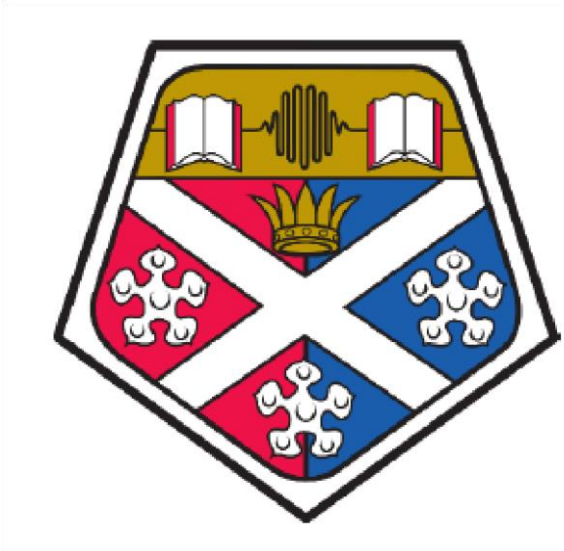
University of Strathclyde
Department of Pure and Applied Chemistry



Developing Chemical Tools for the Cancer Kinome: Targeting DYRK2 in Triple Negative Breast Cancer

A thesis submitted to the University of Strathclyde in part fulfilment of regulations for the degree of Doctor of Philosophy in Chemistry.

Laura Mary Bain

Supervised by Professor Nicholas C. O. Tomkinson

2020

Flyleaf

This thesis is the results 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.

Signed:  Date: 25.09.2020

(Laura Bain)

Acknowledgements

First, I would like to thank my supervisors, Nick Tomkinson and Simon MacKay. I am extremely grateful for their guidance and insight throughout the course of my PhD. And, I am very thankful for their support through this unforeseen crazy year.

Thank you to the members of the Tomkinson group, past and present, for making my time in the group less stressful and mostly, enjoyable. A special thanks must go to Jayde and Steven, who are now two of my closest friends. They have helped keep me sane in the strangest of times.

A special thank you to Louise Young for teaching me the tricks of the biological trade and allowing me to work in her lab to carry out the *in vitro* assay which helped push the project forward. I would also like to extend a thanks to Grainne Abbott and Dr Inga Kruse for welcoming me into the lab and helping with any and all queries. Thank you to Dr Chris Lawson for being so kind as to provide me with some interesting intermediates and teaching me how to use the MacKay H-cube.

Thank you to our collaborators, Professor Joanne Edwards and Dr Laureano de la Vega. Their hard work, enthusiasm and expertise helped drive the project in a very exciting direction.

I would like to extend a special thanks to all of the technical staff in the Thomas Graham building. Thank you to Craig Irving, whose NMR expertise and kindness has been most appreciated throughout my time at Strathclyde. Thank you to Patricia Keating for her effortless help with mass spectrometry. And thank you to Gavin Bain, who went over and above to help with day to day queries and challenges.

Finally, I would like to thank my friends and family. Adele, you shared this experience with me from start to finish, thank you for always having time to listen to me rant and rave. Alana, you are my number one supporter no matter what, I can't thank you enough. Nicole, thank you for always reminding me to have fun! Stuart, thank you for dealing with any and all (definitely not) irrational strops, you're some man. And thank you, most of all, to my family for their unconditional love and support, even in different time zones, throughout my time at Strathclyde and my chemistry journey so far.

Abstract

Protein kinases are one of the largest protein families in the human genome under investigation as therapeutic targets for their involvement in cancer. Their role, phosphorylation, is an essential post-translational modification required for normal function. DYRK2 (Dual-Specificity Tyrosine(Y) Regulated Kinase-2) has been reported to play a promiscuous role in cancer. However, research has revealed that DYRK2 plays an important role in activating the proteasome. In addition, *in vivo* studies have shown that DYRK2 inhibition impedes tumour growth in Triple Negative Breast Cancer (TNBC) mice xenografts. Thus, we were particularly interested in validating DYRK2 as a target for the treatment of TNBC. TNBC is an aggressive, heterogeneous disease and accounts for 10–20% of breast cancer related deaths in women. It is named due to its lack of three key receptors (ER, PR and HER2) and as a result the current treatment options, chemotherapy, surgery and radiotherapy are also aggressive in nature. Thus, there is a clinical demand for the identification of alternative biological targets for TNBC therapy.

Through collaboration with Professor Joanne Edwards at the University of Glasgow and Dr Laureano de la Vega at the University of Dundee, we have reinforced the synergistic relationship between DYRK2 and TNBC. We discovered that DYRK2 expression negatively affects TNBC patient survival and is essential for cancer cell growth *via* the phosphorylation of Heat Shock Factor-1 (HSF1), the master regulator of proteotoxic stress pathways.

This work allowed us to expand upon the SAR of DYRK2 inhibitors and discover a DYRK2 inhibitor, **68** with >10-fold increase in potency compared to the previous lead **CI709** (**68** K_i 19 nM). Pleasingly, **68** is selective for DYRK2 against its closely related isoform DYRK1A (IC_{50} >10 µM) *in vitro.*

68 and closely related analogue **CI709** were tested in TNBC cell models and have been shown to impede cancer cell growth in the MDA-MB-468 cell line. In addition, **68** and **CI709** inhibit the phosphorylation of HSF1 at two significant sites, Ser320 and Ser326, in TNBC cells.

These findings strengthen the reasons for the validation of DYRK2 as a target for investigation in TNBC treatment.

Contents

Abbreviations – Biological Terms

AA -	Amino acid
ADP -	Adenosine di-phosphate
AGC -	Protein kinases A, G and C
AI -	Aromatase inhibitor
AR -	Androgen receptor
ATM -	Ataxia telangiectasia-mutated
ATP -	Adenosine tri-phosphate
BC -	Breast cancer
BCS -	Breast conserving surgery
CAMK -	Calcium/calmodulin-dependant kinase
CDK -	Cyclin dependant kinase
CK1 -	Casein kinase 1
CLK -	CDC2-like kinase
CMGC -	CDK, MAPK, GSK and CLK
CP -	Core particle
CTAD -	C-terminal transactivation domain
CYP450 -	Cytochrome P450
DBD -	DNA-binding domain
DBS -	Double strand breaks
DFS -	Disease free survival
DSF -	Differential scanning fluorimetry
DYRK -	Dual specificity tyrosine(Y) regulated kinase
EGFR -	Epidermal growth factor receptor
ER -	Estrogen receptor
GK -	Gatekeeper
GSK -	Glycogen synthase kinase

HBA - Hydrogen bond acceptor

HBD - Hydrogen bond donor

HDAC - Histone deacetylase

HER - Human epidermal growth factor receptor

HR - Hormone receptor

HS - Heat shock

HSE - Heat shock elements

HSF - Heat shock factor

HSP - Heat shock protein

IDC - Invasive ductal carcinoma

ILC - Invasive lobular carcinoma

KD - Knockdown

KO - Knockout

LAR - Luminal androgen receptor

LZ - Leucine zipper

MAPK - Mitogen-activated protein kinase

mTOR - Mammalian target of rapamycin

OS - Overall survival

PARP - Poly(ADP)ribose polymerase

P-gp - P-glycoprotein

PI3K - Phosphotidylinositol-3-kinase

PR - Progesterone receptor

PTM - Post-translational modification

RGC - Receptor guanylyl cyclase

RD - Regulatory domain

RLU - Relative luminescence units

RTK - Receptor tyrosine kinase

RP - Regulatory particle

SCF - SKp- cullin 1-F-box

SIAH -	Seven in absentia homolog
STE -	Sterile serine/threonine kinases
STK -	Serine/threonine kinase
SSB -	Single strand breaks
TK -	Tyrosine kinase
TKL -	Tyrosine kinase-like
TKR -	Tyrosine kinase receptor
TN -	Triple negative
Ub -	Ubiquitin
VEGFR -	Vascular endothelial growth factor receptor kinase
WT -	Wild type

Abbreviations – Chemistry and Miscellaneous Terms

Ac -	Acyl
Bn -	Benzyl
CDI -	1,1-Carbonyldiimidazole
DCE -	Dichloroethane
DMAP -	*N,N*-dimethylamino pyridine
DMF -	*N,N*-dimethylformamide
DMSO -	Dimethyl sulfoxide
ES -	Excited state
EWG -	Electron withdrawing group
GS -	Ground state
HATU -	Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium
HetAr -	Heteroaromatic
LT -	Life Technologies
MW -	Molecular weight
NCS -	*N*-chlorosuccinimide
NIS -	*N*-iodosuccinimide
NMR -	Nuclear magnetic resonance
PD -	Pharmacodynamics
PK -	Pharmacokinetics
SAR -	Structure activity relationship
SGC -	Structural Genomics Consortium
STAB -	Sodium triacetoxyborohydride
SIPBS -	Strathclyde Institute for Pharmacy and Biomedical Sciences
TBAB -	Tetrabutylammonium bromide
TFA -	Trifluoroacetic acid
THF -	Tetrahydrofuran
TLC -	Thin layer chromatography

Ts - Tosyl

1 Introduction

Proteins are essential to life. 24 hours a day, 7 days a week, a great number of proteins within our body carry out functions crucial for our survival. However, occasionally, some of these proteins either form incorrectly, mutate or become unfit to carry out their role. This unfortunately can have detrimental effects on our lives. With the study of protein structure and function, scientists aim to gain an insight into the precise mechanisms by which these aberrations result in adverse physiological changes and provide methods to restore normal function.

An interesting class of proteins which has been rigorously investigated are the protein kinases. Within the last 18 months there have been >1200 publications in this area in medicinal chemistry journals.[1-4] The next section will introduce this important protein family and explain why they are of high therapeutic value.

1.1 Kinases

Protein kinases are one of the largest studied classes of proteins. To date, there are 518 reported protein kinases. These are divided into 9 main families: AGC, CAMK, CK1, CMGC, STE, TK, TKL, RCG and finally a diverse group of kinases termed 'other'.[5,6] These families are named according to the structure/function of their members and are divided into smaller sub-families which are grouped according to increasing sequence and structural similarities.[6] The set of protein kinases present in the human genome can be depicted by a phylogenetic tree, which is also known as the kinome, **Figure 1.**

AGC - Protein kinases A, G and C
CAMK - Ca^{2+} dependent kinases
CK1 - Casein kinase 1
CMGC - CDKs, MAPKs, GSKs and CLKs
RCG - Receptor guanylate cyclases (not depicted)
STE - Sterile Ser/Thr kinases
TK - Tyrosine kinases
TKL - Tyrosine kinase-like

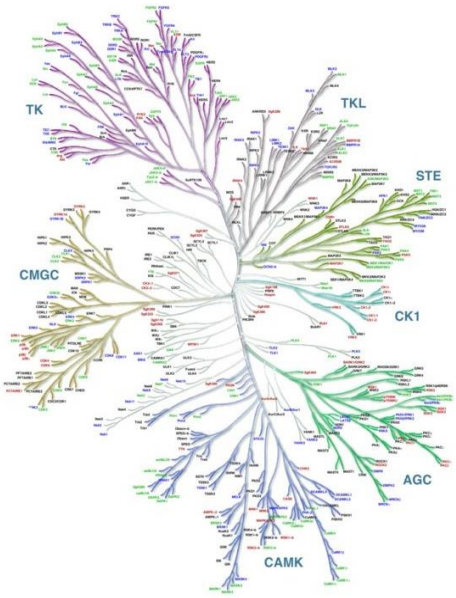


Figure 1: Phylogenetic tree of protein kinases in the human kinome.[7]

Members of each individual family, highlighted in **Figure 1** are largely grouped according to their structural and functional similarities. The structure of the kinase catalytic domain is generally highly conserved across the entire kinome. This is essential as they all perform the same function: phosphorylation.

1.1.1 Function

The role of protein kinases is to catalyse the phosphorylation process. Most commonly, protein kinases phosphorylate amino acid (AA) residues with a terminal hydroxyl group on their protein substrates. These residues include serine, threonine and tyrosine, **Figure 2**.

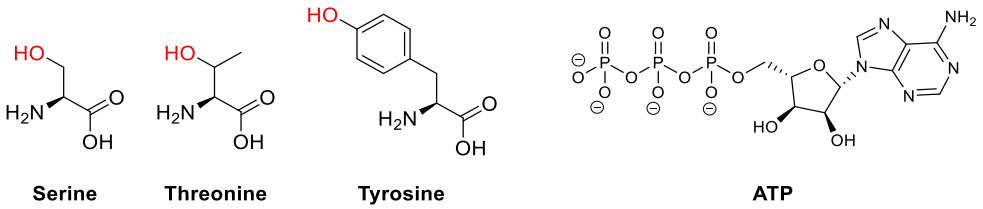


Figure 2: AA residues susceptible to phosphorylation and ATP structure.

Phosphorylation involves the transfer of a phosphate group from the co-factor adenosine triphosphate (ATP), **Figure 2**, to the hydroxyl of the protein substrate. More specifically, the co-factor ATP is bound within the protein kinase active site. This provides a specific orientation of the terminal phosphate of ATP for the transfer to

protein substrates.[8] The structure of protein kinase bound ATP will be discussed in more detail in the following sections. From a simplistic perspective, the phosphate is transferred from the kinase bound ATP to the corresponding hydroxyl bearing residue of its substrate and results in a phosphorylated and negatively charged product and a kinase bound adenosine di-phosphate (ADP) **Figure 3**.

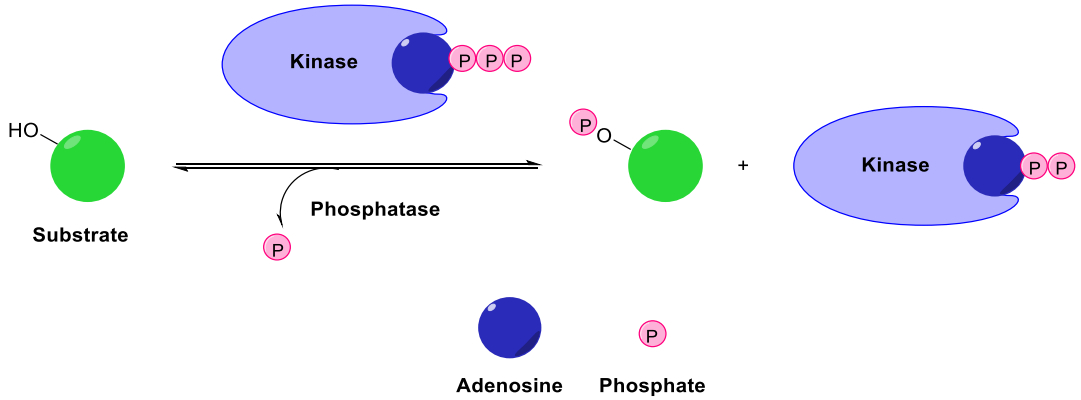


Figure 3: Schematic representation of phosphorylation.

However, the process is not this simple and is highly specific. Protein kinases have distinct protein targets and the interactions between kinase and substrate are driven by residues surrounding the kinase active site and also residues preceding the flanking Ser/Thre/Tyr of the corresponding substrates. This is often referred to as kinase peptide specificity.[9]

The process of phosphorylation is also a reversible process. The removal of a phosphate group is catalysed by a different class of proteins, the protein phosphatases.[10]

The addition or removal of a charged phosphate group can alter the structure of the protein substrate with the potential to influence its function, such as promoting or inhibiting cellular signalling processes that can lead to changes in gene expression. The process of phosphorylation/dephosphorylation is inherently important in cellular function. It is essential for activating signal transduction pathways which maintain homeostasis and in most cases, kinases themselves are activated *via* phosphorylation of the appropriate residues. From a statistical perspective, more than 14,000 of the 21,000 proteins encoded by the human genome are subjected to phosphorylation,[11] which highlights the fundamental importance of kinases.

Protein kinases can be divided into three categories. Firstly, serine/threonine phosphorylating kinases (STKs); these proteins are generally located intracellularly

and involved in but not exclusive to processes such as cell homeostasis and the regulation of the cell cycle.[12,13] Secondly, tyrosine phosphorylating kinases (TKs), these are often, but not exclusively, membrane receptors that mainly promote cellular processes such as cell proliferation, differentiation and migration.[13, 14] Finally, there are kinases which promote phosphorylation of all three residues (serine, threonine and tyrosine) and these are known as dual-specific kinases. These will be discussed in more detail in **Section 1.2**.

It is clear that the catalytic process of phosphorylation is essential for cell regulation. The role of kinases in this process can be explained in more detail through investigation of the kinase structure.

1.1.2 Structure

The general structure of the catalytic domain of protein kinases is well conserved throughout the kinome. The *N*-terminus comprises mainly of β-sheets and the *C*-terminus mostly of α-helices, **Figure 4**.[15]

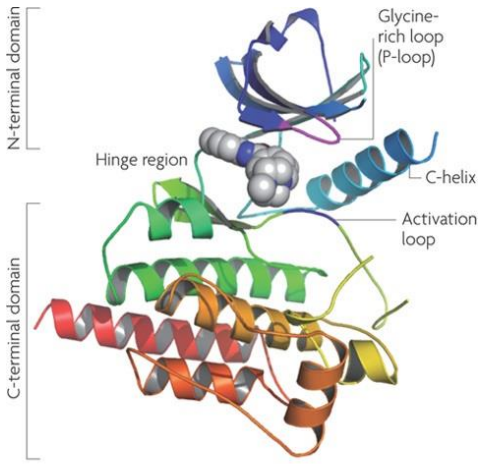


Figure 4: General kinase structure.[16]

These are joined by a motif known as the hinge region which provides the structures flexibility and in turn, helps facilitate phosphorylation by opening and closing to allow ADP to leave and ATP to re-engage after each event. Another important region within the kinase structure is the activation loop. The kinase itself is activated *via* phosphorylation of a residue within the activation loop. The activity of the kinase is governed by the position of the DFG motif. More specifically, the DFG motif is an acronym for the amino acid residues it represents: aspartic acid (D), phenylalanine (F) and glycine (G), respectively. When the DFG motif is 'out' or the activation loop is

closed, the kinase is inactive. When the activation loop is in an open conformation, also referred to as DFG 'in', the kinase is rendered active and can therefore catalyse the phosphorylation process.[17]

Most of the structural differences between the kinases are found in the *N*- and *C*-terminal domains. However, the structure of the ATP-binding domain or the active site is of most interest. This is where subtle differences in structure can aid in the search for selective inhibitors. However, first, we can appreciate the general structure of the ATP-binding site.

1.1.3 ATP binding mode

The structure of ATP is important for binding to the kinase active site and is subsequently essential for the phosphorylation process. We can divide the structure of ATP into three sections: the adenine ring, the ribose sugar and the phosphate groups, **Figure 5.**

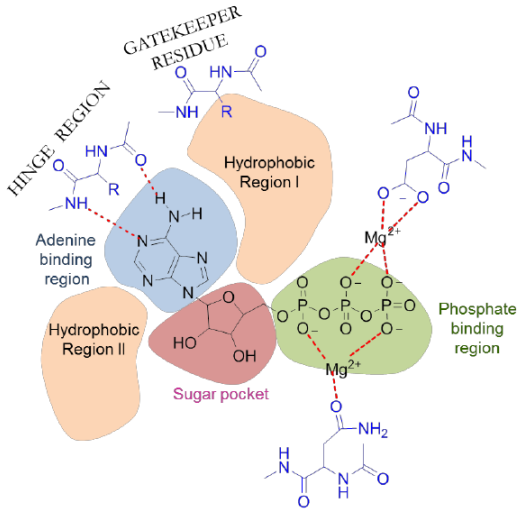


Figure 5: ATP binding mode.[18]

Directing our attention to the adenine ring, two hydrogen bond interactions occur with residues in the hinge region. These residues in this case are AAs adjacent to an important residue within the active site: the gatekeeper (GK). The GK has been described to allow access to the hydrophobic back pocket of the kinase active site and can often aid in selectivity.[19] **Figure 5** is a general representation of the binding interactions between ATP and the active site and therefore the specific AAs of the hinge and GK are not described. From a general perspective, the interactions between the adenine ring and the hinge region are as follows: a hydrogen bond donor

(HBD) of the amine group interacts with a neighbouring hydrogen bond acceptor (HBA) carbonyl on the hinge region (GK+2) and a HBA nitrogen atom, present on the purine scaffold interacts with a backbone amide HBD species (GK+3). Above and below the adenine ring are two hydrophobic regions. Adjoining the adenine ring and the tri-phosphate group is a ribose sugar which fits into a sugar pocket within the active site. Finally, the overall negative charge of the phosphate groups is balanced by Mg^{2+} counter ions present in the active site. The position of the phosphates in the active site is important as the terminal phosphate is directed slightly outside the pocket, primed for phosphorylation of substrate residues.

Kinases are attractive drug targets due to their importance in cell signalling pathways and they are extensively investigated as such. One of the most important starting points when designing kinase inhibitors as tools is the role of ATP mimetics. For example, exploiting the hydrogen bond interactions with the hinge region as described above. However, ATP mimetics are not the only type of kinase inhibitors under investigation: to date, four different types of inhibitor have been identified, **Figure 6**.

1.1.4 Kinase inhibitors

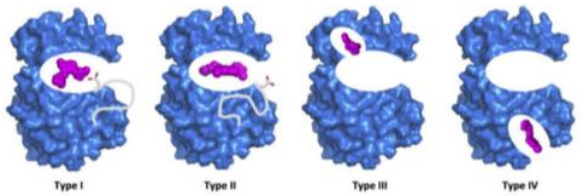
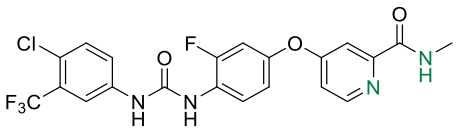


Figure 6: Types of kinase inhibitors.[20]

The inhibitors are categorised as follows: **Type I**: these substrates are ATP competitive. They bind in place of ATP within the active site of the kinase, when it is in its active form (DFG in). **Type II**: these inhibitors bind within the active site when the kinase is in the inactive form (DFG out). **Type III** inhibitors bind to a site adjacent to the active site and are therefore not ATP competitive. Finally, **type IV** inhibitors are the least established, where the molecules described bind to a site unrelated to the active site.[17] This type of interaction is known as allosteric inhibition.

The most common types of kinase inhibitors on the market are **type I**, for example regorafenib, **Figure 7**.



Regorafenib

Figure 7: Type I kinase inhibitor Regorafenib.

Similar to ATP, as discussed in the previous section, regorafenib will interact with the hinge region of the kinase target through hydrogen bond interactions, **Figure 7**. Regorafenib was developed by Bayer in 2012. It targets two different kinases: VEGFR-2 (vascular endothelial growth factor receptor kinase-2) and Tie2 or tyrosine protein kinase receptor. This inhibitor is used in the treatment of colorectal cancer and gastrointestinal stromal tumours.[19] This particular example highlights the selectivity challenge, albeit potentially positive in this instance, which is faced when designing kinase inhibitors.

1.1.5 Selective inhibitors

As briefly mentioned in the previous section, the most common kinase inhibitors on the market are ATP competitive. The ATP site is generally highly conserved among kinases, which poses a challenge for scientists in the design of selective inhibitors. Not only are there >500 protein kinases, but the fact that they are grouped into families with respect to their structural similarities, further decreases the likelihood of finding an inhibitor which will target only one member of a family. Therefore, one of the most important challenges to overcome in the discovery of new kinase inhibitors is selectivity. This is essential because with a lack of selectivity comes potentially negative side effects from inhibiting and/or activating an undesired target.

1.1.6 Kinases and cancer

One of the largest therapeutic areas where kinases are making an impact is oncology. In 2014, Cancer Research UK reported 356,860 cases of cancer, of which 163,444 resulted in death, which corresponds to a poor survival rate of 54%.[21] Furthermore, the most common types of cancer reported were: lung, bowel, breast and prostate cancer, with greater than a fifth of cases being a result of lung cancer.[21] Thus, it is clear that this evolutionary disease demands research attention into understanding its molecular biology to provide treatment options.

The major issue is that, cancer itself is the result of numerous changes within the body that arise as a result of irregularities at the cellular level. Alterations in processes such as cell growth, proliferation and cell death can lead to the uncontrolled growth of cancer cells and subsequently result in tumour formation and progession.[22]

As discussed in **Section 1.1.1**, one of the most important proteins responsible for the regulation of cell signalling pathways is the protein kinase. Thus, control of kinase activity is pivotal. There are a number of processes controlled by kinases, which when subject to abnormalities e.g. mutations or overexpression, promote the growth of cancer cells. One example involves the bodies' response to cell death. Once a cancerous cell is detected, our bodies' primary response would be to invoke cell death. One of the ways in which this is carried out is by receptor tyrosine kinases (RTKs), however, cancerous cells can become resistant to this process.[23] This becomes a serious problem, making it difficult to prevent the spread of cancerous cells and ultimately their potential growth to tumours.

A common feature of tumour cells is their inherent ability to impair the cell cycle of normal functioning cells and subsequently, the growth of healthy cells. This process, the regulation of the cell cycle, is mediated and controlled by another family of kinases: cyclin dependant kinases (CDK).[22]

This is just a glimpse into the importance of kinases in cancer but it is apparent that the control of normal kinase activity is crucial in order for us to tackle this challenging disease.

One kinase in particular, which has been reported to play a significant, yet promiscuous role in cancer, is dual specificity tyrosine(Y) phosphorylation-regulated kinase 2 (DYRK2).

1.2 DYRK2

1.2.1 Subfamily overview

DYRK2 is a member of the CMGC superfamily of protein kinases. This abbreviation stems from the names of the members within this superfamily: cyclin-dependent kinases (CDKs), mitogen-activated protein kinases (MAPKs), glycogen synthase kinases (GSKs) and CDC-like kinases (CDC-like), **Figure 8**.[24]

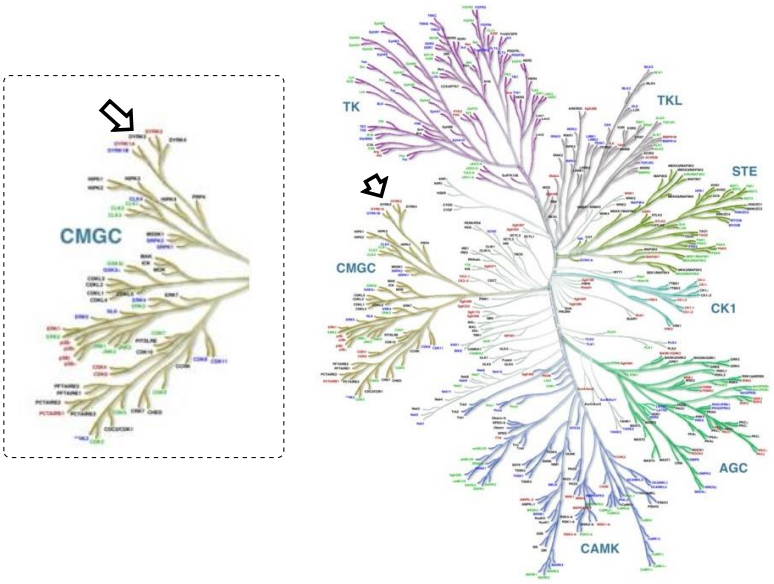


Figure 8: Location of DYRK2 in the human kinome.

DYRK (dual-specificity tyrosine phosphorylation-regulated kinase) is a subfamily within the CDK family and in humans, there are 5 members of the DYRK family: DYRK1A, DYRK1B, DYRK2, DYRK3 and DYRK4.[25] The term dual-specificity comes from the kinase function, where DYRKs auto-phosphorylate a tyrosine residue in their activation loop in order to become primed for phosphorylating their substrates, *via* serine and threonine residues.[26]

The 5 mammalian family members can be further split into two classes. Class I: DYRK1A and DYRK1B and class II: DYRK2, DYRK3 and DYRK4. These classes are divided and described due to their sequence homologies.[27] All DYRK family members share a conserved kinase domain and DYRK homology (DH)-box, **Figure 9**. The class I kinases share domains distinct to the kinase domain and these are the nuclear localisation signal (NLS) that contains a specific amino acid sequence for recognition and transport to the nucleus,[28] in the *N*-terminus and the PEST region, a region rich in proline (P), glutamate (E), serine (S) and threonine (T), situated in the *C*-terminus. The class II DYRKs contain a motif termed the *N*-terminal auto-phosphorylation accessory (NAPA), which is said to aid in the intramolecular auto-phosphorylation of tyrosine residues on the DYRK structures, **Figure 9**.[29]

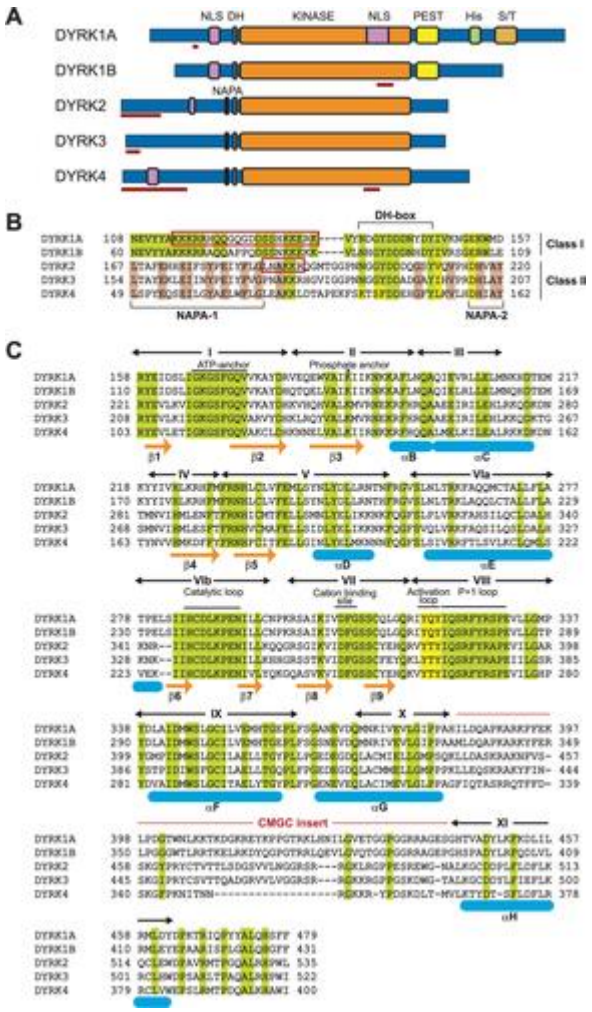


*Figure 9: Representation of the **A**) DYRK family protein structures **B**) the corresponding AA sequences of non-kinase domain regions **C**) the corresponding AA sequences of the kinase domain.*[30]

In **Section 1.1.3** we discussed the structure of the kinase ATP domain and highlighted that the protein kinase active site is highly conserved. Taking a closer look at the AA sequence alignment of the DYRK family kinases this is reinforced. A review by Aranda *et al* discussed the similarities in the kinase domain of the DYRK family kinases (**Figure 9**).[30] More specifically, the sequence homology is highlighted in green in **Figure 9**. From this, it is apparent that the structure of the DYRK kinase domain is, unsurpisingly, highly conserved. This is important to note as it poses a selectivity challenge among the DYRK isoforms when designing DYRK inhibitors. This will be discussed in more detail in a later section.

With respect to the importance of the DYRK family of kinases, this can be discussed briefly as we are primarily interested in DYRK2. The class I DYRKs are primarily nuclear, however, there have been reports of cytosolic DYRK1A and cytoplasmic

DYRK1B.[27] Moreover, DYRK1A has been investigated extensively as a potential drug target, as it has been reported to be a key protein related to the features of Down's syndrome.[31] DYRK1B is overexpressed in some pancreatic cancers and its depletion has been reported to sensitise the cells to apoptosis.[32] Of the class II DYRKs, there is less known of DYRK3 and DYRK4 compared to DYRK2. DYRK3 has been identified as a protein involved in the mediation of endocytosis.[33] The role of DYRK4 is even less known, however, its expression has been localised to the testes and may be of significance to fertility.[34]

One function which is common to the DYRK kinases is their priming abilities. They phosphorylate substrates and activate them for the more efficient phosphorylation of an alternative residue, by other kinases.[35] An important process regulated by the priming DYRKs is the cell cycle, which highlights their importance in disease maintenance and progression.[36]

It is evident that DYRKs are an important class of proteins, however, we were particularly interested in the investigation of DYRK2 as a drug target. Research into DYRK2's significance in cell function is well underway but there are contradictory implications of DYRK2 in cancer progression.

1.2.2 Implications in disease

With respect to function, DYRK2 is an important protein in tumour progression. In response to genotoxic stress, DYRK2 is activated and invokes cell apoptosis by translocating to the nucleus and phosphorylating a serine residue (Ser 46) on p53, a key tumour suppressor protein, **Figure 10**.[37,38]

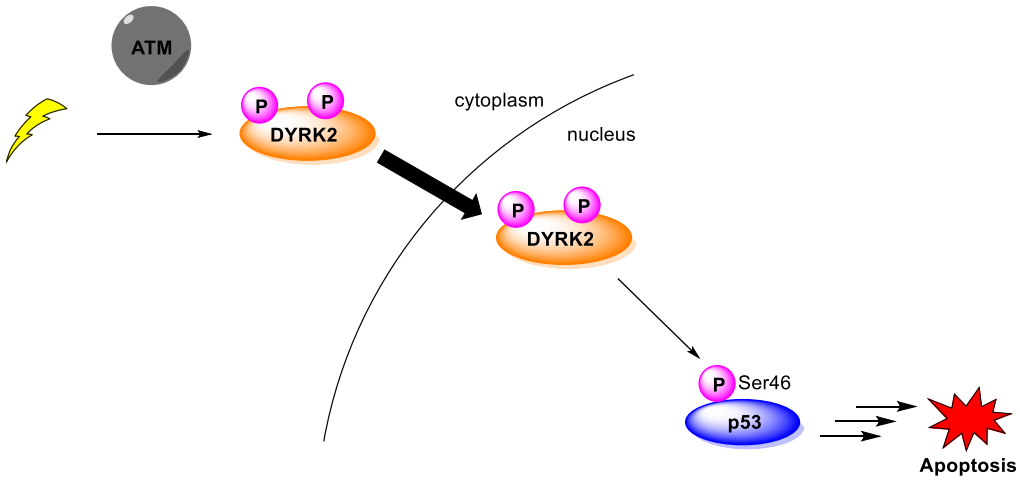


Figure 10: p53 phosphorylation by DYRK2.

This is another example of the importance of kinases in cell maintenance. DYRK2 is primed by another kinase, ataxia telangiectasia-mutated kinase (ATM), which is activated in response to DNA damage, **Figure 10**.[26] This activation switches on the cascade of events leading to cell death.

Another function of DYRK2 is the maintenance of the cell cycle in healthy cells. More specifically, DYRK2 primes c-Jun and c-Myc, important G_1/S transition transcription factors, for phosphorylation by GSK3β. This in turn tags these proteins for poly-ubiquitination by an E3 ubiquitin-ligase, SKp-cullin 1-F-box (SCF), so that they are therefore primed for proteasome degradation. Disruption of this process and accumulation of c-Jun and c-Myc has been reported to contribute to tumour progression.[39]

An interesting study by Morrugares and Calzado *et al.* reported DYRK2 to be a negative regulator of NOTCH1, a member of a receptor family of proteins which plays an important role in the regulation of cell growth and angiogenesis, two key processes in cancer development. Furthermore, they reported that DYRK2 overexpression resulted in a reduction of NOTCH1. This reduction was a result of DYRK2 phosphorylation of NOTCH1 which stimulates its recognition for protein degradation. They also discovered that when DYRK2 expression was depleted, *via* knockdown (KD) experiments, NOTCH1 levels were increased in a number of cancer cell lines.[40] Thus, it is possible that DYRK2 depletion stabilises NOTCH1, by avading protein degradation, which could result in the successful maintenance of cancerous cells.

Although these examples highlight the tumour suppressing capabilities of DYRK2, therefore meaning that DYRK2 inhibition could have potentially negative effects, there is also compelling evidence for the contrary. For example, from a potentially advantageous therapeutic perspective, it was reported that DYRK2 is overexpressed in both lung adenocarcinoma and oesophageal carcinoma.[41] Thus, becoming a potential marker for chemotherapy in non-small cell lung cancer.[42] In addition, an interesting study by Yoshida *et al*., discovered that in hormone receptor positive breast cancers, DYRK2 is a useful predictive marker of sensitivity to the mTOR inhibitor everolimus.[43]

We were interested in the role of DYRK2 in breast cancer (BC) and the effects of small molecule intervention on tumour burden in the aggressive breast cancer subtype triple negative breast cancer (TNBC). BC is a complex disease and the following section addresses the current therapies and research into novel medicines.

1.3 Breast Cancer

BC is the most common cancer of women in the UK and accounts for approximately 15% of reported cancer cases. Cancer Research UK reported 11,400 BC related deaths annually for 2015–2017 and this corresponded to 7% of the cancer related deaths in the UK for this period.[44]

Men face less risk from BC, approximately 100 times less compared to women.[45] There are a number of factors which have been attributed to an increased risk of BC in women. Some of these include: older age, early menarche, late menopause, alcohol consumption, increased body weight and a sedentary lifestyle. Conversely, breast feeding and physical activity have been reported to reduce the risk of breast cancer.[46]

1.3.1 Sub-types

BC can be described and classified according to both histopathological (examination of tissue) and molecular features. First, with respect to tissue examination, there are two common types of invasive carcinoma; invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC). IDC is the most common type of BC in women, with approximately 80% of BC patients having IDC. More specifically, IDC is when the cancer invades the milk duct and can spread to the surrounding fatty tissue of the breast.[47] On the other hand, ILC, accounting for 10–15% of all breast cancers, is less common and is found in the milk producing lobules, which empty into the milk ducts.[48, 49]

BC can be further defined by molecular subtypes, and they are categorised according to the presence and/or absence of important receptors. More specifically, these receptors include the estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor-2 (HER2). The four main molecular subtypes are detailed in **Table 1**.

Table 1: BC subtypes.

	ER	PR	HER2	Cases (%)
Luminal A	+	and/or +	-	>70
Luminal B	±	and/or ±	+	>70
HER2 enriched	-	-	+	13–25
TN/basal like	-	-	-	10–20

The four subtypes are classified according to the presence of hormone receptors (HRs), ER and PR. First, HR-positive, HER2-negative BC tumours can be denoted as luminal A, **Table 1**. Whereas HR-positive, HER2-positive can be described as luminal B, **Table 1**. These are the most common types of breast cancer and contribute to >70% of BC tumours.[50] Next, BC tumours which are HR-negative, HER2-positive are classed as HER2 enriched and finally, tumours which lack all three receptors (HR-negative, HER2-negative) are classed as basal-like or triple negative (TN), **Table 1**.[51] These subtypes are less common and account for 13–25% and 10–20% of BC tumours respectively.[52, 53]

The use of biomarkers is a common tool in identifying and differentiating between breast cancer subtypes and can also be used to identify treatment options. The amplification of a particular gene, Ki67, is a biomarker for cell proliferation. This marker can be used to differentiate between luminal subtypes A and B, where low Ki67 expression is indicative of luminal A and high Ki67 expression suggests the luminal B sub-type. [51, 54]

1.3.2 Biomarkers

In addition to Ki67, from gene sequencing analysis, scientists have identified that each sub-type is enriched with specific genes.[55] Specifically, luminal A and B sub-types have alternative levels of amplification of specific genes. These include ERα, GATA-binding protein and trefoil factor 3, where high expression of these genes is observed in luminal A sub-types and low to moderate expression is found in luminal B sub-types. HER2-enriched BCs are overexpressed with the oncogene ERBB2 and TNBCs are rich in keratins 5/7 and fatty acid binding protein 7.[55, 56]

At a DNA level, these sub-classes exhibit different levels of gene mutations. Luminal sub-types (A and B) have been found to have significantly fewer mutations than the

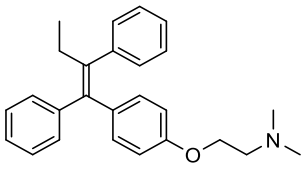
HER2 and TNBCs.[54, 56] Moreover, two of the main mutations are observed in the TP53 and PI3KCA genes,[56] where TP53 encodes for the pivotal tumour suppressor protein, p53.[57] This is particularly significant as mutations of this gene can result in the accumulation of cancerous cells. PI3KCA encodes for PI3K (phosphotidylinositol 3-kinase), an essential protein in cancer maintenance. Activation of this pathway has been reported to be a key process in cancer progression and therapy resistance.[58] These findings are but one example of this complex and heterogeneous disease that is continuously evolving, which highlight the necessary and essential research required to further our understanding of this disease.

These BC sub-types (Luminal, HER2-enriched and TN) help guide clinicians to make informed decisions in terms of the course and length of treatments for patients.[59] Although patient care is becoming more individualised, there are general guidelines when considering treatment options for BC patients. These include tumour size, nodal status (whether the cancer spread to the lymph nodes) and the intrinsic molecular sub-types that are mentioned above.[54]

1.3.3 Therapies

One of the most common approaches to tackling any type of cancer is chemotherapy. There are factors which affect the decision to prescribe chemotherapy and these include: increased Ki67 i.e. increased cell proliferation (luminal B and HER2 enriched) and low hormone receptor status.[51]

HR-positive BCs i.e. luminal A and B have better clinical outcomes, with respect to overall survival (OS) and disease free survival (DFS), compared to HER2-enriched and TNBCs.[54, 60] This is largely due to the presence of the HRs (ER and/or PR) for targeted therapies. Currently, there are a wide range of treatment options for HR-positive BCs with novel biological targets emerging in recent years. However, the most common first line of defence against HR-positive BC is the endocrine (hormone) therapy, tamoxifen **1**, **Figure 11**.[59]



tamoxifen
1

*Figure 11: Structure of tamoxifen **1**.*

In addition to the hormone therapy tamoxifen **1**, aromatase inhibitors (AIs) are used for the treatment of HR-positive BCs.[50, 51, 54] In postmenopausal women, estrogen is primarily produced *via* the aromatase enzyme and therefore the regulation of its production can be targeted through the inhibition of the aromatase enzyme. There are two types of AIs. Type I are steroidal inhibitors such as formestane **2** and exemestane **3**, whereas type II AIs are non-steroidal, for example anastrazole **4** and letrazole **5**, **Figure 12**.[50, 61]

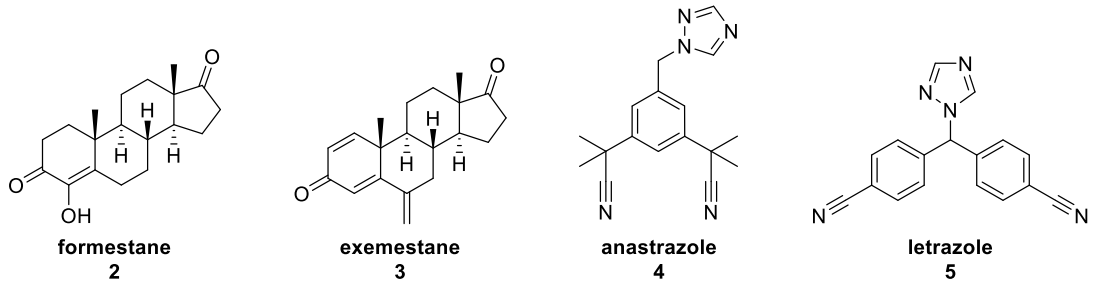


Figure 12: Structure of AIs.

These therapies generally apply to the luminal A subset of HR-positive BCs. The luminal B types require additional treatment as some of them lack ER and are also HER2-positive. Under these circumstances, chemotherapy is often used in addition to/instead of hormone therapy. These include taxanes such as paclitaxel **6** and docetaxel **7** or anthracyclines, **Figure 13**.[51]

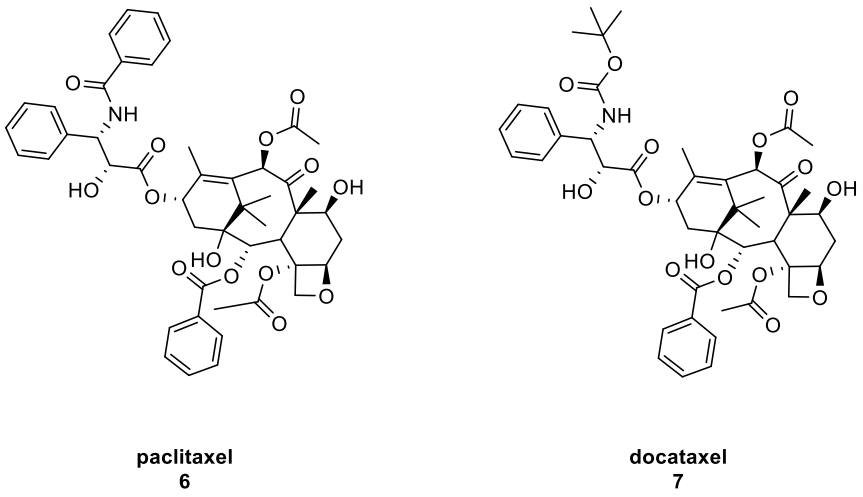


*Figure 13: Examples of taxanes, paclitaxel **6** and docetaxel **7**.*

With respect to HR-negative BC sub-types, HER2-enriched BCs have generally been reported to have poor prognosis and poor clinical outcomes.[54, 55] This sub-class of BC is a result of the overexpression of a particular oncogene, ERBB2/HER2. This gene encodes for the activation of a tyrosine kinase receptor (TKR) pathway, which results

in the promotion of abnormal cellular processes and hence cancer growth.[56, 62] One of the most common methods of treatment for this type of BC is with antibody-based therapies, such as trastuzumab. Trastuzumab treatment increases the OS and DFS of patients with HER2-enriched BC and can be used in combination with chemotherapy in more advanced cases.[56]

One of the biggest challenges in clinical research is the evolution of drug resistance. Unfortunately, some BC patients have developed resistance to hormone therapy and therefore identification of alternative treatment options was required. Scientists have identified novel biological targets which have a significant effect on the regulation of the HRs. Some examples include the mammalian target of rapamycin (mTOR), protein kinase B (also known as AKT), PI3K, cyclin dependent kinases 4 and 6 (CDKs), and histone deacetylases (HDACs).[50, 63] Moreover, with respect to HER2 therapy resistance, some of these proteins are said to be hyper-activated and allow for the regulation of resistance to HR and HER2 receptor therapy.[56] Research into the effect of these proteins is well underway for HR-positive and HER2-enriched BCs. Examples of known inhibitors which are being investigated in clinical trials or are FDA-approved treatments include: everolimus **8** (mTOR), alpelisib **9** (PI3K), palbociclib **10** (CDK4/6), entinostat **11** (HDAC), **Figure 14**. [50, 54, 56]

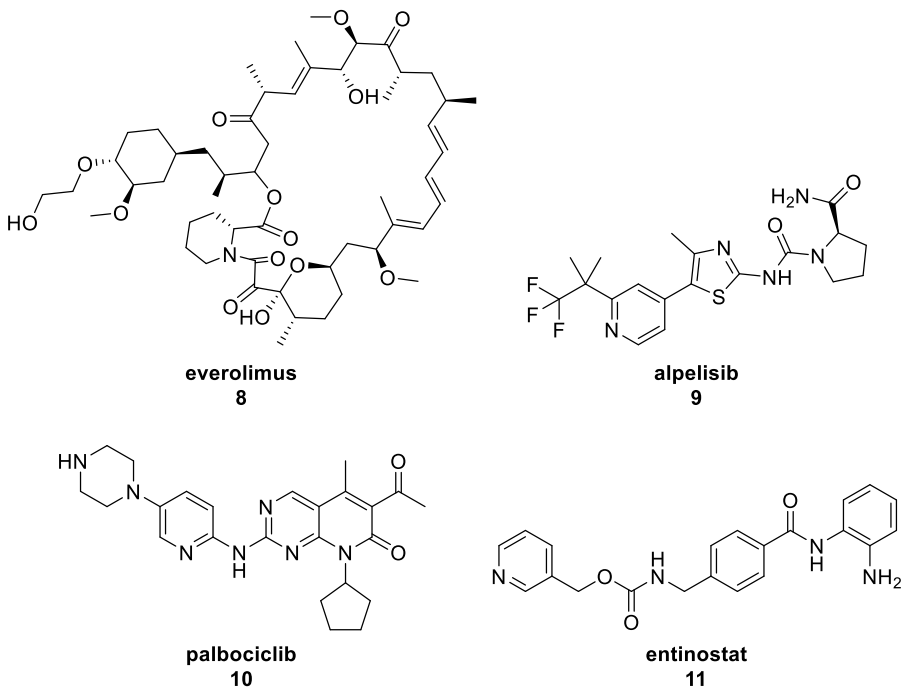


Figure 14: Examples of inhibitors of alternative protein targets.

This is but a brief insight into the understanding of this complex disease and it is apparent that there are successful and promising options for patients with HR-positive

and HER2-enriched BCs. However, the fourth class of BC (TNBC) requires even more attention as this sub-type of BC is associated with the poorest prognosis and clinical outcomes.[55] TNBC has limited therapeutic options due to the lack of receptors for targeted therapies and therefore the most common option for treatment is chemotherapy.[64]

1.4 TNBC

TNBCs account for approximately 10–20% of reported BC cases.[52, 65] TNBC is aggressive in nature and has poorer outcomes with respect to OS and relapse rates, compared to HR-positive and HER2-enriched BCs. There are significant factors which can influence BC and there are also factors which are specifically associated with TNBC. These include race, age and genetics. TNBC is prevalent in younger, premenopausal women and has been reported to be common in African-American and Hispanic women.[52]

This type of BC is associated with a shorter survival rate due its invasiveness with the current most effective treatments being surgery/chemotherapy. This results in an increased rate of relapse and poorer response to secondary treatment. Moreover, depending on the stage of tumour growth and therefore cancer progression, chemotherapy can also be ineffective. This highlights the critical need for the investigation of novel therapeutic targets for this aggressive disease.[52]

TNBC is mainly described as lacking three key receptors; ER-negative, PR-negative and HER2-negative, and within this class there are smaller sub-classes with varying, albeit lower levels of the HRs and HER2. This can be determined *via* IHC, and along with the detection of biomarkers, this can provide a wider understanding of the treatment options.[66]

In the previous section, biomarkers for BC sub-types were highlighted and these are often used as a prognostic tool. There have been reports of overexpression of particular proteins in TNBC tumours and these include: cytokeratins 5/17,[55] EGFR, p53 and Ki67, the proliferative markers. In addition, there are significant mutations observed in TNBCs and these include the germline mutations of the Breast and Ovarian Cancer genes, BRCA1 and BRCA2.[67] These genes encode for tumour suppressor proteins BRCA1/2 and are commonly found in breast and ovarian tissues. Studies have shown that these mutations are often found in TNBC cases (60–80%),[52, 65, 67] thus highlighting the genetic influence on this disease. These mutations disrupt

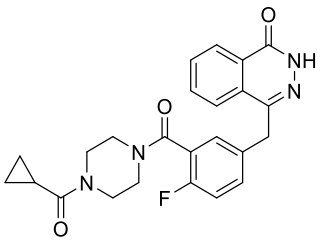
DNA repair and often result in poor outcomes. The further understanding of these mutations has provided a target for potential treatment options.[68]

1.4.1 Current therapies for TNBC

The most common defence against TNBC is chemotherapy e.g. taxanes and anthracyclines. However, depending on the tumour size and metastatic nature of the disease, surgery may be required before or during the course of chemotherapy treatment to ensure the best possible outcome.[52] This may even require a full mastectomy or a breast conserving surgery (BCS), where BCS is followed by radiotherapy to prevent further spread or relapse.[22, 65]

As previously mentioned, a common mutation observed in a high number of TNBCs is the germline BRCA1/2 gene mutation, which disrupts DNA repair.[68] In line with this observation, the use of DNA damaging platinum agents has been investigated for treating patients with this mutation.[69] These include cisplatin and carboplatin. However, there is a long list of negative side effects associated with platinum therapies and therefore, investigations into cis- and carboplatin as viable treatments for BC, has been examined in combination with other drugs to decrease cytotoxicity. For example, the treatment of carboplatin and paclitaxel has shown promise with respect to response rates however, this still requires further investigation to determine the OS benefits.[65] A recent clinical trial investigated the use of carboplatin as an alternative to docetaxel. This trial determined that carboplatin was beneficial to those patients with BRCA mutations, with respect to reduced tumour burden and delayed onset of disease progression.[66]

An alternative target which is currently used for the treatment of ovarian cancer has also shown success with breast cancer. Poly(ADP)ribose polymerase (PARP) is a protein that repairs single strand breaks (SSBs) in DNA, a common feature of BRCA mutations. In cancerous cells, PARP inhibitors are beneficial because they promote the formation of double strand breaks (DBS) which result in cell death.[70] Results from a 2017 phase III clinical trial revealed that treatment with the PARP inhibitor, olaparib **12**, **Figure 15**, provided a significant increase in progression free survival, compared to standard chemotherapy in a cohort of BC patients with HER2-negative status and BRCA mutations, including TNBC.[71]



olaparib
12

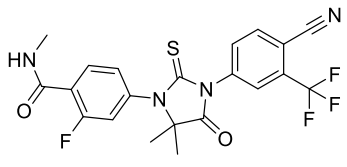
*Figure 15: PARP inhibitor olaparib **12**.*

In 2018, olaparib **12** was approved by the FDA for the treatment of patients with BRCA mutations and HER2-negative BC.[72]

In addition to PARP inhibitors, the current approved treatments for TNBC do not go far beyond chemotherapy. There is an undoubtable need for the identification and investigation of alternative methods to treat TNBC and this research is underway.

1.4.2 Current research

TNBC itself can be divided into molecular subtypes which can be useful as prognostic tools. One of these is the luminal androgen receptor (LAR) subtype.[73] This class of TNBC indicates an androgen receptor (AR) dependence. The AR is a hormone receptor, similar to ER and PR, which promotes the formation of sex hormones, such as testosterone and dihydrotestosterone.[74] TNBC lacks hormone receptors for targeted therapies and thus the presence of the AR provides a gateway for therapeutic investigation. A phase II clinical trial determined the effect of a known AR inhibitor, enzalutamide **13**, **Figure 16**, on patients with AR-positive TNBC.[74]



enzalutamide
13

*Figure 16: AR inhibitor enzalutamide **13**.*

At present, enzalutamide **13** is an approved treatment for castration-resistant prostate cancer.[75] However, with respect to TNBC patients, treatment with enzalutamide resulted in an increase in progression free survival (PFS) of patients with higher levels of AR, providing an insight into the specificity of this inhibitor.[74]

In the previous sections, in response to drug resistance, alternative targets for the treatment of HR-positive and HER2-positive BCs were highlighted, which included the kinases mTOR and CDK.[50] These proteins are also under investigation as tools for TNBC therapy as monotherapies, and in combination with antiandrogens.[74]

TNBCs have increased levels of gene mutations compared to HR-positive BCs. Mutation of the PIK3CA gene, which encodes for the transcription of the PI3K protein, is significant in TNBCs. PI3K is involved in a complex and essential pathway which includes a number of other important kinases, such as AKT and mTOR. This pathway, often termed the PAM (PI3K AKT mTOR) pathway, has been reported to promote the survival of cancerous cells and regulate the cell cycle, thus providing an attractive target for inhibition.[76] With respect to TNBC, targeting this pathway with inhibitors has provided some success. The AKT inhibitor MK-2206 **14** (**Figure 17**) reached clinical trials but was abandoned due to undesirable side effects. The PI3K inhibitor, alpelisib **9** (see **Figure 14**), has shown promise in a phase I/II study. More specifically, in combination with the chemotherapy agent, nab-paclitaxel, an increase in the PFS of HER2-negative (including TNBC) patients, in comparison to patients with no PIK3CA mutation was observed.[77]

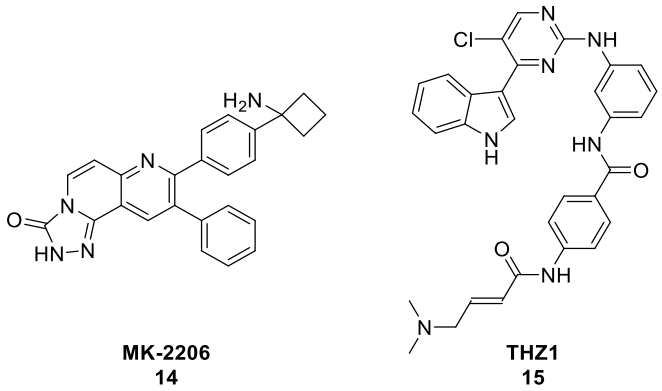


*Figure 17: AKT inhibitor MK-2206 **14** and CDK7 inhibitor THZ1 **15**.*

Another kinase which has been under investigation for potential use as a prognostic tool, is CDK7. A study by Li and colleagues reported that high levels of CDK7 is associated with poor clinical outcomes in TNBC patients.[78] Moreover, they revealed that the inhibition of CDK7 with a known specific inhibitor, THZ1 **15 Figure 17**, resulted in a significant reduction in the cell proliferation of TNBC cell lines.[78]

This is but a brief insight into the research being carried out to further understand this complex disease. Due to its aggressive nature, and the lack of receptors for targeted

therapies, it is necessary to identify and investigate alternative biological targets in order to treat TNBC.

Two targets currently under investigation in TNBC are the proteasome and heat shock factor-1 (HSF1), both of which involve a role for DYRK2. The next section briefly describes the current state of research with these two targets.

1.5 DYRK2 - target validation

1.5.1 The proteasome

DYRK2 plays a vital role in the protein degradation process. An extremely important protein-complex involved in cell maintenance is the 26S proteasome,[79] the main function of which is the targeted degradation of ubiquitin (Ub) tagged proteins, **Figure 18**.

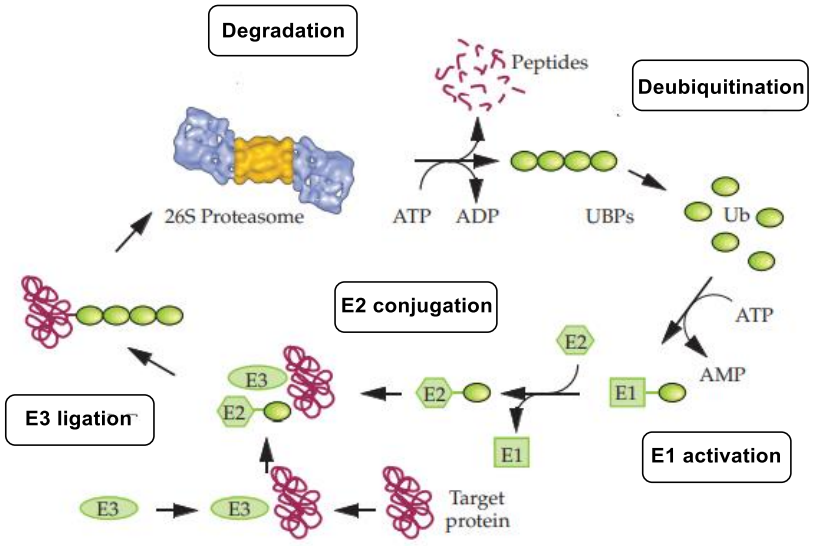


Figure 18: Protein degradation process.[80]

In general, the process of protein degradation can be described in a number of steps. First, the monomeric ubiquitin (Ub) is activated in an ATP dependant manner by an E1 Ub-activating enzyme (E1 activation, **Figure 18**), before the activated Ub is recognised by a second E2 Ub-conjugating enzyme, to which it is transferred by a transthiolation reaction (E2 conjugation, **Figure 18**). Upon interaction with the target protein, an E3 Ub-ligase enzyme, E3 facilitates the binding of Ub to the target protein lysine side chain (E3 ligation, **Figure 18**).[81] Subsequent polyubiquitination by the E3-ligase *via* Lys 44 residues on the Ub monomers targets the protein to the 26S proteasome for degradation. The protein is then broken down, in an ATP dependant

process, into short peptide fragments and the poly-ubiquitin complex is released (degradation, **Figure 18**). This complex can then be broken down into monomeric Ub for the process to begin again (de-ubiquitination, **Figure 18**).[82,83]

The 26S proteasome protein structure is complex, but as like all proteins, it is fit for purpose. Simplistically, it is composed of a 20S core structure or core particle (CP) and a 19S regulatory particle (RP), **Figure 19**.[84] In more detail, the 19S RP can be split into structures described as the base and lid. With respect to the base, there are smaller sub-units, ATPase's (Rpt1–6), which catalyse the unfolding of protein substrates and has been described as the engine of the proteasome framework.[85] With respect to the lid, this contains sub-units (Rpn's) which are responsible for the binding and the release of the Ub chain attached to the protein for degradation.[86]

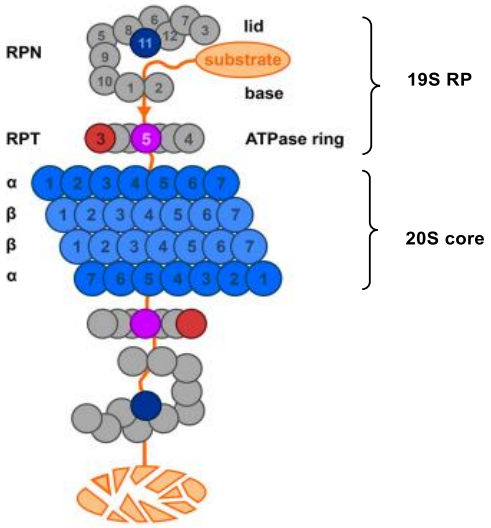


Figure 19: Simple representation of the 26S proteasome.[81]

This is but a simple depiction of the 26S proteasome as it has been widely investigated and is a well-defined protein complex. Nevertheless, this highlights the importance of the 26S proteasome in cell maintenance.

Within cancer the proteasome is of therapeutic value as its normal function is disrupted. Proteasome inhibitors are not only being studied as anti-cancer therapeutics but are used in the clinic as they have been shown to induce apoptosis.[87] The precise mechanism by which apoptosis occurs is unclear, however, we can infer that, proteasome inhibition allows for the accumulation or activation of pro-apoptotic proteins which can promote cancer cell death.[88] Research has shown that the treatment of TNBC tumours with the proteasome inhibitor bortezomib decreases tumour growth, however, proteasome inhibitors alone are not enough to treat these

cancers as they have poor tumour penetration and not all tumours respond to this treatment.[89] Therefore, identifying other target proteins involved in proteasome activity provides the potential for combinative therapy.

Guo and co-workers investigated DYRK2 as a positive regulator of the 26S proteasome. Through extensive investigation they discovered that phosphorylation of the Rpt3 ATPase at Thr25 is essential for proteasomal activity. As briefly highlighted previously, the ATPase subunits are important for protein substrate unfolding and thus activation of such subunits is essential for normal proteasomal function. Furthermore, they found that DYRK2 was the only kinase, along with its closely related family members (1B, 3 and 4), that phosphorylated Rpt3 at Thr25. Additionally, investigating cancer cell growth, in the basal-like TNBC MDA-MB-468 cell line, DYRK2-KO studies confirmed that the loss of DYRK2 resulted in a decrease in tumour volume and an increase in sensitivity to the proteasome inhibitor bortezomib **16**, **Figure 20**.[79]

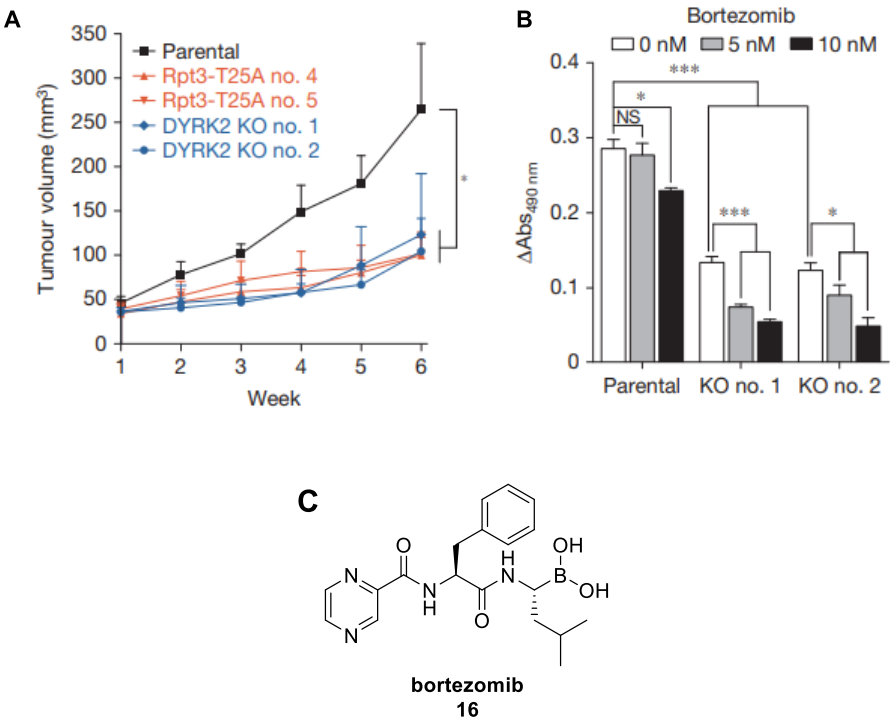
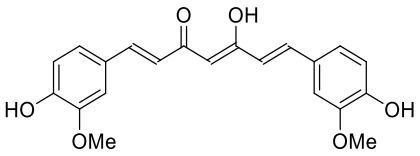


*Figure 20: Results from study by Guo et al., **A**) In vivo results from TNBC tumour growth in mice xenografts. **B**) MDA-MB-468 cells, WT and DYRK2 KO, treated with bortezomib at varying concentrations.[79] C) structure of bortezomib **16**.*

A follow up study by Guo *et al*., investigated the effects of DYRK2 inhibition, with the natural therapeutic curcumin **17**, **Figure 21**, on proteasome activity and tumour burden in TNBC mice xenografts.



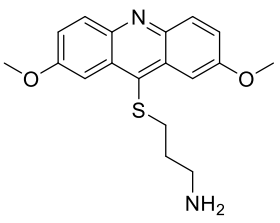
curcumin
17

***DYRK2**: IC_{50} 5 nM*
DYRK1A: 190 nM
DYRK3: 20 nM

*Figure 21: Structure of curcumin **17** and DYRK potencies.*[90]

They discovered that curcumin **17** is a potent inhibitor of DYRK2, with an IC_{50} of 5 nM, however, it is not selective among its closely related isoforms, **Figure 21**. In addition, at 1 µM, curcumin **17** significantly inhibits the activity of other kinases including: RIPK2, MLK1 and 2, PIM3 and PHK. Although curcumin **17** is not a selective inhibitor for DYRK2, it displayed attractive results with respect to proteasome activity and tumour growth. They revealed that curcumin **17** inhibits DYRK2 and reduces proteasome activity in a number of cancer cell lines, including the basal-like TNBC cell lines MDA-MB-436 and 231. Complementary to their previous work, combination treatment with curcumin **17** and the proteasome inhibitor carfilzomib decreased proteasome activity compared to single exposure of each inhibitor. *In vivo* studies with MDA-MB-231 TNBC mice xenografts also showed that treatment with curcumin reduced tumour growth to a similar level to that of the DYRK2 KO results.[90]

The results from this group expanded even further as they developed a novel and potent inhibitor for DYRK2, **18**, however, this compound was not selective for DYRK2 over its family member DYRK3 and PIM3 kinase, **Figure 22**.



LDN192960
18

Kinase	IC_{50} (nM)
DYRK2	13
DYRK1A	122
DYRK3	<3
PIM3	10
CK2	>27000

Figure 22: LDN192960: a potent but non-selective inhibitor for DYRK2.[91]

They identified that **18** reduces 26S proteasome activity in TNBC cells and at 1 µM significantly reduces cell growth in TNBC cell lines. Interestingly, **18** is an effective inhibitor of proteasome activity in bortezomib-resistant multiple myeloma (MM) cells,

where such resistance is common for MM cancer patients.[92] Furthermore, in TNBC mice xenografts, treatment with **18** reduced tumour burden.[91]

Investigations have been carried out by our collaborators at Glasgow and Dundee University, which provided additional evidence for selecting DYRK2 as a target for TNBC therapy. This was achieved by researching the relationship of DYRK2 with heat shock factor-1 (HSF1), the transcription factor which regulates the cellular stress response.

1.5.2 HSF1

Heat shock factor-1 (HSF1), is a transcription factor and a complex protein, which has been described as the master regulator of proteotoxic stress pathways. Most commonly, HSF1 responds to heat shock, however, it is also activated in response to other stress stimuli.[93] These include: irregular pH, the presence of reactive oxygen species, heavy metals and agents such as proteasome inhibitors.[94, 95]

1.5.2.1 Function

Under normal conditions, HSF1 is held in an inactive protein complex in the cytoplasm, and is released upon activation by stress stimuli, **Figure 23**.

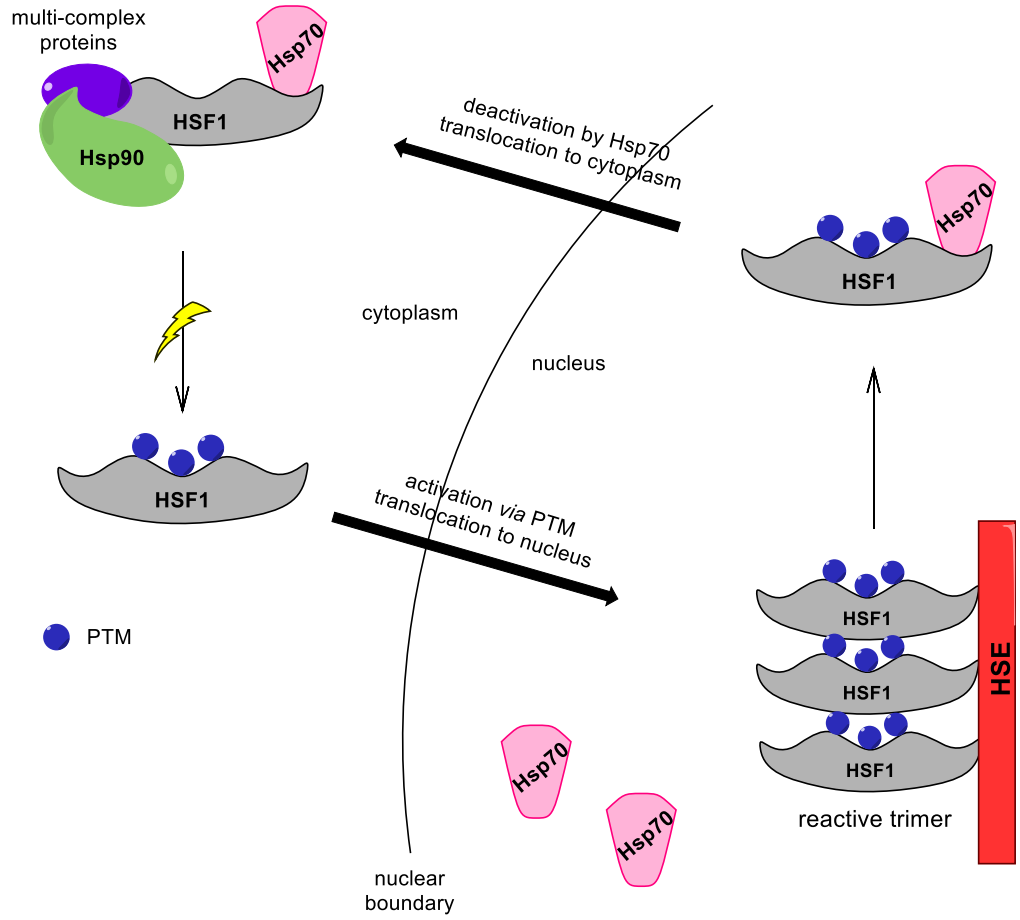


Figure 23: Schematic of HSF1 pathway.

This HSF1 complex also includes heat shock proteins (HSPs), Hsp70 and Hsp90 (**Figure 23**), where HSPs are classified according to their approximate molecular weight.[94] HSPs are also known as molecular chaperones, which respond to alterations in protein folding and function as a defence mechanism against misfolding and aggregation,[94, 96] which is a common response to cellular stress. Hsp90 itself is bound to other co-chaperone proteins, including the p23 protein.[94] p23 plays a role in the regulation of the cell cycle and to help Hsp90 bind to client proteins.[97] Once released from the HSP complex, HSF1 is activated *via* post-translational modifications (PTMs) such as phosphorylation, acetylation or SUMOylation, **Figure 23**.[96] SUMOylation is a process similar to that described for protein degradation *via* ubiquitination (see **Figure 18**, **Section 1.5.1**), but involves the addition of all Ub-like modifiers (SUMOs).[98] These modifications allow HSF1 to translocate from the cytoplasm to the nucleus where it can form a reactive trimer, **Figure 23**. This trimer can then bind to heat shock elements (HSEs) containing repetitive DNA sequences nGAAGn, which allow for the transcription of protective genes and proteins i.e. HSPs.[99] The process is controlled by a negative feedback mechanism. When HSP levels increase, they include Hsp70, which can interact with

HSF1 in the nucleus. This dissociates the trimeric form back to the monomer, allowing reengagement with Hsp90 (and its co-chaperones) to translocate HSF1 out from the nucleus into the cytoplasm, **Figure 23**.

The process of HSF1 activation in response to proteotoxic stress can be described in more detail with respect to its structure.

1.5.2.2 Structure

With respect to the mammalian structure of HSF1, it consists of four main functional domains. These include the *N*-terminal DNA-binding domain (DBD), oligomerisation domains (leucine zipper (LZ) regions), the regulatory domain (RD) and C-terminal transactivation domain (AD), **Figure 24**.[95, 99, 100]

Figure 24: Simplistic representation of HSF1 structure.

At present, there is no complete crystal structure for HSF1, however, there is a co-crystal structure available which represents the human HSF1 DBD bound to DNA. It was subsequently deduced that the DBD (green, **Figure 24**) binds to the major groove of DNA *via* interaction with nGAAGn sequences.[99] Thus, the DBD is essential for binding to HSE in the nucleus, as these contain the same repetitive sequence. In its monomeric form, the DBD and RD can form intramolecular bonds to prevent oligomerisation, however, when stress is induced, these regions can then dissociate and expose one of the oligomerisation domains.[100] To that effect, the oligomerisation domain, also referred to as the LZ region, is responsible for the formation of the reactive trimer *via* intermolecular disulfide bonding of HSF1 monomers.[99] Next, the RD controls the activity of HSF1 as it contains sites which result in the activation or suppression of HSF1. Finally, the fourth region (AD) is the transcriptional activity domain, which carries out the title function.[99]

1.5.2.3 HSF1 activation

HSF1 activity is primarily controlled by PTMs such as phosphorylation and acetylation. Extensive research has therefore been carried out into identifying the sites on HSF1 which undergo PTMs, due to the importance of the HSF1 process in disease

progression. An informative review by Naidu *et. al* describes some of the PTM sites of HSF1 and identifies the enzymes which catalyse these modifications, **Figure 25**.[95]

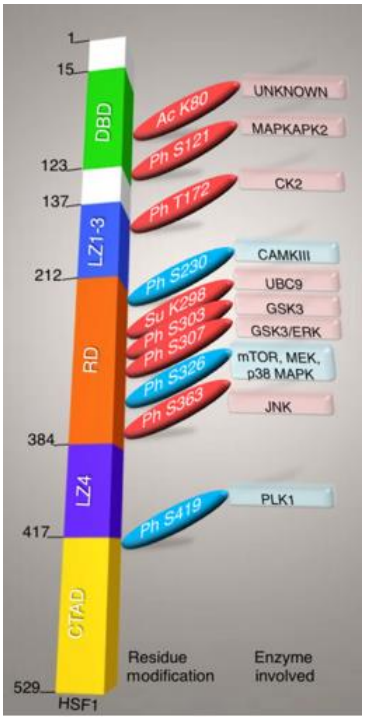


Figure 25: PTM sites of HSF1 (red-inhibiting, blue-activating) Ph = phosphorylation, Ac = acetylation and Su = SUMOylation.[95]

With respect to **Figure 25**, PTM of the residues highlighted in red has been reported to inhibit HSF1 activity, for example phosphorylation of Thr172, and Ser303 by casein kinase-2 (CK2) and glycogen synthase kinase-3 (GSK3) respectively, or SUMOylation of Lys298 by ubiquitin carrier protein-9 (UBC9). Residues highlighted in blue promote HSF1 activity, for example, phosphorylation of serine residues (Ser230, Ser326 and Ser419) activate HSF1 and promote its function, **Figure 25**. Phosphorylation of Ser326 has been reported to be the cornerstone of HSF1 activation and previous work has provided evidence to support that there is more than one kinase involved in this process. These include: mTOR, p38 mitogen activated protein kinases (MAPK) and MEK, an important protein in the MAPK pathway.[95, 101] Studies have shown that deactivation of this site, *via* kinase inhibition or mutation experiments, significantly reduces HSF1 activity.[101] Once more, this further demonstrates the essential work carried out by protein kinases with respect to cell maintenance and survival.

Other PTM sites not included in **Figure 25** that result in the activation of HSF1 include Thr142, phosphorylation by casein kinase-2 (CK2), which promotes the transcriptional

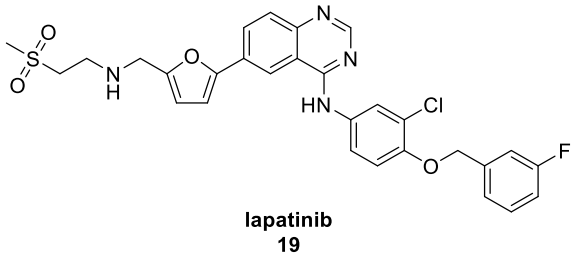
activity of HSF1[101] and Ser320 in the RD, which is phosphorylated by protein kinase A (PKA).[102]

1.5.2.4 Role of HSF1 in BC

The formation and growth of cancerous cells subjects them to numerous forms of stress, which activates HSF1 as an instinctive defence mechanism. Unfortunately, this can aid the survival of these cells *via* the transcription of HSPs. In the previous sections, we have briefly described some of the essential proteins involved in BC and its treatment. Thus, from the description of HSF1 so far, it will come as no surprise that HSF1 has been reported to be of significance in BC and possibly its treatment.[97, 103-106]

Increased HSF1 expression has been reported to be associated with poor prognosis in BC sub-types.[97, 103] Furthermore, increased levels of Hsp70 have been associated with poor response to radiotherapy.[97] We can potentially attribute poor prognosis to the induction of Hsp90. HER2/ERBB2 (one of the molecular sub-types of BC) is a client of Hsp90[97] and in response to cellular stress, Hsp90 is upregulated and maintains the survival of ERBB2 mutants in HER2-positive BCs.

HSF1 has been reported to play an important role in lapatinib **19** (**Figure 26**) resistance in HER2-positive breast cancer. In response to the inhibition of the lapatinib protein targets HER2 and EGFR, receptor tyrosine kinases (RTKs) are upregulated and contribute to the loss of lapatinib's function. HSF1 maintains lapatinib resistance through the release of HSPs. More specifically, HSPs protect and stabilise the RTKs which are expressed in response to lapatinib treatment.[104]



lapatinib
19

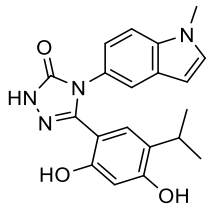
*Figure 26: Structure of EGFR inhibitor lapatinib **19**.*

Hsp70 and Hsp90 maintain the ER, an important prognostic tool for BC, and thus are important in the response to ER mutations.[97, 105] Furthermore, a comprehensive study identified a relationship between HSPs and the BC biomarker BRCA1. Cells with wild type (WT) BRCA1 showed a significant increase in the expression of Hsp27.[106]

It is clear that the heat shock response pathway is critical to BC, and, therefore, provides alternative proteins for investigation as possible therapeutic targets.

1.5.2.5 Targeting HSF1 activity

HSF1 is an attractive target in the investigation of novel therapies for BC due to its upregulation of protective proteins (HSPs). Clinical trials have been and are currently underway into the use of Hsp90 inhibitors for the treatment of BC.[97, 107] Hsp90 plays a pivotal role in the release of HSF1 and thus, is partly responsible for its own activation and the promotion of its cytoprotective capabilities. An example of an Hsp90 inhibitor, which is being examined for TNBC treatment, is ganetespib **20** (STA-9090), **Figure 27**.

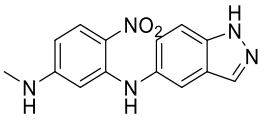


ganetespib
20

Figure 27: Hsp90 inhibitor.

Treatment with ganetespib has been shown to significantly reduce tumour burden in TNBC mice xenografts, and is currently being investigated in a Phase II clinical trial (ENCHANT-1 trial; NCT0167745) for the treatment of HER2-negative BCs, which includes TNBC patients.[107]

Due to the lack of a full length crystal structure of HSF1, targeting HSF1 directly has been a challenge from a medicinal chemistry perspective. However, there have been successes in this area, such as the development of the commercially available HSF1 inhibitor, KRIBB11 **21**, **Figure 28**.



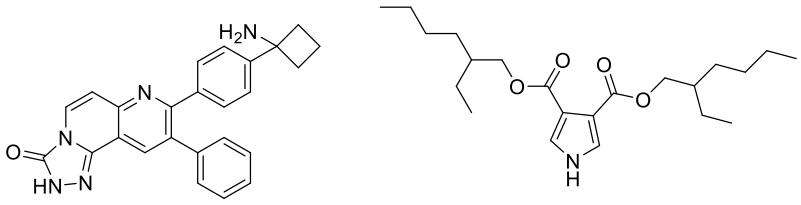
KRIBB11
21

*Figure 28: Structure of HSF1 inhibitor, KRIBB11 **20**.*

KRIBB11 **21** has been reported to bind to the transactivation domain of HSF1 and disrupt the transcription of protective genes.[99, 108] KRIBB11 has provided positive

results both in *in vitro* and *in vivo* settings with respect to reducing cancer cell and tumour growth.[109]

A recent study by Carpenter and co-workers reported that treatment with the HSF1 inhibitor KRIBB11 **21** in combination with the AKT inhibitor MK-2206 **14** (**Figure 29**) significantly reduced tumour burden in TNBC MDA-MB-231 infected mice xenografts, compared to KRIBB11 **21** administered as a single agent.[108]



MK-2206
14

TCCP
22

*Figure 29: Structures of AKT inhibitor MK-2206 **14** and ERK activator TCCP **22**.*

Previously, we mentioned that there are residues in the RD and LZ regions of HSF1 which negatively regulate its activity. A study by Rashmi *et. al* identified a pyrrole based natural product, TCCP **22**, **Figure 29**, which reduces the expression of Hsp90 and Hsp70 in the MDA-MB-231 TNBC cell line.[110] Interestingly, it was found to promote the phosphorylation of HSF1 at Ser307 by EKT, a known inhibitory modification, and thus has a deactivating effect on HSF1.[110]

These findings demonstrate the challenge of targeting HSF1 through small molecule intervention, as there are a number of proteins responsible for its activity. However, from a positive perspective, the engagement of HSF1 with inhibitors is a viable approach for the treatment of TNBC as studies have shown its importance in tumour progression.[107, 110]

Through collaboration with Professor Joanne Edwards at Glasgow University and Dr. Laureano de la Vega at the University of Dundee, we have identified a relationship between TNBC and our target protein DYRK2, and a relationship between DYRK2 and HSF1.

1.5.3 TNBC and DYRK2 – preliminary data

From clinical data, Professor Joanne Edwards at the University of Glasgow identified a relationship between the levels of DYRK2 expression and TNBC patient survival, **Figure 30**.

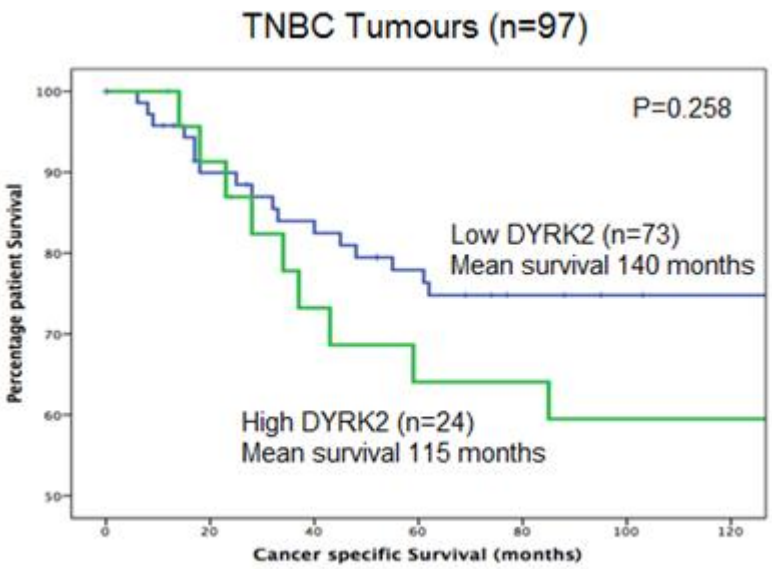


Figure 30: Relationship between DYRK2 expression and TNBC patient survival.

A cohort of 97 TNBC tumour samples were screened and IHC-stained for levels of nuclear DYRK2 expression. The Kaplan-Meier graph shows that low DYRK2 expression trends towards an increased survival rate of TNBC patients (140 months), whereas high DYRK2 expression was associated with a decreased survival rate among TNBC patients (115 months), **Figure 30**. However, it should be noted that due to the small cohort size that this was not a significant result.

With this preliminary clinical data suggesting DYRK2 expression negatively effects TNBC patient survival, we were encouraged to investigate DYRK2 and its relationship with TNBC further. With collaborative efforts from Dr. Laureano de la Vega at the University of Dundee, an important relationship between DYRK2 and HSF1 in TNBC cell lines was identified.

1.5.4 DYRK2 and HSF1 – preliminary data

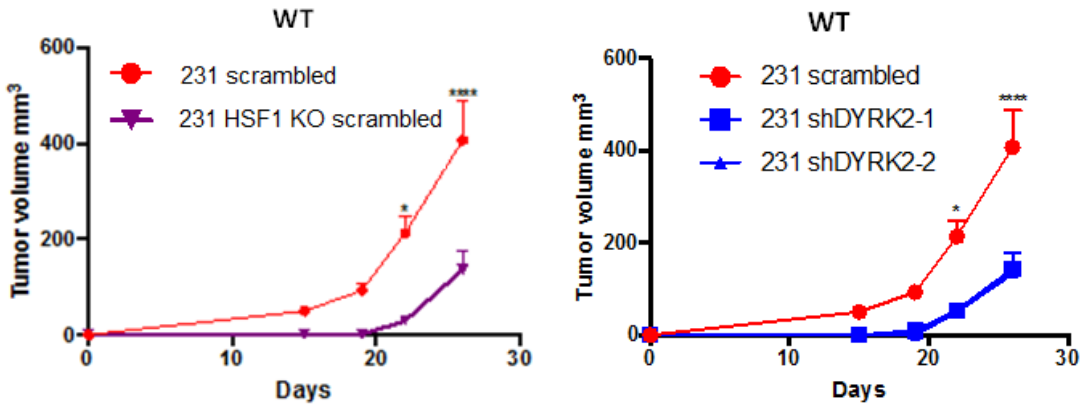


*Figure 31: Tumour burden in mice bearing MDA-MB-231 WT cells infected **LHS**: with sh-scramble or HSF1 KO cells. (all n=3), **RHS**: with sh-scramble or two independent shDYRK2.*

Firstly, to investigate if DYRK2 or HSF1 depletion decreased TNBC tumour burden in mouse xenografts, de la Vega performed a pilot study where WT and HSF1-KO MDA-MB-231 cells infected with either shControl or shDYRK2 RNA which were inoculated sub-cutaneously into the flank of virgin NSG™ mice. From the results highlighted in **Figure 31** (LHS) it is apparent that HSF1 KO cells result in delayed and lower tumour growth after 25 days when infected with shControl, suggesting HSF1 is important for tumour growth. With MDA-MB-231 WT cells infected with either shControl or shDYRK2 RNA in two independent experiments (**Figure 31** (RHS)), the results also showed a significant reduction in the tumour volume after 25 days, implying a role for DYRK2 in tumour growth.

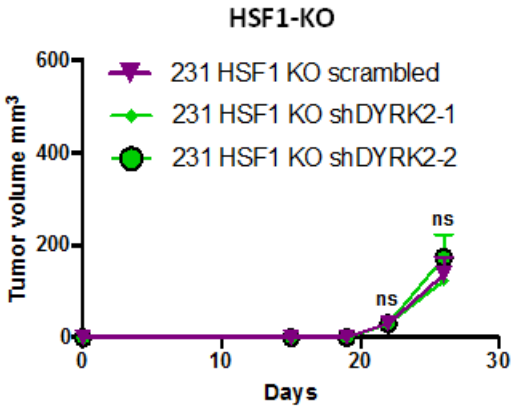


Figure 32: Mice xenografts infected with MDA-MB-231 HSF1 KO and with sh-scrambled or two independent shDYRK2 (n=3).

Finally, comparing the data from the xenograft experiments of HSF1-KO MDA-MB-231 cells infected with shControl or shDYRK2 demonstrated no significant

difference in the effect of DYRK2 depletion on tumour burden in these mice, suggesting an important link between HSF1 and the tumour promoting role of DYRK2, **Figure 32**, and highlighting that it would be advantageous to investigate the inhibition of these proteins.

To determine if DYRK2 was an activator of HSF1, de la Vega examined the relative mRNA expression of Hsp70 (as described earlier, a key protein in the HSF1 pathway) in both the WT and DYRK2-KO MDA-MB-468 TNBC cell line, **Figure 33**.

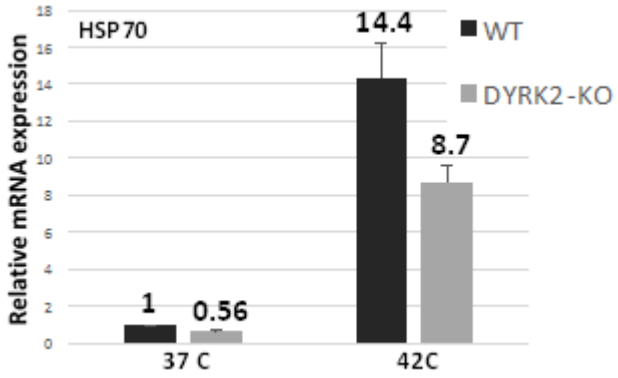


Figure 33: Hsp70 expression in MDA-MB-468 WT and DYRK2 KO at 37ºC and 42ºC.

HSF1 is activated and released from its co-chaperone complex upon proteotoxic stress, before translocation to the nucleus where it can, in turn, promote the expression of Hsp70. Therefore, in order to assess the relative expression of Hsp70, HSF1 must be activated. In this case, heat shock (HS) was the activator and Hsp70 expression was measured at physiological temperature (37 ºC) and HS (42 ºC). From the results (**Figure 33**), it is clear that mRNA Hsp70 expression is increased by heat shock, suggesting induction of protein levels in response to HSF1 activation. Interestingly, in the DYRK2-KO TNBC cell line, there is a significant reduction in Hsp70 expression compared to the WT. Thus, this is suggestive that DYRK2 plays a role in the activation of HSF1 in response to proteotoxic stress.

Through further investigation, de la Vega was able to demonstrate that upon activation by HS, protein levels of HSF1 phosphorylated at two sites Ser320 (p-HSF1 (S320) and Ser326 (p-HSF1 (S326)), were increased, **Figure 34**.

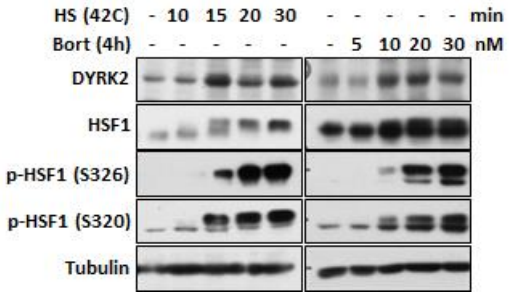


Figure 34: MDA-MB-468 cells were exposed to HS/proteasome inhibitor bortezomib (Bort) for different durations/concentrations.

After 15 min HS exposure (lane 3), it appears that DYRK2 expression is increased, which corresponds with an increase in phosphorylated HSF1 (p-HSF1 (S320) and p-HSF1 (S326)), **Figure 34**. Whilst DYRK2 levels increase to a constant level from 15 min (lanes 3-6), there is a time-dependent increase in p-HSF1 (S326) from 15 to 20 min (lanes 3-4), but not p-HSF1 (S320), which remains constant (lanes 3-6). Treatment with increasing concentrations of the proteasome inhibitor bortezomib as the source of proteotoxic stress delivered similar results. At concentrations ≥10 nM (lane 8), DYRK2 expression is amplified as are phosphorylated HSF1 levels (S326 and S320), **Figure 34**. Therefore, we can suggest that, upon the introduction of stress stimuli, DYRK2 is involved in the phosphorylation of HSF1 at Ser320 and Ser326.

Thus far, we have provided evidence that DYRK2 and HSF1 are essential for tumour growth in TNBC mice xenografts, and upon HS DYRK2 has the potential to be a positive modulator of HSF1 in TNBC cells. To investigate the direct relationship between DYRK2 and HSF1, de la Vega transfected the immortal kidney cell line HEK293T with a DYRK2 analog-sensitive (AS) mutant that was GFP tagged for quantification, **Figure 35**.

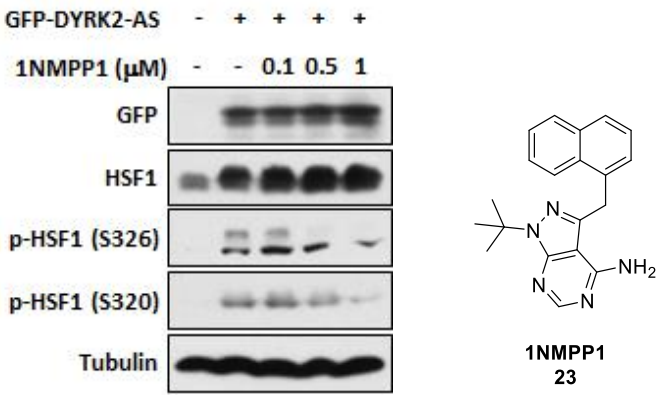


*Figure 35: **LHS**: 293T cells were transfected with the analogue sensitive DYRK2 mutant. 45 hours later, 1NMPP1 (AS inhibitor) was added for 2 hours. **RHS**: structure of AS inhibitor 1NMPP1 **23**.*

AS-kinase mutants (also known as Shokat mutants) are generated by using a space-creating mutation at the GK residue to allow inhibition by a bulky ATP-competitive cell permeable small molecule (in this case 1NMPP1 **23**, **Figure 35**). This enables dissection of kinase-signalling pathways and the physiological role of individual kinases.[111]

Upon introduction of the GFP tagged DYRK2, de la Vega observed an increase in HSF levels and phosphorylation of HSF1 at Ser320 and Ser326 (lane 2). However, upon exposure to the specific inhibitor, 1NMPP1, phosphorylation at these sites was diminished in a concentration-dependent response, **Figure 35** (lanes 3-5), thus demonstrating a synergistic relationship between DYRK2 and HSF1 *via* phosphorylation.

Complementary to this, de la Vega generated kinase dead (KD) flag-tagged DYRK2 mutants to validate the link between DYRK2 and HSF1 phosphorylation in HEK293T cells, **Figure 36**.

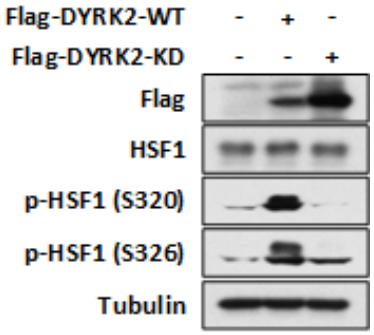


Figure 36: HEK293T cells transfected with flag tagged DYRK2 WT and KD plasmids.

Comparison of cells transfected with the flag-tagged DYRK2 WT or KD mutant revealed that in the former, there is phosphorylation of HSF1 at Ser320 and Ser326, **Figure 36** (lane 2). On the other hand, when transfected with DYRK2-KD, phosphorylation at these two sites was effectively abolished, **Figure 36** (lane 3). Together with the DYRK2-AS experiments, these results provide us with evidence to support that DYRK2 is indeed responsible for the phosphorylation of HSF1 at Ser320 and Ser326.

Finally, de la Vega demonstrated that DYRK2 depletion in the TNBC MDA-MB-468 cell line *via* siRNA impaired HSF1 nuclear accumulation and phosphorylation after induction by proteotoxic stress (HS), **Figure 37**.

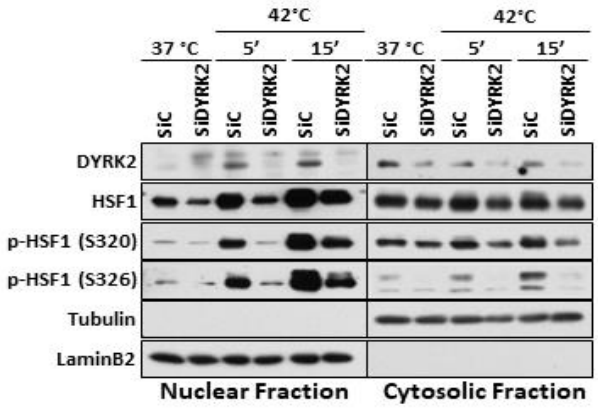


Figure 37: MDA-MB-468 cells transfected with siControl (siC) or siDYRK2 were exposed to heat shock (HS) for 5 and 15 min and nuclear and cytosolic fractions analysed.

The results clearly indicate that DYRK2 depletion reduces nuclear accumulation of phosphorylated HSF1 at Ser230 and Ser236 compared to control after 5 min HS (**Figure 37**, lane 3 vs lane 4), but this is less pronounced after 15 min HS (**Figure 37** lane 6 vs lane 7). This could be due to other kinases known to phosphorylate these residues becoming involved after a longer time period as described earlier in **Figure 25** (e.g. p-Ser326 by MAPK and p-Ser320 by PKA). DYRK2 depletion also reduced overall nuclear levels of HSF1 after HS compared to control siRNA, which implies a role for DYRK2 in the PTM of HSF1 to facilitate nuclear translocation and engage with target genes, also described earlier in **Figure 25**. DYRK2 depletion has less of an effect on cytosolic levels of HSF1, and phosphorylated HSF1 is notably lower, presumably because the PTM of these serine residues results in movement of HSF1 from the cytosol to the nucleus.

The experiments described above performed by de la Vega provide a compelling case for a link between HSF1 and DYRK2, and together with the clinical data from TNBC patient samples provided by Joanne Edwards, suggest that DYRK2 is a viable target in this BC subtype. Developing a selective DYRK2 small molecule inhibitor would help identify whether pharmacological intervention at this kinase provides a way forward towards a new type of therapeutic agent to treat patients with TNBC that overexpress DYRK2.

1.6 DYRK inhibitors

Taking a closer look at the DYRK2 active site and with the aid of the DYRK2 crystal structure from the protein data bank (PDB: 4AZF), we can highlight some of the important residues to target with possible interactions when designing ATP mimetic inhibitors, **Figure 38**.

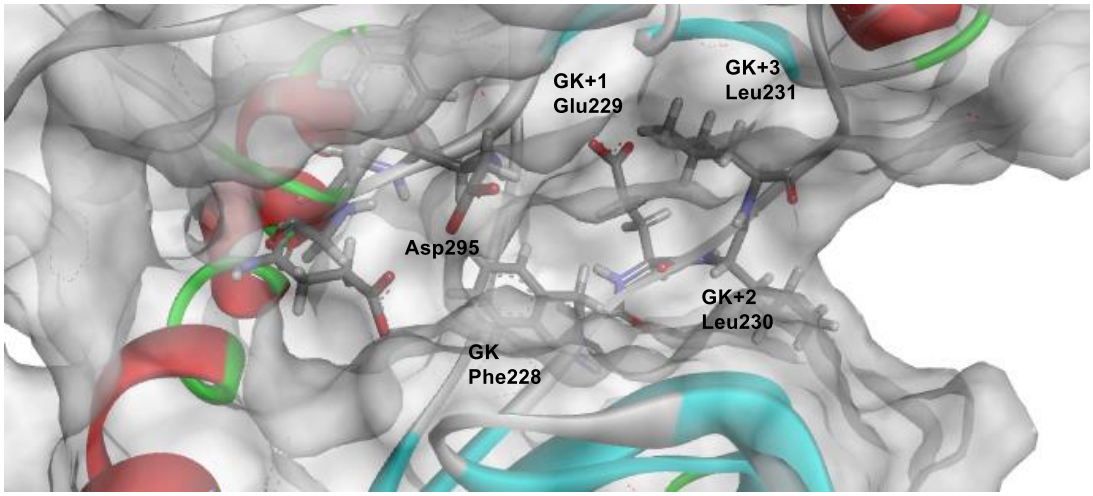


Figure 38: DYRK2 active site and some key AA residues (PDB: 4AZF).

