



MONASH University

Development of the diaryl ether scaffold as inhibitors of *EcDsbA* using a fragment-based approach

Matthew Ross Bentley

Bachelor of Pharmaceutical Sciences (Hons)

A thesis submitted for the degree of *Doctor of Philosophy* at
Monash University in 2019

Department of Medicinal Chemistry
Faculty of Pharmacy and Pharmaceutical Sciences
Monash University

Copyright notice

© Matthew Bentley (2019).

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

Table of contents

Publications during enrolment	vi
Presentations.....	vii
Acknowledgements	viii
Abbreviations	ix
Abstract	xiv
Chapter 1 – Introduction and Thesis Aims.....	1
1.1 Bacterial resistance	2
1.1.1 Antimicrobial resistance	3
1.1.2 Bacterial resistance mechanisms.....	3
1.1.3 Gram-negative bacteria	5
1.2 Bacterial virulence	6
1.2.1 Bacterial adhesion factors	6
1.2.2 Bacterial toxins	7
1.2.3 Secretion systems.....	7
1.3 Targeting bacterial virulence	8
1.3.1 Targeting type II secretion systems	8
1.3.2 Bacterial adhesion as an antivirulence target.....	8
1.4 Disulfide bond forming proteins.....	9
1.4.1 Structure of <i>Ec</i> DsbA	10
1.4.2 Biological evaluation of the DsbA as a drug target	13
1.4.3 Chemical approaches to inhibit Dsb enzymes	14
1.5 Fragment-based drug discovery (FBDD)	15
1.5.1 Nuclear magnetic resonance (NMR)	15
1.5.2 Surface plasmon resonance (SPR).....	21
1.5.3 X-ray crystallography	22
1.6 Ligand efficiency.....	23
1.7 Fragment elaboration.....	24
1.7.1 Growing	24
1.7.2 Merging.....	25
1.7.3 Linking.....	27
1.8 Efficient elaboration of fragments - Off rate screening.....	28
1.9 Fragment screening cascade	32
1.10 Previous work.....	33

1.10.1 Fragment screen	33
1.10.2 Development of phenylthiazoles as inhibitors of <i>EcDsbA</i>	33
1.11 Diaryl ethers a new class of <i>EcDsbA</i> inhibitors	35
1.12 Aims and hypothesis	35
1.13 References	37
Chapter 2 – Substitution of the benzonitrile ring.	43
2.1 Previous work	44
2.2 Screening cascade	46
2.2.1 Solubility and aggregation assessment	46
2.2.2 Binding by ^1H - ^{15}N HSQC NMR	48
2.2.3 Structural characterisation	49
2.3 Crystal structure analysis	50
2.4 Expansion off the benzonitrile ring	52
2.4.1 Design and docking	52
2.4.2 Synthesis	53
2.4.3 Testing R^1 and R^2 expanded compounds	55
2.4.4 Synthesis and testing the <i>para</i> - CF_3 analogue.	59
2.5 Substitution of the <i>para</i> -cyano group	62
2.5.1 Design and docking	62
2.5.2 Synthesis of <i>para</i> -cyano replacement analogues	63
2.5.3 Testing nitrile substituted analogues	68
2.6 Bicyclic diaryl ethers	72
2.6.1 Design and docking	72
2.6.2 Synthesis of bicyclic diaryl ethers	73
2.6.3 Testing bicyclic diaryl ether analogues	74
2.7 Conclusion	75
2.8 References	77
Chapter 3 - Optimisation of the diaryl ether core	78
3.1 Introduction	79
3.2 Replacement of the <i>para</i> -hydroxyl and merging the best analogues	80
3.2.1 Design and docking	80
3.2.2 Synthesis	81
3.2.3 Screening phenol replacement analogues	82
3.3 Further optimisation of the right-hand side	86
3.3.1 Design and synthesis of carboxylic acid analogues	86
3.3.3 Testing diaryl ether carboxylic acid analogues	87

3.4	Isosteric replacement of the ether linker and expansion into the bottom groove.....	90
3.4.1	Introduction and design.....	90
3.4.2	Synthesis and testing of single heavy atom linkers	93
3.4.3	Synthesis of two heavy atom linkers	94
3.4.4	Testing one and two heavy atom linker series	96
3.5	Expansion into the bottom groove using the Petasis reaction.	97
3.5.1	Design and docking.....	97
3.5.2	Synthesis of bottom groove expanded analogues.	99
3.5.3	Testing bottom groove expanded analogues.....	100
3.6	Conclusions	102
3.7	References	104
Chapter 4	– Expansion of the diaryl ethers across the hydrophobic groove.	105
4.1	Introduction	106
4.2	Amidation of the carboxylic acid	107
4.2.1	Design and docking.....	107
4.2.2	Synthesis of expanded carboxylic acid analogues	109
4.2.3	Testing amidation analogues.....	115
4.3	Benzylic expansion and isosteric replacement of the carboxylic acid	119
4.3.1	Acyl sulfonamide design, docking and synthesis	120
4.3.2	Benzylic expansion design and docking.....	121
4.3.3	Synthesis of benzylic expanded analogues	123
4.3.4	Testing benzylic expansion and acyl sulfonamide analogues.....	127
4.4	Conclusion	131
4.5	References	133
Chapter 5	– Rapid elaboration of fragments into leads by X-ray crystallographic screening of parallel chemical libraries (REFiL _X).....	134
5.1	Introduction	135
Chapter 6	– Optimisation of the ethyl pyrazole and expansion across the groove.	143
6.1	Introduction	144
6.2	Isosteric replacement of the ethyl pyrazole	144
6.2.1	Design of pyrazole analogues	144
6.2.2	Replacing the pyrazole ring	146
6.2.3	Synthesis of <i>N,N</i> -dialkyl amide analogues	148
6.2.4	Testing of pyrazole analogues	152
6.3	Alkylation of the pyrazole amide using REFiL _X	155
6.3.1	Introduction, context and design.....	155

6.3.2 Optimisation of plate and soaking conditions.....	156
6.3.3 REFiL _x screen and hit identification	158
6.3.4 Resynthesis of hits	159
6.3.5 Affinity determination and solubility testing.....	163
6.4 Conclusion	166
6.5 References:	167
Chapter 7 – Thesis outcomes and future prospects	168
Chapter 8 –Experimental.....	178
8.1 General experimental.....	179
8.2 General procedures	181
8.3 Chemistry experimental.....	185
8.3.1 Parallel synthesis plate setup	284
8.4.1 Extraction and purification of Double His-Tagged <i>EcDsbA</i> (<i>EcDsbA</i> -DHT)	285
8.4.2 SPR analysis.....	285
8.4.3 Crystallisation and X-ray diffraction experiments.....	286
8.4.4 Assessing solubility of synthesised analogues by q-NMR	287
8.4.5 Validation of hits and affinity determination by ¹ H ¹⁵ N-HSQC	287
8.5.6 Computational docking.....	288
8.5 References	289
Appendices	291
Appendix 1.0: Applications of NMR Spectroscopy in FBDD	292
Appendix 2.1: ¹ H NMR spectra of <i>cis</i> and <i>trans</i> isomers of enol ether 2.59	313
Appendix 2.2: Synthesis conditions for bicyclic cyclopentanone analogue 2.63 via copper catalysed Ullmann coupling.	314
Appendix 4.1: Optimisation of aldol condensation reaction conditions.	315
Appendix 5.1: Rapid elaboration of fragments into leads by X-ray crystallographic screening of parallel chemical libraries (REFiL _x) supplementary information.....	317
Table 2.....	335
Table 3. Electron density maps obtained from crude soaks and co-crystals of hits 3 – 6.	336
Fig 5. X-ray crystal structure and HSQC titrations of oxadiazole 8	339
Fig 6.....	340
Assessing solubility of synthesised analogues by q-NMR.....	346

Publications during enrolment

Bentley M., Doak B. C., Mohanty B., Scanlon M. J. (2018) Applications of NMR Spectroscopy in FBDD. In: Webb G. (eds) Modern Magnetic Resonance. Springer, Cham. (**see appendix 1.0**).

Presentations

Oral presentations

Monash 13th Annual Postgraduate Symposium, Melbourne, Australia, 2018. Bentley, M., *et al.*
Rapid elaboration of fragments into leads using X-ray crystallography (REFiLx).

Fragment-Based Drug Discovery Down Under, Melbourne, Australia, 2019, Bentley, M
Rapid elaboration of fragments into leads by X-ray crystallographic screening of parallel chemical libraries (REFiLx).

Poster presentations

RACI 41st Annual Synthesis Symposium, Melbourne, Australia, 2016. Bentley, M., *et al.* *Development of diaryl ethers as inhibitors of EcDsbA*

Fragment Based Drug Discovery Down Under, Melbourne, Australia, 2017. Bentley, M., *et al.*
Development of diaryl ethers as inhibitors of EcDsbA.

RACI National Centenary Conference, Melbourne, Australia, 2017. Bentley, M., *et al.* *Fragment-based development of diaryl ethers as inhibitors of EcDsbA.*

RACI 42nd Annual Synthesis Symposium, Melbourne, Australia, 2017. Bentley, M., *et al.*
Development of diaryl ethers as inhibitors of EcDsbA.

Fragment Based Lead Discovery, San Diego, United States of America, 2018. **Bentley, M., et al.**
Screening of unpurified reactions by X-ray Crystallography.

RACI Medicinal Chemistry and Chemical Biology Conference, Brisbane, Australia, 2018.
Bentley, M., *et al.* *Rapid elaboration of fragments into leads using X-ray crystallography (REFiLx).*

RACI 43rd Annual Synthesis Symposium, Melbourne, Australia, 2018. Bentley, M., *et al.* *Rapid Elaboration of fragments into leads using X-ray crystallography (REFiLx).*

Acknowledgements

I would like to express my upmost gratitude to my supervisors Prof. Martin Scanlon and Dr Ben Capuano for giving me the opportunity to do research in their groups.

I would like to thank the entire DsbA team including Dr Bradley Doak, Dr Olga Ilyichova, Dr Biswaranjan Mohanty, Prof. Peter Scammells, Wesam Alwan and Rebecca Whitehouse for all the advice and guidance provided. A thanks to our other collaborators Prof. Jenny Martin, A/Prof. Begona Heras and A/Prof Makrina Totsika. Thanks to all the members of the Scanlon and Capuano group and good luck to all current students. Big thanks to Dr Jason Dang for all the NMR and LCMS work he puts in.

A very special thanks to Brad and Ben, for all the many times I came to your office with chemistry problems and we were able to work through them together. Both of you have imparted a great deal of knowledge onto me and I have far greater skills now than when I began. I cannot leave Martins efforts in this area unrecognised either, your continual advice and assistance was extremely helpful. Similarly, thanks to Olga for all her assistance with X-ray crystallography and if there is one thing I have learnt from you it's to never trust an X-ray structure until you see the density.

It was one difficult journey working on DsbA and I can say it was made much easier having both Wes and Bec along for the ride with me. I don't regret at all for choosing this project all those years ago and I have learnt so much. One last thanks to all my friends and family for the continual support.

Abbreviations

^{13}C NMR	Carbon nuclear magnetic resonance
^{19}F NMR	Fluorine nuclear magnetic resonance
^1H NMR	Proton nuclear magnetic resonance
AcOH	Acetic acid
Ac	Acetyl group
AHT	Anhydrotetracycline
APT	Attached proton test
aq.	Aqueous
BBr ₃	Boron tribromide
Boc	<i>tert</i> -Butyloxycarbonyl
Boc ₂ O	Di- <i>tert</i> -butyl dicarbonate
Cat.	Catalytic
CD	β -cyclodextrin
CDCl ₃	Deuterated chloroform
CDI	Carbonyldiimidazole
COMU	(1-Cyano-2-ethoxy-2-oxoethylidenaminoxy)dimethylamino-morpholino-carbenium hexafluorophosphate
CPMG	Carr-Purcell-Meiboom-Gill
CSP	Chemical shift perturbation
DBU	1,8-Diazabicycloundec-7-ene
DCE	1,2-Dichloroethane
DCM	Dichloromethane
DIPEA	<i>N,N</i> -Diisopropylethylamine

DMAP	4-Dimethylaminopyridine
DMF	<i>N,N</i> -Dimethylformamide
DMSO	Dimethyl sulfoxide
DSB	Disulfide bond
DSS	4,4-Dimethyl-4-silapentane-1-sulfonic acid
<i>Ec</i>	<i>Escherichia coli</i>
EDC hydrochloride	<i>N</i> -(3-Dimethylaminopropyl)- <i>N'</i> -ethylcarbodiimide hydrochloride
EDIA	Electron density score for individual atoms
Eq.	Equivalents
Et ₂ O	Diethyl ether
EtOAc	Ethyl acetate
EtOH	Ethanol
EthR	Ethionamide transcriptional repressor
FBDD	Fragment-based drug discovery
FimA	Fimbrial protein A
HA	Heavy atom count
HATU	1-[Bis(dimethylamino)methylene]-1 <i>H</i> -1,2,3-triazolo[4,5- <i>b</i>]pyridinium 3-oxid hexafluorophosphate
H-bond	Hydrogen bond
HCTU	<i>O</i> -(6-Chlorobenzotriazol-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium hexafluorophosphate
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HGT	Horizontal gene transfer
HMBC	Heteronuclear multiple bond correlation

HOBt	Hydroxybenzotriazole
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single-quantum coherence
K_D	Affinity
KOtBu	Potassium <i>tert</i> -butoxide
LCMS	Liquid chromatography-mass spectrometry
LDA	Lithium diisopropylamide
LE	Ligand efficiency
LHS	Left hand side
LiAlH ₄	Lithium aluminium hydride
LLE	Lipophilic ligand efficiency
<i>m/z</i>	Mass-to-charge ratio
MDR	Multidrug resistant
MeCN	Acetonitrile
MEM	2-Methoxyethoxymethyl
MeOH	Methanol
MOM	Methoxymethyl
MS	Mass spectrometry
Ms	Methane sulfonyl group
MW	Molecular weight
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
PBP	Penicillin binding proteins

PDB	Protein data bank
PEG	Polyethylene glycol
Pet spirits	Petroleum spirits or petroleum ether
qNMR	Quantative nuclear magnetic resonance
Rbf	Round bottom flask
REFiL	Rapid elaboration of fragments into leads
REFiL _x	Rapid elaboration of fragments into leads by X-ray crystallography
RHS	Right hand side
RT	Room temperature
SAR	Structure activity relationship
SBDD	Structure based drug design
SEM	2-(Trimethylsilyl)ethoxymethyl
SEM	Standard error of mean
SMILES	Simplified molecular-input line-entry system
SMARTS	SMILES arbitrary target specification
SP	Standard precision
SPR	Surface plasmon resonance
STAB	Sodium triacetoxyborohydride
STD	Saturation-transfer difference
Et ₃ N	Triethylamine
TBAI	Tetrabutylammonium iodide
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TTSS	Type III secretion systems

WaterLOGSY	Water-ligand observed via gradient spectroscopy
XP	Extra precision
δ	Chemical shift
λ	Wavelength

Abstract

The rapid development of bacterial resistance to current antibiotics and limited number of novel compounds in the antibiotic drug pipeline threatens to send the 21st century into a post antibiotic era. Antivirulence agents which do not kill or inhibit bacterial growth have therefore been hypothesised as new approaches to treatment as they are expected to exert less selective pressure towards the development of resistance. This has led to increased interest in identifying suitable targets for the development of antivirulence agents. One such target is DsbA, a dithiol-disulfide oxidoreductase enzyme that catalyses the formation of disulfide bonds in newly synthesised proteins as they are translocated into the periplasm. Multiple studies have shown that bacteria lacking a functional DsbA have attenuated virulence and reduced capacity to cause infection. Thus, DsbA is crucial for bacterial virulence and a novel anti-virulence target.

Ligand detect NMR and HSQC NMR were previously employed to identify several small molecules that bind to *Escherichia coli* DsbA (*EcDsbA*) and inhibit its activity *in vitro*. One chemical class of inhibitors is the diaryl ethers for which an X-ray crystal structure bound to *EcDsbA* was obtained. This afforded a platform for structure-based drug design (SBDD) and fragment-based design of the diaryl ethers as inhibitors of *EcDsbA*. This research focused on exploring the design, synthesis and testing of multiple series of diaryl ethers aiming to develop them as novel anti-virulence agents.

Initial core optimisation afforded multiple X-ray crystal structures of analogues bound to *EcDsbA*, however analogues also demonstrated poor aqueous solubility. Utilising structural information from multiple diaryl ethers a series of merged analogues were designed which yielded improved affinity and solubility. Further expansion of the series by targeting multiple vectors using SBDD was then conducted, affording few analogues with improved affinities. This highlighted the difficulties in inhibiting DsbA's protein-protein interaction active site.

An alternative strategy to SBDD for fragment development was then explored employing parallel microscale chemistry and screening by X-ray crystallography. Diverse reagent libraries were used

with amide coupling chemistry to quickly synthesise a 96-well plate library of analogues. Screening was then conducted on the unpurified reaction products using X-ray crystallography. This rapid elaboration of fragments into leads by X-ray crystallography (REFiLx) method afforded higher affinity analogues and X-ray crystal structures to guide further medicinal chemistry. Optimisation of the best analogues was then conducted and resulted in improved solubility, but no improvements in binding affinity. Additional REFiLx libraries were then attempted, but were met with difficulty in synthesis. During the course of the PhD the highest affinity small molecule binder ($K_D = 62 \pm 9 \mu\text{M}$) to *EcDsbA* was designed, synthesised and characterised.

Chapter 1 – Introduction and Thesis Aims

1.1 Bacterial resistance

Widespread bacterial resistance is a global health problem that threatens to send the 21st century into a post-antibiotic era where common infections could be deadly.¹ The problem of bacterial resistance has arisen due to; the misuse of current antibiotics especially in agricultural and veterinary sectors,²⁻⁴ a lack of interest from big pharmaceutical companies (as antibiotics give poor return on investment)⁵ and the limited number of novel compounds in the antibiotic drug pipeline.⁶ The identification of new drug resistance mechanisms to existing therapeutics has led to the emergence of multidrug resistant (MDR) bacteria or “superbugs”. Such “superbugs” include strains of *Escherichia coli* (*E. coli*)⁷ and *Klebsiella pneumoniae*⁸ which show resistance to 3rd generation cephalosporins, azithromycin resistant strains of *Neisseria gonorrhoeae*⁹ which are becoming more prevalent in Australian hospitals and vancomycin resistant *Enterococci*.⁸ Current treatments for MDR “superbugs” are becoming less effective, thereby further increasing the annual health costs and mortality rates associated with MDR infections (**Figure 1.1**).

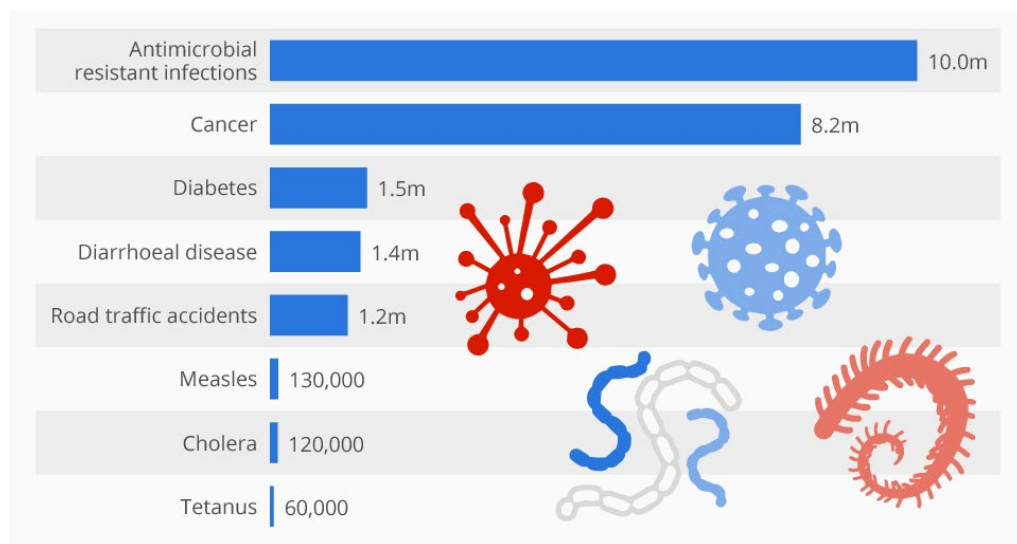


Figure 1.1: Predicted deaths caused by antimicrobial resistant infections in the year 2050. Antimicrobial resistant infections are predicted to overtake cancer as the leading cause of death worldwide. (reproduced from review on antimicrobial resistance, available online <https://amr-review.org/>)¹⁰

Antimicrobial resistant infections currently kill more than 700,000 people annually and are predicted to overtake cancer as the leading cause of death worldwide by 2050.¹¹ A significant global effort is required to overcome the issue of MDR infections.^{8, 12} Consequently, multiple world bodies such as the World Health Organisation and United Nations are calling for new antimicrobials to tackle this global health crisis.^{11, 13, 14}

1.1.1 Antimicrobial resistance

Antimicrobial resistance can be classified as either intrinsic or acquired resistance. Intrinsic bacterial resistance is generally developed slowly over a long period of time and is a result of bacteria responding to their direct surroundings.¹⁵ Intrinsic resistances may include bacteria displaying greater resistance to surrounding substances or conditions such as high levels of radiation, lowered levels of light or extreme pH.¹⁶⁻¹⁸ Acquired resistance however, is developed more quickly through mechanisms such as genetic mutation or horizontal gene transfer (HGT).¹⁹ HGT is a mechanism whereby bacteria share genetic material between organisms. This process can involve the spread of antibiotic resistance genes amongst different pathogens.²⁰ HGT is independent from cell division and rapidly accelerates the spread of resistance mechanisms giving rise to MDR “superbugs”.

1.1.2 Bacterial resistance mechanisms

Bacterial resistance can arise through a number of mechanisms (**Figure 1.2**) including:

- (a) Production of enzymes to chemically alter the antibiotic into a substance that exhibits little to no antimicrobial activity. For example, β -lactam antibiotics are enzymatically cleaved by β -lactamase enzymes in Gram-negative pathogens and aminoglycoside antibiotics are acetylated by acyltransferase enzymes.^{21, 22}
- (b) Mutation or transformation of the target binding site of the antibiotic. For example, fluoroquinolone resistance arises due to a single point amino acid mutation of the

topoisomerase type II enzyme or DNA gyrase enzyme in strains of *E. coli*, this mutation alters the quinolone binding site rendering fluoroquinolone antibiotics inactive.²³

- (c) Active efflux of molecules out of the bacterial cytosol or alteration of membrane permeability to restrict entry. β -lactam and tetracycline antibiotics utilise water-filled *trans*-membrane proteins called porins. These proteins act as a pore where water and other small hydrophilic molecules can diffuse through the outer cell membrane of Gram-negative bacteria.²⁴ Modification or downregulation of these porins in *Pseudomonas* bacteria provides low susceptibility and innate resistance to β -lactams antibiotics.²⁵ Tetracycline resistance arises through active efflux of compound from the cytosol of Gram-negative bacteria such as *Enterobacteriaceae*.²⁶
- (d) Replacing the target enzyme with another that performs a similar function or bypassing the pathway inhibited. β -lactam resistance in methicillin-resistant *Staphylococcus aureus* is caused by replacing the action of penicillin binding proteins (PBP) 1 - 4 with a 5th protein PBP5. The binding site of PBP5 is structurally different for the other 4 proteins hence, β -lactam antibiotics bind weakly to its active site.²⁷

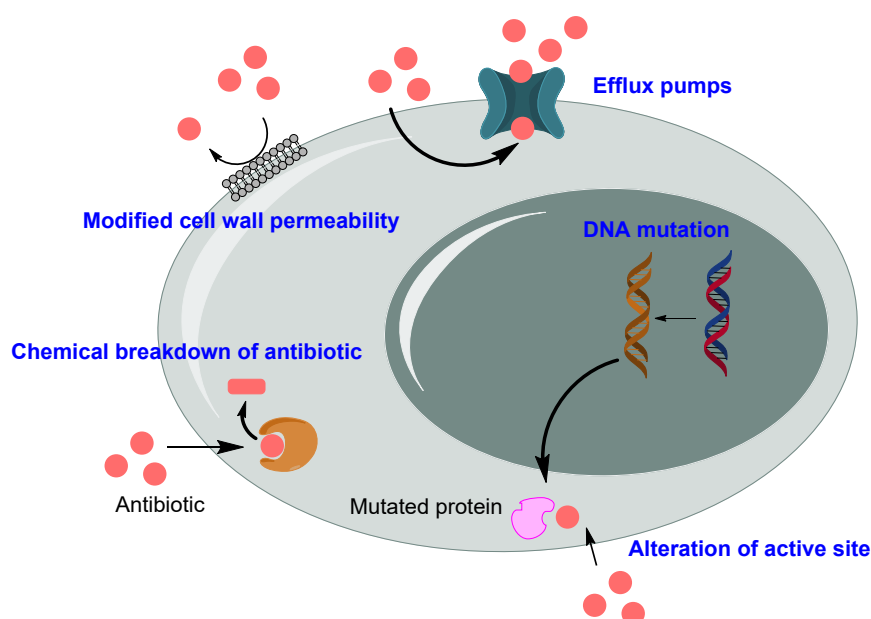


Figure 1.2: Mechanisms of antimicrobial resistance. Major mechanisms of antibiotic resistance are shown including: (a) alteration of the active site, (b) chemical breakdown of the antibiotic, (c) modifications to cell wall permeability and (d) efflux of antibiotics out of the cell.

1.1.3 Gram-negative bacteria

Gram-negative bacteria express an inner phospholipid bilayer and peptidoglycan layer similar to Gram-positive bacteria, but also possess a second outer membrane (**Figure 1.3**).²⁸ This outer membrane is comprised of lipopolysaccharides and is highly effective in limiting the entry of foreign substances, such as antibiotics into the cytosol.²⁹ Consequently, Gram-negative bacterial pathogens (e.g. *E. coli*) have reduced susceptibility to antibiotics due to this outer membrane, making them difficult to treat in a clinical setting. This reduced susceptibility and high levels of resistance to multiple frontline antibiotics has resulted in high morbidity and mortality rates for Gram-negative bacterial infections especially in developing countries.^{30, 31} MDR strains of *Neisseria gonorrhoea*, *Acinetobacter baumannii*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* (*P. aeruginosa*) and *E. coli* are of particular concern.³²

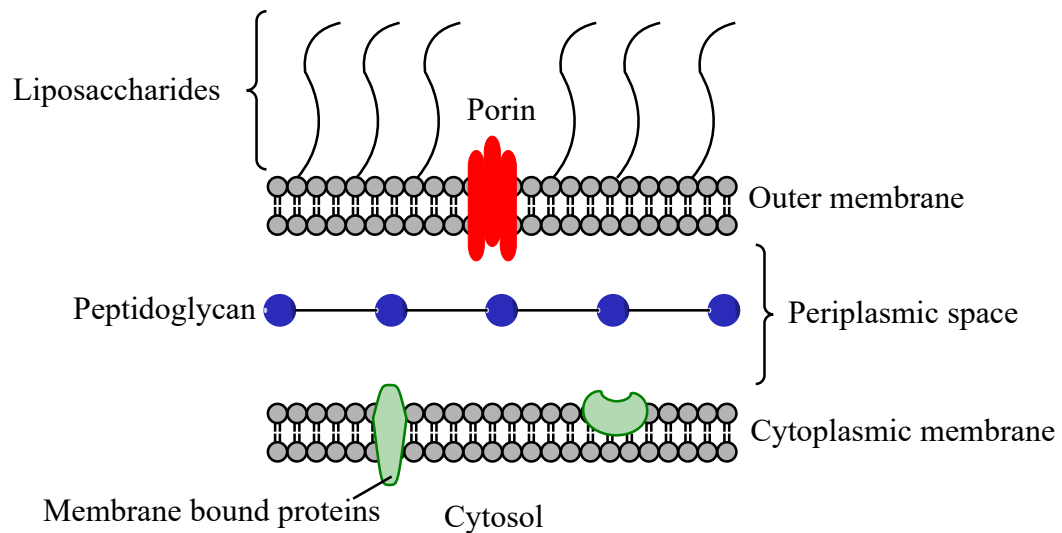


Figure 1.3: Composition of Gram-negative bacterial cell wall.

1.2 Bacterial virulence

Bacteria virulence or pathogenicity is the ability of bacteria to cause infection of the host, which is determined by an array of virulence factors.³³ These virulence factors are usually either membrane bound proteins that exist on the bacteria's cell surface or proteins that are secreted from the periplasm into extracellular space where they can act upon host cells to cause infection.³² Virulence factors are involved in a number of key processes including bacterial adhesion, toxin production and, the assembly of specialised secretion systems, that contribute to pathogenicity.³⁴

1.2.1 Bacterial adhesion factors

Bacterial adhesion is an essential process for establishing and maintaining an infection. Bacteria have developed a number of appendages including fimbriae and pili that enable them to adhere to host cells or tissues.³⁵ One type of adhesion appendage are type 1 pili that form a small 0.1 – 2 μM tail. In *E. coli* these pili are expressed on the cells surface and are composed of ≈ 300 copies of a fimbrial protein A (FimA) subunit that form the body of the appendage. The FimF, FimG and FimH protein subunits are expressed on the tip of the pili and confer the ability to bind to surface proteins expressed on host epithelial cells.³⁶ This binding event is the first step for *E. coli* in establishing a urinary tract infection in humans.³⁷

1.2.2 Bacterial toxins

Bacterial toxins can be classified as either endotoxins which are structures that remain a part of the bacteria or exotoxins which are secreted proteins that act upon host targets. These bacterial toxins are known to cause damage to host cells and organs. In Gram-negative bacteria, lipopolysaccharides both form a part of the outer cell wall and act as an endotoxin upon binding to lipopolysaccharide-sensitive host cells.³⁸ These toxins can also manipulate the host immune system to avoid detection and reduce the host's immune response to an infection.³⁹ Bacteria are able to avoid detection and cause disease by:

- (a) Blocking host neurotransmission pathways, e.g. botulinum toxin produced by *Clostridium botulinum* bacteria inhibits the release of acetylcholine from presynaptic neurons causing muscle paralysis of the lungs and death.⁴⁰
- (b) Directly damaging host cell membranes or immune responders, e.g. diphtheria toxin produced by *Corynebacterium* and anthrolysin O protein produced by *Bacillus anthracis* targets toll like protein 4 and causes macrophage apoptosis.⁴¹⁻⁴³
- (c) Avoiding/supressing the host immune response, e.g. *Staphylococcus aureus* hijack and interact with immunoreceptor tyrosine-based inhibitory receptors on host cells leading to upregulation of these proteins and evasion of the host immune response.^{39, 44}

1.2.3 Secretion systems

Bacteria have developed a number of specialised secretion systems to mediate the transport of toxins and other effector proteins into the cytosol of host cells. Type III secretion systems (TTSS) behave like a syringe, injecting proteins and toxins into the host cells causing tissue damage and cell lysis.⁴⁵ These TTSS are required by a variety of pathogens such as *Salmonella*, *Escherichia*, *Yersinia pseudotuberculosis* and *Pseudomonas* to cause infection.^{46, 47} Inhibition of these virulence factors has the potential to elicit an antivirulence effect and in cases where similar virulence factors are present in different pathogens may exhibit broad spectrum activity against many Gram-negative bacteria.⁴⁸

1.3 Targeting bacterial virulence

Targeting bacterial virulence and virulence factors is a promising strategy as it provides a rich source of novel targets to exploit. Current antibiotics, which target either bacterial growth or bacterial viability place high selection pressure on resistant bacteria which survive the antibiotic treatment and become the dominant phenotype.⁴⁹ This selection for resistant strains has fuelled the emergence of many MDR Gram-negative ‘superbugs’ as previously mentioned.³² Antibiotics targeting virulence rather than growth or viability are hypothesised to exert less selective pressure towards the development of resistant strains, since the ability to cause infection in a host does not confer any survival advantage to the bacteria. Targeting virulence may therefore be a promising strategy to slow down or even combat the emergence of antibiotic resistance.⁵⁰⁻⁵⁴ The use of virulence inhibitors is well documented in the literature and some examples are detailed below.

1.3.1 Targeting type II secretion systems

P. aeruginosa is a Gram-negative bacterial pathogen that is increasingly difficult to treat in a clinical setting due to MDR. High rates of antibiotic resistance are observed due to the intrinsically low outer membrane permeability of these bacteria, as well as the expression of multiple efflux pumps that actively transport antibiotics out of the bacterial cell.⁵⁵ Therefore, targeting virulence mechanisms may be a productive strategy to treat infection. *P. aeruginosa* have developed specialised secretion systems for injecting toxins into endothelial host cells. Zhang *et al.* developed a number of peptide inhibitors targeting type II secretion systems in *P. aeruginosa*⁵⁶ which demonstrated a significant reduction in virulence.

1.3.2 Bacterial adhesion as an antivirulence target.

E. coli are a major cause of bladder infections in humans. Adhesion to the bladder is the first essential step for establishing and maintaining an infection. This process is mediated by pili expressed on the bacterial cell surface. Inhibition of pili formation has attracted attention as a potential antivirulence

strategy. Pinkner *et al.* described the development of a series of bicyclic pyridones (**Figure 1.4**) as potent inhibitors of pilus biogenesis.⁵⁷ These analogues targeted the highly conserved chaperone-usher pathway that is essential in the assembly of bacterial pili and demonstrated that their activity was driven by binding to the chaperone protein PapD.

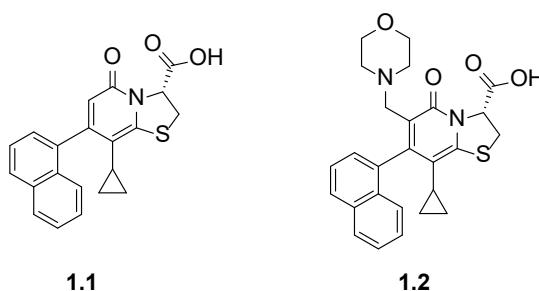


Figure 1.4: Chemical structures of bicyclic pyridones **1.1** and **1.2** as inhibitors of pilus biogenesis.

1.4 Disulfide bond forming proteins

Disulfide bonds (DSB) are present in many proteins that are exported from the bacterial cytoplasm, many of which are virulence factors. The formation of these DSB is essential for the correct folding, stability, function and activity of many of these proteins.^{33, 58, 59} In Gram-negative bacteria such as *E. coli*, *Burkholderia pseudomallei* and *Vibrio cholera*, the formation of DSB is mediated by a family of enzymes that comprise the DSB system. The DSB system consists of two pathways, an oxidative pathway that catalyses formation of DSB and an isomerisation pathway that catalyses reduction and recycling of incorrectly formed DSB. The *E. coli* K-12 strain has a well characterised DSB system. The oxidative pathway includes a highly oxidising periplasmic component, *E. coli* disulfide bond protein A (*EcDsbA*) and a membrane bound partner *E. coli* disulfide bond protein B (*EcDsbB*), which maintains *EcDsbA* in its oxidised form. DsbA is a dithiol-disulfide oxidoreductase enzyme that catalyses the formation of DSB of newly synthesised proteins that are located in the periplasm of Gram-negative bacteria (**Figure 1.5**).⁶⁰ DsbA is considered a master regulator of bacterial virulence and an attractive drug target as bacteria lacking a functional DsbA have been shown to be avirulent.^{61,}

⁶² The highly oxidising nature of *EcDsbA* is due in part to the stability of the reduced form over the oxidised form, which enables DsbA to introduce DSB into newly formed proteins efficiently.⁶³

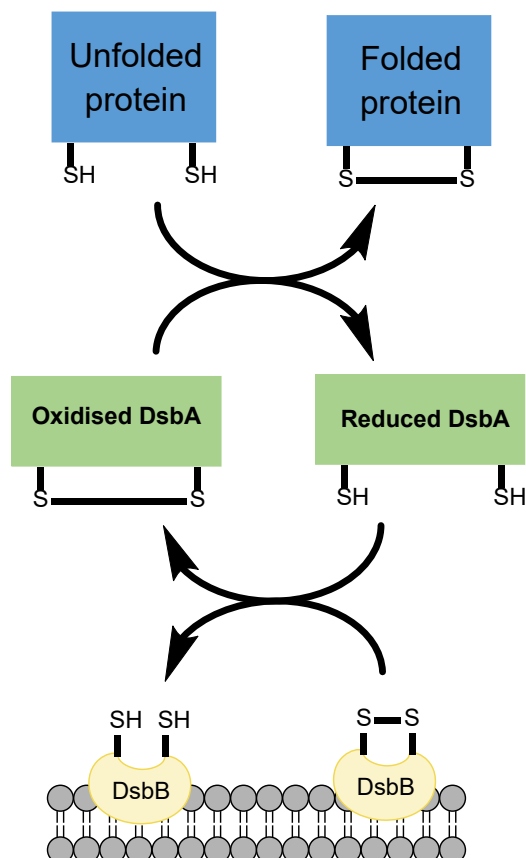


Figure 1.5: The oxidative cycle of DsbA. DsbA catalyses the formation of a disulfide bond in an unfolded substrate protein and is itself reduced in the process. DsbA is then reoxidised by membrane bound DsbB.

1.4.1 Structure of *EcDsbA*

The structure of DsbA consists of a catalytic thioredoxin domain, which is a common structural motif that is present in many oxidoreductase enzymes, as well as an inserted alpha helical domain (**Figure 1.6, a**).^{64, 65} The first DsbA X-ray crystal structure reported was that of *EcDsbA*.²³ The active site was identified as a pair of redox active cysteine residues contained within a $C_{30}P_{31}H_{32}C_{33}$ motif adjacent to a shallow hydrophobic groove, which is a conserved feature of many DsbA enzymes (**Figure 1.6, b**).⁶⁶ A number of other residues have been implicated in the activity, function and stability of DsbA,

including a conserved *cis*-proline residue (Pro151 in *EcDsbA*), which is adjacent to the redox active cysteine residues in the three dimensional structure (**Figure 1.6, b**).³³ Residues from a periplasmic loop in DsbB, as well as substrate peptides bind to DsbA in the shallow groove adjacent to the active site (**Figure 1.6, c**). Small molecule inhibitors binding within this groove could prevent the formation of the DsbA-DsbB complex and therefore may inhibit bacterial virulence.

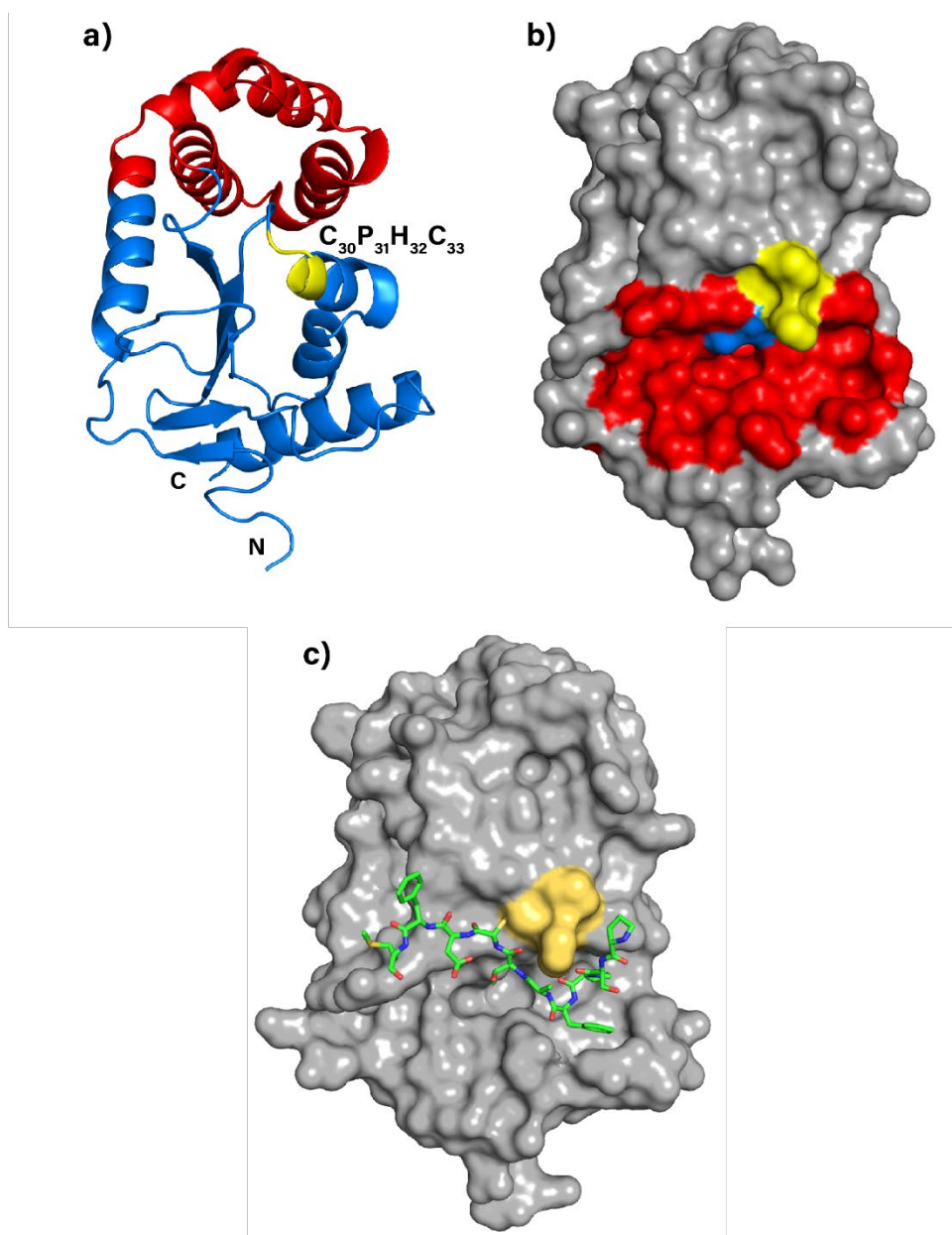


Figure 1.6: a) X-ray crystal structure (PDB code 1FVK) displayed as cartoon with domains highlighted. The active C₃₀P₃₁H₃₂C₃₃ residues are shown in yellow, the thioredoxin domain in blue and the inserted alpha helical domain in red. b) X-ray crystal structure of *EcDsbA* (PDB code 1FVK) shown as a surface representation in grey with the C₃₀P₃₁H₃₂C₃₃ residues shown in yellow, Pro151 is displayed in blue and the hydrophobic groove where substrates and DsbB are thought to bind displayed in red. c) Structure of C33A mutant of *EcDsbA* (grey surface with active site CPHA motif shown as yellow surface) with a periplasmic loop of *EcDsbB* (PDB code 2HI7, residues P₉₈S₉₉P₁₀₀F₁₀₁A₁₀₂T₁₀₃C₁₀₄D₁₀₅F₁₀₆M₁₀₇, shown as greens sticks) binding in the hydrophobic groove of *EcDsbA*.⁶⁵

1.4.2 Biological evaluation of the DsbA as a drug target

Many bacteria including *E. coli*, which lack a functional DsbA have been shown to have attenuated virulence and reduced capacity to cause infection.^{32, 67} This attenuated virulence is attributed to the incorrect folding of disulfide containing proteins, which is accompanied by a range of phenotypes, including preventing bacterial adhesion, which thereby prevents that bacteria from establishing an infection. This prophylactic-like scenario of inhibiting the establishment of an infection may not be useful in a clinical setting where treatment with an antimicrobial agent will commence after the infection is established. More recent data from a study using a conditional knockout of DsbA expression in a uropathogenic *E. coli* (UPEC) strain CFT073 provides more compelling evidence for the effectiveness of DsbA inhibitors.[M. Totsika, unpublished work] The study used a strain of *E. coli* that is only able to express DsbA in the presence of anhydrotetracycline (AHT). A starter culture is grown in the presence of AHT and this is used to infect mice in order to establish a urinary tract infection. After infection, the AHT is washed out and the bacteria no longer produce DsbA. Thus, the infection is established in the presence of an active DsbA, but its activity is lost on AHT washout. This study more accurately replicates infection in a clinical setting where the infection is established prior to administration of the anti-infective agent. Twenty four hours post infection the mice infected with the strain expressing AHT-induced DsbA showed significantly lower titres than those infected with wild type CFT073 in samples recovered from the mouse bladder. This study indicates that DsbA inhibitors would be able to alter the course of a bacterial infection and may be effective in treating bacterial infections. Furthermore, targeting bacterial virulence effectively disarms the bacteria's weaponry (toxins, secretion systems, adhesion mechanisms) rather than directly killing them, allowing the hosts immune system to clear the infection. This way of targeting the bacteria is thought to exert less selective pressure and may potentially lead to decreased rates of resistance development, extending the lifetime of a potential drug.^{32, 50} However, this idea still remains to be proven in a clinical setting.

1.4.3 Chemical approaches to inhibit Dsb enzymes

Research has been reported that describes the design and synthesis of inhibitors of both *EcDsbA* and *EcDsbB*. A high throughput screen of 51,487 compounds found 11 irreversible *EcDsbB* inhibitors targeting the DsbB-DsbA complex *in vitro*.⁶⁸ The compounds were found to covalently modify Cys44 of DsbB, which is an essential residue for catalytic activity of *EcDsbB*. The pyridazinone core highlighted in **1.3** (**Figure 1.7**) was common amongst the 11 hits. Commercially available analogues were purchased and tested. The purchased analogue **1.4** was found to be 23 times more potent than the parent **1.3** and showed broad spectrum DsbB activity against an array of other bacterial strains.

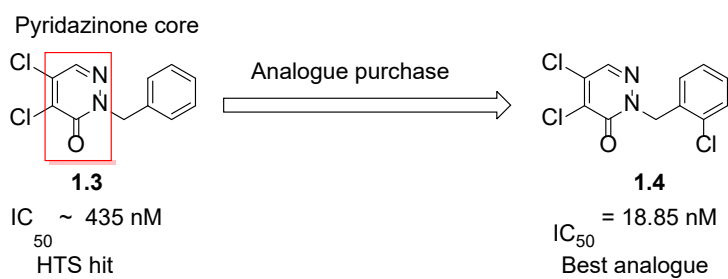


Figure 1.7: Chemical structures of two irreversible *EcDsbB* inhibitors. Left is the parent molecule **1.3** while right is the best related analogue purchased **1.4**.³²

Other literature describes irreversible *EcDsbB* inhibitors targeting the ubiquinone binding site.⁶³ The initial ubiquinone-like hit **1.5** was a weakly binding fragment molecule that was identified from screening a fragment library using a redox sensitive fluorescence assay (**Figure 1.8**). Rational drug design and optimisation produced compound **1.6** as an irreversible *EcDsbB* inhibitor ≈ 1000 times more potent than the initial hit **1.5**.⁶³ Covalent binding of cyclohexenone derivative **1.6** to *EcDsbB* was detected via electrospray ionisation mass spectrometry. Further studies indicated that **1.6** was selective for one of four essential cysteine residues (Cys130) in *EcDsbB*, but showed little selectivity for *EcDsbB* over *EcDsbA* making it a non-selective dual acting irreversible inhibitor of both *EcDsbA* and *EcDsbB*. Irreversible drug inhibitors are becoming increasingly popular due to their increased

residence time on protein targets and therefore potential increased longevity of inhibition of the target.⁶⁹

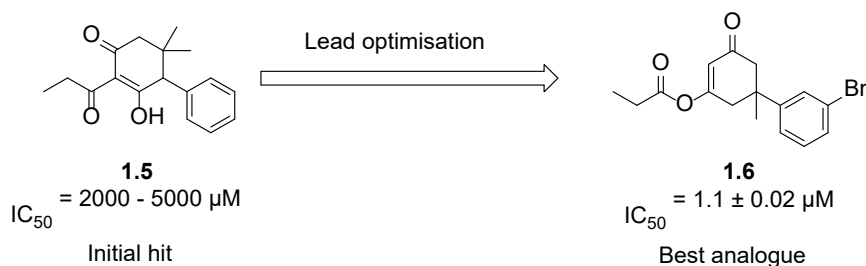


Figure 1.8: Chemical structures of two irreversible *EcDsbA* and *EcDsbB* inhibitors. Left is the weakly binding initial ubiquinone like hit **1.5**. Right is the optimised compound cyclohexenone derivative **1.6**.

1.5 Fragment-based drug discovery (FBDD)

Over the past 20 years fragment-based drug discovery (FBDD) has grown into a widely used method for finding starting points for drug development campaigns.^{70, 71} A fragment screening campaign generally begins with a small library of 500 - 5,000 low molecular weight (usually around 200 Da) compounds that are screened against a target.^{50, 72-74} Due to their small size fragments more efficiently sample chemical space,⁷² which increases the likelihood of finding a hit relative to screens using larger compounds.^{75, 76} Fragments however, form fewer, but efficient interactions with their target, hence fragments generally bind with low affinities typically having equilibrium dissociation constants $K_D > 500 \mu M$.⁷⁷ Consequently, fragment screening requires a variety of specialised methods that are able to detect weak binding including nuclear magnetic resonance (NMR), surface plasmon resonance (SPR) and X-ray crystallography.⁷⁷

1.5.1 Nuclear magnetic resonance (NMR)

NMR is a powerful tool widely used in fragment screening and in the characterisation of hits and developed compounds due to its reliability and sensitivity. A number of ligand-observed NMR

experiments such as Saturation-transfer difference (STD), WaterLOGSY and T2 relaxation-weighted Carr-Purcell-Meiboom-Gill (CPMG) are commonly employed for screening fragment libraries.⁷⁸⁻⁸¹ These NMR experiments are generally performed in mixtures of multiple fragments to increase throughput and are described below.^{77, 82-84}

1.5.1.1 Saturation-transfer difference (STD)

In an STD experiment the protein/target is selectively excited at a unique frequency where only protein signals but no ligand signals are observed (e.g. < 0 ppm). This magnetisation can then be transferred to ligands that bind to the protein through the nuclear Overhauser effect (NOE), which leads to a reduction in the intensity of resonances of bound ligands.⁸⁵ A second comparison spectrum is also acquired in which the protein is not selectively magnetised. A difference spectrum of these two experiments is calculated and only signals from bound ligands will be observed in the resulting ^1H NMR difference spectra (**Figure 1.9**).

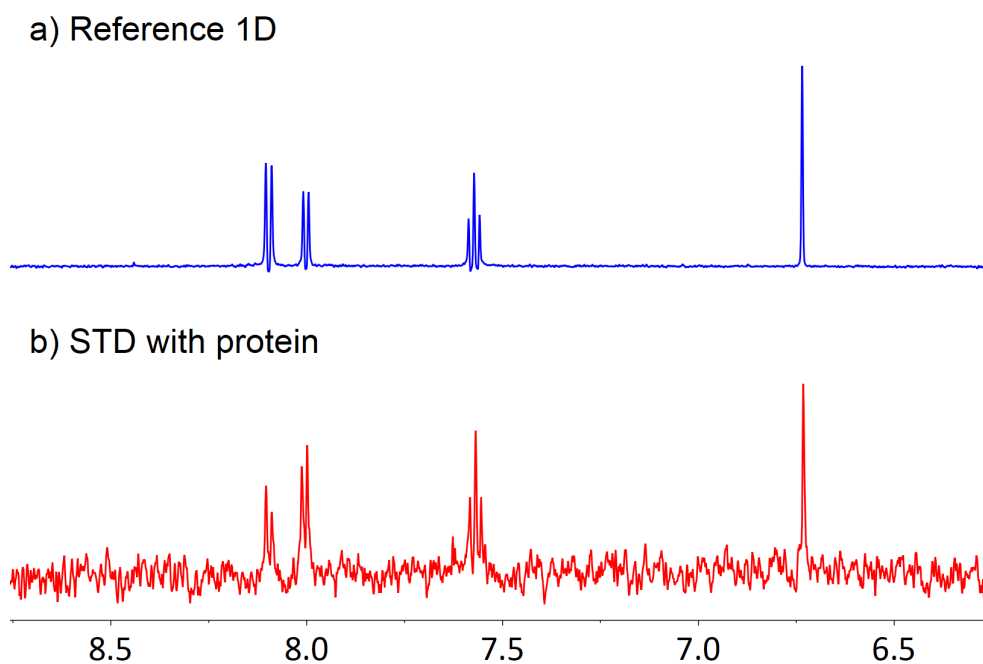


Figure 1.9: STD ^1H NMR spectra. Depicted in panel a) is a reference 1D ^1H NMR spectrum of the ligand, b) difference STD NMR spectrum, the presence of ligand signals in the difference spectrum is indicative of ligand binding to protein.

1.5.1.2 WaterLOGSY

WaterLOGSY experiments make use of NOE transfer from the bulk solvent (water) to the protein and bound ligand.^{79, 86} Free ligands tumble rapidly in solution and give negative NOE signals, whereas, large protein-ligand complexes tumble more slowly and give positive NOE signals. By comparing WaterLOGSY spectra that are recorded in the absence and presence of the protein it is possible to identify binders from a change in the intensity or the sign of the observed NOE (**Figure 1.10**).

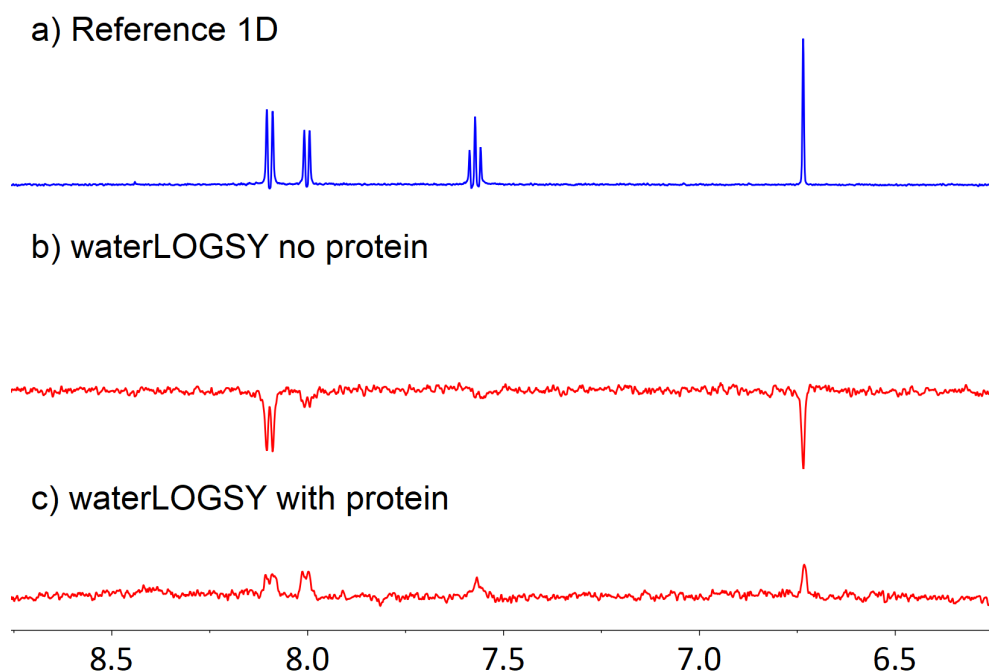


Figure 1.10: WaterLOGSY NMR spectra with a) reference ^1H NMR spectrum of the ligand b) wLOGSY spectrum of the ligand in buffer and c) wLOGSY spectrum of the ligand plus protein, showing a change towards positive signal intensity indicative of ligand binding.

1.5.1.3 Carr-Purcell-Meiboom-Gill (CPMG)

Small molecules in solution tumble quickly and have long T_2 relaxation delays, while larger proteins or protein-ligand complexes have shorter T_2 relaxation delays.⁷⁷ Relaxation filtered 1D NMR experiments (such as CPMG) take advantage of this and apply a relaxation delay to reduce the

intensity of signals resulting from the protein-ligand complex. Hence ligands that bind to the protein show a reduction in ligand signal intensity in the presence of the protein (**Figure 1.11**).

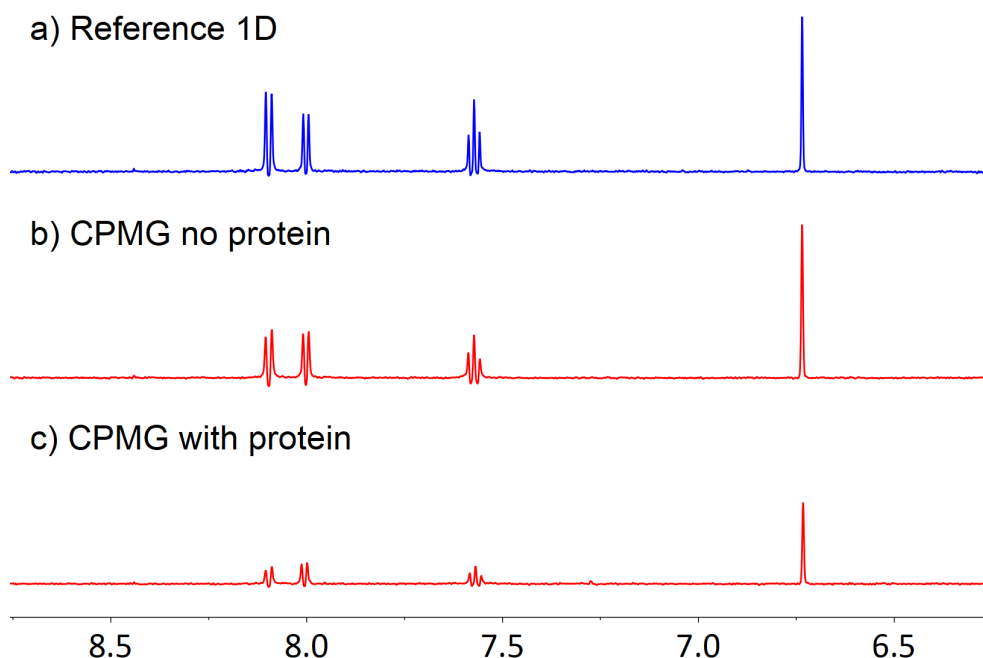


Figure 1.11: CPMG NMR spectra with a) reference ^1H NMR spectrum of the ligand, b) CPMG excited ligand spectrum and c) ligand plus protein CPMG spectrum. A reduction in ligand signal intensity from panel b to c indicates ligand binding.

1.5.1.4 Protein detected NMR

Protein detected NMR methods such as ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) experiments with ^{15}N labelled protein targets can also be employed to detect ligand binding. The chemical shifts of peaks are directly influenced by the electronic environment surrounding them, hence small molecule binding to a protein's surface can alter this environment and this is observed as changes in the position of peaks which are measured as chemical shift perturbations (CSP). ^1H - ^{15}N HSQC experiments measure the CSP of backbone and sidechain amide residues in the absence (APO) and presence of a ligand (HOLO). Both the APO and HOLO HSQC spectra are compared (**Figure 1.12, a**) and CSP are calculated using **equation 1.1**.^{50, 87}

Equation 1.1: Calculation of CSP from change in chemical shift of ^1H and ^{15}N dimensions.

$$\text{CSP} = \sqrt{\Delta H^2 + (0.2 \times \Delta N)^2}$$

Where ΔH and ΔN are the change in chemical shift of ^1H and ^{15}N dimensions of the HSQC spectra.

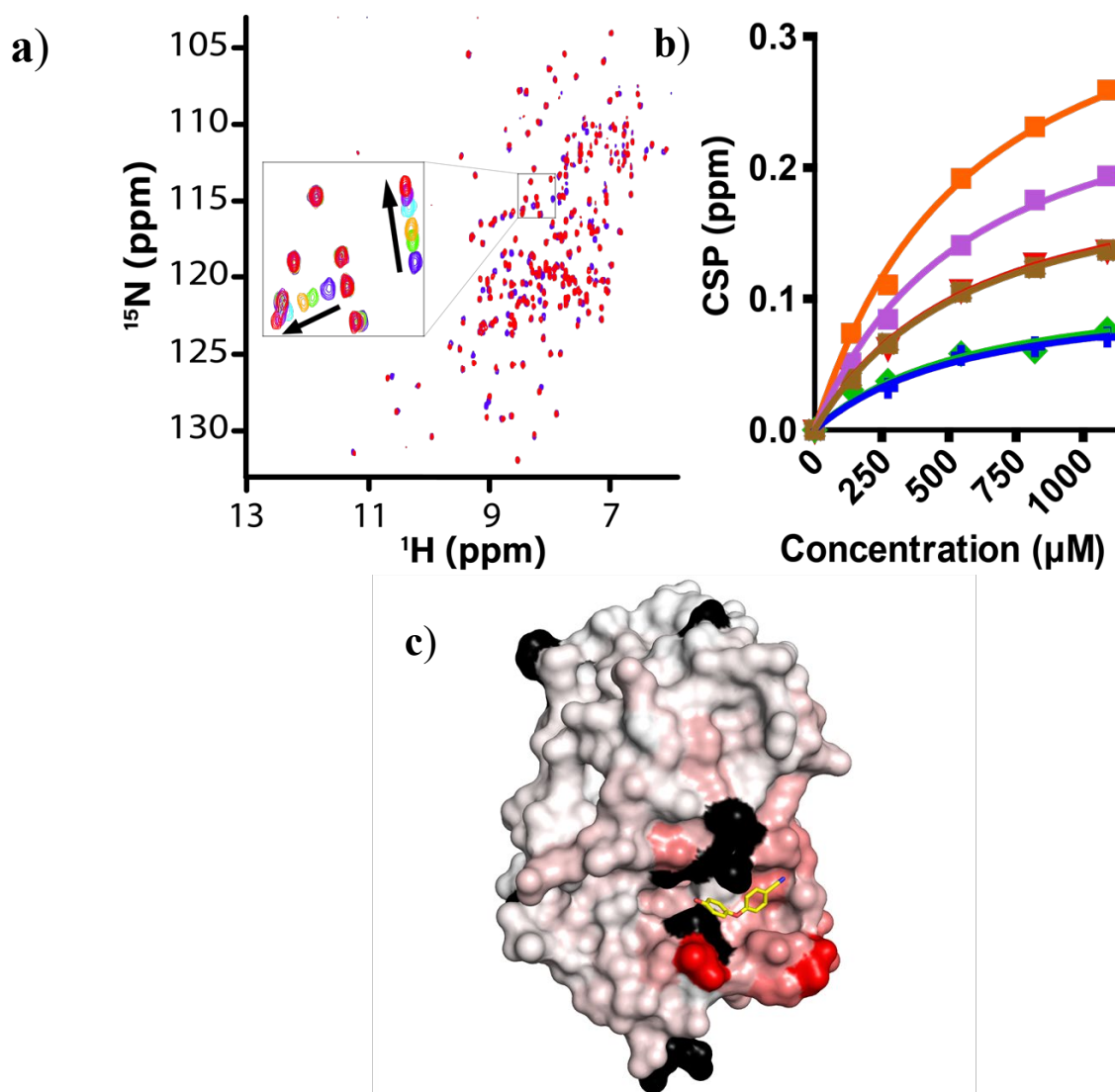


Figure 1.12: a) Analysis of ligand binding to *EcDsbA* by ^1H - ^{15}N HSQC. Spectra were recorded in the presence of ligand at 5 different concentrations. b) HSQC titration plotting ligand concentration (μM) vs CSP (ppm). c) CSP quantified and mapped onto the surface of *EcDsbA* (PDB: 1FVK). Missing residues are coloured in black, perturbed residues are shown in white to red colour gradient correlating to larger perturbations.

The magnitude of a CSP increases in proportion to the occupancy of the binding site, and for 1-site binding recording CSP as a function of ligand concentrations can be used to ascertain binding affinity as a dissociation constant (K_D). As it is often necessary to use relatively high protein concentrations in HSQC experiments, the CSP data are commonly analysed using a single site binding model with ligand depletion as shown in **Equation 1.2**⁸⁸ and applied in **Figure 1.12, b**. If the structure of the protein is known and amide residues assigned, mapping of perturbed residues to the proteins surface can reveal the ligand binding site (**Figure 1.12, c**).

Equation 1.2: Single site binding model equation with ligand depletion

$$CSP_{obs} = DMAX \times \frac{([P] + [L] + K_D) - \sqrt{([P] + [L] + K_D)^2 - (4[P][L])}}{2[P]}$$

Where [P] is total protein concentration, [L] is total ligand concentration, CSP_{obs} is the measured CSP and DMAX is the CSP at 100 % occupancy, which is the extrapolated from the measured CSP.

1.5.1.5 Measurement of ligand solubility

Testing of fragments commonly requires relatively high concentrations (100 – 1000 μ M) of ligand, as fragments generally bind weakly to their target protein ($K_D > 100 \mu$ M) in the early stages of development. Testing at these concentrations brings about concerns of solubility and aggregation in solution. Small molecules can form aggregates in solution invisible to the naked eye that non-specifically bind to proteins and give promiscuous readouts in a multitude of assays.⁸⁹⁻⁹¹ This aggregation can be detected by quantitative 1D 1H NMR.⁹² These NMR spectra are acquired in the presence of known concentration of an internal standard such as 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), which allows for quantitative measurement of the amount of ligand in solution. Additionally, aggregation can manifest as concentration dependant chemical shift changes in the 1D spectra, hence fragments are commonly tested at several concentrations.

1.5.2 Surface plasmon resonance (SPR)

SPR is a biophysical technique that allows for direct measurement of binding events between ligands and proteins. Typically, the protein of interest is immobilised on a thin gold film on the surface of a biosensor chip using one of a number of different approaches. These include biotinylating the protein for capture on a streptavidin surface or using a His-tag on the protein to capture on a Ni^{2+} /NTA surface.^{93, 94} The bottom of the gold surface is irradiated with a beam of polarised light under conditions of total internal reflection. This generates an evanescent wave, which allows plasmons to interact with the surface. The intensity and angle of light reflected from the surface is sensitive to the refractive index of the surface – which is in turn dictated by the mass attached to the surface. This means that the refractive index is highly sensitive to changes at the chip's surface such as those caused by ligand binding events (**Figure 1.13**). The change in refractive index is measured as a change in either the angle or the intensity of reflected light and is measured in arbitrary “response units” or RU. Measurements are made in real time which can be used to characterise both the kinetics (k_{on} and k_{off}) and affinity of binding (K_D).⁹⁵

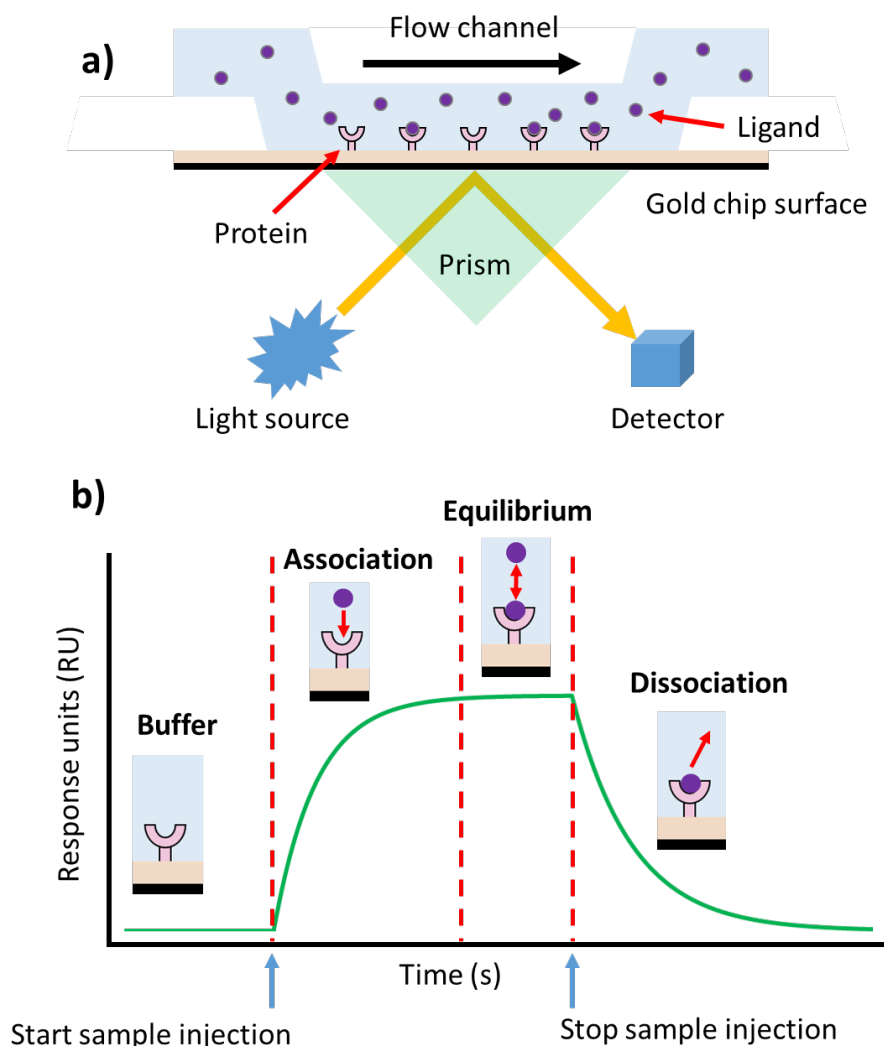


Figure 1.13: a) SPR detects changes in refractive index of polarised light reflected off a gold surface. These changes are as a result of protein-ligand interactions. b) SPR sensorgram depicting the 4 phases of ligand binding, namely: baseline signal, association phase, equilibrium phase and dissociation phase.⁹⁶

1.5.3 X-ray crystallography

X-ray crystallography is a valuable method in FBDD for elucidating structural information of the binding mode of fragments. Analysis of fragment optimisation campaigns identified that between 52 and 70 % of fragment optimisations used structural information derived from X-ray crystallography.^{97, 98} Furthermore, the number of X-ray crystal structures in the protein data bank (PDB) has increased exponentially since the early 2000's. Currently the PDB houses over 135,000 X-ray crystal structures

for a variety of biological targets.⁹⁹ This emphasis on structural data was amplified by the introduction of structure based drug design (SBDD) in the early 1990's and its widespread adoption in the 2000's.^{100, 101} X-ray crystallography plays a vital role in SBDD and is widely considered the “gold standard” for acquiring structural data.^{102, 103} Crystallography is capable of providing information on complexes that interact over a very broad range of affinities from weakly binding fragment hits to tighter binders such as lead-like or drug-like analogues.¹⁰⁴ Hence, the application of X-ray crystallography in FBDD is widespread and many of the methods for fragment development (as described below) are generally guided by structural data. In addition, recent advances in collection and processing of X-ray data have enabled higher throughput, acquisition and analysis of crystallographic data enabling X-ray screening of entire fragment libraries.^{105, 106}

1.6 Ligand efficiency

Ligand efficiency (LE) is a useful metric for ranking compounds based on their affinity with respect to molecular weight (MW) or heavy atom count (HA) and can be calculated using **Equation 1.3**.¹⁰⁷ LE accounts for both a ligand's affinity and size (HA) in search of the best fragment for elaboration.

Equation 1.3: Calculation of ligand efficiency $R = 1.986 \times 10^{-3} \text{ kcal K}^{-1} \text{ mol}^{-1}$, T is temperature in K.

$$LE = \frac{-\Delta G}{HA} = \frac{-RT \ln(K_D)}{HA}$$

Where ΔG is the free energy of binding and HA is heavy atom count with the free energy of binding ΔG often calculated using the affinity (K_D) of a given ligand.

Studies have suggested that optimal development of a fragment can maintain LE throughout the FBDD campaign.¹⁰⁸ A number of additional indices have been proposed similar to LE such as lipophilic ligand efficiency (LLE), see **Equation 1.4**.

Equation 1.4: Equation to calculate LLE

$$\text{LLE} = -\log(\text{IC}_{50}) - \text{LogP}$$

Where LogP is the logarithm partition coefficient between octanol and water.¹⁰⁹

LLE provides a quantitative measurement of affinity vs lipophilicity for optimising drug development. High lipophilicity in drugs has been associated with non-specific binding and off target effects resulting in side effects such as liver toxicity.¹¹⁰⁻¹¹³ Tracking metrics such as LE and LLE allow for monitoring of drug like properties while efficiently elaborating weak binding fragments into higher affinity leads and drugs.¹¹⁴

1.7 Fragment elaboration

The optimisation of fragments usually requires extensive chemical elaboration to build them into lead like and eventual drug like candidates. This process is generally guided by structural data provided by X-ray crystallography or NMR. Fragment development commonly uses three main strategies including: fragment growing, linking and merging.

1.7.1 Growing

Fragment growing is the most widely used method in FBDD and involves expanding a fragment, through the iterative addition of new motifs to form new ligand-target interactions. This involves multiple rounds of design, docking, synthesis and evaluation in order to improve binding affinity and, identify leads with good LE and drug like physicochemical properties.

Vemurafenib is the first FDA approved drug derived from a fragment hit and was developed using a fragment growing strategy.¹¹⁵ Vemurafenib is marketed as a cancer therapeutic for late stage BRAF^{V600} mutant melanoma.^{116, 117} The 7-azaindole fragment **1.7** was identified from a biochemical screen using Plexxikon's fragment scaffold library against kinase Pim-1 (**Figure 1.14**). This compound demonstrated weak binding affinity ($\text{IC}_{50} > 200 \mu\text{M}$) and low kinase specificity. Early

work identified the 3-substituted azaindole **1.7** with greater potency in a biochemical assay ($IC_{50} = 100 \mu M$). Further SBDD of this analogue identified the key lead compound PLX4720 (**1.9**) with an IC_{50} of 13 nM and high target selectivity for the mutant BRAF over the wild type kinase. Further optimisation to improve the preclinical properties of PLX4720 led to the development of Vemurafenib **1.10**, which is a potent ($IC_{50} = 30 \text{ nM}$) and selective BRAF^{V600} inhibitor.

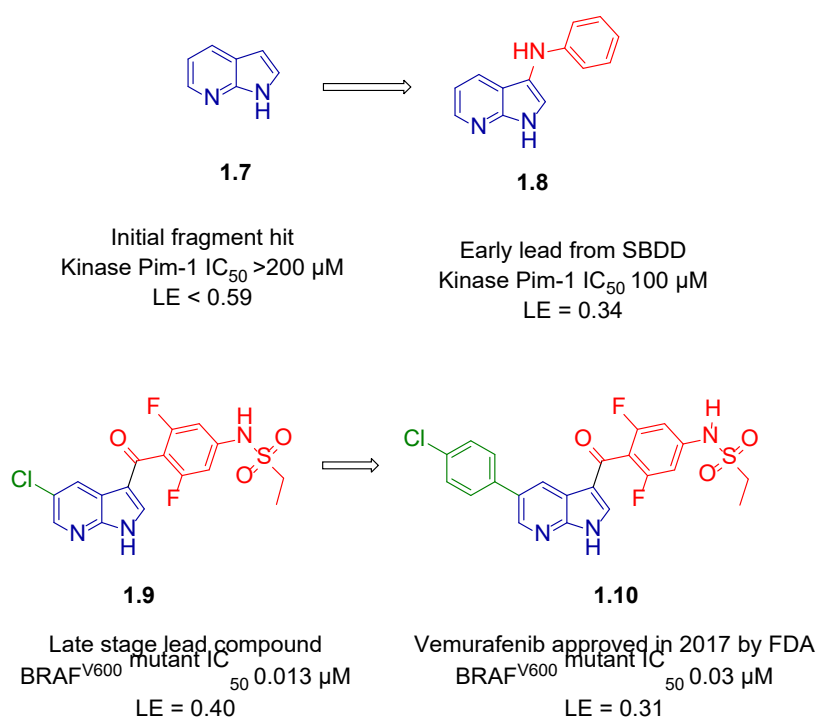


Figure 1.14: Key steps in the development of Vemurafenib, which is the first FDA approved drug derived from a fragment.¹¹⁵

1.7.2 Merging

Fragment merging can be utilised when two fragments demonstrate overlapping, but distinct modes of binding within the same pocket. The merging strategy can combine the beneficial interactions of two fragments to achieve a higher affinity compound, this process is aided by structures of the bound fragments being available. Fragment merging was exemplified in the development of inhibitors of the ethionamide transcriptional repressor (EthR) enzyme from *Mycobacterium tuberculosis*.¹¹⁸ Nikiforov *et al.* identified a number of fragment molecules that crystallised within the same long

internal hydrophobic cavity of EthR (**Figure 1.15, a**). Amide fragment **1.11** and amine fragment **1.12** both bound within this long hydrophobic cavity and their mode of binding was confirmed in X-ray crystal structures. The subsequent merging of these fragments identified analogue **1.13** ($K_i = 35 \mu\text{M}$, **Figure 1.15, b and c**) which showed a 3 – 8-fold higher affinity for EthR than either of the individual fragments (**Figure 1.15, d**).

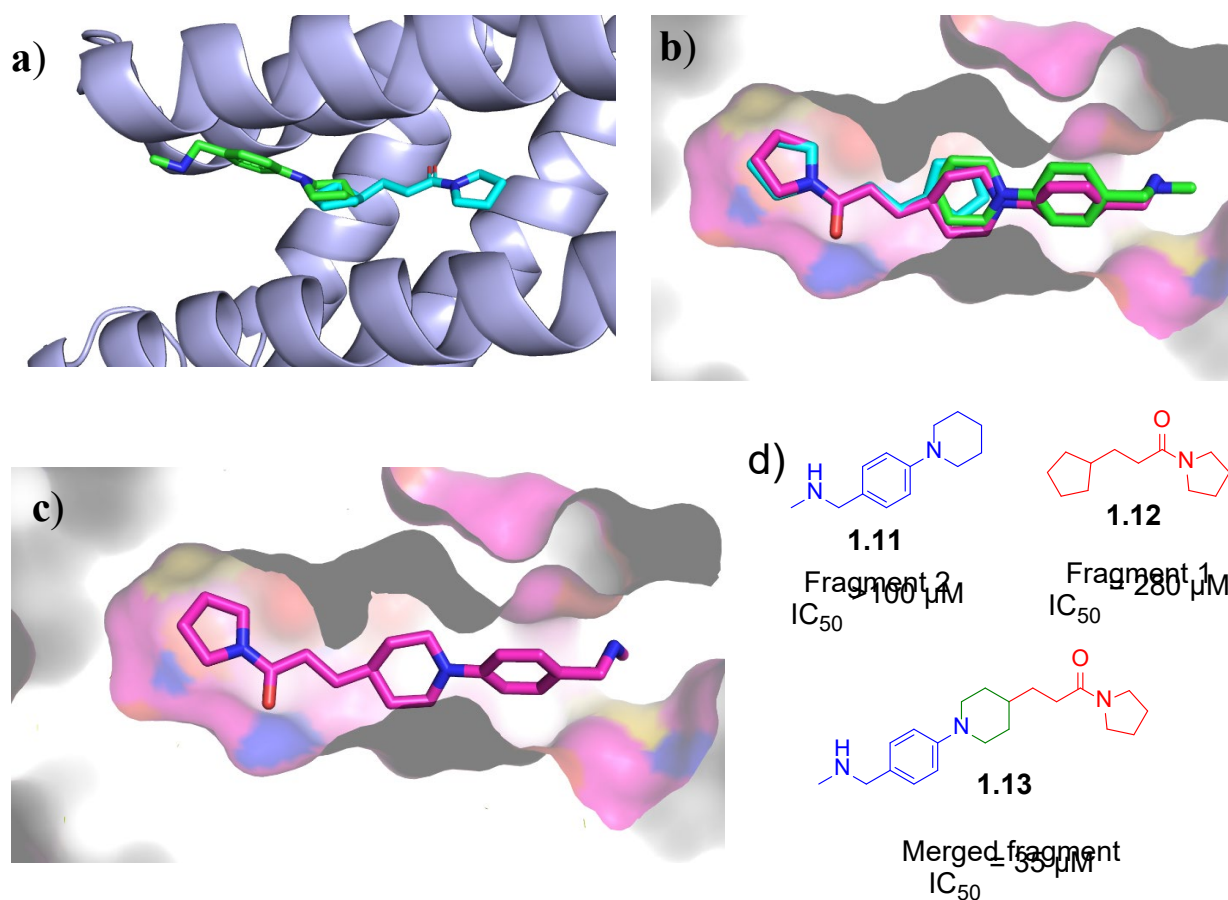


Figure 1.15: a) X-ray crystal structure of amine **1.11** (green sticks) and amide **1.12** (blue sticks) bound to the EthR enzyme (tinted blue cartoon, PDB code: 5F1J and 5F27 respectively). b) X-ray crystal structure of the merged analogue **1.13** (purple sticks) bound to EthR enzyme (purple surface, PDB code: 5EYR). Overlaid are the two fragments **1.11** and **1.12** shown in green and blue sticks, respectively. The proteins surface is coloured by element and has been sliced to reveal the internal cavity. c) Crystal structure of the merged compound **1.13** using the same format and view as panel b. d) 2D structures of amine **1.11**, amide **1.12** and merged analogue **1.13**. K_i values were determined by an SPR assay.¹¹⁸

1.7.3 Linking

A more difficult and less common applied strategy in elaborating fragments is linking. Fragment linking involves tethering two individual fragments that bind in adjacent pockets together via a linker. This strategy is an attractive method due to the potential for large increases in affinity.¹¹⁹ In principle, if the free energy of binding for each fragment is maintained in the linked compounds, the affinity is the product of the two individual binding affinities e.g. 2 fragments that bind with $K_D = 100 \mu\text{M}$, can produce a linked compound with $K_D = 10 \text{ nM}$ if the linker does not bind or interfere with each individual fragment binding. The linker can also contribute to binding, which can produce a superadditivity effect. Superadditivity in FBDD refers to a linked fragment demonstrating a binding free energy that exceeds the sum of the two individual fragments.^{119, 120} However, generally to achieve success in fragment linking, it must be possible to link the fragments without impacting the mode of binding that they adopt before linking. Steric and geometric considerations mean that this is often difficult to achieve. Similarly, this method is heavily guided by X-ray crystallography or NMR structures. This structural information assists in selecting the optimal linker length and the nature of chemical attachment. Sub-optimal linkers that alter the binding mode of either individual fragment can lead to reduced affinity. Interactions through the linker itself are possible further increasing complexity of the system, hence few examples of successful fragment linking are available in the literature.¹¹⁹ A success story implementing fragment linking was achieved by Fusco *et al.* in developing a potent Casein kinase 2 (CK2) inhibitors.¹²¹ Prior optimisation by SBDD identified two CK2 inhibitors benzylamine **1.14** and carboxylic acid **1.15** that bind within two distinct adjacent sites (**Figure 1.16, a**), the αD and ATP binding sites respectively. With the help of X-ray crystal structures these two fragments were linked together via a diamide linker (**Figure 1.16, b**). The resulting linked fragment **1.15** demonstrated far greater affinity for CK2 than either individual fragments ($K_D = 0.32 \mu\text{M}$, **Figure 1.16, c**).

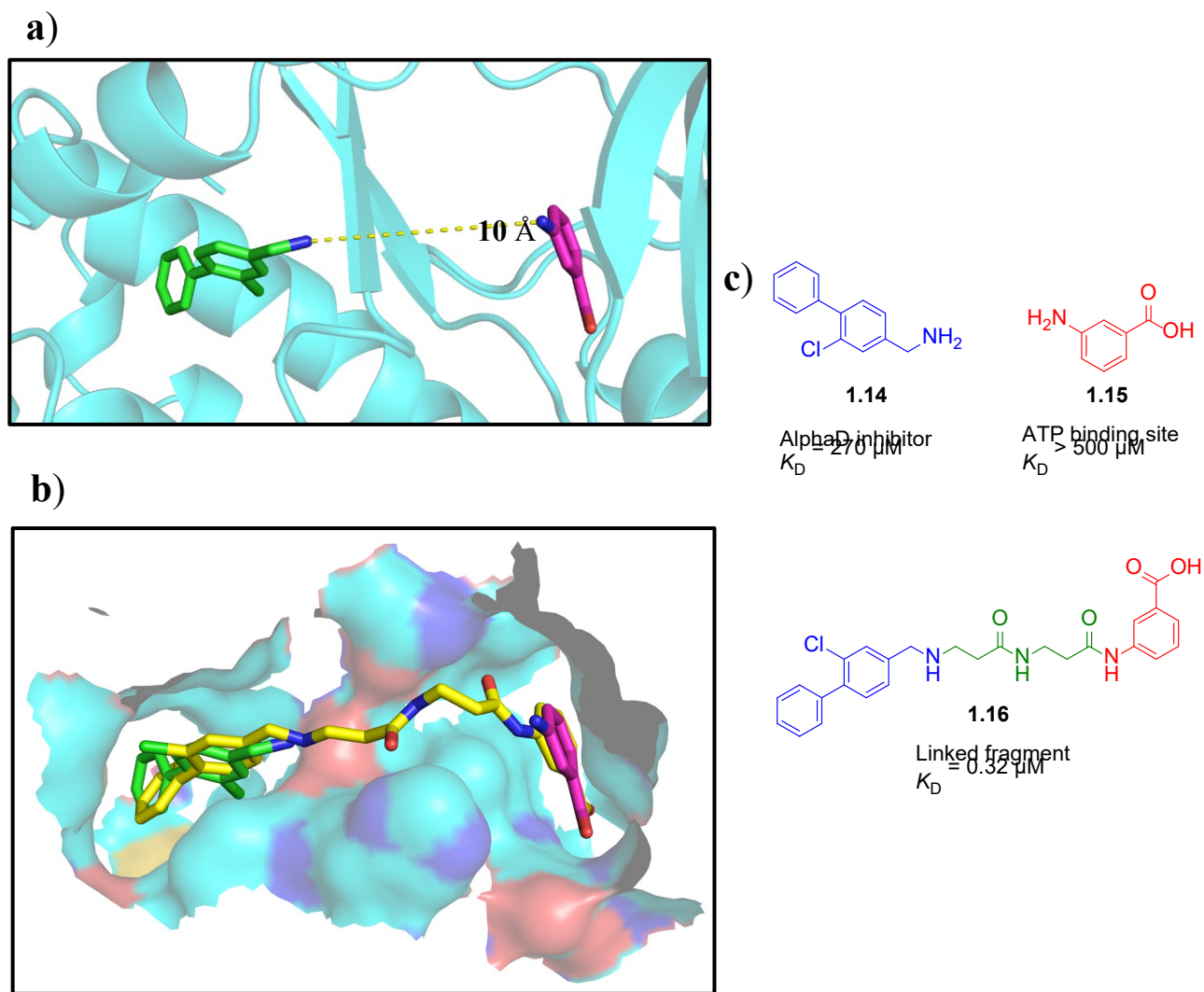


Figure 1.16: a) Fragment molecules benzylamine **1.14** (green sticks) occupying the α D binding pocket and carboxylic acid **1.15** (purple sticks) bound within the adjacent ATP binding site overlaid onto the X-ray crystal structure of the CK2 protein (blue cartoon, PDB code: 5CSP and 5CHS). b) The linked analogue **1.16** (yellow sticks) spans both binding sites and adopts a similar binding mode to the individual fragments (purple and green sticks) bound to CK2 protein (blue ssurface, PDB code:5CU4). c) Structures of the 3 compounds **1.14**, **1.15** and **1.16**

1.8 Efficient elaboration of fragments - Off rate screening

The standard paradigm in medicinal chemistry programs is to design, synthesise, purify and characterise target compounds and then test their binding and/or activity. Purification of synthetic compounds in medicinal chemistry is often time consuming, uses significant amounts of solvent and

can suffer from poor material recovery.¹²² Methods that enable testing of the desired reaction products prior to purification would circumvent this issue and allow for more efficient development of higher affinity analogues. Recently a technique called Off-rate screening (ORS) by SPR was described, which is able to achieve this.¹²³ SPR allows for the measurement of binding interaction kinetics and in ideal cases enables the characterisation of association and dissociation rate constants (**Figure 1.17**). The affinity (K_D) can be determined from the ratio of association and dissociation rate constants as described in **Figure 1.17**. For a series of chemically similar small molecules that bind at the same site on a protein, it is often the case that the association rate constant is controlled by diffusion and that changes in affinity are largely dictated by changes in the dissociation rate constant. Higher affinity compounds will therefore dissociate more slowly from the binding site. ORS by SPR takes advantage of this by monitoring the dissociation phase in the sensorgram of unpurified reaction products. In a mixture containing two components that bind with 10-fold difference in affinity, the dissociation phase of the sensorgram is dominated by the more tightly binding component, even where it is present as a relatively low fraction of the mixture (**Figure 1.18**).¹²³ As the off rate is independent of ligand concentration it can be estimated without measuring the amount of each component that is present in the mixture.¹²³ This means that in a chemical reaction where a starting molecule is converted into a new product, it is possible to identify products that show slower dissociation in the SPR sensorgram than the starting material, without either purification or measurement of the concentration of the product. Additionally, SPR requires only a small amount of sample making it possible to conduct reactions on a small scale and evaluate them without further purification by SPR to identify those which show slower off rates. Consequently, reactions can be performed on microscale in a plate-based format, which allows multiple reactions to be conducted in parallel.

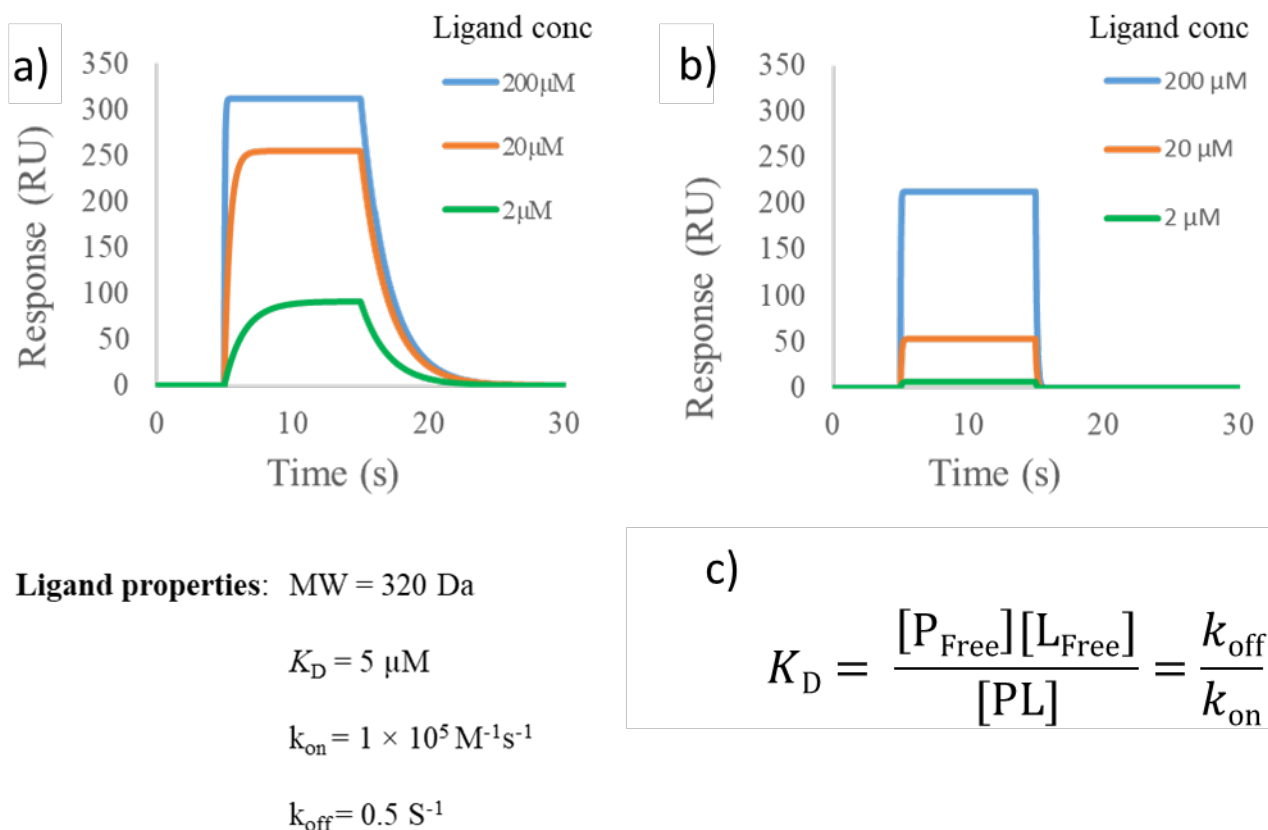


Figure 1.17: a) Theoretical SPR sensorgram of a compound of MW = 320 Da assuming 1 RU per Da at concentrations of 200, 20 and 2 μM of compound and affinity (K_D) of 5 μM with $k_{\text{on}} = 1 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ and $k_{\text{off}} = 0.5 \text{ s}^{-1}$. b) Theoretical SPR sensorgram of another compound with MW = 320 Da assuming 1 RU per Da at concentrations of 200, 20 and 2 μM of compound and affinity (K_D) of 100 μM with $k_{\text{on}} = 1 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ and $k_{\text{off}} = 10 \text{ s}^{-1}$. c) The relationship between equilibrium dissociation constant and kinetic rate constants for a 1-site binding interaction is described in the equation. Where K_D is the ligands affinity, P_{Free} is the concentration of free protein, L_{Free} is the concentration of the ligand, PL is the concentration of protein-ligand complex, k_{off} is the dissociation rate constant and k_{on} association rate constant.

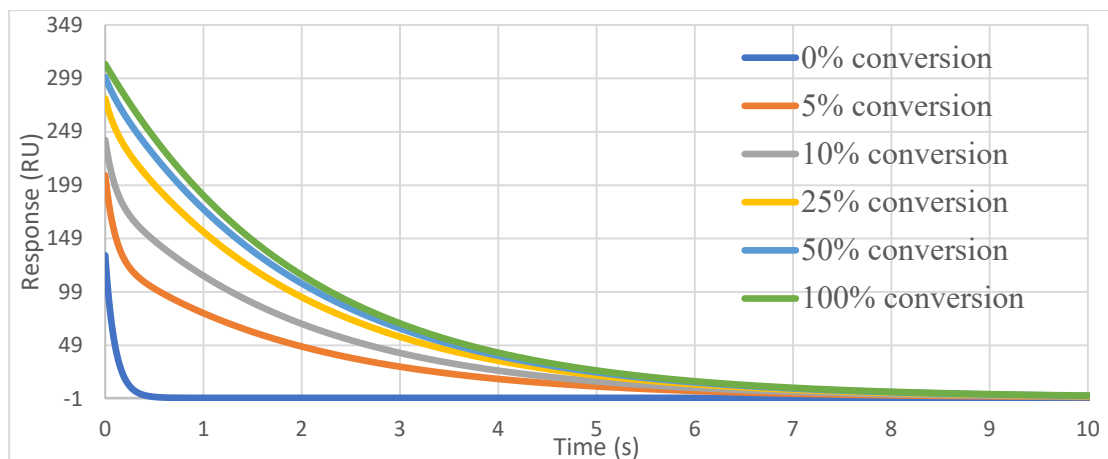


Figure 1.18: Theoretical dissociation phases for a series of sensorgrams for a 2-component mixture containing different amounts of each component to simulate reactions with varying reaction conversions. Sensorgrams were generated to mimic an unpurified reaction mixture containing a weakly binding component of $MW = 200$ Da, $K_D = 100 \mu M$, $k_{off} = 10 s^{-1}$ and a high affinity component with $MW = 320$ Da, affinity $K_D = 5 \mu M$ and $k_{off} = 0.5 s^{-1}$. The observable dissociation phase is dominated by the higher affinity component at concentrations as low as 5% conversion as shown in orange.

In the first reported example of ORS by SPR, a series of PIN1 inhibitors based on a benzimidazole core was synthesised via an amidation reaction using (1-cyano-2-ethoxy-2-oxoethylidenaminoxy)dimethylamino-morpholino-carbenium hexafluorophosphate (COMU) and analysed as crude reaction mixtures using SPR.¹²³ The calculated dissociation rate constants from the reaction mixtures varied by no more than $\pm 15 \%$ of those calculated on the pure samples indicating the accurate measurement of dissociation rate constants was possible in the mixtures, thereby validating the approach. In this way, a series of reactions can be undertaken in parallel, analysed by ORS without purification and only those compounds that show slower dissociation kinetics are resynthesised, purified and fully characterised before a full SPR analysis is performed to generate accurate affinity data. By avoiding the necessity to purify and characterise compounds that do not bind to a protein with higher affinity, ORS provides a cost effective and highly efficient approach to shorten the fragment to lead development process.

1.9 Fragment screening cascade

Our group has developed a screening cascade for identifying potential synthetic targets to take through to screening against *EcDsbA*. This process begins with the design of analogues based on a compound that has been shown to bind to *EcDsbA*. Analogues are ranked by docking and synthesis proceeds for desirable analogues. After synthesis the aqueous solubility of compounds are assessed via a series of quantitative 1D ^1H NMR experiments (as described above) to ensure that the compounds are soluble and free from aggregation under the conditions used for analysis. Binding to *EcDsbA* is then evaluated for soluble analogues via a single point ^1H - ^{15}N HSQC experiments with ^{15}N labelled protein. Analogues exhibiting CSP above a threshold (0.04 ppm) are carried onto HSQC titration experiments for affinity determination. Analogues that exhibit binding via HSQC are analysed via X-ray crystallography by soaking them into *EcDsbA* crystals (**Figure 1.19**).



Figure 1.19: Fragment screening cascade for synthesising and testing *EcDsbA* analogues.

1.10 Previous work

1.10.1 Fragment screen

Work previously published by our group involved the design and synthesis of a series of phenylthiazoles as non-covalent *EcDsbA* inhibitors using a fragment-based approach.⁵⁰ Primary screening was done via STD NMR screening of the MIPS fragment library of 1137 compounds. Screening was performed in mixtures of 3 – 5 fragments at a concentration of 300 μ M in the presence of 10 μ M oxidised *EcDsbA*.⁵⁰ Fragments with observable signals in STD NMR of the mixtures were retested as single compounds and further validated by ^1H - ^{15}N HSQC experiments. Hits from this process could be categorised into 7 distinct chemical classes (**Figure 1.20**) and the best of these fragment hits displayed ligand efficiencies of approximately 0.3. Structural information acquired from these validated fragments found many of them binding within the hydrophobic groove adjacent to the active site.¹²⁴

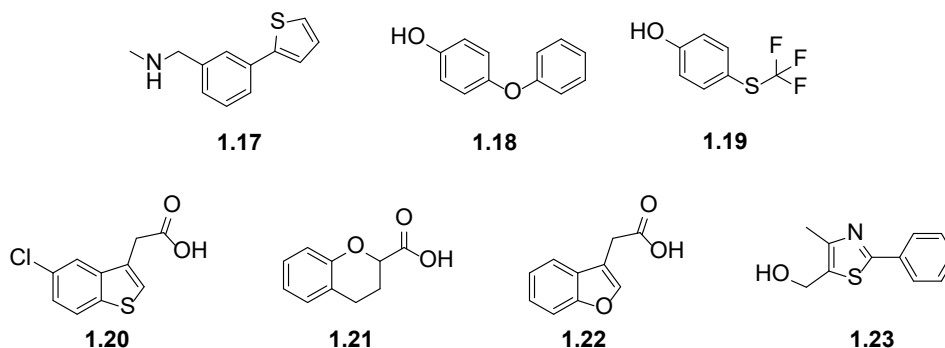


Figure 1.20: 7 Distinct chemical classes of fragment molecules identified from the MIPS fragment screen. Several analogues from these chemical classes displayed >0.04 ppm CSP in ^1H - ^{15}N HSQC experiments.

1.10.2 Development of phenylthiazoles as inhibitors of *EcDsbA*

The development of the phenylthiazole core within our group has been published.⁵⁰ From the initial hit a series of commercially available phenylthiazoles were purchased and tested via HSQC NMR and X-ray crystallography to generate an initial SAR series. The trifluoromethyl substituted

phenylthiazole **1.23** was chosen for further elaboration. Tyrosine and phenylalanine derived phenylthiazoles **1.25** and **1.26** were identified as the best analogues from the series showing significant HSQC perturbations and moderate affinity as determined by SPR (**Figure 1.21**).⁵⁰ Additionally, phenylalanine **1.25** demonstrated activity in a cell-based motility assay with $IC_{50} = 310 \pm 6 \mu M$ and a phenotype consistent with *EcDsbA* inhibition, resulting in a 59 % reduction in *E. coli* motility assay with no reduction in bacterial growth at 600 μM . These findings suggest that DsbA as a target is amenable to the development of small-molecule antivirulence agents. This work focuses on the development of the diaryl ethers **1.18** another chemical class which demonstrated comparable CSP in HSQC experiments to the phenylthiazoles, suggesting that they bind at a similar site on *EcDsbA*.

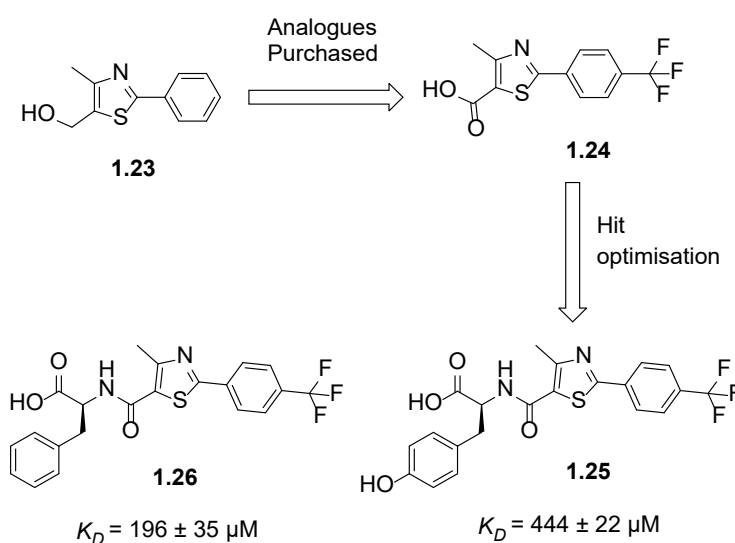


Figure 1.21: Key compounds from a previous FBDD campaign against *EcDsbA* from our group. Phenylthiazole **1.23** was identified from screening the in-house MIPS fragment library, related analogue **1.24** was purchased from a chemical vendor and developed into phenylthiazoles **1.25** and **1.26**. Affinities were determined by 1H - ^{15}N HSQC titrations of ligand with uniformly labelled ^{15}N *EcDsbA*.

1.11 Diaryl ethers a new class of *EcDsbA* inhibitors

Analogue purchasing provided some initial SAR for the diaryl ether core and most notably, the introduction of a *para*-cyano group in **1.27** showed promise. Structural data of **1.27** indicated the potential to expand the diaryl ethers by rational drug design (**Figure 1.22, a**). The structure revealed that the parent diaryl ether **1.27** formed π - π interactions with His32, Phe36 and Phe172 whilst the *para*-cyano functional group picked up a hydrogen bond to Gln35. Together these data suggested the diaryl ether scaffold was amenable to optimisation into more potent DsbA inhibitors.

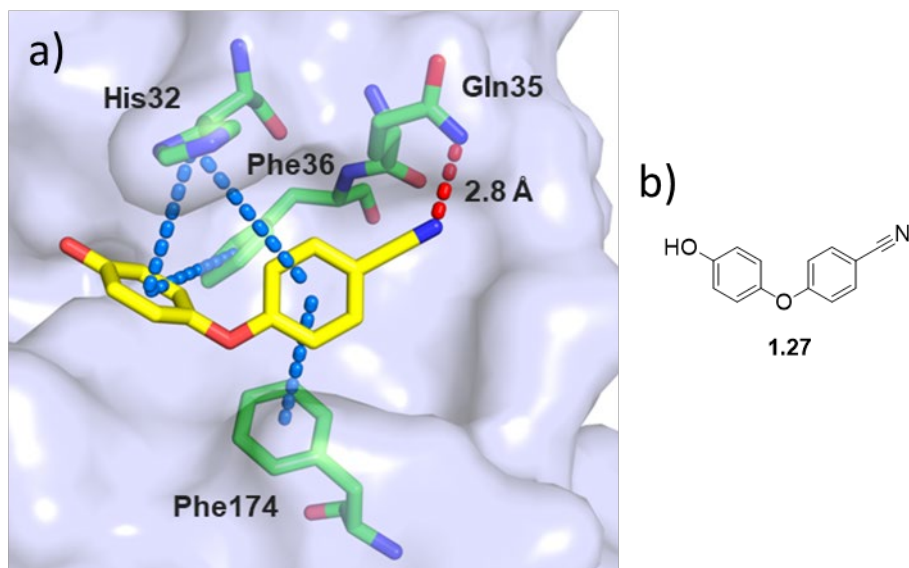


Figure 1.22: a) X-ray crystal structure of the parent diaryl ether **1.27** (yellow sticks) bound to oxidised *EcDsbA*. Residues His32, Gln35, Phe32 and Phe172 highlighted in green sticks and the proteins surface displayed in grey. Crystal structure was provided by Dr Olga Ilyichova. b) Structure of parent diaryl ether **1.27**.

1.12 Aims and hypothesis

The general hypothesis of this research program is that DsbA is a target for the development of compounds that will inhibit bacterial virulence in certain strains of pathogenic Gram-negative bacteria. The aim of my research project was to develop higher affinity compounds that bind to *E. coli* DsbA. Based on the success of our initial screen we hypothesised that the diaryl ether core can

be elaborated into more potent series of DsbA inhibitors. More specifically, the aims of my research were expansion of the diaryl ether core along vectors that were prioritised with the aid of X-ray crystal structures of co-complexes. Compounds were designed and evaluated using computational docking and upon synthesis screened in a number of biophysical binding experiments such as NMR, X-ray crystallography.

Based on this information a series of modifications to the parent diaryl ether (**Figure 1.23**) were designed to:

- Fill the empty pocket on the far RHS of the groove by expanding off the nitrile-substituted ring
- Strengthen interactions with Gln35 by replacing the *para*-cyano functional group with a series of hydrogen bond donors, acceptors and charged functional groups.
- Investigate the effects of replacing the ether linker with other related isosteres.
- Substitute the *para*-hydroxyl functional group with other synthetic handles to allow for more efficient synthetic elaboration across the groove.

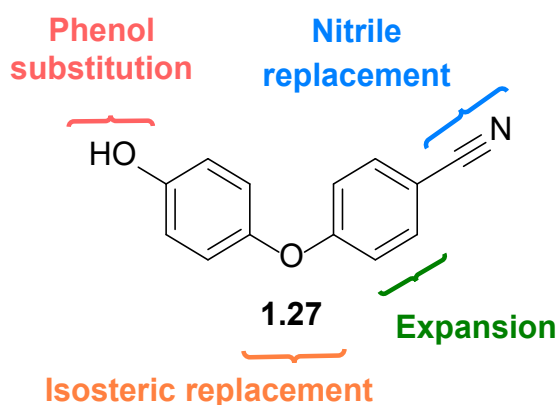


Figure 1.23: Initial design strategy for developing the diaryl ether core.

1.13 References

1. Blair, J. M. A.; Webber, M. A.; Baylay, A. J.; Ogbolu, D. O.; Piddock, L. J. V. Molecular mechanisms of antibiotic resistance. *Nat. Rev. Micro.* **2015**, 13, 42-51.
2. Awad, Y. M.; Kim, K. R.; Kim, S.; Kim, K.; Lee, S. R.; Lee, S. S.; Ok, Y. S. Monitoring antibiotic residues and corresponding antibiotic resistance genes in an agroecosystem. *J. Chem.* **2015**, 2015, 1-7.
3. Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* **2007**, 6, 29-40.
4. World health organisation. *Prioritization of pathogens to guide discovery, research and development of new antibiotics for drug-resistant bacterial infections, including tuberculosis.*; 2017.
5. Drusano, G. L.; Louie, A.; MacGowan, A.; Hope, W. Suppression of emergence of resistance in pathogenic bacteria: keeping our powder dry, part 1. *Antimicrob. Agents Chemother.* **2016**, 60, 1183-93.
6. Mantravadi, P. K.; Kalesh, K. A.; Dobson, R. C. J.; Hudson, A. O.; Parthasarathy, A. The Quest for Novel Antimicrobial Compounds: Emerging Trends in Research, Development, and Technologies. *J. Antibiot.* **2019**, 8, 8.
7. Van Aartsen, J. J.; Moore, C. E.; Parry, C. M.; Turner, P.; Phot, N.; Mao, S.; Suy, K.; Davies, T.; Giess, A.; Sheppard, A. E.; Peto, T. E. A.; Day, N. P. J.; Crook, D. W.; Walker, A. S.; Stoesser, N. Epidemiology of paediatric gastrointestinal colonisation by extended spectrum cephalosporin-resistant *Escherichia coli* and *Klebsiella pneumoniae* isolates in north-west Cambodia. *BMC Microbiol.* **2019**, 19, 59.
8. World Health Organisation, W. *WHO report on surveillance of antibiotic consumption: 2016-2018 early implementation.* ; 2018.
9. Lahra, M.; Enriquez, R.; George, R. *Australian gonococcal surveillance programme annual report*; 2017.
10. O'Neill, J. Review on antimicrobial resistance. <https://amr-review.org/>
11. O'Neill, J. *Antimicrobial resistance: tackling a crisis for the future health and wealth of nations*; 2016.
12. WHO. *Antimicrobial resistance: global report on surveillance* World Health Organisation: 2014; pp 1-5.
13. Tagliabue, A.; Rappuoli, R. Changing priorities in Vaccinology: antibiotic resistance moving to the top. *Front. Immunol.* **2018**, 9, 1068-1068.
14. Group, B. C. Breaking through the wall: a call for concerted action on antibiotics research and development. https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/5_Publikationen/Gesundheit/Berichte/GUARD_Follow_Up_Report_Full_Report_final.pdf (7th May 2019).
15. Cox, G.; Wright, G. D. Intrinsic antibiotic resistance: mechanisms, origins, challenges and solutions. *Int. J. Med. Microbiol.* **2013**, 303, 287-92.
16. Ali, J.; Rafiq, Q. A.; Ratcliffe, E. Antimicrobial resistance mechanisms and potential synthetic treatments. *Future science OA* **2018**, 4, FSO290-FSO290.
17. Russell, A. D. Biocide use and antibiotic resistance: the relevance of laboratory findings to clinical and environmental situations. *Lancet Infect. Dis.* **2003**, 3, 794-803.
18. Finley, R. L.; Collignon, P.; Larsson, D. G.; McEwen, S. A.; Li, X. Z.; Gaze, W. H.; Reid-Smith, R.; Timinouni, M.; Graham, D. W.; Topp, E. The scourge of antibiotic resistance: the important role of the environment. *Clin. Infect. Dis.* **2013**, 57, 704-10.
19. Van Hoek, A.; Mevius, D.; Guerra, B.; Mullany, P.; Roberts, A.; Aarts, H. Acquired antibiotic resistance genes: an overview. *Front. Microbiol.* **2011**, 2, 203-203.
20. Burmeister, A. R. Horizontal Gene Transfer. *Evol. Med. Public Health* **2015**, 2015, 193-194.
21. Poole, K. Resistance to β -lactam antibiotics. *Cell Mol. Life Sci.* **2004**, 61, 2200-2223.

