The Synthesis and Evaluation of WWP2 E3 Ubiquitin Ligase Inhibitors as Anti-Cancer Leads

Jake M. Rigby

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

at

University of East Anglia, Norwich

School of Chemistry

March 2024

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university.

<div align="right">

Jake M. Rigby

March 2024

</div>

Acknowledgements

Firstly I would like to give a big thanks to Dr Tom Storr, Professor Andrew Hemmings, Dr Chris Richards, Professor Richard Stephenson, and Dr Andrew Chandry who all at some point throughout the three years to varying extents supervised and guided the me during this PhD. I thank (Dr to be) Ashley Dudey from the school of Biology, whose tireless efforts to test the compounds synthesised, made this project very productive and in the end fruitful. A special mention for (Dr to be) Tyarna Smith, thank you for this past year together and for your support in and out of the lab over the past two years. I would also like to thank Dr Jacob Gretton for his help and guidance during the initial lab tasks during my MSc at UEA, and (Dr to be) Joshua Hall, for his valuable help in the initial months when I started the PhD project. I would like to thank the school of Chemistry technical and goods teams, for their invaluable support of the research efforts within the school.

I would also like to thank the members of the office on floor 3 who have made the time here just about bearable, both past and present. Thankyou Dr Jack Andrews, (Dr to be) Tyarna Smith, (Dr to be) Dominic Dennis, Ford Fagen, Dr Gareth Hughes, Dr James Sharpe, Dr Adam Wright, and (Dr to be) Rishabh Kumar, for the interesting conversations, arguments, and chats which filled large portions of our break (and sometimes work) time. I would also like to thank the members of the office mentioned above for their help, advice, and enthusiasm to try and answer any problem encountered in the course of each-others research.

I would like to also thank Dr Benjamin Hofmann, Dr Claire Jones, and Dr Giovanni Bressan, and again (Dr to be) Tyarna Smith who have been great friends for the sporadic (and frequent) weekday evening trips to Scholar's bar and about in Norwich, as well as Friday CAP drinks. Giovanni and Ben's 'fresh air' breaks have also been particularly helpful. These outside breaks, evening trips across the campus and about Norwich were very beneficial and their value should not be underestimated.

My family: John and Issac Rigby, and Helen and Robert Bell, I would like to thank for their support over the last 7 and a half years of education. I would also like to thank my Maths and Chemistry A-level teachers, Mrs Jools Burlace and Mrs Beryl Mernagh, whose teaching helped me decide to do Chemistry at university, and perhaps without knowing, taught me very valuable lessons beyond these subjects at school.

Abstract

The ubiquitin-proteasome system (UPS) is the major system involved in the regulation of cellular protein quality and quantity. The UPS does this in a highly selective manner by tagging redundant cellular proteins with ubiquitin, which leads to their degradation by the 26S proteasome. Within the ubiquitin-proteasome system, the WW-domain containing HECT E3 ubiquitin ligase 2 (WWP2) is responsible for the ubiquitin-mediated degradation of key tumour suppressor proteins. WWP2s dysregulation is present in several diseases, most notably in cancer. WWP2 therefore represents a therapeutic target for small-molecule inhibition for cancer therapy.

This doctoral thesis focusses on the generation of small molecule inhibitors of WWP2 as anticancer leads, with initial starting compounds identified from high-throughput screenings. This thesis describes the synthetic work undertaken to produce WWP2 inhibitors *via* a rational drug design approach. The production of focussed libraries surrounding the identified hits was aided by molecular docking studies and, in two cases, co-crystal structures. In the course of this work, the reproduction of 9 hit molecules or molecules of interest for crystal soaking studies and analogue synthesis of 4 different lead compounds was pursued. Approximately 120 derivative compounds have been investigated as part of this aim.

The results of this investigation show little improvement in inhibitory data against WWP2 in any of the hit compounds pursued into analogues synthesis, which points towards some issues surrounding strategy with respect to hit validation. Issues surrounding hit IC_{50} value reproducibility identified impure NCI samples. The identification of PAINs and nuisance compounds as hits investigated also hampered the production of inhibitors of WWP2. The issues identified in the processes prior to hit-to-lead optimisation in the course of this work should be addressed to overcome the investigation of unproductive leads. A more solid experimental foundation with which to choose analogues for synthesis is also desirable. Additionally, several analogue series are currently being investigated as promising WWP1 inhibitors, a close relative of WWP2 also present as a therapeutic target against cancer.

Content

List of Abbreviations and Acronyms

UPS	Ubiquitin-proteasome system
POI	Protein of interest
PTM	Post-translational modification
Ub	Ubiquitin
DUBs	De-ubiquitinating enzymes
UBLs	Ubiquitin-like proteins
PPIs	Protein-Protein interactions
UFD	Ubiquitin-Fold domain
UBC	Ubiquitin-conjugating domain
RBR	RING-between-RING
HECT	Homologous to E6-AP carboxyl terminus
RING	Really Interesting New Gene
IBR	In between RING region
RCC1	Regulator of chromosome condensation 1
RLD	RCC1-like domain
PI3K	Phosphatidylinositol 3-kinase
AKT	Protein Kinase B
PTEN	Phosphatase and tensin homolog
TGFβ	Transforming growth factor β
OCT4	Octamer-binding transcription factor 4
PIs	Proteasome inhibitors
PROTACs	Proteolysis Targeting Chimeras
UAEs	Ubiquitin-activating enzyme
CRLs	Cullin-RING ubiquitin ligases
ABC-DLBCL	Activated B cell-like Diffuse large B-cell lymphoma
GCB-DLBCL	Germinal centre B cell-like Diffuse large B-cell lymphoma
Need-4	Neural precursor cell expressed developmentally downregulated gene 4
AR	Androgen receptor
VHL	Von-Hippel Lindau
MeBS	Methyl Bestatin
CRBN	Cereblon
BRD4	Bromodomain-containing protein 4
C-myc	Avian myelocytomatosis viral oncogene homolog
CRABPs	Cellular retinoic acid binding proteins
TACC3	Transforming acidic coiled-coil-containing protein 3
enDUBs	Engineered DUBs
SSMPB	Sulfosuccinimidyl 4-(p-maleimidophenyl)butyrate
PBS	Phosphate buffer solution
HRP	Horse-radish peroxidase
TMB	Tetramethylbenzidine
NCI	National cancer institute
DEEP	Differential epitope mapping
STD	Saturation transfer difference
NMR	Nuclear magnetic resonance
WWP2 2	WW domain containing protein
LDA	Lithium diisopropylamide
1D	1 dimensional
2D	2 dimensional
RT	Room temperature
NCS	N-Chlorosuccinimide
NBS	N-Bromosuccinimide
CBO	1-chloro-1,2-benzoiodozol-1(3H)-one
CDI	Carbonyldiimidazole
TCDI	Thiocarbonyldiimidazole
IPA	Isopropyl alcohol
HFIP	Hexafluoroisopropanol
TFA	Trifluoroacetic acid
RMSD	Root-mean square deviation
RBF	Round-bottom flask
Sat.	Saturated
Sep.	Separatory
Pdb	Protein data bank
CTAB bromide	Cetyltrimethylammonium
Hex.	Hexanes
Pet. E	Petroleum ether

1 The Ubiquitin-Proteasome system

1.1 Overview & the Players

The ubiquitin-proteasome system (UPS) is one of the two systems responsible for the regulation of cellular protein quantity and quality, the other proteolytic system is called the autophagy-lysosomal system, and both ensure cellular homeostasis.[1] The UPS degrades proteins of interest (POI) that are usually no longer required by the cell or are in some other way redundant in a specific manner. But before the degradation of POI occurs, post-translational modification (PTM) of the protein with a poly-ubiquitin chain has to occur (Figure 1-1, a), ubiquitin (Ub) being a highly conserved 76 amino acid protein (Figure 1-1, b). Apart from the UPS, ubiquitin PTM can affect other downstream processes too. Ubiquitin was first identified in 1975 as being expressed in all eukaryotic cells and was originally labelled as ubiquitous immunopoietic polypeptide.[2] Later work in the 1980s elucidating the functions of Ub and the components of the ubiquitination pathway earned Aaron Ciechanover, Avram Hershko and Irwin Rose the Nobel Prize in Chemistry in 2004 "for the discovery of ubiquitin-mediated protein degradation".[3]
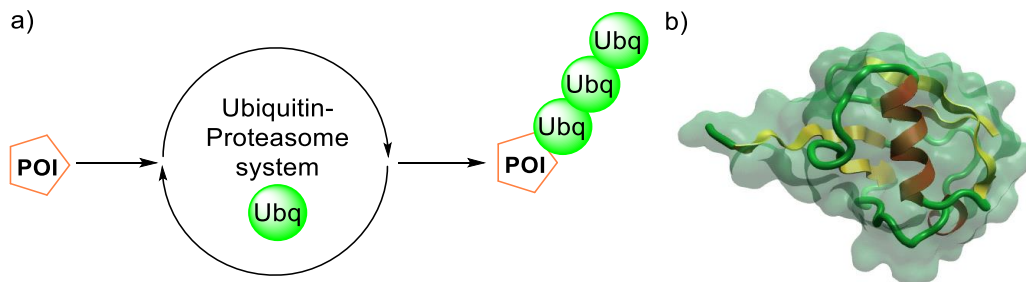
Figure 1-1: a) Simple diagram for the poly-ubiquitination of POIs (Pro.= POI, orange, Ubiquitin = green, Ub) by the UPS. b) crystal structure for ubiquitin (PBD: 1UBQ, visualised using Flare Cresset software)

The UPS system has four essential components for installing Ub and later protein degradation, the E1 ubiquitin-activating enzyme, the E2 ubiquitin-conjugating enzymes, the E3 ubiquitin ligase enzymes and the 26S proteasome.[4] The UPS ubiquitinates target POIs by transferring Ub from E1 to E2 enzymes, which then with the aid of E3 enzymes, attach ubiquitin to the POI (Figure 1-2, a).[5] Poly-ubiquitin chains are formed by the addition of Ub onto POI-Ub adducts, for proteasomal degradation specifically at the Lys48 lysine residue of Ub, although other variations of ubiquitination have other broad downstream cellular responses.[6] The 26S proteasome is a macromolecular intracellular structure which recognises these polyubiquitin chains and degrades the tagged POI into polypeptides whilst deubiquitinating enzymes (DUBs) on the proteasome recycle Ub

(Figure 1-2, b).[5,6] The PTM of POIs with Ub is reversible, saving the poly-ubiquitinated POI from proteasomal degradation, and DUBs are responsible for this too.[6] DUBs can be present as part of the proteasome or present in the cytosol.
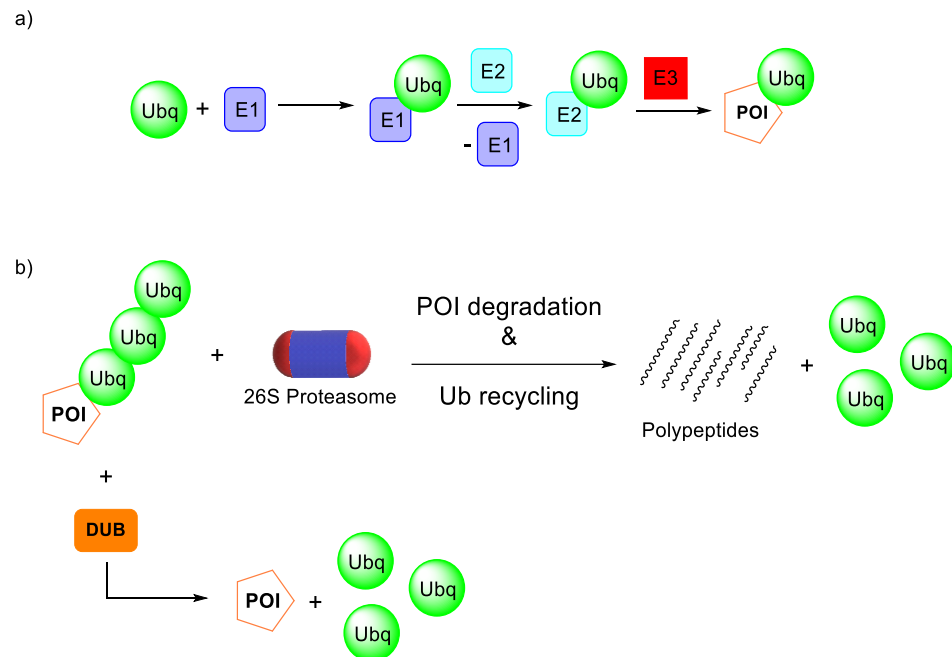
Figure 1-2: a) Simplified overview of POI ubiquitination by the E1, E2 and E3 enzyme cascade, b) Action of DUBs on ubiquitin chains post- and pre-POI degradation.

1.2 <u>Selectivity & Ubiquitination</u>

Obviously, there are many thousands of proteins present within a cell at any one time or cell cycle stage, so the UPS requires high levels of selectivity to ensure the correct proteins are removed. In addition, misfolded proteins are an inevitable by-product of endogenous protein synthesis, and to avoid build-up of non-functioning proteins they also need to be removed from the cell.[7] To achieve this, the hierarchically organised E1-E2-E3 enzyme cascade sequentially transfers ubiquitin from E1 to E2 enzymes, and with the help of E3 enzymes, to the target protein with increasing levels of selectivity occurring at each step, at the end of the cascade each E3 enzyme transfers ubiquitin to several specific POIs. At the start are two E1 enzymes, approximately 40 E2 enzymes and over 600 E3 ligases, enabling high the specificity of ubiquitin attachment to many proteins.[8,9]

Ubiquitination can occur in either a stepwise fashion or in a single attachment of a preformed polyubiquitin chain.[10] Ubiquitin modification of proteins occurs by isopeptide bond formation between the *C*-terminus of ubiquitin and a lysine residue on the POI. For sequential ubiquitin polymerisation, the now attached ubiquitin can support further

ubiquitin attachment at any one of the seven lysine residues present on its surface, as well as its *N*-terminus. Not only that, but further branching at later attached ubiquitin proteins can also occur as well as the attachment of ubiquitin-like proteins (UBls) and further phosphorylation and acetylation of ubiquitin and UBLs.[6,11] This creates a significant diversity in further signalling pathways of poly-ubiquitin/UBL bound POIs, and indeed a broad range of cellular responses have been found to occur upon ubiquitin conjugation, such as protein-protein interactions (PPIs), localisation, enzyme activity, DNA repair, cell cycle progression and cell signalling.[12] Additionally, the autophagy-lysosome system also uses ubiquitin as a signal for protein degradation, so it's easy to see that ubiquitination regulates many functions critical for maintaining the health of the cell.[6] In light of this, in order to keep the contents of this chapter concise, it will concern itself only with the protein degradation resulting from proteasomal degradation because it is the major route by which proteins are regulated within cells, with approximately 80% of all proteins being degraded in this fashion.[13,14]

1.3 <u>The Process of Ubiquitination</u>

E1 enzymes catalyse the first step in the ubiquitination process, and this is the only step requiring energy provided in the form of ATP. Structurally, E1 enzymes have three important domains, the first is made up of two adenylation active site repeats.[5,15] The second is a catalytic cysteine domain which forms an E1-Ub thioester bond in ubiquitin activation. Lastly, the *C*-terminal ubiquitin fold domain (UFD) is responsible for binding the cognate E2 enzymes and aiding in the interaction and transfer of ubiquitin to E2 enzymes.

At the outset of the ubiquitination process, the E1 enzyme conjugates to Mg^{2+}-ATP and ubiquitin (Figure 1-3). It then catalyses the *C*-terminal acyl-adenylation of ubiquitin.[16] A large conformational change is then thought to occur which effectively remodels the enzyme towards thioester bond formation instead of adenylation.[5] The E1 enzyme forms a thioester bond between its catalytic Cys residue and the adenylated ubiquitin *C*-terminus through thiol attack of ubiquitin-AMP.[15] Subsequently, the E1 enzyme interacts with another set of ubiquitin and Mg^{2+}-ATP, so that it has a Ub bound *via* a thioester linkage and another associated non-covalently at the adenylation active site.[15]
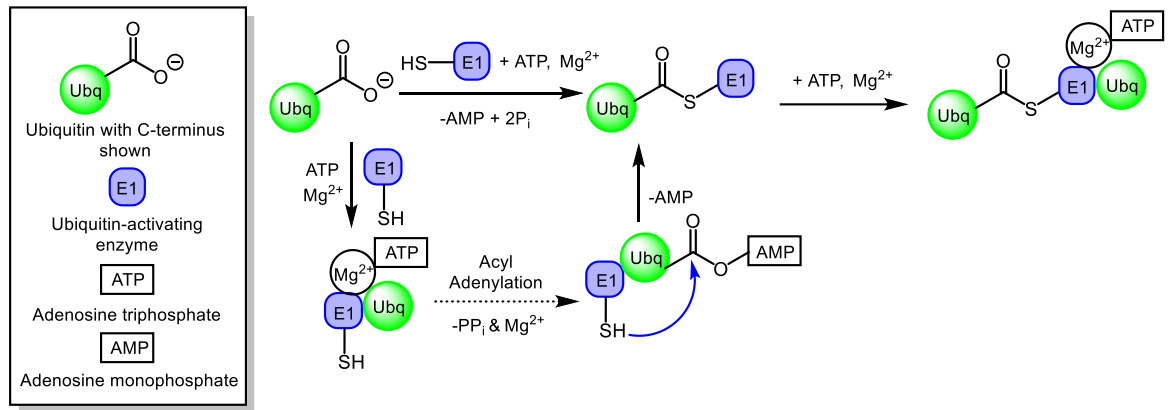
Figure 1-3: Ubiquitin activation by E1 enzymes.

After E1-ubiquitin complex formation, Ub is transferred to the respective E2 conjugating enzyme, through an E1-Ub – E2 interaction and transthioesterification (Figure 1-4). The UFD domain of E1 non-covalently interacts with the E2 enzyme, with structural studies indicating the concave face of the UFD binds to the *N*-terminal helix and part of the core domain of the E2 enzymes.[15] It is proposed the UFD domain undergoes a significant conformational change, bringing the E1 donor and E2 acceptor Cys residues within close proximity of each other for transthioesterification to occur.[15] It is thought that the UFDs conformational flexibility is what allows the E1 enzyme to interact with the numerous E2 enzymes.
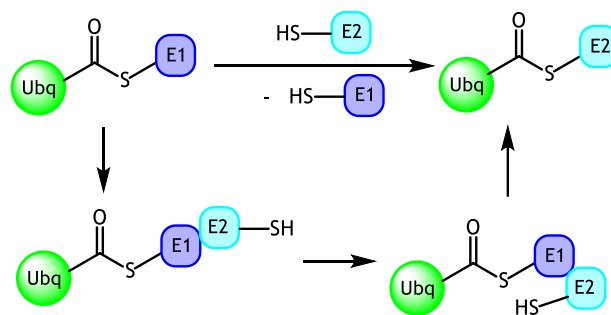
Figure 1-4: Transthioesterification of ubiquitin from E1 to E2 enzymes.

The ubiquitin transfer to E2 may be controlled by a 'thioester switch', which toggles E1-E2 affinity. E1 is thought to display different affinities for E2 enzymes in the E1-free, E1-Ub and doubly ubiquitin-loaded (E1-Ub$_2$) states. E1-Ub$_2$ has nanomolar affinity for cognate E2 enzymes,[17] and free E1 enzymes have lower affinity, suggested by the ability of the E1-free enzyme to separate from E2s during gel filtration.[18] After E2 is charged with ubiquitin, it is apparently released from the E1-E2 complex, as E1 enzymes can undergo multiple reaction cycles in the presence of excess E2 enzyme.[15] The differing affinities would allow the progression of the E1-E2-E3 cascade, as E1-E2 and E2-E3 interactions are mutually exclusive.

All E2 enzymes share a conserved ubiquitin-conjugating (UBC) domain containing the catalytic Cys residue within which the Ub-thioester bond forms. E2 enzymes may have additional N- or C-terminus extensions and/or insertions within the UBC domain which confer additional purposes.[19] Many residues surrounding the active site Cys are highly conserved as several interact with ubiquitin and the E1 enzyme.[20] In contrast, many of the most poorly-conserved residues between E2 enzymes are on another face, which may be what enables specific downstream interactions with the respective E3 enzymes.[21] A study using an E2-ubiquitin complex bound by an oxyester bond (instead of the native thioester bond) found secondary non-covalent interactions between ubiquitin and E2 to be minimal apart from at the very extreme of the C-terminus of ubiquitin.[22] This indicates that perhaps the main interactions for the transfer of ubiquitin from E1 to E2 occur between the two enzymes and is not mediated by ubiquitin, which could explain the specificity observed in these pathways.

The high specificity of E2-E3 interaction and therefore ubiquitination of specific POIs is probably dictated by a small number of individual residues between the E2 and E3 enzymes.[21] For example, in the E2 enzyme UbcH7, which interacts with both E6-AP and c-Cbl E3 ligases, the residues Phe63, Pro97 and Ala98 pack into a hydrophobic groove in each E3.[21]

Ubiquitin is transferred from the E2 enzyme to the target protein in an E3 enzyme-dependent manner, forming an isopeptide bond between the glycine C-terminal residue of ubiquitin and the ε-amino group of a Lys residue on the POI. How this occurs depends on the type of E3 enzyme.[23] E3 enzymes can belong to distinct families, which transfer ubiquitin to the target protein in differing ways. The major families are RING (Really Interesting New Gene), RBR (Ring Between Ring), and HECT (Homologous to E6-AP Carboxyl Terminus) E3 enzymes, which further contain sub-families.

1.3.1 RING E3 Ligases

E3 ligases containing RING domain can be sub-divided into RING and RING-like U-box families, and make up the majority of the over 600 known E3 ligases.[19] Additionally, RINGs not only exist as monomers, but can form and function as homodimers and heterodimers.[24] Structurally, the RING domain contains two loop regions which surround a central α-helix groove, which provides a surface for E2-Ub interaction with RING E3s, specifically with the UBC domain of E2s.[19] RING domains are arranged in a cross-braced structure, and contain eight residues coordinating to two catalytically inert zinc ions, this surface can also recruit POIs. RING ligases only mediate the transfer of ubiquitin to the

POI from E2 without additional thioester bond formation on the E3 enzyme.[19] The spacings between the Zinc coordinating ligands, rather than the primary sequence, is conserved among the RING ligase family.[21] RING-like U-box ligases are structurally similar but without the coordinating zinc species.

From crystallographic studies the RING ligases have been shown to activate the E2-Ub complex towards transfer of Ubiquitin to the POI.[25,26] The conformation of the POI-RING E3-E2-Ub complex geometrically disposes the thioester bond towards nucleophilic attack by the recruited POI Lys residue, whereas the free E2-Ub complex adopts a different, more stable conformation without RING E3 interaction. Using NMR studies, it was shown that generally E2-Ub complexes are flexible and can adopt a variety of conformations. However, upon interaction with RING E3 ligases, a bias occurs where the E2-Ub conformations are restricted to a more closed set conformations, thereby activating the complexes towards Ub transfer.[27]

A model has been constructed for the catalytic transfer of ubiquitin to the POI. Restriction of conformation of the C-terminal tail of Ub and thioester bond to the 'closed' conformation occurs upon formation of the POI-E3-E2-Ub complex (Figure 1-5).[25,27] E2 Asp and Asn residues create a microenvironment capable of partially deprotonating a POI's Lys residue which breaks the E2-ubiqutin thioester bond. They additionally aid in stabilising the anionic tetrahedral intermediate following nucleophile attack, which is followed by formation of the POI-ubiquitin amide bond. Although the specific interactions of cognate E2-Ub complexes, POIs and RING ligases will vary, it is thought that these are general processes leading up to the catalytic ubiquitination of POIs for RING ligases.
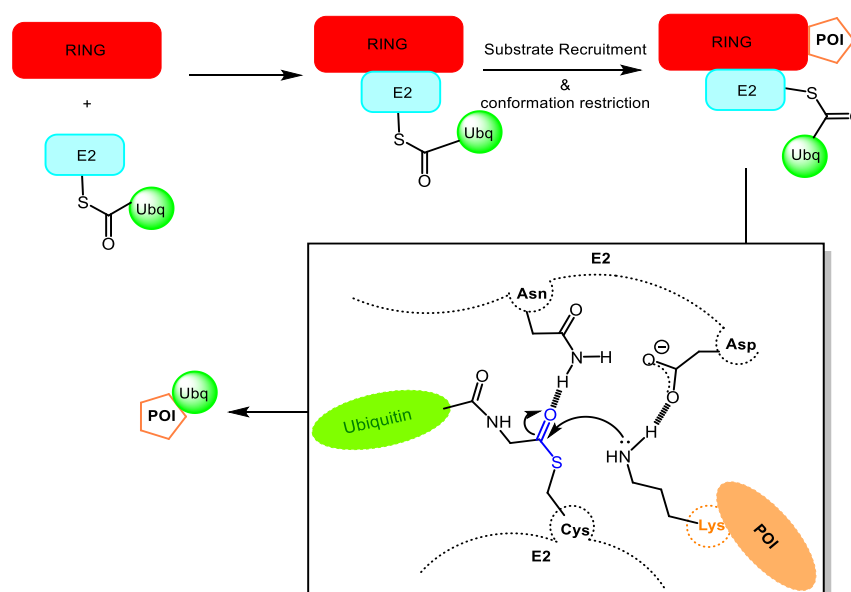
Figure 1-5: RING E3 ubiquitination of the target protein. Substrate recruitment is shown as a direct RING-POI interaction for clarity.

The recruitment of POIs to RING E3 ligases is hugely varied and complex, and can comprise of either direct interactions of POIs with the RING ligase or indirectly within network interactions with HECT E3s and DUBs.[19] Additionally, metabolites have been implicated in POI recognition by RING E3s by either direct RING-metabolite interactions or indirect POI-metabolite interactions. PTM, most commonly phosphorylation, has also been shown as a means of POI ubiquitination as PTM of either E3 or substrate alter ligase activity or POI recognition by the RING ligase. Other PTMs include hydroxylation, glycosylation and acetylation.[19]

1.3.2 <u>HECT E3 Ligases</u>

E3 ligases that have a ≈ 350 aa homologous to E6-AP carboxyl terminus (HECT) domain are called HECT E3 ligases. The HECT domain is located at the C-terminus and their N-terminus are more varied, mediating substrate targeting.[23,28] Approximately 30 HECT ligases are known, with the HECT domain consisting of an N-lobe which recruits the E2-Ub complex, a C-lobe which has the catalytic Cys residue with the last 60 amino acids of the HECT C-lobe controlling the specificity of the ubiquitin chains they interact with.[23] These two lobes are connected by a flexible linker region which is essential for ligase activity.[5]

The HECT family may be split into several sub-families, namely the HERC (6 members), Nedd4 (9 members) and 'other' HECT ligases (13 members).[29] The HERC family is made up of 6 enzymes containing the RLD domain, and can be further split into 2 large HERCs, containing two or three RLD domains, whereas the other HERCs are small HERCs which possess only one RLD domain.[30] The 'other' HECT family do not share a significant commonality and contain a large variety of N-terminal domains. The Nedd4 family is the second largest of the HECT sub-families, with nine members. These Nedd4 enzymes contain 2 to 4 WW domains, a C2 domain and the HECT domain.

Ubiquitination processes utilising HECT domain enzymes undergo an additional transthioesterification event prior to target protein ubiquitination, forming a thioester bond to the HECT E3 ligase before transfer to the substrate (Figure 1-6). In the Nedd4 family of HECT E3 ligases, the E2-Ub complex is thought to interact with the HECT N-lobe primarily through hydrophobic interactions, and the POI is recruited by the N-terminal WW domains. Although more recently the HECT domain has been implicated in substrate recruitment in the Nedd4 family.[10,31] The E2-Ub – N-lobe interactions are thought to bring the catalytic Cys located on the C-lobe and the E2-bound ubiquitin close together for transfer.[31] After Ub transfer to the E3 Cys on the C-lobe a consequent 130° rotation of the

ubiquitin-bound C-lobe (C-Ub), brought about by the flexible linker region, brings the C-Ub lobe into close proximity to the Lys residue of the POI. Through this domain rearrangement a large ubiquitin binding site forms, securing the C-Ub C-terminal tail into place and allowing the Ub transfer to the POI.[5] Because the E3 ligase dissociates slowly from this complex, it can undergo multiple ubiquitination cycles in a processive way, so can become polyubiquitinated ready for proteasomal degradation in one E3-POI encounter.[21] Alternatively, it may undergo ubiquitination in a distributive manner where only one ubiquitin is transferred during an E3-POI encounter.
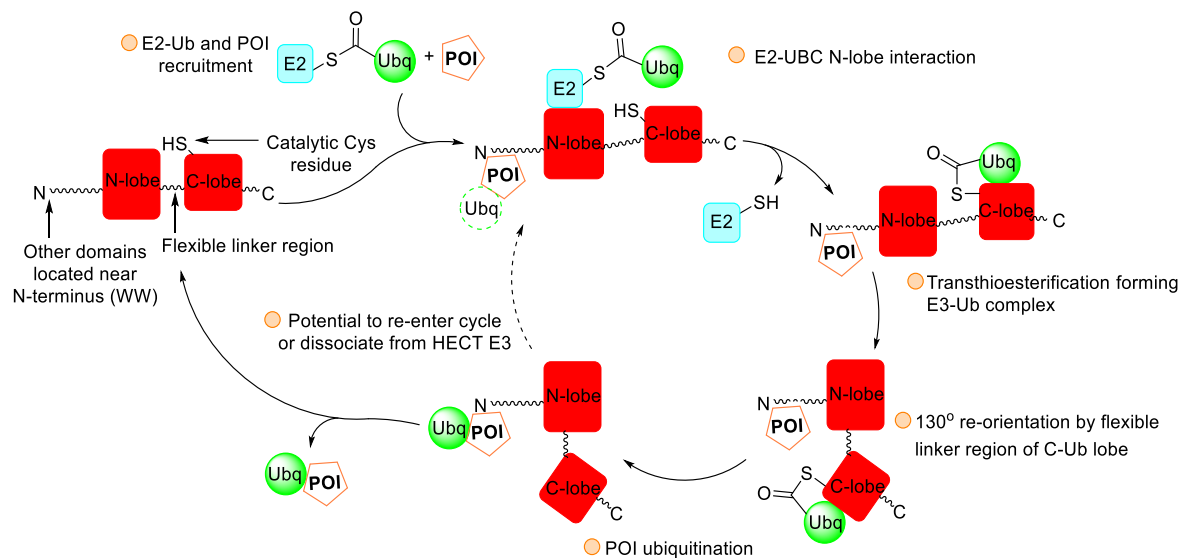
Figure 1-6: Pathway of HECT E3-mediated ubiquitination of the target protein.

1.3.3 <u>RBR E3 Ligases</u>

RBR E3 ligases are the smallest family of E3 ligases, with currently 14 known.[32] A rather important member of the RBR family is Parkin. The structure of Parkin contains a Ub-like domain near its N-terminus and later near its C-terminus a Cys-rich intermediate region between two RING domains, aptly named as RING-between-RING (RBR).[33] RBR ligases are multidomain enzymes containing the RBR domain, itself composed of RING1-IBR-RING2, where this domain is situated in relation of other domains can vary.[34] RING1, IBR and RING2 coordinate two zinc ions respectively. The RING1 subdomain is structurally related to the RING domains of RING ligases, but RING2 subdomain is more structurally similar to IBR. A conserved Cys residue not implicated in Zn ligation is present in RING2 and this site is where Ub binds, as RBRs form covalent complexes with Ub before transfer to POIs. Additionally, RING2 subdomains are thought to bind specific POIs or polyubiquitin chains for Ub transfer, as RBR can also ubiquitinate in a processive manner.

IBR structures can vary greatly between RBR ligases, and are thought to aid in binding ubiquitin during Ub-transfer.[34]

The mechanism of ubiquitin transfer from E2 to RBR E3s then POI is understood to utilise a HECT-RING hybrid mechanism.[34] Very recently, work on RBR ligases HOIL-1, RNF216 and historical work on Parkin, HH9ARI and HOIP revealed a conserved RBR-E2 complex between ligases (Figure 1-7, Boxed).[35] This complex is relatively general between the RBRs investigated and is defined by a conserved RING1 binding site for E2 (of the E2-Ub complex) and a conserved Ub donor binding site (Ub of the E2-Ub complex) in the IBR-RING2 linker with the main interactions stemming from this latter interaction. A stabilisation of an open conformation of the E2-Ub complex with alignment of the E2-Ub and RING2-E3 active sites is thought to occur. These interaction sites produce a E2-RBR E3 complex conformation that allows efficient transfer of ubiquitin from E2 to the RBR E3.[35] Also, an allosteric Ub or UBL allosteric binding site at the RING1-IBR interface for activation by distinct Ub or UBL proteins was identified, which may represent a positive feedback loop to self-activation of RBRs, as binding of Ub or UBLs increases affinity of the RBR for E2-Ub complexes.
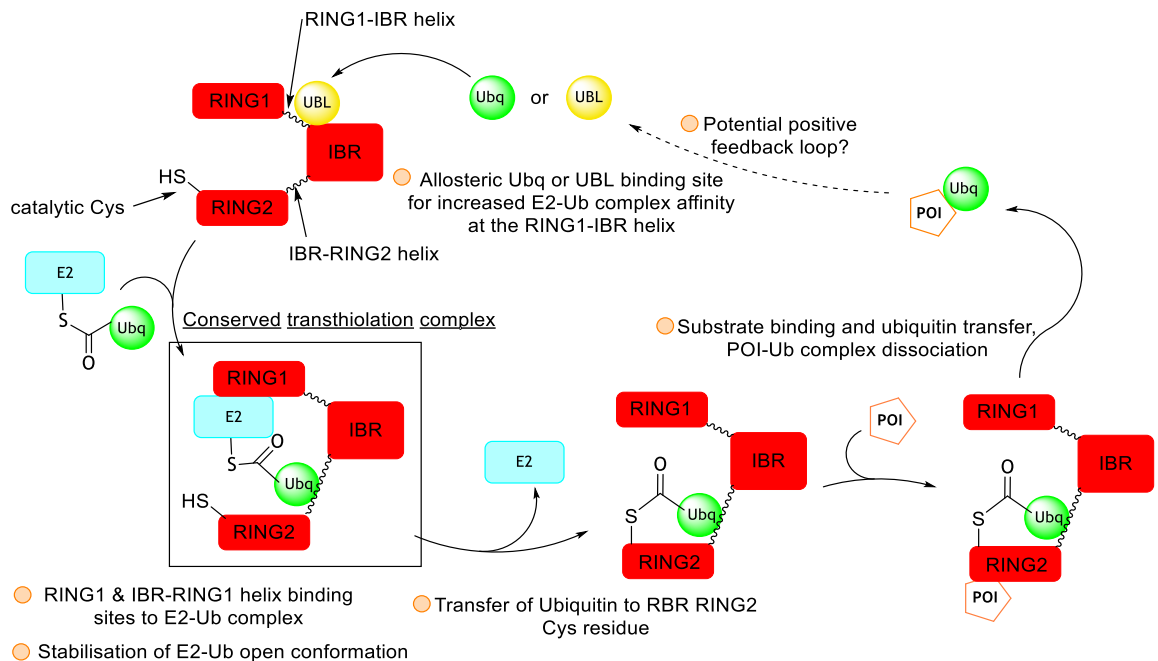
Figure 1-7: Working model for RBR ubiquitin transfer from E2 to substrate.

It is currently thought that transfer of ubiquitin to the substrate occurs by the substrate (or the ubiquitin chain) interacting with the RING2 subdomain of RBR enzymes, although this is not definitive.[34] Structural differences between RING2 subdomains in different RBR enzymes are believed to confer substrate and ubiquitin chain specificity to the ligases.[35]

The poly-ubiquitinated POI product of the ubiquitination cycle may act as the allosteric binder on the RING1-IBR region, amplifying the ubiquitination event.

1.3.4 <u>26S Proteasome</u>

The 26S proteasome is a macromolecular protease, with the common core referred to as the 20S core which is about 700 KDa in size. It is made up of 14 α and 14 β subunits which are arranged in rings of 7, with two α subunit and two β subunit rings. These rings are stacked on top of each other to make the 20S core (Figure 1-8). The two outer rings are made up of the α subunits (α rings) and the two inner rings the β subunits (Figure 1-8, right). A proteasome with the common 20S core and two PA700 (19S) 'caps' is known as the 26S proteasome and is both ATP and ubiquitin dependent. The 19S unit is composed of 19 or more protein units and whose role is to both recognise and remove ubiquitin chains from the target proteins as well as the unfolding and translocation of the protein into the core.[13]
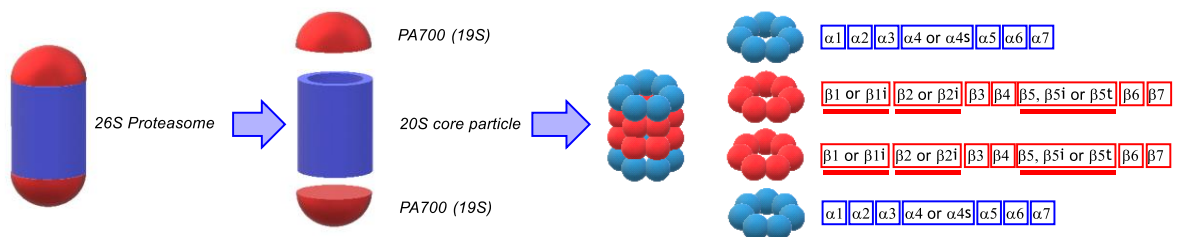
Figure 1-8: Structural outline of the 26S proteasome.

As shown these rings can be made up of various subunits, and do not have to be identical to the other respective ring type. The α rings are structural with their N-termini forming barriers that can prevent proteins from entering or leaving the core. The β rings confer the catalytic activity, specifically the β1, β2 and β5 subunits (or homologs of as underlined in Figure 1-8). These catalytic subunits have an N-terminal threonine (Thr1) present and act as part of a catalytic triad also containing Asp17 and Lys33 to lyse proteins. This occurs in two steps: firstly, by forming an ester bond with the N-terminal portion of the substrate, which is then hydrolysed releasing the product, and regenerating the active Thr1 (Scheme 1-1). The three catalytically active subunits have specific types of amino acids which they cleave. β1 cleaves substrates with acidic residues, whereas β2 cleaves basic residues and β5 hydrophobic residues.[36]
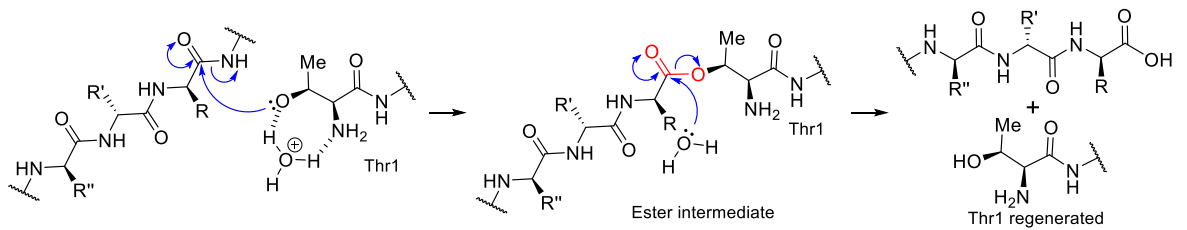
Scheme 1-1: Mechanism for the degradation of proteins.

In summary, for protein degradation to occur the target protein must have undergone polyubiquitination (Figure 1-9).[7] Proteins targeted for ATP-dependent 26S proteasome degradation are tagged with ubiquitin in a highly regulated process.[37] For the target protein to become ubiquitinated, Ub must bind with ubiquitin activating enzyme, which transfers the ubiquitin to an E2 conjugating enzyme. This E2 ligase then binds to an E3 ligase, which is specific for and has recruited the POI. If the E3 ligase is part of the RING ligase family, the ubiquitin is directly transferred from the E2 carrier enzyme to the target protein. If, however, the E3 ligase is part of the HECT or RBR ligase families a thioester bond forms between Ub and the E3 ligase HECT domain (or RING2 domain) first, with removal of the E2 enzyme and transfer to the POI. The polyubiquitinated target protein is recognised by the 26S proteosome, unwound and de-ubiquitinated, passing through the core and degraded into polypeptides for further degradation by proteases.
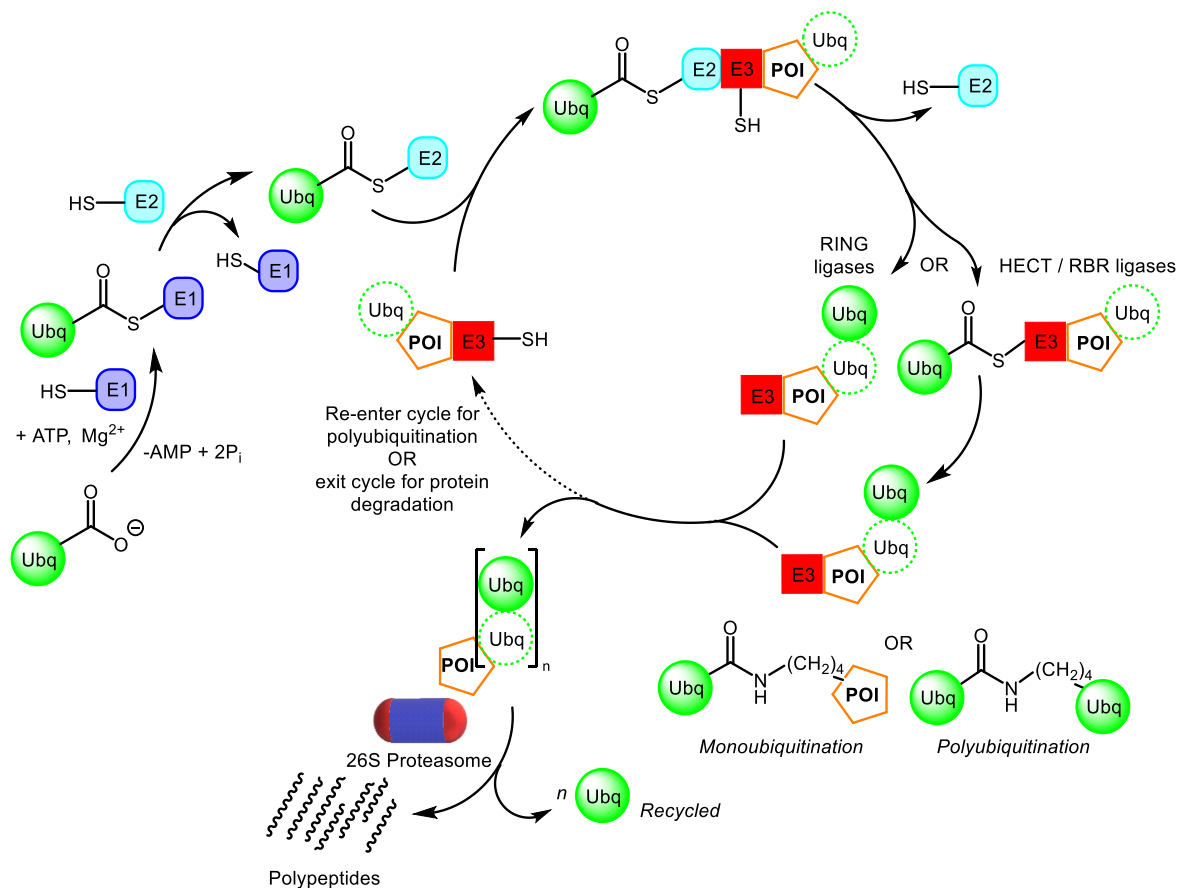
Figure 1-9: Summary of the ubiquitination process from ubiquitin activation and polyubiquitin chain formation to target protein degradation.

1.4 <u>WWP2: A HECT E3 Ligase</u>

The HECT family of E3 ligases is made up of three sub-families, based on the similarities of the HECT domain: The HERC, Nedd4 and 'other' HECT E3s. HERC ligases are classified as such by having one or more regulator of chromosome condensation 1 (RCC-1)-like domains (RLD domains).[38] This family has 6 members which can be further split into two large HERCs and 4 smaller ones.[28] The group classed as 'other HECT ligases' have a varied *N*-terminal domain. The Nedd4 group is the largest subgroup of HECT E3 ligases, containing 9 members, which in addition to the HECT domain contain a C2 domain and two to four WW domains.[28]

The WWP2 ligase, (atrophin-1-interacting protein 2 (AIP2) or WW domain-containing protein 2) is a member of this Nedd4 family, and consists of the HECT domain, the C2 domain and four tandem WW domains. (Figure 1-10). The enzyme was first identified in 1997 and is known to have three isoforms.[39] The full length isoform (WWP2-FL) contains

all domains mentioned previously, the *N*-terminal isoform (WWP2-N) contains only the WW1 and C2 domains, but lack the HECT domain. Finally, the *C*-terminal isoform (WWP2-C) contains the WW4 and HECT domains.[40]
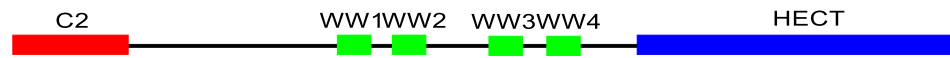
Figure 1-10: Schematic of the WWP2 HECT E3 ligase domains, not to scale.

The C2 domain is responsible for the membrane localisation of the E3 ligase and substrate/ligase trafficking by interacting with calcium ions, phospholipids, and membrane proteins. With a length of $\approx$ 120 residues, it forms 8 folded β-sheets which make up a β-sandwich.[28] The C2 domain is also a key regulator in controlling the ligase activity, by being able to enforce autoinhibition through an interaction with the C-lobe of the HECT domain.[41] Through structural studies, the mechanism of autoinhibition is thought to occur by the C2 domain forcing a conformational change in the HECT domain into an inactive state, which is unable to form the thioester bond to Ub and hence prevents substrate ubiquitination.[42]

The WW domains role is primarily to target protein recruitment and selectivity, they are termed WW because they contain conserved tryptophan residues (single letter code = W) spaced 20 – 22 residues apart.[28] The rest of the domain is made up of three antiparallel β-sheets, which form a hydrophobic channel suited for target protein binding. These domains mainly interact with proline-rich motifs located on POIs, specifically Leu/Pro-Pro-(X, any AA)-Tyr sequences.[43] The HECT domain, as described above, is responsible for the transfer of Ub from an E2 to the POI.

Physiologically, the WWP2 enzyme has been shown to be involved in craniofacial development and chondrogenesis.[44] WWP2 was shown to be highly expressed by cartilage-specific transcription factor Sox9 in cartilage tissues.[45] Furthermore, WWP2 was shown to be an E3 ligase for the homeobox transcription factor Goosecoid, which is known for its importance in craniofacial development.[46] It was found that WWP2 monoubiquitinates Goosecoid, improving its transcriptional activity and thereby increasing the expression of Sox6, a cartilage regulatory protein.

1.4.1 WWP2 & Cancer

The Nedd4 family has more broadly been implicated in cancer, with Nedd4 ligases including WWP2 shown to be critical components and regulators of pathways fundamental to cancer.[10] WWP2 has been shown to be frequently over-expressed in cancer, and so it

also has several substrates which are lost or mutated, including the tumour suppressor protein PTEN, Smads and Oct4.[37,47] In animal models, WWP2's overexpression caused increased tumour outgrowth.[48] Up to 70% of primary prostate cancer tumours lose a copy of the PTEN gene, with cancer genomic analysis of patients with prostate cancer showing almost 40% of patients with castration-resistant prostate cancer have the WWP2 gene amplified.[28]

1.4.1.1 <u>PTEN</u>

PTEN is a lipid phosphatase frequently mutated in cancer and it is a substrate for WWP2 mediated polyubiquitination and proteasomal degradation.[1,44] PTEN under physiological conditions down-regulates the PI3K (phosphatidylinositol 3-kinase) signalling pathway. Increased PI3K levels normally lead to AKT (protein kinase B) activation, which controls several functions such as cell survival and death.[48]

Once activated by membrane receptor tyrosine kinases, PI3K phosphorylates phosphatidylinositol 4,5-bisphosphate and converts it to the 3,4,5-triphosphate. Phosphatidylinositol 3,4,5-triphosphate then accumulates in the cellular membrane and this leads to the activation of AKT.[48] PTEN downregulates PI3K signalling by dephosphorylating phosphatidylinositol 3,4,5-triphosphate back to the bisphosphate, so downregulating AKT signalling. PTEN is a tumour suppressor because this down-regulation reduces downstream oncogenic AKT-mediated signalling.[49] As an example of the oncogenic AKT-mediated signalling, WWP2-mediated depletion of PTEN rendered prostate cancer cell lines resistant to stress-induced cell death.[44,48]

The PTEN Tyr155 residue has been shown to be mutated in several cancers, and that this mutation leads to reduced protein stability and has been shown to play a critical role in WWP2 interaction.[48] Because of this it has been postulated that under normal conditions the residue is phosphorylated and therefore prevents any PTEN-WWP2 interaction, but in cancers with this mutation prevents phosphorylation and therefore allows WWP2-PTEN interaction, polyubiquitination and proteasomal degradation.[48] The degradation of PTEN then allows oncogenic AKT-mediated signalling.

1.4.1.2 <u>TGFβ</u>

Normally, transforming growth factor-beta (TGFβ) signalling plays a key role in regulating tissue homeostasis and modulating the immune response by regulating cell loss and renewal.[40] This pathway is involved in many processes, including cell differentiation, growth and apoptosis among other cellular functions. On the other hand, TGFβ has a

multifunctional role in cancer and late-stage tumours. Notably, acting through Smad transcription factors, it is responsible for epithelial-mesenchymal transition (EMT), a process of converting epithelial cells into mesenchymal cell which are highly invasive and a prerequisite for tumour cell metastasis.[44]

The TGFβ ligand binds with a specific type II receptor, which then phosphorylates a type I receptor after forming a tetrameric complex with another type II and two type I receptors.[50] The type I receptor in turn phosphorylates regulated Smad2 and Smad3 (R-Smads). After phosphorylation, Smad2 and Smad3 can complex with co-Smad (Smad4). This complex then accumulates in the cell nucleus to act as a transcription factor for targeted gene expression. I-Smads (inhibitory Smads) work to supress the activity of R-Smads.

The WWP2 isoforms have been shown to play differing roles in the regulation of TGFβ/Smad signalling activity linked to EMT.[47] The isoforms have been shown to bind differently to different Smads and degrade them, which regulates the levels of activating and inhibitory Smads and therefore gene expression. WWP2-FL is thought to have a role in TGFβ-induced cancer cell metastasis, based on its substrate preference for inhibitory Smad7, which is supported by cell-based EMT experiments.[44] Conversely, it was suggested that WWP2-N could serve as a key suppressor of events involved in the instigation of EMT and metastasis, namely because of its ability to suppress the levels of R-Smads.[44]

1.4.1.3 <u>Oct4</u>

Oct4 (Octamer-binding transcription factor 4) is a homeodomain transcription factor that is involved in the regulation of pluripotency of stem cells.[51] It is usually down-regulated during differentiation, but it has been shown that some malignant cells can regain expression of Oct4.[52] Oct4 also has the capacity to regulate cancer stem cell differentiation and proliferation, with a link found between the expression of Oct4 and cancer prognosis.[53] WWP2 has been shown to regulate Oct4, promoting its degradation during the differentiation of embryonal carcinoma cells but not in undifferentiated embryonal stem cells *via* the 26S proteasome in a dosage-dependent manner.[54] Additionally, by competitively blocking the interaction between WWP2 and Oct4, a decrease in the ubiquitination and later degradation of Oct4 promoted the expansion of tumour initiating cells.[55]

To summarise, TGFβ, PTEN and Oct4 are all involved in key oncogenic signalling pathways linked to cancer cell growth, survival and tumour spread.[56] WWP2 is known as a

regulator of these substrates and therefore WWP2 is an attractive target to inhibit to block these oncogenic pathways and potentially exhibit anti-cancerous effects.

1.5 Small Molecules Targeting the UPS

Practically all enzyme classes within the UPS have been in some way implicated in disease and have been targeted by small molecule inhibitor design. Targeting the UPS is an attractive anti-cancer strategy, and the development of inhibitors has also led to further understanding of specific mechanisms within the UPS. A review of several compounds identified from the literature that are able to inhibit certain parts of the UPS are described below as well as a rather new strategy to highjack the UPS system to target otherwise undruggable proteins.

1.5.1 Proteasome inhibitors

The first synthetic proteasome inhibitors (PIs) contained the protein backbone Gly-Gly, with at the N-terminus a benzyl carbamate and at the C-terminus a leucinal residue (**1.1**, Figure 1-11), which formed a reversible complex with an active site threonine.[57] A later version changing the glycine backbone to isoleucine residues (**1.2**) allowed more potency and cell-permeability.[58] Additionally, the aldehyde group could be substituted for a vinyl sulfone which binds irreversibly (**1.3**).[59] An issue with these compounds was that they inhibited lysosomal and calcium-activated cellular proteases, so were not selective enough for the proteasome to be marketed.[60]
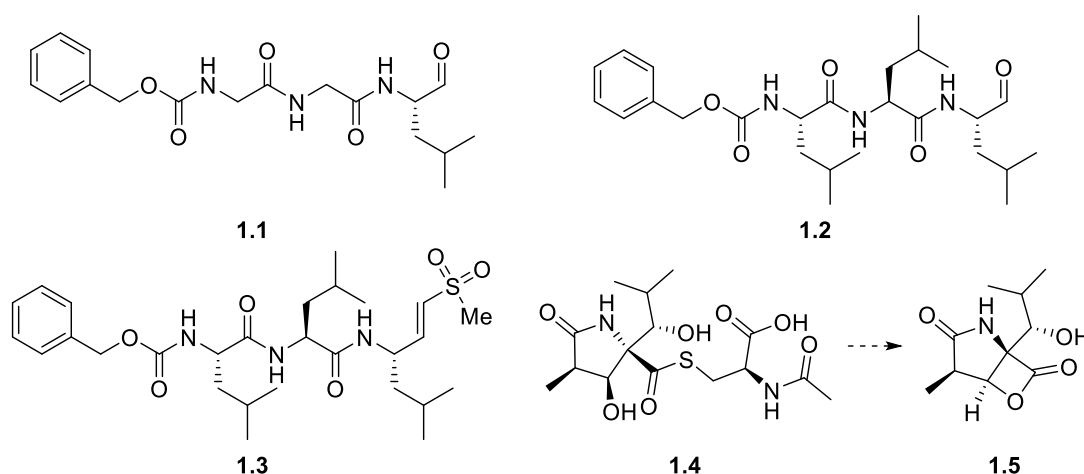
*Figure 1-11: Synthetic peptides (**1.1 – 1.3**), Lactacystin **1.4** and clasto-Lactacystin β-lactone **1.5**, investigated as initial proteasome inhibitors.*

16

A more selective proteasome inhibitor (Lactacystin, **1.4**) was identified from *Streptomyces* extracts, with a metabolite thereof identified as a more potent, selective irreversible inhibitor of the proteasome, called clasto-lactacystin β-lactone, **1.5**.[61,62]

Another natural product identified from *Streptomyces* extracts was epoxomicin **1.6** (Figure 1-12), where the proteasome was found to be its major target, and that it possesses antitumor and anti-inflammatory activities.[60,63] The epoxide functional group forms a covalent bond with the active site threonine, providing irreversible inhibition. Synthetic proteasome inhibitors have been developed (**1.7**, **1.8**), with **1.8** marketed as Bortezomib, which was found to be a potent reversible inhibitor of the proteasome with a rather unusual boronic acid component.[64] Compared to the aldehyde group, which was also investigated on these compounds, the boronate binds more strongly with the catalytic threonine, providing both higher selectivity and potency against the proteasome. In 2003 it was FDA fast-track approved and is used for relapsed refractory myeloma, mantle cell lymphoma as well as new cases of multiple myeloma. Bortezomib represents the first drug targeting the UPS to make it to market.[13]

Epoxomicin, 1.6   1.7   Bortezomib, 1.8
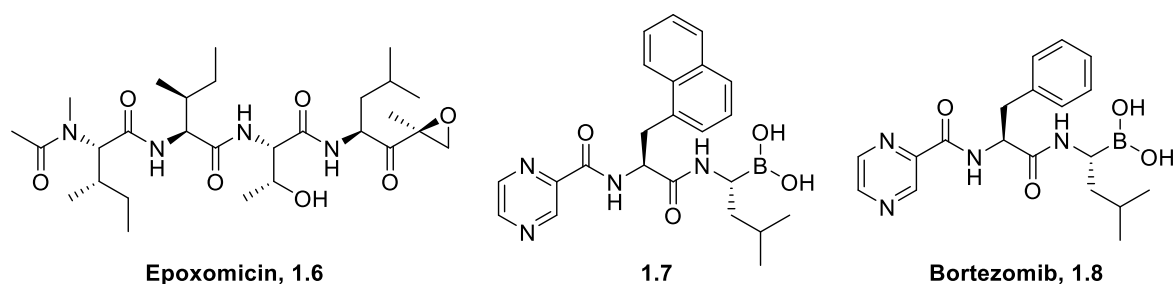
Figure 1-12: Structure of Epoxomicin, analogue of Bortezomib and Bortezomib.

Other FDA-approved drugs such as Carfilzomib also utilised the epoxyketone as the warhead for irreversible inhibition (**1.9**, Figure 1-13).[65] Ixazomib citrate **1.13** was approved as a proteasome inhibitor with the advantage of oral administration over both Bortezomib and carfilzomib, which require subcutaneous or intravenous administration. Just like Bortezomib, Ixazomib contains a boronic acid group, but masked as a boronate moiety which is rapidly hydrolysed into its active form **1.14** upon interaction with plasma.[66] Several other compounds have been developed and are in clinical trials for different types of tumours. These include: delanzomib (**1.10**), marizomib (**1.11**) and oprozomib (**1.12**).[65,67–69] Structural similarities exist between those in clinical trials and those discussed previously, for example, marizomib, isolated from *Salinispora* and otherwise named Salinosporamide A, contains a bicyclic β-lactone-γ-lactam system, as does the active metabolite of Lactacystin. Oprozomib is also structurally related to carfilzomib, with

a peptide-like backbone and epoxyketone warhead. Delanzomib is also related to bortezomib, containing a boronic acid group.
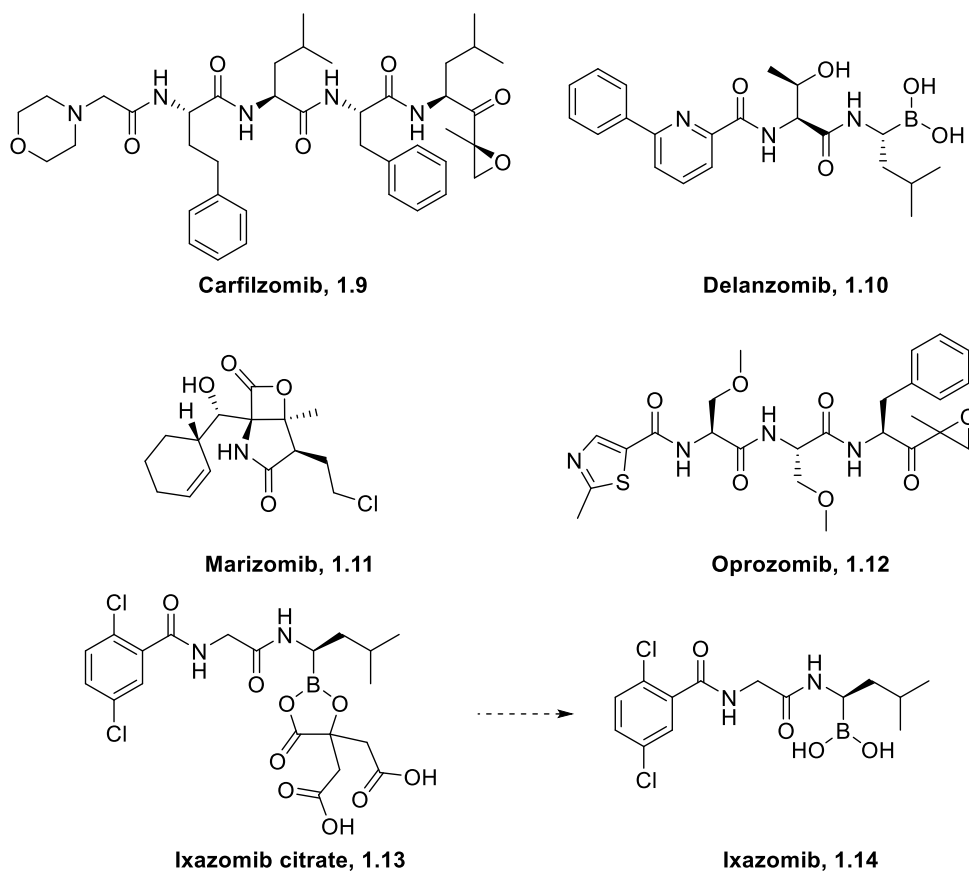
Figure 1-13: Structures of Carfilzomib, Ixazomib citrate and its hydrolysis to Ixazomib, Marizomib, Oprozomib and Delanzomib.

The actual target of these proteasome drugs appears to be the chymotrypsin-like β-5 catalytic subunit within the proteasome, although the targeting of other subunits contributing to their efficacy cannot be ruled out.[13] Lastly, the downstream effects of proteasome inhibition are less well-understood. One potential pathway among others proposed is the accumulation of unfolded and misfolded proteins, which trigger the unfolded protein response followed by apoptosis.[13]

Other small molecules have also been developed that are thought to target the 19S regulatory sections of the proteasome (Figure 1-14). Capzimin **1.15** is one such selective inhibitor, capable of producing the unfolded protein response and blocked the proliferation of cancer cells.[70] In a study aimed at evaluating the therapeutic effect of targeting the proteasome subunit ADRM1 for intrahepatic cholangiocarcinoma (ICC), the authors identified RA-190 (**1.16**), able to significantly reduce ADRM1 levels with ICC cells thereby identifying it as a potential treatment for ICC.[71] Lastly, the utilisation of an assay developed to measure proteasome-mediated protein degradation allowed the discovery of

epidithiodiketopiperazines (ETPs), which blocked the degradation of the model substrates *in vitro*.[72] Further investigation revealed that these ETPs inhibited proteasome function by targeting Rpn11, a proteasomal deubiquitinase as well as other JAMM proteases. Improvement on the ETPs identified SOP11 (**1.17**) which had reduced non-specific effects and induced the unfolded protein response leading to cell death.

Capzimin, 1.15 **RA-190, 1.16** **SOP11, 1.17**
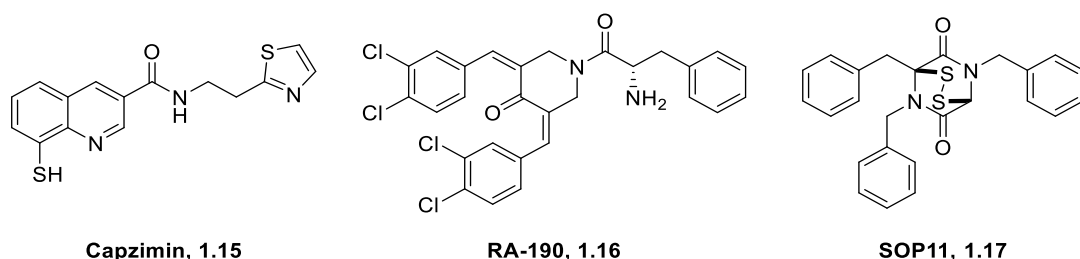
Figure 1-14: Capzimin, SOP11 and RA-190.

1.5.2 E1 inhibitors

Several sulfamate compounds have been identified as E1 inhibitors for potential cancer treatments, mLN4924 (Pevonedistat, **1.18**, Figure 1-15), was found to provide inhibition *via* a novel mode of action, coined substate-assisted inhibition, against NEDD8-activating enzyme.[73–75] This mechanism is thought to arise from the sulfamate NH_2 reacting with the E1-Ub (or E1-Ubl) thioester bond, whilst binding tightly to the adenylation site of the E1 enzyme. This Ub/Ubl-MLN4924 adduct mimics the adenylated Ub (or Ubl) but cannot undergo transthioesterification.

Pevonedistat / MLN4924, 1.18 **1.19**

MLN7243, 1.20 **1.21**
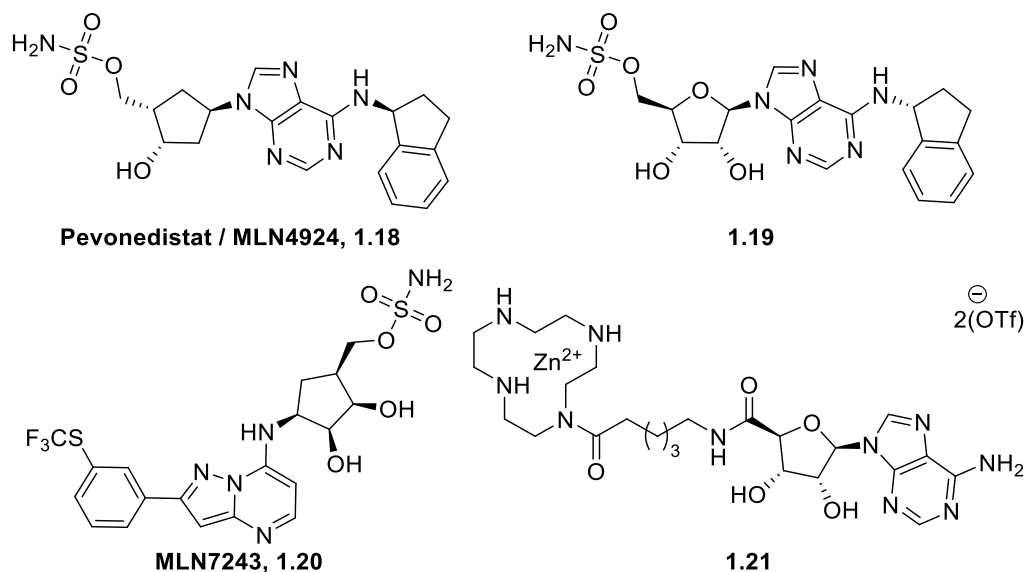
*Figure 1-15: Sulfamate E1 inhibitors (**1.18 – 1.20**) and Zn^{II}-Cyclen complex-containing E1 inhibitor* ***1.21***.

Because all E1 enzymes have the same mechanism of Ub/Ubl activation, other inhibitors based on this core and sulfamate functionality have also been developed, such as **1.19** which is non-specific for E1s, and mLN7243 (**1.20**), a UAE (Ube1) inhibitor.[73,74,76] A selective inhibitor (**1.21**, Figure 1-15) of UFM1-activating enzyme, Uba5, has been developed where the adenosine scaffold is proposed to anchor within the ATP pocket. The Zn[II]-Cyclen complex is thought to interact with the conserved Asp183 residue important for Ub/Ubl adenylation in E1s, and a unique Glu209 residue in Uba5. It was found that the identity of the metal was important for inhibition.[77]

A group screening for inhibitors of the UPS isolated panepophenanthrin (**1.22**, Figure 1-16), a natural product from the fermented broth of the mushroom strain *Panus rudis*.[78] Its structure was determined by NMR and x-ray crystallographic analysis and was shown to inhibit the formation of the E1-Ub complex. However, no significant inhibitory activity was observed within cells up to 50 µgmL^{-1}. Later, a different group developed an enantio- and diastereoselective total synthesis of (+)-panepophenanthrin. After an 8-step synthesis to reach **1.24**, starting from **1.23**, a neat Diels-Alder reaction could be performed to reach the dimer, which undergoes intramolecular hemiacetal formation to provide panepophenanthrin.[79] Several derivatives varying the 2-propanol group could be produced (**1.29a-c**), each of which also providing inhibitory activity against E1.[79] These were accessed in a slightly different manner by coupling **1.25** with **1.26a-c** which forms **1.27a-c** and can be silyl ether deprotected providing **1.28a-c** and subjected to the Diels-Alder reaction to access the desired targets.
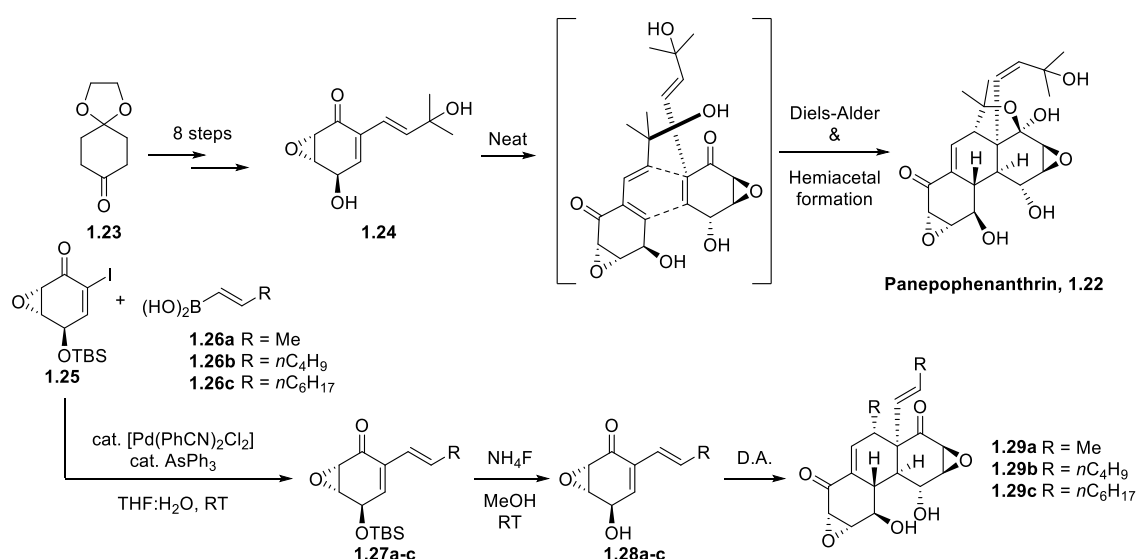
Figure 1-16: Synthesis of Panepophenanthrin and synthetic derivatives as E1 inhibitors.

Himeic acids A-C were isolated from a culture of marine-derived fungus *Aspergillus* sp. isolated from a mussel collected from Toyama bay in the Japan sea, and their structures

determined by spectroscopic analysis (**1.30** – **1.32**, Figure 1-17).[80] Whilst himeic acids B and C were unable to inhibit E1-Ub formation even at 100 µM, himeic acid A was able to produce a 65% inhibition at 50 µM. It was postulated that inhibition stems from thiol attack from the E1 active site Cys onto the imide carbon highlighted, even though Himeic acid C with the same imide system did not provide inhibition.

himeic acid A, 1.30　　　**himeic acid B, 1.31**　　　**himeic acid C, 1.32**
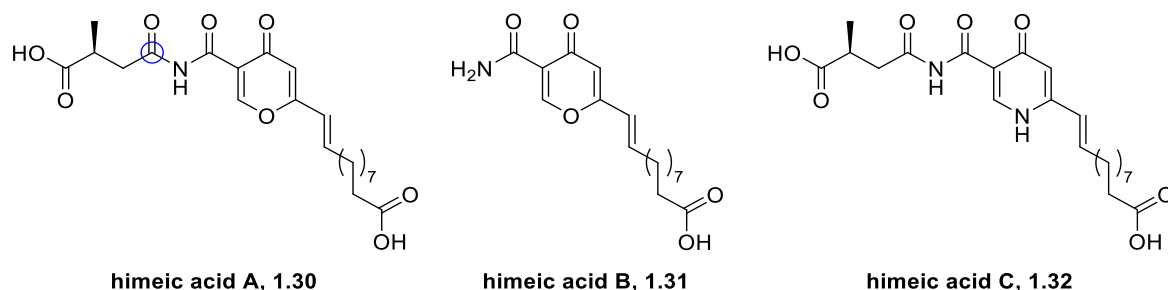
Figure 1-17: Himeic acids A-C.

Largazole (**1.33**, Figure 1-18), originally isolated from an extract of *Symploca* spp. from Key largo, Florida *via* a bioassay-guided fractionation, displayed remarkable antiproliferative activity.[81] Later, largazole and several derivatives (in particular largazole ketone **1.34** and largazole ester **1.35**, among others tested) were identified as able to inhibit ubiquitin conjugation to p27[Kip1] and TRF1 *in vitro*.[82] This was ultimately traced to their ability to inhibit selectively UAE (E1).

Largazole, 1.33　　　**Largazole Ketone, 1.34**　　　**Largazole Ester, 1.35**

Seco-**Largazole, 1.36**　　　**Largazole macrocycle, 1.37**　　　**Largazole thiol, 1.38**
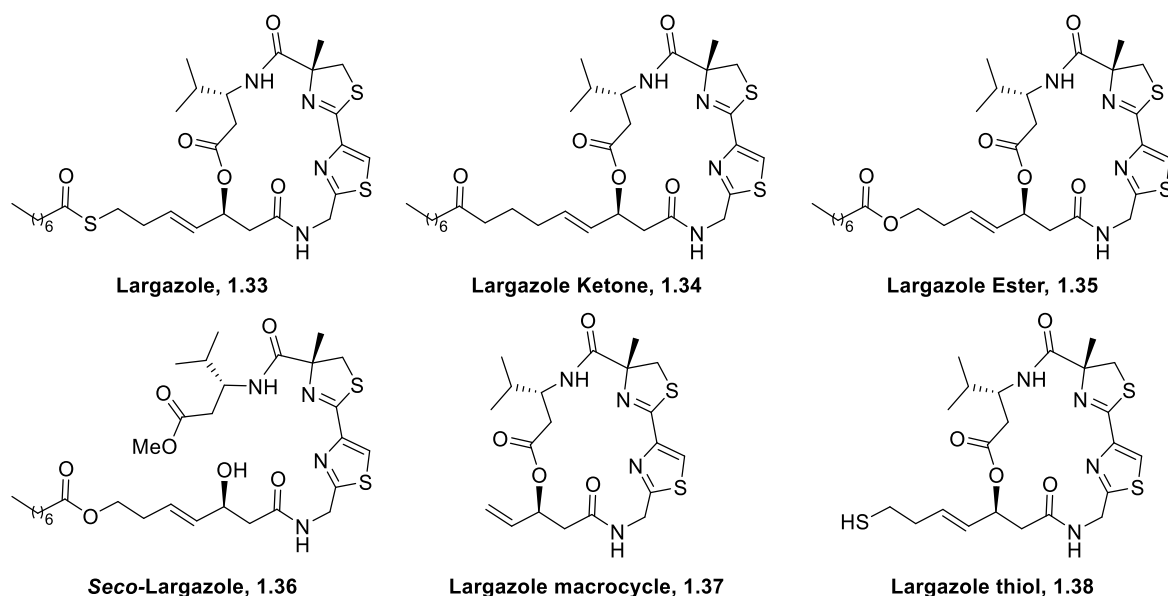
Figure 1-18: Largazole and investigated derivatives thereof.

Another group isolated a series of novel indole alkaloids from a marine sponge *hyrtios reticulatus*, the hyrtioreticulins A-E (**1.39** – **1.43**, Figure 1-19). On the basis of spectroscopic analysis, the researchers identified hyrtioreticulins A, B and E as tetrahydro-β-carboline alkaloids, while hyrtioreticulins C & D were a new azepinoindole-

21

type alkaloids. hyrtioreticulins A and B were found to inhibit E1 with IC_{50} values of 2.4 and 35 μM respectively. By comparison, other naturally occurring E1 inhibitors including himeic acid A, panepophenanthrin (IC_{50} = 40 μM) and largazole (IC_{50} = 50 μgmL^{-1}) provided weaker IC_{50} values.
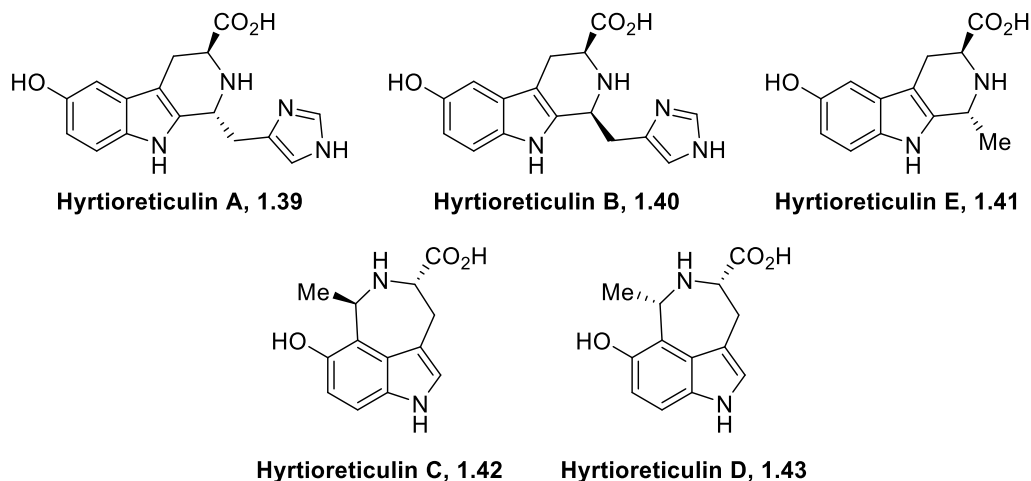
Figure 1-19: Elucidated structures of isolated Hyrtioreticulins A-D.

Lastly, PYR-41 (**1.44**, Figure 1-20) was identified from a high-throughput screening as an E1 inhibitor.[83] Small molecules which provided hits were tested for inhibition of E1-Ub thioester bond formation as well as E2 enzymes. PYR-41 was identified as selective for E1 with an IC_{50} of <10 μmolL^{-1} and has present a Michael acceptor that suggests that it could be open to conjugate addition by the reactive catalytic cysteine thiol. PYR-41 showed seemingly partial inhibition of E3 HECT domains *in vitro.*, although whether actual inhibition was occurring was not further investigated.
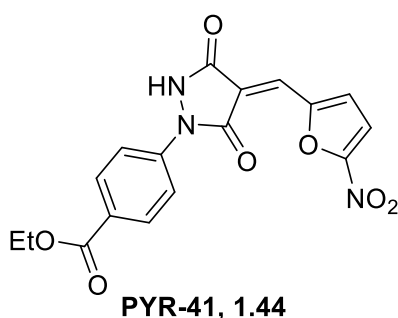
Figure 1-20: Structure of PYR-41 as an E1 inhibitor.

1.5.3 E2 inhibitors

A rather small number of compounds have been identified which inhibit E2 enzymes. The earliest, leucettamol A, isolated from the marine sponge *Leucetta* aff. *microorhaphis,* is

able to inhibit Ubc13-Uev1A interaction and thereby Ubc13 activity due to its requirement to form heterodimeric complex to exhibit activity (**1.45** and **1.46**, Figure 1-21).[84] It was initially thought this molecule could be a lead for anticancer agents that upregulate the activity of the tumour suppressor P53 protein as leucettamol A had an IC_{50} of 50 µgmL^{-1}. Hydrogenation to its saturated derivative improved the IC_{50} to 4 µgmL^{-1}.

Leucettamol A, 1.45

Hydrogenated Leucettamol A, 1.46
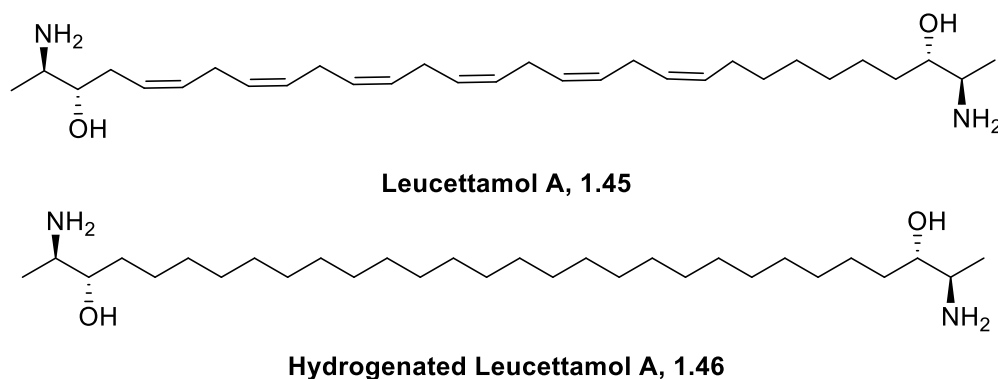
Figure 1-21: The structures of Leucettamol A and its hydrogenated variant.

The E2 enzyme hCdc34 catalyses the ubiquitination of hundreds of proteins in conjugation with the cullin-RING (CRL) superfamily of E3 enzymes. A small-molecule, called CC0651, was identified as being able to selectively inhibit hCdc34 E2 enzyme (**1.47**, Figure 1-22).[85] Further investigation revealed that CC0651 inserts into a cryptic binding pocket removed from the catalytic site, causing the displacement of hCdc34 secondary structure which leads to interference with the discharge of Ub to the acceptor lysine residues. Initial SAR revealed that CC0651 may be improved upon by removing non-essential moieties, although no derivative stood out as a more effective inhibitor. A specific inhibitor of Ubc13-Uev1A is NSC-697923, **1.48**, which acts through inhibiting formation of Ubc13-Ub complex.[86] Initial SAR suggested that the nitrofuran group is important for this inhibition. NSC-697923 was able to exhibit a cytotoxic effect on ABC-DLBCL (activated B cell-like Diffuse large B-cell lymphoma) cells, likely through the inhibition of NF-κB and P53 signalling known to be essential for the survival and proliferation of these cells. **1.48** also inhibits the proliferation and survival of GCB-DLBCL (germinal centre B cell-like DLBCL) cells through a mechanism different to inhibition of NF-κB and P53 signalling.
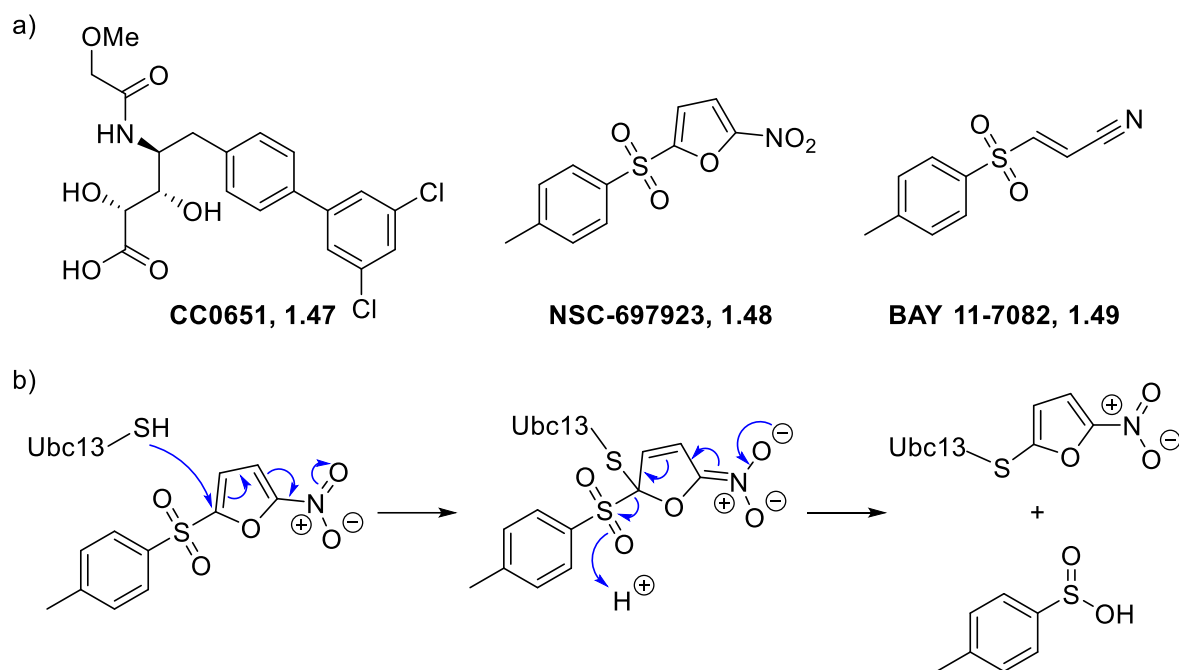
*Figure 1-22: a) Structures of CC0651, NSC-697923 and BAY 11-7082, b) proposed mechanism of action of **1.48**.*

In a later report, it was demonstrated that **1.48** along with another compound, BAY 11-7082 (**1.49**), inhibits Ubc13 *via* covalent attachment through a Michael addition at the Ubc13 active site cysteine.[87] Through the development of a mutant Ubc13 which omitted the cystine, the authors showed that inhibition of cellular DNA damage and NF-κB signalling by NSC-697923 was largely due to said inhibition of Ubc13.

A natural product class of dimeric sterols, manadosterols A and B were also found to inhibit the Ubc13-Uev1A interaction with IC_{50} values of 0.09 and 0.13 µM, respectively (**1.50** & **1.51**, Figure 1-23). [88] These manadosterols were isolated from the marine sponge *Lissodendryx fibrosa* which was collected from Indonesia. This class of natural product is markedly more inhibitory than leucettamol A, which shows only weak inhibitory activity (IC_{50} = 106 µM).

Manadosterol A, 1.50   **Manadosterol B, 1.51**
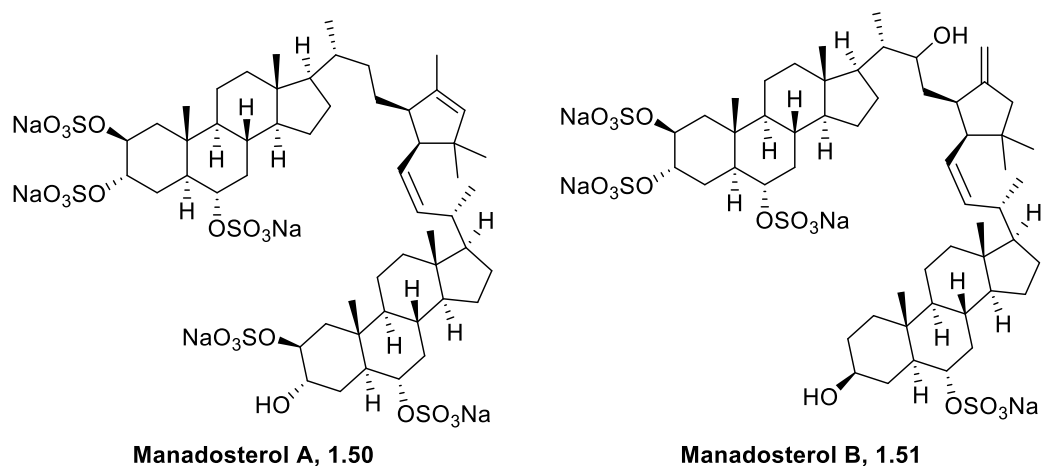
Figure 1-23: The structure of Manadosterols A and B.

A group pursuing the development of a Ubc5Hc inhibitor firstly identified IJ-5 (a herb-derived sesquiterpene lactone, **1.52**) as a new potential anti-rheumatoid arthritis treatment (Figure 1-24). It was demonstrated that this was based upon its ability to inactivate UbcH5C.[89,90] It was also revealed that the mechanism-of-action is that of covalent attack of Cys85 of UbcH5c onto the α-methylene-γ-lactone moiety *via* a conjugate addition. Based upon the information gained, the authors searched other sesquiterpene lactones containing the α-methylene γ-lactone moiety with similar activity to IJ-5.[91] After an extensive medicinal chemistry programme **1.53** was identified as the most robust inhibitor of Ubc5Hc with a K_D = 0.238 µM, higher than that of IJ-5 (2.577 µM).[90] Mass spectrometry and covalent docking simulation both pointed towards the same mechanism-of-action of **1.53** to that of IJ-5.

IJ-5, 1.52   **1.53**
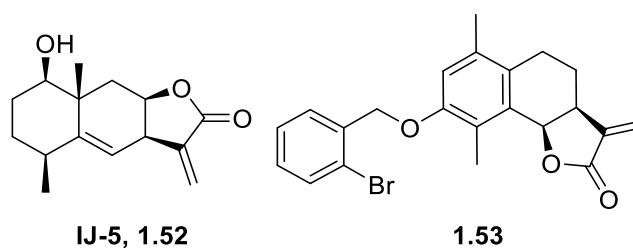
*Figure 1-24: Structures of **1.52** and **1.53** capable of inhibiting Ubc5Hc.*

1.5.4 <u>Targeting E3 Ligases</u>

Targeting E3 ligases is a potentially more selective route with reduced off-target side effects, as compared to the proteasome, E1 and E2 enzymes, less downstream signalling pathways are affected by E3 ligase inhibition due to their higher substrate selectivity.[92] Furthermore, an increasing number of E3 ligases are being identified as crucial regulators

of tumour immune responses.[29,93] Currently, drugs targeting E3 ligases can be divided into four major categories: E3 ligase inhibitors, agonists of E3 ligases, proteolysis targeting chimeras (PROTACs) and molecular glues.[93] As this research project focusses on WWP2, a HECT E3 ligase, drugs targeting the HECT family are discussed. RING and RBR E3 ligase inhibitors and other small molecule strategies have been extensively reviewed.[92,93]

1.5.4.1 <u>E3 ligase Inhibitors</u>

Many small molecules have been investigated for E3 ligase inhibition, with an increasing number being used in clinical trials. However, compared to RING E3 ligases, studies investigating the inhibition of HECT ligases are sparse. Within the HECT family, the neural precursor cell-expressed developmentally downregulated gene 4 (Nedd4) family including Nedd4, ITCH, WWP2, SMURF1 and SMURF2, play key roles in the progression and occurrence of human cancers.[93]

A high throughput screening study attempting to identify ITCH inhibitors demonstrated that clomipramine (**1.54**), an antidepressant drug, was capable of reducing cancer cell growth in breast, prostate, and bladder cancer cell lines (Figure 1-25, a).[94] The authors found that clomipramine inhibits ITCH by blocking ubiquitin transthioesterification in an irreversible manner, and was able to inhibit other HECT E3 ligases, but not RING ligases. Several analogues of Clomipramine were also tested, finding that the chlorine atom is important in providing high inhibitory activity. Based on this and docking simulations, the authors propose an S_NAr-type reaction where the catalytic cysteine reacts at the chlorine position in the ring (Figure 1-25, b).
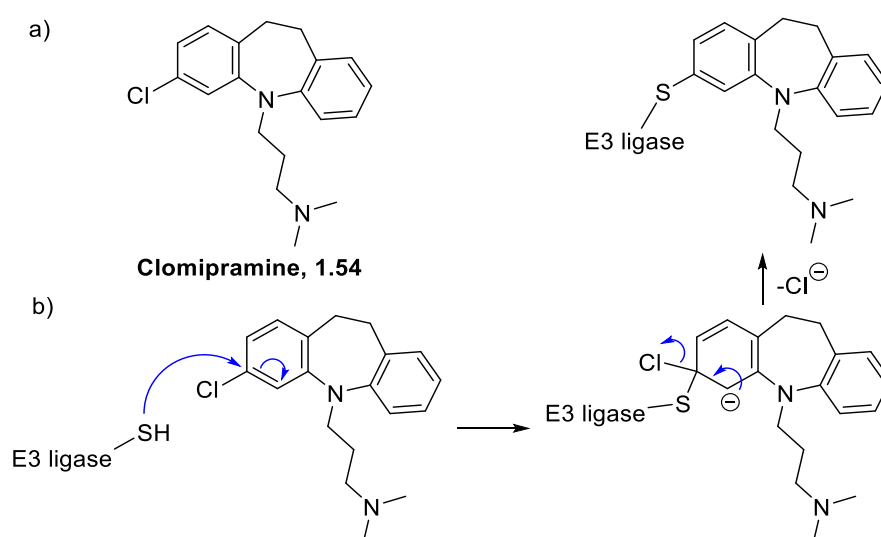
Figure 1-25: Structure of Clomipramine, b) proposed mechanism of action.

Bortezomib has also been identified as capable of downregulating WWP1, SMURF1 and SMURF2 HECT E3 ligases.[95] Bortezomib was shown to reduce the transcription and expression levels of WWP1, SMURF1 and SMURF2 genes in prostate cancer cell lines in a dose-dependent manner, reducing the proliferation of prostate cancer cells.

After developing a method of cell-based high throughput screening to identify modulators of E3 ubiquitin ligases, researchers in 2019 screened for small molecules capable of modulating SMURF1.[96] Having identified a small molecule able to effectively inhibit SMURF1 (HS-152, **1.55**, Figure 1-26), *in vitro* auto-ubiquitination assays demonstrated HS-152 also inhibits SMURF2, indicating it may be an antagonist of the catalytic activity of the HECT domain in SMURF1/2. Additionally, similar to the effects of SMURF1 loss, the authors demonstrate through cell functional assays revealed this inhibition blocks TGFβ-induced EMT in MDCK cells.

HS-152, 1.55
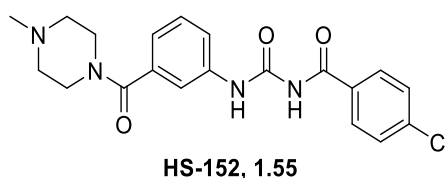
*Figure 1-26: Structure of **1.55**, inhibitor of SMURF1 and SMURF2.*

Nitidine chloride (**1.56**), a benzophenanthridine alkaloid isolated from *Zanthoxylum nitidum* (Roxb.) DC. which is frequently-used in Chinese herbal medicine, has apparent especial beneficial properties in tumours. It has been shown to supress Nedd4 expression at both the mRNA and protein levels in lung cancer cells (Figure 1-27).[97,98] Nitidine chloride suppressed cell viability, migration, invasion and induced apoptosis in lung cancer cells, and mechanistic exploration revealed that nitidine chloride exhibits these effects by reducing the expression of Nedd4. Other natural products, such as paeoniflorin **1.57** and Diosgenin **1.58** which have also been identified as possessing anticancer properties in a variety of cancers have both also been found to downregulate the expression of Nedd4.[99,100]

Nitidine chloride, 1.56   **Paeoniflorin, 1.57**   **Diosgenin, 1.58**
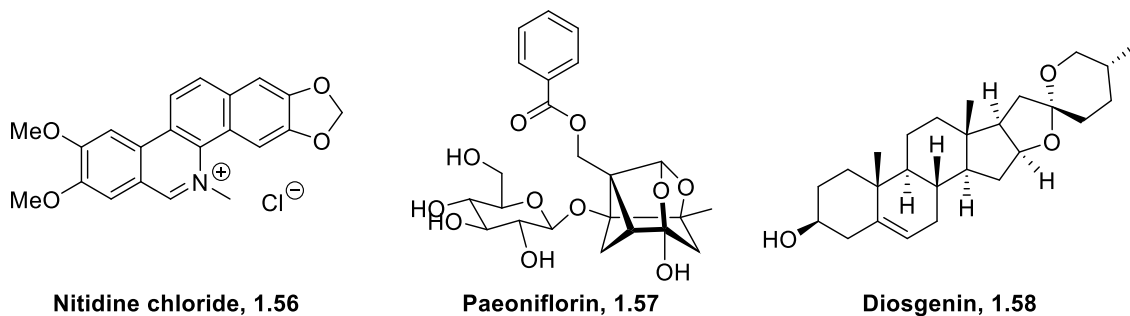
*Figure 1-27: Structures of **1.56**, **1.57** and **1.58** active by downregulating the expression of Nedd4.*

Several bicyclic systems isolated by phage display have been shown to target the E2 binding sites on the HECT domains of SMURF2, Nedd4, Huwe1 and WWP1, acting as inhibitors for these enzymes *in vitro* (**1.59** – **1.62**, Figure 1-28).[101] the four compounds were shown to demonstrate low μM IC$_{50}$ values *via* auto-ubiquitination assay against Smurf2, Nedd4, WWP1 and Ube3c, with **1.59** showing inhibition against Huwe1 too. For Heclin (**1.60**), the most active of those screened, it is thought that upon binding to the HECT domain it changes the normal orientation of the C-lobe making the active site cysteine more prone to oxidation into its reactive sulfinic acid form. Additionally, the conformation change brings it into close proximity with other side chains with which it can react, and in doing so caused inhibition. This model of inhibition is also thought to be likely in cells too, based on the *in vivo* potency observed.
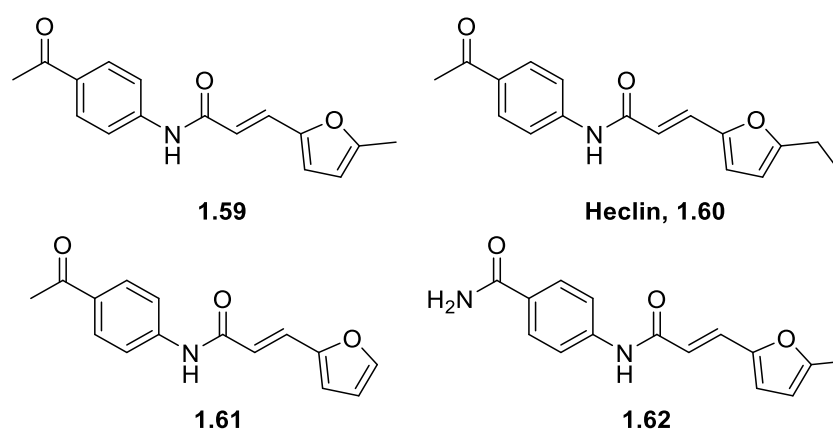
	Smurf2 (μM)	Nedd4 (μM)	WWP1 (μM)	Ube3c (μM)
1.59	7.4	7.1	8.7	8.0
1.60	112	55	88	nd
Heclin, 1.61	6.8	6.3	6.9	nd
1.62	19	24	28	nd

*Figure 1-28: Structures of **1.59** – **1.62** and their associated IC$_{50}$ values against SMURF2, Nedd4, WWP1 and Ube3c.*

To summarise, whilst this is not an exhaustive collection of inhibitors of the proteasome, E1, E2 or E3 HECT ligases, it seems the majority of these compounds are irreversible inhibitors which form covalent bonds with cysteine residues, often at the catalytic cysteine responsible for ubiquitin attachment for the E3 systems.

1.5.4.2 **PROTACs and Variations**

PROTACs have emerged as a potentially novel therapy in cancer treatment, where the PROTAC links a POI to its cognate E3 ligase forming a stable POI/PROTAC/E3 ligase ternary structure which forces ubiquitination and later POI degradation.[93] A PROTAC

structure is made up of a moiety which targets the POI and a moiety which binds the E3 ligase linked together by a linker region.[29] For PROTACs to be effective, the ternary complex formation has to still allow ubiquitination to occur so must not interfere with the catalytic domain of the E3 or the ubiquitin binding site on the POI. Four kinds of E3 ligases are commonly chosen for PROTACs, including VHL, CRBN, IAPs and MDM2.[102]

The first PROTAC to enter clinical trials was ARV-110 (**1.63**, Figure 1-29), an oral small molecule targeting the androgen receptor (AR) for the treatment of metastatic castration-resistant prostate cancer (its E3 ligase target is CRBN) and at present it is currently in phase II trials.[103] The company that developed ARV-110 (Arvinas) also developed ARV-471 (**1.64**) for the treatment of breast cancer, which is now also in phase II clinical trials.[93] ARV-471 targets the ERα POI and CRBN E3 ligase.

ARV-110, 1.63

ARV-471, 1.64
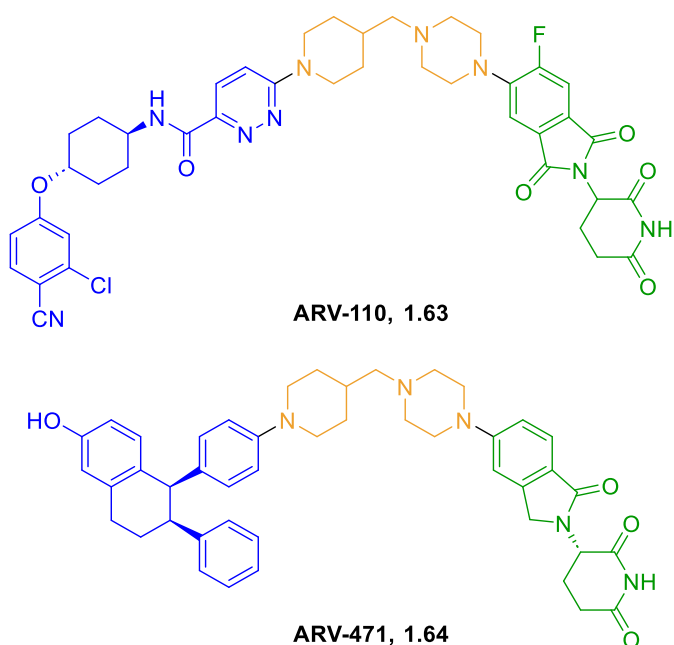
Figure 1-29: Structures of ARV-110 & ARV-471, blue = Androgen receptor binding ligand (ARV-110), ERα binding ligand (ARV-471), orange = link region, green = CRBN recruiter.

Another PROTAC (MZ1, **1.65**, Figure 1-30) that was developed targets the BRD4 protein, which is part of the BET protein family involved in regulating cancer and inflammatory processes, and its E3 degrader, the Von Hippel–Lindau (VHL) protein is also targeted. Researchers discovered that the ternary complex formed between BRD4-MZ1-VHL can mediate the degradation of not only BRD4 but also BRD2 and BRD3.[104] Further development resulted in a bivalent PROTAC AT1 (**1.66**), which selectively targets BRD4, capable of binding two BRD4 proteins resulting in a more efficient degradation ability.[105,106] Lastly, a trivalent PROTAC SIM1 (**1.67**) has been more recently developed with an even higher degradation efficacy and higher levels of anticancer activity.[107]

MZ1, 1.65

AT1, 1.66

SIM1, 1.67
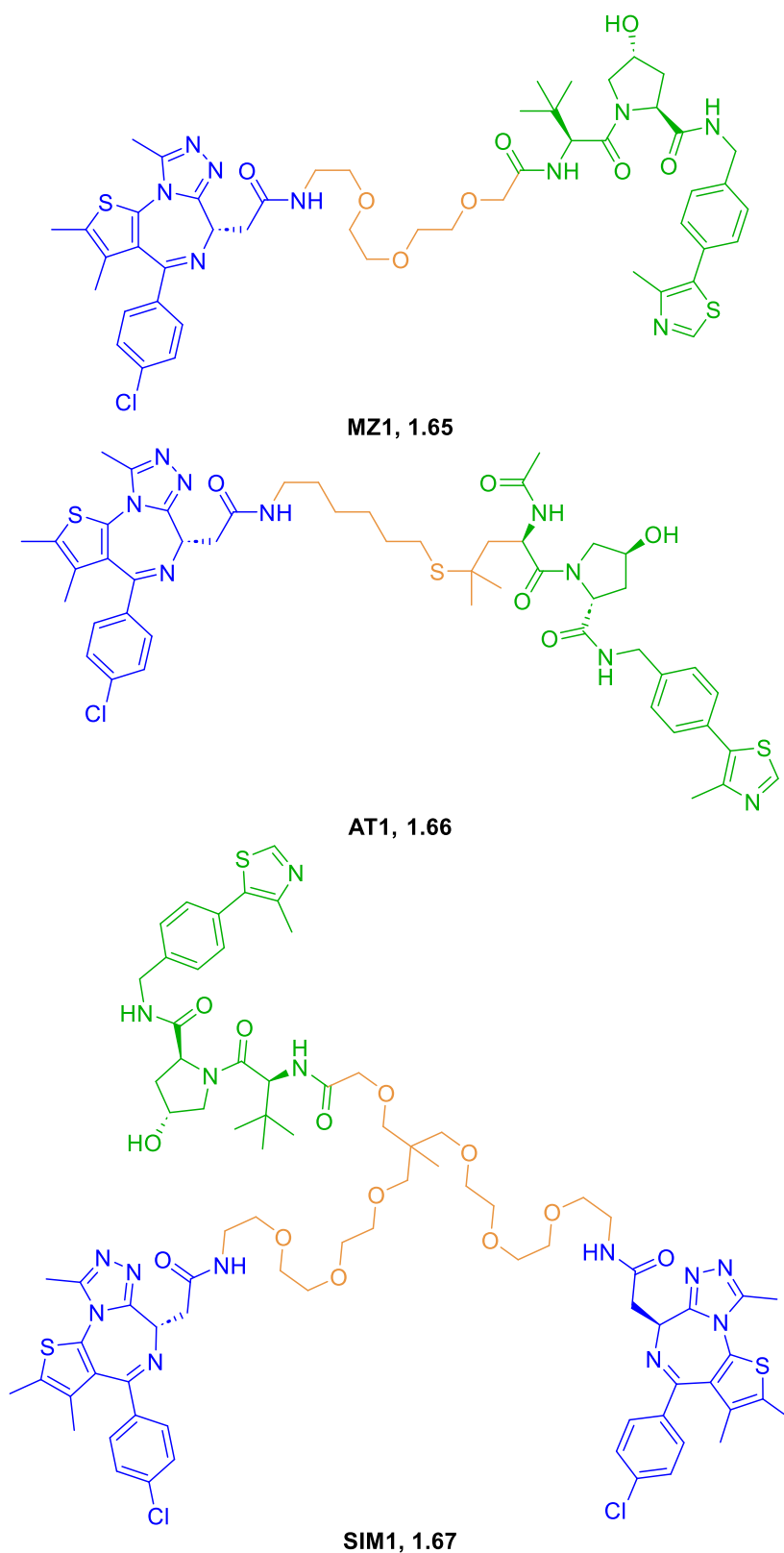
Figure 1-30: Structures of MZ1, AT1 and SIM1, Blue = JQ1, BRD4 inhibitor, orange = linker region, green = VHL inhibitor moiety.

The E3 ligase MDM2 is known to be inhibited by the nutlin class of small-molecule inhibitors. Development of the nutlin-containing PROTAC A1874 (**1.68**, Figure 1-31) could

simultaneously bind MDM2 and BRD4, propagating BRD4 degradation, c-Myc suppression and P53 stabilisation within HCT116 and A375 melanoma cells leading to reduced viability.[108] Compared to JQ1 (**1.69**) alone, A1874 was able to mediate greater c-myc suppression, most-likely owing to the greater impact of BRD4 degradation than inhibition. Additionally, A1874 could exhibit similar inhibition of MDM2 compared with Idasanutlin (**1.70**) leading to similar P53-stabilisation activity. Nutlin-containing PROTACs could be valuable cancer therapeutics which combine both the inhibition and degradation of proto-oncogene (BRD4)/oncogene (c-myc) as well as the upregulation of a P53.
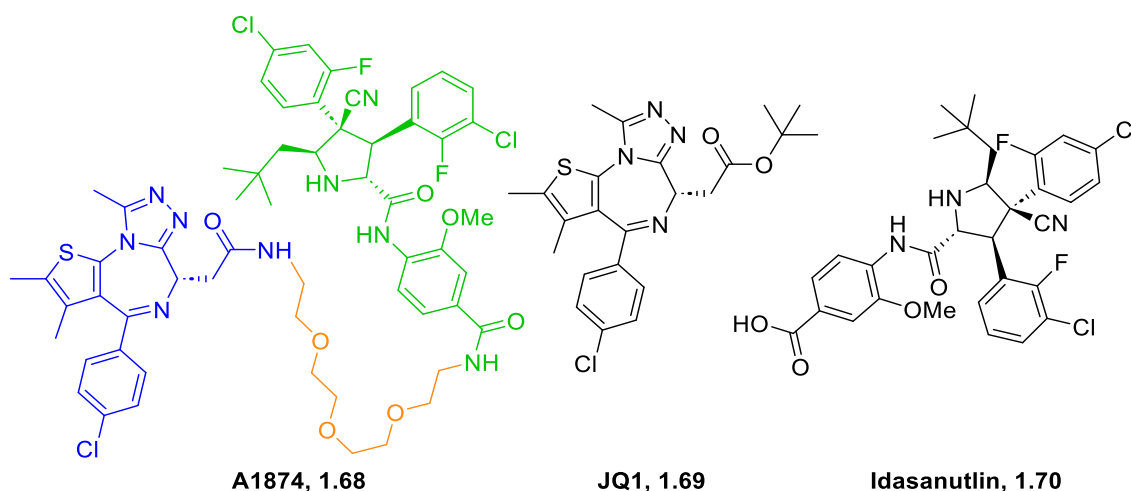
Figure 1-31: Structures of A1874, JQ1 and Idasanutlin. Blue = JQ1, BRD4 inhibitor moiety, orange = linker region, green = MDM2 binder moiety.

Another nutlin-derived PROTAC capable of recruiting both AR and MDM2 is able to provide AR degradation *via* the proteasome.[109] This was achieved by combining a non-steroidal androgen receptor ligand (SARM) with the known MDM2 ligand (-)-nutlin-3 (**1.71**, Figure 1-32). A decrease in AR levels was observed when HeLa cells were treated with this PROTAC at 10 μM for 7 h. It was shown that this degradation was proteasome-dependent and when pretreated with the proteasome inhibitor epoxomicin (10 μM) AR depletion was mitigated. This has implications for the potential of treatments of various cancers with increased levels of AR.

SARM-MDM2 Heterobifunctional PROTAC, 1.71
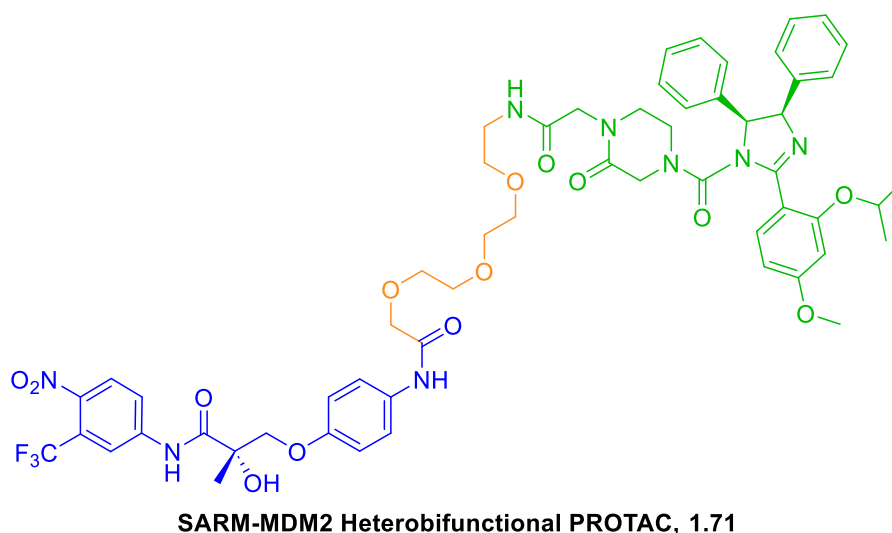
Figure 1-32: Structure of SARM-MDM2 PROTAC, blue = SARM ligand, orange = linker region green = MDM2 binder region.

SNIPERs (Specific and Non-genetic IAP-dependent Protein ERasers) are a variation of PROTACs that have been developed upon the understanding that methyl bestatin (MeBS) interacts with cIAP1 and induces its autoubiquitination. Adding a ligand for the POI (cellular retinoic acid binding proteins, CRABPs) to a methyl residue of MeBS forms a SNIPER which simultaneously induces degradation of the E3 (cIAP1) and the POI (CRABPs, **1.72**, Figure 1-33).[110] SNIPERs are specific for the IAP E3 ligase and several other POIs have been targeted, including ERα, AR, TACC3 and BRD4.[111]

cIAP1- CRABPs SNIPER, 1.72
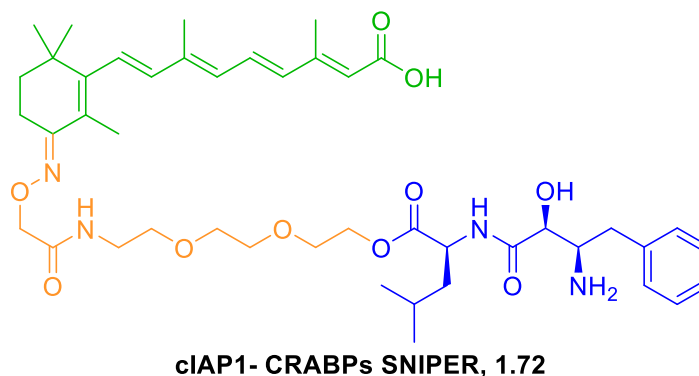
Figure 1-33: SNIPER structure targeting cIAP1 (E3 ligase) and CRABPs POI. Green = CRAPBPs ligand, orange = linker region, blue = methyl bestatin.

Homo-PROTACs hijack two E3 ligases to induce their self-degradation. The first of these was based on the VHL protein which induced dimerization and self-degradation (**1.73**, Figure 1-34), with a later example degrading CBRN (**1.74**) and more recently a Homo-PROTAC pitting both VHL and CBRN against each other for degradation (**1.75**).[112–114] After developing a series of Homo-PROTACs bearing both the VHL and CBRN ligands

with a linker region, the researchers found that in general there was a preferential degradation of CBRN *vs.* VHL.

VHL Homo-PROTAC, 1.73

CRBN Homo-PROTAC, 1.74

VHL-CRBN Homo-PROTAC, 1.75
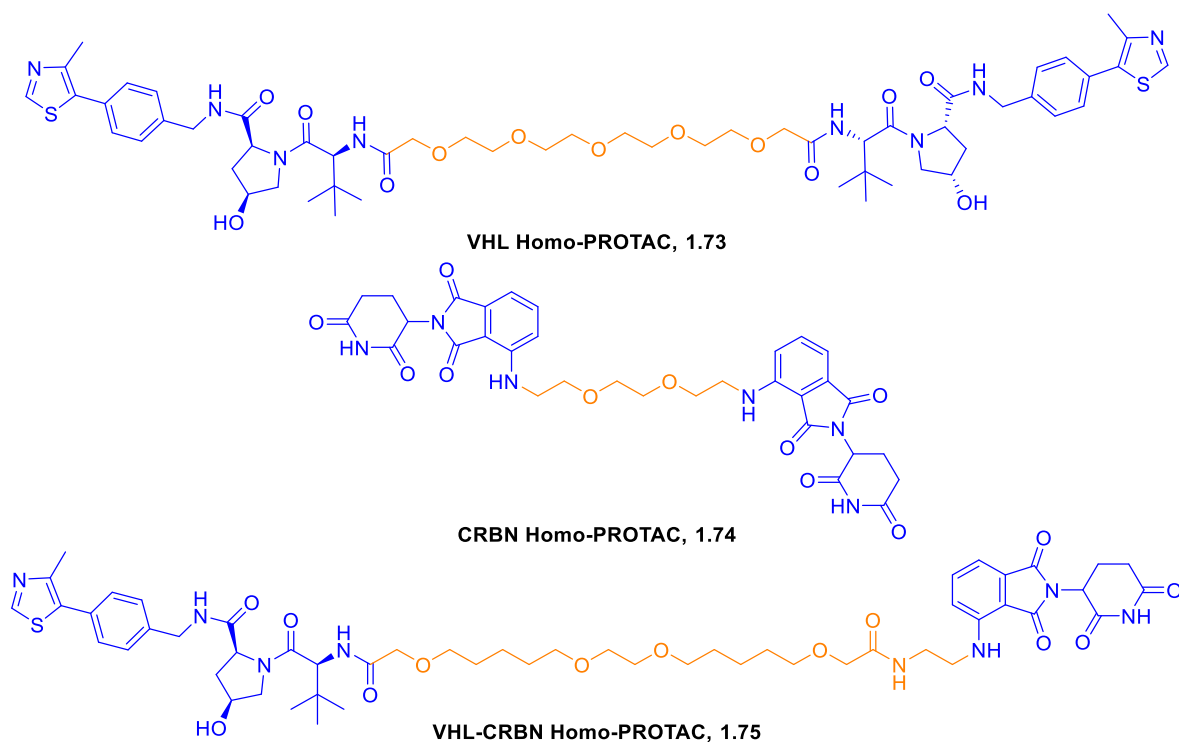
Figure 1-34: Examples of Homo-PROTACs. Blue = E3 recruiters, orange = linker regions.

Recently, PROTACs have been developed to become active *via* light, termed photo or opto-PROTACs, or PHOTACs. They may be employed to directly target tumours *via* photodynamic therapy. One strategy is to employ a photocontrol group (called photocaged PROTACs), which acts to block the formation of complex formation between the PROTAC and its POI and E3 ligase, sometimes by blocking some essential interaction between the ligand and protein (**1.76a**, **1.76b**, Figure 1-35 a).[115, 116] After light irradiation this group can cleave to reveal the active PROTAC molecule. Another strategy is to utilise a reversible activation technology to induce spatiotemporal control of protein degradation by employing a bistable Photo-PROTAC (**1.77a**, **1.77b**, Figure 1-35 b). In the specific example highlighted, the authors took advantage of the critical effects the length of the linker region has on PROTAC activity. By introducing an *ortho*-tetrafluoroazobenzene moiety into the linker region, which can be effectively switched between its *cis* and *trans* isomers, the researchers were able to effectively turn off and on the activity of this PhotoPROTAC.[117] Additionally, variations as to the position of the photo-switchable unit can be made to either restrict binding at the E3 ligase (and the E3 ligand) or the POI (at the POI ligand).[118]

a)

1.76a

hv **Photocaged PROTAC activation**

1.76b

b)

trans, active PROTAC, **1.77a**   530 nm ↕ 415 nm

11 Å

cis, inactive PROTAC, **1.77b**

8 Å
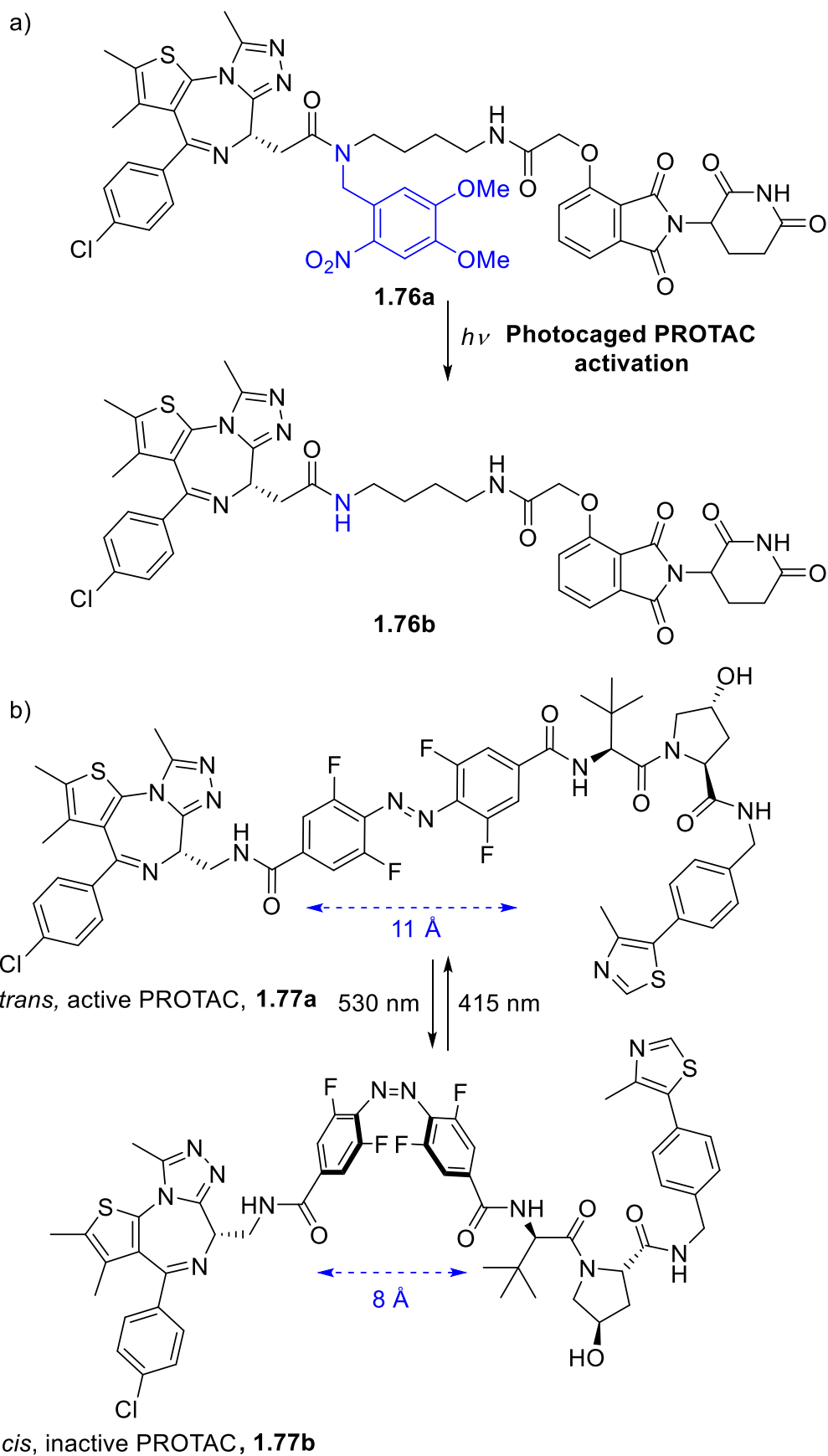
Figure 1-35: a) example of a photocaged PROTAC, b) example of a photo switchable PROTAC.

A final interesting variation is the recruitment of DUBs for protein stabilisation. This has recently been developed where engineered DUBs (enDUBs) facilitate Ubiquitin removal and rescue of POIs. The first enDUB consisted of the catalytic domains of OTUD1 which were fused to a nanobody specific for their target.[119]

Compared to the number of POIs targeted for degradation by PROTACs, the number of E3 ligases targeted is rather small, considering there are over 600 different ligases known.[92] It seems that future directions in PROTAC research is the development of small molecule inhibitors or recruiters of other E3 ligases to enable the degradation of more POIs. Notably, no HECT domain E3 ligase PROTAC has been developed to date.[103] As the mode of ubiquitination is different between the HECT, RING and RBR families of E3 ligases, it would be interesting to understand if functional PROTACs can be designed to target HECT ligases. The opportunistic resurrection of optimized binders of E3 ligases or POIs that failed in clinical trials for the inclusion into PROTACs may provide an alternative avenue into the development of PROTACs.[92]

1.5.5 <u>Literature Investigations on WWP2</u>

Several papers investigating the HECT domain inhibition reference to WWP2 in some way. The first describes etomidate (**1.78,** Figure 1-36), which was shown to inhibit cell proliferation and induce apoptosis in A549 non-small cell lung cancer cells, by downregulating WWP2.[120] It was observed that etomidate, an intravenous anaesthetic was also reported to be tumour suppressive in several types of cancer and the authors sought to elucidate this mechanism-of-action. It was found that etomidate was able to both interact with WWP2 and negatively regulate the expression of WWP2. The actual mechanism of etomidate on WWP2/PTEN on non-small cell lung cancer was not fully investigated. It was also highlighted that the study had limitations, in that only on cell line was used and that multiple cell lines should be utilised for *in vitro* experiments for a more comprehensive and in-depth validation. Another report detailed Nedd4 auto-ubiquitination inhibitors which were found through TR-FRET HTS assay and MS-MS analysis.[121] These were simple compounds capable of covalently binding to Nedd4 Cys182 and Cys867 residues. Mode-of-action studies indicated that the way these compounds inhibit Nedd4 was through blocking the catalytic Cys867 residue. Two of the compounds investigated (**1.79** and **1.80**) were shown to be non-selective for Nedd4, inhibiting both WWP1 and WWP2 as well, although these were not particularly potent (58% inhibition at 200 µM). It was suggested these could be pan-HECT domain inhibitors.

Etomidate, 1.78   **1.79**   **1.80**
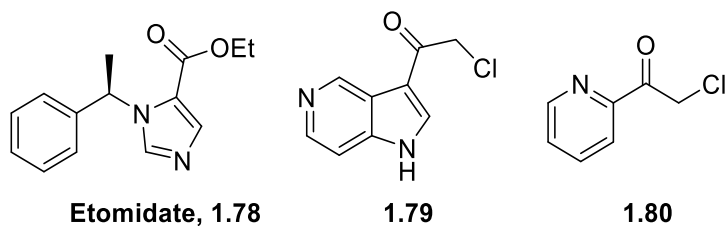
Figure 1-36: Compounds known to interact with WWP2.

Lastly, a computational study published in 2022 to target the HECT E3-E2 binding site identified E2^{UbcH5b}-derived peptide ligands potentially capable of blocking ubiquitin-dependent SARS-CoV2 egression.[122] During the study, the authors characterised the binding pattern of E2^{UbcH5b} to the HECT domains of Nedd4L, WWP1, WWP2, HECW1 and HECW2 through *in silico* analysis, identifying specific E2^{UbcH5b} peptides interacting with the HECT domains. The observation of a similar binding pattern of E2^{UbcH5b} to HECT^{NEDD4L}, HECT^{WWP1} and HECT^{WWP2} identified Trp657 of the HECT^{WWP2} domain was making contacts to Ser94 residue of E2^{Ubc5Hb}. Additionally, the Phe651 residue on HECT^{WWP2} also interacts with E2^{Ubc5Hb} and that E2-to-E3 ubiquitin transfer was mediated by Cys838 HECT^{WWP2} residue. Based off these results, the authors derived two peptides, His55-Phe69 and Asn81-Ala96 from E2^{Ubc5Hb} and tested their *in-silico* binding abilities against all HECT domains investigated within this study. Through Ala-substitutions of several residues, impaired binding validated their interactions with the HECT domain E3 enzymes. Overall, the authors proposed these ligands could selectively target E2-HECT binding and inhibit Ub transfer. This research highlights two peptide derivatives of E2^{Ubc5Hb} that should be investigated further and experimentally validated for HECT-domain inhibition.

Based on the desirability of producing E3 ligase inhibitors as anticancer therapies alone or as part of a PROTAC design, and the evidence surrounding WWP2 as an oncogene with it likely being a promising target, the research group has been investigating the development of WWP2 inhibitors as anticancer leads. The work presented is a continuation of this project, working towards the development of a potent WWP2 small molecule inhibitor.

2 Thesis Aims & Previous work

Information about the IC_{50} assay is provided for reference during later discussions. A description of the high throughput screenings is also provided to understand how the research group chose the hit structures investigated within this thesis. A summary of the previous synthetic work performed within the project is also present, as well as a proposed binding site based upon previous published and unpublished work by the group.[28,37] A proposed mechanism of action is presented based on this prospective binding site. The chapter ends with a description of the overarching aims of the research project and main objectives pursued.

2.1 The Biochemical Assay (IC_{50} Assay)

The details of the bioassay and the results of the work performed for plate coating preparation, biochemical assaying and data processing was provided by our collaborators in the School of Biological Sciences at UEA.

2.1.1 Well Plate Preparation

Initially, 96 well plates were coated with 100 µL of 2% bovine haemoglobin solution containing Na_2CO_3 (0.05 M) and incubated at 4 °C overnight. The wells were washed 3 times with phosphate buffer solution (PBS) before adding the heterobifunctional cross-linker sulfosuccinimidyl 4-(p-maleimidophenyl)butyrate (SSMPB) to a final concentration of 0.1 mM in 50 µL PBS per well and incubating for 1 h. The wells were then washed again thrice with PBS and 50 µL of 10 mM reduced glutathione solution (containing 10 mM Na_2HPO_3, 0.15 M NaCl, 1 M EDTA, pH 7.4) and incubating overnight at RT. Afterwards the plates were washed three times with PBS containing 0.1% Tween (PBST), stored at 4 °C in PBST.

2.1.2 <u>Biochemical Assay Procedure</u>

WWP2 FL (pGEX2T plasmid) pelleted cells were resuspended in 35 mL PBS and sonicated for 8:16 mins, cycle 3 at 50% power on ice. These were then centrifuged (Heraeus Megafuge 1.0R) at 4,300 rcf, 4 °C for 45 min. before 100 μL of cell lysate was added to each well and incubated for 1 hr at RT. E1 (≈ 10ng per well), E2 (≈ 150 ng per well), ubiquitin (100 ng per well) and ATP (1.25 mM) were incubated together in Ub Buffer (25 mM Tris, 100 mM NaCl, 4 mM mgCl$_2$, 50 μM DTT, pH 8.0), at RT for 40 mins.

The plate was washed three times with PBST and tapped dry before 2 μL of the compound was added at desired concentration (in 10% DMSO) followed by 18 μL of reaction mixture, this was incubated for 2 hrs at RT: multiple controls were made up to 20 μL. The plate was washed three times with PBST, before 100 μL of anti-FLAG HRP (Sigma)(1:10000 PBST) was added to each well and incubated for 1 hr at RT. The plate was washed a further three times with PBST, before 100 μL of TMB substrate (Invitrogen) was added to each well and incubated for up to 5 minutes until sufficient colour (blue) change was observed. 100 μL of 1 M HCl was then added to stop the reaction, changing the colour to yellow. The absorbance was read at 450 nm using a Hidex sense microplate reader, with the raw results processed in excel. IC$_{50}$ values were calculated in GraphPad using non-linear regression [inhibitor] vs normalized response.

2.1.3 <u>Discussion</u>

Essentially, the biochemical assay works like this: firstly, to the prepared well-plate WWP2-FL is attached *via* GST, glutathione and SSMPB linkers to bovine haemoglobin, which is attached to the well plate itself (Figure 2-1, a). Afterwards, the attached WWP2-FL is incubated with the UPS machinery (E1, E2, FLAG-tagged ubiquitin, ATP) and the research compound being investigated (Figure 2-1, b). With an inactive compound, which does not bind with WWP2, the UPS machinery will enable the polyubiquitination of WWP2. After a washing step to remove the UPS machinery (and inactive compounds that are not bound), a horseradish peroxidase (HRP) anti-FLAG antibody conjugate is added which recognises and binds to the FLAG-tagged polyubiquitin chain (Figure 2-1, c). After another washing step the bound Horseradish peroxidase is able to catalyse the oxidation of tetramethyl benzidine (TMB) that is added, which changes the colouration of the well plate blue (Figure 2-1, d). After some time HCl solution is added to stop the reaction and turns the well plate yellow. The absorbance of the well plate is then read at 450 nm.
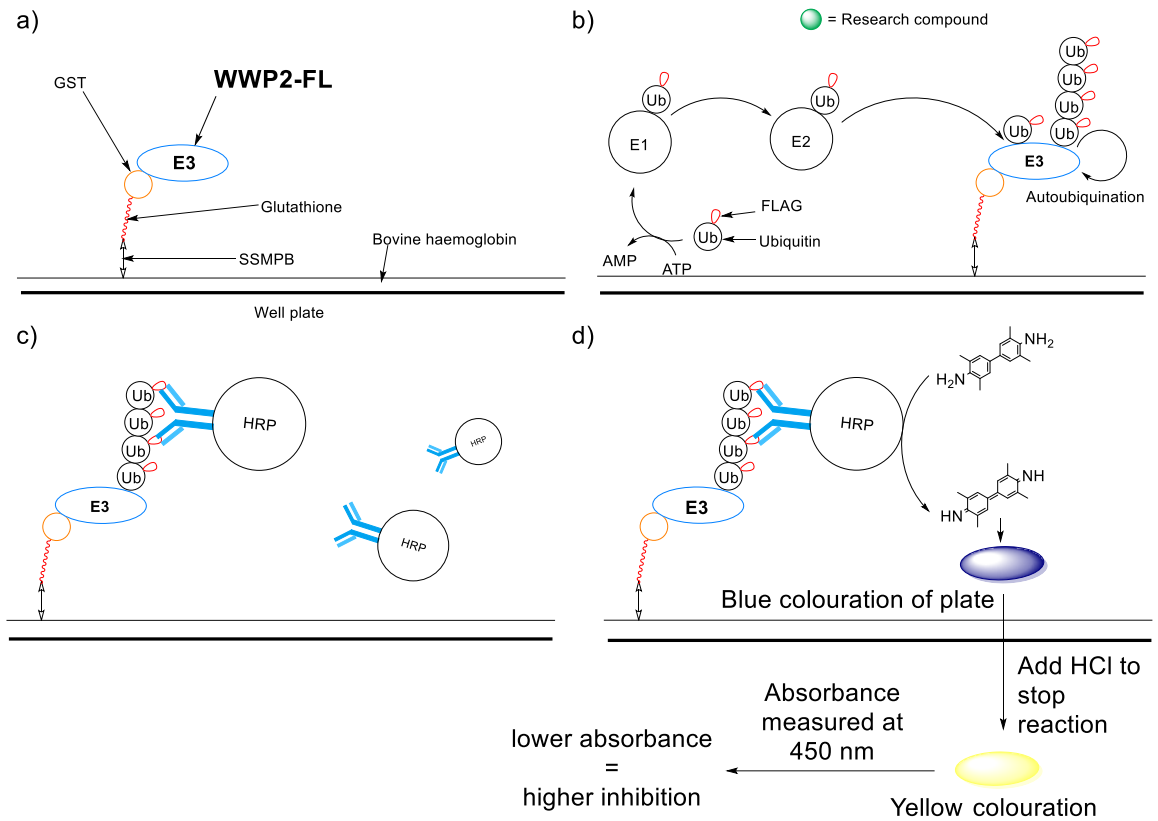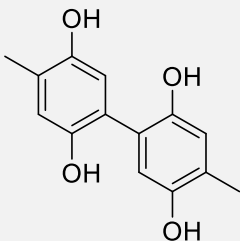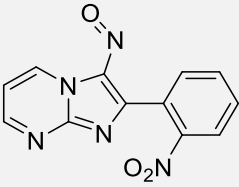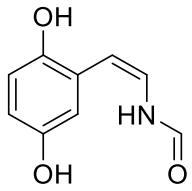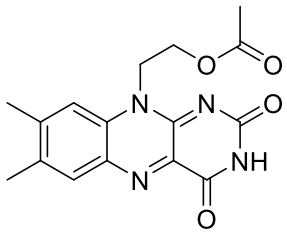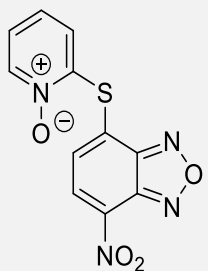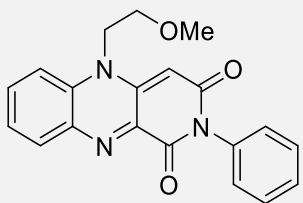
Figure 2-1: Schematic of the biochemical assay.

When the absorbance is higher, the inhibitory effect of the compound is low. This is because the HRP-antibody conjugate was able to bind to the formed polyubiquitin chain because the E3 enzyme was able to undergo polyubiquitination. If the compound is very active for WWP2, limited polyubiquitination occurs, meaning very little HRP-antibody conjugate can bind and oxidise TMB, meaning the later absorbance measured is much lower.

2.2 Hit Identification:

2.2.1 High Throuphut Screening

Several high throuphut screening (HTS) efforts have taken place to identify hits with inhibitory activity against WWP2. Firstly, the initial HTS libraries came from the National Cancer Institute (NCI) body and are the NCI diversity sets V and NCI Approved Oncology Set V. A previous publication from these efforts describes the process of hit identification and the identification of active compounds (Table 2-1). Several were further validated by saturation transfer difference NMR (STD-NMR) and differential epitope mapping (DEEP)-STD NMR.[37] Along with molecular docking studies, the STD-NMR and DEEP-STD NMR data of entry 6 pointed towards a proposed binding site of this ligand in WWP2 (see 2.3 below). Additionally, later research looking at entry 1 identified preliminary co-crystal data which corroborated the previously proposed binding site. All hits were identified form the diversity set, and not the approved oncology set.

Entry	Structure	IC$_{50}$ (µM)	Entry	Structure	IC$_{50}$ (µM)
1		0.38	2		10.28
3		0.77	4		6.51
5		0.83	6		2.29
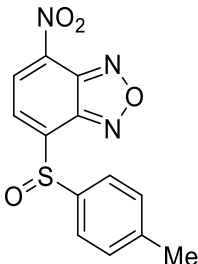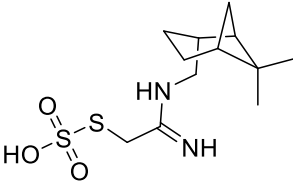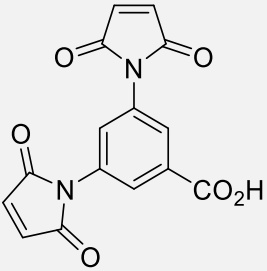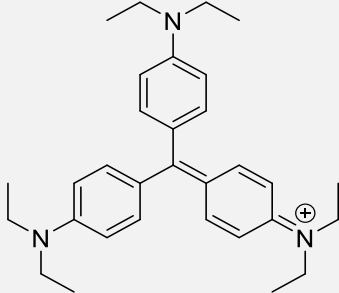
7		1.29	**8**		5.95
9		0.85	**10**		5.39

Table 2-1: Compounds identified from NCI Diversity set V, IC$_{50}$ values are against WWP2.

Both STD-NMR and Carr-Purcell-Meiboom-Gill (CPMG) NMR confirmed that compounds of entries 1, 3, 4, 6, 7 and 8 were showing positive binding results against the HECT domain of WWP2.[37] Whilst the other compounds were triaged due to known PAINs, functional group reactivity, or similarity to known dyes, NSC-288387 (entry 6) and NSC-2805 (entry 1) were taken forward based on both NMR and preliminary cocrystal data by our group.

More recently, another HTS was performed on NCI Diversity set VI which contains 1583 compounds. This HTS was performed against both WWP1 and WWP2, as these have high homology HECT domains. Utilising a differential scanning fluorimetry (DSF) screen against WWP1, 209 compounds showed a significant (3x standard deviation, (SD)) shift from the mean shift in single shot. Further, 26 compounds were confirmed by triplicate as having an average greater than 3x SD within error limits (data not shown). For WWP2, 264 compounds were identified, which was further refined to 33 compounds in the same fashion as above.

For the WWP1 biochemical assay (IC$_{50}$ assay), five compounds from the DSF screenings showed a residual activity (RA) below 50% at 500 µM. Against WWP2, 6 compounds showed an RA < 50% at 500 µM. Within these tests, a bioinformatics screen took place to identify potentially reactive PAINs structures, and these were removed from further investigation. Additionally, compounds of known activity and high RA (40 – 50%) were also removed. This left one compound identified from the WWP2 assay screen and 3 from the WWP1 screen. From these four compounds WWP1 *vs.* WWP2 selectivity tests were

also performed, showing that apart from entry 3 (Table 2-2), all compounds had some degree of activity against both enzymes as detailed below.
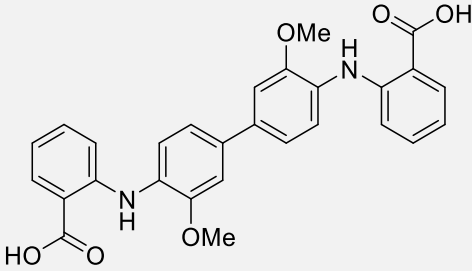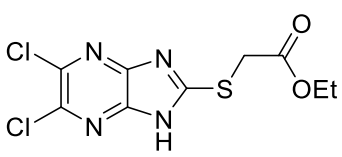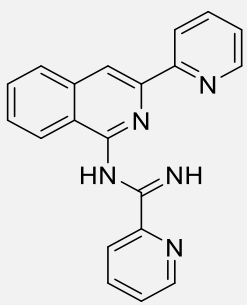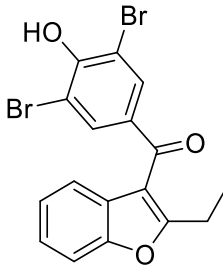
Entry	Structure	WWP1 IC$_{50}$ (µM)	WWP2 IC$_{50}$ (µM)
1		21.7	47.73
2		39.3	145.5
3		19.9	N/A
4		80.1	225.5

Table 2-2: Hit compounds identified from HTS of NCI diversity set VI.

2.2.2 <u>Homolog Inhibitors: Indole-3-carbinol</u>

In addition to these HTS efforts, indole-3-carbinol (I3C, Figure 2-2) was identified as a compound of interest because of its known activity against HECT domains. Several publications describe I3C derivatives of as being active against WWP1 and Nedd4 E3 HECT ligases.[123–127] From the research groups work, I3C is a known binder to WWP2 HECT domain and offers an opportunity to develop this compound further as an inhibitor in a similar fashion to that described in the literature.
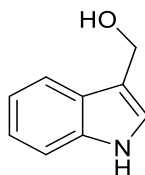
Figure 2-2: Indole-3-carbinol.

Several analogues of I3C have been identified as HECT domain binding against Nedd4-1. Compound **2.1** was identified as the most potent inhibitor in thermal shift assays against Nedd4-1 and an MTT assay against MDA-MD-231 and MCF7 cell lines (17.23 and 19.01 µM, respectively) (Figure 2-3).[123] A range of simple *N*-benzyl and *N*-phenyl I3C analogues were identified as more active than I3C against Nedd4-1 assays with comparable growth inhibition of G361 cells (**2.2** – **2.5**).[126] In a project investigating the production of I3C analogues with improved stability and activity, researchers identified **2.6**, which is resistant to acid-catalysed condensation and exhibits 100-fold higher apoptosis-inducing activity than I3C.[125] Finally, some metabolites of I3C have also been identified as having anticancer properties, such as **2.7** and **2.8** being able to inhibit SMMC-7721 and HepG2 cells in a concentration and time-dependent manner.[124,127]
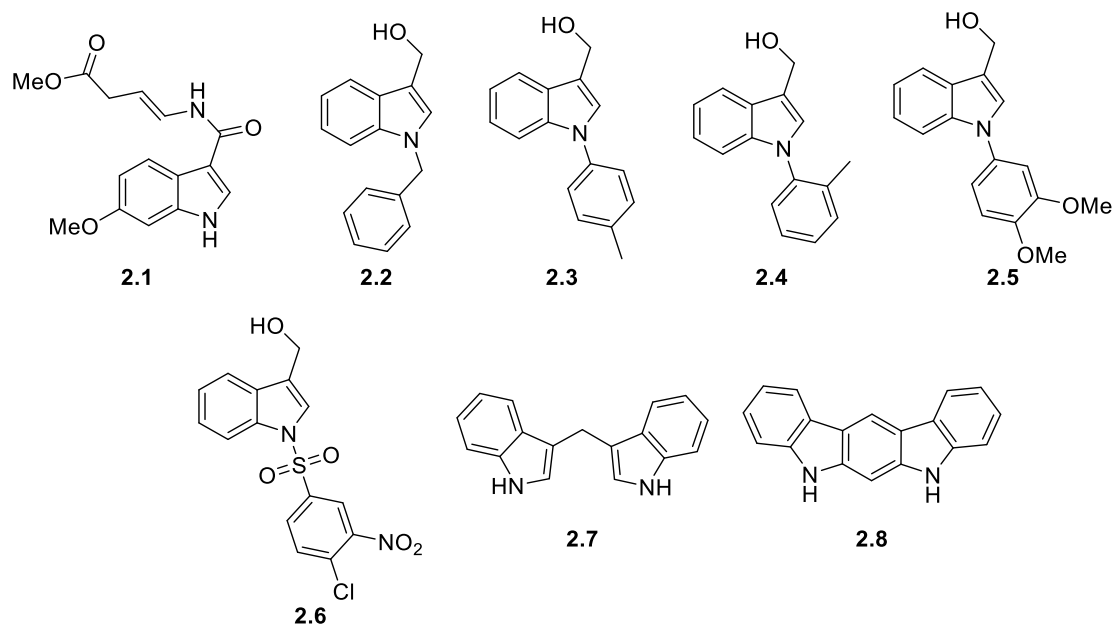
Figure 2-3: Active derivatives of I3C displaying Nedd4-1 activity and anticancer activity.

2.3 <u>Proposed Binding Site</u>

In a previous publication by our group, STD NMR suggested that the isoflavin hit NSC-288387 is recognised most heavily at its *N*-phenyl moiety, with weaker interactions present on the fused phenyl ring and its *N*-alkyl chain (Figure 2-4) upon interaction with HECT[WWP2].[37] These results suggest that NSC-288387 is situated so that the *N*-phenyl moiety (ring B) is further in the HECT domain, with the fused phenyl and *N*-alkyl groups more solvent-exposed. DEEP-STD NMR experiments showed that the amino acid residues interacting with the ligand within the binding site are a Valine methyl group and Arginine β/γ protons. These results aligned remarkably well with the molecular docking ligand pose, with Val806 and Arg803 residues proximal to the fused phenyl system, as shown (Ring A). As a final point, this NMR-validated complex was further supported by molecular dynamic simulations demonstrating the stability of the protein-ligand complex.
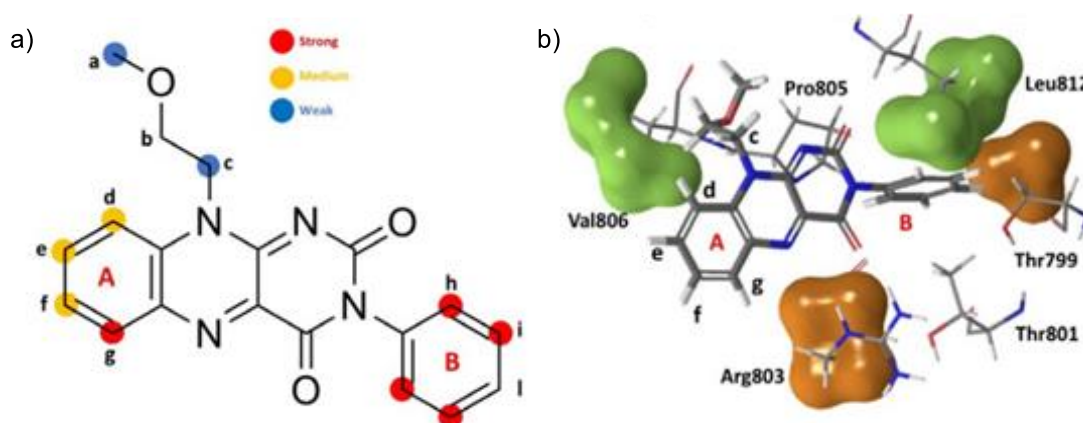
Figure 2-4: a) STD NMR epitope map of NSC-288387, b) ligand pose in proposed binding site highlighting interactions from Val806 and Arg803. Images reproduced from ref.[37]

In the course of a previous PhD student's work, NSC-2805 co-crystal data was obtained which showed the molecule to be bound to the protein in two different sites, one of them being the same proposed binding site of NSC-288387 (Figure 2-5).[28] The other binding site was peripheral and was treated with more caution because it is placed in a gap in the N-lobe which has not been modelled, so the increased electron-density may be a fragment of the polypeptide chain instead of the ligand. Whilst the ligand was fitted into the previously identified binding site from *in silico* modelling and NMR experiments, the quality of the structures left room for ambiguity.[28] This is because there was missing electron-density around a part of a phenyl ring, and no density present around two hydroxy groups in the ring furthest in the binding site.
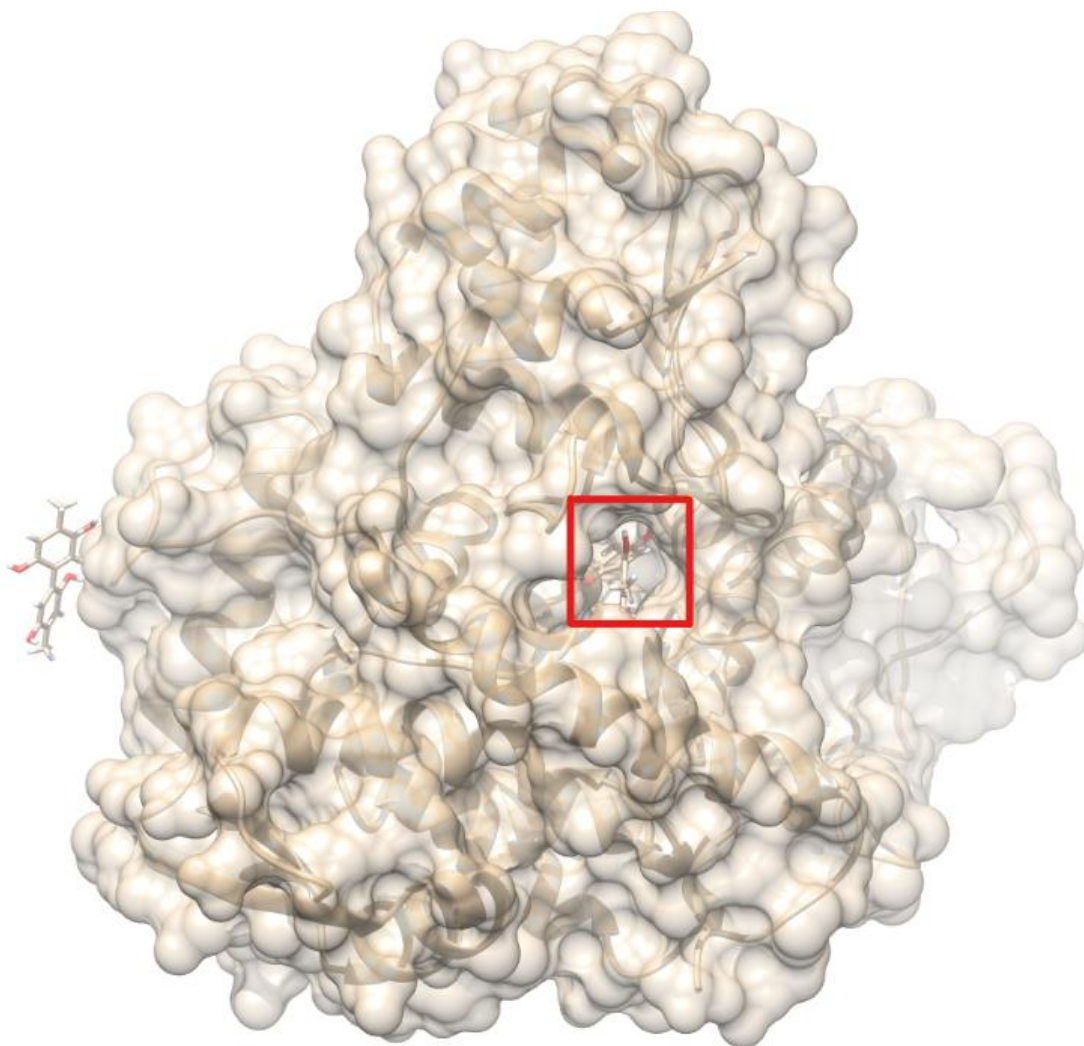
Figure 2-5: Two identified binding sites for NSC-2805, main binding site boxed. Image reproduced from ref.[28]

Two different cocrystal structures were obtained, the first had a resolution of 2.05 Å (Figure 2-6). The ligand is fitted with one ring having two rotamers due to the axial rotation of the symmetrical biphenyl system. This ligand is encompassed by residues Thr801, Cys802, Arg803, Tyr587, Cys588, Asn586, Leu589 and Thr837.[28] π-π interactions are observed with Tyr587 at 4.1 Å, whilst Arg803 forms a cation-π interaction at 5.01 Å. Putative hydrogen bonds with the four hydroxyl units to Tyr587, Thr801, Cys588, Leu589, Tyr581 and Arg803 may exist, but these are at the limits of normal hydrogen bonding distances (2.88 – 3.64 Å).
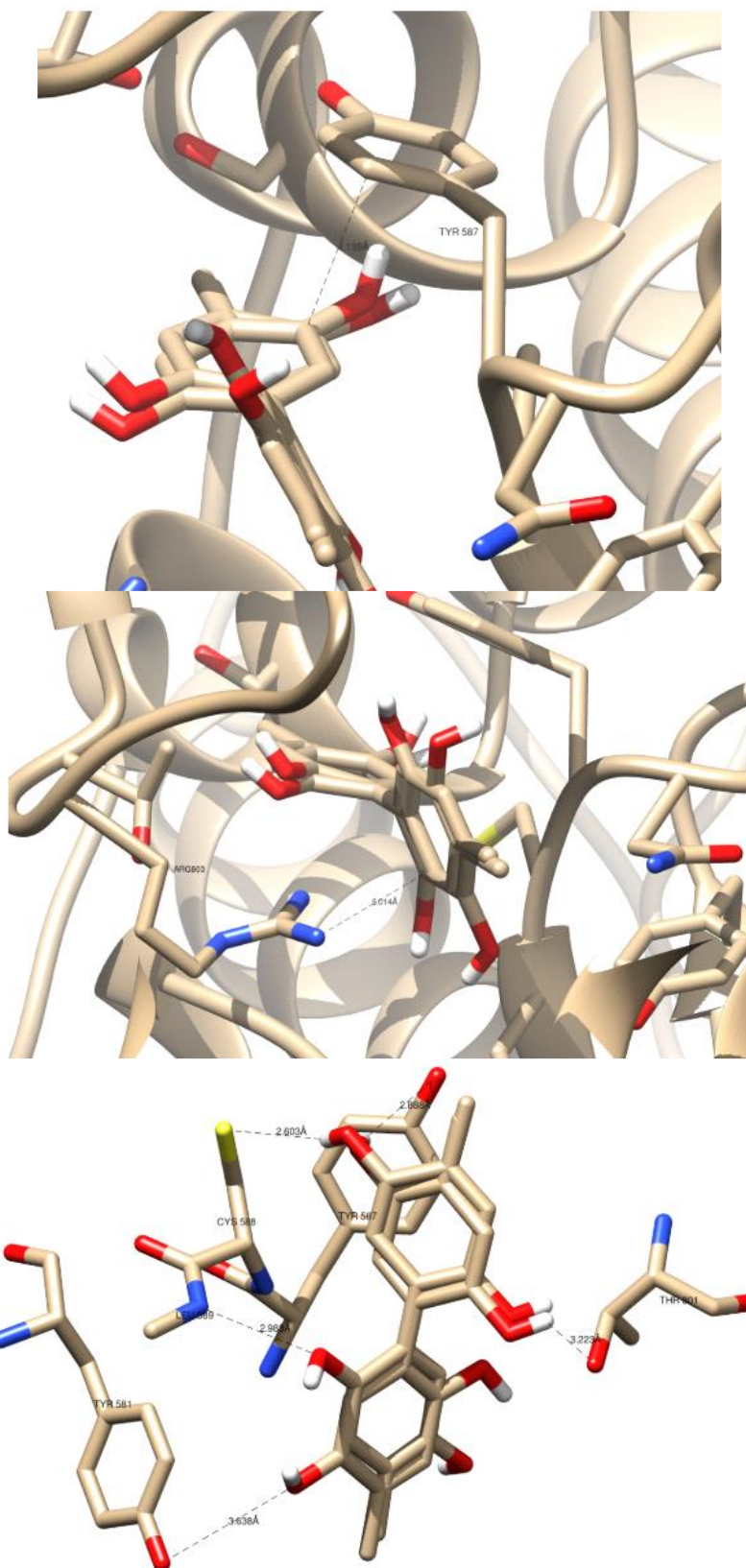
Figure 2-6: Observed π-π interactions (top), cation-π interactions (middle) and hydrogen bonding interactions (bottom) between NSC-2805 and WWP2 HECT. Images reproduced from ref.[28]

The other co-crystal had a higher resolution (2.0 Å), this structure showed a single rotamer which had hydrogen-bond interactions with Tyr587 and Arg803, and three water molecules, with the distances lying in the typical range of 2.5 – 3.0 Å. Additionally, π-π interactions are observed with Tyr587 at 3.8 Å and cation-π interactions at 3.7 Å (Figure 2-7).
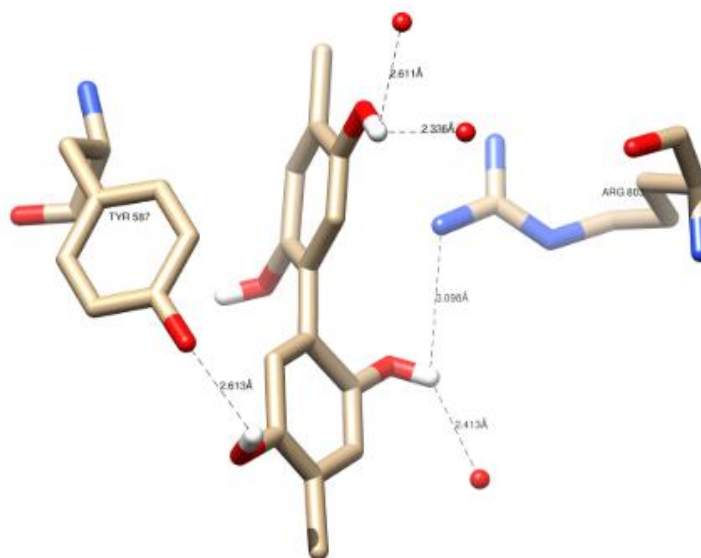
Figure 2-7: Isolated picture of the specific hydrogen bonding interactions of NSC-2805 with Tyr587 and Arg803 residues, as well as water within the protein-ligand co-crystal structure. Image reproduced from ref.[28]

More recent work performed by our biology colleagues also produced a crystal structure of I3C. This crystal structure identified increased electron-density at the same proposed binding site as both NSC-288387 and NSC-2805. The main interactions observed are hydrogen bonding of the carbinol hydroxyl group with Thr627, Ser646 and Tyr587 at a range of 1.6 to 2.3 Å (Figure 2-8). A side-on π-π interaction was also observed between Tyr587 and the indole skeleton at a distance of 3.0 Å. Due to the flexibility of the WW2-2,3-linker section of the protein, the fitting of the indole-3-carbinol into the proposed binding site is based on the increased electron-density within this position.
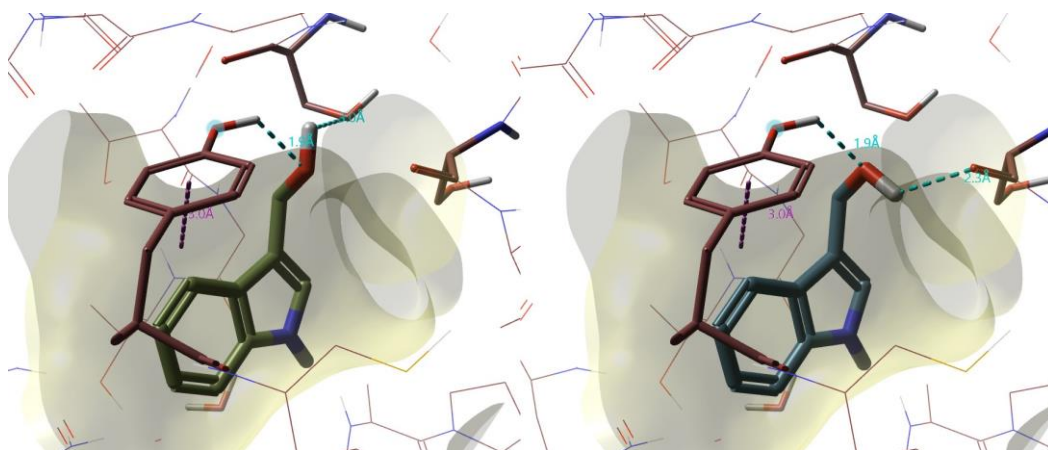
Figure 2-8: Indole-3-carbinol pose within the proposed binding site of the HECT domain.