Some of the key residues include the GK Phe228, a significant residue situated in the hydrophobic back pocket of the active site which can govern selectivity. More specifically, the back pocket of the DYRK2 active site is rich in hydrophobic residues such as Ile212 and Ile294. The hinge region AA residues are Glu229 (GK+1), Leu230 (GK+2) and Leu231 (GK+3), which can interact with inhibitor molecules through

hydrogen bond interactions. In addition, the activation loop residues within DYRK2 are Asp295, Phe296 and Gly297, **Figure 38**.

With respect to known DYRK2 inhibitors, Dixon and co-workers confirmed that curcumin **17** and LDN **18** were ATP competitive inhibitors, as they were able to provide a co-crystal structure of each in the DYRK2 active site, **Figure 39**.

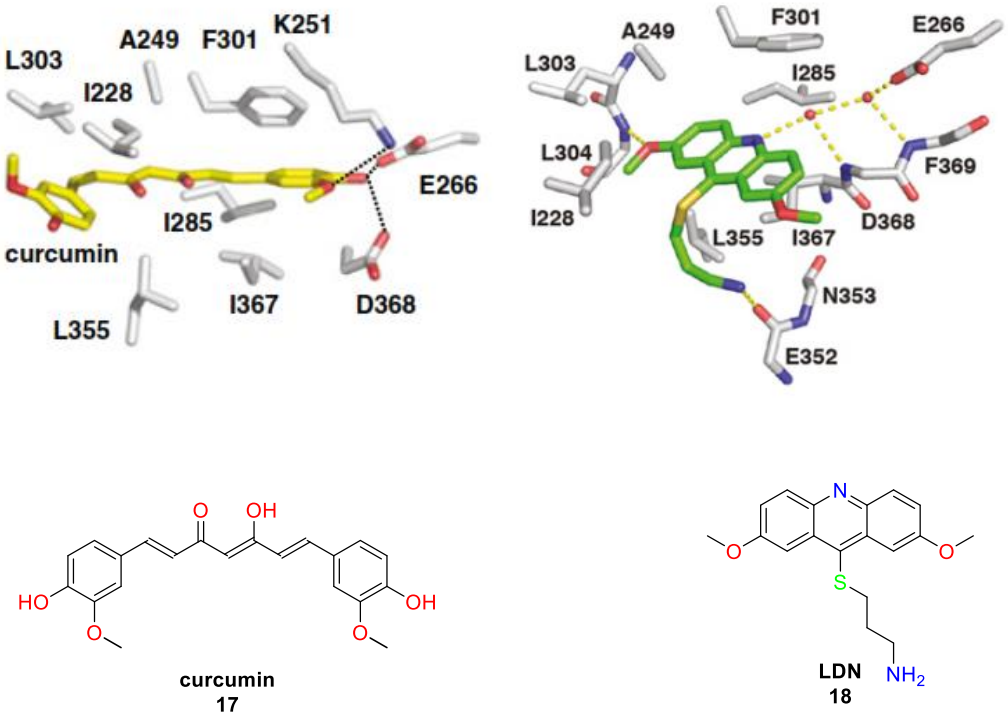


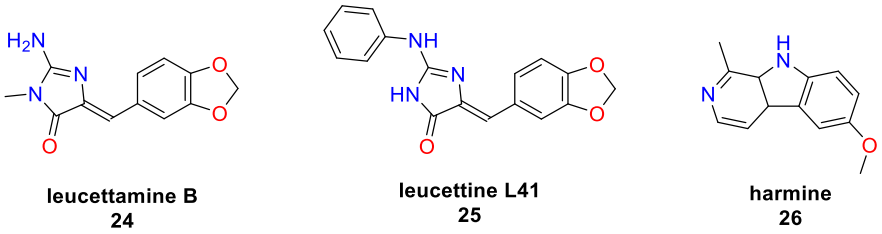
*Figure 39: **LHS**: curcumin **17** in DYRK2 active site (PDB: 6HDR).*

***RHS**: LDN **18** in DYRK2 active site (PDB: 6K0J).*[91]

There are common ATP mimicking interactions that are considered when designing kinase inhibitors, as briefly discussed in **Section 1.1.3**. The most established is the sequential HBD and HBA mimic from the adenine ring of ATP which interacts with the hinge region. However, with respect to curcumin **17**, it does display the adjacent HBD and HBA motifs but they are picking up interactions in a different area in the active site. More specifically, one of the hydroxyl groups makes hydrogen bonding interactions with an aspartic acid (Asp368) and a glutamic acid (Glu266). The neighbouring methoxy group interacts as a HBA with a lysine residue (Lys251) in the active site. In addition, curcumin **17** forms hydrophobic interactions with isoleucine and leucine residues, **Figure 39**. On the other hand, LDN **18** forms interactions with different amino acids in the DYRK2 active site. A methoxy residue of LDN **18** forms a hydrogen bond interaction with Leu303. The pyridyl nitrogen atom forms a through

water interaction with Glu266 and the primary amine moiety forms a hydrogen bonding interaction with Glu352, **Figure 39**. Similarly, LDN **18** is sandwiched between two hydrophobic residues, Leu355 and Ile367.[91]

At present, there are known inhibitors of DYRK2 that are derived from natural products, **Figure 40**.



leucettamine B
24

leucettine L41
25

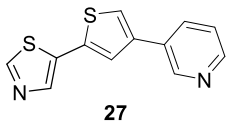
harmine
26

Kinase	Leucettamine B	L41	Harmine
DYRK2	2000	73	120
DYRK1A	600	60	85
DYRK1B	4100	44	28
DYRK3	1400	320	210
DYRK4	>10 000	520	9500
CLK1	66	71	72
CLK4	150	64	50

Figure 40: Natural product derived DYRK2 inhibitors. Associated IC_{50} values in nM.[112]

Mejier and co-workers discovered that DYRK1A inhibition by L41 inhibits neurodegeneration.[112] This also revealed the activity of structurally similar compounds against the rest of the DYRK family. They examined the activity of natural products leucettamine B **24**, a closely related analogue L41 **25** and harmine **26** against a panel of kinases *in vitro* and discovered that L41 **25** was highly potent against all members of the DYRK family, **Figure 40**. The parent alkaloid, leucettamine B **24**, was less active against all DYRK members, whereas harmine **26** was slightly less potent than L41 **25** for DYRK2 but exhibited a similar potency towards the other DYRK isoforms. With respect to L41 **25**, it proved highly potent for other kinases including CLK1 and CLK4, which are also members of the CMGC family **Figure 40**. In addition, L41 **25** displayed moderate activity for GSK-3α/β with IC_{50} values of 210 and 380 nM, respectively,[112] reinforcing the difficult selectivity challenge associated with this target.

Schmitt *et al.* discovered a novel DYRK inhibitor which inhibited DYRK1B in cells and promoted the generation of reactive oxygen species. Through high throughput screening at 10 μM, a hit with a thiophene core was identified, and through SAR iterations a highly potent inhibitor was developed, **27**, **Figure 41**.[113]

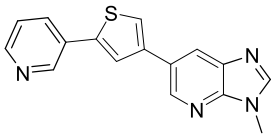


27

DYRK2 *IC_{50} 40 nM*
DYRK1A 100 nM
DYRK1B 70 nM

*Figure 41: Thiophene **27**.*[113]

27 was highly potent against DYRK1A, 1B and 2 with IC_{50} values of ≤100 nM, **Figure *41* 41**. Thiophene **27** was also active against common off target kinases CLK1 (94% inhibition) and CLK4 (99% inhibition) at 5 µM. Through the use of docking template studies, it was believed that the pyridyl and thiazole HBA's are an important feature for binding. Moreover, the high level of aromaticity can allow for potential π-interactions with amino acid residues in the DYRK active site. Interestingly, their SAR studies revealed a less potent compound with DYRK2 selectivity, **28**, **Figure 42**.[113]



28

DYRK2 *IC_{50} 600 nM*
DYRK1A 37% inhibiton (@ 5 µM)
DYRK3 13% inhibition (@ 1 µM)

*Figure 42:Thiophene **28**.*[113]

Although no further investigation has been reported to probe this class of inhibitors for DYRK2, it is an encouraging step towards the development of a selective DYRK2 inhibitor.

From these examples of DYRK inhibitors, it is clear that the DYRK2 active site is tolerant of different scaffolds. However, from a similarity perspective, all of these examples have a molecular weight (MW) of <400 and all but the thiophene series have both HBD and HBA capabilities. A common strategy in drug design is to follow the Lipinski rule of 5. These rules are used to increase the likelihood of oral bioavailability, which include MW<500, logP<5, HBD<5 and HBA<10.[114] Thus, these reported inhibitors fall within this criteria. However, the thiophene series lacks HBD motifs, which decreases the hydrophilicity and could penalise the aqueous solubility of these inhibitors.

To date there are no reported potent and selective inhibitors for DYRK2. The kinase active site is highly conserved and especially between DYRK isoforms, the differences in the active sites are negligible. This challenge was highlighted in **Section 1.2.1** where we briefly discussed the similarities in the AA sequence of the kinase domain of the DYRK family. However, there are some subtle differences which may mean that gaining selectivity between the isoforms is not impossible.

1.6.1 Selectivity

The main differences between DYRK2 and DYRK1A are observed in the *N*-terminal and *C*-terminal lobes. More specifically, the *N*-terminus of DYRK2 (right structure, **Figure 43**) has additional alpha helices.

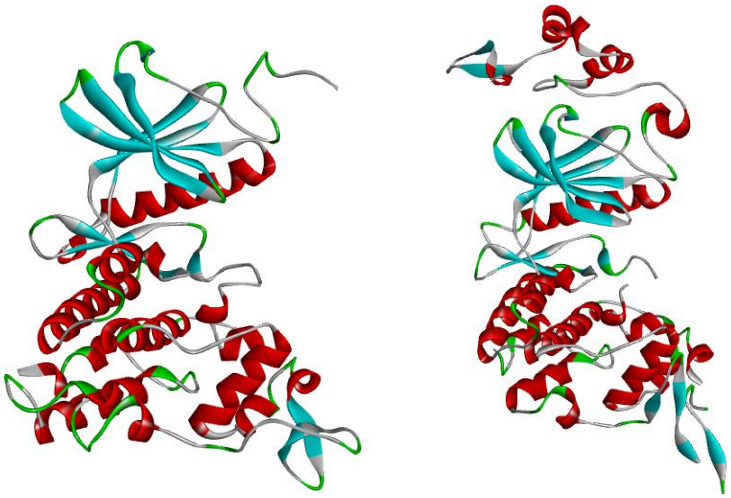


*Figure 43: Protein structures of **LHS**: DYRK1A (PDB: 4AZE) and **RHS**: DYRK2 (PDB: 4AZF).*

However, there have been subtle differences reported within the active site of DYRK1A and DYRK2. In particular, there are three residues that are different between each isoform. In DYRK1A, the Val222 residue is replaced by an isoleucine in DYRK2 (Ile212), and the Met240 residue of the DYRK1A hinge is a leucine (Leu230) in DYRK2. Additionally, the amino acid preceding the DFG motif within the active site, as this is Val306 in DYRK1A and Ile294 in DYRK2.[112] Two out of the three of these differences can be seen in the crystal structures of L41, **25**, bound DYRK1A and DYRK2, **Figure 44**.

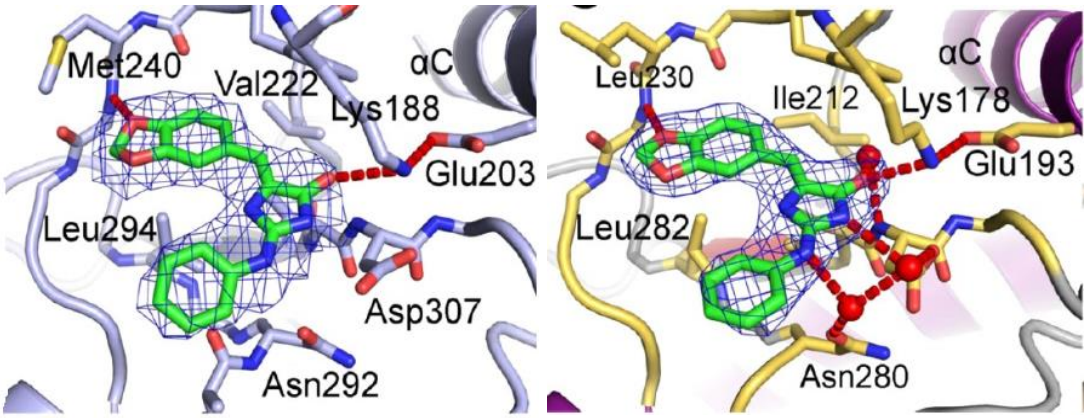


*Figure 44: L41 (**25**) in the active site of **LHS**: DYRK1A (PDB: 4AZE) and **RHS**: DYRK2 (PDB: 4AZF).*[112]

From these differences, we can infer that the DYRK2 active site is slightly smaller in size (DYRK1A Val vs DYRK2 Ile). This can be further visualised with the aid of a superimposed model of the two protein crystal structures, **Figure 45**. Where DYRK1A is depicted in red and DYRK2 is depicted in green.

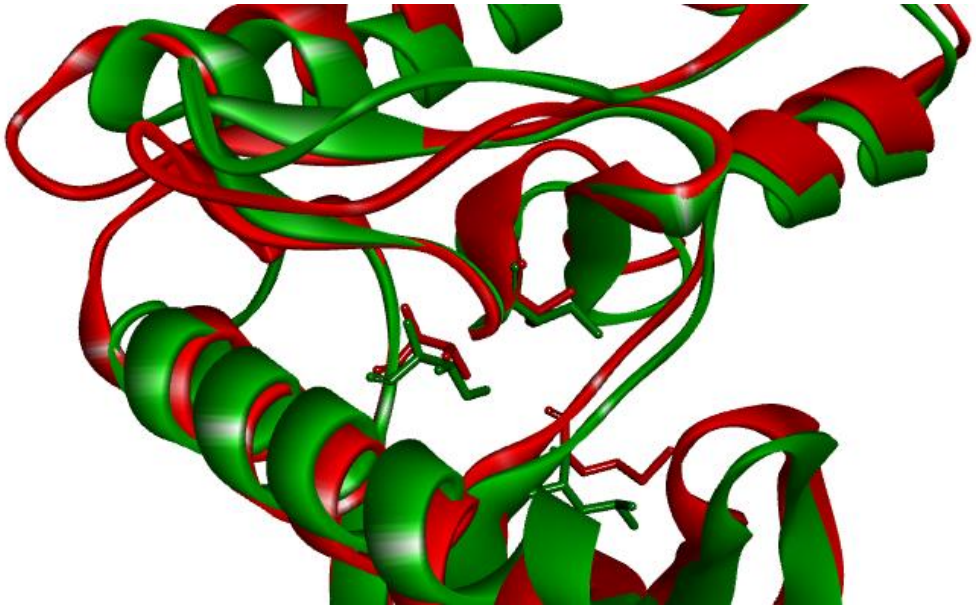


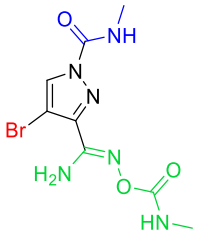
Figure 45: Superimposed DYRK1A (red) and DYRK2 (green) with differing AA residues in the kinase active site.

In addition to the isoform DYRK1A and with respect to the other DYRK family members, DYRK3 and 4, it is unclear what the specific structural differences are between them and DYRK2. As mentioned previously, less is known of these isoforms and there are no reported crystal structures.

Although achieving isoform selectivity appeared challenging, by utilising these subtle structural differences between the isoforms, there was initial promise that selectivity would be possible.

1.7 Previous work

Within the Tomkinson group an SAR profile of inhibitors for DYRK2 was under investigation. Through computational modelling a hit compound was identified, **C10** IC_{50} 2.5 µM, **Figure 46**.



C10
DYRK2 IC_{50} 2.5 µM

*Figure 46: Hit molecule via virtual screening **C10**.*

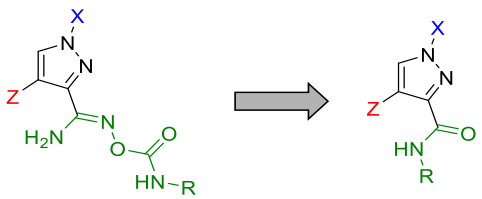
This scaffold was further interrogated and a structure activity relationship (SAR) profile was established. This research was undertaken by a previous PhD student within the group, Camille Indey. Herein, a brief overview of her synthetic approaches to the development of the DYRK2 inhibitor series is described. More specifically, the compounds which were synthesised and assessed externally for their activity towards DYRK2.[115]

The results Camille obtained were measured in the form of thermal stability *via* differential scanning fluorimetry (DSF) and reported as melting temperatures (T_m) from the Structural Genomics Consortium (SGC) in Oxford, and/or assessed by Life Technologies (LT) or Eurofins in the form of single point data (SP) (at 1 µM) or dose response curves (IC_{50}).

With respect to DSF, it is a simple, qualitative method to measure the binding of free small molecular weight compounds to a protein in solution. DSF measures the temperature dependant unfolding of proteins in the presence of a fluorescent dye which has an affinity for the hydrophobic residues of a protein, which are exposed upon its unfolding. The results obtained represent a change in the melting temperature (ΔT_m) which we can associate with binding events of the ligand to the protein.[116] Thus, we can infer that the greater the ΔT_m, the greater the binding affinity of the compound for the protein. Notably, this method does not provide information about the site on the protein where the compound is binding, only that a complex is formed to increase its thermal stability.

From **C10**, investigation began into three positions of diversification. First, the pyrazole *N*-1 substituent (**C10**: highlighted in blue) was explored, **Figure 46.** Alkyl and aryl substituents were investigated, with both secondary and tertiary substitution. Carbamates and amides were also studied, similarly with investigation of the tolerance of both alkyl and aryl groups. However, deletion of the urea altogether resulted in an increase in potency (IC_{50} = 1.1 µM). This was thought to be due to the hydrogen bond donor (HBD) ability of the unsubstituted N-H hydrogen atom *via* interaction with the DYRK2 hinge region.

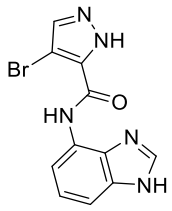
Second, the amidine type functional group was explored (highlighted in green **29**, **Figure 47**). This resulted in a number of derivatives being synthesised. However, complicated synthetic steps resulted in the generation of a simpler target, where the amidine moiety was substituted for an amide, **30**, **Figure 47**.



29 **30**

*Figure 47: Progress of manipulation of **C10**.*

A wide range of amide substituents were investigated, for example, aromatic, heterocyclic, cyclic and alkyl. The most potent functional group at this position was a benzimidazole moiety (**CI639**, IC_{50} = 223 nM, **Figure 48**).



CI639
DYRK2 IC_{50} 223 nM

*Figure 48: **CI639**.*

Third, the importance of the bromine at the 4-position on the pyrazole was briefly investigated. Removal of the bromine atom and substitution with a phenyl group were explored. In both cases, reduced activity was observed (IC_{50} = 2.5 µM and ΔT_m <1 °C). The presence of the bromine proved to be of great importance at this position. It was hypothesised that this could be the result of the electron withdrawing nature of the Br atom, pulling electron density away from the adjacent C-C bond and therefore

polarising the C-H bond and allowing it to act as a potentially weak HBD. On the other hand, from a steric perspective, the bromine atom could be occupying enough space in which to allow the molecule to sit in an optimum geometry for binding.

To summarise, from the initial screening hit **C10**, and through the investigation of three key areas of the scaffold, a new lead series of inhibitors was identified, **Figure 49**.

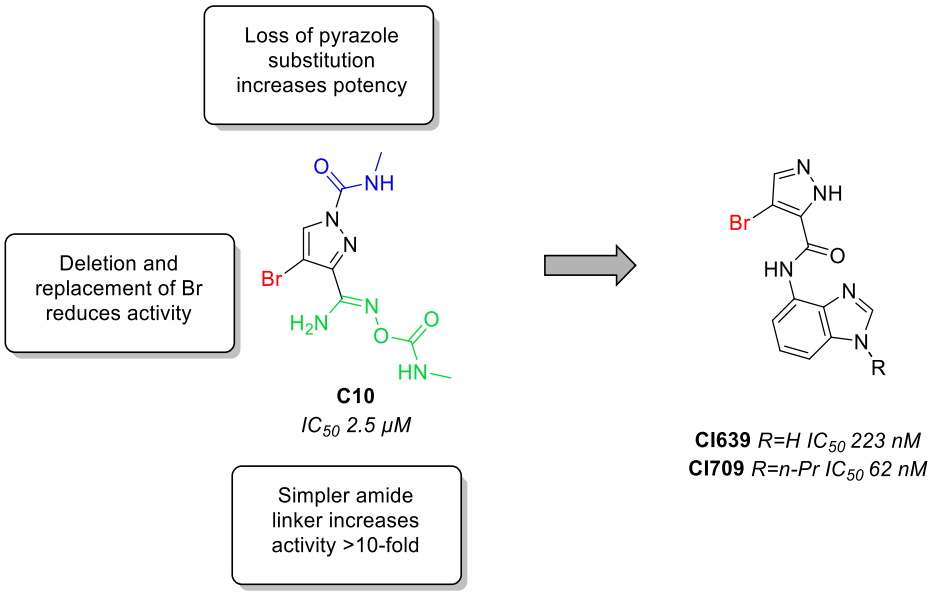


Figure 49: Brief summary of previous SAR.

The new lead series resulted in the removal of the *N*-substitution of the pyrazole and maintenance of the bromide on the pyrazole head group. In addition, the installation of an alternative tail group *via* an amide linker, **Figure 49**, where these modifications resulted in more potent DYRK2 inhibitors.

Through exploring different functionality at these three positions, a new lead compound was generated, **CI709**, **Figure 50**.

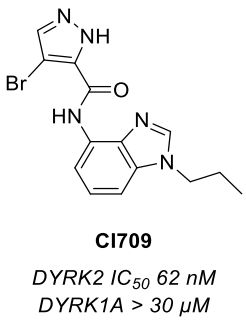


Figure 50: Indey optimised compound.

Based on docking studies that were performed by Camille, the following hypothesis of the possible interactions was formed, **Figure 51**.

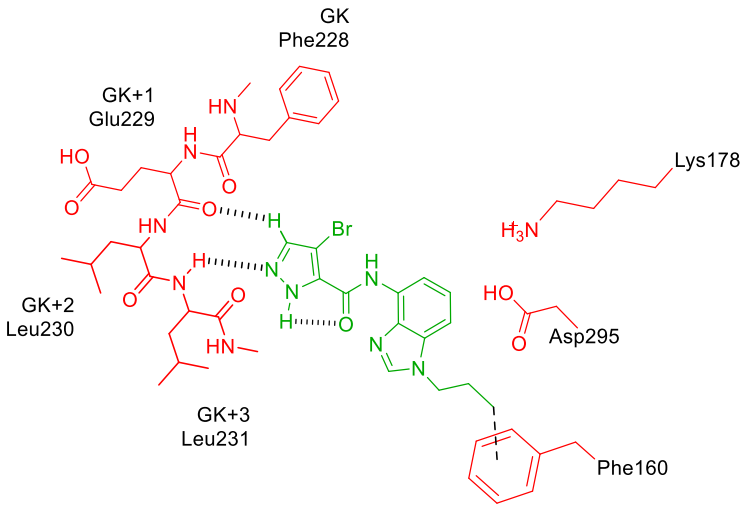


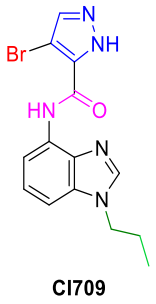
*Figure 51: Possible binding pose of lead compound **CI709** based on previous molecular modelling.*

The retention of the bromine moiety from the original scaffold **C10** was of great importance and, as previously discussed, we postulate that this can be either due to its electronics or its steric size. More specifically, it could potentially be generating a weak HBD out of the adjacent aromatic C-H. In addition, the bromine points into a small selectivity pocket within the active site and by occupying maximal space this could contribute to the molecules increased binding potential. Furthermore, by removing *N*-substitution on the pyrazole ring, it was thought that a hydrogen bond interaction was being made between a nitrogen atom of the pyrazole ring and Leu231 (GK+3) in the DYRK2 active site, **Figure 51**. The importance of the substituent present on the benzimidazole ring was not fully understood. However, with the aid of molecular modelling, it was postulated that there could be a potential interaction with a phenylalanine residue (Phe160) through CH-pi interactions.

Finally, to determine the selectivity of **CI709**, it was screened against a panel of 42 kinases including DYRK1A at 1 µM. Pleasingly, **CI709** proved to be highly selective toward DYRK2. With respect to 41 of the kinases investigated, no more than 32% inhibition was observed (CLK4). As previously mentioned, CLK4 is also a member of the CMCG family and a known off target of the DYRK family.[112,113] The associated selectivity data can be found in **Appendix 10.1**. Furthermore, in order to fully examine the activity of **CI709** and confirm its selectivity, dose response experiments with the highly similar isoform DYRK1A were carried out. **CI709** provided a calculated IC_{50} of >30 µM with respect to DYRK1A. Therefore, **CI709** was 500 fold more selective for DYRK2 over DYRK1A.

1.8 Aims and objectives

A potent and selective inhibitor for DYRK2 had been developed, **CI709**, **Figure 52**. However, the reason for its impressive selectivity profile required investigation. Moreover, in order to assess this inhibitor series in a cell model, an inhibitor with increased potency whilst retaining selectivity was required.



CI709

*Figure 52: Lead compound **CI709**.*

Therefore, the aim of this project was to further interrogate the tolerance of functionality at four areas of the inhibitor scaffold. More specifically; the 4-position of the pyrazole (red), determine the importance of the amide linker (pink), investigate the limit to the substitution pattern on the benzimidazole (green) and to examine the significance of the heterocycle head group (blue), **Figure 52**.

Once synthesised, the **CI709** analogues would be assessed for their biochemical activity, either externally by Life Technologies or *via* in an in-house *in vitro* DYRK2 assay. Successful compounds (K_i <60 nM) would be tested for their *in vitro* physicochemical properties i.e. half-life, solubility and permeability. Moreover, the lead compound would be tested in a TNBC cell based assays, by our collaborator de la Vega, to determine whether target engagement was evident and its effect on cancer cell survival.

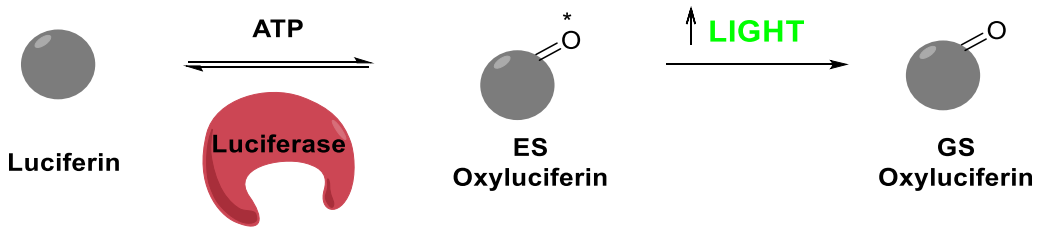
A 10-fold increase in potency whilst maintaining selectivity among kinases, and also identifying what governs the selectivity of this inhibitor series were the main aims at the start of this project. Moreover, we hoped to provide an inhibitor fit for assessment in TNBC cells.

2 Biochemical assay

Through collaboration with biologists within the Strathclyde Institute for Pharmacy and Biomedical Sciences (SIPBS), an *in vitro* screening assay for DYRK2 was developed.

2.1 *In vitro* ADP-Glo™ kinase assay

The Promega ADP-Glo™ kinase assay can be used as a high throughput screening method to determine kinase inhibition by quantifying the amount of ADP released during phosphorylation. The assay is performed in 3 steps. The first step involves the reaction of the kinase enzyme, ATP, substrate ± inhibitor. Next, the ADP-Glo™ reagent is added in an equal volume to the initial assay volume, to terminate the enzyme reaction and remove any excess ATP. Finally, the kinase detection reagent is added, which converts ADP to ATP. This allows the amount of ATP, and therefore the ADP generated in the first step as a result of kinase activity, to be measured *via* a luciferin/luciferase reaction (**Scheme 1**).



Scheme 1: Illustration of luciferin/luciferase reaction.

Luciferase catalyses the reaction between luciferin and ATP to form oxyluciferin in a singlet excited state (ES). The decay of ES oxyluciferin to ground state (GS) oxyluciferin releases a photon and thus emits light.[117] The light generated from this reaction is measured *via* luminescence, which can be correlated to the amount of ADP produced.

The assay kit contains 500 µL of Ultra-Pure ATP (10 mM), 500 µL of ADP (10 mM), 5 mL of ADP-Glo™ kinase reagent, 10 mL of kinase detection reagent and lyophilised kinase detection substrate.

2.2 Optimisation – Investigating K_m and V_{max}

Louise Young and Gillian Berrie developed a robust *in vitro* screening assay for DYRK2 using the Promega ADP-Glo™ kinase assay kit. After this optimisation and the assay was demonstrated to be reproducible, I was given the opportunity to learn the skills and carry out the assay myself. Prior to running the assay, it was necessary to determine the optimum ATP concentration. This was determined from a K_m and V_{max} study of the ATP, substrate and enzyme. In our case, DYRKtide, a synthetic peptide specifically designed for DYRK phosphorylation, was the chosen substrate. The V_{max} is the point at which the enzyme is saturated with substrate (ATP) and the K_m is the concentration of substrate required to achieve half of the V_{max}. Previously, during assay optimisation, the K_m of ATP was determined to be 37 µM for DYRK2.[118]

An interesting review by Knight and Shokat discussed the discrepancies between different reported ATP K_m values for the same kinase. More specifically, differences in K_m have been attributed to the particular assay conditions and the conformation of the kinase. Due to the dynamic equilibrium that exists between the active and inactive forms of protein kinases, this can have an influence on the ATP affinity and therefore the ATP K_m.[119]

To date, there have been no reported ATP K_m values for DYRK2 using the *in vitro* ADP-Glo™ kinase assay. However, ProQuinase have reported an ATP K_m for DYRK2 of 0.71 µM using the FlashPlate™-based radiometric assay,[120] which is significantly less than the 37 µM reported for the developed ADP-Glo™ kinase assay. In order to further highlight the acceptable differences in the K_m of ATP, associated data for its closely related isoforms DYRK1A and DYRK1B have been reported. More specifically, a study by Alexeeva *et al.* reported that the K_m of ATP for DYRK1A and DYRK1B were 118.5 and 80.7 µM, respectively.[121] These were measured using an ATP-regenerative NADH consuming assay.[122] Another study by Ogawa *et al.* reported the K_m of ATP for DYRK1A to be 37 µM. In this case, the K_m value was measured using an *in vitro* radioactivity based assay,[123] giving a significantly lower value than that reported by Alexeeva *et al*. Therefore, it is apparent that observing differences in the K_m of ATP for a particular kinase is not uncommon under different assay conditions.

With this in mind, we began to investigate the K_m and V_{max} of ATP with newly purchased DYRK2 and ATP. As the K_m was determined to be 37 µM, we chose a concentration range of ATP to investigate, examining concentrations of >10×K_m–

0.1×K_m. More specifically, the ATP concentration range of 500–0.5 µM, with half serial dilutions, **Figure 53**.

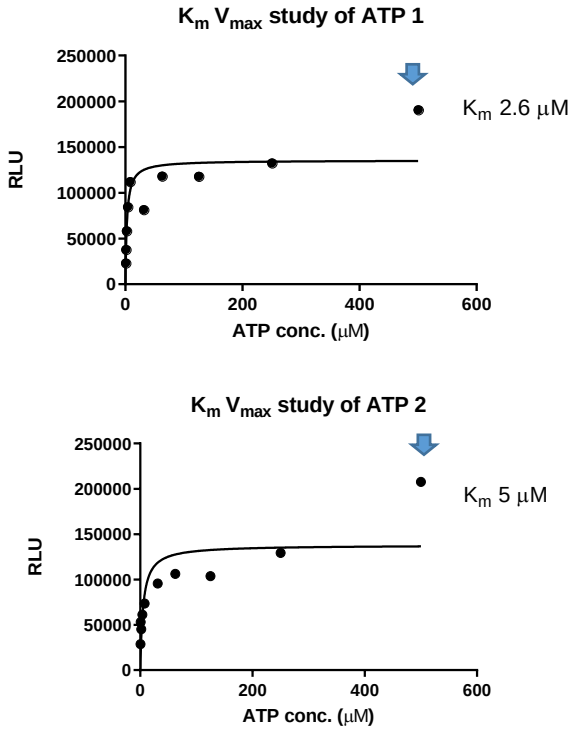


Figure 53: Results from Km Vmax study 500–0.5 µM.

The first experiment was carried out in duplicate and the K_m was determined to be 2.6–5 µM, using Graph Pad prism software, **Figure 53**. Furthermore, at 500 µM the relative luminescence units (RLU) were typically higher and therefore we inferred that this was a concentration of ATP greater than the V_{max}. Consequently, the concentration range of ATP was reduced to 300–0.3 µM.

The experiment was carried out twice more, on separate days and the K_m of ATP was determined to be 6 µM (**Figure 54**).

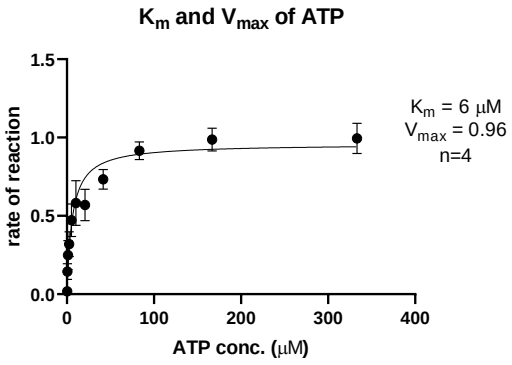


Figure 54: Results from n=4 K_m and Vmax study.

Across the individual experiments, the calculated K_m values ranged from 2.6–26 µM, **Figure 55**.

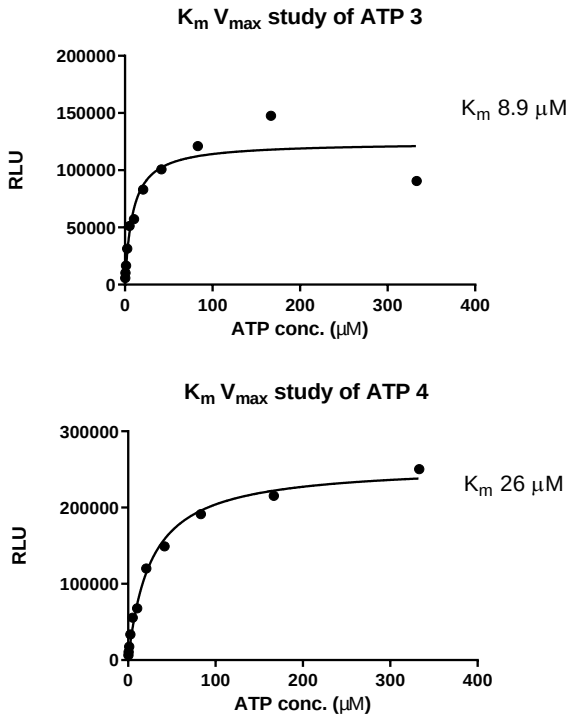


Figure 55: Individual results from K_m Vmax study.

The calculated K_m values were lower than the previous value of K_m of 37 µM. Therefore, we assessed a compound from our library (**LB35**) previously shown to have a K_i value of **19 nM**, with varying concentrations of ATP, to determine if the results were consistent.

We used three concentrations of ATP: 2.5, 5 and 10 µM respectively. Dose response experiments were carried out at 3000–1 nM and the associated K_i values were calculated at each concentration in duplicate using GraphPad prism software, **Figure 56**.

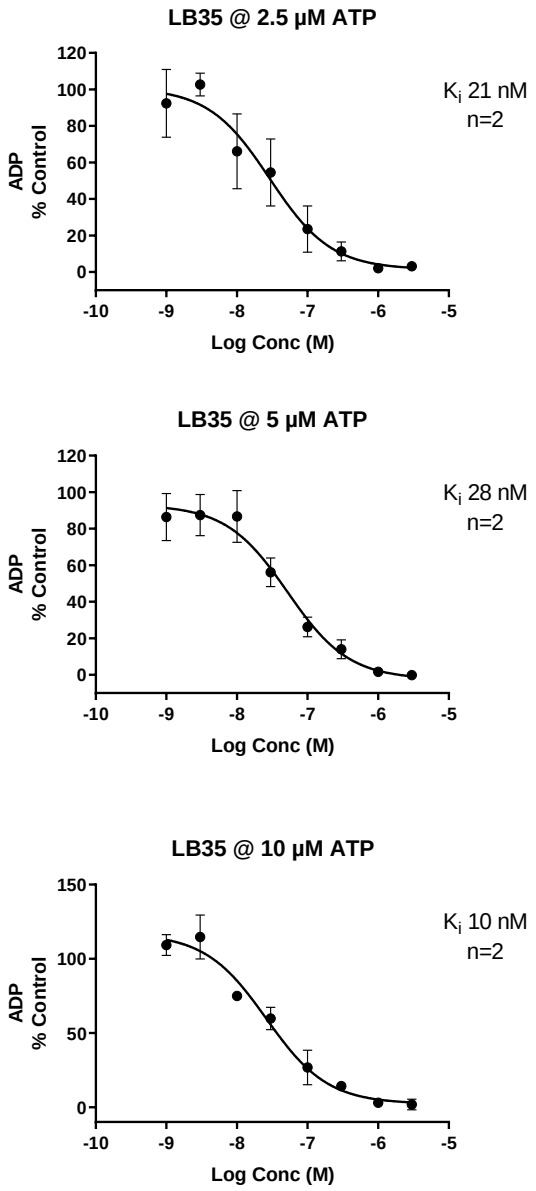


Figure 56: Testing ATP concentrations with LMB035.

From the results, we determined that an ATP concentration of 5 µM for the new batch of DYRK2 gave the most reproducible results. Furthermore, this concentration provided a similar K_i value to that previously reported for **LB35** (28 nM vs 19 nM). At 2.5 µM ATP, a similar K_i, 21 nM **Figure 56**, was observed, however, there appeared to be more variability in the results. At 10 µM ATP the observed K_i for **LB35** was 10 nM, almost half the previous value we reported (K_i 19 nM).

The results for each compound were determined by GraphPad prism software and the associated K_m and concentration of ATP for the experiments were used in order to calculate the K_i values, with respect to the Cheng-Prusoff equation (see **Appendix 10.3**).

2.3 Conclusions

As briefly discussed in the previous section, differences in the K_m of ATP for the same kinase is not unexpected. Although the same assay conditions were adhered to and the same method was followed, a new batch of enzyme was purchased. This batch, when present in the assay reaction mixture, may have a different ratio of the active: inactive form. This could therefore result in a change in affinity for ATP, and hence the observed difference in the K_m. However, it should be noted that DYRK2 is synthesised and purchased as the active form of the enzyme and therefore the affinity for ATP is unchanged. Alternatively, a possible reason for the observed difference in K_m of ATP is the purity of the enzyme. Varying levels of purity could result in a change in the optimum concentration of ATP required for the experiment.

For clarity, it should be noted that within the results section, the analogues stated within **31**–**65**, were assessed externally by Life Technologies and therefore the dose response results were quoted as IC_{50} values. Analogues **31**–**82** were assessed with the first batch of DYRK2 enzyme and finally **50**–**53** and **87**–**92** were assessed with the second batch of enzyme *via* the in-house assay and the results are quoted as K_i values.

3 Results and Discussion

Previous SAR studies generated a potent and selective DYRK2 inhibitor **CI709** (IC_{50} 62 nM), **Figure 57**. However, it was unclear which functionalities contributed to both potency and selectivity. Therefore, one aim of this project was to determine the importance of each substituent whilst improving the potency and maintaining the selectivity.

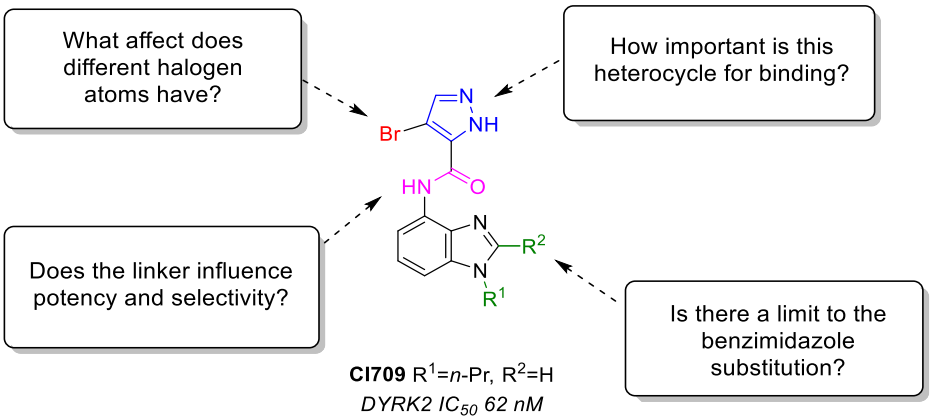


Figure 57: Inhibitor scaffold for development.

From a closer look at the general scaffold of **CI709**, there were four main positions that we aimed to investigate. First, from the previous work, substitution on the NH of the pyrazole (blue, **Figure 57**) diminished DYRK2 activity (single point (SP) data at 1 µM), where methyl (3% inhibition) and methoxyethyl (6% inhibition) analogues were synthesised. Interestingly, previous work had demonstrated that the addition of a urea-type functional group at this position generated active compounds (methyl urea-type IC_{50} 70 nM), however, the unsubstituted pyrazole analogues provided comparable results. Therefore, for ease of synthesis, we did not explore urea functionalisation further. In order to establish whether the HBD/HBA ability of the pyrazole was important for binding and potency, we proposed to replace this moiety with different heterocycles.

In addition to a bromide at the 4-position of the pyrazole (red, **Figure 57**), previous work revealed that substitution with a phenyl group abolished activity (ΔT_m <1 °C). Removal of the bromide also resulted in a less active inhibitor (IC_{50} 2.5 µM), although this inhibitor contained a nitrophenol substituent rather than the benzimidazole in **CI709**, **Figure 57**. Thus, we hypothesised installing different halogens at this position to further interrogate the importance of the bromide.

With respect to the amine partner of the amide, previously a range of substituted aromatics were synthesised and tested. Nitrophenol substituted amines were moderately active (IC_{50} 1.45 µM) and benzoxazole analogues provided similar activity (1.71 µM). However, the benzimidazole amine provided the most active inhibitors, as shown by the unsubstitued benzimidazole (IC_{50} 223 nM), *N*-methyl substituted (IC_{50} 141 nM) and *N*-propyl substituted (**CI709**, IC_{50} 62 nM) derivatives. Finally, arrival at the amide linker from the original hit was determined *via* the search for an analogous series that was easier to synthesise. Pleasingly, this linker choice provided inhibitors of increased potency. In this work, we set out to determine whether the amide was essential for potency and also selectivity.

A summary of the key structures and SAR from previous work is shown in **Appendix 10.2**.

3.1 Model for inhibitor design

A rational model for compound design was followed, as previous docking studies with the existing library had proposed a binding orientation in the DYRK2 active site that matched the SAR series profile, **Figure 58.**

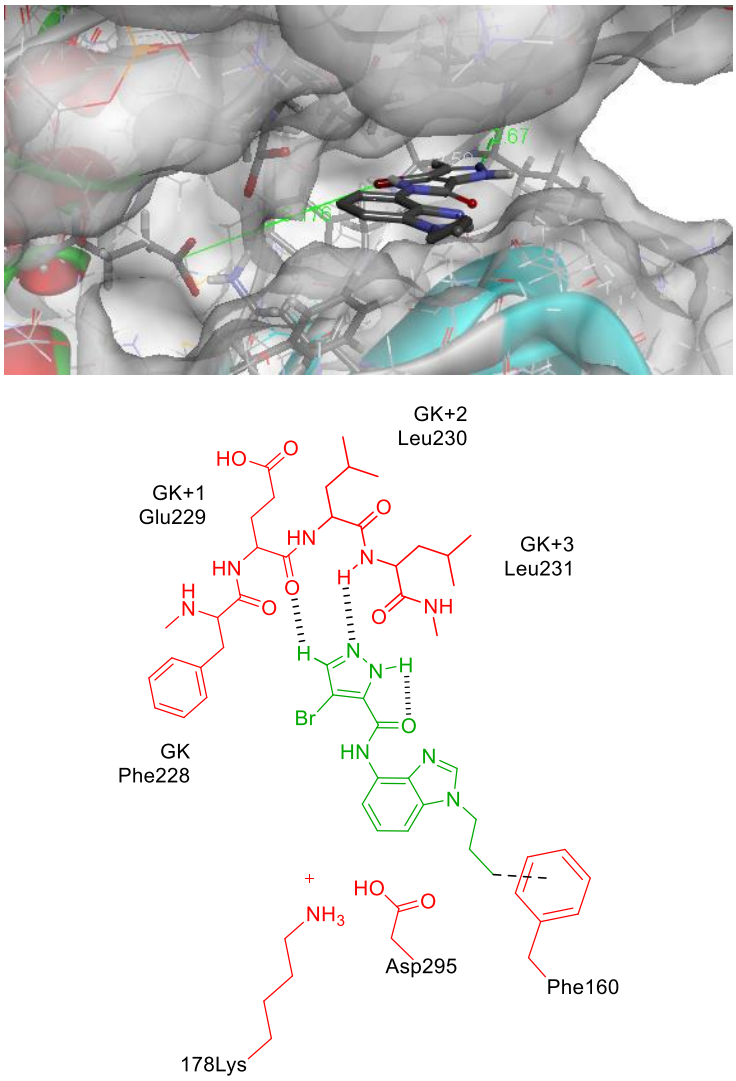


*Figure 58: Model of **CI709** in DYRK2 active site and potential key interactions (PDB: 4AZF).*

This modelled structure provided us with information to reinforce our hypotheses of the key interactions between our scaffold and the DYRK2 active site. First, the pyrazole can bind to the hinge region. Second, the bromine directs into the small hydrophobic back pocket of the DYRK2 active site formed between the GK residue Phe228, Ile212 and Ile294 and finally, the benzimidazole can interact with Phe160 and Asp295, with the *N*-substituent pointing in or out of the site, **Figure 58**.

The next sections will address the approaches to inhibitor design, biochemical evaluation, rationale and the synthesis.

3.2 Head group exploration

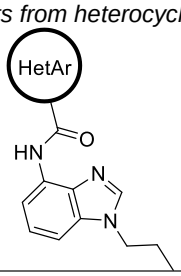
The importance of both the head group heterocycle and the halogen at the 4-position of the pyrazole were investigated.

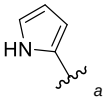
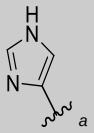
3.2.1 Heterocycle Investigation

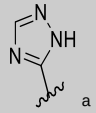
The effect of changing the head group on the activity of this inhibitor series was examined. It should be noted that the synthesis of some of these analogues was carried out by a Masters student within the Tomkinson group (Fiona Keatings) under my supervision. The synthesis of these analogues is described in **Appendix 10.6.2**. The compounds were then assessed externally for DYRK2 inhibitory activity by LT. See **Appendix 10.3** for a full description of the LT biochemical assay. Upon development of the in-house assay, some of the compounds were re-tested, at 1 µM, in order to confirm their activity, **Table 2**.

With respect to the in-house assay, for SP experiments at 1 µM, a threshold of 60% inhibition was set for determining which compounds would be tested in further dose response experiments.

Table 2: Results from heterocycle investigation.



Compound	HetAr	LT results[b] (% inhibition)	In-house results[c] (% inhibition)
31		12	48
32		6	-
33		17	-
34		IC_{50} >10 µM	23

35		4	47
36		1	20
37		19	-
38		9	44

*[a]synthesised by Masters student, [b]LanthaScreen assay n=1 (see **Appendix 10.3**), [c]ADP-Glo™ assay n=3, determined at 1 µM inhibitor concentration*

From these results, it was clear that changing the heterocycle was detrimental to the activity of this inhibitor series. Upon replacement with a single heteroatom heterocycle, furan **31** (LT, 12% inhibition) and pyrrole **32** (LT, 6% inhibition), the activity decreased tremendously. Further investigation with the robust in-house assay confirmed this for **31** (48% inhibition at 1 µM, **Table 2**). The reason for this drop in activity could be attributed to the lack of a second HBD/HBA motif and/or the lack of the bromine at the 4' position. To test the single heteroatom theory, heterocycles with more than 1 heteroatom were tested **33**–**38**, **Table 2**. The introduction of a second HBD/HBA moiety had little or no effect on the activity, imidazole **33** (17% inhibition), oxazole **37** (19% inhibition) and thiazole **38** (LT 9% inhibition, in-house 44% inhibition), **Table 2**. In addition, the methyl substituted imidazole **36** was even less active highlighting the importance of the HBD NH. Furthermore, introducing a third heteroatom, 1,2,4-triazole **35**, provided no significant change in activity (LT 4% inhibition, in-house 47% inhibition, **Table 2**). This was encouraging as it provided us with information on the importance of the HBD and HBA of the head group, and also the significance of the bromine at the 4′ position.

To explore this further, the pyrazole regioisomer **34** was prepared based on the assumption that it could provide a stronger HBD for interacting with the GK+1 Glu229 residue in the hinge region, **Figure 59**.

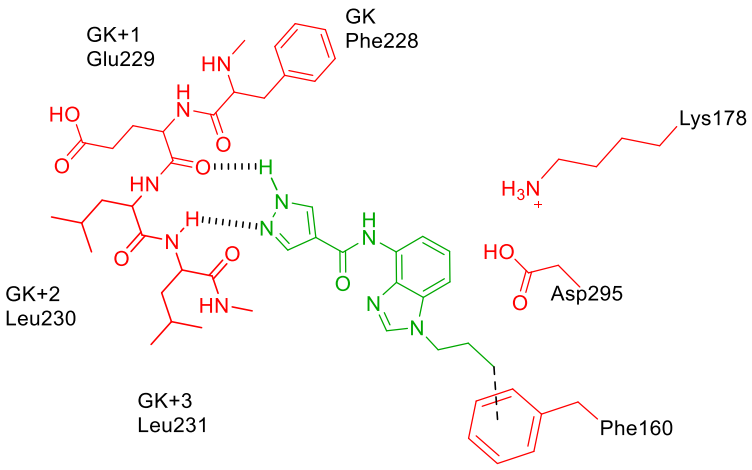


*Figure 59: **34** in the DYRK2 active site.*

However, the poorer activity for **34** (LT > 10 µM and in-house 23% inhibition at 1 µM, **Table 2**) compared with **CI709** suggests that a stronger HBD adjacent to Glu229 (GK+1) is not the driving force for high binding affinity and highlighted the importance of the substitution on the pyrazole, at the 5' position in this case. However, whilst pyrazole **34** provides a stronger HBD **Figure 59**, in order to fully examine this we must introduce the bromide at the 5′ position, but due to time restraints, this work was not carried out.

Next, the importance of the substitution on the original pyrazole head group was examined.

3.2.2 4′ Pyrazole substitution

3.2.2.1 Importance of halogen

We hypothesised that the bromine atom was of optimum size, and fits tightly into a small hydrophobic selectivity pocket bordered by Ile212, Phe228 and Ile294 in the DYRK2 active site. Moreover, the electron withdrawing (EWG) nature of Br also has the ability to polarise the adjacent C-C and therefore C-H bond, providing a weak HBD to interact with the HBA carbonyl group of GK+1 (Glu229), **Figure 60**.

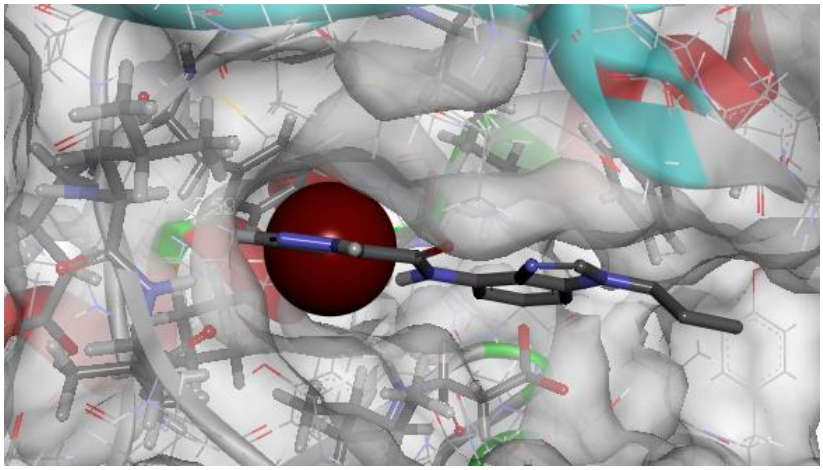
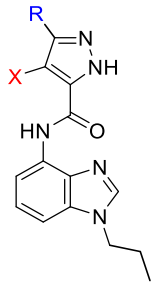


*Figure 60: Model highlighting selectivity pocket with Br of **CI709** shown in CPK format.*

In order to investigate this further, analogues with different halogen atoms at the 4′ position of the pyrazole were synthesised and assessed in the in-house *in vitro* biochemical assay, as described in **Section 9.1**. Interestingly, in the in-house assay, the activity of the lead compound, **CI709** was significantly lower (K_i 210 nM), **Table 3**. However, the previous result (IC_{50} 62 nM) from LT was the result of an n=1 duplicate experiment, whereas the in-house value was more robust.

Table 3: Results for DYRK2 inhibition.



Compound	X	R	DYRK2 K_i (nM)
39	F	Me	12% inhibition[a]
40	Cl	H	335
41 (**CI709**)	Br	H	210
42	I	H	430
43	H	H	>14000

[a]*SP experiment n=3 @ 1 µM*

All experiments were carried out in triplicate and the results are highlighted in **Table 3**. We postulated that the bromine atom could have dual importance. First, from an EWG perspective, we would assume that the order of activity would be as follows

F>Cl>Br>I. However, this sequence could not be fully investigated as the synthesis of the fluoro analogue, without substitution at the 5′ position proved to be challenging. The associated inactivity of the methyl substituted pyrazole **39** does, however, indicate that the C-H is important for binding (12% inhibition at 1 µM, **Table 3**). This is highlighted in the model in **Section 3.1** through its interaction with Glu229, although a methyl in this position is also likely to be sterically unfavourable. However, the activity of the chloro analogue (**40** K_i 335 nM) is less than its bromo counterpart (**41** (**CI709**), K_i 210 nM), which suggests that the activity cannot solely be described by polarisation of the adjacent C-H bond because Cl is more electron-withdrawing than Br. From a steric perspective, the order of activity is Br>Cl>I>H, **Table 3**, which implies that the pocket which is occupied by the halogen is too small for the larger iodide, which resulted in a decrease in activity (**41 CI709** K_i 210 nM vs **42** K_i 430 nM, **Table 3**. We therefore postulated that the bromine atom was the optimum size to occupy the back pocket of the active site, fitting snugly within the hydrophobic cavity formed by the side chains of Ile212, Phe228 and Ile294, **Figure 61**.

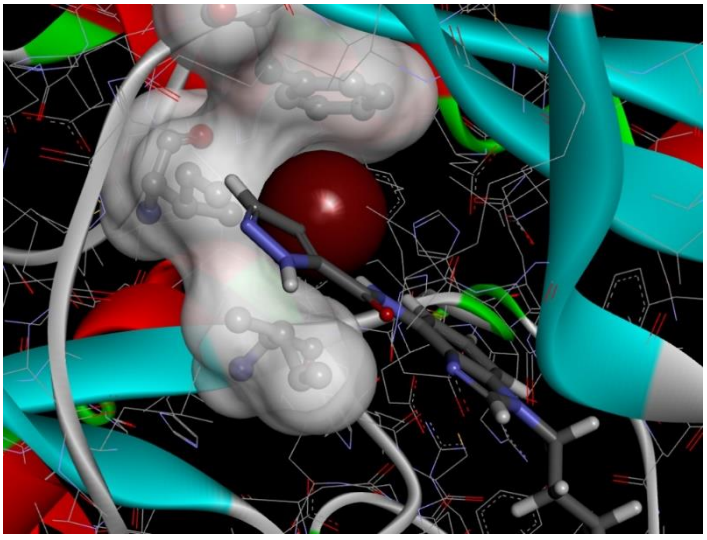


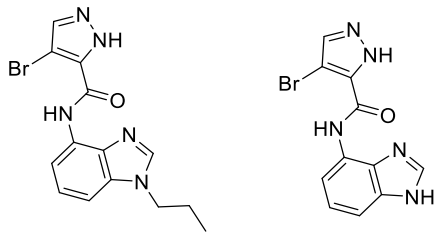
*Figure 61: Illustration of **CI709** snug fit in hydrophobic back pocket of DYRK2 active site.*

The lower activity of **40** and **43** which have smaller substituents (Cl and H respectively) reinforces the importance of a good hydrophobic fit between the ligand and the protein with this region of the binding site. Moreover, it is possible that the Br substituent makes the optimum halogen (σ-hole) interaction with the gatekeeper residue Phe228.[124] A halogen can interact with an acceptor motif which includes oxygen, nitrogen and sulfur atoms on AA residues. The electron cloud of a halogen in a C-X bond contains a region of electro-positivity (σ-hole) which can attract an

interaction with a site of electronegative potential, e.g. oxygen lone pair or the π-electrons of an aromatic side chain in Phe or Tyr.[124, 125]

3.2.2.2 An alternative binding mode

From previous work, we had two active inhibitors, the propyl analogue **CI709** (K_i 210 nM) and the unsubstituted analogue **CI639** (K_i 102 nM), **Figure 62**.



CI709
DYRK2 LT IC_{50} 62 nM
In-house K_i 210 nM

CI639
DYRK2 LT IC_{50} 223 nM
In-house K_i 102 nM

*Figure 62: Comparison of **CI709** and **CI639** potencies.*

CI639 was reported to have an IC_{50} of 223 nM as determined by LT. Upon further evaluation in the in-house assay, the K_i was determined to be 102 nM, **Figure 62**, and therefore, resulted in the order of activity being reversed (**CI639** >**CI709**). Moreover, from previous work, docking studies with **CI639** in the DYRK2 active site revealed an

additional conformation compared to the *n*-propyl analogue **CI709**, **Figure 63** where the benzimidazole had rotated, pointing into the catalytic site.

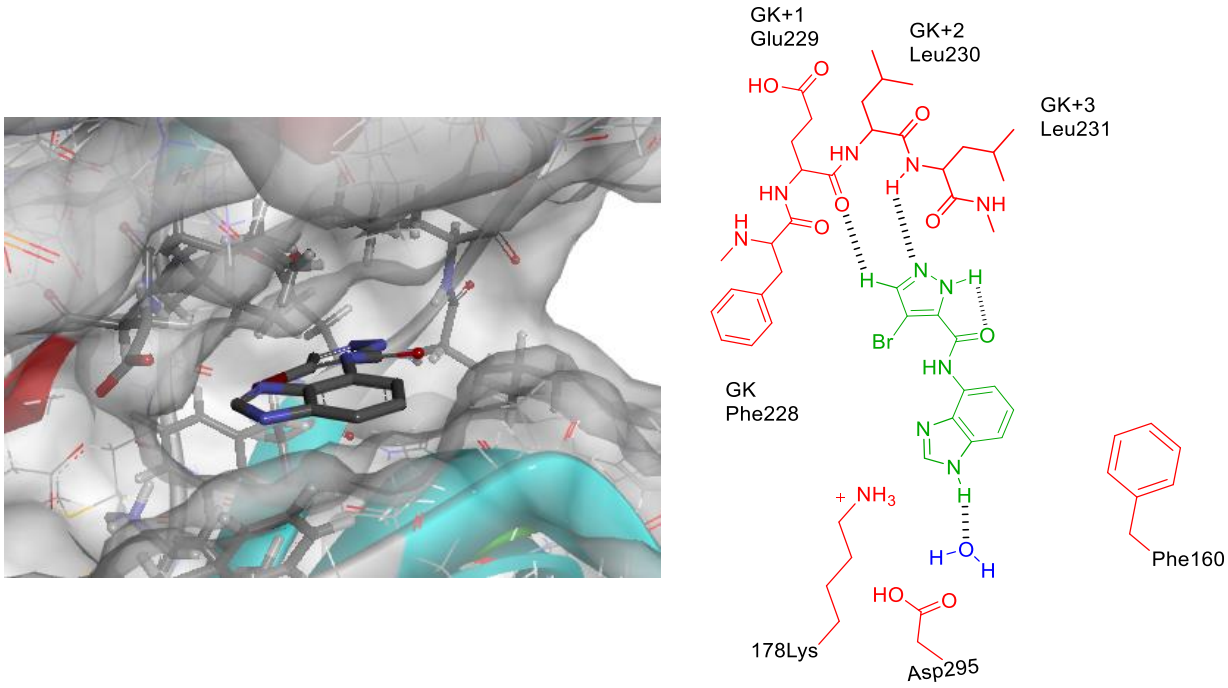
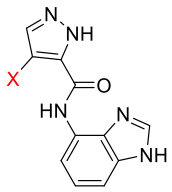


*Figure 63: Model of **CI639** and possible interactions.*

This allowed for the possibility of the NH of the benzimidazole to make a hydrogen bond interaction with a water molecule, **Figure 63**, which could be the reason for the increase in potency. Consequently, we prepared compounds to assess whether potencies across the halogen series were reproduced with unsubstituted benzimidazoles, **Table 4**.

Table 4: Results for CI639 analogues.



Compound	X	DYRK2 K_i (nM)
44	Cl	274
45 (**CI639**)	Br	102
46	I	44%[a]

[a]*result from LT, SP @ 1 µM.*

Pleasingly, these results complemented those shown in **Section 3.2.2.1** for the *n*-propyl benzimidazole series. The bromo analogue **45** (**CI639**) was the most potent (K_i 102 nM), followed by the chloro **44** (K_i 274 nM) and the iodo **46** (44% inhibition from LT). Furthermore, this supports our hypothesis that substituents larger than bromide were less accommodated by the DYRK2 active site and were therefore less active.

Next we looked at replacing the halogen with a similar group in order to unequivocally determine its importance.

3.2.2.3 Ethynyl analogue

A common strategy in medicinal chemistry is the introduction of bioisosteres. Bioisosteres are functional groups of a similar size and electron density which replace an existing group. They are often used to improve the physicochemical properties, such as permeability or to tackle toxicity. Some examples of classical bioisosteres of bromide are: chloride, hydroxyl and thiol.[126] Having already discovered the effect of the chloride analogues **40** and **44** (K_i 335 nM and 274 nM respectively) and with the potential for a challenging route to deliver the hydroxyl and thiol derivatives, we decided to investigate a non-classical ethynyl isostere of halides. The CH of the ethynyl group displays similar electronic properties to the CX bond with the potential to form a σ-hole interaction.[127] The two FDA approved drugs with similar scaffolds, gefitinib **47** and erlotinib **48**, provide an example of where this bioisostere replacement approach has been employed, **Figure 64**.

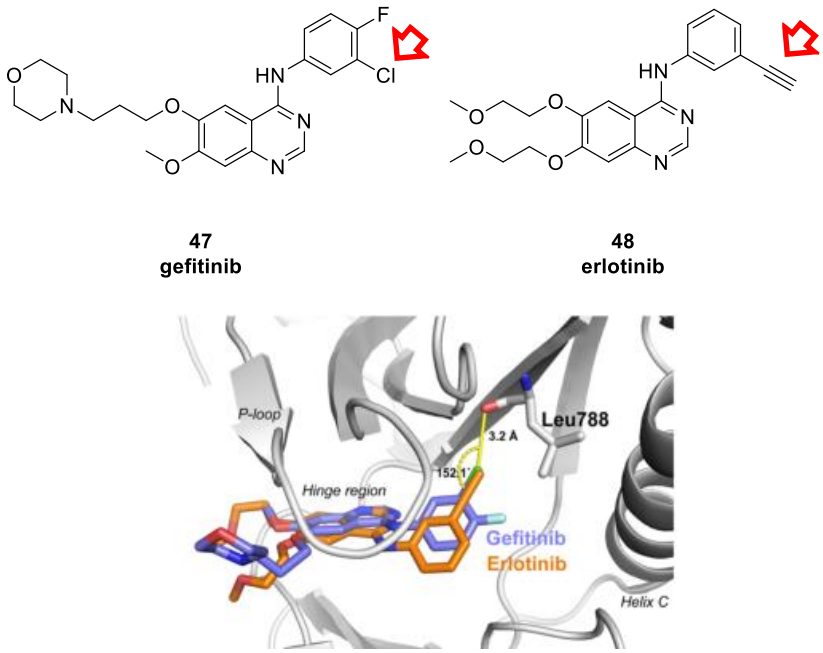
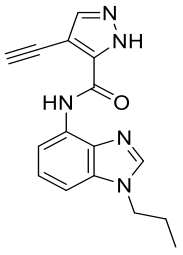


*Figure 64: Structures of gefitinib **47** and erlotinib **48**.*[128]

Gefitinib **47** and erlotinib **48** are EGFR (epidermal growth factor receptor) kinase inhibitors and are used for the treatment of lung cancer.[128] They both share the quinazoline core and a similar substitution pattern, with diversification at each position. However, when overlaid, the crystal structures of both inhibitors in the kinase active site highlights that the chloride and ethynyl are both pointing with the same vector and interacting with the same halogen bond acceptor, the C=O lone pair of Leu788 **Figure 64**.[127]

With this in mind and in order to determine whether a similar replacement of bromide with ethynyl, would maintain or increase the activity in our series, the ethynyl analogue **49** was synthesised and assessed in the *in vitro* assay, **Figure 65**.



49
DYRK2 K_i 510 nM

*Figure 65: Ethynyl analogue **49**.*

The ethynyl analogue was less potent than the bromide derivative (**41 CI709**, K_i 210 nM) and also less potent that the iodide analogue (**42**, K_i 430 nM). Moreover,

similar to the iodo analogue **42**, the ethynyl analogue was likely to occupy too much space in the small pocket of the DYRK2 active site, forcing the pyrazole hinge-binding group of **49** to rotate and become unable to interact with Glu193 or the gatekeeper Phe228, **Figure 66**.

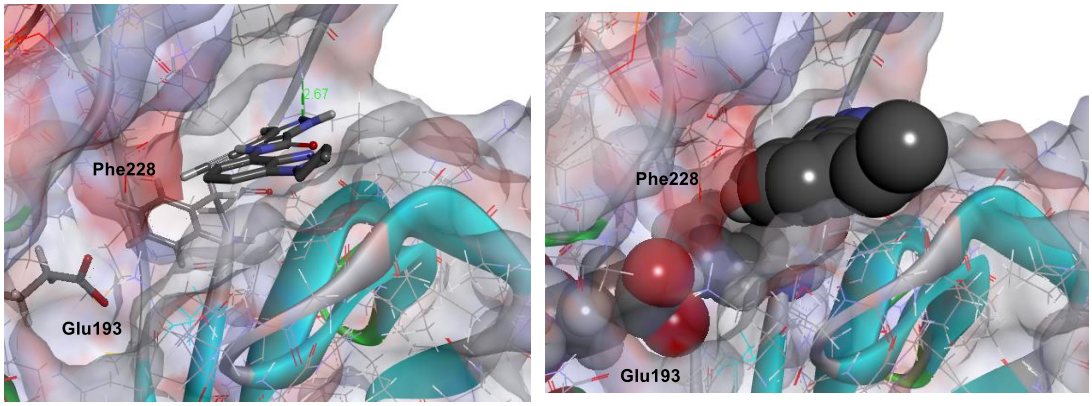


*Figure 66: Potential pose of **49** in the DYRK2 active site.*

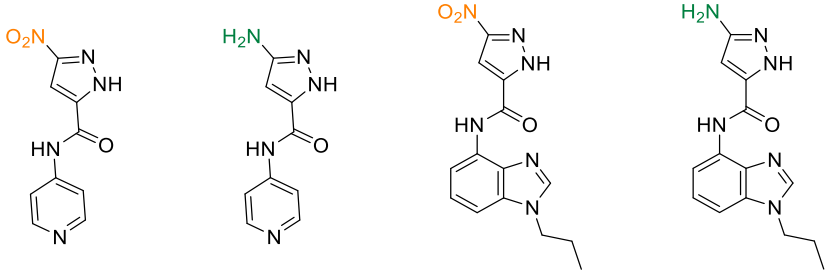
Upon docking with the protein structure and introducing the space fill feature, the CH of the ethyne is almost in direct contact with the pi system of Phe228, **Figure 66**, which could result in an unfavourable steric clash and a different binding pose for **49**.

Having established that a bromide at the 4′ position on the pyrazole head group was pivotal for potency, the next step was to investigate the effect of substituents at the remaining position of the pyrazole ring.

3.2.3 3′ Pyrazole substitution

Our existing SAR had suggested that the CH at the 3′ position made a weak hydrogen bond interaction with oxygen lone pair of GK+1 Glu229 in the hinge of the DYRK2 active site. Whilst the pyrazole regioisomer **34**, with the CH replaced by the stronger HBD NH, had not improved activity, we set out to examine if the introduction of a strong exocyclic HBD motif at this position would increase the affinity for these inhibitors. Our model suggested that an amino group in this position could align more closely with the C=O of Glu229, and in order to prepare the 3′ amino analogue **53**, a nitro precursor was synthesised **52**. To develop the route, for ease of synthesis, pyridine analogues **50** and **51** were synthesised initially, then the methodology was conveyed to the benzimidazole derivatives **52** and **53**. All were evaluated *in vitro* and the associated activities are stated in **Table 5**.

Table 5: Results for 3′ substitution derivatives.

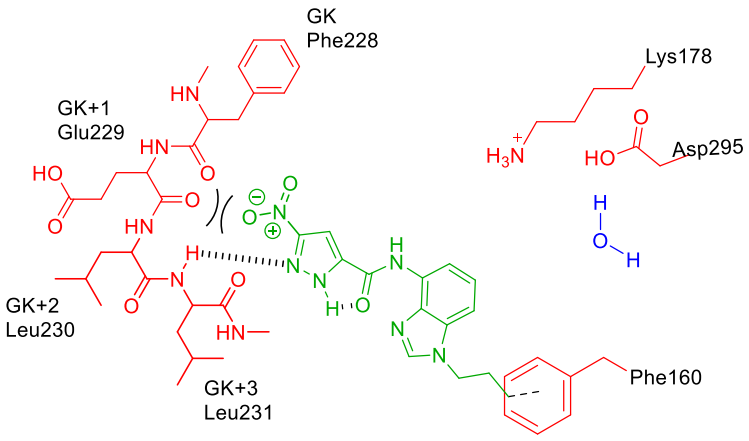


50 **51** **52** **53**

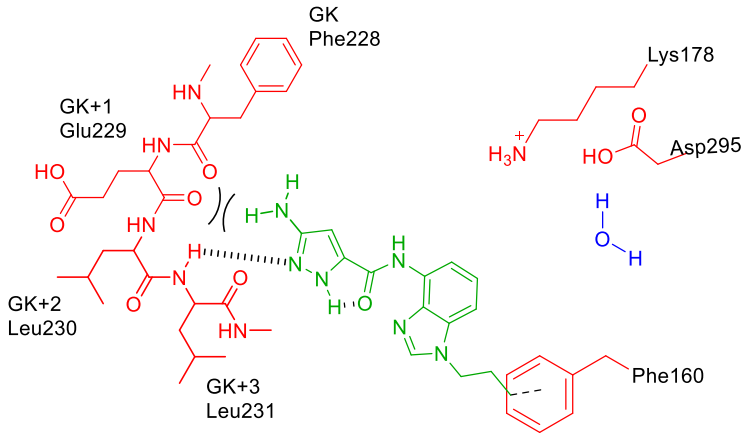
Compound	DYRK2 K_i (nM)
50	368
51	366
52	8% inhibition[a]
53	15% inhibition[a]

[a]*SP experiment @ 1 µM n=3*

Interestingly, the two **CI709** analogues, **52** and **53**, were inactive, **Table 5**. The drop in activity could be due to the loss of the bromine at the 4′ position, which we have shown to be essential for potency. In addition, the introduction of a substituent at the 3′ position may actually be too large and result in a steric clash with the hinge region and a decrease in activity, **Figure 67**. The inactivity of compound **39**, which has a methyl group in this position, would support this explanation.



(52)



(53)

*Figure 67: Model of potentially unfavourable poses of nitro **52** and amino **53**.*

More specifically, with the help of modelling, we recognised that there may be a clash between the C=O lone pair of Glu229 and the oxygen lone pair and/or oxygen lone pair of the nitro of **52**. Similarly, with the amino analogue **53**, the distance between an NH and the oxygen atom of Glu229 is approximately 1.53 Å (hydrogen bond interactions are typically >2.8 Å[129]), thus highlighting a potential steric clash, **Figure 67**. These potentially unfavourable interactions could result in a change in the binding pose and therefore reduction in activity.

With respect to the pyridine analogues **50** and **51**, they displayed similar activity to each other, **Table 5**. We propose that due to the smaller size of these motifs, they have the potential to bind in a number of conformations within the DYRK2 active site, where the pyrazole or the pyridine could interact with the hinge region. If the aminopyrazole does bind to the hinge as we originally proposed, this is probably because the smaller pyridyl tail group allows reorientation of the pyrazole head group to successfully engage its substituents with the HBDs and HBAs of GK+1 and GK+3.

However, because of their small size and number of HBD/HBA moieties, these pyridyl analogues are likely to be promiscuous kinase inhibitors, although they have not been tested against DYRK1A or other kinases to assess their selectivity profile.

This brief investigation highlighted that both HBD and HBA groups are not tolerated at the 3′ position when there is a bulkier tail group present, i.e. the substituted benzimidazole. Further investigation into determining the activity of the unsubstituted benzimidazole analogues of these would provide us with a greater understanding of the tolerance of substitution at this position. Although the loss of activity could also, once more, be due to the loss of the bromide at the 4' position.

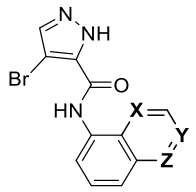
At this stage of our SAR investigation, we had shown that the 4′ bromopyrazole was the optimum motif for the hinge-binding head group, and that the most potent inhibitors contained a benzimidazole tail group i.e. **CI639** and **CI709**. In order to explore the tail group more extensively, we next examined the activity of different nitrogen containing heterocycles in this position.

3.3 Tail group investigation

3.3.1 Alternative heterocycles

Previous work investigated the effect of other 6,5-fused heterocycles such as indole and benzoxazole, at various positions around the phenyl ring (see **Appendix 10.2**). However, these analogues were either inactive or displayed reduced activity (>1.5 μM). Our next step was to examine the activity of a range of nitrogen containing 6,6-fused heterocycles, **Table 6**.

Table 6: Investigating alternative heterocycles.



Compound	X	Y	Z	LT (%)	In-house (K_i nM)
54	CH	CH	N	33	60% inhibition[a]
55	CH	N	CH	21	60% inhibition[a]
56	N	CH	N	-	600
57	N	CH	CH	-	1000

[a]*SP experiment @1 µM n=3*

Unfortunately, all of the investigated 6,6-fused heterocycles were less active than the benzimidazole system **CI639** (K_i 102 nM), **Table 6**. Quinoline **54** and isoquinoline **55** derivatives were both deemed inactive, as they displayed 60% inhibition at 1 µM, which was the threshold set for further dose response experiments, **Table 6**. The 8′ quinolinyl isomer **57** was moderately active (K_i 1 µM) and the quinoxaline analogue **56** was slightly more active (K_i 600 nM), **Table 6**. We hypothesised that the 6,6- fused system was too bulky to bind in a similar fashion to **CI639**, where the heterocycle points into the back of the active site. With the help of modelling, and using the most potent quinoxaline **56** of this series as the example, we modelled a possible conformation of **56** in the DYRK2 active site. For simplicity, only the key amino acid residues with the possibility of binding to this structure were considered, **Figure 68**.

.

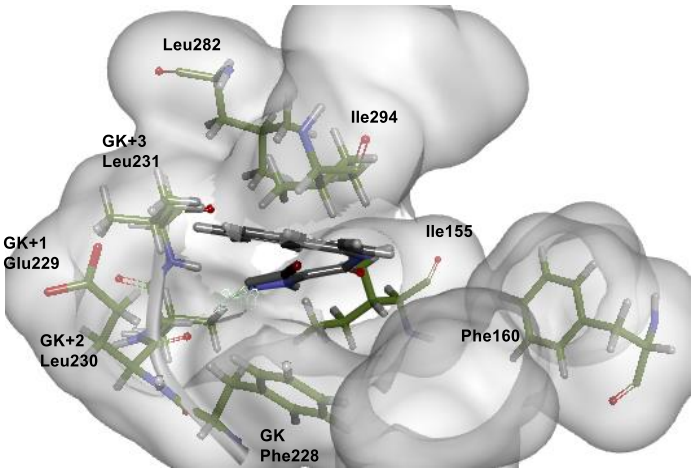


*Figure 68: Possible conformation of quinoxaline **56**.*

One of the possible conformations of the quinoxaline ring of **56** is illustrated in **Figure 68**. The ring points towards the back of the active site towards the hinge, where it can bury itself in the hydrophobic pocket between Leu282 and Ile294, and also has the potential to make pi interactions with the gatekeeper residue Phe228. However, as a result, the pyrazole head group distorts and increases the distance between itself and the hinge, potentially losing these important hydrogen bond interactions. Therefore, this conformation suggests a reason for the loss of activity. However, this rationale is but one possible explanation for the reduction in activity as there are many more possible conformations of this structure.

Furthermore, this does not explain the relationship between the position of the nitrogen atom on the heterocycle and the associated activity. The presence of a nitrogen atom at the 4 position of the ring increases the potency, **Figure 69**.

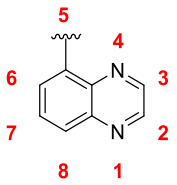


Figure 69: Numbering on the quinoxaline ring.

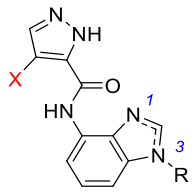
A possible reason for the difference in activity between the inactive 5-quinolinyl **54** and isoquinolyl **55** and the moderately active 8-quinolinyl **57** and quinoxalyl **56** derivatives, is the difference in size of the aromatic C-H vs the N-lone pair. The N-lone pair is smaller in size compared to the aromatic C-H bond and therefore, the C-H analogues distort the structure of the inhibitor which results in a worse fit for the active site.

It is clear that the DYRK2 active site is less accommodating to larger heterocycles compared to the smaller benzimidazole. Therefore, the next step was to examine the effects of different substituents on the benzimidazole.

3.3.2 *N*-1 Investigation

One of the synthetic routes we employed to deliver the substituted benzimidazoles resulted in *N*-1 substituted by-products. Therefore, we decided to assess their activity in order to determine the importance of the *N-3* substitution, **Table 7**.

Table 7: Comparison of N-1 and N-3 substitution.



Compound	X	R	LT (%)	In-house (K_i nM)
58	Cl	*1- n*-propyl	21	-
40	Cl	*3 - n*-propyl	91	335
59	Br	*1* - CH_2CH_2OMe	17	50% inhibition[a]
60	Br	*3* - CH_2CH_2OMe	75	890

[a]*SP experiment @ 1 µM n=3*

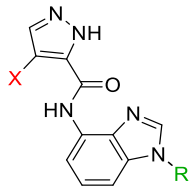
We previously discussed the activity of the chloro analogue **40** (K_i 335 nM) in **Section 3.2.2.1**. Its *N*-1 substituted counterpart **58** was tested externally by LT, and at 1 µM showed little activity (21%), **Table 7**. Next, we envisaged that the introduction of a HBA could increase the potency through interaction with solvent. However, the results indicate that the addition of the terminal methoxy group provided a less potent inhibitor **60** K_i 890 nM, **Table 7**, although as expected, its *N*-1 equivalent **59** was inactive (<60% inhibition at 1 µM), **Table 7**. We postulated that the aliphatic chain at the *N*-1 position cannot be tolerated due to the steric clash from the adjacent hydrophobic Ile155. This could result in a change in the conformation and loss of key interactions, and therefore a loss of activity. Thus, from this brief investigation it was clear that the substitution pattern, specifically *N*-3 substitution on the benzimidazole, was also essential for potency.

3.3.3 *N*-3 Investigation

Following verification of its importance to potency, we thoroughly investigated the extent of the *N-3* substitution on the benzimidazole. A range of different compounds with different substituents were synthesised and their DYRK2 activity was assessed by LT and also *via* the in-house assay.

3.3.3.1 Aliphatic analogues

Table 8: Investigation of aliphatic N-3 substitution.



Compound	X	R	LT (%)	In-house (K_i nM)
61	Br	ethyl	-	483
41 (**CI709**)	Br	*n*-propyl	62 nM[a]	210
62	Cl	*n*-butyl	66	325
63	Br	*n*-butyl	-	46% inhibition[b]
64	Br	*iso*-propyl	-	14000
65	Br	*iso*-butyl	96	64
66	Br	*sec*-butyl	-	433

[a]n=1 experiment, [b]SP experiment @ 1 µM n=3

Firstly, aliphatic analogues of extending and branched chain lengths were prepared and tested **61–66**. This revealed a trend with respect to chain lengths where longer chains (>*n*-propyl) were tolerated when a smaller halogen was present at the 4′ position on the pyrazole (**62** vs **63**, **Table 8**). More specifically, the bromo *n*-butyl analogue **63** was inactive (46% inhibition at 1 µM) compared to its chloro equivalent **62** (K_i 325 nM). This suggested a new binding pose where the chloride could potentially bury itself further into the back pocket of the active site, allowing the longer chain to interact in a similar fashion to the *n*-propyl analogue of **CI709**. An inhibitor with increased potency was also discovered, in the form of the *iso*-butyl analogue **65** (K_i 64 nM, **Table 8**). This isomer, with the longest chain being 3 carbons in length, was 3-fold more potent than the previous lead **41 CI709** (K_i 210 nM). In addition, the *iso*-propyl **64** (K_i 14 µM) and *sec*-butyl **66** (K_i 433 nM) analogues highlighted that

substitution at the C-1 position of the aliphatic chain was less favoured, when compared to their unbranched counterparts; ethyl **61** (K_i 483 nM) and propyl **41** (K_i 210 nM), **Table 8**. From these results, we postulated that the 3 carbon chain was the optimum length for binding, potentially through CH pi and hydrophobic interactions, with Phe160 of the DYRK2 active site, **Figure 70**.

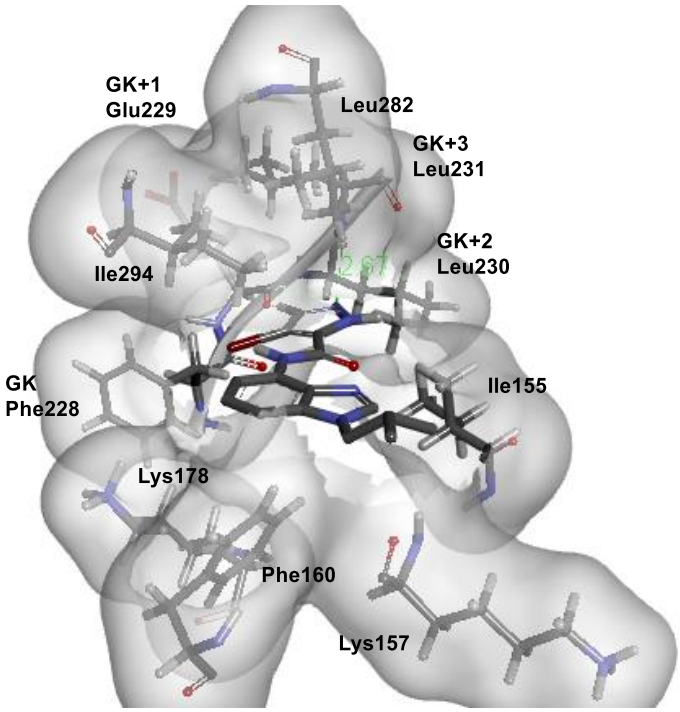
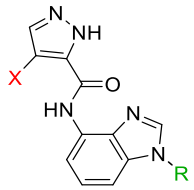


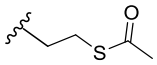
*Figure 70: Possible key interactions between iso-butyl **65** and DYRK2 active site.*

A model of highly potent *iso*-butyl analogue **65** (64 nM), highlighted the potential interactions which may contribute to its increase in activity, **Figure 70**. The addition of the methyl group (*n*-propyl **41** vs *iso*-butyl **65**) increases the overall lipophilicity of the molecule, which is an important property for activity (providing the possibility for hydrophobic interactions) and permeability. More specifically, the calculated logP (CLogP) for the *n*-propyl analogue **CI709** was 2.3, whereas the CLogP for the *iso*-butyl analogue **65** was 3.1. This model also provided us with evidence as to why the branched analogues *iso*-propyl **64** and *sec*-butyl **66** were less active than their linear counterparts (ethyl **61** and *n*-propyl **41**), as the C-1 CH's could possibly interact with Phe160 through C-H pi interactions or *via* a weak H bond interaction with the carbonyl group of Lys157, **Figure 70**. However, in the case of the *iso*-propyl and *sec*-butyl analogues one of these was replaced with a methyl group, potentially disrupting this favourable interaction, resulting in unfavourable steric clashes between the methyl groups and Phe160 or Lys157.

3.3.3.2 Introduction of electronegativity

Table 9: Electronegative N-3 substituted analogues.



Compound	X	R	LT IC_{50} (nM)	In-house (K_i nM)
67	Cl	CH_2CH_2Cl	440	173
68	Br	CH_2CH_2Cl	39	**19**
69	Br	$CH_2CH_2CF_3$	89.9	179
70	Br		1200	-

Next, we decided to introduce an electronegative atom at the 3-position of the carbon chain, in order to provide a handle for further diversification, and for a potential halogen bond interaction with Phe160 or Lys157. The chloro analogues (**67** and **68**) and the trifluoro analogue **69**, were synthesised and assessed *in vitro*. Additionally, we postulated that the trifluoromethyl analogue **69** could provide a weak interaction with the carbonyl carbon of Lys157, through a weak HBD interaction or a weak interaction with the pi system of Phe160. Pleasingly, these analogues were more potent than the *n*-propyl equivalent **CI709 41** (K_i 210 nM), where the 4′ chloro **67** and trifluoro **69** analogues were equipotent (K_i 173 and 179 nM, respectively), **Table 9**. The bis chloride species **67** reinforced the observation that the smaller halogen atom on the pyrazole ring results in a less potent inhibitor, as the 4′ bromo analogue **68** was 10 fold more potent than **CI709** (**41**) (**68** K_i 19 nM vs **41** K_i 210 nM, **Table 9**) and 9-fold more potent than its chloro equivalent **67** (K_i 173 vs 19 nM respectively). This resulted in **68** being the most potent inhibitor of this series. In order to explain these differences in activity, a model of **68** in the DYRK2 active site was considered, **Figure 71**.

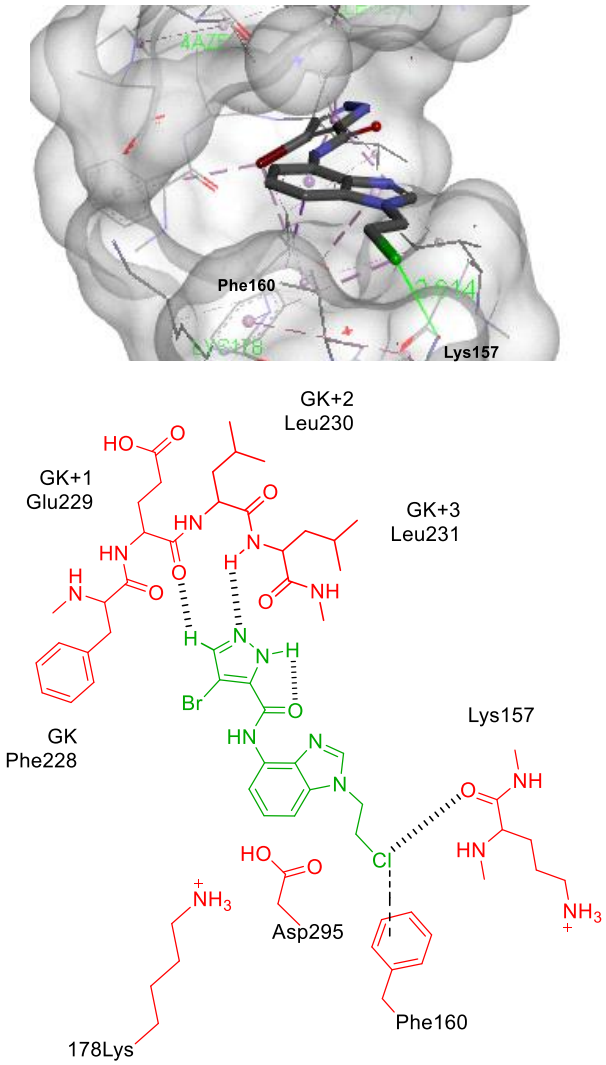


*Figure 71: Model of **68** in DYRK2 active site.*

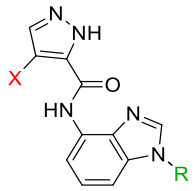
From these modelling studies, we propose that **68** could interact with lone pair of the C=O backbone of Lys157, *via* a halogen bond interaction and Phe160 *via* a van derWaals interactions, **Figure 71**. Interestingly, inhibitor **68** contains an electrophilic handle, which has the potential to react with a nucleophilic cysteine residue and become covalently bound to DYRK2. According to the DYRK2 crystal structure, there are no cysteine residues located in the active site. Lysine residues nearby can be classified as nucleophilic, but are likely to be protonated at physiological pH and therefore, unreactive. From a drug design perspective, the chloroethyl of **68** was used as an electrophilic handle for late stage diversification and from the extended reaction

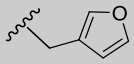
times with a range of nucleophiles (see **Section 6.2.3**), we can infer that **68** is less likely to be a covalent inhibitor.

The final aliphatic derivative, thioacetate **70** was moderately active (IC_{50} 1.2 µM), which reinforced our suggestion that a 3 atom chain length at the *N*-3 position was optimal for potency.

3.3.3.3 Aromatic *N*-3 substituents

Table 10: Investigation of aromatic N-3 substituents.



Compound	X	R	LT (%)	In-house (K_i nM)
71	Cl	Bn	53	-
72	Br	Bn	67	53% inhibition
73	Br	CH_2CH_2Ph	19	50% inhibition
74	Br	CH_2CH_2OBn	27	43% inhibition
75	Br		-	210

In order to fully understand the limit of substitution, longer and bulkier substituents, were examined. More specifically, aromaticity was introduced with the aim of picking up pi-stacking interactions with Phe160. *N*-benzyl analogues **71** and **72** soon revealed the limit to our substitution, where we can infer from the LT results for **71** that the smaller halogen would not result in a tolerance for the bigger phenyl group in the adjacent region of the active site, as seen previously with the *n*-butyl analogues **62** and **63**. The more flexible aromatic analogue **73** was synthesised in the hope that the longer chain could position the phenyl group at an optimum angle to interact with Phe160; however, this analogue was also inactive, **Table 10**, which suggested that this substituent was too bulky. Next, ethoxybenzyl derivative **74**, provided a HBA moiety which could potentially interact with the NH of Lys157. However, the results proved that this analogue was inactive, **Table 10**, thus, further highlighting the spatial restrictions. These results further demonstrated that there was indeed a limit to the size of the substituent at this position on the benzimidazole.

Next, in order to further probe the tolerance of aromaticity, a smaller aromatic 3-furyl analogue **75** was synthesised, which interestingly, was active (K_i 210 nM, **Table 10**). We propose this is due to its smaller size compared to the other aromatic analogues **71**–**74**, which eliminates steric clashes in the binding site. In addition, **75** can form hydrogen bond interactions with AA residues of the DYRK2 active site, **Figure 72**.

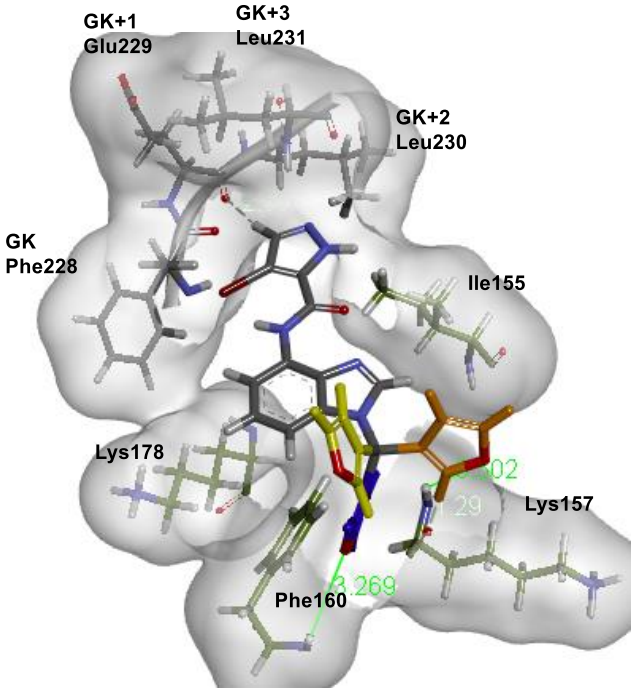
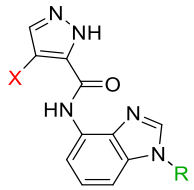


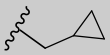
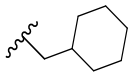
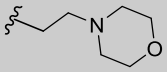
*Figure 72: Possible interactions between 3-furyl analogue **75** and the DYRK2 active site.*

With the aid of modelling studies, three potential positions of the furyl ring were identified and are highlighted in yellow, orange and blue, **Figure 72**. Firstly, with respect to the furyl group conformation highlighted in yellow, it can form edge-to-face aromatic interactions with the neighbouring side chain of Phe160. Alternatively, the furyl conformation shown in orange has the potential to form a hydrogen bond and a weak CH—O interaction with the backbone amide of Lys157 in addition to CH-pi interactions with the side chain of Ile155. Finally, the blue furyl conformation has the potential to form a hydrogen bond interaction with the backbone NH of Phe160, as well as face-to-face pi-stacking interactions between the two aromatic rings. These possible binding interactions, in addition to the essential hinge binding and hydrophobic interactions, resulted in a binding affinity comparable to **CI709** (**41**), (**75** and **CI709** (**41**) both have a K_i 210 nM).

3.3.3.4 Aliphatic cyclic *N*-3 substituents

Table 11: Cyclic N-3 substituents.



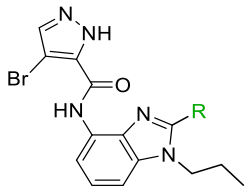
Compound	X	R	LT (%)	In-house (K_i nM)
76	Br		47	55
77	Br		-	2600
78	Br		-6	>14000

We now knew that some level of steric bulk was tolerated (3-furyl **75** 210 nM) at the benzimidazole *N*-3 position. Therefore, we decided to investigate the effect of saturated rings on the activity of this inhibitor series. Pleasingly, the cyclopropyl analogue **76** was 4-fold more potent than the *n*-propyl counterpart **CI709** (K_i 55 nM vs 210 nM respectively, **Table 11**). This activity was comparable to the *iso*-butyl analogue **65** (K_i 64 nM). Thus, cyclopropyl **76** has the potential to bind in a similar fashion to Phe160 *via* CH-pi interactions and possibly *via* a weak H-bond with the C=O of Lys157. Furthermore, this result confirmed that a carbon chain length of three atoms provided the most active inhibitors. Not surprisingly, the cyclohexyl analogue **77** was almost 50-fold less active than the cyclopropyl equivalent (**77** K_i 2.6 µM vs **76** 55 nM, **Table 11**), presumably because the larger and more flexibile hydrophobic cyclohexyl group clashed with the side chains of Phe160 and Lys157. Finally, in order to introduce potential hydrogen bond interactions with this region of the active site, the morpholine analogue **78** was prepared but was completely inactive. This was most likely due to the size of the morpholine ring, as previously exemplified with compounds **73** (CH_2CH_2Ph) and **74** (CH_2CH_2OBn). Taken together, the SAR for the *N*-3 substituents suggested that this group was accommodated at the rear of the active site where there were size limitations. Otherwise, we would have expected compound **78** to be more active, because the water soluble morpholine substituent could have reoriented to project into the solvent.

3.3.4 C-2 Investigation

Another position on the benzimidazole that hadn't been explored was the C-2 position. Through investigation of a new synthetic route to this inhibitor series, the opportunity of introducing diversity at the C-2 position was revealed. Using modelling, we determined that the area of the active site beneath the imidazole region of the benzimidazole group was rich in hydrophobic residues eg. Ile155 and Val163 if the ring projected away from the active site. Conversely, if projected in towards the back of the site, more polar residues were evident, such as Lys173, Glu193 and Asp295. Introducing an alkyl group in this position was synthetically achievable to investigate the effect on potency, **Table 12**.

Table 12: C-2 investigation with alkyl substituents.



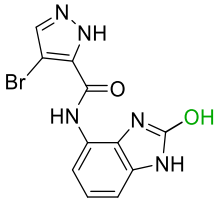
Compound	R	LT (%)	In-house (% inhibition)[a]
79	methyl	27	-
80	ethyl	13	32

[a]*SP experiment @ 1 µM n=3*

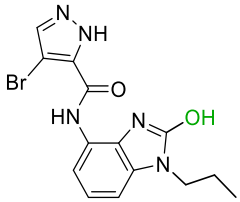
Both methyl **79** and ethyl **80** analogues were inactive, **Table 12**, which again suggested that the benzimidazole projects towards the hydrophilic rear of the site where hydrophobic aliphatic groups cannot be accommodated in this position.

Next, we introduced a polar group in this position by preparing the oxygen analogues **81** and **82**, **Table 13**.

Table 13: C2 investigation with potential HBD.



81

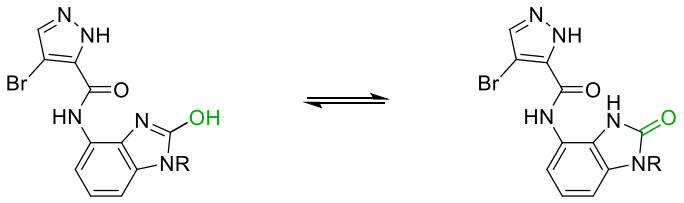


82

Compound	LT (%)	In-house (K_i nM)
81	10	50% inhibition[a]
82	16	>14000

[a]*SP experiment @ 1 µM n=3*

Unfortunately, these analogues were both inactive, **Table 13**. This could be explained by the fact that the benzimidazol-2-ol's **81** and **82** had the potential to tautomerise to their benzimidazol-2-one form **83**, **Scheme 2**, thus converting a HBD to a HBA in this position, and also increasing the size of the heterocyclic group adjacent to the amide.



81 R=H
82 R=*n*-Pr

83

*Scheme 2: Tautomerisation to one form **83**.*

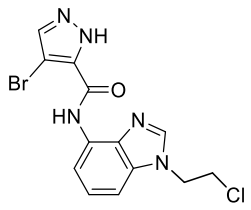
From these results it was clear that the presence of either tautomer is detrimental to activity, **Table 13**. In line with our observations made for the 6,6- fused heterocycles (**Table 6**; **Section 3.3.1**) we suggest this reduction in activity was the result of disfavoured conformations of the 2-one form **83**, **Scheme 2**.

Therefore, from these findings we have shown that substitution at the C-2 position was not tolerated and consequently resulted in a complete loss of activity. It also confirmed that there was very little space at the back of the active site that accommodates the benzimidazole moiety, with only small substituents tolerated at the *N*-3 position of this group.

3.3.5 Brief summary

Through the investigation of functional group tolerance at the *N*-3 position of the benzimidazole, an inhibitor with 10-fold increased activity (**68**, K_i 19 nM) was discovered, in comparison to the previous lead **CI709** (**41**) (41 K_i 210 nM). *Iso*-butyl **65** and cyclopropyl **76** analogues also revealed a further increase in activity (**65** K_i 64 nM and **76** K_i 55 nM) compared with **CI709**. All three of these inhibitors demonstrated the importance for potency of the 3 atom chain length at the *N*-3 position. Whilst improving potency against DYRK2 was one aim of the project, as described in **Section 1.5**, maintaining the impressive selectivity of this series over DYRK1A was also a primary goal.

Pleasingly, upon *in vitro* evaluation with the closely related family member, DYRK1A, **68** was shown to be >500-fold selective for DYRK2, **Figure 73**.



