-4-yl]acetic acid, formic acid salt (2.80)



Ethyl {6-[hydroxy(3-pyridinyl)methyl]-3-oxo-2,3-dihydro-4H-1,4-benzoxazin-4-yl}acetate (70 mg, 0.20 mmol) in acetic acid (5 mL) was cycled (4 times) through the H-cube loaded with 10% Pd on carbon at 80 °C, 30 bar. The solution was blown down at 45 °C under nitrogen. THF (1 mL), EtOH (1 mL) and 2.0 N NaOH (0.5 mL) were added and the reaction was stirred at 20 °C for 6 days. The reaction mixture was neutralised by addition of 2.0 N HCl and was blown down at 45 °C under nitrogen. The crude was purified by MDAP (method A, gradient A) The appropriate fractions were combined and blown down at 45 °C under nitrogen to afford [3-oxo-6-(3-piperidinylmethyl)-2,3-dihydro-4H-1,4-benzoxazin-4-yl]acetic acid as the di-formic acid salt, 5.5 mg, 7.7%.

1H NMR (600 MHz, MeOD) δ 8.23 (s, 1H (overintegrates)), 6.88 (d, $J = 8.1$ Hz, 1H), 6.79 (d, $J = 8.1$ Hz, 1H), 6.73 (s, 1H), 4.63 (s, 2H), 4.51 (d, $J = 16.9$ Hz, 1H), 4.43 (d, $J = 16.9$ Hz, 1H), 3.11 (dd, $J = 13.0, 1.7$ Hz, 1H), 2.84 (ddd, $J = 13.9, 12.8, 2.8$ Hz, 1H), 2.68 (dd, $J = 13.9, 5.8$ Hz, 1H), 2.57 (t, $J = 12.1$ Hz, 1H), 2.46 (dd, $J = 13.9, 8.4$ Hz, 1H), 1.99 – 1.83 (m, 2H), 1.70 - 1.65 (m, 1H), 1.32-1.23 (m, 1H), carboxylic acid OH not observed, 2 protons assumed under MeOH peak; LCMS method A: $C_{16}H_{21}N_2O_4 \cdot CH_2O$ $[M+H]^+ = 305$, found 305, $R_T = 0.48$ min.

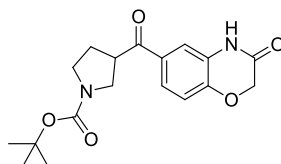
1,1-Dimethylethyl 3-[(3-amino-4-hydroxyphenyl)carbonyl]-1-pyrrolidinecarboxylate (2.86)



A solution of 1,1-dimethylethyl 3-[(2-methyl-1,3-benzoxazol-5-yl)carbonyl]-1-pyrrolidinecarboxylate (200 mg, 0.61 mmol) and 5.0 N HCl (1.1 mL, 36.3 mmol) in EtOH (2 mL) was heated at 100 °C for 1 h under microwave irradiation before concentrating *in vacuo*. The material was dissolved in DCM (4 mL). Et₃N (0.25 mL, 1.82 mmol) and Boc₂O (132 mg, 0.61 mmol) were added and the reaction was stirred at 20 °C under nitrogen for 1 h before concentrating by blow down under nitrogen at 45 °C. The crude was purified by normal phase chromatography (cyclohexane / *t*-butyl methyl ether) to afford 1,1-dimethylethyl 3-[(3-amino-4-hydroxyphenyl) carbonyl]-1-pyrrolidinecarboxylate, as a clear straw coloured gum, 130 mg, 70% (contains minor impurities).

¹H NMR (400 MHz, DMSO-*d*₆) δ 10.82 - 9.71 (br.s, 1H), 7.24 (d, *J* = 2.0 Hz, 1H), 7.19 (dd, *J* = 8.3, 2.0 Hz, 1H), 6.73 (d, *J* = 8.3 Hz, 1H), 5.40 – 4.32 (br.s, 2H), 4.05 – 3.94 (apparent septet, 1H), 3.49 (t, *J* = 9.2 Hz, 1H), 3.38 - 3.27 (m, assumed 3H (masked by water peak)), 2.14 - 2.05 (apparent sextet, 1H), 1.94 (apparent octet, 1H), 1.39 (s, 9 H), contains minor impurities; LCMS method A: C₁₆H₂₁N₂O₄ [M-H]⁻ = 305, found 305, R_T = 0.79 min.

1,1-Dimethylethyl 3-[(3-oxo-3,4-dihydro-2H-1,4-benzoxazin-6-yl)carbonyl]-1-pyrrolidinecarboxylate (2.88)



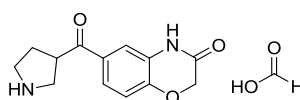
To a solution of 1,1-dimethylethyl 3-[(3-amino-4-hydroxyphenyl)carbonyl]-1-pyrrolidinecarboxylate (130 mg, 0.42 mmol) and K₃PO₄ (90 mg, 0.42 mmol) in MeCN (5 mL) under ice cooling under nitrogen, was added chloroacetyl chloride (34

CONFIDENTIAL – Property of GSK – Do Not Copy

μL , 0.42 mmol). The reaction mixture was stirred at 20 °C under nitrogen for 21 h before drop-wise addition of 5.0 M NaOH(aq) (85 μL , 0.42 mmol) followed by stirring for 1 h. The reaction was quenched with saturated aqueous ammonium chloride solution and was concentrated by blow down under nitrogen at 45 °C. The residue was partitioned between DCM and water and the organic was dried by passing through a hydrophobic frit and was concentrated *in vacuo* to afford 1,1-dimethylethyl 3-[(3-oxo-3,4-dihydro-2*H*-1,4-benzoxazin-6-yl)carbonyl]-1-pyrrolidinecarboxylate, 122 mg, 83%. The compound was used crude in the next step. NMR contains impurities – used crude.

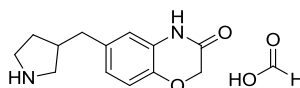
^1H NMR (400 MHz, DMSO- d_6) δ 10.98 (br.s, 1H), 7.67 (dd, $J = 8.3, 2.0$ Hz, 1H), 7.50 (d, $J = 2.0$ Hz, 1H), 7.06 (d, $J = 8.3$ Hz, 1H), 4.69 (s, 2H), 4.34 (br.s, 0.5H, rotamer), 4.12 - 4.01 (br. m, 1.5H rotamer), 3.52 (dd, apparent t, $J = 9.5, 9.5$ Hz, 1H), 3.46 – 3.36 (br. m, 1H, overintegrates), 3.30 (dd, apparent t, $J = 7.0, 7.0$ Hz, assume 1H, under water), 2.18 - 2.07 (m, 1H), 2.00 – 1.90 (m, 1H), 1.40 (s, 9H), contains impurities; LCMS method A: $\text{C}_{18}\text{H}_{21}\text{N}_2\text{O}_5$ $[\text{M}+\text{H}]^+ = 345$, found 345, $R_T = 0.96$ min.

6-(3-Pyrrolidinylcarbonyl)-2*H*-1,4-benzoxazin-3(4*H*)-one, formic acid salt (2.89)



and

6-(3-Pyrrolidinylmethyl)-2*H*-1,4-benzoxazin-3(4*H*)-one, formic acid salt (2.64b)



Et_3SiH (0.25 mL, 1.59 mmol) was added to a solution of 1,1-dimethylethyl 3-[(3-oxo-3,4-dihydro-2*H*-1,4-benzoxazin-6-yl)carbonyl]-1-pyrrolidinecarboxylate (55 mg, 0.16 mmol) in TFA (2 mL). The reaction mixture was stirred at 20 °C for 24 h under nitrogen. The reaction mixture was left sitting at 20 °C over the weekend. The reaction mixture was diluted with water (4 mL) before neutralisation with saturated aqueous sodium bicarbonate solution. The aqueous was extracted with DCM (x 3).

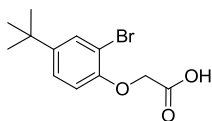
CONFIDENTIAL – Property of GSK – Do Not Copy

The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*. LCMS showed the product remained in the aqueous phase. The aqueous was blown down at 45 °C under nitrogen before dissolving the solid in water and passing through a Biotage 103 cartridge to remove excess salts/inorganics. The cartridge was washed with water. The cartridge was eluted with 1:1 MeOH/MeCN and the organic eluant was combined and blown down at 45 °C under nitrogen to afford a clear brown coloured gum. The gum was purified by reverse phase chromatography (0.1% formic acid in water / MeCN) using a Rediseq 15.5 g C18 aqueous cartridge. The appropriate fractions were combined and blown down at 45°C under nitrogen to afford 6-(3-pyrrolidinylcarbonyl)-2*H*-1,4-benzoxazin-3(4*H*)-one as a clear colourless gum containing minor impurities, 7.6 mg, 16% and 6-(3-pyrrolidinylmethyl)-2*H*-1,4-benzoxazin-3(4*H*)-one, as a clear colourless gum containing minor impurities, 8.3 mg, 19%.

6-(3-pyrrolidinylcarbonyl)-2*H*-1,4-benzoxazin-3(4*H*)-one, formic acid salt (**2.89**). ¹H NMR (600 MHz, MeOD) δ 8.54 (s, 1 H), 7.71 (dd, *J* = 8.4, 1.5 Hz, 1H), 7.57 (d, *J* = 1.5 Hz, 1H), 7.07 (d, *J* = 8.4 Hz, 1H), 4.69 (s, 2H), 4.29 - 4.24 (m, 1H), 3.68 (dd, *J* = 11.9, 5.0 Hz, 1H), 3.44 (dd, *J* = 11.9, 7.9 Hz, 1H), 3.39 - 3.34 (m, 1H), 3.26 (ddd, *J* = 11.6, 7.7, 7.5 Hz, 1H), 2.49 - 2.43 (m, 1 H), 2.16 – 2.08 (m, 1H), contains minor impurities; ¹³C NMR (MeOD, 101 MHz) δ 196.5, 168.9, 164.9, 148.4, 129.6, 127.0, 125.0, 116.2, 115.9, 66.5, 46.5, 44.8, 43.6, 28.8; LCMS method A: C₁₃H₁₅N₂O₃.CH₂O₂ [M+H]⁺ = 247, found 247, R_T = 0.36 min.

6-(3-pyrrolidinylmethyl)-2*H*-1,4-benzoxazin-3(4*H*)-one, formic acid salt (**2.64b**). ¹H NMR (600 MHz, MeOD) δ ppm 8.55 (s, 1H, overintegrates), 6.89 (d, *J* = 8.1 Hz, 1H), 6.83 (dd, *J* = 8.1, 1.5 Hz, 1H), 6.77 (d, *J* = 1.5 Hz, 1H), 4.54 (s, 2H), 3.40 - 3.35 (m, 1H), 3.33 (d, *J* = 7.5 Hz, 1H), 3.27-3.14 (m, 1H), 2.88 (dd, *J* = 11.6, 9.0 Hz, 1H), 2.75 – 2.65 (m, 2H), 2.60 - 2.52 (m, 1H), 2.12 - 2.06 (m, 1H), 1.73 - 1.65 (m, 1H); ¹³C NMR (MeOD, 101 MHz) δ 168.8, 166.2, 142.7, 134.0, 126.9, 123.6, 116.2, 115.8, 66.7, 49.4, 44.8, 39.8, 37.0, 29.6; LCMS method A: C₁₃H₁₇N₂O₂.CH₂O₂ [M+H]⁺ = 233, found 233, R_T = 0.40 min.

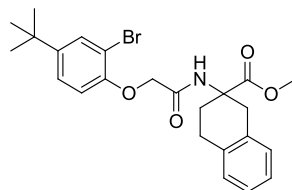
2-(2-Bromo-4-(*tert*-butyl)phenoxy)acetic acid (2.120)



K_2CO_3 (7.3 g, 52.7 mmol) and methyl chloroacetate (1.5 mL, 17.6 mmol) were added to a solution of 2-bromo-4-(*tert*-butyl)phenol (3.0 mL, 17.6 mmol) in DMF (20 mL). The reaction mixture was stirred at 20 °C overnight. The reaction mixture was concentrated *in vacuo* before partitioning between water and DCM (3 x 15 mL). The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*. The crude was dissolved in THF (5 mL) and MeOH (5 mL). 2.0 N NaOH (2 mL) was added and the reaction was stirred at 20 °C under nitrogen for 44 h. The reaction mixture was acidified with 2.0 N HCl and was extracted with EtOAc (3 x 20 mL). The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo* to afford 2-(2-bromo-4-(*tert*-butyl)phenoxy)acetic acid, as a white solid, 1.4 g, 27% (contains minor impurities and was hygroscopic). *Note: Low yield was due to incomplete hydrolysis. Neutralisation of the aqueous phase with saturated sodium bicarbonate solution and extraction with EtOAc (x3) afforded crude ethyl ester intermediate (2.6 g).*

1H NMR (400 MHz, DMSO- d_6) δ 7.52 (d, $J = 2.3$ Hz, 1H), 7.31 (dd, $J = 8.5, 2.3$ Hz, 1H), 6.89 (d, $J = 8.5$ Hz, 1H), 4.74 (s, 2H), 1.24 (s, 9H), carboxylic proton not seen, assumed under water peak as water has shifted down field (contains minor impurities); LCMS: m/z calcd for $C_{12}H_{14}^{79/81}BrO_3$ $[M+H]^+$ = 285/287, found 285/287, $R_T = 1.12$ min.

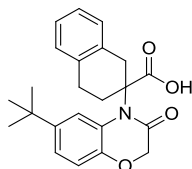
Methyl 2-(2-(2-bromo-4-(*tert*-butyl)phenoxy)acetamido)-1,2,3,4-tetrahydronaphthalene-2-carboxylate (2.121)



1-Chloro-*N,N*,2-trimethylprop-1-en-1-amine (60 μ L, 0.45 mmol) was added to a solution of 2-(2-bromo-4-(*tert*-butyl)phenoxy)acetic acid (130 mg, 0.45 mmol) in THF (3 mL). The reaction mixture was stirred at 20 °C for 15 min before addition of methyl 2-amino-1,2,3,4-tetrahydronaphthalene-2-carboxylate (93 mg, 0.45 mmol) and DIPEA (158 μ L, 0.91 mmol). The reaction mixture was stirred at 20 °C for 15 min before diluting with water. The reaction mixture was extracted with DCM (3 x 10 mL) and the combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*. The residue was purified by MDAP (method A, gradient E) to afford methyl 2-(2-(2-bromo-4-(*tert*-butyl)phenoxy)acetamido)-1,2,3,4-tetrahydronaphthalene-2-carboxylate, as a white solid, 75 mg, 35%.

^1H NMR (400 MHz, CDCl_3) δ ppm 7.44 (d, $J = 2.3$ Hz, 1H), 7.25 - 7.22 (dd, $J = 8.8, 2.3$ Hz, 2H), 7.18 - 7.08 (m, 4H), 6.71 (d, $J = 8.8$ Hz, 1H), 4.43 (apparent d, $J = 5.5, 2\text{H}$), 3.80 (s, 3H), 3.34 (d, $J = 16.6$ Hz, 1H), 3.16 (dd, $J = 16.6, 1.8$ Hz, 1H), 2.95 - 2.84 (m, 2H), 2.63 - 2.56 (m, 1H), 2.29 - 2.17 (m, 1H), 1.27 (s, 9H); LCMS: Method A $\text{C}_{24}\text{H}_{29}^{79/81}\text{BrNO}_4$ $[\text{M}+\text{H}]^+ = 474/476$, found 474/476, $R_T = 1.46$ min.

2-(6-(*tert*-Butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)-1,2,3,4-tetrahydronaphthalene-2-carboxylic acid (2.102)



Methyl 2-(2-(2-bromo-4-(*tert*-butyl)phenoxy)acetamido)-1,2,3,4-tetrahydronaphthalene-2-carboxylate (75 mg, 0.16 mmol), cesium carbonate (77 mg, 0.24 mmol), DBU (28 μ L, 0.19 mmol) and copper iodide (6.0 mg, 32 μ mol) were heated in

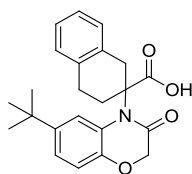
CONFIDENTIAL – Property of GSK – Do Not Copy

DMSO (1.0 mL) under microwave irradiation at 130 °C for 1h. DBU (28 µL, 0.19 mmol) and copper iodide (6.0 mg, 32 µmol) were added and the reaction was heated under microwave irradiation at 130 °C for a further 20 min. EtOAc (~10 mL) was added to the reaction which was extracted with 2N HCl (x 3). The organic was separated, dried and concentrated *in vacuo*. The residue was purified by reverse phase chromatography (10 mM aqueous ammonium bicarbonate / MeCN) to afford 2-(6-(*tert*-butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)-1,2,3,4-tetrahydronaphthalene-2-carboxylic acid, a colourless solid, 4.3 mg, 7%.

¹H NMR (400 MHz, CDCl₃) δ 7.12 - 7.01 (m, 4H), 7.01 – 6.92 (m, 2H), 6.86 (d, *J* = 8.6 Hz, 1H), 4.33 (d, *J* = 15.2 Hz, 1H), 4.20 (d, *J* = 15.2 Hz, 1H), 3.57 (br.s, 2H), 2.83 - 2.72 (m, 2H), 2.69 - 2.58 (m, 1H), 2.29 - 2.17 (m, 1H), 1.15 (s, 9H), carboxylic acid proton not seen; LCMS: method A C₂₃H₂₆NO₄ [M+H]⁺ = 380, found 380, R_T = 1.25 min.

Compound (**2.102**) was reprepared to complete full characterisation; Biochemical screening data was not obtained for this batch as the assay was no longer running.

2-(6-(*tert*-Butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)-1,2,3,4-tetrahydronaphthalene-2-carboxylic acid (2.102**) - Remake for full characterisation.**



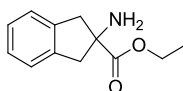
Methyl 2-(2-(2-bromo-4-(*tert*-butyl)phenoxy)acetamido)-1,2,3,4-tetrahydronaphthalene-2-carboxylate (400 mg, 0.84 mmol), cesium carbonate (412 mg, 1.27 mmol), DBU (153 µL, 1.01 mmol) and copper iodide (32 mg, 0.17 mmol) were heated in DMSO (10 mL) under microwave irradiation at 130 °C for 1h. After cooling, the reaction was quenched with 1M HCl (20 mL) solution. The reaction was extracted with EtOAc (2 x 20 mL). The organic was separated and was washed with water (20 mL) and saturated brine (20 mL). The organic was dried over sodium sulphate and

CONFIDENTIAL – Property of GSK – Do Not Copy

was evaporated *in vacuo*. The crude was purified by normal phase chromatography (silica gel; 100 – 200 mesh) eluting with DCM:MeOH (9:1). The compound was purified further by HPLC to afford 2-(6-(*tert*-butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)-1,2,3,4-tetrahydronaphthalene-2-carboxylic acid, a off-white solid, 22 mg, 6.8%. HPLC conditions - Mobile phase A = 10 mM ammonium bicarbonate; Mobile Phase B = MeCN; Column: YMC pak (25 x 150 mm); Flow = 25mL/min; Method: 0/10, 1/10, 10/50, 10.1/100, 15/100, 15.1/10.

¹H NMR (400 MHz, CDCl₃) δ 7.12 - 7.01 (m, 4H), 7.01 – 6.95 (m, 1H), 6.89 (d, *J* = 6.9 Hz, 1H), 6.85 (d, *J* = 8.4 Hz, 1H), 4.34 (d, *J* = 15.4 Hz, 1H), 4.24 (d, *J* = 15.2 Hz, 1H), 3.65 (d, *J* = 17.2 Hz, 1H), 3.55 (d, *J* = 17.2 Hz, 1H), 2.87 - 2.72 (m, 2H), 2.67 - 2.54 (m, 1H), 2.38 - 2.27 (m, 1H) 1.16 (s, 9H), carboxylic acid proton not seen; LCMS: C₂₃H₂₆NO₄ [M+H]⁺ = 380, found 380, R_T = 1.32 min – The UPLC analysis was conducted on an Acquity UPLC CSH C18 column (50mm x 2.1mm i.d. 1.7µm packing diameter) at 40 degrees centigrade, 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 Acetonitrile; ¹³C NMR (151 MHz, DMSO-*d*₆) δ 173.9, 169.7, 145.9, 144.1, 135.5, 134.9, 129.2, 128.5, 127.9, 125.5, 125.4, 120.2, 117.8, 115.7, 69.9, 65.6, 36.8, 34.0, 31.0, 29.3, 25.9; HRMS: m/z calcd for C₂₃H₂₆NO₄ [M+H]⁺ = 380.1856, found 380.1859; IR cm⁻¹ ~3650-2500 (broad), 2949, 1726, 1673, 1656; mp 188-189 °C.

Ethyl 2-amino-2,3-dihydro-1*H*-indene-2-carboxylate (2.125)



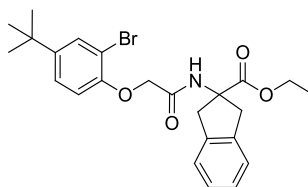
2-Amino-2,3-dihydro-1*H*-indene-2-carboxylic acid (197 mg, 1.11 mmol) was heated in EtOH (10 mL) and concentrated H₂SO₄ (889 µL, 16.7 mmol) was heated at 50 °C for 20 h before leaving to stand at 20 °C for 4 days. The reaction mixture was neutralised by addition of saturated aqueous ammonium bicarbonate solution. The reaction mixture was blown down at 45 °C to remove the ethanol before extracting the aqueous with EtOAc (3 x 10 mL). The combined organics were dried by passing

CONFIDENTIAL – Property of GSK – Do Not Copy

through a hydrophobic frit and were blown down at 45 °C under nitrogen to afford ethyl 2-amino-2,3-dihydro-1*H*-indene-2-carboxylate, 161 mg, 71%.

¹H NMR (400 MHz, CDCl₃) δ ppm 7.25 - 7.18 (m, 4H), 4.24 (q, *J* = 7.1 Hz, 2H), 3.58 (d, *J* = 16.0 Hz, 2H), 2.96 (d, *J* = 16.0 Hz, 2H), 1.30 (t, *J* = 7.1 Hz, 3H), amine protons not seen, contains minor impurities; ¹³C NMR (101 MHz, MeOD) δ ppm 177.5, 141.5, 127.7, 125.5, 65.5, 62.3, 46.5, 14.2; LCMS: method A C₁₂H₁₆NO₂ [M+H]⁺ = 206, found 206, R_T = 0.52 min; Note: HRMS (18 months after synthesis) showed 4 peaks suggesting instability of intermediate.

Ethyl 2-(2-(2-bromo-4-(*tert*-butyl)phenoxy)acetamido)-2,3-dihydro-1*H*-indene-2-carboxylate (2.126)

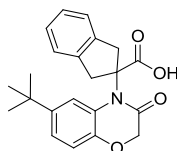


1-Chloro-*N,N*,2-trimethylprop-1-en-1-amine (66 μL, 0.50 mmol) was added to a solution of 2-(2-bromo-4-(*tert*-butyl)phenoxy)acetic acid (130 mg, 0.45 mmol) in THF (3 mL). The reaction mixture was stirred at 20 °C for 30 min before addition of ethyl 2-amino-2,3-dihydro-1*H*-indene-2-carboxylate (93 mg, 0.45 mmol) and DIPEA (0.16 mL, 0.91 mmol). The reaction mixture was stirred at 20 °C for 20 h before dilution with water. The reaction mixture was extracted with EtOAc (x 3). The combined organics were dried by passing through a hydrophobic frit and were concentrated by blow down at 45 °C under nitrogen. The residue was purified by reverse phase chromatography (10 mM aqueous ammonium bicarbonate / MeCN) using a Redisep 30 g C18 gold column to afford ethyl 2-(2-(2-bromo-4-(*tert*-butyl)phenoxy)acetamido)-2,3-dihydro-1*H*-indene-2-carboxylate, as a white powder, 90 mg, 42%.

¹H NMR (400 MHz, CDCl₃) δ 7.57 (d, *J* = 2.4 Hz, 1H), 7.53 (s, 1H), 7.33 (s, 1H), 7.32 (dd, *J* = 8.3, 2.4 Hz, 1H), 7.30 - 7.25 (m, 3H), 6.83 (d, *J* = 8.8 Hz, 1H), 4.53 (s,

2H), 4.31 (q, $J = 7.2$ Hz, 2H), 3.79 (d, $J = 16.4$ Hz, 2H), 3.42 (d, $J = 16.4$ Hz, 2H), 1.34 (s, 9H), 1.33 - 1.30 (t, $J = 7.2$ Hz, 3H); LCMS : Method A $C_{24}H_{29}^{79/81}BrNO_4$ $[M+H]^+ = 474/476$, found 474/476, $R_T = 1.47$ min.

2-(6-(*tert*-Butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)-2,3-dihydro-1*H*-indene-2-carboxylic acid, ammonia salt (2.100)

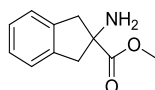


Ethyl 2-(2-(2-bromo-4-(*tert*-butyl)phenoxy)acetamido)-2,3-dihydro-1*H*-indene-2-carboxylate (50 mg, 0.11 mmol), cesium carbonate (52 mg, 0.16 mmol), DBU (19 μ L, 0.13 mmol) and CuI (4.0 mg, 0.02 mmol) were heated in DMSO (0.6 mL) under microwave irradiation at 130 °C for 4.5 h. The reaction was partitioned between EtOAc (~ 10 mL) and 2.0 N HCl (x3). The organics were separated, dried and were concentrated *in vacuo*. The residue was purified by MDAP (method B, gradient B) to afford 2-(6-(*tert*-butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)-2,3-dihydro-1*H*-indene-2-carboxylic acid ammonia salt, as a pale brown solid, 6.6 mg, 16%.