To summarise, from the previously described work it seems firstly as though a single binding site is where the ligands are interacting favourably enough for NMR experiments and crystal studies to identify it (Figure 2-9). Secondly it seems that both Arg803 and Tyr587 are two important interactors identified at this site, owing to their observed interactions from co-crystal experiments.
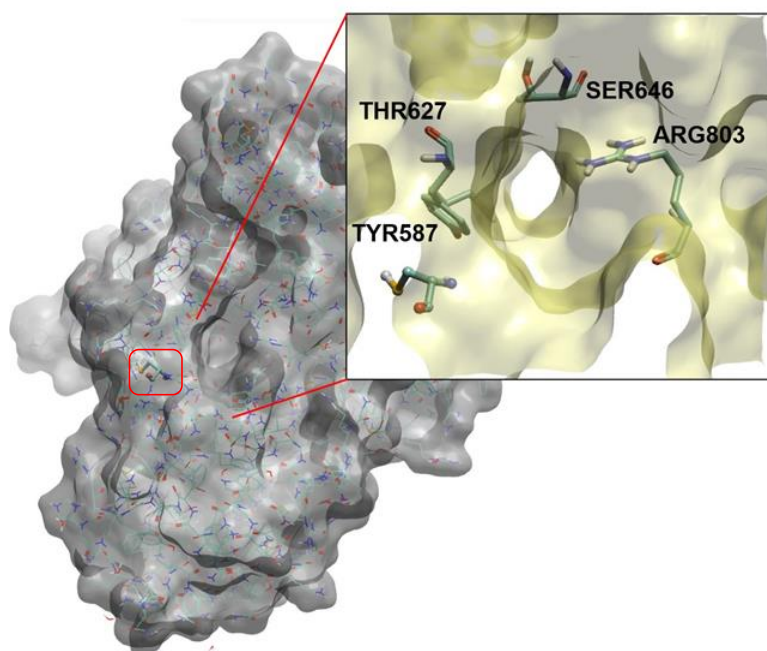
Figure 2-9: Proposed binding site on the WWP2 HECT domain. Zoomed in view highlights important residues as well as the proximally located catalytic Cys838 residue (boxed).

Based upon this information, it was decided that the proposed binding site would be the site to focus the molecular docking work on, and use the important residues from which to design ligands for improved inhibitory activity (Figure 2-9). Finally, it has previously been predicted that this is the most likely binding region due to the mechanistic relevance the

site holds. The site lies close to the hinge loop sitting in the interface of the N and C lobe, therefore targeting this region could potentially hinder the HECT domains intrinsic flexibility required for transferring ubiquitin from the N-lobe bound E2 to the catalytic cysteine in the C-lobe during ubiquitin loading.[28]

2.3.1 <u>Previous synthetic work: NSC-288387 and NSC-2805</u>

Previous synthetic work has been performed on NSC-288387, which identified a derivative with a *N-meta*-chlorophenyl substituent as the most potent inhibitor of WWP2 (Table 2-3, entry 5). These compounds were accessed from the parent compound (entry 3) *via* a Chan-Lam cross coupling with the respective boronic acids. Much of the decisions of which analogues to synthesise was based upon the DEEP-STD NMR investigation against WWP2.[28] Additionally, a CF_3 substituent on the parent isoalloxazine structure was also tested, with only marginal changes to activity taking place.
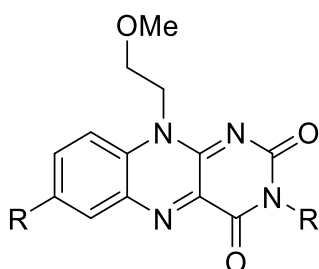
Entry	R	R'	IC_{50} (µM)	Entry	R	R'	IC_{50} (µM)
1	H	(phenyl)	2.29	2	H	(4-OMe-phenyl)	117.8
3	H	H	33.06	4	H	(4-CN-phenyl)	0.31
5	H	(3-Cl-phenyl)	0.23	6	H	(4-CF_3-phenyl)	7.43
7	H	(4-Me-phenyl)	4.26	8	CF_3	H	20.31

9	H		0.55	10	CF$_3$		6.68
11	H		0.55				

Table 2-3: Structures synthesised as part of the investigation into NSC-288387 with the identification of a meta-chlorophenyl derivative as the most potent inhibitor of WWP2.

As an additional point, NSC-288387 was found to be redox active and generate H$_2$O$_2$, although this was thought to occur in negligible amounts to interfere with the assay it was still concerning.[28] It was suggested that for this alloxazine core a scaffold hop to a non-redox active heterocycle should be pursued as this would negate any potential interference from H$_2$O$_2$ production on the measurement of the IC$_{50}$ values. Scaffold hopping can be defined as an extension of traditional bioisosteric replacement, where a replacement or redesign of the core structure moiety whilst retaining pendant groups in a comparable 3D position.[28] Finally, previous work performed on NSC-2805 and at the time when I joined the project, and a discussion of the results of this part is presented in section 5.6.[128]

2.4 <u>Thesis Aims</u>

The overall aim of this continued project was to produce a potent lead inhibitor of WWP2, which would ideally stabilise the perturbed oncogenic signalling pathways overexpressed WWP2 acts upon in cancer cells. This would be achieved by producing analogues with improved binding activity over the identified hits from sections 2.2.1 and 2.2.2 to the proposed HECT domain binding site described in section 2.3. The main objectives therefore are:

1. Resynthesise several chosen hit compounds for co-crystal structure determination and isothermal titration calorimetry analysis against the HECT domain of WWP2.
2. Produce more active analogues of hit compounds by utilising the structural information gained from co-crystal structures and molecular docking studies.
3. Produce a compound with better IC_{50} value compared to the original hit and develop it further to sub-µM potencies.

Specific aims and objectives for hit structures are discussed in more detail in the relevant chapters. Chapter 3 focusses on the resynthesis of hit compounds for co-crystal and isothermal titration calorimetry analysis and introduces the use of molecular docking within this work. Chapters 4, 5 and 6 investigate the synthesis of NSC-217913, both NSC-2805 and NSC-73735, and indole-3-carbinol derivatives respectively for initial structure-activity relationships and derivatives with improved activity (Figure 2-10).

NSC-217913   NSC-2805   NSC-73735 (Redoxal)   Indole-3-carbinol
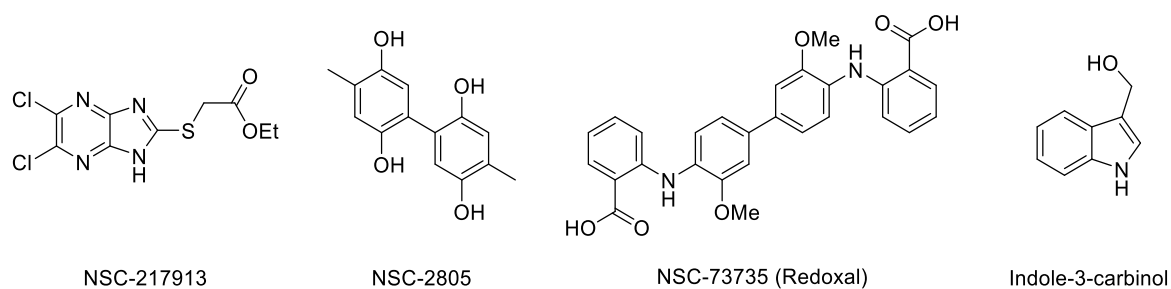
Figure 2-10: Structures of compounds pursued further into analogue synthesis.

3 Efforts Towards a Structural Understanding of Inhibitor Binding

3.1 Introduction

Gaining a co-crystal structure of a hit compound with the target enzyme is an essential requirement for structure-aided drug design.[28] It provides a solid experimental foundation for hit-to-lead optimisation as it provides detailed information on the exact nature of ligand binding to the target enzyme. This chapter describes the synthesis of hit compounds and derivatives for submission to our collaborators in the School of Biology for co-crystal soaking studies.

The utilisation of computational tools such as molecular docking can aid in proposing potential optimisations on hit structures whilst decreasing the number of compounds that have to be synthesised and evaluated experimentally.[129] How molecular docking was used throughout the project is also discussed within this chapter.

3.2 Aims & Objectives

As discussed in section 2.2, several hit molecules have been identified and some have been investigated previously.[28,37] As part of the ongoing aim to develop lead inhibitors of WWP2, gaining a co-crystal structure of hit compounds with WWP2 represents the first objective of this thesis. It was decided within the broader research group to take forward several compounds for resynthesis as part of this objective, which would be of benefit in determining structural modifications required for improved binding against WWP2 (Figure 3-1). Several compounds were identified throughout the project as important enough to attempt to gain a co-crystal structure of and were therefore also targeted for synthesis.
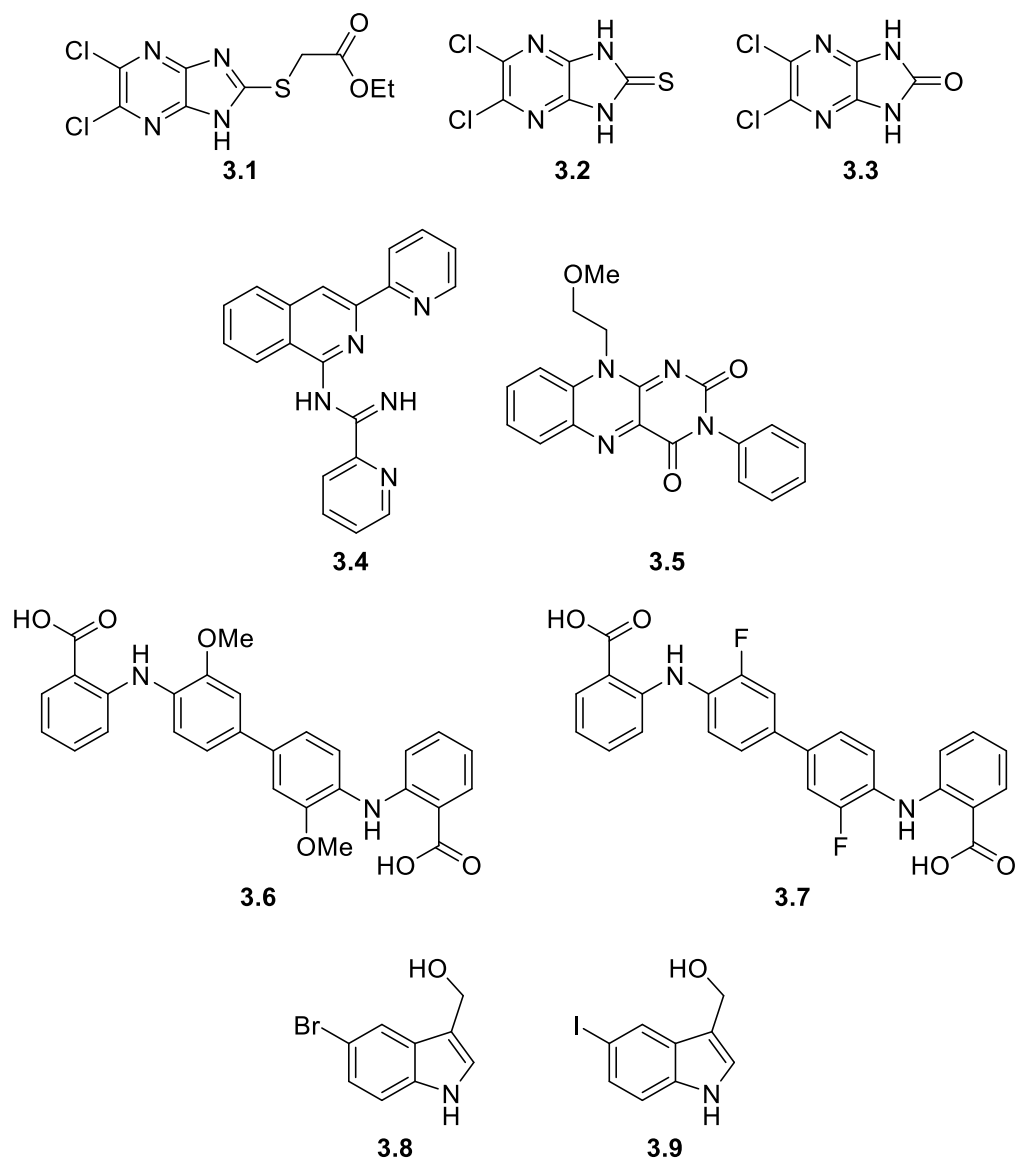
Figure 3-1: Hit compounds targeted for synthesis.

In addition to hit **3.1** (NSC-217913), **3.2** was targeted for co-crystal analysis as this was identified as a molecule with an improved IC_{50} value within this series, and its oxygen analogue **3.3** was also targeted (see section 4.4 for details). Compound **3.4** was also identified as a molecule of interest from HTS, and its synthesis was targeted. Compound **3.5** was identified from NCI diversity set V which had previously been investigated but no crystal structure obtained.[37] For redoxal targets **3.6** and **3.7**, both are to be synthesised to explore how much the methoxy to fluorine substitution affects the binding configuration (see section 5.1.2 for further details). Targets **3.8** and **3.9** were chosen for synthesis as it was hoped that the increased electron density of the halogens within these targets would allow a more accurate orientation of I3C within the binding site (see section 2.3 for discussion). The 5-position was chosen based on the crystal data obtained for I3C and its

proposed orientation (Figure 3-2) shows the phenyl moiety of I3C positioned with its 5-position into void space (highlighted), which was thought would accommodate the halogen atoms without significantly affecting the binding pose. The main objective was to synthesise all compounds in sufficient quantity (at least 100 mg) in high purity for soaking analysis.
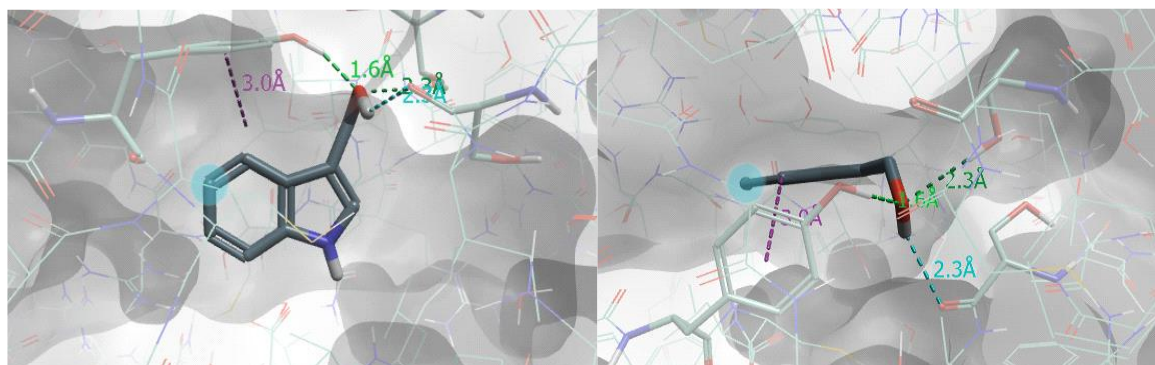
Figure 3-2: Left: side-on view of I3C ligand-protein co-crystal. Right: 90 ° rotation forward of I3C-WWP2 co-crystal side-on view. 5-position highlighted in blue.

In addition to the synthetic work in this chapter attempting to gain a ligand-protein co-crystal structure, molecular docking was utilised to aid in deciding which analogues to target for synthesis and is introduced here. The utilisation of molecular docking represents part of the second main objective of this thesis. Molecular docking is a computational tool that explores the behaviour of small molecules in a target binding site.[130] The orientation of a ligand when bound to a proposed protein binding site is predicted using docking software and is called a ligand pose. Using the sum of shape and electrostatic interactions the used software approximates the potentiality of binding by providing a docking score. As such, it is used as a tool in drug discovery and is essentially routine in several drug discovery phases, including hit-to-lead optimisation. Its implementation into this research project represents an objective of this work.

Cresset Flare® software was chosen over more common molecular docking software such as Autodock® due to its much more intuitive user interface as well as little to no knowledge of coding needed to run the docking calculations to gain poses. A crystal structure of human WWP2 was used from the Protein Data Bank (pdb code = 4y07).[131]

The strategy for utilising Flare® is outlined below. If the hit molecule had a co-crystal structure associated with it the method used to dock the putative derivatives would be validated using redocking and calculating the RMSD of the two poses (Figure 3-3). If not, then the best calculation possible on Flare (very accurate but slow) was performed, followed by visual inspection of the ligand poses and identification of the best ligand pose

using visual inspection parameters. Then, based off how the hit is posed in the binding site, virtual analogues are created with potentially more or stronger binding interactions, and these are docked too. After visual inspection and deciding on the best ligand pose, analogues with the most favourable binding interactions would be taken forward for synthesis and biological evaluation. The experimental results would inform which derivatives have improved activity and further analogue docking based upon these compounds would inform further derivatives to target for synthesis. Successfully implementing this molecular docking flow chart into the decision-making process would mean meeting this objective.
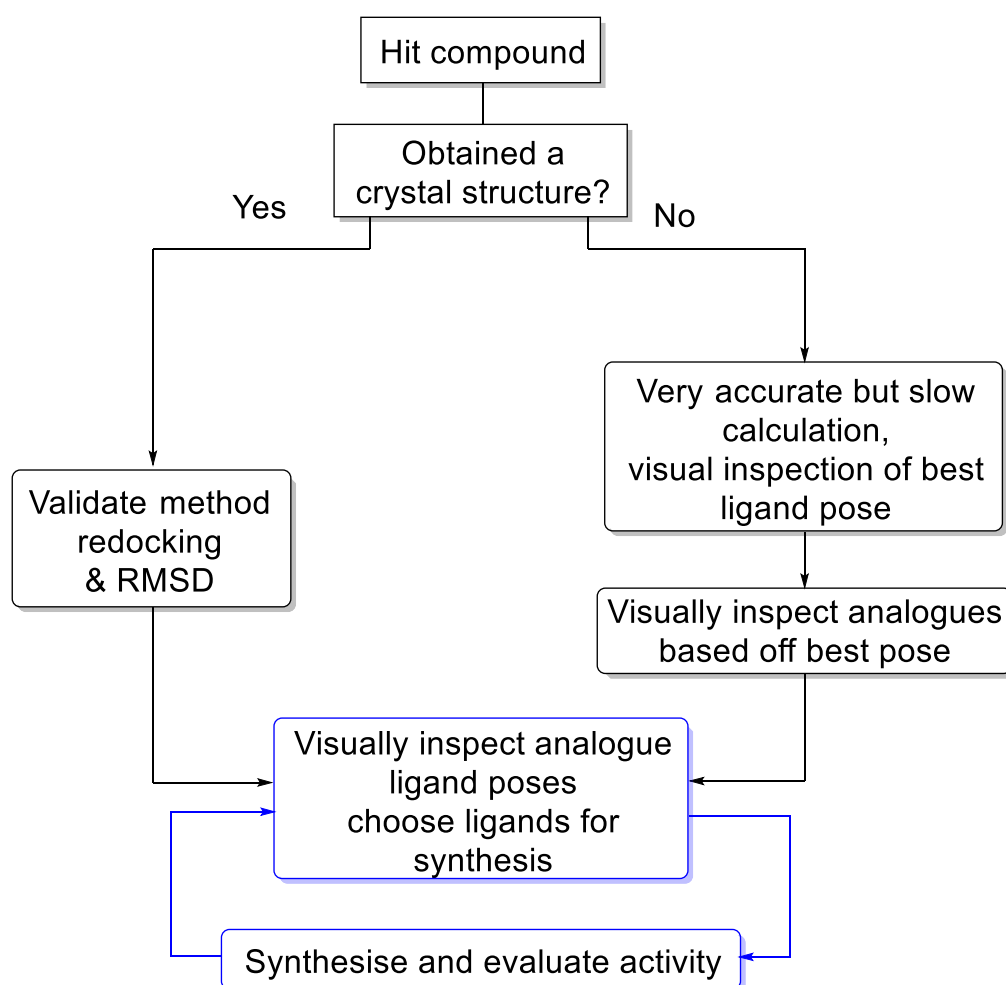
Figure 3-3: Flow chart for the treatment of hits in molecular docking.

3.3 Synthesis I: Synthesis of Hit Compounds and Derivatives

3.3.1 NSC-217913, Thiourea and Urea Derivatives

To avoid restating the synthesis of **3.1**, **3.2** and **3.3**, the production of NSC-217913 **3.1** has been described in section 4.3.1, and from the route described could be synthesised in sufficient quantities for soaking studies (Figure 3-4) for this objective. Additionally, as **3.2** and **3.3** are synthetic intermediates used to produce **3.1**, and their production is also described in section 4.3.1.
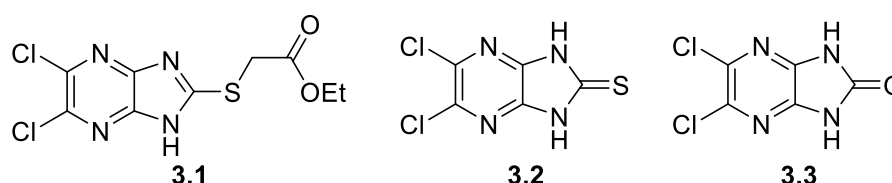
*Figure 3-4: Target compounds **3.1**, **3.2** and **3.3**.*

3.3.2 Isoquinoline – NSC-637578, 3.4

Several medicinal chemistry articles describe the synthesis of 1-amino-3-(2-pyridyl)isoquinoline derivatives for antimycoplasmal activity and 3-(2-pyridyl)isoquinoline derivatives as adenosine A_3 receptor ligands.[132–135] For the synthesis of 1-amino-3-(2-pyridyl)isoquinoline **3.13** as a point of synthetic derivatisation for antimycoplasmal structure-activity relationship studies, Van der Goot developed a method which allowed a 60% yield of **3.13** from 2-methylbenzonitrile **3.11** and 2-cyanopyridine **3.12** (Scheme 3-1) using liquid ammonia and KNH_2. Subsequent reaction of **3.13** with ^nBuLi and a nitrile electrophile provided a series of amidine derivatives **3.14**, including NSC-637578 **3.4**.[135]
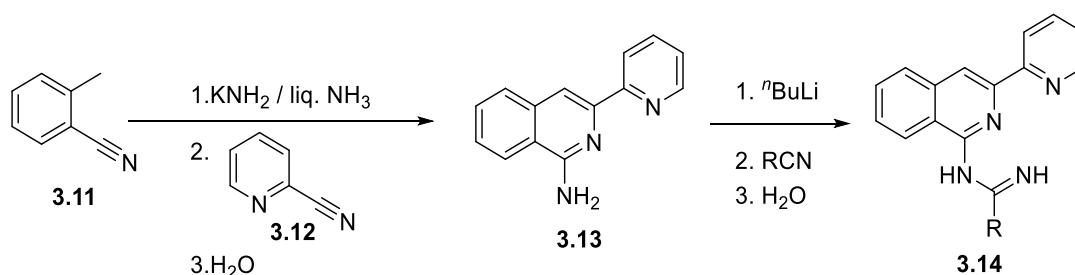
*Scheme 3-1: Previous synthetic scheme for the synthesis of 3-(2-pyridyl)isoquinoline-1-amidine derivatives **3.14**.[133]*

Whilst this route was attractive for potential later derivatisation in the event of a successful co-crystal structure, the use of liquid ammonia and KNH_2 to produce the common intermediate raised safety concerns, which led to the development a different set of

conditions utilising LDA instead (Scheme 3-2). Instead of the desired intermediate **3.13** being isolated, NSC-637578 **3.4** and a by-product **3.15** were isolated in low yield, the structure of **3.15** was fully elucidated by X-ray crystallography, as the location of the 2-methylphenyl group could not be determined by 2D NMR analysis.
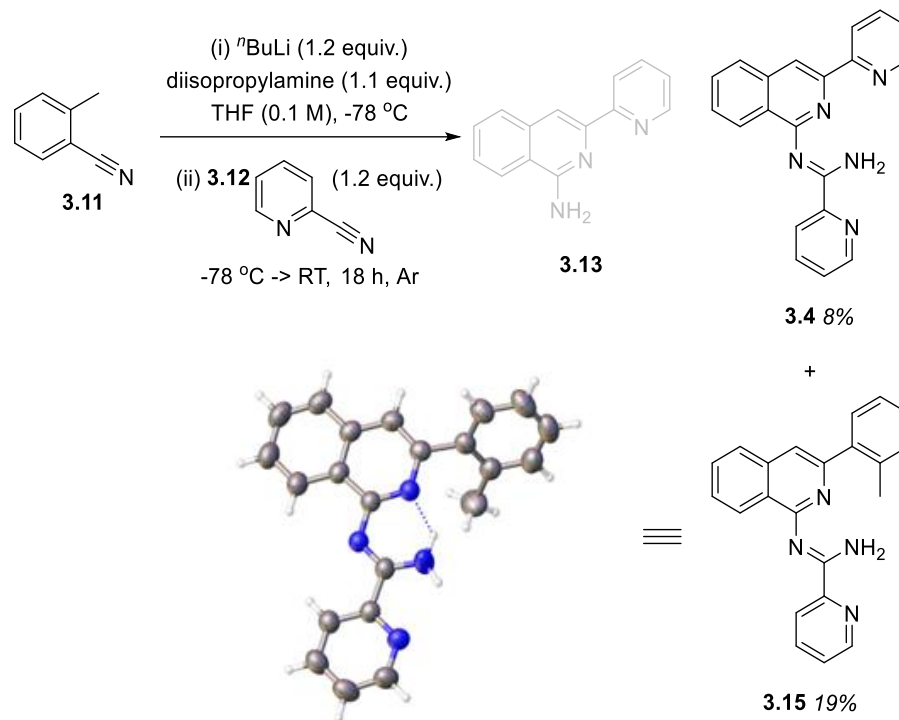
*Scheme 3-2: Initial synthesis of **3.4**.*

Even though **3.4** was synthesised in a single step, it was isolated in very poor yield. By increasing the equivalency of 2-cyanopyridine to 2.2 equiv. and reducing the time LDA has with **3.11** (10 min. *vs.* 20 min.), **3.4** was isolated in an improved 25% yield with reduction in **3.15** by-product (Scheme 3-3).
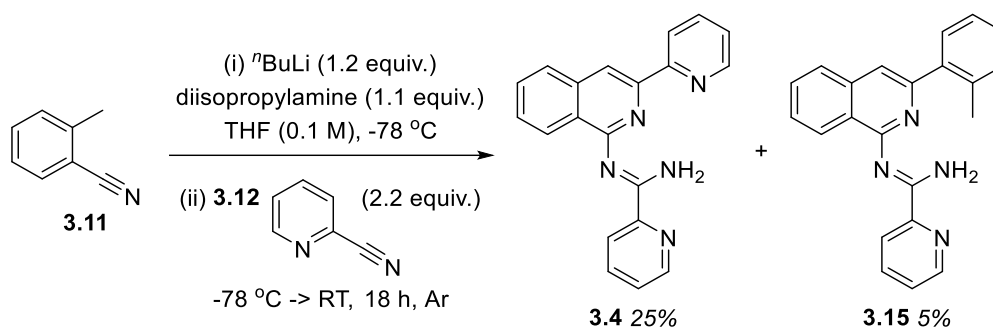
*Scheme 3-3: Improved synthesis of **3.4**.*

From these results it is clear that some kind of cascade reaction is occurring to produce **3.4** and **3.15**. It is thought that initially after LDA formation, deprotonation of 2-methylbenzontirile occurs (Scheme 3-4). Then either path A or B can occur, where ***int-***

3.11 nucleophilically attacks another molecule of **3.11** or **3.12** (added afterwards) at the electrophilic nitrile position. The formed anionic nitrogen of intermediates *int-3.16* and *int-3.17* nucleophilically add intramolecularly to the adjacent nitrile group to cyclise into the isoquinoline ring intermediates *int-3.18* and *int-3.19*. After this, another **3.12** electrophile is nucleophilically attacked intermolecularly to form the deprotonated amidine systems *int-3.20* and *int-3.21*, which then upon workup are quenched and aromatisation occurs (*int-3.22* and *int-3.23*), forming the isoquinoline ring systems and arrival at the final products **3.4** and **3.15**.
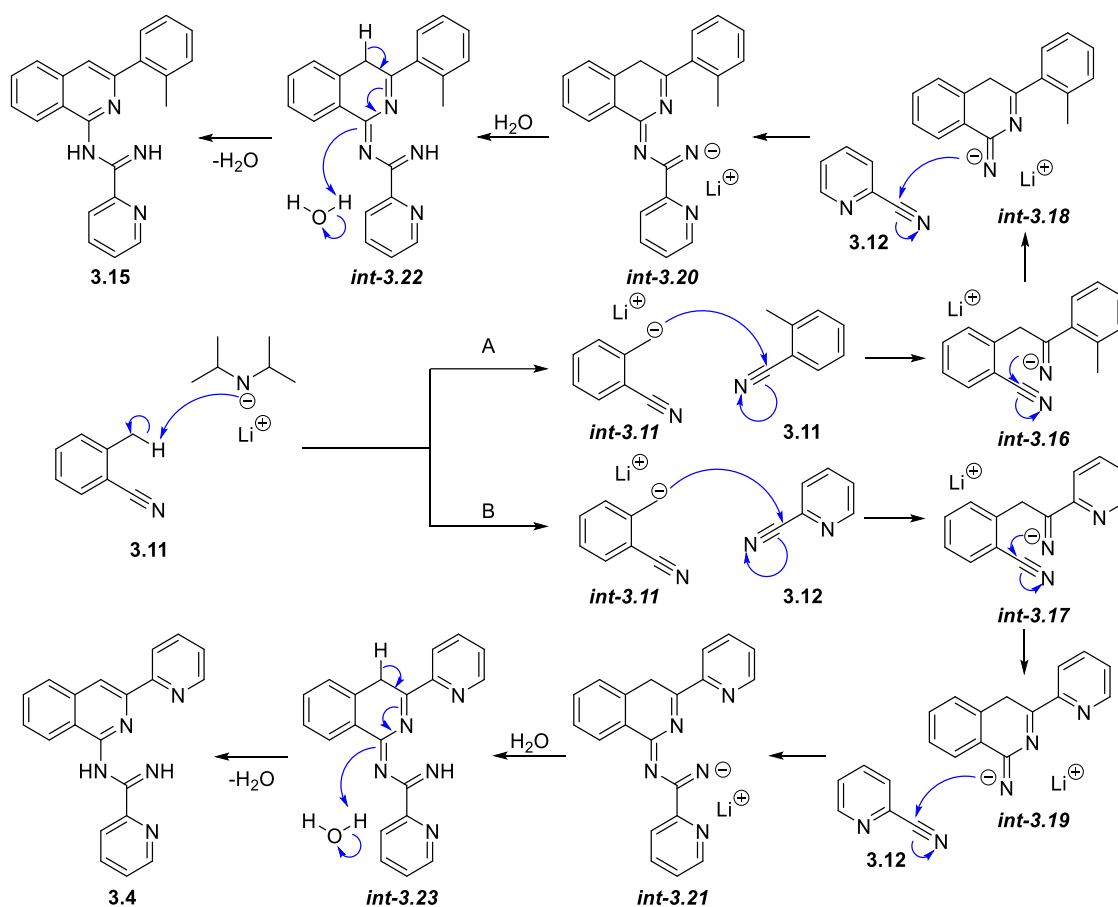
*Scheme 3-4: Postulated shared mechanism for the formation of **3.5** and **3.15**.*

Probably a better way to synthesise **3.4** would be to form LDA, then add a mixture of **3.11** and **3.12** where the stoichiometry of **3.12** is at least 2.0 compared to **3.11**, so that upon deprotonation of **3.11** it would react with 2-cyanopyridine preferentially, and thereby reduce the formation of undesired **3.15**. To conclude, NSC-637578 **3.4** was serendipitously synthesised in one step in low yield. Using a postulated mechanism led to two changes in conditions for an improved and acceptable yield. As 2-methylbenzontirile and 2-cyanopyridine are cheap and readily available, a bulk run of this reaction

(2.56 mmol) allowed access to enough material (211 mg) in sufficient purity (99%+) for full characterisation and soaking studies.

3.3.3 Isoalloxazine – NSC-288387, 3.5

NSC-288387 **3.5** has been previously identified as a weak dual inhibitor of IDO1 and TDO2 enzymes, whose expression in cancer is associated with suppressed immunity and poor patient prognosis.[136] Compound **3.5** has also been previously explored by our group in which DEEP-STD NMR was utilised to understand which structural elements are interacting with WWP2.[37] As an extension of this, gaining a ligand-protein cocrystal structure was targeted in order to investigate this hit further, for example in how best to perform the proposed scaffold hop to circumvent the observed redox activity. However, as with other hits from the HTS, literature surrounding its synthesis was sparse. The previously utilised synthesis was not used (Scheme 3-5) as the final Chan-Lam cross couplings between isoflavin **3.29** and **3.30** were low yielding. It was thought this would hamper bulk synthesis for soaking studies, and if applicable later analogue synthesis.
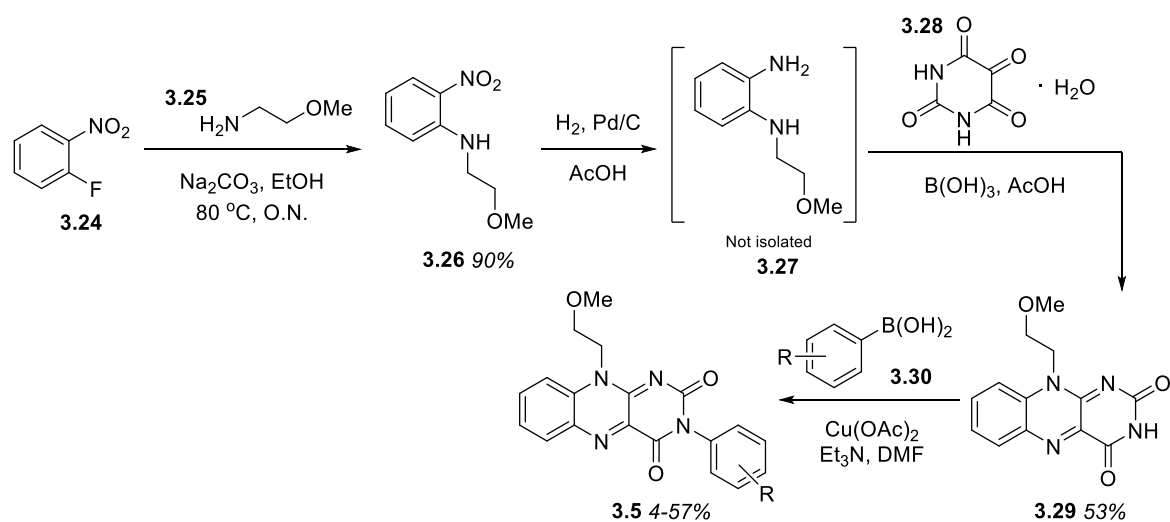
*Scheme 3-5: Previous synthesis of **3.5**.[28]*

Four papers mention NSC-288387's synthesis or an intermediate of, but either had limited experimental detail, or were in a foreign language and only report an analogue.[137–140] It was decided to use these papers nonetheless as the reactions seemed straightforward with, where applicable, reported high yields per step. Additionally, no column chromatography was required as each synthetic intermediate and the final product could be crystallised from polar solvents.

A paper outlining the synthesis of **3.5** is detailed below (Scheme 3-6): Initially, **3.31** is reacted with the primary amine **3.25** to allow for a transamination followed by cyclisation towards 1-alkyl-2-aminopyrazine derivative **3.32**.[137] Urea **3.34** then forms from reaction with phenyl isocyanate, after which cyclisation with Et₃N and aromatisation of **3.35** occurs to provide NSC-288387. It was thought that these transformations would be amenable to bulk runs to access workable quantities of product. This required the synthesis of the morpholino intermediate **3.31**.
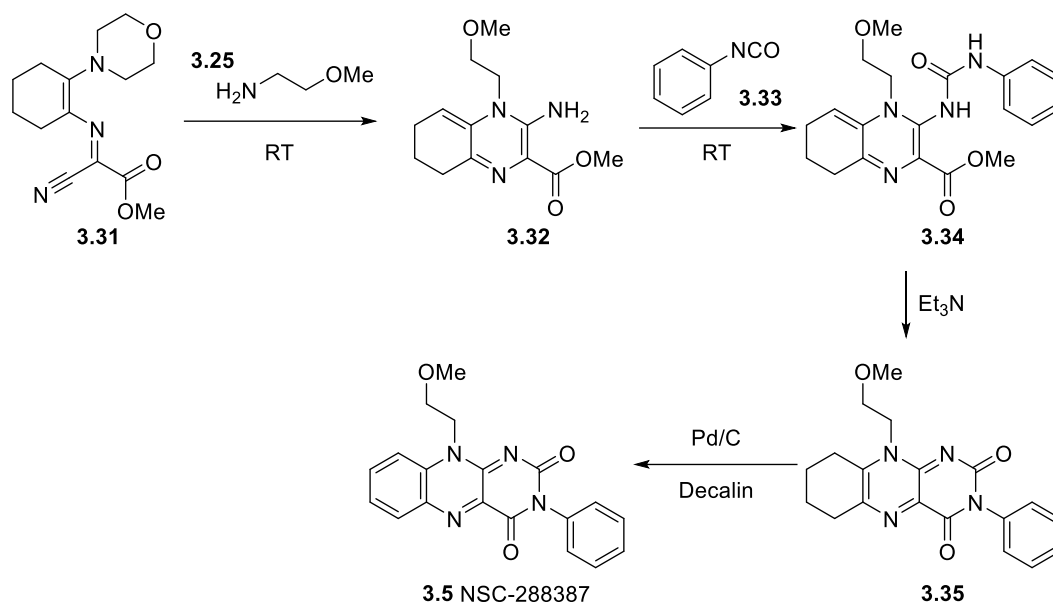
*Scheme 3-6: Synthetic sequence for the synthesis of **3.5**.*

A paper detailing the reactivity of *O*-tosyl oximes revealed that reaction of tosyl oxime **3.36** with 1-morpholinocyclohexene **3.37** would arrive at **3.31** (Scheme 3-7).[138] The enamine was commercially available, and **3.36** was synthesised readily through an oxime formation of commercially available methyl 2-cyanoacetate **3.39** followed by tosylation under standard conditions. Conditions were found for these syntheses from another paper detailing the production of closely related compounds to NSC-288387.[140]
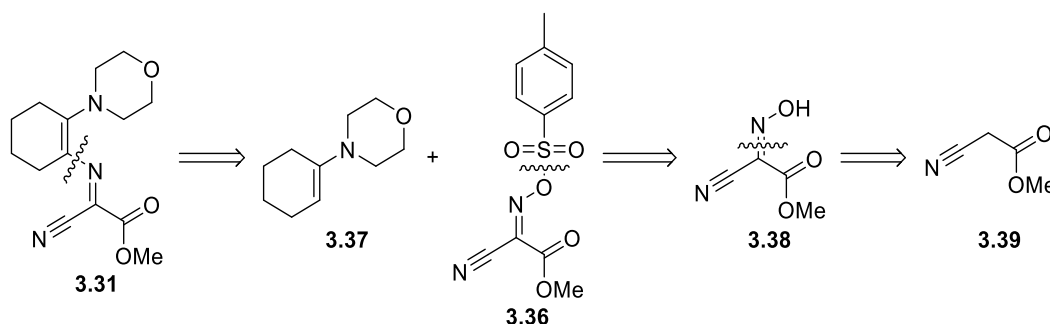
*Scheme 3-7: Retrosynthetic analysis of **3.31**.*

61

The initial synthetic sequence to access derivative **3.31** commenced with oxime formation from **3.39**, which with temperature control proceeded smoothly in high yield (Scheme 3-8). Next, tosylation of the free oxime **3.38** was effective in producing **3.36**, again at low temperature. Both reactions may be performed on decagram scales (27.40 g isolated free oxime **3.38**, 11.80 g isolated tosyl oxime **3.36**) with no diminishment in yields. Finally, reaction with enamine **3.37** to displace the tosyloxy group was performed in dry diethyl ether to access **3.31**. This reaction was again high yielding and amenable to an 11.0 g scale, providing 9.59 g of **3.31** for further reaction.
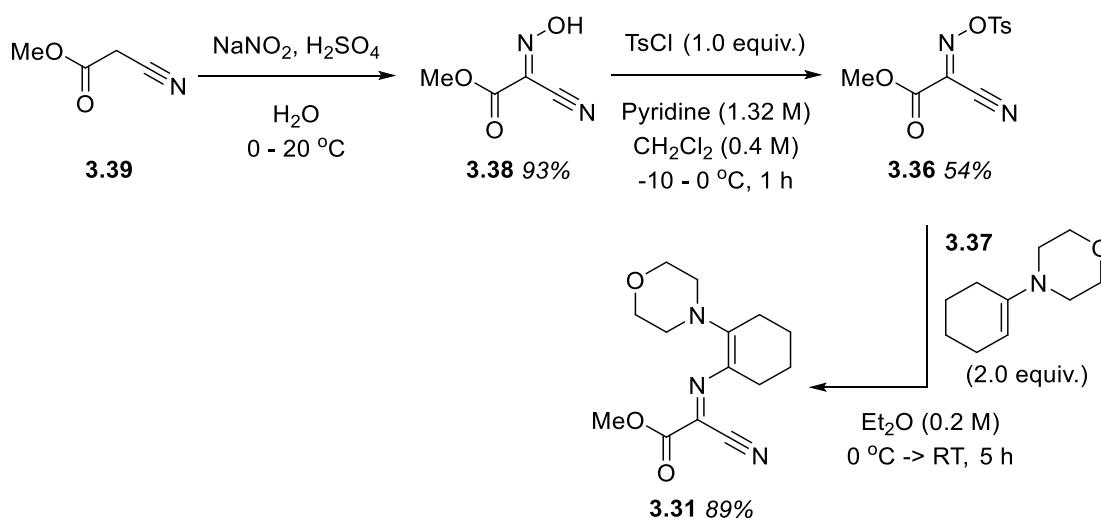
*Scheme 3-8: Synthesis of intermediate **3.31**.*

The next reaction allowed the formation of **3.32** in reasonable yield (Scheme 3-9).[140] It is postulated that a transamination mechanism occurs first through an aza-Michael-type addition and through intermediates ***int-3.40*** and ***int-3.41***, with release of morpholine, providing ***int-3.42***.[137] The secondary amine can then readily attack the cyano group intramolecularly to cyclise and provide ***int-3.43***. Proton transfer then occurs followed by amine reformation by deprotonation to provide **3.32**.[137] The reaction was found to be easy to run and the isolation straightforward by recrystallisation from ethanol.
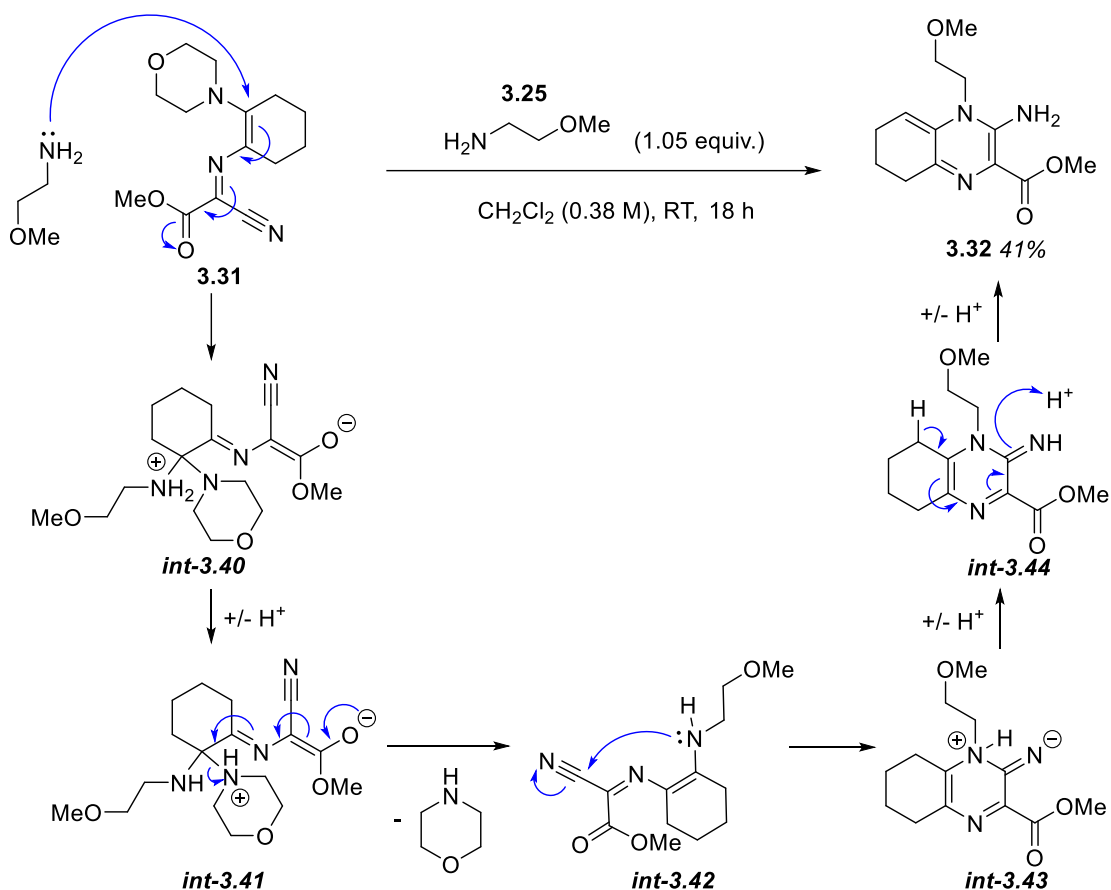
*Scheme 3-9: Postulated mechanism for the formation of **32** from **31**.*

The double bond in the cyclohexyl ring of **3.32** has a vinylic proton which is readily observable by ^{1}H NMR spectroscopy (Figure 3-5). In the starting material **3.31** the cyclohexenyl ring has two overlapping CH_2 positions, and two different CH_2 signals downfield of the overlapping peak. For the proton environments downfield, most are adjacent to the electron-withdrawn methyl 2-imino-2-cyanoacetate moiety, and the relatively electron-rich morpholine group is the least de-shielded of the two. In contrast a much more de-shielded vinylic proton signal is observed as a triplet at 4.63 ppm for **3.32**, and through COSY analysis the other signals, now where none are overlapping, may be deconvoluted and characterised.
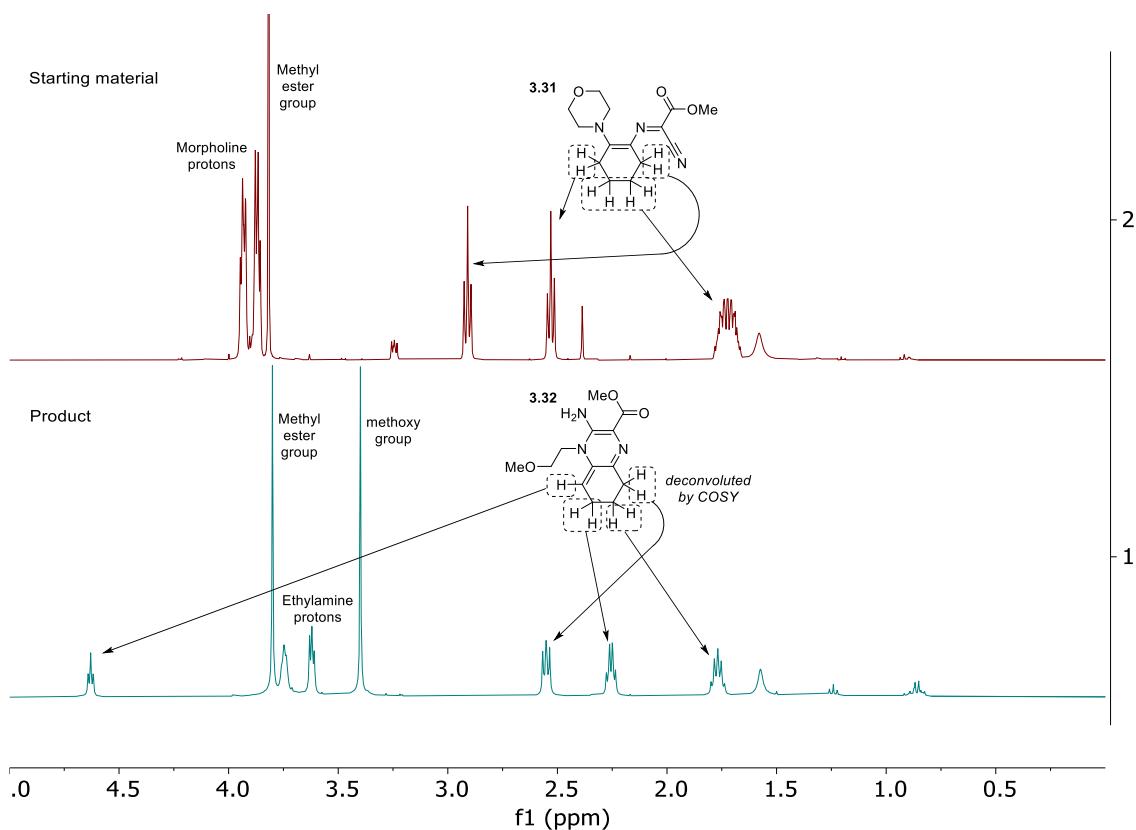
*Figure 3-5: ¹H NMR comparison between **3.31** and **3.32** within the 0.0 – 5.0 ppm region.*

The formation of **3.34** was next targeted, initially the reaction with freshly distilled phenyl isocyanate **3.33** was performed in chloroform with a slight excess of the isocyanate for 18 h (Scheme 3-10). The first run was successful, and the later bulk run (1.8 g scale) provided high yield without much more optimisation. A 75% yield (1.956 g isolated) of **3.34** was furnished by recrystallisation from methanol. In later bulk reactions forming **3.34** the yield became much lower compared to this initial result. These later reactions were using material isolated from sequential recrystallisations of the mother liquor of the bulk reaction forming **3.32**. It became obvious that the mother liquor was providing in later crops a by-product which was overlapping with **3.32** on TLC. It was thought that the by-product must have reduced nucleophilicity as it wasn't reacting with the isocyanate. Therefore, the initial guess was that something was happening to the amino group.
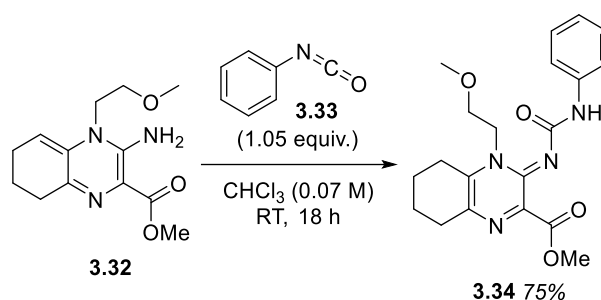
Scheme 3-10: Isocyanate **3.33** reaction with **3.32** to produce **3.34**.

Isolation and ^{1}H NMR analysis (Figure 3-6) of this by-product revealed three things. Firstly, the vinylic signal has disappeared and the alkyl region is similar to that observed for **3.31**. Secondly the methyl ester group has been removed with a concurrent broad singlet at 12.38 ppm appearing. Lastly the amino group, which is initially underneath the chloroform reference peak, has disappeared with a downfield broad singlet forming, integrating to one.
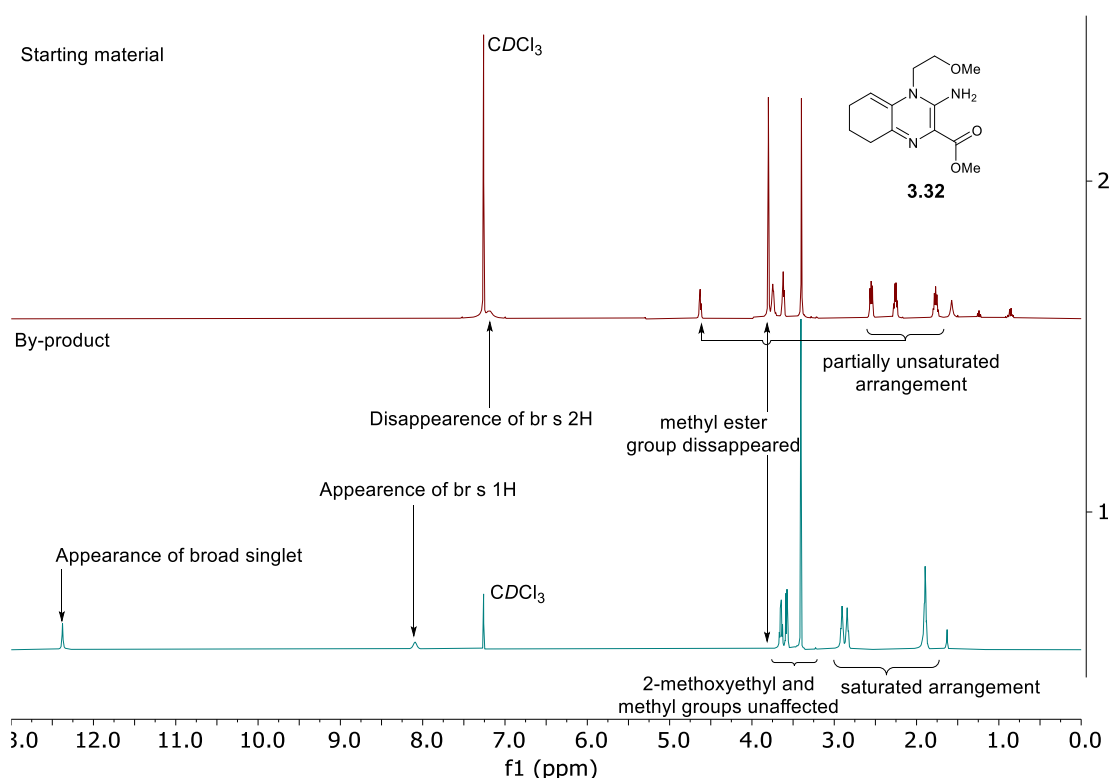
Figure 3-6: ^{1}H NMR (CDCl₃) comparison between **3.32** and by-product.

The initial idea of the structure was one where the methyl ester had been hydrolysed, and where some form of tautomerisation had formed an imine group where the amine group had been (Scheme 3-11). A plausible series of events could be that hydrolysis of the

methyl ester had taken place during either the reaction itself or the later purification by recrystallisation in ethanol.
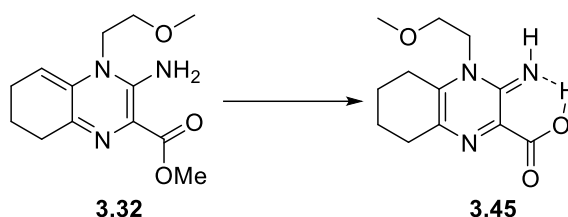
*Scheme 3-11: postulated by-product structure **3.45** based off ^{1}H NMR (CDCl$_3$) analysis.*

The structure is initially appealing as the ^{1}H NMR backs up this idea (Figure 3-6): the carboxylic acid proton being highly de-shielded at 12.38 ppm, and a possible scenario where the imine proton is in a hydrogen bonding arrangement with the carboxylic acid and hence is also de-shielded at 8.09 ppm.

After isolation, the material was crystallised and thankfully an x-ray structure was determined (Scheme 3-12, a). Surprisingly, the x-ray structure determined was in fact the amide **3.46**. A postulated mechanism is provided for the formation of this by-product from the starting materials (Scheme 3-12, b).
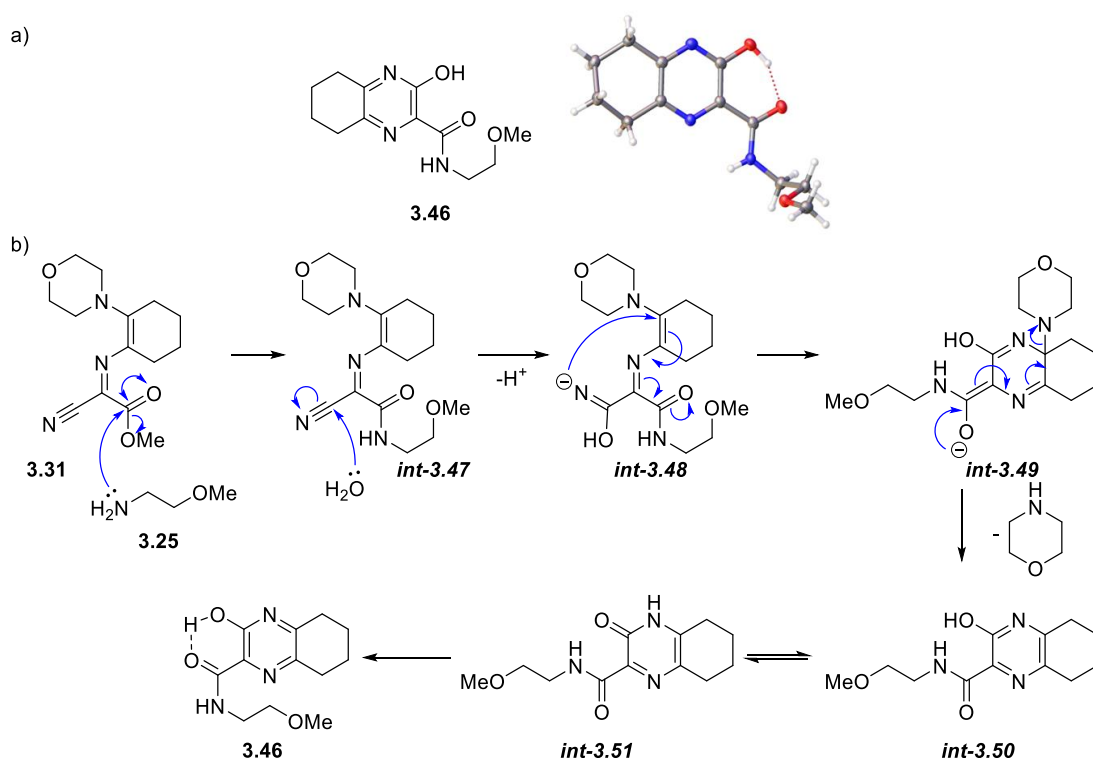
*Scheme 3-12: a) Chemical structure and X-ray crystal structure of by-product **3.46**. b) Proposed mechanism for the formation of **3.46**.*

66

Instead of **3.25** attacking the cyclohexenyl position, it undergoes an amide bond formation with **3.31** to provide ***int-3.47***. This amide could be stable and remain present until the recrystallisation step, where water present in the ethanol attacks the cyano position of ***int-3.47*** to provide anion ***int-3.48***, which is likely to attack intramolecularly *via* a 6-exo-trig cyclisation at the position *ipso* to the morpholine, forming the pyrazine intermediate ***int-3.49***. The anionic intermediate can now rearrange to aromatise the pyrazine core by removing morpholine to provide ***int-3.50***. This is in tautomeric equilibrium with the pyrazinone ***int-3.51***, and these two are essentially forms of the product **3.46**, which is apparently more stable in its hydroxypyrazine form probably because of the intramolecular hydrogen bonding between the hydroxyl and amide carbonyl oxygen.

For the formation of the hydroalloxazine core, initial experimentation from literature conditions were low yielding (Et$_3$N (1.0 equiv.) in MeOH (0.1 M) at 50 °C for 20 min., 22% isolated yield), and the cyclisation of urea **3.34** to form the uracil ring was better achieved using the same reagents at RT for 3 h (monitored by TLC). Initial test aromatisation was successful on a small scale, but attempts to isolate **3.35** proved difficult, so aromatisation was performed in one-pot with the cyclisation as an adaption of the literature procedures.[137] Analysis by ^{1}H NMR of the crude reaction mixture was used to identify when the cyclisation was finished by the disappearance of the methyl ester peak of **3.34** before aromatisation. The cyclisation was performed, followed by a solvent swap to degassed decalin and addition of Pd/C. The reaction was heated to 200 °C for 4 h (monitored by TLC) and after column chromatography and crystallisation NSC-288387 (**3.5**) was provided in 43% yield (348 mg) over two steps (Scheme 3-13). 300 mg of **3.5** was submitted for soaking studies. A crystal suitable for x-ray crystallography was obtained from slow crystallisation in MeOH (Figure 3-7).
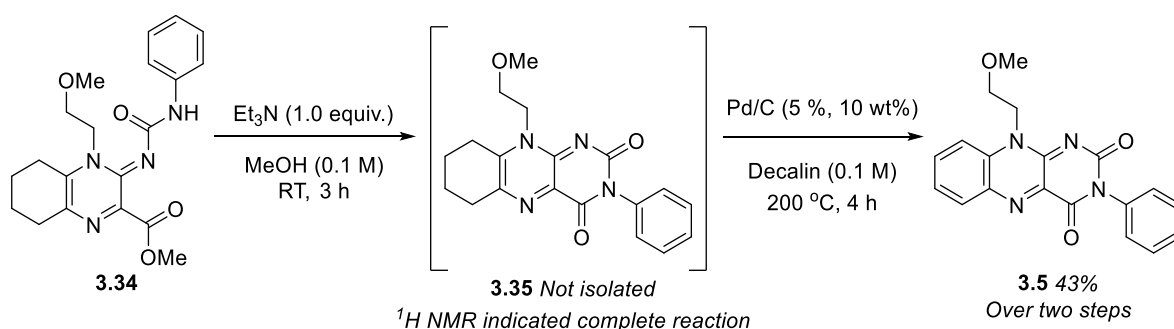
*Scheme 3-13: One-pot cyclisation of **3.34** followed by aromatisation to NSC-288387.*
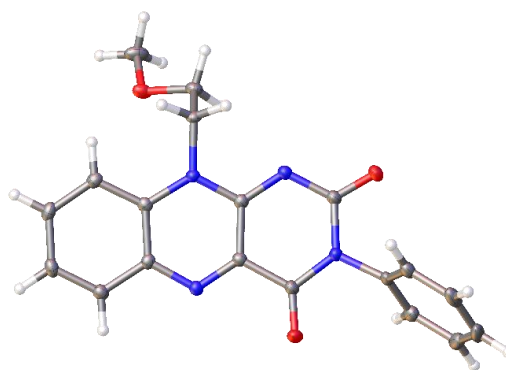
*Figure 3-7: X-ray crystal structure of NSC-288387 **3.5**.*

To summarise, NSC-288387 **3.5** was synthesised in 7 steps, with only the final step requiring column chromatography. Enough material was isolated for soaking studies, and an X-ray crystal structure was obtained from specific crystallisation conditions. Whilst I was able to access NSC-288387 in a highly pure form in a practically straightforward synthetic sequence, efficiency-wise the overall yield is worse than the previously described synthesis (6% *vs.* 27%).

3.3.4 Redoxal – NSC-73735, 3.6 and *F*-Redoxal, 3.7

Using the method presented here (section 3.3.4) the synthesis of both redoxal **3.6** and *F*-redoxal **3.7** in quantities amenable for soaking studies is possible (Scheme 3-14, a). Starting from either 2-methoxyaniline **3.52** or 2-fluoroaniline **3.53** the corresponding azobenzenes **3.54** and **3.55** were formed *via* reaction with MnO_2 under Dean-Stark conditions in toluene (Scheme 3-14, b).[141] Next, the azobenzene compounds **3.54**, **3.55** were reduced using Zn powder with NH_4Cl to the hydrazobenzene intermediates **3.56** and **3.57**.[142] Under acidic conditions, the benzidine rearrangement then occurs, which is thought to be *via* a [5,5] sigmatropic reaction to access the benzidine derivatives **3.58** and **3.59**.[143] Whilst the steps outlined here are low yielding, both 2-fluoro and 2-methoxyaniline are cheap reagents (25 g / £49.70 and 100 g / £28.40 respectively, Sigma Aldrich, accessed 02/06/23) and the reactions easy to setup and run. So sufficient quantities of the benzidine derivatives can be synthesised from these precursors using this route.
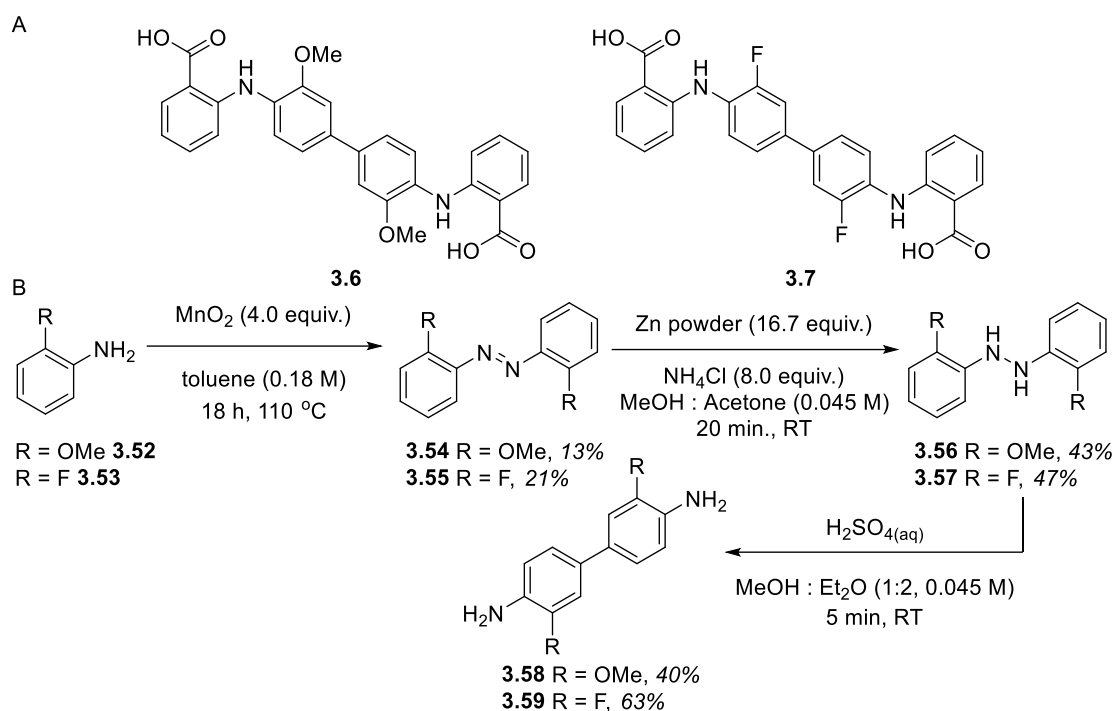
Scheme 3-14: A) Redoxal **3.6** and F-redoxal **3.7**. B) Synthesis of benzidines **3.58** and **3.59** from starting anilines **3.52** and **3.53**.

The benzidine rearrangement used to be a mechanistic puzzle for organic chemists, as it was unknown if the rearrangement occurs through either a concerted mechanism (an allowed sigmatropic suprafacial [5,5] process) or through a π-complex mechanism, which is essentially non-concerted.[144] The ability to solve this was unavailable until heavy atom kinetic isotope measurements were made.[145] It has demonstrated by using isotope-ratio mass spectrometry that the N-N bond breaking was involved in the transition state by the observation of a ^{15}N kinetic isotope effect consistent with calculated values. Measurements of 4,4'-dideuterohydrazobenzene showed an inverse kinetic isotope effect (sp^2 to sp^3), consistent with a concerted mechanism. Finally, measurements using ^{14}C enriched hydrazobenzene showed a primary kinetic isotope effect for the formation of benzidine, showing the C-C bond formation was also occurring during the transition state. Since all this data was consistent with a concerted mechanism, it is thought to undergo the aforementioned sigmatropic [5,5] rearrangement after protonation to **int-3.61** (Scheme 3-15), through **ts-3.62**. Re-aromatisation of **int-3.63** to **3.64** then occurs with loss of 2 protons.
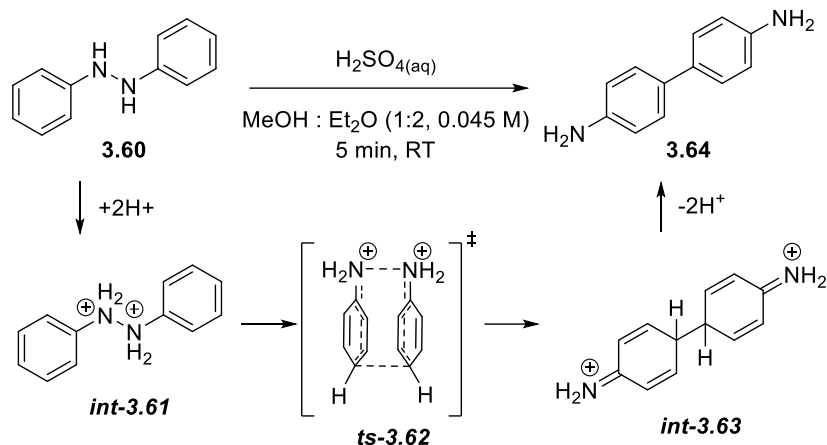
Scheme 3-15: Mechanism of the benzidine rearrangement.

From the benzidine derivatives **3.58** and **3.59**, Buchwald-Hartwig cross-coupling conditions with **3.65** were utilised to access the di-arylated products (Scheme 3-16).[146] Using 2.2 equivalents of methyl 2-bromobenzoate to favour the di-arylated product was attempted. Unfortunately, these reactions were not as successful as hoped for, probably due to some experimental or setup error. Thankfully however, the mono-arylated product of the fluorinated benzidine could be utilised in the fragment analogue screening (not shown here, see section 5.5.1 for details). Even so, enough of the di-arylated products **3.66** and **3.67** were isolated and these methyl esters were hydrolysed to access both redoxal **3.6** and *F*-redoxal **3.7** in sufficient quantities to be submitted for soaking studies (approx. 100 mg each).

(i) R = OMe, LiOH (1.0 M), THF (0.078 M), reflux, 18 h, **3.6** *98 %*
(ii) R = F, LiOH (1.0 M), THF (0.12 M), 50 °C, 18 h, **3.7** *80 %*
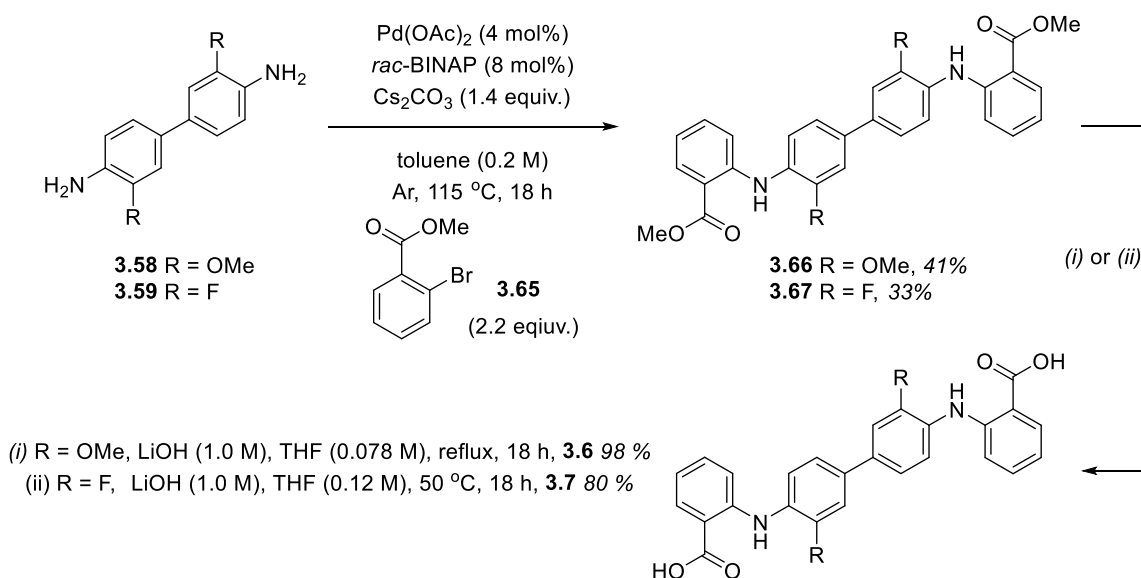
*Scheme 3-16: Buchwald-Hartwig di-arylation reactions on **3.58** and **3.59** followed by saponification to access redoxal **3.6** and F-redoxal **3.7**.*

Crystals suitable for X-ray diffraction were obtained for compounds **3.66** and **3.67**. Their structures are displayed below (Figure 3-8). Compound **3.68** was isolated as described in section 5.5.1 and is present here for comparison. Their biphenyl dihedral angles and diphenylamine angles are also displayed in Table 3-1 and Table 3-2, respectively.

a)

3.67

b)

3.68

c)

3.66
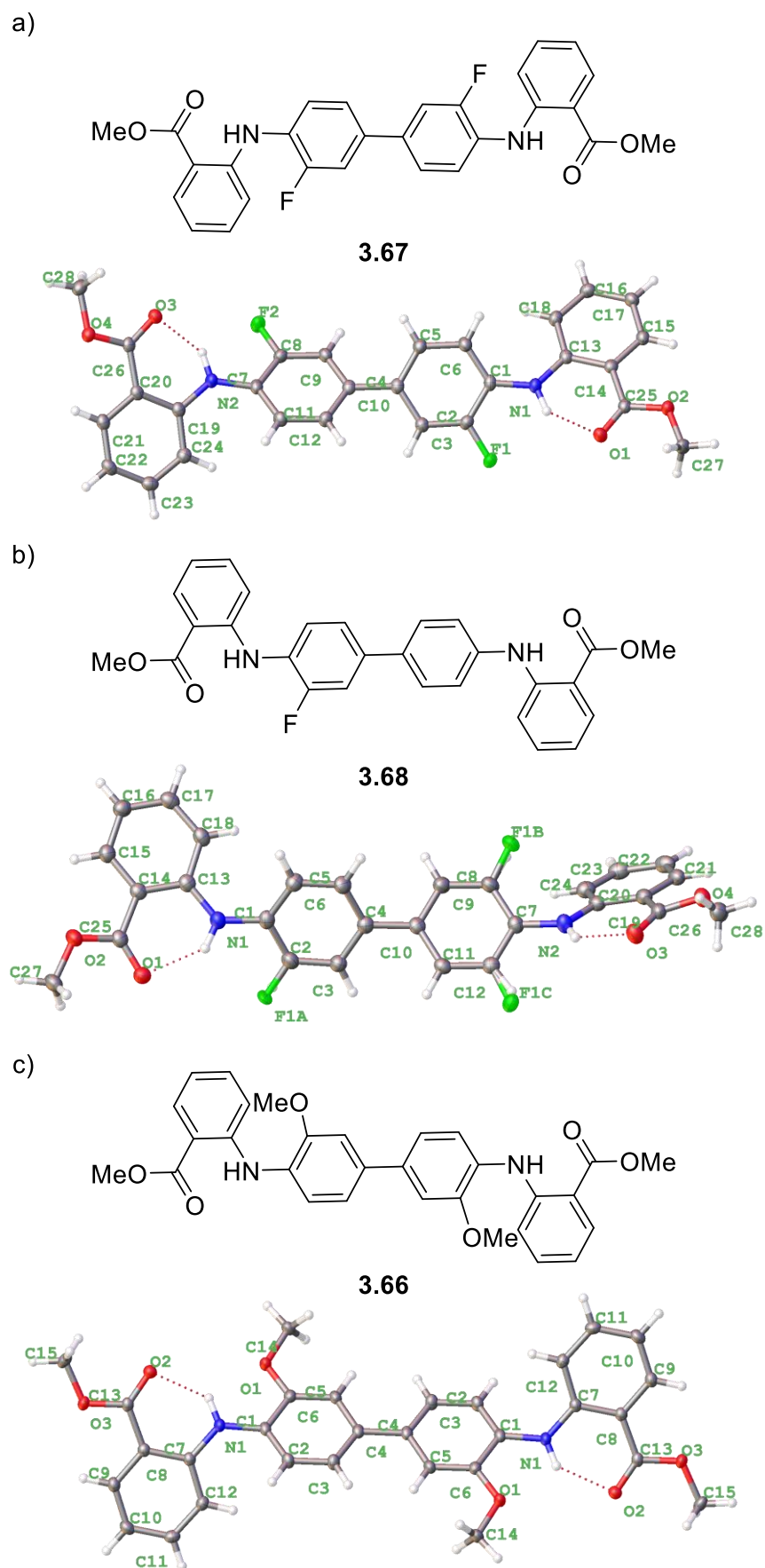
Figure 3-8: X-ray structures of a) **3.67**, *b)* **3.68** *and c)* **3.66**, *with all non-H atoms labelled.*

Compound	A	B	C	D	Torsion Angle (°)
3.67	C3	C4	C10	C11	36.0
3.68	C3	C4	C10	C11	38.4
3.66	C3	C4	C4	C5	0.4

*Table 3-1: Dihedral angles of compounds **3.67**, **3.68**, **3.66**.*

Compound	N1 angle (°)	N2 angle (°)
3.67	130.59 (C1-C13)	130.02 (C7-C19)
3.68	129.23 (C1-C13)	127.25 (C7-C19)
3.66	127.51 (C1-C7)	-

*Table 3-2: Diphenylamine angles of compounds **3.67**, **3.68** and **3.66**.*

Compared to the redoxal precursor **3.66**, whose biphenyl torsion angle is 0.4° (Table 3-1), both fluorinated derivatives have much greater torsion angles. **3.68** as a slightly higher torsion angle compared to **3.67** at 38.4° compared to 36.0°. Additionally, the diphenylamine angles for the fluorinated compounds are slightly higher than that of the redoxal precursor and are not symmetrical. Compound **3.67** has an incrementally higher angle about the nitrogen, but the difference between the two is marginal (at 0.57°). **3.68** on the other hand has different angles between the two nitrogen atoms, one lower than **3.66** at 127.25° and one closer to the angles of **3.67** at 129.23° (Table 3-2).

To conclude, a second synthetic route was devised to allow the synthesis of *F*-redoxal and redoxal in quantities amenable for soaking studies, with the main steps being a benzidine rearrangement and a Buchwald-Hartwig cross-coupling. Overall yields were low but comparable at 3% for *F*-redoxal and 0.9% for redoxal, because all steps were poorly or moderately yielding apart from the hydrolysis of the ester groups.

3.3.5 5-Bromoindole-3-carbinol 3.8, and 5-iodoindole-3-carbinol 3.9

Indole-3-carbinol was originally submitted for co-crystal analysis and is commercially available. However, halogenated indole-3-carbinol derivatives were sought to further improve the quality of the co-crystal structure (See section 3.2 for details).

Starting from 5-haloindoles **3.69** and **3.70**, Vilsmeier-Haack conditions were utilised followed by heating to reflux in base to re-form the enamine double bond (Scheme 3-17).[147] These conditions afforded in good yield both 5-haloindole-3-carboxaldehydes **3.71** and **3.72**. Finally, simple $NaBH_4$ reduction conditions to access the alcohols **3.8** and **3.9** from the aldehydes were performed.[127] As with many indole-3-carbinol derivatives, especially NH-derivatives, careful handling is required due to temperature sensitivity (see relevant experimental sections for details). Elimination of the alcohol functionality and the production of an aza-Michael acceptor is facile in NH-indole-3-carbinol systems. Unfortunately, the 5-iodoindole-3-carbinol **3.9** derivative degraded slowly even at -20 °C upon storage, and so was not suitable for analysis. 5-Bromoindole-3-carbinol **3.8** was stable for several weeks on the benchtop, before slowly discolouring.
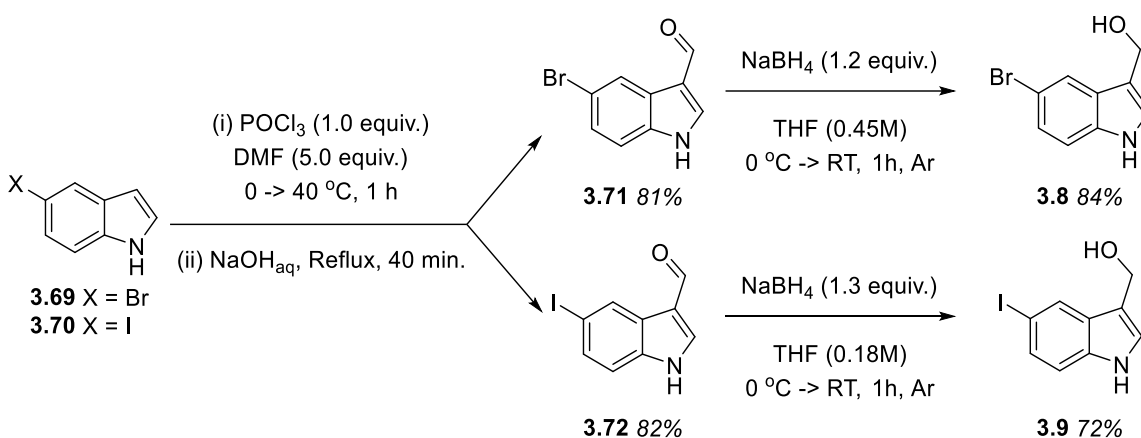
Scheme 3-17: Synthesis of halogenated indole-3-carbinols 8 and 9.

To summarise, both 5-bromoindole-3-carbinol **3.8** and 5-iodoindole-3-carbinol **3.9** were isolated in good overall yields (60% and 68%, respectively) and sufficient quantities for co-crystal structure analysis. Unfortunately, the 5-iodoindole-3-carbinol degraded and was not submitted for analysis. 5-Bromoindole-3-carbinol was stable enough to the submitted.

3.4 Synthetic Conclusions

Overall, all compounds were successfully synthesised, and all apart from 5-iodoinodole-3-carbinol **3.9** were acceptable for submission to the collaborators in the biology department to perform co-crystal analysis against the WWP2 construct (Table 3-3).
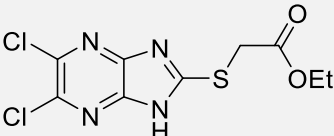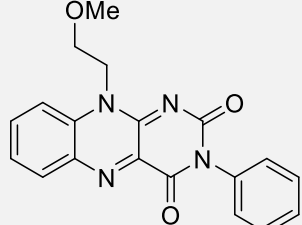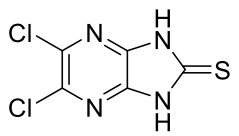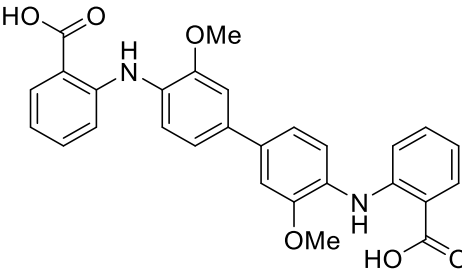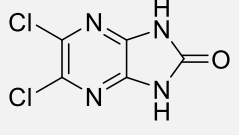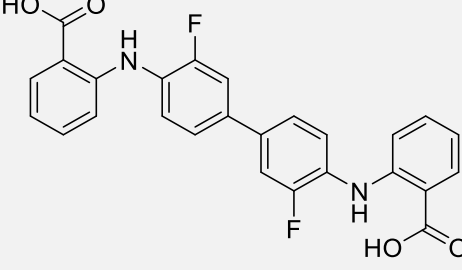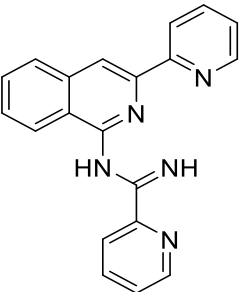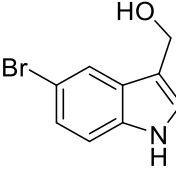
Entry	Structure	Entry	Structure
1		2	
3		4	
5		6	
7		8	

Table 3-3: Compounds 1 - 8 successfully synthesis and submitted for co-crystal analysis.

3.5 Computer-aided Design: Molecular Docking

3.5.1 Introduction to Molecular Docking

Generally, several major steps are involved in the molecular docking process, these include the preparation of the protein structure taken from the protein databank (pdb) and if the active site is unknown, a prediction of the active site is made.[148] This is followed by the preparation of ligands which are introduced into the software by either retrieval from databases, or sketched by the user and their energy minimised. From here, molecular docking may be performed by picking the grid box and running the calculations. The grid box represents the area within which the orientation of the ligand to the protein takes place, *i.e.,* the known or proposed binding site. The results of the docking, ligand poses, may now be analysed, usually with a provided docking score associated with how well the software has predicted the ligands interaction with the protein.

For the purposes of this investigation, Flare® was used to predict potential analogues for synthesis to improve activity by binding more strongly to the proposed binding site within the HECT domain of WWP2. As such, the HECT domain of WWP2 (4Y07) was prepared in a relatively routine fashion (see section 3.6 below) and ligands were sketched and minimised within the software. The analysis of the ligand poses and treatment of the docking scores is discussed in more detail below. Finally, if applicable, methodology validation was also attempted to gain an idea of the accuracy of the molecular docking process. Where co-crystal data was not available for hit structures, individual circumstances are described in section 3.5.5.

3.5.2 Representation of the Binding Site and Grid Box Design

As previously described, the work on structural elucidation of binding of hits against the WWP2 HECT domain has resulted in the identification of a proposed binding site (section 2.3). Usually, an increased entropy upon ligand binding *via* the removal of water from the binding site occurs, as well as a potential increase in enthalpy by strong interactions of the protein with the ligand. However, strongly bound water molecules provide a large enthalpic penalty upon displacement, so instead should be included within the proposed binding site as they may mediate ligand-water-protein interactions (usually, during protein preparation all water molecules are removed before ligand binding takes place).

As such an assessment of known protein crystal structures to identify conserved water-protein interactions has taken place as part of attempting to represent the binding site of

WWP2 more accurately. Full assessment of the behaviour of water is a very complex process and is outside the scope of this project.[149]

An initial screen of WWP2 HECT domain containing crystal structures from the protein data bank was performed.[149] These were collected in Flare and overlaid revealing a single conserved water molecule present in 5TJQ and 4Y07; with points of electron density present also in 5TJ7 and 5TJ8 WWP2 structures which arguably would also be un-orientated water molecules (Figure 3-9, discussed water molecule highlighted). Therefore, this water was presented in as part of the architecture of the site whilst running the molecular docking calculations.
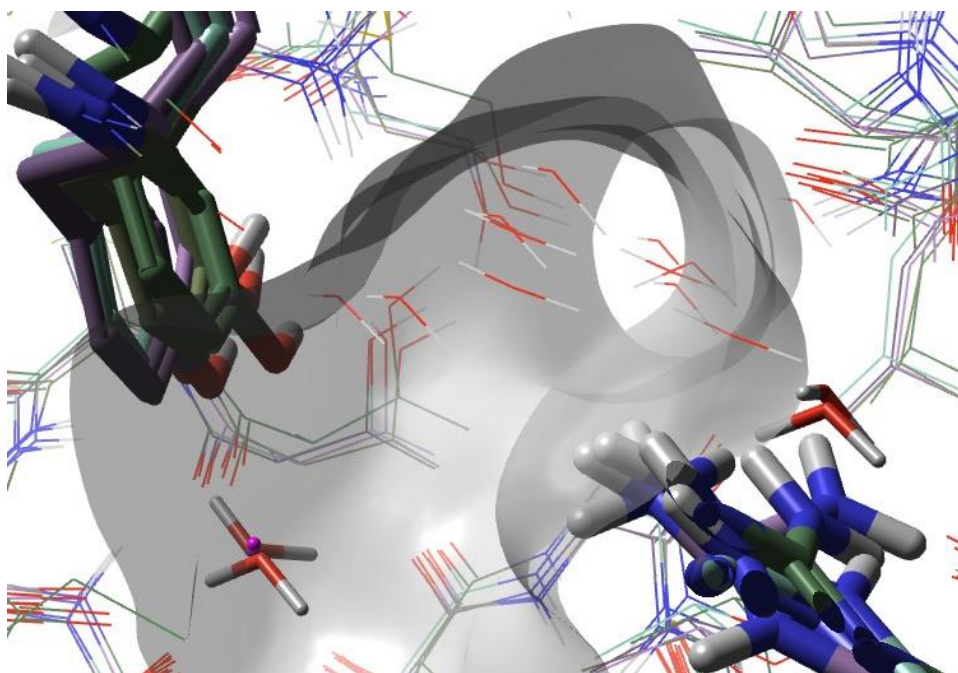
Figure 3-9: Overlaid proposed active site with conserved water molecules highlighted (thick sticks). Also highlighted are Tyr587 and Arg803 (named from pdb file 4Y07, molecular surface also from 4Y07). 5TJQ = WWP2 2,3-linker HECT, 4Y07 = HECT domain of human WWP2, 5TJ7 = WW2-2,3-linker-HECT aa 334-398 linked to 485-865, 5TJ8 = WW2-2,3-linker-HECT (no WW2 observed). Also overlaid was the crystal structure of HECT-2,3-linker-WW2 with I3C, I3C removed for clarity.

After pre-processing of 4Y07 protein crystal structure by the usual methods present in Flare, as well as removing water (apart from this water molecule described identified above) from the proposed binding site, the outline of the gird box was defined.

Usually, the grid box surrounds the known or proposed binding site, as well as additional areas that may have been chosen to probe for additional interactions adjacent to the binding site. Within Flare, the grid box is usually defined at the start of the docking process and consists of picking atoms that will be the limits of the grid box (Figure 3-10).

Residues were picked such that the molecular surface for the active site of the pre-process protein structure (4Y07) was completely encompassed by the grid box.
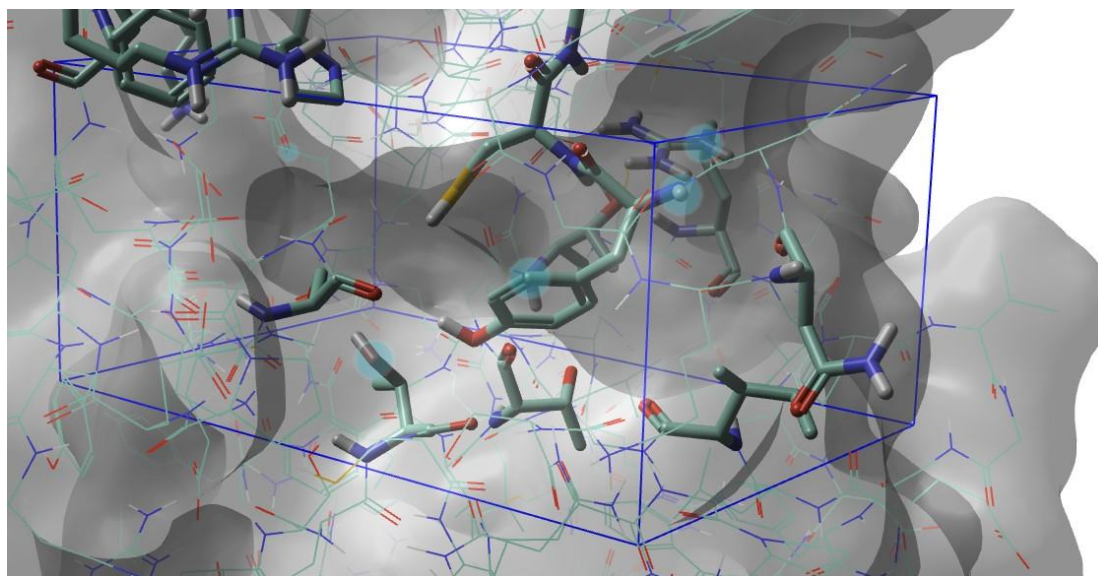
Figure 3-10: Grid box for the proposed active site of HECT WWP2 (4Y07).

3.5.3 <u>Visual Inspection of Ligand Poses – Avoiding Inaccuracies in the Scoring Function</u>

Molecular docking of potential ligands into biological targets provides a favourable ligand-protein configuration called a ligand pose. How well these ligand poses interact with the protein is quantified in a process called scoring as referred to above. However, to achieve realistic timeframes to reach results, several assumptions are made in the scoring function and this represents a significant limitation of the methodology.[149] These limitations are represented by inaccuracies in predicted binding free energies associated with each ligand pose score. Not only is the scoring function itself inaccurate, but limitations are also present due to the induced-fit effects of the protein, which is often treated as a rigid, non-flexible structure.

Due to these limitations a large number of false-positive scores can be produced and this limits molecular docking's applicability as a standalone platform.[149] Later applications such as molecular dynamics, and binding free energy calculations can be used to improve and build upon the results of molecular docking. However, visual inspection of ligand poses is still an important process to perform when deciding which hits and analogues to pursue, with a survey reporting that approximately 50% of 250 virtual screening publications conducted visual inspections for compound prioritisation.[150] In addition to the above, visual

inspection has been utilised to aid in the validation of docking protocols, particularly to evaluate redocking poses supplementing standard measures such as the root-mean-square deviation (RMSD) and native contacts.[149,151]

It seems that visual inspection offers potentially improved accuracy of binding favourability than just comparing scores, so visual inspection for determining ligand pose prioritisation was used. The analysis of these poses would aid in the decision of which analogues were worth synthesising for biological testing. Unfortunately, however, parameters for visual inspection are seldom provided within the literature. Importantly, gaining a consensus on the particular parameters sought during visual inspection is probably one of the best routes to expedite entry and aid in hit-to-lead or lead optimisation programmes.

Thankfully however, a recent survey has collected visual inspection criteria from successful campaigns utilising the visual inspection approach.[149] Twenty parameters were analysed by the group, which represents many variables to consider when contemplating a ligand pose. In addition, usually computational and medicinal scientists both individually and as part of a team meet to discuss and interpret the ligand poses, in order to decide which analogues, for example, to further pursue. Because of the large amount of time required to interpret ligand poses whilst keeping in mind twenty parameters, I decided that only some of the most-important variables available to interpretation (some parameters were not available on the Flare® academic license) as discussed in the article would be included in the parameters utilised in this work.

Interestingly, the importance of the docking scores themselves were valued the least by both experts from industry and academia, highlighting its unreliability. As such it does not form part of this decision-making process. Below are the parameters included in this visual inspection of docking poses to aid in choosing hit analogues to synthesise (Table 3-4). The table below describes the type of interaction looked for and how it will be assessed.

Name	Definition	How to be assessed?
Shape/steric complementarity	Extent of buried non-polar surface area of the ligand within the binding site.	Size of void(s) within binding site not occupied by the ligand.
Similarity to L-P X-ray co-crystal structure	The extent to which the predicted pose resembles the experimentally determined co-crystal structure.	Compare the predicted pose with that of co-crystal structure (if available)

Interaction with specific residues Arg803 & Tyr587	Specific binding site interactions which are conserved between hit compounds	Observation of any interactions with Arg803 or Tyr587. Have been shown as common interactors within this work.
Hydrogen bonding	The number, location and distances between hydrogen bond donors and acceptors are important distinguishing features in hydrogen bonding.	1. Comparison of these against the crystal structure, 2. Frequency of same / similar hydrogen bonding occurring. 3. Local environment of H-bonds 4. Occurring on side chain residues or the backbone of the protein.
Nature of exposed ligand atoms	What type of functionality is protruding into the bulk solution form the ligand when posed in the binding site can dictate how favourable the binding is.	Assessment of the functionality protruding outside of the binding site.
Ligand strain	Ligand strain is adoption of a configuration for optimising its complementarity to the protein. It's an energetic cost upon deformation of the ligand	Identification of unrealistic geometries of functional groups (amides, esters) or overly close contacts between L and P.
Unsatisfied heteroatoms / H-bonder buried within the binding site.	Ligand heteroatoms or hydrogen-bond donors/acceptors that are within the binding site and are not solvent exposed, interacting with the ligand, water, or another part of the protein/ligand.	Visual inspection of the internal void spaces of the docking pose to identify any heteroatoms or hydrogen bond donors/acceptors of the ligand that are not interacting with anything

Table 3-4: Criteria used in the visual inspection of docking poses.

One of the most important visual inspection parameters is steric complementarity. Steric complementarity is simply the buried non-polar surface area of the ligand within the binding site. The extent of buried non-polar surface area is well correlated with free energy of binding as determined by isothermal titration calorimetry (ITC) experiments.[152] The similarity to an acquired ligand-protein x-ray crystal structure is also highly regarded during visual inspection. During the survey the analogy to the crystal structure was deemed the most important parameter by all types of participants irrelevant of whether they were industry or academic workers, and their experience within computer-aided drug design.

Pre-existing knowledge of a protein-ligand complex may aid in measuring the quality of a ligand pose. Available crystal structure data of the biological target can identify important or common interactions between the ligand and the protein, and replication of this in different ligand poses can be utilised as part of the molecular docking visual inspection programme. Motivated by the understanding that Arg803 and Tyr587, identified *via* both NMR experiments and available crystal data, are common interactors between NSC-288387, NSC-2805 and I3C molecules. To account for this I identified them as important residues to look out for interactions with the proposed ligands within the visual inspection parameters.

Hydrogen bonds are easy to identify and are important interactions for binding affinity and selectivity. If applicable, hydrogen bonds commensurate with the bonding make up in co-crystal x-ray structure are routinely considered important during visual inspection. Importantly, the position of hydrogen bonds is also considered, as bonds deeper within binding sites in hydrophobic areas have greater impact on $\Delta G_{binding}$ values as compared to hydrogen bonding interactions near the edge of the binding site at exposed regions. To what part ligand interactions are taking place to the protein are also important. Bonding interactions to side-chain residues of the protein are typically less favourable due to their flexibility and the concomitant entropic cost being greater than bonding to protein backbone structures, which arises from the rigidification of the protein structure upon binding.

The nature of exposed ligand atoms also needs to be considered, as hydrophobic groups are generally considered unfavourable if protruding into the solvent where they may otherwise have potential to interact with the protein. Additionally, polar groups are more favourable as they may engage in polar intermolecular interactions with the solvent. Unsatisfied hydrogen donors and acceptors buried in the binding site additionally need to be carefully considered because of their resultant lowering of binding free energies due to an uncompensated desolvation free energy penalty.

The energetic cost upon ligand deformation during binding is termed ligand strain. Because the software utilised treats the protein as a rigid structure in order to reduce computational cost, protein strain cannot be accounted for. The energetic change upon ligand deformation is typically not optimally represented for by the scoring function, and these limitations may result in unrealistic geometries of functional groups within the ligand or overly close contacts with the protein. Usually, protein flexibility is accounted for later during postprocessing using molecular dynamic simulations. Unfortunately, this was not pursued due to time constraints and would have been computationally very expensive for the number of compounds docked.

3.5.4 <u>Validation of Docking Software – Comparison of I3C Poses Against the Ligand-Protein Co-crystal Structure</u>

A routinely used procedure to assess the performance of a docking program is to calculate the root-mean-square deviation (RMSD) between the docked ligand and a reference ligand pose, referred to as re-docking. Typically, the highest value at which a program is thought to reproduce the correct pose (most commonly compared to crystallographic data) is 2.0 Å.[153] This represents the upper limit when performing re-docking and software which is not able to reproduce the native pose accurately enough should be replaced with either another methodology or software. RMSD values above 3.0 Å are meaningless when looking at docking pose quality. Docking poses above this value have very few, if any, of the binding interactions of the true (crystallographic) binding pose. It is the key binding interactions that really define the binding pose much better than a metric such as RMSD, and therefore a visual inspection of the redocked ligand poses is also considered by comparison to the crystallographic data during re-docking.

RMSD calculations cannot be used for molecules of dissimilar structure to the molecule in the crystal structure as it relies on a one-to-one atomic correspondence. This means the redocking technique to validate the programs methodology for different hits cannot be used against co-crystal structures of different ligands within the crystal structure. Since an x-ray co-crystal structure was obtained for I3C, and to ensure some level of accuracy of docking poses could be obtained, a validation of the protocol described below followed.

Firstly, Flare® allows the user to run four different types of docking calculations, either *quick*, *normal*, *accurate but slow*, or *very accurate but slow*. I chose not to use *quick* as it did not provide different poses to interpret. The observation of key interactions was sought as well as good RMSD values within the quality confinements described above. Two potential orientations of I3C were provided and modelled in the proposed binding site. The only difference between these two is the orientation of the hydroxyl group hydrogen, where the bond fits into two interactions as shown (Figure 3-11). Either the hydrogen of I3C is interacting with either Ser646 (3.0 Å) or Thr627 (2.3 Å). It is quite possible that these two interactions are interchangeable, although the interaction with Thr627 is shorter and could be favoured.

Figure 3-11: Left: orientation of the carbinol such that it is interacting with Ser646, right: its interaction with Thr627.

Next, all calculation methods presented at least two of the three major interactions obtained from the cocrystal data (Figure 3-12). All ligand poses reproduced the Tyr587 hydrogen bonding with the carbinol moiety of I3C at distances (1.8 – 2.2 Å) comparable to the crystallographic data (1.9 Å). Both *normal* and *accurate but slow* calculations have basically identical orientations where a much more face-on orientation is observed for the π-π interactions between the indole and the Tyr587 phenol, at a longer distance compared to the crystallographic data. Whereas the *very accurate but slow* method has closer agreement to the more side-on T-shaped π-π interactions at a very similar distance. The hydrogen bonding interaction with Ser646 was only identified from the *very accurate but slow* calculation at very similar distance. By comparison, the *normal* and *accurate but slow* calculations provided the alternative nearby interaction with Thr627 at shorter distances compared to the crystallographic data, arising from the general difference in orientation of the indole structure. This information is also tabulated along with the associated RMSD values for these poses, with the mean RMSD+/-standard deviation for the *accurate but slow* and *very accurate but slow* calculations (Table 3-5).
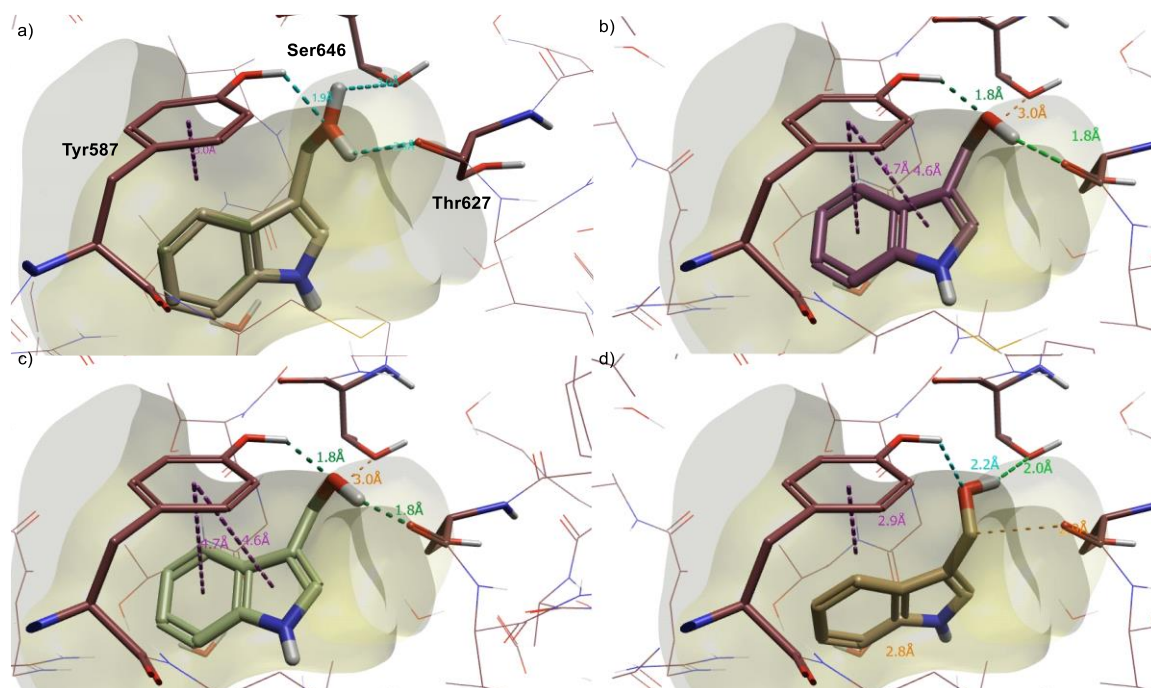
Figure 3-12: a) Both crystallographic poses (overlaid), b) normal docking calculated pose, c) accurate but slow docking calculated pose, d) very accurate but slow docking calculated pose.

Docking Calculation	Interactions observed (distances (Å))	RMSD (Å)	Mean RMSD (Å) (+/- std.dev.)
Crystallographic	Tyr587 OH – O(H)C (1.9) Tyr587 π – indole π side-on T-shaped (3.0) Ser646 C(H)O – HO (3.0) Or Thr627 C=O – HO (2.9)	N/A	N/A
Normal	Tyr587 OH – O(H)C (1.8) Tyr587 π – indole π face-on (4.7, 4.6) Thr627 C=O – HO (1.8)	1.846	N/A
Accurate but slow	Tyr587 OH – O(H)C (1.8) Tyr587 π – indole π face-on (4.7, 4.6) Thr627 C=O – HO (1.8)	1.846	4.2715 (1.14)
Very accurate but slow	Tyr587 OH – O(H)C (2.2) Tyr587 π – indole π side-on T-shaped (2.9) Ser646 C(H)O – HO (3.0)	1.184	3.747(1.44)

Table 3-5: Summary of docking calculation type vs. crystallographic data, observed intermolecular interactions between the protein (LHS) and I3C (RHS).

When overlaid, these ligand poses are essentially occupying the same space in the binding site, but with slightly differing orientations of the indole ring system, which either allows a face-on or side-on interaction with Tyr587, as shown in Figure 3-13. From a both visual and RMSD metric perspectives, based off the aforementioned comparison of both orientations of fitted I3C into the crystallographic data against the molecular docking calculations, there is good agreement between practically all methods presented and both modelled orientations of the crystallographic data. This arguably means that either of the three docking calculations are valid for analogues of I3C with confidence that the poses are accurate.
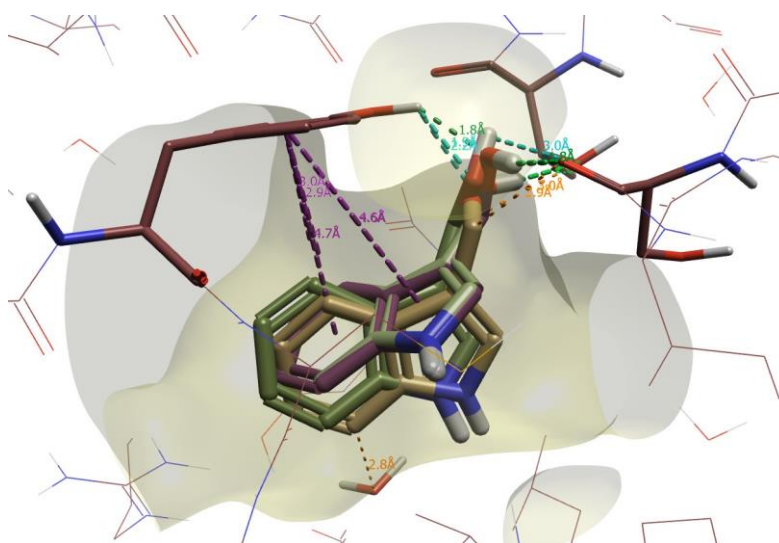
Figure 3-13: Overlaid I3C ligand poses obtained from the various types of calculations performed.

Of course, too much movement away from functionalities which differ significantly at the key interactor site (the carbinol) should be treated with increased caution. Achieving high docking accuracies is a challenging task.[154] This is exacerbated by the understanding that the quality of the co-crystal data is not perfect due to the flexibility of the WW2-2,3-linker section of the protein. This has meant the fitting of the indole-3-carbinol into the proposed binding site is only based off the increased electron-density within this position without a proper understanding of the specific orientation of the ligand within the binding site. This has knock on effects as the molecular docking method has been validated based upon this model, which could mean reproducibility or correlation issues of the ligand poses *vs.* the experimental activity could be later observed.

3.5.5 <u>Treatment of Hits without Co-crystal Structures</u>

How the other hit compounds were treated considering the lack of co-crystal structures was essentially to run the *very accurate but slow* docking calculation and then perform a visual inspection of the produced ligand poses. The previous example highlighted how well this calculation method reproduced the crystallographic data in terms of orientation and interaction distances, so its use would hopefully allow the identification of plausible ligand poses and allow productive comparison against virtual analogue ligand poses. The best ligand poses for these analogues should essentially reproduce the same binding interactions of the non-modified portion of the ligand, as well as the extent of buried surface area. As discussed below, only NSC-217913 utilised this approach, and analogues within the project pursuing NSC-2805 as a hit structure were chosen only utilising co-crystal data analysis. The project pursuing redoxal as a hit followed a different strategy than analogue production and growth as discussed below.

3.5.5.1 <u>NSC-217913</u>

Based on the docking results for NSC-217913 to the proposed binding site, 10 ligand poses were produced but only 5 are discussed as the other 5 were thought to be too unrealistic ligand poses. They have been grouped into similar orientations for easier presentation and comparison.

Two ligand poses orientated the imidazo[4,5-*b*]pyrazine core so that the thioether side chain pointed out of the binding site (Figure 3-14). These orientations were not thought to be particularly favourable due to the lack of steric complementarity with the binding site, essentially situating the ligand at the entrance of the site.[149] These poses have strong cation-π interactions with the Arg803 residue (3.8 – 4.3 Å), with the ethyl thioacetate chain wrapping around to engage in several hydrogen bonding interactions. The hydrogen bonding interactions are situated in a hydrophilic environment (near surface of protein) and may not contribute hugely to the binding energy compared to hydrogen bonds situated in hydrophobic regions (*i.e.*, further into the binding site). Additionally, with the flexible side chain directed outward of the binding site this may engage in interactions with the bulk solution instead of binding. The orientation of pose 1 has some additional interactions as a result of the *NH* imidazole hydrogen bonding with a carbonyl of Asn586 (at 2.4 Å) and sulfur-lone pair interactions with Asn585 (3.7 Å). Pose 2 has an additional π-π interaction occurring with Tyr587 (3.4 Å), although it is not clear if the distance is realistic as the vector describing this interaction isn't pointing at the phenolic residue of

Tyr587. With both poses, only one nitrogen or *NH* is participating in hydrogen bonding as this represents several unsatisfied heteroatoms within the binding site for these poses.
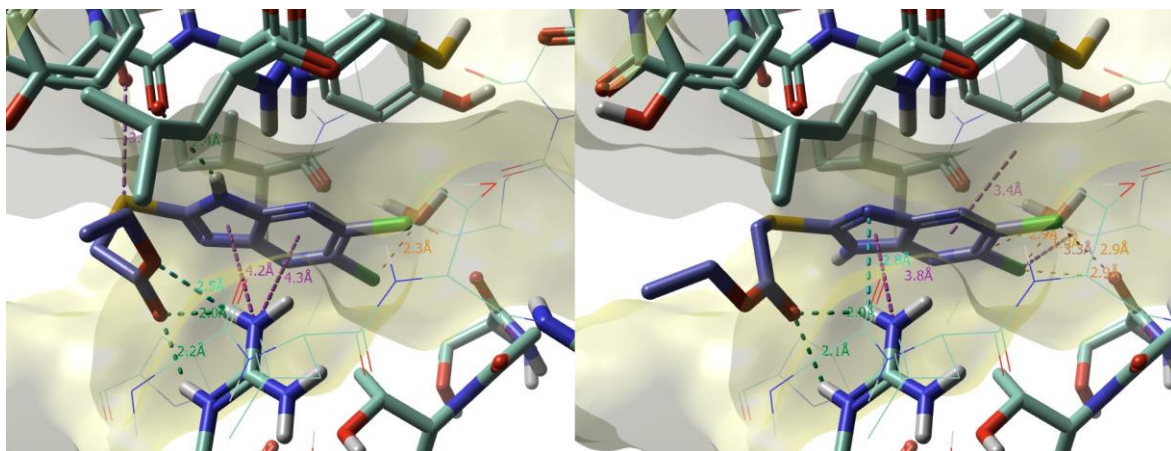
Figure 3-14: Ligand poses 1 (left) and 2 (right).

Sulfur lone-pair interactions are an initially puzzling interaction and are usually thought to be unfavourable interactions due to lone-pair – lone-pair repulsion. However, attraction can sometimes arise between sulfur atoms and oxygen or nitrogen atom lone pairs due to compensatory favourable interactions, such as electrostatic dipole-dipole interactions, and favourable orbital interactions (*e.g.*, $X_{lp} \leftrightarrow \sigma^{*}_{S-Y}$).[155]

Pose 2 has an additional halogen bonding with one of the chlorine atoms of the heterocyclic core and a carbonyl oxygen of an Asn840 residue. As can be seem more easily in Figure 3-15. The interaction is not entirely 'head-on' with respect to the C-Cl bond, and this is not ideal as halogen interactions with nucleophiles typically occur *via* lone-pairs from the nucleophile and the σ-hole of electron-deficiency opposite to the C-Cl bond. The origin of the σ-hole may be explained by a molecular orbital framework of three unshared electrons in a $s^2p^2p^2$ arrangement about the halogen atom, generating an area of electron density about the central region. In contrast, the bonding P orbital forming the C-Cl bond is distorted towards carbon, creating a point of attraction for lone-pairs opposite to the C-Cl bond.[152] At the same time, the distance is at the limits of what is expected for halogen bonding (3.3 Å) and there are clashes between both chlorine atoms and another carbonyl oxygen of the Thr799 residue at 2.9 Å each, probably as a result of non-linear geometries for interaction. Finally, close steric clashes (orange) are observable for this orientation with the conserved water molecule.
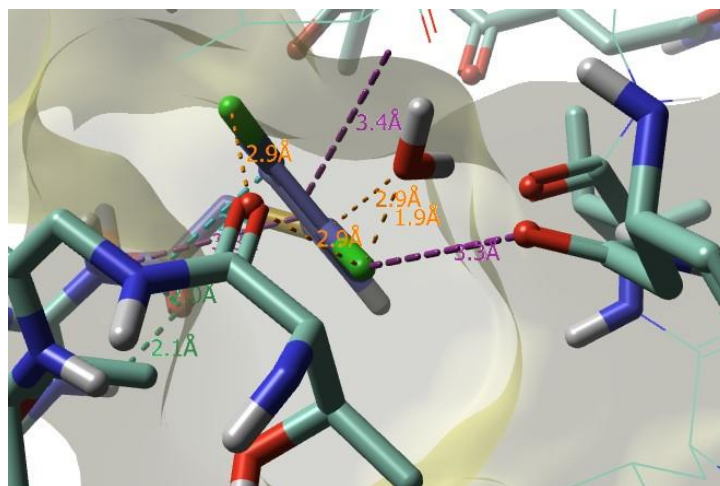
Figure 3-15: Ligand pose 2 highlighting halogen bonding and steric clashes of the heterocycle.

Several poses were found where the orientation of the ligand points the side chain more so into the binding site than outward into the bulk solution (Figure 3-16). These poses are more favourable in terms of buried surface area, whilst maintaining π-π interactions with Arg803, pose 4 to a much greater extent than pose 3 with both heterocycles participating (3.0 – 5.0 Å). Additionally, both poses have observable sulfur lone-pair interactions occurring between the thioether and the residue Ile813 carbonyl oxygen, both at 3.9 Å (additionally the carbonyl oxygen of Thr799 for pose 3 at 3.4 Å). Again, steric clashes are observable with the conserved water residue, and the carbonyl oxygen of the ester moiety (pose 4) or the thioether functionality (pose 3). In pose 4 the ester moiety is not interacting favourably at all and has only unfavourable clashes. Both poses have favourable interactions with Cys588 either *via* the carbonyl of the ester (2.0 Å, pose 3) or the *NH* imidazole (2.5 Å, pose 4).
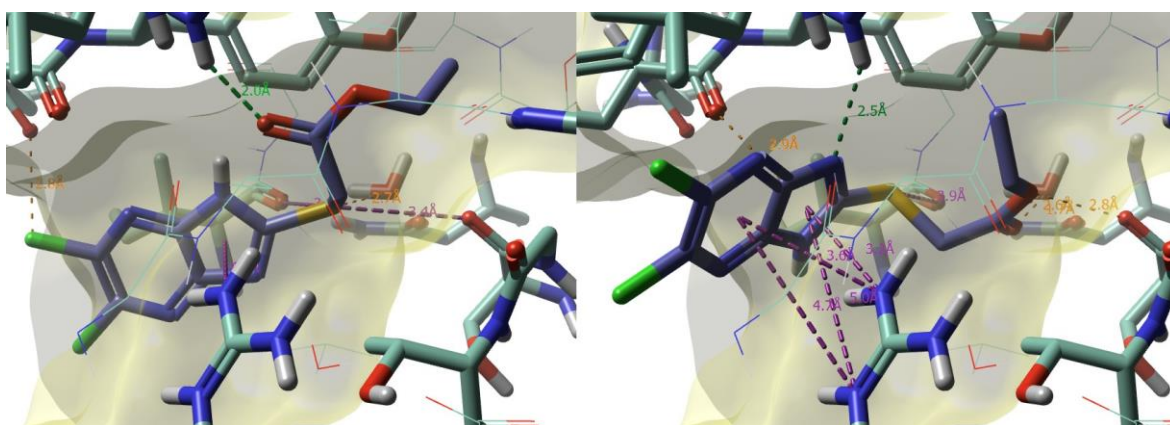
Figure 3-16: Ligand poses 3 (left) and 4 (right).

Finally, pose 5 showed the most promise, with the greatest observable steric complementarity (Figure 3-17) and no steric clashes. No other poses are present without

any steric clashes. The ethyl thioacetate side chain is in a favourable orientation with both oxygens participating in hydrogen bonding interactions with Leu628 at distances of 2.3 Å. The thioether has sulfur lone-pair interactions with Thr799, as does the *NH* imidazole at 4.3 Å and 2.0 Å respectively. Additionally, the imidazo[4,5-*b*]pyrazine core has face-on π-π interactions with phenolic residue of Tyr587 and hydrogen bonding between one pyrazinic nitrogen and Cys588 at 2.1 Å. The face-on π-π interaction is far (4.1 Å) and thought to be at most a weak interaction. However, due to how the core is situated its thought it may be able, in this specific orientation, to interact with both residues Arg803 and Tyr587, perhaps simultaneously. Weak hydrogen bonding is observable between the *NH* imidazole position and the conserved water residue within the binding site. Finally, the two chlorine atoms are exposed to the bulk solution, but it was thought that because halogen interactions can form from both nucleophiles and electrophiles, this would be a favourable interaction with water.[152] Only two buried heteroatoms (both on the heterocyclic core) are not satisfied by interactions within the binding site. However, this may be beneficial as the core would not have to orientate itself in a single conformation with respect to the binding site for a favourable interaction similar to the one shown (as would be the case, for example, with a pyrrolo[3,2-*b*]pyridine core).
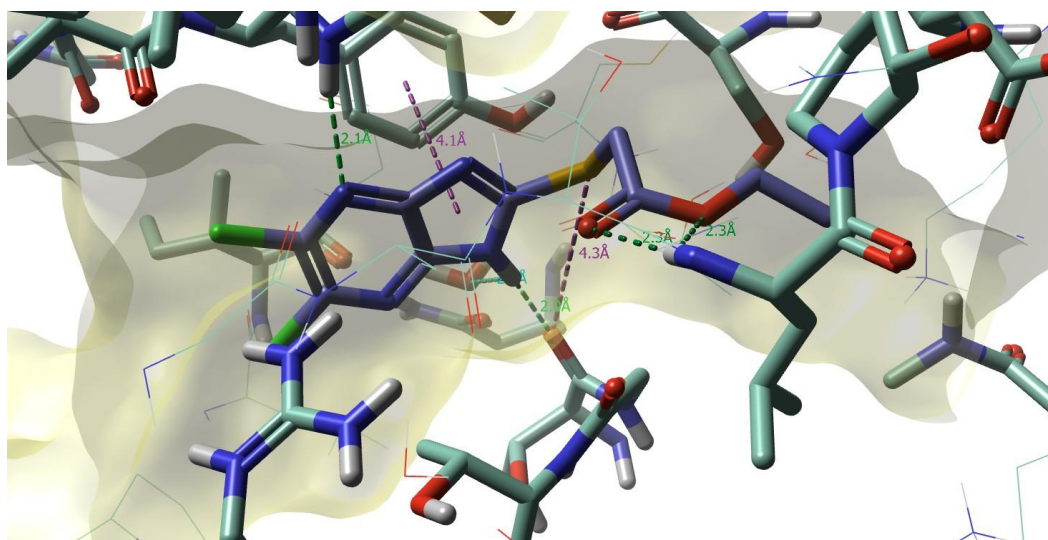
Figure 3-17: Ligand pose 10.

Based on the observations that pose 5 is the most buried ligand, has no steric clashes and only two unsatisfied heteroatoms present on the ligand I took this forward as the template for a plausible orientation of the ligand within the binding site. Several of the hydrogen bonds are occurring on the backbone of the protein structure which is favourable over side chain residue interactions. Lastly, the chlorine atoms pose a potentially favourable but weak halogen bonding opportunity for both electrophilic and nucleophilic interactors present in the bulk solution.

3.5.5.2 Redoxal

For NSC-73735 (Redoxal) 8 poses were generated. All of the ligand poses predicted orientations with huge amounts of steric clashes (orange lines) and very little favourable interactions observed. These observations draw concern because they imply redoxal is too large to actually fit into the binding site. Some of the data surrounding redoxal within Flare® raised some concerns that it would not be a reasonable starting point for further work as a hit structure. Its molecular weight is 484.5 gmol^{-1}, its ClogP = 6.3, suggesting it may provide solubility issues within the assay if made more lipophilic. Its predicted topological polar surface area (TPSA ($Å^2$)) being 117.1, (TPSA is often used as a metric for a molecules ability to permeate cells) as this calculated value is high, suggesting redoxal may have low permeability.

The interactions observed from the best ligand pose is described the as follows, from the outermost point of the proposed binding site to the innermost (Figure 3-18): Hydrogen bonding interactions between the phenol of Tyr581 and the carboxylic acid furthest out of the proposed binding site. Cation-π and π-π interactions between Arg803 and Tyr587 respectively and the left-most phenyl moiety of the benzidine core. Additionally, the inner most benzoic motif has three hydrogen bonding interactions with Arg803, Leu628 and Gly800.
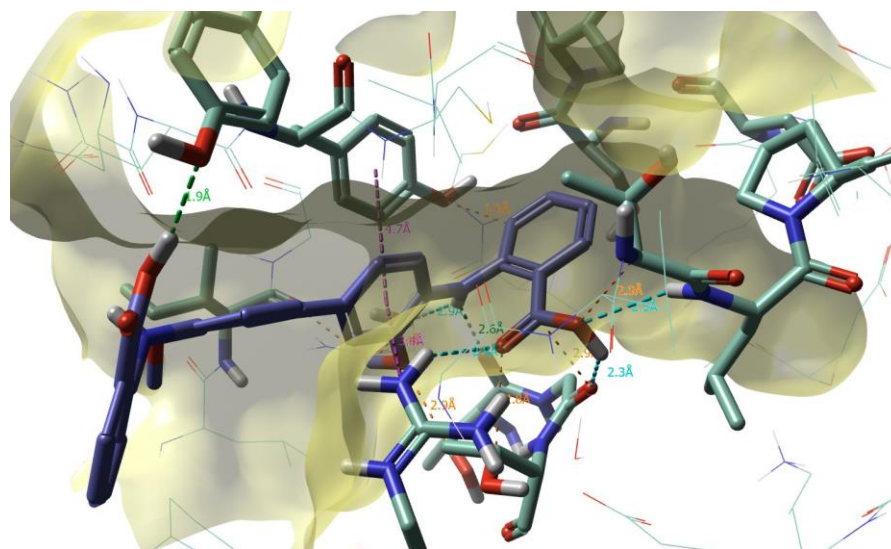
Figure 3-18: Proposed binding of redoxal in WWP2 proposed binding site.

For clarity the binding site has been flipped in Figure 3-19. The inner-most diphenylamine has two hydrogen bonding interactions with a conserved water molecule H_2O903 and Thr799. The proposed binding site has many orange interactions, which are steric or electronic clashes between the ligand and protein. This probably reflects the fact that redoxal is rather ill-fitting into this binding site due to its size. There is also a large void

space further into the binding site which is not filled (other proposed orientations of redoxal provided by Flare® show redoxal interacting within this binding site, but due to the numerous steric clashes this structure was not chosen as a reasonable binding mode).
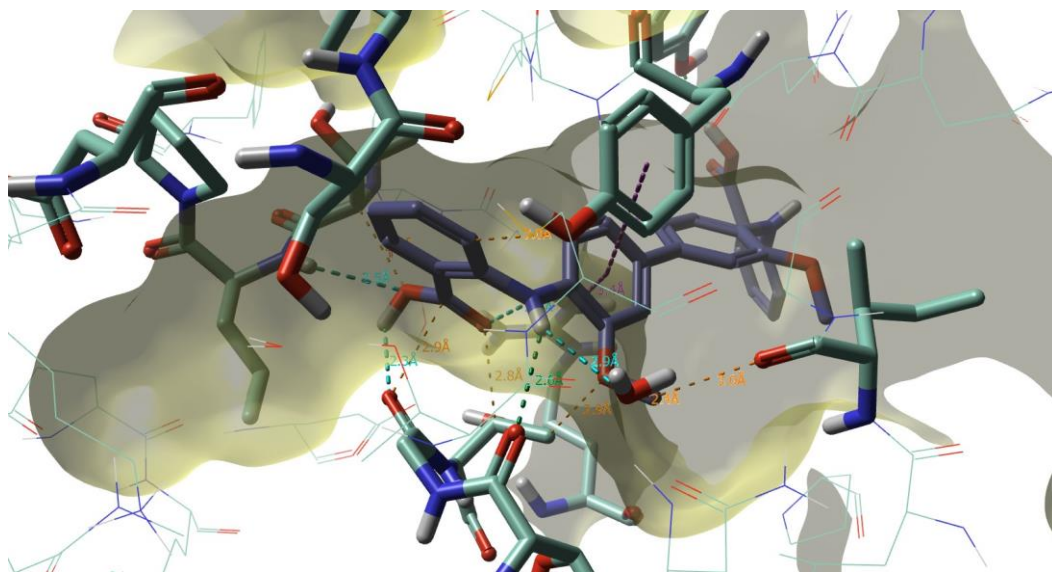
Figure 3-19: A flipped view of redoxal's proposed binding interactions.

These interactions are also tabulated (Table 3-6) with distances associated with each interaction:
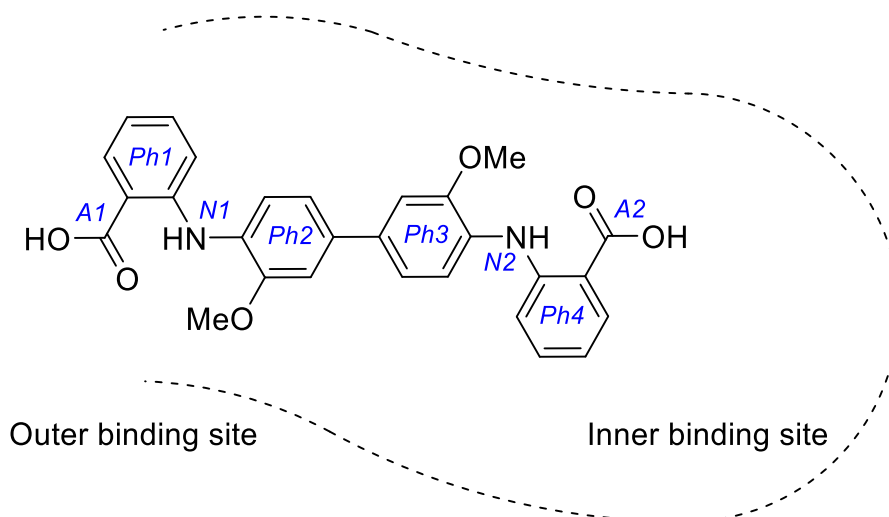
Redoxal	Residue	Interaction type*	Distance (Å)
A1	Tyr581	C-O – H-O	1.9
Ph2	Arg803	π(Ph2) – cation (guanidinium)	3.4

Ph2	Tyr587	π(Ph2) – π(Phenol)	4.7
A2	Leu628	C-O – H-N	2.5
A2	Gly800	O-H – H	2.3
A2	Arg803	C=O – H	2.9
N2	H2O903	N-H – O	2.9
N2	Thr799	N-H – O=C	2.6

Table 3-6: Proposed interactions between redoxal and WWP2 proposed binding site. *Left = redoxal interactor, right = protein residue interactor.

The avoidance of phenolic groups was sought in the redoxal series as it was thought this might lead to PAINs. The methoxy groups were replaced with aryl fluorine motifs, as it was assumed deprotection to phenols at later stages would be pursued for the compound series (*F*-redoxal, Figure 3-20, right).
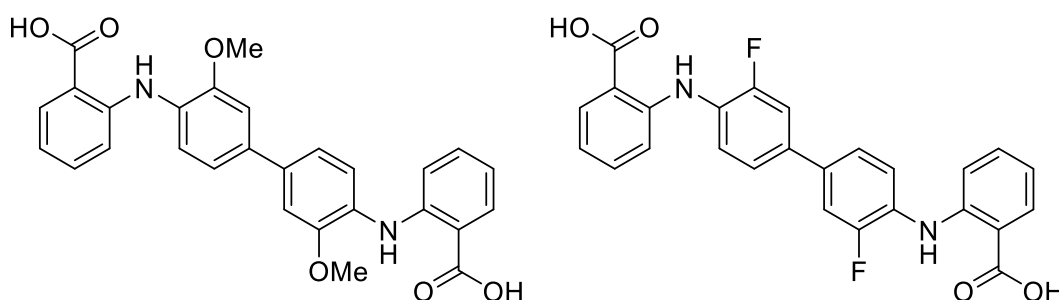
Figure 3-20: redoxal (left) and F-redoxal (right).

Additionally, because of the aforementioned concerns surrounding the size of redoxal, it was thought smaller fragments of *F*-redoxal would be better suited to start from. These would ideally fit better into the proposed binding site without many additional steric or electronic clashes. Additionally, starting from a smaller structure would mean a higher likelihood of producing a molecule with high activity. Usually, improvements to activity come with increases in molecular weights by the addition of interacting functionality, so starting from smaller structures based on *F*-redoxal meant that growth of the structure would not border on or exceed 500 gmol^{-1}. Therefore, *F*-redoxal was taken forward with the intention of Identifying a structural fragment which may have similar activity compared to the original hit, and not based on molecular docking.

3.6 <u>Summation of Calculation Method</u>

Crystallised Human WWP2 HECT domain structure was selected and was retrieved from the protein data bank.[131] Protein crystal structure preparation was done using Cresset Flare software and the water molecules (see section 3.5.2 for exception) were removed. The grid box for docking was defined according to the proposed binding site as discussed in section 3.5.2. Ligands were sketched in the software and their energy minimised, the docking calculations were run in *accurate but slow mode* and default settings.[156] Visual inspection based on the criteria outlined in section 3.5.3 followed, with the virtual analogues with the best ligand poses taken forward to synthesis and biological evaluation.

3.7 <u>Limitations of Visual Inspection</u>

Although the visual inspection of ligand poses has advantages over the use of scoring functions generated by molecular docking software, it still has several drawbacks. Obviously, the experience, intuition and number of experts within the team participating in the visual inspection will have a large impact on the decision making process.[149] A second point is the number of poses that may be investigated visually in a given amount of time. This has greater impact on projects that are investigating many compounds in, for example, virtual screening. Since the use of molecular docking in this project pertains only to the production of several small, focussed libraries of similar structures this issue is less impactful within this work. However, the ability to consistently evaluate ligand poses generated within this project may have become an issue.

Additionally, using visual inspection it is difficult to summarise desolvation free energy and loss of configurational entropy of the ligand upon binding. Scoring functions primarily evaluate enthalpic contributions of binding, and entropic contributions are mostly limited to evaluating ligand desolvation as hydrophobic contacts and configurational entropy as the number of rotatable bonds.[149] As visual inspection cannot properly investigate these entropic contributions, a more accurate summation of the binding thermodynamics is required, usually by coupling with more computationally demanding simulations like molecular dynamics and free energy perturbation.

Protein flexibility plays an important role in the binding of ligands, and even though there are programmes available that can treat proteins as flexible or partially flexible (*e.g.* ensemble docking) most software available treat the protein as a rigid structure.[149] Therefore, visual inspection is based upon this inflexible view of the ligand-protein

complex, similar to a crystal structure, and this serves as a limitation of visual inspection predicated on the limitations of simplified calculations.

Nonetheless, alternatives to using visual inspection to decide upon favourable ligand poses include the use of imperfect scoring functions as described above or require the use of more computationally and time demanding software. Ignoring the use of computational tools in drug design limits the strategy to that of the medicinal chemists' intuition alone and away from rational drug design. Visual inspection offers advantages over scoring function in terms of accuracy of the plausibility of the ligand poses and is not computationally or time expensive to perform for this project. While being imperfect it offers some structural based results compared to intuition alone, so it seemed reasonable to utilise this specific prediction method as part of this project design.

3.8 <u>Experimental</u>

3.8.1 <u>General Experimental</u>

Applies to all experimental sections described in this thesis. Anhydrous THF and Et$_2$O was freshly distilled from sodium and benzophenone. Anhydrous dichloromethane was freshly distilled from calcium hydride powder. Anhydrous pyridine was prepared by fractional distillation over activated 4 Å molecular sieves. Anhydrous toluene, dioxane (1,4-dioxane) and *N,N*-dimethylformamide were prepared by storing over activated 4 Å molecular sieves for at least 24 h. All other solvents and reagents were purchased and used as supplied unless otherwise specifically stated. Thin-layer chromatography was performed on Merck silica gel 60 F254 plates and visualised by UV absorption, purchased from VWR International. Flash column chromatography was carried out using Silica Gel 60 purchased from Material Harvest. 'Concentrated' refers to the removal of volatile organic solvents *via* distillation using a rotary evaporator. 'Dried' refers to pouring onto or adding anhydrous MgSO$_4$ or Na$_2$SO$_4$ to (as specified), followed by filtration. Water refers to deionised water.

Unless specified, all reagents and starting materials were purchased from commercial sources (Sigma-Aldrich (Merck Life Sciences), Fluorochem (Doug Discovery), Fischer Scientific, Alfa Aesar) and used as received. NMR spectra were recorded on 400 or 500 MHz Bruker NMR spectrometer using the deuterated solvent stated in the reported data. ^{1}H, ^{13}C, ^{19}F and ^{31}P NMR samples were prepared by dissolving compound in 0.4 mL – 0.7 mL deuterated solvent. All deuterated solvents were purchased from Cambridge Isotopes and used as received, solvents were stored under 4 Å molecular sieves after opening. All spectra were referenced to the residual solvent peaks of the solvent used.[157] NMR spectra chemical shifts (δ) are reported in ppm and coupling constants (J) reported in hertz (Hz). Abbreviations for NMR splitting are: s (singlet), d (doublet), t (triplet), q (quartet), p (pentet) and m (multiplet). Infrared spectra were recorded using a Perkin Elmer Spectrum Two LITA. High resolution mass spectrometry was performed at the University of East Anglia using a UPLC-HRMS (ACQUITY H-Class PLUS UPLC and Waters SYNAPT XS High Resolution Mass Spectrometer) setup with electrospray ionisation using ca. 1 µgmL^{-1} solution in acetonitrile or methanol. Melting points were recorded on a Büchi Melting Point B-545 using capillary melting point tubes made in house. Rigaku XtaLAB Synergy-S single crystal instrument was used to obtain X-ray crystal structures. The acquisition and structural model refining was performed either by Dr Benjamin Hofmann or Dr Claire Jones using Olex2 processing software.

N-[3-(2-pyridyl)isoquinolin-1-yl]-2-pyridinecarboxamidine[132]
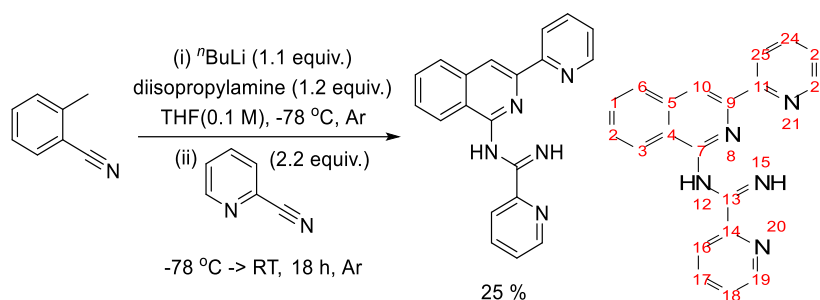
To an oven-dried 3-neck RBF was added under argon diisopropylamine (0.4 mL, 1.1 equiv.) and anhydrous THF (25.6 mL, 0.1 M), the mixture was cooled to -78 °C. To a separate oven dried vial was added under argon 2-cyanopyridine (0.542 mL, 2.2 equiv.) and anhydrous THF (1 mL). To the cooled diisopropylamine solution was added dropwise ^{n}BuLi (2.5 M hexanes, 1.23 mL, 1.2 equiv.) and allowed to stir for 10 min. *o*-Tolunitrile (0.303 mL, 2.56 mmol) was then added dropwise to the cooled solution, with a deep-red colour appearing, stirred for 10 min. The 2-cyanopyridine solution was then transferred *via* syringe dropwise over 5 min. The reaction was then allowed to warm slowly to RT and stirred at RT for 18 h. The mixture was quenched with NH_4Cl sat. soln. (6 mL) and THF removed *in vacuo.* Et_2O (30 mL) was added, and the layers separated. HCl / Et_2O solution was added to the organic phase until no more dark oil separated from the solution (allowed to stand for 20 mins). The Et_2O was then decanted off and the oil was washed with more Et_2O (10 mL) and decanted. To the oil was added NaOH (0.5 M, 20 mL) and Et_2O (20 mL), the mixture was transferred to a sep. funnel, separated and the aqueous layer washed with Et_2O (20 mL). The organic layers were combined, dried ($MgSO_4$) and solvent removed *in vacuo.* The residue was purified by column chromatography (eluting with 8:2 Pet. E : EtOAc) to provide first a light-yellow solid of *N*-[3-(2-methylphenyl)isoquinolin-1-yl]-2-pyridinecarboxamidine and an orange solid. The orange solid was crystallised using $CHCl_3$: Hex. to provide *N*-[3-(2-pyridyl)isoquinolin-1-yl]-2-pyridinecarboxamidine as a light yellow crystalline solid (211 mg, 0.649 mmol, 25%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 10.96 (br s, 1H, **15**), 9.02 (d, *J* = 8.7 Hz, 1H, **19**), 8.84 (dt, *J* = 7.9, 1.1 Hz, 1H, **16**), 8.76 (ddd, *J* = 4.8, 1.9, 0.9 Hz, 1H, **22**), 8.67 (dt, *J* = 4.7, 1.3 Hz, 1H, **25**), 8.42 (d, *J* = 0.8 Hz, 1H, **10**), 8.25 (br dt, *J* = 7.9, 1.1 Hz, 2H, **3**, **12**, br s overlapping with dt, confirmed by D_2O shake), 7.92 (td, *J* = 7.7, 1.7 Hz, 1H, **6**), 7.90 (d, *J* = 7.9 Hz, 1H, **1**), 7.86 (td, *J* = 7.7, 1.8 Hz, 1H, **2**), 7.69 (ddd, *J* = 8.1, 6.8, 1.4 Hz, 1H, **17**), 7.62 (ddd, *J* = 8.3, 6.9, 1.4 Hz, 1H, **18**), 7.45 (ddd, *J* = 7.6, 4.8, 1.2 Hz, 1H, **24**), 7.31 (ddd, *J* = 7.5, 4.8, 1.2 Hz, 1H, **23**).

^{13}C NMR (101 MHz, CDCl$_3$) δ 160.8, **13**, 157.2, **7**, 156.2, **14**, 152.5, **11**, 149.6, **22**, 148.3, **25**, 147.1, **9**, 138.2, **5**, 137.04, **6**, 136.99, **2**, 130.4, **17**, 127.5, **1**, 127.3, **4**, 127.0, **18**, 126.8, **19**, 125.6, **24**, 123.1, **23**, 122.4, **16**, 120.7, **3**, 113.9, **10**.

M.P. 185.8 – 187.7 ^{o}C,

IR (cm^{-1}): 3374 (NH), 3059 (NH), 1614 (C=N). Data is in line with literature data.

N-[3-(2-methylphenyl)isoquinolin-1-yl]-2-pyridinecarboxamidine
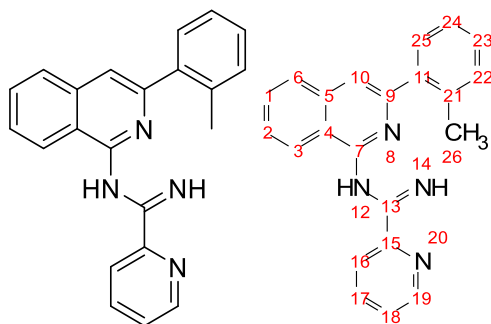
Isolated from the above reaction as a yellow solid, (44.1 mg, 0.13 mmol, 5%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 11.08 (br s, 1H, *14*), 9.06 (d, *J* = 8.3 Hz, 1H, *19*), 8.83 (dt, *J* = 8.0, 1.1 Hz, 1H, *3*), 8.64 (dd, *J* = 4.6, 1.5 Hz, 1H, *6*), 8.10 (br s, 1H, *12*), 7.91 (td, *J* = 7.7, 1.7 Hz, 1H, *2*), 7.78 (d, *J* = 8.1 Hz, 1H, *16*), 7.69 (ddd, *J* = 8.1, 6.8, 1.4 Hz, 1H, *17*), 7.62 (ddd, *J* = 8.2, 6.9, 1.4 Hz, 1H, *18*), 7.59 – 7.54 (m, 1H, *22*), 7.43 (ddd, *J* = 7.4, 4.8, 1.2 Hz, 1H, *1*), 7.40 (s, 1H, *10*), 7.37 – 7.30 (m, 3H, *23*, *24*, *25*), 2.47 (s, 3H, *26*).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 160.4, *13*, 156.4, *7*, 152.6, *15*, 150.6, *9*, 148.3, *6*, 141.6, *4*, 137.9, *5*, 137.0, *2*, 136.2, *11*, 130.8, *25*, 130.0, *17*, 130.0, *22*, 127.9, *24*, 126.7, *19*, 126.5, *16*, 126.4, *18*, 125.9, *23*, 125.5, *1*, 122.4, *3*, 115.9, *10*, 20.8, *26* (21 out of 22 carbon resonances found).

M.P. 133.8 - 135.2 °C,

IR (cm^{-1}): 3412 (NH), 3049 (C(sp^2)-H), 2923 (C-(sp^3)-H), 1611 (C=N).

MS ES+ m/z Calcd for C$_{22}$H$_{18}$N$_4$ (M+H)$^+$: 339.1610, found: 339.1602.

Crystals suitable for XRD were produced by dissolving *N*-[3-(2-methylphenyl)isoquinolin-1-yl]-2-pyridinecarboxamidine in acetone and leaving overnight for the solvent to evaporate. Crystal data available in Appendix 1.
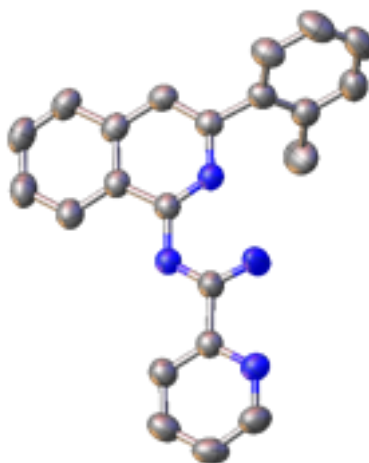
Methyl 2-cyano-2-hydroxyiminoacetate[158]
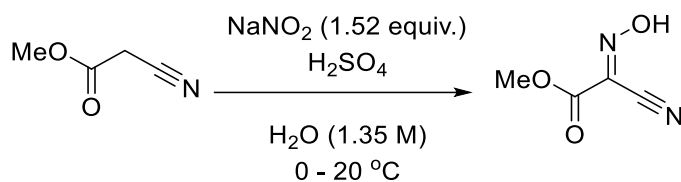
To a 500 mL 3-necked RBF was added NaNO$_2$ (24.0 g, 0.35 mol) dissolved in H$_2$O (60 mL). Methyl cyanoacetate (20.65 mL, 0.23 mol) was added *via* syringe and the mixture vigorously stirred to provide an emulsion which was cooled to 0 °C. H$_2$SO$_4$ (10 mL, 95% in 110 mL H$_2$O) was added dropwise with stirring *via* dropping funnel to maintain temperature below 20 °C. After addition the mixture solidified and more H$_2$SO$_4$ solution was added with agitation to break up the formed mass. The mixture was then cooled to 0 °C and the solid formed was filtered to provide white solid. The filtrate was extracted with Et$_2$O (6x 100 mL), dried (Na$_2$SO$_4$) and solvent removed to provide white crystals which were combined with the previous and dried under high vacuum for 24 h. Provided 27.40 g (0.21 mol, 93%) of white crystalline methyl 2-cyano-2-hydroxyiminoacetate.

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 11.17 (s, 1H), 3.99 (s, 3H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 158.9, 126.8, 107.4, 54.3.

IR (cm^{-1}): 3216 (OH), 2238 (C≡N), 1730 (C=O).

Adapted from literature procedure, data in line with literature data.

Methyl 2-cyano-2-*N*-tosyloxyiminoacetate[140]
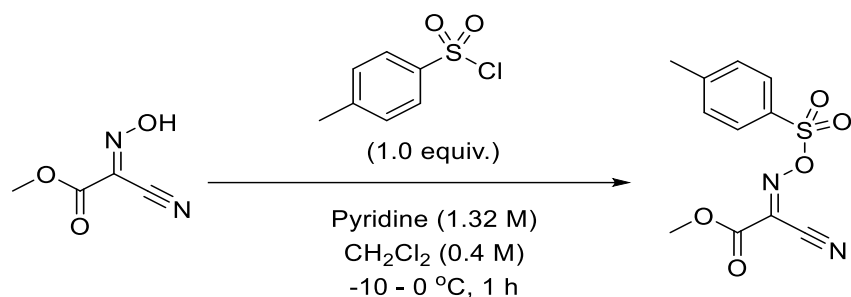
To a 2-necked 50 mL RBF with a thermometer and dropping funnel attached was added methyl 2-cyano-2-hydroxyiminoacetate (1.0 g, 7.8 mmol) and pyridine (6 mL, 1.32 M). The mixture was cooled to -10 °C. To the dropping funnel was added tosyl chloride (1.488 g, 1.0 equiv.) and DCM (15 mL). This was added dropwise over 15 min. to ensure the temperature did not rise above 0 °C. After addition the addition funnel was washed with additional DCM (15 mL) mixture was allowed to stir between -5 – 0 °C for 45 min. Water (6 mL) was added and the mixture transferred to a separatory funnel, washing with more water (10 mL). The layers were separated, and the aqueous layer extracted with DCM (25 mL). The organic layers were combined, dried (MgSO$_4$) and solvent removed under reduced pressure. To the residue was added water (220 mL) and stirred at 0 °C for 1 h. The formed suspension was then filtered and washed with water and then a small volume of methanol. Dried under high vacuum to provide methyl 2-cyano-2-*N*-tosyloxyiminoacetate as a white powder (1.18 g, 4.18 mmol, 54%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 7.96 – 7.89 (m, 2H), 7.45 – 7.38 (m, 2H), 3.96 (s, 3H), 2.48 (s, 3H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 156.7, 147.5, 131.0, 130.4, 130.3, 129.7, 106.2, 54.8, 22.0.

IR (cm^{-1}): 2966 (C(sp^3)-H), 2238 (C≡N), 1741 (C=O).

May be repeated on a 10 g scale with no change in yield. Adapted from literature procedure, data in line with literature data.

Methyl 2-(morpholino-2'-cyclohexene-1'-ylimino)-2-cyanoacetate[140]
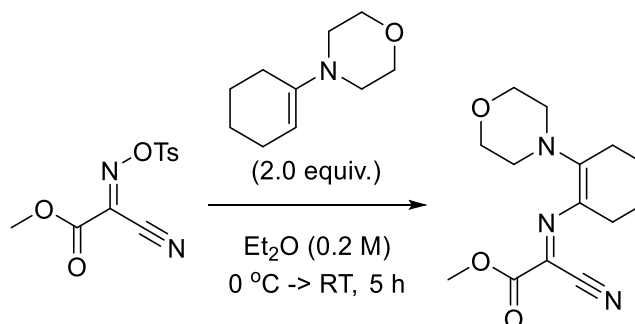
To a 500 mL RBF surmounted with a dropping funnel was added methyl 2-cyano-*N*-tosyloxyiminoacetate (11.0 g, 38.97 mmol) and anhydrous Et$_2$O (194 mL, 0.2 M) under argon, cooled to 0 °C. To a dropping funnel was added 1-morpholinocyclohexene (13.10 mL, 2.0 equiv.) and this was added dropwise over 10 min. The dropping funnel was washed with Et$_2$O (10 mL), and this was added to the reaction mixture. After addition the reaction was allowed to warm to RT and stirred under argon for 5 h. The suspension was filtered and washed with Et$_2$O, then MeOH. Provided methyl 2-(morpholino-2'-cyclohexene-1'-ylimino)-2-cyanoacetate as an orange powder (9.59 g, 34.56 mmol, 89%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 3.96 – 3.91 (m, 4H), 3.87 (ddd, *J* = 6.8, 3.3, 1.1 Hz, 4H), 3.82 (s, 3H), 2.91 (t, *J* = 6.3 Hz, 2H), 2.53 (t, *J* = 6.2 Hz, 2H), 1.82 – 1.65 (m, 4H).

^{13}C NMR (126 MHz, C*D*Cl$_3$) δ 165.3, 162.9, 124.5, 118.2, 97.0, 67.7, 52.6, 52.4, 31.6, 28.4, 22.5, 22.0.

M.P. 124.2 – 125.2 °C.

IR (cm^{-1}): 2935 (C(sp^3)-H), 2172 (C≡N), 1685 (C=O).

MS ES+ m/z Calcd for: C$_{14}$H$_{19}$N$_3$O$_3$ (M+H)$^+$: 278.1505, found (0.9434 average increase): 279.0928.

Followed literature procedure.

1-*N*-(2'-Methoxyethyl)-2-amino-3-methoxycarbonyl-5,6,7-trihydroquinoxaline[137]
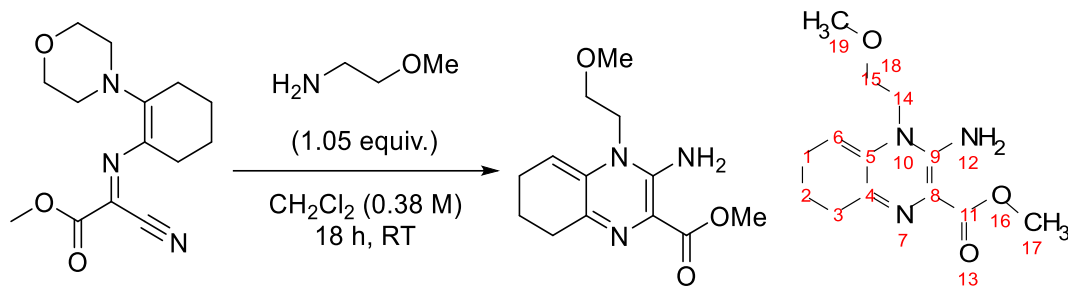
To a 25 mL RBF was added methyl 2-(morpholino-2'-cyclohexene-1'-ylimino)-2-cyanoacetate (700 mg, 2.52 mmol) and DCM (6.63 mL, 0.38 M). To the stirred solution was added 2-methoxyethylamine (0.23 mL, 1.05 equiv.). Allowed to stir at RT for 18 h. The reaction mixture was concentrated under reduced pressure and dried under high vacuum. The residue was recrystallised from EtOH to provide 1-*N*-(2'-methoxyethyl)-2-amino-3-methoxycarbonyl-5,6,7-trihydroquinoxaline as a yellow solid. (226 mg, 0.85 mmol, 41%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 7.19 (s, 2H, **12**), 4.63 (t, J = 4.6 Hz, 1H, **6**), 3.80 (s, 3H, **17**), 3.74 (t, J = 4.4 Hz, 2H, **14**), 3.63 (t, J = 4.4 Hz, 2H, **15**), 3.40 (s, 3H, **19**), 2.59 – 2.51 (m, 2H, **3**), 2.26 (td, J = 6.0, 4.6 Hz, 2H, **1**), 1.77 (p, J = 6.2 Hz, 2H, **2**).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 168.4, **11**, 152.7, 144.8, **4**, 133.4, **5**, 99.1, **6**, 70.4, **15**, 59.4, **19**, 51.4, **17**, 47.6, **14**, 34.4, **3**, 25.1, **1**, 21.6, **2**.

M.P. 149.8 – 151.2 $^{\circ}$C (dec.),

IR (cm^{-1}): 3382 (NH), 3199 (NH), 2944 (C(sp^3)-H), 1624 (C=O).

MS ES+ m/z Calcd for C$_{13}$H$_{19}$N$_3$O$_3$ (M+H)$^+$: 266.1505, found: 266.1501.

Adapted from literature procedure.

3-hydroxy-5,6,7,8-tetrahydro-4*H*-quinoxaline-2-(*N*-2-methoxyethyl)carboxamide
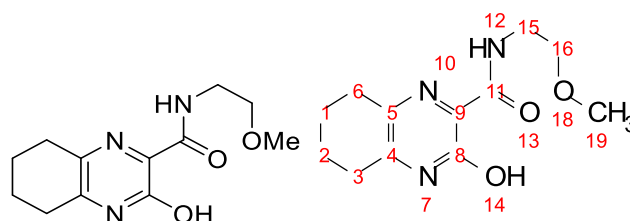
3-Hydroxy-5,6,7,8-tetrahydro-4*H*-quinoxaline-2-(*N*-2-methoxyethyl)carboxamide was isolated from the mother liquor of 1-*N*-(2'-methoxyethyl)-2-amino-3-methoxycarbonyl-5,6,7-trihydroquinoxaline purification upon standing for several days as brown crystals, which was further recrystallised (EtOH) to provide 3-oxo-5,6,7,8-tetrahydro-4*H*-quinoxaline-2-(*N*-2-methoxyethyl)carboxamide as a white crystalline solid.

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 12.38 (br s, 1H, *14*), 8.09 (br s, 1H, *12*), 3.73 – 3.61 (m, 2H, *16*), 3.61 – 3.53 (m, 2H, *15*), 3.40 (s, 3H, *19*), 2.96 – 2.88 (m, 2H, *3*), 2.87 – 2.79 (m, 2H, *6*), 1.95 – 1.84 (m, 4H, *1*, *2*).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 167.7, *11*, 159.8, 157.3, *4*, 143.9, *5*, 71.1, *16*, 59.1, *19*, 39.2, *15*, 32.4, *3*, 31.0, *6*, 22.8, *2*, 22.4, *1* (11 out of 12 carbon resonances observed).

M.P. 144.3 – 146.1 °C.

IR (cm^{-1}): 3286 (OH), 1672 (C=O), 1641 (NH bend).

MS ESI+ m/z Calcd for C$_{12}$H$_{17}$N$_3$O$_3$ (M+H)$^+$: 252.1348, found: 252.1348.

Crystals suitable for x-ray analysis were obtained by a further recrystallisation from EtOH as white crystals. See appendix 1 for details.
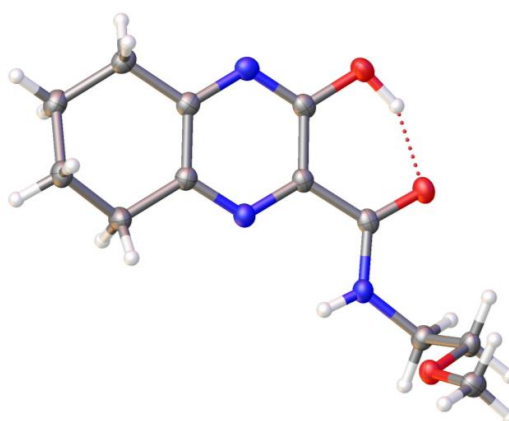
4-(2-methoxy-ethyl)-3-(phenylcarbamoyl-imino)-3,4,5,6,7,8-hexahydro-quinoxaline-2-carboxylic acid methyl ester[137]
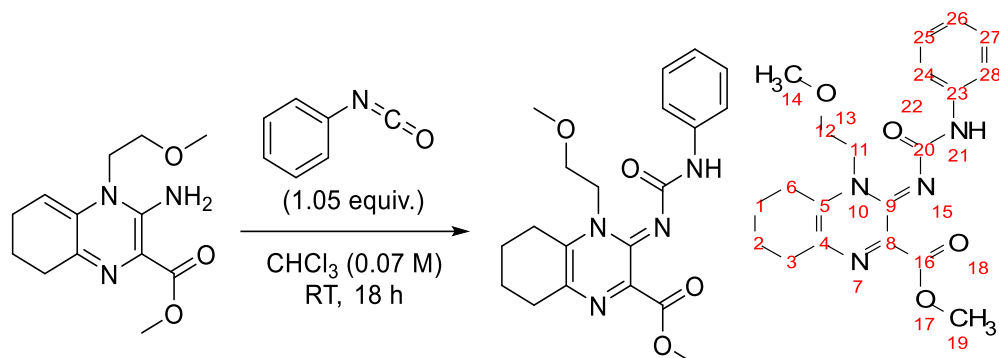
To a 250 mL RBF was added 1-*N*-(2'-methoxyethyl)-2-amino-3-methoxycarbonyl-5,6,7-trihydroquinoxaline (1.80 g, 6.78 mmol) and chloroform (97.0 mL, 0.07 M). To the stirred solution was added phenyl isocyanate (0.774 mL, 1.05 equiv.) *via* syringe. The solution was stirred at RT 18 h. The solvent was removed, and the residue taken up in methanol, heated to boiling, and allowed to cool under a slight nitrogen flow. After crystallisation had started the material was kept in a fridge overnight. The crystallised material was isolated, triturated with hexane and dried. Provided 4-(2-methoxy-ethyl)-3-(phenylcarbamoyl-imino)-3,4,5,6,7,8-hexahydro-quinoxaline-2-carboxylic acid methyl ester as an orange solid (1.95 g, 5.07 mmol, 75%).