68
DYRK2 K_i 19 nM
DYRK1A IC_{50} >10 000 nM

*Figure 73: Potent and selective inhibitor **68**.*

To account for selectivity, we superimposed the crystal structures of DYRK2 and DYRK1A, and examined the ATP-binding site of each isoform. **Figure 74** shows DYRK2 (green) and DYRK1A (red) around this binding site.

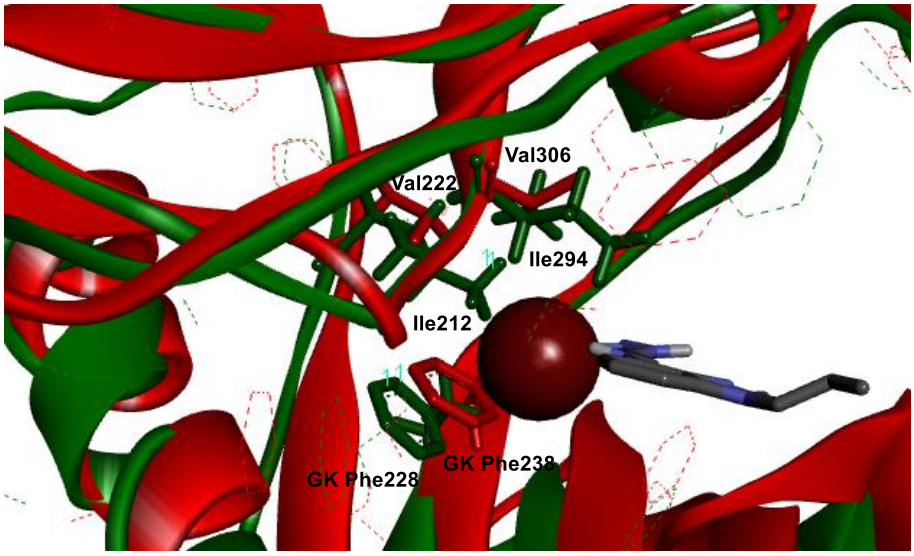


*Figure 74: Superimposed structures of DYRK2 (green) and DYRK1A (red) with **CI709** with Br in CPK format.*

The residues which are highlighted in green (DYRK2) surrounding the CPK formatted Br of **CI709** are the GK Phe228, Ile212 and Ile294. The equivalent residues in red (DYRK1A) are GK Phe238, Val222 and Val306 respectively, **Figure 74**. These smaller residues (Ile vs Val) represent the only significant differences between the two isoforms in the ATP binding site region. The third and final difference is the hinge region residues (DYRK2 Leu230 vs DYRK1A Met240). Notably, Ile212 in DYRK2 is a key residue in the small hydrophobic pocket which accommodates the Br atom throughout the series, and is essential for activity. A possible reason why the series is active in DYRK2 but not DYRK1A is the optimal fit between the larger Ile212 side chain and the Br of the pyrazole head group, which is absent with the smaller Val222 side chain in DYRK1A.

Although this was the most potent inhibitor, it requires further investigation. The compound contains an electrophilic handle, therefore, this has the potential to react with a nucleophilic cysteine or lysine residue of DYRK2 and become irreversibly bound to the protein. This in itself can be advantageous, as shown by the development of osimertinib **84**, afatinib **85** (targeting EGFR mutants in NSCLC) and ibrutinib **86** (targeting Bruton's Tyrosine Kinase (BTK) in B-cell malignancies)[130], **Figure 75**.

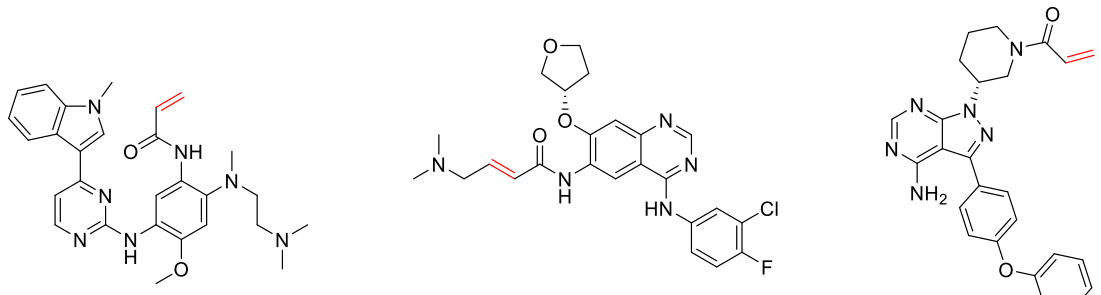


*Figure 75: Known drug molecules with electrophilic sites- osmertinib **84**, afatinib **85** and ibrutinib **86**.*

These kinase inhibitors contain electrophilic functionalities that react with cysteine residues in the active sites of the target kinases through Michael addition. Irreversible binding with the kinase produces a sustained block of the enzyme, allowing lower doses and higher specificity for the target kinase with an appropriately positioned cysteine. Whilst DYRK does not possess such a cysteine itself, an electrophilic chloroethyl group is a recognised toxicophore, which means **68** would need to be assessed for off-target toxicity if this functional group was to be retained.

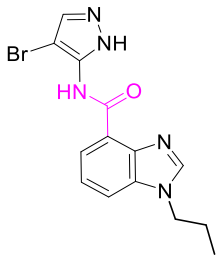
So far, we had established the importance of 1) the pyrazole heterocycle 2) the bromide at the 4′ position of the pyrazole and 3) the substitution pattern on the benzimidazole ring. All of these we believed to be pivotal for potency. The final position left to examine was the amide linker.

3.4 Amide linker

In order to investigate the importance of the amide linker, two strategies were devised. First, the aim was to synthesise the reverse amide analogue of **CI709** (**41**) and determine its activity. Second, we aimed to introduce less flexibility and more lipophilicity within the molecule, and examine the effect that this had on the activity of the inhibitor.

3.4.1 Reverse amide

The reverse amide analogue of **CI709** (**41**) was synthesised and assessed *in vitro*, **Figure 76**.



87
22% inhibition @ 1 μM

*Figure 76: Reverse amide analogue **87**.*

Surprisingly, **87** was inactive (22% inhibition at 1 μM), **Figure 76** and further dose response experiments were not carried out. In order to suggest a reason why **87** was inactive, we had to consider some of the possible conformations it could adopt, **Figure 77**.

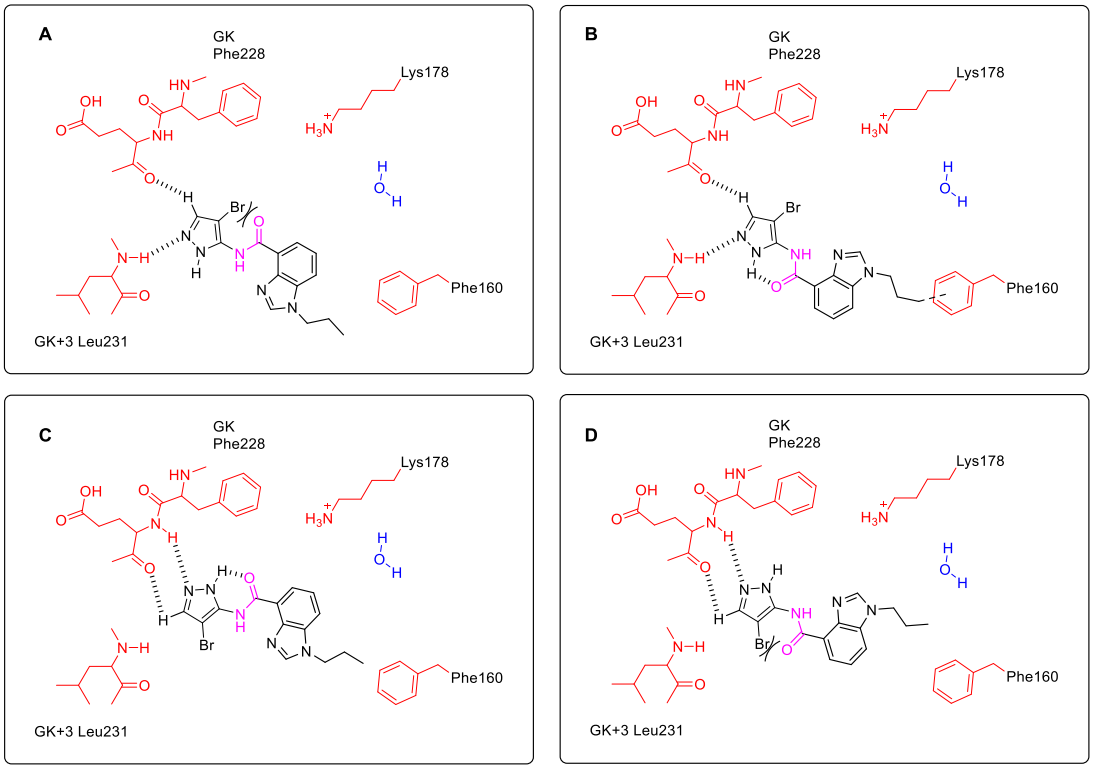


*Figure 77: Possible conformations of reverse amide **87**.*

With the aid of a simplistic model of the important interactions within the DYRK2 active site, we proposed the following explanations as to why **87** was inactive, **Figure 77**. First, if we proposed that **87** adopts a similar geometry to its counterpart **CI709**, **Figure 78**.

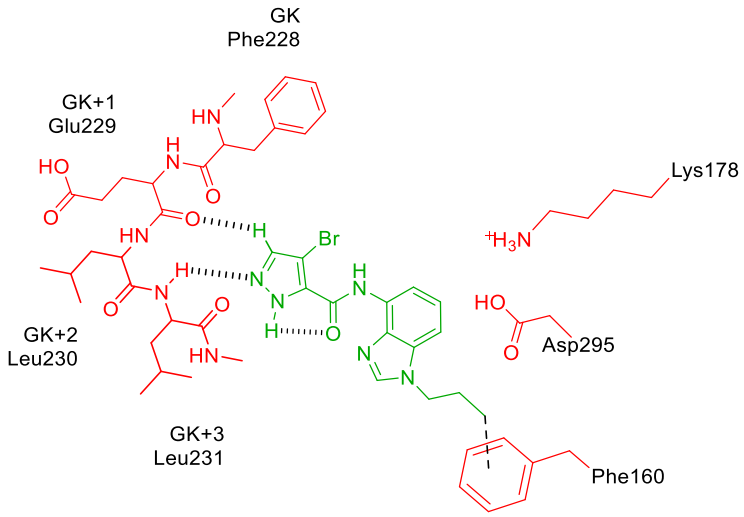


*Figure 78: Reminder of **CI709** model.*

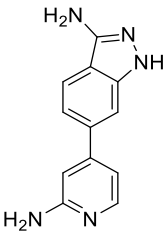
With respect to **A**, **Figure 77**, we encounter an unfavourable conformation where the C-Br bond and oxygen lone pairs are eclipsed, which is not a low energy conformation and is therefore unlikely. The resultant loss of planarity of the scaffold would have a negative impact on binding, leading to inactivity. Next, **B Figure 77**, rotation of the amide bond allowed for the potential intramolecular H bond interaction between the amide oxygen and the pyrazole NH in a 6-membered ring fashion. However, this would bring the amide carbonyl group into closer proximity with the backbone carbonyl of GK+3 Leu231 and the hydrophobic side chain of Ile155 of the P-loop above, leading to repulsive interactions and a loss of activity. Third, **C Figure 77**, highlights the possible conformation if we were to flip **B**, which results in the removal of the bromide from the hydrophobic pocket and replaces it with the polar N-H of the pyrazole. Moreover, the key interactions between **87** and the hinge are potentially elongated and therefore weakened, possibly resulting in reduced activity. Finally, **D Figure 77**, rotation of the amide bond resulted in the unfavourable C-Br, amide oxygen lone-pair interaction and a loss of co-planarity as seen in **A**.

This short sub-series provided an encouraging discovery, as it appeared that the original amide was also a critical part of the scaffold in terms of potency and possibly, selectivity. At present, we are awaiting the DYRK1A results for the reverse amide analogue **87**.

Next, we investigated the replacement of the amide all together in order to explore our hypotheses that the original amide was essential for potency and selectivity.

3.4.2 Hybrid series

Fortunately, we had the opportunity to screen an in-house library of kinase-like structures, where we identified a potent inhibitor **88** with similar functionality to our own series, **Figure 79**.



88
DYRK2 K_i 2 nM
DYRK1A 97% inhibition @ 1 µM

*Figure 79: Kinase-like in house library hit **88**.*

Indazole **88** was 10-fold more potent than our current lead compound **68** (K_i 19 nM) of the pyrazole series but was not selective, showing significant activity against isoform DYRK1A at 1 µM (97% inhibition). To explore whether this improvement in potency could be translated into our own series, but without the loss of selectivity, we decided to incorporate specific features into our inhibitor design. The overlay of both compounds identified similarities and areas for further exploration, **Figure 80**.

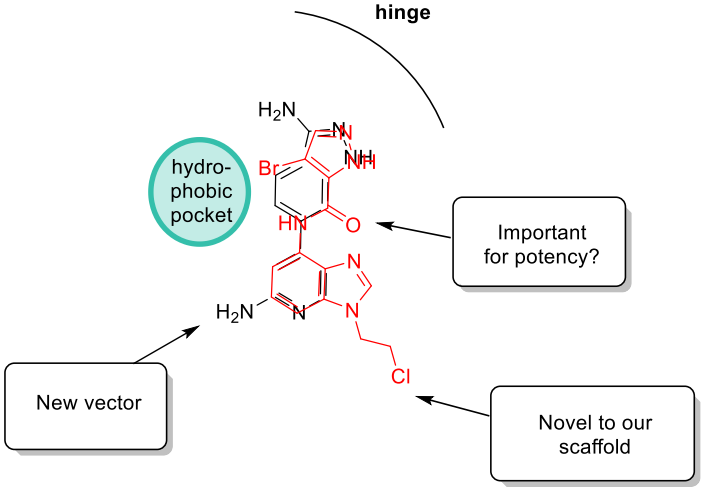
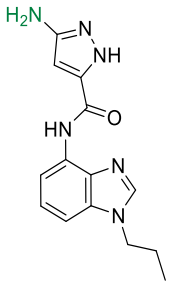


*Figure 80: Overlay of our lead **68** and screening hit **88**.*

First, the amino indazole had all the characteristics of being the hinge binding motif, although within our own series, we identified in **Section 3.2.3** that the 3′ amino pyrazole analogue **53** was inactive, **Figure 81**. This suggested that with a smaller tail

group (aminopyridine vs benzimidazole) and with the additional hydrophobicity (phenyl ring vs amide), this functionality could be tolerated.



53
15% inhibition @ 1 μM

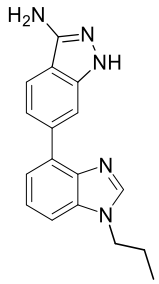
*Figure 81: The inactive amino analogue **53**.*

The overlay also suggested that the phenyl ring of indazole **88** could potentially occupy the hydrophobic pocket and act as a mimic of the bromide of our inhibitor **68**. Finally, the benzimidazole was novel to our scaffold and could play a key role in selectivity. On the other hand, the 2-amino pyridine motif of **88** also had the potential of being a hinge binding motif, which would result in a change in the binding mode compared to the original series. Therefore, we set out to synthesise a number of hybrid derivatives which would allow us to further understand the observed activity and selectivity.

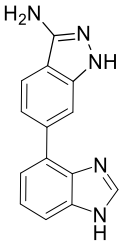
3.4.2.1 Amino indazole analogues

First of all, we synthesised amino indazole derivatives (**89** and **90**) and tested them *in vitro* to determine the importance of the amide and bromide of our series, **Table 14**.

Table 14: Results from hybrid series 1.



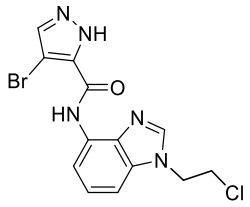
89



90

Compound	DYRK2 K_i (nM)
89	1300
90	758

These analogues were significantly less active than the original amide equivalents, as the *n*-propyl analogue **89** was >6-fold less active than the amide equivalent **CI709** (**41**) (K_i 1300 vs 210 nM respectively). Additionally, the unsubstituted benzimidazole analogue **90** was >7-fold less active than its amide counter-part **CI639** (K_i 758 vs 102 nM respectively). Although these analogues do not represent our most potent inhibitor, **68** (K_i 19 nM) **Figure 82**, they were investigated due to their ease of synthesis and their ability to answer the questions we had posed. Moreover, these results reinforced the importance of the amide and bromide functionalities for potency within the original series.

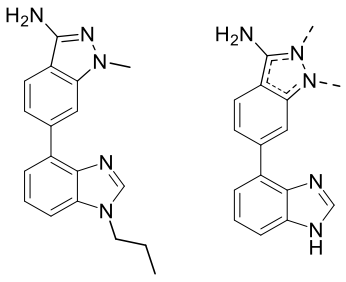


68
DYRK2 K_i 19 nM
DYRK1A > 10 000 nM

*Figure 82: Lead inhibitor **68**.*

To investigate the importance of the indazole N-H for binding, we assessed the activity of methyl substituted analogues **91** and **92**, **Table 15**.

Table 15: Results for substituted indazole analogues.



91 **92**

Compound	DYRK2 K_i (nM)
91	1700
92	47% inhibition[a]

[a]*SP experiment @ 1 µM n=3*

Methyl derivative **91** was not significantly less active than the unsubstituted equivalent **89** (1300 nM), which suggested that the hinge binding motifs were: the unsubstituted

indazole nitrogen atom (HBA) and the 3' amino group (HBD). This could also be explained by the even less active unsubstituted benzimidazole analogue **92** (47% inhibition at 1 µM). This methyl substituted analogue was present as an inseparable mixture (approx. 3:1 N1:N2) and its inactivity could potentially be explained through the loss of a hinge binding nitrogen. Through 2D NOESY NMR experiments (see **Appendix 10.5**), we identified the *n*-propyl analogue **91** as the correct isomer (1′ methyl, **Table 15**). Therefore, we inferred that the major isomer of the unsubstitued benzimidazole analogue was also the 1′ methyl isomer. As a single isomer, we would expect **92** to be somewhat less active or equipotent to its unsubstitued indazole analogue **90** (K_i 758 nM), however, it showed 47% inhibition at 1 µM and, therefore, was not investigated in subsequent dose response experiments. Thus, we believed that this emphasized which indazole nitrogen atom was important for hinge binding, **Figure 83**.

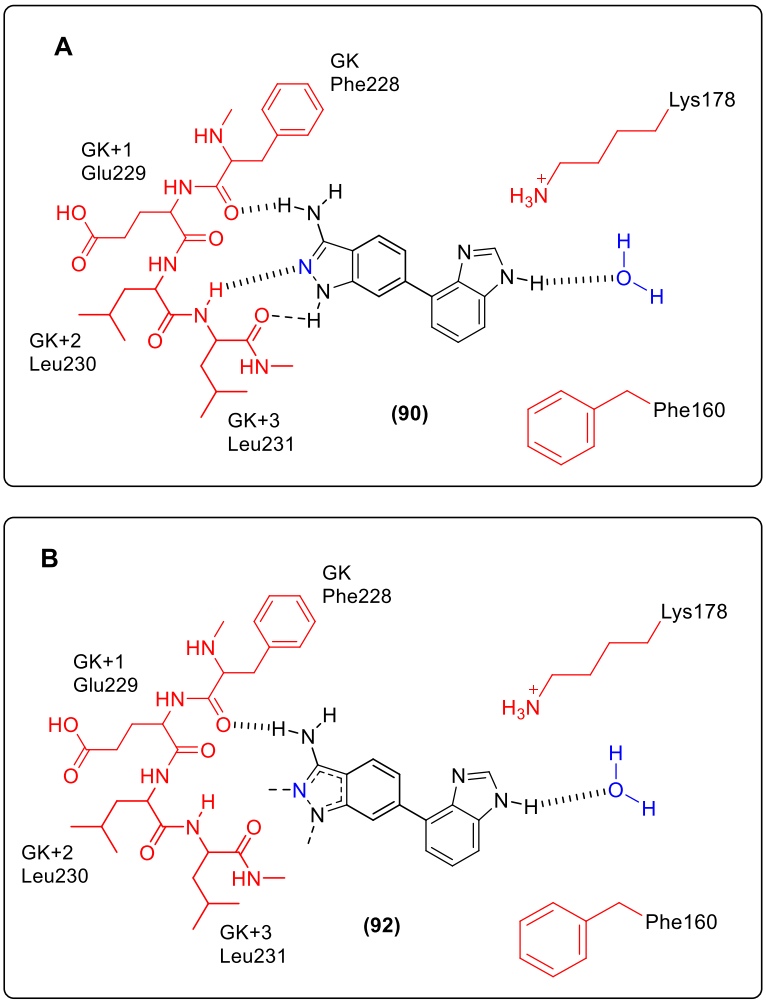
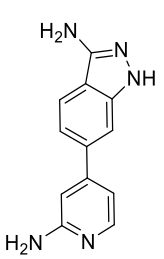


*Figure 83: Potential hinge binding interactions between **A)** unsubstituted **90** and **B)** substituted mixture of isomers **92**.*

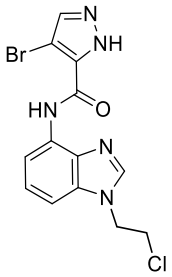
With respect to **A 90**, a nitrogen of the indazole (highlighted in blue) can form a hydrogen bond interaction with Leu231 and the amine can form a hydrogen bond interaction with the gatekeeper Phe228, **Figure 83**. The N-H of the indazole can also form a hydrogen bond interaction with Leu231, however, we believed this was less important due to the similar activity we observed for the methyl substituted and unsubstituted indazoles of the *n*-propyl analogues (**89** K_i 1700 nM and **91** K_i 1300 nM). Next, **B Figure 83**, when we introduce the methyl substituted mixture **92**, we disrupt the hydrogen bond interaction(s) with Leu231 and therefore lose activity. Therefore, we can infer that the corresponding single isomer would have had the key nitrogen (highlighted in blue) available for binding and so may have resulted in an inhibitor of similar potency to its unsubstituted indazole equivalent.

In order to determine if the amide and bromide of our original series were also essential for selectivity, indazoles **89** and **90** should be assessed for their DYRK1A activity.

From this brief investigation into the hybrid series, we have further developed our understanding that the amide and bromide of the original series were essential for potency. Based upon the less active inhibitors **89** (K_i 1300 nM) and **90** (K_i 758 nM), where the aminopyridine has been replaced by a benzimidazole, we can infer that the aminopyridine functionality of the new hit, **88**, is vital for its potency (K_i 2 nM), **Figure 84**.



88
DYRK2 K_i 2 nM
DYRK1A 97% inhibition @ 1 µM



68
DYRK2 K_i 19 nM
DYRK1A IC_{50} > 10 000 nM

*Figure 84: Screening hit **88** and lead **68** potencies.*

3.4.2.2 Substituted pyridine analogues

In order to investigate the effect of introducing the 2-aminopyridine motif into the original series, we set out to synthesise the following analogues, **Figure 85**.

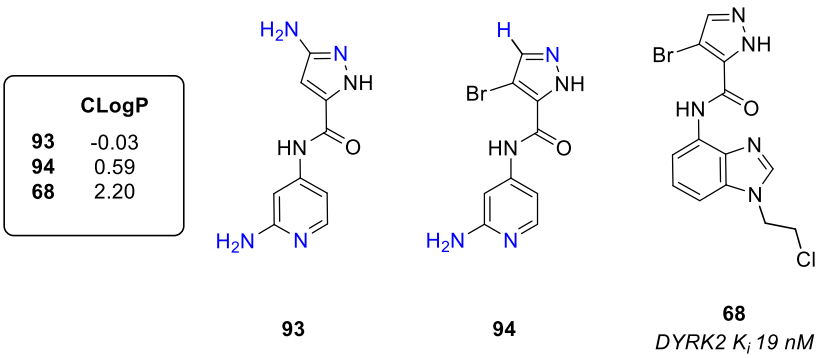


*Figure 85: 2-Amino pyridine analogues **93** and **94** and associated CLogP values.*

These analogues have the ability to bind to the hinge either *via* the head group pyrazole or tail group amino pyridine (highlighted in blue, **Figure 85**). However, with all the evidence that the bromide fits perfectly into the small hydrophobic pocket of the active site, we suspected that the pyrazole of **94** would remain as the hinge binder. Conversely, the hybrid with no bromide **93** could bind in both ways to the hinge. Unfortunately, this feature could be detrimental to selectivity as kinase active sites are well conserved and it was highly likely that **93** would also bind to a number of off-target kinases.

In addition, as we increase the ratio of the number of heteroatoms to carbon atoms, we could be penalised with bioavailability/permeability issues. The calculated logP (CLogP) provided us with a rough indication of the lipophilicity and therefore cell permeability of an inhibitor molecule. With regards to aminopyridines **93** and **94**, the addition of more polarisable groups decreased the CLogP in comparison to our lead compound **68** (CLogP 2.2, **Figure 85**), resulting in a CLogP <1 **Figure 85**, which would likely decrease the permeability. Thus, if these analogues were extremely potent, then their permeability would have to be assessed. However, the 2-aminopyridine analogues have not yet been synthesised.

3.5 SAR summary

From these studies we have discovered an inhibitor with a >10-fold increase in potency for DYRK2 (**68**, K_i 19 nM), which retained selectivity over DYRK1A (IC_{50} >10 µM). Moreover, we have established that there were limits to the substitution

pattern of the inhibitor series, as well as identifying positions which require further investigation, **Figure 86.**

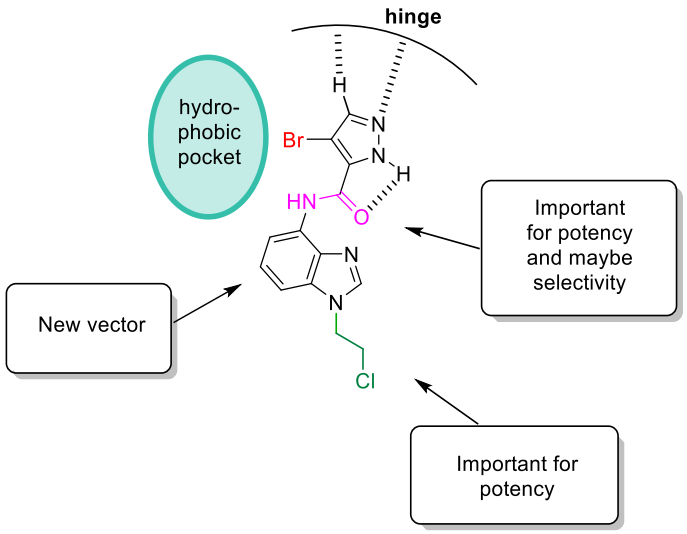


Figure 86: Brief SAR summary.

First, we have provided evidence that confirms our hypothesis that a bromide in the 4-position of the pyrazole was the optimal size to fit in a small hydrophobic pocket in the active site behind the GK residue. With respect to the hinge binding head group, further investigation is required to confirm whether the pyrazole is the best heterocycle at this position. The introduction of a 1,2,3-triazole or isoxazole/isothiazole should be investigated to interrogate this. Next, we extensively examined the effect of benzimidazole substitution on the activity of this inhibitor series, where we identified that inhibitors with a 3 atom chain substitution pattern on the benzimidazole provided the most potent inhibitors. A brief interrogation of the amide linker provided us with evidence that the amide moiety was essential for potency and possibly selectivity, which is yet to be confirmed. Finally, a screening hit provided us with a novel vector to explore on the aryl ring of the benzimidazole, **Figure 86**.

In addition to potency, there were a number of other *in vitro* factors that are highly important and require determination at an early stage in drug discovery projects. As highlighted in the aims and objectives (**Section 1.5**), successful inhibitors (K_i <60 nM) would be assessed for their drug metabolism and pharmacokinetic properties.

4 Physicochemical and pharmacokinetic properties

An important step in the drug discovery process is the physicochemical and pharmacokinetic (PK) assessment of compounds *in vitro.* These parameters are evaluated in order to provide an indication of the *in vivo* capabilities of drug candidates. Physicochemical properties include solubility, lipophilicity (logP/logD), number of HBD/HBA motifs and polarity (topological polar surface area or TPSA). These properties influence the PK parameters of a drug molecule, where PK can be described simply as the effect that the body has on a particular drug.[131] There are a number of important questions that must be addressed in order to progress with a drug candidate, some of which fall under ADME criteria, where ADME is the absorption, distribution, metabolism and excretion processes which drugs are susceptible to within the body. **Figure 87**.

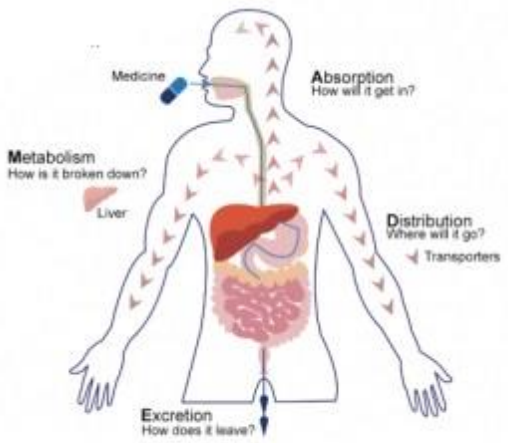


Figure 87: Schematic representation of the ADME process.[132]

Absorption describes the process by which a drug enters the bloodstream (assuming systemic distribution). For example, if administered orally, it must be able to penetrate the intestinal membrane in order to reach general circulation and from there its target. In addition to physiological considerations such as membrane surface area and blood flow, the absorption of a drug molecule is intrinsically associated with its physiochemical properties. Once it reaches general circulation it will be distributed and transported to the target site. Like absorption, the extent of a drug's distribution between the blood, different tissues and organs is largely governed by its physiochemical properties. If administered orally, before the drug reaches general circulation it must pass through the liver, where potentially it can interact with a

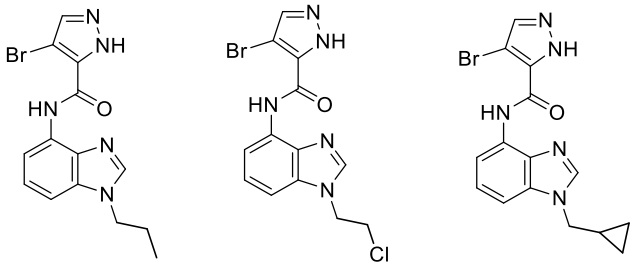
number of metabolising enzymes and be susceptible to biotransformation. These enzymes can be separated into two classes: phase I and phase II. Phase I enzymes introduce water-soluble functionality for recognition by phase II enzymes. For example, phase I cytochrome P450 enzymes (CYP450s) produce water-soluble metabolites *via* oxidation of the parent drug. Phase II enzymes can further modify metabolites by catalysing the conjugation of the metabolites to polar endogenous molecules such as sugars, amino acids and peptides to facilitate their excretion from the body. The primary route of excretion is *via* the kidneys, which are able to clear water soluble metabolites more effectively than the more lipophilic parent drug molecules.[133]

Before assessing these ADME parameters *in vitro*, because they often correlate with the physicochemical properties of the molecule in question, the latter can be determined based on its chemical structure. These properties can be determined experimentally, or be estimated from *in silico* calculations. We can then use these values along with potency, in some cases, to determine whether further *in vitro* ADME studies are necessary.

4.1 Physicochemical properties

In the drug design process there are criteria which medicinal chemists naturally follow to increase the chances of creating not only a potent drug but an orally bioavailable one. A qualitative method for recognising the chances of oral bioavailability, without the need for biological evaluation, is Lipinski's rule of five (Ro5). These criteria include MW <500, LogP <5, HBD's <5 and HBA's <10.[114] LogP is a measurement of hydrophilicity and is the partition coefficient between an organic (octanol) and aqueous phase. According to the Ro5, if one or more of these criteria aren't adhered to, this could lead to poor oral bioavailability.[134] If we examine the two most potent DYRK2 inhibitors from this project, **68** and **76**, along with the previous lead **CI709** (**41**) we can obtain some physicochemical information, **Table 16**.

*Table 16: Physicochemical information for **CI709**, **68** and **76**.*



CI709 (41)
K_i 210 nM

68
K_i 19 nM

76
K_i 55 nM

	CI709 (**41**)	**68**	**76**
MW (g)	348.2	368.6	360.2
TPSA (Å^2)	69.1	69.1	69.1
CLogP	2.7	2.2	2.6
HBD	2	2	2
HBA	3	3	3

All three inhibitors adhere to the MW (<500), CLogP (<5) rules and HB capabilities. Moreover, they have similar and moderate CLogP values (2.2–2.7). The third parameter which is highlighted in the previous table is TPSA. This is a measurement of polarity and is the sum of the surface area occupied by the polar atoms (mainly nitrogen and oxygen) of a molecule, and does not take into account other electronegative atoms.[135] All three inhibitors exhibit the same TPSA value (69.1 Å^2) and are below the acceptable threshold of 140 Å^2, which has been reported to correlate to good oral bioavailability *in vivo*.[136] On the other hand, a comprehensive study by Hughes and co-workers at Pfizer identified that a lower limit of TPSA (<75 Å^2) was associated with adverse toxicological effects *in vivo*.[137] Therefore, as our inhibitors fall under this threshold (69.1 Å^2), progression with this scaffold would require further investigation of potential toxicity at an early stage.

Having shown some of the basic physicochemical criteria were acceptable, we evaluated these compounds further in an *in vitro* setting. More specifically, their aqueous solubility and permeability, as these properties influence oral bioavailability.[138] The most favoured route of drug delivery is oral administration. A review by Zhong *et al*. reported that approximately 62% of FDA approved drugs are

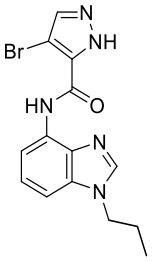
prescribed as oral doses. This is simply due to the ease of administration and patient compliance.[139]

4.2 Solubility and permeability

Finding the sweet spot between designing an orally bioavailable drug molecule which is polar enough to be water-soluble to enable rapid dissolution in the gut fluid but is also hydrophobic enough to be absorbed across the intestinal lipid-based membrane in order to reach the site of action is a perpetual challenge for medicinal chemists.

With respect to aqueous solubility, a quick method for measuring this is the turbidimetric solubility assay. This is a kinetic approach to determining the concentration at which a compound will precipitate in aqueous media. A stock solution in dimethyl sulfoxide (DMSO), of known concentration, is slowly added to aqueous buffer and precipitation is measured *via* UV detection, where the undissolved particles are detected by light scattering.[140] For the purposes of this project, this work was outsourced to Cyprotex Discovery Ltd where they determined the concentration of precipitation of inhibitor **CI709** (**41**), **Table 17**.

*Table 17: Solubility and permeability data for **CI709**.*



CI709 (**41**)
K_i *210 nM*

	Compound 41 (CI709)
Turbidimetric Solubility (µM)	>100
P_{app} A–B (x10^{-6} cms^{-1})[a]	22.7
P_{app} B–A (x10^{-6} cms^{-1})[a]	27.3
Efflux ratio	1.21

[a] *n=2*

Within the literature there are reports of benchmark concentrations which correspond to good aqueous solubility. For example >60—65 µg/mL is a reasonable goal for good solubility.[141,140] Pleasingly, inhibitor **CI709** (**41**) was highly soluble at high concentrations >100 µM in aqueous buffer, **Table 17**. Moreover, this concentration

corresponds to approximately 3500 µg/mL, which greatly exceeds the benchmark, and therefore we can presume that **CI709** (**41**) is highly soluble.

Cell permeability can also be measured *in vitro* and the most common method for this is the Caco-2 permeability assay. This is derived from human colon carcinoma cells and is used to determine the intestinal absorption of drug candidates. This is quantified by the apparent permeability co-efficient (P_{app}), which is the transport rate of the drug at a particular concentration. This is measured in both directions (simply A–B and B–A), **Figure 88**.

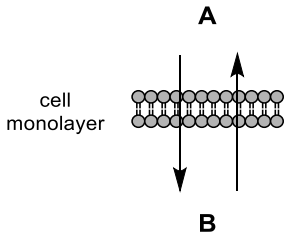


Figure 88: Simple representation of P_{app}.

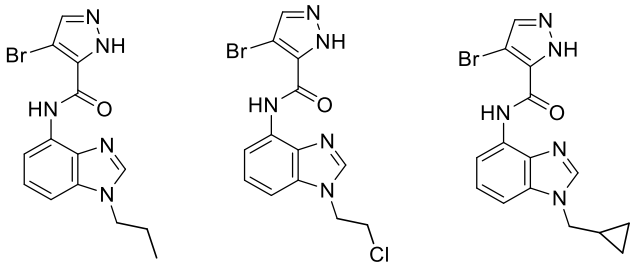
In addition, the efflux ratio can also be calculated which is used to indicate whether a compound undergoes active efflux,[142] where efflux is performed by active transporters such as P-glycoprotein (P-gp) and results in the transport of a substance out of the cell. Cells lining the intestinal tract actively efflux many small molecules that have passively diffused in as a protective mechanism against potential toxins. If a drug molecule under development has good lipophilicity, but is extensively effluxed, its net absorption will be poor and bioavailability low. Both permeability and efflux ratio are therefore important parameters to determine. This work was also outsourced to Cyprotex Discovery Ltd for inhibitor **CI709** (**41**). In addition to testing **CI709** (**41**), reference compounds are also measured, e.g. highly permeable (propranolol) and poorly permeable (talinolol) known drugs to be used as benchmarks when interpreting results. The highly permeable propranolol has P_{app} values of A–B 22.6 $x10^{-6}$ cms^{-1} and B–A 38.3 $x10^{-6}$ cms^{-1} with low efflux, whereas poorly permeable talinolol has P_{app} values of A–B 0.2 $x10^{-6}$ cms^{-1} and B–A 9.7 $x10^{-6}$ cms^{-1}, indicating poor penetration and high efflux. These values represent the amount of substance passing through the intestinal cell (A–B) and are transported back out (B–A). Therefore, from the results highlighted in **Table 17** we can infer that **CI709** (**41**) is highly permeable (P_{app} A–B 22.7 $x10^{-6}$ cms^{-1} and B–A 27.3 $x10^{-6}$ cms^{-1}), with a low efflux ratio (1.21). Cyprotex states that drugs are susceptible to active efflux when the efflux ratio value is >2. Therefore, from these results, we can assume that our inhibitor is not a substrate for P-gp.

We are awaiting the solubility and permeability data for our most potent inhibitor **68**. However, we can infer that as **68** is structurally similar to the previous lead **CI709** (**41**), which is highly soluble and permeable *in vitro*, that **68** will exhibit similar properties.

4.3 Metabolic stability

Commonly, one of the first early stage *in vitro* examinations for prospective drug candidates is drug metabolism. The rate of drug metabolism is a key determinant of its elimination half-life from the body and thus the length of time it is able to engage with its target to elicit the biological response. Ultimately, this determines dose frequency, with rapidly metabolised drugs requiring multiple administrations per day to maintain blood concentrations. This can have a significant negative impact on therapy for the patient, particular regarding compliance. Furthermore, as previously mentioned, the preferred route of administration of a drug is oral administration and the first port of call following absorption from the small intestine is the liver. Consequently, if a drug is extensively metabolised on its first pass through the liver, very little will reach the systemic circulation and its site of action, resulting in poor bioavailability.[143] A measure of a drug's susceptibility to metabolism, in terms of its half-life and rate of hepatic clearance can be determined *in vitro* using liver tissue homogenates (S9 fractions) or hepatocyte cells to give an indication of whether to progress to *in vivo* metabolism studies. The S9 fraction contains a mixture of phase I (majority) and phase II (some) metabolism enzymes.[143] For this project, this work was outsourced to Cyprotex Discovery Ltd where they tested the metabolic stability of inhibitors **CI709** (**41**), **68** and **76** in a mouse S9 fraction assay *in vitro*, **Table 18**.

*Table 18: S9 metabolic stability data for **CI709**, **68** and **76**.*



CI709 (41)
K_i 210 nM

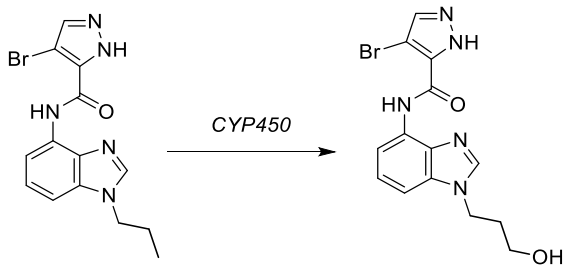
68
K_i 19 nM

76
K_i 55 nM

	CI709 (41)	68	76
CL_{int} (µL/min/mg protein)	61.3	27.5	200
$t_{1/2}$ (mins)	11.3	25.2	3.5

***CI709** (**41**) and **68** n=5 and **76** n=3*

In order to identify an effective dose of a drug it is pivotal to know how much reaches the general circulation and for how long it remains before it is excreted. These are measured by the half-life ($t_{1/2}$) and the intrinsic clearance (CL_{int}), where $t_{1/2}$ is the time it takes the concentration of a drug in the plasma to decrease by 50%.[133] Therefore, longer $t_{1/2}$ and low CL_{int} values are favoured as they infer greater metabolic stability. At this stage, we hoped to achieve a $t_{1/2}$ of >60 mins. **Table 18** illustrates that this property requires further improvement, with all three inhibitors exhibiting values significantly lower than $t_{1/2}$ >60 mins. The most potent inhibitor **68** was also the most metabolically stable ($t_{1/2}$ = 25.2 mins, CL_{int} = 27.5 µL/min/mg of protein) followed by the *n*-propyl analogue **CI709** (**41**) ($t_{1/2}$ 11.3 mins, CL_{int} = 61.3 µL/min/mg of protein) and finally the most unstable is the cyclopropyl analogue **76** ($t_{1/2}$ 3.5 mins, CL_{int} = 200 µL/min/mg of protein). These results highlight that the replacement of an aliphatic CH_3 for the electronegative chloride increases $t_{1/2}$ 2-fold. This increase in stability could suggest that the terminal methyl of **CI709** (**41**) is subject to aliphatic hydroxylation by CYP450 **Scheme 3**, **95**, which is blocked when replaced by the chloride in **68**. Other metabolic transformations elsewhere in the common scaffold are clearly occurring to account for the relatively rapid clearance of **68** (such as the electron rich benzimidazole moiety), but the rapid clearance of both **CI709** (**41**) and **76** suggests that the *N*-3 substituent is a particularly labile site, and needs to be considered if we are to improve the metabolic stability to reach our desired *in vitro* $t_{1/2}$ of >60 mins before committing to *in vivo* PK studies.

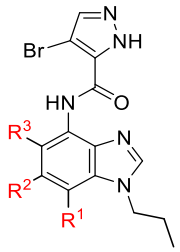


41 (CI709) **95**

*Scheme 3: Potential phase I aliphatic hydroxylation of **CI709** that is blocked in **68**.*

An alternative approach to block this site that avoids the electrophilic chloride is to use the trifluoromethyl bioisostere. However, our previous investigations generated this analogue **69**, which was less potent than **68** (K_i 179 nM and 19 nM respectively; see **Section 3.3.4**), and nicely illustrates the fine line that medicinal chemists have to tread when trying to balance improvements in PK properties with the PD requirement of maintaining potency. An alternative strategy to increase the metabolic stability could be the installation of electron-withdrawing substituents on the LHS of the

benzimidazole to block potential metabolic hotspots on the aryl ring, **96 Figure 89**. Again, these compounds would need to be assessed for potency against DYRK2 to ensure that activity against the clinical target has not been compromised by the changes made to address metabolism.



96

Figure 89: Potential metabolic sites to block.

This data reinforces the importance of investigating *in vitro* PK at an early stage. Our most potent inhibitor **68** (K_i 19 nM) has moderately metabolic stability ($t_{1/2}$ 25.2 mins) and as it contains an electrophilic handle, it has the potential to be toxic. However, from a positive perspective, the *n*-propyl derivative **CI709** (**41**) has impressive aqueous solubility and permeability. Therefore, further investigation is required in order to deliver an inhibitor with increased metabolic stability ($t_{1/2}$ >60 mins) whilst maintaining the solubility and permeability of inhibitor **CI709** (**41**).

Although we face an inherent stability challenge for inhibitors **68** and **CI709** (**41**), we have encouraging evidence to support that they are highly soluble and permeable, if we assume that our most potent inhibitor **68** provides similar results to **CI709** (**41**). Therefore, as we were interested in determining the effect of DYRK2 inhibition in TNBC cells, the next step was to assess their ability to 1) reduce cell growth and 2) inhibit the phosphorylation of HSF1, the master regulator of proteotoxic stress.

5 Assessment of 68 and CI709 (41) in TNBC Cells

DYRK2 (Dual Specificity Tyrosine(Y) Regulated Kinase-2) is a kinase that plays an important role in cancer progression. More specifically, it is significant to TNBC *via* the activation of the 26S proteasome and HSF1.[79] Moreover, inhibition of DYRK2 has been reported to decrease cell proliferation in TNBC cell lines and also reduce tumour burden in TNBC mice xenografts.[90] Due to this, we were interested in investigating DYRK2 as a potential downstream target for the treatment of TNBC with small molecule inhibitors.

We have identified a relationship between DYRK2 and HSF1 in TNBC and HEK293T cells. DYRK2 and HSF1-KO studies, by our collaborator de la Vega, confirmed that both proteins are essential for TNBC tumour growth. In addition, we have provided evidence that DYRK2 is a positive modulator of HSF1 *via* phosphorylation. Moreover, from a clinical study by our collaborator Edwards, we have identified that low DYRK2 expression relates to an increase in the OS of TNBC patients.

These results provided us with the incentive to investigate the effect of DYRK2 inhibitors on 1) cellular growth and 2) the phosphorylation of HSF1, in TNBC cells.

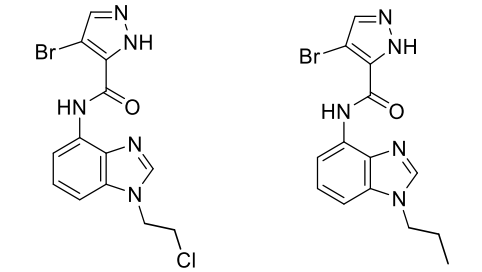


Figure 90: Inhibitors examined in cellular assay.

We chose to investigate the most potent inhibitor **68**, an inhibitor which maintained isoform selectivity, from this project and the previous lead compound, **CI709** (**41**), **Figure 90**. Previous work determined **CI709** to be selective among a panel of 40, kinome representative, kinases (see **Appendix 10.1**). Although we identified inhibitors which were more active than **CI709** (**41**), described in **Section 3** we chose to investigate this analogue to complement the large body of data obtained for this compound (e.g. the selectivity, solubility and permeability data), which we do not have for alternative more active inhibitors. For example, with a permeability of 22.7 $x10^{-6}$ cms^{-1} and an efflux ratio of 1.21, we knew that **CI709** (**41**) could permeate

cells (and by inference its close analogue **68**) to engage with DYRK2 in an intracellular environment.

Our collaborators provided clinical data to support that DYRK2 expression negatively affects TNBC patient survival. Additionally, they identified a relationship between DYRK2 and HSF1 *via* phosphorylation at two significant sites (Ser320 and Ser326). We have reason to believe that DYRK2 stabilises HSF1 *via* phosphorylation in TNBC cells which could lead to tumour growth (see **Section 1.5.4**).

From this work, we discovered an inhibitor with a >10-fold increase in potency **68** (K_i 19 nM) compared to the previous lead compound **CI709** (**41**) (K_i 210 nM), **Figure 90**.

The following sections describe the assessment of these two inhibitors in cell lines, both in wild type and genetically modified MDA-MB-468 TNBC and HEK293T cells. For the figures in this section, lead inhibitor **68** is referred to as **LB35**. This work was carried out by our collaborator, de la Vega.

Initially, we investigated the effect of the inhibitors on the proliferating ability of the TNBC cells.

5.1 Cell proliferation

The basal-like MDA-MB-468 cells were treated with inhibitors **CI709** (**41**) or **LB35** (**68**) at varying concentrations for 5–7days, **Figure 91**.

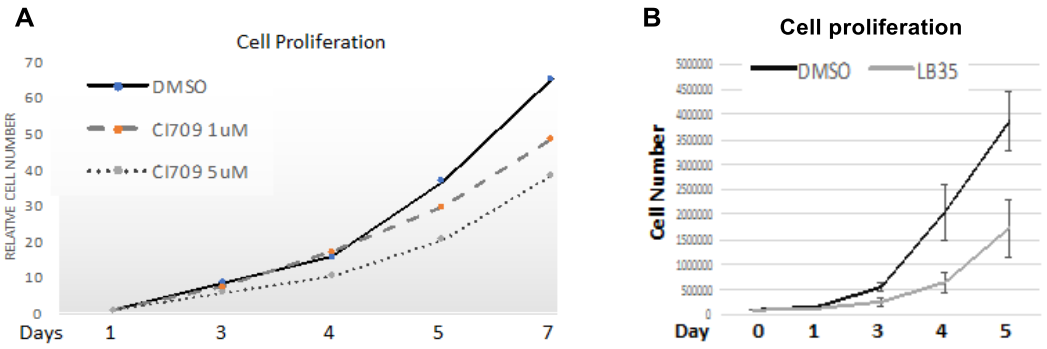


*Figure 91: Effect of **A**) **CI709** at 1 and 5 µM and **B**) LB35 (**68**) at 1 µM on cell count in MDA-MB-468 cells.*

With respect to **A**, **Figure 91**, with **CI709** (**41**) at 1 and 5 µM, the results highlight a significant reduction in the cell count compared to the control after 7 days. For the first 4 days of the experiment, treatment of **CI709** (**41**) at 1 µM resulted in similar growth to the control experiment. However, increasing the concentration of **CI709** (**41**) to 5

µM produced a slower rate of proliferation compared to the control and 1 µM **CI709** (**41**), **Figure 91**.

With respect to **B**, **Figure 91**, treatment with 1 µM **LB35** (**68**) resulted in a substantial reduction in cell count after 5 days compared to the control experiment. Moreover, from day 1 onwards the cell count increased at a notably slower rate with 1 µM **LB35** (**68**) compared to the control. Therefore, a lower concentration of **LB35** (**68**) compared to **CI709** (**41**) was required to achieve results. This is unsurprising as **LB35** (**68**) is 10× more potent than **CI709** (**41**) against the isolated target enzyme (Ki 19 nM vs 210 nM respectively).

Pleasingly, we have determined that in the TNBC cell line MDA-MB-468, treatment with inhibitors **CI709** (**41**) at 5 µM and **LB35** (**68**) at 1 µM results in a significant reduction in cancer cell growth over the course of 7 days.

Next, we investigated the effect of these inhibitors on the relative expression of an important protein in the HSF1 pathway, Hsp70.

5.2 Hsp70 expression

We reported in **Section 1.5.4** that knocking out DYRK2 in MDA-MB-468 cells results in a decrease in Hsp70 expression. Therefore, we would expect a similar result by introducing inhibitors, **CI709** (**41**) and **LB35** (**68**), **Figure 92**.

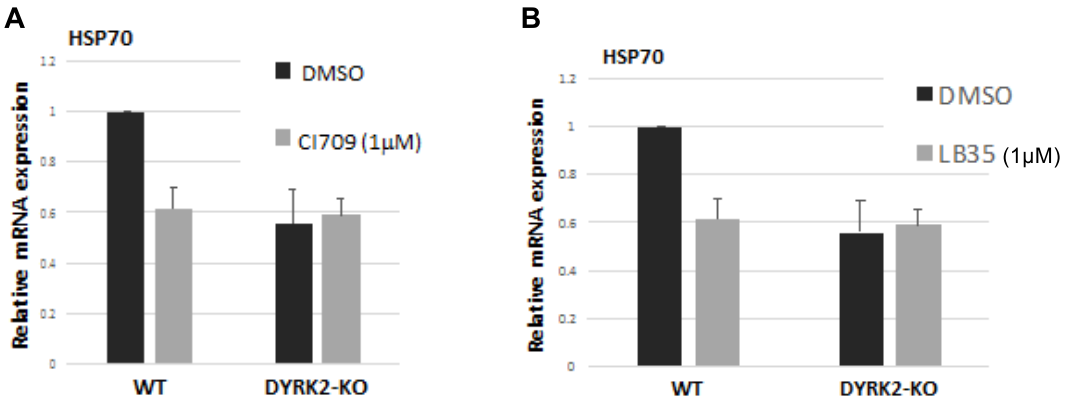


*Figure 92: HSP70 expression in MDA-MB-468 WT/DYRK2-KO cells treated with DMSO or 1 µM **A) CI709** or **B) LB35** (**68**).*

From these results, **A** and **B Figure 92**, we can deduce that both **CI709** (**41**) and **LB35** (**68**) reduce the expression of Hsp70 at the mRNA level in the WT compared to the control to a similar level at the same concentration. The assessment of the inhibitors in the DYRK2 KO cells provides evidence that the effect on HSP70 seen in

the WT cells is through engagement with DYRK2. Essentially, because the inhibitors produce no additional change to the expression of HSP70 mRNA in the absence of the target enzyme, we can infer that reduction of HSP70 levels in both WT and DYRK2 KO cells are through the selective inhibition of DYRK2 by our inhibitors. Moreover, the DYRK2-KO experiments confirm the importance of DYRK2 in the production of HSP70, as the expression is comparable to the inhibited WT by our compounds, **Figure 92**.

5.3 HSF1 phosphorylation

Next, we set out to determine the effect of **CI709** (**41**) and **LB35** (**68**) on HSF1 phosphorylation and its link with DYRK2. We had previously shown that a direct relationship existed between DYRK2 and HSF1 using the transfected immortal kidney cell line HEK293T with a DYRK2 analog-sensitive (AS) mutant, (see **Figure 35** in **Section 1.5.4**). Upon exposure to the specific inhibitor, 1NMPP1 **23**, phosphorylation of HSF1 at Ser320 and Ser326 was diminished in a concentration-dependent response, (**Figure 35** in **Section 1.5.4**) (lanes 3-5), thus demonstrating a synergistic relationship between DYRK2 and HSF1 *via* phosphorylation. Complementary to this, comparison of cells transfected with the flag-tagged DYRK2 WT or KD mutant revealed that DYRK2 was responsible for the phosphorylation of HSF1 at Ser320 and Ser326, **Figure 36**.

To investigate whether our inhibitors could recapitulate these phospho-inhibitory effects in HEK 293T cells transfected with or without GFP-tagged DYRK2, **CI709** (**41**) or **LB35** (**68**) were incubated with the cells for 3 hours, **Figure 93**.

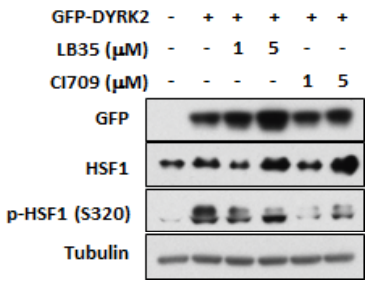


*Figure 93: HEK 293T cells (+/- GFP-DYRK2) were treated with **LB35** (**68**) or **CI709** (**41**) (1 and 5 µM).*

First of all, these results show that upon introduction of GFP-DYRK2, HSF1 is phosphorylated at Ser320, **Figure 93** (lane 2). Next, the treatment with **LB35** (**68**) and **CI709** (**41**) (1 and 5 µM), results in a reduction in the phosphorylation of HSF1 at Ser320 (lanes 3-6, upper blots). Therefore, from these results we can conclude that

both inhibitors elicit inhibitory effects on the phosphorylation of HSF1 at Ser320 in HEK 293T cells.

Previously, we had been able to demonstrate that upon heat shock stress stimulus, protein levels of phosphorylated HSF1 at the two sites (S320 and S326) were increased in MDA-MD-468 cells **Figure 34**. To explore whether our inhibitors could inhibit this effect, MDA-MD-468 cells were treated with DMSO (-) or the inhibitors **CI709** (**41**) or **LB35** (**68**) for 2 hours and then exposed to HS (42 ºC) for 1 hour, **Figure 94**.

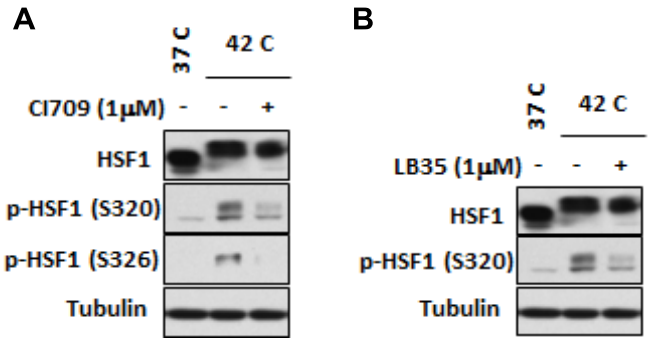


*Figure 94: MDA-MB-468 cells treated with **A**) 1 µM **CI709** (**41**) or **B**) 1 µM **LB35** (**68**) in the presence of HS.*

Both inhibitors significantly reduced phosphorylation of HSF1 at Ser320 when cells were subjected to HS (lane 3, upper blot, **A** and **B**) when compared to control (lane 2, upper blot, **A** and **B**), **Figure 94**. When treated with **CI709** (**41**), **A** phosphorylation of Ser326 is also significantly reduced (lane 3) compared with control (lane 2) (pSer326 for **LB35** (**68**) was not determined). We can therefore conclude that our inhibitors reduce phosphorylation of HSF1 induced by the HS stress stimulus at sites that we have shown from previous experiments are phosphorylated by DYRK2.

5.4 Proteasome importance

In addition, we investigated the effect of **CI709** (**41**) and **LB35** (**68**) on the phosphorylation of a secondary protein, seven in absentia homolog-2 (SIAH2), which is essential in the protein degradation process. More specifically, SIAH2 is an E_3 ubiquitin ligase, which tags proteins for recognition by the proteasome. We previously mentioned, in **Section 1**, that DYRK2 has been reported to play an important role in protein degradation *via* phosphorylation of Rpt3 on the 26S proteasome.[79] In addition, DYRK2 has been reported to activate SIAH2 *via* phosphorylation at five sites, where one of these is Ser28.[144] SIAH2 levels have been reported to be negatively correlated

to the clinical outcomes of BCs, where increased levels of SIAH2 have been observed in aggressive BC subtypes, including TNBC.[145] Studies have also shown that high levels of SIAH2 were associated with reduced DFS (disease free survival).[146]

HEK 293T cells were transfected with or without GFP-tagged DYRK2 and after 48 hours, inhibitors **CI709** (**41**) or **LB35** (**68**) were added for 3 hours, in order to determine the influence of our inhibitors on the phosphorylation of SIAH2, **Figure 95**.

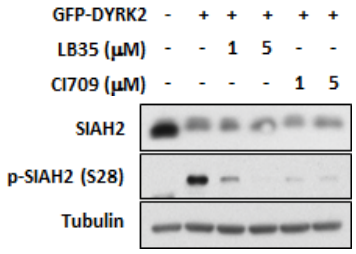


*Figure 95: HEK 293T cells (+/- GFP-DYRK2) were treated with LB35 (**44**) or CI709 (1 and 5 µM).*

From this work, we have confirmed that DYRK2 phosphorylates SIAH2 at Ser28 in HEK 293T cells (lane 2), **Figure 95** and that treatment with **CI709** (**41**) and **LB35** (**68**) at 1 and 5 µM significantly inhibits this process (lanes 3–6), **Figure 95**. Pleasingly, these results complement previous work, which illustrated the importance of DYRK2 in protein degradation. It is possible that DYRK2 plays a number of roles in this process, activating the proteasome and priming SIAH2 for ubiquitination of its protein targets. Disruption of the normal function of the proteasome is common in cancer and can cause accumulation of toxic proteins. Thus, investigation into downstream activators of the degradation process could provide an insight into the treatment options for cancers with high proliferative rates.

5.5 Summary

Through collaboration with de la Vega (University of Dundee) we have been able to identify a relationship between DYRK2 inhibition and TNBC cell growth, and also the activation of HSF1. The treatment of TNBC cells with our inhibitors, **CI709** (**41**) and **LB35** (**68**), has resulted in a significant decrease in cell proliferation and upon HS, these inhibitors have been shown to inhibit the phosphorylation of HSF1 at two sites (Ser320 and Ser326). In addition, investigation into the relationship between DYRK2 and the proteasome has provided us with evidence to support that DYRK2 phosphorylates SIAH2 at Ser28 and that treatment with our inhibitors, **CI709** (**41**) and **LB35** (**68**), impedes the phosphorylation.

6 Chemistry

As described in **Section 3**, the aims of this project were to investigate the effect of manipulating a number of positions on the lead scaffold on both the DYRK2 activity and selectivity against closely related isoform DYRK1A. Moreover, we set out to investigate 4 specific areas of the target scaffold as shown in **Figure 96**: the head group heterocycle and its substitution (blue and red), benzimidazole substitution (green) and the amide linker (pink).

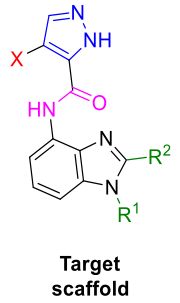
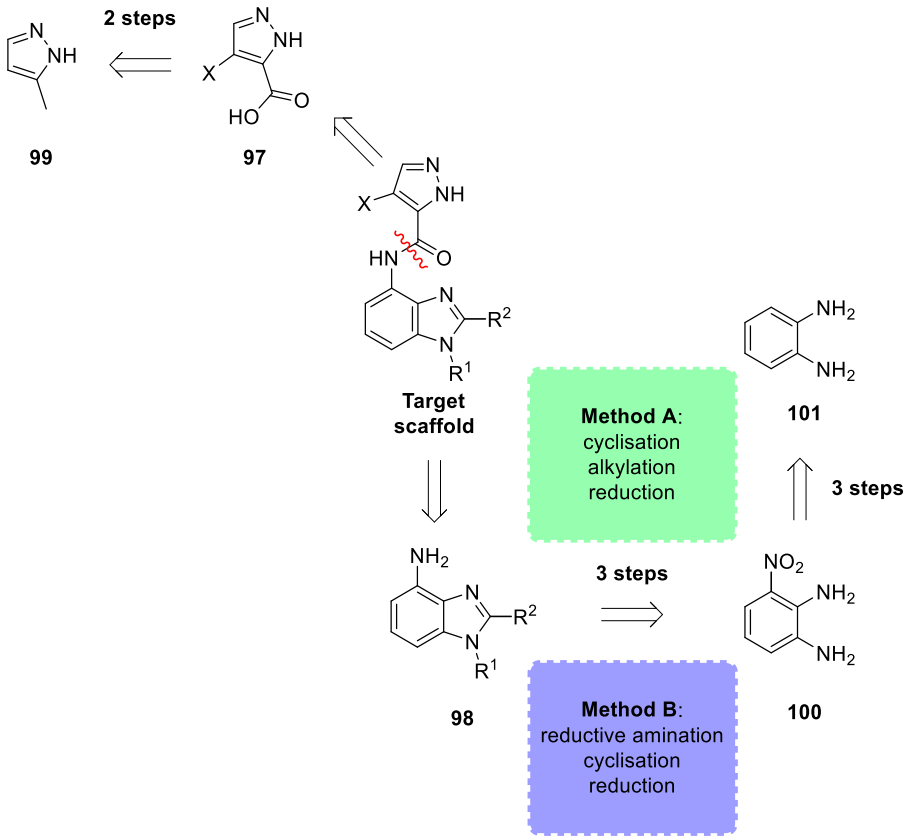


Figure 96: Target scaffold.

In order to approach the synthesis of this target scaffold, if we first consider its retrosynthetic analysis, we can access it through an amide coupling reaction with acid **97** and aniline **98**, **Scheme 4**.



Scheme 4: Retrosynthesis of target scaffold.

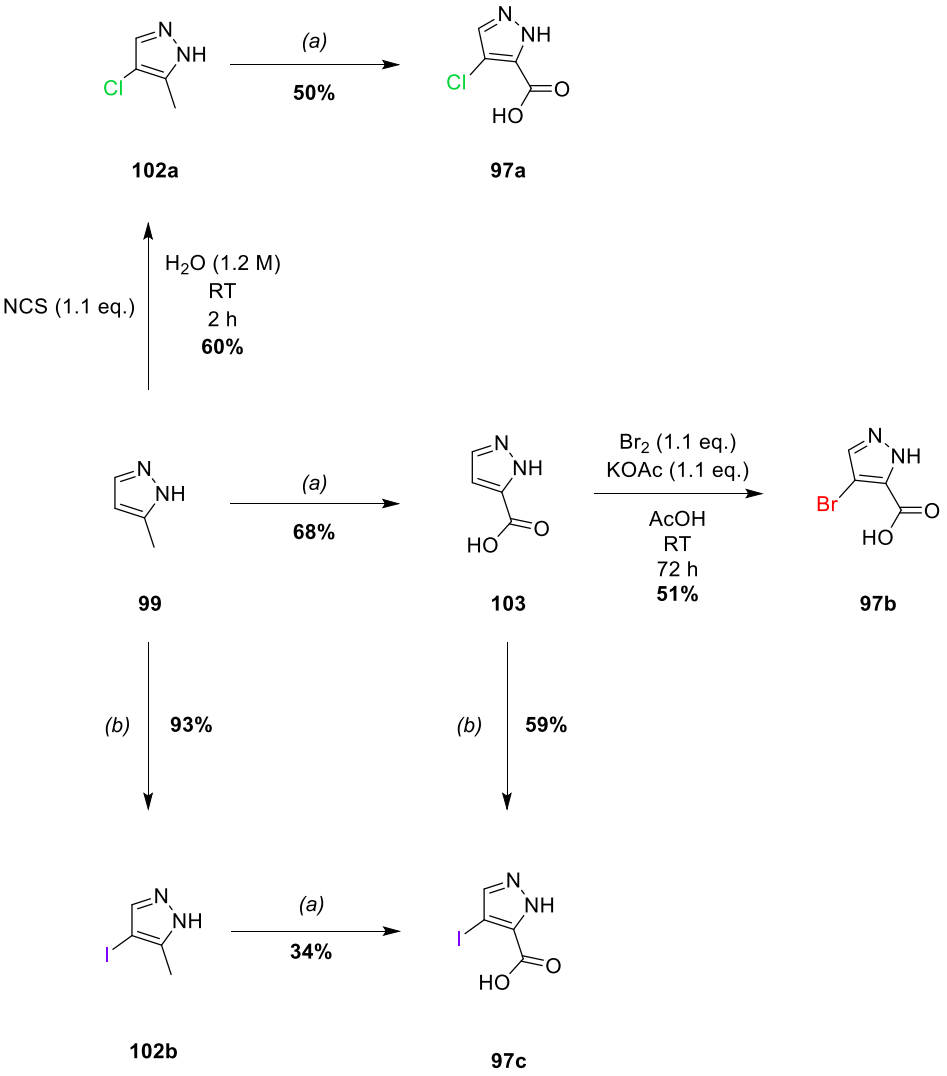
This allowed a tandem synthesis of acid **97** and aniline derivatives **98**. Acid **97** can be synthesised in 2 steps from commercially available 5-methyl-1H-pyrazole **99**. Next, aniline **98** can be accessed in 3 steps from the commercially available starting material 3-nitro-1,2-benzenediamine **100** *via* two methods. Our original strategy (Method A, **Scheme 4**) employed a cyclisation, alkylation and reduction route to furnish the anilines **98**. However, in order to introduce R^2, we devised an alternative 3 step method to deliver the desired anilines **98**. This consisted of a reductive amination, cyclisation and reduction (Method B, **Scheme 4**). 3-Nitro-1,2-benzenediamine **100** (moderately expensive, Sigma Aldrich 5g = £177)[147] can be synthesised *via* 3 steps from inexpensive commercially available starting material *o*-phenylenediamine **101** (Sigma Aldrich 500g = £54.30)[148], **Scheme 4**.

Taking a closer look at the synthesis of acid **97**, it is possible to prepare a number of halogenated analogues in a similar fashion.

6.1 Head group synthesis

6.1.1 Halogenated acids

The synthesis of chloro-, bromo- and iodo-substituted pyrazole carboxylic acids was achieved using common conditions and intermediates, **Scheme 5**.



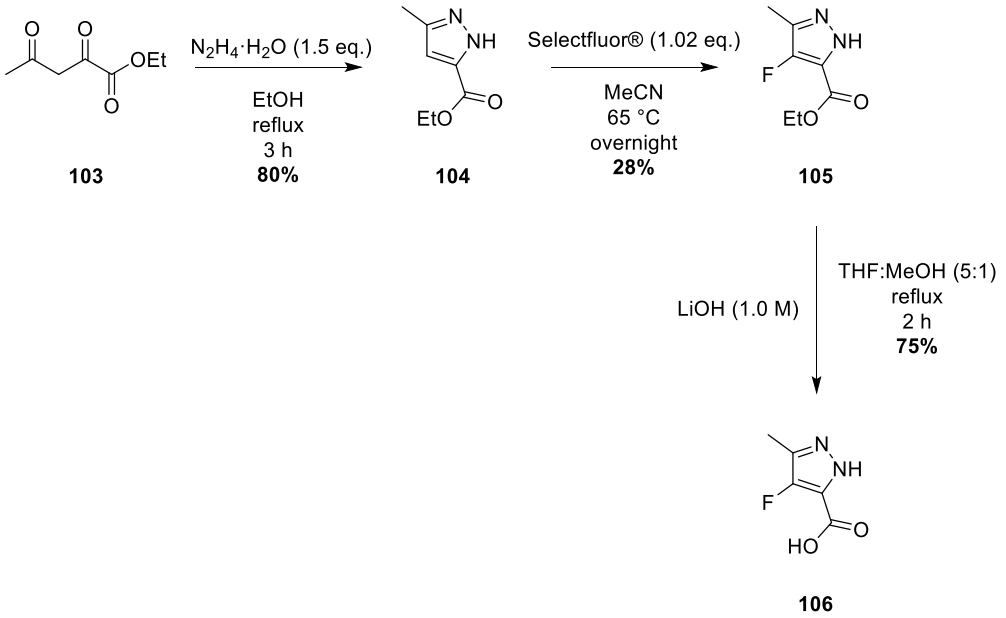
(a) $KMnO_4$ (3.0 eq.), H_2O, reflux, 53–63 h
(b) NIS (1.2 eq.), H_2SO_4, 0 ºC–RT, overnight

Scheme 5: Synthesis of 4' halogenated acids.[149-152]

Within the group, a route to access the chloro derivative had already been established[151], **Scheme 5**. Starting from the commercially available pyrazole **99**, treatment with *N*-chlorosuccinimide (NCS) at room temperature afforded intermediate **102a** (60%). Subsequent treatment with potassium permanganate ($KMnO_4$) afforded the desired carboxylic acid **97a** with an overall yield of 30%. Alternatively, the bromo derivative **97b** was accessed in the reverse order as when the Br was installed in the

first step, the oxidation process did not go to completion. Commercially available pyrazole **99** was treated with $KMnO_4$ to provide carboxylic acid **103** (68%) after 3 days at reflux. Without further purification, acid **103** was treated with bromine in acetic acid at room temperature. This gave the desired bromo acid **97b** *via* 2 steps with an overall yield of 35% (**Scheme 5**). Finally, the iodo derivative **97c** could be synthesised *via* two methods. Similar to the bromo derivative, treatment of pyrazole **99** with $KMnO_4$, followed by *N*-iodosuccinimide (NIS) under acidic conditions afforded the desired iodo substituted acid **97c**.[152] Following this two-step route, we gained access to **97c** with an overall yield of 40%. Alternatively, it was also possible to synthesise **97c** in the reverse order. Iodination of commercially available **99** with NIS provided **102b** in a 93% yield. However, oxidation of this intermediate **102b** with $KMnO_4$ produced the desired acid **97c** in only 34% yield. Following this route, the overall yield was 32% over two steps. Therefore, iodinated acid **97c** can be synthesised *via* one of two routes, preferentially *via* the first route described due to the higher yield obtained.

Unfortunately, the synthesis of the final halogen derivative, the fluoro analogue, has proven to be challenging and therefore has been unsuccessful. However, a structurally similar derivative **106** was synthesised, **Scheme 6**.

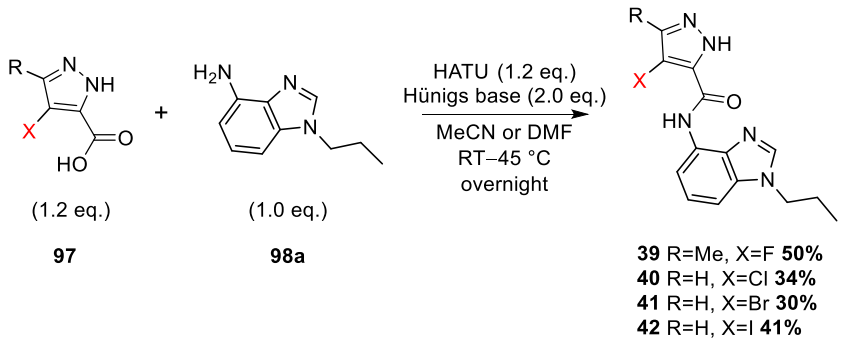


*Scheme 6: Synthesis of substituted fluoro derivative **106**.[153]*

Treatment of di-keto ester **103** with hydrazine hydrate in ethanol (EtOH) afforded intermediate ester **104** in excellent yields (80%) after 3 hours at reflux. Following this, **104** was subjected to the reported fluorination conditions[153] with Selectfluor® and upon purification *via* flash column chromatography, the fluoro ester **105** was isolated

in a 28% yield. Finally, the ester was hydrolysed under basic conditions to afford the acid **106** in a 75% yield, with an overall yield of 17% over 3 steps, **Scheme 6**.

The carboxylic acids were then reacted with the aniline **98a** (synthesis described in **Section 5.2.2**) under basic conditions to generate the desired amide products, **Scheme 7**.



*Scheme 7: Formation of **39**–**42**.*

With the use of the amide coupling reagent HATU, **107a Figure 97**, in the presence of hünigs base, **107b Figure 97**, the reactions of acids **97** with aniline **98a** proceeded at ambient temperatures in either acetonitrile (MeCN) or *N,N*-dimethylformamide (DMF) overnight. Upon purification *via* flash column chromatography, the desired amides (**39**–**42**) were accessed in moderate yields (30–50%), **Scheme 7**.

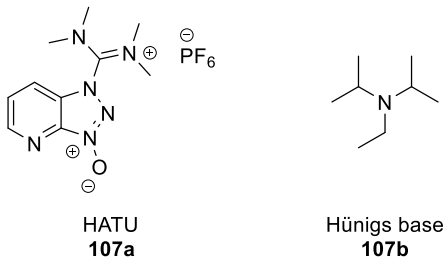
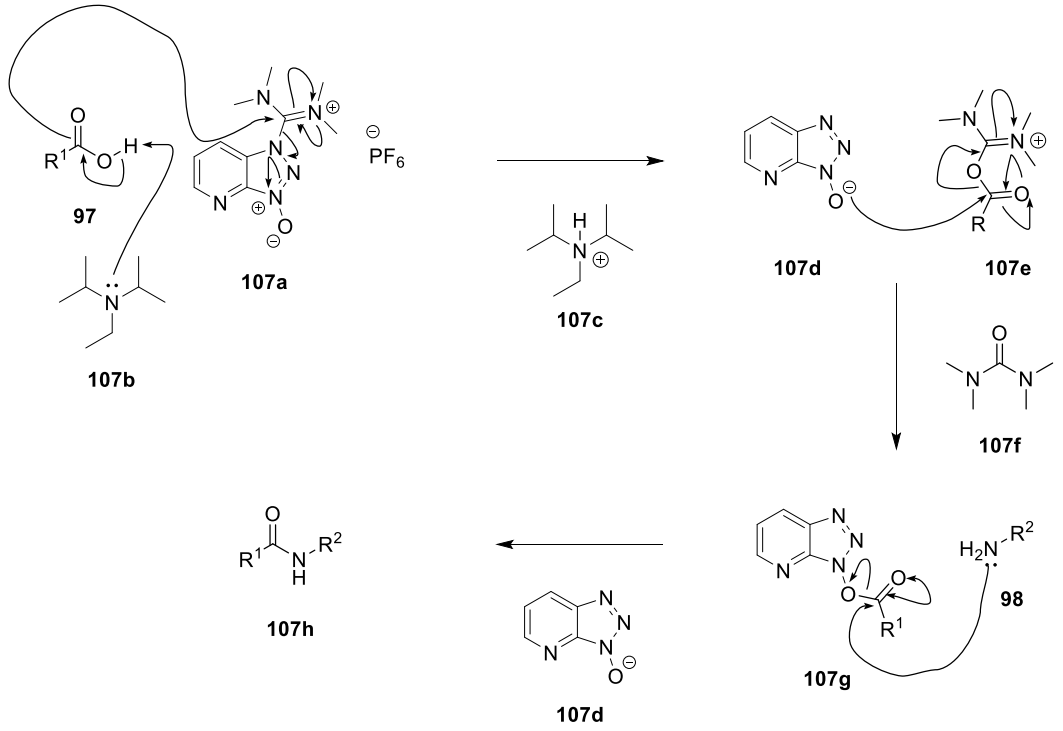


Figure 97: Structure of HATU and hünigs base.

One of the advantages of the amide coupling reagent HATU is that the by-products of the reaction are water soluble, **Scheme 8**.



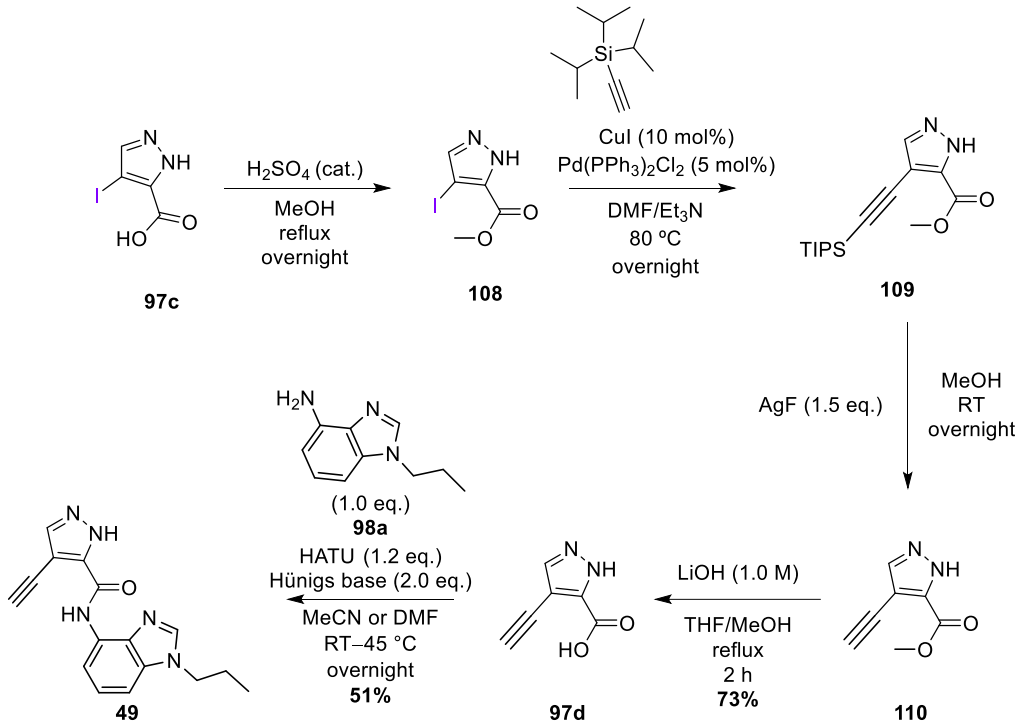
Scheme 8: Mechanism of amide coupling process with HATU.

The nucleophilic carboxylic acid **97** reacts with the electrophilic iminium of HATU **107a** resulting in the release of the pyridinolate **107d** and the formation of iminium **107e**. Hünigs base, **107b**, mops up the proton from the carboxylic acid **97** to form the water soluble protonated species **107c**. Next, the nucleophilic pyridinolate **107d** reacts with the electrophilic iminium centre of **107e** which results in the release of the water soluble dimethyl urea **107f** and the formation of the activated ester **107g**. Finally, the nucleophilic aniline **98** reacts with the electrophilic activated ester **107g**, resulting in the release of the water soluble pyridinolate **107d** and the formation of the desired amide product **107h**.

The investigation of substituents on the 4′ position of the pyrazole led us to the synthesis of the ethynyl acid **97d** (see **Section 3.2.2.3**) as it has been reported to be a bioisostere of halogen atoms. Fortunately, we had access to a number of halogenated pyrazoles primed for a palladium catalysed Sonogashira reaction.

6.1.2 Ethynyl analogue

More specifically, iodo acid **97c** could be manipulated in order to undergo a Sonogashira cross coupling reaction, **Scheme 9**.



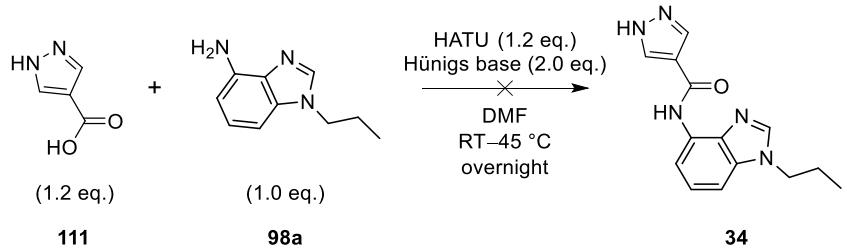
*Scheme 9: Route to ethynyl analogue **49**.*

Ethynyl acid **97d** can be accessed *via* 4 steps from the iodo acid **97c**, **Scheme 9**. The first three steps in the synthesis were carried out by Dr. Chris Lawson (see **Appendix 10.6.3** for experimental procedure). The most acidic proton of **97c** was protected by conversion to the methyl ester **108** under acidic conditions. Ester **108** was then treated with ethynyltriisopropylsilane in the presence of copper(I) iodide and a palladium(II) source to provide the TIPS protected alkyne **109**. Deprotection of the silyl group to provide the unsubstituted alkyne **110** was achieved by treatment with silver fluoride at room temperature. Next, the ester was hydrolysed under basic conditions to provide the desired acid **97d** in good yield (**73**%). The acid was then treated with aniline **98a** and HATU under basic conditions to provide the desired amide **49** in 51% yield after purification by column chromatography.

6.1.3 Alternative heterocycles

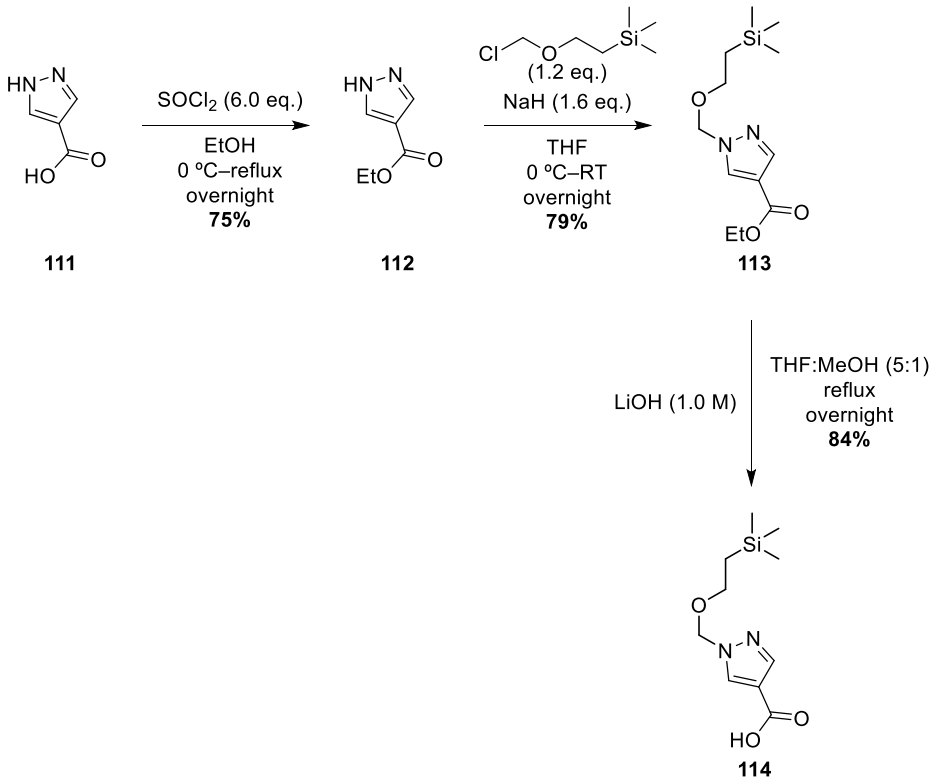
We investigated the effect of replacing the pyrazole head group with a range of different heterocycles. All but one of the inhibitors were synthesised *via* the direct amide coupling with the commercially available acids and the *n*-propyl derived aniline **98a**. This work was carried out by a masters student, Fiona Keatings under my supervision, within the group.[154] The associated experimental procedures for **31**–**33**, **35** and **37**–**38** can be found in **Appendix 10.6.2**.

The direct amide coupling of the pyrazole carboxylic acid isomer **111** and aniline **98a** was unsuccessful, **Scheme 10**. The crude material mostly consisted of unreacted starting materials.



*Scheme 10: Unsuccessful amide coupling of **111** and **98a**.*

We therefore proposed an alternative strategy with the use of a protecting group on the pyrazole nitrogen, **Scheme 11**.

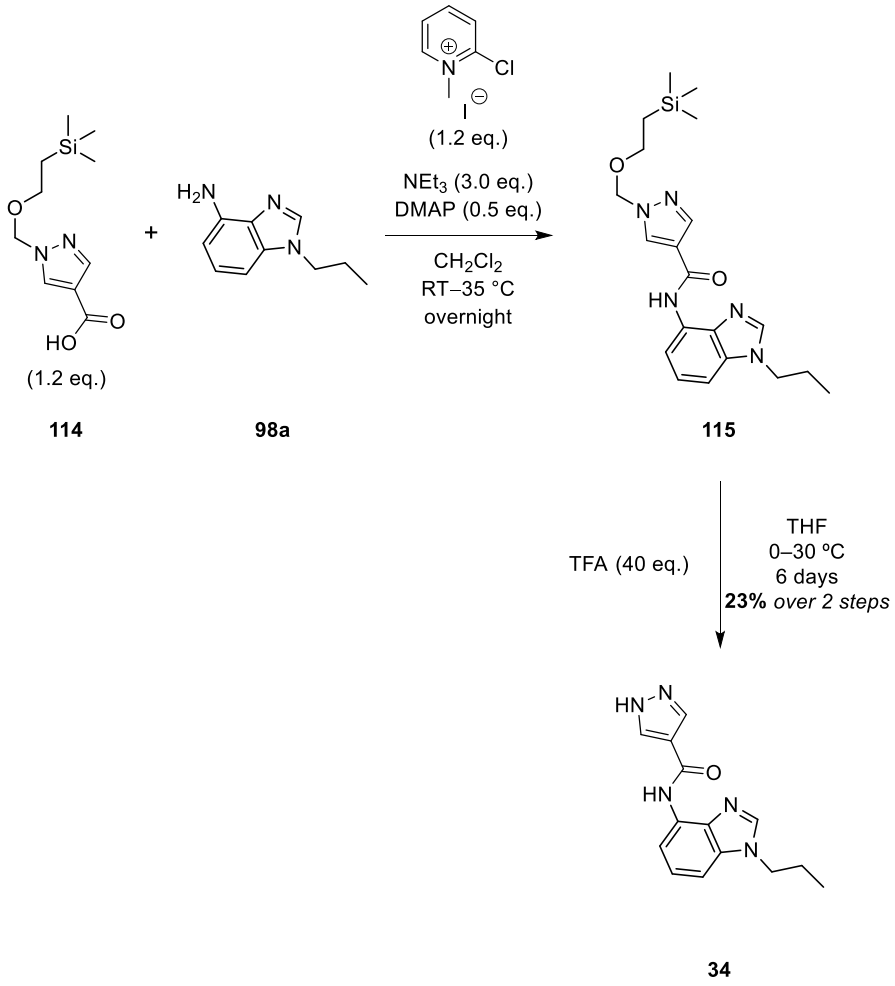


Scheme 11: Protecting group strategy.

Commercially available acid **111** was treated with thionyl chloride ($SOCl_2$) in EtOH at reflux to access ethyl ester **112** in good yield (75%). Next, without further purification, the treatment of ester **112** with 2-(trimethylsilyl)ethoxymethyl chloride (SEM-chloride) under basic conditions provided SEM protected pyrazole **113** in good yield (79%) after purification *via* flash column chromatography, **Scheme 11**. Finally, ester **113** was hydrolysed under basic conditions to provide acid **114** without further purification, in

excellent yield (84%) which was ready for the subsequent amide coupling. Acid **114** was accessed in 3 steps with an overall yield of 50%.

The next step was to subject the acid to amide coupling conditions with the aniline **98a**. The formation of the amide **115** was unsuccessful *via* the conditions using HATU, therefore, we investigated an alternative method, **Scheme 12**.

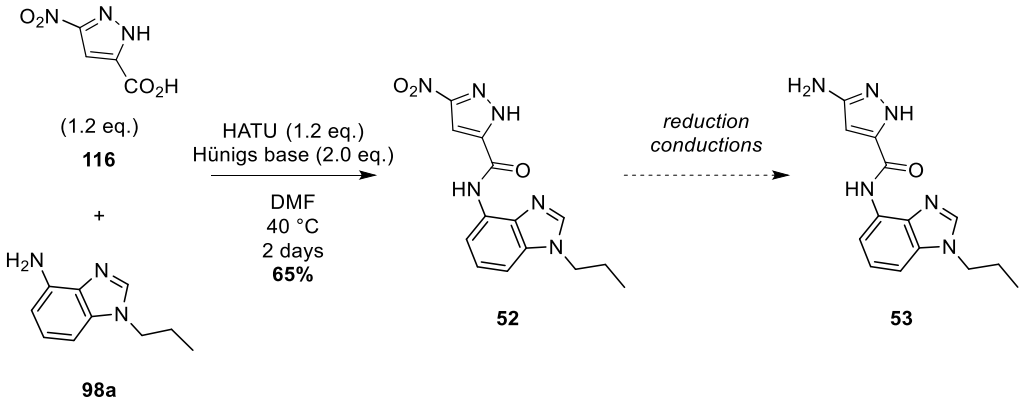


*Scheme 12: Alternative amide coupling to **34**.*

Acid **114** and aniline **98a** were subjected to the amide coupling conditions described above, **Scheme 12**. This resulted in an inseparable mixture of product **115** and aniline **98a**. Upon partial purification *via* flash column chromatography, the mixture was treated with trifluoroacetic acid (TFA) to de-protect the silyl group. After 6 days at 30 ºC and upon purification *via* column chromatography, the desired amide **34** was accessed in 23% over two steps.

6.1.4 3′ Amino pyrazole synthesis

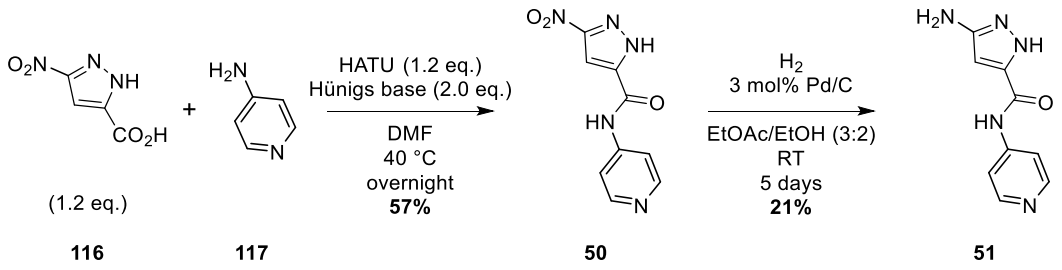
The next series of analogues included substitution at the 3′ position on the pyrazole head group. Starting with commercially available acid **116**, we proposed a two-step amide coupling, reduction process to synthesise the desired amide **Scheme 13**.



*Scheme 13: Route to amino pyrazole analogue **53**.*

Commercially available nitro acid **116** and aniline **98a** were treated with HATU and after 2 days at 40 ºC amide **52** was accessed in a moderate yield (65%). The next step was to reduce the nitro group of compound **52** to the corresponding amine **53**, however, after a brief investigation of reducing conditions, we identified that nitro amide **52** was poorly soluble in the corresponding reduction reaction solvents examined.

Whilst testing reaction conditions, a simpler derivative was chosen to investigate the conditions to avoid wasting material. 4-Amino pyridine **117** was subjected to the two-step amide coupling and reduction method, **Scheme 14**.

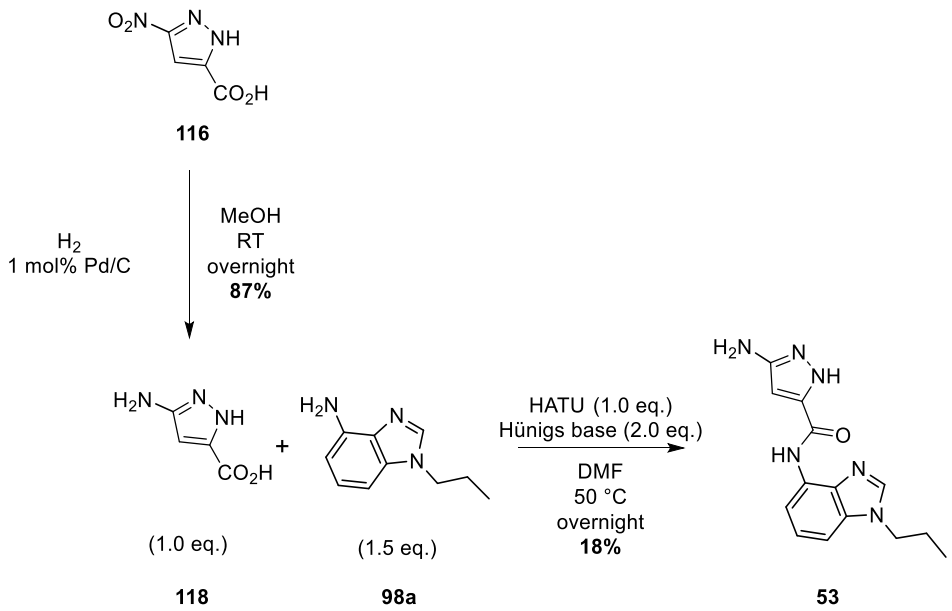


*Scheme 14: Two step route to amino pyrazole derivative **51**.*

Nitro acid **116** and 4-amino pyridine **117** were subjected to the HATU coupling conditions and provided nitro amide **50** in reasonable yield (57%). The next and final step was the reduction to the corresponding amine. Due to the poor solubility of **50** in

methanol (MeOH) or ethyl acetate (EtOAc) for the reduction process, a mixture of EtOAc and EtOH was required along with an extended reaction time of 6 days in order to facilitate the transformation. Upon further purification *via* flash column chromatography **51** was accessed in poor yield (21%). This two-step process provided **51** with an overall yield of 12%.

In order to circumvent the solubility problem of the nitro amide **52**, we proposed an alternative strategy to synthesise the desired amino pyrazole **53**, **Scheme 15**.

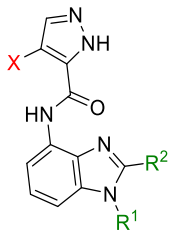


*Scheme 15: Alternative strategy to desired amino pyrazole **53**.*

Treatment of the commercially available acid **116** with palladium on carbon under a hydrogen atmosphere provided acid **118** in good yield (87%). Acid **118** was treated with aniline **98a** and HATU under basic conditions at 50 ºC and this provided the desired amide **53**, after purification *via* flash column chromatography in poor yields (18%). This poor yield could be explained by the competition between the amine of pyrazole **118** and benzimidazole **98a**. Whilst this problem may have been overcome by changes in the stoichiometry of the reaction, sufficient quantities of **53** were obtained through this method in order to examine this molecule within our biological assay.

6.2 Tail group synthesis

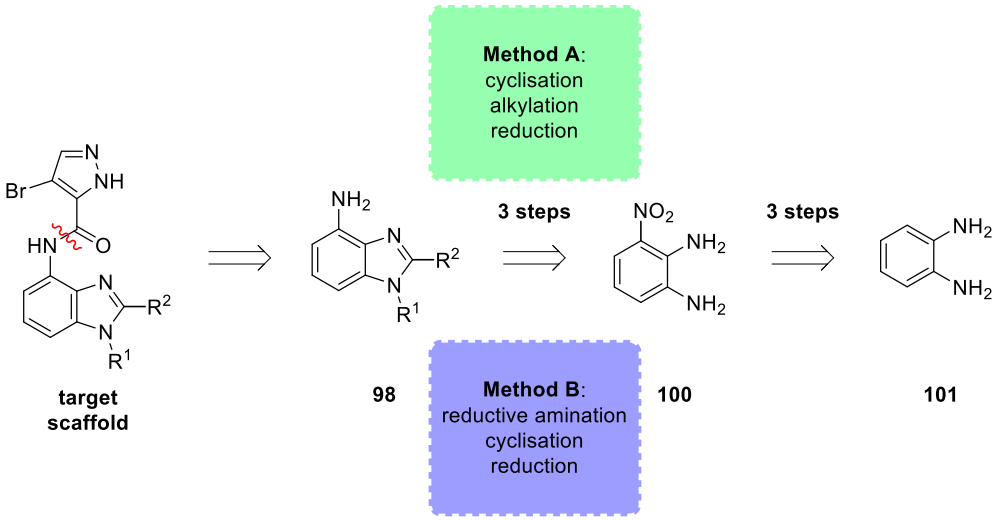
The following section will discuss the approaches to the synthesis of the amino benzimidazole half of the target inhibitor scaffold, **Figure 98**.



target scaffold

Figure 98: Inhibitor scaffold.

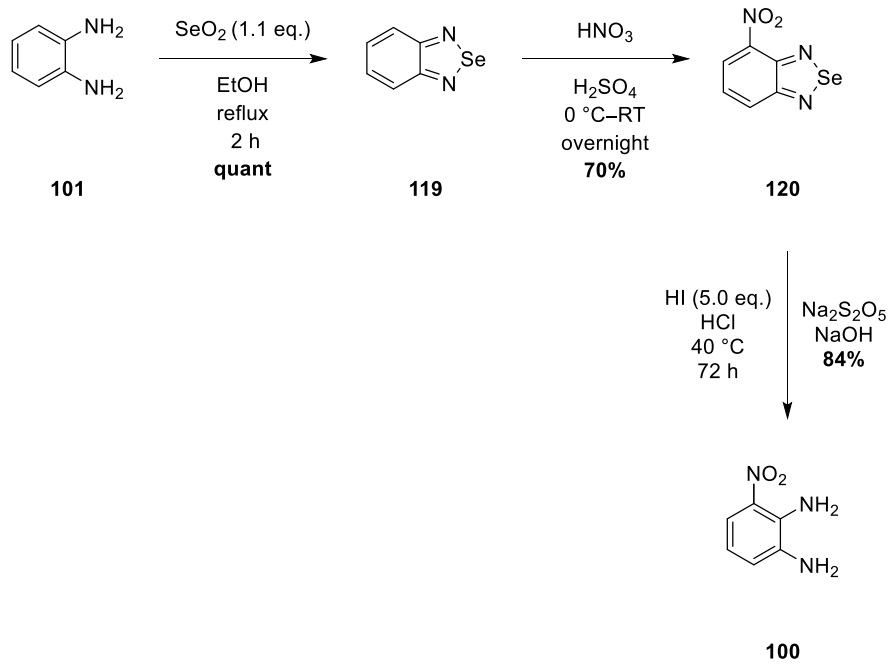
In the retrosynthetic analysis of the target scaffold at the start of **Section 6**, we outlined the two potential routes to synthesising the amino benzimidazoles **98**, **Scheme 16**.



*Scheme 16: Retrosynthesis of **98**.*

Due to the low cost of di-amine **101** compared to nitro derivative **100**, we set out to synthesise **100** on a multi-gram scale. Aershot *et al.*, reported a high yielding, multi-gram, three-step synthesis to nitro aniline **100** from di-amine **101** starting material, **Scheme 17**.[155]

6.2.1 Synthesis of 3-nitro-1,2-benzendiamine 100



*Scheme 17: Three-step route to nitro aniline **100**.*[155]

Commercially available di-amine **101** was treated with selenium dioxide in EtOH for 2 hours to provide selendiazole **119** in quantitative yields. Selendiazole **119** was then treated with nitric acid and, after 12 hours at room temperature, and following subsequent recrystallization in toluene, 4-nitroselendiazole **120** was isolated as a single isomer in good yield (70%). The regiochemistry of this transformation was confirmed by ^{1}H NMR spectroscopy. The final step was the cleavage and de-protection of the selendiazole **120** to the corresponding di-amine **100**, by treatment of **120** with hydriodic acid (HI) at 40 ºC for 3 days. Upon pH adjustement with NaOH, this provided starting material **100** with an overall yield of 59% over 3 steps.