22. Ramirez, M. S.; Tolmasky, M. E. Aminoglycoside modifying enzymes. *Drug Resist. Updat.* **2010**, 13, 151-171.
23. Hooper, D. C.; Jacoby, G. A. Mechanisms of drug resistance: quinolone resistance. *Ann. N. Y. Acad. Sci.* **2015**, 1354, 12-31.
24. Delcour, A. H. Outer membrane permeability and antibiotic resistance. *Biochim. Biophys. Acta* **2009**, 1794, 808-816.
25. Hancock, R. E.; Brinkman, F. S. Function of Pseudomonas porins in uptake and efflux. *Annu. Rev. Microbiol.* **2002**, 56, 17-38.
26. Chopra, I.; Roberts, M. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol. Mol. Biol. Rev.* **2001**, 65, 232-260.
27. Fishovitz, J.; Hermoso, J. A.; Chang, M.; Mobashery, S. Penicillin-binding protein 2a of methicillin-resistant Staphylococcus aureus. *IUBMB life* **2014**, 66, 572-577.
28. Beveridge, T. Structures of Gram-negative cell walls and their derived membrane vesicles. *J. Bacteriol.* **1999**, 181, 4725-4733.
29. Zgurskaya, H.; López, C.; Gnanakaran, S. Permeability barrier of Gram-negative cell envelopes and approaches to bypass it. *ACS Infect. Dis.* **2015**, 1, 512-522.
30. Agyepong, N.; Govinden, U.; Owusu-Ofori, A.; Essack, S. Multidrug-resistant Gram-negative bacterial infections in a teaching hospital in Ghana. *Antimicrob. Resist. Infect. Control* **2018**, 7, 37-37.
31. Hamprecht, A.; Gottig, S. Prevalence of multidrug-resistant Gram-negative bacteria. *Dtsch. Med. Wochenschr.* **2018**, 143, 625-633.
32. Heras, B.; Scanlon, M. J.; Martin, J. L. Targeting virulence not viability in the search for future antibacterials. *Br. J. Clin. Pharmacol.* **2015**, 79, 208-15.
33. Heras, B.; Shouldice, S. R.; Totsika, M.; Scanlon, M. J.; Schembri, M. A.; Martin, J. L. DSB proteins and bacterial pathogenicity. *Nat. Rev. Microbiol.* **2009**, 7, 215-25.
34. Rasko, D. A.; Sperandio, V. Anti-virulence strategies to combat bacteria-mediated disease. *Nat. Rev. Drug Discov.* **2010**, 9, 117-28.
35. Klemm, P.; Schembri, M. Bacterial adhesins: function and structure. *Int. J. Med. Microbiol.* **2000**, 290, 27-35.
36. Sauer, M.; Jakob, R.; Eras, J.; Baday, S.; Eriş, D.; Navarra, G.; Bernèche, S.; Ernst, B.; Maier, T.; Glockshuber, R. Catch-bond mechanism of the bacterial adhesin FimH. *Nat. Commun.* **2016**, 7, 10738.
37. Ronald, A. The etiology of urinary tract infection: traditional and emerging pathogens. *Mediators Inflamm.* **2003**, 49, 71-82.
38. Galanos, C.; Freudenberg, M. Bacterial endotoxins: biological properties and mechanisms of action. *Mediators Inflamm.* **1993**, 2, S11-S16.
39. Van Avondt, K.; Van Sorge, N. M.; Meyaard, L. Bacterial immune evasion through manipulation of host inhibitory immune signaling. *PLoS Pathog.* **2015**, 11, e1004644-e1004644.
40. Lindström, M.; Korkeala, H. Laboratory diagnostics of botulism. *Clin. Microbiol. Rev.* **2006**, 19, 298-314.
41. Collier, R. J. Diphtheria toxin: mode of action and structure. *Bacteriol. Rev.* **1975**, 39, 54-85.
42. Wagner, P. L.; Waldor, M. K. Bacteriophage control of bacterial virulence. *Infect. Immun.* **2002**, 70, 3985-3993.
43. Shannon, J. G.; Ross, C. L.; Koehler, T. M.; Rest, R. F. Characterization of anthrolysin O, the Bacillus anthracis cholesterol-dependent cytolysin. *Infect. Immun.* **2003**, 71, 3183-9.
44. Foster, T. J. Immune evasion by Staphylococci. *Nat. Rev. Microbiol.* **2005**, 3, 948-958.
45. Filloux, A.; Hachani, A.; Bleves, S. The bacterial type VI secretion machine: yet another player for protein transport across membranes. *Microbiology* **2008**, 154, 1570-83.
46. Cornelis, G. R. The type III secretion injectisome. *Nat. Rev. Microbiol.* **2006**, 4, 811-25.
47. Chatterjee, S.; Chaudhury, S.; McShan, A. C.; Kaur, K.; De Guzman, R. N. Structure and biophysics of type III secretion in bacteria. *Biochemistry* **2013**, 52, 2508-2517.

48. Sarkar-Tyson, M.; Atkins, H. S. Antimicrobials for bacterial bioterrorism agents. *Future Microbiol.* 2011/06//, 2011, p 667+.
49. Bitter, W.; Kuijl, C. Targeting bacterial virulence: the coming out of type VII secretion inhibitors. *Cell host & microbe* **2014**, 16, 430-2.
50. Adams, L. A.; Sharma, P.; Mohanty, B.; Ilyichova, O. V.; Mulcair, M. D.; Williams, M. L.; Gleeson, E. C.; Totsika, M.; Doak, B. C.; Caria, S.; Rimmer, K.; Horne, J.; Shouldice, S. R.; Vazirani, M.; Headey, S. J.; Plumb, B. R.; Martin, J. L.; Heras, B.; Simpson, J. S.; Scanlon, M. J. Application of fragment-based screening to the design of inhibitors of Escherichia coli DsbA. *Angew. Chem. Int. Ed. Engl.* **2015**, 54, 2179-84.
51. Aberg, V.; Almqvist, F. Pilicides-small molecules targeting bacterial virulence. *Org. Biomol. Chem.* **2007**, 5, 1827-34.
52. Mühlen, S.; Dersch, P. Anti-virulence strategies to target bacterial infections. In *How to Overcome the Antibiotic Crisis : Facts, Challenges, Technologies and Future Perspectives*, Stadler, M.; Dersch, P., Eds. Springer International Publishing: Cham, 2016; pp 147-183.
53. Zambelloni, R.; Marquez, R.; Roe, A. J. Development of antivirulence compounds: a biochemical review. *Chem. Biol. Drug Des.* **2015**, 85, 43-55.
54. Beckham, K. S.; Roe, A. J. From screen to target: insights and approaches for the development of anti-virulence compounds. *Front. Cell. Infect. Microbiol.* **2014**, 4, 139.
55. Rampioni, G.; Pillai, C. R.; Longo, F.; Bondi, R.; Baldelli, V.; Messina, M.; Imperi, F.; Visca, P.; Leoni, L. Effect of efflux pump inhibition on Pseudomonas aeruginosa transcriptome and virulence. *Sci. Rep.* **2017**, 7, 11392-11392.
56. Zhang, Y.; Faucher, F.; Zhang, W.; Wang, S.; Neville, N.; Poole, K.; Zheng, J.; Jia, Z. Structure-guided disruption of the pseudopilus tip complex inhibits the Type II secretion in Pseudomonas aeruginosa. *PLoS Pathog.* **2018**, 14, e1007343-e1007343.
57. Pinkner, J. S.; Remaut, H.; Buelens, F.; Miller, E.; Aberg, V.; Pemberton, N.; Hedenstrom, M.; Larsson, A.; Seed, P.; Waksman, G.; Hultgren, S. J.; Almqvist, F. Rationally designed small compounds inhibit pilus biogenesis in uropathogenic bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 103, 17897-902.
58. Ren, G.; Champion, M. M.; Huntley, J. F. Identification of disulfide bond isomerase substrates reveals bacterial virulence factors. *Mol. Microbiol.* **2014**, 94, 926-44.
59. Bardwell, J. C. Building bridges: disulphide bond formation in the cell. *Mol. Microbiol.* **1994**, 14, 199-205.
60. Kadokura, H.; Beckwith, J. Detecting folding intermediates of a protein as it passes through the bacterial translocation channel. *Cell* **2009**, 138, 1164-73.
61. Miki, T.; Okada, N.; Danbara, H. Two periplasmic disulfide oxidoreductases, DsbA and SrgA, target outer membrane protein SpiA, a component of the Salmonella pathogenicity island 2 type III secretion system. *J. Biol. Chem.* **2004**, 279, 34631-42.
62. Ireland, P. M.; McMahon, R. M.; Marshall, L. E.; Halili, M.; Furlong, E.; Tay, S.; Martin, J. L.; Sarkar-Tyson, M. Disarming Burkholderia pseudomallei: structural and functional characterization of a disulfide oxidoreductase (DsbA) required for virulence in vivo. *Antioxid. Redox Signal.* **2014**, 20, 606-617.
63. Halili, M. A.; Bachu, P.; Lindahl, F.; Bechara, C.; Mohanty, B.; Reid, R. C.; Scanlon, M. J.; Robinson, C. V.; Fairlie, D. P.; Martin, J. L. Small molecule inhibitors of disulfide bond formation by the bacterial DsbA-DsbB dual enzyme system. *ACS Chem. Biol.* **2015**, 10, 957-64.
64. Martin, J. L. Thioredoxin —a fold for all reasons. *Structure* **1995**, 3, 245-250.
65. Martin, J. L.; Bardwell, J. C.; Kuriyan, J. Crystal structure of the DsbA protein required for disulphide bond formation in vivo. *Nature* **1993**, 365, 464-8.
66. Kadokura, H.; Tian, H.; Zander, T.; Bardwell, J. C.; Beckwith, J. Snapshots of DsbA in action: detection of proteins in the process of oxidative folding. *Science* **2004**, 303, 534-7.
67. Zhang, H. Z.; Donnenberg, M. S. DsbA is required for stability of the type IV pilin of enteropathogenic Escherichia coli. *Mol. Microbiol.* **1996**, 21, 787-97.

68. Landeta, C.; Blazyk, J. L.; Hatahet, F.; Meehan, B. M.; Eser, M.; Myrick, A.; Bronstain, L.; Minami, S.; Arnold, H.; Ke, N.; Rubin, E. J.; Furie, B. C.; Furie, B.; Beckwith, J.; Dutton, R.; Boyd, D. Compounds targeting disulfide bond forming enzyme DsbB of Gram-negative bacteria. *Nat. Chem. Biol.* **2015**, 11, 292-8.
69. González-Bello, C. Designing irreversible inhibitors—worth the effort? *ChemMedChem* **2016**, 11, 22-30.
70. Kumar, A.; Voet, A.; Zhang, K. Y. Fragment based drug design: from experimental to computational approaches. *Curr. Med. Chem.* **2012**, 19, 5128-47.
71. Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discov.* **2007**, 6, 211-9.
72. Mashalidis, E. H.; Śledź, P.; Lang, S.; Abell, C. A three-stage biophysical screening cascade for fragment-based drug discovery. *Nat. Protocols* **2013**, 8, 2309-2324.
73. Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* **2016**, advance online publication.
74. Gozalbes, R.; Carbajo, R. J.; Pineda-Lucena, A. Contributions of computational chemistry and biophysical techniques to fragment-based drug discovery. *Curr. Med. Chem.* **2010**, 17, 1769-94.
75. Hall, R. J.; Mortenson, P. N.; Murray, C. W. Efficient exploration of chemical space by fragment-based screening. *Prog. Biophys. Mol. Biol.* **2014**, 116, 82-91.
76. Erlanson, D.; Zartler, T. How effectively can fragments sample chemical space? . In Erlanson, D., Ed. Online, 2011.
77. Bentley, M.; Doak, B. C.; Mohanty, B.; Scanlon, M. J. Applications of NMR Spectroscopy in FBDD. In *Modern Magnetic Resonance*, Webb, G. A., Ed. Springer International Publishing: Cham, 2018; pp 2211-2231.
78. Hubbard, R. E.; Davis, B.; Chen, I.; Drysdale, M. J. The SeeDs approach: integrating fragments into drug discovery. *Curr. Top. Med. Chem.* **2007**, 7, 1568-81.
79. Dalvit, C.; Pevarello, P.; Tatò, M.; Veronesi, M.; Vulpetti, A.; Sundström, M. Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water*. *J. Biomol. NMR* **2000**, 18, 65-68.
80. Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B. WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability. *J. Biomol. NMR* **2001**, 21, 349-59.
81. Mayer, M.; James, T. L. Detecting ligand binding to a small RNA target via saturation transfer difference NMR experiments in D(2)O and H(2)O. *J. Am. Chem. Soc.* **2002**, 124, 13376-7.
82. Doak, B. C.; Morton, C. J.; Simpson, J. S.; Scanlon, M. J. Design and Evaluation of the Performance of an NMR Screening Fragment Library. *Aust. J. Chem.* **2013**, 66, 1465-1472.
83. Doak, B. C.; Morton, C. J.; Simpson, J. S.; Scanlon, M. J. Assembly of fragment screening libraries. In *Applied Biophysics for Drug Discovery*, 2017; pp 263-283.
84. Taylor, A.; Doak, B. C.; Scanlon, M. J. Design of a Fragment-Screening Library. *Methods Enzymol.* **2018**, 610, 97-115.
85. Viegas, A.; Manso, J.; Nobrega, F. L.; Cabrita, E. J. Saturation-transfer difference (STD) NMR: a simple and fast method for ligand screening and characterization of protein binding. *J. Chem. Educ.* **2011**, 88, 990-994.
86. Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B. WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability. *J. Biomol. NMR* **2001**, 21, 349-359.
87. Ziarek, J. J.; Peterson, F. C.; Lytle, B. L.; Volkman, B. F. Binding site identification and structure determination of protein-ligand complexes by NMR a semiautomated approach. *Methods Enzymol.* **2011**, 493, 241-75.
88. Fielding, L. NMR methods for the determination of protein-ligand dissociation constants. *Curr. Top. Med. Chem.* **2003**, 3, 39-53.

89. Davis, B. J.; Erlanson, D. A. Learning from our mistakes: the 'unknown knowns' in fragment screening. *Bioorg. Med. Chem. Lett.* **2013**, 23, 2844-52.
90. Vom, A.; Headey, S.; Wang, G.; Capuano, B.; Yuriev, E.; Scanlon, M. J.; Simpson, J. S. Detection and prevention of aggregation-based false positives in STD-NMR-based fragment screening. *Aust. J. Chem.* **2013**, 66, 1518-1524.
91. Harner, M. J.; Frank, A. O.; Fesik, S. W. Fragment-based drug discovery using NMR spectroscopy. *J. Biomol. NMR* **2013**, 56, 65-75.
92. Bharti, S. K.; Roy, R. Quantitative ¹H NMR spectroscopy. *Trends Anal. Chem.* **2012**, 35, 5-26.
93. Khan, F.; He, M.; Taussig, M. J. Double-hexahistidine tag with high-affinity binding for protein immobilization, purification, and detection on ni-nitrilotriacetic acid surfaces. *Anal. Chem.* **2006**, 78, 3072-9.
94. Hutsell, S. Q.; Kimple, R. J.; Siderovski, D. P.; Willard, F. S.; Kimple, A. J. High-affinity immobilization of proteins using biotin- and GST-based coupling strategies. *Methods Mol. Biol.* **2010**, 627, 75-90.
95. Hahnefeld, C.; Drewianka, S.; Herberg, F. W. Determination of kinetic data using surface plasmon resonance biosensors. *Methods Mol. Med.* **2004**, 94, 299-320.
96. Cooper, M. A. Optical biosensors in drug discovery. *Nat. Rev. Drug Discov.* **2002**, 1, 515-28.
97. Ferenczy, G. G.; Keserü, G. M. How Are Fragments Optimized? A Retrospective Analysis of 145 Fragment Optimizations. *J. Med. Chem.* **2013**, 56, 2478-2486.
98. Mortenson, P. N.; Erlanson, D. A.; de Esch, I. J. P.; Jahnke, W.; Johnson, C. N. Fragment-to-Lead Medicinal Chemistry Publications in 2017. *J. Med. Chem.* **2019**, 62, 3857-3872.
99. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, 28, 235-242.
100. Roberts, N. A.; Martin, J. A.; Kinchington, D.; Broadhurst, A. V.; Craig, J. C.; Duncan, I. B.; Galpin, S. A.; Handa, B. K.; Kay, J.; Krohn, A.; et al. Rational design of peptide-based HIV proteinase inhibitors. *Science* **1990**, 248, 358-61.
101. Anderson, A. C. The process of structure-based drug design. *Chem. Biol.* **2003**, 10, 787-797.
102. Zheng, H.; Hou, J.; Zimmerman, M. D.; Wlodawer, A.; Minor, W. The future of crystallography in drug discovery. *Expert Opin Drug Discov* **2014**, 9, 125-137.
103. Jhoti, H.; Cleasby, A.; Verdonk, M.; Williams, G. Fragment-based screening using X-ray crystallography and NMR spectroscopy. *Curr. Opin. Chem. Biol.* **2007**, 11, 485-93.
104. Davies, T. G.; Tickle, I. J. Fragment screening using X-ray crystallography. *Top. Curr. Chem.* **2012**, 317, 33-59.
105. Kraft, P.; Bergamaschi, A.; Broennimann, C.; Dinapoli, R.; Eikenberry, E. F.; Henrich, B.; Johnson, I.; Mozzanica, A.; Schlepütz, C. M.; Willmott, P. R.; Schmitt, B. Performance of single-photon-counting PILATUS detector modules. *J. Synchrotron Radiat.* **2009**, 16, 368-375.
106. Collins, P. M.; Douangamath, A.; Talon, R.; Dias, A.; Brandao-Neto, J.; Krojer, T.; von Delft, F. Achieving a Good Crystal System for Crystallographic X-Ray Fragment Screening. *Methods Enzymol.* **2018**, 610, 251-264.
107. Hopkins, A. L.; Keseru, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug Discov.* **2014**, 13, 105-21.
108. Hajduk, P. J. Fragment-based drug design: how big is too big? *J. Med. Chem.* **2006**, 49, 6972-6.
109. Comer, J.; Tam, K. Lipophilicity Profiles: Theory and Measurement. In *Pharmacokinetic Optimization in Drug Research*, pp 275-304.
110. Chen, M.; Borlak, J.; Tong, W. High lipophilicity and high daily dose of oral medications are associated with significant risk for drug-induced liver injury. *Hepatology (Baltimore, Md.)* **2013**, 58, 388-96.
111. Waring, M. J.; Arrowsmith, J.; Leach, A. R.; Leeson, P. D.; Mandrell, S.; Owen, R. M.; Pairaudeau, G.; Pennie, W. D.; Pickett, S. D.; Wang, J.; Wallace, O.; Weir, A. An analysis of the

- attrition of drug candidates from four major pharmaceutical companies. *Nat. Rev. Drug Discov.* **2015**, 14, 475.
112. Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **2000**, 44, 235-49.
113. Hughes, J. D.; Blagg, J.; Price, D. A.; Bailey, S.; Decrescenzo, G. A.; Devraj, R. V.; Ellsworth, E.; Fobian, Y. M.; Gibbs, M. E.; Gilles, R. W.; Greene, N.; Huang, E.; Krieger-Burke, T.; Loesel, J.; Wager, T.; Whiteley, L.; Zhang, Y. Physiochemical drug properties associated with in vivo toxicological outcomes. *Bioorg. Med. Chem. Lett.* **2008**, 18, 4872-5.
114. Kenny, P. W. The nature of ligand efficiency. *J. Cheminformatics* **2019**, 11, 8.
115. Lamoree, B.; Hubbard, R. E. Current perspectives in fragment-based lead discovery (FBLD). *Essays Biochem.* **2017**, 61, 453-464.
116. Oneal, P. A.; Kwitkowski, V.; Luo, L.; Shen, Y. L.; Subramaniam, S.; Shord, S.; Goldberg, K. B.; McKee, A. E.; Kaminskis, E.; Farrell, A.; Pazdur, R. FDA approval summary: Vemurafenib for the treatment of patients with Erdheim-Chester disease with the BRAFV600 mutation. *The oncologist* **2018**, 23, 1520-1524.
117. Bollag, G.; Tsai, J.; Zhang, J.; Zhang, C.; Ibrahim, P.; Nolop, K.; Hirth, P. Vemurafenib: the first drug approved for BRAF-mutant cancer. *Nat. Rev. Drug Discov.* **2012**, 11, 873.
118. Nikiforov, P. O.; Surade, S.; Blaszczyk, M.; Delorme, V.; Brodin, P.; Baulard, A. R.; Blundell, T. L.; Abell, C. A fragment merging approach towards the development of small molecule inhibitors of Mycobacterium tuberculosis EthR for use as ethionamide boosters. *Org. Biomol. Chem.* **2016**, 14, 2318-2326.
119. Ichihara, O.; Barker, J.; Law, R. J.; Whittaker, M. Compound design by fragment-linking. *Mol. Inform.* **2011**, 30, 298-306.
120. Nazaré, M.; Matter, H.; Will, D. W.; Wagner, M.; Urmann, M.; Czech, J.; Schreuder, H.; Bauer, A.; Ritter, K.; Wehner, V. Fragment deconstruction of small, potent factor Xa inhibitors: exploring the superadditivity energetics of fragment linking in protein–ligand complexes. *Angew. Chem. Int. Ed. Engl.* **2012**, 51, 905-911.
121. De Fusco, C.; Brear, P.; Iegre, J.; Georgiou, K. H.; Sore, H. F.; Hyvönen, M.; Spring, D. R. A fragment-based approach leading to the discovery of a novel binding site and the selective CK2 inhibitor CAM4066. *Bioorg. Med. Chem.* **2017**, 25, 3471-3482.
122. Kappe, C. O.; Damm, M. Parallel microwave chemistry in silicon carbide microtiter platforms: a review. *Molec. Divers.* **2012**, 16, 5-25.
123. Murray, J. B.; Roughley, S. D.; Matassova, N.; Brough, P. A. Off-rate screening (ORS) by surface plasmon resonance. An efficient method to kinetically sample hit to lead chemical space from unpurified reaction products. *J. Med. Chem.* **2014**, 57, 2845-2850.
124. Inaba, K.; Murakami, S.; Suzuki, M.; Nakagawa, A.; Yamashita, E.; Okada, K.; Ito, K. Crystal structure of the DsbB-DsbA complex reveals a mechanism of disulfide bond generation. *Cell* **2006**, 127, 789-801.

Chapter 2 – Substitution of the benzonitrile ring.

2.1 Previous work

A screen of the MIPS fragment library identified several distinct chemical classes including phenylthiazoles, benzofurans, diaryl amides and diaryl ethers as fragments that bound to *EcDsbA*. Commercially available analogues of these classes were purchased and tested against *EcDsbA*. This identified parent diaryl ether **2.1** which bound with an affinity of $K_D = 930 \pm 180 \mu\text{M}$ as measured by ^1H - ^{15}N HSQC titrations. The parent **2.1** bound with a moderate ligand efficiency of $0.27 \text{ kcal mol}^{-1} \text{ HAC}^{-1}$ which is consistent with binding within a surface exposed shallow hydrophobic groove and this further emphasised *EcDsbA* as a difficult target. Consequently, ligand efficiency was not used as a criteria to select fragments for progression in this project.

Despite the weak affinity, fragment **2.1** made several key interactions with *EcDsbA* including π - π stacking interactions with sidechains of His32, Phe36, Phe174 and a H-bond from the *para*-cyano group of **2.1** to the sidechain amide of Gln35 (2.8 Å away, **Figure 2.1**).

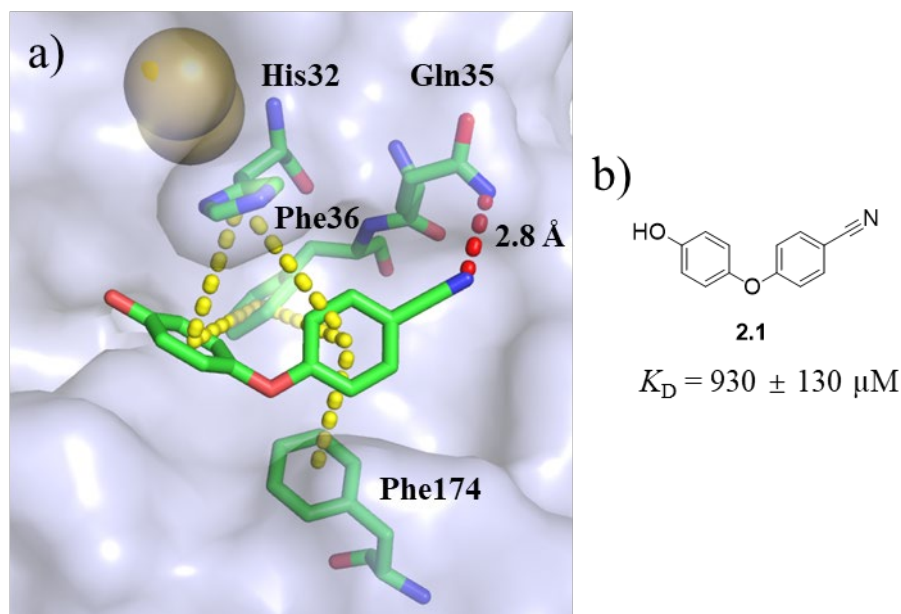


Figure 2.1: a) X-ray crystal structure of parent diaryl ether **2.1** (green sticks) bound to oxidised *EcDsbA* (grey surface) showing key amino acids His32, Gln35, Phe36, Phe174 (green sticks) and active site residues C30 and C33 (yellow sphere). π - π stacking interactions are shown by yellow dashes and H-bonds as red dashes. b) Structure of parent diaryl ether **2.1**.

2.1.1 Project plan

Based on this crystal structure a series modifications to **2.1** was rationally designed (**Figure 2.2, a**) to optimise and elaborate the diaryl ethers. Four series of analogues were designed, these included:

(a) Expanding off the benzonitrile ring along vectors R^1 and R^2 (**Figure 2.2, a**) with small (1 – 2 heavy atoms) functional groups to further fill the groove and improve affinity. These modifications would also reduce the symmetry of the diaryl ethers to decrease the chances of fragments changing binding orientation.

(b) Replacement of the *para*-nitrile substituent was investigated to strengthen the H-bond to Gln35 and potentially improve aqueous solubility.

(c) The *para*-hydroxyl group of the diaryl ether core was replaced with a series of carboxylic acids, amides and alkyl alcohols which had the potential to improve solubility and afford different synthetic handles for expansion into the left-hand side of the binding groove.

(d) Optimisation of the ether linker was investigated to alter the ring electronics and potentially improve the π - π stacking interactions as well as allow for further expansion to the bottom of the protein as depicted in **Figure 2.2b**.

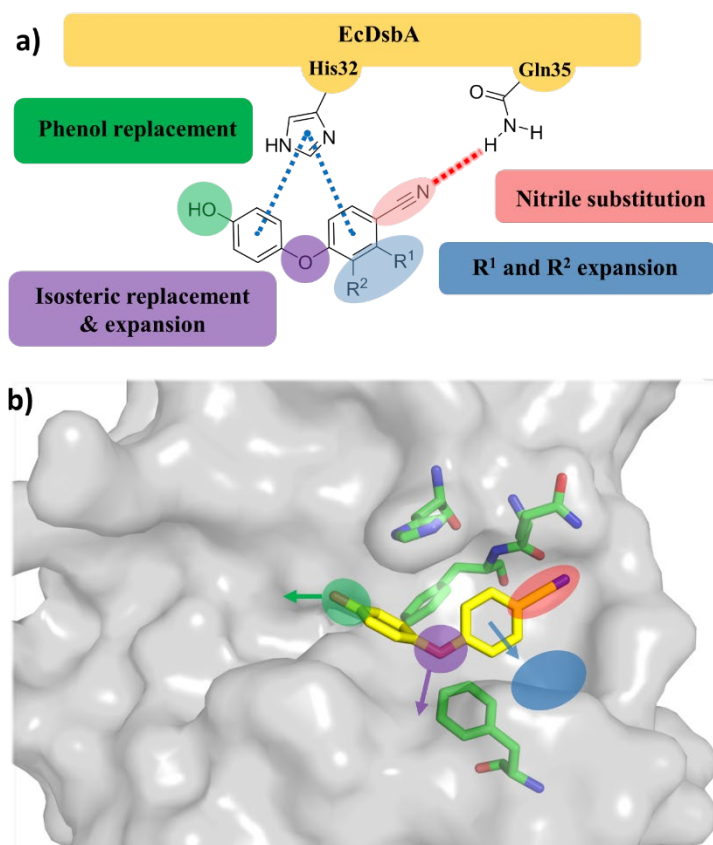


Figure 2.2. a) Schematic depicting the planned analogue series for optimisation of the diaryl ethers. The key π - π stacking and H-bond interactions formed between diaryl ether **2.1** and *EcDsbA* are displayed as dotted lines. b) X-ray crystal structure of diaryl ether **2.1** (yellow sticks) bound to oxidised *EcDsbA*. The regions for replacement and expansion are highlighted and indicted with arrows, respectively.

2.2 Screening cascade

2.2.1 Solubility and aggregation assessment

Since fragments typically bind with low affinity high compound solubility is required to generate reliable affinity and structure activity relationship (SAR) data. Hence evaluation of compounds solubility and aggregation in solution via a semi-quantitative 1D ¹H NMR titration was undertaken. Spectra were recorded with long relaxation delays at 125, 250, 500 and 1000 μ M of ligand in screening buffer with 2% dimethyl sulfoxide (DMSO).¹ DSS was used as an internal standard in the

NMR samples to semi-quantitatively measure the absolute sample concentration. Compounds failed solubility at a given concentration if they:

- (a) failed to show absolute concentration within 3-fold of the desired
- (b) failed to show a doubling in signal integral concordant with doubling sample concentration
- (c) chemical shift changes ≥ 0.004 ppm occurred with increasing sample concentration.

The above results are indicators of insolubility or aggregation.¹ The highest concentration at which the compound passes these solubility criteria was considered the compounds maximum assayable concentration, hereafter referred to as solubility, and above this concentration the reliability of assay data was considered low. Aggregates are complex and come in various shapes and sizes,¹⁻³ as such these solubility estimates do not necessarily indicate a lack of aggregation. Compounds can form aggregates in solution which do not precipitate from solution or show visible chemical shift changes, but are known to interfere with biological assays.¹ Dynamic light scattering is commonly used to detect aggregation in solution however, this method is time consuming and costly. The semi-quantitative 1D ¹H NMR titration assay provides a relatively quick and easily implemented assessment to estimate solubility that can be implemented for the large number of compounds synthesised throughout this project. As an example, **Figure 2.3** shows three solubility overlays: (a) passed at all concentrations up to 1000 μ M, (b) showed doubling in signal intensity up to 500 μ M and precipitated at 1000 μ M. (c) showed significant chemical shift changes suggesting aggregation in solution.

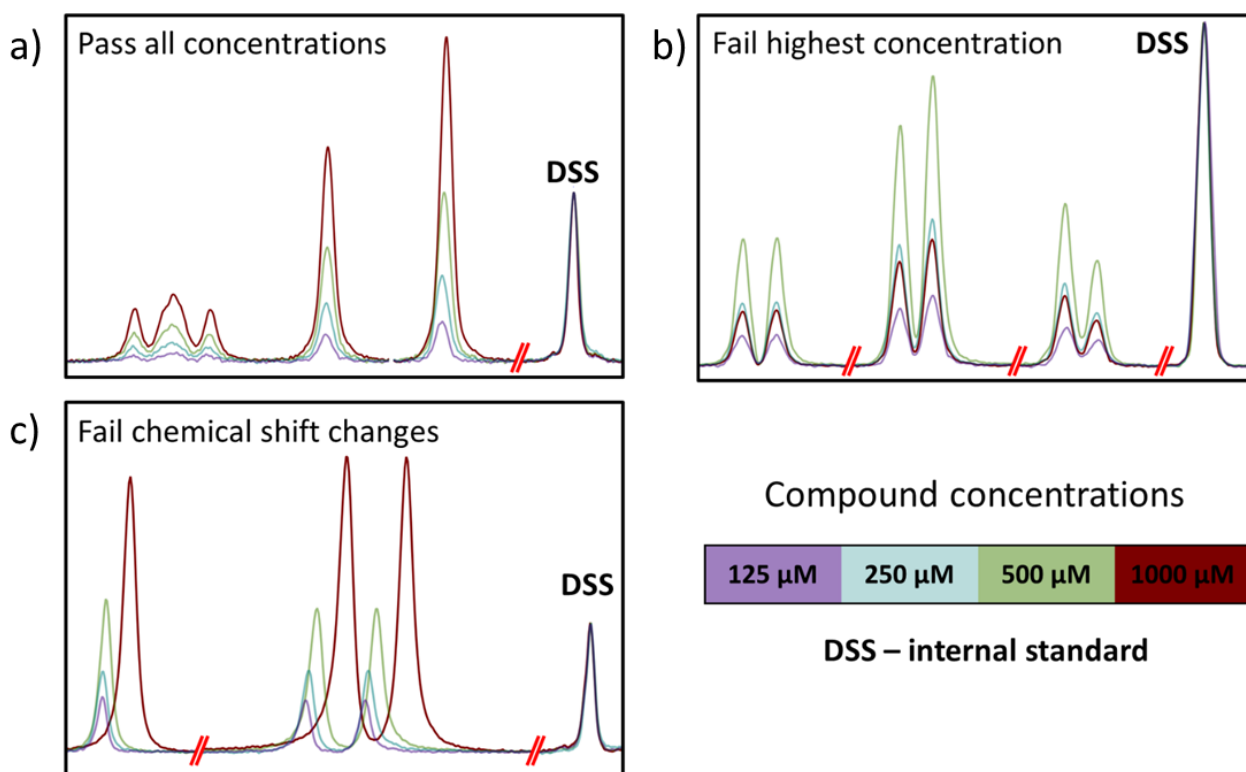


Figure 2.3: Three examples of solubility testing using semi-quantitative 1D ^1H NMR titration spectra. Spectra of a compound at relative concentrations of 125 μM (purple), 250 μM (light blue), 500 μM (green) and 1000 μM (maroon) in 2 % DMSO, 1 M phosphate buffer at pH 7. a) Spectra shows a compound that passed solubility at 1000 μM and below. b) compound passed at 500 μM and below but failed at 1000 μM due to not doubling in peak intensity. c) compound passed at 250 μM and below, but failed at 500 μM and 1000 μM due to chemical shift changes indicating aggregation. NMR spectra have been cropped as indicated by double red lines.

2.2.2 Binding by ^1H - ^{15}N HSQC NMR

Compounds which passed solubility at 250 μM or higher were then tested via single point ^1H - ^{15}N HSQC against ^{15}N -labelled *EcDsbA* at their maximum solubility. ^1H - ^{15}N HSQC were recorded in a HEPES buffer (100 μM ^{15}N *EcDsbA*, 50 mM HEPES, 25 mM NaCl in 2 % DMSO, 10 % D_2O and 88 % H_2O at pH 6.8) as standard protocol within our group. ^1H - ^{15}N HSQC provides information on N-H resonances within a protein, including backbone and side chain N-Hs (**Figure 2.4, a**). Addition of ligand to the protein can induce CSP of the peaks in the HSQC spectrum which is indicative of

ligand binding.⁴ The CSP of residues within the hydrophobic groove of *EcDsbA* were measured and compounds showing weighted CSP >0.04 ppm of multiple residues within the hydrophobic groove were further characterised by recording ^1H - ^{15}N HSQC ligand titrations to determine their affinity (K_D). In the titration experiments, the CSP of residues within the binding site were measured at varying ligand concentrations and the data was fitted to a single site binding model with ligand depletion to calculate affinity (K_D) (Figure 2.4, b and c).

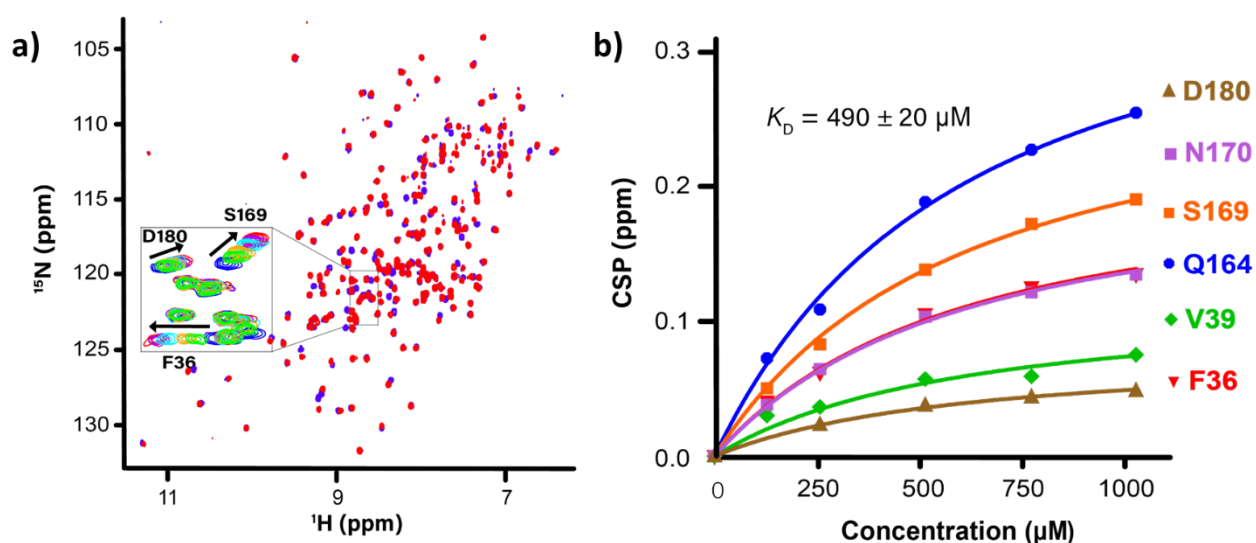


Figure 2.4: a) ^1H - ^{15}N HSQC NMR spectra of oxidized ^{15}N -*EcDsbA* with a ligand at concentrations of 0 μM (blue), 125 μM (green), 250 μM (orange), 500 μM (cyan), 750 μM (purple) and 1000 μM (red). Black arrows indicate the direction of the CSP. b) Analysis of ^1H - ^{15}N HSQC titrations showing the experimental data for several residues. The data was fitted to a single site binding model using the equation 1.2.

2.2.3 Structural characterisation

Analogues which showed binding by HSQC were carried onto X-ray crystallography and soaked into *EcDsbA* crystals in an effort to obtain crystal structures of the complex and visualise their binding mode. Analogues which were judged to be of greatest interest and did not show ligand density by the soaking method were co-crystallised with *EcDsbA* in a further attempt to acquire structural information of the ligand bound. All structures reported had the electron density of their final refined

structure (mtz) and ligand omit maps manually inspected in COOT⁵ to assess the validity of the structural model. Strong electron density for the ligand in the omit electron density maps can verify the presence of ligand in the solved model with greater reliability than inspection of the final electron density maps.⁶ Structures were then categorised as either having full, partial or poor electron density for the ligand. From compounds characterised as having full ligand the electron density and the orientation of the ligand could be unambiguously determined from the omit map and, the final refined structure had electron density for the entire ligand. Partial electron density structures were missing some electron density for the ligand in the final refined structure however the orientation of the core could be determined unambiguously from the omit map. Poor electron density compounds had substantial density missing in the final refined structure and/or the core orientation could not be determined from the omit map. Interpretation of electron densities for compounds was further aided by the use of the Electron density score for individual atoms (EDIA) scoring function.⁷ EDIA scores provide an electron density score for each individual atom. Scores are averaged to give an overall EDIA score for the entire ligand, scores >0.8 were considered well supported by experimental data, decreasing EDIA scores correlate with a decrease in consistency of the model and experimental data.

2.3 Crystal structure analysis

As a number of crystal structures were available for *EcDsbA* as well as the co-structure of diaryl ether **2.1**, computational docking of potential synthetic analogues to the structure was used to aid in design of higher affinity analogues. Initially we set out to define the best protein model for docking calculations of the diaryl ether series. As such, we sought to identify any conserved and structural waters within the binding groove of *EcDsbA* by analysing 8 *EcDsbA* crystal structures from the PDB (1A2J, 1A2M, 1ACV, 1DSB, 1FVK, 1U3A, 3DKS, 4TKY) and the structure of the parent diaryl ether **2.1**. The top 20 conserved waters were visually inspected and 3 highly conserved waters were identified within the hydrophobic groove. These waters were found in 8 of the 9 analysed crystal structures and formed a conserved H-bond network, likely making them integral to the protein

structure and difficult or unfavourable to displace (**Figure 2.5, a**). Literature suggests that displacement of structural waters generally results in a reduction in affinity,⁸ thus expansion into the portion of the hydrophobic groove containing the water filled pocket was not considered as a suitable design strategy. This study also highlighted the flexibility of the sidechains of His32, Gln35, Gln164 and flexibility in the loop forming the bottom of the hydrophobic groove in *EcDsbA* (residues Q164 - M171) possibly making it difficult to rationally design substituents targeting this area (**Figure 2.5, b**). Hence, the PDB model of diaryl ether **2.1** bound to *EcDsbA* with the three structural waters was used as the protein model for all subsequent docking calculations to assess potential analogues.

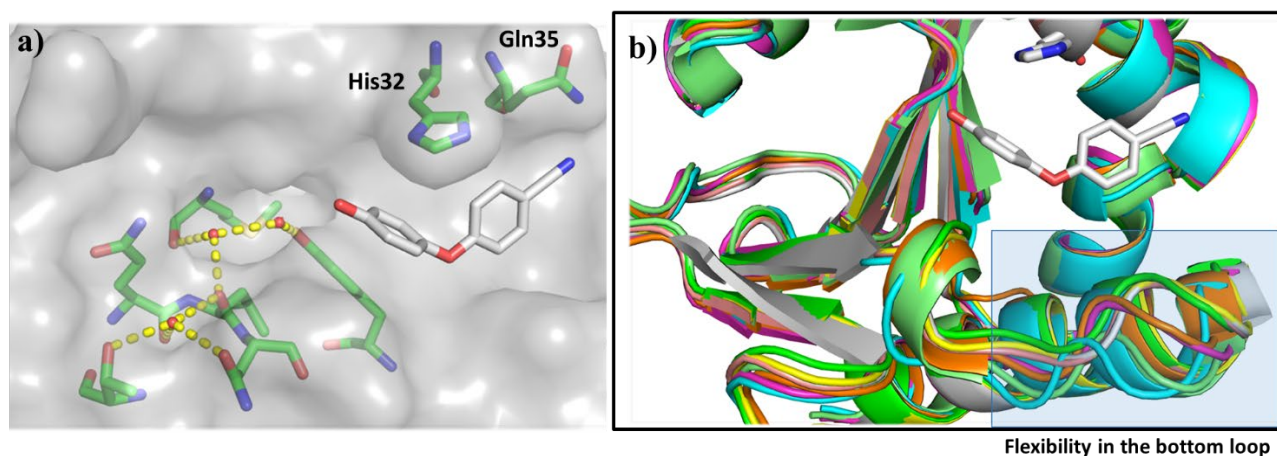


Figure 2.5: a) Crystal structure of the parent diaryl ether **2.1** (grey sticks) bound to oxidised *EcDsbA* with the 3 structural waters in the portion of the hydrophobic groove shown (red spheres) as well as their hydrogen bonding network (yellow dashes). b) Overlay cartoon representations of the 9 crystal structures of *EcDsbA* (PDB ID: 1A2J, 1A2M, 1ACV, 1DSB, 1FVK, 1U3A, 3DKS, 4TKY and parent diaryl ether **2.1** protein structure) showing the flexibility of the bottom loop (blue box) of *EcDsbA*. Ligand and active site His32 are shown as grey sticks for reference.

2.4 Expansion off the benzonitrile ring

2.4.1 Design and docking

A series of diaryl ether analogues exploring substitution of the benzonitrile and phenol rings were docked using Glide⁹ extra precision (XP) mode with default settings against the protein structure derived from parent fragment **2.1** bound to *EcDsbA* (**Figure 2.6, a**). A number of analogues were docked to probe potential interactions with the right-hand side (RHS) of the groove including a series of hydrophobic groups (Me, CF₃, *i*-Pr, Et, cyclopropyl), H-bond donors (OH, NH) and acceptors (OCH₃, CN, CH₂CN). The docked poses were manually inspected for compounds which maintained the binding mode of the parent hit **2.1**, had acceptable geometry (no strained rings or bonds or clashes with the protein) and formed new interactions. The docking displayed few additional polar contacts with a preference for small substitutions off the benzonitrile ring at R¹ or R² positions (**Figure 2.6, b**) while substitution at R³ and R⁴ positions of the phenol ring frequently caused the core to change binding mode in docking calculations.

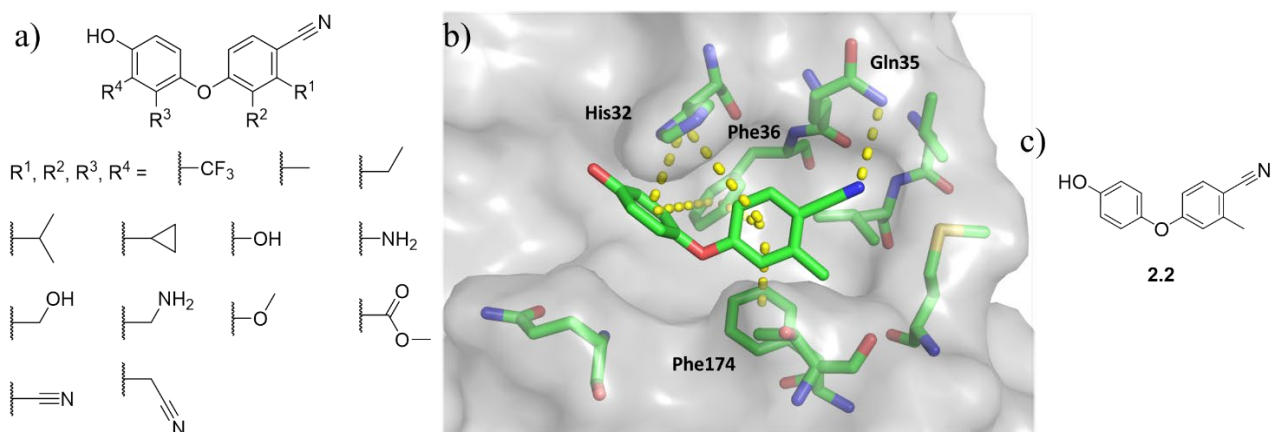


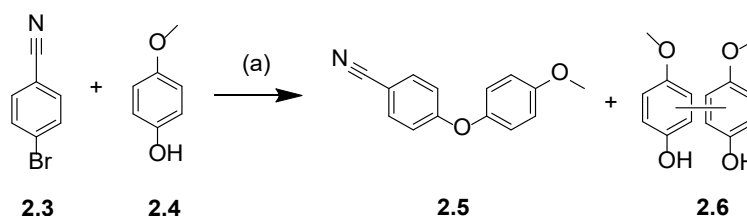
Figure 2.6: a) Diaryl ether substituents explored in computational docking against *EcDsbA*. Compounds were docked with mono-substitution of various functions groups at the R¹, R², R³ or R⁴ positions. b). Docked structure of the 2-methylbenzonitrile derivative **2.2** (green sticks) against to *EcDsbA*. The 2-methyl substituent occupies more of the hydrophobic groove. c) Structure of the 2-methyl substituted diaryl ether **2.2**.

A series of commercially available reagents were then purchased and prioritised for synthesis to probe this region of the groove with small (1 - 2 heavy atoms) functional groups expanding at vectors R¹ and R². These reagents contain a broad range of substituents (H-bond donors, acceptors and hydrophobic functional groups) and were selected to better understand SAR around these positions.

2.4.2 Synthesis

Synthesis of the diaryl ether analogue set was envisaged via an Ullmann coupling reaction allowing for efficient synthesis from pairs of phenols and aryl halides. Initially, optimisation of the Ullmann coupling conditions was conducted between 4-bromobenzonitrile (**2.3**) and 4-methoxyphenol (**2.4**), varying the reaction temperature using *N,N*-dimethylglycine hydrochloride ligand and copper(I) iodide, based on literature.^{10, 11} Preliminary experiments suggested a series of by-product polyphenols shown by the general structure **2.6** were forming via oxidative coupling of phenol **2.4** (Scheme 2.1).¹² The formation of this by-product **2.6** was supported by mass spectrometry *m/z* of 247.1 as the [M+H]⁺ and ¹H NMR spectrum consistent with the tri-substituted phenyl system of **2.6** (δ 6.95 (d, *J* = 8.8 Hz, 2H), 6.87 (dd, *J* = 8.8, 3.0 Hz, 2H), 6.82 (d, *J* = 3.0 Hz, 2H)) of the isolated by-product. Hence, the formation of oxidative by-products was prevented by using degassed solvents (prepared by bubbling N₂ through the solvent for 10 mins) and rigorous vacuum-nitrogen purging of the reaction vessel.

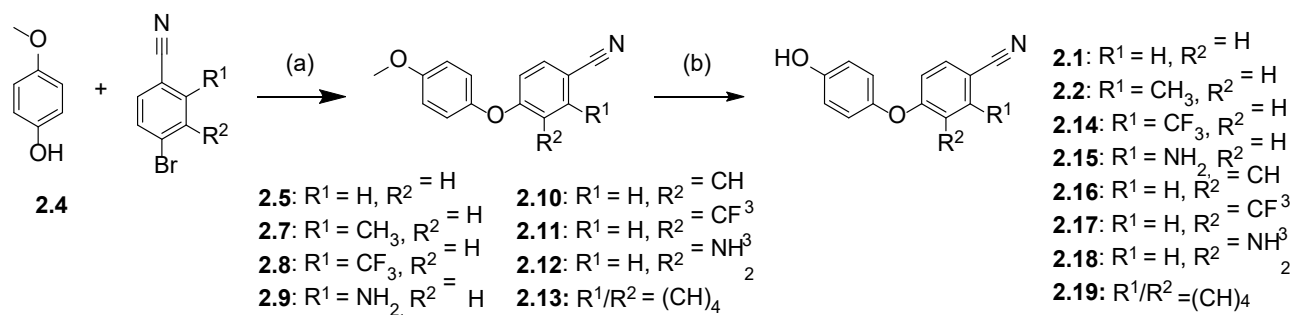
Scheme 2.1: Ullmann coupling of 4-cyanophenyl bromide **2.3** and 4-methoxyphenol **2.4** to form diaryl ether **2.5**.



Reagents and conditions: (a) *N,N*-dimethylglycine hydrochloride, Cs₂CO₃, CuI, 1,4-dioxane, 70 - 120 °C, 18 h.

Using these optimised conditions, a number of diaryl ether analogues **2.5** and **2.7** - **2.13** were synthesised (**Scheme 2.2**) and their formation was supported by the absence of the phenolic OH of **2.4** at δ 5.34 ppm in the ^1H NMR spectra. Literature suggested *N*-arylation or Buchwald-Hartwig coupling is preferentially mediated by palladium catalysts while *O*-arylation or Ullmann couplings are preferred by copper catalysts.¹³ The Ullman coupling of aminophenols **2.9** and **2.12** showed no sign of self-polymerisation of the aryl bromide on LCMS and therefore suggested good selectivity for *O*-arylation over *N*-arylation. In addition, the nitrogen is more electron deficient and sterically hindered than the phenol in these reactions. Subsequent demethylation of the anisole intermediates to phenols **2.1**, **2.2** and **2.14** – **2.19** was achieved using BBR_3 in DCM based on literature procedure¹¹ and was supported by the absence of the characteristic methoxy singlet resonance at δ 3.80 – 3.85 ppm in the ^1H NMR spectra.

Scheme 2.2: Synthesis of R^1/R^2 expanded series of analogues.

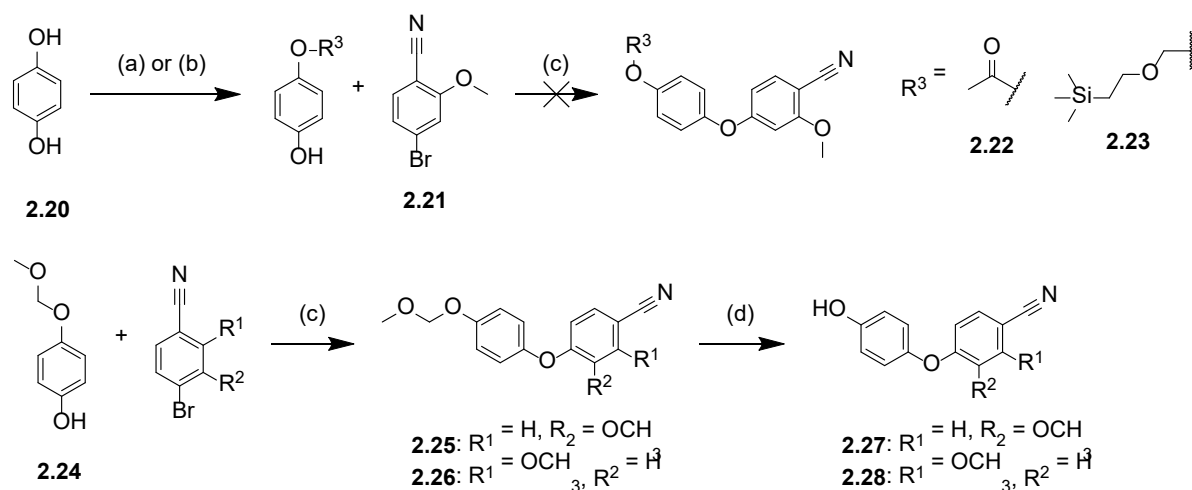


Reagents and conditions: (a) Cs_2CO_3 , CuI, *N,N*-dimethylglycine hydrochloride, 1,4-dioxane, 100 - 120 °C, 18 - 72 h, 19 - 55 %. (b) BBR_3 1 M in heptane, DCM, 0 - 25 °C, 3 - 6 h, 35 - 80 %.

An alternative protecting group to the previously utilised methyl ether was required to synthesise methoxy substituted phenols **2.27** and **2.28** (**Scheme 2.3**) as it was anticipated BBR_3 would provide no selectivity for demethylation of the desired 4-position methoxy group over the slightly more sterically hindered methoxy group at R^1 or R^2 positions. Selective mono-alkylation of hydroquinone (**2.20**) was achieved using a limiting amount of acetyl chloride or 2-(trimethylsilyl)ethoxy]methyl chloride (SEM chloride) to give the protected phenols **2.22** and **2.23**, respectively, in low yields.

However, Ullmann coupling of phenols **2.22** and **2.23** with aryl bromide **2.21** under standard conditions showed only minimal reaction progress by TLC and LCMS. The methoxyethoxymethyl (MEM) protecting group was also investigated, but the alkylation of hydroquinone failed to provide any suitable material for ongoing synthesis. Methoxymethyl (MOM) protected phenol **2.24** was purchased (Supplier: combi blocks) and coupled to the two alkyl bromides to give the methoxy intermediates **2.25** and **2.26** (Scheme 2.3). Deprotection of the MOM protecting group was achieved using methanolic hydrochloride (6 M) and is supported by the disappearance of the characteristic MOM methylene resonance at δ 5.09 ppm in the ^1H NMR spectra.

Scheme 2.3: Synthesis of methoxy derivatives **2.27** and **2.28** using MOM-protected phenol **2.24**.



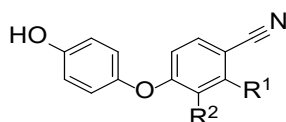
Reagents and conditions: (a) 2-(trimethylsilyl)ethoxymethyl chloride, Et₃N, DCM, 25 °C, 28 h, 26 %. (b) acetyl chloride, THF, 25 °C, 4 h, 27 %. (c) Cs₂CO₃, *N,N*-dimethylglycine hydrochloride, CuI, 1,4-dioxane, 120 °C, 24 h, 46 - 50 %. (d) hydrochloride in MeOH (6 M), H₂O, 80 °C, 16 h, 41 - 85 %.

2.4.3 Testing R¹ and R² expanded compounds

The solubility of synthesised analogues was determined via a 1D ^1H qNMR titration (as discussed above). Analogues were tested in single point ^1H - ^{15}N HSQC, followed by ^1H - ^{15}N HSQC titration and X-ray crystallography for those that showed binding to estimate their affinity and determine their binding mode with oxidised *Ec*DsbA (Table 2.1). Addition of a 3-methyl substituent for example in

2.16, afforded a 2-fold increase in affinity. However, this affinity value is potentially unreliable due to it being close to the maximal solubility of the compound (250 μM). The crystal structure of **2.16** had poor electron density which was well supported by a low EDIA score of 0.67 for the entire ligand. Similarly, the estimated affinity for the 2-methyl analogue **2.2** was solubility limited ($K_D = 600 \pm 90$ μM , solubility 500 μM , **Figure 2.7, a and b**). The refined crystal structure of 2-methyl analogue **2.2** had partial electron density (**Figure 2.7, d**), however the observable density clearly distinguishes the orientation of the core and position of the 2-methyl group (**Figure 2.7, c**), thus the binding mode determined for **2.2** was considered reasonable upon visual inspection.

Table 2.1: Testing of benzonitrile substituted diaryl ether analogues.



Compound number	R ¹	R ²	Solubility max (μM)	Single point HSQC NMR	HSQC NMR Affinity (K_D , μM)*	Electron density for the ligand	EDIA score
2.1	H	H	500	Binding	930 ± 180	-	-
2.2	CH ₃	H	500	Binding	600 ± 90	Partial	0.79
2.14	CF ₃	H	500	Binding	173 ± 35	Full	0.89
2.15	NH ₂	H	250	< 0.04 ppm	-	-	-
2.16	H	CH ₃	250	Binding	236 ± 50	Partial	0.67
2.17	H	CF ₃	500	Binding	330 ± 42	Poor	0.73
2.18	H	NH ₂	1000	Binding	717 ± 217	-	-
2.19	Naphthyl		< 125	-	-	-	-

2.27	OCH ₃	H	250	Binding	380 ± 70	Full	0.85
2.28	H	OCH ₃	250	< 0.04 ppm	-	-	-

*Affinities presented as mean ± SEM.

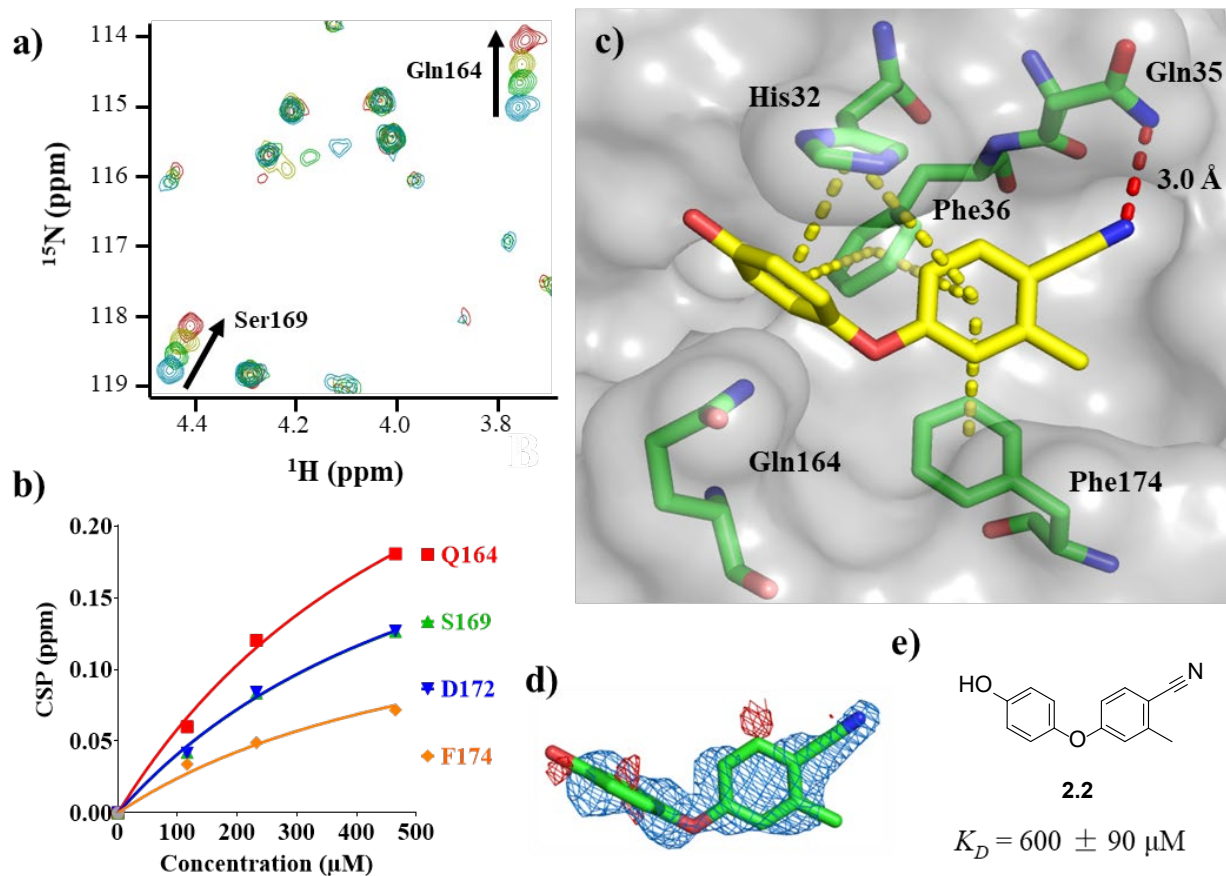


Figure 2.7: a) ^1H - ^{15}N HSQC titration of diaryl ether **2.2** with oxidized ^{15}N *EcDsbA*. Contours for ligand concentrations of 0, 125, 250, 500 μM as shown as blue, green, yellow and red respectively. b) Binding isotherm of 2-methyl analogue **2.2** plotting concentration (μM) vs CSP (ppm) and fit globally using a single site binding model with ligand depletion. c) X-ray crystal structure of 2-methyl analogue **2.2** (yellow sticks) bound to oxidised *EcDsbA* (grey surface). Key π -stacking interactions of **2.2** with His32, Phe36 and Phe174 (green sticks) are displayed as dashed yellow lines and a H-bond to Gln35 depicted by the dashed red line. d) Electron density of 2-methyl analogue **2.2** bound to *EcDsbA* shown as 2mFo-DFc at 1 σ (blue) and mFo-DFc at 3 σ (positive green and negative red). The EDIA score for this ligand was 0.78. e) 2D structure of 2-methyl analogue **2.2**.

Addition of an amino functional group as in 2-amino **2.15** and 3-amino **2.18** both resulted in a drop-in affinity. The 2-CF₃ derivative **2.14** was the most potent analogue from this series with a ≈ 5 -fold increase in affinity ($K_D = 173 \pm 35 \mu\text{M}$) compared to the parent unsubstituted core **2.1**. An X-ray crystal structure of 2-CF₃ derivative **2.14** bound to *EcDsbA* was solved and had full electron density for the ligand (**Figure 2.8, b**), however this fragment analogue flipped vertically in binding orientation placing the ether linker proximal to His32 (**Figure 2.8, a**).

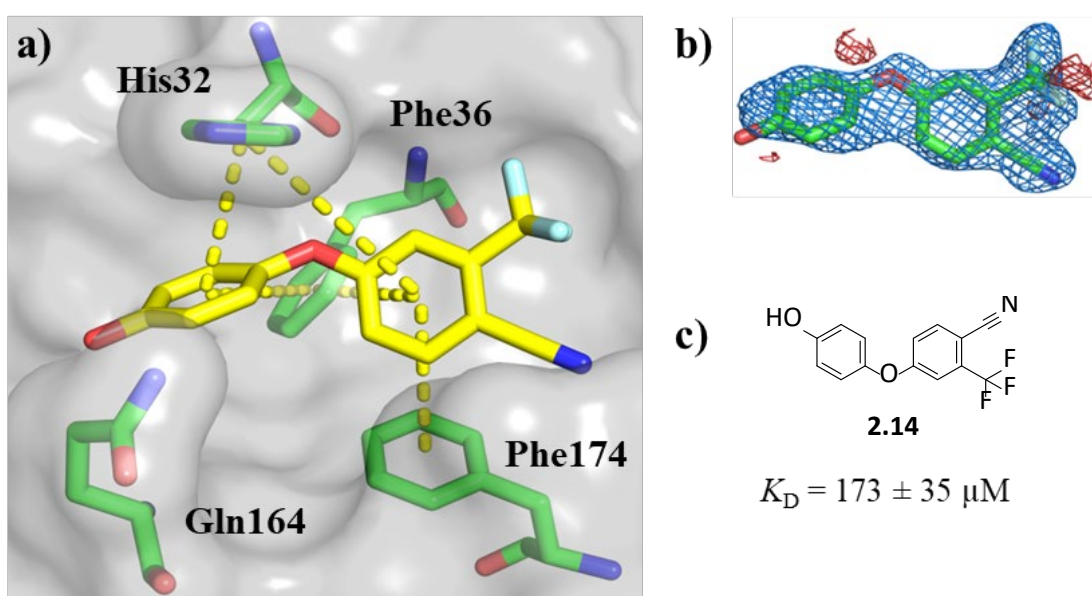


Figure 2.8: a) Crystal structure of 2-trifluoromethyl derivative **2.14** (yellow sticks) bound to oxidised *EcDsbA* (grey surface) with active site residues His32, Phe36, Gln164 and Phe174 (green sticks). π - π stacking interactions shown by yellow dashes. b) Electron density map of **2.14** bound to *EcDsbA* shown as 2mFo-DFc at 1 σ (blue) and mFo-DFc at 3 σ (positive green and negative red). c) 2D structure of trifluoromethyl analogue **2.14**.

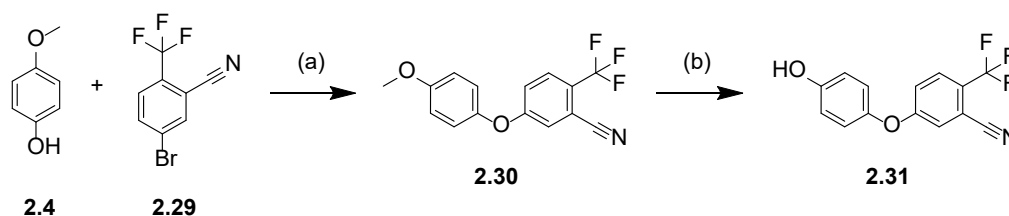
Shifting of the CF₃ group to the 3-position as in **2.17** resulted in a loss of ≈ 2 -fold decrease in affinity ($K_D = 330 \pm 42 \mu\text{M}$). Furthermore, the X-ray crystal structure of 3-CF₃ analogue **2.17** had only partial electron density for the core and no electron density to account for the CF₃ substituent. The 3-substituted methoxy analogue **2.28** showed insufficient CSP to estimate an affinity and was determined to have poor solubility (max 250 μM). In contrast the 2-methoxy isostere **2.27** had its

affinity determined by HSQC titrations to be $K_D = 380 \pm 70 \mu\text{M}$. However, the accuracy of this measurement may be affected by the limited solubility of this compound in aqueous buffer (max 250 μM). The crystal structure of the **2.27** had full electron density and a high EDIA score with the methoxy functional group projecting into the unfilled pocket. Naphthyl derivative **2.19** was synthesised to assess the flexibility of the bottom loop and test the limitations of introducing larger groups in this region, however this analogue had very low solubility and was observed to precipitate from solution at 125 μM and was not tested.

2.4.4 Synthesis and testing the *para*-CF₃ analogue.

Introducing a CF₃ group at the 2-position in the previous series gave compound **2.14**, which afforded the greatest improvement in affinity, however this analogue flipped in binding orientation. We hypothesised that switching the *meta*-CF₃ and *para*-nitrile functional groups of **2.14** would allow the molecule to adopt the parents binding orientation with ether linkage distal to the His32 while maintaining its higher affinity. Glide docking of the swapped substituent analogue **2.31** supported this hypothesis. Thus, reverse trifluoromethyl compound **2.31** was synthesised via the previously optimised route of Ullmann coupling and deprotection (**Scheme 2.4**).

Scheme 2.4: Synthesis of reversed analogue **2.31**.



Reagents and conditions: (a) Cs₂CO₃, *N,N*-dimethylglycine hydrochloride, CuI, 1,4-dioxane, 120 °C, 40 h, 38 %. (b) BBr₃ 1 M in heptane, DCM, 25 °C, 24 h, 76 %.

The reverse trifluoromethyl compound **2.31** was screened via ¹H-¹⁵N HSQC titrations which allowed us to calculate an estimated affinity of $K_D = 250 \pm 190 \mu\text{M}$. Again, it was found this compound had

relatively poor solubility (250 μM). The large error can be attributed to the small number of perturbing residues and solubility limitations of the compound. The crystal structures of both CF_3 containing analogues **2.14** and **2.31** had full electron density and high EDIA scores ($\text{EDIA} > 0.8$) for their respective ligands and when overlaid they show the CF_3 , *para*-cyano group and proximal phenyl ring overlaying into the same space on the proteins surface (**Figure 2.9**). Reverse trifluoromethyl compound **2.31** makes similar π - π stacking interactions with His32, Phe36 and Phe174 to other diaryl ether analogues and forms a potential polar interaction from the sidechain NH of Gln35 to a fluorine of the trifluoromethyl group of **2.31** (2.8 Å away, **Figure 2.9, b**).¹⁴ This in combination with other crystal structures suggests that the CF_3 group was determining the binding orientation. The difference in affinities between the reverse CF_3 analogue **2.31** ($K_D = 250 \pm 190 \mu\text{M}$) and 2-position trifluoromethyl analogue **2.14** ($K_D = 173 \pm 35 \mu\text{M}$) could potentially be as a result of the altered binding mode of **2.14** that places the ether linker proximal to the sidechain of His32.

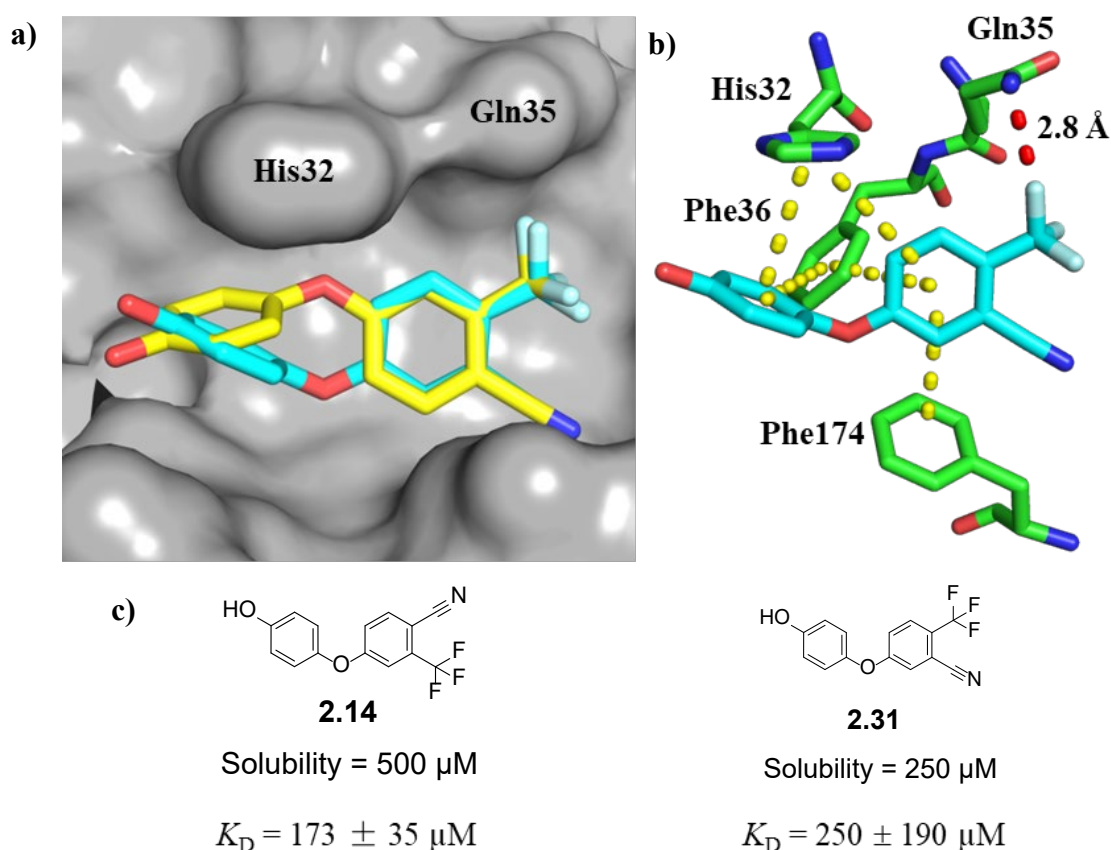


Figure 2.9: a) X-ray crystal structure of CF₃ analogues **2.14** and **2.31** (yellow and blue sticks, respectively) bound to *EcDsbA* (grey surface). Overlay of **2.14** with **2.31** shows the close proximity of the CF₃ group and proximal rings. b) Interactions formed between reverse analogue **2.31** (blue sticks) and key residues His32, Gln35, Phe36 and Phe172 (green sticks) of *EcDsbA*. c) Structures of trifluoromethyl analogues **2.14** and **2.31**

2.4.5 Summary

Over half of the analogues from this series either failed solubility testing or their affinities were unable to be accurately determined thus highlighting the importance of assessing solubility. There appeared to be no preference for substitution off the R¹ over the R² position in terms of affinities. Preliminary X-ray crystallography data suggested that R¹ substitution was more favourable in terms of the quality and quantity of the X-ray structures acquired. The introduction of a trifluoromethyl functional group in **2.14** saw a \approx 5-fold improvement in affinity, however the X-ray structure of this CF₃ analogue had the diaryl ether core flip in binding mode which is a common issue when developing fragments. The

addition of an amino, methyl or methoxy group saw little improvements in affinity and some unexpected losses in solubility. X-ray crystallography provided useful insight into the binding mode of these diaryl ether analogues with the 2-methyl **2.2** and 2-methoxy **2.27** fulfilling the objectives of this series by adopting the same binding orientation as the core while filling the empty pocket. EDIA scores were used to further quantify the quality a crystal structure and they correlated well with manual inspections of electron densities of the omit and refined X-ray structures. The 2-methyl substituted **2.2** was chosen for further elaboration in Chapter 3 as it maintained a similar binding orientation and solubility profile to that of the parent **2.1** while mildly increasing affinity.

2.5 Substitution of the *para*-cyano group

2.5.1 Design and docking

In the X-ray crystal structure of the parent diaryl ether **2.1** the *para*-cyano group forms a H-bond interaction with Gln35, hence substitution of the cyano functional group for a series of other isosteres and H-bond donors was investigated. This series was synthesised and investigated in parallel with the first series of analogues discussed above. Maintaining relatively high aqueous solubility for weakly binding fragments is essential to obtaining reliable affinity data as previously discussed. Hence, a number of carboxylic acids, amines and other polar analogues were prioritised in this series in an effort to improve the solubility of the diaryl ether core while optimising the cyano substituent. Glide docking of a series of diaryl ether analogues exploring this substitution suggested that it was possible to form the H-bond to Gln35 with an ester, alcohol or cyano methyl substituent. Substitution of the *para*-cyano for a carboxylic acid, acetic acid or amide functional group tended to make the diaryl ether core flip horizontally in binding orientation in the docked pose, which resulted in the formation of a new H-bond interaction from the carbonyl of the fragment to the sidechain aromatic NH of His32 as well as a H-bond from the terminal amide NH of the fragment to the backbone carbonyl oxygen of Val150 (**Figure 2.10, a and b**).

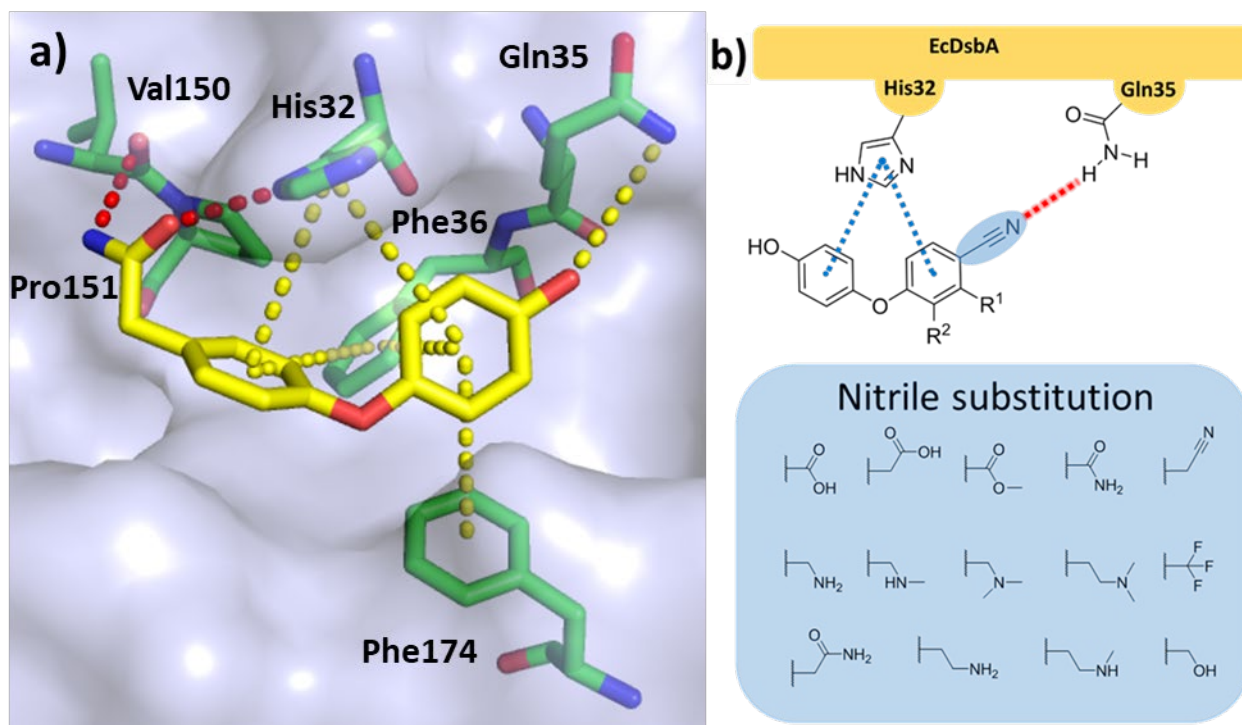
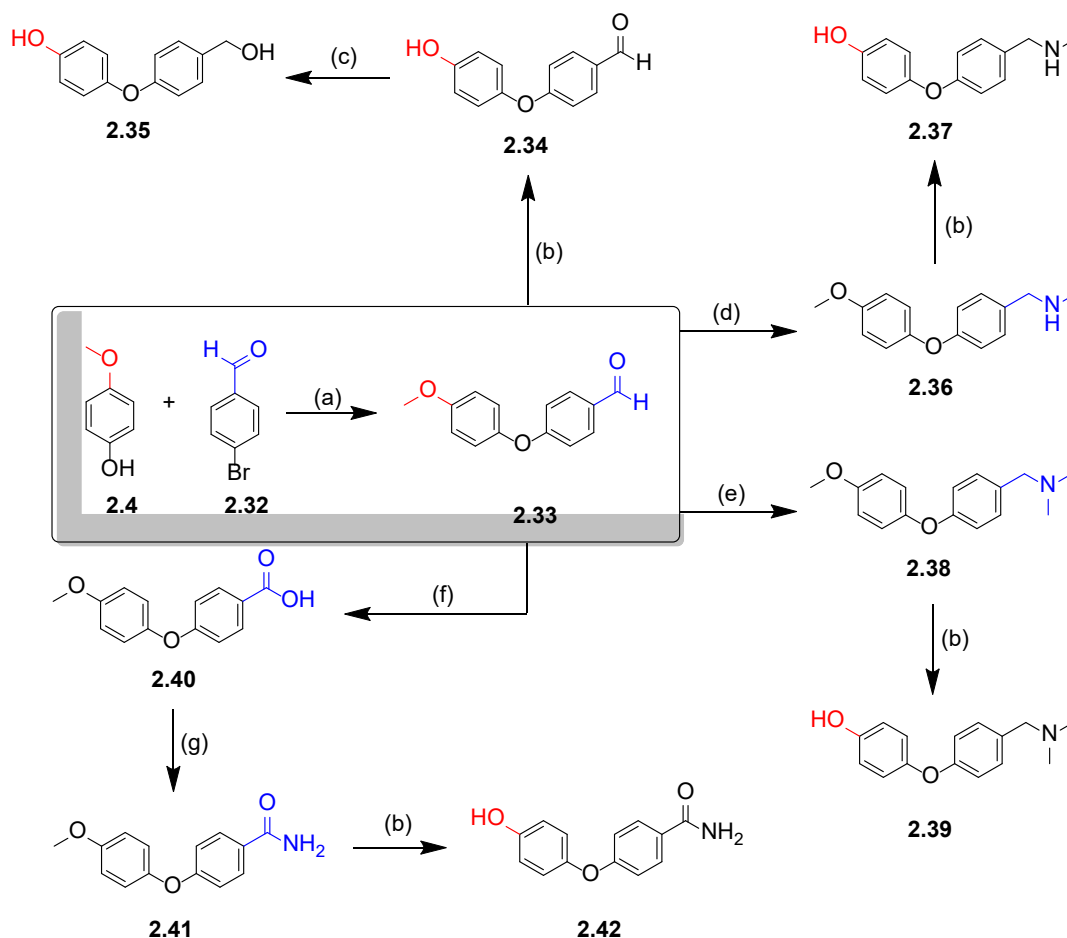


Figure 2.10: a) Docked structure of phenacetamide analogue to *EcDsbA*. Phenacetamide analogue (yellow sticks) and key residues His32, Phe36, Val150, Pro151 and Phe174 (green sticks) are shown. Potential H-bond interactions are shown as red dashed lines and π - π stacking as yellow dashes. b) Target compounds for synthesis in the *para*-nitrile replacement series focusing on improving interactions to Gln35 as depicted by the dashed red line.

2.5.2 Synthesis of *para*-cyano replacement analogues

Many alternative functional groups to the *para*-nitrile were afforded via the key intermediate *para*-benzaldehyde **2.33** (Scheme 2.5). Ullmann coupling of phenol **2.4** with aryl bromide **2.32** gave aldehyde **2.33** in excellent yield, the ^1H NMR spectrum of this analogue was concordant with the literature.¹⁵ Demethylation of **2.33** with BBr_3 followed by reduction of the aldehyde with LiAlH_4 afforded benzyl alcohol **2.35** in low yield after some optimisation (Scheme 2.5). This transformation was supported by the absence of the characteristic aldehyde resonance (δ 9.90 ppm) and appearance of the benzylic protons at δ 4.65 ppm in the ^1H NMR spectrum.

Scheme 2.5: Diaryl ether analogues synthesised from the key aldehyde intermediate **2.33**.

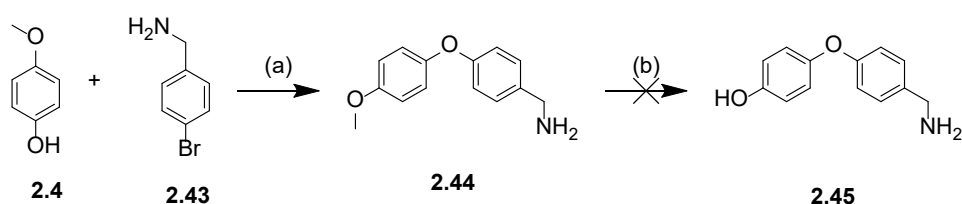
Reagents and conditions: (a) *N,N*-dimethylglycine hydrochloride, CuI, Cs₂CO₃, 1,4-dioxane, 120 °C, 18 h, 72 %. (b) BBr₃ 1 M in heptane, DCM, 6 h, RT, 12 - 52 %. (c) LiAlH₄, THF, reflux, 18 h, 48 %. (d) 33 % methylamine in EtOH, NaHCO₃, MeOH, reflux, 5 h then NaBH₄, MeOH, RT, 16 h, 60 %. (e) dimethylamine, Na(OAc)₃BH, Et₃N, DCE, RT, 3 h, 43 %. (f) NaClO₂, H₂O₂, NaH₂PO₄, THF, reflux, 5 h, 46 %. (g) SOCl₂, catalytic DMF, DIPEA, THF, RT, 16 h then NH₃ (30 % solution in H₂O), THF, RT, 16 h, 27 %.

Reductive amination of aldehyde **2.33** to methylamine **2.36** and dimethylamine **2.38** were achieved using NaBH₄ and Na(AcO)₃BH, respectively. Deprotection of the methyl ethers were achieved under standard conditions using BBr₃ and gave aminophenols **2.37** and **2.39**. Pinnick oxidation of aldehyde **2.33** to benzoic acid **2.40** was accomplished using NaClO₂ and H₂O₂. This transformation was supported by appearance of the characteristic carboxylic acid proton resonance (δ 12.74 ppm in ¹H

NMR spectra recorded in d_6 -DMSO) in accordance with the literature value.¹⁶ Benzoic acid **2.40** was converted to the acid chloride with SOCl_2 and quenched with aqueous NH_3 to give terminal amide **2.41**. This transformation was confirmed by the detection of the two amide protons (δ 7.23 and δ 7.92 – 7.82 ppm) as broad singlets in the ^1H NMR spectrum. Demethylation of amide **2.41** then afforded phenol **2.42** in good yields.

Ullmann coupling of the 4-bromobenzylamine (**2.43**) with phenol **2.4** under standard conditions afforded the desired diaryl ether **2.44** in low yields (Scheme 2.6). Deprotection of the methyl ether using BBr_3 was conducted and the $[\text{M}-\text{NH}_2]^+$ adduct for phenol **2.45** was identified at 199.0 m/z by mass spec. Multiple attempts to purify the crude material failed to give a high purity sample for testing as assessed by 1D ^1H NMR spectrum which contained a number of unidentified aliphatic resonances not belonging to phenol **2.45** at δ 1.30, 1.95 and 3.05 ppm. Thus, methoxy amine **2.44** was screened against *EcDsbA* while ongoing investigation in the synthesis of phenol **2.45** was stopped in preference for higher priority targets. Utilising an alternative protecting group to such as MOM could be investigated for the synthesis of phenol **2.45** in the future.

Scheme 2.6: Attempted synthesis of benzylamine **2.45** via Ullmann coupling and deprotection.

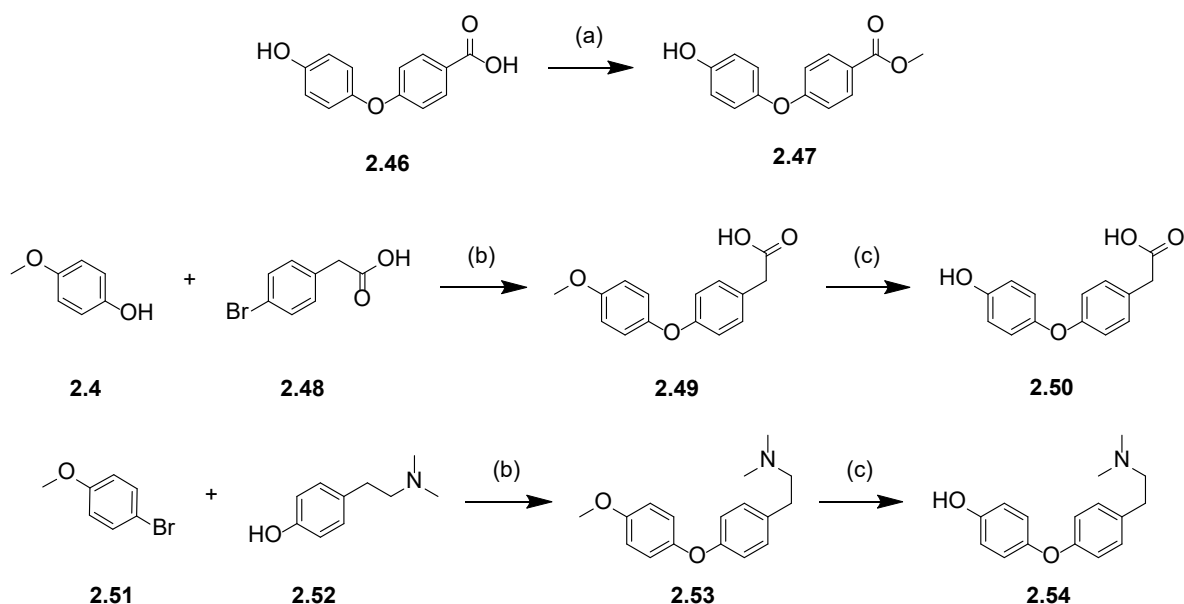


Reagents and conditions (a) *N,N*-dimethylglycine hydrochloride, 1,4-dioxane, 120 °C, 18 h, 23 %.

(b) BBr_3 1 M in heptane, DCM, RT, 18 h, 0 %.

Ester **2.47** was synthesised via Fischer esterification of commercially available carboxylic acid **2.46** (Scheme 2.7). Phenylacetic acid **2.50** and ethyl linked dimethylamino compound **2.54** were synthesised via the optimised Ullmann coupling conditions and methyl ether deprotected in overall poor yields (Scheme 2.7).

Scheme 2.7: Synthesis of ester **2.47**, phenylacetic acid **2.50** and dimethylamine **2.54** via Fischer esterification and the standard Ullmann coupling-deprotection method.

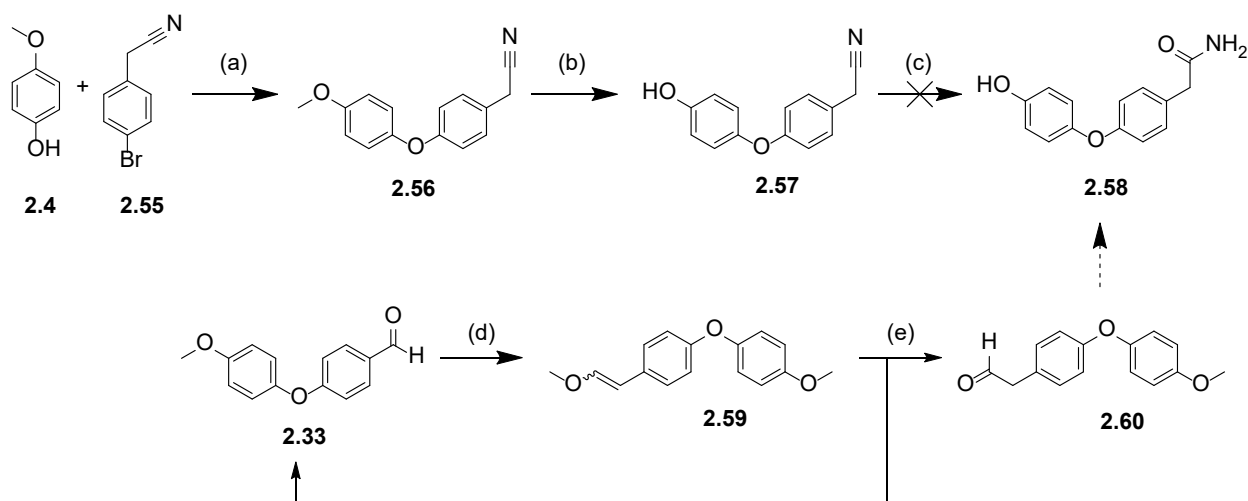


Reagents and conditions: (a) H_2SO_4 , MeOH, reflux 16 h, 51 %. (b) *N,N*-dimethylglycine hydrochloride, Cs_2CO_3 , CuI, 1,4-dioxane, 120 °C, 18 h, 9 - 19 %. (c) BBr_3 1 M in heptane, DCM, RT, 6 h, 5 - 29 %.

Methyl nitrile **2.57** was synthesised via Ullmann coupling and demethylation under the previously optimised conditions (**Scheme 2.8**). Phenylacetic acid **2.50** was synthesised in poor yields for both the Ullmann coupling and deprotection, thus alternative methods to terminal amide **2.58** were investigated involving oxidation of methyl nitrile **2.57** using sodium percarbonate ($2 \text{Na}_2\text{CO}_3 \cdot 3 \text{H}_2\text{O}_2$) following a literature procedure.¹⁷ Analysis of the crude reaction by LCMS suggested minimal amide **2.58** formation (m/z : 244.0 $[\text{M}+\text{H}]^+$) and attempts to purify the crude product returned only starting materials. Consequently, a new pathway to the phenyl acetamide **2.58** was investigated. Wittig homologation of the key intermediate aldehyde **2.33** to enol ether **2.59** (**Scheme 2.8**) with methoxymethyl(triphenyl)phosphonium chloride ($\text{PPh}_3\text{CH}_2\text{OMe}$) ylide was undertaken in an effort to afford the desired enol ether **2.59**. The success of this transformation was supported by the appearance of the alkene resonances (2:3 *cis:trans*, δ 5.12, 6.01 ppm *cis* configuration and δ 5.73, 6.90 ppm *trans* configuration) with characteristic coupling constants of 7.0 Hz and 13.0 Hz

respectively, in the ^1H NMR spectrum (**Appendix 2.1**). Acid mediated hydrolysis of the enol ether of **2.59** to acetaldehyde **2.60** using TFA or hydrochloride gave a 3:2 mixture of the desired aldehyde **2.60** as well as benzaldehyde **2.33**.

Scheme 2.8: Synthesis of methyl nitrile **2.57** and terminal amide **2.58** via oxidation and Wittig homologation.

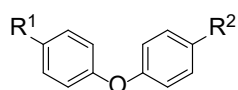


Reagents and conditions: (a) *N,N*-dimethylglycine hydrochloride, Cs_2CO_3 , CuI, 1,4-dioxane, 120°C , 18 h, 54 %. (b) BBr_3 1 M in heptane, DCM, 6 h, RT, 31 %. (c) $2\text{Na}_2\text{CO}_3 \cdot 3\text{H}_2\text{O}_2$, acetone, reflux, 24 h, 0 %. (d) KO t Bu, methoxymethyl(triphenyl)phosphonium chloride, THF, 16 h, RT, 36 %. (e) TFA or hydrochloride, DCM, 72 h, RT, not isolated.

The formation of acetaldehyde **2.60** was supported by the appearance of the aldehyde resonance as a triplet at δ 9.74 ppm and the formation of benzaldehyde **2.33** was confirmed by comparison with authentic product previously obtained. The 3:2 mixture of *trans* to *cis*-enols of **2.59** afforded a 3:2 mixture of acetaldehyde:benzaldehyde. This suggested the *trans* form undergoes acid hydrolysis as expected to form the desired aldehyde **2.60** while the *cis*-enol forms aldehyde **2.33**, though this was not confirmed. Attempts to optimise this deprotection were then halted in preference for higher priority target molecules.

2.5.3 Testing nitrile substituted analogues

A series of 14 substituted analogues were successfully synthesised and screened using the pipeline previously described. Only the benzylamine **2.45** and benzamide **2.58** could not be obtained, however this series provided sufficient information to begin to develop SAR for the core. Substitution of the *para*-nitrile group in parent fragment **2.1** for benzyl alcohol as in **2.35** and methyl nitrile as in **2.57** saw similar and reduced affinity, respectively (**Table 2.2**). The X-ray crystal structure of methyl nitrile **2.57** was obtained with good electron density and placed the extended nitrile functional group projecting close to but not directly towards Gln35 (**Figure 2.11, a and b**). The distance between the amide sidechain of Gln35 and the nitrile functional group of methyl nitrile **2.57** was 2.9 Å which suggested that the H-bond to Gln35 was maintained however the angle appeared less favourable. The series of amines **2.37–2.39**, **2.44**, **2.53** and **2.54** demonstrated weak or no CSP in ¹H-¹⁵N HSQC indicating that the compounds have very low occupancy when tested at their max solubility. An X-ray crystal structure of primary amine **2.44** was obtained, however the ligand could not be unambiguously orientated due to missing electron density.

Table 2.2: Screening and solubility of substituted benzonitrile derivatives.

Compound number	R ¹	R ²	Solubility max (μM)	Single point HSQC NMR	HSQC NMR Affinity (K _D , μM)*	Electron density for the ligand	EDIA score
2.35	OH		1000	Binding	733 ± 200	Poor	0.56
2.37	OH		1000	<0.04 CSP	-	-	-
2.38	OCH ₃		250	<0.04 CSP	-	-	-
2.39	OH		<125	-	-	-	-
2.40	OCH ₃		1000	Binding	1150 ± 400	Partial	0.83
2.42	OH		1000	<0.04 CSP	-	Poor	0.7
2.44	OCH ₃		1000	<0.04 CSP	-	Partial	0.82
2.46	OH		1000	Binding	>2000	Poor	0.72
2.47	OH		<125	-	-	-	-
2.49	OCH ₃		1000	Binding	950 ± 120	Full	0.91
2.50	OH		1000	Binding	>2000	-	-
2.53	OCH ₃		<125	-	-	-	-
2.54	OH		1000	<0.04 CSP	-	-	-
2.57	OH		1000	Binding	1360 ± 400	Full	0.93

*Affinities presented as mean ± SEM.

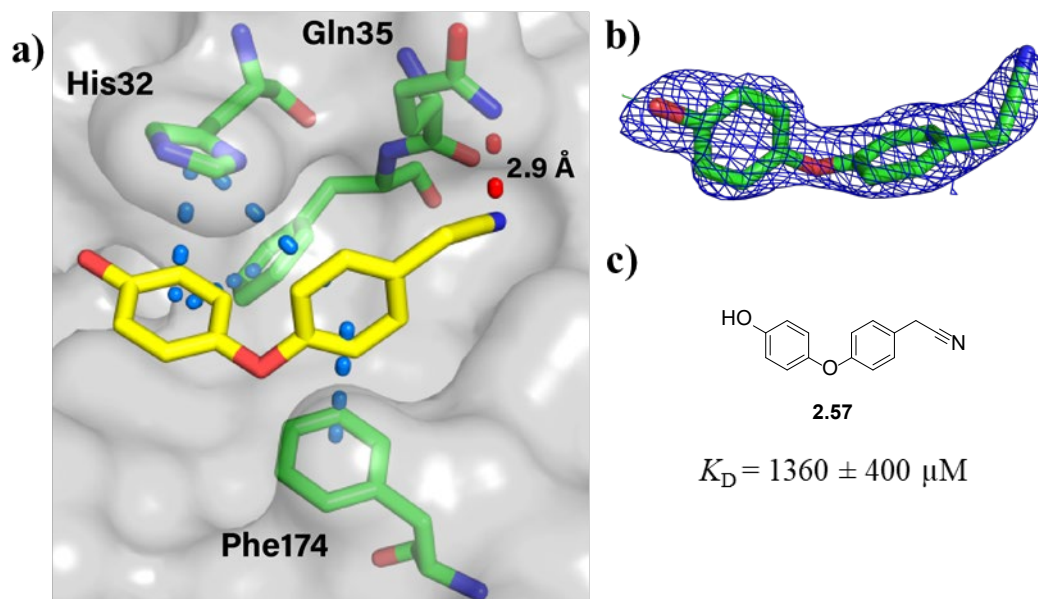


Figure 2.11: X-ray crystal structure of cyanomethyl analogue **2.57** (yellow sticks) bound to oxidised *EcDsbA* (green sticks and grey surface) with interacting residues His32, Gln35, Phe36 (unlabelled) and Phe174 displayed.

Substitution of the benzonitrile for a carboxylic acid functional group in benzoic acids **2.40**, **2.46** and phenylacetic acids **2.49**, **2.50** improved aqueous solubility but these analogues showed weak binding to *EcDsbA* by ^1H - ^{15}N HSQC. Similarly, to primary amine **2.44** the EDIA scores of both benzoic acids **2.40** and **2.46** were reasonable however, the orientation of the core could not be distinguished from the observed electron density in the omit map. An X-ray crystal structure of methoxy phenylacetic acid **2.49** bound to *EcDsbA* was solved with full electron density for the ligand, this was concordant with a high EDIA score (**Figure 2.12, a and b**). The X-ray structure of **2.49** flipped in binding mode and formed a new H-bond interaction to His32 similar to the pose predicted by docking (**Figure 2.10**). This X-ray crystal structure, combined with the moderate affinity ($K_D \approx 950 \mu\text{M}$) inspired the design and synthesis of a number of phenylacetic acid derivatives described in Chapter 3. Ester **2.47** demonstrated low aqueous solubility, while benzamide analogue **2.42** was soluble at $1000 \mu\text{M}$ but bound weakly to *EcDsbA*.

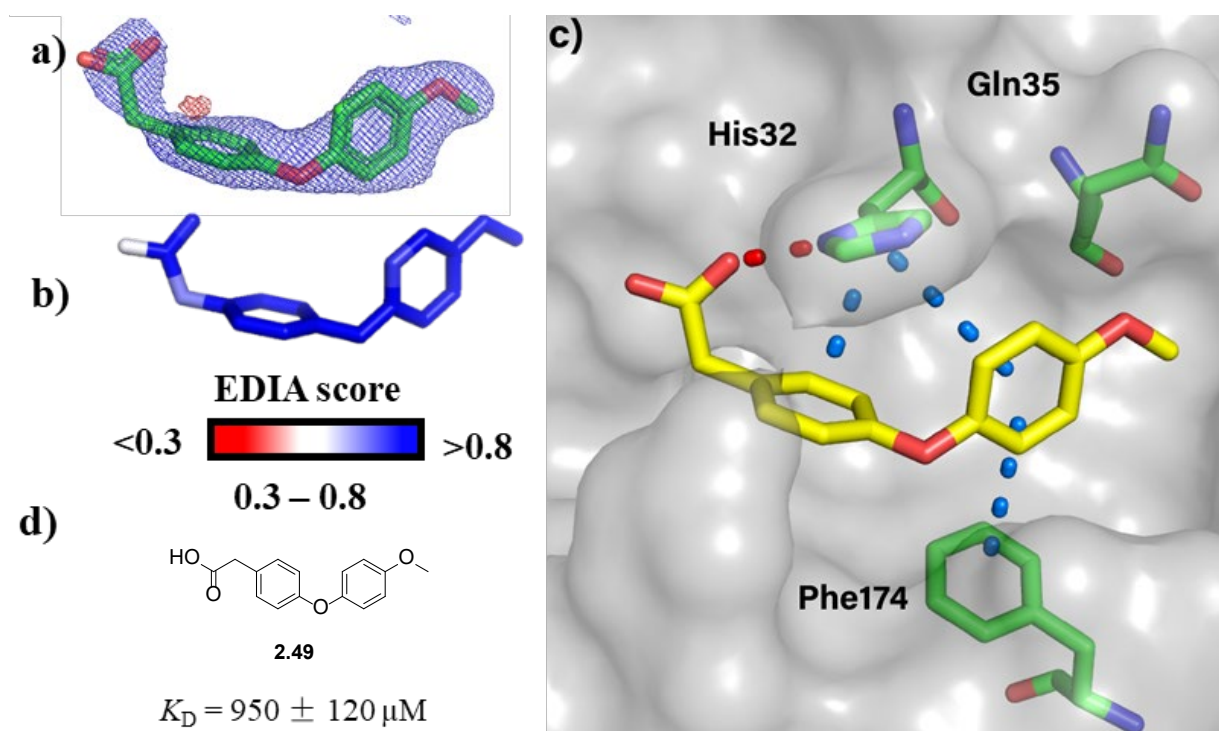


Figure 2.12: a) Electron density map of phenylacetic acid **2.49** bound to *EcDsbA* shown as 2mFo-DFc at 1 σ (blue) and mFo-DFc at 3 σ (positive green and negative red). b) EDIA scores for each individual atom colour coded onto structure of **2.49**, shown as a gradient of red-white-blue (<0.3 to 0.55 to >0.8). c) Crystal structure of phenylacetic acid **2.49** (yellow sticks) bound to oxidised *EcDsbA* (green sticks and grey surface), π -stacking interactions shown by dashed blue lines and H-bonds shown as red dashed lines. d) Structure of phenylacetic acid **2.49**.

2.5.4 Summary

Replacement of the *para*-nitrile with various polar functional groups including acids, amines, amides, esters and alcohols generally improved the solubility of the diaryl ether core with the majority of analogues passing solubility a 1 mM. However, these analogues demonstrated weak CSP and low binding affinities determined by ^1H - ^{15}N HSQC NMR. No analogues in this series displayed superior binding affinity compared to the parent fragment **2.1**, thus the *para*-nitrile remained the most suitable functional group for its balance of moderate solubility and affinity. Phenylacetic acid **2.49** and benzyl alcohol **2.35** showed similar affinity to the parent diaryl ether **2.1** with phenylacetic acid **2.49** horizontally flipped binding mode in its X-ray crystal structure and formed a H-bond interaction with

His32. This interaction was initially predicted by docking and was used to help design a number of analogues described in the next Chapter.

2.6 Bicyclic diaryl ethers

2.6.1 Design and docking

The addition of a methyl group on the benzonitrile ring at the 2 position as in **2.2** elicited a minor increase in affinity (**2.2** $K_D = 600 \mu\text{M}$ vs **2.1** $K_D = 930 \mu\text{M}$), while maintaining reasonable solubility ($500 \mu\text{M}$) and adopting the same binding mode to the parent fragment **2.1**. This 2-position of the benzonitrile ring was chosen for further expansion, however the commercial availability of 2-substituted-1-cyano-4-halobenzene reagents was low. Additionally, such compounds had low synthetic tractability due to unfavourable electronics and sterics of the ring system which made it difficult to access expanded analogues at this position. Hence, we designed a series of constrained lactams, lactones and other carbonyl containing 6-6 and 6-5 bicycles that could mimic the general pharmacophore of *para*-H-bond acceptor and *meta*-hydrophobic group (**Figure 2.13**). More specifically the directly attached carbonyl was designed to mimic the *para*-cyano group as a H-bond acceptor and cyclisation of the ring to fill the empty pocket and form hydrophobic interactions as seen in the crystal structure of 2-methyl analogue **2.2**.

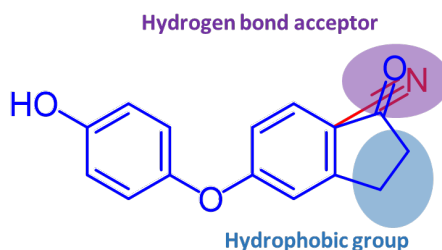
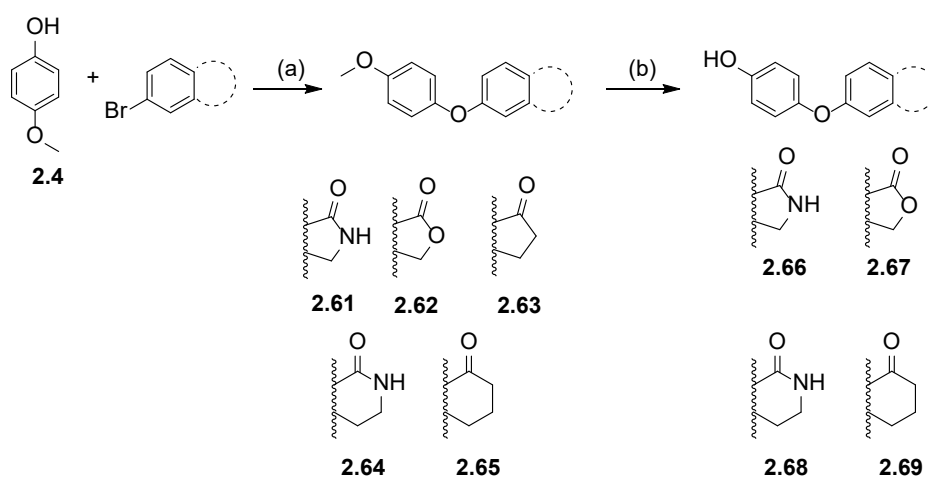


Figure 2.13: General structural motif of a bicyclic diaryl ether analogue with H-bond acceptor (highlighted in purple) to form a H-bond with Gln35 and hydrophobic group (highlighted in blue) to make hydrophobic contacts with EcDsbA.

A list of commercially available reagents were filtered to obtain aryl halides (Br, I) and were virtually enumerated in a chemoinformatics program with *para*-benzoquinone to give a library of diaryl ethers. Analogues containing pan-assay interference compounds (PAINS)¹⁷ and other undesirable functional groups were filtered out using an inhouse maintained set of SMARTS filters derived from the literature¹⁷ in KNIME (version 3.3.1) and the remaining compounds docked against *Ec*DsbA using Glide along with the designed bicyclics described above. The docking results were filtered by an RMSD of < 1 Å from the diaryl ether core which only returned docking poses in a similar binding orientation to the parent molecule. This was required as initial unfiltered docking attempts found > 70 % of the top 200 poses analysed did not maintain binding orientation. The top 200 poses from the RMSD filtered docking were visually inspected and ranked based on maintained interactions, binding orientation and overall fit. This indicated the above designed lactone, lactam and cyclic ketone bicycle series was favoured over other commercially available substituents at this position. Hence the synthesis of a number of lactone, lactam and bicycle analogues was pursued.

2.6.2 Synthesis of bicyclic diaryl ethers

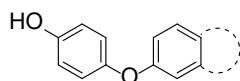
Butyrolactam **2.61**, butyrolactone **2.62** and 6-membered lactam **2.64** were synthesised via copper-mediated Ullmann coupling of phenol **2.4** with their respective aryl bromides (**Scheme 2.9**). Demethylation of the methyl ether was achieved using BBr₃ to give phenols **2.66**, **2.67** and **2.68** in good yields. The Ullmann coupling to give the cyclohexanone analogue **2.65** proceeded in low yield after 14 days at > 100 °C to give cyclohexanone **2.65** which was demethylated to give the corresponding phenol **2.69**. Multiple attempts to synthesise the cyclopentanone **2.63** via the Ullmann coupling reactions proved unsuccessful (**Appendix 2.2**). Optimisation of the reaction conditions by changing the ligand (*N,N*-dimethylglycine hydrochloride or phenanthroline), time (1.5 – 14 d), temperature (95 – 120 °C) solvent (1,4-dioxane or DMF) and purification conditions yielded crudes with new peaks by LCMS, which could not be confirmed as product or isolated in high purity for ongoing synthesis. Hence, the synthesis of cyclopentanone **2.63** was halted.

Scheme 2.9: Synthesis of a series of bicyclic diaryl ethers by Ullmann coupling and deprotection.

Reagents and conditions: (a) *N,N*-dimethylglycine hydrochloride, Cs_2CO_3 , CuI, 1,4-dioxane, 100 - 120 °C, 1 - 14 d, 5 - 64 %. (b) BBr_3 1 M in heptane, DCM, RT, 1 - 24 h, 36 - 71 %.

2.6.3 Testing bicyclic diaryl ether analogues

Solubility of the bicyclic analogues was poor with 5-membered lactam **2.66** and lactone **2.67** precipitating at 125 μM (**Table 2.3**). The 6-membered analogues lactam **2.68** and ketone **2.69** passed solubility testing at 250 and 125 μM , respectively. Cyclic ketone **2.69** showed significant HSQC perturbations however an affinity could not be estimated due to a limited number of residues showing perturbation and the low solubility of the compound.

Table 2.3: Testing, solubility and X-ray crystallography of the bicyclic diaryl ether compounds.

Compound number	R ¹	Solubility max (μM)	Single point HSQC NMR	Electron density for the ligand	EDIA score
2.66		<125	-	-	
2.67		< 125	-	-	
2.68		250	< 0.04 CSP	Poor	0.68
2.69		125	Binding	-	-

2.7 Conclusion

Multiple rounds of design, docking and synthesis focused on optimisation of the benzonitrile ring of the diaryl ether core were investigated. The aqueous solubility of analogues was closely monitored throughout these series with particular emphasis on designing compounds with polar functionalities (carboxylic acids, amines etc). Few SAR conclusions could be drawn from this initial series as improved affinities as determined by HSQC were limited by poor solubility. X-ray crystallography was used to inform decisions for further design and medicinal chemistry. This structural data suggested that further substitution off the benzonitrile ring was favoured by substituents *ortho* to the cyano group such as in 2-methyl substituted analogue **2.2**. Replacement of the *para*-nitrile was attempted to improve aqueous solubility and affinity through improved interactions with Gln35. This substitution improved solubility but also reduced binding affinity as all analogues in the series had

similar or weaker affinity compared to the parent diaryl ether **2.1**. X-ray crystallography demonstrated that multiple analogues including phenylacetic acid **2.49** flipped horizontally in binding orientation and formed a new H-bond interaction to His32 as predicted by the initial docking. A small series of bicyclic diaryl ether analogues were also designed to further explore substituents *ortho* to the cyano group, however, similar to the initial expansion series these analogues demonstrated poor aqueous solubility or weak binding to EcDsbA. The low aqueous solubility of analogues within this Chapter resulted in few analogues having reliable affinity estimates and hence produced inconclusive SAR around this benzonitrile ring, although the data indicated that none of the analogues showed large gains in affinity. X-ray crystallography suggested the 2-methyl substituent as in **2.2** was well tolerated and this substitution pattern was taken forward in the next series where the design focuses on replacing the *para*-hydroxyl with a series of carboxylic acid derivatives in order to improve solubility and potentially form interactions with His32.