^{1}H NMR (400 MHz, Acetone-d_6) δ 8.42 (br s, 1H, **21**), 7.66 (m, 2H, **24**, **28**), 7.23 (m, 2H, **25**, **27**), 6.92 (m, 1H, **26**), 4.42 (t, *J* = 5.3 Hz, 2H, **11**), 3.76 (t, *J* = 5.3 Hz, 2H, **12**), 3.68 (s, 3H, **19**), 3.29 (s, 3H, **14**), 3.02 (t, *J* = 6.2 Hz, 2H, **6**), 2.70 (t, *J* = 6.2 Hz, 2H, **3**), 1.92 – 1.76 (m, 4H, **1**, **2**).

^{13}C NMR (101 MHz, Acetone-d_6) δ 165.2, 160.5, 143.0, 141.0, 139.0, 133.5, 129.3, 122.3, 118.8, 69.4, 59.2, 52.3, 47.3, 30.9, 27.6, 22.8, 22.5 (17 out of 18 carbon resonances observed).

M.P. 167.3 -169.0 °C (dec.),

IR (cm^{-1}): 3325 (NH), 1702 (C=O), 1639 (C=O, urea).

MS ES+ m/z Calcd for C$_{20}$H$_{24}$N$_4$O$_4$ (M+H)$^+$: 385.1876, found: 385.1873.

Procedure adapted from literature.

(Methoxy-2-ethyl)-10-phenyl-3-tetrahydro-6,7,8,9-iso-alloxazine[137]
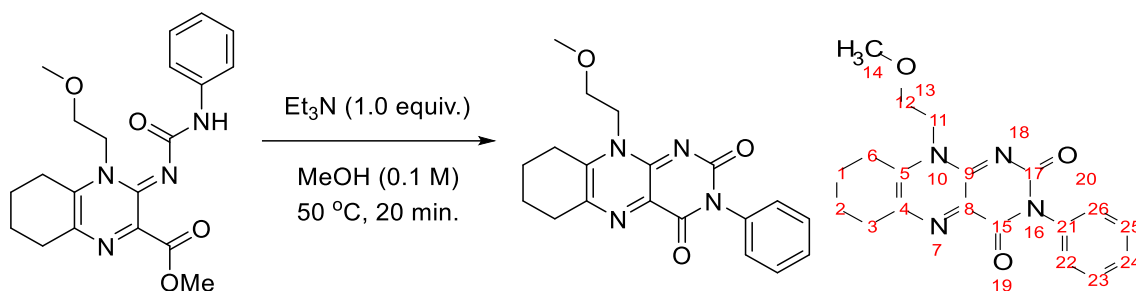
To an 8 mL vial was added 4-(2-methoxy-ethyl)-3-(phenylcarbamoyl-imino)-3,4,5,6,7,8-hexahydro-quinoxaline-2-carboxylic acid methyl ester (100 mg, 0.26 mmol), MeOH (2.60 mL, 0.1 M) and triethylamine (36.3 µL, 1.0 equiv.). The vial was sealed with a suba seal, and a needle was added as a vent. Mixture was heated with stirring at 50 °C for 20 min. The mixture was cooled, and solvent removed *in vacuo.* Extracted with EtOAc (3x 30 mL) from NH$_4$Cl solution (10 mL), organic layers collected and washed with brine (10 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* to provide a glassy orange solid. Attempted crystallisation from MeOH. Isolated 20.8 mg of (Methoxy-2-ethyl)-10-phenyl-3-tetrahydro-6,7,8,9-iso-alloxazine as a dark brown film (22%, 0.059 mmol).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 7.54 – 7.46 (m, 2H, **23**, **25**), 7.44 – 7.38 (m, 1H, **24**), 7.31 – 7.26 (m, 2H, **22**, **26**), 4.73 (t, *J* = 4.7 Hz, 2H, **11**), 3.95 – 3.88 (m, 2H, **12**), 3.30 (s, 3H, **14**), 3.16 (t, *J* = 6.3 Hz, 2H, **3**), 3.07 (t, *J* = 6.2 Hz, 2H, **6**), 2.02 – 1.86 (m, 4H, **1**, **2**).

Followed literature procedure.

NSC-288387, 10-(2-methoxyethyl)-3-phenylbenzo[g]pteridine-2,4(3*H*,10*H*)-dione[137]
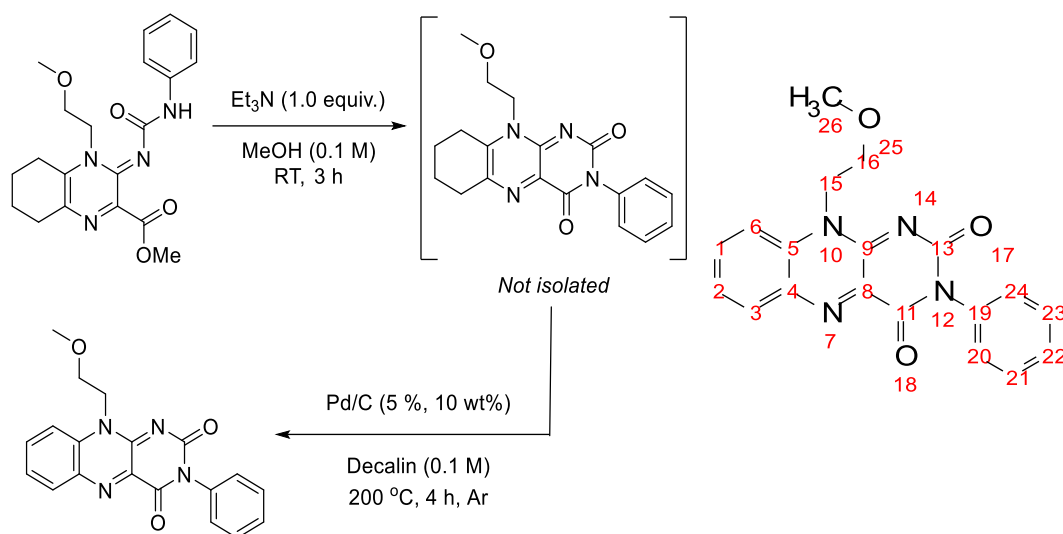
To a 3-necked 100 mL RBF was added 4-(2-methoxy-ethyl)-3-(phenylcarbamoyl-imino)-3,4,5,6,7,8-hexahydro-quinoxaline-2-carboxylic acid methyl ester (1.00 g, 2.30 mmol) and methanol (26.0 mL, 0.1 M). To the stirred suspension was added triethylamine (0.36 mL, 1.0 equiv.) and stirred for 3 h at RT. After the reaction was complete by TLC (EtOAc) and the solvent removed, reaction vessel placed on high vacuum to provide a black glassy solid. Crude ^{1}H NMR confirmed complete consumption of starting material. To the reaction vessel was surmounted an air condenser and added Pd/C (5%, 91.6 mg, 10 wt%) and cycled on a Schlenk line thrice (vacuum / argon). To the reaction vessel was added vigorously degassed *cis*-decalin (26.0 mL, 0.1 M) and the reaction was stirred and heated to 200 °C for 4 h. The reaction was allowed to cool to RT, filtered through a plug of Celite® and washed with hot methanol until no more colour came through. The filtrate was transferred to a separatory funnel and the decalin separated and washed with methanol (100 mL). The methanol layers were collected and concentrated. The residue was purified by column chromatography (8:2 EtOAc : DCM -> 7:2.9:0.1 EtOAc : DCM : MeOH -> 9:1 DCM : MeOH -> 8:2 DCM : MeOH) to provide 10-(2-methoxyethyl)-3-phenylbenzo[g]pteridine-2,4(3*H*,10*H*)-dione as an orange solid. This was recrystallised using MeOH to provide orange crystals (348 mg, 0.99 mmol, 43%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 8.31 (dd, *J* = 8.2, 1.6 Hz, 1H, **6**), 7.95 (dd, *J* = 8.8, 1.3 Hz, 1H, **3**), 7.88 (ddd, *J* = 8.8, 7.0, 1.6 Hz, 1H, **2**), 7.62 (ddd, *J* = 8.2, 7.0, 1.3 Hz, 1H, **1**), 7.57 – 7.48 (m, 2H, **21**, **23**), 7.49 – 7.41 (m, 1H **22**), 7.31 (dd, *J* = 8.3, 1.3 Hz, 2H, **20**, **24**), 4.96 (t, *J* = 5.1 Hz, 2H, **15**), 3.96 (t, *J* = 5.1 Hz, 2H, **16**), 3.31 (s, 3H, **26**).

¹³C NMR (101 MHz, C*D*Cl₃) δ 159.7, *13*, 155.3, *11*, 149.5, *9*, 137.4, *8*, 136.1, *4*, 135.7, *2*, 135.6, *19*, 134.0, *5*, 133.1, *6*, 129.5, *21*, *23*, 128.8, *22*, 128.3, *20*, *24*, 126.7, *1*, 117.0, *3*, 69.7, *16*, 59.4, *26*, 45.8, *15*.

M.P. 234.6 – 235.2 °C.

IR (cm⁻¹): 1672 (C=O, amide), 1641 (C=O amide).

MS ES+ m/z Calcd for $C_{19}H_{16}N_4O_3$ (M+H)$^+$: 350.1330, found: 350.1327.

Adapted from literature procedure. Crystals suitable for XRD were obtained by recrystallising from methanol over a period of one week (3 days at RT, 2 days in a fridge (5 °C) and one day in a freezer (-20 °C), collected by filtration and dried under high vacuum. See appendix 1 for details.
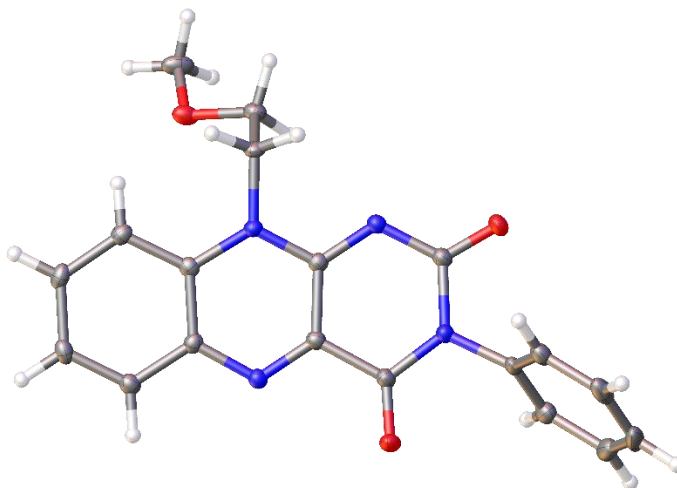
2,2'-dimethoxyazobenzene[159]
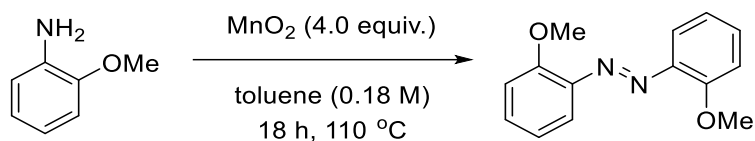
To a 1 L 3-necked RBF was attached a Dean-Stark apparatus surmounted by an air condenser. To the flask was added *o*-anisidine (9.15 mL, 0.081 mol) and toluene (450 mL, 0.18 M). To the stirred solution was added MnO_2 (28.2 g, 4.0 equiv.). With stirring the mixture was set to reflux for 18 h, with occasional removal of the collected water. The mixture was allowed to cool to RT and filtered twice through Celite®. The toluene was removed under reduced pressure. The residue was dissolved in methanol and treated with activated charcoal. Chloroform and acetone were used to aid filtration and were removed under reduced pressure. The methanolic mixture was then cooled in a fridge overnight and the crystals filtered to provide 2,2'-dimethoxyazobenzene (2.55 g, 10.52 mmol, 13%) as dark red crystals. Taken forward without further purification.

^{1}H NMR (500 MHz, C*D*Cl$_3$) δ 7.63 (dd, J = 8.2, 1.7 Hz, 2H), 7.42 (ddd, J = 8.4, 7.3, 1.7 Hz, 2H), 7.08 (dd, J = 8.4, 1.2 Hz, 2H), 7.00 (ddd, J = 8.2, 7.3, 1.2 Hz, 2H), 4.02 (s, 6H).

^{13}C NMR (126 MHz, C*D*Cl$_3$) δ 157.0, 143.1, 132.3, 121.0, 117.7, 112.7, 56.5.

IR (cm^{-1}): 3001, 2839, 1591.

Adapted from literature procedure, data in line with literature data.[141]

2,2'-dimethxoyhydrazobenzene[145]
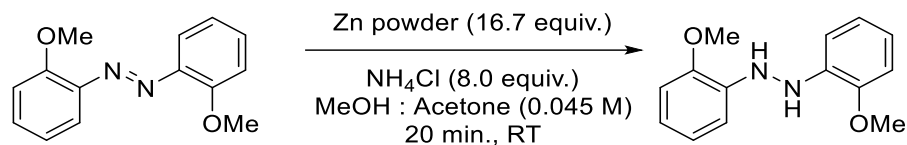
To a 1 L RBF was added 2,2'-dimethoxyazobenzene (3.7 g, 15.27 mmol), methanol : acetone (1:1, 340 mL, 0.045 M) and ammonium chloride (6.53 g). After dissolution, freshly-activated Zn powder (activated by washing with 2.0 M HCl and dried on vacuum filtration by sequential washing with water, methanol, acetone, anhydrous diethyl ether and drying on high vacuum, 16.67 g, 16.7 equiv.) was added. After the disappearance of the starting material colour, TLC indicated complete consumption of starting material and the mixture was poured into ice-water and left to stand for 10 min. The mixture was filtered, and the solid product was dissolved in hot methanol and collected in a separate flask, solvent removed under reduced pressure to provide a brown powder. Taken forward without further purification.

IR (cm^{-1}): 3412 (NH), 1591 (NH bend).

Adapted from literature procedure, data in line with literature data.[142]

3,3'-dimethoxybenzidine
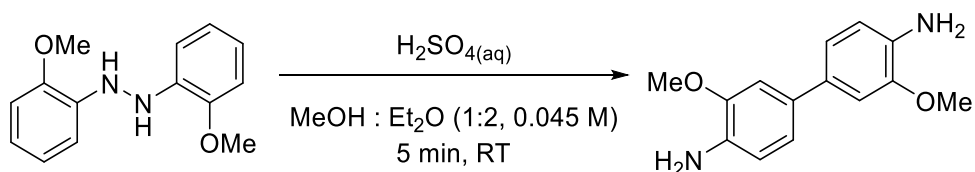
To a 500 mL RBF was added 2,2'-dimethoxyhydrazobenzene (1.59 g, 6.14 mmol), MeOH (68 mL) and Et$_2$O (136 mL). To the stirred solution was added *via* dropping funnel concentrated sulfuric acid in water (5 mL H$_2$SO$_4$ in 10 mL H$_2$O). Immediately a brown precipitate forms, and after addition, TLC indicated complete consumption of starting material (within 5 min. of addition). The precipitate was collected by filtration to provide a white solid. This solid was suspended in ethyl acetate in a separating funnel and washed with NaOH solution (1 M, 50 mL) and separated. The aqueous layer was extracted with ethyl acetate twice (100 mL). The organic layers were dried (MgSO$_4$), and solvent removed under reduced pressure to provide 3,3'-dimethoxybenzidine as a dark grey solid (596 mg, 2.44 mmol, 40%).

^{1}H NMR (500 MHz, C*D*Cl$_3$) *δ* 7.00 – 6.95 (m, 2H), 6.97 (s, 2H), 6.78 – 6.72 (m, 2H), 3.91 (s, 6H), 3.79 (br s, 4H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 147.7, 135.0, 132.8, 119.4, 115.3, 109.4, 55.7.

IR (cm^{-1}): 3429 (NH$_2$), 3399 (NH$_2$), 1618 (NH bend).

M.P. 138 °C (lit),[160] 135.8 – 137.5 °C.

Adapted from literature procedure.[161]

Dimethyl 2,2'-({3,3'-dimethoxy[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate
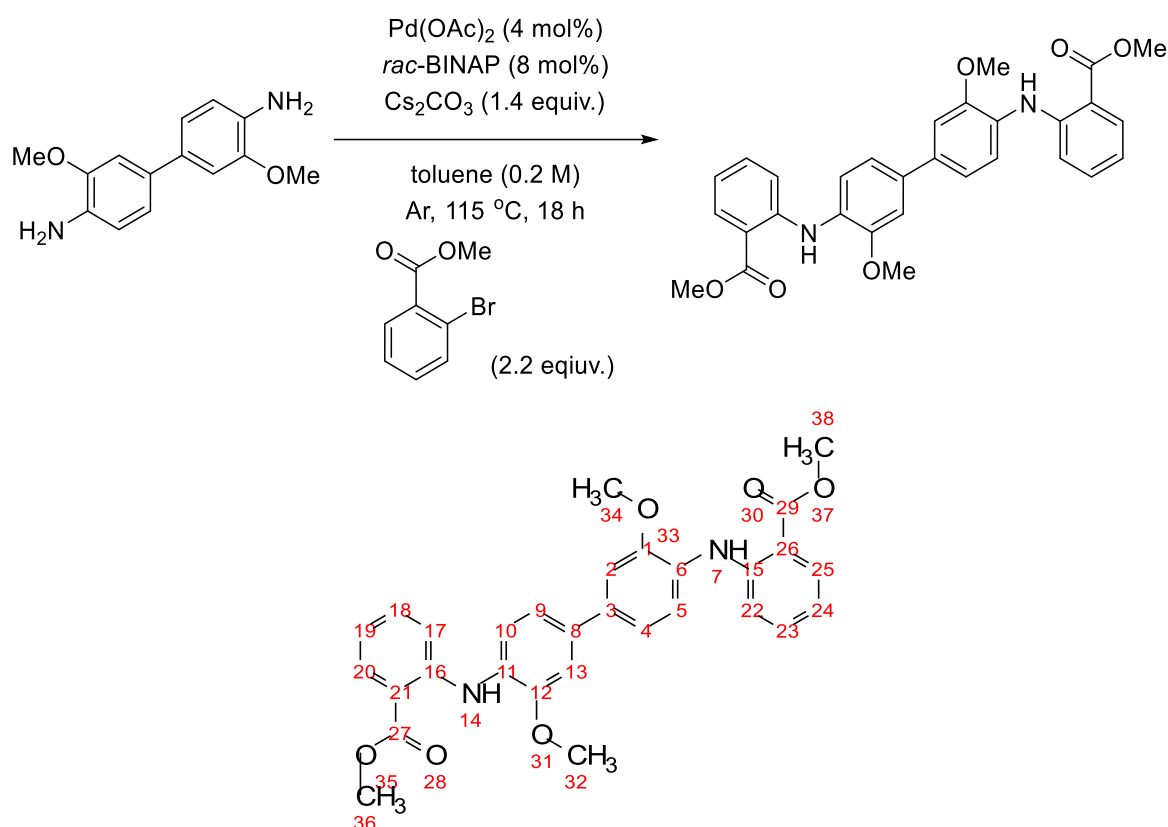
To an oven-dried 3-necked RBF was added under argon 3,3'-dimethoxybenzidine (500 mg, 2.04 mmol) and methyl 2-bromobenzoate (0.286 mL, 2.2 equiv.). Degassed toluene (10.23 mL, 0.2 M) was then added *via* syringe and further degassed for 5 min. Pd(OAc)$_2$ (18.3 mg, 4 mol%), *rac*-BINAP (0.101 g, 8 mol%) and Cs$_2$CO$_3$ (0.933 g, 1.4 equiv.) were then added altogether under an argon flow, the flask was then heated (115 ^{o}C) under argon with stirring for 18 h. The mixture was allowed to cool and filtered through a Celite® pad, washing with EtOAc. The filtrate was concentrated under reduced pressure and subjected to column chromatography (9:1 Tol : EtOAc) to provide a yellow solid which was recrystallised from MeCN to provide methyl 2-([3,3'-dimethoxy-4'-amino-[1,1'-biphenyl]-4-yl]amino)benzoate as an orange crystalline solid (427 mg, 0.83 mmol, 41%).

^{1}H NMR (400 MHz, DMSO-d_6) *δ* 9.49 (br s, 2H, **7, 14**), 7.92 (dd, *J* = 8.0, 1.7 Hz, 2H, **20, 25**), 7.47 (d, *J* = 8.3 Hz, 3H, **5, 10**), 7.46 (ddd, *J* = 8.1, 7.1, 1.7 Hz, 2H, **18, 23**), 7.37 (d, *J* = 2.0 Hz, 2H, **2, 13**), 7.33 (dd, *J* = 8.1, 1.1 Hz, 2H, **17, 22**), 7.29 (dd, *J* = 8.3, 2.0 Hz, 2H, **4, 9**), 6.82 (ddd, *J* = 8.0, 7.1, 1.1 Hz, 2H, **19, 24**), 3.97 (s, 6H, **32, 34**), 3.88 (s, 6H, **36, 38**).

^{13}C NMR (101 MHz, DMSO-d_6) *δ* 168.0, **27, 29**, 150.9, **1, 12**, 146.1, **21, 26**, 135.3, **6, 11**, 134.5, **18, 23**, 131.4, **20, 25**, 128.3, **3, 8**, 119.8, **5, 10**, 118.6, **4, 9**, 117.6, **19, 24**, 114.1, **17, 22**, 112.1, **15, 16**, 109.9, **2, 13**, 55.9, **32, 34**, 52.0, **36, 38**.

M.P. 170.4 – 171.7 °C.

IR (cm^{-1}): 3323 (NH), 1689 (C=O), 1586 (NH bend).

Adapted from literature procedure.[146]

A second fraction (monoarylated product) was collected as a yellow solid (83.2 mg, 0.22 mmol, 10%). Crystals suitable for XRD were obtained by recrystallising from a large volume of MeCN/H$_2$O overnight and drying *in vacuo.* See appendix 1 for details.
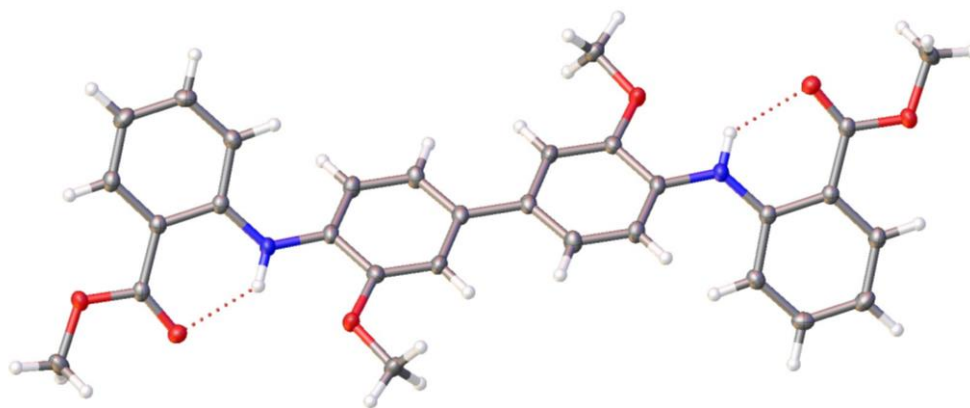
2,2'-({3,3'-dimethoxy[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoic acid – Redoxal
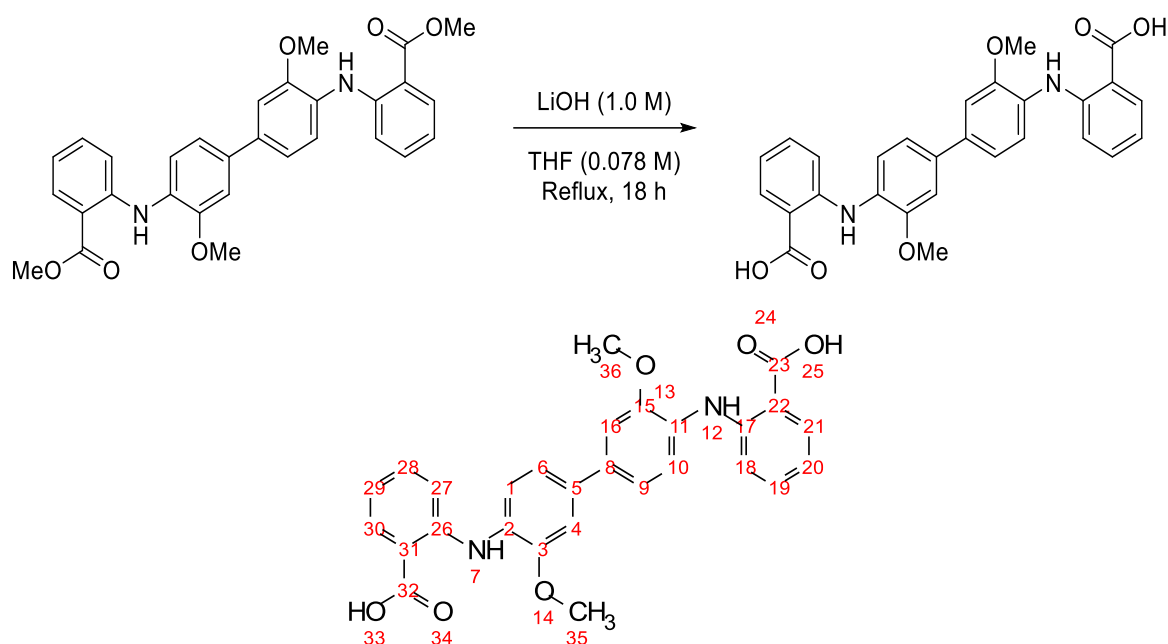
To an 8 mL vial was added dimethyl 2,2'-({3,3'-dimethoxy[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate (200 mg, 0.39 mmol) and lithium hydroxide (24 mg, 1.0 mmol). THF:H_2O (5:1, 0.065 M) was then added and the mixture heated to reflux for 18 h. The mixture was allowed to cool, and solvent removed *in vacuo.* The residue was dissolved partially in H_2O, transferred to a separatory funnel, washing with EtOAc to aid transfer. Extracted with EtOAc (3x 100 mL). The organic layers were collected and dried ($MgSO_4$) and solvent removed. Further drying required the compound to be placed in a 100 °C oil bath *in vacuo.* for 2 days. Provided 2,2'-({3,3'-dimethoxy[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoic acid as a yellow solid (185 mg, 0.38 mmol, 98%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 12.81 (br s, 2H, **25**, **33**), 9.75 (br s, 2H, **7**, **12**), 7.92 (dd, J = 8.0, 1.7 Hz, 2H, **21**, **30**), 7.46 (d, J = 8.2 Hz, 2H, **1**, **10**), 7.40 (ddd, J = 8.6, 7.0, 1.7 Hz, 2H, **19**, **28**), 7.35 (d, J = 2.0 Hz, 2H, **4**, **16**), 7.28 (m, 4H, **6**, **9**, **18**, **27**), 6.78 (ddd, J = 8.0, 7.0, 1.1 Hz, 2H, **20**, **29**), 3.95 (s, 6H, **35**, **36**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 170.0, **23**, **32**, 151.0, **3**, **15**, 146.3, **17**, **26**, 135.0, **5**, **8**, 133.9, **19**, **28**, 131.9, **21**, **30**, 128.6, **2**, **11**, 119.8, **1**, **10**, 118.6, **6**, **9**, 117.3, **20**, **29**, 113.8, **18**, **27**, 109.9, **4**, **16**, 55.9, **35**, **36** (13 out of 14 carbon resonances observed).

M.P. approx. 300 °C (lit.),[162] 270 °C+ (dec.).

IR (cm^{-1}): 3312 (OH), 1672 (C=O), 1585 (NH bend).

MS ES+ m/z Calcd for $C_{28}H_{24}N_2O_6$ (M+H)$^+$: 485.1713, found: 485.1712.

2,2'-difluoroazobenzene[141]
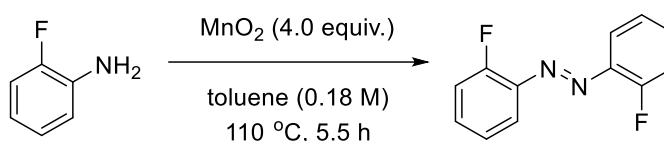
To a 1 L 3-necked RBF was added a stirrer bar and toluene (500 mL, 0.18 M). Attached were two glass stoppers and a Dean-Stark apparatus surmounted by a Liebig condenser. 2-Fluoroaniline (6.69 mL, 10.0 g, 90 mmol) was added by syringe and MnO_2 (dried at 150 °C, 31.3 g) was added and the mixture heated to reflux. The Dean-Stark apparatus was filled with 1 mL H_2O and 9 mL toluene. After 5.5 h the mixture was filtered hot through Celite® and washed with hot toluene until no colour came through the filter. The filtrate was re-filtered to remove remaining MnO_2. The toluene was removed under reduced pressure and to the remaining residue was added EtOH. The suspension was heated to dissolve the material and left to crystallise. Provided dark-red crystals (3.67 g, 1 : 0.77 molar ratio by ^{1}H NMR, 56.3% by weight product, 21%, by-product was 2,2'-difluorohydrazobenzene). The mother liquor was concentrated but further attempts to form crystals resulted in a tarry material.

M.P. 99.6 – 101 °C (lit.),[163] 98.4 – 101.3 °C.

IR (cm[-1]): 2922, 1602, 1588.

^{1}H NMR (500 MHz, C*D*Cl$_3$) δ 7.80 (td, J = 7.8, 1.8 Hz, 1H), 7.48 (dddd, J = 8.3, 7.3, 5.0, 1.8 Hz, 1H), 7.30 – 7.26 (m, 1H), 7.23 (dddd, J = 8.0, 7.3, 1.3, 0.7 Hz, 1H).

^{13}C NMR (126 MHz, C*D*Cl$_3$) δ 160.5 (d, J = 258.5 Hz), 141.0, 133.2 (d, J = 8.5 Hz), 124.5, 118.1, 117.2 (d, J = 19.7 Hz).

^{19}F NMR (471 MHz, C*D*Cl$_3$) δ -124.18 (ddd, J = 10.6, 7.6, 5.0 Hz).

Prepared by adaptation of literature procedure, data in line with literature data.[141]

2,2'-difluorohydrazobenzene[164]
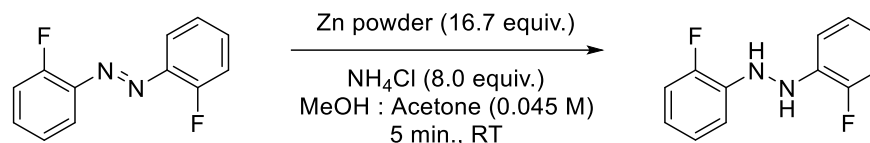
To a 1 L 3-necked RBF was added 2,2'-difluoroazobenzene (3.6 g, 16.5 mmol) and a 1:1 mixture of MeOH : Acetone (360 mL, 0.045 M). NH_4Cl (7.06 g, 8.0 equiv.) was added, and freshly activated Zn dust (18.01 g, 16.7 equiv.) added with rapid stirring. The red colour changed into yellow within 5 min. and TLC (10% EtOAc / Hex.) indicated complete reaction. The mixture was poured into ice water (300 mL) and stirred for 30 min. The suspension was then filtered and the solid washed with toluene (collected in a separate flask). The toluene filtrate was concentrated and EtOH added. The mixture was crystallised from EtOH to provide dark, transparent crystals of 2,2'-difluorohydrazobenzene (2.0 g, 0.16 : 1.0 molar ratio of azobenzene to hydrazobenzene *via* ^{1}H NMR, 86.3% by weight product, 47%). Taken through without further purification.

^{1}H NMR (500 MHz, CDCl$_3$) δ 7.08 – 6.96 (m, 6H), 6.82 – 6.74 (m, 2H), 5.86 (br s, 2H).

^{13}C NMR (126 MHz, CDCl$_3$) δ 150.8 (d, $J = 239.3$ Hz), 136.6 (d, $J = 10.1$ Hz), 124.9 (d, $J = 3.6$ Hz), 119.9 (d, $J = 7.1$ Hz), 115.1 (d, $J = 17.7$ Hz), 113.8 (d, $J = 2.7$ Hz).

^{19}F NMR (471 MHz, CDCl$_3$) δ -134.76 – -137.59 (m).

IR (cm^{-1}): 3361 (NH), 1617 (NH bend).

Prepared by adaptation of literature procedure, data in line with literature data.[142]

3,3'-difluorobenzidine[161]
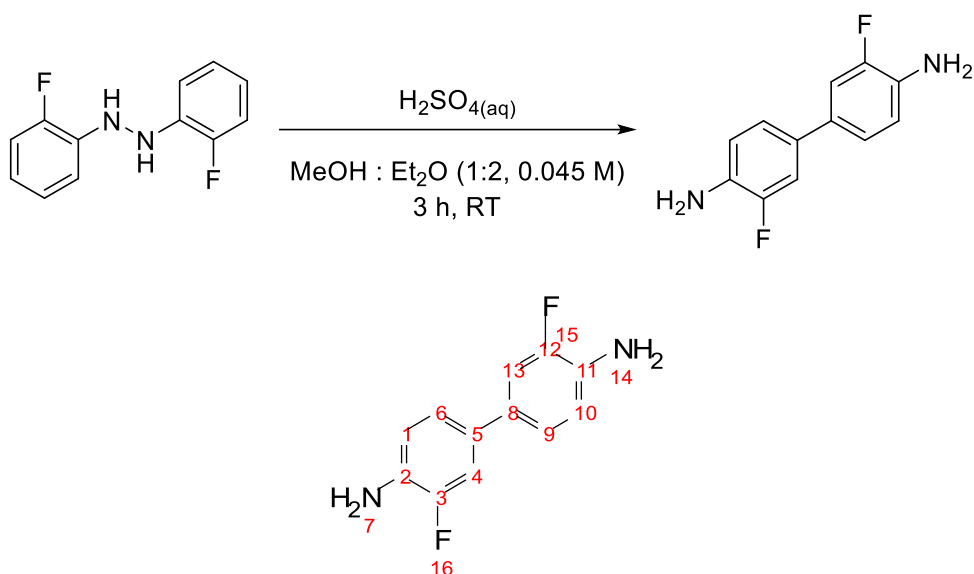
To a 500 mL RBF was added 2,2'-difluorohydrazobenzene (2.0 g, 9.09 mmol) and Et_2O (200 mL, 0.045 M). To this solution was added MeOH (100 mL) and slow addition of H_2SO_4 solution (3.45 M, 25 mL). White solid began to precipitate immediately upon addition of acid, and the starting material was monitored by TLC (10% EtOAc / Hexane). After 3 h the starting material had been consumed and the mixture filtered, washing with Et_2O. The white solid was dried under high vacuum to provide 2.86 g of sulfate salt. The material was free-based by adding the solid to a separatory funnel containing 1 M NaOH solution (100 mL) and extracting with chloroform (4x 50 mL) the organic layers were dried ($MgSO_4$) and concentrated to provide 3,3'-difluorobenzidine (1.26 g, 5.72 mmol, 63%) as a light-purple crystalline solid. Could be recrystallised from EtOH.

^{1}H NMR (400 MHz, CD_3CN) δ 7.20 (dd, J = 13.1, 2.1 Hz, 2H, **4**, **13**), 7.14 (ddd, J = 8.2, 2.1, 0.7 Hz, 2H, **6**, **9**), 6.82 (dd, J = 9.7, 8.2 Hz, 2H, **1**, **10**), 4.19 (br s, 4H, **7**, **14**).

^{13}C NMR (101 MHz, CD_3CN) δ 152.2 (d, J = 236.4 Hz, **3**, **12**), 135.0 (d, J = 13.1 Hz, **2**, **11**), 130.5, **5**, **8**, 122.6 (d, J = 2.9 Hz, **6**, **9**), 117.4 (d, J = 4.6 Hz, **1**, **10**), 113.1 (d, J = 19.5 Hz, **4**, **13**).

^{19}F NMR (376 MHz, CD_3CN) δ -136.97 (ddd, J = 13.1, 9.7, 0.7 Hz).

M.P. 119 °C (lit.),[165] 119.3 – 121.2 °C.

IR (cm^{-1}): 3420 (NH), 3301 (NH), 1639 (NH bend).

Prepared by adapting a literature procedure, data matches with literature values.[143]

Dimethyl 2,2'-({3,3'-difluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate
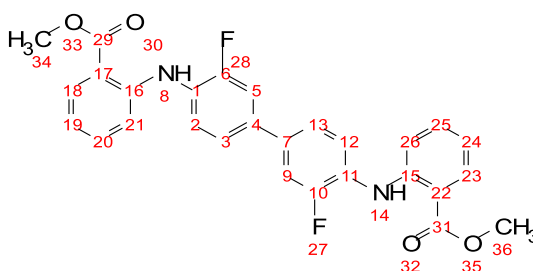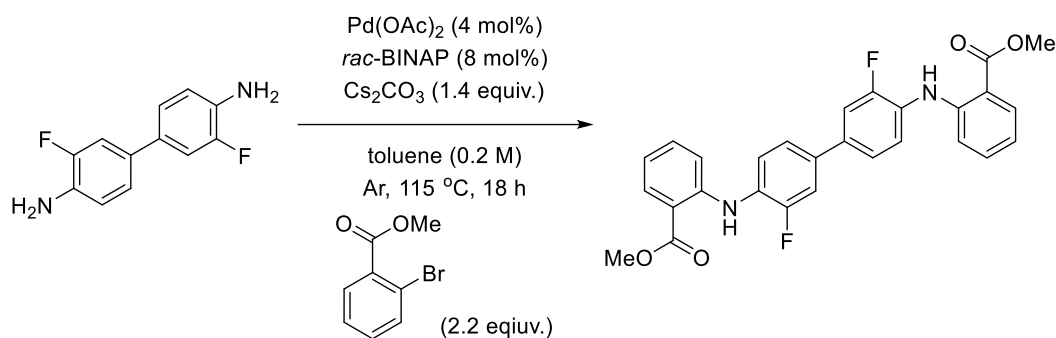
To an oven-dried 3-necked 100 mL RBF was added a stirrer bar and fitted with an oven-dried air condenser and two septa. To this was added under argon 3,3'-difluorobenzidine (1.0 g, 4.54 mmol) and methyl 2-bromobenzoate (1.27 mL, 2.0 equiv.). Degassed, dried toluene (22.7 mL, 0.2 M) was added *via* syringe and the mixture was further degassed for another 10 min. Pd(OAc)$_2$ (40.7 mg, 4 mol%), *rac*-BINAP (226 mg, 8 mol%) and Cs$_2$CO$_3$ (2.07 g, 1.4 equiv.) were then added under an argon flow. The mixture was stirred at 115 °C for 18 h under argon. The mixture was allowed to cool to RT and filtered through a pad of Celite®, washing with EtOAc. The solvent was removed, and the mixture purified by column chromatography, eluting with toluene to provide dimethyl 2,2'-({3,3'-difluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate as a light-yellow solid (725 mg, 1.51 mmol, 33%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 9.43 (br s, 2H, **8**, **14**), 7.94 (dd, *J* = 8.0, 1.7 Hz, 2H, **18**, **23**), 7.75 (dd, *J* = 13.2, 1.5 Hz, 2H, **5**, **9**), 7.62 – 7.54 (m, 4H, **2**, **3**, **12**, **13**), 7.49 (ddd, *J* = 8.7, 7.2, 1.7 Hz, 2H, **20**, **25**), 7.19 (dt, *J* = 8.5, 1.2 Hz, 2H, **21**, **26**), 6.90 (ddd, *J* = 8.1, 7.1, 1.1 Hz, 2H, **19**, **24**), 3.88 (s, 6H, **34**, **36**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 168.1, **29**, **31**, 155.0 (d, *J* = 243.4 Hz, **6**, **10**), 145.7, **17**, **22**, 134.7, **20**, **25**, 134.5 (dd, *J* = 7.6 2.2 Hz, **4**, **7**), 131.3, **18**, **23**, 127.6 (d, *J* = 11.7 Hz, **1**, **11**), 122.9, **3**, **13**, 122.6, **2, 12**, 118.5, **19**, **24**, 114.6, **21**, **26**, 113.9 (d, *J* = 20.9 Hz, **5**, **9**), 112.6, **15**, **16**, 52.2, **34**, **36**.

^{19}F NMR (376 MHz, DMSO-d_6) δ -125.51 – -125.63 (m).

M.P. 157.0 – 159.2 °C.

IR (cm^{-1}): 3254 (NH), 1688 (C=O).

MS ES+ m/z calcd for $C_{28}H_{22}F_2N_2O_4$ (M+H)$^+$: 489.1626, found: 489.1656.

Method adapted from literature.[146] May be recrystallised from MeCN/H_2O to provide light yellow crystals suitable for XRD, see appendix 1.
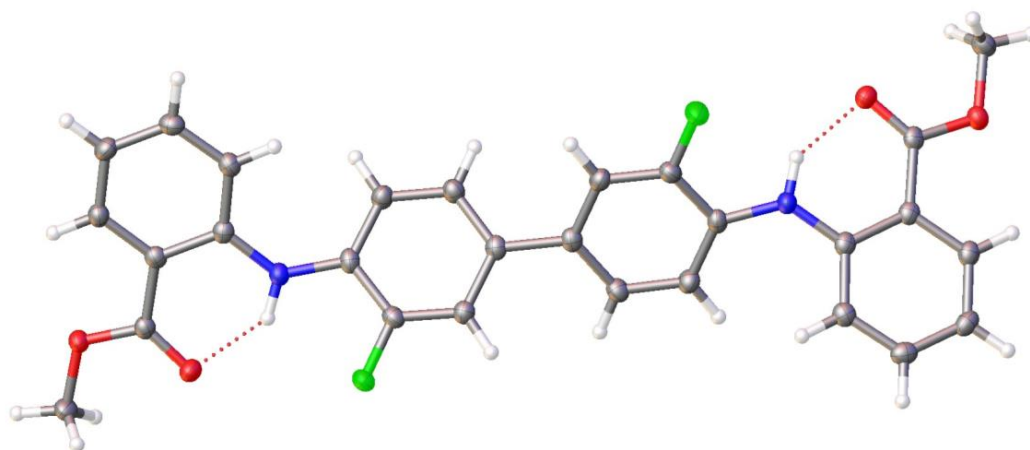
2,2'-({3,3'-difluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoic acid
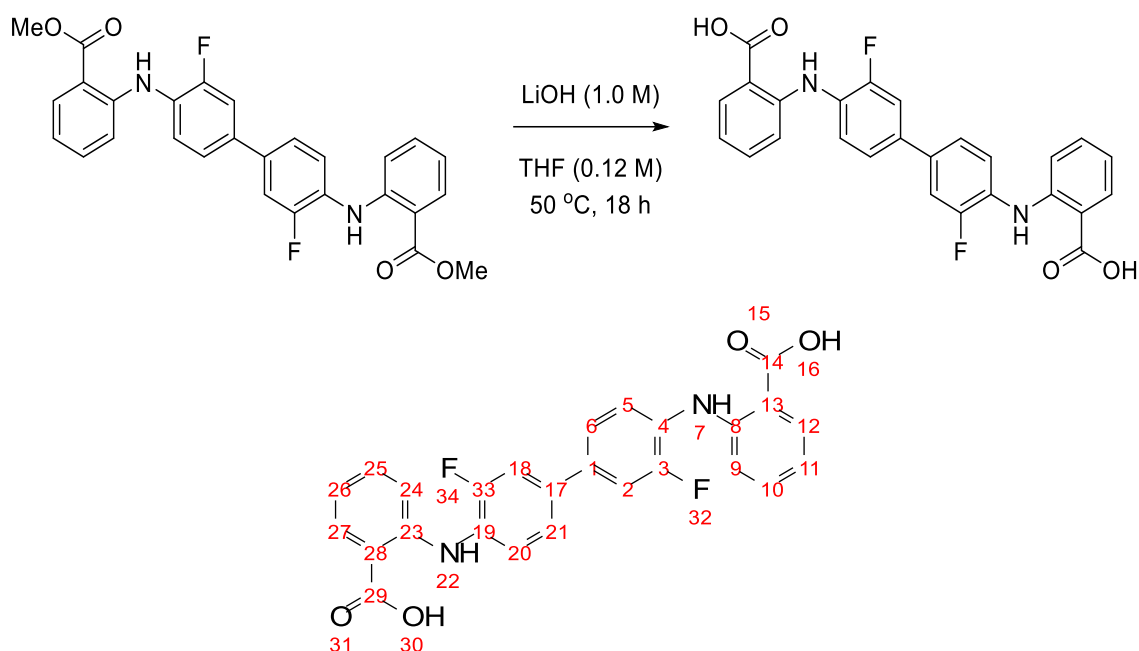
To a 10 mL RBF was added methyl dimethyl 2,2'-({3,3'-difluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate (300 mg, 0.614 mmol) and THF (5 mL, 0.12 M). To the stirred solution LiOH solution (1.0 M, 1.0 mL) was added and the mixture heated to reflux over 18 h. The mixture was cooled to RT, transferred to a separatory funnel with the aid of water and acidified (1.0 M, HCl$_{aq}$). The suspension was then extracted with EtOAc (3x 20 mL), the organic layers were collected, dried (MgSO$_4$) and solvent removed. Provided 2,2'-({3,3'-difluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoic acid as a yellow solid (228 mg, 0.49 mmol, 80%).

^{1}H NMR (500 MHz, DMSO-d_6) δ 13.25 (br s, 2H, **16**, **30**), 9.78 (br s, 2H, **7**, **22**), 7.95 (dd, J = 8.0, 1.7 Hz, 1H, **12**, **27**), 7.79 – 7.70 (m, 2H, **2**, **18**), 7.65 – 7.54 (m, 4H, **5**, **6**, **20**, **21**), 7.47 (ddd, J = 8.6, 7.1, 1.7 Hz, 2H, **10**, **25**), 7.20 (dd, J = 8.5, 1.2 Hz, 2H, **9**, **24**), 6.87 (ddd, J = 8.0, 7.1, 1.1 Hz, 2H, **11**, **26**).

^{13}C NMR (126 MHz, DMSO-d_6) δ 167.0, **14**, **29**, 154.8 (d, J = 242.9 Hz, **3**, **33**), 145.9, **8**, **23**, 134.3, **10**, **25**, 134.2, **4**, **19**, 131.8, **12**, **27**, 127.8 (d, J = 11.6 Hz, **1**, **17**), 122.5 (d, J = 17.9 Hz, **5**, **6**, **20**, **21**), 118.3, **11**, **26**, 114.2, **9**, **24**, 113.8 (d, J = 20.8 Hz, **2**, **18**), 113.4, **13**, **28**.

^{19}F NMR (471 MHz, DMSO-d_6) δ -125.82 – -125.99 (m, **32**, **34**).

M.P. 292.5 °C (dec.),

IR (cm^{-1}): 3258 (OH), 1689 (C=O).

MS ES+ m/z calcd for C$_{26}$H$_{18}$F$_2$N$_2$O$_4$ (M+H)$^+$: 461.1313, found: 461.1320.

119

5-bromoindole-3-carboxaldehyde[147]
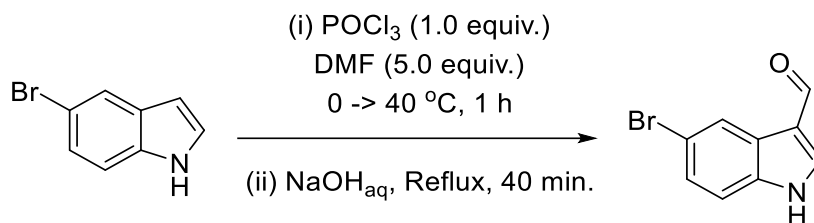
To a 50 mL RBF was added 5-bromoindole (1.005 g, 5.10 mmol), DMF (1.86 mL, 25.50 mmol, 5.0 equiv.) and cooled to 0 °C. To the cooled, stirred solution was added dropwise POCl$_3$ (0.47 mL, 1.0 equiv.). After addition the mixture was heated to 40 °C for 1 h. KOH (20 mL, 1.0 M) solution was added, and the mixture heated to reflux for 1.2 h. After allowing to cool the reaction mixture was transferred to a sep. funnel and washed with EtOAc (3x 20 mL). The organic layers were collected and washed with brine (20 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* The residue was purified by column chromatography eluting with (5:5 EtOAc : Hex.) to provide 5-bromoindole-3-carboxyaldehyde as an off-white solid (930 mg, 4.15 mmol, 81%).

^{1}H NMR (500 MHz, DMSO-d_6) *δ* 12.30 (br s, 1H), 9.92 (s, 1H), 8.34 (s, 1H), 8.21 (d, *J* = 2.1 Hz, 1H), 7.49 (d, *J* = 8.6 Hz, 1H), 7.39 (dd, *J* = 8.6, 2.1 Hz, 1H).

^{13}C NMR (126 MHz, DMSO-d_6) *δ* 185.1, 139.3, 136.0, 126.0, 122.91, 122.88, 117.4, 114.8, 114.6.

IR (cm^{-1}): 3183 (NH), 1634 (C=O).

Followed literature procedure, data is in line with literature values.

5-bromoindole-3-carbinol[127]
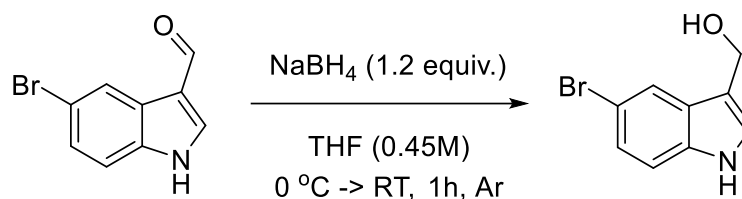
To an oven-dry 8 mL vial was added 5-bromoindole-3-carboxyaldehyde (0.4 g, 1.78 mmol) and anhydrous THF (4.0 mL, 0.45 M) under argon, the solution was cooled to 0 °C and NaBH$_4$ (80.8 mg, 1.2 equiv.) was added in one batch and the mixture allowed to warm to RT and stirred for 2 h. The mixture was quenched with water (3 mL) and transferred to a sep. funnel, where more water (7 mL) was added. The mixture was extracted with EtOAc (2x 10 mL), the organic layers were collected and washed with brine (10 mL), dried (MgSO$_4$) and solvent removed to provide an off-white solid of 5-bromoindole-3-carbinol (334 mg, 1.47 mmol, 84%).

^{1}H NMR (400 MHz, CD_3CN) δ 9.29 (br s, 1H), 7.81 (d, $J = 1.9$ Hz, 1H), 7.35 (dd, $J = 8.7$, 0.6 Hz, 1H), 7.28 – 7.21 (m, 2H), 4.70 (dd, $J = 5.5$, 0.7 Hz, 2H), 2.84 (br td, $J = 5.6$, 1.4 Hz, 1H).

^{13}C NMR (101 MHz, CD_3CN) δ 135.9, 129.2, 125.4, 124.7, 122.0, 116.6, 113.8, 112.4, 56.3.

IR (cm^{-1}): 3519 (OH), 3338 (NH).

Followed literature procedure, data is in line with literature values.

5-iodoindole-3-carboxyaldehyde[166]

(i) POCl$_3$ (1.0 equiv.)
DMF (5.0 equiv.)
0 -> 40 °C, 1 h

(ii) NaOH$_{aq}$, Reflux, 40 min.
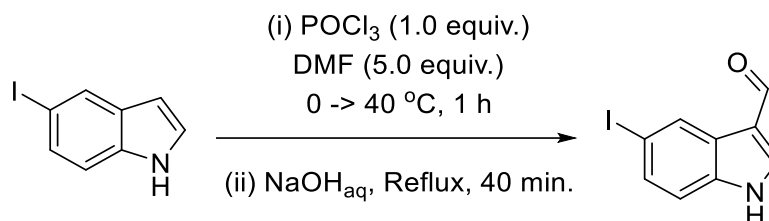
To a 50 mL RBF was added 5-iodoindole (1.0 g, 4.11 mmol) and DMF (1.6 mL, 5.0 equiv.). POCl$_3$ (0.385 mL) was added dropwise to the stirred solution at 0 °C. The mixture was heated to 40 °C for 1 h. Ice was added, and KOH$_{aq}$ (2.0 M, 20 mL) was added, and the mixture heated to reflux for 40 min. The mixture was cooled to RT and extracted with EtOAc (3x 20 mL). The organic layers were collected and dried (MgSO$_4$), solvent removed and the residue columned (eluting with 40% -> 60% EtOAc / Hexane) to provide 5-iodoindole-3-carboxyaldehyde as an orange solid (917 mg, 3.38 mmol, 82%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 12.27 (br s, 1H), 9.91 (s, 1H), 8.43 (d, J = 1.8 Hz, 1H), 8.29 (s, 1H), 7.54 (dd, J = 8.5, 1.8 Hz, 1H), 7.37 (d, J = 8.5 Hz, 1H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 185.1, 138.9, 136.2, 131.5, 129.1, 126.6, 117.1, 114.9, 86.6.

M.P. 228.4 – 232.1 °C.

IR (cm^{-1}): 3106 (NH), 2891 (CH aldehyde), 2810 (CH aldehyde), 1614 (C=O).

Literature procedure followed, data in line with literature values.[147]

5-iodoindole-3-carbinol
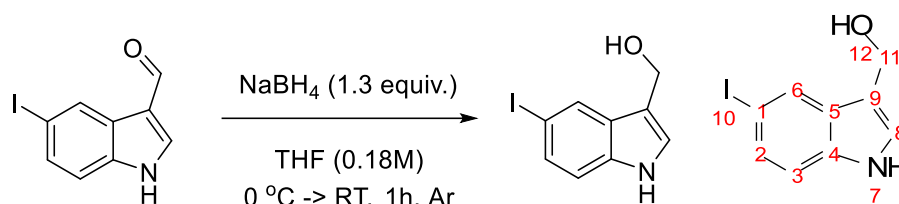
To a 3-necked 50mL RBF was added under argon 5-iodoindole-3-carboxyaldehyde (500 mg, 1.84 mmol) and anhydrous THF (10 mL, 0.18 M). NaBH$_4$ (90 mg, 1.3 equiv.) was added batchwise under argon flow to the stirred solution at 0 $^{\circ}$C, and afterwards the reaction allowed to warm to RT. After 1 h the reaction was considered complete by TLC (7:3 EtOAc : Hex., both aldehyde and alcohol have the same R$_f$ values in this system but heating the plate prior to running the plate degrades the product and not the starting material). H$_2$O (10 mL) was added slowly, and the mixture extracted with EtOAc (3x 30 mL), the organic layers were washed with brine (10 mL) and solvent removed under reduced pressure below 30 $^{\circ}$C. The solid residue was washed with hexane and dried under high vacuum to provide 5-iodoindole-3-carbinol as a white solid (361 mg, 1.32 mmol, 72%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 11.05 (br s, 1H, **7**), 7.95 (d, $J = 1.6$ Hz, 1H, **6**), 7.32 (dd, $J = 8.4, 1.6$ Hz, 1H, **2**), 7.24 (d, $J = 2.4$ Hz, 1H, **8**), 7.21 (d, $J = 8.4$ Hz, 1H, **3**), 4.80 (br t, $J = 5.4$ Hz, 1H, **12**), 4.59 (d, $J = 5.4$ Hz, 2H, **11**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 135.5, **4**, 129.4, **5**, 128.8, **2**, 127.5, **6**, 124.4, **8**, 115.5, **9**, 113.9, **3**, 82.2, **1**, 55.2, **11**.

IR (cm^{-1}): 3523 (OH), 3311 (NH).

M.P. 99.0 - 100.1 $^{\circ}$C (dec.)

Adapted from literature procedure.[127]

4 Efforts Towards the Binding Optimisation of NSC-217913: Analogue Synthesis

4.1 Introduction

NSC-217913 (**4.1**) was taken forward for analogue library synthesis firstly because of the apparent ease of resynthesis for co-crystal studies, and secondly for the perceived ease of library synthesis and optimisation (Figure 4-1). Additionally, the molecule seemed to have favourable properties as predicted by Flare. The predicted ClogP was 2.4 and its molecular weight = 307.2 gmol^{-1}, both suggesting a favourable hit as derivatisation and growth would not result in too large or lipophilic derivatives.
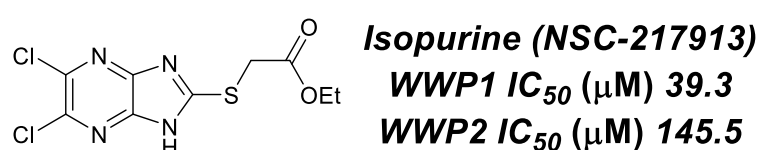
Figure 4-1: NSC-217913 hit structure **4.1** and associated IC$_{50}$ data.

Compound **4.1** has only been mentioned occasionally within the literature, with its synthesis described in 1981.[167] Briefly, starting from 2,3-diamino-5,6-dichloropyrazine **4.2**, cyclourea formation by trans-amidation with urea (**4.3**) occurred in 38% yield, forming **4.4** (Scheme 4-1). Following this, thionation with pentaphosphorus decasulfide provided the cyclothiourea **4.5**. To finish the synthesis, S-alkylation was performed in 39% yield with ethyl chloroacetate **4.6** to arrive at NSC-217913 **4.1** in an overall yield of 6%.
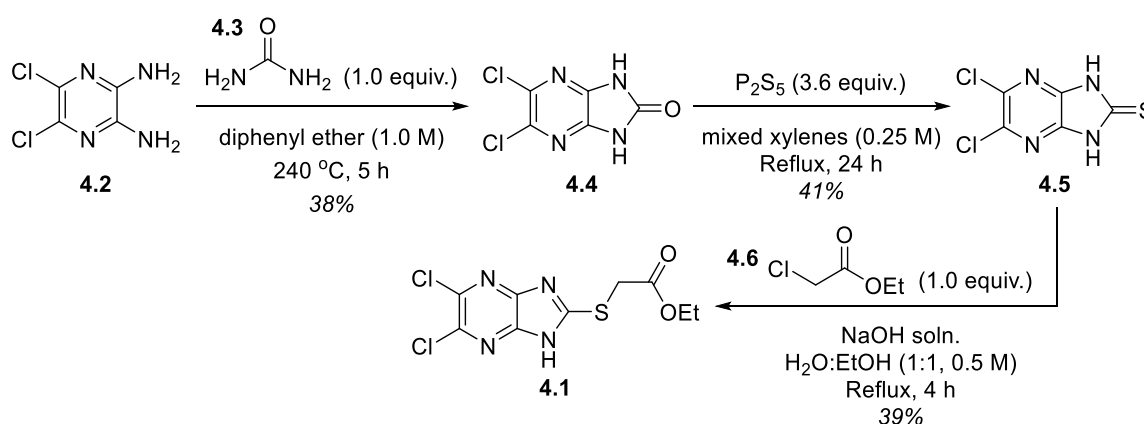
Scheme 4-1: Literature synthesis of NSC-217913 **4.1**.

Compound **4.1** was also mentioned in a paper detailing pyruvate kinase inhibitors of *Cryptosporidium parvum*, although not as an active compound.[168] Through a multifaceted strategy against *Schistosoma mansoni* lysine deacetylase 8 enzyme (SmKDAC8), a target for the prevention of bilharzia, **4.1** was identified from NCI diversity set V. Having been

identified from virtual screening as an interesting compound, it was only able to inhibit some hKDAC isoenzymes and then only weakly and was not pursued further.[169]

Additionally, because **4.1** has only minimal research surrounding it, there was an opportunity to explore the production of novel compounds with potentially interesting activities based upon its structure. The imidazo[4,5-*b*]pyrazine heterocyclic core has received comparably little interest as a heterocyclic core within medicinal chemistry research (several papers do describe its uses) than its related isomeric neighbour purine (imidazo[4,5-*d*]pyrimidine), and deaza-isomer imidazo[4,5-*b*]pyridine and further related imidazo[4,5-*c*]pyridine core (Figure 4-2), which have received extensive interest.[170–175]

a)

Imidazo[4,5-*c*]pyridine   Imidazo[4,5-*b*]pyridine   Imidazo[4,5-*b*]pyrazine   Imidazo[4,5-*d*]pyrimidine
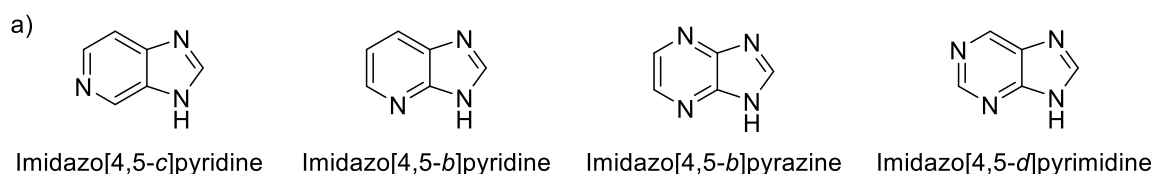
Figure 4-2: a) Fused 6,5 heterocyclic cores structurally related to imidazo[4,5-b]pyrazine.

4.2 <u>Aims and Objectives</u>

The general objective of this body of work was to produce a compound based upon NSC-217913 which is an inhibitor of WWP2 with improved activity over NSC-217913 itself. The two objectives of this chapter therefore were to first resynthesise NSC-217913 and secondly produce derivatives to help understand what essential interactions are present, and if any putative additional interactions are available to provide better binding affinity. This addresses the bulk of the work in this project, within the aim of producing an analogue with improved inhibitory activity against WWP2.

The second objective was split into three smaller objectives and were achieved by producing three subset analogue libraries, the first of which was to produce variations of the ethyl thioacetate moiety to assess the limitations of depth and functionality of the side chain in the putative binding domain. Within this objective the synthesis of **4.1** was also achieved. Figure 4-3 shows void space is clearly visible from the RHS of the binding site, this model was taken from the molecular docking results of NSC-217913 as discussed in section 3.5.5.1. Based on the predicted orientation of **4.1** within the proposed binding site, there was potential for side-chain growth into the void space and derivatives which look at exploring additional interactions here were targeted. Compounds with reduced functionality were also targeted, aimed at investigating what essential interactions the ethyl thioacetate side chain is making for inhibition.
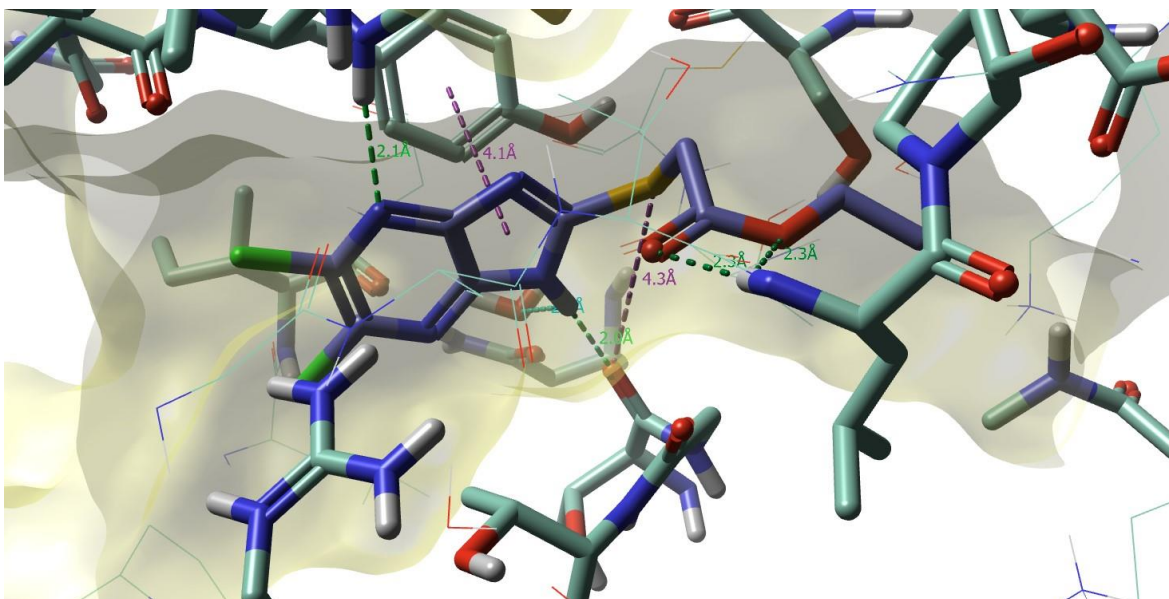
*Figure 4-3: Favoured molecular docking ligand pose of NSC-217913 **4.1** in the proposed binding site of WWP2.*

The second subset of analogues varied the heterocycle, which allowed an understanding of the position and electronics of the nitrogen atoms in the 5,6-fused system, any of which may have produced more favourable π-π and hydrogen bonding interactions than those described in Figure 4-3 above. This objective comprised of two parts: the first of which kept the imidazo-portion of the heterocycle constant but varied the 6-membered ring. This allowed an efficient synthetic plan to be executed which aimed to, at least partially, circumvent the tedious and time-consuming nature of heterocycle variation by telescoping routes from similar heterocycle starting materials (Figure 4-4).

Diamino-heterocycle functionalisation   Commercial Library Starting materials
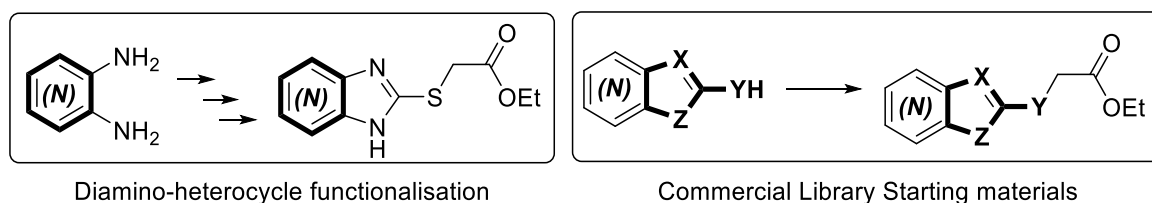
Figure 4-4: Sub-divisions of heterocyclic variation project.

The second subpart was to purchase commercial materials which expedited the production of different heterocycles. Fused pyrazine derivatives each require different starting materials which results in an inefficient synthetic plan, so purchasing commercially available heterocycles with several constant 6-membered azacycles but with variation on the 5-membered moiety was more efficient. Further variations away from the 5,6-fused ring system were also sought. These heterocycles should be amenable to functionalisation with an ethyl thioacetate moiety (or other, more active appendage if identified) for comparison.

126

Lastly, the third subset was based on finding new interactions further away from the interactions as predicted on Flare®. A series of compounds with functionality emanating away from the heterocycle towards the outer region of the putative binding domain were planned to be synthesised (Figure 4-5). This would be achieved by cross-coupling or S$_N$Ar reactions of the dichlorinated imidazo[4,5-b]pyrazine core with several heterocycles, again with the ethyl thioacetate chain still present if this portion has not been improved upon in previous library syntheses.
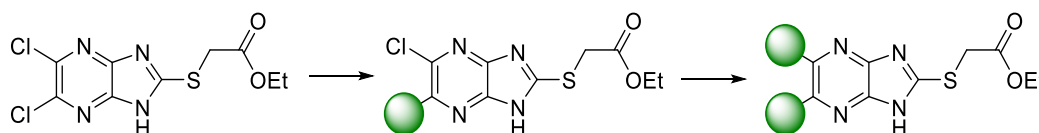
Figure 4-5: Functionalisation of the chlorine positions on the 5,6-dichloroimidazo[4,5-b]pyrazine core.

4.3 Results and Discussion

4.3.1 Synthesis II: NSC-217913 Production and Side-chain Variation

Following the aims and objectives set out in section 4.2, the first objective (variations of the ethyl thioacetate moiety) was targeted. Literature pointed to a general synthetic precursor **4.5** which was used as a common advanced intermediate for the synthesis of ethyl thioacetate variations library. This work therefore shows the research towards an improved synthesis of **4.5** as a common intermediate over historical routes, with the synthesis of NSC-217913 detailed as part of the library of compounds produced afterwards.

4.3.1.1 Synthetic Work Towards 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)thione

The single literature procedure found to synthesise 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)thione (**4.5**) involves the thionation from 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)one (**4.4**), which in turn is produced from 2,3-diamino-5,6-dichloropyrazine **4.2** *via* trans-amidation with urea.[167] However, the routes used to obtain **4.2** include firstly accessing tetrachloropyrazine (**4.7**) and then double amination at high temperature for a prolonged period. The production of **4.7** has previously been achieved either by vapour-phase chlorination of 2-methylpyrazine in carbon tetrachloride at 545 °C for 16 seconds, or treatment of a number of pyrazine or piperazine precursors (chloropyrazine, 2,5-diketopiperazine, pyrazine-2,3-dicarboxylic acid, 2-hydroxypyrazine) with phosphorus

127

pentachloride in an autoclave at 310 - 330 °C.[176,177] The formation of tetrachloropyrazine is followed by amination in a sealed vessel to obtain 2,3-diamino-5,6-dichloropyrazine **4.2** by sequential S_NAr using ammonium hydroxide.[178] Because of the safety issues and lack of an autoclave, an alternative route to access **4.2** was sought.

A relatively safer chlorination at lower temperatures used 2,5-diketopiperazine **4.8**, which was produced by heating glycine **4.9** in ethylene glycol for 18 h (Scheme 4-2).[179] From adaption of a literature procedure, **4.8** was subjected to heating at 90 °C with PCl_5 in $POCl_3$ for 18 h.[180] Distillation of the phosphorus reagents and subjecting the crude mixture to chromatography provided **4.7** in 32% yield. After this, **4.7** was subjected to amination conditions over two days to provide **4.2** in 31% yield.
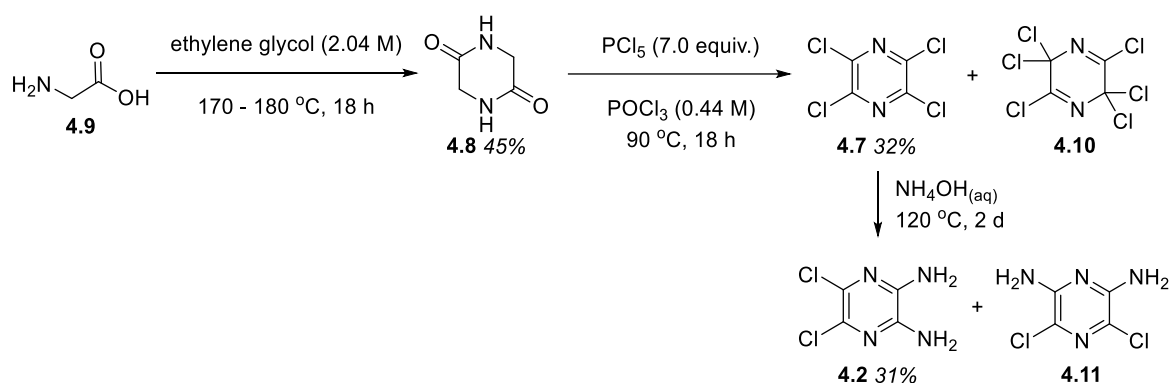
*Scheme 4-2: Synthetic route towards 2,3-diamino-5,6-dichloropyrazine **4.2**.*

Unfortunately, a hexa-chloropyrazine by-product **4.10**, most-likely produced from dichlorination of the carbonyl moiety by PCl_5, reduces the yield and complicates purification.[181] Amination of tetrachloropyrazine was also complicated by the formation of the 2,6-diamino-3,5-dichloropyrazine **4.11** regioisomer, which forms in approximately equal amounts and the formation of which cannot be supressed. This regioisomer formation can be explained by the electronics surrounding intermediate 2-amino-3,5,6-trichloropyrazine **4.12** towards S_NAr with ammonium hydroxide (Scheme 4-3).
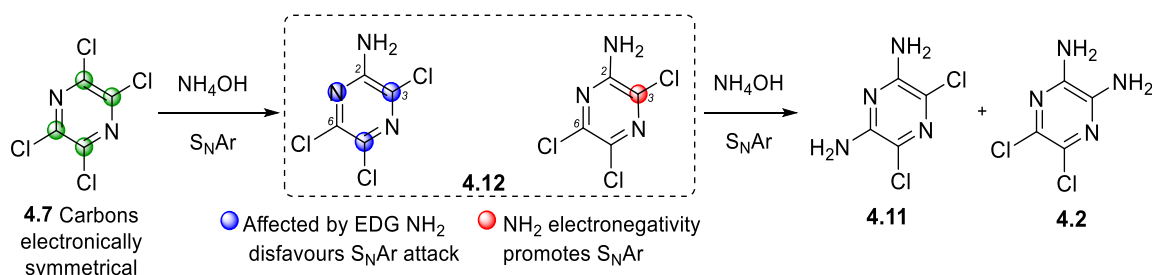
Scheme 4-3: Rationalisation for the two diamino-dichloropyrazine regioisomers are observed.

After initial S_NAr to form the mono-aminated intermediate **4.12** from **4.7**, the electronics of the amine group favours both the 6-position and the 3-positions towards secondary S_NAr

attack (Scheme 4-3). The 3-position is affected by the electron donation of the amine, but because of its position being *ortho* to the amine, it is relatively electron-deficient as the nitrogen atoms electronegativity increases the reactivity of the 3-position towards S_NAr.

The overall yield to produce **4.2** as starting material is 4.5%. In the two reactions leading up to its production, by-products **4.10** and **4.11** are produced which cannot be suppressed under the conditions utilised. All three steps are low yielding and require high temperatures, long reaction times. It was considered generally this route was not amenable to larger-scale production of **4.2** for library synthesis and so was abandoned in favour of a more selective and higher yielding route.

More success was had when starting from 2-aminopyrazine **4.13**, which after dichlorination with *N*-chlorosuccinimide (NCS) for 2 hours in THF produced 2-amino-3,5-dichloropyrazine **4.14** in 97% yield (Scheme 4-4) on multigram scale.[182] Compound **4.14** was subjected to similar amination conditions as above, and this allowed the regioselective introduction of an amine group at the 3-position in 64% yield as the sole product (**4.15**), as the 5-position is deactivated towards S_NAr.[183] The introduction of a chlorine atom at the 6-position would complete the synthesis of **4.2** *via* this route.
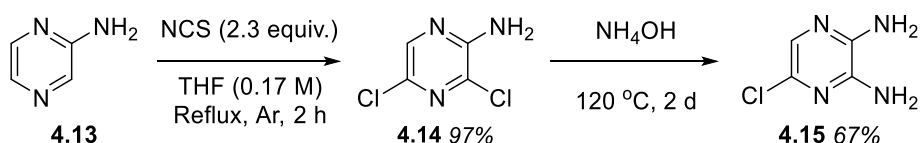
Scheme 4-4: synthetic route towards **4.15**.

Several attempts to accomplish this using NCS failed to provide acceptable results, with significant degradation occurring in nearly all test reactions in a variety of conditions (Table 4-1). The product was only isolated when TLC confirmed product was present, as it would be unproductive to workup and purify if the TLC was inconclusive or showed practically no product present. Initially, leaving NCS at 1.0 equiv. overnight in THF at RT lead to significant degradation of material, with no yield isolated (Entry 1). Decreasing the molarity to 0.075 M from 0.15 M of THF did not improve yields (data not shown). The addition of acid and heating for 30 min. allowed the isolation of **4.2** in 12% yield. Improvement to this was found when heating at reflux with 1.3 equiv. of NCS (entry 3), although further improvements were not found using NCS.

4.15   4.2   4.16 CBO
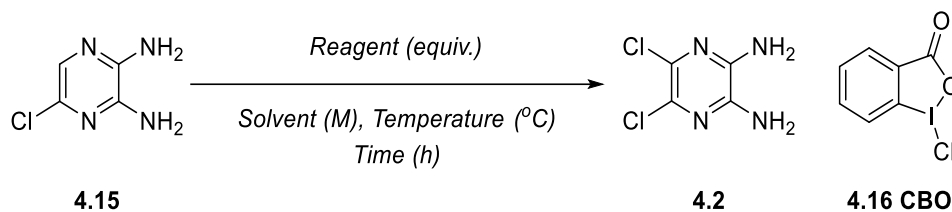
Entry	Reagent(s) (eq.)	Solvent (M)	Temp. (°C)	Time*	Yield (isolated)
1	NCS (1.0)	THF (0.15)	RT	18 h	N.D.
2	NCS (1.0) + HCl (2 M, 1 drop)	THF (0.075)	40	30 min.	12%
3	NCS (1.3) + HCl (2 M, 5 drops)	THF (0.15)	Reflux	30 min.	29%
4	CBO (1.0)	DMF (0.15)	RT	5 min.	38%
5	CBO (1.2)	DMF (0.15)	-15	N/A	N.D.
6	CBO (1.2)	DMF (0.15)	-41	N/A	N.D.
7	CBO (1.2)	DMF (0.075)	RT	2.5 min.	63%

*Table 4-1: Representative conditions tested for chlorination of 2-amino-3,5-dichloropyrazine 4.15. *Monitored by TLC. All test reactions conducted on 0.345 mmol scale.*

The use of the hypervalent iodine reagent 1-chloro-1,2-benzoiodozol-3(1*H*)-one (**4.16**, CBO) as an electrophilic chlorination regent of amine-containing nitrogen heterocycles later piqued our interest as a simple to prepare, bench-stable reagent (Scheme 4-5), shown to chlorinate 2-aminopyrazine **4.13** in high yields.[184] Compound **4.16** was efficiently prepared by reaction of 2-iodobenzoic acid with sodium chlorite in acidic aqueous solvent.
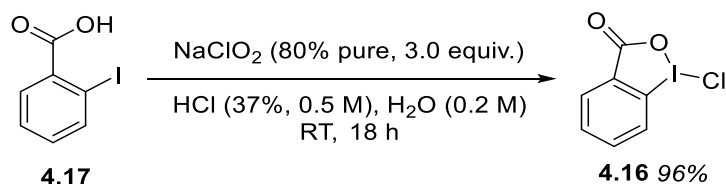
Scheme 4-5: Synthesis of CBO.

Initial attempts adapting the literature procedure (time reduced to 5 min.) produced a low but improved yield (38%, entry 4). A fast darkening of the reaction media of all tests attempted indicated a rapid reaction was occurring upon addition of the chlorination reagent. Attempts to reduce the temperature and therefore rate of reaction were

unsuccessful, as temperature control was difficult to regulate. Reducing the reaction time to 2.5 minutes under more dilute conditions improved yields (Table 4-1, entry 7), with these reaction conditions allowing acceptable and reproducible yields (63%) on gram scale when the reaction time was reduced to 1.5 minutes (timed) at room temperature. Increasing to decagram (10 g) scale provided a diminished yield of 56%. No further optimisation of this reaction was performed in the interests of time and because all that was required was an acceptable yielding reaction for the production of **4.2** for library synthesis. It seems that the reaction is very time sensitive and if this reaction is investigated again the movement towards a higher yielding reaction that does not seemingly degrade product over time would be ideal. DMF is teratogenic and in larger scale reactions such quantities of DMF represents a significant risk. Additionally, other chlorination reagents should be experimented with as only NCS and CBO were used in these optimisation experiments. The overall yield from this route is 41% over three steps which represents a marked improvement over the previous route, providing only 4.5% overall yield.

Whilst the production of **4.2** was ongoing, ring forming and thionation tests with the synthetically less-expensive 2,3-diamino-5-chloropyrazine **4.15** were investigated. The route proposed by the literature procedure for NSC-217913 synthesis, *i.e.,* formation of **4.4** using urea at high temperatures (240 °C) was unappealing due to the reported low yields (38%).[167] A test reaction with carbonyldiimidazole (CDI) provided high (79%) yield and much shorter reaction times with straightforward workup and purification (Scheme 4-6). This could be somewhat reproduced on **4.2** but with increased reaction temperatures and equivalency required to reach a 67% yield of **4.4**. The longer reaction times and higher temperatures were required probably due to the less-activated nature of the tetrasubstituted pyrazine ring system.
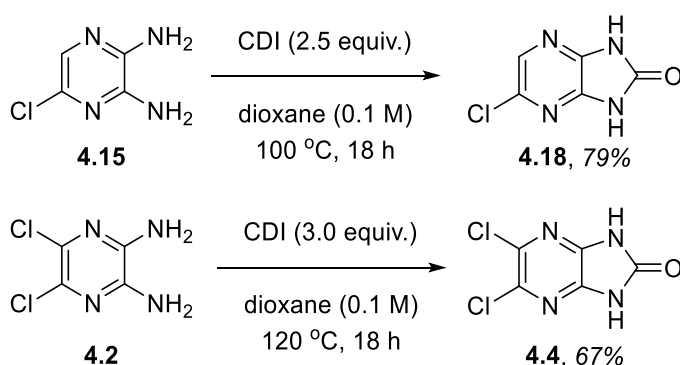
Scheme 4-6: production of **4.18** and **4.4**

Next, thionation conditions were surveyed to find appropriate conditions for the formation of imidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione variants **4.5** and **4.19**. Lawesson's reagent **4.20**

(Scheme 4-7) has long been the reagent of choice for such transformations, however the reactions are hampered by side-reactions and by-products, making purification difficult. Alternatively, the use of the precursor P_4S_{10} **4.21** utilised to synthesise Lawesson's reagent itself may also be used, as in the original report detailing NSC-217913's synthesis, although similar purification issues became apparent when using this reagent as well. A report detailing the use of a pyridine-phosphorus pentasulfide complex **4.22** ($Py_2P_2S_5$) for the clean transformation of amides and urea derivatives to thioamides and thioureas was interesting as the by-products were all reported to be salts and therefore purification is potentially much easier.[185] Synthesis is straightforward by mixing phosphorus pentasulfide (**4.21**) with dry pyridine at reflux and then filtering the crystals after cooling. The reactions performed were reported to require minimal purification and, apparently, had minimal handling issues with the usually malodorous properties of thiation reagents being completely avoided with this non-volatile salt, if sufficiently pure and stored correctly. Both Lawesson's reagent **4.20** and **4.22** were synthesised directly from P_4S_{10} by heating to reflux in anisole or pyridine respectively in good yields.[185,186]
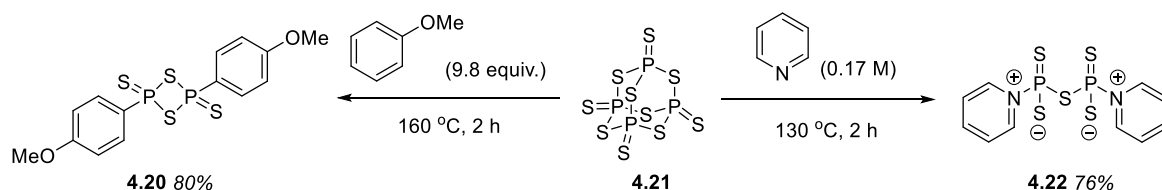
*Scheme 4-7: Tested thionation reagents **4.20 – 4.22** and their synthesis.*

Likewise with the previous ring closure to **4.4**, initial thionation experimentation was also carried out using the mono-chlorinated isomer **4.18**. The added benefit of using **4.18** is its single aromatic proton, which can be used in ^{1}H NMR analysis to monitor the reaction progression. The N*H* protons, which would also shift if a reaction had occurred, would not allow accurate integrations to be obtained. The aim of these exploratory experiments was identifying conditions with high conversion, ease of reaction setup and ease of purification due to the malodourous nature of these sulfur-containing materials (Table 4-2).
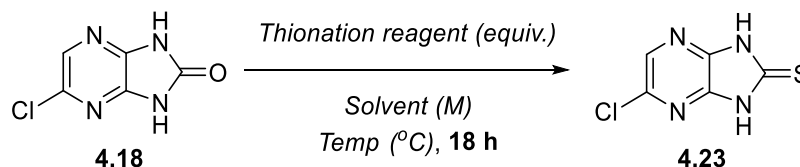
Entry	Reagent (equiv.)	Solvent (M)	Temp. (°C)	Conversion (%)*
1	**4.20** (1.2)	Dioxane (0.5)	100	100
2	**4.21** (3.6)	Pyridine (0.25)	115	88

3	**4.22** (0.3)	Pyridine (0.5)	115	0
4	**4.22** (1.5)	Pyridine (0.5)	130	84
5	**4.22** (2.0)	Pyridine (0.5)	130	100**

Table 4-2: Thionation conditions tested, reactions performed on a 0.346 mmol scale. *By ^{1}H NMR, **high-purity by ^{1}H NMR (DMSO-d$_6$) of the reaction mixture.

Conditions were found which utilise **4.20** in dioxane, and 100% conversion was obtained (entry 1), but because of the large amounts of impurities present in the crude reaction mixture that would have hampered isolation, different conditions were pursued.[187] Using the same thionation conditions as the literature procedure for the synthesis of NSC-217913 (and identical lengthy workup) did not allow for a significant quantity of product to be isolated (entry 2), even though the ^{1}H NMR spectrum after workup identified approximately 88% conversion. Using phosphorus pentasulfide-pyridine complex **4.22** in pyridine was initially unsuccessful (entry 3, Table 4-2) but this reaction setup was scrutinised more closely because of the very clean ^{1}H NMR spectrum of starting material obtained after workup (Figure 4-6). Compound **4.22** equivalency was started at 0.3, as **4.22** can theoretically react up to four times. Increasing $Py_2P_2S_5$ equivalency to 1.5 was much more successful, the product was only contaminated with a small amount of starting material and trace pyridine (which could be azeotropically removed with toluene) for an 84% conversion (Entry 4). 100% conversion was obtained when 2.0 equiv. was used under otherwise identical conditions for a practically pure by ^{1}H NMR crude material (Entry 5). These conditions allowed 5-chloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione **4.23** to be accessed in 85% isolated yield.
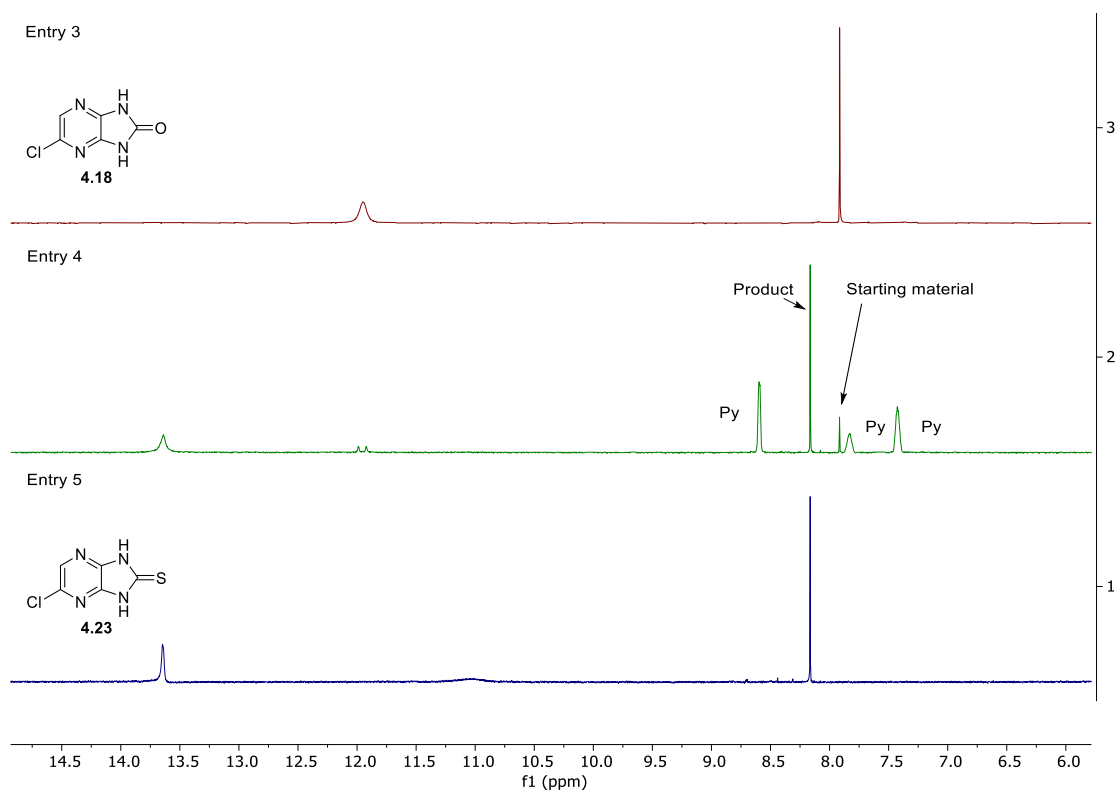
Figure 4-6: Crude ^1H NMRs (DMSO-d_6) showing reagent **4.22** thionation progress, Py denotes pyridine. Entries correspond to Table 4-2 entries.

Compound **4.4** was exposed to these conditions for a reduced 48% isolated yield of **4.5**. It is thought this result is because of the electronically more deficient system hampering the initial step of the thionation mechanism. Initially, the pyridine-phosphorus pentasulfide complex **4.22** may dissociate into the pyridinium-mono(dithiophosphine) ylide species ***int-4.24*** or the bis(dithiophosphine) ylide species ***int-4.24'*** (Scheme 4-8). The report that detailed the use of **4.22** in the first place does not comment about the proposed mechanism of thionation of this complex, although it is believed the mechanism to be practically identical to the mechanism of thionation using Lawesson's reagent.[188]
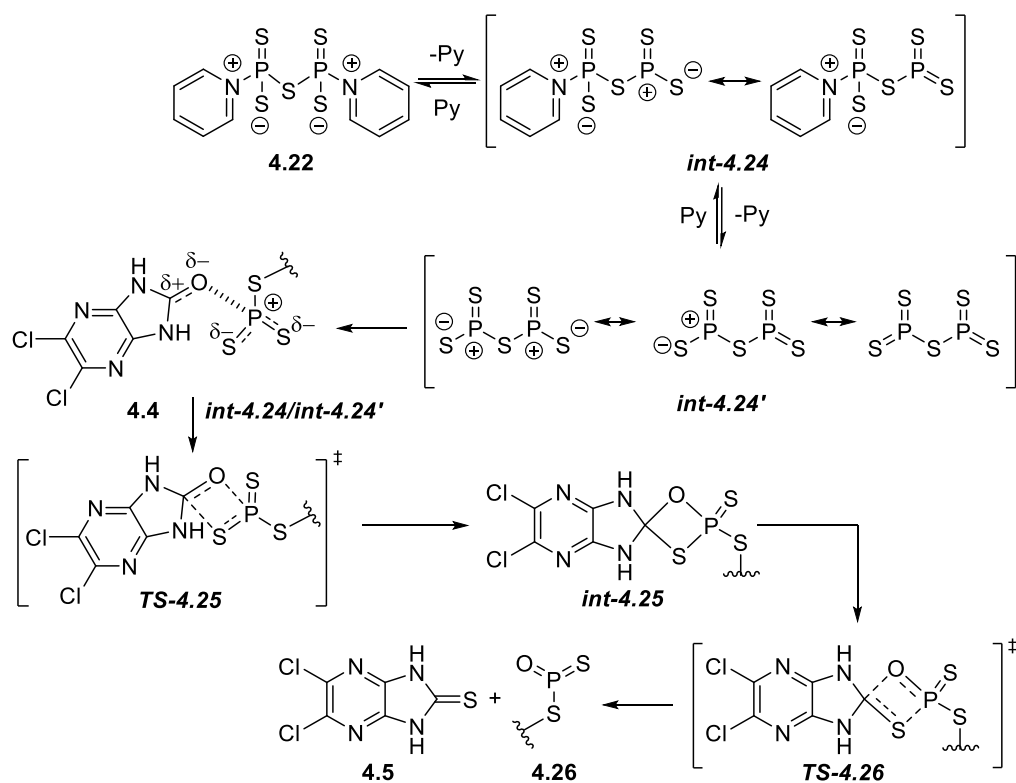
*Scheme 4-8: Proposed mechanism of thionation of **4.5**.*

According to a computational mechanistic study on the thionation of carbonyl compounds using Lawesson's reagent, an initial electrostatic encounter takes place between either mono- or bis-(dithiophosphine) yildes ***int-4.24*** or ***int-4.24'*** and carbonyl oxygen.[188] This interaction is only electrostatic and no covalent bond forms. A concerted cycloaddition through transition state ***TS-4.25*** of the urea moiety of **4.4** with either mono- or bis-(dithiophosphine) yildes ***int-4.24*** or ***int-4.24'*** leads to a thiaoxaphosphetane intermediate ***int-4.25***. which itself then undergoes a concerted cycloreversion to form **4.5** and an oxathiophosphine ylide **4.26**, through ***TS.4.26***. The oxathiophosphine ylide may also be a reactive species in the thionation and would operate in an identical way to the dithiophosphine yildes.

Attempts made to form **4.5** directly from **4.2** using carbon disulfide in the presence of sodium hydroxide or potassium hydroxide were unsuccessful (Table 4-3, entries 1 & 2) returning only starting material. It was thought that because of the low-boiling nature of carbon disulfide and the electron-withdrawn amine groups mean that higher temperatures and a large excess would be required for the reactions to proceed. Attempting this reaction with potassium ethyl xanthogenate too only returned starting material under the conditions shown (entry 3).[189] Thiocarbonyl diimidazole (TCDI) may also allow the direct formation of **4.5** from **4.2**, as the reaction with CDI had been successfully implemented in the same manner and is much safer than using CS_2. Initial test reactions proved

135

successful, although longer reaction times (36 h), increased stoichiometry of TCDI and the use of a higher-boiling solvent were required to drive the reaction forward and obtained only acceptable yields, with the highest stated in entry 4.
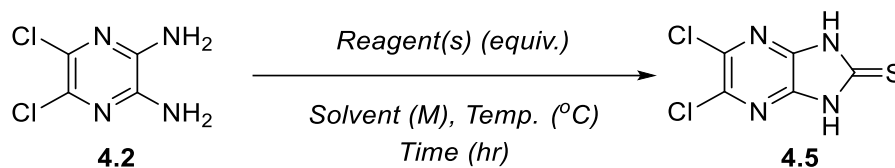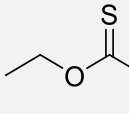
Entry	Reagent(s) (equiv.)	Solvent (M)	Temp. (°C)	Time (h)	Isolated yield
1	CS$_2$ (1.15), NaOH (1.0)	EtOH:H$_2$O (5:1, 0.17)	50, then 80	2, then 18	N.R.
2	CS$_2$ (1.15) KOH (1.0)	EtOH:H$_2$O (5:1, 0.17)	50, then 80	2, then 18	N.R.
3	[ethyl xanthate structure] SK (3.0), KOH (1.0)	Dioxane:H$_2$O (10:1, 0.17)	120	36	N.R.
4	TCDI (2.0)	Dioxane (0.2)	120	36	34%

*Table 4-3: Attempts towards formation of **4.5**. N.R. = no reaction.*

The low yield likely stems from low reactivity, probably due to the decreased electrophilicity of the carbonyl position of TCDI compared to CDI. The TCDI route was used in bulk runs to access NSC-217913 **4.1** and analogues over the thionation route as it was one step less and provided a more efficient route to access final products. However, the overall yields were comparable for either route, with **4.4** formation and thionation providing 32% yield over two 18 h reactions for total time of about 3 days for reaction and purification (Scheme 4-9). Using TCDI provided 34% yield for one 36 h reaction, with total time was about two days for reaction and purification. The overall synthesis of 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione **4.5** over the 4 steps described is 13%. A sufficient quantity of **4.5** was produced to allow work to begin on the RHS analogue library.
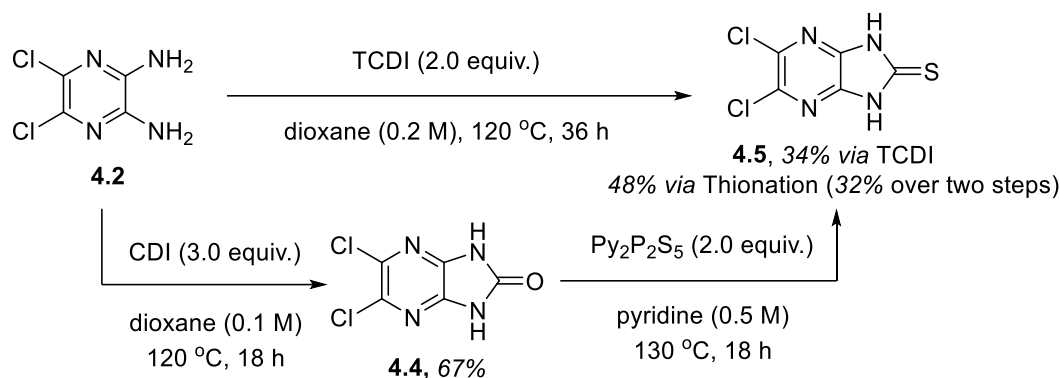
Scheme 4-9: routes to access **4.5**.

To conclude on the synthesis efforts towards **4.5**, a new synthesis of **4.2** has been developed. Initially utilising the inherent regioselectivity of 3,5-dichloro-2-aminopyrazine **4.15** towards amination, which overcomes the previous issues associated with regioselectivity with tetrachloropyrazine. 1-Chloro-1,2-benzoiodozol-3(1*H*)-one **4.16** was utilised as a key reagent for the installation of the final chlorine atom on the pyrazine ring where NCS failed. Finally, either direct thione ring closure using TCDI or cyclourea formation followed by thionation was performed to access **4.5**.

4.3.1.2 Synthesis of NSC-217913 and Side-chain Diversification from a Common Core

The choice of some variants for library synthesis was made on the basis of molecular docking of virtual compounds. These were derivatives with additional functionality on the ethyl acetate side (Section 4.2). This was performed in order to target compounds with a potential to improve the inhibitory activity against WWP2 and to identify limitations in the structures which could be accommodated within the binding site. Several compounds were also targeted with a reduced amount of functionality in order to investigate whether some functional groups (like the ester) were actually essential for binding at all (Figure 4-7). These were chosen without using molecular docking studies. It was hoped that one or more compounds from this series would provide improved activity, from which more thorough SAR may be developed.
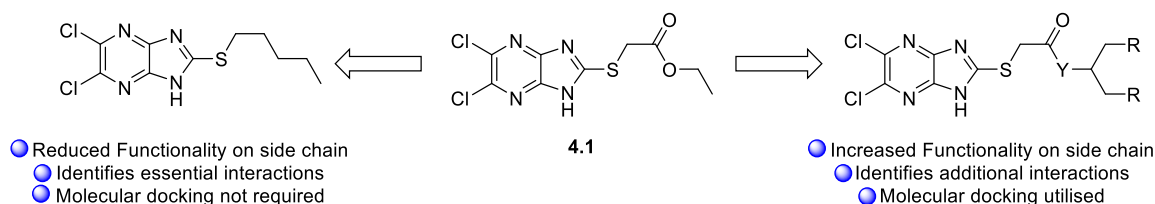
Figure 4-7: Plan for side-chain library synthesis.

As described in section 3.5.5.1, the virtual compounds investigated using the Flare®
software reproducing the ligand pose of NSC-217913 were chosen for synthesis. For
virtual analogues with larger side chains this did not have to be as similar (apart from the
heterocyclic interactions), as long as the interactions seemed reasonable. These
compounds would be synthesised as they could aid in understanding the limits of what
would reasonably fit into the binding site. Of the compounds investigated by Flare®
docking, the compounds presented below were chosen for synthesis (Figure 4-8) as they
provided favourable ligand poses as a result of the visual inspection. The target library,
apart from providing sensible ligand poses, represents a reasonable starting point for
structure-activity relationship (SAR) studies. If some analogues provided improved activity
a more rigorous SAR could be pursued surrounding that structure.
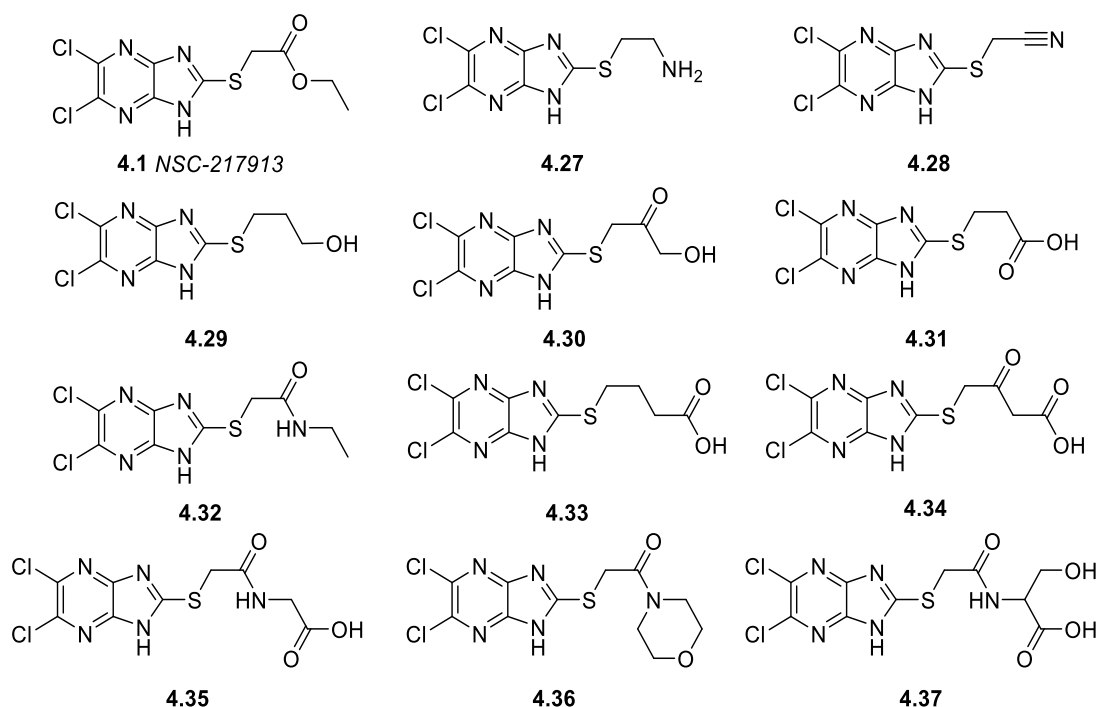
Figure 4-8: Target structures with different or increased side-chain functionality.

The compounds with a reduced amount of functionality are presented below (Figure 4-9).
These included ester hydrolysis and ketone derivatives of NSC-217913. Linear and cyclic
alkane side-chains were chosen to investigate if the presence of heteroatoms is required
for activity or not. The synthetic precursor of this series was also included to see if the
side chain was required at all. Lastly, removal of the sulfur in the thioether functional
group for a carbon atom would determine if the sulfur was important for activity.
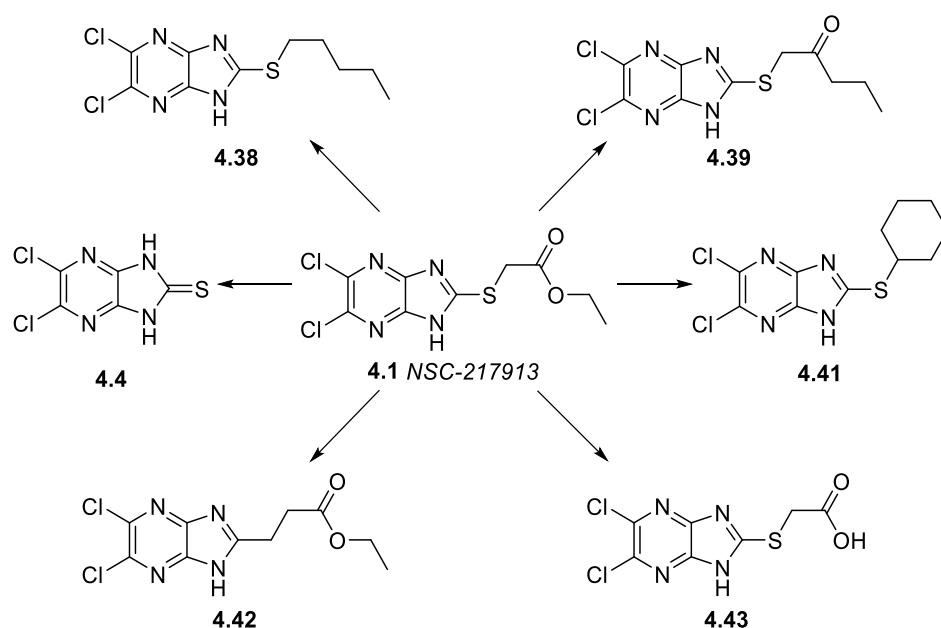
Figure 4-9: Targeted analogues with reduced functionality.

During computational work predicting additional positive interactions, an α-hydroxyketone analogue **4.30** presented a potentially beneficial binding pose. Fortunately, α-halo-α'-hydroxyacetones have been studied previously and was produced as an alkylation partner to **4.5**.[190] Initially, pyruvic acid **4.44** was chlorinated using sulfuryl chloride at RT for 2.5 days to provide **4.45** (Scheme 4-10).[191] The resulting crude material was reacted without further purification with trimethyl orthoformate in methanol with sulfuric acid at reflux for 2 days to produce **4.46** and non-chlorinated methyl 2,2-dimethoxypropionate **4.47** from the unreacted **4.44**. This mixture was purified by vacuum distillation to isolate the methyl 2,2-dimethoxy-3-chloroproionate **4.46** in 19% yield. Lastly, ester reduction to produce 2,2-dimethoxy-3-chloropropanol **4.48** was performed in 34%. The low overall yield of 6% is presumably because of aqueous solubility of the methyl ester **4.46** and alcohol **4.48**, as well as potential degradation during distillation. It was thought that this compound would be better suited to the base-promoted *S*-alkylation over the free deprotected acetone, because it was reported that these ketones readily form acetals which would be difficult to isolate from the reaction medium.
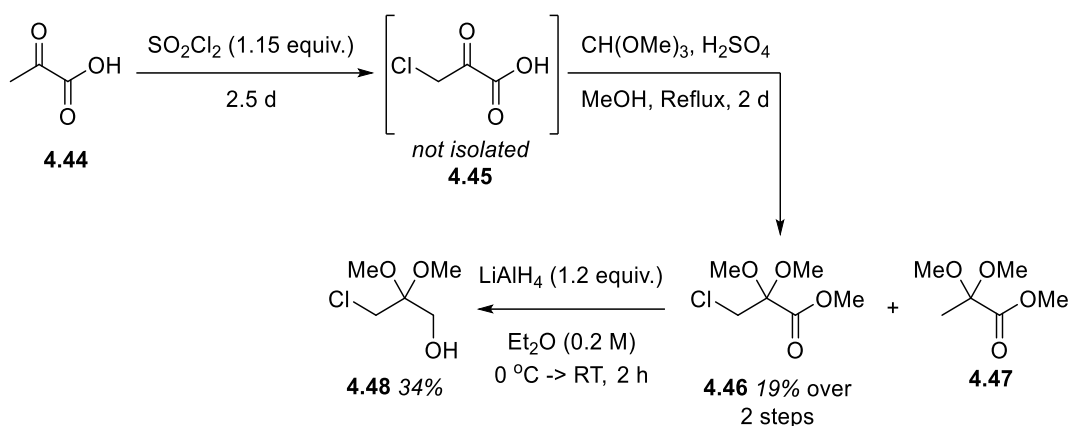
*Scheme 4-10: Synthesis of **4.48**.*

Amino acid variants were also tested computationally as it was initially thought that additional complementarity could be imparted using amino acid groups. The two best (glycine and serine) derivatives were targeted, with the (*S*)-serine precursor utilised as it was readily available. A morpholine derivative was also targeted (Scheme 4-11) as this provided a favourable binding pose. The protected amino acids and morpholine intermediates **4.49**, **4.50** and **4.51** were all reacted in a similar manner to provide a synthetic handle for *S*-alkylation. Each was reacted with bromoacetyl bromide **4.52** to give 41 - 52% yields of the α-bromoamides **4.53**, **4.54** and **4.55**. Finally, 1-bromopentan-2-one **4.56** alkylation partner as also synthesised, to provide a precursor to an analogue where the ester functionality was replaced for a ketone.
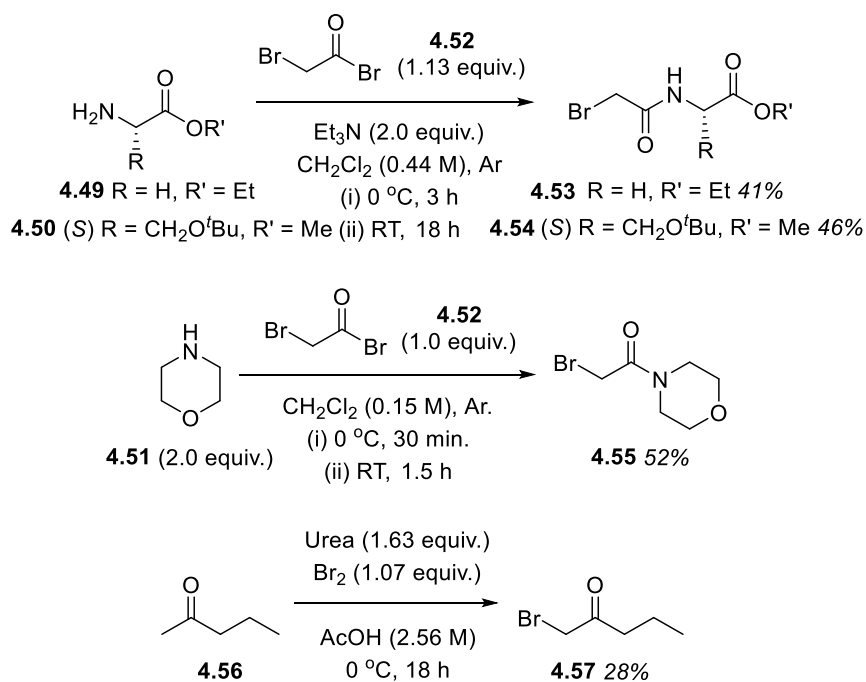
*Scheme 4-11: Synthesis of specific non-commercial alkylation partners **4.53 – 4.55, 4.57**.*

After these were synthesised, the *S*-alkylation with a series of alkyl halides and **4.5** provided final products **4.1**, **4.28**, **4.31**, **4.36**, **4.38**, **4.39**, **4.41**, and **4.58 – 4.61** (Scheme 4-12).[167] The conditions to access this library are general, and adapted from literature conditions.[189] Under basic conditions the reactions are selective for the sulfur, and it has been shown that under neutral conditions and at higher temperatures unselective *N*-alkylations (on both pyrazine and imidazole rings) can occur with *S*-alkylations, which would hamper yields and purifications.[192] Molecule **4.29** could not be purified and contained impurities, probably originating from impure commercial 3-bromopropan-1-ol. The sterics surrounding cyclohexyl bromide reduces ease of approach of the relatively large nucleophile **4.5**, and under RT conditions, no reaction occurred. Increasing the temperature to reflux overnight allowed the formation of **4.41** in low yield. Reaction of **4.5** with **4.48** was also unsuccessful, as under the basic conditions of the reaction, **4.48** may have undergone either an intramolecular cyclisation to form a 4-membered oxetane ring, or an intermolecular reaction with a hydroxide anion, to produce a 2,2-methoxyproan-1,3-diol or (-3-methoxy-1-ol) product. The ^{1}H NMR spectrum of the crude reaction material was inconclusive. All the other compounds reacted smoothly to provide reaction mixtures which were easily purified by column chromatography and further prepared for biological testing or further deprotection and derivatisation.
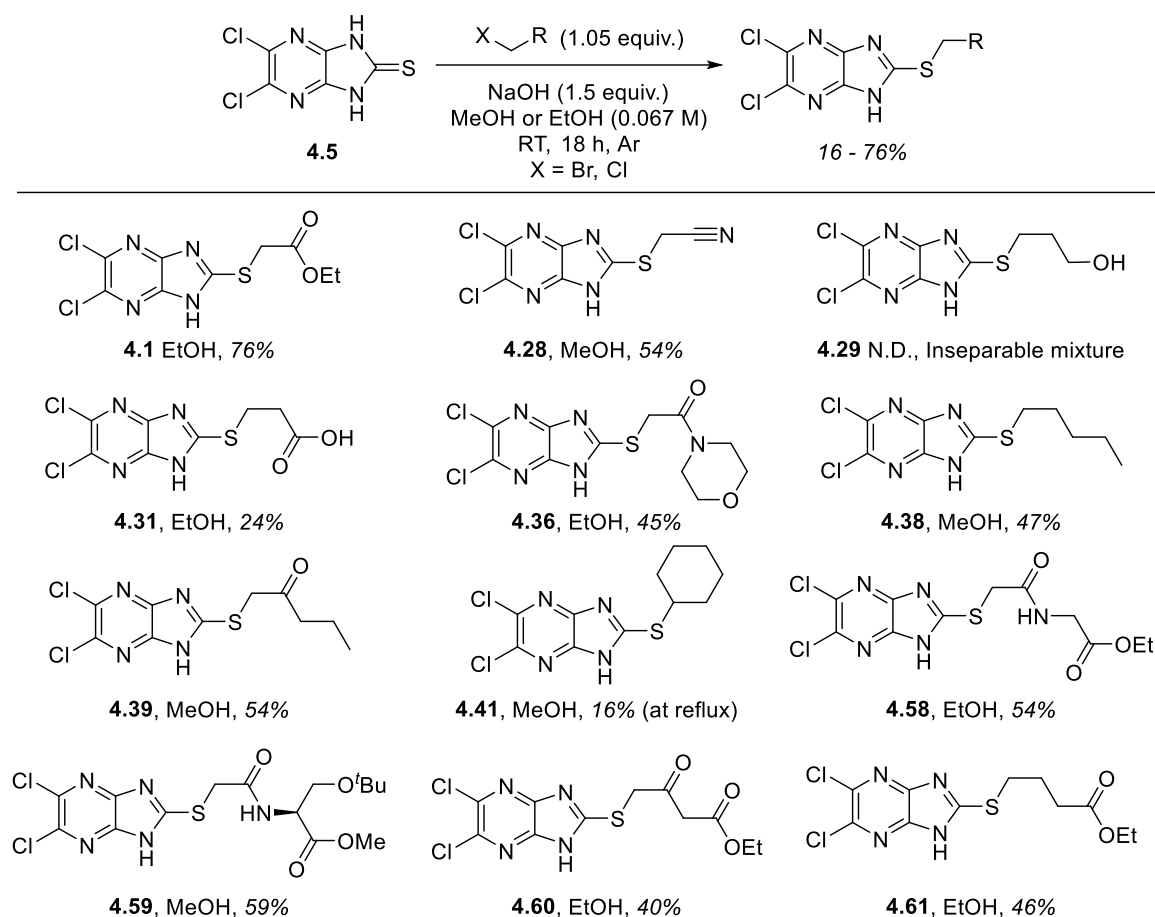
Scheme 4-12: S-alkylation for ethyl thioacetate variation library synthesis.

As mentioned in the paper detailing NSC-217913's synthesis, the introduction of ketone groups α to the thioether elicits an equilibrium between the free aminoketone and cyclic hemiaminol in the dissolved state. However, no such behaviour is observed in other α-carbonyl containing compounds where the carbonyl moiety is part of an ester or amide. This most-likely results from the partial double bond character and therefore reduced electrophilicity these functional groups impose, so that the carbonyl position is not so readily attacked by the imidazole nitrogen. This behaviour has been observed in benzimidazole compounds, and other systems containing this structural relationship.[193]

The two compounds **4.39** and **4.60** in this library have this motif and this behaviour is seen in the ^{1}H NMR spectrum. In both cases, the tricyclic hemiaminol system is not dominant, with the free aminoketone present in the larger ratios (**4.39**: 80:20, **4.60**: 76:24 free chain to hemiaminol, molar ratio in solution by ^{1}H NMR, CDCl$_3$). An example of an annotated spectrum for **4.39** is provided (Figure 4-10). The cyclic hemiaminol state renders the chain protons and thiazole ring protons diastereotopic, with recognisable constants observed for ^2J couplings (for **4.39** thiazole ring protons A, ^2J (Hz) = 12.3).

142
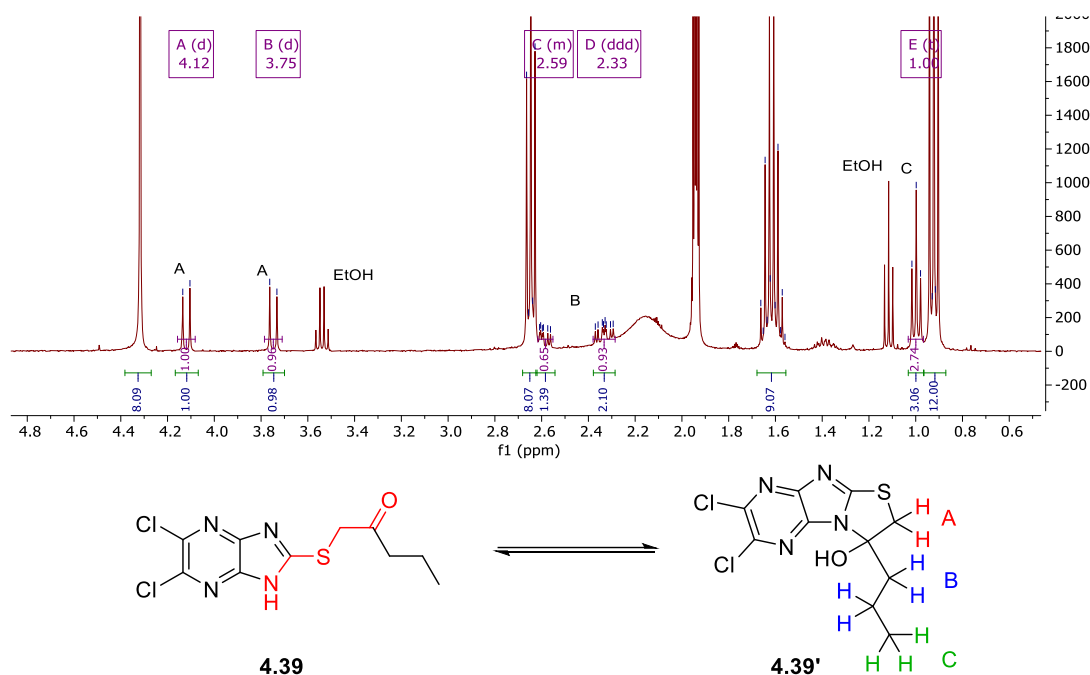
*Figure 4-10: Example of ketone-hemiaminol tautomerism observed in **4.39** by ^{1}H NMR.*

Several of these compounds were functionalised further, compound **4.59** was deprotected using phosphoric acid (Scheme 4-13) to provide **4.62** in reasonable yield.[194] After workup the material was taken forward for saponification as ^{1}H NMR analysis showed that essentially full deprotection had taken place and the material was pure enough to be used in the next step without further purification.
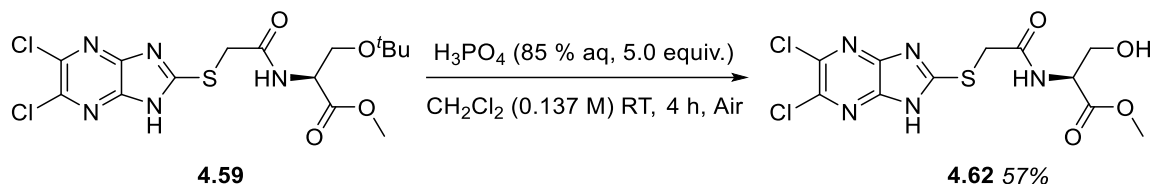
*Scheme 4-13: tert-Butyl ether deprotection of **4.59**.*

As several molecules in this series were either methyl or ethyl esters, a general saponification procedure to the free carboxylic acids was undertaken (Scheme 4-14). A procedure which does this under mild (RT) conditions was sought and one which uses a methanol : DCM solvent mixture was chosen for ease of workup and purification.[195] Compounds **4.33** - **4.35**, **4.37** & **4.43** were prepared from **4.1**, **4.58**, **4.60**, **4.61** and **4.62** in acceptable to excellent yields, except for **4.34**, which after isolation degraded as expected, most-likely through a concerted decarboxylation.
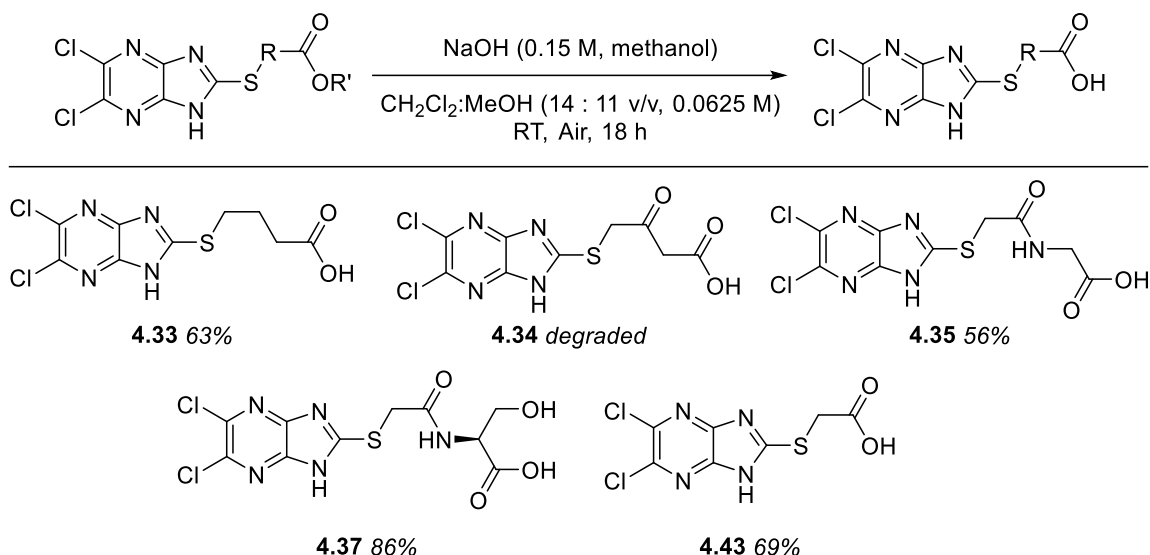
Scheme 4-14: Saponification of ester products

Optical rotation of **4.37** was found to be $[\alpha]^{25}_{589}$ = 115.38, indicating full racemisation had not occurred at the proton α to the acid and amide; this proton would be reasonably acidic, and so under harsher basic conditions may epimerise. Partial racemisation cannot be ruled out. Additionally, from **4.43**, an ethyl amide bond formation was pursued (Scheme 4-15), whereby the acetic acid derivative was treated with triphenylphosphine oxide and oxalyl chloride, forming a chlorophosphonium chloride active intermediate, which could react with the acid moiety to provide an acid chloride intermediate. From here this was reacted with ethylamine solution to provide **4.32**.[196] Unfortunately, whilst this was characterised as the desired compound, the mixture was impure with something unknown.
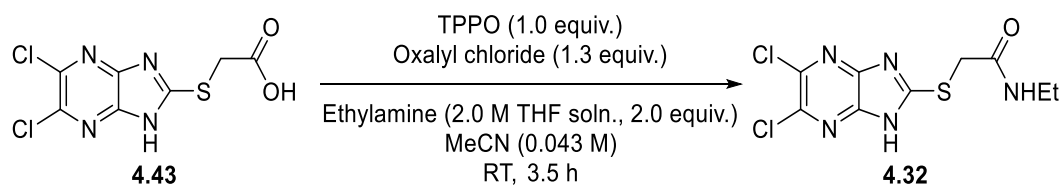
Scheme 4-15: Compound **4.32** production by amide bond formation.

From nitrile **4.28**, a reduction was attempted to provide the ethyl amine variant **4.27** as shown in Scheme 4-16. Instead of the desired reduction of the nitrile occurring, a reduction at the α-position to the nitrile occurred instead, providing mainly the synthetic precursor **4.5**. This is not wholly surprising, as the electron deficient 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione would arguably be rather stabilised as an anion through resonance and the aforementioned electron deficiency of the diazine heterocycle, and therefore act as a good leaving group.
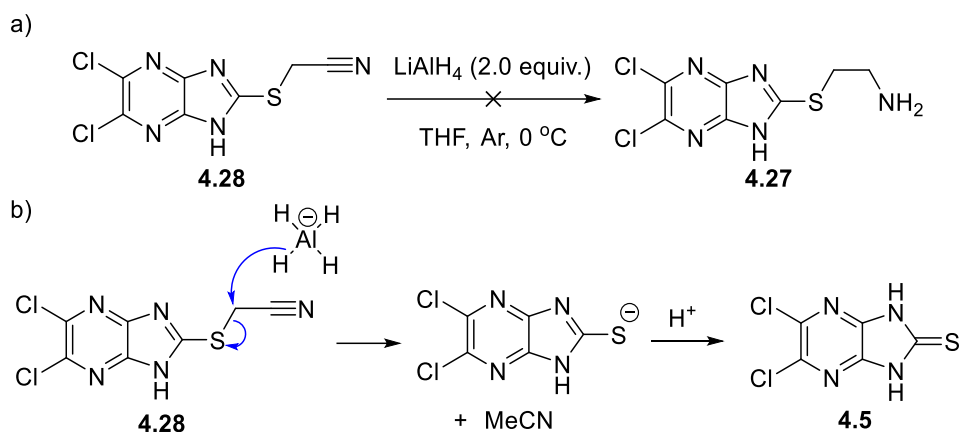
Scheme 4-16: a) Attempted nitrile reduction to form **4.27**, b) postulated reaction to form **4.5** under the reaction conditions.

Lastly, the synthesis of **4.42** was attempted, and it was thought this could be achieved by subjecting **4.2** with succinic anhydride in mixed xylenes at reflux, followed by esterification as described by Scheme 4-17.[197] However, even when heated for 48 h, the major product observed and partially isolated was the amide intermediate **4.63**, it seems the consequent cyclisation to the imidazole ring system forming **4.64** occurred in only very minor amounts. Later attempts with longer reaction times also proved unsuccessful and further attempts were not pursued.
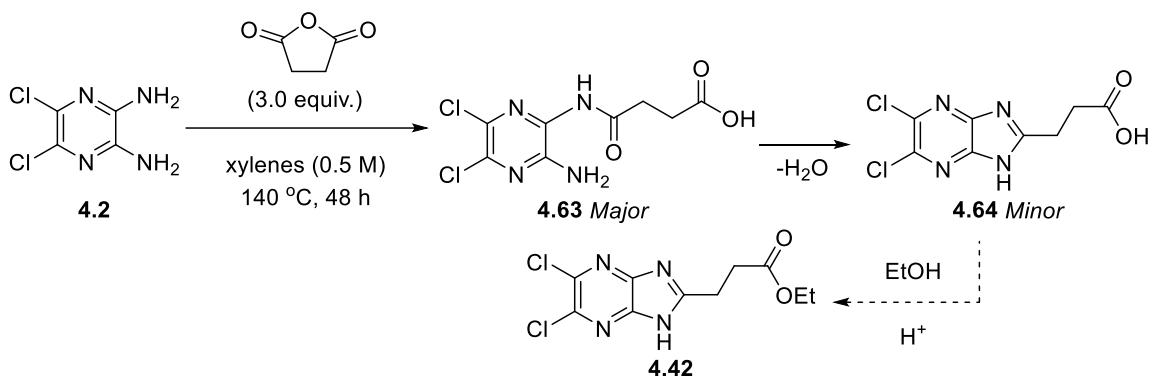
Scheme 4-17: Attempted synthesis of **4.42**.

In conclusion the objective to synthesise ethyl thioacetate derivatives of the NSC-217913 series and NSC-217913 itself was successfully met. Even though not all planned compounds from both libraries of compounds with reduced functionality and increased functionality based upon the Flare® docking software were produced, the vast majority have. Below is a summary of the compounds synthesised (Figure 4-11). The majority of these compounds are novel, with the exception of **4.5** and **4.1**, for which only limited historical characterisation data was available.
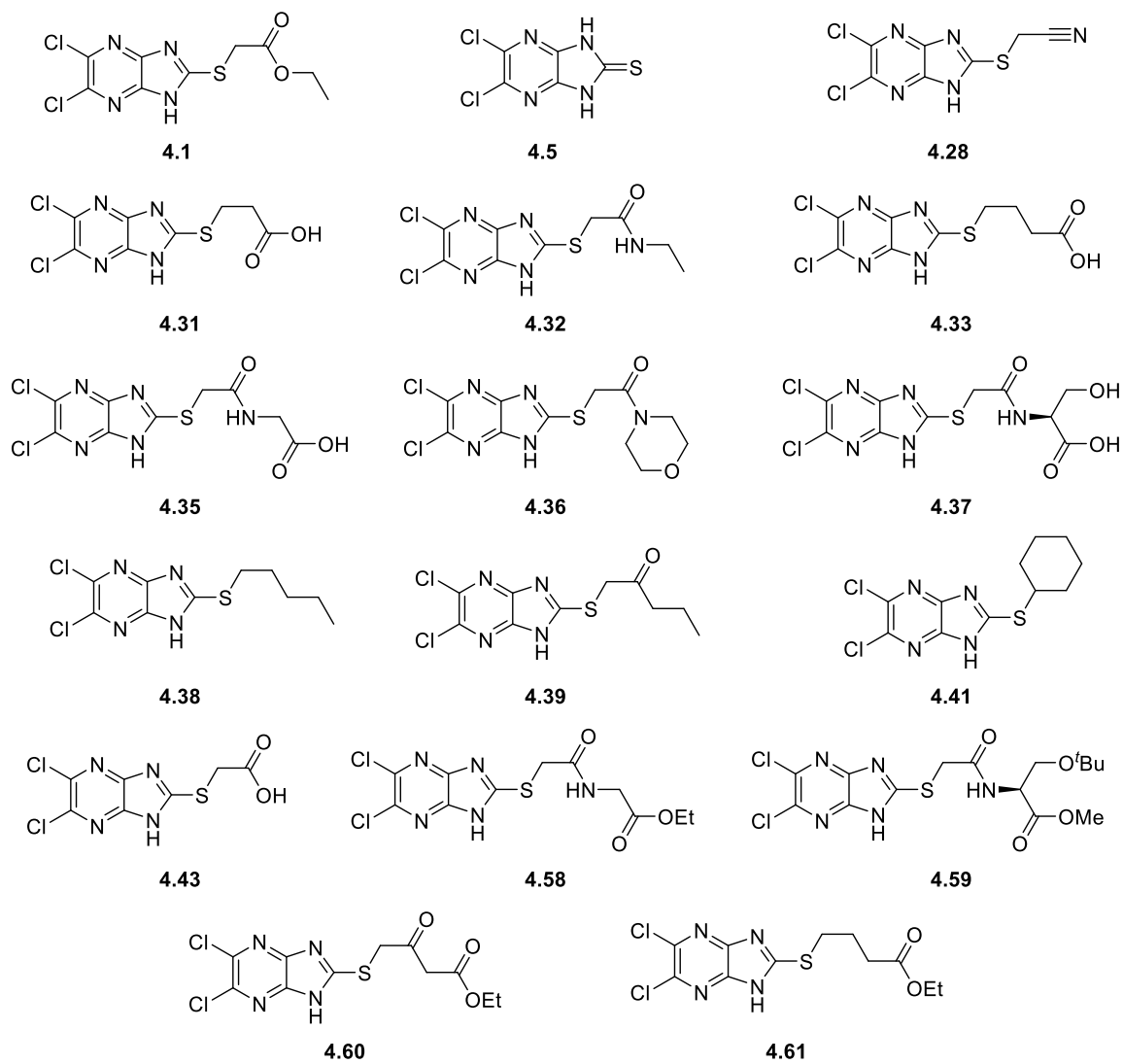
Figure 4-11: Summary of compounds synthesised for this series.

146

4.3.2 Synthesis III: Variation of the Heterocycle

4.3.2.1 6-Membered Ring Variations

The next objective was to vary the heterocyclic group of the structure, and this was further split into three parts for ease of library generation. The first subgroup was synthesised from diamino (di)azines, through a simple TCDI ring closure followed by *S*-alkylation. As a result, **4.65** – **4.70** derivatives were planned for synthesis in this manner (Figure 4-12).
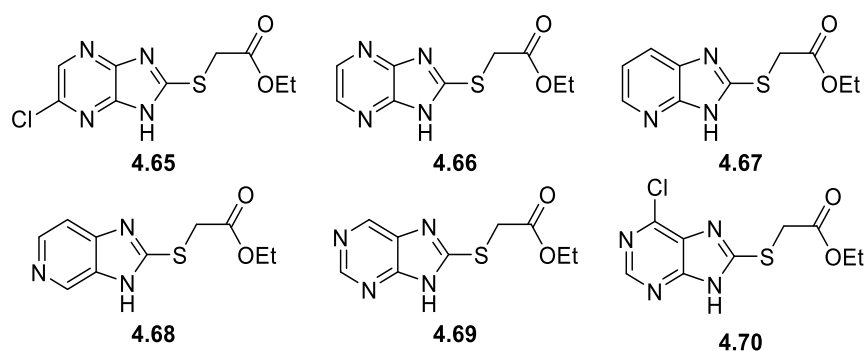
Figure 4-12: Planned heterocycle variations.

The second approach was to produce compounds with an imidazo[1,2-*a*]pyrazine system. No literature was present for the synthesis of derivatives of the **4.71** system (Figure 4-13), but synthesis of derivatives of structure shown **4.72** (boxed) was achievable by literature methods. Direct comparison is still available between this system and the previously discussed heterocyclic variant targets (Figure 4-12) as a derivative with the imidazo[4,5-*b*]pyrazine heterocycle without the thioether linkage could be used as a comparison between heterocycle libraries. Therefore, the plan changed slightly with **4.72**, **4.73** - **4.76** targeted instead.
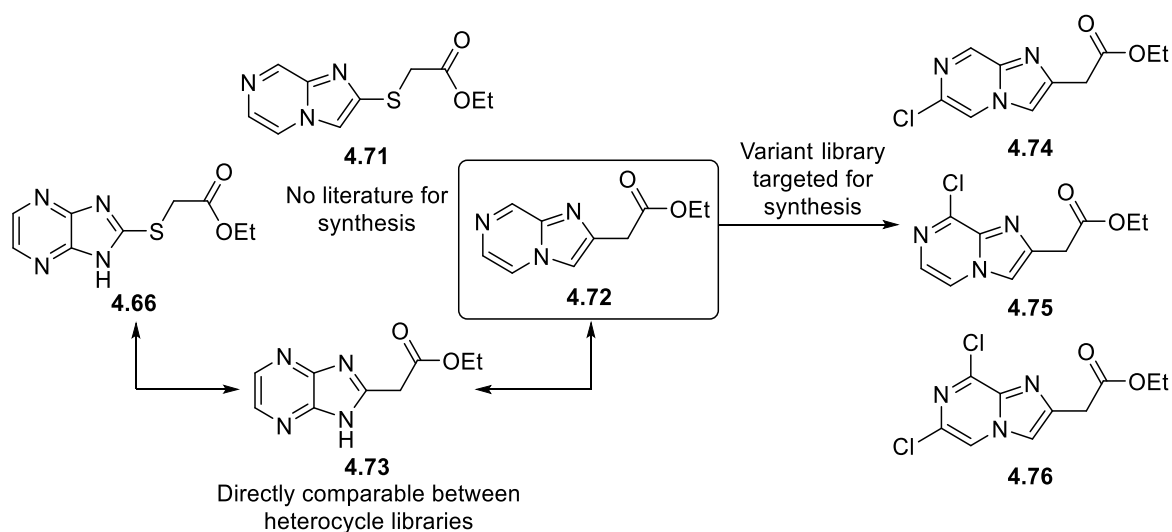
Figure 4-13: Plan to directly compare imidazo[4,5-b]pyrazine thioethers with imidazo[1,2-a]pyrazine variations targeted for synthesis.

Also planned for synthesis was a small library of imidazo[4,5-*d*]pyrimidine analogues **4.77** – **4.81** (Scheme 4-18, a). A general telescoped pathway was planned for a quick library production (Scheme 4-18, b). First, the pyrimidine starting materials would be reacted to form the nitroso intermediates, followed by sodium dithionite reduction to the 4,5-amino intermediates. These would then be reacted with carbon disulfide or TCDI to form the purine-8-thione skeletons, which would then be alkylated in the usual manner to form the thioether final products.
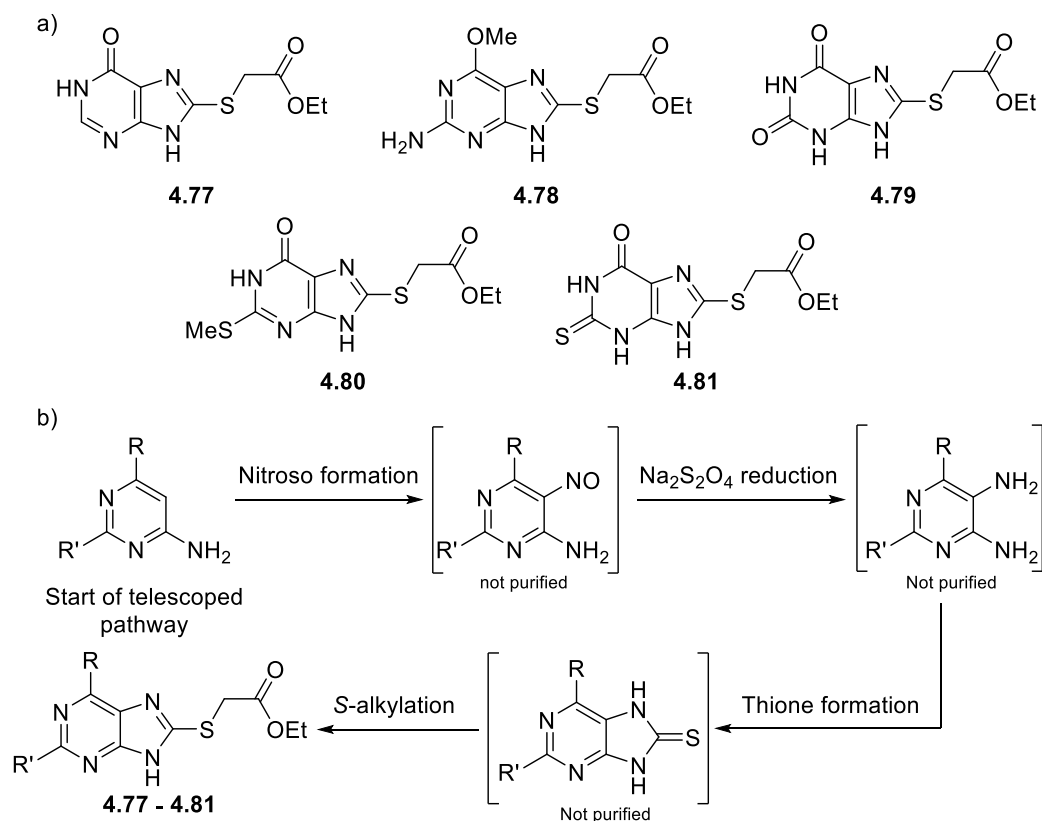
Scheme 4-18: Synthetic plan for purine derivatives synthesis.

To achieve the first part of this objective as outlined above, several diamino-(di)azines were commercially available, but 2,3-diaminopyrazine **4.82** had to be synthesised. Compound **4.15** was therefore subjected to several reductive dechlorination conditions to access **4.82**, and the experiments to achieve this is outlined in Table 4-4 below.
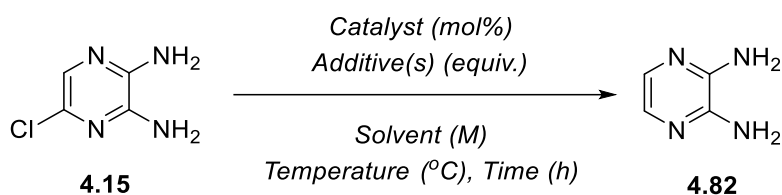
Entry	Catalyst (mol%)	Additive(s) (equiv.)	Solvent (M)	Temperature (°C)	Time (h)	Conversion* (%)
1	Pd/C (10, 5 wt%)	$H_{2\,(g)}$ (1 atm)	EtOH (0.16)	45	18	N.R.
2	Pd/C (10, 10 wt%)	Et_3N (10) HCO_2H (5)	Acetone (0.1)	75	4	N.R.
3	Pd/C (10, 10 wt%)	$H_{2\,(g)}$ (40 psi)	EtOH (0.16)	RT	18	N.R
4	Pd/C (10, 10 wt%)**	$H_{2\,(g)}$ (40 psi)	EtOH (0.16)	RT	18	N.R
5	$Pd(PPh_3)_4$ (5)	$NaHCO_2$ (3.0)	DMF (0.2)	100	2	N.R.
6	Pd/C (10, 10 wt%)	$H_{2\,(g)}$ (50 psi) MgO (2.5)	EtOH:H_2O (1:1, 0.34)	RT	18	25%
7	Pd/C (10, 10 wt%)	$H_{2\,(g)}$ (60 psi) MgO (2.5)	EtOH:H_2O (1:1, 0.34)	RT	72	34%

*Table 4-4: Tested conditions for reductive dichlorination of **4.15**. *As determined by ^{1}H NMR.*

***Activated by drying in a desiccator overnight, Pd/C otherwise un-activated.*

Initially, normal hydrogenation conditions were tested (entry 1) but no reaction was observed by ^{1}H NMR spectroscopy of the reaction mixture. Alternative transfer hydrogenation conditions (entry 2) were also attempted but again with no success.[198] Increasing the pressure of hydrogen gas (entry 3) was also non-productive. Changing to different transfer hydrogenation conditions utilising tetrakis(triphenylphosphine)palladium[0] and sodium formate was again unsuccessful as well.[199] The addition of MgO as an additive was attempted as it was observed that pyrimidine systems were reductively de-halogenated under similar conditions.[200] In the first run a 25% conversion was observed,

and attempts to drive this reaction towards higher conversions was met with an improved 34% conversion when the pressure was increased to 60 psi and time increased to 72 h.

It was thought that perhaps additional improvements could be had when increasing the equivalency of MgO in the reaction, as the pyrazine nitrogen coordination with the Mg atom in MgO may activate the proximal C-Cl bond towards reaction on the surface of Pd/C. 4 reactions, one under the conditions of entry 7 and three adapted from entry 7 (time changed to 24 h) provided conversions between 13 - 34%. Combining the crude materials from these reactions for purification provided a yield of 33%, based upon the initial amount of **4.15** added to these reactions. Additionally, 4,5-diamino-6-chloropyrimidine **4.84** was produced *via* a nucleophilic aromatic substitution of 5-amino-4,6-dichloropyrimidine **4.83** with ammonium hydroxide at 100 °C overnight, Scheme 4-19.
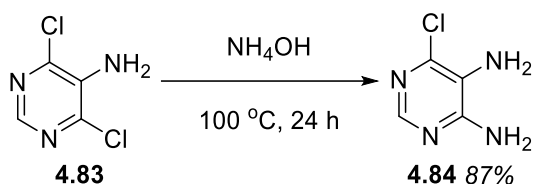
*Scheme 4-19: production of **4.84**.*

Commercial 4,5-diaminopyrimidine and 2,3-diaminopyridine, along with the diamino precursors **4.15**, **4.82**, **4.84** were reacted with TCDI to provide the thione products (**4.23**, **4.85** – **4.88**) under similar but more mild conditions to the formation of **4.5**. Compound **4.23** has been previously discussed above and its *S*-alkylation is included here as a heterocyclic variant. Simple *S*-alkylation using ethyl bromoacetate was performed utilising the same conditions as described previously to access heterocyclic derivatives **4.65** – **4.67**, **4.69** and **4.70** (Scheme 4-20).
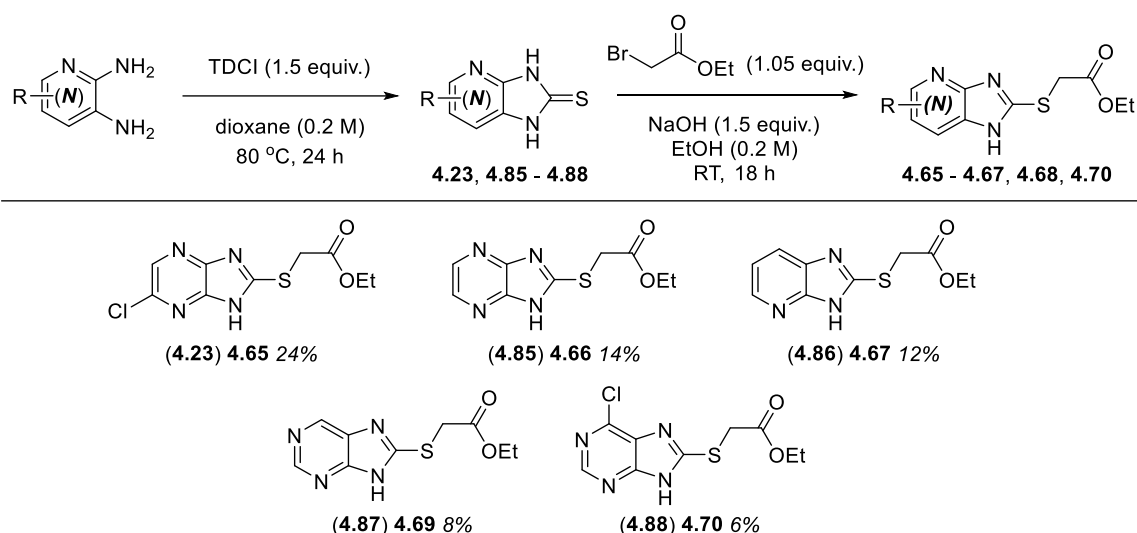
*Scheme 4-20: Synthesis of heterocyclic intermediates **4.23, 4.85 – 4.88** (respective thiones are bracketed) and thioether final products **4.65 – 4.67, 4.69** and **4.70**, yields over two steps.*

3,4-Diaminopyridine **4.89** was subjected to CS$_2$ ring closure conditions after the use of TCDI failed, but thione **4.90** could not be isolated in the normal manner and the crude product was taken through without purification after workup. As observed by ^{1}H NMR spectroscopy this reaction occurred with 76% conversion. Unexpectedly, after subjecting the material to thioether formation conditions, producing **4.91** followed by basic conditions, **4.92** was isolated doubly alkylated at both the sulfur and pyridine nitrogen position as confirmed by HMBC correlations. This additional alkylation likely occurred from the increased nucleophilicity of the pyridine nitrogen arising from the 1,4-positioning of the amino group highlighted below (Scheme 4-21).
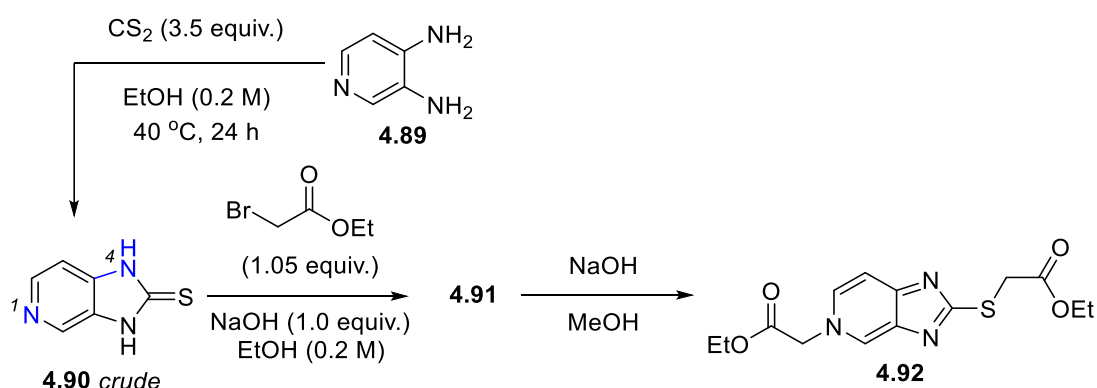
*Scheme 4-21: synthesis of **4.92**.*

Surprisingly, after isolation of **4.91** by column chromatography an additional broad singlet at 14.97 ppm was observed by ^{1}H NMR spectroscopy compared to the final material of this sequence, suggestive of an *NH*-proton, whilst other signals observed to have a slight shift (Figure 4-14). It is postulated that this material could be the salt shown in Figure

4-15, with either a bromide or chloride as counterion. Arguably, the column conditions support a marginally more polar material compared to the final product of this synthetic sequence (5% MeOH in DCM compared to 3 -> 4% MeOH in DCM). An additional piece of evidence is a comparison between specific signals in the two ^{1}H NMR spectra below. All signals in the aromatic region have a significant shift up-field going from **4.91** to **4.92**, along with the signals corresponding to the methylene positions. This could arguably stem from the pyridinium cation being electron-poor in its salt form when compared to after treatment with base.
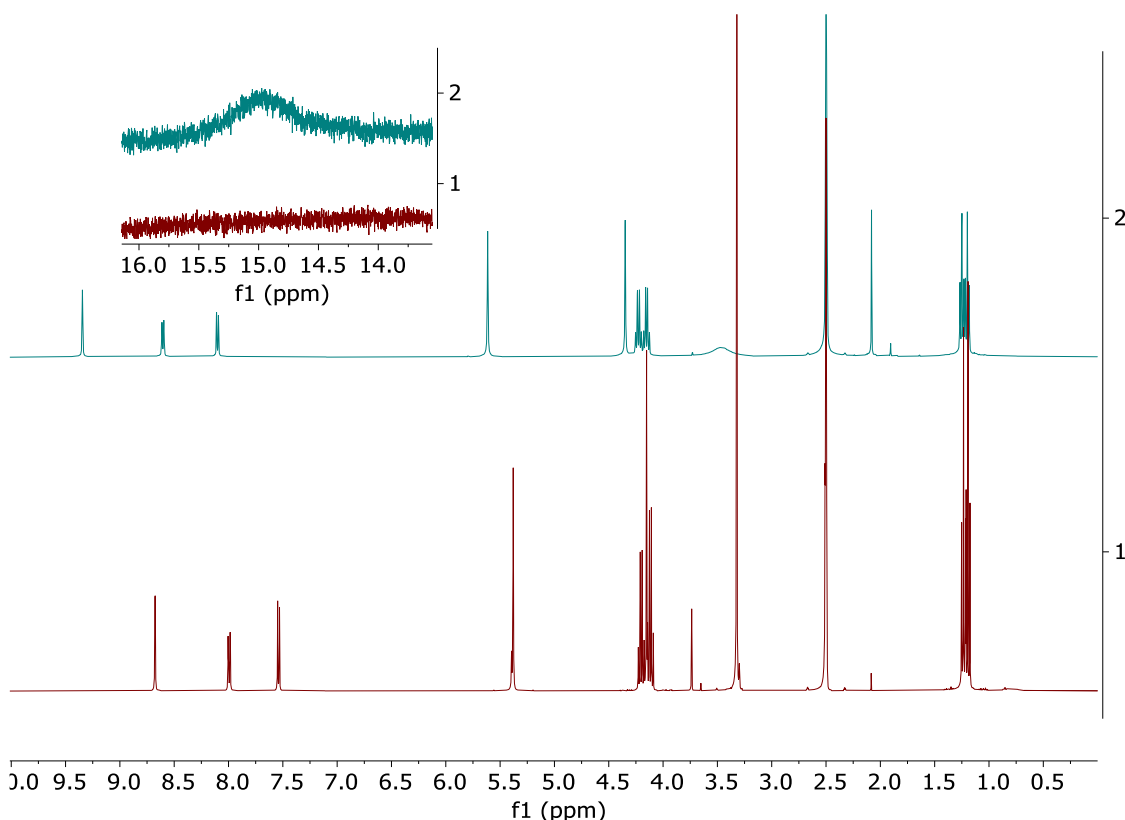
*Figure 4-14: Comparison of ^{1}H NMR spectra (DMSO-d$_6$) of **4.91** (top) and final product **4.92** shown in Scheme 4-21 (bottom), highlighted is the presence of the broad singlet at 14.97 ppm.*
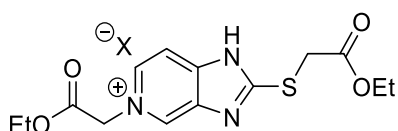
*Figure 4-15: Postulated structure of **4.91**.*

Pursuing the library of imidazo[1,2-*a*]pyrazine variants meant subjecting **4.13** to chlorination with 1.1 equiv. of NCS overnight in THF at 40 °C. This reaction provided three chlorinated derivatives **4.93**, **4.94** and **4.14** (Scheme 4-22).
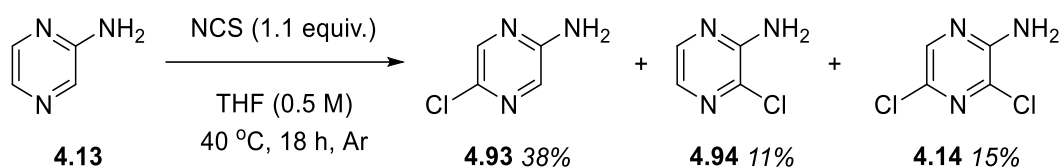
*Scheme 4-22: Production of chlorinated isomers **4.14**, **4.93** and **4.94**.*

From here, with **4.13**, **4.14**, **4.93** and **4.94** were subjected to reflux in ethanol with ethyl 4-chloroacetoacetate **4.95** to provide the ethyl imidazo[1,2-*a*]pyrazine-2-acetate derivates **4.72**, **4.74**, **4.76** and **4.96**. These compounds were rather difficult to purify, requiring sequential column chromatography followed by recrystallisation to provide high purity samples fit for biological analysis (Scheme 4-23). Unfortunately, **4.75** was not produced, instead **4.96** was obtained, most-likely forming through **4.75** *via* an S$_N$Ar mechanism with the residual water in the EtOH. Attempts to form **4.75** from **4.96** using POCl$_3$ were unsuccessful (not shown).
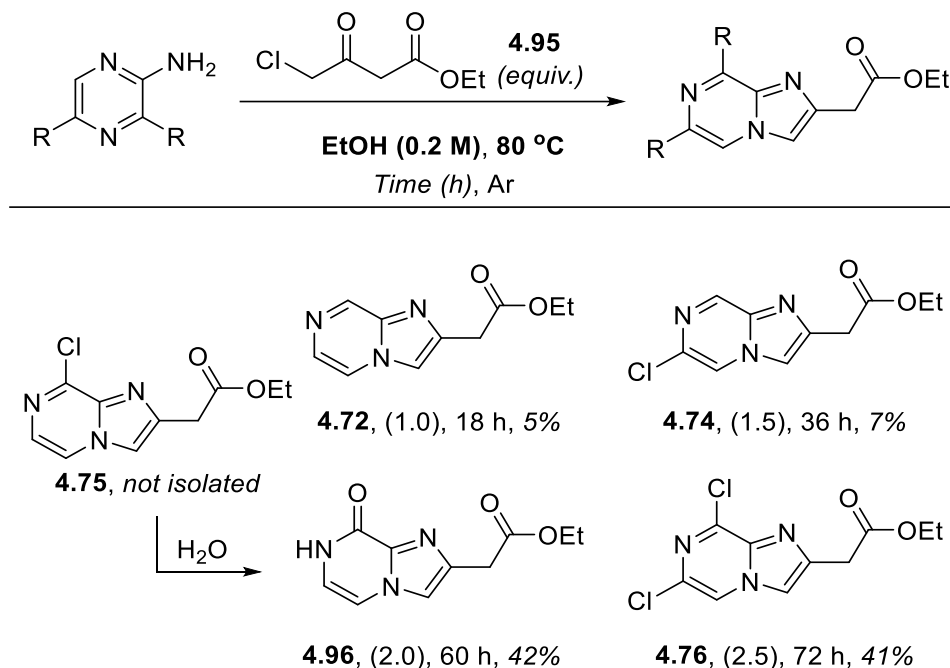
*Scheme 4-23: Synthesis of imidazo[1,2-a]pyrazine derivatives **4.72, 4.74, 4.76** and **4.96**. Bracketed is **4.95** equivalency.*

The mechanism for the formation of the imidazo[1,2-*a*]pyrazine core could be as shown in Scheme 4-24. Initially, the ring nitrogen *ortho* to the amino group in **4.13** would attack the α-position of **4.95** to produce ***int-4.97***, which would then rapidly undergo an intramolecular cyclisation with proton transfer to ***int-4.98***. After this, arguably an aromatisation would occur from loss of the amino groups remaining proton to neutralise the charged ***int-4.98***, providing ***int-4.99***. Finally, elimination of the tertiary alcohol would complete aromatisation and produce the final imidazo[1,2-*a*]pyrazine core of **4.72**.
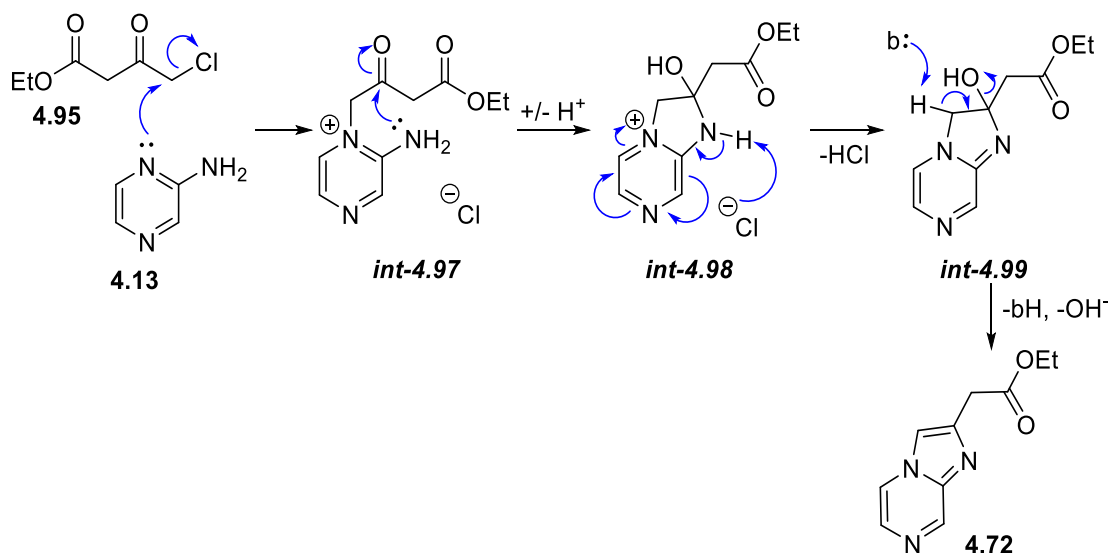
Scheme 4-24: Proposed mechanism for the formation of imidazo[1,2-a]pyrazine heterocycles.