1H NMR (600 MHz, DMSO- d_6) δ 13.04 (br.s, 1H), 7.26 (d, $J = 3.5$ Hz, 2H), 7.24 - 7.19 (m, 2H), 7.04 - 6.97 (m, 2H), 6.71 (br.s, 1H), 4.52 (s, 2H), 3.86 (d, $J = 17.6$ Hz, 2H), 3.44 (d, $J = 17.8$ Hz, 2H), 1.02 (s, 9H); LCMS (method A): m/z calcd for $C_{22}H_{24}NO_4.H_3N$ $[M+H]^+ = 366$, found 366, $R_T = 1.24$ min.

Compound (2.100) (the free acid) was reprepared to complete full characterisation using a slightly modified synthetic route; Biochemical screening data was not obtained for this batch as the assay was no longer running.

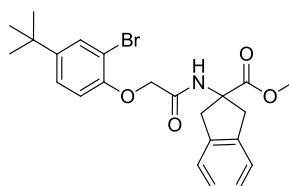
Methyl 2-amino-2,3-dihydro-1H-indene-2-carboxylate – For remake of (2.100)



A solution of 2-amino-2,3-dihydro-1H-indene-2-carboxylic acid, hydrochloride (2.3 g, 10.8 mmol) in MeOH (35 mL) was cooled to 0 °C. SOCl₂ (1.18 mL, 16.2 mmol) was added. The reaction mixture was stirred at 80°C for 12 h before concentrating *in vacuo*. The crude was triturated with diethyl ether (10 mL) and dried under vacuum to afford methyl 2-amino-2,3-dihydro-1H-indene-2-carboxylate, 2.2 g, 88%.

LCMS: C₁₁H₁₃NO₂ [M+H]⁺ = 192, found 192, R_T = 2.82 min; Column = Xbridge C18 (50 mm x 4.6 mm, 2.5 μm); Mobile Phase A = 5 mM ammonium acetate in water; Mobile phase B = MeCN; Time (min)/%B: 0/5, 0.5/5, 1/15, 3.3/98, 5.5/5, 6/5; Column Temp = 35 °C; Flow rate = 1.3 mL/min.

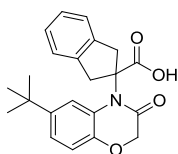
Methyl 2-(2-(2-bromo-4-(tert-butyl)phenoxy)acetamido)-2,3-dihydro-1H-indene-2-carboxylate - For remake of (2.100)



To a solution of 2-(2-bromo-4-(tert-butyl)phenoxy)acetic acid (300 mg, 1.05 mmol) in THF (10 mL) was added methyl 2-amino-2,3-dihydro-1H-indene-2-carboxylate (200 mg, 1.05 mmol) and 1-chloro-N,N,2-trimethylprop-1-en-1-amine (140 mg, 1.05 mmol). The reaction mixture was stirred under nitrogen at RT for 24 h. Water (20 mL) was added to the reaction mixture. The reaction was extracted with EtOAc (2 x 10 mL). The organic phase was separated, washed with water (5 mL) and saturated brine (5 mL), dried over sodium sulfate and concentrated *in vacuo*. The crude was purified by normal phase chromatography (silica gel; 100 – 200 mesh) eluting with Hexane:EtOAc (4:1) to afford methyl 2-(2-(2-bromo-4-(tert-butyl)phenoxy)acetamido)-2,3-dihydro-1H-indene-2-carboxylate, as a brown gum, 50 mg, 7.2%.

LCMS: $C_{23}H_{27}^{79/81}BrNO_4$ $[M+H]^+ = 460/462$, found 460/462, $R_T = 3.00$ min; Column = Aquity BEH C18 (50 mm x 2.1 mm, 1.7 μ m); Mobile phase A = 0.1 % formic acid in water; Mobile phase B = 0.1 % formic acid in MeCN; Time (min)/%B: 0/3, 0.4/3, 3.2/98, 3.8/98, 4.2/3, 4.5/3; Column Temp = 35 °C; Flow rate = 0.6 mL/min.

Remake of 2-(6-(*tert*-Butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)-2,3-dihydro-1*H*-indene-2-carboxylic acid (2.100)



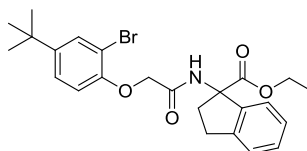
Methyl 2-(2-(2-bromo-4-(*tert*-butyl)phenoxy)acetamido)-2,3-dihydro-1*H*-indene-2-carboxylate (200 mg, 0.43 mmol) and cesium carbonate (255 mg, 0.78 mmol) in 1,4-dioxane (3 mL) were stirred under nitrogen at 25 °C. The reaction mixture was degassed with argon for 15 min. $Pd_2(dba)_3$ (40 mg, 0.04 mmol) and 2-dicyclohexylphosphino-2',4',6'-triisopropylbiphenyl (41 mg, 0.09 mmol) were added and the reaction mixture was stirred at 90 °C for 16 h. The reaction mixture was filtered and was washed with EtOAc (5 mL). The filtrate was washed with water (5 mL). The organic layer was dried with sodium sulphate and concentrated *in vacuo*. The crude was purified by prep TLC (silica plate) eluting with 7% EtOAc. THF (4 mL), water (2 mL) and LiOH (17 mg, 0.73 mmol) were added to the crude. The reaction mixture was stirred under nitrogen at 25 °C for 24 h. Ether (8 mL) was added and the layers were separated. The pH of the aqueous layer was adjusted to 2 with 1N HCl. The aqueous layer was extracted with EtOAc (8 mL). The organic layer was separated and dried with sodium sulphate before concentrating *in vacuo* to afford 2-(6-(*tert*-butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)-2,3-dihydro-1*H*-indene-2-carboxylic acid ammonia salt, as a pale brown solid, 41 mg, 25%.

1H NMR (400 MHz, DMSO- d_6) δ 12.83 (s, 1H, underintegrates), 7.34 – 7.28 (m, 2H), 7.28 – 7.20 (m, 2H), 7.04 (s, 1H), 6.60 (br.s, 1H), 4.57 (s, 2H), 3.88 (d, $J = 18.1$

CONFIDENTIAL – Property of GSK – Do Not Copy

Hz, 2H), 3.48 (d, $J = 18.1$ Hz, 2H), 1.00 (s, 9H), residual MeOH observed; ^{13}C NMR (101 MHz, DMSO- d_6) δ 174.1, 170.4, 145.7, 145.5, 140.1, 128.0, 127.6, 124.8, 121.1, 117.3, 116.0, 69.7, 68.2, 42.9, 34.5, 31.3; LCMS: m/z calcd for $\text{C}_{22}\text{H}_{24}\text{NO}_4$ $[\text{M}+\text{H}]^+ = 366$, found 366, $R_T = 1.30$ min – The UPLC analysis was conducted on an Acquity UPLC CSH C18 column (50 mm x 2.1mm i.d. 1.7 μm packing diameter) at 40 degrees centigrade, 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 Acetonitrile; HRMS: m/z calcd for $\text{C}_{22}\text{H}_{24}\text{NO}_4$ $[\text{M}+\text{H}]^+ = 366.1700$, found 366.1701; IR cm^{-1} 3671, 3159, 2973, 1744, 1658; mp 225-227 °C.

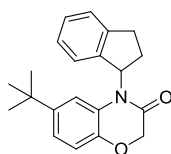
Ethyl 1-(2-(2-bromo-4-(*tert*-butyl)phenoxy)acetamido)-2,3-dihydro-1*H*-indene-1-carboxylate (2.130)



1-Chloro-*N,N*,2-trimethylprop-1-en-1-amine (43 μL , 0.33 mmol) was added to a solution of 2-(2-bromo-4-(*tert*-butyl)phenoxy)acetic acid (84 mg, 0.29 mmol) in THF (2 mL). The reaction mixture was stirred at 20 °C for 30 min under nitrogen before addition of ethyl 1-amino-2,3-dihydro-1*H*-indene-1-carboxylate (60 mg, 0.29 mmol) and DIPEA (102 μL , 0.59 mmol). The reaction mixture was stirred at 20 °C for 62 h before heating at 45 °C for 5 h. The reaction mixture was cooled, diluted with water (10 mL) and extracted with DCM (3 x 10 mL). The combined organics were dried by passing through a hydrophobic frit and were concentrated by blow down under nitrogen at 45 °C. The residue was purified by MDAP (method A, gradient E). The appropriate fractions were neutralised by addition of saturated aqueous sodium bicarbonate solution. The fractions were blown down under nitrogen at 45 °C to remove the volatile organic solvent. The aqueous was extracted with DCM (3 x 5 mL). The combined organics were dried by passing through a hydrophobic frit and were blown down at 45 °C under nitrogen to afford ethyl 1-(2-(2-bromo-4-(*tert*-butyl)phenoxy)acetamido)-2,3-dihydro-1*H*-indene-1-carboxylate, as a pale brown solid, 34 mg, 23%.

¹H NMR (400 MHz, CDCl₃) δ 7.73 (s, 1H), 7.53 (d, *J* = 2.3 Hz, 1H), 7.37 (d, *J* = 7.5 Hz, 1H), 7.29 - 7.18 (m, 4H), 6.79 (d, *J* = 8.8 Hz, 1H), 4.49 (s, 2H), 4.21 - 4.09 (m, 2H), 3.19 - 3.07 (m, 3H), 2.45 - 2.36 (m, 1H), 1.27 (s, 9H), 1.18 (t, *J* = 7.2 Hz, 3H); LCMS: Method A C₂₄H₂₉^{79/81}BrNO₄ [M+H]⁺ = 474/476, found 474/476, R_T = 1.49 min.

6-(*tert*-Butyl)-4-(2,3-dihydro-1*H*-inden-1-yl)-2*H*-benzo[*b*][1,4]oxazin-3(4*H*)-one (2.131)

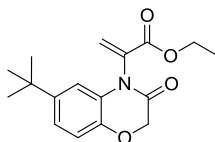


1-(2-(2-Bromo-4-(*tert*-butyl)phenoxy)acetamido)-2,3-dihydro-1*H*-indene-1-carboxylate (34 mg, 0.07 mmol), cesium carbonate (35 mg, 0.11 mmol), DBU (13 μL, 86 μmol) and CuI (2.7 mg, 14 μmol) were heated in DMSO (0.6 mL) under microwave irradiation at 130 °C for 8 h. The reaction mixture was partitioned between water and DCM. The aqueous was washed with DCM (10 mL x 2). The combined organics were washed with brine and dried by passing through a hydrophobic frit. The organic was concentrated by blown down under nitrogen at 45 °C. The residue was purified by MDAP (method A, gradient E). The appropriate fractions were neutralised by addition of saturated aqueous sodium bicarbonate solution. The fractions were blown down under nitrogen at 45 °C to remove the volatile organic solvent before extracting with DCM (3 x 5 mL). The combined organics were dried by passing through a hydrophobic frit and were blown down at 45 °C under nitrogen to afford 6-(*tert*-butyl)-4-(2,3-dihydro-1*H*-inden-1-yl)-2*H*-benzo[*b*][1,4]oxazin-3(4*H*)-one, 5.4 mg, 23%.

¹H NMR (400 MHz, CDCl₃) δ 7.35 (d, *J* = 7.5 Hz, 1H), 7.24 (t, *J* = 7.5 Hz, 1H), 7.14 (t, *J* = 7.5 Hz, 1H), 7.05 (d, *J* = 7.5 Hz, 1H), 6.90 (s, 2H), 6.77 (br.t, *J* = 8.9 Hz, 1H), 6.45 (s, 1H), 4.67 (s, 2H), 3.23 (ddd, *J* = 16.3, 10.1, 2.5 Hz, 1H), 3.17 - 3.07 (m, 1H),

2.64 - 2.55 (m, 1H), 2.43 - 2.32 (m, 1H), 0.96 (s, 9H), contains minor impurities;
LCMS: Method A C₂₁H₂₄NO₂ [M+H]⁺ = 322, found 322, R_T = 1.43 min.

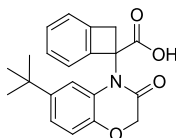
Ethyl 2-(6-(*tert*-butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)acrylate (2.143)



Ethyl propiolate (0.20 mL, 1.95 mmol) in DCM (5 mL) was added dropwise over 10 min to an ice cooled solution of 6-(*tert*-butyl)-2*H*-benzo[*b*][1,4]oxazin-3(4*H*)-one (0.40 g, 1.95 mmol) and triphenylphosphine (0.51 g, 1.95 mmol) in DCM (10 mL) under nitrogen. The reaction mixture was warmed to 20 °C and was stirred under nitrogen for 3.5 h before concentrating *in vacuo*. The crude mixture was purified by normal phase chromatography (cyclohexane / *t*-butyl methyl ether) using a Redisep silicycle 24 g column to afford ethyl 2-(6-(*tert*-butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)acrylate, a brown gum, 384 mg, 65%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 7.06 (dd, *J* = 6.9, 2.3 Hz, 1H), 6.99 (d, *J* = 6.9 Hz, 1H), 6.77 (d, *J* = 0.5 Hz, 1H), 6.71 (d, *J* = 2.0 Hz, 1H), 6.19 (d, *J* = 0.5 Hz, 1H), 4.70 (s, 2H), 4.18 (q, *J* = 7.1 Hz, 2H), 1.21 (s, 9H), 1.15 (t, *J* = 7.1 Hz, 3H); LCMS : Method A C₁₇H₂₂NO₄ [M+H]⁺ = 304, found 304, R_T = 1.22 min.

7-(6-(*tert*-Butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)bicyclo[4.2.0]octa-1(6*9*),2,4-triene-7-carboxylic acid (2.101)



To a mixture of ethyl 2-(6-(*tert*-butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)acrylate (167 mg, 0.55 mmol) and cesium fluoride (209 mg, 1.38 mmol) in anhydrous MeCN (5 mL) under nitrogen in a sealed vessel, was added 2-(trimethylsilyl)phenyl trifluoromethanesulfonate (134 μL, 0.55 mmol). The reaction was stirred at 20 °C for 68 h before acidifying with 2.0 M HCl (0.83 mL, 1.66

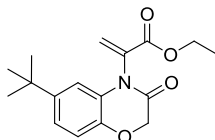
CONFIDENTIAL – Property of GSK – Do Not Copy

mmol). The reaction was immediately basified by addition of saturated aqueous sodium bicarbonate solution. The reaction mixture was extracted with EtOAc (3 x 5 mL). The combined organics were dried by passing through a hydrophobic frit and were concentrated by blow down at 45 °C under nitrogen. The crude was dissolved in EtOH (1 mL) and THF (3 mL). 2.0 N NaOH (1 mL) was added and the reaction mixture was stirred at RT for 94 h. The reaction mixture was blown down under nitrogen at 45 °C to remove the volatile solvents. The aqueous was acidified with 2.0 N HCl (5 mL) and was immediately extracted with DCM (3 x 5 mL). The combined organics were dried by passing through a hydrophobic frit and were concentrated by blow down at 45 °C under nitrogen. The residue was purified by MDAP (method B, gradient B). The material was further purified using MDAP (method A, gradient D). The fractions were neutralised by addition of saturated aqueous sodium bicarbonate solution before removing the volatiles by blown down at 45 °C under nitrogen. The aqueous was extracted with DCM (x 3). The combined organics were dried by passing through a hydrophobic frit and were concentrated by blow down at 45 °C under nitrogen to afford 7-(6-(*tert*-butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)bicyclo[4.2.0]octa-1(6),2,4-triene-7-carboxylic acid, as a colourless gum, 6.5 mg, 3.4% over 2 steps.

¹H NMR (400 MHz, CDCl₃) δ 7.71 (d, *J* = 7.5 Hz, 1H), 7.37 (apparent t, *J* = 7.5 Hz, 1H), 7.30 (apparent t, *J* = 7.3 Hz, 1H), 7.15 (d, *J* = 7.3 Hz, 1H), 7.09 (dd, *J* = 8.4, 2.2 Hz, 1H), 7.01 (d, *J* = 2.2 Hz, 1H), 6.97 (d, *J* = 8.4 Hz, 1H), 4.74 (d, *J* = 15.9 Hz, 1H), 4.52 (d, *J* = 15.9 Hz, 1H), 4.20 (d, *J* = 14.2 Hz, 1H), 3.52 (d, *J* = 14.2 Hz, 1H) 1.33 (s, 9H), carboxylic acid proton not seen, contains minor aliphatic impurities; ¹³C NMR (101 MHz, CDCl₃) δ 172.1, 171.6, 146.4, 144.2, 142.4, 141.1, 130.6, 128.1, 128.0, 126.6, 122.9, 121.7, 116.8, 115.2, 69.2, 67.8, 41.0, 34.6, 31.4; LCMS : (method A) C₂₁H₂₂NO₄ [M+H]⁺ = 352, found 352, R_T = 1.04 min; mp 87-89 °C (decomposed).

Compound (**2.101**) was reprepared to complete full characterisation using a slightly modified synthetic route; Biochemical screening data was not obtained for this batch as the assay was no longer running.

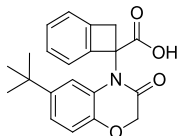
Remake of ethyl 2-(6-(*tert*-butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)acrylate (2.143)



Ethyl propiolate (1.1 g, 10.7 mmol) was added dropwise over 10 min to an ice cooled solution of 6-(*tert*-butyl)-2*H*-benzo[*b*][1,4]oxazin-3(4*H*)-one (2.2 g, 10.7 mmol) and triphenylphosphine (2.8 g, 10.7 mmol) in DCM (5 mL) under nitrogen at 0 °C. The reaction mixture was warmed to 25 °C and was stirred under nitrogen for 12 h before concentrating *in vacuo*. The crude mixture was purified by normal phase chromatography eluting with 5% EtOAc in hexane to afford ethyl 2-(6-(*tert*-butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)acrylate, a white solid, 2.0 g, 61%.

¹H NMR (400 MHz, CDCl₃) δ ppm 7.03 (dd, *J* = 8.3, 2.20 Hz, 1H) 6.95 (d, *J* = 8.3 Hz, 1H) 6.80 (s, 1H) 6.75 (d, *J* = 2.0 Hz, 1H) 5.99 (s, 1H) 4.66 (s, 2H) 4.24 (q, *J* = 7.2 Hz, 2H) 1.25 (s, 9H) 1.21 - 1.25 (t, *J* = 7.2 Hz, 3H); LCMS: C₁₇H₂₂NO₄ [M+H]⁺ = 304, found 304, R_T = 3.84 min; Column = Xbridge C18 (50 mm x 4.6 mm, 2.5 μm); Mobile Phase A = 5 mM ammonium acetate in water; Mobile phase B = MeCN; Time (min)/%B: 0/5, 0.5/5, 1/15, 3.3/98, 5.5/5, 6/5; Column Temp = 35 °C; Flow rate = 1.3 mL/min.

Remake of 7-(6-(*tert*-Butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)bicyclo[4.2.0]octa-1(69),2,4-triene-7-carboxylic acid (2.101)



To a mixture of ethyl 2-(6-(*tert*-butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)acrylate (1.0 g, 3.5 mmol) and cesium fluoride (0.53 g, 3.5 mmol) in MeCN (8 mL) under nitrogen in a sealed vessel, was added 2-(trimethylsilyl)phenyl trifluoromethanesulfonate (1.0 g, 3.5 mmol). The reaction was stirred at 25 °C for 68

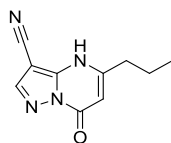
CONFIDENTIAL – Property of GSK – Do Not Copy

h before acidifying with 2N HCl to pH 1. The reaction was immediately basified by addition of saturated aqueous sodium bicarbonate solution to pH 9. The reaction mixture was extracted with EtOAc (3 x 5 mL). The combined organics were dried by sodium sulfate and were concentrated by blow down at 45 °C. The crude material was dissolved in EtOH (6 mL) and THF (9 mL). 2.0 N NaOH (8 mL) was added and the reaction mixture was stirred at RT for 94 h. The reaction mixture was blown down under nitrogen at 45 °C to remove the volatile solvents. The aqueous was acidified with 2.0 N HCl to pH 1 and was immediately extracted with DCM (3 x 5 mL). The combined organics were dried by sodium sulphate and were concentrated by blow down at 45 °C under nitrogen. The residue was purified by prep-HPLC to afford 7-(6-(*tert*-butyl)-3-oxo-2*H*-benzo[*b*][1,4]oxazin-4(3*H*)-yl)bicyclo[4.2.0]octa-1(6),2,4-triene-7-carboxylic acid, as a colourless gum, 160 mg, 12% over 2 steps. HPLC conditions - Mobile phase A = 10 mM ammonium bicarbonate (aq); Mobile Phase B = MeCN; Column: C18 (30 x 150 mm); Flow = 16mL/min; Method: 0/10, 1/10, 10/50.

¹H NMR (400 MHz, CDCl₃) δ 7.71 (d, *J* = 7.6 Hz, 1H), 7.38 (apparent t, *J* = 7.5 Hz, 1H), 7.31 (apparent t, *J* = 7.3 Hz, 1H), 7.15 (d, *J* = 7.3 Hz, 1H), 7.09 (dd, *J* = 8.3, 2.2 Hz, 1H), 7.01 (d, *J* = 2.0 Hz, 1H), 6.97 (d, *J* = 8.3 Hz, 1H), 4.74 (d, *J* = 15.9 Hz, 1H), 4.52 (d, *J* = 15.9 Hz, 1H), 4.20 (d, *J* = 14.2 Hz, 1H), 3.52 (d, *J* = 14.2 Hz, 1H) 1.33 (s, 9H), carboxylic acid proton not seen, contains minor aliphatic impurities; LCMS: method A C₂₁H₂₂NO₄ [M+H]⁺ = 352, found 352, R_T = 1.28 min – The UPLC analysis was conducted on an Acquity UPLC CSH C18 column (50 mm x 2.1mm i.d. 1.7µm packing diameter) at 40 degrees centigrade, 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 Acetonitrile; HRMS: m/z calcd for C₂₁H₂₂NO₄ [M+H]⁺ = 352.1543, found 352.1549; IR cm⁻¹ 3675, 2972, 2902, 1702.

6.2 Experimental Details – CD38

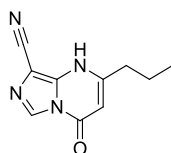
7-Oxo-5-propyl-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.10)



5-Amino-1*H*-pyrazole-4-carbonitrile (105 mg, 0.97 mmol) was heated with ethyl 3-oxohexanoate (214 mg, 1.35 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 17 h. The reaction mixture was blown down at 45 °C under nitrogen. The crude was purified by reverse phase chromatography (MDAP method A, gradient A) to afford 7-oxo-5-propyl-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as a off white solid, 79 mg, 40%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 13.19 (br.s, 1H), 8.34 (s, 1H), 5.84 (s, 1H), 2.57 (t, *J* = 7.3 Hz, 2H), 1.70 (apparent sextet, *J* = 7.5 Hz, 2H), 0.94 (t, *J* = 7.3 Hz, 3H); LCMS method A: C₁₀H₁₁N₄O [M+H]⁺ = 203, found 203, R_T = 0.56 min.

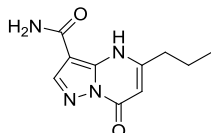
4-Oxo-2-propyl-1,4-dihydroimidazo[1,5-*a*]pyrimidine-8-carbonitrile (4.13)



4-Amino-1*H*-imidazole-5-carbonitrile (200 mg, 1.85 mmol), ethyl 3-oxohexanoate (0.36 mL, 2.22 mmol), acetic acid (4 mL) and water (2 mL) were heated at 100 °C for 4 hours. Additional ethyl 3-oxohexanoate (0.36 mL, 2.22 mmol) was added, and the reaction mixture was heated at 100 °C for a further 5 days. The reaction mixture was cooled to room temperature, and concentrated *in vacuo*. The crude was purified using a 50 g C18 aqueous cartridge using a gradient of 0 – 50% methanol in water containing 0.1% formic acid. The appropriate fractions were concentrated *in vacuo* to afford 4-oxo-2-propyl-1,4-dihydroimidazo[1,5-*a*]pyrimidine-8-carbonitrile, as a cream solid, 35 mg, 9.3%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 13.01 (br.s, 1H), 8.24 (s, 1H), 5.67 (s, 1H), 2.57 (t, *J* = 7.8 Hz, 2H), 1.73 - 1.63 (m, 2H), 0.94 (t, *J* = 7.3 Hz, 3H); LCMS method A: C₁₀H₁₁N₄O [M+H]⁺ = 203, found 203, R_T = 0.61 min.

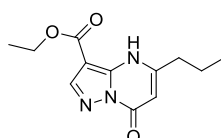
7-Oxo-5-propyl-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carboxamide (4.19)



3-Amino-1*H*-pyrazole-4-carboxamide (122 mg, 0.97 mmol) was heated with ethyl 3-oxohexanoate (214 mg, 1.35 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 17 h. The reaction mixture was blown down under nitrogen at 45 °C. The crude was purified by reverse phase chromatography (MDAP method A, gradient A) to afford 7-oxo-5-propyl-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carboxamide, as a white solid, 132 mg, 62%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 11.49 (s, 1H), 8.31 (s, 1H), 7.78 (br.s, 1H), 7.25 (br.s, 1H), 5.74 (br.s, 1H), 2.69 (t, *J* = 7.6 Hz, 2H), 1.68 - 1.59 (m, 2H), 0.93 (t, *J* = 7.3 Hz, 3H); LCMS method A: C₁₀H₁₃N₄O₂ [M+H]⁺ = 221, found 221, R_T = 0.50 min.

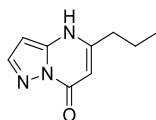
Ethyl 7-oxo-5-propyl-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carboxylate (4.20)



Ethyl 3-imino-2,3-dihydro-1*H*-pyrazole-4-carboxylate (150 mg, 0.97 mmol) was heated with ethyl 3-oxohexanoate (214 mg, 1.35 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 17 h. The reaction mixture was blown down under nitrogen at 45 °C. The crude was purified by reverse phase chromatography (MDAP method A, gradient B) to afford 7-oxo-5-propyl-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carboxamide, as a white solid, 152 mg, 63%.