It should be noted that the original reported conditions, for the final cleavage step to furnish **100**, used a very large excess of HI. HI is an extremely toxic and corrosive substance. Therefore, due to the large scale of the reactions carried out, we decided to investigate the effect of reducing the equivalents of HI on the progress of the reaction, **Table 19**.

Table 19: Optimisation of cleavage reaction.

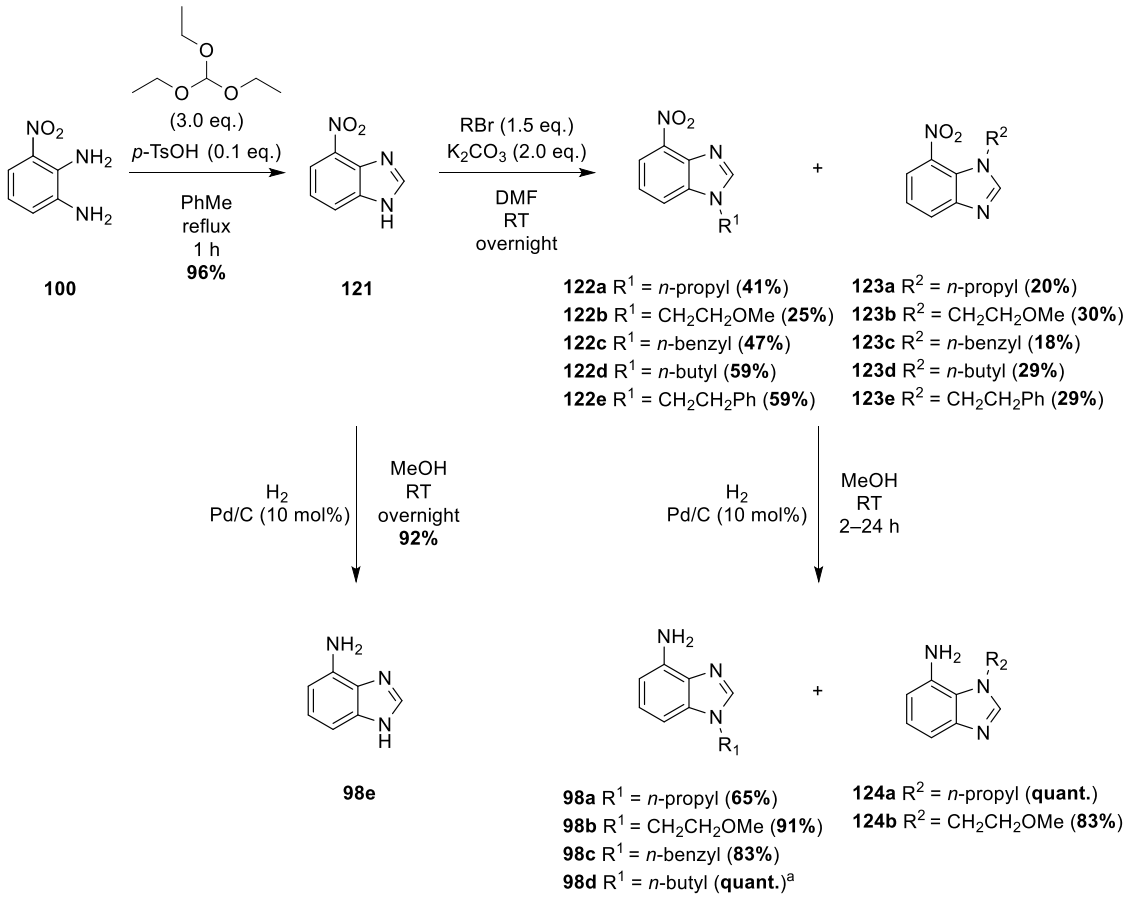
	Reported Conditions[155]	Optimised Conditions
HI (eq.)	12	5.0
HCl (M)	0.4	0.1
Associated yield (%)	54	84
Temp. (ºC)	25	40
Time (h)	2	72

From a brief optimisation, we discovered that by decreasing the equivalents by >50%, we could achieve a greater yield (84% vs 54%), although this required a longer reaction time (72 vs 2 h) and an increase in temperature (40 vs 25 ºC). These optimised conditions resulted in a slower process, but had the advantage of requiring a smaller quantity of a toxic reagent that could implemented in subsequent reactions.

Upon the multi-gram synthesis of starting material **100**, the next step was to synthesise the benzimidazole anilines *via* one of two routes.

6.2.2 Method A – *via* alkylation

Within the group, a three-step route to the substituted amino benzimidazoles was under investigation, **Scheme 18**.

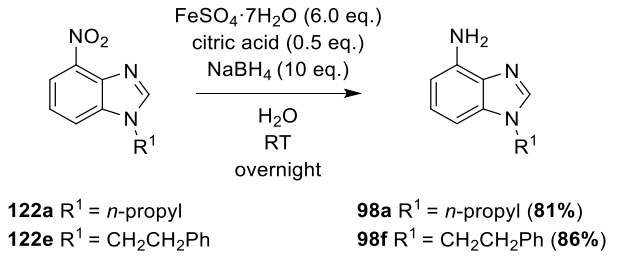


*Scheme 18: Alkylation strategy to anilines **98a–e** and **124a/b**. ([a]taken forward as a mixture)*

Di-amine **100** was treated with triethylorthoformate under acidic conditions, and after 1 hour at reflux, benzimidazole **121** was accessed in near quantitative yields (96%). The second step was the treatment of **121** with a range of alkylating agents under basic conditions to provide, upon separation of regioisomers, a series of alkylated nitrobenzimidazoles **122a**–**e** and **123a**–**e** in moderate yields (18–59%). Assignment of the regiochemistry was determined by 2D NOESY NMR experiments (see **Appendix 10.5**). The final step in the synthetic sequence was the treatment of nitro species (**121**, **122a**–**e** and **123a**–**e**) with 10% palladium on carbon under a hydrogen atmosphere. This generated the aniline derivatives **98a**–**e, 124a** and **124b** in good to quantitative yields (65%–quant.). Upon biological evaluation of the *N*-1 substituted final products, we soon realised that this isomer resulted in inactive inhibitors (**Section 3.3.2**). Therefore, once separated, isomers **123c**–**e** were not reduced.

With respect to the nitro reduction step, in some cases, alternative by-products were observed. First of all, the reduction of *N*-1 benzyl derivative **123c** resulted in full recovery of the de-benzylated product **98e**. This was not surprising as hydrogenation

conditions are also used in the cleavage of benzyl groups.[156] With close monitoring of the reaction *via* TLC analysis it was possible to isolate aniline **97c** (83%) before benzyl cleavage had taken place. Second, the reduction of the *n*-butyl derivative **122d** resulted in a 9:1 mixture of the desired alkylated aniline **98d** to de-alkylated product **98e**. These compounds were inseparable by column chromatography and therefore, the procedure was altered with a reduced catalyst loading of 5 mol%. The reaction was monitored at 15 minute intervals *via* TLC analysis, and within 30 mins no starting material and only one product was observed. However, upon analysis by ^{1}H NMR spectroscopy a 9:1 mixture of products was still observed. We therefore attempted the reduction using milder conditions, **Scheme 19**.

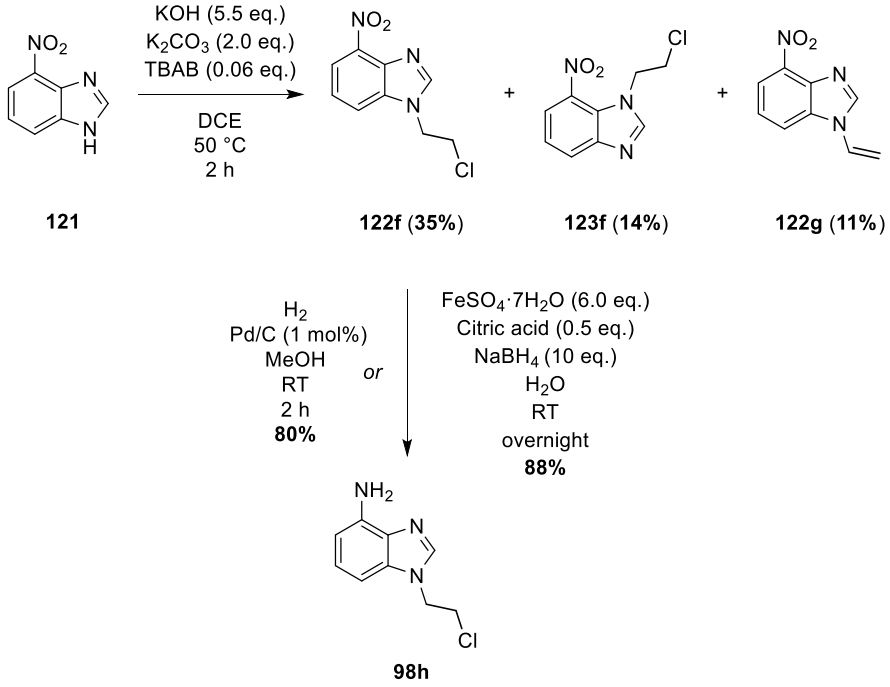


Scheme 19: Investigation of alternative reduction conditions.[157]

Dey and co-workers described a reduction method using an iron salt, **Scheme 19**.[157] They reported a number of examples where reduction of the nitro group occurred preferentially in the presence of other sensitive functionalities, such as benzyl groups and alkenes. The reported conditions were 3.0 eq. of $FeSO_4$, 0.3 eq. citric acid and 5.0 eq. of $NaBH_4$, however, when subjected to these conditions, only 50% conversion to products **98a** and **98f** (confirmed by ^{1}H NMR spectroscopy) was observed, after stirring at room temperature overnight. Therefore, we decided to implement an increase in equivalents of the reagents in order to drive the reaction to completion, **Scheme 19**. Upon basic work up, this mild alternative reduction method provided anilines **98a** and **98f** in excellent yields (81 and 86% respectively) without further

purification. This method was then used as the first choice nitro reduction method for subsequent intermediates.

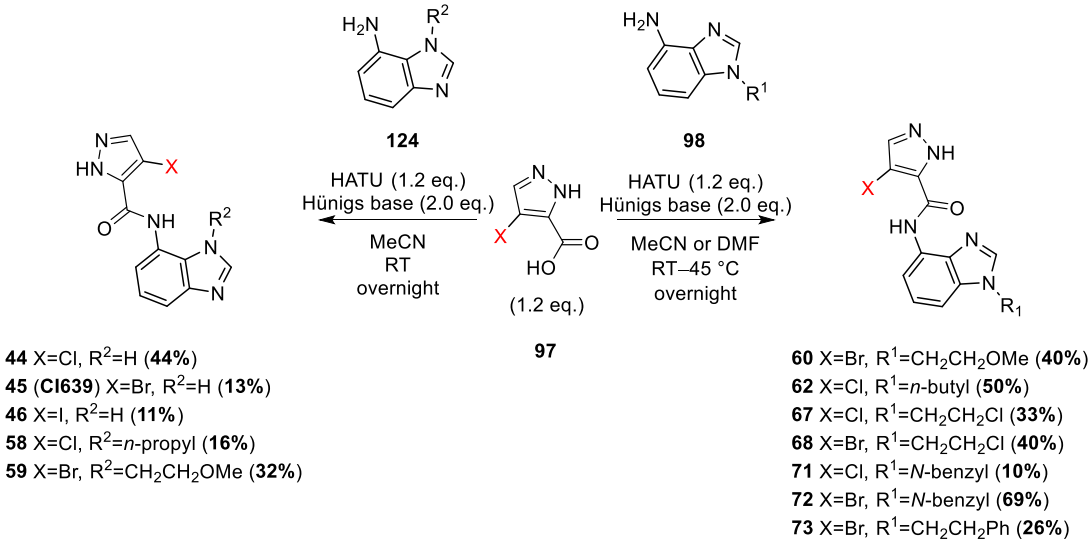
A similar approach was taken to synthesise aniline **98h** from benzimidazole intermediate **121**, **Scheme 20**.



*Scheme 20: Synthesis of amino benzimidazole **98h**.*

Compound **121** was treated with tetrabutylammoniumbromide (TBAB) under basic conditions at 50 ºC in dichloroethane to generate the alkylating agent *in situ*. This produced a mixture of products which were separated *via* column chromatography including the desired product **122f**, the *N*-1 substituted isomer **123f** and the elimination product **122g** in moderate yields (11–35%). Elucidation of the structures of **122f**, **123f** and **122g** was confirmed by ^{1}H NMR spectroscopy and 2D NMR experiments (see **Appendix 10.5**). Identification of the structure of **122f** was confirmed by the interaction of the aliphatic chain and the aromatic CHs of the benzimidazole motif through 2D NOESY NMR experiments. The desired aniline **98h** could be accessed *via* two reduction methods. Nitro **122f** was subjected to palladium hydrogenation conditions for 2 hours to provide aniline **98h** in good yields (80%). Alternatively, this intermediate could also be accessed *via* the milder reduction method with iron sulfate and sodium borohydride in aqueous media. Upon basic work up, this generated aniline **98h** in excellent yields without the need for further purification (88%).

The final step in the process to synthesising the inhibitors was an amide coupling reaction with the corresponding acid partner, **Scheme 21**.



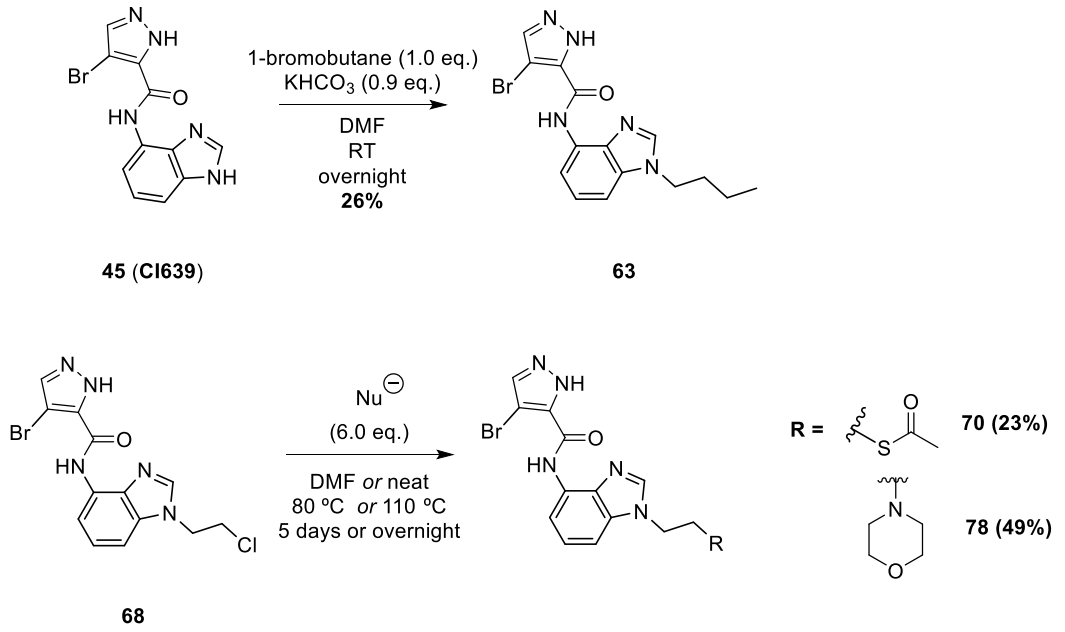
Scheme 21: Amide coupling process to access inhibitors.

Acid **97** was reacted with anilines **124** or **98** and HATU under basic conditions at ambient temperatures to generate the desired amide products **60**, **62**, **67**, **68** and **71–73** in a wide range of yields (10–69%). This range of yields are a result of the amide coupling process most often not proceeding to completion and thus the mass balance can be associated with recovery of starting material anilines. These amides were then assessed *in vitro* and their DYRK2 activity was reported and rationalised, as described in **Section 3.3.4**.

A beneficial approach to compound design is the installation of functional handles for late stage diversification. This can allow medicinal chemists to synthesise an intermediate in bulk and then streamline the synthesis of a wide range of derivatives to build a robust SAR profile of inhibitors.

6.2.3 Late stage diversification

We proposed that derivatives **63**, **70** and **78** could be accessed through late stage diversification, **Scheme 22**.



*Scheme 22: Synthesis of **63**, **70** and **78** via late stage diversification.*

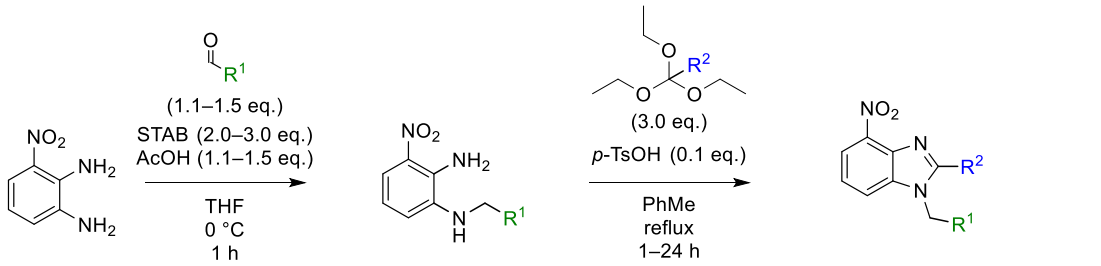
Benzimidazole **45** (**CI639**) was treated with 1-bromobutane and sub-stoichiometric amounts of the weak base potassium hydrogen carbonate ($KHCO_3$) to avoid deprotonation of the pyrazole and subsequent reaction with the alkylating agent. After stirring at room temperature overnight and upon purification *via* column chromatography, product **63** was isolated in poor yield (26%), **Scheme 22**. However, due to the potential competition between nucleophiles and the low yield of the previous step to generate the starting material **45**, this approach wasn't pursued further.

Benzimidazole **68** contained an electrophilic handle, which we proposed to use to introduce a HBD motif. Therefore, **68** was reacted with potassium thioacetate for 5 days at 80 ºC. Upon purification *via* column chromatography, **70** was isolated in poor yield (23%), **Scheme 22**. Notably, **70** was moderately soluble in MeOH and due to the small scale of the reaction (10 mg of **70** was recovered), we were therefore unable to subject it to further hydrolysis. In addition, in order to investigate the limit to the substitution of the *N*-3 position of the benzimidazole, we proposed the synthesis of the ionisable morpholine analogue **78**. Chloride **68** was reacted with morpholine at reflux overnight. Upon purification *via* column chromatography **78** was isolated in moderate yield (49%). Upon testing these inhibitors *in vitro*, we identified a limit to the substitution on the benzimidazole. More specifically, carbon chains of >3 resulted in less active inhibitors (see **Section 3.3.4**). Therefore, this route was not pursued any further.

6.2.4 Method B – *via* reductive amination

The previous approach to install substitution on the benzimidazole ring resulted in the generation of *N*-1 alkylated by-products that required removal by purification. We therefore explored an alternative route to circumvent this problem. Moreover, we devised a strategy which could also allow us to install functionality at the C-2 position of the benzimidazole, **Scheme 23**.

6.2.4.1 Reductive amination route with aldehydes



100

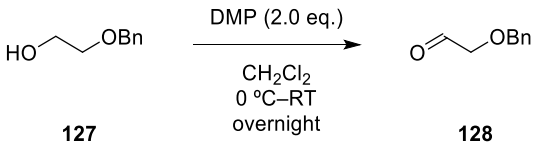
125a R^1=*iso*-Pr (**50%**)
125b R^1=C_6H_{11} (**66%**)
125c R^1=CH_2CF_3 (**79%**)
125d R^1=C_3H_5 (**55%**)
125e R^1=CH_2OBn (**32%**)[a]
125f R^1=CH_2CH_3 (**54%**)
125g R^1=CH_3 (**70%**)
125h R^1=3-furyl (**31%**)[b]

126a R^1=*iso*-Pr, R^2=H (**95%**)
126b R^1=C_6H_{11}, R^2=H (**70%**)
126c R^1=CH_2CF_3, R^2=H (**87%**)
126d R^1=C_3H_5, R^2=H (**77%**)
126e R^1=CH_2OBn, R^2=H (**80%**)
126f R^1=CH_2CH_3, R^2=methyl (**78%**)
126g R^1=CH_2CH_3, R^2=ethyl (**77%**)
126h R^1=CH_3, R^2=H (**55%**)
126i R^1=3-furyl, R^2=H (**80%**)

*Scheme 23: Nitrobenzimidazoles **126a–i** via reductive amination ([a]aldehyde synthesised and used immediately and [b]$NaBH_4$ used as reducing agent).*

Erion *et al*., reported reductive amination conditions to install an *iso*-butyl substituted amine on a similar scaffold to nitro di-amine **100**.[158] However, their solvent of choice was dichloroethane (DCE) in which substrate **100** was found to be poorly soluble and the reaction did not proceed. Pleasingly, without a solvent screen, we discovered that the reaction proceeded in THF, **Scheme 23**.

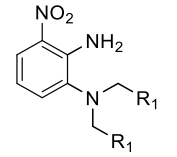
The nitro amine **100** was treated with a range of commercially available aldehydes and the mild reducing agent sodium(triacetoxy)borohydride (STAB) at 0 ºC to generate a range of substituted anilines **125a**–**h** in moderate to good yields (31–79%). The aldehyde required to generate the *O*-benzyl derivative **125e** was prepared using the method reported by Schoenberger and Trauner for an alternative alcohol,[159] **Scheme 24**.



*Scheme 24: Formation of aldehyde **128**.*

Commercially available alcohol **127** was treated with oxidising agent Dess-Martin Periodinane (DMP) at 0 ºC in dichloromethane overnight. Upon aqueous work-up, aldehyde **128** was generated and its structure was confirmed by ^{1}H NMR spectroscopy. Compound **128** was used immediately in the next step without further purification. The regioselectivity of the products of the reductive amination process (**125a**–**h**, **Scheme 23**) was confirmed by 2D NOESY NMR experiments (see **Appendix 10.5**).

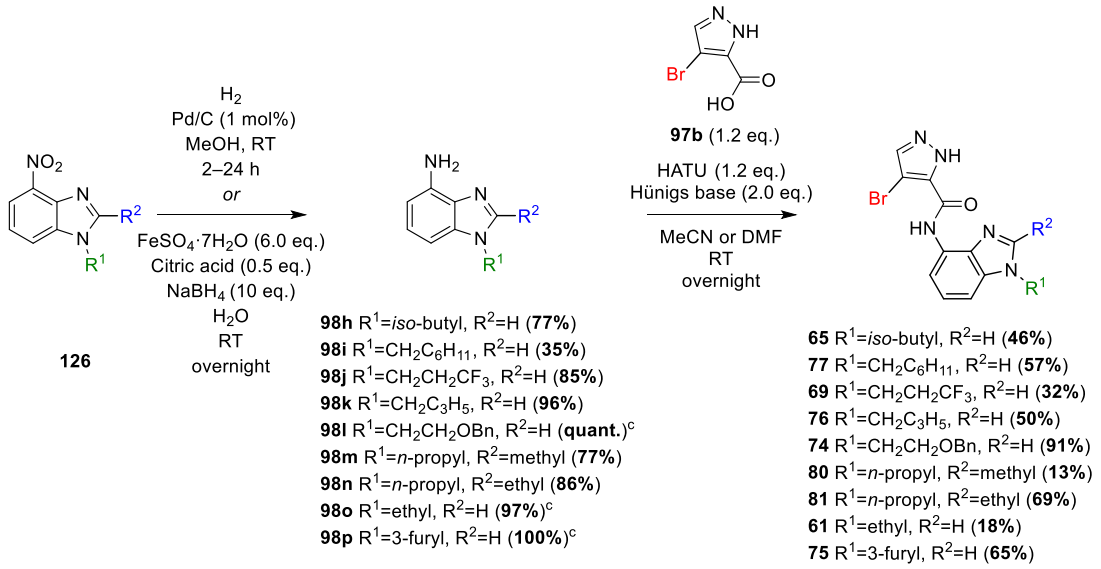
It was interesting to note that depending on the size of the R group, a reduction in equivalents of the reagents was required in order to avoid the formation of an undesired bis-alkylation product, **Figure 99**. More specifically, 1.5 eq. of propionaldehyde and isobutyraldehyde resulted in the formation of a significant amount of **125i** and **125j** respectively.



125i R_1 = CH_2CH_3
125j R_1 = *i*-Pr

*Figure 99: Bis-alkylation product **125i** and **125j** from reductive amination.*

The next step, was the cyclisation reaction with the newly synthesised anilines **125a**–**h** and the subsequent orthoformate species under acidic conditions, **Scheme 23**. Upon further purification, a range of substituted benzimidazoles **126a**–**i** were generated in good to excellent yields (55–95%). Moreover, the treatment of aniline **125f** with triethylorthoacetate or the orthopropionate species gave rise to the C-2 substituted benzimidazoles **126f** and **126g** in excellent yields (77% and 78%). The final steps in the synthesis were the reduction to the corresponding anilines and the amide coupling process with acid **97b**, **Scheme 25**.

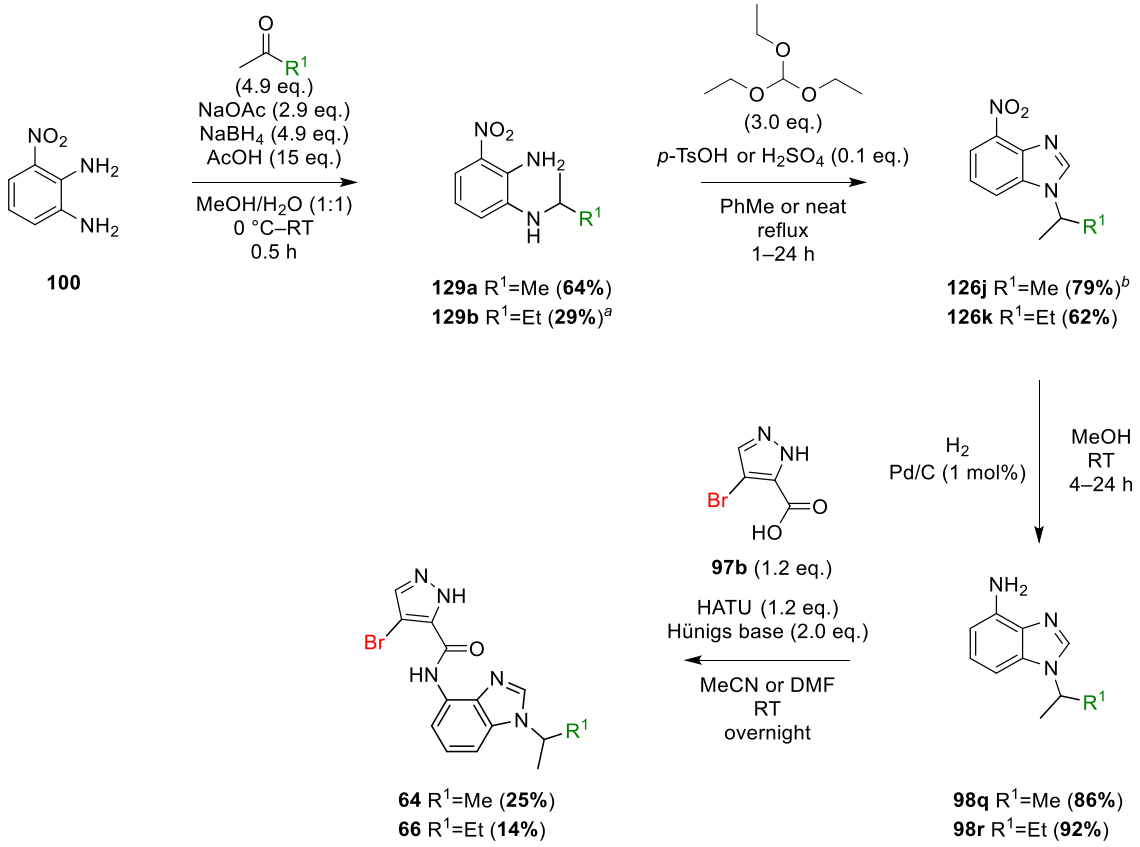


Scheme 25: Formation of inhibitors ([c] Pd method).

Nitro benzimidazoles **126** were subjected to reduction conditions to generate the anilines (**98h–p**) in excellent to quantitative yields (77%–quant.), **Scheme 25**. Some derivatives required the stronger palladium catalysed hydrogenation conditions in order to generate the aniline, including the 3-furyl derivative **98p**. For the *O*-benzyl derivative **98l**, we envisaged a simultaneous nitro reduction and benzyl de-protection to expose a HBD motif at the C-3 position on the carbon chain, however, this one-pot process proved challenging and requires further investigation. More specifically, the reaction stops at the reduction of the nitro and the benzyl group remains intact. A possible explanation for this could be that the aniline poisons the palladium catalyst, stalling the reaction at the aniline product. The full conversion of the furyl derivative to the corresponding amine product was not achievable *via* the mild iron reduction method. Therefore, nitro precursor **126i** was treated with H_2 and palladium on carbon at room temperature for 4 hours to generate the desired aniline **98p** in quantitative yields, **Scheme 25**. Finally, in order to access the final compounds for biological evaluation, anilines **98h–p** were reacted with pyrazole acid **97b** and amide coupling reagent HATU under basic conditions at ambient temperatures to provide the desired amides in reasonable to excellent yields (13–91%). Similar to previous comments, the reaction with acid **97b** and anilines **98h–p** does not always proceed to completion and therefore resulted in a wide range of yields being observed.

In addition to the branched *iso*-butyl derivative **65**, branched derivatives **64** and **66** were synthesised *via* a reductive amination process with the corresponding ketones, **Scheme 26**.

6.2.4.2 Reductive amination with ketones

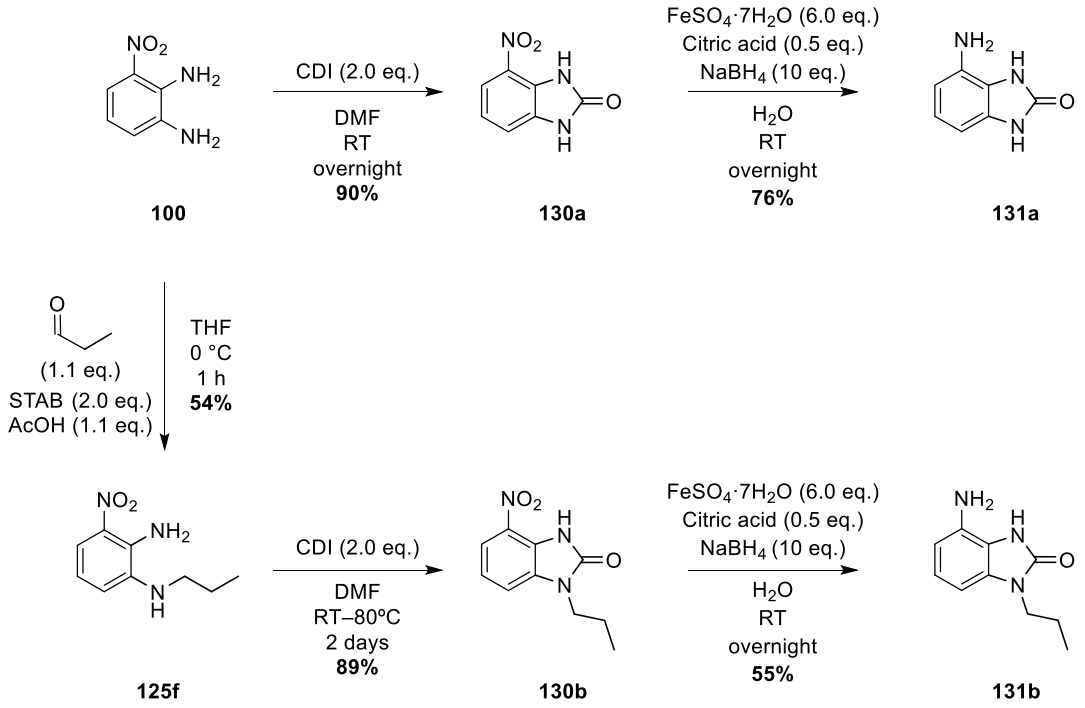


Scheme 26: Reductive amination route with ketones ([a]2× reagents, [b]neat reaction).

Compound **100** was reacted with the relevant ketone and an excess of sodium borohydride in a mixture of MeOH and H_2O. After 30 mins at room temperature, and upon purification *via* column chromatography, this provided substituted anilines **129a** and **129b** in moderate yields (29–64%). Next, cyclisation with triethylorthoformate generated the corresponding benzimidazoles **126j** and **126k** in good yields (62–79%). The *iso*-propyl analogue **126j** required neat conditions at reflux (130 ºC) with a catalytic amount of H_2SO_4 to facilitate the reaction. We observed that standard conditions involving the solvent toluene and a milder acid (*p*-TsOH), resulted in an incomplete reaction. In the penultimate step, treatment with palladium on carbon under a H_2 atmosphere provided desired anilines **98q** and **98r** in excellent yields without the need for further purification (86–92%). Finally, these anilines were reacted with pyrazole acid **97b** and HATU at room temperature and upon purification *via* column chromatography, provided the desired target inhibitors **64** and **66**, but in poor yields (14% and 25% respectively).

6.2.4.3 Further C-2 substitution

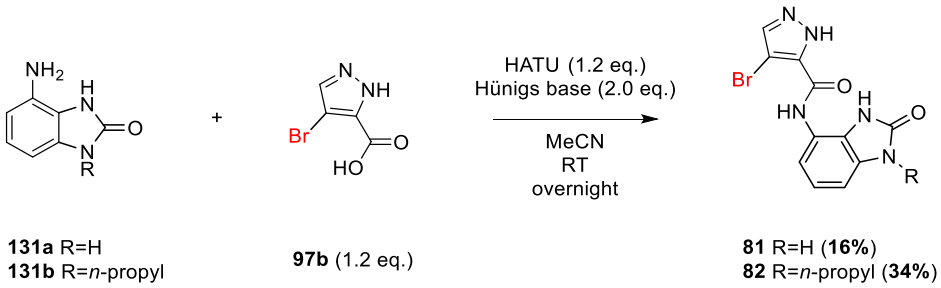
As described in **Section 3.3.4**, we hypothesised that the introduction of a HBD could allow interaction with hydrophilic residues within the DYRK2 active site, which could lead to an increase in the potency of our inhibitors. In order to synthesise these derivatives, we proposed a route using 1,1-carbonyldiimidazole (CDI), **Scheme 27**.



*Scheme 27: Route to urea-type amines **131a** and **131b**.*

Common intermediate **100** was treated with CDI in DMF at room temperature overnight, to generate the corresponding benzimidazol-2-one **130a** in excellent yield after purification through a basic work-up (90%). Next, benzimidazol-2-one **130a** was reacted with the iron salt cocktail of reagents in aqueous media to give **131a** in good yield (76%). Aniline **131a** was accessed in two steps with an overall yield of 68%.

With respect to the *n*-propyl analogue **131b**, di-amine **100** was reacted with propionaldehyde and STAB in THF at 0 ºC for 1 hour to provide substituted amine **125f** in reasonable yield (54%). Treatment with CDI at 80 ºC for 2 days, provided the corresponding *n*-propyl benzimidazol-2-one analogue **130b** in excellent yield (89%). Finally, nitro derivative **130b** was subjected to reductive conditions at room temperature overnight to access the desired aniline **131b** in moderate yield (55%). Aniline **131b** was synthesised with an overall yield of 26% over 3 steps. The final step in the synthesis of our target inhibitors was the amide coupling reaction with HATU, **Scheme 28**.



131a R=H
131b R=*n*-propyl

97b (1.2 eq.)

81 R=H (**16%**)
82 R=*n*-propyl (**34%**)

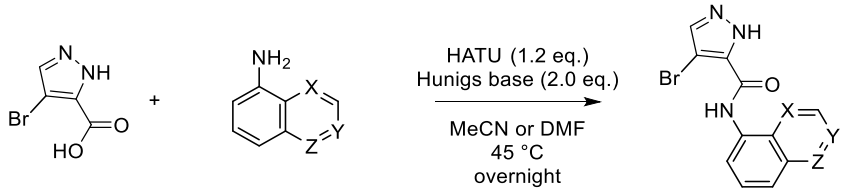
*Scheme 28: Amide coupling to access **81** and **82**.*

Anilines **131a** and **131b** were treated with pyrazole acid **97b** and HATU at ambient temperatures overnight in MeCN. Upon purification *via* column chromatography or, with respect to the unsubstituted analogue **81**, *via* trituration in H_2O, the desired amides **81** and **82** were accessed in reasonable yields (16–34%).

After extensively exploring the substitution on the benzimidazole tail group, we next investigated the introduction of different heterocycles in this region of the inhibitor.

6.2.5 Alternative heterocycles

A range of amine containing 6,6- fused heterocycles was evaluated *in vitro*, most of which were synthesised *via* the direct amide coupling with the pyrazole acid **97b** and the commercially available anilines **132a**, **132b** and **132d** (**Scheme 29**). However, the quinoxaline aniline **132c** required synthesis from the common intermediate nitro di-amine **100**, **Scheme 30**, using the method reported by Zhou *et al*.[160]



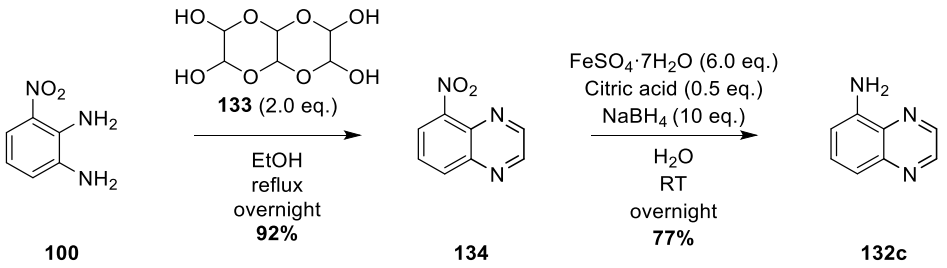
97b (1.2 eq.)

132a X=CH, Y=CH, Z=N
132b X=CH, Y=N, Z=CH
132c X=N, Y=CH, Z=N
132d X=N, Y=CH, Z=CH

54 X=CH, Y=CH, Z=N (**22%**)
55 X=CH, Y=N, Z=CH (**4%**)
56 X=N, Y=CH, Z=N (**21%**)
57 X=N, Y=CH, Z=CH (**61%**)

*Scheme 29: Formation of 6,6- fused heterocycle derivatives **54–57**.*

Acid **97b** was reacted with amines **132 a–d** and HATU in either MeCN or DMF at 45 °C overnight. Upon purification by column chromatography, amides **54–57** were accessed in poor to moderate yields (4–61%). The low yield (4%) for analogue **55** could be a result of the poor nucleophilicity of the isoquinoline starting material.



*Scheme 30: Synthesis of amino quinoxaline **132c**.*[160]

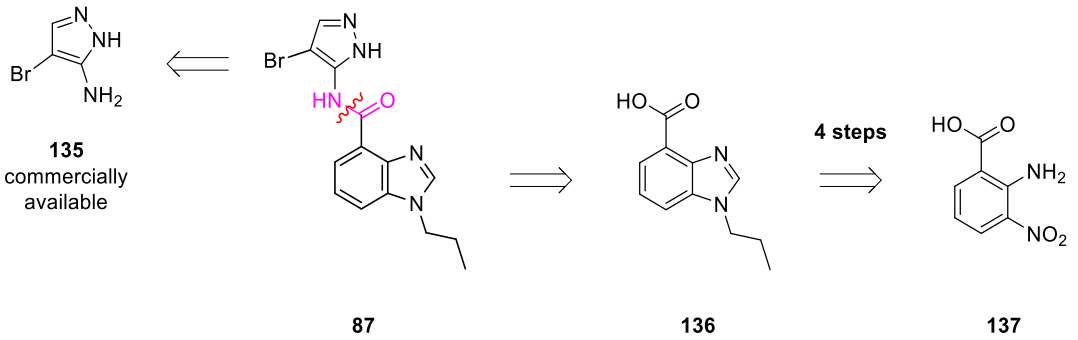
To generate the amino quinoxaline monomer **132c**, diamine **100** was reacted with glyoxal trimer dihydrate **133** in EtOH at reflux overnight. Upon purification *via* column chromatography, this generated nitro quinoxaline **134** in excellent yield (92%). Next, the nitro group was subjected to the mild reducing conditions with iron sulfate at room temperature overnight. Upon basic work-up, aniline **132c** was isolated in good yield (77%). Finally, the last step to access the inhibitor molecules was the amide coupling reaction with acid **97b**, **Scheme 29**.

A synthesis of alternative heterocyclic structures was undertaken through a final amide coupling step from common intermediate nitro di-amine **100**. This gave rise to the synthesis of a range of derivatives to investigate the tolerance of functionality at 3 positions. More specifically, the head group heterocycle, the substitution on the head group and the substitution on the benzimidazole. The final area for investigation within this study was the amide linker.

6.3 Amide linker

The first approach to modification of the amide linker was to synthesise the reverse amide analogue **87** of a potent inhibitor. For ease of synthesis, and to complement the large body of results, we set out to synthesise the **CI709** (**41**) analogue **87**, **Scheme 31**.

6.3.1 Reverse amide

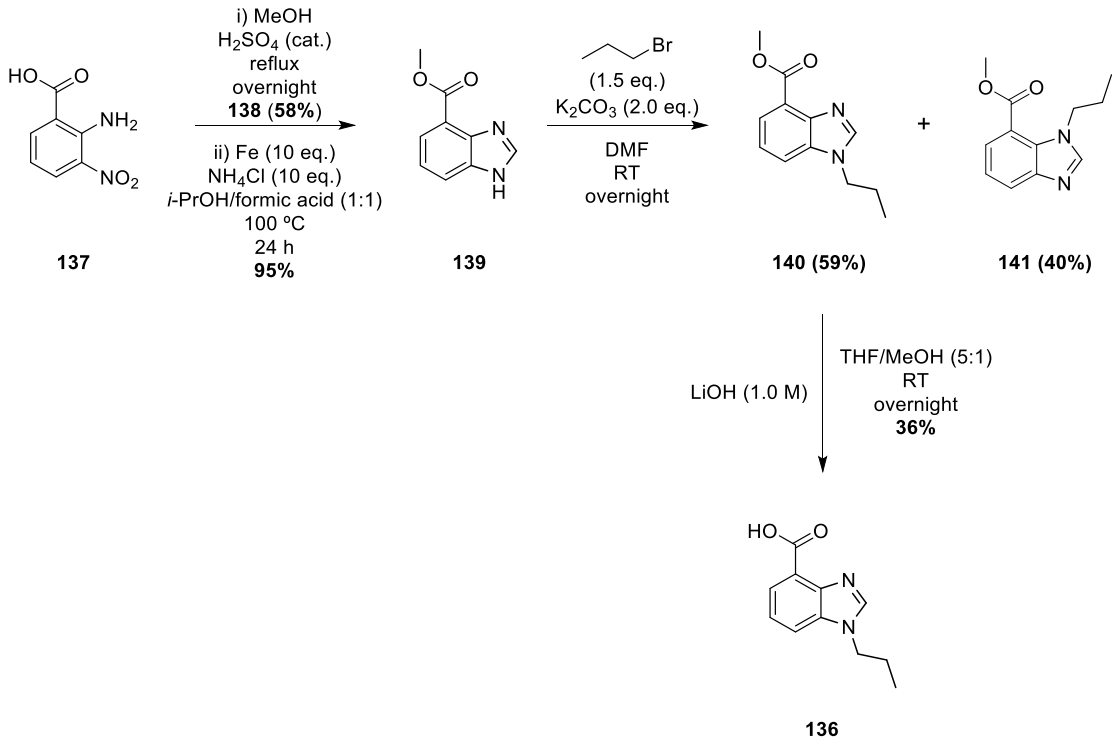


*Scheme 31: Retrosynthetic analysis of reverse amide **87**.*

From a retrosynthetic perspective, the final step in the synthesis would be an amide coupling with relatively inexpensive commercially available amino pyrazole **135** and benzimidazole acid **136**. Acid **136** can be accessed in 4 steps from commercially available benzoic acid **137** with the use of a similar strategy to that adopted in the previous sections.

6.3.1.1 Acid synthesis

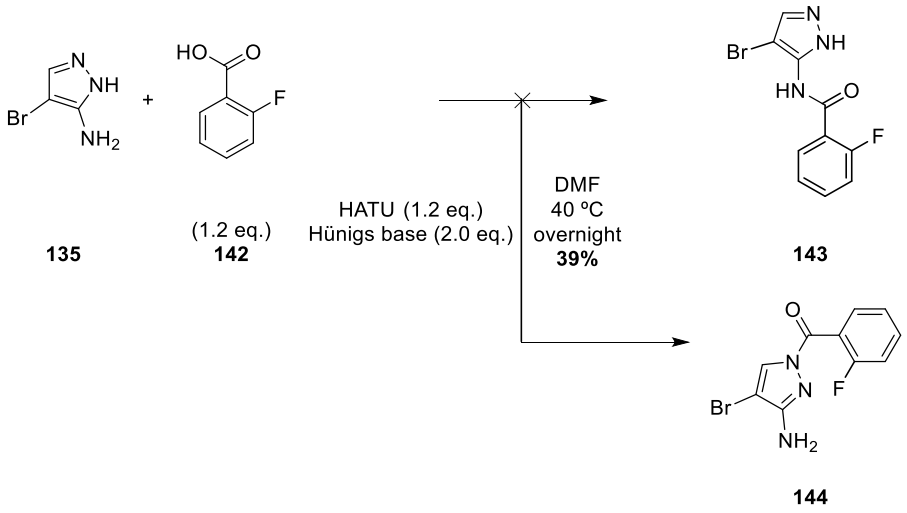
Acid **136** can be synthesised in 4 steps from commercially available benzoic acid **137**, **Scheme 32**.



*Scheme 32: Synthesis of acid **136**.*

Benzoic acid **137** was treated with an excess of MeOH under acidic conditions and after stirring at reflux overnight, the methyl ester intermediate **138** was generated in moderate yield (58%). Hanan *et al*. had reported a mild one-pot reduction/cyclisation procedure to gain access to a range of substituted benzimidazoles.[161] The methyl ester intermediate **138** was treated with iron, ammonium chloride and formic acid in isopropanol and heated to 100 ºC for 24 hours. Upon filtration and basic work-up, benzimidazole **139** was isolated in excellent yield (95%). Benzimidazole **139** was then reacted with alkylating agent *n*-propyl bromide under basic conditions at room temperature overnight to provide regioisomeric **140** and **141** in reasonable yield (59% and 40% respectively). Elucidation of connectivity was determined by 2D NOESY NMR experiments (see **Appendix 10.5**). Finally, in order to access the acid for the final coupling procedure, benzimidazole **140** was subjected to basic hydrolysis conditions in a THF/MeOH mixture at room temperature overnight. Upon pH alteration and aqueous extraction, acid **136** was isolated in moderate yield (36%). The low yield in this case can be attributed to the acidity of the species. The pKa of the acid product is approximately 2.0 whereas the pKa of the benzimidazole amines are roughly 5.0 and 12 (substituted amine). Therefore, in the pH range of 2–12 there will be a mixture of ionised products, resulting in them being primarily present in the aqueous phase. Acid **136** was accessed in 4 steps with an overall yield of 12%.

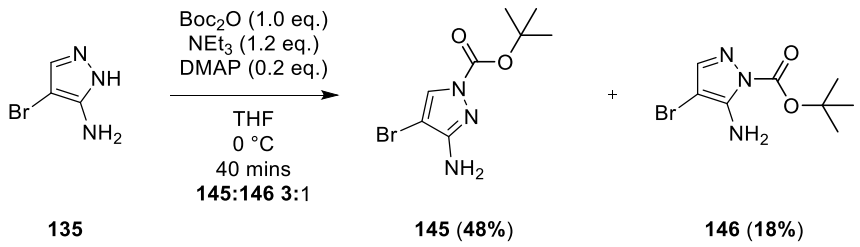
Upon successful synthesis of the acid **136**, we proposed a direct amide coupling with commercially available amino pyrazole **135** to furnish the desired reverse amide product **87**. Previously, we have observed that the pyrazole NH is more reactive than the exo-cyclic amine on the pyrazole. Therefore, we chose initially to investigate the amide coupling with a sacrificial commercially available acid **142**, **Scheme 33**.



*Scheme 33: Formation of undesired amide coupling product **144**.*

Commercially available amino pyrazole **135** was reacted with commercially available benzoic acid **142** and HATU at 40 ºC overnight. Upon precipitation and trituration in H_2O, amide **144** was isolated in moderate yield (39%). This was primarily determined by the presence of a peak in the ^1H NMR spectrum at 5.90 ppm which we believed to be the NH_2 of (**144**). Similar to previous amide couplings, this reaction did not go to completion. This investigation reinforced the differences in the reactivity's of the pyrazole amines. In order to combat this, we proposed a protecting group strategy, **Scheme 34**.

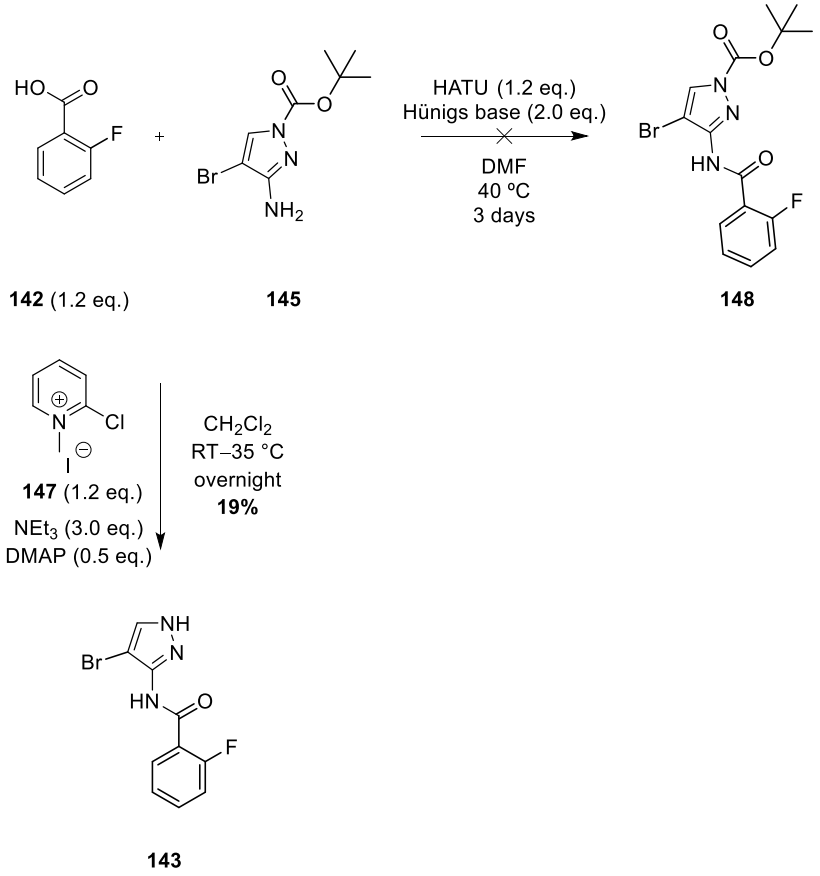
6.3.1.2 Amine synthesis



*Scheme 34: Boc protection of **135**.*

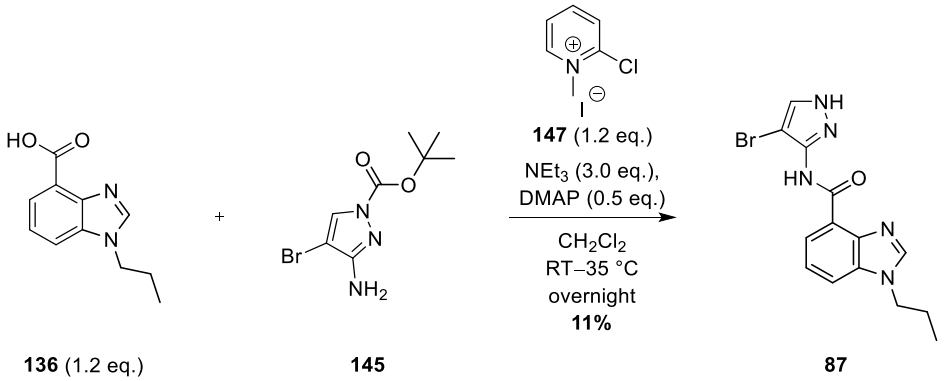
A common amine protecting group strategy is the formation of a *tert*-butyl carbamate (Boc group). Amino pyrazole **135** was treated with *tert*-butyl di-carbonate and a catalytic amount of *N*,*N*-dimethylaminopyridine (DMAP) under basic conditions at 0 ºC for 40 minutes. Upon further purification by column chromatography, regioisomers **145** and **146** were isolated in reasonable yield (48% and 18% respectively). Furthermore, elucidation of the regioisomeric identity of the products was confirmed by 2D NOESY NMR experiments (see **Appendix 10.5**). Identification of isomer **146** was rationalised by the NOESY relationship between the NH_2 and *t*-Bu methyl groups. The next step was then to examine the protected pyrazole in an amide coupling reaction. From a steric perspective, isomer **145** provided less steric hindrance around the reacting amine compared to the alternative regioisomer **146**. Therefore, **145** was taken forward and reacted with acid **136** to determine its reactivity, **Scheme 35**.

6.3.1.3 Amide coupling investigation



*Scheme 35: Amide coupling with protected pyrazole **145**.*

Unfortunately, reacting acid **142** and amine **145** under HATU amide coupling conditions at 40 ºC for 3 days to gain access to **148** was unsuccessful. More specifically, upon aqueous work up ^{1}H NMR spectroscopy indicated a mixture of starting materials. We hypothesised that the active intermediate formed between acid **142** and HATU was too sterically hindered for the approach of the (also bulky) amine **145** and would therefore require more energy to facilitate the reaction. Therefore, we investigated the reaction of the pyridinium salt **147**, acid **142** and amine **145**, in dichloromethane at 35 ºC overnight, **Scheme 35**. Upon acidic work-up and subsequent purification *via* flash column chromatography, the desired amide and pleasingly, the de-protected pyrazole product **143** was generated in reasonable yield (19%). It should be noted that upon the preparation of **143**, comparison of the spectra of **143** and **144** confirmed our hypothesis of regioselectivity. Additionally, **144** was insoluble in chloroform whereas **143** was soluble. This procedure was then followed for the formation of the reverse amide analogue **87**, **Scheme 36**.



*Scheme 36: Formation of desired reverse amide **87**.*

Carboxylic acid **136** was reacted with protected pyrazole **145** and pyridinium salt **147** at 35 ºC overnight. Upon acidic work-up, purification by column chromatography and trituration in hexane, de-protected amide **87** was accessed in low yield (11%).

Through the use of a tandem synthesis we could access the reverse amide inhibitor molecule **87** in 5 linear steps. Unfortunately, this analogue was inactive *in vitro* and the search for an inhibitor with increased potency was on-going. Therefore, with the help of a short screen of in-house kinase-like structures a similar scaffold with increased potency (**Section 3.4.2**) was identified, **Figure 100**.

6.3.2 Hybrid series

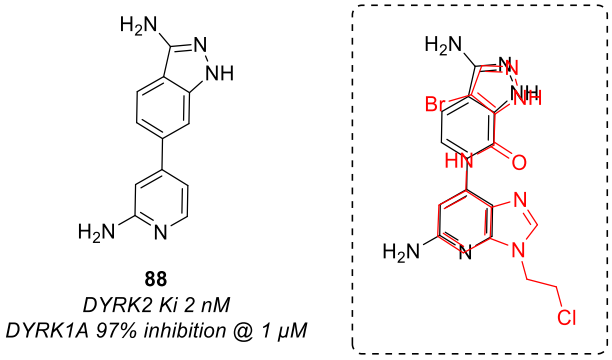
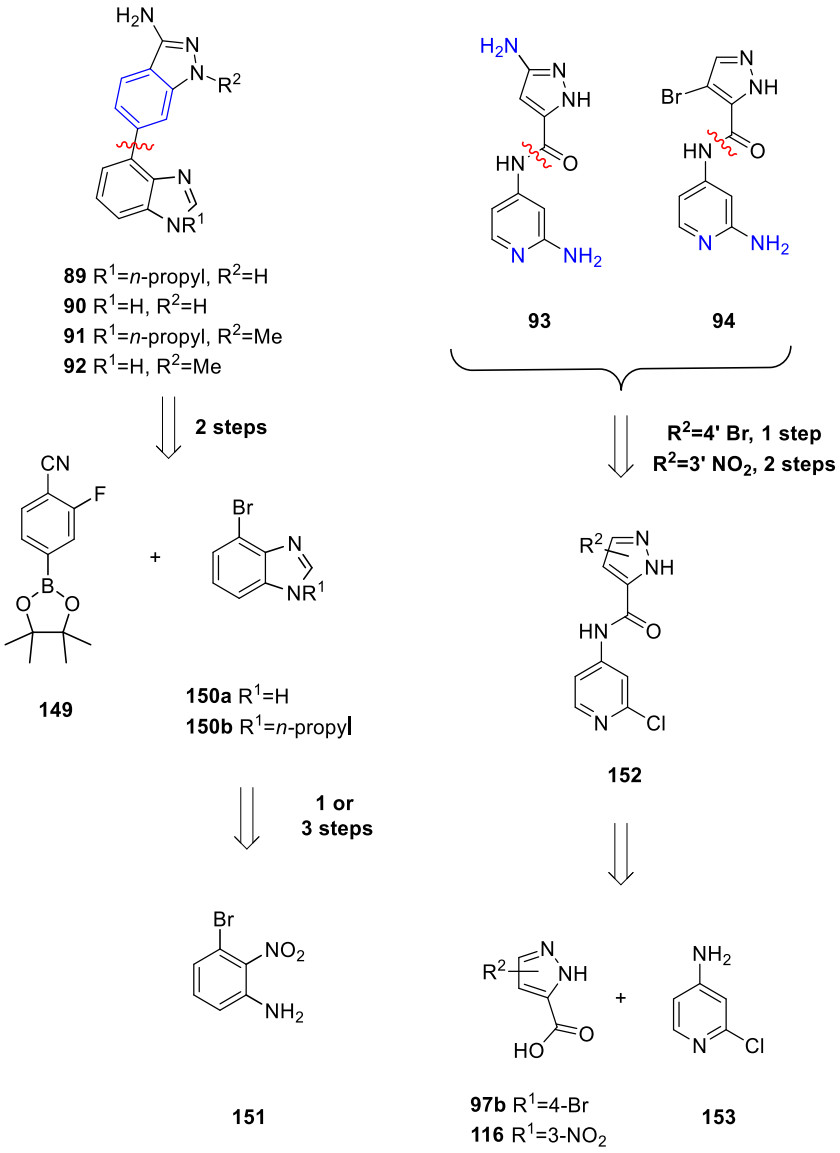


*Figure 100: Hit from screening **88** and overlay with **68** (LB35).*

If we overlay this screen hit **88** and the most potent inhibitor **68** there are some structural similarities. More specifically, the head group pyrazole of **68** is replaced with the indazole motif of **88**. Therefore, to investigate the specificities of the activity and lack of selectivity, we proposed the synthesis of a number of hybrid analogues.

More specifically, we set out to synthesise six hybrid derivatives **89**–**94**, **Scheme 37**.



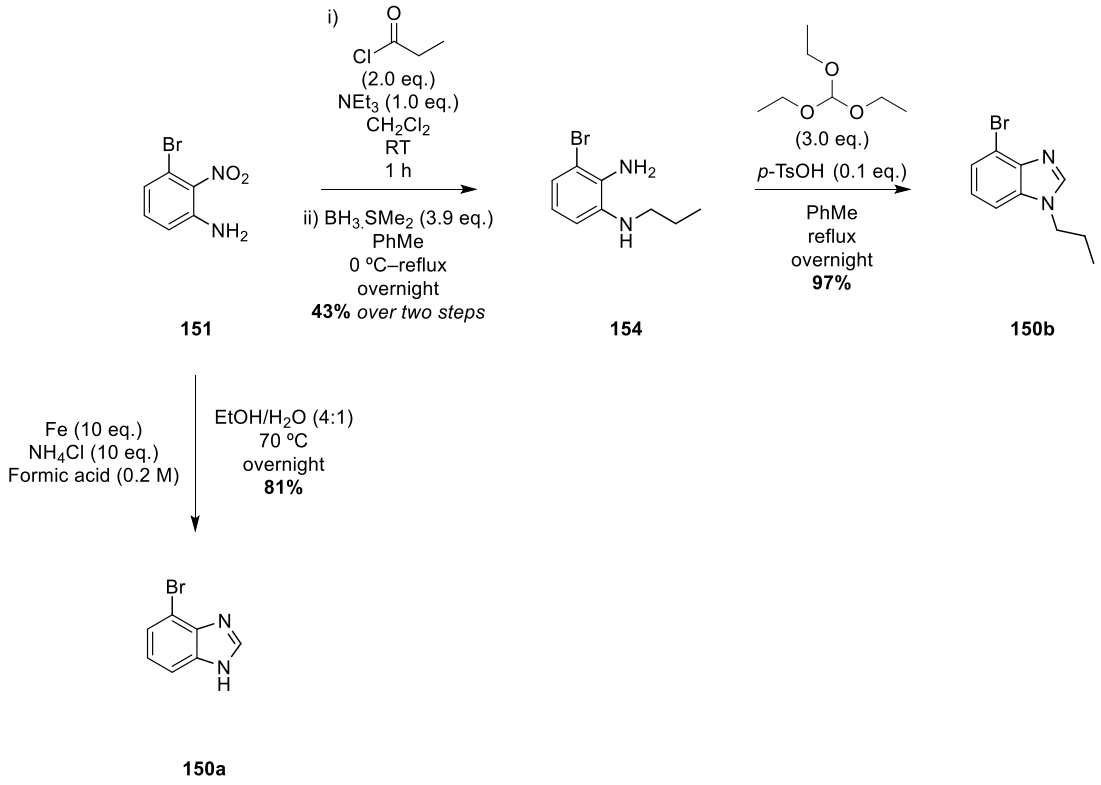
*Scheme 37: Retrosynthetic analysis of hybrid species **89**–**94**.*

We proposed indazole hybrid species **89**–**92,** where the amide linker was replaced with a phenyl ring and the amino functionality at the 3′ position on the indazole ring has been introduced. From a retrosynthetic perspective; indazoles **89**–**92** can be accessed in two steps; more specifically, *via* a transition metal catalysed cross-coupling reaction with the associated boronic ester **149** and bromides **150a** and **150b**, followed by the subsequent cyclisation reaction with the appropriate hydrazine reagent. The bromides **150a** and **150b** can be synthesised in a number of steps depending on their substitution pattern. The unsubstituted benzimidazole analogue **150a** can be accessed in one reaction *via* a one-pot reduction/cyclisation from commercially available aniline **151**. Whereas, the *n*-propyl derivative **150b** can be synthesised in three steps *via* an acylation, reduction and cyclisation process, from commercially available aniline **151**, **Scheme 37**.

The hybrid amino pyridine analogues **93** and **94** maintain the amide linker and contain the alternative pyridyl tail group. We proposed that these derivatives could be accessed *via* one or two steps from the chloride counterpart **152**. In the case of analogue **93,** a nitro reduction followed by an S_NAr reaction with an ammonium equivalent would afford the desired compound. In turn, chloride **152** could be synthesised from the amide coupling reaction of acids **97b** or **116** and amine **153**.

6.3.2.1 Amino indazole synthesis

The bromides for the penultimate Suzuki-Miyaura cross coupling reaction were synthesised *via* a 1 or 3 step sequence, **Scheme 38**.

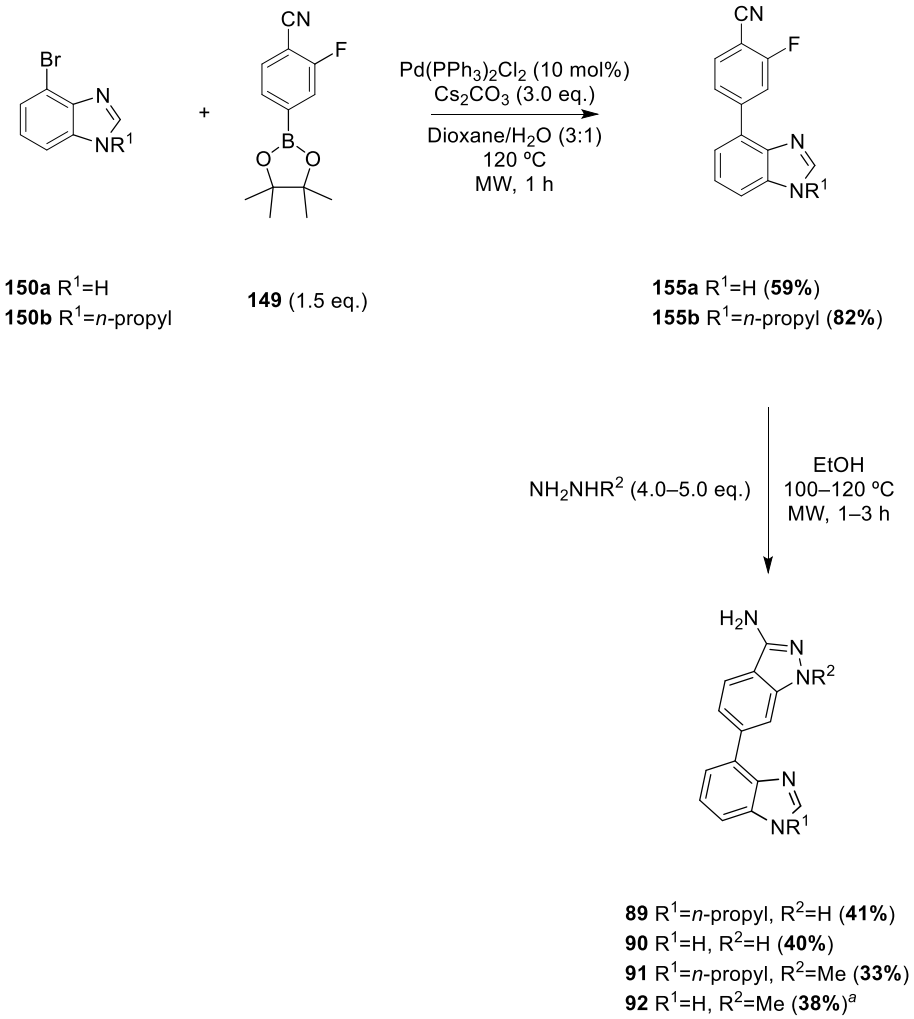


*Scheme 38: Formation of bromides **150a** and **150b**.*

To access the unsubstituted benzimidazole analogue **150a**, commercially available aniline **151** was treated with iron and formic acid in an EtOH/H_2O mixture at 70 ºC overnight. Upon basic work-up, bromide **150a** was isolated in excellent yield (81%). To gain access to the substituted analogue **150b**, compound **151** was treated with propionyl chloride under basic conditions at room temperature for 1 hour. Upon aqueous work-up, the crude material was reacted with borane dimethyl sulfide complex at 0 ºC. Increasing the temperature to reflux in toluene allowed for the global

amide and nitro reduction to furnish the aniline **154** in moderate yield over the two steps (43%). Finally, aniline **154** was treated with triethylorthoformate under acidic conditions in toluene at reflux overnight. After purification *via* column chromatography, bromo benzimidazole **150b** was isolated in excellent yield (97%).

The next step in the route to synthesise indazoles **89**–**92**, was the palladium catalysed cross coupling between the bromides **150a** and **150b** with commercially available pinacol ester **149**, **Scheme 39**.



*Scheme 39: Formation of indazoles **89**–**92** ([a]isolated as an inseparable mixture of isomers).*

The bromide **150a** or **150b** and BPin **149** were treated with palladium bistriphenylphosphine dichloride and subjected to microwave irradiation at 120 °C for 1 hour. Upon purification *via* column chromatography and further trituration in hexanes, intermediates **155a** and **155b** were isolated in moderate to excellent yields (59% and 82% respectively). Finally, the last step in the generation of the indazole hybrid analogues was the cyclisation reaction with aryl fluorides (**155a** and **155b**) and

the appropriate hydrazine reagent. As described in **Section 3.4.2**, in order to determine the importance of the indazole NH, we proposed the synthesis of methyl substituted analogues **91** and **92**. Therefore, the fluorides were treated with hydrazine hydrate or methyl hydrazine in EtOH and subjected to microwave irradiation at 100–120 ºC for 1–3 hours. Upon purification *via* flash column chromatography and trituration in hexanes, indazoles **89**–**92** were isolated in moderate yields (33–41%). The methyl indazole analogue **92** was isolated as an inseparable mixture of isomers (3:1 **92a**:**92b**), whereas the *n*-propyl derivative **91** was isolated as a single isomer, **Figure 101**.

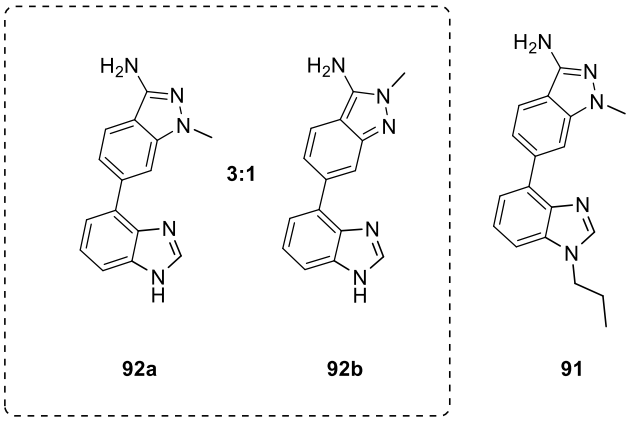


*Figure 101: Mixture of regioisomers **92a**:**92b** and single isomer **91**.*

Identification of the structure of the single isomer **91** was confirmed by 2D NMR experiments (see **Appendix 10.5**). The rationale behind the identification of **91** as the targeted isomer was the presence of a relationship between the N-CH_3 and the isolated aromatic CH of the indazole ring. From this, we inferred that the dominant isomer was **92a**. A possible rationalisation as to why we observe a mixture vs a single isomer is based on the 3D structures of the intermediates **155a** and **155b**, which were generated using Axon MarvinSketch software, **Figure 102**.

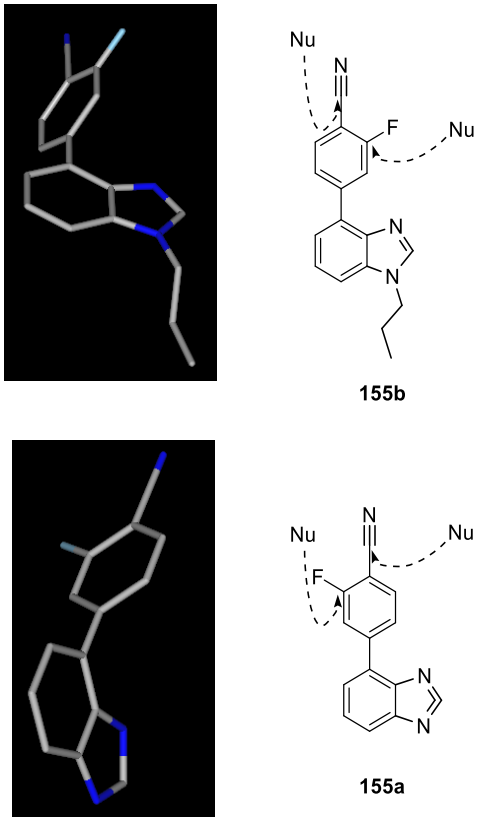


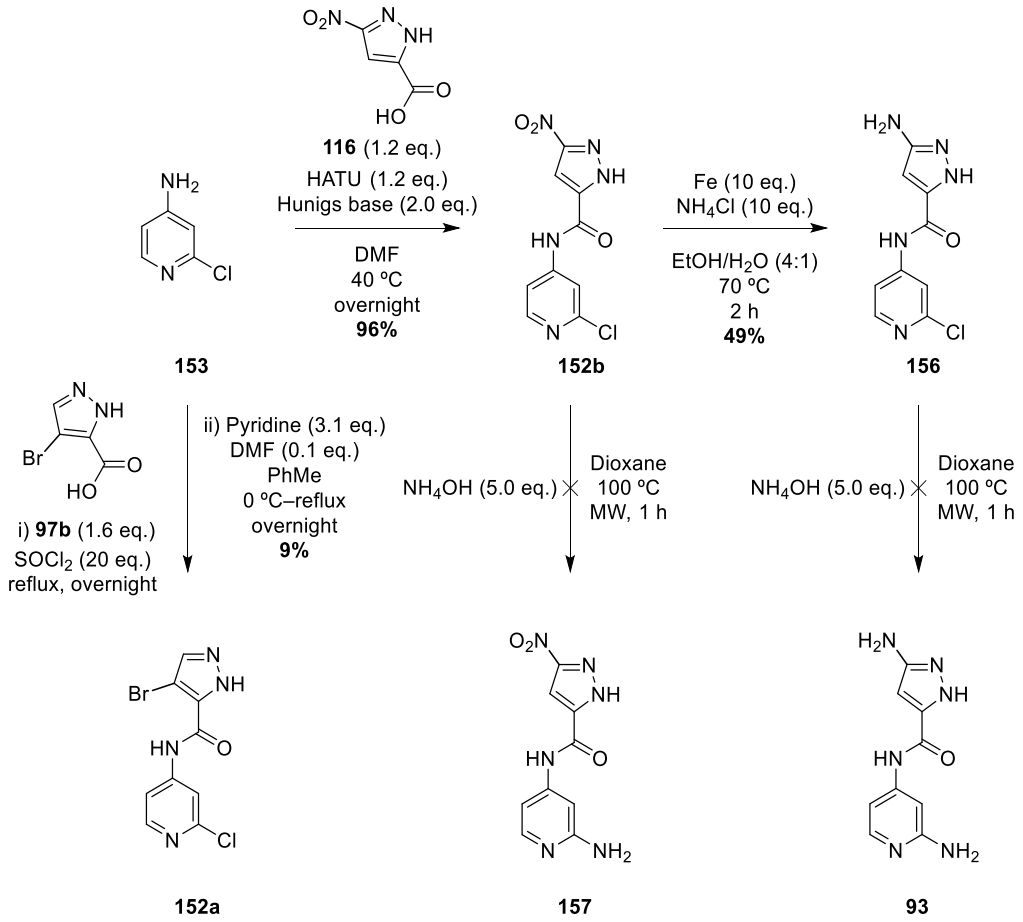
Figure 102: Possible reasoning for regioselectivity when cyclising with methyl hydrazine.

The hydrazine nucleophile can attack two electrophilic centres: the nitrile carbon centre or the fluorine substituted carbon, **Figure 102**. From investigation into the low energy 3D structures of **155a** and **155b**, the unsubstituted analogue **155a** adopts a different conformation to **155b Figure 102**. This provides less steric encumbrance from the benzimidazole. Therefore, compared to its *n*-propyl counter-part, the C-F bond bears less steric restriction which may contribute to the mixture of isomers observed, (3:1 **92a**:**92b**, see previous **Figure 101**). Assuming that the primary nucleophilic attack was *via* the methyl substituted NH of the hydrazine reagent, attack into the electrophilic C-F bond occurs first to provide the isolated single isomer **91**. Therefore, with respect to the intermediate **155a** the mixture of isomers, **92a** and **92b**, was a result of the competition between the electrophilic sites.

The next series of hybrid analogues were designed to introduce the amino pyridine tail.

6.3.2.2 Amino pyridine analogues

We proposed two and three step routes to access the amino pyridine derivatives **93** and **94** from commercially available 4-amino-2-chloro pyridine **153**, **Scheme 40**.



Scheme 40: Original strategy to amino pyridine derivatives.

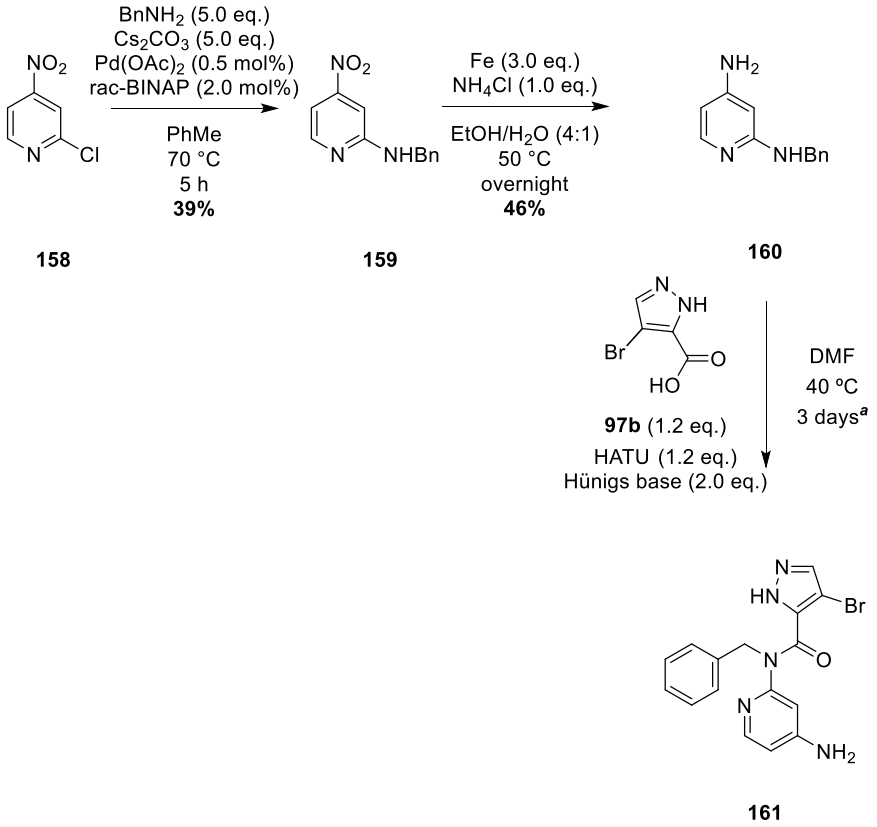
Chloro pyridine **153** was reacted with commercially available nitro acid **116** and HATU in DMF at 40 ºC overnight. Upon trituration in H_2O, amide **152b** was isolated in excellent yield (96%). Next, we investigated the reaction of amide **152b** with ammonium hydroxide under microwave irradiation at 100 ºC, **Scheme 40**. Unfortunately, this did not generate the desired amino pyridine product, **157**, but resulted in a basic cleavage of the amide bond, as we observed the presence of chloro pyridine **153** in the crude mixture, as confirmed by ^{1}H NMR spectroscopy. This reaction was then investigated at ambient temperature, but only starting material was observed after stirring overnight, which was determined *via* TLC analysis and ^{1}H NMR spectroscopy. Therefore, we synthesised the amino pyrazole analogue **156**. Nitro derivative **152b** was treated with iron and ammonium chloride in an EtOH/H_2O mixture at 70 ºC for 2 hours. Upon basic work-up and subsequent purification by column chromatography, the amino pyrazole analogue **156** was isolated in moderate yield (49%). This analogue was then subjected to the amination conditions, but unfortunately, after 1 hour of microwave irradiation at 100 ºC, only starting material **156** was observed *via* TLC analysis. Thus, the introduction of the 2-amino pyridine group and formation of **93** *via* a late stage and S_NAr approach proved to be more

challenging than envisaged. The electron withdrawing substituents on the pyrazole activate an alternative electrophilic centre at high temperatures and out compete the desired electrophilic site on the pyridine ring. Therefore, an alternative route to the 2-amino pyridines was designed.

Whilst investigating this route, the synthesis of the bromo containing hybrid **94** was also underway. However, the general amide coupling reaction between 4-amino-2-chloro pyridine **153** and HATU proved unsuccessful. In order to generate amide **152a**, acid **97b** was treated with thionyl chloride and stirred at reflux overnight to generate the acid chloride. The excess $SOCl_2$ was removed and the crude material was then treated with a catalytic amount of DMF, pyridine and amine **153** in toluene at reflux overnight. Upon aqueous work-up and recrystallisation in MeOH, amide **152a** was isolated in poor yield (9%). As the previous investigation highlighted the difficulty of the S_NAr reaction, this intermediate was not investigated further.

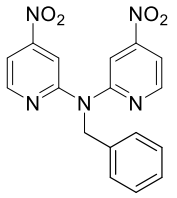
6.3.2.3 Alternative strategy to amino pyridines

We proposed an alternative strategy to synthesising the amino pyridine analogues from commercially available 4-nitro-2-chloro-pyridine **158**, **Scheme 41**.



Scheme 41: New strategy to the synthesis of 2-amino pyridines ([a]partially purified).

By using of another palladium catalysed cross coupling process (a Buchwald-Hartwig amination reaction), we could install a benzyl protected amine at the 2-position of the pyridine. 4-Nitro-2-chloropyridine **158** was treated with a catalytic amount of palladium acetate, a bis-phosphine ligand and benzyl amine in toluene for 5 hours at 70 ºC. This generated benzylamine **159** in reasonable yield (39%). However, this process also generated the di-addition by-product **162**, **Figure 103**.



162

*Figure 103: Di-addition product **162** from Buchwald process.*

The conditions presented in **Scheme 41**, are the optimised conditions which were developed to alter the ratio of mono **159** to di-addition product **162**.

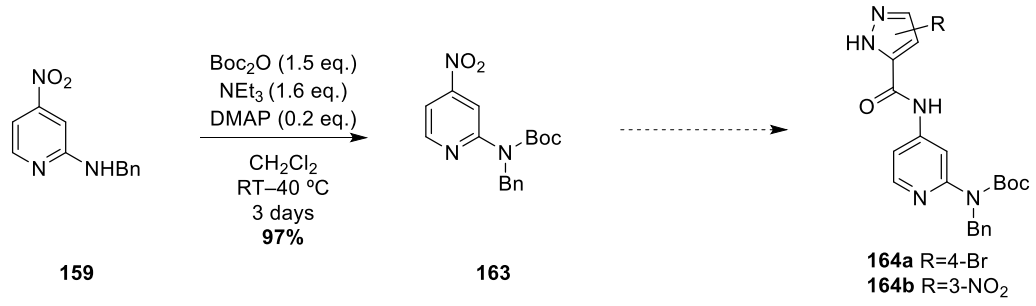
Table 20: Buchwald conditions investigation.

	Original Conditions	**Optimised Conditions**
$BnNH_2$ (eq.)	1.5	5.0
Cs_2CO_3 (eq.)	1.5	5.0
Molarity of PhMe	0.1	0.1
Temp. (ºC)	80	70
Time (h)	1	5
Ratio 159:162	1:2	3:1

The original conditions used a slight excess of amine and Cs_2CO_3 base, **Table 20**. Upon stirring at 80 ºC for 1 hour and subsequent aqueous work-up, a 1:2 mixture of mono **159** to di-addition **162** products was observed and confirmed by ^1H NMR spectroscopy. In order to combat this, we increased the equivalents of amine and base, and decreased the temperature (70 ºC) and after 5 hours and upon subsequent work-up, we observed a 3:1 ratio of mono **159** to di-addition **162** products, which could be separated by column chromatography.

The next step was to reduce the nitro group to the corresponding amine prior to the amide coupling process. Nitro derivative **159** was treated with iron and ammonium chloride in an EtOH/H_2O mixture at 50 ºC overnight. Upon basic work-up and purification *via* flash column chromatography, amine **160** was isolated in moderate yield (46%). From here, we proposed an amide coupling reaction followed by a cleavage of the benzyl group to deliver our desired product. In addition, with the 3′ nitro pyrazole analogue, we envisaged a global nitro and benzyl reduction under hydrogenation conditions. However, when amine **160** was reacted with acid **97b** and HATU, the major product formed was *via* the amine at the 2-position of the pyridine, **161**, **Scheme 41**. The presence of a peak at 4.13 ppm in the ^{1}H NMR spectrum was believed to be the NH_2 which was confirmed to be a XH peak through its disappearance upon addition of D_2O. Further investigation with the aid of 2D NMR experiments provided us with evidence that **161** was indeed the isolated product. More specifically, a HMBC NMR experiment highlighted a relationship between the ^{13}C peak of the amide carbonyl carbon and the ^{1}H of the CH_2 of the benzyl motif (3 bond correlation). This relationship would not be observed if the 4-amino pyridine product was isolated (7-bond correlation). See **Appendix 10.5** for the associated spectra.

Both of the pyridyl amines are weakly nucleophilic, however, due to the inductive effect of the benzyl group, the ortho amine is possibly slightly more nucleophilic. Therefore, in order to circumvent this problem, we proposed the addition of a secondary protecting group, **Scheme 42**.



Scheme 42: Protecting group strategy for 2-amino pyridine derivatives.

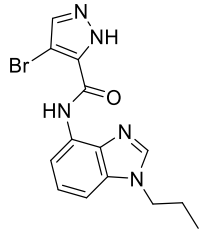
The 4-nitro species **159** was treated with *tert*-butyl dicarbonate, triethylamine and a catalytic amount of DMAP. After stirring at room temperature overnight, TLC analysis confirmed no new products had formed. Therefore, the temperature was increased and the reaction stirred for a further 3 days. Upon work-up and trituration in hexanes, the 4-nitropyridine derivative **163** was isolated in excellent yield (97%).

The final steps to furnish these desired 2-aminopyridine have yet to be performed. We hope that the products and the possible intermediates will provide more insight into the SAR of these hybrid species.

7 Conclusions

The protein kinases are an important class of therapeutic targets for small molecule inhibitors. They play essential roles in cancer progression and are therefore of high therapeutic value. DYRK2 is an important kinase in tumour growth *via* the induction of apoptosis and maintenance of the cell cycle.[38,39] More recently, DYRK2 has been identified as a key priming kinase for the 26S proteasome and there is evidence to support that DYRK2 inhibition reduces tumour burden in TNBC mice xenografts.[79,91] Through collaboration with scientists at Glasgow University and the University of Dundee we have identified a relationship between DYRK2, our target kinase, and TNBC, and furthermore, DYRK2 and the master regulator of proteotoxic stress, HSF1. We discovered that low DYRK2 expression is associated with an increased survival rate in TNBC patients. Secondly, KO studies revealed that both DYRK2 and HSF1 are essential for TNBC cell growth. In addition, we believe that DYRK2 stabilises HSF1 *via* phosphorylation, and we have identified two significant phosphorylation sites Ser320 and Ser326, which have been reported to promote HSF1 activity.

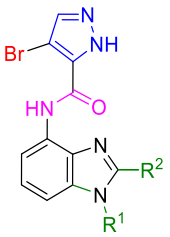
To date there are no published selective inhibitors for DYRK2. Previous work within the group revealed a potent and selective inhibitor for DYRK2 (**CI709** (**41**) K_i 210 nM, **Figure 104**).



CI709
DYRK2 K_i 210 nM
DYRK1A IC_{50} >30 µM

*Figure 104: Previous lead **CI709**.*

The reasons for the impressive activity and potency were unclear and required further investigation. Thus, the aims of this project were to identify and explore the importance of all areas of the successful inhibitor scaffold, whilst increasing the potency and maintaining the selectivity for assessment in TNBC cells. In doing so, a selective inhibitor with increased potency (>10 fold) was discovered (**68** 19 nM), **Figure 105**.



68
$R^1 = CH_2CH_2Cl$, $R^2 = H$
DYRK2 K_i 19 nM
DYRK1A IC_{50} > 10 µM

*Figure 105: Hit compound **68**.*

We investigated the diversification of four areas on the inhibitor scaffold: the head group heterocycle (blue, **Figure 105**), the substitution on the head group (red, **Figure *105***), the substitution on the benzimidazole tail group (green, **Figure 105**) and finally the amide linker (pink, **Figure 105**). Once synthesised, the analogues were evaluated *in vitro*. This was carried out *via* two methods: outsourcing to Life Technologies for single point (SP) at 1 µM or dose-response data (IC_{50}) and/or through an in-house biochemical assay (SP at 1 µM or dose-response data in the form of K_i).

7.1 Inhibitor overview

We identified that the pyrazole head group provided the most potent inhibitors, as replacement with alternative heterocycles resulted in inactivity at 1 µM. The pyrazole acts as a hinge binding motif providing the typical HBD/HBA interactions with the hinge region of the active site. Investigation of alternative two heteroatom containing systems to mimic this, e.g. oxazole and thiazole derivatives, were inactive at 1 µM. However, these derivatives did not contain the bromo substitution similar to that of **CI709** (**41**), therefore, leading us to determine the importance of substitution on inhibitor potency. The synthesis of these analogues was not considered during this project and thus requires further examination. Moreover, we determined that the bromide at the 4′ position on the pyrazole ring was essential for potency. Substitution with other halogens (chloro and iodo analogues) resulted in less active inhibitors, chloro **40** K_i 335 nM and iodo **42** 430 nM. In addition, replacing the bromide with a possible bioisostere, the ethynyl group, also resulted in a less active species **49** K_i 510 nM. Moreover, deletion of substitution at this position results in complete loss of DYRK2 activity. We believe that the bromide is an optimal size to occupy a small hydrophobic pocket in the DYRK2 active site, **Figure 106**.

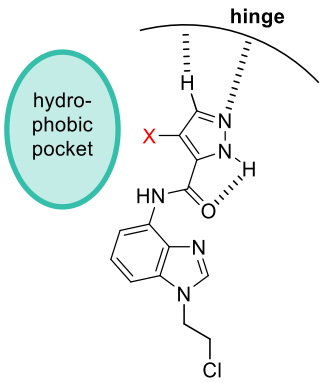
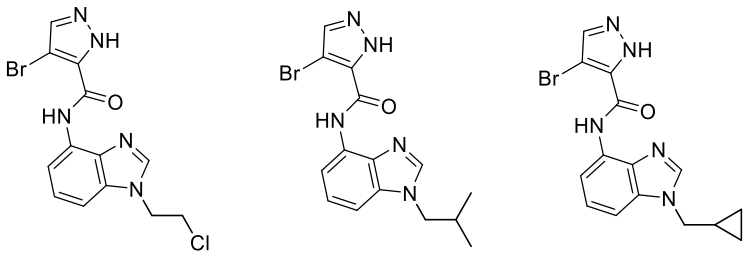


Figure 106: Representation of key interactions.

Further investigation into substitution at the 3′ position on the pyrazole revealed that the introduction of a HBA (nitro) or HBD (amino) resulted in an inactive species at 1 µM. This was an interesting result as we postulated that the introduction of a strong HBD motif could interact with the hinge region, however, modelling studies revealed possible repulsive interactions and therefore a change in conformation of the inhibitor could contribute to the loss of activity. More specifically, the amine at the 3-position is <2 Å away from Glu229 (GK+1) of the hinge and could result in a clash.

We extensively investigated substitution on the benzimidazole tail group. Replacement of the benzimidazole with larger ring systems resulted in less active inhibitors. More specifically, quinoline derivative **57** (K_i 1 µM) and quinoxaline derivative **56** (K_i 600 nM) were >30-fold less active than inhibitor **68** (19 nM). These results highlighted the importance of the substitution pattern on the benzimidazole for potency. In addition, we discovered that substitution at the *N*-1 position of the benzimidazole provided inactive inhibitors. Substitution at this position may result in repulsive interactions with hydrophobic residues, including Ile155 in the DYRK2 active site. Furthermore, investigation of the substitution pattern on the *N*-3 position of the benzimidazole revealed the most potent inhibitor of this series (**68**, K_i 19 nM). A range of functionality was explored, both aliphatic and aromatic and also the introduction of HBA motifs and electronegative atoms. We determined that a three atom chain length was optimal for potency, **68** 19 nM, **65** 64 nM and **76** 55 nM, **Figure 107**.



68
K_i 19 nM

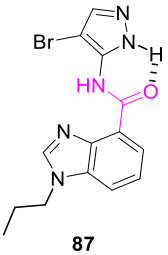
65
K_i 64 nM

76
K_i 55 nM

Figure 107: Most potent inhibitors of benzimidazole substitution series.

We believe that the three atom chain length directs the final atom to a prime position for interaction with Phe160 and Lys157 residues. In addition, we investigated the introduction of functionality at the C-2 position of the benzimidazole. Both hydrophobic (methyl and ethyl) and hydrophilic (keto/enol) groups were introduced, however, this resulted in inactive analogues. These substitution patterns we believe alter the conformation of the entire structure and therefore result in the loss of key interactions.

The final area of investigation was the amide linker. We determined that the reverse amide analogue **87** was inactive, possibly due to favourable intramolecular interactions resulting in a change in binding mode, **Figure 108**.

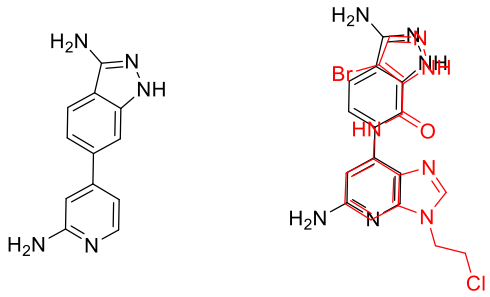


87
Inactive @ 1 µM

*Figure 108: Reverse amide analogue **87**.*

In order to maintain the key hinge binding interactions and the position of the bromide in the back pocket, the benzimidazole tail could flip which would be unfavourable due to the restricted space of the pocket. Therefore, inhibitor **87** must adopt an alternative conformation to avoid repulsive interactions and thus results in an inactive inhibitor.

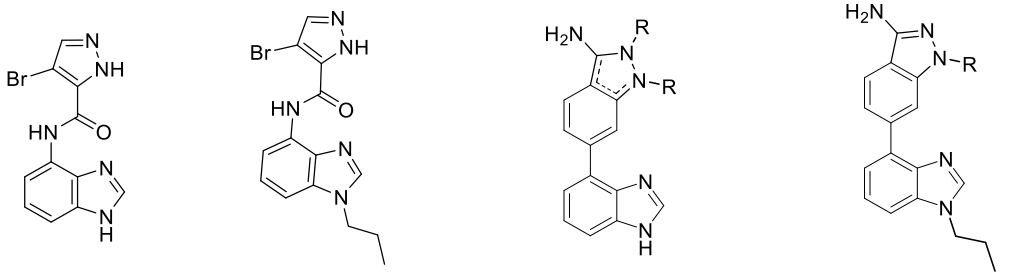
In addition, after a small in-house screen of kinase like structures we identified an inhibitor scaffold similar to ours with impressive DYRK2 activity, **88** K_i 2 nM, **Figure 109**. However, this compound was not selective for DYRK2 over its closely related family member DYRK1A at 1 µM, **Figure 109**.



88
DYRK2:K_i 2 nM
DYRK1A 1μM: 97% ± 0

*Figure 109: Screening hit **88** and overlay with our scaffold.*

Therefore, we designed hybrid molecules to determine the key areas of affinity for each scaffold. We identified that indazole analogues **89**–**92** were less active than the corresponding bromo amide analogues **CI709** (**41**) and **CI639** (**45**), **Figure 110**.



CI639 (45)
K_i 102 nM

CI709 (41)
K_i 210 nM

90 R=H *K_i 758 nM*
92 R=Me *inactive @ 1 μM*

89 R=H *K_i 1300 nM*
91 R=Me *K_i 1700 nM*

*Figure 110: Indazole hybrids and the original scaffolds **CI639** (**45**) and **CI709** (**41**).*

This provided us with more evidence of the size restrictions within the DYRK2 active site. The introduction of the rigid indazole may result in the benzimidazole bottom half adopting a different conformation and thus resulting in a loss of binding interactions. This is complimented by the unsubstituted benzimidazole analogue **90** (758 nM), **Figure 110**. Moreover, through the biological assessment of the methyl substituted indazoles **91** and **92**, we identified the motif required for hinge binding.

In addition to the indazole hybrid species **89**–**92**, we began to investigate the amino pyridine hybrid analogues **93** and **94**, **Figure 111**. The synthesis of these is currently underway.

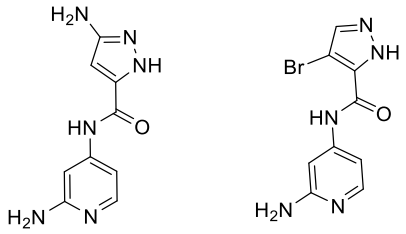
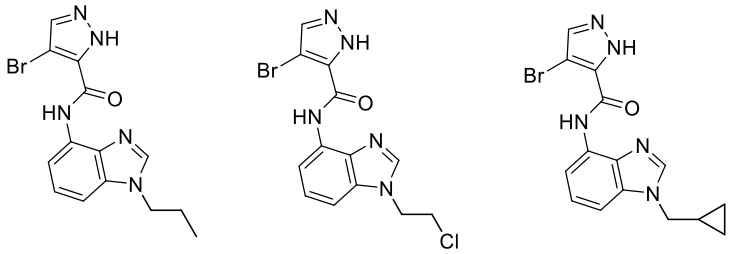


Figure 111: Amino pyridine hybrid species.

7.2 DMPK properties

From the investigation of the *in vitro* drug metabolism and pharmacokinetic properties of our most potent inhibitors we discovered an encouraging permeability and solubility profile but also identified a metabolic stability challenge, **Table 21**.

Table 21: Summary of physicochemical and PK properties.



Properties	CI709 (41)	68	76
K_i (nM)	210	19	55
MW (g)	348.2	368.6	360.2
TPSA (Å^2)	69.1	69.1	69.1
CLogP	2.7	2.2	2.6
$t_{1/2}$ (mins)	11.3	25.2	3.5
CL_{int} (µL/min/mg of protein)	61.3	27.5	200
P_{app} A–B (10^{-6} cms^{-1})	22.7	-	-
P_{app} B–A (10^{-6} cms^{-1})	27.3	-	-
Efflux Ratio	1.21	-	-
Turbidimetric solubility (µM)	100	-	-

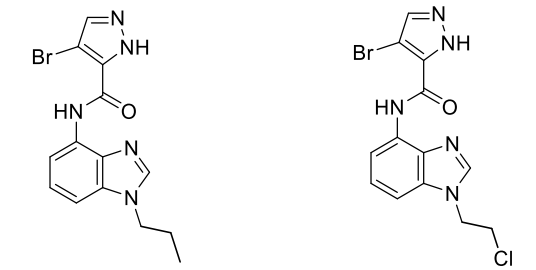
These inhibitors exhibit good physicochemical properties which adhere to the 'Rule of Five' associated with an increased chance of oral bioavailability, MW <500, CLogP <5 and HB capabilities, **Table 21**.[114] In addition, the associated TPSA values (69.1 $Å^2$) are agreeably within the reported threshold for good bioavailability (<140 $Å^2$). However, a lower limit (<75 $Å^2$) has been associated with toxicity *in vivo* and therefore should be considered at an early stage.

The previous lead, **CI709** (**41**) was discovered to have an encouraging Caco-2 permeability profile, **Table 21**. Where it is comparable to the highly permeable drug propranolol, which has P_{app} values of A–B 22.6 and B–A 38.3 10^{-6} cms^{-1} respectively. Moreover, the efflux ratio provides us with promising evidence that **CI709** (**41**) will not undergo active efflux. In addition, the associated turbidimetric solubility for **CI709** (**41**) is also encouraging as it has revealed that the chances of precipitation in aqueous media requires a high concentration of inhibitor >100 µM, **Table 21**. Moreover, we can potentially infer that these properties would remain for similar structures of this inhibitor scaffold.

Finally, inhibitors **CI709** (**41**), **68** and **76** were externally investigated for their metabolic stability *in vitro* in the S9 liver fraction assay. We identified that the cyclopropyl analogue **76** was metabolically unstable ($t_{1/2}$ = 3.5 mins, CL_{int} = 200 µL/min/mg of protein). The *n*-propyl analogue **CI709** (**41**) was more stable ($t_{1/2}$ = 11.3 mins CL_{int} = 61.3 µL/min/mg of protein) and the chloro analogue **68** was the most metabolically stable ($t_{1/2}$ = 25.2 mins CL_{int} = 27.5 µL/min/mg of protein). However, provided that the chloro analogue **68** has an electrophilic handle, it has the potential to be a covalent inhibitor and react with nucleophilic residues in the DYRK2 active site such as a lysine residue. This could result in an increase in the $t_{1/2}$ due to the stronger binding interactions and a slower metabolic breakdown of **68**. Moreover, from these results we postulate that the aliphatic substitution on the benzimidazole is oxidised in first pass metabolism, possibly by CYP450 enzymes. Therefore, in order to progress with this scaffold, further investigation into determining the specific metabolites would be worthwhile.