From the literature used to produce these compounds it was known that a 5*H*-pyrrolo[2,3-*b*]pyrazine by-product was possible, but this was only observed for **4.72** (Scheme 4-25, **4.100**). Some details were provided on the postulated mechanism, but the literature was vague in describing how exactly reagent isomerisation of ethyl 4-chloroacetoacetate could occur.
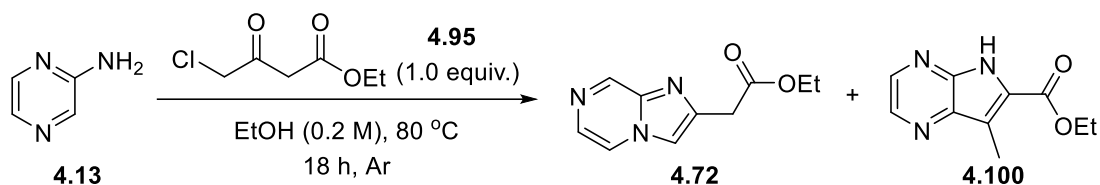
*Scheme 4-25: Formation of **4.72** with by-product **4.100**.*

Firstly, after partial reagent isomerisation, however this occurs (**4.101**), instead of **4.13** endocyclic nitrogen attack as described previously for imidazo[1,2-*a*]pyrazine core formation, the primary amine nucleophilically displaces the chlorine at the 2-position of **4.101** providing ***int-4.102*** (Scheme 4-26). Then, attack of the pyrazinic ring *ortho* to the amino group occurs on the acetyl group intramolecularly with protonation, probably from the solvent, giving ***int-4.103***. Neutralisation of the positive amine and re-aromatisation of the pyrazine ring occurs by removal of the proton from the position *ortho* to the amino group providing ***int-4.104***. Lastly, elimination of the alcohol group forms the pyrrole ring system and the product **4.100**. Unfortunately, this material was isolated impure and in very small quantity. It is unlikely that any of the other derivatives used to produce this library could also form this by-product due to the presence of the chlorine atom at the 3-position, apart from **4.74**, although this was never observed.
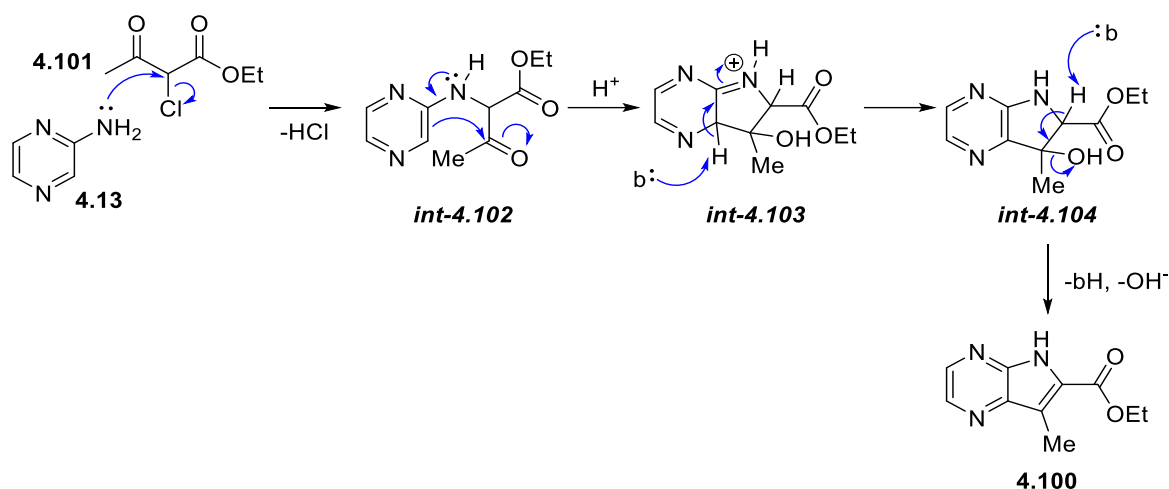
*Scheme 4-26: Postulated mechanism for the formation of by-product **4.100**.*

To finish this small series, variant **4.73** was planned for synthesis (Scheme 4-27). This would entail an initial amide bond formation followed by cyclisation and formation of the imidazole ring with acid chloride **4.105**, which could be produced easily from potassium ethyl malonate. Unfortunately, in the interests of time on this part of the project these reactions were not undertaken, instead purine derivatives were next targeted for synthesis.
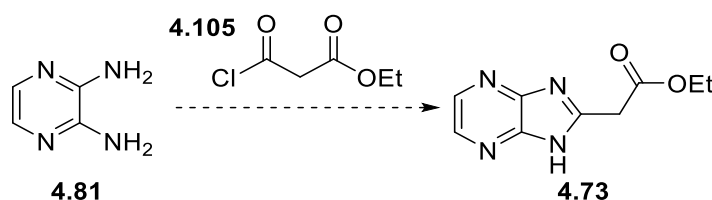
*Scheme 4-27: Planned synthesis for **4.73**.*

To pursue the purine series, this initially meant some modification of commercially bought aminopyrimidines was required as shown (Scheme 4-28). This included a *S*-methylation of **4.106**, which occurred in 91% yield providing **4.107**, and a Raney-Ni reduction provided **4.108** in 85% yield.[201] Incorporation of a methoxy group *via* S_NAr of **4.109** occurred in 34% yield to give **4.110**.
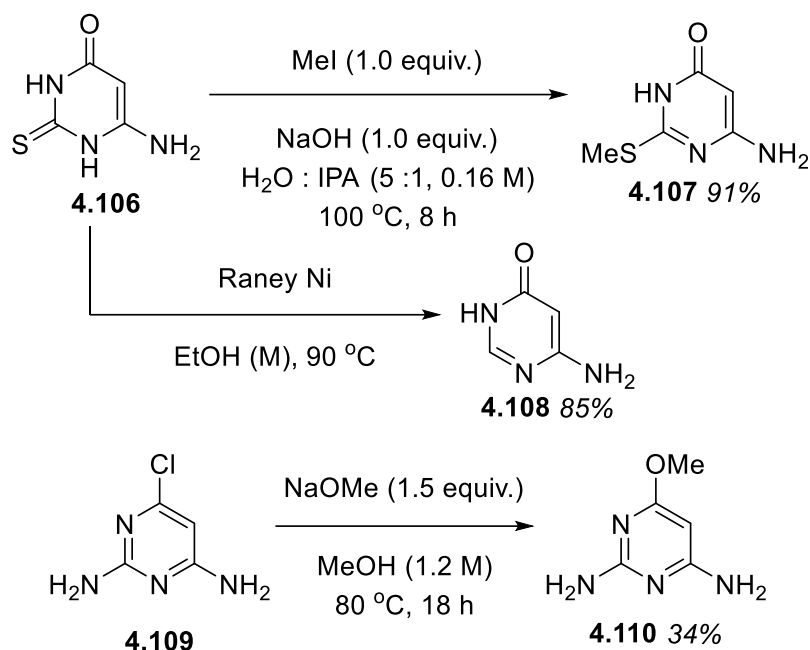
Scheme 4-28: Synthesis of pyrimidine precursors **4.107** , **4.108** and **4.110**.

Following the synthesis of these pyrimidine precursors, nitroso formation was conducted using sodium nitrite in an acetic acid-water mixture at various temperatures as indicated (Scheme 4-29). Unfortunately, several compounds provided inconclusive results or degraded upon isolation (**4.111**, **4.113**, **4. 115**). Even with these results, compounds **4.111** and **4.115** were taken forward anyway as it was thought any water content would shift or change the observed signals due to all protons being exchangeable in these molecules.
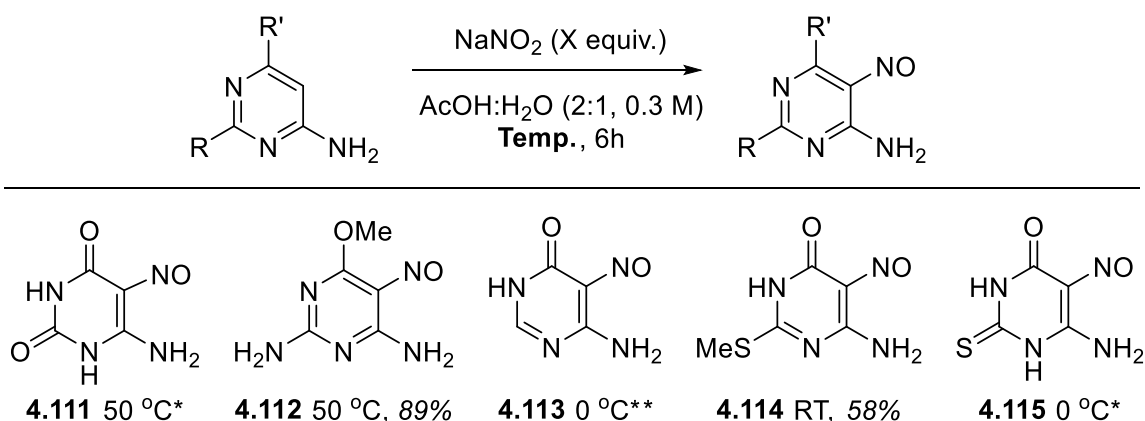
Scheme 4-29: *no crude ^1H NMR signals indicated product formation.** degraded at RT.

With the nitroso compounds in hand, sodium dithionite reduction was performed to access the diamino intermediates, followed by CS_2 ring closure to access the purine-8-thione ring systems (Scheme 4-30). Molecule **4.119** was thought to not have formed, as no signals corresponding to the desired compound could be found in the crude ^1H NMR spectrum from CS_2 ring closure. Additionally, compound **4.117** was thought to have formed due to

the 5 carbon resonances observed by ^{13}C NMR spectroscopy. However, this material was very insoluble and attempts to use it under the *S*-alkylation conditions failed.
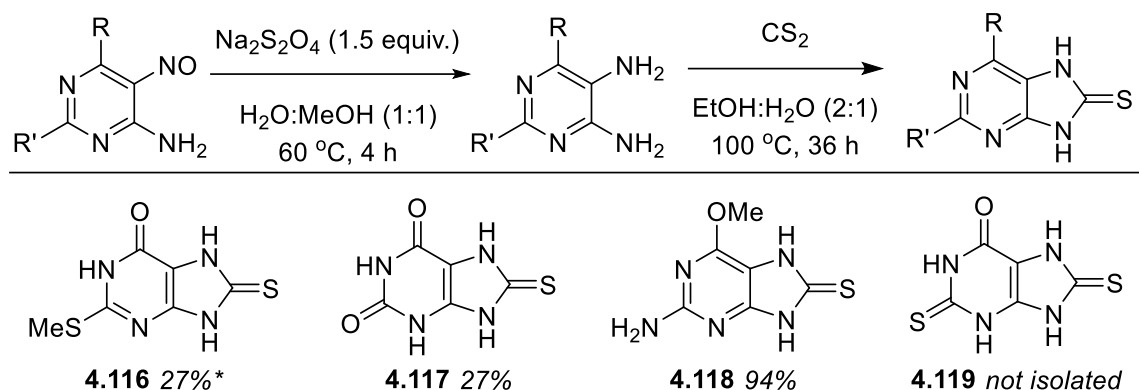
*Scheme 4-30: General telescoped nitroso reduction and thione formation reactions. Yields are over two steps. *Nitroso reduction of **4.116** performed with Na₂S₂O₄ (1.5 equiv.), NaOH (3.0 equiv.) in water (0.3 M) at RT.*

S-Alkylation was successfully performed on **4.116** utilising the now standard reaction conditions as above to give **4.120** (Scheme 4-31). Unfortunately, harsher conditions were needed for **4.118**, and these resulted in saponification of the resulting ester, providing **4.121**. Heating **4.121** in EtOH to reflux with concentrated H₂SO₄ allowed isolation of the ethyl ester product **4.122**. Extensive purification of both compounds was required, needing sequential column chromatography and recrystallisation to provide material pure enough for bioassay experiments. This was probably a result of the route being telescoped and minimal purification being performed over 4 steps.
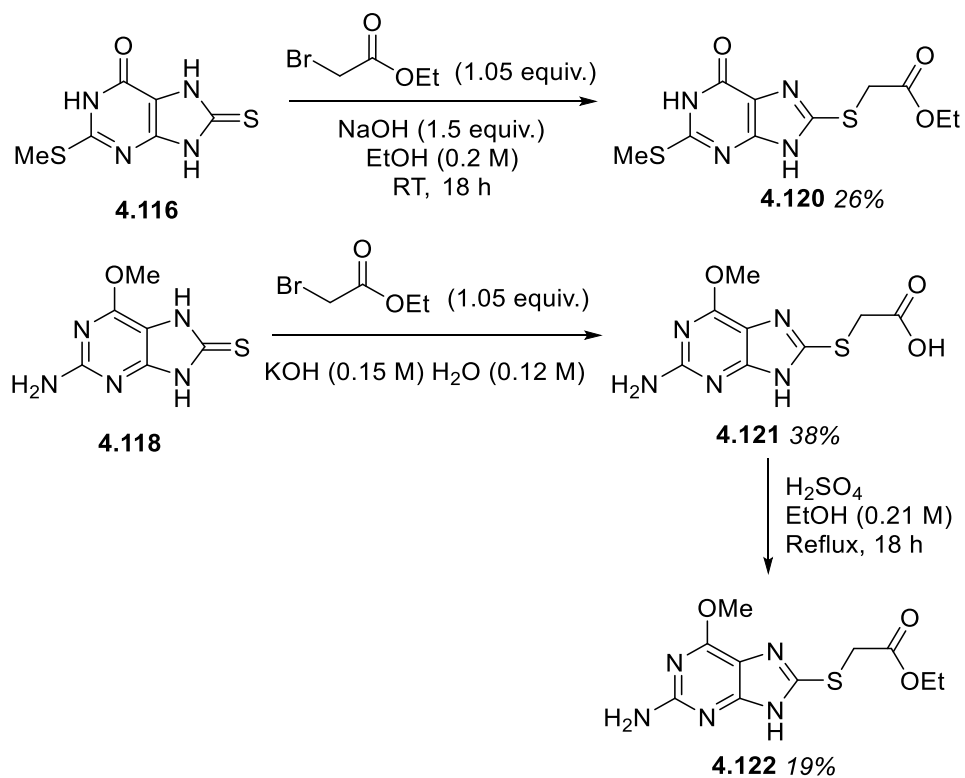
*Scheme 4-31: Synthesis of final products **4.120** & **4.122**.*

The summary of the heterocycle variations successfully synthesised is presented below (Figure 4-16). This library was much more difficult to synthesise compared to the sidechain variations. This is primarily because the practicality of synthesising variants differing in a side chain bearing one heterocycle is vastly easier compared to essentially *de novo* synthesis of each variant bearing different heterocycles with a consistent side chain, which may only be installed in the final stages of the synthetic sequences. Nonetheless, the results summarised below show that this part of the objective has been met, the majority of compounds targeted have been successfully synthesised. The heterocyclic variations where derivatives of the purine core bearing various groups was rather unsuccessful within this series, with only two compounds synthesised.
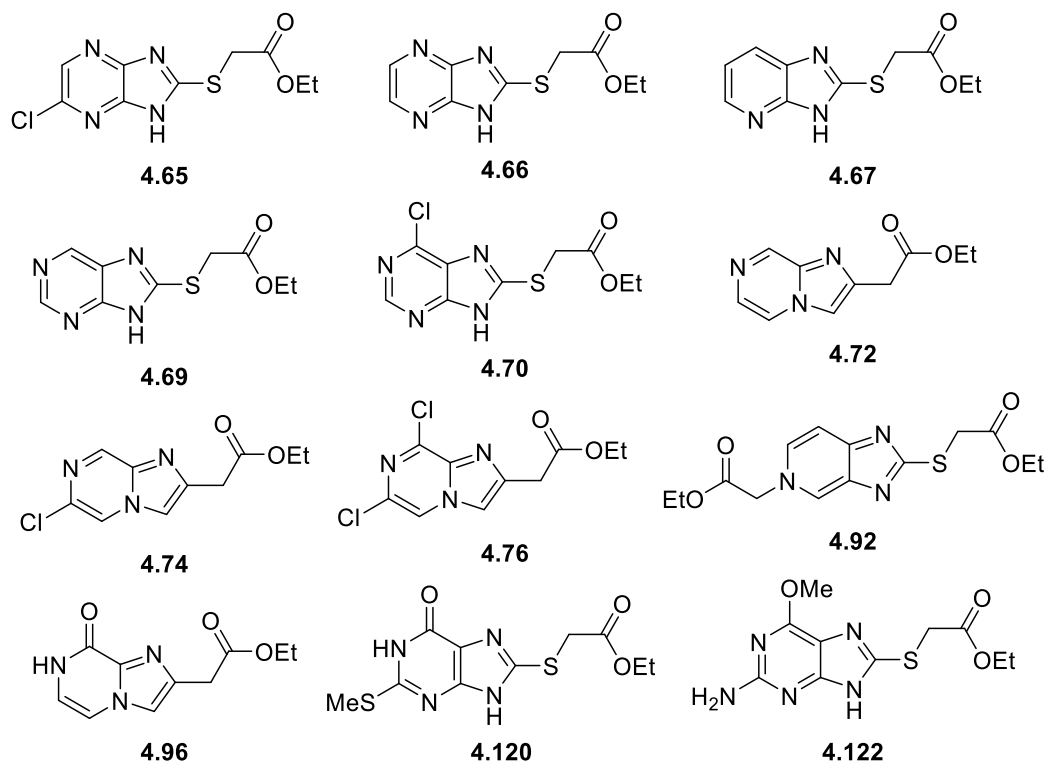
Figure 4-16: Summary of compounds synthesised as part of the heterocyclic variation series.

4.4 Initial Biological Results

All compounds isolated in sufficient purity from the *S*-alkylation series were passed on for testing in the biochemical assay and the results are presented below. Firstly, only a few active compounds displayed IC$_{50}$ values even though all were soluble at 10% DMSO in 1 mM and the majority at 10 mM (see table for exceptions). Unexpectantly, resynthesised **4.1** was not able to corroborate the NCI NSC-217913 sample hit IC$_{50}$ value and was not soluble in 10% DMSO at 10 mM. The carboxylic acids **4.43** and **4.31** (Table 4-5, entries 2 and 4 respectively) along with the ethyl butyrate ester (**4.61**) and acetylmorpholine (**4.36**) derivatives (entries 6 and 12 respectively) provided much worse IC$_{50}$ values than the original hit, for unknown reasons. Entry 10 (**4.5**) provided an improved IC$_{50}$ value over the original hit by roughly a factor of 10 (145.5 *vs.* 14.6 µM). This represents a marked improvement of inhibitory activity against WWP2. Compounds synthesised as part of the heterocyclic variation series are also presented below. None of the compounds provided an IC$_{50}$ value, even though all samples were soluble in the assay.

Entry	Structure	IC$_{50}$ (µM)	Entry	Structure	IC$_{50}$ (µM)
1 (4.1)		N/A*	2 (4.43)		666.4
3 (4.38)		N/A*	4 (4.31)		603.3
5 (4.39)		N/A	6 (4.61)		269.2*
7 (4.28)		N/A	8 (4.59)		N/A*
9 (4.58)		N/A*	10 (4.5)		14.6

11 **(4.35)**		N/A	**12** **(4.36)**		621.8
13 **(4.37)**		N/A	**14** **(4.66)**		N/A
15 **(4.69)**		N/A	**16** **(4.74)**		N/A
17 **(4.67)**		N/A	**18** **(4.96)**		N/A
19 **(4.70)**		N/A	**20** **(4.76)**		N/A
21 **(4.80)**		N/A	**22** **(4.92)**		N/A
23 **(4.72)**		N/A			

*Table 4-5: Biological results for the S-alkylation compound library. *Denotes not soluble in 10% DMSO at 10 mM.*

As a result of the rather unsuccessful bioassay results, these compounds were subjected to DSF assay to see if they are interacting with WWP2 at all. The compounds **4.74**, **4.76**, **4.96**, **4.58**, **4.35** and **4.92** all stabilised WWP2 (Table 4-6). Compound **4.5** (entry 10, Table 4-5) did not provide any significant change in the melting point of WWP2, even though it provided a significant inhibitory effect from the IC_{50} assay. This does not necessarily mean the compound is not interacting with WWP2, as it could be that upon its interaction with the enzyme it neither stabilises nor destabilises WWP2.
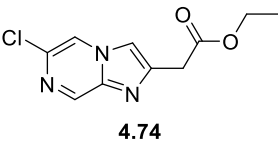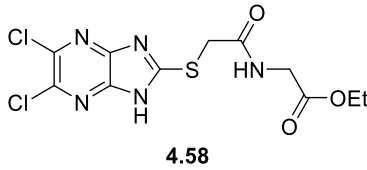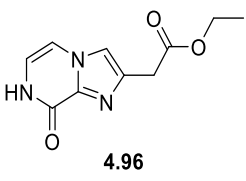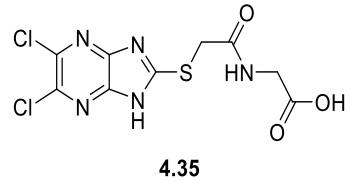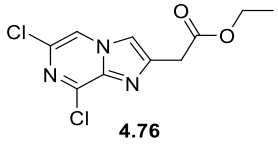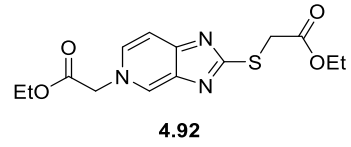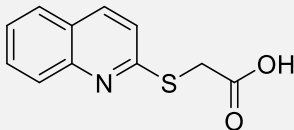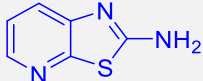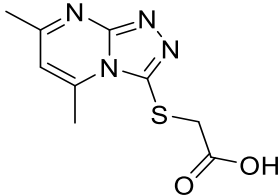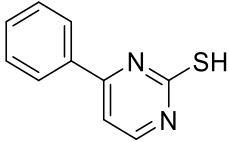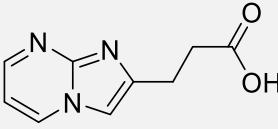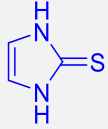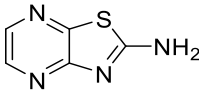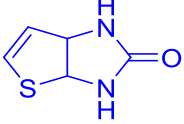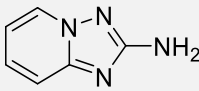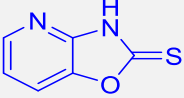
Structure	4.74	4.58
DSF Result	**Stabilising**	**Stabilising**
Structure	4.96	4.35
DSF result	**Stabilising**	**Stabilising**
Structure	4.76	4.92
DSF result	**Stabilising**	**Stabilising**

Table 4-6: DSF results from the S-alkylation and heterocyclic variation series.

4.4.1 <u>Commercial Heterocycles</u>

During this investigation a small library of commercial heterocycles were purchased as this expedited the investigation into heterocyclic variation. Table 4-7 describes the heterocycles purchased, all were confirmed by ^{1}H and ^{13}C NMR to be pure enough for bioassay without further purification. These were initially planned to have an ethyl acetate side chain by functionalising either the exocyclic heteroatom or sulfanyl-acetic acid motif. Based on the initial results presented above on this project it was decided that this was not necessary as this portion of the molecule did not apparently provide the activity observed. As with the *S*-alkylation products and heterocyclic variations, the majority provided no activity. All compounds were soluble 10% DMSO at 1 mM. The majority of the IC$_{50}$ values are provided by the heterocycles containing the cyclothiourea or cyclourea moieties, although an amine was also shown to be active (entries 2, 6, 8, 10 and 12). This is in alignment with the *S*-alkylation results (Table 4-5, entry 10).

Entry	Structure	IC$_{50}$ (µM)	Entry	Structure	IC$_{50}$ (µM)
1		N/A	2		292.8
3		N/A	4		N/A
5		N/A	6		78.9
7		N/A	8		152.6
9		N/A	10		32.5

11	Br–[thiazolo-pyrazine]–NH$_2$	N/A	**12**	[methyl-thiazolo-pyridine-thione]	49.3
13	[isoxazolo-pyrazine-NH$_2$]	N/A			

Table 4-7: Commercial heterocycles biological results.

At this stage, two things can be drawn from these results, firstly, the re-synthesised NSC-217913 material is not active and has not reproduced the activity observed from the NCI sample of NSC-217913. This is a cause for concern as it implies that either the NCI sample is impure or has degraded, or what I synthesised was not the desired product. However, full characterisation was possible for this material with all evidence pointing towards the desired product being synthesised. Further analysis of the original NCI sample purity is presented later (section 4.7). The second point is that activity is coming from the cyclo(thio)urea motif of these heterocycles, which are providing activity better than NSC-217913. It was decided to further investigate compounds of this type, focussing on **4.5** due to it providing the highest activity.

4.5 <u>Synthesis IV: Further Investigation of the Thione Group</u>

Four objectives were investigated as part of the further investigations into the active thiourea compounds. Firstly, functionalising the sulfur atoms with a small methyl group was targeted on all the molecules bearing the cyclothiourea motif (Figure 4-17). This was to see if any groups on the sulfur were tolerated or whether just the free sulfur was required for activity. Secondly, substituting the sulfur atom for oxygen and carbon were targeted, in order to see if the sulfur could be substituted at all to maintain activity. The targeted molecules within these two objectives (apart from **4.4** targeted in the second objective) would also investigate whether loss of an NH proton would be significantly impacting on inhibitory activity. The next objective was the third initial objective described in section 4.2, but modified to incorporate the compounds with the free thiourea functional group to see if additional interactions were possible away from the pyrazine ring (Figure 4-17). Restricting the side chain by cyclising with the proximal imidazole amine was also investigated as the final objective, to see if the second free N-H was also an important for providing activity. Installing carbonyl groups on this restricted side chain would provide different orientations the carbonyl would sit with respect to the rest of the heterocycle and

may offer information on which orientation is optimal for binding. Again, these objectives would be successful if compounds from each objective could be synthesised and submitted for bioassay.
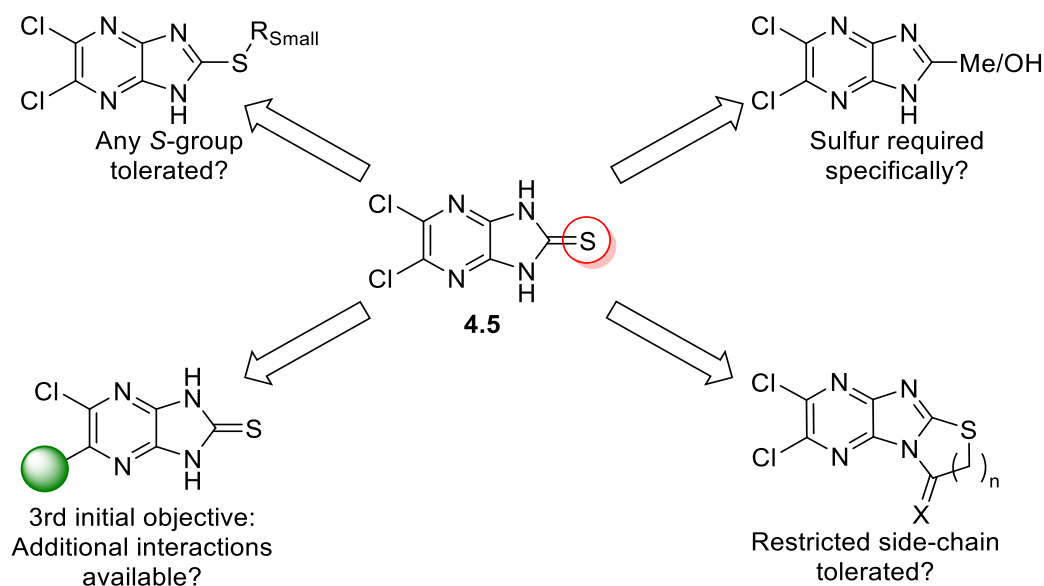
Figure 4-17: Summary of further investigation objectives.

4.5.1 Small Thioether Chains

To achieve the first objective, the active heterocycles were subjected to the alkylation conditions previously described utilising methyl iodide, as this is the smallest carbon moiety that can be attached. All compounds except **4.126** were *S*-methylated in acceptable yields with methyl iodide in ethanol (**4.123** – **4.125**, Scheme 4-32). A different set of conditions were found for **4.126**, providing **4.127** 61% yield.[202]
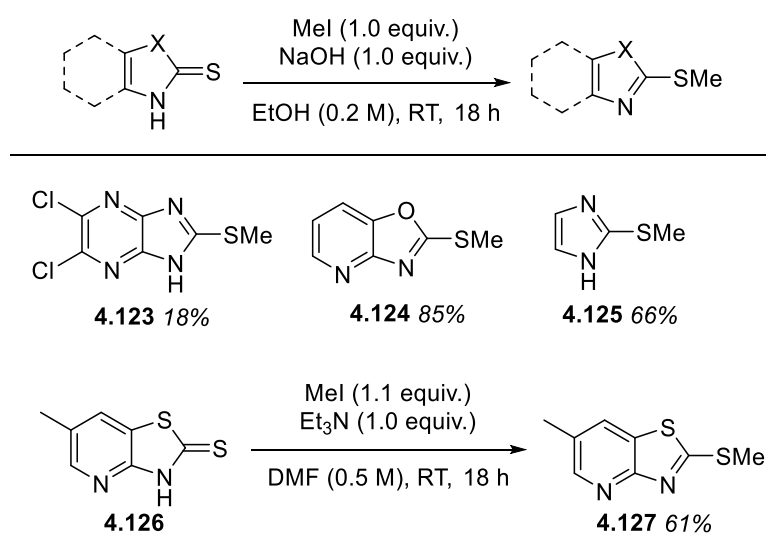
Scheme 4-32: Methylation of active cyclothiourea heterocycles.

4.5.2 Substituting Sulfur

The next objective had already been partially achieved as compound **4.4** had already been synthesised as described in section 4.3.1.1. Compound **4.2** was reacted with acetic anhydride to access the carbon derivative **4.128** in moderate yield (Scheme 4-33).[197]
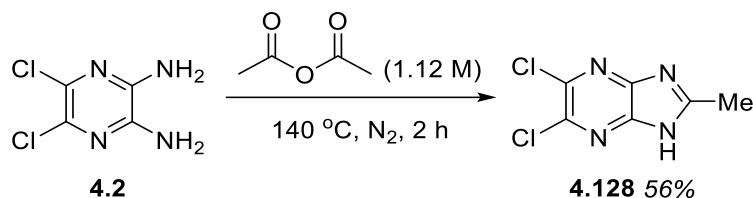
Scheme 4-33: Synthesis of 2-methyl-5,6-dichloro-1H-imidazo[4,5-b]pyrazine.

4.5.3 Potential for Additional Interactions?

In order to pursue the third main objective for this compound series, conditions had to be researched for cross-coupling, but early on it was found that thione and thiol functionalities are some of the worst tolerated under Pd cross-coupling conditions. This is because of their strong ability to coordinate Pd and hence result in catalyst poisoning by forming stable off-cycle Pd resting states, with only thiophenes showing competency in the cross-coupling reactions.[203] Because compound **4.65** was readily available, a test reaction was attempted by coupling with phenylboronic acid.[204] This resulted in majority hydrolysis occurring (forming **4.130**) and what is thought to be a Liebeskind–Srogl cross-coupling occurring (Scheme 4-34) to provide **4.129** in 5%. In this type of coupling the heteroarylthioether is activated (usually in the presence of a copper additive, but in this case absent) and the palladium centre oxidatively adds to the heteroaryl carbon – sulfur bond.[205] This would be followed by a transmetallation with the boronic acid and reductive elimination occurring to form the product isolated in this case in trace amounts.
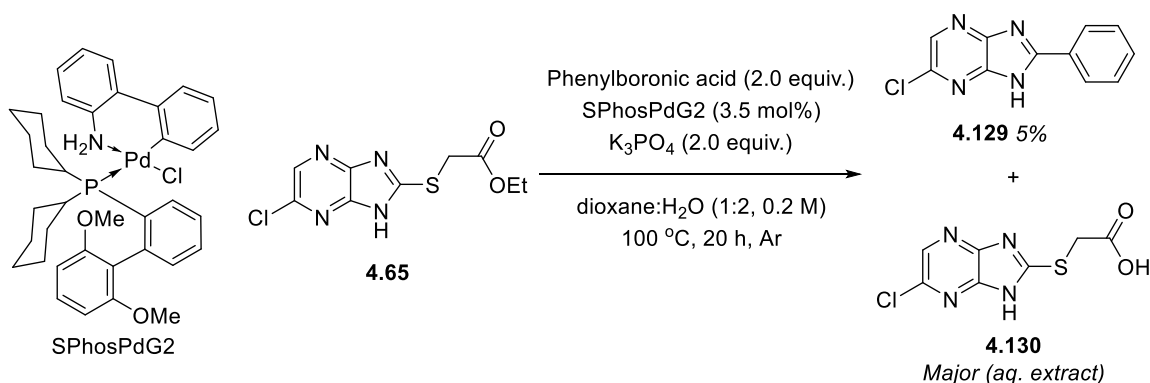
Scheme 4-34: Attempted S-M cross-coupling with phenylboronic acid.

If time permitted, the more obvious method to access these derivatives would start with the resynthesis of **4.4**, followed by cross-coupling and then thionation (Scheme 4-35**,** a), as literature conditions are present for reactions tolerating the cyclourea motif.[174,206] Additionally, some protecting groups have been devised for thiols which allow Suzuki-Miyaura couplings to be performed (Scheme 4-35, b).[203,207] A potential issue with these however is that these protecting groups have only been trialled on aromatic thiols without acidic imidazole protons present, so may not be a reliable route for target compound synthesis.
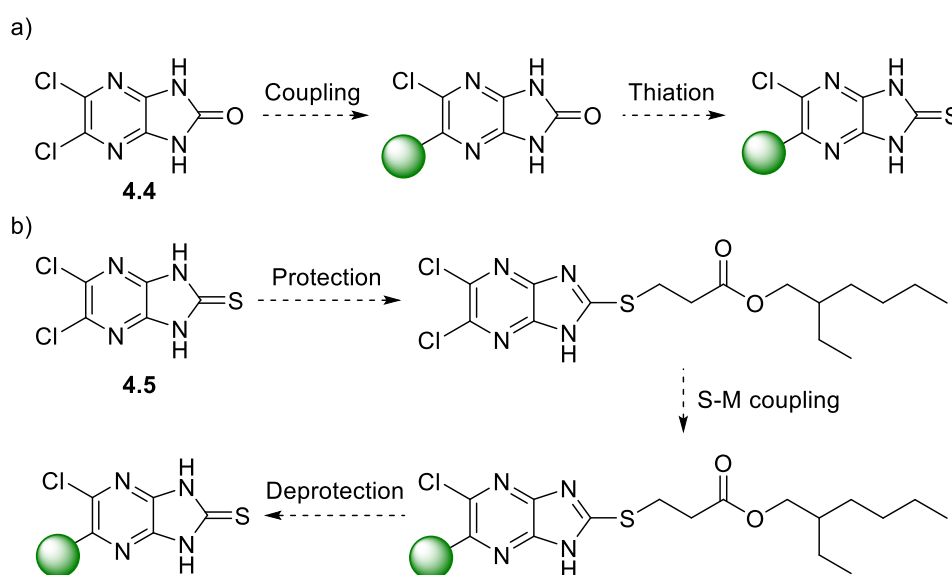
Scheme 4-35: a) Cyclourea cross-coupling followed by thionation strategy, b) Protecting group strategy for cross-coupling.

This meant attention turned to S$_N$Ar chemistry, with very harsh conditions required to enact the desired reactions. Initially, heating **4.18** in neat morpholine at 160 °C for 4 h under argon led to the undesired sulfur substitution (**4.131**) shown in Scheme 4-36.[208] Minor amounts of desired **4.132** were isolated, with the majority starting material left unreacted. Later attempts overnight on larger scales did not improve the yield, and due to purification issues a lower yield (5%) of **4.132** was obtained in quality sufficient for bioassay. Opportunistically, **4.131** could also be utilised in the investigation as sulfur substitutions with larger saturated heterocycles.
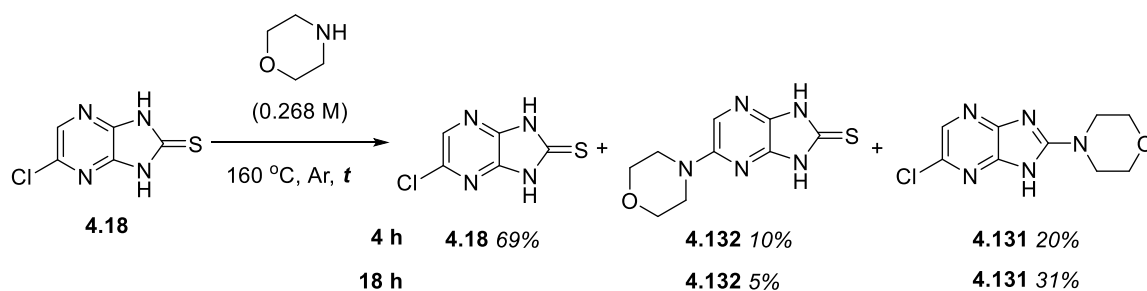
*Scheme 4-36: S$_N$Ar synthesis of morpholino derivatives **4.131** and **4.132**.*

Changing the nucleophile to piperidine provided similar results, allowing access to both **4.133** and **4.134** (Scheme 4-37) but using *N*-methylpiperazine made the isolation of products **4.135** and **4.136** very difficult. By TLC, it seemed that in this case, the vast majority was the product of substitution at the thione and not the desired pyrazine position, in this case isolation was not pursued.
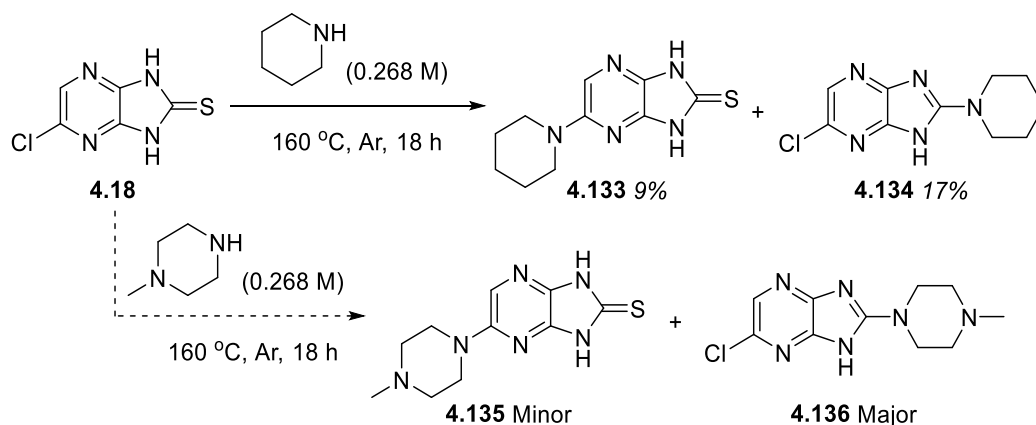
Scheme 4-37: Synthesis and attempted synthesis of piperidine and N-methylpiperazine derivatives
***4.133 – 4.136**.*

4.5.4 <u>Restricting the Side Chain</u>

The last objective for the further scrutiny of activity surrounding the thione functionality was side chain restriction. Installation of alkyl chains was successful and 5, 6 and 7-membered rings were installed as shown in Scheme 4-38 to provide products **4.138** – **4.140**.[209] These three reactions are all favourable intramolecular S$_N$2 exo-tet cyclisations following Baldwin's rule. Additionally, the ease of cyclisation differs due to the size of the ring being formed. The ease of ring cyclisation occurring goes from 5 > 6 > 3 > 7 > 4, and this is reflected in the yields of the reactions diminishing from **4.138** to **4.140**.
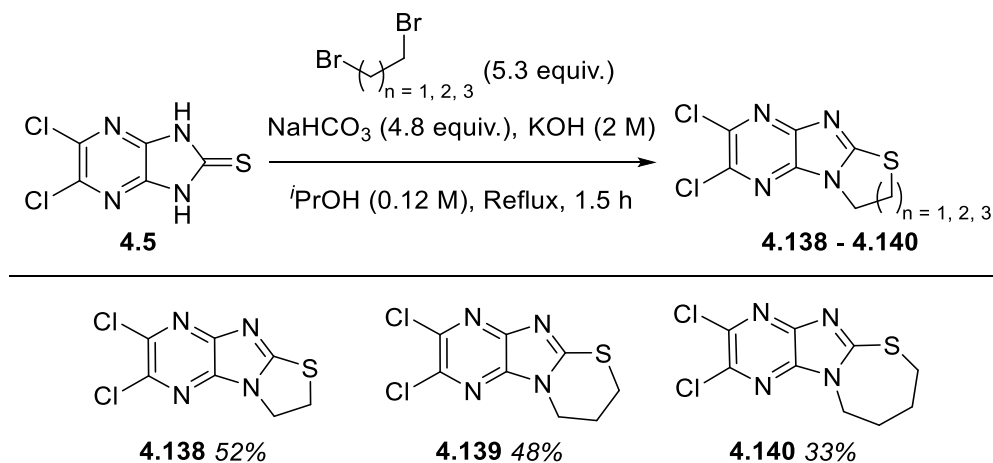
Scheme 4-38: Synthesis of restricted side chain variations **4.138 – 4.140**.

Attempts to install carbonyl functional groups within this side chain were generally unsuccessful (Scheme 4-39). Heating **4.5** with neat chloroacetic acid at 100 °C only provided the *S*-alkylated intermediate **4.43**, with no subsequent cyclised product **4.141** observed. Treating this carboxylic acid with acetic anhydride provided inconclusive results, with attempted isolation proving again difficult. Synthesis of the thioester regioisomer **4.142** was not pursued in the interests of time.
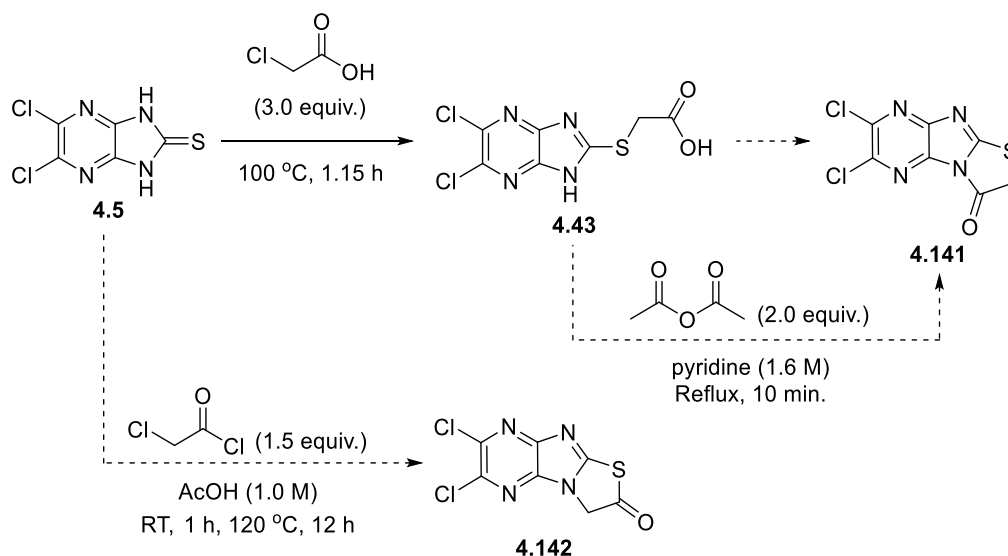
Scheme 4-39: Attempted synthesis of restricted side chains bearing carbonyl groups.

To conclude the synthesis part of the project was partially successful, and the compounds synthesised for this part of the project are summarised in Figure 4-18. Whilst the *S*-methylations were successful, the issues surrounding the synthesis of growth vectors from the pyrazine ring *via* cross-coupling reactions meant several compounds that were targeted could not be produced in the way hoped for, for efficient library production. If time permitted, an investigation into whether the thiol protecting groups would work under

cross-coupling conditions for cyclothioureas, or if the cross-coupling-thionation route would be amenable to synthesis of the desired compounds would have been conducted. If successful, this would represent one or two ways to access a class of compounds which are otherwise not tolerated in Pd cross-coupling reactions directly and to the best of my knowledge has not been described. The difficulties surrounding isolation of the restricted side chain compounds bearing carbonyl functional groups meant these could not be investigated to explore the potential for these compounds to act as inhibitors of WWP2.
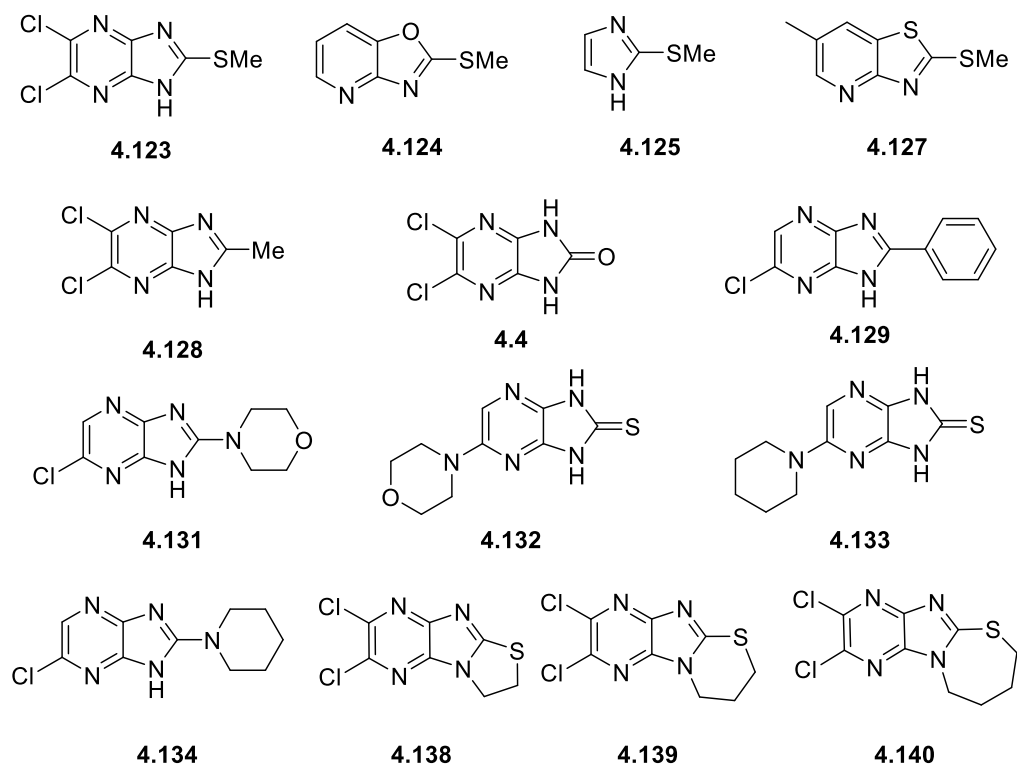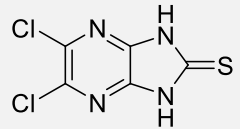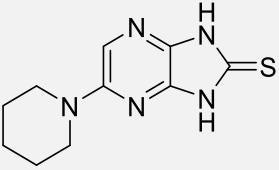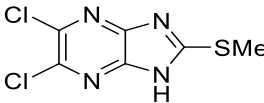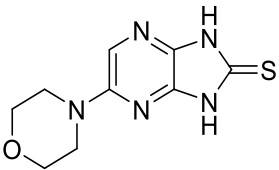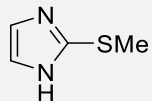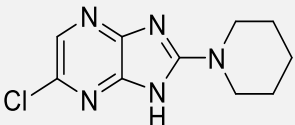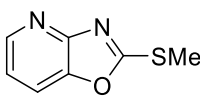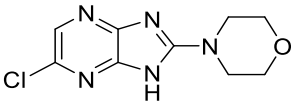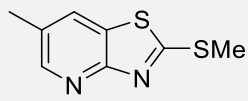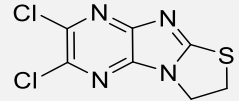
Figure 4-18: Summary of compounds synthesised as part of further investigations into the thione compounds.

4.6 <u>Biological Results of the Thione Investigation</u>

As described in Table 4-8, all methyl groups completely removed activity (entries 3, 5, 7 and 9). Substituting the sulfur for either the methyl or oxygen groups removed activity too (entries 11, 13). Substitutions of the saturated nitrogen heterocycles in place of the sulfur (**4.134** & **4.131**) also displayed no activity (entries 6 and 8). The morpholine and piperidine derivatives with the thione group present (**4.132** and **4.133**) provided much higher activities, even though these are apparently insoluble at 10% DMSO in both 10 mM and 1 mM concentrations (entries 2, 4). All restricted side chain variations were too lipophilic and were not soluble in the bioassay. To summarise these results, the free thione seems to be essential for the activity found in the bioassay for these compounds.

Entry	Structure	IC$_{50}$ (µM)	Entry	Structure	IC$_{50}$ (µM)
1 (4.5)		14.6	2 (4.133)		<1**
3 (4.123)		N/A*	4 (4.132)		<1**
5 (4.125)		N/A	6 (4.134)		N/A*
7 (4.124)		N/A	8 (4.131)		N/A
9 (4.127)		N/A	10 (4.138)		N/A**

171
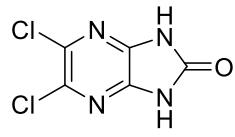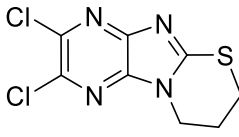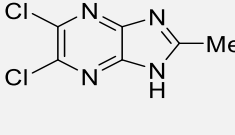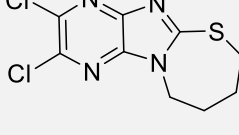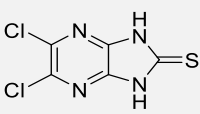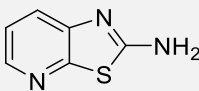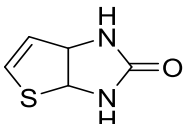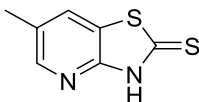
11 **(4.4)**		N/A	**12** **(4.139)**		N/A**
13 **(4.128)**		N/A	**14** **(4.140)**		N/A**

*Table 4-8: Biological results for compound submitted as part of the further investigations. * Not soluble in 10% DMSO at 10 mM. ** not soluble in 10% DMSO at 1 mM.*

Due to the fact that even compounds with minimal functionality are able to provide high IC_{50} values is a cause to be concerned about (e.g. commercial heterocycle precursor to analogue **4.125**, Table 4-7, entry 6), especially when **4.125** (Table 4-8, entry 5) provides no IC_{50} activity. The fact that there is not a readily observable SAR between these active compounds suggests these compounds were exhibiting some form of nuisance behaviour. Another point is the discrepancy between the NCI hit IC_{50} value and the resynthesised compounds result of no activity. Reactivity against E1 and / or E2 in the bioassay can result in high IC_{50} values that would be false positives and therefore misleading against E3. E1 cross-reactivity tests were not carried out until later in the project for these compounds, and these are presented below in Table 4-9, showing all the compounds that showed high IC_{50} values in the bioassay are in fact reactive with E1. These results show that these compounds are active at some point against the UPS system under study and may not be solely active against the E1 enzyme.

Entry	Structure	Cross react E1?	Entry	Structure	Cross react E1?
1		Yes	2		Yes
3		Yes	4		Yes
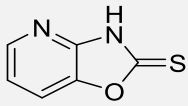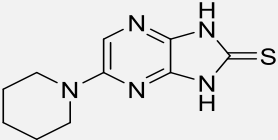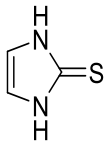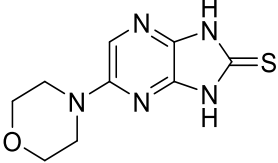
5		Yes	6		Yes
7		Yes	8		Yes

Table 4-9: E1 cross reactivity of active thione compounds.

4.7 <u>NMR & HRMS Purity Analysis of the NCI Hit Sample</u>

4.7.1 <u>NMR Analysis</u>

As seen in Figure 4-19, the original sample (known as NSC-217913 in this section) for initial screenings (top) and synthesised sample of NSC-217913 (known as **4.1** in this section, bottom) are practically identical. All the required peaks except the *NH* imidazole signal are present and are in good agreement with one another. The NSC-217913 sample does have a solvent impurity of DMSO (2.54 ppm), but this is most-likely from incomplete solvent removal when the sample was handled for NMR analysis.
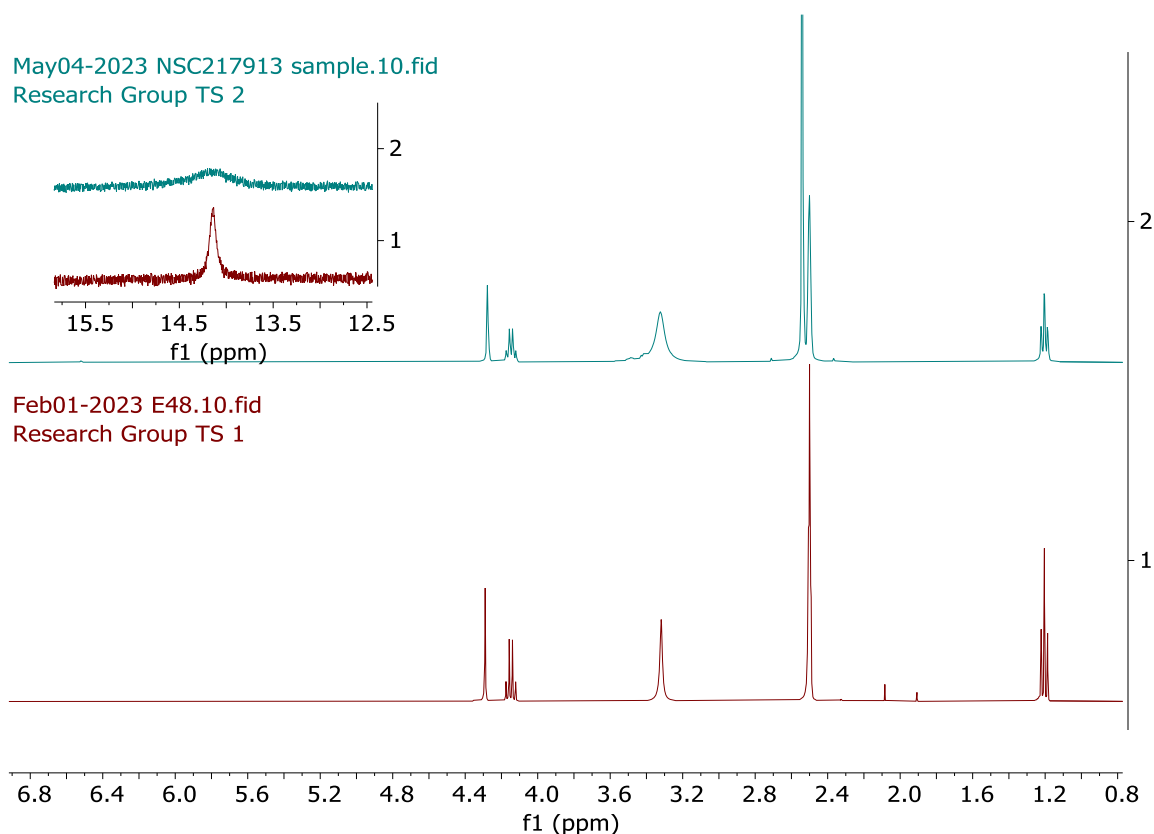
*Figure 4-19: ^{1}H NMR (DMSO-d$_6$) comparison between NSC-217913 vs. **4.1**.*

Compound **4.1** does have two minor known impurities of acetone (2.08 ppm) and acetic acid (1.91 ppm) as shown in Figure 4-19. Compared to that, NSC-217913 does have minor impurities present, which have been uncharacterizable: δ (ppm) 7.70, 7.22, 6.52, 5.16, 4.50, 3.49, 3.42, 1.98, 1.88 and 1.70 (Figure 4-20). It is difficult to suggest what these impurities could be from the ^{1}H NMR spectrum, and in what quantities they are present.
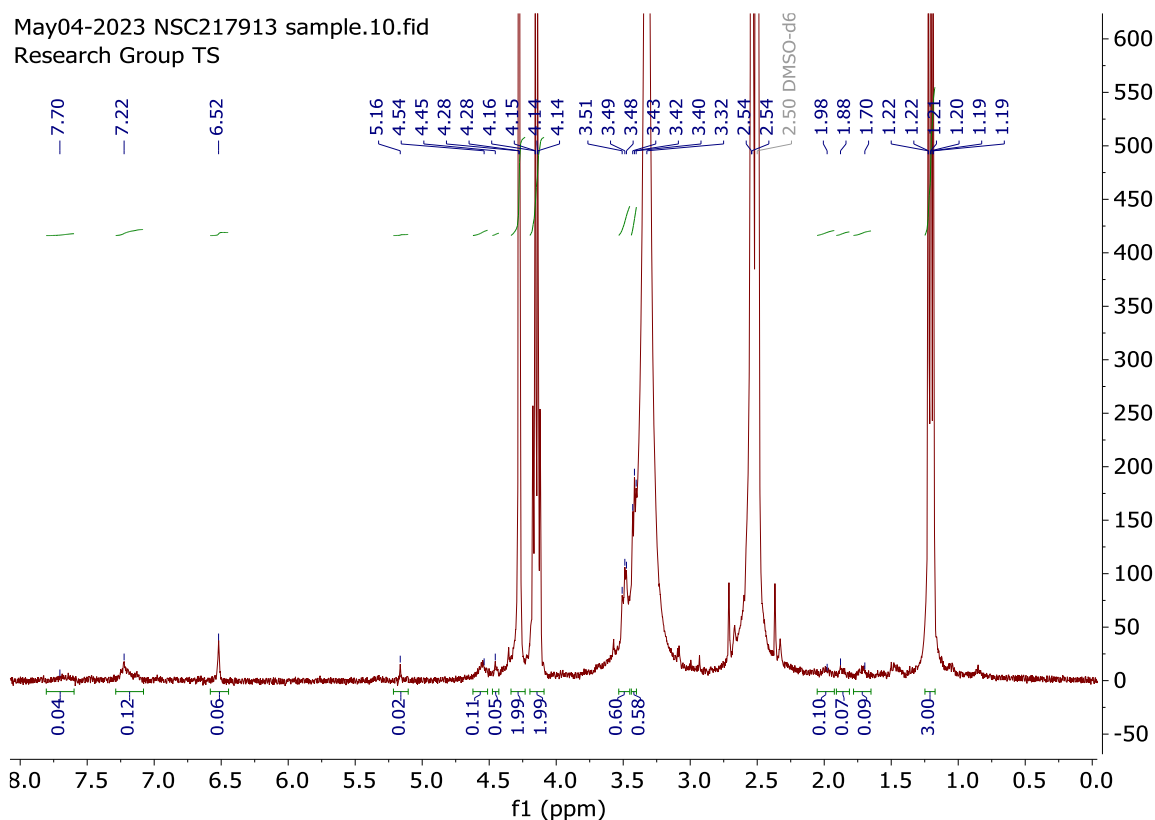
Figure 4-20: ^1H NMR (DMSO-d_6) spectrum (0.0 – 8.0 ppm) of NSC-217913.

4.7.2 <u>**HRMS Analysis**</u>

Because HRMS is so sensitive, a lot of impurities can be identified that are not part of the original NCI sample or resynthesised sample, but come from for example, plasticisers leached from bottle lids and syringes into the solvent used for HRMS studies during sample preparation. It was thought that a blank run would help sort some of these impurities out and help identify actual impurities present in the NCI sample.

Unfortunately, after subjecting the samples to UPLC-HRMS, using the subtraction tool on the Masslynx software provided little gain in removing the materials present in the blank from the materials actually part of the NCI sample. This meant a more manual comparison was required. Below are the two chromatograms obtained from the blank MeCN and the NCI sample of NSC-217913 (Figure 4-21).
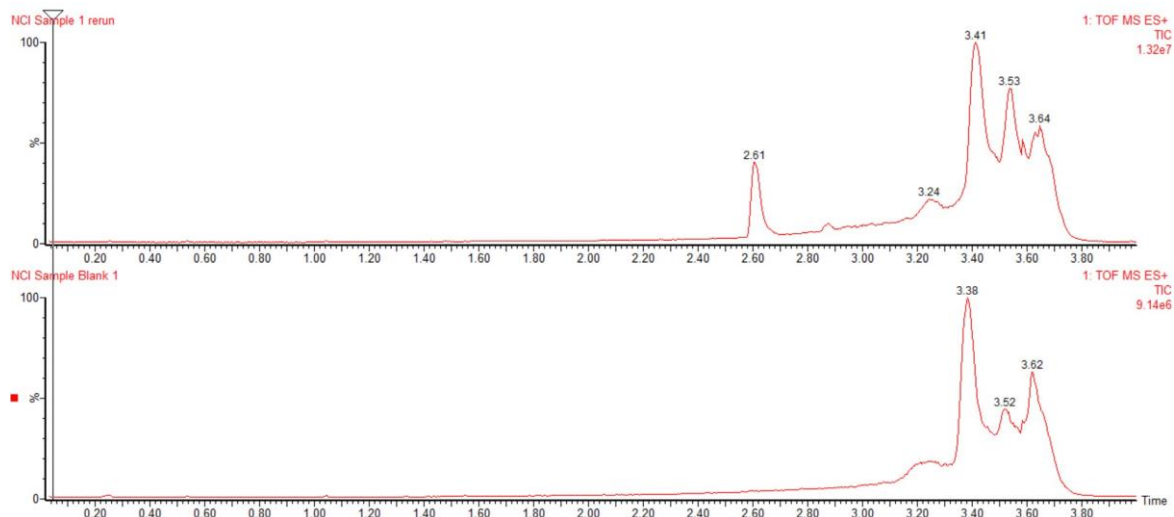
Figure 4-21: Chromatograms of NSC-217913 (top) and blank sample (bottom) in MeCN.

Firstly, it is obvious that there is a large number of impurities present at the end of the chromatogram in both samples (approx. 3.24 min. to 3.80 min.), the most striking differences between the chromatograms is the peak observed at 2.61 min. followed by a peak at approx. 2.88 min. (not labelled). Additionally, there are several more peaks present at the end of the chromatogram compared to the blank sample (in between 3.53 min. and 3.64 min., top).

The peak at 2.61 corresponds to the (M+H) molecular ion of NSC-217913 as shown in the mass spectrum below (Figure 4-22). Signals at 306.9816, 307.9844, 308.9789, 310.9759 correspond to the isotopes and their relative abundances within the molecule, the pattern is unique because of the two chlorine atoms. Presented also is the (M+Na) *m/z* signals at 328.9638 *etc.*
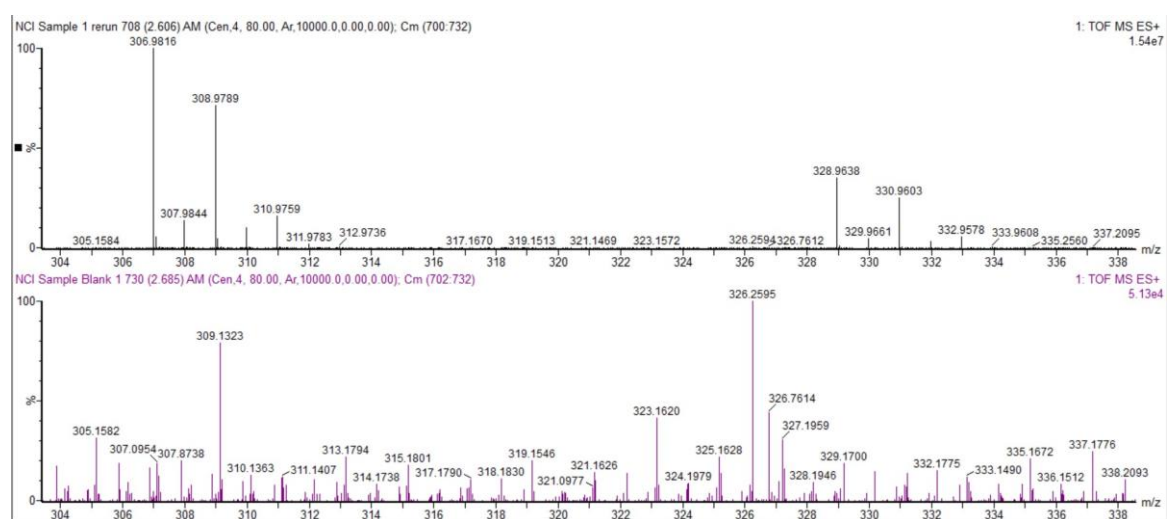
Figure 4-22: Mass spectrum at 2.606 (top, NCI sample) and 2.685 (bottom, blank).

Next the peak at 2.88 was investigated and it was observed (Figure 4-23) that an impurity is present at 393.0209 and 415.0035, which is not present in the blank sample. The repeating *m/z* signals below these are probably some form of polymeric impurity absent in the original sample.
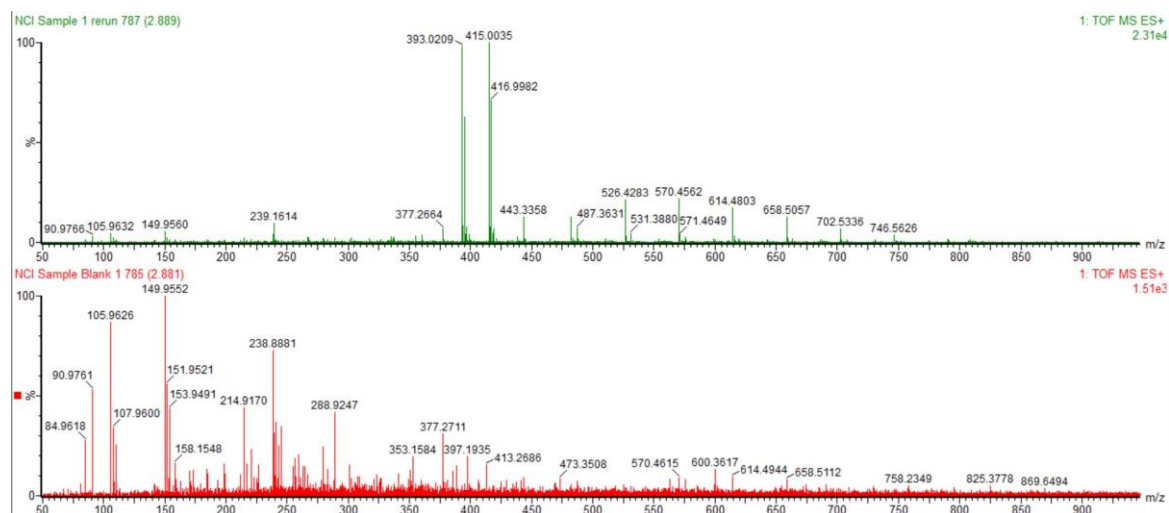
Figure 4-23: Identified difference from chromatogram peak at 2.88 min.

A closer look shows that these signals have the same isotope pattern as that observed for NSC-217913 (Figure 4-24). Additionally, the *m/z* signal at 415.0002 has a difference from the 393.0179 signal of 21.9823, suggesting that it was the M+Na ion signal.
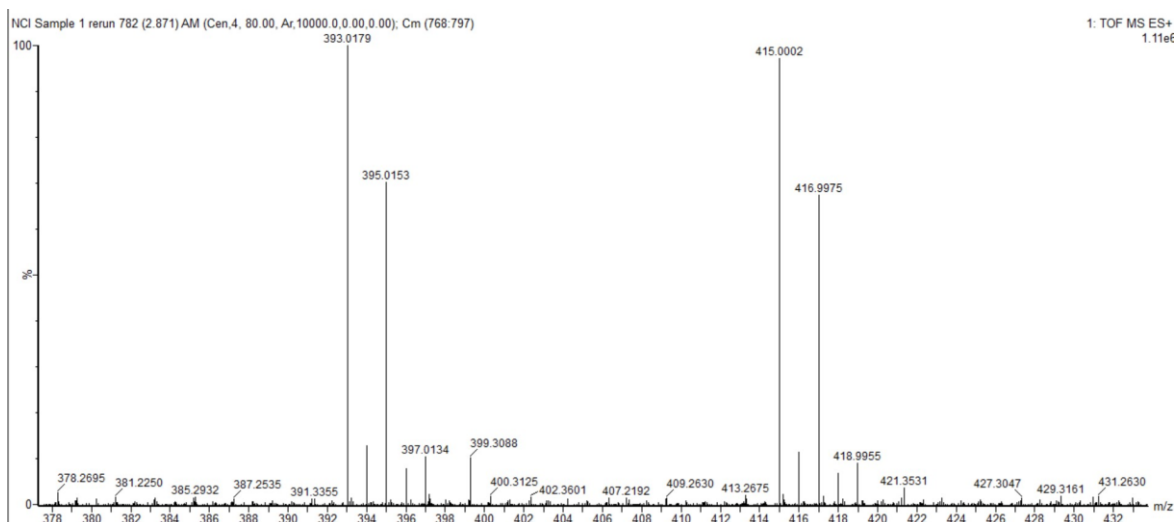
Figure 4-24: Zoomed in view of identified impurities at 2.88 min.

Thankfully, using the predictive isotope model on the Masslynx software used for HRMS data processing, screening of potential compounds that match this *m/z* could be undertaken. Firstly, it was thought that an impurity with two chlorine atoms present within this NCI sample would also have the same heterocyclic core, as this is the only known route to access these compounds. Looking further into the original preparation (it is

unknown if the NCI sample was prepared from this procedure, but lack of alternative literature precedent points towards this) of NSC-217913, which was by reflux of **4.5** in EtOH with ethyl chloroacetate, an additional *N*-alkylation may have occurred after *S*-alkylation, since this is known to occur slower than *S*-alkylation and more readily at higher temperatures.[192] Using the isotope prediction of a doubly *S*- and *N*-alkylation product, the *m/z* spectrum matches well with the observed *m/z* signal (Figure 4-25). Therefore, it is thought the likely impurity in with NCI sample is either **4.143** or **4.144** shown in Figure 4-26. It is known that the *N*-alkylation is unselective due to the highly-delocalised nature of the electrons within these heteroaromatic rings. In a study investigating the *N*-alkylation of imidazo[4,5-*b*]pyrazine ring systems, it was found that in DMF with K_2CO_3 and a benzyl bromide alkylation partner, *N*-alkylation of imidazo[4,5-*b*]pyrazine system occurred mostly at the *N-4* position over the *N-1* position (46% *vs.* 8% respectively).[172] Whilst these conditions are not in any way similar to the *S*-alkylation conditions utilised in the literature synthesis, it could provide some information as to the regioselectivity of this additional ethyl acetate moiety. The other signals investigated were also present in the blank sample and were not thought to be contaminants of the NCI sample.
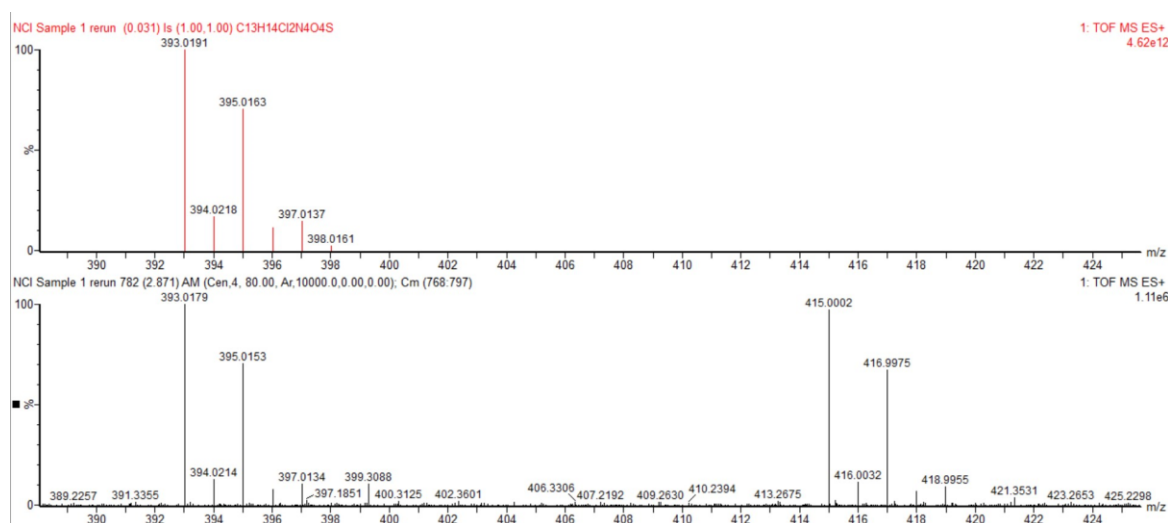
Figure 4-25: Matched predicted doubly-alkylated material with (M+H) ion observed.

4.143   **4.144**
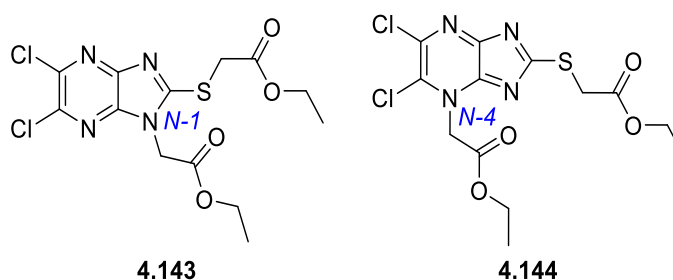
*Figure 4-26: postulated impurities **4.143** and **4.144** in NCI sample based off HRMS analysis.*

4.8 Conclusions

To conclude this chapter, firstly two routes have been devised which allow the preparation of the 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione heterocyclic core. Either through a thionation route starting from **4.4**, or from direct cyclothiourea ring closure from **4.2** with TCDI the thione may be prepared (Scheme 4-40). Both of these routes required the starting 2,3-diamino-5,6-dichlorpyrazine system, which through a novel route removed problems surrounding by-product formation and improved yields over the historical route. The improved route consists of performing a dichlorination on **4.13**, followed by a regioselective amination to form **4.15** exclusively. This route was made successful by discovering the rapid chlorination of **4.15** with CBO in DMF. Whilst this route provided a 63% yield which was amenable for this work, a decreased yield was observed on a 10 g scale and more broadly offers an opportunity to improve yields by further optimisation.
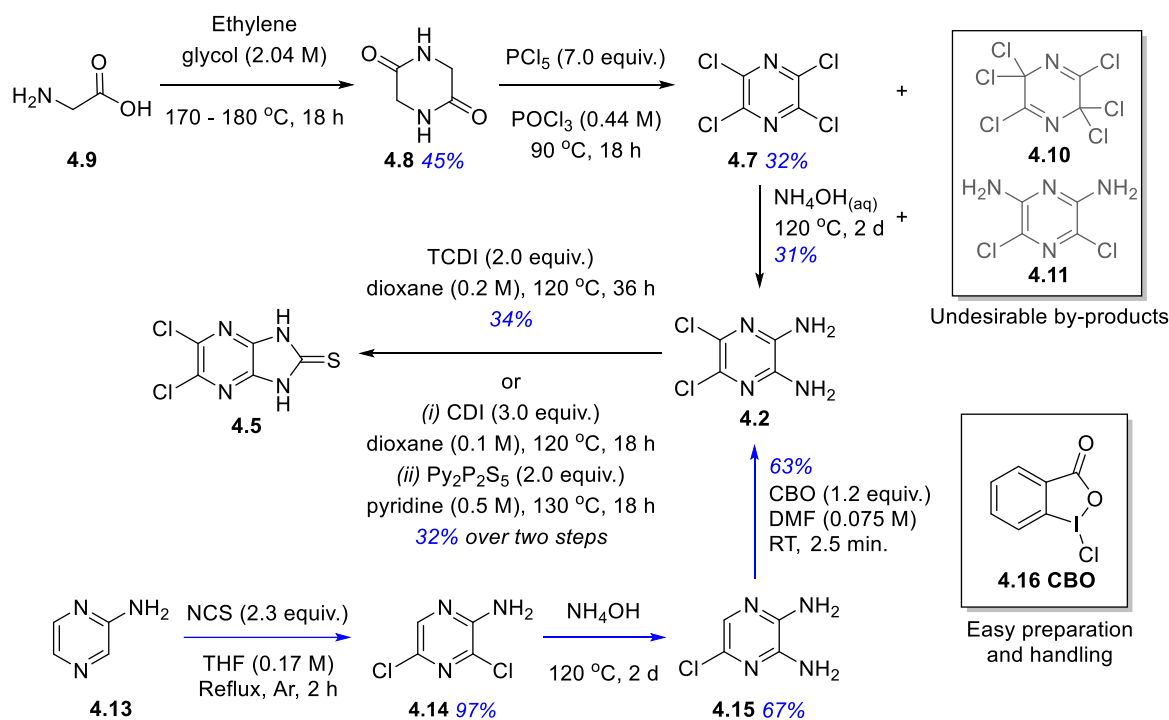
*Scheme 4-40: Summary of routes to 5,6-dichloroimidazo[4,5-b]pyrazine-2(1,3H)-thione **4.5**.*

Secondly, a range of *S*-alkylated products were synthesised and allowed for an efficient production of a library, which was partly chosen based on molecular docking studies on these compounds against the proposed binding site of WWP2. Furthermore, several heterocyclic variations were produced, which along with the commercial heterocycles purchased represents 24 different heterocyclic variations examined in this series. From these series very little biological data was obtained, and compound **4.5** represented the only molecule of the imidazo[4,5-*b*]pyrazine class which provided better activity, although

other heterocycles bearing the cyclothiourea were also more active than NSC-217913 too. Both the first and second objectives of this series were successfully met.

Further investigation of the cyclothiourea motif found that only the free thione group was able to provide activity in the bioassay, and these compounds were in fact active against the E1 enzyme. The strong IC_{50} data obtained stems from the UPS being inhibited prior to E3 autoubiquitination at one of the E1-involved steps, essentially providing a false positive data against WWP2 (see section 2.1.3 for assay explanation). The DSF screen showed these compounds did not alter the stability of WWP2 itself, but other compounds were identified that were interacting with WWP2 to either stabilise or destabilise the enzyme in some way.

HRMS analysis of the NCI sample predicted that a di-alkylated material was present, either **4.143** or **4.144**. It is unknown which regioisomer it was, or if it was a mixture of both. Some form of impurity was described *via* NMR analysis, but this was not characterizable. It is currently unclear whether the NMR data corroborates the HRMS data, but the number and chemical shifts do not match what would be predicted for these identified impurities. This of course represents a problem since it is unknown what compounds present and are providing the activity, and therefore worth pursuing into analogue synthesis. Future work confirming if these di-alkylated materials are inhibitory against WWP2 would be prudent.

Either way, the IC_{50} data obtained from the NCI sample could not be reproduced by **4.1**, and two hypotheses can be proposed. Firstly, some form of impurity as identified by HRMS could be responsible for the activity observed. The resynthesis and biological testing of the compounds identified would confirm if this was the case. The second idea is that the NCI sample had partially hydrolysed at the acetate position upon storage, providing the free thione **4.5** in very small quantities which then provided the observed IC_{50} assay result for the NCI sample. An issue with this is idea is that no *m/z* signal was observed in the mass spectrum for **4.5**, but seeing as there were a lot of impurities present arising from sample preparation it would be difficult to fully elucidate all impurities within the NCI sample *via* HRMS.

Finally, whilst the two main objectives were successfully met by synthesising NSC-217913 and analogues thereof, the aim of finding a more potent molecule was not met as only compounds were found to be interacting with WWP2 *via* DSF experiments or were providing worse IC_{50} values. It seems as though, because of the issues surrounding the purity of NSC-217913 and the difficulty in reproducing the NCI samples IC_{50} data, that this hit probably required further validation it against WWP2 prior to the analogue synthesis.

4.9 Experimental

2,6-diketopiperazine[179]
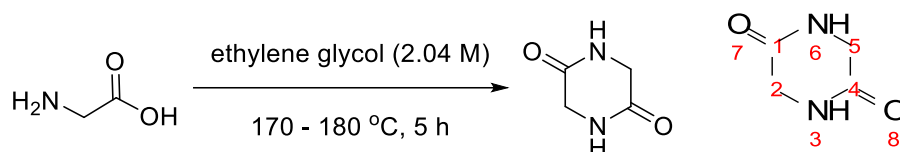
To a 500 mL RBF was added a stirrer bar, glycine (20 g, 0.266 mol) and ethylene glycol (130 mL, 2.04 M). The mixture was heated to 170-180 °C for 5 h, left to cool to RT and placed in a freezer overnight. The mixture was filtered and washed with methanol to yield a fine brown solid. This was recrystallised from water to give 2,6-diketopiperazine as a brown crystalline solid (6.88 g, 60.3 mmol, 45%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.00 (br s, 2H, **3**, **6**), 3.70 (d, $J = 2.0$ Hz, 4H, **2**, **5**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 166.1, **1**, **4**, 44.3, **2**, **5**.

IR (cm^{-1}) 3161 (NH), 1664 (C=O).

Data is in accordance with literature data.

Tetrachloropyrazine[180]
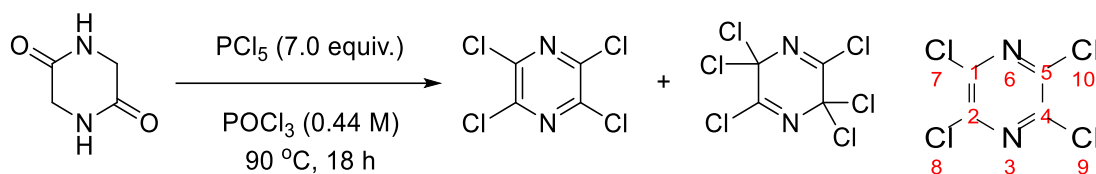
To a 250 mL RBF was added 2,6-diketopiperazine (5.00 g, 0.044 mol), phosphorus pentachloride (63 g, 7 equiv.), phosphorus oxychloride (100 mL, 0.44 M) and a stirrer bar, this was placed under nitrogen atmosphere. The vessel was heated at 90 °C for 18 h. The unreacted phosphorus oxychloride and phosphorus pentachloride were distilled off at 200 °C (increasing in 10 °C increments). The distillation continued until no more liquid distillate came off (approx. 4 h total, 3 h at 200 °C). The vessel was allowed to cool to RT and the solid reaction mixture was dissolved in hexane (50 mL) and aqueous HCl (1 M, 50 mL) was added to the reaction vessel, the mixture was transferred to a separatory funnel and the aqueous layer was extracted with hexane (3x 30 mL). The organic layers were collected and filtered through a pad of Florisil® (100-200 mesh), the filtrate was washed with aqueous NaOH (1 M, 20 mL) and brine (20 mL). The hexane was then removed *in vacuo.* The residue was purified *via* column chromatography (*n*-hexane) to yield tetrachloropyrazine as a white crystalline solid (3.0 g, 0.014 mol, 32%).

^{13}C NMR (101 MHz, DMSO-d_6) $δ$ 143.8, *1*, *2*, *4*, *5*.

IR (cm^{-1}) 1736.

M.P. approx. 100 °C (lit.), 99.6 – 100.3 °C.

Data is in accordance with literature data.[180]

2,3,3,5,6,6-Hexachloro-3,6-dihydropyrazine: Isolated as an off-white solid.

IR (cm^{-1}) 3201, 2834, 1725, 1645.

M.P. approx. 100°C (lit.) 102 – 103 °C.

Data in line with literature data.

2,3-diamino-5,6-dichloropyrazine
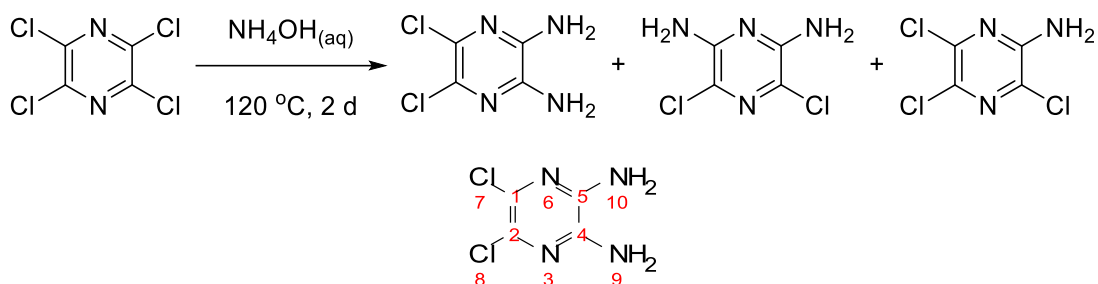
To five 30 mL microwave vials was added in equal portions tetrachloropyrazine (2.5 g, 11.50 mmol, 0.5 g/vial, 2.3 mmol/vial), a stirrer bar and ammonium hydroxide solution (25%, 12 mL/vial). The vials were sealed with a Teflon cap and stirred at 120 °C for two days in a sand bath. The vials were allowed to cool to RT and diluted with EtOAc, transferred to a separatory funnel, and neutralised with acetic acid. The organic layer was separated, and the aqueous layer extracted with EtOAc (3x 20 mL). The organic layers were collected and treated with NaHCO$_3$ (sat. soln., 4x 30 mL) until gas evolution was no longer observed. Organic layer was dried (Na$_2$SO$_4$), and solvent removed *in vacuo.* The solid recovered was subjected to column chromatography (8:2 -> 7:3 Hexane : EtOAc) and yielded the desired 2,3-diamino-5,6-dichloropyrazine as an off-white solid (641 mg, 3.58 mmol, 31%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 6.55 (br s, 4H, **9**, **10**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 142.6, **4**, **5**, 126.2, **1**, **2**.

IR (cm^{-1}) 3435 (NH), 3294 (NH).

M.P. 275.7 – 276.1 °C.

MS ES+ m/z calcd for C$_4$H$_4$^{35}Cl^{37}ClN$_4$ (M+H)$^+$: 179.9909, found: 179.9910.

Isolated **2,6-diamino-3,5-dichloropyrazine** as an off-white solid:

^{1}H NMR (400 MHz, DMSO-d_6) δ 6.33 (br s, 4H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 150.1, 113.9.

IR (cm^{-1}) 3490 (NH), 3311 (NH).

M.P. 141.8 – 142.7 °C.

MS ES+ m/z calcd for C$_4$H$_4$^{37}Cl$_2$N$_4$ (M+H)$^+$: 182.9833, found: 182.9835.

Also isolated **2-amino-3,5,6-trichloropyrazine** as an off-white solid:

^{1}H NMR (400 MHz, DMSO-d_6) δ 7.46 (br s, 2H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 151.3, 141.7, 128.5, 126.6.

IR (cm[-1]) 3476 (NH), 3288 (NH).

M.P. 230.0-231.1 °C (dec.).

Adapted from literature procedure.[167]

2-amino-3,5-dichloropyrazine
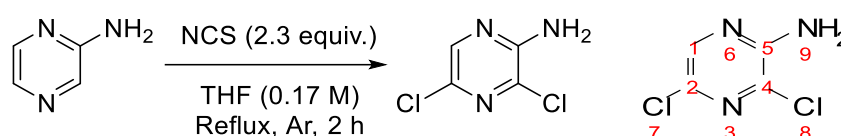
To a 1 L 3-neck RBF attached with two glass stoppers and a reflux condenser, was added 2-aminopyrazine (8.08 g, 85.0 mmol), a stirrer bar and placed under an argon atmosphere with low stirring. THF (500 mL, over mol. sieves, 0.17 M) was then added. *N*-chlorosuccinimide (26.11 g, 195.4 mmol, 2.30 equiv.) was added in one batch to the solution, the mixture warmed to 80 °C and monitored by TLC (7:3 Hex. : EtOAc). After 2 h TLC indicated full consumption of starting material, the vessel was cooled to RT. The reaction mixture was filtered through a pad of celite® and washed with excess EtOAc. The solvent was reduced *in vacuo.* and the residue dry loaded and purified by column chromatography (8:2 -> 7:3 Hexane : EtOAc) to provide 2-amino-3,5-dichloropyrazine (13.48 g, 82.2 mmol, 97%) as yellow needle-like crystals.

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 7.97 (s, 1H, *1*), 5.01 (br s, 2H, *9*).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 150.5, *5*, 140.2, *1*, 134.8, *2*, 131.6, *4*.

IR (cm[-1]) 3435 (NH), 3167 (NH).

M.P. 115 – 118 °C (lit.), 126.9 – 127.3 °C.

Adapted from literature procedure.[182]

2,3-diamino-5-choropyrazine
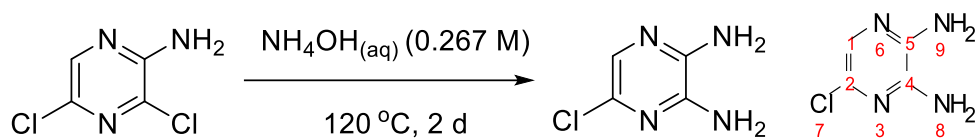
To eighteen 30 mL microwave vials was added 2-amino-3,5-dichloropyrazine (total of 9.00 g (54.9 mmol), approx. 0.5 g per vial) and a stirrer bar. To each vial was added ammonium hydroxide solution (12.5 mL, 0.276 M) and sealed. Vials were placed in a sand bath and heated with stirring to 120 °C (sand temp.) for 2 days. The vessels were allowed to cool to RT, contents transferred to a separatory funnel and extracted with EtOAc (8x 100 mL). The organic extracts were collected and washed with brine (100 mL) and dried (MgSO$_4$). The solvent was removed *in vacuo.*, the material dry loaded (acetone) and purified by column chromatography (5:5 -> 2:8 -> 0:1 Hex. : EtOAc) to provide 2,3-diamino-5-chloropyrazine (5.06 g, 35.0 mmol, 64%) as a beige solid. Starting 2-amino-3,5-dichloropyrazine was also recovered (664 mg, 4.05 mmol, 7%) as yellow needle-like crystals.

^{1}H NMR (400 MHz, Acetone-d_6) *δ* 7.24 (s, 1H, **1**), 5.72 (br s, 2H, **8**), 5.47 (br s, 2H, **9**).

^{13}C NMR (101 MHz, Acetone-d_6) *δ* 144.5, **5**, 143.5, **4**, 134.3, **2**, 127.7, **1**.

IR (cm^{-1}) 3440 (NH), 3375 (NH), 1646 (NH bend).

M.P. 222.7 – 223.1 °C (deg.).

This reaction was repeated on a 23.5 g (0.143 mol) scale. Eighteen 30 mL microwave vials were charged with 1.25 g of 2-amino-3,5-dichloropyrazine each and NH$_4$OH (25%, 12.5 mL) for two days at 120 °C in sealed vials. Provided 15.72 g of 2,3-diamino-5-chloropyrazine (67%) after purification in the same manner as above. Adapted from literature procedure.[167]

1-chloro-1,2-benziodoxol-3(1*H*)-one[210]
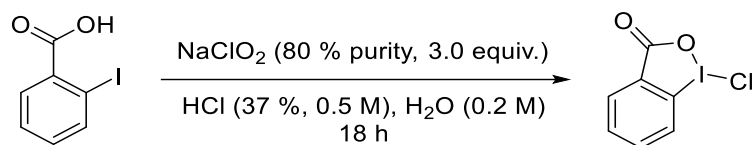
To a 250 mL RBF was added 2-iodobenzoic acid (2.0 g, 8.06 mmol), sodium chlorite (2.19 g, 24.2 mmol, 3.0 equiv.), a stirrer bar and H_2O (40 mL, 0.2 M). The suspension was stirred and to an attached dropping funnel was added concentrated hydrochloric acid (37%, 16.2 mL, 0.5 M). This was added dropwise at RT with vigorous stirring, which was continued for 18 h after addition. The precipitated solid was filtered and washed with H_2O and Pet E., the solid was dried under vacuum to provide 1-chloro-1,2-benziodoxol-3(1*H*)-one as a light-yellow powder (2.18 g, 7.72 mmol, 96%).

^{1}H NMR (500 MHz, C*D*Cl$_3$) *δ* 8.30 – 8.24 (m, 1H), 8.24 – 8.18 (m, 1H), 8.00 (tdd, *J* = 6.9, 3.2, 1.6 Hz, 1H), 7.80 (tt, *J* = 6.0, 2.9 Hz, 1H).

^{13}C NMR (126 MHz, C*D*Cl$_3$) *δ* 167.3, 136.8, 133.6, 132.0, 128.8, 127.0, 117.2.

IR (cm^{-1}): 1724 (C=O).

M.P. 172 – 173 °C (lit.), [210] 173.4 – 175.0 °C.

Data matches literature.

2,3-diamino-5,6-dichloropyrazine
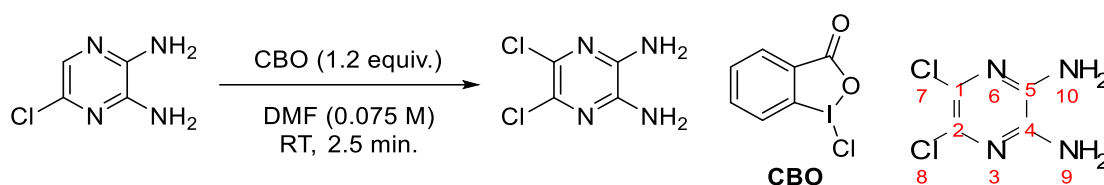
To an 8 mL microwave vial was added 2,3-diamino-5-chloropyrazine (50 mg, 0.345 mmol) and a stirrer bar. Anhydrous DMF (4.6 mL, 0.075 M) was added *via* syringe. 1-Chloro-1,2-benziodoxol-3(1*H*)-one (CBO, 0.117 g, 1.2 equiv.) was added in one batch with vigorous stirring. Timing was started with this addition for 2.5 minutes, with a significant colour change from light yellow to black occurring. The reaction mixture was quenched immediately following the 2.5-minute mark with NaHCO$_3$ sat. solution and the reaction mixture transferred to a separatory funnel, washing with water. This was extracted with EtOAc (3x 15 mL), the organic layers were collected and washed with water (2x 15 mL), dried (Na$_2$SO$_4$) and the solvent removed *in vacuo.* The residue was purified using column chromatography (6:4 Pet. E : EtOAc) to provide 2,3-diamino-5,6-dichloropyrazine as a beige solid (38.9 mg, 0.21 mmol, 63%).

^{1}H NMR (400 MHz, DMSO-d_6) *δ* 6.55 (br s, 4H, **9**, **10**.).

^{13}C NMR (101 MHz, DMSO-d_6) *δ* 142.6, **4**, **5**, 126.2, **1**, **2**.

IR (cm^{-1}) 3435 (NH), 3202 (NH), 1625 (NH bend).

M.P. 275.7 – 276.1 ^{o}C.

MS ES+ m/z calcd for C$_4$H$_4$^{35}Cl^{37}ClN$_4$ (M+H)$^+$: 179.9909, found: 179.9910.

May be repeated on a 1 g scale with no diminishment in yield (63%) and a 10 g scale with a diminished yield of 56%. Adapted from literature conditions.[184]

5-chloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-one
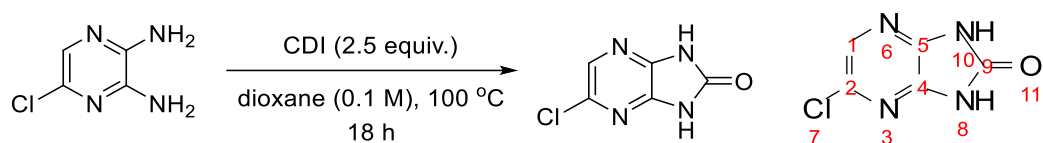
To an 8 mL microwave vial was added 2,3-diamino-5-chloropyrazine (50 mg, 0.346 mmol) and dioxane (0.346 mL, 0.1 M). This was stirred at 80 °C and carbonyldiimidazole (CDI, 84.3 mg, 1.5 equiv.). This was stirred at 100 °C for 6 h and an additional equivalent of CDI (56 mg) was added. Allowed to stir at temperature for 18 h. The reaction was allowed to cool to RT and the solvent removed *in vacuo.* The residue was taken up in EtOAc and washed with HCl (1 M, 2x 10 mL), sat. NaHCO$_3$ (1x 10 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* to provide 5-chloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-one as a light orange solid (39.6 mg, 0.23 mmol, 67%).

^{1}H NMR (400 MHz, DMSO-*d*$_6$) *δ* 11.96 (br s, 2H, **8**, **10**), 7.91 (s, 1H, **1**).

^{13}C NMR (101 MHz, DMSO-*d*$_6$) *δ* 154.1, 139.8, 139.4, 138.1, 132.0.

IR (cm^{-1}) 3198 (NH), 1737 (C=O), 1610 (NH bend).

M.P. 279.8 – 280.2 °C.

MS ESI+ calcd for C$_5$H$_3$^{37}ClN$_4$O (M+H)$^+$: 172.0094 found: 172.0095.

Was repeated on a 1 g scale, provided 765.4 mg (5.44 mmol) 79% yield.

5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-one
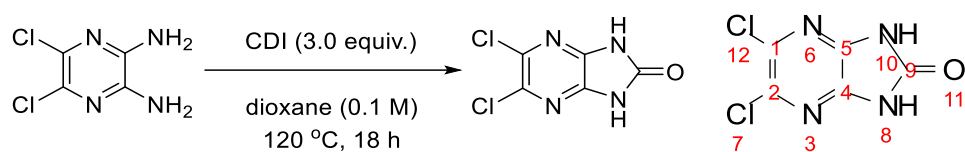
To a 250 mL RBF was added 2,3-diamino-5,6-dichloropyrazine (1.46 g, 8.14 mmol) and a stirrer bar. Dioxane (82 mL, 0.1 M) was added *via* syringe and the mixture was heated to 80 °C to dissolve. To the stirred solution was added carbonyldiimidazole (CDI, 1.98 g, 1.50 equiv.) and the temperature raised to 120 °C at reflux with vigorous stirring. An additional portion of CDI (1.5 equiv.) was added at the 5 h mark. After full consumption of starting material (5:5 Pet. E : EtOAc TLC) dioxane was removed *in vacuo.* and the residue taken up in EtOAc (100 mL), washed with HCl (2.0 M, 20 mL). The aqueous layer was extracted with EtOAc (3x 20 mL). The organic layers were combined and washed with HCl (2 M, 2x 10 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* Provided a brown solid of 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-one (1.12 g, 5.47 mmol, 67%).

^{1}H NMR (400 MHz, DMSO-d_6) *δ* 12.15 (br s, 2H, *8*, *10*).

^{13}C NMR (101 MHz, DMSO-d_6) *δ* 154.3, *9*, 139.1, *4*, *5*, 134.2, *1*, *2*.

M.P. 349.0 - 351 °C (deg.).

IR (cm^{-1}) 3125 (NH), 1698 (C=O), 1596 (NH bend).

MS ES+ m/z calcd for C$_5$H$_2$^{35}Cl$_2$N$_4$O (M+H)$^+$: 204.9684, found: 204.9685.

Melting point in-line with literature data.[167]

Lawesson's reagent[186]
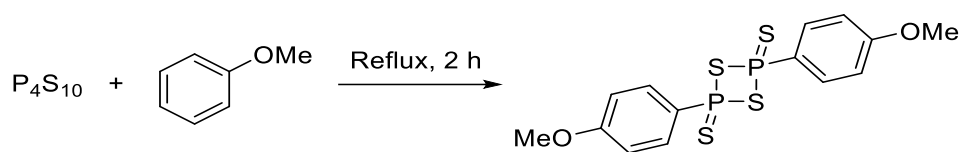
To a 25 mL RBF was added a stirrer bar, phosphorus pentasulfide (5.0 g, 11.25 mmol) and anisole (12.22 mL, 110.25 mmol). This was attached to a reflux condenser which was attached to a bubbler through cold bleach (for H_2S evolution). The mixture was stirred vigorously and heated to 160 °C for 2 h, afterwards it was allowed to cool to RT, the solid mass was broken up and filtered, washing with dry diethyl ether. Provided a malodourous yellow crystalline solid (7.27 g, 18.0 mmol, 80%). May be recrystallised from toluene to provide light-yellow crystals.

M.P. 228 °C (lit.),[211] 215 °C.

Pyridin-1-ium-1-yl[pyridin-1-ium-1-yl(sulfido)phosphinothioyl]sulfanyl-sulfido-thioxo-phosphane[185]
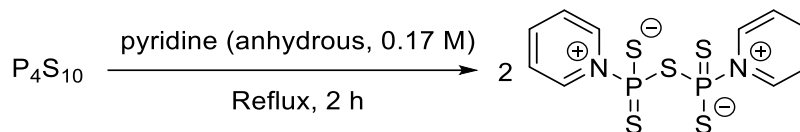
Anhydrous pyridine was freshly prepared by fractional distillation under argon. To a heated (80 °C) stirred solution of anhydrous pyridine (200 mL, 0.17 M), phosphorus pentasulfide (15 g, 33.74 mmol) was added portion-wise and after addition the mixture was heated to reflux for 2 h under argon. The mixture was allowed to cool for 1 h and the formed crystals filtered off. They were washed with acetonitrile (over mol. sieves), transferred to a beaker and placed in a desiccator alongside a beaker of concentrated H_2SO_4 under vacuum to remove any traces of pyridine. Provided light-yellow crystalline solid of pyridin-1-ium-1-yl[pyridin-1-ium-1-yl(sulfido)phosphinothioyl]sulfanyl-sulfido-thioxo-phosphane (19.4 g, 51.0 mmol, 76%).

$^{31}P\{^1H\}$ NMR (162 MHz, (pyridine)) δ 104.02 (s).

IR (cm^{-1}) 3087, 3040, 1608, 1450.

M.P. 165.3 – 168.0 °C.

Two other peaks are observable in the phosphorus spectrum, most-likely those of $Py_2P_2S_4O$ and P_2S_5, data in-line with literature.[185]

5-chloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione
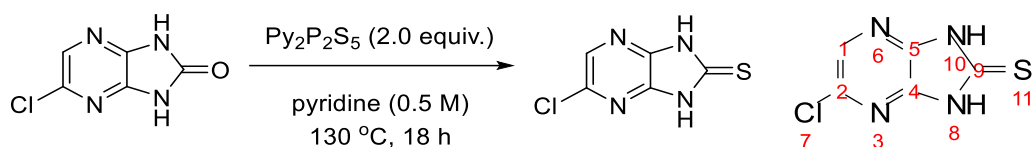
To a 25 mL RBF was added a stirrer bar, 5-chloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-one (656.2 mg, 3.847 mmol), pyridine-phosphorous pentasulfide complex (2.93 g, 2.0 equiv.) and dry pyridine (7.7 mL, 0.5 M). An air condenser was added, and the mixture was stirred at 130 °C for 18 h. The mixture was allowed to cool, and the pyridine removed under reduced pressure. The residue was dissolved in water and transferred to a separatory funnel, washing with EtOAc (sonication aids transfer). The mixture was acidified with HCl (2 M, 20 mL), the layers were separated, and the aqueous layer extracted with EtOAc (3x 20 mL). The organic layers were collected and washed with HCl (2 M, 2x 10 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* Toluene was added to aid the removal of residual pyridine. Provided 5-chloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione as a yellow powder (835 mg, wet, calculated with solvents: 608 mg, 3.26 mmol, 85%).

Adapted from literature conditions.[185]

May be alternatively prepared:
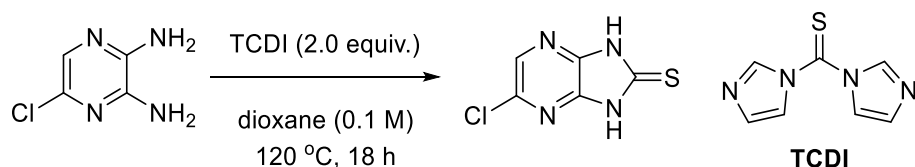
To a 100 mL RBF was added 2,3-diamino-5-chloropyrazine (0.5 g, 3.458 mmol). To the RBF was attached an air condenser, and the system placed under argon atmosphere. Dioxane (35 mL, 0.1 M, anhydrous) was added *via* syringe and the mixture was heated to 80 °C to dissolve starting material. Thiocarbonyl diimidazole (TCDI, 0.65 g, 1.05 equiv.) was added in one batch and the mixture heated with stirring to 120 °C. Monitored by TLC, an additional 0.5 equiv. of TCDI was added after 4 h (0.308 g). Heated to reflux for 18 h. After TLC indicated complete reaction, mixture was allowed to cool to RT, dioxane removed *in vacuo.* The residue was taken up in EtOAc, washed with 1 M HCl (20 mL). The organic layer was removed and the aqueous extracted with EtOAc (3x 30 mL). The organic layers were collected and dried (MgSO$_4$) and solvent removed *in vacuo.* to provide a brown solid. Pure by ^{1}H NMR (593 mg, 3.17 mmol, 92%).

^{1}H NMR (400 MHz, DMSO-*d*$_6$) *δ* 13.63 (br s, 2H, 8, **10**), 8.16 (d, *J* = 1.2 Hz, 1H, **1**).

^{13}C NMR (101 MHz, DMSO-*d*$_6$) *δ* 172.9, 140.4, 140.3, 140.0, 134.4.

IR (cm^{-1}) 2768 (SH), 1616 (NH bend).

M.P. 281.7 – 282.9 °C.

MS ES+ m/z calcd for $C_5H_3{}^{35}ClN_4S$ (M+H)$^+$: 186.9845, found: 186.9844.

5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione
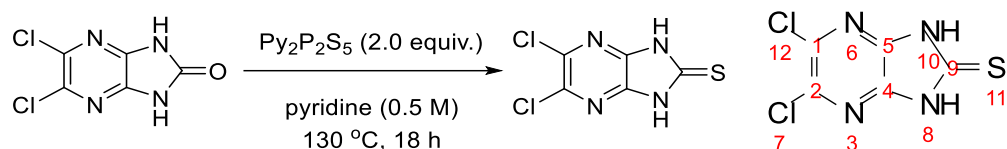
To a 50 mL RBF was added a stirrer bar, 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-one (1.09 g, 5.31 mmol), pyridine-phosphorous pentasulfide complex (4.04 g, 2.0 equiv.), and dry pyridine (10.63 mL, 0.5 M). An air condenser was attached, and the mixture was heated to 130 °C with stirring for 18 h. After allowing to cool to RT, the pyridine was removed *in vacuo.*, the residue was taken up in water and transferred to a separatory funnel, washing with EtOAc. The aqueous layer was acidified with HCl (2 M, 10 mL) and extracted with EtOAc (3x 20 mL). The organic layers were collected and dried (MgSO$_4$) and toluene was added (to azeotropically remove residual pyridine). The solvent was removed *in vacuo.* to afford 1.15 g crude material. This was dry loaded and purified by column chromatography (7:3 Pet. E: EtOAc) to provide a bright orange / yellow solid of 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (435 mg isolated (37%), isolated + mixture yield 48%).

Adapted from literature procedure.[185]

May be alternatively prepared:
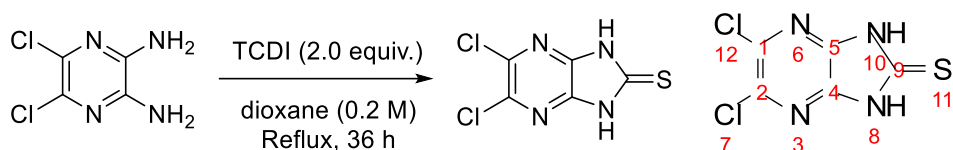
To a 100 mL RBF was added 2,3-diamino-5,6-dichloropyrazine (1.99 g, 11.11 mmol), a stirrer bar and dioxane (56 mL, 0.2 M). This was stirred to dissolve by heating to 80 °C. TCDI (1.98 g, 1.0 equiv.) was added, and the mixture heated to reflux for 24 h. An additional equivalent of TCDI (1.98 g) was added and heating was continued for 12 h. The mixture was allowed to cool to RT and the dioxane removed *in vacuo.* To the residue was added HCl (1 M, 20 mL) and EtOAc (20 mL). The mixture was transferred to a separatory funnel and separated. The aqueous layer was extracted twice with EtOAc (2x 20 mL). The

organic layers were collected and dried (MgSO$_4$) and solvent removed *in vacuo.* The crude material was purified by column chromatography, eluting with 8 : 2 Pet. E : EtOAc to provide a yellow solid of 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (828 mg, 3.74 mmol, 34%).

^{1}H NMR (400 MHz, DMSO-*d*$_6$) *δ* 13.77 (br s, 2H, **8**, **10**).

^{13}C NMR (101 MHz, DMSO-*d*$_6$) *δ* 174.2, **9**, 140.5, **4**, **5**, 136.8, **1**, **2**.

M.P. 360 °C (deg.).

IR (cm^{-1}) 3140 (NH), 1603.

MS ES+ m/z calcd for C$_5$H$_2$^{35}Cl$_2$N$_4$S (M+H)$^+$: 220.9455, found: 220.9455.

Methyl 3-chloro-2,2-dimethoxypropanoate[190]
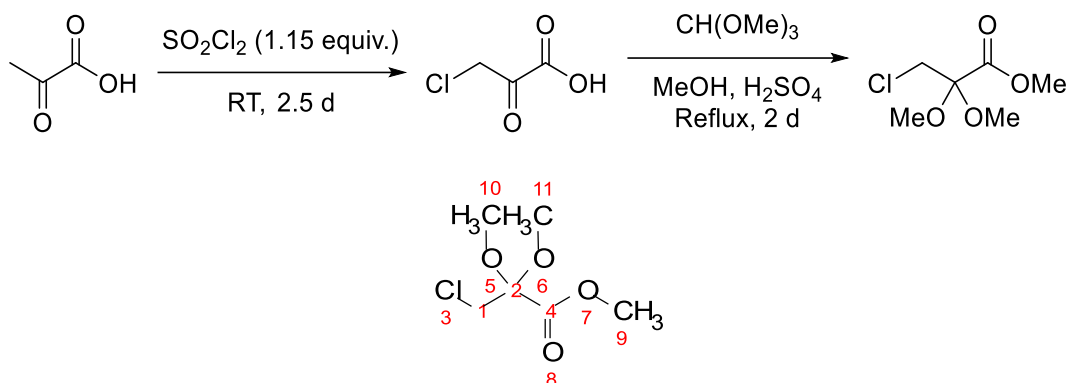
To freshly-distilled pyruvic acid (1.544 g, 17.53 mmol) in a 25 mL RBF was added a stirrer bar and sulfuryl chloride (1.47 mL, 1.15 equiv.), the temperature was regulated using a water bath. This was removed 20 min. after addition and the mixture allowed to stir for 2.5 d at RT. The reaction mixture was concentrated *in vacuo.* and placed in a desiccator *in vacuo.* to yield after a day a waxy yellow solid. This was taken through without purification.

To a 250 mL RBF was transferred the previous reaction mixture and methanol (50 mL, over mol. sieves). Trimethyl orthoformate (15 mL) and sulfuric acid (1.5 mL) were added subsequently, and the mixture set to reflux for 2 d. This was left to cool to RT and poured into NaHCO$_3$ sat. soln. (150 mL) and extracted with chloroform (3x 50 mL). The chloroform extracts were collected and washed with water (100 mL), dried (Na$_2$SO$_4$) and solvent removed *in vacuo.* to yield a colourless oil. Crude ^{1}H NMR indicated a mixture of methyl 3-chloro-2,2-dimethoxypropanoate and methyl 2,2,-dimethoxypropanoate. The crude oil was vacuum distilled (T = 125 °C atmospheric pressure) to provide methyl 3-chloro-2,2-dimethoxypropanoate (593 mg, 3.24 mmol, 19%) as a colourless oil, taken forward without further purification.

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 3.85 (s, 3H, **9**), 3.76 (s, 2H, **1**), 3.32 (s, 6H, **10**, **11**).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 167.6, 101.1, 53.1, 50.5, 42.4.

IR (cm^{-1}): 3462, 2955, 2839, 2739, 2438.

3-chloro-2,2-dimethoxy-1-propanol[190]
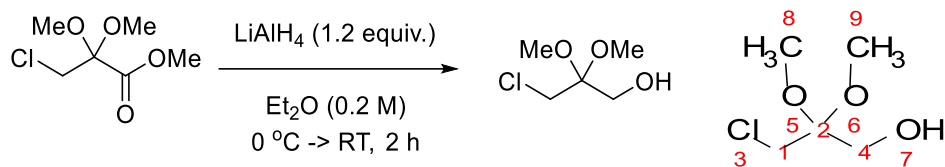
To an oven-dried 8 mL vial was added methyl 2,2-dimethoxy-3-chloropropionate (0.5 g, 2.74 mmol), and this was placed under argon atmosphere. To a separate 30 mL vial was added LiAlH$_4$ (0.125 g, 1.2 equiv.), which was also placed under argon. Anhydrous diethyl ether (10 mL) was added to the 30 mL vial and 4 mL to the vial containing methyl 2,2-dimethoxy-3-chloropropionate. The 30 mL vial was cooled to 0 °C and the methyl 2,2-dimethoxy-3-chloropropionate solution added dropwise *via* syringe with stirring. After addition the vial was warmed to RT and stirring was continued for 3 h. EtOAc (1 mL) was added to quench, followed by water (3 mL), stirring was continued until no more gas evolution was observed. The mixture was then acidified to pH ≈ 2 with 2.0 M HCl followed by rapid neutralisation with sat. soln. NaHCO$_3$. The aqueous layer was separated and extracted with Et$_2$O (3x 10 mL). The organic layers were collected and washed with water (10 mL), dried (Na$_2$SO$_4$) and solvent removed *in vacuo.* to provide 3-chloro-2,2-dimethoxy-1-propanol a yellow oil (> 95% purity by ^{1}H NMR, 142 mg, 0.92 mmol, 34%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 3.72 (d, *J* = 0.6 Hz, 2H, **4**), 3.61 (d, *J* = 0.6 Hz, 2H, **1**), 3.30 (s, 6H, **8**, **9**).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 101.0, 59.7, 48.9, 41.4.

IR (cm^{-1}) 3411 (OH), 2947, 2836, 1636.

N-(bromoacetyl)glycine ethyl ester[212]
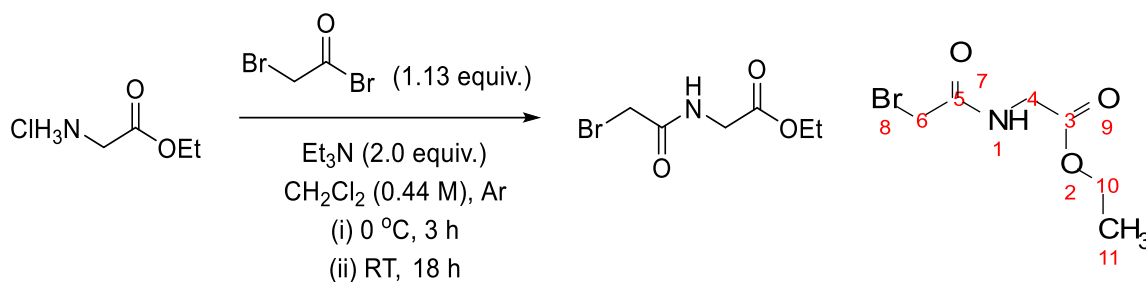
To an oven-dried 8 mL vial was added glycine ethyl ester hydrochloride (250 mg, 1.79 mmol) and placed under argon atmosphere. DCM (4.07 mL, 0.44 M) was then added. Triethylamine (0.5 mL, 2.0 equiv.) was added to the suspension with stirring and the vial cooled to 0 °C. Bromoacetyl bromide (0.179 mL, 1.13 equiv.) was added dropwise *via* syringe and the mixture allowed to stir at 0 °C for 3 h, then warmed to RT and left to stir for 18 h. The reaction was quenched with water (4 mL), transferred to a sep. funnel and more DCM (10 mL) added. The aqueous layer was removed, and the DCM extracted with 1 M HCl (10 mL), water (10 mL), sat. soln. NaHCO$_3$ (10 mL) and brine (10 mL). The organic layer was dried (MgSO$_4$), and the solvent removed *in vacuo.* to provide a brown solid. Partially purified by column chromatography using CHCl$_3$ -> EtOAc : CHCl$_3$ (1:1) to provide a light-orange solid of ethyl *α*-bromoacetamidoacetate (166 mg, 0.74 mmol, 41%). Used in the next step without further purification. Alternatively, the crude was recrystallised from EtOAc to provide an off-white crystalline solid after washing with a small amount of EtOAc.

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 7.03 (br d, *J* = 46.7 Hz, 1H, **1**), 4.27 (q, *J* = 7.2 Hz, 2H, **10**), 4.14 – 4.06 (m, 2H, **4**), 3.94 (s, 2H, **6**), 1.33 (t, *J* = 7.2 Hz, 3H, **11**).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 169.3, 165.8, 62.0, 42.1, 28.7, 14.3.

IR (cm^{-1}) 3264 (NH), 2987, 2944 (C(sp^3)-H), 1736 (C=O ester), 1643 (C=O amide).

M.P. 68 – 69 °C (lit.),[213] 65 – 68 °C.

Data in-line with literature values.

(+)-(2*S*)-methyl 2-(2-bromoacetylamino)-3-(*tert*-butoxy)propanoate
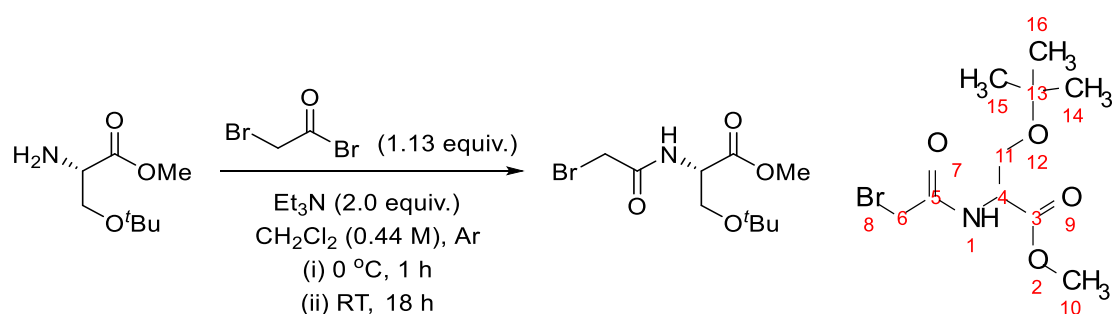
To an 8 mL vial under argon was added (*S*)-H-serine(O*t*Bu) methyl ester (0.25 g, 1.18 mmol), DCM (2.7 mL, 0.44 M) and triethylamine (0.33 mL, 2.0 equiv.). This was cooled to 0 °C and bromoacetyl bromide (0.116 mL, 1.13 equiv.) added dropwise *via* syringe, left to warm to RT over 1 h and left to stir at RT for 18 h. The reaction was quenched with water (4 mL), transferred to a sep. funnel and DCM (10 mL) added. The aqueous layer was removed, and the DCM extracted with 1 M HCl (10 mL), water (10 mL), sat. soln. NaHCO$_3$ (10 mL) and brine (10 mL). The organic layer was dried (MgSO$_4$), and the solvent removed *in vacuo.* to provide a brown solid. Purified by column chromatography (7:3 Pet.E : EtOAc) and collected (+)-(2*S*)-methyl 2-(2-bromoacetylamino)-3-(*tert*-butoxy)propanoate as a yellow oil (161 mg, 0.54 mmol, 46%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 7.20 (br s, 1H, **1**), 4.66 (dt, *J* = 8.3, 3.0 Hz, 1H, **4**), 3.94 (d, *J* = 13.6 Hz, 1H, **6''**), 3.89 (d, *J* = 13.6 Hz, 1H, **6'**), 3.84 (dd, *J* = 9.1, 3.0 Hz, 1H, **11''**), 3.76 (s, 3H, **10**), 3.59 (dd, *J* = 9.1, 3.0 Hz, 1H, **11'**), 1.15 (s, 9H, **14**, **15**, **16**).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 170.5, **3**, 165.5, **5**, 73.8, **13**, 61.7, **11**, 53.5, **4**, 52.7, **10**, 28.9, **6**, 27.4, **14**, **15**, **16**.

IR (cm^{-1}) 3313 (NH), 1746 (C=O ester), 1658 (C=O amide).

Adapted from literature procedure.[212]

N-(bromoacetyl)morpholine[214]
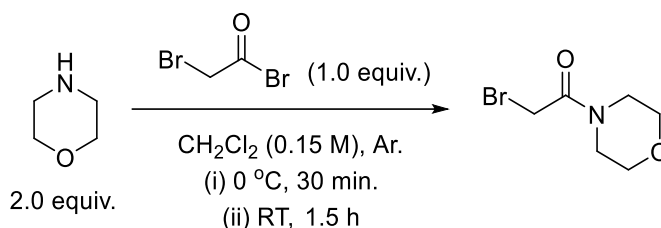
To a 25 mL RBF was added DCM (8.25 mL) and a stirrer bar. This was cooled to 0 °C and placed under argon. To this was added bromoacetyl bromide (0.107 mL, 1.0 equiv.) with stirring. To this solution was added morpholine (0.214 mL, 2.0 equiv.) and left to stir at 0 °C for 30 min. The mixture is allowed to warm to RT and stirred for 1.5 h. A sat. solution of NH_4Cl was added (10 mL), and the mixture extracted with diethyl ether (3x 10 mL). The organic layers were collected and dried (Na_2SO_4) and solvent removed *in vacuo.* to provide an off-white oil which was a mixture of bromoacetic acid and product. Re-dissolved in diethyl ether (approx. 10 mL) and washed with sat. soln. Na_2CO_3 (2x 10 mL). The ether was dried (Na_2SO_4), and solvent removed *in vacuo.* to provide a colourless oil of N-(bromoacetyl)morpholine (132 mg, 0.63 mmol, 52%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 3.87 (s, 2H), 3.78 – 3.72 (m, 2H), 3.72 – 3.68 (m, 2H), 3.67 – 3.61 (m, 2H), 3.57 – 3.48 (m, 2H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 165.5, 66.8, 66.5, 47.3, 42.6, 25.5.

IR (cm^{-1}): 3005, 2866 (C(sp^3)-H), 1650 (C=O).

M.P. approx. 36 °C (melted when handled). Data in-line with literature values.

1-bromopentan-2-one[215]
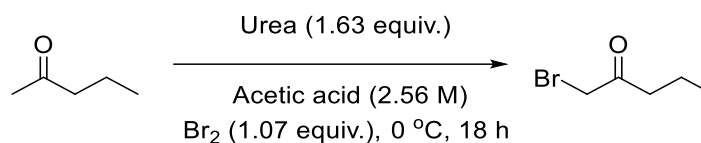
To an oven-dried 25 mL RBF was added urea (1.14 g, 1.63 equiv., 19 mmol), acetic acid (5 mL, 2.56 M) and a stirrer bar. This was cooled to 0 °C and stirred. 2-Pentanone (1.23 mL, 11.6 mmol) was added and then bromine (0.63 mL, 1.07 equiv.) was added at 0 °C with vigorous stirring. The reaction was stirred at RT for 18 h. Water (20 mL) was added, and the mixture transferred to a separatory funnel, extracted with DCM (3x 20 mL). The organic layers were collected and washed with sat. soln. Na_2CO_3 until no more gas evolution was observed. The organic layers were dried ($MgSO_4$), and solvent removed *in vacuo.* The residue was purified *via* column chromatography (3% Et_2O in Hex.) to yield 1-bromopentan-2-one (542 mg, 3.28 mmol, 28%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 3.88 (s, 2H), 2.63 (t, *J* = 7.3 Hz, 2H), 1.65 (h, *J* = 7.3 Hz, 2H), 0.94 (t, *J* = 7.3 Hz, 3H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 202.2, 41.8, 34.4, 17.5, 13.7.

Data in-line with literature values.

Ethyl 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)-acetate
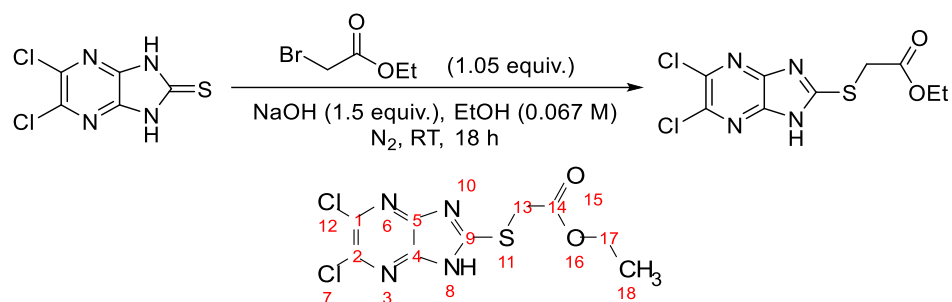
To an 8 mL vial was added 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (100 mg, 0.452 mmol) and sodium hydroxide (28 mg, 1.5 equiv.). The vial was then sealed with a suba seal and wrapped with parafilm. Ethanol (6.75 mL, 0.067 M) was added *via* syringe and the mixture stirred until all solids had dissolved. Ethyl bromoacetate (53 µL, 1.05 equiv.) was added *via* microsyringe and the reaction was left to stir for 18 h. The solvent was removed under reduced pressure and the solid redissolved in NH₄Cl solution and EtOAc (*ca.* 5 mL each) and the mixture transferred to a separatory funnel. The organic layer was removed, and the aqueous layer extracted with EtOAc (3x 10 mL). The organic layers were collected and dried (MgSO₄) and solvent removed *in vacuo*. The crude material was purified by column chromatography eluting with Pet. E. : EtOAc (7:3) to provide an off-white solid (106 mg, 0.344 mmol, 76%). A sample for biological testing was prepared by triturating the solid with cold *n*-hexane (*ca.* 4 mL) and recrystallising from ethanol to provide a white solid of ethyl 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)-acetate.

^{1}H NMR (400 MHz, C*D*₃CN) *δ* 4.19 (q, *J* = 7.1 Hz, 2H, **17**), 4.17 (s, 2H, **13**), 1.24 (t, *J* = 7.1 Hz, 3H, **18**).

^{13}C NMR (126 MHz, C*D*₃CN) *δ* 168.9, **14**, 160.7, **9**, 139.5, **1**, **2**, 62.9, **17**, 34.3, **13**, 14.4, **18** (6 out of 7 possible carbon resonances observed).

M.P. 167.6 – 168.1 °C.

IR (cm^{-1}) 3084 (NH), 1731 (C=O), 1595 (NH bend).

MS ES+ m/z calcd for $C_9H_8{}^{35}Cl^{37}ClN_4O_2S$ (M+H)$^{+}$ 308.9794, found: 308.9792.

Adapted from literature procedure.[167]

2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetonitrile
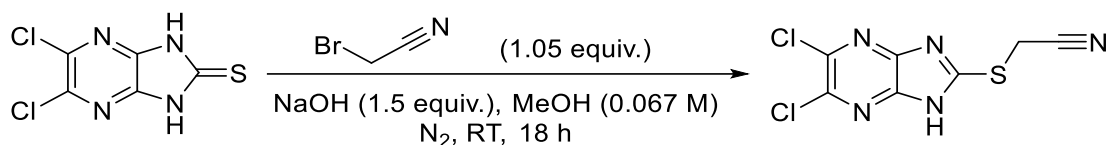
To an 8 mL vial was added 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (100 mg, 0.452 mmol) and sodium hydroxide (28 mg, 1.5 equiv.). The vial was then sealed with a suba seal and wrapped with parafilm. Methanol (6.75 mL, 0.067 M) was added *via* syringe and the mixture stirred until all solids had dissolved. Bromoacetonitrile (33 µL, 1.05 equiv.) was added *via* microsyringe, and the reaction was left to stir for 18 h. The solvent was removed under reduced pressure and the solid redissolved in ammonium chloride solution and ethanol (*ca.* 5 mL each) and the mixture transferred to a separatory funnel. Separation of the single-phase mixture was achieved by adding a small amount of brine solution. The organic layer was removed, and the aqueous layer extracted with EtOAc (3x 10 mL). The organic layers were collected and dried (MgSO$_4$) and solvent removed *in vacuo.* The crude material was purified by column chromatography eluting with Pet. E. : EtOAc (7:3) to provide an off-white solid (63.3 mg, 0.243 mmol, 54%). A sample for biological testing was prepared by triturating the solid with cold *n*-hexane (*ca.* 4 mL) and recrystallising from ethanol to provide a light yellow-white solid of 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetonitrile.

^{1}H NMR (400 MHz, C*D*$_3$CN) *δ* 4.25 (s, 2H).

^{13}C NMR (101 MHz, C*D*$_3$CN) *δ* 158.3, 140.0, 117.5, 18.0 (4 out of 5 carbon resonances observed).

M.P. 213.5 – 216 °C (deg.).

IR (cm^{-1}) 3100 (NH), 2248 (C≡N).

MS ES+ m/z calcd for C$_7$H$_3$^{35}Cl$_2$N$_5$S (M+H)$^+$: 260.9584, found: 260.9583.

Adapted from literature procedure.[167]

2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)propionic acid
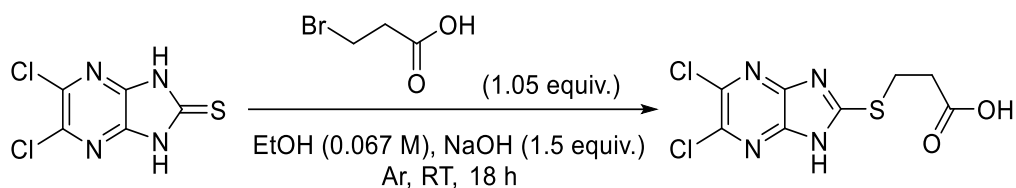
To an 8 mL vial was added a stirrer bar, 5,6-dichloroimidaazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (75 mg, 0.34 mmol), sodium hydroxide (20. 3 mg, 1.5 equiv.) and ethanol (over mol. sieves, 5.0 mL), stirred to dissolve. 3-bromopropionic acid (54.5 mg, 1.05 equiv.) was added and the vial placed under argon, left to stir for 18 h at RT. The solvent was removed *in vacuo.* and the residue dissolved in water (5 mL), extracted twice with diethyl ether (2x 10 mL), the aqueous layer was then acidified to pH approx. 2 and re-extracted with EtOAc (3x 10 mL), the EtOAc layers were then dried (MgSO$_4$) and solvent removed *in vacuo.* Further purified by column chromatography, eluting with Pet. E : EtOAc (6:4 with 0.4% acetic acid) to provide 2-({5,6-dichloro-1*H*-imidazo[4,5-b]pyrazine-2-yl}sulfanyl)propionic acid (23.8 mg, 0.08 mmol, 24%) as a yellow solid.

^{1}H NMR (400 MHz, C*D*$_3$CN) *δ* 3.55 (t, *J* = 6.8 Hz, 2H), 2.87 (t, *J* = 6.8 Hz, 2H). O*H* and N*H* signals not observed.

^{13}C NMR (101 MHz, C*D*$_3$CN) *δ* 173.0, 161.9, 139.2, 34.4, 27.4 (5 out of 6 carbon resonances observed).

M.P. 205.1 – 206.2 °C (deg.).

IR (cm^{-1}): 3111 (NH), 3023 (OH), 1721 (C=O).

MS ES+ m/z calcd for C$_8$H$_6$^{35}Cl$_2$N$_4$O$_2$S (M+H)$^+$: 293.9689, found: 293.9688.

Adapted from literature procedure.[167]

2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetylmorpholine
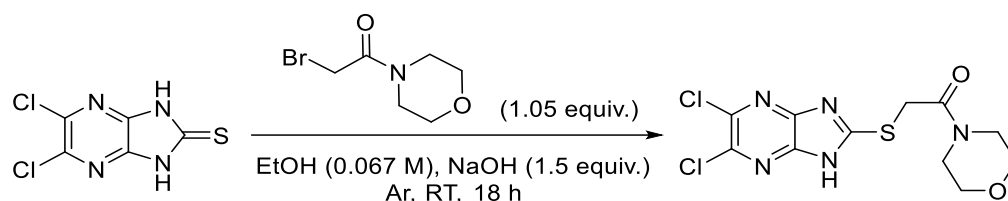
To an 8 mL vial was added a stirrer bar, 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (75 mg, 0.34 mmol), sodium hydroxide (20.3 mg, 1.5 equiv.). Ethanol (over mol. sieves, 5.0 mL) was added *via* syringe. *N*-(bromoacetyl)morpholine (74.1 mg, 1.05 equiv.) was added in one batch and the reaction placed under argon atmosphere. Left to stir at RT for 18 h. Afterwards, the solvent is removed *in vacuo.* and the residue redissolved in sat. soln. NH$_4$Cl (10 mL), and EtOAc (10 mL), transferred to a separatory funnel and the organic layer removed. Aqueous layer extracted with EtOAc (3x 10 mL), organic layers collected and dried (MgSO$_4$) and solvent removed *in vacuo.* The residue purified by column chromatography (6:4 Pet. E : EtOAc -> EtOAc) to provide an off-white solid of 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetylmorpholine (53.4 mg, 0.153 mmol, 45%).

^{1}H NMR (400 MHz, C*D*$_3$CN) *δ* 4.37 (s, 2H), 3.73 – 3.66 (m, 2H), 3.65 – 3.61 (m, 2H), 3.61 – 3.54 (m, 4H).

^{13}C NMR (101 MHz, C*D*$_3$CN) *δ* 166.8, 161.7, 139.3, 67.1, 67.1, 47.3, 43.4, 35.6 (8 out of 9 carbon resonances observed).

M.P. 214.4 – 215.1 °C.

IR (cm^{-1}) 3143 (NH), 2771, 1623 (C=O).

MS ES+ m/z calcd for C$_{11}$H$_{11}$^{35}Cl$_2$N$_5$O$_2$S (M+H)$^+$: 349.0113, found: 349.0112.

Adapted from literature procedure.[167]

2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)pentane
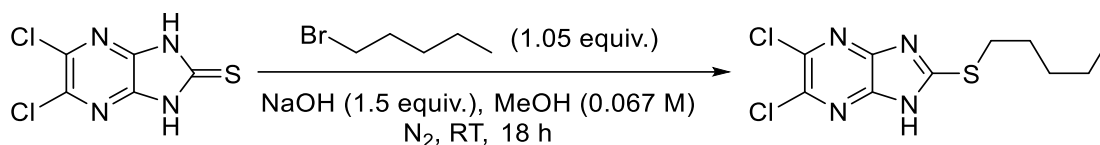
To an 8 mL vial was added 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (100 mg, 0.452 mmol) and sodium hydroxide (28 mg, 1.5 equiv.). The vial was then sealed with a suba seal and wrapped with parafilm. Methanol (6.75 mL, 0.067 M) was added *via* syringe and the mixture stirred until all solids had dissolved. 1-Bromopentane (59 µL, 1.05 equiv.) was added *via* microsyringe, and the reaction was left to stir for 18 h. The solvent was removed under reduced pressure and the solid redissolved in NH$_4$Cl solution and EtOAc (*ca.* 5 mL each) and the mixture transferred to a separatory funnel. The organic layer was removed, and aqueous layer extracted with EtOAc (3x 10 mL). The organic layers were collected and dried (MgSO$_4$) and solvent removed *in vacuo.* The crude material was purified by column chromatography eluting with Pet. E. : EtOAc (95:5) to provide 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)pentane as a white solid (62.1 mg, 0.213 mmol, 47%). A sample for biological testing was prepared by triturating the solid with cold *n*-hexane (*ca.* 4 mL).

^{1}H NMR (500 MHz, C*D*$_3$CN) *δ* 11.16 (br s, 1H), 3.43 – 3.31 (m, 2H), 1.86 – 1.76 (m, 2H), 1.50 – 1.40 (m, 2H), 1.40 – 1.29 (m, 2H), 0.91 (t, *J* = 7.2 Hz, 3H).

^{13}C NMR (126 MHz, C*D*$_3$CN) *δ* 162.5, 145.0, 139.0, 32.2, 31.4, 29.8, 22.8, 14.2.

M.P. 145.9 – 146.3 ^{o}C.

IR (cm^{-1}) 3088 (NH), 2959, 2930, 2861 (C(sp^3)-H).

MS ES+ m/z calcd for C$_{10}$H$_{12}$^{35}Cl^{37}ClN$_4$S (M+H)$^+$: 293.0208, found: 293.0209.

Adapted from literature procedure.[167]

2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)pentan-2-one

To an 8 mL vial was added 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (100 mg, 0.452 mmol) and sodium hydroxide (28 mg, 1.5 equiv.). The vial was then sealed with a suba seal and wrapped with parafilm. Methanol (6.75 mL, 0.067 M) was added *via* syringe and the mixture stirred until all solids had dissolved. 1-Bromopentan-2-one (70 µL, 1.05 equiv.) was added *via* microsyringe, and the reaction was left to stir for 18 h. The solvent was removed under reduced pressure and the solid redissolved in NH_4Cl solution and EtOAc (*ca.* 5 mL each) and the mixture transferred to a separatory funnel. The organic layer was removed, and the aqueous layer extracted with EtOAc (3x 10 mL). The organic layers were collected and dried ($MgSO_4$) and solvent removed *in vacuo.* The crude material was purified by column chromatography eluting with Pet. E. : EtOAc (8:2) to provide 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)pentan-2-one as a white solid (74.7 mg, 0.245 mmol, 54%). Consistent with the previous report of imidazo[4,5-b]pyrazine-2-thioacetoketone compounds, I believe this compound is in equilibrium with its tricyclic hemiaminol, as observed by ¹H NMR and IR. The ratio of free ketone to tricyclic alcohol is approx. 8:2 (molar).
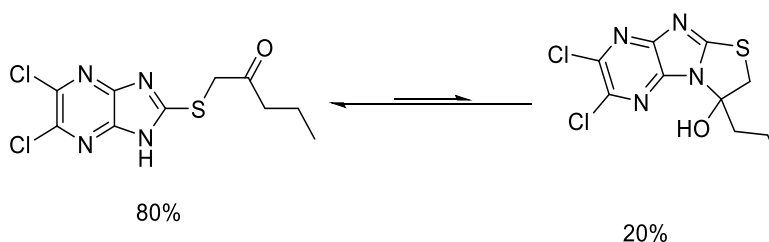
M.P. 151.2 – 151.9 °C.

IR (cm⁻¹) 1743 (C=O), 1584 (NH bend).

MS ES+ m/z calcd for $C_{10}H_{10}{}^{35}Cl_2N_4OS$ (M+H)⁺: 305.0031, found: 305.0032.

Free ketone: ¹H NMR (400 MHz, C*D*₃CN) *δ* 11.11 (br s, 1H, **8**), 4.32 (s, 2H, **13**), 2.65 (t, *J* = 7.3 Hz, 2H, **16**), 1.62 (h, *J* = 7.3 Hz, 2H, **17**), 0.92 (t, *J* = 7.3 Hz, 3H, **18**).

^{13}C NMR (101 MHz, C*D*₃CN) δ 204.2, *14*, 161.3, *9*, 139.3, *1*, *2*, 44.0, *16*, 42.1, *13*, 17.9, *17*, 13.8, *18* (7 out of 8 carbon resonances observed).

Hemiaminol: ^{1}H NMR (400 MHz, C*D*₃CN) δ 4.12 (d, J = 12.3 Hz, 1H), 3.75 (d, J = 12.3 Hz, 1H), 2.62 – 2.55 (m, 2H), 2.33 (ddd, J = 13.9, 12.0, 4.6 Hz, 2H), 1.00 (t, J = 7.4 Hz, 3H).

Observed carbon peaks for tricyclic alcohol (alkyl peaks only): ^{13}C NMR (101 MHz, C*D*₃CN) δ 45.9, 40.6, 18.0, 13.8. Adapted from literature procedure.[167]

2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)cyclohexane
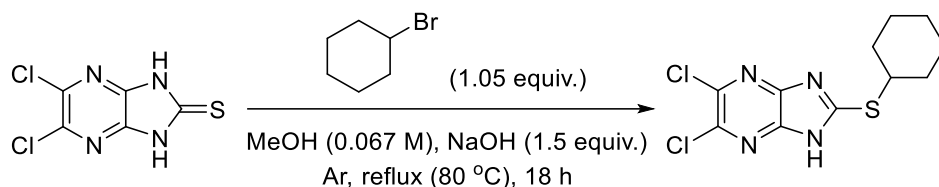
To a 25 mL RBF was added 5,6-dichloroimidazo[4,5-b]pyrazine-2(1,3*H*)-thione (75 mg, 0.339 mmol), sodium hydroxide (20 mg, 1.5 equiv.) and was placed under argon atmosphere with attached air condenser. Methanol (5 mL, 0.067 M) was added *via* syringe and stirred to dissolve. Cyclohexyl bromide (0.207 mL, 5.0 equiv.) was added *via* syringe and the solution stirred at reflux for 18 h. The mixture was allowed to cool to RT and solvent removed *in vacuo.* The residue was re-dissolved in NH₄Cl soln. and EtOAc (5 mL each), separated and the aqueous layer extracted with EtOAc (3x 10 mL). The organic layers were collected and dried (MgSO₄) and solvent removed *in vacuo.* The crude was purified by column chromatography (9:1 Pet E : EtOAc) to provide 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)cyclohexane as a white solid (16.4 mg, 0.054 mmol, 16%).

^{1}H NMR (400 MHz, C*D*Cl₃) δ 10.15 (br s, 1H), 4.16 (tt, J = 10.4, 3.8 Hz, 1H), 2.22 (dt, J = 13.7, 6.8 Hz, 2H), 1.79 (dt, J = 13.3, 4.3 Hz, 2H), 1.71 – 1.58 (m, 3H), 1.48 (tt, J = 11.0, 3.3 Hz, 2H), 1.41 – 1.30 (m, 1H).

^{13}C NMR (101 MHz, C*D*Cl₃) δ 160.7, 46.5, 33.2, 25.9, 25.4 (5 out of 7 carbon resonances observed).

M.P. 200.8 – 202.1 °C.

IR (cm^{-1}): 3120 (NH).

MS ES+ m/z calcd for C₁₁H₁₂^{35}Cl₂N₄S (M+H)$^+$: 304.0275, found: 304.0275.

Adapted from literature procedure.[167]

Ethyl 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetamidoacetate
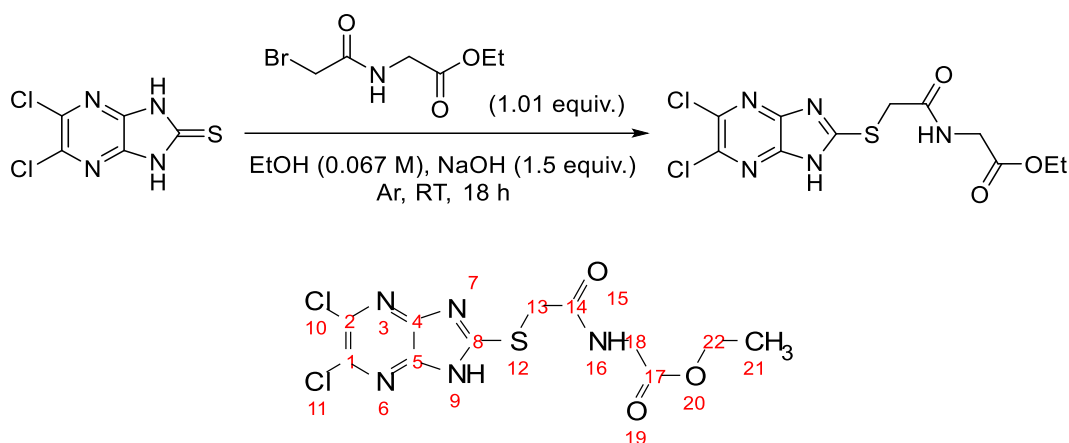
To a 25 mL RBF was added a stirrer bar, 5,6-dichlorimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (150 mg, 0.678 mmol), sodium hydroxide (40.7 mg, 1.5 equiv.) and placed under argon atmosphere. Ethanol (over mol. sieves, 5.0 mL) was added with stirring to dissolve starting material. Ethanol (over mol. sieves, 5.0 mL) was added to a separate vial containing ethyl bromo-acetamidoacetate (0.154 g, 1.01 equiv.) and stirred to dissolve. This solution was then transferred to the 25 mL RBF *via* syringe and left to stir at RT for 18 h. The solvent was removed *in vacuo.* and the residue redissolved in sat. soln. NH₄Cl (10 mL), and EtOAc (10 mL), transferred to a separatory funnel and the organic layer removed. Aqueous layer extracted with EtOAc (3x 10 mL), organic layers collected and dried (MgSO₄) and solvent removed *in vacuo.* and residue purified by column chromatography (6:4 Pet. E : EtOAc -> EtOAc) to provide a white fluffy powder of ethyl 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetamidoacetate (133 mg, 0.36 mmol, 54%).

¹H NMR (400 MHz, C*D*₃CN) *δ* 11.66 (br s, 1H, **9**), 7.41 (br s, 1H, **16**), 4.12 (q, J = 7.1 Hz, 2H, **22**), 4.07 (s, 2H, **13**), 3.92 (d, J = 5.8 Hz, 2H, **18**), 1.19 (t, J = 7.1 Hz, 3H, **21**).

¹³C NMR (101 MHz, C*D*₃CN) *δ* 169.9, **17**, 168.7, **14**, 160.8, **8**, 139.2, **1**, **2**, 61.6, **22**, 42.0, **18**, 34.8, **13**, 14.0, **21**. (8 out of 9 carbon resonances observed)*.*

M.P. 187.7 – 188.3 °C.

IR (cm⁻¹) 3292 (NH), 3219 (NH), 1717 (C=O ester), 1648 (C=O amide).

MS ES+ m/z calcd. for C₁₁H₁₁³⁵Cl₂N₅O₃S (M+H)⁺: 365.0062, found: 365.0063.

Adapted from literature procedure.[167]

Methyl 3-(*tert*-butoxy)-2(*S*)-[2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetamido]propionate
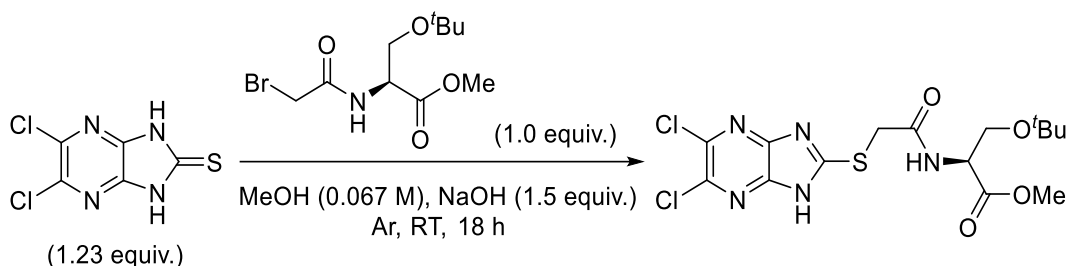
To a 20 mL vial was added 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (150 mg, 0.678 mmol, 1.23 equiv.), sodium hydroxide (40 mg, 1.5 equiv.) and placed under argon atmosphere. To a separate vial was added (+)-(2*S*)-methyl 2-(2-bromoacetylamino)-3-(*tert*-butoxy)propanoate (167 mg, 0.55 mmol) and this was dissolved with methanol (5 mL). This solution was transferred *via* syringe to the reaction vial and stirred at RT for 18 h. The solvent is removed *in vacuo.* and the residue taken up in ammonium chloride and EtOAc (5 mL each). The layers were separated and the aqueous extracted with EtOAc (3x 10 mL). The organic layers were collected and dried (MgSO₄) and solvent removed under reduced pressure. The crude material was purified by column chromatography (6:4 -> 5:5, Pet E. : EtOAc) to provide methyl 3-(*tert*-butoxy)-2-(*S*)-[2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetamido]propionate as an off-white solid (173 mg, 0.397 mmol, 59%).

^{1}H NMR (400 MHz, Acetone-d_6) δ 12.93 (br s, 1H), 8.06 (br d, J = 8.5 Hz, 1H), 4.63 (dt, J = 8.5, 3.5 Hz, 1H), 4.24 (s, 2H), 3.80 (dd, J = 9.2, 3.5 Hz, 1H), 3.66 (s, 3H), 3.59 (dd, J = 9.2, 3.5 Hz, 1H), 1.08 (s, 9H).

^{13}C NMR (101 MHz, Acetone-d_6) δ 171.0, 168.3, 161.7, 139.2, 73.8, 62.6, 54.4, 52.4, 35.1, 27.5 (10 out of 11 carbon resonances observed).

M.P. 145.7 – 146.2 °C.

IR (cm^{-1}) 3272 (NH), 1756 (C=O ester), 1719 (C=O amide).

MS ES+ m/z calcd for $C_{15}H_{19}{}^{35}Cl_2N_5O_4S$ (M+H)$^+$: 436.0613, found: 436.0623.

Adapted from literature procedure.[167]

Ethyl 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetoacetonate
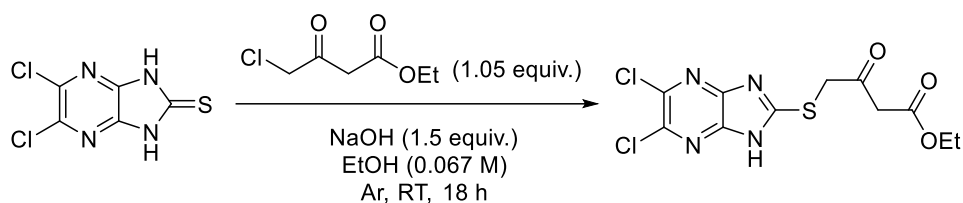
To an 8 mL RBF was added a stirrer bar, 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (100 mg, 0.452 mmol), sodium hydroxide (27 mg, 1.5 equiv.) and placed under an argon atmosphere. Ethanol (6.75 mL, 0.067 M) was added *via* syringe to dissolve materials. Ethyl 4-chloroacetoacetate (64.2 µL, 1.05 equiv.) was added *via* microsyringe and the reaction allowed to stir at RT for 18 h. Afterwards, the solvent was removed under reduced pressure, and the residue taken up in EtOAc and NH$_4$Cl soln. (5 mL each). The layers were separated, and the aqueous layer extracted with EtOAc (3x 10 mL). The organic layers were collected and dried (MgSO$_4$) and solvent removed under reduced pressure. The crude was purified by column chromatography (7:3 -> 6:4 Pet E : EtOAc) to provide ethyl 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetoacetonate as a beige solid (63.0 mg, 0.018 mmol, 39%). Again, I believe this compound is in equilibrium with its cyclised hemiaminol, in approx. 76 : 23 mol ratio.
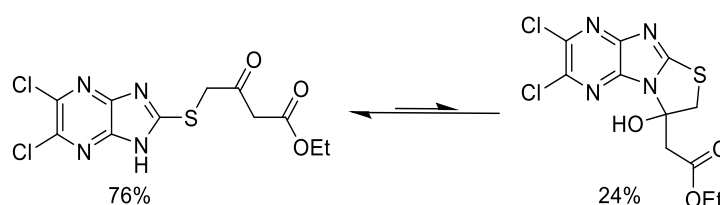
IR (cm^{-1}) 3157 (NH), 1736 (C=O).

M.P. 179.4 – 180.0 °C.

MS ES+ m/z calcd for C$_{11}$H$_{10}$^{35}Cl$_2$N$_4$O$_3$S (M+H)$^+$: 348.9929, found (-1.0208 average neutral loss): 347.9720.

Free β-ketoester: ^{1}H NMR (400 MHz, 2:1 v/v DMSO-d_6 : CD_3CN) δ 13.92 (br s, 1H), 4.43 (s, 2H), 4.12 (q, *J* = 7.1 Hz, 2H), 3.75 (s, 2H), 1.19 (t, *J* = 7.1 Hz, 3H).

^{13}C NMR (101 MHz, CD_3CN) δ 197.7, 167.7, 161.4, 61.9, 48.9, 41.8, 14.5 (7 out of 9 carbon resonances found).

Cyclised hemiaminal: ^{1}H NMR (400 MHz, 2:1 v/v DMSO-d_6 : CD_3CN) δ 4.38 (d, *J* = 12.4 Hz, 1H), 4.06 (q, *J* = 7.1 Hz, 2H), 3.96 (d, *J* = 12.4 Hz, 1H), 3.66 (d, *J* = 16.0 Hz, 1H), 3.43 (d, *J* = 16.0 Hz, 1H), 1.14 (t, *J* = 7.1 Hz, 3H). No ^{13}C signals observed. Adapted from literature conditions.[167]

Ethyl 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)butyrate
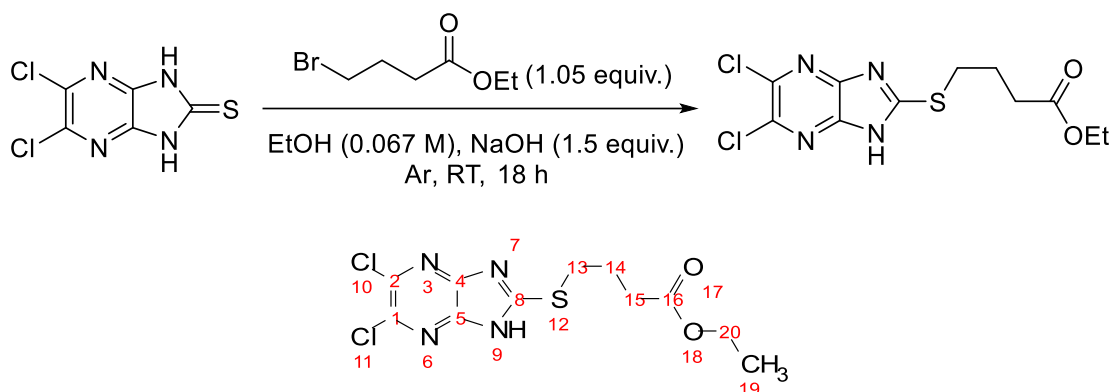
To a 25 mL RBF was added a stirrer bar, 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (100 mg, 0.452 mmol) and sodium hydroxide (27.4 mg, 1.5 equiv.) and placed under argon atmosphere. Ethanol (over mol. sieves, 6.75 mL) was added with stirring. Ethyl 4-bromobutyrate (67.9 µL, 1.05 equiv.) was added via syringe with stirring. Allowed to stir at RT for 18 h. The solvent was removed under reduced pressure and the solid redissolved in sat. soln. NH$_4$Cl (10 mL) and EtOAc (10 mL). Transferred to a separatory funnel and the organic layer separated. The aqueous layer was extracted with EtOAc (3x 10 mL), the organic layers were collected and dried (MgSO$_4$). EtOAc was removed *in vacuo.* to provide a beige solid which was purified by column chromatography, eluting with 8:2 Pet. E : EtOAc to provide a colourless oil which solidifies over time into a light-orange solid. Further purified by triturating with *n*-hexane (5 mL) and ice-cold ethanol (5 mL) to provide a white solid of ethyl 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)butyrate (70.0 mg, 0.20 mmol, 46%).

1H NMR (400 MHz, C*D*$_3$CN) δ 11.19 (br s, 1H, **9**), 4.11 (q, J = 7.2 Hz, 1H, **20**), 3.39 (t, J = 7.3 Hz, 1H, **13**), 2.47 (t, J = 7.3 Hz, 1H, **15**), 2.09 (p, J = 7.3 Hz, 1H, **14**), 1.21 (t, J = 7.2 Hz, 1H, **19**).

^{13}C NMR (101 MHz, C*D*$_3$CN) *δ* 173.7, 162.1, 139.1, 61.3, 33.3, 31.3, 25.7, 14.5 (8 out of 9 carbon resonances observed).

M.P. 120.3 – 121.3 °C.

IR (cm^{-1}): 3151 (NH), 1691 (C=O).

MS ES+ m/z calcd for C$_{11}$H$_{12}$^{35}Cl$_2$N$_4$O$_2$S (M+H)$^+$: 335.0136, found: 335.0135.

Adapted from literature procedure.[167]

Methyl (3-hydroxy)-2(*S*)-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)propanioate
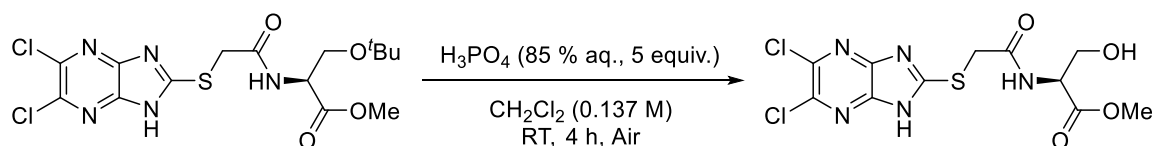
To an 8 mL vial was added a stirrer bar and methyl (3-*tert*-butoxy)-2(*S*)-[(5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-sulfanyl)acetamido]propanioate (60 mg, 0.137 mmol). DCM (1 mL, 0.137 M) was added *via* syringe. Once the material had dissolved, phosphoric acid (79 µL, 85% aq. 5 equiv.) was added dropwise with stirring at RT. Left to stir at RT over 4 h, the reaction was diluted with water (5 mL) and extracted with EtOAc (3x 10 mL). The aqueous phase was neutralised (NaHCO$_3$) and re-extracted with acetonitrile (3x 10 mL) the organic layers were collected and dried (MgSO$_4$) and solvent removed *in vacuo.* to provide methyl (3-hydroxy)-2(*S*)-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)propanioate as an off-white solid (29.7 mg, 0.078 mmol, 57%).