^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 11.78 (br.s, 1H), 8.17 (s, 1H), 5.84 (s, 1H), 4.31 (q, $J = 7.1$ Hz, 2H), 2.72 - 2.67 (t, $J = 7.7$ Hz, 2H), 1.70 - 1.61 (m, 2H), 1.32 (t, $J = 7.1$ Hz, 3H), 0.94 (t, $J = 7.3$ Hz, 3H); LCMS method A: $\text{C}_{12}\text{H}_{14}\text{N}_3\text{O}_3$ $[\text{M}+\text{H}]^- = 248$, found 248, $R_T = 0.74$ min.

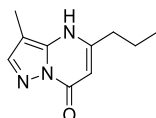
5-Propylpyrazolo[1,5-*a*]pyrimidin-7(4*H*)-one (4.22)



1*H*-Pyrazol-5-amine (81 mg, 0.98 mmol) was heated with ethyl 3-oxohexanoate (216 mg, 1.37 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 17 h. The reaction mixture was blown down under nitrogen at 45 °C. The crude was purified by reverse phase chromatography (MDAP method A, gradient A) to afford 5-propylpyrazolo[1,5-*a*]pyrimidin-7(4*H*)-one, as an off white solid, 52 mg, 30%.

^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 12.17 (br.s, 1H), 7.81 (d, $J = 1.9$ Hz, 1H), 6.09 (d, $J = 1.9$ Hz, 1H), 5.56 (s, 1H), 2.52 (t, $J = 7.6$ Hz, assumed 2H as overlaps with DMSO), 1.71 - 1.62 (m, 2H), 0.92 (t, $J = 7.3$ Hz, 3H); LCMS method A: $\text{C}_9\text{H}_{12}\text{N}_3\text{O}$ $[\text{M}+\text{H}]^+ = 178$, found 178, $R_T = 0.50$ min.

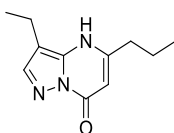
3-Methyl-5-propylpyrazolo[1,5-*a*]pyrimidin-7(4*H*)-one (4.23)



4-Methyl-1*H*-pyrazol-5-amine (95 mg, 0.98 mmol) was heated with ethyl 3-oxohexanoate (216 mg, 1.37 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 17 h. The reaction mixture was blown down under nitrogen at 45 °C. The crude was purified by reverse phase chromatography (MDAP method A, gradient A) to afford 3-methyl-5-propylpyrazolo[1,5-*a*]pyrimidin-7(4*H*)-one, as an off white solid, 80 mg, 43%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 11.81 (br.s, 1H), 7.67 (s, 1H), 5.52 (s, 1H), 2.56 - 2.52 (t, *J* = 7.8 Hz, assumed 2H as overlaps with DMSO), 2.13 (s, 3H), 1.72 - 1.63 (m, 2H), 0.94 (t, *J* = 7.3 Hz, 3H); LCMS method A: C₁₀H₁₄N₃O [M+H]⁺ = 192, found 192, R_T = 0.56 min.

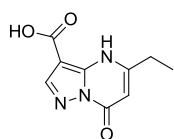
3-Ethyl-5-propylpyrazolo[1,5-*a*]pyrimidin-7(4*H*)-one (4.24)



4-Ethyl-1*H*-pyrazol-5-amine (108 mg, 0.98 mmol) was heated with ethyl 3-oxohexanoate (216 mg, 1.37 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 17 h. The reaction mixture was blown down under nitrogen at 45 °C. The crude was purified by reverse phase chromatography (MDAP method A, gradient A, extended run) to afford 3-ethyl-5-propylpyrazolo[1,5-*a*]pyrimidin-7(4*H*)-one, as an off white solid, 44 mg, 22%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 11.79 (br.s, 1H), 7.73 (s, 1H), 5.52 (s, 1H), 2.60 - 2.50 (m, assumed 4H as overlaps with DMSO), 1.72 - 1.63 (m, 2H), 1.17 (t, *J* = 7.6 Hz, 3H), 0.94 (t, *J* = 7.3 Hz, 3H); LCMS method A: C₁₁H₁₆N₃O [M+H]⁺ = 206, found 206, R_T = 0.64 min.

5-Ethyl-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carboxylic acid (4.31)

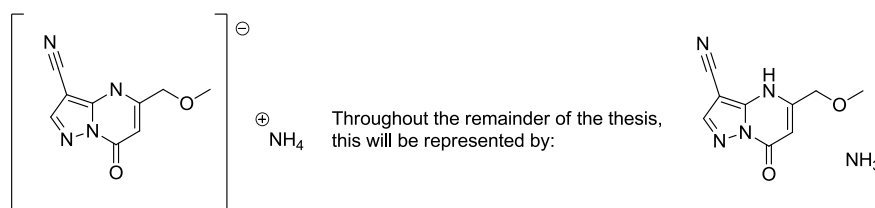


5-Ethyl-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (62 mg, 0.33 mmol) was suspended in water (1 mL) and 10 M NaOH (2 mL). The reaction mixture was heated at reflux for 23 hours. The reaction mixture was acidified to pH1 by dropwise addition into 5.0 N HCl. The aqueous solution was applied to a Biotage isolate 103 cartridge (500 mg). The cartridge was eluted with 10 x column volumes of water. The compound was eluted by washing the cartridge with 10 x column volumes of 1:1 MeOH:MeCN. The organic was concentrated by blow down under

nitrogen at 45 °C and dried in the vacuum oven to afford 5-ethyl-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carboxylic acid, as a white solid, 51 mg, 75%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.12 (s, 1H), 5.79 (s, 1H), 2.73 (q, *J* = 7.5 Hz, 2H), 1.20 (t, *J* = 7.4 Hz, 3H), exchangeable proton very broad at ~ 12–13 ppm; LCMS method A: C₉H₁₀N₃O₃ [M+H]⁺ = 208, found 208, R_T = 0.42 min.

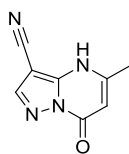
5-(Methoxymethyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, ammonia salt (4.36b)



Methyl 4-methoxy-3-oxobutanoate (379 mg, 2.59 mmol) was heated with 5-amino-1*H*-pyrazole-4-carbonitrile (200 mg, 1.85 mmol) in water (2 mL) and acetic acid (4 mL) at 100 °C for 42 hours. The reaction mixture was left to cool to room temperature and was concentrated *in vacuo*. The crude was purified using reverse phase chromatography (MDAP method B, gradient A, methanol in place of acetonitrile, with at column dilution) to afford 5-(methoxymethyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile ammonia salt, as a pale yellow solid, 233 mg, 57%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.00 (s, 1H), 7.09 (s, 4H), 5.62 (s, 1H), 4.22 (s, 2H), 3.33 (s, 3H); LCMS method A: C₉H₉N₄O₂.H₃N [M+H]⁺ = 205, found 205, R_T = 0.45 min.

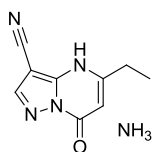
5-Methyl-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.37)



5-Amino-1*H*-pyrazole-4-carbonitrile (200 mg, 1.85 mmol) was heated with ethyl 3-oxobutanoate (0.33 mL, 2.59 mmol) in water (3.00 mL) and acetic acid (6 mL) at 100 °C for 42 hours. The reaction mixture was left to cool to room temperature and was concentrated *in vacuo*. The crude was pre-absorbed onto diatomaceous earth and was purified by reverse phase chromatography (10 mM ammonium bicarbonate in aqueous solution (solvent A), 20% methanol in 10 mM ammonium bicarbonate in aqueous solution (solvent B)), using a Rediseq 100 g C18 gold aqueous column to afford 5-methyl-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as a white solid, 38 mg, 12%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.03 (s, 1H), 7.10 (br.s, 1H - overintegrates), 5.51 (s, 1H), 2.19 (s, 3H), very broad peaks are observed at 12.02, 7.49, 6.27 totalling to approximately 1H; LCMS method B: C₈H₇N₄O [M+H]⁺ = 175, found 175, R_T = 0.38 min

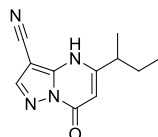
5-Ethyl-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, ammonia salt (4.38b)



5-Amino-1*H*-pyrazole-4-carbonitrile (200 mg, 1.85 mmol) was heated with ethyl 3-oxopentanoate (373 mg, 2.59 mmol) in water (2 mL) and acetic acid (4 mL) at 100 °C for 42 hours. The reaction mixture was left to cool to room temperature and was concentrated *in vacuo*. The crude was purified using reverse phase chromatography (MDAP method B, gradient A, methanol in place of acetonitrile, with at column dilution) to afford 5-ethyl-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile ammonia salt, as a white solid, 75 mg, 20%.

^1H NMR (400 MHz, DMSO- d_6) δ 7.96 (s, 1H), 7.06 (br.s, 4H), 5.45 (s, 1H), 2.43 (q, $J = 7.6$ Hz, 2H), 1.16 (t, $J = 7.6$ Hz, 3H); LCMS method B: $\text{C}_9\text{H}_9\text{N}_4\text{O} \cdot \text{H}_3\text{N}$ $[\text{M}+\text{H}]^+ = 189$, found 189, $R_T = 0.45$ min.

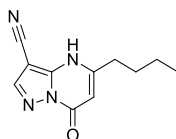
5-(*sec*-Butyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.39)



5-Amino-1*H*-pyrazole-4-carbonitrile (45 mg, 0.42 mmol) was heated with ethyl 4-methyl-3-oxohexanoate (100 mg, 0.58 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 18 h. The reaction mixture was blown down under nitrogen at 45 °C. The crude purified by reverse phase chromatography (0.1% formic acid in water (solvent A), acetonitrile (solvent B)), using a Rediseq 50 g C18 gold aqueous column to afford 5-(*sec*-butyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as an off white solid, 14 mg, 16%.

^1H NMR (400 MHz, DMSO- d_6) δ 13.08 (s, 1H), 8.37 (s, 1H), 5.86 (s, 1H), 2.75 - 2.65 (m, 1H), 1.76 - 1.65 (m, 1H), 1.64 - 1.53 (m, 1H), 1.24 (d, $J = 6.8$ Hz, 3H), 0.87 (t, $J = 7.3$ Hz, 3H); LCMS method A: $\text{C}_{11}\text{H}_{13}\text{N}_4\text{O}$ $[\text{M}+\text{H}]^+ = 217$, found 217, $R_T = 0.63$ min.

5-Butyl-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.40)

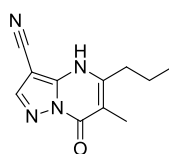


Ethyl 3-oxoheptanoate (478 mg, 2.78 mmol) was heated with 5-amino-1*H*-pyrazole-4-carbonitrile (200 mg, 1.85 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 72 hours. The reaction mixture was cooled and was blown down at 45 °C under nitrogen. MeOH (10 mL) was added, followed by water (20 mL). The precipitate was collected by filtration and was dried by blow down at 45 °C under nitrogen. The resultant solid was purified by reverse phase chromatography (MDAP method A,

gradient B) to afford 5-butyl-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as a white powder, 27 mg, 7%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 13.21 (br.s, 1H), 8.35 (s, 1H), 5.85 (s, 1H), 2.59 (t, *J* = 7.6 Hz, 2H), 1.67 - 1.59 (m, 2H), 1.40 - 1.31 (m, 2H), 0.92 (t, *J* = 7.3 Hz, 3H); LCMS method A: C₁₁H₁₃N₄O [M+H]⁺ = 217, found 217, R_T = 0.66 min.

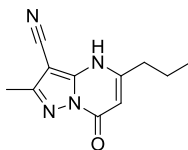
6-Methyl-7-oxo-5-propyl-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.47)



Methyl 2-methyl-3-oxohexanoate (439 mg, 2.78 mmol) was heated with 5-amino-1*H*-pyrazole-4-carbonitrile (200 mg, 1.85 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 72 h. The reaction mixture was left to cool. The precipitated solid was collected by filtration and was washed with water (10 mL) and MeOH (10 mL). The solid was dried by blow down at 45 °C under nitrogen to afford 6-methyl-7-oxo-5-propyl-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as a cream solid, 133 mg, 33%. Note: Additional product in the filtrate was not purified due to sufficient quantity for biochemical screening.

¹H NMR (400 MHz, DMSO-*d*₆) δ 13.00 (s, 1H), 8.34 (s, 1H), 2.68 - 2.61 (m, 2H), 2.01 (s, 3H), 1.68 - 1.58 (m, 2H), 0.97 (t, *J* = 7.3 Hz, 3H); LCMS method A: C₁₁H₁₃N₄O [M+H]⁺ = 217, found 217, R_T = 0.64 min.

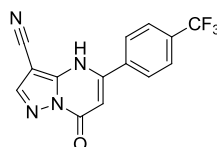
2-Methyl-7-oxo-5-propyl-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.49)



5-Amino-3-methyl-1*H*-pyrazole-4-carbonitrile (118 mg, 0.97 mmol) was heated with ethyl 3-oxohexanoate (214 mg, 1.35 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 17 h. The reaction mixture was cooled and was blown down at 45 °C under nitrogen. The crude was purified by reverse phase chromatography (MDAP method A, gradient A) to afford 2-methyl-7-oxo-5-propyl-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as an off white solid, 106 mg, 51%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 13.07 (s, 1H), 5.80 (s, 1H), 2.58 - 2.52 (m, 2H), 2.37 (s, 3H), 1.71 - 1.62 (m, 2H), 0.93 (t, *J* = 7.4 Hz, 3H); LCMS method A: C₁₁H₁₃N₄O [M+H]⁺ = 217, found 217, R_T = 0.61 min.

7-Oxo-5-(4-(trifluoromethyl)phenyl)-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.81)

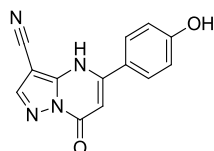


Methyl 3-oxo-3-(4-(trifluoromethyl)phenyl)propanoate (191 mg, 0.78 mmol) was heated with 5-amino-1*H*-pyrazole-4-carbonitrile (60 mg, 0.56 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 18 hours. The reaction mixture was blown down under nitrogen at 45 °C. The crude was purified by reverse phase chromatography (MDAP method A, gradient B) to afford 7-oxo-5-(4-(trifluoromethyl)phenyl)-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as a cream solid, 12 mg, 7%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.43 (s, 1H), 8.07 (d, *J* = 8.3 Hz, 2H), 7.94 (d, *J* = 8.3 Hz, 2H), 6.35 (s, 1H), exchangeable NH not observed, very broad due to

tautomerism; LCMS method A: $C_{14}H_8F_3N_4O$ $[M+H]^+ = 305$, found 305, $R_T = 0.85$ min.

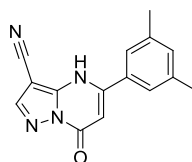
5-(4-Hydroxyphenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.82)



Methyl 3-(4-hydroxyphenyl)-3-oxopropanoate (151 mg, 0.78 mmol) was heated with 5-amino-1*H*-pyrazole-4-carbonitrile (60 mg, 0.56 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 18 h. The reaction mixture was blown down under nitrogen at 45 °C. The crude was purified by reverse phase chromatography (MDAP method A, gradient A) to afford 5-(4-hydroxyphenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as a off white solid with ~ 0.15 mole of acetophenone impurity, 12 mg, 9%.

1H NMR (600 MHz, $DMSO-d_6$) δ 13.29 (br.s, 1H), 10.17 (s, 1H), 8.39 (s, 1H), 7.72 (d, $J = 8.5$, 2H), 6.93 (d, $J = 8.8$, 2H), 6.16 (s, 1H); ^{13}C NMR (151 MHz, $DMSO-d_6$) δ ppm 159.8, 156.0, 153.2, 147.0, 144.4, 128.8, 124.1, 115.0, 113.9, 94.0, signals broadened by tautomerism resulting in one carbon signal missing; LCMS method A: $C_{13}H_9N_4O_2$ $[M+H]^+ = 253$, found 253, $R_T = 0.55$ min; HRMS: m/z calcd for $C_{13}H_9N_4O_2$ $[M+H]^+ = 253.0720$, found 253.0712; IR cm^{-1} 3280, 3084, 2245, 1708; mp not melted at 319 °C.

5-(3,5-Dimethylphenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, ammonia salt (4.91b)



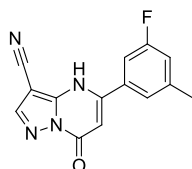
A solution of potassium carbonate (320 mg, 2.31 mmol) in water (1 mL) was added in a single portion to a stirred suspension of 5-chloro-7-oxo-4,7-dihydropyrazolo[1,5-

CONFIDENTIAL – Property of GSK – Do Not Copy

a]pyrimidine-3-carbonitrile (150 mg, 0.77 mmol), (3,5-dimethylphenyl)boronic acid (173 mg, 1.16 mmol), dicyclohexyl(2',6'-dimethoxy-[1,1'-biphenyl]-2-yl)phosphine (S-Phos) (63 mg, 0.15 mmol) and palladium(II) acetate (27 mg, 0.12 mmol) in acetonitrile (1.5 mL). The resultant suspension was purged with nitrogen for 10 mins and then the vial was sealed. The reaction was heated in a microwave at 100 °C for 2 h. Upon cooling to room temperature the suspension was diluted with MeOH (30 mL) and the remaining solid was removed by filtration. The filtrate was blown down at 45 °C under nitrogen to afford crude (412 mg). The crude (100 mg) was purified by reverse phase chromatography (MDAP method B, gradient B) to afford 5-(3,5-dimethylphenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (ammonia salt), as a white solid, 7.2 mg. *Note: Unable to calculate the yield as only a portion of the crude was purified.*

¹H NMR (400 MHz, MeOD) δ 8.19 (s, 1H), 7.49 (s, 2H), 7.15 (s, 1H), 6.25 (s, 1H), 2.39 (s, 6H), exchangeable proton not observed; LCMS method A: C₁₅H₁₁N₄O.NH₃ [M+H]⁺ = 264, found 264, R_T = 0.85 min.

5-(3-Fluoro-5-methylphenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.92)



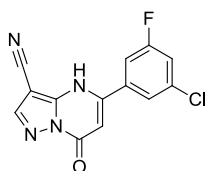
To a solution of 5-chloro-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (120 mg, 0.62 mmol), (3-fluoro-5-methylphenyl)boronic acid (142 mg, 0.93 mmol), palladium(II) acetate (22 mg, 0.096 mmol) and S-Phos (50 mg, 0.12 mmol) in acetonitrile (1.5 mL) was added a solution of potassium carbonate (256 mg, 1.85 mmol) in water (1 mL). The mixture was de-gassed with nitrogen for 10 mins, then heated under microwave irradiation to 100 °C for 3 h. The reaction mixture was dissolved in MeOH (3 mL) and filtered through a hydrophobic frit. The reaction mixture was dissolved in MeOH (~ 50 mL) and filtered through a hydrophobic frit before concentrating *in vacuo*. The crude purified by reverse phase chromatography

CONFIDENTIAL – Property of GSK – Do Not Copy

(0.1% formic acid in water (solvent A), acetonitrile (solvent B)) to afford 5-(3-fluoro-5-methylphenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, 39 mg, 23%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.41 (br.s, 1H), 7.56 – 7.49 (m, 2H), 7.29 (d, *J* = 9.8 Hz, 1H), 6.30 (br.s, 1H), 2.43 (s, 3H), exchangeable proton not observed, very broad; LCMS method A: C₁₃H₁₀ClFN₄O [M+H]⁺ = 269, found 269, R_T = 0.78 min.

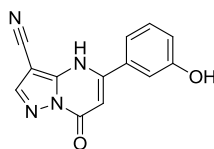
5-(3-Chloro-5-fluorophenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.93)



To a solution of 5-chloro-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (120 mg, 0.62 mmol), (3-chloro-5-fluorophenyl)boronic acid (161 mg, 0.93 mmol), palladium(II) acetate (22 mg, 0.10 mmol) and S-Phos (51 mg, 0.12 mmol) in acetonitrile (1.5 mL) was added a solution of potassium carbonate (256 mg, 1.90 mmol) in water (1 mL). The mixture was de-gassed with nitrogen for 10 mins, then heated under microwave irradiation to 100 °C for 3 h. The reaction mixture was dissolved in MeOH (50 mL) and filtered through a hydrophobic frit. Diatomaceous earth was added and the solvent was removed *in vacuo*. The crude purified by reverse phase chromatography (0.1% formic acid in water (solvent A), acetonitrile (solvent B)) to afford 5-(3-chloro-5-fluorophenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, 15 mg, 8%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.40 (s, 1H), 7.86 (s, 1H), 7.81 – 7.75 (m, 1H), 7.68 (d, *J* = 8.6 Hz, 1H), 6.38 (s, 1H), exchangeable proton not observed; LCMS method A: C₁₃H₇³⁵ClFN₄O [M+H]⁺ = 289, found 289, RT = 0.80 min.

5-(3-Hydroxyphenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.94)

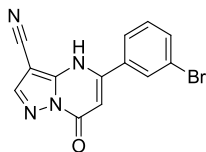


Palladium(II) acetate (27 mg, 0.12 mmol) and dicyclohexyl(2',6'-dimethoxy-[1,1'-biphenyl]-2-yl)phosphine (42 mg, 0.10 mmol) were suspended in acetonitrile (1 mL) and water (0.5 mL). The suspension was purged with nitrogen for 10 mins before sealing and heating under microwave irradiation for 5 mins at 100 °C.

Potassium carbonate (213 mg, 1.54 mmol), 5-chloro-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (100 mg, 0.51 mmol) and (3-hydroxyphenyl)boronic acid (106 mg, 0.77 mmol) were added and the reaction mixture was heated under microwave irradiation for 1.5 h at 100 °C. The reaction mixture was partitioned between DCM (10 mL) and water (5 mL). The organic layer was separated and dried by passing through a hydrophobic frit and blown down under nitrogen at 45 °C. The crude was purified by normal phase chromatography (12 g silica column) eluting with a gradient of 10 – 80% EtOAc in cyclohexane. The appropriate fractions were combined and concentrated *in vacuo* to afford 5-(3-hydroxyphenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as a cream solid, 10 mg, 8%.

¹H NMR (400 MHz, DMSO-*d*₆, 303 K) δ 13.51 (br.s, ~ 0.4H), 9.83 (br.s, ~0.6H), 8.37 (s, 1H), 7.35 (t, *J* = 7.8 Hz, 1H), 7.31 – 7.18 (m, 1H), 7.29 – 7.18 (m, 2H), 6.15 (s, 1H), underintegration of exchangeables due to broadening resulting from tautomerism; ¹H NMR (400 MHz, DMSO-*d*₆, 393 K) 8.22 (br.s, 1H), 7.34 (t, *J* = 7.9 Hz, 1H), 7.38 (s, 2H), 6.99 (d, *J* = 7.9 Hz, 1H), 6.17 (s, 1H (underintegrates)), both exchangeables not observed due to broadening from tautomerism; LCMS method A: C₁₃H₉N₄O₂ [M+H]⁺ = 253, found 253, R_T = 0.56 min.

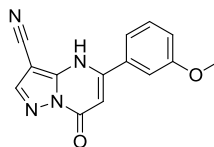
5-(3-Bromophenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.103)



5-Amino-1*H*-pyrazole-4-carbonitrile (60 mg, 0.56 mmol) was heated with ethyl 3-(3-bromophenyl)-3-oxopropanoate (200 mg, 0.78 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 18 h. The reaction mixture was blown down under nitrogen at 45 °C. The crude was purified by reverse phase chromatography (MDAP method A, gradient B) to afford 5-(3-bromophenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as an off white solid, 27 mg, 15%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.36 (s, 1H), 8.09 (s, 1H), 7.87 (d, *J* = 7.8 Hz, 1H), 7.77 (d, *J* = 8.1 Hz, 1H), 7.51 (t, *J* = 7.8 Hz, 1H), 6.28 (s, 1H), exchangeable NH broadened due to tautomerism, contains 20% ammonium salt (unclear of its origin); LCMS method A: C₁₃H₈^{79/81}BrN₄O [M+H]⁺ = 315/317, found 315/317, R_T = 0.79 min.

5-(3-Methoxyphenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.104)

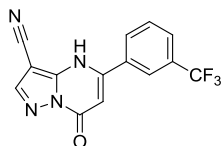


5-Amino-1*H*-pyrazole-4-carbonitrile (60 mg, 0.56 mmol) was heated with ethyl 3-(3-methoxyphenyl)-3-oxopropanoate (173 mg, 0.78 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 18 h. The reaction mixture was blown down under nitrogen at 45 °C. The crude was purified by reverse phase chromatography (MDAP method A, gradient B) to afford 5-(3-methoxyphenyl)-7-oxo-4,7-dihydropyrazolo [1,5-*a*]pyrimidine-3-carbonitrile, as a cream solid, 7.6 mg, 5%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.40 (s, 1H), 7.49 (t, *J* = 7.9 Hz, 1H), 7.42 (s, 1H), 7.39 (s, 1H), 7.17 (dd, *J* = 7.9, 1.8 Hz, 1H), 6.28 (s, 1H), 3.86 (s, 3H), exchangeable

NH not observed, contains 20% ammonium salt (unclear of its origin); LCMS method A: $C_{14}H_{11}N_4O_2$ $[M+H]^+ = 267$, found 267, $R_T = 0.70$ min (93 % by UV).

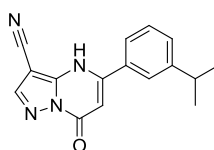
7-Oxo-5-(3-(trifluoromethyl)phenyl)-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.105)



5-Amino-1*H*-pyrazole-4-carbonitrile (200 mg, 1.85 mmol) was heated with ethyl 3-(3-methoxyphenyl)-3-oxopropanoate (674 mg, 2.59 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 20 h. The reaction mixture was blown down under nitrogen at 45 °C. The crude purified by reverse phase chromatography (0.1% formic acid in water (solvent A), acetonitrile (solvent B)), using a Rediseq 50 g C18 gold aqueous column to afford 7-oxo-5-(3-(trifluoromethyl)phenyl)-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as a cream solid, 173 mg, 31%.