7.3 Investigation of inhibitors in TNBC cells

Our collaborator identified a negative relationship between DYRK2 expression and tumour burden in TNBC mice xenografts and also determined that DYRK2 expression was an important kinase in HSF1 activation. In order to further examine the relationship between DYRK2 and TNBC cell proliferation and DYRK2 and HSF1 phosphorylation, we set out to determine the effects of two inhibitors, **CI709** (**41**) and **LB35** (**68**), in TNBC cells, **Figure 112**.



CI709 (41)
K_i 210 nM
DYRK1A IC_{50} >30 µM

LB35 (68)
K_i 19 nM
DYRK1A IC_{50} >10 µM

*Figure 112: Reminder of inhibitors **CI709** (**41**) and **LB35** (**68**).*

More specifically, we investigated the ability of these inhibitors to 1) influence cell growth and 2) inhibit the phosphorylation of HSF1. We have found that **CI709** (5 µM) and **LB35** (1 µM) significantly reduced cell growth in the MDA-MB-468 cell line. In addition, through investigation of the expression of an important protein in the HSF1 pathway, Hsp70, we have provided more evidence to reinforce that there is a relationship between DYRK2 and HSF1. The treatment of 468 cells with both **CI709** and **LB35** appreciably reduced Hsp70 expression, where KO studies confirmed that the reduction was due to the DYRK2 inhibition. Next, with respect to HSF1 phosphorylation, upon HS, we observed that in 468 cells, **CI709** inhibits the phosphorylation of HSF1 at Ser320 and Ser326 and **LB35** inhibits the phosphorylation of HSF1 at Ser320.

We also investigated the effect of the inhibitors in healthy cells, 293T in particular, on HSF1 phosphorylation at Ser320. Treatment of the cells with **CI709** and **LB35** (1 µM) resulted in the loss of phosphorylation at Ser320.

Finally, we briefly examined the relationship between DYRK2 and the proteasome *via* the E3 Ub ligase SIAH2, where DYRK2 phosphorylates SIAH2 at Ser28.[144] We observed that treatment with both **CI709** and **LB35** (at 1 and 5 µM) resulted in the loss of phosphorylation at Ser28 in 293T cells. These findings provide us with more

evidence to support the association between DYRK2 and protein degradation. Inhibition of DYRK2 could impair the proteasome *via* a number of pathways, for example, impairing the function of the proteasome itself (inhibiting phosphorylation of Rpt3)[79] or *via* disrupting the activation of SIAH2 which tags proteins for poly-ubiquitination, and recognition by the proteasome. Therefore, inhibiting either of these pathways could prove beneficial in the treatment of cancer.

7.4 SAR summary

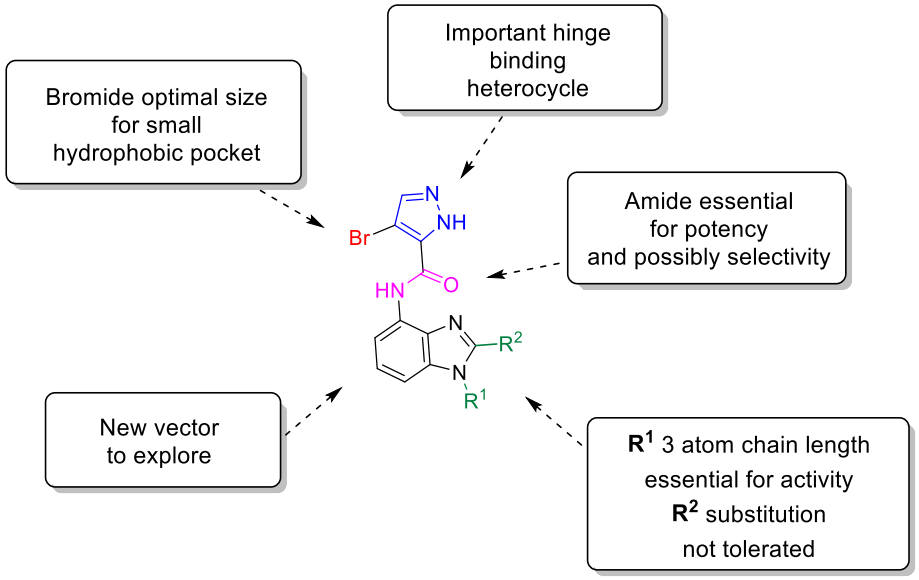


Figure 113: Key SAR highlights.

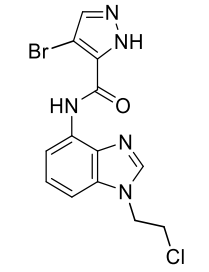
We have identified that the bromide at the 4′ position on the head group pyrazole is essential for potency potentially through its ability to bury itself into a small hydrophobic pocket within the DYRK2 active site. However, further investigation of the pyrazole head group with alternative hinge binding heterocycles, including the bromide substitution could result in an inhibitor with increased binding affinity. We have extensively investigated the substitution of the benzimidazole portion of the inhibitor scaffold and a trend has been revealed. More specifically, smaller aliphatic chains ≤ 3 atoms in length provided the most potent inhibitors. Where the n=3 chain length provided the three most potent inhibitors of this series (<70 nM). However, with the replacement of the bromide with the smaller halogen, a longer chain on the benzimidazole was tolerated n=4 (**62** K_i 325 nM). Substitution at the R^2 position, (**Figure 113**) was not tolerated. Next, we examined the importance of the amide linker and from a brief investigation we have determined that it is also pivotal for potency. Moreover, the activity of the alternative linker analogues is yet to be assessed against the closely related isoform DYRK1A in order for us to determine the rationale behind the impressive selectivity of this series. Finally, as a new hit was identified through an

in-house screen, this revealed an amino pyridine motif to be extremely active for DYRK2 (**88** K_i 2 nM) but not selective. Therefore, investigating the addition of substitution on the aryl ring of the benzimidazole could provide an inhibitor with improved potency and metabolic stability. Finally, the results from de la Vega have provided us with convincing evidence that DYRK2 inhibition reduces cell proliferation and Hsp70 expression in TNBC cells. In addition, we have evidence to support that **CI709** (**41**) and **LB35** (**68**) inhibits the phosphorylation of HSF1 in TNBC cells, including phosphorylation at Ser326, which has been recognised as an essential modification in HSF1 activation. Therefore, from these exciting and preliminary findings, we believe that DYRK2 requires further investigation as a viable target for the treatment of TNBC.

8 Future Work

8.1 Determine the toxicity of 68

The most potent inhibitor of this series **68** (K_i 19 nM), **Figure 114**, contains an electrophilic handle, which could react with nucleophilic residues on our target protein, such as cysteine and lysine. From the amino acid sequence there are no cysteine residues in the DYRK2 active site, however, retaining the chloroethyl motif could result in toxicity. Therefore, if we wish to continue with this functionality on our inhibitor scaffold, we must investigate the toxicity profile of this inhibitor.

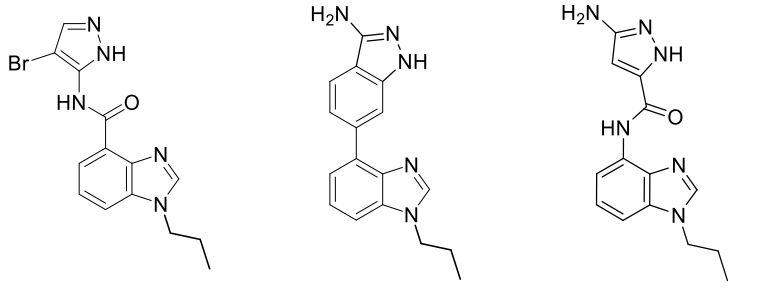


68
DYRK2 K_i 19 nM
DYRK1A IC_{50} >10 µM

*Figure 114: Lead compound **68**.*

8.2 Selectivity profile

Next, we would investigate the selectivity over DYRK1A associated with the inhibitors highlighted below, **Figure 115**.



87
22% inhibition @ 1 µM

89
K_i 1300 nM

53
15% inhibition @ 1 µM

Figure 115: Inhibitors to be assessed for DYRK1A activity.

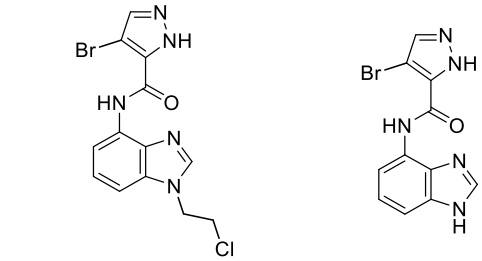
The DYRK1A active site is more accommodating than the DYRK2 active site, with only three different AA residues. Our investigation revealed reverse amide **87** and amino pyrazole **53** to be inactive at 1 µM. In addition, our studies identified that

indazole **89** was >50 fold less active than our lead species (1300 vs 19 nM). Therefore, determining the DYRK1A activity of these species would provide us with an understanding of which functionality is essential for the DYRK2 selectivity.

Moreover, the previous lead **CI709** (**41**) was screened against a kinase panel for its selectivity among a representative group of kinases from the kinome. In order to confirm that our new lead **68** is also highly selective, it would be advantageous to determine its selectivity profile among the same group of kinases

8.3 Improve metabolic stability

Our SAR studies have revealed a potent and selective DYRK2 inhibitor **68**. However, upon investigation of the *in vitro* PK parameters, we discovered that the metabolic stability of **68** requires improvement ($t_{1/2}$ = 25.2 mins). We believe that the initial site of metabolism is the aliphatic chain on the benzimidazole tail, **Figure 116**. The chloro analogue **68** blocks this site from aliphatic hydroxylation and therefore results in an improved $t_{1/2}$ in comparison to **CI709** (**41**) ($t_{1/2}$ = 11.3 mins). Investigation of the metabolic stability of the unsubstituted **CI639** (**45**) would provide us with an indication of the metabolic stability if we were to remove this motif from our inhibitor series, **Figure 116**.



68
K_i 19 nM
$t_{1/2}$ 25.2 mins
CL_{int} 27.5 µL/min/mg of protein

CI639 (45)
K_i 102 nM

Figure 116: Metabolic stability investigation.

In addition, introducing further substitution onto the benzimidazole ring should be investigated to aid in the improvement of the metabolic stability of this inhibitor series.

8.4 Further investigation in TNBC cells

In order to further understand the effect of our inhibitors **CI709** (**41**) and **LB35** (**68**), we would seek to investigate these inhibitors in alternative TNBC cell lines and compare with the current data. This would be beneficial as not all TNBC tumours have

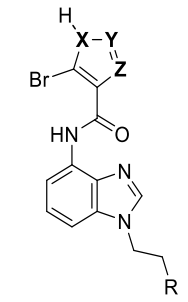
the same molecular framework and, therefore, similar results for different cells would highlight any advantages of DYRK2 inhibition. These cell lines could include MDA-MB-321 and/or MCF107A cells.

In order to further exploit the relationship between DYRK2 and the proteasome, investigating the effect of our inhibitors on SIAH2 phosphorylation in TNBC cells would provide us with more insight into the benefits of targeting the proteasome. Previous work has shown that SIAH2 levels are higher in breast cancer cells in comparison to normal cells.[145] Thus, determining the effect of our inhibitors on SIAH2 phosphorylation in TNBC would provide us with further understanding of the relevance of our inhibitors.

8.5 Consolidate SAR

In order to fully exploit the understanding that we have developed for our inhibitor series this far, there are some areas of the scaffold that could be investigated further.

The exploration of the head group with an alternative heterocycle that also includes the bromide substitution **165**, **Figure 117**.



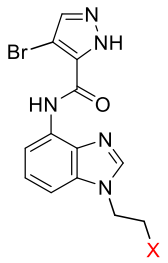
165
X=N, Y/Z= C,N,O or S

Figure 117: Further investigation of the head group.

Investigation of these analogues would provide us with a further understanding of the key hinge binding atoms and whether an intramolecular hydrogen bonding interaction between the head group and amide carbonyl group is important for inducing a stable conformation.

We have provided evidence that the most potent inhibitors include benzimidazole substitution with a three atom aliphatic chain. We have postulated that the introduction of a HBD motif at the third atom position could interact with Lys157 and possibly result in an inhibitor with increased binding affinity. Our efforts thus far to deliver an inhibitor

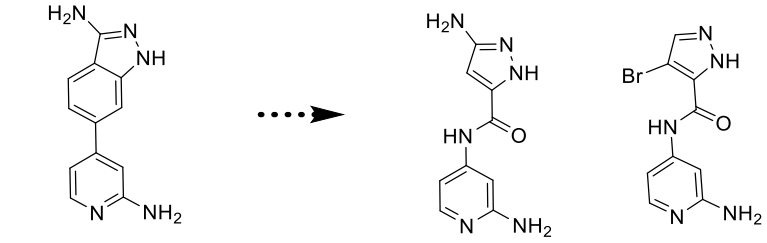
of this nature have been unsuccessful, however, we have identified a promising strategy to access this analogue which requires investigation, **Figure 118**.



166
X=OH,SH or NH_2

Figure 118: Introduction of a HBD.

We have also identified an alternative hit scaffold with impressive DYRK2 activity **88** (K_i 2 nM) and the synthesis of hybrid analogues has begun. The synthesis and evaluation of the 2-amino pyridine analogues **93** and **94**, would provide us with important information about which functionality is responsible for the impressive selectivity of our series, **Figure 119**.



88
DYRK2:K_i 2 nM
DYRK1A 1µM: 97% ± 0

93

94

Figure 119: Hybrid analogues.

In addition, the overlay of the original scaffold with the new scaffold has highlighted a new vector for exploration, **Figure 120**.

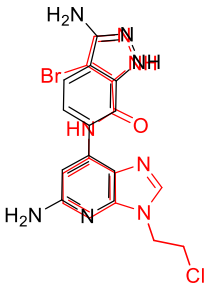


*Figure 120: Overlay of original **68** (red) and new **88** (black) scaffolds.*

With respect to the addition of the amino group on the aryl ring, this has the potential to interact with Asp295 in the DYRK2 active site and may provide us with an inhibitor with increased potency. Additionally, the introduction of a polar group is also attractive for PK purposes with respect to improved solubility and bioavailability. However, it also provides a handle for phase II metabolism.

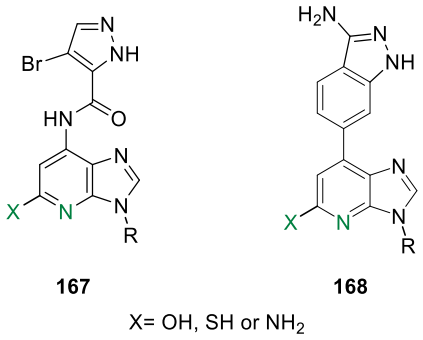


Figure 121: Exploring a new vector.

The replacement of the benzimidazole with the pyridyl equivalent **167** and **168** would also be beneficial from an aqueous solubility perspective, **Figure 121**. The addition of a heteroatom would increase the TPSA and decrease the CLogP. Although, this could provide a permeability penalty as the inhibitors become more hydrophilic in nature. Moreover, the introduction of further complexity can also result in a synthetic challenge.

Throughout the investigation of these highlighted vectors: (1) determining if our hit species is a covalent inhibitor (2) the DYRK1A activity of the less active species, (3) the functionality to block metabolic 'hot spots' on our inhibitor scaffold, and (4) consolidating our SAR we should achieve a better understanding of the nature of our DYRK2 selectivity. Moreover, we should be able to deliver a potent inhibitor species with improved metabolic stability ($t_{1/2}$ >60 mins) whilst maintaining good permeability and solubility properties. Finally, successful inhibitors will be assessed in a TNBC cell model by our collaborator to determine their effect on cancer cell maintenance.

9 Experimental

9.1 DYRK2 Biochemical Assay

The associated DYRK2 inhibitory constants (K_i) were determined using a luminescence detection ADP-Glo™ Kinase assay. The protocol provided by Promega was optimised by Louise Young, Gillian Berrie and the author, Laura Bain.

DYRK2 enzyme 3ng/4µL (Promega, Southampton, UK) was incubated with DYRK specific substrate DYRKtide (25 µM, Genscript, Leiden, Netherlands) and ATP (15 µM or 5 µM see **Section 2**) in assay buffer (40 mM Tris-HCl, 20 mM $MgCl_2$, 50 µM and DTT 0.1 mg/mL BSA (pH 7.5)) in a V-bottomed 96 well plate in the presence and absence of test compound. The assay plate was then incubated for 60 mins at room temperature in the absence of light. The kinase reaction was then quenched with the ADP-Glo™ reagent and the assay plate was incubated for a further 40 mins at room temperature, again in the absence of light. Thereafter, the Kinase-detection reagent was added to promote the luciferase reaction and to allow the luminescence reading, and the assay plate was incubated for a final 30 mins at room temperature, in the absence of light. The resulting mixture was transferred to a white half-area flat-bottomed 96 well plate for analysis. The relative luminescence units were read on a Hidex plate reader (LabLogic, Sheffield, UK) and converted to the amount of ADP produced as a percentage of the control experiments. The apparent K_i values were calculated using GraphPad Prism software for three individual experiments *via* the Cheng-Prusoff equation (see **Appendix 10.4**).

9.2 Chemistry – General

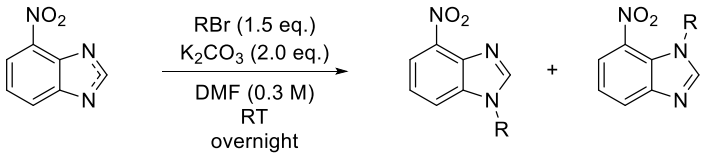
All reagents were obtained from commercial sources, such as Sigma-Aldrich and Alfa Aesar, and used without purification. All solvents were obtained through the SPS system at the University of Strathclyde. Microwave reactions were carried out using a BIOTAGE initiator microwave. Reactions monitored by TLC were done so using Machery-Nagel pre-coated TLC sheets coated in 0.20 mm silica gel 60 with UV_{254} fluorescent indicator. Purification was performed by flash chromatography, with chromatography grade silica 60 Å particle size 35–70 micron from Fisher Scientific, using the solvent systems stated.

^{1}H and ^{13}C NMR were carried out using a Bruker Avance 3 (^{1}H 400 MHz and ^{13}C 101 MHz) and a Bruker Avance 500 (^{1}H 500 MHz and ^{13}C 125 MHz) as stated. The spectra were recorded in the deuterated solvent $CDCl_3$ (chloroform), DMSO-d_6 (dimethylsulfoxide) and acetone-d_6. Multiplicities were indicated as follows: s (singlet); d (doublet); t (triplet); dd (doublet of doublets); m (multiplet) *etc.* Coupling constants (*J*) were given in Hertz (Hz). Chemical shifts were reported in parts per million (ppm) relative to tetramethylsilane (TMS) (δ = 0.00). Peaks with δ values of 7.26 and 77.40 ppm (^{13}C NMR) correspond to the residual solvent peak for $CDCl_3$; 2.50 and 39.52 for DMSO-d_6, and 2.05 and 29.84 for acetone-d_6 (^{1}H and ^{13}C NMR respectively).

Mass spectrometry was carried out on an Agilent 6130 liquid chromatograph with electrospray mass spectrometer using a 5–100% acetonitrile in water with ammonium hydroxide or formic acid solvent system. Infrared spectra were recorded in the range 4000–600 cm^{-1} on a Shimadzu IRAffinity-1 equipped with an ATR accessory. Melting points were determined on a Stuart SMP11 and are uncorrected.

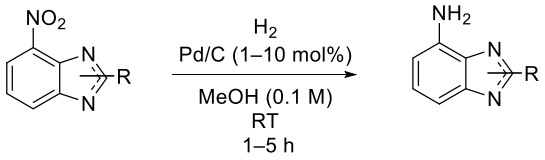
9.3 General procedures

General Procedure **A**: Benzimidazole Alkylation



To a microwave vial was added nitrobenzimidazole (1.0 eq.) and potassium carbonate (K_2CO_3) (2.0 eq.). The vial was sealed and anhydrous *N,N*-dimethylformamide (DMF) was added (0.3 M) followed by the addition of the alkylating agent (RBr) (1.5 eq.). The resulting mixture was allowed to stir at room temperature overnight. Upon completion, the mixture was diluted with EtOAc (10 mL) and washed with water (20 mL). The aqueous layer was extracted with ethyl acetate (EtOAc) (3×20 mL), and the combined organics were dried over magnesium sulfate ($MgSO_4$), filtered and concentrated *in vacuo*. The crude material was purified *via* flash column chromatography to isolate the regioisomers.

General Procedure **B**: Hydrogenation of nitrobenzimidazoles

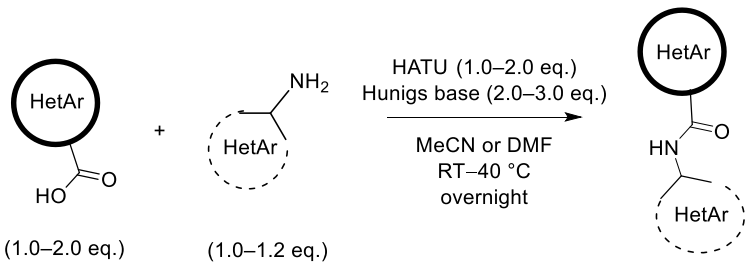


To a flame-dried microwave vial purged with argon was added Pd/C (1–10 mol%) followed by a solution of nitro compound (1.0 eq.) in the appropriate solvent (0.1 M). The mixture was then exposed to a H_2 atmosphere by the application of vacuum followed by a H_2 balloon. This sequence was repeated 5 times. H_2 was then bubbled through the mixture for 5–10 mins through a needle immersed in the solution. Following this, another H_2 balloon was pierced through a septum in the flask (unsubmerged) and the reaction was allowed to stir at room temperature for up to 5 hours. Upon completion, the H_2 balloon was removed and the resulting solution was diluted with the appropriate solvent (5–10 mL) and filtered through a plug of Celite®, washing with further solvent (20 mL). The filtrate was then concentrated *in vacuo* to afford the product which was used in the next step without further purification.

General Procedure **C**: Fe Reduction of nitrobenzimidazoles

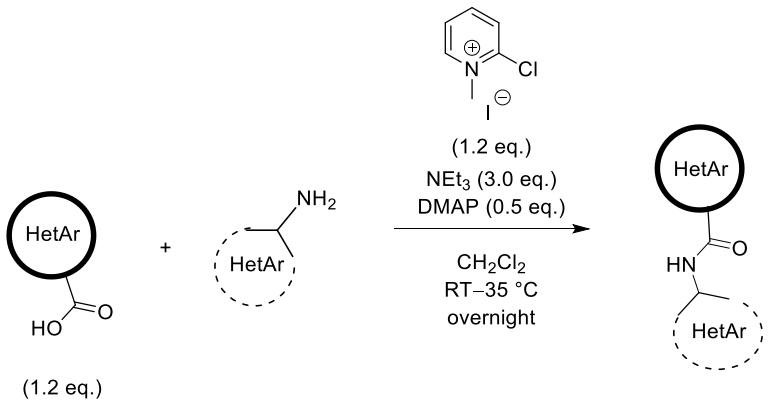
To a round bottomed flask was added $FeSO_4 \cdot 7H_2O$ (6.0 eq.) and citric acid (0.5 eq.) in H_2O (0.01 M) and the reaction mixture stirred at room temperature for 5 minutes. To the stirring solution was added $NaBH_4$ (10 eq.) portion-wise and the resulting mixture was stirred vigorously for a further 5 minutes at room temperature. The resultant slurry was decanted into a beaker and discarded. The flask was then rinsed with H_2O (2×20 mL) leaving behind a minimal amount of solid Fe surrounding the stirrer bar in the flask. The nitro compound (1.0 eq.) was then taken up in H_2O (1–2 mL) and added to the flask as a suspension. The reaction mixture was allowed to stir vigorously at room temperature overnight. Upon completion, the mixture was filtered through Celite® and washed with water (2×30 mL) and saturated sodium hydrogen carbonate ($NaHCO_3$) solution (2×30 mL). The aqueous layer was then extracted with EtOAc (4×30 mL). The combined organics were washed with brine (20 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo* to afford amine product which was used in the next step without further purification.

General Procedure **D**: Amide Coupling with HATU



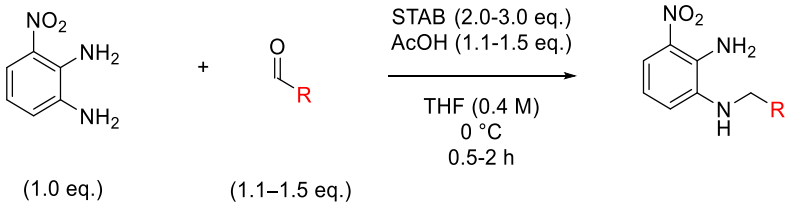
To a flame dried microwave vial purged with argon was added acid substrate (1.0–2.0 eq.) and HATU (1.0–2.0 eq.) followed by a solution of the amine (1.0–1.2 eq.) in the appropriate solvent. To the resulting solution was added Hünigs base (2.0–3.0 eq.) and the reaction mixture was allowed to stir at 25–40 °C overnight. Upon completion, the mixture was diluted with EtOAc (10 mL) then washed with H_2O (3×10 mL) and brine (10 mL). The organics were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography provided the desired amide.

General Procedure **E** – Amide Coupling with 2-chloro-1-methylpyridin-1-ium iodide



To a flame dried microwave vial purged with argon was added the acid substrate (1.2 eq.) in anhydrous CH_2Cl_2 (0.3 M). The vial was then wrapped in aluminium foil and placed in an oil bath at room temperature for 5 mins. To the resulting solution was added 2-chloro-1-methylpyridin-1-ium iodide (1.2 eq.), *N*,*N*-dimethylpyridin-4-amine (DMAP) (0.5 eq.), trimethylamine (NEt_3) (3.0 eq.) and amine substrate (1.0 eq.). The reaction mixture was then allowed to stir at 35 °C overnight. Upon completion, the reaction mixture was allowed to cool to room temperature before quenching with 10 % citric acid solution (2–5 mL) and stirring for 10 minutes at room temperature. The layers were separated and the aqueous extracted with CH_2Cl_2 (3×10 mL). The combined organics were then washed with brine (10 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography afforded the desired amide product.

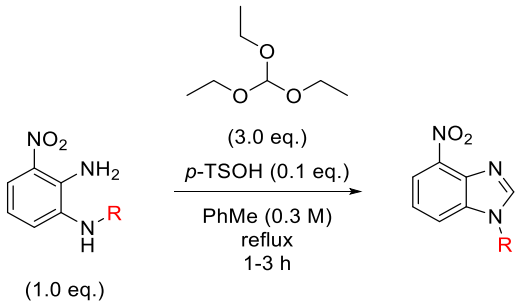
General Procedure **F** - Reductive Amination of 3-Nitro-1,2-benzenediamine



To a flame dried vial purged with argon was added 3-nitrobenzene-1,2-diamine (1.0 eq.) in anhydrous THF (0.4 M). The resulting mixture was cooled to 0 °C before the addition of the appropriate aldehyde (1.1–1.5 eq.), sodium triacetoxyborohydride (STAB) (2.0–3.0 eq.) and acetic acid (AcOH) (1.1–1.5 eq.). The resulting solution was allowed to stir at 0 °C until completion, monitoring by TLC. Upon completion, the mixture was quenched with saturated $NaHCO_3$ solution (10 mL) and allowed to stir at room temperature for 15 mins. The layers were then separated and the aqueous

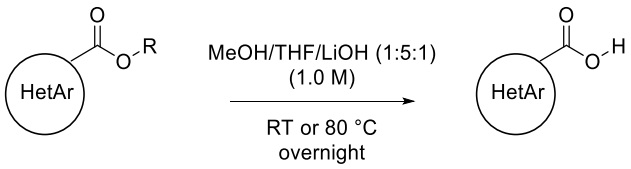
extracted with CH_2Cl_2 (3×10 mL). The combined organics were washed with brine (10 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography provided the desired products.

General Procedure **G**: Ring Closure to Benzimidazoles



To a solution of nitro compound (1.0 eq.) in toluene (0.3 M) was added triethylorthoformate (3.0 eq.) and *p*-toluenesulfonic acid (0.1 eq.).The reaction mixture was heated to reflux for 1–3 hours, monitoring by TLC. Upon completion the reaction was cooled to room temperature, concentrated *in vacuo* and purified *via* flash column chromatography to provide the target benzimidazole.

General Procedure **H**: Ester hydrolysis

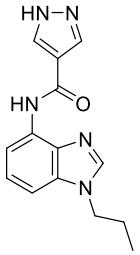


To a round bottomed flask was added the ester (1.0 eq.) in a mixture of 1:5:1 MeOH/ THF/1M LiOH (0.03 M) and stirred at RT or 80°C overnight. Upon completion, the mixture was allowed to cool to room temperature and then concentrated *in vacuo*. The resulting residue was dissolved in water (10 mL), cooled to 0°C and acidified to pH 1 with 2M aqueous HCl solution. A precipitate formed and was collected *via* filtration, washing with H_2O (2×10 mL) to afford the appropriate acid.

9.4 Inhibitors

9.4.1 Alternative head group heterocycle analogues

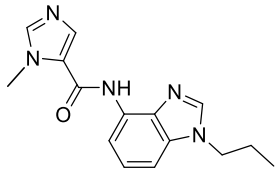
N-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-4-carboxamide **34**



According to **General Procedure E;** acid **114** (0.27 mmol, 65 mg), 2-chloro-1-methylpyridin-1-ium iodide (0.25 mmol, 69 mg), DMAP (0.11 mmol, 13 mg), NEt_3 (0.33 mmol, 90 µL) and amine **98a** (0.22 mmol, 39 mg) were stirred in anhydrous CH_2Cl_2 (0.7 mL) at 35 °C overnight. Partial purification *via* flash column chromatography in 99:1 CH_2Cl_2/MeOH afforded an inseparable mixture of product amide and starting material amine. The mixture was carried onto the next step without further purification.

To a flame dried microwave vial was added mixture (51 mg) in anhydrous THF (0.65 mL). The mixture was cooled to 0 °C and trifluoroacetic acid (TFA) (393 µL) was added dropwise. The resulting mixture was heated to 30°C and stirred at this temperature for 6 days. Upon completion, the reaction was diluted with EtOAc (5 mL) and washed with saturated aqueous $NaHCO_3$ solution (5 mL). The aqueous was then extracted with EtOAc (3×5 mL) and the combined organics washed with brine (5 mL), dried over $MgSO_4$ and concentrated *in vacuo*. Further purification *via* flash column chromatography in 95:5 to 90:10 CH_2Cl_2/MeOH afforded *title compound* **34** as a white solid (0.05 mmol, 14 mg, 23% over two steps). **MP** decomp. >164 °C; **IR** (ATR, cm^{-1}) 3406, 3090, 2922, 1658, 1495; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.28 (s, 1H, NH), 9.51 (s, 1H, C(O)NH), 8.50 (s, 1H, ArH), 8.24 (s, 1H, NCHN), 8.09 (s, 1H, ArH), 7.81 (d, *J* 8.0 Hz, 1H, ArH), 7.37 (d, *J* 8.0 Hz, 1H, ArH), 7.22 (app t, *J* 8.0 Hz, 1H, ArH) 4.23 (t, *J* 7.0 Hz, 2H, $CH_2CH_2CH_3$), 1.82 (app sextet, *J* 7.0 Hz, 2H, $CH_2CH_2CH_3$), 0.84 (t, *J* 7.0 Hz, 3H, $CH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 160.8, 143.1, 139.0, 135.6, 134.3, 130.3, 129.5, 122.4, 117.8, 113.5, 106.1, 45.8, 22.7, 10.9; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{15}N_5O$ 269.1 found 270.0 $[M+H]^+$.

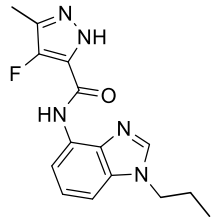
1-Methyl-*N*-(1-propyl-1*H*-benzo[d]imidazol-4-yl)-1*H*-imidazole-5-carboxamide **36**



According to **General Procedure D**; 1-methyl-1*H*-imidazole-5-carboxylic acid (0.58 mmol, 73 mg), HATU (0.58 mmol, 221 mg), amine **98a** (0.29 mmol, 50 mg) and Hünigs base (0.87 mmol, 0.1 mL) in anhydrous MeCN (3.2 mL) were stirred at room temperature overnight. Further purification *via* flash column chromatography in 95:5 EtOAc/MeOH afforded the *title compound* **36** as a white solid (0.21 mmol, 59 mg, 72%). **MP** 140–142 °C; **IR** (ATR, cm^{-1}) 3156, 3098, 2959, 1655, 1621; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 9.70 (s, 1H, C(O)NH), 8.25 (s, 1H, NCHNCH$_3$), 7.89 (s, 1H, NCHNR), 7.85 (s, 1H, NCHCC(O)), 7.73 (d, *J* 8.0 Hz, 1H, ArH), 7.40 (d, *J* 8.0 Hz, 1H, ArH), 7.23 (app t, *J* 8.0 Hz, 1H, ArH), 4.23 (t, *J* 8.0 Hz, 2H, NCH$_2$CH$_2$CH$_3$), 3.88 (s, 3H, NCH$_3$), 1.82 (app sept, *J* 8.0 Hz, 2H, NCH$_2$CH$_2$CH$_3$), 0.83 (t, *J* 8.0 Hz, 3H, NCH$_2$CH$_2$CH$_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.4, 143.3, 142.6, 136.0, 134.3, 132.9, 129.0, 125.8, 122.4, 113.8, 106.7, 45.8, 33.7, 22.7, 10.9; **LRMS** (LCMS-ESI) m/z calc. for $C_{15}H_{17}N_5O$ 283.14 found 284.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{15}H_{18}N_5O$ 284.1506 found 284.1506 $[M+H]^+$.

9.4.2 4′ Halogen pyrazole analogues

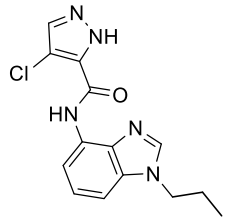
4-Fluoro-3-methyl-*N*-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **39**



According to **General Procedure D**; acid **106** (0.17 mmol, 25 mg), HATU (0.17 mmol, 65 mg), amine **98a** (0.14 mmol, 25 mg) and Hünigs base (0.28 mmol, 50 µL) in anhydrous DMF (0.5 mL) were stirred at 45°C overnight. Further purification *via* flash column chromatography in 9:1 EtOAc/petroleum ether 40–60 and further trituration in hexanes afforded the *title compound* **39** as a white solid (0.07 mmol, 21 mg, 50%). **MP** decomp. >180 °C; **IR** (ATR, cm^{-1}) 3389, 3113, 2879, 1688, 1498; **^{1}H NMR**

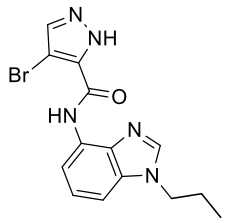
(400 MHz, DMSO-d_6) δ 13.31 (br s, 1H, NH), 9.66 (br s, 1H, C(O)NH), 8.25 (s, 1H, NCHN), 8.12 (d, *J* 8.0 Hz, 1H, ArH), 7.35 (d, *J* 8.0 Hz, 1H, ArH), 7.24 (app t, *J* 8.0 Hz, 1H, ArH), 4.23 (t, J 7.0 Hz, 2H, $NCH_2CH_2CH_3$), 2.26 (s, 3H, CH_3), 1.82 (app sextet, *J* 7.0 Hz, 2H, $NCH_2CH_2CH_3$), 0.85 (t, *J* 7.0 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.1 (d, $J_{C\text{-}F}$ 3.1 Hz), 146.8, 144.3, 143.1, 133.9 (d, $J_{C\text{-}F}$ 6.9 Hz), 131.4 (d, $J_{C\text{-}F}$ 3.9 Hz), 129.0, 126.4 (d, $J_{C\text{-}F}$ 25.1 Hz), 122.9, 109.4, 105.8, 45.9, 22.8, 10.9, 7.4 (d, $J_{C\text{-}F}$ 2.0 Hz); **^{19}F NMR** (376 MHz, DMSO-d_6) δ -176.05 (s, 1F, CF); **LRMS** (LCMS-ESI) *m/z* calc. for $C_{15}H_{16}FN_5O$ 301.1 found 302.1 $[M+H]^+$.

4-Chloro-*N*-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **40**



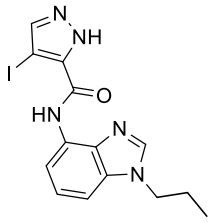
According to **General Procedure D**; acid **97a** (0.34 mmol, 50 mg), HATU (0.34 mmol, 129 mg), amine **98a** (0.29 mmol, 50 mg) and Hünigs base (0.58 mmol, 0.1 mL) were stirred in anhydrous acetonitrile (MeCN) (1.0 mL) at room temperature overnight. Upon completion, H_2O was added (0.5 mL) and a precipitate formed. The vial was then sonicated for 5 minutes and placed in the fridge for 1 hour. The resulting solids were filtered, and washed with H_2O (5 mL) to afford the *title compound* **40** as a pink solid (0.1 mmol, 35 mg, 34%). **MP** 207–209 °C; **IR** (ATR, cm^{-1}) 3389, 3134, 2874, 1682, 1532, 1496; **^{1}H NMR**, (400 MHz, DMSO-d_6) δ 13.85 (s, 1H, NH), 9.82 (br s, 1H, C(O)NH), 8.25 (s, 1H. NCHNR), 8.22 (br s, 1H, NCHCCl), 8.13 (d, *J* 8.0 Hz, 1H, ArH), 7.36 (d, *J* 8.0 Hz, 1H, ArH), 7.25 (app t, *J* 8.0 Hz, 1H, ArH), 4.23 (t, *J* 7.3 Hz, 2H, $NCH_2CH_2CH_3$), 1.83 (app sextet, *J* 7.3 Hz, 2H, $NCH_2CH_2CH_3$), 0.85 (t, *J* 7.3 Hz, 3H, CH_3); **^{13}C NMR**, (101 MHz, DMSO-d_6) δ 158.2, 143.2, 140.0, 133.9, 130.6, 130.0, 122.9, 109.5, 105.9, 104.1, 97.9, 45.9, 22.8, 10.9; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{14}{}^{35}ClN_5O$ 303.1 found 304.0 $[M+H]^+$ **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{15}{}^{35}ClN_5O$ 304.0960 found 304.0961 $[M+H]^+$.

4-Bromo-*N*-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **41** (**CI709**)



According to **General Procedure D**; acid **97b** (0.34 mmol, 65 mg), HATU (0.34 mmol, 129 mg), amine **98a** (0.29 mmol, 50 mg) and Hünigs base (0.58 mmol, 0.1 mL) were stirred in anhydrous acetonitrile (MeCN) (1.0 mL) at 40°C overnight. Further purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **41** as a white solid (0.1 mmol, 30 mg, 30%). **MP** 228 °C; **IR** (ATR, cm^{-1}) 2981, 2882, 1680, 1485, 1252; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.90 (s, 1H, NH), 9.86 (s, 1H, C(O)NH) , 8.26–8.24 (m, 2H, NCHN and NCHCBr), 8.13 (d, *J* 8.0 Hz, 1H, ArH), 7.37 (d, *J* 8.0 Hz, 1H, ArH), 7.25 (app t, *J* 8.0 Hz, 1H, ArH), 4.24 (t, J 7.0 Hz, 2H, $NCH_2CH_2CH_3$) 1.83 (sextet, *J* 7.0 Hz, 2H, $NCH_2CH_2CH_3$), 0.86 (t, J 7.0 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.3, 143.2, 141.2, 133.9, 134.0, 129.1, 122.9, 109.5, 106.9, 93.1, 45.9, 22.8, 10.9, 1 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{14}{}^{79}BrN_5O$ 347.0 found 348.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{15}{}^{79}BrN_5O$ 348.0454 found 348.0446 $[M+H]^+$.

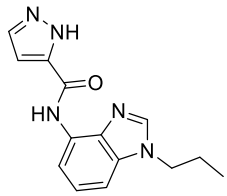
4-Iodo-*N*-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **42**



According to **General Procedure D**; acid **97c** (0.46 mmol, 109 mg), HATU (0.46 mmol, 175 mg), amine **98a** (0.2 mmol, 40 mg, 1.0 eq.) and Hünigs base (0.69 mmol, 0.1 mL) in anhydrous MeCN (2.5 mL) were stirred at room temperature overnight. Purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **42** as a white solid (0.1 mmol, 37 mg, 41%). **MP** decomp. >200°C; **IR** (ATR, cm^{-1}) 3383, 3106, 2875, 1679, 1527, 1495; **^{1}H NMR**, (400 MHz, DMSO-d_6) δ 13.88 (s, 1H, ArNH), 9.90 (s, 1H, C(O)NH), 8.24 (s, 1H, NCHN), 8.17 (s, 1H, NCHCl), 8.13 (d, *J* 7.8 Hz, 1H, ArH), 7.35 (d, *J* 7.8 Hz, 1H, ArH), 7.25 (app t, *J* 7.8 Hz, 1H, ArH), 4.23 (t, *J* 7.3 Hz, 1H, $NCH_2CH_2CH_3$), 1.83 (app sextet, *J* 7.3 Hz, 2H,

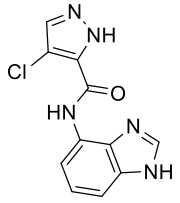
$NCH_2C\underline{H}_2CH_3$), 0.85 (t, *J* 7.3 Hz, 3H, $C\underline{H}_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.8, 143.6, 143.11, 137.5, 134.1, 133.9, 129.2, 122.9, 109.4, 105.8, 58.0, 45.9, 22.76, 10.9; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{14}IN_5O$ 395.0 found 395.9 $[M+H]^+$; **HRMS** (ESI, +ve) *m/z* calc. for $C_{14}H_{14}IN_5O$ 396.0316 found 396.0316 $[M+H]^+$.

N-(1-propyl-1H-benzo[d]imidazol-4-yl)-1H-pyrazole-5-carboxamide **43**



According to **General Procedure D**; acid **103** (0.57 mmol, 64 mg), HATU (0.57 mmol, 217 mg), amine **98a** (0.29 mmol, 50 mg) and Hünigs base (0.87 mmol, 0.15 mL) in anhydrous MeCN (3.2 mL) were stirred at room temperature overnight. Further purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **43** as a white solid (0.17 mmol, 46 mg, 59%). **MP** 218–219 °C; **IR** (ATR, cm^{-1}) 3387,3049, 2924, 1680, 1532, 1495; **^{1}H NMR** (500 MHz, DMSO-d_6) δ 13.51 (s, 1H, N$\underline{H}$), 9.88 (s, 1H, C(O)N$\underline{H}$), 8.25 (s, 1H, NC$\underline{H}$N), 8.16 (d, *J* 8.0 Hz, 1H, Ar*H*), 7.96–7.95 (m, 1H, Ar$\underline{H}$), 7.34 (d, *J* 8.0 Hz, 1H, Ar$\underline{H}$), 7.25 (app t, *J* 8.0 Hz, 1H, Ar$\underline{H}$) 6.83–6.82 (m, 1H, ArH), 4.23 (t, J 7.0 Hz, 2H, $NC\underline{H}_2CH_2CH_3$), 1.85, 1.83 (app sextet, *J* 7.0 Hz, 2H, $NCH_2C\underline{H}_2CH_3$), 0.85 (t, *J* 7.0 Hz. 3H, CH_3); **^{13}C NMR** (126 MHz, DMSO-d_6) δ 159.6, 146.2, 143.1, 133.9, 133.8, 131.1, 129.3, 122.9, 109.4, 105.6, 105.3, 45.9, 22.8, 10.9; **LRMS** (LCMS-ESI) m/z calc. for $C_{14}H_{15}N_5O$ 269.1 found 270.0 $[M+H]^+$.

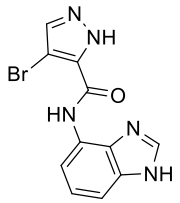
N-(1*H*-benzo[*d*]imidazol-4-yl)-4-chloro-1*H*-pyrazole-5-carboxamide **44**



According to **General Procedure D**; acid **97a** (0.62 mmol, 91 mg), HATU (0.62 mmol, 235 mg), amine **98e** (0.56 mmol, 75 mg) and Hünigs base (1.1 mmol, 0.2 mL) in anhydrous MeCN (5.6 mL) were stirred at 30 °C overnight. Further purification *via* flash column chromatography in 98:2 EtOAc/MeOH afforded the *title compound* **44** as a pale green solid (0.3 mmol, 65 mg, 44%). **MP** decomp. >190 °C; **IR** (ATR, cm^{-1}) 3370, 3225, 2918, 1681, 1540, 1488; **^{1}H NMR**, (400 MHz, DMSO-d_6) δ 13.83 (s, 1H, N$\underline{H}$), 12.62 (s, 1H, NCHN$\underline{H}$), 9.83 (s, 1H, C(O)N$\underline{H}$), 8.22 (br s, 2H, NC$\underline{H}$CCl/NC$\underline{H}$NH),

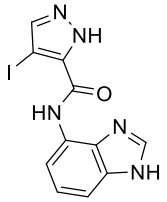
8.11 (d, *J* 7.9 Hz, 1H, ArH), 7.28 (d, *J* 7.9 Hz, 1H, ArH), 7.21 (app t, *J* 7.9 Hz, 1H, ArH); δ **^{13}C NMR**, (101 MHz, DMSO-d_6) δ 158.2, 141.1, 140.1, 133.5, 133.2, 130.5, 128.8, 122.9, 109.6, 109.3, 107.0; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{11}H_8{}^{35}ClN_5O$ 261.0 found 259.9 $[M-H]^-$ **HRMS** (ESI, +ve) m/z calc. for $C_{11}H_9{}^{35}ClN_5O$ 262.0496 found 262.0492 $[M+H]^+$.

N-(1*H*-benzo[*d*]imidazol-4-yl)-4-bromo-1*H*-pyrazole-5-carboxamide **45** (**CI639**)



According to **General Procedure D**; acid **97b** (1.9 mmol, 361 mg), HATU (1.9 mmol, 719 mg), amine **98e** (2.1 mmol, 277 mg) and Hünigs base (3.8 mmol, 0.7 mL) in MeCN (4 mL) were stirred at room temperature overnight. Further purification *via* flash column chromatography in 98:2 EtOAc/MeOH afforded the *title compound* **45** as a white solid (0.3 mmol, 75 mg, 13%). **MP** >250 °C; **IR** (ATR, cm^{-1}) 3383, 3198, 3144, 2934, 1701, 1673, 1537, 1415; **^{1}H NMR**, (500 MHz, DMSO-d_6) δ 13.89 (br s, 1H, NH), 12.62 (s, 1H, NCHNH), 9.87 (s, 1H, C(O)NH), 8.22 (s, 2H, NCHCBr and NCHNH), 8.11 (d, *J* 7.8 Hz, 1H, ArH), 7.28 (d, *J* 7.8 Hz, 1H, ArH), 7.20 (app t, *J* 7.8 Hz, 1H, ArH); **^{13}C NMR** (126 MHz, DMSO-d_6) δ 141.1, 133.5, 133.2, 132.9, 128.9, 122.9, 109.2, 107.0, 93.0, 2 C missing **LRMS** (LCMS-ESI) *m/z* calc. for $C_{11}H_8{}^{79}BrN_5O$ 305.0 found 307.9 $[M+2H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{11}H_9{}^{79}BrN_5O$ 305.9990 found 305.9992 $[M+H]^+$.

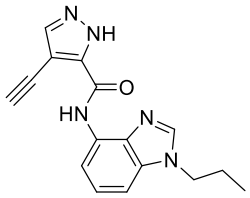
N-(1*H*-benzo[*d*]imidazol-4-yl)-4-iodo-1*H*-pyrazole-5-carboxamide **46**



According to **General Procedure D**; acid **97c** (0.75 mmol, 178 mg), HATU (0.75 mmol, 285 mg), amine **98e** (0.38 mmol, 50 mg) and Hünigs base (1.12 mmol, 0.2 mL) in anhydrous MeCN (4.2 mL) were stirred at room temperature overnight. Further purification *via* flash column chromatography in 92:8 CH_2Cl_2/MeOH afforded the *title compound* **46** as a brown solid (0.04 mmol, 15 mg, 11%). **MP** 244–246 °C; **IR** (ATR, cm^{-1}) 3379, 3126, 3107, 2923, 1726, 1673; **^{1}H NMR** (400 MHz, DMSO-d_6)

δ 13.88 (s, 1H, NH), 12.62 (s, 1H, NCHNH), 9.91 (s, 1H, C(O)NH), 8.22 (s, 1H, NCHN), 8.17 (s, 1H, NCHC(I)), 8.11 (d, *J* 8.0 Hz. 1H, ArH), 7.27 (d, *J* 8.0 Hz, 1H, Ar**^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.9, 141.1, 137.5, 133.5, 133.2, 129.0, 124.1, 122.9, 109.2, 106.9, 58.0; **LRMS** (LCMS-ESI) m/z calc. for $C_{11}H_8IN_5O$ 352.98 found 353.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{11}H_9IN_5O$ 353.9846 found 353.9848 $[M+H]^+$.

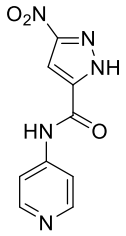
4-Ethynyl-*N*-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **49**



According to **General Procedure D**; 4-ethynyl-1H-pyrazole-3-carboxylic acid **97d** (0.17 mmol, 25 mg), HATU (0.17 mmol, 65 mg), amine **98a** (0.14 mmol, 25 mg) and Hünigs base (0.28 mmol, 50 µL) in MeCN (0.5 mL) were stirred at 40 °C overnight. Purification *via* flash column chromatography in 3:2 EtOAc: hexanes to 100% EtOAc afforded the *title compound* **49** as a white solid (0.07 mmol, 21 mg, 51%). **MP** 160–162 °C; **IR** (ATR, cm^{-1}) 3387, 3304, 3134, 2915, 2848, 2123, 1582, 1498; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 10.94 (br s, 1H, NNH), 10.06 (br s, 1H, C(O)NH), 8.36 (d, *J* 8.0 Hz, 1H, ArH), 7.84 (s, 1H, NCHN), 7.81 (s, 1H, NCHCC), 7.33 (app t, *J* 8.0 Hz, 1H, ArH),7.19 (d, *J* 8.0 Hz, 1H, ArH) 4.16 (t, *J* 7.3 Hz, 2H, $NCH_2CH_2CH_3$), 3.68 (br s, 1H, CCH), 1.94 (app sextet, *J* 7.3 Hz, 2H, $NCH_2CH_2CH_3$), 0.97 (t, *J* 7.3 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.3, 145.7, 143.2, 135.8, 133.9, 129.2, 122.9, 109.5, 105.9, 101.8, 83.6, 74.9, 45.91, 22.8, 11.0, 1 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{16}H_{15}N_5O$ 293.1 found 294.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{16}H_{16}N_5O$ 294.1349 found 294.1351 $[M+H]^+$.

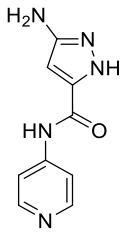
9.4.3 3′ HBD pyrazole analogues

3-Nitro-*N*-(pyridin-4-yl)-1*H*-pyrazole-5-carboxamide **50**



According to **General Procedure D**; 3-nitro-1*H*-pyrazole-5-carboxylic acid **116** (0.64 mmol, 100 mg), HATU (0.64 mmol, 243 mg), 4-aminopyridine **117** (0.53 mmol, 50 mg) and Hünigs base (1.06 mmol, 0.18 mL) in anhydrous DMF (1.8 mL) were stirred at room temperature overnight. Upon completion, H_2O was added (1 mL) and a white precipitate formed. The solids were filtered to provide the *title compound* **50** as a white solid (0.30 mmol, 70 mg, 57%). **MP** decomp. >240 °C; **IR** (ATR, cm^{-1}) 3132, 2455, 1717, 1590, 1508; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 10.75 (s, 1H, C(O)NH), 8.53–8.52 (m, 2H, 2×ArH), 7.87 (s, 1H, CHCNO$_2$), 7.76–7.75 (m, 2H, 2×ArH); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 156.9, 155.7, 150.4, 145.0, 139.0, 114.1, 102.8 ; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_7N_5O_3$ 233.1 found 232.1 $[M-H]^+$.

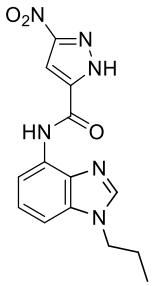
3-Amino-*N*-(pyridin-4-yl)-1*H*-pyrazole-5-carboxamide **51**



To a flame dried microwave vial purged with argon was added nitro amide **50** (0.30 mmol, 70 mg, 1.0 eq.) in a mixture of anhydrous 5:2 EtOAc/EtOH (5.3 mL). To the resulting solution was added 3 mol% Pd/C (0.04 mmol, 4 mg, 0.3 eq.). Hydrogen was introduced as described in **General Procedure B** and the reaction was allowed to stir at room temperature for 5 days. Upon completion, the reaction mixture was filtered through a plug of Celite®, washed with EtOH (3×10 mL) and concentrated *in vacuo.* Further purification *via* flash column chromatography 9:1 CH_2Cl_2/MeOH provided the *title compound* **51** as a green solid (0.06 mmol, 13 mg, 21%). **MP** 216–218 °C; **IR** (ATR, cm^{-1}) 3332, 3119, 2647, 2749, 1675, 1590; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 12.20 (s, 1H, NNH), 10.18 (s, 1H, C(O)NH), 8.42 (d, *J* 4.5 Hz, 2H, 2×ArH),

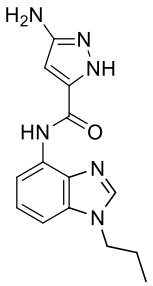
7.81 (d, J 4.5 Hz. 2H, 2×ArH), 5.75 (s, 1H, CHCNH$_2$), 5.22 (br s, 2H, NH$_2$); **^{13}C NMR** (101 MHz, DMSO-*d*$_6$) δ 161.9, 150.1, 149.6, 145.7, 113.8, 87.8, 1 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_9N_5O$ is 203.1 found 204.2 $[M+H]^+$.

3-Nitro-*N*-(3-propyl-3*H*-imidazo[4,5-*b*]pyridin-7-yl)-1*H*-pyrazole-5-carboxamide **52**



According to **General Procedure D**; 3-nitro-1*H*-pyrazole-5-carboxylic acid **116** (0.67 mmol, 107 mg), HATU (0.67 mmol, 243 mg), amine **98a** (0.57 mmol, 100 mg) and Hünigs base (1.14 mmol, 0.2 mL) in anhydrous DMF (1.9 mL) were stirred at room temperature for 2 days. Upon completion, H_2O was added (1 mL) and a white precipitate formed. The solids were filtered to provide the *title compound* **52** as a yellow solid (0.37 mmol, 116 mg, 65%). **MP** 198–200 °C; **IR** (ATR, cm^{-1}) 3183, 3126, 2971, 1676, 1546; **^{1}H NMR** (400 MHz, DMSO-*d*$_6$) δ 10.56 (s, 1H, C(O)NH), 8.31 (s, 1H, NCHN), 8.01 (s,1H, CHCNO$_2$), 7.70 (d, *J* 8.0 Hz, 1H, ArH) , 7.49 (d, *J* 8.0 Hz, 1H, ArH), 7.29 (app t, *J* 8.0 Hz, 1H, ArH) 4.25 (t, J 7.0 Hz, 2H, NCH$_2$CH$_2$CH$_3$),1.83 (app sextet, *J* 7.0 Hz, 2H, NCH$_2$CH$_2$CH$_3$), 0.85 (t, *J* 7.0 Hz, 3H, CH$_3$); **^{13}C NMR** (101 MHz, DMSO-*d*$_6$) δ 156.2, 155.6, 143.6, 139.7, 136.5, 134.6, 128.0, 122.4, 115.4, 107.8, 102.5, 45.9, 22.7, 10.9; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{14}N_7O_3$ 314.11 found 313.2 $[M-H]^+$.

3-Amino-*N*-(3-propyl-3*H*-imidazo[4,5-*b*]pyridin-7-yl)-1*H*-pyrazole-5-carboxamide **53**

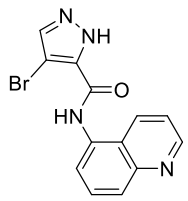


According to **General Procedure D**; 3-amino-1*H*-pyrazole-5-carboxylic acid **118** (0.21 mmol, 27 mg), HATU (0.21 mmol, 80 mg), amine **98a** (0.31 mmol, 55 mg) and Hünigs base (0.42 mmol, 70 µL) in anhydrous DMF (0.7 mL) were stirred at 50 ºC

overnight. Purification *via* flash column chromatography in 97:3 to 9:1 CH_2Cl_2: MeOH provided the crude material. The resulting solid was taken up in EtOAc (15 mL) and washed with H_2O (3×10 mL), brine (10 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo* to provide the *title compound* **53** as a white solid (0.04 mmol, 11 mg, 18%). **MP** decomp. >170 °C; **IR** (ATR, cm^{-1}) 3363, 3319, 3142, 2967, 2929, 1600, 1543; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.35 (s, 1H, C(O)NH), 8.26 (d, *J* 8.0 Hz, 1H, ArH), 7.91 (s, 1H, NCHN), 7.30 (app t, *J* 8.0 Hz, 1H, ArH), 7.14 (d, *J* 8.0 Hz, 1H, ArH), 6.18 (s, 1H, CHCNH$_2$), 4.13 (t, *J* 7.2 Hz, 2H, NCH$_2$CH$_2$CH$_3$) 1.92 (app sextet, *J* 7.2 Hz, 2H, NCH$_2$CH$_2$CH$_3$), 0.95 (t, *J* 7.2 Hz, 3H, CH$_3$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 159.2, 151.5, 142.0, 134.4, 134.1, 129.6, 124.0, 111.8, 105.6, 91.4, 47.2, 23.3, 11.4, 1 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{16}N_6O$ 284.1 found 285.3 $[M+H]^+$.

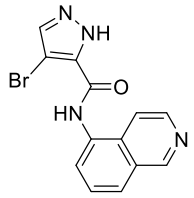
9.4.4 Alternative tail group heterocycle analogues

4-Bromo-*N*-(quinolin-5-yl)-1*H*-pyrazole-5-carboxamide **54**



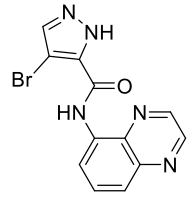
According to **General Procedure D**; acid **97b** (0.42 mmol, 79 mg), HATU (0.42 mmol, 160 mg), quinolin-5-amine **132a** (0.35 mmol, 50 mg) and Hünigs base (0.70 mmol, 0.12 mL) in MeCN (1.2 mL) were stirred at 40 °C overnight. Purification *via* flash column chromatography in 98:2 CH_2Cl_2: MeOH afforded the *title compound* **54** as a brown solid (0.08 mmol, 24 mg, 22%). **MP** decomp >238 °C; **IR** (ATR, cm^{-1}) 3406, 3105, 2888, 1693, 1532, 1496; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.86 (s, 1H, NNH), 10.43 (s, 1H, C(O)NH), 8.93 (s, 1H, NCHCBr), 8.34–8.22 (m, 2H, 2×ArH), 7.95–7.57 (m, 4H, 4×ArH); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 160.3, 150.6, 148.1, 142.1, 133.6, 132.2, 131.8, 129.0, 126.9, 123.9, 123.2, 121.1, 93.2; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{13}H_9{}^{79}BrN_4O$ 316.0 found 316.9 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{13}H_{10}{}^{79}BrN_4O$ 317.0033 found 317.0035 $[M+H]^+$.

4-Bromo-*N*-(isoquinolin-5-yl)-1*H*-pyrazole-5-carboxamide **55**



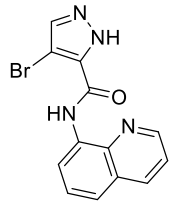
According to **General Procedure D**; acid **97b** (0.42 mmol, 79 mg.), HATU (0.42 mmol, 160 mg) isoquinolin-5-amine **132b** (0.35 mmol, 50 mg) and Hünigs base (0.70 mmol, 0.12 mL) in MeCN (1.2 mL) were stirred at 40 °C overnight. Purification *via* flash column chromatography in 95:5 CH_2Cl_2: MeOH afforded the *title compound* **55** as a pink solid (0.01 mmol, 10 mg, 4%). **MP** decomp >220 °C; **IR** (ATR, cm^{-1}) 3395, 3175, 2958, 2921, 1680, 1521; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.85 (s, 1H, NNH), 10.36 (s, 1H, C(O)NH), 9.35 (s, 1H, NCHCBr), 8.56–8.53 (m, 1H, ArH), 8.21–8.20 (m, 1H, ArH), 8.03 (d, *J* 8.0 Hz, 1H, ArH), 7.98–7.96 (m, 1H, ArH), 7.82–7.78 (m, 1H, ArH), 7.72 (app t, *J* 8.0 Hz, 1H, ArH); **^{13}C NMR** (151 MHz, DMSO-d_6) δ 152.5, 142.8, 132.5, 132.2, 130.8, 128.7, 127.2, 126.4, 125.3, 115.9, 93.2, 2 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{13}H_9{}^{79}BrN_4O$ 316.0 found 316.9 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{13}H_{10}{}^{79}BrN_4O$ 317.0033 found 317.0036 $[M+H]^+$.

4-Bromo-*N*-(quinoxalin-5-yl)-1*H*-pyrazole-5-carboxamide **56**



According to **General Procedure D**; acid **97b** (0.35 mmol, 67 mg), HATU (0.35.mmol, 133 mg), quinazoline **132c** (0.29 mmol, 42 mg) and Hünigs base (0.14 mmol, 100 µL) in DMF (1.0 mL) were stirred at 45 °C overnight. Purification *via* flash column chromatography in 7:3 petroleum ether 40–60/EtOAc afforded the *title compound* **31** as a brown solid (0.06 mmol, 19 mg, 21%). **MP** 267–269 °C; **IR** (ATR, cm^{-1}) 3344, 3101, 2958, 1677, 1526, 1489; ^{1}H NMR (400 MHz, DMSO-d_6) δ 13.94 (br s, 1H, NH), 10.97 (br s, 1H, C(O)NH), 9.09 (d, *J* 1.8 Hz, 1H, ArH), 9.00 (d, *J* 1.8 Hz, 1H, ArH), 8.80 (dd, *J* 8.0, 1.0 Hz, 1H, ArH), 8.26 (s, 1H, NCHCBr), 7.89 (app t, *J* 8.0 Hz, 1H, ArH), 7.82 (dd, *J* 8.0, 1.0 Hz, 1H, ArH); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.6, 146.5, 143.8, 142.4, 141.2, 134.2, 133.0, 132.6, 130.9, 122.8, 115.9, 93.3; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{12}H_8{}^{79}BrN_5O$ 316.9 found 317.9 $[M+H]^+$.

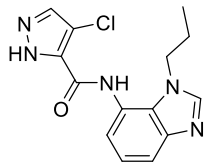
4-Bromo-*N*-(quinolin-8-yl)-1*H*-pyrazole-5-carboxamide **57**



According to **General Procedure D**; acid **97b** (0.42 mmol, 80 mg), HATU (0.42 mmol, 160 mg), quinolin-8-amine **132d** (0.35 mmol, 50 mg) and Hünigs base (0.70 mmol, 120 µL) in DMF (1.2 mL) were stirred at 45 °C overnight. The following day a second portion of reagents were added: acid **97b** (0.35 mmol, 67 mg) and HATU (0.35 mmol, 133 mg) and the resulting mixture was stirred overnight at 45 °C. Purification *via* flash column chromatography in 9:1 EtOAc/hexanes afforded the *title compound* **57** as an off white solid (0.21 mmol, 68 mg, 61%). **MP** decomp >200 °C; **IR** (ATR, cm^{-1}) 3296, 3181, 2952, 2922, 1643, 1539; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.90 (br s, 1H, NNH) 11.20 (br s, 1H, C(O)NH), 8.98–8.96 (m, 1H, ArH), 8.78–8.76 (m, 1H, ArH), 8.46–8.44 (m, 1H, ArH), 8.25 (s, 1H, NCHCBr), 7.72–7.62 (m, 3H, 3×ArH); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.6, 149.1, 141.4, 137.8, 136.7, 133.9, 132.9, 127.8, 127.1, 122.3, 121.8, 115.4, 93.2; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{13}H_9{}^{79}BrN_5O$ 316.0 found 316.9 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{13}H_{10}{}^{79}BrN_5O$ 317.0033 found 317.0037 $[M+H]^+$.

9.4.5 *N*-1 Tail group derivatives

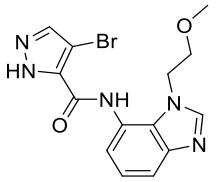
4-Chloro-*N*-(1-propyl-1*H*-benzo[*d*]imidazol-7-yl)-1*H*-pyrazole-5-carboxamide **58**



According to **General Procedure D**; acid **97a** (0.31 mmol, 45 mg), HATU (0.31 mmol, 118 mg), amine **123a** (0.29 mmol, 50 mg) and Hünigs base (0.6 mmol, 0.1 mL) in anhydrous MeCN (0.6 mL) were stirred at room temperature overnight. Further purification *via* flash column chromatography in 97:3 CH_2Cl_2/MeOH afforded the *title compound* **58** as a white solid (0.1 mmol, 14 mg, 16%). **MP** 176–178 °C; **IR** (ATR, cm^{-1}) 3638, 3203, 2967, 1653, 1534, 1468; **^{1}H NMR**, (400 MHz, DMSO-d_6) δ 13.77 (br s, 1H, NH), 10.23 (s, 1H, C(O)NH), 8.75 (s, 1H, NCHNR), 8.17 (s, 1H, NCHCCl), 7.59–7.57 (m, 1H, ArH), 7.22–7.13 (m, 2H, 2×ArH), 4.25 (t, *J* 7.2 Hz, 2H,

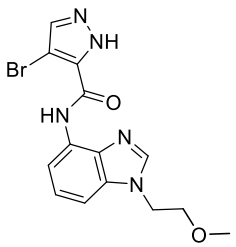
$NCH_2CH_2CH_3$), 1.66 (app sextet, *J* 7.2 Hz, 2H, $NCH_2CH_2CH_3$), 0.68 (t, *J* 7.2 Hz, 3H, CH_3); **^{13}C NMR**, (101 MHz, DMSO-d_6) δ 160.6, 145.6, 145.3, 140.7, 129.8, 129.4, 122.5, 121.3, 118.1, 109.4, 53.4, 47.1, 24.3, 10.7; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{14}{}^{35}ClN_5O$ 303.1 found 304.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{15}{}^{35}ClN_5O$ 304.0960 found 364.0965 $[M+H]^+$.

4-Bromo-*N*-(1-(2-methoxyethyl)-1*H*-benzo[d]imidazol-7-yl)-1*H*-pyrazole-5-carboxamide **59**



According to **General Procedure D**; carboxylic acid **97b** (0.26 mmol, 50 mg), HATU (0.26 mmol, 100 mg), amine **123b** (0.32 mmol, 70 mg) and Hünigs base (0.52 mmol, 0.1 mL, 2.0 eq) in MeCN (2 mL) were stirred at room temperature overnight. Further purification *via* flash column chromatography in 98:2 CH_2Cl_2/MeOH afforded the *title compound* **59** as an off-white solid (0.1 mmol, 26 mg, 32%). **MP** decomp. >196 °C; **IR** (ATR, cm^{-1}) 3201, 3093, 2922, 1669, 1593, 1495; **^{1}H NMR** (500 MHz, DMSO-d_6), δ 13.79 (s, 1H, NH), 10.30 (s, 1H, C(O)NH), 8.18 (s, 1H, NCHCBr), 8.12 (s, 1H, NCHNR), 7.57–7.55 (m, 1H, ArH), 7.23–7.18 (m, 2H, ArH), 4.48 (t, *J* 5.0 Hz, 2H, NCH_2CH_2O), 3.60 (t, *J* 5.0 Hz, 2H, NCH_2CH_2O), 3.17 (s, 3H, OCH_3); **^{13}C NMR**, (125 MHz, DMSO-d_6) δ 160.5, 145.6, 145.4, 142.0, 132.1, 129.2, 122.1, 121.8, 121.3, 117.8, 93.1, 71.1, 58.1, 45.3; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{14}{}^{79}BrN_5O_2$ 363.0 found 363.9 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{15}{}^{79}BrN_5O_2$ 364.0404 found 364.0404 $[M+H]^+$.

4-Bromo-*N*-(1-(2-methoxyethyl)-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **60**



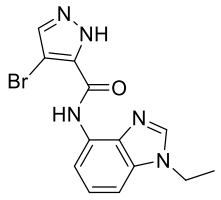
According to **General Procedure D**; carboxylic acid **97b** (0.26 mmol, 50 mg), HATU (0.26 mmol, 100 mg), amine **98b** (0.3 mmol, 70 mg) and Hünigs base (0.5 mmol,

0.1 mL) in MeCN (2 mL) were stirred at room temperature overnight. Further purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **60** as a white solid (0.1 mmol, 38 mg, 40%). **MP** 138–140 °C; **IR** (ATR, cm^{-1}) 3118, 2919, 2850, 1672, 1538, 1469; **^{1}H NMR**, (500 MHz, DMSO-d_6) δ 13.89 (s, 1H, NH), 9.84 (s, 1H, C(O)NH), 8.22 (s, 1H, NCHNR), 8.19 (s, 1H, NCHCBr), 8.13 (d, *J* 7.5 Hz, 1H, ArH), 7.37 (d, *J* 7.5 Hz, 1H, ArH), 7.25 (app t, *J* 7.5 Hz, 1H, ArH), 4.41 (t, *J* 5.0 Hz, 2H, NCH_2CH_2O), 3.69 (t, *J* 5.0 Hz, 2H, NCH_2CH_2O), 3.23 (s, 3H, OCH_3); **^{13}C NMR**, (125 MHz, DMSO-d_6) δ 158.4 143.5, 139.1, 134.0, 133.8, 132.9, 129.1, 124.1, 122.9, 109.5, 93.1, 70.2, 58.1, 44.3; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{14}{}^{79}BrN_5O_2$ 363.0 found 363.9 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{15}{}^{79}BrN_5O_2$ 364.0404 found 364.0405 $[M+H]^+$.

9.4.6 *N*-3 Tail group derivatives

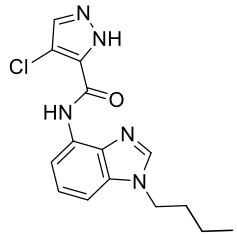
9.4.6.1 Aliphatic analogues

4-Bromo-*N*-(1-ethyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **61**



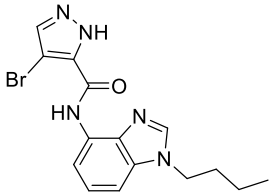
According to **General Procedure D**; acid **97b** (0.20 mmol, 38 mg), HATU (0.20 mmol, 76 mg) amine (0.17 mmol, 27 mg) and Hünigs base (0.34 mmol, 60 µL) in MeCN (0.6 mL) were stirred at 45 °C overnight. The following day a second portion of reagents were added, acid **98o** (0.17 mmol, 32 mg) and HATU (0.17 mmol, 35 mg) and the resulting mixture was stirred overnight at 45 °C. Purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **61** as a white solid (0.03 mmol, 10 mg, 18%). **MP** decomp. >220 °C; **IR** (ATR, cm^{-1}) 3388, 3119, 2923, 1682, 1501; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 9.85 (s, 1H, C(O)NH), 8.26 (s, 1H, NCHN), 8.21 (s, 1H, NCHCBr), 8.13 (d, *J* 8.0 Hz, 1H, ArH), 7.36 (d, *J* 8.0 Hz, 1H, ArH), 7.25 (app t, *J* 8.0 Hz, 1H, ArH), 4.30 (q, *J* 7.2 Hz, 2H, CH_2),1.43 (t, J 7.2 Hz, 3H, CH_3); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.4, 142.6, 134.0, 133.6, 133.0, 129.1, 122.9, 109.5, 105.8, 93.1, 15.28, 2 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{13}H_{12}{}^{79}BrN_5O$ 333.0 found 333.9 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{13}H_{13}{}^{79}BrN_5O$ 334.0298 found 334.0303 $[M+H]^+$.

N-(1-butyl-1*H*-benzo[*d*]imidazol-4-yl)-4-chloro-1*H*-pyrazole-5-carboxamide **62**



According to **General Procedure D**; acid **97a** (0.32 mmol, 47 mg), HATU (0.32 mmol, 122 mg), amine **98d** (0.26 mmol, 50 mg) and Hünigs base (0.5 mmol, 0.1 mL) in anhydrous MeCN (0.6 mL) were stirred at room temperature overnight. Further purification *via* flash column chromatography in 9:1:0.1 EtOAc/hexanes/MeOH afforded the *title compound* **62** as a brown solid (0.1 mmol, 40 mg, 50%). **MP** 78–80 °C; **IR** (ATR, cm^{-1}) 3637, 3137, 2918, 1668, 1627, 1582 **^{1}H NMR**, (400 MHz, DMSO-d_6) δ 13.84 (s, 1H, NH), 9.82 (s, 1H, C(O)NH), 8.25 (s, 1H, NCHNR), 8.22 (s, 1H, NCHCCl), 8.13 (d, *J* 8.0 Hz, 1H, ArH), 7.35 (d, *J* 8.0 Hz, 1H, ArH), 7.25 (app t, *J* 8.0 Hz, 1H, ArH), 4.28–4.25 (m, 2H, $NCH_2CH_2CH_2$), 1.82–1.75 (m, 2H, $NCH_2CH_2CH_2$), 1.30–1.21 (m, 2H, $NCH_2CH_2CH_2CH_3$), 0.91–0.87 (m, 3H, CH_3); **^{13}C NMR**, (101 MHz, DMSO-d_6) δ 162.3, 158.2, 143.1, 140.0, 133.9 130.6, 129.1, 122.9, 109.6, 109.5, 105.9, 44.1, 31.5, 19.3, 13.4; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{15}H_{16}{}^{35}ClN_5O$ 317.1 found 318.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{15}H_{17}{}^{35}ClN_5O$ 318.1116 found 318.1117 $[M+H]^+$.

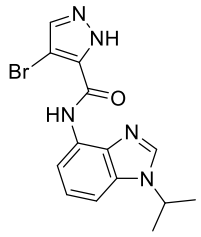
4-Bromo-*N*-(1-butyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **63**



To a microwave vial was added amide **45** (0.2 mmol, 65 mg, 1.0 eq.) and $KHCO_3$ (0.2 mmol, 19 mg, 0.9 eq.) in DMF (0.8 mL, 0.3 M). The resulting mixture was stirred at room temperature for 5 minutes before the addition of 1-bromobutane (0.2 mmol, 20 µL, 1.0 eq.). The reaction mixture was then stirred at room temperature overnight. Upon completion, the mixture was diluted with EtOAc (10 mL) and washed with water (3×10 mL) and brine (10 mL). The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography in 4:1 EtOAc/petroleum ether 40–60 with 0.1 % v/v Et_3N, afforded the *title compound* **63** as a white solid (0.1 mmol, 20 mg, 26%). **MP** 218–220 °C; **IR** (ATR, cm^{-1}) 3582,

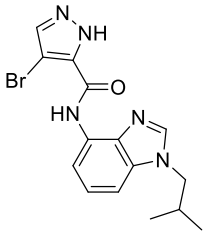
3360,2957, 2931, 1668, 1532; **^{1}H NMR**, (500 MHz, DMSO-*d_6*) δ 12.63 (s, 1H, NNH), 9.75 (s, 1H, C(O)NH), 8.24 (s, 2H, NCHCBr and NCHNR), 8.11 (d, *J* 7.9 Hz, 1H, ArH), 7.28 (d, *J* 7.9 Hz, 1H, ArH), 7.20 (app t, *J* 7.9 Hz, 1H, ArH), 4.25 (t, *J* 7.4 Hz, 2H, $NCH_2CH_2CH_2$), 1.86 (app quint, *J* 7.4 Hz, 2H, $NCH_2CH_2CH_2$), 1.33–1.26 (m, 2H, $NCH_2CH_2CH_2CH_3$), 0.92 (t, *J* 7.4 Hz, 3H, CH_3); **^{13}C NMR**, (100 MHz, DMSO-*d_6*) δ 158.0, 141.1, 140.8, 134.1, 133.4, 133.2, 128.8, 122.9, 109.3, 107.0, 92.8, 52.4, 31.5, 19.1, 13.4; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{15}H_{16}{}^{79}BrN_5O$ 361.1 found 362.0 $[M+H]^+$ **HRMS** (ESI, +ve) m/z calc. for $C_{15}H_{17}{}^{79}BrN_5O$ 362.0611 found 362.0613 $[M+H]^+$.

4-Bromo-*N*-(1-isopropyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **64**



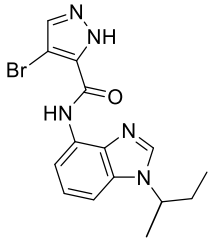
According to **General Procedure D**; acid **97b** (0.29 mmol, 55 mg), HATU (0.29 mmol, 110 mg), amine **98q** (0.24 mmol, 42 mg) and Hünigs base (0.48 mmol, 80 µL) in MeCN (0.8 mL) were stirred at 45 °C overnight. Purification *via* flash column chromatography 9:1 EtOAc/petroleum ether 40–60 afforded the *title compound* **64** as a white solid (0.06 mmol, 21 mg, 25%). **MP** 186–188 °C; **IR** (ATR, cm^{-1}) 3376, 3116, 2868, 1688, 1530, 1487; **^{1}H NMR** (400 MHz, DMSO-*d_6*) δ 13.90 (br s, 1H, NNH), 9.86 (br s, 1H, C(O)NH), 8.34 (s, 1H, NCHN), 8.22 (s, 1H, NCHCBr), 8.13 (d, *J* 8.0 Hz, 1H, ArH), 7.39 (d, *J* 8.0 Hz, 1H, ArH), 7.24 (app t, *J* 8.0 Hz, 1H, ArH), 4.77 (septet, J 6.7 Hz, 1H. $NCH(CH_3)_2$), 1.55 (d, J 6.7 Hz, 6H, 2×CH_3); **^{13}C NMR** (101 MHz, DMSO-*d_6*) δ 158.4, 141.3, 140.7, 134.1, 133.1, 132.9, 129.2, 122.8, 109.5, 106.3, 93.1, 47.5, 22.2; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{14}{}^{79}BrN_5O$ 347.0 found 348.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{15}{}^{79}BrN_5O$ 348.0454 found 348.0459 $[M+H]^+$.

4-Bromo-*N*-(1-isobutyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **65**



According to **General Procedure D**; acid **97b** (0.25 mmol, 48 mg), HATU (0.25 mmol, 95 mg), amine **98h** (0.21 mmol, 40 mg) and Hünigs base (0.42 mmol, 70 µL) in MeCN (0.7 mL) were stirred at room temperature overnight. Further purification *via* flash column chromatography in 4:1 EtOAc: hexanes afforded the *title compound* **65** as a white solid (0.1 mmol, 35 mg, 46%). **MP** 183–185 °C; **IR** (ATR, cm^{-1}) 3639, 3388, 2903, 2869, 1685, 1499; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.90 (br s, 1H, NH), 9.86 (br s, 1H, C(O)NH),, 8.24 (s, 1H, NCHN) , 8.22 (s, 1H, NCHCBr), 8.13 (d, *J* 8.0 Hz, 1H, ArH), 7.37 (d, *J* 8.0 Hz, 1H, ArH), 7.25 (app t, *J* 8.0 Hz, 1H, ArH), 4.09 (d, *J* 7.0 Hz, 2H, NCH_2*i*Pr) 2.17 (app septet, *J* 7.0 Hz, 1H, $CH_2CH(CH_3)_2$), 0.88 (d, *J* 7.0 Hz, 6H, 2×CH_3); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.4, 143.4, 141.3, 134.1, 133.8, 132.9, 129.1, 122.9, 109.5, 106.1, 93.1, 51.4, 28.7, 19.7; **LRMS** (LCMS-ESI) *m/z* calc.for $C_{15}H_{16}{}^{79}BrN_5O$ 361.1 found 362.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{15}H_{17}{}^{79}BrN_5O$ 362.0611 found 362.0614 $[M+H]^+$.

4-Bromo-*N*-(1-(*sec*-butyl)-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **66**

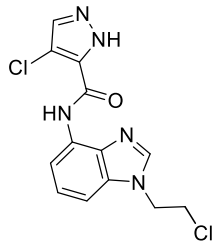


According to **General Procedure D**; acid **97b** (0.24 mmol, 46 mg), HATU (0.24 mmol, 91 mg), amine **98r** (0.20 mmol, 38 mg) and Hünigs base (0.40 mmol, 70 µL) in MeCN (0.6 mL) were stirred at 45 °C overnight. Purification *via* flash column chromatography in 4:1 EtOAc/petroleum ether afforded the *title compound* **66** as a brown solid (0.10 mmol, 10 mg, 14%).**MP** 86–88 °C; **IR** (ATR, cm^{-1}) 3413, 2919, 1693, 1489; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.89 (s, 1H, NNH), 9.87 (s, 1H, C(O)NH), 8.33 (s, 1H, NCHN), 8.23 (s, 1H, NCHCBr), 8.12 (d, *J* 8.0 Hz, 1H, ArH), 7.39 (d, *J* 8.0 Hz, 1H, ArH), 7.23 (app t, *J* 8.0 Hz. 1H, ArH), 4.54 (app q, *J* 7.0 Hz, 1H, $CH_3CHCH_2CH_3$), 1.96—1.85 (m, 2H, $CH(H)CH_3$), 1.55 (d, *J* 7.0 Hz, 3H, $CHCH_3$) 0.75 (t, J 7.3 Hz, 3H,

CH(H)C$\underline{H_3}$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.4, 141.4, 141.3, 134.0, 133.4, 132.9, 129.2, 122.8, 109.5, 106.3, 93.1, 53.2, 28.8, 20.3, 10.5; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{15}H_{16}{}^{79}BrN_5O$ 361.1 found 361.9 $[M+H]^+$.

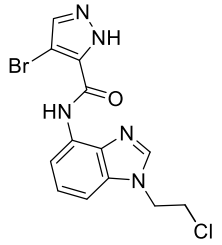
9.4.6.2 Electronegative analogues

4-Chloro-*N*-(1-(2-chloroethyl)-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **67**



According to **General Procedure D**; acid **97a** (0.36 mmol, 53 mg), HATU (0.4 mmol, 137 mg), amine **98g** (0.2 mmol, 35 mg) and Hünigs base (0.5 mmol, 0.1 mL) in anhydrous MeCN (2.0 mL) were stirred at room temperature overnight. Purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **67** as a white solid (0.1 mmol, 19 mg, 33%). **MP** 182–184°C; **IR** (ATR, cm^{-1}) 3639, 3386, 2973, 1690, 1532, 1421; **^{1}H NMR**, (400 MHz, DMSO-d_6) δ 13.84 (s, 1H, N$\underline{H}$), 9.82 (s, 1H, C(O)N$\underline{H}$), 8.29 (s, 1H, NC$\underline{H}$NR), 8.23 (s, 1H, NC$\underline{H}$CCl), 8.15 (d, *J* 8.0Hz, 1H, Ar$\underline{H}$), 7.43 (d, *J* 8.0 Hz, 1H, Ar$\underline{H}$), 7.27 (app t, *J* 8.0 Hz, 1H, Ar$\underline{H}$), 4.66 (t, *J* 5.7 Hz, 2H, NC$\underline{H_2}$$CH_2$Cl), 4.06 (t, *J* 5.7 Hz, 2H, $NCH_2C\underline{H_2}$Cl); **^{13}C NMR**, (100 MHz, DMSO-d_6) δ 158.2, 143.6, 140.0, 133.8, 133.7, 130.6, 129.1, 123.3, 109.8, 109.6, 107.0, 59.7, 43.5; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{13}H_{11}{}^{35}ClN_5O$ 323.0 found 324.0; **HRMS** (ESI, +ve) *m/z* calc. for $C_{13}H_{12}{}^{35}Cl_2N_5O$ 324.0413 found 324.0417 $[M+H]^+$.

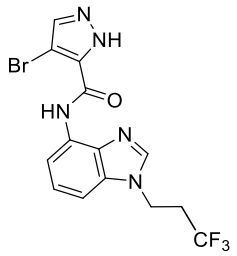
4-Bromo-*N*-(1-(2-chloroethyl)-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **68**



According to **General Procedure D**; acid **97b** (0.25 mmol, 48 mg), HATU (0.21 mmol, 160 mg), amine **98g** (0.21 mmol, 44 mg) and Hünigs base (0.42 mmol, 0.1 mL) in

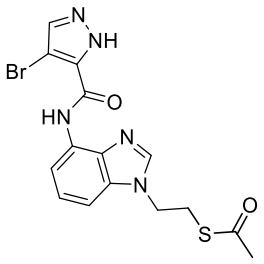
MeCN (2.4 mL) were stirred at room temperature overnight. Further purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **68** as a white solid (0.1 mmol, 31 mg, 40%). **MP** 178–180 °C; **IR** (ATR, cm^{-1}) 3244, 3164, 2966, 1626, 1555; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.89 (br s, 1H, NH), 9.86 (s, 1H, C(O)NH), 8.29 (s, 1H, NCHNR) 8.22 (s, 1H, NCHCBr), 8.14 (d, *J* 8.0 Hz, 1H, ArH), 7.42 (d, *J* 8.0 Hz, 1H, ArH), 7.27 (app t, *J* 8.0 Hz, 1H, ArH), 4.66 (t, *J* 5.7 Hz, 2H, NCH_2CH_2Cl), 4.06 (t, *J* 5.7 Hz, 2H, NCH_2CH_2Cl);**^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.4, 143.5, 141.3, 135.5, 133.7, 132.8, 129.1, 123.1, 109.8, 105.9, 93.1, 46.0, 43.5; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{13}H_{11}{}^{79}Br^{35}ClN_5O$ 367.0 found 367.8 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{13}H_{12}{}^{79}Br^{35}ClN_5O$ 367.9908 found 367.9912 $[M+H]^+$.

4-Bromo-*N*-(1-(3,3,3-trifluoropropyl)-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **69**



According to **General Procedure D**; acid **97b** (0.21 mmol, 40 mg), HATU (0.21 mmol, 80 mg), amine **98j** (0.18 mmol, 40 mg) and Hünigs base (0.36 mmol, 60 µL) in MeCN (0.6 mL) were stirred at room temperature overnight. Purification *via* flash column chromatography in 4:1 EtOAc/hexanes afforded the *title compound* **69** as a white solid (0.06 mmol, 23 mg, 32%). **MP** 180–182 °C; **IR** (ATR, cm^{-1}) 3383, 3180, 2960, 2916, 1688, 1536, 1499; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.92 (s, 1H, NNH), 9.84 (s, 1H, C(O)NH), 8.29 (s, 1H, NCHN), 8.21 (s, 1H, NCHCBr), 8.14 (d, *J* 8.0 Hz, 1H, ArH), 7.39 (d, *J* 8.0 Hz, 1H, ArH),7.28 (app t, *J* 8.0 Hz, 1H, ArH), 4.58–4.56 (m, 2H, $NCH_2CH_2CF_3$), 2.97–2.89 (m, 2H, $NCH_2CH_2CF_3$);**^{13}C NMR** (151 MHz, DMSO-d_6) δ 158.4, 143.2, 141.2, 133.7, 133.5, 132.9, 129.2, 126.4 (q, J_{C-F} 277.8 Hz), 123.2, 109.8, 105.7, 93.1, 37.8–37.7 (m), 32.7 (q, J_{C-F} 27.6 Hz); **^{19}F NMR** (376 MHz, DMSO-d_6) δ -63.74 (t, J 11.2 Hz, 3F, CF_3); **LRMS** (LCMS-ESI) *m/z* calc. $C_{14}H_{11}{}^{79}BrF_3N_5O$ for 401.1 found 401.9 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{12}{}^{79}BrF_3N_5O$ 402.0172 found 402.0173 $[M+H]^+$.

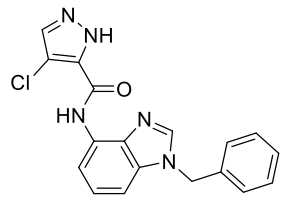
S-(2-(4-(4-bromo-1*H*-pyrazole-5-carboxamido)-1*H*-benzo[*d*]imidazol-1-yl)ethyl) ethanethioate **70**



To a microwave vial was added chloro derivative **68** (0.05 mmol, 20 mg, 1.0 eq.) and DMF (0.5 mL, 0.1 M). To the solution was added potassium thioacetate (0.3 mmol, 37 mg, 6.0 eq.) and the resulting mixture was heated to 80 °C and allowed to stir at this temperature for 5 days. Upon completion the reaction mixture was diluted with EtOAc (10 mL) and washed with water (3×10 mL) and brine (10 mL). The organics were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **70** (0.01 mmol, 5 mg, 23%). **MP** decomp. >220 °C; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.90 (s, 1, NH), 9.84 (s, 1H, C(O)NH), 8.23 (s, 2H, NCHNR and NCHCBr), 8.14 (d, *J* 8.0 Hz, 1H, ArH), 7.42 (d, *J* 8.0 Hz, 1H, ArH), 7.28 (app t, *J* 8.0 Hz, 1H, ArH), 4.45– 4.42 (m, 2H, CH_2), 3.33–3.31 (m, 2H, CH_2), 2.33 (s, 3H, CH_3); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 194.8, 158.4, 143.2, 141.3, 133.8, 133.7, 132.9, 129.1, 123.1, 109.7, 105.8, 93.1, 43.7, 30.5, 28.5; **LRMS** (LCMS-ESI) m/z calc. $C_{15}H_{14}{}^{79}BrN_5O_2S$ 407.1 found 405.8 $[M-H]^-$; **HRMS** (ESI, +ve) m/z calc. for $C_{15}H_{15}{}^{79}BrN_5O_2S$ 408.0130 found 408.0127 $[M+H]^+$.

9.4.6.3 Aromatic analogues

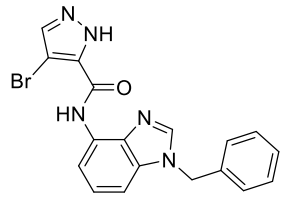
N-(1-benzyl-1*H*-benzo[*d*]imidazol-4-yl)-4-chloro-1*H*-pyrazole-5-carboxamide **71**



According to **General Procedure D**; acid **97a** (0.24 mmol, 35 mg), HATU (0.24 mmol, 91 mg), amine **98c** (0.22 mmol, 50 mg) and Hünigs base (0.44 mmol, 0.1 mL) in MeCN (1.0 mL) were stirred at room temperature overnight. Further purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **71** as a

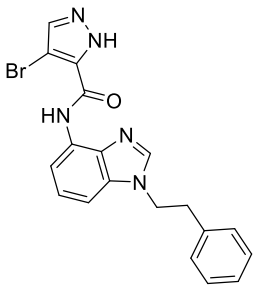
white solid (0.02 mmol, 8 mg, 10%). **MP** (decomp. >150°C; **IR** (ATR, cm^{-1}) 3372, 3134, 2961, 1688, 1539, 1498; **^{1}H NMR**, (500 MHz, Acetone-d_6) δ 12.87 (br s, 1H, NH), 9.96 (s, 1H, C(O)NH), 8.32–8.30 (m, 1H, ArH), 8.25 (s, 1H, NCHNR), 8.06 (br s, 1H, NCHCCl), 7.36–7.29 (m, 5H, ArH), 7.24–7.22 (m, 2H, ArH), 5.57 (s, 2H, CH_2); **^{13}C NMR** (125 MHz, Acetone-d_6) δ 159.2, 143.7, 137.8, 135.6, 135.1, 131.0, 129.7, 128.8, 128.3, 128.2, 124.2, 110.9, 106.5, 49.3, 2 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{18}H_{14}{}^{35}ClN_5O$ 351.1 found 352.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{18}H_{15}{}^{35}ClN_5O$ 352.0960 found 352.0961 $[M+H]^+$.

N-(1-benzyl-1*H*-benzo[*d*]imidazol-4-yl)-4-bromo-1*H*-pyrazole-5-carboxamide **72**



According to **General Procedure D**; acid **97b** (0.19 mmol, 36 mg), HATU (0.19 mmol, 72 mg, 1.2 eq.), amine **98c** (0.16 mmol, 35 mg) and Hünigs base (0.32 mmol, 60 µL) in MeCN (0.5 mL) were stirred at 40 °C overnight. Purification *via* flash column chromatography in 9:1 EtOAc/hexanes afforded the *title compound* **72** as a white solid (0.11 mmol, 43 mg, 69%). **MP** decomp. >214 °C; **IR** (ATR, cm^{-1}) 3640, 3348, 3218, 2963, 1667, 1557, 1418; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.92 (br s, 1H, NNH), 9.87 (br s, 1H, C(O)NH), 8.45 (s, 1H, NCHN), 8.24 (s, 1H, NCHCBr), 8.11 (d, *J* 8.0 Hz, 1H, ArH), 7.35–7.25 (m, 6H, 6×ArH), 7.20 (app t, *J* 8.0 Hz, 1H, ArH), 5.53 (s, 2H, CH_2); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.4, 143.4, 141.2, 136.9, 134.0, 133.7, 132.9, 129.2, 128.8, 127.8, 127.5, 123.2, 109.7, 106.2, 93.1, 47.9; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{18}H_{14}{}^{81}BrN_5O$ 395.0 found 394.0 $[M-H]^-$.

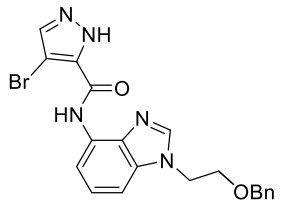
4-Bromo-*N*-(1-phenethyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **73**



According to **General Procedure D**; acid **97b** (0.23 mmol, 44 mg), HATU (0.2 mmol, 87 mg), amine **98f** (0.2 mmol, 45 mg) and Hünigs base (0.4 mmol, 0.1 mL) in MeCN

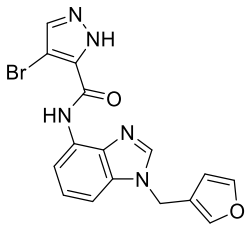
(2.0 mL) were stirred at 40 °C overnight. Purification *via* flash column chromatography in 4:1 EtOAc/hexanes afforded the *title compound* **73** as a white solid (0.1 mmol, 20 mg, 26%). **MP** 157–159°C; **IR** (ATR, cm^{-1}) 3377, 3106, 2923, 1685, 1538, 1493; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.89 (s, 1H, NNH), 9.81 (s, 1H, C(O)NH), 8.22 (s, 1H, NCHN), 8.13–8.11 (m, 1H, ArH), 8.05 (s, 1H, NCHCBr), 7.38–7.36 (m, 1H, ArH), 7.27–7.16 (m, 6H, 6×ArH), 4.52 (t, *J* 7.2 Hz, 2H, NCH_2CH_2Ph), 3.23 (t, *J* 7.2 Hz, 2H, NCH_2CH_2Ph); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.4, 143.0, 141.3, 138.0, 133.8, 133.7, 132.9, 129.0, 128.7, 128.4, 126.5, 122.9, 109.5, 105.9, 93.1, 45.7, 35.4; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{19}H_{16}{}^{79}BrN_5O$ 409.1 found 410.0 $[M+H]^+$. **HRMS** (ESI, +ve) m/z calc. for $C_{19}H_{17}{}^{79}BrN_5O$ 410.0611 found 410.0611 $[M+H]^+$.

N-(1-(2-(Benzyloxy)ethyl)-1*H*-benzo[*d*]imidazol-4-yl)-4-bromo-1*H*-pyrazole-5-carboxamide **74**



To a microwave vial was added acid **97b** (0.12 mmol, 23 mg, 1.2 eq.) and HATU (0.12 mmol, 46 mg, 1.2 eq.) in MeCN (0.4 mL). To the resulting solution was added amine **98l** (0.10 mmol, 27 mg, 1.0 eq.) and Hünigs base (0.2 mmol, 30 µL, 2.0 eq.) and the mixture was heated to 40 °C and stirred overnight. The following day a solution of acid **97b** (0.1 mmol, 19 mg, 1.0 eq.) and HATU (0.1 mmol, 38 mg, 1.0 eq.) in MeCN (50 µL) were added and the resulting mixture was stirred overnight at 40 °C. Upon completion, the reaction mixture was diluted with EtOAc (10 mL) and washed with H_2O (3×10 mL) and brine (10mL). The organics were dried over $MgSO_4$, filtered and concentrated *in vacuo*. This afforded the *title compound* **74** without any further purification as a brown solid (0.09 mmol, 40 mg, 91%). **MP** 78–80 °C; **IR** (ATR, cm^{-1}) 3359, 2961, 2917, 1673, 1623, 1537, 1495; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.91 (s, 1H, NH), 9.85 (s, 1H, C(O)NH), 8.22 (s, 2H, NCHN and NCHCBr), 8.13 (d, *J* 8.0 Hz, 1H, ArH), 7.37 (d, *J* 8.0 Hz, 1H, ArH), 7.30–7.17 (m, 6H, 6×ArH), 4.49 (t, *J* 5.0 Hz, 2H, NCH_2CH_2O), 4.47 (s, 2H, CH_2Ph), 3.80 (t, *J* 5.0 Hz, 2H, NCH_2CH_2O); **^{13}C NMR** (151 MHz, DMSO-d_6) δ 143.5, 141.2, 138.0, 134.0, 133.8, 132.9, 129.0, 128.2, 127.4, 127.2, 122.9, 109.5, 106.0, 93.1, 71.8, 68.0, 44.5, 38.2; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{20}H_{18}{}^{79}BrN_5O_2$ 439.1 found 441.9 $[M+3H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{20}H_{19}{}^{79}BrN_5O_2$ 440.0717 found 440.0715 $[M+H]^+$.

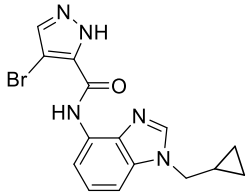
4-Bromo-*N*-(1-(furan-3-ylmethyl)-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **75**



According to **General Procedure D**; acid **97b** (0.32 mmol, 61 mg), HATU (0.32 mmol, 122 mg), amine **98p** (0.15 mmol, 31 mg) and Hünigs base (0.32 mmol, 60 µL) in MeCN (0.6 mL) were stirred at 45 °C overnight. Purification *via* flash column chromatography in 4:1 EtOAc/petroleum ether afforded the *title compound* **75** as a white solid (0.10 mmol, 40 mg, 65%). **MP** decomp. >180 °C; **IR** (ATR, cm^{-1}) 3357, 3196, 3136, 2917, 1660, 1539; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.89 (br s, 1H, NNH), 9.85 (br s, 1H, C(O)NH), 8.34 (s, 1H, NCHN), 8.22 (s, 1H, NCHCBr), 8.12 (d, *J* 8.0 Hz, 1H, ArH), 7.81 (s, 1H, ArH), 7.61–7.60 (m, 1H, ArH), 7.36 (d, *J* 8.0 Hz, 1H, ArH), 7.23 (app t, *J* 8.0 Hz, 1H, ArH), 6.47–6.46 (m, 1H, ArH), 5.35 (s, 2H, CH_2); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.4, 144.0, 143.0, 141.3, 141.1, 133.9, 133.6, 132.9, 129.1, 123.0, 121.0, 110.3, 109.6, 106.0, 93.1, 1 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{16}H_{12}{}^{79}BrN_5O_2$ 385.0 found 386.0 $[M+H]^+$.

9.4.6.4 Aliphatic cyclic analogues

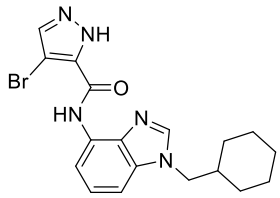
4-Bromo-*N*-(1-(cyclopropylmethyl)-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **76**



According to **General Procedure D**; acid **97b** (0.26 mmol, 49 mg), HATU (0.26 mmol, 99 mg) amine **98k** (0.18 mmol, 34 mg) and Hünigs base (0.42 mmol, 70 µL) in MeCN (0.6 mL) were stirred at 30 °C overnight. Purification *via* flash column chromatography in 4:1 EtOAc/hexanes afforded the *title compound* **76** as a white solid (0.1 mmol, 34 mg, 50%).**MP** 185–186 °C; **IR** (ATR, cm^{-1}) 3384, 3128, 3100, 2869, 2800, 1683, 1495; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.91 (br s, 1H, NNH), 9.85 (s, 1H, C(O)NH), 8.27 (s, 1H, NCHN), 8.21 (s, 1H, NCHCBr), 8.12 (d, *J* 8.0 Hz, 1H, ArH), 7.39 (d, *J* 8.0

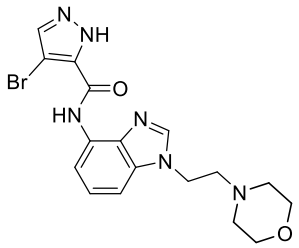
Hz, 1H, ArH), 7.25 (app t, *J* 8.0 Hz, 1H, ArH), 4.14–4.12 (m, 2H, NCH_2CH), 1.34–1.22 (m, 1H, NCH_2CH), 0.57–0.52 (m, 2H, C(H)H), 0.44-0.41 (m, 2H, C(H)H);**^{13}C NMR** (151 MHz, DMSO-d_6) δ 158.4, 142.8, 141.3, 133.9, 133.8, 132.9, 129.1, 122.9, 109.5, 106.0, 93.1, 48.7, 11.4, 3.9; **LRMS** (LCMS-ESI) *m/z* calc. $C_{15}H_{14}{}^{79}BrN_5O$ for 359.0 found 360.8 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{15}H_{15}{}^{79}BrN_5O$ 360.0454 found 360.0457 $[M+H]^+$.