^{1}H NMR (400 MHz, 1:6 v/v DMSO-d_6 : CD_3CN) *δ* 8.17 (s, 1H), 4.43 (dt, *J* = 8.0, 4.1 Hz, 1H), 4.15 (d, *J* = 2.9 Hz, 2H), 3.80 (dd, *J* = 11.3, 4.3 Hz, 1H), 3.68 (dd, *J* = 11.3, 3.9 Hz, 1H), 3.64 (s, 3H).

^{13}C NMR (101 MHz, 1:6 v/v DMSO-d_6 : CD_3CN) *δ* 171.6, 168.2, 162.2, 138.7, 62.5, 56.2, 52.7, 35.3 (8 out of 9 carbon resonances observed).

Followed literature procedure.[194]

General procedure 4.1: Ester hydrolysis

({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetic acid
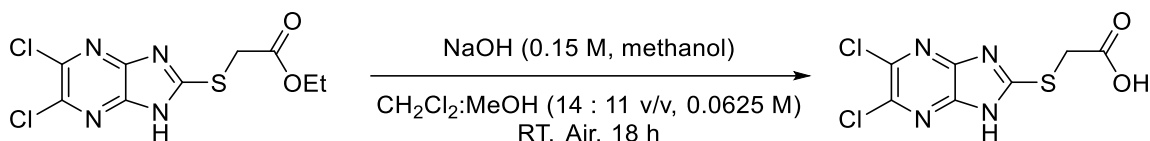
A stock solution of NaOH (0.25 M) (NaOH (50 mg) in methanol (5 mL)) and DCM : MeOH (9:1 v/v, 10 mL) were prepared. To an 8 mL vial was added ethyl 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)-acetate (50 mg, 0.162 mmol) and a stirrer bar. DCM : MeOH (9:1 v/v, 1.62 mL) solution was added to reach a concentration of 0.1 M. Methanolic NaOH solution was added (0.97 mL) to reach a NaOH concentration of 0.15 M and a final reagent concentration of 0.0625 M. The mixture was left to stir at RT for 18 h. Afterwards, TLC analysis (7:3 Pet E: EtOAc) indicated complete consumption of starting material, solvents were removed under reduced pressure and the residue dissolved in water (5 mL). Extracted with Et$_2$O (2x 10 mL) and the aqueous layer acidified (2 M HCl, 5 mL) and re-extracted with Et$_2$O (3x 10 mL), these extracts were collected and dried (MgSO$_4$) and solvent removed under reduced pressure to provide ({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetic acid as a light-yellow solid (31.4 mg, 0.11 mmol, 69%).

^{1}H NMR (400 MHz, 1:2 v/v DMSO-d_6 : CD_3CN) $δ$ 13.24 (br s, 1H), 4.14 (s, 2H).

^{13}C NMR (101 MHz, 1:2 v/v DMSO-d_6 : CD_3CN) $δ$ 169.6, 161.2, 144.8, 139.5, 34.2.

M.P. 219.1 – 220.7 °C.

IR (cm^{-1}) 3111 (NH), 3027 (OH), 1740 (C=O).

MS ES+ m/z calcd for C$_7$H$_4$^{35}Cl$_2$N$_4$O$_2$S (M+H)$^+$: 278.9510, found (-1.979 average neutral loss): 276.9719.

Adapted from literature procedure.[195]

4-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)butanoic acid
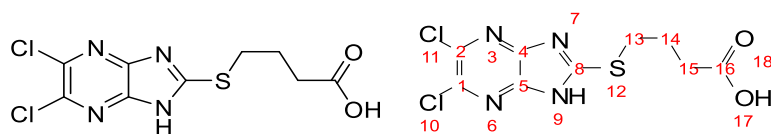
Followed general procedure 4.1. Ethyl 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)butyrate (50 mg, 0.155 mmol), DCM : MeOH solution (1.40 mL), NaOH solution (0.25 M, 0.93 mL). Provided 4-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)butanoic acid as a light-yellow solid (30.0 mg, 0.097 mmol, 63%).

^{1}H NMR (400 MHz, 1:2 v/v DMSO-d_6 : CD_3CN) δ 12.27 (br s, 1H, **17**), 3.37 (t, J = 7.2 Hz, 2H, **13**), 2.40 (t, J = 7.3 Hz, 2H, **15**), 2.02 (h, J = 7.3 Hz, 2H, **14**).

^{13}C NMR (101 MHz, 1:2 v/v DMSO-d_6 : CD_3CN) δ 174.6, 162.2, 144.0, 137.9, 32.8, 30.9, 25.2.

MS ES+ m/z calcd for C$_9$H$_8$^{35}Cl^{37}ClN$_4$O$_2$S (M+H)$^+$: 310.9766, found: 310.9765.

Adapted from literature procedure.[195]

2-({5,6-dichloro-1H-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetamidoacetic acid
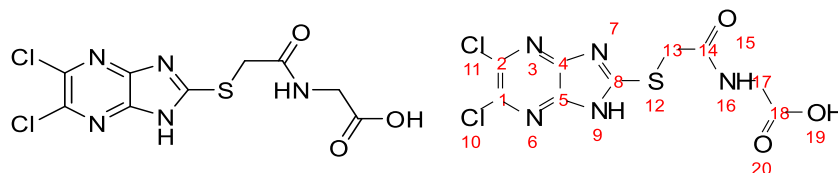
Followed general procedure 4.1. Ethyl 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetamidoacetate (50 mg, 0.137 mmol), DCM : MeOH solution (1.37 mL). NaOH (0.25 M, 0.82 mL). Provided 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetamidoacetic acid as a light-yellow solid (26.0 mg, 0.077 mmol, 56%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 14.11 (br s, 1H, **9**), 12.66 (br s, 1H, **19**), 8.65 (t, J = 5.8 Hz, 1H, **16**), 4.23 (s, 2H, **13**), 3.81 (d, J = 5.8 Hz, 2H, **17**).

^{13}C NMR (101 MHz, DMSO) δ 170.9, 166.7, 161.0, 41.1, 34.6 (5 out of 7 carbon resonances observed).

M.P. 237.8 – 238.5 °C (deg.)

IR (cm^{-1}) 3347 (NH), 3259 (OH), 1716 (C=O acid), 1649 (C=O amide).

MS ES+ m/z calcd for C$_9$H$_7$^{35}Cl$_2$N$_5$O$_3$S (M+H)$^+$: 335.9725, found: 335.9718.

Adapted from literature procedure.[195]

3-hydroxy-2(*S*)-[({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetamido]propanoic acid
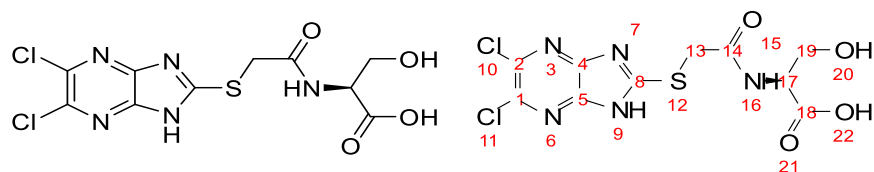
Followed general procedure 4.1. Methyl 3-hydroxy-2(*S*)-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)propanioate (29.7 mg, 0.078 mmol), DCM : MeOH solution (0.781 mL, 0.1 M), NaOH solution (0.468 mL, 0.25 M). Provided 3-hydroxy-2(*S*)-[({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetamido]propionic acid as a white solid (24.6 mg, 0.067 mmol, 86%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 14.04 (br s, 1H), 12.69 (br s, 1H), 8.57 (d, *J* = 7.9 Hz, 1H), 5.03 (br s, 1H), 4.29 (dt, *J* = 7.9, 4.6 Hz, 1H), 4.26 (s, 2H), 3.71 (dd, *J* = 10.9, 5.0 Hz, 1H), 3.63 (dd, *J* = 10.8, 4.2 Hz, 1H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 171.6, 166.5, 161.1, 61.2, 55.1, 34.7 (6 out of 8 carbon resonances observed).

M.P. 139.3 – 140.2 °C (deg.).

IR (cm^{-1}): 3326 (OH), 1721 (C=O acid), 1657 (C=O amide).

$[\alpha]^{25}_{589}$ = 115.38 (EtOAc, 0.026 M, *l* = 1 cm, α = 0.03).

MS ES+ m/z calcd for $C_{10}H_9{}^{35}Cl_2N_4O_4S$ (M+H)$^+$: 351.9800, found (1.992 average neutral loss): 349.9878.

Adapted from literature procedure.[195]

2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)-*N*-ethylacetamide
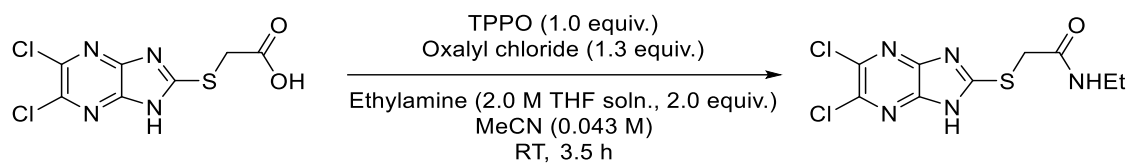
To a reaction vial was added triphenylphosphine oxide (29.7 mg, 1.0 equiv.) and MeCN (1 mL), with stirring was added freshly distilled oxalyl chloride (11.9 µL, 1.3 equiv.). The mixture was stirred for 5 min., then ({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetic acid (30 mg, 0.107 mmol) dissolved in MeCN (1.5 mL) was added *via* syringe. Ethylamine (2.0 M in THF, 0.107 mL, 2.0 equiv.) was added dropwise *via* syringe. Stirring was continued for 3.5 h, then EtOAc (5 mL) was added, and the mixture transferred to a sep. funnel and washed twice with NaHCO$_3$ sat. soln. (10 mL). The organic layer was collected and dried (MgSO$_4$) and solvent removed *in vacuo.* Purified by preparative TLC (500 µm thick) using two solvent elutions (20%, then 30% EtOAc / Hexane) to provide 2-({5,6-dichloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)-*N*-ethylacetamide as a mixture with something as yet uncharacterised as a white solid (20.9 mg, 0.068 mmol, 64%).

M.P. 157.3 – 160.0 °C.

^{1}H NMR (400 MHz, DMSO-*d$_6$*) *δ* 14.16 (br s, 1H), 8.72 (br s, 1H), 4.28 (s, 2H), 4.15 (q, *J* = 7.1 Hz, 2H), 1.20 (t, *J* = 7.1 Hz, 3H).

^{13}C NMR (101 MHz, DMSO-*d$_6$*) *δ* 168.3, 62.0, 33.5, 14.9, 14.5 (4 out of 7 carbon resonances found).

IR (cm^{-1}): 3225 (NH), 1725 (C=O)

MS ES+ m/z calcd for C$_9$H$_9$^{35}Cl$_2$N$_5$OS (M+H)$^+$: 304.9668, found (0.9836 average loss): 303.9827.

4-((3-amino-5,6-dichloro-pyrazin-2-yl)amino)-4-oxobutanoic acid
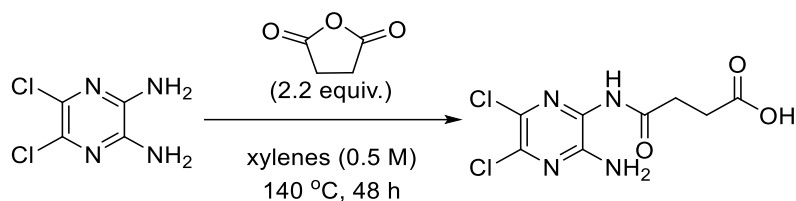
To an 8 mL vial was added 2,3-diamino-5,6-dichloropyrazine (28 mg, 0.156 mmol), succinic anhydride (20 mg, 1.2 equiv.) and mixed xylenes (0.51 mL, 0.305 M). The vial was sealed and heated to 140 °C for 18 h. The vial was allowed to cool and after TLC (EtOAc) indicated incomplete reaction, an additional equivalent of succinic anhydride was added, the vial re-sealed and heated to 140 °C for an additional 24 h. The reaction was allowed to cool to RT, solvent removed *in vacuo.* and dissolved in NaOH (2.0 M, 5 mL). The solution was extracted with Et$_2$O (3x 10 mL), the aqueous layer was acidified to pH approx. 2 with HCl (2.0 M) and re-extracted with Et$_2$O (3x 10 mL), these organic washing were collected and dried (MgSO$_4$) and solvent removed to provide an impure beige solid.

^{1}H NMR (400 MHz, C*D*$_3$CN) *δ* 11.95 (br s, 1H), 9.97 (br s, 1H), 6.17 (br s, 2H), 2.64 (ddd, *J* = 6.7, 5.5, 1.6 Hz, 2H), 2.58 (ddd, *J* = 7.8, 5.5, 1.6 Hz, 2H).

Heterocyclic variation, core preparations:

2,3-diaminopyrazine
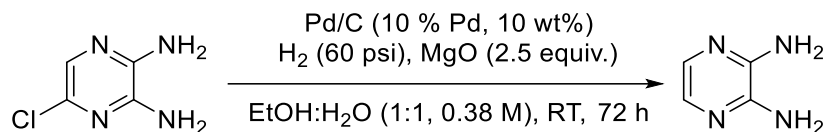
To a pressure vessel was added 2,3-diamino-5-chloropyrazine (1.0 g, 6.97 mmol), a stirrer bar, palladium on carbon (0.1 g, 10 wt. equiv., 10% Pd, un-activated), magnesium oxide (0.697 g, 17.3 mmol, 2.50 equiv., 325 mesh) and an argon-degassed ethanol : water mixture (1:1, 18.0 mL, 0.38 M). The autoclave was sealed, and the mixture degassed twice with H_2 at 40 psi, and once at 60 psi. The mixture was then stirred at 60 psi (dynamic pressure) for 72 h at RT. The vessel was depressurised, and the contents filtered through a pad of Celite®, washing with acetone. The filtrate was then dried (MgSO$_4$) and solvent removed *in vacuo.* to provide a crude mixture as a light-beige powder (approx. 34% conversion, determined by ^{1}H NMR). Four of these reactions of this type were performed with varying degrees of conversion (13 – 34%, three at 24 h, one at 72 h). These were combined and purified in a single batch of approx. 3.26 g crude, using column chromatography (95:5 DCM : MeOH) to provide 2,3-diaminopyrazine (810 mg, 7.35 mmol, 33%) as a dark orange crystalline powder.

M.P. (205.6 °C lit.), 204.8 – 206.4 °C.

^{1}H NMR (400 MHz, DMSO-d_6) δ 7.13 (s, 2H), 5.83 (br s, 4H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 144.0, 129.0.

IR (cm^{-1}) 3310 (NH), 3130 (NH), 1644, 1598.

Procedure adapted from the literature.[200]

4,5-diamino-6-chloropyrimidine
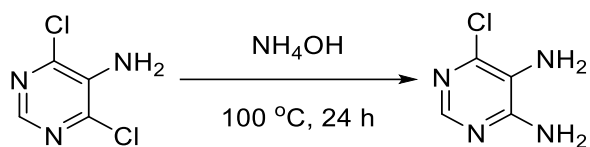
To four 30 mL microwave vials was added 5-amino-4,6-dichloropyrimidine (1.0 g, 6.09 mmol, 0.25 g per vial), a stirrer bar and ammonium hydroxide solution (25% soln., 12.5 mL per vial, 0.12 M). The vials were sealed and heated to 100 °C in a sand bath over 24 h. The vials were allowed to cool to RT and the contents transferred to a separatory funnel, washing with EtOAc. The organic and aqueous layers were separated, and the aqueous layer extracted with EtOAc (6x 20 mL). The organic layers were collected and dried (MgSO$_4$) to provide 4,5-diamino-6-chloropyrimidine (762 mg, 5.27 mmol, 87%) as an off-white solid. If the crude was impure, it was triturated with DCM to provide 4,5-diamino-6-chloropyrimidine as an off-white solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 7.63 (s, 1H), 6.72 (br s, 2H), 4.94 (br s, 2H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 153.5, 145.8, 137.5, 123.1.

M.p. 253.6 – 255.6 °C (deg.).

IR (cm^{-1}): 3331 (NH), 3269 (NH), 1671 (NH bend).

Followed literature procedure.[216]

Six-membered ring variations: Thione formation

General procedure 4.3:

To a 50 mL RBF was added the diamino-heterocycle, a stirrer bar and TCDI (1.0 equiv.). Dioxane (0.2 M) was added, and the mixture heated to 80 °C for 18 h. Afterwards, TLC indicated incomplete reaction and additional TCDI (0.408 g, 0.5 equiv.) was added, and the mixture further heated at 80 °C for 6 h. After TLC indicated complete reaction, allowed to cool to RT and dioxane removed under reduced pressure. The residue was taken up in EtOAc and HCl (1.0 M) (10 mL each). The layers separated and the aqueous layer extracted with EtOAc (3x 20 mL). The organic layers were collected and dried (MgSO₄) and solvent removed to provide crude material that is further purified on an individual basis.

Imidazo[4,5-*b*]pyridine-2(1,3*H*)-thione
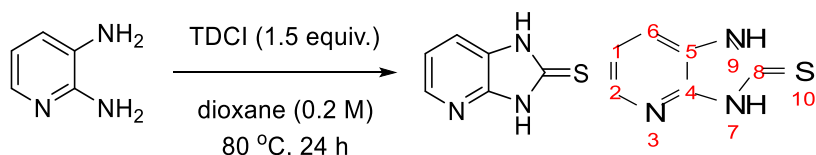
Synthesised following general procedure 4.3. 2,3-diaminopyridine (0.5 g, 4.58 mmol), total TCDI (1.22 g, 1.5 equiv.). Recrystallised from EtOH : H₂O (9:1) to provide light-beige crystals of imidazo[4,5-*b*]pyridine-2(1,3*H*)-thione (261 mg, 1.73 mmol, 37%).

M.P. (322 – 324 °C lit.), 321.8 – 323.8 °C.

^{1}H NMR (400 MHz, DMSO-d_6) δ 13.11 (br s, 1H, **7**), 12.70 (br s, 1H, **9**), 8.10 (dd, $J = 5.0$, 1.4 Hz, 1H, **2**), 7.47 (dd, $J = 7.9$, 1.4 Hz, 1H, **6**), 7.12 (dd, $J = 7.9$, 5.0 Hz, 1H, **1**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 169.8, **8**, 146.5, **4**, 142.3, **2**, 125.4, **5**, 118.1, **1**, 116.2, **6**.

IR (cm^{-1}): 3141 (NH), 3065 (NH), 2572 (SH), 1614 (NH bend).

MS ES+ m/z calcd for C₆H₅N₃S (M+H)$^{+}$: 152.0282, found: 152.0306.

Imidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione
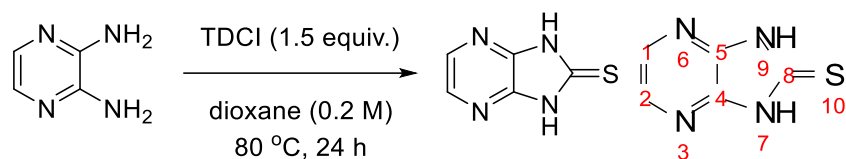
Synthesised following general procedure 4.3. 2,3-diaminopyrazine (0.5 g, 4.08 mmol), total TCDI (1.213 g, 1.5 equiv.). Recrystallised from EtOH : H_2O (9:1) to provide orange crystals of imidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (314 mg, 2.06 mmol, 50%).

M.P. (326.5 – 329 °C lit.), 322.5 – 327.8 °C.

^{1}H NMR (400 MHz, DMSO-d_6) δ 13.44 (br s, 2H, **7**, **9**), 8.07 (s, 2H, **1**, **2**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 172.3 , **8,** 141.0 , **4**, **5**, 136.6, **1**, **2**.

IR (cm^{-1}): 3388 (NH), 2743 (SH), 1621 (NH bend).

MS ES+ m/z calcd for $C_5H_4N_4S$ (M+H)$^+$: 153.0235, found: 153.0235.

Imidazo[4,5-*d*]pyrimidine-8(7,9*H*)-thione
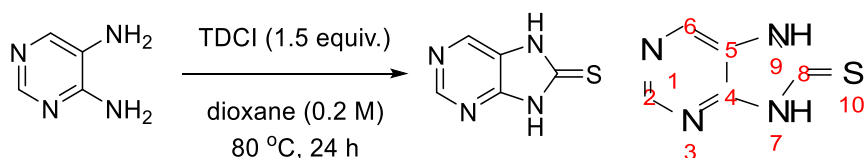
Synthesised following general procedure 4.3. 4,5-diaminopyrimidine (0.25 g, 2.27 mmol), total TCDI (0.606 g, 1.5 equiv.). Recrystallised from H_2O to provide an orange powder of imidazo[4,5-*d*]pyrimidine-8(7,9*H*)-thione (77.0 mg, 0.5 mmol, 22%).

M.P. (312 – 318 °C lit.), 310.5 – 314.4 °C (dec.).

^{1}H NMR (400 MHz, DMSO-d_6) δ 13.68 (br s, 1H, **7**), 13.15 (br s, 1H, **9**), 8.75 (s, 1H, **2**), 8.50 (s, 1H, **6**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 172.0, 152.4, 151.2, 133.4, 125.0.

IR (cm^{-1}): 3429 (NH), 2763 (SH), 1614 (NH bend).

MS ES+ m/z calcd for $C_5H_4N_4S$ (M+H)$^+$: 153.0235, found: 153.0237.

6-Chloroimidazo[4,5-*d*]pyrimidine-8(7,9*H*)-thione
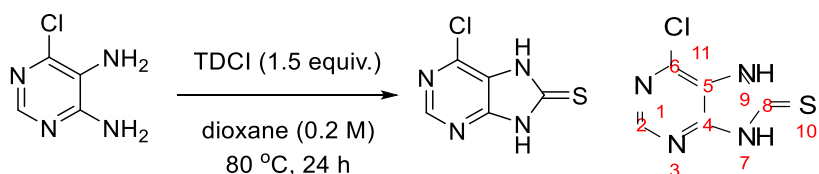
Synthesised following general procedure 4.3. 4,5-diamino-6-chloropyrimidine (0.5 g, 3.45 mmol), total TCDI (0.616 g, 1.5 equiv.). Dioxane (17.25 mL, 0.2 M). Provided 6-chloroimidazo[4,5-*d*]pyrimidine-8(7,9*H*)-thione as a yellow solid (173 mg, 0.92 mmol, 27%)

M.P. 205.2 °C (dec.).

^{1}H NMR (400 MHz, DMSO-d_6) *δ* 13.71 (br s, 2H, **7**, **9**), 8.52 (s, 1H, **2**).

^{13}C NMR (101 MHz, DMSO-d_6) *δ* 172.0, 152.9, 151.5, 136.0, 123.0.

IR (cm^{-1}): 2782 (SH), 1602 (NH bend).

MS ES+ m/z calcd for $C_5H_3^{37}ClN_4S$ (M+H)$^+$: 188.9815, found: 188.9814.

Six-membered ring variations: *S*-alkylation
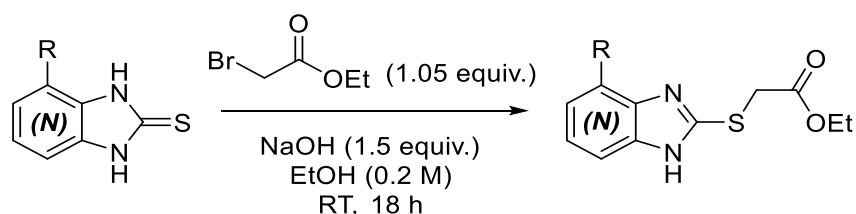
General procedure 4.4:

To an 8 mL vial was added imidazo-2-thione derivative and the respective amount of sodium hydroxide (1.5 equiv.). EtOH (0.2 M) was added *via* syringe with stirring. Ethyl bromoacetate (1.05 equiv.) was added *via* microsyringe. The vial was sealed with a septum and an exit needle. Left to stir for 18 h at RT. Afterwards, the ethanol was removed *in vacuo.* and the residue was taken up in sat. soln. NH$_4$Cl and EtOAc (*ca.* 5 mL each). Transferred to a separatory funnel and the layers separated. The aqueous layer was extracted with EtOAc (2x 10 mL), and organic layers collected and dried (MgSO$_4$). Solvent removed under reduced pressure to provide a residue, further purified on an individual basis.

Adapted from literature procedure.[167]

Ethyl 2-({5-chloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetate
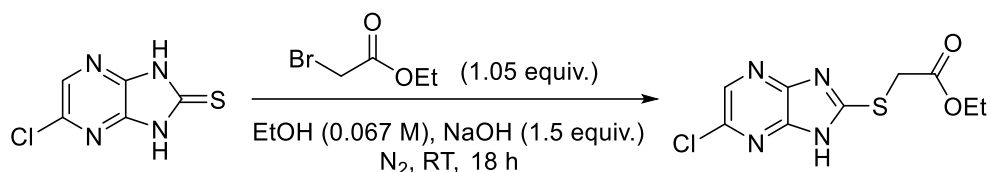
To a 50 mL RBF was added 5-chloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (442 mg, 2.37 mmol), sodium hydroxide (0.14 g, 1.5 equiv.) and ethanol (35 mL, 0.067 M). To the stirred solution was added ethyl bromoacetate (0.275 mL, 1.05 equiv.) and the reaction left to stir for 18 h. The ethanol was removed *in vacuo.* Sat. soln. of NH$_4$Cl and EtOAc were added to dissolve the solid residue (10 mL each), and this was transferred to a separatory funnel, separated and the aqueous layer extracted with EtOAc (3x 30 mL). The organic layers were collected and dried, loaded onto silica and purified by column chromatography eluting with 8:2 -> 7:3 Pet.E : EtOAc to provide ethyl 2-({5-chloro-1*H*-imidazo[4,5-*b*]pyrazine-2-yl}sulfanyl)acetate as an off-white powder. A sample for biological testing was prepared by triturating with *n*-hexane after recrystallisation from ethanol. (170 mg, 0.62 mmol, 26%).

^{1}H NMR (500 MHz, C*D*$_3$CN) *δ* 11.15 (br s, 1H), 8.23 (s, 1H), 4.19 (q, *J* = 7.1 Hz, 2H), 4.17 (s, 2H), 1.24 (t, *J* = 7.1 Hz, 3H).

^{13}C NMR (126 MHz, C*D*$_3$CN) *δ* 169.0, 159.4, 142.6, 137.2, 62.8, 34.2, 14.4 (7 out of 9 carbon resonances observed).

M.P. 118.7 – 119.2 °C.

IR (cm^{-1}): 1737 (C=O).

MS ES+ m/z calcd for C$_9$H$_9$^{35}ClN$_4$O$_2$S (M+H)$^+$: 273.0213, found: 273.0213.

Adapted from literature procedure.[167]

Ethyl 2-({1*H*-imidazo[4,5-*b*]pyrazine}sulfanyl-2-yl)acetate
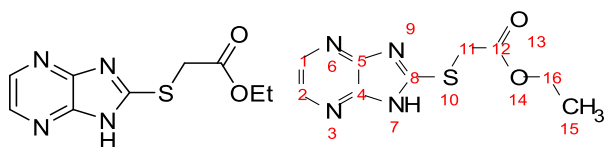
Prepared according to general procedure 4.4. Imidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (0.1 g, 0.66 mmol), ethyl bromoacetate (76.5 μL), sodium hydroxide (39.4 mg), EtOH (3.3 mL). Purified by column chromatography using 7:3 -> 5:5 -> 4:6 Pet.E : EtOAc to provide ethyl 2-({1*H*-imidazo[4,5-*b*]pyrazine}sulfanyl-2-yl)acetate a white solid. Recrystallised from EtOH with a hot filtration (41.9 mg, 0.18 mmol, 27%).

M.P.: 177.4 – 175.2 °C.

^{1}H NMR (400 MHz, DMSO-d_6) $δ$ 13.68 (br s, 1H, **7**), 8.25 (s, 2H, **1**, **2**), 4.29 (s, 2H, **11**), 4.15 (q, $J = 7.1$ Hz, 2H, **16**), 1.20 (t, $J = 7.1$ Hz, 3H, **15**).

^{13}C NMR (101 MHz, DMSO-d_6) $δ$ 168.1, **12**, 157.4, **8**, 146.0, **4**, **5**, 137.4, **1**, **2**, 61.4, **16**, 32.8, **11**, 14.0, **13**.

IR (cm^{-1}): 1733 (C=O).

MS ES+ m/z calcd for $C_9H_{10}N_4O_2S$ (M+H)$^+$: 239.0603, found: 239.0628.

Ethyl 2-({1*H*-imidazo[4,5-*b*]pyridine}sulfan-2-yl)acetate
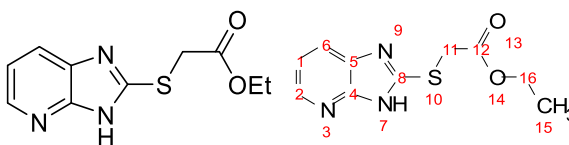
Prepared according to general procedure 4.4. Imidazo[4,5-*b*]pyridine-2(1,3*H*)-thione (0.1 g, 0.66 mmol), ethyl bromoacetate (77 µL), sodium hydroxide (39.6 mg), EtOH (3.3 mL). Purified by column chromatography using 5:5 -> 6:4 EtOAc : Pet.E. Recrystallised from EtOH to provide ethyl 2-({1*H*-imidazo[4,5-*b*]pyridine}sulfan-2-yl)acetate a white solid (51.4 mg, 0.22 mmol, 33%).

M.P. 157.6 – 160.2 °C.

^{1}H NMR (400 MHz, DMSO-d_6) δ 13.10 (br s, 1H, **7**), 8.20 (dd, $J = 4.9$, 1.5 Hz, 1H, **2**), 7.82 (dd, $J = 7.9$, 1.5 Hz, 1H, **6**), 7.15 (dd, $J = 7.9$, 4.9 Hz, 1H, **1**), 4.24 (s, 2H, **11**), 4.13 (q, $J = 7.1$ Hz, 2H, **16**), 1.18 (t, $J = 7.1$ Hz, 3H, **15**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 168.4, **12**, 152.4, **8**, 142.5, **2**, 117.4, **1**, 61.2, **16**, 32.9, **11**, 14.0, **15**.

IR (cm^{-1}): 3003 (NH), 1744 (C=O).

MS ES+ m/z calcd for $C_{10}H_{11}N_3O_2S$ (M+H)$^+$: 239.0676, found: 239.0675.

Ethyl 2-({9*H*-imidazo[4,5-*d*]pyrimidine}sulfan-8-yl)acetate
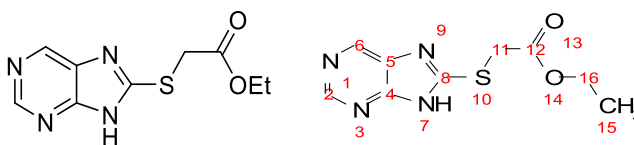
Prepared according to general procedure 4.4. Imidazo[4,5-*b*]pyrimidine-2(1,3*H*)-thione (0.075 g, 0.474 mmol), ethyl bromoacetate (55.2 µL), sodium hydroxide (28 mg), EtOH (2.4 mL). Purified by column chromatography using 5:5 -> 4:6 Pet.E : EtOAc to provide ethyl 2-({9*H*-imidazo[4,5-*d*]pyrimidine}sulfan-8-yl)acetate as a white solid (41.9 mg, 0.18 mmol, 37%).

M.P. 140.3 – 141.0 °C.

^{1}H NMR (400 MHz, DMSO-d_6) *δ* 13.62 (br s, 1H, **7**), 8.88 (s, 1H, **2**), 8.78 (s, 1H, **6**), 4.27 (s, 2H, **11**), 4.14 (q, *J* = 7.1 Hz, 2H, **16**), 1.19 (t, *J* = 7.1 Hz, 3H, **15**).

^{13}C NMR (101 MHz, DMSO-d_6) *δ* 168.0, **12**, 156.2, **8**, 150.8, **6**, 141.4, **2**, 131.97*, **4**, **5**, 61.1, **16**, 32.8, **11**, 13.8, **15**.

IR (cm^{-1}): 1742 (C=O).

MS ES+ m/z calcd for $C_9H_{10}N_4O_2S$ (M+H)$^+$: 240.0626, found: 240.0632.

HMBC correlations of the purine aromatic protons with a broad singlet at 131.97 ppm, which could be carbons 4 and 5.

Ethyl 2-({6-chloro-9*H*-imidazo[4,5-*d*]pyrimidine}sulfan-8-yl)acetate
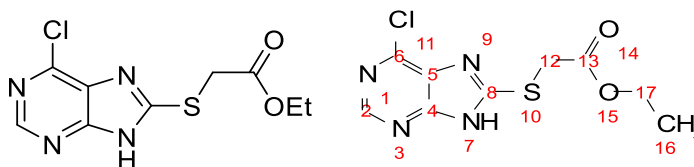
Prepared according to general procedure 4.4. 6-chloroimidazo[4,5-*d*]pyrimidine-8(7,9*H*)-thione (0.1 g, 0.535 mmol), ethyl bromoacetate (62.3 µL), sodium hydroxide (32 mg). EtOH (2.7 mL). Purified by column chromatography using 5:5 Pet. E : EtOAc. Recrystallised from EtOH to provide ethyl 2-({6-chloro-9*H*-imidazo[4,5-*d*]pyrimidine}sulfan-8-yl)acetate a white solid (35.5 mg, 0.13 mmol, 24%).

M.P. 151.5 – 153.3 °C.

^{1}H NMR (400 MHz, DMSO-d_6) δ 14.09 (br s, 1H, **7**), 8.61 (s, 1H, **2**), 4.29 (s, 2H, **12**), 4.15 (q, J = 7.1 Hz, 2H, **17**), 1.20 (t, J = 7.1 Hz, 3H, **16**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 168.5, **13**, 155.5, **8**, 151.2, **2**, 61.8, **17**, 33.5, **12**, 14.5, **16** (6 out of 9 carbon resonances observed).

IR (cm^{-1}): 3062 (NH), 1727 (C=O).

MS ES+ m/z calcd for $C_9H_9{}^{35}ClN_4O_2S$ (M+H)$^+$: 274.0237, found: 274.0244.

Imidazo[4,5-*c*]pyridine-2(1,3*H*)-thione
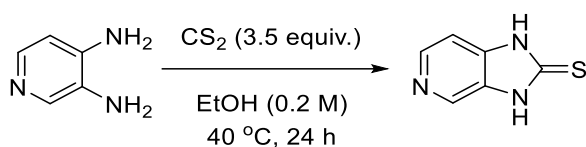
To a 50 mL RBF was added 3,4-diaminopyridine (0.25 g, 2.29 mmol) and a stirrer bar. To this was added EtOH (11.5 mL, 0.2 M) and stirred to dissolve. To the stirred solution was added CS$_2$ (5 mL, 3.5 equiv.) and the mixture heated with stirring to 40 °C. The mixture was allowed to cool to RT then 0 °C. The solid precipitate was collected and filtered to provide imidazo[4,5-*c*]pyridine-2(1,3*H*)-thione an off-white solid.

Taken through as crude material (ca. 76% conversion as indicated by ^{1}H NMR.

^{1}H NMR (400 MHz, DMSO-d_6) δ 12.80 (br s, 2H), 8.37 (s, 1H), 8.24 (d, J = 5.2 Hz, 1H), 7.17 (d, J = 5.2 Hz, 1H).

Adapted from literature procedure.[189]

4.91
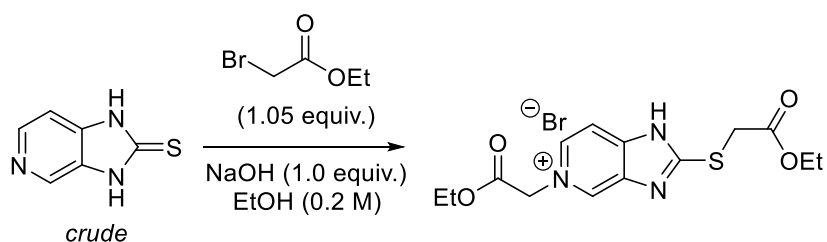
To the crude imidazo[4,5-*c*]pyridine (0.392 g) was added EtOH (13 mL), sodium hydroxide (1.0 equiv. 130 mg, accounting for the 76% conversion) and ethyl bromoacetate (1.05 equiv., 0.187 mL, accounting for the 76% conversion). Stirred at RT for 18 h. The solvent was removed, and residue taken up in NH$_4$Cl sat. soln. and EtOAc (*ca.* 5 mL each) the layers were separated and the aqueous extracted with EtOAc (2x 10 mL). Organic layers collected and dried (MgSO$_4$) and solvent removed to provide an oil. Purified using column chromatography (5% MeOH / DCM) to isolate a white solid which was recrystallised from EtOH and triturated using EtOH to provide the apparent hydrobromide (or chloride) salts.

M.P. 146.8 – 147.3 °C.

^{1}H NMR (400 MHz, DMSO-d_6) δ 14.98 (br s, 1H), 9.34 (s, 1H), 8.60 (dt, J = 6.9, 1.3 Hz, 1H), 8.10 (d, J = 6.8 Hz, 1H), 5.61 (s, 3H), 4.35 (s, 3H), 4.23 (qd, J = 7.1, 1.0 Hz, 3H), 4.15 (qd, J = 7.1, 1.0 Hz, 3H), 1.25 (td, J = 7.1, 1.0 Hz, 4H), 1.20 (td, J = 7.1, 1.0 Hz, 4H).

¹³C NMR (101 MHz, DMSO-d₆) δ 167.9, 167.1, 138.2, 62.1, 61.5, 59.4, 33.5, 14.00, 13.96 (9 out of 14 carbon resonances observed).

IR (cm⁻¹): 1730 (C=O).

Adapted from literature procedure.[167]

Ethyl 2-[({imidazo[4,5-*c*]pyridine}sulfan-2-yl-5*N*-(ethyl-2-acetate))]acetate
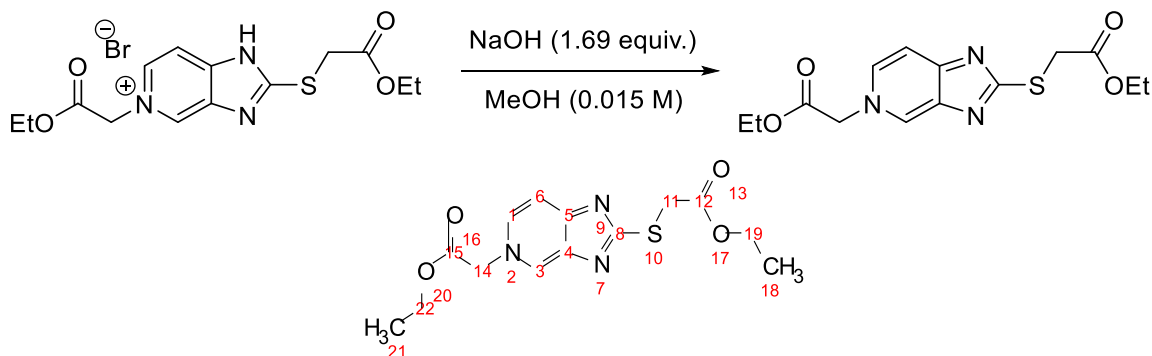
Crude material from the EtOH trituration (30 mg) was dissolved in MeOH (*ca.* 5 mL) and sodium hydroxide was added (5 mg). Left to stir for 18 h at RT. The mixture was diluted with H₂O and EtOAc, transferred to a separatory funnel and extracted with EtOAc (3x 10 mL). Organic layers were collected and dried (MgSO₄). Crude material was purified by column chromatography (3 -> 4% MeOH / DCM) to provide product as a white solid (24.5 mg, 0.075 mmol).

M.P. 146.8 – 147.3 °C.

¹H NMR (400 MHz, DMSO-d₆) δ 8.68 (d, *J* = 1.5 Hz, 1H, **3**), 7.99 (dd, *J* = 6.9, 1.5 Hz, 1H, **1**), 7.54 (d, *J* = 6.9 Hz, 1H, **6**), 5.38 (s, 2H, **14**), 4.20 (q, *J* = 7.1 Hz, 2H, **22**), 4.15 (s, 2H, **11**), 4.12 (q, *J* = 7.1 Hz, 2H, **19**), 1.23 (t, *J* = 7.1 Hz, 2H, **21**), 1.19 (t, *J* = 7.1 Hz, 3H, **18**).

¹³C NMR (101 MHz, DMSO-d₆) δ 172.1, **8**, 169.5, **12**, 168.0, **15**, 156.4, **4**, 144.9, **5**, 132.8, **1**, 130.2, **3**, 109.9, **6**, 61.9, **22**, 61.1, **19**, 58.4, **14**, 33.5, **11**, 14.3, **21**, 14.2, **18**.

IR (cm⁻¹): 1730 (C=O).

MS ES+ m/z calcd for C₁₄H₁₇N₃O₄S (M+H)⁺: 325.1046, found: 325.1067.

2-amino-4-chloropyrazine, 2-amino-3-chloropyrazine & 2-amino-3,5-dichloropyrazine[182]
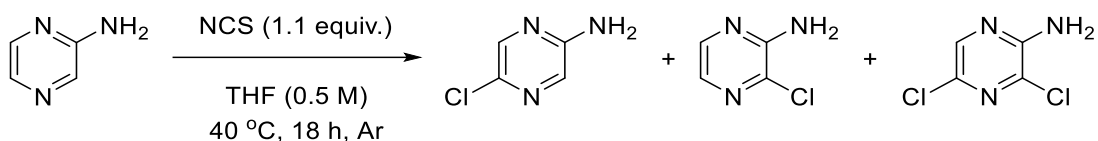
To an oven-dried 50 mL RBF was added 2-aminopyrazine (1.0 g, 10.51 mmol) and a stirrer bar. This was purged with argon and anhydrous THF (21 mL, 0.5 M) was added *via* syringe. To the stirred solution NCS (1.544 g, 1.1 equiv.) was then added in one portion with stirring until the majority dissolved. The reaction was then heated to 40 °C for 18 h. The reaction was allowed to cool to RT and filtered through a pad of Celite®, washing with EtOAc. The filtrate was then washed with $NaHCO_3$ sat. soln. (3x 30 mL), dried (Na_2SO_4) and dry loaded onto silica gel. Purified by column chromatography (9:1 -> 8:2 -> 7:3 Hexane : EtOAc) to yield:

2-amino-5-chloropyrazine as a yellow powder (518 mg, 3.99 mmol, 38%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 8.01 (d, $J = 1.5$ Hz, 1H), 7.78 (d, $J = 1.5$ Hz, 1H), 4.58 (br s, 2H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 153.2, 141.4, 137.8, 131.0

2-amino-3-chloropyrazine as a light-yellow powder (148 mg, 1.14 mmol, 11%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 7.93 (d, $J = 2.7$ Hz, 1H), 7.72 (d, $J = 2.7$ Hz, 1H), 5.09 (br s, 2H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 151.7, 140.6, 134.4, 133.1.

2-amino-3,5-dichloropyrazine as white fluffy crystals (133 mg, 0.81 mmol, 15%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 7.97 (s, 1H), 5.01 (br s, 2H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 150.5, 140.2, 134.8, 131.6.

IR (cm^{-1}) 3435, 3291, 3167, 1620, 1562.

Ethyl imidazo[1,2-a]pyrazine-2-acetate
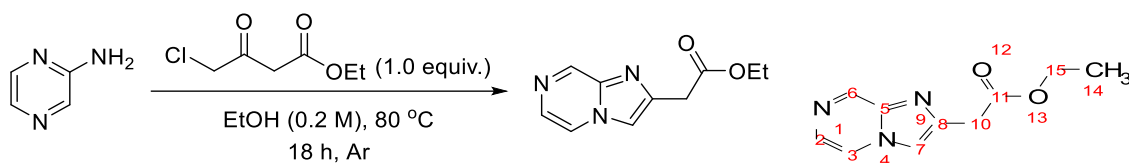
To a 100 mL RBF was added 2-aminopyrazine (500 mg, 5.25 mmol) and EtOH (26.5 mL, over mol. sieves, 0.2 M). to the stirred solution, ethyl 4-chloroacetoacetate (0.70 mL, 1.0 equiv.) was then added dropwise *via* syringe and the mixture heated to reflux for 18 h. The EtOH was removed under reduced pressure and to the residue was added NaHCO$_3$ sat. soln. (20 mL) and this was extracted with CHCl$_3$ (4x 20 mL). The organic layers were collected and dried (MgSO$_4$). Solvent removed under reduced pressure and the residue was subjected to sequential column chromatography (1% EtOH in CHCl$_3$ -> 3% EtOH in CHCl$_3$, then 7:2.5:0.5 EtOAc : Hex : Tol.) and the product (0.1349 g) was recrystallised from hexane to provide ethyl imidazo[1,2-a]pyrazine-2-acetate as white crystals (49.4 mg, 0.24 mmol, 5%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 9.03 (d, J = 0.9 Hz, 1H, **6**), 8.02 (dd, J = 4.6, 1.5 Hz, 1H, **2**), 7.86 (d, J = 4.6 Hz, 1H, **3**), 7.73 (d, J = 0.6 Hz, 1H, **7**), 4.21 (q, J = 7.1 Hz, 2H, **15**), 3.92 (d, J = 0.7 Hz, 2H, **10**), 1.29 (t, J = 7.1 Hz, 3H, **14**).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 170.5, **11**, 143.5, **6**, 142.3, **5**, 140.3, **8**, 129.7, **3**, 118.8, **2**, 112.3, **7**, 61.4, **15**, 35.3, **10**, 14.3, **14**.

IR (cm^{-1}): 1714 (C=O).

Adapted from literature procedure.[217]

7-methyl-5*H*-pyrrolo[2,3-*b*]pyrazine-6-carboxlic acid ethyl ester
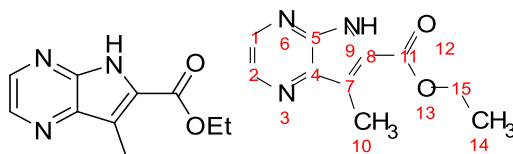
Second column of the above reaction isolated ethyl 7-methyl-5*H*-pyrrolo[2,3-*b*]pyrazine-6-acetate (6.9 mg, impure).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 10.29 (br s, 1H, **9**), 8.61 (d, J = 2.7 Hz, 1H, **2**), 8.21 (d, J = 2.7 Hz, 1H, **1**), 4.49 (q, J = 7.1 Hz, 2H, **15**), 2.90 (s, 3H, **10**), 1.45 (t, J = 7.1 Hz, 3H, **14**).

Ethyl 6-chloroimidazo[1,2-*a*]pyrazine-2-acetate
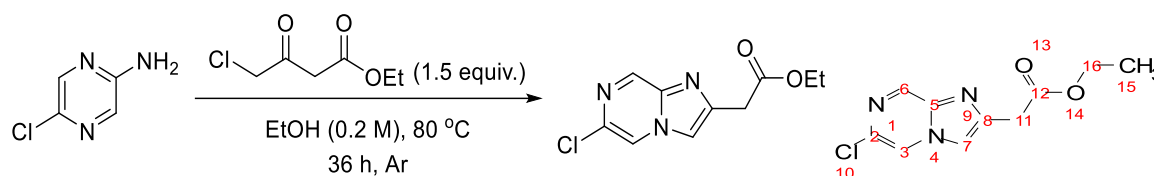
To an oven-dried 3-necked 100 mL RBF was attached a glass stopper, a reflux condenser and a suba seal. To this was added a stirrer bar, 2-amino-5-chloropyrazine (0.305 g, 2.35 mmol) and the vessel was placed under an atmosphere of argon. EtOH (over mol. sieves, 0.2 M, 12.0 mL) was added *via* syringe. To this stirred solution was added ethyl 4-chloroacetoacetate (0.47 mL, 1.5 equiv.) and the mixture heated to reflux for 36 h. The mixture was allowed to cool to RT and the EtOH removed *in vacuo.* The residue was taken up in NaHCO$_3$ sat. soln. (25 mL) and extracted with chloroform (3x 20 mL). The organic layers were collected, washed with water (20 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* The residue was then loaded onto silica and purified twice *via* column chromatography (1% MeOH in DCM) and (3.5:6.0:0.5 EtOAc : Hex. : Tol.), and then triturated with cold Et$_2$O to afford ethyl 6-chloroimidazo[1,2-*a*]pyrazine-2-aceate as a beige crystalline solid (42 mg, 0.17 mmol, 7%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 8.85 (dd, J = 1.4, 0.7 Hz, 1H, **3**), 8.13 (d, J = 1.4 Hz, 1H, **6**), 7.77 (d, J = 0.7 Hz, 1H, **7**), 4.23 (q, J = 7.1 Hz, 2H, **16**), 3.93 (d, J = 0.7 Hz, 2H, **11**), 1.30 (t, J = 7.1 Hz, 3H, **15**).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 170.0, **12**, 143.8, **5**, 141.7, **3**, 139.0, **8**, 135.2, **2**, 116.5, **6**, 113.0, **7**, 61.6, **16**, 35.1, **11**, 14.3, **15**.

IR (cm^{-1}): 1715 (C=O)

M.P. 100.2 – 100.4 °C.

MS ES+ Calcd for C$_{10}$H$_{10}$^{35}ClN$_3$O$_2$ (M+H)$^+$: 240.0540, found: 240.0534.

Adapted from literature procedure.[217]

Ethyl 2-aceto-dihydroimidazo[1,2-*a*]pyrazin-8(7*H*)-one
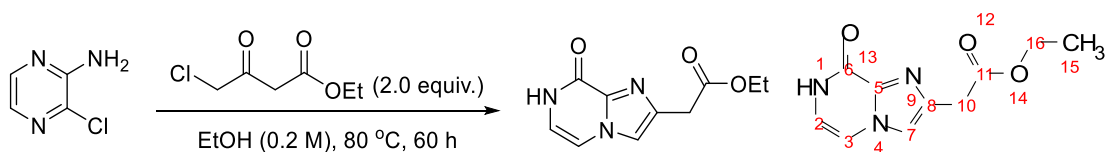
To a 10 mL RBF was added 2-amino-3-chloropyrazine (105.8 mg, 0.837 mmol) and EtOH (0.2 M, 4.2 mL). Ethyl 4-chloroacetoacetate (0.226 mL, 2 equiv.) was added and the mixture was heated to reflux for 60 h. The mixture was allowed to cool to RT and the solvent removed *in vacuo.* The residue was redissolved in NaHCO$_3$ sat. soln. (20 mL) and extracted with DCM (3x 20 mL). The organic washings were collected and washed with brine (20 mL) dried (Na$_2$SO$_4$), and solvent removed *in vacuo.* Purified *via* column chromatography (1% MeOH in DCM) to yield ethyl 2-acetoimidazo[1,2-*a*]pyrazin-8(7*H*)-one (78 mg, 0.35 mmol, 42%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 11.31 (br s, 1H, **1**), 7.50 (d, J = 0.8 Hz, 1H, **7**), 7.13 (d, J = 5.6 Hz, 1H, **3**), 6.86 (s, 1H, **2**), 4.20 (q, J = 7.1 Hz, 2H, **16**), 3.87 (d, J = 0.8 Hz, 2H, **10**), 1.29 (t, J = 7.1 Hz, 3H, **15**).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 170.7, 155.7, 139.9, 136.9, 116.7, 115.6, 107.8, 61.3, 34.9, 14.3.

IR (cm^{-1}): 3043 (NH), 1723 (C=O).

M.P. 190.3 – 195.0 °C.

MS ES+ calcd. For C$_{10}$H$_{11}$N$_3$O$_3$ (M+H)$^+$: 223.0907, found: 223.0904.

Adapted from literature procedure.[217]

Ethyl 6,8-Dichloroimidazo[1,2-*a*]pyrazine-2-acetate
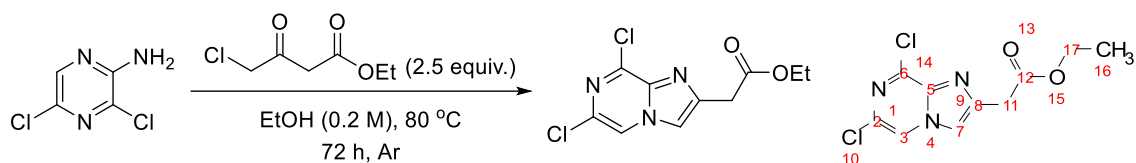
To a 50 mL RBF was added 2-amino-3,5-dichloropyrazine (0.19 g, 1.16 mmol) and EtOH (over mol. sieves, 6 mL, 0.2 M). To the stirred solution was added ethyl 4-chloroacetoacetate (0.235 mL, 1.5 equiv.) This was heated to reflux at 80 °C for 18 h. Afterwards, additional ethyl 4-chloroacetoacetate (0.16 mL, 1.0 equiv.) was added and reflux was continued for 48 h. After allowing the mixture to cool to RT, the solvent was removed *in vacuo.* and dissolved in NaHCO$_3$ sat. soln. (20 mL). This was extracted with DCM (3x 20 mL), the organic layers were collected, washed with brine (20 mL), dried (Na$_2$SO$_4$) and solvent removed *in vacuo.* the residue was purified *via* column chromatography (9:1 -> 7:3 Hexane : EtOAc) to yield ethyl 6,8-dichloroimidazo[1,2-*a*]pyrazine-2-acetate as a beige solid (132 mg, 0.48 mmol, 41%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 8.09 (s, 1H, **3**), 7.87 (t, J = 0.7 Hz, 1H, **7**), 4.22 (q, J = 7.1 Hz, 2H, **17**), 3.97 (d, J = 0.7 Hz, 2H, **11**), 1.30 (t, J = 7.1 Hz, 3H, **16**).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 170.0, **12**, 144.2, 141.7, 136.6, 132.4, 116.2, **3**, 115.1, **7**, 61.6, **17**, 35.2, **11**, 14.3, **16**.

IR (cm^{-1}): 1716 (C=O).

M.P. 96.8 – 97.1 °C.

MS ES+ Calcd for C$_{10}$H$_9$^{37}Cl$_2$N$_3$O$_2$ (M+H)$^+$: 278.0096, found: 278.0090.

Adapted from literature procedure.[217]

2-thiomethyl-4-aminopyrimidin-6(1*H*)-one
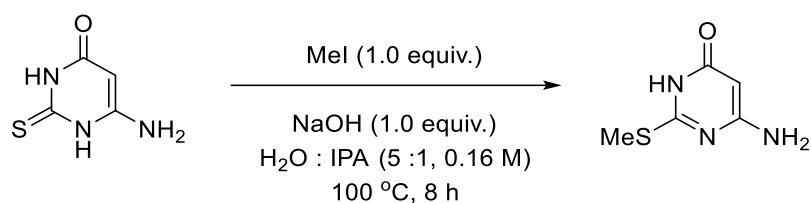
To a 250 mL RBF was added 2(1*H*)-thioxo-4(3*H*)-oxo-6-aminopyrimidine (2.0 g, 13.96 mmol) and sodium hydroxide (0.55 g, 1.0 equiv.). IPA : H$_2$O (1:5, 0.16 M, 87 mL) was added and the mixture heated to 100 ^{o}C with stirring. After full dissolution, iodomethane (0.896 mL, 1.0 equiv.) was added *via* syringe and the mixture left at reflux for 8 h. The mixture was cooled to RT and then cooled to 0 ^{o}C. After 10 min. the precipitate was filtered, collected, and dried under vacuum to provide 2-thiomethyl-4-aminopyrimidin-6(1*H*)-one as a white solid (1.99 g, 12.7 mmol, 91%).

^{1}H NMR (DMSO-d_6, 400 MHz) δ 11.47 (br s, 1H), 6.43 (br s, 2H), 4.89 (s, 1H), 2.41 (s, 3H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 164.3, 163.5, 162.9, 81.2, 12.6.

IR (cm^{-1}): 3462 (NH), 3209 (NH), 1596 (C=O amide).

M.P. (266 – 267 ^{o}C lit.) 266.7 – 268.9 ^{o}C (dec.).

Adapted from literature procedure.[218]

4-aminopyrimidin-6(1*H*)-one
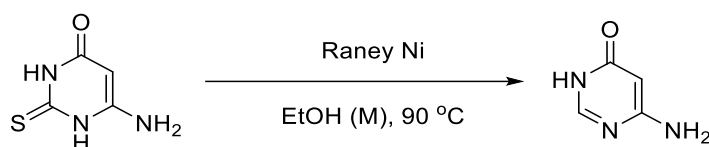
To a 3-necked 50 mL RBF was added 6-oxo-4-amino-2-mercaptopyrimidine (1.5 g, 0.01 mol) and ethanol (23 mL, 0.46 M) and heated to 80 °C. To this was added 4 glass pipettes of Raney nickel suspension (approx. 6.8 g total) and the mixture heated to 100 °C for 2 h. After TLC analysis another pipette's worth of Raney nickel solution was added and the mixture allowed to reflux for a further 1h. The mixture was filtered whilst hot on Celite®, washing with hot ethanol (100 mL). Filtrate was concentrated to provide an off-white solid of 4-aminopyrimidin-6(1*H*)-one (932 mg, 86.93 mmol, 85%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 11.40 (br s, 1H), 7.76 (s, 1H), 6.38 (br s, 2H), 4.96 (s, 1H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 164.6, 162.5, 150.1, 85.8. Followed literature procedure.[201]

2,4-diamino-6-methoxypyrimidine[189]
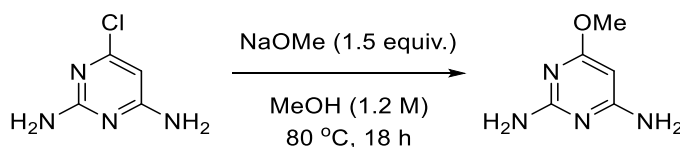
To a 50 mL RBF was added 2,4-diamino-6-chloropyrimidine (3.0 g, 0.2 mol), methanol (22.9 mL, 1.2 M). To the stirred solution was added sodium methoxide (1.68 g, 1.5 equiv.) and the mixture heated to reflux for 18 h. The mixture was allowed to cool to RT and neutralised with HCl (2 M) to pH = 6. The mixture was extracted with EtOAc (3x 10 mL), the organic layers were dried (MgSO$_4$), and solvent removed *in vacuo.* Purified by column chromatography (8 : 2 EtOAc : Pet. E) to provide 2,4-diamino-6-methoxypyrimidine as a white solid (950 mg, 6.77 mmol, 34%).

M.P. 166 – 169 °C (lit.), 165.7 – 167.1 °C.

^{1}H NMR (DMSO-d_6, 400 MHz) δ 5.99 (br s, 2H), 5.87 (br s, 2H), 5.03 (s, 1H), 3.67 (s, 3H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 170.4, 166.0, 163.0, 75.8, 52.4.

IR (cm^{-1}): 3486 (NH), 3321 (NH), 3134 (NH).

Followed literature procedure, data in-line with literature data.

Nitrosopyrimidine formation
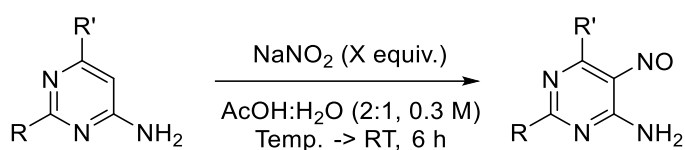
Crude ^{1}H NMR was used to confirm if the reaction had occurred before moving onto the next step.

2-methylthio-4-amino-5-nitrosopyrimidin-6(1*H*)-one
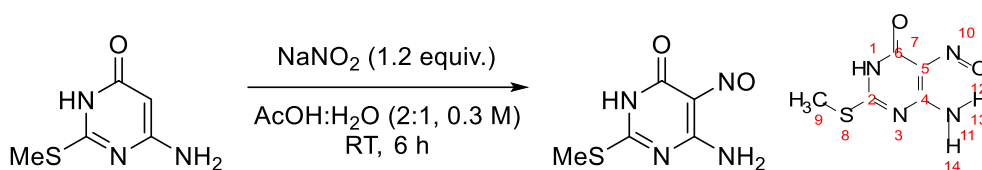
To a 25 mL RBF was added 2-thiomethyl-4-aminopyrimidin-6(1*H*)-one (700 mg, 4.45 mmol), a stirrer bar and AcOH:H$_2$O (2:1 v/v, 14.8mL 0.3 M). NaNO$_2$ solution (0.368 g, 1.2 equiv., in 2.4 mL H$_2$O) was added *via* a dropping funnel at RT and stirred for 6 h. The mixture was transferred to a fridge and left for 18 h. Filtration of the precipitated solid provided 2-methylthio-4-amino-5-nitrosopyrimidin-6(1*H*)-one as a purple solid (477 mg, 2.56 mmol, 58%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 12.71 (br s, 1H, **1**), 11.26 (s, 1H, **13**), 9.06 (s, 1H, **14**), 2.54 (s, 3H, **9**).

Adapted from literature procedure.[189]

2,4(1,3*H*)-oxo-5-nitroso-6-aminopyrimidine
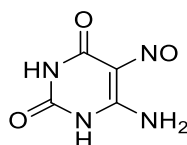
To a 50 mL RBF was added 2,4(1,3*H*)-dioxo-6-aminopyrimidine (1.40 g, 11.01 mmol) and AcOH:H$_2$O (36.7 mL), NaNO$_2$ (0.911 g, 1.2 equiv., in 6.1 mL H$_2$O) was added *via* dropping funnel at 50 °C and the reaction stirred for 6 h. Afterwards, once the reaction had cooled to RT the precipitate was filtered to provide 2,4-oxo-5-nitroso-6-aminopyrimidine as an orange solid (2.055 g).

No ^{1}H NMR peaks indicative of product by comparison with literature data.

4(3*H*)-oxo-5-nitroso-6-aminopyrimidine
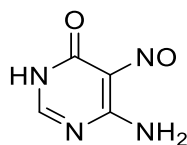
Reaction product degraded at room temperature after isolation of a blue solid.

2,4-diamino-5-nitroso-6-methoxypyrimidne
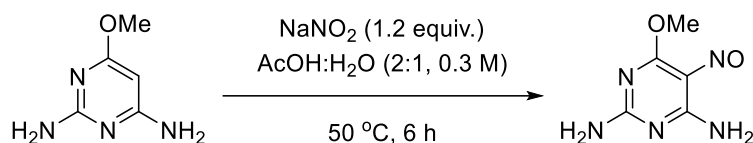
To a 50 mL RBF was added 2,4-diamino-6-methoxypyrimidine (1.40 g, 9.99 mmol) and AcOH:H_2O (33.3 mL), $NaNO_2$ (0.827 g, 1.2 equiv., in 5.5 mL H_2O) was then added *via* dropping funnel at 50 °C and allowed to react for 6 h. The precipitate was filtered to provide 2,4-diamino-5-nitroso-6-methoxypyrimidne as a violet solid (1.49 g, 8.84 mmol, 89%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 10.07 (br s, 1H), 8.00 (br s, 1H), 7.85 (br s, 1H), 7.79 (br s, 1H), 4.05 (s, 3H).

2(3*H*)-thioxo-4-amino-5-nitrosopyrimidin-6(1*H*)-one
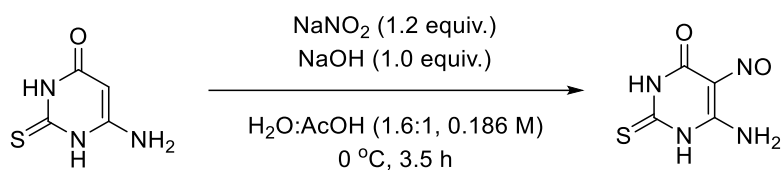
To a cooled (0 °C) 50 mL RBF containing 2(3*H*)-thioxo-4-amino-pyrimidin-6(1*H*)-one (700 mg, 4.89 mmol), NaOH (0.195 g, 1.0 equiv.) and $NaNO_2$ (0.404 g, 1.2 equiv.) in H_2O (16.3 mL) was added dropwise *via* a dropping funnel AcOH (10 mL). After addition the mixture was allowed to warm to RT with stirring for 3.5 h. The red precipitate was collected and dried under vacuum to provided 2(3*H*)-thioxo-4-amino-5-nitrosopyrimidin-6(1*H*)-one as a brick red solid (693 mg).

No ^{1}H NMR peaks indicative of product by comparison with literature data.

Adapted from literature procedure.[219]

General procedure 4.2: Telescoped nitroso reduction and thione formation:
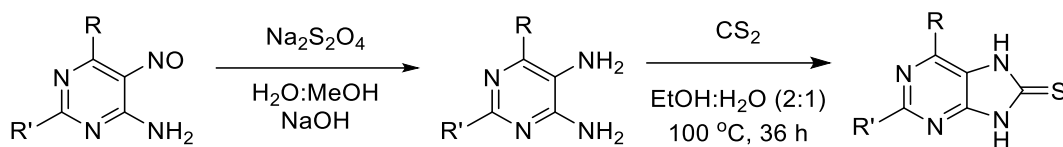
Nitroso reduction: To an RBF was added the crude 5-nitroso-6-aminopyrimidine, to which was added solvent (either H_2O 0.3 M, or H_2O:MeOH (1:1, 0.17 M)) and in some cases sodium hydroxide (3 equiv.). With rapid stirring was added $Na_2S_2O_4$ (1.5 equiv.) was added in one batch and stirred at temperature for 6 h. If a complete colour change did not occur after 2 h, more $Na_2S_2O_4$ (0.5 equiv.) was added and continued to react. The mixtures were placed in a fridge overnight. Afterwards, the solids were filtered, washed with water, and dried on high vacuum.

Thione formation: Depending on the amount of diaminopyrimidine crude, two solvent volumes were used, 10 mL (for less than 1.0 g) or 30 mL (for more than 1.0 g) (EtOH:H_2O, 2:1). CS_2 was added *via* syringe in either 1.5 mL or 3.0 mL for 10 mL and 30 mL reactions respectively. $NaHCO_3$ (1 g) was then added, and the mixtures were heated to reflux for 36 h. The reactions were allowed to cool, and solvent was removed *in vacuo.* and a small amount of water was added to the residue, after which the mixture was acidified with acetic acid. The precipitates were filtered and dried *in vacuo.*

Procedure adapted from the literature.[189]

2-thiomethyl-6-oxo(1*H*)purine-8(7,9*H*)-thione
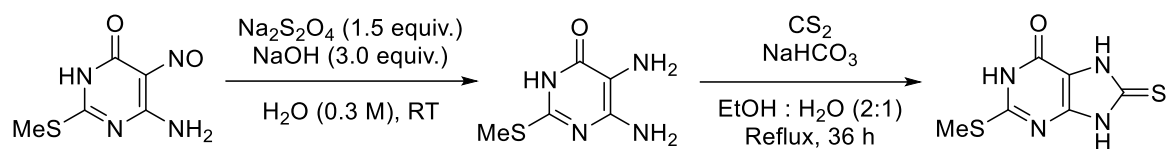
Followed general procedure 4.2. <u>Nitroso reduction:</u> Sodium hydroxide (0.30 g, 3.0 equiv.), H_2O as solvent (8.54 mL) $Na_2S_2O_4$ (1.5 equiv., 0.669 g). Reaction carried out at RT. The filtered material was washed with water and dried *in vacuo.* to provide the diamino compound as a beige solid.

<u>CS_2 Ring closure:</u> Provided 2-thiomethyl-6-oxo(5*H*)purine-8(7,9*H*)-thione as a white solid (148 mg, 0.694 mmol, 27% over two steps).

M.P. (275 °C decomp., lit.), 280 °C (decomp.).

^{1}H NMR (400 MHz, DMSO-d_6) δ 13.03 (br s, 1H), 12.90 (br s, 2H), 2.50 (s, 3H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 165.5, 158.1, 151.0, 148.0, 109.9, 13.1.

IR (cm^{-1}): 3396 (NH), 3104 (NH), 1692 (C=O).

MS ES+ m/z calcd for $C_6H_8N_4OS_2$ (M+H)$^{+}$: 215.0061, found: 215.0070.

8(7,9*H*)-thioxoimidazo[4,5-*d*]pyrimidin-2,6(1,3*H*)-dione
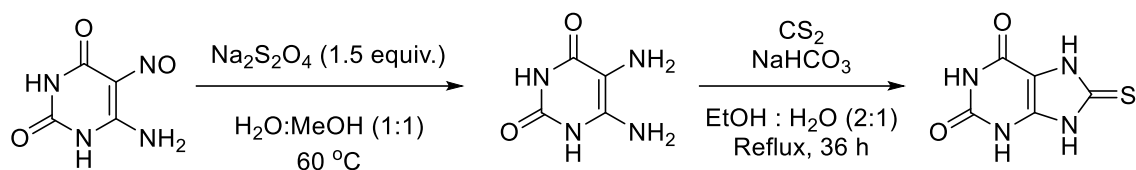
Followed general procedure 4.2.

Nitroso reduction: H$_2$O:MeOH (1:1) was used as solvent (77.4 mL), Na$_2$S$_2$O$_4$ (1.5 equiv., 3.43 g). Reaction carried out at 60 °C. The filtered material was washed with water and dried *in vacuo.* to provide the diamino compound as a beige solid.

CS$_2$ ring closure: After the reaction the mixture was filtered and the solid dried under high vacuum to provide 8(7,9*H*)-thioxoimidazo[4,5-*d*]pyrimidin-2,6(1,3*H*)-dione as a light-yellow solid (555 mg, 3.01 mmol, 27% over three steps).

^{13}C NMR (101 MHz, DMSO-*d*$_6$) *δ* 166.1, 154.5, 153.0, 150.8, 103.5.

IR (cm^{-1}): 3505 (NH), 3366 (NH), 1671 (C=O).

MS ES+ m/z calcd for C$_5$H$_6$N$_4$O$_2$S (M+H)$^+$: 185.0133, found: 185.0136.

8(7,9*H*)-thioxo-2-amino-6-methoxyimidazo[4,5-*d*]pyrimidine
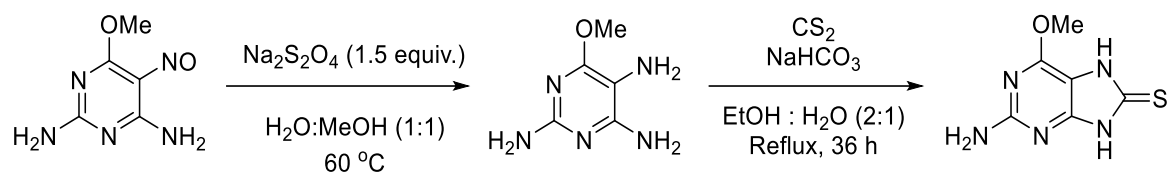
Followed general procedure 4.2.

Nitroso reduction: H_2O:MeOH (1:1) was used as solvent (52.0 mL) $Na_2S_2O_4$ (1.5 equiv., 2.30 g). Reaction carried out at 60 °C. Solvent removed *in vacuo.* and solid dried on high vacuum to provide the diamino compound as a beige solid.

CS_2 Ring closure: After ring closure the reaction mixture was filtered and the filtered solid was dried under high vacuum to provide 8(7,9*H*)-thioxo-2-amino-6-methoxyimidazo[4,5-*d*]pyrimidine as an off-white solid (1.63 g, 8.28 mmol, 94% over two steps).

M.P. 340 °C (decomp.).

^{1}H NMR (DMSO-d_6, 400 MHz) δ 12.6 (br s, 2H), 6.4 (br s, 2H), 3.9 (s, 3H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 166.6, 159.6, 154.5, 153.1, 103.0, 53.2.

IR (cm^{-1}): 3397 (NH), 3344 (NH), 3226 (NH), 1612 (NH bend).

MS ES+ m/z calcd for $C_5H_7N_5OS$ (M+H)$^+$: 198.0450, found: 198.0452.

Ethyl 2-({2-thiomethyl-6(1*H*)-oxo-9*H*-imidazo[4,5-*d*]pyrimidine}sulfan-8-yl)acetate
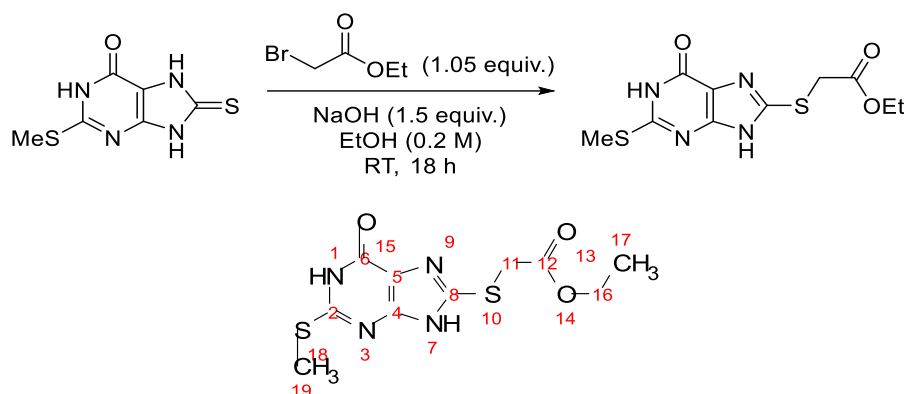
To an 8 mL vial was added 2-thiomethyl-6-oxo(5*H*)purine-8(7,9*H*)-thione (0.1 g, 0.466 mmol), sodium hydroxide (32.6 mg) and EtOH (2.7 mL 0.2 M). Ethyl bromoacetate (54.4 µL, 1.05 equiv.) was added *via* microsyringe. The vial was sealed with a supa seal and left to stir at RT for 18 h. The solvent was removed under reduced pressure and the residue redissolved in EtOAc and NH₄Cl soln. (10 mL each). The layers were separated, and the aqueous layer extracted with EtOAc (3x 10 mL). Organic layers collected and dried (MgSO₄) and solvent removed under reduced pressure. Purified by sequential column chromatography 5:5 -> 6:4 EtOAc : Pet.E, then 4% MeOH in DCM and recrystallised from EtOH : H₂O (1:1). Provided ethyl 2-({2-thiomethyl-6(1*H*)-oxo-9*H*-imidazo[4,5-*d*]pyrimidine}sulfan-8-yl)acetate as a white solid (37 mg, 0.12 mmol, 26%).

¹H NMR (400 MHz, DMSO-*d₆*) *δ* 13.40 (br s, 1H, **7**), 12.55 (br s, 1H, **1**), 4.13 (s, 2H, **11**), 4.11 (d, *J* = 7.1 Hz, 2H, **16**), 2.50 (s, 3H, **19**), 1.18 (t, *J* = 7.1 Hz, 3H, **17**).

¹³C NMR (101 MHz, DMSO-*d₆*) *δ* 168.4, 155.9, 61.2, 33.3, 14.0, 13.0 (6 out of 10 carbon resonances observed).

M.P. 144.2 – 145.3 °C.

IR (cm⁻¹): 3371 (NH), 1720 (C=O ester), 1645 (C=O amide).

MS ES+ m/z calcd for C₁₀H₁₄N₄O₃S₂ (M+H)⁺: 303.0400, found: 303.0402.

Adapted from literature procedure.[167]

2-({2-amino-6-methoxy-9*H*-imidazo[4,5-*d*]pyrimidine}sulan-8-yl)acetic acid
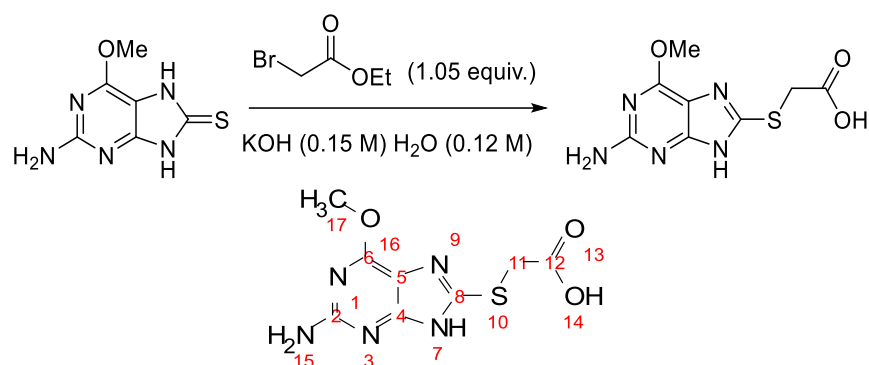
To a 100 mL RBF was added 8(7,9*H*)-thioxo-2-amino-6-methoxyimidazo[4,5-*d*]pyrimidine (0.6 g, 3.04 mmol) and a KOH aqueous solution (1.5 M soln., 25.35 mL, 0.12 M). Ethyl bromoacetate (0.353 mL, 1.05 equiv.) was added *via* syringe and stirred at RT for 18 h. The reaction mixture was acidified using acetic acid to pH = 6 and left to allow precipitation at 0 -10 °C for 20 min. Solid collected, triturated with water, and dried *in vacuo.* to provide 2-({2-amino-6-methoxy-1*H*-imidazo[4,5-*d*]pyrimidine}sulfan-8-yl)acetic acid (370 mg, 1.45 mmol, 38%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 14.70 (br s, 1H, *14*), 6.07 (br s, 2H, *15*), 3.91 (s, 3H, *17*), 3.46 (s, 2H, *11*).

^{13}C NMR (101 MHz, DMSO-d_6) δ 171.6, 159.4, 158.3, 53.3, 37.9 (5 out of 8 carbon resonances observed).

M.P. 280.5 °C (dec.).

IR (cm^{-1}): 3349 (NH), 3230 (OH acid), 1649 (C=O).

MS ES+ m/z calcd for $C_7H_9N_5O_3S$ (M+H)$^+$: 257.0526, found: 257.0532.

Adapted from literature procedure.[189]

Ethyl 2-({2-amino-6-methoxy-9*H*-imidazo[4,5-*d*]pyrimidine}sulan-8-yl)acetate
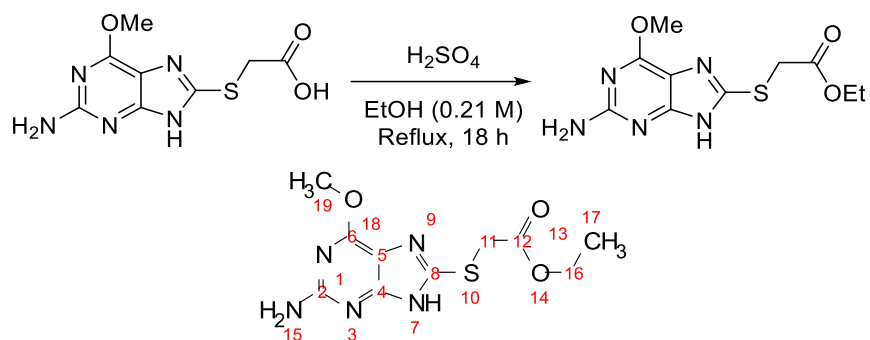
2-({2-amino-6-methoxy-1*H*-imiadazo[4,5-*d*]pyrimidine}sulfan-8-yl)acetic acid (300 mg, 1.05 mmol) was dissolved in EtOH (5 mL, 0.21 M) and with stirring concentrated sulfuric acid (5 mL) was added dropwise. The mixture was stirred at reflux for 18 h then allowed to cool, solvent was removed, and the residue purified by column chromatography (5% MeOH / DCM) to provide a white solid which was subsequently recrystallised from water to provide ethyl 2-({2-amino-6-methoxy-1*H*-imidazo[4,5-*d*]pyrimidine}sulfan-8-yl)acetate as a white solid (61.4 mg, 0.197 mmol, 19%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 12.61 (br s, 1H, **7**), 6.27 (br s, 2H, **15**), 4.10 (q, J = 7.1 Hz, 2H, **16**), 4.09 (s, 2H, **11**), 3.92 (s, 3H, **19**), 1.17 (t, J = 7.1 Hz, 3H, **17**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 168.5, 159.3, 144.0, 61.2, 53.0, 33.1, 14.0 (7 out of 10 carbon resonances observed).

M.P. 225.0 – 225.9 °C.

IR (cm^{-1}): 3488 (NH), 3339 (NH), 3221 (NH), 1732 (C=O).

MS ES+ m/z calcd for $C_{10}H_{13}N_5O_3S$ (M+H)$^+$: 284.0817, found: 284.0818.

5,6-dichloro-2-(methylsulfanyl)-1*H*-imidazo[4,5-*b*]pyrazine
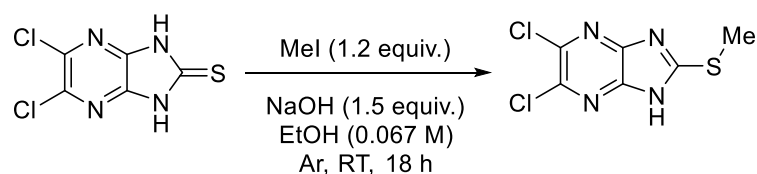
To an 8 mL vial was added a stirrer bar, 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (100 mg, 0.452 mmol), sodium hydroxide (27 mg, 1.5 equiv.) and placed under argon atmosphere. Ethanol (6.75 mL, 0.067 M) was added *via* syringe and stirred to dissolve materials. Iodomethane (34 µL, 1.2 equiv.) was added *via* microsyringe. The reaction was left to stir at RT for 18 h. The solvent was removed under reduced pressure and the residue taken up in EtOAc and NH$_4$Cl (5 mL each). The layers were separated and the aqueous extracted with EtOAc (3x 10 mL). The organic layers were collected and dried (MgSO$_4$) and solvent removed under reduced pressure. The crude material was purified by sequential column chromatography (7:3 Pet E : EtOAc, then 1% -> 2% MeOH in DCM) to provide 5,6-dichloro-2-(methylsulfanyl)-1*H*-imidazo[4,5-*b*]pyrazine as a yellow solid (18.7 mg, 0.08 mmol, 18%). May alternatively be recrystallised from EtOH, triturated with cold EtOH.

^{1}H NMR (400 MHz, DMSO-*d*$_6$) *δ* 14.01 (br s, 1H), 2.75 (s, 3H).

^{13}C NMR (101 MHz, DMSO-*d*$_6$) *δ* 163.0, 137.3, 14.0 (3 out of 4 carbon resonances observed).

M.P. > 200 °C

IR (cm^{-1}): 3009 (NH).

MS ES+ Calcd for C$_6$H$_4$^{35}Cl$_2$N$_4$S (M+H)$^+$: 235.9632, found: 235.9630.

Adapted from literature procedure.[167]

2-(methylsulfanyl)oxazolo[4,5-*b*]pyridine
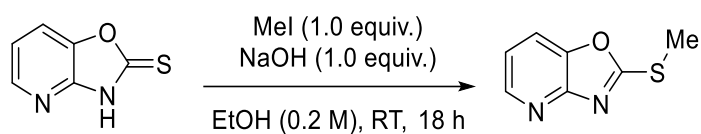
To an 8 mL vial was added oxazolo[4,5-*b*]pyridine-2(3*H*)-thione (100 mg, 0.657 mmol), sodium hydroxide (26.2 mg, 1.0 equiv.) and EtOH (3.28 mL, 0.2 M). To the stirred solution was added MeI *via* syringe (42 μL, 1.0 equiv.) and the mixture was left to stir at RT for 18 h. The solvent was removed *in vacuo.* and the residue redissolved in EtOAc (15 mL), transferred to a sep. funnel and washed with H_2O (15 mL). The phases were separated, and the aqueous layer extracted with EtOAc (2x 15 mL). The organic layers were collected, washed with brine (15 mL) and dried ($MgSO_4$). The solvent was removed under reduced pressure and the residue purified by column chromatography (7 : 3 Pet. E. : EtOAc) to provide 2-(methylsulfanyl)oxazolo[4,5-*b*]pyridine as a white solid (93.3 mg, 0.56 mmol, 85%).

^{1}H NMR (500 MHz, C*D*Cl₃) *δ* 8.46 (dt, *J* = 5.0, 1.3 Hz, 1H), 7.69 (dt, *J* = 8.1, 1.3 Hz, 1H), 7.17 (ddd, *J* = 8.1, 5.0, 1.1 Hz, 1H), 2.81 (d, *J* = 1.1 Hz, 3H).

^{13}C NMR (126 MHz, C*D*Cl₃) *δ* 170.5, 156.3, 146.0, 144.4, 119.0, 117.2, 14.9.

IR (cm^{-1}): 2819, 2623, 1548, 1097.

M.P. (lit.) 63 ᵒC, 63.5 – 63.9 ᵒC

Adapted from literature procedure.[167]

2-(methylsulfanyl)-1*H*-imidazole
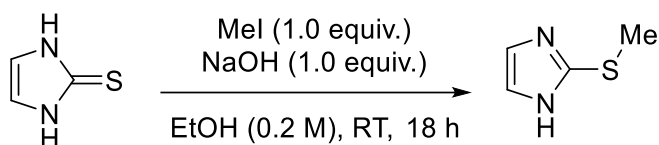
To an 8 mL vial was added imidazo-2(1,3*H*)-thione (100 mg, 0.998 mmol), sodium hydroxide (40.0 mg, 1.0 equiv.) and EtOH (2.75 mL, 0.2 M). To the stirred solution was added MeI *via* syringe (34 µL, 1.0 equiv.) and the mixture was left to stir at RT for 18 h. The solvent was removed *in vacuo.* and the residue redissolved in EtOAc (15 mL), transferred to a sep. funnel and washed with H_2O (15 mL). The phases were separated, and the aqueous layer extracted with EtOAc (2x 15 mL). The organic layers were collected, washed with brine (15 mL) and dried (MgSO$_4$). The solvent was removed under reduced pressure and the residue purified by column chromatography (6 : 4 Pet. E. : EtOAc) to provide 2-(methylsulfanyl)-1*H*-imidazole as a white solid (74.9 mg, 0.65 mmol, 66%).

^{1}H NMR (500 MHz, C*D*Cl$_3$) δ 9.26 (br s, 1H), 7.08 (s, 2H), 2.60 (s, 3H).

^{13}C NMR (126 MHz, C*D*Cl$_3$) δ 141.8, 16.8 (2 out of 3 carbon resonances observed).

IR (cm^{-1}): 2928, 1613.

M.P. (lit.) 138 – 139 °C, 138.2 – 139.4 °C.

Adapted from literature procedure.[167]

7-methyl-2-(methylsulfanyl)thiazolo[4,5-*b*]pyridine
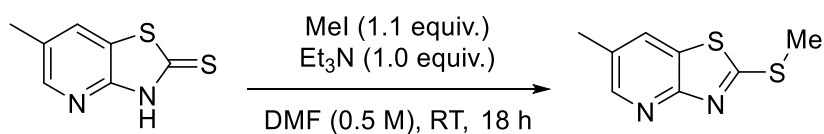
To an 8 mL vial was added 5-methyl thiazolo[4,5-*b*]pyridine-2(3*H*)-thione (100 mg, 0.536 mmol) and purged with Ar. Dry DMF (1.0 mL, 0.5 M) was added *via* syringe along with Et_3N (74.8 µL, 1.0 equiv.) and MeI (36.7 µL, 1.1 equiv.) and allowed to stir at RT for 18 h. Afterwards H_2O (6 mL) was added, and the precipitate filtered. The solid was washed with H_2O. The solid was redissolved in EtOAc (10 mL) and dried ($MgSO_4$) and solvent removed. The residue was purified by column chromatography (5 : 5 EtOAc : Hex.) to provide 6-methyl-2-(methylsulfanyl)thiazolo[4,5-*b*]pyridine as a white crystalline solid (65.8 mg, 0.33 mmol, 61%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 8.43 (app dd, J = 2.2, 0.7 Hz, 1H), 7.88 (dq, J = 2.2, 0.7 Hz, 1H), 2.85 (s, 3H), 2.45 (t, J = 0.7 Hz, 3H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 171.5, 162.4, 148.5, 130.0, 128.9, 128.8, 18.5, 16.0.

IR (cm^{-1}): 2623, 1548.

M.P. 115 – 118 °C (lit.) 123.5 – 124.2 °C

Adapted from literature procedure.[202]

5,6-dichloro-2-methyl-1*H*-imidazo[4,5-*b*]pyrazine
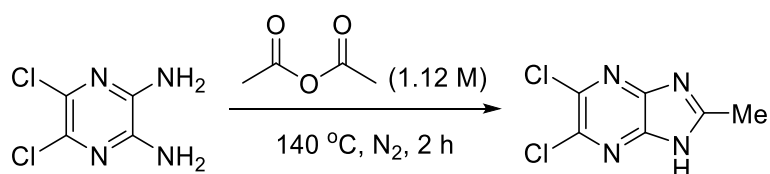
To an 8 mL vial was added 2,3-diamino-5,6-dichloropyrazine (100 mg, 0.558 mmol) and acetic anhydride (1 mL, 1.12 M). This was flushed with N_2 and sealed. The vial was placed in a preheated sand bath (140 °C) for 2 h. The vessel was allowed to cool to RT, diluted with NaOH solution (1 M, 10 mL) and transferred to a separatory funnel. Acidified with HCl solution (1 M, 10 mL) and extracted with EtOAc (3x 20 mL). the organic layers were collected, washed with brine (10 mL) and dried (MgSO$_4$) and solvent removed *in vacuo.* to the residue was added water (60 mL) and activated carbon, this was heated to boiling, filtered, and allowed to cool to RT. The formed crystals were isolated by filtration and dried in an oven for 10 min. (150 °C). Provided 5,6-dichloro-2-methyl-1*H*-imidazo[4,5-b]pyrazine as a white solid (63.6 mg, 0.31 mmol, 56%).

^{1}H NMR (400 MHz, DMSO-d_6) *δ* 10.39 (br s, 1H), 2.09 (s, 3H).

^{13}C NMR (101 MHz, DMSO-d_6) *δ* 168.8, 138.5, 137.5, 23.4.

M.P. > 200 °C

Adapted from literature procedure.[177]

2-phenyl-5-chloro-1*H*-imidazo[4,5-*b*]pyrazine
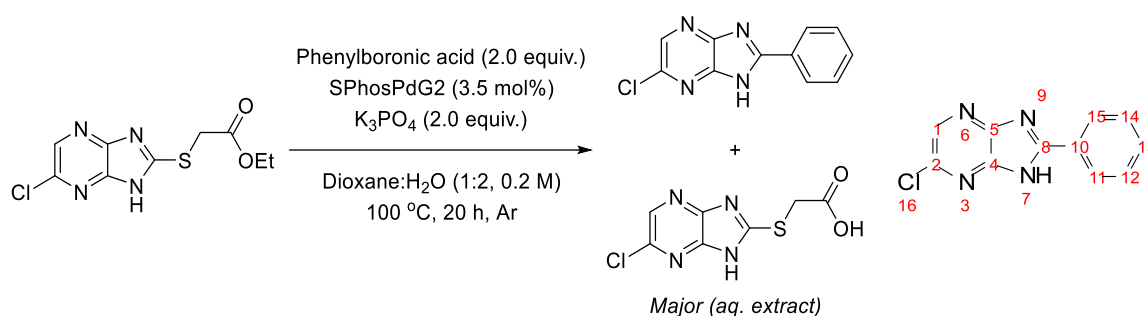
To an oven-dried vial was added SPhosPdG2 (9.3 mg, 3.5 mol%), phenylboronic acid (89.4 mg, 2.0 equiv.), K$_3$PO$_4$ (155.67 mg, 2.0 equiv.) and ethyl 2-({5-chloro-1H-imidazo[4,5-b]pyrazine-2-yl}sulfanyl)acetate (100 mg, 0.366 mmol). The vial was sealed and cycled on a Schlenk line (argon / vacuum) three times. Degassed dioxane : water mixture (1:2, 0.31 mL, 0.2 M) was added to the sealed vial *via* syringe and the mixture heated to 100 °C for 20 h. The cooled mixture was passed through a Celite® pad, washing with excess EtOAc. The filtrate was concentrated, and the residue purified by column chromatography (7:3 Hex. : EtOAc to provide 2-phenyl-5-chloro-1*H*-imidazo[4,5-*b*]pyrazine as a white solid (4 mg, 0.009 mmol, 4.7%).

^{1}H NMR (400 MHz, DMSO-*d*$_6$) *δ* 14.19 (br s, 1H, **7**), 8.48 (s, 1H, **1**), 8.38 – 8.13 (m, 2H, **11**, **15**), 7.70 – 7.58 (m, 3H, **12**, **13**, **14**).

^{13}C NMR (101 MHz, DMSO-*d*$_6$) *δ* 156.9 , **8**, 141.6, **2**, 137.0, **1**, 131.8, **13**, 129.2, **12**, **14**, 128.7, **10**, 127.3, **11**, **15** (7 out of 9 carbon resonances observed).

The Celite® pad was suspended in water and acidified with acetic acid, and this was extracted with EtOAc (3x 10 mL). The organic layers were collected and dried (MgSO$_4$) and concentrated. Analysis of this residue indicated the hydrolysed ({5-chloro-1H-imidazo[4,5-b]pyrazine-2-yl}sulfanyl)acetic acid was present by ^{1}H NMR.

Method adapted from literature procedure.[204]

5-*N*-morpholino-1*H*-imidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione
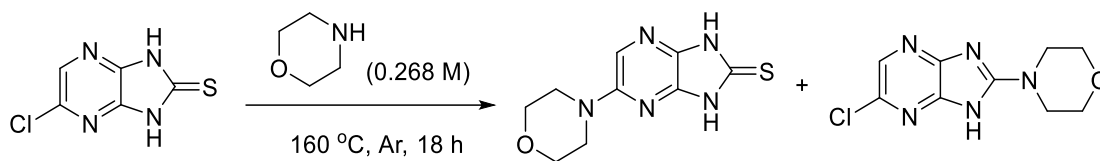
To an 8 mL vial was added 5-chloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (100 mg, 0.536 mmol) and morpholine (2 mL, 0.268 M), the vial was flushed with argon and sealed. The vial was then heated to 160 °C for 18 h. Afterward the vessel was allowed to cool to RT and transferred to a separatory funnel, washing with water to aid transfer (10 mL). Extracted with EtOAc (3x 10 mL), organic layers were collected and dried (MgSO$_4$) and solvent removed to provide an orange oil residue. Purified by column chromatography (6:4 DCM : EtOAc) to provide 2-*N*-morpholine-5-chloro-1*H*-imidazo[4,5-*b*]pyrazine (40.2 mg, 0.16 mmol, 31%) and 5-*N*-morpholino-1*H*-imidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (6.7 mg, 0.028 mmol, 5%). Adapted from literature conditions.[208]

5-*N*-morpholino-1*H*-imidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione:

^{1}H NMR (400 MHz, DMSO-d_6) δ 13.02 (br s, 2H), 7.76 (s, 1H), 3.75 – 3.68 (m, 4H), 3.39 (t, J = 4.9 Hz, 4H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 169.2, 152.2, 138.2, 132.9, 122.1, 65.7, 45.6.

M.P. 250 °C + (deg.).

IR (cm^{-1}): 3408 (NH), 2936 (C(sp^3)-H), 2852, 1193.

MS ES+ m/z Calcd for C$_9$H$_{11}$N$_5$OS (M+H)$^+$: 239.0785, found: 239.0781.

2-*N*-morpholine-5-chloro-1*H*-imidazo[4,5-*b*]pyrazine:

^{1}H NMR (400 MHz, DMSO-d_6) δ 12.58 (br s, 1H), 7.91 (s, 1H), 3.73 (dd, J = 5.8, 3.7 Hz, 4H), 3.65 (dd, J = 5.7, 3.7 Hz, 4H).

^{13}C NMR (101 MHz, DMSO) δ 159.1, 65.4, 45.4 (3 out of 7 carbon resonances observed).

M.P. 200 °C + (deg.).

IR (cm^{-1}): 2944 (C(sp^3)-H), 2853, 1283.

MS ES+ m/z Calcd for C$_9$H$_{10}$^{37}ClN$_5$O (M+H)$^+$: 243.0647, found: 243.0644.

Adapted from literature procedure.[208]

5-*N*-piperidyl-1*H*-imidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione
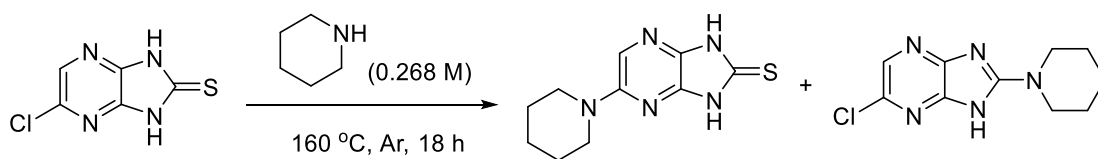
To an 8 mL vial was added 5-chloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (150 mg, 0.803 mmol) and morpholine (2.81 mL, 0.268 M), the vial was flushed with argon and sealed. The vial was then heated to 160 °C for 18 h. Afterward the vessel was allowed to cool to RT and transferred to a separatory funnel, washing with water to aid transfer (10 mL). Extracted with EtOAc (3x 10 mL), organic layers were collected and dried ($MgSO_4$) and solvent removed to provide an orange oil residue. Purified by column chromatography (9:1 –> 8:2 DCM : EtOAc) to provide 5-*N*-piperidyl-1*H*-imidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione as a yellow solid (16.8 mg, 0.0714 mmol, 9%) and 2-*N*-piperidine-5-chloro-1*H*-imidazo[4,5-*b*]pyrazine as a beige solid (31.8 mg, 0.133 mmol, 17%). Adapted from literature conditions.[208]

5-*N*-piperidyl-1*H*-imidazo[4,5-b]pyrazine-2(1,3*H*)-thione:

^{1}H NMR (400 MHz, DMSO-d_6) δ 12.92 (br s, 2H), 7.74 (s, 1H), 3.50 – 3.44 (m, 4H), 1.62 – 1.53 (m, 6H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 168.6, 152.3, 122.2, 46.12, 24.8, 24.0 (6 out of 8 carbon resonances observed).

M.P. 250 °C + (deg.).

IR (cm^{-1}): 3389 (NH), 2852 (C(sp^3)-H), 1608, 1255.

MS ES+ m/z Calcd for $C_{10}H_{13}N_5S$ (M+H)$^+$: 236.0970, found: 236.0964.

2-*N*-piperidine-5-chloro-1*H*-imidazo[4,5-*b*]pyrazine:

^{1}H NMR (400 MHz, DMSO-d_6) δ 12.40 (br s, 1H), 7.83 (s, 1H), 3.64 (t, *J* = 5.2 Hz, 4H), 1.70 – 1.54 (m, 6H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 159.0, 141.9, 128.7, 46.1, 24.9, 23.6 (6 out of 8 carbon resonances observed). 159.0 and 128.7 signals identified through HMBC and HSQC.

M.P. 250 °C + (deg.).

IR (cm^{-1}): 3394 (NH), 2962 (C(sp^3)-H), 2851, 1274.

MS ES+ m/z Calcd for $C_{10}H_{12}{}^{35}ClN_5$ (M+H)$^+$: 238.0860, found 238.0855.

Adapted literature procedure.[208]

5,6-dichloro-2,3-dihydro-1,3-thiazolo[3,2-*a*]imidazo[4,5-*b*]pyrazine
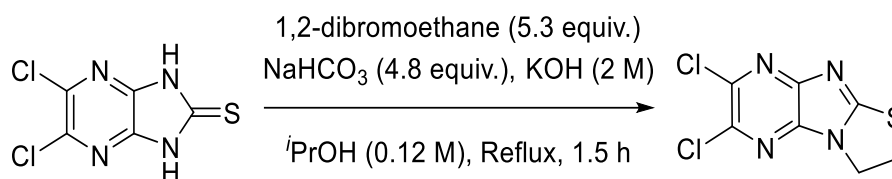
To a 5 mL RBF was added 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (100 mg, 0.452 mmol), KOH soln. (0.226 mL, 2.0 M) and ^{i}PrOH (2.23 mL). Stirred for 20 min., to a separate 5 mL RBF was added NaHCO$_3$ (182.4 mg, 4.8 equiv.) and ^{i}PrOH (1.38 mL), to this stirred suspension was added 1,2-dibromoethane (0.206 mL, 5.3 equiv.) and to this was surmounted an air condenser. The mixture was then heated to reflux with vigorous stirring. The thione solution was then taken up in a syringe and added to the refluxing dibromoethane suspension dropwise over 30 minutes from the top of the air condenser. After addition the reaction was monitored by TLC. After 1 h of reflux TLC indicated complete consumption of starting thione, the mixture was allowed to cool to RT. The reaction mixture was then transferred to a 100 mL RBF and KOH soln. (2 M, 10 mL) was then added, with acetone to aid transfer. Organic solvents and dibromoethane were then removed *in vacuo.*, the mixture transferred to a separatory funnel and extracted with EtOAc (3x 20 mL). The organic layers were collected and washed with dilute KOH solution, water, and then dried (MgSO$_4$). Solvent removed *in vacuo.* The residue was purified by column chromatography (1:1 Hexane : EtOAc) to provide light yellow crystals of 5,6-dichloro-2,3-dihydro-1,3-thiazolo[3,2-*a*]imidazo[4,5-*b*]pyrazine (58.5 mg, 0.23 mmol, 52%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 4.47 (dd, J = 8.1, 6.8 Hz, 2H), 4.09 (dd, J = 8.1, 6.8 Hz, 2H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 168.5, 151.9, 138.7, 137.3, 136.0, 44.0, 35.0.

IR (cm^{-1}): 2922 (C(sp^3)-H), 1580.

MS ES+ m/z Calcd for C$_7$H$_4$^{35}Cl$_2$N$_4$S (M+H)$^+$: 246.9612, found: 246.9602.

M.P. 274.6 – 276.8 ^{o}C (deg.).

Adapted from literature procedure.[209]

6,7-dichloro-2,3,4-trihydro-1,3-thiazino[3,2-*a*]imidazo[4,5-*b*]pyrazine
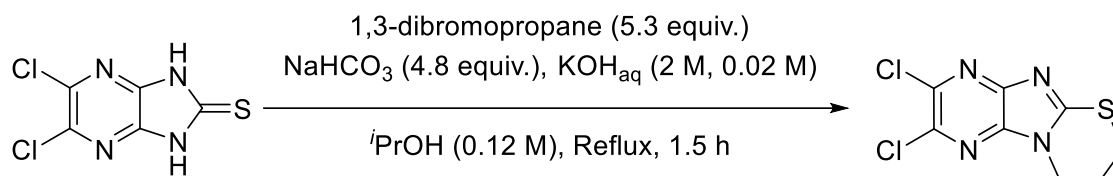
1,3-dibromopropane (5.3 equiv.)
NaHCO$_3$ (4.8 equiv.), KOH$_{aq}$ (2 M, 0.02 M)

^{i}PrOH (0.12 M), Reflux, 1.5 h

To a 5 mL RBF was added 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (100 mg, 0.452 mmol), KOH soln. (0.226 mL of 2.0 M) and ^{i}PrOH (2.23 mL). Stirred for 20 min. to dissolve thione. To a separate 5 mL RBF was added NaHCO$_3$ (182.4 mg, 4.8 equiv.) and ^{i}PrOH (1.38 mL), to this stirred suspension was added 1,3-dibromopropane (0.243, 5.3 equiv.) and to this was surmounted an air condenser. The mixture was then heated to reflux with vigorous stirring. The thione solution was then taken up in a syringe and added to the refluxing dibromopropane suspension dropwise over 30 minutes from the top of the air condenser. After addition the reaction was monitored by TLC. After 1 h of reflux TLC indicated complete consumption of starting thione, allowed to cool to RT. The reaction mixture was diluted with H$_2$O, transferred to a separatory funnel, washing with EtOAc to aid transfer. The mixture was extracted with EtOAc (3x 20 mL), the organic layers were collected, washed with brine (10 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* The residue was purified by column chromatography (6:4 Hex. : EtOAc) to isolate 6,7-dichloro-2,3,4-trihydro-1,3-thiazino[3,2-*a*]imidazo[4,5-*b*]pyrazine as an off-white solid (56.7 mg, 0.21 mmol, 48%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 4.25 (t, J = 5.8 Hz, 2H), 3.45 – 3.36 (m, 2H), 2.35 (ddt, J = 11.5, 8.5, 4.4 Hz, 2H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 158.2, 146.6, 139.3, 138.1, 135.7, 41.7, 25.5, 21.5.

IR (cm^{-1}): 2945 (C(sp^3)-H).

MS ES+ m/z Calcd for C$_8$H$_6$^{35}Cl$_2$N$_4$S (M+H)$^+$: 260.9768, found: 260.9763.

M.P. 249.9 – 250.4 °C (deg.)

Adapted from literature procedure.[209]

7,8-dichloro-2,3,4,5-tetrahydro-1,3-thiazepino[3,2-*a*]imidazo[4,5-*b*]pyrazine
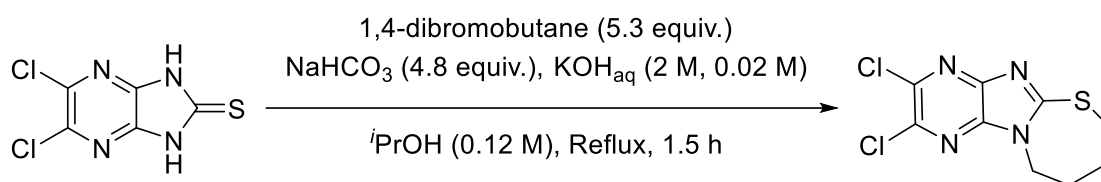
1,4-dibromobutane (5.3 equiv.)
NaHCO$_3$ (4.8 equiv.), KOH$_{aq}$ (2 M, 0.02 M)

^{i}PrOH (0.12 M), Reflux, 1.5 h

To a 5 mL RBF was added 5,6-dichloroimidazo[4,5-*b*]pyrazine-2(1,3*H*)-thione (100 mg, 0.452 mmol), KOH soln. (0.226 mL of 2.0 M) and ^{i}PrOH (2.23 mL). Stirred for 20 min. to dissolve thione. To a separate 5 mL RBF was added NaHCO$_3$ (182.4 mg, 4.8 equiv.) and ^{i}PrOH (1.38 mL), to this stirred suspension was added 1,4-dibromobutane (0.281 mL, 5.3 equiv.) and to this was surmounted an air condenser. The mixture was then heated to reflux with vigorous stirring. The thione solution was then taken up in a syringe and added to the refluxing dibromobutane suspension dropwise over 30 min. from the top of the air condenser. After addition the reaction was monitored by TLC. After 1 h of reflux TLC indicated complete consumption of starting thione, allowed to cool to RT. The reaction mixture was diluted with H$_2$O, transferred to a separatory funnel, washing with EtOAc to aid transfer. The mixture was extracted with EtOAc (3x 20 mL), the organic layers were collected, washed with brine (10 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* The residue was purified by column chromatography (7:3 Hex. : EtOAc) to isolate 7,8-dichloro-2,3,4,5-tetrahydro-1,3-thiazepino[3,2-*a*]imidazo[4,5-*b*]pyrazine as a crystalline off-white solid (40.5 mg, 0.14 mmol, 33%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 4.46 – 4.26 (m, 2H), 3.23 – 3.11 (m, 2H), 2.20 – 2.11 (m, 2H), 1.94 (ddd, J = 11.3, 7.3, 3.6 Hz, 2H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 161.2, 146.0, 139.6, 138.7, 137.7, 44.0, 31.8, 30.6, 26.1.

IR (cm^{-1}): 2924 (C(sp^3)-H), 2904.

MS ES+ m/z Calcd for C$_9$H$_8$^{35}Cl$_2$N$_4$S (M+H)$^+$: 275.9948, found: 275.9942.

M.P. 162.5 – 164.2 ^{o}C

Adapted from literature procedure.[209]

5 Biaryl Hits: Hydroquinone (NSC-2805) & Redoxal (NSC-73735)

5.1 Introduction

Biaryl systems are known to form core scaffolds within pharmacologically active compounds.[220] Since both NSC-2805 **5.1** (Hydroquinone) and NSC-73735 **5.2** (Redoxal) contain as a core moiety the biphenyl system, the work surrounding them is described in this chapter. A significant part of the section on NSC-2805 analogue synthesis was performed by (Dr to be) Joshua Hall, so several of the starting materials used in this section were synthesised previously in this project and are not described experimentally. Where appropriate, a reference to this is given to this work.[128] The compounds submitted for bioassay by Joshua have also been discussed in this chapter with the compounds whose synthesis is described here for a fuller picture of the work done in the NSC-2805 series. Below is a summary of the data associated with each hit compound (Figure 5-1).

Hydroquinone (NSC-2805)
WWP2 IC$_{50}$ (μM) 0.23

Redoxal (NSC-73735)
WWP1 IC$_{50}$ (μM) 21.7
WWP2 IC$_{50}$ (μM) 47.7
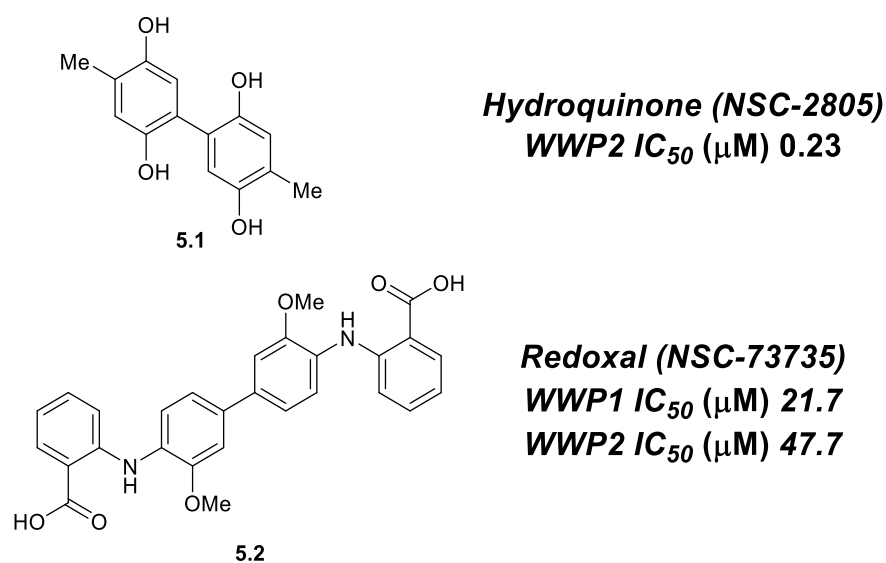
Figure 5-1: Hits NSC-2805 and redoxal.

5.1.1 Hydroquinone (NSC-2805): Introduction

Compound **5.1** has been investigated mainly in the context of virtual and experimental HTS campaigns against various targets for cancer therapy from the NCI diversity sets. For example, NSC-2805 has been identified as a Werner syndrome helicase inhibitor, displaying antiproliferative effects *in vitro*.[221] Additionally, NSC-2805 targets the vitronectin-binding site of the Urokinase receptor with sub-micromolar activity to block cancer cell invasion in a dose-dependent fashion.[222] A drug screen against gastroenteropancreatic neuroendocrine carcinoma tumour spheroids highlighted NSC-

2805 as active against both cell lines tested (Nec913 and Nec1452), as well as having a broad range of activity against epigenetic regulators (15 in total).[223] Lastly, a screen against α-methylacyl-Coenzyme A racemase (AMACR), a cancer biomarker identified NSC-2805 as a strong inhibitor.[224]

It is clear that NSC-2805 has a range of activity against several anti-cancer targets, with WWP2 being a recently identified target. Unfortunately, no synthetic analogues were pursued within the previous research on NSC-2805 and to the best of our knowledge this represents the first foray into hit-to-lead optimisation of NSC-2805 against a cancer target. NSC-2805 was previously synthesised by Joshua as described; using DDQ and BF_3 etherate in DCM at low temperatures 2,5-dimethoxytoluene **5.3** was oxidatively coupled to form **5.4** in high yield (Scheme 5-1).[128] From here, either BBr_3 demethylation or CAN oxidation to the quinone (to **5.5**) may be performed, with subsequent aqueous sodium dithionite workup required in both cases, reducing the quinone to **5.1**, all in quantitative yield. This product was taken forward by our colleagues in the biology department for co-crystal soaking studies, providing a structure as described in 2.3.
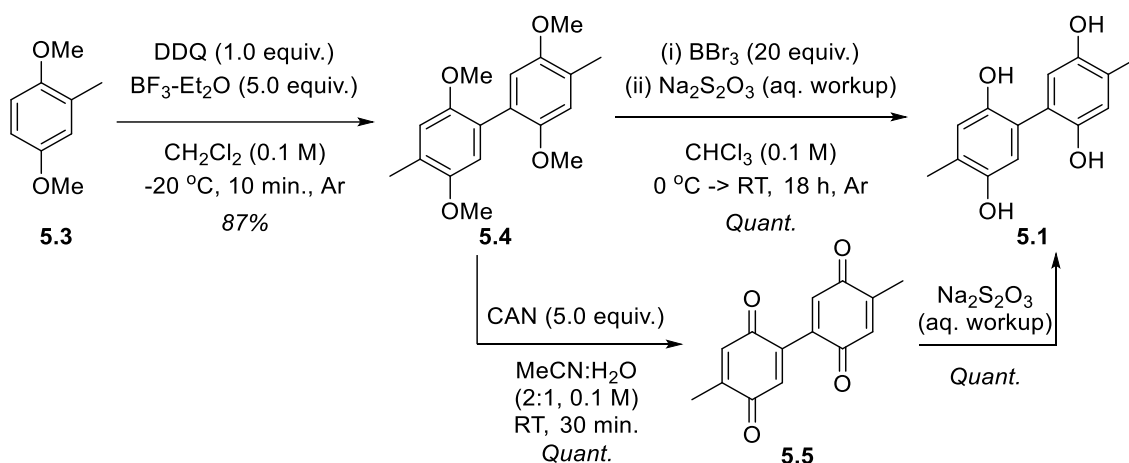
*Scheme 5-1: Previous synthesis of NSC-2805 **5.1** within our group.*

5.1.2 <u>Redoxal (NSC-73735): Introduction – Moving to *F*-Redoxal</u>

Redoxal was first reported as a redox indicator in alkaline media in 1960.[225] Redoxal has received sporadic interest since then as a bioactive molecule in HTS.[226–232] For example it was found to be active as a *M. tuberculosis* DPPS (*cis*-decaprenyl diphosphate synthase) inhibitor with antimicrobial activity against *M. smegmatis, K. pneumoniae, A. baumannii, P. aeruginosa, S. aureus* and *C. difficile*. Additionally, the authors found that redoxal was an inhibitor of undecaprenyl diphosphate synthase (UPPS) and farnesyl diphosphate synthase (FPPS), and a weak protonophore uncoupler.[226] As discussed in section 3.5.5.2, based on how Flare predicted its binding and the concerns surrounding working from

redoxal, a fragment library of *F*-redoxal was to be synthesised to find a smaller active compound.

5.2 <u>Hydroquinone (NSC-2805): Aims and Objectives</u>

The aim of this project was to improve upon the binding of NSC-2805 to WWP2 by developing rational analogues of NSC-2805. These derivatives have been chosen based on a protein-ligand co-crystal structure between NSC-2805 and WWP2.[28] There were two objectives for this project:

Firstly, to explore the potential for a hydrogen-bonding interaction between a hydroxyl group of ring A (Figure 5-2, left, highlighted) and the guanidine residue on the proximally located Arg803 subunit (Figure 5-2, middle). Crystal structure data shows two water molecules are present between Arg803 and this hydroxyl group.[28] Therefore, increasing this phenol motif by a single carbon and introducing a polar functionality could displace these water molecules and introduce direct hydrogen-bonding interactions between the guanidine side chain and ligand. This objective requires the preparation of a library of compounds with an additional carbon at this highlighted position bearing polar functional groups instead of the phenol functional group.
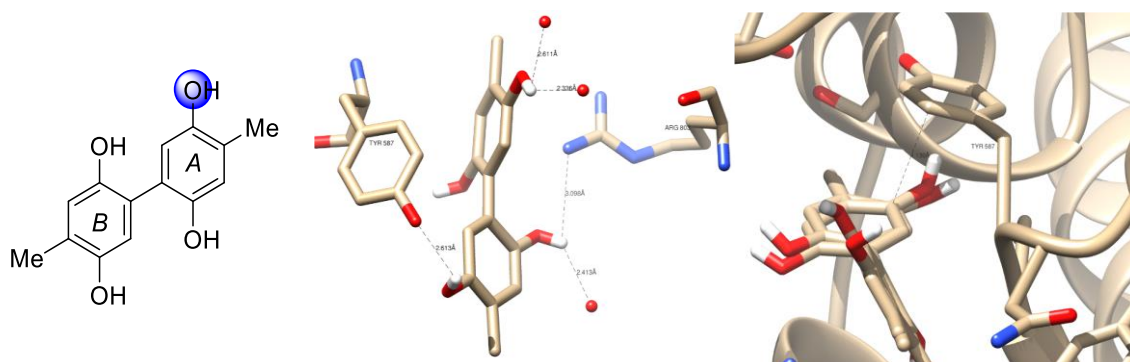
Figure 5-2: Left: Annotated structure of NSC-2805. Middle and Right: co-crystal structures of NSC-2805-WWP2 HECT domain depicting differing interactions. Crystal structure images reproduced from ref.[28]

Secondly, substituting ring B for an electron-deficient heterocycle to try improving the apparent π-π stacking interactions between ring B of NSC-2805 and the phenolic residue of the proximally located Tyr587 subunit, as seen on the ligand-protein complex crystal structure (Figure 5-2, right). It is predicted that the more electron-deficient aromatics would display stronger inhibition due to the additional polar interactions between the electron-rich phenolic residue and the electron deficient partners within the π-π stacking

interactions. A series of ring B variations bearing other aromatic and heterocyclic derivatives are targeted for synthesis.

5.3 *F*-Redoxal: Aims and Objectives

The main aim of the *F*-redoxal section of this chapter was to produce a fragment of *F*-redoxal with similar activity from which further optimisation into an inhibitor of WWP2 with sub-µM activity may be achieved. This means the main objective for this project was to produce a fragment library. A synthetic plan was required that would allow the efficient production of fragment analogues of *F*-redoxal as exemplified in Figure 5-3.
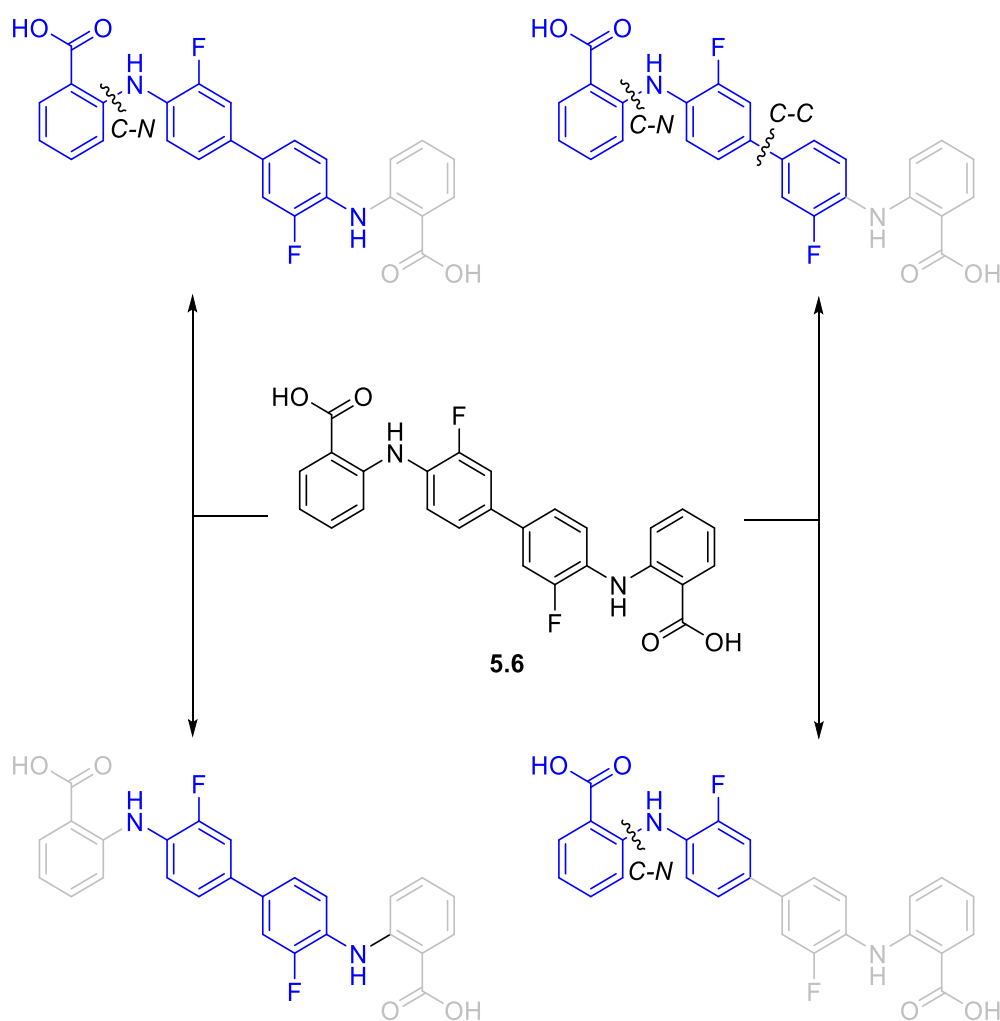
Figure 5-3: Fragments (highlighted in blue) of F-Redoxal targeted for synthesis. Indicated bond cleavage sites for synthesis of fragments.

Literature methods describing redoxal's synthesis are essentially unavailable, so an apparently novel route was devised instead. As a general outline, it was thought that coupling reactions of halogenated methyl esters and anilines followed by saponification

would allow access to all the desired fragments (Figure 5-3). The largest fragment, and *F*-redoxal itself was synthesised through benzidine core production followed by a Buchwald-Hartwig (B-H) C-N coupling reaction. The two smaller fragments may be accessed again through just a Buchwald-Hartwig cross-coupling or a Suzuki-Miyaura cross-coupling followed by a Buchwald-Hartwig cross-coupling. Afterwards, all compounds may be saponified to access the derivative fragments of *F*-redoxal.

For this objective to be successful, the synthesis of a library of *F*-redoxal must be achieved. The biological results from this small library should inform which fragments provide activity against WWP2, if not *F*-redoxal itself. From here, if successful, the second objective is to synthesise a larger library of molecules based on the most active fragment of the initial series.

5.4 <u>Hydroquinone (NSC-2805): Results and Discussion</u>

5.4.1 <u>Synthesis V: Improving Hydrophilic Interactions</u>

To accomplish the first goal for this project the potential to create or improve hydrophilic interactions between the hydroxyl group at the highlighted position in ring A (Figure 5-2, left) and the guanidine residue on Arg803 (Figure 5-2, middle), was pursued by synthesising derivatives with an extra carbon at the position highlighted bearing different polar functional groups. It was hypothesised that the distance between the hydroxyl group at position highlighted and Arg803 was too great for strong hydrogen bonding interactions, and an additional carbon linkage was necessary. All targets are accessible from the methoxy-protected bi-aryl carboxylic acid **5.7** through functional group interconversions (Figure 5-4). Compound **5.7** was produced previously *via* a Suzuki-Miyaura cross-coupling between 2,5-dimethoxy-4-methylphenylboronic acid and 3-bromo-4-methoxy-6-methylbenzoic acid.[128]
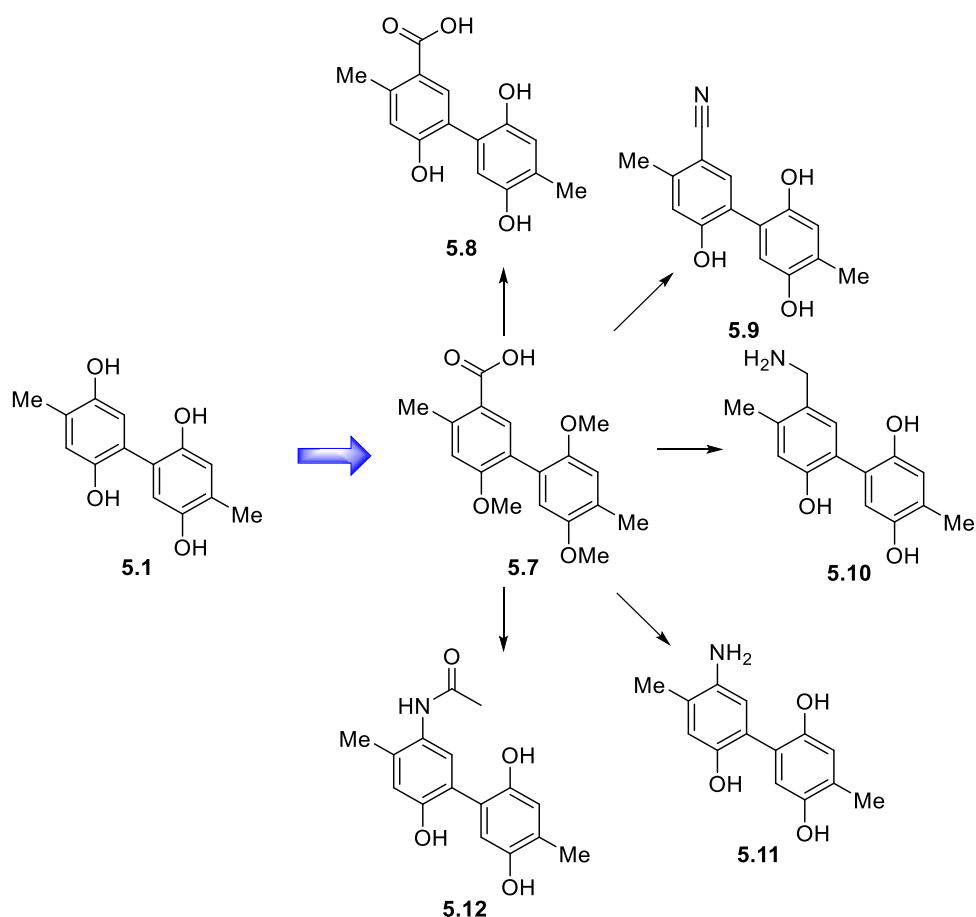
Figure 5-4: Synthetic Hydroquinone targets 5.8 – 5.12 for hydroquinone objective 1.

Starting from **5.7** the synthetic targets **5.8** and **5.9** were synthesised (Scheme 5-2). Amide formation occurred in 54% yield through the *in-situ* formation of the acid chloride *via* reaction with thionyl chloride at 50 °C with subsequent quenching in cold ammonium hydroxide solution to give **5.13**. Some starting **5.7** was recovered from the aqueous layer post-work up through neutralisation with acid as well as from the column of the crude reaction material, allowing recycling and production of more **5.13** in similar yields. Amide **5.13** was converted to **5.14** using catalytic amounts of triphenylphosphine under conditions of a catalytic Appel reaction.[233,234] Finally, demethylation of **5.14** and **5.7** using excess BBr$_3$ allowed the production of targets **5.8** and **5.9**.[235]
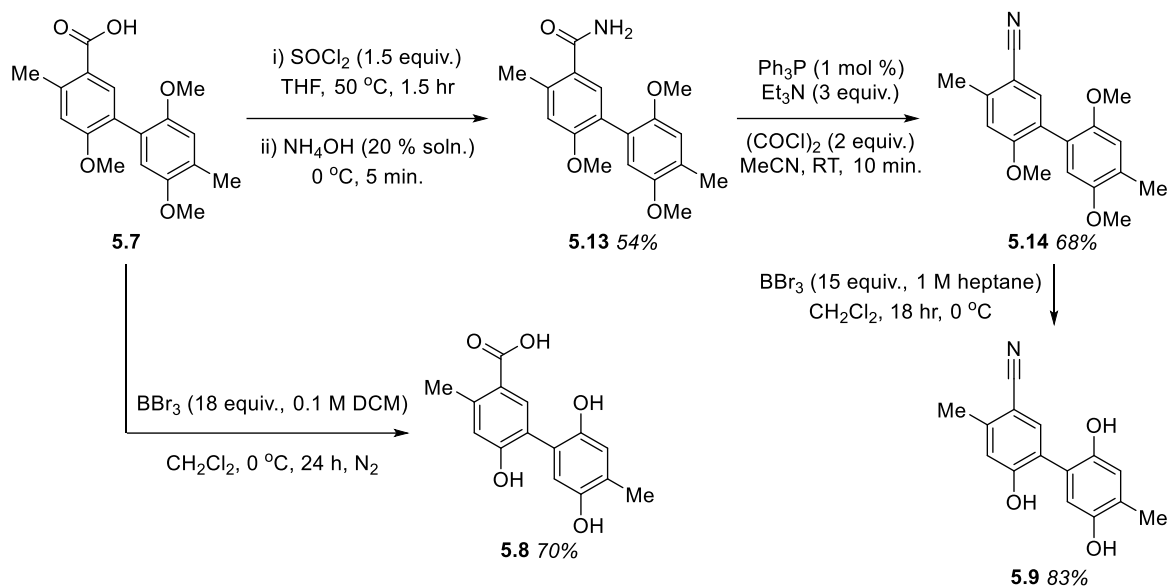
*Scheme 5-2: Synthetic routes to targets **5.8** and **5.9**.*

The amide dehydration is an application of an organophosphorus catalysis. Initially, triphenylphosphine **5.15** is thought to react with oxalyl chloride **5.16** to form a chlorophosphonium chloride salt ***int-5.19*** (through ***int-5.17*** and ***int-5.18*** with release of carbon monoxide) as the active reagent (Scheme 5-3). This is then thought to react with **5.13** as shown, which then can be deprotonated twice in sequence (***int-5.20*** and ***int-5.21***) with triethylamine to provide triphenylphosphine oxide **5.23** through ***int-5.22*** and the desired nitrile product **5.14**. The triphenylphosphine oxide may then react with oxalyl chloride through ***int-5.24*** and ***int-5.25*** to regenerate the chlorophosphonium chloride ***int-5.17***, releasing carbon dioxide and carbon monoxide as shown. This reaction was very effective, providing 68% isolated yield from a 10-minute reaction utilising only 1 mol% of triphenylphosphine.
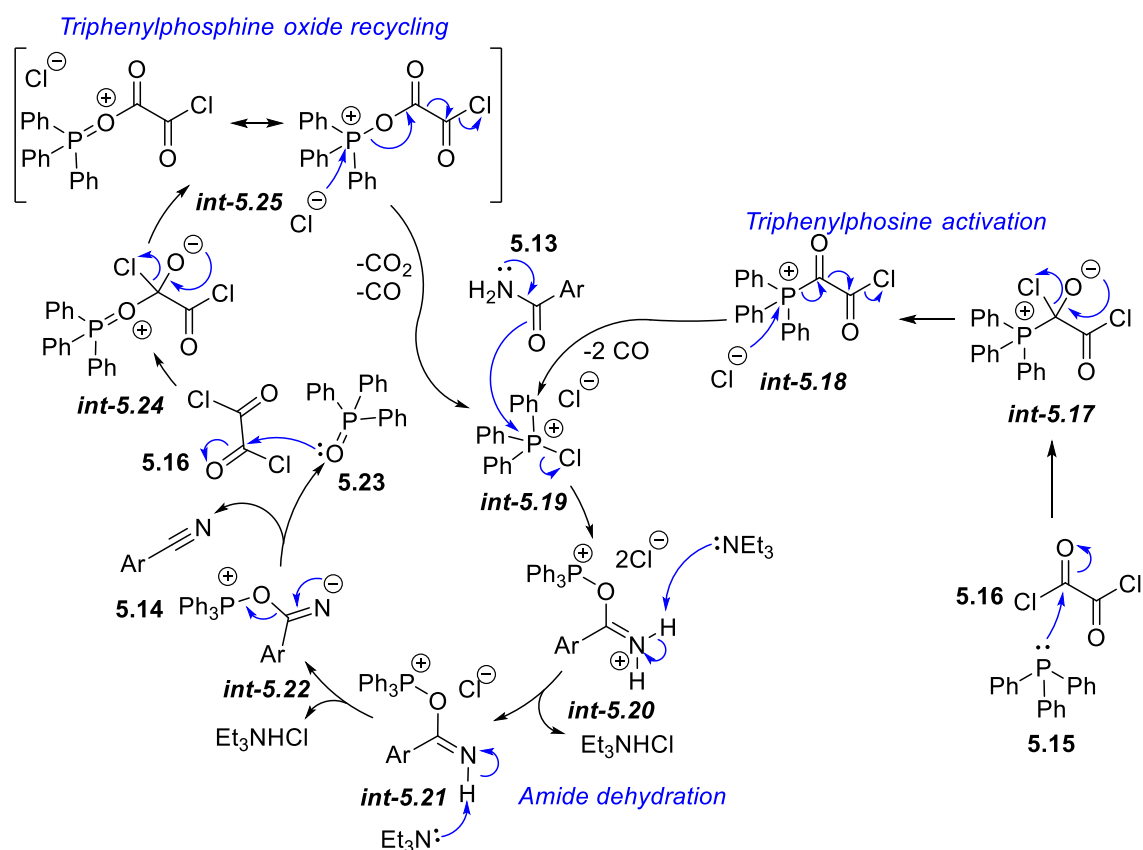
Scheme 5-3: Proposed mechanism for the Ph₃P catalysed dehydration of amide 5.13.

Reduction of **5.14** with LiAlH₄ allowed the quantitative conversion to **5.26** (Scheme 5-4). The subsequent demethylation towards **5.10**, appeared to be successful as the crude ^{1}H NMR spectrum showed no observable methoxy peaks corresponding to the starting material (Figure 5-5). Unfortunately, no phenol signals were identified, and this was a cause for concern as all other products had sharp singlets associated with the OH hydrogens. It was decided to not further pursue purification due to the small amount of crude material (22 mg) isolated from a 76 mg scale reaction.
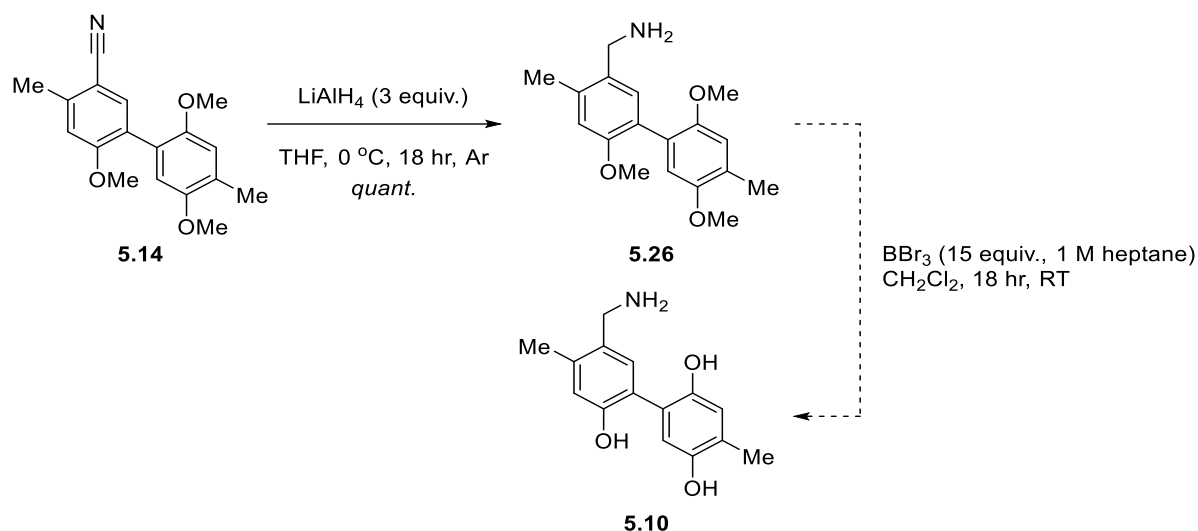
Scheme 5-4: Attempted synthetic route to target **5.10**.
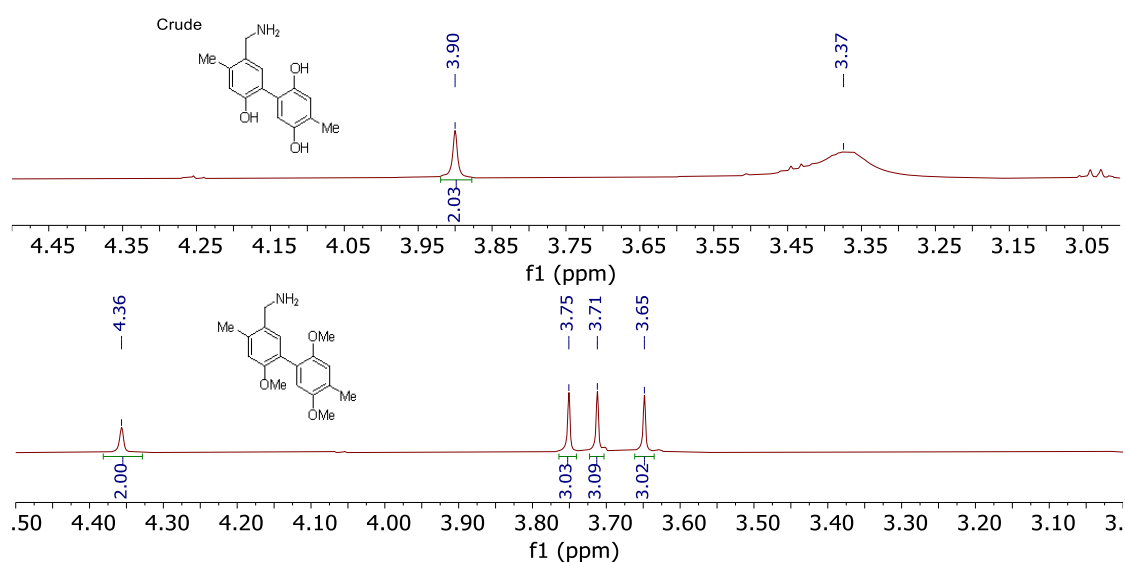
*Figure 5-5: ^{1}H NMR spectrum (3.0 – 4.5 ppm) of reaction mixture of target **5.10** (top) & synthetic intermediate **5.26** (bottom) describing the disappearance of methoxy signals.*

A different route was attempted to reach **5.10** (Scheme 5-5). Demethylation of **5.14** was attempted first with partial purification and direct transfer to the LiAlH$_4$ reduction. Unfortunately, no desired product was observable by ^{1}H NMR spectroscopy, and this is most likely due to such small amounts used in this attempt (9.9 mg).
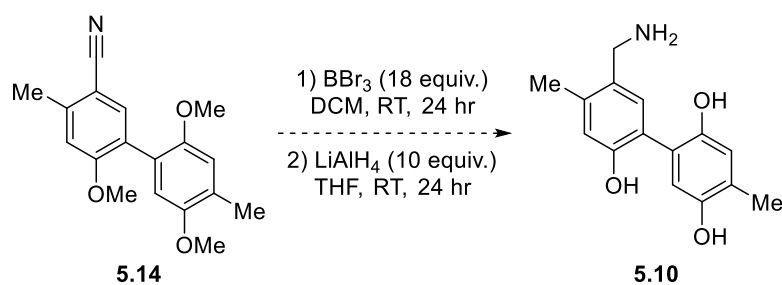
Scheme 5-5: Alternative unsuccessful route toward target 5.10.

Subjecting acid **5.7** to Schmidt reaction conditions to access the remaining two synthetic targets **5.11** and **5.12** was unsuccessful (Scheme 5-6), with 2D TLC showing many reaction products, all of which decomposed over time. Further investigation was not pursued.
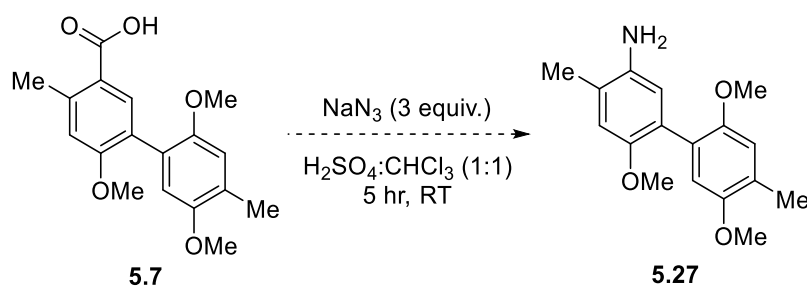
Scheme 5-6: Attempted Schmidt reaction to access 5.27.

5.4.2 Synthesis VI: Improving π-π Stacking Interactions

A small library of synthetic targets was chosen to investigate the potential to improve the observed π-π interactions between the B ring and the phenolic residue of the Tyr587, to accomplish the second objective of this project (**5.28** – **5.32**, Figure 5-6). The targets chosen include a variety of electron-deficient nitrogen-containing heteroaromatic rings, which would theoretically interact more strongly with the relatively electron-rich phenolic residue of Tyr587 compared to the hydroquinone ring B system.
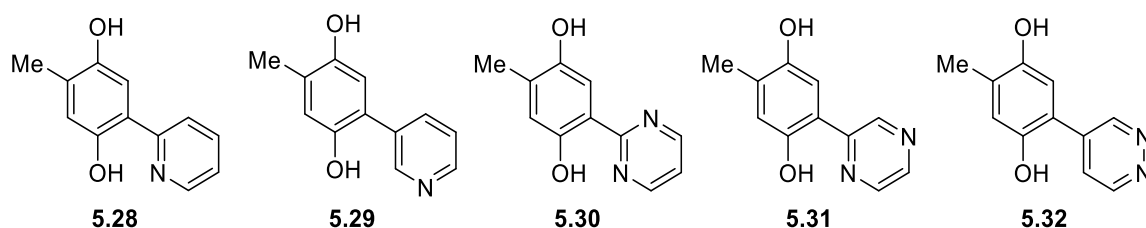
Figure 5-6: Synthetic targets for π-π interaction investigation.

To access these targets, coupling the common 2,5-dihydroxytoluene motif to a library of heterocyclic halides was chosen as the most efficient route. Therefore, **5.33** was efficiently brominated under acidic conditions in the presence of *N*-bromosuccinimide to provide **5.34** (Scheme 5-7). This was then borylated *via* metal-halogen exchange and subsequent tri-isopropyl borate treatment and hydrolysis allowed the generation of **5.35** in good yield.
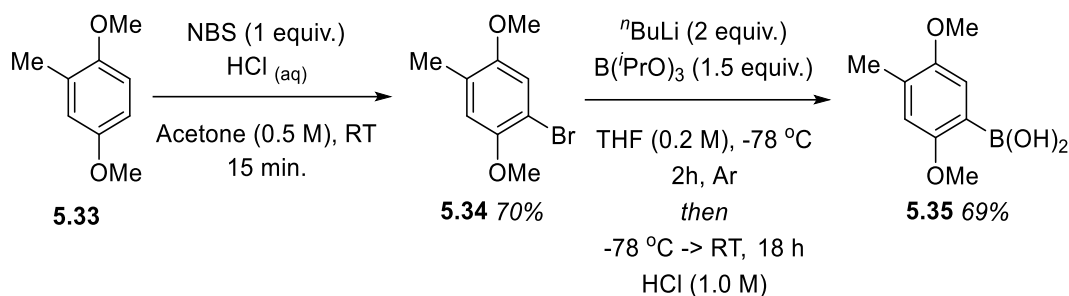
Scheme 5-7: Synthetic route towards coupling partner 5.35.

Demethylation of **5.35** was attempted using CAN oxidation to **5.36**, with subsequent reduction using $Na_2S_2O_4$ to **5.37**. The reaction was unsuccessful as no signal was observed in the ^{11}B NMR spectrum of the material isolated after reaction.
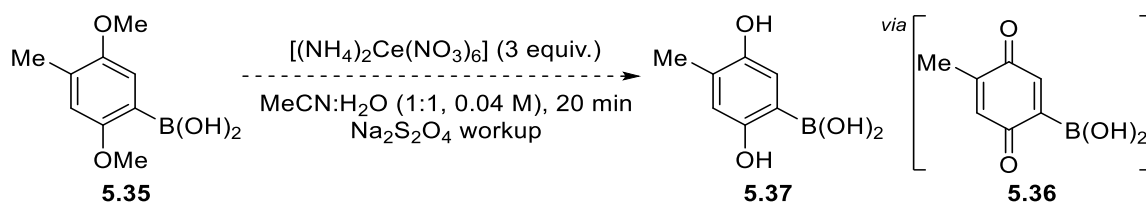
Scheme 5-8: Attempted CAN demethylation and Na₂S₂O₄ reduction to 5.37, with attempted cross-coupling reaction under Suzuki conditions.

Using an alternative route, **5.35** successfully coupled under Suzuki conditions with all the heterocyclic bromides (Scheme 5-9). This allowed straightforward access to the heterocyclic methoxy-protected intermediates **5.43 – 5.47**.
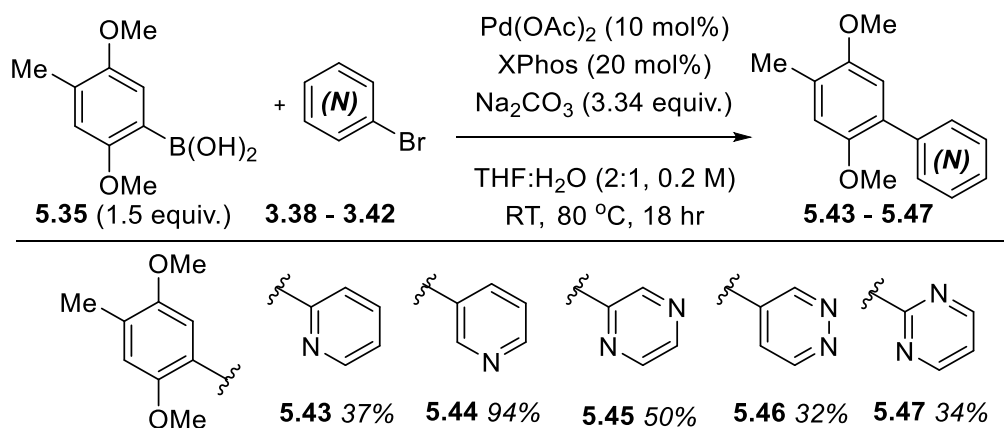
Scheme 5-9: Access to compounds 5.43 – 5.47.

Palladium catalysed C-C bond forming reactions are routinely used in academia and industry, highlighting its popularity is the fact that in 2014 it was the 5th most frequently performed reaction and Akira Suzuki's Nobel prize in 2010 (along with Richard F. Heck and Ei-ichi Negishi for their contribution to the discovery and development of palladium-catalysed cross-couplings in organic synthesis).[236,237]

The conditions require a palladium source, ligand (or precatalyst of these two), base, sparged solvents including water, as well as the aryl halide and aryl boron species (for $C(sp^2)$-$C(sp^2)$ cross-coupling). It is well-known that usually Pd^0 species are the catalytically active species in Suzuki-Miyaura cross-couplings and can be generated from either *in-situ* mixing of excess ligand with the Pd^0 source or from L_4Pd^0 sources. Additionally, Pd^{II} sources such as $L_2Pd^{II}X_2$ can be utilised, although instead of ligand dissociated as with L_4Pd^0 sources, a reductive elimination process must occur to form the catalytically active L_2Pd^0 species.[238] However, it was shown that L_1Pd^0 species are the more active species, with experimental analysis demonstrating that sterically bulky ligands favour the monoligated Pd^0 species.[239] The generation of mono-coordinated Pd^0 species can be formed in several different ways from precatalysts, with different research groups developing divergent precatalyst designs. These include: $[L_1Pd^IX_2]_2$ (X = Br, I) and $[L_1Pd^{II}X_2]_2$ (X = Br, Cl) dimers **5.48** & **5.49**, Pd^{II} palladacycles like **5.50** and **5.51**, and L_1Pd-based acyclic precatalysts including Pd(π-*R*-allyl)(NHC)Cl complexes like **5.52** and phosphine-based π-allyl precatalysts **5.53** (Figure 5-7).[238] Each pre-catalyst group have their own routes to generate the catalytically active species.[240]
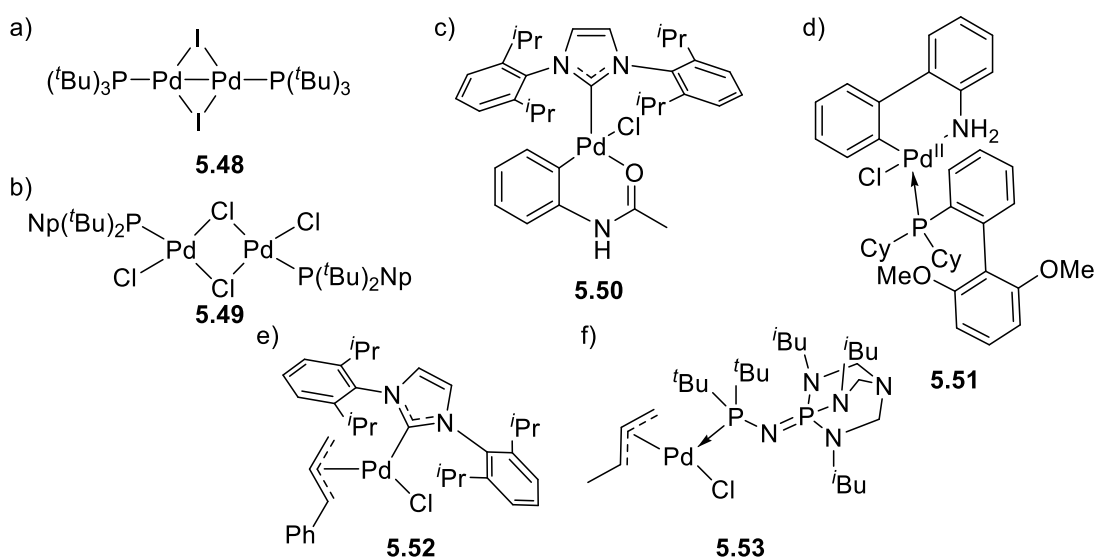
*Figure 5-7: Examples of Pd precatalysts, a) Pd[I] dimeric species **5.48**, b) Pd[II] dimeric species **5.49**, c) Pd[II] palladacycle NHC ligated species **5.50**, d) phosphine-based Pd[II] palladacycle **5.51**, e) Pd[II](π-R-allyl)(NHC)Cl precatalyst **5.52**, f) phosphine-based π-allyl Pd[II] precatalyst **5.53**.*

No matter the way the catalytically active species is formed, the catalytic cycle proceeds in three general steps: (i) oxidative addition between the Pd^0 species and an aryl(pseudo)halide, providing a *cis*-isomer which can rapidly isomerise into the *trans*-isomer (ii) transmetalation between the boron species and the Pd^{II} intermediate and (iii) reductive elimination from the Pd^{II} species to form the desired C-C bond in the product and regenerate the Pd^0 species.[241] In the first step (Scheme 5-10), the oxidative addition of *int-5.55* to the aryl halide **5.54** provides *int-5.56*, which in the presence of two ligands is in the *cis*-conformation, and rapidly isomerises to the *trans*-organopalladium intermediate. Transmetalation is thought to occur *via* firstly a ligand exchange between the palladium-bound halide *int-5.56* and a hydroxide ion formed by the base and water present in the reaction, forming *int-5.57*. This hydroxopalladium intermediate is then thought to interact with the boronic acid species **5.58**, forming a '8-B-4' intermediate *int-5.59* with a Pd-O-B linkage, and based on experimental observations and computational calculations this then undergoes transmetalation to form *int-5.60*.[240] After this reductive elimination occurs, forming the carbon-carbon bond between the coupling partners (forming **5.61**) and regeneration of the Pd^0 species *int-5.55*.
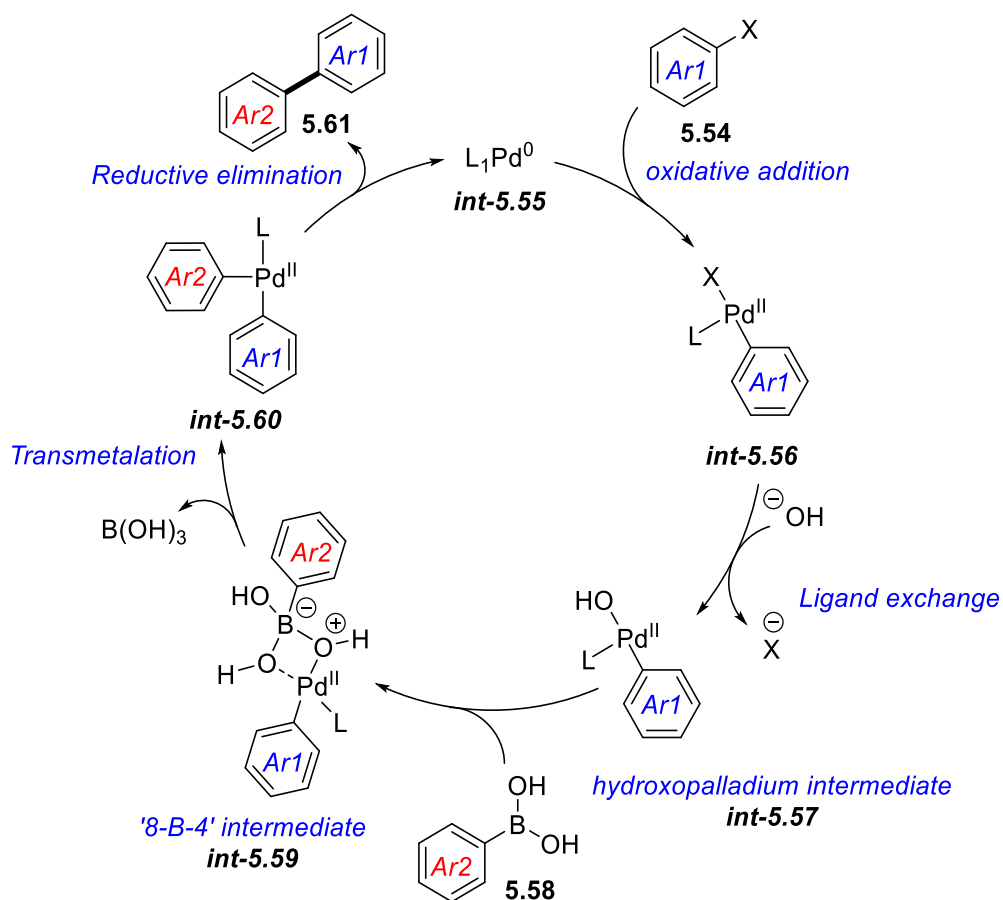
Scheme 5-10: Catalytic cycle of the Suzuki-Miyaura cross-coupling reaction.

A major problem encountered in Suzuki cross-couplings is catalyst poisoning with highly basic coordinating substrates, such as heteroaromatic compounds often found in drug molecules, which produce off-cycle resting states that are non-productive in the catalytic process. These species have been shown by computational and experimental methods to provide their inhibitory effect by their coordinating ability, and not the energetics surrounding oxidative addition or reductive elimination.[204] The researchers found that not only were singly substrate-ligated Pd species acting as Pd reservoirs, but the dimeric forms were also possible, depending on the substrates, and that these states were energetically favourable compared to the on-cycle species they are in equilibrium with. In light of these results, the authors demonstrated that increased loadings of the boron species and elevated temperatures to be favourable, as well as extended reaction times. This is because the increased concentration of the boron species shifts the equilibrium back to on-cycle, as these are reacting faster with the higher concentration of the boron species. Elevated temperatures increase the rate of exchange between equilibrium positions and so additionally favours this process.

The choice of ligand may be important in these systems, as the use of bidentate chelating ligands would reduce the ability of the systems to form the off-cycle substrate-ligated resting states, most likely due to steric constraints around the Pd centre and the electron-rich nature of these ligands.[242] However, the development of dialkylbiarylphosphine ligands by the Buchwald group have also found use as highly robust ligands for a variety of cross-coupling reactions, not just for heteroaromatic substrates.[243] In Figure 5-8 below the structural features of this ligand class impacting on the efficiency of the catalysts is presented. The bulky and electron-rich behaviour of these ligands are thought to stabilise the L_1Pd^0 species, with oxidative addition being shown to occur faster for the monoligated species than for more highly coordinating complexes.[243] This is thought to be a direct result of the overall smaller Pd complex, allowing substrates to approach closer, and hence the oxidative addition to be more facile.[244] It is thought that the transmetallation step can be rationalised by a similar approach. Reductive elimination has additionally been shown to occur faster in the monoligated form than a bis-ligated form.[245] The lower ring *ortho* substitutions (R_3 and R_4) negate the formation of palladacycles and increase the ligand size, thereby improving the activity and stability of the Pd complexes and further stabilising the L_1Pd^0 active species.[246]
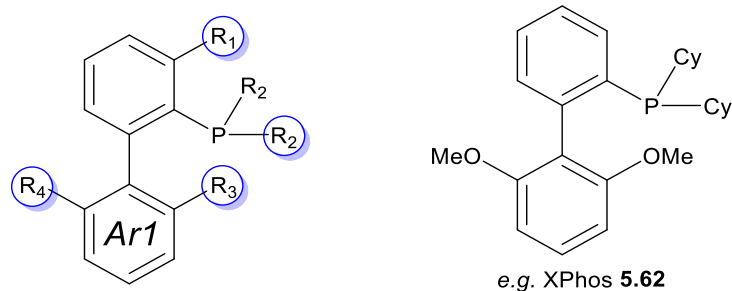
· R_1: Fixes PR_2 over bottom of ring, enhancing rate of reductive elimination
· R_2: Increased size enhances rate of reductive elimination, alkyl groups
  increase electron density at P, enhancing oxidative addition.
· R_3, R_4: If large, increases $[L_1Pd^0]$ by stabilisation & prevents
  cyclometalation, increases stability
· *Ar1*: Increases size of ligand, slows rate of [O] by O_2, allows stabilising
  ipso-Pd interactions, promotes reductive elimination

Figure 5-8: Structural features of dialkylbiarylphosphine ligands.