1H NMR (400 MHz, $DMSO-d_6$) δ 8.44 (s, 1H), 8.20 (s, 1H), 8.15 (d, $J = 8.1$ Hz, 1H), 7.97 (d, $J = 7.7$ Hz, 1H), 7.81 (t, $J = 7.7$ Hz, 1H), 6.38 (s, 1H), exchangeable NH not observed; LCMS method A: $C_{14}H_8F_3N_4O$ $[M+H]^+ = 305$, found 305, $R_T = 0.83$ min.

5-(3-Isopropylphenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.111)

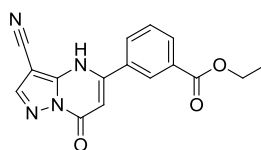


A solution of potassium carbonate (320 mg, 2.31 mmol) in water (1 mL) was added in a single portion to a stirred suspension of 5-chloro-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (150 mg, 0.77 mmol), (3-isopropylphenyl)boronic acid (190 mg, 1.16 mmol), dicyclohexyl(2',6'-dimethoxy-[1,1'-biphenyl]-2-yl)phosphine (*S*-Phos) (63 mg, 0.15 mmol) and palladium(II) acetate (27 mg, 0.12 mmol) in acetonitrile (1.5 mL). The resultant suspension was purged with nitrogen for 10 mins and then the vial was sealed. The reaction was heated in a microwave at 100 °C for 2

h. Upon cooling to room temperature the suspension was diluted with MeOH (30 mL) and the remaining solid was removed by filtration. The filtrate was blown down at 45 °C under nitrogen. The crude purified by reverse phase chromatography (0.1% formic acid in water (solvent A), acetonitrile (solvent B)), using a Rediseq 50 g C18 gold aqueous column. The material was re-purified by reverse phase chromatography (MDAP method A, gradient B) to afford 7-oxo-5-(3-(trifluoromethyl)phenyl)-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as an off white solid, 69 mg, 32%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 13.49 (s, 1H), 8.41 (s, 1H), 7.70 (s, 1H), 7.68 - 7.63 (m, 1H), 7.49 (s, 1H), 7.48 (s, 1H), 6.25 (s, 1H), 3.06 – 2.96 (sept, *J* = 6.9 Hz, 1H), 1.27 (d, *J* = 6.8 Hz, 6H), contains traces of ammonium salt (unclear of its origin); LCMS method A: C₁₆H₁₅N₄O [M+H]⁺ = 279, found 279, R_T = 0.91 min.

Ethyl 3-(3-cyano-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidin-5-yl)benzoate
(4.112)

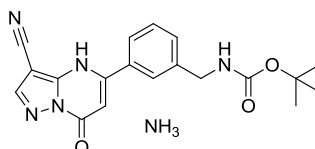


A solution of potassium carbonate (320 mg, 2.31 mmol) in water (1 mL) was added in a single portion to a stirred suspension of 5-chloro-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (150 mg, 0.77 mmol), ethyl 3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzoate (319 mg, 1.16 mmol), dicyclohexyl(2',6'-dimethoxy-[1,1'-biphenyl]-2-yl)phosphine (S-Phos) (63 mg, 0.15 mmol) and palladium(II) acetate (27 mg, 0.12 mmol) in acetonitrile (1.5 mL). The resultant suspension was purged with nitrogen for 10 mins and then the vial was sealed. The reaction was heated in a microwave at 100 °C for 2 h. Upon cooling to room temperature the suspension was diluted with MeOH (30 mL) and the remaining solid was removed by filtration. The filtrate was blown down at 45 °C under nitrogen to afford crude compound (317 mg). A portion of the crude product (100 mg) was purified by reverse phase chromatography (MDAP method A, gradient B) to afford ethyl 3-(3-

cyano-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidin-5-yl)benzoate, as a white solid, 9.4 mg. *Note: Unable to calculate the yield as only a portion of the crude was purified.*

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.42 (s, 1H), 8.38 (s, 1H), 8.15 (d, *J* = 7.6 Hz, 1H), 8.11 (d, *J* = 8.1 Hz, 1H), 7.72 (t, *J* = 7.8 Hz, 1H), 6.30 (s, 1H), 4.39 (q, *J* = 7.1 Hz, 2H), 1.36 (t, *J* = 7.1 Hz, 3H), exchangeable NH not observed; LCMS method A: C₁₆H₁₃N₄O₃ [M+H]⁺ = 309, found 309, R_T = 0.78 min.

***tert*-Butyl 3-(3-cyano-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidin-5-yl)benzylcarbamate ammonium salt (4.113b)**



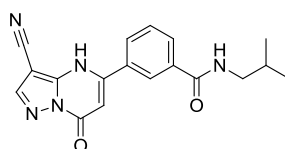
A solution of potassium carbonate (320 mg, 2.31 mmol) in water (1 mL) was added in a single portion to a stirred suspension of 5-chloro-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (150 mg, 0.77 mmol), *tert*-butyl 3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzylcarbamate (385 mg, 1.16 mmol), dicyclohexyl(2',6'-dimethoxy-[1,1'-biphenyl]-2-yl)phosphine (S-Phos) (63 mg, 0.15 mmol) and palladium(II) acetate (27 mg, 0.12 mmol) in acetonitrile (1.5 mL). The resultant suspension was purged with nitrogen for 10 mins and then the vial was sealed. The reaction was heated in a microwave at 100 °C for 2 h. Upon cooling to room temperature the suspension was diluted with MeOH (30 mL) and the remaining solid was removed by filtration. The filtrate was blown down at 45 °C under nitrogen. The crude purified by reverse phase chromatography (0.1% formic acid in water (solvent A), acetonitrile (solvent B)), using a Redisep 50 g C18 gold aqueous column. The material was re-purified by reverse phase chromatography (MDAP method A, gradient B) and then by reverse phase chromatography (10 mM ammonium bicarbonate in water (solvent A), acetonitrile (solvent B)), using a Redisep 30 g C18

CONFIDENTIAL – Property of GSK – Do Not Copy

gold aqueous column to afford *tert*-Butyl 3-(3-cyano-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidin-5-yl)benzylcarbamate ammonium salt, as a white solid, 94 mg, 32%.

¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 8.15 (s, 1H), 7.84 (s, 1H), 7.80 (d, *J* = 7.6 Hz, 1H), 7.49 - 7.41 (m, 1H), 7.41 - 7.37 (m, 1H), 7.30 (d, *J* = 7.3 Hz, 1H), 7.09 (t, *J* = 50.1 Hz, 4H, ammonium) 6.12 (s, 1H), 4.20 (d, *J* = 6.1 Hz, 2H), 1.40 (s, 9H); LCMS method A: C₁₉H₂₀N₅O₃.NH₃ [M+H]⁺ = 366, found 366, R_T = 0.76 min.

3-(3-Cyano-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidin-5-yl)-*N*-isobutylbenzamide (4.114)

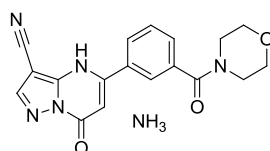


A solution of potassium carbonate (320 mg, 2.31 mmol) in water (1 mL) was added in a single portion to a stirred suspension of 5-chloro-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (150 mg, 0.77 mmol), (3-(isobutylcarbamoyl)phenyl)boronic acid (256 mg, 1.16 mmol), dicyclohexyl(2',6'-dimethoxy-[1,1'-biphenyl]-2-yl)phosphine (S-Phos) (63 mg, 0.15 mmol) and palladium(II) acetate (27 mg, 0.12 mmol) in acetonitrile (1.5 mL). The resultant suspension was purged with nitrogen for 10 mins and then the vial was sealed. The reaction was heated in a microwave at 100 °C for 2 h. Upon cooling to room temperature the suspension was diluted with MeOH (30 mL) and the remaining solid was removed by filtration. The filtrate was blown down at 45 °C under nitrogen. The crude was purified by reverse phase chromatography (0.1% formic acid in water (solvent A), acetonitrile (solvent B)), using a Redisep 50 g C18 gold aqueous column to afford 3-(3-cyano-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidin-5-yl)-*N*-isobutylbenzamide, as an off white solid, 122 mg, 47%.

¹H NMR (400 MHz, DMSO-*d*₆) ppm 14.21-13.33 (br.s, 1H (underintegrates)), 8.65 (br.t, *J* = 5.7 Hz, 1H), 8.43 (s, 1H), 8.29 (s, 1H), 8.05 (d, *J* = 7.8 Hz, 1H), 8.00 (d, *J* =

8.1 Hz, 1H), 7.66 (t, $J = 7.8$ Hz, 1H), 6.37 (s, 1H), 3.17 - 3.11 (m, 2H), 1.93 - 1.81 (m, 1H), 0.92 (d, $J = 6.7$ Hz, 6H); LCMS method A: $C_{18}H_{18}N_5O_2$ $[M+H]^+ = 336$, found 336, $R_T = 0.74$ min.

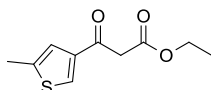
5-(3-(Morpholine-4-carbonyl)phenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, ammonium salt (4.115b)



A solution of potassium carbonate (320 mg, 2.31 mmol) in water (1 mL) was added in a single portion to a stirred suspension of 5-chloro-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (150 mg, 0.77 mmol), (3-(morpholine-4-carbonyl)phenyl)boronic acid (272 mg, 1.16 mmol), dicyclohexyl(2',6'-dimethoxy-[1,1'-biphenyl]-2-yl)phosphine (*S*-Phos) (63 mg, 0.15 mmol) and palladium(II) acetate (27 mg, 0.12 mmol) in acetonitrile (1.5 mL). The resultant suspension was purged with nitrogen for 10 mins and then the vial was sealed. The reaction was heated in a microwave at 100 °C for 2 h. Upon cooling to room temperature the suspension was diluted with MeOH (30 mL) and the remaining solid was removed by filtration. The filtrate was blown down at 45 °C under nitrogen to afford crude compound (509 mg). A portion of the crude (100 mg) was purified by reverse phase chromatography (MDAP method A, gradient A) and then by reverse phase chromatography (MDAP method B, gradient A) to afford 5-(3-(morpholine-4-carbonyl)phenyl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, ammonium salt, as a white solid, 8.7 mg. *Note: Unable to calculate the yield as only a portion of the crude was purified.*

1H NMR (400 MHz, MeOD) δ 8.16 (s, 1H), 8.11 - 8.05 (m, 2H), 7.58 (t, $J = 7.8$ Hz, 1H), 7.51 (apparent dt, $J = 7.6, 1.3$ Hz, 1H), 6.35 (s, 1H), 3.88 - 3.47 (br.m., 8H), NH not observed; LCMS method A: $C_{18}H_{16}N_5O_3.NH_3$ $[M+H]^+ = 350$ found 350, $R_T = 0.56$ min.

Ethyl 3-(5-methylthiophen-3-yl)-3-oxopropanoate (4.125)



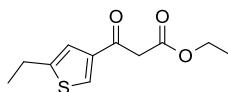
CDI (242 mg, 1.49 mmol) was added to a solution of 5-methylthiophene-3-carboxylic acid (200 mg, 1.41 mmol) in THF (5 mL) at RT under nitrogen. The solution was stirred at RT for 5 h.

Meanwhile, magnesium ethanolate (177 mg, 1.55 mmol) was added to a solution of 3-ethoxy-3-oxopropanoic acid (0.37 mL, 3.09 mmol) in THF (5 mL) at RT under nitrogen. The solution was stirred for 4.5 h. The reaction mixture was concentrated *in vacuo* to give a white solid. The white solid was suspended in THF (3 mL) and was added dropwise to the activated acid solution. The combined solution was stirred for 3 h.

Saturated aqueous sodium bicarbonate solution (10 mL) and ethyl acetate (10 mL) were added. The organic layer was separated and the aqueous portion was washed with further ethyl acetate (3 x 20 mL). The combined organic fractions were passed through a hydrophobic frit and concentrated *in vacuo* to afford a clear oil (350 mg). The product was taken crude into the next reaction.

LCMS method A: C₁₀H₁₃O₃S [M+H]⁺ = 213, found 213, R_T = 0.94 min.

Ethyl 3-(5-ethylthiophen-3-yl)-3-oxopropanoate (4.126)



CDI (220 mg, 1.36 mmol) was added to a solution of 5-ethylthiophene-3-carboxylic acid (200 mg, 1.28 mmol) in THF (5 mL) at RT under nitrogen. The solution was stirred at RT under nitrogen for 5 h.

Meanwhile, magnesium ethanolate (161 mg, 1.41 mmol) was added to a solution of 3-ethoxy-3-oxopropanoic acid (0.33 mL, 2.82 mmol) in THF (5 mL) at RT under

CONFIDENTIAL – Property of GSK – Do Not Copy

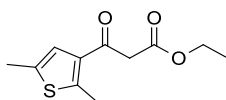
nitrogen. The solution was stirred for 4.5 h. The reaction mixture was concentrated *in vacuo* to give a white solid. The white solid was suspended in THF (3 mL) and was added dropwise to the activated acid solution. The combined solution was stirred for 3 h.

Saturated aqueous sodium bicarbonate solution (10 mL) and ethyl acetate (10 mL) were added. The organic layer was separated and the aqueous portion was washed with further ethyl acetate (3 x 20 mL). The combined organic fractions were passed through a hydrophobic frit and concentrated *in vacuo* to afford a dark yellow oil.

The crude was purified by normal phase chromatography (24 g silica column) eluting with a gradient of 0 – 20% EtOAc in cyclohexane. The appropriate fractions were combined and concentrated *in vacuo* to afford ethyl 3-(5-ethylthiophen-3-yl)-3-oxopropanoate (137 mg), as a yellow oil containing impurities. Used without further purification.

LCMS method A: C₁₁H₁₅O₃S [M+H]⁺ = 227, found 227, R_T = 1.02 min.

Ethyl 3-(2,5-dimethylthiophen-3-yl)-3-oxopropanoate (4.127)



CDI (220 mg, 1.36 mmol) was added to a solution of 2,5-dimethylthiophene-3-carboxylic acid (200 mg, 1.28 mmol) in THF (5 mL) at RT under nitrogen. The solution was stirred at RT under nitrogen for 5 h.

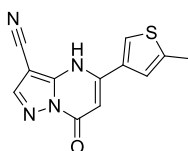
Meanwhile, magnesium ethanolate (161 mg, 1.41 mmol) was added to a solution of 3-ethoxy-3-oxopropanoic acid (0.33 mL, 2.82 mmol) in THF (5 mL) at RT under nitrogen. The solution was stirred for 4.5 h. The reaction mixture was concentrated *in vacuo* to give a white solid. The white solid was suspended in THF (3 mL) and was added dropwise to the activated acid solution. The combined solution was stirred for 19 h.

CONFIDENTIAL – Property of GSK – Do Not Copy

Saturated aqueous sodium bicarbonate solution (10 mL) and ethyl acetate (10 mL) were added. The organic layer was separated and the aqueous portion was washed with further ethyl acetate (3 x 20 mL). The combined organic fractions were passed through a hydrophobic frit and concentrated *in vacuo* to afford a dark brown oil (196 mg). The product was taken crude into the next reaction.

LCMS method A: C₁₁H₁₅O₃S [M+H]⁺ = 227, found 227, R_T = 1.05 min.

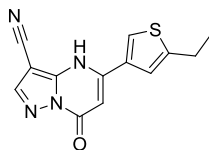
5-(5-Methylthiophen-3-yl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.116)



5-Amino-1*H*-pyrazole-4-carbonitrile (153 mg, 1.41 mmol) was heated with ethyl 3-(5-methylthiophen-3-yl)-3-oxopropanoate (300 mg, 1.41 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 17 h. The reaction mixture was concentrated *in vacuo*. The crude purified by reverse phase chromatography (0.1% formic acid in water (solvent A), acetonitrile (solvent B)), using a Rediseq 50 g C18 gold aqueous column to afford 5-(5-methylthiophen-3-yl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as a cream solid, 23 mg, 6.4 %.

¹H NMR (400 MHz, DMSO-*d*₆) δ 13.20 (s, 1H), 8.40 (s, 1H), 8.19 (s, 1H), 7.44 (s, 1H), 6.32 (s, 1H), 2.50 - 2.48 (m, assumed 3H, under DMSO solvent peak); LCMS method A: C₁₂H₉N₄OS [M+H]⁺ = 257, found 257, R_T = 0.71 min.

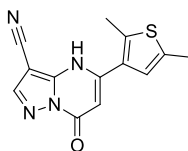
5-(5-Ethylthiophen-3-yl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.117)



5-Amino-1*H*-pyrazole-4-carbonitrile (60 mg, 0.56 mmol) was heated with ethyl 3-(5-ethylthiophen-3-yl)-3-oxopropanoate (137 mg, 0.60 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 42 h. The reaction mixture was concentrated *in vacuo*. The crude purified by reverse phase chromatography (0.1% formic acid in water (solvent A), acetonitrile (solvent B)), using a Rediseq 50 g C18 gold aqueous column to afford 5-(5-methylthiophen-3-yl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as a cream solid, 13 mg, 8.8%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 13.19 (s, 1H), 8.40 (s, 1H), 8.22 (d, *J* = 1.2 Hz, 1H), 7.48 (d, *J* = 1.2 Hz, 1H), 6.35 (s, 1H), 2.87 (q, *J* = 7.5 Hz, 2H), 1.30 (t, *J* = 7.5 Hz, 3H); LCMS method A: C₁₃H₁₁N₄OS [M+H]⁺ = 271, found 271, R_T = 0.82 min.

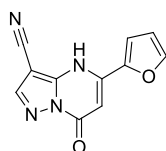
5-(2,5-Dimethylthiophen-3-yl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.118)



5-Amino-1*H*-pyrazole-4-carbonitrile (90 mg, 0.83 mmol) was heated with ethyl 3-(2,5-dimethylthiophen-3-yl)-3-oxopropanoate (196 mg, 0.87 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 66 h. The reaction mixture was concentrated by blow down at 45 °C. The crude was purified by reverse phase chromatography (MDAP method A, gradient B) to afford 5-(2,5-dimethylthiophen-3-yl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as a white solid, 30 mg, 13%.

^1H NMR (400 MHz, MeOD) δ 8.28 (s, 1H), 6.92 (d, J = 1.0 Hz, 1H), 6.01 (s, 1H), 2.56 (s, 3H), 2.49 (s, 3H), exchangeable NH not observed; LCMS method A: $\text{C}_{13}\text{H}_{11}\text{N}_4\text{OS}$ $[\text{M}+\text{H}]^+ = 271$, found 271, $R_T = 0.78$ min.

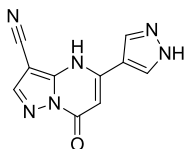
5-(Furan-2-yl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.120)



5-Amino-1*H*-pyrazole-4-carbonitrile (100 mg, 0.93 mmol) was heated with commercially available ethyl 3-(furan-2-yl)-3-oxopropanoate (196 mg, 1.08 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 4 days. The reaction mixture was concentrated *in vacuo*. The crude was purified by reverse phase chromatography (MDAP method A, gradient A) to afford 5-(furan-2-yl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as an off white solid, 17 mg, 8%.

^1H NMR (400 MHz, DMSO- d_6) δ 14.21 – 12.89 (br.s, assumed 1H (underintegrates)), 8.39 (s, 1H), 8.06 (d, J = 1.2 Hz, 1H), 7.57 (d, J = 3.4 Hz, 1H), 6.80 (dd, J = 3.7, 1.7 Hz, 1H), 6.29 (s, 1H); LCMS method A: $\text{C}_{11}\text{H}_7\text{N}_4\text{O}_2$ $[\text{M}+\text{H}]^+ = 227$, found 227, $R_T = 0.52$ min.

7-Oxo-5-(1*H*-pyrazol-4-yl)-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.119)

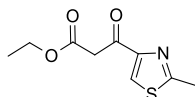


A solution of potassium carbonate (320 mg, 2.31 mmol) in water (1 mL) was added in a single portion to a stirred suspension of 5-chloro-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (150 mg, 0.77 mmol), dicyclohexyl(2',6'-dimethoxy-[1,1'-biphenyl]-2-yl)phosphine (S-Phos) (63 mg, 0.15 mmol) and palladium(II) acetate (27 mg, 0.12 mmol) in acetonitrile (1.5 mL). The resultant suspension was

purged with nitrogen for 10 mins and then the vial was sealed. The reaction was heated in a microwave at 100 °C for 1.5 h. Upon cooling to room temperature the suspension was diluted with MeOH (30 mL) and the remaining solid was removed by filtration. The filtrate was blown down at 45 °C under nitrogen. The crude purified by reverse phase chromatography (0.1% formic acid in water (solvent A), MeOH (solvent B)), using a Redisep 50 g C18 gold aqueous column to afford 7-oxo-5-(1*H*-pyrazol-4-yl)-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, 1.7 mg, 1%.

¹H NMR (400 MHz, MeOD) δ 8.21 (s, 1H), 8.16 (s, 2H), 6.25 (s, 1H), two exchangeable NH not observed; LCMS method A: C₁₀H₇N₆O [M+H]⁺ = 227, found 227, R_T = 0.43 min.

Ethyl 3-(2-methylthiazol-4-yl)-3-oxopropanoate (4.134)



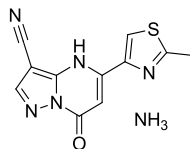
2-Methylthiazole-4-carboxylic acid (215 mg, 1.51 mmol) was dissolved in anhydrous THF (5 mL). CDI (259 mg, 1.60 mmol) was added to the reaction mixture and the solution was stirred under nitrogen for 5 h.

Meanwhile, 3-ethoxy-3-oxopropanoic acid (437 mg, 3.31 mmol) was dissolved in THF (5 mL). Magnesium ethanolate (189 mg, 1.66 mmol) was added in a single portion. The reaction mixture was stirred at room temperature under nitrogen for 4.5 h. The reaction mixture was concentrated *in vacuo* and redissolved in anhydrous THF (25 mL). The solution was added dropwise to of the activated acid. The reaction mixtures were left stirring at room temperature under nitrogen for 19 h. Saturated aqueous sodium bicarbonate solution (20 mL) was added to the reaction mixture, which was extracted with EtOAc (3 x 20 mL). The organic was combined, dried by passing through a hydrophobic frit and concentrated *in vacuo* to afford ethyl 3-(2-methylthiazol-4-yl)-3-oxopropanoate (248 mg). The product was taken crude into the next reaction.

CONFIDENTIAL – Property of GSK – Do Not Copy

LCMS method A: C₉H₁₂NO₃S [M+H]⁺ = 214, found 214, R_T = 0.73 min.

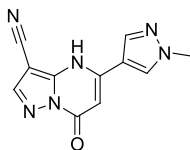
5-(2-Methylthiazol-4-yl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, ammonia salt (4.132b)



5-Amino-1*H*-pyrazole-4-carbonitrile (100 mg, 0.93 mmol) was heated with ethyl 3-(2-methylthiazol-4-yl)-3-oxopropanoate (248 mg, 1.16 mmol) in acetic acid (2 mL) and water (1 mL) at 100 °C for 72 h. The reaction mixture was concentrated by blow down at 45 °C. The crude was purified by reverse phase chromatography (MDAP method B, gradient A) and repurified by reverse phase chromatography (MDAP method B, gradient A, extended run) to afford 5-(2-methylthiazol-4-yl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile ammonia salt, as an off white solid, 8.2 mg, 3%.

¹H NMR (600 MHz, DMSO-*d*₆) δ 8.63 (s, 1H), 8.43 (s, 1H), 6.57(s, 1H), 2.78 (s, 3H), spectra run in the presence of a drop of deuterated HCl to speed tautomerism, exchangeable not observed; ¹³C NMR (151 MHz, DMSO-*d*₆) δ ppm 167.3, 155.4, 145.6, 145.0, 144.2, 144.1, 123.1, 112.8, 96.2, 75.8, 18.8, spectra run in the presence of a drop of deuterated HCl to speed tautomerism; LCMS method B: C₁₁H₈N₅OS.H₃N [M+H]⁺ = 258, found 258, R_T = 0.54 min; HRMS: m/z calcd for C₁₁H₈N₅OS [M+H]⁺ = 258.0444, found 258.0439; mp - very gradual decomposition observed.

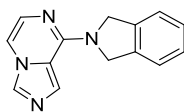
5-(1-Methyl-1*H*-pyrazol-4-yl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (4.130)



A solution of potassium carbonate (266 mg, 1.93 mmol) in water (1 mL) was added in a single portion to a stirred suspension of 5-chloro-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile (125 mg, 0.64 mmol), 1-methyl-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-1*H*-pyrazole (200 mg, 0.96 mmol), dicyclohexyl(2',6'-dimethoxy-[1,1'-biphenyl]-2-yl)phosphine (S-Phos) (53 mg, 0.13 mmol) and palladium(II) acetate (27 mg, 0.12 mmol) in acetonitrile (1.5 mL). The resultant suspension was purged with nitrogen for 10 mins and then the vial was sealed. The reaction was heated in a microwave at 100 °C for 1.5 h. Upon cooling to room temperature the suspension was diluted with MeOH (30 mL) and the remaining solid was removed by filtration. The filtrate was blown down at 45 °C under nitrogen. The crude purified by normal phase chromatography (isopropyl acetate (solvent A), 30% MeOH in isopropyl acetate (solvent B)), using a silica 12 g column. The sample was re-purified by reverse phase chromatography (MDAP method A, gradient A) to afford 5-(1-methyl-1*H*-pyrazol-4-yl)-7-oxo-4,7-dihydropyrazolo[1,5-*a*]pyrimidine-3-carbonitrile, as a white solid, 11 mg, 7%.