4-Bromo-*N*-(1-(cyclohexylmethyl)-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **77**



According to **General Procedure D**; acid **97b** (0.08 mmol, 16 mg), HATU (0.08.mmol, 30 mg), amine **98i** (0.07 mmol, 16 mg) and Hünigs base (0.14 mmol, 20 µL) in MeCN (0.3 mL) were stirred at 45 °C overnight. The following day a second portion of reagents were added, acid **97b** (0.07 mmol, 13 mg) and HATU (0.07 mmol, 27 mg) in MeCN (50 µL) and the resulting mixture was stirred overnight at 40 °C. Purification *via* flash column chromatography in 4:1 EtOAc/petroleum ether 40–60 afforded *title compound* **77** as a white solid (0.04 mmol, 16 mg, 57%). **MP** 177–179 °C; **IR** (ATR, cm^{-1}) 3640, 3356, 2930, 2917, 1623, 1539.**^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.90 (br s, 1H, NNH), 9.85 (s, 1H, C(O)NH), 8.22–8.21 (m, 2H, NCHN and NCHCBr), 8.12 (d, *J* 8.0 Hz, 1H, ArH), 7.35 (d, *J* 8.0 Hz, 1H, ArH), 7.24 (app t, *J* 8.0 Hz, 1H, ArH), 4.11 (d, *J* 7.0 Hz, 2H, NCH_2CH), 1.85–1.80 (m, 1H, $NCH_2CH(CH_2)_5$), 1.67–0.97 (m, 10H, $NCH_2CH(CH_2)_5$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.4, 143.5, 141.3, 134.2, 133.8, 132.9, 129.1, 122.9, 109.4, 106.1, 93.1, 50.2, 37.8, 30.0, 25.8, 25.1; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{18}H_{20}{}^{79}BrN_5O$ 401.1 found 402.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{18}H_{21}{}^{79}BrN_5O$ 402.0924 found 402.0924 $[M+H]^+$.

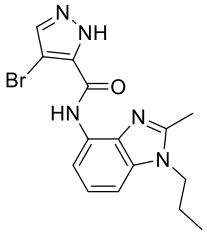
4-Bromo-*N*-(1-(2-morpholinoethyl)-1*H*-benzo[d]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **78**



To a microwave vial was added chloro derivative **68** (0.05 mmol, 20 mg, 1.0 eq.) and morpholine (20 µL, 4.5 eq.). The resulting mixture was heated to reflux and stirred overnight. Upon completion, the reaction mixture was diluted with EtOAc (10 mL) and washed with water (3×10 mL) and brine (10 mL). The organics were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography in 98:2—90:10 EtOAc/MeOH afforded the *title compound* **78** (0.03 mmol, 11 mg, 49%). **IR** (ATR, cm^{-1}) 3355, 3103, 2846, 1618, 1532; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 9.85 (s. 1H, C(O)NH), 8.24 (s, 1H, NCHN), 8.21 (s, 1H, NCHCBr), 8.12 (d, J 8.0 Hz, 1H, ArH) ,7.37 (d, J 8.0 Hz, 1H, ArH), 7.25 (app t, *J* 8.0 Hz, 1H, ArH) 4.38 (t, J 6.2 Hz, 2H, NCH_2CH_2N), 3.54– 3.51 (m, 4H, 2×CH_2), 2.69 (t, J 6.2 Hz, NCH_2CH_2N), 2.44–2.41 (m, 4H, 2×CH_2); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.0, 143.5, 133.9, 133.7, 129.0, 122.8, 109.5, 105.9, 93.1, 66.2, 57.2, 54.9, 53.1, 41.5, 3 C missing; **LRMS** (LCMS-ESI) m/z calc. $C_{17}H_{19}{}^{79}BrN_6O2$ 418.0 found 418.9 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{17}H_{20}{}^{79}BrN_6O_2$ 419.0826 found 419.0826 $[M+H]^+$.

9.4.7 C-2 Tail group derivatives

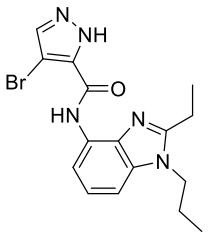
4-Bromo-*N*-(2-methyl-1-propyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **79**



According to **General Procedure D;** acid **97b** (0.21 mmol, 40 mg), HATU (0.21 mmol, 80 mg), amine **98m** (0.17 mmol, 33 mg, 1.0 eq) and Hünigs base (0.34 mmol, 60 µL) in MeCN (0.6 mL) were stirred at 40 °C overnight. Purification *via* flash column

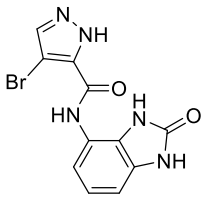
chromatography in 9:1 EtOAc/hexanes afforded the *title compound* **79** as a white solid (0.02 mmol, 8 mg, 13%). **MP** 188–190 °C; **IR** (ATR, cm^{-1}) 3365, 2964, 2925, 1669, 1416; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.87 (s, 1H, NNH), 9.79 (s, 1H, C(O)NH), 8.22 (s, 1H, NCHCBr), 8.07 (d, *J* 7.9 Hz, 1H, ArH), 7.26 (d, *J* 7.9 Hz, 1H, ArH), 7.16 (app t, *J* 7.9 Hz, 1H, ArH) , 4.16 (t, *J* 7.1Hz, 2H, $NCH_2CH_2CH_3$), 2.58 (s, 3H, $NCCH_3$), 1.75 (app sextet, *J* 7.2 Hz, 2H, $NCH_2CH_2CH_3$) 0.88 (t, *J* 7.2 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.3, 151.0, 141.3, 135.1, 132.9, 132.6, 128.2, 122.0, 109.3, 105.4, 93.0, 44.7, 22.5, 13.4, 11.0; **LRMS** (LCMS-ESI) *m/z* calc. for. $C_{15}H_{16}{}^{79}BrN_5O$ 361.0 found 362.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{15}H_{17}{}^{79}BrN_5O$ 362.0616 found 362.0612 $[M+H]^+$.

4-Bromo-*N*-(2-ethyl-1-propyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **80**



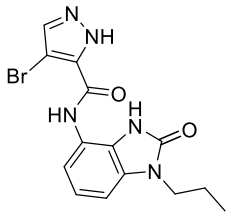
According to **General Procedure D**; acid **97b** (0.25 mmol, 47 mg), HATU (0.25 mmol, 95 mg), amine **98m** (0.21 mmol, 42 mg) and Hünigs base (0.42 mmol, 70 µL) in MeCN (0.7 mL) were stirred at 40 °C overnight. Purification *via* flash column chromatography in 9:1 EtOAc/hexanes afforded the *title compound* **80** as a white solid (0.15 mmol, 55 mg, 69%). **MP** 90–92 °C; **IR** (ATR, cm^{-1}) 3456, 3147, 2967, 1621, 1416; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.90 (s, 1H, NNH), 9.81 (s, 1H, C(O)NH), 8.24 (s, 1H, NCHCBr), 8.09 (d, *J* 7.7 Hz, 1H, ArH), 7.27 (d, *J* 7.7 Hz, 1H, ArH), 7.17 (app t, *J* 7.7 Hz, 1H, ArH), 4.15 (t, *J* 7.2 Hz, 2H, $NCH_2CH_2CH_3$), 2.91 (q, *J* 7.5 Hz, 2H, $NCCH_2CH_3$), 1.74 (app sextet, *J* 7.2 Hz, 2H, $NCH_2CH_2CH_3$), 1.36 (t, *J* 7.5 Hz, 3H, $NCCH_2CH_3$), 0.88 (t, J 7.2 Hz, 3H, $NCH_2CH_2CH_3$).**^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.3, 155.3, 141.3, 135.2, 132.9, 132.6, 128.3, 122.1, 109.3, 105.6, 93.1, 44.5, 22.8, 20.0, 12.0, 11.1; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{16}H_{18}{}^{79}BrN_5O$ 375.1 found 376.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{16}H_{19}{}^{79}BrN_5O$ 376.0767 found 376.0769 $[M+H]^+$.

4-Bromo-*N*-(2-oxo-2,3-dihydro-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **81**



To a microwave vial was added acid **97b** (0.40 mmol, 77 mg, 1.2 eq.) and HATU (0.40 mmol, 152 mg, 1.2 eq.). The vial was sealed and to the mixture was added a solution of amine **131a** (0.34 mmol, 36 mg, 1.0 eq.) in MeCN (1.1 mL, 0.3 M) and Hünigs base (0.68 mmol, 120 µL, 2.0 eq.). The resulting mixture was allowed to stir at 40 °C overnight. Upon completion, the reaction mixture was diluted with EtOAc (1 mL) and a precipitate formed. The solids were filtered and washed with H_2O (3×5 mL) to provide the *title compound* **81** as a white solid (0.06 mmol, 18 mg, 16%). **MP** >250 °C; **IR** (ATR, cm^{-1}) 3336, 3195, 3128, 2799, 1714, 1668, 1558; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.79 (br s, 1H, NNH), 10.66 (br s, 1H, NHC(O)NH), 10.37 (br s, 1H, NHC(O)NH), 9.80 (s, 1H, CC(O)NH), 8.15 (s, 1H, NCHCBr), 7.26 (d, *J* 8.0 Hz, 1H, ArH), 6.91 (app t, *J* 8.0 Hz, 1H, ArH), 6.77 (d, *J* 8.0 Hz, 1H, ArH); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 159.3, 154.7, 132.1, 130.6, 122.5, 120.3, 115.4, 105.4, 93.3, 2 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{11}H_8{}^{79}BrN_5O_2$ 321.0 found 322.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{11}H_9{}^{79}BrN_5O_2$ 321.9940 found 321.9933 $[M+H]^+$.

4-Bromo-*N*-(2-oxo-1-propyl-2,3-dihydro-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrazole-5-carboxamide **82**

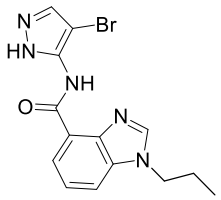


According to **General Procedure D**; acid **97b** (0.09 mmol, 18 mg), HATU (0.09 mmol, 34 mg), amine **131b** (0.08 mmol, 15 mg) and Hünigs base (0.16 mmol, 30 µL) in MeCN (0.3 mL) were stirred at 30 °C overnight. A second portion of acid **97b** (0.0 mmol, 18 mg) and HATU (0.09 mmol, 34 mg) was added and the resulting mixture was heated to 40 ºC overnight. Purification *via* flash column chromatography in 4:1 EtOAc/hexanes afforded the *title compound* **82** as a white solid (0.03 mmol, 10 mg,

34%). **IR** (ATR, cm^{-1}) 3158, 2960, 2928, 1675, 1651, 1539; **^{1}H NMR** (500 MHz, DMSO-d_6) δ 13.77 (s, 1H, NNH), 10.61 (s, 1H, NC(O)NH), 9.84 (s, 1H, C(O)NH), 8.17 (s, 1H, ArH) 7.30–7.29 (m, 1H, ArH), 6.99–6.98 (m, 2H, 2×ArH), 3.75 (t, *J* 7.0 Hz, 2H, $NCH_2CH_2CH_3$), 1.66 (app sextet, *J* 7.0 Hz, 2H, $NCH_2CH_2CH_3$), 0.88 (t, *J* 7.0 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (151 MHz, DMSO-d_6) δ 153.8, 132.0, 131.2, 126.4, 124.1, 121.2, 120.3, 115.9, 104.6, 41.5, 21.2, 13.5, 11.1; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{14}{}^{79}BrN_5O_2$ 363.0 found 364.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{15}{}^{79}BrN_5O_2$ 364.0404 found 364.0407 $[M+H]^+$.

9.4.8 Reverse amide analogue

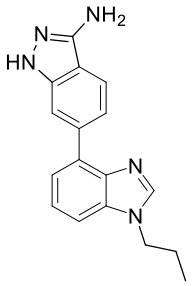
N-(4-bromo-1H-pyrazol-5-yl)-1-propyl-1H-benzo[d]imidazole-4-carboxamide **87**



According to **General Procedure E**; acid **136** (0.26 mmol, 36 mg), 2-chloro-1-methylpyridin-1-ium iodide (0.26 mmol, 66 mg), DMAP (0.11 mmol, 14 mg), NEt_3 (0.63 mmol, 90 µL) and amine **145** (0.21 mmol, 55 mg) in anhydrous CH_2Cl_2 (0.7 mL) were stirred at 35 °C overnight. Further purification *via* flash column chromatography 3:2 40–60 petroleum ether/EtOAc, followed by trituration in hexanes afforded the *title compound* **87** as a white solid (0.02 mmol, 8 mg, 11%). **MP** 187–189 °C; **IR** (ATR, cm^{-1}) 3290, 3107, 3037, 2967, 1609, 1498; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 12.73 (br s, 1H, NH), 8.22–8.21 (m, 1H, ArH), 8.10 (s, 1H, NCHN), 7.67–7.65 (m, 1H, ArH), 7.52 (s, 1H, CHCBr) 7.50–7.46 (m, 1H, ArH), 4.24 (t, *J* 7.2 Hz, 2H, $NCH_2CH_2CH_3$), 1.98 (app sextet, *J* 7.2 Hz, 2H, $NCH_2CH_2CH_3$), 1.00 (t, *J* 7.2 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 164.3, 143.3, 141.6, 139.0, 138.7, 134.2, 124.4, 123.3, 121.7, 114.8, 79.3, 47.4, 23.4, 11.5; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{14}{}^{79}BrN_5O$ 347.0 found 348.1 $[M+H]^+$.

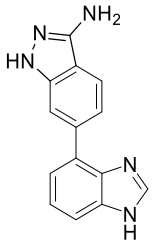
9.4.9 Hybrid series

6-(1-Propyl-1*H*-benzo[*d*]imidazol-4-yl)-1H-indazol-3-amine **89**



To a microwave vial was added the fluoro nitrile analogue **155b** (0.11 mmol, 30 mg, 1.0 eq.) in EtOH (0.6 mL, 0.2 M). To the resulting solution was added hydrazine hydrate (0.41 mmol, 20µL, 3.8 eq.) and the mixture was subjected to microwave irradiation at 100°C for 2 hours. To the mixture was added a second portion of hydrazine hydrate (0.41 mmol, 20 µL) and the resulting suspension was heated to 100°C in an oil bath for 3 hours. The mixture was then cooled to room temperature whereupon a precipitate formed. The solids were filtered and washed with H_2O (2×5 mL) to provide the *title compound* **89** as a white solid (0.05 mmol, 13 mg, 41%). **MP** 222–224 °C; **IR** (ATR, cm^{-1}) 3268, 3173, 2980, 2939, 1581, 1495; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.43 (s, 1H, NH), 8.30 (s, 1H, ArH), 8.16–8.14 (m, 1H, ArH), 7.76–7.74 (m, 1H, ArH), 7.60–7.58 (m, 1H, ArH), 7.55–7.49 (m, 2H, 2×ArH), 7.36– 7.32 (m, 1H, ArH), 5.34 (br s, 2H, NH_2), 4.26 (t, J 7.0 Hz, 2H, $NCH_2CH_2CH_3$), 1.87 (app sextet, *J* 7.3 Hz, 2H, $NCH_2CH_2CH_3$), 0.87 (t, J 7.3 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 149.1, 144.0, 141.9, 141.1, 135.8, 134.6, 131.8, 122.6, 120.6, 119.8, 118.7, 113.1, 110.1, 109.5, 45.7, 22.8, 11.0; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{17}H_{17}N_5$ 291.2 found 292.1 $\{M+H\}^+$.

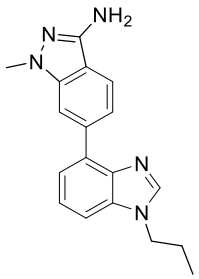
6-(1*H*-benzo[*d*]imidazol-4-yl)-1H-indazol-3-amine **90**



To a microwave vial was added the fluoro nitrile analogue **155a** (0.08 mmol, 16 mg, 1.0 eq.) in EtOH (0.4 mL, 0.2 M). To the solution was added hydrazine hydrate (0.34 mmol, 16 µL, 4.0 eq.) and the mixture was subjected to microwave irradiation at 120°C

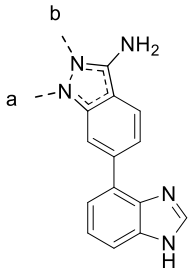
for 2 hours. The resulting mixture was concentrated *in vacuo*. The residue was then taken up in a 1:1 mixture of CH_2Cl_2: hexanes (10 mL) and filtered to provide the *title compound* **90** as a brown solid (0.03 mmol, 8 mg, 40%). **MP** decomp. >190 °C; **IR** (ATR, cm^{-1}) 3119, 2961, 2923, 2853, 1625, 1426; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.43 (br s, 0.36H, NCHNH), 8.26–8.17 (m, 2H, 2×ArH), 7.80–7.70 (m, 2H, 2×ArH), 7.52–7.45 (m, 2H, 2×ArH), 7.31–7.27 (m, 1H, ArH), 5.34 (br s, 2H, NH_2); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 149.1, 141.9, 140.6, 136.0, 134.1, 131.4, 122.6, 120.4, 119.8, 118.7, 118.2, 113.1, 110.8, 110.0; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{11}N_5$ 249.1 found 250.1 $[M+H]^+$.

1-Methyl-6-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-indazol-3-amine **91**



To a microwave vial was added the fluoro nitrile analogue **155b** (0.07 mmol, 20 mg, 1.0 eq.) in EtOH (0.18 mL, 0.4 M). To the resulting solution was added methyl hydrazine (0.36 mmol, 19 µL, 5.0 eq.) and the reaction was heated to reflux overnight. Upon completion, the reaction mixture was diluted with EtOAc (5 mL) and washed with H_2O (5 mL). The aqueous was extracted with EtOAc (4×5 mL), washed with brine (5 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography in 100% hexanes to 95:5 CH_2Cl_2/MeOH and subsequent trituration in hexanes provided the *title compound* **91** as an orange solid (0.02 mmol, 7 mg, 33%). **MP** 63–65 °C; **IR** (ATR, cm^{-1}) 3304, 3196, 2902, 2927, 1618, 1497; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.19 (s, 1H, NCHN), 7.85 (s, 1H, ArH), 7.65–7.63 (m, 1H, ArH) 7.57–7.50 (m, 2H, 2×ArH), 7.44–7.43 (m, 2H, 2×ArH), 4.24 (t, J 7.1 Hz, 2H, $NCH_2CH_2CH_3$), 3.92 (s, 3H, NCH_3), 1.98 (app sextet, J 7.1 Hz, 2H, $NCH_2CH_2CH_3$), 1.01 (t, J 7.1 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 146.9, 143.0, 142.2, 137.1, 134.4, 133.3, 123.7, 122.8, 120.4, 119.7, 114.0, 109.6, 109.3, 47.3, 35.2, 23.3, 11.5, 1 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{18}H_{19}N_5$ 305.2 found 306.3 $[M+H]^+$.

6-(1*H*-benzo[*d*]imidazol-4-yl)-1-methyl-1*H*-indazol-3-amine: 6-(1*H*-benzo[*d*]imidazol-4-yl)-2-methyl-2*H*-indazol-3-amine **92a/b** (3:1 mixture)

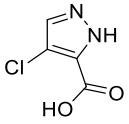


To a microwave vial was added the fluoro nitrile analogue **155a** (0.08 mmol, 20 mg, 1.0 eq.) in EtOH (0.5 mL, 0.2 M). To the resulting solution was added methyl hydrazine (0.42 mmol, 22 µL, 5.0 eq.) and the reaction was subjected to microwave irradiation at 120°C for 2 hours Upon completion, the reaction mixture was diluted with CH_2Cl_2 (10 mL) and washed with H_2O (10 mL). The aqueous was extracted with CH_2Cl_2 (4×5 mL), washed with brine (5 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo.* The resulting residue was triturated in hexanes and filtered to provide a mixture of two regioisomers **92a/92b** in a 3:1 ratio as an orange solid (0.03 mmol, 8 mg, 38%). **MP** 170–172 °C; **IR** (ATR, cm^{-1}) 3306, 3169, 3064, 2820, 1621; **^{1}H NMR** (400 MHz, Acetone-d_6) δ 8.22 (s, 1H, ArH), 8.20 (s, 0.3H, ArH), 7.76–7.30 (m, 7.8H, 7.8×ArH), 4.04 (s, 3H, CH_3), 3.84 (s, 1H, 0.33×CH_3) **^{13}C NMR** (101 MHz, DMSO-d_6) δ 148.3, 143.7, 142.7, 142.0, 141.5, 140.7, 136.6, 136.1, 134.1, 133.5, 131.5, 131.2, 122.5, 122.3, 121.7, 121.0, 120.6, 119.9, 119.0, 118.2, 118.0, 113.7, 113.4, 110.9, 108.7, 108.1, 54.9, 34.5; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{15}H_{13}N_5$ 263.1 found 264.1 $[M+H]^+$.

9.5 Synthesis of intermediates

9.5.1 Acids and anilines

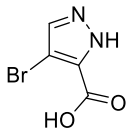
4-Chloro-1*H*-pyrazole-5-carboxylic acid **97a**[162]



To a solution of 4-chloro-5-methyl-*1H*-pyrazole **102a** (8.6 mmol, 1.0 g, 1.0 eq), which was previously synthesised by Fiona Keatings, in H_2O (43 mL, 0.2 M) was added $KMnO_4$ (25.8 mmol, 4.1 g, 3.0 eq). The mixture was heated to reflux and stirred for 40

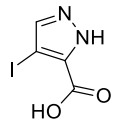
hours. Upon completion the solids were filtered and the volume of the filtrate reduced to approximately a 10th *in vacuo*. The filtrate was then treated with 0.2 M HCl until pH 2 was reached. A precipitate formed and was filtered to provide the *title compound* **97a** as a white solid (4.3 mmol, 629 mg, 50%). **MP lit**[162] 250°C; **MP** decomp. 223–225°C; **IR** (ATR, cm^{-1}) 3357, 3137, 2747, 1694; **^{1}H NMR**, (400 MHz, DMSO-d_6) δ 13.57 (br s, 2H, NH and COOH), 7.91 (br s, 1H, ArH); **^{13}C NMR**, (101 MHz, DMSO-d_6) δ 161.4, 131.9, 129.3, 111.5; **LRMS** (LCMS-ESI) *m/z* calc. for $C_4H_3{}^{35}ClN_2O_2$ 146.0 found 145.0 $[M-H]^-$.

4-Bromo-1*H*-pyrazole-5-carboxylic acid **97b**[162]



To a solution of acid **103** (17.5 mmol, 2.0 g, 1.0 eq) in AcOH (50 mL, 0.4 M) was added bromine (19.3 mmol, 1.0 mL, 1.1 eq) dropwise followed by potassium acetate (KOAc) (19.3 mmol, 1.9 g, 1.1 eq) portion-wise. The resulting mixture was stirred for 72 hours at room temperature. Upon completion, the mixture was concentrated *in vacuo* before being diluted with EtOAc (50 mL) and washed with water (3×50 mL) and brine (50mL). The combined organics were dried over $MgSO_4$, filtered and concentrated *in vacuo* to provide the *title compound* **97b** as a white solid (9.1 mmol, 1.7 g, 51%). **MP lit**[162] 250–252°C **MP** decomp. 240°C; **IR** (ATR, cm^{-1}) 3350, 3155, 2504, 1690; **^{1}H NMR**, (400 MHz, DMSO-d_6) δ 14.49 (br s, 1H, NH), 13.79 (br s, 1H, COOH), 7.97 (br s, 1H, ArH); **^{13}C NMR**, (101 MHz, DMSO-d_6) δ161.0, 136.4, 95.6, 1 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_4H_3{}^{79}BrN_2O_2$ 190.0, found 189.0 $[M-H]^-$.

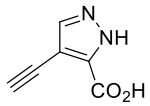
4-Iodo-1*H*-pyrazole-5-carboxylic acid **97c**[152]



1H-pyrazole-5-carboxylic acid **103** (0.9 mmol, 100 mg, 1.0 eq) was dissolved in 50% v/v H_2SO_4 (0.3 mL, 3.4 M), cooled to 0 °C and NIS (1.1 mmol, 236 mg, 1.2 eq) was added. The resultant mixture was allowed to stir for 10 minutes at 0°C then 3 hours at room temperature. The mixture was then dispersed in water (2 mL) and allowed to stir for a further 48 hours. The resulting emulsion was suspended in water (2 mL) and stirred overnight at room temperature. The precipitated solids were filtered off and then suspended in boiling H_2O (3 mL). After treatment with saturated sodium

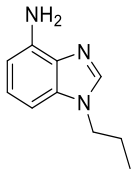
thiosulfate ($Na_2S_2O_3$) (0.1 mL) the solution was allowed to cool to room temperature and the resulting solids were filtered to provide the *title compound* **97c** as a white solid (0.5 mmol, 123 mg, 59%). **MP lit**[152] 240°C; **MP** 240–242°C; **IR** (ATR, cm^{-1}) 3189, 3135, 2350, 1713; **^{1}H NMR**, (400 MHz, DMSO-d_6) δ 13.39 (br s, 2H, NH and COOH), 7.88 (s, 1H, ArH); **^{13}C NMR**, (101 MHz, DMSO-d_6) δ 161.6, 142.4, 61.9, 1 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_4H_3IN_2O_2$ 237.9 found 238.9 $[M+H]^+$.

4-Ethynyl-1*H*-pyrazole-5-carboxylic acid **97d**



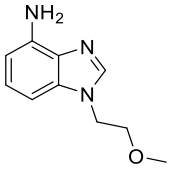
According to **General Procedure H**, ester **110** (0.4 mmol, 60 mg) was dissolved in 1:5:1 MeOH/THF/1M LiOH (14 mL) and heated to 80 °C overnight. A white precipitate formed and was filtered to afford the *title compound* **97d** as a white solid (0.3 mmol, 40 mg, 73%). **MP** decomp. >130 °C; **IR** (ATR, cm^{-1}) 3296, 3135, 2919, 1691; **^{1}H NMR** (400 MHz, DMSO) δ 13.51 (s, 2H, COOH and NH), 8.01 (s, 1H, ArH), 4.10 (s, CH); **^{13}C NMR** (101 MHz, DMSO) δ 162.0, 143.5, 136.0, 103.3, 83.8, 75.03; **LRMS** (LCMS-ESI) *m/z* calc. for $C_6H_4N_2O_2$ 136.0 found 137.0 $[M+H]^+$.

1-Propyl-1*H*-benzo[*d*]imidazol-4-amine **98a**



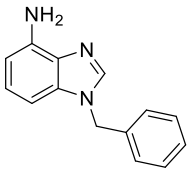
According to **General Procedure B**; 10 mol% Pd/C (0.1 mmol, 70 mg) and nitrobenzimidazole **122a** (1.2 mmol, 250 mg) in MeOH (12 mL) were stirred at room temperature for 5 hours. The *title compound* **98a** was isolated without further purification as a beige solid (0.8 mmol, 138 mg, 65%). **MP** 94–96°C; **IR** (ATR, cm^{-1}) 3409, 3147, 2924, 1610, 1493; **^{1}H NMR**, (400 MHz, $CDCl_3$) δ 7.75 (s, 1H, NCHNR), 7.09 (app t, *J* 8.0 Hz, 1H, ArH), 6.78 (dd, *J* 8.0, 0.7 Hz, 1H, ArH), 6.53 (dd, *J* 8.0, 0.7 Hz, 1H, ArH), 4.34 (br s, 2H, NH_2), 4.09 (t, *J* 7.2 Hz, 2H, $NCH_2CH_2CH_3$), 1.90 (app sextet, J 7.2 Hz, 2H, $NCH_2CH_2CH_3$), 0.96 (t, *J* 7.2 Hz, 3H, CH_3); **^{13}C NMR**, (125 MHz, $CDCl_3$) δ 140.9, 139.2, 134.8, 133.2, 124.1, 105.9, 99.8, 47.0, 23.3, 11.5; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{13}N_3$ 175.1 found 176.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{10}H_{14}N_3$ 176.1182 found 176.1181 $[M+H]^+$.

1-(2-Methoxyethyl)-1*H*-benzo[*d*]imidazol-4-amine **98b**



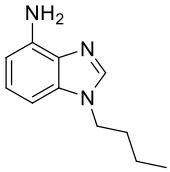
According to **General Procedure B**; 10 mol% Pd/C (0.1 mmol, 80 mg) and nitrobenzimidazole **122b** (0.6 mmol, 142 mg) in MeOH (7.0 mL) were stirred at room temperature overnight. The *title compound* **98b** was isolated without any further purification as a beige solid (0.6 mmol, 110 mg, 91%). **MP** 94–96 °C; **IR** (ATR, cm^{-1}) 3428, 3194, 2919, 1593, 1495; **^{1}H NMR**, (400 MHz, $CDCl_3$) δ 7.82 (s, 1H, NC<u>H</u>NR), 7.09 (app t, *J* 8.0 Hz, 1H, Ar<u>H</u>), 6.77 (dd, *J* 8.0, 0.7 Hz, 1H, Ar<u>H</u>), 6.54 (dd, *J* 8.0, 0.7 Hz. 1H, Ar<u>H</u>), 4.36 (br s, 2H, N<u>H$_2$</u>), 4.26 (t, *J* 5.3 Hz, 2H, NC<u>H$_2$</u>CH_2O), 3.70 (t, *J* 5.3 Hz, 2H, NCH_2C<u>H$_2$</u>O), 3.31 (s, 3H, OC<u>H$_3$</u>); **^{13}C NMR**, (101 MHz, $CDCl_3$) δ 141.6, 139.2, 134.7, 133.0, 124.1, 105.9, 99.4, 70.7, 59.2, 45.2; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{13}N_3O$ 191.1 found 192.1 $[M+H]^+$.

1-Benzyl-1*H*-benzo[*d*]imidazol-4-amine **98c**



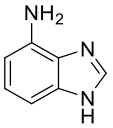
According to **General Procedure** B; 10 mol% Pd/C (0.1 mmol, 90 mg) and nitrobenzimidazole **122c** (0.8 mmol, 200 mg) in MeOH (9 mL) were stirred at room temperature for 5 hours. The *title compound* **98c** was isolated without further purification as a beige solid (0.7 mmol, 145 mg, 83%). **MP** 113–115 °C; **IR** (ATR, cm^{-1}) 3402, 3307, 2919, 1632, 1595, 1455; **^{1}H NMR**, (500 MHz, $CDCl_3$) δ 7.81 (s, 1H, NC<u>H</u>NR), 7.35–7.30 (m, 3H, 3×Ar<u>H</u>), 7.19–7.17 (m, 2H, 2×Ar<u>H</u>), 7.05 (app t, *J* 8.0 Hz, 1H, Ar<u>H</u>), 6.69 (d, *J* 8.0 Hz, 1H, Ar<u>H</u>), 6.54 (d, *J* 8.0 Hz, 1H, Ar<u>H</u>), 5.31 (s, 2H, C<u>H$_2$</u>), 4.36 (s, 2H, N<u>H$_2$</u>); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 141.1, 139.2, 135.9, 135.0, 133.2, 129.1, 128.3, 127.2, 124.4, 106.1, 100.0, 49.01; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{13}N_3$ 223.1 found 224.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{14}N_3$ 224.1182 found 224.1183 $[M+H]^+$.

1-Butyl-1*H*-benzo[*d*]imidazol-4-amine **98d**



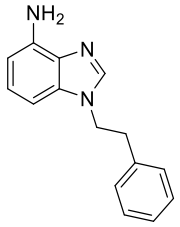
According to **General Procedure** B; 10 mol% Pd/C (0.1 mmol, 70 mg) and nitrobenzimidazole **122d** (0.7 mmol, 150 mg) in MeOH (8 mL) were stirred at room temperature for 1.5 hours. The *title compound* was carried onto the next step without further purification as a 9:1 mixture of **98d**:**98e** and as a beige solid (0.7 mmol, 145 mg, 83%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.74 (s, 1H, NCHN), 7.09 (app t, J 8.0 Hz, 1H, ArH), 6.78 (dd, J Hz, 1H, ArH) 6.53 (dd, J Hz, 1H, ArH), 4.35 (br s, 2H, NH_2), 4.13–4.07 (m, 2H, $NCH_2CH_2CH_2CH_3$), 1.88–1.81 (m, 2H, $NCH_2CH_2CH_2CH_3$), 1.39–1.31 (m, 2H, $NCH_2CH_2CH_2CH_3$), 0.97–0.93 (m, 3H, CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 140.9, 139.2, 134.8, 133.2, 124.0, 105.8, 99.7, 45.1, 32.0, 20.1, 13.7; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{11}H_{15}N_3$ 189.1 found 190.1 $[M+H]^+$.

1*H*-benzo[*d*]imidazol-4-amine **98e**[163]



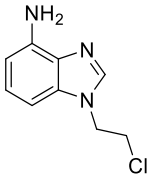
According to **General Procedure B**; 1 mol% Pd/C (0.6 mmol, 65 mg) and nitrobenzimidazole **121** (6.1 mmol, 1 g) in MeOH (68 mL) were stirred at room temperature overnight. The *title compound* **98e** was isolated without further purification as a dark green solid (5.7 mmol, 754 mg, 92%). **MP** 112–114 °C; **IR** (ATR, cm^{-1}) 3432, 3348, 3077, 2815, 1610, 1429; **^{1}H NMR**, (500 MHz, $CDCl_3$) δ 7.92 (s, 1H, NCHNH), 7.09 (app t, *J* 8.0 Hz, 1H, ArH), 6.90 (d, *J* 8.0 Hz, 1H, ArH), 6.56 (d, *J* 8.0 Hz, 1H, ArH), 3.49 (s, 2H, NH_2); **^{13}C NMR**, (125 MHz, $CDCl_3$) δ 138.6, 138.5, 134.0, 131.8, 124.6, 106.4, 101.6; **LRMS** (LCMS-ESI) *m/z* calc. for $C_7H_7N_3$ 133.1 found 134.1 $[M+H]^+$

1-Phenethyl-1*H*-benzo[*d*]imidazol-4-amine **98f**



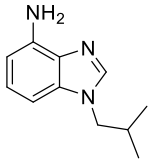
According to **General Procedure C**; $FeSO_4.7H_2O$ (2.3 mmol, 650 mg), citric acid (0.2 mmol, 41 mg), $NaBH_4$ (3.9 mmol, 149 mg) and nitrobenzimidazole compound **122e** (0.4 mmol, 100 mg) in H_2O (45 mL) were stirred overnight at room temperature. The *title compound* **98f** was isolated without any further purification as an orange oil (0.33 mmol, 60 mg, 81%). **IR** (ATR, cm^{-1}) 3450, 3352, 3026, 2917, 1613, 1495; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.48 (s, 1H, NCHNR), 7.30–7.22 (m, 3H, 3×ArH), 7.12 (app t, *J* 8.0 Hz, 1H, ArH), 7.06 (m, 2H, 2×ArH), 6.82 (dd, *J* 8.0, 0.6 Hz, 1H, ArH), 6.55 (dd, *J* 8.0, 0.6 Hz, 1H, ArH), 4.35 (t, *J* 7.1 Hz, 2H, NCH_2CH_2Ph), 3.13 (t, *J* 7.1 Hz, 2H, NCH_2CH_2Ph); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 140.7, 139.1, 137.7, 134.3, 132.5, 129.0, 128.8, 127.2, 124.4, 106.2, 99.5, 47.1, 36.3; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{15}H_{15}N_3$ 237.1 found 238.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{15}H_{15}N_3$ 238.1339 found 238.1340 $[M+H]^+$.

1-(2-Chloroethyl)-1*H*-benzo[*d*]imidazol-4-amine **98g**



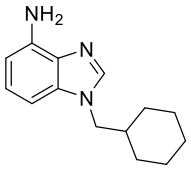
According to **General Procedure B**; 1 mol% Pd/C (0.04 mmol, 5 mg) and nitrobenzimidazole **122f** (0.4 mmol, 83 mg) in MeOH (5 mL) were stirred at room temperature for 2 hours. The *title compound* **98g** was isolated without further purification as an orange oil (0.3 mmol, 58 mg, 80%). **IR** (ATR, cm^{-1}) 2958, 2921, 2850, 1513, 1433; **^{1}H NMR**, (400 MHz, $CDCl_3$) δ 7.84 (s, 1H, NCHNR), 7.12 (app t, *J* 8.0 Hz, 1H, ArH), 6.75 (dd, *J* 8.0, 0.6 Hz, 1H, ArH), 6.56 (dd, *J* 8.0, 0.6 Hz, 1H, ArH), 4.47 (t, J 6.2 Hz, 2H, NCH_2CH_2Cl), 3.85 (t, *J* 6.2 Hz, 2H, NCH_2CH_2Cl); **^{13}C NMR**, (101 MHz, $CDCl_3$) δ 141.2, 139.5, 134.2, 133.2, 124.6, 106.3, 98.9, 46.9, 42.2; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_{10}{}^{35}ClN_3$ 195.1 found 196.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_9H_{10}{}^{35}ClN_3$ 196.0636 found 196.0636 $[M+H]^+$.

1-*Iso*butyl-1*H*-benzo[*d*]imidazol-4-amine **98h**



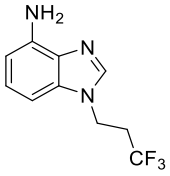
According to **General Procedure C;** $FeSO_4.7H_2O$ (2.3 mmol, 651 mg), citric acid (0.2 mmol, 41 mg), $NaBH_4$ (3.9 mmol, 149 mg) and nitrobenzimidazole **126a** (0.4 mmol, 100 mg) in H_2O (45 mL) were stirred overnight at room temperature. The *title compound* **98h** was isolated without any further purification as an orange oil (0.33 mmol, 57 mg, 77%). **IR** (ATR, cm^{-1}) 3425, 3156, 2962, 2873. 1740, 1612; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.73 (s, 1H, NC<u>H</u>N), 7.08 (app t, *J* 8.0 Hz, 1H, Ar<u>H</u>), 6.76 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 6.53 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 4.35 (br s, 2H, N<u>H$_2$</u>), 3.91 (d, *J* 7.0 Hz, 2H, NC<u>H$_2$</u>^{i}Pr), 2.22 (app septet, *J* 7.0 Hz, 1H, NCH_2C<u>H</u>$(CH_3)_2$), 0.95 (d, *J* 7.0 Hz, 6H, 2×C<u>H$_3$</u>); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 141.3, 139.1, 135.0, 133.2, 124.0, 105.8, 99.9, 52.9, 29.1, 20.3; **LRMS** (LCMS-ESI) *m/z* calc. $C_{11}H_{15}N_3$ for 189.1 found 190.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{11}H_{16}N_3$ 190.1339 found 190.1338 $[M+H]^+$.

1-(Cyclohexylmethyl)-1*H*-benzo[*d*]imidazol-4-amine **98i**



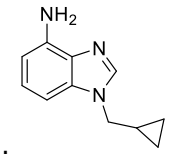
According to **General Procedure B**; 1 mol% Pd/C (0.02 mmol, 2 mg) and alkyl nitrobenzimidazole **126b** (0.2 mmol, 60 mg) in anhydrous EtOAc (2.6 mL) were stirred at room temperature overnight. Further purification via flash column chromatography in 100% EtOAc provided the *title compound* **98i** as a brown oil (0.07 mmol, 16 mg, 35%). **IR** (ATR, cm^{-1}) 3476, 3369, 2921, 2850, 1615, 1489; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.71 (s, 1H, NCHN), 7.09 (app t, J 8.0 Hz, 1H, ArH), 6.77 (dd, J 8.0, 0.8 Hz, 1H, ArH) 6.53 (dd, J 8.0, 0.8 Hz, 1H, ArH) 4.34 (br s, 2H, NH2), 3.94–3.92 (m, 2H, CH2), 1.91–1.81 (m, 1H, CH), 1.73–1.64 (m, 4H, Ch 2×C(<u>H)H</u>), 1.33–0.95 (m, 6H, Ch 3×C(<u>H)H</u>); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 140.0, 138.6, 134.5, 125.2, 106.8, 99.7, 52.2, 38.2, 30.9, 29.8, 26.2, 25.6.; **LRMS** (LCMS-ESI) *m/z* calc. $C_{14}H_{19}N_3$ for 229.1 found 230.1 $[M+H]^+$.

1-(3,3,3-Trifluoropropyl)-1*H*-benzo[*d*]imidazol-4-amine **98j**



According to **General Procedure C**, $FeSO_4.7H_2O$ (1.4 mmol, 386 mg), citric acid (0.1 mmol, 25 mg), $NaBH_4$ (2.3 mmol, 86 mg) and nitrobenzimidazole **126c** (0.2 mmol, 60 mg) in H_2O (23 mL) were stirred overnight at room temperature. The *title compound* **98j** was isolated without any further purification as a brown solid (0.8 mmol, 45 mg, 85%). **MP** 71–72 °C; **IR** (ATR, cm^{-1}) 3388, 3316, 2957, 2920, 1636, 1599, 1495; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.76 (s, 1H, NC<u>H</u>N), 7.13 (app t, *J* 8.0 Hz, 1H, Ar<u>H</u>), 6.74 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 6.56 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 4.42–4.38 (m, 2H, NC<u>H$_2$</u>), 2.73–2.61 (m, 2H, C<u>H$_2$</u>CF_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 140.5, 139.6, 134.1, 133.2, 128.0 (q, $J_{C\text{-}F}$ 208 Hz), 124.8, 106.4, 98.8, 38.6–38.4 (m) 34.3 (q, $J_{C\text{-}F}$ 29 Hz); **^{19}F NMR** (376 MHz, $CDCl_3$) δ -65.58 (t, *J* 10.3 Hz, 3F, C<u>F$_3$</u>); **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{10}F_3N_3$ 229.1 found 230.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{10}H_{11}F_3N_3$ 230.0900 found 230.0901 $[M+H]^+$.

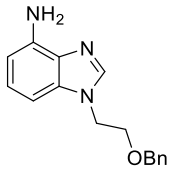
1-(Cyclopropylmethyl)-1*H*-benzo[*d*]imidazol-4-amine **98k**



.

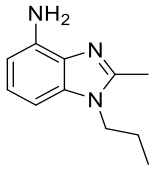
According to **General Procedure C** $FeSO_4{\cdot}7H_2O$ (1.16 mmol, 323 mg) citric acid (0.1 mmol, 20 mg), $NaBH_4$ (1.90 mmol, 71 mg) and nitrobenzimidazole compound **126d** (0.19 mmol, 42 mg) in H_2O (19 mL) were stirred at room temperature overnight. The *title compound* **98k** was isolated without any further purification as a purple solid (0.18 mmol, 34 mg, 96%). **MP** 78-80 °C; **IR** (ATR, cm^{-1}) 3415, 3357, 3151, 2960, 1597, 1491;**^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.87 (s, 1H, NC<u>H</u>N), 7.10 (app t, *J* 8.0 Hz, 1H, Ar<u>H</u>), 6.80 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 6.54 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 4.35 (br s, 2H, N<u>H$_2$</u>), 3.97–3.95 (m, 2H, NC<u>H$_2$</u>CH), 1.35–1.24 (m, 1H, NCH_2C<u>H</u>), 0.71–0.66 (m, 2H, Cp C(<u>H</u>)<u>H</u>), 0.42–0.38 (m, 2H, Cp C(<u>H</u>)<u>H</u>); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 140.5, 139.1, 135.0, 133.2, 124.1, 105.9, 99.7, 49.9, 11.1, 4.5; **LRMS** (LCMS-ESI) *m/z* calc. $C_{11}H_{13}N_3$ for 187.1 found 188.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{11}H_{14}N_3$ 188.1182 found 188.1181 $[M+H]^+$.

1-(2-(Benzyloxy)ethyl)-1*H*-benzo[*d*]imidazol-4-amine **98l**



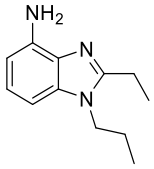
According to **General Procedure B**; 1 mol% Pd/C (0.02 mmol, 2 mg) and nitrobenzimidazole **126e** (0.2 mmol, 50 mg) in anhydrous EtOAc (2.0 mL) were stirred at room temperature overnight. The *title compound* **98l** was isolated without further purification as a brown solid (0.2 mmol, 45 mg, 100%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.87 (s, 1H, NC<u>H</u>N), 7.33–7.26 (m, 3H, 3×Ar<u>H</u>), 7.23–7.20 (m, 2H, 2×Ar<u>H</u>), 7.08 (app t, *J* 8.0 Hz, 1H, Ar<u>H</u>), 6.76 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 6.54 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 4.47 (s, 2H, OC<u>H_2</u>Ph), 4.34 (br s, 2H, N<u>H_2</u>), 4.30 (t, *J* 5.4 Hz, 2H, NC<u>H_2</u>CH_2OR), 3.79 (t, *J* 5.4 Hz, 2H, NCH_2C<u>H_2</u>OR); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 141.0, 138.5, 137.0, 134.1, 132.4, 128.0, 127.4, 127.1, 123.5, 105.3, 98.8, 72.9, 67.6, 44.7; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{16}H_{17}N_3O$ 267.1 found 268.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{16}H_{18}N_3O$ 268.1444 found 268.1446 $[M+H]^+$.

2-Methyl-1-propyl-1*H*-benzo[*d*]imidazol-4-amine **98m**



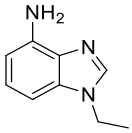
According to **General Procedure C;** $FeSO_4.7H_2O$ (1.1 mmol, 304 mg), citric acid (0.1 mmol, 19 mg), $NaBH_4$ (1.8 mmol, 68 mg) and nitrobenzimidazole **126f** (0.2 mmol, 40 mg) in H_2O (18 mL) were stirred overnight at room temperature. The *title compound* **98m** was isolated without any further purification as an orange oil (0.33 mmol, 57 mg, 77%). **IR** (ATR, cm^{-1}) 3348, 3330, 3311, 2960, 2874, 1621, 1403; **^{1}H NMR** (400 MHz, $CDCl_3$) δ.7.02 (app t, *J* 8.0 Hz, 1H, Ar<u>H</u>), 6.70 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>),6.50 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 4.29 (br s, 2H, N<u>H_2</u>), 4.02 (t, *J* 7.3 Hz, 2H, NC<u>H_2</u>CH_2CH_3), 2.59 (s, 3H, NCC<u>H_3</u>), 1.83 (app sextet, *J* 7.3 Hz, 2H, NCH_2C<u>H_2</u>CH_3) 0.96 (t, *J* 7.3 Hz, 3H, NCH_2CH_2C<u>H_3</u>); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 159.9, 149.3, 138.2, 136.0, 123.2, 106.1, 99.6, 45.7, 23.1, 13.9, 11.5; **LRMS** (LCMS-ESI) *m/z* calc.for $C_{11}H_{15}N_3$ 189.1 found 190.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{11}H_{16}N_3$ 190.1339 found 190.1337 $[M+H]^+$.

2-Ethyl-1-propyl-1*H*-benzo[*d*]imidazol-4-amine **98n**



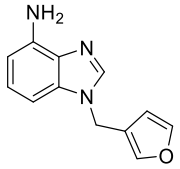
According to **General Procedure C**; $FeSO_4.7H_2O$ (1.4 mmol, 400 mg), citric acid (0.1 mmol, 25 mg), $NaBH_4$ (2.4 mmol, 91 mg) and nitrobenzimidazole **126g** (0.2 mmol, 56 mg) in H_2O (24 mL) were stirred overnight at room temperature. Upon subsequent work up, a second portion of $FeSO_4.7H_2O$ (1.4 mmol, 400 mg), citric acid (0.1 mmol, 25 mg) and $NaBH_4$ (2.4 mmol, 91 mg) were added and the crude material was stirred in H_2O (3.5 mL) at room temperature overnight. The *title compound* **98n** was isolated without any further purification as a dark red oil (0.21 mmol, 42 mg, 86%). **IR** (ATR, cm^{-1}) 3447, 3348, 3198, 2963, 2872, 1619, 1504; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.01 (app t, *J* 8.0 Hz, 1H, ArH), 6.70 (dd, *J* 8.0, 0.8 Hz, 1H, ArH), 6.49 (dd, *J* 8.0, 0.8 Hz, 1H, ArH), 4.29 (br s, 2H, NH_2), 4.01 (t, *J* 7.4 Hz, 2H, $NCH_2CH_2CH_3$), 2.88 (q, *J* 7.6 Hz, 2H $NCCH_2CH_3$), 1.83 (app sextet, *J* 7.4 Hz, 2H, $NCH_2CH_2CH_3$), 1.43 (t, *J* 7.6 Hz, 3H, $NCCH_2CH_3$) 0.97 (t, *J* 7.4 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 154.1, 138.3, 136.0, 131.9, 123.0, 106.0, 99.7, 45.5, 23.3, 21.0, 12.7, 11.6; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{12}H_{17}N_3$ 203.1 found 204.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{12}H_{18}N_3$ 204.1495 found 204.1495 $[M+H]^+$.

1-Ethyl-1*H*-benzo[*d*]imidazol-4-amine **98o**



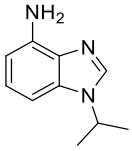
According to **General Procedure B**; 1 mol% Pd/C (0.02 mmol, 2 mg) and nitrobenzimidazole **126h** (0.25 mmol, 47 mg) in EtOAc (2.8 mL) were stirred at room temperature overnight. The *title compound* **98o** was isolated without further purification as a pink solid (0.24 mmol, 39 mg, 97%). **MP** 80–82 °C; **IR** (ATR, cm^{-1}) 3409, 3196, 3069, 2928, 1597, 1496; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.78 (s, 1H, NCHN), 7.10 (app t, *J* 8.0 Hz, 1H, ArH), 6.79 (dd, *J* 8.0, 0.8 Hz, 1H, ArH), 6.54 (dd, *J* 8.0, 0.8 Hz, 1H, ArH), 4.34 (br s, 2H, NH_2), 4.18 (q, *J* 7.3 Hz, 2H, CH_2) 1.52 (t, J 7.3 Hz, 3H, CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 140.2, 139.2, 134.5, 133.2, 124.1, 105.9, 99.6, 40.1, 15.4; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_{11}N_3$ 161.1 found 162.1 $[M+H]^+$.

1-(Furan-3-ylmethyl)-1*H*-benzo[*d*]imidazol-4-amine **98p**



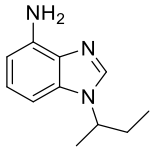
According to **General Procedure B**; 1 mol% Pd/C (0.02 mmol, 2 mg) and nitrobenzimidazole **126i** (0.21 mmol, 50 mg) in EtOAc (2.3 mL) were stirred at room temperature for 4 hours. The *title compound* **98p** was isolated without further purification as a brown solid (0.21 mmol, 45 mg, 100 %). **MP** 107–108 °C; **IR** (ATR, cm^{-1}) 3395, 3290, 3192, 3082, 2917, 1595, 1491; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.81 (s, 1H, NC<u>H</u>N), 7.42–7.40 (m, 2H, 2×Ar<u>H</u>), 7.10 (app t, *J* 8.0 Hz, 1H, Ar<u>H</u>) 6.77 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 6.55 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>) 6.30–6.29 (m, 1H, Ar<u>H</u>), 5.15 (s, 2H, C<u>H_2</u>), 4.43 (br s, 2H, N<u>H_2</u>); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 144.2, 140.6, 140.5, 139.2, 134.7, 133.2, 124.4, 120.5, 110.0, 106.1, 99.7, 40.3; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{12}H_{11}N_3O$ 213.1 found 214.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{12}H_{12}N_3O$ 214.0975 found 214.0975 $[M+H]^+$.

1-*iso*Propyl-1*H*-benzo[*d*]imidazol-4-amine **98q**



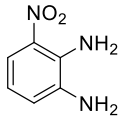
According to **General Procedure B**; 1 mol% Pd/C (0.04 mmol, 4 mg) and nitrobenzimidazole **126j** (0.4 mmol, 77 mg) in EtOAc (5.0 mL) were stirred at room temperature overnight. The *title compound* **98q** was isolated without further purification as a dark brown oil (0.3 mmol, 45 mg, 86%). **IR** (ATR, cm^{-1}) 3445, 3352, 2976, 2930, 1619, 1487; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.86 (s, 1H, NC<u>H</u>N), 7.09 (app t, *J* 8.0 Hz, 1H, Ar<u>H</u>), 6.81 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 6.53 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 4.57 (septet, *J* 6.8 Hz, 1H, C<u>H</u>), 4.35 (br s, 2H, N<u>H_2</u>), 1.61 (d, *J* 6.8 Hz, 6H, 2×C<u>H_3</u>); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 139.2, 138.0, 134.1, 132.9, 124.1, 106.0, 100.2, 48.0, 22.7; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{13}N_3$ 175.1 found 176.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{10}H_{14}N_3$ 176.1182 found 176.1180 $[M+H]^+$.

1-(*sec*-Butyl)-1*H*-benzo[*d*]imidazol-4-amine **98r**



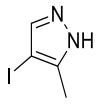
According to **General Procedure B**; 1 mol% Pd/C (0.02 mmol, 2 mg) and nitrobenzimidazole **126k** (0.24 mmol, 50 mg) in EtOAc (2.7 mL) were stirred at room temperature for 4 hours. The *title compound* **98r** was isolated without further purification as a red oil (0.22 mmol, 42 mg, 92%). **IR** (ATR, cm^{-1}) 2956, 2924, 2854, 1597, 1489; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.82 (s, 1H, NCHN), 7.07 (app t, *J* 8.0 Hz, 1H, ArH), 6.80 (dd, *J* 8.0. 0.5 Hz, 1H, ArH), 6.52 (dd, *J* 8.0, 0.5 Hz, 1H, ArH), 4.36–4.26 (m, 1H. CH), 2.06–1.86 (m, 2H, CH(H)CH_3), 1.59 (d, *J* 7.0 Hz, 3H, CHCH_3) 0.87 (t, J 7.0 Hz, 3H, CH_2CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 139.2, 138.8, 134.3, 133.0, 124.0, 105.9, 100.3, 54.1, 29.6, 20.7, 10.9; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{11}H_{15}N_3$ 189.1 found 190.1 $[M+H]^+$.

3-Nitrobenzene-1,2-diamine **100**[155]



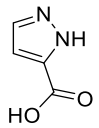
To a round bottomed flask was added nitroselendiazole **120** (31.2 mmol, 7.1 g, 1.0 eq.) in conc. HCl (310 mL, 0.1 M). To the resulting mixture was added 57% w/w HI (72.9 mmol, 6.9 mL, 4.0 eq.) at room temperature. The resulting mixture was then heated to 45 °C and allowed to stir at this temperature for 3 days. Upon completion, the mixture was cooled to room temperature and a 15% w/w solution of $Na_2S_2O_5$ (100 mL, 0.2 M) was added. The resulting mixture was heated to 90 °C and allowed to stir at this temperature for 1 hour. The resulting suspension was hot filtered and the filtrate was cooled to <5 °C in an ice bath. The filtrate was then collected and basified to pH 8 with a 40% NaOH solution. The aqueous was extracted with EtOAc (7×50 mL) and the organics were dried over $MgSO_4$ and concentrated *in vacuo* to provide *title compound* **100** as a red solid (15.6 mmol, 2.35 g, 50%). **MP lit**[164] 158–159 °C; **MP** 155—157 °C ; **IR** (ATR, cm^{-1}) 3459, 3438, 3312, 3100, 1642, 1521; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 7.30 (dd, *J* 7.5, 1.4 Hz, 1H, ArH), 6.89 (br s, 2H, NH_2), 6.77, (dd. *J* 7.5, 1.4 Hz, 1H, ArH) 6.45 (app t, *J* 7.5 Hz, 1H, ArH), 5.22 (br s, 2H, NH_2); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 137.6, 135.1, 131.1, 117.6, 115.7, 113.0; **LRMS** (LCMS-ESI) m/z calc. for $C_6H_7N_3O_2$ 153.1 found 154.1 $[M+H]^+$.

4-Iodo-5-methyl-1*H*-pyrazole **102b**[165]



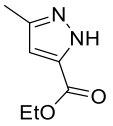
5-Methyl-*1H*-pyrazole **99** (1.2 mmol, 100 mg, 1.0 eq) was dissolved in 50% v/v H_2SO_4 (0.4 mL, 3.4 M), cooled to 0 °C and *N*-iodosuccinimide (NIS) (1.5 mmol, 329 mg, 1.2 eq) was added. The resulting mixture was allowed to stir for 10 minutes at 0 °C then 3 hours at room temperature. The mixture was then dispersed in water (2 mL) and allowed to stir for a further 48 hours. The aqueous was extracted with EtOAc (3x10 mL) and the combined organics were washed with brine (10 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo* to provide the *title compound* **102b** as a yellow solid (1.1 mmol, 236 mg, 93%). **MP lit**[166] 110–111 °C; **MP** 96–98°C; **IR** (ATR, cm^{-1}) 3350, 3155, 2504, 1690; **^{1}H NMR**, (400 MHz, DMSO-d_6) δ 12.82 (br s, 1H, NH), 7.59 (br s, 1H, ArH), 2.16 (s, 3H, CH_3); **^{13}C NMR**, (101 MHz, DMSO-d_6) δ 179.4, 59.2, 29.5, 11.7; **LRMS** (LCMS-ESI) *m/z* calc. for $C_4H_5IN_2$ 208.0 found 208.9 $[M+H]^+$.

1*H*-pyrazole-5-carboxylic acid **103**[167]



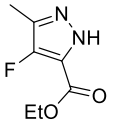
To a solution of 5-methyl-*1H*-pyrazole **99** (122 mmol, 10.0 g, 1.0 eq) in H_2O (608 mL, 0.2 M) was added potassium permanganate ($KMnO_4$) (365 mmol, 57.6 g, 3.0 eq). The resulting mixture was heated to reflux and stirred for 72 hours. Upon completion, the solids were filtered and the volume of the filtrate reduced to approximately a 10th *in vacuo*. The filtrate was then acidified with 0.2 M HCl to pH 2. A precipitate formed and was filtered to provide the *title compound* **103** as a white solid (83 mmol, 9.5 g, 69%). **MP lit**[168] 216–217 °C **MP** 213–215 °C; **IR** (ATR, cm^{-1}) 3241, 3128, 2936, 1694; **^{1}H NMR**, (400 MHz, DMSO-d_6) δ 13.14 (s, 2H, COOH and NH), 7.74 (d, *J* 2.0 Hz, 1H, ArH), 6.71 (d, *J* 2.2 Hz, 1H, ArH); **^{13}C NMR**, (101 MHz, DMSO-d_6) δ 162.6, 140.9, 122.9, 107.8; **LRMS** (LCMS-ESI) *m/z* calc. for $C_4H_4N_2O_2$ 112.0 found 113.1 $[M+H]^+$.

Ethyl 3-methyl-1*H*-pyrazole-5-carboxylate **104**[153]



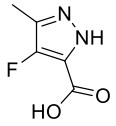
To a solution of hydrazine monohydrate (4.7 mmol, 230 µL, 1.5 eq) in EtOH (20 mL, 0.1 M) was added ethyl 2,4-dioxopentanoate (3.2 mmol, 500 mg, 1.0 eq.) portion-wise. The resulting mixture was heated to reflux for 3 hours. Upon completion the mixture was concentrated *in vacuo*. The resulting residue was dissolved in EtOAc (10 mL) and washed with water (3×10 mL) and brine (10 mL). The organics were then dried over $MgSO_4$, filtered and concentrated *in vacuo* to provide *title compound* **86** as a dark orange solid (2.5 mmol, 392 mg, 80%). **MP lit**[169] 80–82°C**; MP** 80–82 °C; **IR** (ATR, cm^{-1}) 3208, 2999, 1722, 1580; **^{1}H NMR**, (500 MHz, $CDCl_3$) δ 10.39 (br s, 1H, NH), 6.60 (s, 1H, ArH), 4.38 (q, *J* 7.1 Hz, 2H, CH_2), 2.35 (s, 3H, CH_3), 1.39 (t, *J* 7.1 Hz, 3H, CH_2CH_3); **^{13}C NMR**, (101 MHz, $CDCl_3$) δ 162.1, 143.3, 141.8, 107.4, 61.0, 14.4, 11.6; **LRMS** (LCMS-ESI) *m/z* calc. for $C_7H_{10}N_2O_2$ 154.1 found 155.2 $[M+H]^+$.

Ethyl 4-fluoro-3-methyl-1*H*-pyrazole-5-carboxylate **105**[153]



To a solution of **104** (0.7 mmol, 100 mg, 1.0 eq) in MeCN (22 mL, 0.03 M) was added Selectfluor© (0.7 mmol, 235 mg, 1.0 eq) and the resulting mixture was heated to reflux and stirred overnight. The mixture was concentrated *in vacuo* before dilution with CH_2Cl_2 (20 mL). The organics were washed with 1 M HCl solution (2×10 mL) and brine (10 mL) followed by drying over $MgSO_4$, filtration and concentration *in vacuo*. The crude material was purified *via* flash column chromatography in 3:2 hexanes/EtOAc to afford the *title compound* **105** as a white solid (0.2 mmol, 30 mg, 27 %). **MP** 109—111 °C; **IR** (ATR, cm^{-1}) 3230, 3156, 1956, 2882, 1733, 1285, 1160; **^{19}F** ($CDCl_3$, 400 MHz) δ -169.9 **^{1}H NMR**, (400 MHz, $CDCl_3$) δ 10.72 (br s, 1H, NH), 4.41 (q, *J* 6.6 Hz, 2H, OCH_2CH_3), 2.89 (d, *J* 4.0 Hz, 3H, $NCCH_3$), 1.39 (t, *J* 8.0 Hz, 3H, OCH_2CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 159.1, 149.1, 146.6, 61.6, 29.9, 14.4, 9.3; **^{19}F NMR** (376 MHz, $CDCl_3$) δ -169.32 (s,1F); **LRMS** (LCMS-ESI) *m/z* calc. for $C_7H_9FN_2O_2$ 172.1 found 173.1 $[M+H]^+$.

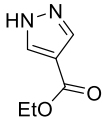
4-Fluoro-3-methyl-1*H*-pyrazole-5-carboxylic acid **106**[153]



According to **General Procedure H**, ester **105** (0.5 mmol, 85 mg) was dissolved in 1:5:1 MeOH/THF/1M LiOH (17 mL) and heated to 80 °C for 6 hours. The resulting aqueous solution was extracted with EtOAc (3×10 mL) and the combined organics were dried over $MgSO_4$, filtered and concentrated *in vacuo* to afford the *title compound* **106** as an off-white solid (0.4 mmol, 53 mg, 75%). **MP** decomp. >190 °C; **IR** (ATR, cm^{-1}) 3213, 2920, 2951, 1713; **^{1}H NMR**, ($CDCl_3$, 400 MHz) δ 13.18 (br s, 1H, COOH), 2.17 (d, *J* 4.0 Hz, 3H, CH_3); **^{13}C NMR**, ($CDCl_3$, 100 MHz) δ 160.6, 148.2, 145.7, 29.0, 8.3; **^{19}F** ($CDCl_3$, 400 MHz) δ -71.1 (s, 1F); **LRMS** (LCMS-ESI) *m/z* calc. for $C_5H_5FN_2O_2$ 144.0 found 144.9 $[M+H]^+$.

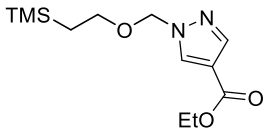
9.5.2 Alternative head group heterocycles

Ethyl 1*H*-pyrazole-4-carboxylate **112**[167]



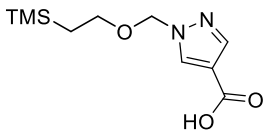
To a round bottomed flask was added 1*H*-pyrazole-4-carboxylic acid **111** (8.9 mmol, 1.0 g, 1.0 eq.) and EtOH (15 mL). The solution was cooled to 0 °C and $SOCl_2$ (53.5 mmol, 3.9 mL, 6.0 eq.) was added dropwise. The resulting mixture was then heated to reflux overnight. Upon completion, the reaction mixture was allowed to cool to room temperature and concentrated *in vacuo.* The resulting residue was taken up in EtOAc (20 mL) and washed with saturated $NaHCO_3$ solution (20 mL) and brine (20 mL). The organics were dried over $MgSO_4$, filtered and concentrated *in vacuo* to afford the *title compound* **112** as an orange solid (6.7 mmol, 935 mg, 75%) which was used in the next step without further purification. **MP lit**[170] 72–74°C; **MP** 71-73°C; **IR** (ATR, cm^{-1}) 3125, 2943, 2831, 1723, 1705, 1526; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 11.10 (br s, 1H, NH), 8.07 (s, 2H, 2×ArH), 4.32 (q, *J* 7.2 Hz, 2H, CH_2), 1.36 (t, *J* 7.2 Hz, 3H, CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 163.3, 136.9, 115.4, 60.5, 14.5; **LRMS** (LCMS-ESI) *m/z* calc. for $C_6H_8N_2O_2$ 140.1 found 141.1 $[M+H]^+$.

Ethyl 1-((2-(trimethylsilyl)ethoxy)methyl)-1*H*-pyrazole-4-carboxylate **113**[171]



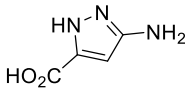
To a flame dried round bottomed flask purged with argon was added NaH (5.7 mmol, 137 mg, 1.6 eq.) and anhydrous THF (8.2 mL). The resulting solution was cooled to 0 °C and treated dropwise with a solution of pyrazole ester **112** (3.6 mmol, 500 mg, 1.0 eq.) in anhydrous THF (2.0 mL). The resulting mixture was allowed to stir at room temperature for 2 hours, before cooling to 0 °C. A solution of 2-(trimethylsilyl)ethoxymethyl chloride (SEM-Cl) (4.3 mmol, 0.8 mL, 1.2 eq.) in anhydrous THF (2.0 mL) was then added dropwise and the resulting mixture was allowed to warm to room temperature overnight. Upon completion, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3×20 mL). The combined organics were dried over $MgSO_4$, filtered and concentrated *in vacuo.* Further purification *via* flash column chromatography 85:15 petroleum ether 40–60/EtOAc afforded the *title compound* **113** as a colourless oil (2.8 mmol, 758 mg, 79%). **IR** (ATR, cm^{-1}) 3131, 2958, 2894, 1716, 1556; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.05 (s, 1H, ArH), 7.94 (s, 1H, ArH), 5.43 (s, 2H, NCH_2OR), 4.31 (q, *J* 7.1 Hz, 2H, OCH_2CH_3) 3.60–3.56 (m, 2H, OCH_2CH_2Si), 1.35 (t, *J* 7.1 Hz, 3H, OCH_2CH_3), 0.94–0.87 (m, 2H, OCH_2CH_2Si), -0.02 (s, 9H, $Si(CH_3)_3$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 163.1, 141.4, 133.0, 116.5, 80.8, 67.4, 60.5, 17.9, 14.5, -1.3; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{12}H_{22}N_2O_3Si$ 270.1 found 271.1 $[M+H]^+$.

1-((2-(Trimethylsilyl)ethoxy)methyl)-1H-pyrazole-4-carboxylic acid **114**[172]



According to **General Procedure H**, ester **113** (0.92 mmol, 250 mg) was dissolved in 1:5:1 MeOH/THF/1M LiOH (31 mL) and heated to 80 °C overnight. A white precipitate formed and was filtered to afford the *title compound* **114** as a white solid (0.77 mmol, 187 mg, 84%). **MP** 104–106 °C; **IR** (ATR, cm^{-1}) 3125. 3090, 2952, 2541, 1699, 1565; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.12 (s, 1H, ArH), 7.99 (s, 1H, ArH), 5.45 (s, 2H, NCH_2O), 3.62–3.58 (m, 2H, OCH_2CH_2Si), 0.95–0.91 (m, 2H, OCH_2CH_2Si), -0.01 (s, 9H, $Si(CH_3)_3$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 168.1, 142.0, 133.9, 115.5, 80.8, 67.6, 17.9, -1.3; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{18}N_2O_3Si$ 242.1 found 243.1 $[M+H]^+$.

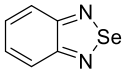
3-Amino-1*H*-pyrazole-5-carboxylic acid **118**[173]



According to **General Procedure B**, 10 mol% Pd/C (0.16 mmol, 17 mg) and 3-nitro-1*H*-pyrazole-5-carboxylic acid **116** (1.59 mmol, 250 mg, 1.0 eq.) in MeOH (18 mL) were stirred at room temperature overnight. The *title compound* **118** was used in the next step without further purification as a purple solid (1.38 mmol, 175 mg, 87%). **MP lit:** 245–246 °C; **MP** decomp. >240 °C; **IR** (ATR, cm^{-1}) 3455, 3322, 3154, 2926, 2853, 1726, 1619; **^{1}H NMR** (400 MHz, DMSO) δ 5.76 (s. 1H, CH); **^{13}C NMR** (101 MHz, DMSO) δ 162.1, 152.9, 138.3, 92.2; **LRMS** (LCMS-ESI) *m/z* calc. for $C_4H_5N_3O_2$ 127.0 found 128.2 $[M+H]^+$.

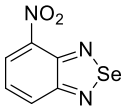
9.5.3 Synthesis of tail group intermediates

Benzo[*c*][1,2,5]selenadiazole **119**[155]



To a round bottomed flask was added *o*-phenylenediamine **101** (92.5 mmol, 10 g, 1.0 eq) and SeO_2 (101.7 mmol, 11.3 g, 1.1 eq) in EtOH (58 mL, 1.6 M) at room temperature. The resulting mixture was heated to reflux and stirred at this temperature for 2 hours. Upon cooling to room temperature a precipitate was formed. The solids were filtered and collected to afford *title compound* **119** as a beige solid, which was used in the next step without further purification (92.4 mmol, 17.3 g, 99%) **MP lit**: 75 °C; **MP** 77–78 °C; **IR** (ATR, cm^{-1}) 3081, 3037, 1506, 1476; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 7.87–7.83 (m, 2H, 2×ArH), 7.56–7.51 (m, 2H, 2×ArH); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 160.4, 150.4, 141.0, 130.0, 127.3, 127.0; **LRMS** (LCMS-ESI) m/z calc. $C_6H_4N_2Se$ 184.0 found 183.1 $[M-H]^-$.

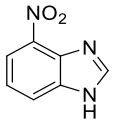
4-Nitrobenzo[*c*][1,2,5]selenadiazole **120**[155]



To a round bottomed flask was added selenadiazole **119** (57.3 mmol, 10.7 g) in H_2SO_4 (30 mL, 1.9 M). The mixture was cooled to 0 °C before the addition of a 1:1 mixture of 70% HNO_3 and conc. H_2SO_4 (14 mL, 4.1 M). The reaction mixture was allowed to

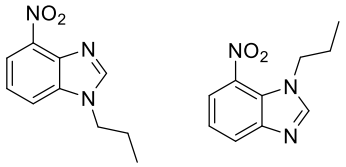
warm to room temperature overnight. The resulting mixture was poured onto ice and the resulting yellow precipitate was collected *via* filtration to provide crude material. Further purification *via* recrystallisation in toluene afforded the *title compound* **120** as a yellow solid (40.0 mmol, 9.1 g, 70%). **MP lit**[174] 214—216°C; **MP** 218–220°C; **IR** (ATR, cm^{-1}) 3092, 2850, 2826, 1519, 15004, 1429; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.47 (dd, *J* 7.3, 1.0 Hz, 1H, ArH), 8.29 (dd, *J* 7.3, 1.0 Hz, 1H, ArH), 7.75 (app t, *J* 7.3 Hz, 1H, ArH); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 160.4, 150.4, 141.0, 130.0, 127.3, 127.0; **LRMS** (LCMS-ESI) m/z calc. $C_6H_3N_3O_2Se$ 228.9 found 229.1 $[M]^-$.

4-Nitro-1*H*-benzimidazole **121**[155]



According to **General Procedure G**; 3-nitro-1,2-benzenediamine **100** (19.6 mmol, 3.0 g), triethylorthoformate (58.8 mmol, 9.8 mL) and *p*-TSOH (2.0 mmol, 344 mg) in toluene (45 mL) were stirred at reflux for 2 hours. Upon completion, the resulting suspension was allowed to cool to room temperature before further cooling to 0 °C. The resulting solids were filtered and washed with toluene (50 mL), to afford *the title compound* **121** as a brown solid (18.7 mmol, 3.1 g, 96%). **121** was used in the next step without any further purification. **MP** 242–244 °C; **IR** (ATR, cm^{-1}) 3053, 2768, 1550, 1489, 1338, 1294; **^{1}H NMR**, (400 MHz, DMSO-d_6) δ 8.48 (s, 1H, NCHN), 8.17–8.16 (m, 2H, 2×ArH), 7.45–7.41 (m, 1H, ArH); **^{13}C NMR**, (101 MHz, DMSO-d_6) δ 145.2, 133.7, 127.8, 126.5, 125.5, 121.4, 119.1; **LRMS** (LCMS-ESI) *m/z* calc. for $C_7H_5N_3O_2$ 163.0 found 164.1 $[M+H]^+$

4-Nitro-1-propyl-1*H*-benzo[*d*]imidazole **122a** and 7-nitro-1-propyl-1*H*-benzo[*d*]imidazole **123a**

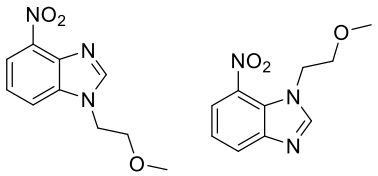


122a **123a**

According to **General Procedure A**; nitrobenzimidazole **121** (3.1 mmol, 500 mg), K_2CO_3 (6.2 mmol, 849 mg) and 1-bromopropane (4.6 mmol, 0.42 mL) were added to a microwave vial in DMF (12 mL) at room temperature overnight Further purification *via* flash column chromatography in 3:2 petroleum ether 40–60/EtOAc isolated *title*

regioisomer **122a** (1.3 mmol, 260 mg, 41%) and **123a,** (0.6 mmol, 128 mg, 20%) as orange oils and respectively. **122a**: **IR** (ATR, cm^{-1}) 2964, 1514, 1338[115]; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.17 (dd, *J* 8.0, 0.9 Hz, 1H, ArH), 8.14 (s, 1H, NCHNR), 7.73 (dd, *J* 8.0, 0.9 Hz, 1H, ArH), 7.41 (app t, *J* 8.0 Hz, 1H, ArH), 4.23 (t, *J* 7.4 Hz, 2H, $NCH_2CH_2CH_3$), 1.96 (app sextet, J 7.4 Hz, 2H, $NCH_2CH_2CH_3$), 0.99 (t, *J* 7.4 Hz, 3H, CH_3) **^{13}C NMR**, (101 MHz, $CDCl_3$,) δ 146.4, 139.7, 137.5, 136.8, 122.2, 119.5, 116.3, 47.4, 23.4, 11.4; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{11}N_3O_2$ 205.2 found 206.0 $[M+H]^+$.**123a: ^{1}H NMR**, (500 MHz, $CDCl_3$) δ 8.10 (dd, *J* 8.0, 0.8 Hz, 1H, ArH), 8.03, (m, 1H, ArH), 7.99 (s, 1H, NCHNR), 7.36 (app t, *J* 8.0 Hz, 1H, ArH), 4.42 (t, *J* 7.4 Hz, 2H, $NCH_2CH_2CH_3$), 1.72 (app sextet, *J* 7.4 Hz, 2H, $NCH_2CH_2CH_3$), 0.88 (t, *J* 7.4 Hz, 3H, CH_3); **^{13}C NMR**, (101 MHz, $CDCl_3$) δ 148.2, 147.5, 136.4, 127.3, 125.9, 121.5, 121.1, 50.5, 24.5, 11.0; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{11}N_3O_2$ 205.1 found 206.2 $[M+H]^+$.

1-(2-methoxyethyl)-4-nitro-1*H*-benzo[*d*]imidazole **122b** and 1-(2-methoxyethyl)-7-nitro-1*H*-benzo[*d*]imidazole **123b**

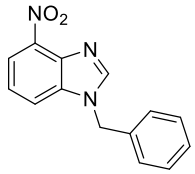


122b **123b**

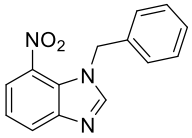
According to **General Procedure A**; nitrobenzimidazole **121** (3.1 mmol, 500 mg), K_2CO_3 (6.2 mmol, 849 mg) and 1-bromo-2-methoxyethane (4.6 mmol, 0.43 mL) were added to a microwave vial in DMF (12 mL) at room temperature overnight. The crude material was purified *via* column chromatography in 100% EtOAc to afford the isolated *title regioisomer* **122b**, (0.8 mmol, 166 mg, 25%) as a brown oil, and **123b**, (0.9 mmol, 202 mg, 30%) as a yellow solid and respectively. **122b: IR** (ATR, cm^{-1}) 3100, 2925, 2843, 1526, 1500; **^{1}H NMR**, (400 MHz, $CDCl_3$) δ 8.18 (s, 1H, NCHNR), 8.08 (d, *J* 8.1 Hz, 1H, ArH), 7.76 (d, *J* 8.1 Hz, 1H, ArH), 7.34 (app t, *J* 8.1 Hz, 1H, ArH), 4.40 (t, *J* 5.0 Hz, 2H, $NCH_2CH_2OCH_3$), 3.71 (t, *J* 5.0 Hz, 2H, $NCH_2CH_2OCH_3$), 3.27 (s, 1H, OCH_3) **^{13}C NMR** (101 MHz, $CDCl_3$) δ 147.0, 139.2, 137.1, 136.8, 122.0, 119.2, 116.6, 70.7, 59.0, 45.6; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{11}N_3O_3$ 221.1 found 222.0; **HRMS** (ESI, +ve) *m/z* calc. for $C_{10}H_{12}N_3O_3$ 222.0873 found 222.0874 $[M+H]^+$; **123b: MP** 83–85 °C; **IR** (ATR, cm^{-1}) 3096, 2921, 1523, 1426, 1357; **^{1}H NMR** (400 MHz, $CDCl_3$), δ 8.09–8.06 (m, 1H, ArH), 8.03 (s, 1H, NCHNR), 8.00–7.98 (m, 1H, ArH), 7.35–7.31 (m, 1H, ArH), 4.61 (t, *J* 4.7 Hz, 2H, $NCH_2CH_2OCH_3$) 3.62 (t, *J* 4.7 Hz, 2H,

$NCH_2C\underline{H}_2OCH_3$), 3.22 (s, 3H, $OC\underline{H}_3$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 148.1, 147.9, 142.8, 136.7, 127.1, 121.4, 121.0, 71.7, 59.3, 48.7; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{11}N_3O_3$ 221.1 found 222.0 $[M-H]^-$.

1-Benzyl-4-nitro-1*H*-benzo[*d*]imidazole **122c** and 1-benzyl-7-nitro-1*H*-benzo[*d*]imidazole **123c**



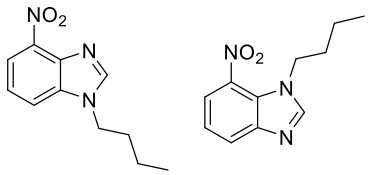
122c



123c

According to **General Procedure A**; nitrobenzimidazole **121** (3.1 mmol, 500 mg), K_2CO_3 (6.2 mmol, 849 mg) and benzyl bromide (4.6 mmol, 0.56 mL) were added to a microwave vial in DMF (12 mL) at room temperature overnight. Further purification *via* column chromatography in 3:2 EtOAc/hexanes isolated the *title regioisomer* **122c**, (1.4 mmol, 363 mg, 47%) as a beige solid and **123c**, (0.6 mmol, 143 mg, 18%) as a brown solid, and respectively. **122c: MP** 98–100 °C; **IR** (ATR, cm^{-1}) 3103, 2917, 1571, 1521, 1495, 1299; **^{1}H NMR**, (500 MHz, $CDCl_3$) δ 8.21 (s, 1H, NC$\underline{H}$NR), 8.17 (d, *J* 8.1 Hz, 1H, Ar$\underline{H}$), 7.60 (d, *J* 8.1 Hz, 1H, Ar$\underline{H}$), 7.39–7.31 (m, 4H, 4×Ar$\underline{H}$), 7.19–7.17 (m, 2H, 2×Ar$\underline{H}$), 5.45 (s, 2H, $C\underline{H}_2$); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 146.6, 139.7, 137.8, 136.7, 134.5, 129.5, 129.0, 127.4, 122.5, 119.7, 116.8, 49.6; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{11}N_3O_2$ 253.1 found 254.0 $[M+H]^+$; **123c: MP** 175–177 °C; **IR** (ATR, cm^{-1}) 3032, 2958, 1578, 1491, 1346; **^{1}H NMR**, (500 MHz $CDCl_3$) δ 8.13–8.12 (m, 1H, Ar$\underline{H}$), 8.11 (s, 1H, NC$\underline{H}$NR), 7.92 (m, 1H, Ar$\underline{H}$), 7.37–7.33 (m, 1H, Ar$\underline{H}$), 7.30–7.24 (m, 3H, 3×Ar$\underline{H}$), 7.00–6.97 (m, 2H, 2×Ar$\underline{H}$), 5.67 (s, 2H, $C\underline{H}_2$); **^{13}C NMR**, (125 MHz, $CDCl_3$) δ 148.2, 147.8, 136.9, 134.6, 129.9, 129.2, 128.5, 127.6, 127.3, 126.9, 126.0, 121.8, 121.1, 52.1; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{11}N_3O_2$ 253.1 found 254.0 $[M+H]^+$; **HRMS** (ESI, +ve) *m/z* calc. for $C_{14}H_{12}N_3O_2$ 254.0927 found 254.0927 $[M+H]^+$

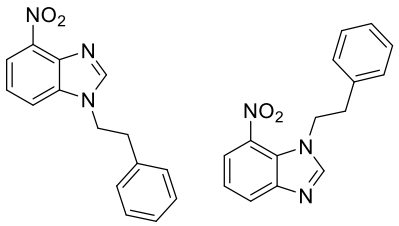
1-butyl-4-nitro-1*H*-benzo[*d*]imidazole **122d** and 1-butyl-7-nitro-1*H*-benzo[*d*]imidazole **123d**



122d **123d**

According to **General Procedure A**; nitrobenzimidazole **118** (1.8 mmol, 300 mg), K_2CO_3 (3.7 mmol, 509 mg) and 1-butyl bromide (2.8 mmol, 0.3 mL) were added to a microwave vial in DMF (7 mL) at room temperature overnight. Further purification *via* flash column chromatography in 3:1:0.1 EtOAc/petroleum ether 40–60/MeOH isolated *title regioisomer* **122d,** (1.1 mmol, 240 mg, 59%) as a pale yellow solid and **123d**, (0.5 mmol, 116 mg, 29%) as a brown oil and respectively. **122d: MP** 74–76 °C; **IR** (ATR, cm^{-1}) 3373, 3094, 2927, 1625, 1579, 1344, 1333; **^{1}H NMR**, (400 MHz, $CDCl_3$) δ 8.13 (dd, *J* 8.1, 0.9 Hz, 1H, ArH), 8.12 (s, 1H, NCHNR), 7.72 (dd, *J* 8.1, 0.9 Hz, 1H, ArH), 7.38 (app t, *J* 8.1 Hz, 1H, ArH), 4.27–4.23 (m, 2H, $NCH_2CH_2CH_2CH_3$), 1.92–1.84 (m, 2H, $NCH_2CH_2CH_2CH_3$), 1.41-1.32 (m, 2H, $NCH_2CH_2CH_2CH_3$), 0.97–0.94 (m, 3H, CH_3); **^{13}C NMR**, (101 MHz, $CDCl_3$) δ 146.3, 139.5, 137.4, 136.7, 122.1, 119.4, 116.4, 45.5, 31.9, 20.1, 13.6; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{11}H_{13}N_3O_2$ 219.1 found 220.2 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{11}H_{14}N_3O_2$ 220.1081 found 220.1080 $[M+H]^+$; **123d: IR** (ATR, cm^{-1}) 3087, 2957, 2871, 1623, 1501, 1346; **^{1}H NMR**, (400 MHz, $CDCl_3$) δ 8.10 (dd, *J* 8.0, 1.0 Hz, 1H, ArH), 8.03 (dd, *J* 8.0, 1.0 Hz, 1H, ArH), 7.98 (s, 1H, NCHNR), 7.36 (app t, *J* 8.0 Hz, 1H, ArH), 4.46–4.42 (m, 2H, $NCH_2CH_2CH_2CH_3$), 1.68-1.64 (m, 2H, $NCH_2CH_2CH_2CH_3$), 1.31–1.25 (m, 2H, $NCH_2CH_2CH_2CH_3$), 0.92–0.89 (m, 3H, CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 148.1, 147.4, 136.6, 127.3, 125.9, 121.5, 121.0, 48.7, 33.2, 19.8, 13.7; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{11}H_{13}N_3O_2$ 219.1 found 220.2 $[M+H]^+$; **HRMS** (ESI, +ve) *m/z* calc. for $C_{11}H_{14}N_3O_2$ 220.1081 found 220.1080 $[M+H]^+$

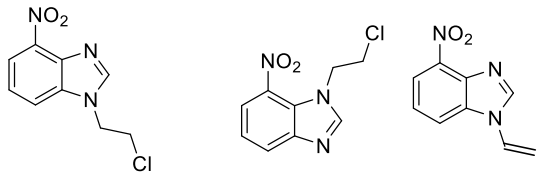
4-Nitro-1-phenethyl-1*H*-benzo[*d*]imidazole **122e** and 7-nitro-1-phenethyl-1*H*-benzo[*d*]imidazole **123e**



122e **123e**

According to **General Procedure A**; nitrobenzimidazole **121** (1.8 mmol, 300 mg), K_2CO_3 (3.7 mmol, 509 mg) and (2-bromoethyl)benzene (2.8 mmol, 0.4 mL) were added to a microwave vial in DMF (7 mL) at room temperature overnight. Further purification *via* flash column chromatography in 3:1:0.1 EtOAc/petroleum ether 40–60/MeOH isolated *title regioisomer* **122e,** (1.1 mmol, 240 mg, 59%) as a beige solid and **123e** as orange crystals (0.5 mmol, 116 mg, 29%).**122e: MP** 144–146°C; **IR** (ATR, cm^{-1}) 3102, 2933, 1521, 1497, 1350; **^{1}H NMR**, (400 MHz, $CDCl_3$), δ 8.13 (dd, *J* 8.1, 0.8 Hz, 1H, ArH), 7.84 (s, 1H, NCHNR), 7.60 (dd, *J* 8.1, 0.8 Hz, 1H, ArH), 7.35 (app t, *J* 8.1 Hz, 1H, ArH), 7.27–7.20 (m, 3H, 3×ArH), 6.99–6.96 (m, 2H, 2×ArH), 4.49 (t, *J* 6.8 Hz, 2H, NCH_2CH_2Ar), 3.15 (t, *J* 6.8 Hz, 2H, NCH_2CH_2Ar); **^{13}C NMR**, (101 MHz, $CDCl_3$) δ 146.4, 139.6, 137.3, 136.9, 136.5, 129.2, 128.7, 127.5, 122.2, 119.5, 116.2, 47.3, 36.5; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{15}H_{13}N_3O_2$ 267.1 found 268.2 $[M+H]^+$; **HRMS** (ESI, +ve) *m/z* calc. for $C_{15}H_{14}N_3O_2$ 268.1081 found 268.1083 $[M+H]^+$; **123e: MP** 84–86°C; **IR** (ATR, cm^{-1}) 3094, 2923, 1534, 1510, 1313; **^{1}H NMR**, (400 MHz, $CDCl_3$) δ 8.10-8.06 (m, 2H, 2×ArH), 7.69 (s, 1H, NCHNR), 7.39–7.34 (m, 1H, ArH), 7.25–7.19 (m, 3H, 3×ArH), 6.99-6.97 (m, 2H, 2×ArH), 4.69 (t, *J* 7.2 Hz, 2H, NCH_2CH_2Ar), 3.00 (t, *J* 7.2 Hz, 2H, NCH_2CH_2Ar); **^{13}C NMR**, (101 MHz, $CDCl_3$) δ 148.1, 147.5, 137.1, 136.5, 129.0, 128.2, 127.9, 127.3, 125.9, 121.6, 121.2, 50.5, 37.8; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{15}H_{13}N_3O_2$ 267.1 found 268.2 $[M+H]^+$; **HRMS** (ESI, +ve) *m/z* calc. for $C_{15}H_{14}N_3O_2$ 268.1081 found 268.1083 $[M+H]^+$.

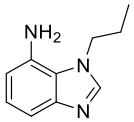
1-(2-Chloroethyl)-4-nitro-1*H*-benzo[*d*]imidazole **122f**, 1-(2-chloroethyl)-7-nitro-1*H*-benzo[*d*]imidazole **123f** and 4-nitro-1-vinyl-1*H*-benzo[*d*]imidazole **122g**



122f **123f** **122g**

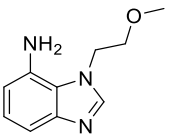
To a mixture of tetrabutylammonium bromide (TBAB) (0.1 mmol, 20 mg, 0.02 eq.), K_2CO_3 (6.5 mmol, 891 mg, 2.1 eq.) and KOH (20.6 mmol, 1.2 g, 6.7 eq.) in dichloroethane (8.2 mL, 0.38 M) was added nitrobenzimidazole **118** (3.1 mmol, 500 mg, 1.0 eq.). The resulting mixture was heated to 50 °C and allowed to stir overnight. Upon completion, the mixture was cooled to room temperature and the solids filtered off. The filtrate was diluted with CH_2Cl_2 (20 mL), washed with water (3×20 mL) and brine (20 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography 4:1 EtOAc/hexanes isolated the *title regioisomer* **122f**, (1.1 mmol, 240 mg, 35%) as a beige solid, **123f,** (0.4 mmol, 100 mg, 14%) as an orange solid and **122g** (0.3 mmol, 65 mg, 11%) as a beige solid. **122f: MP** 150–152°C; **IR** (ATR, cm^{-1}) 3099, 2921, 1519, 1493, 1340; **^{1}H NMR**, (400 MHz, $CDCl_3$) δ 8.23 (s, 1H, NC<u>H</u>NR), 8.21 (d, *J* 6.5 Hz, 1H, Ar<u>H</u>), 7.75 (d, *J* 6.5 Hz, 1H, Ar<u>H</u>), 7.46 (app t, *J* 6.5 Hz, 1H, Ar<u>H</u>), 4.62 (t, *J* 4.9 Hz, 2H, NC<u>H</u>$_2$CH$_2$Cl), 3.90 (t, *J* 4.9 Hz, 2H, NCH$_2$C<u>H</u>$_2$Cl); **^{13}C NMR**, (101 MHz, $CDCl_3$) δ 146.7, 139.8, 137.4, 136.4, 122.7, 119.8, 116.0, 47.2, 42.3; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_8{}^{35}ClN_3O_2$ 225.0 found 226.0 $[M+H]^+$; **123f: MP** 97–99 °C; **IR** (ATR, cm^{-1}) 3097, 2956, 1538, 1496, 1355; **^{1}H NMR**, (400 MHz, $CDCl_3$) δ 8.13 (m, 2H, 2×Ar<u>H</u>), 8.08 (s, 1H, NC<u>H</u>NR), 7.41–7.37 (m, 1H, Ar<u>H</u>), 4.84 (t, *J* 5.4 Hz, 2H, NC<u>H</u>$_2$CH$_2$Cl), 3.85 (t, *J* 5.4 Hz, 2H, NCH$_2$C<u>H</u>$_2$Cl); **^{13}C NMR**, (101 MHz $CDCl_3$) δ 148.3, 145.1, 136.2, 131.5, 128.0, 122.0, 121.7, 50.7, 44.1; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_8{}^{35}ClN_3O_2$ 225.0 found 226.0 $[M+H]^+$. **122g**:**MP** 136–138 ºC; **IR** (ATR, cm^{-1}) 3111, 3079, 1647, 1525, 1476, 1337; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.35 (s, 1H, NC<u>H</u>N), 8.16 (dd, *J* 8.0, 0.8 Hz, 1H, ArH), 7.84 (dd, *J* 8.0, 0.8 Hz, 1H, ArH), 7.45 (app t, *J* 8.0 Hz, 1H, ArH) 7.15 (dd, J 15.8, 8.8 Hz, 1H, C<u>H</u>), 5.64 (dd, J 15.8, 2.0 Hz, 1H, C=C<u>H</u>) 5.29 (dd, J 8.8, 2.0 Hz, 1H, C=C<u>H</u>); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 143.7, 139.7, 137.4, 135.3, 127.4, 123.3, 120.1, 116.7, 106.5; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_7N_3O_2$ 189.0 found 190.0 $[M+H]^+$.

1-Propyl-1*H*-benzo[*d*]imidazol-7-amine **124a**



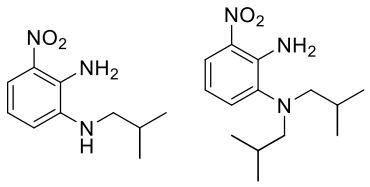
According to **General Procedure B**; 10 mol% Pd/C (0.1 mmol, 7 mg) and nitrobenzimidazole **123a** (0.6 mmol, 128 mg) in MeOH (6.8 mL) were stirred at room temperature for 5 hours. The *title compound* **124a** was isolated without further purification as a brown solid (0.6 mmol, 106 mg, 100%). **MP** 75–77 °C; **IR** (ATR, cm^{-1}) 3317, 3207, 2983, 2928, 1595, 1495; **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.75 (s, 1H, NC<u>H</u>NR), 7.32–7.30 (m 1H, Ar<u>H</u>), 7.07–7.04 (m, 1H, Ar<u>H</u>), 6.60–6.58 (m, 1H, Ar<u>H</u>), 4.31 (t, *J* 7.5 Hz, 2H, NC<u>H_2</u>CH_2CH_3), 3.72 (br s, 2H, N<u>H_2</u>), 1.96 (app sextet, J 7.5 Hz, 2H, NCH_2C<u>H_2</u>CH_3), 1.98 (t, *J* 7.5 Hz, 3H, C<u>H_3</u>); **^{13}C NMR**, ($CDCl_3$, 125 MHz) δ 146.3, 144.0, 132.4, 124.2, 122.9, 112.4, 111.2, 48.8, 25.7, 11.2; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{13}N_3$ 175.1 found 176.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{10}H_{13}N_3$ 176.1182 found 176.1181 $[M+H]^+$.

1-(2-Methoxyethyl)-1*H*-benzo[*d*]imidazol-7-amine **124b**



According to **General Procedure B**; 10 mol% Pd/C (0.1 mmol, 10 mg) and nitrobenzimidazole **123b** (0.9 mmol, 190 mg) in MeOH (10 mL) were stirred at room temperature for 4 hours. Further purification *via* flash column chromatography in 9:1 CH_2Cl_2/MeOH afforded the *title compound* **133** as a purple solid (0.7 mmol, 135 mg, 83%). **MP** 98–100 °C; **IR** (ATR, cm^{-1}) 3352, 3205, 2833, 1593, 1491; **^{1}H NMR**, (500 MHz, $CDCl_3$) δ 7.75 (s, 1H, NC<u>H</u>NR), 7.28–7.26 (m, 1H, Ar<u>H</u>), 7.05–7.02 (m, 1H, Ar<u>H</u>), 6.56–6.55 (m, 1H, Ar<u>H</u>), 4.61 (t, *J* 4.8 Hz, 2H, NC<u>H_2</u>CH_2O), 4.30 (br s, 2H, N<u>H_2</u>), 3.76 (t, *J* 4.8 Hz, 2H, NCH_2C<u>H_2</u>O), 3.33 (s, 3H, OC<u>H_3</u>); **^{13}C NMR**, (151 MHz, $CDCl_3$) δ 145.5, 143.5, 133.7, 125.4, 122.9, 112.0, 111.1, 73.2, 59.4, 46.7; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{13}N_3O$ 191.1 found 192.8 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{10}H_{14}N_3O$ 192.1131 found 192.1130 $[M+H]^+$

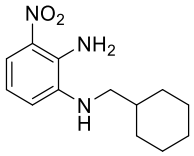
N[1]-isobutyl-3-nitrobenzene-1,2-diamine **125a** and *N*[1],*N*[1]-diisobutyl-3-nitrobenzene-1,2-diamine **125j**



125a **125j**

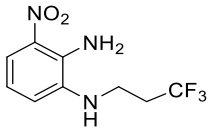
According to **General Procedure F**; 3-nitro-1,2-benzenediamine **100** (0.82 mmol, 125 mg), isobutyraldehyde (1.22 mmol, 0.11 mL), STAB (4.10 mmol, 867 mg) and AcOH (1.22 mmol, 70 µL) in THF (2.0 mL) at 0 °C for 1 hour. Further purification *via* flash column chromatography 9:1 to 3:2 hexanes/EtOAc afforded *title compound* **125a** as a red oil (0.41 mmol, 80 mg, 50%). **IR** (ATR, cm^{-1}) 3474, 3366, 2951, 2892, 2866, 1521, 1496; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.69–7.66 (m, 1H, ArH) 6.86–6.84 (m, 1H, ArH), 6.73–6.69 (m, 1H, ArH), 5.93 (br s, 2H, NH_2), 3.11 (br s, 1H, NH), 2.92 (d, *J* 6.7 Hz, 2H, CH_2), 1.94 (app septet, *J* 6.7 Hz, 1H, CH), 1.05 (d, *J* 6.7 Hz, 6H, 2×CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.3, 136.6, 133.6, 117.8, 117.2, 116.5, 53.0, 28.3, 20.8; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{15}N_3O_2$ 209.1 found 210.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{10}H_{16}N_3O_2$ 210.1237 found 210.1238 $[M+H]^+$;**125j** was isolated as a dark red oil (0.18 mmol, 50 mg, 23%); **IR** (ATR, cm^{-1}) 3490, 3366, 2957, 2812, 1514, 1467; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.93–7.90 (m, 1H, ArH), 7.26–7.24 (m, 1H, ArH), 6.63– 6.59 m, 1H, ArH), 2.60 (d, *J* 6.7 Hz, 4H, 2×CH_2), 1.75 (app septet, *J* 6.7 Hz, 2H, 2×CH), 0.92 (d, *J* 6.7 Hz, 12H, 4×CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 143.0, 141.0, 141.0, 132.5, 128.7, 122.4, 115.0, 63.1, 26.3, 3 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{23}N_3O_2$ 265.2 found 266.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{24}N_3O_2$ 266.1863 found 266.1864 $[M+H]^+$.

N[1]-(cyclohexylmethyl)-3-nitrobenzene-1,2-diamine **125b**



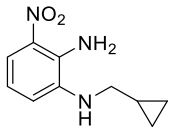
According to **General Procedure F**; 3-nitro-1,2-benzenediamine **100** (0.65 mmol, 100 mg), cyclohexanecarbaldehyde (0.98 mmol, 0.12 mL, 1.5 eq.), STAB (1.95 mmol, 413 mg) and AcOH (0.98 mmol, 0.06 mL, 1.5 eq.) were reacted in anhydrous THF (1.7 mL) at 0 °C for 1 hour. Further purification *via* flash column chromatography 9:1 hexanes/EtOAc afforded the *title compound* **125b** as a red oil (0.43 mmol, 106 mg, 66%). **IR** (ATR, cm^{-1}) 3476, 3369, 2919, 2848, 1623, 1522; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.68–7.65 (m, 1H, ArH), 6.85–6.93 (m, 1H, ArH), 6.72–6.68 (m, 1H, ArH), 5.92 (br s, 2H, NH_2), 3.12 (br s, 1H, NH), 2.94 (app t, J 6.0 Hz, 2H, CH_2), 1.89-1.69 (m, 6H, 3×CH_2), 1.67–1.58 (m, 1H, $NHCH_2CH$), 1.35-1.16 (m, 4H, 2×CH_2), 1.09–1.00 (m, 2H, Ch C(H)H); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.7, 136.5, 133.5, 117.3, 117.2, 116.1, 51.7, 37.8, 31.6, 26.7, 26.1; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{13}H_{19}N_3O_2$ 249.2 found 250.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{13}H_{20}N_3O_2$ 250.1550 found 250.1552 $[M+H]^+$.