2.8 References

1. LaPlante, S. R.; Carson, R.; Gillard, J.; Aubry, N.; Coulombe, R.; Bordeleau, S.; Bonneau, P.; Little, M.; O'Meara, J.; Beaulieu, P. L. Compound aggregation in drug discovery: implementing a practical NMR assay for medicinal chemists. *J. Med. Chem.* **2013**, 56, 5142-50.
2. Davis, B. J.; Erlanson, D. A. Learning from our mistakes: the 'unknown knowns' in fragment screening. *Bioorg. Med. Chem. Lett.* **2013**, 23, 2844-52.
3. Bentley, M.; Doak, B. C.; Mohanty, B.; Scanlon, M. J. *Applications of NMR Spectroscopy in FBDD*. 2nd ed.; Modern Magnetic Resonance, 2017; p 22.
4. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, 73, 1-16.
5. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **2010**, 66, 486-501.
6. Liebschner, D.; Afonine, P. V.; Moriarty, N. W.; Poon, B. K.; Sobolev, O. V.; Terwilliger, T. C.; Adams, P. D. Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr. D* **2017**, 73, 148-157.
7. Meyder, A.; Nittinger, E.; Lange, G.; Klein, R.; Rarey, M. Estimating Electron Density Support for Individual Atoms and Molecular Fragments in X-ray Structures. *J. Chem. Inf. Model.* **2017**, 57, 2437-2447.
8. Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Energetics of Displacing Water Molecules from Protein Binding Sites: Consequences for Ligand Optimization. *J. Am. Chem. Soc.* **2009**, 131, 15403-15411.
9. Repasky, M. P.; Shelley, M.; Friesner, R. A. Flexible ligand docking with Glide. *Curr. Protoc. Bioinformatics* **2007**, Chapter 8, Unit 8.12.
10. Naidu, A. B.; Jaseer, E. A.; Sekar, G. General, Mild, and Intermolecular Ullmann-Type Synthesis of Diaryl and Alkyl Aryl Ethers Catalyzed by Diol-Copper(I) Complex. *J. Org. Chem.* **2009**, 74, 3675-3679.
11. Sandanayaka, V.; Singh, J.; Gurney, M.; Mamat, B.; Yu, P.; Bedel, L.; Zhao, L. Biaryl substituted heterocycle inhibitors of LTA4H for treating inflammation. 2007.
12. Malkowsky, I. M.; Rommel, C. E.; Fröhlich, R.; Griesbach, U.; Pütter, H.; Waldvogel, S. R. Novel Template-Directed Anodic Phenol-Coupling Reaction. *Chem. Eur. J.* **2006**, 12, 7482-7488.
13. Shafir, A.; Lichtor, P. A.; Buchwald, S. L. N- versus O-Arylation of Aminoalcohols: Orthogonal Selectivity in Copper-Based Catalysts. *J. Am. Chem. Soc.* **2007**, 129, 3490-3491.
14. Dalvit, C.; Invernizzi, C.; Vulpetti, A. Fluorine as a hydrogen-bond acceptor: experimental evidence and computational calculations. *Chemistry* **2014**, 20, 11058-68.
15. Zheng, X.; Ding, J.; Chen, J.; Gao, W.; Liu, M.; Wu, H. The Coupling of Arylboronic Acids with Nitroarenes Catalyzed by Rhodium. *Org. Lett.* **2011**, 13, 1726-1729.
16. Jung, N.; Bräse, S. Diaryl Ether and Diaryl Thioether Syntheses on Solid Supports via Copper (I)-Mediated Coupling. *J. Comb. Chem.* **2009**, 11, 47-71.
17. Kabalka, G. W.; Deshpande, S. M.; Wadgaonkar, P. P.; Chatla, N. The Transformation of Nitriles into Amides Using Sodium Percarbonate. *Synth. Commun.* **1990**, 20, 1445-1451.

Chapter 3 - Optimisation of the diaryl ether core.

3.1 Introduction

Two key analogues were identified from Chapter 2 the 2-methyl benzonitrile **3.1** and methoxy phenylacetic acid **3.2** (**Figure 3.1**). X-ray crystallography data for these two analogues provided the basis for replacing the *para*-hydroxy substituent and merging these key analogues. The LHS of the hydrophobic groove is significantly larger than the RHS. Although to this point, efforts to optimise expansions into the LHS had not been particularly successful. Consequently, we hypothesised that optimising the current scaffolds and exploring other potential vectors prior to expansion to the LHS would provide the best chance of success. Producing compounds with a single low energy binding pose is desirable to reduce the chance of fragments reorientating themselves within the binding groove which can complicate interpretation of the SAR as we look ahead to exploring the LHS of the groove. Hence, this Chapter details the optimisation of substituents on the phenol ring, merging of the best analogues from the previous Chapter as well as incorporating previous SAR generated around the benzonitrile ring. Further investigations into the effects of isosteric replacement of the ether linker to enable expansion to a potential bottom vector of the binding site are also discussed.



Figure 3.1: a) Chemical structure and key features of 2-methyl benzonitrile **3.1**. b) Chemical structure and key properties of *para*-phenylacetic acid **3.2**.

3.2 Replacement of the *para*-hydroxyl and merging the best analogues

3.2.1 Design and docking

The solubility of the diaryl ethers described above were poor thus restricting our ability to accurately determine the binding affinity for many analogues. The work in Chapter 2 suggested that a phenylacetic acid functional group (as in compound **3.2**, **Figure 3.1**) could form an additional H-bond interaction with the sidechain of His32. We therefore investigated substitution of the *para*-hydroxyl for a series of polar functional groups to improve affinity and solubility of the diaryl ethers. Analogues that contained a synthetic handle (alcohol, sulfonic acid, carboxylic acid, amide etc) for expansion were prioritised. Furthermore, addition of the 2-methyl substituent on the benzonitrile ring improved affinity. We hypothesised that merging of these two analogues through substituting the *para*-hydroxy of the 2-methyl analogue **3.1** for a series of carboxylic acid functional groups (**Figure 3.2, a**) may improve the analogues solubility and potentially form a H-bond interaction with His32.

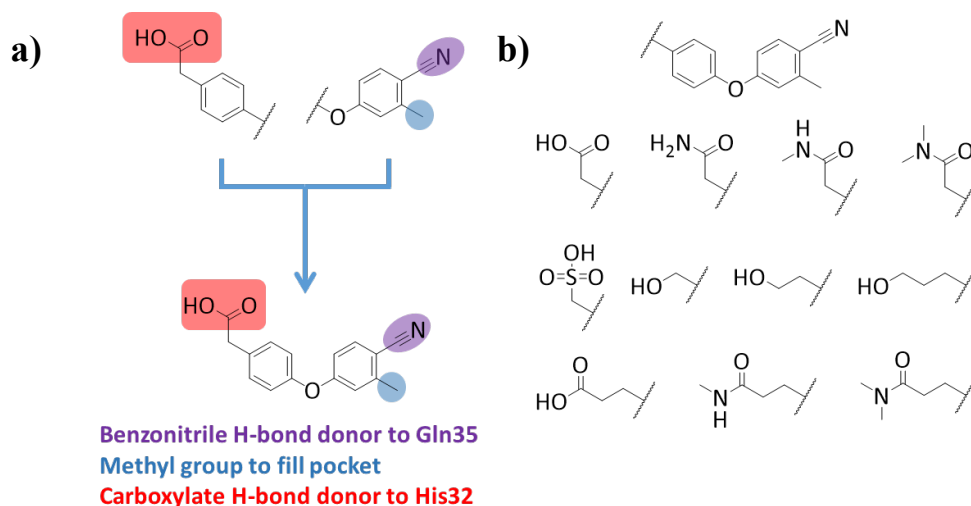


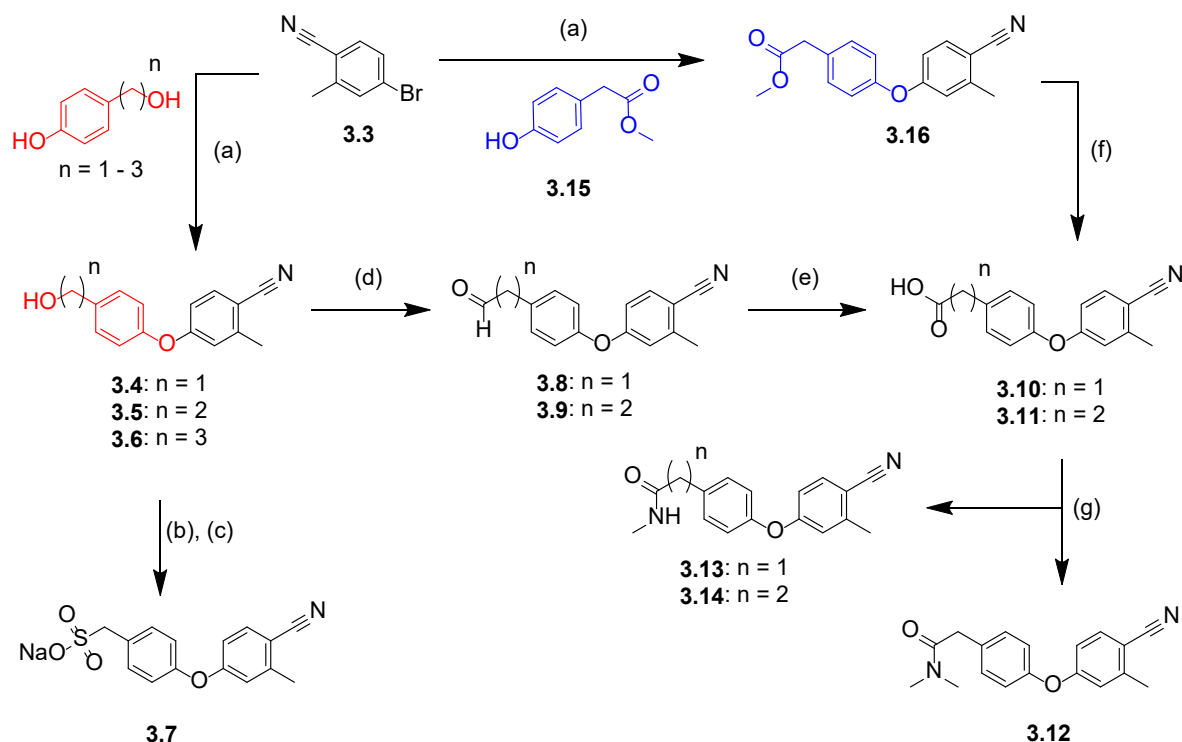
Figure 3.2: a) Merging of the best features of methyl nitrile **3.1** and phenylacetic acid **3.2** identified in Chapter 2. b) Series of polar substituents selected as replacements to the *para*-hydroxy group.

Glide docking was performed as previously described in chapter 2 on a series of carboxylic acid, alcohol and amide derivatives of varying lengths (**Figure 3.2, b**). The docking suggested that all but the directly attached benzoic acid analogues made the H-bond interaction with His32 while

maintaining a similar core orientation to other diaryl ether analogues. Hence, a series of carboxylic acid and amide analogues were synthesised to test this hypothesis.

3.2.2 Synthesis

The diaryl ether core was synthesised via the previously optimised copper-mediated Ullmann coupling reaction to give 1, 2 and 3 carbon linked alkyl alcohols **3.4** – **3.6** in moderate to high yields (**Scheme 3.1**). Bromination of benzyl alcohol **3.4** was achieved using NBS/PPh₃ followed by nucleophilic displacement with Na₂SO₃ then gave sulfonic acid **3.7** with an overall yield of 23 % over the 2 steps. Dess-Martin periodinane was used to selective oxidise the 2 and 3 carbon linked alcohols **3.5** and **3.6** to aldehyde intermediates **3.8** and **3.9**, respectively. This was supported by the appearance of an aldehyde resonance ($\delta \approx 9.8$ ppm (t, $J = 1.3$ Hz)) in the ¹H NMR spectra. Pinnick oxidation of aldehydes **3.8** and **3.9** using NaClO₂ and H₂O₂ gave phenylacetic acid **3.10** and propanoic acid **3.11**, respectively. Phenylacetic acid **3.10** however, was more efficiently synthesised in 2 steps via Ullmann coupling of methyl ester protected phenol **3.15** with aryl bromide **3.3** to give diaryl ether ester **3.16**, followed by base-mediated ester hydrolysis. All further synthesis of phenylacetic acid **3.10** was achieved via this Ullmann coupling and ester hydrolysis pathway. Amide coupling of carboxylic acids **3.10** and **3.11** with methylamine hydrochloride or dimethylamine using *O*-(6-chlorobenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HCTU) as the coupling agent then gave dimethyl amide **3.12** and methyl amides **3.13** and **3.14** in moderate yields.

Scheme 3.1: Synthesis of phenol replacement analogues.

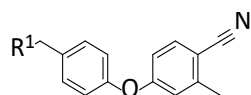
Reagents and conditions: (a) *N,N*-dimethylglycine hydrochloride, Cs₂CO₃, CuI, DMF, 120 °C, 18 h, 48 - 78 %. (b) NBS, PPh₃, DCM, 40 °C, 5 h. (c) Na₂SO₃, EtOH/H₂O, 40 °C, 20 h, 23 % over two steps. (d) Dess-Martin periodinane, DCM, RT, 16 h, 38 - 69 %. (e) NaClO₂, KH₂PO₄, H₂O₂, MeCN, RT, 3 h, 44 - 46 %. (f) LiOH, THF/H₂O, RT, 16 h, 95 %. (g) CH₃NH₂ hydrochloride or (CH₃)₂NH, HCTU, Et₃N, MeCN, RT, 16 h, 50 - 56 %.

3.2.3 Screening phenol replacement analogues

These analogues were screened via the solubility, HSQC and X-ray crystallography pipeline previously described in Chapter 2. Introduction of the benzyl alcohol as in **3.4** improved the overall solubility of the core ($\geq 1000 \mu\text{M}$) while slightly improving the affinity 1.6-fold ($K_D = 380 \pm 35 \mu\text{M}$, **Table 3.1**) compared to the phenol variant **3.1** ($K_D = 600 \pm 90 \mu\text{M}$). Extension to phenethyl alcohol **3.5** reduced solubility while maintaining the slight improvement in affinity. The propyl alcohol as in **3.6** comparatively improved affinity ($K_D = 90 \pm 10 \mu\text{M}$) over the shorter chain alcohols but significantly reduced solubility ($\leq 250 \mu\text{M}$). The X-ray crystal structures of these alcohols **3.4** – **3.6**

were missing electron density for their expansions from the core and provided no insight into why the propyl alcohol **3.6** had higher affinity than the other related analogues.

Table 3.1: Testing, solubility and X-ray crystallography data for the phenol replacement analogues.



Compound number	R ¹	Solubility max (μ M)	Single point HSQC NMR	HSQC NMR Affinity (K_D , μ M)*	Electron density for the ligand	EDIA score
3.4		1000	Binding	380 ± 40	Partial	0.81
3.5		500	Binding	430 ± 120	Partial	0.87
3.6		250	Binding	90 ± 10	Partial	0.75
3.7		1000	Binding	530 ± 30	Full	0.93
3.10		1000	Binding	$490 \pm 20^{**}$	Full	0.86
3.11		1000	Binding	430 ± 30	Partial	0.77
3.12		125	-	-	-	-
3.13		250	Binding	640 ± 250	-	-
3.14		<125	-	-	-	-

*Affinities presented as mean \pm SEM. **Affinity determined in phosphate buffer.

Sulfonic acid **3.7** and phenylacetic acid **3.10** analogues demonstrated good aqueous solubility (≥ 1000 μ M) but saw no significant improvement in affinity, despite the X-ray crystal structures of both analogues adopting a similar binding mode to the 2-methyl analogue **3.1** and form a new H-bond to

His32 (**Figure 3.3, a**). Extension of the phenylacetic acid to a propanoic acid in **3.11** maintained solubility but showed no improvements in affinity, however the obtained crystal structure of the propanoic acid **3.11** bound to *EcDsbA* was missing electron density for the carboxylic acid expansion. *N*-methyl amide **3.12** and **3.14** as well as dimethyl amide **3.13** had significantly reduced aqueous solubility, likely due to loss of the negative charge of the carboxylic acid, and could not have their K_D values determined by HSQC titrations.

The affinity measurement for 2-methyl phenol **3.1** was not accurately able to be determined due to poor solubility. Substitution of the phenolic OH for an acetic or sulfonic acid functional group (as in **3.10** and **3.7**) improved the solubility of the core (500 μ M to 1000 μ M) and therefore strengthened our confidence in the estimated affinity for these analogues (**3.10** shown in **Figure 3.3, d and e**). Both analogues **3.10** and **3.7** displayed strong electron density in their X-ray crystal structures (**3.10** shown in **Figure 3.3, a and c**). In addition to HSQC NMR based affinity screening several of the most promising diaryl ether analogues were analysed via SPR. The SPR assay to this point was limited by the lack of a suitable positive control, therefore the most promising analogues were investigated to both validate the SPR assay and further confirm the binding of key compounds. Phenylacetic acid **3.10** was tested against *EcDsbA* in the SPR assay and had an affinity of $K_D = 308 \pm 50 \mu$ M (**Figure 3.3, f and g**), concordant with the affinity estimated by ^1H - ^{15}N HSQC titrations.

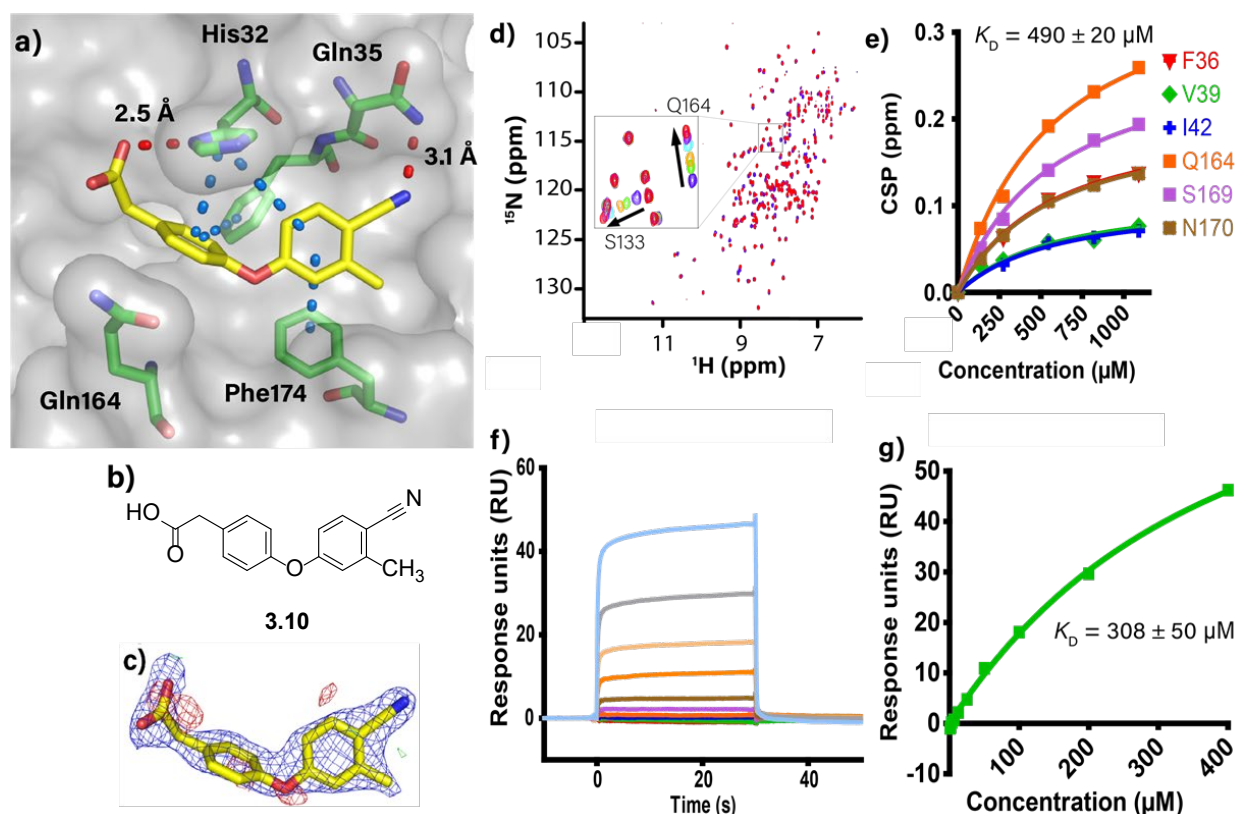


Figure 3.3: a) X-ray crystal structure of phenylacetic acid **3.10** (yellow sticks) bound to oxidised *EcDsbA* (grey surface and green sticks for key interacting residues). H-bond interactions are shown as red dashed lines and π - π stacking interactions are shown as blue dashed lines. b) 2D structure of phenylacetic acid **3.10**. c) Electron density map of **3.10** shown as 2mFo-DFc at 1 σ (blue) and mFo-DFc at 3 σ (positive green and negative red). d) ^1H - ^{15}N HSQC spectra of oxidised u- ^{15}N *EcDsbA* with phenylacetic acid **3.10** at concentrations of 0 μM (blue), 125 μM (green), 250 μM (orange), 500 μM (cyan), 750 μM (purple) and 1000 μM (red). e) Dose-response plot of CSP (ppm) vs ligand concentration (μM) for selected residues in the HSQC titration experiments. Affinity of **3.10** was calculated by fitting a single site binding model with ligand depletion displayed as $K_D \pm$ error of fit. f) Raw SPR sensorgrams of phenylacetic acid **3.10** binding to *EcDsbA* displayed at concentrations of 0.8 μM (red), 1.56 μM (green), 3.12 μM (blue), 6.25 μM (orange), 12.5 μM (purple), 25 μM (brown), 50 μM (dark orange), 100 μM (tan), 200 μM (grey) and 400 μM (blue). g) Dose-response plot of steady-state SPR response (RU) for phenylacetic acid **3.10**. Affinity is calculated by fitting to a single site binding model displayed as $K_D \pm$ error of fit.

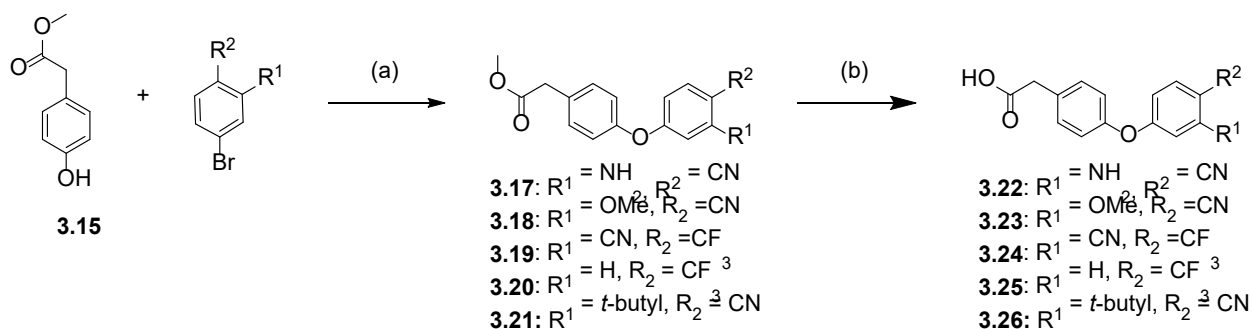
3.2.4 Summary

Replacement of the *para*-hydroxyl with a series of alcohols, carboxylic acids and amides saw minor increases in affinity. H-bond interactions with His32 were observed in the X-ray structures of several analogues and the solubility of analogues was notably improved. The propyl alcohol **3.6** was the most potent analogue from this series, however due to a lack of structural data and poor solubility it was not further pursued. The most interesting analogues in the series was phenylacetic acid **3.10** and the corresponding sulfonic acid **3.7** as both analogues maintained affinity ($K_D = 490 \pm 20 \mu\text{M}$ and $K_D = 530 \pm 30 \mu\text{M}$, respectively) improved solubility ($1000 \mu\text{M}$) and formed H-bonds with both His32 and Gln35 as hypothesised.

3.3 Further optimisation of the right-hand side

3.3.1 Design and synthesis of carboxylic acid analogues

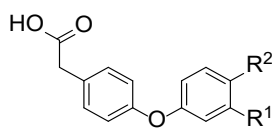
The initial series of compounds **2.1**, **2.2**, **2.14** – **2.19** investigating substitution of the benzonitrile ring (R^1 and R^2 series in Chapter 2) produced a number of insoluble analogues. Consequently, the estimated affinities of these analogues were often solubility limited. Given the improved solubility observed when incorporating a *para*-phenyl acetic acid as in **3.10**, we sought to synthesise a select set of substituents on the benzonitrile ring. These included methyl, trifluoromethyl, amino and methoxy substituents as matched pairs to the initial series of phenols. Substitution for a bulkier *tert*-butyl functional group was also investigated to assess the flexibility of the bottom loop of EcDsbA. We anticipated the incorporation of the carboxylic acid substituent would improve solubility and allow us to better understand the SAR around substitution of the benzonitrile ring. These carboxylic acid analogues **3.22** – **3.26** were synthesised via the previously optimised Ullmann coupling and ester hydrolysis method (Scheme 3.2).

Scheme 3.2: Synthesis of phenylacetic acid diaryl ether analogues.

Reagents and conditions: (a) *N,N*-dimethylglycine hydrochloride, Cs_2CO_3 , CuI, DMF, 100 °C, 24 h, 16 - 50 %. (b) NaOH, THF, H_2O , 25 - 50 °C, 24 - 72 h, 9 - 65 %.

3.3.3 Testing diaryl ether carboxylic acid analogues

These analogues were screened via the solubility, HSQC and X-ray crystallography pipeline previously described. All carboxylic acid containing analogues in this series (**3.22** – **3.26**) demonstrated comparable solubility ($\geq 1000 \mu\text{M}$) to the cyano-methyl derivative **3.10** enabling analysis of their affinities (**Table 3.2**). Substitution of the methyl for an amino (as in **3.22**), replacement of the *para*-nitrile and *meta*-methyl groups with *para*- CF_3 and *meta*-nitrile (as in **3.24**) or removal of the methyl and replacement of the nitrile with a CF_3 (as in **3.25**) all displayed similar binding affinities to the parent cyano-methyl derivative **3.10**. Substitution of the methyl group for a bulkier and more lipophilic *tert*-butyl functional group (as in **3.26**) improved affinity ≈ 2 -fold ($K_D = 265 \pm 10 \mu\text{M}$) while maintaining high aqueous solubility. However, this analogue was not further pursued due to a lack of structural data and the increased lipophilicity (cLogP = 2.84 for **3.10** vs cLogP = 4.63 for **3.26**, as calculated in vortex v5.2.1500) of the *tert*-butyl compound which were considered too great a penalty to pay for the 2-fold increase in affinity.

Table 3.2: Testing results for linker replacement and RHS optimisation.

Compound number	R ¹	R ²	Solubility max (μ M)	Single point HSQC NMR	HSQC NMR Affinity (K_D , μ M)*	Electron density for the ligand	EDIA score
3.10	CH ₃	CN	1000	Binding	490 \pm 20**	Full	0.86
3.22	NH ₂	CN	1000	Binding	578 \pm 75	-	-
3.23	OCH ₃	CN	1000	Binding	900 \pm 140	Full	0.96
3.24	CN	CF ₃	1000	Binding	377 \pm 40	Partial	0.89
3.25	H	CF ₃	1000	Binding	394 \pm 45	Poor	0.77
3.26	<i>t</i> -butyl	CN	1000	Binding	265 \pm 10	-	-

*Affinities presented as mean \pm SEM **Affinity determined in phosphate buffer.

Substitution of the methyl group for a methoxy functional group (as in **3.23**) reduced affinity \approx 2-fold (K_D = 900 \pm 140 μ M) compared to the parent carboxylic acid **3.10**. An X-ray crystal structure of methoxy analogue **3.23** was acquired with full ligand density and it was observed to adopt a similar binding mode to the methyl analogue **3.10**, with the methoxy functional group filling the hydrophobic groove (**Figure 3.4**). This analogue displayed a H-bond interaction to His32 (2.3 Å) but the H-bond to Gln35 (3.6 Å) and π - π stacking with Phe36 at the top and back of the were lost which is a possible cause of the 2-fold drop in affinity.

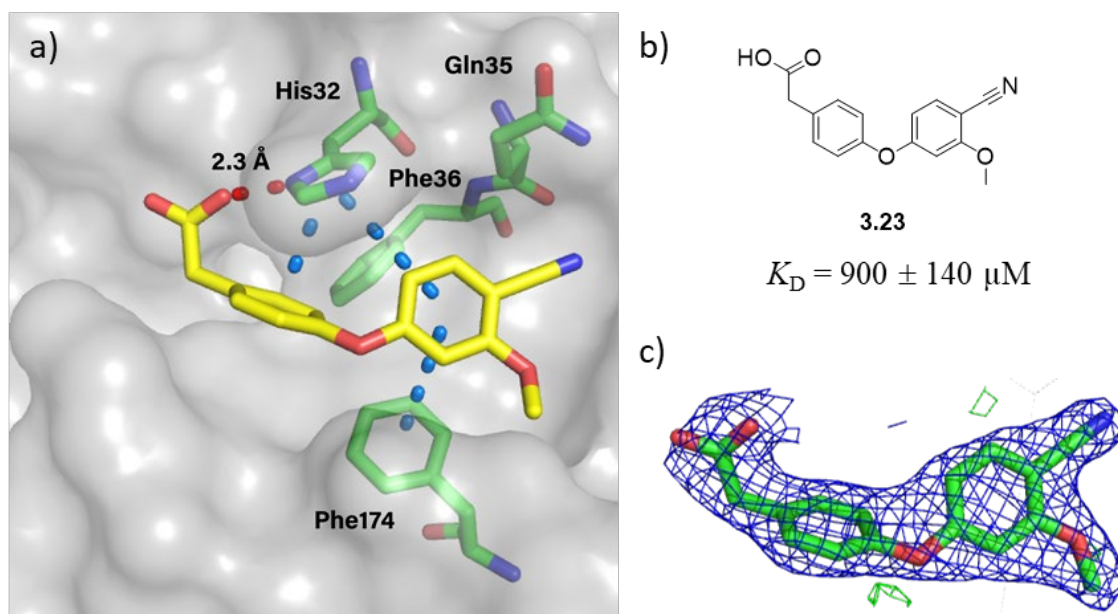


Figure 3.4: a) X-ray crystal structure of methoxy derivative **3.23** (yellow sticks) bound to oxidised *EcDsbA* (grey surface) with key residues His32, Gln35, Phe35 and Phe174 (green sticks). H-bond interactions are shown as red dashed lines and π - π stacking interactions are shown by blue dashed lines. b) Structure of methoxy phenylacetic acid analogue **3.23**. c) Electron density map of **3.23** shown as 2mFo-DFc at 1 σ (blue) and mFo-DFc at 3 σ (positive green and negative red).

3.3.4 Summary

A number of benzonitrile and trifluoromethyl analogues containing the more polar *para*-phenylacetic acid motif were synthesised and tested. All phenylacetic acid analogues showed improved solubility (≥ 1 mM) compared to their phenol counterparts. Whilst, substitution of the 2-methyl substituent for a *tert*-butyl group (as in **3.26**) demonstrated a ≈ 2 -fold increase in affinity, however there was a lack of structural data and a poor LogP trade off associated with this modification. The lack of improved affinity with other substituents led us to retain the methyl group in this position in future analogue design.

3.4 Isosteric replacement of the ether linker and expansion into the bottom groove

3.4.1 Introduction and design

Isosteric replacement of the ether linker was investigated to alter the electronics of the diaryl ring system, which we hypothesised could potentially strengthen the π - π stacking interactions with His32, Phe36 and Phe174. Additionally, previous work had serendipitously revealed that triphenylphosphine oxide (**3.27**), a common by-product from palladium coupling reactions (e.g. Suzuki, Sonagashira) that is notoriously difficult to remove¹, binds to *Ec*DsbA and occupies a novel pocket at the bottom of the hydrophobic groove between Gln164 and Phe174 (**Figure 3.5, a**). The X-ray crystal structure of **3.27** suggested how a vector may be accessed on the RHS of the groove below where the ether linkage currently binds (**Figure 3.5, b**). This will be referred to as the bottom vector. Triphenylphosphine oxide forms a π - π stacking interaction with Phe174 and its phenyl ring sits adjacent to the sidechain amide of Gln164 (3.7 – 4.4 Å away) forming possible weak electrostatic interactions with the NH or C=O of the amide (**Figure 3.5, a**).

In order to access this vector from the current diaryl ethers (**Figure 3.5, b**) replacement of the ether linker was required as extension off the ether oxygen was not possible. Hence, isosteric replacement of the ether linker for other 1 or 2 heavy atom linkers that can be further elaborated (e.g. amino, acyl, amide, benzylamine and benzyl alcohol) was investigated.

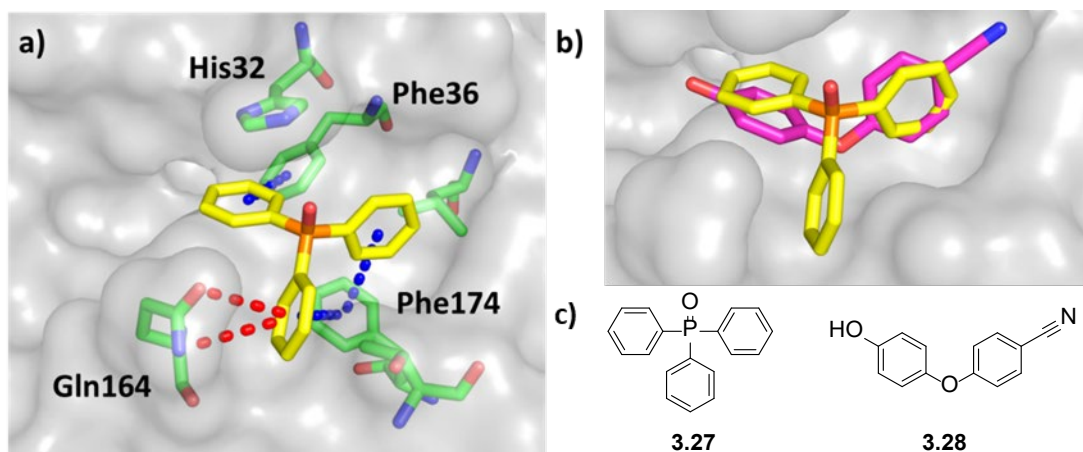
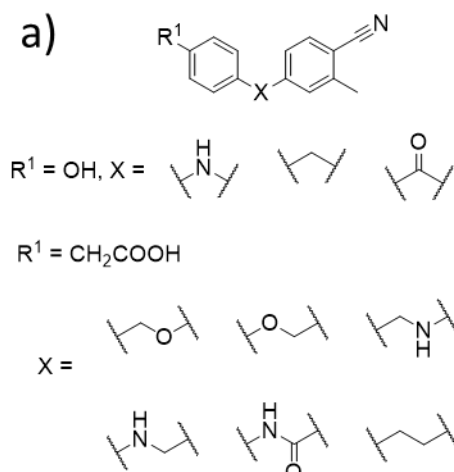


Figure 3.5: a) X-ray crystal structure of triphenylphosphine oxide **3.27** (yellow sticks) bound to oxidised *EcDsbA* (grey surface) showing a phenyl ring occupying the bottom of the hydrophobic groove. Key residues are highlighted in green sticks, π - π stacking interactions are shown as blue dashed lines while potential polar contacts in red. b) X-ray crystal structure of parent diaryl ether **3.28** (pink sticks) overlaid with triphenylphosphine oxide **3.27** (yellow sticks) displaying the potential vector for expansion off the diaryl ethers. c) Structures of triphenylphosphine oxide **3.27** and parent diaryl ether **3.28**.

Glide docking of 1 - 2 heavy atom linkers containing a phenol or acid functional groups on the LHS ring revealed that these molecules maintained binding mode (**Figure 3.6, a**). Expansion off the 1 – 2 heavy atom linkers with aryl, alkyl and charged functional groups was next investigated. This introduced a chiral centre into the molecules and both enantiomers were enumerated and docked using Glide. The docked structures projected their expansions along the bottom vector (**Figure 3.6, b**), making potential interactions with this region of the groove. For two heavy atom linkers, expansion proximal to the phenylacetic acid ring was favoured compared to expansion from the linker atom proximal to the benzonitrile ring. A number of two atom linkers that were synthetically accessible to this type of expansion (CH_2O , CH_2NH , CONH , NHCH_2 , NHCO) were then designed and synthesised. The commercial availability of 2-methyl benzonitrile starting material for the chemistry described below was limited hence the 2-methyl group was not incorporated in these two-analogue series.

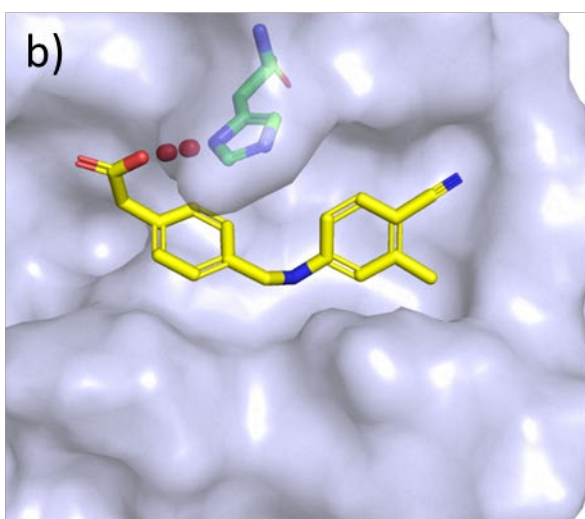
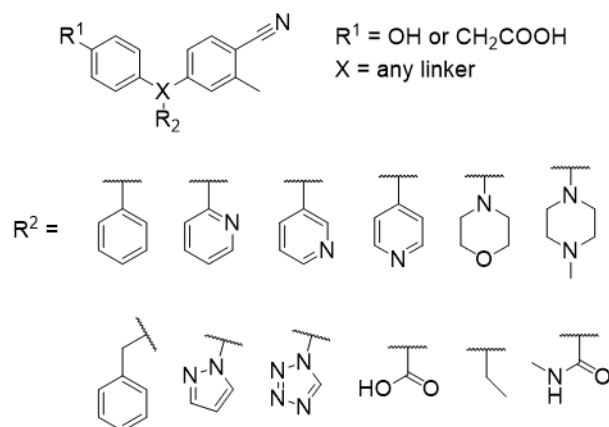
Docking linkers

a)

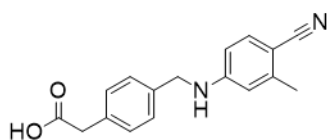


Docking linkers + expansion

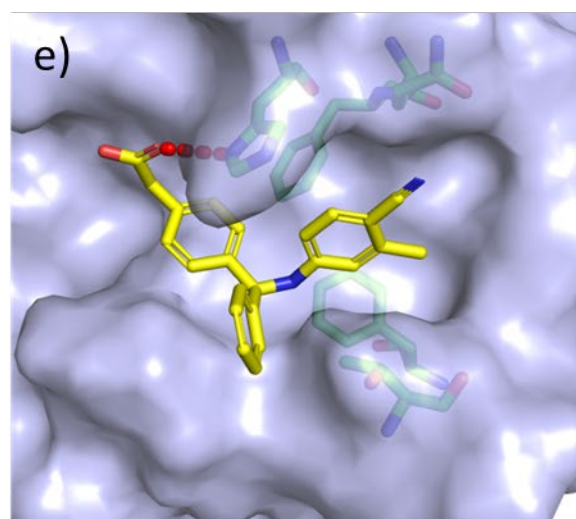
d)



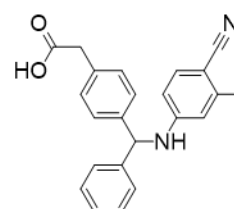
c)



3.29



f)



3.30

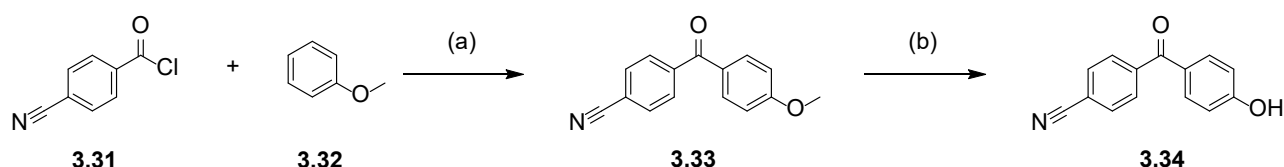
Figure 3.6: Docking of 1 and 2 heavy atom linkers and their expansions against *EcDsbA* using Glide.

a) list of linkers replacing the ether that were designed and docked against *EcDsbA*. b) Docked structure of the amino methyl linked **3.29** against *EcDsbA*. The binding orientation and H-bond interaction to His32 (red dashes) are maintained. c) 2D structure of amino methyl linked **3.29**. d) List of expanded analogues probing potential interactions with the bottom groove region of *EcDsbA*. e) Docked structure of the expanded amino methyl analogue **3.30** showing its phenyl expansion projecting into the bottom groove. H-bond to His32 was maintained as shown by red-dashes. f) 2D structure of expanded amino methyl analogue **3.30**.

3.4.2 Synthesis and testing of single heavy atom linkers

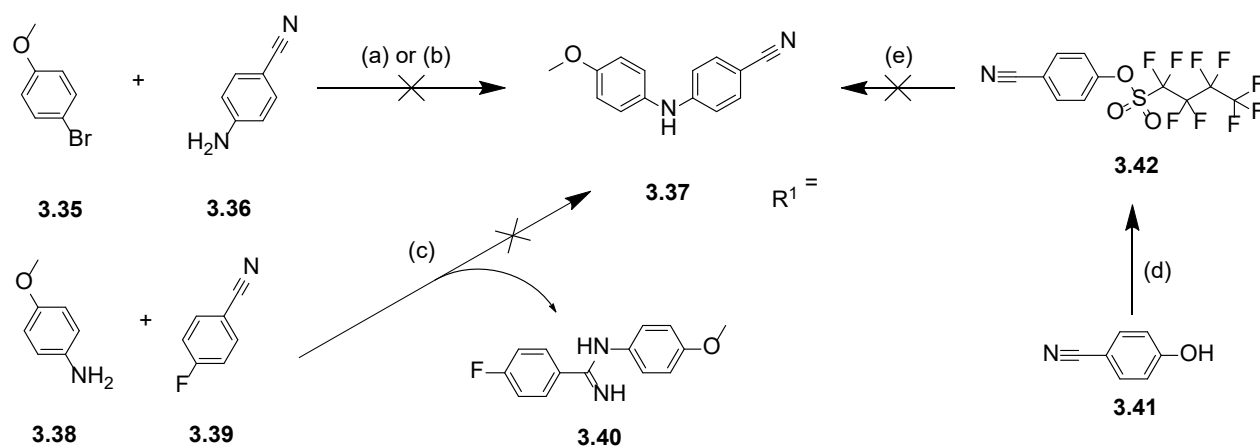
Initially we targeted the synthesis of the one heavy atom amino and acyl linkers. Friedel-Crafts acylation of anisole (**3.32**) with 4-cyano benzoyl chloride (**3.31**) (**Scheme 3.3**) gave the acyl linked benzophenone **3.33** with 99 % purity by HPLC purity but < 80 % purity by ^1H NMR with anisole **3.32** a major impurity. Deprotection of the methyl ether was achieved using BBr_3 to give the desired acyl linked phenol **3.34**. ^1H NMR spectrum of benzophenone **3.34** was in accordance with the literature² showing four doublet resonances at 7.65 – 7.90 ppm consistent with the benzophenone structure.

Scheme 3.3: Synthesis of single heavy atom linker.



Reagents and conditions: (a) AlCl_3 , DCM, RT, 16 h, 19 %. (b) BBr_3 1 M in heptane, DCM, RT, 16 h, 35 %.

In order to access the amino linked analogue **3.37** a screen of Ullmann type and Buchwald-Hartwig coupling reactions was undertaken (**Scheme 3.4**), by varying the solvent (DMF, DMSO), metal ligand (*N,N*-dimethylglycine hydrochloride, phenanthroline, *L*-proline, (2-biphenyl)di-*tert*-butylphosphine) and catalytic metal (CuI , $\text{Pd}_2(\text{dba})_3$) (**Scheme 3.4**). These conditions proved unsuccessful and gave low conversion as estimated from the relative intensities of the UV peaks of reactant **3.35** and desired product **3.37** observed in the LCMS data and the desired product **3.37** could not be purified and isolated. Nucleophilic aromatic substitution of 4-fluorobenzonitrile (**3.39**) with *para*-anisdine (**3.38**) and KO^tBu base gave amidine **3.40** as the major product via attack of the nitrile. The formation of this by-product was supported by a peak in ^{19}F NMR spectrum corresponding to the unreacted fluorine, a corresponding LCMS mass $m/z = 245.0$ $[\text{M}+\text{H}]^+$ and fluorine splitting in the ^1H NMR spectrum of the isolated compound consistent with the amidine structure.

Scheme 3.4: Synthesis of 1 heavy atom linkers.

Reagents and conditions: (a) CuI, Cs₂CO₃, *N,N*-dimethylglycine hydrochloride or phenanthroline or L-proline, DMF or DMSO or 1,4-dioxane, 120 °C, 18 h - 3 d, 0 %. (b) Pd₂(dba)₃, (2-Biphenyl)di-*tert*-butylphosphine, KO*t*Bu, toluene, 120 °C, 18 h, 0 %. (c) KO*t*Bu, DMSO, RT, 25 h, 0 %. (d) perfluoro-1-butanefluoride, DMAP, DIPEA, 0 – 25 °C, 18 h, 78 %. (e) Pd₂DBA₃, 4,5-bis(diphenylphosphino)-9,9-dimethylxanthene, toluene, 80 °C, 16 h, 0 %.

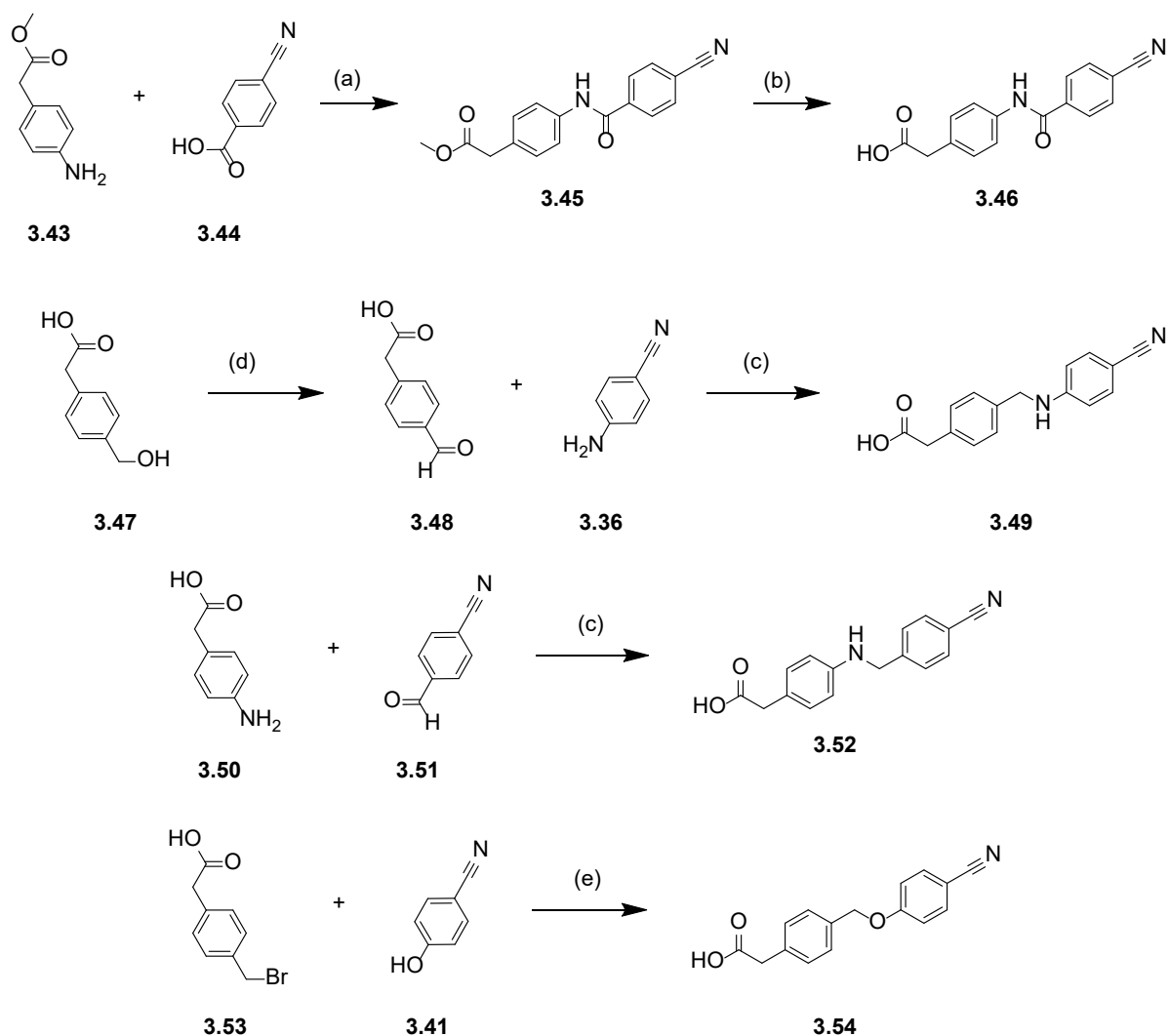
Following a literature procedure³ nonaflate **3.42** was synthesised using 4-cyanophenol (**3.41**) and perfluoro-1-butanefluoride with catalytic DMAP and DIPEA and subjected to Buchwald-Hartwig coupling conditions with *para*-anisidine (**3.38**). Some conversion was obtained as observed by LCMS with a new UV vis peak being detected and a corresponding LCMS mass $m/z = 245$ [M+H]⁺ however, this product could not be purified. Due to the problems with this chemistry further optimisation was halted to focus on higher priority synthetic targets.

3.4.3 Synthesis of two heavy atom linkers

Access to the two heavy atom linked analogues was initiated by 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate (HATU) mediated amide coupling of methyl 4-aminophenylacetate (**3.43**) with 4-cyanobenzoic acid (**3.44**) to give the amide linked intermediate **3.45** in moderate yields (**Scheme 3.5**). Hydrolysis of the methyl ester using 2 M aq. NaOH was found to hydrolyse both the biaryl amide and ester functionalities yielding a 1:1 ratio of

the desired carboxylic acid product **3.46** and multiple species of amide cleaved by-products. Lowering the reaction temperature to 0 °C and reducing the number of molar eq. of NaOH did not improve the ratio of desired compound to by-products. Hence, substitution of NaOH for the weaker K₂CO₃ base was attempted and afforded improved selectivity for ester over amide hydrolysis. Low yields were obtained due to co-elution of carboxylic acid **3.46** with 4-cyanobenzoic acid (**3.44**) during column chromatography purification.

Scheme 3.5: Synthesis of a series of two heavy atom linkers.



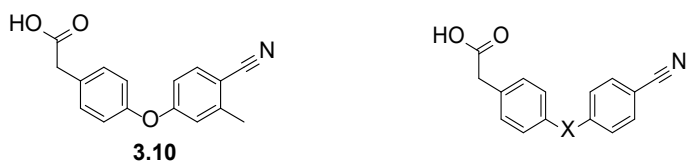
Reagents and conditions: (a) HATU, DIPEA, DMF, RT, 18 h, 27 %. (b) K₂CO₃, H₂O, MeOH, RT, 72 h, 2.2 %. (c) STAB, AcOH, 1,2-DCE, RT, 18 h, 10 - 29 %. (d) Dess-Martin periodinane, DCM, RT, 18 h, 49 %. (e) K₂CO₃, KI, MeCN, RT, 96 h, 49 %.

Dess-Martin oxidation of 2-(4-(hydroxymethyl)phenyl)acetic acid (**3.47**) was conducted following a literature method⁴ to give benzaldehyde **3.48**. Benzaldehyde **3.48** and the purchased 4-cyanobenzaldehyde (**3.51**) were both then coupled via reductive amination using Na(AcO)₃BH with their corresponding anilines **3.36** and **3.50** to give the desired amino linked analogues **3.49** and **3.52**, respectively. Finally, two heavy atom ether linked **3.54** was synthesised by a Williamson ether reaction of 4-cyanophenol (**3.41**) with alkyl bromide **3.53** in moderate yields.

3.4.4 Testing one and two heavy atom linker series

Replacement of the ether linker with a ketone (as in benzophenone **3.34**) gave poor solubility (250 μ M) and little to no binding by ¹H-¹⁵N HSQC NMR. Substitution of the ether linker with various two heavy atom linkers on the acetic acid substituted series maintained good aqueous solubility with all analogues passing at 1 mM (Table 3.3). Amide linked **3.46** and amino methyl linked **3.52** showed only small CSP (<0.04 ppm) in single point ¹H-¹⁵N HSQC NMR experiments consistent with very weak binding and an affinity was not determined. The methyl amino linked **3.49** and methyl ether linked **3.54** demonstrated 3-fold loss and similar affinity compared to the parent diaryl ether **3.10**, respectively. While not showing increased affinity the amino linked **3.49** and ether linked **3.54** maintained solubility and introduced new synthetic handles which allow for further expansion into the bottom groove.

Table 3.3: Testing two heavy atom linkers which allow expansion into the bottom groove. Affinities of analogues were determined in phosphate buffer.



Compound number	X	Solubility max (μM)	Single point HSQC NMR	HSQC NMR Affinity (K_D , μM)*	Electron density for the ligand
3.10	See above	1000	Binding	490 ± 20	Partial
3.46		1000	< 0.04 CSP	-	-
3.49		1000	Binding	1500 ± 300	-
3.52		1000	< 0.04 CSP	-	-
3.54		1000	Binding	700 ± 80	-

*Affinities presented as mean \pm SEM.

3.5 Expansion into the bottom groove using the Petasis reaction.

3.5.1 Design and docking

The previous docking studies suggested that expansion off the linker proximal to the phenylacetic acid ring was preferred over expansions proximal to the benzonitrile ring. We therefore sought to quickly explore the potential for expansion at this vector in the two heavy atom linked series. The Petasis reaction is a multi-component coupling reaction of a boronic acid, amine and an aldehyde that proceeds under facile reaction conditions (**Figure 3.7**).⁵ It was hypothesised that this chemistry could

afford the amino acid core (**Figure 3.8**) as well as installing a carboxylic acid handle for further elaboration to provide access into the bottom groove of *EcDsbA*.

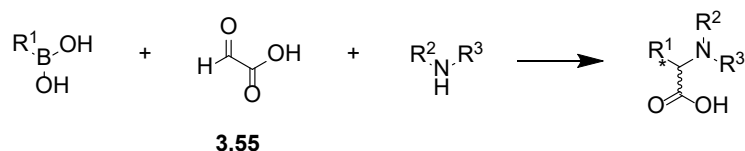


Figure 3.7: General reaction scheme for the Petasis reaction.

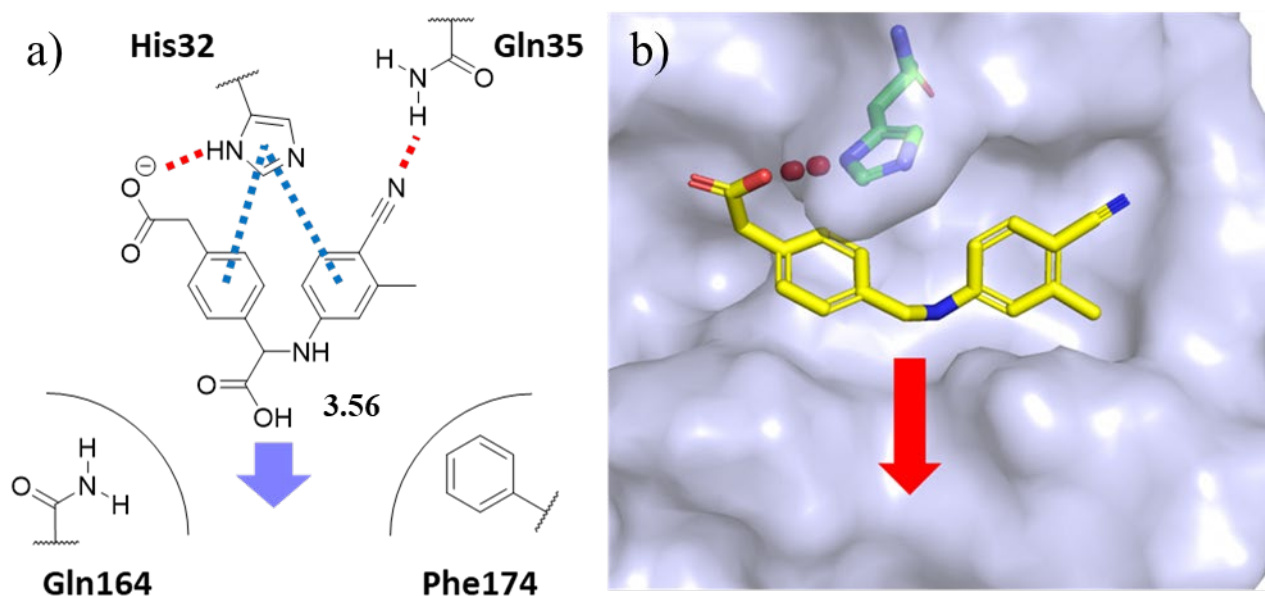


Figure 3.8: a) Expansion strategy for accessing the bottom groove. Suggested interactions formed by amino linked carboxylic acid compound **3.56**. Purple arrow indicating region for expansion off the carboxylic acid of **3.56**. b) Docked structure of amino linked diaryl analogue **3.29** against *EcDsbA* with red arrow depicting bottom groove.

The bottom groove of *EcDsbA* contained a number of potential sites to form polar interactions, including backbone carbonyls and side chain alcohols. Enumeration of amides formed by virtual enumeration as previously described with the carboxylic acid of **3.56**, allowed for docking of a variety of aryl polar and non-polar functional groups (**Figure 3.9**). The aim was to both mimic the PPh_3O crystal structure and make potential new polar interactions in the bottom groove region of the pocket.

The docked structures often adopted the binding mode proposed in **Figure 3.8**. Thus, a subset of analogues were prioritised for synthesis to investigate this vector. To simplify the synthesis of analogues and quickly evaluate the vector for developability a *para*-methoxy boronic acid was utilised in the Petasis coupling instead of phenylacetic acid to simplify the subsequent amide coupling reactions.

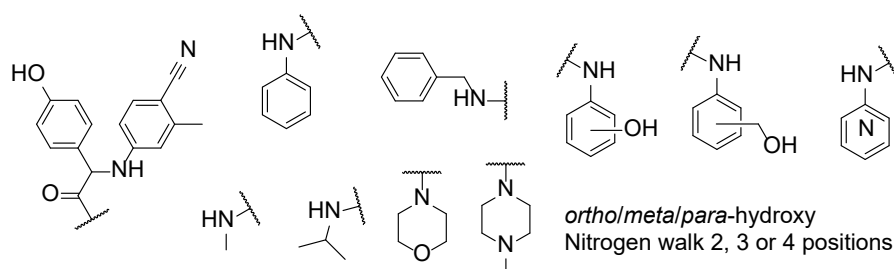
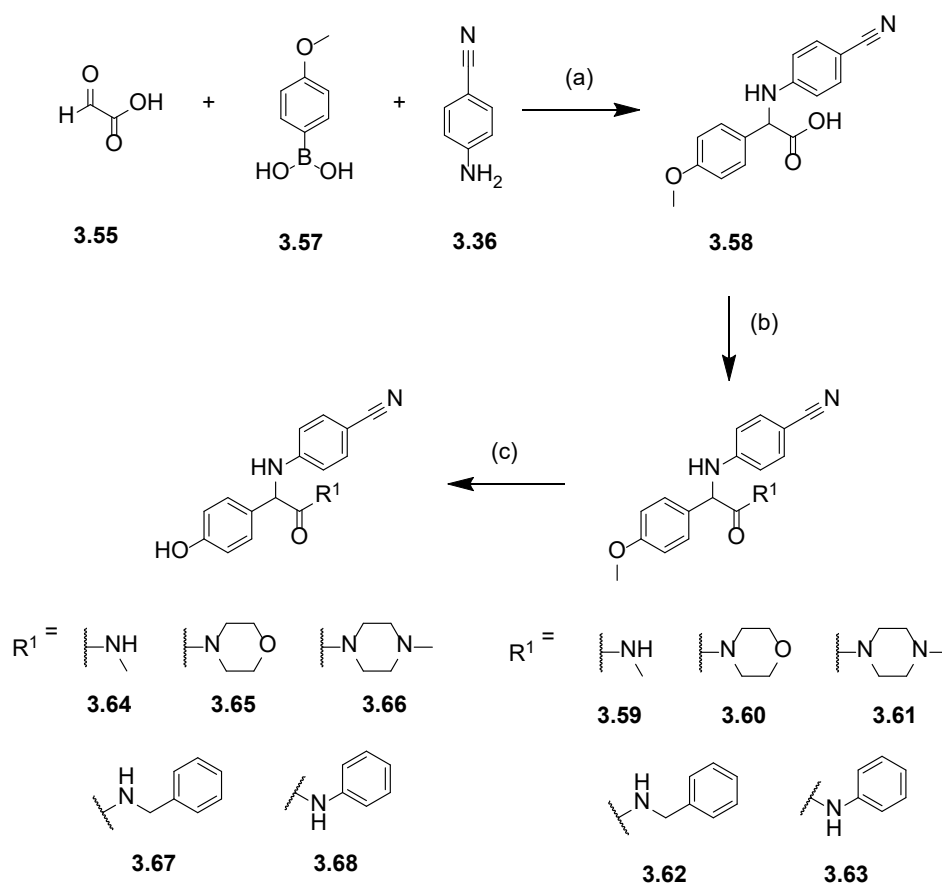


Figure 3.9: Analogues based on the amino linked core that were docked against *EcDsbA* to explore the bottom vector.

3.5.2 Synthesis of bottom groove expanded analogues.

The Petasis multi-component coupling of glyoxylic acid (**3.55**), 4-methoxyphenyl boronic acid (**3.57**) and 4-cyanoaniline (**3.36**) proceeded based on literature conditions⁶ to give the amino acid **3.58** in excellent yields (**Scheme 3.6**). This transformation was supported by the appearance of the chiral benzylic proton resonance at 5.08 ppm in the ¹H NMR spectrum. Amidation of the carboxylic acid of **3.58** with methyl amine, morpholine, *N*-methylpiperazine, benzylamine and aniline using HATU and Et₃N gave amides **3.59-3.63**, respectively. Deprotection of methyl ether in the morpholine **3.60**, benzylamine **3.62** and aniline **3.63** analogues proceeded under the previously optimised conditions with BBr₃ to give phenols **3.65**, **3.67** and **3.68**, respectively. Demethylation of the methyl ether in *N*-methyl amide **3.59** and *N*-methylpiperazine **3.61** to phenols **3.64** and **3.66** was unsuccessful due to the formation of by-products and difficulty upon purification. Hence, the phenol and methyl ether intermediates were tested concurrently with other compounds from this series as reported below (**Table 3.4**).

Scheme 3.6: Synthesis of bottom groove expanded analogues.

Reagents and conditions: (a) MeCN, reflux, 1 h, 92 %. (b) HATU, Et₃N, DMF, RT, 72 h, 12 - 72 %.

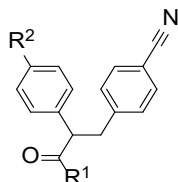
(c) BBr₃ 1 M in heptane, DCM, 0 - 25 °C, 4 - 24 h, 23 - 47 %.

3.5.3 Testing bottom groove expanded analogues

The carboxylic acid **3.58**, methyl amine **3.59**, *N*-methylpiperazine **3.61** and morpholine **3.65** analogues displayed high solubilities ($\geq 1000 \mu\text{M}$) but only small CSP which were not sufficient for affinity determination (**Table 3.4**). Methoxy intermediates morpholine **3.50** benzylamine **3.62** and aniline **3.63** had poor solubility and were unable to be screened via ¹H-¹⁵N HSQC NMR. The two aryl substituted phenol analogues, phenyl **3.68** and benzyl **3.68**, more closely resemble triphenylphosphine oxide with 3 phenyl rings and show binding to *Ec*DsbA in ¹H-¹⁵N HSQC NMR, perturbing multiple residues > 0.04 ppm. However, similar to PPh₃O, they showed low aqueous solubility and their affinities could not be determined accurately. Incorporation of a phenyl or benzyl

functional group without a significant improvement in affinity is unfavourable from a drug design perspective (increasing LogP and reducing aqueous solubility etc.).

Table 3.4: Testing, solubility and X-ray crystallography results for the Petasis derivatives.



Compound number	R ¹	R ²	Solubility max (μM)	Single point HSQC NMR	Electron density for the ligand	EDIA score
3.58	OH	OCH ₃	1000	< 0.04 CSP	-	-
3.59		OCH ₃	1000	< 0.04 CSP	-	-
3.60		OCH ₃	< 125	-	-	-
3.61		OCH ₃	1000	< 0.04 CSP	-	-
3.62		OCH ₃	< 125	-	-	-
3.63		OCH ₃	< 125	-	-	-
3.65		OH	1000	<0.04 CSP	-	-
3.67		OH	125	Binding	-	-
3.68		OH	125	Binding	Partial	0.79

*Affinities presented as mean ± SEM.

An X-ray crystal structure of **3.68** with partial electron density for the ligand was obtained with phenyl amide **3.68** bound to *EcDsbA* (**Figure 3.10, a and b**). This structure positioned **3.67** in an orientation where it was solved to be making contacts with the second chain of *EcDsbA* in the asymmetric unit

of the crystal. The weak binding obtained for these analogues, and potentially altered binding mode for **3.68** did not support the ongoing exploration of the bottom groove of *EcDsbA*. Therefore, no further attempts were made to generate the phenol analogues **3.64** and **3.66**.

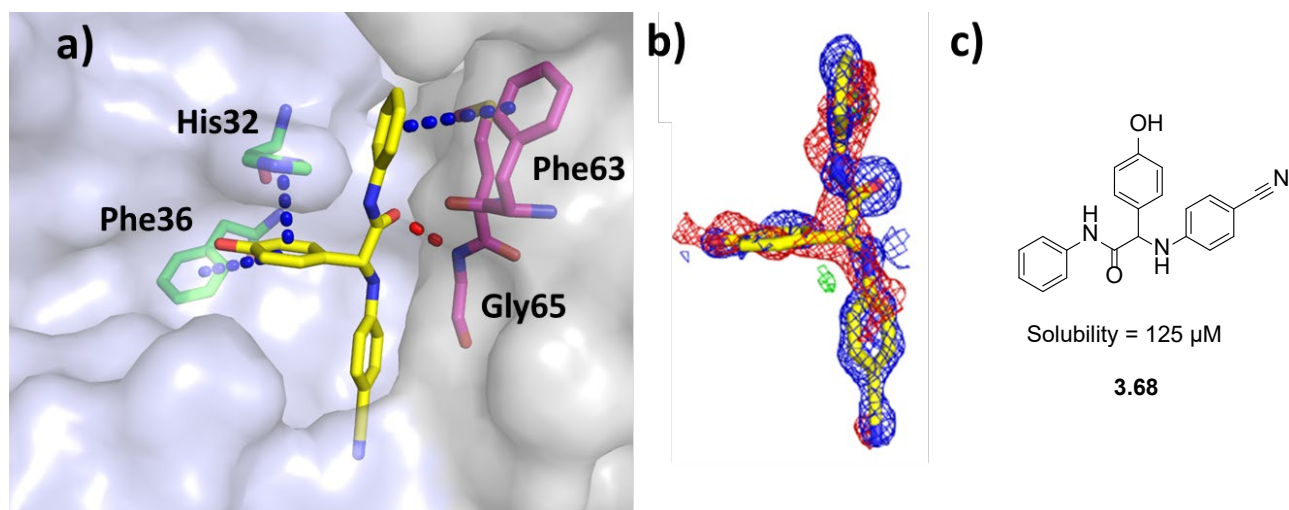


Figure 3.10: a) X-ray crystal structure of phenyl amide **3.68** (yellow sticks) bound to oxidised *EcDsbA* with both chains in the asymmetric unit displayed (chain A blue tint and chain B in grey). Key residues are shown for both chains in green and purple sticks, respectively. π - π stacking interactions are shown as blue dashed lines and hydrogen bonds as red dashed lines. b) Electron density map of phenyl analogue **3.68** shown as 2mFo-DFc at 1 σ (blue) and mFo-DFc at 3 σ (positive green and negative red). c) 2D structure of phenyl amide **3.68**.

3.6 Conclusions

With the aid of crystal structures obtained the diaryl ether core was further explored through merging of the best analogues and expanded upon via Petasis chemistry. Merging of the *para*-phenylacetic acid functional group with the 2-methyl benzonitrile gave a series of analogues with comparable affinity to the phenol **3.1** but higher aqueous solubility (max 1000 μ M). Structural data for 2-methyl phenylacetic acid derivative **3.10** and related isostere sulfonic acid **3.7** made a H-bond interaction to the sidechain of His32 through the carboxylate and sulfonate, respectively, while also maintaining the H-bond from the cyano to the sidechain of Gln35. The SAR of the methyl substituent of the

benzonitrile ring was further investigated with replacement to a *tert*-butyl (analogue **3.26**) demonstrating \approx 2-fold increase in affinity. However, this small increase in affinity is outweighed by the large increase in cLogP associated with this modification, hence the 2-methyl substituent phenylacetic acid remained preferred for further elaboration. Isosteric replacement of the ether linker with one and two heavy atom linkers was investigated and demonstrated methyl ether **3.54** and methyl amino **3.49** linkers retained binding to *EcDsbA*. The methyl amino linker was then further investigated by a series of Petasis products exploring the bottom of the hydrophobic groove. From a preliminary series of analogues only the least soluble and most lipophilic phenyl amide **3.68** and benzyl amide **3.67** derivatives showed binding to *EcDsbA*. Similar to the results obtained in Chapter 2, it was observed that designing analogues that display improved binding to *EcDsbA* while maintaining aqueous solubility was challenging.

Identification of the merged 2-methyl phenylacetic acid analogue **3.10** gave an improvement in solubility while making an additional H-bond with full electron density for the ligand by X-ray crystallography. Having synthesised and tested >60 analogues that explored the RHS of the groove, we sought to further expand the diaryl ethers across into the left hand side (LHS) of the groove. This phenylacetic acid analogue **3.10** was chosen as the core from which to explore this expansion.

3.7 References

1. Batesky, D. C.; Goldfogel, M. J.; Weix, D. J. Removal of Triphenylphosphine Oxide by Precipitation with Zinc Chloride in Polar Solvents. *J. Org. Chem.* **2017**, 82, 9931-9936.
2. Wood, P. M.; Woo, L. W. L.; Labrosse, J.-R.; Trusselle, M. N.; Abbate, S.; Longhi, G.; Castiglioni, E.; Lebon, F.; Purohit, A.; Reed, M. J.; Potter, B. V. L. Chiral Aromatase and Dual Aromatase–Steroid Sulfatase Inhibitors from the Letrozole Template: Synthesis, Absolute Configuration, and In Vitro Activity. *J. Med. Chem.* **2008**, 51, 4226-4238.
3. Anderson, K. W.; Mendez-Perez, M.; Priego, J.; Buchwald, S. L. Palladium-Catalyzed Amination of Aryl Nonaflates. *J. Org. Chem.* **2003**, 68, 9563-9573.
4. Jimenez-Moreno, E.; Gomez, A. M.; Bastida, A.; Corzana, F.; Jimenez-Oses, G.; Jimenez-Barbero, J.; Asensio, J. L. Modulating weak interactions for molecular recognition: a dynamic combinatorial analysis for assessing the contribution of electrostatics to the stability of CH- π bonds in water. *Angew. Chem. Int. Ed. Engl.* **2015**, 54, 4344-8.
5. Petasis, N. A.; Akritopoulou, I. The boronic acid mannich reaction: A new method for the synthesis of geometrically pure allylamines. *Tetrahedron Lett.* **1993**, 34, 583-586.
6. Glunz, P. W.; Mueller, L.; Cheney, D. L.; Ladziata, V.; Zou, Y.; Wurtz, N. R.; Wei, A.; Wong, P. C.; Wexler, R. R.; Priestley, E. S. Atropisomer Control in Macrocyclic Factor VIIa Inhibitors. *J. Med. Chem.* **2016**, 59, 4007-4018.

Chapter 4 – Expansion of the diaryl ethers across the hydrophobic groove.

4.1 Introduction

Introduction of a *para*-phenylacetic acid group (as in **4.1**) improved aqueous solubility (1000 μM) and maintained affinity ($K_D = 490 \pm 20 \mu\text{M}$). A systematic approach to generating SAR for this fragment identified few analogues bound with greater affinity than the methyl substituted acetic acid **4.1**, and those which had higher affinity lacked structural data. Thus, methyl substituted phenylacetic acid **4.1** was chosen to further explore the LHS of the groove. The X-ray crystal structure of carboxylic acid **4.1** projects the carboxylate towards the LHS of the groove providing a potential vector for expansion (**Figure 4.1, b**). Amidation off this position with methylamine **4.2** and dimethylamine **4.3** reduced solubility (250 μM and 125 μM respectively).

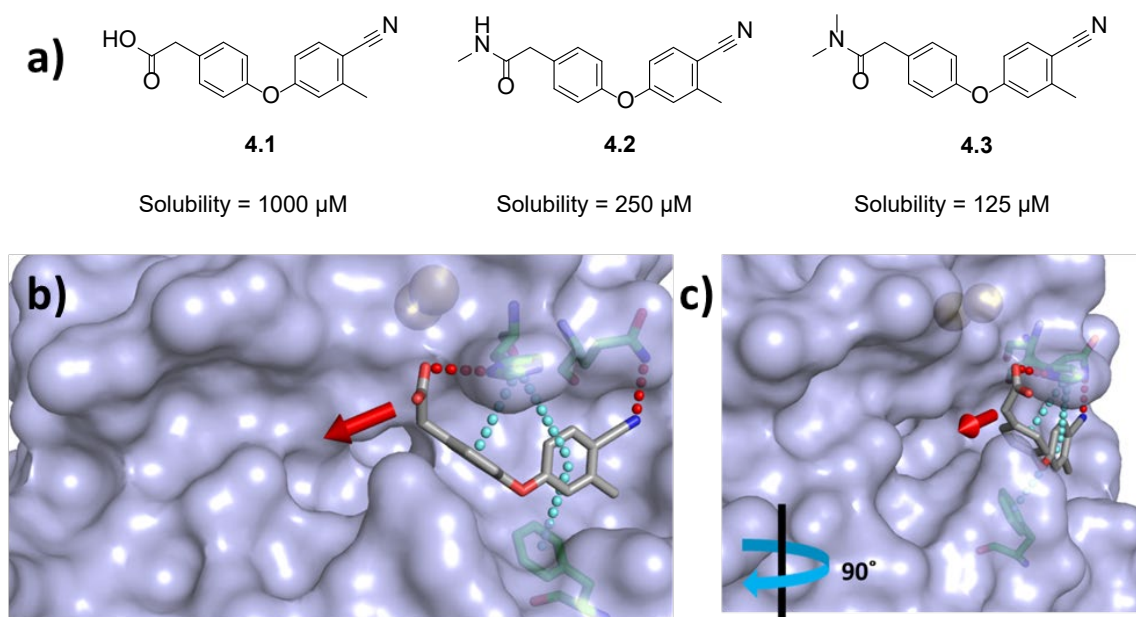


Figure 4.1: a) 2D structures of phenylacetic acid **4.1**, *N*-methyl amide **4.2** and *N,N*-dimethyl amide **4.3**. b) X-ray crystal structure of phenylacetic acid **4.1** (grey sticks) bound to *EcDsbA* (tinted blue surface) showing the proposed vectors for expansion (red arrow). Expansion vector off the carboxylic acid of **4.1** indicated by a red arrow. c) X-ray crystal structure of carboxylic acid **4.1** rotated 90° from the view in (b) showing the expansion vector off the benzylic carbon of **4.1** as indicated by a red arrow.

Hence, we sought to design polar extensions off the carboxylic acid in order to maintain reasonable solubility. We aimed to do this using 3 main chemistries a) amidation of the carboxylic acid with a series of extended carboxylic acid analogues, b) isosteric replacement of the carboxylate with a series of acyl sulfonamides and c) alkylation of the benzylic position (**Figure 4.1, c**).

4.2 Amidation of the carboxylic acid

4.2.1 Design and docking

The LHS of the binding groove in *EcDsbA* contains a number of surface exposed polar amino acid side chains and backbone carbonyls where H-bonding interactions and other polar contacts could be formed. These include Gln146, Arg148, Pro151, Gln160 and Asn162 (**Figure 4.2, a**). The sidechain guanidine of Arg148 is charged at physiological pH and resides ≈ 11 Å from the carboxylate of **4.1**. We hypothesised that forming a strong ionic interaction with the Arg148 positively charged sidechain could improve affinity of the diaryl ethers. Such an ionic interaction required a negatively charged group such as a carboxylate, which would most likely aid in maintaining high solubility. Electron density for the side chain of Arg148 is often absent in X-ray crystal structures of *EcDsbA* due to its flexibility, indicating the interaction may be difficult to obtain. Therefore, we designed and tested a series of carboxylic acids linked by 4 – 8 heavy atoms to the diaryl ether **4.1** (**Figure 4.2, b**).

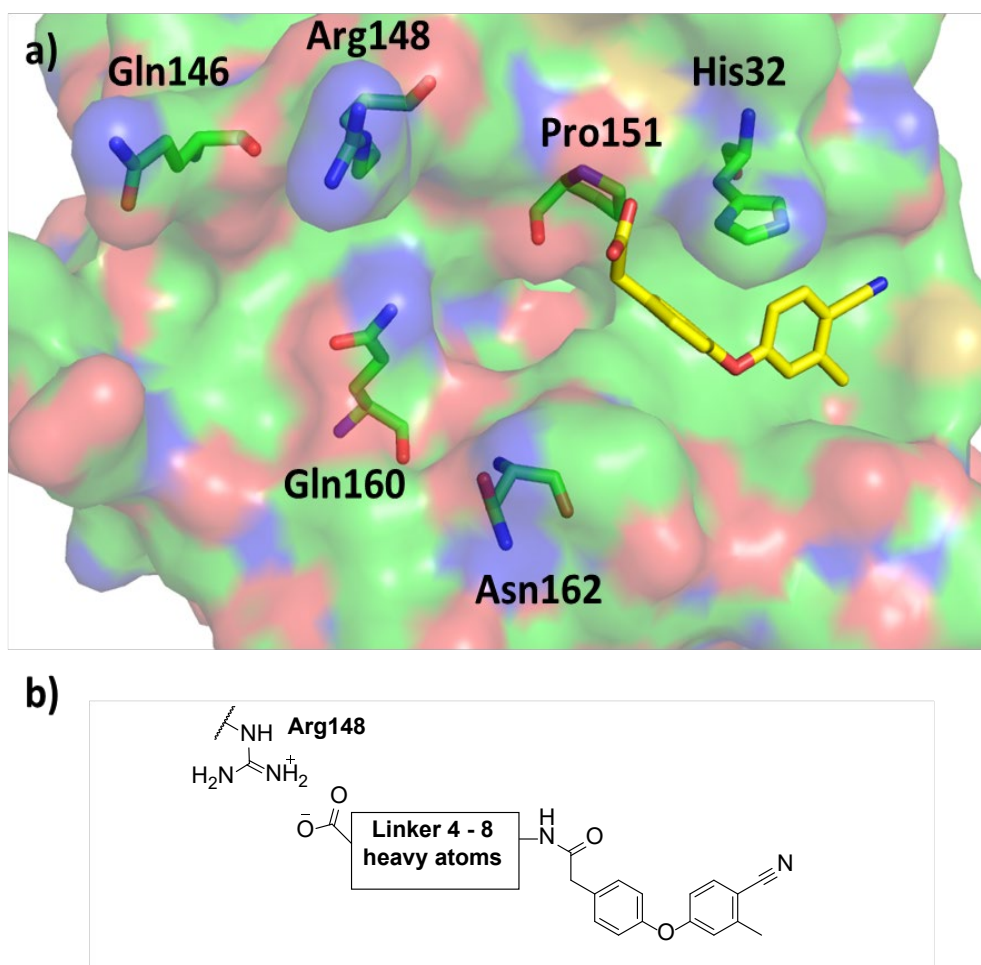


Figure 4.2: a) X-ray crystal structure of carboxylic acid **4.1** (yellow sticks) bound to oxidised *EcDsbA* (green surface). A number of key residues on the LHS of the groove are shown including: Gln146, Arg148, Pro151, Gln160 and Asn162 along with the active site His32 (green sticks). b) General structure of the linked diaryl ether analogues designed to target Arg148.

A list of commercially available reagents were filtered for amino acids of varying lengths and virtually enumerated with 2-methyl carboxylic acid core **4.1** to give a virtual library of amino acid extended analogues. Compounds containing PAINS and other undesirable functional groups were filtered out and the remaining compounds were computationally docked against *EcDsbA* using Glide to assess their potential to form ionic interactions with Arg148. The variable acid substituents projected into the LHS of the *EcDsbA* groove forming ionic and H-bond interactions with Arg148, Gln160 and Asn162 however, these analogues did not maintain the H-bond interaction of the carbonyl to His32. Nevertheless, a series of aromatic and aliphatic rings, straight chain and glycine-based

linkers were purchased and coupled to the carboxylic acid core **4.1** to explore potential interactions (Figure 4.3).

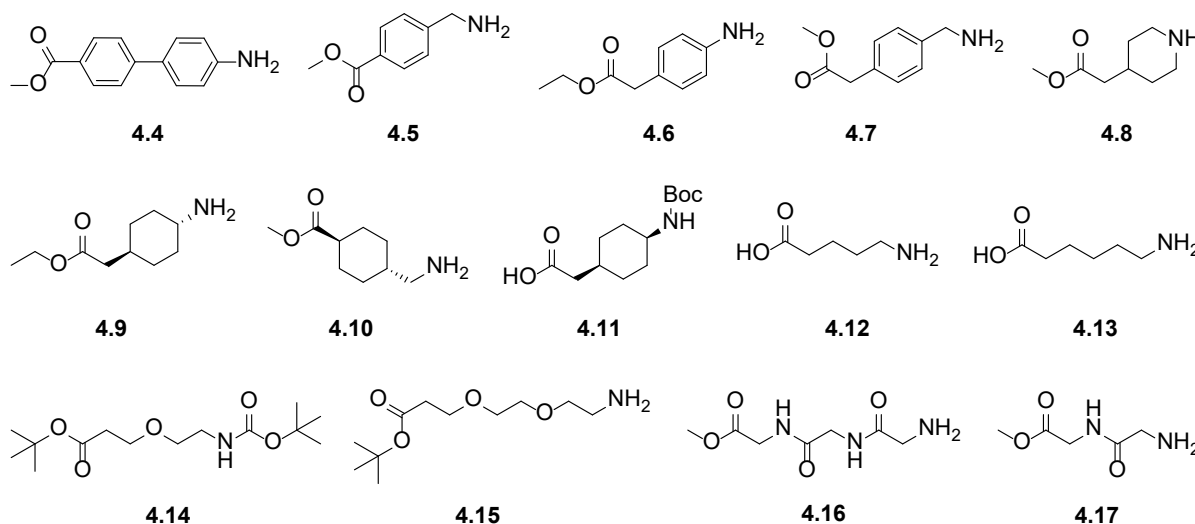
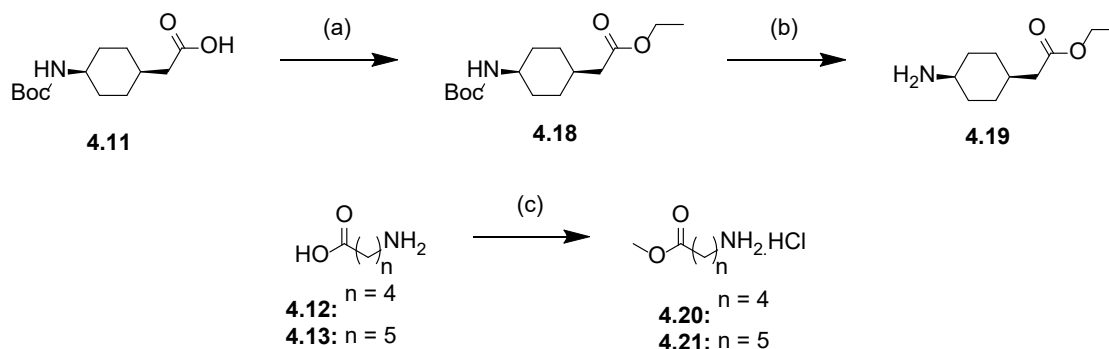


Figure 4.3: Amino acids reagents **4.4** – **4.17** identified from enumeration and docking studies which have the potential to make interactions on the LHS of the groove. These reagents were purchased from commercial vendors with protecting groups as shown.

4.2.2 Synthesis of expanded carboxylic acid analogues

Some linkers required additional protecting group manipulation before coupling to carboxylic acid fragment **4.1** (Scheme 4.1). Consequently, following a literature procedure¹ esterification of Boc-*cis*-1,4-aminocyclohexylacetic acid (**4.11**) with ethyl iodide and K₂CO₃ followed by Boc deprotection with TFA gave ethyl-*cis*-aminocyclohexyl acetate (**4.19**) consistent with literature.¹ Fischer esterification of 5-aminovaleric acid (**4.12**) and 6-aminocaproic acid (**4.13**) with SOCl₂ in MeOH gave amino esters **4.20** and **4.21** as the hydrochloride salts, respectively. ¹H NMR characterisation of these two esters was in accordance with the literature showing a singlet resonance at 3.7 ppm corresponding to the methyl ester in the ¹H NMR spectra.^{2,3}

Scheme 4.1: Synthesis of a series of amino ester precursors for coupling to the diaryl ether core.

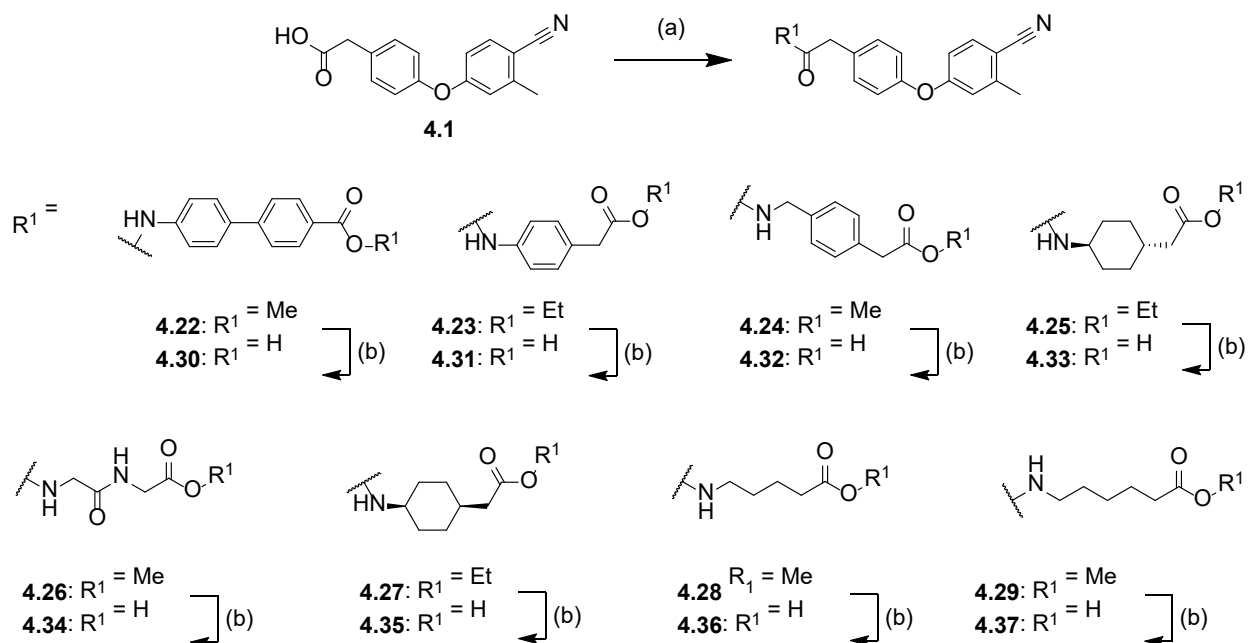


Reagents and conditions: (a) EtI, K₂CO₃, MeCN, 50 °C, 24 h, 96 %. (b) TFA, DCM, RT, 16 h, 47 %.

(c) SOCl₂, MeOH, reflux, 16 h, 80 - 94 %.

Phenylacetic acid **4.1** was synthesised via the previously optimised route of Ullmann coupling and ester hydrolysis (Chapter 3). Synthesis of the linked carboxylic acid analogues was achieved by HATU mediated amidation of carboxylic acid **4.1** with biaryl amine **4.4**, aniline **4.6**, benzylamine **4.7**, *trans*-cyclohexyl amine **4.9**, diglycine **4.17**, *cis*-cyclohexyl amine **4.19** and straight chain amines **4.20** and **4.21** to generate amides **4.22** - **4.29**, respectively (**Scheme 4.2**). Ester hydrolysis of these analogues was achieved using LiOH and MeOH or THF and gave carboxylic acids **4.30** - **4.37**, respectively.

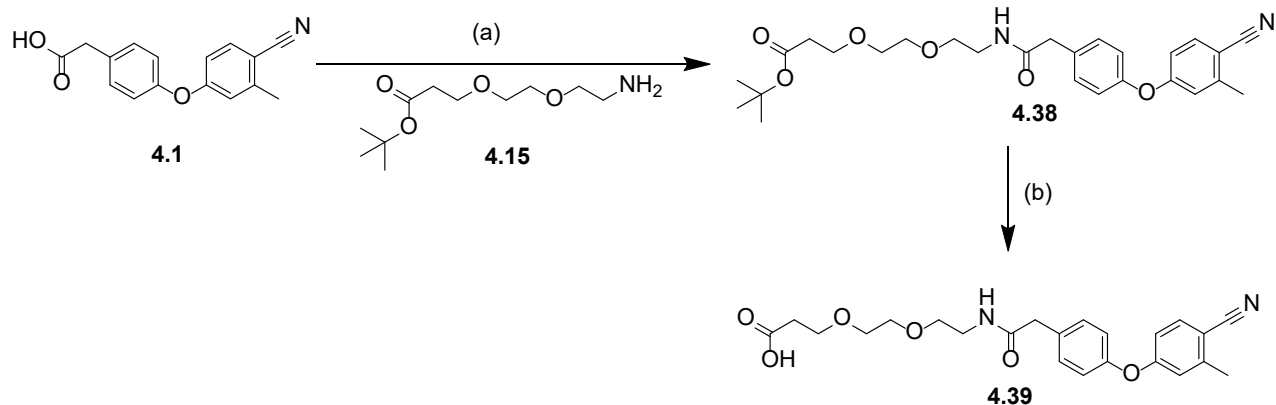
Scheme 4.2: Synthesis of a series of expanded carboxylic acid based diaryl ethers.



Reagents and conditions: (a) amine (R^1), HATU, Et_3N , DMF, RT, 24 h, 23 - 65 %. (b) LiOH, THF, H_2O , RT, 6 - 24 h, 12 - 55 %.

The PEG linked analogue **4.39** was synthesised via amide coupling of carboxylic acid **4.1** with amine **4.15** followed by acid mediated ester hydrolysis using TFA in DCM to give PEG linked carboxylic acid **4.39** (Scheme 4.3). This transformation was supported by ^1H NMR data showing a loss of the *tert*-butyl signal at δ 1.42 ppm in the ^1H NMR spectrum.

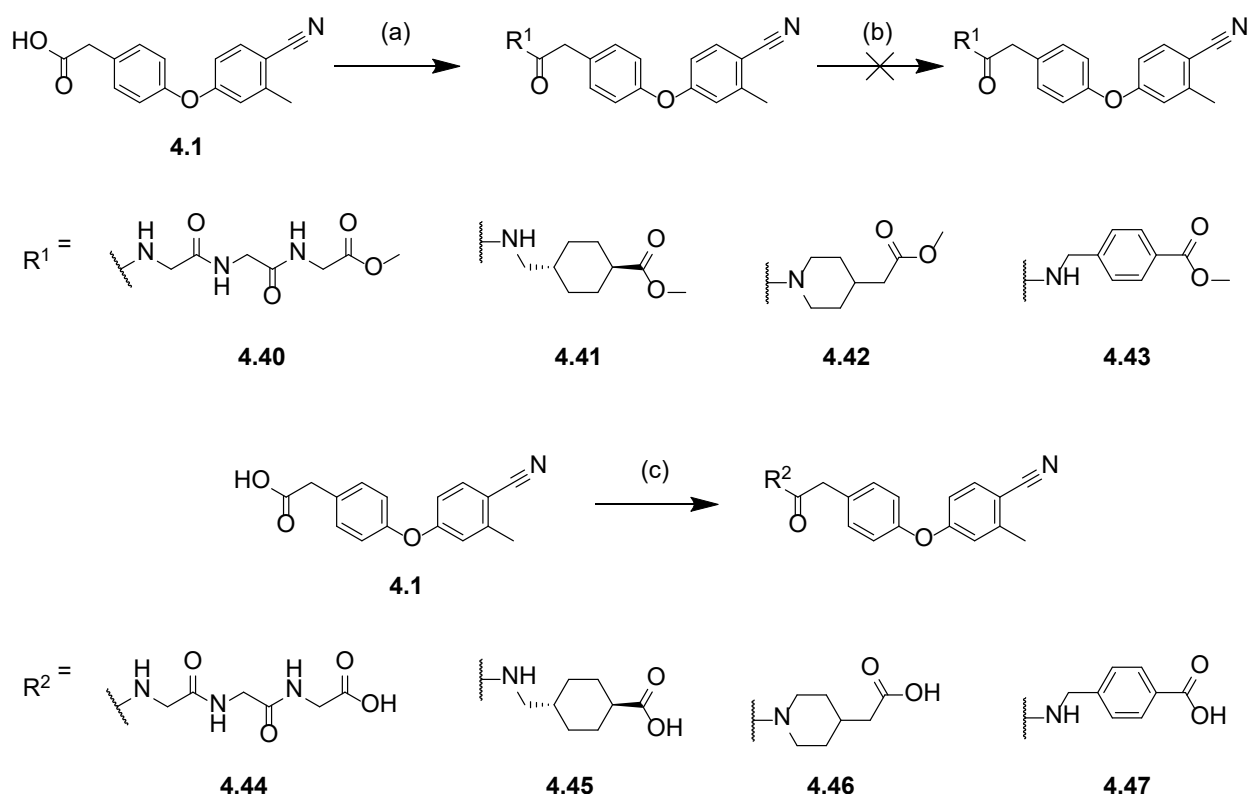
Scheme 4.3: Synthesis of peg linked analogue **4.38**.



Reagents and conditions: (a) HATU, Et_3N , DMF, RT, 24 h, 46 %. (b) TFA, DCM, R.T, 18 h, 69 %.

Similarly, amidation of phenylacetic acid **4.1** with triglycine **4.16**, benzylamine **4.5**, piperidine **4.8**, and methyl *trans*-cyclohexyl carboxylate **4.10** gave a amides **4.40** – **4.43** (**Scheme 4.4**). However, deprotection of the esters of amides **4.40** – **4.43** using LiOH and MeOH showed no product by TLC, crude ^1H NMR or LCMS. A number of reaction conditions were trialled in an effort to optimise hydrolysis conditions of these amides, including DMSO/NaOH co-solvent, iso-propanol/NaOH and increasing equivalents of LiOH/NaOH (5 – 100 eq.) base at room temperature. No reaction was observed after 24 h under these conditions, therefore reactions were conducted at elevated temperature (50 – 100 C) however this yielded only phenylacetic acid **4.1** indicating hydrolysis of amides **4.41** – **4.43**. Thus, an alternate synthetic route to these linked carboxylic acids via acid chloride mediated amidation was investigated. The acid chloride was synthesised from phenyl acetic acid **4.1** using oxalyl chloride and catalytic DMF and was supported by 1D ^1H NMR where the benzylic protons shifted from 3.68 ppm to 4.16 ppm, consistent with the greater electron withdrawing power of acid chloride. Dropwise addition of the commercially available unprotected amino acids at low temperature (0 °C) yielded amides **4.44** – **4.47** in low yields.

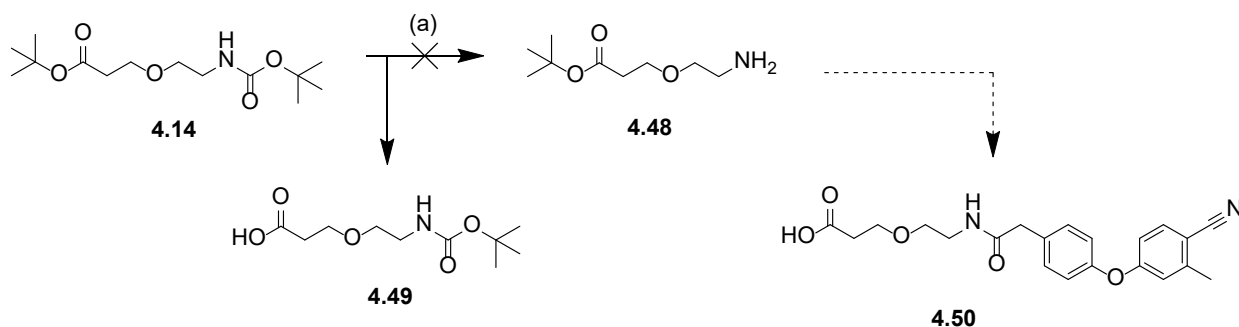
Scheme 4.4: Synthesis of a series of expanded carboxylic acid based diaryl ethers.



Reagents and conditions: (a) amine (R¹), HATU, Et₃N, DMF, RT, 24 h, 23 - 65 %. (b) LiOH or NaOH, THF or DMSO or isopropanol, H₂O, 25 – 100 °C, 6 - 24 h, 0 %. (c) (COCl)₂, amine (R²), DIPEA, DCM/DMF, 0 – 25 °C, 24 - 72 h, 3 - 25 %.

The synthesis of PEG linked analogue **4.50** was investigated using a number of synthetic routes (**Scheme 4.5**). Selective deprotection of the Boc protecting group of the purchased *tert*-butyl-*N*-Boc protected amino acid **4.14** to yield amine **4.48** was attempted following a literature procedure.⁴ A crude ¹H NMR of the reaction showed a loss of a single *tert*-butyl resonance however, ¹³C NMR displayed two carbonyl resonances at 158.8 and 171.0, indicating that the *tert*-butyl ester was removed to give carboxylic acid **4.49**.

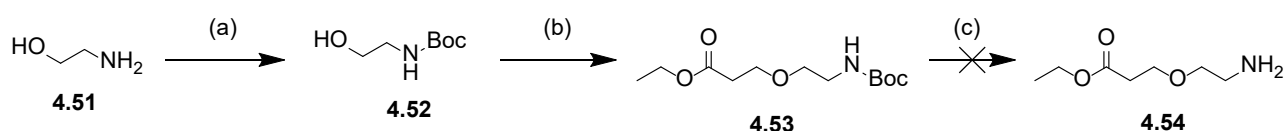
Scheme 4.5: Synthesis of amino ester intermediate **4.48**.



Reagents and conditions: (a) TFA, DCM, 0 °C, 2 h, 0 %.

An alternative route to access ethyl ester **4.54** was detailed in the literature (**Scheme 4.6**).^{5, 6} Boc protection of ethanolamine **4.51** using Boc₂O gave alcohol **4.52**. Michael addition of Boc-protected alcohol **4.52** to ethyl acrylate in the presence of catalytic sodium metal then gave the differentially protected *N*-Boc amino ethyl ester **4.53**. Cleavage of the *N*-Boc protecting group with TFA in DCM to give amine **4.54** was tracked via TLC and visualised using a ninhydrin indicating product was formed however, could not be isolated by flash column chromatography.

Scheme 4.6: Synthesis of PEG linked amino ester intermediate **4.54**.

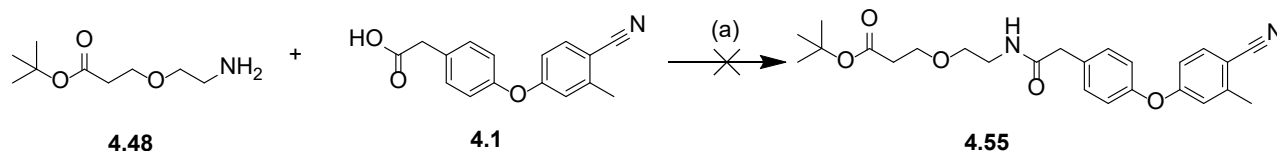


Reagents and conditions: (a) Boc₂O, DCM, RT, 16 h, 43 %. (b) ethyl acrylate, Cat. Na metal, THF, RT, 2 h, 24 %. (c) TFA, DCM, 0 °C, 2 h, 0 %.

The previously described route of acid chloride formation and amine coupling was also investigated with the commercially available amino acid **4.48** (**Scheme 4.7**). Crude ¹H NMR and LCMS showed no product being formed and purification of the reaction returned starting material phenylacetic acid **4.1** and unknown by-products. Amino acid **4.48** and related analogues were expensive to purchase

making optimisation of the reaction conditions costly, thus the synthesis of PEG linked acid **4.50** was not pursued further.

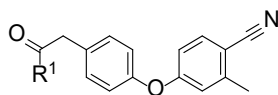
Scheme 4.7: Synthesis of PEG linked carboxylic acid analogue **4.55**.



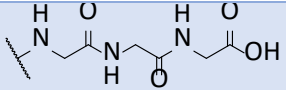
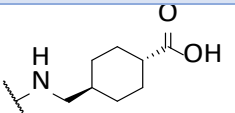
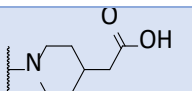
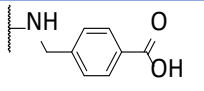
Reagents and conditions: (a) (COCl)₂, DCM, DMF, DIPEA, 0 – 25 °C, 24 h, 0 %.

4.2.3 Testing amidation analogues

These analogues were screened via the solubility, HSQC and X-ray crystallography pipeline previously described. Alkyl linked analogues **4.36**, **4.37** and PEG linked **4.39** maintained high aqueous solubility (max 1000 μ M) but failed to show any improvement in affinity (**Table 4.1**). Rigidifying the linker as in diglycine **4.34** had only a minor effect on binding affinity ($K_D = 308 \pm 31$ μ M) and should be considered with caution as it was tested at 60 % purity due to purification issues. Addition of a 3rd glycine sidechain in triglycine amide **4.44**, reduced aqueous solubility (<125 μ M) and was not further screened. Aromatic ring linked analogues biphenyl **4.30** and benzoic acid **4.47** showed poor aqueous solubility (≤ 125 μ M) while extension to a phenylacetic acid sidechain in acetamide **4.31** and benzyl amide **4.32** displayed moderate aqueous solubility (500 μ M), however all 4 analogues demonstrated no significant improvement in affinity compared to the parent acid **4.1**. The series of saturated cyclohexyl analogue *trans*-cyclohexylacetic acid **4.33**, *cis*-cyclohexylacetic acid **4.35**, *trans*-cyclohexyl carboxylic acid **4.45** and piperidine **4.46** demonstrated notable improvements in solubility over the aromatic analogues although they had similar or weaker affinities compared to the phenyl acetic acid **4.1**.

Table 4.1: Testing data for the expanded amides

Compound number	R ¹	Solubility max (μM)	Single point HSQC NMR	HSQC NMR Affinity (K _b , μM)*	Electron density for the ligand	EDIA score
4.1	OH	1000	Binding	490 ± 20**	Full	0.86
4.30		< 125	-	-	-	-
4.31		500	Binding	715 ± 110	-	-
4.32		500	Binding	349 ± 40	-	-
4.33		1000	Binding	400 ± 25	-	-
4.34		1000	Binding	308 ± 31	-	-
4.35		500	Binding	>2000	-	-
4.36		1000	Binding	650 ± 40	Partial	0.75
4.37		1000	Binding	921 ± 121	-	-
4.39		1000	Binding	595 ± 55	Partial	0.65

4.44		< 125	-	-	Partial	0.61
4.45		500	Binding	475 ± 65	-	-
4.46		1000	Binding	484 ± 38	-	-
4.47		125	-	-	-	-

*Affinities presented as mean ± SEM. **Affinity determined in phosphate buffer.

Limited structural data in the form of X-ray crystal structures of analogues bound to *EcDsbA* were acquired for this series. Those X-ray structures acquired displayed poor electron density for their respective amide extensions upon visual inspection and low EDIA scores. Notably alkyl linker **4.36** and PEG linker **4.39** had electron density for the diaryl ether core and lack density for their respective amidated extension, likely due to the flexibility of the alkyl linkers (**Figure 4.4, a**). The perturbed residues in the ¹H-¹⁵N HSQC spectra of analogues within this series was consistent with known groove binders and did not show any additional perturbations of residues in the LHS of the groove. (**Figure 4.4, b**).

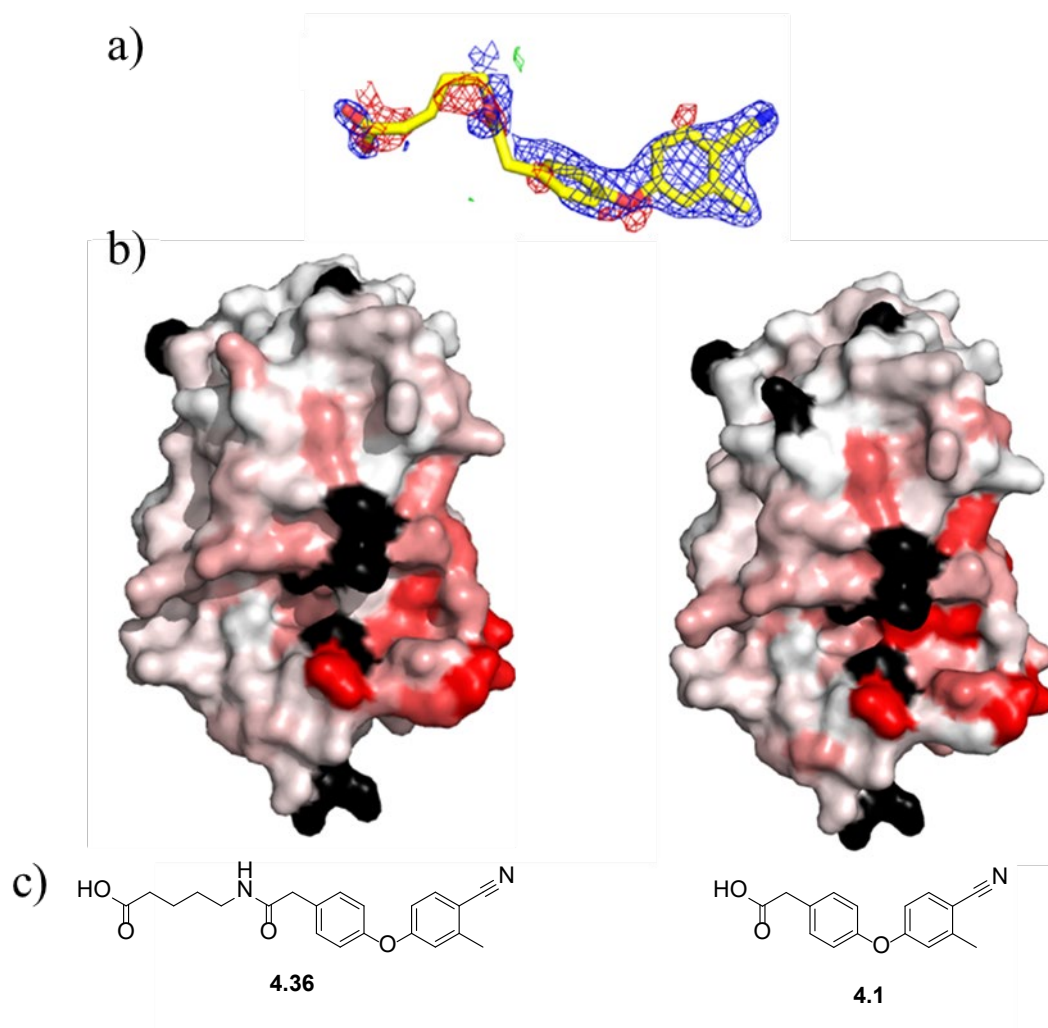


Figure 4.4: a) Electron density map of straight chain linker **4.36**, shown as 2mFo-DFc at 1 σ (blue) and mFo-DFc at 3 σ (positive green and negative red). b) HSQC perturbations of carboxylic acid **4.36** (left) and **4.1** (right) binding to *EcDsbA* are mapped onto the surface of *EcDsbA* (PDB code: 1FVK), darker red colour indicated larger perturbations in that residue, residues that are unassigned in the protein plus ligand spectra are shown in grey and residues unassigned in both apo and plus ligand spectra are displayed in black. The observed CSP are similar to a known groove binder carboxylic acid **4.1** which suggested that **4.36** is binding within the hydrophobic groove, however no significant additional CSP are observed, suggesting that the extension does not interact strongly with *EcDsbA*. c) 2D structures of straight chain linker **4.36** and phenylacetic **4.1**.

4.2.4 Summary

Using SBDD a number of expanded carboxylic acid diaryl ether analogues were designed, docked and synthesised to assess their potential to form an ionic interaction with Arg148. Aqueous solubility was a major concern in the design of this library of analogues as amidation of the carboxylic acid of **4.1** was known to reduce solubility. Therefore, the amines that were used to form the amide with **4.1** were selected to maintain a carboxylic acid in the elaborated compound, and it was found that most derivatives in this series were soluble at $\geq 500 \mu\text{M}$. However, the affinities measured by HSQC titrations revealed that the elaborated compounds did not significantly improve in binding affinity compared to the parent phenylacetic acid analogue **4.1**. The lack of improved or additional CSP in the HSQC spectra, the relatively low affinities of the compounds and the missing electron density observed in the crystallographic data for the synthesised analogues indicate that interactions with Arg148 or other residues on the LHS of the groove were not formed. Although the docking data suggested that these compounds were capable of interacting with the side chain of Arg148, it was hypothesised the $\approx 11 \text{ \AA}$ distance between the Arg148 sidechain and the carboxylic acid of **4.1** along with the flexible sidechain of Arg148 was to not currently amenable to new interactions with this series. Hence, further efforts to expand to the LHS of the groove focused on targeting residues closer to the carboxylic acid of **4.1** in an attempt to generate new interactions. The following section details a series of analogues targeting Pro151 which expanded from the carboxylic acid and benzylic position of **4.1**.

4.3 Benzylic expansion and isosteric replacement of the carboxylic acid

Expansion off the benzylic position or isosteric replacement of the carboxylic acid of **4.1** with an acyl sulfonamide both provide an alternative method to expand across the groove, whilst retaining a polar functional group to maintain solubility. These two series of analogues were investigated in parallel. Both expansion approaches maintain a functional group that is ionisable at physiological pH (COO^- for benzylic and CON-SO_2 for the acyl sulfonamides) and are therefore were expected to maintain

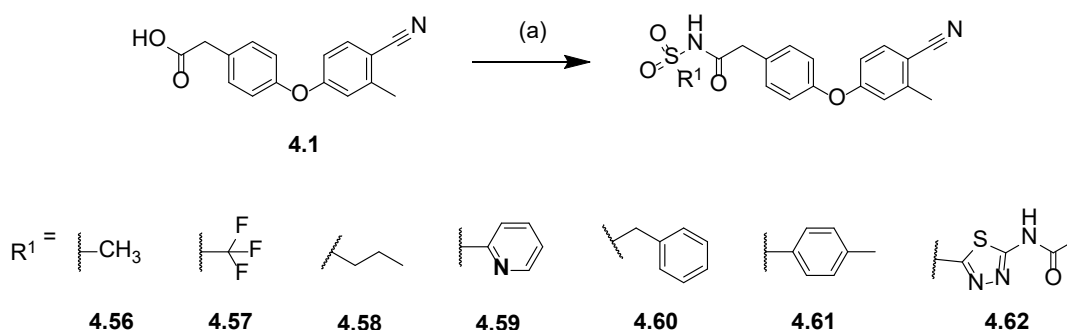
good solubility circumventing some of the solubility problems that had been observed with earlier analogues in the diaryl ether series.

4.3.1 Acyl sulfonamide design, docking and synthesis

Acyl sulfonamides can be accessed synthetically via a coupling reaction of a phenyl acetic acid as in **4.1** and a terminal sulfonamide. Similar to previous series, a library of commercially available reagents from eMolecules (>2,000,000 reagents) were filtered to select for terminal sulfonamides. These reagents were then enumerated as acyl sulfonamides onto the diaryl ether core of **4.1** and docked against *EcDsbA* using Glide. The top 200 docked poses ranked by Glide score were manually inspected. This analysis indicated that using this vector for fragment growth was a viable strategy, with most analogues maintaining the binding mode for the diaryl ether core and projecting their respective substituents across the groove. Consequently, a series of in house and commercially available aliphatic, aromatic and heteroaromatic terminal sulfonamides were selected for synthesis to explore if acylsulfonamides were tolerated and could be used to access the LHS of the groove.

The acyl sulfonamide analogues **4.56** – **4.62** were then synthesised via EDC mediated coupling with 1 eq. DMAP in low to moderate yields (**Scheme 4.8**).