¹H NMR (600 MHz, DMSO-*d*₆) δ 13.37-12.52 (m, 1H), 8.38 (br.s, 1H), 8.21 (br.s, 1H), 8.08 (br.s, 1H), 6.12 (br.s, 1H), 3.91 (s, 3H), broad signals due to tautomerism, exchangeable proton observed but split between two positions (13.09 and 12.71), trace fomic acid present; ¹³C NMR (151 MHz, DMSO-*D*₆) δ ppm 156.4, 144.4, 138.0, 130.8 114.6, 92.8, 75.1, 39.3, signals broadened by tautomerism resulting in some carbon signals missing from 1D ¹³C spectra; LCMS method A: C₁₁H₉N₆O [M+H]⁺ = 241 found 241, R_T = 0.47 min; HRMS: m/z calcd for C₁₁H₉N₆O [M+H]⁺ = 241.0832, found 241.0824; insufficient material for mp and IR.

8-(Isoindolin-2-yl)imidazo[1,5-*a*]pyrazine (4.179)

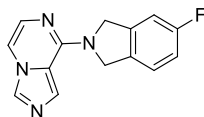


8-Chloroimidazo[1,5-*a*]pyrazine (50 mg, 0.33 mmol) was heated with isoindoline (39 mg, 0.33 mmol) and K₂CO₃ (180 mg, 1.30 mmol) in anhydrous DMSO (0.5 mL) at 60 °C for 5 h. The reaction mixture was diluted with MeOH (5 mL). The milky suspension was pipetted off (leaving crystalline solid) and the suspension was blown down under nitrogen at 45 °C.

The crude was pre-absorbed onto diatomaceous earth and was purified by reverse phase chromatography 10 mM ammonium bicarbonate in aqueous solution (solvent A), acetonitrile (solvent B)), using a Redisep 50 g C18 gold aqueous column to afford 8-(isoindolin-2-yl)imidazo[1,5-*a*]pyrazine, as a pale brown solid, 47 mg, 61%.

¹H NMR (400 MHz, MeOD) δ 8.33 (s, 1H), 8.02 (s, 1H), 7.57 (d, *J* = 4.9 Hz, 1H), 7.43 (dd, *J* = 5.4, 3.2 Hz, 2H), 7.37 - 7.32 (m, 2H), 7.08 (d, *J* = 4.9 Hz, 1H), 5.17 (s, 4H), weak NMR, very poor solubility; ¹³C NMR (101 MHz, DMSO-*d*₆) δ ppm 149.4, 130.6, 127.9, 127.3, 124.4, 122.8, 121.7, 118.8, 106.7, 54.0; LCMS method A: C₁₄H₁₃N₄ [M+H]⁺ = 237, found 237, R_T = 0.84 min; HRMS: *m/z* calcd for C₁₄H₁₃N₄ [M+H]⁺ = 237.1135, found 237.1128; mp 250 – 252 °C; IR cm⁻¹ 2920 (broad and complex), 2578 (broad).

8-(5-Fluoroisoindolin-2-yl)imidazo[1,5-*a*]pyrazine (4.180)



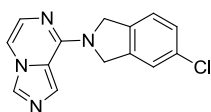
8-Chloroimidazo[1,5-*a*]pyrazine (50 mg, 0.33 mmol) was heated with 5-fluoroisoindoline hydrochloride (74 mg, 0.42 mmol) and K₂CO₃ (117 mg, 0.85 mmol) in anhydrous DMSO (0.4 mL) at 60 °C overnight in a sealed vial. The resultant solid was attempted to be dissolved in DMSO. The insoluble material was collected by vacuum filtration and was washed with water before drying in a

CONFIDENTIAL – Property of GSK – Do Not Copy

desiccator under vacuum to afford 8-(5-fluoroisindolin-2-yl)imidazo[1,5-*a*]pyrazine, 46 mg, 55%.

^1H NMR (400 MHz, DMSO- d_6) δ 8.40 (s, 1H), 7.93 (s, 1H), 7.67 (d, $J = 4.8$ Hz, 1H), 7.47 (dd, $J = 8.3, 5.3$ Hz, 1H), 7.28 (dd, $J = 9.1, 2.3$ Hz, 1H), 7.18 (t.d, $J = 9.0, 2.4$ Hz, 1H), 7.09 (d, $J = 4.8$ Hz, 1H), 5.10 (s, 4H); LCMS method B: $\text{C}_{14}\text{H}_{12}\text{FN}_4$ $[\text{M}+\text{H}]^+ = 255$, found 255, $R_T = 0.85$ min.

8-(5-Chloroisindolin-2-yl)imidazo[1,5-*a*]pyrazine (4.181)

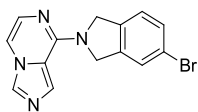


8-Chloroimidazo[1,5-*a*]pyrazine (50 mg, 0.33 mmol) was heated with 5-chloroisindoline (50 mg, 0.33 mmol) and K_2CO_3 (225 mg, 1.63 mmol) in anhydrous DMSO (0.5 mL) at 60 °C for 17 h. The reaction mixture was cooled and was diluted with MeOH (3 mL). The milky suspension was pipetted off (leaving crystalline solid) and was blown down under nitrogen at 45 °C.

The crude was pre-absorbed onto diatomaceous earth and was purified by reverse phase chromatography 10 mM ammonium bicarbonate in aqueous solution (solvent A), acetonitrile (solvent B)), using a Rediseq 50 g C18 gold aqueous column to afford 8-(5-chloroisindolin-2-yl)imidazo[1,5-*a*]pyrazine, as a white solid, 15 mg, 17%.

^1H NMR (400 MHz, DMSO- d_6) δ 8.41 (s, 1H), 7.93 (s, 1H), 7.68 (d, $J = 4.6$ Hz, 1H), 7.53 (s, 1H), 7.49 - 7.45 (m, 1H), 7.43 - 7.39 (m, 1H), 7.09 (d, $J = 4.6$ Hz, 1H), 5.10 (br.s, 4H), poor solubility; LCMS method A: $\text{C}_{14}\text{H}_{12}^{35}\text{ClN}_4$ $[\text{M}+\text{H}]^+ = 271$, found 271, $R_T = 0.96$ min.

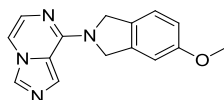
8-(5-Bromoisindolin-2-yl)imidazo[1,5-*a*]pyrazine (4.174)



8-Chloroimidazo[1,5-*a*]pyrazine (50 mg, 0.33 mmol) was heated with 5-bromoisindoline hydrochloride (122 mg, 0.52 mmol) and K₂CO₃ (117 mg, 0.85 mmol) in DMSO (0.4 mL) at 60 °C overnight in a sealed vial during which a precipitate was formed. Additional DMSO was added and the reaction mixture was stirred for a further 2 h. The remaining solid was collected by filtration and was washed with water before drying in a desiccator overnight to afford 8-(5-bromoisindolin-2-yl)imidazo[1,5-*a*]pyrazine, 82 mg, 80%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.41 (s, 1H), 7.93 (s, 1H), 7.69-7.65 (m, 2H), 7.54 (dd, *J* = 8.1, 1.8 Hz, 1H), 7.41 (d, *J* = 8.1 Hz, 1H), 7.09 (d, *J* = 4.8 Hz, 1H), 5.09 (br.s, 4H); LCMS method B: C₁₄H₁₂^{79/81}BrN₄ [M+H]⁺ = 315/317, found 315/317, R_T = 0.98 min.

8-(5-Methoxyisindolin-2-yl)imidazo[1,5-*a*]pyrazine (4.182)



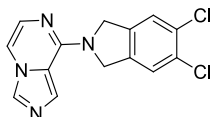
8-Chloroimidazo[1,5-*a*]pyrazine (50 mg, 0.33 mmol) was heated with 5-methoxyisindoline hydrochloride (60 mg, 0.33 mmol) and K₂CO₃ (135 mg, 0.98 mmol) in anhydrous DMSO (0.4 mL) at 60 °C for 18 h. The reaction mixture was cooled and was diluted with MeOH (5 mL). The solution was filtered to remove remaining solid and the organic was blown down under nitrogen at 45 °C.

The sample was purified by reverse phase chromatography (MDAP method B, gradient B) to afford 8-(5-methoxyisindolin-2-yl)imidazo[1,5-*a*]pyrazine, as a light grey solid, 15 mg, 17%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.39 (s, 1H), 7.93 (s, 1H), 7.65 (d, *J* = 4.9 Hz, 1H), 7.34 (d, *J* = 8.5 Hz, 1H), 7.08 (d, *J* = 4.9 Hz, 1H), 7.03 (d, *J* = 2.1 Hz, 1H), 6.91 (dd,

$J = 8.4, 2.3$ Hz, 1H), 5.05 (s, 4H), 3.78 (s, 3H); LCMS method B: $C_{15}H_{15}N_4O$ $[M+H]^+ = 267$, found 267, $R_T = 0.86$ min.

8-(5,6-Dichloroisindolin-2-yl)imidazo[1,5-*a*]pyrazine (4.184)

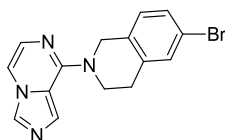


8-Chloroimidazo[1,5-*a*]pyrazine (50 mg, 0.33 mmol) was heated with 5,6-dichloroisindoline hydrochloride (73 mg, 0.33 mmol) and K_2CO_3 (135 mg, 0.98 mmol) in DMSO (0.4 mL) at 60 °C for 18 h. The reaction mixture was cooled and was diluted with MeOH (5 mL). The solution was pipetted off (leaving crystalline solid) and was blown down under nitrogen at 45 °C.

The crude was pre-absorbed onto diatomaceous earth and was purified by reverse phase chromatography (0.1% formic acid in water (solvent A), acetonitrile (solvent B)), using a Rediseq 30 g C18 gold aqueous column to afford 8-(5,6-dichloroisindolin-2-yl)imidazo[1,5-*a*]pyrazine, as a dark grey/brown solid, 4.3 mg, 4.3%.

1H NMR (400 MHz, $DMSO-d_6$) δ 8.42 (s, 1H), 7.89 (s, 1H), 7.73 (s, 2H), 7.69 (d, $J = 4.9$ Hz, 1H), 7.10 (d, $J = 4.9$ Hz, 1H), 5.11 (s, 4H), weak NMR due to poor solubility; LCMS method A: $C_{14}H_{11}^{35}Cl_2N_4$ $[M+H]^+ = 305/307$, found 305/307, $R_T = 0.69$ min.

6-Bromo-2-(imidazo[1,5-*a*]pyrazin-8-yl)-1,2,3,4-tetrahydroisoquinoline (4.175)



8-Chloroimidazo[1,5-*a*]pyrazine (50 mg, 0.33 mmol) was heated with 3-bromo-5,6,7,8-tetrahydro-1,6-naphthyridine hydrochloride (90 mg, 0.42 mmol) and potassium carbonate (72 mg, 0.52 mmol) in anhydrous DMSO (0.4 mL) at 60 °C overnight in a sealed vial. Water (a few drops) was added and the reaction mixture

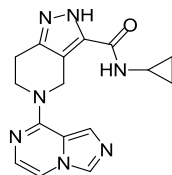
CONFIDENTIAL – Property of GSK – Do Not Copy

was stirred before partitioning between EtOAc and water. The organic was dried through a hydrophobic frit and was concentrated *in vacuo*.

The sample was purified by reverse phase chromatography (MDAP method B) to afford 6-bromo-2-(imidazo[1,5-a]pyrazin-8-yl)-1,2,3,4-tetrahydroisoquinoline, as a solid, 51 mg, 47%, contains minor impurities.

^1H NMR (400 MHz, DMSO- d_6) δ 8.41 (s, 1H), 7.94 (s, 1H), 7.73 (dd, $J = 4.8, 1.0$ Hz, 1H), 7.59 (d, $J = 1.9$ Hz, 1H), 7.39 (dd, $J = 8.2, 2.0$ Hz, 1H), 7.19 (d, $J = 8.3$ Hz, 1H), 7.10 (d, $J = 4.8$ Hz, 1H), 4.95 (s, 2H), 4.03 (t, $J = 5.9$ Hz, 2H), 2.94 (t, $J = 5.9$ Hz, 2H) contains minor impurities; LCMS method B: $\text{C}_{15}\text{H}_{14}^{79/81}\text{BrN}_4$ $[\text{M}+\text{H}]^+ = 329/331$, found 329/331, $R_T = 1.06$ min.

***N*-Cyclopropyl-5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide (4.192)**



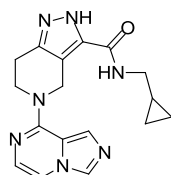
A solution of cyclopropanamine (30 μ L, 0.42 mmol), 5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxylic acid (60 mg, 0.21 mmol), triethylamine (88 μ L, 0.63 mmol) and T3P (50% in EtOAc) (188 μ L, 0.32 mmol) in 2-MeTHF (1 mL) was stirred at room temperature for 20 h. Additional T3P (50% in EtOAc) (188 μ L, 0.32 mmol), triethylamine (88 μ L, 0.63 mmol) and cyclopropanamine (30 μ L, 0.42 mmol) were added and the reaction mixture was stirred for a further 24 h at room temperature before heating at 40 $^{\circ}$ C for 4 h. The reaction mixture was concentrated *in vacuo* and was partitioned between saturated aqueous sodium bicarbonate (5 mL) and EtOAc (5 mL). The aqueous was washed with EtOAc (3 x 5 mL) and DCM (2 x 5 mL). The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*.

The sample was purified by reverse phase chromatography (MDAP method B, gradient A) to afford. *N*-cyclopropyl-5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide, as a white solid, 5.0 mg, 7.3%.

^1H NMR (600 MHz, DMSO- d_6) δ 12.98 (br.s, 1H), 8.42 (s, 1H), 8.06 (d, J = 4.4 Hz, 1H), 7.82 (s, 1H), 7.76 (d, J = 4.8 Hz, 1H), 7.12 (d, J = 4.8 Hz, 1H), 4.95 (s, 2H), 4.05 (t, J = 5.7 Hz, 2H), 2.91 (t, J = 5.5 Hz, 2H), 2.85 - 2.68 (m, 1H), 0.76 - 0.41 (m, 4H); ^{13}C NMR (151 MHz, DMSO- d_6) δ ppm 151.4, 138.7, 130.9, 126.8, 123.6, 118.6, 114.0, 108.2, 43.7, 43.5, 22.2, 21.4, 5.6, signals broadened by tautomerism resulting in some carbon signals missing from 1D ^{13}C spectra; LCMS method B: $\text{C}_{16}\text{H}_{18}\text{N}_7\text{O}$ $[\text{M}+\text{H}]^+ = 324$, found 324, $R_T = 0.63$ min; HRMS: m/z calcd for

$C_{16}H_{18}N_7O$ $[M+H]^+ = 324.1567$, found 324.1566, insufficient material for melting point determination.

***N*-(Cyclopropylmethyl)-5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide (4.193)**



A solution of cyclopropylmethanamine (37 μ L, 0.42 mmol), 5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxylic acid (60 mg, 0.21mmol), triethylamine (88 μ L, 0.63 mmol) and T3P (50% in EtOAc) (188 μ L, 0.32 mmol) in 2-MeTHF (1 mL) was stirred at room temperature for 20 h. Additional T3P (50% in EtOAc) (188 μ L, 0.32 mmol), triethylamine (88 μ L, 0.63 mmol) and cyclopropylmethanamine (37 μ L, 0.42 mmol) were added and the reaction mixture was stirred for a further 24 h. The reaction mixture was concentrated *in vacuo* and was partitioned between saturated aqueous sodium bicarbonate (5 mL) and EtOAc (5 mL). The aqueous was washed with EtOAc (3 x 5 mL). The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*.

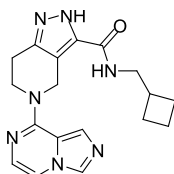
The sample was purified by reverse phase chromatography (MDAP method B, gradient A) to afford *N*-(cyclopropylmethyl)-5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide, as a white solid, 9.0 mg, 13%.

1H NMR (600 MHz, DMSO- d_6) δ 12.99 (br.s, 1H), 8.42 (s, 1H), 8.06 (t, $J = 6.1$ Hz, 1H), 7.81 (s, 1H), 7.76 (d, $J = 4.8$ Hz, 1H), 7.12 (d, $J = 4.8$ Hz, 1H), 4.95 (s, 2H), 4.05 (t, $J = 5.7$ Hz, 2H), 3.11 (t, $J = 6.4$ Hz, 2H), 2.92 (t, $J = 5.7$ Hz, 2H), 1.13 – 0.96 (m, 1H), 0.51 – 0.15 (m, 4H); ^{13}C NMR (151 MHz, DMSO- d_6) δ ppm 161.7, 151.2, 138.2, 130.5, 126.6, 123.3, 118.5, 113.9, 107.9, 43.4, 43.2, 42.1, 21.0, 10.7, 2.9,

CONFIDENTIAL – Property of GSK – Do Not Copy

signals broadened by tautomerism resulting in one carbon signal being missing from 1D ^{13}C spectra; LCMS method A: $\text{C}_{17}\text{H}_{20}\text{N}_7\text{O}$ $[\text{M}+\text{H}]^+ = 338$, found 338, $R_T = 0.50$ min; HRMS: m/z calcd for $\text{C}_{17}\text{H}_{20}\text{N}_7\text{O}$ $[\text{M}+\text{H}]^+ = 338.1724$, found 338.1718; insufficient material for mp and IR.

***N*-(Cyclobutylmethyl)-5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide (4.194)**



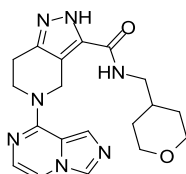
A solution of cyclobutylmethanamine hydrochloride (51 mg, 0.42 mmol), 5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxylic acid (60 mg, 0.21 mmol), triethylamine (88 μL , 0.63 mmol) and T3P (50% in EtOAc) (188 μL , 0.32 mmol) in 2-MeTHF (1 mL) was stirred at room temperature for 20 h. Additional T3P (50% in EtOAc) (188 μL , 0.32 mmol), triethylamine (88 μL , 0.63 mmol) and cyclobutylmethanamine hydrochloride (51 mg, 0.42 mmol) were added and the reaction mixture was stirred for a further 24 h at room temperature before heating at 40 $^\circ\text{C}$ for 4 h. The reaction mixture was concentrated *in vacuo* and was partitioned between saturated aqueous sodium bicarbonate (5 mL) and EtOAc (5 mL). The aqueous was washed with EtOAc (3 x 5 mL) and DCM (2 x 5 mL). The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*.

The sample was purified by reverse phase chromatography (MDAP method B, gradient B) to afford *N*-(cyclobutylmethyl)-5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide, as a white solid, 10.7 mg, 14%.

^1H NMR (600 MHz, $\text{DMSO-}d_6$) δ 13.24 – 12.63 (m, 1H), 8.42 (s, 1H), 7.98 (t, $J = 6.1$ Hz, 1H), 7.81 (s, 1H), 7.76 (d, $J = 4.8$ Hz, 1H), 7.12 (d, $J = 4.8$ Hz, 1H), 4.94 (s,

2H), 4.04 (t, $J = 5.7$ Hz, 2H), 3.26 (t, $J = 6.6$ Hz, 2H), 2.91 (t, $J = 5.5$ Hz, 2H), 2.54 (s, 1H), 1.97 (dd, $J = 10.3, 7.0$ Hz, 2H), 1.88 - 1.79 (m, 2H), 1.77 - 1.63 (m, 2H); ^{13}C NMR (151 MHz, DMSO- d_6) δ ppm 162.0, 151.2, 138.6, 130.6, 126.6, 123.4, 118.6, 114.0, 107.9, 43.5, 43.3, 43.1, 34.5, 24.9, 21.2, 17.4, signals broadened by tautomerism resulting in a carbon signal being missing from 1D ^{13}C spectra; LCMS method B: $\text{C}_{18}\text{H}_{22}\text{N}_7\text{O}$ $[\text{M}+\text{H}]^+ = 352$, found 352, $R_T = 0.79$ min; HRMS: m/z calcd for $\text{C}_{18}\text{H}_{22}\text{N}_7\text{O}$ $[\text{M}+\text{H}]^+ = 352.1880$, found 352.1876; IR cm^{-1} 3212, 2960, 1633; mp 236 – 238 °C.

5-(Imidazo[1,5-*a*]pyrazin-8-yl)-*N*-((tetrahydro-2*H*-pyran-4-yl)methyl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide (4.195)

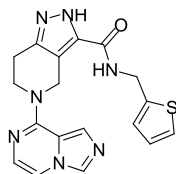


A solution (tetrahydro-2*H*-pyran-4-yl)methanamine (49 mg, 0.42 mmol), 5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxylic acid (60 mg, 0.21 mmol), triethylamine (88 μL , 0.63 mmol) and T3P (50% in EtOAc) (188 μL , 0.32 mmol) in 2-MeTHF (1 mL) was stirred at room temperature for 20 h. Additional T3P (50% in EtOAc) (188 μL , 0.32 mmol), triethylamine (88 μL , 0.63 mmol) and (tetrahydro-2*H*-pyran-4-yl)methanamine (49 mg, 0.42 mmol) were added and the reaction mixture was stirred for a further 24 h at room temperature. The reaction mixture was concentrated *in vacuo* and was partitioned between saturated aqueous sodium bicarbonate (5 mL) and EtOAc (5 mL). The aqueous was washed with EtOAc (3 x 5 mL). The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*.

The sample was purified by reverse phase chromatography (MDAP method B, gradient A) to afford 5-(imidazo[1,5-*a*]pyrazin-8-yl)-*N*-((tetrahydro-2*H*-pyran-4-yl)methyl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide, as an off white powder, 14 mg, 18%.

^1H NMR (600 MHz, $\text{DMSO-}d_6$) δ 13.03 (br.s, 1H), 8.42 (s, 1H), 8.08 (t, $J = 6.1$ Hz, 1H), 7.81 (s, 1H), 7.76 (d, $J = 4.8$ Hz, 1H), 7.12 (d, $J = 4.8$ Hz, 1H), 4.95 (s, 2H), 4.05 (t, $J = 5.7$ Hz, 2H), 3.84 (dd, $J = 11.2, 2.4$ Hz, 2H), 3.26 (td, $J = 11.6, 1.7$ Hz, 2H), 3.15 – 3.10 (m, 2H), 2.91 (t, $J = 5.3$ Hz, 2H), 1.79 (ddt, $J = 11.1, 7.4, 3.9$ Hz, 1H), 1.56 (d, $J = 12.5$ Hz, 2H), 1.19 (qd, $J = 12.2, 4.4$ Hz, 2H), contains residual methanol; ^{13}C NMR (151 MHz, $\text{DMSO-}d_6$) δ ppm 162.3, 151.3, 141.2, 138.2, 130.5, 126.5, 123.3, 118.6, 113.8, 107.7, 66.4, 43.4, 43.5, 43.1, 34.5, 30.1, 21.0; LCMS method A: $\text{C}_{19}\text{H}_{24}\text{N}_7\text{O}_2$ $[\text{M}+\text{H}]^+ = 382$ found 382, $R_T = 0.63$ min; HRMS to be run; IR cm^{-1} 3383, 3092, 2908, 2831, 1653; mp 248 – 249 °C.

5-(Imidazo[1,5-a]pyrazin-8-yl)-*N*-(thiophen-2-ylmethyl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide (4.196)



A solution thiophen-2-ylmethanamine (48 mg, 0.42 mmol), 5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxylic acid (60 mg, 0.21mmol), triethylamine (88 μL , 0.63 mmol) and T3P (50% in EtOAc) (188 μL , 0.32 mmol) in 2-MeTHF (1 mL) was stirred at room temperature for 20 h. Additional T3P (50% in EtOAc) (188 μL , 0.32 mmol), triethylamine (88 μL , 0.63 mmol) and thiophen-2-ylmethanamine (48 mg, 0.42 mmol) were added and the reaction mixture was stirred for a further 24 h at room temperature before heating at 40 °C for 4 h. The reaction mixture was concentrated *in vacuo* and was partitioned between saturated aqueous sodium bicarbonate (5 mL) and EtOAc (5 mL). The aqueous was washed with EtOAc (3 x 5 mL) and DCM (2 x 5 mL). The combined organics were dried by passing through a hydrophobic frit and were concentrated *in vacuo*.

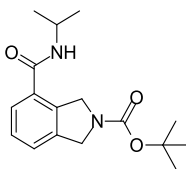
The sample was purified by reverse phase chromatography (MDAP method B, gradient B) to afford 5-(imidazo[1,5-*a*]pyrazin-8-yl)-*N*-(thiophen-2-ylmethyl)-

CONFIDENTIAL – Property of GSK – Do Not Copy

4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide, as a white solid, 7.1 mg, 9%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 13.05 (s, 1H), 8.68 (br.t., *J* = 6.0 Hz, 1H), 8.41 (s, 1H), 7.81 (s, 1H), 7.75 (dd, *J* = 4.6, 0.7 Hz, 1H), 7.36 (dd, *J* = 5.0, 1.1 Hz, 1H), 7.11 (d, *J* = 4.9 Hz, 1H), 6.99 (d, *J* = 2.7 Hz, 1H), 6.95 (dd, *J* = 5.1, 3.4 Hz, 1H), 4.94 (s, 2H), 4.57 (d, *J* = 6.1 Hz, 2H), 4.04 (t, *J* = 5.5 Hz, 2H), 2.91 (t, *J* = 5.5 Hz, 2H), contains residual deuterated methanol; LCMS method B: C₁₈H₁₈N₇OS [M+H]⁺ = 380, found 380, R_T = 0.77 min.