Several of the Suzuki cross-coupling products were demethylated using a CAN oxidation from the 1,4-dimethoxy motif to *para*-quinone intermediates through a single electron transfer process. After this reaction and during the workup treatment with sodium dithionite reduces these to the *para*-hydroquinone motif, successfully demethylating to reach **5.28**, **5.29** and **5.31** (Scheme 5-11). The final two compounds **5.30** and **5.32** were unsuccessful in the reaction and did not yield the demethylated products.
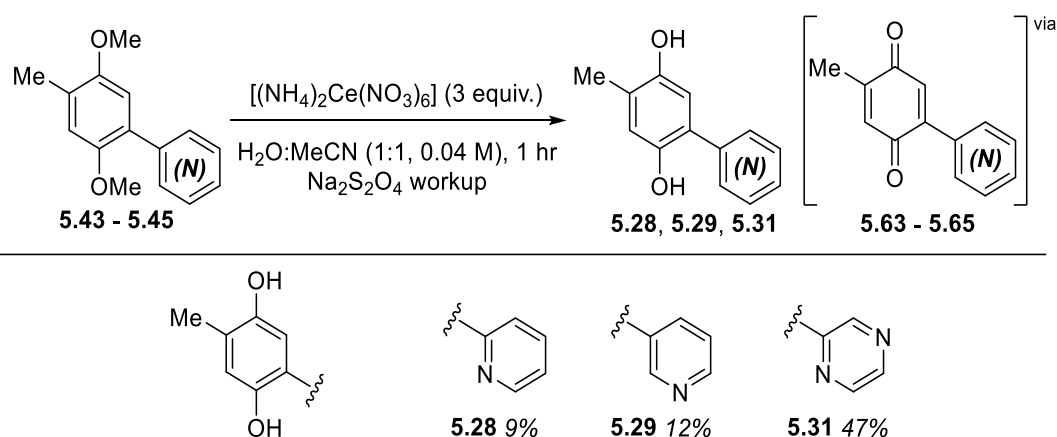
*Scheme 5-11: Access to synthetic targets **5.45 – 5.47**, demethylation via CAN oxidation to **5.63 – 5.65** followed by sodium dithionite reduction.*

As the oxidative demethylation using CAN has been utilised for this project, its mechanism is discussed. It is thought to take place *via* first a single electron transfer (SET) of ***int-5.66*** to generate a radical cation (***int-5.67***, Scheme 5-12). This can then be attacked by a water molecule forming ***int-5.68***, another resonance form is shown as the tertiary radical ***int-5.69***. CAN can then abstract another electron through SET to form a tertiary cation ***int-***

5.70, which is then attacked by water to form the di-hemiketal intermediate *int-5.71*. This postulated mechanism is based on oxygen isotope enrichment studies, which have shown the oxygen from the solvent is the one incorporated into the product.[247] Finally, elimination of two equivalents of methanol forms the quinone product *int-5.72*.
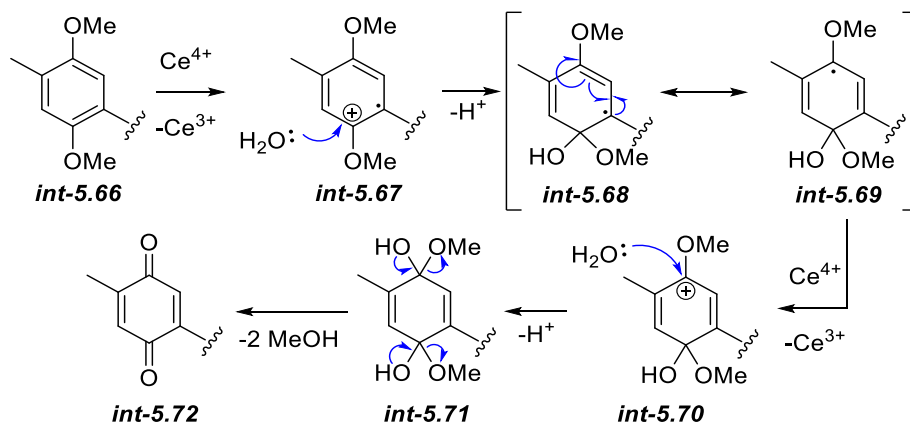
Scheme 5-12: Mechanism for oxidative demethylation using CAN.

To complete the synthesis, the pyridazine and pyrimidine-containing bi-aryls **5.48** and **5.49** were demethylated using BBr$_3$ in DCM overnight using the same method as previously employed, successfully delivering the last two synthetic targets **5.30** and **5.32** in 86% and 38% yields, respectively (Scheme 5-13).
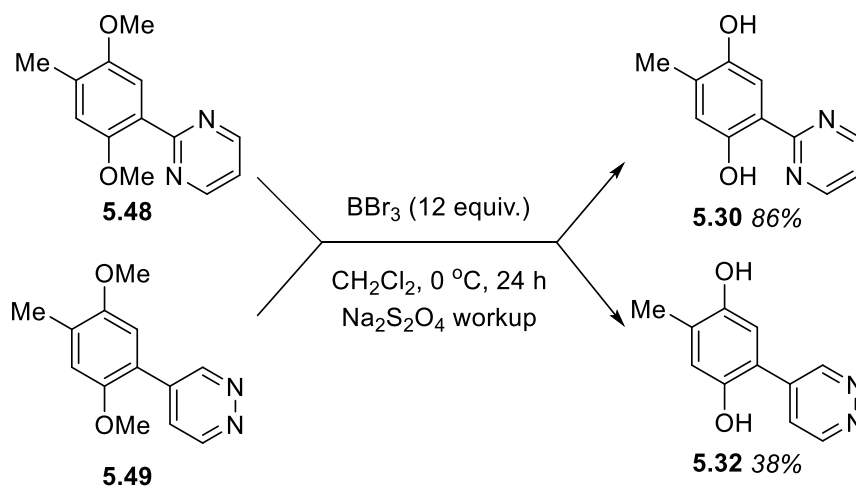
*Scheme 5-13: BBr$_3$ deprotection of **5.48** and **5.49** to give synthetic targets **5.30** and **5.32**.*

5.5 *F*-Redoxal results and discussion.

5.5.1 Synthesis VII: Coupling Reactions.

To start the synthesis of the *F*-redoxal fragment series, a literature procedure which allowed the synthesis of three variously fluorinated benzidine compounds was found

starting from 4-bromo-2-fluoroaniline **5.73** (Scheme 5-14).[165] This homocoupling method gave an opportunity to explore the effects of the aryl fluorine atoms on activity of both the mono- and di-arylated benzidine products, starting from a single step. Cetyltrimethylammonium bromide ($[(C_{16}H_{33})N(CH_3)_3]Br$, CTAB) acted as a phase-transfer catalyst, but the poor yields of **5.74**, **5.75** and **5.76** may have resulted from the insolubility of starting material in the reaction medium or reductive dehalogenation of the starting aryl bromide by action of palladium with sodium formate. Due to the starting material being commercially available and relatively cheap, starting on a 10 g scale (52.6 mmol, £42.50/10 g) provided enough material of all three benzidine derivatives for use in the next steps.
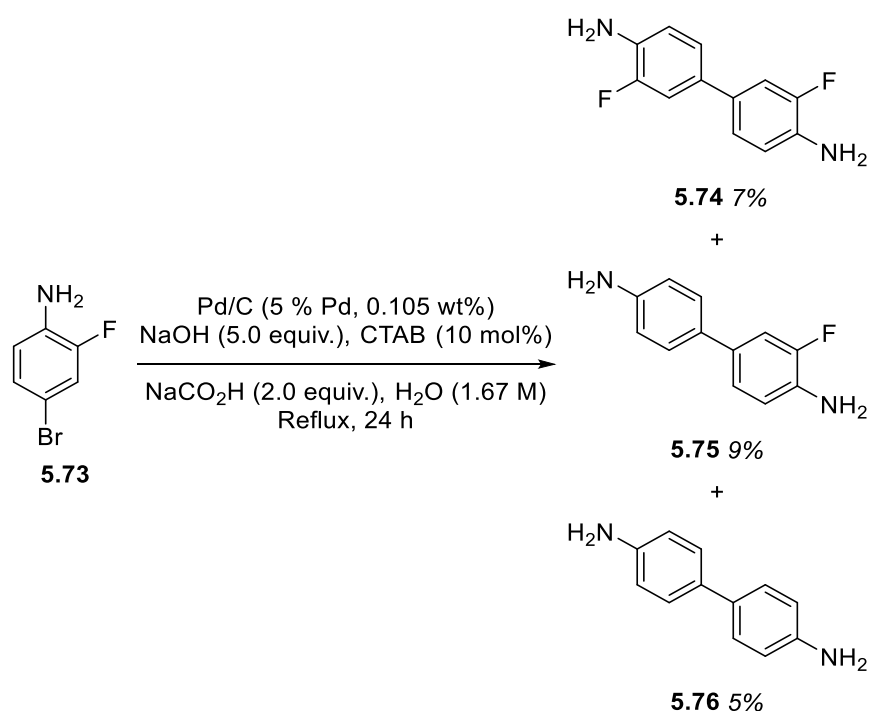
*Scheme 5-14: synthesis of fluorinated benzidine compounds **5.74 – 5.76**.*

With the benzidine compounds in hand, Buchwald-Hartwig C-N cross-coupling reactions were performed adapting a literature method as a general procedure (Scheme 5-15).[146] Choosing 1.5 equiv. of aryl bromide **5.77** was to theoretically provide even mixtures of both the mono- and di-arylated amines for library synthesis. Initial purification conditions were elution with EtOAc : hexane on SiO_2 and performed on the reaction mixture of **5.74**, this unfortunately led to severe streaking and required further purification on preparative TLC, hence providing lower yields of **5.78** and **5.79**. Switching the elution solvent to toluene on SiO_2 provided efficient separation for the other two reaction mixtures. Unfortunately, di-arylated benzidine **5.82** was not formed under the reaction conditions, for

unknown reasons. With the five methyl esters or diesters in hand the final hydrolysis step could now be performed.
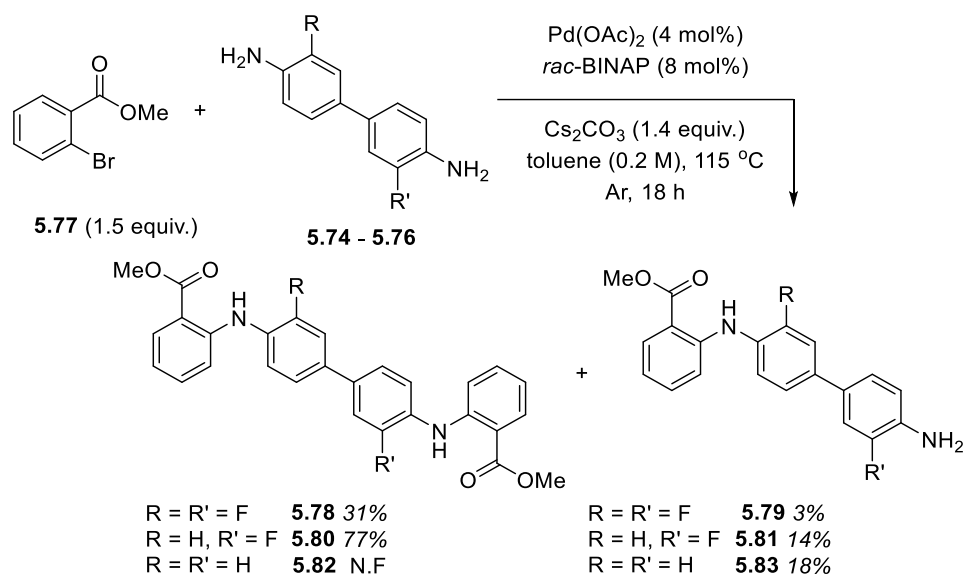
*Scheme 5-15: Buchwald-Hartwig cross-couplings of partially-fluorinated benzidine derivatives **5.78** – **5.81** & **5.83**. N.F. = Not formed*

For the mono-fluorinated benzidine derivative **5.81**, ^{1}H NMR spectroscopy is indicative of a single isomer present, and ^{1}H-^{19}F HMBC and ^{1}H-^{19}F HOESY NMR experiments indicate that both through bond and through space the aniline functional group is interacting with the fluorine atom, whereas the 2° diarylamine is not (Figure 5-9). Therefore, it is thought the mono-arylated mono-fluorine isomer is the one described above. The other isomer **5.81'** was not observed on TLC nor isolated during column chromatography. This could be rationalised by the fluorine reducing the nucleophilicity of the proximal amine, which reduces its co-ordination to the Pd-centre compared to the other amine present on the bi-aryl motif.
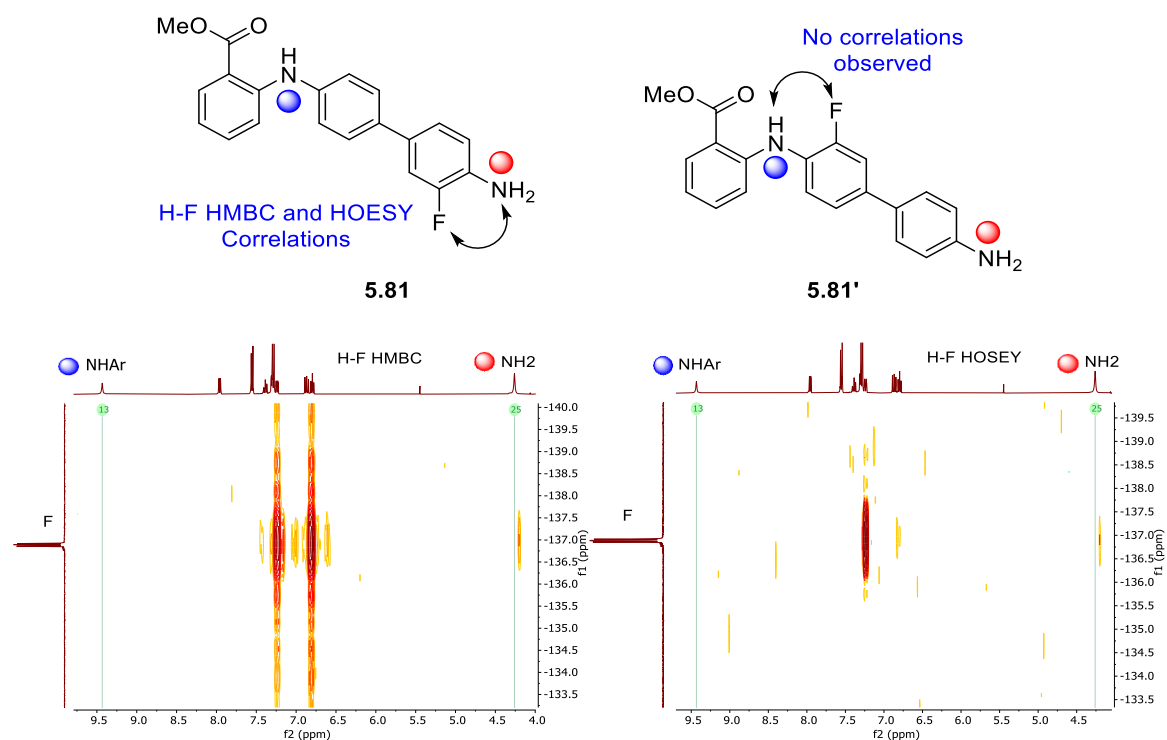
*Figure 5-9: Fluorinated regioisomer **5.81** determination via 2D H-F NMR (CD$_3$CN).*

The Buchwald-Hartwig C-N cross-coupling reaction is an important reaction in modern organic chemistry, exemplified by several reviews, and with notable applications in heterocycle and natural product synthesis.[248,249] The mechanism of the reaction has also been explored extensively and is discussed only briefly here.[250] Initially precatalyst Pd(OAc)$_2$ and ligand *rac*-BINAP are added individually to form *in-situ* the active catalyst **5.85** after Pd^{II} reduction and ligand dissociation, rather than as a single precatalyst. Reductive elimination of some kind to form Pd^0 from Pd^{II} occurs, and this varies depending on the conditions and ligands present, in this case most-likely from oxidation of one of the BINAP phosphine groups, providing **int-5.84**. For these specific reaction conditions, the use of two equivalents of *rac*-BINAP ligand to palladium means that to enter the catalytic cycle, ligand dissociation must occur to form a mono(diphosphino)palladium$^{(0)}$ species from the bis(diphosphino)palladium$^{(0)}$ species (Scheme 5-16). Oxidative addition reaction between **int-5.85** and aryl bromide **int-5.86** forms the Pd^{II} complex **int-5.87**, with coordination of the amine nitrogen *via* ligand exchange, producing **int-5.88**. Due to the now enhanced acidity of the amine proton(s) resulting from palladium coordination, they may be deprotonated by the base to form Pd^{II} amido complex **int-5.89** (in this case Cs$_2$CO$_3$ is sufficient), which can then undergo reductive elimination to form the desired C-N bond **5.90** with regeneration of the active Pd^0 complex.
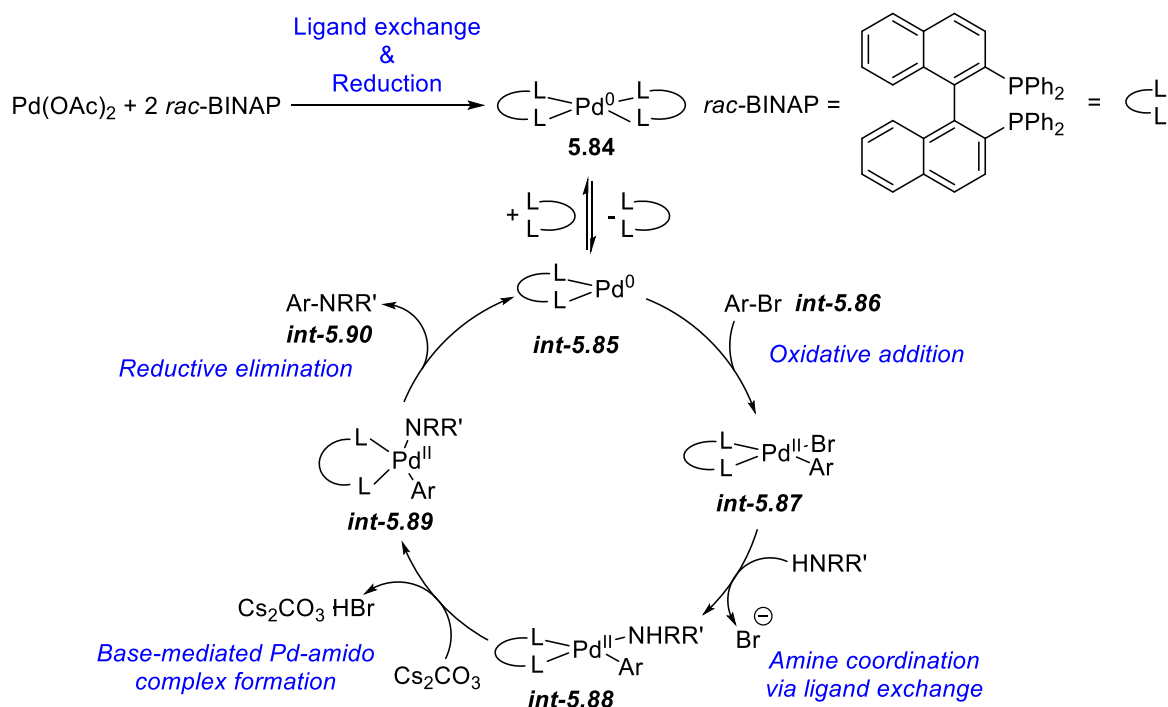
Scheme 5-16: Buchwald-Hartwig amination mechanism.

The smaller fragments of *F*-redoxal were synthesised as follows. The monomeric methyl ester fragment of *F*-redoxal **5.92** was synthesised *via* the above Buchwald-Hartwig amination conditions between **5.77** and 2-fluoroaniline **5.91** in excellent yield (Scheme 5-17).
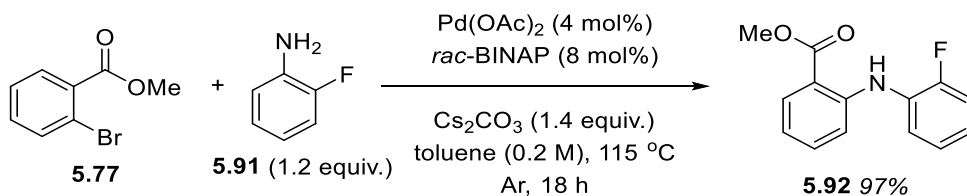
*Scheme 5-17: Buchwald Hartwig amination between **5.77** and **5.91**.*

Likewise, the methyl ester fragment **5.95** was produced by two cross-coupling reactions (Scheme 5-18). Suzuki-Miyaura cross-coupling of 3-fluoroboronic acid **5.93** and 2-fluoro-4-bromoaniline **5.73** was successful by adapting a literature procedure with dppf as the ligand.[251] This was followed by a Buchwald-Hartwig cross-coupling under identical conditions to the above procedures between the formed 4-(3-fluorophenyl)-2-fluoroaniline **5.94** and **5.77**. Both of these reactions provided good yields, for an overall yield of 51%.
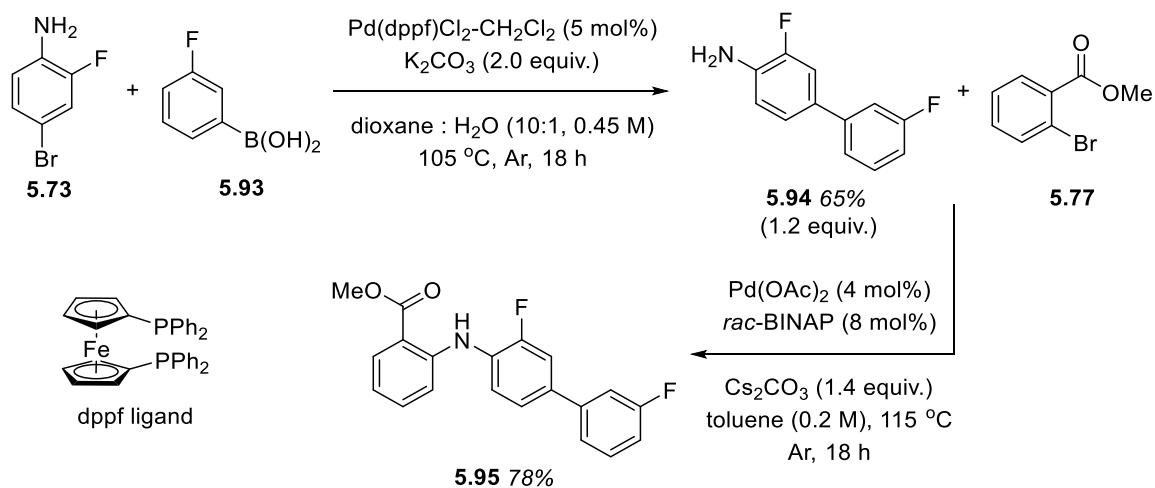
*Scheme 5-18: Synthetic pathway to **5.95**.*

5.5.2 Synthesis VIII: Saponifications

Saponification of the methyl ester products from the cross-coupling reactions allowed access to the final carboxylic acid products (Scheme 5-19). Due to issues with solubility in the initial reaction solvent, other conditions were used where appropriate to achieve hydrolysis, which meant access to **5.96** – **5.102**.
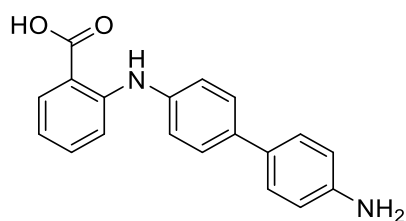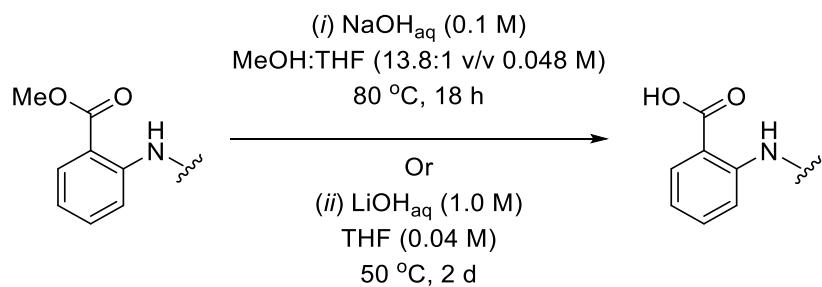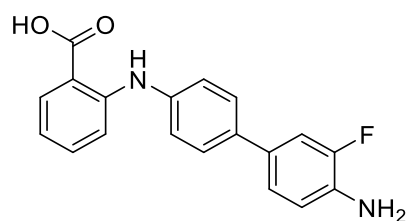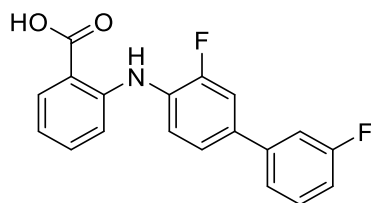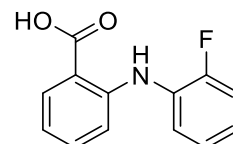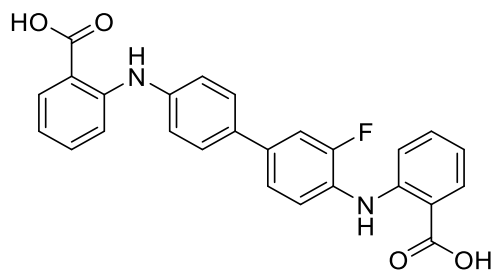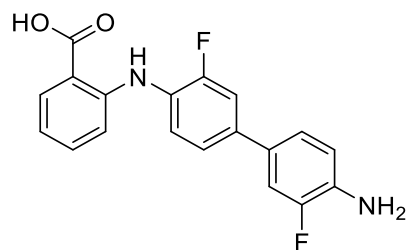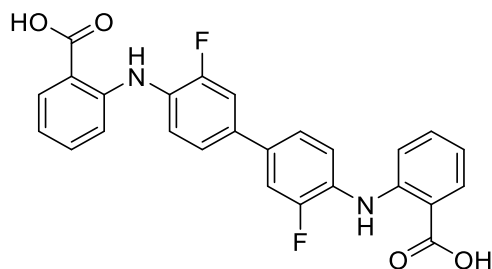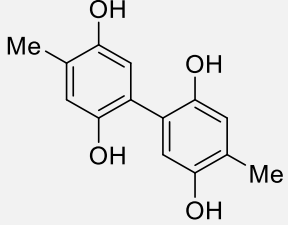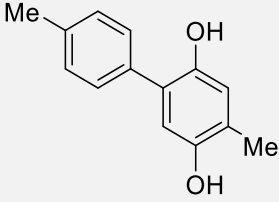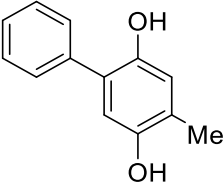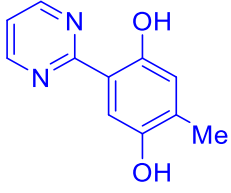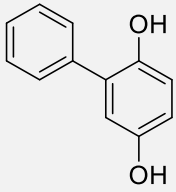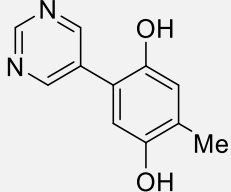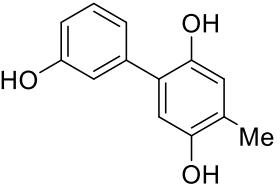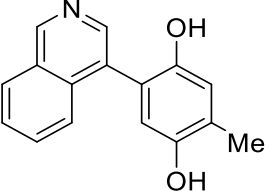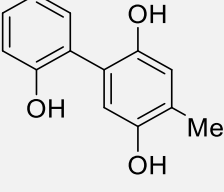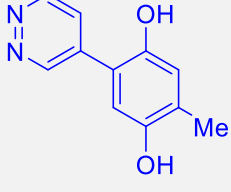
Scheme 5-19: Hydrolysis conditions and products *5.6, 5.96 – 5.101*.

5.6 Hydroquinone (NSC-2805): Biological Results and Discussion

Of the compounds successfully synthesised five were submitted for testing (highlighted in blue and numbered), with the results and all the other active compounds previously submitted presented below (Table 5-1). As this work is part of a larger body of work conducted by a previous PhD student, the synthesis of the majority of compounds tabulated has not been described.[128]

Entry	Compound structure	IC$_{50}$ (µM)	Entry	Compound structure	IC$_{50}$ (µM)
1		0.23	2		0.32 – 2.89
3		0.15 – 1.19	4 (5.30)		5.03 +/- 33.57
5		1.75 -1.80	6		0.29 +/- 0.05
7		0.11 – 1.80	8		0.80 +/- 9.53
9		0.58 – 1.18	10 (5.32)		0.42 +/- 1.33

#	Structure	Value	#	Structure	Value
11	4-Cl-phenyl, OH, OH, Me	0.48 – 13.93	12 (5.9)	N≡C, Me, OH, OH, Me, OH	0.60 +/- 1.04
13	F, F, F, OH, Me, OH	1.16 +/- 0.21	14	O=CH, Me, OH, OH, Me, OH	2.59 +/- 0.31
15	S (thiophene), OH, Me, OH	19.78 +/- 5.64	16	OH, Me, OH, OH, Me, OH	6.68 +/- 0.71
17	S (thiophene), OH, Me, OH	15.25 +/- 2.12	18	O=, NH₂, Me, OH, OH, Me, OH	2.95 +/- 0.65
19	N (pyridine), OH, Me, OH	1.20 +/- 2.01	20 (5.8)	O=, OH (COOH), Me, OH, OH, Me, OH	0.64 +/- 0.93
21	N, N (pyrimidine), OH, Me, OMe	11.1 +/- 1.56	22	F, F, F, O=, NH (pyridinone)	131.1 – 400.3
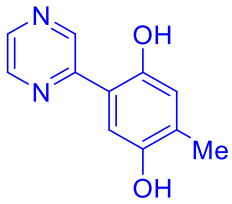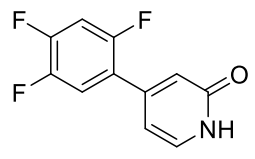
23 **(5.31)**		0.26 +/- 1.09	**24**		327.3 – 458.7

Table 5-1: Biological results of NSC-2805 series.

All of the submitted compounds are shown to be active compounds with the most active being **5.31** with an IC_{50} value of 0.26 μM (entry 23), although the value was imprecise compared to other compounds of this series. The compounds synthesised for the purpose of exploring the possible extra hydrogen bonding exhibit similar ranges (0.6 – 1.04 μM and 0.64 – 0.93 μM). Compared to the other compounds within this series, IC_{50} scales from nitrile = carboxylic acid < aldehyde < amide < benzylic alcohol.

The compounds synthesised for the second objective (exploration of π-π interactions) display much more varied IC_{50} ranges (0.26 – 33.57 μM). The thiophene regioisomers display worse results than for 6-membered nitrogen heterocycle bearing molecules (entries 15 & 17 *vs.* entries 4, 6, 10, and 23). This could be rationalised as the hypothesis holding, that more-electron-deficient aromatics are capable of stronger π-π stacking interactions with the phenolic moiety of Tyr587 within the proposed binding site of WWP2. In contrast, electron-rich aromatics like thiophene will have less-favourable interactions. Unfortunately however, all of the derivatives submitted were not as active against WWP2 compared to NSC-2805 (0.23 μM), which has as its ring B substituent an electron-rich *para*-hydroquinone motif, which contradicts this point. This highlights only a potential understanding of an SAR with these compound series, as the associated reactivity of the *para*-hydroquinone motif rules out an understanding that these IC_{50} values are a true reflection of the activity of these compounds.

A major cause for concern in the testing of these compounds was raised because all the compounds found to be active in the low micro-molar range had present a *para*-hydroquinone motif, its oxidised quinone moiety is increasingly known to cause pan-assay interference (PAINs).[252,253] PAINs are motifs which appear on multiple assays as false positives, which are often reactive with elements of the assay itself, and represent un-optimisable hits in hit-to-lead optimisation.[28,254] Many of these compounds had their methoxy-protected analogues submitted for testing, which were found to not be soluble in the assay (not shown). The partially methylated compound submitted had slightly worse results (entry 21) than similar diazine heterocycle bearing the free *para*-hydroquinone derivatives (entries 4, 6, 10 & 23), and is instead has a closer IC_{50} to the thiophene heterocycles, perhaps because oxidation and further reactivity is slower. Lastly, entries 22

and 24, which don't contain any *para*-hydroquinone motif display significantly worse IC$_{50}$ values compared to anything else in the series.

Pan-assay interference was confirmed for NSC-2805 where it was soaked during the auto-ubiquitination step (normal assay step) and afterwards, with the results showing assay interference. Additionally, NSC-2805 was found to be reactive towards E1 in the bioassay from a cross-reactivity screen (data not shown).

Although the results confirm NSC-2805's PAINs properties with respect to interference with the assay setup and methodology, the crystal data also suggests that it can somewhat inhibit WWP2 activity. This gives reason to believe that the IC$_{50}$ values for the compounds bearing the hydroquinone motif are inaccurate but not wholly a result of assay interference, as there is potential that this motif is interacting with components of the assay.

5.6.1 <u>Hydroquinone (NSC-2805): Bioisosteric Investigation</u>

From these results it was decided that compounds bearing unreactive isosteres of the *para*-hydroquinone motif oxygen atoms should be synthesised and tested, as this should remove any reactivity associated with the hydroquinone whilst still exhibiting some inhibition in the assay (Figure 5-10). Compound **5.102** was targeted based on of some preliminary molecular docking studies.[128] Compounds **5.103** – **5.104** are bound to an isoquinoline structure, as this was shown to be a relatively active structure previously (Table 5-1, entry 16). Additionally, a *meta*-dihydroxyphenyl variant bearing a carboxylic acid was also targeted based on docking analysis. These targets should have higher IC$_{50}$ values compared to the hydroquinone-containing molecules, displaying both the fact that the hydroquinone motif is reacting with the assay and altering the IC$_{50}$ values, and that these compounds are still inhibitory against WWP2.
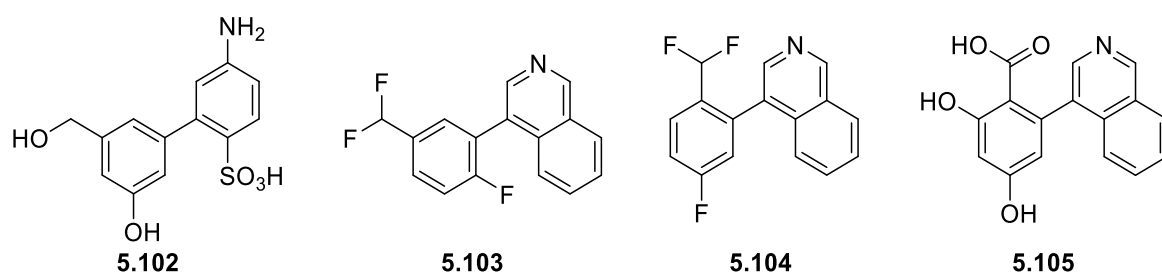
Figure 5-10: Targeted isoquinoline compounds.

A difluoromethyl motif was chosen to mimic a heteroatom positioned where an oxygen atom was (Figure 5-11). The difluoromethyl group acts as a lipophilic isostere of an aniline

group, rather than a phenol.[255] The hydrogen atom neighbouring the two fluorine atoms will be significantly electron-withdrawn, and would be able to participate in hydrogen bonding, somewhat analogous to protons bound to electron negative heteroatoms such as nitrogen and oxygen.
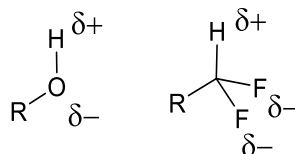
Figure 5-11: Comparison of hydroxyl vs. difluoromethyl groups dipoles for bioisosteric replacement.

To begin, two regioisomeric 4-fluoro-(2 or 3)-bromobenzaldehydes were treated with DAST to provide the difluoromethyl-fluoroaryl bromides **5.106** and **5.107** in quantitative yields (not shown).[128] Next, Miyaura borylations were achieved on both isomers to provide **5.108** and **5.109** (Scheme 5-20). Afterwards, Suzuki cross-coupling was used to couple 4-bromoisoquinoline **5.110** and access both regioisomers **5.103** and **5.104**, which could be fully characterised, purified and submitted for biological assay.
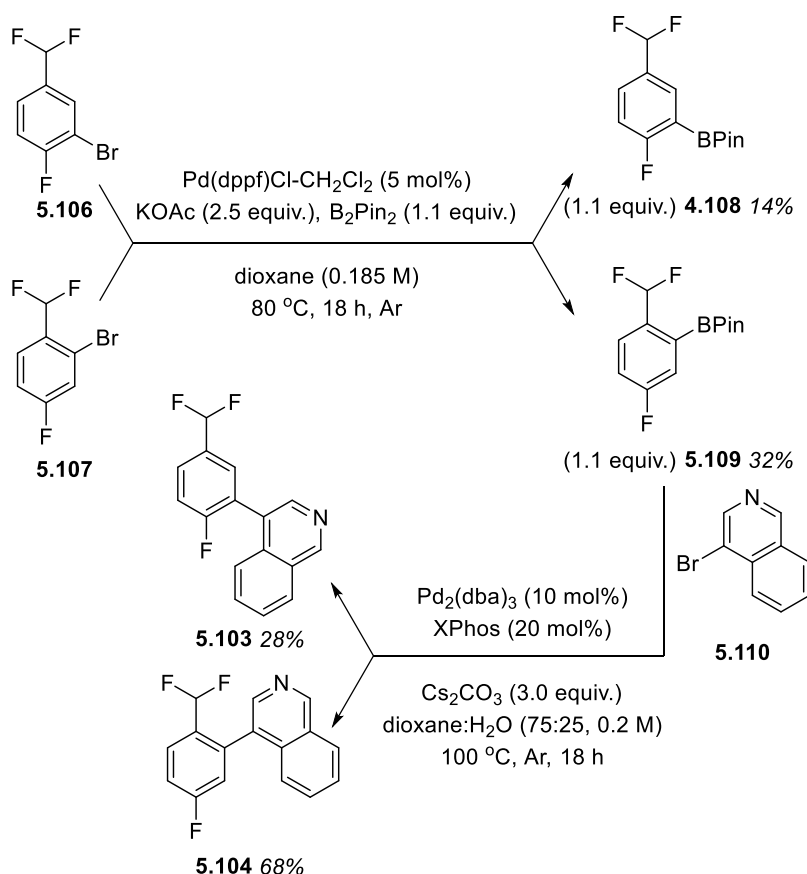
*Scheme 5-20: Miyaura borylation followed by Suzuki-Miyaura cross-couplings to access **5.103** & **5.104**.*

The Miyaura borylation is a useful tool in a large toolkit of borylation reactions. Mechanistically, the borylation process is thought to proceed as follows, utilising a bidentate ligand.[256] Initially, after formation of the Pd⁰ species **int-5.111** (as discussed above), oxidative addition occurs with **int-5.112** to form **int-5.113** (Scheme 5-21). The acetate group of K_2CO_3 displaces the halide bound to the palladium centre to form acetoxy intermediate **int-5.114**, for which there is experimental evidence for. There is not experimental evidence for initial coordination of the acetate with B_2Pin_2. This species then undergoes a transmetalation with B_2Pin_2 **5.115** to form **int-5.117** with release of by-product **int-5.116**. Reductive elimination of **int-5.117** occurs afterwards, to provide the borylated product **5.118** and the reformed catalyst **int-5.111**. The use of K_2CO_3 is essential in this reaction in order to not form diarylated products and to accelerate the reaction. This rate increase is thought to come from increasing the rate of the transmetalation step because of both the high reactivity of the Pd-O bond, which is a combination of soft Lewis acid and hard Lewis base, and the high oxophilicity of the boron centre.
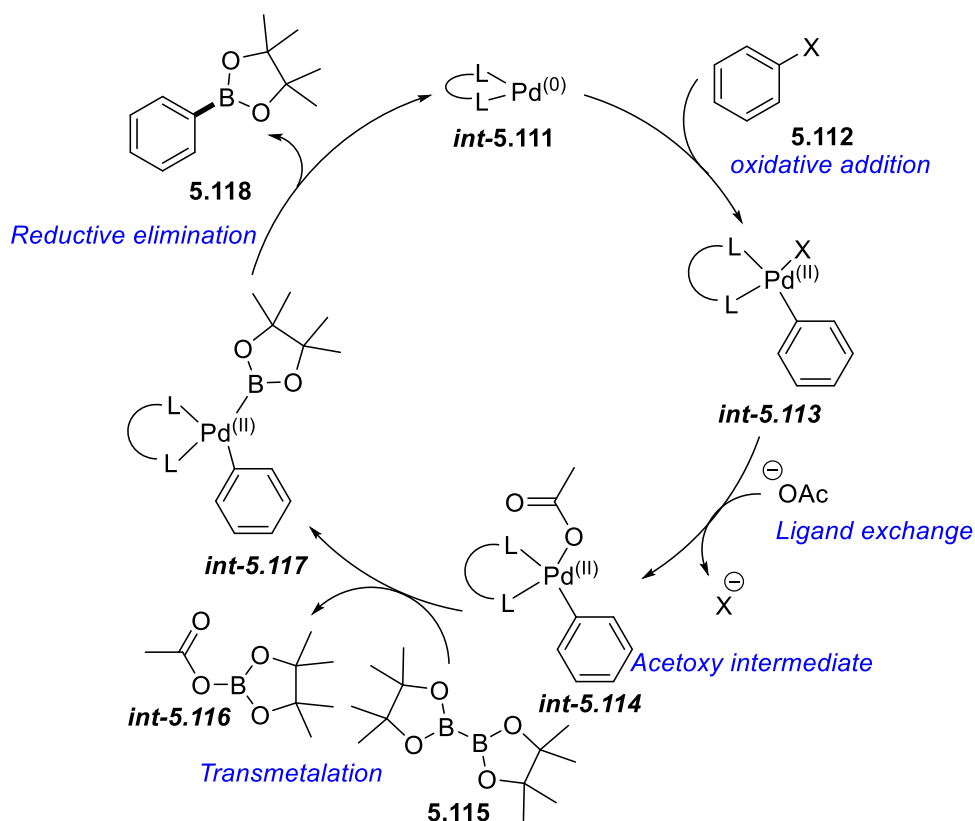
Scheme 5-21: Mechanism for the Miyaura borylation.

Interestingly, 4-(2-difluoromethyl-5-fluorophenyl)isoquinoline **5.104** displays chemically non-equivalent diastereotopic fluorine atoms in the difluoromethyl motif, as observed by {¹H}¹⁹F NMR (Figure 5-12) which shows suggestive ^2J F-F coupling constants of 303.3 Hz,

and perhaps ^6J F-F coupling between the difluoromethyl and *para*-fluorine atoms of 3.2 Hz and 3.9 Hz. ^{1}H-^{19}F HOESY suggests some form of specific configuration exists, as only one fluorine signal from the difluoromethyl motif displays a spatial interaction with the signal for the hydrogen atom located between the aryl-aryl linkage and the nitrogen atom of the isoquinoline system, as indicated in Figure 5-12. This likely stems from a restricted rotation about the aryl linkage highlighted in Figure 5-12, giving rise to atropisomerism within this molecule, but not its regioisomer, 4-(5-difluoromethyl-2-fluorophenyl)isoquinoline. Unfortunately, both compounds were also insoluble in the bioassay.
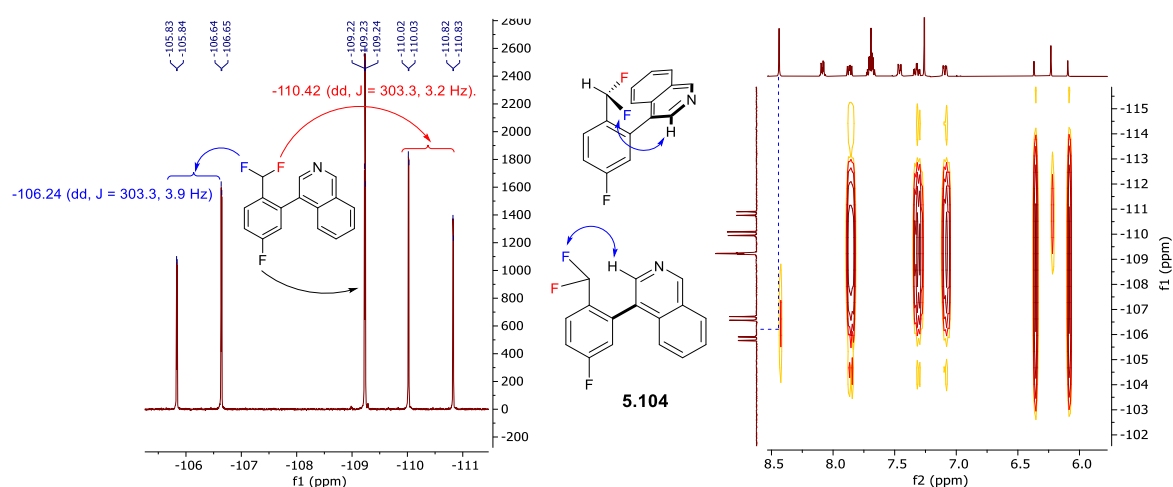
*Figure 5-12: {^{1}H}^{19}F NMR (CDCl$_3$, left) of **5.104**, middle: visual (top = 3D, bottom = 2D) representations of through space interactions observed from ^{1}H-^{19}F HOESY NMR (CDCl$_3$, right).*

For the isoquinoline derivative **5.104**, work was started where 3,5-dimethoxybromobenzene **5.119** was treated under Vilsmeier-Haack conditions to formylate the 2-position exclusively in 73% yield providing **5.120** (Scheme 5-22). Simple treatment with potassium permanganate oxidised the aldehyde to the carboxylic acid **5.121**, which was methylated in high yield, producing **5.122**. From here, another Miyaura borylation and Suzuki cross-coupling sequence was planned similar to the fluorinated derivatives to access **5.125** through **5.124**, the product of which could be demethylated and saponified to access the final target **5.104**. In the interests of time however, focus shifted away from both this synthesis and the synthesis of **5.102**.
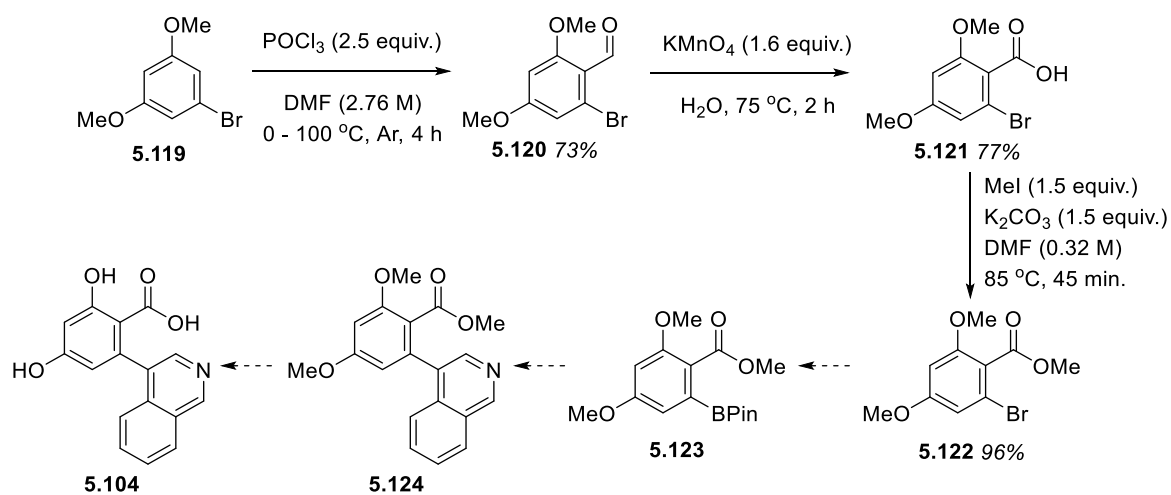
*Scheme 5-22: Incomplete synthesis of **5.104** starting from **5.119**.*

5.7 *F*-Redoxal Biological Results and Discussion

Overall, in addition to the carboxylic acid final products, methyl ester and benzidine intermediates were also submitted for bioassay (Table 5-2). Below is a summary of the submitted compounds and biological results. The NCI sample provided an IC_{50} value of 47.7 µM from the HTS.

Entry	Structure	IC_{50} (µM)	Entry	Structure	IC_{50} (µM)
1 (5.76)		N/A	2 (5.95)		N/A*
3 (5.75)		N/A	4 (5.80)		N/A*
5 (5.74)		N/A**	6 (5.78)		N/A
7 (5.92)		N/A*	8 (5.100)		N/A*
9 (5.99)		N/A**	10 (5.98)		368.5*
11 (5.83)		N/A*	12 (5.97)		N/A*
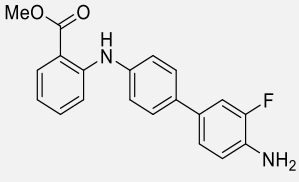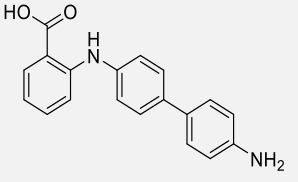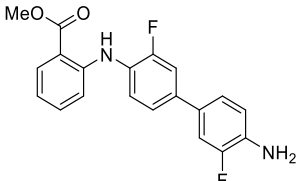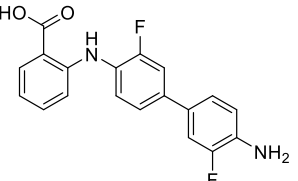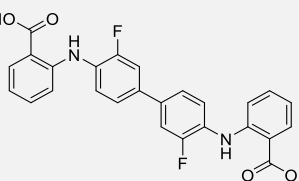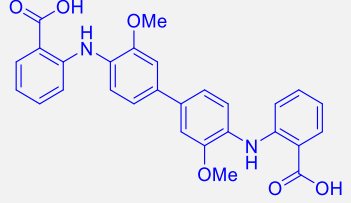
13 **(5.81)**		N/A*	**14** **(5.96)**			N/A*
15 **(5.79)**		N/A*	**16** **(5.101)**			N/A*
17 **(5.6)**		N/A*	**18** **(5.2)**			N/A*

*Table 5-2: Submitted compound for bioassay from the F-redoxal series. *Insolubility at 10 mM (10% DMSO). **Partial insolubility at 10 mM (10% DMSO).*

Practically all compounds were inactive in the bioassay, apart from entry 10 (**5.98**, Table 5-2) with a worse IC_{50} value of 368.5 µM compared to the original hit value of 47.7 µM. This is a strange result considering the NCI sample provided an IC_{50} value. It is also surprising that the majority of compounds synthesised here do not dissolve in the assay media, and the NCI sample of redoxal does, considering it has a log*P* of 5.95.[226] It was assumed that since redoxal provided an IC_{50} value in the HTS that it dissolved and therefore the similar fragment compounds would behave in similar fashion and provide some form of activity. However, also included (entry 18) is the resynthesised redoxal from section 3.3.4, which was unable to reproduce the IC_{50} value from the HTS and was also not soluble at 10 mM (10% DMSO). This discrepancy between the NCI sample and resynthesised material meant analysis of the NCI sample purity was undertaken (section 5.7.1).

In addition to IC_{50} assays, differential scanning fluorimetry (DSF) was also performed (Table 5-3), revealing that entries 10, 11 & 12 (Table 5-2), provide some thermal shift whilst the rest of the compounds tested do not. This is somewhat structurally in line with IC_{50} data in that entry 10 has also provided an IC_{50} value. This may point towards these *N*-biphenyl-*N*-phenyl scaffolds ability to interact with WWP2, even though the other two compounds with this structure did not provide IC_{50} values.
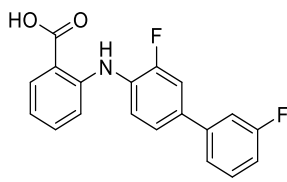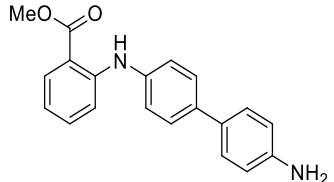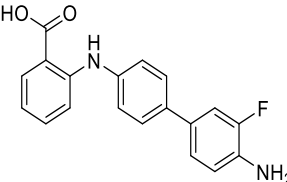
Structure			
DSF result	Destabilising	Destabilising	Stabilising

Table 5-3: DSF results for the F-redoxal series.

5.7.1 NMR & HRMS Purity Analysis of the NCI Hit Sample

5.7.1.1 NMR Analysis

As redoxal was resynthesised as part of the objective of this research into gaining a co-crystal structure with WWP2 (see section 3.3.4 for details), but also did not provide any IC_{50} data due to insolubility in the assay. This indicates some issue with the purity of the original sample used for HTS. ^{1}H NMR analysis of this sample provided an essentially identical aromatic region to the re-synthesised redoxal sample (Figure 5-13).
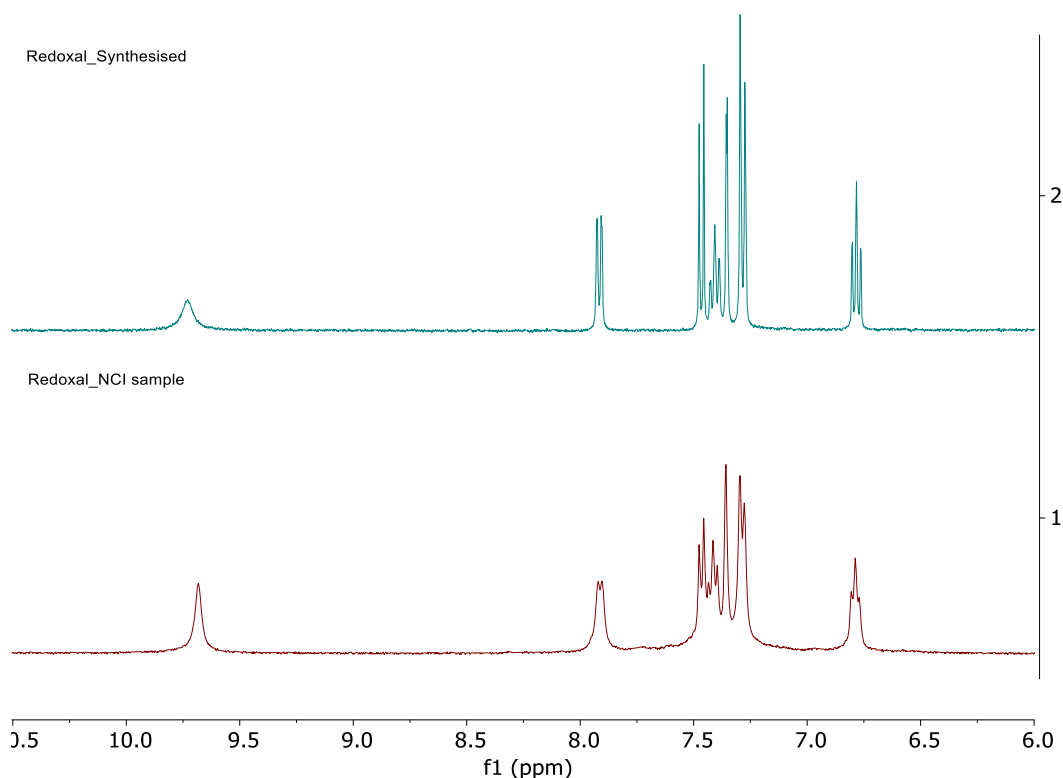
Figure 5-13: Comparison of resynthesised redoxal vs. the NCI original sample of redoxal.

The aliphatic region does contain some impurities, as shown below (Figure 5-14). The DMSO is expected as not all the DMSO was removed without heating *in vacuo*., which

288

may have affected the sample. Impurities at 2.99 ppm, 3.41 ppm, 3.48 ppm and 4.55 ppm could not be attributed to anything common as impurities (the common NMR impurities table was consulted).[157] Additionally, the DMSO-d_6 used was fresh and did not contain any impurities. Compared to the original sample, the molar equivalencies of these impurities are quite high, sometimes greater than 1:1 (for the signals at 3.41 ppm and 3.48 ppm, assuming they are single protons). It doesn't seem that these impurities are of a single compound as their integrations do not match, although the signal at 4.55 ppm seems like a broad singlet and the peak at 3.41 ppm is bordering on the H_2O signal, so the integrations are skewed and therefore inaccurate. This was seen as a cause for concern as I was unable to identify what the impurities are and in what quantities they are present.
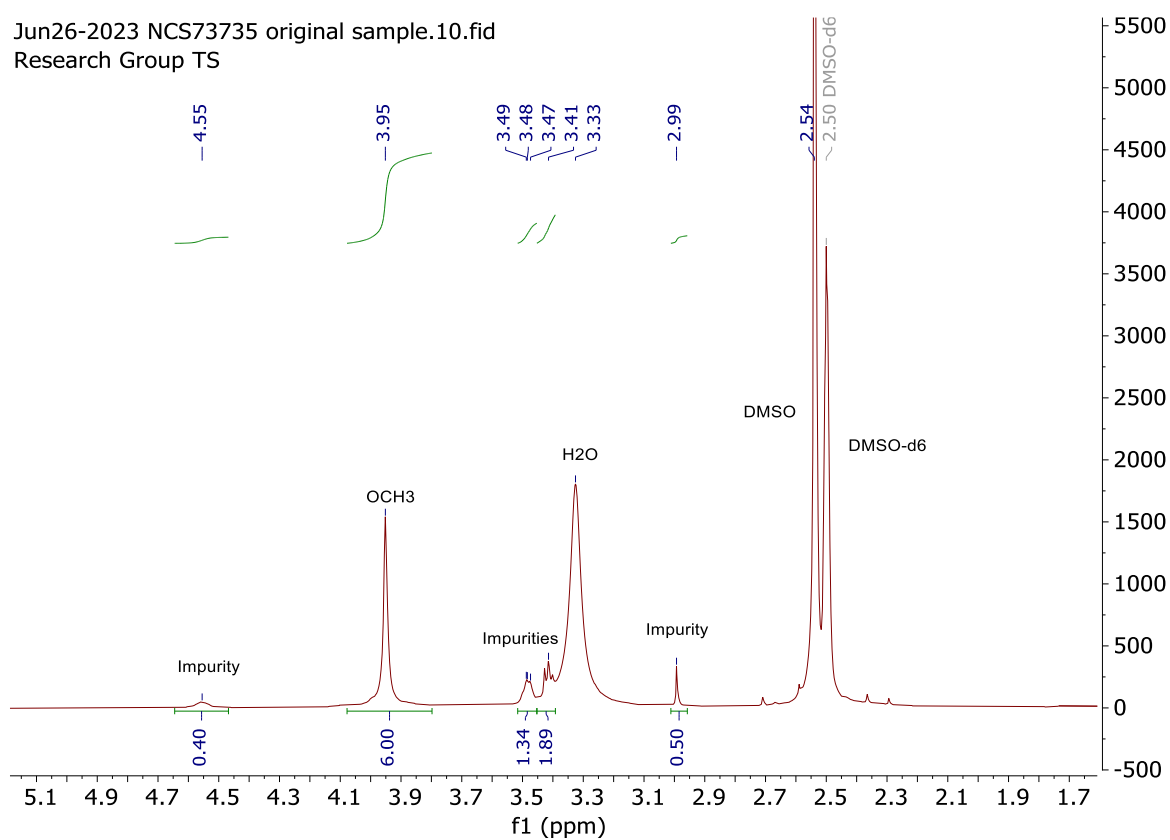
Figure 5-14: ^1H NMR spectra up-field region of the NCI original sample of redoxal.

5.7.1.2 Visual Analysis

A visual inspection of the two NMR samples side-by-side offers a rather striking observation: something is certainly different about the two samples (Figure 5-15). The resynthesised version is a transparent yellow solution (right) and the NCI diversity set VI sample, an opaque black solution (left).
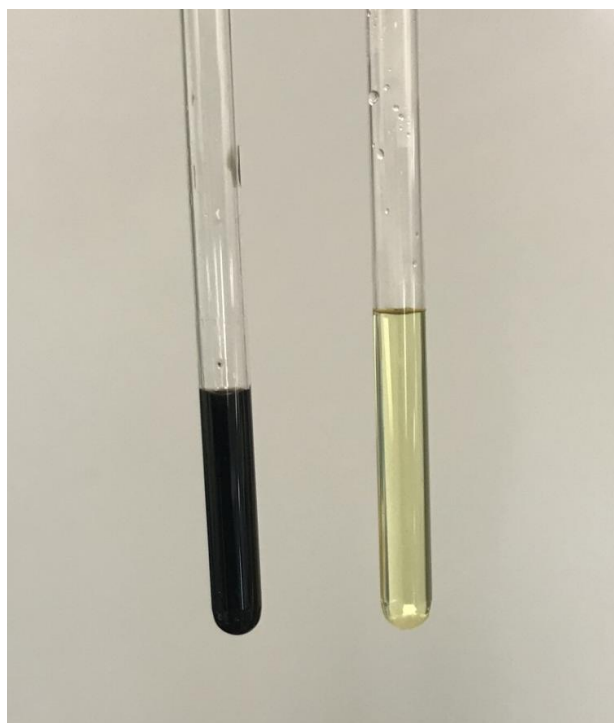
Figure 5-15: Left: NCI original sample of redoxal. Right: re-synthesised sample of redoxal.

5.7.1.3 UPLC-HRMS Analysis

It is important to note that black material was filtered out of solution during the HRMS sample preparation stage as it was not soluble in MeCN used for sample preparation. Presented below are the chromatograms of both the NCI sample of redoxal and the blank (Figure 5-16) from subjecting the samples to UPLC-HRMS. These samples had the same purity issues as those of NSC-217913 in section 4.7.2. Comparing the two it is clear that there is an additional peak at 3.47 min., and several between 2.60 and approx. 3.10 min. There is additionally a peak at 3.64 min. which seems to have a shoulder at approx. 3.62 min. so the area in between 3.50 and 3.64 is worth investigating further. The peaks that are practically identical are not investigated because they are most-likely the impurities present in the sample from sample preparation. Additionally, the areas investigated that do not show any observable differences are not presented.
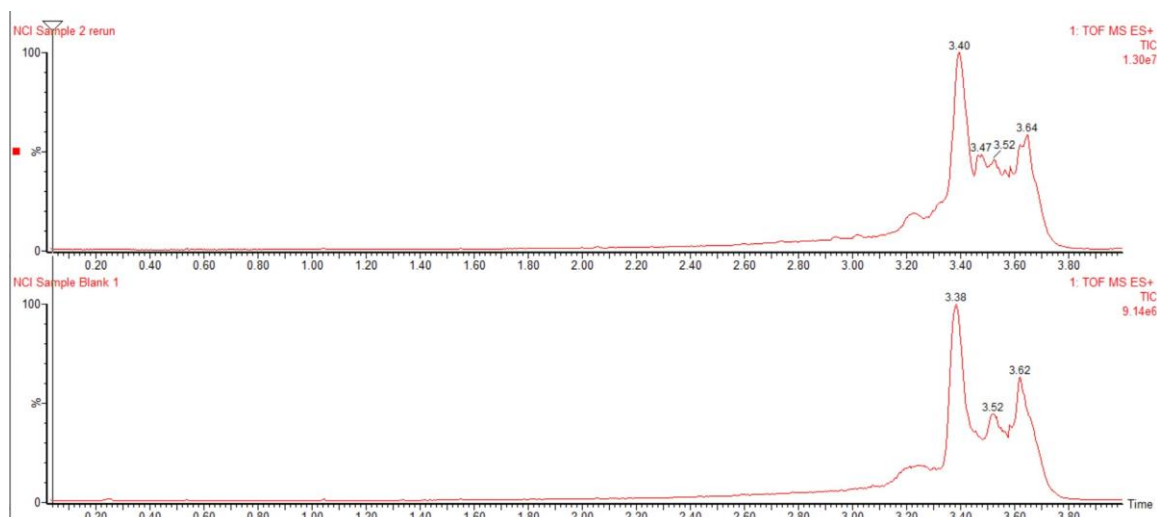
Figure 5-16: Chromatograms of the NCI sample of redoxal (top) and the blank sample (bottom).

The peak at 3.47 min. corresponds to the (M+H) peak of redoxal with an *m/z* of 485.1939 (Figure 5-17, bottom), which when compared to the blank (top) is not present, as expected.
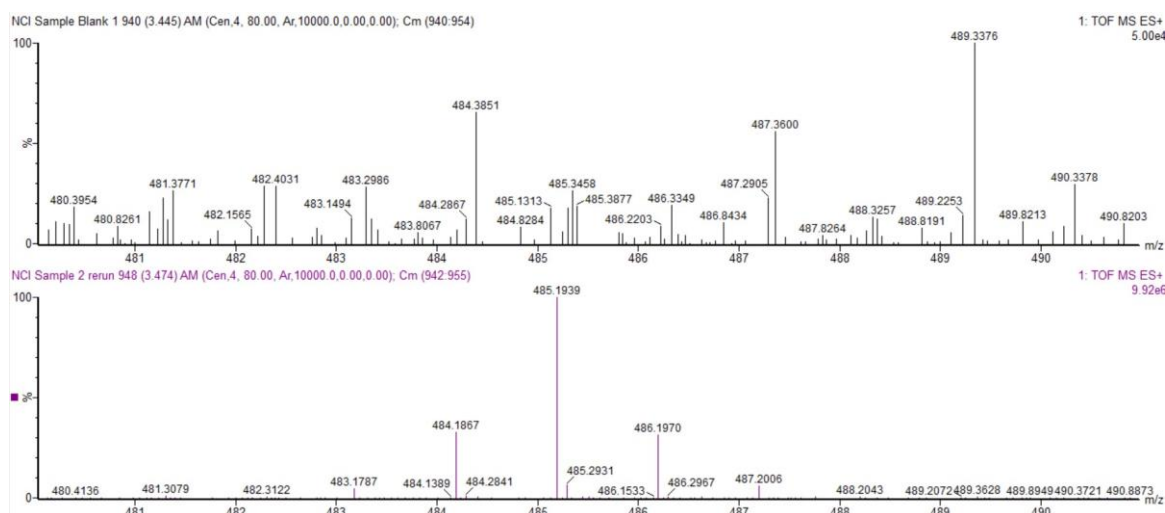
Figure 5-17: Mass spectrums of blank sample at 3.445 min. (top), and NCI redoxal sample (bottom).

Next, the signals at 2.77, 3.01, 3.58 min. that were identified as differences in the chromatograms were investigated. Firstly, as can be seen in Figure 5-18, *m/z* peaks of 214.9180 and 238.8849 – 239.1271 in the blank sample can be identified in the NCI sample, but at lower intensities. The peak at 231.1106 *m/z* seen in the NCI sample is not present in the blank sample and has been identified as a potential impurity within the NCI sample.
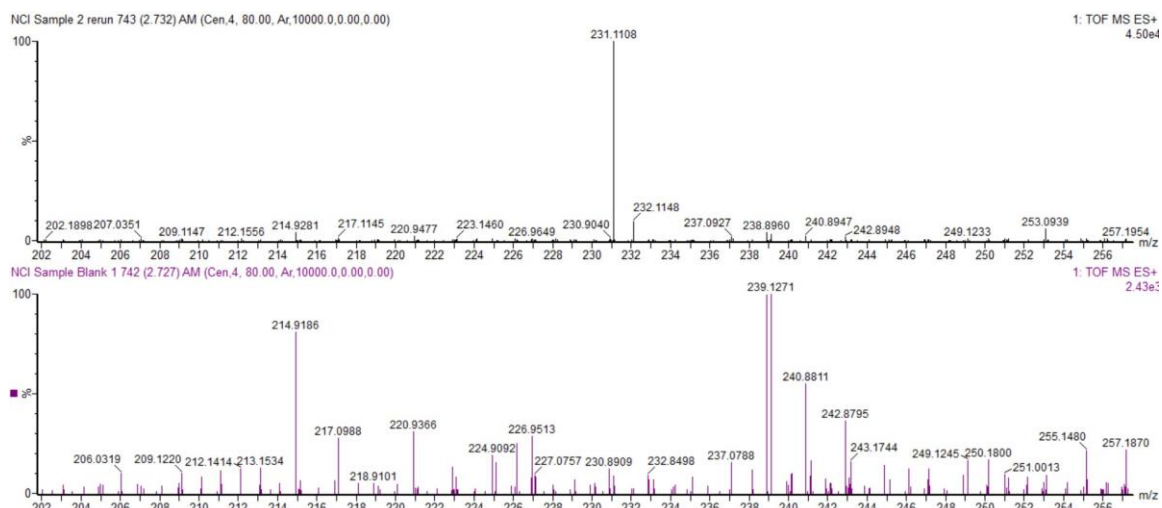
Figure 5-18: Mass spectra of blank (bottom) and NCI sample (top) at 2.7 min.

Next, the masses at 3.01 min. pointed towards potential impurities at 375.1525, 393.1637, 406.3487, 415.1472 *m/z* (Figure 5-19). Peaks at 367.2641, 388.4133 and 437.2158 are thought to correspond to the peaks in the blank sample at 367.2463, 388.3943 and 437.2625 *m/z*. The signal at 415.1472 may correspond to the M+Na ion of 393.1637
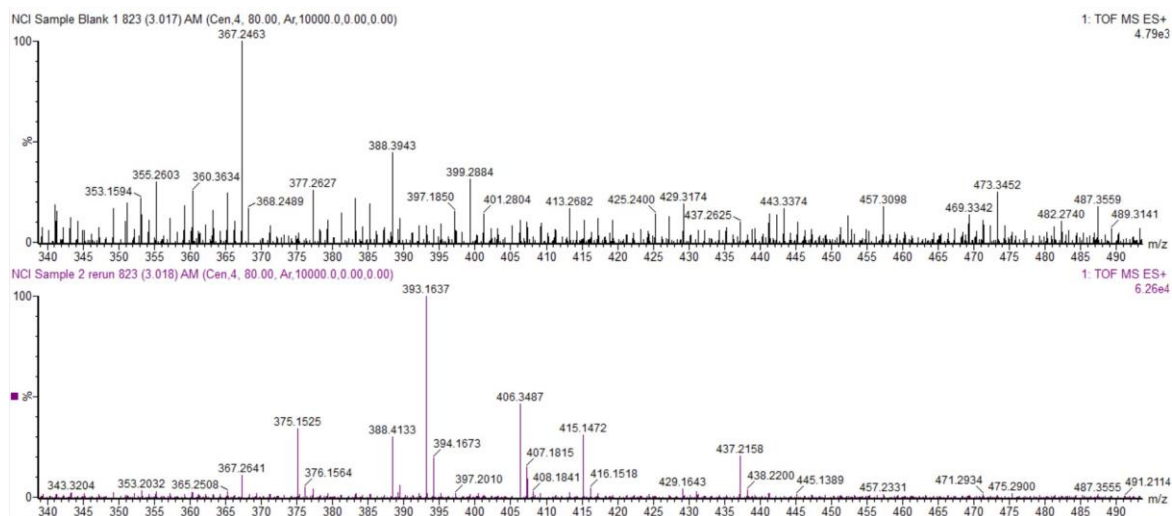
Figure 5-19: Mass spectra of the blank sample (top) and the NCI redoxal sample (bottom) at 3.017 min.

Finally, investigating the differences in the chromatograms at 3.58 min. found two signals at 254.2575 and 276.2438 *m/z* which were not present in the blank sample (Figure 5-20). Its thought these correspond to the same impurity as they are the M+H and M+Na signals, respectively.
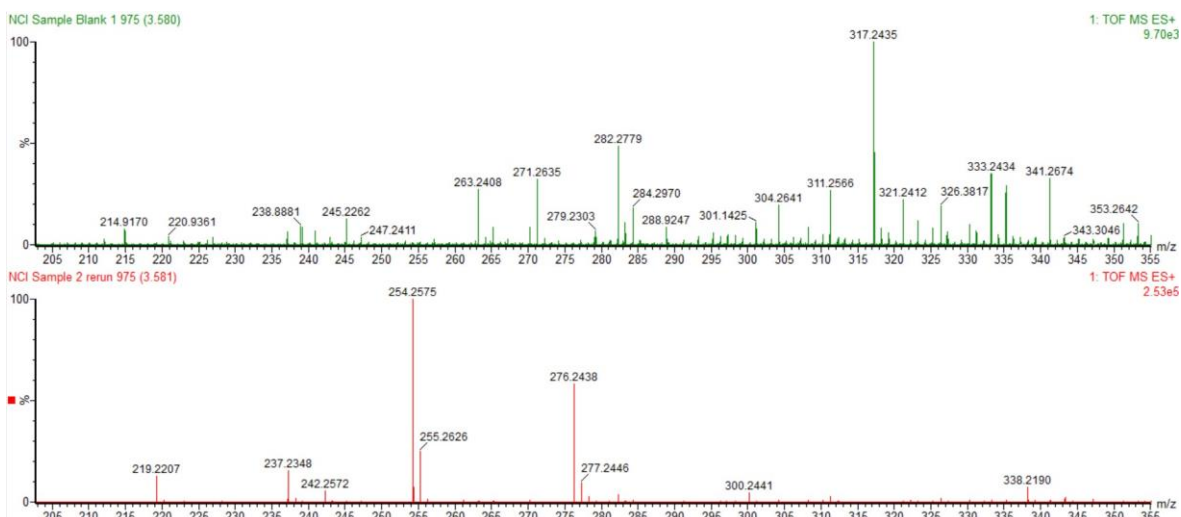
Figure 5-20: Mass spectra of the blank sample (top) and NCI sample of redoxal (bottom) at 3.58 min.

To summarise, the *m/z* signals found that may be impurities in the NCI sample are 375.1525, 393.1637, 406.3487, 415.1472, 254.2575 and 231.1108. It is difficult to identify what these impurities are since no literature synthesis is known, and none point to derivatives of redoxal. Using a fragmentation approach and comparison against databases using Waters UNIFI software would improve the chances of identification of potential impurities. However, in the interests of time the identification of these impurities was not pursued further. Overall, the NCI sample of redoxal was not pure and the potential impurities present, whatever they are, are probably what provided the NCI sample activity in the HTS.

5.8 <u>Hydroquinone (NSC-2805): Conclusions</u>

Based on the above results and previous discussion, gaining an accurate understanding of the structural features underlying inhibition from this biaryl system proved difficult. Whilst the synthesis of these compounds was successful, perhaps a prior warning of the reactivity of the *para*-hydroquinone motif was overlooked during isolation of the final products. After the last reaction on each individual derivative, which always consisted of deprotection of the 1,4-dimethoxyphenyl systems **5.126**, reductive workup with sodium dithionite was required as it was observed on several occasions that these compounds form highly-coloured precipitates in the absence of the sodium dithionite workup. It was later found that this is a result of quinhydrone formation **5.129** (Scheme 5-23).[257]
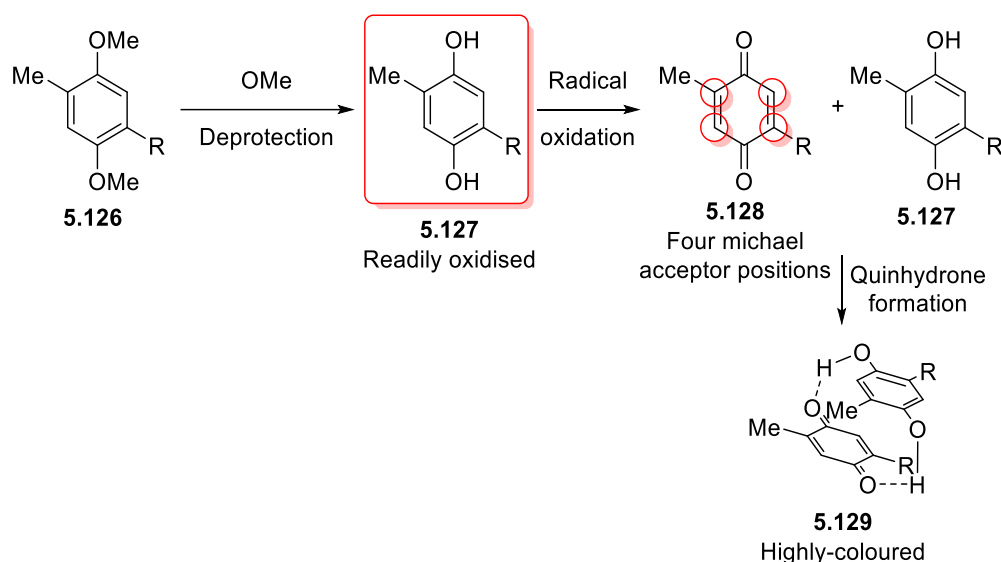
Scheme 5-23: Reactivity associated with the para-hydroquinone motif.

These quinhydrone products form due to radical oxidation of the *para*-hydroquinone motif **5.127** into the *para*-quinone product, which upon interaction with another unoxidized molecule of *para*-hydroquinone, form strong non-covalent hydrogen bonding and π-π staking interactions. Not only does this show the ready reactivity of these species, but the production of the well-known PAINs quinone moiety **5.128**, which has present on its structure four Michael-acceptor positions capable of further reaction.

Producing bioisosteres was partially successful, as from the compounds produced (entries 22 & 24, Table 5-1), predicably worse IC_{50} values were obtained. These compounds that moved away from the *para*-hydroquinone motif also moved away from its inherent reactivity, including the partially methoxy protected derivative (Entry 11, Table 5-1), and so provided much worse IC_{50} values. This reinforces the idea that Hit PAIN structures are inherently unoptimizable and these results actually describe a structure interference relationship.[258]

Starting from structures containing PAINs in rational drug design is contradictory to the rational strategy implemented in hit-to-lead optimisation. As soon as movement away from the reactive motifs to stable structures occurs (in order for SAR to be developed and rational changes to the structures of these molecules implemented), the activity worsens as the readout observed from the reactivity is no longer present, and the overall process is not one of rational optimisation. In some cases, it may be that these compounds are only reactive, and not interacting with the biological target at all. In this case the crystal data provides some understanding that NSC-2805 does indeed interact with WWP2. Because of the PAINs issue surrounding this compound series, further investigation was not

pursued as improving upon the hit IC_{50} and further optimising the structure to more active derivatives was unsuccessful.

5.9 *F*-Redoxal: Conclusions

Based on the discussion surrounding the *F*-redoxal fragment library, whilst the objective to synthesise a library of fragment derivatives of *F*-redoxal was met as a library of 18 compounds was successfully synthesised, the biological results proved disappointing as only limited biological data could be gained.

The HTS sample of redoxal provided an IC_{50} value, which was assumed would translate to both the fragments being soluble in the assay and also provide some IC_{50} data. But, subjecting the resynthesised redoxal from section 3.3.4 to bioassay showed that it was unable to dissolve at 10 mM (10% DMSO) as well as being unable to provide any IC_{50} data. This meant the assumption was wrong and that something else must be providing the NCI sample IC_{50} data observed from the HTS.

From the bioassay results, the single IC_{50} value gained for **5.98** may mean it is a true inhibitor, as this compound was additionally found to be interacting in some way with WWP2 *via* a DSF experiment (Figure 5-21). It could be that whilst this compound along with the other two are interacting in some way with WWP2 as evidenced by the DSF results, due to their lipophilicity alone they may be 'sticking' to WWP2 and are not binding in a productive fashion. The aim of producing a fragment of *F*-redoxal with similar activity was not reached as this fragment has much-reduced activity (IC_{50} = 368.5 *vs*. 47.7 µM) . Although, compound **5.98** should be investigated further with additional experiments to confirm its interaction with WWP2 (*e.g.,* E1 counter-screens, STD NMR) and may offer a different starting point for lead design, albeit from a low activity and relatively high molecular weight (325.21 gmol^{-1}) starting point.
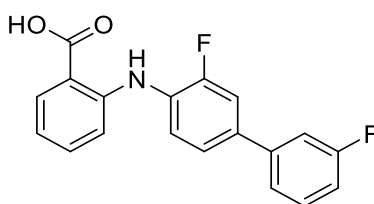
*Figure 5-21: Compound **5.98** with potential for further investigation.*

The general fragment inactivity along with the inability to reproduce the IC_{50} value gained from the NCI sample led to the examination of the NCI sample purity, which provided the rather clear result that the sample contained impurities. The exact structure of the impurities is currently unknown but considering the results from subjecting resynthesised

redoxal to the bioassay, it is predicted that the obtained IC_{50} value from HTS originated from the impurities somehow interacting with the assay at some point, and not redoxal.

5.10 Experimental

5.10.1 Hydroquinone

3-(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-6-methylbenzoic acid
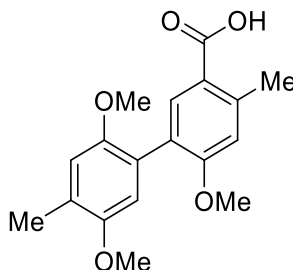
Crude material from a Suzuki cross-coupling was provided by Mr Joshua Hall. 3-(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-6-methylbenzoic acid was purified *via* column chromatography (2:3 Et_2O:Hexane) as a white solid.

^{1}H NMR (500 MHz, Acetone-d_6) δ 10.82 (br s, 1H), 7.86 (s, 1H), 6.96 (s, 1H), 6.88 (s, 1H), 6.76 (s, 1H), 3.82 (s, 3H), 3.78 (s, 3H), 3.67 (s, 3H), 2.65 (s, 3H), 2.22 (s, 3H).

^{13}C NMR (126 MHz, Acetone-d_6) δ 168.4, 160.8, 152.4, 151.9, 143.0, 135.4, 127.0, 126.4, 125.7, 121.9, 115.5, 114.8, 114.5, 56.6, 56.2, 56.0, 22.4, 16.4.

IR (cm^{-1}): 3080 (OH), 1673 (C=O).

M.P. 184.8 °C (dec.).

MS ES+ m/z Calcd for $C_{18}H_{20}O_5$ (M+H)$^+$: 319.1448, found: 319.1447.

3-(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-6-methylbenzamide
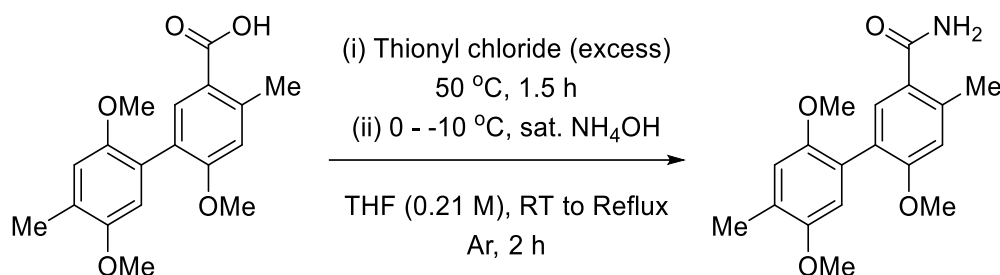
To a 10 mL RBF was added 3-(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-6-methylbenzoic acid (0.20 g, 0.633 mmol) and dry THF (3 mL, 0.21 M). To the solution was added thionyl chloride (5 drops, excess) and heated at 50 °C for 1.5 h. The mixture was allowed to cool to RT, then cooled to 0 - -10 °C (acetone, ice) poured into a cooled, stirred solution of ammonium hydroxide (20%, 3 mL) and left for approx. 5 min. The mixture was then transferred to a separatory funnel and extracted with DCM (4x 20 mL), dried (Na$_2$SO$_4$) and the solvent removed *in vacuo.* Purified *via* column chromatography (Et$_2$O -> 3% TEA in Et$_2$O) to yield a white solid, 3-(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-6-methylbenzamide (108 mg, 0.34 mmol, 54%). The aqueous layer was acidified to afford a white precipitate, confirmed to be starting material (30 mg, 15%).

^{1}H NMR (500 MHz, Acetone-d_6) δ 7.36 (s, 1H), 6.96 (s, 1H), 6.99 (br s, 1H, *NH*), 6.86 (d, *J* = 0.7 Hz, 1H), 6.75 (s, 1H), 6.41 (br s, 1H, *NH*), 3.78 (s, 3H), 3.77 (s, 3H), 3.67 (s, 3H), 2.52 (d, *J* = 0.7 Hz, 3H), 2.21 (s, 3H).

^{13}C NMR (126 MHz, Acetone-d_6) δ 171.1, 158.9, 152.4, 151.9, 148.2, 138.7, 131.7, 129.0, 126.6, 126.0, 115.4, 114.7, 114.4, 56.6, 56.2, 55.9, 20.7, 16.4.

IR (cm^{-1}): 3409 (NH), 3256 (NH), 1652 (C=O).

M.P.: 196.9 – 197.7 °C.

MS ES+ m/z Calcd for C$_{18}$H$_{21}$NO$_4$ (M+H)$^+$: 318.1608, found: 318.1603.

Prepared by adapting a literature procedure.[259]

3(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-6-methylbenzonitrile
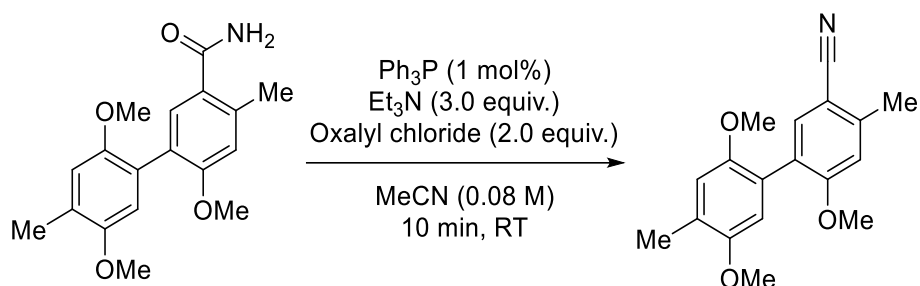
To an oven-dried RBF (25 mL) was added 3-(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-6-methylbenzamide (0.20 g, 0.63 mmol), triphenylphosphine (2.2 mg, 1 mol%), triethylamine (0.26 mL, 3.0 equiv.) and MeCN (4 mL, over mol. sieves, 0.08 M). To this stirred mixture was added dropwise oxalyl chloride (0.107 mL, 2.0 equiv.) and this was left to stir for 10 min. Reaction was monitored by TLC (2:3 Et$_2$O : Hexane), after confirmation of starting material consumption (10 min.), the reaction mixture was filtered and washed with excess MeCN. The solvent was removed *in vacuo.* and the residue purified by column chromatography (2:3 Et$_2$O : Hexane). Yielded a white solid of 3-(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-6-methylbenzonitrile (128 mg, 0.43 mmol, 68%).

^{1}H NMR (500 MHz, Acetone-d_6) δ 7.45 (s, 1H), 7.12 (s, 1H), 6.90 (s, 1H), 6.77 (s, 1H), 3.84 (s, 3H), 3.78 (s, 3H), 3.69 (s, 3H), 2.54 (s, 3H), 2.22 (s, 3H).

^{13}C NMR (126 MHz, Acetone-d_6) δ 161.4, 152.4, 151.8, 143.9, 135.8, 127.8, 127.5, 124.3, 119.0, 115.4, 114.4, 113.8, 104.5, 56.6, 56.23, 56.21, 20.6, 16.41.

MS ES+ m/z Calcd for C$_{18}$H$_{19}$NO$_3$ (M+H)$^+$: 300.1503, found: 300.1496.

Prepared by adapting a literature procedure.[233,234]

3(2',5'-dihydroxy-4'-methylphenyl)-4-hydroxy-6-methylbenzoic acid
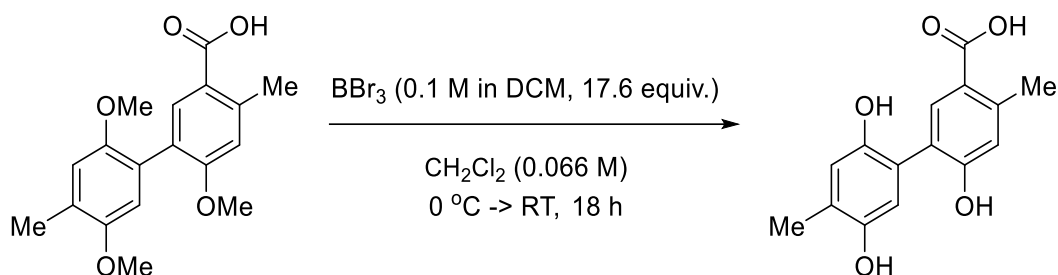
To a 10 mL RBF was added 3-(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-5-methylbenzoic acid (0.0994 g, 0.314 mmol), DCM (4.7 mL, 0.066 M) and cooled to 0 °C. BBr_3 (3.9 mL, 0.1 M in DCM, 0.39 mmol) was added dropwise with stirring and this was allowed to warm to RT and stirred for 18 h. The mixture was quenched carefully with H_2O (10 mL) and extracted with EtOAc (4x 10 mL), dried (Na_2SO_4) and solvent removed *in vacuo.* The residue was dissolved in EtOAc (10 mL) was transferred to a separatory funnel and a liberal amount of $Na_2S_2O_4$ was added with H_2O (10 mL). This was then extracted with EtOAc (3x 15 mL). The organic layers were collected and dried (Na_2SO_4) and solvent concentrated *in vacuo.*, towards the end of the solvent removal the remaining solvent was removed under a stream of nitrogen and finally placed on high vacuum. Solidified as a light beige solid, 3-(2',5'-dihydroxy-4'-methylphenyl)-4-hydroxy-6-methylbenzoic acid (60 mg, 0.22 mmol, 70%).

^{1}H NMR (500 MHz, CD_3CN) δ 9.16 (br s, 1H), 7.82 (s, 1H), 7.58 (br s, 1H), 6.83 (s, 1H), 6.73 (br s, 1H), 6.71 (s, 1H), 6.64 (s, 1H), 6.43 (br s, 1H), 2.55 (s, 3H), 2.17 (s, 3H).

^{13}C NMR (126 MHz, CD_3CN) δ 168.4, 158.2, 150.0, 147.2, 142.7, 136.0, 126.1, 124.7, 123.6, 122.2, 120.6, 119.5, 117.9, 21.9, 16.0.

IR: 3410 (OH), 3306 (OH), 2923, 1695 (C=O).

M.P. > 200 °C

MS ES+ m/z Calcd for $C_{15}H_{14}O_5$ (M+H)$^+$: 275.0919, found: 275.0921.

Prepared by adapting a literature procedure.[260]

3-(2',5'-dihydroxy-4'-methylphenyl)-4-hydroxy-6-methylbenzonitrile
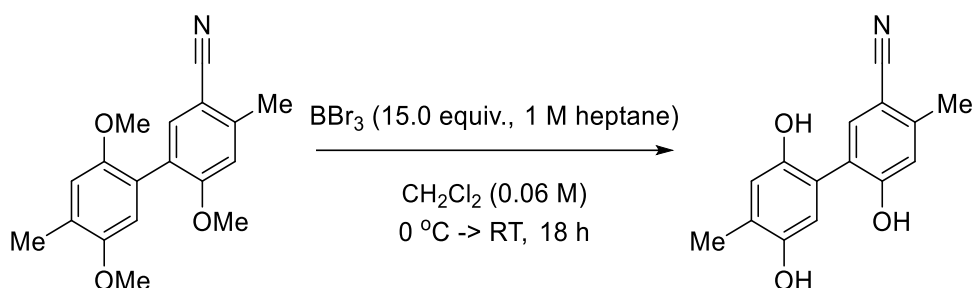
To an 8 mL microwave vial was added 3-(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-6-methylbenzonitrile (29.7 mg, 0.1 mmol), DCM (0.2 mL) and a stirrer bar, this was cooled to 0 °C. With stirring, BBr_3 (1.51 mL, 15.0 equiv., 1 M heptane) was added dropwise *via* syringe (total concentration 0.06M). After addition this was allowed to warm to RT and stirred for 18 h. The reaction mixture was cooled to 0 °C and quenched with water (≈ 1 mL). The reaction mixture was transferred to a separatory funnel and extracted with EtOAc (3x 10 mL). The organic layers were combined, washed with brine (10 mL), dried ($MgSO_4$) and solvent removed *in vacuo.* to yield 3-(2',5'-dihydroxy-4'-methylphenyl)-4-hydroxy-6-methylbenzonitrile as a beige solid (21 mg, 0.082 mmol, 83%).

^{1}H NMR (500 MHz, Acetone-d_6) δ 8.86 (br s, 1H), 8.12 (br s, 1H), 7.79 (br s, 1H), 7.52 (s, 1H), 6.96 (d, $J = 0.8$ Hz, 1H), 6.77 (d, $J = 0.7$ Hz, 1H), 6.74 (s, 1H), 2.46 (d, $J = 0.8$ Hz, 3H), 2.19 (d, $J = 0.7$ Hz, 3H).

^{13}C NMR (126 MHz, Acetone-d_6) δ 159.0, 149.8, 147.4, 143.2, 136.5, 126.6, 126.0, 122.1, 119.33, 119.27, 119.0, 117.9, 104.7, 20.0, 16.1.

IR (cm^{-1}): 3309 (OH), 2222 (C≡N).

MS ES+ m/z Calcd for $C_{15}H_{13}NO_3$ (M+H)$^+$: 257.1006, found: 257.1005.

Prepared by adapting a literature procedure.[260]

3-(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-6-methylbenzylamine
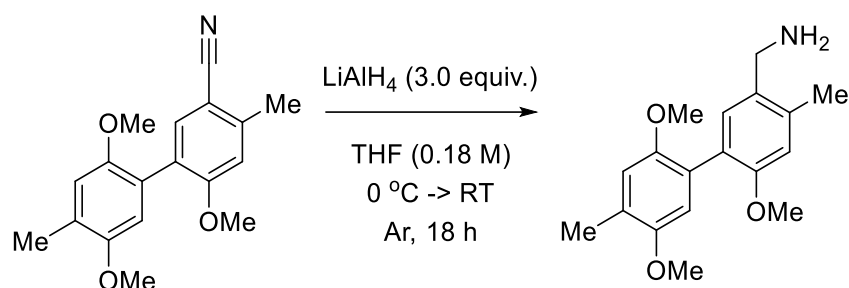
To a microwave vial was added LiAlH$_4$ (39.8 mg, 1 mmol), a stirrer bar and anhydrous THF (1.8 mL), this was cooled to 0 °C and with a funnel was added 3-(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-6-methylbenzonitrile (88.4 mg, 0.3 mmol) in one batch. This was warmed to RT and left to stir for 18 h. The reaction mixture was quenched at 0 °C with water (0.5 mL), NaOH solution (5 M, 1 mL) and again with water (0.5 mL). The mixture was transferred to a separatory funnel and extracted with EtOAc (3x 10 mL). The organic layers were collected and washed with brine (10 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* Gave a colourless oil of 3-(2',5'-dimethoxy-4'-methylphenyl)-4-methoxy-6-methylbenzylamine in quantitative yield (101 mg, 0.3 mmol, quant.).

^{1}H NMR (500 MHz, Acetone-d_6) δ 7.14 (s, 1H), 6.85 (s, 1H), 6.84 (s, 1H), 6.73 (s, 1H), 4.36 (s, 2H), 3.75 (s, 3H), 3.71 (s, 3H), 3.65 (s, 3H), 3.02 (br s, 2H), 2.35 (s, 3H), 2.23 (s, 3H).

^{13}C NMR (126 MHz, Acetone-d_6) δ 156.6, 152.3, 152.0, 137.0, 132.2, 131.6, 127.2, 126.2, 126.1, 115.5, 114.8, 114.0, 56.6, 56.2, 56,0, 53.4, 19.5, 16.4.

MS ES+ m/z, No M+H$^+$ found, Calcd. for C$_{18}$H$_{21}$O$_3$ (-NH$_2$): 286.1525, found 286.1526.

Prepared by adapting a literature procedure.[261]

4-bromo-2,5-dimethoxytoluene
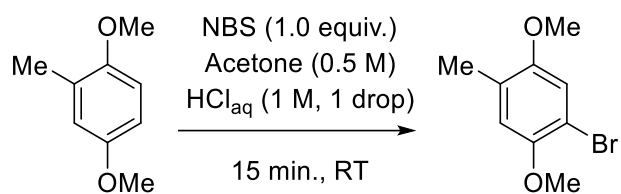
To a 250 mL RBF was added a stirrer bar, 2,5-dimethoxytoluene (4.93 g, 32.41 mmol), NBS (5.78 g, 32.41 mmol) and acetone (65 mL, 0.5 M). To this stirred solution was added aqueous HCl (1 M, 1 mL), this was left to stir for 15 min. The acetone is removed *in vacuo.* and the solid is triturated with cold Pet. E on a fritted glass funnel. Yielded light pink crystals of 4-bromo-2,5-dimethoxytoluene (5.29 g, 22.89 mmol, 70%).

^{1}H NMR (500 MHz, C*D*Cl$_3$) δ 6.99 (s, 1H), 6.74 (d, J = 0.8 Hz, 1H), 3.84 (s, 3H), 3.78 (s, 3H), 2.18 (d, J = 0.8 Hz, 3H).

^{13}C NMR (126 MHz, C*D*Cl$_3$) δ 152.4, 149.9, 127.0, 115.6, 115.4, 108.2, 57.1, 56.2, 16.5.

IR (cm^{-1}): 2981.

Data in-line with literature data.[262]

2,5-dimethoxy-4-methylphenylboronic acid
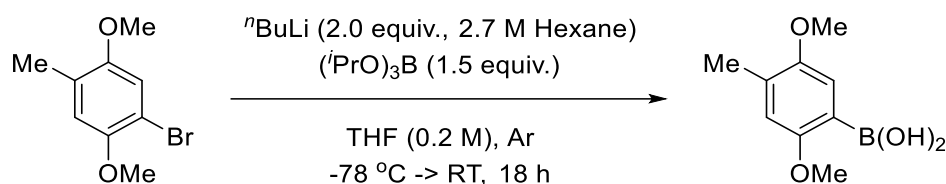
To an oven-dried 100 mL RBF was added a stirrer bar, 2,5-dimethoxy-4-bromotoluene (2.24 g, 9.69 mmol). This was placed under an argon atmosphere and anhydrous THF (50 mL) was added. After dissolution of 2,5-dimethoxy-4-bromotoluene the vessel was cooled to -78 °C. ^{n}BuLi (8.6 mL, 2.0 equiv., 2.7 M toluene) was added dropwise *via* syringe and left to stir. After 1-hour triisopropyl borate (3.4 mL, 1.5 equiv.) was added dropwise and was left to stir for 1 hour. The vessel was allowed to warm to RT and stirred for 18 h. The reaction was quenched with HCl (1 M, 40 mL), transferred to a separatory funnel, and extracted with EtOAc (3x 50 mL). The organic layers were collected and washed with brine (30 mL), dried (MgSO$_4$) and the solvent removed *in vacuo.* Yielded an orange oil, adding petroleum ether caused a yellow solid to precipitate. This was collected and triturated with petroleum ether, then *n*-hexane to give white fluffy needle-like crystals of 2,5-dimethoxy-4-methylphenylboronic acid (1.31 g, 6.68 mmol, 69%).

^{1}H NMR (500 MHz, C*D*Cl$_3$) δ 7.26 (s, 1H), 6.75 (d, J = 0.8 Hz, 1H), 5.83 (br, s, 2H), 3.87 (s, 3H), 3.84 (s, 3H), 2.26 (d, J = 0.8 Hz, 3H).

^{13}C NMR (126 MHz, C*D*Cl$_3$) δ 158.8, 152.2, 131.9, 117.4, 113.4, 56.2, 56.0, 17.0.

^{11}B NMR (128 MHz, C*D*Cl$_3$) δ 29.31.

IR: 3004 (OH), 2954.

M.P. 108 – 110 °C (lit.),[263] 117.6 – 119.3 °C.

Prepared by adapting literature procedure, data in-line with literature data.[264]

General procedure 5.1: 2-(2',5'-dimethoxy-4'-methylphenyl)-pyridine
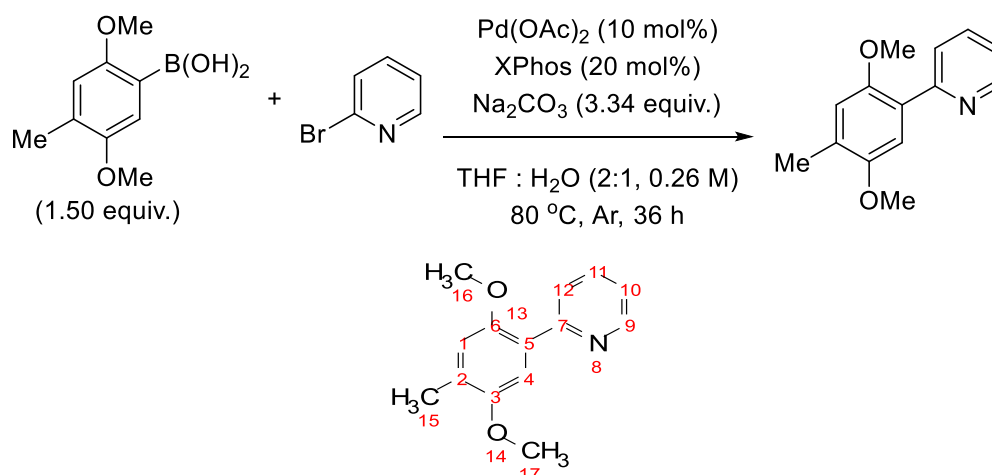
A stock solution of active catalyst was prepared by addition of Pd(OAc)$_2$ (10 mol%/reaction) and XPhos (20 mol%/reaction) in a 30 mL microwave vial. This was sealed and placed under an argon atmosphere. Anhydrous THF (1 mL/reaction) was added, and the mixture was degassed using an argon balloon with stirring for 20 min. To a separate 8 mL microwave vial was added 2,5-dimethoxy-4-methylphenylboronic acid (0.233 g, 1.18 mmol, 1.5 equiv.), 2-bromopyridine (0.125 g, 75.4 µL, 0.79 mmol), anhydrous sodium carbonate (0.28 g, 2.64 mmol) and a stirrer bar. The vial was sealed with a supa seal, and an argon atmosphere was applied. To this water (1 mL) and anhydrous THF (1 mL) were added, and the mixture was degassed using an argon balloon for 20 min. To this was added the activated catalyst solution *via* syringe (1 mL/reaction) and under an argon flow the vessel was sealed. This was heated to 80 °C for 36 h. Afterward the reaction vessel is left to cool to RT, the reaction mixture was transferred to a separatory funnel and extracted with EtOAc (3x 10 mL). The organic layers were collected and washed with brine (10 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* The crude material was purified *via* column chromatography (Hexane : EtOAc, 1:0 -> 9:1 -> 5:1) to yield 2-(2'-5'-dimethoxy-4'-methylphenyl)pyridine as a yellow oil (69.4 mg, 0.30 mmol, 37%).

^{1}H NMR (500 MHz, C*D*Cl$_3$) *δ* 8.69 (ddd, *J* = 4.9, 1.9, 1.1 Hz, 1H, **9**), 7.88 (ddd, *J* = 7.9, 1.2 Hz, 1H, **12**), 7.69 (ddd, *J* = 7.9, 1.9 Hz, 1H, **11**), 7.36 (s, 1H, **4**), 7.19 (ddd, *J* = 7.6, 4.9, 1.2 Hz, 1H, **10**), 6.83 (d, *J* = 0.9 Hz, 1H, **1**), 3.88 (s, 3H, **17**), 3.81 (s, 3H, **16**), 2.28 (s, 3H, **15**).

^{13}C NMR (126 MHz, C*D*Cl$_3$) *δ* 156.1, 152.4, 151.0, 149.4, 135.9, 128.5, 126.8, 125.3, 121.6, 115.3, 112.8, 56.6, 56.1, 16.6.

IR (cm^{-1}): 2926.

Prepared by adapting a literature procedure.[243]

3-(2',5'-dimethoxy-4'-methylphenyl)-pyridine
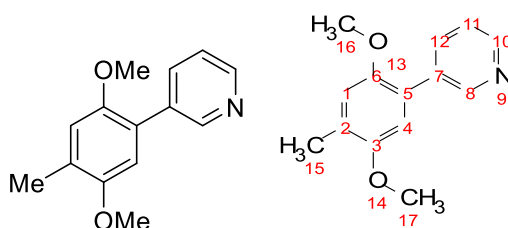
Prepared *via* general procedure 5.1 using the active catalyst stock solution as described above. 3-Bromopyridine (0.0773 mL, 0.79 mmol), 2,5-dimethoxy-4-methylphenylboronic acid (0.233 g, 1.18 mmol), anhydrous sodium carbonate (0.28 g, 3.34 equiv.), THF : H_2O (2:1, 0.2 M). Purified *via* column chromatography (9:1 -> 5:1 Hex : EtOAc). Yielded 3-(2',5'-dimethoxy-4'-methylphenyl)-pyridine as a pale yellow crystalline solid (172 mg, 0.75 mmol, 94%).

^{1}H NMR (500 MHz, C*D*Cl$_3$) *δ* 8.77 (s, 1H, **8**), 8.54 (d, *J* = 4.8 Hz, 1H, **10**), 7.87 (ddd, *J* = 7.8, 2.3, 1.7 Hz, 1H, **12**), 7.33 (dd, *J* = 7.8, 4.8 Hz, 1H, **11**), 6.84 (d, *J* = 0.9 Hz, 1H, **1**), 6.80 (s, 1H, **4**), 3.83 (s, 3H, **16**), 3.76 (s, 3H, **17**), 2.29 (s, 3H, **15**).

^{13}C NMR (126 MHz, C*D*Cl$_3$) *δ* 152.3, 150.4, 150.2, 147.8, 137.0, 128.0, 126.7, 124.7, 123.1, 115.1, 112.9, 56.5, 56.2, 16.5.

IR (cm^{-1}): 2936, 2845, 1513.

MS ES+ m/z Calcd for $C_{14}H_{15}NO_2$ (M+H)$^+$: 230.1181, found: 230.1173.

(2',5'-dimethoxy-4'-methylphenyl)-pyrazine
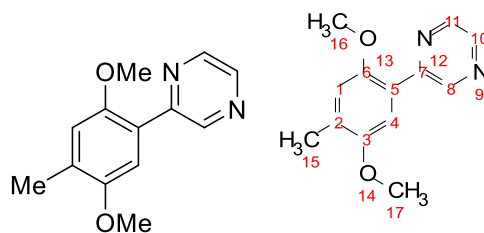
Prepared *via* general procedure 5.1 using the active catalyst stock solution as described above. 2-Bromopyrazine (0.071 mL, 0.79 mmol), 2,5-dimethoxy-4-methylphenyl boronic acid (0.231 g, 1.18 mmol, 1.5 equiv.), anhydrous sodium carbonate (0.28 g, 3.34 equiv.), THF : H_2O (2:1, 0.2 M). Purified *via* column chromatography (Hex : EtOAc 9:1 -> 4:1). Yielded 2-(2',5'-dimethoxy-4'-methylphenyl)-pyrazine as a yellow crystalline solid (91 mg, 0.39 mmol, 50%).

^{1}H NMR (500 MHz, C*D*Cl$_3$) δ 9.22 (d, J = 1.6 Hz, 1H, **8**), 8.63 (dd, J = 2.6, 1.6 Hz, 1H, **10**), 8.43 (d, J = 2.6 Hz, 1H, **11**), 7.40 (s, 1H, **4**), 6.86 (d, J = 0.8 Hz, 1H, **1**), 3.88 (s, 3H, **17**), 3.85 (s, 3H, **16**), 2.30 (d, J = 0.8 Hz, 3H, **15**).

^{13}C NMR (126 MHz, C*D*Cl$_3$) δ 152.5, 152.0, 151.2, 146.7, 144.0, 141.9, 130.0, 123.3, 115.0, 112.4, 56.4, 56.1, 16.7.

2-(2',5'-dimethoxy-4'-methylphenyl)-pyrimidine
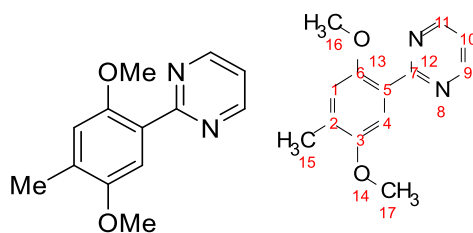
Anhydrous THF (1 mL) was added to an 8 mL vial and degassed with argon. To a separate 8 mL vial was added water (1 mL) and anhydrous THF (1 mL) and the mixture was degassed using an argon balloon for 20 min. The active catalyst was prepared by addition of $Pd(OAc)_2$ (10 mol%, 17.7 mg) and Xphos (20 mol%, 74.9 mg) in an 8 mL microwave vial which was backfilled with argon. To a separate 8 mL microwave vial was added 2,5-dihydroxy-4-methylphenylboronic acid (0.233 g, 1.18 mmol, 1.5 equiv.), 2-bromopyrimidine (0.125 g, 0.79 mmol), anhydrous sodium carbonate (0.28 g, 2.64 mmol) and a stirrer bar. The vial was sealed with a super seal and an argon atmosphere was applied. The degassed THF was added *via* syringe to the catalyst mixture, and the THF:H_2O mixture was transferred *via* syringe to the reagent mixture, both were stirred for 10 min. To the reagent mixture was added the activated catalyst solution and under an argon flow the vessel was sealed with a microwave cap. This was heated at 80 °C for 18 h. Afterward the reaction vessel is left to cool to RT, transferred to a separatory funnel and extracted with EtOAc (3x 10 mL). The organic layers were collected and washed with brine (10 mL), dried (MgSO₄) and solvent removed *in vacuo.* The crude material was purified *via* column chromatography (2:1 -> 0:1 Hexane : EtOAc) to yield 2-(2',5'-dimethoxy-4'-methylphenyl)-pyrimidine as a yellow solid (43 mg, 0.18 mmol, 24%).

¹H NMR (500 MHz, C*D*Cl₃) *δ* 8.84 (d, *J* = 4.9 Hz, 2H, **9**, **11**), 7.29 (s, 1H, **4**), 7.18 (t, *J* = 4.9 Hz, 1H, **10**), 6.86 (s, 1H, **1**), 3.85 (s, 3H, **16**), 3.83 (s, 3H, **17**), 2.28 (d, *J* = 0.8 Hz, 3H, **15**).

¹³C NMR (126 MHz, C*D*Cl₃) *δ* 165.9, 157.0, 152.0, 151.8, 130.0, 125.9, 118.5, 115.9, 113.4, 57.1, 56.0, 16.7.

MS ES+ m/z Calcd for $C_{13}H_{14}N_2O$ (M+H)⁺: 232.1164, found: 232.1161.

Prepared by adapting a literature procedure.[243]

4-(2',5'-dimethoxy-4'-methylphenyl)-pyridazine
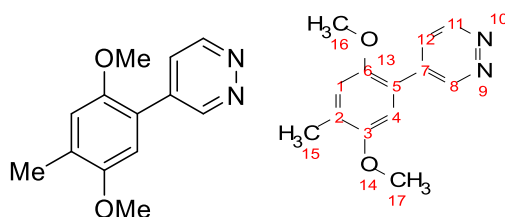
Prepared *via* general procedure 5.1 using the active catalyst stock solution as described above. 2-Bromopyridazine (0.125 g, 0.79 mmol), 2,5-dimethoxy-4-methylphenyl boronic acid (0.231 g, 1.18 mmol, 1.5 equiv.), anhydrous sodium carbonate (0.28 g, 3.34 equiv.). THF : H$_2$O (2:1, 0.2 M). Purified *via* column chromatography (4:1 -> 1:1 Hex : EtOAc). Yielded 3-(2',5'-dimethoxy-4'-methylphenyl)-pyridazine as an orange crystalline solid (39.1 mg, 0.17 mmol, 22%).

^{1}H NMR (500 MHz, C*D*Cl$_3$) *δ* 9.08 (dd, *J* = 4.8, 1.7 Hz, 1H, **8**), 8.12 (dd, *J* = 8.7, 1.7 Hz, 1H, **11**), 7.59 (s, 1H, **4**), 7.45 (dd, *J* = 8.7, 4.8 Hz, 1H, **12**), 6.85 (d, *J* = 0.9 Hz, 1H, **1**), 3.89 (s, 3H, **16**), 3.82 (s, 3H, **17**), 2.30 (d, *J* = 0.9 Hz, 3H, **15**).

^{13}C NMR (126 MHz, C*D*Cl$_3$) *δ* 159.3, 152.3, 151.2, 151.1, 149.3, 129.9, 128.5, 125.5, 114.9, 112.4, 56.3, 56.0, 16.6.

IR (cm^{-1}): 2923, 2833, 1578.

MS ES+ m/z Calcd for C$_{13}$H$_{14}$N$_2$O$_2$ (M+H)$^+$: 231.1134, found: 231.1131.

General procedure 5.2: Demethylation of *via* CAN

3-(2',5'-dihydroxy-4'-methylphenyl)pyridine
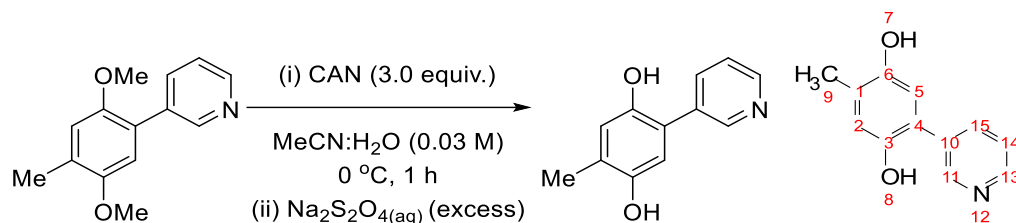
To a 30 mL vial was added 3-(2',5'-dimethoxy-4'-methylphenyl)pyridine (0.162 g, 0.7 mmol) and MeCN (8.75 mL). To a separate vial was added CAN (1.15 g, 3.0 equiv.) and H_2O (8.75 mL). Both were cooled to 0 °C. The CAN solution was added in one batch to the MeCN solution and stirred at 0 °C for 1 hr at which TLC (1:1 Pet E:EtOAc) showed full starting material consumption. MeCN was removed under reduced pressure and the aqueous mixture neutralised with $NaHCO_3$ to ≈ pH 7 and extracted with EtOAc (3x 10 mL). The organic layers were collected and treated with liberal amounts of sodium dithionite and shaken, dried ($MgSO_4$), filtered, solvent degassed with argon and then removed under vacuum. Product was purified *via* column chromatography (1% EtOH in DCM) to yield a green oil of 3-(2',5'-dihydroxy-4'-methylphenyl)pyridine (17 mg, 0.08 mmol, 12%).

1H NMR (500 MHz, DMSO-d6) *δ* 8.90 (br s, 1H, **7**), 8.74 (br s, 1H, **8**), 8.67 (d, J = 2.3 Hz, 1H, **11**), 8.44 (dd, J = 4.8, 1.7 Hz, 1H, **15**), 7.87 (dt, J = 7.9, 2.0 Hz, 1H, **13**), 7.39 (dd, J = 7.9, 4.8 Hz, 1H, **14**), 6.70 (s, 1H, **2**), 6.68 (s, 1H, **5**), 2.09 (s, 3H, **9**).

^{13}C NMR (126 MHz, DMSO-d_6) *δ* 149.3, **11**, 148.4, **6**, 147.**1**, 15, 146.7, **3**, 136.0, **13**, 134.4, **10**, 125.0, **4**, 123.1, **14**, 121.7, **1**, 118.3, **2**, 115.6, **5**, 15.9, **9** (134.4 identified *via* HMBC analysis).

IR (cm^{-1}): 3202 (OH), 1191 (C-O).

Not enough material for further analysis. Prepared by adapting a literature procedure.[265]

2-(2',5'-dihydroxy-4'-methylphenyl)pyridine
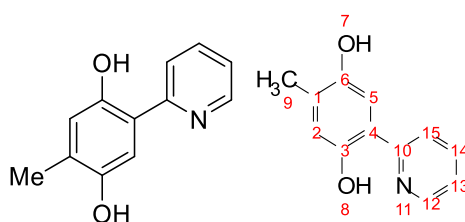
Prepared *via* general procedure 5.2. 2-(2',5'-dimethoxy-4'-methylphenyl)pyridine (69.4 mg, 0.30 mmol), CAN (493 mg, 0.91 mmol), MeCN (3.78 mL), H_2O (3.78 mL). Purified *via* column chromatography (1% EtOH in DCM) to yield an off-white solid of 2-(2',5'-dihydroxy-4'-methylphenyl)pyridine (3 mg, 0.015 mmol, 9%).

^{1}H NMR (500 MHz, DMSO-d_6) $δ$ 13.18 (br s, 1H, **8**), 8.79 (br s, 1H, **7**), 8.57 (ddd, J = 5.0, 1.9, 1.1 Hz, 1H, **12**), 7.98 (ddd, J = 8.3, 7.4, 1.9 Hz, 1H, **14**), 7.89 (dt, J = 8.3, 1.1 Hz, 1H, **15**), 7.37 (ddd, J = 7.4, 5.0, 1.1 Hz, 1H, **13**), 7.28 (s, 1H, **5**), 6.66 (d, J = 0.9 Hz, 1H, **2**), 2.12 (d, J = 0.9 Hz, 3H, **9**).

^{13}C NMR (126 MHz, DMSO-d_6) $δ$ 156.9, **10**, 151.7, **3**, 147.8, **6**, 146.5, **12**, 138.5, **14**, 129.0, **4**, 121.7, **13**, 119.5, **2**, 119.2, **15**, 116.3, **1**, 111.5, **5**, 16.1, **9**.

Not enough material for further analysis.

2-(2',5'-dihydroxy-4'-methylphenyl)pyrazine
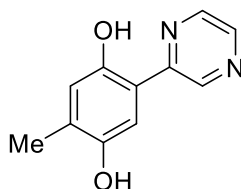
Prepared *via* general procedure 5.2. 2-(2',5'-dimethoxy-4'-methylphenyl)pyrazine (91 mg, 0.39 mmol), CAN (648 mg, 1.18 mmol), MeCN (4.93 mL), H_2O (4.93 mL). Isolated as a brown solid of 2-(2',5'-dihydroxy-4'-methylphenyl)pyrazine (38 mg, 0.19 mmol, 47%).

^{1}H NMR (500 MHz, DMSO-d_6) $δ$ 10.68 (br s, 1H), 9.25 (d, J = 1.6 Hz, 1H), 8.88 (br s, 1H), 8.63 (dd, J = 2.6, 1.6 Hz, 1H), 8.52 (d, J = 2.6 Hz, 1H), 7.37 (s, 1H), 6.72 (s, 1H), 2.13 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) $δ$ 152.1, 150.2, 148.8, 144.1, 143.0, 142.3, 129.3, 119.5, 117.9, 113.7, 16.6.

IR (cm^{-1}): 3157 (OH), 1143 (C-O). Not enough material for further analysis.

4-(2',5'-dihydroxy-4'-methylphenyl)pyridazine
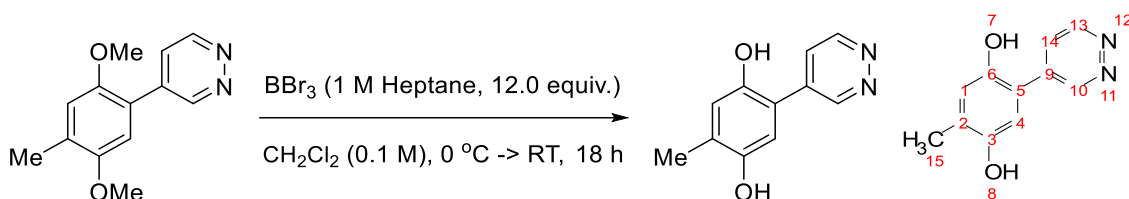
To a vial was added 4-(2',5'-dimethoxy-4'-methylphenyl)-pyridazine (66.9 mg, 0.29 mmol) with 2,5-dimethoxy-4-methylphenylboronic acid impurity (10.3 mg, 0.05 mmol), DCM (3.4 mL, 0.1 M) and a stirrer bar. This was cooled to 0 °C and BBr$_3$ (1 M in heptane, 3.9 mL, 3.48 mmol, 12 equiv.) was added dropwise with stirring. After addition, the mixture was left to stir at 0 °C for 5 min. and then allowed to warm to RT, left to stir for 18 h. The mixture is cooled to 0 °C and quenched with water (≈ 5 mL), neutralised to ≈ pH 7 with NaHCO$_3$ sat. solution, transferred to a separatory funnel and extracted with EtOAc (3x 20 mL). The organic layers were collected and treated with Na$_2$S$_2$O$_4$, dried (MgSO$_4$), filtered, degassed with argon for ≈ 10 min. and solvent removed *in vacuo.* to give a bright yellow solid of 4-(2',5'-dihydroxy-4'-methylphenyl)pyridazine (38.1 mg, 0.19 mmol, 38%).

^{1}H NMR (500 MHz, DMSO-d_6) δ 11.85 (br s, 1H, **7**), 9.13 (dd, J = 4.8, 1.5 Hz, 1H, **10**), 8.91 (br s, 1H, **8**), 8.20 (dd, J = 8.8, 1.5 Hz, 1H, **13**), 7.82 (dd, J = 8.8, 4.8 Hz, 1H, **14**), 7.31 (s, 1H, **4**), 6.75 (s, 1H, **1**), 2.15 (d, J = 0.8 Hz, 3H, **15**).

^{13}C NMR (126 MHz, DMSO-d_6) δ 159.9, 150.7, 149.7, **10**, 148.2, 129.6, 128.2, **14**, 124.9, **13**, 119.3, **1**, 115.9, 112.6, **4**, 16.2, **15**.

IR (cm^{-1}): 3400 (OH), 3245 (OH).

MS ES+ m/z Calcd for C$_{11}$H$_{10}$N$_2$O$_2$ (M+H)$^+$: 203.0820, found: 203.0822.

Prepared by adapting a literature procedure.[260]

2-(2',5'-dimethoxy-4'-methylphenyl)pyrimidine
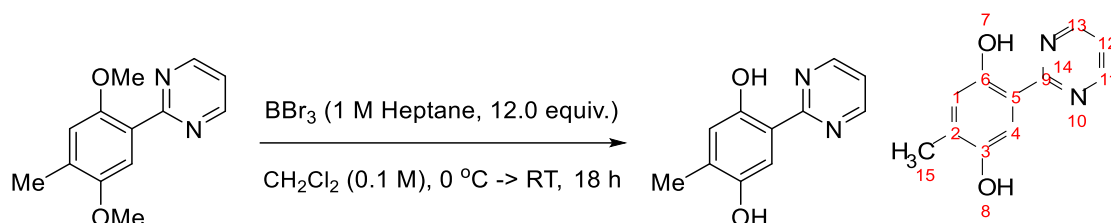
To a vial was added 2-(2',5'-dimethoxy-4'-methylphenyl)-pyrimidine (79.7 mg, 0.34 mmol), a stirrer bar and DCM (3.4 mL, 0.1 M). this was cooled to 0 °C and BBr$_3$ (1 M heptane, 4.2 mL, 12 equiv.) was added dropwise with stirring. The reaction mixture was left stirring for ≈ 5 min. and then allowed to warm to RT and left stirring for 18 h. The reaction mixture was cooled to 0 °C and quenched with water (≈ 5 mL), neutralised to ≈ pH 7 with NaHCO$_3$ (sat. solution), transferred to a separatory funnel and extracted with EtOAc (3x 20 mL). The organic layers were collected and treated with Na$_2$S$_2$O$_4$, dried (MgSO$_4$), filtered, degassed with argon for ≈ 10 min. and solvent removed *in vacuo.* to yield a yellow solid. This was purified *via* column chromatography (2% EtOH in DCM) to yield bright yellow needle-like fluffy crystals 2-(2',5'-dimethoxy-4'-methylphenyl)pyrimidine (60.5 mg, 0.30 mmol, 86%).

^{1}H NMR (500 MHz, DMSO-d_6) δ 12.44 (br s, 1H, **7**), 8.96 (br s, 1H, **8**), 8.91 (d, J = 4.9 Hz, 2H, **11**, **13**), 7.80 (s, 1H, **4**), 7.44 (t, J = 4.9 Hz, 1H, **12**), 6.70 (s, 1H, **1**), 2.14 (s, 3H, **15**).

^{13}C NMR (126 MHz, DMSO-d_6) 13C NMR (126 MHz, DMSO) δ 164.0, **9**, 156.9, **11**, **13**, 152.80, **6**, 147.9, **3**, 131.16, **5**, 119.14, **1**, 118.74, **12**, 115.70, **2**, 112.6, **4**, 16.3, **15**.

IR (cm^{-1}): 3166 (OH).

M.P.: 200 °C+

MS ES+ m/z Calcd for C$_{11}$H$_{10}$N$_2$O$_2$ (M+H)$^+$: 204.0851, found: 204.0846.

Prepared by adapting a literature procedure.[260]

2-fluoro-5-difluoromethylphenyl boronic acid pinacol ester
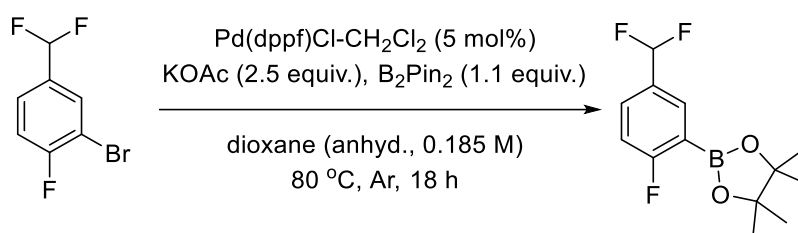
To an oven-dried Schlenk tube was added a stirrer bar, B_2Pin_2 (0.267 g, 1.1 equiv.) and KOAc (0.235 g, 2.50 equiv.). The mixture was cycled on a Schlenk line thrice (argon/vacuum) and degassed, anhydrous dioxane (5.2 mL, 0.185 M) added *via* syringe. 2-Fluoro-5-difluoromethylbromobenzene (0.216 g, 0.96 mmol) was added *via* syringe. Under an argon flow, Pd(dppf)Cl-CH_2Cl_2 (35 mg, 5 mol%) was added and the tube sealed and placed in a pre-heated oil bath (80 °C) and left to stir for 18 h. Afterward the tube was removed and left to cool to RT, the reaction mixture was filtered on celite, washed with EtOAc. The filtrate was dry loaded on silica and purified by column chromatography (Pet. E -> 5:1 Pet. E : EtOAc) which provided a clear oil which solidified over time to an off-white crystalline solid (43% by mass of product as determined by ^{1}H NMR, 37.3 mg, 14%, main impurity was thought to be pinacol, some aldehyde precursor impurity was also present).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 7.88 (t, J = 3.5 Hz, 1H), 7.64 – 7.56 (m, 1H), 7.11 (t, J = 8.7 Hz, 1H), 6.63 (t, J = 56.4 Hz, 1H), 1.37 (s, 12H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 168.4 (d, J = 254.6 Hz), 136.9 – 132.8 (m), 130.6 (d, J = 9.7 Hz), 115.9 (d, J = 25.1 Hz), 114.3 (t, J = 238.6 Hz), 84.3, 24.8 (7 out of 9 carbon resonances observed).

^{19}F NMR (376 MHz, C*D*Cl$_3$) δ -99.21 (ddt, J = 8.6, 5.8, 2.9 Hz, 1F), -109.46 (dd, J = 56.4, 3.3 Hz, 2F).

Prepared by adapting literature procedure.[266]

2-difluoromethyl-5-fluorophenyl boronic acid pinacol ester
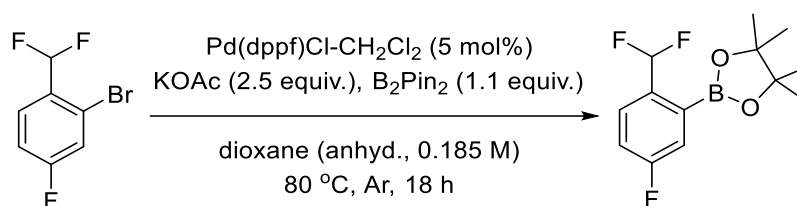
To an oven-dried Schlenk tube was added a stirrer bar, B_2Pin_2 (0.381 g, 1.10 equiv.) and KOAc (0.335 g, 2.5 equiv.). The mixture was cycled on a Schlenk line thrice (argon/vacuum) and degassed, anhydrous dioxane (7.4 mL, 0.185 M) added *via* syringe. 2-Difluoromethyl-5-fluorobromobenzene (0.308 g, 1.37 mmol) was added *via* syringe. Under an argon flow, Pd(dppf)Cl-CH$_2$Cl$_2$ (50 mg, 5 mol%) was added and the tube sealed and placed in a preheated oil bath (80 °C) and left to stir for 18 h. The tube was removed from the oil bath and allowed to cool to RT. The reaction mixture was filtered on celite®, washing with EtOAc. The filtrated was dry loaded onto silica and purified by column chromatography (Pet.E -> 1% EtOAc/Pet. E -> 2% Pet. E/EtOAc). Provided 2-difluoromethyl-5-fluorophenylboronic acid pinacol ester as a colourless oil (118 mg, 0.43 mmol, 32%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 7.71 (dd, $J = 8.6$, 5.1 Hz, 1H), 7.55 (ddd, $J = 8.9$, 2.9, 1.4 Hz, 1H), 7.33 (t, $J = 56.1$ Hz, 1H), 7.21 (td, $J = 8.5$, 2.8 Hz, 1H), 1.35 (s, 12H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 163.46 (d, $J = 250.7$ Hz), 136.23, 127.34 (q, $J = 6.9$ Hz), 122.39 (d, $J = 20.5$ Hz), 118.45 (d, $J = 21.7$ Hz), 113.42 (t, $J = 236.1$ Hz), 84.73, 24.93 (8 out of 9 carbon resonances observed).

^{19}F NMR (376 MHz, C*D*Cl$_3$) δ -109.77 (dd, $J = 55.8$, 3.7 Hz, 1F), -111.32 (tq, $J = 8.7$, 4.2 Hz, 2F).

^{11}B NMR (128 MHz, C*D*Cl$_3$) δ 30.56.

Prepared by following literature procedure.[266]

4-(2-fluoro-5,5-difluoromethylphenyl)isoquinoline
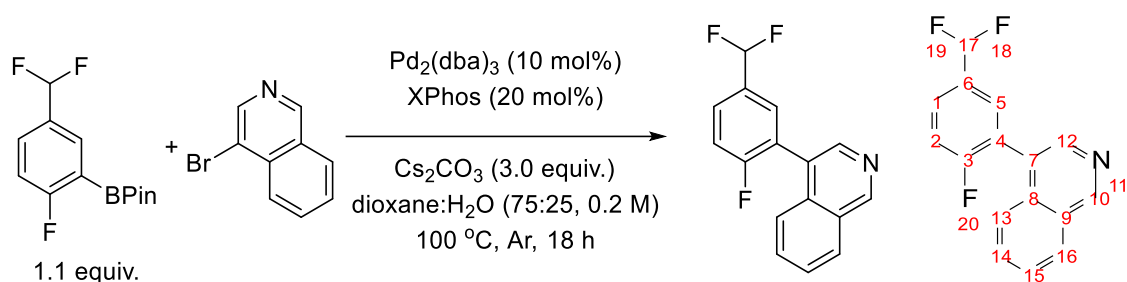
To an oven-dried 8 mL vial was added a stirrer bar, Pd$_2$(dba)$_3$ (20.5 mg, 10 mol%), XPhos (34 mg, 20 mol%), Cs$_2$CO$_3$ (0.376 g, 3.0 equiv.) and 4-bromoisoquinoline (83 mg, 0.40 mmol). This was placed under argon and cycled with vacuum/argon thrice. A degassed mixture of dioxane : water (3:1, 1.8 mL, 0.2 M) was added to an argon-filled microwave vial of 2-fluoro-5-difluoromethylphenylboronic acid pinacol ester mixture (96 mg, 1.1 equiv.), and further degassed with argon for 5 min. This mixture was transferred *via* syringe to the reaction vial, which was sealed under argon and placed in a pre-heated oil bath (100 °C) for 18 h. After allowing to cool to RT the reaction mixture was transferred to a separatory funnel and diluted with water. This was extracted with EtOAc (3x 20 mL), the organic layers were collected and washed with brine (10 mL), dried (MgSO$_4$) and solvent removed under reduced pressure. The residue was purified by column chromatography (9:1 Hexane : EtOAc). Further purified using 5% EtOAc in Hexane to provide 4-(2'-fluoro-5'-difluoromethylphenyl)isoquinoline as a transparent oil which crystallised into a white solid over time (10.5 mg, 0.04 mmol, 28%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 9.33 (s, 1H, **10**), 8.50 (s, 1H, **12**), 8.11 – 8.03 (m, 1H, **16**), 7.75 – 7.56 (m, 5H, **1**, **5**, **13**, **14**, **15**), 7.40 – 7.30 (m, 1H, **2**), 6.72 (t, *J* = 56.6 Hz, 1H, **17**).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 161.7 (d, *J* = 251.9 Hz, **3**), 153.4, **10**, 143.7, **12**, 134.3, **8**, 131.2, **15**, 131.4 – 130.7 (m, **6**), 130.2 – 130.0 (m, **5**), 128.4, **9**, 128.2, **16**, 128.0 – 127.8 (m, **1**), 127.7 , **13**, 126.7, **7**, 125.4 (d, *J* = 16.5 Hz, **4**), 124.6 (d, *J* = 1.7 Hz, **14**), 116.7 (d, *J* = 22.7 Hz, **2**), 114.1 (t, *J* = 239.4 Hz, **17**).

^{19}F NMR (376 MHz, C*D*Cl$_3$) *δ* -109.73 – -109.77 (m), -109.77 – -109.83 (m), -109.88 – -109.93 (m).

MS ES+ m/z Calcd for C$_{16}$H$_{10}$F$_3$N (M+H)$^+$: 275.0876 found: 275.0874.

Prepared by adapting literature procedure.[267]

4-(2,2-difluoromethyl-5-fluorophenyl)isoquinoline
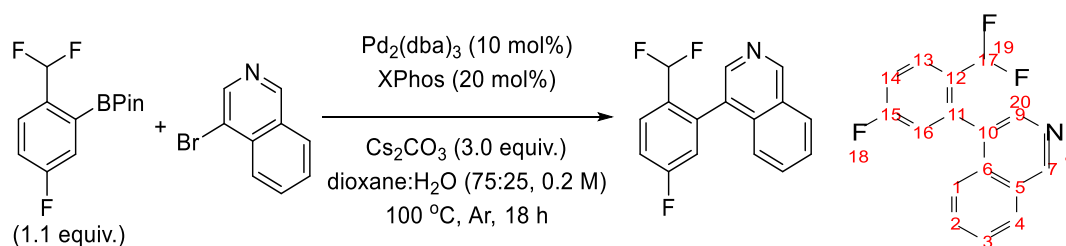
To an oven-dried 8 mL vial was added a stirrer bar, Pd$_2$(dba)$_3$ (25.0 mg, 10 mol%), XPhos (41 mg, 20 mol%), Cs$_2$CO$_3$ (0.459 g, 3.0 equiv.) and 4-bromoisoquinoline (81 mg, 0.4 mmol). This was placed under argon and cycled with vacuum thrice. An argon-degassed mixture of dioxane : water (75:25, 2.17 mL, 0.2 M) was added to an argon-filled microwave vial of 2-difluoromethyl-5-fluorophenylboronic acid pinacol ester mixture (0.118 g, 1.1 equiv.), and further degassed for 5 minutes with argon. This mixture was transferred *via* syringe to the reaction vial, sealed and placed in a pre-heated oil bath (100 °C) for 18 h. After allowing to cool to RT the reaction mixture was transferred to a separatory funnel and diluted with water. This was extracted with EtOAc (3x 20 mL), the organic layers were collected and washed with brine (10 mL), dried (MgSO$_4$) and solvent removed under reduced pressure. The residue was purified by column chromatography (9:1 Hexane : EtOAc). Purified further by placing onto a short silica column, washing with *n*-hexane, and then eluting with EtOAc to provide 4-(2'-difluoromethyl-5'-fluorophenyl)isoquinoline (69.5 mg, 0.25 mmol, 65%) as light-yellow oil.

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 9.35 (d, J = 0.9 Hz, 1H, **7**), 8.44 (s, 1H, **9**), 8.09 (ddd, J = 6.4, 2.2, 0.7 Hz, 1H, **4**), 7.87 (dd, J = 8.7, 5.6 Hz, 1H, **13**), 7.75 – 7.66 (m, 2H, **2**, **3**), 7.51 – 7.42 (m, 1H, **1**), 7.32 (td, J = 8.3, 2.6 Hz, 1H, **14**), 7.09 (ddt, J = 8.9, 2.5, 1.2 Hz, 1H, **16**), 6.23 (t, J = 54.9 Hz, 1H, **17**).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 163.6 (d, J = 249.5 Hz, **15**), 153.5, **7**, 143.2, **9**, 138.2 (dd, J = 7.1, 2.2 Hz, **11**), 134.7, **6**, 131.5, **3**, 130.0 (td, J = 22.7, 4.1 Hz, **12**), 128.5, **10**, 128.4 (dd, J = 10.3, 5.6 Hz, **13**), 128.3, **5**, 128.2, **4**, 128.0, **2**, 124.5, **1**, 118.5 (d, J = 22.1 Hz, **16**), 116.3 (d, J = 21.6 Hz, **14**), 112.4 (t, J = 237.4 Hz, **17**).

^{19}F NMR (376 MHz, C*D*Cl$_3$) δ -106.24 (ddd, J = 303.2, 54.9, 3.8 Hz), -109.17 – -109.29 (m), -110.42 (ddd, J = 303.5, 54.8, 3.2 Hz).

IR (cm^{-1}) 3037, 1615, 1588.

MS ES+ m/z C$_{16}$H$_{10}$F$_3$N (M+H)$^+$: 276.0909, found: 276.0905. Prepared by adapting literature procedure.[267]

2,4-dimethoxy-6-bromobenzaldehyde
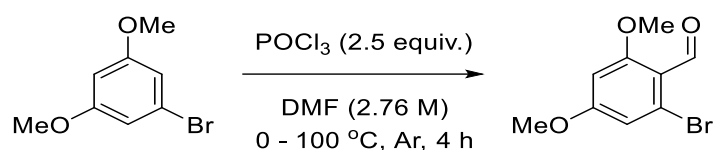
To an oven-dried 30 mL vial was added a stirrer bar and 3,5-dimethoxybromobenzene (2.0 g, 9.21 mmol), this was placed under an argon atmosphere. Anhydrous DMF (over mol. sieves, 3.34 mL, 2.76 M) was added *via* syringe, after dissolution this was subsequently cooled to 0 °C and POCl$_3$ (2.15 mL, 2.50 equiv.) was added dropwise *via* syringe. The mixture was then allowed to warm to RT and the vial sealed. The vial was then placed into a sand bath and heated to 100 °C for 4 h. The vial was allowed to cool to RT and the contents were poured onto ice-water (250 mL). The mixture was allowed to stand overnight after which the formed solid was filtered, washed with water, and dried under vacuum. The solid was recrystallised from Hex : EtOAc (2:1) to provide light-yellow needle-like crystals of 2,4-dimethoxy-6-bromobenzaldehyde (1.64 g, 6.70 mmol, 73%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 10.32 (s, 1H), 6.79 (d, *J* = 2.3 Hz, 1H), 6.44 (d, *J* = 2.3 Hz, 1H), 3.89 (s, 3H), 3.87 (s, 3H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 189.3, 164.6, 163.8, 127.6, 117.1, 111.7, 98.4, 56.3, 56.0.

IR (cm^{-1}) 2937, 2870 (C-H aldehyde), 2832 (C-H aldehyde), 1681 (C=O).

M.P. 88.0 – 89.4 °C.

Prepared following literature procedure.[268]

2,4-dimethoxy-6-bromobenzoic acid
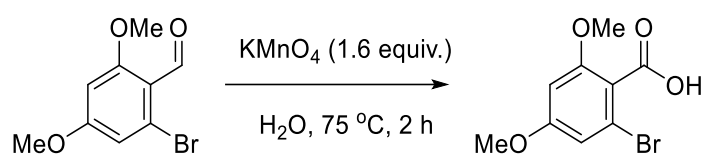
To a suspension of 2,4-dimethoxy-6-bromobenzaldehyde (0.30 g, 1.22 mmol) in water (11 mL, 0.11 M) at 75 °C was added batchwise an aqueous solution of KMnO$_4$ (0.309 g, 1.95 mmol, 1.6 equiv., 16.5 mL H$_2$O) within 20 min. Stirring was continued for 2 h after which the solution was basified with KOH (2 M) to approx. pH = 12. The mixture was filtered through celite®, and the filtrate acidified with HCl (1.0 M) to approximately pH = 2. The filtrate was then extracted with DCM (3x 30 mL), the organic layers were collected and dried (MgSO$_4$) and solvent removed under vacuum to yield 2,4-dimethoxy-6-bromobenzoic acid as a white powder (245 mg, 0.94 mmol, 77%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 6.75 (d, J = 2.2 Hz, 1H), 6.45 (d, J = 2.2 Hz, 1H), 3.86 (s, 3H), 3.83 (s, 3H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 168.7, 162.2, 158.9, 121.7, 117.1, 109.9, 98.4, 56.5, 55.9.

IR (cm^{-1}) 2943, 2842, 1692, 1593.

M.P. 153.0 – 154.6 °C.

Prepared by following literature procedure.[268]

Methyl 2,4-dimethoxy-6-bromobenzoate
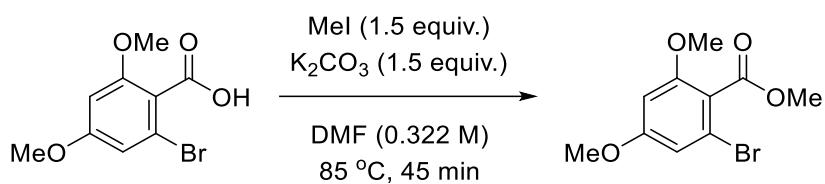
To an 8 mL microwave vial was added 2,4-dimethoxy-6-bromobenzoic acid (0.2 g, 0.766 mmol), K_2CO_3 (0.169 g, 1.49 mmol, 1.50 equiv.) and a stirrer bar. To this was added anhydrous DMF (over mol. sieves, 2.53 mL, 0.322 M) and iodomethane (76.6 µL, 1.49 mmol, 1.50 equiv.), both *via* syringe. The vial was sealed and heated to 85 °C for 45 min. The mixture was cooled and quenched slowly with HCl (1 M, 5 mL), transferred to a separatory funnel, and extracted with EtOAc (3x 20 mL). The organic layers were collected, washed with water, dried (MgSO$_4$) and solvent removed *in vacuo.* to provide a dark red oil which solidified under vacuum to a brown solid. (202 mg, 0.74 mmol, 96%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 6.68 (d, J = 2.1 Hz, 1H), 6.41 (d, J = 2.1 Hz, 1H), 3.91 (s, 3H), 3.80 (s, 3H), 3.80 (s, 3H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 167.0, 161.8, 158.5, 120.7, 119.1, 109.0, 98.3, 56.3, 55.9, 52.8.

IR (cm^{-1}) 3092 (OH), 1717 (C=O).

M.P. 57.4 – 58.4 °C

Prepared by following literature procedure.[268]

5.10.2   *F*-Redoxal

3,3'-difluorobenzidine, 3-fluorobenzidine and benzidine
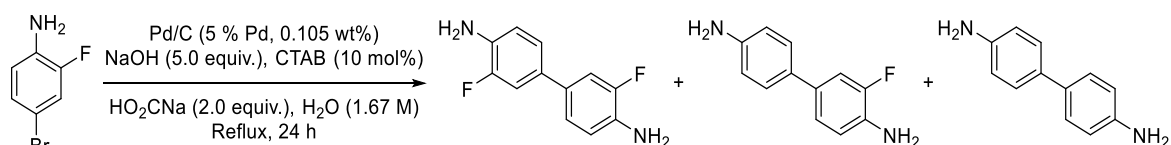
To a 100 mL RBF was added sodium hydroxide (10.52 g, 5.0 equiv.) and water (31.5 mL, 1.67 M). To this solution was added 4-bromo-2-fluoroaniline (10 g, 52.6 mmol), cetyltrimethylammonium bromide (CTAB, 1.92 g, 0.1 equiv.), Pd/C (5% Pd, 1.0 g, 10.5 wt%.) and HO_2CNa (3.58 g, 1.0 equiv.). The mixture was heated to reflux for 4 h and an additional portion of HO_2CNa (3.58 g, 1.0 equiv.) was added. The mixture was allowed to reflux for 18 h. Afterwards, the cooled mixture was filtered through a pad of Celite®, washing with EtOAc (50 mL). The filtrate was transferred to a separatory funnel and washed twice with H_2O (20 mL). The organic layer was dried ($MgSO_4$), and the solvent removed *in vacuo.* The crude material was purified by column chromatography, eluting with DCM -> DCM : MeOH (75 : 25). Chromatography provided 2,2'-difluorobenzidine as a yellow solid (387 mg, 1.74 mmol, 7%) and 2-fluorobenzidine as an off-white solid (468 mg, 2.31 mmol, 9%). Benzidine was further purified after this column. Followed literature procedure, all compound characterisations agree with literature values.[165]

3,3'-difluorobenzidine
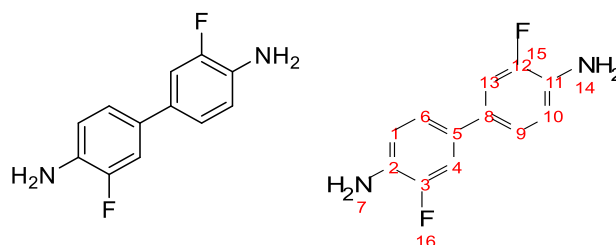
^{1}H NMR (400 MHz, CD_3CN) *δ* 7.20 (dd, J = 13.0, 2.1 Hz, 2H, **4**, **13**), 7.14 (ddd, J = 8.2, 2.1, 0.7 Hz, 2H, **6**, **9**), 6.82 (dd, J = 9.6, 8.2 Hz, 2H, **1**, **10**), 4.19 (s, 4H, **7**, **14**).

^{13}C NMR (101 MHz, CD_3CN) *δ* 152.15 (d, J = 236.4 Hz, **3**, **12**), 135.01 (d, J = 13.1 Hz, **2**, **11**), 130.49, **5**, **8**, 122.60 (d, J = 2.9 Hz, **6**, **9**), 117.42 (d, J = 4.6 Hz, **1**, **10**), 113.10 (d, J = 19.5 Hz, **4**, **13**).

^{19}F NMR (376 MHz, CD_3CN) *δ* -136.97 (ddd, J = 13.0, 9.6, 0.7 Hz).

M.P. 119.3 – 121.2 ℃.

IR (cm^{-1}): 3420 (NH), 3301 (NH), 1639 (NH bend).

MS ES+ m/z calcd for $C_{12}H_{10}F_2N_2$ (M+H)$^+$: 221.0890, found: 221.0890.

3-Fluorobenzidine
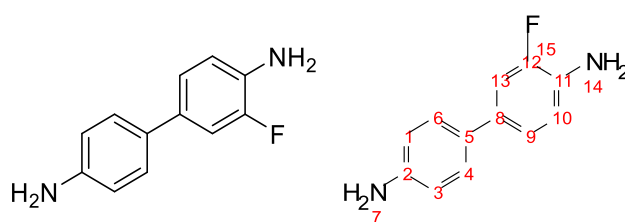
^{1}H NMR (400 MHz, CD_3CN) δ 7.29 (d, J = 8.5 Hz, 2H, *1*, *3*), 7.18 (dd, J = 13.1, 2.1 Hz, 1H, *13*), 7.12 (dd, J = 8.2, 2.1 Hz, 1H, *9*), 6.82 (dd, J = 9.6, 8.2 Hz, 1H, *10*), 6.67 (d, J = 8.5 Hz, 2H, *4*, *6*), 4.15 (br s, 4H, *7*, *14*).

^{13}C NMR (101 MHz, CD_3CN) δ 152.7 (d, J = 236.1 Hz, *12*), 148.0, *2*, 134.8 (d, J = 13.2 Hz, *11*), 132.3 (d, J = 6.5 Hz, *8*), 130.0 (d, J = 1.9 Hz, *5*), 127.8, *1*, *3*, 122.8 (d, J = 2.9 Hz, *9*), 117.9 (d, J = 4.5 Hz, *10*), 115.8, *4*, *6*, 113.3 (d, J = 19.3 Hz, *13*).

^{19}F NMR (376 MHz, CD_3CN) δ -136.93 (dd, J = 13.1, 9.6 Hz).

M.P. 99.8 – 100.9 ^{o}C.

IR (cm^{-1}): 3410 (NH), 3298 (NH), 3200 (NH), 1606 (NH Bend).

MS ES+ m/z calcd for C$_{12}$H$_{11}$FN$_2$ (M+H)$^+$: 203.0984, found 203.0983.

Benzidine
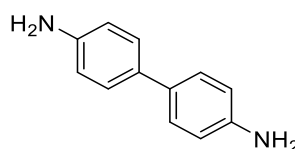
Further purified by dissolving the fractions containing benzidine in a small amount of DCM, adding it to a short plug of silica and eluting with DCM. Collected material was crystallised by dissolving in a small amount of EtOH and slowing adding hexane to induce crystallisation. Provided benzidine as light-yellow crystals (227 mg, 1.23 mmol, 5%).

^{1}H NMR (400 MHz, CD_3CN) δ 7.32 – 7.24 (m, 4H), 6.70 – 6.62 (m, 4H), 4.10 (s, 4H).

^{13}C NMR (101 MHz, CD_3CN) δ 147.5, 131.4, 127.7, 115.8.

M.P. 123.4 – 124.5 ^{o}C.

IR (cm^{-1}): 3401 (NH), 3327 (NH), 3193 (NH), 3032 (C(sp^2)-H), 1605 (NH bend).

MS ES+ m/z calcd for C$_{12}$H$_{12}$N$_2$ (M+H)$^+$: 185.1079, found: 185.1078.

General procedure 5.3: Buchwald-Hartwig C-N cross-coupling

R = R' = F
R = F, R' = H
R = R' = H

R = R' = F
R = F, R' = H
R = R' = H

+

R = R' = F
R = H, R' = F
R = R' = H
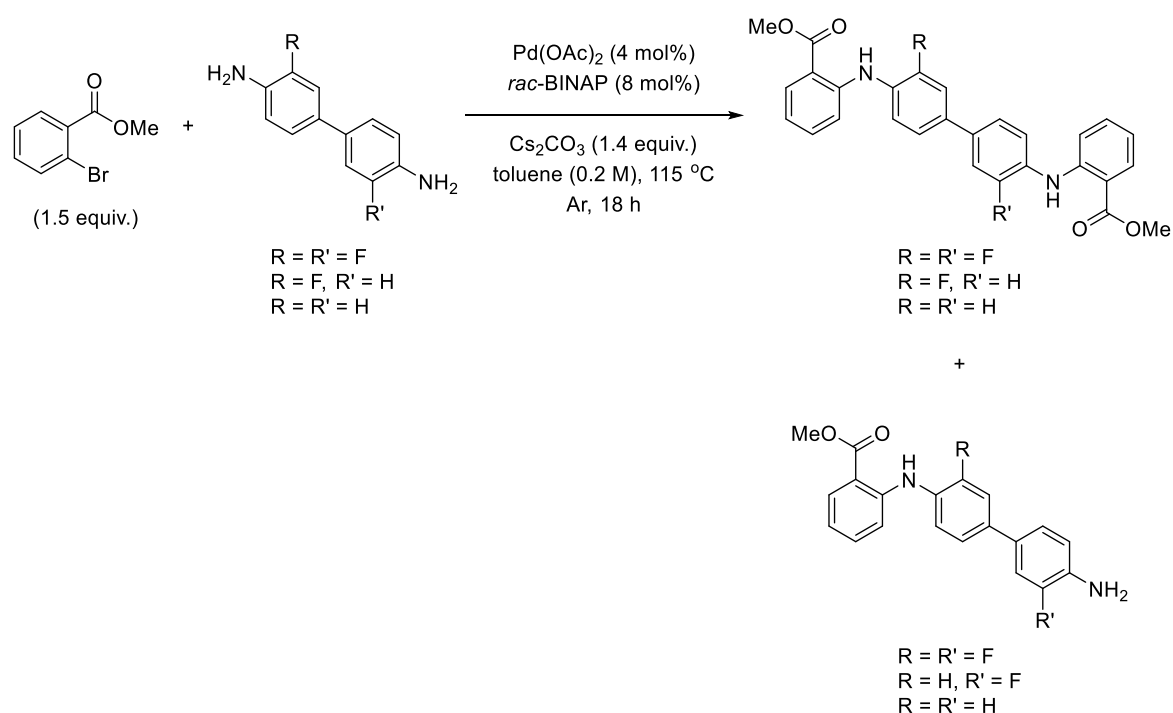
To an oven-dried 3-necked 50 mL RBF was added a stirrer bar and fitted with an oven-dried air condenser and two septa. To this was added under argon the respective benzidine (1.0 equiv.) and methyl 2-bromobenzoate (1.5 equiv.). Degassed, dried toluene (0.2 M) was added *via* syringe and the mixture was degassed for a further 10 min. Pd(OAc)$_2$ (4 mol%), *rac*-BINAP (8 mol%) and Cs$_2$CO$_3$ (1.4 equiv.) were then added under an argon flow. The mixture was stirred at 115 °C for 18 h under argon. Afterwards, the mixture was allowed to cool to RT and filtered through a pad of Celite®, washing with EtOAc. The solvent was removed, and the residue purified by column chromatography to provide both mono- and diarylated products. Adapted literature procedure.[146]

Dimethyl 2,2'-({3,3'-difluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate
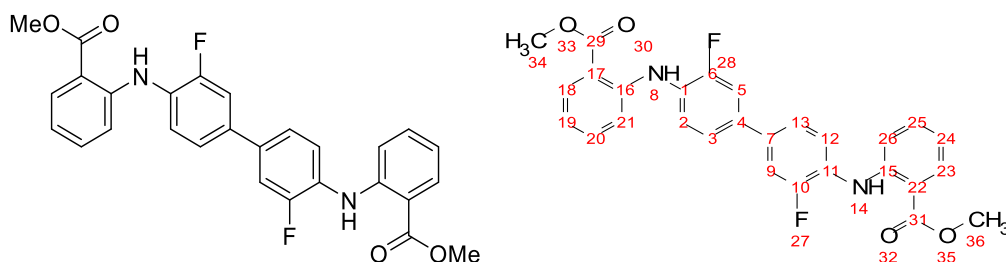
Prepared according to general procedure 5.3. 3,3'-difluorobenzidine (300 mg, 1.36 mmol), methyl 2-bromobenzoate (0.286 mL, 1.5 equiv.), Pd(OAc)$_2$ (12.3 mg, 4 mol%), *rac*-BINAP (67.9 mg, 8 mol%), Cs$_2$CO$_3$ (0.621 g, 1.4 equiv.), toluene (6.81 mL, 0.2 M). Residue was purified by column chromatography using 10% EtOAc in Hex. to provide dimethyl 2,2'-({3,3'-difluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate as a yellow solid (194 mg, 0.4 mmol, 31%) and a second fraction, P.T.O.

^{1}H NMR (400 MHz, DMSO-d_6) δ 9.43 (br s, 2H, **8**, **14**), 7.94 (dd, *J* = 8.0, 1.7 Hz, 2H, **18**, **23**), 7.75 (dd, *J* = 13.2, 1.5 Hz, 2H, **5**, **9**), 7.62 – 7.54 (m, 4H, **2**, **3**, **12**, **13**), 7.49 (ddd, *J* = 8.6, 7.2, 1.7 Hz, 2H, **20**, **25**), 7.19 (dd, *J* = 8.6, 1.3 Hz, 2H, **21**, **26**), 6.90 (ddd, *J* = 8.0, 7.2, 1.3 Hz, 2H, **19**, **24**), 3.88 (s, 6H, **34**, **36**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 168.1, **29**, **31**, 155.0 (d, *J* = 243.4 Hz, **6**, **10**), 145.7, **17**, **22**, 134.7, **20**, **25**, 134.5 (dd, *J* = 7.6 2.2 Hz, **4**, **7**), 131.3, **18**, **23**, 127.6 (d, *J* = 11.7 Hz, **1**, **11**), 122.9, **3**, **13**, 122.6, **2, 12**, 118.5, **19**, **24**, 114.6, **21**, **26**, 113.9 (d, *J* = 20.9 Hz, **5**, **9**), 112.6, **15**, **16**, 52.2, **34**, **36**.

^{19}F NMR (376 MHz, DMSO-d_6) δ -125.51 – -125.63 (m), **27**, **28**.

M.P. 157.0 – 159.2 ºC,

IR (cm^{-1}): 3254 (NH), 1688 (C=O).

MS ES+ m/z calcd for C$_{28}$H$_{22}$F$_2$N$_2$O$_4$ (M+H)$^+$: 489.1626, found: 489.1656.

Methyl 2-({3,3'-difluoro-4'-amino-[1,1'-biphenyl]-4-yl}amino)benzoate
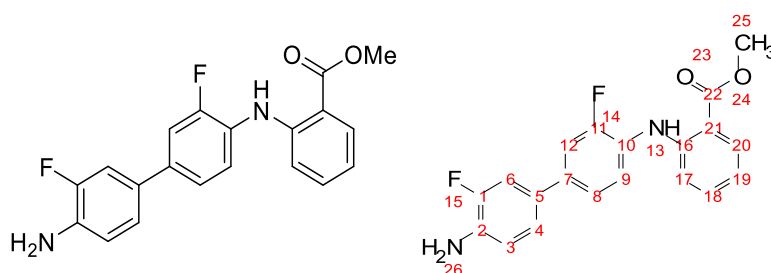
Second fraction was further purified from the above column by preparative TLC (500 μm thickness, glass-backed, 3 runs, one in 2.5%, then 5%, then 10% EtOAc / Hex.) to isolate methyl 2-({3,3'-difluoro-4'-amino-[1,1'-biphenyl]-4-yl}amino)benzoate as a yellow solid (13.6 mg, 0.04 mmol, 3%).

Methyl 2-([3,3'-difluoro-4'-amino-[1,1'-biphenyl]-4-yl]amino)benzoate was also isolated from a reaction described in section 3.3.4. (389 mg, 1.10 mmol, 24%) as a yellow gum, which was solidified by dissolving in a small volume of Et_2O and cyclohexane, slow evaporation of Et_2O causes precipitation as a yellow solid.