Scheme 4.8: Synthesis of a series of acyl sulfonamide diaryl ether analogues.



Reagents and conditions: (a) sulfonamide, EDC hydrochloride, DMAP, DCM, RT, 24 h, 8 - 58 %.

4.3.2 Benzylic expansion design and docking

Analysis of the crystal structure of phenylacetic acid **4.1** bound to *Ec*DsbA identified the backbone carbonyl of Pro151 as a possible target for interactions via expansion at the benzylic position. This residue has been implicated in the activity, function and stability of *Ec*DsbA⁷ and is conserved in the DsbA enzymes of a number of *Gram*-negative bacterial strains.⁷ The Pro151 carbonyl oxygen is positioned 4.6 Å from the benzylic carbon of phenylacetic acid **4.1** (**Figure 4.5**). H-bond interactions to Pro151 had been observed previously in the X-ray crystal structures of a number of phenylthiazole analogues synthesised internally within the group (W. Alwan unpublished work) thus making it an established and desirable target for additional interactions and improved affinity. It was envisioned that a H-bond to Pro151 could be formed through the addition of a hydrogen bond donor (NH, OH, CONH) 1 – 2 heavy atoms from the benzylic carbon of **4.1**.

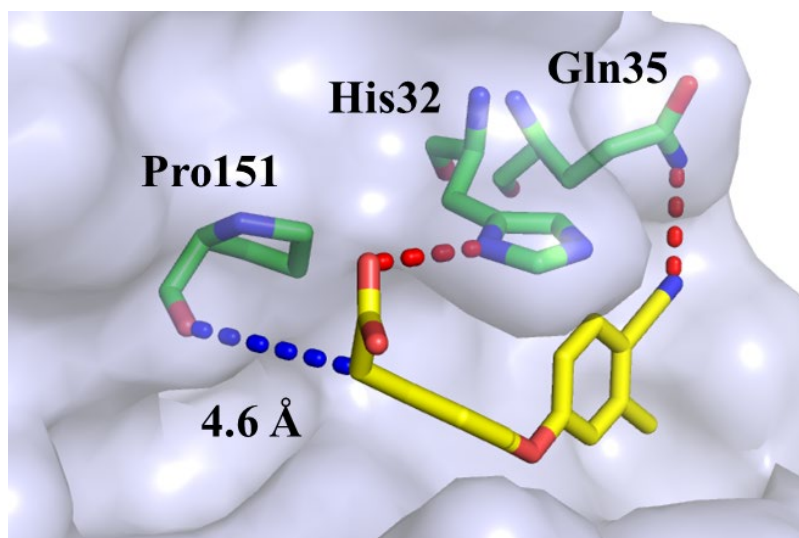


Figure 4.5: X-ray crystal structure of **4.1** (yellow sticks) bound to oxidised *Ec*DsbA (tinted blue surface) showing distance between the benzylic carbon of **4.1** and the backbone carbonyl of Pro151 (green sticks). Key active site residues His32, Gln35 and Pro151 are shown as green sticks. H-bonds shown as red dashed lines, with the distance between the benzylic position and carbonyl of Pro151 shown as a blue dashed line.

A number of synthetically accessible analogues with a variety of different H-bond donors at different distances and geometries from the fragment core of **4.1**, were designed and docked using Glide against *Ec*DsbA (**Figure 4.6, a**). Three different core structures were used to enumerate compounds for docking in an effort to determine the optimal length and geometry between the fragment core and H-bond donor. Many of the docked compounds formed a H-bond interaction with Pro151 and some larger analogues also made a H-bond with Gly149, which is further away. The docking suggested that a H-bond donor either directly attached to the benzylic carbon of carboxylic acid **4.1** or with a single heavy atom spacer between the benzylic carbon and the donor were able to make a H-bond with the carbonyl of Pro151 while maintaining the interaction with His32 (**Figure 4.6, b**). The 3rd core based upon the amino acid tyrosine demonstrated a less favourable geometry for Pro151 interactions, however analogues of this core were easily accessible, hence a subset of expanded diaryl ether analogues based on all 3 cores (amino acid core, β -amino acid core and tyrosine core) were selected for synthesis to test this hypothesis.

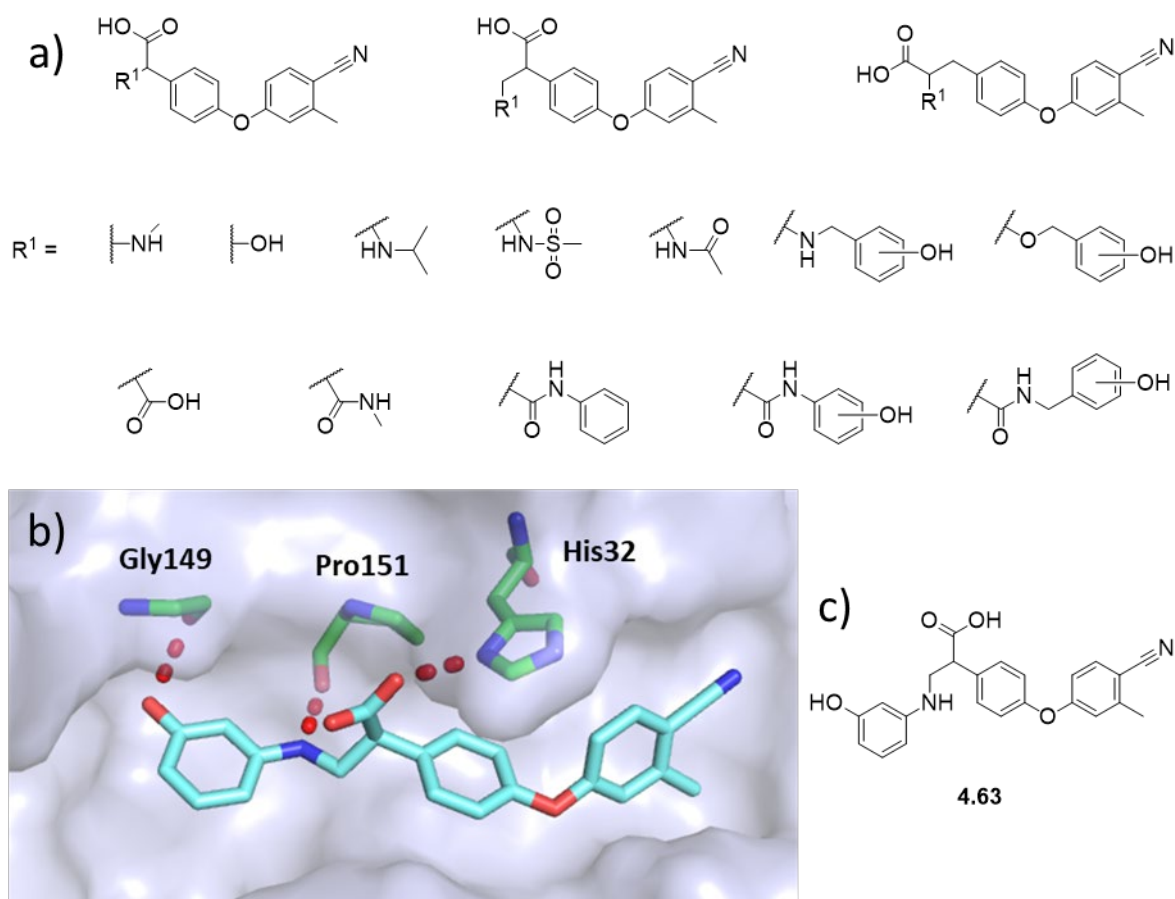
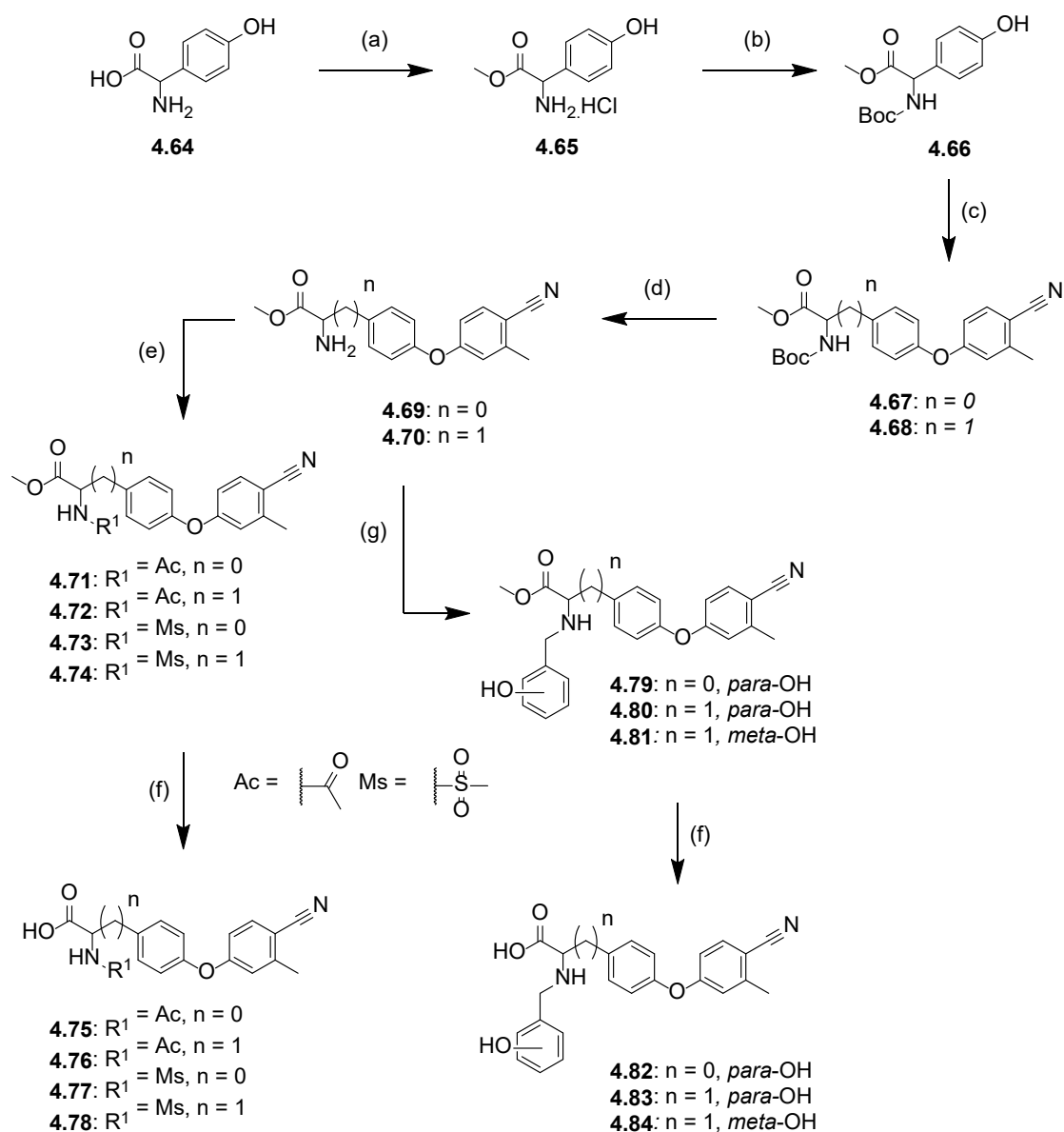


Figure 4.6: a) List of cores and substituents expanding from the benzylic position of **4.1** that were designed, docked against *EcDsbA* and chosen for synthesis (as depicted in the red box). b) Docked structure of a benzylic expanded analogue **4.63** (grey sticks) against *EcDsbA* (grey surface), showing potential interactions with His32, Gly149 and Pro151 (green sticks). H-bonds shown as red dashed lines. c) Structure of docked compound **4.63**.

4.3.3 Synthesis of benzylic expanded analogues

Initially we targeted synthesis of the α -amino acid like cores which have *N*-acyl, *N*-sulfonyl or *N*-alkyl substituents. Fischer esterification of 2-amino-2-(4-hydroxyphenyl)acetic acid (**4.64**) using SOCl_2 and MeOH achieved the ester intermediate **4.65** as the hydrochloride salt (**Scheme 4.9**). *N*-Boc protection of amino ester **4.65** was achieved with Boc_2O in 1,4-dioxane and H_2O following a literature procedure⁸ to give the diprotected phenol **4.66**. Ullmann coupling and subsequent *N*-Boc-deprotection using TFA then gave the key amino ester diaryl ethers **4.69** and **4.70**.

Scheme 4.9: Synthesis of a series of benzylic expanded diaryl ether analogues.



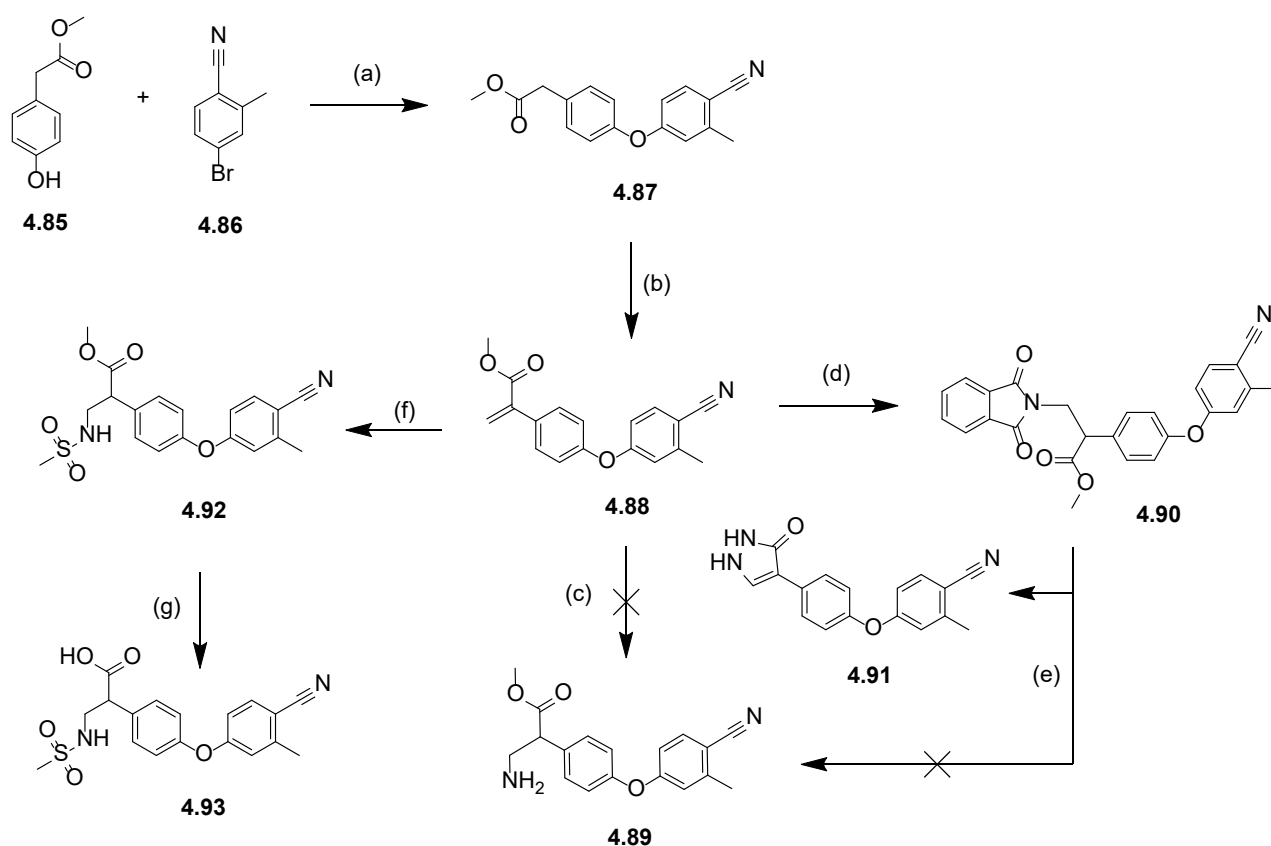
Reagents and conditions: (a) SOCl_2 , MeOH, RT, 72 h, 43 %. (b) Boc_2O , 1,4-dioxane, H_2O , RT 18 h, 54 %. (c) 4-bromo-2-methylbenzonitrile, *N,N*-dimethylglycine hydrochloride, Cs_2CO_3 , DMF, 80 °C, 24 h, 42 - 45 %. (d) TFA, DCM, RT, 24 h, 45 - 91 %. (e) CH_3COCl or $\text{CH}_3\text{SO}_2\text{Cl}$, DCM, RT, 18 h, 46 - 100 %. (f) NaOH, MeOH, H_2O , RT, 18 h, 72 - 96 %. (g) $\text{Na}(\text{AcO})_3\text{BH}$, AcOH, DCE, RT, 18 h, 40 - 73 %.

Acetylation and sulfonylation of amines **4.69** and **4.70** using acetyl chloride and methane sulfonyl chloride afforded amides **4.71** and **4.72** and sulfonamides **4.73** and **4.74**, respectively. Base mediated ester hydrolysis of esters **4.71** – **4.74** then gave carboxylic acids **4.75** – **4.78**. Reductive amination of

amino esters **4.69** and **4.70** with *para* or *meta*-hydroxybenzaldehyde mediated by, Na(AcO)₃BH and catalytic AcOH afforded alkylated amino esters **4.79** – **4.81** in good yields. These transformations were supported by the appearance a new resonance corresponding to the benzylic protons at δ 3.6 ppm in the ¹H NMR spectra. Finally, ester hydrolysis of **4.79** – **4.81** using NaOH/MeOH gave the amino acid products **4.82** – **4.84**.

The commercial availability of suitable β -phenylalanine precursors containing the required phenol functional group for Ullmann coupling was limited, hence an alternative synthetic route was devised to synthesise the key amino methyl substituted fragment **4.89** (**Scheme 4.10**). Ullmann coupling of phenol **4.85** and aryl bromide **4.86** generated the core ester **4.87** as previously described in Chapter 3. Aldol condensation of ester **4.87** with aqueous formaldehyde solution using K₂CO₃ then gave acrylate **4.88**. The reaction conditions of this aldol condensation were optimised by screening reaction solvent (toluene, DMF, CDCl₃), time (5, 24, 48 h), temperature (50, 70 °C), equivalents of formaldehyde (3, 10 eq.) and formaldehyde source (aqueous formaldehyde, solid paraformaldehyde) to improve reaction conversion (estimated by crude NMR) while minimising ester hydrolysis to acid **4.1** (**Appendix 4.1**). Michael addition of NH₃ to acrylate **4.88** showed good conversion to product **4.89** by LCMS with observation of the expected ion at m/z : 310.9 [M+H]⁺, however purification via flash column chromatography isolated the desired product in moderate yields (~44 %) but low purity (< 60 %). Phthalimide was investigated as an alternative nucleophile to NH₃⁹ and Michael addition of phthalimide mediated by a non-nucleophilic base, DBU, achieved the coupled phthalimide intermediate **4.90** in low yield. This transformation was supported by the loss of the alkene resonances at δ 5.85 and 6.30 ppm in ¹H NMR spectrum. Deprotection of the phthalimide protecting group using hydrazine produced a mixture of by-product **4.91** (\approx 25 %), starting material **4.90** (\approx 20 %) and the desired amine **4.89** (\approx 50 %), as estimated by LCMS. The formation of by-product **4.91** was supported by a LCMS mass as well as crude ¹H NMR and ¹³C NMR which indicated no degradation of the benzonitrile, a loss of both the aliphatic CH and CH₂ resonances of **4.90** and a new singlet at δ 7.63 ppm that may correspond to the aromatic CH resonance of **4.91**. The formation of by-product **4.91**

was postulated to occur via displacement of the phthalimide with hydrazine or base mediated elimination and Michael addition followed by ring cyclisation. The low conversion and difficulty in purification of the desired amino product then led us to pursue alternative nucleophiles for the acrylate to install the sulfonamide and acetamide substituents. Michael addition of methane sulfonamide to acrylate **4.88** with DBU base successfully generated sulfonamide intermediate **4.92** followed by base mediated ester hydrolysis gave the desired screening compound sulfonamide **4.93**. Michael addition of acetamide to acrylate **4.88** using DBU or KO^tBu base failed to provide any material for ongoing synthesis and therefore the sulfonamide was tested as a representative analogue from this series.

Scheme 4.10: Synthesis of analogues based on the β -amino acid diaryl ether core.

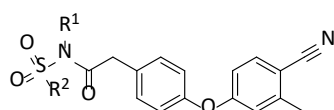
Reagents and conditions: (a) *N,N*-dimethylglycine hydrochloride, Cs₂CO₃, CuI, DMF, 100 °C, 24 h, 56 %. (b) Aq. formaldehyde, tetrabutylammonium iodide (TBAI), K₂CO₃, toluene, 70 °C, 24 h, 35 %. (c) NH₃ in MeOH, MeOH, 0 - 25 °C, 48 h, 0 %. (d) phthalamide, DBU, MeCN, RT, 24 h, 16 %. (e) hydrazine monohydrate, MeOH, DCM, 50 °C, 18 h, 0 %. (f) methanesulfonamide, DBU, MeCN, RT, 8 d, 41 %. (g) NaOH, MeOH, H₂O, RT, 1 h, 85 %.

4.3.4 Testing benzylic expansion and acyl sulfonamide analogues

The acyl sulfonamides were tested for solubility and then binding by single point ¹H-¹⁵N HSQC showing all analogues in this series bound to *Ec*DsbA. This was then followed by ¹H-¹⁵N HSQC titrations to determine their affinity and X-ray crystallography for bound pose as described previously. Isosteric replacement of the carboxylic acid for small acyl sulfonamides such as methyl 4.56, trifluomethyl 4.57, propyl 4.58 or a subset of the larger aromatic substituted acyl sulfonamides pyridinyl 4.59 and benzyl 4.60 maintained high aqueous solubility and resulted only minor increase

in affinity ($K_D = 300 - 371 \mu\text{M}$) compared to the phenylacetic acid **4.1** ($K_D = 490 \mu\text{M}$). While introduction of tolyl or thiadiazole substituents as in **4.61** and **4.62** to the acyl sulfonamide reduced aqueous solubility (max $250 \mu\text{M}$) and prevented an accurate estimate of their binding affinity by HSQC titrations. Despite soaking all 6 analogues to *EcDsbA* crystals, no crystallographic data of sufficient quality was obtained for ligands in this series.

Table 4.2: Testing of acyl sulfonamide analogues.

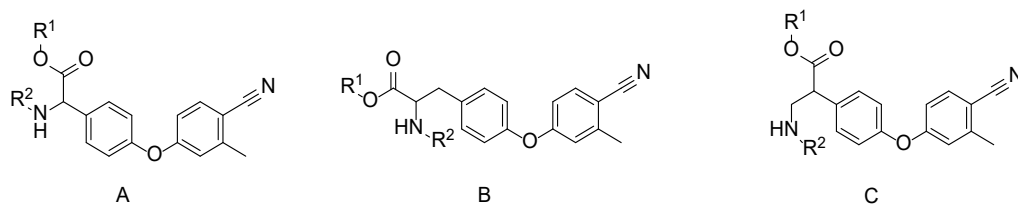


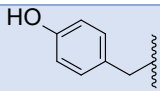
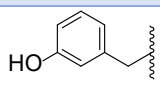
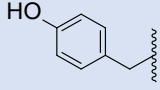
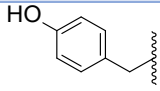
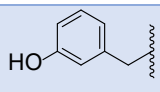
Compound number	R ¹	R ²	Solubility max (μM)	Single point HSQC NMR	HSQC NMR Affinity (K_D , μM)*	Electron Density for the ligand	EDIA score
4.56	H		1000	Binding	330 ± 21	Poor	0.68
4.57	H		1000	Binding	330 ± 27	-	-
4.58	H		1000	Binding	371 ± 13	-	-
4.59	H		1000	Binding	300 ± 12	-	-
4.60	H		1000	Binding	350 ± 20	-	-
4.61	H		250	Binding	200 ± 30	-	-
4.62	H		250	Binding	250 ± 30	-	-

*Affinities presented as mean \pm SEM.

Up to this point in the project the solubility of compounds was assessed in 50 mM sodium phosphate, 25 mM NaCl, pH 7.4 and ^1H - ^{15}N HSQC spectra were recorded with 100 μM ^{15}N EcDsbA in 50 mM HEPES buffer, 25 mM NaCl, 2 % DMSO, 10 % D_2O and 88 % H_2O at pH 6.8. Investigation by other members of the group had indicated 2 - 4-fold discrepancies in solubility between HEPES and phosphate buffer hence, HSQC experiments detailed below and all subsequent ^1H - ^{15}N HSQC experiments were performed using 50 mM phosphate, 25 mM NaCl in 2 % DMSO, 10 % D_2O and 88 % H_2O at pH 6.8.

Screening of the benzylic position expanded analogues via HSQC and X-ray crystallography was then conducted and revealed binding for all compounds with solubility $\geq 250 \mu\text{M}$. Benzylamine ester **4.69** demonstrated poor aqueous solubility (max 125 μM), whereas the tyrosine-based amino ester (**4.70**) passed solubility at 1000 μM but had a 3-fold reduced affinity ($K_D = 1400 \pm 150 \mu\text{M}$) compared to the parent phenylacetic acid **4.1** ($K_D = 500 \pm 20 \mu\text{M}$). Amino acid analogues **4.80** – **4.84** demonstrated low solubility in neat DMSO, 2 M aq. hydrochloride and 1 M phosphate buffer (pH 7) at room temperature, thus these analogues were not tested by ^1H - ^{15}N HSQC. The amide and sulfonamide derivatives **4.75** – **4.78** and **4.93** displayed no significant improvements in affinity over the parent carboxylic acid **4.1**, suggesting that interactions with Pro151 were not formed. Furthermore, X-ray crystallographic analysis of these expanded analogues gave no electron density for the ligands, which did not allow for detailed analysis of the binding mode.

Table 4.3: Testing results for the benzylic expansion series.

Compound number	Core	R ¹	R ²	Solubility max (μ M)	Single point HSQC NMR	HSQC NMR Affinity (K_D , μ M)*	Electron Density for the ligand
4.69	A	CH ₃	H	125	-	-	-
4.70	B	CH ₃	H	1000	Binding	1435 \pm 145	-
4.75	A	H	COCH ₃	500	Binding	722 \pm 116	-
4.76	B	H	COCH ₃	1000	Binding	1153 \pm 152	-
4.77	A	H	SO ₂ CH ₃	125	-	-	-
4.78	B	H	SO ₂ CH ₃	1000	Binding	328 \pm 15	-
4.80	B	CH ₃		< 125	-	-	-
4.81	B	CH ₃		< 125	-	-	-
4.82	A	H		< 125	-	-	-
4.83	B	H		< 125	-	-	-
4.84	B	H		< 125	-	-	-

4.93	C	H	SO ₂ CH ₃	1500	Binding	578 ± 30	-
------	---	---	---------------------------------	------	---------	----------	---

*Affinities presented as mean ± SEM. Affinities were determined in phosphate buffer.

4.3.5 Summary

Replacement of the carboxylic acid with an acyl sulfonamide had little impact on affinity or solubility of the compounds tested within this series. In general, the series displayed flat SAR with the majority of the compounds demonstrating affinities of $K_D \approx 300 - 400 \mu\text{M}$ where these could be measured from the HSQC data. The benzylic expanded compounds designed to target polar interactions with Pro151 and Gly149 were supported by docking calculations, but failed to significantly improve binding affinities across the series. This suggested that no new interactions were formed due to the lack of affinity gains and structural information from this elaboration strategy, hence this analogue series was not developed further.

4.4 Conclusion

Using traditional SBDD guided by X-ray crystallography and computational docking three different series of analogues were designed and synthesised to investigate expansion of the phenylacetic acid diaryl ether core **4.1** across to the LHS of the groove. However, after synthesis of 32 compounds probing this vector no significant improvements in affinity were observed and no new structural information was obtained to suggest what expansion was favoured at this vector. Hence, further elaboration of compounds at this vector guided by SBDD was hypothesised to be difficult and time consuming.

To this point in the project, >90 diaryl ether analogues had been designed, synthesised and tested and few analogues had demonstrated a measurable affinity of $K_D < 300 \mu\text{M}$. Inhibition of protein-protein interaction binding sites is frequently difficult due to the binding interfaces often being relatively flat, featureless and not ideal for binding of small molecules.¹⁰ *EcDsbA* is no exception to this observation, as demonstrated by the challenges faced in elaborating the initial lead fragment. In this case the SBDD

strategy employed here had failed provide any lead compounds with significantly higher affinity.¹¹

Hence, a decision was taken to pursue expansion via an alternate strategy.

4.5 References

1. Shonberg, J.; Herenbrink, C. K.; López, L.; Christopoulos, A.; Scammells, P. J.; Capuano, B.; Lane, J. R. A structure–activity analysis of biased agonism at the dopamine D2 receptor. *J. Med. Chem.* **2013**, 56, 9199-9221.
2. Perrone, M. G.; Vitale, P.; Panella, A.; Ferorelli, S.; Contino, M.; Lavecchia, A.; Scilimati, A. Isoxazole-Based-Scaffold Inhibitors Targeting Cyclooxygenases (COXs). *ChemMedChem* **2016**, 11, 1172-1187.
3. Romulus, J.; Weck, M. Single-Chain Polymer Self-Assembly Using Complementary Hydrogen Bonding Units. *Macromol. Rapid Commun.* **2013**, 34, 1518-1523.
4. Gibson, F. S.; Bergmeier, S. C.; Rapoport, H. Selective Removal of an N-BOC Protecting Group in the Presence of a tert-Butyl Ester and Other Acid-Sensitive Groups. *J. Org. Chem.* **1994**, 59, 3216-3218.
5. Varala, R.; Nuvula, S.; Adapa, S. R. Molecular Iodine-Catalyzed Facile Procedure for N-Boc Protection of Amines. *J. Org. Chem.* **2006**, 71, 8283-8286.
6. Reddy, D. S.; Vander Velde, D.; Aubé, J. Synthesis and Conformational Studies of Dipeptides Constrained by Disubstituted 3-(Aminoethoxy)propionic Acid Linkers. *J. Org. Chem.* **2004**, 69, 1716-1719.
7. Heras, B.; Shouldice, S. R.; Totsika, M.; Scanlon, M. J.; Schembri, M. A.; Martin, J. L. DSB proteins and bacterial pathogenicity. *Nat. Rev. Microbiol.* **2009**, 7, 215-25.
8. Chang, D. J.; Jeong, M. Y.; Song, J.; Jin, C. Y.; Suh, Y.-G.; Kim, H.-J.; Min, K. H. Discovery of small molecules that enhance astrocyte differentiation in rat fetal neural stem cells. *Bioorg. Med. Chem. Lett.* **2011**, 21, 7050-7053.
9. Ginsburg, D. Practical considerations in preparation of amines In *Concerning Amines*, Ginsburg, D., Ed. Pergamon: 1967; pp 17-62.
10. Arkin, M. R.; Tang, Y.; Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chem. Biol.* **2014**, 21, 1102-1114.
11. Ran, X.; Gestwicki, J. E. Inhibitors of protein-protein interactions (PPIs): an analysis of scaffold choices and buried surface area. *Curr. Opin. Chem. Biol.* **2018**, 44, 75-86.

**Chapter 5 – Rapid elaboration of fragments
into leads by X-ray crystallographic screening
of parallel chemical libraries (REFiL_x)**

RESEARCH ARTICLE

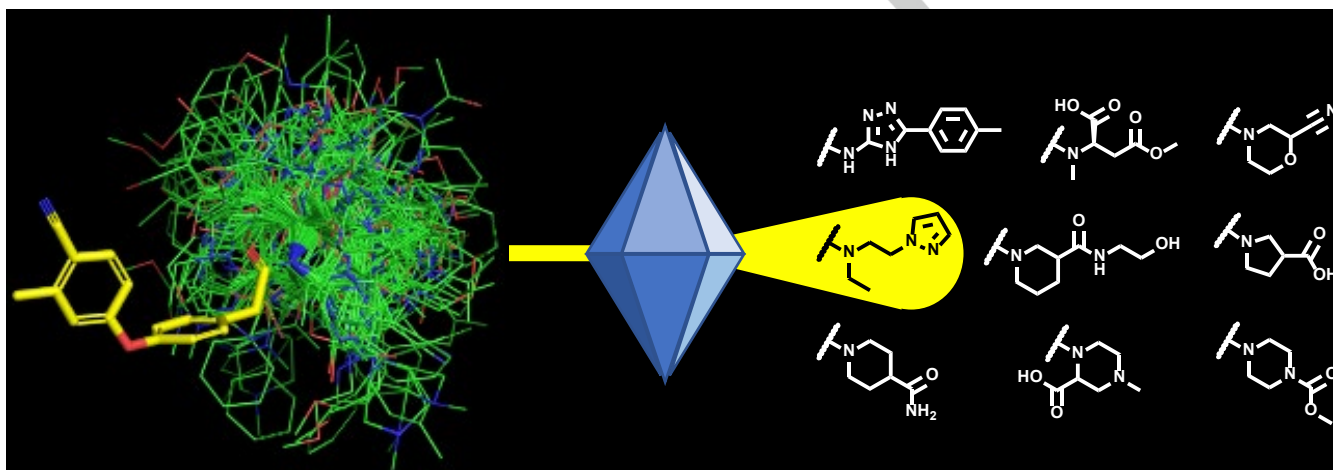
5.1 Introduction

The work detailed below is a manuscript that will be submitted to *Angewante Chemie International Edition*. This paper details our novel approach to more efficiently elaborate fragments into higher affinity binders. As previously discussed, interactions on the RHS of the groove of *EcDsbA* had been thoroughly optimised whilst expansion into the LHS of the groove targeted multiple polar interactions but afforded few analogues with improved affinity. We therefore sought to use an alternate approach with structurally diverse chemical libraries of analogues obtained by 96 well plate-based parallel synthesis to more efficiently and thoroughly explore the LHS of the groove. One implementation of parallel synthesis adopted within our lab is the use of parallel chemistry and the screening of chemical libraries as unpurified reactions using a method called off rate screening (ORS) by surface plasmon resonance (SPR).² We are unable to use ORS by SPR with *EcDsbA* due to affinity constraints as discussed in the manuscript. Hence, we sought to apply the same parallel chemistry approach, but to use X-ray crystallography as a readout. We hypothesised this parallel screening approach using X-ray crystallography would enable us to apply parallel chemistry earlier in the fragment to lead process and thereby more efficiently develop higher affinity binders. This manuscript details our implementation of this method. The phenylacetic core **4.1** was chosen to exemplify this Rapid Elaboration of Fragments into Leads by X-ray crystallography (REFiLx) method as it has good solubility, moderate affinity, an X-ray crystal structure of the bound ligand and a suitable vector for expansion off the carboxylic acid. Though this vector had been explored in chapter 4, the scale and diversity of the parallel libraries would enable greater chemical space coverage, unbiased by computational and SBDD approaches. We hypothesised this would more efficiently probe the LHS of the groove for interactions and identify analogues with improved affinity, whilst also yielding additional structural information. Experimental methods for this manuscript are located in **Appendix 5.1**. I conducted all aspects of the chemistry, LCMS, NMR analysis, affinity determination and assisted Dr Olga Ilyichova in optimising and acquiring the crystallographic data.

Rapid elaboration of fragments into leads by X-ray crystallographic screening of parallel chemical libraries (REFiLx)

Matthew R. Bentley,^{+[a]} Olga V. Ilyichova,^{+[a]} Geqing Wang,^[b] Martin L. Williams,^[a] Gaurav Sharma,^[a] Wesam S. Alwan,^[a] Rebecca L. Whitehouse,^[a] Biswaranjan Mohanty,^[a] Peter J. Scammells,^[a] Begoña Heras,^[b] Jennifer L. Martin,^[c] Makrina Totsika,^[d] Ben Capuano,^[a,e] Bradley C. Doak*^[a] and Martin J. Scanlon*^[a,e]

We implemented a fragment development approach that enables the screening of libraries of unpurified reaction products by X-ray crystallography to identify higher affinity analogues without purification. Using this approach, termed REFiLx, we gained >8-fold increase in affinity for our challenging protein target *EcDsbA*.



Rapid elaboration of fragments into leads by X-ray crystallographic screening of parallel chemical libraries (REFiL_x)

Matthew R. Bentley,^{†[a]} Olga V. Ilyichova,^{†[a]} Geqing Wang,^[b] Martin L. Williams,^[a] Gaurav Sharma,^[a] Wesam S. Alwan,^[a] Rebecca L. Whitehouse,^[a] Biswaranjan Mohanty,^[a] Peter J. Scammells,^[a] Begoña Heras,^[b] Jennifer L. Martin,^[c] Makrina Totsika,^[d] Ben Capuano,^[a,e] Bradley C. Doak^{*[a]} and Martin J. Scanlon^{*[a,e]}

Abstract:

Fragment hits often bind to their target with weak affinity, requiring sensitive biophysical techniques and structural data to guide their elaboration into higher affinity molecules. Rapid fragment development approaches analysing libraries of unpurified reactions or *in situ* libraries have been reported but require specific chemistries and/or affinity ranges to be employed. Herein we describe the identification of higher affinity analogues from a library of parallel microscale reactions analysed by X-ray crystallography. This rapid elaboration of fragments into leads by X-ray crystallography (REFiL_x) method was applied to a challenging protein-protein interaction target, yielding an 8-fold improvement in affinity and detailed structural information from a single round of REFiL_x. This method can be applied to targets amenable to X-ray crystallography in a wide affinity range, allowing for more efficient development of low affinity fragment hits.

Introduction

Over the past 20 years, fragment-based screening has become established as a robust method to find starting points for drug development.^[1] Due to their small size, fragment hits typically form a small number of interactions with their target protein which results in low affinity binding. As a consequence, significant medicinal chemistry effort is required to produce higher affinity lead compounds.^[2] This design-synthesis-evaluation cycle required to elaborate fragment hits can become a bottleneck to development.

approach, compounds with higher affinity are detected by virtue of their slower dissociation kinetics in SPR sensorgrams. The dissociation rate is independent of free analyte concentration, which means that this method does not require the concentration of a compound to be determined before it is tested. Moreover, since weakly binding compounds dissociate rapidly, the dissociation phase of the sensorgram for a mixture is dominated by the most slowly dissociating component, even when it is a minor component of the mixture.^[3] This allows compounds to be tested without extensive purification. In addition, SPR requires only small quantities of material, allowing moderately sized libraries of compounds to be synthesised on small scale for testing by ORS without extensive purification. However, this approach requires that the dissociation rate constant for the test compounds be slow enough to be resolved by SPR. Typically this requires that the dissociation rate constant, $k_{off} < 1 \text{ s}^{-1}$. Assuming an association rate constant in the range of $k_{on} = 1 \times 10^5 - 1 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$, this produces a threshold whereby the equilibrium dissociation constant for the products must be $K_D < 1\text{--}10 \text{ }\mu\text{M}$ before ORS can be employed. As noted above, it is common for the initial fragment hits, particularly for difficult targets, to bind with weak affinity, i.e. often $K_D > 500 \text{ }\mu\text{M}$. Therefore, in order for ORS to be successful, a significant improvement (> 50 -fold) would be required to meet the affinity threshold whereby dissociation kinetics are likely to be observed. Among other techniques, X-ray crystallography has been used previously in fragment development with dynamic combinatorial chemistry (DCC) to identify more potent compounds *in situ*.^[4] However, DCC is limited in scope by the small number of reactions suitable for *in situ* chemistry. Therefore, we sought to develop a method applicable to a broader range of chemistries^[5] (harsher conditions and organic solvents), as well as be applicable to development of weak affinity fragments not amenable to ORS development.

Advantages of using X-ray crystallography include that it can be used to characterise binding across a broad range of affinities, including weakly binding compounds, and that it provides important structural information for subsequent chemical elaboration. Furthermore, recent advances in X-ray crystallography data collection and analysis, such as automation of crystallisation experiments, development of single-photon-counting detectors, pipeline processing software and multi-dataset analysis methods for identification of ligand binding, have enabled significantly higher throughput and more routine screening of libraries.^[6] The drawbacks of X-ray crystallography screening may include false positives/negatives, ambiguity in binding mode due to crystal contacts, and errors in interpretation of electron density. Despite these challenges, it was anticipated that screening unpurified reaction products by X-ray crystallography could identify compounds of higher affinity and provide key structural information to enable the rapid elaboration

- [a] Mr M. R. Bentley[†], Dr O. V. Ilyichova[†], Dr M. L. Williams, Dr G. Sharma, Mr W. S. Alwan, Ms R. L. Whitehouse, Dr B. Mohanty, Prof. P. J. Scammells, Dr B. Capuano, Dr B. C. Doak, Prof. M. J. Scanlon.
Medicinal Chemistry, Monash Institute of Pharmaceutical sciences, Monash University
381 Royal Parade, Parkville, VIC 3052 (Australia)
E-mail: martin.scanlon@monash.edu
bradley.doak@monash.edu
- [b] Dr G. Wang, A/Prof. B. Heras.
Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, VIC 3086, Australia
- [c] Prof. J. L. Martin
Griffith institute for drug discovery, building N75, Brisbane Innovation Park, Don Young Road, Nathan QLD 4111
- [d] A/Prof M. Totsika
School of Biomedical Sciences, Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, QLD 4059, Australia
- [e] ARC Training Centre for Fragment Based Design, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, VIC 3052, Australia
- [f] This research was undertaken on the MX1 and MX2 beamline at the Australian Synchrotron, part of ANSTO. We acknowledge the CSIRO Collaborative Crystallisation Centre (www.csiro.au/C3).
- [g] Work was carried out in collaboration with the Monash Fragment Platform.
- [†] These authors contributed equally to the work.
- [*] Corresponding authors.

Therefore, methods that enable a more rapid and systematic approach to fragment elaboration are needed.

One such method, termed off-rate screening (ORS) by surface plasmon resonance (SPR), was described by Murray et al.^[3] In this

RESEARCH ARTICLE

of fragments into leads for targets to which ORS and other techniques are not applicable.

Herein we report an initial proof-of-principle study that we performed on *Escherichia coli* DsbA (*EcDsbA*). We have previously undertaken a fragment screen against *EcDsbA* and demonstrated that fragment binding is amenable to analysis by X-ray crystallography.^[7] *EcDsbA* is a disulfide oxidoreductase enzyme responsible for catalysing the formation of disulfide bonds in diverse substrate proteins, including secreted proteins and virulence factors.^[8] Therefore, inhibitors of *EcDsbA* have been suggested as anti-virulence agents.^[9] *EcDsbA* can be considered a “difficult” target for drug development since its binding site is a shallow, extended hydrophobic groove that accommodates a diverse array of substrate proteins. In addition, due to its hydrophobic nature there are relatively few opportunities for specific polar interactions that enable the generation of high affinity small molecule binders. Therefore, we

aimed to develop inhibitors of *EcDsbA* based on a previously identified diaryl ether core, using this rapid elaboration of fragments into leads by X-ray crystallography (REFiL_X).^[7]

Results and Discussion

Diaryl ether **1** was identified from a fragment screen and demonstrated weak binding to *EcDsbA* ($K_D > 1000 \mu\text{M}$, **Figure 1b**).^[7] Structure-based approaches to elaborate fragment hit **1** led to the synthesis of phenylacetic acid **2**, which was found to bind to *EcDsbA* with an improved affinity of $K_D = 490 \pm 20 \mu\text{M}$ (**Figure 1b**). Analysis of the X-ray crystal structure of **2** bound to *EcDsbA* suggests that the acetate moiety forms a charge reinforced hydrogen bonding interaction with His32's imidazole NH in the active site of *EcDsbA*, while the nitrile group participates in a hydrogen bond with Gln35's sidechain amide NH. The diaryl ether core fills a hydrophobic pocket

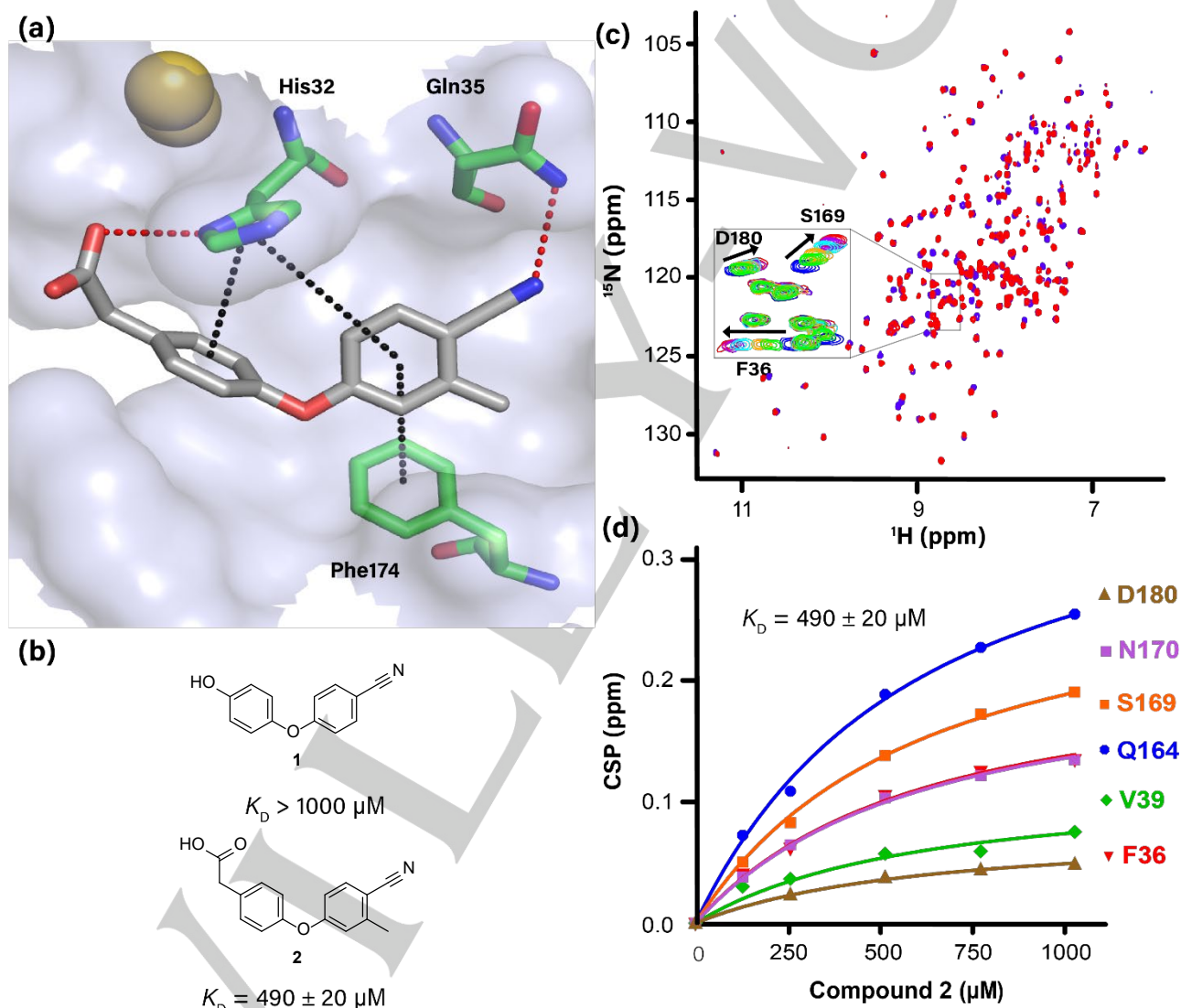


Figure 1. Structure and biophysical characterisation of phenylacetic acid **2** binding to *EcDsbA*. (a) X-ray co-crystal structure of **2** (shown as grey sticks) bound to oxidised *EcDsbA* (shown as blue-grey surface with key residues as green sticks and active site cysteines (Cys30 and Cys33) as yellow spheres). Red dashed lines represent hydrogen-bond and polar interactions, black dashed lines show π -stacking interactions. (b) Structure of diaryl ether hit **1** and phenylacetic acid **2**. Affinities of **1** and **2** were calculated from NMR data using a single site binding model with ligand depletion, which is displayed as $K_D \pm$ error of fit. (c) ^1H - ^{15}N HSQC spectra showing titration of phenylacetic acid **2** against oxidised $u\text{-}^{15}\text{N}$ *EcDsbA* (100 μM) at concentrations of 0, 125, 250, 500, 750 & 1000 μM shown in blue, green, orange, cyan, purple & red respectively. (d) Analysis of ^1H - ^{15}N HSQC titrations with **2** plotting concentration (μM) vs chemical shift perturbation (CSP) in ppm.

RESEARCH ARTICLE

adjacent to the active site of *EcDsbA* and contributes π -stacking interactions with the side chains of His32 and Phe174 (**Figure 1a**).

To explore the development of analogues expanding from the carboxylic acid of **2**, a library of 93 diverse amine reagents (**SI Table 1**) was designed to maximise the coverage of 3D pharmacophore space around the expansion vector whilst at the same time balancing physicochemical properties (clogP, PSA, flexibility, complexity etc.) and synthetic tractability. The library was designed to add 5–12 heavy atoms to the parent fragment, which, assuming a group efficiency of $0.3 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{HAC}^{-1}$, would potentially yield ≈ 10 to 100-fold improvement in affinity.^[10] To maximise diversity of the library protecting groups were used for functional groups that were likely to interfere with the amide formation chemistry to be employed, which required deprotection prior to screening. Due to the wide range of reactivities expected for the amine reagents in the library, we first optimised the reaction conditions to identify those that gave the best yields across a small test set of representative amines. A series of 45 reaction conditions were trialled for 3 amines of varying reactivity, using different coupling reagents, equivalents of amine, temperature and solvent (**SI Fig 1**). Simultaneously, crystals of *EcDsbA* were evaluated for their sensitivity to possible reaction conditions (reagents, solvents, and by-products) as well as optimisation of soaking times (details in **SI**). The most robust set of conditions identified from the chemistry and X-ray crystallography were then used to react the 93 diverse amine reagents with phenylacetic acid **2** (**Figure 2**). Reactions were performed on microscale (2 μmol) in parallel 96 well plate format. Upon completion of the reactions and required deprotections (TFA treatment for acid labile and NaOH treatment for base labile protecting groups) the reactions were evaporated and dissolved in DMSO to generate a stock. The stocks

were made at a notional concentration of 100 mM assuming 100 % reaction conversion without further purification. The reactions were analysed via LCMS to identify the presence of the desired product and estimate purity. For 65 of the 93 reactions the product was unambiguously identified by both UV/Vis and mass spectrometry (MS). Of the remaining 38 reactions, 3 only identified the desired product mass in MS while 15 showed only an additional peak in the UV/Vis trace consistent with where the desired product was expected. For all reaction where a UV/Vis peak was observed for the product purities were calculated by measuring the area of the new peak in divided by the total area of all peaks, excluding the HATU peak. The remaining 4 reactions showed no evidence of product by either MS or UV. The purities ranged from not observed to 90 %, with an average of 24 % across all 93 reactions with significant amounts of carboxylic acid starting material **2** observed in the majority of low purity samples. The importance of purity/conversion in the subsequent X-ray crystallography screening was investigated further, as described below. Each DMSO stock was diluted 10 fold and soaked into *EcDsbA* crystals using previously described soaking and cryoprotection conditions.^[7] Data were collected at MX1 and MX2 beamlines at the Australian Synchrotron.^[11] All datasets were processed with the automated data processing pipeline implemented at the beamline.^[12] Briefly, each dataset was indexed, integrated and scaled with *xdsmc* and *Aimless*.^[13] The output statistical description of the data was manually inspected and datasets with poor statistics were rejected at this stage (see **SI** for a full description of rejection parameters). In total 90 of the 93 samples had datasets with acceptable collection statistics and high resolution data within the range of 1.8–2.5 Å (**SI Fig 3 and 4**) with 9 of these requiring lower sample concentration and shorter soaking times. The remaining 3

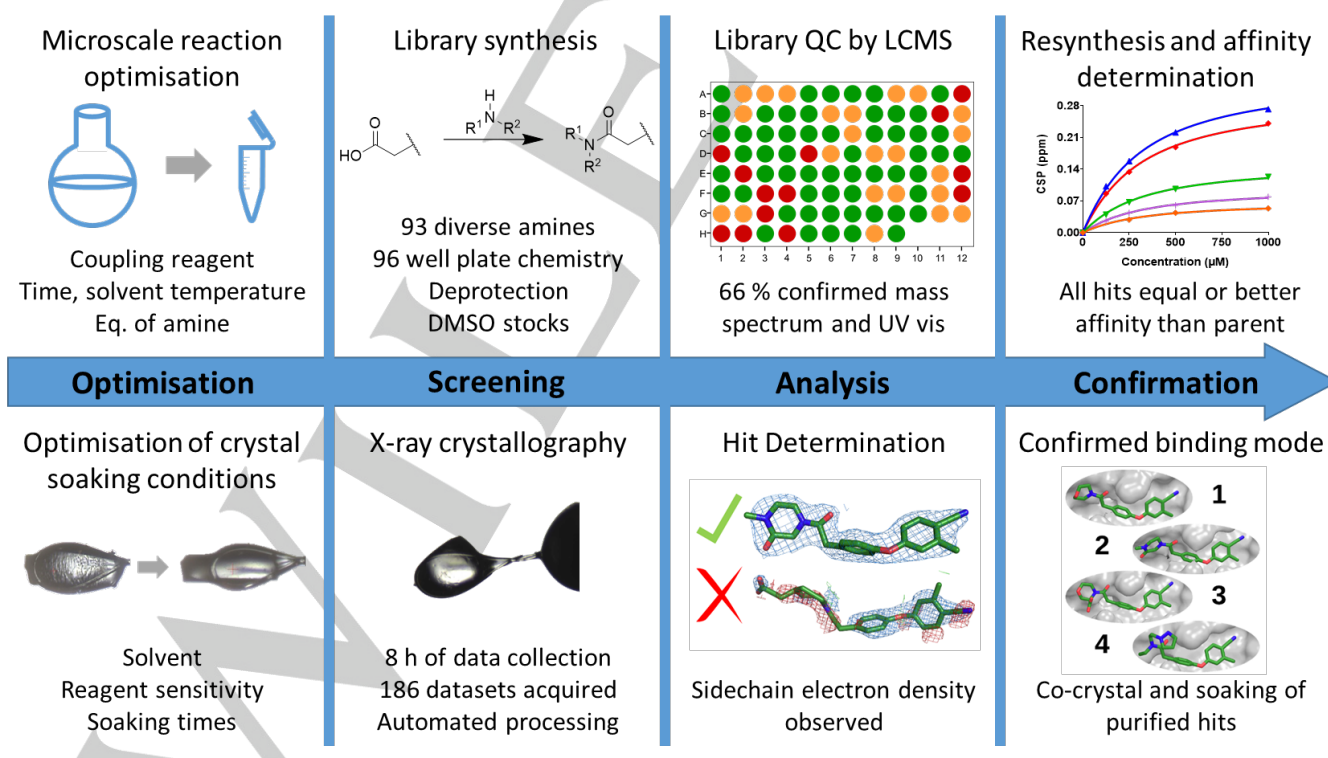


Figure 2. REFIL_x workflow diagram. Optimisation: reaction conditions optimised on microscale in conjunction with optimising X-ray soaking conditions and assessing crystal stability. Screening: microscale reactions with a library of diverse reagents were setup in a 96 well plate, protecting groups were cleaved and DMSO stocks prepared. DMSO stocks soaked into *EcDsbA* crystals, data collected at synchrotron and automated processing implemented. Analysis: reaction conversions estimated by LCMS. X-ray hits determined based on additional electron density observed in the data. Confirmation: hits resynthesised as pure compounds for full characterisation and affinity determination. Pure compounds were co-crystallised with *EcDsbA* to obtain structures of their complexes.

RESEARCH ARTICLE

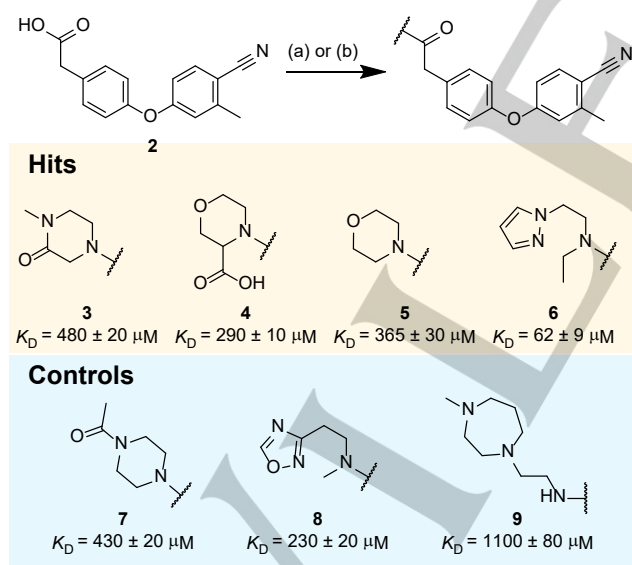
samples did not produce datasets of sufficient quality after multiple soaking and collection attempts. All datasets were refined automatically using the Auto-Rickshaw molecular replacement multi-dataset pipeline^[14] and inspected manually. In addition the datasets were analysed using Pan-Dataset Density Analysis (PanDDA)^[6a] to automatically identify bound ligands. 82 of the 90 analysed samples showed electron density for only the diaryl ether core present in the parent phenylacetic acid **2**. Four samples were identified as hits due to the presence of additional electron density that allowed modelling of their corresponding amide expansions from the parent (**Figure 2** and **Scheme 1** compounds **3–6**). These four hits were resynthesised and purified for full characterisation, K_D determination and validation by X-ray crystallography.

The solubility of resynthesised analogues was determined by recording a quantitative 1D ^1H NMR of the compound in aqueous buffer at multiple concentrations as previously described (**SI Table 2**).^[15] Binding to *EcDsbA* was assessed by recording ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) spectra of oxidised ^{15}N -labelled *EcDsbA* with increasing amounts of ligand. The titrations were continued until the ligand concentration reached either 1 mM or, where the ligands were not soluble at 1 mM, the maximum concentration at which they were deemed soluble. Affinities were determined by measuring chemical shift perturbations (CSP) of backbone amide resonances as a function of ligand concentration and fitting the data to a single site binding model. All of the hits demonstrated similar or better affinity for *EcDsbA* compared to the starting phenylacetic acid **2**, with the highest affinity analogue pyrazole **6** showing an 8-fold improvement in affinity ($K_D = 62 \pm 9 \mu\text{M}$, **Figure 3**). All purified compounds were analysed by co-crystallisation and soaking which gave similar electron density maps to those obtained from the crude products in the library (**SI Table 3**). X-ray analysis of pyrazole **6** bound to *EcDsbA* found similar polar

stacking interaction with the imidazole ring of His32. The additional interactions observed in this unexpected binding mode most likely underpin the increase in affinity of pyrazole **6** over the parent **2** in binding to *EcDsbA*.

To assess the effects of reaction conversion and purity on the outcome of the REFIL_X screening strategy, several additional compounds were resynthesised and purified (**Scheme 1** compounds **7, 8** and **9**). The library analogue oxadiazole **8** is structurally similar to the highest affinity hit pyrazole **6** and gave only low reaction purity in the library (6 %) as estimated by LCMS, hence it was resynthesised to test for false negatives in the screen. Screening by ^1H - ^{15}N HQSC determined its K_D to be $230 \pm 20 \mu\text{M}$ (**SI Fig 5**), a 2-fold improvement over the parent **2**. A crystal structure was also determined for oxadiazole **8** using the purified material (**SI Fig 6**). The X-ray structure of **8** has partial electron density, however would be sufficient to identify oxadiazole **8** as a possible hit using the purified compound. Oxadiazole **8** was not identified as a hit by REFIL_X despite its improved affinity compared to parent **2** which can be attributed to its low purity in the library synthesis and highlights one of the limitations of screening unpurified reaction products by X-ray crystallography. Additional high conversion products **7** and low conversion product **9** that were not identified as hits were confirmed as having similar or weaker affinity for *EcDsbA* compared to parent **2** (**SI Table 2**).

These limitations were further investigated via a series of soaking experiments containing both parent **2** and higher affinity hit pyrazole **6**, which were prepared from the purified compounds as mixtures at relative concentrations corresponding to conversions of 6, 13 and 28 %. These samples were made in the absence or presence of a faux background amidation reaction with unrelated reactants (**SI Fig 6**). Density for the pyrazole extension was observed at all conversions in the absence of background faux amidation reaction, however decreased density was observed for the pyrazole sidechain in the presence of faux background reactions and this was sensitive to conversion (**SI Fig 7**). This suggests that X-ray screening can produce false negatives from crude reactions at low conversions (<28 %) and is consistent with the outcome observed with oxadiazole **8**. Nonetheless, in the current dataset 39 samples in the library achieved conversion/purity of >28 % by LCMS.



Scheme 1. Synthesis of hit compounds (top row) and control analogues (bottom row). Reagents and conditions: (a) HATU, Et₃N, amine, DMF, MeCN, 25 °C, 18 h. (b) (COCl)₂, DMF cat., CDCl₃, 0 °C, 5 h, 30 %, then amine, Et₃N, CDCl₃, 25 °C, 24 h, 4 – 64 %. Purity of library synthesised samples estimated by LCMS for hits **3** = 74 %, **4** = 55 %, **5** = 83 %, and **6** = 62 % as well as control analogues **7** = 86 %, **8** = 6 %, and **9** = 0 %.

and π -stacking interactions to those observed for the parent **2** (**Figure 3a**) with additional water-mediated interactions of the amide carbonyl oxygen of pyrazole **6** with the backbone NH of His32. The pyrazole sidechain of **6** projects out from the shallow groove and folds back towards the active site of *EcDsbA*, where it forms an additional π -

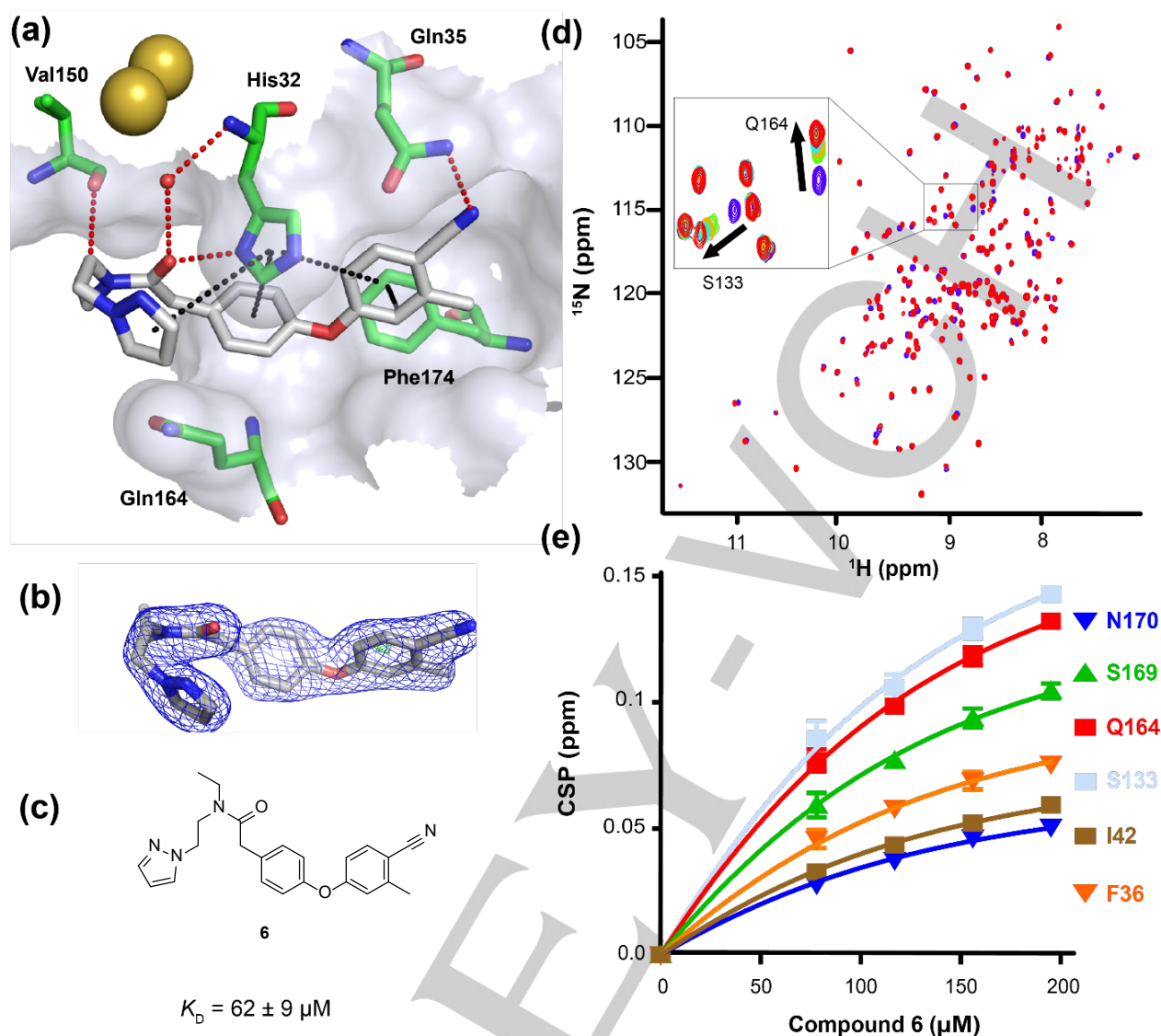


Figure 3. (a) X-ray co-crystal structure of pyrazole **6** bound to oxidised *EcDsbA*. Red dashed lines represent hydrogen bond and polar interactions, black dashed lines show π -stacking interactions. Amino acids are displayed in green sticks, active site cysteines (Cys30 and Cys33) are shown as yellow spheres and pyrazole **6** is shown as grey sticks. (b) Electron density maps of pyrazole **6**, 2mFo-DFc map is contoured at 1 σ (blue) and mFo-DFc maps are shown at $\pm 3 \sigma$ (positive green and negative red). (c) Structure of pyrazole **6** with its affinity estimated by ^1H - ^{15}N HSQC titrations. (d) ^1H - ^{15}N HSQC spectra of pyrazole **6** against oxidised $u\text{-}^{15}\text{N}$ *EcDsbA* at concentrations of 0, 100, 150, 200 & 250 μM shown in blue, green, orange, cyan & red respectively. (e) ^1H - ^{15}N HSQC titrations of pyrazole **6** plotting concentration (μM) vs CSP (ppm). Data from three replicate titration experiments are plotted as mean \pm SEM, fitted to a single site binding model with ligand depletion (see SI for full details).

Conclusion

We have implemented a fragment development approach which uses X-ray crystallography to screen chemical libraries synthesised in parallel on microscale without extensive purification to identify compounds that bind to a protein target. Using this approach, termed REFIL_X, we synthesised a library of 93 amides, which we soaked into crystals of *EcDsbA*. We identified 4 hits which gave clear electron density in the X-ray data for the expanded side chain with the highest affinity hit being pyrazole **6** with an affinity of $K_D = 62 \pm 9 \mu\text{M}$. This approach is complementary to the strategy of ORS by SPR,^[3] that uses X-ray crystallography as the readout to identify higher affinity binders. An advantage of REFIL_X is that it can operate over a wider affinity range. For example – although pyrazole **6** provided an 8-fold affinity gain over the parent compound, the dissociation kinetics for pyrazole **6** would be too fast to measure by SPR. Current limitations of REFIL_X, include the availability of a suitably robust crystal system that is amenable to soaking, and the requirement for reasonable

conversion in the parallel synthesis ($\geq \approx 28\%$ in the current study). However, REFIL_X can also provide structural information on the most promising compounds in a library to guide further design of high affinity ligands. We therefore expect that REFIL_X will provide a useful tool in the armamentarium for early-stage fragment elaboration – particularly against challenging targets such as *EcDsbA* where the initial fragments bind with very low affinity.

Experimental Section

Full details are presented in the supporting information.

Acknowledgements

We thank Monash University and NHMRC (Project grant ID 1009785) for funding.

Keywords: Fragment-based design • Medicinal chemistry • Parallel synthesis • X-ray crystallography • *EcDsbA* • antibiotics • Drug discovery

RESEARCH ARTICLE

- [1] D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, H. Jhoti, *Nat. Rev. Drug Discov.* **2016**, advance online publication.
- [2] G. M. Keserü, D. A. Erlanson, G. G. Ferenczy, M. M. Hann, C. W. Murray, S. D. Pickett, *Journal of Medicinal Chemistry* **2016**, 59, 8189-8206.
- [3] J. B. Murray, S. D. Roughley, N. Matassova, P. A. Brough, *J. Med. Chem.* **2014**, 57, 2845-2850.
- [4] M. S. Congreve, D. J. Davis, L. Devine, C. Granata, M. O'Reilly, P. G. Wyatt, H. Jhoti, *Angewandte Chemie* **2003**, 115, 4617-4620.
- [5] S. D. Roughley, A. M. Jordan, *J. Med. Chem.* **2011**, 54, 3451-3479.
- [6] aN. M. Pearce, T. Krojer, A. R. Bradley, P. Collins, R. P. Nowak, R. Talon, B. D. Marsden, S. Kelm, J. Shi, C. M. Deane, F. von Delft, *Nature Communications* **2017**, 8, 15123; bP. Kraft, A. Bergamaschi, C. Broennimann, R. Dinapoli, E. F. Eikenberry, B. Henrich, I. Johnson, A. Mozzanica, C. M. Schlepütz, P. R. Willmott, B. Schmitt, *J. Synchrotron Radiat.* **2009**, 16, 368-375; cR. Dinapoli, A. Bergamaschi, B. Henrich, R. Horisberger, I. Johnson, A. Mozzanica, E. Schmid, B. Schmitt, A. Schreiber, X. Shi, G. Theidel, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **2011**, 650, 79-83.
- [7] L. A. Adams, P. Sharma, B. Mohanty, O. V. Ilyichova, M. D. Mulcair, M. L. Williams, E. C. Gleeson, M. Totsika, B. C. Doak, S. Caria, K. Rimmer, J. Horne, S. R. Shouldice, M. Vazirani, S. J. Headey, B. R. Plumb, J. L. Martin, B. Heras, J. S. Simpson, M. J. Scanlon, *Angew. Chem. Int. Ed. Engl.* **2015**, 54, 2179-2184.
- [8] B. Heras, S. R. Shouldice, M. Totsika, M. J. Scanlon, M. A. Schembri, J. L. Martin, *Nat. Rev. Microbiol.* **2009**, 7, 215-225.
- [9] B. Heras, M. J. Scanlon, J. L. Martin, *Br. J. Clin. Pharmacol.* **2015**, 79, 208-215.
- [10] A. L. Hopkins, G. M. Keserü, P. D. Leeson, D. C. Rees, C. H. Reynolds, *Nat. Rev. Drug Discov.* **2014**, 13, 105-121.
- [11] T. M. McPhillips, S. E. McPhillips, H. J. Chiu, A. E. Cohen, A. M. Deacon, P. J. Ellis, E. Garman, A. Gonzalez, N. K. Sauter, R. P. Phizackerley, S. M. Soltis, P. Kuhn, *Journal of synchrotron radiation* **2002**, 9, 401-406.
- [12] aN. P. Cowieson, D. Aragao, M. Clift, D. J. Ericsson, C. Gee, S. J. Harrop, N. Mudie, S. Panjikar, J. R. Price, A. Riboldi-Tunnicliffe, R. Williamson, T. Caradoc-Davies, *Journal of synchrotron radiation* **2015**, 22, 187-190; bD. Aragão, J. Aishima, H. Cherukuvada, R. Clarken, M. Clift, N. P. Cowieson, D. J. Ericsson, C. L. Gee, S. Macedo, N. Mudie, S. Panjikar, J. R. Price, A. Riboldi-Tunnicliffe, R. Rostan, R. Williamson, T. T. Caradoc-Davies, *Journal of synchrotron radiation* **2018**, 25, 885-891.
- [13] aW. Kabsch, *Acta Crystallographica Section D* **2010**, 66, 133-144; bP. R. Evans, G. N. Murshudov, *Acta crystallographica. Section D, Biological crystallography* **2013**, 69, 1204-1214; cP. Legrand; dP. Legrand, **2017**.
- [14] aS. Panjikar, V. Parthasarathy, V. S. Lamzin, M. S. Weiss, P. A. Tucker, *Acta crystallographica. Section D, Biological crystallography* **2005**, 61, 449-457; bS. Panjikar, V. Parthasarathy, V. S. Lamzin, M. S. Weiss, P. A. Tucker, *Acta crystallographica. Section D, Biological crystallography* **2009**, 65, 1089-1097.
- [15] S. R. LaPlante, R. Carson, J. Gillard, N. Aubry, R. Coulombe, S. Bordeleau, P. Bonneau, M. Little, J. O'Meara, P. L. Beaulieu, *J. Med. Chem.* **2013**, 56, 5142-5150.

Chapter 6 – Optimisation of the ethyl pyrazole and expansion across the groove.

6.1 Introduction

The previous chapter detailed the growth of the diaryl ethers using a novel screening method REFIL_X. This entailed the screening of a series of unpurified reaction products by X-ray crystallography. Resynthesis and testing of hits identified ethyl pyrazole **6.1** as the highest affinity small molecule *EcDsbA* inhibitor that we had generated at this stage. This method utilised an in-house library of diverse amine reagents that were specifically designed to maximise coverage of pharmacophoric elements in 3D space around the vector. This reagent library was designed to probe different binding motifs (i.e. HBA, HBD, aryl, negative, positive charges) at varying 3D geometries from the attachment point and therefore be suitable for use in any medicinal chemistry program. As a consequence, it conducts a coarse screen covering a diverse chemical space. Thus, having identified fragment **6.1** as a hit we wanted to generate closer structural analogues to further optimise this binder. We therefore investigated a) replacement of the pyrazole ring, b) removal of the *N*-ethyl group and c) expansion of this analogue by additional round of parallel synthesis and screening.

6.2 Isosteric replacement of the ethyl pyrazole

6.2.1 Design of pyrazole analogues

The X-ray co-crystal structure of ethyl pyrazole **6.1** bound to *EcDsbA* positions the pyrazole ring proximal to the highly flexible Gln164 sidechain (**Figure 6.1**, ~3.5 Å away). Although a H-bond between the pyrazole ring of **6.1** and the sidechain amide of Gln164 is unlikely at this distance, we hypothesised that a H-bond to Gln164 would be possible via replacement of the pyrazole ring with related nitrogen-containing heterocycles.

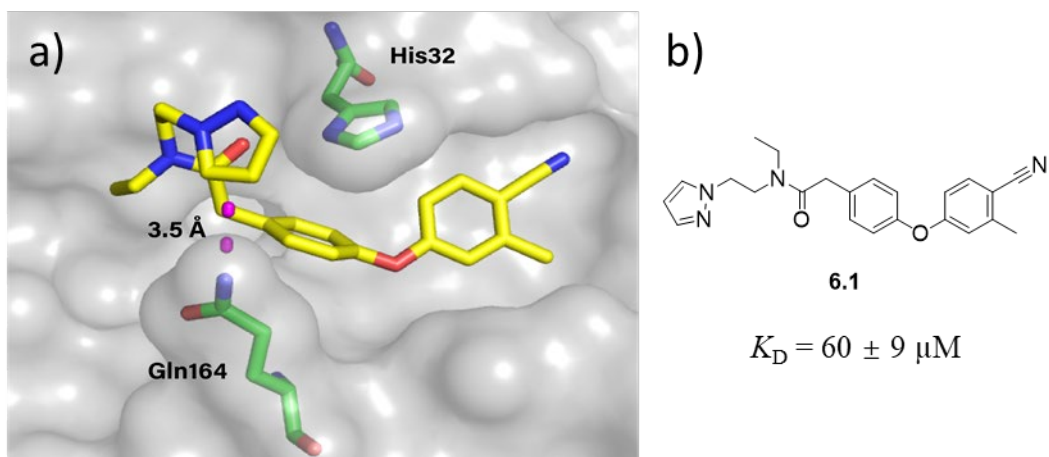


Figure 6.1: a) X-ray crystal structure of pyrazole **6.1** (yellow sticks) bound to *EcDsbA* (grey surface) with key residues Gln164 and His 32 (green sticks). The distance between the pyrazole ring (C4 and C5) and the sidechain amide of Gln164 is shown as a purple dashed line. b) 2D structure of pyrazole **6.1**.

Analogues with suitable commercially available amine precursors were prioritised to maximise the efficiency in exploring this hypothesis. As few *N*-ethyl aminoethyl heterocycle reagents were commercially available in sufficient quantities for synthesis, a series of aminoethyl heterocycles were purchased (**Figure 6.2**) and the desired *N*-ethyl derivatives were to be accessed synthetically through alkylation chemistry. Additionally, mono-alkyl amide intermediates were of interest to test against *EcDsbA* and subsequently formed a second series assessing the importance of the *N*-ethyl functional group of **6.1** for binding to *EcDsbA*.

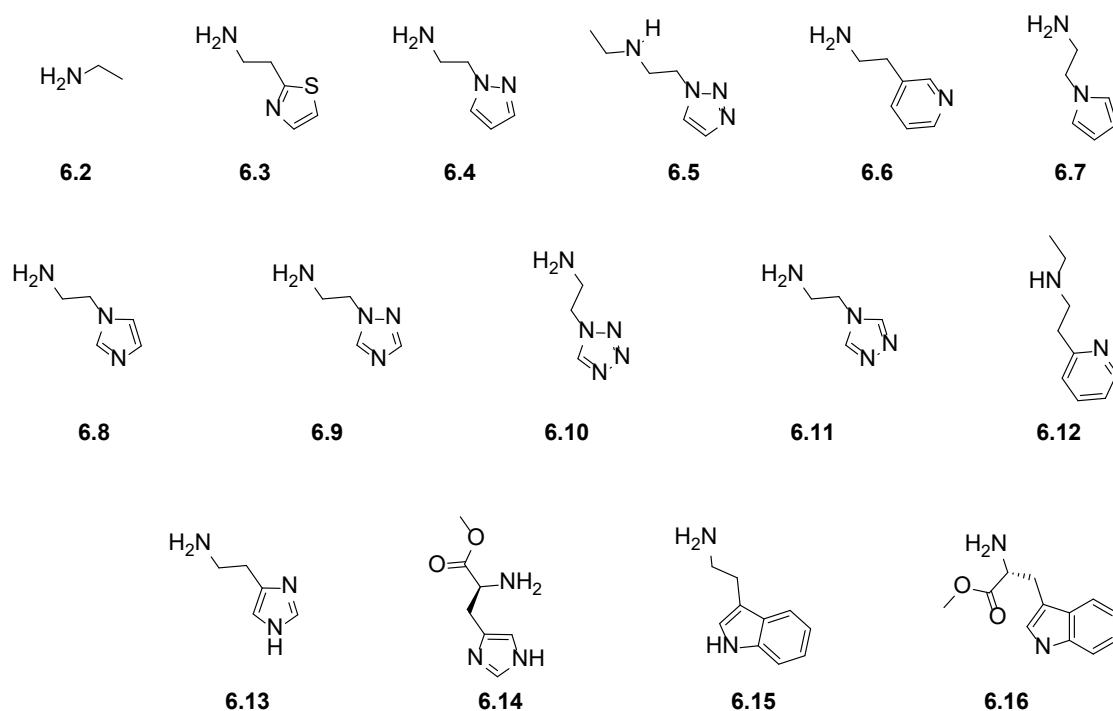
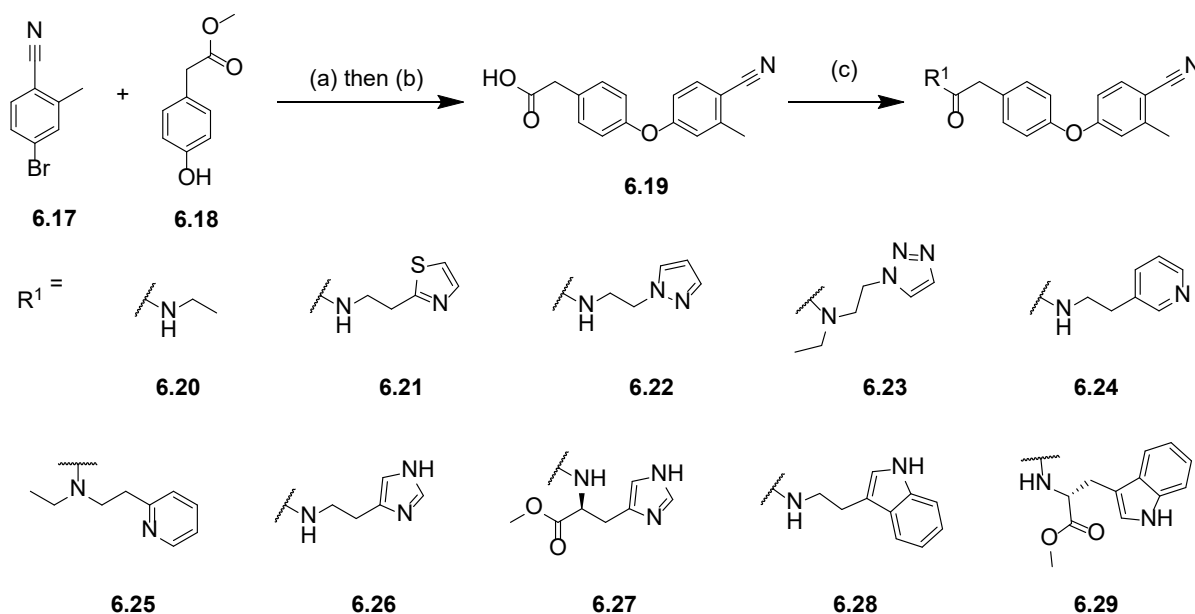


Figure 6.2: A series of amino ethyl heterocycle reagents exploring replacement of the pyrazole ring of **6.1**.

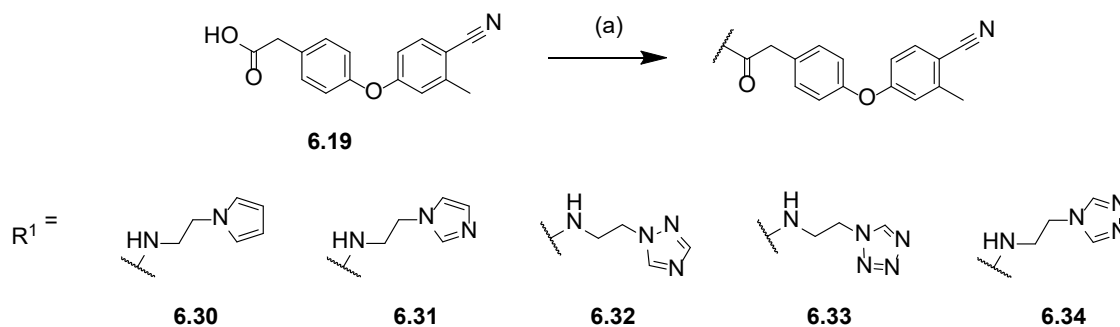
6.2.2 Replacing the pyrazole ring

Key phenylacetic acid precursor **6.19** was synthesised via the previously optimised route of Ullmann coupling and ester hydrolysis (see Chapter 3). Access to the heterocycle analogues was initiated through HATU-mediated amide coupling of carboxylic acid **6.19** with amino heterocycles **6.2** - **6.6** and **6.12** - **6.16** in the presence of Et₃N to afford the corresponding amides **6.20** – **6.29** (Scheme 6.1).

Scheme 6.1: Synthesis of heteroaryl amide analogues series via HATU mediated amidation.

Reagents and conditions: (a) *N,N*-dimethylglycine hydrochloride, CuI, Cs₂CO₃, DMF, 105 °C, 24 h, 56 %. (b) LiOH, MeOH, H₂O, R.T, 6 h, 95 %. (c) HATU, Et₃N, DMF, R.T, 16 h, 4 – 33 %.

Attempts to couple amines **6.7** – **6.11** to phenylacetic acid **6.19** failed due to low reaction conversion (as determined by crude NMR, LCMS and TLC) and difficulty in purification of the amides by column chromatography. To resolve this, the amide coupling conditions between amine **6.4** with phenylacetic acid **6.19** were optimised. This optimisation investigated different; solvent (DMF, MeCN, DCM), base (DBU, Et₃N, K₂CO₃) and coupling agents (EDC/hydroxybenzotriazole (HOBt), HATU). A combination of EDC/HOBt in DMF with Et₃N afforded the highest conversion (>95 %) estimated by LCMS. Hence, a number of amides **6.30** – **6.34** were synthesised using these optimised conditions (**Scheme 6.2**).

Scheme 6.2: Synthesis of a series of pyrazole analogues via EDC mediated amide coupling.

Reagents and conditions: EDC hydrochloride, HOBt, Et₃N, DMF, 2 – 18 h, 36 -70 %.

6.2.3 Synthesis of *N,N*-dialkyl amide analogues

As the corresponding *N*-ethyl-amines were often not commercially available and to assess the importance of the *N*-ethyl group of pyrazole **6.1** a pathway for synthesis of the *N,N*-dialkyl analogues from the corresponding mono-alkyl amines was also investigated. Alkylation was envisioned to be accessible by either reductive amination to the secondary amine then amide coupling or selective alkylation of the mono-alkyl amide NH (**Figure 6.3**). The direct alkylation method utilised late stage derivatisation of each analogue, but raised concerns around non-selective alkylation of the amide (NH pK_a \approx 26)¹ and benzylic position (benzylic CH pK_a \approx 27)² hence, reductive amination was initially prioritised for investigation.

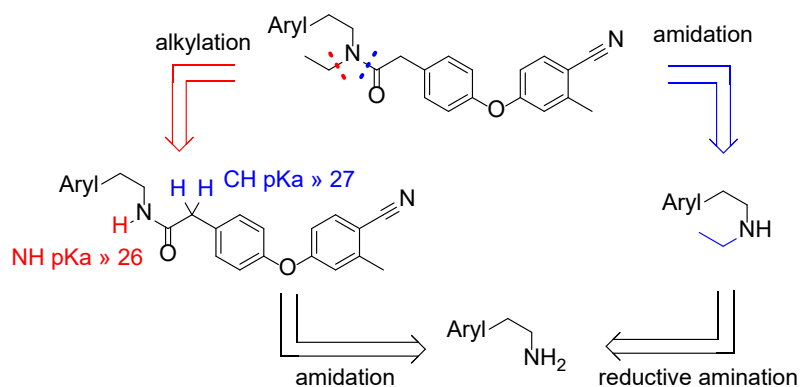
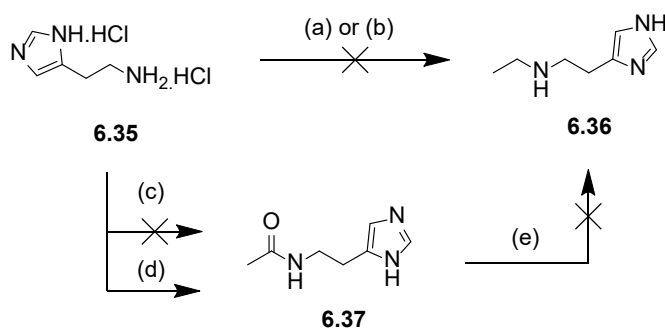


Figure 6.3: Retrosynthetic analysis of the *N,N*-dialkyl derivatives of pyrazole **6.1**. Estimated pKa values of an acetamide NH (red) and activated benzylic protons (blue) are shown. Values were sourced from Bordwell pKa table (<https://www.chem.wisc.edu/areas/reich/pkatable/>)

Histamine dihydrochloride (**6.35**, **Scheme 6.3**) was used as an initial case to optimise the synthesis of *N,N*-dialkyl analogues. Acid catalysed reductive amination of histamine (**6.35**) with acetaldehyde using $\text{Na}(\text{AcO})_3\text{BH}$ in DCM showed no reaction progress toward the desired secondary amine **6.36** by ^1H NMR or LCMS (**Scheme 6.3**). Consequently, a higher boiling point solvent (MeOH) and addition of a drying agent Na_2SO_4 were used in attempt to increase imine formation. After overnight reflux NaBH_4 was added to reduce the imine and any remaining aldehyde starting material. Analysis of the crude reaction mixture by LCMS, TLC and NMR identified only histamine **6.35** starting materials hence, an alternative acylation and reduction approach to access these *N*-ethyl amines was investigated.

Scheme 6.3: Synthesis to access the key dialkyl amine **6.36**.

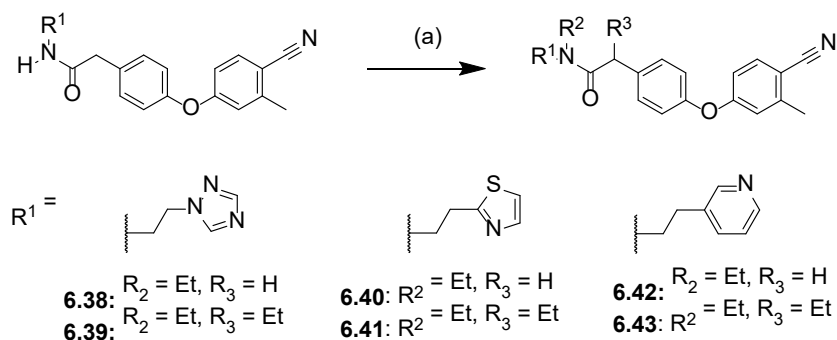
Reagents and conditions: (a) acetaldehyde, AcOH, Na(AcO)₃BH, DCM, RT to reflux, 18 h, 0 % (b) acetaldehyde, AcOH, NaBH₄, Na₂SO₄, MeOH, reflux, 18 h, 0 % (c) acetyl chloride, DBU, Et₃N or K₂CO₃, MeCN, RT, 48 h, 0 %. (d) acetic anhydride, K₂CO₃, EtOH, RT, 18 h, >100 %. (e) LiAlH₄, THF, 0 – 25 °C, 3 h, 0 %.

Acetylation of histamine **6.35** with acetyl chloride (**Scheme 6.3**) showed no reaction progress by crude ¹H NMR after testing multiple bases (Et₃N, DBU) and solvents (MeCN, DMF). It is possible that failure of these reactions may be due to quenching of the electrophile by residual water within the reaction or in the reagent bottle. As such, acetic anhydride was used as an alternative electrophile which gave amide **6.37** in quantitative yield with K₂CO₃ and potassium acetate likely contaminants. Attempts to purify the amide **6.37** via trituration with water also dissolved the desired product thus, amide **6.37** was used as a crude mixture in the following reaction. Reduction of the amide bond using LiAlH₄ gave no indication of the desired *N*-ethyl product by TLC, ¹H NMR or LCMS. The excess acetate and carbonate salt was suspected to cause this reaction to fail. At this point the acylation and reduction of histamine **6.35** was halted as direct alkylation of the amides intermediates was yielding more promising results.

As previously mentioned, direct *N*-alkylation of phenyl acetamides such as **6.22** and related analogues was not attempted due to concerns of non-selective alkylation. However due to the difficulty in *N*-ethylation of the small mono-alkyl amines this pathway was investigated. Thiazole **6.21**, pyridine **6.24** and 1,2,4-triazole **6.32** analogues were selected to synthesise the desired *N*-ethyl analogues due

to their initial screening results as discussed below. Optimisation of alkylation conditions on these 3 analogues was investigated concurrently and included variation of solvent (THF, DMF), equivalents of NaH (2, 3, 6 eq.), ethyl iodide (2, 3, 6 eq.), reaction time (3 – 24 h) and temperature (0 – 40 °C). The equivalents of NaH most significantly impacted reaction progress, with reactions containing a significant excess of NaH (>5 eq.) gave higher conversion to the desired product by LCMS. Hence, using these optimised conditions a number of *N*-dialkyl amides were synthesised (**Scheme 6.4**). Alkylation of the thiazole amide **6.21** and pyridine amide **6.24** produced the desired amide products (thiazole **6.40** and pyridine **6.42**, respectively) as well as the amide and benzylic dialkylated by-products (thiazole **6.41** and pyridine **6.43**, respectively, **Scheme 6.4**). Under the same conditions triazole **6.32** gave only dialkylated by-product **6.39**, however selective monoalkylation product **6.38** was obtained by reducing the amount of ethyl iodide from 6 to 1.1 equivalents.

Scheme 6.4: Synthesis of a series of *N*-ethyl amides by selective alkylation chemistry.



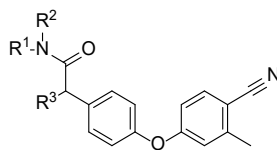
Reagents and conditions: (a) NaH, EtI, THF, 2 - 6 h, 10 – 75 %. Both mono and dialkyl products were isolated from the same reaction.

The by-products were isolated in high purity and were also screened alongside the desired *N*-ethyl analogues. The formation of these amide alkylated products was supported by the disappearance of the resonance corresponding to the amide proton at $\delta \sim 6.38$ ppm in the ^1H NMR spectrum and retention of the resonances corresponding to the two benzylic protons as rotameric singlets at $\delta \sim 3.70$ and 3.60 ppm respectively.

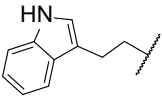
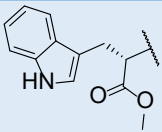
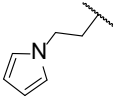
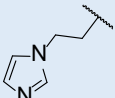
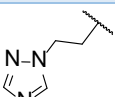
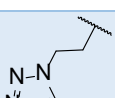
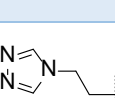
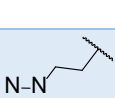
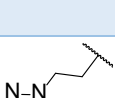
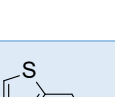
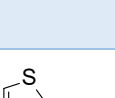
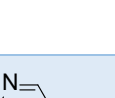
6.2.4 Testing of pyrazole analogues

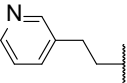
Heterocycle analogues were tested for solubility via 1D ^1H NMR and binding to *EcDsbA* via ^1H - ^{15}N HSQC as detailed previously. Removal of the *N*-ethyl functional group of the parent pyrazole **6.1** as in amide **6.22** caused a >4-fold loss in affinity, while removal of the ethyl pyrazole sidechain as in ethyl amide **6.20** resulted in a decrease in solubility, which suggested that the pyrazole functional group was important for binding. When the pyrazole was substituted for a heteroaromatic group with fewer hetero atoms (as in **6.28** – **6.30**) the solubility was reduced ($< 125\ \mu\text{M}$) and these analogues were not further screened. Exchanging the heteroaromatic system and removal of the *N*-ethyl group (as in **6.21**, **6.22**, **6.24**, **6.26**, **6.27** and **6.31** – **6.34**) gave analogues with similar or increased solubility, but reduced affinity compared to parent pyrazole **6.1**. Among the *N*-ethyl analogues only 1,2,4-triazole **6.39** which also bares an ethyl group on the benzylic position retained affinity ($K_{\text{D}} = 97 \pm 12\ \mu\text{M}$), while all other analogues had reduced affinity for *EcDsbA*. However, a bound structure of the 1,2,4-triazole **6.39** could not be obtained by soaking or co-crystallisation.

Inclusion of additional heteroatoms in the *N*-ethyl heteroaryl substituent were designed to probe interactions with Gln164, however no increase in binding affinity was observed across the series. It was therefore speculated that no new interactions with Gln164 were formed. Furthermore, X-ray crystallographic analysis of these expanded analogues gave no electron density for the ligands, which did not allow for detailed analysis of the binding mode.

Table 6.1: Testing results for a series of pyrazole analogues.

Compound number	R ¹	R ²	R ³	Solubility max (μM)	Single point HSQC NMR	HSQC NMR Affinity (K _D , μM)*	Electron Density for the ligand
6.1		Et	H	250	Binding	63 ± 9	Full
6.20	H	Et	H	125	< 0.04 CSP	-	-
6.21		H	H	250	Binding	250 ± 76	-
6.22		H	H	250	Binding	250 ± 45	-
6.23		Et	H	500	Binding	272 ± 15	-
6.24		H	H	250	Binding	163 ± 45	-
6.25		Et	H	<125	-	-	-
6.26		H	H	500	Binding	646 ± 110	-
6.27		H	H	1000	Binding	211 ± 13	-

6.28		H	H	<125	-	-	-
6.29		H	H	<125	-	-	-
6.30		H	H	<125	-	-	-
6.31		H	H	1000	Binding	635 ± 31	-
6.32		H	H	1000	Binding	564 ± 30	-
6.33		H	H	250	Binding	760 ± 470	-
6.34		H	H	1000	Binding	680 ± 43	-
6.38		Et	H	1000	Binding	440 ± 24	-
6.39		Et	Et	250	Binding	97 ± 12	-
6.40		Et	H	<125	-	-	-
6.41		Et	Et	<125	-	-	-
6.42		Et	H	125	Binding	240 ± 40	-

6.43		Et	Et	<125	-	-	-
------	---	----	----	------	---	---	---

*Affinities presented as mean \pm SEM.

6.2.5 Summary

A series of commercially available ethyl linked amino heterocycles were designed and synthesised to form a new H-bond interaction with Gln164. In total 23 analogues were made with 7 compounds failing solubility testing at the lowest concentration of 125 μ M. The importance of the *N*-ethyl amide of pyrazole **6.1** was investigated by synthesising a number of amides with a free NH and *N*-ethyl amides. Removal of the *N*-ethyl functional group of the parent pyrazole and other *N*-ethyl analogues reduced affinity > 4-fold for *EcDsbA*. The best analogue from this series was triazole **6.39** having both the *N*-ethyl and benzylic ethyl groups which displayed a comparable affinity to the *N*-ethyl pyrazole **6.1**. Benzylic alkylation as in **6.39** introduced a chiral centre, therefore isolation and testing of the individual enantiomers is an avenue for future work. The presence of both the pyrazole ring and *N*-ethyl amide were favoured for binding with all other modifications resulting in a loss in affinity. The investigation of selective amide alkylation also provide a synthetic pathway with a late stage diversity alkylation step which may be used for further exploration of this vector.

6.3 Alkylation of the pyrazole amide using REFiLx

6.3.1 Introduction, context and design

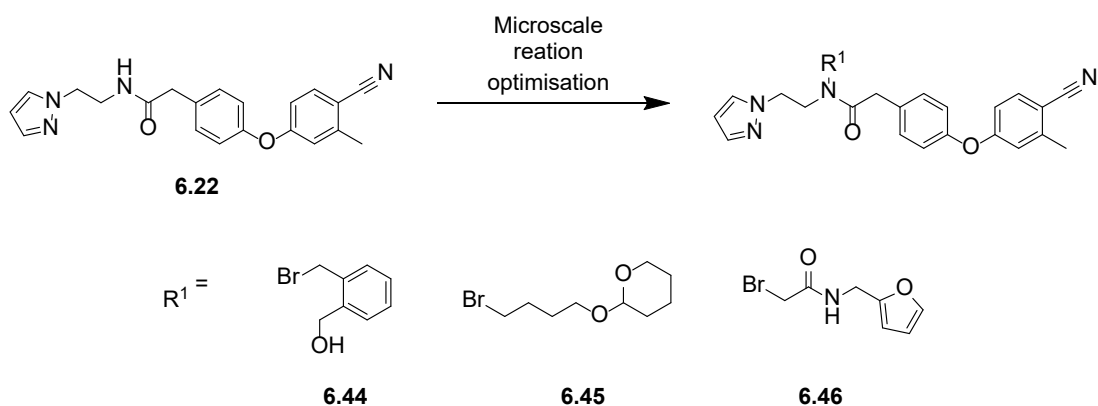
Expansion across the hydrophobic groove has been investigated by a number of chemistries (amidation, benzylic alkylation and acyl sulfonamides) using SBDD with few analogues demonstrating any improvements in affinity (see Chapter 4). Obtaining X-ray crystal structures of these analogues which expand across the groove also proved challenging. Analysis of the X-ray crystal structure of pyrazole **6.1** bound to *EcDsbA* (**Figure 6.1**) shows the *N*-ethyl function group projecting across the groove making only weak interactions with *EcDsbA*. However, the

corresponding amide lacking the *N*-ethyl group (as in **6.22**) was significantly weaker in affinity suggesting the ethyl may contribute to conformational restriction of the compound to improve its affinity. The further expansion of this *N*-ethyl group or substitution and formation of new interactions could provide compounds of higher affinity than ethyl pyrazole **6.1**. Access to these amide substituted analogues was envisaged through alkylation of pyrazole **6.22**. As previous SBDD had failed an additional REFiL_X workflow was proposed using an in-house alkyl halide reagent library. This reagent library was designed using the same principles as the amine reagent library requiring a balance of properties with maximal coverage of chemical space (described in Chapter 5).

6.3.2 Optimisation of plate and soaking conditions

Similar to the previous REFiL_X screen optimisation of microscale reaction conditions for this alkylation were required. Alkylation reaction conditions were optimised in a 96 well plate on microscale (100 μ L volume, 5 μ mol) varying solvent (THF, Et₂O, DMF), alkyl halide reactivity (**6.44**, **6.45**, **6.46**, **Scheme 6.5**) and equivalents of alkyl halide (1.1, 2, 4 eq.). The anion of amide **6.22** was formed under dry conditions in a single round bottom flask on larger scale (~100 mg) and dispensed into each well. LCMS of the reactions at 1, 2, 4 and 5.5 h showed minimal reaction progress in all tested conditions (maximum 20 % conversion). It was hypothesised that residual water in the reaction solvent/reagent stocks was quenching the amide anion and limiting reaction progress in the 96 well plate format. This quenching issue is often encountered when using strong bases and was further exacerbated by performing the reactions on microscale where only trace amounts of water could compromise the reaction. Hence, the optimisation on microscale was repeated with excess NaH (>10 eq.) scooped into each individual well of the plate and the reactions tracked by LCMS. Analysis of the LCMS data suggested the presence of excess NaH, 2 eq. of electrophile and 2 - 3 h reaction time were sufficient to observe some desired product. Furthermore, DMF was identified as the preferred solvent as Et₂O was too volatile to work with in such small quantities and THF formed a gel like substance that was difficult to work with.

Scheme 6.5: Optimisation of microscale reaction conditions for the alkylation of amide **6.22** with a number of electrophiles **6.44–6.46**.



Reagents and conditions: Alkyl bromide (**6.44 – 6.46**, 1.1, 2 or 4 eq.), NaH, DMF, THF or Et₂O, R.T., 5.5 h.

The use of excess NaH in the reaction mixture necessitated quenching of the reactions prior to soaking into crystals of *Ec*DsbA for REFiL_X. Standard quenching methods using ice water were deemed unsuitable since the excess production of hydroxide (OH⁻) could alter the pH of the crystallography soaking buffer and may denature the protein or dissolve the crystals. Consequently, alternate quenching conditions were tested for compatibility with subsequent screening by X-ray crystallography. AcOH and pH 7.0 phosphate buffer were proposed as viable alternatives. These alternatives were evaluated via test reactions which were setup using the optimised reaction conditions described above and quenched with double volume (200 μL) of either 50:50 H₂O:AcOH or 1 M pH 7.0 phosphate buffer. The reactions were evaporated and samples soaked into *Ec*DsbA crystals. The faux reactions were spiked with a known *Ec*DsbA binder ethyl pyrazole **6.1** at 10 mM and data collected at the synchrotron. AcOH quenching method was preferred and gave suitable X-ray crystallography results, whilst phosphate buffer quenching gave poor resolution crystallographic data.

6.3.3 REFiL_x screen and hit identification

The in-house library of 93 diverse alkyl halides were used to synthesise a library of pyrazole **6.1** analogues. Analysis by LCMS revealed poor conversion rates compared to those observed with the first amidation plate (**Figure 6.4**) despite microscale reaction optimisation. The average conversion for the alkylation plate was estimated at 6 % (compared to 24 % for the amidation) with 76 % of reactions giving no indication of product formation by LCMS. Although, these results suggest that few reactions were successful the poor results in reaction optimisation and a few alternatives lead us to progress these reactions to crystal soaking.

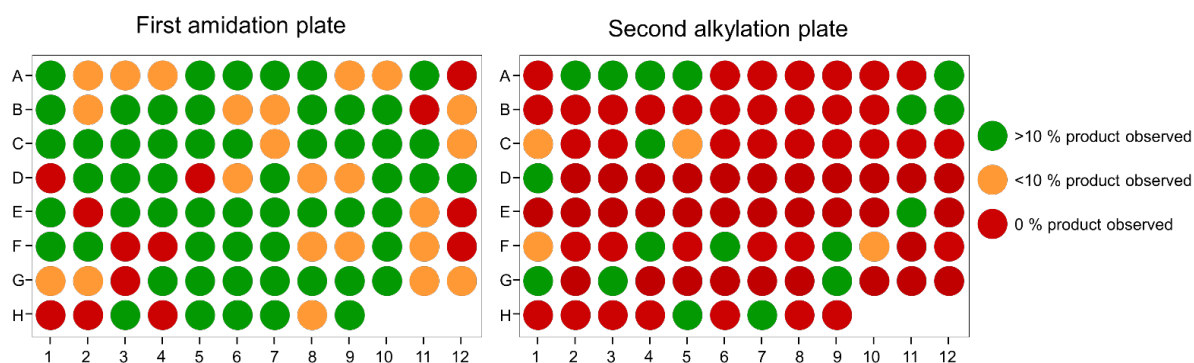


Figure 6.4: Reaction conversions for the amidation plate (left) described in Chapter 5 and the current alkylation plate (right) as estimated by LCMS. Wells are colour coded based on the estimated purity of product observed by LCMS, red = no product detected, orange < 10 % and green > 10 %.

Analysis of the crystallographic data revealed that very few wells displayed any electron density corresponding to a ligand and those which did had only partial electron density and none had density for the diverse alkyl extension. This was in contrast to the first screen, in which electron density was observed for the starting material in the majority of wells, and hits were identified where additional electron density for the extension could be observed in the data. Herein, we investigated wells where any electron density corresponding to a ligand was observed. Using this criteria four analogues **6.47** – **6.50** were resynthesised to acquire full characterisation data (**Figure 6.5**).

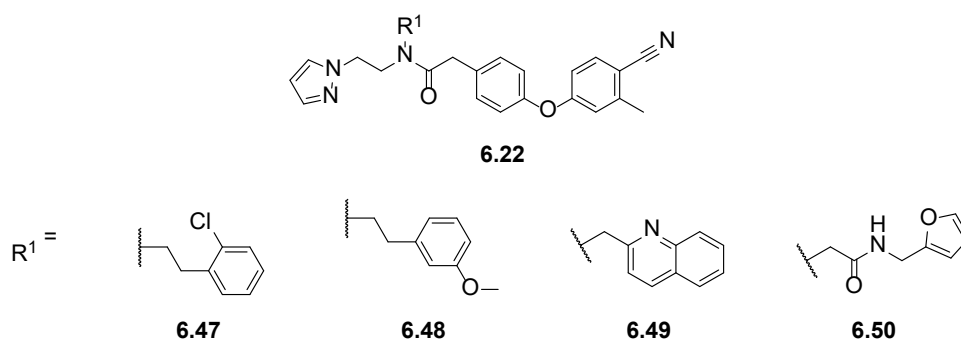
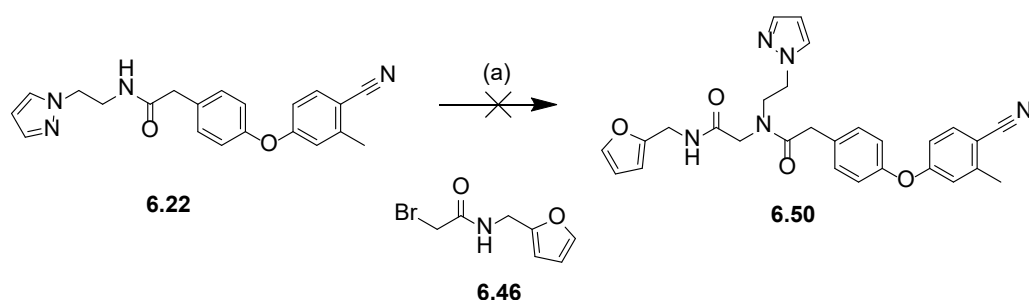


Figure 6.5: Hit compounds identified from the second REFILX screen.

6.3.4 Resynthesis of hits

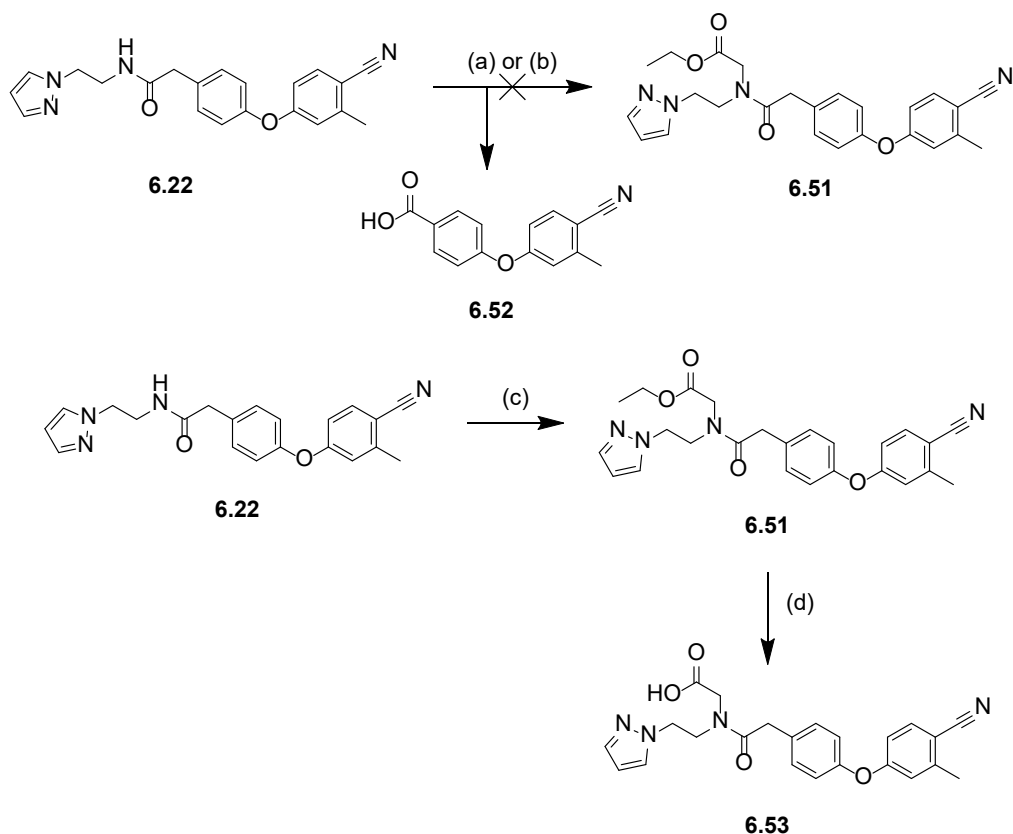
Initially, attempts were made to synthesise the identified hits on mg scale via similar alkylation reaction conditions as those used on microscale (see above). However, this did not yield the desired products and additional parallel optimisation was required to identify suitable reaction conditions for synthesis of each analogue. Described here is one of the optimisations followed by the final synthesis conditions of all analogues.

The alkylation of amide pyrazole **6.22** with furan alkyl bromide **6.46** (Scheme 6.6) showed minimal product formed by LCMS (m/z : 496.0 [M-H]⁻). Additionally, this alkylation reaction produced multiple by-products consistent with di and tri alkylated structures (m/z : 633.0 [M-H]⁻ and (m/z : 770.7 [M-H]⁻).

Scheme 6.6: Synthesis of furan **6.50** and proposed by-products.

Reagents and conditions: (a) Aryl bromide **6.46**, NaH, DMF, 0 °C – R.T, 5 h, 0 %.

Hence, in an effort to avoid over alkylation of furan **6.50** the alkyl halide was changed to ethyl chloroacetate (**Scheme 6.7**). Optimisation of time (2 – 72 h), temperature (0 – 70 °C), base (NaH, LDA, Cs₂CO₃) and solvent (DMF or THF) failed to show any reaction towards ester **6.51** with ethyl chloroacetate by either LCMS or ¹H NMR. The ethyl chloroacetate electrophile was then changed to the more reactive ethyl bromoacetate and the optimisation repeated. Addition of KI, 18-crown-6 ether, as well as increased reaction time (2 – 72 h), temperature (0 – 70 °C), use of strong base (NaH, LDA) gave no significant improvements in reaction conversion. A major by-product benzoic acid **6.52** was identified from the crude reaction by LCMS (*m/z*: 253.9 [M+H]⁺) and isolated by column chromatography. The formation of carboxylic acid **6.52** by-product was supported by the ¹H NMR spectrum which showed an electron withdrawn doublet at 8.14 ppm corresponding to the protons *ortho* to the carboxylic acid (**Appendix 6.1**). The benzoic acid **6.52** was identified as a major by-product in the synthesis of all 4 hits. It was hypothesised that the formation of this by-product may be mediated by the degradation of DMF in the presence of NaH as supported by the literature.³ The exact mechanism of this oxidation remains unknown however, upon repeating the reaction in THF, the ester intermediate **6.51** was successfully isolated and the benzoic acid by-product was not observed via LCMS or ¹H NMR. The REFIL_X alkylation library was synthesised using DMF and NaH which produced benzoic acid **6.52** as the major product during resynthesis. However, this by-product was not identified in by LCMS QC of the REFIL_X parallel library. Base mediated ester hydrolysis of ester **6.51** afforded carboxylic acid **6.53** as a screening compound.