***tert*-Butyl 4-(isopropylcarbamoyl)isoindoline-2-carboxylate (4.204)**



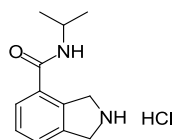
Propan-2-amine (112 mg, 1.90 mmol) was added to a solution of 2-(*tert*-butoxycarbonyl)isoindoline-4-carboxylic acid (250 mg, 0.95 mmol), 2,4,6-tripropyl-1,3,5,2,4,6-trioxatriphosphinane 2,4,6-trioxide (50% in EtOAc) (0.85 mL, 1.42 mmol) and triethylamine (0.40 mL, 2.85 mmol) in 2-MeTHF (4 mL). The reaction mixture was stirred at room temperature for 18 h. Additional propan-2-amine (112 mg, 1.90 mmol), 2,4,6-tripropyl-1,3,5,2,4,6-trioxatriphosphinane 2,4,6-trioxide (50% in EtOAc) (0.85 mL, 1.42 mmol) and triethylamine (0.40 mL, 2.85 mmol) were added. The reaction mixture was stirred at room temperature for 5 h. The reaction was worked up by addition of EtOAc (4 mL). The organic was washed with water (2 x 5 mL) and 2M HCl (5 mL). The organic was separated and dried by passing through a hydrophobic frit and was blown down under nitrogen.

The sample was purified by reverse phase chromatography (MDAP method B, gradient C) to afford *tert*-butyl 4-(isopropylcarbamoyl)isoindoline-2-carboxylate, 160 mg, 55%.

CONFIDENTIAL – Property of GSK – Do Not Copy

^1H NMR (400 MHz, DMSO- d_6) δ 8.24 - 8.16 (m, 1H), 7.64 (t, $J = 7.3$ Hz, 1H), 7.45 (t, $J = 7.3$ Hz, 1H), 7.38 (t, $J = 7.3$ Hz, 1H), 4.77 (d, $J = 12.1$ Hz, 2H), 4.59 (d, $J = 11.3$ Hz, 2H), 4.13 - 4.01 (m, 1H), 1.46 (s, 9H), 1.16 (d, $J = 6.6$ Hz, 6H); LCMS method B: $\text{C}_{17}\text{H}_{23}\text{N}_2\text{O}_3$ $[\text{M}-\text{H}]^- = 303$, found 303, $R_T = 1.06$ min.

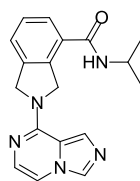
N-Isopropylisoindoline-4-carboxamide hydrochloride (4.205)



tert-Butyl 4-(isopropylcarbamoyl)isoindoline-2-carboxylate (160 mg, 0.53 mmol) was stirred in DCM (2 mL) and 4.0 N HCl in dioxane (64 μL , 0.26 mmol) at room temperature for 24 h. The reaction mixture was blown down under nitrogen at 45 $^\circ\text{C}$ to afford *N*-isopropylisoindoline-4-carboxamide hydrochloride, as a white solid, 112 mg, 88%.

^1H NMR (400 MHz, DMSO- d_6) δ 9.74 (s, 2H), 8.39 (d, $J = 7.5$ Hz, 1H), 7.80 (d, $J = 7.2$ Hz, 1H), 7.55 (d, $J = 7.2$ Hz, 1H), 7.48 (t, $J = 7.5$ Hz, 1H), 4.71 (s, 2H), 4.50 (s, 2H), 4.15 - 4.03 (m, 1H), 1.17 (d, $J = 6.6$ Hz, 6H); LCMS method B: $\text{C}_{12}\text{H}_{17}\text{N}_2\text{O}\cdot\text{HCl}$ $[\text{M}+\text{H}]^+ = 205$, found 205, $R_T = 0.56$ min.

2-(Imidazo[1,5-*a*]pyrazin-8-yl)-*N*-isopropylisoindoline-4-carboxamide (4.198)



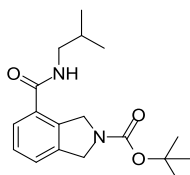
N-Isopropylisoindoline-4-carboxamide hydrochloride (78 mg, 0.33 mmol) was heated with 8-chloroimidazo[1,5-*a*]pyrazine (50 mg, 0.33 mmol) and K_2CO_3 (180 mg, 1.30 mmol) in DMSO at 60 $^\circ\text{C}$ for 5 h. The reaction mixture was diluted with MeOH (2 mL). The solution was carefully pipetted into another vial (leaving the insoluble inorganics behind) and blown down under nitrogen at 45 $^\circ\text{C}$. The crude was pre-absorbed onto diatomaceous earth and was purified by reverse phase chromatography 10 mM ammonium bicarbonate in aqueous solution (solvent A),

CONFIDENTIAL – Property of GSK – Do Not Copy

acetonitrile (solvent B)), using a Rediseq 50 g C18 gold aqueous column to afford 2-(imidazo[1,5-*a*]pyrazin-8-yl)-*N*-isopropylisoindoline-4-carboxamide, as a light tan coloured solid, 16 mg, 15%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.41 (s, 1H), 8.28 (d, *J* = 7.8 Hz, 1H), 7.86 (s, 1H), 7.73 (d, *J* = 7.6 Hz, 1H), 7.67 (d, *J* = 4.9 Hz, 1H), 7.58 (d, *J* = 7.6 Hz, 1H), 7.46 (t, *J* = 7.6 Hz, 1H), 7.10 (d, *J* = 4.6 Hz, 1H), 5.30 (s, 2H), 5.16 (s, 2H), 4.18 - 4.08 (m, 1H), 1.20 (d, *J* = 6.6 Hz, 6H); LCMS method B: C₁₈H₂₀N₅O [M+H]⁺ = 322, found 322, R_T = 0.78 min.

***tert*-Butyl 4-(isobutylcarbamoyl)isoindoline-2-carboxylate (4.206)**



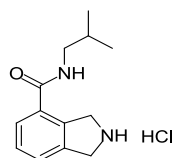
2-Methylpropan-1-amine (139 mg, 0.19 mmol) was added to a solution of 2-(*tert*-butoxycarbonyl)isoindoline-4-carboxylic acid (250 mg, 0.95 mmol), 2,4,6-tripropyl-1,3,5,2,4,6-trioxatriphosphinane 2,4,6-trioxide (50% in EtOAc) (0.85 mL, 1.42 mmol) and triethylamine (0.40 mL, 2.85 mmol) in 2-MeTHF (4 mL). The reaction mixture was stirred at room temperature for 18 h. EtOAc (4 mL) was added and the organic phase was washed with water (2 x 5 mL) and 2M HCl (5 mL). The organic was separated and dried by passing through a hydrophobic frit and was blown down under nitrogen.

The sample was purified by reverse phase chromatography (MDAP method B, gradient C) to afford *tert*-butyl 4-(isobutylcarbamoyl)isoindoline-2-carboxylate, as a white solid, 82 mg, 27%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.48-8.37 (br.m, 1H), 7.64 (dd, *J* = 7.3, 2.9 Hz, 1H), 7.46 (t, *J* = 6.7 Hz, 1H), 7.39 (t, *J* = 7.5 Hz, 1H), 4.77 (d, *J* = 13.7 Hz, 2H), 4.59 (d, *J* = 10.3 Hz, 2H), 3.09 - 3.04 (m, 2H), 1.88 - 1.78 (m, 1H), 1.46 (s, 9H), 0.89 (d, *J*

= 6.6 Hz, 6H); LCMS method B: $C_{18}H_{25}N_2O_3$ $[M-H]^- = 317$, found 317, $R_T = 1.14$ min.

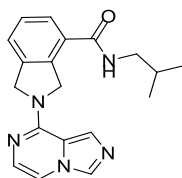
***N*-Isobutylisoindoline-4-carboxamide, hydrochloride (4.207)**



tert-Butyl 4-(isobutylcarbamoyl)isoindoline-2-carboxylate, (82 mg, 0.26 mmol) was stirred in DCM (2 mL) and 4.0 N HCl in dioxane (64 μ L, 0.26 mmol) at room temperature for 24 h. The reaction mixture was blown down under nitrogen at 45 °C to afford *N*-isobutylisoindoline-4-carboxamide hydrochloride as a white solid, 61 mg, 92%.

1H NMR (400 MHz, $DMSO-d_6$) δ 9.54 (s, 2H), 8.63 (t, $J = 5.6$ Hz, 1H), 7.80 (d, $J = 7.5$ Hz, 1H), 7.59 - 7.54 (m, 1H), 7.49 (t, $J = 7.5$ Hz, 1H), 4.71 (s, 2H), 4.50 (s, 2H), 3.12 - 3.05 (m, 2 H), 1.90 - 1.79 (m, 1H), 0.89 (d, $J = 6.8$ Hz, 6H); LCMS method B: $C_{13}H_{19}N_2O.HCl$ $[M+H]^+ = 219$, found 219, $R_T = 0.66$ min.

2-(Imidazo[1,5-*a*]pyrazin-8-yl)-*N*-isobutylisoindoline-4-carboxamide (4.199)



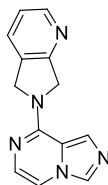
N-Isobutylisoindoline-4-carboxamide hydrochloride (83 mg, 0.33 mmol) was heated with 8-chloroimidazo[1,5-*a*]pyrazine (50 mg, 0.33 mmol) and K_2CO_3 (180 mg, 1.30 mmol) in DMSO at 60 °C for 5 h. The reaction mixture was diluted with MeOH (2 mL). The solution was carefully pipetted into another vial (leaving the insoluble inorganics behind) and blown down under nitrogen at 45 °C. The crude was pre-absorbed onto diatomaceous earth and was purified by reverse phase chromatography 10 mM ammonium bicarbonate in aqueous solution (solvent A), acetonitrile (solvent B)), using a Rediseq 50 g C18 gold aqueous column to afford 2-

CONFIDENTIAL – Property of GSK – Do Not Copy

(imidazo[1,5-*a*]pyrazin-8-yl)-*N*-isobutylisoindoline-4-carboxamide, as a red brown solid, 36 mg, 33%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.51 (t, *J* = 5.5 Hz, 1H), 8.41 (s, 1H), 7.86 (s, 1H), 7.73 (d, *J* = 7.6 Hz, 1H), 7.67 (d, *J* = 4.6 Hz, 1H), 7.59 (d, *J* = 7.6 Hz, 1H), 7.47 (t, *J* = 7.6 Hz, 1H), 7.10 (d, *J* = 4.6 Hz, 1H), 5.29 (br.s, 2H), 5.17 (br.s, 2H), 3.12 (apparent t, *J* = 6.4 Hz, 2H), 1.94 – 1.80 (m, 1H), 0.92 (d, *J* = 6.6 Hz, 6H); ¹³C NMR (101 MHz, DMSO-*d*₆) δ ppm 166.4, 149.3, 130.6, 130.1, 128.0, 127.5, 125.7, 125.2, 124.1, 124.1, 123.9, 118.8, 106.6, 54.8, 53.4, 46.5, 28.1, 20.2; LCMS method B: C₁₉H₂₂N₅O [M+H]⁺ = 336, found 336, R_T = 0.86 min; HRMS: *m/z* calcd for C₁₉H₂₂N₅O [M+H]⁺ = 336.1819, found 336.1814; IR cm⁻¹ 3277, 2955, 1632; mp – slow decomposition.

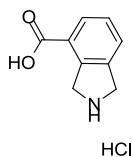
8-(5*H*-Pyrrolo[3,4-*b*]pyridin-6(7*H*)-yl)imidazo[1,5-*a*]pyrazine (4.201)



8-Chloroimidazo[1,5-*a*]pyrazine (50 mg, 0.33 mmol) was heated with 6,7-dihydro-5*H*-pyrrolo[3,4-*b*]pyridine (39 mg, 0.33 mmol) with K₂CO₃ (180 mg, 1.30 mmol) in DMSO (0.5 mL) at 60 °C for 71 h. The reaction mixture was diluted with MeOH (2 mL). The solution was carefully pipetted into another vial (leaving the insoluble inorganics behind) and blown down under nitrogen at 45 °C. The sample was purified by reverse phase chromatography (MDAP method B, gradient B) to afford 8-(5*H*-Pyrrolo[3,4-*b*]pyridin-6(7*H*)-yl)imidazo[1,5-*a*]pyrazine, as a pale brown solid, 12 mg, 16%.

¹H NMR (400 MHz, MeOD) δ 8.51 (d, *J* = 4.9 Hz, 1H), 8.36 (s, 1H), 8.02 (s, 1H), 7.92 (d, *J* = 7.6 Hz, 1H), 7.61 (d, *J* = 4.9 Hz, 1H), 7.41 (dd, *J* = 7.6, 4.9 Hz, 1H), 7.12 (d, *J* = 4.9 Hz, 1H), 5.25 (s, 2H), 5.18 (s, 2H), poor NMR due to solubility; LCMS method B: C₁₃H₁₂N₅ [M+H]⁺ = 238, found 238, R_T = 0.61 min (93.7% by UV).

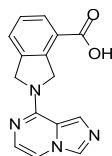
Isoindoline-4-carboxylic acid, hydrochloride (4.208)



2-(*tert*-Butoxycarbonyl)isoindoline-4-carboxylic acid (1 g, 3.80 mmol) was suspended in DCM (10 mL). 4.0 N HCl in dioxane (3.8 mL, 15.2 mmol) was added and the reaction mixture was stirred at room temperature for 2 h. The reaction mixture was heated at 45 °C for 20 h and was then left sitting at room temperature for a further 72 h. The reaction mixture was concentrated *in vacuo* to afford isoindoline-4-carboxylic acid hydrochloride, 702 mg, 93%. Isoindoline-4-carboxylic acid hydrochloride was taken onto the next step without purification.

¹H NMR (400 MHz, DMSO-*d*₆) δ 13.40 (s, 1H), 9.79 (s, 2H), 7.91 (d, *J* = 7.8 Hz, 1H), 7.67 (d, *J* = 7.8 Hz, 1H), 7.53 (t, *J* = 7.8 Hz, 1H), 4.76 (s, 2H), 4.54 (s, 2H), NMR contains minor impurities as no purification was carried out; LCMS – product too polar to be observed, no starting material remaining.

2-(Imidazo[1,5-*a*]pyrazin-8-yl)isoindoline-4-carboxylic acid (4.209)

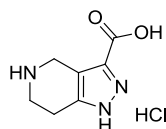


Isoindoline-4-carboxylic acid hydrochloride (65 mg, 0.33 mmol) was heated with 8-chloroimidazo[1,5-*a*]pyrazine (50 mg, 0.33 mmol) and K₂CO₃ (180 mg, 1.30 mmol) in DMSO at 60 °C for 5 h. The reaction mixture was diluted with MeOH (2 mL). The solution was carefully pipetted into another vial (leaving the insoluble inorganics behind) and blown down under nitrogen at 45 °C. The crude was pre-absorbed onto diatomaceous earth and was purified by reverse phase chromatography 10 mM ammonium bicarbonate in aqueous solution (solvent A), acetonitrile (solvent B)), using a Rediseq 50 g C18 gold aqueous column to afford 2-(imidazo[1,5-*a*]pyrazin-8-yl)isoindoline-4-carboxylic acid, as a dark red brown solid containing minor impurities, 50 mg, 55%.

CONFIDENTIAL – Property of GSK – Do Not Copy

^1H NMR (600 MHz, $\text{DMSO-}d_6$) δ 13.72 – 12.80 (m, 1H), 8.42 (s, 1H), 8.07 – 7.83 (m, 2H), 7.75 – 7.62 (m, 2H), 7.50 (t, $J = 7.7$ Hz, 1H), 7.12 (d, $J = 4.8$ Hz, 1H), 5.60 – 5.27 (m, 2H), 5.27 – 4.95 (m, 2H); ^{13}C NMR (151 MHz, $\text{DMSO-}d_6$) δ ppm 167.0, 149.1, 138.1, 130.1, 128.5, 127.6, 127.5, 126.6, 123.8, 118.6, 106.3, 55.1, 53.2, signals broadened by tautomerism resulting in carbons being missing from the spectrum; LCMS method B: $\text{C}_{15}\text{H}_{13}\text{N}_4\text{O}_2$ $[\text{M}+\text{H}]^+ = 281$ found 281, $R_T = 0.50$ min. HRMS: failed due to poor solubility causing material to precipitate from solution whilst queuing for analysis; IR cm^{-1} 3109 (weak), 3004 (weak), 2843 (weak), 2466, 1696; mp – decomposed at 296 °C.

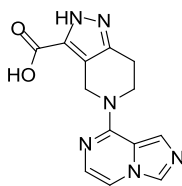
4,5,6,7-Tetrahydro-1H-pyrazolo[4,3-c]pyridine-3-carboxylic acid, hydrochloride (4.239)



5-(*tert*-Butoxycarbonyl)-4,5,6,7-tetrahydro-1H-pyrazolo[4,3-c]pyridine-3-carboxylic acid (5.2 g, 19.5 mmol) was dissolved in 2-propanol (30 mL). Hydrogen chloride (5.0-6.0 M in 2-propanol) (19.5 mL, 97.0 mmol) was added and the reaction mixture was heated at 65 °C for 18 h. The reaction mixture was cooled and the white solid was collected by filtration. The solid was washed with 2-propanol (20 mL). The solid was dried at 45 °C in a vacuum oven to afford 4,5,6,7-tetrahydro-1H-pyrazolo[4,3-c]pyridine-3-carboxylic acid hydrochloride as a white solid, 3.5 g, 87%.

^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 13.42 (s, 2H), 9.63 (s, 2H), 4.21 (s, 2H), 3.35 (s, 2H), 2.94 (t, $J = 6.0$ Hz, 2H), NMR contains minor impurities as no purification was carried out; LCMS – product too polar to be observed, no starting material remaining.

5-(Imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxylic acid (4.187)



4,5,6,7-Tetrahydro-1*H*-pyrazolo[4,3-*c*]pyridine-3-carboxylic acid hydrochloride (3.3 g, 16.3 mmol), 8-chloroimidazo[1,5-*a*]pyrazine (2.5 g, 16.3 mmol) and sodium bicarbonate (6.8 g, 81.0 mmol) were heated at 70 °C in DMSO (20 mL) under nitrogen for 44 h.

Sodium bicarbonate (2.5 g, 29.8 mmol) was added and the reaction mixture was heated under nitrogen at 80 °C for a further 24 h. The reaction mixture was cooled and MeCN (50 mL) was added to the reaction. The solid was collected by filtration, washed with additional MeCN (30 mL) and dried under vacuum. The crude was pre-absorbed onto diatomaceous earth and was purified by reverse phase chromatography 10 mM ammonium bicarbonate in aqueous solution (solvent A), acetonitrile (solvent B)), using a Redisep 150 g C18 gold aqueous column to afford 5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxylic acid in two batches. Batch one as a tan coloured solid, 2.8 g, 61%; Batch 2 as a dark brown coloured solid, 1.6 g, 35%.

Batch One:

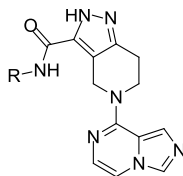
¹H NMR (400 MHz, DMSO-*d*₆) δ 12.32 (s, 1H), 8.39 (s, 1H), 7.79 (s, 1H), 7.71 (dd, *J* = 4.7, 0.6 Hz, 1H), 7.09 (d, *J* = 4.7 Hz, 1H), 4.95 (s, 2H), 4.02 (t, *J* = 5.6 Hz, 2H), 2.79 (t, *J* = 5.6 Hz, 2H), only one exchangeable observed, assumed too broad; LCMS method B: C₁₃H₁₃N₆O₂ [M+H]⁺ = 285, found 285, R_T = 0.41 min.

Batch Two:

¹H NMR (400 MHz, DMSO-*d*₆) δ 12.32 (s, 1H), 8.39 (s, 1H), 7.79 (s, 1H), 7.71 (dd, *J* = 4.7, 0.62 Hz, 1H), 7.09 (d, *J* = 4.7 Hz, 1H), 4.95 (s, 2H), 4.02 (t, *J* = 5.6 Hz, 2H),

2.79 (t, $J = 5.6$ Hz, 2H), only one exchangeable observed, assumed too broad; LCMS method B: $C_{13}H_{13}N_6O_2$ $[M+H]^+ = 285$, found 285, $R_T = 0.41$ min.

General array procedure used by arrays team for synthesis of (4.230) and (4.222).

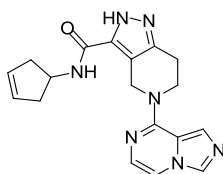


A stock solution was prepared containing 5-(Imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxylic acid (1.0 g, 3.6 mmol), HATU (1.4 g, 3.6 mmol), and DIPEA (1.9 mL, 10.8 mmol) in DMF (9.6 mL). An aliquot of this solution (480 μ L each) was added to a set of pre-weighed amine monomers (24 x 0.2 mmol) in micronic vials. The reaction mixtures were capped, shaken to aid dispersion and stood at room temp for 18 h.

To each of these reaction mixtures PyBOP (1 eq) was added and the reactions stood at room temp for 18 h.

To aid solubility H_2O (0.8 mL) and a few drops of ammonia hydroxide were added to each vial. The samples were purified directly by reverse phase chromatography (MDAP method B, gradient B) to afford the desired products.

Characterisation for *N*-(Cyclopent-3-en-1-yl)-5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide (4.222)

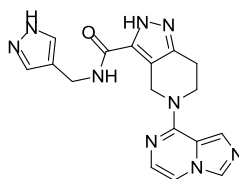


1H NMR (400 MHz, $DMSO-d_6$) δ 13.05 (s, 1H), 8.41 (s, 1H), 7.97 (d, $J = 7.6$ Hz, 1H), 7.81 (s, 1H), 7.75 (dd, $J = 4.5, 2.1$ Hz, 1H), 7.11 (d, $J = 4.5$ Hz, 1H), 5.72 (s,

CONFIDENTIAL – Property of GSK – Do Not Copy

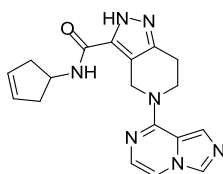
2H), 4.94 (s, 2H), 4.58 - 4.47 (m, 1H), 4.04 (t, $J = 5.7$ Hz, 2H), 2.90 (t, $J = 5.3$ Hz, 2H), 2.70 - 2.57 (m, 2H), 2.37 - 2.28 (m, 2H); LCMS method A: $C_{18}H_{20}N_7O$ $[M+H]^+$ = 350, found 350, $R_T = 0.50$ min.

Characterisation for *N*-((1*H*-Pyrazol-4-yl)methyl)-5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide (4.230)



1H NMR (600 MHz, $DMSO-d_6$) δ 13.02 (s, 0.5H), 12.61 (br.s, 0.5H), 8.41 (s, 1H), 8.31 (t, $J = 5.9$ Hz, 1H), 7.81 (s, 1H), 7.75 (d, $J = 4.5$ Hz, 1H), 7.52 (br.s, 2H), 7.11 (d, $J = 4.5$ Hz, 1H), 5.75 (s, 1H), 4.94 (s, 2H), 4.29 (s, 2H), 4.10 (br.s, 1H), 4.04 (t, $J = 5.5$ Hz, 2H), 2.87 - 2.92 (m, 2H), contains unidentified impurities; LCMS method A: $C_{17}H_{18}N_9O$ $[M+H]^+$ = 364, found 364, $R_T = 0.35$ min.

Remake of *N*-(Cyclopent-3-en-1-yl)-5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide (4.222)



Cyclopent-2-enamine hydrochloride (42 mg, 0.35 mmol) was added to a solution of 5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxylic acid (100 mg, 0.35 mmol), T3P (50% in DMF) (448 mg, 0.70 mmol) and DIPEA (0.31 mL, 1.76 mmol) in EtOAc (3 mL). DMF (1 mL) was added to aid with solubility. The reaction mixture was stirred at RT for 18 h. The reaction mixtures were warmed to 50 °C for 8 h. T3P (50% in DMF) (448 mg, 0.70 mmol) and DIPEA

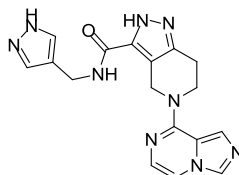
CONFIDENTIAL – Property of GSK – Do Not Copy

(0.31 mL, 1.76 mmol) were added and the reaction mixture was left sitting at RT for 16 h. The reaction mixtures was heated at 50 °C for 24 h before concentrating *in vacuo*.

The reaction mixtures were purified by reverse phase chromatography on a Redisep C18 aqueous 100 g cartridge. The crude was loaded onto to the column in 1:2:2 water:MeOH:DMSO (5 mL). The column was eluted with a gradient of 0 - 100% MeOH in water containing 10 mM ammonium bicarbonate. The appropriate fractions were combined and were concentrated in vacuo. The compound was repurified by reverse phase chromatography (Method B, Gradient B). The appropriate fractions were combined and concentrated *in vacuo* to afford *N*-(Cyclopent-3-en-1-yl)-5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide, as a cream solid, 27 mg, 22%.

¹H NMR (600 MHz, DMSO-*d*₆) δ 13.01 (br.s, 1H), 8.42 (s, 1H), 7.97 (br.s, 1H), 7.81 (s, 1H), 7.75 (d, *J* = 4.8 Hz, 1H), 7.12 (d, *J* = 4.8 Hz, 1H), 5.73 (s, 2H), 4.95 (s, 2H), 4.57 - 4.49 (m, 1H), 4.04 (t, *J* = 5.4 Hz, 2H), 2.91 (t, *J* = 5.3 Hz, 2H), 2.64 (dd, *J* = 15.0, 8.1 Hz, 2H), 2.34 (dd, *J* = 14.7, 5.5 Hz, 2H); ¹³C NMR (151 MHz, DMSO-*d*₆) 162.1, 151.4, 141.4, 138.5, 131.0, 129.0, 126.8, 123.6, 118.6, 114.0, 108.2, 48.2, 43.7, 43.5, 21.4, some of the signals are broadened by tautomerism; LCMS (Acquity UPLC CSH C18 column) method B: C₁₈H₂₀N₇O [M+H]⁺ = 350, found 350, R_T = 0.76 min; HRMS: *m/z* calcd for C₁₈H₂₀N₇O [M+H]⁺ = 350.1724, found 350.1712; IR cm⁻¹ 3225, 3107, 2988, 2847, 1625; mp 234 – 236 °C (decomposed).