3-Nitro-*N*[1]-(3,3,3-trifluoropropyl)benzene-1,2-diamine **125c**



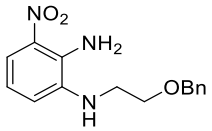
According to **General Procedure F**, 3-nitro-1,2-benzenediamine **100** (0.65 mmol, 100 mg), 3,3,3-trifluoropropanal (0.98 mmol, 80 µL, 1.5 eq.), STAB (1.95 mmol, 413 mg, 3.0 eq.) and AcOH (0.98 mmol, 60 µL, 1.5 eq.) were reacted in anhydrous THF (0.4 M, 1.7 mL) at 0 °C. Further purification *via* flash column chromatography 3:2 hexanes: EtOAc afforded the *title compound* **125c** as a red solid (0.51 mmol, 128 mg, 79%). **MP** 73–74 °C; **IR** (ATR, cm^{-1}) 3463, 3352, 3101, 2952, 1634, 1526; **^{1}H NMR** ($CDCl_3$, 400 MHz) δ 7.77–7.75 (m, 1H, ArH), 6.91–6.89 (m, 1H, ArH), 6.75–6.71 (m, 1H, ArH), 5.99 (br s, 2H, NH_2), 3.45–3.41 (m, 2H, $NHCH_2$), 3.21 (br s, 1H, NH), 2.55–2.44 (m, 2H, CH_2CF_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 137.3, 136.9, 133.7, 126.6 (q, $J_{C\text{-}F}$ 278 Hz) 118.7, 117.9, 117.0, 38.5–38.4 (m), 33.9 (q, $J_{C\text{-}F}$ 28.5 Hz); **^{19}F NMR** ($CDCl_3$, 376 MHz) δ -64.78 (t, *J* 10.7 Hz, 3F, CF_3); **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_{10}F_3N_3$ 229.1 found 230.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc for $C_9H_{11}F_3N_3$ 230.0900 found 230.0901 $[M+H]^+$.

N[1]-(Cyclopropylmethyl)-3-nitrobenzene-1,2-diamine **125d**



To a flame dried flask purged with argon was added 3-nitro-1,2-benzenediamine **100** (0.65 mmol, 100 mg, 1.0 eq.) in anhydrous THF (0.9 mL, 0.4 M). The reaction mixture was cooled to 0 °C and cyclopropanecarbaldehyde (0.17 mmol, 12 µL, 0.5 eq.), STAB (1.02 mmol, 216 mg, 1.0 eq.) and AcOH (0.17 mmol, 10 µL, 0.5 eq.) were added sequentially. After 30 minutes a second portion of aldehyde (12 µL) and STAB (216 mg) were added to the cooled solution. After a further 30 minutes at 0 °C the reaction mixture was slowly quenched with $NaHCO_3$ (2 mL) and allowed to stir at room temperature for 20 mins. The organics were extracted with CH_2Cl_2 (3 × 20 mL), washed with brine (10 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography in 85:15 hexanes/EtOAc afforded the *title compound* **125d** as a red oil (0.36 mmol, 74 mg, 55%). **IR** (ATR, cm^{-1}) 3316, 2972, 2925, 2880, 1454; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.69–7.67 (m, 1H, ArH), 6.82–6.80 (m, 1H, ArH), 6.71–6.67 (m, 1H, ArH), 5.98 (br s, 2H, NH_2), 3.24 (br s, 1H, NH), 2.96–2.93 (m, 2H, $NHCH_2$), 1.18–1.12 (m, 1H, $NHCH_2CH$), 0.64–0.59 (m, 2H, Cp C(H)H), 0.29–0.26 (m, 2H, Cp C(H)H); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.4, 136.1, 133.5, 117.4, 117.2, 116.4, 50.4, 11.0, 3.8; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{13}N_3O_2$ 207.1 found 208.1 $[M+H]^+$.

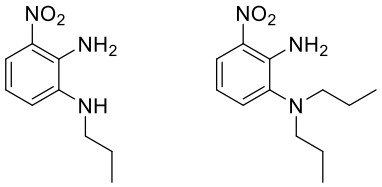
N[1]-(2-(Benzyloxy)ethyl)-3-nitrobenzene-1,2-diamine **125e**



According to **General Procedure F**, 3-nitro-1,2-benzenediamine **100** (0.66 mmol, 101 mg), 2-(benzyloxy)acetaldehyde **128** (0.66 mmol, 99 mg), STAB (1.32 mmol, 280 mg) and AcOH (0.66 mmol, 40 µL) were stirred in THF (1.7 mL) at 0 °C. Purification *via* flash column chromatography 4:1 hexanes/EtOAc afforded the *title compound* **125e** as a red oil (0.21 mmol, 60 mg, 32%). **IR** (ATR, cm^{-1}) 3130, 3020, 2905, 2800, 1519, 1335; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.71–7.69 (m, 1H, ArH), 7.38–7.32 (m, 5H, $OCH_2C_6H_5$), 6.85–6.83 (m, 1H, ArH), 6.69–6.65 (m, 1H, ArH), 6.02 (br s, 2H, NH_2), 4.59 (s, 2H, OCH_2Ph), 3.78–3.75 (m, 2H, NCH_2CH_2O), 3.58 (br s, 1H, NH), 3.29–3.27 (m, 2H, NCH_2CH_2O); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.0, 137.9, 137.2, 133.3, 128.7,

128.1, 128.0, 118.4, 117.0, 116.8, 73.40, 68.5, 44.9;**LRMS** (LCMS-ESI) *m/z* calc.for $C_{15}H_{17}N_3O_3$ 287.1 found 288.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{15}H_{18}N_3O_3$ 288.1343 found 288.1345 $[M+H]^+$.

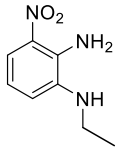
3-Nitro-*N*[1]-propylbenzene-1,2-diamine **125f** and 3-nitro-*N*[1],*N*[1]-dipropylbenzene-1,2-diamine **125i**



125f **125i**

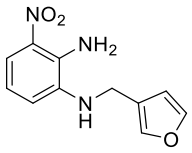
According to **General Procedure F**, 3-nitro-1,2-benzenediamine **100** (1.31 mmol, 200 mg), propionaldehyde (1.44 mmol, 100 µL), STAB (2.62 mmol, 555 mg) and AcOH (1.44 mmol, 80 µL) were stirred in anhydrous THF (3.3 mL) at 0 °C .Purification *via* flash column chromatography in 9:1 hexanes/EtOAc, afforded the *title compound* **125f** as a black solid (0.71 mmol, 139 mg, 54%). **MP** 82–84 °C; **IR** (ATR, cm^{-1}) 3466, 3358, 2953, 2871, 1517, 1368; **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.68 (d, *J* 7.9 Hz, 1H, ArH), 6.86 (d, *J* 7.9 Hz, 1H, ArH), 6.71 (app t, *J* 7.9 Hz, 1H, ArH), 5.93 (br s, 2H, NH_2), 3.09–3.07 (m, 2H, $NHCH_2CH_2CH_3$), 1.75–1.69 (m, 2H, $NHCH_2CH_2CH_3$), 1.07–1.04 (m, 3H, $NHCH_2CH_2CH_3$);**^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.5, 136.5, 133.5, 117.4, 117.2, 116.3, 46.9, 22.9, 11.9; **LRMS** (LCMS-ESI) *m/z* calc. $C_9H_{13}N_3O_2$ for 195.1 found 196.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_9H_{14}N_3O_2$ 196.1081 found 196.1080 $[M+H]^+$; **125i** was isolated as a dark red oil (0.60 mmol, 142 mg, 46%). **IR** (ATR, cm^{-1}) 3466, 3358, 2953, 2871, 1517; **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.92 (d, *J* 8.8 Hz, 1H, ArH), 7.23 (d, *J* 6.7 Hz, 1H, ArH), 6.81 (br s, 2H, NH_2) , 6.63 (app t, *J* 7.6 Hz, 1H, ArH), 2.83 (t, *J* 7.4 Hz, 4H, $N(CH_2CH_2CH_3)_2$) 1.43 (sextet, *J* 7.4 Hz, 4H, $N(CH_2CH_2CH_3)_2$), 0.87 (t, *J* 7.4 Hz, 6H. $N(CH_2CH_2CH_3)_2$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 143.7, 140.5, 132.4, 129.1, 122.2, 114.9, 56.5, 20.6, 11.8; **LRMS** (LCMS-ESI) *m/z* calc. for $12H_{19}N_3O_2$ 237.1 found 238.2 $[M+H]^+$.

N[1]-Ethyl-3-nitrobenzene-1,2-diamine **125g**



According to **General Procedure F**, 3-nitro-1,2-benzenediamine **100** (0.65 mmol, 100 mg), acetaldehyde (0.52 mmol, 29 µL, 0.8 eq.), STAB (1.30 mmol, 276 mg, 2.0 eq.) and AcOH (0.52 mmol, 30 µL, 0.8 eq.) were reacted in THF (2.2 mL) at 0 °C. Further purification *via* flash column chromatography in 85:15 petroleum ether 40–60/EtOAc afforded the *title compound* **125g** as a dark red solid. (0.46 mmol, 83 mg, 70 %). **MP** 105–106 °C; **IR** (ATR, cm^{-1}) 3467, 3385, 3361, 2967, 2871, 1522; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.70–7.67 (m, 1H, ArH), 6.89–6.85 (m, 1H, ArH), 6.73–6.69 (m, 1H, ArH), 5.95 (br s, 2H, NH_2), 3.16 (q, *J* 7.1 Hz, 2H, NCH_2CH_3) 1.34 (t, *J* 7.1 Hz, 3H, NCH_2CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.2, 136.5, 133.5, 117.5, 117.2, 116.5, 39.7, 15.0; **LRMS** (LCMS-ESI) *m/z* calc. for $C_8H_{11}N_3O_2$ 181.1 found 182.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_8H_{12}N_3O_2$ 182.0924 found 182.0923 $[M+H]^+$.

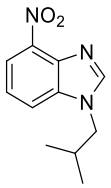
N[1]-(Furan-3-ylmethyl)-3-nitrobenzene-1,2-diamine **125h**



To a flame dried microwave vial purged with argon was added 3-nitro-1,2-benzenediamine **100** (0.98 mmol, 150 mg) in anhydrous THF (3.3 mL, 0.3 M). The resulting solution was cooled to 0 °C before the addition of furan-3-carbaldehyde (1.96 mmol, 170 µL, 2.0 eq.). The resulting mixture was then allowed to warm to room temperature. To the mixture was added $NaBH_4$ (2.94 mmol, 111 mg, 3.0 eq.) portion-wise, followed by AcOH (1.96 mmol, 110 µL, 2.0 eq.). Upon completion, the resulting mixture was slowly quenched with sat. aqueous $NaHCO_3$ solution (10 mL). The aqueous was extracted with CH_2Cl_2 (4×10 mL) and washed with brine (10 mL). The combined organics were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography in 4:1 petroleum ether 40–60/EtOAc afforded the *title compound* **125h** as a dark red solid (0.30 mmol, 69 mg, 31%). **MP** 84–86 °C **IR** (ATR, cm^{-1}) 3471, 3365, 3311, 2922, 2844, 1621, 1521; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.73–7.71 (m, 1H, ArH), 7.44–7.43 (m, 2H, 2×ArH), 6.93–6.91 (m, 1H. ArH), 6.73–6.69 (m, 1H, ArH), 6.45–6.44 (m, 1H, ArH), 5.97 (br s, 2H,

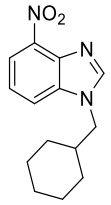
NH_2), 4.17 (s, 2H, CH_2), 3.29 (br s, 1H. NH); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 143.8, 140.4, 137.7, 136.8, 133.5, 122.6, 118.2, 117.1, 110.4, 40.4, 1 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{11}H_{11}N_3O_3$ 233.1 found 234.1 $[M+H]^+$.

1-Isobutyl-4-nitro-1*H*-benzo[*d*]imidazole **126a**



According to **General Procedure G**; alkyl nitrobenzenediamine **125a** (0.41 mmol, 80 mg), triethylorthoformate (0.97 mmol, 0.20 mL) and *p*-TSOH (0.04 mmol, 8 mg) in toluene (1 mL) were stirred at reflux for 1 hour. Further purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **126a** as a yellow oil (0.39 mmol, 85 mg, 95%). **IR** (ATR, cm^{-1}) 3081, 2960, 2873, 1519, 1495; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.16 (d, *J* 6.4 Hz, 1H, ArH), 8.11 (s, 1H, NCHNR), 7.72 (d, *J* 6.4 Hz, 1H, ArH), 7.40 (app t, *J* 6.4 Hz, 1H, ArH), 4.05 (d, *J* 5.9 Hz, 2H, CH_2), 2.27–2.19 (m, 1H, CH), 0.98 (d, J 5.3 Hz, 6H, 2×CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 146.7, 139.7, 137.4, 137.0, 122.2, 119.5, 116.5, 53.2, 29.4, 20.2; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{11}H_{13}N_3O_2$ 219.1 found 220.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{11}H_{14}N_3O_2$ 220.1081 found 220.1079 $[M+H]^+$.

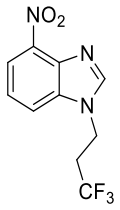
1-(Cyclohexylmethyl)-4-nitro-1*H*-benzo[*d*]imidazole **126b**



According to **General Procedure G**; alkyl nitrobenzenediamine **125b** (0.53 mmol, 133 mg), triethylorthoformate (1.60 mmol, 0.27 mL) and *p*-TSOH (0.05 mmol, 10 mg) in toluene (2 mL) were stirred at reflux for 1 hour. Further purification *via* flash column chromatography in 3:2 EtOAc/hexanes afforded the *title compound* **126b** as a yellow solid (0.19 mmol, 96 mg, 70%). **MP** 111-113 °C; **IR** (ATR, cm^{-1}) 3100, 2920, 2853, 1625, 1519, 1452; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.17 (dd, *J* 8.1, 0.8 Hz, 1H, ArH) 8.09 (s, 1H, NCHNR), 7.72 (dd, *J* 8.1, 0.8 Hz, 1H, ArH), 7.41 (app t, *J* 8.1 Hz. 1H, ArH), 4.10–4.08 (m, 2H, NCH_2), 1.91–1.80 (m, 1H, NCH_2CH), 1.77–1.62 (m, 6H, Ch 3×C(H)H), 1.28–0.97 (m, 4H, Ch 2×C(H)H); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 146.8,

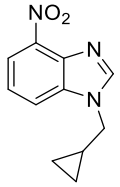
139.6, 137.2, 137.0, 122.2, 119.5, 116.6, 52.1, 38.6, 30.9, 26.1, 25.6; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{17}N_3O_2$ 259.1 found 260.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{17}N_3O_2$ 260.1394 found 260.1396 $[M+H]^+$.

4-Nitro-1-(3,3,3-trifluoropropyl)-1*H*-benzo[*d*]imidazole **126c**



According to **General Procedure G**; alkyl nitrobenzenediamine **125c** (0.40 mmol, 100 mg), triethylorthoformate (1.20 mmol, 0.20 mL) and *p*-TsOH (0.05 mmol, 10 mg) in toluene (0.9 mL) were stirred at reflux for 1 hour. Further purification *via* flash column chromatography 100% EtOAc afforded the *title compound* **126c** as a pale yellow solid (0.35 mmol, 90 mg, 87%). **MP** 95–96 °C; **IR** (ATR, cm^{-1}) 3145, 3106, 2983, 1519, 1452; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.22 (dd, *J* 8.1, 0.9 Hz, 1H, ArH), 8.17 (s, 1H, NCHN), 7.72 (dd, *J*, 8.1, 0.9 Hz, 1H, ArH), 7.48 (app t, *J* 8.1 Hz, 1H, ArH), 4.58–4.55 (m, 2H, NCH_2), 2.79–2.68 (m, 2H, CH_2CF_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 146.1, 140.0, 137.5, 136.0, 126.7, 123.0, 120.0, 38.9–38.8 (m), 34.5 (q, $J_{C\text{-}F}$ 28.5 Hz), 29.8;**^{19}F NMR** (376 MHz, $CDCl_3$) δ -65.43 (t, *J* 10.8 Hz, 3F, CF_3); **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_8F_3N_3O_2$ 259.1 found 260.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{10}H_9F_3N_3O_2$ 260.0641 found 260.0644 $[M+H]^+$.

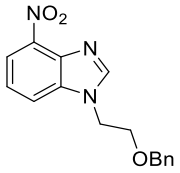
1-(Cyclopropylmethyl)-4-nitro-1*H*-benzo[*d*]imidazole **126d**



According to **General Procedure G**; alkyl nitrobenzenediamine **125d** (0.30 mmol, 63 mg), triethylorthoformate (0.91 mmol, 0.15 mL) and *p*-TsOH (0.05 mmol, 10 mg) in toluene (0.7 mL) were stirred at reflux for 1 hour. Purification *via* flash column chromatography in 4:1 EtOAc/hexanes afforded the *title compound* **126d** as a yellow solid (0.23 mmol, 50 mg, 77%). **MP** 84–85 °C; **IR** (ATR, cm^{-1}) 3094, 2951, 2923, 2853, 1519, 1495, 1365; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.25 (s, 1H, NCHN), 8.17 (dd, *J* 8.1, 0.9 Hz, 1H, ArH), 7.75 (dd, *J* 8.1, 0.9 Hz, 1H, ArH), 7.41 (app t, *J* 8.1 Hz, 1H, ArH), 4.12–4.10 (m, 2H, NCH_2CH), 1.36–1.28 (m, 1H, $CH_2CH(CH_2)_2$), 0.79–0.75 (m, 2H,Cp

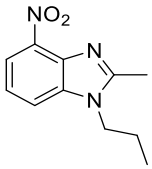
C(H)H), 0.47–0.44 (m, 2H, Cp C(H)H);**^{13}C NMR** (101 MHz, $CDCl_3$) δ 145.9, 139.7 137.5, 136.9, 122.2, 119.5, 116.3, 50.4, 10.9, 4.7; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{11}H_{11}N_3O_2$ 217.1 found 218.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{11}H_{12}N_3O_2$ 218.0924 found 218.0925 $[M+H]^+$.

1-(2-(Benzyloxy)ethyl)-4-nitro-1*H*-benzo[*d*]imidazole **126e**



According to **General Procedure G**; alkyl nitrobenzenediamine **125e** (0.19 mmol, 55 mg), triethylorthoformate (0.57 mmol, 0.10 mL) and *p*-TsOH (0.02 mmol, 4 mg) in toluene (0.6 mL) were stirred at reflux for 1 hour. Further purification *via* flash column chromatography in 9:1 EtOAc/hexanes afforded the *title compound* **126e** as a yellow solid (0.15 mmol, 45 mg, 80%). **MP** 108–110 °C; **IR** (ATR, cm^{-1}) 3130, 3020, 2905, 2800, 1519, 1335; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.25 (s, 1H, NCHN), 8.17 (dd, *J* 8.0, 0.9 Hz, 1H, ArH), 7.73 (dd, *J* 8.0, 0.9 Hz, 1H, ArH), 7.37 (app t, *J* 8.0 Hz, 1H, ArH), 7.29–7.25 (m, 3H, 3×ArH), 7.16–7.13 (m, 2H, 2×ArH), 4.47 (s, 2H, OCH_2Ph), 4.44 (t, *J* 5.0 Hz, 2H, NCH_2CH_2O), 3.82 (t, *J* 5.0 Hz, 2H, NCH_2CH_2O); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 147.1, 139.6, 137.3, 137.1, 136.9, 128.7, 128.2, 127.8, 122.2, 119.5, 116.5, 73.7, 68.1, 45.9; **LRMS** (LCMS-ESI) *m/z* calc.for $C_{16}H_{15}N_3O_3$ 297.1 found 298.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{16}H_{16}N_3O_3$ 298.1186 found 298.1189 $[M+H]^+$.

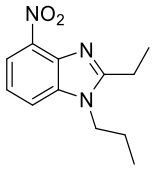
2-Methyl-4-nitro-1-propyl-1*H*-benzo[*d*]imidazole **126f**



To a microwave vial was added alkyl nitrobenzenediamine **125f** (0.32 mmol, 63 mg, 1.0 eq.), triethylorthoacetate (0.93 mmol, 0.17 mL, 3.0 eq.) and *p*-TsOH (0.03 mmol, 6 mg, 0.1 eq.) in toluene (1.0 mL, 0.3 M). The resulting mixture was heated to reflux for 1 hour. Upon completion the resulting mixture was concentrated *in vacuo* and purified *via* flash column chromatography in 100% EtOAc to afford the *title compound* **126f** as a yellow solid (0.25 mmol, 55 mg, 78%). **MP** 86–87 °C; **IR** (ATR, cm^{-1}) 3106, 2925, 2877, 1523, 1506, 1333; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.10 (dd, *J* 8.0, 0.9 Hz, 1H, ArH), 7.60 (dd, *J* 8.0, 0.9 Hz, 1H, ArH), 7.31 (app t, *J* 8.0 Hz, 1H, ArH), 4.15 (t, *J* 7.4 Hz, 2H, $NCH_2CH_2CH_3$) 2.73 (s, 3H, $NCCH_3$), 1.86 (app sextet, *J* 7.4 Hz, 2H,

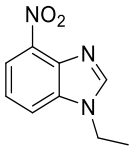
$NCH_2C\underline{H_2}CH_3$), 0.99 (t, *J* 7.4 Hz, 3H, $NCH_2CH_2C\underline{H_3}$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 156.0, 138.7, 138.2, 136.7, 121.2, 119.1, 115.6, 46.0, 23.2, 14.5, 11.4; **LRMS** (LCMS-ESI) *m/z* calc.for $C_{11}H_{13}N_3O_2$ 219.1 found 220.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{11}H_{14}N_3O_2$ 220.1081 found 220.1080 $[M+H]^+$.

2-Ethyl-4-nitro-1-propyl-1*H*-benzo[*d*]imidazole **126g**



To a microwave vial was added alkyl nitrobenzenediamine **125f** (0.31 mmol, 60 mg, 1.0 eq.), triethylorthopropionate (1.84 mmol, 0.37 mL, 6.0 eq.) and *p*-TsOH (0.03 mmol, 6 mg, 0.1 eq.) in toluene (1.0 mL, 0.3 M). The resulting mixture was heated to reflux overnight. Upon completion, the mixture was purified *via* flash column chromatography in 1:1 EtOAc/hexanes to afford the *title compound* **126g** as a yellow oil (0.24 mmol, 56 mg, 77%). **IR** (ATR, cm^{-1}) 2965, 2934, 1874, 1517, 1472; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.09 (dd, *J* 8.0, 0.9 Hz, 1H, Ar$\underline{H}$), 7.60 (dd, *J* 8.0, 0.9 Hz, 1H, Ar$\underline{H}$), 7.31 (app t, *J* 8.0 Hz, 1H, ArH), 4.15 (t, *J* 7.4 Hz, 2H, $NC\underline{H_2}CH_2CH_3$), 3.03 (q, *J* 7.4 Hz, 2H, $NCC\underline{H_2}CH_3$), 1.91–1.82 (m, 2H, $NCH_2C\underline{H_2}CH_3$) 1.50 (t, *J* 7.4 Hz, 3H, $NCCH_2C\underline{H_3}$) 1.00 (t, *J* 7.4 Hz, 3H, $NCH_2CH_2C\underline{H_3}$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 160.5, 138.8, 138.1, 136.3, 121.3, 119.2, 115.7, 45.9, 23.4, 21.3, 12.6, 11.5; **LRMS** (LCMS-ESI) *m/z* calc.for $C_{12}H_{15}N_3O_2$ 233.1 found 234.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{12}H_{16}N_3O_2$ 234.1237 found 234.1237 $[M+H]^+$.

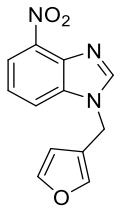
1-Ethyl-4-nitro-1*H*-benzo[*d*]imidazole **126h**



According to **General Procedure G**; alkyl nitrobenzenediamine **125g** (0.45 mmol, 81 mg), triethylorthoformate (1.34 mmol, .0.25 mL) and *p*-TsOH (0.05 mmol, 10 mg) in toluene (1.5 mL) were stirred at reflux for 1 hour. Further purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **126h** as a yellow oil (0.25 mmol, 47 mg, 55%); **IR** (ATR, cm^{-1}) 3079, 2973, 1521, 1500, 1366; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.18–8.16 (m, 2H, Ar$\underline{H}$ and NC$\underline{H}$N), 7.75–7.73 (m, 1H, Ar$\underline{H}$), 7.43–7.39 (m, 1H, Ar$\underline{H}$), 4.33 (q, *J* 7.3 Hz, 2H, C$\underline{H_2}$), 1.59 (t, *J* 7.3 Hz, 3H, C$\underline{H_3}$);**^{13}C**

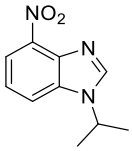
NMR (101 MHz, $CDCl_3$) δ 145.7, 139.7, 137.6, 136.6, 122.2, 119.5, 116.2, 40.6, 15.4; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_9N_3O_3$ 191.0 found 192.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_9H_{10}N_3O_3$ 192.0766 found 192.0768 $[M+H]^+$.

1-(Furan-3-ylmethyl)-4-nitro-1*H*-benzo[*d*]imidazole **126i**



According to **General Procedure G**; alkyl nitrobenzenediamine **125h** (0.32 mmol, 75 mg), triethylorthoformate (0.96 mmol, 0.16 mL) and *p*-TsOH (0.03 mmol, 6 mg) in toluene (1.1 mL) were stirred at reflux for 2 hours. Purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **126i** as a yellow solid (0.26 mmol, 62 mg, 80%). **MP** 113-115 °C; **IR** (ATR, cm^{-1}) 3106, 3056, 1517, 1493; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.19–8.17 (m, 2H, 2×ArH), 7.71 (d, *J* 8.1 Hz, 1H, ArH), 7.47–7.44 (m, 2H, 2×ArH), 7.40 (app t, *J* 8.1 Hz, 1H, ArH), 6.28–6.27 (m, 1H, ArH), 5.30 (s, 2H, CH_2); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 146.1, 144.9, 140.8, 139.8, 137.7, 136.6, 122.5, 119.7, 119.4, 116.5, 109.6, 41.0; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{12}H_9N_3O_3$ 243.2 found 244.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{12}H_{10}N_3O_3$ 244.0717 found 244.0716 $[M+H]^+$.

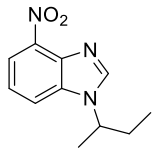
1-*iso*Propyl-4-nitro-1*H*-benzo[*d*]imidazole **126j**



To a microwave vial was added alkyl nitrobenzenediamine **129a** (0.28 mmol, 55 mg, 1.0 eq.) and triethylorthoformate (0.90 mmol, 0.15 mL, 3.2 eq.). To the resulting mixture was added a drop of conc. H_2SO_4, before heating to 130 °C for 2 h. Upon completion, the reaction was washed with H_2O (5 mL), extracted with CH_2Cl_2 (3×10 mL), dried over Mg_2SO_4, filtered and concentrated *in vacuo* to provide the *title compound* **126j** as a yellow solid (0.2 mmol, 42 mg, 73 %). **126j** was carried onto the next step without any further purification. **MP** 65–67 ºC; **IR** (ATR, cm^{-1}) 3086, 2960, 2848, 1519, 1344; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.22 (s, 1H, NCHN), 8.18 (dd, *J* 8.1, 1.0 Hz, 1H, ArH), 7.76 (dd, *J* 8.1, 1.0 Hz, 1H, ArH) 7.41 (app t, *J* 8.1 Hz, 1H, ArH), 4.72 (septet, *J* 6.8 Hz, 1H, CH), 1.68 (d, *J* 6.8 Hz, 6H, 2×CH_3); **^{13}C NMR** (101 MHz,

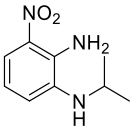
$CDCl_3$) δ 143.7, 139.6, 137.5, 136.2, 122.1, 119.5, 116.7, 48.7, 22.8; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{11}N_3O_2$ 205.1 found 206.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{10}H_{12}N_3O_2$ 206.0924 found 206.0924 $[M+H]^+$.

1-(*sec*-Butyl)-4-nitro-1*H*-benzo[*d*]imidazole **126k**



According to **General Procedure G**; alkyl nitrobenzenediamine **129b** (0.38 mmol, 79 mg), triethylorthoformate (1.13 mmol, 0.19 mL) and *p*-TsOH (0.04 mmol, 8 mg) were stirred at reflux overnight. Purification *via* flash column chromatography in 9:1 EtOAc/petroleum ether afforded the *title compound* **126k** as an orange oil (0.24 mmol, 52 mg, 62%). **IR** (ATR, cm^{-1}) 3088, 2974, 2932, 1527; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.22 (s, 1H, NC<u>H</u>N), 8.17 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 7.76 (dd, *J* 8.0, 0.8 Hz, 1H, Ar<u>H</u>), 7.40 (app t, *J* 8.0 Hz, 1H, Ar<u>H</u>), 4.46 (app sextet, *J* 7.0 Hz, 1H, C<u>H</u>), 2.08– 1.95 (m, 2H, C<u>H</u>(<u>H</u>)CH_3), 1.66 (d, *J* 7.0 Hz, 3H, CHC<u>H$_3$</u>), 0.91 (t, *J* 7.0 Hz, 3H, CH_2C<u>H$_3$</u>); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 144.3, 139.7, 137.5, 136.4, 122.1, 119.6, 116.7, 54.7, 29.8, 20.8, 10.8; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{11}H_{13}N_3O_2$ 219.0 found 220.1 $[M+H]^+$.

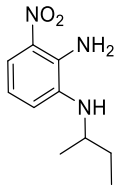
N^1-*iso*propyl-3-nitrobenzene-1,2-diamine **129a**



To a round bottomed flask was added 3-nitro-1,2,-benzenediamine **100** (0.65 mmol, 100 mg, 1.0 eq.), sodium acetate (NaOAc) (1.90 mmol, 156 mg, 2.9 eq.), AcOH (0.6 mL) and acetone (3.19 mmol, 0.23 mL, 4.9 eq.) in a 1:1 mixture of H_2O/MeOH (1.5 mL). The resulting mixture was cooled to 0°C before the portion-wise addition of $NaBH_4$ (3.19 mmol, 121 mg, 4.9 eq.). The resulting mixture was allowed to warm to room temperature and stirred for 30 mins. Upon completion, the mixture was slowly quenched with the dropwise addition of sat. aqueous $NaHCO_3$ solution (10 mL). The aqueous was extracted with EtOAc (3×10 mL), and the combined organics were washed with brine (10 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography 9:1 hexanes: EtOAc afforded the *title compound* **129a** as a dark red solid (0.42 mmol, 81 mg, 64%); **MP** 59–60 °C;

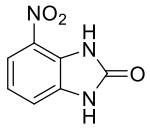
IR (ATR, cm^{-1}) 3463, 3357, 2965, 2928, 1619, 1515; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.70–7.68 (m, 1H, ArH), 6.90–6.88 (m, 1H, ArH), 6.71–6.67 (m, 1H, ArH), 6.01 (br s, 2H, NH_2), 3.56 (app septet, *J* 6.2 Hz, 1H, CH) 2.91 (br s, 1H, NH), 1.25 (d, *J* 6.2 Hz, 6H, 2×CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 137.5, 133.6, 137.1, 119.6, 116.9, 116.9, 45.2, 23.0; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_{13}N_3O_2$ 195.1 found 196.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_9H_{14}N_3O_2$ 196.1081 found 196.1081 $[M+H]^+$.

N^1-(*sec*-Butyl)-3-nitrobenzene-1,2-diamine **129b**



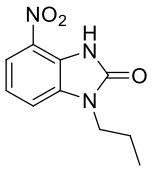
To a round bottomed flask was added 3-nitro-1,2,-benzenediamine **100** (0.65 mmol, 100 mg, 1.0 eq.), sodium acetate (1.90 mmol, 312 mg, 5.9 eq.) and acetic acid (1.2 mL) in a 1:1 mixture of H_2O/MeOH (4.3 mL). The resulting mixture was cooled to 0°C before the addition of 2-butanone (6.4 mmol, 0.6 mL, 9.8 eq.). The mixture was then stirred at 0 °C for 5–10 mins before the portion-wise addition of $NaBH_4$ (6.36 mmol, 241 mg, 9.8 eq.). The mixture was allowed to warm to RT and stirred for 1 hour. Upon completion, the mixture was slowly quenched with the dropwise addition of sat. aqueous $NaHCO_3$ solution (10 mL). The resulting mixture was extracted with EtOAc (3×10 mL) and the organics were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification via flash column chromatography 4:1 hexanes:/EtOAc afforded the *title compound* **129b** as a dark red oil (0.19 mmol, 40 mg, 29 %). **IR** (ATR, cm^{-1}) 3469, 3359, 2965, 2932, 2874, 1521; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.70–7.67 (m, 1H, ArH), 6.89–6.88 (m, 1H, ArH), 6.69–6.66 (m, 1H, ArH), 6.01 (br s, 2H, NH_2), 3.34 (app sextet, J 6.2 Hz, 1H, CH), 1.67–1.62 (m, 1H, CH(H)CH_3), 1.57–1.47 (m, 1H, CH(H)CH_3), 1.21 (d, *J* 6.2 Hz, 3H, CHCH_3), 1.01–0.97 (m, 3H, CH_2CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 137.4, 133.7, 119.6, 116.9, 51.0, 29.6, 20.0, 10.5, 2 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{15}N_3O_2$ 209.1 found 210.1 $[M+H]^+$.

4-Nitro-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one **130a**[175]



To a microwave vial was added 3-nitro-1,2-benzenediamine **100** (1.63 mmol, 250 mg, 1.0 eq.) and DMF (5.4 mL, 0.3 M). The resulting mixture was cooled to 0 °C and 1,1′-carbonyldiimidazole (CDI) (3.26 mmol, 529 mg, 2.0 eq) was added. The resulting mixture was allowed to warm to room temperature overnight. Upon completion, the reaction mixture was diluted with EtOAc (10 mL) and washed with sat. aqueous $NaHCO_3$ (20 mL). A yellow precipitate formed and was collected *via* filtration. This provided the *title compound* **130a** without further purification as a yellow solid (1.46 mmol, 262 mg, 90%). **MP** decomp. >300 °C; **IR** (ATR, cm^{-1}) 3145, 1739, 1683, 1527, 1316; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.45 (br s, 2H, 2×NH), 7.74 (d, *J* 8.0 Hz, 1H, ArH), 7.31 (d, *J* 8.0 Hz, 1H, ArH), 7.12 (app t, *J* 8.0 Hz, 1H, ArH); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 155.1, 132.4, 130.1, 125.9, 120.4, 115.3, 114.3; **LRMS** (LCMS-ESI) m/z calc. for $C_7H_5N_3O_3$ 179.1 found 178.1 $[M-H]^-$.

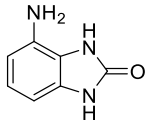
4-Nitro-1-propyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one **130b**



To a microwave vial was added amine **125f** (0.31 mmol, 60 mg, 1.0 eq.) in DMF (1.0 mL, 0.3 M). The resulting mixture was cooled to 0 °C before the addition of CDI (0.62 mmol, 100 mg, 2.0 eq.). The resulting mixture was allowed warm to room temperature overnight. A second portion of CDI was added (0.62 mmol, 100 mg) and the resulting mixture was heated to 80 °C and allowed to stir at this temperature overnight. Upon completion, the reaction mixture was diluted with EtOAc (20 mL) and washed with $NaHCO_3$ (10 mL) and brine (10 mL). The organics were dried over $MgSO_4$, filtered and concentrated *in vacuo* to provide the *title compound* **130b** as an off white solid (0.28 mmol, 61 mg, 89%) and this was used in the next step without further purification. **MP** 171–173 °C; **IR** (ATR, cm^{-1}) 3133, 2957, 2871, 1694, 1525, 1324; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.78 (s, 1H, C(O)NH), 7.78 (dd, *J* 8.0, 0.8 Hz, 1H, ArH), 7.58 (dd, *J* 8.0, 0.8 Hz, 1H, ArH) 7.19 (app t, *J* 8.0 Hz, 1H, ArH), 3.83 (t, *J* 7.0 Hz, 2H, $NCH_2CH_2CH_3$), 1.67 (app sextet, *J* 7.0 Hz, 2H, $NCH_2CH_2CH_3$) 0.87 (t, *J*

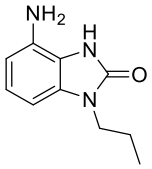
7.0 Hz, 3H, $NCH_2CH_2C\underline{H_3}$); **^{13}C NMR** (101 MHz, DMSO-*d_6*) δ 154.2, 133.0, 130.3, 124.6, 120.4, 115.5, 113.6, 41.8, 21.0, 10.9; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{11}N_3O_3$ 221.1 found 221.9 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{10}H_{12}N_3O_3$ 222.0873 found 222.0874 $[M+H]^+$.

4-Amino-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one **131a**[176]



According to **General Procedure C**; $FeSO_4.7H_2O$ (5.3 mmol, 1.5 g), citric acid (0.5 mmol, 92 mg), $NaBH_4$ (8.8 mmol, 333 mg) and nitrobenzimidazole **130a** (0.9 mmol, 157 mg) in H_2O (88 mL) were stirred overnight at room temperature. The *title compound* **131a** was isolated without any further purification as a beige solid (0.8 mmol, 100 mg, 76%). **MP** decomp. >240 °C; **IR** (ATR, cm^{-1}) 3427, 3344, 3130, 3020, 2918, 2847, 1633, 1519; **^{1}H NMR** (400 MHz, DMSO-*d_6*) δ 10.31 (br s, 1H, N$\underline{H}$), 9.97 (br s, 1H, N$\underline{H}$), 6.67–6.63 (m, 1H, ArH), 6.26–6.21 (m, 2H, 2×Ar$\underline{H}$), 4.83 (br s, 2H, $N\underline{H_2}$); **^{13}C NMR** (101 MHz, DMSO-*d_6*) δ 154.8, 131.4, 130.0, 121.2, 116.0, 106.7, 97.9; **LRMS** (LCMS-ESI) m/z calc. $C_7H_7N_3O$ 149.1 found 149.9 $[M+H]^+$.

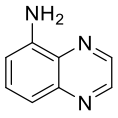
4-Amino-1-propyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one **131b**



According to **General Procedure C;** $FeSO_4.7H_2O$ (1.3 mmol, 358 mg), citric acid (0.1 mmol, 22 mg), $NaBH_4$ (2.1 mmol, 79 mg) and nitrobenzimidazole **130b** (0.2 mmol, 51 mg) in H_2O (21 mL) were stirred overnight at room temperature. To the resulting reaction mixture was added a second portion of reagents $FeSO_4.7H_2O$ (1.3 mmol, 358 mg), citric acid (0.1 mmol, 22 mg) and $NaBH_4$ (2.1 mmol, 79 mg) in H_2O (21 mL) in the same manner as described in **General Procedure C**. The *title compound* **131b** was isolated without any further purification as a brown solid (0.12 mmol, 22 mg, 55%). **MP** 206–208 °C; **IR** (ATR, cm^{-1}) 3432, 3339, 3133, 2957, 2871, 1694, 1525; **^{1}H NMR** (400 MHz, DMSO-*d_6*) δ 10.18 (br s, 1H, C(O)N$\underline{H}$), 6.72 (app t. *J* 8.0 Hz, 1H, Ar$\underline{H}$), 6.38 (d, *J* 8.0 Hz, 1H, Ar$\underline{H}$), 6.30 (d, *J* 8.0 Hz, 1H, Ar$\underline{H}$), 4.89 (br s, 2H, $N\underline{H_2}$) 3.66 (t, *J* 7.0 Hz, 2H, $NC\underline{H_2}CH_2CH_3$), 1.62 (app sextet, *J* 7.0 Hz, 2H, $NCH_2C\underline{H_2}CH_3$) 0.85 (t, *J* 7.0 Hz, 3H, $NCH_2CH_2C\underline{H_3}$); **^{13}C NMR** (151 MHz, DMSO) δ 153.8, 131.6,

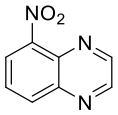
130.7, 121.3, 114.4, 107.2, 97.0, 41.5, 21.3, 11.0; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{13}N_3O$ 191.1 found 192.1 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{10}H_{14}N_3O$ 192.1131 found 192.1131 $[M+H]^+$.

Quinoxalin-5-amine **132c**[160]



According to **General Procedure C;** $FeSO_4.7H_2O$ (2.4 mmol, 667 mg), citric acid (0.2 mmol, 42 mg), $NaBH_4$ (4.0 mmol, 151 mg) and nitroquinoxaline **134** (0.4 mmol, 70 mg) in H_2O (18 mL) were stirred overnight at room temperature. The *title compound* **132c** was isolated without any further purification as a yellow solid (0.33 mmol, 57 mg, 77%); **MP lit**[160] 87–90 °C; **MP** 84–85 °C; **IR** (ATR, cm^{-1}) 3445, 3339, 3183, 2960, 1615, 1472; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.82 (d, *J* 1.8 Hz, 1H, ArH), 8.67 (d, *J* 1.8 Hz, 1H, ArH) 7.56 (app t, *J* 8.0 Hz, 1H, ArH), 7.44 (dd, *J* 8.0, 1.2 Hz, 1H, ArH), 6.96 (dd, *J* 8.0, 1.2 Hz, 1H, ArH) 4.99 (br s, 2H, NH_2); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 145.4, 144.4, 144.0, 141.8, 133.4, 131.2, 117.6, 110.4 **LRMS** (LCMS-ESI) *m/z* calc. for $C_8H_7N_3$ 145.1 found 146.1 $[M+H]^+$.

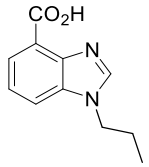
5-Nitroquinoxaline **134**[160]



To a microwave vial was added 3-nitro-1,2,-benzenediamine **100** (0.65 mmol, 100 mg, 1.0 eq.) and glyoxal trimer dihydrate (1.31 mmol, 274 mg, 2.0 eq.) in EtOH (1.6 mL, 0.4 M). The resulting mixture was heated to reflux overnight. Upon completion, the mixture was concentrated *in vacuo*, diluted with EtOAc (10 mL) and washed with H_2O (10 mL) and brine (10 mL). The organics were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography in 2:1 hexanes/EtOAc afforded the *title compound* **134** as a coral solid (0.60 mmol, 105 mg, 92%). **MP lit** [160] 90–92 °C; **MP** 90–91 °C; **IR** (ATR, cm^{-1}) 3058, 3013, 2854, 1521, 1489; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 9.15–9.11 (m, 2H, 2×ArH), 8.45–8.40 (m, 2H, 2×ArH), 8.05–8.01 (m, 1H, ArH); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 147.7, 147.6, 147.2, 141.9, 133.7, 133.4, 129.3, 124.3; **LRMS** (LCMS-ESI) *m/z* calc. for $C_8H_5N_3O_2$ 175.0 found 176.1 $[M+H]^+$.

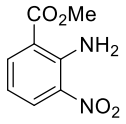
9.5.4 Reverse amide intermediates

1-Propyl-1H-benzo[*d*]imidazole-4-carboxylic acid **136**



To a round bottomed flask was added ester **140** (0.80 mmol, 175 mg) in a mixture of 1:5:1 MeOH/THF/1M LiOH (26.7 mL) and stirred at room temperature overnight. The mixture was then concentrated *in vacuo* and diluted with H_2O (10 mL) and cooled to 0°C. The mixture was then acidified to pH 2 with 1M aqueous HCl. The organics were extracted with CH_2Cl_2 (5×20 mL), dried over $MgSO_4$ and concentrated *in vacuo* to provide acid **136** as a green solid (0.29 mmol, 59 mg, 36%). The *title compound* **16** was used in the next step without further purification. **MP** 68–70 °C; **IR** (ATR, cm^{-1}) 3458, 3385 (br), 2958, 1710, 1505; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.13 (d, *J* 8.0 Hz, 1H, ArH), 8.02 (s, 1H, NC<u>H</u>N), 7.65 (d, *J* 8.0 Hz, 1H, ArH), 7.46 (app t, *J* 8.0 Hz, 1H, Ar<u>H</u>), 4.24 (t, *J* 7.3 Hz, 2H, NC<u>H$_2$</u>CH_2CH_3), 1.98 (app sextet, *J* 7.3 Hz, 2H, NCH_2C<u>H$_2$</u>CH_3), 1.00 (t, *J* 7.3 Hz, NCH_2CH_2C<u>H$_3$</u>); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.5, 143.1, 142.6, 133.7, 125.6, 123.7, 120.0, 115.0, 47.6, 23.3, 11.4; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{11}H_{12}N_2O_2$ 204.1 found 205.1 $[M+H]^+$.

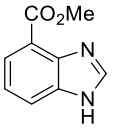
Methyl 2-amino-3-nitrobenzoate **138**[177]



To a round bottomed flask was added 2-amino-3-nitrobenzoic acid **137** (5.49 mmol, 1.0 g) in MeOH (12.5 mL, 0.4 M). To the resulting mixture was added dropwise H_2SO_4 (1 mL) and heated to reflux overnight. The mixture was then cooled to room temperature and concentrated *in vacuo*. The resulting residue was taken up in CH_2Cl_2 (50 mL) and washed with saturated aqueous $NaHCO_3$ (30 mL). The layers were separated and the aqueous was extracted with CH_2Cl_2 (4×20 mL). The combined organics were then washed with brine (20 mL), filtered through a phase separator and concentrated *in vacuo*. Further purification *via* flash column chromatography 9:1 petroleum ether 40–60/EtOAc provided the *title compound* **138** as a yellow solid (3.19 mmol, 626 mg, 58%). **MP lit**[177] 95–96 °C; **MP** 94–95 °C; **IR** (ATR, cm^{-1}) 3455, 3309,

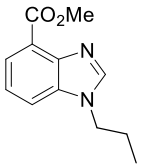
2964, 1698, 1552; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.38 (dd, *J* 8.0, 1.5 Hz, 1H, ArH), 8.24 (dd, *J* 8.0, 1.5 Hz, 1H, ArH), 6.65 (app t, *J* 8.0 Hz, 1H, ArH), 3.92 (s, 3H, CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 167.6, 147.5, 139.5, 133.4, 132.4, 114.6, 114.1, 52.4; **LRMS** (LCMS-ESI) *m/z* calc. for $C_8H_8N_2O_4$ 196.1 found 197.1 $[M+H]^+$.

Methyl 1*H*-benzo[*d*]imidazole-4-carboxylate **139**[161]

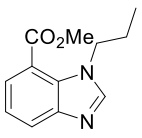


To a round bottomed flask was added methyl ester **138** (2.59 mmol, 508 mg, 1.0 eq.), iron powder (25.9 mmol, 1.45 g, 10 eq.), ammonium chloride (NH_4Cl) (25.9 mmol, 1.39 g, 10 eq.) and *i*-PrOH (13 mL, 0.2 M). To the resulting mixture was added formic acid (13 mL, 0.2 M) and the reaction was heated to 100°C for 24 hours. Upon completion, the resulting mixture was allowed to cool to room temperature and diluted with *i*-PrOH (10 mL). The resulting suspension was filtered, washed with *i*-PrOH (3×10 mL) and the filtrate was concentrated *in vacuo*. The resulting residue was taken up in CH_2Cl_2 (30 mL), washed with $NaHCO_3$ (20 mL) and brine (20 mL), dried over $MgSO_4$ and concentrated *in vacuo* to provide the *title compound* **139** (2.46 mmol, 434 mg, 95%) as a brown solid which was used in next step without further purification. **MP lit**[178] decomp. >250 °C; **MP** 214–216 °C; **IR** (ATR, cm-1) 3131, 3007, 2800, 1701, 1599; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.16 (s, 1H, NCHN), 8.05 (dd, *J* 7.8, 0.9 Hz, 1H, ArH), 7.96 (dd, *J* 7.8, 0.9 Hz, 1H, ArH) 7.34 (app t, *J* 7.8 Hz, 1H, ArH), 4.01 (s, 3H, CH_3); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 167.1, 143.8, 141.5, 133.5, 125.7, 125.4, 122.0, 113.9, 52.4; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_8N_2O_2$ 176.1 found 177.1 $[M+H]^+$.

Methyl 1-propyl-1*H*-benzo[*d*]imidazole-4-carboxylate **140** and methyl 1-propyl-1*H*-benzo[*d*]imidazole-7-carboxylate **141**



140

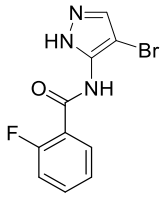


141

According to **General Procedure A**; benzimidazole **139** (1.70 mmol, 300 mg), *n*-propyl bromide (1.87 mmol, 0.16 mL) and K_2CO_3 (3.4 mmol, 470 mg) in DMF (5.7 mL)

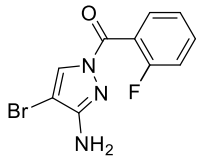
were stirred at room temperature overnight. Further purification *via* flash column chromatography in 1:1 petroleum ether/EtOAc then 9:1 EtOAc/MeOH isolated *title isomer* **140** (1.00 mmol, 218 mg, 59%) and **141** (0.68 mmol, 149 mg, 40%) respectively as brown oils. **140: IR** (ATR, cm^{-1}) 3401, 2964, 2882, 1707, 1498; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.05 (s, 1H, NC<u>H</u>N), 8.01 (dd, *J* 7.8, 1.1 Hz, 1H, Ar<u>H</u>), 7.61 (dd, J 7.8 Hz, 1.1 Hz, 1H, ArH), 7.36 (app t, *J* 7.8 Hz, 1H, ArH), 4.18 (t, *J* 7.0 Hz, 2H, NC<u>H$_2$</u>CH$_2$CH$_3$), 4.06 (s, 3H, C<u>H$_3$</u>), 1.92 (app sextet, J 7.0 Hz, 2H, NCH$_2$C<u>H$_2$</u>CH$_3$), 0.96 (t, J 7.0 Hz, 3H, NCH$_2$CH$_2$C<u>H$_3$</u>); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 167.2, 144.7, 142.7, 135.1, 125.7, 122.3, 122.0, 114.5, 52.6, 47.1, 23.3, 11.4; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{12}H_{14}N_2O_2$ 218.1 found 219.1 $[M+H]^+$.**141: IR** (ATR, cm^{-1}) 3417, 2967, 1714, 1501; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.00 (dd, *J* 8.0,1.0 Hz, 1H, Ar<u>H</u>), 7.92 (s, 1H, NC<u>H</u>N), 7.84 (dd, *J* 8.0, 1.0 Hz, 1H, ArH), 7.29 (app t, *J* 8.0 Hz, 1H, Ar<u>H</u>), 4.50 (t, *J* 7.1 Hz, 2H, NC<u>H$_2$</u>CH$_2$CH$_3$), 3.97 (s, 3H, C<u>H$_3$</u>), 1.73 (app sextet, *J* 7.1 Hz, 2H, NCH$_2$C<u>H$_2$</u>CH$_3$), 0.86 (t, *J* 7.1 Hz, 3H, NCH$_2$CH$_2$C<u>H$_3$</u>); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.9, 146.3, 131.7, 126.4, 125.4, 121.4, 116.8, 52.5, 49.8, 24.1, 11.1, 1 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{12}H_{14}N_2O_2$ 218.1 found 219.1 $[M+H]^+$.

<u>*N*-(4-bromo-1H-pyrazol-5-yl)-2-fluorobenzamide</u> **143**



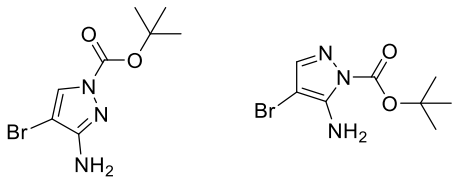
According to **General Procedure E**; 2-fluorobenzoic acid **142** (0.46 mmol, 64 mg), 2-chloro-1-methylpyridin-1-ium iodide (0.46 mmol, 118 mg), DMAP (0.19 mmol, 23 mg), NEt_3 (1.14 mmol, 160 µL) and amine **145** (0.38 mmol, 100 mg) in anhydrous CH_2Cl_2 (1.3 mL) were stirred at 35 °C overnight. Further purification *via* flash column chromatography in 4:1 PhMe/EtOAc and subsequent trituration in hexanes provided the *title compound* **143** as a white solid (0.07 mmol, 20 mg, 19%) **MP** 142–144 °C; **IR** (ATR, cm^{-1}) 3265, 3157, 2923, 1660, 1555; ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.05–9.00 (m, 1H, C(O)N<u>H</u>), 8.25–8.20 (m, 1H, Ar<u>H</u>), 7.64–7.58 (m, 1H, Ar<u>H</u>), 7.55 (s, 1H, NC<u>H</u>CBr), 7.39–7.35 (m, 1H, Ar<u>H</u>), 7.28–7.22 (m, 1H, ArH); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 162.1, 161.2 (d. $J_{C\text{-}F}$ 3.0 Hz), 159.7, 138.0, 137.6, 135.3 (d, $J_{C\text{-}F}$ 9.6 Hz), 132.5, 125.6 (d, $J_{C\text{-}F}$ 2.8 Hz), 119.0 (d, $J_{C\text{-}F}$ 10 Hz), 116.6 (d, $J_{C\text{-}F}$ 24.9 Hz); **^{19}F NMR** (376 MHz, $CDCl_3$) δ -112.0–112.1 (m, 1F); **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_7BrFN_3O$ 283.0 found 282.1 $[M-H]^+$.

(3-Amino-4-bromo-1*H*-pyrazol-1-yl)(2-fluorophenyl)methanone **144**



To a flame dried microwave vial was added 2-fluorobenzoic acid **142** (0.37 mmol, 52 mg, 1.2 eq.) and HATU (0.37 mmol, 141 mg, 1.2 eq.). To the mixture was added anhydrous DMF (1.0 mL, 0.3 M), 4-bromo-1H-pyrazol-5-amine (0.31 mmol, 50 mg, 1.0 eq.) and Hünigs base (0.62 mmol, 0.11 mL, 2.0 eq.). The resulting mixture was heated to 40 °C and stirred at this temperature overnight. Upon completion, H_2O (2 mL) was added and a precipitate formed. The solids were filtered to afford the *title compound* **144** as a white solid (0.12 mmol, 35 mg, 39%). **MP** decomp. >150 °C; **IR** (ATR, cm^{-1}) 3449, 3289, 3140, 1706, 1374.**^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.50 (s, 1H, NCHCBr), 7.63–7.58 (m, 2H, 2×ArH), 7.36, 7.34–7.31 (m, 2H, 2×ArH) 5.90 (br s, 2H, NH_2); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 161.5, 159.9, 157.4, 156.6, 133.0 (d,$J_{C\text{-}F}$ 8.5 Hz), 130.1 (d, $J_{C\text{-}F,}$ 11.7 Hz), 124.3 (d, $J_{C\text{-}F,}$ 2.9 Hz), 121.9 (d, $J_{C\text{-}F,}$ 15.6 Hz), 115.8 (d, $J_{C\text{-}F,}$ 14.0 Hz), 93.1; **^{19}F NMR** (376 MHz, DMSO) δ -113.46 (s, 1F); **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_7{}^{79}BrFN_3O$ 283.0 found 284.0 $[M+H]^+$.

tert-Butyl 3-amino-4-bromo-1*H*-pyrazole-1-carboxylate **145** and tert-*butyl* 5-amino-4-bromo-1*H*-pyrazole-1-carboxylate **146**



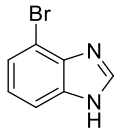
145 **146**

To a flame dried microwave vial purged with argon was added 4-bromo-1*H*-pyrazol-3-amine (3.09 mmol, 500 mg, 1.0 eq.), DMAP (0.62 mmol, 76 mg, 0.2 eq.) and anhydrous THF (7.7 mL, 0.3 M). The resulting solution was cooled to 0°C before the addition of NEt_3 (3.71 mmol, 0.52 mL, 1.2 eq.) and di-*tert*-butyl dicarbonate (3.09 mmol, 659 mg, 1.0 eq.). The mixture was allowed to stir at 0°C for 40 mins before quenching with 1M HCl (5 mL). The resulting suspension was then allowed to warm to room temperature. The layers were separated and the aqueous was extracted with CH_2Cl_2 (3×10 mL). The combined organics were washed with saturated aqueous $NaHCO_3$ solution (10 mL), brine (10 mL), dried over $MgSO_4$, filtered and concentrated

in vacuo. Further purification *via* flash column chromatography in 75:25 40–60 petroleum ether/EtOAc isolated the *title isomer* **145** (1.49 mmol, 391 mg, 48%) and **146** (0.56 mmol, 147 mg, 18%) respectively as white solids. **145**: **MP** 127–129 °C; **IR** (ATR, cm^{-1}) 3446, 3262, 3148, 2983, 1752, 1612; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.88 (s, 1H, CH), 4.12 (br s, 2H, NH_2), 1.60 (s, 9H, $C(CH_3)_3$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 155.1, 147.2, 131.2, 88.5, 85.2, 28.1 **LRMS** (LCMS-ESI) *m/z* calc. for $C_8H_{12}{}^{79}BrN_3O_2$ 261.01 found 205.1 $[M+H\text{-}t\text{Bu}]^+$.**146**: **MP** 126–128 °C; **IR** (ATR, cm^{-1}) 3512, 3281, 3221, 3173, 1767, 1619; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.38 (s, 1H, ArH), 5.39 (br s, 2H, NH_2), 1.65 (s, 9H, $C(CH_3)_3$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 149.9, 147.1, 142.8, 86.2, 75.7, 28.1; **LRMS** (LCMS-ESI) *m/z* calc. for $C_8H_{12}{}^{81}BrN_3O_2$ 263.0 found 162.1 $[M\text{-}C(O)t\text{Bu}]^-$.

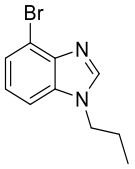
9.5.5 Hybrid series

4-Bromo-1*H*-benzo[*d*]imidazole **150a**[179]



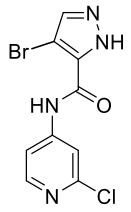
To a microwave vial was added 3-bromo-2-nitroaniline **151** (1.44 mmol 313 mg, 1.0 eq.), iron powder (14.4 mmol, 806 mg, 10 eq.) and ammonium chloride (14.4 mmol, 770 mg, 10 eq.) in a mixture of 4:1 EtOH: H_2O (2.2 mL, 0.7 M). The resulting suspension was heated to 70 °C for 2 hours. The mixture was then cooled to room temperature and formic acid (7.2 mL, 0.2 M) in EtOH (5 mL, 0.3 M) was added. The mixture was then heated to 70°C and allowed to stir at this temperature overnight. Upon completion, the mixture was cooled to room temperature and the solids were filtered and washed with EtOH (3×15 mL). The filtrate was concentrated *in vacuo* and taken up in CH_2Cl_2 (20 mL). The organics were washed with saturated $NaHCO_3$ solution (10 mL) and brine (10 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo* to provide the *title compound* **150a** as a brown solid (1.17 mmol, 231 mg, 81%). **150a** was carried onto the next step without further purification. **MP lit**[180] 168 °C; **MP** 159–160 °C; **IR** (ATR, cm^{-1}) 3021, 2791, 2751, 1579; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.61 (br s, 1H, NH), 8.12 (s, 1H, NCHN), 7.64 (m, 1H, ArH), 7.47 (d, *J* 7.5 Hz, 1H, ArH), 7.19 (app t, *J* 7.5 Hz, 1H, ArH); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 141.0, 138.2, 137.1, 126.1, 124.3, 115.1, 108.5; **LRMS** (LCMS-ESI) *m/z* calc. for $C_7H_5{}^{79}BrN_2$ 196.0 found 196.9 $[M+H]^+$.

4-Bromo-1-propyl-1*H*-benzo[*d*]imidazole **150b**



According to **General Procedure G**: amine **154** (0.87 mmol, 199 mg), triethylorthoformate (2.60 mmol, 0.43 mL) and *p*-TsOH (0.09 mmol, 15 mg) in PhMe (2.9 mL) were stirred at reflux overnight. Further purification *via* flash column chromatography 1:1 petroleum ether: EtOAc provided the *title compound* **150b** as an orange oil (0.84 mmol, 200 mg, 97%). **IR** (ATR, cm^{-1}) 3078, 2967, 2933, 2882, 1501; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.95 (s, 1H, ArH), 7.47 (dd, *J* 8.0, 0.8 Hz, 1H, ArH), 7.35 (dd. *J* 8.0, 0.8 Hz, 1H, ArH), 7.17 (app t, *J* 8.0 Hz, 1H. ArH), 4.14 (t, *J* 7.0 Hz, 2H, $NCH_2CH_2CH_3$), 1.92 (app sextet, *J* 7.0 Hz, 2H, $NCH_2CH_2CH_3$) 0.96 (t, J 7.0 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 143.5, 142.7, 134.5, 125.3, 123.9, 114.1, 109.2, 47.3, 23.3, 11.4; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{10}H_{11}{}^{79}BrN_2$ 238.0 found 240.9 $[M+2H]^+$.

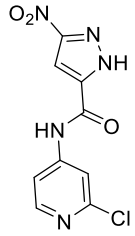
4-Bromo-*N*-(2-chloropyridin-4-yl)-1*H*-pyrazole-5-carboxamide **152a**



To a microwave vial was added acid **97b** (0.52 mmol, 200 mg, 1.0 eq.) and thionyl chloride (20.4 mmol, 1.56 mL, 20.4 eq.). The resulting mixture was heated to reflux overnight. The mixture was then concentrated *in vacuo* and taken up in toluene (10.5 mL, 0.1 M), cooled to 0°C and pyridine (3.22 mmol, 0.27 mL, 3.1 eq.) and DMF (0.3 mmol, 23 µL, 0.1 eq.) were added. To a separate microwave vial was added 2-chloropyridin-4-amine (0.70 mmol, 90 mg, 0.6 eq.) in toluene (0.5 mL, 2.0 M) and the resulting mixture was cooled to 0°C. To this mixture was added dropwise the mixture of acid chloride in toluene at 0°C. The mixture was then heated to reflux and stirred overnight. Upon completion, the mixture was quenched with H_2O (5 mL) and a precipitate formed. The solids were filtered and washed with H_2O (3×5 mL) and EtOAc (3×5 mL) to provide the *title compound* **152a** as a white solid (0.1 mmol, 27 mg, 9%). **MP** decomp. >240 °C; **IR** (ATR, cm^{-1}) 3357, 3132, 2904, 2809, 1717, 1581; **^{1}H NMR** (400 MHz, DMSO) δ 13.95 (br s, 1H, NNH), 10.81 (br s, 1H, C(O)NH), 8.28 (d, *J*

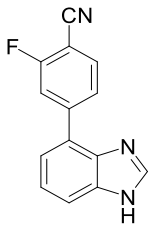
5.6 Hz, 1H, ArH), 8.22 (s, 1H, CHCBr) , 7.99 (d *J* 1.5 Hz, 1H, ArH), 7.81 (dd, *J* 5.6, 1.5 Hz, 1H, ArH); **^{13}C NMR** (101 MHz, DMSO) δ 160.4, 150.9, 150.2, 148.1, 141.4, 132.6, 113.3, 113.2, 93.7; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_6{}^{79}Br^{35}ClN_4O$ 299.9 found 301.0 $[M+H]^+$.

N-(2-Chloropyridin-4-yl)-3-nitro-1*H*-pyrazole-5-carboxamide **152b**



According to **General Procedure D**; 3-nitro-1*H*-pyrazole-5-carboxylic acid **116** (1.87 mmol, 294 mg), HATU (1.87 mmol, 711 mg), chloropyridin-4-amine **153** (1.56 mmol, 200 mg) and Hünigs base (3.12 mmol, 0.54 mL) in anhydrous DMF (0.5 mL) were stirred at 40°C overnight. Upon completion, the mixture was cooled to 0 °C and H_2O (1 mL) was added and a precipitate formed. The solids were filtered to provide the *title compound* **152b** as a white solid (1.50 mmol, 402 mg, 96%).**MP** decomp. >183 °C; **IR** (ATR, cm^{-1}) 3563, 3135, 2708, 2654, 1685, 1517; **^{1}H NMR** (400 MHz, DMSO) δ 15.15 (s, 1H, NNH), 10.91 (s. 1H, C(O)NH), 8.37 (d, *J* 5.6 Hz, 1H, ArH), 7.89 (d. *J* 1.8 Hz, 1H, ArH), 7.86 (s, 1H, ArH), 7.69 (dd, *J* 5.6, 1.8 Hz, 1H, ArH); **^{13}C NMR** (101 MHz, DMSO) δ 162.3, 156.9, 155.7, 151.0, 150.6, 147.5, 138.5, 113.4, 103.0; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_6{}^{35}ClN_5O_3$ 267.0 found 266.0 $[M-H]^-$.

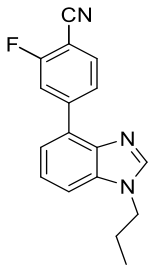
4-(1*H*-benzo[*d*]imidazol-4-yl)-2-fluorobenzonitrile **155a**



To a microwave vial was added bromo benzimidazole **150a** (0.51 mmol, 100 mg, 1.0 eq.), 2-fluoro-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzonitrile (0.76 mmol, 188 mg, 1.5 eq.) and Cs_2CO_3 (1.53 mmol, 498 mg, 3.0 eq.) in a 3:1 mixture of dioxane/ H_2O (5.1 mL, 0.1 M). The resulting mixture was degassed before the addition of palladium bistriphenylphosphine dichloride ($Pd(PPh_3)_2Cl_2$) (0.03 mmol, 19mg, 0.1 eq.). The resulting mixture was degassed once more before being subjected to microwave irradiation at 120°C for 1 hour. Upon completion, the mixture was filtered

through a plug of Celite® followed by washing with CH_2Cl_2 (3×10 mL). The filtrate was washed with H_2O (10 mL) and the aqueous was extracted with CH_2Cl_2 (2 × 10 mL). The combined organics were washed with brine (10 mL), dried over $MgSO_4$ and concentrated *in vacuo*. Further purification *via* flash column chromatography in 4:1 petroleum ether/EtOAc afforded the *title compound* **155a** as a brown solid (0.30 mmol, 71 mg, 59%).**MP** 148–150 °C; **IR** (ATR, cm^{-1}) 3132, 3097, 2964, 2233, 1622, 1558; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.47 (br s, 1H, NH), 8.14 (s, 1H, NCHN), 8.05–7.88 (m, 2H, 2×ArH), 7.74–7.71 (m, 1H, ArH), 7.63–7.55 (m, 1H, ArH), 7.50–7.40 (m, 2H, 2×ArH); **^{13}C NMR** (126 MHz, DMSO-d_6) δ 162.5 (d, $J_{C\text{-}F}$ 264 Hz), 145.8 (d, $J_{C\text{-}F}$ 8.9 Hz) 142.8, 133.6, 125.2 (d, $J_{C\text{-}F}$ 2.4 Hz), 122.6, 121.0, 116.0 (d, $J_{C\text{-}F}$ 20.8 Hz), 114.3, 113.4, 98.0 (d, $J_{C\text{-}F}$ 14.5 Hz), 3 C missing; **^{19}F NMR** (376 MHz, DMSO-d_6) δ -108.82 (s, 1F); **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_8FN_3$ 237.1 found 238.1 $[M+H]^+$.

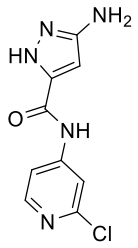
2-Fluoro-4-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)benzonitrile **155b**



To a microwave vial was added bromo benzimidazole **150b** (0.27 mmol, 65 mg, 1.0 eq.), 2-fluoro-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzonitrile (0.40 mmol, 100 mg, 1.5 eq.) and Cs_2CO_3 (0.81 mmol, 264 mg, 3.0 eq.) in a 3:1 mixture of dioxane/ H_2O (2.7 mL, 0.1 M). The resulting mixture was degassed with N_2 before the addition of palladium bistriphenylphosphine dichloride ($Pd(PPh_3)_2Cl_2$) (0.03 mmol, 19mg, 0.1 eq.). The resulting mixture was degassed with N_2 once more before being subjected to microwave irradiation at 120°C for 1 hour. Upon completion, the mixture was filtered through a plug of Celite® and washed with CH_2Cl_2 (3×10 mL). The filtrate was washed with H_2O (10 mL) and the aqueous was extracted with CH_2Cl_2 (2×10 mL), washed with brine (10 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography in 9:1 to 4:1 petroleum ether/EtOAc afforded the *title compound* **155b** as a white solid (0.22 mmol, 62 mg, 82%). **MP** 104–106 °C; **IR** (ATR, cm^{-1}) 3081, 2967, 2876, 2236,1619, 1492; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.05 (s, 1H, ArH), 8.00–7.95 (m, 2H, 2×ArH), 7.74–7.70 (m, 1H, ArH), 7.51–7.40 (m, 3H, 3×ArH), 4.22 (t, *J* 7.0 Hz, 2H, $NCH_2CH_2CH_3$), 1.97 (app sextet, *J* 7.0 Hz, 2H, $NCH_2CH_2CH_3$), 1.00 (t, *J* 7.0 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 163.4 (d, $J_{C\text{-}F}$ 263 Hz),145.9 (d, $J_{C\text{-}F}$ 8.5 Hz), 143.6, 141.2, 134.8, 133.4, 129.6, 125.5

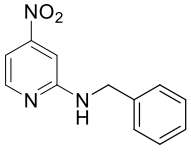
(d, $J_{C\text{-}F}$ 2.8 Hz), 123.5, 121.9, 117.0 (d, $J_{C\text{-}F}$ 18.2 Hz), 114.6, 111.1 , 99.8 (d, $J_{C\text{-}F}$ 15.6 Hz), 47.2, 23.3, 11.5; **^{19}F NMR** (376 MHz, $CDCl_3$) δ -106.5—-106.6 (m, 1F);. **LRMS** (LCMS-ESI) *m/z* calc. for $C_{17}H_{14}FN_3$ 279.1 found 280.1 $[M+H]^+$.

3-Amino-*N*-(2-chloropyridin-4-yl)-1*H*-pyrazole-5-carboxamide **156**



To a microwave vial was added nitro amide **152b** (0.37 mmol, 100 mg, 1.0 eq.), iron powder (3.7 mmol, 207 mg, 10 eq.) and ammonium chloride (3.7 mmol, 198 mg, 10 eq.) in a 4:1 mixture of EtOH: H_2O (0.6 mL). The resulting mixture was heated to 70 °C for 2 hours. Upon completion, the reaction mixture was cooled to room temperature and filtered through a pad of Celite®. The Celite® was washed with EtOH (2×20 mL) and the filtrate was concentrated *in vacuo*. The resulting residue was suspended in EtOAc (20 mL) and washed with H_2O (10 mL). The layers were separated and the aqueous was extracted with EtOAc (3×10 mL), washed with brine (10 mL), dried over $MgSO_4$ and concentrated *in vacuo* to provide the *title compound* **156** as a white solid (0.18 mmol, 43 mg, 49%). **MP** 212–214 °C; **IR** (ATR, cm^{-1}) 3439, 3299, 3182, 1687, 1575; **^{1}H NMR** (400 MHz, DMSO) δ 12.25 (s, 1H, NNH), 10.45 (s, 1H, C(O)NH), 8.24 (d, *J* 5.0 Hz, 1H, ArH), 7.99 (s, 1H, ArH), 7.82 (d, *J* 5.0 Hz, 1H, ArH), 5.74 (s, 1H, CHCNH$_2$), 5.26 (s, 2H, NH$_2$); **^{13}C NMR** (101 MHz, DMSO) δ 162.0, 150.8, 150.1, 149.7, 148.4, 145.6, 113.2, 113.0, 87.8; **LRMS** (LCMS-ESI) *m/z* calc. for $C_9H_8{}^{35}ClN_5O$ 237.0 found 238.0 $[M+H]^+$.

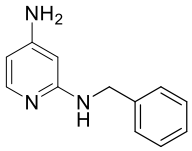
N-Benzyl-4-nitropyridin-2-amine **159**



To a flame dried flask purged with argon was added 2-chloro-4-nitropyridine **158** (4.74 mmol, 750 mg, 1.0 eq.), Cs_2CO_3 (23.7 mmol, 8.19 g, 5.0 eq.), palladium acetate (0.02 mmol, 5 mg, 0.5 mol%), *rac*-BINAP (0.09 mmol, 59 mg, 2 mol%) and benzylamine (23.7 mmol, 0.17 mL, 5.0 eq.) in anhydrous toluene (47.4 mL, 0.1 M). The resulting mixture was degassed for 20 mins and then heated to 70°C overnight. The mixture

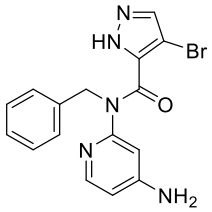
was then cooled to room temperature and quenched with H_2O (10 mL). The layers were separated and the aqueous was extracted with CH_2Cl_2 (4×20 mL), washed with brine (20 mL), dried over $MgSO_4$ and concentrated *in vacuo*. Further purification *via* flash column chromatography in 95:5 PhMe/EtOAc afforded the *title compound* **159** as a yellow solid (1.86 mmol, 426 mg, 39%). **MP** 110–112 °C; **IR** (ATR, cm^{-1}) 3126, 3091, 2939, 1574, 1530; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.31–8.30 (m, 1H, ArH), 7.37–7.25 (m, 6H, 6×ArH), 7.09–7.08 (m, 1H, ArH), 5.39 (br s, 1H, NH), 4.59 (d, *J* 5.7 Hz, 2H, CH_2); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 160.0, 155.7, 150.6, 138.0, 129.0, 127.9, 127.6, 105.5, 99.8, 46.5; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{12}H_{11}N_3O_2$ 229.1 found 230.1 $[M+H]^+$.

N^2-Benzylpyridine-2,4-diamine **160**



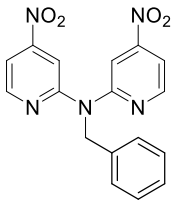
To a microwave vial was added nitro benzylamine **159** (0.87 mmol, 200 mg, 1.0 eq.), iron powder (8.72 mmol, 487 mg, 10.0 eq.) and NH_4Cl (8.72 mmol, 466 mg, 10.0 eq.) in a mixture of 4:1 EtOH: H_2O (1.3 mL, 0.7 M). The resulting mixture was heated to 50°C overnight. The mixture was then filtered through Celite®. The Celite® was washed with EtOH (10 mL) and concentrated *in vacuo*. The residue was taken up in CH_2Cl_2 (10 mL) and washed with $NaHCO_3$ (10 mL). The aqueous was extracted with CH_2Cl_2 (5×10 mL), dried over $MgSO_4$ and concentrated *in vacuo*. Further purification via flash column chromatography in 1:1 petroleum ether/EtOAc to 9:1 CH_2Cl_2 provided the *title compound* **160** as a purple solid (0.40 mmol, 80 mg, 46%). **MP** 120–122 °C**; IR** (ATR, cm^{-1}) 3439, 3246, 3307, 1606, 1577; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.76 (d, *J* 5.8 Hz, 1H, ArH), 7.37–7.24 (m, 5H, 5×ArH), 5.97 (dd, *J* 5.8, 2.0 Hz, 1H, ArH), 5.58 (d, *J* 2.0 Hz, 1H, ArH), 5.30 (br s, 1H, NH), 4.44 (d, *J* 5.8 Hz, 2H, CH_2), 3.97 (br s, 2H, NH_2); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 160.0, 154.4, 148.8, 139.4, 128.8, 127.4, 127.3, 102.2, 90.3, 46.7; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{12}H_{13}N_3$ 199.1 found 200.1 $[M+H]^+$.

N-(4-Aminopyridin-2-yl)-*N*-benzyl-4-bromo-1*H*-pyrazole-5-carboxamide **161**



According to **General Procedure D**; acid **97b** (0.30 mmol, 57 mg), HATU (0.30 mmol, 114 mg), amine **160** (0.25 mmol, 45 mg) and Hünigs base (0.50 mmol, 0.09 mL) in anhydrous DMF (0.8 mL) were stirred at 40 °C overnight. Partial purification *via* flash column chromatography in 100% EtOAc afforded the *title compound* **161** as a beige solid (0.02 mmol, 9 mg, 10%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.05–8.04 (m, 1H, ArH), 7.44 (s, 1H, ArH), 7.39–7.36 (m, 2H, 2×ArH), 7.32–7.24 (m, 5H, 5×ArH), 6.36–6.34 (m, 1H, ArH), 6.10–6.09 (m, 1H, ArH), 5.25 (s, 2H, CH_2), 4.13 (br s, 2H, NH_2); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 155.6, 154.4, 149.2, 137.2, 128.6, 128.5, 127.1, 108.6, 106.7, 96.7, 95.0, 51.78, 29.9, 3 C missing; **LRMS** (LCMS-ESI) *m/z* calc. $C_{16}H_{14}{}^{79}BrN_5O$ 371.0 found 370.0 $[M-H]^-$.

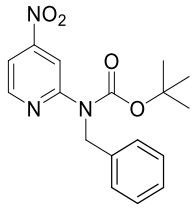
N-benzyl-4-nitro-*N*-(3-nitrophenyl)pyridin-2-amine **162**



To a flame dried flask purged with argon was added 2-chloro-4-nitropyridine **158** (1.26 mmol, 200 mg, 1.0 eq.), Cs_2CO_3 (1.51 mmol, 492 mg, 1.2 eq.), palladium acetate (0.01 mmol, 2 mg, 0.5 mol%), *rac*-BINAP (0.03 mmol, 16 mg, 2 mol%) and benzylamine (1.51 mmol, 0.17 mL, 1.2 eq.) in anhydrous toluene (12.6 mL, 0.1 M). The resulting mixture was degassed for 20 mins and then heated to 80°C overnight. The mixture was then cooled to room temperature and quenched with H_2O (10 mL). The layers were separated and the aqueous was extracted with CH_2Cl_2 (4×20 mL), washed with brine (20 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. Further purification *via* flash column chromatography in 95:5 PhMe/EtOAc afforded the *title compound* **162** as a yellow solid (0.46 mmol, 162 mg, 38%). **MP** 90–92 °C; **IR** (ATR, cm^{-1}) 3237, 3034, 2964, 1568, 1527; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.59 (dd, *J* 5.4, 0.4 Hz, 2H, 2×ArH) 8.02 (dd, *J* 1.9, 0.4 Hz, 2H, 2×ArH), 7.65 (dd, *J* 5.4, 1.9 Hz, 2H, 2×ArH) 7.33–7.22 (m, 6H, 6×ArH), 5.59 (s, 2H, CH_2); **^{13}C NMR** (101 MHz, $CDCl_3$)

δ 157.9, 155.5, 150.4, 136.9, 129.1, 127.7, 127.0, 110.7, 107.8, 52.2; **LRMS** (LCMS-ESI) *m/z* calc. $C_{18}H_{14}N_4O_4$ 350.1 found 352.0 $[M+2H]^+$.

tert-Butyl benzyl(4-nitropyridin-2-yl)carbamate **163**



To a flame dried microwave vial purged with argon was added nitro benzylamine **159** (1.76 mmol, 403 mg, 1.0 eq.) and DMAP (0.34 mmol, 42 mg, 0.2 eq.) in anhydrous CH_2Cl_2 (4.4 mL, 0.4 M). The resulting mixture was cooled to 0°C before the addition of Et_3N (2.82 mmol, 0.39 mL, 1.6 eq.) and di-*tert*-butyl dicarbonate (2.64 mmol, 576 mg, 1.5 eq.). The mixture was then heated to 40°C for 3 days. Upon completion, the reaction mixture was cooled to 0°C and quenched with 1M aqueous HCl solution (5 mL) and allowed to warm to room temperature. The organics were extracted with CH_2Cl_2 (3 × 10 mL), washed with sat. aqueous $NaHCO_3$ solution (10 mL), brine (10 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The *title compound* **163** was isolated without any further purification as a light yellow solid (1.70 mmol, 560 mg, 97%). **MP** 50–52 °C; **IR** (ATR, cm^{-1}) 3141, 3007, 2969, 1716, 1571; **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.66 (d, 2.0 Hz, 1H, ArH), 8.56 (dd, J 5.5, 0.5 Hz, 1H, ArH), 7.66 (dd, J 5.5, 2.0 Hz, 1H, ArH) 7.29–7.20 (m, 5H, 5×ArH), 5.28 (s, 2H, CH_2), 1.45 (s, 9H, $OC(CH_3)_3$); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 156.5, 154.9, 153.9, 149.3, 138.8, 128.5, 127.2, 127.2, 111.6, 111.3, 83.0, 49.8, 28.2; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{17}H_{19}N_3O_4$ 329.1 found 330.1, 274.1 and 230.1 $[M+H]^+$, $[M–t\text{Bu}]^+$ and $[M–C(O)Ot\text{Bu}]^+$, respectively.

10 Appendix

10.1 Kinome selectivity profile for CI709

CMGC Family	% Inhibition
DYRK2	87%
DYRK1A	8%
DYRK1B	7%
DYRK3	12%
DYRK4	1%
HIPK1	3%
HIPK2	2%
HIPK4	5%
CLK1	-1%
CLK2	-3%
CLK3	3%
CLK4	32%
GSKα	3%
GSKβ	5%
CDK1/Cyclin B	4%
CDK2/Cyclin A	1%
ICK	2%
CDK9/Cyclin T1	-2%
MAPK1 (ERK2)	6%

TK Family	% Inhibition
EGFR	-5%
ABL1	-3%
MET (cMET)	-1%
SRC	14%
JAK2	-16%

AGC Family	% Inhibition
ROCK1	-8%
AKT1 (PKBα)	1%

CAMK Family	% Inhibition
PIM1	-4%
PIM2	-6%
MAP2K2 (MEK2)	5%
CHEK1 (Chk1)	-2%

TK1 Family	% Inhibition
ALK	6%
BRAF	-1%
IRAK1	0%

STE Family	% Inhibition
PAK1	15%
SLK	3%

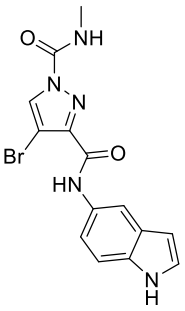
CK1 Family	% Inhibition
CSNK2A1 (CK2 α1)	-1%
CSNK1A1 (CK2 δ1)	-3%

Other	% Inhibition
Aurora A	1%
Aurora B	-11%
PLK1	6%
Wee1	2%
Haspin	24%

Atypical kinases	% Inhibition
FRAP 1 (mTOR)	-2%

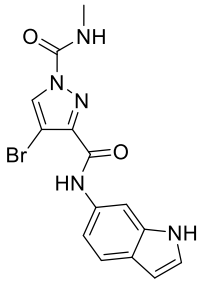
Legend
<20%
20%-50%
50-80%
>80%

10.2 SAR summary from previous work



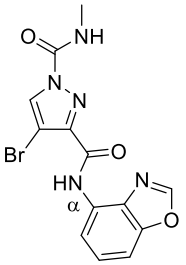
CI483

DSF DYRK2 :2.2°C
DSF DYRK1A:1.1°C
DYRK2 1μM:ND
DYRK1A 1μM:ND



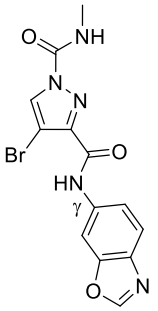
CI494

DSF DYRK2 :2.8°C
DSF DYRK1A:1.4°C
DYRK2 1μM:24%
DYRK1A 1μM:ND



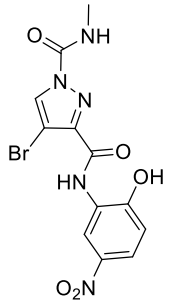
CI503

DSF DYRK2 :2.6°C
DSF DYRK1A:0.4°C
DYRK2 1μM:43%
DYRK1A 1μM:ND
IC50 = 1.71 μM



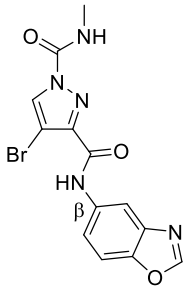
CI508

DSF DYRK2 :2.2°C
DSF DYRK1A:0.4°C
DYRK2 1μM:26%
DYRK1A 1μM:ND



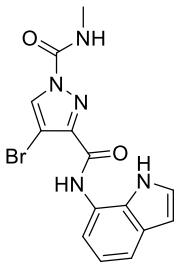
CI511

DSF DYRK2 :2.7°C
DSF DYRK1A:0.9°C
DYRK2 1μM:64%
DYRK1A 1μM:-13%
IC50 = 1.45 uM



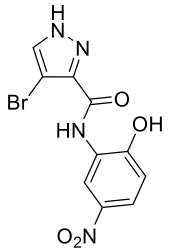
CI524

DSF DYRK2 :1.8°C
DSF DYRK1A:0.2°C
DYRK2 1μM:11%
DYRK1A 1μM:ND



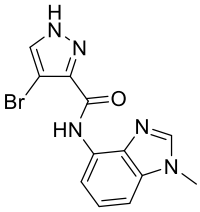
CI592

DSF DYRK2 :0.9°C
DSF DYRK1A:-0.2°C
DYRK2 1μM:4%
DYRK1A 1μM:ND



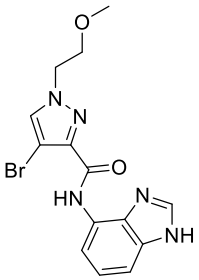
CI608

DSF DYRK2 :2.1°C
DSF DYRK1A:0.3°C
DYRK2 1μM:51%
DYRK1A 1μM:8%
IC50 = 1.1 μM



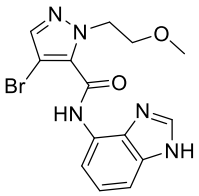
CI641

DSF DYRK2 :3.7°C
DSF DYRK1A:0.1°C
DYRK2 1μM:77%
DYRK1A 1μM:0%
IC50 = 141 nM



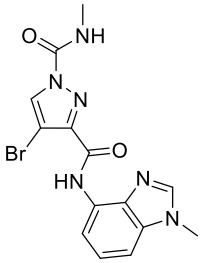
CI665

DSF DYRK2 :0.1°C
DSF DYRK1A:-0.4°C
DYRK2 1μM:6%
DYRK1A 1μM:ND



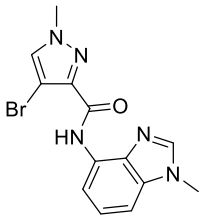
CI673

DSF DYRK2 :0.0°C
DSF DYRK1A:-0.1°C
DYRK2 1μM-4%
DYRK1A 1μM:ND



CI688

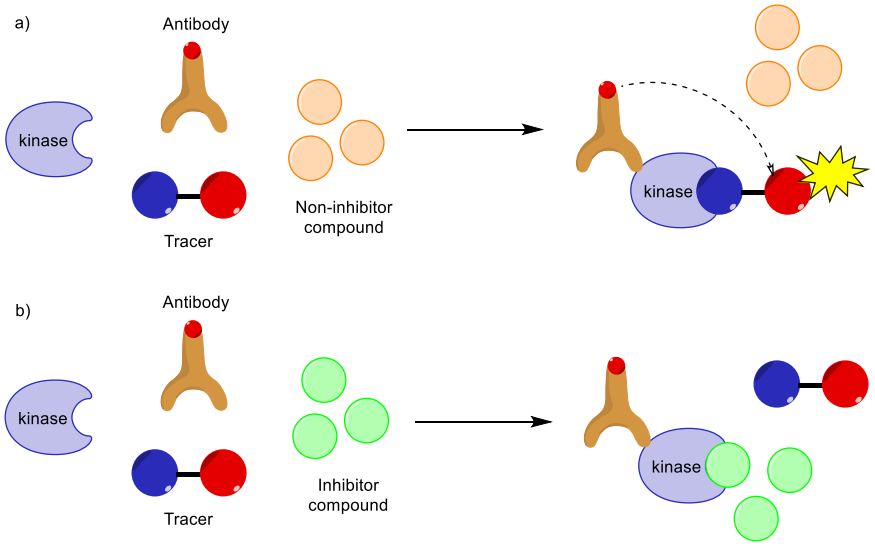
DSF DYRK2 :2.1°C
DSF DYRK1A:0.6°C
DYRK2 1μM:81%
DYRK1A 1μM:8%
IC50 = 70 nM



CI705

DSF DYRK2 :-0.1°C
DSF DYRK1A:-0.3°C
DYRK2 1μM:3%
DYRK1A 1μM:ND

10.3 Life Technologies LanthaScreen binding assay description



Scheme 43: Schematic representation of kinase binding assay

Life Technologies mix the desired kinase (DYRK2) with a europium (Eu) labelled anti-tag antibody, a fluorescent tracer (AlexaFluor™ 647) and the compound under investigation. **Scheme 43** describes the outcome of two scenarios: a) when the compound doesn't bind to the kinase and b) when the compound does bind to the kinase. If the compound does not bind to the kinase, the tracer molecule binds in its absence. In turn, this results in the Eu labelled antibody and the tracer being within close proximity of one another, resulting in an energy transfer which gives off a high fluorescence resonance energy transfer (FRET) reading. However, if the compound does bind to the kinase, (i.e. displays inhibitory properties) then the tracer cannot bind and therefore it is unable to interact with the antibody. This produces a low FRET reading (**Scheme 43**). The results are quantified as a % displacement with respect to a control reading (0% displacement for a non-inhibitor) and a known inhibitor (100% displacement). The following equations describe how these results are calculated:

$$\%\ Displacement = \left(\frac{ER_{0\%\ displ.} - ER_{sample}}{ER_{0\%\ displ.} - ER_{100\%\ displ.}}\right)100$$

$$Emission\ Ratio\ (ER) = \frac{tracer\ emission}{europium\ tag\ emission}$$

10.4 Cheng-Prusoff equation[181]

$$K_i = \frac{IC_{50}}{1 + \frac{S}{K_m}}$$

K_i = inhibitory constant

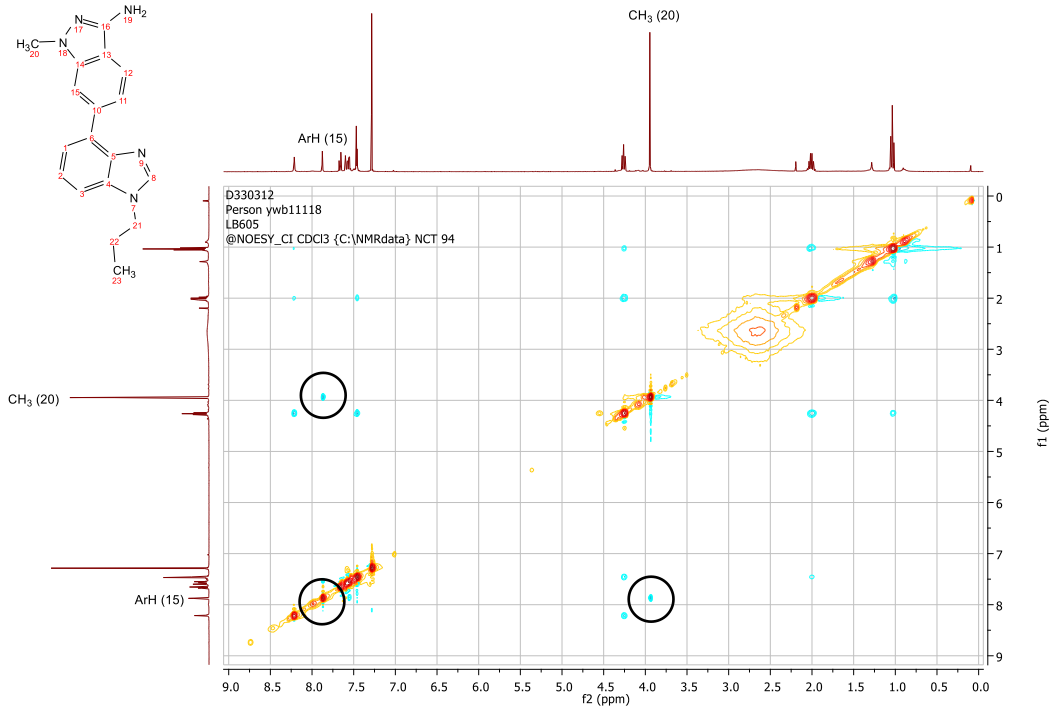
IC_{50} = concentration at 50% inhibition

S = substrate concentration

K_m = Michaelis constant for the enzyme substrate

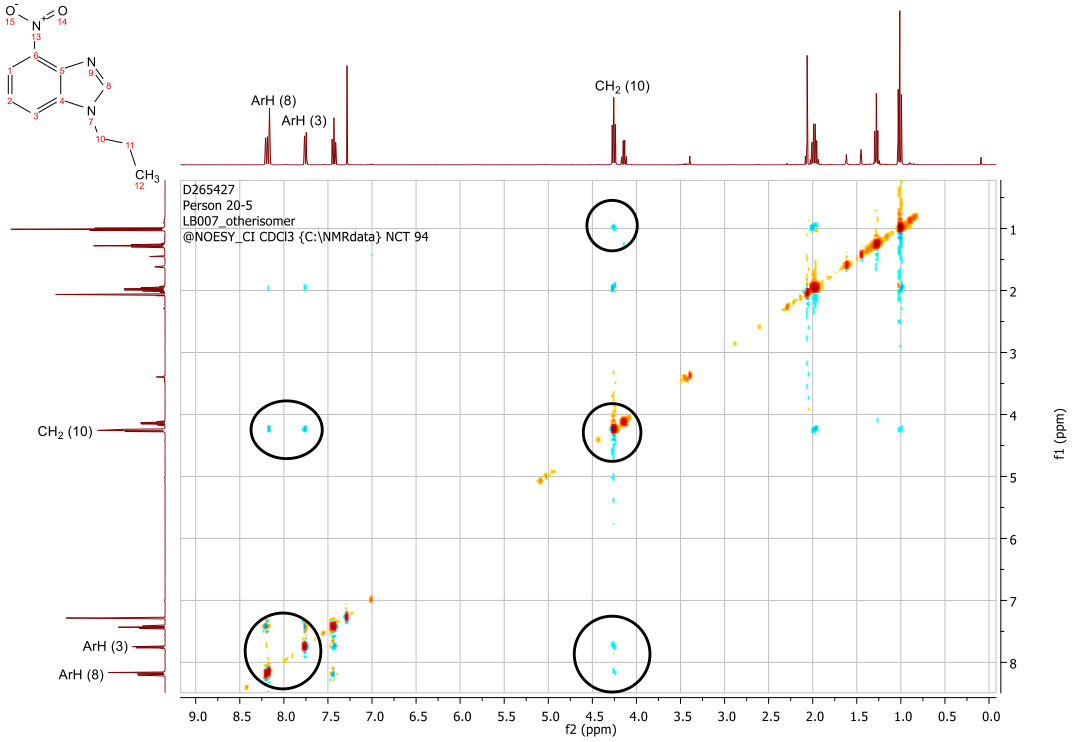
10.5 2D NMR spectra

10.5.1 1-Methyl-6-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-indazol-3-amine **91**



NOESY relationship between ArH (#15 in above structure) and N-CH_3 (20).

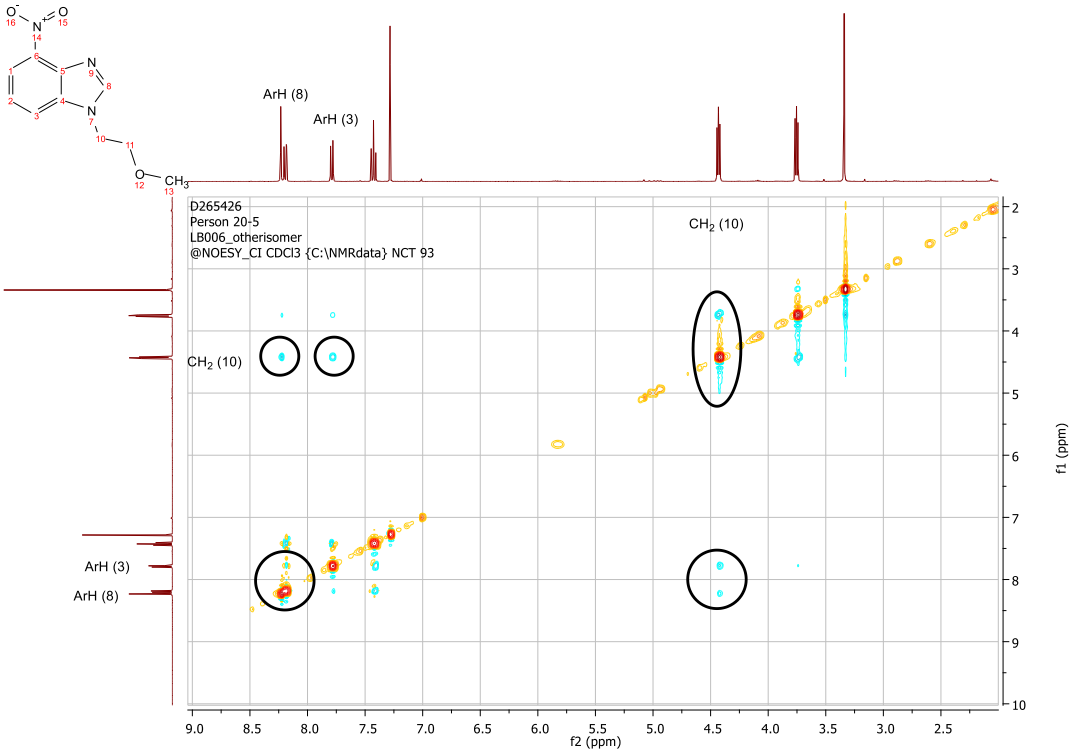
10.5.2 4-Nitro-1-propyl-1*H*-benzo[*d*]imidazole **122a**



NOESY relationship between N-C<u>H</u>$_2$ (#10 in above structure) and Ar<u>H</u> (3).

NOESY relationship between N-C<u>H</u>$_2$ (10) and Ar<u>H</u> (8).

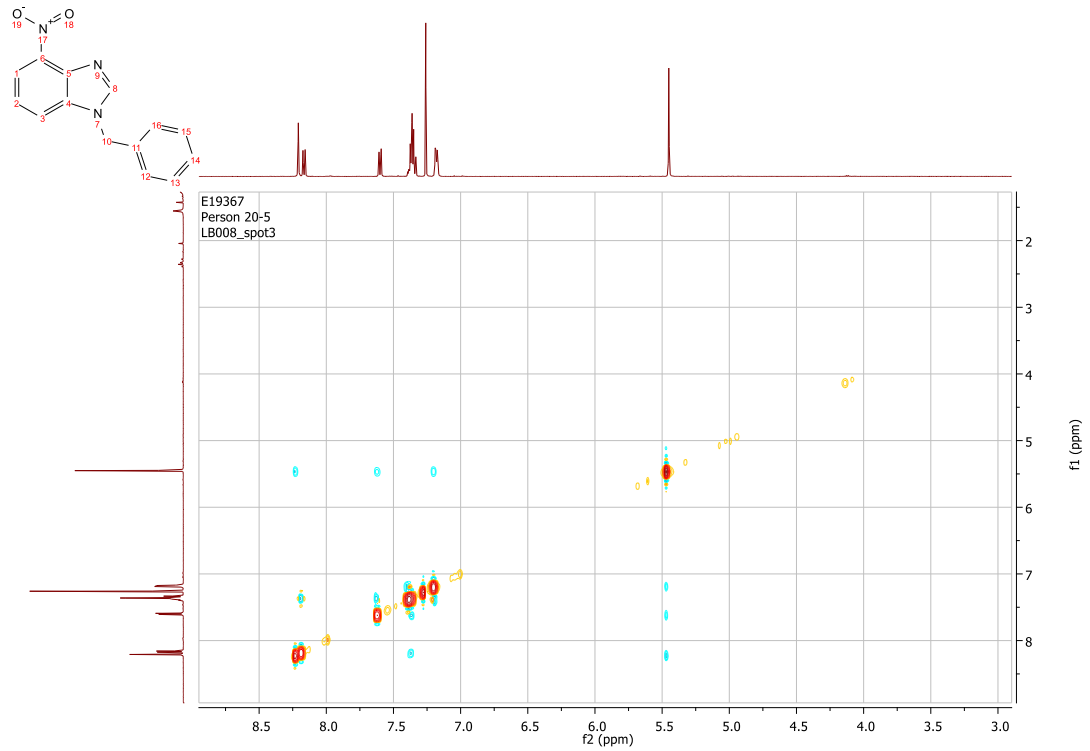
10.5.3 1-(2-Methoxyethyl)-4-nitro-1*H*-benzo[*d*]imidazole **122b**



NOESY relationship between ArH (3) and CH_2 (10).