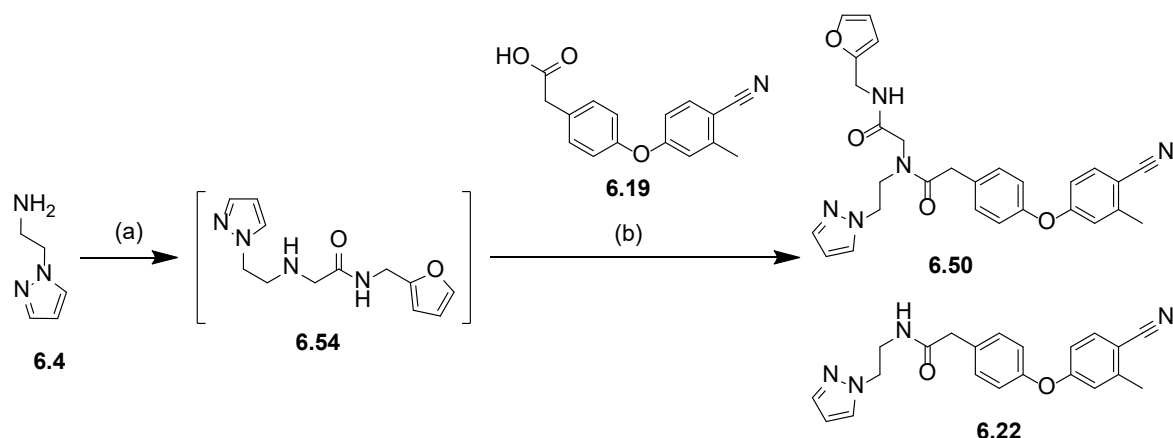
Scheme 6.7: Synthesis of carboxylic acid **6.53**.

Reagents and conditions: (a) ethyl chloroacetate, KI, NaH or LDA or Cs_2CO_3 , DMF or THF, 2 – 72 h, 0 – 70 °C. (b) ethyl bromoacetate, KI, 18-crown-6, NaH, DMF, 2 – 72 h, 0 – 70 °C. (c) ethyl bromoacetate, KI, 18-crown-6, NaH, THF, 4 d, reflux, 43 %. (d) NaOH, MeOH, H_2O , 4h, RT, 61 %.

An alternative approach to synthesise furan **6.50** was investigated in conjunction with the optimisation described above. We hypothesised that, alkylation of the amine **6.4** followed by amidation with phenylacetic acid **6.19** could afford furan **6.50** in a single step with any dialkylated by-products not participating in further amidation reactions (**Scheme 6.8**). Hence, this crude mixture of amines was directly coupled to carboxylic acid **6.19** via EDC mediated amide coupling and gave a mixture of amides **6.22** and the desired product **6.50** as determined by LCMS (m/z : 361.2 $[\text{M}+\text{H}]^+$ for **6.22** and m/z : 498.2 $[\text{M}+\text{H}]^+$ for **6.50**). This indicated that unconsumed starting material **6.4** was also present after alkylation and amide **6.22** was also isolated. The desired furan product **6.50** was then purified by flash column chromatography followed by preparative HPLC to acquire material of high purity for testing. The identity of furan **6.50** was confirmed by disappearance of the amide proton

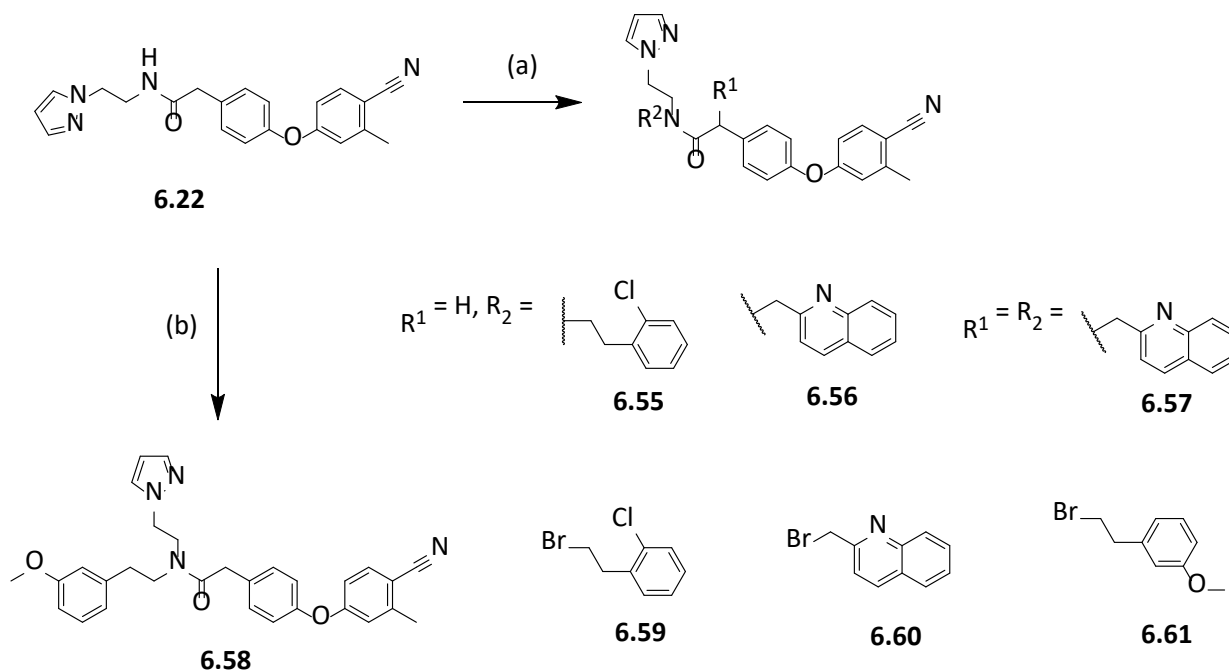
of **6.22** at δ 6.61 ppm and appearance of a number of electron withdrawn aliphatic protons (δ 2.9 – 4.5 ppm) integrating to 10 protons in the ^1H NMR spectrum.

Scheme 6.8: Synthesis of furan **6.50** via one pot alkylation and amidation reaction.



Reagents and conditions: (a) 2-bromo-N-(furan-2-ylmethyl)acetamide (**6.46**), Et_3N , DMF, 15 min, RT (b) EDC hydrochloride, HOBT, Et_3N , DMF, 0.25 - 4 h, RT, 29 %. Reactions (a) and (b) stirred for 15 min at RT in separate RBF's then combined and stirred for 4h.

As stated above the synthesis of all alkylation hits were optimised in parallel. The remaining four analogues (**6.55** - **6.58**) were isolated from varying alkylation reaction conditions of amide **6.22** with NaH/DMF or LDA/THF in low yields (**Scheme 6.9**). The poor yields may be attributed to the degradation of the starting material amide **6.22** to benzoic acid by-product **6.52**, however a sufficient quantity of desired compound was isolated for screening. During optimisation of the alkylation of amide **6.22** with bromoalkanes chlorophenyl **6.59**, quinoline **6.60** and methoxy phenyl **6.61** products alkylated at the benzylic position (**6.55**, **6.56**, **6.58**) and dialkylated at both the benzylic position and the amide (as in **6.57**) were isolated similar to previous alkylation chemistry. These selectively benzylic alkylation products were supported by the loss of the benzylic singlet at ~ 3 ppm and retention of the broad singlet amide proton at ~ 6.6 ppm in the ^1H NMR spectra.

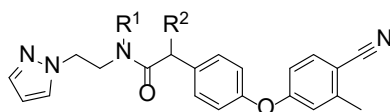
Scheme 6.9: Resynthesis of hits compounds via alkylation of the amide NH.

Reagents and conditions: (a) Alkyl bromide (**6.59** – **6.60**), NaH, DMF, 0 - 45 °C, 4 – 6 h, 1 – 40 %.

(b) 1-(2-bromoethyl)-3-methoxybenzene (**6.61**), LDA, THF, 6 h, RT, 5 %.

6.3.5 Affinity determination and solubility testing

The solubility of diaryl ether analogues was assessed by a series of 1D ^1H NMR titration experiments and with the exception of carboxylic acid **6.53** all analogues precipitated from solution at the lowest concentration tested of 125 μM (Table 6.2). Solubility was subsequently tested at lower concentrations of 80, 40, 20 and 10 μM by qNMR, however all compounds also failed at 10 μM .

Table 6.2: Solubility of alkylated products in aqueous buffer in the presence and absence of β -cyclodextrin. Affinities of REFIL_X hits using alkyl halide library of reagents.

Compound number	R ¹	R ²	Solubility max (μM)	Solubility with β -cyclodextrin (μM)	HSQC NMR Affinity (K _D , μM)*	Electron Density for the ligand
6.1	H	Et	250	500	86 ± 8 **	Full
6.50		H	< 10	150	Weak binder **	Not tested
6.51		H	< 10	200	200 ± 52 **	Not tested
6.53		H	1000	-	200 ± 7	Not tested
6.55	H		< 10	< 100	-	Not tested
6.56	H		< 10	< 100	-	Not tested
6.57			< 10	< 100	-	Not tested
6.58	H		< 10	< 100	-	Not tested

*Affinities presented as mean ± SEM. **Affinity determined with β -cyclodextrin.

β -cyclodextrin (CD) was therefore used in an attempt to increase their aqueous solubility and enable affinity assessment by ¹H-¹⁵N HSQC. CD is used in the pharmaceutical industry as a drug carrier system^{4, 5} to aid in solubilising highly lipophilic compounds in aqueous media for delivery. CD is a cyclic oligosaccharide that has a hydrophilic surface and hydrophobic interior where lipophilic small

molecules can bind commonly in a 1:1 binding complex that is highly water soluble. The solubility of *N*-ethyl pyrazole **6.1** as well as poorly soluble analogues **6.50**, **6.51** and **6.55** - **6.58** were tested at 100 μM with 1 equivalent of CD in solution. Compounds with observable signals in the 1D ^1H NMR spectra at the frequency of the free ligand were then repeated at higher concentration to more accurately estimate their maximum solubility in the presence of CD. The use of CD was successful in improving the aqueous solubility of furan **6.50** (max. 150 μM), ester **6.51** (max. 200 μM) and pyrazole **6.1** (max. 500 μM). These compounds were assessed for binding to *EcDsbA* via ^1H - ^{15}N HSQC spectra (Table 6.2).

Calculating binding affinity in this system is more complex than with simple titrations of *EcDsbA* with soluble ligand, since binding of the ligand to CD is significant. However, to provide an initial estimate of the binding affinity to *EcDsbA*, a simple single site 1:1 which does not account for binding of the ligand to CD was used to estimate affinities. The control compound pyrazole **6.1** demonstrated a similar affinity in the presence and absence of CD suggesting that binding to CD is negligible or did not affect its affinity for *EcDsbA*. Furan analogue **6.50** displayed only small CSP's and when titrated against *EcDsbA* and only a linear correlation of CSP's to ligand concentration suggesting weak binding. Ester **6.51** displayed moderate binding affinity ($K_D = 200 \pm 51 \mu\text{M}$) with the measured K_D value being close to the maximal solubility for this compound.

6.3.6 Summary

Alkylation off the amide of pyrazole **6.22** was explored using the REFIL_X workflow with microscale reactions and a quenching method optimised for X-ray soaking. The library however, gave poor conversion as assessed by LCMS and only four reactions gave crystal structures with partial electron density. The resynthesised compounds had low aqueous solubility and the affinities could only be estimated in the presence of CD for a subset of the analogues which were found to be of lower affinity than pyrazole **6.1**. The failure to identify hits of higher affinity could be due to the low reaction conversions as estimated by LCMS. It was hypothesised that the combination of DMF and NaH in

the reaction led to by-product formation as was observed when resynthesising the hits at mg scale. This significantly reduced the number of reactions with observable desired product and consequently limited the ability to find higher affinity analogues via the REFIL_X workflow.

6.4 Conclusion

This work focused on further elucidating SAR around pyrazole **6.1** and expansion through an alkylation expansion REFIL_X library. Substitution of the pyrazole for other heteroaromatic rings assessed possible H-bond interactions with Gln164. A number of amide *N*-ethyl and free NH matched pair analogues were also synthesised to assess the importance of the *N*-ethyl functional group. No analogues in these two series showed appreciable improvements in binding affinity, hence elaboration of pyrazole **6.1** was investigated. The *N*-ethyl vector of pyrazole **6.1** was then explored via alkylation chemistry using the REFIL_X workflow. Despite optimising the microscale reaction conditions LCMS estimates of reaction conversion for the library were low. In total four hits were identified based on partial electron density being observed in the crystallographic data and all hits were resynthesised for solubility and affinity determination. Further investigation during resynthesis identified the combination of DMF and NaH as the likely cause of low conversions in the microscale parallel synthesis library. These hits demonstrated low aqueous solubility and were tested against *EcDsbA* in the presence of cyclodextrin. Some hits demonstrated improved solubility with CD but had reduced affinity for *EcDsbA* compared to pyrazole **6.1**.

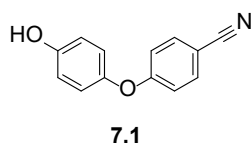
6.5 References:

1. Bordwell, F. G.; Harrelson, J. A.; Lynch, T. Y. Homolytic bond dissociation energies for the cleavage of .alpha.-nitrogen-hydrogen bonds in carboxamides, sulfonamides, and their derivatives. The question of synergism in nitrogen-centered radicals. *J. Org. Chem.* **1990**, 55, 3337-3341.
2. Bordwell, F. G.; Fried, H. E. Acidities of the hydrogen-carbon protons in carboxylic esters, amides, and nitriles. *J. Org. Chem.* **1981**, 46, 4327-4331.
3. Hesek, D.; Lee, M.; Noll, B. C.; Fisher, J. F.; Mobashery, S. Complications from dual roles of sodium hydride as a base and as a reducing agent. *J. Org. Chem.* **2009**, 74, 2567-2570.
4. Uekama, K.; Hirayama, F.; Irie, T. Cyclodextrin drug carrier systems. *Chem. Rev.* **1998**, 98, 2045-2076.
5. Gidwani, B.; Vyas, A. A comprehensive review on cyclodextrin-based carriers for delivery of chemotherapeutic cytotoxic anticancer drugs. *BioMed Res. Int.* **2015**, 2015, 198268-198268.

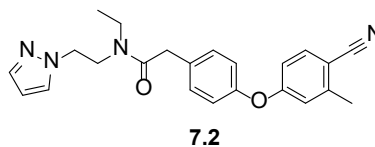
Chapter 7 – Thesis outcomes and future prospects

The principal aim of this research was to develop the diaryl ether fragments into potent *Ec*DsbA inhibitors through rational drug design. Progress was made towards higher affinity inhibitors however further improvements are still required.

The diaryl ethers were identified from a fragment screen against *Ec*DsbA. Analogue purchasing was conducted, which found diaryl ether **7.1** with an affinity of $K_D \approx 1$ mM (**Figure 7.1**), which bound in a hydrophobic groove on the surface of *Ec*DsbA. Throughout the course of my PhD the diaryl ethers were optimised using SBDD and REFIL_X which achieved a > 15-fold improvement in affinity with pyrazole **7.2** while maintaining moderate aqueous solubility (**Figure 7.1**).



$K_D \approx 1$ mM, solubility 500 μ M



$K_D = 60 \pm 9$ μ M, solubility = 250 μ M

Figure 7.1: Structure of parent diaryl ether **7.1** and best compound pyrazole **7.2**.

Early work focused on core optimisation of the nitrile ring of **7.1** with the goal of filling the hydrophobic RHS of the groove. This design strategy aimed to form a consistent binding conformation of the fragment to the RHS of the groove, hypothesised to then enable expansion across the more open LHS of the groove. The first two series of analogues investigated additional substituents on benzonitrile ring and substitution of the *para*-cyano group with a number of H-bond donors, acceptors, hydrophobic and fused ring systems (**Figure 7.2**). Generating clear SAR around modification of the nitrile ring was challenging as analogues in this series displayed limited solubility. However, numerous X-ray crystal structures were acquired from this series and were used to further medicinal chemistry efforts. Substitution proximal to the cyano group appeared to be the most favourable modification and the 2-methyl analogue **7.3** was chosen for further elaboration. Modification of the cyano substituent caused many analogues to change binding mode. However, X-ray crystallography identified the binding mode of *para*-phenylacetic acid **7.4** with a new H-bond

interaction with His32. This structure guided the design of a series of phenol replaced analogues which were described in Chapter 3. A series of bicyclic diaryl ethers aimed to fuse the 2-methyl and H-bond acceptor on the benzonitrile ring, however these bicycles demonstrated poor aqueous solubility and weak binding to *EcDsbA*. Constant monitoring of compound solubility was essential in acquiring reliable affinity data for the diaryl ether analogues. Attempts to improve solubility generally resulted in losses in affinity as demonstrated by the replacement of the *para*-cyano group. The 2-methyl benzonitrile **7.3** maintained binding mode, improved affinity and was selected for further elaboration.

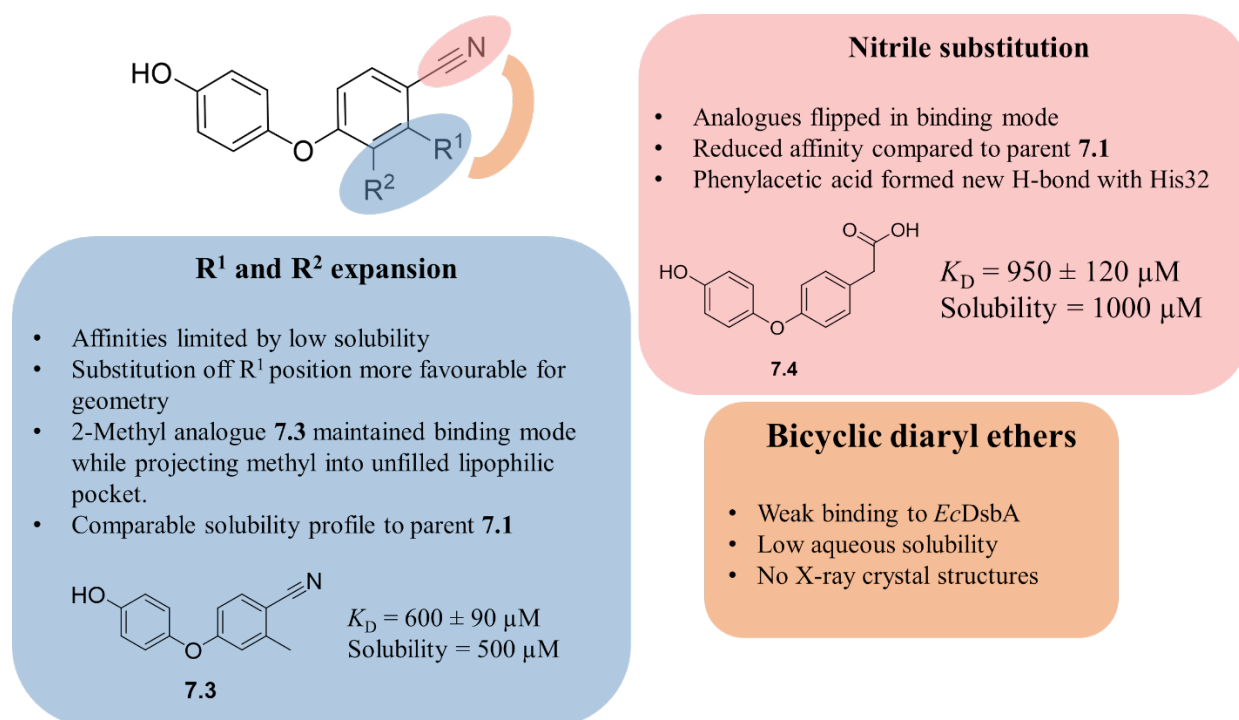


Figure 7.2: Summary of SAR generated and key analogues when exploring the RHS of the groove in Chapter 2.

A second generation of diaryl ethers were developed by merging of the best analogues elucidated in Chapter 2, as well as continued optimisation of binding to the RHS of the groove by reassessing SAR around nitrile ring substituents. In addition, replacing the ether linker to also enable expansion into the bottom groove was investigated (**Figure 7.3**). Merging of the 2-methyl benzonitrile **7.3** and *para*-

phenylacetic acid **7.4** led to the discovery of phenylacetic acid **7.5** which formed a new H-bond interaction with His32 and, had improved solubility and comparable affinity to the phenol analogue **7.3**. Compounds with higher aqueous solubility ($\geq 500 \mu\text{M}$) were prioritised for further development to maintain desirable physicochemical properties for development. The improved solubility of **7.5** enabled further exploration of the SAR around the benzonitrile ring. Substitution of the 2-methyl functional group with a bulkier *tert*-butyl group proximal to the nitrile in **7.6** increased affinity 2-fold however, **7.6** lacked structural data for further elaboration. Isosteric replacement of the ether linker for a methyl amino linker and expansion into the bottom groove was not tolerated. Having synthesised and tested > 60 analogues exploring and optimising the RHS of the binding site the next step was expansion across the LHS of the groove with the 2-methyl phenylacetic acid **7.5**.

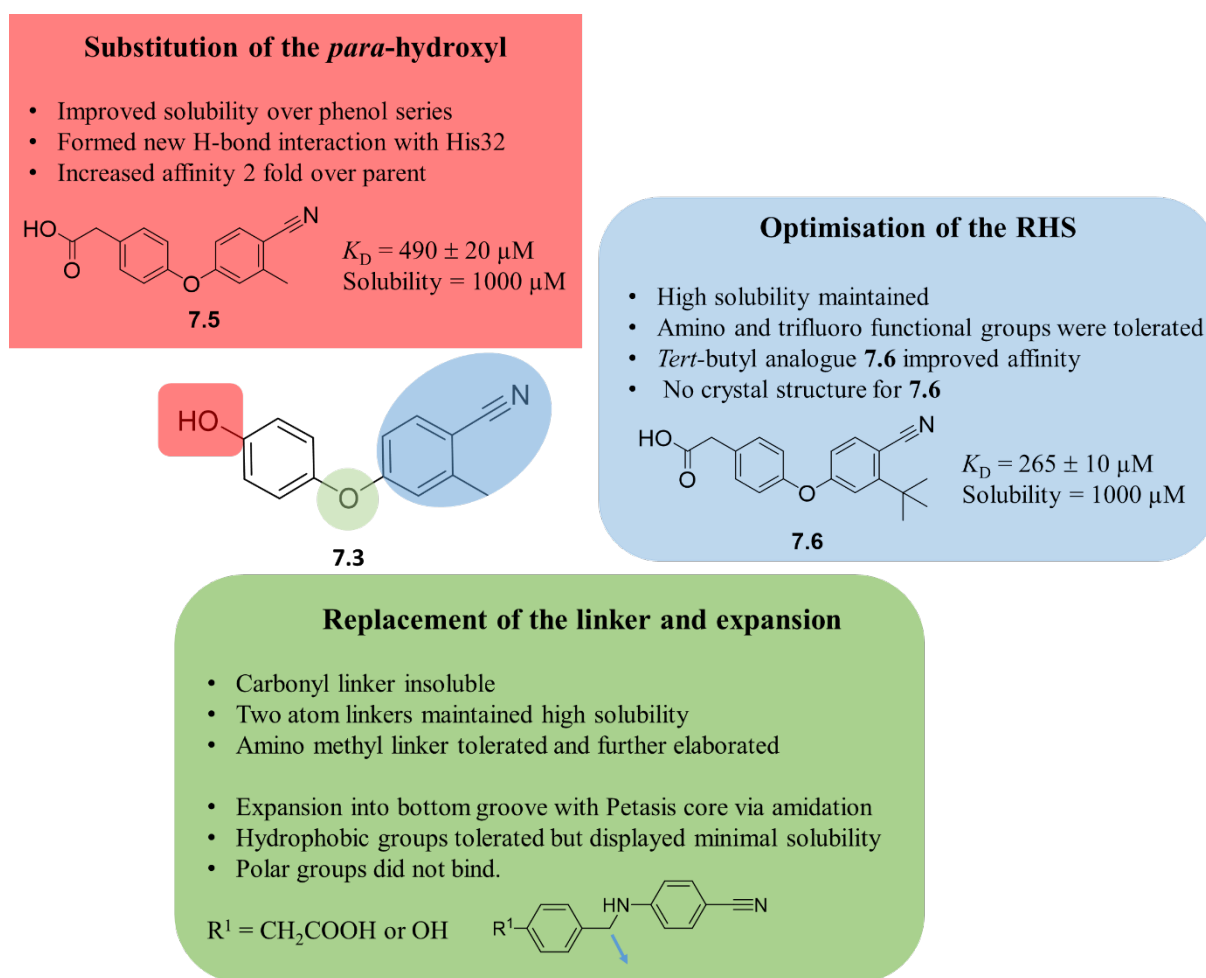


Figure 7.3: Summary of SAR generated and key analogues when exploring the RHS and bottom vector of the groove in Chapter 3.

Maintaining high aqueous solubility relative to affinity remained at the forefront of our design strategy, hence polar functional groups that would explore the LHS of the groove were prioritised. These expansions primarily aimed at making polar contacts to Arg148 and Pro151 through the amidation of carboxylic acid of **7.5** and expansion at the benzylic position. Replacement of the carboxylic acid with a series of acyl sulfonamides was also explored (**Figure 7.4**) to maintain a negative charge while also expanding to the LHS of the groove. Expansion via amidation and benzylic substitution failed to yield the desired ligand-protein interactions and generated few analogues with improved affinity. Additionally, the limited number of X-ray crystal structures obtained, hampered our ability to guide medicinal chemistry efforts. Isosteric replacement of the carboxylic acid with an expanded acyl sulfonamide improved affinity of the core, however the acyl sulfonamides had flat SAR and no new structural information. To this point in the exploration of the diaryl ether series over 90 analogues were synthesised, characterised and tested against *EcDsbA* using SBDD. A new H-bond was formed and multiple vectors explored, however few analogues had measured affinities with $K_D < 200 \mu\text{M}$. The potentially higher affinity analogues were hampered by poor solubility or lack of structural information by X-ray crystallography. *EcDsbA*'s binding site is a shallow hydrophobic groove with minimal opportunity for polar interactions and, development of inhibitors using traditional SBDD was challenging and failed to provide significant improvement in affinity (K_D). Hence, we investigated an alternate workflow for efficiently elaborating fragments into leads using parallel synthesis and screening via X-ray crystallography.

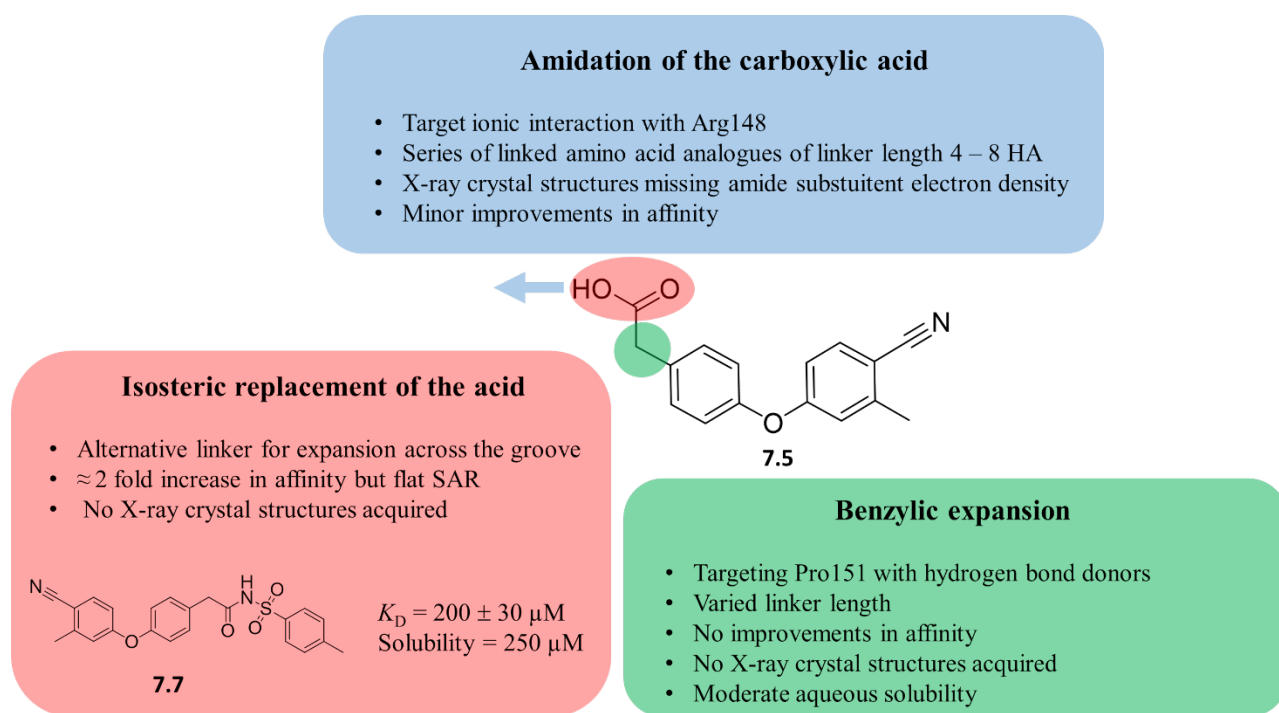


Figure 7.4: Summary of design, synthesised analogues and SAR generated in Chapter 4 to explore expansion into the LHS of the groove of *EcDsbA*.

A parallel chemistry and screening approach was implemented to more rapidly identify higher affinity compounds by circumventing extensive purification and characterisation of analogues which is time consuming. Methods such as ORS by SPR¹ circumvent compound purification and are able to test unpurified reaction products and identifying higher affinity analogues by examining specific kinetic binding properties. However, ORS by SPR could not be implemented for the diaryl ether series with *EcDsbA*, hence we investigated a complementary approach termed REFIL_X that uses X-ray crystallography to identify higher affinity compounds from a library of unpurified reaction mixtures (**Figure 7.5**). Our novel REFIL_X screening method identified 4 hit compounds, the best of which was the ethyl pyrazole **7.2** with a notable 8-fold improvement in affinity ($K_D = 62 \pm 9 \mu\text{M}$, Solubility = $250 \mu\text{M}$) over the phenylacetic acid **7.5**. Elucidation of its binding mode via X-ray crystallography demonstrated an unexpected binding mode and this analogue represents a significant improvement in binding affinity for *EcDsbA*. Having identified a higher affinity analogue, the next step was to optimise the interactions formed between pyrazole **7.2** and *EcDsbA*.

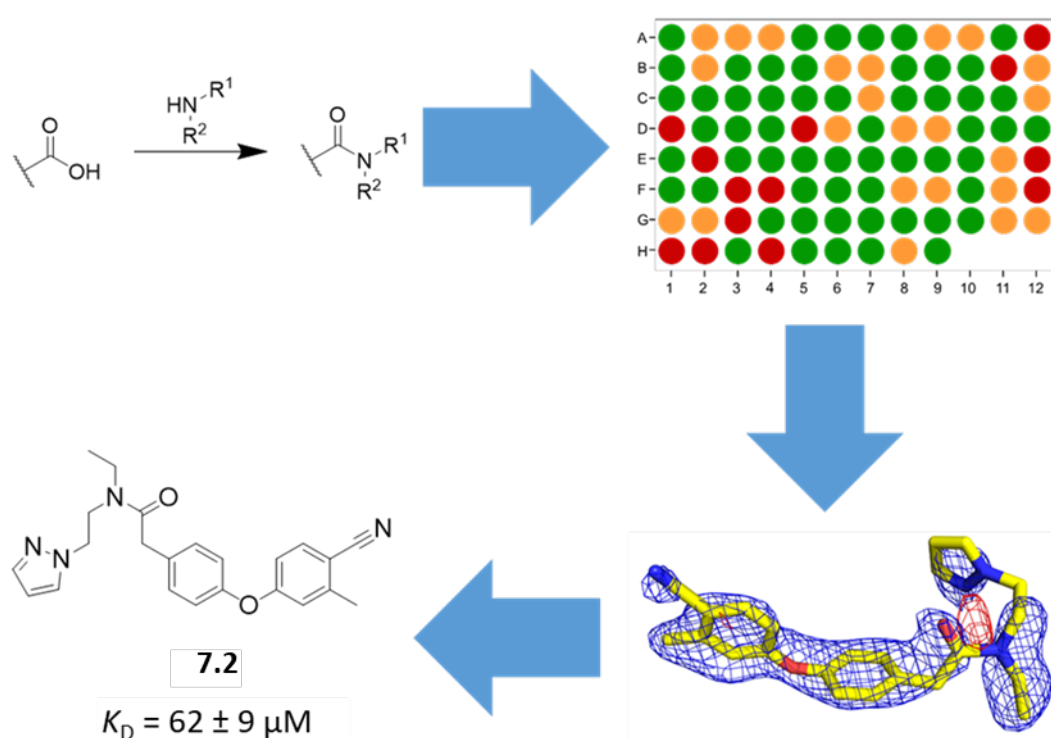


Figure 7.5: Summary of REFILX workflow. Amidation reaction was performed in parallel on microscale with a diverse amine reagent library, assessed by LCMS and soaked into crystals of *EcDsbA*. Hits were identified, resynthesised and gave pyrazole 7.2.

Having identified a higher affinity ethyl pyrazole amide substituent from a diverse library screen we set out to investigate and optimise the substituent. A series of heteroaromatic ring replacements of the pyrazole ring as well as *N*-ethyl amide vs NH amide analogues were synthesised to further optimise the ethyl pyrazole compound. Structural modifications to the pyrazole ring saw a notable loss in binding affinity for *EcDsbA* with the exception of the 1,2,4-triazole 7.8 which demonstrated comparable affinity to pyrazole 7.2 (**Figure 7.6**). Removal of the *N*-ethyl functional group reduced affinity compared to pyrazole 7.2 and was similar or reduced for all other analogues tested. The *N*-ethyl group was then further explored using alkylation chemistry in a second REFILX workflow. In contrast to the work in Chapter 5 this parallel synthesis library had low conversion and few reactions displayed density in the crystallography screen. All resynthesised analogues in this series failed solubility at 10 μM in standard *EcDsbA* screening buffer and as a result affinities could only be

estimated for two analogues in the presence of cyclodextrin (CD) and demonstrated lower affinity than the parent pyrazole **7.2**. LCMS QC and subsequent investigation during resynthesis suggested that despite optimisation of the microscale chemistry the presence of DMF and excess NaH in the parallel library synthesis may have caused significant degradation of starting materials and products limiting the number of successful reactions.

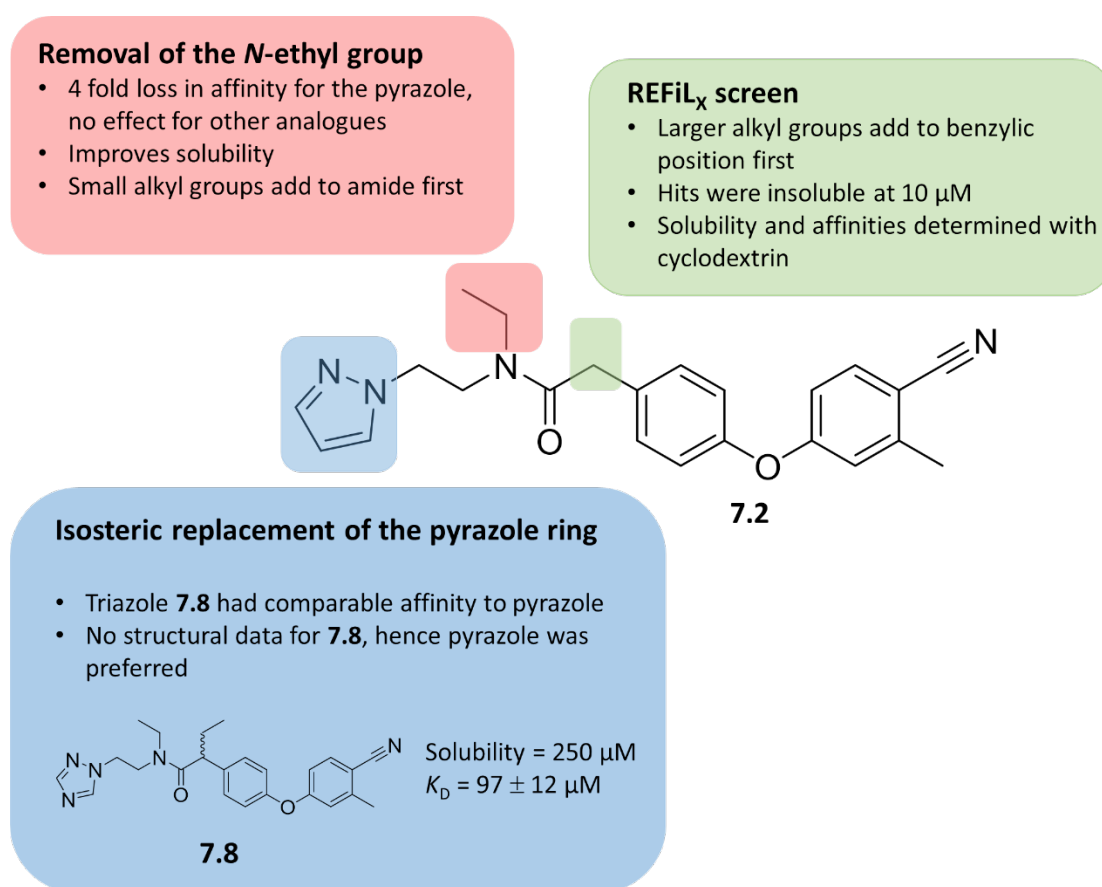


Figure 7.6: Summary of synthesised analogues and SAR generated around pyrazole **7.2** in Chapter 6.

Future prospects

Ethyl pyrazole **7.2** is the highest affinity small molecule *EcDsbA* identified to date and further optimisation of this core could afford a lead-like compound. Further elaboration and optimisation could be achieved by repeating the REFiL_x alkylation screen with THF to prevent the degradation of the amide starting material, improve reaction conversions and may enable a better understanding of

SAR around this vector. Additionally, altering the ethyl linker from the amide to the pyrazole ring was not explored and could be investigated to improve affinity. Optimisation of the pyrazole ring identified the chiral 1,2,4-triazole **7.8** which had comparable affinity to the ethyl pyrazole **7.2**. Hence, isolation and testing of its individual enantiomers could also yield slight improvements in affinity. Introduction of the *tert*-butyl functional group on the nitrile ring in **7.9** improved affinity 2-fold over the methyl substituent (**Figure 7.7**) and may provide additional affinity when incorporated into higher affinity more elaborated analogues **7.2** and **7.8**. Replacement of the *N*-ethyl sidechain of **7.2** with an acetic acid as in **7.10** was tolerated losing only 4-fold in affinity but provides the opportunity for further elaboration via amidation chemistry using the diverse amine reagent library and the REFIL_X or ORS by SPR workflows.

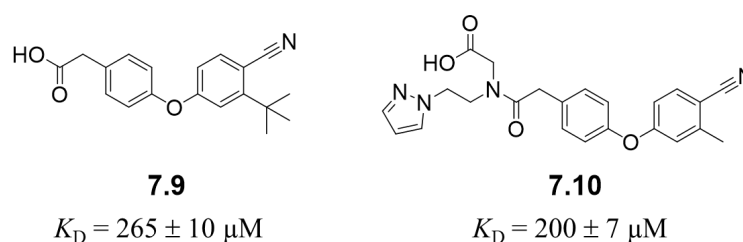


Figure 7.7: Structures and affinities of *tert*-butyl phenylacetic acid **7.9** and pyrazole carboxylic acid **7.10**.

The parent diaryl ether **7.1** was identified from a fragment screen and bound to *Ec*DsbA with an affinity of $K_D \approx 1 \text{ mM}$. Through SBDD and utilising the novel method of REFIL_X the diaryl ethers were improved by > 15 -fold in affinity while maintaining moderate aqueous solubility. The difficulties of working on *Ec*DsbA were highlighted throughout and further development would require close monitoring of aqueous solubility while exploring via SBDD and the use of efficient development workflows such as REFIL_X to identify improved affinity analogues.

References

1. Murray, J. B.; Roughley, S. D.; Matassova, N.; Brough, P. A. Off-rate screening (ORS) by surface plasmon resonance. An efficient method to kinetically sample hit to lead chemical space from unpurified reaction products. *J. Med. Chem.* **2014**, 57, 2845-2850.

Chapter 8 –Experimental

8.1 General experimental

^1H , ^{13}C and ^{19}F NMR spectra were acquired at 400 and 100 and 377 MHz respectively on an Avance III Nanobay Brüker spectrometer coupled to a BACS 60 automatic sample changer. Data were processed using MestReNova v6.0.2 software. ^1H and ^{13}C spectra were referenced with the residual proteo-solvent peak. ^1H NMR shifts: 7.26 ppm CDCl_3 , 2.50 ppm $(\text{CD}_3)_2\text{SO}$, 2.05 ppm $(\text{CD}_3)_2\text{CO}$, 1.94 ppm CD_3CN , 3.31 ppm CD_3OD and 0.030 ppm for DSS in D_2O . ^{13}C NMR shifts: 77.16 ppm CDCl_3 , 39.52 ppm $(\text{CD}_3)_2\text{SO}$, 1.32 and 118.26 ppm CD_3CN and 49.00 CD_3OD .¹ ^{19}F NMR were referenced from the ^1H NMR spectra. The data for all spectra are reported in the following format: (multiplicity, coupling constant J (Hz), integral, assignment), where multiplicity is defined as; s = singlet, d = doublet, t = triplet, q = quartet, p = pentet and m = multiplet. Subsequent abbreviations also include app = apparent and b = broad. For all ^{13}C Attached proton test (APT) spectra (C) = quaternary carbon, (CH) = methine carbon, (CH_2) = methylene carbon, (CH_3) = methyl carbon, and $(\text{C}=\text{O})$ = carbonyl carbon. Overlapped carbons were confirmed by HSQC or HMBC where possible. Rotameric compounds had all ^1H NMR and ^{13}C NMR resonances reported and their rotameric ratio specified.

High resolution mass spectrometry (HRMS) were acquired on an Agilent 6224 TOF LC/MS Mass Spectrometer coupled to an Agilent 1290 Infinity or a H₂O_s LCT TOF LC/MS Mass Spectrometer coupled to a 2795 Alliance Separations module. For the Agilent 6224 TOF LC/MS data were mass corrected via a dual-spray electrospray ionisation (ESI) source with the following *parameters*: Electrospray Ionisation, Drying gas-flow: 11 L/ min; Nebuliser: 45 psi; Drying gas temperature: 325°C; Capillary Voltage (Vcap): 4000 V; Fragmentor: 160 V; Skimmer: 65 V; OCT RFV: 750 V; Scan range acquired: 100–1500 m/z . Internal Reference ions: Positive Ion Mode = m/z = 121.050873 and 922.009798. Acquisition and analysis were performed using Agilent Mass Hunter Data Acquisition software vB.05.00 Build 5.0.5042.2 and Mass Hunter Qualitative Analysis vB.05.00 Build 5.0.519.13, respectively. For the H₂O_s LCT TOF LC/MS data were acquired and mass corrected via a dual-spray Leucine Enkephaline reference sample with the following *parameters*:

Electrospray Ionisation Desolvation gas flow: 550 L/Hr; Desolvation temperature: 250 °C; Source temperature: 110 °C; Capillary Voltage: 2400 V; Sample cone voltage: 60 V; Scan range acquired: 100–1500 m/z Scan time: 1 s Internal Reference ions: Positive Ion Mode = m/z = 556.2771. Acquisition and analysis were performed using Masslynx software v4.1. All Cl and Br containing compounds mass spectra are quoted as highest abundant isotope ^{35}Cl and ^{79}Br .

Liquid chromatography mass spectroscopy (LCMS) was detected on an Agilent 6100 Series Single Quad LC/MS (equipped with a 1200 Series G1311A Quaternary Pump, G1329A Thermostatted Autosampler, and a G1314B Variable Wavelength Detector) and the data was processed using LC/MSD Chemstation Rev.B.04.01 SP1 coupled with Easy Access Software. The system was equipped with a Reverse Phase Luna C8(2) (5 μm , 50 \times 4.6 mm, 100 Å) column maintained at 30 °C. All data was acquired and reference mass corrected via dual-spray electrospray ionisation (ESI) source. Each scan or data point on the total ion chromatogram (TIC) is average of 13700 transients, producing one spectrum per second. Mass spectra were created by averaging the scans across each peak and background subtracted against the first 10 sec of the TIC. Data acquisition was carried out using the Agilent Mass Hunter Data Acquisition software version B.05.00 Build 5.0.5042 and analysis was performed using Mass Hunter Qualitative Analysis version B.05.00 Build 5.0.519.13. LCMS standard method was a gradient of 5 % buffer B to 100 % buffer B over 6 min at 0.5 mL/ min. Buffer A = 99.9 % H_2O , 0.1 % TFA. Buffer B = 99.9 % MeCN, 0.1 % TFA.

The purity and retention time of final products were determined on an Analytical Reverse-Phase HPLC system fitted with an Agilent Zorbex SB-C18 Rapid resolution HT 2.1 mm \times 50 mm, 1.8 μm column using the specified gradient of buffer A: 99.9 % H_2O , 0.1 % TFA and buffer B: 99.9 % MeCN, 0.1 % TFA. Hydrophobic method 95 % A and 5 % B over 0.1 minutes, a gradient of 80 % to 100 % buffer B gradient over 9 minutes and 100 % buffer B over 2 minutes at a flow rate of 1 mL/ min. Standard method 95 % A and 5 % B over 0.1 minutes, a gradient of 5 % to 100 % buffer B gradient over 9 minutes and 100 % buffer B over 2 minutes at a flow rate of 1 mL/ min monitored at both 214 and 254 nm using a H₂O 996 Photodiode Array Detector. The enantiomeric excess (*ee*) of

enantiopure analogues was determined on an Agilent analytical normal phase HPLC. Equip with a lux 5 μ amylose 2 (5 μ m, 150 \times 4.6 mm) column. The standard method was a dual solvent isocratic method using buffer a 99.9 % EtOH, 0.1 % TFA and buffer b 100 % pet spirits at a flow rate of 2 mL per min, with a total run time of 20 min.

Thin layer chromatography (TLC) was carried out routinely on silica gel 60 F₂₅₄ pre-coated plates (0.25 mm, Merck). TLC plates were visualised under UV illumination at 254 nm and/or with the aid of bromocresol green, 2,4-dinitrophenylhydrazine (2,4-*DNP*), vanillin, ferric chloride, iodine or potassium permanganate solutions. Column chromatography was conducted using Davisil silica gel LC60A (40-63 μ m) or Buchi reveleris X2 flash instrument equip with ELSD and UV detection at 254, 265 and 280 nm on reveleris HP silica (20 μ m) cartridges 4 – 40 g according to manufacturer instructions. Dry solvents MeCN, THF, EtOH, DCM and DMF were dried using an Mbraun solvent purification system (SPS-800) according to the manufacturer's instructions. NMR solvents were stored over Acros 3 Å, 8 - 12 mesh molecular sieves. Cooling baths at -20 °C are 1:3:1 mixture of NaCl, ice and H₂O; -78 °C baths are a mixture of dry ice and acetone. Solvents were degassed by bubbling high purity N₂ gas through the solvent for 5 min. All other solvents were reagent grade and used as required.

8.2 General procedures

General procedure 1: Ullmann coupling

Aryl halide (1 eq.), phenol (1.5 eq.), Cs₂CO₃ (2 eq.), *N,N*-dimethylglycine hydrochloride (0.2 eq.) and copper(I) iodide (0.1 eq.) were sealed in an oven dried crimp sealed vial and purged with N₂ (3 times). Degassed 1,4-dioxane was added and the reaction mixture heated at 120 °C for 18 – 96 h. * **Note the reaction was conducted above standard pressure and caution should be used with suitable protective equipment.** The reaction was then cooled and diluted with H₂O and EtOAc. The layers were separated and the aqueous layer further extracted with EtOAc. The combined organic layers

were washed with NaOH (aq., 1 M) thrice, brine and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

General procedure 2: Ether deprotection or demethylation

To a solution of methyl ether (1 eq.) dissolved in dry DCM at 0 °C was added BBr₃ in heptane (3 – 5 eq.) dropwise. The reaction was stirred for 0.5 h at 0 °C, then allowed to warm to room temperature slowly and stirred for an additional 2 - 4 h. The reaction mixture was then quenched with phosphate buffer (aq., 1 M, pH 7) and diluted with DCM. The layers were separated and the aqueous layer further extracted with DCM. The combined organic layers were washed with brine and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

General procedure 3: Dess-Martin oxidation

To a solution of alcohol (1 eq.) dissolved in dry DCM was added Dess-Martin periodinane (1.5 eq.) at 0 °C. The reaction mixture was stirred for 0.5 h at 0 °C then allowed to warm to room temperature and stirred for an additional 18 h. After this time, the reaction was quenched with a 1:1 mixture of Na₂S₂O₃ (aq., sat.) and NaHCO₃ (aq., sat.) and diluted with DCM. The layers were separated and the aqueous layer further extracted with DCM twice. The combined organic layers were washed with brine and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

General procedure 4: Amide coupling reaction with peptide coupling reagents

To a solution of carboxylic acid (1 eq.), HCTU or HATU (1 eq.) and amine (2 eq.) dissolved in dry MeCN or DMF was added DIPEA or Et₃N (3 eq.) dropwise under N₂ at room temperature. The reaction mixture was then stirred for 18 - 72 h under N₂. After this time, the solvent was removed under vacuum and the residue dissolved in EtOAc and washed with hydrochloride (aq., 1 M), NaHCO₃ (aq., sat.), brine and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

General procedure 5: Ester hydrolysis

Ester (1 eq.) and LiOH (5 eq.) were dissolved in a 1:1 mixture of THF and H₂O and stirred for 18 h at room temperature. After this time, the THF was removed under vacuum and the aqueous layer acidified with hydrochloride (aq., 2 M) and diluted in EtOAc. The layers were separated and the

aqueous layer further extracted with EtOAc twice. The combined organic layers were washed with brine and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

General procedure 6: Ullmann coupling 2

Aryl halide (1.00 eq.), phenol (1.5 eq.), Cs₂CO₃ (2 eq.), *N,N*-dimethylglycine hydrochloride (0.4 eq.) and copper(I) iodide (0.2 eq.) were sealed in a crimp sealed 25 mL vial and purged with N₂ (3 times). Degassed DMF was added and the reaction mixture heated at 100 °C for 20 h. The reaction was then cooled and diluted with H₂O and Et₂O. The layers were separated and the aqueous layer further extracted with Et₂O twice. The combined organic layers were washed with H₂O thrice, brine thrice and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

General procedure 7: Ester hydrolysis with hydroxide

To a solution of ester (1 eq) dissolved in a 2:1 mixture of MeOH and NaOH (aq., 2 M) at room temperature. The reaction was stirred overnight at room temperature. After this time, the MeOH was removed under vacuum and the aqueous layer acidified with hydrochloride (aq., 2 M) and diluted in EtOAc. The layers were separated and the aqueous layer further extracted with EtOAc twice. The combined organic layers were washed with acidic brine (pH 0 - 1) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

General procedure 8: Acid chloride formation

Oxalyl chloride (3 eq.) was dissolved in dry DCM. DMF (cat. 2 - 3 drops) was added at 0 °C and the reaction mixture stirred for 30 min under N₂. After this time, carboxylic acid (1 eq.) was added and the reaction stirred for a further 4 - 6 h at 0 °C. The reaction mixture was evaporated to dryness and the residue dissolved in dry DMF. Amine (2 – 5 eq.) and DIPEA (3 - 5 eq.) were added and the reaction mixture was stirred at room temperature for 24 - 72 h. The reaction mixture was then diluted in EtOAc and washed with hydrochloride (aq., 0.1 M) and NaHCO₃ (aq., sat.). The layers were separated and the aqueous further extracted with EtOAc twice. The combined organic layers were washed brine thrice and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

General procedure 9: acyl sulfonamides via EDC coupling

To a solution of carboxylic acid (1 eq.), EDC hydrochloride (1.5 eq.) and DMAP (1 eq.) dissolved in dry DCM under N₂ was added DIPEA (3.5 eq.). The reaction mixture was stirred for 15 min at room temperature. After this time, sulfonamide (1.5 eq.) was added and the reaction stirred for a further 24 h at room temperature. The reaction was then diluted in DCM and hydrochloride aq., 1 M). The layers were separated and the aqueous further extracted with DCM twice. The combined organic layers were washed with acidic brine (pH 4) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

General procedure 10: Reductive amination with sodium triacetoxyborohydride (STAB)

To a solution of aldehyde (1 eq.) and amine (1.5 – 3.0 eq.) dissolved in anhydrous 1,2-DCE was added glacial AcOH (3 drops) at room temperature under N₂. The reaction mixture was stirred for 4 - 48 h at room temperature. Na(AcO)₃BH (1.5 eq.) was added and the reaction mixture stirred for a further 18 h. After this time, the reaction was filtered and the solution diluted in DCM and AcOH (aq., pH 4). The layers were separated and the aqueous layer further extracted with DCM twice. The combined organic layers were washed with acidic brine (aq., pH 4) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

General procedure 11: Peptide coupling procedure 2 with HATU

To a solution of carboxylic acid (1 eq.), HATU (1 eq.) dissolved in dry DMF was added Et₃N (2 - 5 eq.) under N₂ at room temperature. The reaction mixture was stirred for 15 min then and amine (2 eq.) added and the reaction stirred for a further 18 - 72 h under N₂. After this time, the reaction was diluted in EtOAc and H₂O. Layers were separated and the organic layer washed with H₂O twice, NaHCO₃ (aq., sat.) thrice, brine and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

General procedure 12: Amide coupling with EDC hydrochloride and HOBt

To a solution of carboxylic acid (1 eq.), EDC hydrochloride (1.5 eq.) and DMAP (1 eq.) dissolved in dry DMF under N₂ was added Et₃N (3.5 eq.). The reaction mixture was stirred for 15 min at room temperature. After this time, sulfonamide (1.5 eq.) was added and the reaction stirred for a further 24 h at room temperature. The reaction was then diluted in EtOAc and H₂O. The layers were separated

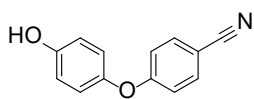
and the organic washed with H₂O twice, brine and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

General procedure 13: Alkylation with strong base

To a solution of carboxylic acid (1 eq.) dissolved in dry THF was added NaH (60 % dispersion in mineral oil, 5 – 10 eq.) at 0 °C under N₂. The reaction was stirred for 15 min on ice then alkyl halide (6.0 eq.) was added and the reaction stirred for a further 4 h at room temperature. After this time, the reaction mixture was quenched with H₂O and diluted in EtOAc. The layers were separated and the organic washed with H₂O twice, brine and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

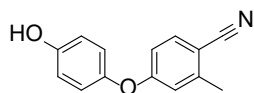
8.3 Chemistry experimental

4-(4-Hydroxyphenoxy)benzonitrile (1.27, 2.1, 3.28, 7.1)¹



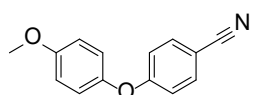
To a solution of ether **2.5** (55 mg, 1.0 eq., 0.24 mmol) dissolved in dry DCM (3 mL) was added BBr₃ in hexane (1.0 M, 1.2 mL, 5.0 eq., 1.2 mmol) dropwise at -78 °C. The reaction was stirred for 1 h at -78 °C, then allowed to warm to room temperature and stirred for an additional 2 h. The reaction mixture was then quenched with hydrochloride (aq., 0.5 M, 12 mL) and diluted with DCM (8 mL). The layers were separated and the aqueous layer further extracted with DCM (8 mL). The combined organic layers were extracted with NaOH (aq., 1 M, 2 × 8 mL). The basic layer was acidified to pH 3 using saturated citric acid (aq.) causing a white precipitate. The mixture was filtered and the precipitate washed with H₂O and dried under vacuum to give **2.1** as a white solid (42 mg, 81 %). ¹H NMR (401 MHz, CDCl₃) δ 7.60 – 7.56 (m, 2H), 6.98 – 6.93 (m, 4H), 6.90 – 6.85 (m, 2H), 4.91 (s, 1H). ¹³C NMR (101 MHz, CDCl₃) APT δ 162.69 (C), 153.30 (C), 148.08 (C), 134.25 (CH), 122.11 (CH), 119.05 (C), 117.33 (CH), 116.91 (CH), 105.34 (C). LCMS (*m/z*): 209.9 [M-H]⁻, *t_R* = 3.29 min, standard method. HRMS (*m/z*): C₁₃H₉NO₂ requires 212.0706 [M+H]⁺; found 212.0704. HPLC: *t_R* = 2.36 min, >99 %, hydrophobic method.

4-(4-Hydroxyphenoxy)-2-methylbenzonitrile (2.2, 3.1, 7.2)



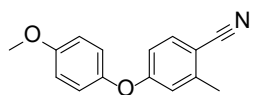
Prepared according to **general procedure 2** with ether **2.7** (260 mg, 1.00 eq., 1.08 mmol), DCM (8 mL) and BBr₃ in hexane (1.0 M, 4.30 mL, 4.00 eq., 4.30 mmol) at -78 °C. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **2.2** as a white solid (189 mg, 77 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.51 (d, *J* = 8.5 Hz, 1H), 6.97 – 6.91 (m, 2H), 6.91 – 6.85 (m, 2H), 6.80 (d, *J* = 2.3 Hz, 1H), 6.76 (dd, *J* = 8.5, 2.3 Hz, 1H), 5.23 (s, 1H), 2.48 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 162.51, 153.20, 148.11, 144.57, 134.46, 122.11, 118.44, 118.18, 116.84, 114.71, 105.78, 20.76. **LCMS** (*m/z*): 224.0 [M-H]⁻, *t*_R = 3.36 min, standard method. **HRMS** (*m/z*): C₁₄H₁₁NO₂ requires 226.0863 [M+H]⁺; found 226.0857. **HPLC**: *t*_R = 7.51 min, >99 %, standard method.

4-(4-Methoxyphenoxy)benzonitrile (**2.5**)²



4-Iodobenzonitrile (300 mg, 1.00 eq., 1.31 mmol), 4-methoxyphenol (244 mg, 1.50 eq., 1.96 mmol), Cs₂CO₃ (854 mg, 2.00 eq., 2.62 mmol) and *N,N*-dimethylglycine hydrochloride (18 mg, 0.10 eq., 0.13 mmol) were dissolved in degassed 1,4-dioxane (3 mL) and purged with N₂ (3 times). Copper(I) iodide (12 mg, 0.050 eq., 0.070 mmol) was added and the reaction mixture stirred for 18 h at 90 °C. The reaction was then cooled, diluted in EtOAc (8 mL) and washed with H₂O (8 mL), NaOH (aq., 1 M, 2 × 8 mL), brine (8 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was then recrystallised from MeOH/H₂O to give **2.5** as a pink crystalline solid (95 mg, 32 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.59 – 7.55 (m, 2H), 7.03 – 6.98 (m, 2H), 6.97 – 6.91 (m, 4H), 3.83 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 162.51, 157.04, 147.90, 134.08, 121.80, 118.94, 117.16, 115.25, 105.30, 55.68. **LCMS** (*m/z*): 226.0 [M+H]⁺, *t*_R = 3.52 min, standard method. **HRMS** (*m/z*): C₁₄H₁₁NO₂ requires 226.0863 [M+H]⁺; found 226.0857. **HPLC**: *t*_R = 2.66 min, >99 %, hydrophobic method.²

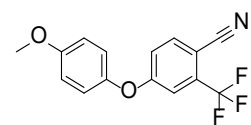
4-(4-Methoxyphenoxy)-2-methylbenzonitrile (**2.7**)



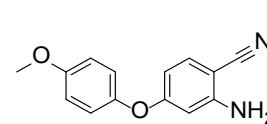
Prepared according to **general procedure 1** with 4-bromo-2-methylbenzonitrile (500 mg, 1.00 eq., 2.55 mmol), 4-methoxyphenol (475 mg, 1.50 eq., 3.83 mmol), Cs₂CO₃ (1.66 g, 2.00 eq., 5.10 mmol), *N,N*-dimethylglycine hydrochloride (142 mg,

0.40 eq., 1.02 mmol), copper(I) iodide (98 mg, 0.20 eq., 0.51 mmol) and degassed 1,4-dioxane (5 mL). The residue was purified by flash column chromatography using a gradient of 90:10 pet spirits:DCM to 0:100 pet spirits:DCM to give **2.7** as a white solid (439 mg, 72 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.49 (d, *J* = 8.5 Hz, 1H), 7.01 – 6.89 (m, 4H), 6.80 (d, *J* = 2.4 Hz, 1H), 6.76 (dd, *J* = 8.5, 2.4 Hz, 1H), 3.82 (s, 3H), 2.47 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 162.45, 157.06, 148.17, 144.50, 134.42, 121.93, 118.43, 118.18, 115.32, 114.71, 106.01, 55.81, 20.77. **LCMS** (*m/z*): 240.0 [M+H]⁺, *t_R* = 3.59 min, standard method. **HRMS** (*m/z*): C₁₅H₁₃NO₂ requires 240.1019 [M+H]⁺; found 240.1009. **HPLC**: *t_R* = 7.79 min, >99 %, standard method.

4-(4-Methoxyphenoxy)-2-(trifluoromethyl)benzonitrile (**2.8**)³

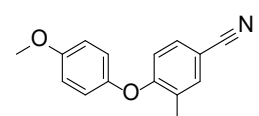
 Prepared according to **general procedure 1** with 4-iodo-2-(trifluoromethyl)benzonitrile (1.30 g, 1.00 eq., 4.37 mmol), 4-methoxyphenol (813 mg, 1.50 eq., 6.55 mmol), Cs₂CO₃ (2.85 g, 2.00 eq., 8.73 mmol), *N,N*-dimethylglycine hydrochloride (245 mg, 0.40 eq., 1.75 mmol), copper(I) iodide (167 mg, 0.20 eq., 0.870 mmol) and degassed 1,4-dioxane (15 mL). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **2.8** as a white solid (666 mg, 52 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.72 (d, *J* = 8.6 Hz, 1H), 7.29 (d, *J* = 2.5 Hz, 1H), 7.09 (dd, *J* = 8.6, 2.4 Hz, 1H), 7.05 – 6.93 (m, 4H), 3.84 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 162.55, 157.62, 147.17, 136.75, 134.96 (q, ²*J*_{C-F} = 32.9 Hz), 122.15 (q, ¹*J*_{C-F} = 274 Hz), 121.90, 119.23, 115.69, 115.61, 115.34 (q, ³*J*_{C-F} = 4.8 Hz), 102.60 (q, ³*J*_{C-F} = 1.9 Hz), 55.76. **¹⁹F NMR** (377 MHz, CDCl₃) δ -62.89. **LCMS** (*m/z*): 292.0 [M-H]⁻, *t_R* = 3.62 min, standard method. **HRMS** (*m/z*): C₁₅H₁₀F₃NO₂ requires 294.0736 [M+H]⁺; found 294.0741. **HPLC**: *t_R* = 8.85 min, >99 %, standard method.

2-Amino-4-(4-methoxyphenoxy)benzonitrile (**2.9**)

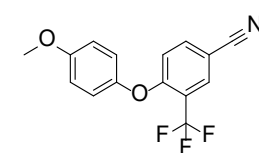
 Prepared according to **general procedure 1** with 2-amino-4-bromobenzonitrile (500 mg, 1.00 eq., 2.54 mmol), 4-methoxyphenol (473 mg, 1.00 eq., 3.81 mmol), Cs₂CO₃ (1.65 g, 2.00 eq., 5.08 mmol), *N,N*-dimethylglycine hydrochloride (142 mg, 0.400 eq., 1.02 mmol), copper(I) iodide (49 mg, 0.20 eq., 0.25 mmol) and degassed 1,4-dioxane (3

mL). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **2.9** as a white solid (114 mg, 19 %). **¹H NMR** (400 MHz, CDCl₃) δ 7.30 (d, *J* = 8.7 Hz, 1H), 7.03 – 6.97 (m, 2H), 6.95 – 6.89 (m, 2H), 6.31 (dd, *J* = 8.7, 2.3 Hz, 1H), 6.17 (d, *J* = 2.3 Hz, 1H), 4.35 (bs, 2H), 3.82 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) APT δ 163.95 (C), 156.99 (C), 151.49 (C), 147.09 (C), 134.18 (CH), 121.09 (CH), 117.93 (C), 115.19 (CH), 107.80 (CH), 102.17 (CH), 89.98 (C), 55.80 (CH₃). **LCMS** (*m/z*): 240.9 [M+H]⁺, *t_R* = 3.42 min, standard method. **HRMS** (*m/z*): C₁₄H₁₃N₂O₂ requires 241.0972 [M+H]⁺; 241.0976. **HPLC**: *t_R* = 6.74 min, >99 %, standard method.

4-(4-Methoxyphenoxy)-3-methylbenzonitrile (2.10)

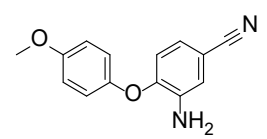
 Prepared according to **general procedure 1** with 4-bromo-3-methylbenzonitrile (500 mg, 1.00 eq., 2.55 mmol), 4-methoxyphenol (475 mg, 1.50 eq., 3.83 mmol), Cs₂CO₃ (1.66 g, 2.00 eq., 5.10 mmol), *N,N*-dimethylglycine hydrochloride (142 mg, 0.40 eq., 1.02 mmol), copper(I) iodide (98 mg, 0.20 eq., 0.51 mmol) and degassed 1,4-dioxane (5 mL). The residue was purified by flash column chromatography using a gradient of 90:10 pet spirits:DCM to 0:100 pet spirits:DCM to give **2.10** as a colourless oil (333 mg, 54 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.49 (d, *J* = 1.2 Hz, 1H), 7.36 (dd, *J* = 8.5, 2.2 Hz, 1H), 6.98 – 6.89 (m, 4H), 6.67 (d, *J* = 8.5 Hz, 1H), 3.82 (s, 3H), 2.35 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 160.74, 156.84, 148.63, 134.93, 131.62, 129.27, 121.47, 119.26, 115.65, 115.31, 105.34, 55.80, 16.22. **LCMS** (*m/z*): 238.0 [M-H]⁻, *t_R* = 3.61 min, standard method. **HRMS** (*m/z*): C₁₅H₁₃NO₂ requires 240.1019 [M+H]⁺; found 240.1013. **HPLC**: *t_R* = 7.99 min, >99 %, standard method.

4-(4-Methoxyphenoxy)-3-(trifluoromethyl)benzonitrile (2.11)

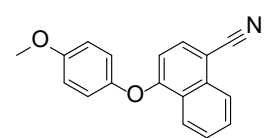
 Prepared according to **general procedure 1** with 4-iodo-3-(trifluoromethyl)benzonitrile (975 mg, 1.00 eq., 3.28 mmol), 4-methoxyphenol (921 mg, 2.25 eq., 7.42 mmol), Cs₂CO₃ (3.23 g, 3.00 eq., 9.90 mmol), *N,N*-dimethylglycine hydrochloride (276 mg, 0.60 eq., 1.98 mmol), copper(I) iodide (189 mg, 0.30 eq., 0.99 mmol) and degassed 1,4-dioxane (8 mL). The residue was purified by flash column

chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **2.11** as a white solid (319 mg, 34 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.91 (d, *J* = 1.8 Hz, 1H), 7.66 (dd, *J* = 8.7, 2.0 Hz, 1H), 7.04 – 6.99 (m, 2H), 6.96 – 6.90 (m, 2H), 6.87 (d, *J* = 8.7 Hz, 1H), 3.81 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) APT δ 160.31 (q, ³*J*_{C-F} = 1.5 Hz), 157.56, 147.16, 137.19, 131.57 (q, ³*J*_{C-F} = 5.2 Hz), 122.35 (q, ¹*J*_{C-F} = 273 Hz), 121.87, 120.77 (q, ²*J*_{C-F} = 32.5 Hz), 117.58, 117.02, 115.43, 105.28, 55.63. **¹⁹F-NMR** (377 MHz, CDCl₃) δ -62.89. **LCMS**: (*m/z*): did not ionise, *t*_R = 3.59 min, standard method. **HRMS** (*m/z*): C₁₅H₁₀F₃NO₂ requires 294.0736 [M+H]⁺; found 294.0736. **HPLC**: *t*_R = 7.79 min, >99 %, standard method.

3-Amino-4-(4-methoxyphenoxy)benzonitrile (**2.12**)

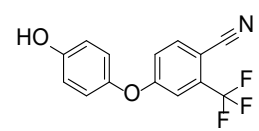
 Prepared according to **general procedure 1** with 3-amino-4-iodo-benzonitrile (975 mg, 1.00 eq., 4.00 mmol), 4-methoxyphenol (921 mg, 2.25 eq., 7.42 mmol), Cs₂CO₃ (3.22 g, 2.50 eq., 9.90 mmol), *N,N*-dimethylglycine hydrochloride (276 mg, 0.50 eq., 1.98 mmol), copper(I) iodide (189 mg, 0.25 eq., 0.990 mmol) and degassed 1,4-dioxane (10 mL). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 60:40 pet spirits:EtOAc. The residue was further purified by flash column chromatography using a gradient of 85:15 pet spirits:EtOAc to 70:30 pet spirits:EtOAc to give **2.12** as a colourless oil (302 mg, 31 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.01 – 6.96 (m, 3H), 6.93 – 6.89 (m, 3H), 6.63 (d, *J* = 8.3 Hz, 1H), 4.05 (bs, 2H), 3.81 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) APT δ 156.68 (C), 149.35 (C), 148.38 (C), 137.99 (C), 122.83 (CH), 121.04 (CH), 119.41 (C), 118.11 (CH), 115.95 (CH), 115.17 (CH), 105.98 (C), 55.70 (CH₃). **LCMS** (*m/z*): 241.0 [M+H]⁺, *t*_R = 3.36 min, standard method. **HRMS** (*m/z*): C₁₄H₁₂N₂O₂ requires 241.0972 [M+H]⁺; found 241.0978. **HPLC**: *t*_R = 6.47 min, >99 %, standard method.

4-(4-Methoxyphenoxy)-1-naphthonitrile (**2.13**)

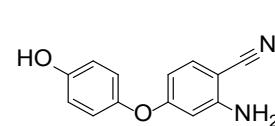
 Prepared according to **general procedure 1** with 4-bromonaphthalene-1-carbonitrile (150 mg, 1.00 eq., 0.650 mmol), 4-methoxyphenol (120 mg, 1.50 eq., 0.97 mmol), Cs₂CO₃ (632 mg, 2.00 eq., 1.94 mmol), *N,N*-dimethylglycine hydrochloride (36 mg,

0.40 eq., 0.26 mmol), copper(I) iodide (25 mg, 0.20 eq., 0.13 mmol) and degassed 1,4-dioxane (3 mL). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 90:10 pet spirits:EtOAc to give **2.13** as a white solid (63 mg, 35 %). **¹H NMR** (401 MHz, CDCl₃) δ 8.49 (ddd, *J* = 8.4, 1.3, 0.7 Hz, 1H), 8.23 (ddd, *J* = 8.4, 1.2, 0.7 Hz, 1H), 7.78 – 7.73 (m, 2H), 7.66 (ddd, *J* = 8.2, 6.9, 1.2 Hz, 1H), 7.12 – 7.07 (m, 2H), 7.01 – 6.95 (m, 2H), 6.65 (d, *J* = 8.1 Hz, 1H), 3.85 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 159.56, 157.30, 148.28, 134.11, 133.75, 129.40, 127.31, 125.57, 125.23, 122.96, 122.18, 118.34, 115.47, 108.20, 103.27, 55.86. **LCMS** (*m/z*): Did not ionise, *t_R* = 3.67 min, standard method. **HRMS** (*m/z*): C₁₈H₁₃NO₂ requires 276.1019 [M+H]⁺; found 276.1015. **HPLC**: *t_R* = 8.22 min, >99 %, standard method.

4-(4-Hydroxyphenoxy)-2-(trifluoromethyl)benzonitrile (**2.14**)

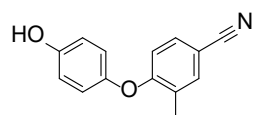
 Prepared according to **general procedure 2** with ether **2.8** (666 mg, 1.00 eq., 2.27 mmol), DCM (15 mL) and BBr₃ in heptane (1.0 M, 6.81 mL, 3.00 eq., 6.81 mmol) at 0 °C. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **2.14** as a white solid (260 mg, 41 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.66 (d, *J* = 8.6 Hz, 1H), 7.22 (d, *J* = 2.5 Hz, 1H), 7.02 (dd, *J* = 8.6, 2.4 Hz, 1H), 6.94 – 6.87 (m, 2H), 6.87 – 6.80 (m, 2H), 4.87 (s, 1H). **¹³C NMR** (101 MHz, CDCl₃) δ 162.67, 153.85, 147.19, 136.83, 135.12 (q, ²*J*_{C-F} = 33.0 Hz), 122.14 (q, ¹*J*_{C-F} = 274 Hz), 122.12, 119.26, 115.75, 115.41 (q, ³*J*_{C-F} = 4.8 Hz), 102.41, 102.39. **LCMS** (*m/z*): 277.9 [M-H]⁻, *t_R* = 3.40 min, standard method. **HRMS** (*m/z*): C₁₄H₈F₃NO₂ requires 280.0580 [M+H]⁺; found 280.0580. **HPLC**: *t_R* = 6.71 min, >99 %, standard method.

2-Amino-4-(4-hydroxyphenoxy)benzonitrile (**2.15**)

 Prepared according to **general procedure 2** with ether **2.9** (460 mg, 1.00 eq., 1.91 mmol), DCM (10 mL) and BBr₃ in DCM (1 M, 4.79 mL, 3.00 eq., 4.79 mmol) at -78 °C. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to pet spirits:EtOAc (50:50). The residue was then recrystallised from DCM/pet spirits to give **2.15** as a white solid (191 mg, 44 %). **¹H NMR** (401 MHz, (CD₃)₂CO) δ 8.33 (s, 1H),

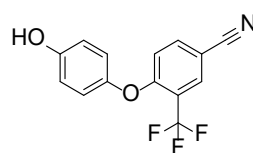
7.28 (dd, $J = 8.5, 0.5$ Hz, 1H), 6.93 – 6.87 (m, 2H), 6.87 – 6.82 (m, 2H), 6.23 (d, $J = 2.3$ Hz, 1H), 6.20 (dd, $J = 8.5, 2.4$ Hz, 1H), 5.48 (bs, 2H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 163.41, 154.61, 153.44, 146.15, 134.35, 121.93, 118.23, 116.37, 105.72, 100.87, 87.50. LCMS (m/z): 227.0 $[\text{M}+\text{H}]^+$, $t_{\text{R}} = 3.16$ min, standard method. HRMS (m/z): $\text{C}_{13}\text{H}_{10}\text{N}_2\text{O}_2$ requires 227.0815 $[\text{M}+\text{H}]^+$; found 227.0813. HPLC: $t_{\text{R}} = 5.71$ min, >99 %, standard method.

4-(4-Hydroxyphenoxy)-3-methylbenzonitrile (2.16)

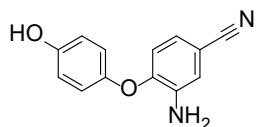


Prepared according to **general procedure 2** with ether **2.10** (250 mg, 1.00 eq., 1.10 mmol), DCM (10 mL) and BBr_3 in hexane (1.0 M, 4.4 mL, 4.0 eq., 4.4 mmol) at -78°C . The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **2.16** as a white solid (193 mg, 79 %). ^1H NMR (401 MHz, CDCl_3) δ 7.50 (d, $J = 1.2$ Hz, 1H), 7.37 (dd, $J = 8.5, 1.2$ Hz, 1H), 6.93 – 6.84 (m, 4H), 6.68 (d, $J = 8.5$ Hz, 1H), 5.17 (s, 1H), 2.35 (s, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 162.48, 153.13, 148.18, 144.57, 134.46, 122.12, 118.44, 118.19, 116.84, 114.72, 105.86, 20.77. LCMS (m/z): 226.0 $[\text{M}+\text{H}]^+$, $t_{\text{R}} = 3.38$ min, standard method. HRMS (m/z): $\text{C}_{14}\text{H}_{11}\text{NO}_2$ requires 226.0863 $[\text{M}+\text{H}]^+$; found 226.0861. HPLC: $t_{\text{R}} = 7.69$ min, >99 %, standard method.

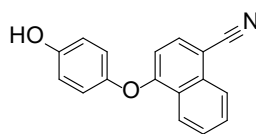
4-(4-Hydroxyphenoxy)-3-(trifluoromethyl)benzonitrile (2.17)



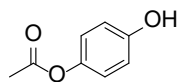
Prepared according to **general procedure 2** with ether **2.11** (279 mg, 1.00 eq., 0.950 mmol), DCM (5 mL) and BBr_3 in heptane (1.0 M, 2.85 mL, 3.00 eq., 2.85 mmol) at 0°C . The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **2.17** as a white solid (92 mg, 35 %). ^1H NMR (401 MHz, CDCl_3) δ 7.94 (d, $J = 2.0$ Hz, 1H), 7.67 (dd, $J = 8.7, 2.0$ Hz, 1H), 7.01 – 6.96 (m, 2H), 6.92 – 6.85 (m, 3H), 4.87 (s, 1H). ^{13}C NMR (101 MHz, CDCl_3) δ 160.49 (q, $^3J_{\text{C-F}} = 1.5$ Hz), 153.80, 147.36, 137.30, 131.81 (q, $^3J_{\text{C-F}} = 5.2$ Hz), 122.55, 122.41 (q, $^1J_{\text{C-F}} = 273.2$ Hz), 121.08 (q, $^2J_{\text{C-F}} = 32.6$ Hz), 117.73, 117.12, 117.09, 105.30. LCMS (m/z): 277.9 $[\text{M}-\text{H}]^-$, $t_{\text{R}} = 3.39$ min, standard method. HRMS (m/z): $\text{C}_{14}\text{H}_8\text{F}_3\text{NO}_2$ requires 280.0580 $[\text{M}+\text{H}]^+$; found 280.0580. HPLC: $t_{\text{R}} = 6.83$ min, >99 %, standard method.

3-Amino-4-(4-hydroxyphenoxy)benzonitrile (2.18)

Prepared according to **general procedure 2** with ether **2.12** (302 mg, 1.00 eq., 1.26 mmol), DCM (10 mL) and BBr₃ in heptane (1.0 M, 3.77 mL, 3.00 eq., 3.77 mmol) at -78 °C. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc. The residue was triturated with pet spirits to give **2.18** as a yellow solid (33 mg, 11 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.01 (d, *J* = 1.9 Hz, 1H), 6.96 – 6.91 (m, 3H), 6.87 – 6.83 (m, 2H), 6.64 (d, *J* = 8.3 Hz, 1H), 4.68 (s, 1H), 4.10 (bs, 2H). **¹³C NMR** (101 MHz, CDCl₃) δ 152.84, 149.50, 148.55, 137.96, 123.14, 121.38, 119.43, 118.29, 116.79, 116.07, 106.17. **LCMS** (*m/z*): 227.0 [M+H]⁺, *t_R* = 3.12 min, standard method. **HRMS** (*m/z*): C₁₃H₁₀N₂O₂ requires 227.0815 [M+H]⁺; found 227.0818. **HPLC**: *t_R* = 5.20 min, >99 %, standard method.

4-(4-Hydroxyphenoxy)-1-naphthonitrile (2.19)

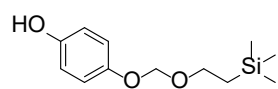
Prepared according to **general procedure 2** with ether **2.13** (60 mg, 1.0 eq., 0.22 mmol), DCM (5 mL) and BBr₃ in heptane (1.0 M, 0.97 mL, 4.5 eq., 0.97 mmol) at 0 °C. The residue was purified by flash column chromatography using a gradient of 90:10 pet spirits:EtOAc to 65:35 pet spirits:EtOAc to give **2.19** as a white solid (31 mg, 54 %). **¹H NMR** (401 MHz, CD₃CN) δ 8.48 (ddd, *J* = 8.4, 1.3, 0.7 Hz, 1H), 8.17 (ddd, *J* = 8.4, 1.2, 0.8 Hz, 1H), 7.85 (d, *J* = 8.2 Hz, 1H), 7.81 (ddd, *J* = 8.3, 7.0, 1.4 Hz, 1H), 7.72 (ddd, *J* = 8.2, 6.9, 1.2 Hz, 1H), 7.09 – 7.04 (m, 2H), 7.09 – 6.89 (bs, 1H), 6.94 – 6.89 (m, 2H), 6.70 (d, *J* = 8.2 Hz, 1H). **¹³C NMR** (101 MHz, CD₃CN) δ 160.47, 155.58, 148.45, 135.16, 134.72, 130.51, 128.36, 126.24, 125.67, 123.66, 123.13, 118.81, 117.63, 109.22, 103.61. **LCMS** (*m/z*): 259.9 [M-H]⁻, *t_R* = 3.43 min, standard method. **HRMS** (*m/z*): C₁₇H₁₁NO₂ requires 262.0863 [M+H]⁺; found 262.0862. **HPLC**: *t_R* = 7.10 min, >99 %, standard method.

4-Hydroxyphenyl acetate (2.22)

To a solution of hydroquinone (560 mg, 2.0 eq. 5.1 mmol) and DIPEA (0.89 mL, 2.0 eq. 5.1 mmol) was added acetyl chloride (0.18mL, 1.0 eq. 2.6 mmol) dropwise at 0 °C. The reaction mixture was stirred for 0.5 h at 0 °C, then 4 h at room temperature. After this time,

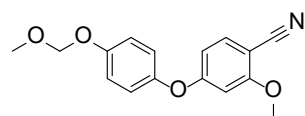
the solvent was removed under vacuum and the residue dissolved in EtOAc (20 mL). The organic layer was washed with H₂O (15 mL), brine (15 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 50:50 pet spirits:EtOAc to give **2.22** as a pink solid (105 mg, 27 %). ¹H NMR(401 MHz, CDCl₃) δ 6.92 – 6.82 (m, 2H), 6.74 – 6.64 (m, 2H), 6.39 (s, 1H), 2.28 (s, 3H). ¹³C NMR(101 MHz, CDCl₃) δ 171.22, 153.75, 143.85, 122.37, 116.22, 21.19. LCMS (*m/z*): 211.2 [M+CH₃COO-H]⁺. HPLC: *t*_R= 7.42 min, 94 %, standard method.⁴

4-((2-(Trimethylsilyl)ethoxy)methoxy)phenol (**2.23**)



To a solution of hydroquinone (1.0 g, 1.0 eq. 9.1 mmol) and DIPEA (3.2 mL, 2.0 eq. 18.2 mmol) dissolved in MeCN (30 mL) was added 2-(trimethylsilyl)ethoxymethyl chloride (1.8 mL, 1.1 eq. 10 mmol) dropwise at 0 °C. The reaction mixture was stirred for 0.5 h at 0 °C, then 28 h at room temperature. After this time, the solvent was removed under vacuum and the residue dissolved in EtOAc (50 mL). The organic layer was washed with NaHSO₄ (aq., 1 M, 30 mL), H₂O (2 x 50 mL), brine (30 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 85:15 pet spirits:EtOAc to give **2.23** as a clear oil (568 mg, 26 %). ¹H NMR (401 MHz, CDCl₃) δ 6.93 – 6.86 (m, 2H), 6.74 – 6.69 (m, 2H), 5.96 (s, 1H), 5.15 (s, 2H), 3.80 – 3.75 (m, 2H), 1.02 – 0.91 (m, 2H), -0.00 (s, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 152.62, 152.17, 119.41, 117.57, 95.24, 67.63, 19.49, 0.00. LCMS (*m/z*): 239.0 [M-H]⁺, *t*_R = 3.56 min, 95 %, standard method. HPLC: *t*_R= 7.42 min, >95 %, standard method.

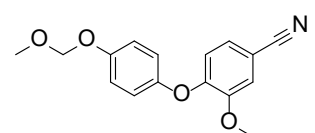
2-Methoxy-4-(4-(methoxymethoxy)phenoxy)benzonitrile (**2.25**)



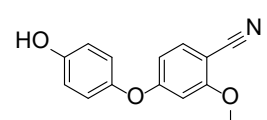
Prepared according to **general procedure 1** with 4-bromo-2-methoxybenzonitrile (25 mg, 1.0 eq., 0.12 mmol), 4-(methoxymethoxy)phenol (0.02 mL, 1.2 eq., 0.14 mmol), Cs₂CO₃ (0.23 g, 2.0 eq., 0.71 mmol), *N,N*-dimethylglycine hydrochloride (7.0 mg, 0.40 eq., 0.04 mmol), copper(I) iodide (5.0 mg, 0.20 eq., 0.02 mmol) and degassed 1,4-dioxane (3 mL). The residue was purified by flash column chromatography

using a gradient of 100:0 pet spirits:EtOAc to 80:20 pet spirits:EtOAc. The residue was dissolved in C hydrochloride₃ (15 mL) and washed with NaOH (aq., 1 M, 2 × 10 mL), brine and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness to give **2.25** as a colourless oil (15.6 mg, 46 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.36 (d, *J* = 8.6 Hz, 1H), 7.05 – 6.89 (m, 4H), 6.47 (d, *J* = 1.6 Hz, 1H), 6.38 (dd, *J* = 8.5, 1.8 Hz, 1H), 5.10 (s, 2H), 3.79 (s, 3H), 3.44 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 164.05, 163.16, 154.78, 149.05, 135.02, 121.85, 117.94, 116.75, 108.92, 100.76, 95.44, 95.02, 56.23, 56.22. **LCMS** (*m/z*): did not ionise, *t_R* = 3.44 min, standard method. **HRMS** (*m/z*): C₁₆H₁₅NO₄ requires 286.1074 [M+H]⁺; found 286.1079. **HPLC**: *t_R* = 7.17 min, >99 %, standard method.

3-Methoxy-4-(4-(methoxymethoxy)phenoxy)benzonitrile (**2.26**)

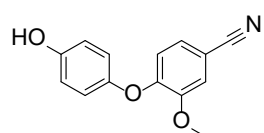
 Prepared according to **general procedure 1** with 4-bromo-3-methoxybenzonitrile (25 mg, 1.0 eq., 0.12 mmol), 4-(methoxymethoxy)phenol (0.02 mL, 1.2 eq., 0.14 mmol), Cs₂CO₃ (0.23 g, 6.0 eq., 0.71 mmol), *N,N*-dimethylglycine hydrochloride (7.0 mg, 0.40 eq., 0.05 mmol), copper(I) iodide (5.0 mg, 0.20 eq., 0.02 mmol) and degassed 1,4-dioxane (5 mL). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 80:20 pet spirits:EtOAc to give **2.26** as a colourless oil (17 mg, 50 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.12 – 7.08 (m, 2H), 7.00 – 6.96 (m, 2H), 6.92 – 6.88 (m, 2H), 6.70 (d, *J* = 8.2 Hz, 1H), 5.09 (s, 2H), 3.86 (s, 3H), 3.43 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 154.48, 151.87, 150.28, 149.67, 126.12, 121.12, 119.03, 117.87, 117.29, 115.44, 106.15, 95.06, 56.43, 56.19. **LCMS** (*m/z*): did not ionise, *t_R* = 3.40 min, standard method. **HRMS** (*m/z*): C₁₆H₁₅NO₄ requires 286.1074 [M+H]⁺; found 286.1073. **HPLC**: *t_R* = 6.95 min, >99 %, standard method.

4-(4-Hydroxyphenoxy)-2-methoxybenzonitrile (**2.27**)

 To a solution of ether **2.25** (16 mg, 1.0 eq., 0.05 mmol) dissolved in 50:50 MeOH/H₂O (8 mL) was added hydrochloride (36 % w/v, 0.3 mL) dropwise at room temperature. The reaction mixture was heated at 80 °C for 18 h. After this time the reaction was cooled to room temperature and diluted in H₂O (15 mL) and EtOAc (25 mL). The layers were

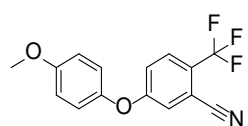
separated and the organic washed with H₂O (20 mL), brine (20 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 70:30 pet spirits:EtOAc to give **2.27** as a white solid (11 mg, 85 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.44 (d, *J* = 8.6 Hz, 1H), 6.99 – 6.93 (m, 2H), 6.91 – 6.85 (m, 2H), 6.52 (d, *J* = 2.2 Hz, 1H), 6.45 (dd, *J* = 8.6, 2.2 Hz, 1H), 5.14 (bs, 1H), 3.85 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 164.15, 163.04, 153.04, 147.94, 134.91, 122.03, 116.71, 116.67, 108.68, 100.41, 95.09, 56.06. **LCMS** (*m/z*): 239.9 [M-H]⁻, *t_R* = 3.30 min, standard method. **HRMS** (*m/z*): C₁₄H₁₁NO₃ requires 242.0812 [M+H]⁺; found 242.0814. **HPLC**: *t_R* = 6.11 min, >99 %, standard method.

4-(4-Hydroxyphenoxy)-3-methoxybenzonitrile (**2.28**)



To a solution of ether **2.26** (17 mg, 1.0 eq., 0.058 mmol) dissolved in 50:50 MeOH/H₂O (8 mL) was added hydrochloride (aq., 36 % w/v, 0.3 mL) dropwise at room temperature. The reaction mixture was stirred at 80 °C for 18 h. After this time the reaction was cooled and diluted with H₂O (15 mL) and EtOAc (25 mL). The layers were separated and the organic washed with H₂O (20 mL), brine (20 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 70:30 pet spirits:EtOAc to give **2.28** as a white solid (6 mg, 42 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.22 – 7.12 (m, 2H), 6.98 – 6.90 (m, 2H), 6.89 – 6.82 (m, 2H), 6.74 (d, *J* = 8.2 Hz, 1H), 5.07 (s, 1H), 3.94 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 152.94, 152.17, 150.10, 148.52, 126.18, 121.54, 119.06, 116.83, 116.76, 115.33, 105.78, 56.42. **LCMS** (*m/z*): 239.9 [M-H]⁻, *t_R* = 3.20 min, standard method. **HRMS** (*m/z*): C₁₄H₁₁NO₃ requires 242.0812 [M+H]⁺; found 242.0813. **HPLC**: *t_R* = 4.02 min, >99 %, standard method.

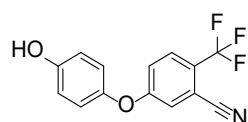
5-(4-Methoxyphenoxy)-2-(trifluoromethyl)benzonitrile (**2.30**)⁵



Prepared according to **general procedure 1** with 5-bromo-2-(trifluoromethyl)benzonitrile (50 mg, 1.0 eq., 0.20 mmol), 4-methoxyphenol (37 mg, 1.5 eq., 0.30 mmol), Cs₂CO₃ (130 mg, 2.00 eq., 0.399 mmol), *N,N*-dimethylglycine

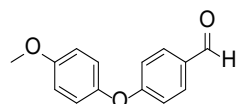
hydrochloride (11 mg, 0.40 eq., 0.08 mmol), copper(I) iodide (8.0 mg, 0.20 eq., 0.04 mmol) and degassed 1,4-dioxane (5 mL). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 80:20 pet spirits:EtOAc to give **2.30** as a colourless oil (22 mg, 38 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.68 (d, *J* = 8.8 Hz, 1H), 7.28 (d, *J* = 2.6 Hz, 1H), 7.19 (dd, *J* = 8.8, 2.6 Hz, 1H), 7.03 – 6.99 (m, 2H), 6.98 – 6.94 (m, 2H), 3.84 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 161.77, 157.59, 147.40, 128.73 (q, ³*J*_{C-F} = 4.6 Hz), 126.11 (q, ²*J*_{C-F} = 33.1 Hz), 122.65 (q, ¹*J*_{C-F} = 273 Hz), 122.19, 121.92, 120.44, 115.65, 115.25, 111.75 (q, ³*J*_{C-F} = 2.1 Hz), 55.84. **LCMS**: (*m/z*): did not ionise, *t*_R = 3.71 min, standard method. **HRMS** (*m/z*): C₁₅H₁₀F₃NO₂ requires 294.0736 [M+H]⁺; found 294.0748. **HPLC**: *t*_R = 7.78 min, >99 %, standard method.

5-(4-Hydroxyphenoxy)-2-(trifluoromethyl)benzonitrile (**2.31**)



Prepared according to **general procedure 2** with ether **2.30** (15 mg, 1.0 eq., 0.05 mmol), DCM (2 mL) and BBr₃ in heptane (1.0 M, 0.30 mL, 6.0 eq., 0.30 mmol) at 0 °C. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 80:20 pet spirits:EtOAc to give **2.31** as a white solid (11 mg, 76 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.61 (d, *J* = 8.8 Hz, 1H), 7.21 (d, *J* = 2.4 Hz, 1H), 7.13 (dd, *J* = 8.8, 2.0 Hz, 1H), 6.92 – 6.86 (m, 2H), 6.86 – 6.81 (m, 2H), 5.00 (bs, 1H). **¹³C NMR** (101 MHz, CDCl₃) δ 161.71 (q, ³*J*_{C-F} = 0.8 Hz), 153.65, 147.50, 128.76 (q, ³*J*_{C-F} = 4.6 Hz), 126.19 (q, ²*J*_{C-F} = 33.2 Hz), 122.64 (q, ¹*J*_{C-F} = 273 Hz), 122.22, 122.12, 120.52, 117.18, 115.25, 111.68 (q, ³*J*_{C-F} = 2.0 Hz). **LCMS** (*m/z*): 277.9 [M+H]⁺, *t*_R = 3.40 min, standard method. **HRMS** (*m/z*): C₁₄H₈F₃NO₂ requires 280.0580 [M+H]⁺; found 280.0576. **HPLC**: *t*_R = 6.84 min, >99 %, standard method.

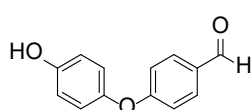
4-(4-Methoxyphenoxy)benzaldehyde (**2.33**)⁶



Prepared according to **general procedure 1** with 4-bromobenzaldehyde (2.00 g, 1.00 eq., 10.8 mmol), 4-methoxyphenol (2.01 g, 1.50 eq., 16.2 mmol), Cs₂CO₃ (7.04 g, 2.00 eq., 21.6 mmol), *N,N*-dimethylglycine hydrochloride (604 mg, 0.40 eq., 4.32 mmol), copper(I) iodide (414 mg, 0.20 eq., 2.16 mmol) and degassed 1,4-dioxane (15 mL). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet

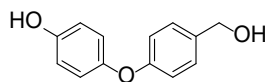
spirits:EtOAc to give **2.33** as a yellow solid (1.86 g, 75 %). **¹H NMR** (401 MHz, CDCl₃) δ 9.90 (s, 1H), 7.84 – 7.79 (m, 2H), 7.05 – 6.97 (m, 4H), 6.96 – 6.90 (m, 2H), 3.82 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) APT δ 190.73 (CH), 164.11 (C), 156.91 (C), 148.21 (C), 131.93 (CH), 130.92 (C), 121.82 (CH), 116.80 (CH), 115.18 (CH), 55.66 (CH₃). **LCMS** (*m/z*): 229.0 [M+H]⁺, *t_R* = 3.44 min, standard method. **HRMS** (*m/z*): C₁₄H₁₂O₃ requires 229.0859 [M+H]⁺; found 229.0859. **HPLC**: *t_R* = 6.92 min, >99 %, standard method.⁶

4-(4-Hydroxyphenoxy)benzaldehyde (**2.34**)⁷



Prepared according to **general procedure 2** with aldehyde **2.33** (300 mg, 1.00 eq., 1.31 mmol), DCM (5 mL) and BBr₃ in heptane (1.0 M, 3.94 mL, 3.0 eq., 3.94 mmol) at 0 °C. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 80:20 pet spirits:EtOAc. The residue was further purified by flash column chromatography using a gradient of 88:12 pet spirits:EtOAc to 78:22 pet spirits:EtOAc to give **2.34** as a white solid (34 mg, 12 %). **¹H NMR** (401 MHz, CDCl₃) δ 9.94 (s, 1H), 7.89 – 7.81 (m, 2H), 7.08 – 6.97 (m, 4H), 6.94 – 6.87 (m, 2H), 4.72 (bs, 1H). **LCMS** (*m/z*): 214.9 [M+H]⁺, *t_R* = 3.18 min, standard method. **HRMS** (*m/z*): C₁₃H₁₀O₃ requires 215.0703 [M+H]⁺; found 215.0703. **HPLC**: *t_R* = 5.78 min, >99 %, standard method.

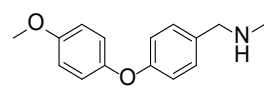
4-(4-(Hydroxymethyl)phenoxy)phenol (**2.35**)



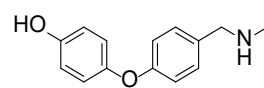
Aldehyde **2.34** (35 mg, 1.0 eq., 0.16 mmol) was added to an oven dried multineck round bottom flask connected to a reflux condenser and purged with N₂ (twice). Dry THF (5 mL) was added, followed by dropwise addition of LiAlH₄ in THF (1 M, 0.25 mL, 1.5 eq., 0.25 mmol) at room temperature. The reaction mixture was then heated at reflux for 18 h. After this time, the reaction was cooled to room temperature, quenched with 1 M citric acid (aq., 20 mL) and diluted in DCM (20 mL). The layers were separated and the aqueous layer further extracted with DCM (2 × 40 mL). The combined organic layers were washed with brine (40 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 80:20 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to

give **2.35** as a white solid (17 mg, 48 %). **¹H NMR** (401 MHz, CD₃CN) δ 7.30 – 7.25 (m, 2H), 6.91 – 6.86 (m, 4H), 6.85 (bs, 1H), 6.83 – 6.78 (m, 2H), 4.51 (d, J = 3.0 Hz, 2H), 3.10 (t, J = 5.4 Hz, 1H). **¹³C NMR** (101 MHz, CD₃CN) δ 158.80, 154.27, 150.47, 137.24, 129.37, 121.87, 118.12, 117.20, 64.28. **LCMS** (m/z): 215.0 [M-H]⁻, t_R = 3.02 min, standard method. **HRMS** (m/z): C₁₃H₁₂O₃ requires 239.0679 [M+Na]⁺; found 239.0675. **HPLC**: t_R = 4.95 min, >99 %, standard method.

1-(4-(4-Methoxyphenoxy)phenyl)-*N*-methylethanamine (**2.36**)

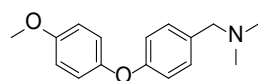
 Aldehyde **2.33** (600 mg, 1.00 eq., 2.63 mmol), methylamine in EtOH (33 %, 0.43 mL, 4.0 eq., 11 mmol) and NaHCO₃ (265 mg, 1.20 eq., 3.15 mmol) were dissolved in MeOH (3 mL) and heated at reflux for 5 h. The reaction was then cooled and NaBH₄ (199 mg, 2.00 eq., 5.26 mmol) was added at 0 °C. The reaction mixture was stirred for 1 h at 0 °C, then allowed to warm to room temperature and stirred for an additional 18 h. The reaction mixture was then quenched with H₂O (25 mL) and diluted in EtOAc (30 mL). The layers were separated and the aqueous layer further extracted with EtOAc (2 × 30 mL). The combined organic layers were washed with brine (20 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 DCM:MeOH to 65:35 DCM:MeOH to give **2.36** as a colourless oil (381 mg, 60 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.26 – 7.20 (m, 2H), 6.99 – 6.93 (m, 2H), 6.93 – 6.83 (m, 4H), 3.78 (s, 3H), 3.69 (s, 2H), 2.44 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 157.49, 155.86, 150.39, 134.37, 129.49, 120.66, 117.64, 114.87, 55.65, 55.51, 36.02. **LCMS** (m/z): 213.0 [M-CH₃NH₂]⁺, t_R = 2.842 min, standard method. **HRMS** (m/z): C₁₅H₁₇NO₂ requires 244.1332 [M+H]⁺; found 244.1336. **HPLC**: t_R = 5.13 min, >99 %, standard method.

4-(4-((Methylamino)methyl)phenoxy)phenol (**2.37**)

 Prepared according to **general procedure 2** with ether **2.36** (384 mg, 1.00 eq., 1.58 mmol), DCM (15 mL) and BBr₃ in heptane (1.0 M, 4.73 mL, 3.00 eq., 4.73 mmol) at 0 °C. The residue was purified by flash column chromatography using a gradient of 100:0 EtOAc:MeOH to 70:30 EtOAc:MeOH to give **2.37** as a white solid (49 mg, 14 %). **¹H NMR**

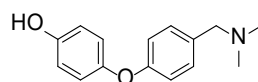
(401 MHz, d_6 -DMSO) δ 9.40 (bs, 1H), 8.46 (bs, 1H), 7.46 – 7.41 (m, 2H), 6.96 – 6.87 (m, 4H), 6.83 – 6.78 (m, 2H), 4.06 (s, 2H) 2.53 (s, 3H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 159.00, 154.14, 147.37, 131.61, 125.87, 121.17, 116.78, 116.35, 50.85, 32.10. **LCMS** (m/z): 199.0 $[\text{M}-\text{CH}_3\text{NH}_2]^+$, t_{R} = 2.68 min, standard method. **HRMS** (m/z): $\text{C}_{14}\text{H}_{15}\text{NO}_2$ requires 230.1176 $[\text{M}+\text{H}]^+$; found 230.1180. **HPLC**: t_{R} = 4.19 min, >99 %, standard method.

1-(4-(4-Methoxyphenoxy)phenyl)-*N,N*-dimethylmethanamine (2.38)



To a solution of aldehyde **2.33** (579 mg, 1.00 eq., 2.54 mmol), *N,N*-dimethylamine hydrochloride (414 mg, 2.00 eq., 5.07 mmol) and Et_3N (0.71 mL, 2.0 eq., 5.1 mmol) dissolved in anhydrous 1,2-DCE (10 mL) was added $\text{Na}(\text{CH}_3\text{COO})_3\text{BH}$ (807 mg, 1.50 eq., 3.81 mmol) at room temperature. The reaction mixture was stirred at room temperature for 3 h. After this time the reaction was diluted in DCM (15 mL) and washed with NaHCO_3 (aq., sat., 3×20 mL), brine (20 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 93:5:2 EtOAc:MeOH:6 M NH_3 in MeOH to 78:20:2 EtOAc:MeOH:6 M NH_3 in MeOH to give **2.38** as a colourless oil (280 mg, 43 %). ^1H NMR (401 MHz, CDCl_3) δ 7.25 – 7.20 (m, 2H), 6.99 – 6.95 (m, 2H), 6.92 – 6.84 (m, 4H), 3.77 (s, 3H), 3.37 (s, 2H), 2.23 (s, 6H). ^{13}C NMR (101 MHz, CDCl_3) δ 157.66, 155.96, 150.47, 133.14, 130.48, 120.80, 117.55, 114.97, 63.86, 55.79, 45.42. **LCMS** (m/z): 258.1 $[\text{M}+\text{H}]^+$, t_{R} = 2.86 min, standard method. **HRMS** (m/z): $\text{C}_{16}\text{H}_{19}\text{NO}_2$ requires 258.1489 $[\text{M}+\text{H}]^+$; found 258.1485. **HPLC**: t_{R} = 5.34 min, >99 %, standard method.

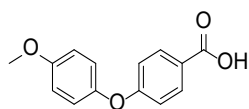
4-(4-((Dimethylamino)methyl)phenoxy)phenol (2.39)



Prepared according to **general procedure 2** with ether **2.38** (1.24 g, 1.00 eq., 4.82 mmol), DCM (30 mL) and BBr_3 in heptane (1.0 M, 9.64 mL, 2.00 eq., 9.64 mmol) at -78 °C. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 80:20 pet spirits:EtOAc. The residue was further purified by trituration with C hydrochloride₃ to give **2.39** as a white solid (293 mg, 25 %). ^1H NMR (401 MHz, d_6 -DMSO) δ 9.42 (s, 1H), 7.54 – 7.47 (m, 2H), 6.97 – 6.91 (m, 4H), 6.84 – 6.78 (m, 2H), 4.57 (s, 2H), 2.83 (s,

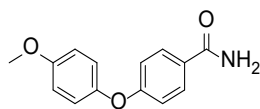
6H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 159.84, 154.35, 146.90, 135.14, 122.13, 121.50, 116.38 (2 carbons), 61.60, 45.39. LCMS (m/z): did not ionise, t_R = 3.58 min, standard method. HRMS (m/z): $\text{C}_{15}\text{H}_{17}\text{NO}_2$ requires 244.1332 $[\text{M}+\text{H}]^+$; found 244.1337. HPLC: t_R = 7.79 min, >99 %, standard method.

4-(4-Methoxyphenoxy)benzoic acid (**2.40**)⁸



To a solution of aldehyde **2.33** (1.38 g, 1.00 eq., 6.06 mmol), H_2O_2 (30 %, 1.00 mL, 1.60 eq., 9.71 mmol), NaH_2PO_4 (764 mg, 1.05 eq., 6.37 mmol) in dry MeCN (40 mL) at 0 °C was added NaClO_2 (2.74 g, 5.0 eq., 30.3 mmol) dissolved in H_2O (15 mL). The reaction mixture was stirred at 0 °C for 15 min, then heated at reflux for 3 h. The MeCN was removed under vacuum and the remaining aqueous layer was extracted with EtOAc (3 \times 50 mL). The combined organic layers were washed with brine (40 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **2.40** as a white solid (684 mg, 46 %). ^1H NMR (401 MHz, d_6 -DMSO) δ 12.74 (s, 1H), 7.96 – 7.85 (m, 2H), 7.12 – 7.06 (m, 2H), 7.05 – 6.98 (m, 2H), 6.98 – 6.92 (m, 2H), 3.77 (s, 3H). ^{13}C NMR (101 MHz, d_6 -DMSO) APT δ 167.25 (C), 162.52 (C), 156.81 (C), 148.45 (C), 132.07 (CH), 125.06 (C), 122.12 (CH), 116.64 (CH), 115.76 (CH), 55.92 (CH_3). LCMS (m/z): 245.0 $[\text{M}+\text{H}]^+$, t_R = 3.35 min, standard method. HRMS (m/z): $\text{C}_{14}\text{H}_{12}\text{O}_4$ requires 245.0808 $[\text{M}+\text{H}]^+$; found 245.0812. HPLC: t_R = 6.45 min, >99 %, standard method.

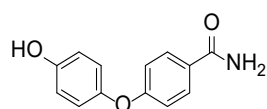
4-(4-Methoxyphenoxy)benzamide (**2.41**)⁹



To a solution of carboxylic acid **2.40** (500 mg, 1.00 eq., 2.05 mmol) and DIPEA (1.10 mL, 3.00 eq., 6.10 mmol) dissolved in dry THF (10 mL) was added DMF (3 drops) followed by SOCl_2 (1.79 mL, 12.0 eq., 24.6 mmol) dropwise at room temperature. The reaction mixture was heated at reflux for 18 h. The reaction mixture was evaporated to dryness under vacuum and the residue dissolved in dry THF (5 mL). The acid chloride was added dropwise to a solution of 30 % ammonia (5 mL) at 0 °C. The reaction was stirred for 1 h at 0 °C, then allowed to

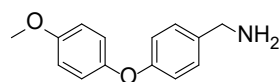
warm to room temperature and stirred for an additional 60 h. The THF was removed under vacuum, causing precipitation. The solution was filtered and the precipitate washed with H₂O. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **2.41** as a white solid (132 mg, 27 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 7.92 – 7.82 (m, 3H), 7.23 (bs, 1H), 7.09 – 7.03 (m, 2H), 7.02 – 6.96 (m, 2H), 6.95 – 6.88 (m, 2H), 3.77 (s, 3H). **¹³C NMR** (101 MHz, *d*₆-DMSO:CDCl₃, 95:5) δ 167.20, 160.63, 156.09, 148.41, 129.56, 128.26, 121.31, 116.04, 115.18, 55.40. **LCMS** (*m/z*): 243.9 [M+H]⁺, *t*_R = 3.14 min, standard method. **HRMS** (*m/z*): C₁₄H₁₃NO₃ requires 244.0968 [M+H]⁺; found 244.0973. **HPLC**: *t*_R = 5.75 min, >99 %, standard method.

4-(4-Hydroxyphenoxy)benzamide (**2.42**)⁹



Prepared according to **general procedure 2** with ether **2.41** (75 mg, 1.0 eq., 0.31 mmol), DCM (3 mL) and BBr₃ in heptane (1.0 M, 0.92 mL, 3.0 eq., 0.92 mmol) at 0 °C. The residue was purified by flash column chromatography using a gradient of 50:50 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **2.42** as a white solid (37 mg, 52 %). **¹H NMR** (401 MHz, CD₃CN) δ 7.80 – 7.75 (m, 2H), 6.97 – 6.90 (m, 4H), 6.88 – 6.82 (m, 2H), 6.64 (bs, 1H), 5.84 (bs, 1H). **¹³C NMR** (101 MHz, CD₃CN) δ 168.99, 162.78, 154.99, 149.15, 130.46, 128.68, 122.66, 117.40, 117.18. **LCMS** (*m/z*): 229.9 [M+H]⁺, *t*_R = 2.92 min, standard method. **HRMS** (*m/z*): C₁₃H₁₁NO₃ requires 230.0812 [M+H]⁺; found 230.0814. **HPLC**: *t*_R = 4.63 min, >99 %, standard method.

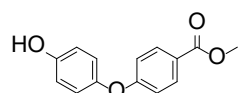
(4-(4-Methoxyphenoxy)phenyl)methanamine (**2.44**)



Prepared according to **general procedure 1** with 4-methoxyphenol (2.00 g, 2.00 eq., 16.1 mmol), Cs₂CO₃ (5.25 g, 2.00 eq., 5.08 mmol), *N,N*-dimethylglycine hydrochloride (300 mg, 0.40 eq., 2.15 mmol), copper(I) iodide (205 mg, 0.20 eq., 1.07 mmol), (4-bromophenyl)methanamine (0.68 mL, 1.0 eq., 5.4 mmol) and degassed 1,4-dioxane (10 mL). The residue was purified by flash column chromatography using a gradient of 99:0:1 EtOAc:MeOH:Et₃N to 94:5:1 EtOAc:MeOH:triethylamine to give **2.44** as a brown solid (322 mg,

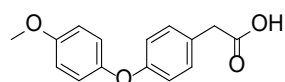
26 %). **¹H NMR** (401 MHz, CD₃CN) δ 7.30 – 7.25 (m, 2H), 6.98 – 6.85 (m, 6H), 3.77 (s, 3H), 3.74 (s, 2H). **¹³C NMR** (101 MHz, CD₃CN) δ 158.02, 156.98, 151.42, 139.56, 129.51, 121.48, 118.47, 115.91, 56.27, 46.22. **LCMS** (m/z): 213.0 [M-NH₂]⁺, t_R = 2.84 min, standard method. **HRMS** (m/z): C₁₄H₁₅NO₂ requires 252.0995 [M+Na]⁺; found 252.0995. **HPLC**: t_R = 4.97 min, >95 %, standard method.

Methyl 4-(4-hydroxyphenoxy)benzoate (**2.47**)¹⁰



To a solution of 4-(4-hydroxyphenoxy)benzoic acid (22 mg, 1.0 eq., 0.10 mmol) dissolved in MeOH (5 mL) was added conc. H₂SO₄ (5 drops) at room temperature. The reaction was stirred at room temperature for 10 min, then 18 h at reflux. Additional H₂SO₄ (3 drops) was added and the reaction refluxed for a further 18 h. After this time, the solution was evaporated to dryness under vacuum and the residue dissolved in DCM (10 mL) and washed with phosphate buffer (aq. 1.0 M, pH 7, 15 mL), brine (10 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **2.47** as a white solid (12 mg, 51 %). **¹H NMR** (401 MHz, CDCl₃) δ 8.01 – 7.94 (m, 2H), 6.99 – 6.90 (m, 4H), 6.90 – 6.83 (m, 2H), 5.23 (bs, 1H), 3.90 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 167.08, 162.98, 152.89, 148.77, 131.81, 123.94, 121.99, 116.72, 116.49, 52.21. **LCMS** (m/z): 242.9 [M-H]⁻, t_R = 3.28 min, standard method. **HRMS** (m/z): C₁₄H₁₃O₄ requires 245.0808 [M+H]⁺; found 245.0812. **HPLC**: t_R = 6.35 min, >99 %, standard method.