^{1}H NMR (400 MHz, CD_3CN) δ 9.39 (br s, 1H, **13**), 7.97 (ddd, J = 8.0, 1.7, 0.5 Hz, 1H, **20**), 7.47 (app t, J = 8.5 Hz, 1H, **9**), 7.42 (dd, J = 12.7, 1.6 Hz, 1H, **12**), 7.43 – 7.35 (m, 2H, **17**, **19**), 7.31 (dd, J = 12.9, 2.1 Hz, 1H, **6**), 7.25 (ddd, J = 8.2, 2.1, 0.7 Hz, 1H, **4**), 7.12 (dt, J = 8.5, 1.6 Hz, 1H, **8**), 6.87 (dd, J = 9.5, 8.2 Hz, 1H, **3**), 6.85 – 6.81 (m, 1H, **18**), 4.32 (br s, 2H, **26**), 3.89 (s, 3H, **25**).

^{13}C NMR (101 MHz, CD_3CN) δ 169.7, **22**, 157.0 (d, J = 243.2 Hz, **11**), 152.5 (d, J = 236.9 Hz, **1**), 148.0, **16**, 137.8 (dd, J = 7.5, 2.1 Hz, **5**), 136.5 (d, J = 12.9 Hz, **2**), 135.4, **19**, 132.5, **20**, 129.6 (dd, J = 6.6, 2.1 Hz, **7**), 127.9 (d, J = 12.0 Hz, **10**), 125.0 (d, J = 2.3 Hz, **9**), 123.7 (d, J = 2.9 Hz, **4**), 122.9 (d, J = 3.1 Hz, **17**), 118.9, **18**, 117.8 (d, J = 4.6 Hz, **3**), 115.1 (d, J = 1.1 Hz, **8**), 114.4 (d, J = 21.1 Hz, **12**), 114.0 (d, J = 19.8 Hz, **6**), 113.6, **21**, 52.7, **25**.

^{19}F NMR (376 MHz, CD_3CN) δ -126.83 (dd, J = 12.7, 8.5 Hz, **14**), -136.78 (dd, J = 12.9, 9.5 Hz, **15**).

M.P. 77.8 – 78.8 °C.

IR (cm^{-1}): 3401 (NH), 3314 (NH), 1677 (C=O).

MS ES+ m/z calcd for $C_{20}H_{16}F_2N_2O_2$ (M+H)$^+$: 355.1258, found: 355.1296.

Dimethyl 2,2'-({3-fluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate
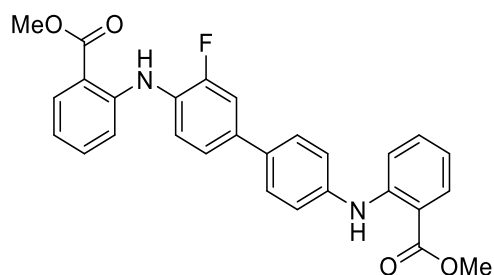
Prepared according to general procedure 5.3. 3-fluorobenzidine (400 mg, 1.98 mmol), methyl 2-bromobenzoate (0.416 mL, 1.5 equiv.), Pd(OAc)$_2$ (17.7 mg, 4 mol%), *rac*-BINAP (98.5 mg, 8 mol%), Cs$_2$CO$_3$ (0.90 g, 1.4 equiv.), toluene (9.88 mL, 0.2 M). Eluted using toluene provided dimethyl 2,2'-({3-fluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate as a yellow solid (678 mg, 1.44 mmol, 77%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 9.40 (br s, 1H), 9.38 (br s, 1H), 7.97 – 7.89 (m, 2H), 7.77 – 7.62 (m, 3H), 7.60 – 7.50 (m, 2H), 7.46 (t, J = 7.6 Hz, 2H), 7.34 (dd, J = 8.7, 3.4 Hz, 3H), 7.15 (d, J = 8.5 Hz, 1H), 6.92 – 6.80 (m, 2H), 3.88 (s, 3H), 3.87 (s, 3H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 168.6, 168.5, 155.7 (d, J = 242.8 Hz), 146.5 (d, J = 4.2 Hz), 140.7, 136.6 (d, J = 7.1 Hz), 135.0 (d, J = 18.8 Hz), 133.3, 131.8 (d, J = 11.6 Hz), 127.9, 127.3 (d, J = 11.8 Hz), 123.9, 122.8, 121.6, 118.7 (d, J = 6.7 Hz), 115.3, 114.8, 114.0 (d, J = 20.6 Hz), 113.2, 112.8, 52.6, 52.5 (22 out of 28 carbon resonances observed).

^{19}F NMR (376 MHz, DMSO-d_6) δ -125.35 (dd, J = 12.5, 7.7 Hz).

M.P. 148.2 – 149.2 $^{\circ}$C,

IR (cm^{-1}): 3307 (NH), 3266 (NH), 1679 (C=O).

MS ES+ m/z calcd for C$_{28}$H$_{23}$FN$_2$O$_4$ (M+H)$^+$: 472.1752, found: 472.1764.

May be recrystallised from MeCN to provide crystals suitable for X-ray diffraction, see appendix 1. 3 Fluorine atoms present due to 3-component disorder, the molecule was arranged in 3 different orientations within the lattice structure.
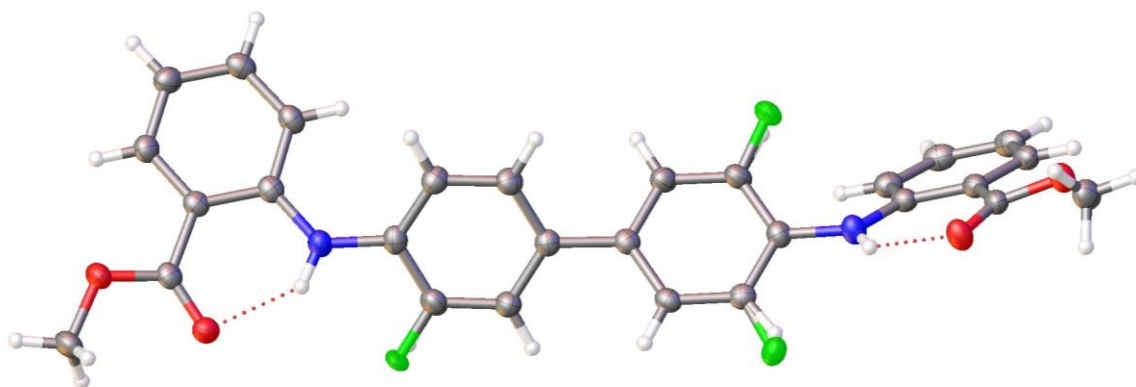
Methyl 2-([3'-fluoro-4'-amino-[1,1'-biphenyl]-4-yl]amino)benzoate
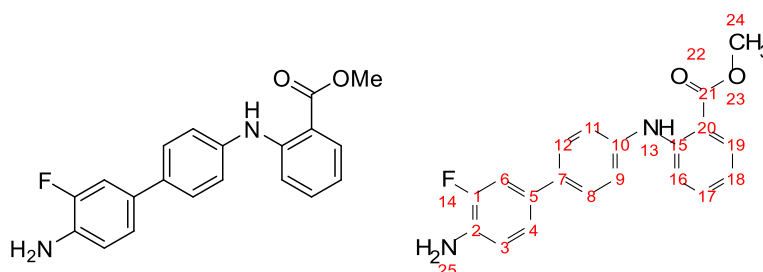
Isolated from the above reaction. A single regioisomer was isolated predominantly, methyl 2-([3'-fluoro-4'-amino-[1,1'-biphenyl]-4-yl]amino)benzoate (by ^{1}H-^{19}F HMBC and [1]H-^{19}F HOSEY analysis), isolated as a yellow solid (88.6 mg, 0.26 mmol, 14%).

^{1}H NMR (400 MHz, CD_3CN) δ 9.43 (s, 1H, **13**), 7.96 (ddd, J = 8.1, 1.7, 0.5 Hz, 1H, **19**), 7.60 – 7.51 (m, 2H, **8**, **12**), 7.39 (ddd, J = 8.7, 7.1, 1.7 Hz, 1H, **17**), 7.33 – 7.26 (m, 4H, **6**, **9**, **11**, **16**), 7.24 (ddd, J = 8.2, 2.1, 0.7 Hz, 1H, **4**), 6.87 (dd, J = 9.5, 8.2 Hz, 1H, **3**), 6.80 (ddd, J = 8.1, 7.1, 1.2 Hz, 1H, **18**), 4.26 (br s, 2H, **25**), 3.88 (s, 3H, **24**).

^{13}C NMR (101 MHz, CD_3CN) δ 169.7, **21**, 152.6 (d, J = 236.6 Hz, **1**), 148.5, **20**, 140.5, **15**, 136.1 (d, J = 2.1 Hz, **7**), 135.9 (d, J = 13.1 Hz, **2**), 135.3, **17**, 132.6, **19**, 131.0 (d, J = 6.4 Hz, **5**), 127.9, **8**, **12**, 123.5 (d, J = 2.9 Hz, **4**), 123.3, **16**, 118.5, **18**, 117.9 (d, J = 4.5 Hz, **3**), 115.2, **9**, **11**, 113.9 (d, J = 19.5 Hz, **6**), 113.3, **10**, 52.6, **24**.

^{19}F NMR (376 MHz, CD_3CN) δ -136.89 (dd, J = 13.0, 9.5 Hz, **14**).

M.P. 142.3 – 143.8 $^{\circ}$C.

IR (cm^{-1}): 3409 (NH), 3314 (NH), 1673 (C=O).

MS ES+ m/z calcd for C$_{20}$H$_{17}$FN$_2$O$_2$ (M+H)$^+$: 337.1352, found: 337.1352.

Methyl 2-([4'-amino[1,1'-biphenyl]]amino)benzoate
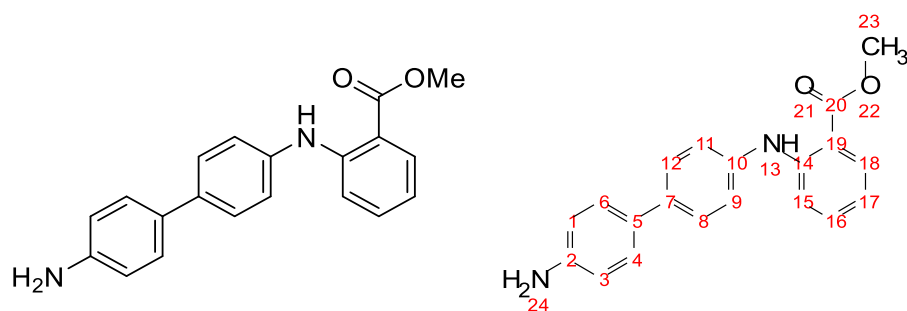
Prepared according to general procedure 5.3. Benzidine (200 mg, 1.08 mmol), methyl 2-bromobenzoate (0.230 mL, 1.5 equiv.), Pd(OAc)$_2$ (9.8 mg, 4 mol%), *rac*-BINAP (54.0 mg, 8 mol%), Cs$_2$CO$_3$ (0.495 g, 1.4 equiv.), toluene (5.4 mL, 0.2 M). Eluted using toluene to isolate methyl 2-([4'-amino-[1,1'-biphenyl]-4-yl]amino)benzoate as a yellow solid (58.8 mg, 0.18 mmol, 18%).

^{1}H NMR (400 MHz, C*D*$_3$CN) δ 9.41 (br s, 1H, *13*), 7.95 (ddd, *J* = 8.1, 1.7, 0.5 Hz, 1H, *18*), 7.58 – 7.50 (m, 2H, *8*, *12*), 7.41 – 7.38 (m, 2H, *4*, *6*), 7.37 (ddd, *J* = 7.1, 1.7, 0.5 Hz, 1H, *16*), 7.30 – 7.25 (m, 3H, *9*, *11*, *15*), 6.78 (ddd, *J* = 8.1, 7.1, 1.2 Hz, 1H, *17*), 6.73 – 6.69 (m, 2H, *1*, *3*), 4.22 (br s, 2H, *24*), 3.88 (s, 3H, *23*).

^{13}C NMR (101 MHz, C*D*$_3$CN) δ 169.7, *20*, 148.7, *14*, 148.5, *2*, 139.9, *10*, 137.5, *7*, 135.3, *16*, 132.5, *18*, 130.1, *5*, 128.2, *4*, *6*, 127.7, *8*, *12*, 123.6, *9*, *11*, 115.8, *1*, *3*, 115.1, *15*, 113.1, *19*, 52.6, *23*. *17* Carbon signal is behind 118.32 solvent signal (HSQC).

M.P. 152.2 – 153.2 °C,

IR (cm^{-1}): 3424 (NH), 3325 (NH), 1677 (C=O).

MS ES+ m/z calcd for C$_{20}$H$_{18}$N$_2$O$_2$ (M+H)$^+$: 319.1447, found: 319.1435.

Methyl 2-{(2'-fluorophenyl)amino}benzoate
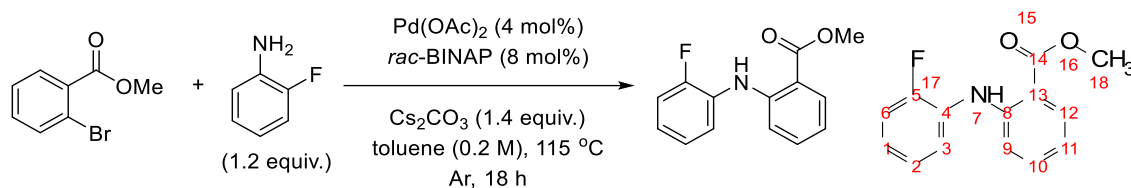
To an oven-dried 50 mL 2-necked RBF was attached a reflux condenser and two septa. The setup placed under an argon atmosphere and methyl 2-bromobenzoate (0.28 mL, 2.0 mmol) and 2-fluoroaniline (0.232 mL, 2.4 mmol) were added *via* syringe. Toluene (dried over mol. sieves and degassed with argon, 10 mL, 0.2 M) was added *via* syringe and degassing was continued with argon for a further 10 min. Pd(OAc)$_2$ (18 mg, 4 mol%), *rac*-BINAP (99 mg, 8 mol%) and Cs$_2$CO$_3$ (0.91 g, 1.4 equiv.) were added under an argon flow. The mixture was heated to 115 °C for 18 h with stirring. The mixture was allowed to cool to RT and filtered through a Celite® pad, washing with EtOAc (30 mL). The solvents were removed under reduced pressure, and the residue was purified by column chromatography (5% EtOAc in Hex.) to provide methyl 2-(2'-fluorophenylamino)benzoate as a light-yellow oil which solidifies into an orange solid (475 mg, 1.93 mmol, 97%).

^{1}H NMR (400 MHz, C*D*$_3$CN) *δ* 9.37 (br s, 1H, **7**), 7.97 (ddd, *J* = 8.0, 1.7, 0.4 Hz, 1H, **12**), 7.47 (ddd, *J* = 8.1, 1.8 Hz, 1H, **3**), 7.39 (ddd, *J* = 9.0, 7.2, 1.7 Hz, 1H, **11**), 7.25 – 7.11 (m, 3H, **1**, **2**, **6**), 7.09 (ddd, *J* = 8.5, 1.3 Hz, 1H, **10**), 6.83 (ddd, *J* = 8.2, 7.2, 1.1 Hz, 1H, **9**), 3.89 (s, 3H, **18**).

^{13}C NMR (101 MHz, C*D*$_3$CN) *δ* 169.7, **14**, 156.8 (d, *J* = 243.5 Hz, **5**), 148.0, **13**, 135.3, **11**, 132.5, **12**, 129.6 (d, *J* = 11.7 Hz, **4**), 125.8 (d, *J* = 5.7 Hz, **1**), 125.7 (d, *J* = 1.7 Hz, **2**), 124.9 (d, *J* = 1.7 Hz, **3**), 119.0, **9**, 117.0 (d, *J* = 20.1 Hz, **6**), 115.1, **10**, 113.6, **8**, 52.7, **18**.

^{19}F NMR (376 MHz, C*D*$_3$CN) *δ* -127.22 (ddd, *J* = 12.8, 7.0, 2.7 Hz).

M.P. 77.5 – 78.2 °C.

IR (cm^{-1}): 3316 (NH), 1682 (C=O).

MS ES+ m/z calcd for C$_{14}$H$_{12}$FNO$_2$ (M+H)$^+$: 246.0930, found: 246.0946.

Prepared according to literature procedure, data matches literature values.[146]

4-(3'-fluorophenyl)-2-fluoroaniline
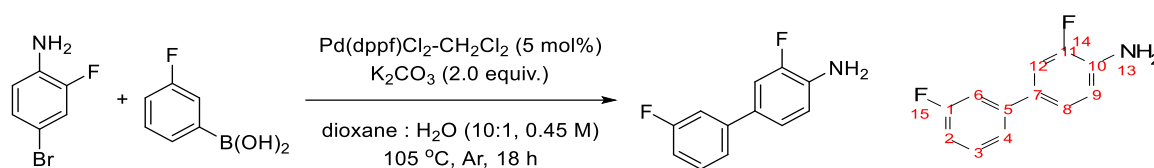
To an oven-dried Schlenk flask was added under argon 4-bromo-2-fluoroaniline (0.38 g, 2.0 mmol), 3-fluorophenylboronic acid (0.279 g, 1.0 equiv.) and a stirrer bar. This mixture was cycled three times on vacuum / argon. A degassed mixture of dioxane : H_2O was added to the flask *via* syringe (10 : 1, 4.44 mL, 0.45 M). Under a flow of argon was added K_2CO_3 (0.552 g, 2.0 equiv.) and Pd(dppf)Cl-CH_2Cl_2 (81.6 mg, 5 mol%) in one batch. The flask was sealed under argon and placed in a pre-heated oil bath at 105 ^{o}C for 18 h. The cooled mixture was filtered through a pad of Celite®, washing with EtOAc (20 mL). The filtrate was collected, and solvent removed *in vacuo.* The crude material was purified by column chromatography (10% EtOAc / Pet. E) to isolate 4-(3'-fluorophenyl)-2-fluoroaniline as a yellow oil (267 mg, 1.30 mmol, 65%). Solidifies in the freezer (-20 ^{o}C) into a waxy orange solid.

^{1}H NMR (400 MHz, CD_3CN) δ 7.42 – 7.37 (m, 2H, **3**, **4**), 7.36 – 7.27 (m, 2H, **6**, **12**), 7.25 (ddd, J = 8.3, 2.1, 0.7 Hz, 1H, **8**), 7.07 – 6.96 (m, 1H, **2**), 6.87 (dd, J = 9.4, 8.2 Hz, 1H, **9**), 4.34 (br s, 2H, **13**).

^{13}C NMR (101 MHz, CD_3CN) δ 164.3 (d, J = 242.8 Hz, **1**), 152.4 (d, J = 236.9 Hz, **11**), 143.5 (dd, J = 8.0, 2.1 Hz, **5**), 136.8 (d, J = 12.9 Hz, **10**), 131.6 (d, J = 8.7 Hz, **3**), 129.8 (dd, J = 6.5, 2.4 Hz, **7**), 124.0 (d, J = 2.9 Hz, **8**), 122.9 (d, J = 2.6 Hz, **4**), 117.7 (d, J = 4.5 Hz, **9**), 114.4 (d, J = 19.8 Hz, **12**), 114.1 (d, J = 21.4 Hz, **2**), 113.6 (d, J = 22.5 Hz, **6**).

^{19}F NMR (376 MHz, CD_3CN) δ -114.88 – -115.04 (m, **15**), -136.89 (ddd, J = 12.8, 9.4, 2.5 Hz, **14**).

IR (cm^{-1}): 3469 (NH), 3384 (NH).

MS ES+ m/z calcd for $C_{12}H_9F_2N$ (M+H)$^+$: 206.0781, found: 206.0783.

Prepared by adaptation of literature procedure.[251]

Methyl 2-({3,3'-difluoro-[1,1'-biphenyl]-4-yl}amino)benzoate
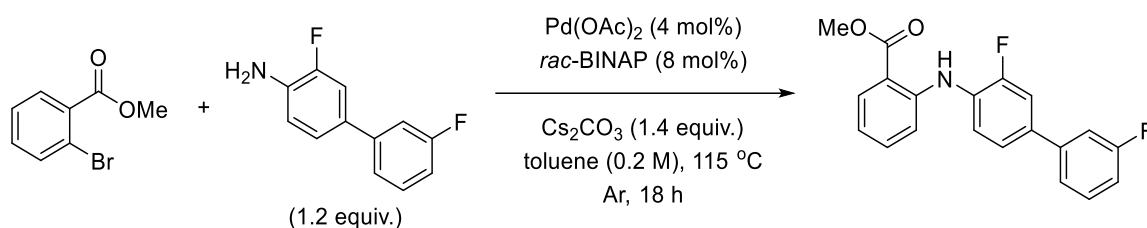
To an oven-dried 2-necked 50 mL RBF was added a stirrer bar and fitted with an oven-dried reflux condenser. To this was added under argon 4-(3'-fluorophenyl)-2-fluoroaniline (0.20 g, 0.97 mmol, 1.2 equiv.) and methyl 2-bromobenzoate (0.11 mL, 0.77 mmol). Degassed, dried toluene (5 mL, 0.2 M) was added *via* syringe and the mixture was degassed for a further 10 min. Pd(OAc)$_2$ (9 mg, 4 mol%), *rac*-BINAP (49 mg, 8 mol%) and Cs$_2$CO$_3$ (0.44 g, 1.4 equiv.) were then added under an argon flow. The mixture was stirred at 115 °C for 18 h under argon. Afterwards, the mixture was allowed to cool to RT and filtered through a pad of Celite®, washing with EtOAc (30 mL). The solvent was removed, and the mixture purified by sequential column chromatography, eluting first 5 -> 10% EtOAc / hexane and then 20% toluene / hexane to afford methyl 2-({3,3'-difluoro-[1,1'-biphenyl]-4-yl}amino)benzoate as a light green oil which crystallises to an off-white solid (206 mg, 0.61 mmol, 78%).

^{1}H NMR (400 MHz, CD_3CN) δ 9.49 (br s, 1H), 7.99 (ddd, J = 8.0, 1.7, 0.4 Hz, 1H), 7.57 (t, J = 8.5 Hz, 1H), 7.54 (dd, J = 12.4, 2.1 Hz, 1H), 7.50 – 7.39 (m, 5H), 7.23 (dt, J = 8.4, 1.2 Hz, 1H), 7.16 – 7.06 (m, 1H), 6.88 (ddd, J = 8.1, 7.1, 1.1 Hz, 1H), 3.90 (s, 3H).

^{13}C NMR (101 MHz, CD_3CN) δ 169.7, 164.2 (d, J = 243.3 Hz), 156.5 (d, J = 243.3 Hz), 147.3, 142.5, 136.5, 135.4, 132.5, 131.8 (d, J = 8.5 Hz), 129.7 (d, J = 11.7 Hz), 124.0 (d, J = 3.2 Hz), 123.9 (d, J = 2.2 Hz), 123.5 (d, J = 2.8 Hz), 119.4, 115.6 (d, J = 1.1 Hz), 115.4 (d, J = 21.2 Hz), 115.2 (d, J = 21.4 Hz), 114.3 (d, J = 20.4 Hz), 114.1, 52.7.

^{19}F NMR (376 MHz, CD_3CN) δ -114.44 – -114.71 (m), -127.18 – -127.64 (m).

M.P. 91.5 -95.1 °C.

IR (cm^{-1}): 3314 (NH), 1680 (C=O).

MS ES+ m/z calcd for C$_{20}$H$_{15}$F$_2$NO$_2$ (M+H)$^+$: 341.1182, found: 341.1181.

Prepared by adaptation of literature procedure.[146]

General procedure 5.4: Hydrolysis of methyl esters
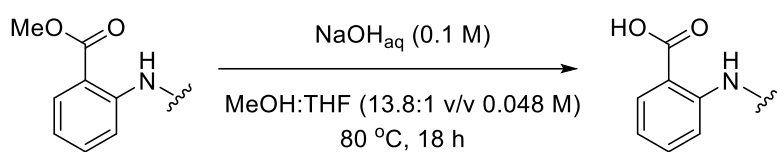
To an RBF was added the corresponding methyl ester or bis(methyl ester) and MeOH (0.1 M). To this was added a NaOH solution (0.1 M, 0.1 M). This was stirred at 85 °C. If the material was not dissolved, THF (0.5 M) was added, and the mixture stirred at 85 °C for 18 h. Afterwards, the reaction was allowed to cool to RT, transferred to a sep. funnel, acidified (1 M HCl) and extracted with DCM (3x 10 mL), organic layers collected and dried (Na₂SO₄) and solvent removed *in vacuo.* Residues were triturated with toluene and Et₂O, dried under high vacuum.

2(2'-fluorophenyl)-aminobenzoic acid
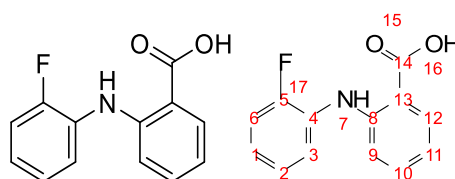
Prepared according to general procedure 5.4. Methyl 2(2'-fluorophenyl)aminobenzoate (0.2 g, 0.815 mmol), 8.15 mL MeOH (0.1 M), 8.15 mL NaOH soln. (0.1 M). Final concentration 0.048 M. Provided 2(2'-fluorophenyl)-aminobenzoic acid as a white crystalline solid (41.4 mg, 0.18 mmol, 22%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 13.20 (br s, 1H, **16**), 9.66 (br s, 1H, **7**), 7.92 (dd, J = 8.0, 1.7 Hz, 1H, **12**), 7.51 (app td, J = 8.2, 7.9, 1.8 Hz, 1H, **3**), 7.42 (ddd, J = 8.6, 7.1, 1.7 Hz, 1H, **10**), 7.31 (ddd, J = 11.3, 8.0, 1.6 Hz, 1H, **6**), 7.20 (app td, J = 7.9, 7.8, 1.6 Hz, 1H, **2**), 7.14 (dddd, J = 8.0, 7.8, 5.0, 1.8 Hz, 1H, **1**), 7.08 (dt, J = 8.6, 1.1 Hz, 1H, **9**), 6.83 (ddd, J = 8.0, 7.1, 1.1 Hz, 1H, **11**).

^{13}C NMR (101 MHz, DMSO-d_6) δ 170.0, **14**, 154.9 (d, J = 243.3 Hz, **5**), 146.4, **13**, 134.3, **10**, 131.8, **12**, 128.3 (d, J = 11.4 Hz, **4**), 124.9 (d, J = 3.6 Hz, **2**), 124.4 (d, J = 7.7 Hz, **1**), 123.1, **3**, 117.9, **11**, 116.1 (d, J = 19.5 Hz, **6**), 113.7, **9**, 113.0, **8**.

^{19}F NMR (376 MHz, DMSO-d_6) δ -125.95 – -126.06 (m) **17**.

M.P. 182.6 – 182.9 ºC (lit.),[269] 184.3 – 185.2 ºC.

IR (cm^{-1}): 3318 (OH), 1647 (C=O).

TOF MS ES+ m/z calcd for $C_{13}H_{10}FNO_2$ (M+H)$^+$: 233.0806, found: 233.0805.

Data in accordance with literature.[269]

2-([3,3'-difluoro-[1,1'-biphenyl]-4-yl]amino)benzoic acid
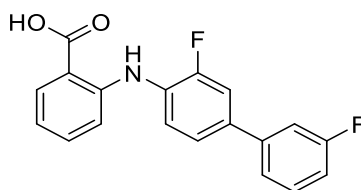
Prepared according to general procedure 5.4. Methyl 2-([3,3'-difluoro-[1,1'-biphenyl]-4-yl]amino)benzoate (210 mg, 0.294 mmol), MeOH (2.95 mL, 0.1 M), 0.59 mL THF (0.5 M), NaOH (0.1 M, 2.95 mL, 0.1 M). Final concentration 0.048 M. Provided 2-([3,3'-difluoro-[1,1'-biphenyl]-4-yl]amino)benzoic acid as a white crystalline solid (34.7 mg, 0.15 mmol, 50%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 13.28 (br s, 1H), 9.81 (br s, 1H), 7.95 (dd, J = 8.0, 1.7 Hz, 1H), 7.74 (dd, J = 12.7, 1.9 Hz, 1H), 7.58 (ddd, J = 7.9, 6.0, 2.7 Hz, 4H), 7.54 – 7.43 (m, 2H), 7.26 – 7.14 (m, 2H), 6.88 (t, J = 8.1 Hz, 1H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 170.0, 162.8 (d, J = 243.3 Hz), 154.7 (d, J = 242.9 Hz), 145.7, 140.9 (d, J = 8.1 Hz), 134.3, 134.2 (d, J = 7.4 Hz), 131.8, 130.9 (d, J = 8.7 Hz), 128.4 (d, J = 11.7 Hz), 123.0 (d, J = 3.0 Hz), 122.3 (d, J = 2.6 Hz), 122.1 (d, J = 2.1 Hz), 118.4, 114.4 (d, J = 6.1 Hz), 114.1 (d, J = 7.1 Hz), 113.5, 113.0 (d, J = 22.4 Hz) (18 out of 19 carbon resonances observed).

^{19}F NMR (376 MHz, DMSO-d_6) δ -112.70 (ddd, J = 11.1, 9.0, 6.2 Hz), -126.06 – -126.16 (m).

M.P. 185.4 °C (dec.).

IR (cm^{-1}): 3310(OH), 1668 (C=O).

MS ES+ m/z calcd for $C_{19}H_{13}F_2NO_2$ (M+H)$^+$: 327.1025, found: 327.1021.

2-([3'-fluoro-4'-amino-[1,1'-biphenyl]-4-yl]amino)benzoic acid
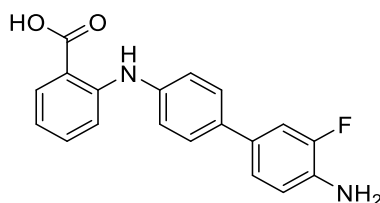
Prepared according to general procedure 5.4. Methyl 2-([3'-fluoro-4'-amino-[1,1'-biphenyl]-4-yl]amino)benzoate (50 mg, 0.155 mmol), MeOH (1.55 mL, 0.1 M), THF (0.31 mL, 0.5 M), NaOH (0.1 M, 1.55 mL, 0.1 M). Final concentration 0.048 M. Provided 2-([3'-fluoro-4'-amino-[1,1'-biphenyl]-4-yl]amino)benzoic acid as a beige solid, (28.9 mg, 0.085 mmol, 55%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 13.09 (br s, 1H), 9.69 (br s, 1H), 7.91 (dd, J = 8.0, 1.7 Hz, 1H), 7.61 – 7.53 (m, 2H), 7.40 (ddd, J = 8.7, 7.1, 1.8 Hz, 1H), 7.32 (dd, J = 13.2, 2.1 Hz, 1H), 7.29 – 7.24 (m, 3H), 7.22 (dd, J = 8.3, 2.2 Hz, 1H), 6.87 – 6.75 (m, 2H), 5.23 (br s, 2H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 169.9, 150.9 (d, J = 236.5 Hz), 146.9, 138.9, 135.5 (d, J = 13.4 Hz), 134.1, 131.9, 127.8, 126.5, 122.1, 121.6, 117.4, 116.5 (d, J = 5.2 Hz), 113.9, 112.5 (d, J = 19.0 Hz) (15 out of 17 carbon resonances observed).

^{19}F NMR (376 MHz, DMSO-d_6) δ -134.74 (dd, J = 12.5, 9.4 Hz).

M.P. 225.6 – 227.9 ^{o}C.

IR (cm^{-1}): 3345 (NH), 3268 (OH), 1670 (C=O).

MS ES+ m/z calcd for $C_{19}H_{15}FN_2O_2$ (M+H)$^+$: 323.1196, (-3.9752 neutral loss), found: 319.1446.

2-([4'-amino[1,1'-biphenyl]]amino)benzoic acid
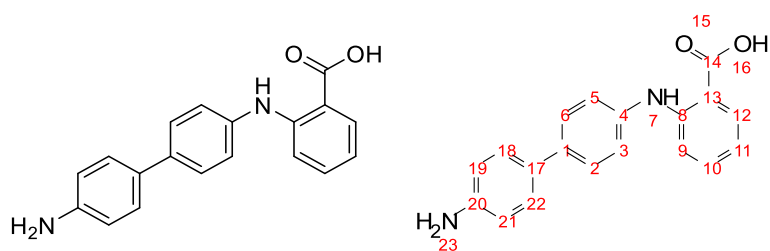
Prepared according to general procedure 5.4. Methyl 2-([4'-amino[1,1'-biphenyl]]amino)benzoate (30 mg, 0.098 mmol), MeOH (1 mL, 0.1 M), THF (0.2 mL, 0.5 M), NaOH (0.1 M, 1 mL), Final concentration 0.048 M. Provided 2-([4'-amino[1,1'-biphenyl]]amino)benzoic acid as a brown solid (15 mg, 0.05 mmol, 51%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 9.72 (br s, 1H, **7**), 7.90 (dd, J = 8.0, 1.7 Hz, 1H, **12**), 7.56 – 7.48 (m, 2H, **19**, **21**), 7.38 (ddd, J = 8.7, 7.0, 1.7 Hz, 1H, **10**), 7.36 – 7.32 (m, 2H, **3**, **5**), 7.26 – 7.20 (m, 3H, **9**, **18**, **22**), 6.76 (ddd, J = 8.0, 7.0, 1.1 Hz, 1H, **11**), 6.67 – 6.59 (m, 2H, **2**, **6**) (O*H* and N*H*$_2$ signals not observed).

^{13}C NMR (101 MHz, DMSO-d_6) δ 170.5, **14**, 148.5, **20**, 147.7, **4**, 138.7, **17**, 136.2, **1**, 134.6, **10**, 132.4, **12**, 127.5, **8**, 127.2, **3**, **5**, 126.7, **19**, **21**, 122.4, **18**, **22**, 117.6, **11**, 114.7, **2**, **6**, 114.2, **9**, 113.0, **13**.

M.P. 235 °C (dec.).

IR (cm^{-1}): 3230 (NH), 1623 (C=O), 1584.

MS ES+ m/z calcd for C$_{19}$H$_{16}$N$_2$O$_2$ (M+H)$^+$: 306.1322, found: 306.1324.

General procedure 5.5: Hydrolysis of methyl esters
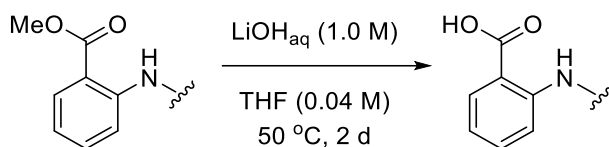
To a 25 mL RBF was added methyl ester or bis(methyl ester) (100 mg) and LiOH solution (aq. 1.0 M, 3 mL). THF was then added (6 mL) and the mixture heated to 50 °C for 2 days. After allowing to cool to RT, the reaction mixture was transferred to a sep. funnel and acidified with HCl (1.0 M), extracted with DCM (3x 10 mL), dried (Na_2SO_4) and solvent removed *in vacuo*.

2,2'-({3-fluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoic acid
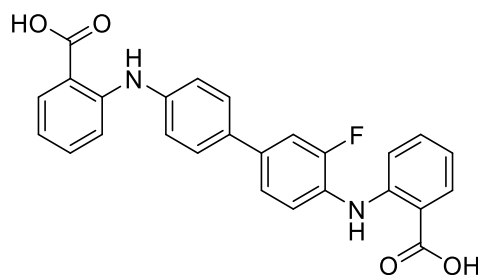
Prepared according to general procedure 5.5. Methyl 2,2'-({3-fluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate (0.226 mmol). Provided 2,2'-({3-fluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoic acid (37.9 mg, 0.085 mmol, 39%) as a yellow solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 13.15 (br s, 2H), 9.75 (br s, 2H), 7.93 (ddd, J = 7.9, 6.2, 1.7 Hz, 2H), 7.72 – 7.68 (m, 2H), 7.65 (dd, J = 12.6, 2.1 Hz, 1H), 7.57 (t, J = 8.5 Hz, 1H), 7.52 (dd, J = 8.4, 2.1 Hz, 1H), 7.44 (tdd, J = 8.3, 7.1, 1.7 Hz, 2H), 7.36 – 7.30 (m, 3H), 7.17 (dt, J = 8.5, 1.2 Hz, 1H), 6.87 – 6.84 (m, 1H), 6.85 – 6.79 (m, 1H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 170.0, 169.9, 155.1 (d, J = 243.1 Hz), 146.4, 146.2, 140.3, 135.7 (d, J = 7.3 Hz), 134.2 (d, J = 14.3 Hz), 132.7, 131.9 (d, J = 13.9 Hz), 127.3, 127.1 (d, J = 11.8 Hz), 122.9, 122.3 (d, J = 3.0 Hz), 121.1, 118.1, 117.9, 114.3, 114.0, 113.5 (d, J = 20.6 Hz), 113.3 (21 out of 24 carbon resonances found).

^{19}F NMR (471 MHz, DMSO-d_6) δ -125.70 (dd, J = 12.5, 8.7 Hz).

M.P. 152.2 – 153.5 °C.

IR (cm^{-1}): 3328 (OH), 1650 (C=O).

MS ES+ m/z calcd for $C_{26}H_{19}FN_2O_4$ (M+H)$^+$: 443.1407, found: 443.1412.

2,2'-({3,3'-difluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoic acid
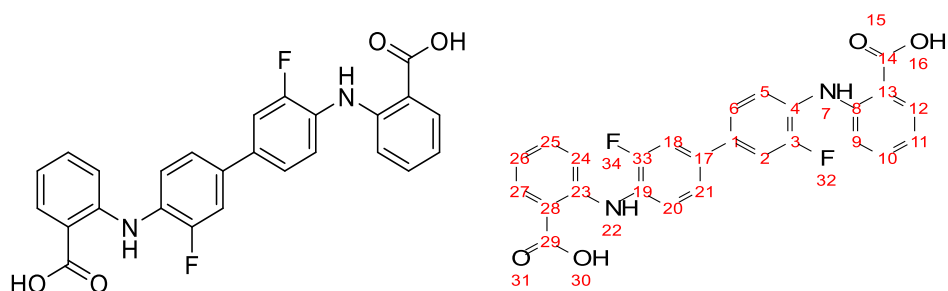
Prepared according to general procedure 5.5. Methyl 2,2'-({3,3'-difluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-methylbenzoate (0.208 mmol). Provided 2,2'-({3,3'-difluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoic acid (61.2 mg, 0.13 mmol, 61%) as a yellow solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 13.25 (br s, 2H, **16**, **30**), 9.78 (br s, 2H, **7**, **22**), 7.95 (dd, J = 8.0, 1.7 Hz, 2H, **12**, **27**), 7.79 – 7.70 (m, 2H, **2**, **18**), 7.65 – 7.54 (m, 4H, **5**, **6**, **20**, **21**), 7.47 (ddd, J = 8.6, 7.1, 1.7 Hz, 2H, **10**, **25**), 7.20 (dd, J = 8.6, 1.2 Hz, 2H, **9**, **24**), 6.87 (ddd, J = 8.0, 7.1, 1.2 Hz, 2H, **11**, **26**).

^{13}C NMR (126 MHz, DMSO-d_6) δ 167.0 , **14**, **29**, 154.8 (d, J = 242.9 Hz, **3**, **33**), 146.0 , **8**, **23**, 134.3 , **10**, **25**, 134.2 , **4**, **19**, 131.8 , **12**, **27**, 127.8 (d, J = 11.6 Hz, **1**, **17**), 122.5 (d, J = 17.9 Hz, **5**, **6**, **20**, **21**), 118.3 , **11**, **26**, 114.2 , **9**, **24**, 113.8 (d, J = 20.8 Hz, **2**, **18**), 113.4 , **13**, **28**.

^{19}F NMR (471 MHz, DMSO-d_6) δ -125.82 – -125.99 (m), **32**, **34**.

M.P. 292.5 °C (dec.).

IR (cm^{-1}): 3258 (OH), 1689 (C=O).

MS ES+ m/z calcd for $C_{26}H_{18}F_2N_2O_4$ (M+H)$^+$: 461.1313, found: 461.1320.

2-([3,3'-difluoro-4'-amino-[1,1'-biphenyl]-4-yl]amino)benzoic acid
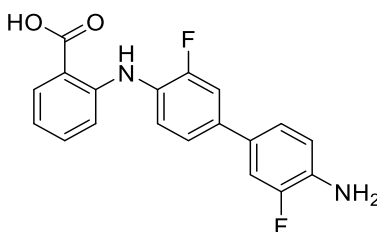
Prepared according to general procedure 5.5, provided 2-([3,3'-difluoro-4'-amino-[1,1'-biphenyl]-4-yl]amino)benzoic acid as a brown solid. (73.4 mg, 0.22 mmol, 76%).

^{1}H NMR (500 MHz, DMSO-d_6) δ 13.13 (br s, 1H), 9.84 (br s, 1H), 7.92 (dd, J = 8.0, 1.7 Hz, 1H), 7.55 (dd, J = 12.8, 2.1 Hz, 1H), 7.49 (t, J = 8.5 Hz, 1H), 7.41 (m, 3H), 7.28 (dd, J = 8.3, 2.1 Hz, 1H), 7.10 (d, J = 8.4 Hz, 1H), 6.87 – 6.77 (m, 2H), 5.32 (br s, 2H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 170.1, 155.2 (d, J = 242.7 Hz), 150.9 (d, J = 236.8 Hz), 146.4, 136.1 (d, J = 13.0 Hz), 135.8 (d, J = 7.4 Hz), 134.0, 131.8, 126.5, 126.3, 123.0, 122.4 (d, J = 2.5 Hz), 121.6 (d, J = 2.9 Hz), 117.8, 116.4 (d, J = 5.2 Hz), 113.7 (d, J = 19.9 Hz), 113.0 (d, J = 19.0 Hz), 112.7 (d, J = 19.0 Hz) (17 out of 19 carbon resonances found).

^{19}F NMR (376 MHz, DMSO-d_6) δ -125.24 – -126.08 (m), -134.71 (dd, J = 13.2, 9.6 Hz).

M.P. > 200 ºC (dec.).

IR (cm^{-1}): 3345 (NH), 3269 (OH), 1669 (C=O).

MS ES+ Calcd for $C_{19}H_{14}F_2N_2O_2$ (M+H)$^+$: 343.1161, found: 343.1163.

6 Efforts Towards the Binding Optimisation of Indole-3-Carbinol

6.1 Introduction

Indole-3-carbinol (I3C, **6.2**) has been identified as a phytochemical derived from the hydrolysis of glycobrassicin (Figure 6-1, a, **6.1**).[123] This natural product is found in the *Brassica* family of vegetables such as broccoli (*Brassica oleracea var. italica*), Brussel sprouts (*Brassica oleracea var. gemmifera*) and cabbage (*Brassica oleracea var. capitata*). I3C and its metabolites are known chemo-preventative agents and have a broad range of activity against multiple aspects of cancer cell-cycle progression. As discussed previously I3C has been investigated against multiple cancer targets (see section 2.2.2).[270] The metabolites of I3C are produced *via* an acid-catalysed dehydration from the vinyl hemiaminol moiety to provide a reactive aza-Michael acceptor (Figure 6-1, b, **6.3**), which is able to react with more I3C to provide several oligomeric and cyclic metabolites in acidic media (Figure 6-1, c, **6.4** – **6.8**), each of which demonstrate their own pharmacological activities.[270]
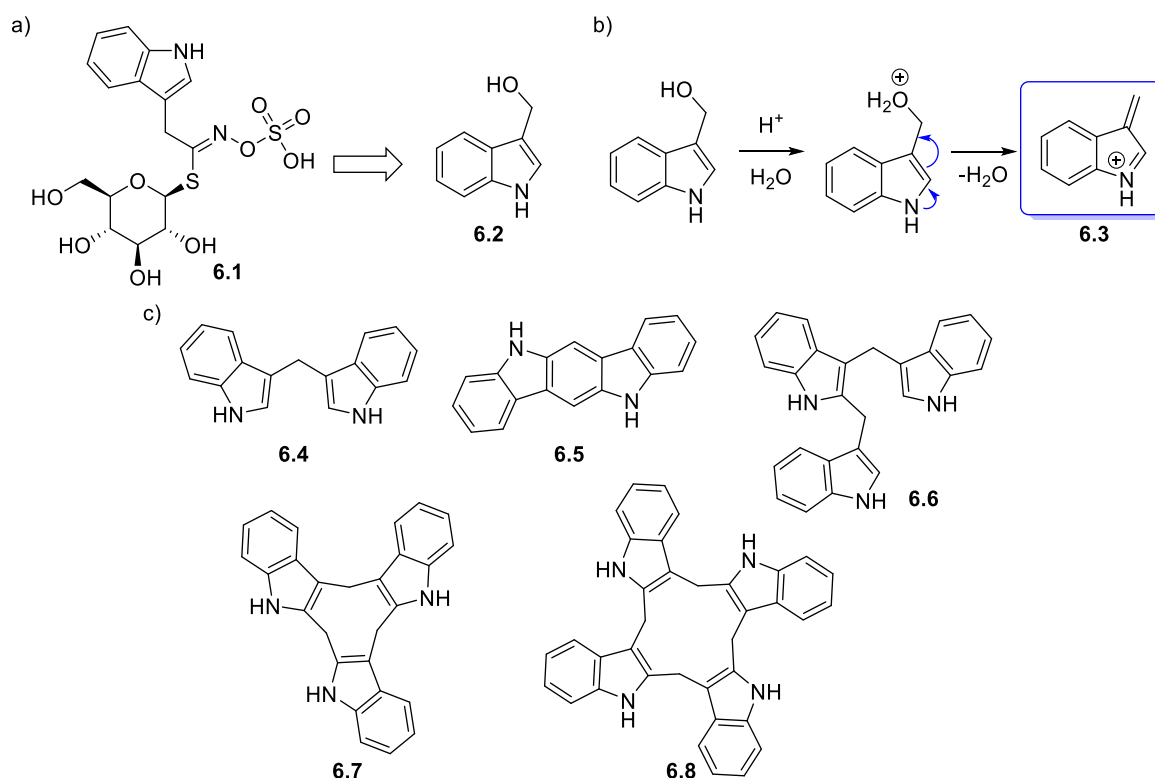
Figure 6-1: a) glycobrassicin, b) formation of aza-Michael acceptor from I3C, c) acid-catalysed oligomerisation products.

Based upon this information and the previous work performed on both I3C and its metabolites as discussed in section 2.2.2, it was decided that I3C would be an interesting natural product to investigate as a WWP2 HECT domain inhibitor. Seeing as there is high homology between the HECT domains of NEDD4-1 and WWP2, it seemed reasonable to also investigate some of the compounds synthesised from investigations into NEDD4-1 inhibition. Additionally, 3,3'-diindolylmethane (DIM), a metabolite of I3C, seemed like opportunistic synthetic targets to probe their activity against WWP2.

6.2 <u>Aims and Objectives</u>

As previously described, the synthesis of derivatives with improved IC_{50} values is the second main objective in this thesis. This would be achieved by splitting it into three smaller objectives. For I3C, a known binder, the choice of analogues would be decided by analysis of the co-crystal structure obtained from the soaking studies and / or the choice of analogues from validated molecular docking methodology (see section 3.5.4 for details).

As shown in Figure 6-2, the main interactions between I3C and the HECT domain of WWP2 are focussed on the alcohol functional group. The alcohol is participating in three hydrogen bonding interactions with Thr627, Tyr587 and Ser646. The indole system is participating in side-on π-π interactions with Tyr587.
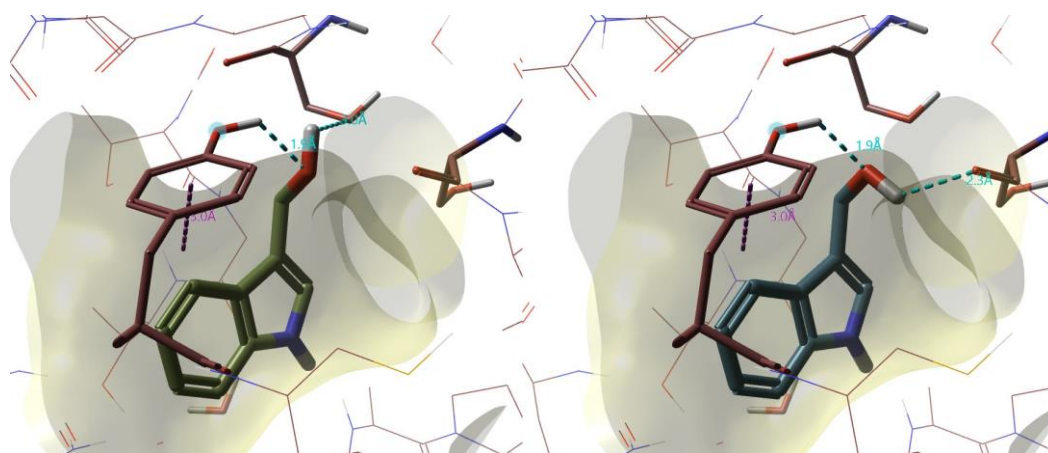
Figure 6-2: I3C bound to WWP2 with important residues highlighted.

As a first objective, exploration of the importance of the alcohol functional group, by creating analogues which vary in a polar functional group at the 3-methyl position, as well as removing functionality to probe its importance was investigated. Secondly, there is a significant amount of void space surrounding the phenyl ring of the indole system, and the

functionalisation of this with aromatic groups was targeted to explore further aromatic binding possibilities, which constitutes the second objective.

Initially, synthesising compounds previously identified as NEDD4-1 HECT domain binders was considered. The functionalisation of the indole nitrogen with phenyl, benzyl, and sulfonamide groups was pursued, as these have been demonstrated in the literature to be effective binders to Nedd4-1 HECT domains. Lastly the metabolite 3,3'-diindolylmethane (DIM, **6.4**) was also targeted for synthesis due also to its known anti-cancer properties.[124] These targets are more exploratory as, based on the crystal structure data, there is predicted a larger steric clash here and if these compounds do bind within this binding site, it is likely in a significantly different orientation than I3C.

6.3 Results and discussion

6.3.1 Synthesis IX: Modification of the 3-position

For the first objective, a series of compounds were targeted with varying functional groups present where the 1° alcohol is situated (Figure 6-3). These could be compared to I3C, to further understand the importance of polar interactions as this position.
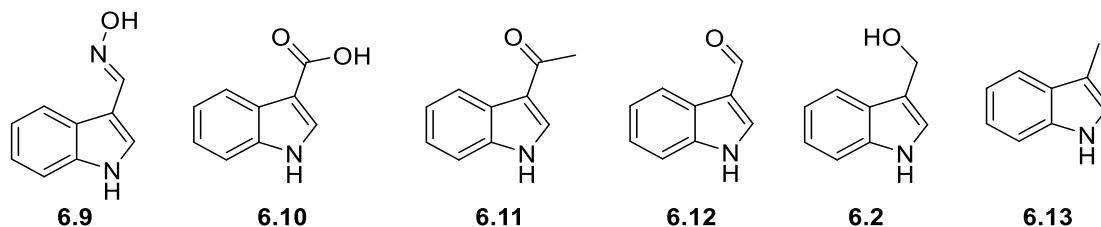
Figure 6-3: Targeted I3C derivatives varying at the 3-methyl position.

Several variations of the 3-position were obtained either from their availability in the laboratory or from commercial vendors and could also be submitted for WWP2 inhibition assays (Figure 6-4, **6.9** – **6.15**). Also, several compounds with additional functionalisation about the indole skeleton were obtained from the laboratory and also submitted (**6.16** – **6.20**). This library acquisition meant successfully meeting the first objective to vary the 3-position.
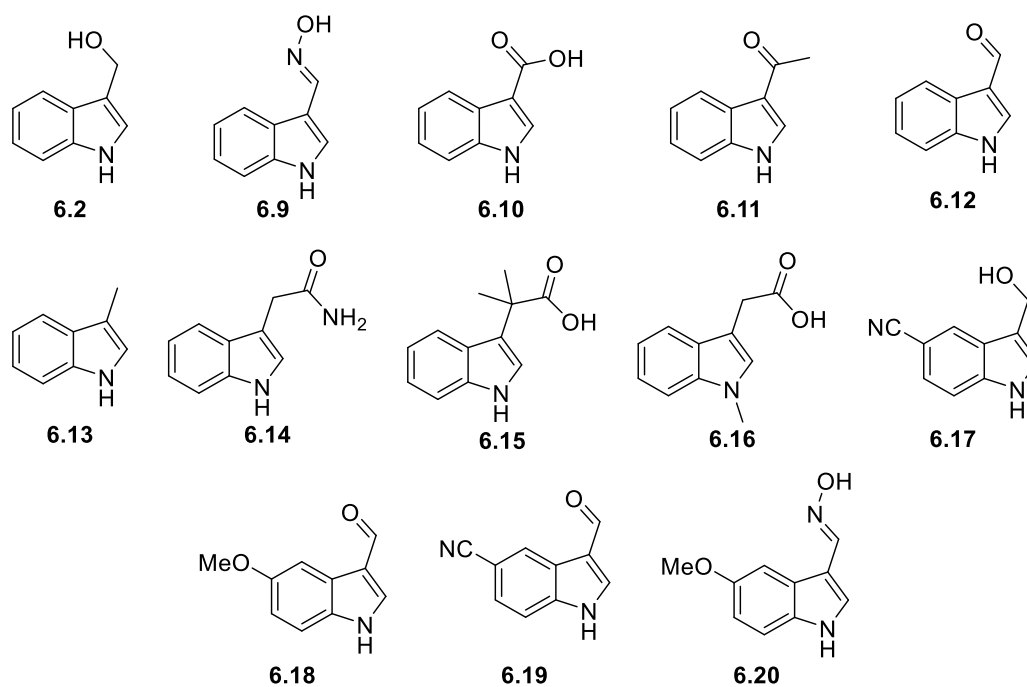
Figure 6-4: Summary of I3C analogues with 3-position variations.

6.3.2 Synthesis X: Exploration of Void Space Surrounding the I3C Phenyl Ring

The second objective of this project was to explore the void space surrounding I3C in the hopes of gaining additional non-polar / aromatic interactions. Therefore, phenyl derivatives were planned for synthesis (**6.21** – **6.24**, Figure 6-5) where the phenyl group was attached to positions 4, 5, 6 and 7 of the indole system.
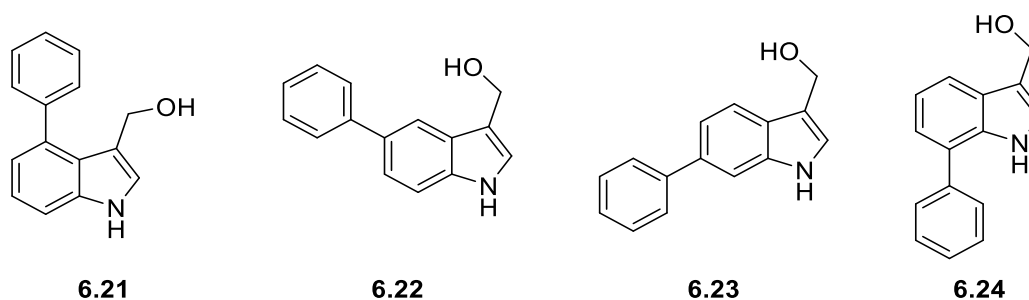
Figure 6-5: Target compounds of the second objective to explore void space.

Halogenated or borylated indole precursors at the 5-, 6- and 7-positions are commercially available and under the Suzuki cross-coupling conditions shown (Scheme 6-1) with either bromobenzene or phenyl boronic acid, the synthetic intermediates **6.30** – **6.32** were readily produced in high yield.[204]

6.25 X = 5-Br   Y = **6.28** Br or
6.26 X = 6-B(OH)$_2$ **6.29** B(OH)$_2$
6.27 X = 7-Br

(Boronic acid in 1.5 equiv.)

6.30 *quant.*   **6.31** *92%*   **6.32** *84%*
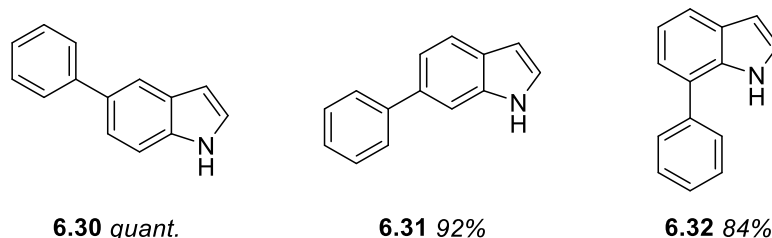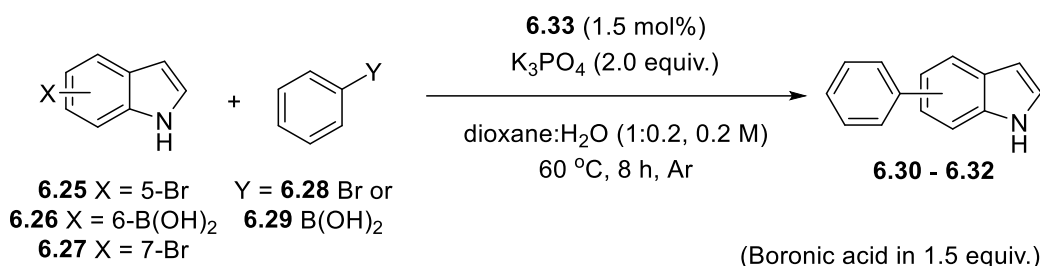
Scheme 6-1: Synthesis of intermediates 6.30 – 6.32.

The use of a palladium precatalyst such as **6.33** was probably the reason for these high yields, as ***int-6.34***, formed after amido complex formation by the action of a base on **6.33**, is able to undergo facile reductive elimination from the stable Pd^{II} palladacycle, with loss of carbazole **6.35** to the catalytically active Pd^{0}L$_1$ species **6.36** (Scheme 6-2). It has been shown that monoligated Pd^0 species are the most reactive in Suzuki cross-coupling reactions, as discussed, see section 5.4.2.[240]

SPhosPdG2, **6.33**   ***int-6.34***   **6.35**   catalyitcally active species **6.36**
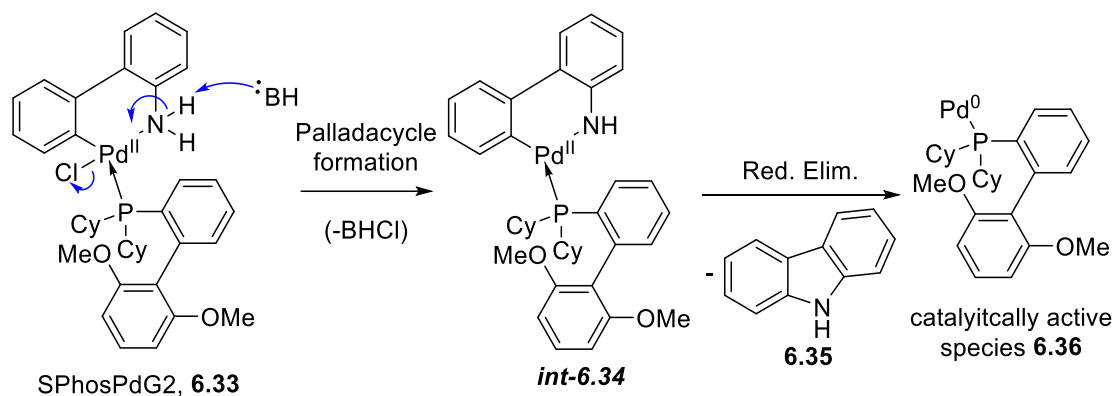
Scheme 6-2: Formation of catalytically active Pd^{0}L$_1$ species.

Methodologies utilising the indole skeleton are extremely common, and newer methods such as C-H activation now allow practically all positions to be activated towards functionalisation in some way.[271] They offer advantages over more classical reactions, as no pre-functionalisation is necessary because the otherwise inert C(sp^2)-H bond is activated, usually under transition metal (TM) catalysis. In order to imbue regioselectivity, directing groups (DG) are often used. These are Lewis basic functional groups which can co-ordinate to TM centres, directing them to specific C-H bonds, creating metallacycles *via* C-H activation, which can undergo further reaction to functionalise the activated C-H

344

position. In this case, a method has been reported for the C-4 C-H functionalisation of indole-3-carboxaldehydes, where the aldehyde is acting as the directing group. Under Pd catalysis utilising iodobenzene **6.37** as coupling partner, 4-phenylindole-3-carboxaldehyde **6.38** was produced in reasonable yield (Scheme 6-3).[272]
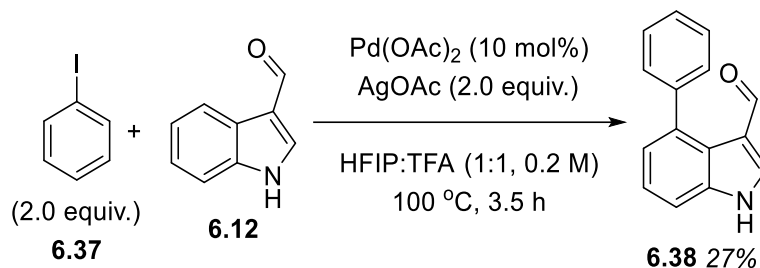
Scheme 6-3: Production of 4-phenylindole-3-carboxaldehyde **6.38** from a C-H activation methodology.

The proposed mechanism is thought to begin by oxidative addition of **6.37** to the *in-situ* formed Pd^0 centre forming **int-6.40** (Scheme 6-4). AgOAc could arguably abstract the iodide ligand from **int-6.40**, forming the coordinatively unsaturated **int-6.41** and insoluble AgI. Co-ordination of **6.12** to **int-6.41** can form **int-6.42**, putting the C-4 C-H bond into a position proximal to the Pd^{II} centre. This would be followed by a concerted metalation deprotonation (CMD) event to form the 6-membered palladacycle **int-6.43**. From here, reductive elimination can occur to forge the C(sp^2)-C(sp^2) bond in **6.38** and form Pd^0, which can then dissociate from the aldehyde, providing **6.38**.
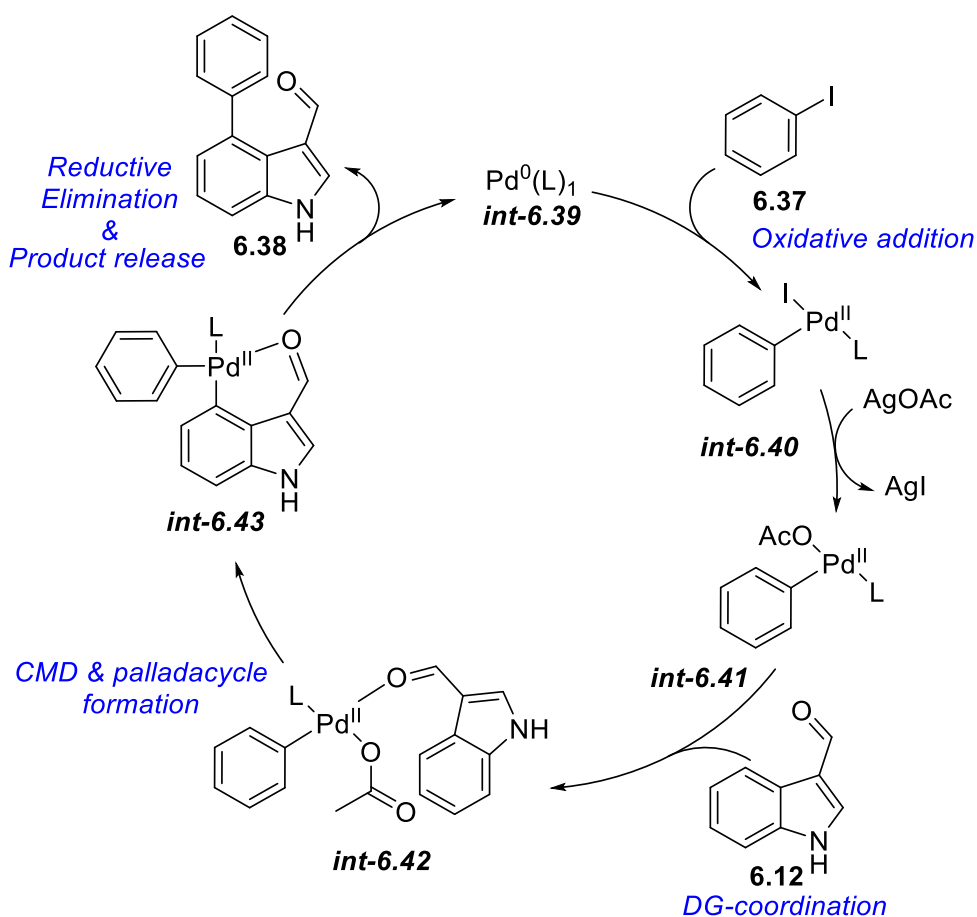
Scheme 6-4: Postulated mechanism for C-4 C-H activation of indole-3-carboxaldehyde.

From here, for phenylindoles **6.30**, **6.31** and **6.32**, Vilsmeier-Haack formylation was conducted in high yields and products taken forward without further purification post ^{1}H NMR analysis (Scheme 6-5).[273] After this, simple reduction of the aldehydes **6.38**, **6.44** – **6.46** to the alcohols was performed using NaBH$_4$ in THF:EtOH, providing all the desired phenylindole-3-carbinols **6.21** – **6.24** in acceptable yields and purity for bioassay. This represents successfully meeting the second objective to produce a small library of arylated indole compounds to explore the void space surrounding I3C in the WWP2 co-crystal structure.

6.21 52%
(*14%* over 2 steps)

6.22 53% (over 2 steps)
53% (over 3 steps)

6.23 84% (over 2 steps)
77% (over 3 steps)

6.24 98% (over 2 steps)
82% (over 3 steps)
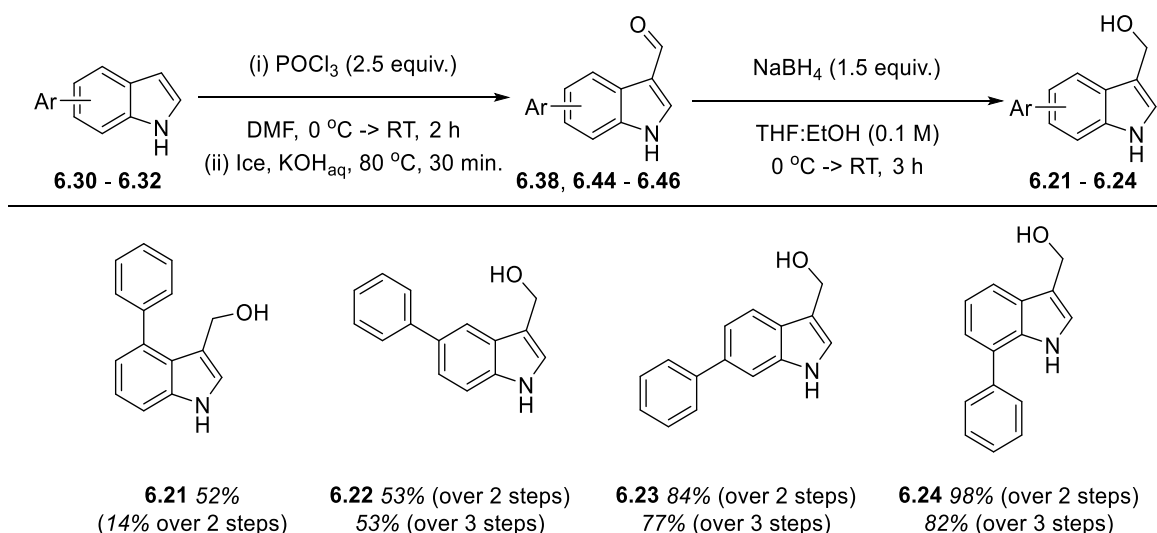
*Scheme 6-5: Vilsmeier-Haack formylation followed by reduction to provide **6.21** – **6.24** final products.*

6.3.3 <u>Synthesis XI: Synthesis of I3C Nedd4-1 HECT Domain Inhibitor Variations</u>

Several *N*-functionalised I3C derivatives were targeted for synthesis, as *N*-functionalised indole-3-carbinol compounds were shown to be inhibitory towards similar HECT domain E3 enzymes such as Nedd4-1. These included *N*-phenyl **6.47**, *N*-benzyl **6.48**, and sulfonamide derivatives **6.49**, which were targeted for synthesis (Figure 6-6). Additionally, DIM **6.50** was also targeted along with some phenyl hydrazones **6.51**, which were identified from an undergraduate student's project.[274]

6.47   **6.48**   **6.49**   **6.50**   **6.51**
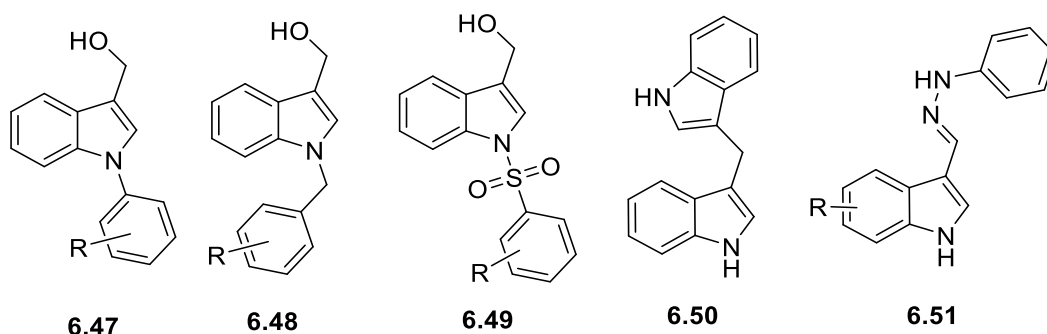
Figure 6-6: I3C analogues and metabolite derivative targeted as part of the 3rd objective in this project.

As part of this series, the *N*-phenyl derivative was synthesised *via* firstly an Ullman cross-coupling using Cu₂O in DMF over 3 days for an acceptable yield (Scheme 6-6, **6.53**).[126]

347

Subsequently, simple reduction was performed to provide the desired *N*-phenyl derivative **6.47a**.
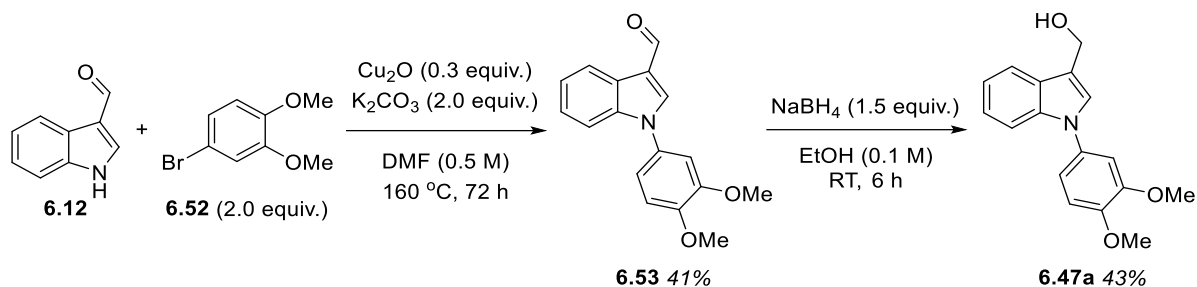
*Scheme 6-6: Ullman cross-coupling followed by reduction to access **6.47a**.*

Next, *N*-benzyl indole-3-carbinol **6.48a** and *N*-3,4-dimethoxybenzyl indole-3-carbinol **6.48b** (Scheme 6-7, a) were targeted for synthesis. Methylating vanillin under standard conditions provided **6.55** in high yield (Scheme 6-7, b).[275] With **6.55** in hand, the aldehyde was reduced, and the benzyl alcohol **6.56** converted to the benzyl bromide **6.57** under Appel conditions.[276,277]
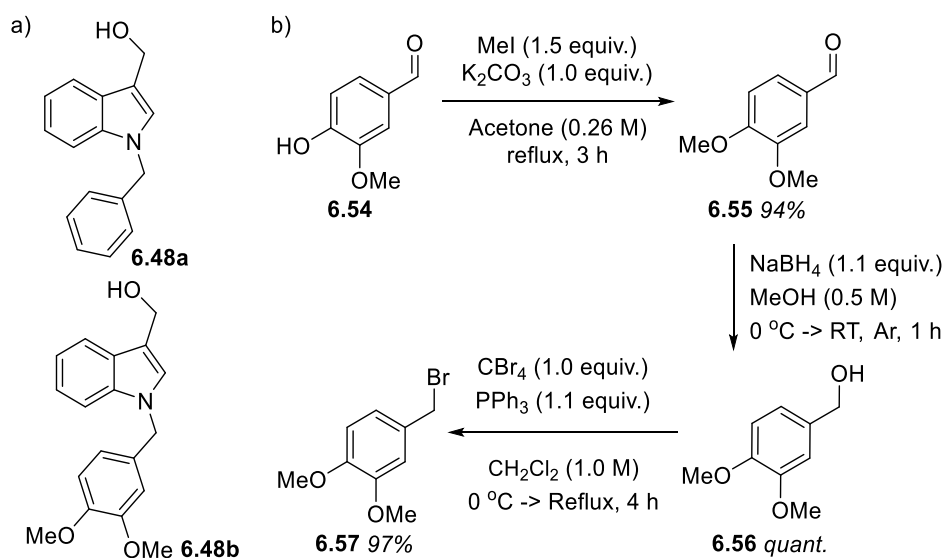
*Scheme 6-7: a) Targeted N-benzyl derivatives **6.48a** & **6.48b**, b) synthesis of **6.57** from **6.54**.*

The Appel reaction proceeds by attack of the phosphorus lone pairs in triphenylphosphine **6.58** onto tetrabromomethane **6.59** to form a bromophosphonium cation and tribromocarbanion, existing as a tight ion pair ***int-6.60*** (Scheme 6-8). Deprotonation of the benzyl alcohol **6.56** provides the alkoxide ***int-6.61*** and bromoform **6.62**. The alkoxide can then attack the bromophosphonium cation of ***int-6.60*** (***int-6.60'***) in an S$_N$2 fashion displacing a bromide anion, which can then displace the oxygen atom of the formed ***int-6.63***, forming the desired benzyl bromide **6.57** and triphenylphosphine oxide.
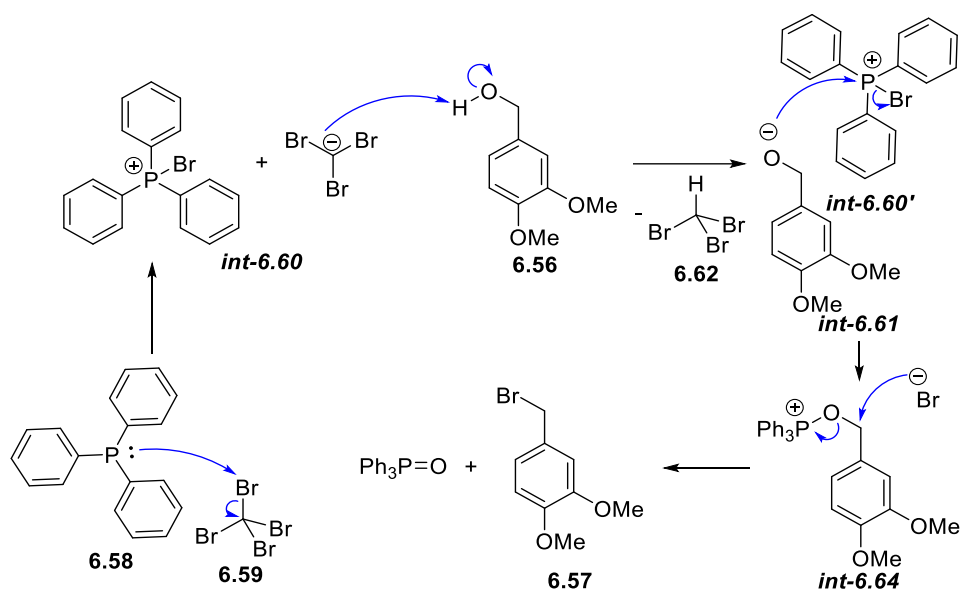
Scheme 6-8: Appel reaction mechanism.

After the synthesis of the benzyl bromide derivative **6.57**, this was reacted with indole-3-carboxaldehyde **6.12** to form the *N*-benzyl derivative **6.64** (Scheme 6-9).[278] Additionally, benzyl bromide **6.65** was also reacted under identical conditions. Both *N*-functionalised indole-3-carboxaldehdyes **6.64** and **6.66** were then reduced to from the desired *N*-functionalised indole-3-carbinols **6.48a** & **6.48b** in good yields for both steps.
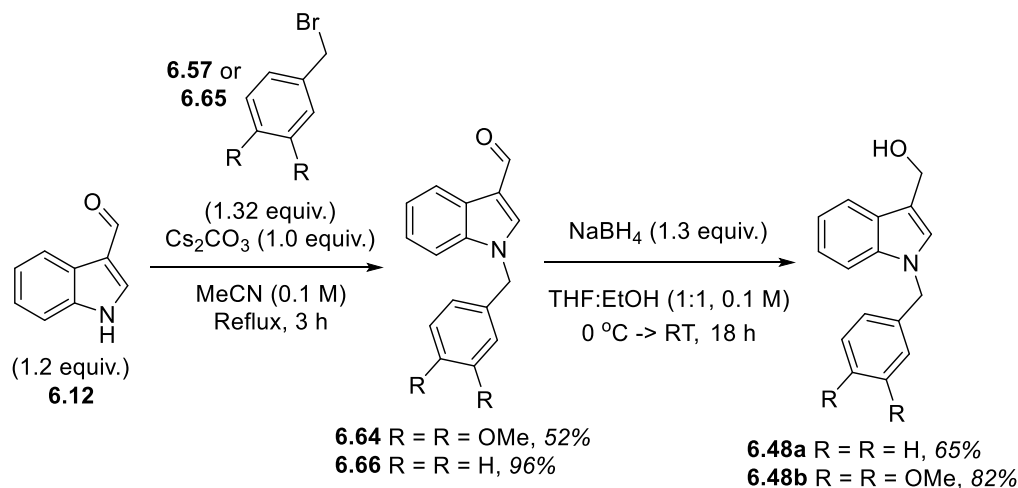
Scheme 6-9: Benzyl bromide reactions followed by reduction to access 6.48a and 6.48b.

Targeting the sulfonamide series, initial conditions utilising potassium carbonate in DCM at reflux were not particularly high yielding for **6.68c** (Scheme 6-10), although enough material was obtained for **6.68a** and **6.68b**.[279] Using phase transfer catalysis with **6.69c** and **6.69d** allowed higher yields of **6.68c** and **6.68d** to move forwards in these syntheses.[280] Afterwards, the reduction step was high yielding in all cases to provide four *N*-sulfonamide indole-3-carbinol derivatives **6.49a** – **6.49d** in excellent yields.
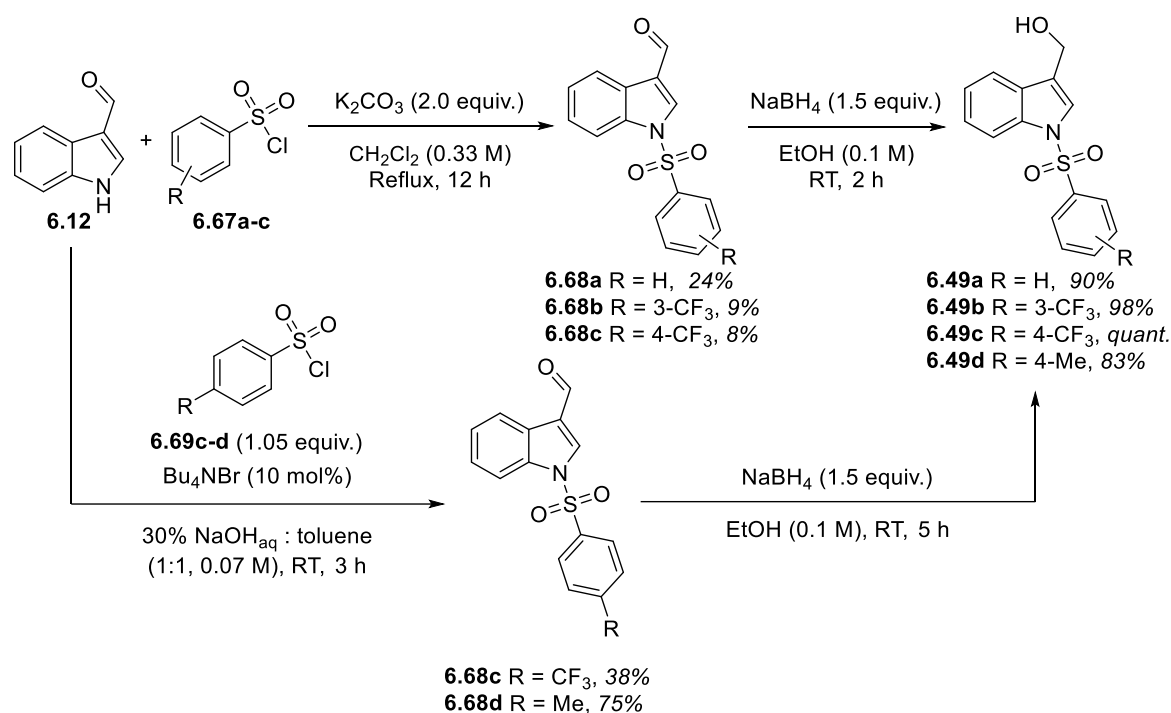
Scheme 6-10: Sulfonamide formation and reduction steps to produce sulfonamide I3C derivatives ***6.49a – 6.49d****.*

3,3'-Diindolylmethane **6.50** was also produced *via* a literature procedure from **6.2** (Scheme 6-11).[281] 3,3'-Diindolylmethane is a known metabolite of indole-3-carbinol and was produced under basic conditions at reflux followed by recrystallisation.
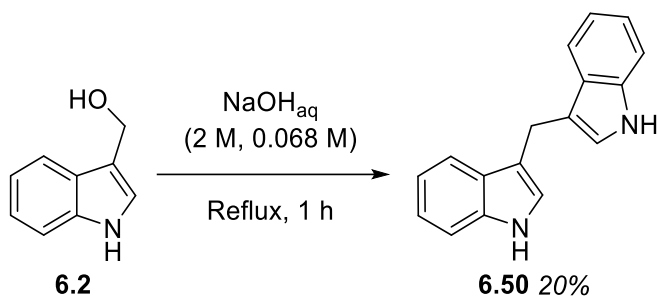
Scheme 6-11: Production of DIM ***6.50****.*

Some phenylhydrazones (**6.51a-c**) were also synthesised by an undergraduate student in a simple condensation reaction of indole-3-carboxaldehyde and derivatives (**6.12a-c**) with *N*-phenylhydrazine **6.70** (Scheme 6-12).[274]
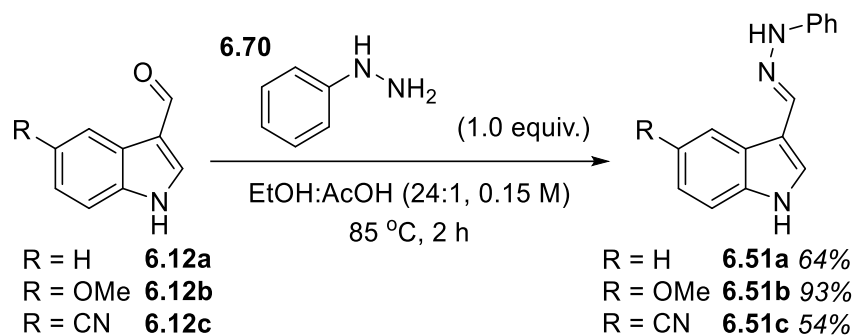
Scheme 6-12: Synthesis of Indole-3-carboxaldehyde phenylhydrazones **6.51a-c**.

Below is a summary of the compounds synthesised (Figure 6-7). In total of 11 derivatives were synthesised in good yield for this objective, all in sufficient purity for biological analysis.
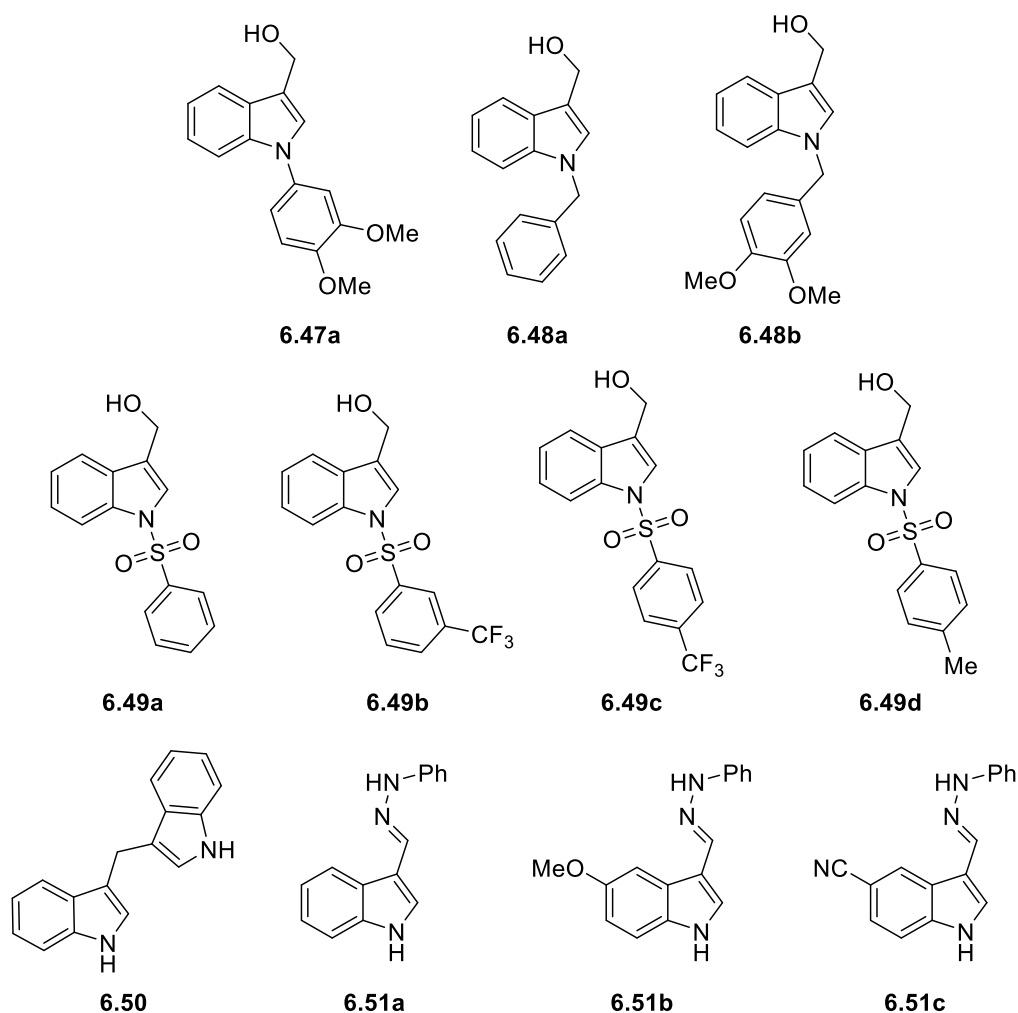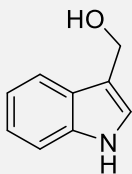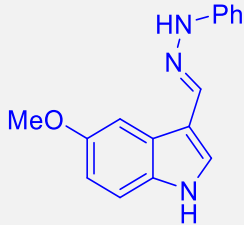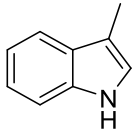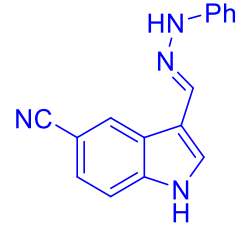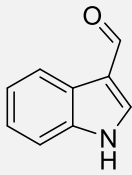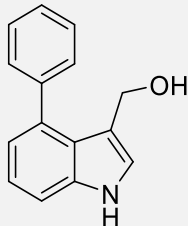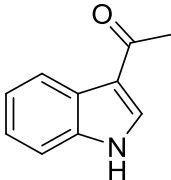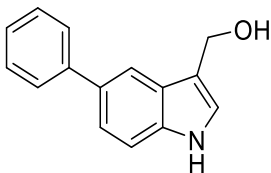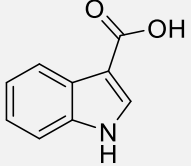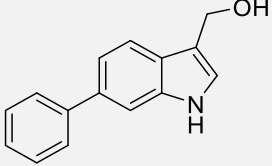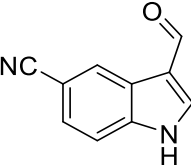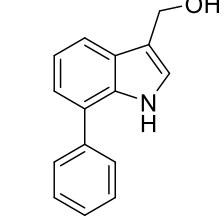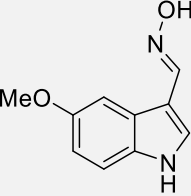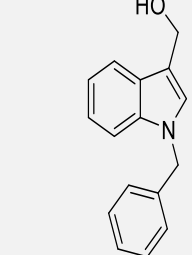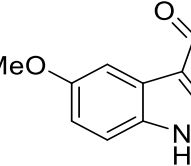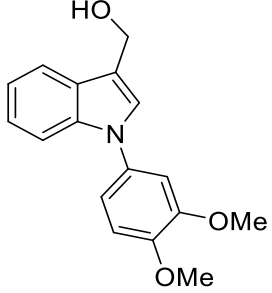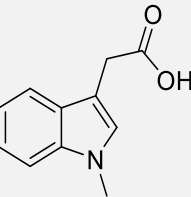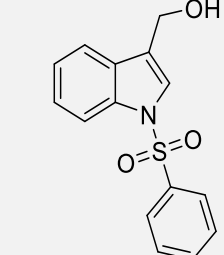
Figure 6-7: Summary of N-functionalised I3C derivatives synthesised.

6.4 Biological results

Presented in Table 6-1 are the biological results from the I3C analogue series. Indole-3-carbinol itself is a known binder to the HECT domain (see section 2.3) but not inhibitory as observed from the biochemical assay. As can be seen, only the phenylhydrazone derivatives provided IC_{50} values of the three objectives, which represents an improvement over I3C (Entries 2, 4 and 27), although entries 2 and 4 did not dissolve at 1 mM in 10% DMSO and entry 27 at 10 mM, so these results need to be treated with caution. Additionally, entries 24 (**6.49c**) and 28 (**6.50**) were observed to be inhibitory at 1 mM but did not provide IC_{50} values. Most of these compounds dissolved (see exceptions in Table 6-1 caption), but did not provide any inhibitory activity, so further DSF tests were conducted.

Entry	Structure	IC_{50} (µM)	Entry	Structure	IC_{50} (µM)
1 (6.2)		Binder	2 (6.51b)		388.1**
3 (6.13)		N/A	4 (6.51c)		245.6*
5 (6.12)		N/A	6 (6.21)		N/A
7 (6.11)		N/A*	8 (6.22)		N/A

9 **(6.10)**	(indole-3-carboxylic acid)	N/A	**10** **(6.23)**	(6-phenyl-1H-indol-3-yl)methanol	N/A
11 **(6.19)**	5-cyano-1H-indole-3-carbaldehyde	N/A	**12** **(6.24)**	(7-phenyl-1H-indol-3-yl)methanol	N/A
13 **(6.20)**	5-methoxy-1H-indole-3-carbaldehyde oxime	N/A	**14** **(6.48a)**	(1-benzyl-1H-indol-3-yl)methanol	N/A
15 **(6.18)**	5-methoxy-1H-indole-3-carbaldehyde	N/A	**16** **(6.47a)**	(1-(3,4-dimethoxyphenyl)-1H-indol-3-yl)methanol	N/A*
17 **(6.16)**	2-(1-methyl-1H-indol-3-yl)acetic acid	N/A	**18** **(6.49a)**	(1-(phenylsulfonyl)-1H-indol-3-yl)methanol	N/A
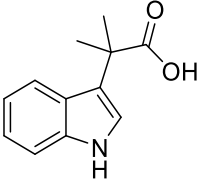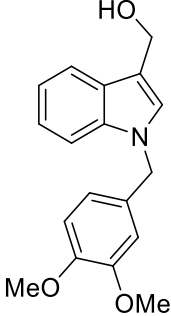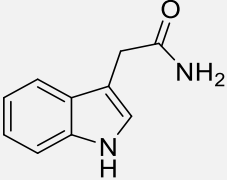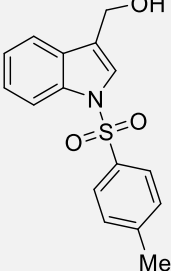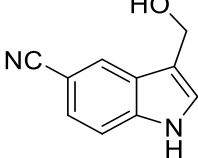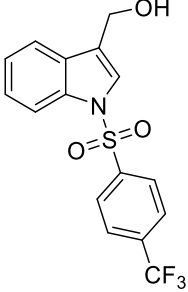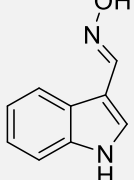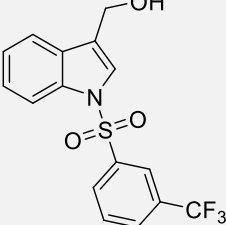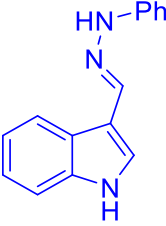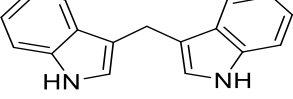
19 **(6.15)**		N/A	**20** **(6.48b)**		N/A
21 **(6.14)**		N/A	**22** **(6.49d)**		N/A*
23 **(6.17)**		N/A	**24** **(6.49c)**		N/A*,‡
25 **(6.9)**		N/A	**26** **(6.49b)**		N/A**
27 **(6.51a)**		486.8**	**28** **(6.50)**		N/A*,†

*Table 6-1: Biological results of I3C series. *Not soluble at 10 mM (10% DMSO), **Not soluble at 1 mM (10% DMSO), †Slight inhibition at 1 mM, ‡40% inhibition at 1 mM*

Table 6-2 shows the three compounds that provided stabilising DSF results, only one of which provided an IC_{50} value (entry 16, Table 6-1).
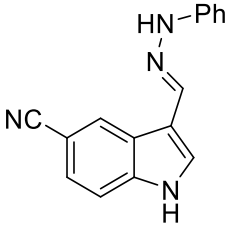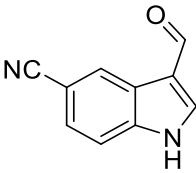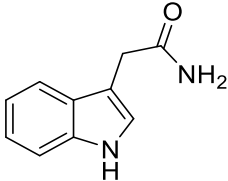
Structure			
DSF results	Stabilising	Stabilising	Stabilising

Table 6-2: DSF results from the I3C series.

6.5 I3C Conclusions

To summarise, for the first objective the acquisition of commercial indole-3-carbinol variants at the 3-position was readily achieved in high purity for bioassay. Production of phenyl moieties for the second objective was also successful by employing cross-coupling conditions followed by standard formylations and reductions. The last objective looking into producing literature HECT domain inhibitors, a metabolite and phenylhydrazone derivatives was also largely successful.

From the biological results, the vast majority of compounds did not provide any IC_{50} values and those that did also had solubility issues, which could be a cause for concern. One sulfonamide derivative (entry 24, Table 6-1) and DIM (entry 28, Table 6-1) were observed to have some inhibitory activity at high concentrations (1 mM), but also had associated solubility issues. If the phenylhydrazones and sulfonamide derivatives do bind where it has been proposed, then they most likely orient themselves significantly differently to I3C, to accommodate the large phenylhydrazone and arylsulfonamide moieties. DIM also most-likely binds in a significantly different orientation due to the absence of the hydroxyl group, loosing those thought to be important interactions observed for I3C.

It is now understood that, at least from DSF experiments that some I3C derivatives do interact with the WWP2 HECT domain construct. However, trusting the IC_{50} values associated with entries 2, 4 and 27 and the apparent observed inhibition of entries 24 and 28 (Table 6-1) requires further validation to rule out potential nuisance behaviour. It therefore cannot be said that an improvement upon the binding of I3C has been achieved and an understanding of the SAR could not be obtained to any real extent due to lack of data.

6.6 <u>Experimental</u>

<u>5-Phenyl-1*H*-indole</u>
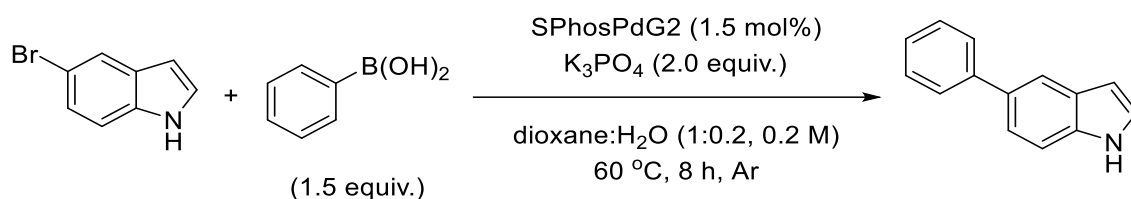
To a microwave vial was added 5-bromoindole (200 mg, 1.02 mmol), phenyl boronic acid (186.6 mg, 1.53 mmol, 1.5 equiv.), K_3PO_4 (433 mg, 2.0 equiv.) and SPhosPdG2 (11.0 mg, 1.5 mol%). This was sealed with a suba seal, evacuated under vacuum, and backfilled with argon thrice. Afterwards, a degassed dioxane:water mixture (1:1, 5.1 mL, 0.2 M) was added *via* syringe. The vial was then sealed with a Teflon cap under argon flow and placed in a preheated metal block at 60 °C for 18 h. The cooled reaction mixture was filtered through a Celite® pad, washing with EtOAc (50 mL). The filtrate was then concentrated under reduced pressure and the crude material was purified by column chromatography (9:1 Hex. : EtOAc) to provide an off-white solid of 5-phenyl-1*H*-indole. (210 mg, 1.08 mmol, quant.).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 8.17 (br s, 1H), 7.87 (q, *J* = 1.1 Hz, 1H), 7.70 – 7.62 (m, 2H), 7.49 – 7.40 (m, 4H), 7.31 (ddt, *J* = 7.8, 6.8, 1.3 Hz, 1H), 7.26 – 7.23 (m, 1H), 6.62 (dd, *J* = 3.2, 2.0 Hz, 1H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 142.7, 135.5, 133.6, 128.8, 128.5, 127.5, 126.5, 125.0, 122.1, 119.4, 111.4, 103.2.

IR (cm^{-1}): 3406 (NH).

Followed literature procedure.[204] Data in-line with literature.[282]

6-phenyl-1*H*-indole
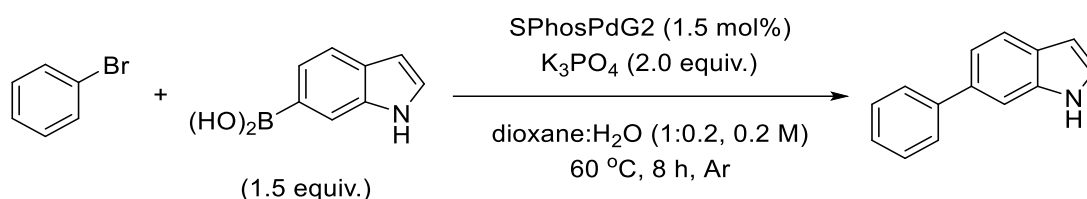
To a microwave vial was added bromobenzene (0.107 mL, 1.02 mmol), 6-indolylboronic acid (246.3 mg, 1.53 mmol, 1.5 equiv.), K_3PO_4 (433 mg, 2.0 equiv.) and SPhosPdG2 (11.0 mg, 1.5 mol%). This was sealed with a suba seal, evacuated under vacuum, and backfilled with argon three times. Afterwards, a degassed dioxane : water mixture (1:1, 5.1 mL, 0.2 M) was added *via* syringe. The vial was then sealed with a Teflon cap and placed in a preheated metal block at 60 °C for 18 h. The cooled reaction mixture was filtered through a Celite® pad, washing with EtOAc (50 mL). The filtrate was concentrated under reduced pressure and the crude was purified by column chromatography (9:1 Hex. : EtOAc) to provide a white crystalline solid of 6-phenyl-1*H*-indole. (184 mg, 0.95 mmol, 94%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 8.19 (br s, 1H), 7.71 (dt, *J* = 8.2, 0.8 Hz, 1H), 7.69 – 7.64 (m, 2H), 7.61 (dt, *J* = 1.6, 0.8 Hz, 1H), 7.48 – 7.42 (m, 2H), 7.40 (dd, *J* = 8.2, 1.6 Hz, 1H), 7.36 – 7.30 (m, 1H), 7.25 (dd, *J* = 3.2, 2.4 Hz, 1H), 6.59 (ddd, *J* = 3.1, 2.1, 1.0 Hz, 1H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 142.5, 136.5, 135.8, 128.8, 127.5, 127.4, 126.7, 124.9, 121.0, 120.0, 109.7, 102.7.

IR (cm^{-1}): 3379 (NH).

Followed literature procedure.[204] Data in-line with literature.[283]

7-phenyl-1*H*-indole
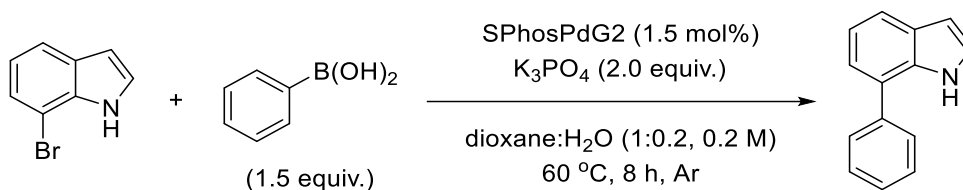
To a microwave vial was added 7-bromoindole (200 mg, 1.02 mmol), phenyl boronic acid (186.6 mg, 1.53 mmol, 1.5 equiv.), K$_3$PO$_4$ (433 mg, 2.0 equiv.) and SPhosPdG2 (11.0 mg, 1.5 mol%). This was sealed with a suba seal, evacuated under vacuum, and backfilled with argon three times. Afterwards, a degassed dioxane : water mixture (1:1, 5.1 mL, 0.2 M) was added *via* syringe. The vial was then sealed with a Teflon cap and placed in a preheated metal block at 60 °C for 18 h. The cooled reaction mixture was filtered through a Celite® pad, washing with EtOAc (50 mL). The filtrate was concentrated under reduced pressure and the crude material was purified by column chromatography (2% EtOAc in Hex.) to provide a colourless oil of 7-phenyl-1*H*-indole. (162 mg, 0.84 mmol, 82%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 8.43 (br s, 1H), 7.73 – 7.65 (m, 3H), 7.55 (t, *J* = 7.6 Hz, 2H), 7.44 (t, *J* = 7.4 Hz, 1H), 7.34 – 7.20 (m, 3H), 6.69 – 6.63 (m, 1H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 139.4, 133.9, 129.3, 128.40, 128.37, 127.5, 125.7, 124.5, 122.0, 120.5, 120.2, 103.2.

IR (cm^{-1}): 3424 (NH).

Followed literature procedure, data in accordance with literature values.[204]

4-phenyl-1*H*-indole-3-carboxaldehyde[272]
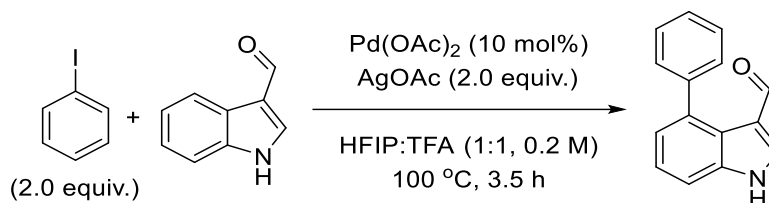
To a 30 mL microwave vial was added in air indole-3-carboxaldehyde (200 mg, 1.37 mmol), Pd(OAc)$_2$ (30.1 mg, 10 mol%), and AgOAc (444.1 mg, 2.0 equiv.). To this was added HFIP and TFA simultaneously (3.43 mL each, total 6.86 mL, 0.2 M) and iodobenzene (0.310 mL, 2.0 equiv.). This was placed in a pre-heated sand bath at 100 °C for 4 h. The cooled solution was filtered through a pad of Celite®, washing with EtOAc (75 mL), the solvent was removed, and the crude material purified by column chromatography (8:2 -> 6:4 Hexane : EtOAc) to provide 4-phenyl-1*H*-indole-3-carboxaldehyde a brown solid. (83.1 mg, 0.38 mmol, 27%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 9.49 (d, J = 1.0 Hz, 1H), 8.92 (br s, 1H), 8.03 (d, J = 3.1 Hz, 1H), 7.55 – 7.49 (m, 2H), 7.49 – 7.43 (m, 3H), 7.43 – 7.38 (m, 1H), 7.34 (dd, J = 8.1, 7.3 Hz, 1H), 7.19 (dd, J = 7.3, 1.0 Hz, 1H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 187.0, 142.0 136.9, 135.9, 130.7, 129.1, 128.8, 127.9, 124.3, 124.0, 123.6, 119.7, 111.1.

Followed literature procedure, data in accordance with literature data.

General procedure 6.1: formylation of aryl-indole derivatives
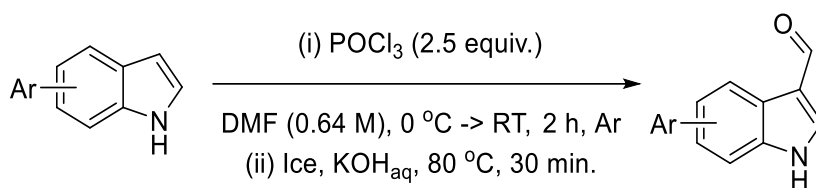
To a 25 mL RBF was added under argon anhydrous DMF (0.9 mL) and cooled to 0 °C. POCl$_3$ (0.180 mL, 2.5 equiv.) was then added dropwise, and the solution stirred for 5 min. To a separate vial was added the respective indole (0.150 g, 0.77 mmol) and DMF (0.3 mL). The indole solution was added to the POCl$_3$ solution dropwise *via* syringe with vigorous stirring. After addition the reaction mixture was allowed to warm to RT and stirred for 2 h. Afterwards ice was added and KOH$_{aq}$ solution (2 M, 10 mL) and the mixture heated to 80 °C for 30 min. The cooled mixture was extracted with EtOAc (3x 20 mL), organic layers were collected and dried (MgSO$_4$) and solvent removed *in vacuo.* Unless otherwise stated the materials were taken through to the next step without further purification. Crude ^{1}H NMR was used to confirm complete reaction. Followed literature procedure.[273]

5-phenyl-1*H*-indole-3-carboxaldehyde[284]
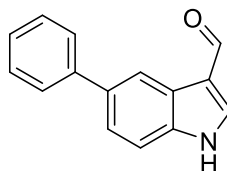
Followed general procedure 6.1. 5-phenyl-1*H*-indole (150 mg, 0.774 mmol). Purified by passing through a plug of silica, eluting with EtOAc to provide 5-phenyl-1*H*-indole-3-carboxaldehyde as a white solid.

^{1}H NMR (400 MHz, C*D*$_3$CN) *δ* 10.02 (s, 1H), 8.41 (dd, *J* = 1.7, 0.9 Hz, 1H), 8.04 (br s, 1H), 7.71 – 7.68 (m, 2H), 7.62 (dd, *J* = 8.5, 0.9 Hz, 1H), 7.59 (dd, *J* = 8.5, 1.7 Hz, 1H), 7.50 – 7.44 (m, 2H), 7.38 – 7.33 (m, 1H).

Data in-line with literature.

6-phenyl-1*H*-indole-3-carboxaldehyde
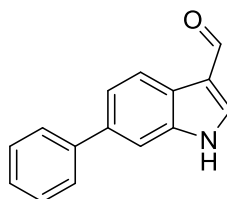
Followed general procedure 6.1. 6-phenyl-1*H*-indole (150 mg, 0.77 mmol). Provided 6-phenyl-1*H*-indole-3-carboxaldehyde as an off-white solid.

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 10.09 (s, 1H), 8.73 (br s, 1H), 8.37 (dt, *J* = 8.3, 0.8 Hz, 1H), 7.88 (d, *J* = 3.0 Hz, 1H), 7.68 – 7.61 (m, 3H), 7.59 (dd, *J* = 8.3, 1.6 Hz, 1H), 7.50 – 7.43 (m, 2H), 7.39 – 7.33 (m, 1H).

7-phenyl-1*H*-indole-3-carboxaldehyde
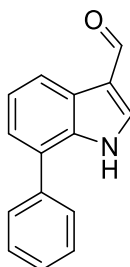
Followed general procedure 6.1. 7-phenyl-1*H*-indole (131.6 mg, 0.68 mmol). Provided a colourless oil of 7-phenyl-1*H*-indole-3-carboxaldehyde that slowly solidified into an off-white solid.

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 10.09 (s, 1H), 9.15 (br s, 1H), 8.32 (dd, *J* = 8.0, 1.5 Hz, 1H), 7.86 (d, *J* = 3.1 Hz, 1H), 7.62 – 7.56 (m, 2H), 7.56 – 7.48 (m, 2H), 7.48 – 7.30 (m, 3H).

General procedure 6.2: Reduction of arylindole-3-carboxaldehydes
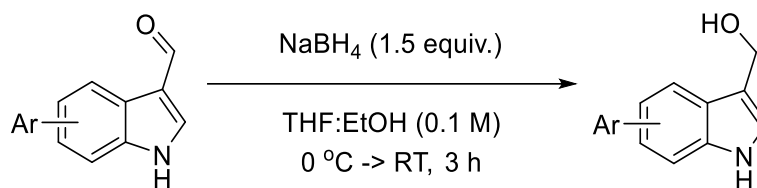
To a 25 mL RBF was added the (crude) arylindole-3-carboxaldehyde, anhydrous THF and EtOH (1:1, 0.1 M). To the stirred solution at 0 °C was added in one batch NaBH₄ (1.5 equiv.), the mixture was stirred for 5 minutes at this temperature then allowed to warm to RT. Left to stir for 3 h. Upon complete reaction (TLC) the reaction was quenched with water (10 mL) and transferred to a separatory funnel. Brine (5 mL) was added, and the mixture extracted with EtOAc (3x 20 mL). The organic layers were collected and washed with brine (20 mL), dried (MgSO₄) and solvent removed *in vacuo.* Further purification details are mentioned under the relevant molecule.

4-phenyl-1*H*-indole-3-carbinol
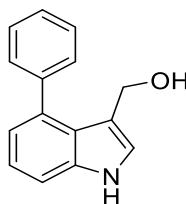
Followed general procedure 6.2. 4-phenyl-1*H*-indole-3-carboxaldehyde (83 mg, 0.373 mmol), THF:EtOH (0.1 M, 4 mL), NaBH₄ (21.3 mg, 1.5 equiv.). Purified by column chromatography (6:4 Hexane : EtOAc) to provide 4-phenyl-1*H*-indole-3-carbinol as an orange semisolid (42.8 mg, 0.192 mmol, 52%).

¹H NMR (400 MHz, DMSO-*d₆*) *δ* 11.07 (br s, 1H), 7.47 – 7.33 (m, 6H), 7.26 (d, *J* = 2.4 Hz, 1H), 7.13 – 7.07 (m, 1H), 6.80 (d, *J* = 7.1 Hz, 1H), 4.40 (br t, *J* = 5.1 Hz, 1H), 4.10 (d, *J* = 5.1 Hz, 2H).

¹³C NMR (101 MHz, DMSO-*d₆*) *δ* 142.0, 136.9, 134.6, 129.0, 127.7, 126.8, 124.0, 123.3, 120.7, 120.0, 116.6, 110.7, 57.1.

IR (cm⁻¹): 3550 (OH), 3402 (NH).

HRMS: ESI+ Calc. for C₁₅H₁₂N (-OH): 208.1035 found: 208.1025.

5-phenyl-1*H*-indole-3-carbinol
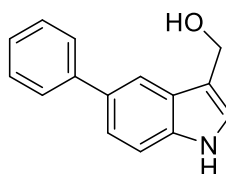
Followed general procedure 6.2. 5-phenyl-1*H*-indole-3-carboxaldehyde (171.7 mg, 0.77 mmol), THF:EtOH (0.1 M, 7.8 mL), NaBH$_4$ (44 mg, 1.5 equiv.). Provided 5-phenyl-1*H*-indole-3-carbinol as an off-white solid (91.8 mg, 0.41 mmol, 53% over two steps).

^{1}H NMR (400 MHz, DMSO-*d$_6$*) *δ* 10.92 (br s, 1H), 7.88 (s, 1H), 7.67 (d, *J* = 7.7 Hz, 2H), 7.48 – 7.35 (m, 4H), 7.33 – 7.24 (m, 2H), 4.78 (br t, *J* = 5.5 Hz, 1H), 4.68 (d, *J* = 5.5 Hz, 2H).

^{13}C NMR (101 MHz, DMSO-*d$_6$*) *δ* 142.0, 136.0, 130.9, 128.7, 127.2, 126.6, 126.1, 124.1, 120.4, 117.1, 116.5, 111.7, 55.4.

IR (cm^{-1}): 3534 (OH), 3265 (NH).

M.P. 99.9 – 102.4 °C.

HRMS, ESI+ Calc. for C$_{15}$H$_{12}$N (-OH) = 207.1002, found: 207.1001.

6-phenyl-1*H*-indole-3-carbinol
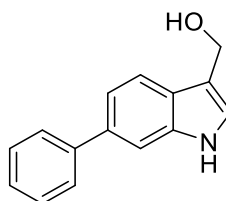
Followed general procedure 6.2. 6-phenyl-1*H*-indole-3-carboxaldehyde (171.7 mg, 0.77 mmol), THF:EtOH (0.1 M, 7.8 mL), NaBH$_4$ (44 mg, 1.5 equiv.). Provided 6-phenyl-1*H*-indole-3-carbinol an off-white solid (144 mg, 0.64 mmol, 84% over two steps).

^{1}H NMR (400 MHz, DMSO-d_6) δ 10.95 (br s, 1H), 7.70 – 7.63 (m, 3H), 7.59 (d, J = 1.6 Hz, 1H), 7.53 – 7.40 (m, 2H), 7.38 – 7.25 (m, 3H), 4.75 (br t, J = 5.4 Hz, 1H), 4.65 (d, J = 5.4 Hz, 2H).

^{13}C NMR (101 MHz, DMSO-d_6) δ 141.7, 137.0, 133.5, 128.8, 126.7, 126.4, 126.1, 124.2, 119.4, 117.8, 109.3, 55.4 (12 out of 13 carbon resonances found)

IR (cm^{-1}): 3377 (OH), 3290 (NH).

M.P. 200 °C+ (dec.)

MS ES+ m/z Calc. for C$_{15}$H$_{12}$N (-OH): 208.1035, found: 208.1027.

7-phenyl-1*H*-indole-3-carbinol
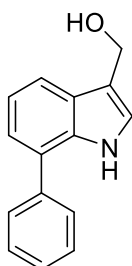
Followed general procedure 6.2. 7-phenyl-1*H*-indole-3-carboxaldehyde (150.6 mg, 0.68 mmol), THF:EtOH (0.1 M, 6.8 mL), NaBH₄ (38.6 mg, 1.5 equiv.). The crude was redissolved in EtOAc (30 mL) and washed with 30% brine in water solution (3x 10 mL). Provided 7-phenyl-1*H*-indole-3-carbinol as an off-white gum (169 mg, 0.76 mmol, 98% over two steps).

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.76 (br s, 1H), 7.66 – 7.59 (m, 3H), 7.52 (t, J = 7.6 Hz, 2H), 7.41 (t, J = 7.4 Hz, 1H), 7.21 (d, J = 2.5 Hz, 1H), 7.15 – 7.07 (m, 2H), 4.76 (br t, J = 5.3 Hz, 1H), 4.67 (d, J = 5.3 Hz, 2H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 138.9, 133.5, 128.9, 128.2, 127.6, 127.1, 125.2, 124.1, 121.2, 119.1, 118.4, 116.5, 55.4.

IR (cm^{-1}): 3419 (OH), 3305 (NH).

MS ES+ m/z Calc. for $C_{15}H_{12}N$ (-OH): 207.1002, found: 207.0997.

N-(3,4-dimethoxyphenyl)indole-3-carboxaldehyde[126]
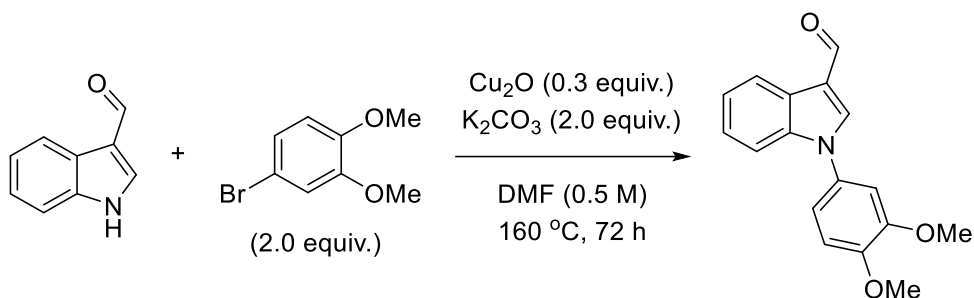
To an oven-dried vial was added indole-3-carboxaldehyde (200 mg, 1.37 mmol), K$_2$CO$_3$ (378.6 mg, 2.0 equiv.) and Cu$_2$O (59 mg, 0.3 equiv.). This was sealed and cycled on a Schlenk line thrice (vacuum/argon). Dry DMF (2.74 mL, 0.5 M) and 3,4-dimethoxybromobenzene (0.422 mL, 2.0 equiv.) were then added *via* syringe. The vial was then sealed with a crimp cap and heated 160 °C for 72 h. The reaction was allowed to cool to RT, filtered through a plug of Celite®, washing with EtOAc (30 mL). The solvent was then removed from the filtrate, the residue taken up in EtOAc (20 mL) and washed successively with 2.5% NH$_4$OH solution, 1 M HCl solution, and brine (20 mL each). The organic layer was dried (MgSO$_4$), and solvent removed *in vacuo.* Residue purified by sequential column chromatography first eluting with 3:7 EtOAc : Hex. then with DCM to provide *N*-(3,4-dimethoxyphenyl)indole-3-carboxyaldehyde as an off-white solid (167 mg, 0.59 mmol, 41%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 10.11 (s, 1H), 8.41 − 8.34 (m, 1H), 7.89 (s, 1H), 7.45 − 7.42 (m, 1H), 7.35 (ddd, *J* = 7.4, 6.5, 1.5 Hz, 2H), 7.07 (dd, *J* = 8.5, 2.4 Hz, 1H), 7.04 − 7.00 (m, 2H), 3.98 (s, 3H), 3.93 (s, 3H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 184.9, 149.8, 149.1, 138.4, 138.0, 131.1, 125.4, 124.5, 123.4, 122.2, 119.3, 117.4, 111.6, 111.0, 108.8, 56.2 (both methoxy signals are at 56.2 *via* HSQC).

IR (cm^{-1}): 1647 (C=O).

Data in line to literature data.[126]

N-(3,4-dimethoxyphenyl)indole-3-carbinol
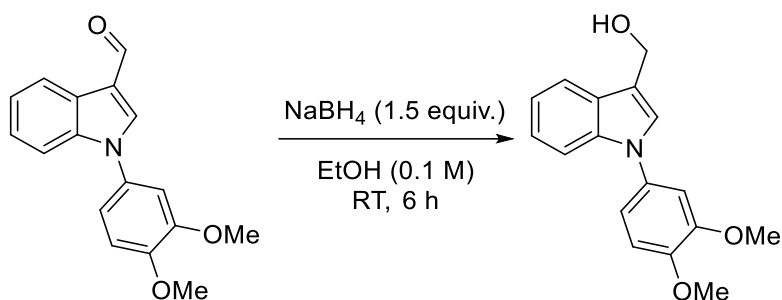
To a 25 mL RBF was added N-(3,4-dimethoxyphenyl)indole-3-carboxaldhyde (100 mg, 0.338 mmol) and ethanol (3.4 mL, 0.1 M). To this stirred suspension was added sodium borohydride (20 mg, 1.5 equiv.). Allowed to stir for 6 h and monitored by TLC. After 6 h water (10 mL) was added, and the mixture extracted with EtOAc (a small amount of brine was added to induce phase separation, 3x 20 mL). The organic layers were collected and dried (MgSO$_4$) and solvent removed under *vacuo.* The residue was purified by column chromatography (9:1 CHCl$_3$: EtOAc) to provide N-(3,4-dimethoxyphenyl)indole-3-carbinol as a white solid (43.5 mg, 0.15 mmol, 43%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 7.82 – 7.77 (m, 1H), 7.50 – 7.45 (m, 1H), 7.32 (d, $J = 0.7$ Hz, 1H), 7.26 – 7.18 (m, 2H), 7.05 – 7.01 (m, 1H), 7.00 (s, 1H), 6.98 (d, $J = 6.0$ Hz, 1H), 4.96 (dd, $J = 5.5, 0.7$ Hz, 2H), 3.96 (s, 3H), 3.91 (s, 3H), 1.50 (t, $J = 5.5$ Hz, 1H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 149.8, 148.1, 137.1, 132.8, 127.6, 127.2, 122.9, 120.5, 119.5, 117.0, 116.7, 111.7, 110.8, 108.9, 57.4, 56.34, 56.26.

IR (cm^{-1}): 3276 (OH). Data in line with literature.[126]

3,4-dimethoxybenzaldehyde
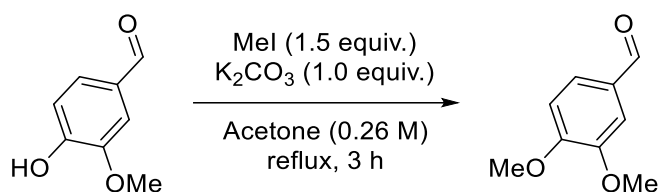
To a 500 mL RBF was added Vanillin (5.0 g, 32.86 mmol) and acetone (126 mL, 0.26 M). To the stirred solution was added potassium carbonate (4.54 g, 1.0 equiv.) and iodomethane (3.07 mL, 1.5 equiv.) *via* syringe. The mixture was heated to reflux for 3 h and monitored by TLC. After complete reaction, the mixture was allowed to cool to RT and the solvent removed. Residue taken up in EtOAc and 2.0 M NaOH soln. (50 mL each). The organic layer was separated and washed with water (2x 20 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* to provide 3,4-dimethoxybenzaldehyde as a white solid (5.14 g, 30.93 mmol, 94%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 9.85 (s, 1H), 7.45 (dd, J = 8.2, 1.9 Hz, 1H), 7.40 (d, J = 1.9 Hz, 1H), 6.97 (d, J = 8.2 Hz, 1H), 3.96 (s, 3H), 3.94 (s, 3H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 191.0, 154.6, 149.8, 130.3, 127.0, 110.5, 109.1, 56.3, 56.1.

M.P.:43 – 45 °C (lit.),[285] 42.3 – 43.7 °C.

Adapted from literature procedure, data matches literature data.[275]

3,4-dimethoxybenzyl alcohol
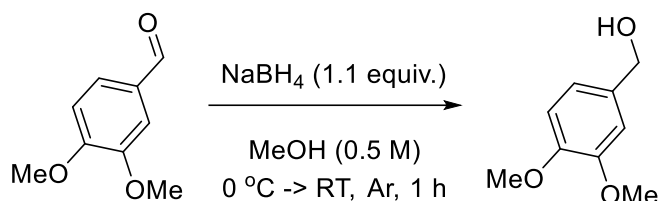
To a 250 mL RBF was added 3,4-dimethoxybenzaldehyde (5.0 g, 30.1 mmol) and methanol (60.2 mL, 0.5 M). The solution was placed under argon and cooled to 0 °C. Crushed NaBH$_4$ (1.25 g, 1.1 equiv.) was then added portion-wise over the space of 5 min. Stirring was continued at 0 °C for 10 min. The mixture was warmed to RT and allowed to stir for 1 h. After TLC indicated complete reaction, water (30 mL) was added, and methanol removed under reduced pressure. The mixture was extracted with EtOAc (3x 350 mL), organic layers collected and dried (MgSO$_4$) and solvent removed *in vacuo.* to provide 3,4-dimethoxybenzyl alcohol as a colourless oil (5.13 g, quant.).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 6.93 (d, J = 1.5 Hz, 1H), 6.89 (dd, J = 8.1, 1.5 Hz, 1H), 6.84 (d, J = 8.1 Hz, 1H), 4.62 (d, J = 5.6 Hz, 2H), 3.89 (s, 3H), 3.88 (s, 3H), 1.67 (br t, J = 5.6 Hz, 1H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 149.3, 148.8, 133.7, 119.5, 111.2, 110.6, 65.5, 56.1, 56.0.

IR (cm^{-1}): 3479 (OH).

Adapted from literature procedure, data is in line with literature.[276]

3,4-dimethoxybenzyl bromide
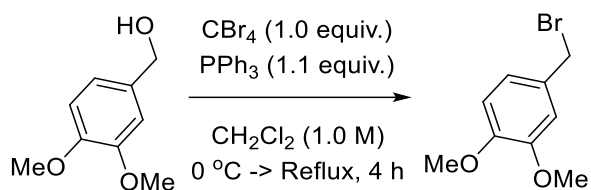
To a 100 mL RBF was added 3,4-dimethoxybenzyl alcohol (5.1 g, 30.3 mmol), carbon tetrabromide (10.75 g, 1.0 equiv.) and DCM (30 mL, 1.0 M). To this cooled (0 °C) solution was added batchwise over 30 min. triphenylphosphine (8.74 g, 1.1 equiv.). This was allowed to warm to RT and then heated to reflux for 3.5 hours. After TLC indicated complete consumption of SM, the mixture was cooled to RT and DCM removed under reduced pressure. To the residue was added Et_2O (300 mL), the mixture shaken and then filtered to removed most of the triphenylphosphine oxide. The filtrate was concentrated and passed through a plug of silica (7:3 Hex : EtOAc) to provide 3,4-dimethoxybenzyl bromide as a white solid (6.43 g, 27.82 mmol, 97%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 6.95 (dd, J = 8.2, 2.1 Hz, 1H), 6.91 (d, J = 2.1 Hz, 1H), 6.81 (d, J = 8.2 Hz, 1H), 4.50 (s, 2H), 3.90 (s, 3H), 3.88 (s, 3H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 149.4, 149.3, 130.4, 121.7, 112.2, 111.2, 56.08, 56.05, 34.5.

Adapted from literature procedure, data is in line with literature.[277]

N-benzylindole-3-carboxaldehyde
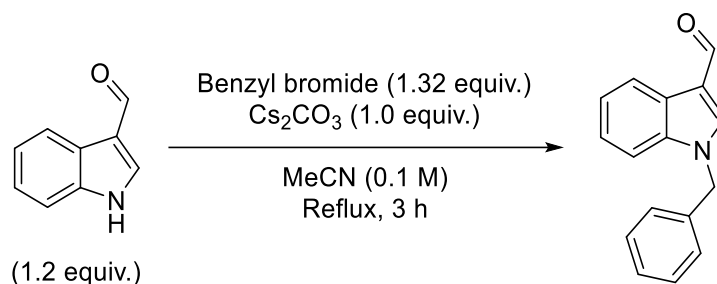
To a 50 mL RBF was added indole-3-carboxaldehyde (200 mg, 1.38 mmol), MeCN (13.8 mL, 0.1 M) and Cs_2CO_3 (372.6 mg, 1.0 equiv.). Heated to reflux for 2 h. Afterwards, benzyl bromide (0.18 mL, 1.32 equiv.) was added *via* syringe and the mixture continued to reflux for 1 h. After completion of the reaction by TLC the mixture was allowed to cool to RT and solvent was removed under reduced pressure. The residue was taken up in water (50 mL) and extracted with EtOAc (3x 50 mL). The organic layers were collected and dried ($MgSO_4$), solvent was removed *in vacuo.* The residue was purified by column chromatography (8:2 Hexane : EtOAc) to provide *N*-benzylindole-3-carboxyaldehyde as a white solid (312 mg, 1.33 mmol, 96%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 10.01 (s, 1H), 8.36 – 8.31 (m, 1H), 7.72 (s, 1H), 7.37 – 7.29 (m, 6H), 7.21 – 7.17 (m, 2H), 5.37 (s, 2H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 184.6, 138.4, 137.5, 135.3, 129.2, 128.4, 127.2, 125.5, 124.2, 123.1, 122.2, 118.6, 110.4, 51.0.

IR (cm^{-1}): 2815 (C-H aldehyde), 1650 (C=O).

Followed literature procedure, data matches literature.[278]

N-benzylindole-3-carbinol
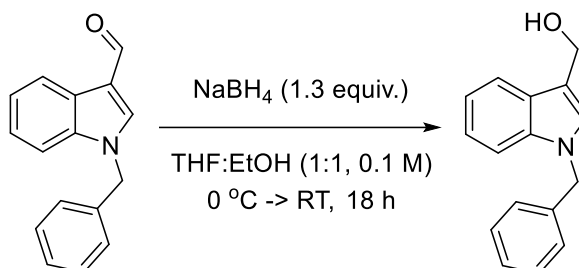
To a 25 mL RBF was added *N*-benzylindole-3-carboxaldehyde (300 mg, 1.275 mmol), anhydrous THF (6.4 mL) and EtOH (6.4 mL) to reach a total concentration of 0.1 M. NaBH$_4$ (62.7 mg, 1.3 equiv.) was added in one batch at 0 °C and the mixture allowed to warm to RT and stirred for 18 h. Afterwards, water (20 mL) was added to the suspension and the solvents were removed *in vacuo.* The aqueous suspension was then transferred to a separatory funnel and extracted with EtOAc (3x 20 mL). The organic layers were collected and washed with Brine (20 mL), dried (MgSO$_4$) and solvent removed under reduced pressure to provide a white solid. Purified by column chromatography (8:2 -> 7:3 Hex. : EtOAc) to provide *N*-benzylindole-3-carbinol as a white solid (197 mg, 0.83 mmol, 65%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 7.75 (dt, *J* = 7.5, 1.0 Hz, 1H), 7.33 – 7.26 (m, 4H), 7.21 (ddd, *J* = 8.2, 7.0, 1.4 Hz, 1H), 7.19 – 7.15 (m, 1H), 7.15 – 7.11 (m, 3H), 5.30 (s, 2H), 4.89 (d, *J* = 5.5 Hz, 2H), 1.42 (d, *J* = 5.5 Hz, 1H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 137.4, 137.0, 129.0, 127.9, 127.4, 127.14, 127.08, 122.4, 119.9, 119.4, 115.6, 110.0, 57.4, 50.2.

IR (cm^{-1}): 3367 (OH).

Data matches that of the literature.[286]

N-(3,4-dimethoxybenzyl)indole-3-carboxaldehyde
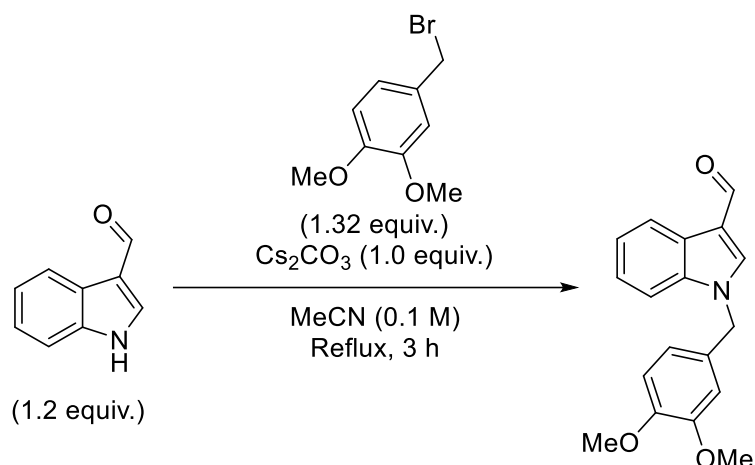
To a 50 mL RBF was added 1*H*-indole-3-carboxaldehyde (200 mg, 1.37 mmol, 1.2 equiv.), Cs_2CO_3 (327.6 mg, 1.0 equiv.) and MeCN (13.8 mL, 0.1 M). This was heated to reflux for 2 h. Afterwards, 3,4-dimethoxybenzylbromide (356.9 mg, 1.32 equiv.) was added and the mixture again heated to reflux for 1 h. After allowing to cool to RT and solvent removed *in vacuo.* Residue taken up in EtOAc and water (50 mL each), separated and the aqueous layer extracted with EtOAc (2x 50 mL). The organic layers were combined and dried ($MgSO_4$), solvent removed under reduced pressure. The residue was purified by column chromatography (100% $CHCl_3$ -> 95:5 $CHCl_3$: EtOAc) to provide *N*-(3,4-dimethoxybenzyl)indole-3-carboxaldehyde as a white solid (209 mg, 0.70 mmol, 52%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 10.00 (s, 1H), 8.37 – 8.29 (m, 1H), 7.68 (s, 1H), 7.41 – 7.30 (m, 3H), 6.85 (d, *J* = 8.2 Hz, 1H), 6.77 (dd, *J* = 8.2, 2.1 Hz, 1H), 6.72 (d, *J* = 2.1 Hz, 1H), 5.29 (s, 2H), 3.88 (s, 3H), 3.81 (s, 3H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 184.7, 149.7, 149.4, 138.3, 137.7, 127.6, 125.7, 124.3, 123.2, 122.3, 120.3, 118.6, 111.6, 110.7, 110.5, 56.1, 50.9.

M.P. 250 °C (dec.)

IR (cm^{-1}): 2887 (C-H aldehyde).

Adapted from literature procedure.[278]

N-(3,4-dimethoxybenzyl)indole-3-carbinol
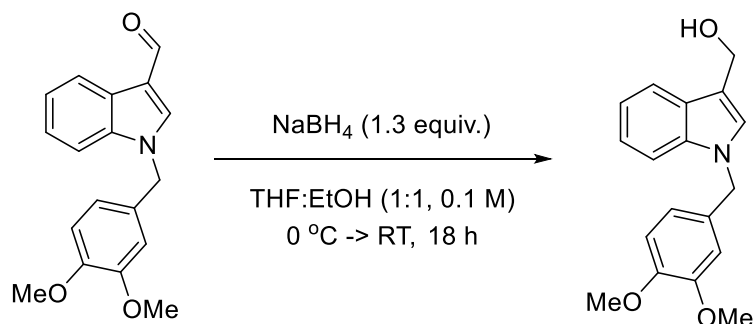
To a 50 mL RBF was added *N*-(3,4-dimethoxybenzyl)indole-3-carboxaldehyde (180 mg, 0.609 mmol) and a THF:EtOH mixture (1:1, 6.1 mL, 0.1 M). To the stirred, cooled (0 °C) solution was added NaBH$_4$ (30 mg, 1.3 equiv.). The mixture was then allowed to warm to RT and stirred for 18 h. After TLC indicated complete reaction (PMA stain), water (5 mL) was added to the mixture and the THF:EtOH solvent was removed under reduced pressure. To the suspension was added EtOAc (20 mL) and the layers separated. The aqueous layer was extracted with EtOAc (2x 20 mL). The organic layers were combined and washed with brine (20 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* to provide a colourless oil. Purified by sequential column chromatography (1% MeOH in CHCl$_3$, then CHCl$_3$ -> 9:1 CHCl$_3$:EtOAc) to provide *N*-(3,4-dimethoxybenzyl)indole-3-carbinol as a colourless oil (148 mg, 0.50 mmol, 82%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 7.75 (dt, *J* = 7.7, 0.9 Hz, 1H), 7.33 (dt, *J* = 8.2, 1.0 Hz, 1H), 7.22 (ddd, *J* = 8.2, 7.0, 1.3 Hz, 1H), 7.16 (ddd, *J* = 8.0, 7.0, 1.1 Hz, 1H), 7.11 (s, 1H), 6.81 – 6.78 (m, 1H), 6.72 – 6.68 (m, 2H), 5.22 (s, 2H), 4.88 (s, 2H), 3.85 (s, 3H), 3.80 (s, 3H), 1.42 (br s, 1H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 149.4, 148.8, 146.0, 137.1, 129.7, 127.4, 127.0, 122.3, 119.9, 119.7, 119.4, 115.5, 111.4, 110.5, 110.0, 57.4, 56.1, 50.0.

IR (cm^{-1}): 2999, 2932, 2834.

MS ES+ Calc. for C$_{18}$H$_{19}$NO$_3$ (M+H)$^+$: 297.1365, found: 297.1362.

General procedure 6.3: Sulfonamide formation[287]
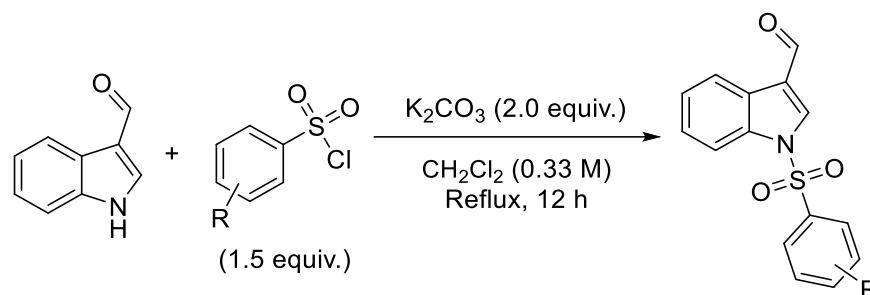
To a 10 mL RBF was added indole-3-carboxaldehyde (200 mg, 1.37 mmol), K$_2$CO$_3$ (379 mg, 2.0 equiv.) and DCM (4.15 mL, 0.33 M). To this was added the respective sulfonyl chloride (1.5 equiv.). The reaction mixture was heated to reflux for 18 h. Afterwards water was added (10 mL) and the mixture transferred to a separatory funnel and extracted with chloroform (3x 20 mL). The organic layers were collected and washed with brine (20 mL), dried (MgSO$_4$) and solvent removed *in vacuo.* The residues were purified by column chromatography.

General procedure 6.4: Phase transfer catalysed sulfonamide formation
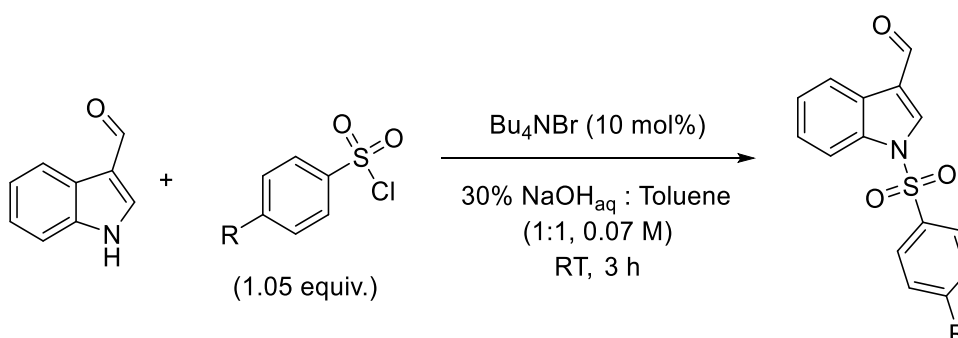
To an RBF was added indole-3-carboxaldehyde and toluene (0.14 M). To the stirred suspension was added 30% NaOH solution (0.14 M) and Bu$_4$NBr (10 mol%). Afterwards, the respective sulfonyl chloride (1.05 equiv.) was added, and the mixture vigorously stirred for 3 h. The mixture was transferred to a sep. funnel and separated, the aqueous phase washed with toluene (20 mL). The organic layers were collected and dried (MgSO$_4$) and solvent removed. Purified by column chromatography to provide the respective *N-*(arylsuflonyl)jndole-3-carboxaldehyde. Adapted from literature conditions.[280]

N-(benzenesulfonyl)indole-3-carboxaldehyde
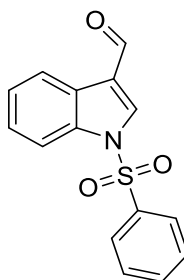
Followed general procedure 6.3. Benzenesulfonyl chloride (0.26 mL, 1.5 equiv.). Elution with 8:2 Hexane : EtOAc provided *N*-(benzenesulfonyl)indole-3-carboxaldehyde as a white solid (91.9 mg, 0.32 mmol, 24%).

^{1}H NMR (400 MHz, Acetone-d_6) *δ* 10.16 (s, 1H), 8.70 (s, 1H), 8.21 (dt, *J* = 7.8, 1.1 Hz, 1H), 8.17 – 8.13 (m, 2H), 8.04 (dt, *J* = 8.4, 1.0 Hz, 1H), 7.80 – 7.71 (m, 1H), 7.66 (ddd, *J* = 8.0, 6.4, 1.3 Hz, 2H), 7.47 (ddd, *J* = 8.5, 7.3, 1.4 Hz, 1H), 7.40 (td, *J* = 7.6, 1.1 Hz, 1H).

^{13}C NMR (101 MHz, Acetone-d_6) *δ* 186.7, 138.4, 138.2, 136.0, 136.0, 130.9, 128.2, 127.3, 127.1, 125.9, 123.4, 123.1, 114.3.

IR (cm^{-1}): 3129, 2846 (C-H aldehyde), 1676 (C=O).

Data in line with literature data.[288]

N-(4-trifluoromethylbenzenesuflonyl)indole-3-carboxaldehyde
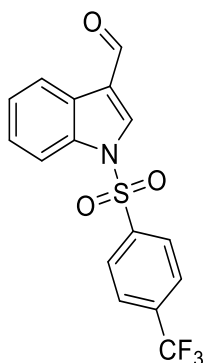
Followed general procedure 6.3. 4-trifluoromethylbenzenesulfonyl chloride (0.502g, 1.5 equiv.). Elution with 85:15 Hexane : EtOAc to provide *N*-(4-trifluorobenzenesulfonyl)indole-3-carboxaldehyde as a white solid (37.2 mg, 8%).

May be alternatively prepared by following general procedure 6.4. Indole-3-carboxaldehyde (282.1 mg, 1.94 mmol), toluene (6.9 mL), 30% NaOH solution (9.2 mL), Bu$_4$NBr (62.5 mg) 4-trifluoromethylbenzenesulfonyl chloride (500 mg, 2.04 mmol). Column chromatography (9:1 Hex. : EtOAc) provided *N*-(4-trifluoromethylbenzenesuflonyl)jndole-3-carboxaldehyde as a white solid (0.260 g, 0.74 mmol, 38%).

^{1}H NMR (400 MHz, Acetone-*d$_6$*) *δ* 10.16 (s, 1H), 8.73 (s, 1H), 8.39 (d, *J* = 8.3 Hz, 2H), 8.22 (dt, *J* = 7.8, 1.0 Hz, 1H), 8.06 (dt, *J* = 8.3, 1.0 Hz, 1H), 8.02 (d, *J* = 8.1 Hz, 2H), 7.49 (ddd, *J* = 8.4, 7.3, 1.4 Hz, 1H), 7.42 (td, *J* = 7.5, 1.1 Hz, 1H).

^{13}C NMR (101 MHz, Acetone-*d$_6$*) *δ* 186.7, 141.8, 138.3, 136.2, 136.0, 129.3, 128.1 (q, *J* = 3.8 Hz), 127.4 (d, *J* = 3.0 Hz), 126.2, 123.8, 123.3, 114.2 (12 out of 13 carbon resonances found).

^{19}F NMR (376 MHz, Acetone-*d$_6$*) *δ* -64.00.

IR (cm^{-1}): 3107, 2923 (C-H aldehyde), 1666 (C=O).

M.P. 124.9 – 126.1 °C.

MS ES+ m/z Calc. for C$_{16}$H$_{10}$F$_3$NO$_3$S (M+H)$^+$: 356.0403, Found: 356.0392.

N-(3-trifluorobenzenesulfonyl)indole-3-carboxaldehyde
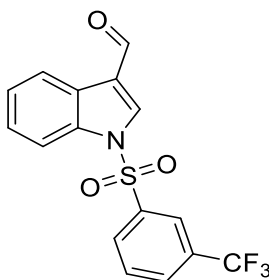
Followed general procedure 6.3. 3-trifluoromethylbenzenesulfonyl chloride (0.33 mL, 1.5 equiv.). Elution with 8:2 Hexane : EtOAc to provide N-(3-trifluorobenzenesulfonyl)indole-3-carboxaldehyde as a white solid (41.1 mg, 0.12 mmol, 9%).

^{1}H NMR (400 MHz, Acetone-d_6) δ 10.16 (s, 1H), 8.77 (s, 1H), 8.50 − 8.43 (m, 2H), 8.22 (dt, J = 7.7, 1.1 Hz, 1H), 8.11 (ddt, J = 9.3, 8.4, 0.9 Hz, 2H), 7.98 − 7.90 (m, 1H), 7.50 (ddd, J = 8.5, 7.3, 1.4 Hz, 1H), 7.43 (ddd, J = 8.3, 7.3, 1.0 Hz, 1H).

^{13}C NMR (101 MHz, Acetone-d_6) δ 186.7, 139.3, 138.4, 136.0, 132.8 − 132.6 (m), 132.5, 132.3, 132.1, 127.4 (d, J = 2.3 Hz), 126.2, 125.1 (d, J = 4.0 Hz), 123.8, 123.3, 122.7, 114.2.

^{19}F NMR (376 MHz, Acetone-d_6) δ -63.48

IR (cm^{-1}): 3143, 1666 (C=O).

M.P. 122.5 − 124.8 °C.

MS ES+ m/z Calc. for $C_{16}H_{10}F_3NO_3S$ (M+H)$^+$: 354.0412, found: 354.0400.

N-(4-methylbenzenesulfonyl)indole-3-carboxaldehyde
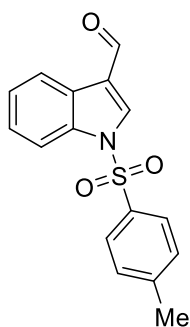
Followed general procedure 6.4. Indole-3-carboxaldehyde (200 mg, 1.37 mmol), toluene (9.84 mL), 30% NaOH soln. (9.84 mL), Bu$_4$NBr (44.4 mg), tosyl chloride (275.8 mg). Purified by column chromatography (9:1 Hex. : EtOAc) to provide *N*-(4-methylbenzenesulfonyl)indole-3-carboxaldehdye as a white solid (293 mg, 0.98 mmol, 75%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 10.10 (s, 1H), 8.25 (ddd, *J* = 7.5, 1.6, 0.8 Hz, 1H), 8.23 (s, 1H), 7.97 – 7.93 (m, 1H), 7.88 – 7.83 (m, 2H), 7.44 – 7.39 (m, 1H), 7.36 (td, *J* = 7.6, 1.2 Hz, 1H), 7.32 – 7.27 (m, 2H), 2.38 (s, 3H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 185.5, 146.3, 136.3, 135.4, 134.5, 130.5, 127.4, 126.5, 125.2, 122.8, 122.5, 113.4, 21.8 (13 out of 14 carbon resonances found).

IR (cm^{-1}): 3133, 2847 (C-H aldehyde), 1661 (C=O).

Data in line with literature data.[126]

General procedure 6.5: *N*-(aryl)indole-3-carboxaldehyde reduction
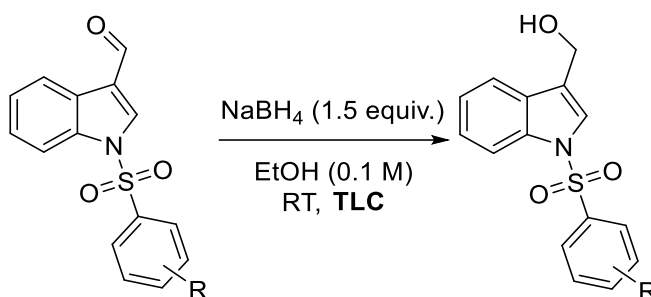
To a vial was added the respective *N*-functionalised indole-3-carboxaldehyde and ethanol (0.1 M). To the stirred solution was added NaBH$_4$ (1.5 equiv.) and allowed to stir at RT, monitored by TLC. H$_2$O (3 mL) was added to quench the reaction after completion, the mixture was transferred to a separatory funnel and brine was added (10 mL), the mixture extracted with EtOAc (3x 20 mL), the organic layers collected and washed with brine (10 mL), dried (MgSO$_4$) and solvent removed under reduced pressure at 27 - 28 °C (if necessary). Purified by column chromatography (7:3 Hexane : EtOAc), followed by trituration with hexane.

N-(benzenesulfonyl)indole-3-carbinol
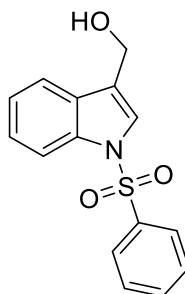
Followed general procedure 6.5. *N*-(benzenesulfonyl)indole-3-carboxaldehyde (70 mg, 0.243 mmol), NaBH$_4$ (14 mg, 1.5 equiv.), EtOH (2.43 mL). Reaction time 2 h. Provided a light-orange solid of *N*-(benzenesulfonyl)indole-3-carbinol (63.3 mg, 0.22 mmol, 90%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) *δ* 7.93 (dt, *J* = 8.3, 0.9 Hz, 1H), 7.85 – 7.79 (m, 2H), 7.54 (dt, *J* = 7.8, 1.0 Hz, 1H), 7.50 – 7.44 (m, 2H), 7.41 – 7.33 (m, 2H), 7.28 (ddd, *J* = 8.4, 7.2, 1.3 Hz, 1H), 7.24 – 7.15 (m, 1H), 4.76 (dd, *J* = 5.5, 1.0 Hz, 2H), 1.53 (br t, *J* = 5.5 Hz, 1H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) *δ* 138.4, 135.6, 134.0, 129.6, 129.5, 127.0, 125.3, 123.9, 123.6, 122.6, 120.1, 113.9, 57.3.

IR (cm^{-1}): 3334 (OH). Data in line with literature data.[288]

_N_-(4-trifluoromethylbenzenesulfonyl)indole-3-carbinol
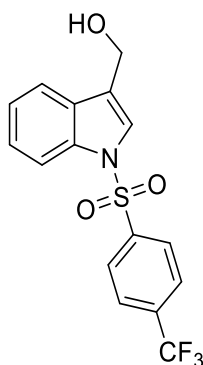
Followed general procedure 6.5. _N_-(4-trifluoromethylbenzenesulfonyl)indole-3-carboxaldehyde (30 mg, 0.085 mmol), NaBH$_4$ (5 mg, 1.5 equiv.)n EtOH (1 mL). Reaction time 2 h. Provided an off-white solid of _N_-(4-trifluoromethylbenzenesulfonyl)indole-3-carbinol (30 mg, quant.).

^{1}H NMR (400 MHz, C*D*Cl$_3$) _δ_ 8.04 – 7.97 (m, 3H), 7.73 – 7.67 (m, 2H), 7.65 – 7.59 (m, 1H), 7.54 (d, _J_ = 1.1 Hz, 1H), 7.38 (ddd, _J_ = 8.5, 7.3, 1.3 Hz, 1H), 7.29 (ddd, _J_ = 8.3, 7.3, 1.1 Hz, 1H), 4.84 (dd, _J_ = 5.7, 1.1 Hz, 2H), 1.62 (t, _J_ = 5.7 Hz, 1H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) _δ_ 141.63 135.6, 135.4, 129.7, 127.5, 126.7 (q, J = 3.9 Hz), 125.6, 124.0, 123.6, 123.5, 120.3, 113.8, 57.3 (13 out of 14 carbon resonances observed).

^{19}F NMR (376 MHz, C*D*Cl$_3$) _δ_ -63.37.

IR (cm^{-1}): 3246 (OH).

M.P. 164.1 – 165.5 $^{\circ}$C.

MS ESI+ Calc. for C$_{16}$H$_{11}$F$_3$NO$_2$S (-OH): 339.0493, found: 339.0493.

_N_-(3-trifluoromethylbenzenesulfonyl)indole-3-carbinol
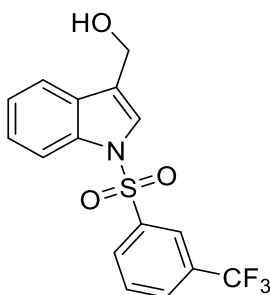
Followed general procedure 6.5. _N_-(3-trifluoromethylbenzenesulfonyl)indole-3-carboxaldehyde (30 mg, 0.085 mmol), NaBH$_4$ (5 mg, 1.5 equiv.), EtOH (1 mL). Reaction time 2 h. Provided an off-white solid of _N_-(3-trifluoromethylbenzenesulfonyl)indole-3-carbinol (29.5 mg, 0.08 mmol, 98%).

^{1}H NMR (400 MHz, C_DCl$_3$) δ 8.18 (tt, J = 1.9, 0.7 Hz, 1H), 8.07 – 8.03 (m, 1H), 8.00 (dt, J = 8.3, 0.9 Hz, 1H), 7.80 (ddt, J = 7.8, 1.8, 1.0 Hz, 1H), 7.65 – 7.56 (m, 2H), 7.54 (t, J = 1.1 Hz, 1H), 7.38 (ddd, J = 8.4, 7.3, 1.3 Hz, 1H), 7.29 (ddd, J = 8.2, 7.2, 1.0 Hz, 1H), 4.84 (dd, J = 5.7, 1.1 Hz, 2H), 1.61 (t, J = 5.7 Hz, 1H).

^{13}C NMR (101 MHz, C_DCl$_3$) δ 139.4, 135.5, 132.4, 130.7 (d, J = 3.6 Hz), 130.4, 130.0, 129.7, 125.6, 124.1 (d, J = 3.9 Hz), 124.0, 123.5, 123.5, 120.3, 113.73, 57.3 (15 out of 16 carbon resonances observed).

^{19}F NMR (376 MHz, C_DCl$_3$) δ -62.96.

IR (cm^{-1}): 3310 (OH).

M.P. 116.8 – 122.3 °C.

MS ES+ Calc. for C$_{16}$H$_{11}$F$_3$NO$_2$S (-OH): 339.0493, found: 339.0489.

_N_-(4-methylbenzenesulfonyl)indole-3-carbinol
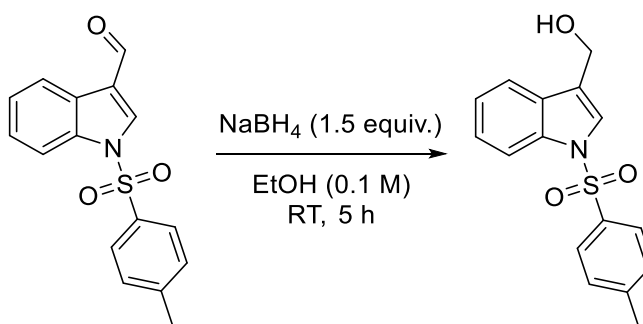
Adapted from general procedure 6.5. _N_-(4-methylbenzenesulfonyl)indole-3-carboxaldehyde (200 mg, 0.668 mmol), EtOH (6.68 mL, 0.1 M), sodium borohydride (38 mg, 1.5 equiv.). 5 h reaction time. Routine workup provided _N_-(4-methylbenzenesulfonyl)indole-3-carbinol as a colourless oil (168 mg, 0.56 mmol, 83%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 7.99 (dt, J = 8.4, 0.9 Hz, 1H), 7.81 – 7.73 (m, 2H), 7.61 (dt, J = 7.8, 1.1 Hz, 1H), 7.55 (d, J = 1.1 Hz, 1H), 7.34 (ddd, J = 8.4, 7.2, 1.3 Hz, 1H), 7.28 – 7.20 (m, 3H), 4.82 (dd, J = 5.7, 1.0 Hz, 2H), 2.34 (s, 3H), 1.59 (br t, J = 5.7 Hz, 1H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 145.2, 135.6, 135.4, 130.1, 129.6, 127.0, 125.2, 124.0, 123.5, 122.4, 120.0, 113.9, 57.4, 21.7.

IR (cm^{-1}): 3384 (OH).

Data in line with literature data.[126]

3,3'-Diindolylmethane
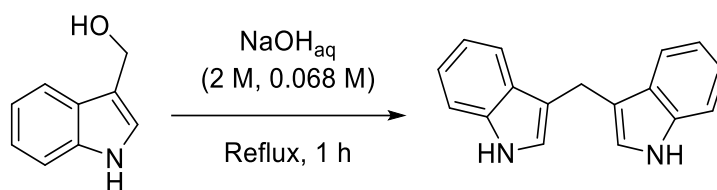
To a 250 mL RBF was added indole-3-carbinol (1.0 g, 6.78 mmol) and 2.0 M NaOH solution (100 mL, 0.068 M). The mixture was heated to reflux for 1 h. The mixture was cooled and neutralised using CO_2 gas (dry ice). The precipitate was filtered and dried, then subjected to column chromatography (8:2 Hex. : EtOAc) to provide 3,3'-diindolylmethane as an off-white solid (341 mg, 1.38 mmol, 20%).

^{1}H NMR (400 MHz, C*D*Cl$_3$) δ 7.90 (br s, 2H), 7.63 (dd, J = 8.0, 1.1 Hz, 2H), 7.36 (dd, J = 8.0, 1.1 Hz, 2H), 7.19 (ddd, J = 8.2, 7.0, 1.2 Hz, 2H), 7.10 (ddd, J = 8.0, 7.0, 1.0 Hz, 2H), 6.97 – 6.92 (m, 2H), 4.25 (s, 2H).

^{13}C NMR (101 MHz, C*D*Cl$_3$) δ 136.6, 127.7, 122.3, 122.1, 119.4, 119.3, 115.9, 111.2, 21.4.

IR (cm^{-1}): 3391 (NH).

Followed literature procedure, data matches that of the literature.[281]

7 <u>Overall Conclusions & Future Work</u>

7.1 <u>Synthetic</u>

All hit compounds targeted in chapter 3 were successfully synthesised, with notably and to the best of my knowledge, new syntheses devised for redoxal and *F*-redoxal and the successful production NSC-637578 *via* a novel one-step route. Also, a literature route was effectively followed and where appropriate optimised for the production of NSC-288387. Finally, full characterisation data is provided for all of these compounds including an x-ray crystal structure for NSC-288387 and redoxal and *F*-redoxal methyl ester precursors.