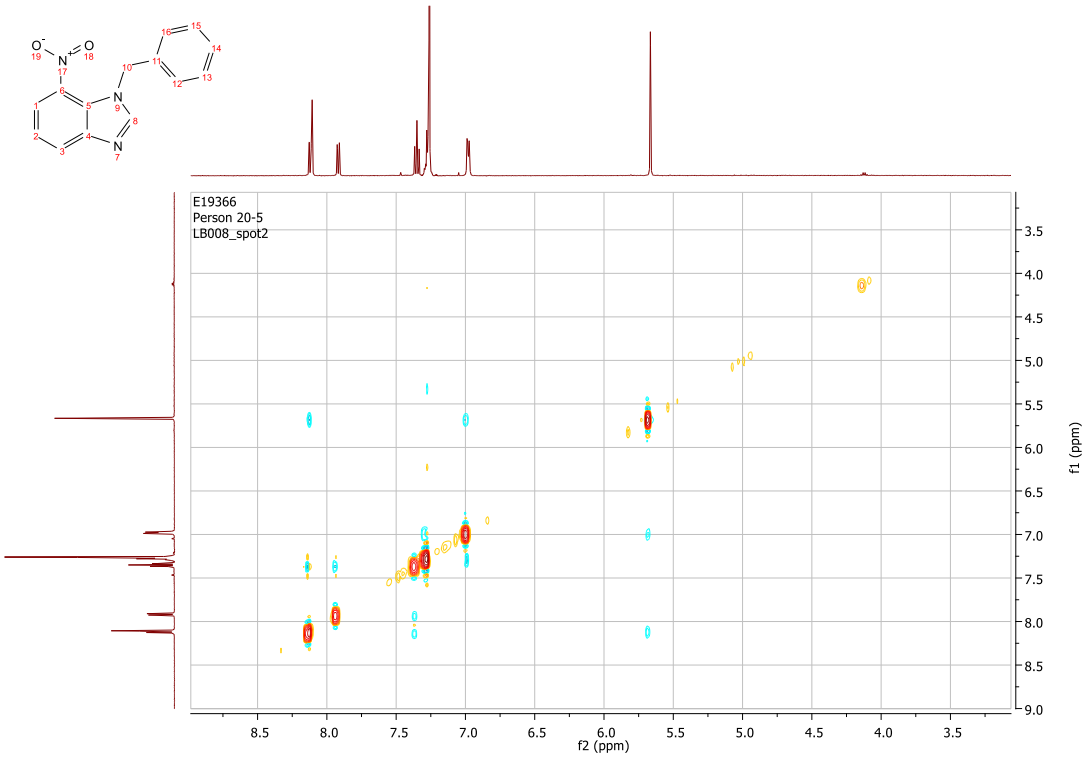
NOESY relationship between CH_2 (10) and ArH (8).

10.5.4 1-Benzyl-4-nitro-1*H*-benzo[*d*]imidazole **122c**

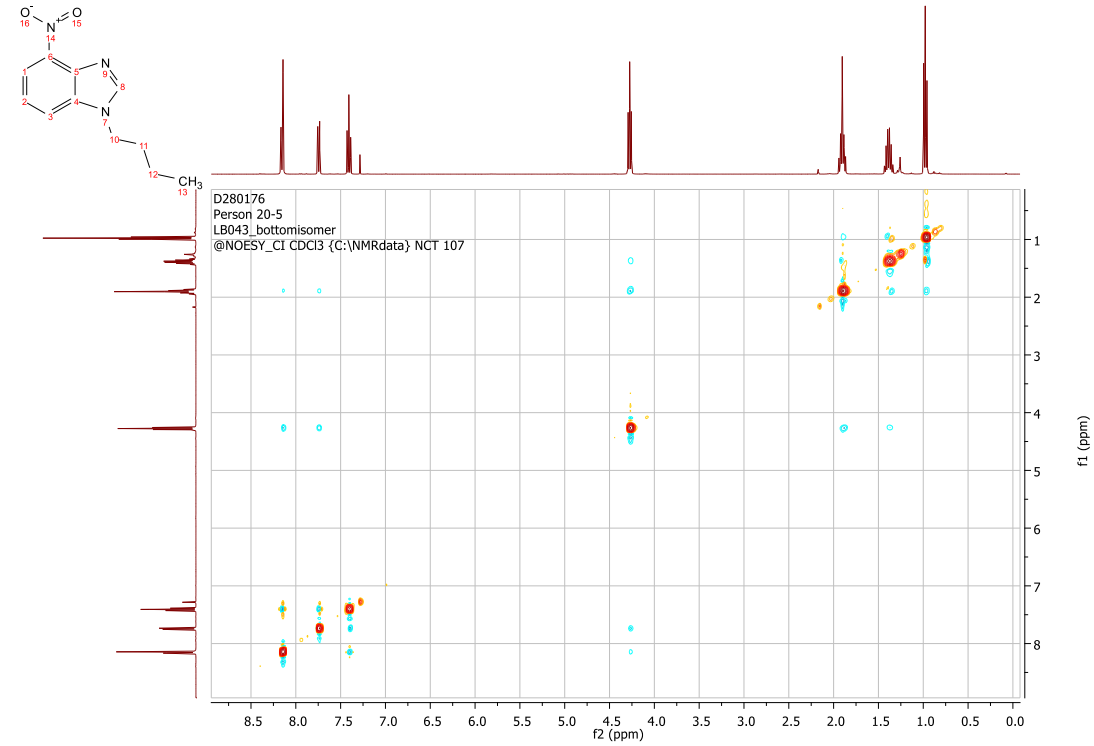


NOESY relationship between ArH (3) and CH_2 (10).

10.5.5 1-Benzyl-7-nitro-1*H*-benzo[*d*]imidazole **123c**

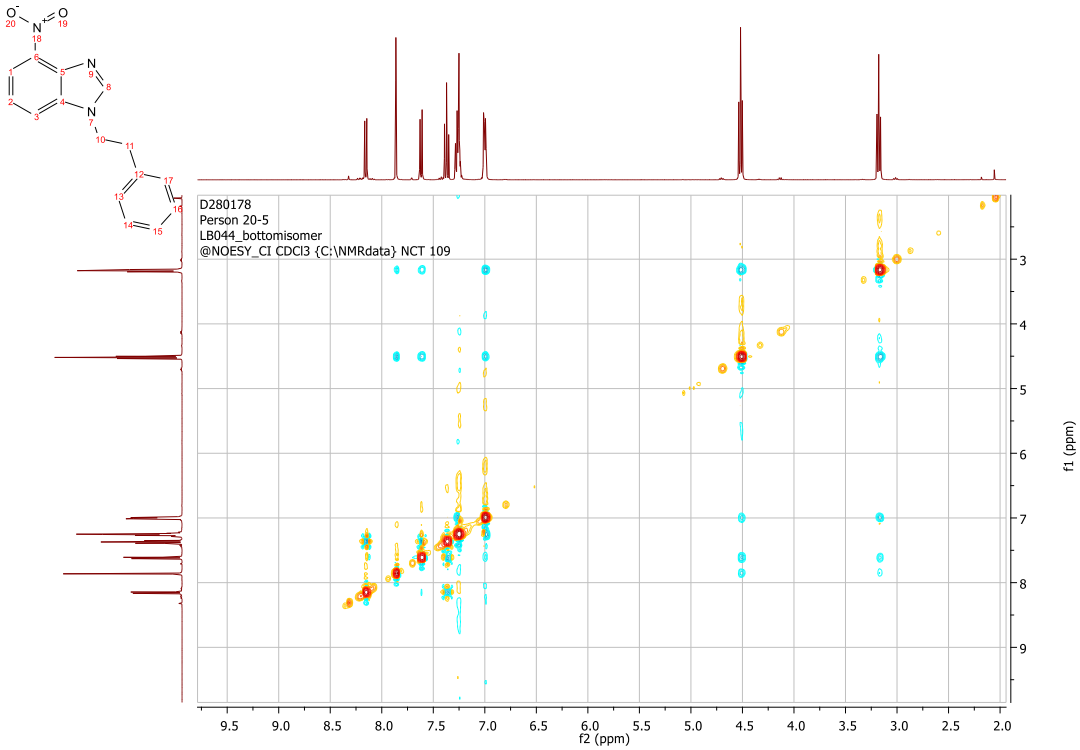


10.5.6 1-Butyl-4-nitro-1*H*-benzo[*d*]imidazole **122d**



NOESY relationship between ArH (3) and CH_2 (10).

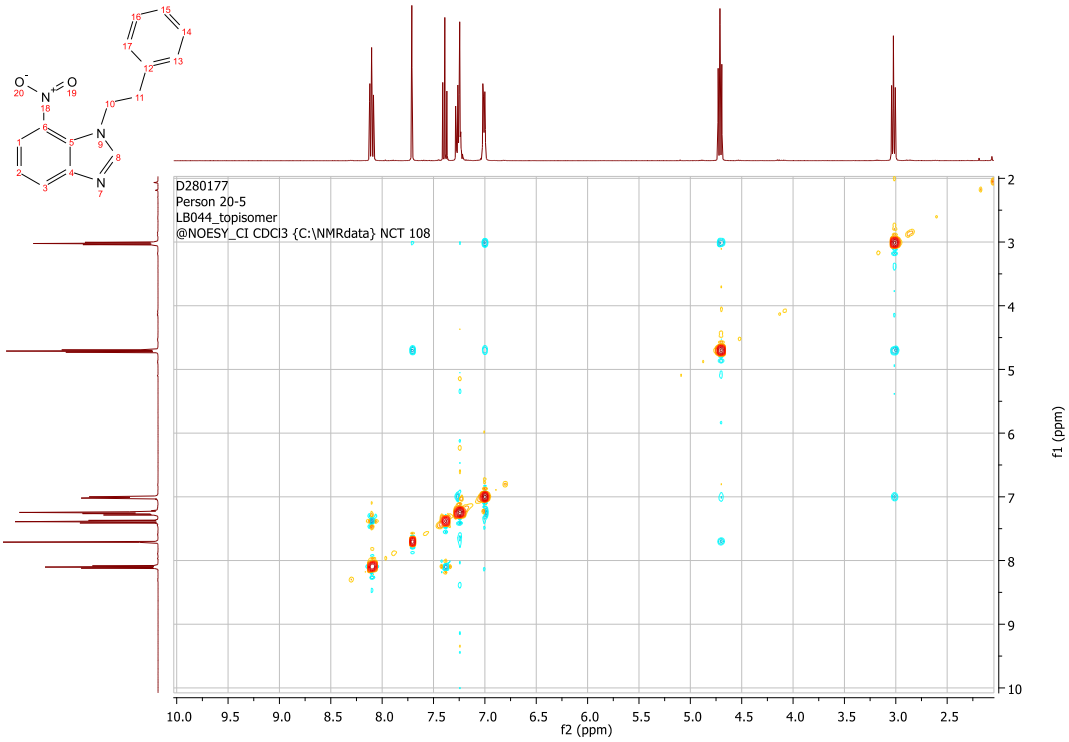
10.5.7 4-Nitro-1-phenethyl-1*H*-benzo[*d*]imidazole **122e**



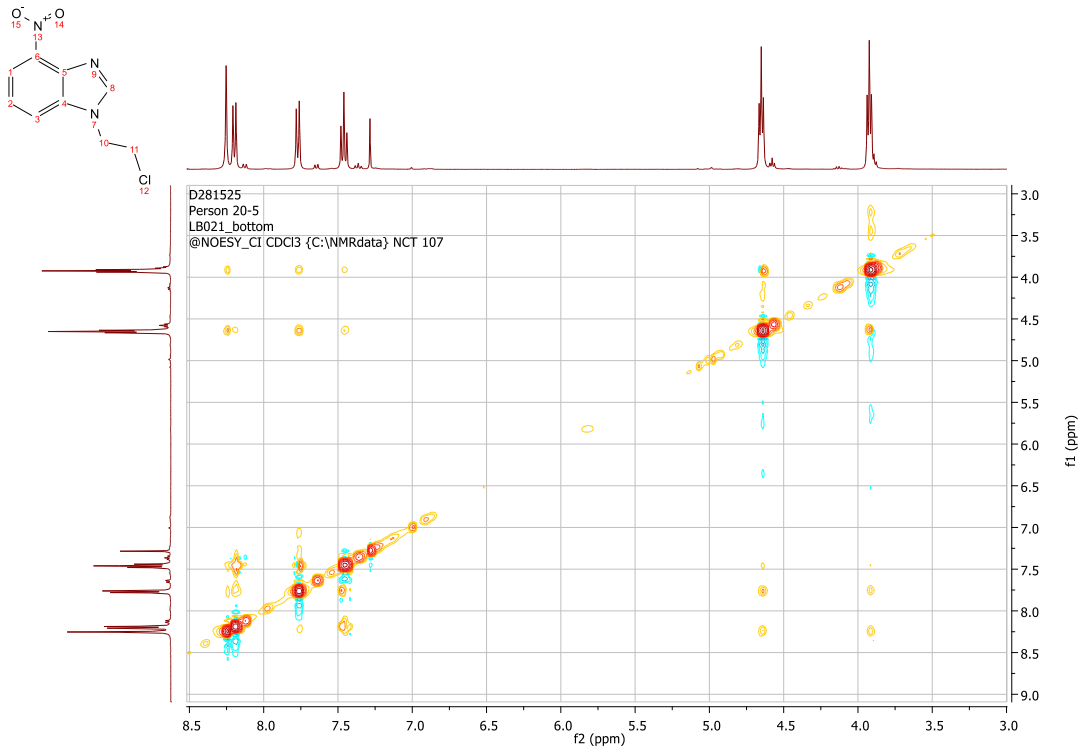
NOESY relationship between ArH (3) and CH_2 (10).

NOESY relationship between ArH (3) and CH_2 (11).

10.5.8 7-Nitro-1-phenethyl-1*H*-benzo[*d*]imidazole **123e**



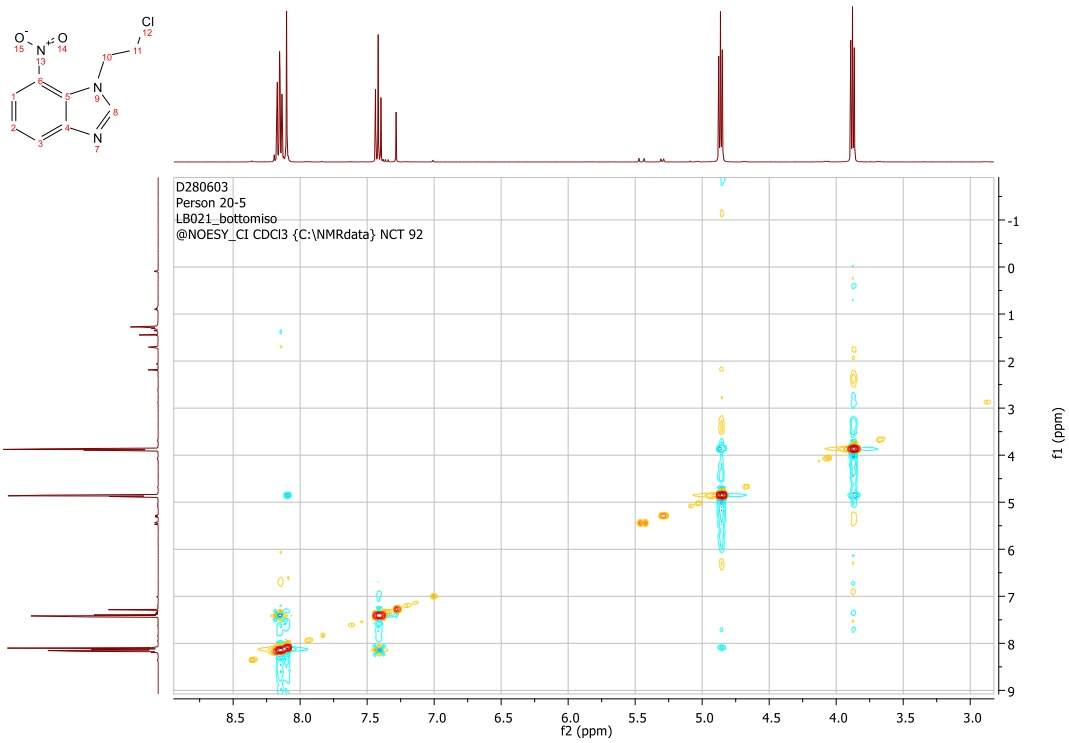
10.5.9 1-(2-Chloroethyl)-4-nitro-1H-benzo[d]imidazole **122f**



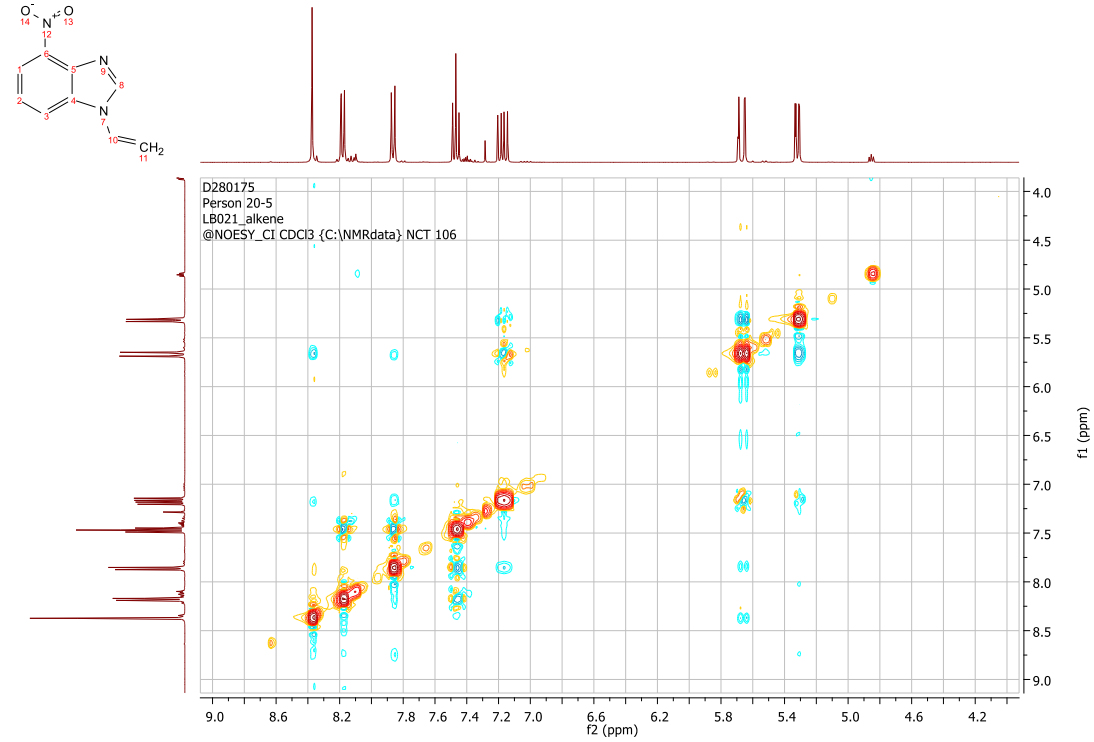
NOESY relationship between ArH (3) and CH_2 (10).

NOESY relationship between ArH (3) and CH_2 (11).

10.5.10 1-(2-Chloroethyl)-7-nitro-1*H*-benzo[*d*]imidazole **123f**

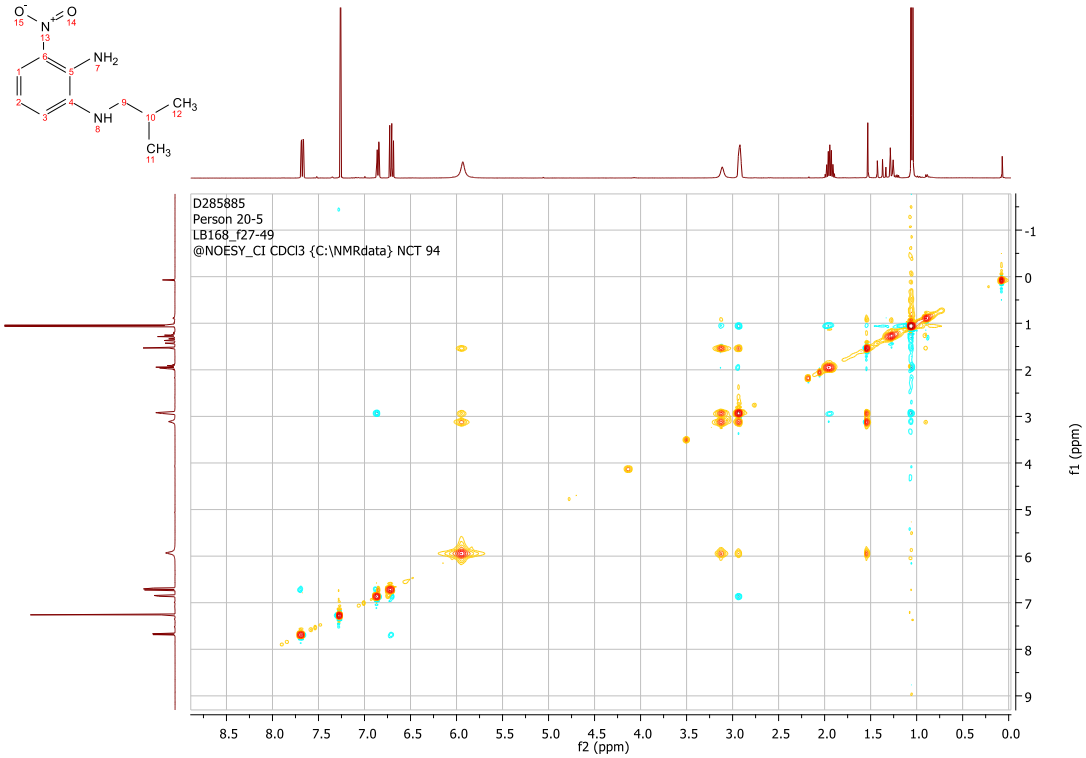


10.5.11 4-Nitro-1-vinyl-1*H*-benzo[*d*]imidazole **122g**



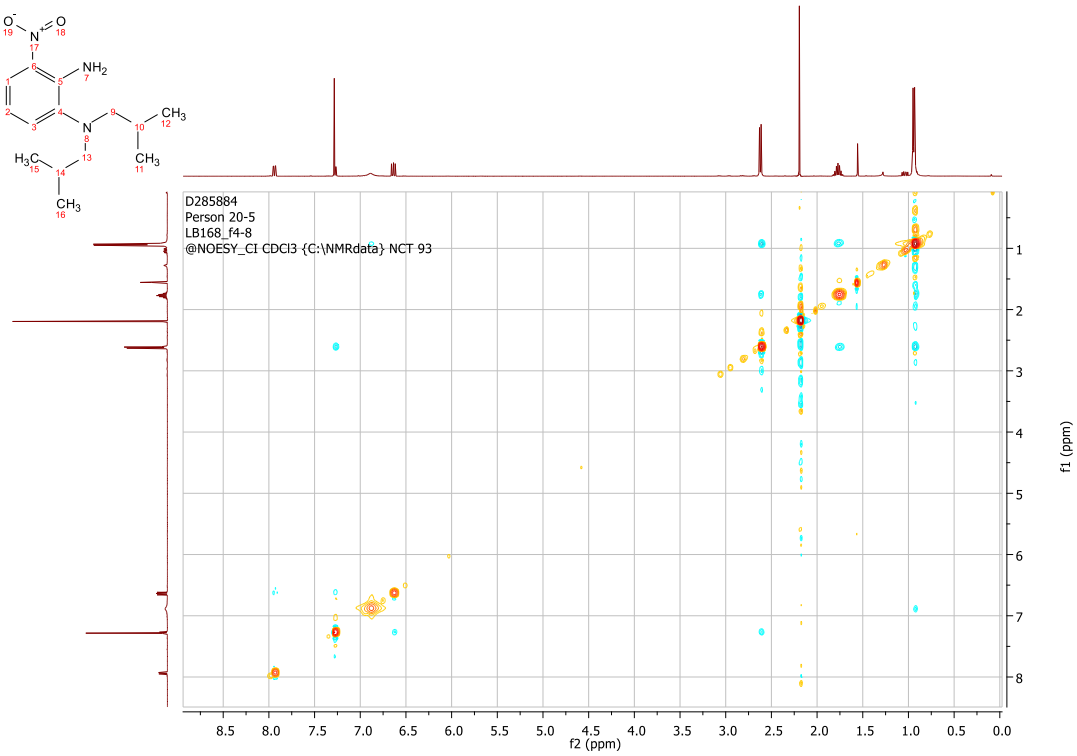
NOESY relationship between ArH (3) and CH_2 (10).

10.5.12 N^1-isobutyl-3-nitrobenzene-1,2-diamine **125a**



NOESY relationship between ArH (3) and CH_2 (9).

10.5.13 N^1,N^1-diisobutyl-3-nitrobenzene-1,2-diamine **125j**

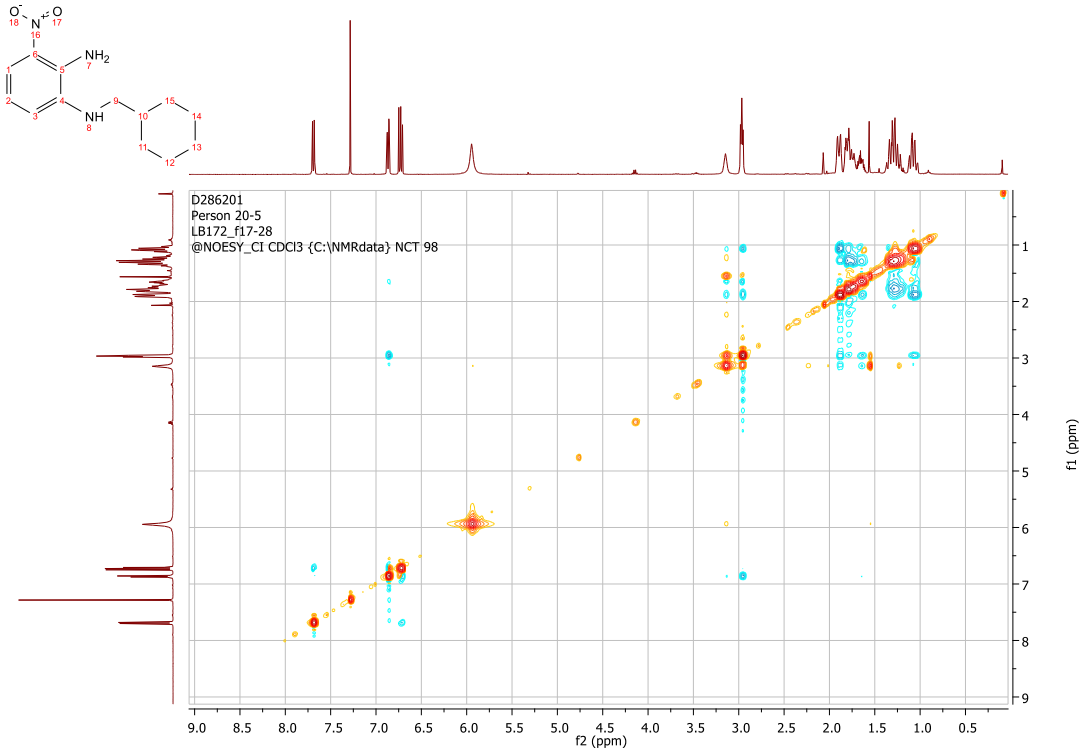


NOESY relationship between ArH (3) and CH_2 (9/13).

NOESY relationship between ArH (3) and CH (10/14).

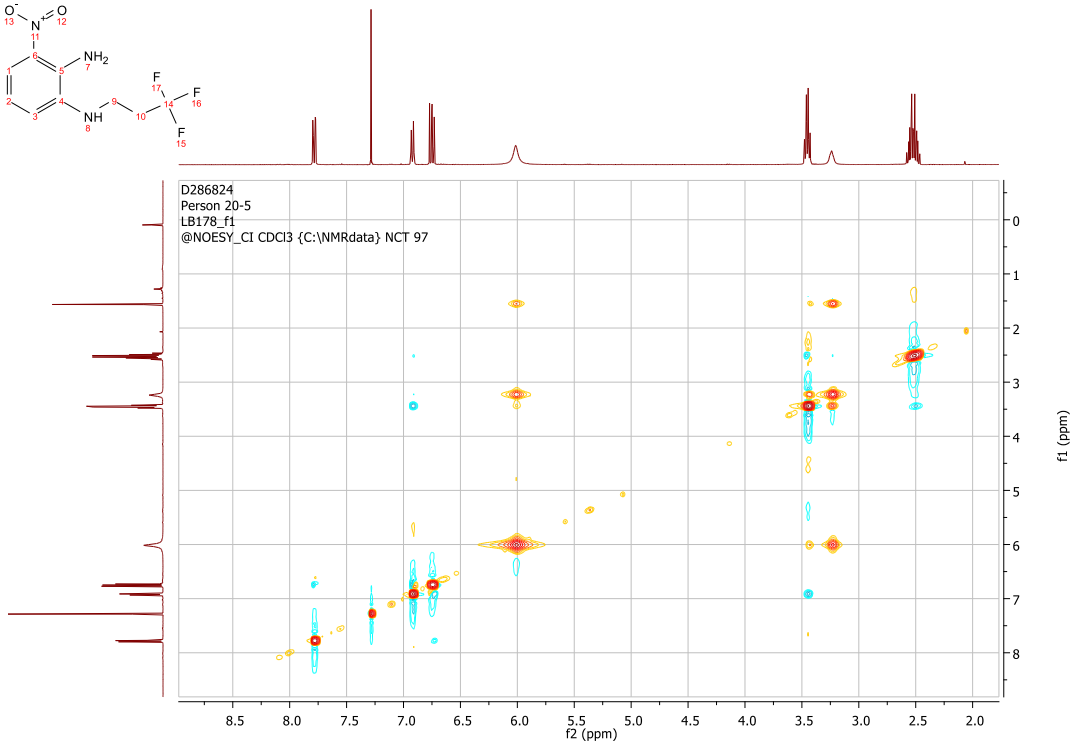
NOESY relationship between ArH (3) and CH_3 (11/12 or 15/16).

10.5.14 N^1-(cyclohexylmethyl)-3-nitrobenzene-1,2-diamine **125b**



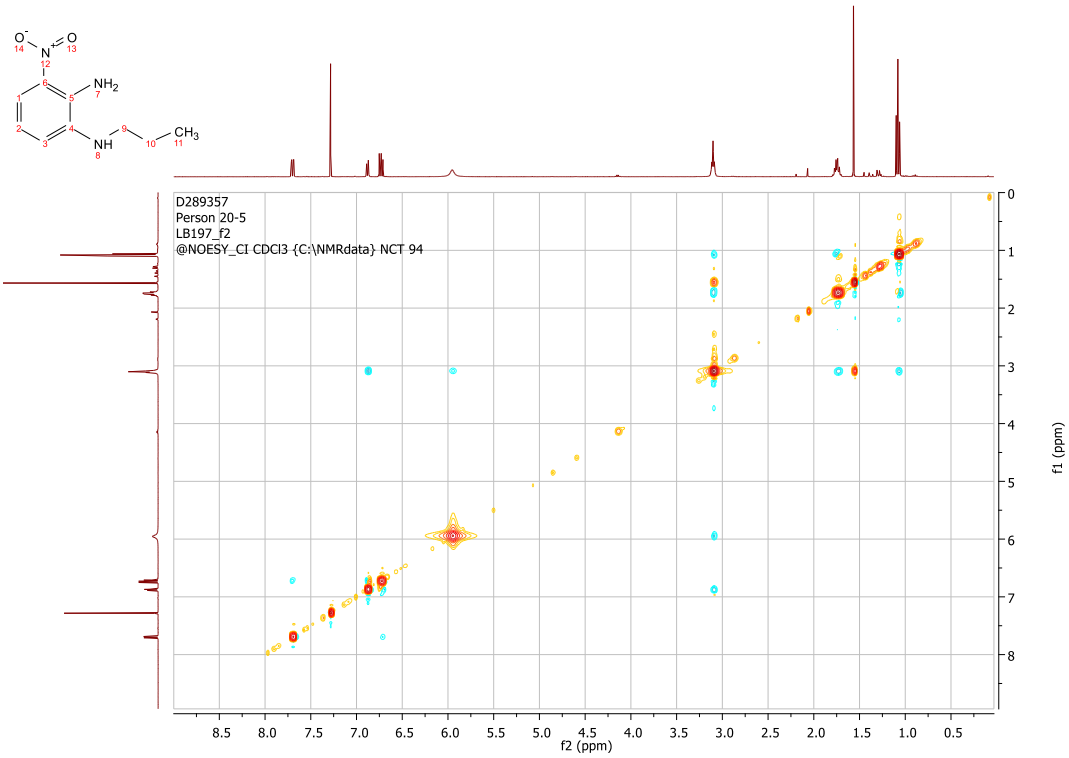
NOESY relationship between ArH (3) and CH_2 (9).

10.5.15 3-Nitro-*N*[1]-(3,3,3-trifluoropropyl)benzene-1,2-diamine **125c**



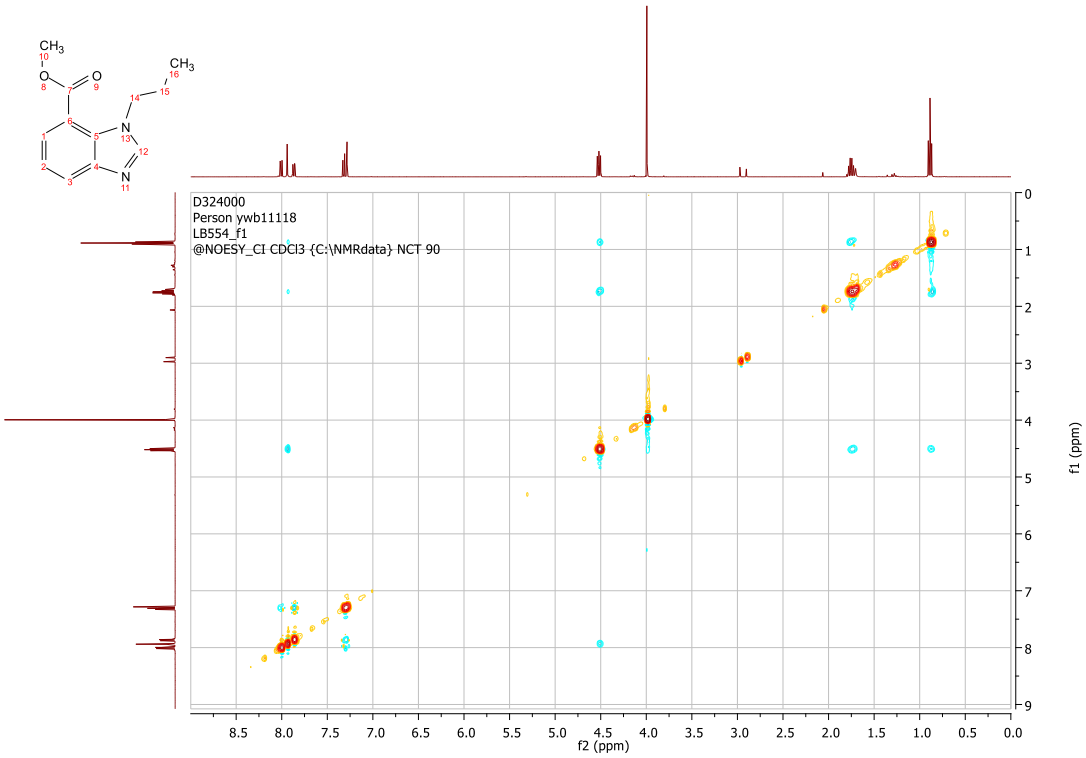
NOESY relationship between ArH (3) and CH_2 (9).

10.5.16 3-Nitro-N^1-propylbenzene-1,2-diamine **125f**

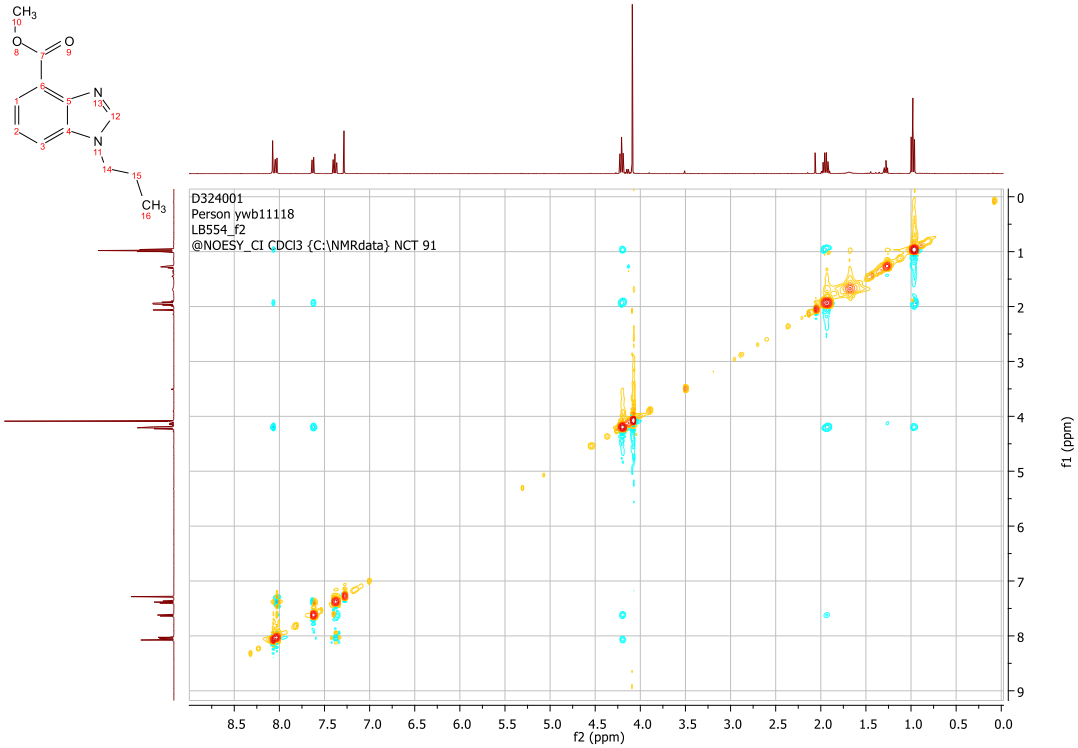


NOESY relationship between ArH (3) and CH_2 (9).

10.5.17 Methyl 1-propyl-1*H*-benzo[*d*]imidazole-7-carboxylate **141**

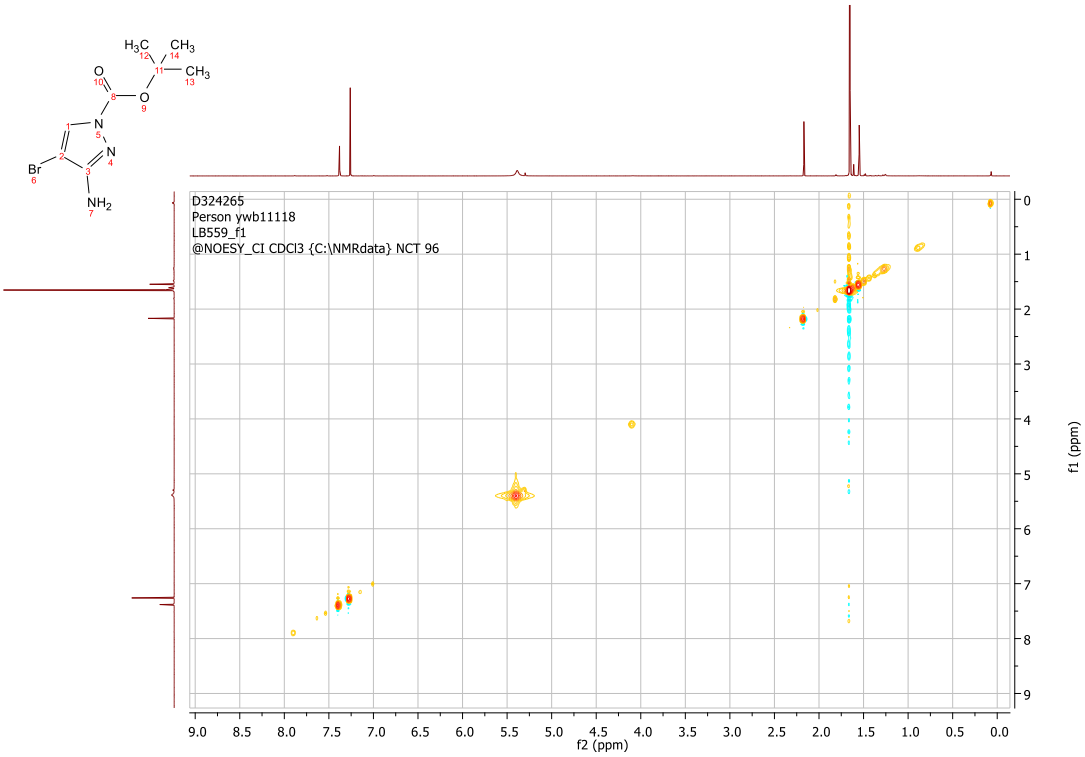


10.5.18 Methyl 1-propyl-1*H*-benzo[*d*]imidazole-4-carboxylate **140**

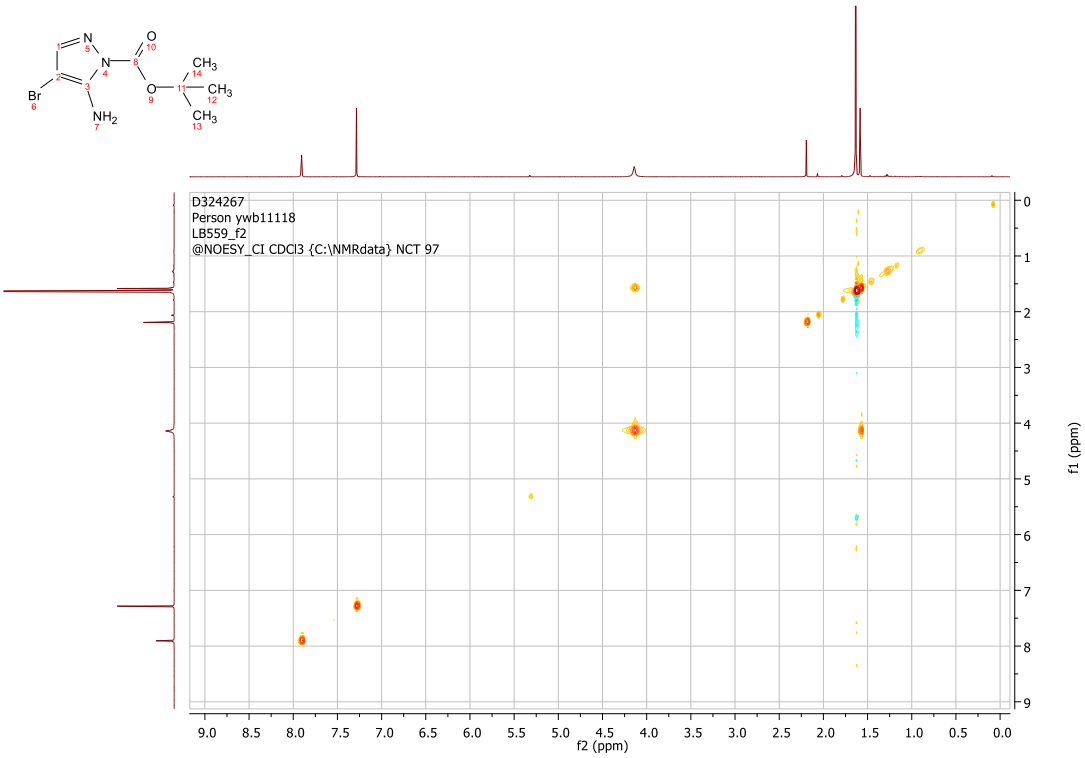


NOESY relationship between ArH (3) and CH_2 (14).

10.5.19 *tert*-Butyl 3-amino-4-bromo-1*H*-pyrazole-1-carboxylate **145**

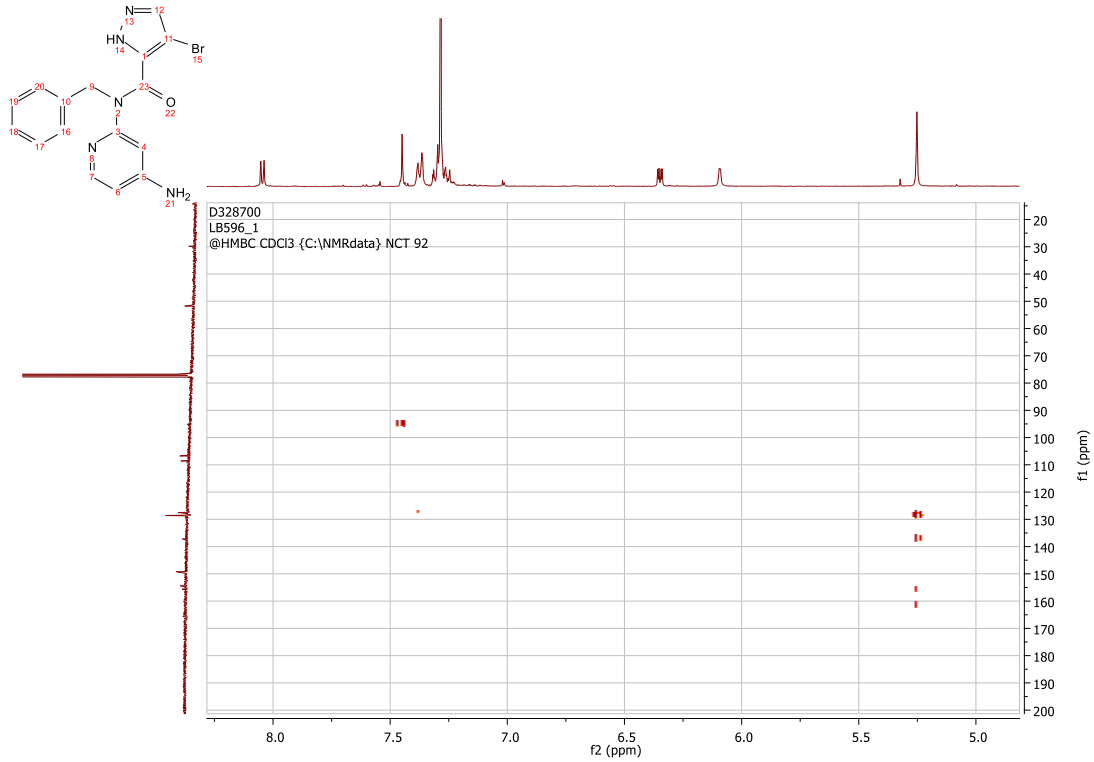


10.5.20 tert-*Butyl* 5-amino-4-bromo-1*H*-pyrazole-1-carboxylate **146**



NOESY relationship between NH_2 (7) and CH_3 (12/13/14).

10.5.21 *N*-(4-aminopyridin-2-yl)-benzyl-4-bromo-1*H*-pyrazole-5-carboxamide **161**



HMBC correlation between CH_2 (9) and C23.

10.6 Experimental procedures

10.6.1 Cell based work carried out by Laureano de la Vega – materials and methods

Cell culture

MDA-MB-231 were obtained from the culture collection at Public Health England. 293T, HeLa, and MDA-MB-468 cell lines were obtained from ATCC. All cell lines were grown in DMEM containing 10% FBS at 37°C and 5% CO2 and routinely tested for mycoplasma. Cells are passaged once 70-90% confluency is reached, and are maintained in culture for no more than 20 passages. Freshly thawed cells are passaged 2-3 times before used.

Quantitative real time PCR (rt-qPCR)

RNA was extracted using RNeasy kit (QIAGEN). 500 ng of RNA per sample was reversetranscribed to cDNA using Omniscript RT kit (QIAGEN) supplemented with RNase inhibitor (QIAGEN) according to the manufacturer's instructions. Resulting cDNA was analysed using TaqMan Universal Master Mix II (Life technologies). Gene expression was determined using an Applied Biosystems 7300 Real-Time PCR system by the comparative ΔΔCT method. All experiments were performed at least in triplicates and data were normalized to the housekeeping gene β-actin. Probes were obtained from Applied Biosystems. When applicable, the differences between groups were determined by 2 way ANOVA. Analyses were performed using GraphPad Prism (GraphPad Software); a P value of <0.05 was considered significant. *P ≤ 0.01, *** P ≤ 0.001.

Lentivirus production

HEK293T cells were transfected using Lipofectamine 2000 (Invitrogen) with the empty vector (pLL3.7) or the lentiviral shDYRK2-1 or shDYRK2-2 together with the packaging vectors (psPAX2 and pMD2.G) and cultivated in OptiMEM medium (Invitrogen). The next day the cells were further grown in DMEM complete medium and one day later the lentivirus-containing supernatant was collected, filtered and used to transduce cells.

Cell transfections

On the day prior to transfection, cells were plated to the required cell density (70-90% confluency). Lipofectamine 2000 and Lipofectamine RNAiMAX (Invitrogen) were used for plasmid DNA and siRNA respectively. The plasmid DNA/siRNA and lipofectamine were individually diluted in Optimem (Gibco) and incubated for 10 minutes at room temperature. Diluted DNA/siRNA was added to the diluted Lipofectamine solution (1:1 ratio) and further incubated for 15 minutes. DNA-lipid complex was added to the cells and incubated overnight in a humidified incubator at 37°C and 5% CO2. The next morning, the medium was replaced with fresh medium and cells were incubated 36 hours more prior lysis.

Cell lysis and western blotting

Cells were washed and harvested in ice-cold phosphate-buffered saline (PBS) and lysed in either SDS buffer or RIPA buffer [50 mM Tris-HCl pH 7.5, 150 mM NaCl, 2 mM EDTA, 1% NP40, 0.5% sodium deoxycholate, 0.5 mM Na3VO4, 50 mM NaF, 2 µg/ml leupeptine, 2 µg/ml aprotinin, 0.05mM pefabloc]. Cells directly lysed in SDS were boiled for two minutes, sonicated and boiled again for another 5 minutes. Cells lysed in RIPA buffer were sonicated and lysates were cleared by centrifugation for 15 minutes at 4°C. Protein concentration established using the BCA assay (Pierce). Supernatant was mixed with SDS sample buffer and boiled for 5 minutes. Equal amounts of protein were separated by SDS-PAGE, followed by semidry blotting to a polyvinylidene difluoride membrane (Thermo Scientific). After blocking of the membrane with 5% (w/v) TBST nonfat dry milk, primary antibodies were added. Appropriate secondary antibodies coupled to horseradish peroxidase were detected by enhanced chemiluminescence using Clarity™ Western ECL Blotting Substrate (BIO-Rad).

Immunoprecipitation

Cells were washed with PBS and lysed in IP buffer (50 mM Hepes pH 7.5, 50 mM NaCl, 1% (v/v) Triton X-100, 2 mM EDTA, 10 mM sodium fluoride, 0.5 mM sodium orthovanadate, leupeptine (10 µg/ml), aprotinin (10 µg/ml), and 1 mM PMSF), followed by a sonication step. The extract was centrifuged and the supernatant was transferred to a new tube. The immunoprecipitation was performed with 2 µg of precipitating antibodies together with 50 µl of Dynabeads™ Protein G. Tubes were rotated for 30 min on a spinning wheel at 4°C. The immunoprecipitates were washed 3x with PBS/0,01% Tween-20 and eluted by boiling in 1 x SDS sample buffer. Equal amounts of protein were separated by SDS-PAGE.

CRISPR-edited cell lines

The endogenous DYRK2 and/or HSF1 genes were knocked-out by transfecting cells with pLentiCRISPR-v2 (which codes for Cas9, and a puromycin cassette) containing gRNAs against the first exon of the short DYRK2 isoform or against the fourth exon of HSF1. For MDA-MB231, HeLa and 293T DYRK2-KO cells the gRNA sequence used was GCTTGCCAGTGGTGCCAGAG and for MDA-MB-468 DYRK2-KO cells the target sequence was CGCTCACGGACAGATCCAGG. Additionally, we also tested some of our results in MDA-MB-231 cells in which we almost completely remove the DYRK2 ORF by using two gRNAs (N-term sequence GCTTGCCAGTGGTGCCAGAG and C-term sequence GAAGCTGAGCTAGAAGGTGG). For HSF1- KO cells the gRNA sequence used was AAGTACTTCAAGCACAACAA. Control cells were transfected with the empty pLentiCRISPRV2 vector. After transfection, cells were exposed to 2 µg/ml of puromycin for two days followed by a medium exchange. Surviving cells were clonally selected (in the case of control cells were used as pool population) by serial dilution, and positive clones were identified by genomic analysis and western blot. At least two clones for each cell line were used for the experiments.

Figure 91: Cells were seeded at lowest concentration. The next day, the cells were treated with either DMSO (-) or the inhibitors as indicated. Samples were taken every day and the numbers of cells were measured using a cell viability protocol in a nucleocounter (based on the PI-exclusion method).

Figure 92: The cells were treated with the inhibitors. 3-4 h later cells were collected and mRNA extracted as indicated in above materials and methods (M&M). Hsp70 expression was calculated as indicated in M&M using GADPH as a housekeeping gene.

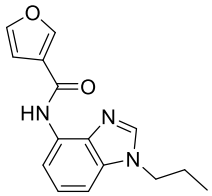
Figure 93: The cells were incubated with either DMSO (-) or the inhibitors. 4 hours later cells were either kept at 37 ºC or exposed to 42C (HS). After 1 h the cells were collected, lysed and the levels of the indicated proteins were measured by western blotting (WB).

Figure 94: 293T cells were transfected with empty plasmid or with DYRK2-GFP as indicated. 36 hours later the cells were treated with the indicated concentrations of LB35 or CI709. 4 hours later cells were collected, lysed and the levels of HSF1 and phospho-HSF1 were measured by WB (see M&M for the WB protocol).

Figure 95: 293T cells were transfected with SIAH2 plasmid together with empty plasmid or with DYRK2-GFP as indicated. 36 hours later cells were treated with the indicated concentrations of LB35 or CI709. 4 hours later cells were collected, lysed and the levels of HSF1 and phospho-HSF1 were measured by WB (see M&M for the WB protocol).

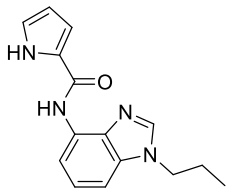
10.6.2 Compounds synthesised by Fiona Keatings

N-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)furan-3-carboxamide **31**



According to **General Procedure D**; furan-3-carboxylic acid (0.34 mmol, 38 mg), HATU (0.29 mmol, 108 mg), amine **98a** (0.29 mmol, 50 mg) and Hünigs base (0.57 mmol, 0.09 mL) in anhydrous MeCN (1 mL) were stirred at 25°C for 48 h. Further purification *via* flash column chromatography in 100% EtOAc afforded *title compound* **31** as a beige solid (0.16 mmol, 44 mg, 57%). **MP** 156–157 ºC; **IR** (ATR, cm^{-1}) 3396, 3097, 2960, 1660, 1524, 1424; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 9.67 (s, 1H, C(O)NH), 8.51 (s, 1H, ArH), 8.25 (s, 1H, ArH), 8.01 (d, J 6.6 Hz, 1H, ArH), 7.76 (d, *J* 7.8 Hz, 1H, ArH), 7.40 (d, *J* 7.8 Hz, 1H, ArH), 7.24 (app t, *J* 7.8 Hz, 1H, ArH), 7.04 (d, *J* 6.6 Hz, 1H, ArH), 4.24 (t, *J* 7.6 Hz, 2H, $NCH_2CH_2CH_3$), 1.86–1.79 (m, 2H, $NCH_2CH_2CH_3$), 0.85 (t, *J* 7.6 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 160.4, 144.3, 143.2, 135.9, 134.3, 129.1, 122.4, 114.0, 109.2, 106.6, 45.8, 22.7, 10.9; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{15}H_{15}N_3O_2$ 269.1 found 270.0 $[M+H]^+$.

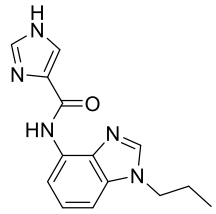
N-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-pyrrole-2-carboxamide **32**



According to **General Procedure D**; 1*H*-pyrrole-2-carboxylic acid (0.34 mmol, 38 mg), HATU (0.29 mmol, 108 mg), amine **98a** (0.29 mmol, 50 mg) and Hünigs base (0.57 mmol, 0.09 mL) in anhydrous MeCN (1 mL) were stirred at 25°C for 48 h. Further purification *via* flash column chromatography in 4:1 EtOAc/petroleum ether 40–60 afforded *title compound* **32** as a beige solid (0.14 mmol, 38 mg, 50%). **MP** 163–164

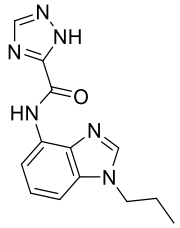
ºC; **IR** (ATR, cm^{-1}) 3255, 2950, 1623, 1554, 1428; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 9.67 (s, 1H, C(O)NH), 8.51 (s, 1H, ArH), 8.25 (s, 1H, ArH), 7.76 (d, *J* 7.8 Hz, 1H, ArH), 7.40 (d, *J* 8.0 Hz, 1H, ArH), 7.24 (app t, *J* 8.0 Hz, 1H, ArH), 7.03 (d, *J* 8.0 Hz, 1H, ArH), 4.23 (t, *J* 7.0 Hz, 2H, $NCH_2CH_2CH_3$), 1.86–1.80 (m, 2H, $NCH_2CH_2CH_3$), 0.84 (t, *J* 7.0 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.8, 143.0, 135.2, 134.2, 129.6, 125.9, 122.7, 112.4, 111.3, 109.2, 105.8, 45.8, 22.7, 10.9; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{15}H_{16}N_4O$ 268.1 found 269.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{15}H_{17}N_4O$ found 269.1390 $[M+H]^+$.

N-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)-1*H*-imidazole-4-carboxamide **33**



According to **General Procedure D**; 1*H*-imidazole-4-carboxylic acid (0.34 mmol, 38 mg), HATU (0.29 mmol, 108 mg), amine **98a** (0.29 mmol, 50 mg) and Hünigs base (0.57 mmol, 0.09 mL) in anhydrous MeCN (1 mL) were stirred at 25°C for 72 h. Further purification *via* flash column chromatography in 96:4 EtOAc/MeOH afforded *title compound* **33** as a beige solid (0.06 mmol, 15 mg, 19%). **MP** 159–161 ºC; **IR** (ATR, cm^{-1}) 3387, 3049, 2924, 1680, 1532, 1495; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 12.71 (s, 1H, NH), 10.05 (s, 1H, C(O)NH), 8.24 (s, 1H, ArH), 8.18 (d, *J* 8.0 Hz, 1H, ArH), 7.31 (d, *J* 8.0 Hz, 1H, ArH), 7.22 (app t, *J* 8.0 Hz, 1H, ArH), 4.22 (t, *J* 7.4 Hz, 2H, $NCH_2CH_2CH_3$), 1.86–1.80 (m, 2H, $NCH_2CH_2CH_3$), 0.85 (t, *J* 7.4 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 160.2, 142.9, 136.0, 133.9, 129.7, 122.9, 120.0, 109.1, 105.3, 46.6, 22.8, 10.9; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{15}N_5O$ 269.1 found 270.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{16}N_5O$ found 270.1344 $[M+H]^+$.

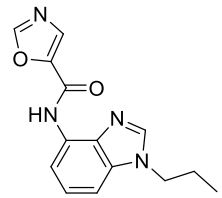
N-(1-propyl-1*H*-benzo[d]imidazol-4-yl)-1*H*-1,2,4-triazole-5-carboxamide **35**



According to **General Procedure D**; 1*H*-1,2,4-triazole-5-carboxylic acid (0.34 mmol, 38 mg), HATU (0.29 mmol, 108 mg), amine **98a** (0.29 mmol, 50 mg) and Hünigs base

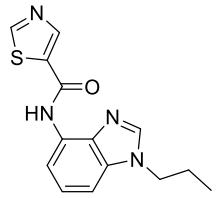
(0.57 mmol, 0.09 mL) in anhydrous MeCN (1 mL) were stirred at 25°C for 48 h. Upon completion, the crude material was triturated in H_2O and collected *via* filtration to provide the *title compound* **35** as a beige solid (0.16 mmol, 42 mg, 55%). **MP** >250 ºC; **IR** (ATR, cm^{-1}) 3331, 3134, 2921, 1673, 1548, 1491; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 14.84 (s, 1H, NH), 10.08 (s, 1H, C(O)NH), 8.77 (s, 1H, ArH), 8.28 (s, 1H, ArH), 8.14 (d, *J* 8.0 Hz, 1H, ArH), 7.40 (d, *J* 8.0 Hz, 1H, ArH), 7.27 (app t, *J* 8.0 Hz, 1H, ArH), 4.23 (t, *J* 6.8 Hz, 2H, $NCH_2CH_2CH_3$), 1.86–1.79 (m, 2H, $NCH_2CH_2CH_3$), 0.84 (t, *J* 6.8 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 143.4, 133.9, 133.9 128.6, 122.9, 122.9, 109.9, 106.4, 45.9, 22.7, 10.9, 2 C missing; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{13}H_{14}N_6O$ 270.1 found 271.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{13}H_{15}N_6O$ found 271.1295 $[M+H]^+$.

N-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)oxazole-5-carboxamide **37**



According to **General Procedure D**; oxazole-5-carboxylic acid (0.43 mmol, 48 mg), HATU (0.29 mmol, 108 mg), amine **98a** (0.29 mmol, 50 mg) and Hünigs base (0.57 mmol, 0.09 mL) in anhydrous MeCN (1 mL) were stirred at 25°C for 72 h. Further purification *via* flash column chromatography in 98:2 EtOAc/MeOH afforded *title compound* **37** as a beige solid (0.12 mmol, 31 mg, 40%).**MP** 164–166 ºC; **IR** (ATR, cm^{-1}) 3084, 2961, 1682, 1589, 1487; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 9.88 (s, 1H, C(O)NH), 8.67 (s, 1H, ArH), 8.28 (s, 1H, ArH), 8.08 (s, 1H, ArH), 7.83 (d, *J* 8.0 Hz, 1H, ArH), 7.45 (d, *J* 8.0 Hz, 1H, ArH), 7.26 (app t, *J* 8.0 Hz, 1H, ArH), 4.24 (t, *J* 7.0 Hz, 2H, $NCH_2CH_2CH_3$), 1.87–1.80 (m, 2H, $NCH_2CH_2CH_3$), 0.85 (t, *J* 7.0 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 154.8, 153.7, 145.1, 143.5, 135.5, 134.3, 130.3, 128.3, 122.6, 113.1, 107.2, 45.9, 22.7, 10.9; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{14}N_4O_2$ 270.1 found 271.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{14}N_4O_2$ 271.1190 found 271.1192 $[M+H]^+$.

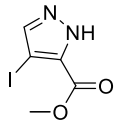
N-(1-propyl-1*H*-benzo[*d*]imidazol-4-yl)thiazole-5-carboxamide **38**



According to **General Procedure D**; thiazole-5-carboxylic acid (0.34 mmol, 44 mg), HATU (0.29 mmol, 108 mg), amine **98a** (0.29 mmol, 50 mg) and Hünigs base (0.57 mmol, 0.09 mL) in anhydrous MeCN (1 mL) were stirred at 25°C for 24 h. After this time, a second portion of thiazole-5-carboxylic acid (0.14 mmol, 18 mg) and HATU (0.10 mmol, 36 mg) was added to the reaction mixture and the resulting mixture was allowed to stir at 25 ºC for 24 h. Further purification *via* flash column chromatography in 97:3 EtOAc/MeOH afforded *title compound* **38** as a beige solid (0.11 mmol, 31 mg, 38%).**MP** 177–178 ºC; **IR** (ATR, cm^{-1}) 3073, 2917, 2848, 1649, 1426; **^{1}H NMR** (400 MHz, DMSO-d_6) δ 10.36 (s, 1H, C(O)NH), 9.31 (s, 1H, ArH), 8.82 (s, 1H, ArH), 8.27 (s, 1H, ArH), 7.66 (d, *J* 8.4 Hz, 1H, ArH), 7.46 (d, *J* 8.4 Hz, 1H, ArH), 7.26 (app t, *J* 8.4 Hz, 1H, ArH), 4.24 (t, *J* 7.1 Hz, 2H, $NCH_2CH_2CH_3$), 1.86–1.70 (m, 2H, $NCH_2CH_2CH_3$), 0.84 (t, *J* 7.1 Hz, 3H, $NCH_2CH_2CH_3$); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.6, 142.9, 136.6, 135.7, 134.5, 128.6, 122.3, 115.2, 107.4, 45.8, 22.7, 10.9; **LRMS** (LCMS-ESI) *m/z* calc. for $C_{14}H_{14}N_4OS$ 286.1 found 287.0 $[M+H]^+$; **HRMS** (ESI, +ve) m/z calc. for $C_{14}H_{14}N_4OS$: pending.

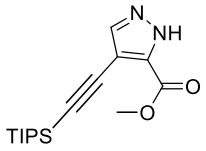
10.6.3 Compounds synthesised by Chris Lawson

Methyl 4-iodo-1*H*-pyrazole-5-carboxylate **108**



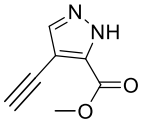
To a solution of 4-iodo-1*H*-pyrazole-5-carboxylic acid **97c** in methanol was added concentrated H_2SO_4 (1 mL) and the resultant solution stirred at reflux for 18 hours. The reaction was then cooled, concentrated under reduced pressure and partitioned between EtOAc and aqueous sat. aqueous $NaHCO_3$. The aqueous layer was extracted with EtOAc and the combined organic layers concentrated under reduced pressure to afford the target compound as a pale yellow solid which was used without further purification.**^{1}H NMR** (500 MHz, DMSO-d_6) δ 13.87 (s, 1H), 8.01 (d, *J* 14.1 Hz, 1H), 3.82 (s, 3H).

Methyl 4-((triisopropylsilyl)ethynyl)-1*H*-pyrazole-5-carboxylate **109**



A microwave vial was charged with ester **108,** CuI (10 mol%) and palladium bistriphenylphosphine dichloride (5 mol%). The vial was then flushed with Ar and to the mixture was added a degassed solution of (DMF/Et_3N; 4:1, 5 mL) and the resulting mixture was heated to 80°C overnight. The reaction was then cooled and poured into ice/water (50 mL) and the aqueous was extracted with EtOAc, dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. Purification by flash column chromatography on silica gel eluting with a gradient of 0-50% EtOAc/Petroleum Ether afforded the target compound in 76% yield as a light yellow solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.91 (s, 1H), 3.98 (s, 3H), 1.17 (s, 25H).

Methyl 4-ethynyl-1*H*-pyrazole-5-carboxylate **110**



To a 20 mL flask charged with a degassed solution of methyl 4-((triisopropylsilyl)ethynyl)-1*H*-pyrazole-5-carboxylate **109** (1.0 eq.) in methanol, AgF (1.5 eq.) was added and the resulting solution stirred in the dark at room temperature overnight. An aqueous 1M HCl solution (3.0 eq) was then added and stirred for a further hour and filtered. The filtrate was then extracted with EtOAc, the organic layer washed with brine, dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. Purification by flash column chromatography on silica gel eluting with a gradient of 0-50% (20% MeOH in EtOAc)/petroleum ether afforded the target compound in 69% yield as a light yellow solid. ^{1}H NMR (400 MHz, DMSO-d_6) δ 13.98 (d, *J* 188.3 Hz, 1H), 8.22 (s, 1H), 4.13 (s, 1H), 3.82 (s, 3H).

11 References

1. https://pubs.acs.org/action/doSearch?AllField=protein+kinases&SeriesKey=jmcmar&startPage=&Earliest=%5B20190912%20TO%2020200912%5D (accessed Sep 2020).

2. https://pubs.rsc.org/en/journals/journalissues/md?_ga=2.148567605.385558139.1599104621499575321.1587722890#!recentarticles&adv (accessed Sep 2020).

3. https://pubs.rsc.org/en/results/journals?Category=journal&AllText=protein%20kinases&IncludeReference=false&SelectJournal=True&ArtRefJournalName=MedChemComm%3B&DateRange=false&SelectDate=false&PriceCode=False&OpenAccess=false&JournalName=MedChemComm (accessed Sep 2020).

4. https://www.sciencedirect.com/search?qs=protein%20kinases&pub=European%20Journal%20of%20Medicinal%20Chemistry&cid=271932 (accessed Sep 2020).

5. Roskoski Jr, R., *Pharma. Res.* **2015,** *100*, 1.

6. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S., *Science* **2002,** *298*, 1912.

7. Chartier, M.; Chenard, T.; Barker, J.; Najmanovich, R., *Peer J.* **2013,** *126*, 1.

8. Hunter, T., *Methods in Enzymology*. Academic Press. Inc.: 1991; Vol. 200.

9. Bradley, D.; Beltrao, P., *PLOS Bio.* **2019,** *17*, e3000341.

10. Cheng, H.; Qu, R. Z.; Paudel, H.; Zhu, H., *Enzyme Res.* **2011,** *2011*, 1.

11. Ardito, F.; Giuliani, M.; Perrone, D.; Troiano, G.; Lo Muzio, L., *Int. J. Mol. Med.* **2017,** *40*, 271.

12. Edelman, A. M.; Blumenthal, D. K.; Krebs, E. G., *Ann. Rev. Biochem.* **1987,** *56*, 567.

13. Capra, M.; Nuciforo, P. G.; Confalonieri, S.; Quarto, M.; Bianchi, M.; Gishizky, M. L.; Draetta, G. F.; Fiore, P. P. D., *Cancer Res.* **2006,** *66*, 8147.

14. Schlessinger, J.; Ullrich, A., *Neuron* **1992,** *9*, 383.

15. Fabbro, D.; Cowan-Jacob, S. W.; Möbitz, H.; Martiny-Baron, G., *Kinase Inhibitors; Methods and Protocols*. SpringerSBM: 2012; Vol. 795.

16. Chico, L. K.; Van Edik, L. J.; Watterson, D. M., *Nature Reviews* **2009,** *8*, 905.

17. Zhang, J.; Yang, P. L.; Gray, N. S., *Nat. Rev. Cancer.* **2009,** *9*, 29.

18. www.researchgate.net/publication/2669056691/fig9/AS:295816536313870@1447539566954/Figure-28-Illustration-of-important-regions-and-interactions-with-ATP-binding-site.png (accessed Jul 2017).

19. Fabbro, D.; Cowan-Jacob, S. W.; Moebitz, H., *Br. J. Pharmacol.* **2015,** *172*, 2675.

20. Wu, P.; Neilson, T. E.; Clausen, M. H., *Trends in Pharmological Sciences* **2015,** *36*, 422.

21. Cancer Statistics for the UK. http://www.cancerresearchuk.org/health-professional/cancer-statistics (accessed Jul 2017).

22. Matthews, D. J.; Gerritsen, M. E., Targeting Protein Kinases for Cancer Therapy. Wiley: Hoboken, 2010; p. 1 online resource (798 p.).

23. Vouri, M.; Hafizi, S., *Cancer Res.* **2017,** *77*.

24. Jain, P.; Karthikeyan, C.; Moorthy, N. S. H. N.; Waiker, D. K.; Jain, A. K.; Trivedi, P., *Curr. Drug Targets,* **2014,** *15*, 1.

25. Chaikaud, A.; Diharce, J.; Schroder, M.; Foucourt, A.; Leblond, B.; Casagrande, A. S.; Desire, L.; Bonnet, P.; Knapp, S.; Besson, T.,. *J. Med. Chem.* **2016,** *59*, 10315

26. Yoshida, S.; Yoshida, K., *FEBS Lett.* **2019,** *593*, 2953.

27. Singh, R.; Lauth, M., *J. Dev. Biol.* **2017,** *5*, 1.

28. Lange, A.; Mills, R. E.; Lange, C. J.; Stewart, M.; Devine, S. E.; Corbett, A. H., *J. Biol. Chem.* **2007,** *282*, 5101.

29. Kinstrie, R.; Luebbering, N.; Miranda-Saavedra, D.; Sibbet, G.; Han, J.; Lochhead, P. A.; Cleghon, V., *Sci. Signal.* **2010,** *3*, ra16.

30. Aranda, S.; Laguna, A.; de la Luna, S., *FASEB* **2011,** *25*, 449.

31. Epstein, C. J., *Nature* **2006,** *441*, 582.

32. Jin, K.; Park, S.; Ewton, D. Z.; Friedman, E., *Cancer Res.* **2007,** *67*, 7247.

33. Pelkmans, L.; Fava, E.; Grabner, H.; Hannus, M.; Habermann, B.; Krausz, E.; Zerial, M., *Nature* **2005,** *436*, 78.

34. Sacher, F.; Moller, C.; Bone, W.; Gottwald, U.; Fritsch, M., *Mol. Cell. Endocrin.* **2007,** *267*, 80.

35. Becker, W., *Cell Cycle* **2012,** *11*, 3389.

36. Soppa, U.; Becker, W., *Curr. Bio.* **2015,** *25*, R488.

37. Yamashita, S. I.; Chujo, M.; Tokuishi, K.; Anami, K.; Miyawaki, M.; Yamamoto, S.; Kawahara, M., *J. Thoricacs Cardiovas. Sur.,* **2009***, 138*, 1303.

38. Taira, N.; Nihira, K.; Yamaguchi, T.; Miki, Y.; Yoshida, K., *Mol. Cell* **2007,** *25*, 725.

39. Taira, N.; Mimoto, R.; Kurata, M.; Yamaguchi, T.; Kitagawa, M.; Miki, Y.; Yoshida, K., *J. Clin. Invest.* **2012,** *122*, 859.

40. Morrugares, R.; Correa-Saez, A.; Moreno, R.; Garrido-Rodriguez, M.; Munoz, E.; de la Vega, L.; Calzado, M. A., *Cell. Mol. Life Sci.* **2019,** *1*, 1.

41. Miller, C. T.; Aggarwal, S.; Lin, T. K.; Dagenais, S. L.; Contreras, J. I.; Orringer, M. B.; Glover, T. W.; Beer, D. G.; L., L., *Cancer Res.* **2003,** *63*, 4136.

42. Yamashita, S. I.; Chujo, M.; Moroga, T.; Anami, K.; Tokuishi, K.; Miyawaka, M.; Kawano, Y.; Takeno, S.; Yamamoto, S.; Kawahara, K., *Anticancer Res.* **2009,** *29*, 2753.

43. Mimoto, R.; Nihari, N. T.; Hirooka, S.; Takeyama, H.; Yoshida, K.,. *Cancer Lett.* **2017,** *384*, 27.

44. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer (accessed Feb 2020).

45. Makki, J., *Clin. Med. Insig. Path.* **2015,** *8*, 23.

46. O'Sullivan, C. C.; Loprinzi, C. L.; Haddad, T. C., *Mayo Clin. Proc.* **2018,** *93*, 794.

47. Togacar, M.; Ergen, B.; Coewert, Z., *Medical Hypothesis* **2020,** *135*, 109503.

48. Du, T.; Zhu, L.; Levine, K. M.; Tasdemir, N.; Lee, A. V.; Vignali, D. A. A.; Van Houten, B.; Tseng, G. C.; Oesterreich, S., *Scientific Reports* **2018,** *8*, 7205.

49. https://www.breastcancer.org/symptoms/types/ilc (accessed Feb 2020).

50. Rugo, H. S.; Rumble, B.; Macrae, E.; Barton, D. L.; Connolly, H. K.; Dickler, M. N.; Fallowfield, L.; Fowble, B.; Ingle, J. N.; Jahanzeb, M.; Johnson, S. R. D.; Korde, L. A.; KHatcheressian, J. L.; Mehta, R. S.; Muss, H. B.; Burstein, H. J., *J. Clin. Oncol.* **2016,** *2016*, 1.

51. Goldhirsch, A.; Wood, W. C.; Coates, A. S.; Gelber, R. D.; Thurlimann, B.; -J., S. H., *Ann. Oncol.* **2011,** *22*, 1736.

52. Kumar, P.; Aggarwal, R., *Arch. Gynecol. Obstet.* **2016,** *293*, 247.

53. Cronin, K. A.; Harlan, L. C.; Dodd, K. W.; Abrams, J. S.; Ballard-Barbash, R., *Cancer Invest.* **2010,** *28* 963.

54. Prat, A.; Pineda, E.; Adamo, B.; Galvan, P.; Fernandez, A.; Gaba, L.; Diez, M.; Viladot, M.; Arance, A.; Munoz, M., *The Breast* **2015,** *2015*, 1.

55. Sorlie, T.; Perou, C. M.; Tibshirani, R.; Aas, T.; Geisler, S.; Johnson, H.; Hastie, T.; Eisen, M. B.; can de Rijn, M.; Jeffrey, S. S.; Thorsen, T.; Quise, H.; Matese, J. C.; Brown, P. O.; Botstein, D.; Lonning, P. E.; Borresen-Dale, A., *PNAS* **2001,** *98*, 10869.

56. Godoy-Ortiz, A.; Sanchez-Munoz, A.; Rosario Chica Parrado, M.; Alvarez, M.; Ribelles, N.; Rueda Dominguez, A.; Alba, E., *Front. Oncol.* **2019,** *9*, 1.

57. Aubrey, B. J.; Strasser, A.; Kelly, G. L., *Cold Spring Harb. Perspect. Med.* **2016,** *6*.

58. Yang, J.; Ma, X.; Wei, Y.; Peng, Y.; Wei, X., *Mol. Cancer* **2019,** *18*, 1.

59. Tremont, A.; Lu, J.; Cole, J. T., *Ochsner J.* **2017,** *17*, 405.

60. Yang, S. X.; Polley, E. C., *Breast Can. Res. Treat.* **2019,** *175*, 287.

61. Schneider, R.; Barakat, A.; Pippen, J.; Osborne, C., *Breast Can. Tar. Ther.* **2011,** *3*, 113.

62. Baselga, J., *The Oncologist* **2002,** *7*, 2.

63. Yang, S. X.; Costantino, J. P.; Kim, C.; Mamounas, E. P.; Nguyen, D.; Joeng, J.; Wolmark, N.; Kidwell, K.; Paik, S.; M., S. S., *J. Clin. Oncol.* **2010,** *28*, 2974.

64. Denkert, C.; Liedtke, C.; Tutt, A.; von Minckwitz, G., *Lancet* **2017,** *389*, 2430.

65. Wang, C.; Kar, S.; Lai, X.; Cai, W.; Arfuso, F.; Sethi, G.; Lobie, P. E.; Goh, B. C.; Lim, L. H. K.; Hartman, M.; Chan, C. W.; Lee, S. C.; Tan, S. H.; Kumar, A. P., *Cancer Treat. Rev.* **2018,** *62*, 29.

66. Tutt, A.; Tovey, H.; Cheang, M. C. U.; Kernaghan, S.; Kilburn, L.; Gazinska, P.; Owen, J.; Abraham, J.; Barrett, S.; Barrett-Lee, P.; Brown, R.; Chan, S.; Dowsett, M.; Flanagan, J. M.; Fox, L.; Grigoriadis, A.; Gutin, A.; Harper-Wynne, C.; Hatton, M. Q.; Hoadley, K. A.; Parikh, J.; Parker, P.; Perou, C. M.; Roylance, R.; Shah, V.; Shaw, A.; Smith, I. E.; Timms, K. M.; Wardley, A. M.; Wilson, G.; Gillett, C.; Lanchbury, J. S.; Ashworth, A.; Rahman, N.; Harries, M.; Ellis, P.; Pinder, S. E.; Bliss, J. M., *Nat. Med.* **2018,** *24*, 628.

67. Afghani, A.; Telli, M. L.; Kurian, A. W., *Curr. Probl. Cancer* 2016, *40*, 130.

68. Bianchini, G.; Balki, J. M.; Mayer, I. A.; Sanders, M. E.; Gianni, L., *Nat. Rev. Clin. Oncol.* **2016,** *13*, 674.

69. Pandy, J. G.; Balolong-Garcia, J. C.; Cruz0Ordinario, M. V. B.; Que, F. V. F., *BMC Cancer* **2019,** *2019*, 1.

70. Geenen, J. J. J.; Linn, S. C.; Beijnen, J. H.; Schellens, J. H. M., *Clin Pharmacokinet.* **2018,** *57*, 427.

71. Robson, M.; Im, S.; Senkus, E.; Xu, B.; Domchek, S. M.; Masuda, N.; Delaloge, S.; Li, W.; Tung, N.; Armstrong, A.; Wu, W.; Goessi, C.; Runswick, S.; Conte, P., *New Eng. J. Med.* **2017,** *377*, 523.

72. McCann, K. E.; Hurvitz, S. A., *Drugs in Context* **2018,** *7*, 1.

73. Ding, Y. C.; Steele, L.; Warden, C.; Wilczynski, S.; Mortimer, J.; Yuan, Y.; Neuhausen, S. L., *Oncotarget* **2019,** *10*, 198.

74. Rampurwala, M.; Wisinski, K. B.; O'Regan, R., *Clin. Adv. Hematol. Oncol.* **2016,** *14*, 186.

75. Beer, T. M.; Armstrong, A. J.; Rathkopf, D. E.; Loriot, Y.; Sternberg, C. N.; Higano, C. S.; Iverson, P.; Bhattacharya, S.; Carles, J.; Chowdhury, S.; Davis, I. D.; de Bono, J. S.; Evans, C. P.; Fizazi, K.; Joshua, A. M.; Kim, C.-S.; Kimura, G.; Mainwaring, P.; Mansbach, H.; Miller, K.; Noonberg, S. B.; Perabo, F.; Phung, D.; Saad, F.; Scher, H. I.; Taplin, M.-E.; Venner, P. M.; Tombal, B., *New Eng. J. Med.* **2014,** *371*, 424.

76. Khan, M. A.; Jain, V. K.; Rizwanullah, M.; Ahmad, J.; Jain, K., *Drug Disc. Tod.* **2019,** *24*, 2181.

77. Chan, J. J.; Tan, T. J. Y.; Dent, R. A., *Ther. Adv. Med. Oncol.* **2019,** *11*, 1.

78. Li, B.; Chonghaile, T. N.; Fan, Y.; Madden, S. F.; Klinger, R.; O'Connor, A. E.; Walsh, L.; O'Hurley, G.; Udupi, G. M.; Joseph, J.; Tarrant, F.; Conroy, E.; Gaber, A.; Chin, S.-F.; Bardwell, H. A.; Provenzano, E.; Crown, J.; Dubois, T.; Linn, S.; Jirstrom, K.; Caldas, C.; O'Connor, D. P.; Gallagher, W. M., *Cancer Res.* **2017,** *77*, 3834.

79. Guo, X.; Wang, X.; Wang, Z.; Banerjee, S.; Yang, J.; Huang, L.; Dixon, J. E., *Nat. Cell Biol.* **2016,** *18*, 202.

80. Voges, D.; Zwickl, P.; Baumeister, W., *Annu. Rev. Biochem.* **1999,** *68*, 1015.

81. Gentier, R. J.; van Leeuwen, F. W., *Front. Mol. NeuroSci.* **2015,** *8*, 47.

82. Jentsch, S., *Ann. Rev. Genet.* **1992,** *26*, 179.

83. Lecker, S. H.; Goldberg, A. L.; Mitch, W. E., *J. Am. Soc. Nephrol.* **2006,** *17*, 1807.

84. Coux, O.; Tanaka, K.; Goldberg, A. L., *Annu. Rev. Biochem.* **1996,** *65*, 801.

85. Bard, J. A. M.; Goodall, E. A.; Greene, E. R.; Jonsson, E.; Dong, K. C.; Martin, A., *Annu. Rev. Biochem.* **2018,** *87*, 697.

86. de Poot, S. A. H.; Tian, G.; Finley, D., *J. Mol. Biol.* **2018,** *429*, 3525.

87. Dou, Q. P.; Zonder, J. A., *Curr. Cancer Drug Targets* **2014,** *14*, 517.

88. Almond, J. B.; Cohen, G. M., *Leukemia* **2002,** *16*, 433.

89. Petroccca, F.; Altschuler, G.; Tan, S. M.; Mendillo, M. L.; Yan, H.; Jerry, D. J.; Kung, A. L.; Hide, W.; Ince, T. A.; Lieberman, J., **2013,** *24*, 182.

90. Banerjee, S.; Ji, C.; Mayfield, J. E.; Goel, A.; Xiao, J.; Dixon, J. E.; Gio, X., *PNAS* **2018,** *115*, 8155.

91. Banerjee, S.; Wei, T.; Wang, J.; Lee, J. J.; Gutierrez, H. L.; Chapman, O.; Wiley, S. E.; Mayfield, J. E.; Tandon, V.; Juarez, E. F.; Chaves, L.; Liang, R.; Sah, R. L.; Costello, C.; Mesirov, J. P.; de la Vega, L.; Cooper, K. L.; Dixon, J. E.; Xiao, J.; Lei, X., *PNAS* **2019,** *116*, 24881.

92. Zaal, E. A.; Wu, W.; Jensen, G.; Zweegman, S.; Cloos, J.; Berkers, C. R., *Can. Metab.* **2017,** *5*, 1.

93. Flick, K.; Kaiser, P., *Semin. Cell Dev. Biol.* **2012,** *23*, 515.

94. Mazaira, G. I.; Daneri-Becerra, C.; Zgajinar, N. R.; Lutufo, C. M.; Galigniana, M. D., *Biochem. Soc. Trans.* **2018,** *46*, 51.

95. Naidu, S. D.; Dinkova-Kostova, A. T., *FEBS J.* **2017,** *284*, 1606.

96. Huang, C.; Wu, J.; Xu, L.; Wang, J.; Chen, Z.; Yang, R., *Eur. J. Pharmacol.* **2018,** *822*, 69.

97. Della Salda, L.; Romanucci, M., *J. Can. Ther.* **2012,** *3*, 755.

98. Wilkinson, K. A.; Henley, J. M., *Biochem. J.* **2012,** *428*, 133.

99. Sharma, C.; Seo, Y. H., *Molecules* **2018,** *23*, 1.

100. Neudegger, T.; Verghese, J.; Hayer-Hartl, M.; Hartl, F. U.; Bracher, A., *Nat. Struc. Mol. Bio.* **2016,** *23*, 1.

101. Naidu, S. D.; Sutherland, C.; Zhang, Y.; Risco, A.; de la Vega, L.; Caunt, C. J.; Hastie, C. J.; Lamont, D. J.; Torrente, L.; Chowdhry, S.; Benjamin, I. J.; Keyse, S. M.; Cuenda, A.; Dinkova-Kostova, A. T., *Mol. Cell Bio.* **2016,** *36*, 2403.

102. Murshid, A.; Chou, S.; Zhang, Y.; Bjarti, A.; Calderwood, S. K., *PLOS One* **2010,** *5*, e13830.

103. Calderwood, S. K., *Curr. Mol. Med.* **2012,** *12*, 1102.

104. Yallowitz, A.; Ghaleb, A.; Garcia, L.; Alexandrova, E. M.; Marchenko, N., *Cell Death and Disease* **2018,** *9*, 1.

105. Cicooca, D. R.; Calderwood, S. K., *Cell Stress Chap.* **2005,** *10*, 86.

106. Ma, Y. X.; Fan, S.; Xiong, J.; Yuan, R.; Meng, Q.; Gao, M.; Goldberg, I. D.; Fuqua, S. A.; Pestell, R. G.; Rosen, E. M., *Oncogene* **2003,** *22*, 10.

107. Hwang, S.-Y.; Park, S.; Kwon, Y., *Pharmacol. Therap.* **2019,** *199*, 30.

108. Carpenter, R. L.; Sirkinsoon, S.; Zhu, D.; Rimkus, T.; Harrison, A.; Anderson, A.; Paw, I.; Qasem, S.; Zing, F.; Liu, Y.; Chan, M.; Metheny-Barlow, L.; Pasche, B. C.; Debinski, W.; Watabe, K.; Lo, H., *Oncotarget* **2017,** *8*, 1.

109. Carpenter, R. L.; Goekmen-Polar, Y., *Curr. Can. Drug Tar.* **2019,** *19*, 515.

110. Rashmi, K. C.; Atreya, H. S.; Raj, M. H.; Salimath, B. P.; Aparna, H. S., *Cell Stress Chap.* **2017,** *22*, 751.

111. Bishop, A. C.; Ubersax, J. A.; Petsch, D. T.; Matheos, D. P.; Grayk, N. S.; Blethrow, J.; Shimizu§, E.; Tsien, J. Z.; Schultzk, P. G.; Rose, M. D.; Wood, J. L.; Morgan, D. O.; Shokat, K. M., *Nature* **2000,** *407*, 395.

112. Tahtouh, T.; Elkins, J. M.; Filippakopoulos, P.; Soundararajan, M.; Burgy, G.; Durieu, E.; Cochet, C.; Schmid, R. S.; Lo, D. C.; Delhommel, F.; Oberholzer, A. E.; Pearl, L. H.; Carreaux, F.; Bazureau, J.; Knapp, S.; Mejier, L., *J. Med. Chem.* **2012,** *55*, 9312.

113. Schmitt, C.; Kail, D.; Mariano, M.; Empting, M.; Weber, N.; Paul, T.; Hartmann, R. W.; Engel, M., *PLOS One* **2014,** *9*, e87851.

114. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., *Adv. Drug Del. Rev.* **2001,** *46*, 3.

115. Tomkinson, N. C. O., PhD Thesis from Camille Roxanne Chevalier Indey. University of Strathclyde, P. a. A. C., Glasgow, UK., Ed.

116. Neisen, F. H.; Berglund, H.; Vedadi, M., *Nat. Protocol.* **2007,** *2*, 2212.

117. Marques, S. M.; Esteves da Silva, J. C. G., *Life* **2009,** *61*, 6.

118. Mackay, S. P. Masters Thesis of Gillian Berrie. University of Strathclyde, 2018.

119. Knight, Z. A.; Shokat, K. M., *Chem. Biol.* **2005,** *12*, 621.

120. http://www.mch.estranky.sk/file/86/130813_atp-km_ver_02_2013.pdf (accessed Mar 2020).

121. Alexeeva, M.; Aberg, E.; Engh, R. A.; Rothweiler, U., *Acta Cryst.* **2015,** *D71*, 1207.

122. Cook, P. F.; Neville, M. E.; Vrana, K. E.; Hartl, T.; Roskoski Jr, R., *ACS. Biochem.* **1982,** *21*, 5794.

123. Ogawa, Y.; Nonaka, Y.; Goto, T.; Ohnishi, E.; Hiramatsu, T.; Yoshida, M.; Ikura, T.; Onogi, H.; Shibuya, H.; Hosoya, T.; Ito, N.; Hagiwara, M., *Nat. Commun.* **2010,** *86*, 1.

124. Sirimulla, S.; Bailey, J. B.; Vegesna, R.; Narayan, M., *J. Chem. Inf. Model.* **2013,** *53*, 2781.

125. Wang, H.; Wang, W.; Jun Jin, W., *Chem. Rev.* **2016,** *116*, 5072.

126. Patani, G. A.; LaVoie, E. J., *Chem. Rev.* **1996,** *96*, 3147.

127. Wilcken, R.; Zimmermann, M. O.; Bauer, M. R.; Rutherford, T. J.; Fersht, A. R.; Joerger, A. C.; Boeckler, F. M., *ACS Chem. Biol.* **2015,** *10*, 2725.

128. Krawczyk, m. P.; Kowalski, D. M.; Ramlau, R.; Kalinka-Warzocha, E.; Winiarczyk, K.; Stencel, K.; Powrozek, T.; Reszka, K.; Wojas-Krawczyk, K.; Bryl, M.; Wojcik-Superczynska, M.; Glogowski, M.; Barinow-Wojewodzki, A.; Milanowski, J.; Krzakowski, M., *Oncol. Lett.* **2017,** *13*, 4433.

129. Bissantz, C.; Kuhn, B.; Stahl, M., *J . Med. Chem.* **2010,** *53*, 5061.

130. Zhao, Z.; Liu, Q.; Bliven, S.; Xie, L.; Bourne, P. E., . *J. Med. Chem.* **2017,** *60*, 2879.

131. Kumar, R.; Sharma, A.; Kumar Tiwari, R., *Introduction to Drug Designing and Development*. Nova Science Publishers Inc.: New York, 2014.

132. https://www.medicosite.com/pharmacokinetics-definition/ (accessed Sep 2020).

133. Uetrecht, J.; Trager, W., *Drug Metabolism: Chemical and Enzymatic Aspects*. CRC Press LLC: Florida, 2007.

134. Mullard, A., *Nat. Rev. Drug Dis.* **2018,** *17*, 777.

135. Caron, G.; Ermondi, G., *Fut. Med. Chem.* **2016,** *8*, 2013.

136. Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D., *J. Med. Chem.* **2002,** *45*, 2615.

137. Hughes, J. D.; Blagg, J.; Price, D. A.; Bailey, S.; A., D. G.; Devraj, R. V.; Ellsworth, E.; Fobian, Y. M.; Gibbs, M. E.; Gilles, R. W.; Greene, N.; Huang, E.; Krieger-Burke, T.; Loesel, J.; Wager, T.; Whiteley, L.; Zhang, Y., *Bioorg. Med. Chem. Lett.* **2008,** *18*, 4872.

138. Savjani, K. T.; Gajjar, A. K.; Savjani, J. K., *ISRN Pharmaceutics* **2012,** *2012*.

139. Zhong, H.; Chan, G.; Hu, Y.; H., H.; Ouyang, D., *Pharmaceutics* **2018,** *10*, 1.

140. Kerns, E. H.; Di, L.; Carter, G. T., *Curr. Drug Met.* **2008,** *9*, 879.

141. Lipinksi, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., *Adv. Drug Del. Rev.* **2001,** *46*, 3.

142. van Breemen, R.; Li, Y., *Expert Opin. Durg Metab. Toxicol.* **2005,** *1*, 175.

143. Richardson, S. J.; Bai, A.; Kulkarni, A. A.; Moghaddam, M. F., *Drug Met. Lett.* **2016,** *10*, 83.

144. Perez, M.; Garcia-Limones, C.; Zapico, I.; Marina, A.; Schmitz, M. L.; Munoz, E.; Calzado, M. A., *J. Mol. Cell Bio.* **2012,** *4*, 316.

145. Knauer, S. K.; Mahendrarajah, N.; Roos, W. P.; Kraemer, O. H., *Cyto. Growth Fac. Rev.* **2015,** *26*, 405.

146. Confalonieri, S.; Quarto, M.; Goisis, G.; Nuciforo, P.; Donzelli, M.; Jodice, G.; Pelosi, G.; Viale, G.; Pece, S.; Di Fiore, P., *Oncogene* **2009,** *28*, 2959.

147. https://www.sigmaaldrich.com/catalog/product/aldrich/n21308?lang=en&region=GB (accessed April 2020).

148. https://www.sigmaaldrich.com/catalog/substance/ophenylenediamine108149554511?lang=en&region=GB (accessed April 2020).

149. Fatin-Rouge, N.; Toth, E.; Perret, D.; Backer, R. H.; Merbach, A. E.; Buenzli, J. G., *J. Am. Chem. Soc.* **2000,** *122*, 10810.

150. Anderson, D. R.; Volkmann, R. A.; Mennite, F. S.; Fanger, C. 2017.

151. Zhao, Z.; Wang, Z., *Synth. Commun.* **2007,** *37*, 137.

152. Lyalin, B. V.; Petrosyan, V. A., *Russ. Chem. Bull.* **2013,** *62*, 1044.

153. Skinner, P. J.; Cherrier, M. C.; Webb, P. J.; Shin, Y.-J.; Gharbaoui, T.; Lindstram, A.; Hong, V.; Tamura, S. Y.; Dang, H. T.; Pride, C. C.; Chen, R.; Richman, J. G.; Connolly, D. T.; Semple, G.. *Bioorg. Med. Chem. Lett.* **2007,** *17*, 5620.

154. Tomkinson, N. C. O. Masters Thesis of Fiona Keatings. University of Strathclyde, 2016.

155. Gadakh, B.; Vondenhoff, G.; Lescrinier, E.; Rozenski, J.; Froeyen, M.; Aerschot, A. V., *Bioorg. Med. Chem.* **2014,** *22*, 2875.

156. Greene, T. W.; Wuts, P. G. M., *Protective Groups in Organic Chemistry*. Third ed.; John Wiley & Sons, Inc.: New York, 1999; p 779.

157. Dey, R.; Mukherjee, N.; Ahammed, S.; Ranu, C. B., *Chem. Comm.*: **2012**; 48, 7982.

158. Erion, M. D.; Dang, Q.; Reddy, M. R.; Kasibhatla, S. R.; Huang, J.; Lipscomb, W. N.; van Poelje, P. D., *J. Am. Chem. Soc.* **2007,** *129*, 15480.

159. Schoenberger, M.; Trauner, D., *Angew. Chem. Int. Ed.* **2014,** *53*, 3264.

160. Zhou, D.; Zhou, P.; Evrard, D. A.; Meagher, K.; Webb, M.; Harrison, B. L.; Huryn, D. M.; Golembieski, J.; Hornby, G. A.; Schechter, L. E.; Smith, D. L.; Andree, T. H.; Mewshaw, R. E., **2008**.

161. Hanan, E. J.; Chan, B. K.; Estrada, A. A.; Shore, D. G.; Lyssikatos, J. P., *Synlett* **2010,** *18*, 2759.

162. Janin, Y. L., *J. Het. Chem.* **2013,** *50*, 1410.

163. Tanakit, A.; Rouffet, M.; Martin, D. P.; Cohen, S. M.,. *Dalton Trans.* **2012,** *41*, 6507.

164. Elvidge, J. A.; Newbold, G. T.; Percival, A.; Sencia, I. R., *J. Chem. Soc.* **1965,** *942*, 5119.

165. Rodrıguez-Franco, M. I.; Dorronsoro, I.; Hernandez-Higueras, A. I.; Antequera, G., *Tet. Lett.* **2001**, 863.

166. Lyalin, B. V.; Petrosyan, V. A.; Ugrak, B. I., *Russ. Chem. Bull. Int. Ed.* **2010,** *59*, 1549.

167. Jones, R. G.; Mann, M. J., *J. Am. Chem. Soc.* **1953,** *75*, 4048.

168. Jones, R. B.; Mann, M. J., *J. Am. Chem. Soc.* **1953,** *75*, 4049.

169. Wang, Z.-X.; Qin, H.-L., *Green Chem.* **2004,** *6*, 90.

170. Holzer, W.; Seiringer, G., *J. Heterocyclic Chem.* **1993,** *30*, 865.

171. Huang, Q.; Tran, G.; Pardo, D. G.; Tsuchiya, T.; Hillebrand, S.; Vors, J.; Cossy, J., *Tetrahedron* **2015,** *71*, 7250.

172. Van der Plas, S. E.; Kelgtermans, H.; De Munck, T.; Martina, S. L. X.; Dropsit, S.; Quinton, E.; De Blieck, A.; Joannesse, C.; Tomaskovic, L.; Jans, M.; Christophe, T.; van der Aar, E.; Borgonovi, M.; Nelles, L.; Gees, M.; Stouten, P.; Van Der Schueren, J.; Mammoliti, O.; Conrath, K.; Andrews, M., *J. Med. Chem.* **2018,** *61*, 1425.

173. Dalinger, I. L.; Vatsadse, I. A.; Shkineva, T. K.; Popova, G. P.; Ugrak, B. I.; Shevelev, S. A., *Russ. Chem. Bull. Int. Ed.* **2010,** *59*, 1631.

174. Wright, S. W.; McClure, L. D., *J. Het. Chem.* **2004,** *41*, 1023.

175. Soerensen, U. S.; Grunnet, M.; Bentzem, B. H.; Christophersen, P.; Diness, J. G.; Skibsbye, L.; Stroebaek, D., WO 2013104577, 2013.

176. Napoletano, M.; Trevisiani, M.; Pavani, M. G.; Fruttarolo, F., EP 2377850, 2011.

177. Gur, Z. T.; Caliskan, B.; Garscha, U.; Olgac, A.; Schubert, U. S.; Gerstmeier, J.; Werz, O.; Banoglu, E., *Eur. J. Med. Chem.* **2018,** *150*, 876.

178. Williams, A.; Salvadori, G., J. Chem. Soc. Perk. Trans. 2. **1972**, 883.

179. Duncan, K. W.; Chesworth, R.; Munchhof, M. J.; Jin, L., WO2014100695, 2014.

180. Montanari, F.; Passerini, R., *Bollettino Scientifico della Facolta di Chimica Industriale di Bologna* **1953,** *11*, 42.

181. Cheng, H. S., *J. Pharmacol. Toxicol. Met.* **2002,** *46*, 61.