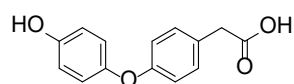
2-(4-(4-Methoxyphenoxy)phenyl)acetic acid (**2.49**, **3.2**, **7.3**)¹¹



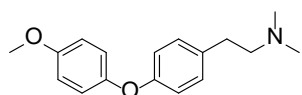
4-Bromophenylacetic acid (500 mg, 1.00 eq. 2.32 mmol), 4-methoxyphenol (433 mg, 1.50 eq. 3.49 mmol), Cs₂CO₃ (2.61 g, 3.45 eq. 8.02 mmol), *N,N*-dimethylglycine hydrochloride (75 mg, 0.23 eq. 0.53 mmol) and copper(I) iodide (51 mg, 0.12 eq. 0.27 mmol) were sealed in a 25 mL RBF with a reflux condenser attached and purged with N₂ (3 times). Degassed 1,4-dioxane (10 mL) was added and the reaction mixture refluxed for 4 days. After this time, the reaction was cooled to room temperature and diluted in hydrochloride (aq., 1 M, 30 mL)

and EtOAc (30 mL). The layers were separated and the aqueous layer further extracted with EtOAc (2×30 mL). The combined organic layers were dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using 75:25 pet spirits:EtOAc. The residue was further purified by flash column chromatography using a gradient of 90:10 pet spirits:EtOAc to 80:20 petroleum spirits:EtOAc. To achieve adequate purity the residue was further purified by flash column chromatography using a gradient of 89:10:1 pet spirits:EtOAc:AcOH to 79:20:1 pet spirits:EtOAc:AcOH to give **2.49** as a yellow solid (54 mg, 9.0 %). **^1H NMR** (401 MHz, CD_3OD) δ 7.25 – 7.19 (m, 2H), 6.97 – 6.89 (m, 4H), 6.89 – 6.83 (m, 2H), 3.78 (s, 3H), 3.56 (s, 2H). **^{13}C NMR** (101 MHz, CD_3OD) δ 175.68, 158.90, 157.44, 151.60, 131.66, 130.14, 121.64, 118.57, 115.93, 56.05, 41.06. **LCMS** (m/z): 257.0 $[\text{M}-\text{H}]^-$, t_{R} = 3.29 min, standard method. **HRMS** (m/z): $\text{C}_{15}\text{H}_{14}\text{O}_4$ requires 259.0965 $[\text{M}+\text{H}]^+$; found 259.0963. **HPLC**: t_{R} = 5.78 min, >99 %, standard method.¹²

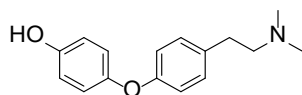
2-(4-(4-Hydroxyphenoxy)phenyl)acetic acid (**2.50**)¹²



To a solution of ether **2.49** (21 mg, 1.0 eq., 81 mmol) dissolved in dry DCM (1 mL) at $-78\text{ }^\circ\text{C}$ was added BBr_3 in heptane (1 M, 0.25 mL, 3.0 eq., 243 mmol) dropwise. The reaction was stirred for 0.5 h at $-78\text{ }^\circ\text{C}$, then allowed to warm to room temperature and stirred for an additional 3 h. The reaction mixture was then quenched with water (5 mL) and diluted with EtOAc (5 mL). The layers were separated and the aqueous layer further extracted with EtOAc (2×5 mL). The combined organic layers were dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 25:75 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **2.50** as a yellow solid (5.8 mg, 29 %). **^1H NMR** (401 MHz, CD_3OD) δ 7.23 – 7.16 (m, 2H), 6.89 – 6.79 (m, 4H), 6.80 – 6.72 (m, 2H), 3.55 (s, 2H). **^{13}C NMR** (101 MHz, CD_3OD) δ 175.89, 159.25, 154.90, 150.56, 131.59, 129.96, 121.91, 118.27, 117.15, 41.17. **LCMS** (m/z): $[\text{M}-\text{H}]^-$, t_{R} = 3.58 min, standard method. **HRMS** (m/z): $\text{C}_{14}\text{H}_{12}\text{O}_4$ requires 244.0730 $[\text{M}+\text{H}]^+$; found 244.0733. **HPLC**: t_{R} = 4.73 min, >95 %, standard method.

2-(4-(4-Methoxyphenoxy)phenyl)-*N,N*-dimethylethan-1-amine (2.53)

Prepared according to **general procedure 1** with 4-[2-(dimethylamino)ethyl]phenol (1.58 g, 1.50 eq., 9.56 mmol), Cs₂CO₃ (4.15 g, 2.00 eq., 12.8 mmol), *N,N*-dimethylglycine hydrochloride (356 mg, 0.400 eq., 2.55 mmol), copper(I) iodide (243 mg, 0.20 eq., 1.27 mmol), 4-bromoanisole (0.80 mL, 1.00 eq., 6.37 mmol) and degassed 1,4-dioxane (15 mL). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 0:100 pet spirits:EtOAc. The fractions were pooled, evaporated to dryness and taken up into hydrochloride (aq., 1 M, 50 mL) and washed with EtOAc (2 × 40 mL). The aqueous layer was basified to pH 14 using NaOH (aq., 1 M, 100 mL) causing precipitation. The solution was then extracted with DCM (3 × 30 mL) and the combined organic layers were evaporated to dryness and purified through a silica plug using DCM, then evaporated to dryness to give **2.53** as a colourless oil (330 mg, 19 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.15 – 7.10 (m, 2H), 6.98 – 6.93 (m, 2H), 6.89 – 6.84 (m, 4H), 3.79 (s, 3H), 2.77 – 2.71 (m, 2H), 2.55 – 2.48 (m, 2H), 2.29 (s, 6H). **¹³C NMR** (101 MHz, CDCl₃) APT δ 156.81 (C), 155.88 (C), 150.64 (C), 134.63 (C), 129.83 (CH), 120.69 (CH), 117.87 (CH), 114.95 (CH), 61.83 (CH₂), 55.79 (CH₃), 45.62 (CH₃), 33.71 (CH₂). **LCMS** (*m/z*): 272.0 [M+H]⁺, *t_R* = 2.90 min, standard method. **HRMS** (*m/z*): C₁₇H₂₁NO₂ requires 272.1645 [M+H]⁺; found 272.1643. **HPLC**: *t_R* = 5.58 min, >99 %, standard method.

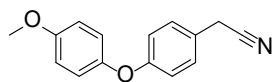
4-(4-(2-(Dimethylamino)ethyl)phenoxy)phenol (2.54)

Prepared according to **general procedure 2** with ether **2.53** (300 mg, 1.0 eq., 1.1 mmol), DCM (10 mL) and BBr₃ in DCM (1.0 M, 3.3 mL, 3.0 eq., 3.3 mmol) at -78 °C. The residue was purified by flash column chromatography using a gradient of 80:20 pet spirits:EtOAc to 0:100 pet spirits:EtOAc. The residue was further purified by trituration with cold pet spirits and DCM to give **2.54** as a white solid (15 mg, 5.3 %). **¹H NMR** (401 MHz, CD₃OD) δ 7.27 – 7.24 (m, 2H), 6.90 – 6.82 (m, 4H), 6.81 – 6.75 (m, 2H), 3.38 – 3.33 (m, 2H), 3.02 (dd, *J* = 9.6, 6.9 Hz, 2H), 2.93 (s, 6H). **¹³C NMR** (101 MHz, CD₃OD) δ 159.70, 155.11, 150.29, 131.09, 130.84, 122.03, 118.70, 117.22, 59.98, 31.02, 29.62. **LCMS** (*m/z*): 258.0 [M+H]⁺, *t_R* = 2.79

min, standard method. **HRMS** (m/z): $C_{16}H_{19}NO_2$ requires 258.1489 $[M+H]^+$; found 258.1485.

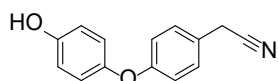
HPLC: t_R = 4.61 min, >99 %, standard method.

2-(4-(4-Methoxyphenoxy)phenyl)acetonitrile (**2.56**)



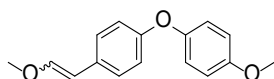
Prepared according to **general procedure 1** with 2-(4-bromophenyl)acetonitrile (500 mg, 1.00 eq., 2.55 mmol), 4-methoxyphenol (475 mg, 1.50 eq., 3.83 mmol), CS_2CO_3 (1.66 g, 2.00 eq., 5.10 mmol), *N,N*-dimethylglycine hydrochloride (143 mg, 0.400 eq., 1.02 mmol), copper(I) iodide (97 mg, 0.20 eq., 0.51 mmol) and degassed 1,4-dioxane (5 mL). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 70:30 pet spirits:EtOAc to give **2.56** as a yellow solid (330 mg, 54 %). **1H NMR** (401 MHz, $CDCl_3$) δ 7.25 – 7.21 (m, 2H), 7.00 – 6.87 (m, 6H), 3.81 (s, 3H), 3.70 (s, 2H). **^{13}C NMR** (101 MHz, $CDCl_3$) APT δ 158.66 (C), 156.35 (C), 149.78 (C), 129.40 (CH), 123.73 (C), 121.13 (CH), 118.14 (C), 118.14 (CH), 115.13 (CH), 55.81 (CH_3), 23.05 (CH_2). **LCMS** (m/z): 237.7 $[M-H]^-$, t_R = 3.48 min, standard method. **HRMS** (m/z): $C_{15}H_{13}NO_2$ requires 240.1019 $[M+H]^+$; found 240.1014. **HPLC**: t_R = 7.02 min, >99 %, standard method.

2-(4-(4-Hydroxyphenoxy)phenyl)acetonitrile (**2.57**)



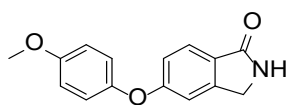
Prepared according to **general procedure 2** with ether **2.56** (188 mg, 1.00 eq., 0.790 mmol), DCM (8 mL) and BBr_3 in hexane (1.0 M, 2.36 mL, 3.00 eq., 2.36 mmol) at $-78^\circ C$. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 50:50 pet spirits:EtOAc. The residue was then recrystallised from DCM/pet spirits to give **2.57** as a white solid (55 mg, 31 %). **1H NMR** (401 MHz, $CDCl_3$) δ 7.25 – 7.21 (m, 2H), 6.95 – 6.89 (m, 4H), 6.85 – 6.80 (m, 2H), 5.40 (bs, 1H), 3.71 (s, 2H). **^{13}C NMR** (101 MHz, $CDCl_3$) APT δ 158.53 (C), 152.29 (C), 149.56 (C), 129.31 (CH), 123.46 (C), 121.21 (CH), 118.12 (C), 118.00 (CH), 116.50 (CH), 22.92 (CH_2). **LCMS** (m/z): 223.9 $[M-H]^-$, t_R = 3.26 min, standard method. **HRMS** (m/z): $C_{14}H_{11}NO_2$ requires 226.0863 $[M+H]^+$; found 226.0851. **HPLC**: t_R = 5.91 min, >99 %, standard method.

1-Methoxy-4-(4-(2-methoxyvinyl)phenoxy)benzene (**2.59**).



To a solution of aldehyde **2.33** (280 mg, 1.00 eq., 1.23 mmol), methoxymethyl(triphenyl)phosphonium chloride (631 mg, 1.50 eq., 1.84 mmol) dissolved in dry THF (6 mL) was added KOtBu (206 mg, 1.50 eq., 1.84 mmol) at room temperature. The reaction mixture was stirred for 18 h at room temperature. The reaction was quenched with saturated NH₄Cl (10 mL) and diluted in Et₂O (25 mL). The layers were separated and the aqueous further extracted with Et₂O (25 mL). The combined organic layers were washed with brine (20 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 80:20 pet spirits:EtOAc to give **2.59** as a 2:3 mixture of *cis:trans* isomers (114 mg, 36 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.48 – 7.41 (m, 0.8H), 7.12 – 7.05 (m, 1.2H), 6.93 – 6.85 (m, 2.6H), 6.84 – 6.76 (m, 4.0H), 6.01 (d, *J* = 7.0 Hz, 0.4H), 5.72 (d, *J* = 13.0 Hz, 0.6H), 5.12 (d, *J* = 7.0 Hz, 0.4H), 3.73 (d, *J* = 0.6 Hz, 3H), 3.69 (s, 1.2H), 3.60 (s, 1.8H). **¹³C NMR** (101 MHz, CDCl₃) δ 156.47, 156.19, 155.75, 155.71, 150.54, 148.15, 146.97, 130.90, 130.64, 129.47, 126.26, 120.46, 120.45, 118.12, 117.64, 114.83, 114.79, 105.05, 104.46, 60.57, 56.52, 55.67. **LCMS** (*m/z*): Did not ionise [M+H]⁺, *t_R* = 3.612 min, standard method. **HRMS** (*m/z*): C₁₆H₁₆O₃ requires 257.1172 [M+H]⁺; found 257.1173. **HPLC**: *t_R* = 7.86 and 7.98 min, >99 %, standard method.

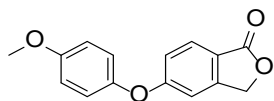
5-(4-Methoxyphenoxy)isoindolin-1-one (**2.61**)



5-Bromoisindolin-1-one (50 mg, 1.0 eq., 0.24 mmol), 4-methoxyphenol (132 mg, 1.50 eq., 1.06 mmol), Cs₂CO₃ (230 mg, 3.00 eq., 0.71 mmol), *N,N*-dimethylglycine hydrochloride (13 mg, 0.40 eq., 0.09 mmol) and copper(I) iodide (9 mg, 0.20 eq., 0.05 mmol) were sealed in a crimp sealed 8 mL vial and purged with N₂ (3 times). Degassed 1,4-dioxane (5 mL) was added and the reaction mixture heated at 120 °C for 10 days. After this time, the reaction was worked up according to **general procedure 1**. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 90:10 pet spirits:EtOAc to give **2.61** as a white solid (10 mg, 17 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.77 (d, *J* = 8.4 Hz, 1H), 7.15 (bs, 1H), 7.06 – 7.00 (m, 3H), 6.95 – 6.90 (m, 3H), 4.37 (s, 2H), 3.82 (s, 3H). **¹³C NMR** (101 MHz,

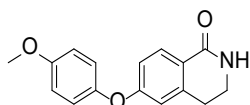
CDCl_3) δ 171.64, 162.59, 156.78, 149.01, 146.06, 126.06, 125.37, 121.77, 117.51, 115.25, 111.10, 55.80, 45.56. **LCMS**: (m/z): 256.0 $[\text{M}+\text{H}]^+$, t_R = 3.14 min, standard method. **HPLC**: t_R = 5.83 min, >99 %, standard method.

5-(4-Methoxyphenoxy)isobenzofuran-1(3H)-one (2.62)



5-Bromo-3*H*-isobenzofuran-1-one (151 mg, 1.00 eq., 0.707 mmol), 4-methoxyphenol (132 mg, 1.50 eq., 1.06 mmol), Cs_2CO_3 (461 mg, 2.00 eq., 1.42 mmol), *N,N*-dimethylglycine hydrochloride (39 mg, 0.40 eq., 0.28 mmol) and copper(I) iodide (27 mg, 0.20 eq., 0.14 mmol) were sealed in a crimp sealed 8 mL vial and purged with N_2 (3 times). Degassed 1,4-dioxane (5 mL) was added and the reaction mixture heated at 105 °C for 14 days. * **Note the reaction was conducted above standard pressure and caution should be used with suitable protective equipment.** The reaction was then cooled and diluted with H_2O (20 mL) and EtOAc (20 mL). The layers were separated and the aqueous layer further extracted with EtOAc (20 mL). The combined organic layers were washed with brine (20 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 95:5 pet spirits:EtOAc to 70:30 pet spirits:EtOAc to give **2.62** as a brown solid (36 mg, 20 %). **^1H NMR** (401 MHz, CDCl_3) δ 7.80 (d, J = 8.5 Hz, 1H), 7.09 – 6.97 (m, 3H), 6.96 – 6.90 (m, 2H), 6.87 (d, J = 1.4 Hz, 1H), 5.19 (s, 2H), 3.82 (s, 3H). **^{13}C NMR** (101 MHz, CDCl_3) δ 170.69, 164.43, 157.12, 149.28, 148.18, 127.49, 121.95, 119.24, 118.45, 115.35, 109.09, 69.17, 55.76. **LCMS** (m/z): 256.9 $[\text{M}+\text{H}]^+$, t_R = 3.38 min, standard method. **HRMS** (m/z): $\text{C}_{15}\text{H}_{12}\text{O}_4$ requires 257.0808 $[\text{M}+\text{H}]^+$; found 257.0815. **HPLC**: t_R = 6.49 min, >99 %, standard method.

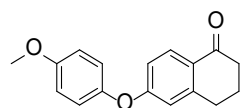
6-(4-Methoxyphenoxy)-3,4-dihydroisoquinolin-1(2H)-one (2.64)



Prepared according to **general procedure 1** with 6-bromo-3,4-dihydro-2*H*-isoquinolin-1-one (150 mg, 1.00 eq., 0.664 mmol), 4-methoxyphenol (124 mg, 1.50 eq., 0.995 mmol), Cs_2CO_3 (432 mg, 2.00 eq., 1.32 mmol), *N,N*-dimethylglycine hydrochloride (37 mg, 0.40 eq., 0.27 mmol), copper(I) iodide (25 mg, 0.20 eq., 0.13 mmol) and degassed 1,4-

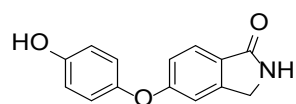
dioxane (5 mL). The residue was purified by flash column chromatography using a gradient of 50:50 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **2.64** as a brown solid (114 mg, 64 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.97 (d, *J* = 8.6 Hz, 1H), 7.45 (bs, 1H), 7.02 – 6.95 (m, 2H), 6.93 – 6.86 (m, 2H), 6.83 (dd, *J* = 8.6, 2.5 Hz, 1H), 6.69 (d, *J* = 2.4 Hz, 1H), 3.79 (s, 3H), 3.52 (td, *J* = 6.7, 2.8 Hz, 2H), 2.88 (t, *J* = 6.6 Hz, 2H). **¹³C NMR** (101 MHz, CDCl₃) δ 166.60, 161.93, 156.56, 148.72, 141.24, 130.04, 123.15, 121.65, 115.32, 115.05, 114.96, 55.65, 40.07, 28.57. **LCMS** (*m/z*): 269.9 [M+H]⁺, *t_R* = 3.22 min, standard method. **HRMS** (*m/z*): C₁₆H₁₅NO₃ requires 270.1125 [M+H]⁺; found 270.1130. **HPLC**: *t_R* = 6.16 min, >99 %, standard method.

6-(4-Methoxyphenoxy)-3,4-dihydronaphthalen-1(2*H*)-one (2.65)



6-Hydroxy-1-tetralone (190 mg, 1.46 eq., 1.17 mmol), Cs₂CO₃ (523 mg, 2.00 eq., 1.60 mmol), *N,N*-dimethylglycine hydrochloride (45 mg, 0.40 eq., 0.32 mmol) and copper(I) iodide (31 mg, 0.20 eq., 0.16 mmol) were sealed in a crimp sealed 8 mL vial and purged with N₂ (3 times). 4-Bromoanisole (0.100 mL, 1.00 eq., 0.802 mmol) dissolved in degassed 1,4-dioxane (5 mL) was added and the reaction mixture heated at 105 °C for 14 days. After this time, the reaction was worked up according to **general procedure 1**. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 87:13 pet spirits:EtOAc to give **2.65** as a white solid (11 mg, 4.9 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.99 (d, *J* = 8.7 Hz, 1H), 7.04 – 6.98 (m, 2H), 6.96 – 6.89 (m, 2H), 6.82 (dd, *J* = 8.7, 2.5 Hz, 1H), 6.69 (d, *J* = 2.4 Hz, 1H), 3.83 (s, 3H), 2.87 (app t, *J* = 6.1 Hz, 2H), 2.63 – 2.58 (m, 2H), 2.14 – 2.06 (m, 2H). **¹³C NMR** (101 MHz, CDCl₃) δ 197.28, 163.12, 156.84, 148.53, 147.13, 129.85, 127.44, 121.93, 115.58, 115.40, 115.20, 55.78, 39.05, 30.16, 23.42. **LCMS** (*m/z*): 269.0 [M+H]⁺, *t_R* = 3.51 min, standard method. **HRMS** (*m/z*): C₁₇H₁₆O₃ requires 269.1172 [M+H]⁺; found 269.1178. **HPLC**: *t_R* = 7.29 min, >99 %, standard method.

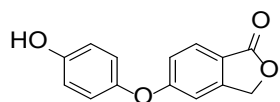
5-(4-Hydroxyphenoxy)isoindolin-1-one (2.66)



Prepared according to **general procedure 2** with ether **2.61** (10 mg, 1.0 eq., 0.04 mmol), DCM (4 mL) and BBr₃ in hexane (1.0 M, 0.20 mL, 5.1 eq., 0.20

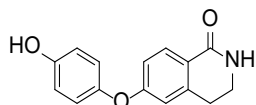
mmol) at -20 °C. The residue was purified by flash column chromatography using a gradient of 100:0 EtOAc:MeOH to 95:5 EtOAc:MeOH. The residue was triturated with cold pet spirits (5 mL) and dried under vacuum to give **2.66** as a white solid (5 mg, 53 %). **¹H NMR** (401 MHz, CD₃OD) δ 7.70 (d, *J* = 8.3 Hz, 1H), 7.03 (dd, *J* = 8.4, 2.2 Hz, 1H), 7.01 – 6.99 (m, 1H), 6.96 – 6.91 (m, 2H), 6.86 – 6.81 (m, 2H), 4.37 (s, 2H). **¹³C NMR** (101 MHz, CD₃OD) δ 173.41, 164.37, 155.84, 149.26, 148.14, 126.79, 125.81, 122.75, 118.20, 117.45, 112.01, 46.64. **LCMS** (*m/z*): 241.9 [M+H]⁺, *t_R* = 2.94 min, standard method. **HRMS** (*m/z*): C₁₄H₁₁NO₃ requires 242.0812 [M+H]⁺; found 242.0808. **HPLC**: *t_R* = 4.75 min, >99 %, standard method.

5-(4-Hydroxyphenoxy)isobenzofuran-1(3*H*)-one (2.67)



Prepared according to **general procedure 2** with ether **2.62** (60 mg, 1.0 eq., 0.23 mmol), DCM (4 mL) and BBr₃ in heptane (1.0 M, 0.42 mL, 3.0 eq., 0.42 mmol) at -20 °C. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 55:45 pet spirits:EtOAc to give **2.67** as a white solid (40 mg, 71 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 9.53 (s, 1H), 7.80 (d, *J* = 8.4 Hz, 1H), 7.10 (dd, *J* = 8.5, 2.2 Hz, 1H), 7.05 – 6.97 (m, 3H), 6.88 – 6.83 (m, 2H), 5.30 (s, 2H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 170.03, 163.85, 154.88, 150.17, 146.23, 126.87, 121.91, 118.39, 117.87, 116.55, 109.29, 69.37. **LCMS** (*m/z*): 243.0 [M+H]⁺, *t_R* = 3.13 min, standard method. **HRMS** (*m/z*): C₁₄H₁₀O₄ requires 243.0652 [M+H]⁺; found 243.0655. **HPLC**: *t_R* = 5.39 min, >99 %, standard method.

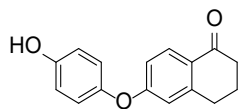
6-(4-Hydroxyphenoxy)-3,4-dihydroisoquinolin-1(2*H*)-one (2.68)



Prepared according to **general procedure 2** with ether **2.64** (29 mg, 1.0 eq., 0.11 mmol), DCM (4 mL) and BBr₃ in heptane (1.0 M, 0.32 mL, 3.0 eq., 0.32 mmol) at -20 °C. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **2.68** as a white solid (19 mg, 69 %). **¹H NMR** (401 MHz, CD₃OD) δ 7.86 (d, *J* = 8.6 Hz, 1H), 6.94 – 6.87 (m, 2H), 6.85 – 6.79 (m, 3H), 6.74 (d, *J* = 2.4 Hz, 1H), 3.50 – 3.42 (m, 2H), 2.89 (t, *J* = 6.7 Hz, 2H). **¹³C NMR** (101 MHz, CD₃OD) δ 168.21, 164.03, 155.92, 148.93, 143.27, 130.74, 123.61, 122.79, 117.44, 116.06, 115.83, 40.79, 29.26. **LCMS**

(m/z): 253.8 $[M-H]^-$, t_R = 2.10 min, standard method. **HRMS** (m/z): $C_{15}H_{13}NO_3$ requires 256.0968 $[M+H]^+$; found 256.0972. **HPLC**: t_R = 5.05 min, >99 %, standard method.

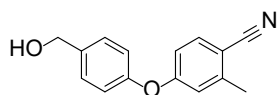
6-(4-Hydroxyphenoxy)-3,4-dihydronaphthalen-1(2H)-one (2.69)



Prepared according to **general procedure 2** with ether **2.65** (93 mg, 1.0 eq., 0.35 mmol), DCM (5 mL) and BBr_3 in heptane (1.0 M, 1.0 mL, 3.0 eq., 1.0 mmol) at

0 °C. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 50:50 pet spirits:EtOAc. The residue was further purified by trituration with cold pet spirits and MeOH to give **2.69** as a white solid (8 mg, 9.0 %). **1H NMR** (401 MHz, d_6 -DMSO) δ 9.51 (s, 1H), 7.84 (d, J = 8.7 Hz, 1H), 6.99 – 6.91 (m, 2H), 6.86 – 6.77 (m, 3H), 6.74 (d, J = 2.5 Hz, 1H), 2.86 (app t, J = 6.0 Hz, 2H), 2.56 – 2.51 (m, 2H), 2.04 – 1.95 (m, 2H). **^{13}C NMR** (101 MHz, d_6 -DMSO) δ 196.13, 162.59, 154.64, 147.31, 146.31, 129.06, 126.78, 121.80, 116.45, 115.07, 114.72, 38.37, 29.16, 22.83. **LCMS** (m/z): 255.0 $[M+H]^+$, t_R = 3.24 min, standard method. **HRMS** (m/z): $C_{16}H_{14}O_3$ requires 255.1016 $[M+H]^+$; found 255.1020. **HPLC**: t_R = 6.15 min, >99 %, standard method.

4-(4-(Hydroxymethyl)phenoxy)-2-methylbenzonitrile (3.4)

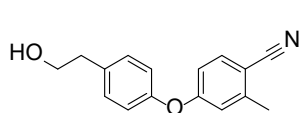


Prepared according to **general procedure 1** with 4-bromo-2-methylbenzonitrile (100 mg, 1.00 eq., 0.510 mmol), 4-hydroxybenzyl alcohol

(95 mg, 1.5 eq., 0.76 mmol), copper(I) iodide (19 mg, 0.20 eq., 0.10 mmol), *N,N*-dimethylglycine hydrochloride (28 mg, 0.40 eq., 0.20 mmol), CS_2CO_3 (332 mg, 2.00 eq., 1.02 mmol) and degassed DMF (5 mL). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **3.4** as a colourless oil (63 mg, 52 %). **1H NMR** (401 MHz, $CDCl_3$) δ 7.43 (d, J = 8.5 Hz, 1H), 7.34 – 7.29 (m, 2H), 6.98 – 6.93 (m, 2H), 6.77 (d, J = 2.4 Hz, 1H), 6.72 (dd, J = 8.5, 2.3 Hz, 1H), 4.62 (s, 2H), 2.39 (s, 3H), 2.19 (s, 1H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 161.51, 154.38, 144.56, 137.74, 134.41, 128.96, 120.49, 118.85, 118.25, 115.37, 106.41, 64.60, 20.67. **LCMS** (m/z): 237.9 $[M-H]^-$, t_R = 3.26 min, standard method. **HRMS** (m/z):

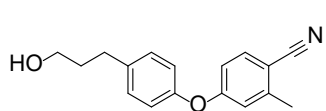
$C_{15}H_{13}NO_2$ requires 240.1019 $[M+H]^+$; found 240.1010. **HPLC**: t_R = 6.19 min, >99 %, standard method.

4-(4-(2-Hydroxyethyl)phenoxy)-2-methylbenzonitrile (3.5)



Prepared according to **general procedure 1** with 4-bromo-2-methylbenzonitrile (800 mg, 1.00 eq., 4.10 mmol), 4-hydroxyphenethyl alcohol (840 mg, 1.50 eq., 6.12 mmol), CS_2CO_3 (2.65 g, 2.00 eq., 8.16 mmol), *N,N*-dimethylglycine hydrochloride (228 mg, 0.40 eq., 1.63 mmol), copper(I) iodide (155 mg, 0.20 eq., 0.816 mmol) and degassed 1,4-dioxane (10 mL). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **3.5** as a colourless oil (807 mg, 78 %). **1H NMR** (401 MHz, $CDCl_3$) δ 7.32 (d, J = 8.5 Hz, 1H), 7.10 – 7.02 (m, 2H), 6.83 – 6.76 (m, 2H), 6.66 (d, J = 2.4 Hz, 1H), 6.60 (dd, J = 8.5, 2.4 Hz, 1H), 3.69 (t, J = 6.6 Hz, 2H), 2.69 (t, J = 6.6 Hz, 2H), 2.29 (s, 3H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 161.62, 153.51, 144.51, 135.53, 134.39, 130.75, 120.54, 118.79, 118.29, 115.28, 106.31, 63.58, 38.54, 20.69. **LCMS**: (m/z): 254.0 $[M+H]^+$, t_R = 3.42 min, standard method. **HRMS** (m/z): $C_{16}H_{15}NO_2$ requires 254.1176 $[M+H]^+$; found 254.1166. **HPLC**: t_R = 6.54 min, >99 %, standard method.

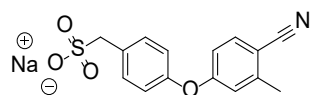
4-(4-(3-Hydroxypropyl)phenoxy)-2-methylbenzonitrile (3.6)



Prepared according to **general procedure 1** with 3-(4-hydroxyphenyl)-1-propanol (116 mg, 1.50 eq., 0.765 mmol), 4-bromo-2-methylbenzonitrile (100 mg, 1.00 eq., 0.510 mmol), copper(I) iodide (19 mg, 0.20 eq., 0.10 mmol), *N,N*-dimethylglycine hydrochloride (28 mg, 0.40 eq., 0.20 mmol), CS_2CO_3 (332 mg, 2.00 eq., 1.02 mmol) and degassed DMF (5 mL). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **3.6** as a colourless oil (104 mg, 76 %). **1H NMR** (401 MHz, $CDCl_3$) δ 7.49 (d, J = 8.5 Hz, 1H), 7.24 – 7.19 (m, 2H), 6.98 – 6.93 (m, 2H), 6.83 (d, J = 2.4 Hz, 1H), 6.77 (dd, J = 8.5, 2.2 Hz, 1H), 3.68 (t, J = 6.4 Hz, 2H), 2.75 – 2.67 (m, 2H), 2.46 (s, 3H), 1.99 (s, 1H), 1.94 – 1.85 (m, 2H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 161.75, 152.88, 144.43, 138.79, 134.32, 130.08, 120.40, 118.64, 118.29, 115.10, 106.07, 61.98, 34.23, 31.43, 20.63. **LCMS** (m/z):

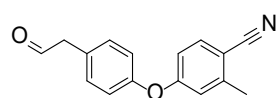
267.9 $[M+H]^+$, $t_R = 3.38$ min, standard method. **HRMS** (m/z): $C_{17}H_{17}NO_2$ requires 268.1332 $[M+H]^+$; found 268.1326. **HPLC**: $t_R = 6.75$ min, >99 %, standard method.

4-(4-Cyano-3-methylphenoxy)phenylmethanesulfonate sodium salt (3.7)



To a solution of alcohol **3.4** (178 mg, 1.00 eq., 0.744 mmol) dissolved in dry DCM (5 mL) was added *N*-bromosuccinimide (265 mg, 2.00 eq., 1.49 mmol) and PPh_3 (390 mg, 2.00 eq., 1.49 mmol) at room temperature. The reaction was stirred for 0.5 h at room temperature then 5 h at 40 °C. After this time the reaction mixture was diluted with H_2O (10 mL) and DCM (20 mL). The layers were separated and the aqueous layer further extracted with DCM (20 mL). The combined organic layers were washed with brine (10 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 85:15 pet spirits:EtOAc to give the bromo intermediate which was used without further purification. To a solution of 4-[4-(bromomethyl)phenoxy]-2-methylbenzonitrile (100 mg, 1.0 eq., 0.33 mmol) dissolved in 1:1 EtOH/ H_2O (4 mL) was added Na_2SO_3 (104 mg, 2.5 eq., 0.827 mmol) at room temperature. The reaction was stirred for 0.5 h at room temperature then 5 h at 40 °C. The solvent was removed under vacuum and the residue was dissolved in hot EtOH and filtered to remove inorganic salts. The filtrate was concentrated under vacuum and further purified by trituration with C hydrochloride₃ giving **3.7** as a light brown solid (48 mg, 44 %). **1H NMR** (400 MHz, D_2O) δ 7.68 (d, $J = 8.6$ Hz, 1H), 7.47 (d, $J = 8.5$ Hz, 2H), 7.13 (d, $J = 8.5$ Hz, 2H), 6.99 – 6.95 (m, 2H), 4.19 (s, 2H), 2.44 (s, 3H). **^{13}C NMR** (101 MHz, D_2O) δ 164.01, 157.20, 147.95, 137.45, 134.88, 131.26, 123.10, 121.93, 121.37, 118.21, 108.01, 58.88, 22.34. **LCMS** (m/z): 301.9 $[M-H]^-$, $t_R = 3.36$ min, standard method. **HRMS** (m/z): $C_{15}H_{13}NO_4S$ requires 304.0638 $[M+H]^+$; found 304.0633. **HPLC**: $t_R = 6.66$ min, 96 %, standard method.

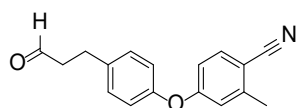
2-Methyl-4-(4-(2-oxoethyl)phenoxy)benzonitrile (3.8)



Prepared according to **general procedure 3** with alcohol **3.5** (543 mg, 1.00 eq., 2.14 mmol), DCM (20 mL) and Dess-Martin periodinane (1.04 g 1.15 eq.,

2.45 mmol). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **3.8** as a colourless oil (373 mg, 69 %). **¹H NMR** (401 MHz, CDCl₃) δ 9.78 (t, *J* = 2.1 Hz, 1H), 7.52 (d, *J* = 8.5 Hz, 1H), 7.27 – 7.22 (m, 2H), 7.07 – 7.02 (m, 2H), 6.87 (d, *J* = 2.5 Hz, 1H), 6.81 (dd, *J* = 8.5, 2.5 Hz, 1H), 3.73 (d, *J* = 2.1 Hz, 2H), 2.47 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 198.96, 161.11, 154.27, 144.42, 134.28, 131.35, 128.47, 120.65, 118.90, 118.04, 115.38, 106.50, 49.64, 20.52. **LCMS** (*m/z*): 249.9 [M-H]⁻, *t_R* = 3.40 min, standard method. **HRMS** (*m/z*): C₁₆H₁₃NO₂ requires 252.1019 [M+H]⁺; found 252.1018. **HPLC**: *t_R* = 6.63 min, 94 %, standard method.

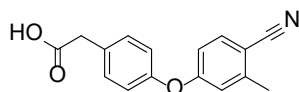
2-Methyl-4-(4-(3-oxopropyl)phenoxy)benzonitrile (**3.9**)



Prepared according to **general procedure 3** with alcohol **3.6** (554 mg, 1.00 eq., 2.07 mmol), DCM (20 mL) and Dess-Martin periodinane (1.32 g, 1.50

eq., 3.11 mmol). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **3.9** as a colourless oil (170 mg, 31 %). **¹H NMR** (401 MHz, CDCl₃) δ 9.80 (t, *J* = 1.3 Hz, 1H), 7.49 – 7.44 (d, *J* = 8.5 Hz, 1H), 7.23 – 7.17 (m, 2H), 6.96 – 6.92 (m, 2H), 6.83 – 6.80 (dd, *J* = 8.4, 2.3 Hz, 1H), 6.77 – 6.74 (d, *J* = 2.4 Hz, 1H), 2.94 (t, *J* = 7.4 Hz, 2H), 2.81 – 2.76 (m, 2H), 2.44 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 201.31, 161.45, 153.18, 144.33, 137.13, 134.22, 129.95, 120.43, 118.63, 118.11, 115.11, 106.17, 45.10, 27.30, 20.51. **LCMS** (*m/z*): did not ionise, *t_R* = 3.36 min, standard method. **HRMS** (*m/z*): C₁₇H₁₅NO₂ requires 266.1176 [M+H]⁺; found 266.1176. **HPLC**: *t_R* = 6.61 min, 94 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetic acid (**3.10**, **4.1**, **5.1**, **6.19**, **7.4**)

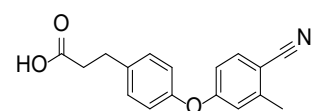


Method A: To a solution of aldehyde **3.8** (300 mg, 1.00 eq., 1.20 mmol) and H₂O₂ (30 %, 2.0 mL, 16 eq., 19.6 mmol) dissolved in dry MeCN (20 mL) was added NaH₂PO₄ (179 mg, 1.10 eq., 1.31 mmol) and NaClO₂ (540 mg, 5.0 eq., 6.00 mmol) dissolved in H₂O (15 mL) at 10 °C. The reaction mixture was stirred for 1 h at 10 °C. After this time, Na₂SO₃ (1.60 g, 10.6 eq., 12.7 mmol) was added to quench any unreacted oxidant and the MeCN was removed under vacuum. The remaining aqueous layer was acidified with hydrochloride (aq., 1 M, 25

mL) and extracted with EtOAc (3×30 mL). The combined organic layers were washed with brine (40 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **3.10** as a white solid (149 mg, 47 %).

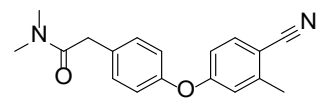
Method B: Ester **3.16** (1.02 g, 1.00 eq., 3.64 mmol) and LiOH (440 mg, 5.00 eq., 18.2 mmol) were dissolved in a 1:1 THF:H₂O mixture (16 mL) and stirred for 18 h at room temperature. After this time, the THF was removed under vacuum and the aqueous layer acidified with hydrochloride (aq., 36 % w/v, 0.5 mL) and diluted in EtOAc (30 mL). The layers were separated and the aqueous layer further extracted with EtOAc (2×30 mL). The combined organic layers were washed with brine (20 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness to give **3.10** as a white solid (925 mg, 95 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 12.38 (bs, 1H), 7.75 (d, *J* = 8.6 Hz, 1H), 7.37 – 7.32 (m, 2H), 7.10 – 7.05 (m, 2H), 7.03 (d, *J* = 2.5 Hz, 1H), 6.88 (dd, *J* = 8.6, 2.5 Hz, 1H), 3.60 (s, 2H), 2.44 (s, 3H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 172.63, 160.95, 153.20, 144.36, 134.76, 131.84, 131.35, 119.96, 118.74, 117.94, 115.37, 105.66, 49.33, 20.00. **LCMS** (*m/z*): 265.9 [M-H]⁻, *t*_R = 3.28 min, standard method. **HRMS** (*m/z*): C₁₆H₁₃NO₃ requires 268.0968 [M+H]⁺; found 268.0965. **HPLC**: *t*_R = 6.31 min, >99 %, standard method.

3-(4-(4-Cyano-3-methylphenoxy)phenyl)propanoic acid (**3.11**)

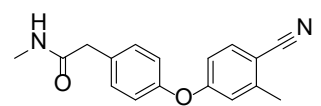
 To a solution of aldehyde **3.9** (140 mg, 1.00 eq., 0.528 mmol) and H₂O₂ (30 %, 0.5 mL, 10.4 eq., 5.5 mmol) dissolved in dry MeCN (20 mL) was added KH₂PO₄ (79 mg, 1.1 eq., 0.58 mmol) and NaClO₂ (239 mg, 5.00 eq., 2.64 mmol) dissolved in H₂O (15 mL) at 10 °C. The reaction mixture was stirred for 0.5 h at 10 °C then allowed to warm to room temperature and stirred for an additional 3 h. Na₂SO₃ (1.60 g, 24.1 eq., 12.7 mmol) was added to quenched any unreacted oxidant. The MeCN was removed under vacuum and the remaining aqueous layer was acidified with hydrochloride (aq., 1 M, 25 mL) and diluted with EtOAc (30 mL). The layers were separated and the aqueous layer further extracted with EtOAc (2×30 mL). The combined organic layers were washed with brine (40 mL) and dried over anhydrous Na_2SO_4 , filtered

and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **3.11** as a white solid (149 mg, 47 %). **¹H NMR** (401 MHz, CD₃OD) δ 7.60 (d, J = 8.6 Hz, 1H), 7.33 – 7.29 (m, 2H), 7.02 – 6.97 (m, 2H), 6.92 (d, J = 2.4 Hz, 1H), 6.84 (dd, J = 8.6, 2.1 Hz, 1H), 2.98 – 2.90 (m, 2H), 2.65 – 2.59 (m, 2H), 2.46 (s, 3H). **¹³C NMR** (101 MHz, CD₃OD) δ 176.56, 163.33, 154.68, 145.73, 139.25, 135.56, 131.20, 121.49, 119.72, 118.97, 116.30, 107.07, 36.67, 31.31, 20.52. **LCMS** (m/z): 279.9 [M-H]⁻, t_R = 3.28 min, standard method. **HRMS** (m/z): C₁₇H₁₅NO₃ requires 282.1125 [M+H]⁺; found 282.1126. **HPLC**: t_R = 6.66 min, >99 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N,N*-dimethylacetamide (**3.12**, 4.3)

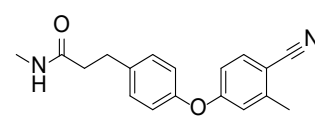
 Prepared according to **general procedure 4** with **3.10** (37 mg, 1.0 eq., 0.14 mmol), HCTU (57 mg, 1.0 eq., 0.14 mmol), dimethylamine hydrochloride (12 mg, 1.1 eq., 0.15 mmol), MeCN (5 mL) and DIPEA (0.10 mL, 4.2 eq., 0.57 mmol). The residue was purified by flash column chromatography using a gradient of 100:0 DCM:MeOH to 95:5 DCM:MeOH to give **3.12** as a white solid (23 mg, 56 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.51 (d, J = 8.5 Hz, 1H), 7.30 – 7.26 (m, 2H), 7.02 – 6.97 (m, 2H), 6.86 (d, J = 2.4 Hz, 1H), 6.81 (dd, J = 8.5, 2.3 Hz, 1H), 3.71 (s, 2H), 3.06 (s, 3H), 2.99 (s, 3H), 2.48 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 170.90, 161.53, 153.81, 144.56, 134.43, 132.00, 130.85, 120.59, 118.90, 118.32, 115.43, 106.50, 40.03, 37.82, 35.82, 20.73. **LCMS** (m/z): 294.9 [M+H]⁺, t_R = 3.29 min, standard method. **HRMS** (m/z): C₁₈H₁₈N₂O₂ requires 295.1441 [M+H]⁺; found 295.1441. **HPLC**: t_R = 6.45 min, >99 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N*-methylacetamide (**3.13**, 4.2)

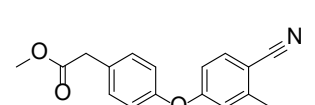
 Prepared according to **general procedure 4** with carboxylic acid **3.10** (51 mg, 1.0 eq., 0.19 mmol), HCTU (79 mg, 1.0 eq., 0.19 mmol), methylamine hydrochloride (29 mg, 2.2 eq., 0.43 mmol), MeCN (5 mL) and DIPEA (0.10 mL, 3.0 eq., 0.57 mmol). The residue was purified by flash column chromatography using a gradient of 100:0 DCM:MeOH to 92:8 DCM:MeOH to give **3.13** as a white solid (27 mg, 50 %). **¹H NMR** (401 MHz, CD₃OD) δ 7.61

(d, $J = 8.6$ Hz, 1H), 7.39 – 7.34 (m, 2H), 7.06 – 7.02 (m, 2H), 6.96 (d, $J = 2.4$ Hz, 1H), 6.87 (dd, $J = 8.6, 2.5$ Hz, 1H), 3.53 (s, 2H), 2.75 (s, 3H), 2.48 (s, 3H). ^{13}C NMR (101 MHz, CD_3OD) δ 173.01, 161.75, 153.94, 144.37, 134.19, 132.49, 130.63, 120.15, 118.44, 117.54, 115.06, 105.85, 41.58, 25.16, 19.13. **LCMS** (m/z): 280.9 $[\text{M}+\text{H}]^+$, $t_{\text{R}} = 3.22$ min, standard method. **HRMS** (m/z): $\text{C}_{17}\text{H}_{16}\text{N}_2\text{O}_2$ requires 281.1285 $[\text{M}+\text{H}]^+$; found 281.1282. **HPLC**: $t_{\text{R}} = 6.79$ min, >99 %, standard method.

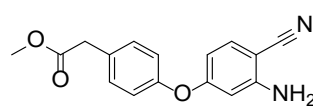
3-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N*-methylpropanamide (3.14)

 Prepared according to **general procedure 4** with carboxylic acid **3.11** (21 mg, 1.0 eq., 0.08 mmol), HCTU (31 mg, 1.0 eq., 0.08 mmol), methylamine hydrochloride (11 mg, 2.3 eq., 0.17 mmol), MeCN (2 mL) and DIPEA (0.04 mL, 3.0 eq., 0.22 mmol). The residue was purified by flash column chromatography using a gradient of 80:20 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **3.14** as a white solid (11 mg, 50 %). ^1H NMR (401 MHz, CDCl_3) δ 7.51 (d, $J = 8.5$ Hz, 1H), 7.25 – 7.20 (m, 2H), 6.99 – 6.94 (m, 2H), 6.83 (d, $J = 2.4$ Hz, 1H), 6.78 (dd, $J = 8.5, 2.4$ Hz, 1H), 5.41 (bs, 1H), 3.02 – 2.96 (m, 2H), 2.80 (d, $J = 4.8$ Hz, 3H), 2.51 – 2.45 (m, 5H). ^{13}C NMR (101 MHz, CDCl_3) δ 172.54, 161.69, 153.36, 144.55, 137.91, 134.44, 130.14, 120.58, 118.82, 118.35, 115.28, 106.41, 38.39, 31.07, 26.47, 20.76. **LCMS** (m/z): 295.0 $[\text{M}+\text{H}]^+$, $t_{\text{R}} = 3.23$ min, standard method. **HRMS** (m/z): $\text{C}_{18}\text{H}_{18}\text{N}_2\text{O}_2$ requires 295.1441 $[\text{M}+\text{H}]^+$; found 295.1441. **HPLC**: $t_{\text{R}} = 6.42$ min, 98 %, standard method.

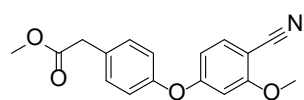
Methyl 2-(4-(4-cyano-3-methylphenoxy)phenyl)acetate (3.16, 4.87)

 4-Bromo-2-methylbenzonitrile (500 mg, 1.00 eq., 2.55 mmol), methyl 2-(4-hydroxyphenyl)acetate (636 mg, 1.50 eq., 3.83 mmol), Cs_2CO_3 (1.66 g, 2.00 eq., 5.10 mmol), *N,N*-dimethylglycine hydrochloride (142 mg, 0.40 eq., 1.02 mmol) and copper(I) iodide (97 mg, 0.20 eq., 0.51 mmol) were sealed in a crimp sealed 25 mL vial and purged with N_2 (3 times). Degassed DMF (10 mL) was added and the reaction mixture heated at 105 °C for 24 h. The reaction was then cooled and diluted with H_2O (8 mL) and EtOAc (15 mL). The layers were separated and the aqueous layer further extracted with EtOAc (15 mL). The combined organic layers were

washed brine (10 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **3.16** as a white solid (401 mg, 56 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.45 (d, J = 8.5 Hz, 1H), 7.26 – 7.21 (m, 2H), 6.96 – 6.91 (m, 2H), 6.79 (d, J = 2.3 Hz, 1H), 6.74 (dd, J = 8.5, 2.4 Hz, 1H), 3.65 (s, 3H), 3.57 (s, 2H), 2.41 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 171.98, 161.40, 154.19, 144.56, 134.43, 131.15, 130.74, 120.50, 118.99, 118.26, 115.49, 106.61, 52.25, 40.46, 20.71. **LCMS** (m/z): 279.9 [M-H]⁻, t_R = 3.54 min, standard method. **HRMS** (m/z): C₁₇H₁₅NO₃ requires 282.1125 [M+H]⁺; found 282.1124. **HPLC**: t_R = 7.34 min, >99 %, standard method.

Methyl 2-(4-(3-amino-4-cyanophenoxy)phenyl)acetate (3.17)

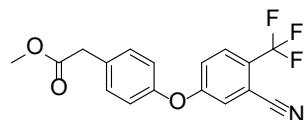
Prepared according to **general procedure 6** with 2-amino-4-bromobenzonitrile (983 mg, 1.00 eq., 4.99 mmol), methyl 2-(4-hydroxyphenyl)acetate (1.23 g, 1.49 eq., 7.43 mmol), Cs₂CO₃ (3.25 g, 2.00 eq., 9.98 mmol), *N,N*-dimethylglycine hydrochloride (281 mg, 0.404 eq., 2.01 mmol), copper(I) iodide (194 mg, 0.204 eq., 1.02 mmol) and degassed DMF (15 mL). The residue was purified by flash column chromatography using a gradient of 95:5 pet spirits:EtOAc to 55:45 pet spirits:EtOAc to give **3.17** product as a yellow solid (350 mg, 25 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.32 – 7.27 (m, 3H), 7.03 – 6.98 (m, 2H), 6.34 (dd, *J* = 8.7, 2.3 Hz, 1H), 6.23 (d, *J* = 2.2 Hz, 1H), 4.41 (bs, 2H), 3.71 (s, 3H), 3.63 (s, 2H). **¹³C NMR** (101 MHz, CDCl₃) δ 172.07, 162.91, 154.14, 151.57, 134.21, 131.04, 130.64, 120.68, 117.79, 108.36, 103.04, 90.46, 52.26, 40.47. **LCMS** (*m/z*): 282.9 [M+H]⁺, *t_R* = 3.31 min, standard method. **HRMS** (*m/z*): C₁₆H₁₄N₂O₃ requires 283.1077 [M+H]⁺; found 283.1079. **HPLC**: *t_R* = 6.31 min, >99 %, standard method.

Methyl 2-(4-(4-cyano-3-methoxyphenoxy)phenyl)acetate (3.18)

Prepared according to **general procedure 6** with 4-bromo-2-methoxybenzonitrile (292 mg, 1.00 eq., 1.38 mmol), methyl 2-(4-hydroxyphenyl)acetate (343 mg, 1.50 eq., 2.07 mmol), Cs₂CO₃ (1.0 g, 2.2 eq., 3.1 mmol), *N,N*-dimethylglycine hydrochloride (95 mg, 0.49 eq., 0.68 mmol), copper(I) iodide (65 mg, 0.25 eq., 0.34 mmol) and degassed DMF (8 mL). The residue was purified by flash column chromatography using a gradient of 95:5 pet spirits:EtOAc to 20:80 pet spirits:EtOAc. The residue was further purified by flash column chromatography using a gradient of 25:75 pet spirits:DCM to 0:100 pet spirits:DCM to give **3.18** as a white solid (65 mg, 16 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.43 (d, *J* = 8.6 Hz, 1H), 7.33 – 7.28 (m, 2H), 7.03 – 6.98 (m, 2H), 6.57 (d, *J* = 2.2 Hz, 1H), 6.48 (dd, *J* = 8.6, 2.2 Hz, 1H), 3.84 (s, 3H), 3.70 (s, 3H), 3.63 (s, 2H). **¹³C NMR** (101 MHz, CDCl₃) δ 171.88, 163.16, 163.06, 153.96, 134.95, 131.12, 130.88, 120.44, 116.57, 109.49, 101.35, 95.76, 56.19, 52.19, 40.38. **LCMS**

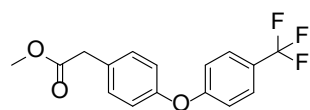
(m/z): 298.0 $[M+H]^+$, t_R = 3.51 min, standard method. **HRMS** (m/z): $C_{17}H_{15}NO_4$ requires 298.1074 $[M+H]^+$; found 298.1079. **HPLC**: t_R = 6.75 min, >99 %, standard method.

Methyl 2-(4-(3-cyano-4-(trifluoromethyl)phenoxy)phenyl)acetate (3.19)



Prepared according to **general procedure 6** with 5-bromo-2-(trifluoromethyl)benzonitrile (381 mg, 1.00 eq., 1.52 mmol), copper(I) iodide (58 mg, 0.20 eq., 0.31 mmol), *N,N*-dimethylglycine hydrochloride (85 mg, 0.40 eq., 0.61 mmol), CS_2CO_3 (992 mg, 2.00 eq., 3.05 mmol) and methyl 2-(4-hydroxyphenyl)acetate (380 mg, 1.50 eq., 2.28 mmol) and DMF (8 mL). The residue was purified by flash column chromatography using a gradient of 98:2 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **3.19** as a white solid (84 mg, 16 %). **1H NMR** (401 MHz, $CDCl_3$) δ 7.70 (d, J = 8.8 Hz, 1H), 7.38 – 7.35 (m, 2H), 7.34 (d, J = 2.5 Hz, 1H), 7.26 – 7.22 (m, 1H), 7.06 – 7.00 (m, 2H), 3.73 (s, 3H), 3.66 (s, 2H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 171.80, 160.87 (q, $^4J_{C-F}$, J = 0.9 Hz), 153.39, 131.76, 131.54, 128.79 (q, $^3J_{C-F}$, J = 4.6 Hz), 126.55 (q, $^2J_{C-F}$, J = 33.0 Hz), 122.90, 122.56 (q, $^1J_{C-F}$, J = 273 Hz), 121.10, 120.62, 115.09, 111.82 (q, $^3J_{C-F}$, J = 2.2 Hz), 52.29, 40.45. **LCMS** (m/z): did not ionise, t_R = 3.59 min, standard method. **HRMS** (m/z): $C_{17}H_{12}F_3NO_3$ requires 336.0842 $[M+H]^+$; found 336.0848. **HPLC**: t_R = 7.39 min, >99 %, standard method.

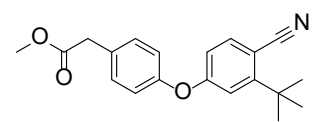
Methyl 2-(4-(4-(trifluoromethyl)phenoxy)phenyl)acetate (3.20)



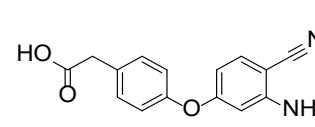
Prepared according to **general procedure 6** with 4-bromobenzotrifluoride (0.62 mL, d = 1.61 g/mL, 1.00 eq., 4.44 mmol), methyl 2-(4-hydroxyphenyl)acetate (1.11 g, 1.50 eq., 6.67 mmol), CS_2CO_3 (2.90 g, 2.00 eq., 8.89 mmol), *N,N*-dimethylglycine hydrochloride (248 mg, 0.40 eq., 1.78 mmol), copper(I) iodide (169 mg, 0.20 eq., 0.89 mmol) and degassed DMF (10 mL). The residue was purified by flash column chromatography using a gradient of 98:2 pet spirits:EtOAc to 90:10 pet spirits:EtOAc to give **3.20** as a white solid (691 mg, 50 %). **1H NMR** (401 MHz, $CDCl_3$) δ 7.57 (d, J = 8.5 Hz, 2H), 7.34 – 7.28 (m, 2H), 7.05 (d, J = 8.5 Hz, 2H), 7.03 – 6.99 (m, 2H), 3.72 (s, 3H), 3.64 (s, 2H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 172.04, 160.54 (q, $^3J_{C-F}$ = 1.3 Hz), 154.95, 131.06, 130.28, 127.22 (q, $^3J_{C-F}$ = 3.7 Hz), 125.02 (q, $^2J_{C-F}$

$f = 32.7$ Hz), 124.31 (q, $^1J_{C-F} = 271$ Hz), 120.08, 118.00, 52.19, 40.46. **LCMS** (m/z): 308.8 $[M-H]^-$, $t_R = 3.69$ min, standard method. **HRMS** (m/z): $C_{16}H_{13}F_3O_3$ requires 311.0890 $[M+H]^+$; found 311.0893. **HPLC**: $t_R = 7.57$ min, >99 %, standard method.

Methyl 2-(4-(3-(*tert*-butyl)-4-cyanophenoxy)phenyl)acetate (3.21)

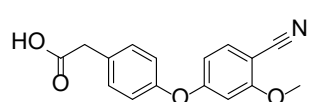
 Prepared according to **general procedure 6** with 4-bromo-2-*tert*-butylbenzonitrile (0.50 mg, 1.00 eq., 2.1 mmol), methyl 2-(4-hydroxyphenyl)acetate (0.52 g, 1.50 eq., 3.1 mmol), CS_2CO_3 (1.4 g, 2.0 eq., 4.2 mmol), *N,N*-dimethylglycine hydrochloride (120 mg, 0.40 eq., 0.84 mmol) and copper(I) iodide (80 mg, 0.20 eq., 0.42 mmol) degassed dry DMF (5 mL). The residue was purified by flash column chromatography using 90:10 pet spirits:EtOAc to give **3.21** as a colourless oil (264 mg, 39 %). **1H NMR** (401 MHz, $CDCl_3$) δ 7.58 (d, $J = 8.5$ Hz, 1H), 7.34 – 7.28 (m, 2H), 7.12 (d, $J = 2.5$ Hz, 1H), 7.04 – 6.97 (m, 2H), 6.72 (dd, $J = 8.5, 2.5$ Hz, 1H), 3.72 (s, 3H), 3.64 (s, 2H), 1.49 (s, 9H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 171.98, 161.24, 156.68, 154.23, 137.48, 131.14, 130.69, 120.42, 120.39, 116.47, 114.27, 104.48, 52.25, 40.49, 35.85, 30.09. **LCMS** (m/z): 324.0 $[M+H]^+$, $t_R = 3.69$ min, standard method. **HRMS** (m/z): $C_{20}H_{21}NO_3$ requires 324.1594 $[M+H]^+$; found 324.1599. **HPLC**: $t_R = 7.38$ min, >99 %, standard method.

2-(4-(3-Amino-4-cyanophenoxy)phenyl)acetic acid (3.22)

 Prepared according to **general procedure 5** with ester **3.17** (341 mg, 1.00 eq., 1.21 mmol) and LiOH (519 mg, 10.0 eq., 12.1 mmol) in a 1:1 THF:H₂O mixture (20 mL). The residue was purified by flash column chromatography using a gradient of 70:30 pet spirits:EtOAc to 30:70 pet spirits:EtOAc to give **3.22** as a white solid (28 mg, 8.6 %). **1H NMR** (401 MHz, CD_3OD) δ 7.35–7.32 (m, 2H), 7.30 (d, $J = 8.6$ Hz, 1H), 7.05 – 6.98 (m, 2H), 6.30 (d, $J = 2.2$ Hz, 1H), 6.27 (dd, $J = 8.6, 2.3$ Hz, 1H), 3.62 (s, 2H). **^{13}C NMR** (101 MHz, CD_3OD) δ 175.45, 164.47, 155.51, 154.69, 135.25, 132.72, 132.13, 121.49, 118.96, 108.11, 103.63, 90.14, 41.16. **LCMS** (m/z): 268.9 $[M+H]^+$, $t_R = 3.12$ min, standard method. **HRMS** (m/z):

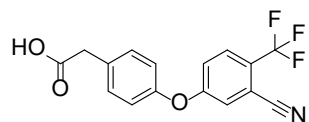
$C_{15}H_{13}N_2O_3$ requires 269.0921 $[M+H]^+$; found 269.0924. **HPLC**: t_R = 5.48 min, >99 %, standard method.

2-(4-(4-Cyano-3-methoxyphenoxy)phenyl)acetic acid (3.23)

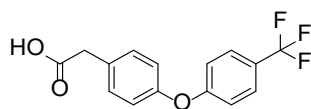


Prepared according to **general procedure 5** with ester **3.18** (65 mg, 1.0 eq., 0.22 mmol), LiOH (87 mg, 9.5 eq., 2.1 mmol) in a 1:1 THF:H₂O mixture (3 mL). The residue was purified by flash column chromatography using a gradient of 80:20 pet spirits:EtOAc to 30:70 pet spirits:EtOAc. The residue was recrystallised from MeOH and triturated with cold Et₂O to give **3.23** as a white solid (17 mg, 27 %). **¹H NMR** (401 MHz, CD₃CN) δ 7.57 (d, J = 8.6 Hz, 1H), 7.39 – 7.34 (m, 2H), 7.11 – 7.06 (m, 2H), 6.77 (d, J = 2.2 Hz, 1H), 6.55 (dd, J = 8.6, 2.2 Hz, 1H), 3.89 (s, 3H), 3.66 (s, 2H). **¹³C NMR** (101 MHz, CD₃CN) δ 173.10, 164.15, 164.14, 154.99, 136.11, 132.56, 132.27, 121.00, 117.33, 110.57, 102.87, 96.28, 57.17, 40.31. **LCMS** (m/z): 281.8 $[M-H]^-$, t_R = 3.22 min, standard method. **HRMS** (m/z): $C_{16}H_{13}NO_4$ requires 284.0917 $[M+H]^+$; found 284.0917. **HPLC**: t_R = 5.92 min, >99 %, standard method.

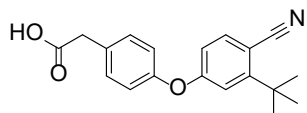
2-(4-(3-Cyano-4-(trifluoromethyl)phenoxy)phenyl)acetic acid (3.24)



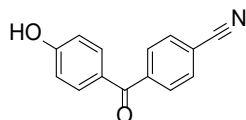
Ester **3.19** (73 mg, 1.0 eq., 0.22 mmol) and LiOH (52 mg, 10 eq., 1.1 mmol) were dissolved in a 1:1 THF:Water mixture (2 mL) and stirred for 3 days at 50 °C. After this time, the THF was removed under vacuum and the aqueous layer acidified with hydrochloride (aq., 36 % w/v, 5 drops) and worked up according to **general procedure 5**. The residue was purified by flash column chromatography using a gradient of 50:50 pet spirits:EtOAc to 30:70 pet spirits:EtOAc to give **3.24** as a white solid (25 mg, 36 %). **¹H NMR** (401 MHz, CD₃CN) δ 7.82 (d, J = 8.8 Hz, 1H), 7.52 (d, J = 2.5 Hz, 1H), 7.40 – 7.35 (m, 2H), 7.32 – 7.28 (m, 1H), 7.11 – 7.05 (m, 2H), 3.65 (s, 2H). **¹³C NMR** (101 MHz, CD₃CN) δ 173.06, 161.83 (q, $^4J_{C-F}$, J = 1.0 Hz), 154.52, 133.21, 132.52, 130.12 (q, $^3J_{C-F}$, J = 4.8 Hz), 126.63 (q, $^2J_{C-F}$, J = 32.7 Hz), 124.77, 123.92 (q, $^1J_{C-F}$, J = 271 Hz), 122.48, 121.11, 116.04, 112.25 (q, $^3J_{C-F}$, J = 2.2 Hz), 40.35. **LCMS** (m/z): 640.8 $[2M-H]^-$, t_R = 3.33 min, standard method. **HRMS** (m/z): $C_{16}H_{10}F_3NO_3$ requires 360.0244 $[M+K]^+$; found 360.0247. **HPLC**: t_R = 6.54 min, >99 %, standard method.

2-(4-(4-(Trifluoromethyl)phenoxy)phenyl)acetic acid (3.25)¹³

Prepared according to **general procedure 7** with ester **3.20** (560 mg, 1.00 eq., 1.80 mmol), MeOH (10 mL) and NaOH (aq., 2 M 5 mL, 5.55 eq., 10 mmol). The residue was purified by flash column chromatography 60:40 pet spirits:EtOAc to give **3.25** as a white solid (349 mg, 65 %). **¹H NMR** (401 MHz, CDCl₃) δ 11.28 (s, 1H), 7.58 (d, *J* = 8.5 Hz, 2H), 7.31 (d, *J* = 8.6 Hz, 2H), 7.09 – 6.99 (m, 4H), 3.67 (s, 2H). **¹³C NMR** (101 MHz, CDCl₃) δ 177.99, 160.42 (q, ³*J*_{C-F} = 1.3 Hz), 155.27, 131.20, 129.44, 127.29 (q, ³*J*_{C-F} = 3.8 Hz), 125.19 (q, ²*J*_{C-F} = 32.8 Hz), 124.31 (q, ¹*J*_{C-F} = 272 Hz), 120.10, 118.15, 40.44. **LCMS** (*m/z*): 251.0 [M-COOH]⁻, *t*_R = 3.49 min, standard method. **HRMS** (*m/z*): C₁₅H₁₁F₃O₃ requires 297.0733 [M+H]⁺; found 297.0736. **HPLC**: *t*_R = 6.76 min, >99 %, standard method.

2-(4-(3-(*tert*-Butyl)-4-cyanophenoxy)phenyl)acetic acid (3.26)

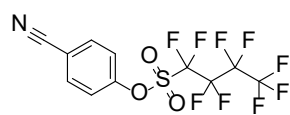
Prepared according to **general procedure 7** with ester **3.21** (250 mg, 1.00 eq., 0.77 mmol), THF (3 mL) and NaOH (aq., 2 M, 1.5 mL, 3.9 eq., 3 mmol). The residue was purified by flash column chromatography using 50:50 pet spirits:EtOAc to give the product **3.26** as a colourless oil (176 mg, 74 %). **¹H NMR** (401 MHz, CDCl₃) δ 10.79 (bs, 1H), 7.58 (d, *J* = 8.5 Hz, 1H), 7.37 – 7.29 (m, 2H), 7.13 (d, *J* = 2.4 Hz, 1H), 7.06 – 6.99 (m, 2H), 6.73 (dd, *J* = 8.5, 2.4 Hz, 1H), 3.67 (s, 2H), 1.49 (s, 9H). **¹³C NMR** (101 MHz, CDCl₃) δ 177.64, 161.14, 156.74, 154.48, 137.51, 131.27, 129.88, 120.40, 120.38, 116.57, 114.34, 104.56, 40.41, 35.86, 30.10. **LCMS** (*m/z*): 348.9 [M+CH₃CN]⁻, *t*_R = 3.46 min, standard method. **HRMS** (*m/z*): C₁₉H₁₉NO₃ requires 310.1438 [M+H]⁺; found 310.1425. **HPLC**: *t*_R = 6.63 min, >99 %, standard method.

4-(4-Methoxybenzoyl)benzonitrile (3.34)¹⁴

To a solution of 4-cyanobenzoyl chloride (500 mg, 1.0 eq., 3.0 mmol) and anisole (1.0 mL, 3 eq., 9.1 mmol) dissolved in dry DCM (15 mL) was added AlCl₃ (410 mg, 1.0 eq., 3.1 mmol) at 0 ° C. The reaction was stirred for 1 h at 0 ° C, then allowed to warm to room temperature and stirred for an additional 18 h. After this time the reaction mixture was poured onto ice water (15 mL) and diluted with DCM (30 mL). The layers were separated and the

aqueous layer further extracted with DCM (2×30 mL). The combined organic layers were washed with brine and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 95:5 pet spirits:EtOAc to 78:22 pet spirits:EtOAc. The residue was further purified by silica plug 75:25 pet spirits:EtOAc. The ether intermediate **3.33** was confirmed by ^1H NMR resonances at 7.77 – 7.68 (m, 6H) and 6.93 – 6.88 (m, 2H) ppm at <80 % purity.¹⁵ The materials were used in ongoing synthesis without further purification. The reaction was continued according to **general procedure 2** with the above sample of 4-(4-methoxybenzoyl)benzonitrile (42 mg, 1.0 eq., 0.18 mmol), dry DCM (6 mL) and BBr_3 in heptane (1.0 M, 0.53 mL, 3.0 eq., 0.53 mmol). The reaction was stirred for 0.5 h at 0 °C, then 18 h at reflux. The reaction was quenched with H_2O (10 mL) and extracted with DCM (2×10 mL). The combined organic layers were washed with brine (10 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 70:30 pet spirits:EtOAc to give **3.34** as a brown solid (14 mg, 2.1 % over 2 steps). ^1H NMR (401 MHz, CD_3CN) δ 7.88 – 7.83 (m, 2H), 7.82 – 7.77 (m, 2H), 7.74 – 7.68 (m, 2H), 6.95 – 6.90 (m, 2H). ^{13}C NMR (101 MHz, CD_3CN) δ 194.64, 162.82, 143.35, 133.77, 133.27, 130.68, 129.39, 119.17, 116.27, 115.59. LCMS (m/z): 222.0 $[\text{M}-\text{H}]^-$, t_{R} = 3.18 min, standard method. HRMS (m/z): $\text{C}_{14}\text{H}_9\text{NO}_2$ requires 224.0706 $[\text{M}+\text{H}]^+$; found 224.0704. HPLC: t_{R} = 5.90 min, 97 %, standard method.

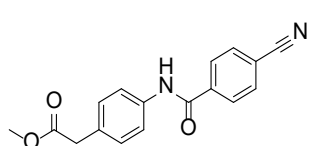
4-Cyanophenyl 1,1,2,2,3,3,4,4,4-nonafluorobutane-1-sulfonate (**3.42**)



4-Cyanophenol (500 mg, 1.00 eq., 4.20 mmol), DMAP (26 mg, 0.05 eq., 0.21 mmol) and DIPEA (0.877 mL, 1.20 eq., 5.04 mmol) were dissolved in dry DCM (10 mL) under N_2 . Perfluoro-1-butanesulfonyl fluoride (0.829 mL, 1.10 eq., 4.62 mmol) was added dropwise to the reaction at 0 °C. The reaction was stirred for 0.5 h at 0 °C, then allowed to warm to room temperature and stirred for an additional 16 h. The reaction mixture was then quenched with H_2O (80 mL) and diluted in DCM (40 mL). The layers were separated and the aqueous layer further extracted with DCM (40 mL). The combined organic layers were washed with brine and dried

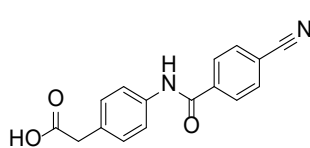
over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 90:10 pet spirits:EtOAc to give **3.42** as a white solid (1.3 g, 77%). **^1H NMR** (401 MHz, CDCl_3) δ 7.82 – 7.76 (m, 2H), 7.46 – 7.41 (m, 2H). **^{13}C NMR** (101 MHz, CDCl_3) δ 152.36, 134.61, 122.72, 117.22, 113.02. **^{19}F NMR** (377 MHz, CDCl_3) δ -80.72 (tt, J = 9.7, 2.0 Hz), -108.44 – -108.55 (m), -120.79 – -120.93 (m), -125.79 – -125.95 (m). **LCMS** (m/z): 399.7 $[\text{M}-\text{H}]^-$, t_{R} = 3.68 min, >99 %, standard method. **HPLC**: t_{R} = 8.144 min, >99 %, standard method.

Methyl 2-(4-(4-cyanobenzamido)phenyl)acetate (**3.45**)



Prepared according to **general procedure 4** with 4-cyanobenzoic acid (500 mg, 1.00 eq., 3.40 mmol), HATU (1.36 g, 1.05 eq., 3.57 mmol), DIPEA (1.78 mL, 3.0 eq., 10.2 mmol), dry DMF (15 mL) and methyl 2-(4-aminophenyl)acetate (1.12 g, 2.0 eq., 6.80 mmol). The residue was purified by flash column chromatography using a gradient of 80:20 pet spirits:EtOAc to 34:66 pet spirits:EtOAc. The residue was further triturated with Et_2O and dried under vacuum to give **3.45** as a white solid (315 mg, 27 %). **^1H NMR** (401 MHz, d_6 -DMSO) δ 10.47 (s, 1H), 8.13 – 8.08 (m, 2H), 8.04 – 8.00 (m, 2H), 7.71 (d, J = 8.5 Hz, 2H), 7.26 (d, J = 8.6 Hz, 2H), 3.66 (s, 2H), 3.62 (s, 3H). **^{13}C NMR** (101 MHz, d_6 -DMSO) δ 171.69, 164.07, 138.94, 137.49, 132.47, 130.07, 129.62, 128.52, 120.47, 118.33, 113.83, 51.69, 39.60 (confirmed via HSQC). **LCMS** (m/z): 294.9 $[\text{M}+\text{H}]^+$, t_{R} = 3.16 min, standard method. **HRMS** (m/z): $\text{C}_{17}\text{H}_{14}\text{N}_2\text{O}_3$ requires 295.1077 $[\text{M}+\text{H}]^+$; found 295.1080. **HPLC**: t_{R} = 5.63 min, >99 %, standard method.

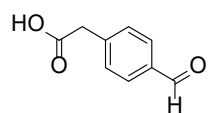
2-(4-(4-Cyanobenzamido)phenyl)acetic acid (**3.46**)



Ester **3.45** (317 mg, 1.00 eq., 0.81 mmol) was dissolved in a 1:2 sat. K_2CO_3 :MeOH mixture (18 mL) and stirred for 6 days at room temperature. After this time the MeOH was removed under vacuum and the aqueous layer acidified with hydrochloride (aq., 1 M, 20 mL, Caution: vigorous bubbling) and diluted in EtOAc (40 mL). The layers were separated and the aqueous layer further extracted with EtOAc (2 ×

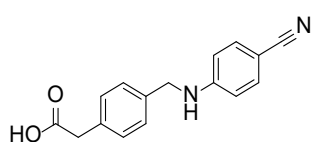
40 mL). The combined organic layers were dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 EtOAc:MeOH to 90:10 EtOAc:MeOH to give **3.46** as a white solid (21 mg, 7.0 %). **^1H NMR** (400 MHz, d_6 -DMSO) δ 10.47 (s, 1H), 8.14 – 8.08 (m, 2H), 8.06 – 8.00 (m, 2H), 7.71 (d, J = 8.4 Hz, 2H), 7.26 (d, J = 8.4 Hz, 2H), 3.56 (s, 2H). **^{13}C NMR** (101 MHz, d_6 -DMSO) δ 173.08 (confirmed via HMBC), 164.50, 139.43, 137.72, 132.94, 131.37, 130.09, 128.99, 120.85, 118.80, 114.28, 40.39 (confirmed via HSQC). **LCMS** (m/z): 278.9 $[\text{M}-\text{H}]^-$, t_{R} = 3.93 min, standard method. **HRMS** (m/z): $\text{C}_{16}\text{H}_{12}\text{N}_2\text{O}_3$ requires 281.0921 $[\text{M}+\text{H}]^+$; found 281.0921. **HPLC**: t_{R} = 4.69 min, >99 %, standard method.

2-(4-Formylphenyl)acetic acid (**3.48**)¹⁶



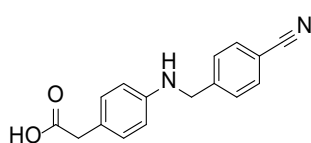
To a solution of 2-[4-(hydroxymethyl)phenyl]acetic acid (1.00 g, 1.00 eq., 6.02 mmol) dissolved in a 1:1 mixture of dry DCM/THF (30 mL) was added Dess-Martin periodinane (2.81 g, 1.1 eq., 6.62 mmol) under N_2 . The reaction mixture was then stirred for 5 h at room temperature. After this time the reaction was quenched with a 1:3 mixture of sat. $\text{Na}_2\text{S}_2\text{O}_3$ (aq., 20 mL) and NaHCO_3 (aq., sat., 60 mL) and diluted with DCM (80 mL). The solution was acidified with hydrochloride (aq., 6 M, 30 mL, caution vigorous bubbling) and sat. citric acid (aq., 50 mL). The layers were separated and the aqueous layer further extracted with DCM (2×50 mL). The combined organic layers were filtered, washed with brine (20 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 90:10 pet spirits:EtOAc to 50:50 pet spirits:EtOAc to give **3.48** as a white solid (488 mg, 49 %). **^1H NMR** (401 MHz, d_6 -DMSO) δ 9.99 (s, 1H), 7.89 – 7.83 (m, 2H), 7.50 (d, J = 8.0 Hz, 2H), 3.71 (s, 2H). **^{13}C NMR** (101 MHz, d_6 -DMSO) δ 192.77, 172.04, 142.11, 134.82, 130.31, 129.48, 40.68. **LCMS** (m/z): 119.1 $[\text{M}-\text{COOH}]^-$, t_{R} = 2.79 min, standard method. **HPLC**: t_{R} = 3.68 min, >99 %, standard method.

2-(4-(((4-Cyanophenyl)amino)methyl)phenyl)acetic acid (**3.49**)



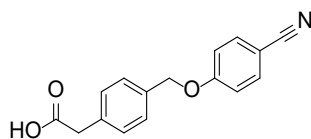
Prepared according to **general procedure 10** with aldehyde **3.48** (300 mg, 1.0 eq., 1.83 mmol), 4-aminobenzonitrile (324 mg, 1.50 eq., 2.74 mmol), anhydrous 1,2-DCE (18 mL) and Na(CH₃COO)₃BH (581 mg, 1.5 eq., 2.74 mmol). The residue purified by flash column chromatography using a gradient of 70:30 pet spirits:EtOAc to 50:50 pet spirits:EtOAc to give **3.49** as a white solid (140 mg, 29 %). **¹H NMR** (401 MHz, CD₃OD) δ 7.39 – 7.34 (m, 2H), 7.30 (d, *J* = 8.3 Hz, 2H), 7.25 (d, *J* = 8.2 Hz, 2H), 6.67 – 6.61 (m, 2H), 4.36 (s, 2H), 3.58 (s, 2H). **¹³C NMR** (101 MHz, CD₃OD) δ 175.66, 153.77, 138.87, 134.99, 134.51, 130.63, 128.34, 121.62, 113.42, 98.02, 47.31, 41.60. **LCMS** (*m/z*): 264.9 [M-H]⁻, *t_R* = 3.14 min, standard method. **HRMS** (*m/z*): C₁₆H₁₄N₂O₂ requires 267.1128 [M+H]⁺; found 267.1127. **HPLC**: *t_R* = 5.50 min, > 99 %, standard method.

2-(4-((4-Cyanobenzyl)amino)phenyl)acetic acid (**3.52**)



Prepared according to **general procedure 10** with 4-cyanobenzaldehyde (250 mg, 1.0 eq., 1.91 mmol), 4-aminophenylacetic acid (432 mg, 1.50 eq., 2.86 mmol), anhydrous 1,2-DCE (10 mL) and Na(CH₃COO)₃BH (606 mg, 1.5 eq., 2.86 mmol). The solution was filtered, basified with Et₃N (4 mL) and the residue purified by flash column chromatography using a gradient of 70:30 pet spirits:EtOAc to 0:100 pet spirits:EtOAc with 0.5 % AcOH added at 0:100 pet spirits:EtOAc. The fractions containing precipitate were filtered to give **3.52** as a white solid (51 mg, 10 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 7.80 – 7.75 (m, 2H), 7.53 (d, *J* = 8.4 Hz, 2H), 6.91 (d, *J* = 8.5 Hz, 2H), 6.47 (d, *J* = 8.5 Hz, 2H), 6.32 (t, *J* = 6.3 Hz, 2H), 4.35 (d, *J* = 6.2 Hz, 2H), 3.33 (s, 2H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 173.29, 146.86, 146.81, 132.23, 129.82, 127.92, 122.30, 118.97, 112.17, 109.30, 46.08, 39.52 (Confirmed via HSQC). **LCMS** (*m/z*): 266.9 [M+H]⁺, *t_R* = 3.16 min, standard method. **HRMS** (*m/z*): C₁₆H₁₄N₂O₂ requires 267.1128 [M+H]⁺; found 267.1128. **HPLC**: *t_R* = 4.64 min, 99 %, standard method.

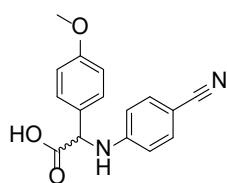
2-(4-((4-Cyanophenoxy)methyl)phenyl)acetic acid (**3.54**)



To a solution of 4-cyanophenol (100 mg, 1.0 eq., 0.84 mmol), 4-(bromomethyl)phenylacetic acid (288 mg, 1.5 eq., 1.26 mmol) and KI (279

mg, 2.0 eq., 1.68 mmol) dissolved in dry MeCN (5 mL) was added anhydrous K_2CO_3 (464 mg, 4.0 eq., 3.36 mmol) at 0 °C. The reaction mixture was stirred for 4 h on ice, then 4 days at room temperature. The reaction was diluted in EtOAc causing precipitation. The solid was filtered and triturated with Et_2O to give **3.54** as a white solid (111 mg, 49 %). 1H NMR (401 MHz, $CDCl_3$) δ 7.61 – 7.56 (m, 2H), 7.39 (d, J = 8.2 Hz, 2H), 7.35 – 7.30 (m, 2H), 7.03 – 6.98 (m, 2H), 5.10 (s, 2H), 3.67 (s, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 177.11, 162.03, 135.05, 134.18, 133.63, 129.95, 127.91, 119.27, 115.72, 104.44, 70.05, 40.76. LCMS (m/z): 259.9 $[M-H]^-$, t_R = 3.16 min, standard method. HRMS (m/z): $C_{16}H_{13}NO_3$ requires 268.0968 $[M+H]^+$; found 268.0967. HPLC: t_R = 5.81 min, 99 %, standard method.

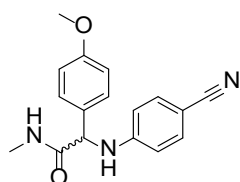
2-((4-Cyanophenyl)amino)-2-(4-methoxyphenyl)acetic acid (**3.58**)



4-Aminobenzonitrile (1.20 g, 1.00 eq., 10.0 mmol), 4-methoxyphenylboronic acid (1.52 g, 1.00 eq., 10 mmol) and glyoxylic acid monohydrate (1.00 g, 1.10 eq., 10.9 mmol) were dissolved in MeCN (30 mL) and heated at reflux for 1 h.

After this time, the solvent was removed under vacuum and the residue purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc. Two samples of **3.58** were isolated in 99 % purity (600 mg, 21 %) and 95 % purity (2.00 g, 71 %) for a combined yield of 92 %. Analytical and screening data are reported for the higher purity sample. 1H NMR (401 MHz, CD_3OD) δ 7.43 – 7.38 (m, 2H), 7.38 – 7.32 (m, 2H), 6.92 – 6.87 (m, 2H), 6.68 – 6.63 (m, 2H), 5.08 (s, 1H), 3.75 (s, 3H). ^{13}C NMR (101 MHz, CD_3OD) δ 174.34, 161.14, 151.86, 134.42, 130.54, 129.70, 121.32, 115.18, 114.18, 99.14, 60.36, 55.72. LCMS (m/z): 281.0 $[M-H]^-$, t_R = 3.29 min, standard method. HRMS (m/z): $C_{16}H_{14}N_2O_3$ requires 283.1077 $[M+H]^+$; found 283.1074. HPLC: t_R = 5.76 min, >99 %, standard method.

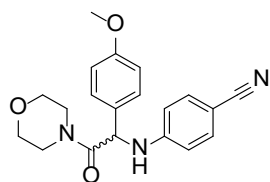
2-((4-Cyanophenyl)amino)-2-(4-methoxyphenyl)-*N*-methylacetamide (**3.59**)



Prepared according to **general procedure 4** with carboxylic acid **3.58** (150 mg, 1.00 eq., 0.531 mmol), methylamine hydrochloride (72 mg, 2.0 eq., 1.1 mmol), HATU (202 mg, 1.00 eq., 0.531 mmol), DMF (5 mL) and DIPEA (0.278 mL,

3.00 eq., 1.60 mmol). The residue was purified by flash column chromatography using a gradient of 75:25 pet spirits:EtOAc to 25:75 pet spirits:EtOAc to give **3.59** as a yellow oil (88 mg, 56 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.38 – 7.30 (m, 4H), 6.94 – 6.89 (m, 2H), 6.55 – 6.51 (m, 2H), 5.90 – 5.88 (m, 1H), 5.67 (d, *J* = 3.2 Hz, 1H), 4.75 (d, *J* = 3.5 Hz, 1H), 3.80 (s, 3H), 2.81 (d, *J* = 4.9 Hz, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 170.84, 160.11, 149.79, 133.72, 129.96, 128.46, 120.25, 115.07, 113.53, 100.10, 61.34, 55.48, 26.87. **LCMS** (*m/z*): 294.0 [M-H]⁻, *t_R* = 3.12 min, standard method. **HRMS** (*m/z*): C₁₇H₁₇N₃O₂ requires 318.1213 [M+Na]⁺; found 318.1214. **HPLC**: *t_R* = 5.51 min, >99 %, standard method.

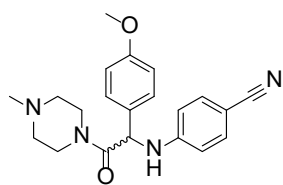
4-((1-(4-Methoxyphenyl)-2-morpholino-2-oxoethyl)amino)benzonitrile (**3.60**)



Prepared according to **general procedure 4** with carboxylic acid **3.58** (150 mg, 1.00 eq., 0.531 mmol), HATU (202 mg, 1.00 eq., 0.531 mmol), DMF (5 mL), DIPEA (0.278 mL, 3.00 eq., 1.60 mmol) and morpholine (92 μL, 2.00

eq., 1.06 mmol). The residue was purified by flash column chromatography using a gradient of 90:10 pet spirits:EtOAc to 25:75 pet spirits:EtOAc to give **3.60** as a white foam (134 mg, 72 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.41 – 7.32 (m, 4H), 6.95 – 6.88 (m, 2H), 6.62 – 6.56 (m, 2H), 6.06 (d, *J* = 6.4 Hz, 1H), 5.20 (d, *J* = 6.5 Hz, 1H), 3.81 (s, 3H), 3.78 – 3.44 (m, 7H), 3.25 – 3.15 (m, 1H). **¹³C NMR** (101 MHz, CDCl₃) δ 168.69, 159.81, 149.22, 133.77, 128.98, 128.96, 120.41, 114.81, 113.06, 99.34, 66.73, 66.20, 56.58, 55.44, 46.02, 43.16. **LCMS** (*m/z*): 350.0 [M-H]⁻, *t_R* = 3.19 min, standard method. **HRMS** (*m/z*): C₂₀H₂₁N₃O₃ requires 352.1656 [M+H]⁺; found 352.1651. **HPLC**: *t_R* = 5.89 min, >99 %, standard method.

4-((1-(4-Methoxyphenyl)-2-(4-methylpiperazin-1-yl)-2-oxoethyl)amino)benzonitrile (**3.61**)

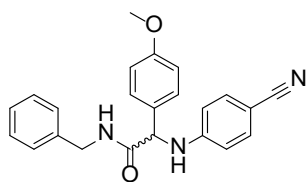


Prepared according to **general procedure 4** with carboxylic acid **3.58** (150 mg, 1.00 eq., 0.531 mmol), HATU (202 mg, 1.00 eq., 0.531 mmol), DMF (5 mL), DIPEA (0.278 mL, 3.00 eq., 1.60 mmol) and *N*-methylpiperazine (0.12

mL, 2.0 eq., 1.1 mmol). The residue was purified by flash column chromatography using a gradient of 100:0 EtOAc:MeOH to 85:15 EtOAc:MeOH to give **3.61** as a colourless oil (120 mg, 62 %). **¹H**

NMR (401 MHz, CDCl_3) δ 7.37 – 7.29 (m, 4H), 6.90 – 6.85 (m, 2H), 6.59 – 6.53 (m, 2H), 6.06 (d, J = 6.5 Hz, 1H), 5.19 (d, J = 6.6 Hz, 1H), 3.78 (s, 3H), 3.77 – 3.66 (m, 1H), 3.62 – 3.43 (m, 3H), 2.46 – 2.15 (m, 2H), 2.22 (m, 1H), 2.21 (s, 3H), 1.89 (m, 1H). **^{13}C NMR** (101 MHz, CDCl_3) δ 168.43, 159.68, 149.28, 133.73, 129.23, 128.94, 120.46, 114.69, 113.01, 99.12, 56.49, 55.42, 54.53 (2 carbons), 45.99, 45.40, 42.70. **LCMS** (m/z): 362.9 $[\text{M}+\text{H}]^+$, t_R = 2.82 min, standard method. **HRMS** (m/z): $\text{C}_{21}\text{H}_{24}\text{N}_4\text{O}_2$ requires 751.3691 $[2\text{M}+\text{Na}]^+$; found 751.3690. **HPLC**: t_R = 5.02 min, >99 %, standard method.

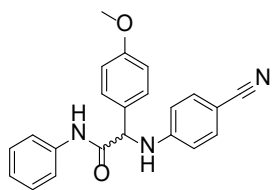
***N*-Benzyl-2-((4-cyanophenyl)amino)-2-(4-methoxyphenyl)acetamide (3.62)**



Prepared according to **general procedure 4** with carboxylic acid **3.58** (500 mg, 1.00 eq., 1.77 mmol), HATU (673 mg, 1.00 eq., 1.77 mmol), DMF (5 mL), benzylamine (0.387 mL, 2.00 eq., 3.54 mmol) and DIPEA (0.342 mL,

3.00 eq., 1.96 mmol). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc. The residue was further purified by trituration with C hydrochloride₃ to give **3.62** as a white solid (81 mg, 12 %). **^1H NMR** (401 MHz, d_6 -DMSO) δ 8.79 (t, J = 5.9 Hz, 1H), 7.48 – 7.40 (m, 4H), 7.30 – 7.18 (m, 4H), 7.16 – 7.11 (m, 2H), 6.95 – 6.90 (m, 2H), 6.72 (d, J = 8.8 Hz, 2H), 5.08 (d, J = 7.1 Hz, 1H), 4.34 – 4.22 (m, 2H), 3.74 (s, 3H). **^{13}C NMR** (101 MHz, d_6 -DMSO) δ 170.26, 158.93, 150.78, 139.00, 133.19, 130.45, 128.41, 128.25, 127.12, 126.85, 120.41, 113.88, 112.90, 96.72, 59.26, 55.12, 42.18. **LCMS** (m/z): 370.0 $[\text{M}-\text{H}]^-$, t_R = 3.33 min, standard method. **HRMS** (m/z): $\text{C}_{23}\text{H}_{21}\text{N}_3\text{O}_2$ requires 765.3160 $[2\text{M}+\text{Na}]^+$; found 765.3160. **HPLC**: t_R = 6.55 min, >99 %, standard method.

2-((4-Cyanophenyl)amino)-2-(4-methoxyphenyl)-*N*-phenylacetamide (3.63)

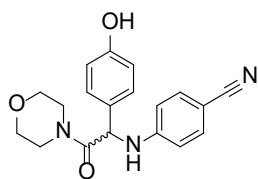


Prepared according to **general procedure 4** with carboxylic acid **3.58** (150 mg, 1.00 eq., 0.531 mmol), HATU (202 mg, 1.00 eq., 0.531 mmol), DMF (5 mL), DIPEA (0.278 mL, 3.00 eq., 1.60 mmol) and aniline (0.10 mL, 2.0 eq.,

1.1 mmol). The residue was purified by flash column chromatography using a gradient of 100:0 DCM:EtOAc to 75:25 DCM:EtOAc to give **3.63** as a white foam (62 mg, 33 %). **^1H NMR** (401 MHz,

CDCl₃) δ 8.15 (s, 1H), 7.47 – 7.43 (m, 2H), 7.41 – 7.37 (m, 2H), 7.37 – 7.32 (m, 2H), 7.31 – 7.25 (m, 2H), 7.10 (t, J = 7.4 Hz, 1H), 6.92 – 6.86 (m, 2H), 6.60 – 6.54 (m, 2H), 5.59 (d, J = 4.0 Hz, 1H), 4.90 (d, J = 4.0 Hz, 1H), 3.76 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 168.78, 160.13, 149.64, 137.12, 133.77, 129.35, 129.14, 128.52, 125.09, 120.27, 120.23, 115.00, 113.56, 100.20, 62.04, 55.41. **LCMS** (m/z): 355.9 [M-H]⁻, t_R = 3.37 min, standard method. **HRMS** (m/z): C₂₂H₁₉N₃O₂ requires 380.1369 [M+Na]⁺; found 380.1368. **HPLC**: t_R = 6.75 min, >99 %, standard method.

4-((1-(4-Hydroxyphenyl)-2-morpholino-2-oxoethyl)amino)benzonitrile (**3.65**)

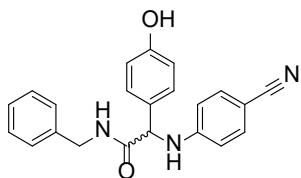


Prepared according to **general procedure 2** with ether **3.60** (102 mg, 1.00 eq., 0.290 mmol), DCM (5 mL) and BBr₃ in heptane (2.90 mL, 10.0 eq., 2.90 mmol) at 0 °C. The residue was purified by flash column chromatography using a

gradient of 100:0 EtOAc:MeOH to 90:10 EtOAc:MeOH to give **3.65** as a white solid (23 mg, 23 %).

¹H NMR (401 MHz, CD₃CN) δ 7.41 – 7.35 (m, 2H), 7.32 – 7.26 (m, 2H), 7.14 (s, 1H), 6.82 – 6.77 (m, 2H), 6.73 – 6.68 (m, 2H), 6.18 (d, J = 7.2 Hz, 1H), 5.40 (d, J = 7.3 Hz, 1H), 3.65 – 3.42 (m, 6H), 3.42 – 3.31 (m, 1H), 3.24 – 3.14 (m, 1H). **¹³C NMR** (101 MHz, CD₃CN) δ 169.65, 157.87, 151.08, 134.37, 130.18, 129.69, 121.10, 116.56, 114.05, 99.11, 67.21, 66.91, 56.62, 46.73, 43.57. **LCMS** (m/z): 335.9 [M-H]⁻, t_R = 3.03 min, standard method. **HRMS** (m/z): C₁₉H₁₉N₃O₃ requires 338.1499 [M+H]⁺; found 338.1495. **HPLC**: t_R = 5.00 min, >99 %, standard method.

N-Benzyl-2-((4-cyanophenyl)amino)-2-(4-hydroxyphenyl)acetamide (**3.67**)

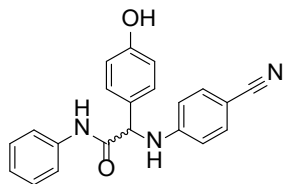


Prepared according to **general procedure 2** with ether **3.62** (69 mg, 1.0 eq., 0.19 mmol), DCM (4 mL) and BBr₃ in heptane (1.9 mL, 10 eq., 1.9 mmol) at 0 °C. The residue was purified by flash column chromatography using a

gradient of 100:0 pet spirits:EtOAc to 50:50 pet spirits:EtOAc giving **3.67** as a colourless oil (23 mg, 35 %). **¹H NMR** (401 MHz, CD₃OD) δ 7.41 – 7.36 (m, 2H), 7.34 – 7.28 (m, 2H), 7.25 – 7.18 (m, 3H), 7.11 (dd, J = 7.5, 1.8 Hz, 2H), 6.81 – 6.76 (m, 2H), 6.70 – 6.65 (m, 2H), 4.91 (s, 1H), 4.41 (d, J = 15.0 Hz, 1H), 4.32 (d, J = 15.0 Hz, 1H). **¹³C NMR** (101 MHz, CD₃OD) δ 173.78, 158.83, 152.18, 139.68, 134.45, 129.87, 129.80, 129.41, 128.42, 128.18, 121.27, 116.62, 114.28, 99.50, 62.49, 44.06.

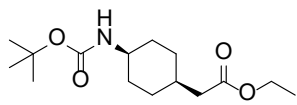
LCMS (m/z): 355.9 $[M-H]^-$, t_R = 3.19 min, standard method. **HRMS** (m/z): $C_{22}H_{19}N_3O_2$ requires 358.1550 $[M+H]^+$; found 358.1546. **HPLC**: t_R = 5.90 min, 90 %, standard method.

2-((4-Cyanophenyl)amino)-2-(4-hydroxyphenyl)-*N*-phenylacetamide (3.68)



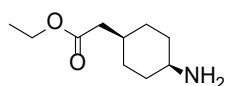
Prepared according to **general procedure 2** with ether **3.63** (42 mg, 1.0 eq., 0.12 mmol), DCM (3 mL) and BBr_3 in heptane (1.2 mL, 10 eq., 1.2 mmol) at 0 °C. The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 50:50 pet spirits:EtOAc giving **3.68** as a brown solid (19 mg, 47 %). **1H NMR** (401 MHz, d_6 -DMSO) δ 9.49 (s, 1H), 7.59 (dd, J = 8.6, 1.1 Hz, 2H), 7.48 (d, J = 8.9 Hz, 2H), 7.38 (t, J = 7.2 Hz, 2H), 7.34 – 7.23 (m, 3H), 7.06 (t, J = 7.4 Hz, 1H), 6.79 – 6.73 (m, 4H), 5.23 – 5.11 (d, J = 4.1 Hz, 1H), 4.87 (d, J = 4.1 Hz, 1H). **LCMS** (m/z): 342.1 $[M-H]^-$, t_R = 3.28 min, standard method. **HPLC**: t_R = 6.05 min, 90 %, standard method.

Ethyl 2-((1*s*,4*s*)-4-((*tert*-butoxycarbonyl)amino)cyclohexyl)acetate (4.18)¹⁷



To a solution of boc-*cis*-1,4-aminocyclohexyl acetic acid (909 mg, 1.00 eq., 3.53 mmol) and K_2CO_3 (732 mg, 1.50 eq., 5.30 mmol) dissolved in MeCN (20 mL) was added iodoethane (311 μ L, 1.1 eq., 388 mmol) dropwise at room temperature. The reaction was then stirred overnight at 50 °C. After this time, the MeCN was removed under vacuum and the residue dissolved in EtOAc (30 mL) and diluted in H_2O (20 mL). The layers were separated and the organic washed with $NaHCO_3$ (aq., sat., 20 mL) and brine (30 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness to give **4.18** as a colourless oil (970 mg, 96 %). **1H NMR** (401 MHz, $CDCl_3$) δ 4.60 (bs, 1H), 4.12 (q, J = 7.1 Hz, 2H), 3.71 (m, 1H), 2.23 (d, J = 7.3 Hz, 2H), 1.97 – 1.85 (m, 1H), 1.70 – 1.58 (m, 6H), 1.44 (s, 9H), 1.25 (t, J = 7.1 Hz, 3H), 1.30 – 1.17 (m, 2H). **LCMS** (m/z): 307.9 $[M+Na]^+$, t_R = 3.83 min, standard method. **HPLC**: t_R = 6.79 min, >99 %, standard method.

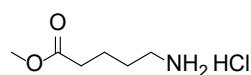
Ethyl 2-((1*s*,4*s*)-4-aminocyclohexyl)acetate (4.19)¹⁸



N-Boc amine **4.18** (952 mg, 1.00 eq., 3.34 mmol) dissolved in DCM (25 mL) was added TFA (5 mL) dropwise at RT. After 16 h stirring, the solution was diluted

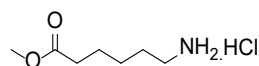
with DCM (100 mL) and H₂O (50 mL), followed by dropwise addition of 33 % NH₄OH (aq., 3 mL or until pH 10). The product was further extracted with DCM (2 × 50 mL). The combined organic extracts washed with brine (40 mL), dried over Na₂SO₄ and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 DCM:MeOH to 75:25 DCM:MeOH to give **4.19** as a colourless oil (290 mg, 47 %). **¹H NMR** (401 MHz, CDCl₃) δ 4.10 (q, *J* = 7.1 Hz, 2H), 2.97 (dt, *J* = 8.8, 4.5 Hz, 1H), 2.25 (d, *J* = 7.3 Hz, 2H), 2.19 (bs, 2H), 1.99 – 1.89 (m, 1H), 1.65 – 1.37 (m, 8H), 1.23 (t, *J* = 7.1 Hz, 3H). **LCMS** (*m/z*): 186.20 [M+H]⁺, *t_R* = 0.63 min, standard method. **HPLC**: *t_R* = 3.293 min, 96 %, standard method.

Methyl 5-aminopentanoate hydrochloride (**4.20**)¹⁹



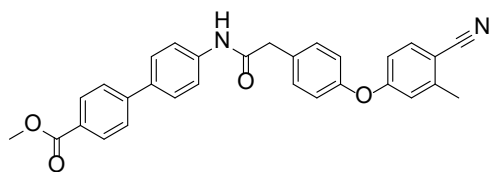
To a solution of 5-aminovaleric acid (501 mg, 1.00 eq., 4.28 mmol) dissolved in MeOH (5 mL) was added SOCl₂ (0.372 mL, 1.20 eq., 5.13 mmol) dropwise at 0 °C. The reaction was stirred for 0.5 h at 0 °C, then allowed to warm to room temperature and stirred for an additional 24 h. After this time the solvent was removed under vacuum to give **4.20** as a white solid (570 mg, 73 %). **¹H NMR** (401 MHz, CD₃OD) δ 3.67 (s, 3H), 2.98 – 2.90 (m, 2H), 2.45 – 2.38 (m, 2H), 1.73 – 1.65 (m, 4H). **¹³C NMR** (101 MHz, CD₃OD) δ 175.23, 52.11, 40.41, 33.93, 27.92, 22.70. **LCMS** (*m/z*): 132.1 [M+H]⁺, *t_R* = 0.57 min, standard method. **HPLC**: *t_R* = 1.55 min, >99 %, standard method.

Methyl 6-aminohexanoate hydrochloride (**4.21**)²⁰



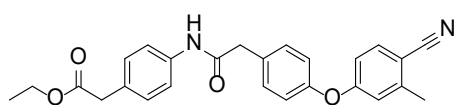
To a solution of 6-aminocaproic acid (525 mg, 1.00 eq., 4.00 mmol) dissolved in MeOH (5 mL) was added SOCl₂ (0.311 mL, 1.20 eq., 4.80 mmol) dropwise at 0 °C. The reaction was stirred for 0.5 h at 0 °C, then allowed to warm to room temperature and stirred for an additional 24 h. After this time the solvent was removed under vacuum to give **4.21** as a white solid (685 mg, 94 %). **¹H NMR** (401 MHz, CD₃OD) δ 3.66 (s, 3H), 2.97 – 2.88 (m, 2H), 2.37 (t, *J* = 7.3 Hz, 2H), 1.72 – 1.60 (m, 4H), 1.48 – 1.36 (m, 2H). **¹³C NMR** (101 MHz, CD₃OD) δ 175.62, 52.04, 40.57, 34.36, 28.23, 26.86, 25.37. **LCMS** (*m/z*): 146.1 [M+H]⁺, *t_R* = 0.45 min, standard method. **HPLC**: *t_R* = 1.60 min, 98 %, standard method.

Methyl 4'-(2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamido)-[1,1'-biphenyl]-4-carboxylate (4.22)



Prepared according to **general procedure 4** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (250 mg, 1.00 eq., 0.935 mmol), HATU (357 mg, 1.00 eq., 0.935 mmol), DIPEA (0.58 mL, 3.5 eq., 3.3 mmol), dry DMF (4 mL) and methyl 4-(4-aminophenyl)benzoate (524 mg, 2.47 eq., 2.31 mmol). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 30:70 pet spirits:EtOAc. The residue was then recrystallised from MeOH and dried to give **4.22** as a white solid (181 mg, 41 %). **¹H NMR** (401 MHz, CDCl₃) δ 8.07 (d, *J* = 8.3 Hz, 2H), 7.67 (bs, 1H), 7.63 – 7.56 (m, 6H), 7.52 (d, *J* = 8.5 Hz, 1H), 7.38 (d, *J* = 8.4 Hz, 2H), 7.05 (d, *J* = 8.4 Hz, 2H), 6.87 (d, *J* = 1.9 Hz, 1H), 6.82 (dd, *J* = 8.5, 2.2 Hz, 1H), 3.93 (s, 3H), 3.75 (s, 2H), 2.49 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 169.15, 167.11, 161.18, 154.56, 144.81, 144.65, 137.95, 136.09, 134.48, 131.29, 131.13, 130.24, 128.83, 127.89, 126.73, 120.79, 120.32, 119.16, 118.25, 115.60, 106.75, 52.26, 43.97, 20.72. **LCMS** (*m/z*): 475.0 [M-H]⁻, *t_R* = 3.59 min, standard method. **HRMS** (*m/z*): C₃₀H₂₄N₂O₄ requires 477.1809 [M+H]⁺; found 477.1813. **HPLC**: *t_R* = 7.78 min, >99 %, standard method.

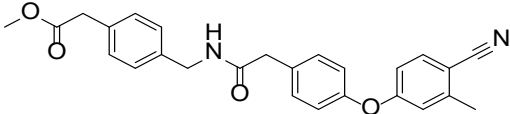
Ethyl 2-(4-(2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamido)phenyl)acetate (4.23)



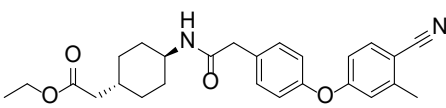
Prepared according to **general procedure 4** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.00 eq., 0.748 mmol), HATU (285 mg, 1.00 eq., 0.748 mmol), ethyl 2-(4-aminophenyl)acetate (268 mg, 2 eq., 1.50 mmol), dry DMF (5 mL) and DIPEA (0.456 mL, 3.00 eq., 2.62 mmol). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 75:25 pet spirits:EtOAc to give **4.23** as a yellow solid (98 mg, 31 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.79 (bs, 1H), 7.51 (d, *J* = 8.5 Hz, 1H), 7.45 – 7.40 (m, 2H), 7.37 – 7.32 (m, 2H), 7.21 – 7.16 (m, 2H), 7.04 – 6.99 (m, 2H), 6.86 (d, *J* = 2.2 Hz, 1H), 6.80 (dd, *J* = 8.5, 2.4 Hz, 1H), 4.13 (q, *J* = 7.1 Hz, 2H), 3.66 (s, 2H), 3.55 (s, 2H), 2.48 (s, 3H), 1.23 (t, *J* = 7.1 Hz, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 171.78,

169.12, 161.24, 154.34, 144.59, 136.90, 134.43, 131.39, 131.18, 130.23, 129.85, 120.67, 120.12, 119.05, 118.25, 115.53, 106.57, 61.00, 43.75, 40.82, 20.66, 14.23. **LCMS** (m/z): 426.9 $[M+H]^+$, t_R = 3.47 min, standard method. **HRMS** (m/z): $C_{26}H_{24}N_4O_4$ requires 429.1809 $[M+H]^+$; found 429.1813. **HPLC**: t_R = 7.24 min, >99 %, standard method.

Methyl 2-((2-((4-(4-cyano-3-methylphenoxy)phenyl)acetamido)methyl)phenyl)acetate (4.24)

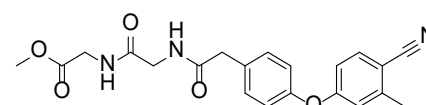
 Prepared according to **general procedure 4** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (209 mg, 1.00 eq., 0.782 mmol), HATU (303 mg, 1.02 eq., 0.797 mmol), DIPEA (0.42 mL, 3.0 eq., 2.3 mmol), dry DMF (5 mL) and methyl 2-[4-(aminomethyl)phenyl]acetate hydrochloride (338 mg, 2.00 eq., 1.57 mmol). The residue was purified by flash column chromatography using a gradient of 80:20 pet spirits:EtOAc to 30:70 pet spirits:EtOAc. The residue was then recrystallised from MeOH and dried to give **4.24** as a white solid (167 mg, 50 %). **¹H NMR** (401 MHz, d_6 -DMSO) δ 8.55 (t, J = 5.8 Hz, 1H), 7.75 (d, J = 8.6 Hz, 1H), 7.39 – 7.33 (m, 2H), 7.23 – 7.15 (m, 4H), 7.10 – 7.05 (m, 2H), 7.02 (d, J = 2.4 Hz, 1H), 6.87 (dd, J = 8.6, 2.3 Hz, 1H), 4.25 (d, J = 5.8 Hz, 2H), 3.64 (s, 2H), 3.60 (s, 3H), 3.50 (s, 2H), 2.44 (s, 3H). **¹³C NMR** (101 MHz, d_6 -DMSO) δ 171.63, 169.97, 161.05, 153.02, 144.33, 137.96, 134.72, 133.22, 132.86, 130.94, 129.26, 127.27, 120.02, 118.64, 117.94, 115.28, 105.59, 51.66, 41.94, 41.55, 40.29, 19.99. **LCMS** (m/z): 428.9 $[M+H]^+$, t_R = 3.37 min, standard method. **HRMS** (m/z): $C_{26}H_{24}N_2O_4$ requires 429.1809 $[M+H]^+$; found 429.1818. **HPLC**: t_R = 6.76 min, >99 %, standard method.

Ethyl 2-((1*r*,4*r*)-4-(2-((4-(4-cyano-3-methylphenoxy)phenyl)acetamido)cyclohexyl)acetate (4.25)

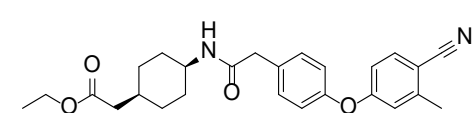
 Prepared according to **general procedure 4** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (213 mg, 1.00 eq., 0.797 mmol), HATU (303 mg, 1.00 eq., 0.797 mmol), ethyl *trans*-2-(4-aminocyclohexyl)acetate hydrochloride (396 mg, 2.24 eq., 1.79 mmol), DIPEA (0.486 mL, 3.5 eq., 2.79 mmol) and DMF (5 mL). The residue was purified by flash column chromatography using a gradient of 90:10 pet spirits:EtOAc to 40:60 pet spirits:EtOAc to give **4.25** as a white solid (134 mg, 39 %). **¹H NMR** (401

MHz, CDCl₃) δ 7.52 (d, J = 8.5 Hz, 1H), 7.30 – 7.26 (m, 2H), 7.04 – 6.96 (m, 2H), 6.85 (d, J = 2.4 Hz, 1H), 6.81 (dd, J = 8.5, 2.3 Hz, 1H), 5.35 (d, J = 7.1 Hz, 1H), 4.10 (q, J = 7.1 Hz, 2H), 3.77 – 3.66 (m, 1H), 3.51 (s, 2H), 2.48 (s, 3H), 2.17 (d, J = 6.8 Hz, 2H), 2.01 – 1.88 (m, 2H), 1.82 – 1.74 (m, 2H), 1.74 – 1.66 (m, 1H), 1.23 (t, J = 7.1 Hz, 3H), 1.16 – 1.02 (m, 4H). ¹³C NMR (101 MHz, CDCl₃) δ 172.93, 169.93, 161.29, 154.24, 144.60, 134.46, 131.89, 131.09, 120.68, 119.06, 118.23, 115.54, 106.69, 60.35, 48.56, 43.17, 41.46, 34.03, 32.76, 31.51, 20.73, 14.36. LCMS (m/z): 434.9 [M+H]⁺, t_R = 3.46 min, standard method. HRMS (m/z): C₂₆H₃₀N₂O₄ requires 435.2278 [M+H]⁺; found 435.2282. HPLC: t_R = 7.09 min, >99 %, standard method.