In chapter 4, the vast majority of target derivatives of NSC-217913 were synthesised, with the common 5,6-dichloroimiazo[4,5-*b*]pyrazine-2(1,3*H*)-thione core produced in an improved fashion over the known literature method (section 4.3.1.1). The synthesis of heterocyclic variations was expediated by the purchasing of commercial heterocycles, and several variations of the 5,6-fused heterocyclic system could also be produced, with further variations produced from the imidazo[1,2-*a*]pyrazine and imidazo[4,5-*d*]pyrimidine core (section 4.3.2).

Within chapter 5, the production of NSC-2805 variants targeting hydrophilic interactions and π-π interactions was largely successful. Production of bioisosteric analogues of an NSC-2805 isoquinoline analogue worked *via* a borylation cross-coupling approach. The production of a library of *F*-redoxal variants was also successful, in part by the implementation of an efficient synthetic plan of cross-coupling reactions followed by saponifications.

Chapter 6 synthetic objectives were successfully met by first producing phenylhydrazone variants and the testing of commercial derivatives varying at the 3-poisition. Synthesis of both phenyl derivatives about the indole skeleton as well as the functionalisation at the indole nitrogen was also successful.

Overall, this represents successfully meeting both the first and part of the second main objectives set out at the beginning of this thesis.

7.2 <u>Computational</u>

The implementation of molecular docking into the workflow for analogue synthesis as part of the second main objective was successful. Analogues which were predicted to have improved binding affinity over their respective hit compound were targeted for synthesis. The choice of target derivatives was based upon the visual inspection parameters set out

in section 3.5.3. The docking protocol developed was successfully conducted per the objectives set out in sections 3.5.4 and 3.5.5 for the respective hit compounds, and these were largely easy to run and the results easy to interpret.

Unfortunately, definite conclusions as to the success of the methods used within molecular docking are difficult to arrive at, due to the general lack of biological data. This is because no clear comparisons can be made between the analogues predicted to have improved binding and their experimental data. So, it cannot be said whether the use of visual inspection or the parameters used, methodology validation or protocol used for molecular docking were able to accurately predict compounds with improved binding affinity. However, the implementation of molecular docking has been successful, and this represents meeting part of the second main objective of this thesis.

7.3 <u>Biological</u>

The biological results obtained from NSC-217913 series was disappointing as most compounds did not provide IC_{50} values. The thione precursor to the *S*-alkylation library was found to be 10-fold more active than the original hit IC_{50}, and this derivative was pursued further. Findings from these investigations showed that only the cyclothiourea was able to provide these IC_{50} values and any variation away from it resulted in no activity. Later cross-reactivity testing found these compounds to be reactive in some way with the E1 enzyme within the bioassay. This, in the absence of other data showing these compounds binding or interacting with WWP2, means these compounds have to be treated as PAINs as they are interacting in unproductive ways with other parts of the bioassay and providing false positives.

Another problem encountered within chapter 4 was the original NCI hit providing an IC_{50} but not the resynthesised material. This led to a purity analysis, finding that again the NCI sample was impure. Based on HRMS data, it is predicted that either one or both doubly alkylated NSC-217913 (see section 4.7.2) is present in the original sample and may have been contributing to the activity observed in the HTS, again further investigation would be required here.

For the library synthesised for NSC-2805 optimisation, it is now understood that the *para*-hydroquinone motif is reacting with assay components, but based off co-crystal data is also binding to WWP2. The hydroquinone motif represents a common PAINs moiety, and the IC_{50} values gained should be treated with caution as it cannot be elucidated how much actual inhibitory effect these compounds have *vs.* the effect of the hydroquinone groups

reactivity has on the outcome of the bioassays. It is well known that PAINs hits are essentially unoptimisable within modern hit-to-lead programmes and although it was postulated that these compounds could be further optimised within this research, derivatives of NSC-2805 which do not have the *para*-hydroquinone motif have much less apparent activity and do not represent compounds of interest that would be pursued further in a rational design sense. So, even though NSC-2805 may bind in some way to WWP2, it should not be investigated further as the likelihood of producing compounds of similar structure and IC_{50} values without the reactivity of the *para*-hydroquinone moiety is highly unlikely.

For the *F*-redoxal series many compounds were found to be inactive, this time because of their insolubility. This was initially strange considering that the redoxal hit was able to provide an IC_{50} value, but upon subjecting the resynthesised material made in section 3.3.4 to bioassay and finding no IC_{50} was also most-likely as a result of insolubility. A purity analysis of the original NCI sample of redoxal resulted, finding that several impurities were present based upon NMR, visual inspection of a DMSO sample, and mass spectral analysis with the conclusion being that likely the hit stemmed from these impurities and not redoxal itself.

The I3C series studied showed little activity against WWP2, but some compounds were found to be marginally active within the bioassay. An issue with these compounds however is their observed insolubility. Additionally, it seems that when starting from a binder without known inhibitory activity, it is difficult to optimise it into the range where inhibitory activity is quantified by the methods used.

7.4 **Final Discussion**

With the overall aim of this project to develop a potent inhibitor of WWP2 in mind, this has not been reached. Even though both the first and second main objectives have been successfully met, no derivative of any hit has been shown to provide higher activity compared to their parent hit compound, and this represents not meeting the last objective set out in section 2.4. For each hit compound pursued into analogue synthesis, several observations may be drawn:

- NCI sample which provided hits were impure.
- Compounds have been pursued which contain known PAINs motifs.
- Motifs have been identified that interact non-productively with components of the bioassay.

- The vast majority of analogues synthesised have not provided IC$_{50}$ data.

Within the framework of PAINs, molecules such as Redoxal (484.51 gmol^{-1}, Log$P$: 5.95) which display as hitters but otherwise seem unoptimisable are best treated with caution.[226–232,289] No attempts within this literature have been made to optimise its structure by chemical methods. In this case, redoxal was inadvertently treated as unoptimisable, as the implemented strategy was to essentially break it into various fragments from which to build a more unique structure from. It is unfortunate that only after the library synthesis was purity analysis undertaken, and the same can be said of NSC-217913, where initially a purity analysis was not considered until the same discrepancy arose between the resynthesised material and the NCI sample. It seems that, based upon these findings, purity analysis of the HTS samples is required prior to taking hits forward, at least with NCI libraries. Additionally, literature research is also required so that potentially frequent hitters like redoxal can be identified as hitters with difficult to optimise structures, again prior to the decision to take them forward for analogue synthesis.

Having said this, from the library synthesis of redoxal, compound **7.1** was identified which displays activity in both DSF and biochemical assays and may offer a different starting point from which to further investigate inhibition of WWP2 (Figure 7-1). The same may be said of **7.2** & **7.3** from the NSC-2805 library, although only identified from the biochemical assay and served to display the structure-interference relationship of the *para*-hydroquinone. Caution is again required here and further experiments to validate these materials should be sought prior to synthetic investigation, such as implementing biophysical techniques. Additionally, an investigation into the activity of the NSC-217913 NCI sample impurity, whichever regioisomer it is, may offer movement to potentially active novel compounds.

7.1, *368.5 μM*   **7.2**, *131.1 - 400.3 μM*   **7.3**, *327.3 - 458.7 μM*

7.4   **7.5**
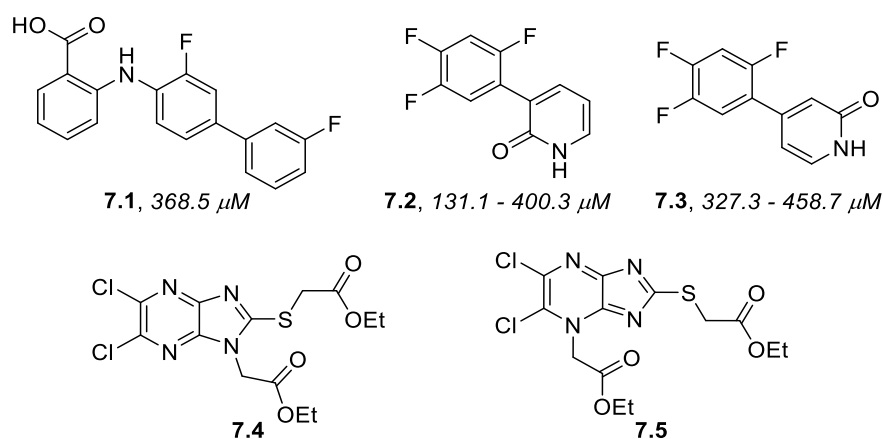
Figure 7-1: Structures identified from this investigation that may warrant further investigation as potential WWP2 inhibitors.

Structural similarities between **7.1** and **7.2** & **7.3** may well offer opportunities for combination (Figure 7-2, a). Also, the bioisosteric replacement of the hydroxyl groups was relatively underexplored, whilst others have utilised cyclourea and indazoles, for example (**7.8**, **7.9**).[203,204] The surveying of different bioisosteres of the *para*-hydroquinone may provide additional starting points for optimisation of inhibitory molecules against WWP2.

a)

7.6   **7.7**

b)

NSC-2805   **7.8**   **7.9**
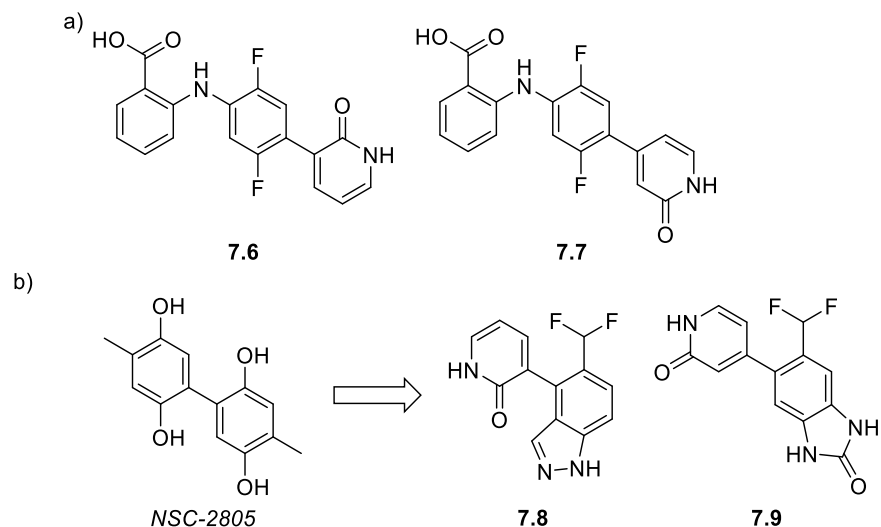
*Figure 7-2: a) Example combinations of **7.1**, **7.2** and **7.3**, b) bioisosteric replacement of hydroxyl groups for NSC-2805.*

A major hurdle in drug discovery is the optimisation of primary hits from screening campaigns to produce lead compounds.[129] From this standpoint, compounds have been identified and pursued which contain known PAINs motifs which interact non-productively with the assay during the course of this work. The issues surrounding how PAINs moieties have been included as legitimate hits during the HTS stage need to be addressed to improve the chances of successfully optimising hit compounds. In order to do this, when

HTS is again conducted improvements should be incorporated into the process immediately after HTS and later hit validation stages. Additionally, utilising a library for HTS that contains less nuisance compounds would be very desirable, for example one commonly used by the pharmaceutical industry.

For the vast majority of hit compounds identified within this project it can be argued that they may be PAINs.[252,253,289,292,293] It is therefore important to rule out any potential reactivity or interference of these compounds within the HTS and bioassay. Knowledge and experimental-based methods for ruling out PAINs compounds are well known and these may be incorporated following hit identification into a hit validation stage, including counter-screens against E1 and E2 enzymes, prior to analogue synthesis.[252] A fuller understanding of the chemical basis of PAINs motifs at this stage would result in a lower likelihood of selecting PAINs compounds which would be unproductive to attempt to optimise.

It is understood that due to the nature of the biochemical assay, inactivation of E1 or E2 would readout as inhibition of WWP2 and whilst cross-reactivity tests exist for both, the timing of these tests is important. As the assay stands, if a derivative is showing activity, especially improved activity, an understanding of its WWP2 specificity with respect to the other assay components is required to fully understand if the inhibitory data associated with it can be trusted. As such cross-reactivity tests are required for each supposedly active derivative submitted due to the issues surrounding the assay readout reliability, perhaps a look at assay redevelopment is warranted. Having gained a better understanding of the sensitivity the biochemical assay has towards different mechanisms of interference; improvements could result in the mitigation of unproductive but supposedly active derivatives requiring verification. A test that asks a more particular research question would result in a reduction in the number of tests that verify the result, which would in turn reduce the time and necessity for validation of otherwise rather ambiguous results.

A fuller picture of how potential hits are interacting with the WWP2 HECT domain is highly desirable to understand better the nature of their binding, and this could be gained from additional biophysical experiments such as NMR, SPR and ITC experiments, for example. These experiments would not only further validate these compounds as hit molecules, but also provide important kinetic, thermodynamic, and visual information for analogue design.[294] To make best use of this information, these experiments have to occur prior to the decision to pursue analogues if more success in the identification of lead inhibitors is desired.

Depending upon the results of the co-crystal structure analysis, additional understanding as to the orientation of ligands within the binding site may be gained, which may inform future development, or provide evidence supporting reasoning as to why so little success was had. If NSC-288387 provides a high-quality co-crystal structure, a more specific understanding of how to perform the scaffold hop is probable, as proposed previously to remove the concerns surrounding reactivity of the alloxazine scaffold whilst maintaining activity.[28] Additionally, at this time these libraries are under investigation as WWP1 inhibitors, which may prove fruitful against this other promising cancer target.

8 <u>References</u>

1   L. Deng, T. Meng, L. Chen, W. Wei and P. Wang, *Signal Transduct. Target. Ther.*, 2020, **5**, 28.

2   G. Goldstein, M. Scheid, U. Hammerling, D. H. Schlesinger, H. D. Niall and E. A. Boyse, *Proc. Natl. Acad. Sci. U. S. A.*, 1975, **72**, 11–15.

3   The Nobel Prize in Chemistry in 2004, https://www.nobelprize.org/prizes/chemistry/2004/summary/, (accessed 26 January 2024).

4   S. Niazi, M. Purohit and J. H. Niazi, *Eur. J. Med. Chem.*, 2018, **158**, 7–24.

5   G. Kleiger and T. Mayor, *Trends Cell Biol.*, 2014, **24**, 352–359.

6   N. J. Schauer, R. S. Magin, X. Liu, L. M. Doherty and S. J. Buhrlage, *J. Med. Chem.*, 2020, **63**, 2731–2750.

7   D. Nandi, P. Tahiliani, A. Kumar and D. Chandu, *J. Biosci.*, 2006, **31**, 137–155.

8   S. M. Ge, S. Li, S. J. Chen, X. Y. Kong, W. Shi, H. J. Zhang, Y. H. Meng, L. Q. Shi, X. Liu, M. Wang, C. K. Zhang, C. Y. Chiu and C. Y. Lee, *23rd Int. Disp. Work. conjunction with Asia Display, IDW/AD*, 2016, **1**, 62–64.

9   M. J. Clague, C. Heride and S. Urbé, *Trends Cell Biol.*, 2015, **25**, 417–426.

10   X. Zou, G. Levy-Cohen and M. Blank, *Biochim. Biophys. Acta - Rev. Cancer*, 2015, **1856**, 91–106.

11   V. K. Chaugule and H. Walden, *Biochem. Soc. Trans.*, 2016, **44**, 212–227.

12   C. Grabbe, K. Husnjak and I. Dikic, *Nat. Rev. Mol. Cell Biol.*, 2011, **12**, 295–307.

13   L. D. Fricker, *Annu. Rev. Pharmacol. Toxicol.*, 2020, **60**, 457–476.

14   M. J. Clague and S. Urbé, *Cell*, 2010, **143**, 682–685.

15   B. A. Schulman and J. Wade Harper, *Nat. Rev. Mol. Cell Biol.*, 2009, **10**, 319–331.

16   Z. Tokgöz, R. N. Bohnsack and A. L. Haas, *J. Biol. Chem.*, 2006, **281**, 14729–14737.

17   C. M. Pickart and I. A. Rose, *J. Biol. Chem.*, 1985, **260**, 1573–1581.

18   A. L. Haas, P. M. Bright and V. E. Jackson, *J. Biol. Chem.*, 1988, **263**, 13268–13275.

19   M. B. Metzger, J. N. Pruneda, R. E. Klevit and A. M. Weissman, *Biochim. Biophys. Acta - Mol. Cell Res.*, 2014, **1843**, 47–60.

20   I. Lee and H. Schindelin, *Cell*, 2008, **134**, 268–278.

21   C. M. Pickart, *Annu. Rev. Biochem.*, 2001, **70**, 503–533.

22   T. Miura, W. Klaus, B. Gsell, C. Miyamoto and H. Senn, *J. Mol. Biol.*, 1999, **290**, 213–228.

23   M. B. Metzger, V. A. Hristova and A. M. Weissman, *J. Cell Sci.*, 2012, **125**, 531–537.

24   P. S. Brzovic, P. Rajagopal, D. W. Hoyt, M. C. King and R. E. Klevit, *Nat. Struct. Biol.*, 2001, **8**, 833–837.

25   A. Plechanovov, E. G. Jaffray, M. H. Tatham, J. H. Naismith and R. T. Hay, *Nature*, 2012, **489**, 115–120.

26   H. Dou, L. Buetow, G. J. Sibbet, K. Cameron and D. T. Huang, *Nat. Struct. Mol. Biol.*, 2012, **19**, 876–883.

27   J. N. Pruneda, P. J. Littlefield, S. E. Soss, K. A. Nordquist, W. J. Chazin, P. S. Brzovic and R. E. Klevit, *Mol. Cell*, 2012, **47**, 933–942.

28   G. R. Hughes, PhD thesis, University of East Anglia, 2019.

29   Q. Yang, J. Zhao, D. Chen and Y. Wang, *Mol. Biomed.*, 2021, **2**, 23.

30   S. Sánchez-Tena, M. Cubillos-Rojas, T. Schneider and J. L. Rosa, *Cell. Mol. Life Sci.*, 2016, **73**, 1955–1968.

31   H. B. Kamadurai, Y. Qiu, A. Deng, J. S. Harrison, C. MacDonald, M. Actis, P. Rodrigues, D. J. Miller, J. Souphron, S. M. Lewis, I. Kurinov, N. Fujii, M. Hammel, R. Piper, B. Kuhlman and B. A. Schulman, *Elife*, 2013, **2**, 1–26.

32   I. Marín, J. I. Lucas, A. C. Gradilla and A. Ferrús, *Physiol. Genomics*, 2004, **17**, 253–263.

33   I. Marín and A. Ferrúst, *Mol. Biol. Evol.*, 2002, **19**, 2039–2050.

34   K. K. Dove and R. E. Klevit, *J. Mol. Biol.*, 2017, **429**, 3363–3375.

35   X. S. Wang, T. R. Cotton, S. J. Trevelyan, L. W. Richardson, W. T. Lee, J. Silke and B. C. Lechtenberg, *Nat. Commun.*, 2023, **14**, 168.

36   K. A. NUSSBAUM, T. P. DICK, W. KEILHOLZ, R. G. T. SCHILD and E. Al., *Adv. Exp. Med. Biol.*, 2017, **995**, 53–71.

37   J. E. Watt, G. R. Hughes, S. Walpole, G. R. Stephenson, P. C. B. Page, A. M. Hemmings, J. Angulo and A. Chantry, *Chem. - A Eur. J. Commun.*, 2018, **24**, 17677–17680.

38   J. Sala-Gaston, A. Martinez-Martinez, L. Pedrazza, L. F. Lorenzo-Martín, R. Caloto, X. R. Bustelo, F. Ventura and J. L. Rosa, *Cancers (Basel).*, 2020, **12**, 1–13.

39   G. Pirozzi, S. J. McConnell, A. J. Uveges, J. M. Carter, A. B. Sparks, B. K. Kay and D. M. Fowlkes, *J. Biol. Chem.*, 1997, **272**, 14611–14616.

40   S. M. Soond and A. Chantry, *BioEssays*, 2011, **33**, 749–758.

41   S. Wiesner, A. A. Ogunjimi, H. R. Wang, D. Rotin, F. Sicheri, J. L. Wrana and J. D. Forman-Kay, *Cell*, 2007, **130**, 651–662.

42   S. Mari, N. Ruetalo, E. Maspero, M. C. Stoffregen, S. Pasqualato, S. Polo and S. Wiesner, *Structure*, 2014, **22**, 1639–1649.

43   H. I. Chen and M. Sudol, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 7819–7823.

44   A. Chantry, *Cell Cycle*, 2011, **10**, 2437–2439.

45   W. Zou, X. Chen, J. H. Shim, Z. Huang, N. Brady, D. Hu, R. Drapp, K. Sigrist, L. H. Glimcher and D. Jones, *Nat. Cell Biol.*, 2011, **13**, 59–65.

46   Y. Nakamura, K. Yamamoto, X. He, B. Otsuki, Y. Kim, H. Murao, T. Soeda, N. Tsumaki, J. M. Deng, Z. Zhang, R. R. Behringer, B. de Crombrugghe, J. H. Postlethwait, M. L. Warman, T. Nakamura and H. Akiyama, *Nat. Commun.*, 2011, **2**, 251.

47   S. M. Soond, P. G. Smith, L. Wahl, T. E. Swingler, I. M. Clark, A. M. Hemmings and A. Chantry, *Biochim. Biophys. Acta - Mol. Basis Dis.*, 2013, **1832**, 2127–2135.

48   S. Maddika, S. Kavela, N. Rani, V. R. Palicharla, J. L. Pokorny, J. N. Sarkaria and J. Chen, *Nat. Cell Biol.*, 2011, **13**, 728–733.

49   V. Stambolic, A. Suzuki, J. L. De la Pompa, G. M. Brothers, C. Mirtsos, T. Sasaki, J. Ruland, J. M. Penninger, D. P. Siderovski and T. W. Mak, *Cell*, 1998, **95**, 29–39.

50   I. Fabregat, J. Fernando, J. Mainez and P. Sancho, *Curr. Pharm. Des.*, 2014, **20**, 2934–2947.

51   L. A. Boyer, I. L. Tong, M. F. Cole, S. E. Johnstone, S. S. Levine, J. P. Zucker, M. G. Guenther, R. M. Kumar, H. L. Murray, R. G. Jenner, D. K. Gifford, D. A. Melton, R. Jaenisch and R. A. Young, *Cell*, 2005, **122**, 947–956.

52   S. Gidekel, G. Pizov, Y. Bergman and E. Pikarsky, *Cancer Cell*, 2003, **4**, 361–370.

53   S. K. Saha, Y. Jeong, S. Cho and S. G. Cho, *Sci. Rep.*, 2018, **8**, 1–19.

54   B. Liao and Y. Jin, *Cell Res.*, 2010, **20**, 332–344.

55   Y. W. Qian, Y. Chen, W. Yang, J. Fu, J. Cao, Y. Bin Ren, J. J. Zhu, B. Su, T. Luo, X. F. Zhao, R. Y. Dai, J. J. Li, W. Sun, M. C. Wu, G. S. Feng and H. Y. Wang, *Gastroenterology*, 2012, **142**, 1547–1558.

56   W. Chen, X. Jiang and Z. Luo, *Pathol. Oncol. Res.*, 2014, **20**, 799–803.

57   S. Wilk and M. Orlowski, *J. Neurochem.*, 1983, **40**, 842–849.

58   M. Hayashi, M. Inomata, Y. Saito, H. Ito and S. Kawashima, *BBA - Mol. Cell Res.*, 1991, **1094**, 249–256.

59   M. Bogyo, J. S. McMaster, M. Gaczynska, D. Tortorella, A. L. Goldberg and H. Ploegh, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 6629–6634.

60   L. Meng, R. Mohan, B. H. B. Kwok, M. Elofsson, N. Sin and C. M. Crews, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 10403–10408.

61   S. Ōmura, T. Fujimoto, K. Otoguro, K. Matsuzaki, R. Moriguchi, H. Tanaka and Y. Sasaki, *J. Antibiot. (Tokyo).*, 1991, **44**, 113–116.

62   S. Ōmura and A. Crump, *J. Antibiot. (Tokyo).*, 2019, **72**, 189–201.

63   M. Hanada, K. Sugawara, K. Kaneta, S. Toda, Y. Nishiyama, K. Tomita, H. Yamamoto, M. Konishi and T. Oki, *J. Antibiot. (Tokyo).*, 1992, **45**, 1746–1752.

64   J. Adams, M. Behnke, S. Chen, A. A. Cruickshank, L. R. Dick, L. Grenier, J. M. Klunder, Y.-T. Ma, L. Plamondon and R. L. Stein, *ChemInform*, 1998, **29**, 333–338.

65   D. J. Kuhn, Q. Chen, P. M. Voorhees, J. S. Strader, K. D. Shenk, C. M. Sun, S. D. Demo, M. K. Bennett, F. W. B. Van Leeuwen, A. A. Chanan-Khan and R. Z. Orlowski, *Blood*, 2007, **110**, 3281–3290.

66   E. Kupperman, E. C. Lee, Y. Cao, B. Bannerman, M. Fitzgerald, A. Berger, J. Yu, Y. Yang, P. Hales, F. Bruzzese, J. Liu, J. Blank, K. Garcia, C. Tsu, L. Dick, P. Fleming, L. Yu, M. Manfredi, M. Rolfe and J. Bolen, *Cancer Res.*, 2010, **70**, 1970–1980.

67   L. R. Dick and P. E. Fleming, *Drug Discov. Today*, 2010, **15**, 243–249.

68   J. Park, J. Cho and E. J. Song, *Arch. Pharm. Res.*, 2020, **43**, 1144–1161.

69   J. Adams and M. Kauffman, *Cancer Invest.*, 2004, **22**, 304–311.

70   J. Li, T. Yakushi, F. Parlati, A. L. MacKinnon, C. Perez, Y. Ma, K. P. Carter, S. Colayco, G. Magnuson, B. Brown, K. Nguyen, S. Vasile, E. Suyama, L. H. Smith, E. Sergienko, A. B. Pinkerton, T. D. Y. Chung, A. E. Palmer, I. Pass, S. Hess, S. M. Cohen and R. J. Deshaies, *Nat. Chem. Biol.*, 2017, **13**, 486–493.

71   G. Y. Yu, X. Wang, S. S. Zheng, X. M. Gao, Q. A. Jia, W. W. Zhu, L. Lu, H. L. Jia, J. H. Chen, Q. Z. Dong, M. Lu and L. X. Qin, *Cell. Physiol. Biochem.*, 2018, **47**, 1152–1166.

72   J. Li, Y. Zhang, B. Da Silva Sil Dos Santos, F. Wang, Y. Ma, C. Perez, Y. Yang, J. Peng, S. M. Cohen, T. F. Chou, S. T. Hilton and R. J. Deshaies, *Cell Chem. Biol.*, 2018, **25**, 1350-1358.e9.

73   J. J. Chen, C. A. Tsu, J. M. Gavin, M. A. Milhollen, F. J. Bruzzese, W. D. Mallender, M. D. Sintchak, N. J. Bump, X. Yang, J. Ma, H. K. Loke, Q. Xu, P. Li, N. F. Bence, J. E. Brownell and L. R. Dick, *J. Biol. Chem.*, 2011, **286**, 40867–40877.

74   J. M. Gavin, K. Hoar, Q. Xu, J. Ma, Y. Lin, J. Chen, W. Chen, F. J. Bruzzese, S. Harrison, W. D. Mallender, N. J. Bump, M. D. Sintchak, N. F. Bence, P. Li, L. R. Dick, A. E. Gould and J. J. Chen, *J. Biol. Chem.*, 2014, **289**, 22648–22658.

75   T. A. Soucy, P. G. Smith, M. A. Milhollen, A. J. Berger, J. M. Gavin, S. Adhikari, J. E. Brownell, K. E. Burke, D. P. Cardin, S. Critchley, C. A. Cullis, A. Doucette, J. J. Garnsey, J. L. Gaulin, R. E. Gershman, A. R. Lublinsky, A. McDonald, H. Mizutani, U. Narayanan, E. J. Olhava, S. Peluso, M. Rezaei, M. D. Sintchak, T. Talreja, M. P. Thomas, T. Traore, S. Vyskocil, G. S. Weatherhead, J. Yu, J. Zhang, L. R. Dick, C. F. Claiborne, M. Rolfe, J. B. Bolen and S. P. Langston, *Nature*, 2009, **458**, 732–736.

76   M. L. Hyer, M. A. Milhollen, J. Ciavarri, P. Fleming, T. Traore, D. Sappal, J. Huck, J. Shi, J. Gavin, J. Brownell, Y. Yang, B. Stringer, R. Griffin, F. Bruzzese, T. Soucy, J. Duffy, C. Rabino, J. Riceberg, K. Hoar, A. Lublinsky, S. Menon, M. Sintchak, N. Bump, S. M. Pulukuri, S. Langston, S. Tirrell, M. Kuranda, P. Veiby, J. Newcomb, P. Li, J. T. Wu, J. Powe, L. R. DIck, P. Greenspan, K. Galvin, M. Manfredi, C. Claiborne, B. S. Amidon and N. F. Bence, *Nat. Med.*, 2018, **24**, 186–193.

77   R. Sara, S. Paiva, M. Bancerz, M. Geletu, A. M. Lewis, J. Chen, Y. Cai, J. L. Lukkarila, H. Li and P. T. Gunning, *Bioorg. Med. Chem. Lett.*, 2016, **26**, 4542–4547.

78   R. Sekizawa, S. Ikeno, H. Nakamura, H. Naganawa, S. Matsui, H. Iinuma and T. Takeuchi, *J. Nat. Prod.*, 2002, **65**, 1491–1493.

79   M. Matsuzawa, H. Kakeya, J. Yamaguchi, M. Shoji, R. Onose, H. Osada and Y. Hayashi, *Chem. - An Asian J.*, 2006, **1**, 845–851.

80   S. Tsukamoto, H. Hirota, M. Imachi, M. Fujimuro, H. Onuki, T. Ohta and H. Yokosawa, *Bioorganic Med. Chem. Lett.*, 2005, **15**, 191–194.

81   J. Hong and H. Luesch, *Nat. Prod. Rep.*, 2012, **29**, 449–456.

82   D. Ungermannova, S. J. Parker, C. G. Nasveschuk, W. Wang, B. Quade, G. Zhang, R. D. Kuchta, A. J. Phillips and X. Liu, *PLoS One*, 2012, **7**, e29208.

83   Y. Yang, J. Kitagaki, R. M. Dai, C. T. Yien, K. L. Lorick, R. L. Ludwig, S. A. Pierre, J. P. Jensen, I. V. Davydov, P. Oberoi, C. C. H. Li, J. H. Kenten, J. A. Beutler, K. H. Vousden and A. M. Weissman, *Cancer Res.*, 2007, **67**, 9472–9481.

84   S. Tsukamoto, T. Takeuchi, H. Rotinsulu, R. E. P. Mangindaan, R. W. M. van

Soest, K. Ukai, H. Kobayashi, M. Namikoshi, T. Ohta and H. Yokosawa, *Bioorganic Med. Chem. Lett.*, 2008, **18**, 6319–6320.

85   D. F. Ceccarelli, X. Tang, B. Pelletier, S. Orlicky, W. Xie, V. Plantevin, D. Neculai, Y. C. Chou, A. Ogunjimi, A. Al-Hakim, X. Varelas, J. Koszela, G. A. Wasney, M. Vedadi, S. Dhe-Paganon, S. Cox, S. Xu, A. Lopez-Girona, F. Mercurio, J. Wrana, D. Durocher, S. Meloche, D. R. Webb, M. Tyers and F. Sicheri, *Cell*, 2011, **145**, 1075–1087.

86   M. Pulvino, Y. Liang, D. Oleksyn, M. DeRan, E. Van Pelt, J. Shapiro, I. Sanz, L. Chen and J. Zhao, *Blood*, 2012, **120**, 1668–1677.

87   C. D. Hodge, R. A. Edwards, C. J. Markin, D. McDonald, M. Pulvino, M. S. Y. Huen, J. Zhao, L. Spyracopoulos, M. J. Hendzel and J. N. M. Glover, *ACS Chem. Biol.*, 2015, **10**, 1718–1728.

88   S. Ushiyama, H. Umaoka, H. Kato, Y. Suwa, H. Morioka, H. Rotinsulu, F. Losung, R. E. P. Mangindaan, N. J. De Voogd, H. Yokosawa and S. Tsukamoto, *J. Nat. Prod.*, 2012, **75**, 1495–1499.

89   F. Bohlmann, J. Jakupovic, L. Müller and A. Schuster, *Angew. Chemie Int. Ed. English*, 1981, **20**, 292–293.

90   L. Liu, Y. Hua, D. Wang, L. Shan, Y. Zhang, J. Zhu, H. Jin, H. Li, Z. Hu and W. Zhang, *Chem. Biol.*, 2014, **21**, 1341–1350.

91   H. Chen, G. Wu, S. Gao, R. Guo, Z. Zhao, H. Yuan, S. Liu, J. Wu, X. Lu, X. Yuan, Z. Yu, X. Zu, N. Xie, N. Yang, Z. Hu, Q. Sun and W. Zhang, *J. Med. Chem.*, 2017, **60**, 6828–6852.

92   G. Laplante and W. Zhang, *Cancers (Basel).*, 2021, **13**, 1–43.

93   P. Ye, X. Chi, J. Cha, S. Luo, G. Yang, X. Yan and W. Yang, *Cells*, 2021, **10**, 2–34.

94   M. Rossi, B. Rotblat, K. Ansell, I. Amelio, M. Caraglia, G. Misso, F. Bernassola, C. N. Cavasotto, R. A. Knight, A. Ciechanover and G. Melino, *Cell Death Dis.*, 2014, **5**, 1–12.

95   Z. Wang, J. Wang, X. Li, L. Xing, Y. Ding, P. Shi, Y. Zhang, S. Guo, X. Shu and B. Shan, *Int. J. Oncol.*, 2014, **45**, 1469–1478.

96   M. Tian, T. Zeng, M. Liu, S. Han, H. Lin, Q. Lin, L. Li, T. Jiang, G. Li, H. Lin, T. Zhang, Q. Kang, X. Deng and H. R. Wang, *J. Biol. Chem.*, 2019, **294**, 2880–2891.

97   Q. Lu, S. Luo, Z. Shi, M. Yu, W. Guo and C. Li, *Front. Pharmacol.*, 2022, **13**, 1–16.

98   J. Zhang, R. Cao, C. Lian, T. Cao, Y. Shi, J. Ma, P. Wang and J. Xia, *Aging (Albany. NY).*, 2021, **13**, 782–793.

99   J. Zhang, J. J. Xie, S. J. Zhou, J. Chen, Q. Hu, J. X. Pu and J. L. Lu, *Am. J. Transl. Res.*, 2019, **11**, 3461–3471.

100   Y. Chen, R. Zhang, W. Zhao, M. Lv, M. Chen, Y. Yan and S. Feng, *Am. J. Transl. Res.*, 2019, **11**, 7579–7590.

101   T. Mund, M. J. Lewis, S. Maslen and H. R. Pelham, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 16736–16741.

102   M. Pettersson and C. M. Crews, *Drug Discov. Today Technol.*, 2019, **31**, 15–27.

103   G. R. Hughes, A. P. Dudey, A. M. Hemmings and A. Chantry, *Drug Discov. Today*, 2021, **26**, 2377–2383.

104 M. Zengerle, K. H. Chan and A. Ciulli, *ACS Chem. Biol.*, 2015, **10**, 1770–1777.

105 M. S. Gadd, A. Testa, X. Lucas, K. H. Chan, W. Chen, D. J. Lamont, M. Zengerle and A. Ciulli, *Nat. Chem. Biol.*, 2017, **13**, 514–521.

106 M. Tanaka, J. M. Roberts, H. S. Seo, A. Souza, J. Paulk, T. G. Scott, S. L. Deangelo, S. Dhe-Paganon and J. E. Bradner, *Nat. Chem. Biol.*, 2016, **12**, 1089–1096.

107 S. Imaide, K. M. Riching, N. Makukhin, V. Vetma, C. Whitworth, S. J. Hughes, N. Trainor, S. D. Mahan, N. Murphy, A. D. Cowan, K. H. Chan, C. Craigon, A. Testa, C. Maniaci, M. Urh, D. L. Daniels and A. Ciulli, *Nat. Chem. Biol.*, 2021, **17**, 1157–1167.

108 J. Hines, S. Lartigue, H. Dong, Y. Qian and C. M. Crews, *Cancer Res.*, 2019, **79**, 251–262.

109 A. R. Schneekloth, M. Pucheault, H. S. Tae and C. M. Crews, *Bioorganic Med. Chem. Lett.*, 2008, **18**, 5904–5908.

110 Y. Itoh, M. Ishikawa, M. Naito and Y. Hashimoto, *J. Am. Chem. Soc.*, 2010, **132**, 5820–5826.

111 M. Naito, N. Ohoka and N. Shibata, *Drug Discov. Today Technol.*, 2019, **31**, 35–42.

112 C. Maniaci, S. J. Hughes, A. Testa, W. Chen, D. J. Lamont, S. Rocha, D. R. Alessi, R. Romeo and A. Ciulli, *Nat. Commun.*, 2017, **8**, 830.

113 C. Steinebach, S. Lindner, N. D. Udeshi, D. C. Mani, H. Kehm, S. Köpff, S. A. Carr, M. Gütschow and J. Krönke, *ACS Chem. Biol.*, 2018, **13**, 2771–2782.

114 M. Girardini, C. Maniaci, S. J. Hughes, A. Testa and A. Ciulli, *Bioorganic Med. Chem.*, 2019, **27**, 2466–2479.

115 J. Liu, H. Chen, L. Ma, Z. He, D. Wang, Y. Liu, Q. Lin, T. Zhang, N. Gray, H. Ü. Kaniskan, J. Jin and W. Wei, *Sci. Adv.*, 2020, **6**, 1–12.

116 G. Xue, K. Wang, D. Zhou, H. Zhong and Z. Pan, *J. Am. Chem. Soc.*, 2019, **141**, 18370–18374.

117 P. Pfaff, K. T. G. Samarasinghe, C. M. Crews and E. M. Carreira, *ACS Cent. Sci.*, 2019, **5**, 1682–1690.

118 M. Reynders and D. Trauner, *Methods Mol. Biol.*, 2021, **2365**, 315–329.

119 S. A. Kanner, Z. Shuja, P. Choudhury, A. Jain and H. M. Colecraft, *Nat. Methods*, 2020, **17**, 1245–1253.

120 D. Li, J. Zhang, L. Yin, Z. Jin, X. Chen and X. Meng, *Exp. Ther. Med.*, 2021, **22**, 1–9.

121 D. Yong, S. R. Green, P. Ghiabi, V. Santhakumar and M. Vedadi, *Sci. Rep.*, 2023, **13**, 1–15.

122 S. Zahid, M. Gul, S. Shafique and S. Rashid, *Comput. Biol. Med.*, 2022, **146**, 105660.

123 W. Zheng, Z. Shi, X. Zhang, W. Wu, Z. Yuan, L. Zhao, Q. Li, Z. Qiu, C. Zhang, B. Chu, Z. Liu, W. M. Chen and Y. Jiang, *Phytochem. Lett.*, 2023, **54**, 7–13.

124 Y. Jiang, Y. Fang, Y. Ye, X. Xu, B. Wang, J. Gu, M. Aschner, J. Chen and R. Lu, *Front. Pharmacol.*, 2019, **10**, 1–23.

125 J. R. Weng, C. H. Tsai, S. K. Kulp, D. Wang, C. H. Lin, H. C. Yang, Y. Ma, A. Sargeant, C. F. Chiu, M. H. Tsai and C. S. Chen, *Cancer Res.*, 2007, **67**, 7815–7824.

126 J. G. Quirit, S. N. Lavrenov, K. Poindexter, J. Xu, C. Kyauk, K. A. Durkin, I. Aronchik, T. Tomasiak, Y. A. Solomatin, M. N. Preobrazhenskaya and G. L. Firestone, *Biochem. Pharmacol.*, 2017, **127**, 13–27.

127 US 2004/0043965 A1, 2004.

128 J. Y. Hall, Work towards a PhD thesis, University of East Anglia., 2021.

129 L. Hoffer, C. Muller, P. Roche and X. Morelli, *Mol. Inform.*, 2018, 37, 1800059.

130 G. M. Morris and M. Lim-Wilby, *Methods Mol. Biol.*, 2008, **443**, 365–382.

131 W. Gong, X. Zhang, W. Zhang, J. Li and Z. Li, *Acta Crystallogr. Sect. Struct. Biol. Commun.*, 2015, **71**, 1251–1257.

132 M. A. H. De Zwart, H. van der Goot and H. Timmerman, *J. Med. Chem.*, 1988, **31**, 716–722.

133 J. E. Van Muijlwijk-Koezen, H. Timmerman, R. Link, H. Van der Goot and A. P. Ijzerman, *J. Med. Chem.*, 1998, **41**, 3987–3993.

134 J. E. Van Muijlwijk-Koezen, H. Timmerman, R. Link, H. Van der Goot and A. P. Uzerman, *J. Med. Chem.*, 1998, **41**, 3994–4000.

135 M. A. H. De Zwart, G. Henk Van der and H. Timmerman, *J. Med. Chem.*, 1989, **32**, 487–493.

136 R. Capochiani de Iudicibus, P. Tomek, B. D. Palmer, S. M. Tijono, J. U. Flanagan and L. M. Ching, *Bioorganic Med. Chem.*, 2021, **39**, 116160.

137 D. Clerin, A. Lacroix and J.-P. Fleury, *Tetrahedron Lett.*, 1976, 2899–2902.

138 J. P. Schoeni and J. P. Fleury, *Tetrahedron*, 1975, **31**, 671–678.

139 J. Fleury, J. Schoeni and D. Clerin, *Helv. Chim. Acta*, 1975, **58**, 2018–2026.

140 M. Lang, J. Schoeni, C. Pont and J. Fleury, *Helv. Chim. Acta*, 1986, **69**, 793–802.

141 O. H. Wheeler and D. Gonzalez, *Tetrahedron*, 1964, **20**, 189–193.

142 H. M. Nanjundaswamy and M. A. Pasha, *J. Chem. Res.*, 2005, **12**, 772–774.

143 S. Savard and P. D. Josephy, *Carcinogenesis*, 1986, **7**, 1239–1241.

144 H. J. Shine, H. Zmuda and H. P. Koon, *J. Am. Chem. Soc.*, 1981, **103**, 955–956.

145 W. Subotkowski, L. Kupczyk-Subotkowska and H. J. Shine, *J. Am. Chem. Soc.*, 1993, **115**, 5073–5076.

146 I. Sungwienwong, J. J. Ferrie, J. V. Jun, C. Liu, T. M. Barrett, Z. M. Hostetler, N. Ieda, A. Hendricks, A. K. Muthusamy, R. M. Kohli, D. M. Chenoweth, G. A. Petersson and E. J. Petersson, *J. Phys. Org. Chem.*, 2018, **31**, 1–10.

147 C. I. Canché Chay, R. G. Cansino, C. I. Espitia Pinzón, R. O. Torres-Ochoa and R. Martínez, *Mar. Drugs*, 2014, **12**, 1757–1772.

148 K. K. Chaudhary and N. Mishra, *A Rev. Mol. Docking Nov. Tool Drug Discov. JSM Chem*, 2016, **4**, 1029.

149   A. Fischer, M. Smieško, M. Sellner and M. A. Lill, *J. Med. Chem.*, 2021, **64**, 2489–2500.

150   P. Ripphausen, D. Stumpfe and J. Bajorath, *Future Med. Chem.*, 2012, **4**, 603–613.

151   M. J. Hartshorn, *J. Comput. Aided. Mol. Des.*, 2002, **16**, 871–881.

152   C. Bissantz, B. Kuhn and M. Stahl, *J. Med. Chem.*, 2010, **53**, 5061–5084.

153   H. Rimac, M. Grishina and V. Potemkin, *J. Chem. Inf. Model.*, 2021, **61**, 1801–1813.

154   G. Wu, H. D. Robertson, C. L. Brooks III and M. Veith, *J. Comput. Chem.*, 2003, **24**, 1549–1562.

155   B. M. Hudson, E. Nguyen and D. J. Tantillo, *Org. Biomol. Chem.*, 2016, **14**, 3975–3980.

156   T. Cheeseright, M. Mackey, S. Rose and A. Vinter, *J. Chem. Inf. Model.*, 2006, **46**, 665–676.

157   G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw and K. I. Goldberg, *Organometallics*, 2010, **29**, 2176–2179.

158   L. Li, A. El Khoury, B. O. N. Clement, C. Wu and P. G. Harran, *Org. Lett.*, 2022, **24**, 2607–2612.

159   A. C. Sousa, S. R. Baptista, L. O. Martins and M. P. Robalo, *Chem. - An Asian J.*, 2019, **14**, 187–193.

160   C. K. Ingold and H. V. Kidd, *J. Chem. Soc.*, 1933, 984–986.

161   V. R. Holland, B. C. Saunders, F. L. Rose and A. L. Walpole, *Tetrahedron*, 1974, **30**, 3299–3302.

162   N. S. Frumina and I. S. Mustafin, *Zhurnal Anal. Khimii*, 1960, **15**, 671–675.

163   N. Sakai, S. Asama, S. Anai and T. Konakahara, *Tetrahedron*, 2014, **70**, 2027–2033.

164   H. Zhou, R. Fan, J. Yang, X. Sun, X. Liu and X. C. Wang, *J. Org. Chem.*, 2022, **87**, 14536–14543.

165   T. Ikeda, I. Aprahamian and J. F. Stoddart, *Org. Lett.*, 2007, **9**, 1481–1484.

166   Z. Lei, F. Xue, B. Wang, S. Wang, Y. Zhang, Y. Xia, W. Jin and C. Liu, *Green Chem.*, 2022, **25**, 348–356.

167   Y. C. Tong, *J. Heterocycl. Chem.*, 1981, **18**, 22–24.

168   S. M. Khan, X. Zhang and W. H. Witola, *Front. Microbiol.*, 2022, **12**, 1–14.

169   F. Ballante, D. R. Reddy, N. J. Zhou and G. R. Marshall, *Bioorganic Med. Chem.*, 2017, **25**, 2105–2132.

170   M. Krause, H. Foks and K. Gobis, *Molecules*, 2017, **22**, 399.

171   M. Legraverend and D. S. Grierson, *Bioorganic Med. Chem.*, 2006, **14**, 3987–4006.

172   C. Karaaslan, F. Doganc, M. Alp, A. Koc, A. Z. Karabay and H. Göker, *J. Mol. Struct.*, 2020, **1205**, 1–13.

173   D. S. Mortensen, S. M. Perrin-Ninkovic, R. Harris, B. G. S. Lee, G. Shevlin, M.

Hickman, G. Khambatta, R. R. Bisonette, K. E. Fultz and S. Sankar, *Bioorganic Med. Chem. Lett.*, 2011, **21**, 6793–6799.

174   F. Zhao, J. Zhang, L. Zhang, Y. Hao, C. Shi, G. Xia, J. Yu and Y. Liu, *Bioorganic Med. Chem.*, 2016, **24**, 4281–4290.

175   J. E. Payne, C. Bonnefous, K. T. Symons, P. M. Nguyen, M. Sablad, N. Rozenkrants, Y. Zhang, L. Wang, N. Yazdani, A. K. Shiau, S. A. Noble, P. Rix, T. S. Rao, C. A. Hassig and N. D. Smith, *J. Med. Chem.*, 2010, **53**, 7739–7755.

176   C. G. Allison, R. D. Chambers, J. A. H. MacBride and W. K. R. Musgrave, *J. Chem. Soc. C Org.*, 1970, 1023–1029.

177   Y. C. Tong, *J. Heterocycl. Chem*, 1975, **12**, 1127–1131.

178   G. W. H. Cheeseman and E. S. G. Werstiuk, *Adv. Heterocycl. Chem.*, 1972, **14**, 99–209.

179   S. Oda, T. Shimizu, T. Katayama, H. Yoshikawa and T. Hatakeyama, *Org. Lett.*, 2019, **21**, 1770–1773.

180   J. Fleischhauer, S. Zahn, R. Beckert, U. W. Grummt, E. Birckner and H. Görls, *Chem. - A Eur. J.*, 2012, **18**, 4549–4557.

181   F. Da Paixao Soares, E. Groaz and P. Herdewijn, *Molecules*, 2018, **23**, 1457.

182   E. Lizano, J. Grima and M. D. Pujol, *Synlett*, 2019, **30**, 2000–2003.

183   O. Vitse, F. Laurent, T. M. Pocock, V. Bénézech, L. Zanik, K. R. F. Elliott, G. Subra, K. Portet, J. Bompart, J. P. Chapat, R. C. Small, A. Michel and P. A. Bonnet, *Bioorganic Med. Chem.*, 1999, **7**, 1059–1065.

184   M. Wang, Y. Zhang, T. Wang, C. Wang, D. Xue and J. Xiao, *Org. Lett.*, 2016, **18**, 1976–1979.

185   J. Bergman, B. Pettersson, V. Hasimbegovic and P. H. Svensson, *J. Org. Chem.*, 2011, **76**, 1546–1553.

186   I. Thomsen, K. Clausen, S. Scheibye, S.-O. Lawesson, H. C. Heathcock, M. Sanner and T. Rosen, *Org. Synth.*, 1984, **62**, 158.

187   K. Felczak, M. Bretner, T. Kulikowskit and D. Shugar, *Nucleosides and Nucleotides*, 1993, **12**, 245–261.

188   L. Legnani, L. Toma, P. Caramella, M. A. Chiacchio, S. Giofrè, I. Delso, T. Tejero and P. Merino, *J. Org. Chem.*, 2016, **81**, 7733–7740.

189   L. Chang, S. Y. Lee, P. Leonczak, J. Rozenski, S. De Jonghe, T. Hanck, C. E. Müller and P. Herdewijn, *J. Med. Chem.*, 2014, **57**, 10080–10100.

190   R. V. J. Chari and J. W. Kozarich, *J. Org. Chem.*, 1982, **47**, 2355–2358.

191   J. E. Cragoe, M. C. Robb, J. Cason and D. J. Willett, *Org. Synth.*, 1960, **40**, 54.

192   T. Deligeorgiev, S. Kaloyanova, N. Lesev and J. J. Vaquero, *Ultrason. Sonochem.*, 2010, **17**, 783–788.

193   A. E. Alpers and A. Taurins, *Can. J. Chem.*, 1967, **45**, 2903–2912.

194   B. Li, M. Berliner, R. Buzon, C. K. F. Chiu, S. T. Colgan, T. Kaneko, N. Keene, W. Kissel, T. Le, K. R. Leeman, B. Marquez, R. Morris, L. Newell, S. Wunderwald, M. Witt, J. Weaver, Z. Zhang and Z. Zhang, *J. Org. Chem.*, 2006, **71**, 9045–9050.

195 V. Theodorou, K. Skobridis, A. G. Tzakos and V. Ragoussis, *Tetrahedron Lett.*, 2007, **48**, 8230–8233.

196 L. Jiang, J. Yu, F. Niu, D. Zhang and X. Sun, *Heteroat. Chem.*, 2017, **28**, 1–8.

197 Y. C. Tong, *J. Heterocycl. Chem.*, 1975, **12**, 1127–1131.

198 R. Saito, K. Ishibashi, M. Noumi, S. Uno, S. Higashi, M. Goto, S. Kuwahara and T. Komatsu, *Chem. Pharm. Bull.*, 2019, **67**, 556–565.

199 A. Ohta and Y. Akita, *Heterocycles*, 1981, **16**, 1325.

200 D. L. Ladd, *J. Heterocycl. Chem.*, 1982, **19**, 917–921.

201 A. Holy, I. Votruba, G. Andrei, R. Snoeck, L. Naesens, E. De Clercq and J. Balzarini, *J. Med. Chem.*, 2002, **45**, 1918–1929.

202 L. Bethge, D. V. Jarikote and O. Seitz, *Bioorganic Med. Chem.*, 2008, **16**, 114–125.

203 B. Zeysing, C. Gosch and A. Terfort, *Org. Lett.*, 2000, **2**, 1843–1845.

204 M. A. Düfert, K. L. Billingsley and S. L. Buchwald, *J. Am. Chem. Soc.*, 2013, **135**, 12877–12885.

205 H. G. Cheng, H. Chen, Y. Liu and Q. Zhou, *Asian J. Org. Chem.*, 2018, **7**, 490–508.

206 D. S. Mortensen, S. M. Perrin-Ninkovic, R. Harris, B. G. S. Lee, G. Shevlin, M. Hickman, G. Khambatta, R. R. Bisonette, K. E. Fultz and S. Sankar, *Bioorganic Med. Chem. Lett.*, 2011, **21**, 6793–6799.

207 T. Itoh and T. Mase, *J. Org. Chem.*, 2006, **71**, 2203–2206.

208 M. H. Norman, N. Chen, Z. Chen, C. Fotsch, C. Hale, N. Han, R. Hurt, T. Jenkins, J. Kincaid, L. Liu, Y. Lu, O. Moreno, V. J. Santora, J. D. Sonnenberg and W. Karbon, *J. Med. Chem.*, 2000, **43**, 4288–4312.

209 A. N. Krasovskii and P. M. Kochergin, 1968, **34**, 78–86.

210 X. Q. Li and C. Zhang, *Synthesis (Stuttg).*, 2009, **7**, 1163–1169.

211 W. E. Van Zyl and J. P. Fackler, *Phosphorus, Sulfur Silicon Relat. Elem.*, 2000, **167**, 117–132.

212 J. E. Meegan, X. Yang, R. Rungsirisakun, S. C. Cosgrove, R. J. Bushby, A. Sadeghpour, M. Rappolt, R. Brydson and R. J. Ansell, *Soft Matter*, 2017, **13**, 5922–5932.

213 R. Kataky, D. Parker, A. Teasdale, J. P. Hutchinson and H. J. Buschmann, *J. Chem. Soc. Perkin Trans. 2*, 1992, **8**, 1347–1351.

214 L. Roiser and M. Waser, *Org. Lett.*, 2017, **19**, 2338–2341.

215 P. D. Kiser, J. Zhang, M. Badiee, J. Kinoshita, N. S. Peachey, G. P. Tochtrop and K. Palczewski, *J. Pharmacol. Exp. Ther.*, 2017, **362**, 131–145.

216 N. Ibrahim, L. Mouawad and M. Legraverend, *Eur. J. Med. Chem.*, 2010, **45**, 3389–3393.

217 C. Sablayrolles, G. H. Cros, J. C. Milhavet, M. Boucard, J. J. Serrano and J. H. Mcneill, *J. Med. Chem.*, 1984, **27**, 206–212.

218 A. A. Abu-Hashem and M. A. Gouda, *J. Het*, 2017, **54**, 850–858.

219   C. L. Gibson, S. La Rosa and C. J. Suckling, *Org. Biomol. Chem.*, 2003, **1**, 1909–1918.

220   Z. J. Jain, P. S. Gide and R. S. Kankate, *Arab. J. Chem.*, 2017, **10**, S2051–S2066.

221   M. Aggarwal, J. A. Sommers, R. H. Shoemaker and R. M. Brosh, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 1525–1530.

222   V. E. A. Rea, A. Lavecchia, C. Di Giovanni, F. W. Rossi, A. Gorrasi, A. Pesapane, A. De Paulis, P. Ragno and N. Montuori, *Mol. Cancer Ther.*, 2013, **12**, 1402–1416.

223   G. V. Beyer, S. Hueser, R. Li, D. Manika, M. Lee, C. H. F. Chan, J. R. Howe and P. H. Ear, *Surg. (United States)*, 2024, **175**, 605–612.

224   B. A. P. Wilson, H. Wang, B. A. Nacev, R. C. Mease, J. O. Liu, M. G. Pomper and W. B. Isaacs, *Mol. Cancer Ther.*, 2011, **10**, 825–838.

225   N. S. Frumina and I. S. Mustafin, *Zh. Anal. Khim.*, 1960, **15**, 671–675.

226   J. Song, S. R. Malwal, N. Baig, L. A. Schurig-Briccio, Z. Gao, G. S. Vaidya, K. Yang, N. S. Abutaleb, M. N. Seleem, R. B. Gennis, T. V. Pogorelov, E. Oldfield and X. Feng, *ACS Infect. Dis.*, 2020, **6**, 2979–2993.

227   E. Pery, A. Sheehy, N. Miranda Nebane, V. Misra, M. K. Mankowski, L. Rasmussen, E. Lucile White, R. G. Ptak and D. Gabuzda, *Virology*, 2015, **484**, 276–287.

228   W. Knecht and M. Löffler, *FEBS Lett.*, 2000, **467**, 27–30.

229   Q. Bao and J. Zhou, *Eur. J. Med. Chem.*, 2023, **250**, 115118.

230   X. Y. Yan, J. F. Leng, T. T. Chen, Y. J. Zhao, L. Y. Kong and Y. Yin, *Eur. J. Med. Chem.*, 2022, **237**, 114372.

231   D. Chaudhary, A. Singh, M. Marzuki, A. Ghosh, S. Kidwai, T. P. Gosain, K. Chawla, S. K. Gupta, N. Agarwal, S. Saha, Y. Kumar, K. G. Thakur, A. Singhal and R. Singh, *Sci. Rep.*, 2022, **12**, 1–16.

232   A. M. Beekman, M. A. O'Connell and L. A. Howell, *ChemMedChem*, 2016, **11**, 840–844.

233   K. V. Rajendran, L. Kennedy, C. T. O'Connor, E. Bergin and D. G. Gilheany, *Tetrahedron Lett.*, 2013, **54**, 7009–7012.

234   S. A. Shipilovskikh, V. Y. Vaganov, E. I. Denisova, A. E. Rubtsov and A. V. Malkov, *Org. Lett.*, 2018, **20**, 728–731.

235   T. M. Kosak, H. A. Conrad, A. L. Korich and R. L. Lord, *European J. Org. Chem.*, 2015, **2015**, 7460–7467.

236   D. G. Brown and J. Boström, *J. Med. Chem.*, 2016, **59**, 4443–4458.

237   N. Miyaura, T. Yanagi and A. Suzuki, *Synth. Commun.*, 1981, **11**, 513–519.

238   M. C. D'Alterio, E. Casals-Cruanas, V. N. Tzouraz, G. Talarico, S. P. Nolan and A. Poater, *Chem. - A Eur. J.*, 2021, **27**, 13481–13493.

239   U. Christmann and R. Vilar, *Angew. Chemie - Int. Ed.*, 2005, 44, 366–374.

240   S. J. Firsan, V. Sivakumar and T. J. Colacot, *Chem. Rev.*, 2022, **122**, 16983–17027.

241   V. F. Slagt, A. H. M. De Vries, J. G. De Vries and R. M. Kellogg, *Org. Process Res.*

Dev., 2010, **14**, 30–47.

242   T. Itoh and T. Mase, *Tetrahedron Lett.*, 2005, **46**, 3573–3577.

243   R. Martin and S. L. Buchwald, *Acc. Chem. Res.*, 2008, **41**, 1461–1473.

244   F. Barrios-Landeros and J. F. Hartwig, *J. Am. Chem. Soc.*, 2005, **127**, 6944–6945.

245   J. F. Hartwig, *Inorg. Chem.*, 2007, **46**, 1936–1947.

246   E. R. Strieter and S. L. Buchwald, *Angew. Chemie - Int. Ed.*, 2006, **45**, 925–928.

247   P. Jacob, P. S. Gallery, A. T. Shulgin and N. Castagnoli, *J. Org. Chem.*, 1976, **41**, 3627–3629.

248   R. Dorel, C. P. Grugel and A. M. Haydl, *Angew. Chemie - Int. Ed.*, 2019, **58**, 17118–17129.

249   M. M. Heravi, Z. Kheilkordi, V. Zadsirjan, M. Heydari and M. Malmir, *J. Organomet. Chem.*, 2018, **861**, 17–104.

250   J. F. Hartwig, *Synlett*, 1997, **4**, 329–340.

251   Z. Zuo, X. Liu, X. Qian, T. Zeng, N. Sang, H. Liu, Y. Zhou, L. Tao, X. Zhou, N. Su, Y. Yu, Q. Chen, Y. Luo and Y. Zhao, *J. Med. Chem.*, 2020, **63**, 7633–7652.

252   J. B. Baell and J. W. M. Nissink, *ACS Chem. Biol.*, 2018, **13**, 36–44.

253   J. B. Baell and G. A. Holloway, *J. Med. Chem.*, 2010, **53**, 2719–2740.

254   Baell J., *Nature*, 2014, **513**, 481–483.

255   Y. Zafrani, D. Yeffet, G. Sod-Moriah, A. Berliner, D. Amir, D. Marciano, E. Gershonov and S. Saphier, *J. Med. Chem.*, 2017, **60**, 797–804.

256   T. Ishiyama, M. Murata and N. Miyaura, *J. Org. Chem.*, 1995, **60**, 7508–7510.

257   J. Morey and J. M. Saá, *Tetrahedron*, 1993, **49**, 105–112.

258   J. Bajorath, *Expert Opin. Drug Discov.*, 2021, **16**, 719–721.

259   W. Chen, B. Feng, S. Han, P. Wang, W. Chen, Y. Zang, J. Li and Y. Hu, *Bioorganic Med. Chem. Lett.*, 2022, **58**, 128526.

260   R. H. Pawle, A. Agarwal, S. Malveira, Z. C. Smith and S. W. Thomas, *Macromolecules*, 2014, **47**, 2250–2256.

261   L. H. Amundsen and L. S. Nelson, *J. Am. Chem. Soc.*, 1951, **73**, 242–244.

262   R. Klein, S. N. Sunassee and M. T. Davies-Coleman, *J. Chem. Res.*, 2009, **8**, 468–472.

263   M. Veguillas, M. C. Redondo, I. García, M. Ribagorda and M. Carmen Carreño, *Chem. - A Eur. J.*, 2010, **16**, 3707–3719.

264   M. Yoshida, Y. Shoji and K. Shishido, *Org. Lett.*, 2009, **11**, 1441–1443.

265   J. I. Peyton, S. Callery, P, T. Shulgin, A and N. Castagnoli, *J. Org. Chem.*, 1976, **41**, 3627–3629.

266   P. . Brown, Sean, J. . Dransfield, Paul, J. Houze, J. . Kohn, Todd, J. Liu, J. Medina, V. Pattaropong, W. Shen, M. Vimolratana, Y. Wang, M. Yu and L. Zhu, 2009, WO2005086661A2.

267 Y. S. Cho, H. Angove, C. Brain, C. H. Chen, H. Cheng, R. Cheng, R. Chopra, K. Chung, M. Congreve, C. Dagostin, D. J. Davis, R. Feltell, J. Giraldes, S. D. Hiscock, S. Kim, S. Kovats, B. Lagu, K. Lewry, A. Loo, Y. Lu, M. Luzzio, W. Maniara, R. Mcmenamin, P. N. Mortenson, R. Benning, M. O. Reilly, D. C. Rees, J. Shen, T. Smith, Y. Wang, G. Williams, A. J. Woolford, W. Wrona, M. Xu, F. Yang and S. Howard, *ACS Med. Chem. Lett.*, 2012, **3**, 445–449.

268 D. Kohler and J. Podlech, *European J. Org. Chem.*, 2019, **2019**, 1748–1753.

269 S. Chakrabarty, D. A. Monlish, M. Gupta, T. D. Wright, V. T. Hoang, M. Fedak, I. Chopra, P. T. Flaherty, J. Madura, S. Mannepelli, M. E. Burow and J. E. Cavanaugh, *Bioorganic Med. Chem. Lett.*, 2018, **28**, 2294–2301.

270 J. R. Weng, C. H. Tsai, S. K. Kulp and C. S. Chen, *Cancer Lett.*, 2008, **262**, 153–163.

271 A. H. Sandtorv, *Adv. Synth. Catal.*, 2015, **357**, 2403–2435.

272 Y. Taskesenligil, M. Aslan, T. Cogurcu and N. Saracoglu, *J. Org. Chem.*, 2023, **88**, 1299–1318.

273 M. Z. Zhang, N. Mulholland, A. Seville, G. Hough, N. Smith, H. Q. Dong, W. H. Zhang and Y. C. Gu, *Tetrahedron*, 2021, **79**, 131835.

274 C. Foord, BSc Final year project, University of East Anglia, 2022.

275 R. Iioka, K. Yorozu, Y. Sakai, R. Kawai, N. Hatae, K. Takashima, G. Tanabe, H. Wasada and M. Yoshimatsu, *European J. Org. Chem.*, 2021, **2021**, 1553–1558.

276 Y. Sawama, M. Masuda, S. Asai, R. Goto, S. Nagata, S. Nishimura, Y. Monguchi and H. Sajiki, *Org. Lett.*, 2015, **17**, 434–437.

277 C. J. Martin, A. T. L. Lee, R. W. Adams and D. A. Leigh, *J. Am. Chem. Soc.*, 2017, **139**, 11998–112002.

278 Y. Sawama, Y. Miki and H. Sajiki, *Synlett*, 2020, **31**, 699–702.

279 D. Damour, J. P. Pulicani, M. Vuilhorgne and S. Mignani, *Synlett*, 1999, **6**, 786–788.

280 P. Kutschy, M. Dzurilla, M. Takasugi, M. Török, I. Achbergerová, R. Homzová and M. Rácová, *Tetrahedron*, 1998, **54**, 3549–3566.

281 W. R. Chao, D. Yean, K. Amin, C. Green and L. Jong, *J. Med. Chem.*, 2007, **50**, 3412–3415.

282 Y. C. Hu, D. W. Ji, C. Y. Zhao, H. Zheng and Q. A. Chen, *Angew. Chemie - Int. Ed.*, 2019, **58**, 5438–5442.

283 S. Wang, J. Chen, F. Zhang, Y. Zhao, X. Wu and R. Chen, *Chem. Commun.*, 2023, **59**, 6568–6571.

284 S. Gandhi and B. Baire, *J. Org. Chem.*, 2019, **84**, 3904–3918.

285 S. Azeez, P. Sureshbabu, S. Sabiah and J. Kandasamy, *Org. Biomol. Chem.*, 2022, **20**, 2048–2053.

286 A. Arcadi, A. Calcaterra, M. Chiarini, G. Fabrizi, A. Fochetti, A. Goggiamani, A. Iazzetti, F. Marrone, V. Marsicano and A. Serraiocco, *Synth.*, 2021, **54**, 741–753.

287 D. Damour, J. P. Pulicani, M. Vuilhorgne and S. Mignani, *Synlett*, 1999, 786–788.

288 G. W. Gribble, D. J. Keavy, D. A. Davis, M. G. Saulnier, B. Pelcman, T. C. Barden, M. P. Sibi, E. R. Olson and J. J. BelBruno, *J. Org. Chem.*, 1992, **57**, 5878–5891.

289 J. B. Baell, *J. Nat. Prod.*, 2016, **79**, 616–628.

290 Z. Li, X. Bao, X. Bai, G. Zhang, J. Wang, M. Zhu, Y. Wang, J. Shang, C. Sheng, D. Zhang and Y. Wang, *Sci. Rep.*, 2019, **9**, 1–7.

291 J. Lüken, G. Goerges, N. Ritter, P. Disse, J. A. Schreiber, J. Schmidt, B. Frehland, D. Schepmann, G. Seebohm and B. Wünsch, *J. Med. Chem.*, 2023, **66**, 11573–11588.

292 E. Gilberg, D. Stumpfe and J. Bajorath, *RSC Adv.*, 2017, **7**, 35638–35647.

293 J. L. Dahlin, J. W. M. Nissink, J. M. Strasser, S. Francis, L. Higgins, H. Zhou, Z. Zhang and M. A. Walters, *J. Med. Chem.*, 2015, **58**, 2091–2113.

294 A. J. Ruben, Y. Kiso and E. Freire, *Chem. Biol. Drug Des.*, 2006, **67**, 2–4.

9 Appendix 1: XRD data

9.1 Table 9.1: Crystal data and structure refinement for single X-ray structures 9.1, 9.2, 9.3, 9.4, 9.5 and 9.6.

Identification code	9.1	9.2	9.3
Empirical formula	$C_{28}H_{22}F_2N_2O_4$	$C_{15}H_{14}NO_3$	$C_{28}H_{23}FN_2O_4$
Formula weight	488.47	256.27	470.48
Temperature/K	100.01(13)	100.00(10)	100.00(10)
Crystal system	triclinic	triclinic	monoclinic
Space group	P-1	P-1	P2$_1$/n
a/Å	7.2847(2)	7.36260(10)	14.5737(2)
b/Å	9.6461(3)	9.5766(2)	9.85540(10)
c/Å	16.6316(5)	10.0288(2)	15.9886(2)
α/Å	100.367(2)	101.832(2)	90
β/Å	98.058(2)	111.080(2)	105.5850(10)
γ/Å	97.290(2)	103.725(2)	90
Volume/Å	1124.29(6)	607.18(2)	2212.00(5)
Z	2	2	4
ρ_{calc}g/cm^3	1.443	1.402	1.413
μ/mm^{-1}	0.905	0.805	0.830
F(000)	508.0	270.0	984.0
Crystal size/mm^3	0.122 × 0.102 × 0.091	0.278 × 0.209 × 0.137	0.598 × 0.226 × 0.145
Radiation	Cu Kα (λ = 1.54184)	Cu Kα (λ = 1.54184)	Cu Kα (λ = 1.54184)
2Θ range for data collection/°	5.478 to 153.904	9.978 to 154.168	7.292 to 153.76
Index ranges	-6 ≤ h ≤ 9, -11 ≤ k ≤ 12, -20 ≤ l ≤ 19	-9 ≤ h ≤ 9, -11 ≤ k ≤ 7, -11 ≤ l ≤ 12	-18 ≤ h ≤ 18, -8 ≤ k ≤ 12, -18 ≤ l ≤ 19
Reflections collected	13387	6932	31321
Independent reflections	4400 [R$_{int}$ = 0.0254,	2387 [R$_{int}$ = 0.0349,	4525 [R$_{int}$ = 0.0437,

	$R_{sigma} = 0.0264]$	$R_{sigma} = 0.0335]$	$R_{sigma} = 0.0249]$
Data/restraints/parameters	4400/0/335	2387/0/178	4525/0/344
Goodness-of-fit on F^2	1.053	1.034	1.088
Final R indexes [I>=2 σ (I)]	$R_1 = 0.0369$, $wR_2 = 0.0964$	$R_1 = 0.0405$, $wR_2 = 0.1079$	$R_1 = 0.0431$, $wR_2 = 0.1109$
Final R indexes [all data]	$R_1 = 0.0413$, $wR_2 = 0.0995$	$R_1 = 0.0439$, $wR_2 = 0.1110$	$R_1 = 0.0457$, $wR_2 = 0.1126$
Largest diff. peak/hole / e	0.55/-0.28	0.19/-0.30	0.30/-0.32

Table 9.1 continued

Identification code	9.4	9.5	9.6
Empirical formula	$C_{12}H_{17}N_3O_3$	$C_{19}H_{16}N_4O_3$	$C_{22}H_{18}N_4$
Formula weight	251.28	348.36	338.40
Temperature/K	99.99(10)	100.00(10)	274(6)
Crystal system	triclinic	monoclinic	triclinic
Space group	P-1	P2$_1$/c	P-1
a/Å	7.5676(4)	9.68270(10)	6.9762(5)
b/Å	8.7128(5)	22.9931(2)	11.3568(8)
c/Å	10.0000(6)	7.24450(10)	12.5353(10)
α/Å	102.528(5)	90	115.896(7)
β/Å	108.486(5)	99.8270(10)	91.041(6)
γ/Å	93.767(5)	90	97.282(6)
Volume/Å	603.97(6)	1589.22(3)	883.20(12)
Z	2	4	2
ρ_{calc}g/cm^3	1.382	1.456	1.272
μ/mm^{-1}	0.835	0.836	0.607
F(000)	268.0	728.0	356.0
Crystal size/mm^3	0.513 × 0.272 × 0.187	0.24 × 0.122 × 0.059	0.21927 × 0.13047 × 0.05793
Radiation	Cu Kα (λ = 1.54184)	Cu Kα (λ = 1.54184)	Cu Kα (λ = 1.54184)
2Θ range for data collection/°	9.634 to 153.49	7.69 to 153.482	7.866 to 152.982
Index ranges	-9 ≤ h ≤ 9, -10 ≤ k ≤ 10, -12 ≤ l ≤ 7	-8 ≤ h ≤ 12, -27 ≤ k ≤ 29, -9 ≤ l ≤ 8	-6 ≤ h ≤ 8, -14 ≤ k ≤ 13, -14 ≤ l ≤ 15
Reflections collected	6115	11474	10833
Independent reflections	2367 [R$_{int}$ = 0.0322, R$_{sigma}$ = 0.0361]	3106 [R$_{int}$ = 0.0298, R$_{sigma}$ = 0.0259]	3463 [R$_{int}$ = 0.0313, R$_{sigma}$ = 0.0342]

Data/restraints/parameters	2367/0/173	3106/0/237	3463/2/245
Goodness-of-fit on F^2	1.054	1.030	1.072
Final R indexes [I>=2 σ (I)]	$R_1 = 0.0380$, $wR_2 = 0.1004$	$R_1 = 0.0348$, $wR_2 = 0.0909$	$R_1 = 0.0396$, $wR_2 = 0.1150$
Final R indexes [all data]	$R_1 = 0.0407$, $wR_2 = 0.1024$	$R_1 = 0.0386$, $wR_2 = 0.0939$	$R_1 = 0.0473$, $wR_2 = 0.1216$
Largest diff. peak/hole / e	0.29/-0.21	0.27/-0.21	0.14/-0.13

9.2 Dimethyl 2,2'-({3,3'-difluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate, 9.1
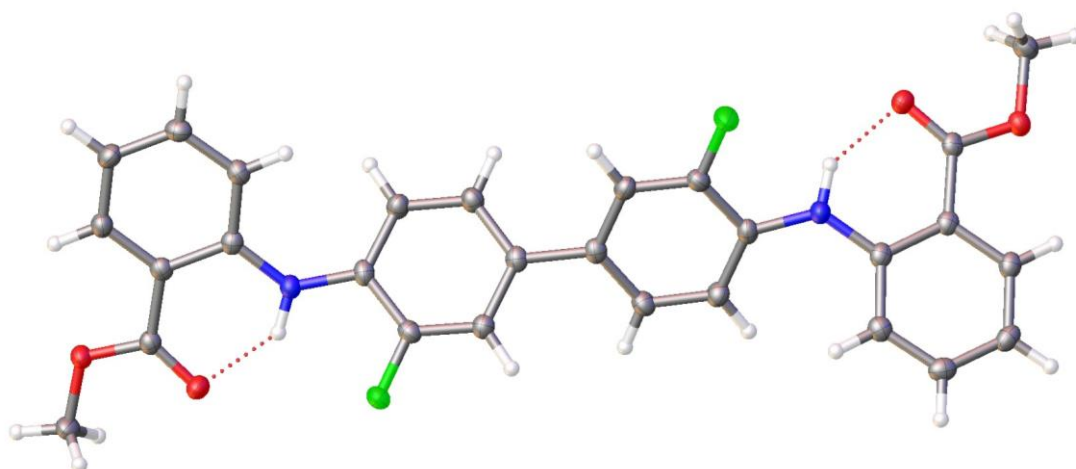
Figure 9.1: X-ray crystal structure of Dimethyl 2,2'-({3,3'-difluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate.

Table 9.1.2: Fractional Atomic Coordinates (×10⁴) and Equivalent Isotropic Displacement Parameters (Å²×10³) for 9.2. U_{eq} is defined as 1/3 of the trace of the orthogonalised U_{IJ} tensor.

Atom	x	y	z	U(eq)
F1	885.9(11)	3368.8(8)	3669.4(4)	26.45(19)
F2	1713.0(12)	4455.4(8)	8803.7(5)	31.5(2)
O1	2755.4(13)	5209.0(9)	2220.8(6)	25.7(2)
O2	3331.4(13)	6918.3(9)	1500.6(5)	23.2(2)
O3	3794.0(13)	2535.2(10)	10335.0(6)	24.4(2)
O4	3999.8(12)	869.4(10)	11107.3(5)	22.8(2)
N1	2638.5(15)	6018.6(12)	3814.7(7)	21.9(2)
N2	1971.1(16)	1742.2(12)	8795.6(7)	24.3(2)
C1	2579.0(17)	5483.7(14)	4534.4(8)	20.4(3)
C2	1660.1(17)	4089.2(14)	4450.1(8)	20.8(3)
C3	1495.4(17)	3412.8(13)	5102.3(8)	20.5(3)
C4	2289.7(17)	4128.8(14)	5908.8(8)	20.8(3)
C5	3275.8(17)	5503.1(14)	6006.2(8)	22.3(3)
C6	3433.6(17)	6167.0(14)	5338.4(8)	22.1(3)
C7	1924.7(17)	2192.5(14)	8044.8(8)	20.8(3)
C8	1852.1(18)	3636.8(14)	8063.6(8)	22.0(3)
C9	1950.0(17)	4274.1(14)	7394.7(8)	21.5(3)
C10	2116.0(16)	3449.7(14)	6631.3(8)	20.4(3)
C11	2150.6(18)	1997.0(14)	6590.5(8)	22.4(3)
C12	2067.2(18)	1373.1(14)	7277.7(8)	22.5(3)
C13	2826.6(16)	7403.7(13)	3695.6(8)	19.5(3)
C14	2998.9(16)	7663.9(13)	2891.5(8)	19.9(3)

Table 9.1.2: Fractional Atomic Coordinates (×10⁴) and Equivalent Isotropic Displacement Parameters (Å²×10³) for 9.2. U$_{eq}$ is defined as 1/3 of the trace of the orthogonalised U$_{IJ}$ tensor.

Atom	x	y	z	U(eq)
C15	3154.7(17)	9059.9(14)	2757.3(8)	22.0(3)
C16	3158.7(18)	10207.5(14)	3391.9(8)	24.1(3)
C17	2986.5(18)	9951.7(14)	4174.8(8)	23.9(3)
C18	2811.7(17)	8583.2(14)	4327.4(8)	22.3(3)
C19	1609.7(17)	382.6(13)	8952.4(8)	20.3(3)
C20	2296.8(16)	133.7(13)	9749.4(8)	19.5(3)
C21	1854.1(17)	-1225.4(14)	9922.9(8)	22.0(3)
C22	772.1(18)	-2328.8(14)	9334.1(9)	24.5(3)
C23	87.1(18)	-2078.0(14)	8556.4(9)	24.4(3)
C24	495.7(17)	-746.7(14)	8367.5(8)	22.5(3)
C25	3011.0(16)	6475.4(14)	2192.2(8)	20.0(3)
C26	3410.7(17)	1294.8(13)	10400.9(8)	19.6(3)
C27	3370(2)	5792.9(14)	800.1(8)	26.5(3)
C28	5054.8(19)	1999.6(15)	11760.2(8)	25.9(3)

Table 9.1.3: Anisotropic Displacement Parameters (Å²×10³) for 9.1. The Anisotropic displacement factor exponent takes the form: -2π²[h²a*²U$_{11}$+2hka*b*U$_{12}$+…].

Atom	U$_{11}$	U$_{22}$	U$_{33}$	U$_{23}$	U$_{13}$	U$_{12}$
F1	37.4(4)	23.5(4)	16.8(4)	2.9(3)	2.7(3)	2.2(3)
F2	50.3(5)	24.1(4)	20.4(4)	3.1(3)	8.8(3)	6.1(4)
O1	35.7(5)	19.0(5)	22.6(5)	4.7(4)	7.4(4)	1.1(4)
O2	32.3(5)	20.4(5)	17.9(4)	4.6(4)	7.3(4)	3.1(4)
O3	30.2(5)	20.3(5)	22.3(5)	5.4(4)	5.0(4)	0.4(4)
O4	24.3(4)	24.6(5)	19.6(5)	6.4(4)	2.3(3)	2.7(4)
N1	27.7(6)	20.6(6)	18.2(6)	5.5(4)	4.9(4)	3.3(4)
N2	34.4(6)	19.7(6)	17.9(6)	5.6(4)	2.6(5)	1.2(5)
C1	20.0(6)	24.3(6)	20.1(6)	8.7(5)	5.7(5)	6.7(5)
C2	22.3(6)	22.9(6)	17.3(6)	2.7(5)	3.0(5)	5.8(5)
C3	20.1(6)	20.9(6)	21.9(6)	6.0(5)	4.7(5)	4.2(5)
C4	18.7(6)	24.1(7)	21.7(6)	7.7(5)	4.4(5)	6.0(5)
C5	20.7(6)	25.5(7)	20.5(6)	6.0(5)	2.1(5)	2.8(5)
C6	20.4(6)	22.5(6)	23.5(7)	7.2(5)	3.1(5)	1.3(5)
C7	19.7(6)	23.7(7)	19.8(6)	7.9(5)	2.7(5)	2.0(5)
C8	23.6(6)	23.0(6)	18.1(6)	2.5(5)	3.1(5)	2.0(5)
C9	22.4(6)	19.4(6)	22.8(6)	6.3(5)	2.4(5)	2.5(5)
C10	16.4(6)	24.7(6)	20.8(6)	7.1(5)	2.7(5)	3.0(5)
C11	23.4(6)	25.5(7)	19.5(6)	4.7(5)	5.4(5)	5.7(5)
C12	24.6(6)	20.7(6)	23.7(7)	6.5(5)	4.3(5)	5.6(5)
C13	16.1(5)	21.8(6)	21.4(6)	6.5(5)	2.7(5)	3.1(5)

Table 9.1.3: Anisotropic Displacement Parameters ($Å^2 \times 10^3$) for 9.1. The Anisotropic displacement factor exponent takes the form: $-2\pi^2[h^2a^{*2}U_{11}+2hka^*b^*U_{12}+\ldots]$.

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
C14	17.1(6)	22.1(6)	19.9(6)	4.4(5)	2.4(5)	1.8(5)
C15	22.7(6)	23.0(6)	20.4(6)	6.6(5)	2.7(5)	2.2(5)
C16	26.9(6)	18.9(6)	26.2(7)	5.6(5)	3.1(5)	3.0(5)
C17	24.6(6)	22.8(7)	22.8(7)	0.7(5)	2.7(5)	4.8(5)
C18	22.3(6)	25.8(7)	19.4(6)	5.3(5)	3.9(5)	4.4(5)
C19	19.7(6)	21.4(6)	21.9(6)	6.1(5)	7.3(5)	4.8(5)
C20	18.7(6)	20.8(6)	21.3(6)	5.7(5)	7.6(5)	5.1(5)
C21	22.1(6)	23.1(6)	23.9(6)	8.2(5)	8.2(5)	5.6(5)
C22	25.1(6)	19.1(6)	32.1(7)	7.3(5)	11.0(5)	3.4(5)
C23	20.7(6)	23.2(7)	27.7(7)	0.8(5)	6.7(5)	1.5(5)
C24	20.2(6)	25.6(7)	22.5(6)	5.4(5)	4.6(5)	4.6(5)
C25	17.9(6)	22.9(6)	19.2(6)	5.7(5)	2.7(5)	1.4(5)
C26	18.6(6)	23.0(6)	19.8(6)	6.7(5)	7.9(5)	4.9(5)
C27	36.4(7)	23.0(7)	19.7(7)	2.2(5)	8.3(5)	1.9(5)
C28	24.9(6)	29.6(7)	21.4(7)	4.4(5)	1.2(5)	1.4(5)

Table 9.1.4: Bond Lengths for 9.1.

Atom	Atom	Length/Å	Atom	Atom	Length/Å
F1	C2	1.3629(15)	C7	C12	1.3994(18)
F2	C8	1.3617(15)	C8	C9	1.3704(18)
O1	C25	1.2226(15)	C9	C10	1.4011(18)
O2	C25	1.3367(15)	C10	C11	1.3943(18)
O2	C27	1.4487(15)	C11	C12	1.3892(17)
O3	C26	1.2217(15)	C13	C14	1.4238(17)
O4	C26	1.3452(15)	C13	C18	1.4053(18)
O4	C28	1.4441(16)	C14	C15	1.3973(17)
N1	C1	1.3901(16)	C14	C25	1.4807(17)
N1	C13	1.3782(16)	C15	C16	1.3846(19)
N2	C7	1.3915(16)	C16	C17	1.3876(18)
N2	C19	1.3827(16)	C17	C18	1.3821(18)
C1	C2	1.3980(18)	C19	C20	1.4228(18)
C1	C6	1.3983(18)	C19	C24	1.4016(19)
C2	C3	1.3741(17)	C20	C21	1.4004(17)
C3	C4	1.3995(18)	C20	C26	1.4741(18)
C4	C5	1.3957(18)	C21	C22	1.3816(19)
C4	C10	1.4811(17)	C22	C23	1.394(2)
C5	C6	1.3894(17)	C23	C24	1.3823(18)
C7	C8	1.3958(18)			

Table 9.1.5: Bond Angles for 9.1.

Atom	Atom	Atom	Angle/°	Atom	Atom	Atom	Angle/°
C25	O2	C27	115.09(10)	C11	C12	C7	120.62(12)
C26	O4	C28	114.21(10)	N1	C13	C14	119.00(11)
C13	N1	C1	130.59(11)	N1	C13	C18	123.19(11)
C19	N2	C7	130.02(12)	C18	C13	C14	117.80(11)
N1	C1	C2	117.33(11)	C13	C14	C25	120.99(11)
N1	C1	C6	126.78(12)	C15	C14	C13	119.67(12)
C2	C1	C6	115.79(11)	C15	C14	C25	119.34(11)
F1	C2	C1	117.31(11)	C16	C15	C14	121.51(12)
F1	C2	C3	118.64(11)	C15	C16	C17	118.74(12)
C3	C2	C1	124.05(12)	C18	C17	C16	121.27(12)
C2	C3	C4	119.62(12)	C17	C18	C13	121.00(12)
C3	C4	C10	121.41(12)	N2	C19	C20	119.16(12)
C5	C4	C3	117.44(11)	N2	C19	C24	122.19(12)
C5	C4	C10	121.15(12)	C24	C19	C20	118.54(11)
C6	C5	C4	122.06(12)	C19	C20	C26	120.99(11)
C5	C6	C1	120.93(12)	C21	C20	C19	119.27(12)
N2	C7	C8	116.88(12)	C21	C20	C26	119.70(11)
N2	C7	C12	126.99(12)	C22	C21	C20	121.26(12)
C8	C7	C12	115.99(11)	C21	C22	C23	119.31(12)
F2	C8	C7	117.03(11)	C24	C23	C22	120.75(12)
F2	C8	C9	118.72(11)	C23	C24	C19	120.85(12)
C9	C8	C7	124.24(12)	O1	C25	O2	121.82(11)
C8	C9	C10	119.41(12)	O1	C25	C14	125.12(11)
C9	C10	C4	120.18(12)	O2	C25	C14	113.05(11)
C11	C10	C4	122.27(12)	O3	C26	O4	121.18(12)
C11	C10	C9	117.54(11)	O3	C26	C20	125.42(11)
C12	C11	C10	122.18(12)	O4	C26	C20	113.40(10)

Table 9.1.6: Torsion Angles for 9.1.

A	B	C	D	Angle/°	A	B	C	D	Angle/°
F1	C2	C3	C4	179.60(10)	C10	C4	C5	C6	-179.09(11)
F2	C8	C9	C10	179.25(11)	C10	C11	C12	C7	0.64(19)
N1	C1	C2	F1	-0.39(16)	C12	C7	C8	F2	179.99(11)
N1	C1	C2	C3	179.62(11)	C12	C7	C8	C9	-1.39(19)
N1	C1	C6	C5	-179.43(12)	C13	N1	C1	C2	155.08(12)
N1	C13	C14	C15	178.79(11)	C13	N1	C1	C6	-28.7(2)
N1	C13	C14	C25	-1.06(17)	C13	C14	C15	C16	0.44(19)
N1	C13	C18	C17	-179.34(12)	C13	C14	C25	O1	5.16(19)
N2	C7	C8	F2	-4.05(17)	C13	C14	C25	O2	-175.02(10)
N2	C7	C8	C9	174.57(12)	C14	C13	C18	C17	-0.93(18)
N2	C7	C12	C11	-174.76(12)	C14	C15	C16	C17	-0.57(19)
N2	C19	C20	C21	176.86(11)	C15	C14	C25	O1	-174.69(12)
N2	C19	C20	C26	-1.10(17)	C15	C14	C25	O2	5.13(16)
N2	C19	C24	C23	-176.94(12)	C15	C16	C17	C18	-0.1(2)
C1	N1	C13	C14	174.94(12)	C16	C17	C18	C13	0.8(2)
C1	N1	C13	C18	-6.7(2)	C18	C13	C14	C15	0.31(17)
C1	C2	C3	C4	-0.42(19)	C18	C13	C14	C25	-179.54(11)
C2	C1	C6	C5	-3.12(18)	C19	N2	C7	C8	163.05(12)
C2	C3	C4	C5	-1.96(18)	C19	N2	C7	C12	-21.5(2)
C2	C3	C4	C10	178.88(11)	C19	C20	C21	C22	0.32(18)
C3	C4	C5	C6	1.75(18)	C19	C20	C26	O3	2.10(19)
C3	C4	C10	C9	-145.40(12)	C19	C20	C26	O4	-178.01(10)
C3	C4	C10	C11	36.02(18)	C20	C19	C24	C23	-0.65(18)
C4	C5	C6	C1	0.9(2)	C20	C21	C22	C23	-0.90(19)
C4	C10	C11	C12	177.22(12)	C21	C20	C26	O3	-175.85(12)
C5	C4	C10	C9	35.47(17)	C21	C20	C26	O4	4.04(16)
C5	C4	C10	C11	-143.11(13)	C21	C22	C23	C24	0.70(19)
C6	C1	C2	F1	-177.06(10)	C22	C23	C24	C19	0.08(19)
C6	C1	C2	C3	2.95(18)	C24	C19	C20	C21	0.45(17)
C7	N2	C19	C20	160.55(12)	C24	C19	C20	C26	-177.50(11)
C7	N2	C19	C24	-23.2(2)	C25	C14	C15	C16	-179.71(11)
C7	C8	C9	C10	0.7(2)	C26	C20	C21	C22	178.30(11)
C8	C7	C12	C11	0.73(18)	C27	O2	C25	O1	-0.61(17)
C8	C9	C10	C4	-177.89(11)	C27	O2	C25	C14	179.56(10)
C8	C9	C10	C11	0.76(18)	C28	O4	C26	O3	1.34(16)
C9	C10	C11	C12	-1.40(18)	C28	O4	C26	C20	-178.55(10)

Table 9.1.7: Hydrogen Atom Coordinates (Å×10⁴) and Isotropic Displacement Parameters (Å²×10³) for 9.1.

Atom	x	y	z	U(eq)
H1	2460(20)	5364(19)	3340(11)	36(5)
H2	2480(20)	2400(20)	9252(12)	38(5)
H3	845.99	2464.51	5006.34	25
H5	3857.43	6000.07	6546.14	27
H6	4131.91	7098.5	5429.21	26
H9	1905.79	5266.03	7448.14	26
H11	2233.93	1414.56	6075.72	27
H12	2107.19	380.89	7226.39	27
H15	3260.45	9225.45	2218.44	26
H16	3276.96	11152.15	3293.21	29
H17	2988.88	10732.21	4614.32	29
H18	2679.4	8437.76	4867.64	27
H21	2306.46	-1391.92	10455.77	26
H22	497.8	-3248.98	9457.96	29
H23	-669.24	-2830.56	8151.5	29
H24	15.22	-595.22	7834.43	27
H27A	3723.78	6214.75	339.24	40
H27B	2124.22	5211.88	627.56	40
H27C	4289.33	5190.6	959.56	40
H28A	6153.56	2441.92	11567.2	39
H28B	5467.46	1604.55	12246.08	39
H28C	4259.29	2719.28	11910.37	39

9.3 Dimethyl 2,2'-({3,3'-dimethoxy[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate, 9.2

Figure 9.2: X-ray crystals structure of Dimethyl 2,2'-({3,3'-dimethoxy[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate.

Table 9.2.1: Fractional Atomic Coordinates ($\times 10^4$) and Equivalent Isotropic Displacement Parameters ($\text{Å}^2 \times 10^3$) for 9.2. U_{eq} is defined as 1/3 of the trace of the orthogonalised U_{IJ} tensor.

Atom	x	y	z	U(eq)
O1	3655.4(14)	5728.4(9)	880.7(9)	23.2(2)
O2	7323.6(14)	4573.3(10)	4003.5(10)	21.9(2)
O3	8156.6(14)	4353.6(9)	6315.9(10)	22.2(2)
N1	6608.1(17)	7132.8(12)	3644.1(12)	19.6(2)
C1	6197.8(19)	7987.2(14)	2663.8(14)	18.4(3)
C2	7342.3(19)	9500.4(14)	3016.8(14)	20.7(3)
C3	6877.4(19)	10286.9(14)	1990.7(14)	21.0(3)
C4	5270.2(19)	9581.2(14)	547.8(13)	18.4(3)
C5	4170.5(19)	8025.4(14)	155.7(14)	20.1(3)
C6	4622.4(19)	7245.5(13)	1183.7(14)	19.1(3)
C7	7100.9(18)	7578.3(13)	5165.2(13)	17.0(3)
C8	7585.1(18)	6593.9(13)	6017.1(13)	17.5(3)
C9	7972.3(19)	7036.1(14)	7544.7(14)	20.6(3)
C10	7949(2)	8420.8(15)	8259.1(14)	22.9(3)
C11	7498(2)	9394.0(14)	7424.1(15)	22.0(3)
C12	7076.0(19)	8986.2(14)	5917.5(14)	19.9(3)
C13	7656.6(18)	5098.7(13)	5315.7(14)	17.9(3)
C14	2142(2)	4881.8(14)	-627.0(14)	24.1(3)
C15	8262(2)	2884.4(14)	5736.2(15)	25.1(3)

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
O1	31.8(5)	16.8(4)	15.0(5)	5.7(3)	4.6(4)	6.4(4)
O2	28.8(5)	19.2(4)	17.1(5)	5.5(3)	9.1(4)	8.6(4)
O3	29.3(5)	16.9(4)	19.6(5)	8.6(3)	7.7(4)	8.7(4)
N1	26.2(6)	18.0(5)	15.5(5)	7.3(4)	7.2(4)	9.9(4)
C1	21.8(6)	22.1(6)	16.2(6)	9.0(5)	9.3(5)	11.5(5)
C2	19.1(6)	23.7(6)	17.7(6)	8.5(5)	5.4(5)	7.0(5)
C3	19.9(6)	21.6(6)	21.5(6)	10.5(5)	8.1(5)	5.3(5)
C4	21.6(6)	21.8(6)	17.7(6)	9.2(5)	11.1(5)	10.9(5)
C5	24.8(6)	21.3(6)	14.5(6)	6.3(5)	7.2(5)	9.7(5)
C6	23.9(6)	18.3(6)	17.7(6)	6.7(5)	9.8(5)	9.4(5)
C7	15.7(5)	18.3(6)	16.4(6)	6.7(5)	6.2(5)	4.9(4)
C8	16.8(6)	18.1(6)	17.3(6)	6.8(5)	6.7(5)	5.3(5)
C9	21.8(6)	22.3(6)	18.3(6)	9.3(5)	8.0(5)	7.1(5)
C10	25.6(6)	25.6(6)	16.4(6)	4.8(5)	9.9(5)	6.8(5)
C11	22.6(6)	19.0(6)	23.2(7)	3.0(5)	10.6(5)	6.8(5)
C12	20.9(6)	18.4(6)	21.9(6)	8.7(5)	9.2(5)	7.4(5)
C13	16.1(6)	18.4(6)	17.8(6)	7.7(5)	5.7(5)	4.4(4)
C14	30.1(7)	20.9(6)	15.2(6)	4.1(5)	5.9(5)	5.8(5)
C15	27.1(7)	15.6(6)	28.7(7)	8.5(5)	7.3(6)	7.4(5)

Table 9.2.2: Anisotropic Displacement Parameters (Å2×103) for 9.2. The Anisotropic displacement factor exponent takes the form: -2π2[h2a*2U11+2hka*b*U12+…].

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
O1	31.8(5)	16.8(4)	15.0(5)	5.7(3)	4.6(4)	6.4(4)
O2	28.8(5)	19.2(4)	17.1(5)	5.5(3)	9.1(4)	8.6(4)
O3	29.3(5)	16.9(4)	19.6(5)	8.6(3)	7.7(4)	8.7(4)
N1	26.2(6)	18.0(5)	15.5(5)	7.3(4)	7.2(4)	9.9(4)
C1	21.8(6)	22.1(6)	16.2(6)	9.0(5)	9.3(5)	11.5(5)
C2	19.1(6)	23.7(6)	17.7(6)	8.5(5)	5.4(5)	7.0(5)
C3	19.9(6)	21.6(6)	21.5(6)	10.5(5)	8.1(5)	5.3(5)
C4	21.6(6)	21.8(6)	17.7(6)	9.2(5)	11.1(5)	10.9(5)
C5	24.8(6)	21.3(6)	14.5(6)	6.3(5)	7.2(5)	9.7(5)
C6	23.9(6)	18.3(6)	17.7(6)	6.7(5)	9.8(5)	9.4(5)
C7	15.7(5)	18.3(6)	16.4(6)	6.7(5)	6.2(5)	4.9(4)
C8	16.8(6)	18.1(6)	17.3(6)	6.8(5)	6.7(5)	5.3(5)
C9	21.8(6)	22.3(6)	18.3(6)	9.3(5)	8.0(5)	7.1(5)
C10	25.6(6)	25.6(6)	16.4(6)	4.8(5)	9.9(5)	6.8(5)
C11	22.6(6)	19.0(6)	23.2(7)	3.0(5)	10.6(5)	6.8(5)
C12	20.9(6)	18.4(6)	21.9(6)	8.7(5)	9.2(5)	7.4(5)
C13	16.1(6)	18.4(6)	17.8(6)	7.7(5)	5.7(5)	4.4(4)
C14	30.1(7)	20.9(6)	15.2(6)	4.1(5)	5.9(5)	5.8(5)
C15	27.1(7)	15.6(6)	28.7(7)	8.5(5)	7.3(6)	7.4(5)

Table 9.2.3: Bond Lengths for 9.2.

Atom	Atom	Length/Å	Atom	Atom	Length/Å
O1	C6	1.3721(14)	C4	C4[1]	1.486(2)
O1	C14	1.4288(15)	C4	C5	1.4101(17)
O2	C13	1.2177(15)	C5	C6	1.3851(17)
O3	C13	1.3430(14)	C7	C8	1.4200(16)
O3	C15	1.4394(15)	C7	C12	1.4114(17)
N1	C1	1.4003(15)	C8	C9	1.4036(17)
N1	C7	1.3792(16)	C8	C13	1.4822(16)
C1	C2	1.3896(18)	C9	C10	1.3791(18)
C1	C6	1.4112(17)	C10	C11	1.3954(18)
C2	C3	1.3894(17)	C11	C12	1.3764(18)
C3	C4	1.3924(17)			

Table 9.2.4: Bond Angles for 9.2.

Atom	Atom	Atom	Angle/°	Atom	Atom	Atom	Angle/°
C6	O1	C14	117.31(9)	C5	C6	C1	120.93(11)
C13	O3	C15	115.71(10)	N1	C7	C8	120.31(11)
C7	N1	C1	127.51(10)	N1	C7	C12	121.92(11)
N1	C1	C6	118.31(11)	C12	C7	C8	117.73(11)
C2	C1	N1	124.16(11)	C7	C8	C13	121.17(11)
C2	C1	C6	117.36(11)	C9	C8	C7	119.36(11)
C3	C2	C1	121.63(11)	C9	C8	C13	119.46(11)
C2	C3	C4	121.35(12)	C10	C9	C8	121.92(11)
C3	C4	C4[1]	121.83(14)	C9	C10	C11	118.60(12)
C3	C4	C5	117.31(11)	C12	C11	C10	121.02(12)
C5	C4	C4[1]	120.86(14)	C11	C12	C7	121.35(11)
C6	C5	C4	121.28(11)	O2	C13	O3	122.35(11)
O1	C6	C1	114.26(10)	O2	C13	C8	125.97(11)
O1	C6	C5	124.80(11)	O3	C13	C8	111.68(10)

Table 9.2.5: Torsion Angles for 9.2.

A	B	C	D	Angle/°	A	B	C	D	Angle/°
N1	C1	C2	C3	179.23(11)	C7	N1	C1	C2	43.24(19)
N1	C1	C6	O1	0.07(16)	C7	N1	C1	C6	-141.62(13)
N1	C1	C6	C5	-178.90(11)	C7	C8	C9	C10	-1.69(19)
N1	C7	C8	C9	-176.57(11)	C7	C8	C13	O2	-0.1(2)
N1	C7	C8	C13	2.47(18)	C7	C8	C13	O3	179.53(10)
N1	C7	C12	C11	177.61(12)	C8	C7	C12	C11	-0.24(18)
C1	N1	C7	C8	-176.52(11)	C8	C9	C10	C11	0.91(19)
C1	N1	C7	C12	5.68(19)	C9	C8	C13	O2	178.91(12)
C1	C2	C3	C4	-1.5(2)	C9	C8	C13	O3	-1.43(16)
C2	C1	C6	O1	175.54(11)	C9	C10	C11	C12	0.22(19)
C2	C1	C6	C5	-3.42(18)	C10	C11	C12	C7	-0.55(19)
C2	C3	C4	C4^1	178.67(13)	C12	C7	C8	C9	1.32(17)
C2	C3	C4	C5	-1.72(18)	C12	C7	C8	C13	-179.64(10)
C3	C4	C5	C6	2.31(18)	C13	C8	C9	C10	179.25(11)
C4^1	C4	C5	C6	-178.07(13)	C14	O1	C6	C1	-175.77(10)
C4	C5	C6	O1	-178.58(11)	C14	O1	C6	C5	3.15(18)
C4	C5	C6	C1	0.27(19)	C15	O3	C13	O2	-0.40(17)
C6	C1	C2	C3	4.04(18)	C15	O3	C13	C8	179.92(10)

Table 9.2.6: Hydrogen Atom Coordinates (Å×10⁴) and Isotropic Displacement Parameters (Å²×10³) for 9.2.

Atom	x	y	z	U(eq)
H1	6270(30)	6150(20)	3213(19)	32(4)
H2	8470.53	10009.67	3983.7	25
H3	7672.46	11327.53	2279.43	25
H5	3098.73	7502.03	-832.82	24
H9	8258.59	6361.81	8102.74	25
H10	8234.84	8706.11	9297.78	28
H11	7481.63	10352.91	7902.8	26
H12	6761.48	9666.35	5373.02	24
H14A	2748.64	5032.36	-1333.23	36
H14B	950.62	5230.99	-853.04	36
H14C	1690.32	3801.52	-728.74	36
H15A	7016.49	2295.03	4798.19	38
H15B	8351.56	2349.96	6480.31	38
H15C	9488.68	3008.07	5535.51	38

9.4 Dimethyl 2,2'-({3-fluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate 9.3
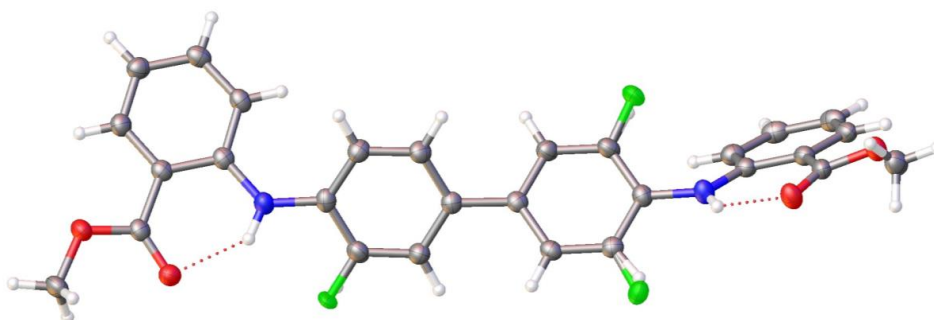
Figure 9.3: X-ray crystal structure of 9.4. Dimethyl 2,2'-({3-fluoro[1,1'-biphenyl]-4,4'-diyl}diimino)bis-benzoate.

Table 9.3.1: Fractional Atomic Coordinates (×10⁴) and Equivalent Isotropic Displacement Parameters (Å²×10³) for 9.3. U_{eq} is defined as 1/3 of the trace of the orthogonalised U_{IJ} tensor.

Atom	x	y	z	U(eq)
F1A	3087.2(13)	2483.7(17)	6023.3(13)	28.8(4)
F1B	3884.7(18)	10804(2)	6547.2(17)	30.7(5)
F1C	6452(4)	7817(6)	7488(4)	29.6(11)
O1	1245.6(7)	43.7(11)	6227.0(7)	29.8(2)
O2	-183.7(8)	-837.9(10)	6183.5(7)	31.9(3)
O3	7102.2(8)	11750.7(10)	8566.9(7)	31.8(3)
O4	7368.1(8)	13964.9(10)	8413.9(7)	30.3(2)
N1	1349.1(9)	2767.7(12)	6210.5(8)	27.5(3)
N2	5878.9(9)	10505.9(13)	7232.8(9)	28.0(3)
C1	1973.3(10)	3873.3(15)	6362.3(9)	25.8(3)
C2	2865.8(11)	3703.7(15)	6205.5(10)	29.4(3)
C3	3528.8(11)	4732.4(15)	6318.9(10)	27.3(3)
C4	3317.5(10)	6013.0(15)	6589.5(9)	24.4(3)
C5	2448.1(11)	6175.2(15)	6788.2(9)	27.2(3)
C6	1795.7(11)	5128.8(16)	6691.3(9)	27.5(3)
C7	5246.3(10)	9408.0(14)	6977.5(9)	24.7(3)
C8	4268.7(11)	9577.6(15)	6626.8(10)	28.1(3)
C9	3647.9(11)	8486.1(15)	6480.5(10)	27.1(3)
C10	3989.3(10)	7166.9(14)	6685.9(9)	24.1(3)
C11	4972.2(10)	6986.2(14)	7015.6(9)	24.0(3)
C12	5585.6(11)	8085.1(15)	7156.0(9)	25.1(3)
C13	364.2(10)	2757.1(15)	5990.2(9)	24.9(3)
C14	-124.7(10)	1502.3(14)	5942.5(9)	24.2(3)
C15	-1123.9(11)	1491.5(15)	5717.9(10)	27.6(3)
C16	-1650.3(11)	2671.1(16)	5539.3(10)	29.7(3)
C17	-1172.0(11)	3898.7(15)	5570.0(9)	28.5(3)

Table 9.3.1: Fractional Atomic Coordinates (×10⁴) and Equivalent Isotropic Displacement Parameters (Å²×10³) for 9.3. U_{eq} is defined as 1/3 of the trace of the orthogonalised U_{IJ} tensor.

Atom	x	y	z	U(eq)
C18	-187.2(11)	3942.8(15)	5788.5(9)	27.1(3)
C19	5911.6(10)	11679.1(14)	6776.2(10)	25.4(3)
C20	6454.4(10)	12812.1(14)	7185.5(10)	25.5(3)
C21	6495.8(11)	13981.8(15)	6694.7(10)	28.6(3)
C22	6029.0(11)	14060.0(17)	5826.6(11)	32.6(3)
C23	5499.2(11)	12948.9(17)	5424.6(10)	32.0(3)
C24	5437.5(11)	11787.9(16)	5888.2(10)	28.6(3)
C25	393.3(10)	201.3(14)	6129.8(9)	25.3(3)
C26	6990.3(10)	12762.8(14)	8109.9(10)	26.1(3)
C27	273.1(12)	-2144.4(15)	6372.7(11)	33.9(4)
C28	7928.6(12)	13959.1(17)	9304.5(11)	34.5(4)

Table 9.3.2: Anisotropic Displacement Parameters (Å²×10³) for 9.3. The Anisotropic displacement factor exponent takes the form: $-2\pi^2[h^2a^{*2}U_{11}+2hka^*b^*U_{12}+\ldots]$.

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
F1A	26.1(9)	15.0(8)	47.0(11)	-11.6(8)	12.6(8)	0.3(7)
F1B	32.2(13)	17.5(11)	41.7(14)	2.7(10)	8.8(11)	3.9(9)
F1C	24(2)	30(2)	34(2)	4.0(18)	6.3(18)	-3.7(18)
O1	26.5(5)	26.4(5)	35.0(6)	-2.4(4)	6.0(4)	1.0(4)
O2	31.5(6)	20.8(5)	45.0(7)	0.2(4)	12.8(5)	0.6(4)
O3	35.5(6)	24.2(5)	31.3(6)	3.4(4)	1.3(5)	-2.0(4)
O4	32.3(6)	24.6(5)	32.8(6)	-1.8(4)	6.5(5)	-3.3(4)
N1	26.5(6)	22.0(6)	32.2(7)	1.2(5)	5.0(5)	-1.3(5)
N2	29.9(7)	23.7(6)	27.0(6)	2.3(5)	1.7(5)	-2.6(5)
C1	27.8(7)	25.1(7)	22.3(7)	2.8(5)	3.3(6)	-3.1(6)
C2	31.0(8)	22.8(7)	32.8(8)	1.1(6)	6.0(6)	1.9(6)
C3	27.0(7)	26.0(7)	29.5(7)	0.1(6)	8.4(6)	0.5(6)
C4	26.4(7)	25.8(7)	20.5(7)	1.0(5)	5.2(5)	-0.7(5)
C5	29.2(7)	27.7(7)	24.0(7)	-4.7(6)	6.0(6)	-0.8(6)
C6	27.0(7)	32.1(8)	23.7(7)	-3.1(6)	7.5(6)	-2.1(6)
C7	28.5(7)	23.4(7)	22.2(7)	-0.7(5)	6.7(6)	-1.8(6)
C8	30.2(8)	22.1(7)	30.2(8)	-1.7(6)	5.3(6)	4.0(6)
C9	25.4(7)	26.4(7)	27.9(7)	-2.4(6)	4.4(6)	1.4(6)
C10	27.8(7)	24.4(7)	20.5(7)	-2.5(5)	7.4(5)	-0.2(6)
C11	28.6(7)	22.2(7)	21.8(7)	-0.2(5)	7.9(6)	2.6(5)
C12	26.5(7)	27.5(7)	22.1(7)	0.8(5)	7.9(6)	1.2(6)
C13	29.0(7)	25.5(7)	19.8(6)	-1.6(5)	5.8(5)	-0.7(6)
C14	27.9(7)	23.9(7)	20.5(7)	-1.6(5)	5.8(5)	0.6(6)
C15	28.8(7)	26.3(7)	27.1(7)	-3.0(6)	6.2(6)	-2.2(6)

Table 9.3.2: Anisotropic Displacement Parameters ($Å^2×10^3$) for 9.3. The Anisotropic displacement factor exponent takes the form: $-2\pi^2[h^2a^{*2}U_{11}+2hka^*b^*U_{12}+\ldots]$.

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
C16	26.8(7)	31.6(8)	29.1(8)	-3.3(6)	4.7(6)	1.6(6)
C17	34.3(8)	25.9(7)	23.4(7)	-1.2(6)	4.5(6)	4.8(6)
C18	34.2(8)	22.8(7)	23.0(7)	-0.7(5)	5.4(6)	-1.1(6)
C19	24.9(7)	23.8(7)	28.7(7)	1.6(6)	9.1(6)	2.9(5)
C20	24.7(7)	23.3(7)	29.5(7)	2.1(6)	9.0(6)	1.9(5)
C21	26.8(7)	23.8(7)	37.1(8)	4.1(6)	11.6(6)	2.0(6)
C22	32.7(8)	32.6(8)	35.8(8)	11.7(7)	15.0(7)	6.3(6)
C23	31.4(8)	39.9(9)	26.0(7)	6.0(6)	10.0(6)	8.0(7)
C24	27.7(7)	31.3(8)	27.2(7)	-1.4(6)	7.9(6)	2.6(6)
C25	29.7(7)	23.3(7)	22.4(7)	-3.7(5)	6.4(6)	-1.8(6)
C26	23.8(7)	21.6(7)	33.5(8)	0.0(6)	8.7(6)	-0.7(5)
C27	38.0(9)	21.3(7)	44.8(9)	1.7(6)	14.8(7)	2.9(6)
C28	33.5(8)	31.3(8)	35.1(8)	-5.0(7)	2.8(7)	-2.1(6)

Table 9.3.3: Bond Lengths for 9.3.

Atom	Atom	Length/Å	Atom	Atom	Length/Å
F1A	C2	1.298(2)	C7	C8	1.394(2)
F1B	C8	1.324(3)	C7	C12	1.396(2)
F1C	C12	1.259(6)	C8	C9	1.384(2)
O1	C25	1.2197(18)	C9	C10	1.399(2)
O2	C25	1.3424(18)	C10	C11	1.399(2)
O2	C27	1.4437(18)	C11	C12	1.384(2)
O3	C26	1.2212(18)	C13	C14	1.419(2)
O4	C26	1.3413(18)	C13	C18	1.406(2)
O4	C28	1.4387(19)	C14	C15	1.403(2)
N1	C1	1.3982(19)	C14	C25	1.477(2)
N1	C13	1.3830(19)	C15	C16	1.380(2)
N2	C7	1.4089(19)	C16	C17	1.391(2)
N2	C19	1.3751(19)	C17	C18	1.384(2)
C1	C2	1.399(2)	C19	C20	1.422(2)
C1	C6	1.396(2)	C19	C24	1.406(2)
C2	C3	1.378(2)	C20	C21	1.405(2)
C3	C4	1.395(2)	C20	C26	1.475(2)
C4	C5	1.396(2)	C21	C22	1.374(2)
C4	C10	1.481(2)	C22	C23	1.393(2)
C5	C6	1.383(2)	C23	C24	1.379(2)

Table 9.3.4: Bond Angles for 9.3.

Atom	Atom	Atom	Angle/°	Atom	Atom	Atom	Angle/°
C25	O2	C27	115.62(12)	F1C	C12	C11	115.8(3)
C26	O4	C28	115.07(12)	C11	C12	C7	121.32(14)
C13	N1	C1	129.23(13)	N1	C13	C14	119.47(13)
C19	N2	C7	127.25(13)	N1	C13	C18	122.84(13)
N1	C1	C2	117.95(13)	C18	C13	C14	117.67(13)
C6	C1	N1	125.34(14)	C13	C14	C25	121.58(13)
C6	C1	C2	116.67(13)	C15	C14	C13	119.47(13)
F1A	C2	C1	116.73(15)	C15	C14	C25	118.95(13)
F1A	C2	C3	120.20(16)	C16	C15	C14	121.85(14)
C3	C2	C1	122.76(14)	C15	C16	C17	118.70(14)
C2	C3	C4	120.06(14)	C18	C17	C16	120.83(14)
C3	C4	C5	117.62(13)	C17	C18	C13	121.47(14)
C3	C4	C10	122.15(13)	N2	C19	C20	120.68(13)
C5	C4	C10	120.22(13)	N2	C19	C24	121.50(14)
C6	C5	C4	121.87(14)	C24	C19	C20	117.80(13)
C5	C6	C1	120.78(14)	C19	C20	C26	120.96(13)
C8	C7	N2	122.89(13)	C21	C20	C19	119.29(14)
C8	C7	C12	117.49(13)	C21	C20	C26	119.73(13)
C12	C7	N2	119.32(13)	C22	C21	C20	121.75(15)
F1B	C8	C7	120.54(17)	C21	C22	C23	118.98(14)
F1B	C8	C9	116.94(17)	C24	C23	C22	120.80(15)
C9	C8	C7	121.72(13)	C23	C24	C19	121.37(15)
C8	C9	C10	120.49(14)	O1	C25	O2	121.92(13)
C9	C10	C4	120.37(13)	O1	C25	C14	125.72(13)
C11	C10	C4	121.52(13)	O2	C25	C14	112.36(12)
C11	C10	C9	118.05(13)	O3	C26	O4	121.93(14)
C12	C11	C10	120.87(13)	O3	C26	C20	125.17(13)
F1C	C12	C7	122.8(3)	O4	C26	C20	112.90(12)

Table 9.3.5: Torsion Angles for 9.3.

A	B	C	D	Angle/°	A	B	C	D	Angle/°
F1A	C2	C3	C4	-174.34(17)	C10	C4	C5	C6	178.81(13)
F1B	C8	C9	C10	-169.62(18)	C10	C11	C12	F1C	177.2(3)
N1	C1	C2	F1A	-7.5(2)	C10	C11	C12	C7	0.2(2)
N1	C1	C2	C3	178.84(14)	C12	C7	C8	F1B	171.20(18)
N1	C1	C6	C5	-177.64(14)	C12	C7	C8	C9	1.7(2)
N1	C13	C14	C15	-179.92(13)	C13	N1	C1	C2	-152.72(15)
N1	C13	C14	C25	0.5(2)	C13	N1	C1	C6	29.9(2)
N1	C13	C18	C17	179.97(13)	C13	C14	C15	C16	0.1(2)
N2	C7	C8	F1B	-2.5(3)	C13	C14	C25	O1	-10.0(2)
N2	C7	C8	C9	-171.92(14)	C13	C14	C25	O2	170.18(12)
N2	C7	C12	F1C	-4.7(4)	C14	C13	C18	C17	1.5(2)
N2	C7	C12	C11	171.93(13)	C14	C15	C16	C17	1.1(2)
N2	C19	C20	C21	178.15(13)	C15	C14	C25	O1	170.42(14)
N2	C19	C20	C26	-0.4(2)	C15	C14	C25	O2	-9.40(19)
N2	C19	C24	C23	-177.67(14)	C15	C16	C17	C18	-1.0(2)
C1	N1	C13	C14	-174.40(14)	C16	C17	C18	C13	-0.3(2)
C1	N1	C13	C18	7.1(2)	C18	C13	C14	C15	-1.4(2)
C1	C2	C3	C4	-0.9(2)	C18	C13	C14	C25	179.07(13)
C2	C1	C6	C5	4.9(2)	C19	N2	C7	C8	-45.9(2)
C2	C3	C4	C5	3.9(2)	C19	N2	C7	C12	140.60(15)
C2	C3	C4	C10	-177.43(14)	C19	C20	C21	C22	-0.3(2)
C3	C4	C5	C6	-2.5(2)	C19	C20	C26	O3	8.2(2)
C3	C4	C10	C9	144.25(15)	C19	C20	C26	O4	-172.63(13)
C3	C4	C10	C11	-38.4(2)	C20	C19	C24	C23	0.4(2)
C4	C5	C6	C1	-2.0(2)	C20	C21	C22	C23	0.1(2)
C4	C10	C11	C12	-175.68(13)	C21	C20	C26	O3	-170.38(14)
C5	C4	C10	C9	-37.1(2)	C21	C20	C26	O4	8.82(19)
C5	C4	C10	C11	140.24(14)	C21	C22	C23	C24	0.3(2)
C6	C1	C2	F1A	170.13(16)	C22	C23	C24	C19	-0.6(2)
C6	C1	C2	C3	-3.5(2)	C24	C19	C20	C21	0.0(2)
C7	N2	C19	C20	167.05(14)	C24	C19	C20	C26	-178.53(13)
C7	N2	C19	C24	-14.9(2)	C25	C14	C15	C16	179.68(13)
C7	C8	C9	C10	0.2(2)	C26	C20	C21	C22	178.25(14)
C8	C7	C12	F1C	-178.6(3)	C27	O2	C25	O1	0.7(2)

Table 9.3.5: Torsion Angles for 9.3.

A	B	C	D	Angle/°	A	B	C	D	Angle/°
C8	C7	C12	C11	-2.0(2)	C27	O2	C25	C14	-179.50(12)
C8	C9	C10	C4	175.49(13)	C28	O4	C26	O3	1.5(2)
C8	C9	C10	C11	-1.9(2)	C28	O4	C26	C20	-177.76(12)
C9	C10	C11	C12	1.7(2)					

Table 9.3.6: Hydrogen Atom Coordinates (Å×10⁴) and Isotropic Displacement Parameters (Å²×10³) for 9.3.

Atom	x	y	z	U(eq)
H1	1621.78	1964.05	6262.64	33
H2	6293(14)	10433(19)	7751(13)	35(5)
H00R	3021.42	2843.53	6012.9	35
H3	4130.03	4569.16	6212.66	33
H5	2300.31	7028.65	6995.8	33
H6	1219.75	5266.14	6850.57	33
H00Q	4021.91	10465.26	6484.49	34
H9	2985.35	8633.84	6238.75	33
H11	5222.11	6097.18	7144.82	29
H007	6280(40)	8050(60)	7470(40)	130(20)
H15	-1447.15	648.32	5687.82	33
H16	-2326.24	2645.16	5398.24	36
H17	-1525.81	4717.31	5439.35	34
H18	123.3	4793.42	5802.48	33
H21	6856.25	14738.93	6970.39	34
H22	6066.86	14859.24	5505.32	39
H23	5176.59	12991.55	4823.95	38
H24	5066.94	11046.45	5601.75	34
H27A	645.23	-2330.93	5958.16	51
H27B	-212.54	-2849.49	6324.72	51
H27C	696.58	-2139.75	6963.68	51
H28A	8444.26	13294.43	9377.11	52
H28B	7523.19	13715.77	9681.09	52
H28C	8200.25	14863.19	9461.89	52

Table 9.3.7: Atomic Occupancy for 9.3.

Atom	Occupancy	Atom	Occupancy	Atom	Occupancy
F1A	0.45	F1B	0.35	F1C	0.2
H00R	0.55	H00Q	0.65	H007	0.8

9.5 3-hydroxy-5,6,7,8-tetrahydro-4*H*-quinoxaline-2-(*N*-2-methoxyethyl)carboxamide, 9.4
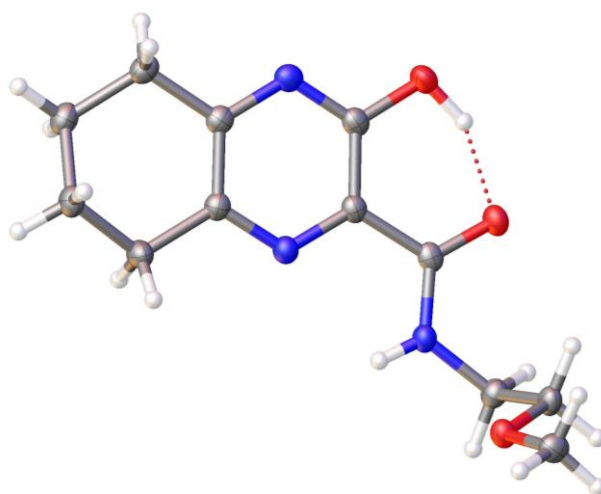
Figure 9.4: X-ray crystal structure of 3-hydroxy-5,6,7,8-tetrahydro-4*H*-quinoxaline-2-(*N*-2-methoxyethyl)carboxamide.

Table 9.4.1: Fractional Atomic Coordinates (×10^4) and Equivalent Isotropic Displacement Parameters (Å^2×10^3) for 9.4. Ueq is defined as 1/3 of the trace of the orthogonalised UIJ tensor.

Atom	x	y	z	U(eq)
O1	8123.6(12)	1918.0(10)	5016.4(10)	26.0(2)
O2	7595.1(12)	1680.3(10)	2299.4(9)	23.5(2)
O3	3371.6(11)	3063.6(10)	-993.8(8)	21.6(2)
N1	7960.0(13)	4304.6(12)	6415.1(10)	21.9(2)
N2	7234.8(13)	5620.3(12)	4025.6(10)	20.1(2)
N3	6924.8(13)	3807.6(12)	1357.9(10)	20.4(2)
C1	7861.6(15)	3439.7(14)	5114.1(13)	20.6(3)
C2	7489.1(15)	4102.1(14)	3905.0(12)	19.2(3)
C3	7361.6(15)	6510.5(14)	5334.6(12)	20.2(3)
C4	7139.8(18)	8235.1(15)	5457.8(13)	25.0(3)
C5	6843.9(17)	9051.7(15)	6869.3(13)	24.6(3)
C6	8175.8(17)	8588.0(15)	8170.8(13)	25.6(3)
C7	7840.6(17)	6801.2(15)	8011.3(12)	23.4(3)
C8	7720.9(15)	5832.3(14)	6531.0(12)	20.3(3)
C9	7340.4(15)	3103.7(14)	2446.9(12)	19.5(3)
C10	6681.7(16)	2975.3(14)	-131.4(12)	22.7(3)
C11	4740.9(16)	2016.2(14)	-932.3(12)	22.5(3)
C12	1515.0(17)	2250.7(15)	-1830.1(13)	25.6(3)

Table 9.4.2: Anisotropic Displacement Parameters ($\text{Å}^2\times10^3$) for 9.4. The Anisotropic displacement factor exponent takes the form: $-2\pi^2[h^2a^{*2}U_{11}+2hka^*b^*U_{12}+\ldots]$.

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
O1	30.1(5)	19.1(5)	26.7(5)	7.9(3)	4.9(4)	6.3(3)
O2	23.8(4)	18.3(4)	27.6(4)	4.3(3)	8.0(3)	5.7(3)
O3	19.7(4)	19.1(4)	23.6(4)	4.4(3)	4.4(3)	4.6(3)
N1	18.2(5)	23.6(5)	22.3(5)	6.7(4)	4.2(4)	2.5(4)
N2	19.1(5)	19.2(5)	21.3(5)	4.1(4)	6.5(4)	3.7(4)
N3	21.0(5)	17.8(5)	20.8(5)	2.9(4)	6.1(4)	4.2(4)
C1	15.4(5)	19.9(6)	24.1(6)	5.5(4)	3.6(4)	1.9(4)
C2	15.1(5)	18.4(6)	22.3(6)	4.1(4)	4.5(4)	2.0(4)
C3	17.2(5)	21.1(6)	21.1(5)	3.9(4)	5.9(4)	2.9(4)
C4	32.1(6)	20.8(6)	24.1(6)	5.7(5)	11.4(5)	7.0(5)
C5	25.5(6)	22.0(6)	26.2(6)	2.7(5)	10.6(5)	5.4(5)
C6	23.6(6)	27.5(7)	23.3(6)	1.2(5)	8.2(5)	2.8(5)
C7	20.9(6)	28.0(7)	20.0(6)	4.3(5)	6.5(4)	3.1(5)
C8	14.9(5)	23.3(6)	21.9(6)	4.9(4)	5.6(4)	2.3(4)
C9	13.3(5)	18.3(6)	25.4(6)	3.9(4)	5.6(4)	1.4(4)
C10	22.8(6)	22.7(6)	21.8(6)	2.9(5)	8.1(4)	5.3(5)
C11	24.4(6)	19.1(6)	21.7(5)	2.2(4)	5.9(4)	6.7(5)
C12	21.8(6)	26.3(7)	25.4(6)	7.1(5)	3.6(5)	1.4(5)

Table 9.4.3: Bond Lengths for 9.4.

Atom	Atom	Length/Å	Atom	Atom	Length/Å
O1	C1	1.3406(14)	C1	C2	1.4080(17)
O2	C9	1.2518(14)	C2	C9	1.4927(16)
O3	C11	1.4207(14)	C3	C4	1.5057(16)
O3	C12	1.4291(14)	C3	C8	1.4057(17)
N1	C1	1.3312(16)	C4	C5	1.5279(16)
N1	C8	1.3401(16)	C5	C6	1.5241(17)
N2	C2	1.3343(15)	C6	C7	1.5259(18)
N2	C3	1.3396(15)	C7	C8	1.5082(16)
N3	C9	1.3263(16)	C10	C11	1.5164(16)
N3	C10	1.4553(15)			

Table 9.4.4: Bond Angles for 9.4.

Atom	Atom	Atom	Angle/°	Atom	Atom	Atom	Angle/°
C11	O3	C12	111.70(9)	C3	C4	C5	113.44(10)
C1	N1	C8	117.38(10)	C6	C5	C4	111.02(10)
C2	N2	C3	118.29(10)	C5	C6	C7	111.26(10)
C9	N3	C10	122.68(10)	C8	C7	C6	112.95(10)
O1	C1	C2	122.11(10)	N1	C8	C3	122.05(10)
N1	C1	O1	116.78(10)	N1	C8	C7	116.89(10)
N1	C1	C2	121.11(11)	C3	C8	C7	121.06(11)
N2	C2	C1	121.18(10)	O2	C9	N3	123.46(11)
N2	C2	C9	118.48(10)	O2	C9	C2	120.13(10)
C1	C2	C9	120.34(11)	N3	C9	C2	116.41(10)
N2	C3	C4	117.40(10)	N3	C10	C11	112.73(9)
N2	C3	C8	119.98(11)	O3	C11	C10	108.83(9)
C8	C3	C4	122.61(10)				

Table 9.4.5: Torsion Angles for 9.4.

A	B	C	D	Angle/°	A	B	C	D	Angle/°
O1	C1	C2	N2	-179.91(9)	C3	N2	C2	C1	0.67(16)
O1	C1	C2	C9	1.05(17)	C3	N2	C2	C9	179.72(9)
N1	C1	C2	N2	0.44(17)	C3	C4	C5	C6	43.28(14)
N1	C1	C2	C9	-178.60(9)	C4	C3	C8	N1	-178.38(10)
N2	C2	C9	O2	179.08(10)	C4	C3	C8	C7	2.06(17)
N2	C2	C9	N3	-1.23(15)	C4	C5	C6	C7	-60.89(13)
N2	C3	C4	C5	166.46(10)	C5	C6	C7	C8	47.64(13)
N2	C3	C8	N1	0.51(17)	C6	C7	C8	N1	161.77(10)
N2	C3	C8	C7	-179.04(9)	C6	C7	C8	C3	-18.65(15)
N3	C10	C11	O3	60.69(12)	C8	N1	C1	O1	179.30(9)
C1	N1	C8	C3	0.58(16)	C8	N1	C1	C2	-1.04(16)
C1	N1	C8	C7	-179.84(9)	C8	C3	C4	C5	-14.62(16)
C1	C2	C9	O2	-1.87(16)	C9	N3	C10	C11	81.38(13)
C1	C2	C9	N3	177.83(10)	C10	N3	C9	O2	1.43(17)
C2	N2	C3	C4	177.83(10)	C10	N3	C9	C2	-178.26(9)
C2	N2	C3	C8	-1.12(16)	C12	O3	C11	C10	175.69(9)

Table 9.4.6: Hydrogen Atom Coordinates (Å×10⁴) and Isotropic Displacement Parameters (Å²×10³) for 9.4.

Atom	x	y	z	U(eq)
H1	7960(30)	1520(30)	4020(30)	66(6)
H3	6750(20)	4770(20)	1506(17)	33(4)
H4A	6050.91	8316.54	4621.23	30
H4B	8277.96	8803.83	5398.31	30
H5A	7065.77	10218.49	7017.55	30
H5B	5525.39	8746.29	6792.78	30
H6A	7984.67	9159.12	9073.12	31
H6B	9494.17	8910.35	8256.54	31
H7A	8878.99	6514.57	8765.23	28
H7B	6653.9	6529.36	8183.56	28
H10A	7641.88	2254.2	-116.54	27
H10B	6893.36	3760.19	-667.72	27
H11A	4648.07	1455.36	-1930.85	27
H11B	4518.72	1212.68	-416.95	27
H12A	1459.9	1761.34	-2824.82	38
H12B	611.31	3009	-1856.54	38
H12C	1200.63	1424.24	-1382.39	38

9.6 10-(2-methoxyethyl)-3-phenylbenzo[g]pteridine-2,4(3*H*,10*H*)-dione, 9.5
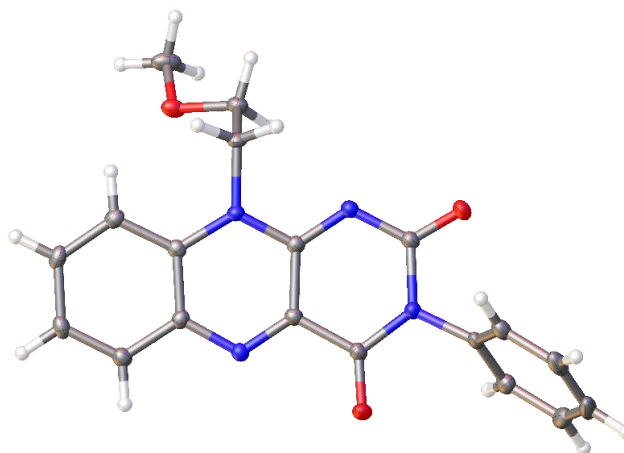
Figure 9.5: X-Ray crystal structure of NSC-288387 - 10-(2-methoxyethyl)-3-phenylbenzo[g]pteridine-2,4(3*H*,10*H*)-dione.

Table 9.5.1: Fractional Atomic Coordinates (×10⁴) and Equivalent Isotropic Displacement Parameters (Å²×10³) for 9.5. U$_{eq}$ is defined as 1/3 of the trace of the orthogonalised U$_{IJ}$ tensor.

Atom	x	y	z	U(eq)
O1	1520.0(8)	3840.1(4)	2938.0(14)	25.4(2)
O2	5268.1(9)	5007.3(4)	2598.9(13)	23.2(2)
O3	8826.7(8)	2656.9(4)	4096.8(11)	20.9(2)
N1	3365.4(10)	4430.9(4)	2630.6(14)	16.2(2)
N2	5660.4(10)	4036.3(4)	2460.8(13)	16.6(2)
N3	5915.5(9)	3044.3(4)	2174.4(13)	14.1(2)
N4	3183.6(10)	2873.3(4)	2869.2(13)	15.4(2)
C1	2749.9(12)	3896.5(5)	2808.9(16)	16.9(2)
C2	4824.0(12)	4514.0(5)	2574.9(16)	16.8(3)
C3	5116.3(11)	3515.7(5)	2466.8(15)	14.7(2)
C4	5410.6(11)	2479.5(5)	2280.7(15)	14.8(2)
C5	6202.2(12)	1984.8(5)	2024.5(16)	17.2(2)
C6	5636.3(13)	1437.0(5)	2167.3(16)	18.8(3)
C7	4281.2(13)	1363.3(5)	2577.7(16)	19.1(3)
C8	3494.7(12)	1844.0(5)	2826.3(16)	17.8(3)
C9	4033.3(12)	2407.8(5)	2674.7(15)	15.2(2)
C10	3699.3(11)	3390.9(5)	2749.7(15)	15.1(2)
C11	2441.2(11)	4930.9(5)	2370.7(17)	17.6(3)
C12	2120.2(13)	5171.8(5)	598.7(18)	22.0(3)
C13	1117.9(14)	5608.3(6)	277(2)	29.2(3)
C14	461.7(14)	5800.1(6)	1720(2)	33.0(3)
C15	810.5(14)	5563.0(6)	3496(2)	30.5(3)

Table 9.5.1: Fractional Atomic Coordinates (×10⁴) and Equivalent Isotropic Displacement Parameters (Å²×10³) for 9.5. U_{eq} is defined as 1/3 of the trace of the orthogonalised U_{IJ} tensor.

Atom	x	y	z	U(eq)
C16	1815.9(13)	5124.9(5)	3838.6(19)	22.8(3)
C17	7307.4(11)	3163.7(5)	1673.1(16)	16.4(2)
C18	8464.2(12)	3218.1(5)	3349.0(17)	18.7(3)
C19	10140.5(13)	2663.7(6)	5347.0(18)	26.2(3)

Table 9.5.2: Anisotropic Displacement Parameters (Å²×10³) for 9.5. The Anisotropic displacement factor exponent takes the form: -$2\pi^2[h^2a^{*2}U_{11}+2hka^*b^*U_{12}+\dots]$.

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
O1	13.1(4)	17.3(4)	47.1(6)	3.8(4)	8.7(4)	0.0(3)
O2	18.4(4)	13.7(4)	37.7(5)	1.3(3)	5.2(4)	-3.0(3)
O3	13.9(4)	25.2(5)	22.8(4)	5.4(3)	0.8(3)	-2.3(3)
N1	13.1(5)	12.3(5)	23.4(5)	1.5(4)	3.6(4)	0.2(4)
N2	14.0(5)	14.4(5)	21.5(5)	0.5(4)	3.5(4)	-1.2(4)
N3	11.5(5)	14.3(5)	16.9(5)	-0.1(4)	3.3(4)	0.0(4)
N4	14.4(5)	13.8(5)	17.8(5)	1.2(4)	1.9(4)	-0.1(4)
C1	15.2(6)	13.9(6)	21.8(6)	1.8(4)	3.3(4)	-0.9(4)
C2	14.6(5)	15.1(6)	20.4(6)	1.1(4)	2.6(4)	-1.5(4)
C3	13.7(5)	15.3(6)	14.7(5)	0.6(4)	1.1(4)	0.2(4)
C4	14.7(5)	15.7(6)	13.1(5)	-0.3(4)	-0.1(4)	-1.2(4)
C5	15.4(6)	19.2(6)	16.7(5)	-1.8(4)	1.9(4)	1.7(5)
C6	21.4(6)	16.1(6)	17.5(5)	-2.6(4)	-0.1(4)	5.1(5)
C7	22.8(6)	13.2(6)	19.9(6)	0.2(4)	-0.7(5)	-1.6(5)
C8	16.2(5)	17.0(6)	19.4(6)	1.2(4)	1.2(4)	-1.3(5)
C9	14.5(5)	15.2(6)	15.3(5)	0.1(4)	0.7(4)	1.5(4)
C10	13.4(5)	15.5(6)	16.1(5)	0.9(4)	1.8(4)	-0.6(4)
C11	12.3(5)	11.0(5)	29.1(6)	0.6(5)	2.0(4)	-2.1(4)
C12	18.9(6)	16.2(6)	30.1(7)	3.1(5)	1.8(5)	-3.6(5)
C13	21.3(6)	18.2(6)	44.4(8)	10.4(6)	-5.0(6)	-3.9(5)
C14	17.0(6)	13.3(6)	66.6(10)	1.7(6)	1.2(6)	1.8(5)
C15	21.0(6)	18.7(6)	53.6(9)	-10.8(6)	11.1(6)	-1.9(5)
C16	19.6(6)	17.3(6)	31.7(7)	-3.4(5)	5.6(5)	-2.8(5)
C17	13.2(5)	18.4(6)	18.8(6)	1.1(4)	6.5(4)	-0.6(4)
C18	14.1(5)	19.9(6)	22.6(6)	0.4(5)	4.9(4)	-3.0(4)
C19	14.2(6)	39.6(8)	23.9(6)	7.1(6)	0.8(5)	-2.1(5)

Table 9.5.3: Bond Lengths for 9.5.

Atom	Atom	Length/Å	Atom	Atom	Length/Å
O1	C1	1.2172(14)	C3	C10	1.4501(15)
O2	C2	1.2121(14)	C4	C5	1.4016(16)
O3	C18	1.4202(14)	C4	C9	1.4203(16)
O3	C19	1.4302(14)	C5	C6	1.3845(17)
N1	C1	1.3813(15)	C6	C7	1.4043(17)
N1	C2	1.4325(14)	C7	C8	1.3714(17)
N1	C11	1.4493(14)	C8	C9	1.4084(16)
N2	C2	1.3755(15)	C11	C12	1.3839(17)
N2	C3	1.3080(15)	C11	C16	1.3843(17)
N3	C3	1.3692(15)	C12	C13	1.3878(18)
N3	C4	1.3943(15)	C13	C14	1.385(2)
N3	C17	1.4801(14)	C14	C15	1.385(2)
N4	C9	1.3720(15)	C15	C16	1.3934(18)
N4	C10	1.2991(15)	C17	C18	1.5100(16)
C1	C10	1.4873(16)			

Table 9.5.4: Bond Angles for 9.5.

Atom	Atom	Atom	Angle/°	Atom	Atom	Atom	Angle/°
C18	O3	C19	111.53(9)	C6	C5	C4	119.74(11)
C1	N1	C2	124.36(10)	C5	C6	C7	121.44(11)
C1	N1	C11	116.81(9)	C8	C7	C6	119.37(11)
C2	N1	C11	118.71(9)	C7	C8	C9	120.67(11)
C3	N2	C2	119.27(10)	N4	C9	C4	122.01(10)
C3	N3	C4	121.07(9)	N4	C9	C8	118.26(10)
C3	N3	C17	116.97(9)	C8	C9	C4	119.70(10)
C4	N3	C17	121.92(9)	N4	C10	C1	117.88(10)
C10	N4	C9	117.63(10)	N4	C10	C3	125.06(11)
O1	C1	N1	123.05(11)	C3	C10	C1	116.98(10)
O1	C1	C10	122.45(10)	C12	C11	N1	118.29(11)
N1	C1	C10	114.45(10)	C12	C11	C16	121.68(11)
O2	C2	N1	118.26(10)	C16	C11	N1	119.81(11)
O2	C2	N2	122.48(10)	C11	C12	C13	119.01(12)
N2	C2	N1	119.25(10)	C14	C13	C12	120.10(13)
N2	C3	N3	119.03(10)	C13	C14	C15	120.35(12)
N2	C3	C10	124.96(11)	C14	C15	C16	120.15(13)
N3	C3	C10	116.01(10)	C11	C16	C15	118.68(12)
N3	C4	C5	122.94(10)	N3	C17	C18	113.55(9)
N3	C4	C9	117.99(10)	O3	C18	C17	109.42(9)
C5	C4	C9	119.07(10)				

Table 9.5.5: Torsion Angles for 9.5.

A	B	C	D	Angle/°	A	B	C	D	Angle/°
O1	C1	C10	N4	-1.02(17)	C4	N3	C3	C10	-4.37(15)
O1	C1	C10	C3	175.96(11)	C4	N3	C17	C18	-92.98(12)
N1	C1	C10	N4	-178.57(10)	C4	C5	C6	C7	-0.42(17)
N1	C1	C10	C3	-1.60(15)	C5	C4	C9	N4	-177.45(10)
N1	C11	C12	C13	172.90(11)	C5	C4	C9	C8	0.86(16)
N1	C11	C16	C15	-172.76(11)	C5	C6	C7	C8	0.60(17)
N2	C3	C10	N4	-175.60(11)	C6	C7	C8	C9	-0.03(17)
N2	C3	C10	C1	7.67(17)	C7	C8	C9	N4	177.68(10)
N3	C3	C10	N4	4.88(17)	C7	C8	C9	C4	-0.70(17)
N3	C3	C10	C1	-171.85(9)	C9	N4	C10	C1	175.28(10)
N3	C4	C5	C6	179.42(10)	C9	N4	C10	C3	-1.43(17)
N3	C4	C9	N4	2.81(16)	C9	C4	C5	C6	-0.31(16)
N3	C4	C9	C8	-178.88(10)	C10	N4	C9	C4	-2.51(16)
N3	C17	C18	O3	75.33(12)	C10	N4	C9	C8	179.16(10)
C1	N1	C2	O2	-172.91(11)	C11	N1	C1	O1	-7.33(17)
C1	N1	C2	N2	8.32(17)	C11	N1	C1	C10	170.20(10)
C1	N1	C11	C12	-104.50(12)	C11	N1	C2	O2	11.05(16)
C1	N1	C11	C16	70.24(14)	C11	N1	C2	N2	-167.72(10)
C2	N1	C1	O1	176.56(11)	C11	C12	C13	C14	0.47(19)
C2	N1	C1	C10	-5.91(16)	C12	C11	C16	C15	1.80(18)
C2	N1	C11	C12	71.84(14)	C12	C13	C14	C15	0.7(2)
C2	N1	C11	C16	-113.42(12)	C13	C14	C15	C16	-0.6(2)
C2	N2	C3	N3	173.85(10)	C14	C15	C16	C11	-0.59(19)
C2	N2	C3	C10	-5.65(17)	C16	C11	C12	C13	-1.75(18)
C3	N2	C2	O2	179.08(11)	C17	N3	C3	N2	-6.18(15)
C3	N2	C2	N1	-2.20(16)	C17	N3	C3	C10	173.37(9)
C3	N3	C4	C5	-178.83(10)	C17	N3	C4	C5	3.55(16)
C3	N3	C4	C9	0.90(15)	C17	N3	C4	C9	-176.72(10)
C3	N3	C17	C18	89.30(12)	C19	O3	C18	C17	164.55(9)
C4	N3	C3	N2	176.09(10)					

Table 9.5.6: Hydrogen Atom Coordinates ($Å×10^4$) and Isotropic Displacement Parameters ($Å^2×10^3$) for 9.5.

Atom	x	y	z	U(eq)
H5	7123.54	2025.21	1753.93	21
H6	6176.56	1103.69	1983.53	23
H7	3912.71	984.12	2682.2	23
H8	2577.24	1796.31	3103.59	21
H12	2579.04	5040.36	-383.06	26
H13	881.72	5775.71	-934.39	35
H14	-231.37	6095.93	1490.84	40
H15	363.27	5699.3	4482.84	37
H16	2067.24	4962.49	5055.45	27
H17A	7255.06	3529.1	941.58	20
H17B	7545.51	2845.83	860.59	20
H18A	9294.54	3403.99	2969.44	22
H18B	8146.87	3465.4	4314.42	22
H19A	10872.26	2809.92	4685.88	39
H19B	10377.58	2268.2	5797.5	39
H19C	10071.86	2917.46	6414.38	39

9.7 *N*-[3-(2-methylphenyl)isoquinolin-1-yl]-2-pyridinecarboxamidine, 9.6
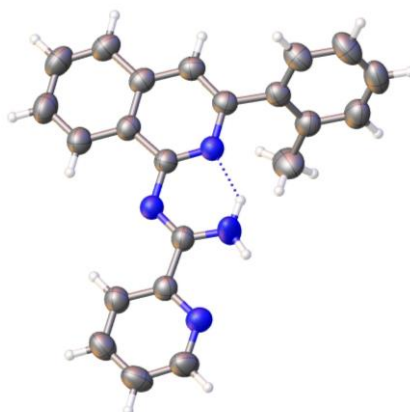
Figure 9.6: X-ray crystal structure of *N*-[3-(2-methylphenyl)isoquinolin-1-yl]-2-pyridinecarboxamidine.

Table 9.6.1 Fractional Atomic Coordinates (×10^4) and Equivalent Isotropic Displacement Parameters (Å^2×10^3) for 9.6. U$_{eq}$ is defined as 1/3 of the trace of the orthogonalised U$_{IJ}$ tensor.

Atom	x	y	z	U(eq)
N2	4498.4(13)	2765.1(10)	5391.8(8)	52.3(2)
N18	6635.2(13)	3883.0(10)	7144.3(8)	53.2(3)
N21	11155.0(14)	5404.2(11)	6692.4(9)	60.2(3)
C3	4870.1(16)	3163.8(12)	6549.9(10)	49.8(3)
C22	9824.2(16)	5050.6(12)	7296.4(10)	51.2(3)
C1	2758.5(16)	2028.4(11)	4826.3(10)	50.9(3)
C10	1317.3(16)	1713.6(12)	5418.6(11)	54.9(3)
C4	3464.1(16)	2828.0(12)	7239.0(10)	51.3(3)
C11	2579.7(16)	1595.3(11)	3516.8(10)	52.5(3)
N20	7727(2)	4205.8(17)	5515.4(11)	86.6(5)
C9	1646.7(17)	2095.3(12)	6650.7(11)	53.6(3)
C19	7927.7(16)	4323.8(12)	6612.7(10)	53.6(3)
C16	3915.2(17)	859.0(13)	2785.2(11)	58.9(3)
C23	10159.7(18)	5332.3(13)	8480.2(11)	58.8(3)
C8	261.1(19)	1733.2(14)	7305.7(13)	64.5(3)
C5	3836.8(19)	3181.5(14)	8457.0(12)	61.7(3)
C12	1074(2)	1922.4(14)	3015.8(12)	66.8(4)
C6	2470(2)	2797.3(16)	9056.7(13)	71.5(4)
C15	3676(2)	501.4(15)	1575.2(12)	71.8(4)
C26	12856.7(19)	6059.6(15)	7280.5(14)	69.9(4)
C24	11926(2)	6025.8(15)	9070.1(13)	69.6(4)
C14	2193(2)	850.1(16)	1092.8(13)	77.7(4)
C7	688(2)	2064.9(16)	8476.3(14)	72.8(4)
C25	13298(2)	6400.2(16)	8461.5(14)	74.2(4)
C13	885(2)	1557.8(16)	1807.3(14)	80.0(4)

439

Table 9.6.1 Fractional Atomic Coordinates (×10⁴) and Equivalent Isotropic Displacement Parameters (Å²×10³) for 9.6. U$_{eq}$ is defined as 1/3 of the trace of the orthogonalised U$_{IJ}$ tensor.

Atom	x	y	z	U(eq)
C17	5533(2)	415.2(18)	3261.4(15)	85.6(5)

Table 9.6.2 Anisotropic Displacement Parameters (Å²×10³) for 9.6. The Anisotropic displacement factor exponent takes the form: -2π²[h²a*²U$_{11}$+2hka*b*U$_{12}$+...].

Atom	U$_{11}$	U$_{22}$	U$_{33}$	U$_{23}$	U$_{13}$	U$_{12}$
N2	50.7(5)	57.7(6)	48.0(5)	24.7(5)	2.7(4)	1.5(4)
N18	49.0(5)	62.9(6)	48.7(5)	27.5(5)	3.6(4)	1.5(4)
N21	53.3(5)	67.7(7)	61.5(6)	32.9(5)	3.9(4)	-1.2(5)
C3	49.9(6)	53.2(6)	49.0(6)	25.5(5)	4.8(5)	6.6(5)
C22	53.1(6)	49.6(6)	50.8(6)	22.8(5)	5.0(5)	4.5(5)
C1	50.2(6)	48.2(6)	54.0(6)	23.1(5)	2.5(5)	5.3(5)
C10	49.0(6)	53.2(7)	60.1(7)	24.4(6)	2.9(5)	2.3(5)
C4	52.7(6)	53.7(6)	53.5(6)	28.9(5)	8.9(5)	9.1(5)
C11	54.3(6)	47.9(6)	53.2(6)	23.4(5)	-1.5(5)	-2.0(5)
N20	73.7(8)	124.7(12)	55.2(7)	48.2(7)	-6.1(6)	-36.4(8)
C9	53.1(6)	50.6(6)	61.4(7)	28.3(6)	11.0(5)	7.8(5)
C19	55.0(6)	57.7(7)	46.7(6)	23.4(5)	4.7(5)	2.3(5)
C16	57.6(6)	58.4(7)	56.3(7)	23.4(6)	4.7(5)	1.0(5)
C23	62.0(7)	61.5(7)	54.4(7)	28.4(6)	2.4(5)	4.4(5)
C8	59.7(7)	63.0(8)	74.8(9)	34.9(7)	16.8(6)	4.6(6)
C5	63.2(7)	72.8(8)	57.8(7)	36.8(6)	7.6(5)	10.0(6)
C12	70.4(8)	61.9(8)	64.2(8)	24.7(6)	-6.9(6)	10.2(6)
C6	79.1(9)	87.5(10)	64.0(8)	46.6(8)	18.5(7)	17.4(7)
C15	75.3(8)	71.9(9)	56.2(7)	21.0(7)	8.7(6)	-2.3(7)
C26	55.4(7)	82.0(10)	77.9(9)	45.1(8)	-0.8(6)	-6.0(6)
C24	72.1(8)	74.8(9)	60.9(8)	32.7(7)	-10.5(6)	0.5(7)
C14	98.4(11)	75.0(9)	52.2(7)	28.0(7)	-7.2(7)	-10.8(8)
C7	75.4(9)	79.5(10)	78.9(9)	48.1(8)	28.8(7)	12.3(7)
C25	61.2(7)	77.5(9)	84.4(10)	42.3(8)	-16.8(7)	-8.8(7)
C13	94.0(10)	76.2(10)	68.6(9)	34.2(8)	-21.6(8)	4.7(8)
C17	78.4(9)	93.7(12)	80.1(10)	28.5(9)	11.2(8)	34.2(9)

Table 9.6.3 Bond Lengths for 9.6.

Atom	Atom	Length/Å	Atom	Atom	Length/Å
N2	C3	1.3274(14)	C11	C12	1.3835(17)
N2	C1	1.3672(14)	N20	C19	1.3255(16)
N18	C3	1.3852(15)	C9	C8	1.4145(17)
N18	C19	1.3063(15)	C16	C15	1.3889(18)
N21	C22	1.3391(15)	C16	C17	1.506(2)
N21	C26	1.3310(17)	C23	C24	1.3777(18)
C3	C4	1.4412(16)	C8	C7	1.363(2)
C22	C19	1.4934(16)	C5	C6	1.3680(19)
C22	C23	1.3831(16)	C12	C13	1.384(2)
C1	C10	1.3630(16)	C6	C7	1.394(2)
C1	C11	1.4912(16)	C15	C14	1.372(2)
C10	C9	1.4153(17)	C26	C25	1.374(2)
C4	C9	1.4164(17)	C24	C25	1.373(2)
C4	C5	1.4085(17)	C14	C13	1.367(2)
C11	C16	1.4025(18)			

Table 9.6.4 Bond Angles for 9.6.

Atom	Atom	Atom	Angle/°	Atom	Atom	Atom	Angle/°
C3	N2	C1	120.44(10)	C8	C9	C10	122.60(11)
C19	N18	C3	120.89(10)	C8	C9	C4	118.83(11)
C26	N21	C22	117.23(11)	N18	C19	C22	117.86(10)
N2	C3	N18	121.91(10)	N18	C19	N20	126.51(11)
N2	C3	C4	121.11(10)	N20	C19	C22	115.62(11)
N18	C3	C4	116.95(10)	C11	C16	C17	122.34(12)
N21	C22	C19	115.88(10)	C15	C16	C11	117.91(12)
N21	C22	C23	122.82(11)	C15	C16	C17	119.72(13)
C23	C22	C19	121.31(11)	C24	C23	C22	118.67(12)
N2	C1	C11	114.57(10)	C7	C8	C9	120.34(13)
C10	C1	N2	122.29(11)	C6	C5	C4	120.22(13)
C10	C1	C11	123.13(10)	C11	C12	C13	121.11(14)
C1	C10	C9	119.57(11)	C5	C6	C7	120.60(13)
C9	C4	C3	117.96(10)	C14	C15	C16	121.96(14)
C5	C4	C3	122.79(11)	N21	C26	C25	123.72(13)
C5	C4	C9	119.25(11)	C25	C24	C23	119.01(13)
C16	C11	C1	120.92(10)	C13	C14	C15	119.94(13)
C12	C11	C1	119.57(11)	C8	C7	C6	120.73(12)
C12	C11	C16	119.51(11)	C24	C25	C26	118.54(12)
C10	C9	C4	118.53(11)	C14	C13	C12	119.55(13)

Table 9.6.5 Torsion Angles for 9.6.

A	B	C	D	Angle/°	A	B	C	D	Angle/°
N2	C3	C4	C9	-2.01(17)	C10	C1	C11	C16	-122.98(13)
N2	C3	C4	C5	177.08(11)	C10	C1	C11	C12	56.93(16)
N2	C1	C10	C9	-3.43(18)	C10	C9	C8	C7	176.46(12)
N2	C1	C11	C16	56.63(15)	C4	C9	C8	C7	-1.20(19)
N2	C1	C11	C12	-123.46(12)	C4	C5	C6	C7	-0.7(2)
N18	C3	C4	C9	179.86(10)	C11	C1	C10	C9	176.14(10)
N18	C3	C4	C5	-1.05(18)	C11	C16	C15	C14	0.2(2)
N21	C22	C19	N18	175.52(11)	C11	C12	C13	C14	0.7(2)
N21	C22	C19	N20	-5.17(18)	C9	C4	C5	C6	1.28(19)
N21	C22	C23	C24	1.13(19)	C9	C8	C7	C6	1.9(2)
N21	C26	C25	C24	1.3(2)	C19	N18	C3	N2	7.69(18)
C3	N2	C1	C10	2.47(17)	C19	N18	C3	C4	-174.19(10)
C3	N2	C1	C11	-177.14(10)	C19	C22	C23	C24	-178.89(12)
C3	N18	C19	C22	-177.89(10)	C16	C11	C12	C13	-1.3(2)
C3	N18	C19	N20	2.9(2)	C16	C15	C14	C13	-0.7(2)
C3	C4	C9	C10	1.01(17)	C23	C22	C19	N18	-4.46(18)
C3	C4	C9	C8	178.76(10)	C23	C22	C19	N20	174.85(13)
C3	C4	C5	C6	-177.80(12)	C23	C24	C25	C26	-0.4(2)
C22	N21	C26	C25	-0.9(2)	C5	C4	C9	C10	-178.12(11)
C22	C23	C24	C25	-0.7(2)	C5	C4	C9	C8	-0.36(18)
C1	N2	C3	N18	178.37(10)	C5	C6	C7	C8	-0.9(2)
C1	N2	C3	C4	0.33(17)	C12	C11	C16	C15	0.84(18)
C1	C10	C9	C4	1.60(17)	C12	C11	C16	C17	-177.14(13)
C1	C10	C9	C8	-176.07(11)	C15	C14	C13	C12	0.3(2)
C1	C11	C16	C15	-179.25(11)	C26	N21	C22	C19	179.66(11)
C1	C11	C16	C17	2.77(19)	C26	N21	C22	C23	-0.35(19)
C1	C11	C12	C13	178.80(13)	C17	C16	C15	C14	178.19(14)

Table 9.6.6 Hydrogen Atom Coordinates ($\text{Å}\times10^4$) and Isotropic Displacement Parameters ($\text{Å}^2\times10^3$) for 9.6.

Atom	x	y	z	U(eq)
H10	99.48	1241.15	5007.12	66
H23	9191.35	5053.08	8878.63	71
H8	-973.51	1256.37	6927.58	77
H5	5041.1	3688.32	8863.5	74
H12	154.99	2405.44	3509.8	80
H6	2738.03	3031.75	9876.18	86
H15	4565.7	0.54	1066.36	86
H26	13821.09	6305.87	6860.14	84
H24	12191.61	6241.78	9885.81	83
H14	2075.77	599.89	262.86	93
H7	-239.39	1794.08	8901.57	87
H25	14524.55	6884.15	8848.2	89
H13	-147.66	1797.69	1476.53	96
H17A	6547.84	1172.4	3698.07	128
H17B	6078.79	-267.83	2598.41	128
H17C	5026.79	52.31	3797.68	128
H20A	8690(20)	4586(19)	5282(17)	103(6)
H20B	6580(20)	3737(17)	5087(15)	92(5)