Remake of *N*-((1*H*-Pyrazol-4-yl)methyl)-5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide (4.230)



(1*H*-Pyrazol-4-yl)methanamine dihydrochloride (127 mg, 0.75 mmol) was added to a solution of 5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxylic acid (213 mg, 0.75 mmol), T3P (50% in DMF) (715 mg, 1.12 mmol) and DIPEA (0.79 mL, 4.50 mmol). The reaction mixture was stirred at 60 °C for 28 h and was left sitting at room temperature for a further 40 h. The reaction mixture was refluxed under nitrogen for a further 4 h. The reaction mixture was concentrated *in vacuo*. The crude was purified MDAP (Method A, Gradient A). The appropriate fractions were combined and concentrated *in vacuo* to afford *N*-((1*H*-pyrazol-4-yl)methyl)-5-(imidazo[1,5-*a*]pyrazin-8-yl)-4,5,6,7-tetrahydro-2*H*-pyrazolo[4,3-*c*]pyridine-3-carboxamide as a colourless glass/gum, 31 mg, 11%.

¹H NMR (400 MHz, DMSO-*d*₆) δ 8.41 (s, 1H), 8.39 (s, 1H underintegrates), 8.32 (t, *J* = 6.0 Hz, 1H), 7.81 (s, 1H), 7.75 (d, *J* = 4.6 Hz, 1H), 7.52 (s, 2H), 7.11 (d, *J* = 4.6 Hz, 1H), 4.95 (s, 2H), 4.29 (d, *J* = 6.1 Hz, 2H), 4.04 (t, *J* = 5.5 Hz, 2H), 2.90 (t, *J* = 5.5 Hz, 2H), one exchangeable proton not observed, contains minor grease impurities in aliphatic region; ¹³C NMR (101 MHz, DMSO-*d*₆) 164.9, 161.8, 151.5, 130.9, 126.9, 123.7, 118.6, 118.2, 114.1, 108.2, 43.7, 43.6, 32.5, 21.5, two signals are not observed, broadened by tautomerism; LCMS (Acquity UPLC CSH C18 column) method B: C₁₇H₁₈N₉O [M+H]⁺ = 364, found 364, R_T = 0.60 min; HRMS: *m/z* calcd for C₁₇H₁₈N₉O [M+H]⁺ = 364.1629, found 364.1619; IR cm⁻¹ 3114, 2902, 1636; mp 228 – 230 °C (decomposed).

7 References

1. Leeson, P. D.; Springthorpe, B. *Nat. Rev. Drug Discovery* **2007**, *6*, 881-890.
2. Arnott, J. A.; Planey, S. L. *Expert Opin. Drug Discovery* **2012**, *7*, 863-875.
3. Murray, C. W.; Rees, D. C. *Nat. Chem.* **2009**, *1*, 187-192.
4. Waring, M. J. *Expert Opin. Drug Discovery* **2010**, *5*, 235-248.
5. Muthas, D.; Boyer, S.; Hasselgren, C. *Med. Chem. Commun.* **2013**, *4*, 1058-1065.
6. Leo, A.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525-616.
7. Barbato, F.; Cappello, B.; Miro, A.; La Rotonda, M. I.; Quaglia, F. *Farmaco* **1998**, *53*, 655-661.
8. Valko, K. *J. Chromatogr. A* **2004**, *1037*, 299-310.
9. Bhattachar, S. N.; Deschenes, L. A.; Wesley, J. A. *Drug Discovery Today* **2006**, *11*, 1012-1018.
10. van de Waterbeemd, H.; Gifford, E. *Nat. Rev. Drug Discovery* **2003**, *2*, 192-204.
11. El-Kattan, A.; Varma, M. (2012). Oral Absorption, Intestinal Metabolism and Human Oral Bioavailability, Topics on Drug Metabolism, Dr. James Paxton (Ed.), ISBN: 978-953-51-0099-7, InTech, Available from: <http://www.intechopen.com/books/topics-on-drug-metabolism/oral-absorption-intestinalmetabolism-and-human-oral-bioavailability>. **22/11/14**.
12. Young, R. J.; Green, D. V. S.; Luscombe, C. N.; Hill, A. P. *Drug Discovery Today* **2011**, *16*, 822-830.

13. Bembenek, S. D.; Tounge, B. A.; Reynolds, C. H. *Drug Discovery Today* **2009**, *14*, 278-283.
14. Wager, T. T.; Chandrasekaran, R. Y.; Hou, X.; Troutman, M. D.; Verhoest, P. R.; Villalobos, A.; Will, Y. *ACS Chem. Neurosci.* **2010**, *1*, 420-434.
15. Mortenson, P. N.; Murray, C. W. *J. Comput. -Aided Mol. Des.* **2011**, *25*, 663-667.
16. Murray, C. W.; Erlanson, D. A.; Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H.; Richmond, N. J. *ACS Med. Chem. Lett.* **2014**, *5*, 616-618.
17. Schultes, S.; de Graaf, C.; Haaksma, E. E. J.; de Esch, I. J. P.; Leurs, R.; Krämer, O. *Drug Discovery Today: Technol.* **2010**, *7*, 157-162.
18. Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. *Nat. Rev. Drug Discovery* **2014**, *13*, 105-121.
19. Chessari, G.; Woodhead, A. J. *Drug Discovery Today* **2009**, *14*, 668-675.
20. Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. *Nat. Rev. Drug Discovery* **2004**, *3*, 660-672.
21. Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J. *J. Med. Chem.* **2008**, *51*, 3661-3680.
22. Franzini, R. M.; Neri, D.; Scheuermann, J. *Acc. Chem. Res.* **2014**, *47*, 1247-1255.
23. Clark, M. A. *Curr. Opin. Chem. Biol.* **2010**, *14*, 396-403.
24. Kollmann, C. S.; Bai, X.; Tsai, C.-H.; Yang, H.; Lind, K. E.; Skinner, S. R.; Zhu, Z.; Israel, D. I.; Cuzzo, J. W.; Morgan, B. A.; Yuki, K.; Xie, C.; Springer, T. A.; Shimaoka, M.; Evindar, G. *Bioorg. Med. Chem.* **2014**, *22*, 2353-2365.

25. Deng, H.; O'Keefe, H.; Davie, C. P.; Lind, K. E.; Acharya, R. A.; Franklin, G. J.; Larkin, J.; Matico, R.; Neeb, M.; Thompson, M. M.; Lohr, T.; Gross, J. W.; Centrella, P. A.; O'Donovan, G. K.; Bedard, K. L.; van Vloten, K.; Mataruse, S.; Skinner, S. R.; Belyanskaya, S. L.; Carpenter, T. Y.; Shearer, T. W.; Clark, M. A.; Cuozzo, J. W.; Arico-Muendel, C. C.; Morgan, B. A. *J. Med. Chem.* **2012**, *55*, 7061-7079.
26. Disch, J. S.; Evindar, G.; Chiu, C. H.; Blum, C. A.; Dai, H.; Jin, L.; Schuman, E.; Lind, K. E.; Belyanskaya, S. L.; Deng, J.; Coppo, F.; Aquilani, L.; Graybill, T. L.; Cuozzo, J. W.; Lavu, S.; Mao, C.; Vlasuk, G. P.; Perni, R. B. *J. Med. Chem.* **2013**, *56*, 3666-3679.
27. Thalji, R. K.; McAtee, J. J.; Belyanskaya, S.; Brandt, M.; Brown, G. D.; Costell, M. H.; Ding, Y.; Dodson, J. W.; Eisennagel, S. H.; Fries, R. E.; Gross, J. W.; Harpel, M. R.; Holt, D. A.; Israel, D. I.; Jolivette, L. J.; Krosky, D.; Li, H.; Lu, Q.; Mandichak, T.; Roethke, T.; Schnackenberg, C. G.; Schwartz, B.; Shewchuk, L. M.; Xie, W.; Behm, D. J.; Douglas, S. A.; Shaw, A. L.; Marino, J. P., Jr. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 3584-3588.
28. Hu, L.; Magesh, S.; Chen, L.; Wang, L.; Lewis, T. A.; Chen, Y.; Khodier, C.; Inoyama, D.; Beamer, L. J.; Emge, T. J.; Shen, J.; Kerrigan, J. E.; Kong, A.-N. T.; Dandapani, S.; Palmer, M.; Schreiber, S. L.; Munoz, B. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 3039-3043.
29. Hajduk, P. J. *J. Med. Chem.* **2006**, *49*, 6972-6976.
30. Churcher, I. GlaxoSmithKline, *Unpublished results*.
31. Blum, L. C.; Reymond, J. L. *J. Am. Chem. Soc.* **2009**, *131*, 8732-8733.
32. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. *Drug Discov. Today* **2003**, *8*, 876-877.
33. Orita, M.; Warizaya, M.; Amano, Y.; Ohno, K.; Niimi, T. *Expert Opin. Drug Discovery* **2009**, *4*, 1125-1144.

34. Carr, R. A.; Congreve, M.; Murray, C. W.; Rees, D. C. *Drug Discovery Today* **2005**, *10*, 987-992.
35. Baurin, N.; Aboul-Ela, F.; Barril, X.; Davis, B.; Drysdale, M.; Dymock, B.; Finch, H.; Fromont, C.; Richardson, C.; Simmonite, H.; Hubbard, R. E. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2157-2166.
36. Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B. *J. Biomol. NMR* **2001**, *21*, 349-359.
37. Neumann, T.; Junker, H. D.; Schmidt, K.; Sekul, R. *Curr. Top. Med. Chem.* **2007**, *7*, 1630-1642.
38. Ward, W. H.; Holdgate, G. A. *Prog. Med. Chem.* **2001**, *38*, 309-376.
39. Taylor, J. D.; Gilbert, P. J.; Williams, M. A.; Pitt, W. R.; Ladbury, J. E. *Proteins: Struct. Funct. Bioinf.* **2007**, *67*, 981-990.
40. Pantoliano, M. W.; Petrella, E. C.; Kwasnoski, J. D.; Lobanov, V. S.; Myslik, J.; Graf, E.; Carver, T.; Asel, E.; Springer, B. A.; Lane, P.; Salemme, F. R. *J. Biomol. Screening* **2001**, *6*, 429-440.
41. Hofstadler, S. A.; Sannes-Lowery, K. A. *Nat. Rev. Drug Discovery* **2006**, *5*, 585-595.
42. Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531-1534.
43. Bonser, S.; Leesnitzer, L. GlaxoSmithKline, *Unpublished results*.
44. Degorce, F.; Card, A.; Soh, S.; Trinquet, E.; Knapik, G. P.; Xie, B. *Curr. Chem. Genomics* **2009**, *3*, 22-32.
45. http://www.htrf.com/technology/htrftheory/tr_basics/ **22/11/14**.

46. <http://www.perkinelmer.com/CMSResources/Images/44-72337LUcamp5.jpg>
22/11/14.
47. Kokko, T.; Kokko, L.; Soukka, T. *J. Fluoresc.* **2009**, *19*, 159-164.
48. Horton, R. A.; Vogel, K. W. *J. Biomol. Screening* **2010**, *15*, 1008-1015.
49. Zhang, R.; Monsma, F. *Curr. Opin. Drug Discovery Dev.* **2010**, *13*, 389-402.
50. Lo, M.-C.; Aulabaugh, A.; Jin, G.; Cowling, R.; Bard, J.; Malamas, M.; Ellestad, G. *Anal. Biochem.* **2004**, *332*, 153-159.
51. Perspicace, S.; Banner, D.; Benz, J.; Müller, F.; Schlatter, D.; Huber, W. *J. Biomol. Screening* **2009**, *14*, 337-349.
52. http://employees.csbsju.edu/hjakubowski/classes/ch331/bind/olbinde_rveq.html **22/11/14.**
53. Huber, W.; Mueller, F. *Curr. Pharm. Des.* **2006**, *12*, 3999-4021.
54. <http://www.endocytosis.org/techniqs/ITC.html> **22/11/14.**
55. Zartler, E. R.; Shapiro, M. J. *Curr. Pharm. Des.* **2006**, *12*, 3963-3972.
56. Lepre, C. A.; Moore, J. M.; Peng, J. W. *Chem. Rev.* **2004**, *104*, 3641-3675.
57. Barker, J.; Hesterkamp, T.; Schade, M.; Whittaker, M. *Innovations Pharm. Technol.* **2007**, *23*, 19-22.
58. Peng, J. W.; Lepre, C. A.; Fejzo, J.; Abdul-Manan, N.; Moore, J. M. *Methods Enzymol.* **2001**, *338*, 202-230.
59. Sahakyan, A. B.; Vranken, W. F.; Cavalli, A.; Vendruscolo, M. *J. Biomol. NMR* **2011**, *50*, 331-346.
60. Carvalho, A. L.; Trincão, J.; Romão, M. J. *Methods Mol. Biol.* **2010**, *572*, 31-56.

61. Blundell, T. L.; Jhoti, H.; Abell, C. *Nat. Rev. Drug Discovery* **2002**, *1*, 45-54.
62. Sharff, A.; Jhoti, H. *Curr. Opin. Chem. Biol.* **2003**, *7*, 340-345.
63. Hartshorn, M. J.; Murray, C. W.; Cleasby, A.; Frederickson, M.; Tickle, I. J.; Jhoti, H. *J. Med. Chem.* **2005**, *48*, 403-413.
64. Bosch, J.; Robien, M. A.; Mehlin, C.; Boni, E.; Riechers, A.; Buckner, F. S.; Van Voorhis, W. C.; Myler, P. J.; Worthey, E. A.; DeTitta, G.; Luft, J. R.; Lauricella, A.; Gulde, S.; Anderson, L. A.; Kalyuzhniy, O.; Neely, H. M.; Ross, J.; Earnest, T. N.; Soltis, M.; Schoenfeld, L.; Zucker, F.; Merritt, E. A.; Fan, E.; Verlinde, C. L. M. J.; Hol, W. G. J. *J. Med. Chem.* **2006**, *49*, 5939-5946.
65. Hubbard, R. E.; Davis, B.; Chen, I.; Drysdale, M. J. *Curr. Top. Med. Chem.* **2007**, *7*, 1568-1581.
66. Olsson, T. S. G.; Williams, M. A.; Pitt, W. R.; Ladbury, J. E. *J. Mol. Biol.* **2008**, *384*, 1002-1017.
67. Shimokhina, N.; Bronowska, A.; Homans, S. W. *Angew. Chem. Int. Ed.* **2006**, *45*, 6374-6376.
68. Boehm, H. J.; Klebe, G. *Angew. Chem. Int. Ed.* **1996**, *35*, 2588-2614.
69. Mack, H.; Pfeiffer, T.; Hornberger, W.; Böhm, H. J.; Höffken, H. W. *J. Enzyme Inhib.* **1995**, *9*, 73-86.
70. Leung, C. S.; Leung, S. S. F.; Tirado-Rives, J.; Jorgensen, W. L. *J. Med. Chem.* **2012**, *55*, 4489-4500.
71. Loladze, V. V.; Ermolenko, D. N.; Makhatadze, G. I. *J. Mol. Biol.* **2002**, *320*, 343-357.
72. Bissantz, C.; Kuhn, B.; Stahl, M. *J. Med. Chem.* **2010**, *53*, 5061-5084.

73. Freire, E. *Protein Rev.* **2005**, *3*, 291-307.
74. Freire, E. *Drug Discovery Today* **2008**, *13*, 869-874.
75. Edink, E.; Jansen, C.; Leurs, R.; de Esch, I. J. P. *Drug Discovery Today: Technol.* **2010**, *7*, E189-E201.
76. Garbett, N. C.; Chaires, J. B. *Expert Opin. Drug Discovery* **2012**, *7*, 299-314.
77. Mills, J. E. J.; Dean, P. M. *J. Comput. -Aided Mol. Des.* **1996**, *10*, 607-622.
78. Clark, T.; Hennemann, M.; Murray, J. S.; Politzer, P. *J. Mol. Model.* **2007**, *13*, 291-296.
79. Panigrahi, S. K.; Desiraju, G. R. *Proteins: Struct. Funct. Bioinf.* **2007**, *67*, 128-141.
80. Imai, Y. N.; Inoue, Y.; Yamamoto, Y. *J. Med. Chem.* **2007**, *50*, 1189-1196.
81. Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem. Int. Ed.* **2003**, *42*, 1210-1250.
82. http://www.chem.tamu.edu/faculty/faculty_detail.php?ID=9005 **26/01/15**.
83. Hajduk, P. J.; Greer, J. *Nat. Rev. Drug Discovery* **2007**, *6*, 211-219.
84. Woodhead, A. J.; Angove, H.; Carr, M. G.; Chessari, G.; Congreve, M.; Coyle, J. E.; Cosme, J.; Graham, B.; Day, P. J.; Downham, R.; Fazal, L.; Feltell, R.; Figueroa, E.; Frederickson, M.; Lewis, J.; McMenamin, R.; Murray, C. W.; O'Brien, M. A.; Parra, L.; Patel, S.; Phillips, T.; Rees, D. C.; Rich, S.; Smith, D.-M.; Trewartha, G.; Vinkovic, M.; Williams, B.; Woolford, A. J.-A. *J. Med. Chem.* **2010**, *53*, 5956-5969.
85. Wells, J. A.; McClendon, C. L. *Nature* **2007**, *450*, 1001-1009.
86. <http://practicalfragments.blogspot.co.uk/2013/01/fragments-in-clinic-2013-edition.html> **26/01/15**.

87. <http://practicalfragments.blogspot.com/2011/08/first-fragment-based-drug-approved.html> **26/01/15**.
88. Bollag, G.; Hirth, P.; Tsai, J.; Zhang, J.; Ibrahim, P. N.; Cho, H.; Spevak, W.; Zhang, C.; Zhang, Y.; Habets, G.; Burton, E. A.; Wong, B.; Tsang, G.; West, B. L.; Powell, B.; Shellooe, R.; Marimuthu, A.; Nguyen, H.; Zhang, K. Y. J.; Artis, D. R.; Schlessinger, J.; Su, F.; Higgins, B.; Iyer, R.; D'Andrea, K.; Koehler, A.; Stumm, M.; Lin, P. S.; Lee, R. J.; Grippo, J.; Puzanov, I.; Kim, K. B.; Ribas, A.; McArthur, G. A.; Sosman, J. A.; Chapman, P. B.; Flaherty, K. T.; Xu, X.; Nathanson, K. L.; Nolop, K. *Nature* **2010**, *467*, 596-599.
89. Tsai, J.; Lee, J. T.; Wang, W.; Zhang, J.; Cho, H.; Mamo, S.; Bremer, R.; Gillette, S.; Kong, J.; Haass, N. K.; Sproesser, K.; Li, L.; Smalley, K. S. M.; Fong, D.; Zhu, Y.-L.; Marimuthu, A.; Nguyen, H.; Lam, B.; Liu, J.; Cheung, I.; Rice, J.; Suzuki, Y.; Luu, C.; Settachatgul, C.; Shellooe, R.; Cantwell, J.; Kim, S.-H.; Schlessinger, J.; Zhang, K. Y. J.; West, B. L.; Powell, B.; Habets, G.; Zhang, C.; Ibrahim, P. N.; Hirth, P.; Artis, D. R.; Herlyn, M.; Bollag, G. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3041-3046.
90. Bower, J. F.; Pannifer, A. *Curr. Pharm. Des.* **2012**, *18*, 4685-4696.
91. Christ, F.; Voet, A.; Marchand, A.; Nicolet, S.; Desimmie, B. A.; Marchand, D.; Bardiot, D.; Van der Veken, N. J.; Van Remoortel, B.; Strelkov, S. V.; De Maeyer, M.; Chaltin, P.; Debyser, Z. *Nat. Chem. Biol.* **2010**, *6*, 442-448.
92. Greene, W. C.; Debyser, Z.; Ikeda, Y.; Freed, E. O.; Stephens, E.; Yonemoto, W.; Buckheit, R. W.; Esté, J. A.; Cihlar, T. *Antiviral Res.* **2008**, *80*, 251-265.
93. Marchand, C.; Maddali, K.; Metifiot, M.; Pommier, Y. *Curr. Top. Med. Chem.* **2009**, *9*, 1016-1037.
94. Max, B.; Sherer, R. *Clin. Infect. Dis.* **2000**, *30*, S96-S116.
95. Pace, P.; Rowley, M. *Curr. Opin. Drug Discovery Dev.* **2008**, *11*, 471-479.

96. De Luca, L.; Ferro, S.; Morreale, F.; Chimirri, A. *Mini-Rev. Med. Chem.* **2011**, *11*, 714-727.
97. Rhodes, D. I.; Peat, T. S.; Vandegraaff, N.; Jeevarajah, D.; Newman, J.; Martyn, J.; Coates, J. A. V.; Ede, N. J.; Rea, P.; Deadman, J. J. *Chem. Bio. Chem.* **2011**, *12*, 2311-2315.
98. Johnson, B. C.; Métifiot, M.; Pommier, Y.; Hughes, S. H. *Antimicrob. Agents Chemother.* **2012**, *56*, 411-419.
99. De Luca, L.; Ferro, S.; Morreale, F.; De Grazia, S.; Chimirri, A. *Chem. Med. Chem.* **2011**, *6*, 1184-1191.
100. Pohlhaus, D. GlaxoSmithKline, *Unpublished results*.
101. Christ, F.; Debyser, Z. *Virology* **2013**, *435*, 102-109.
102. De Luca, L.; Gitto, R.; Christ, F.; Ferro, S.; De Grazia, S.; Morreale, F.; Debyser, Z.; Chimirri, A. *Antiviral Res.* **2011**, *92*, 102-107.
103. Arkin, M. R.; Wells, J. A. *Nat. Rev. Drug Discovery* **2004**, *3*, 301-317.
104. Dietrich, U.; Dürr, R.; Koch, J. *Curr. Pharm. Biotechnol.* **2013**, *14*, 501-512.
105. Glas, A.; Bier, D.; Hahne, G.; Rademacher, C.; Ottmann, C.; Grossmann, T. *N. Angew. Chem. Int. Ed.* **2014**, *53*, 2489-2493.
106. Zhang, Z.; Lin, Z.; Zhou, Z.; Shen, H. C.; Yan, S. F.; Mayweg, A. V.; Xu, Z.; Qin, N.; Wong, J. C.; Zhang, Z.; Rong, Y.; Fry, D. C.; Hu, T. *ACS Med. Chem. Lett.* **2014**, *5*, 993-998.
107. Hilinski, G. J.; Kim, Y.-W.; Hong, J.; Kutchukian, P. S.; Crenshaw, C. M.; Berkovitch, S. S.; Chang, A.; Ham, S.; Verdine, G. L. *J. Am. Chem. Soc.* **2014**, *136*, 12314-12322.

108. Hayouka, Z.; Hurevich, M.; Levin, A.; Benyamini, H.; Iosub, A.; Maes, M.; Shalev, D. E.; Loyter, A.; Gilon, C.; Friedler, A. *Bioorg. Med. Chem.* **2010**, *18*, 8388-8395.
109. De Luca, L.; Morreale, F.; Christ, F.; Debyser, Z.; Ferro, S.; Gitto, R. *Eur. J. Med. Chem.* **2013**, *68*, 405-411.
110. De Luca, L.; Ferro, S.; Morreale, F.; Christ, F.; Debyser, Z.; Chimirri, A.; Gitto, R. *J. Enzyme Inhib. Med. Chem.* **2013**, *28*, 1002-1009.
111. Serrao, E.; Debnath, B.; Otake, H.; Kuang, Y.; Christ, F.; Debyser, Z.; Neamati, N. *J. Med. Chem.* **2013**, *56*, 2311-2322.
112. Rhodes, D. I.; Peat, T. S.; Vandegraaff, N.; Jeevarajah, D.; Le, G.; Jones, E. D.; Smith, J. A.; Coates, J. A. V.; Winfield, L. J.; Thienthong, N.; Newman, J.; Lucent, D.; Ryan, J. H.; Savage, G. P.; Francis, C. L.; Deadman, J. J. *Antiviral Chem. Chemother.* **2011**, *21*, 155-168.
113. Wielens, J.; Headey, S. J.; Deadman, J. J.; Rhodes, D. I.; Parker, M. W.; Chalmers, D. K.; Scanlon, M. J. *Chem. Med. Chem.* **2011**, *6*, 258-261.
114. Molteni, V.; Greenwald, J.; Rhodes, D.; Hwang, Y.; Kwiatkowski, W.; Bushman, F. D.; Siegel, J. S.; Choe, S. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2001**, *57*, 536-544.
115. <http://www.rcsb.org/pdb/explore/images.do?structureId=1HYV> **09/06/15**.
116. Loregian, A.; Palù, G. *J. Cell. Physiol.* **2005**, *204*, 750-762.
117. Hayouka, Z.; Levin, A.; Maes, M.; Hadas, E.; Shalev, D. E.; Volsky, D. J.; Loyter, A.; Friedler, A. *Biochem. Biophys. Res. Commun.* **2010**, *394*, 260-265.