Methyl (2-(4-(4-cyano-3-methylphenoxy)phenyl)acetyl)glycylglycinate (4.26)

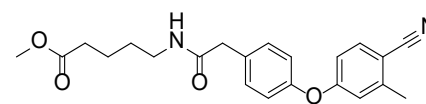
 Prepared according to **general procedure 4** with 2-(4-(4-cyano-3-methylphenoxy)phenyl)acetic acid (150 mg, 1.00 eq., 0.561 mmol), HATU (213 mg, 1.00 eq., 0.561 mmol) and methyl 2-[(2-aminoacetyl)amino]acetate hydrochloride (256 mg, 2.50 eq., 1.40 mmol), DMF (5 mL) and DIPEA (0.342 mL, 3.00 eq., 1.96 mmol). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 90:10 pet spirits:EtOAc. The residue was further purified by trituration with cold Et₂O to give **4.26** as a white solid (51 mg, 23 %). ¹H NMR (401 MHz, *d*₆-DMSO) δ 8.34 (t, J = 5.8 Hz, 2H), 7.74 (d, J = 8.6 Hz, 1H), 7.38 – 7.33 (m, 2H), 7.09 – 7.03 (m, 2H), 7.01 (d, J = 2.4 Hz, 1H), 6.88 (dd, J = 8.6, 2.4 Hz, 1H), 3.86 (d, J = 5.9 Hz, 2H), 3.76 (d, J = 5.8 Hz, 2H), 3.63 (s, 3H), 3.51 (s, 2H), 2.43 (s, 3H). ¹³C NMR (101 MHz, *d*₆-DMSO) δ 170.36, 170.25, 169.43, 161.05, 153.01, 144.34, 134.73, 133.03, 131.04, 119.97, 118.64, 117.95, 115.31, 105.58, 51.71, 41.78, 41.26, 40.52, 19.99. LCMS (m/z): 396.0 [M+H]⁺, t_R = 3.20 min, standard method. HRMS (m/z): C₂₁H₂₁N₃O₅ requires 396.1554 [M+H]⁺; found 396.1555. HPLC: t_R = 5.77 min, >99 %, standard method.

Ethyl 2-((1*s*,4*s*)-4-(2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamido)cyclohexyl)acetate (4.27)

 Prepared according to **general procedure 4** with 2-[4-(4-cyano-3-methylphenoxy)phenyl]acetic acid (300 mg, 1.00 eq., 1.12 mmol), HATU (427 mg, 1.00 eq., 1.12 mmol), DIPEA (0.59 mL, 3.0 eq., 3.37 mmol), dry

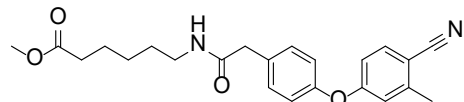
DMF (11 mL) and amine **4.19** (285 mg, 1.37 eq., 1.54 mmol). The residue was purified by flash column chromatography using a gradient of 90:10 pet spirits:EtOAc to 40:60 pet spirits:EtOAc to give **4.27** as a white solid (280 mg, 57 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.54 (d, *J* = 8.5 Hz, 1H), 7.32 – 7.28 (m, 2H), 7.06 – 7.01 (m, 2H), 6.87 (d, *J* = 2.3 Hz, 1H), 6.81 (dd, *J* = 8.5, 2.4 Hz, 1H), 5.51 (d, *J* = 7.3 Hz, 1H), 4.11 (q, *J* = 7.1 Hz, 2H), 4.05 – 3.97 (m, 1H), 3.55 (s, 2H), 2.49 (s, 3H), 2.18 (d, *J* = 7.2 Hz, 2H), 1.97 – 1.85 (m, 1H), 1.66 – 1.56 (m, 6H), 1.24 (t, *J* = 7.1 Hz, 3H), 1.13 – 1.00 (m, 2H). **¹³C NMR** (101 MHz, CDCl₃) δ 172.58, 169.76, 161.18, 154.25, 144.54, 134.39, 131.85, 131.05, 120.71, 119.01, 118.12, 115.39, 106.67, 60.33, 45.34, 43.22, 40.29, 32.55, 28.98, 27.67, 20.64, 14.28. **LCMS** (*m/z*): 434.9 [M+H]⁺, *t_R* = 3.48 min, standard method. **HRMS** (*m/z*): C₂₆H₃₀N₂O₄ requires 435.2278 [M+H]⁺; found 435.2279. **HPLC**: *t_R* = 7.10 min, >99 %, standard method.

Methyl 5-(2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamido)pentanoate (4.28)

 Prepared according to **general procedure 4** with 2-(4-(4-cyano-3-methylphenoxy)phenyl)acetic acid (150 mg, 1.00 eq., 0.561 mmol), amine **4.20** (235 mg, 2.50 eq., 1.40 mmol), HATU (213 mg, 1.00 eq., 0.561 mmol), DMF (5 mL) and DIPEA (0.342 mL, 3.50 eq., 1.96 mmol). The residue was purified by flash column chromatography using a gradient of 75:25 pet spirits:EtOAc to 25:75 pet spirits:EtOAc. The residue was dissolved in DCM (20 mL) and wash with H₂O (2 × 30 mL), hydrochloride (aq., 0.1 M, 2 × 30 mL) brine (10 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness to give **4.28** as a white solid (139 mg, 65 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.46 (d, *J* = 8.5 Hz, 1H), 7.26 – 7.21 (m, 2H), 6.98 – 6.93 (m, 2H), 6.80 (d, *J* = 2.4 Hz, 1H), 6.75 (dd, *J* = 8.5, 2.3 Hz, 1H), 5.59 (bs, 1H), 3.59 (s, 3H), 3.48 (s, 2H), 3.18 (dt, *J* = 6.7, 5.9 Hz, 2H), 2.42 (s, 3H), 2.25 (t, *J* = 7.1 Hz, 2H), 1.59 – 1.49 (m, 2H), 1.49 – 1.40 (m, 2H). **¹³C NMR** (101 MHz, CDCl₃) δ 173.99, 170.75, 161.28, 154.30, 144.60, 134.46, 131.79, 131.20, 120.73, 119.09, 118.24, 115.54, 106.70, 51.70, 43.11, 39.33, 33.49, 29.01, 22.04, 20.72. **LCMS** (*m/z*): 380.9 [M+H]⁺, *t_R* = 3.30 min, standard method. **HRMS** (*m/z*):

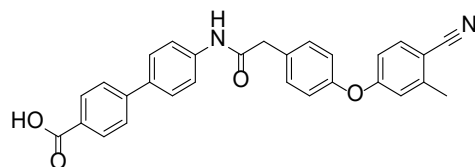
$C_{22}H_{24}N_2O_4$ requires 381.1809 $[M+H]^+$; found 381.1813. **HPLC**: t_R = 6.54 min, 95 %, standard method.

Methyl 6-(2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamido)hexanoate (4.29)



Prepared according to **general procedure 4** with 2-(4-(4-cyano-3-methylphenoxy)phenyl)acetic acid (150 mg, 1.00 eq., 0.561 mmol), amine **4.21** (255 mg, 2.50 eq., 1.40 mmol), HATU (213 mg, 1.00 eq., 0.561 mmol), DMF (5 mL) and DIPEA (0.342 mL, 3.50 eq., 1.96 mmol). The residue was purified by flash column chromatography using a gradient of 75:25 pet spirits:EtOAc to 25:75 pet spirits:EtOAc. The residue was dissolved in DCM (20 mL) and further washed with H_2O (2×30 mL), hydrochloride (aq., 0.1 M, 2×30 mL), brine (10 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness to give **4.29** as a white solid (126 mg, 57 %). **1H NMR** (401 MHz, $CDCl_3$) δ 7.51 (d, J = 8.5 Hz, 1H), 7.31 – 7.26 (m, 2H), 7.03 – 6.98 (m, 2H), 6.85 (d, J = 2.4 Hz, 1H), 6.80 (dd, J = 8.5, 2.3 Hz, 1H), 5.68 (bs, 1H), 3.64 (s, 3H), 3.53 (s, 2H), 3.23 (dt, J = 7.1, 5.9 Hz, 2H), 2.47 (s, 3H), 2.27 (t, J = 7.4 Hz, 2H), 1.64 – 1.56 (m, 2H), 1.51 – 1.44 (m, 2H), 1.33 – 1.23 (m, 2H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 174.07, 170.69, 161.26, 154.22, 144.57, 134.42, 131.87, 131.14, 120.66, 119.04, 118.20, 115.50, 106.64, 51.59, 43.05, 39.51, 33.86, 29.21, 26.33, 24.45, 20.68. **LCMS** (m/z): 395.0 $[M+H]^+$, t_R = 3.34 min, standard method. **HRMS** (m/z): $C_{23}H_{26}N_2O_4$ requires 395.1965 $[M+H]^+$; found 395.1973. **HPLC**: t_R = 6.74 min, >99 %, standard method.

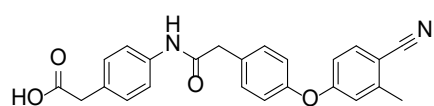
4'-(2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetamido)-[1,1'-biphenyl]-4-carboxylic acid (4.30)



Prepared according to **general procedure 5** with ester **4.22** (161 mg, 1.00 eq., 0.338 mmol) and LiOH (161 mg, 11.1 eq., 3.75 mmol) in a 1:1 THF: H_2O mixture (6 mL) for 4 days at 50 °C. The product **4.30** was isolated as a white solid (46 mg, 29 %) after workup. **1H NMR** (401 MHz, d_6 -DMSO) δ 10.36 (s, 1H), 8.03 – 7.96 (m, 2H), 7.82 – 7.69 (m, 7H), 7.47 – 7.41 (m, 2H), 7.14 – 7.09 (m, 2H), 7.04 (d, J = 2.3 Hz, 1H), 6.90 (dd, J = 8.6, 2.5 Hz, 1H), 3.72 (s, 2H), 2.44 (s, 3H). **^{13}C NMR** (101 MHz, d_6 -DMSO) δ 174.40, 172.40, 166.19, 158.46, 149.58, 148.96, 144.65, 139.98,

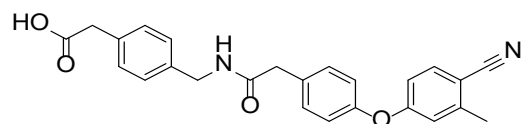
138.83, 137.86, 136.31, 135.18, 134.40, 132.54, 131.44, 125.30, 124.70, 123.95, 123.16, 120.59, 110.88, 47.76, 25.21. **LCMS** (m/z): 460.9 $[M-H]^-$, t_R = 3.42 min, standard method. **HRMS** (m/z): $C_{29}H_{22}N_2O_4$ requires 463.1652 $[M+H]^+$; found 463.1653. **HPLC**: t_R = 7.00 min, >99 %, standard method.

2-(4-(2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetamido)phenyl)acetic acid (4.31)



Ester **4.23** (85 mg, 1.0 eq., 0.20 mmol) and LiOH (50 mg, 6.0 eq., 1.2 mmol) were dissolved in a 1:1 DMSO:H₂O mixture (2 mL) and stirred for 2 h at 70 °C. After this time, the aqueous layer was acidified with hydrochloride (aq., 36 % w/v, 5 drops) and diluted in Et₂O (30 mL) and H₂O (30 mL). The layers were separated and the aqueous layer further extracted with Et₂O (2 × 30 mL). The combined organic layers were washed with brine (2 × 5 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 50:50 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **4.31** as a white solid (41 mg, 51 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 10.17 (s, 1H), 7.75 (d, J = 8.6 Hz, 1H), 7.54 (d, J = 8.5 Hz, 2H), 7.42 (d, J = 8.6 Hz, 2H), 7.18 (d, J = 8.5 Hz, 2H), 7.12 – 7.07 (m, 2H), 7.04 (d, J = 2.4 Hz, 1H), 6.89 (dd, J = 8.6, 2.5 Hz, 1H), 3.66 (s, 2H), 3.49 (s, 2H), 2.44 (s, 3H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 172.80, 168.87, 160.98, 153.18, 144.36, 137.71, 134.74, 132.79, 131.03, 129.92, 129.63, 120.05, 119.02, 118.71, 117.94, 115.37, 105.63, 42.48, 28.99, 19.99. **LCMS** (m/z): 354.9 $[M-CO_2]^-$, t_R = 3.85 min, standard method. **HRMS** (m/z): $C_{24}H_{21}N_2O_4$ requires 401.1496 $[M+H]^+$; found 401.1499. **HPLC**: t_R = 6.31 min, >99 %, standard method.

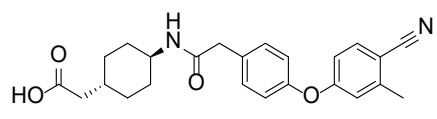
2-(4-((2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetamido)methyl)phenyl)acetic acid (4.32)



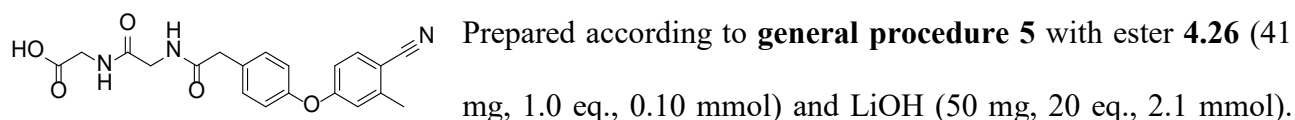
Ester **4.24** (157 mg, 1.00 eq., 0.366 mmol) and LiOH (154 mg, 10.0 eq., 3.66 mmol) were dissolved in a 1:1:1 THF:DMSO:H₂O mixture (6 mL) and stirred for 3 h at 86 °C. After this time, the THF was removed under vacuum and the aqueous layer acidified with hydrochloride (aq., 1 M, 15 mL). The aqueous layer was diluted with H₂O (40 mL) and Et₂O (30 mL). The layers were separated and the aqueous

layer further extracted with Et₂O (2 × 30 mL). The combined organic layers were washed with brine (2 × 10 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 50:50 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **4.32** as a white solid (84 mg, 55 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 12.28 (bs, 1H), 8.54 (t, *J* = 5.8 Hz, 1H), 7.75 (d, *J* = 8.6 Hz, 1H), 7.38 – 7.34 (m, 2H), 7.21 – 7.15 (m, 4H), 7.10 – 7.05 (m, 2H), 7.02 (d, *J* = 2.4 Hz, 1H), 6.88 (dd, *J* = 8.6, 2.5 Hz, 1H), 4.25 (d, *J* = 5.8 Hz, 2H), 3.53 (s, 2H), 3.50 (s, 2H), 2.44 (s, 3H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 172.72, 169.96, 161.05, 153.02, 144.34, 137.67, 134.73, 133.56, 133.23, 130.95, 129.29, 127.17, 120.02, 118.64, 117.95, 115.30, 105.59, 41.97, 41.55, 40.32, 20.01. **LCMS** (*m/z*): 412.9 [M-H]⁻, *t*_R = 3.25 min, standard method. **HRMS** (*m/z*): C₂₅H₂₂N₂O₄ requires 415.1652 [M+H]⁺; found 415.1662. **HPLC**: *t*_R = 6.19 min, >99 %, standard method.

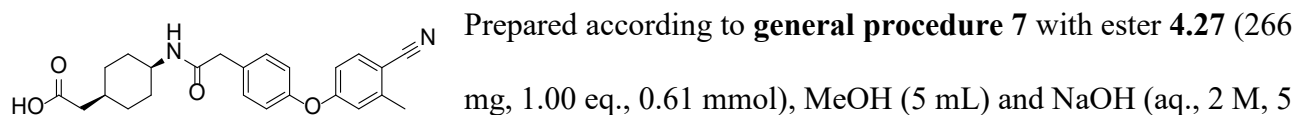
2-((1*r*,4*r*)-4-(2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetamido)cyclohexyl)acetic acid (4.33**)**

 Ester **4.25** (126 mg, 1.00 eq., 0.29 mmol) and LiOH·H₂O (118 mg, 9.70 eq., 2.81 mmol) were dissolved in a 4:4:1

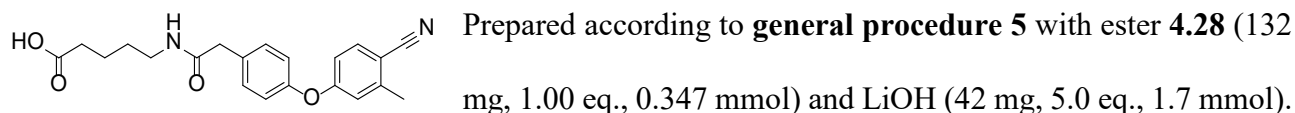
THF:H₂O:DMSO mixture (6 mL) and refluxed for 24 h. The reaction was worked up according to **general procedure 7** and purified by flash column chromatography using a gradient of 40:60 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **4.33** as a white solid (36 mg, 31 %). **¹H NMR** (600 MHz, *d*₆-DMSO) δ 7.94 (d, *J* = 7.8 Hz, 1H), 7.74 (d, *J* = 8.6 Hz, 1H), 7.33 – 7.29 (m, 2H), 7.07 – 7.03 (m, 2H), 7.01 (d, *J* = 2.2 Hz, 1H), 6.87 (dd, *J* = 8.6, 2.4 Hz, 1H), 3.45 (dt, *J* = 15.5, 7.9, 3.9 Hz, 1H), 3.39 (s, 2H), 2.43 (s, 3H), 2.09 (d, *J* = 6.9 Hz, 2H), 1.80 – 1.75 (m, 2H), 1.73 – 1.70 (m, 2H), 1.64 – 1.55 (m, 1H), 1.17 (qd, *J* = 12.7, 3.0 Hz, 2H), 1.00 (qd, *J* = 13.1, 2.9 Hz, 2H). **¹³C NMR** (101 MHz, DMSO) δ 173.73, 168.98, 161.05, 152.94, 144.34, 134.74, 133.50, 130.77, 119.97, 118.66, 117.96, 115.30, 105.59, 47.68, 41.62, 41.04, 33.59, 32.15, 31.09, 20.01. **LCMS** (*m/z*): 406.9 [M+H]⁺, *t*_R = 3.247 min, standard method. **HRMS** (*m/z*): C₂₄H₂₆N₂O₄ requires 407.1965 [M+H]⁺; found 407.1975. **HPLC**: *t*_R = 6.15 min, >99 %, standard method.

(2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetyl)glycylglycine (4.34)

The residue was further purified by trituration with C hydrochloride₃ giving **4.34** as a white solid (4.3 mg, 11 %, 60 mol % purity). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 8.33 (t, *J* = 5.9 Hz, 1H), 8.17 (bs, 1H), 7.74 (d, *J* = 8.6 Hz, 1H), 7.37 – 7.35 (m, 2H), 7.07 – 7.05 (m, 2H), 7.02 (d, *J* = 2.2 Hz, 1H), 6.88 (dd, *J* = 8.5, 2.2 Hz, 1H), 3.75 (app d, *J* = 5.7 Hz, 4H), 3.51 (s, 2H), 2.43 (s, 3H). **LCMS** (*m/z*): 379.9 [M-H]⁻, *t*_R = 3.15 min, standard method. **HRMS** (*m/z*): C₂₀H₁₉N₃O₅ requires 382.1397 [M+H]⁺; found 382.1403. **HPLC**: *t*_R = 5.40 min, 97 %, standard method.

2-((1s,4s)-4-(2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetamido)cyclohexyl)acetic acid (4.35)

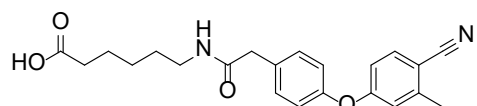
¹H NMR (401 MHz, *d*₆-DMSO) δ 12.05 (bs, 1H), 7.91 (d, *J* = 7.5 Hz, 1H), 7.74 (d, *J* = 8.6 Hz, 1H), 7.37 – 7.29 (m, 2H), 7.09 – 7.03 (m, 2H), 7.01 (d, *J* = 2.3 Hz, 1H), 6.87 (dd, *J* = 8.6, 2.4 Hz, 1H), 3.76 (bs, 1H), 3.45 (s, 2H), 2.43 (s, 3H), 2.16 (d, *J* = 7.3 Hz, 2H), 1.83 – 1.72 (m, 1H), 1.57 – 1.41 (m, 6H), 1.40 – 1.27 (m, 2H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 173.80, 169.28, 161.08, 152.95, 144.35, 134.74, 133.65, 130.82, 119.99, 118.67, 117.97, 115.30, 105.61, 44.78, 41.54, 39.66 (confirmed via HSQC), 32.14, 28.66, 27.19, 20.02. **LCMS** (*m/z*): 406.9 [M+H]⁺, *t*_R = 3.22 min, standard method. **HRMS** (*m/z*): C₂₄H₂₆N₂O₄ requires 407.1965 [M+H]⁺; found 407.1970. **HPLC**: *t*_R = 6.18 min, >99 %, standard method.

5-(2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetamido)pentanoic acid (4.36)

The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc

to 0:100 pet spirits:EtOAc to give **4.36** as a white solid (27 mg, 21 %). **¹H NMR** (401 MHz, CD₃CN) δ 9.17 (bs, 1H), 7.61 (d, J = 8.6 Hz, 1H), 7.35 – 7.30 (m, 2H), 7.05 – 6.99 (m, 2H), 6.94 (d, J = 2.3 Hz, 1H), 6.84 (dd, J = 8.6, 2.4 Hz, 1H), 6.54 (bs, 1H) 3.45 (s, 2H), 3.14 (app dd, J = 12.7, 6.6 Hz, 2H), 2.45 (s, 3H), 2.27 (t, J = 7.3 Hz, 2H), 1.59 – 1.41 (m, 4H). **¹³C NMR** (101 MHz, CD₃CN) δ 175.03, 171.54, 162.42, 154.78, 145.64, 135.53, 134.09, 131.98, 121.18, 119.81, 118.89, 116.25, 107.26, 43.06, 39.59, 33.76, 29.61, 22.82, 20.68. **LCMS** (m/z): 366.9 [M+H]⁺, t_R = 3.17 min, standard method. **HRMS** (m/z): C₂₁H₂₂N₂O₄ requires 367.1652 [M+H]⁺; found 367.1656. **HPLC**: t_R = 5.91 min, >99 %, standard method.

6-(2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetamido)hexanoic acid (4.37)

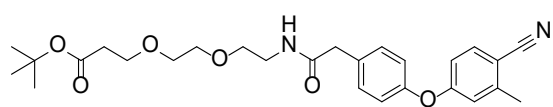


Prepared according to **general procedure 4** with ester **4.29** (120 mg, 1.00 eq., 0.304 mmol) and LiOH (36 mg, 5.0 eq., 1.5 mmol). The residue was purified by flash column chromatography using a gradient of 100:0 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **4.37** as a white solid (14 mg, 12 %). **¹H NMR** (401 MHz, CD₃CN) δ 7.61 (d, J = 8.6 Hz, 1H), 7.35 – 7.29 (m, 2H), 7.05 – 7.00 (m, 2H), 6.95 (d, J = 2.3 Hz, 1H), 6.84 (dd, J = 8.6, 2.4 Hz, 1H), 6.51 (bs, 1H), 3.44 (s, 2H), 3.12 (dt, J = 6.9, 6.0 Hz, 2H), 2.45 (s, 3H), 2.24 (t, J = 7.4 Hz, 2H), 1.59 – 1.50 (m, 2H), 1.49 – 1.38 (m, 2H), 1.34 – 1.24 (m, 2H). **¹³C NMR** (101 MHz, CD₃CN) δ 175.15, 171.49, 162.41, 154.78, 145.64, 135.53, 134.12, 131.97, 121.16, 119.82, 118.88, 116.25, 107.27, 43.06, 39.84, 34.12, 29.82, 26.96, 25.21, 20.68. **LCMS** (m/z): 380.9 [M+H]⁺, t_R = 3.20 min, standard method. **HRMS** (m/z): C₂₂H₂₄N₂O₄ requires 381.1809 [M+H]⁺; found 381.1815. **HPLC**: t_R = 6.08 min, >99 %, standard method.

***Tert*-Butyl**

3-(2-(2-(2-(4-(4-cyano-3-

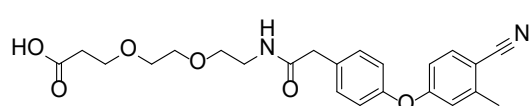
methylphenoxy)phenyl)acetamido)ethoxy)ethoxy)propanoate (4.38)



Prepared according to **general procedure 4** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (222 mg, 1.00 eq., 0.831 mmol), *tert*-butyl 3-[2-(2-aminoethoxy)ethoxy]propanoate (480 mg, 2.48 eq., 2.08 mmol), HATU (324 mg, 1.03 eq., 0.852 mmol), DMF (4 mL) and DIPEA (0.518 mL, 3.5 eq.,

2.91 mmol). The residue was purified by flash column chromatography using a gradient of 50:50 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **4.38** as a colourless oil (183 mg, 46 %). **¹H NMR** (401 MHz, CD₃CN) δ 7.61 (d, J = 8.6 Hz, 1H), 7.36 – 7.31 (m, 2H), 7.06 – 7.01 (m, 2H), 6.95 (d, J = 2.4 Hz, 1H), 6.85 (dd, J = 8.6, 2.5 Hz, 1H), 6.60 (bs, 1H), 3.63 (t, J = 6.2 Hz, 2H), 3.51 (s, 4H), 3.49 – 3.44 (m, 4H), 3.30 (app q, J = 5.6 Hz, 2H), 2.46 (s, 3H), 2.41 (t, J = 6.2 Hz, 2H), 1.42 (s, 9H). **¹³C NMR** (101 MHz, CD₃CN) δ 172.15, 171.81, 162.38, 154.82, 145.64, 135.53, 133.94, 131.99, 121.15, 119.85, 118.86, 116.28, 107.31, 81.14, 70.90, 70.86, 70.46, 67.48, 42.93, 40.08, 36.97, 28.30, 20.69. **LCMS** (m/z): 482.9 [M+H]⁺, t_R = 3.42 min, standard method. **HRMS** (m/z): C₂₇H₃₄N₂O₆ requires 482.2394 [M]⁺; found 482.2417. **HPLC**: t_R = 7.06 min, >99 %, standard method.

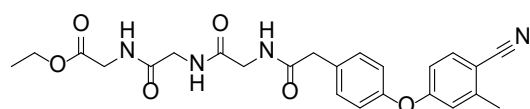
3-(2-(2-(2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetamido)ethoxy)ethoxy)propanoic acid (4.39)



Ester **4.38** (163 mg, 1.00 eq., 0.338 mmol) and TFA (231 mg, 6.0 eq., 2.03 mmol) were dissolved DCM (5 mL) and

stirred for 18 h at room temperature. After this time, the solvent was removed under vacuum and the residue purified by flash column chromatography using a gradient of 50:50 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **4.39** as a colourless oil (100 mg, 69 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.52 (d, J = 8.5 Hz, 1H), 7.35 – 7.30 (m, 2H), 7.02 – 6.97 (m, 2H), 6.86 (d, J = 2.4 Hz, 1H), 6.81 (dd, J = 8.5, 2.4 Hz, 1H), 6.34 (t, J = 5.4 Hz, 1H), 3.76 (t, J = 6.0 Hz, 2H), 3.63 – 3.56 (m, 6H), 3.54 (t, J = 5.1 Hz, 2H), 3.44 (t, J = 5.3 Hz, 2H), 2.62 (t, J = 6.0 Hz, 2H), 2.48 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 174.56, 171.60, 161.41, 154.19, 144.63, 134.47, 131.91, 131.25, 120.67, 119.03, 118.29, 115.50, 106.63, 70.50, 69.96, 69.89, 66.52, 42.80, 39.67, 34.88, 20.74. **LCMS** (m/z): 426.9 [M-H]⁻, t_R = 3.22 min, standard method. **HRMS** (m/z): C₂₃H₂₆N₂O₆ requires 427.1864 [M+H]⁺; found 427.1870. **HPLC**: t_R = 5.74 min, >99 %, standard method.

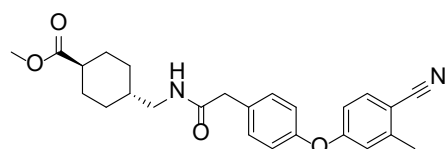
Ethyl (2-(4-(4-cyano-3-methylphenoxy)phenyl)acetyl)glycylglycylglycinate (4.40)



2-[4-(4-Cyano-3-methyl-phenoxy)phenyl]acetic acid (150 mg, 1.00 eq., 0.561 mmol), HATU (213 mg, 1.00

eq., 0.561 mmol) and ethyl 2-[[2-[(2-aminoacetyl)amino]acetyl]amino]acetate hydrochloride (285 mg, 2.00 eq., 1.12 mmol) dissolved in dry DMF (5 mL) was added DIPEA (0.342 mL, 3.00 eq., 1.96 mmol) dropwise under N₂ at room temperature. The reaction mixture was then stirred for 16 h at 25 °C. After this time, the reaction was diluted in EtOAc (30 mL) and filtered. The precipitate was washed with cold EtOAc to give **4.40** as a white solid (163 mg, 62 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 8.34 (t, *J* = 5.6 Hz, 1H), 8.28 (t, *J* = 5.9 Hz, 1H), 8.22 (t, *J* = 5.9 Hz, 1H), 7.75 (d, *J* = 8.6 Hz, 1H), 7.39 – 7.34 (m, 2H), 7.09 – 7.04 (m, 2H), 7.02 (d, *J* = 2.5 Hz, 1H), 6.88 (dd, *J* = 8.6, 2.3 Hz, 1H), 4.09 (q, *J* = 7.1 Hz, 2H), 3.82 (d, *J* = 5.9 Hz, 2H), 3.77 (d, *J* = 5.6 Hz, 2H), 3.76 (d, *J* = 5.7 Hz, 2H), 3.53 (s, 2H), 2.44 (s, 3H), 1.19 (t, *J* = 7.1 Hz, 3H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 170.51, 169.67, 169.34, 169.21, 161.05, 153.02, 144.34, 134.73, 133.04, 131.04, 119.98, 118.65, 117.95, 115.31, 105.58, 60.43, 42.17, 41.71, 41.23, 40.62, 20.00, 14.06. **LCMS** (*m/z*): 466.9 [M+H]⁺, *t*_R = 3.13 min, standard method. **HRMS** (*m/z*): C₂₄H₂₆N₄O₆ requires 467.1925 [M+H]⁺; found 467.1934. **HPLC**: *t*_R = 5.79 min, >99 %, standard method.

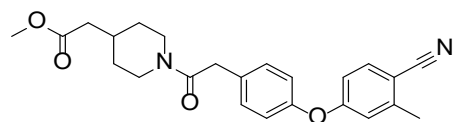
Methyl (1*r*,4*r*)-4-((2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamido)methyl)cyclohexane-1-carboxylate (4.41**)**



Prepared according to **general procedure 4** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (201 mg, 1.00 eq., 0.752 mmol), HATU (294 mg, 1.03 eq., 0.773 mmol), methyl *trans*-4-(aminomethyl)cyclohexane-1-carboxylate (313 mg, 2.00 eq., 1.51 mmol), DIPEA (0.458 mL, 3.5 eq., 2.63 mmol) and dry DMF (5 mL). The residue was purified by flash column chromatography using a gradient of 80:20 pet spirits:EtOAc to 40:60 pet spirits:EtOAc to give **4.41** as a white solid (129 mg, 41 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.53 (d, *J* = 8.5 Hz, 1H), 7.32 – 7.27 (m, 2H), 7.06 – 7.00 (m, 2H), 6.86 (d, *J* = 2.4 Hz, 1H), 6.83 – 6.77 (m, 1H), 5.55 (t, *J* = 5.6 Hz, 1H), 3.65 (s, 3H), 3.56 (s, 2H), 3.10 (t, *J* = 6.5 Hz, 2H), 2.49 (s, 3H), 2.20 (tt, *J* = 12.2, 3.6 Hz, 1H), 2.02 – 1.92 (m, 2H), 1.79 – 1.68 (m, 2H), 1.49 – 1.41 (m, 1H), 1.37 (qd, *J* = 12.8, 3.3 Hz, 2H), 0.92 (qd, *J* = 13.1, 3.4 Hz, 2H). **¹³C NMR** (101 MHz, CDCl₃) δ 176.29, 170.80, 161.26, 154.37, 144.63, 134.48, 131.82,

131.19, 120.77, 119.11, 118.22, 115.54, 106.77, 51.70, 45.63, 43.19, 43.16, 37.41, 29.82, 28.48, 20.74. **LCMS** (m/z): 420.9 $[M+H]^+$, t_R = 3.43 min, standard method. **HRMS** (m/z): $C_{25}H_{28}N_2O_4$ requires 421.2122 $[M+H]^+$; found 421.2125. **HPLC**: t_R = 6.81 min, >99 %, standard method.

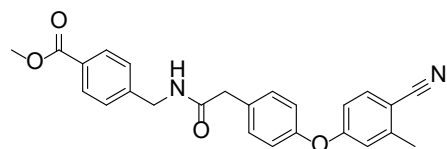
Methyl 2-(1-(2-(4-(4-cyano-3-methylphenoxy)phenyl)acetyl)piperidin-4-yl)acetate (4.42)



Prepared according to **general procedure 4** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (203 mg, 1.00 eq., 0.760

mmol), HATU (289 mg, 1.00 eq., 0.760 mmol), DIPEA (0.463 mL, 3.5 eq., 2.66 mmol), dry DMF (5 mL), methyl 2-(4-piperidyl)acetate (246 mg, 2.06 eq., 1.56 mmol). The residue was purified by flash column chromatography using a gradient of 60:40 petroleum spirits:EtOAc to 20:80 petroleum spirits:ethyl acetate to give **4.42** as a colourless oil (75 mg, 25 %). **1H NMR** (401 MHz, $CDCl_3$) δ 7.52 (d, J = 8.5 Hz, 1H), 7.29 – 7.24 (m, 2H), 7.02 – 6.96 (m, 2H), 6.85 (d, J = 2.4 Hz, 1H), 6.82 – 6.77 (m, 1H), 4.67 – 4.59 (m, 1H), 3.92 – 3.85 (m, 1H), 3.72 (s, 2H), 3.66 (s, 3H), 3.08 – 2.99 (m, 1H), 2.61 (td, J = 13.0, 2.7 Hz, 1H), 2.48 (s, 3H), 2.29 – 2.16 (m, 2H), 2.07 – 1.94 (m, 1H), 1.74 (t, J = 11.9 Hz, 2H), 1.14 (qd, J = 12.4, 4.3 Hz, 1H), 0.99 (qd, J = 12.5, 4.2 Hz, 1H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 172.70, 169.15, 161.51, 153.82, 144.57, 134.45, 132.14, 130.66, 120.66, 118.92, 118.30, 115.41, 106.55, 51.72, 46.27, 42.13, 40.68, 40.18, 33.10, 32.34, 31.65, 20.74. **LCMS** (m/z): 406.9 $[M+H]^+$, t_R = 3.40 min, standard method. **HRMS** (m/z): $C_{24}H_{26}N_2O_4$ requires 407.1965 $[M+H]^+$; found 407.1969. **HPLC**: t_R = 6.83 min, >99 %, standard method.

Methyl 4-((2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamido)methyl)benzoate (4.43)

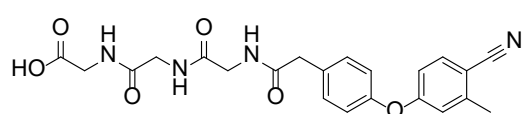


Prepared according to **general procedure 4** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.0 eq., 0.748 mmol), HATU (285 mg, 1.0 eq., 0.748 mmol), methyl 4-

(aminomethyl)benzoate hydrochloride (302 mg, 2.0 eq., 1.50 mmol), dry DMF (5 mL) and DIPEA (0.74 mL, 3.5 eq., 2.6 mmol). The residue was purified by flash column chromatography using a gradient of 70:30 petroleum spirits:ethyl acetate to 30:70 petroleum spirits:ethyl acetate to give **4.43** as a white solid (50 mg, 16 %). **1H NMR** (401 MHz, d_6 -DMSO) δ 8.67 (t, J = 5.9 Hz, 1H), 7.90 (app

d, $J = 8.3$ Hz, 2H), 7.75 (d, $J = 8.6$ Hz, 1H), 7.39 – 7.33 (m, 4H), 7.11 – 7.05 (m, 2H), 7.01 (d, $J = 2.3$ Hz, 1H), 6.87 (dd, $J = 8.6, 2.4$ Hz, 1H), 4.36 (d, $J = 6.0$ Hz, 2H), 3.84 (s, 3H), 3.53 (s, 2H), 2.43 (s, 3H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 170.24, 166.08, 161.06, 153.07, 145.20, 144.34, 134.72, 133.11, 130.96, 129.21, 128.13, 127.30, 120.08, 118.64, 117.95, 115.28, 105.60, 52.07, 41.94, 41.56, 20.00. LCMS (m/z): 414.9 $[\text{M}+\text{H}]^+$, $t_R = 3.48$ min, standard method. HRMS (m/z): $\text{C}_{25}\text{H}_{22}\text{N}_2\text{O}_4$ requires 415.1652 $[\text{M}+\text{H}]^+$; found 415.1661. HPLC: $t_R = 6.82$ min, >99 %, standard method.

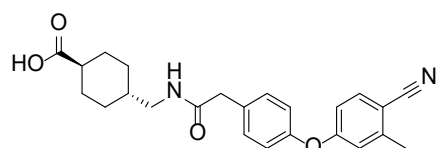
(2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetyl)glycylglycylglycine (4.44)



Prepared according to **general procedure 8** with oxalyl chloride (0.32 mL, 5.0 eq., 3.74 mmol), dry DCM (4 mL)

and 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.0 eq., 0.75 mmol) stirred for 6 h at 0 °C, then DMF (6 mL) and triglycine (184 mg, 1.3 eq., 0.97 mmol) at room temperature for 24 h. The residue was purified by flash column chromatography using a gradient of 50:50:0 pet spirits:EtOAc:MeOH to 0:60:40 pet spirits:EtOAc:MeOH with 0.5 % AcOH added at 60:40 EtOAc:MeOH. The residue was dissolved in DMF (5 mL), filtered and the solution evaporated to dryness to give **4.44** as a brown solid (5 mg, 1.52 %). ^1H NMR (401 MHz, d_6 -DMSO) δ 8.41 (t, $J = 5.6$ Hz, 1H), 8.26 (t, $J = 5.8$ Hz, 1H), 7.74 (d, $J = 8.6$ Hz, 1H), 7.65 (bs, 1H), 7.36 (d, $J = 8.5$ Hz, 2H), 7.05 (d, $J = 8.6$ Hz, 2H), 7.01 (d, $J = 2.3$ Hz, 1H), 6.88 (dd, $J = 8.6, 2.4$ Hz, 1H), 3.76 (d, $J = 5.7$ Hz, 2H), 3.71 (d, $J = 5.9$ Hz, 2H), 3.52 (s, 2H), 3.48 (bs, 2H), 2.43 (s, 3H). ^{13}C NMR could not be obtained due to low solubility in neat DMSO and insufficient material. LCMS (m/z): 438.8 $[\text{M}+\text{H}]^+$, $t_R = 3.01$ min, standard method. HRMS (m/z): $\text{C}_{22}\text{H}_{22}\text{N}_4\text{O}_6$ requires 439.1612 $[\text{M}+\text{H}]^+$; found 439.1615. HPLC: $t_R = 5.17$ min, >99 %, standard method.

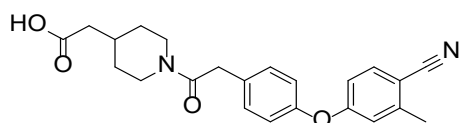
(1*r*,4*r*)-4-((2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetamido)methyl)cyclohexane-1-carboxylic acid (4.45)



Prepared according to **general procedure 8** with oxalyl chloride (0.2 mL, 3.12 eq., 2.34 mmol), dry DCM (8 mL), 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.0 eq., 0.74

mmol), dry DMF (10 mL) and trans-4-(aminomethyl)cyclohexanecarboxylic acid (592 mg, 5.0 eq., 3.74 mmol). The residue was purified by flash column chromatography using a gradient of 60:40 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **4.45** as a white solid (115 mg, 38 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 8.04 (t, *J* = 5.7 Hz, 1H), 7.74 (d, *J* = 8.6 Hz, 1H), 7.35 – 7.31 (m, 2H), 7.09 – 7.03 (m, 2H), 7.00 (d, *J* = 2.4 Hz, 1H), 6.86 (dd, *J* = 8.5, 2.4 Hz, 1H), 3.42 (s, 2H), 2.90 (t, *J* = 6.2 Hz, 2H), 2.43 (s, 3H), 2.09 (tt, *J* = 12.2, 3.5 Hz, 1H), 1.87 (dd, *J* = 13.4, 2.6 Hz, 2H), 1.70 (dd, *J* = 13.1, 2.6 Hz, 2H), 1.39 – 1.12 (m, 3H), 0.89 (qd, *J* = 12.9, 3.1 Hz, 2H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 176.74, 169.94, 161.08, 152.94, 144.33, 134.72, 133.50, 130.85, 120.00, 118.61, 117.95, 115.25, 105.57, 44.72, 42.49, 41.63, 37.00, 29.40, 28.28, 20.01. **LCMS** (*m/z*): 406.9 [M+H]⁺, *t*_R = 3.27 min, standard method. **HRMS** (*m/z*): C₂₄H₂₆N₂O₄ requires 407.1965 [M+H]⁺; found 407.1968. **HPLC**: *t*_R = 6.08 min, 97 %, standard method.

2-(1-(2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetyl)piperidin-4-yl)acetic acid (4.46)

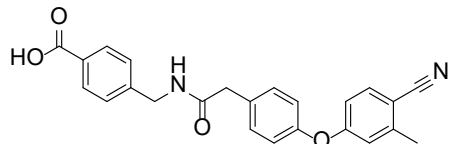


Prepared according to **general procedure 8** with oxalyl chloride (0.64 mL, 5.0 eq., 7.48 mmol), dry DCM (5 mL) and

2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (400 mg, 1.0 eq., 1.50 mmol) stirred for 6 h at 0 °C. Then dry DMF (6 mL) and 2-(4-piperidyl)acetic acid hydrochloride (3.50 mg, 1.3 eq., 1.95 mmol) at room temperature for 24 h. The residue was purified by flash column chromatography using a gradient of 100:0 DCM:MeOH to 95:5 DCM:MeOH to give **4.46** as a white solid (78 mg, 13 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 12.09 (s, 1H), 7.75 (d, *J* = 8.6 Hz, 1H), 7.34 – 7.26 (m, 2H), 7.10 – 7.04 (m, 2H), 7.01 (d, *J* = 2.4 Hz, 1H), 6.86 (dd, *J* = 8.6, 2.5 Hz, 1H), 4.36 (app d, *J* = 13.1 Hz, 1H), 3.94 (app d, *J* = 13.3 Hz, 1H), 3.71 (s, 2H), 3.05 – 2.95 (m, 1H), 2.56 (td, *J* = 12.7, 2.1 Hz, 1H), 2.43 (s, 3H), 2.12 (d, *J* = 7.0 Hz, 2H), 1.93 – 1.81 (m, 1H), 1.64 (app t, *J* = 11.5 Hz, 2H), 1.04 – 0.86 (m, 2H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 173.35, 168.40, 161.06, 152.90, 144.33, 134.73, 132.95, 130.91, 120.09, 118.67, 117.94, 115.23, 105.60, 45.44, 41.27, 40.36, 38.81 (confirmed via HSQC), 31.15, 20.01. **LCMS** (*m/z*): 393.0 [M+H]⁺, *t*_R = 3.25 min, standard method. **HRMS** (*m/z*):

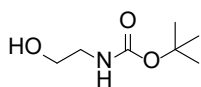
$C_{23}H_{24}N_2O_4$ requires 393.1809 $[M+H]^+$; found 393.1816. **HPLC**: t_R = 6.04 min, >99 %, standard method.

4-((2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetamido)methyl)benzoic acid (4.47)



Prepared according to **general procedure 8** with oxalyl chloride (0.24 mL, 3.00 eq., 2.81 mmol), dry DCM (10 mL), 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (250 mg, 1.00 eq., 0.94 mmol) stirred for 4 h at 0 °C, then dry DMF (10 mL), 4-(aminomethyl)benzoic acid (282 mg, 2.0 eq., 1.87 mmol) and DIPEA (0.50 mL, 3.0 eq., 2.81 mmol) at room temperature for 72 h. The residue was purified by flash column chromatography a gradient of 59:40:1 pet spirits:EtOAc:AcOH to 0:99:1 pet spirits:EtOAc:AcOH to give **4.47** as a white solid (66 mg, 18 %). **¹H NMR** (400 MHz, d_6 -DMSO) δ 8.66 (t, J = 5.9 Hz, 1H), 7.87 (d, J = 8.2 Hz, 2H), 7.75 (d, J = 8.6 Hz, 1H), 7.40 – 7.35 (m, 2H), 7.33 (d, J = 8.3 Hz, 2H), 7.11 – 7.05 (m, 2H), 7.01 (d, J = 2.4 Hz, 1H), 6.87 (dd, J = 8.6, 2.4 Hz, 1H), 4.35 (d, J = 5.9 Hz, 2H), 3.52 (s, 2H), 2.43 (s, 3H). **¹³C NMR** (101 MHz, d_6 -DMSO) δ 170.23, 167.19, 161.08, 153.08, 144.65, 144.36, 134.74, 133.14, 130.98, 129.37, 127.15, 120.09, 118.65, 117.96, 115.29, 105.61, 41.99, 41.59, 20.02. **Missing one carbon resonance, not found via HSQC or HMBC experiments.** **LCMS** (m/z): 400.9 $[M+H]^+$, t_R = 3.29 min, standard method. **HRMS** (m/z): $C_{24}H_{20}N_2O_4$ requires 401.1496 $[M+H]^+$; found 401.1491. **HPLC**: t_R = 6.10 min, >99 %, standard method.

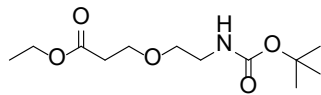
tert-Butyl (2-hydroxyethyl)carbamate (4.52)



To a solution of ethanolamine (1.00 g, 1.00 eq., 16.4 mmol) dissolved in dry DCM (80 mL) was added Boc_2O (3.75 g, 1.05 eq., 17.2 mmol) dissolved in DCM (20 mL) dropwise at 0 °C. The reaction was allowed to warm to room temperature and stirred for 24 h. After this time, the reaction was diluted in H_2O (50 mL). The layers were separated and the organic washed with H_2O (50 mL) and brine (30 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness to give the product as a colourless oil (712 mg, 26 %). **¹H NMR** (401 MHz, $CDCl_3$) δ 5.29 (s, 1H), 3.67 (s, 1H), 3.59 (t, J = 4.9 Hz, 2H), 3.19 (d, J = 5.0 Hz, 2H), 1.37 (s, 9H). **¹³C NMR** (101

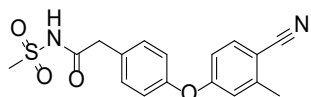
MHz, CDCl₃) δ 156.92, 79.68, 62.29, 43.15, 28.45. **LCMS**(*m/z*): did not ionise. **HPLC**: no absorbance at 214 nM, standard method.²¹

Ethyl 3-(2-((tert-butoxycarbonyl)amino)ethoxy)propanoate (4.53).



To a solution of ethyl acrylate (1.13 mL, 1.5 eq., 10.6 mmol) dissolved in dry THF (50 mL) was added sodium metal (96 mg, 0.57 eq., 4.0 mmol) under nitrogen. The reaction mixture was stirred for 30 min at room temperature. After this time, amine **4.52** (1.14 g, 1.00 eq., 7.07 mmol) was added and the solution stirred for a further 18h. The reaction was quenched with glacial AcOH (0.5 mL) and diluted in H₂O (50 mL). The THF was removed under vacuum and the aqueous solvent diluted in EtOAc (80 mL). The layers were separated and the aqueous further extracted with EtOAc (2 x 80 mL). The combined organic layers were washed with brine (30 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 10:90 pet spirits:EtOAc to 30:70 pet spirits:EtOAc to give **4.53** as a colourless oil (349 mg, 19 %). **¹H NMR** (401 MHz, CDCl₃) δ 4.94 (bs, 1H), 4.14 (q, *J* = 7.1 Hz, 2H), 3.69 (t, *J* = 6.2 Hz, 2H), 3.52 – 3.45 (m, 2H), 3.27 (dd, *J* = 10.2, 5.1 Hz, 2H), 2.54 (t, *J* = 6.2 Hz, 2H), 1.42 (bs, 9H), 1.25 (t, *J* = 7.1 Hz, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 171.69, 156.06, 79.29, 70.02, 66.29, 60.68, 40.31, 35.17, 28.49, 14.31. **LCMS**(*m/z*): 283.9 [M+Na]⁺, *t*_R = 3.25 min, standard method. **HPLC**: no activity at 214.4 and 254.4 nm.

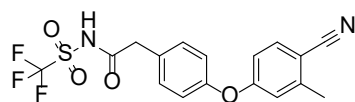
2-(4-(4-Cyano-3-methylphenoxy)phenyl)-N-(methylsulfonyl)acetamide (4.56)



Prepared according to **general procedure 9** with 2-[4-(4-cyano-3-methylphenoxy)phenyl]acetic acid (200 mg, 1.00 eq., 0.75 mmol), EDC hydrochloride (215 mg, 1.5 eq., 1.12 mmol), DMAP (91 mg, 1.0 eq., 0.75 mmol), dry DCM (8 mL), DIPEA (0.46 mL, 3.5 eq., 2.62 mmol) and methane sulfonamide (107 mg, 1.5 eq., 1.12 mmol). The residue was purified by flash column chromatography using a gradient of 60:40 pet spirits:EtOAc to 45:55 pet spirits:EtOAc to give **4.56** as a white solid (135 mg, 52 %). **¹H NMR** (401 MHz, CD₃OD) δ 7.60 (d, *J* = 8.6 Hz, 1H), 7.38 – 7.32 (m, 2H), 7.07 – 7.01 (m, 2H), 6.95 (d, *J* = 2.3 Hz, 1H), 6.86 (dd, *J* = 8.6, 2.3 Hz, 1H), 3.66 (s, 2H), 3.24 (s, 3H), 2.46 (s, 3H). **¹³C NMR** (101 MHz, CD₃OD) δ

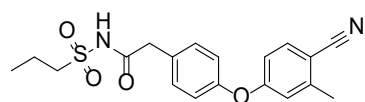
172.60, 162.96, 155.67, 145.79, 135.60, 132.40, 131.74, 121.48, 119.97, 118.92, 116.55, 107.34, 42.75, 41.32, 20.51. **LCMS** (m/z): 344.9 $[M+H]^+$, t_R = 3.29 min, standard method. **HRMS** (m/z): $C_{17}H_{16}N_2O_4S$ requires 345.0904 $[M+H]^+$; found 345.0911. **HPLC**: t_R = 5.93 min, >99 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N*-((trifluoromethyl)sulfonyl)acetamide (4.57)



Prepared according to **general procedure 9** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.00 eq., 0.75 mmol), EDC hydrochloride (215 mg, 1.5 eq., 1.12 mmol), DMAP (91 mg, 1.0 eq., 0.75 mmol), dry DCM (8 mL), DIPEA (0.46 mL, 3.5 eq., 2.62 mmol) and trifluoromethane sulfonamide (167 mg, 1.5 eq., 1.12 mmol). The residue was purified by flash column chromatography using a gradient of 60:40 pet spirits:EtOAc to 30:70 pet spirits:EtOAc to give **4.57** as a white foam (162 mg, 54 %). **¹H NMR** (401 MHz, CD₃OD) δ 7.59 (d, J = 8.6 Hz, 1H), 7.38 – 7.33 (m, 2H), 7.02 – 6.97 (m, 2H), 6.92 (d, J = 2.5 Hz, 1H), 6.86 (dd, J = 8.6, 2.5 Hz, 1H), 3.56 (s, 2H), 2.46 (s, 3H). **¹³C NMR** (101 MHz, CD₃OD) δ 180.54, 163.35, 154.85, 145.68, 135.51, 134.72, 132.29, 121.90 (q, $^1J_{C-F}$ = 323.2 Hz), 121.25, 119.61, 118.98, 116.30, 106.94, 46.58, 20.50. **LCMS** (m/z): 396.8 $[M-H]^-$, t_R = 3.88 min, standard method. **HRMS** (m/z): $C_{17}H_{13}F_3N_2O_4S$ requires 399.0621 $[M+H]^+$; found 399.0622. **HPLC**: t_R = 6.34 min, >99 %, standard method.

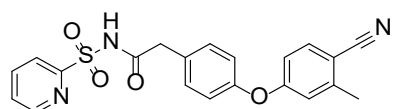
2-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N*-(propylsulfonyl)acetamide (4.58)



Prepared according to **general procedure 9** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (300 mg, 1.00 eq., 1.12 mmol), EDC hydrochloride (323 mg, 1.5 eq., 1.68 mmol), DMAP (137 mg, 1.0 eq., 1.12 mmol), dry MeCN (5 mL), DIPEA (0.6 mL, 3.07 eq., 3.44 mmol) and propane-1-sulfonamide (276 mg, 2.0 eq., 2.24 mmol). The residue was purified by flash column chromatography using 50:50 pet spirits:EtOAc. The residue was further purified by flash column chromatography using 25:75 acetone:toluene to give **4.58** as a white solid (56 mg, 13 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.47 (d, J = 8.5 Hz, 1H), 7.26 – 7.21 (m, 2H), 7.00 – 6.94 (m, 2H), 6.81 (d, J = 2.3 Hz, 1H), 6.76 (dd, J = 8.5, 2.4 Hz, 1H),

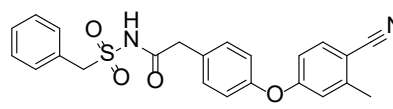
3.61 (s, 2H), 3.39 – 3.30 (m, 2H), 2.43 (s, 3H), 1.81 – 1.70 (m, 2H), 0.98 (t, $J = 7.5$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 169.96, 160.87, 154.95, 144.63, 134.44, 131.25, 128.87, 120.68, 119.25, 118.10, 115.70, 106.89, 55.05, 42.84, 20.64, 16.92, 12.75. LCMS (m/z): 373.0 $[\text{M}+\text{H}]^+$, $t_R = 4.15$ min, standard method. HRMS (m/z): $\text{C}_{19}\text{H}_{20}\text{N}_2\text{O}_4\text{S}$ requires 373.1217 $[\text{M}+\text{H}]^+$; found 373.1212. HPLC: $t_R = 6.20$ min, >99 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N*-(pyridin-2-ylsulfonyl)acetamide (4.59)



Prepared according to **general procedure 9** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (300 mg, 1.00 eq., 1.12 mmol), EDC hydrochloride (323 mg, 1.5 eq., 1.68 mmol), DMAP (137 mg, 1.0 eq., 1.12 mmol), dry MeCN (5 mL), DIPEA (0.6 mL, 3.07 eq., 3.44 mmol) and pyridine-2-sulfonamide (355 mg, 2.0 eq., 2.24 mmol). The residue was purified by flash column chromatography using 25:75 acetone:toluene to give **4.59** as a white solid (100 mg, 21 %). ^1H NMR (401 MHz, d_6 -DMSO) δ 8.74 (ddd, $J = 4.7, 1.6, 0.9$ Hz, 1H), 8.13 (td, $J = 7.7, 1.7$ Hz, 1H), 8.08 (ddd, $J = 8.9, 6.6, 1.4$ Hz, 1H), 7.76 (d, $J = 8.6$ Hz, 1H), 7.72 (ddd, $J = 7.4, 4.7, 1.3$ Hz, 1H), 7.29 – 7.23 (m, 2H), 7.08 – 7.03 (m, 2H), 7.02 (d, $J = 2.4$ Hz, 1H), 6.86 (dd, $J = 8.6, 2.4$ Hz, 1H), 3.64 (s, 2H), 2.44 (s, 3H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 170.05, 160.82, 155.68, 153.48, 150.10, 144.36, 138.73, 134.74, 131.28, 130.55, 127.93, 123.26, 119.99, 118.82, 117.91, 115.41, 105.75, 41.16, 19.99. LCMS (m/z): 407.9 $[\text{M}+\text{H}]^+$, $t_R = 3.85$ min, standard method. HRMS (m/z): $\text{C}_{21}\text{H}_{17}\text{N}_3\text{O}_4\text{S}$ requires 408.1013 $[\text{M}+\text{H}]^+$; found 408.1026. HPLC: $t_R = 6.03$ min, >99 %, standard method.

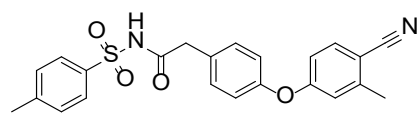
N-(Benzylsulfonyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamide (4.60)



Prepared according to **general procedure 9** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (293 mg, 1.00 eq., 1.10 mmol), EDC hydrochloride (315 mg, 1.5 eq., 1.64 mmol), DMAP (134 mg, 1.0 eq., 1.10 mmol), dry MeCN (5 mL), DIPEA (0.60 mL, 3.1 eq., 3.44 mmol) and α -toluenesulfonamide (375 mg, 2.0 eq., 2.20 mmol). The residue was purified by flash column chromatography using a gradient of 33:67 pet spirits:EtOAc to 40:60 pet spirits:EtOAc. The residue was further triturated with cold Et_2O (4 mL) to

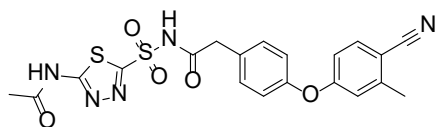
give **4.60** as a white solid (58 mg, 12 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.55 (d, *J* = 8.5 Hz, 1H), 7.43 (bs, *J* = 6.0 Hz, 1H), 7.42 – 7.28 (m, 5H), 7.19 – 7.11 (m, 2H), 7.03 – 6.97 (m, 2H), 6.88 (d, *J* = 2.4 Hz, 1H), 6.83 (dd, *J* = 8.5, 2.2 Hz, 1H), 4.65 (s, 2H), 3.59 (s, 2H), 2.51 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 169.91, 160.93, 155.10, 144.75, 134.56, 131.36, 130.75, 129.46, 129.12, 128.61, 127.90, 120.76, 119.38, 118.18, 115.81, 107.04, 58.93, 42.80, 20.75. **LCMS** (*m/z*): 418.9 [M-H]⁻, *t_R* = 4.20 min, standard method. **HRMS** (*m/z*): C₂₃H₂₀N₂O₄S requires 421.1217 [M+H]⁺; found 421.1214. **HPLC**: *t_R* = 6.64 min, >99 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N*-tosylacetamide (**4.61**)



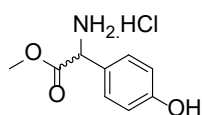
To a solution of 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (300 mg, 1.00 eq., 1.12 mmol), EDC hydrochloride (323 mg, 1.5 eq., 1.68 mmol) and DMAP (137 mg, 1.0 eq., 1.12 mmol) dissolved in dry MeCN (5 mL) under N₂ was added DIPEA (0.59 mL, 3.0 eq., 3.37 mmol). The reaction mixture was stirred for 15 min at room temperature. After this time, p-toluenesulfonamide (384 mg, 2.0 eq., 2.24 mmol) was added and the reaction stirred for a further 24 h at room temperature. The reaction was then diluted in DCM (30 mL) and hydrochloride (aq., 1 M, 20 mL). The layers were separated and the aqueous further extracted with DCM (2 × 30 mL). The combined organic layers were washed with acidic brine (aq., pH 4, 20 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using 50:50 pet spirits:EtOAc. The residue was further purified by trituration with cold C hydrochloride₃ to give **4.61** as a white solid (55 mg, 11 %). **¹H NMR** (401 MHz, DMSO-*d*₆) δ 12.29 (s, 1H), 7.80 (d, *J* = 8.3 Hz, 2H), 7.75 (d, *J* = 8.6 Hz, 1H), 7.41 (d, *J* = 8.0 Hz, 2H), 7.28 – 7.20 (m, 2H), 7.08 – 6.99 (m, 3H), 6.86 (dd, *J* = 8.6, 2.5 Hz, 1H), 3.58 (s, 2H), 2.44 (s, 3H), 2.39 (s, 3H). **¹³C NMR** (101 MHz, DMSO) δ 169.32, 160.82, 153.50, 144.38, 144.31, 136.36, 134.76, 131.32, 130.58, 129.55, 127.59, 119.97, 118.86, 117.93, 115.45, 105.77, 41.17, 21.08, 20.00. **LCMS** (*m/z*): 420.8 [M+H]⁺, *t_R* = 3.60 min, standard method. **HRMS** (*m/z*): C₂₃H₂₀N₂O₄S requires 421.1217 [M+H]⁺; found 421.1216. **HPLC**: *t_R* = 6.68 min, >99 %, standard method.

***N*-((5-Acetamido-1,3,4-thiadiazol-2-yl)sulfonyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamide (4.62)**

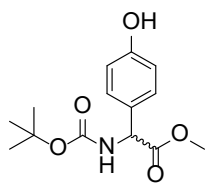


Prepared according to **general procedure 9** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (60 mg, 1.00 eq., 0.22 mmol), EDC hydrochloride (65 mg, 1.5 eq., 0.34 mmol), DMAP (28 mg, 1.0 eq., 0.22 mmol), dry MeCN (2 mL), DIPEA (0.12 mL, 3.0 eq., 0.67 mmol) and acetazolamide (100 mg, 2.0 eq., 0.45 mmol). The residue was purified by flash column chromatography using a gradient of 20:80:0 pet spirits:EtOAc:MeOH to 0:98:2 pet spirits:EtOAc:MeOH to give **4.62** as a white solid (8 mg, 7.6 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 7.73 (d, *J* = 8.6 Hz, 1H), 7.32 – 7.25 (m, 2H), 7.05 – 6.99 (m, 3H), 6.88 (dd, *J* = 8.6, 2.4 Hz, 1H), 3.45 (s, 2H), 2.43 (s, 3H), 2.20 (s, 3H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 175.01, 169.55, 161.54, 161.50, 153.31, 144.79, 135.19, 133.66, 131.74, 120.22, 119.11, 118.45, 115.79, 105.99, 44.44, 22.80, 20.45 (**Quaternary carbon of thiadiazole missing, could not be obtained through HSQC or HMBC**). **LCMS** (*m/z*): 469.8 [M-H]⁻, *t*_R = 4.18 min, standard method. **HRMS** (*m/z*): C₂₀H₁₇N₅O₅S₂ requires 472.0744 [M+H]⁺; found 472.0747. **HPLC**: *t*_R = 5.95 min, >99 %, standard method.

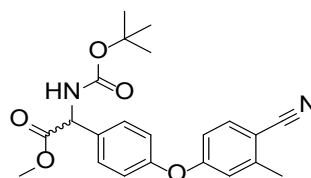
Methyl 2-amino-2-(4-hydroxyphenyl)acetate hydrochloride (4.65)²²



To a solution of DL-*p*-hydroxyphenyl glycine (3.0 g, 1.0 eq., 17.9 mmol) dissolved in MeOH (60 mL) was added SOCl₂ (1.82 mL, 1.4 eq., 25.1 mmol) dropwise at 0°C. The reaction was warmed to room temperature and stirred for a further 3 days. After this time the solvent was removed under vacuum and the solid triturated with cold MeOH to give the hydrochloride salt of **4.65** as a pale green solid (1.70 g, 43 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 9.93 (s, 1H), 8.92 (s, 3H), 7.28 (d, *J* = 8.4 Hz, 2H), 6.83 (d, *J* = 8.4 Hz, 2H), 5.10 (s, 1H), 3.70 (s, 3H). **LCMS** (*m/z*): 162.8 [M-NH₂]⁺, *t*_R = 0.797 min, standard method. **HPLC**: *t*_R = 1.82 min, >99 %, standard method.

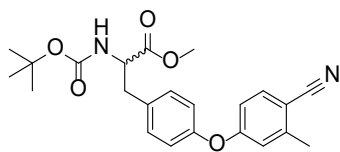
Methyl 2-((*tert*-butoxycarbonyl)amino)-2-(4-hydroxyphenyl)acetate (4.66)²³

To a solution of amine **4.65** (1.68 g, 1.00 eq., 7.71 mmol) and di-*tert*-butyl dicarbonate (1.77 g, 1.05 eq., 8.10 mmol) dissolved in dry DCM (30 mL) was added DIPEA (4.70 mL, 3.5 eq., 27.0 mmol) dropwise under N₂. The reaction mixture was then stirred for 18 h. After this time, the reaction was diluted in DCM (50 mL) and H₂O (50 mL). The layers were separated and the organic washed with H₂O (2 × 100 mL), brine (3 × 50 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 80:20 pet spirits:EtOAc to 67:33 pet spirits:EtOAc to give **4.66** as a white solid (1.18 g, 54 %). ¹H NMR (401 MHz, CD₃OD) δ 7.20 – 7.15 (m, 2H), 6.78 – 6.73 (m, 2H), 5.09 (s, 1H), 3.68 (s, 3H), 1.44 (s, 9H). LCMS (*m/z*): did not ionise, *t*_R = 3.23 min, standard method. HRMS (*m/z*): C₁₄H₁₉NO₅ requires 280.1190 [M-H]⁻; found 280.1195. HPLC: *t*_R = 5.26 min, >99 %, standard method.

Methyl 2-((*tert*-butoxycarbonyl)amino)-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetate (4.67)

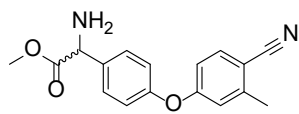
Prepared according to **general procedure 6** with 4-bromo-2-methylbenzonitrile (667 mg, 1.00 eq., 3.40 mmol), phenol **4.66** (1.05 g, 1.1 eq., 3.74 mmol), Cs₂CO₃ (2.22 g, 2.00 eq., 6.80 mmol), *N,N*-dimethylglycine hydrochloride (190 mg, 0.40 eq., 1.36 mmol) and copper(I) iodide (130 mg, 0.20 eq., 0.68 mmol) and degassed DMF (15 mL) at 80 °C for 24 h. The residue was purified by flash column chromatography using a gradient of 90:10 pet spirits:EtOAc to 80:20 pet spirits:EtOAc to give **4.67** as a colourless oil (609 mg, 45 %). ¹H NMR (401 MHz, CDCl₃) δ 7.54 (d, *J* = 8.5 Hz, 1H), 7.42 – 7.37 (m, 2H), 7.05 – 6.99 (m, 2H), 6.88 (d, *J* = 1.8 Hz, 1H), 6.83 (dd, *J* = 8.5, 2.3 Hz, 1H), 5.60 (d, *J* = 5.7 Hz, 1H), 5.34 (d, *J* = 7.0 Hz, 1H), 3.75 (s, 3H), 2.50 (s, 3H), 1.44 (bs, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 171.52, 160.99, 155.30, 154.92, 144.65, 134.48, 133.61, 129.05, 120.53, 119.32, 118.16, 115.78, 106.96, 80.50, 57.07, 52.96, 28.41, 20.72. LCMS (*m/z*): did not ionise, *t*_R = 3.52 min, standard method. HRMS (*m/z*): C₂₂H₂₄N₂O₅ requires 419.1577 [M+H]⁺; found 419.1580. HPLC: *t*_R = 7.22 min, >99 %, standard method.

Methyl 2-((*tert*-butoxycarbonyl)amino)-3-(4-(4-cyano-3-methylphenoxy)phenyl)propanoate (4.68)



Prepared according to **general procedure 6** with *N*-(*tert*-butoxycarbonyl)-*L*-tyrosine methyl ester (3.90 g, 1.50 eq., 13.2 mmol), 4-bromo-2-methylbenzonitrile (1.73 g, 1.00 eq., 8.82 mmol), Cs₂CO₃ (7.18 g, 2.50 eq., 22.1 mmol), *N,N*-dimethylglycine hydrochloride (493 mg, 0.40 eq., 3.53 mmol), copper(I) iodide (336 mg, 0.20 eq., 1.76 mmol) and dry DMF (17 mL) at 80 °C for 18 h. The residue was purified by flash column chromatography using a gradient of 90:10 pet spirits:EtOAc to 70:30 pet spirits:EtOAc to give **4.68** as a white solid (1.50 g, 42 %). ¹H NMR (401 MHz, CDCl₃) δ 7.52 (d, *J* = 8.5 Hz, 1H), 7.18 – 7.13 (m, 2H), 6.99 – 6.94 (m, 2H), 6.84 (d, *J* = 2.3 Hz, 1H), 6.79 (dd, *J* = 8.5, 2.3 Hz, 1H), 5.04 (d, *J* = 7.8 Hz, 1H), 4.59 (dd, *J* = 13.2, 6.2 Hz, 1H), 3.73 (s, 3H), 3.14 (dd, *J* = 13.8, 5.4 Hz, 1H), 3.01 (dd, *J* = 13.8, 6.4 Hz, 1H), 2.49 (s, 3H), 1.41 (s, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 172.32, 161.44, 155.15, 154.13, 144.56, 134.44, 133.01, 131.14, 120.45, 119.04, 118.27, 115.45, 106.61, 80.16, 54.57, 52.43, 38.02, 28.41, 20.73. LCMS (*m/z*): 408.9 [M-H]⁻, *t*_R = 3.59 min, standard method. HRMS (*m/z*): C₂₃H₂₆N₂O₅ requires 433.1734 [M+Na]⁺; found 433.1744. HPLC: *t*_R = 7.26 min, >99 %, standard method.

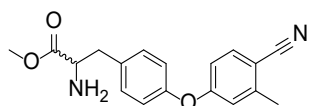
Methyl 2-amino-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetate (4.69)



N-Boc amine **4.67** (589 mg, 1.00 eq., 1.49 mmol) was dissolved in 3:1 DCM:TFA (8 mL) mixture and stirred overnight at room temperature. After this time, the reaction mixture was diluted in DCM (30 mL) and basified with NaHCO₃ (aq., sat., 20 mL, caution vigorous bubbling). The layers were then separated and the aqueous extracted with DCM (2 × 40 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered and evaporated to dryness to give **4.69** as a yellow oil (267 mg, 61 %). ¹H NMR (401 MHz, CDCl₃) δ 7.53 (d, *J* = 8.5 Hz, 1H), 7.45 – 7.40 (m, 2H), 7.06 – 7.01 (m, 2H), 6.87 (d, *J* = 2.4 Hz, 1H), 6.82 (dd, *J* = 8.7, 2.3 Hz, 1H), 4.67 (bs, 1H), 3.74 (s, 3H), 2.50 (s, 3H), 2.21 (bs, 2H). ¹³C NMR (101 MHz, CDCl₃) δ 174.30, 161.15, 154.97, 144.61, 136.57, 134.45, 128.82, 120.56, 119.15, 118.19, 115.63, 106.81,

58.05, 52.66, 20.70. **LCMS** (m/z): 279.9 $[M-NH_2]^+$, t_R = 3.05 min, standard method. **HRMS** (m/z): $C_{17}H_{16}N_2O_3$ requires 297.1234 $[M+H]^+$; found 297.1227. **HPLC**: t_R = 4.98 min, >99 %, standard method.

Methyl 2-amino-3-(4-(4-cyano-3-methylphenoxy)phenyl)propanoate (4.70)

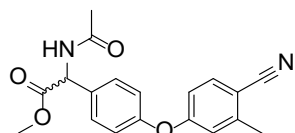


N-Boc amine **4.68** (1.35 g, 1.00 eq., 3.28 mmol) was dissolved in 3:1 DCM:TFA (20 mL) mixture and stirred overnight at room temperature.

After this time, the reaction mixture was diluted in DCM (50 mL) and basified with $NaHCO_3$ (aq., sat., 50 mL, caution vigorous bubbling). The layers were then separated and the aqueous extracted with DCM (2×40 mL). The combined organic layers were dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 50:50 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **4.70** as a colourless oil (907 mg, 89 %).

1H NMR (401 MHz, d_6 -DMSO) δ 7.75 (d, J = 8.6 Hz, 1H), 7.30 – 7.25 (m, 2H), 7.07 – 7.02 (m, 2H), 7.01 (d, J = 2.4 Hz, 1H), 6.89 – 6.84 (m, 1H), 3.60 (s, 3H), 3.60 – 3.56 (m, 1H), 2.90 (dd, J = 13.4, 6.1 Hz, 1H), 2.79 (dd, J = 13.4, 7.5 Hz, 1H), 2.44 (s, 3H), 1.81 (bs, 2H). **^{13}C NMR** (101 MHz, d_6 -DMSO) δ 175.41, 161.09, 152.89, 144.30, 134.91, 134.74, 131.12, 119.91, 118.63, 117.94, 115.18, 105.53, 55.71, 51.36, 20.00. **LCMS** (m/z): 310.9 $[M+H]^+$, t_R = 3.09 min, standard method. **HRMS** (m/z): $C_{18}H_{18}N_2O_3$ requires 311.1390 $[M+H]^+$; found 311.1389. **HPLC**: t_R = 5.20 min, >99 %, standard method.

Methyl 2-acetamido-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetate (4.71)

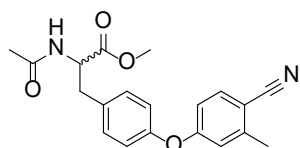


To a solution of amine **4.69** (110 mg, 1.00 eq., 0.371 mmol) and DIPEA (0.19 mL, 3.00 eq., 1.11 mmol) dissolved in dry DCM (3 mL) was added acetyl chloride (29 μ L, 1.1 eq., 0.41 mmol) dropwise at 0 °C under N_2 . The

reaction mixture was then stirred for 6 days. After this time, the reaction was diluted in DCM (20 mL) and hydrochloride (aq., pH 3, 15 mL). The layers were separated and the aqueous further extracted with DCM (2×30 mL). The combined organic layers were washed with brine (20 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash

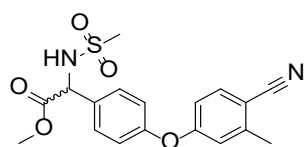
column chromatography using a gradient of 50:50 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **4.71** as a white solid (58 mg, 46 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.53 (d, *J* = 8.5 Hz, 1H), 7.42 – 7.35 (m, 2H), 7.04 – 6.98 (m, 2H), 6.87 (d, *J* = 2.3 Hz, 1H), 6.81 (dd, *J* = 8.5, 2.2 Hz, 1H), 6.61 (d, *J* = 6.9 Hz, 1H), 5.60 (d, *J* = 7.1 Hz, 1H), 3.75 (s, 3H), 2.49 (s, 3H), 2.05 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 171.43, 169.52, 160.91, 155.38, 144.68, 134.49, 133.22, 129.20, 120.55, 119.37, 118.16, 115.80, 107.02, 55.86, 53.09, 23.21, 20.72. **LCMS** (*m/z*): 339.8 [M+H]⁺, *t_R* = 3.20 min, standard method. **HRMS** (*m/z*): C₁₉H₁₈N₂O₄ requires 339.1339 [M+H]⁺; found 339.1340. **HPLC**: *t_R* = 5.93 min, >99 %, standard method.

Methyl 2-acetamido-3-(4-(4-cyano-3-methylphenoxy)phenyl)propanoate (**4.72**)



To a solution of amine **4.70** (110 mg, 1.00 eq., 0.354 mmol) and DIPEA (0.12 mL, 2.00 eq., 0.71 mmol) dissolved in dry DCM (4 mL) was added acetyl chloride (38 μL, 1.50 eq., 0.53 mmol) dropwise at 0 °C under N₂. The reaction mixture was then stirred for 18 h. After this time, the reaction was diluted in DCM (15 mL) and hydrochloride (aq., pH 3, 15 mL). The layers were separated and the aqueous further extracted with DCM (2 × 15 mL). The combined organic layers were washed with brine (20 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness to give **4.72** as a white solid (117 mg, 95 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.53 (d, *J* = 8.5 Hz, 1H), 7.15 – 7.09 (m, 2H), 6.99 – 6.93 (m, 2H), 6.85 (d, *J* = 2.4 Hz, 1H), 6.79 (dd, *J* = 8.5, 2.4 Hz, 1H), 5.96 (d, *J* = 7.6 Hz, 1H), 4.89 (dt, *J* = 7.7, 5.8 Hz, 1H), 3.74 (s, 3H), 3.18 (dd, *J* = 13.9, 5.9 Hz, 1H), 3.08 (dd, *J* = 13.9, 5.7 Hz, 1H), 2.49 (s, 3H), 2.01 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 172.11, 169.70, 161.33, 154.25, 144.58, 134.45, 132.76, 131.07, 120.41, 119.11, 118.24, 115.49, 106.67, 53.31, 52.53, 37.43, 23.28, 20.73. **LCMS** (*m/z*): 352.9 [M+H]⁺, *t_R* = 3.30 min, standard method. **HRMS** (*m/z*): C₂₀H₂₀N₂O₄ requires 353.1496 [M+H]⁺; found 353.1496. **HPLC**: *t_R* = 6.13 min, >99 %, standard method.

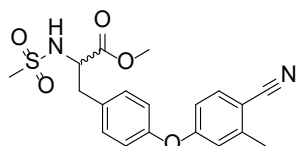
Methyl 2-(4-(4-cyano-3-methylphenoxy)phenyl)-2-(methylsulfonamido)acetate (**4.73**)



To a solution of amine **4.69** (200 mg, 1.00 eq., 0.674 mmol) and DIPEA (0.36 mL, 3.0 eq., 2.0 mmol) dissolved in dry DCM (5 mL) was added

methanesulfonyl chloride (55 μ L, 1.05 eq., 0.71 mmol) dropwise at 0 °C under N₂. The reaction mixture was then stirred for 18 h. After this time, additional methane sulfonyl chloride (38 μ L, 0.7 eq., 0.50 mmol) was added and the reaction stirred for a further 4 h. The reaction mixture was then diluted in H₂O (15 mL) and DCM (20 mL). The layers were separated and the organic layer was washed with H₂O (2 \times 15 mL), brine (10 mL) and dried over Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using 66:34 pet spirits:EtOAc to give **4.73** as a colourless oil (154 mg, 61 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.54 (d, J = 8.5 Hz, 1H), 7.44 – 7.38 (m, 2H), 7.08 – 7.01 (m, 2H), 6.88 (d, J = 2.4 Hz, 1H), 6.82 (dd, J = 8.4, 2.3 Hz, 1H), 5.76 (d, J = 6.9 Hz, 1H), 5.24 (d, J = 7.0 Hz, 1H), 3.77 (s, 3H), 2.83 (s, 3H), 2.49 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 170.78, 160.57, 155.87, 144.72, 134.52, 132.15, 129.36, 120.53, 119.58, 118.05, 115.93, 107.26, 58.97, 53.43, 42.17, 20.70. **LCMS** (m/z): 372.8 [M-H]⁻, t_R = 3.27 min, standard method. **HRMS** (m/z): C₁₈H₁₈N₂O₅S requires 397.0829 [M+H]⁺; found 397.0831. **HPLC**: t_R = 6.08 min, >99 %, standard method.

Methyl 3-(4-(4-cyano-3-methylphenoxy)phenyl)-2-(methylsulfonyl)propanoate (**4.74**)

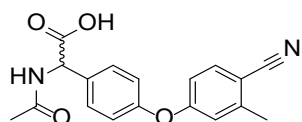


To a solution of amine **4.70** (110 mg, 1.00 eq., 0.354 mmol) and DIPEA (0.12 mL, 2.00 eq., 0.71 mmol) dissolved in dry DCM (4 mL) was added methanesulfonyl chloride (41 μ L, 1.50 eq., 0.53 mmol) dropwise at 0 °C

under N₂. The reaction mixture was then stirred for 18 h. After this time, the reaction was diluted in DCM (15 mL) and hydrochloride (aq., pH 3, 15 mL). The layers were separated and the aqueous further extracted with DCM (2 \times 15 mL). The combined organic layers were washed with brine (20 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 80:20 pet spirits:EtOAc to 50:50 pet spirits:EtOAc to give **4.74** as a white solid (109 mg, 79 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.54 (d, J = 8.5 Hz, 1H), 7.24 – 7.19 (m, 2H), 7.03 – 6.98 (m, 2H), 6.85 (d, J = 2.4 Hz, 1H), 6.80 (dd, J = 8.5, 2.4 Hz, 1H), 4.82 (d, J = 9.0 Hz, 1H), 4.42 (ddd, J = 9.1, 6.8, 5.3 Hz, 1H), 3.81 (s, 3H), 3.18 (dd, J = 13.9, 5.2 Hz, 1H), 3.07 (dd, J = 13.9, 6.8 Hz, 1H), 2.82 (s, 3H), 2.50 (s, 3H). **¹³C NMR** (101 MHz,

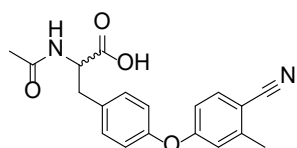
CDCl_3) δ 171.84, 161.23, 154.60, 144.66, 134.51, 131.98, 131.39, 120.63, 119.16, 118.24, 115.58, 106.83, 57.10, 53.00, 41.50, 38.95, 20.76. **LCMS** (m/z): 389.9 $[\text{M}-\text{H}]^-$, t_R = 3.39 min, standard method. **HRMS** (m/z): $\text{C}_{19}\text{H}_{20}\text{N}_2\text{O}_5\text{S}$ requires 389.1166 $[\text{M}+\text{H}]^+$; found 389.1147. **HPLC**: t_R = 6.35 min, >99 %, standard method.

2-Acetamido-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetic acid (**4.75**)

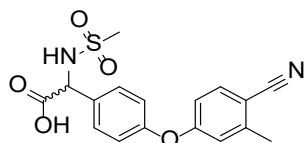


Prepared according to **general procedure 7** with ester **4.71** (58 mg, 1.0 eq., 0.17 mmol), MeOH (2 mL) NaOH (aq., 2 M, 1 mL, 11.7 eq., 2 mmol). The product **4.75** was isolated as a white solid (40 mg, 72 %) after workup. **^1H NMR** (401 MHz, CD_3OD) δ 7.62 (d, J = 8.6 Hz, 1H), 7.52 – 7.46 (m, 2H), 7.13 – 7.05 (m, 2H), 6.97 (d, J = 2.4 Hz, 1H), 6.89 (dd, J = 8.6, 2.2 Hz, 1H), 5.47 (s, 1H), 2.47 (s, 3H), 2.02 (s, 3H). **^{13}C NMR** (101 MHz, CD_3OD) δ 173.49, 172.86, 162.72, 156.57, 145.87, 135.65, 134.91, 130.80, 121.49, 120.20, 118.86, 116.78, 107.63, 57.58, 22.29, 20.51. **LCMS** (m/z): 324.9 $[\text{M}+\text{H}]^+$, t_R = 3.08 min, standard method. **HRMS** (m/z): $\text{C}_{18}\text{H}_{16}\text{N}_2\text{O}_4$ requires 325.1183 $[\text{M}+\text{H}]^+$; found 325.1186. **HPLC**: t_R = 5.35 min, >99 %, standard method.

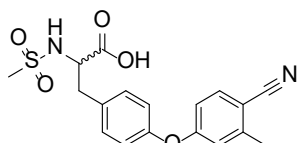
2-Acetamido-3-(4-(4-cyano-3-methylphenoxy)phenyl)propanoic acid (**4.76**)



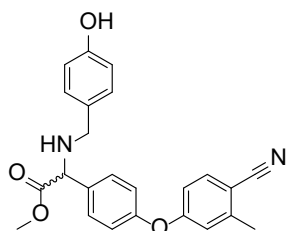
Prepared according to **general procedure 7** with ester **4.72** (115 mg, 1.00 eq., 0.326 mmol), MeOH (2 mL) and 2 M NaOH (aq., 1 mL, 6.1 eq., 2 mmol). The combined organic layers were dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness to give **4.76** as a white solid (106 mg, 96 %). **^1H NMR** (401 MHz, CDCl_3) δ 8.36 (bs, 1H), 7.52 (d, J = 8.5 Hz, 1H), 7.19 (d, J = 8.5 Hz, 2H), 6.97 (d, J = 8.5 Hz, 2H), 6.85 (d, J = 2.2 Hz, 1H), 6.78 (dd, J = 8.5, 2.3 Hz, 1H), 6.18 (d, J = 7.5 Hz, 1H), 4.87 (dd, J = 13.3, 6.0 Hz, 1H), 3.25 (dd, J = 14.1, 5.6 Hz, 1H), 3.11 (dd, J = 14.1, 6.2 Hz, 1H), 2.48 (s, 3H), 2.03 (s, 3H). **^{13}C NMR** (101 MHz, CDCl_3) δ 173.94, 171.58, 161.32, 154.23, 144.61, 134.46, 132.62, 131.16, 120.44, 119.11, 118.19, 115.41, 106.51, 53.60, 36.79, 22.91, 20.68. **LCMS** (m/z): 338.9 $[\text{M}+\text{H}]^+$, t_R = 3.22 min, standard method. **HRMS** (m/z): $\text{C}_{19}\text{H}_{18}\text{N}_2\text{O}_4$ requires 339.1339 $[\text{M}+\text{H}]^+$; found 339.1339. **HPLC**: t_R = 5.55 min, >99 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-2-(methylsulfonamido)acetic acid (4.77)

Prepared according to **general procedure 7** with ester **4.73** (150 mg, 1.0 eq., 0.40 mmol), MeOH (2 mL) and NaOH (aq., 2 M, 1 mL, 5.0 eq., 2 mmol). The combined organic layers were dried over anhydrous Na₂SO₄, filtered and evaporated to dryness to give **4.77** as a white solid (129 mg, 89 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.56 (d, *J* = 8.5 Hz, 1H), 7.49 – 7.41 (m, 2H), 7.11 – 7.03 (m, 2H), 6.90 (d, *J* = 2.3 Hz, 1H), 6.83 (dd, *J* = 8.5, 2.3 Hz, 1H), 5.72 (d, *J* = 6.7 Hz, 1H), 5.29 (d, *J* = 6.6 Hz, 1H), 2.86 (s, 3H), 2.51 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 173.60, 160.57, 156.19, 144.88, 134.63, 131.57, 129.48, 120.68, 119.72, 118.04, 116.04, 107.38, 58.83, 42.39, 20.78. **LCMS** (*m/z*): did not ionise, *t_R* = 3.31 min, standard method. **HRMS** (*m/z*): C₁₇H₁₆N₂O₅S requires 361.0853 [M+H]⁺; found 361.0855. **HPLC**: *t_R* = 5.49 min, >99 %, standard method.

3-(4-(4-Cyano-3-methylphenoxy)phenyl)-2-(methylsulfonamido)propanoic acid (4.78)

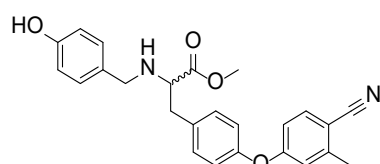
Prepared according to **general procedure 7** with ester **4.74** (100 mg, 1.00 eq., 0.257 mmol), MeOH (2 mL) and 2 M NaOH (aq., 1 mL, 7.8 eq., 2 mmol). The dried organic layers were evaporated to dryness to give **4.78** as a colourless oil (85 mg, 88 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.54 (d, *J* = 8.5 Hz, 1H), 7.31 – 7.22 (m, 2H), 7.02 (d, *J* = 8.6 Hz, 2H), 6.85 (d, *J* = 2.3 Hz, 1H), 6.80 (dd, *J* = 8.5, 2.4 Hz, 1H), 4.90 (d, *J* = 9.1 Hz, 1H), 4.46 (dd, *J* = 13.1, 8.2 Hz, 1H), 3.26 (dd, *J* = 14.0, 5.0 Hz, 1H), 3.07 (dd, *J* = 14.0, 7.4 Hz, 1H), 2.81 (s, 3H), 2.50 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 174.14, 161.05, 154.61, 144.59, 134.41, 131.66, 131.34, 120.59, 119.89, 119.07, 115.47, 106.76, 56.95, 41.54, 38.46, 20.65. **LCMS** (*m/z*): 372.9 [M-H]⁻, *t_R* = 3.36 min, standard method. **HRMS** (*m/z*): C₁₈H₁₈N₂O₅S requires 375.1009 [M+H]⁺; found 375.1010. **HPLC**: *t_R* = 5.80 min, >99 %, standard method.

Methyl 2-(4-(4-cyano-3-methylphenoxy)phenyl)-2-((4-hydroxybenzyl)amino)acetate (4.79)

Prepared according to **general procedure 10** with amine **4.69** (266 mg, 1.1 eq., 0.90 mmol), 4-hydroxybenzaldehyde (100 mg, 1.00 eq., 0.82 mmol), anhydrous 1,2-DCE (5 mL) and Na(CH₃COO)₃BH (260 mg, 1.5 eq., 1.23

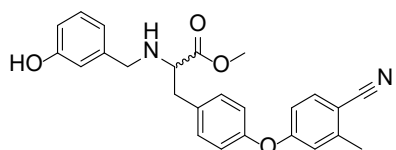
mmol). The residue purified by flash column chromatography using a gradient of 50:50 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **4.79** as a colourless oil (176 mg, 53 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.51 (d, *J* = 8.6 Hz, 1H), 7.40 (d, *J* = 8.6 Hz, 2H), 7.12 (d, *J* = 8.5 Hz, 2H), 7.03 – 6.98 (m, 2H), 6.87 (d, *J* = 2.3 Hz, 1H), 6.79 (dd, *J* = 8.6, 2.3 Hz, 1H), 6.76 – 6.68 (m, 2H), 4.77 (bs, 1H), 4.42 (s, 1H), 3.71 (s, 3H), 3.66 (s, 2H), 2.48 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 173.42, 161.08, 155.63, 155.01, 144.61, 134.44, 134.17, 130.26, 129.83, 129.46, 120.42, 119.22, 118.15, 115.61, 115.61, 106.59, 63.64, 52.55, 51.04, 20.63. **LCMS** (*m/z*): 402.9 [M+H]⁺, *t*_R = 2.96 min, standard method. **HRMS** (*m/z*): C₂₄H₂₂N₂O₄ requires 403.1652 [M+H]⁺; found 403.1656. **HPLC**: *t*_R = 5.45 min, >99 %, standard method.

Methyl 3-(4-(4-cyano-3-methylphenoxy)phenyl)-2-((4-hydroxybenzyl)amino)propanoate (4.80)



Prepared according to **general procedure 10** with **4.70** (90 mg, 1.0 eq., 0.29 mmol), 4-hydroxybenzaldehyde (53 mg, 1.50 eq., 0.44 mmol), anhydrous 1,2-DCE (3 mL) and Na(CH₃COO)₃BH (92 mg,

1.5 eq., 0.44 mmol). The reaction was diluted in DCM (20 mL) and phosphate buffer (aq., 1 M, pH 7, 20 mL). The layers were separated and the aqueous layer further extracted with DCM (2 × 20 mL). The combined organic layers were washed with brine (20 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue purified by flash column chromatography using a gradient of 80:20 pet spirits:EtOAc to 50:50 pet spirits:EtOAc to give **4.80** as a colourless oil (73 mg, 74 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.52 (d, *J* = 8.5 Hz, 1H), 7.21 – 7.16 (m, 2H), 7.11 – 7.07 (m, 2H), 6.98 – 6.93 (m, 2H), 6.84 (d, *J* = 2.4 Hz, 1H), 6.79 (dd, *J* = 8.4, 2.3 Hz, 1H), 6.75 – 6.70 (m, 2H), 3.75 (d, *J* = 12.9 Hz, 1H), 3.68 (s, 3H), 3.57 (d, *J* = 12.9 Hz, 1H), 3.53 (t, *J* = 6.8 Hz, 1H), 2.98 – 2.94 (m, 2H), 2.49 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 175.06, 161.59, 155.28, 153.82, 144.59, 134.47, 134.08, 131.06, 131.05, 129.75, 120.42, 118.92, 118.31, 115.51, 115.33, 106.42, 61.95, 51.96, 51.61, 38.98, 20.74. **LCMS** (*m/z*): 416.9 [M+H]⁺, *t*_R = 3.14 min, standard method. **HRMS** (*m/z*): C₂₅H₂₄N₂O₄ requires 417.1809 [M+H]⁺; found 417.1790. **HPLC**: *t*_R = 5.65 min, >99 %, standard method.

Methyl 3-(4-(4-cyano-3-methylphenoxy)phenyl)-2-((3-hydroxybenzyl)amino)propanoate (4.81)

Prepared according to **general procedure 10** with amine **4.70** (240

mg, 1.0 eq., 0.77 mmol), 3-hydroxybenzaldehyde (142 mg, 1.50 eq.,

1.16 mmol), anhydrous 1,2-DCE (5 mL) and Na(CH₃COO)₃BH (246

mg, 1.5 eq., 0.44 mmol). The residue purified by flash column chromatography using a gradient of

80:20 pet spirits:EtOAc to 60:40 pet spirits:EtOAc to give **4.81** as a white foam (129 mg, 40 %). **¹H**

NMR (401 MHz, CDCl₃) δ 7.52 (d, *J* = 8.5 Hz, 1H), 7.22 – 7.18 (m, 2H), 7.14 (t, *J* = 7.7 Hz, 1H),

6.99 – 6.94 (m, 2H), 6.85 (d, *J* = 2.4 Hz, 1H), 6.82 – 6.75 (m, 2H), 6.72 – 6.66 (m, 2H), 3.79 (d, *J* =

13.4 Hz, 1H), 3.69 (s, 3H), 3.60 (d, *J* = 13.4 Hz, 1H), 3.53 (t, *J* = 6.8 Hz, 1H), 3.02 – 2.91 (m, 2H),

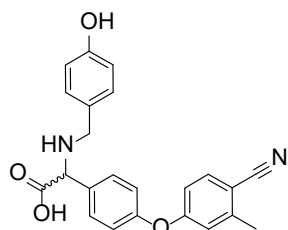
2.49 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 175.06, 161.55, 156.15, 153.84, 144.59, 141.15, 134.47,

134.11, 131.10, 129.71, 120.39, 120.38, 118.96, 118.31, 115.36, 115.24, 114.42, 106.44, 62.03,

51.98, 51.88, 39.01, 20.73. **LCMS** (*m/z*): 416.9 [M+H]⁺, *t_R* = 3.13 min, standard method. **HRMS**

(*m/z*): C₂₅H₂₄N₂O₄ requires 417.1809 [M+H]⁺; found 417.1810. **HPLC**: *t_R* = 5.69 min, >99 %,

standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-2-((4-hydroxybenzyl)amino)acetic acid (4.82)

Ester **4.79** (170 mg, 1.0 eq., 0.42 mmol) was dissolved in MeOH (2 mL) and

2 M NaOH (aq., 1 mL, 4.8 eq., 2 mmol) and stirred for 20 h at room

temperature. After this time, the reaction was diluted in phosphate buffer

(aq., 1 M, pH 7, 10 mL) and filtered to give **4.82** as a white solid (246 mg,

150 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 7.74 (d, *J* = 8.6 Hz, 1H), 7.44 (d, *J* = 8.6 Hz, 2H), 7.11

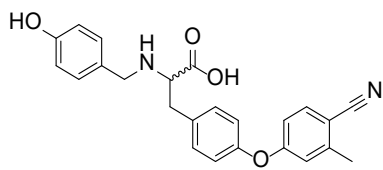
(d, *J* = 8.5 Hz, 2H), 7.04 – 6.98 (m, 3H), 6.87 (dd, *J* = 8.6, 2.3 Hz, 1H), 6.70 (d, *J* = 8.5 Hz, 2H), 3.94

(s, 1H), 3.61 (d, *J* = 12.9 Hz, 1H), 3.56 (d, *J* = 12.9 Hz, 1H), 2.44 (s, 3H). CNMR not obtained due

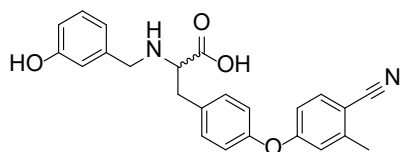
to low DMSO and H₂O solubility. **LCMS** (*m/z*): 388.9 [M+H]⁺, *t_R* = 3.08 min, standard method.

HRMS (*m/z*): C₂₃H₂₀N₂O₄ requires 389.1496 [M+H]⁺; found 389.1496. **HPLC**: *t_R* = 5.08

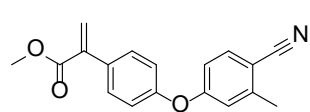
min, >99 %, standard method.

3-(4-(4-Cyano-3-methylphenoxy)phenyl)-2-((4-hydroxybenzyl)amino)propanoic acid (4.83)

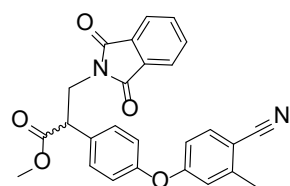
Ester **4.80** (86 mg, 1.00 eq., 0.21 mmol) was dissolved in MeOH (2 mL) and NaOH (aq., 2 M, 1 mL, 9.5 eq., 2 mmol) and stirred for 24 h at room temperature. After this time, the MeOH was removed under vacuum. The residue diluted in phosphate buffer (aq., 1 M, pH 7, 10 mL) and filtered to give **4.83** as a white solid (39 mg, 47 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 7.75 (d, *J* = 8.6 Hz, 1H), 7.31 (d, *J* = 8.5 Hz, 2H), 7.05 (d, *J* = 8.5 Hz, 2H), 7.02 (d, *J* = 8.5 Hz, 2H), 6.99 (d, *J* = 2.3 Hz, 1H), 6.86 (dd, *J* = 8.6, 2.4 Hz, 1H), 6.68 (d, *J* = 8.5 Hz, 2H), 3.70 (d, *J* = 13.0 Hz, 1H), 3.57 (d, *J* = 13.1 Hz, 1H), 3.26 (t, *J* = 6.5 Hz, 1H), 2.99 (dd, *J* = 13.8, 5.7 Hz, 1H), 2.86 (dd, *J* = 13.8, 7.3 Hz, 1H), 2.43 (s, 3H). **CNMR** not obtained, compound insoluble in H₂O and DMSO. **LCMS** (*m/z*): 402.9 [M+H]⁺, *t*_R = 3.30 min, standard method. **HRMS** (*m/z*): C₂₄H₂₂N₂O₄ requires 403.1652 [M+H]⁺; found 403.1652. **HPLC**: *t*_R = 5.35 min, >99 %, standard method.

3-(4-(4-Cyano-3-methylphenoxy)phenyl)-2-((3-hydroxybenzyl)amino)propanoic acid (4.84)

Ester **4.81** (129 mg, 1.00 eq., 0.310 mmol) was dissolved in MeOH (2 mL) and NaOH (aq., 2 M, 1 mL, 6.5 eq., 2 mmol) and stirred for 24 h at room temperature. After this time the MeOH was removed under vacuum and the residue diluted in EtOAc (15 mL) and phosphate buffer (aq., 1 M, pH 7, 10 mL). The solution was then filtered to give **4.84** as a white solid (52 mg, 42 %). **¹H NMR** (400 MHz, *d*₆-DMSO) δ 7.74 (d, *J* = 8.6 Hz, 1H), 7.31 (d, *J* = 8.5 Hz, 2H), 7.09 – 7.04 (m, 1H), 7.02 (d, *J* = 8.5 Hz, 2H), 7.00 (d, *J* = 2.3 Hz, 1H), 6.86 (dd, *J* = 8.6, 2.4 Hz, 1H), 6.69 (bs, 1H), 6.64 (t, *J* = 8.1 Hz, 2H), 3.70 (d, *J* = 13.5 Hz, 1H), 3.53 (d, *J* = 13.5 Hz, 1H), 3.30 – 3.23 (m, 1H), 2.97 (dd, *J* = 13.6, 5.8 Hz, 1H), 2.85 (dd, *J* = 13.5, 7.3 Hz, 1H), 2.43 (s, 3H). **CNMR** not obtained, compound insoluble in H₂O and DMSO. **LCMS** (*m/z*): 402.9 [M+H]⁺, *t*_R = 3.30 min, standard method. **HRMS** (*m/z*): C₂₄H₂₂N₂O₄ requires 403.1652 [M+H]⁺; found 403.1656. **HPLC**: *t*_R = 5.35 min, >99 %, standard method.

Methyl 2-(4-(4-cyano-3-methylphenoxy)phenyl)acrylate (4.88)

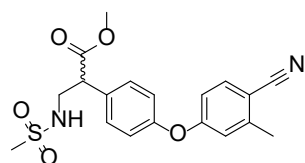
Methyl 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetate (2.03 g, 1.00 eq., 7.22 mmol), tetrabutylammonium iodide (133 mg, 0.05 eq., 0.36 mmol) and K_2CO_3 (3.00 g, 3.00 eq., 21.7 mmol) were dissolved in dry toluene (25 mL). Formaldehyde solution 38 % (aq., 2.0 mL, 10.0 eq., 72.6 mmol) was added and the reaction mixture stirred vigorously at 70 °C for 24 h. The reaction was then cooled to room temperature and diluted in EtOAc (80 mL) and hydrochloride (aq., 1 M, 30 mL). The layers were separated and the aqueous further extracted with EtOAc (2 × 50 mL). The combined organic layers were washed with brine (20 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using 9:1 pet spirits:EtOAc to give **4.88** as a white solid (750 mg, 35 %). 1H NMR (401 MHz, $CDCl_3$) δ 7.46 (d, J = 8.5 Hz, 1H), 7.41 – 7.35 (m, 2H), 6.98 – 6.92 (m, 2H), 6.82 (d, J = 2.3 Hz, 1H), 6.77 (dd, J = 8.5, 2.2 Hz, 1H), 6.30 (d, J = 1.0 Hz, 1H), 5.85 (d, J = 1.0 Hz, 1H), 3.76 (s, 3H), 2.42 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 167.12, 161.10, 155.16, 144.59, 140.25, 134.43, 133.40, 130.25, 127.03, 119.83, 119.21, 118.18, 115.73, 106.81, 52.37, 20.68. **LCMS:** (m/z): did not ionise, t_R = 3.52 min, standard method. **HRMS** (m/z): $C_{18}H_{15}NO_3$ requires 294.1125 $[M+H]^+$; found 294.1125. **HPLC:** t_R = 7.19 min, >99 %, standard method.

Methyl 2-(4-(4-cyano-3-methylphenoxy)phenyl)-3-(1,3-dioxoisindolin-2-yl)propanoate (4.90)

Alkene **4.88** (200 mg, 1.0 eq., 0.68 mmol), phthalimide (120 mg, 1.2 eq., 0.82 mmol) and DBU (42 μ L, 0.4 eq., 0.28 mmol) were dissolved in MeCN (4 mL) and stirred overnight at room temperature. After this time, the solvent was removed under vacuum and the residue purified by flash column chromatography using a gradient of 100:0 DCM:methanol to 80:20 DCM:methanol. The residue was further purified by trituration with cold methanol to give **4.90** as a white solid (47 mg, 16 %). 1H NMR (401 MHz, $CDCl_3$) δ 7.84 – 7.75 (m, 2H), 7.74 – 7.66 (m, 2H), 7.49 (d, J = 8.6 Hz, 1H), 7.38 – 7.32 (m, 2H), 6.99 – 6.91 (m, 2H), 6.80 (d, J = 2.2 Hz, 1H), 6.72 (dd, J = 8.5, 2.2 Hz, 1H), 4.32 – 4.12 (m, 3H), 3.70 (s, 3H), 2.47 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.05, 167.92, 161.23, 154.79, 144.56,

134.41, 134.23, 132.29, 131.82, 130.28, 123.44, 120.69, 119.07, 118.22, 115.46, 106.71, 52.59, 48.74, 40.57, 20.73. **LCMS** (m/z): 440.8 $[M+H]^+$, t_R = 3.824 min, standard method. **HRMS** (m/z): $C_{26}H_{20}N_2O_5$ requires 441.1445 $[M+H]^+$; found 441.1438. **HPLC**: t_R = 7.26 min, >99 %, standard method.

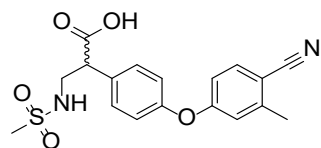
Methyl 2-(4-(4-cyano-3-methylphenoxy)phenyl)-3-(methylsulfonamido)propanoate (4.92)



Alkene **4.88** (236 mg, 1.0 eq., 0.80 mmol), methane sulfonamide (230 mg, 3.0 eq., 0.82 mmol) and DBU (360 μ L, 3.0 eq., 2.41 mmol) were dissolved in MeCN (4 mL) and stirred for 8 days at room temperature. After this time,

the solvent was removed under vacuum and the residue dissolved in DCM (20 mL) and washed with H_2O (20 mL). The layers were separated and the aqueous layer further extracted with DCM (2×30 mL). The combined organic layers were washed with brine (20 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using 50:50 pet spirits:EtOAc to give **4.92** as a colourless oil (128 mg, 41 %). **1H NMR** (400 MHz, $CDCl_3$) δ 7.54 (d, J = 8.5 Hz, 1H), 7.31 – 7.27 (m, 2H), 7.06 – 7.00 (m, 2H), 6.88 (d, J = 2.4 Hz, 1H), 6.82 (dd, J = 8.5, 2.4 Hz, 1H), 4.80 (t, J = 6.7 Hz, 1H), 3.95 (dd, J = 8.9, 5.6 Hz, 1H), 3.73 (s, 3H), 3.66 (ddd, J = 13.6, 8.9, 6.3 Hz, 1H), 3.50 – 3.41 (m, 1H), 2.93 (s, 3H), 2.50 (s, 3H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 172.86, 160.93, 155.15, 144.68, 134.50, 132.10, 129.93, 120.71, 119.36, 118.12, 115.76, 107.04, 52.65, 51.63, 45.91, 40.73, 20.70. **LCMS** (m/z): Did not ionise. **HRMS** (m/z): $C_{19}H_{20}N_2O_5S$ requires 339.1166 $[M+H]^+$; found 339.1172. **HPLC**: t_R = 6.13 min, >99 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-3-(methylsulfonamido)propanoic acid (4.93)

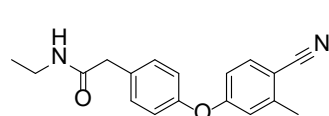


Prepared according to **general procedure 7** with ester **4.92** (110 mg, 1.00 eq., 0.28 mmol), MeOH (2 mL), 2 M NaOH (aq., 1 mL, 7.14 eq., 2 mmol).

The combined organic layers were dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness to give **4.93** as a pink solid (90 mg, 85 %). **1H NMR** (401 MHz, d_6 -DMSO) δ 7.77 (d, J = 8.6 Hz, 1H), 7.43 – 7.35 (m, 2H), 7.19 (t, J = 5.9 Hz, 1H), 7.16 – 7.08 (m, 2H), 7.06 (d,

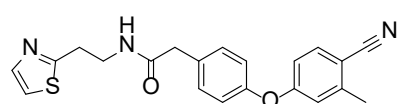
$J = 2.3$ Hz, 1H), 6.90 (dd, $J = 8.5, 2.4$ Hz, 1H), 3.82 – 3.72 (m, 1H), 3.56 (ddd, $J = 13.1, 8.4, 6.2$ Hz, 1H), 3.23 (dt, $J = 12.7, 6.2$ Hz, 1H), 2.86 (s, 3H), 2.45 (s, 3H). ^{13}C NMR (101 MHz, d_6 -DMSO) δ 173.13, 160.69, 153.97, 144.42, 134.78, 133.58, 130.04, 120.16, 119.05, 117.92, 115.58, 105.91, 51.03, 45.26, 39.34 (confirmed via HSQC), 20.01. LCMS (m/z): 746.8 $[2\text{M}-\text{H}]^-$, $t_R = 3.37$ min, standard method. HRMS (m/z): $\text{C}_{18}\text{H}_{18}\text{N}_2\text{O}_5\text{S}$ requires 397.0829 $[\text{M}+\text{Na}]^+$; found 397.0837. HPLC: $t_R = 5.51$ min, >99 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N*-ethylacetamide (6.20)



Prepared according to **general procedure 11** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.00 eq., 0.75 mmol), HATU (569 mg, 2.0 eq., 1.5 mmol), dry DMF (3 mL), Et_3N (0.20 mL, 2.0 eq., 1.5 mmol) and ethylamine (2 M in THF, 56 mL, 1.5 eq., 1.1 mmol) The residue was purified by flash column chromatography using 66:34 pet spirits:EtOAc to give **6.20** as a white solid (70 mg, 32 %). ^1H NMR (401 MHz, CDCl_3) δ 7.50 (d, $J = 8.5$ Hz, 1H), 7.29 (d, $J = 8.1$ Hz, 2H), 6.99 (d, $J = 8.1$ Hz, 2H), 6.84 (d, $J = 2.5$ Hz, 1H), 6.79 (dd, $J = 8.6, 2.5$ Hz, 1H), 5.76 (bs, 1H), 3.52 (s, 2H), 3.26 (app p, $J = 7.1$ Hz, 2H), 2.46 (s, 3H), 1.09 (t, $J = 7.2$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 170.60, 161.27, 154.14, 144.53, 134.38, 131.91, 131.12, 120.62, 118.98, 118.19, 115.47, 106.55, 42.97, 34.66, 20.65, 14.82. LCMS (m/z): 295.1 $[\text{M}+\text{H}]^+$, $t_R = 3.55$ min, standard method. HRMS (m/z): $\text{C}_{18}\text{H}_{18}\text{N}_2\text{O}_2$ requires 295.1441 $[\text{M}+\text{H}]^+$; found 295.1454. HPLC: $t_R = 5.81$ min, >99 %, standard method.

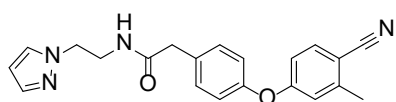
2-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N*-(2-(thiazol-2-yl)ethyl)acetamide (6.21)



Prepared according to **general procedure 11** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.00 eq., 0.75 mmol), HATU (569 mg, 2.0 eq., 1.50 mmol), dry DMF (2 mL), DBU (0.22 mL, 3.0 eq., 1.50 mmol) and 2-thiazol-2-ylethanamine (144 mg, 1.50 eq., 1.12 mmol). The residue was purified by flash column chromatography using a gradient of 20:80:0 pet spirits:EtOAc:MeOH to 0:99:1 pet spirits:EtOAc:MeOH give **6.21** as a colourless oil (14 mg, 5.0 %). ^1H NMR (401 MHz,

CDCl₃) δ 7.54 (d, J = 3.4 Hz, 1H), 7.46 (d, J = 8.5 Hz, 1H), 7.22 – 7.15 (m, 2H), 7.14 (d, J = 3.3 Hz, 1H), 6.97 – 6.89 (m, 2H), 6.80 (d, J = 2.5 Hz, 1H), 6.74 (dd, J = 8.6, 2.5 Hz, 1H), 6.38 (bs, 1H), 3.62 (q, J = 5.9 Hz, 2H), 3.48 (s, 2H), 3.12 (t, J = 6.1 Hz, 2H), 2.42 (s, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 170.83, 167.65 (Confirmed by HMBC), 161.42, 154.24, 144.64, 142.48, 134.48, 131.70, 131.34, 120.79, 119.00, 118.84, 118.27, 115.46, 106.70, 43.18, 38.68, 32.38, 20.76. LCMS (m/z): 377.9 [M+H]⁺, t_R = 3.55 min, standard method. HRMS (m/z): C₂₁H₁₉N₃O₂S requires 378.1271 [M+H]⁺; found 378.1272. HPLC: t_R = 5.60 min, >99 %, standard method.

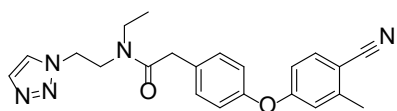
***N*-(2-(1H-pyrazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamide (6.22)**



Prepared according to **general procedure 11** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.00 eq., 0.75 mmol),

HATU (569 mg, 2.0 eq., 1.50 mmol), dry DMF (2 mL), DBU (0.22 mL, 2.0 eq., 1.50 mmol) and 2-pyrazol-1-ylethanamine (125 mg, 1.50 eq., 1.12 mmol). The residue was purified by flash column chromatography using a gradient of 25:75:0 pet spirits:EtOAc:MeOH to 0:99:1 pet spirits:EtOAc:MeOH give **6.22** as an orange oil (39 mg, 15 %). ¹H NMR (401 MHz, CD₃CN) δ 7.60 (d, J = 8.6 Hz, 1H), 7.43 – 7.40 (m, 2H), 7.34 – 7.21 (m, 2H), 7.06 – 6.99 (m, 2H), 6.96 – 6.92 (m, 1H), 6.84 (dd, J = 8.6, 2.5 Hz, 1H), 6.61 (bs, 1H), 6.19 (t, J = 2.1 Hz, 1H), 4.26 – 4.09 (m, 2H), 3.60 – 3.47 (m, 2H), 3.43 (s, 2H), 2.45 (s, 3H). ¹³C NMR (101 MHz, CD₃CN) δ 171.71, 162.39, 154.81, 145.63, 140.08, 135.51, 133.68, 132.08, 130.90, 121.20, 119.78, 118.87, 116.23, 107.26, 106.01, 51.53, 42.94, 40.53, 20.68. LCMS (m/z): 361.0 [M+H]⁺, t_R = 3.78 min, standard method. HRMS (m/z): C₂₁H₂₀N₄O₂ requires 361.1659 [M+H]⁺; found 361.1664. HPLC: t_R = 6.88 min, >99 %, standard method.