118. Biron, E.; Chatterjee, J.; Ovadia, O.; Langenegger, D.; Brueggen, J.; Hoyer, D.; Schmid, H. A.; Jelinek, R.; Gilon, C.; Hoffman, A.; Kessler, H. *Angew. Chem. Int. Ed.* **2008**, *47*, 2595-2599.
119. Du, L.; Zhao, Y.; Chen, J.; Yang, L.; Zheng, Y.; Tang, Y.; Shen, X.; Jiang, H. *Biochem. Biophys. Res. Commun.* **2008**, *375*, 139-144.
120. Christ, F.; Shaw, S.; Demeulemeester, J.; Desimmie, B. A.; Marchand, A.; Butler, S.; Smets, W.; Chaltin, P.; Westby, M.; Debyser, Z.; Pickford, C. *Antimicrob. Agents Chemother.* **2012**, *56*, 4365-4374.
121. Fader, L. D.; Malenfant, E.; Parisien, M.; Carson, R.; Bilodeau, F.; Landry, S.; Pesant, M.; Brochu, C.; Morin, S.; Chabot, C.; Halmos, T.; Bousquet, Y.; Bailey, M. D.; Kawai, S. H.; Coulombe, R.; LaPlante, S.; Jakalian, A.; Bhardwaj, P. K.; Wernic, D.; Schroeder, P.; Amad, M.; Edwards, P.; Garneau, M.; Duan, J.; Cordingley, M.; Bethell, R.; Mason, S. W.; Bös, M.; Bonneau, P.; Poupart, M.-A.; Faucher, A. M.; Simoneau, B.; Fenwick, C.; Yoakim, C.; Tsantrizos, Y. *ACS Med. Chem. Lett.* **2014**, *5*, 422-427.
122. Demeulemeester, J.; Chaltin, P.; Marchand, A.; De Maeyer, M.; Debyser, Z.; Christ, F. *Expert Opin. Ther. Pat.* **2014**, *24*, 609-632.
123. Fenwick, C.; Amad, M.; Bailey, M. D.; Bethell, R.; Bos, M.; Bonneau, P.; Cordingley, M.; Coulombe, R.; Duan, J.; Edwards, P.; Fader, L. D.; Faucher, A. M.; Garneau, M.; Jakalian, A.; Kawai, S.; Lamorte, L.; LaPlante, S.; Luo, L.; Mason, S.; Poupart, M.-A.; Rioux, N.; Schroeder, P.; Simoneau, B.; Tremblay, S.; Tsantrizos, Y.; Witvrouw, M.; Yoakim, C. *Antimicrob. Agents Chemother.* **2014**, *58*, 3233-3244, 13.
124. Li, W.; De Croos, P.; Fandrick, K. R.; Gao, J. J.; Haddad, N.; Lu, Z.-H.; Qu, B.; Rodriguez, S.; Senanayake, C. H.; Zhang, Y.; Tang, W. Patent WO2012138670, 2012.
125. Naidu, B. N.; Patel, M. Patent WO2013025584, 2013.

CONFIDENTIAL – Property of GSK – Do Not Copy

126. Pendri, A.; Li, G.; Gerritz, S.; Langley, D. R.; Trainor, G. L.; Meanwell, N. A. Patent WO2012033735, 2012.
127. Zheng, Z. B.; D'Andrea, S.; Langley, D. R.; Naidu, B. N. Patent WO2013134142, 2013.
128. Babaoglu, K.; Bjornson, K. L.; Hrvatin, P.; Lansdon, E.; Link, J. O.; Liu, H.; McFadden, R.; Mitchell, M. L.; Qi, Y.; Roethle, P. A.; Xu, L. Patent WO2013103724, 2013.
129. De La Rosa, M. A.; Haydar, S. N.; Johns, B. A.; Velthuisen, E. J. Patent WO2012102985, 2012.
130. Babaoglu, K.; Brizgys, G.; Cha, J.; Chen, X.; Guo, H.; Halcomb, R. L.; Han, X.; Huang, R.; Liu, H.; McFadden, R.; Mitchell, M. L.; Qi, Y.; Roethle, P. A.; Xu, L.; Yang, H. Patent WO2013159064, 2013.
131. Babaoglu, K.; Bjornson, K.; Guo, H.; Halcomb, R. L.; Link, J. O.; Liu, H.; Mitchell, M. L.; Sun, J.; Vivian, R. W.; Xu, L. Patent WO2012003498, 2012.
132. Babaoglu, K.; Brizgys, G.; Guo, H.; Hrvatin, P.; Lansdon, E.; Link, J. O.; Liu, H.; Mcfadden, R.; Mitchell, M. L.; Qi, Y.; Roethle, P. A.; Vivian, R. W.; Xu, L.; Yang, H. Patent WO2013103738, 2013.
133. Walker, M. A.; Sorenson, M. E.; Naidu, B. N.; Peese, K.; Johnson, B. L.; Patel, M. Patent WO2013123148, 2013.
134. Chaltin, P.; Christ, F.; Debyser, Z.; De Maeyer, M.; Marchand, A.; Marchand, D.; Voet, A. Patent WO2012065963, 2012.
135. Bardiot, D.; Chaltin, P.; Christ, F.; Debyser, Z.; De Maeyer, M.; Marchand, A.; Marchand, D.; Voet, A. Patent WO2010130842, 2010.
136. Tsantrizos, Y. S.; Boes, M.; Brochu, C.; Fenwick, C.; Malenfant, E.; Mason, S.; Pesant, M. Patent WO2007131350, 2007.

CONFIDENTIAL – Property of GSK – Do Not Copy

137. Tsantrizos, Y. S.; Bailey, M. D.; Bhardwaj, P. K.; Brochu, C.; Edwards, P. J.; Fader, L.; Jakalian, A.; Kawai, S.; Parisien, M.; Poupart, M.-A.; Simoneau, B. Patent WO2009062308, 2009.
138. Carson, R. J.; Fader, L.; Kawai, S.; Landry, S.; Tsantrizos, Y. S.; Brochu, C.; Morin, S.; Parisien, M.; Simoneau, B. Patent WO2009062288, 2009.
139. Li, Z. J.; Li, Z.; Luo, L.; Offerdahl, T.; Kim, S.; Yang, B.-S. Patent WO2012138669, 2012.
140. Tsantrizos, Y. S.; Bailey, M. D.; Bilodeau, F.; Carson, R. J.; Coulombe, R.; Fader, L.; Halmos, T.; Kawai, S.; Landry, S.; Laplante, S.; Morin, S.; Parisien, M.; Poupart, M.-A.; Simoneau, B. Patent WO2009062285, 2009.
141. Carlens, G.; Chaltin, P.; Christ, F.; Debyser, Z.; Marchand, A.; Marchand, D.; Voet, A.; De Maeyer, M. Patent WO2011076765, 2011.
142. Hattori, K.; Kurihara, N.; Iwaki, T.; Inoue, T.; Akiyama, T.; Hasegawa, Y. Patent WO2013002357, 2013.
143. Hattori, K.; Kurihara, N.; Iwaki, T.; Inoue, T.; Akiyama, T.; Hasegawa, Y. Patent WO2013002357, 2012.
144. Leach, A. GlaxoSmithKline, *Unpublished results*.
145. Green, D. GlaxoSmithKline, *Unpublished results*.
146. Hann, M.; Pickett, S.; Churcher, I. GlaxoSmithKline, *Unpublished results*.
147. Briand, J. GlaxoSmithKline, *Unpublished results*.
148. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug. Deliv. Rev.* **2001**, *46*, 3-26.
149. Spitzfaden, C. GlaxoSmithKline, *Unpublished results*.
150. Nolte, R.; Wang, L. GlaxoSmithKline, *Unpublished results*.

CONFIDENTIAL – Property of GSK – Do Not Copy

151. Consoler, T. GlaxoSmithKline, *Unpublished results*.
152. Huddler, D. GlaxoSmithKline, *Unpublished results*.
153. Leesnitzer, L. GlaxoSmithKline, *Unpublished results*.
154. White, G.; Williams, S. GlaxoSmithKline, *Unpublished results*.
155. Baldwin, I.; Fillmore, M.; Mitchel, C. GlaxoSmithKline, *Unpublished results*.
156. Baldwin, I.; Fillmore, M.; White, G. GlaxoSmithKline, *Unpublished results*.
157. Atkinson, P. J.; Bromidge, S. M.; Duxon, M. S.; Gaster, L. M.; Hadley, M. S.; Hammond, B.; Johnson, C. N.; Middlemiss, D. N.; North, S. E.; Price, G. W.; Rami, H. K.; Riley, G. J.; Scott, C. M.; Shaw, T. E.; Starr, K. R.; Stemp, G.; Thewlis, K. M.; Thomas, D. R.; Thompson, M.; Vong Antonio, K. K.; Watson, J. M. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 737-741.
158. Lovell, P. J.; Blaney, F. E.; Goodacre, C. J.; Scott, C. M.; Smith, P. W.; Starr, K. R.; Thewlis, K. M.; Vong, A. K. K.; Ward, S. E.; Watson, J. M. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1033-1036.
159. Hall, H. K., Jr. *J. Am. Chem. Soc.* **1957**, *79*, 5441-5444.
160. Bromidge, S. M.; Bertani, B.; Borriello, M.; Faedo, S.; Gordon, L. J.; Granci, E.; Hill, M.; Marshall, H. R.; Stasi, L. P.; Zucchelli, V.; Merlo, G.; Vesentini, A.; Watson, J. M.; Zonzini, L. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5653-5656.
161. Trost, B. M.; Fandrick, D. R. *Org. Lett.* **2005**, *7*, 823-826.
162. Krasovskiy, A.; Knochel, P. *Angew. Chem. Int. Ed.* **2004**, *43*, 3333-3336.
163. Imamoto, T.; Takiyama, N.; Nakamura, K.; Hatajima, T.; Kamiya, Y. *J. Am. Chem. Soc.* **1989**, *111*, 4392-4398.

CONFIDENTIAL – Property of GSK – Do Not Copy

164. Conlon, D. A.; Kumke, D.; Moeder, C.; Hardiman, M.; Hutson, G.; Sailer, L. *Adv. Synth. Catal.* **2004**, *346*, 1307-1315.
165. Wall, I. GlaxoSmithKline, *Unpublished results*.
166. Young, R.; Gummer, L. GlaxoSmithKline, *Unpublished results*.
167. Reynisson, J.; McDonald, E. *J. Comput. -Aided Mol. Des.* **2004**, *18*, 421-431.
168. Huang, M.-Z.; Luo, F.-X.; Mo, H.-B.; Ren, Y.-G.; Wang, X.-G.; Ou, X.-M.; Lei, M.-X.; Liu, A.-P.; Huang, L.; Xu, M.-C. *J. Agric. Food Chem.* **2009**, *57*, 9585-9592.
169. Feng, E.; Huang, H.; Zhou, Y.; Ye, D.; Jiang, H.; Liu, H. *J. Org. Chem.* **2009**, *74*, 2846-2849.
170. Kitamura, Y.; Hashimoto, A.; Yoshikawa, S.; Odaira, J. i.; Furuta, T.; Kan, T.; Tanaka, K. *Synlett* **2006**, 115-117.
171. Baldwin, I. GlaxoSmithKline, *Unpublished results*.
172. Fillmore, M. GlaxoSmithKline, *Unpublished results*.
173. Sloan, L. GlaxoSmithKline, *Unpublished results*.
174. Blackburn, T.; Ramtohul, Y. K. *Synlett* **2008**, 1159-1164.
175. Mohtat, B.; Jabbar, S.; Ghasemi, A.; Yavari, I. *J. Chem. Res.* **2008**, 601-603.
176. Fu, Z.; Li, X.; Merz, K. M. *J. Comput. Chem.* **2011**, *32*, 2587-2597.
177. De Luca, L.; Ferro, S.; Gitto, R.; Barreca, M. L.; Agnello, S.; Christ, F.; Debyser, Z.; Chimirri, A. *Bioorg. Med. Chem.* **2010**, *18*, 7515-7521.
178. Wang, T. GlaxoSmithKline, *Unpublished results*.
179. Goodford, P. J. *J. Med. Chem.* **1985**, *28*, 849.

CONFIDENTIAL – Property of GSK – Do Not Copy

180. von Itzstein, M.; Wu, W.-Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Phan Tho Van; Smythe, M. L.; White, H. F.; Oliver, S. W.; Colman, P. M.; Varghese, J. N.; Ryan, D. M.; Woods, J. M.; Bethell, R. C.; Hotham, V. J.; Cameron, J. M.; Penn, C. R. *Nature* **1993**, *363*, 418-423.
181. Bhan, A. K.; Reinherz, E. L.; Poppema, S.; McCluskey, R. T.; Schlossman, S. F. *J. Exp. Med.* **1980**, *152*, 771-782.
182. Chini, E. N. *Curr. Pharm. Des.* **2009**, *15*, 57-63.
183. Malavasi, F.; Deaglio, S.; Funaro, A.; Ferrero, E.; Horenstein, A. L.; Ortolan, E.; Vaisitti, T.; Aydin, S. *Physiol. Rev.* **2008**, *88*, 841-886.
184. Graeff, R.; Liu, Q.; Kriksunov, I. A.; Kotaka, M.; Oppenheimer, N.; Hao, Q.; Lee, H. C. *J. Biol. Chem.* **2009**, *284*, 27629-27636.
185. Egea, P. F.; Muller-Steffner, H.; Kuhn, I.; Cakir-Kiefer, C.; Oppenheimer, N. J.; Stroud, R. M.; Kellenberger, E.; Schuber, F. *PLoS One* **2012**, *7*, e34918.
186. D'Arena, G.; Nunziata, G.; Coppola, G.; Vigliotti, M. L.; Tartarone, A.; Carpinelli, N.; Matera, R.; Bisogno, R. C.; Pistolese, G.; Di Renzo, N. *Blood* **2002**, *100*, 3052-3053.
187. Graeff, R.; Liu, Q.; Kriksunov, I. A.; Hao, Q.; Lee, H. C. *J. Biol. Chem.* **2006**, *281*, 28951-28957.
188. Guse, A. H. *J. Mol. Med. (Berlin)* **2000**, *78*, 26-35.
189. Liang, M.; Chini, E. N.; Cheng, J.; Dousa, T. P. *Arch. Biochem. Biophys.* **1999**, *371*, 317-325.
190. Lee, H. C. *Biol. Chem.* **1999**, *380*, 785-793.
191. Chini, E. N.; Dousa, T. P. *Biochem. Biophys. Res. Commun.* **1995**, *209*, 167-174.

CONFIDENTIAL – Property of GSK – Do Not Copy

192. Schmid, F.; Bruhn, S.; Weber, K.; Mittruecker, H. W.; Guse, A. H. *FEBS Lett.* **2011**, 585, 3544-3548.
193. Hoshino, S.-I.; Kukimoto, I.; Kontani, K.; Inoue, S.-I.; Kanda, Y.; Malavasi, F.; Katada, T. *J. Immunol.* **1997**, 158, 741-747.
194. De Flora, A.; Guida, L.; Franco, L.; Zocchi, E.; Pestarino, M.; Usai, C.; Marchetti, C.; Fedeles, E.; Fontana, G.; Raiteri, M. *Biochem. J.* **1996**, 320, 665-672.
195. Liu, Q.; Graeff, R.; Kriksunov, I. A.; Jiang, H.; Zhang, B.; Oppenheimer, N.; Lin, H.; Potter, B. V. L.; Lee, H. C.; Hao, Q. *J. Biol. Chem.* **2009**, 284, 27637-27645.
196. Preugschat, F. GlaxoSmithKline, *Unpublished results*.
197. Schewchuk-Chapman, L. GlaxoSmithKline, *Unpublished results*.
198. Matsui, K.; Ezoe, S.; Oritani, K.; Shibata, M.; Tokunaga, M.; Fujita, N.; Tanimura, A.; Sudo, T.; Tanaka, H.; McBurney, M. W.; Matsumura, I.; Kanakura, Y. *Biochem. Biophys. Res. Commun.* **2012**, 418, 811-817.
199. Hersher, R. *Nat. Med. (N. Y. , NY, U. S.)* **2012**, 18, 1446.
200. Damle, R. N.; Wasil, T.; Fais, F.; Ghiotto, F.; Valetto, A.; Allen, S. L.; Buchbinder, A.; Budman, D.; Dittmar, K.; Kolitz, J.; Lichtman, S. M.; Schulman, P.; Vinciguerra, V. P.; Rai, K. R.; Ferrarini, M.; Chiorazzi, N. *Blood* **1999**, 94, 1840-1847.
201. Yagi, K.; Kanatsuka, A. *Naibunpi, Tonyobyoka* **1998**, 6, 74-77.
202. Chini, E. N.; Barbosa, M. T. P.; Soares, S. M. Patent US20090196825, 2009.
203. Aksoy, P.; Escande, C.; White, T. A.; Thompson, M.; Soares, S.; Benech, J. C.; Chini, E. N. *Biochem. Biophys. Res. Commun.* **2006**, 349, 353-359.

CONFIDENTIAL – Property of GSK – Do Not Copy

204. Escande, C.; Nin, V.; Price, N. L.; Capellini, V.; Gomes, A. P.; Barbosa, M. T.; O'Neil, L.; White, T. A.; Sinclair, D. A.; Chini, E. N. *Diabetes* **2013**, *62*, 1084-1093.
205. Barbosa, M. T.; Soares, S. M.; Novak, C. M.; Sinclair, D.; Levine, J. A.; Aksoy, P.; Chini, E. N. *FASEB J.* **2007**, *21*, 3629-3639, 10.
206. Kellenberger, E.; Kuhn, I.; Schuber, F.; Muller-Steffner, H. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 3939-3942.
207. Sinclair, D. A.; Price, N. L.; Chini, E.; Clardy, J. C.; Cao, S. Patent WO2013002879, 2013.
208. Haffner, C. GlaxoSmithKline, *Unpublished results*.
209. Chen, S. GlaxoSmithKline, *Unpublished results*.
210. Felitsky, D. GlaxoSmithKline, *Unpublished results*.
211. Fillmore, M.; Young, R. GlaxoSmithKline, *Unpublished results*.
212. Stuart, D. GlaxoSmithKline, *Unpublished results*.
213. Vimal, M. GlaxoSmithKline, *Unpublished results*.
214. Hughes, J. P.; Rees, S.; Kalindjian, S. B.; Philpott, K. L. *Br. J. Pharmacol.* **2011**, *162*, 1239-1249.
215. Kranz, M. GlaxoSmithKline, *Unpublished results*.
216. Turner, D. R.; Edwards, A. J.; Piltz, R. O. *CrystEngComm* **2012**, *14*, 6447-6451.
217. McDowell, C. A. *Nature (London, U. K.)* **1947**, *159*, 508-509.
218. Walsh, A. D. *Trans. Faraday Soc.* **1949**, *45*, 179-190.

CONFIDENTIAL – Property of GSK – Do Not Copy

219. Horner, K. E.; Karadakov, P. B. *J. Org. Chem.* **2013**, *78*, 8037-8043.
220. Humphreys, P. GlaxoSmithKline, *Unpublished results*.
221. Martin, R.; Buchwald, S. L. *Acc. Chem. Res.* **2008**, *41*, 1461-1473.
222. Hoffmann, M.; Kumar, G.; Schafer, P.; Cedzik, D.; Capone, L.; Fong, K.-L.; Gu, Z.; Heller, D.; Feng, H.; Surapaneni, S.; Laskin, O.; Wu, A. *Xenobiotica* **2011**, *41*, 1063-1075.
223. Brooks, D. W.; Lu, L. D. L.; Masamune, S. *Angew. Chem.* **1979**, *91*, 76-77.
224. Laurence, C.; Brameld, K. A.; Graton, J.; Le Questel, J. Y.; Renault, E. *J. Med. Chem.* **2009**, *52*, 4073-4086.
225. Freeman, D. GlaxoSmithKline, *Unpublished results*.
226. Redman, J. GlaxoSmithKline, *Unpublished results*.
227. Ironmonger, A. GlaxoSmithKline, *Unpublished results*.
228. Glauder, J. *Spec. Chem. Mag.* **2004**, *24*, 30-31.
229. Garcia, A. L. L. *Synlett* **2007**, 1328-1329.
230. Dunetz, J. R.; Xiang, Y.; Baldwin, A.; Ringling, J. *Org Lett* **2011**, *13*, 5048-5051.
231. http://www.euticals.com/attachments/075_EUTICALS_T3P-Coupling-Agents_2012_final_web.pdf **26/01/15**.
232. Whitehurst, B. GlaxoSmithKline, *Unpublished results*.
233. www.biotage.com **26/01/15**.
234. www.thalesnano.com/products/h-cube **26/01/15**.
235. www.stuart-equipment.com **26/01/15**.

8 Appendices

FDA Approval	Brand Name	Generic Name	Manufacturer
<i>Nucleoside Reverse Transcriptase Inhibitors (NRTIs)</i>			
1987	Retrovir	zidovudine, azidothymidine, AZT, ZDV	GlaxoSmithKline
1991	Videx	didanosine, dideoxyinosine, ddI	Bristol-Myers Squibb
1992	Hivid	zalcitabine, dideoxycytidine, ddC	Roche Pharmaceuticals
1994	Zerit	stavudine, d4T	Bristol-Myers Squibb
1995	Epivir	lamivudine, 3TC	GlaxoSmithKline
1997	Combivir	lamivudine + zidovudine	GlaxoSmithKline
1998	Ziagen	abacavir sulfate, ABC	GlaxoSmithKline
2000	Trizivir	abacavir + lamivudine + zidovudine	GlaxoSmithKline
2000	Videx EC	enteric coated didanosine, ddI EC	Bristol-Myers Squibb
2001	Viread	tenofovir disoproxil fumarate, TDF	Gilead Sciences
2001	Edurant	rilpivirine	Tibate Therapeutics
2003	Emtriva	emtricitabine, FTC	Gilead Sciences
2004	Epzicom	abacavir + lamivudine	GlaxoSmithKline
2004	Truvada	emtricitabine + tenofovir disoproxil fumarate	Gilead Sciences
2011	Viramure XR	nevirapine, NVP	Boehringer Ingelheim
<i>Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs)</i>			
1996	Viramune	nevirapine, NVP	Boehringer Ingelheim
1997	Rescriptor	delavirdine, DLV	Pfizer
1998	Sustiva	efavirenz, EFV	Bristol-Myers Squibb
2008	Intelence	etravirine	Tibotec Therapeutics
<i>Protease Inhibitors (PIs)</i>			
1995	Invirase	saquinavir mesylate, SQV	Roche Pharmaceuticals
1996	Norvir	ritonavir, RTV	Abbott Laboratories
1996	Crixivan	indinavir, IDV	Merck
1997	Viracept	nelfinavir mesylate, NFV	Pfizer
1997	Fortovase	saquinavir (no longer marketed)	Roche Pharmaceuticals
1999	Agenerase	amprenavir, APV	GlaxoSmithKline
2000	Kaletra	lopinavir + ritonavir, LPV/RTV	Abbott Laboratories
2003	Reyataz	atazanavir sulfate, ATV	Bristol-Myers Squibb
2003	Lexiva	fosamprenavir calcium, FOS-APV	GlaxoSmithKline
2005	Aptivus	triplanavir, TPV	Boehringer Ingelheim
2006	Prezista	darunavir	Tibotec Therapeutics

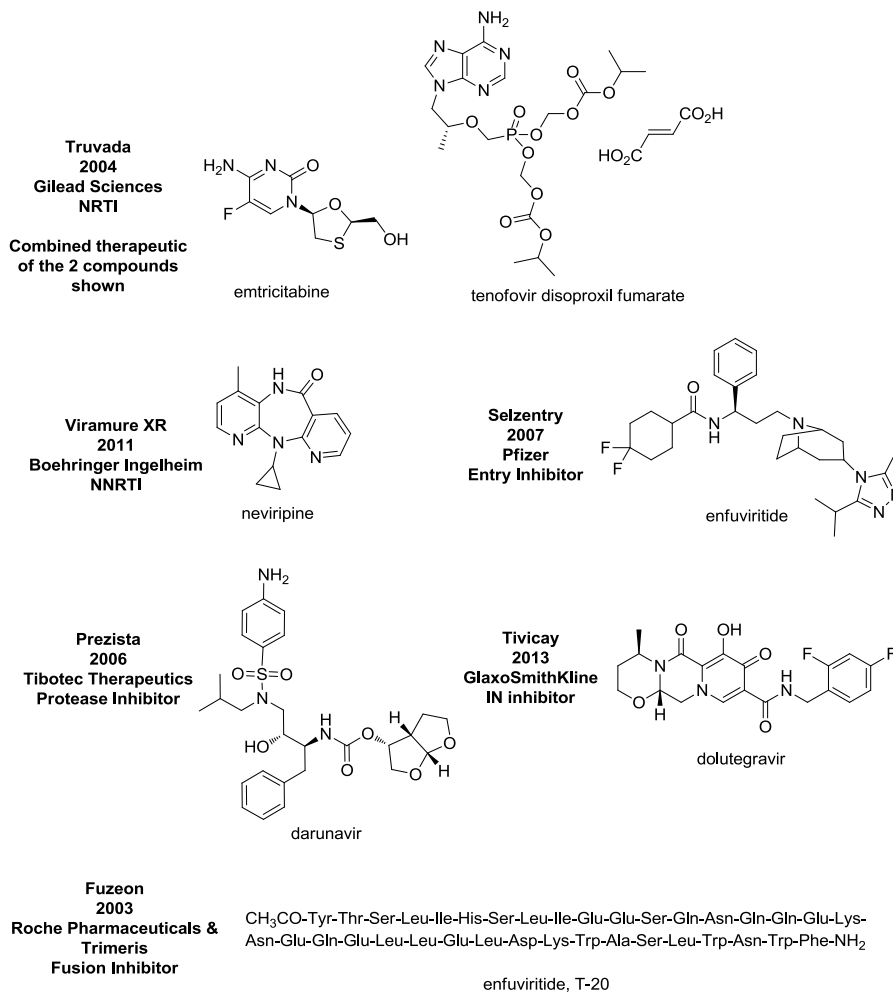
Table continued on next page...

CONFIDENTIAL – Property of GSK – Do Not Copy

FDA Approval	Brand Name	Generic Name	Manufacturer
<i>Fusion Inhibitors</i>			
2003	Fuzeon	enfuvirtide, T-20	Roche Pharmaceuticals & Trimeris
<i>Entry Inhibitors</i>			
2007	Selzentry	maraviroc	Pfizer
<i>IN Inhibitors</i>			
2007	Isentress	raltegravir	Merck
2013	Tivicay	dolutegravir	GlaxoSmithKline
<i>Multi-class Combination Products (HAART)</i>			
2006	Atripla	etavirenz, emtricitabine, tenofovir disoproxil fumarate	Bristol-Myers Squibb Gilead Sciences
2011	Complera	emtricitabine, rilpivine, tenofovir disoproxil fumarate	Gilead Sciences
2012	Stribild	elvitegravir, cobicistat, emtricitabine, tenofovir disoproxil fumarate	Gilead Sciences

Appendix i – FDA approved HIV therapies – updated August 2014.⁹³

CONFIDENTIAL – Property of GSK – Do Not Copy



Appendix ii – Chemical structures of an approved HIV therapy from each of the target classes.