***N*-(2-(1H-1,2,3-triazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)-*N*-ethylacetamide (6.23)**

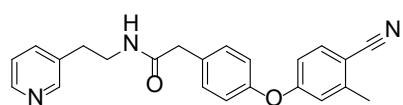


Prepared according to **general procedure 11** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.00 eq., 0.75 mmol),

HATU (569 mg, 2.0 eq., 1.50 mmol), dry DMF (2 mL), DBU (0.44 mL, 4.0 eq., 3.00 mmol) and *N*-

ethyl-2-(triazol-1-yl)ethanamine dihydrochloride (191 mg, 1.20 eq., 0.90 mmol). The residue was purified by flash column chromatography using a gradient of 20:80 pet spirits:EtOAc to 0:100 pet spirits:EtOAc. The residue was further purified by flash column chromatography using a gradient of 25:75 acetone:pet spirits to 33:67 acetone:pet spirits to give **6.23** as an orange oil (18 mg, 6.2 %). **¹H NMR** (401 MHz, CD₃CN) δ 7.80 (d, *J* = 1.1 Hz, 0.3H), 7.69 (d, *J* = 1.0 Hz, 1H), 7.64 – 7.59 (m, 1.7H), 7.29 – 7.23 (m, 1.4H), 7.15 – 7.11 (m, 0.6H), 7.05 – 7.01 (m, 1.4H), 7.01 – 6.97 (m, 0.6H), 6.95 (d, *J* = 2.5 Hz, 0.6H), 6.93 (d, *J* = 2.4 Hz, 0.4H), 6.87 – 6.81 (m, 1H), 4.55 (app td, *J* = 6.1, 1.6 Hz, 2H), 3.80 (t, *J* = 6.0 Hz, 0.6H), 3.74 (t, *J* = 6.2 Hz, 1.4H), 3.67 (s, 1.4H), 3.31 (q, *J* = 7.1 Hz, 0.6H), 3.21 (s, 0.6H), 3.16 (q, *J* = 7.1 Hz, 1.4H), 2.46 (s, 2.1H), 2.45 (s, 0.9H), 1.06 (t, *J* = 7.1 Hz, 0.9H), 0.99 (t, *J* = 7.1 Hz, 2.1H). **Spectra were consistent with rotamers in a 7:3 ratio.** **¹³C NMR** (101 MHz, CD₃CN) δ 171.87, 171.19, 162.46, 154.65, 154.62, 145.65, 135.54, 135.52, 134.54, 134.21, 133.77, 133.66, 132.08, 132.07, 125.79, 121.08, 121.00, 119.79, 118.89, 116.25, 107.26, 49.06, 48.32, 48.09, 46.70, 44.21, 41.10, 39.90, 39.35, 20.69, 14.16, 12.85. Multiple carbon resonances were confirmed by analysing the 2D HMBC NMR spectra. **LCMS** (*m/z*): 390.1 [M+H]⁺, *t_R* = 2.16 min, fragment LCMS. **HRMS** (*m/z*): C₂₂H₂₃N₅O₂ requires 390.1925 [M+H]⁺; found 390.1921. **HPLC**: *t_R* = 5.94 min, >99 %, standard method.

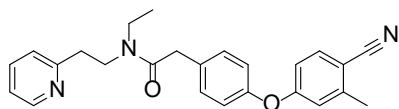
2-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N*-(2-(pyridin-3-yl)ethyl)acetamide (**6.24**)



Prepared according to **general procedure 11** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.00 eq., 0.75 mmol), HATU (569 mg, 2.0 eq., 1.50 mmol), dry DMF (2 mL), DBU (0.22 mL, 2.0 eq., 1.50 mmol) and 2-(3-pyridyl)ethanamine (137 mg, 1.50 eq., 1.12 mmol). The residue was purified by flash column chromatography using a gradient of 99:1 DCM:MeOH to 97:3 DCM:MeOH give **6.24** as a white solid (93 mg, 33 %). **¹H NMR** (401 MHz, CDCl₃) δ 8.45 (dd, *J* = 4.8, 1.7 Hz, 1H), 8.36 (dd, *J* = 2.2, 0.8 Hz, 1H), 7.53 (d, *J* = 8.5 Hz, 1H), 7.47 – 7.41 (m, 1H), 7.25 – 7.17 (m, 3H), 7.03 – 6.97 (m, 2H), 6.90 – 6.85 (m, 1H), 6.81 (dd, *J* = 8.6, 2.5 Hz, 1H), 5.58 (bs, 1H), 3.55 – 3.45 (m, 4H), 2.80 (t, *J* = 6.9 Hz, 2H), 2.49 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 170.86, 161.19, 154.34, 150.10, 148.04,

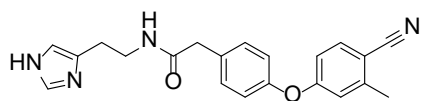
144.60, 136.37, 134.44, 134.32, 131.47, 131.12, 123.60, 120.69, 119.10, 118.20, 115.54, 106.71, 42.99, 40.53, 32.85, 20.69. **LCMS** (m/z): 372.0 $[M+H]^+$, t_R = 3.10 min, standard method. **HRMS** (m/z): $C_{23}H_{21}N_3O_2$ requires 372.1707 $[M+H]^+$; found 372.1723. **HPLC**: t_R = 5.12 min, >99 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N*-ethyl-*N*-(2-(pyridin-2-yl)ethyl)acetamide (6.25)

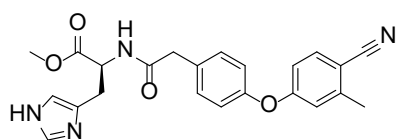


Prepared according to **general procedure 11** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (180 mg, 1.00 eq., 0.67 mmol),

HATU (512 mg, 2.0 eq., 1.35 mmol), dry DMF (2 mL), DBU (0.20 mL, 2.0 eq., 1.35 mmol) and *N*-ethyl-2-(2-pyridyl)ethanamine (126 mg, 1.25 eq., 0.84 mmol). The residue was purified by flash column chromatography using a gradient of 100:0 EtOAc:MeOH to 95:5 EtOAc:MeOH give **6.25** as a colourless oil (39 mg, 15 %). **1H NMR** (401 MHz, $CDCl_3$) δ 8.57 (ddd, J = 4.9, 1.9, 0.9 Hz, 0.6H), 8.50 (ddd, J = 4.9, 1.9, 0.9 Hz, 0.4H), 7.62 (td, J = 7.7, 1.8 Hz, 0.4H), 7.56 (td, J = 7.7, 1.9 Hz, 0.6H), 7.50 (dd, J = 8.5, 7.3 Hz, 1H), 7.30 – 7.22 (m, 1.2H), 7.23 – 7.19 (m, 0.8H), 7.19 – 7.14 (m, 1H), 7.15 – 7.06 (m, 1H), 7.03 – 6.93 (m, 2H), 6.84 (dd, J = 6.1, 2.5 Hz, 1H), 6.79 (ddd, J = 8.5, 5.9, 2.5 Hz, 1H), 3.77 – 3.69 (m, 2H), 3.68 (s, 1.2H), 3.56 (s, 0.8H), 3.43 (q, J = 7.1 Hz, 0.8H), 3.29 (q, J = 7.1 Hz, 1.2H), 3.09 – 3.03 (m, 1H), 2.99 (t, J = 7.3 Hz, 1H), 2.47 (s, 1.7H), 2.45 (s, 1.3H), 1.15 (t, J = 7.1 Hz, 1.5H), 1.11 (t, J = 7.1 Hz, 1.5H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.45, 170.40, 161.50, 161.48, 159.22, 158.17, 153.70, 153.63, 149.76, 149.28, 144.47, 144.44, 136.76, 136.58, 134.35, 134.32, 132.31, 132.20, 130.80, 123.71, 123.68, 121.95, 121.55, 120.51, 120.44, 118.79, 118.78, 118.25, 118.23, 115.34, 106.40, 106.34, 47.46, 46.24, 43.45, 40.70, 39.76, 39.61, 37.64, 36.38, 20.66, 20.63, 14.28, 12.87. **1H NMR and ^{13}C NMR spectra were consistent with rotamers in a 3:2 ratio.** **LCMS** (m/z): 400.0 $[M+H]^+$, t_R = 3.23 min, standard method. **HRMS** (m/z): $C_{25}H_{25}N_3O_2$ requires 400.2020 $[M+H]^+$; found 400.2036. **HPLC**: t_R = 5.58 min, >99 %, standard method.

***N*-(2-(1H-imidazol-4-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamide (6.26)**

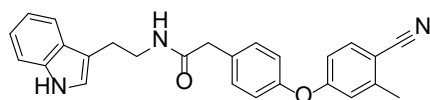
Prepared according to **general procedure 11** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.00 eq., 0.75 mmol), HATU (569 mg, 2.0 eq., 1.50 mmol) dry DMF (5 mL) Et₃N (0.41 mL, 4.1 eq., 3.00 mmol) and histamine dihydrochloride (207 mg, 1.50 eq., 1.12 mmol). The residue was purified by flash column chromatography with deactivated silica (Et₃N washed) and a gradient of 90:10 EtOAc:MeOH to 85:15 EtOAc:MeOH give **6.26** as an orange oil (11 mg, 4.1 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 11.80 (bs, 1H), 8.13 (bs, 1H), 7.75 (d, *J* = 8.6 Hz, 1H), 7.51 (s, 1H), 7.37 – 7.26 (m, 2H), 7.08 – 7.04 (m, 2H), 7.02 (d, *J* = 2.4 Hz, 1H), 6.88 (dd, *J* = 8.6, 2.5 Hz, 1H), 6.82 (bs, 1H), 3.42 (s, 2H), 3.29 (td, *J* = 7.4, 5.6 Hz, 2H), 2.64 (bs, 2H), 2.44 (s, 3H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 170.31, 161.54, 153.42, 144.81, 138.22, 135.21, 135.10, 133.80, 131.35, 120.45, 119.10, 118.43, 115.77, 112.22, 106.05, 42.09, 38.48, 27.93, 20.47. Multiple carbon resonances were confirmed by analysing the 2D HMBC NMR spectra. **LCMS** (*m/z*): 361.0 [M+H]⁺, *t*_R = 3.12 min, standard method. **HRMS** (*m/z*): C₂₁H₂₀N₄O₂ requires 361.1659 [M+H]⁺; found 361.1677. **HPLC**: *t*_R = 5.00 min, >99 %, standard method.

Methyl (2-(4-(4-cyano-3-methylphenoxy)phenyl)acetyl)-L-histidinate (6.27)

Prepared according to **general procedure 11** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.00 eq., 0.75 mmol), HATU (569 mg, 2.0 eq., 1.5 mmol), dry DMF (3 mL) Et₃N (0.4 mL, 4.0 eq., 3.0 mmol) and L-histidine methyl ester dihydrochloride (271 mg, 1.5 eq., 1.1 mmol). The residue was purified by flash column chromatography using a gradient of 67:33:0 pet spirits:EtOAc:MeOH to 0:95:5 pet spirits:EtOAc:MeOH give **6.27** as a colourless oil (27 mg, 8.6 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.91 (s, 1H), 7.47 – 7.35 (m, 2H), 7.28 (d, *J* = 7.6 Hz, 1H), 7.25 – 7.16 (m, 2H), 6.97 – 6.85 (m, 2H), 6.77 (d, *J* = 2.4 Hz, 1H), 6.71 (dd, *J* = 8.5, 2.5 Hz, 1H), 6.68 (s, 1H), 4.77 – 4.63 (m, 1H), 3.59 (s, 3H), 3.51 (s, 2H), 3.11 – 2.92 (m, 2H), 2.38 (s, 3H). **¹³C NMR** (101 MHz, CD₃OD) δ 173.68, 173.29, 163.14, 155.25, 145.75, 136.34, 135.57, 134.43, 133.58, 132.05, 121.48, 119.80, 118.94,

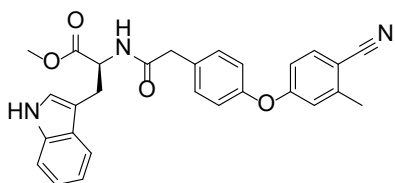
118.07, 116.41, 107.19, 54.07, 52.79, 42.60, 29.95, 20.52. **LCMS** (m/z): 419.0 $[M+H]^+$, t_R = 3.28 min, standard method. **HRMS** (m/z): $C_{23}H_{11}N_4O_4Na$ requires 441.1533 $[M+Na]^+$; found 441.1536. **HPLC**: t_R = 5.10 min, >99 %, standard method.

***N*-(2-(1*H*-indol-3-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamide (6.28)**



Prepared according to **general procedure 11** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.00 eq., 0.75 mmol), HATU (569 mg, 2.0 eq., 1.5 mmol), dry DMF (2 mL) Et_3N (0.22 mL, 2.0 eq., 1.5 mmol) and tryptamine (180 mg, 1.5 eq., 1.1 mmol). The residue was purified by flash column chromatography using a gradient of 75:25 pet spirits:EtOAc to 40:60 pet spirits:EtOAc give **6.28** as a yellow oil (74 mg, 24 %). **1H NMR** (401 MHz, CD_3CN) δ 9.17 (bs, 1H), 7.60 – 7.51 (m, 2H), 7.54 (d, J = 5.7 Hz, 1H), 7.36 (d, J = 8.2 Hz, 1H), 7.28 – 7.20 (m, 2H), 7.10 (ddd, J = 8.2, 6.9, 1.2 Hz, 1H), 7.04 – 6.99 (m, 2H), 6.99 – 6.94 (m, 2H), 6.91 (d, J = 2.5 Hz, 1H), 6.78 (dd, J = 8.6, 2.5 Hz, 1H), 6.62 (t, J = 5.9 Hz, 1H), 3.48 – 3.38 (m, 4H), 2.93 – 2.84 (m, 2H), 2.42 (s, 3H). **^{13}C NMR** (101 MHz, CD_3CN) δ 171.65, 162.33, 154.73, 145.59, 137.54, 135.49, 133.87, 131.97, 128.47, 123.66, 122.41, 121.12, 119.82, 119.75, 119.42, 118.90, 116.20, 113.31, 112.31, 107.22, 43.07, 40.81, 25.93, 20.69. **LCMS** (m/z): 410.0 $[M+H]^+$, t_R = 3.60 min, standard method. **HRMS** (m/z): $C_{26}H_{23}N_3O_2$ requires 410.1863 $[M+H]^+$; found 410.1872. **HPLC**: t_R = 6.52 min, >99 %, standard method.

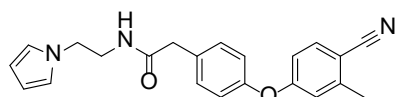
Methyl (2-(4-(4-cyano-3-methylphenoxy)phenyl)acetyl)-L-tryptophanate (6.29)



Prepared according to **general procedure 11** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (200 mg, 1.00 eq., 0.75 mmol), HATU (569 mg, 2.0 eq., 1.5 mmol), dry DMF (2 mL), Et_3N (0.22 mL, 2.0 eq., 1.5 mmol) and L-tryptophan methyl ester hydrochloride (285 mg, 1.5 eq., 1.1 mmol). The residue was purified by flash column chromatography using 99:1 DCM:MeOH to give **6.29** as a colourless oil (87 mg, 25 %). **1H NMR** (401 MHz, $CDCl_3$) δ 8.05 (bs, 1H), 7.51 (d, J = 8.5 Hz, 1H), 7.47 (d, J = 7.9 Hz, 1H), 7.35 (dt, J = 8.2, 0.9 Hz, 1H), 7.19 (ddd, J = 8.2, 7.0, 1.2 Hz, 1H),

7.19 – 7.11 (m, 2H), 7.10 (ddd, $J = 8.0, 7.0, 1.0$ Hz, 1H), 6.97 – 6.89 (m, 2H), 6.86 – 6.80 (m, 2H), 6.77 (dd, $J = 8.6, 2.3$ Hz, 1H), 5.96 (d, $J = 7.8$ Hz, 1H), 4.94 (dt, $J = 7.8, 5.3$ Hz, 1H), 3.70 (s, 3H), 3.51 (s, 2H), 3.34 (dd, $J = 14.8, 5.4$ Hz, 1H), 3.27 (dd, $J = 14.8, 5.2$ Hz, 1H), 2.48 (s, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 172.30, 170.44, 161.30, 154.07, 144.52, 136.15, 134.37, 131.21, 131.13, 127.68, 122.80, 122.24, 120.63, 119.71, 118.87, 118.46, 118.22, 115.35, 111.40, 109.67, 106.46, 53.18, 52.44, 42.69, 27.40, 20.61. LCMS (m/z): 468.0 $[\text{M}+\text{H}]^+$, $t_{\text{R}} = 3.79$ min, standard method. HRMS (m/z): $\text{C}_{28}\text{H}_{25}\text{N}_3\text{O}_4$ requires 468.1918 $[\text{M}+\text{H}]^+$; found 468.1930. HPLC: $t_{\text{R}} = 6.56$ min, >99 %, standard method.

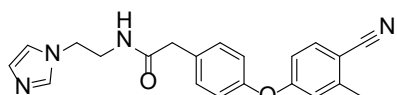
***N*-(2-(1H-Pyrrol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamide (6.30)**



Prepared according to **general procedure 12** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (150 mg, 1.00 eq., 0.56 mmol),

EDC hydrochloride (215 mg, 2.0 eq., 1.12 mmol), HOBt (152 mg, 2.0 eq., 1.12 mmol), dry DMF (2 mL), Et_3N (0.230 mL, 3.0 eq., 1.68 mmol) and 2-pyrrol-1-ylethanamine (93 mg, 1.50 eq., 0.84 mmol). The residue was purified by flash column chromatography using a gradient of 50:50 pet spirits:EtOAc to 34:66 pet spirits:EtOAc to give **6.30** as a white solid (126 mg, 62 %). ^1H NMR (401 MHz, CDCl_3) δ 7.54 (d, $J = 8.5$ Hz, 1H), 7.25 – 7.19 (m, 2H), 7.07 – 6.96 (m, 2H), 6.88 (d, $J = 2.5$ Hz, 1H), 6.81 (dd, $J = 8.6, 2.5$ Hz, 1H), 6.52 (app t, $J = 2.1$ Hz, 2H), 6.12 (app t, $J = 2.1$ Hz, 2H), 5.44 (bs, 1H), 4.04 – 3.96 (m, 2H), 3.63 – 3.49 (m, 4H), 2.50 (s, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 171.02, 161.23, 154.49, 144.66, 134.50, 131.25, 131.22, 120.82, 120.68, 119.15, 118.24, 115.59, 108.92, 106.83, 48.78, 43.07, 41.23, 20.76. LCMS (m/z): 360.0 $[\text{M}+\text{H}]^+$, $t_{\text{R}} = 3.37$ min, standard method. HRMS (m/z): $\text{C}_{22}\text{H}_{21}\text{N}_3\text{O}_2$ requires 360.1707 $[\text{M}+\text{H}]^+$; found 360.1699. HPLC: $t_{\text{R}} = 6.22$ min, >99 %, standard method.

***N*-(2-(1H-imidazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamide (6.31)**

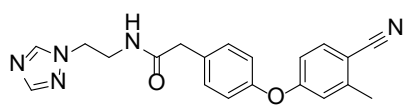


Prepared according to **general procedure 12** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (150 mg, 1.00 eq., 0.561

mmol), EDC hydrochloride (215 mg, 2.0 eq., 1.12 mmol), HOBt (152 mg, 2.0 eq., 1.12 mmol), dry

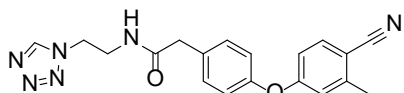
DMF (3 mL), Et₃N (0.38 mL, 5.0 eq., 2.80 mmol) and 2-imidazol-1-ylethanamine dihydrochloride (155 mg, 1.50 eq., 0.84 mmol). The residue was purified by flash column chromatography using a gradient of 90:10 EtOAc:MeOH to 80:20 EtOAc:MeOH to give **6.31** as a colourless oil (125 mg, 62 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.54 (d, *J* = 8.6 Hz, 1H), 7.31 – 7.22 (m, 3H), 7.05 – 7.00 (m, 2H), 6.98 (bs, 1H), 6.92 – 6.86 (m, 2H), 6.85 – 6.79 (m, 2H), 4.17 – 4.02 (m, 2H), 3.64 – 3.51 (m, 4H), 2.50 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 171.49, 161.21, 154.34, 144.59, 137.05, 134.44, 131.49, 131.04, 129.47, 120.68, 119.17, 119.07, 118.20, 115.53, 106.67, 46.24, 42.72, 40.77, 20.68. **LCMS** (*m/z*): 361.0 [M+H]⁺, *t_R* = 2.95 min, standard method. **HRMS** (*m/z*): C₂₁H₂₀N₄O₂ requires 361.1659 [M+H]⁺; found 361.1663. **HPLC**: *t_R* = 5.01 min, >99 %, standard method.

***N*-(2-(1H-1,2,4-triazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamide (6.32)**



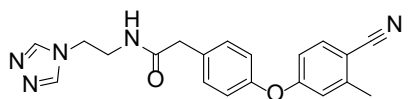
Prepared according to **general procedure 12** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (150 mg, 1.00 eq., 0.561 mmol), EDC hydrochloride (215 mg, 2.0 eq., 1.12 mmol), HOBT (152 mg, 2.0 eq., 1.12 mmol), dry DMF (3 mL) Et₃N (0.38 mL, 5.0 eq., 2.80 mmol) and 2-(1,2,4-triazol-1-yl)ethanamine dihydrochloride (156 mg, 1.50 eq., 0.84 mmol). The residue was purified by flash column chromatography using a gradient of 90:10 EtOAc:MeOH to 80:15 EtOAc:MeOH to give **6.32** as a colourless oil (141 mg, 70 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.93 (s, 1H), 7.85 (s, 1H), 7.51 (d, *J* = 8.5 Hz, 1H), 7.25 – 7.16 (m, 2H), 7.03 – 6.97 (m, 2H), 6.86 (d, *J* = 2.5 Hz, 1H), 6.80 (dd, *J* = 8.5, 2.5 Hz, 1H), 6.23 – 6.13 (m, 1H), 4.34 – 4.24 (m, 2H), 3.67 (q, *J* = 5.8 Hz, 2H), 3.51 (s, 2H), 2.47 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 171.36, 161.17, 154.43, 152.40, 144.60, 143.72, 134.45, 131.16, 131.09, 120.73, 119.09, 118.18, 115.54, 106.70, 48.42, 42.84, 39.46, 20.67. **LCMS** (*m/z*): 362.0 [M+H]⁺, *t_R* = 3.14 min, standard method. **HRMS** (*m/z*): C₂₀H₁₉N₅O₂ requires 362.1612 [M+H]⁺; found 362.1615. **HPLC**: *t_R* = 5.38 min, >99 %, standard method.

***N*-(2-(1H-tetrazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamide (6.33)**



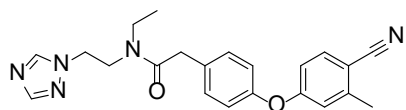
Prepared according to **general procedure 12** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (150 mg, 1.00 eq., 0.56 mmol), EDC hydrochloride (215 mg, 2.0 eq., 1.12 mmol), HOBt (152 mg, 2.0 eq., 1.12 mmol), dry DMF (2 mL) Et₃N (0.306 mL, 4.0 eq., 2.24 mmol) and 2-(tetrazol-1-yl)ethanamine hydrochloride (126 mg, 1.50 eq., 0.84 mmol). The residue was purified by flash column chromatography using EtOAc to give **6.33** as a white solid (125 mg, 61 %). **¹H NMR** (401 MHz, CDCl₃) δ 8.47 (s, 1H), 7.52 (d, *J* = 8.5 Hz, 1H), 7.33 – 7.24 (m, 2H), 7.05 – 6.93 (m, 3H), 6.88 (d, *J* = 2.5 Hz, 1H), 6.81 (dd, *J* = 8.6, 2.5 Hz, 1H), 4.63 – 4.54 (m, 2H), 3.76 (q, *J* = 5.8 Hz, 2H), 3.57 (s, 2H), 2.48 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 171.96, 161.15, 154.27, 144.55, 143.23, 134.40, 131.33, 130.99, 120.61, 119.00, 118.17, 115.46, 106.50, 47.51, 42.58, 39.30, 20.59. **LCMS** (*m/z*): 362.9 [M+H]⁺, *t_R* = 3.37 min, standard method. **HRMS** (*m/z*): C₁₉H₁₈N₆O₂ requires 363.1564 [M+H]⁺; found 363.1572. **HPLC**: *t_R* = 5.46 min, >99 %, standard method.

***N*-(2-(4H-1,2,4-triazol-4-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamide (6.34)**



Prepared according to **general procedure 12** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (150 mg, 1.00 eq., 0.56 mmol), EDC hydrochloride (215 mg, 2.0 eq., 1.12 mmol), HOBt (152 mg, 2.0 eq., 1.12 mmol), dry DMF (2 mL), Et₃N (0.382 mL, 5.0 eq., 2.80 mmol) and 2-(1,2,4-triazol-4-yl)ethanamine dihydrochloride (140 mg, 1.34 eq., 0.76 mmol). The residue was purified by flash column chromatography using a gradient of 90:10 EtOAc:MeOH to 75:25 EtOAc:MeOH to give **6.34** as a white solid (74 mg, 36 %). **¹H NMR** (401 MHz, *d*₆-DMSO) δ 8.37 (s, 2H), 8.20 (t, *J* = 5.7 Hz, 1H), 7.75 (d, *J* = 8.6 Hz, 1H), 7.34 – 7.24 (m, 2H), 7.09 – 7.01 (m, 3H), 6.90 (dd, *J* = 8.6, 2.5 Hz, 1H), 4.14 – 4.08 (m, 2H), 3.46 – 3.35 (m, 4H), 2.44 (s, 3H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 170.51, 161.05, 153.05, 144.37, 143.20, 134.71, 132.90, 130.87, 120.05, 118.62, 117.96, 115.38, 105.58, 43.68, 41.51, 39.83 (confirmed by HMBC), 19.97. **LCMS** (*m/z*): 362.0 [M+H]⁺, *t_R* = 3.31 min, standard method. **HRMS** (*m/z*): C₂₀H₁₉N₅O₂ requires 362.1612 [M+H]⁺; found 362.1607. **HPLC**: *t_R* = 5.11 min, >99 %, standard method.

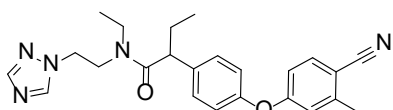
***N*-(2-(1*H*-1,2,4-triazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)-*N*-ethylacetamide (6.38)**



Prepared according to **general procedure 13** 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]-*N*-[2-(1,2,4-triazol-1-yl)ethyl]acetamide

(193 mg, 1.0 eq., 0.53 mmol) dry THF (5 mL) NaH (60 % dispersion in mineral oil, 106 mg, 5.0 eq., 2.67 mmol) and ethyl iodide (47 μ L, 1.1 eq., 0.59 mmol). The residue was purified by flash column chromatography using a gradient of 0:100 MeOH:EtOAc to 10:90 MeOH:EtOAc. The residue was further purified by C18 reverse phase column chromatography using a gradient of 10:90 MeOH:H₂O to 100:0 MeOH:H₂O give **6.38** as a colourless oil (20 mg, 9.6 %). **¹H NMR** (401 MHz, CDCl₃) δ 8.01 (app d, J = 3.2 Hz, 0.3H), 7.93 (s, 0.85H), 7.83 (s, 0.85H), 7.56 – 7.47 (m, 1.0H), 7.30 – 7.22 (m, 1.8H), 7.21 – 7.14 (m, 0.2H), 7.07 – 6.95 (m, 2.0H), 6.88 (d, J = 2.4 Hz, 0.90H), 6.87 – 6.75 (m, 1.1H), 4.43 (t, J = 5.9 Hz, 1.75H), 4.26 (t, J = 6.1 Hz, 0.25H), 3.78 (t, J = 6.2 Hz, 0.25H), 3.72 (t, J = 5.9 Hz, 1.75H), 3.66 (s, 1.75H), 3.46 – 3.29 (m, 0.50H), 3.02 (q, J = 7.1 Hz, 1.75H), 2.48 (s, 2.6H), 2.47 (s, 0.4H), 1.14 (t, J = 7.1 Hz, 0.4H), 0.98 (t, J = 7.1 Hz, 2.6H). **¹³C NMR** (101 MHz, CDCl₃) δ 171.40, 161.36, 154.16, 154.00, 152.45, 144.60, 144.56, 143.94, 134.46, 134.42, 131.48, 131.46, 130.79, 130.72, 120.68, 120.60, 119.00, 118.96, 118.26, 115.52, 115.47, 106.64, 106.60, 48.10, 47.27, 47.06, 46.72, 44.49, 40.92, 39.77, 39.40, 20.69, 14.05, 12.78. **LCMS** (m/z): 390.0 [M+H]⁺, t_R = 3.54 min, standard method. **HRMS** (m/z): C₂₂H₂₃N₅O₂ requires 390.1925 [M+H]⁺; found 390.1913. **HPLC**: t_R = 5.84 min, >99 %, standard method.

***N*-(2-(1*H*-1,2,4-triazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)-*N*-ethylbutanamide (6.39)**

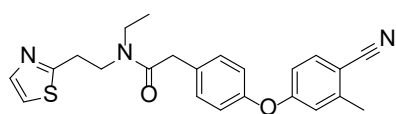


Prepared according to **general procedure 13** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]-*N*-[2-(1,2,4-triazol-1-yl)ethyl]acetamide

(116 mg, 1.0 eq., 0.32 mmol), dry THF (3 mL), NaH (60 % dispersion in mineral oil, 64 mg, 5.0 eq., 1.60 mmol) and ethyl iodide (154 μ L, 6.0 eq., 1.93 mmol). The residue was purified by flash column chromatography using EtOAc to give **6.39** as a colourless oil (101 mg, 75 %). **¹H NMR** (401 MHz,

CDCl₃) δ 7.99 (s, 0.1H), 7.92 (s, 0.2H), 7.86 (s, 0.9H), 7.56 – 7.46 (m, 1.8H), 7.35 – 7.25 (m, 2.0H), 7.04 – 6.95 (m, 1.8H), 6.98 – 6.92 (m, 0.2H), 6.89 (d, J = 2.4 Hz, 0.9H), 6.86 – 6.78 (m, 1.0H), 6.77 (dd, J = 8.6, 2.5 Hz, 0.1H), 4.45 – 4.29 (m, 1.7H), 4.24 – 4.00 (m, 0.3H), 3.94 – 3.76 (m, 1.0H), 3.57 – 3.39 (m, 2.0H), 3.24 – 3.07 (m, 1.1H), 2.65 (dq, J = 14.3, 7.1 Hz, 0.9H), 2.47 (s, 2.5H), 2.46 (s, 0.4H), 2.14 – 2.02 (m, 1.0H), 1.96 – 1.85 (m, 0.1H), 1.74 (dp, J = 14.3, 7.3 Hz, 0.9H), 1.06 (t, J = 7.1 Hz, 0.3H), 0.89 (t, J = 7.1 Hz, 2.7H), 0.86 (t, J = 7.3 Hz, 2.7H), 0.74 (t, J = 7.3 Hz, 0.3H). ¹³C NMR (101 MHz, CDCl₃) δ 173.58, 161.25, 154.27, 152.36, 144.60, 143.84, 136.63, 134.46, 129.55, 120.66, 119.07, 118.21, 115.51, 106.64, 77.48, 77.16, 76.84, 49.87, 47.27, 46.78, 43.79, 28.21, 20.62, 14.30, 12.51. **Spectra were consistent with rotamers in a 9:1 ratio.** LCMS (m/z): 418.1 [M+H]⁺, t_R = 3.68 min, standard method. HRMS (m/z): C₂₄H₂₇N₅O₂ requires 418.2238 [M+H]⁺; found 418.2232. HPLC: t_R = 6.41 min, >99 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-N-ethyl-N-(2-(thiazol-2-yl)ethyl)acetamide (6.40)



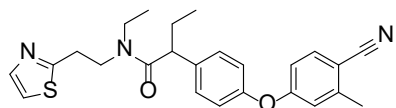
Prepared according to **general procedure 13** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]-N-(2-thiazol-2-ylethyl)acetamide (140

mg, 1.0 eq., 0.37 mmol), dry THF (3 mL), NaH (60 % dispersion in mineral oil, 74 mg, 5.0 eq., 1.85 mmol) and ethyl iodide (179 μ L, 6.0 eq., 2.2 mmol). The residue was purified by flash column chromatography using a gradient of 50:50 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give **6.40** as a white solid (19.3 mg, 13 %). ¹H NMR (401 MHz, CDCl₃) δ 7.74 (d, J = 3.3 Hz, 0.4H), 7.68 (d, J = 3.4 Hz, 0.6H), 7.50 (app dd, J = 8.5, 5.5 Hz, 1H), 7.30 – 7.25 (m, 1.6H), 7.25 – 7.21 (m, 0.75H), 7.20 (d, J = 3.3 Hz, 0.65H), 7.02 – 6.95 (m, 2H), 6.85 (app dd, J = 4.9, 2.4 Hz, 1H), 6.84 – 6.75 (m, 1H), 3.80 – 3.72 (m, 2H), 3.70 (s, 1.3H), 3.60 (s, 0.7H), 3.45 (q, J = 7.1 Hz, 0.7H), 3.35 – 3.26 (m, 2.6H), 3.23 – 3.15 (m, 0.7H), 2.48 (s, 2H), 2.46 (s, 1H), 1.16 (t, J = 7.1 Hz, 1H), 1.11 (t, J = 7.1 Hz, 2H). **Spectra consistent with rotamers in a 2:1 ratio.** ¹³C NMR (101 MHz, CDCl₃) δ 170.68, 170.45, 167.69, 166.12, 161.51, 161.48, 153.82, 153.81, 144.53, 144.52, 143.00, 142.54, 134.40, 134.39, 132.03, 132.02, 130.87, 130.80, 120.59, 120.58, 118.88, 118.86, 118.28, 115.42, 115.40, 106.49, 106.47, 47.40, 46.36, 43.85, 40.83, 39.90, 39.80, 32.51, 31.38, 20.71, 20.69, 14.28,

12.93. Multiple carbon resonances were confirmed by analysing the 2D HMBC NMR spectra.

LCMS (m/z): 406.0 $[M+H]^+$, t_R = 3.38 min, standard method. **HRMS** (m/z): $C_{23}H_{23}N_3O_2S$ requires 406.1584 $[M+H]^+$; found 406.1591. **HPLC**: t_R = 6.18 min, >99 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N*-ethyl-*N*-(2-(thiazol-2-yl)ethyl)butanamide (6.41)

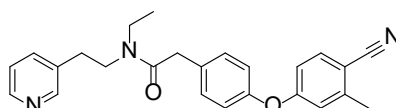


Product **6.41** isolated from the above alkylation reaction as a colourless oil (4.31 mg, 2.7 %). **1H NMR** (401 MHz, $CDCl_3$) δ 7.75

(d, J = 3.3 Hz, 0.35H), 7.67 (d, J = 3.4 Hz, 0.65H), 7.52 (app dd, J = 8.5, 5.3 Hz, 1.0H), 7.38 – 7.30 (m, 2H), 7.25 (d, J = 3.4 Hz, 0.35H), 7.18 (d, J = 3.3 Hz, 0.65H), 7.02 – 6.95 (m, 2H), 6.86 (app d, J = 2.5 Hz, 1H), 6.81 – 6.77 (m, 1H), 3.86 – 3.42 (m, 3.7H), 3.41 – 3.22 (m, 2.3H), 3.17 (dq, J = 14.3, 7.3 Hz, 1H), 2.98 (ddd, J = 14.6, 9.1, 5.2 Hz, 0.4H), 2.49 (s, 2H), 2.48 (s, 1H), 2.12 (dq, J = 13.5, 7.4 Hz, 0.8H), 2.01 (dq, J = 14.6, 7.3 Hz, 0.5H), 1.83 – 1.63 (m, 0.7H), 1.12 (t, J = 7.1 Hz, 1.1H), 1.03 (t, J = 7.1 Hz, 2H), 0.90 (t, J = 7.3 Hz, 2.1H), 0.82 (t, J = 7.3 Hz, 1H). **Spectra consistent with rotamers in a 2:1 ratio.**

Insufficient material was isolated to acquire ^{13}C NMR NMR spectra. **LCMS** (m/z): 434.0 $[M+H]^+$, t_R = 3.35 min, standard method. **HRMS** (m/z): $C_{25}H_{27}N_3O_2S$ requires 434.1897 $[M+H]^+$; found 434.1905. **HPLC**: t_R = 6.75 min, >99 %, standard method.

2-(4-(4-cyano-3-methylphenoxy)phenyl)-*N*-ethyl-*N*-(2-(pyridin-3-yl)ethyl)acetamide (6.42)

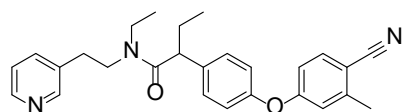


Prepared according to **general procedure 13** with 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]-*N*-[2-(3-pyridyl)ethyl]acetamide (118

mg, 1.0 eq., 0.32 mmol), dry THF (3 mL), NaH (60 % dispersion in mineral oil, 64 mg, 5.0 eq., 1.60 mmol) and ethyl iodide (153 μ L, 6.0 eq., 2.2 mmol). The residue was purified by flash column chromatography using a gradient of 0:100 MeOH:EtOAc to 1:99 MeOH:EtOAc give **6.42** as a colourless oil (41 mg, 32 %). **1H NMR** (401 MHz, $CDCl_3$) δ 8.51 (dd, J = 4.8, 1.6 Hz, 0.3H), 8.46 – 8.39 (m, 1.7H), 7.55 – 7.47 (m, 1.7H), 7.44 (dt, J = 7.8, 2.0 Hz, 0.3H), 7.28 – 7.22 (m, 1.7H), 7.22 – 7.15 (m, 1.3H), 7.04 – 6.94 (m, 2.0H), 6.87 – 6.75 (m, 2.0H), 3.69 (s, 1.4H), 3.59 – 3.48 (m, 2.6H), 3.45 (q, J = 7.1 Hz, 0.6H), 3.27 (q, J = 7.2 Hz, 1.4H), 2.92 – 2.84 (m, 1.4H), 2.79 (t, J = 7.5 Hz, 0.6H), 2.47 (s, 2.1H), 2.45 (s, 0.9H), 1.16 (t, J = 7.1 Hz, 1.0H), 1.12 (t, J = 7.1 Hz, 2.0H). **Spectra**

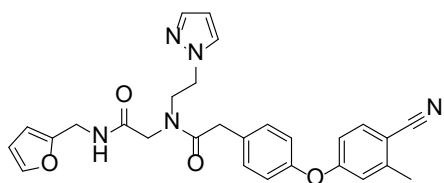
consistent with rotamers in a 2:1 ratio. ^{13}C NMR (101 MHz, CDCl_3) δ 170.55, 170.15, 161.46, 161.39, 153.88, 153.84, 150.12, 150.06, 148.47, 147.92, 144.51, 136.51, 136.33, 134.63, 134.39, 133.60, 131.97, 131.85, 130.81, 130.66, 123.71, 123.53, 120.59, 118.87, 118.26, 115.43, 115.41, 106.50, 106.47, 49.03, 47.46, 43.56, 40.85, 39.98, 39.79, 32.68, 31.29, 20.68, 20.66, 14.30, 12.88. **LCMS** (m/z): 400.0 $[\text{M}+\text{H}]^+$, t_{R} = 3.19 min, standard method. **HRMS** (m/z): $\text{C}_{25}\text{H}_{25}\text{N}_3\text{O}_2$ requires 400.2020 $[\text{M}+\text{H}]^+$; found 400.2023. **HPLC**: t_{R} = 5.52 min, >99 %, standard method.

2-(4-(4-cyano-3-methylphenoxy)phenyl)-N-ethyl-N-(2-(pyridin-3-yl)ethyl)butanamide (6.43)



Product **6.43** isolated from the above alkylation reaction as a white solid (36 mg, 27 %). ^1H NMR (401 MHz, CDCl_3) δ 8.56 – 8.29 (m, 2H), 7.52 (d, J = 8.4 Hz, 0.85H), 7.50 – 7.46 (m, 0.85H), 7.42 (dt, J = 7.8, 2.0 Hz, 0.3H), 7.36 – 7.25 (m, 2.0H), 7.27 – 7.20 (m, 0.3H), 7.16 (dd, J = 7.8, 4.8 Hz, 0.7H), 7.03 – 6.94 (m, 2.0H), 6.87 (d, J = 2.4 Hz, 0.7H), 6.84 (d, J = 2.4 Hz, 0.3H), 6.85 – 6.74 (m, 1.0H), 3.72 – 3.26 (m, 4.3H), 3.07 (dq, J = 14.5, 7.1 Hz, 0.7H), 2.91 – 2.73 (m, 1.7H), 2.60 (ddt, J = 13.9, 8.9, 4.6 Hz, 0.3H), 2.48 (s, 2.1H), 2.46 (s, 0.9H), 2.18 – 1.92 (m, 1.6H), 1.85 – 1.57 (m, 1.1H), 1.11 (t, J = 7.1 Hz, 0.9H), 1.03 (t, J = 7.1 Hz, 2.1H), 0.88 (t, J = 7.3 Hz, 2.1H), 0.79 (t, J = 7.3 Hz, 0.9H). **Spectra consistent with rotamers in a 7:3 ratio.** ^{13}C NMR (101 MHz, CDCl_3) δ 172.89, 172.39, 161.42, 161.28, 154.06, 153.98, 150.10, 150.04, 148.42, 147.83, 144.57, 137.18, 137.11, 136.61, 136.32, 134.90 (**Confirmed via HMBC**), 134.46, 129.69, 129.62, 120.59, 119.11, 119.05, 118.30, 118.26, 115.47, 115.43, 106.68, 106.59, 50.22, 49.94, 48.53, 47.69, 43.11, 41.16, 32.96, 31.34, 28.78, 28.63, 20.74, 20.72, 14.60, 12.83, 12.61, 12.54. **LCMS** (m/z): 428.0 $[\text{M}+\text{H}]^+$, t_{R} = 3.30 min, standard method. **HRMS** (m/z): $\text{C}_{27}\text{H}_{29}\text{N}_3\text{O}_2$ requires 428.2333 $[\text{M}+\text{H}]^+$; found 428.2337. **HPLC**: t_{R} = 5.99 min, >99 %, standard method.

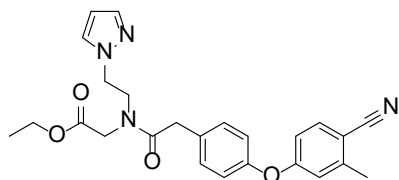
***N*-(2-(1H-pyrazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)-*N*-(2-((furan-2-ylmethyl)amino)-2-oxoethyl)acetamide (6.50)**



Mixture 1: To a solution of 2-pyrazol-1-ylethanamine (127 mg, 4.0 eq., 1.14 mmol) and Et₃N (200 μ L, 7.5 eq., 1.7 mmol) dissolved in DMF (1.5 mL) was added a solution of 2-bromo-*N*-(2-furylmethyl)acetamide (50 mg, 1.0 eq., 0.23 mmol) dissolved in DMF (1 mL) dropwise. **Mixture 2:** To a separate oven dried RBF was added 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]acetic acid (306 mg, 5.0 eq., 1.15 mmol), EDC hydrochloride (440 mg, 10 eq., 2.29 mmol), HOBt (309 mg, 10 eq., 2.29 mmol) and Et₃N (200 μ L, 7.5 eq., 1.7 mmol) in DMF (2.5 mL). Both reactions were stirred for 15 min at room temperature. After this time, **mixture 1** was added dropwise to **mixture 2** and stirred for 4 h at room temperature. The reaction was diluted in EtOAc (20 mL) and H₂O (20 mL). The layers were separated and the aqueous further extracted with EtOAc (20 mL), the combined organic layers were washed with hydrochloride (aq., 2 M, 20 mL), brine (10 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 99:1 DCM:MeOH to 98:2 DCM:MeOH. The residue was further purified by reverse phase preparative HPLC using a C8 silica column and a gradient of 70:30 H₂O:MeCN to 20:80 H₂O:MeCN over 12 min to give **6.50** as a white solid (33 mg, 29 %). **¹H NMR** (401 MHz, CDCl₃) δ 7.93 (bs, 0.3H), 7.56 (d, *J* = 1.8 Hz, 0.65H), 7.45 – 7.47 (m, 1.1H), 7.44 – 7.36 (m, 0.95H), 7.36 – 7.29 (m, 1.9H), 7.16 (d, *J* = 8.3 Hz, 0.70H), 7.05 (d, *J* = 8.4 Hz, 1.3H), 7.00 – 6.90 (m, 2.0H), 6.88 – 6.82 (m, 1.0H), 6.82 – 6.76 (m, 1.0H), 6.32 (dd, *J* = 3.2, 1.8 Hz, 0.35H), 6.29 (dd, *J* = 3.2, 1.8 Hz, 0.65H), 6.28 – 6.21 (m, 1.3H), 6.20 (d, *J* = 3.2 Hz, 0.65H), 4.48 – 4.30 (m, 4H), 4.01 (s, 1.4H), 3.91 – 3.79 (m, 2.7H), 3.57 (s, 0.7H), 3.00 (s, 1.4H), 2.48 (s, 1.1H), 2.48 (s, 1.9H). **¹³C NMR** (101 MHz, CDCl₃) δ 172.65, 172.36, 169.04, 168.29, 161.42, 161.36, 154.05, 154.01, 151.07, 150.88, 144.58, 142.46, 142.30, 140.48, 139.70, 139.67, 134.44, 131.15, 131.12, 131.10, 130.97, 120.51, 120.45, 119.03, 118.97, 118.28, 115.51, 115.47, 110.66, 110.57, 108.04, 107.73, 106.61, 106.60, 106.33, 106.28, 53.53, 52.66, 51.42, 50.17, 49.71, 49.46, 39.66, 38.27, 36.60, 36.57, 20.73. **Spectra**

consistent with rotamers in a 7:3 ratio. LCMS (m/z): 498.2 $[M+H]^+$, t_R = 3.29 min, standard method. HRMS (m/z): $C_{28}H_{27}N_5O_4$ requires 498.2136 $[M+H]^+$; found 498.2152. HPLC: t_R = 5.67 min, 99 %, standard method.

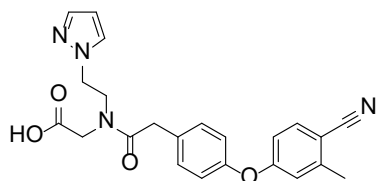
Ethyl *N*-(2-(1H-pyrazol-1-yl)ethyl)-*N*-(2-(4-(4-cyano-3-methylphenoxy)phenyl)acetyl)glycinate (6.51)



In an oven dried RBF cooled to room temperature was added NaH (60 % dispersion in mineral oil, (133 mg, 6.0 eq., 3.33 mmol) under N_2 . Excess mineral oil was removed by trituration with dry pet spirits (3 mL). THF (4 mL) was added and the reaction cooled to 0 °C, 2-[4-(4-cyano-3-methylphenoxy)phenyl]-*N*-(2-pyrazol-1-ylethyl)acetamide (300 mg, 1.0 eq., 0.83 mmol) was added and the reaction stirred for 15 min at 0 °C. Ethyl bromoacetate (146 mg, 1.2 eq., 0.66 mmol), KI (138 mg, 1.0 eq., 0.83 mmol) and 18-crown-6 (242 mg, 1.1 eq., 0.92 mmol) were added and the reaction stirred for a further 30 min on ice. The ice bath was removed, the reaction warmed to room temperature slowly and heated at reflux overnight. After this time, additional NaH (60 % dispersion, 200 mg, 6 eq., 5.0 mmol) and ethyl bromoacetate (146 mg, 1.2 eq., 0.66 mmol) were added and the reaction heated at reflux for a further 3 days. After this time, the reaction mixture was filtered and the filtrate diluted in EtOAc (20 mL) and H_2O (20 mL). The layers were separated and the aqueous further extracted with EtOAc (20 mL), the combined organic layers were washed with H_2O (15 mL), brine (10 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using 66:34 EtOAc:pet spirits to give **6.51** as a yellow oil (160 mg, 43 %). 1H NMR (401 MHz, $CDCl_3$) δ 7.60 (d, J = 2.5 Hz, 1.1H), 7.55 – 7.43 (m, 1.4H), 7.27 – 7.18 (m, 1.3H), 7.19 – 7.11 (m, 1.1H), 7.03 – 6.89 (m, 2.0H), 6.88 – 6.72 (m, 2.0H), 6.29 (t, J = 2.1 Hz, 0.6H), 6.18 (t, J = 2.1 Hz, 0.4H), 4.34 (dd, J = 6.4, 5.3 Hz, 0.8H), 4.29 (dd, J = 6.0, 4.7 Hz, 1.2H), 4.19 (q, J = 7.2 Hz, 1.6H), 4.12 (q, J = 7.1 Hz, 1.2H), 3.99 (s, 1.2H), 3.84 (td, J = 5.6, 2.1 Hz, 1.9H), 3.67 (s, 0.8H), 3.57 (s, 0.9H), 3.12 (s, 1.4H), 2.48 (s, 1.6H), 2.46 (s, 1.6H), 1.26 (t, J = 7.1 Hz, 1.6H), 1.23 (d, J = 7.2 Hz, 1.4H). Spectra consistent with rotamers in a 3:2 ratio. ^{13}C NMR (101

MHz, CDCl₃) δ 171.85, 171.74, 169.48, 169.06, 161.45, 161.40, 154.05, 153.85, 144.56, 144.51, 140.64, 140.10, 134.41, 134.38, 131.35, 131.25, 131.00, 130.97, 130.95, 130.40, 120.59, 120.46, 118.90, 118.87, 118.28, 118.23, 115.43, 115.42, 106.58, 106.45, 106.22, 105.64, 61.81, 61.50, 50.96, 50.64, 50.33, 49.90, 48.99, 39.84, 38.29, 20.69, 20.67, 14.23, 14.22. **LCMS** (m/z): 477.2 [M+H]⁺, t_R = 3.92 min, standard method. **HRMS** (m/z): C₂₅H₂₆N₄O₄ requires 477.2027 [M+H]⁺; found 477.2042. **HPLC**: t_R = 6.22 min, 99 %, standard method.

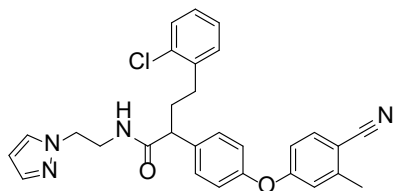
***N*-(2-(1H-pyrazol-1-yl)ethyl)-*N*-(2-(4-(4-cyano-3-methylphenoxy)phenyl)acetyl)glycine (6.53)**



Ester **6.51** (154 mg, 1.0 eq., 0.34 mmol) dissolved in MeOH (2 mL) was added NaOH (aq., 2 M, 1 mL, 5.8 eq., 2.0 mmol). The reaction was stirred for 4 h at room temperature. After this time, the reaction mixture was acidified with hydrochloride (aq., 2 M, 3 mL) and diluted in EtOAc (15 mL). The layers were separated and the aqueous layer further extracted with EtOAc (2 × 15 mL). Combined organic layers were dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified via flash column chromatography using a gradient of 98:2 DCM:MeOH to 90:10 DCM:MeOH to give **6.53** as a white solid (88 mg, 61 %). **¹H NMR** (401 MHz, CDCl₃) δ 9.51 (s, 1.0H), 7.55 (d, J = 1.9 Hz, 0.45H), 7.47 (d, J = 2.1 Hz, 1.0H), 7.44 (d, J = 3.9 Hz, 0.5H), 7.42 (d, J = 3.9 Hz, 0.5H), 7.30 (d, J = 2.3 Hz, 0.55H), 7.20 – 7.12 (m, 1.1H), 7.11 – 7.03 (m, 0.9H), 6.96 – 6.84 (m, 2.0H), 6.79 (d, J = 2.5 Hz, 0.55H), 6.76 (d, J = 2.4 Hz, 0.45H), 6.76 – 6.66 (m, 1.0H), 6.24 (t, J = 2.1 Hz, 0.45H), 6.17 (t, J = 2.2 Hz, 0.55H), 4.36 (t, J = 6.7 Hz, 1.1H), 4.25 (t, J = 5.6 Hz, 0.9H), 3.98 (s, 0.9H), 3.81 (s, 1.1H), 3.77 (dt, J = 13.4, 6.2 Hz, 2.0H), 3.56 (s, 1.1H), 3.13 (s, 0.9H), 2.40 (s, 1.6H), 2.39 (s, 1.4H). **Spectra consistent with rotamers in a 1:1 ratio.** **¹³C NMR** (101 MHz, CDCl₃) δ 172.51, 172.20, 171.75, 171.28, 161.38, 161.36, 154.10, 154.04, 144.59, 144.58, 140.37, 139.40, 134.44, 131.31, 131.11, 131.02, 131.00, 130.98, 130.90, 120.62, 120.54, 118.98, 118.96, 118.24, 115.41, 106.58, 106.56, 106.50, 106.07, 50.98, 50.36, 50.18, 49.08, 49.00, 48.86, 39.80, 38.44, 20.70, 20.68. **LCMS** (m/z): 419.2 [M+H]⁺, t_R = 3.13 min, standard method. **HRMS** (m/z):

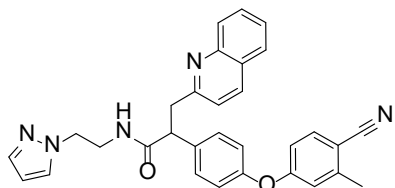
$C_{23}H_{22}N_4O_4$ requires 419.1714 $[M+H]^+$; found 419.1721. **HPLC**: t_R = 5.56 min, 99 %, standard method.

***N*-(2-(1H-pyrazol-1-yl)ethyl)-4-(2-chlorophenyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)butanamide (6.55)**



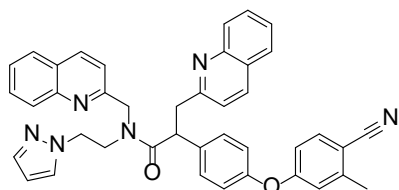
In an oven dried RBF cooled to room temperature was added NaH (60 % dispersion in mineral oil, (133 mg, 6.0 eq., 3.33 mmol) under N_2 . Excess mineral oil was removed by trituration with dry pet spirits (3 mL). DMF (4 mL) was added and the reaction cooled to 0 °C, 2-[4-(4-cyano-3-methylphenoxy)phenyl]-*N*-(2-pyrazol-1-ylethyl)acetamide (200 mg, 1.0 eq., 0.55 mmol) was added and the reaction stirred for 15 min at 0 °C. 1-(2-Bromoethyl)-2-chloro-benzene (146 mg, 1.2 eq., 0.66 mmol) was added and the reaction stirred for a further 30 min on ice. The ice bath was removed, the reaction warmed to room temperature slowly and heated at 60 °C for 6 h. After this time, the reaction was quenched with H_2O (15 mL) and diluted in EtOAc (20 mL). The layers were separated and the aqueous further extracted with EtOAc (20 mL), the combined organic layers were washed with H_2O (15 mL), $NaHCO_3$ (aq., sat., 2 × 20 mL), brine (10 mL), dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 50:50 pet spirits:EtOAc to 33:67 pet spirits:EtOAc to give **6.55** as a clear oil (2.3 mg, 0.83 %). **1H NMR** (401 MHz, CD_3OD) δ 7.59 (d, J = 8.5 Hz, 1H), 7.46 (d, J = 2.0 Hz, 1H), 7.42 – 7.36 (m, 2H), 7.37 – 7.29 (m, 2H), 7.26 – 7.11 (m, 3H), 7.08 – 6.98 (m, 2H), 6.94 (d, J = 2.6 Hz, 1H), 6.85 (dd, J = 8.6, 2.5 Hz, 1H), 6.17 (t, J = 2.2 Hz, 1H), 4.24 (t, J = 5.9 Hz, 2H), 3.58 (t, J = 5.9 Hz, 2H), 3.48 (t, J = 7.6 Hz, 1H), 2.65 (t, J = 7.9 Hz, 2H), 2.45 (s, 3H), 2.30 (ddd, J = 15.6, 8.2, 6.3 Hz, 1H), 2.12 – 1.98 (m, 1H). **LCMS** (m/z): 499.2 $[M+H]^+$, t_R = 4.27 min, standard method. **HRMS** (m/z): $C_{29}H_{27}ClN_4O_2$ requires 499.1895 $[M+H]^+$; found 499.1909. **HPLC**: t_R = 7.12 min, 85 %, standard method.

***N*-(2-(1H-pyrazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)-3-(quinolin-2-yl)propenamide (6.56)**



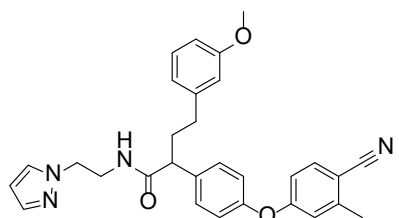
In an oven dried RBF cooled to room temperature was added NaH (60 % dispersion in mineral oil, 33 mg, 6.0 eq., 0.83 mmol) under N₂. Mineral oil was removed by trituration with dry pet spirits. DMF (4 mL) was added and the reaction cooled to 0 °C, 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]-N-(2-pyrazol-1-ylethyl)acetamide (50 mg, 1.0 eq., 0.14 mmol) was added and the reaction stirred for 15 min at 0 °C. 2-(Bromomethyl)quinolone (46 mg, 1.5 eq., 2.1 mmol) was added and the reaction stirred for a further 30 min on ice. The ice bath was removed, the reaction warmed to room temperature slowly and heated at 45 °C for 4.5 h. After this time, the reaction was quenched with H₂O (15 mL) and diluted in EtOAc (20 mL). The layers were separated and the aqueous further extracted with EtOAc (20 mL), the combined organic layers were washed with NaHCO₃ (aq., sat., 2 × 20 mL), brine (10 mL), dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 25:75:0 pet spirits:EtOAc:MeOH to 0:90:10 pet spirits:EtOAc:MeOH to give **6.56** as a clear oil (18 mg, 5 %). ¹H NMR (401 MHz, CDCl₃) δ 8.01 (d, *J* = 8.5 Hz, 1H), 7.97 (dd, *J* = 8.5, 1.1 Hz, 1H), 7.77 (dd, *J* = 8.1, 1.4 Hz, 1H), 7.68 (ddd, *J* = 8.4, 6.9, 1.5 Hz, 1H), 7.53 – 7.47 (m, 2H), 7.43 – 7.36 (m, 3H), 7.23 (d, *J* = 8.4 Hz, 1H), 7.00 – 6.93 (m, 3H), 6.85 – 6.82 (m, 1H), 6.77 (ddd, *J* = 8.5, 2.5, 0.6 Hz, 1H), 6.61 – 6.53 (m, 1H), 6.02 (t, *J* = 2.1 Hz, 1H), 4.27 (dd, *J* = 9.0, 5.9 Hz, 1H), 4.18 – 4.02 (m, 2H), 3.81 (dd, *J* = 14.6, 9.0 Hz, 1H), 3.68 – 3.51 (m, 2H), 3.35 (dd, *J* = 14.6, 6.0 Hz, 1H), 2.47 (s, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 173.32, 161.36, 159.60, 154.23, 147.67, 144.56, 139.91, 136.50, 136.41, 134.43, 129.92, 129.74, 129.71, 128.66, 127.72, 126.99, 126.24, 122.45, 120.53, 119.01, 118.26, 115.46, 106.61, 105.45, 51.52, 50.82, 42.20, 40.02, 20.73. LCMS (*m/z*): 499.8 [M+H]⁺, *t_R* = 3.95 min, standard method. HRMS (*m/z*): C₃₁H₂₇N₅O₂ requires 502.2238 [M+H]⁺; found 502.2244. HPLC: *t_R* = 5.52 min, >99 %, standard method.

***N*-(2-(1H-pyrazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)-3-(quinolin-2-yl)-N-(quinolin-2-ylmethyl)propenamide (6.57)**



Product **6.57** isolated from the above reaction as a clear oil (115 mg, 32 %). **¹H NMR** (401 MHz, CDCl₃) δ 8.02 (t, *J* = 8.9 Hz, 1H), 7.86 (t, *J* = 7.3 Hz, 1.4H), 7.78 (d, *J* = 8.0 Hz, 1.6H), 7.68 – 7.58 (m, 3.0H), 7.60 – 7.54 (m, 1.6H), 7.54 – 7.43 (m, 3.7H), 7.40 – 7.32 (m, 2.4H), 7.29 (d, *J* = 8.4 Hz, 0.4H), 7.12 (d, *J* = 2.3 Hz, 0.4H), 7.06 – 7.00 (m, 0.8H), 6.91 – 6.84 (m, 2.0H), 6.83 (d, *J* = 2.4 Hz, 0.6H), 6.83 – 6.76 (m, 1.0H), 6.71 (d, *J* = 2.3 Hz, 0.5H), 6.69 (dd, *J* = 8.6, 2.5 Hz, 0.6H), 6.10 (t, *J* = 2.1 Hz, 0.4H), 5.90 (t, *J* = 2.1 Hz, 0.6H), 4.98 (dd, *J* = 10.3, 4.5 Hz, 0.4H), 4.80 – 4.72 (m, 1.0H), 4.63 – 4.52 (m, 1.5H), 4.55 – 4.45 (m, 1.5H), 4.35 (d, *J* = 17.7 Hz, 0.4H), 4.32 – 4.17 (m, 1.6H), 4.10 – 3.98 (m, 0.5H), 3.99 – 3.84 (m, 1.9H), 3.77 – 3.62 (m, 0.6H), 3.35 (dd, *J* = 15.0, 4.5 Hz, 0.4H), 3.22 (dd, *J* = 14.2, 4.4 Hz, 0.6H), 2.48 (s, 1.8H), 2.47 (s, 1.2H). **Spectra consistent with rotamers in a 1:1 ratio.** **¹³C NMR** (101 MHz, CDCl₃) δ 173.78, 173.08, 161.20, 161.16, 160.06, 159.95, 157.79, 157.05, 154.30, 154.09, 147.66, 147.53, 147.46, 147.34, 144.51, 144.48, 140.25, 139.92, 136.87, 136.66, 136.35, 136.12, 135.98, 134.38, 134.35, 130.04, 129.98, 129.91, 129.87, 129.81, 129.53, 129.50, 129.45, 128.94, 128.84, 128.63, 128.45, 127.73, 127.64, 127.56, 127.52, 127.25, 127.12, 127.00, 126.95, 126.46, 126.31, 126.07, 126.05, 122.68, 122.60, 120.72, 120.41, 120.06, 119.07, 119.04, 118.20, 118.17, 115.45, 115.26, 106.64, 106.61, 105.71, 105.37, 54.59, 53.51, 52.66, 50.50, 49.56, 48.89, 48.83, 47.75, 46.61, 44.09, 43.92, 41.11, 24.09, 20.68. **LCMS** (*m/z*): 644.0 [M+H]⁺, *t_R* = 4.13 min, standard method. **HRMS** (*m/z*): C₄₁H₃₄N₆O₂ requires 643.2816 [M+H]⁺; found 643.2813. **HPLC**: *t_R* = 5.77 min, >99 %, standard method.

***N*-(2-(1H-pyrazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)-4-(3-methoxyphenyl)butanamide (6.58)**



To a solution of 2-[4-(4-cyano-3-methyl-phenoxy)phenyl]-N-(2-pyrazol-1-ylethyl)acetamide (50 mg, 1.0 eq., 0.14 mmol) was added LDA (1 M in THF, 0.21 mL, 1.5 eq., 0.21 mmol) dropwise at 0 °C.

The reaction was stirred on ice for 1h then slowly warmed to room temperature and stirred for a further 5 h. Additional LDA (1 M in THF, 0.4 mL, 3.0 eq., 0.42 mmol) was added and the reaction stirred at room temp for 1h. After this time, the reaction was diluted in EtOAc (20 mL) and quenched with H₂O (20 mL). Layers were separated and the aqueous further extracted with EtOAc (20 mL). The combined organic layers were washed with H₂O (20 mL), brine (10 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using 98:2 DCM:MeOH to give **6.58** as a colourless oil (3.54 mg, 5.2 %). ¹H NMR (401 MHz, CDCl₃) δ 7.46 (d, *J* = 8.6 Hz, 1H), 7.40 (dd, *J* = 1.9, 0.7 Hz, 1H), 7.24 – 7.20 (m, 2H), 7.15 – 7.09 (m, 2H), 6.96 – 6.89 (m, 2H), 6.81 (dt, *J* = 2.5, 0.6 Hz, 1H), 6.74 (ddd, *J* = 8.6, 2.5, 0.6 Hz, 1H), 6.70 – 6.63 (m, 2H), 6.64 – 6.59 (m, 1H), 6.12 (t, *J* = 2.1 Hz, 1H), 6.06 (t, *J* = 5.7 Hz, 1H), 4.25 – 4.05 (m, 2H), 3.72 (s, 3H), 3.60 (qd, *J* = 5.6, 1.3 Hz, 2H), 3.21 (t, *J* = 7.4 Hz, 1H), 2.51 – 2.36 (m, 6H), 2.06 – 1.93 (m, 1H). LCMS (*m/z*): 495.2 [M+H]⁺, *t*_R = 4.19 min, standard method. HRMS (*m/z*): C₃₀H₃₀N₄O₃ requires 495.2391 [M+H]⁺; found 495.2407. HPLC: *t*_R = 6.83 min, >95 %, standard method.

8.3.1 Parallel synthesis plate setup

In a glass coated 96-well plate, NaH (60 % dispersion in mineral oil, 10 – 20 mg) was scooped each well under N₂. pet spirits (100 μL) was added to these wells, mixed and decanted with a micro pipette. Additional pet spirits (100 μL) was added and decanted again, then *N*-(2-(1H-pyrazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)acetamide (**6.22**) (65 mM, 80 μL 1.0 eq., 5 μmol) dissolved in dry DMF was added and the reaction stirred for 15 min at room temperature. After this time, alkyl halide (500 mM, 20 μL, 2.0 eq., 10 μmol) was added and the reaction stirred for a further 3 h at room temperature. Reactions were quenched with AcOH (aq., 50 %, 50 μL) and the solvent removed under vacuum. Some wells required Boc deprotection or ester hydrolysis using 1:1 DCM:TFA (100 μL) or THF (50 μL) and NaOH (aq., 0.6 M, 50 μL, 3.0 eq., 30 μmol). The plate was covered and left at room temperature for a further 3 h without agitation. The solvent was then removed under vacuum and 50

μL of DMSO added to each well for a final concentration of 100 mM. A 1 μL aliquot was taken from each well and diluted in MeCN (49 μL) to 2 mM and analysed on LCMS 1. The remaining DMSO samples was stored at -78 °C until crystallographic soaking.

8.4 Experimental methods for assessing binding

Please see REFIL_X supplementary information (appendix 5.0) for X-ray crystallography data processing procedures.

8.4.1 Extraction and purification of Double His-Tagged *EcDsbA* (*EcDsbA*-DHT)

Frozen cell pellets were thawed and resuspended in an equal volume of 50 mM HEPES (pH 8.0), 150 mM NaCl, 10 mM imidazole supplemented with complete, EDTA-free Protease Inhibitor Cocktail tablets (Roche) and the cells were disrupted by sonication (10 × 30 s, 50 % duty cycle). After centrifugation at 50000×g at 4 °C for 30 min, the supernatant was 0.45 μm syringe filtered (Millex, Millipore) and loaded onto a HisTrap HP 5 mL column (GE Healthcare). After a 25 mM imidazole column wash to elute impurities, the target protein was eluted using a gradient from 25 – 500 mM imidazole. Fractions were analysed by SDS-PAGE and those containing target protein were pooled and concentrated to 10 mL using an Amicon centrifugal diafiltration unit (3000 MWCO, Millipore), buffer exchanged to 50 mM HEPES (pH 6.8), 50 mM NaCl using a HiPrep 2610 desalting column (GE Healthcare) and loaded onto a MonoQ HR 10/10 column (GE Healthcare). In these buffer conditions, the target protein eluted in the flowthrough fraction and impurities were eluted from the anion exchange column using a gradient from 0 – 1 M NaCl.

8.4.2 SPR analysis

Fragment binding to *EcDsbA* was measured using a BIAcore S200 (GE Healthcare). Double- His-tagged (DHT-) *EcDsbA* was immobilised on the surface of an NTA sensor chip, according to manufacturer's instructions. Following 3 injections 90 s each of 0.5 M EDTA, the chip surface was

activated with a 120 s injection of 5 mM NiCl₂, followed by a 360 s injection of 10 μM DHT-*EcDsbA*. The final immobilization level of DHT-*EcDsbA* was typically 19000-21000 RU. Binding assays were carried out at 20 °C using a flow-rate of 50 μL/min in running buffer (12 mM Phosphate, 137 mM NaCl, pH 7.4, 50 μM EDTA, 0.001 % Tween20, 2 % DMSO-*d*₆). 100 mM fragment stocks were diluted in running buffer to obtain 400 μM working concentrations, and from this a concentration series of 2-fold dilutions was prepared for dose-response studies. Raw sensorgram data were reduced, solvent-corrected, and double-referenced using BIAEvaluation Software (GE Healthcare). Dose-response determinations were evaluated using 1:1 steady-state binding model. *K*_d values were determined by BIAEvaluation software assuming 1:1 binding; data and fit are plotted in GraphPad Prism for clarity. Errors were determined as the Standard Error in the dose-response fit, as calculated by Biacore S200 Evaluation Software.

8.4.3 Crystallisation and X-ray diffraction experiments

The *EcDsbA*-compound complexes were prepared by crystal soaking using 10 mM final concentration of an inhibitor to maximise occupancy of the binding site. The previously published crystallization conditions of oxidized *EcDsbA* were used as a guide for obtaining diffraction quality crystals.²⁴ Briefly, protein was crystallised using hanging drop vapour diffusion. 1 μL of 30 mg/mL *EcDsbA* was mixed with an equal volume of crystallization buffer (11-13 % PEG 8000, 5-7.5 % glycerol, 1 mM CuCl₂, 100 mM sodium cacodylate pH 6.1) and equilibrated against 0.5 mL of reservoir buffer at 18°C. Typically, crystals appeared within one to three days and grew to average dimensions of 0.6 mm × 0.4 mm × 0.2 mm. For soaking experiments, crystals were transferred into 2 μL drops of 24 % PEG 8000, 22 % glycerol, 100 mM sodium cacodylate pH 6.1 containing a compound of interest at the final concentration of 10 mM (2-5 % of DMSO-*d*₆) and incubated for 2 hours. Crystals were mounted on loops and flash-cooled in liquid nitrogen.

8.4.4 Assessing solubility of synthesised analogues by q-NMR

Purified compounds had their solubility and aggregation evaluated via a 1D ^1H -qNMR titration in aqueous buffer. Compounds were accurately weighed and dissolved in d_6 -DMSO at 100 mM. D_2O buffer was prepared with 50 mM sodium phosphate, 25 mM NaCl at pH 7.4 with 100 μM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) internal standard. Serial dilution of the d_6 -DMSO stock was used to prepare a 4-point two-fold dilution series with final concentrations of 125, 250, 500 and 1000 μM with 2 % d_6 -DMSO in phosphate buffer and a total volume of 600 μL . NMR spectra were recorded on a Bruker 400 MHz spectrometer with $\text{TD} = 32 \text{ K}$, $\text{D1} = 5 \text{ s}$, 64 scans and water suppression at 4.7 ppm. NMR spectra were processed in MNOVA and referenced to the DSS peak at 0.030 ppm. Compounds failed solubility at a given concentration if they did not show a doubling in intensity concordant with doubling concentration or showed any chemical shift changes $>0.004 \text{ ppm}$ (1.6 Hz) from the lowest recorded concentration. Ligand concentrations were calculated relative to the DSS integral in the 1D ^1H -qNMR spectra and used to calculate the stock concentration which was used in K_D determination experiments.

8.4.5 Validation of hits and affinity determination by ^1H ^{15}N -HSQC

Hits from the REFIL_X screen were further validated by measuring CSP in ^1H - ^{15}N HSQC NMR. A reference ^1H - ^{15}N HSQC spectrum of uniformly labelled ^{15}N oxidised *EcDsbA* (100 μM protein, 2% d_6 -DMSO, 50 mM sodium phosphate, 25 mM NaCl, pH 6.8 in 10 % D_2O , 90% H_2O) was acquired. A series of ^1H - ^{15}N HSQC spectra were acquired under the same conditions in the presence of a sample fragment at and below its maximum solubility. CSP induced by fragment binding were calculated using the following equation

$$\text{CSP} = \sqrt{\Delta H^2 + (0.2 \times \Delta N)^2}$$

where ΔH and ΔN are the measured change in chemical shift of the backbone amide protons in the absence (*apo*) and presence of the fragment.

Backbone amide assignments were taken from the biological magnetic resonance data bank with a BMRB code of 19838. All assigned peaks showing >0.040 ppm shift within the ^1H - ^{15}N HSQC spectra at their highest concentration were used for affinity determination. CSP below 0.023 ppm shift, ambiguous merging peaks, poorly fitting residues and residues with less than 4 data points were excluded. Affinities were determined using the following equation where $[P]$ is protein concentration and $[L]$ is ligand concentration.

$$Y = \frac{dmax}{2 \times [P]} \times ([P] + [L] + K_D) - \sqrt{([P] + [L] + K_D)^2 - 4 \times [P] \times [L]}$$

8.5.6 Computational docking

Compounds were prepared for docking using Ligprep extension in Maestro (version 11.8.12) where ionisation, stereoisomers and 3D conformers were enumerated and sampled for each ligand as input for docking. The protein model (parent diaryl ether **1.27** crystal structure) was prepared using the protein preparation wizard extension in Maestro where steric clashes and missing side chains were corrected followed by removal of non-structural waters, optimisation of the hydrogen bonding network and restrained minimisation of the 3D structure using default settings. Compounds were docked using Glide (version 68014) XP docking with default parameters. Docked structures were then manually analysed.

8.5 References

1. Tashiro, M.; Itoh, T.; Fukata, G. Cyclodienones. X. Reaction of halo-cyclohexadien-1-ones with phenols in the presence of α -picoline and preparation of 4-hydroxy- and 2-hydroxyphenyl aryl ethers. *Bull. Chem. Soc. Jpn.* **1984**, 57, 416-420.
2. Naidu, A. B.; Jaseer, E. A.; Sekar, G. General, Mild, and Intermolecular Ullmann-Type Synthesis of Diaryl and Alkyl Aryl Ethers Catalyzed by Diol-Copper(I) Complex. *J. Org. Chem.* **2009**, 74, 3675-3679.
3. Mitchell, L. H.; Hu, L.; Nguyen, M.; Fakhoury, S.; Smith, Y.; Iula, D.; Kostlan, C.; Carroll, M.; Dettling, D.; Du, D.; Pocalyko, D.; Wade, K.; Lefker, B. Diphenyl ethers as androgen receptor antagonists for the topical suppression of sebum production. *Bioorg. Med. Chem. Lett.* **2009**, 19, 2176-2178.
4. Coombes, C. L.; Moody, C. J. First Syntheses of 2,2-Dimethyl-7-(2'-methylbut-3'-en-2'-yl)-2H-chromen-6-ol and 2-(3'-Methylbut-2'-enyl)-5-(2'-methylbut-3'-en-2'-yl)-1,4-benzoquinone, Novel Prenylated Quinone Derivatives from the New Zealand Brown Alga *Perithalia capillaris*. *J. Org. Chem.* **2008**, 73, 6758-6762.
5. Sawyer, J. S.; Schmittling, E. A.; Palkowitz, J. A.; Smith, W. J. Synthesis of diaryl ethers, diaryl thioethers, and diarylamines mediated by potassium fluoride-alumina and 18-crown-6: expansion of scope and utility. *J. Org. Chem.* **1998**, 63, 6338-6343.
6. Łysek, R.; Schütz, C.; Vogel, P. (1S,2S,3R,6R)-6-Aminocyclohex-4-ene-1,2,3-triol (= (-)-conduramine B-1) is a selective inhibitor of α -mannosidases. Its inhibitory activity is enhanced by N-benylation. *Helv. Chim. Acta* **2005**, 88, 2788-2811.
7. Nishizawa, R.; Nishiyama, T.; Hisaichi, K.; Hirai, K.; Habashita, H.; Takaoka, Y.; Tada, H.; Sagawa, K.; Shibayama, S.; Maeda, K.; Mitsuya, H.; Nakai, H.; Fukushima, D.; Toda, M. Discovery of orally available spirodiketopiperazine-based CCR5 antagonists. *Bioorg. Med. Chem. Lett.* **2010**, 18, 5208-5223.
8. Jung, N.; Bräse, S. Diaryl ether and diaryl thioether syntheses on solid supports via copper (I)-mediated coupling. *J. Comb. Chem.* **2009**, 11, 47-71.
9. Murthy, S.; Desantis, J.; Verheugd, P.; Maksimainen, M. M.; Venkannagari, H.; Massari, S.; Ashok, Y.; Obaji, E.; Nkizinkinko, Y.; Lüscher, B.; Tabarrini, O.; Lehtiö, L. 4-(Phenoxy) and 4-(benzyloxy)benzamides as potent and selective inhibitors of mono-ADP-ribosyltransferase PARP10/ARTD10. *Eur. J. Med. Chem.* **2018**, 156, 93-102.
10. Haga, N.; Takayanagi, H. Mechanisms of the photochemical rearrangement of diphenyl ethers. *J. Org. Chem.* **1996**, 61, 735-745.
11. Breyholz, H.; Schäfers, M.; Wagner, S.; Hölteke, C.; Faust, A.; Rabeneck, H.; Levkau, B.; Schober, O.; Kopka, K. C-5-Disubstituted barbiturates as potential molecular probes for noninvasive matrix metalloproteinase imaging. *J. Med. Chem.* **2005**, 48, 3400-3409.
12. Breyholz, H.-J.; Schäfers, M.; Wagner, S.; Hölteke, C.; Faust, A.; Rabeneck, H.; Levkau, B.; Schober, O.; Kopka, K. C-5-Disubstituted Barbiturates as Potential Molecular Probes for Noninvasive Matrix Metalloproteinase Imaging. *J. Med. Chem.* **2005**, 48, 3400-3409.
13. Samaritoni, J. G.; Arndt, L.; Bruce, T. J.; Dripps, J. E.; Gifford, J.; Hatton, C. J.; Hendrix, W. H.; Schoonover, J. R.; Johnson, G. W.; Hegde, V. B.; Thornburgh, S. N-Alkyl-N-(5-isothiazolyl)- and N-(Alkylisothiazolin-5-ylidene)-phenylacetamides. Synthesis and biological activity. *J. Agric. Food Chem.* **1997**, 45, 1920-1930.
14. Wood, P. M.; Woo, L. W. L.; Labrosse, J.; Trusselle, M. N.; Abbate, S.; Longhi, G.; Castiglioni, E.; Lebon, F.; Purohit, A.; Reed, M. J.; Potter, B. V. L. Chiral aromatase and dual aromatase-steroid sulfatase inhibitors from the letrozole template: synthesis, absolute configuration, and in vitro activity. *J. Med. Chem.* **2008**, 51, 4226-4238.

15. Duplais, C.; Bures, F.; Sapountzis, I.; Korn, T. J.; Cahiez, G.; Knochel, P. An Efficient Synthesis of Diaryl Ketones by Iron-Catalyzed Arylation of Aroyl Cyanides. *Angew. Chem. Int. Ed. Engl.* **2004**, 43, 2968-2970.
16. Chiang, Y.; Kolmakov, K.; Kresge, A. J. Ketonization of the unusually acidic elongated enol generated by flash photolytic decarboxylation of p-formylphenylacetic acid in aqueous solution. *Can. J. Chem.* **2008**, 86, 101-104.
17. Shonberg, J.; Herenbrink, C. K.; López, L.; Christopoulos, A.; Scammells, P. J.; Capuano, B.; Lane, J. R. A structure–activity analysis of biased agonism at the dopamine D2 receptor. *J. Med. Chem.* **2013**, 56, 9199-9221.
18. Mistry, S. N.; Shonberg, J.; Draper-Joyce, C. J.; Klein Herenbrink, C.; Michino, M.; Shi, L.; Christopoulos, A.; Capuano, B.; Scammells, P. J.; Lane, J. R. Discovery of a Novel Class of Negative Allosteric Modulator of the Dopamine D2 Receptor Through Fragmentation of a Bitopic Ligand. *J. Med. Chem.* **2015**, 58, 6819-6843.
19. Perrone, M. G.; Vitale, P.; Panella, A.; Ferorelli, S.; Contino, M.; Lavecchia, A.; Scilimati, A. Isoxazole-Based-Scaffold Inhibitors Targeting Cyclooxygenases (COXs). *ChemMedChem* **2016**, 11, 1172-1187.
20. Romulus, J.; Weck, M. Single-Chain Polymer Self-Assembly Using Complementary Hydrogen Bonding Units. *Macromol. Rapid Commun.* **2013**, 34, 1518-1523.
21. Varala, R.; Nuvula, S.; Adapa, S. R. Molecular Iodine-Catalyzed Facile Procedure for N-Boc Protection of Amines. *J. Org. Chem.* **2006**, 71, 8283-8286.
22. Goldup, S. M.; Leigh, D. A.; McBurney, R. T.; McGonigal, P. R.; Plant, A. Ligand-assisted nickel-catalysed sp³–sp³ homocoupling of unactivated alkyl bromides and its application to the active template synthesis of rotaxanes. *Chem. Sci.* **2010**, 1, 383-386.
23. Mohammat, M. F.; Shaameri, Z.; Hamzah, A. S.; Fun, H.; Chantrapromma, S. 4-Hydroxy-5-(4-methoxyphenyl)pyrrolidin-2-one. *Acta Crystallogr. E* **2008**, 64, o578-o579.
24. Adams, L. A.; Sharma, P.; Mohanty, B.; Ilyichova, O. V.; Mulcair, M. D.; Williams, M. L.; Gleeson, E. C.; Totsika, M.; Doak, B. C.; Caria, S.; Rimmer, K.; Horne, J.; Shouldice, S. R.; Vazirani, M.; Headey, S. J.; Plumb, B. R.; Martin, J. L.; Heras, B.; Simpson, J. S.; Scanlon, M. J. Application of fragment-based screening to the design of inhibitors of Escherichia coli DsbA. *Angew. Chem. Int. Ed. Engl.* **2015**, 54, 2179-84.

Appendices



Applications of NMR Spectroscopy in FBDD

111

Matthew Bentley, Bradley C. Doak, Biswaranjan Mohanty, and Martin J. Scanlon

Contents

Introduction	2212
Library Design and Quality Control	2213
Library Screening	2215
Mixture Design for NMR	2215
¹ H Ligand-Detect Mixtures/Singleton/Competition	2217
¹⁹ F Detection	2219
Protein Detect	2219
Affinity Determination	2221
Protein-Detected Approaches	2221
Ligand-Detected Approaches	2225
Structure Modeling	2225
Ligand-Observed NMR	2226
Protein-Observed NMR	2227
Chemical Shift Perturbation Mapping	2227
NOE-Based Protein-Ligand Complexes	2227
NMR ²	2228
Concluding Remarks	2228
References	2229

Abstract

Fragment-based drug design (FBDD) has become firmly established as a viable approach to the identification of starting points for the development of potent and selective compounds that modulate protein activity. As of 2017, the United States Food and Drug Administration have approved two molecules derived from FBDD for therapeutic use, many more are in advanced clinical trials, and the

M. Bentley · B. C. Doak · B. Mohanty · M. J. Scanlon (✉)
 Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Australia
 e-mail: matthew.bentley@monash.edu; bradley.doak@monash.edu; biswaranjan.mohanty@monash.edu; martin.scanlon@monash.edu

© Springer International Publishing AG, part of Springer Nature 2018
 G. A. Webb (ed.), *Modern Magnetic Resonance*,
https://doi.org/10.1007/978-3-319-28388-3_127

2211

technology has been embraced by both academia and industry. The starting point for FBDD is the identification of very small molecules – “fragments” – that bind to a protein of interest. Due to their small size, fragments are able to sample chemical space more efficiently than larger molecules. This increases the likelihood of finding a “hit” – i.e., a fragment that binds to the desired protein. However, their small size also dictates that a fragment is likely to bind to its target protein with low affinity. For this reason, screening is generally carried out using biophysical binding assays rather than biochemical activity assays. Nuclear magnetic resonance (NMR) spectroscopy is an extremely powerful approach for detecting weak interactions. In fact the first implementation of FBDD employed NMR to characterize binding, and NMR remains a mainstay of many FBDD screening campaigns. NMR has a broader application in supporting programs of FBDD – it provides an essential component for quality control of the compounds in fragment screening libraries, it can be used to assess solubility, aggregation, and in addition to its role in screening to find fragment hits, it can be used to rank hits and assess their suitability for crystallographic structure determination in complex with a target protein. Where crystallography is not possible, several NMR-based approaches have been developed to determine structures of fragment-protein structures. In the current chapter we review these myriad applications of NMR in FBDD.

Keywords

Compound aggregation · Library screening · Structure modeling · Ligand observed NMR · Mixture design · Protein detection · WaterLOGSY spectra · Fragment-based drug design

Introduction

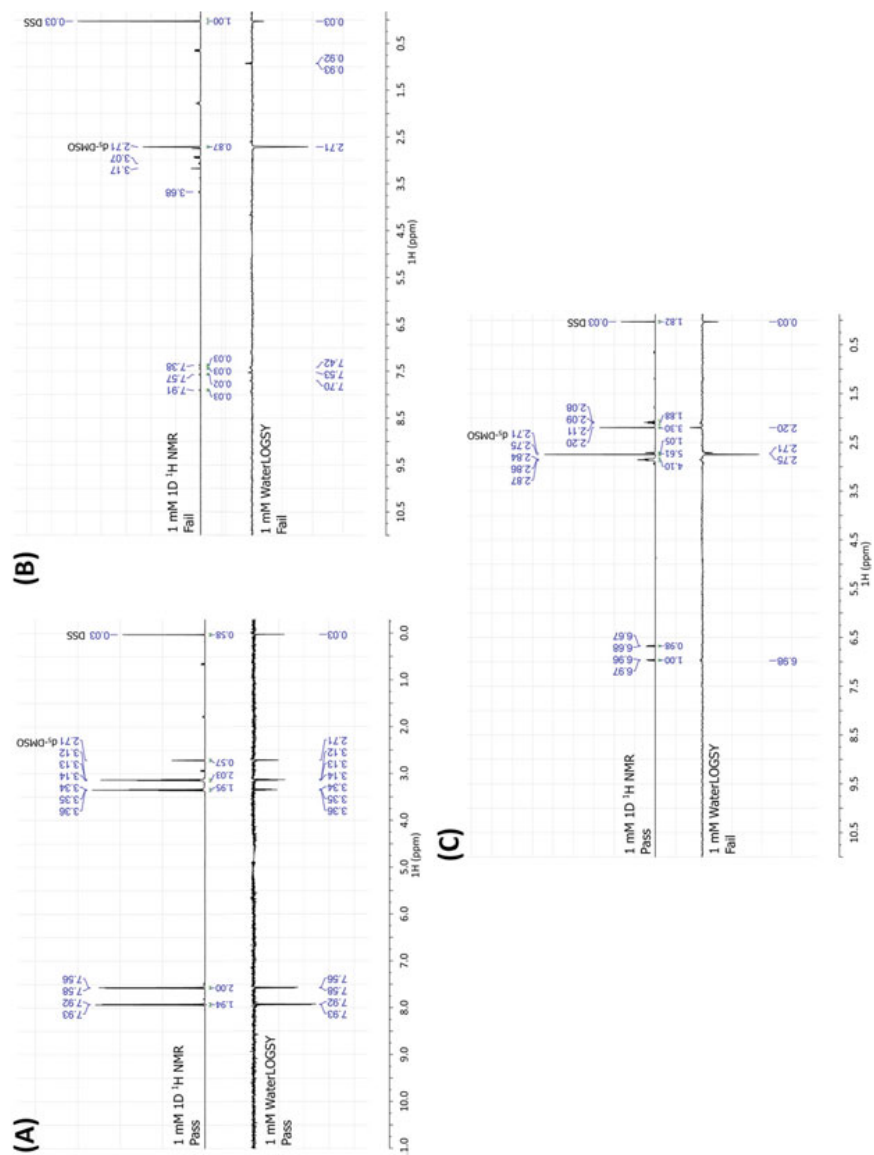
The first practical demonstration of fragment-based drug design (FBDD) was reported in a landmark paper by researchers from Abbott Laboratories [1]. The paper described the design and synthesis small molecules that bound with high affinity to FK506 binding protein. Perhaps more importantly, it also provided an outline of the various steps that are required in the process of implementing FBDD. Although many different experimental approaches to the practice of FBDD have been developed in the intervening period [2], the fundamental steps described in the original paper have been broadly conserved. These involve assembling a library of fragments, screening the library against the target to identify hits, and then implementing a program of medicinal chemistry to optimize the hits and improve their affinity and potency. This latter step is almost always supported by structural data describing the mode of binding of a fragment to its target.

Nuclear magnetic resonance (NMR) spectroscopy has been utilized in different ways in each of these stages of FBDD. The process of fragment library design generally involves selecting a diverse set of fragments that possess a desired profile of physicochemical properties and has been described and reviewed in detail

elsewhere [3–8]. An important requirement is that the fragments are of suitable purity, and that they are soluble and do not aggregate under the conditions that are to be used for screening. NMR provides a robust means to determine compound identity and purity, it can be used to determine concentration in solution [9] and monitor aggregation [10]. NMR is also widely used for screening fragment libraries. In a series of polls conducted through the *Practical Fragments* blog over several years, NMR spectroscopy was consistently identified as one of the most widely used techniques for finding hits. NMR binding assays are usually segregated into protein-detected approaches – such as the one employed by Abbott researchers in their original paper [1], or ligand-detected approaches, where changes in the spectroscopic properties of the ligand are used to infer binding [11–13]. Most commonly, ligand-detected experiments are performed with ^1H detection although ^{19}F -NMR can also be used [14]. These various NMR-based methods can also be used to identify other relevant information including the binding-site of fragments on the protein target, the affinity of the binding interaction, and the likelihood of obtaining a crystal structure of the fragment-protein complex [15–17]. X-ray crystallography remains the preferred approach for generating structural data for fragments bound to their target due to considerations of speed and resolution, although where crystallography is not feasible, NMR can be used to generate structural data (exemplified in [1]), and a number of approaches have been described to improve the throughput and ease of NMR-based structure determinations [18, 19]. In the current chapter we will review the application of several of these NMR-based approaches in the implementation of FBDD.

Library Design and Quality Control

The design and assembly of a fragment library underpins any successful fragment-based screening campaign and generally involves selecting chemically diverse fragments with desired physicochemical profiles. For a general-purpose screening library the goals are typically to maximize diversity and coverage of chemical space while selecting fragments that are amenable to chemical elaboration to improve their affinity and potency [3, 7, 8]. Here we focus on some of the quality control (QC) measures that can be undertaken to assess the suitability of fragments for inclusion in a screening library. These methods are also applicable to fragment analogues and larger elaborated compounds that may be synthesized and tested during the elaboration of hits. Since fragments often bind their target protein with low affinity ($K_D \approx 100 \mu\text{M} - 10 \text{ mM}$) it is necessary to screen them at relatively high concentration (typically $100 \mu\text{M} - 1 \text{ mM}$ screening concentration). Therefore it is essential to confirm the fragment identity, purity, and solubility under the conditions used for screening [20]. Much of this information can be obtained via analysis of the aqueous 1D ^1H -NMR spectrum of a sample (Fig. 1). Identity and impurities are quickly identified by visual inspection of the spectra by an experienced medicinal chemist. Inclusion of an internal standard such as DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) and suitable acquisition parameters allows for quantitative calculation of the compound concentration/amount which can be compared to the expected



concentration to identify solubility, inorganic impurities, or incorrect stock concentrations that lead to observed low concentration samples [9, 21]. It has been reported that between 15 and 40% of compounds fail the QC process, with the majority of failures due to solubility/amount or aggregation (Fig. 1) [7, 8]. An additional problem can arise due to the fact that many small fragments have appreciable volatility, which can lead to sample loss. In such cases there may be no resonances observed in a 1D ^1H -NMR spectra. This can be distinguished from sample loss during liquid transfer by the presence of stock solvent peak/s, e.g., for fragments that are stored as a concentrated stock solution in $^2\text{H}_6$ -DMSO, the 1D ^1H -NMR spectrum of a dilute sample in aqueous buffer will contain a resonance for $^2\text{H}_5$ -DMSO, the intensity of which reflects the volume of the stock solution that was added to the sample. In addition to 1D ^1H -NMR it is common to record liquid chromatography-mass spectrometry (LCMS) data, which also helps to resolve ambiguities for fragments with complex 1D ^1H -NMR spectra or fragments that adopt multiple slowly exchanging conformations in solution.

Compound aggregation is a difficult property to calculate computationally; indeed databases of aggregators [22] are among the best resources available for the prospective assessment of aggregation propensity. Many small organic compounds are able to associate and form invisible aggregates in solution, which can often result in false positives on screening. These can be time consuming to follow up, sometimes difficult to identify, and hence costly during screening campaigns [20, 23]. Aggregation can be observed in 1D ^1H -NMR titrations both through concentration-dependent changes in chemical shift and via measurement of signal intensity, which does not increase proportionally with compound concentration when aggregates are formed [10]. Alternatively, water-ligand observed via gradient spectroscopy (WaterLOGSY) [11] NMR can be used to assess the rotational correlation time (τ_c) of compounds, where positive phase signals with respect to the water signal are indicative of larger, slower tumbling aggregates (Fig. 1).

Library Screening

Mixture Design for NMR

Ligand-observed NMR spectroscopy has endured as one of the most commonly used fragment screening methods due to its reliability, wide applicability, and sensitivity [8]. Typically fragments are screened via one or more of saturation-transfer difference (STD)-NMR, WaterLOGSY, T2 relaxation-weighted experiments such as the Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence or 1D ^{19}F -NMR spectroscopy [15, 21]. Ligand-observed NMR screening, for both ^1H -detected and ^{19}F -detected NMR spectra, is generally performed in mixtures, which provides greater throughput. In such cases the mixtures must be designed to enable straightforward analysis of which fragment(s) in the mixture is binding. Different approaches using simulated annealing and genetic algorithms to design optimized mixtures for fragment screening have been reported [24, 25], and this process is available in at least one commercially available

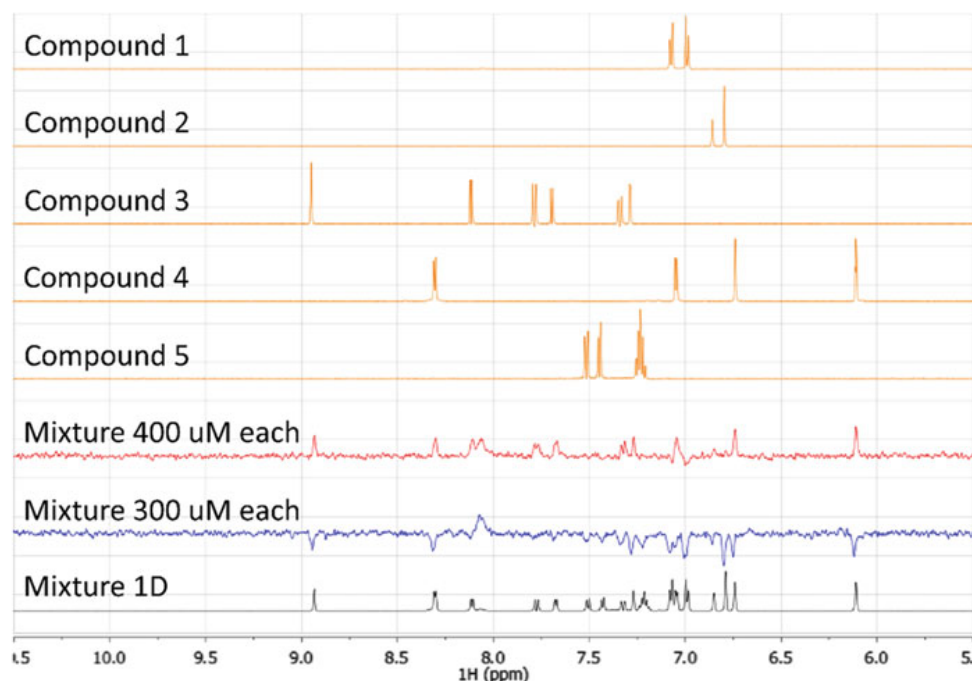


Fig. 2 Mixture design and the effect of concentration on mixture aggregation assessed by water-ligand observed via gradient spectroscopy (WaterLOGSY). 1D ^1H -NMR of the mixture (*black, bottom*), composed of five compounds (Compound 1–5, orange, top) and WaterLOGSY of the same mixture at 300 μM (*blue*) and 400 μM (*red*) of each compound

software (MNova, MestreLab Research, Santiago de Compostela, Spain). Here the aim is to create mixtures with minimal spectral overlap to facilitate reliable and unambiguous interpretation of ligand-detected NMR experiments via comparison with reference spectra of the individual compounds. Additionally, chemical incompatibilities (electrophiles and nucleophiles) as well as acid-base balance can also be taken into account to optimize mixture combinations [26]. A balance must be struck between increasing the efficiency of screening by having a high number of components in the mixture and increasing the chances of spectral overlap and/or mixture effects, which can complicate spectral analysis and may generate false positives. In order to deconvolute the binding components of the mixture, reference spectra of each compound are required and are commonly recorded during initial QC of the compounds. Mixtures of 5–15 compounds have been used for STD, CPMG, or WaterLOGSY at concentration of 250–500 μM for each compound [4, 21, 27]. It is necessary to optimize the concentration of fragments in the mixture even though they may be soluble at high concentration as single fragments, as mixture effects can lead to aggregation (Fig. 2). This can be demonstrated by recording WaterLOGSY spectra of the same mixture at different concentrations, e.g., Fig. 2 shows WaterLOGSY spectra of the same fragments in a mixture recorded at two different concentrations (300 and 400 μM). In this case aggregation is observed at the higher concentration.

¹H Ligand-Detect Mixtures/Singleton/Competition

Ligand-detected ¹H NMR experiments have the advantages of being applicable to almost any soluble protein, and unlike protein-detected experiments the quality of the spectra are far less sensitive to the size of the protein. The spectra are relatively quick to acquire and they are able to detect even very weakly binding compounds (K_D up to ≈ 10 mM). The ability to screen mixtures increases screening efficiency, and direct observation of the ligand NMR spectra during screening generates QC data to confirm the identity, purity, and concentration of each fragment at the time of screening. STD-NMR remains one of the most commonly used experiments for ligand-observed screening [13]. STD experiments observe differences in the intensity of ligand resonances in ¹H-NMR spectra that are acquired either with or without selective excitation of the protein. Selective excitation is possible as protein spectra generally contain resonances at frequencies (e.g., $\delta \leq 0$ ppm) where fragment signals are almost never observed. Following selective saturation of the protein (in the “on-resonance experiment”), spin-diffusion results in magnetization being transferred throughout the protein, from where it can also be transferred to bound fragments via the nuclear Overhauser effect (NOE). In the absence of this selective excitation (the “off-resonance experiment”), there is no NOE from protein to fragment, which results in a difference in the intensity of ligand signals in the on-resonance and off-resonance spectra (Fig. 3).

WaterLOGSY experiments also make use of intermolecular NOEs between the protein and fragment to detect binding, but in this case it is the bulk water in the sample that is the source of magnetization transfer [11, 28]. WaterLOGSY takes advantage of the difference in τ_c between small proteins that tumble rapidly and show positive NOEs (which are conventionally phased as negative peaks in the spectra) and larger proteins that tumble slowly and show negative NOEs (i.e., positive peaks in the spectra). For screening via WaterLOGSY two spectra are usually acquired – in the absence and presence of protein. Fragments are expected to show negative peaks in the absence of protein, which for fragments that bind become less negative or positive in the presence of the protein (Fig. 3). Noteworthy in WaterLOGSY are exchangeable protons (on the ligand or protein), which always show positive phase signals due to direct exchange of protons with the bulk water and should not be interpreted as binding or aggregation. WaterLOGSY also requires the presence of H₂O and therefore spectra cannot be acquired in 100% D₂O, which in turn requires effective water suppression in order to reliably detect small changes in the spectra. Standard water suppression schemes that are available on modern NMR spectrometers are generally sufficient and in practice ligand-detect NMR experiments are routinely conducted for samples in H₂O [15, 29].

Small molecules which tumble quickly in solution have long T₂ relaxation times, whereas larger proteins have short T₂ relaxation times [21]. Therefore, T₂ filtered experiments such as the Carr-Purcell-Meiboom-Gill (CPMG) [30, 31] can be used to identify fragments that bind to a protein via changes in their T₂ relaxation. Experiments are recorded in the absence and presence of protein, and fragments which bind to the protein show faster T₂ relaxation in the protein-containing spectrum, which is manifest as a reduction in the intensity of the fragment resonances in the spectrum (Fig. 3).

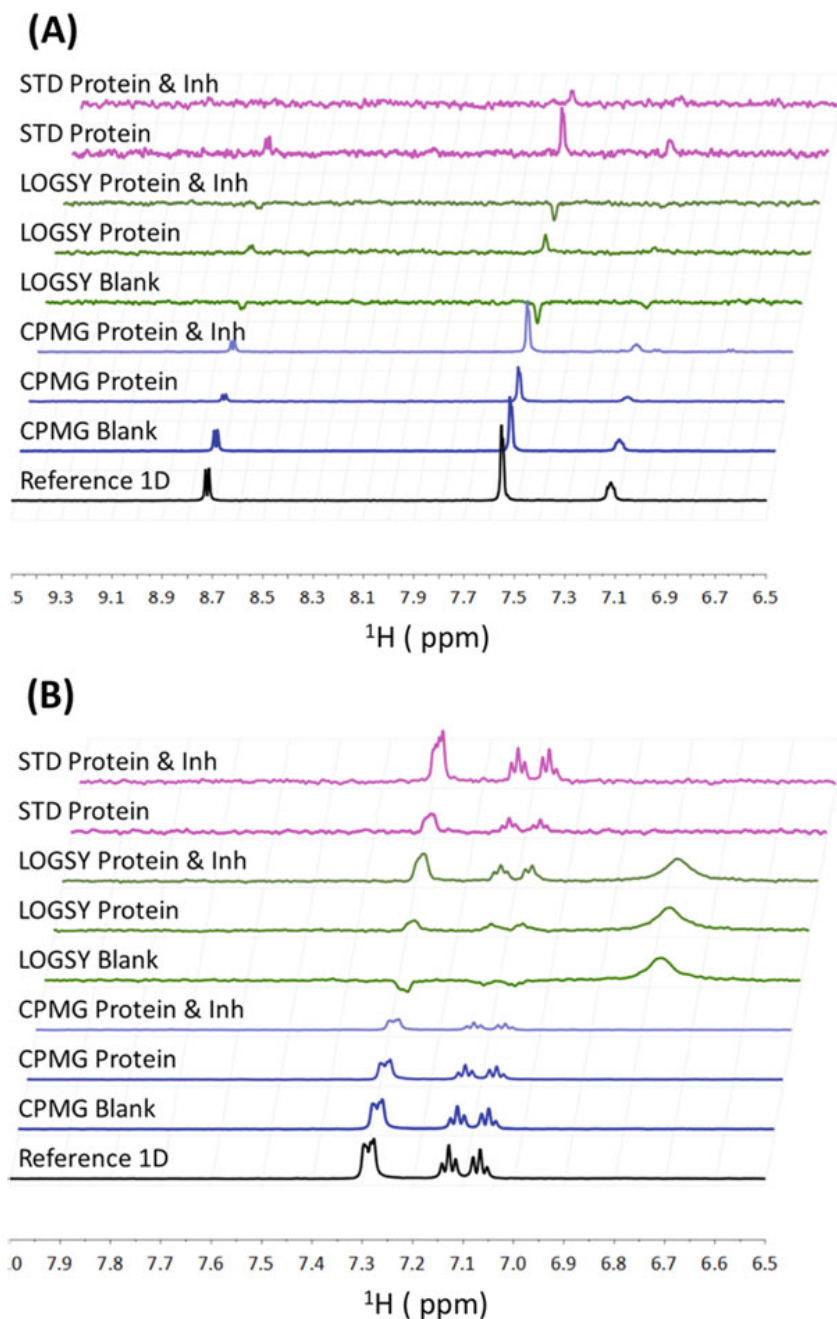


Fig. 3 Ligand-detected NMR analysis of fragment binding of a (a) competitive hit and (b) noncompetitive hit. Saturation transfer difference (STD) are shown in *purple*, Water-ligand observed via gradient spectroscopy (LOGSY) are shown in *green*, and Carr-Purcell-Meiboom-Gill (CPMG) are shown in *blue* with reference 1D ^1H -NMR shown in *black*. Spectra are labeled according to ligand in buffer as “Blank,” ligand with Protein target as “Protein,” and ligand with Protein and competitor/inhibitor as “Protein & Inh”

In each of the ligand-detected experiments above, competition with a known higher affinity ligand can be conducted to identify fragments that bind competitively at a specific site (Fig. 3). In the presence of inhibitor the intensity of the signal observed in the ligand-detected spectrum reverts to that observed for free ligand. Hence, the presence of a saturating concentration of a high-affinity competitor (usually $>10 \times K_D$) results in a loss of signal intensity in the STD-difference experiment, signals with more negative phase in the WaterLOGSY, and an increase in intensity in the CPMG experiment (Fig. 3).

This approach can be used to eliminate nonspecific binders such as aggregates and promiscuous ligands binding at multiple sites on the protein [2, 21, 32]; however, it may also result in genuine second-site binders being discarded.

An additional advantage of running all three of these ligand-detect experiments with and without competitor is that false positives are unlikely to arise in all three experiments due to the different mechanisms by which binding is detected. Additionally, validation by more techniques has been linked to successful structure determination by X-ray crystallography providing a means of prioritizing primary screening hits for further investigations [32].

¹⁹F Detection

¹⁹F-NMR is becoming increasingly popular in FBDD. ¹⁹F detection affords high sensitivity, enables quantitative analysis of protein-ligand binding and, due to the absence of ¹⁹F in natural biomolecules, provides background-free signal detection [12]. ¹⁹F-NMR can be used to observe proteins (into which ¹⁹F-containing residues can be incorporated) but is more commonly used as a means to characterize the binding of fragments that contain ¹⁹F [12, 33, 34]. One limitation of ¹⁹F detection is that optimal sensitivity requires a dedicated cryoprobe that is capable of ¹⁹F detection with ¹H decoupling, and although these are commercially available they are not yet in common usage. However, ¹⁹F detection provides a number of unique advantages. For example, membrane proteins almost invariably require additives including lipids and/or detergents to maintain their solubility and stability; these additives contain protons, which are often present at relatively high concentrations and therefore generate strong signals that interfere in many of the ¹H ligand-detected experiments that are commonly used in FBDD screening. In addition, the large chemical shift range that is observed for ¹⁹F in fragments means that signal overlap is rarely problematic (Fig. 4).

Protein Detect

In ligand-detected approaches, changes in the NMR spectrum of the ligand are monitored upon addition/saturation of protein. It is also possible to detect fragment binding by monitoring changes in the NMR spectrum of the protein upon addition of ligand. This was the approach taken in the original SAR-by-NMR method to identify ligand binding [1]. Protein-detected methods have the advantage that they can also

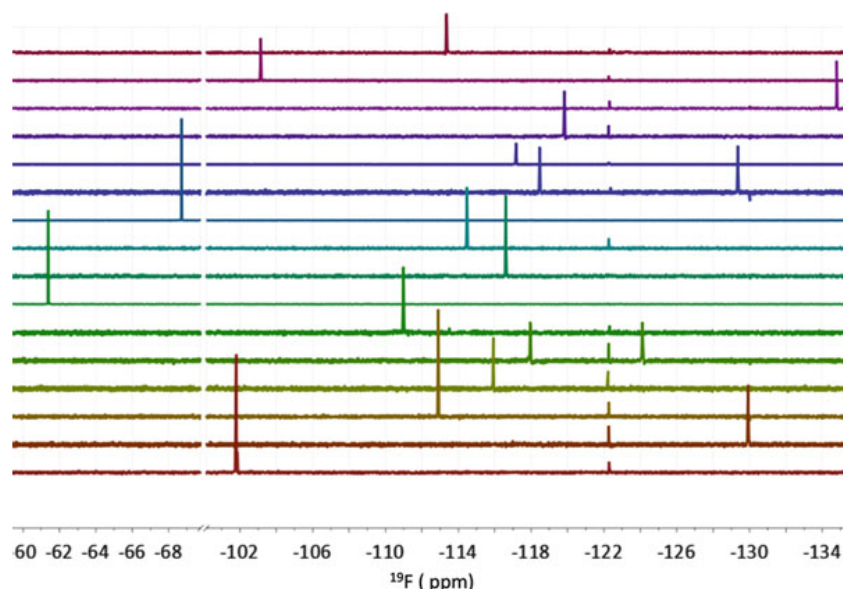


Fig. 4 Mixture design for ^{19}F -NMR fragment screening. Overlay of 16 ^{19}F -containing fragments depicting the excellent chemical shift dispersion that is typically observed in 1D ^{19}F -NMR spectra

reveal the location of the fragment-binding site (if sequence-specific resonance assignments are available for the protein) and the affinity of binding (vide infra). A limitation of protein-detected NMR is that it becomes less suitable with increasing protein size and it requires large amounts of a suitably isotope-labeled protein sample, which increases cost. Where ligand-detected experiments typically measure changes in the intensity of peaks in a spectrum, protein-detected experiments measure changes in the resonance frequency of one or more peaks in the spectrum

so-called *chemical shift perturbations* (CSP). For fragments that usually bind weakly ($K_D > 10 \mu\text{M}$) and exchange rapidly between the free and bound state, a consequence is that in protein-detected experiments the CSP is dictated by the fractional occupancy of the ligand at the binding site, whereas this is not always the case with ligand-detected experiments. For example, in an STD-NMR experiment it is possible to modulate the intensity of signal in the difference experiment by changing the power or duration of the pulse train that is used for protein saturation. For this reason it can be more difficult to identify very weakly binding fragments using protein-detected methods, which may in turn influence the design of a fragment screening library for protein-detected NMR. This is due in part to the fact that fragment hits are often ranked not on the basis of their affinity, but their “ligand efficiency” (LE) [35–37]. This is commonly expressed as

$$LE = \frac{\Delta G}{HAC}$$

where ΔG is the free energy change on binding and HAC is the heavy atom count in the fragment. A molecule with a $K_D = 10 \text{ nM}$ having 36 heavy atoms has a

Table 1 Relationship between the number of heavy atoms (HAC) and equilibrium dissociation constant (K_D) for fragments that bind with similar ligand efficiency (LE). Also shown is the occupancy that would be expected for a fragment screened at a concentration of 1 mM in a protein-detected experiment with a protein concentration of 100 μ M

HAC	LE (kcal HAC ⁻¹ mol ⁻¹)	K_D (mM)	Occupancy @ 1 mM
8	0.3	17	0.05
9	0.3	10	0.09
10	0.3	6.2	0.14
11	0.3	3.8	0.21
12	0.3	2.3	0.30
13	0.3	1.4	0.41
14	0.3	0.82	0.53
15	0.3	0.50	0.65
16	0.3	0.30	0.76

LE \approx 0.3 kcal HAC⁻¹ mol⁻¹, and this is often used as an arbitrary indicative value for compounds possessing “good” LE. Hence small fragments with few heavy atoms can have a LE = 0.3 kcal HAC⁻¹ mol⁻¹ even though they bind with a relatively high K_D (Table 1). Since it is common practice to screen fragments at similar concentrations (e.g., 1 mM) this results in smaller fragments having lower occupancy, which may translate to CSP that are not observable under the experimental conditions.

Affinity Determination

NMR titration data can be analyzed to determine the equilibrium dissociation constant (K_D) for a fragment binding to its protein target. To do so it is common to record a series of NMR spectra at a fixed protein concentration and increasing ligand concentrations and measure chemical shift perturbations (CSP) for resonances in the spectra. These can be recorded as **protein-detected** experiments, where CSP for protein resonances are measured, or **ligand-detected** experiments, where CSP for ligand resonances are measured, as a function of increasing ligand concentration. Figure 5 presents simulated data showing the measured CSP as a function of ligand concentration for a protein-detected experiment (Fig. 5A) and a ligand-detected experiment (Fig. 5B) assuming a binding stoichiometry of 1:1, a K_D of 1 mM, and a maximum CSP of 1 ppm in each case.

Protein-Detected Approaches

Protein-detected experiments are commonly undertaken by recording a series of ¹H, ¹⁵N-HSQC experiments on the protein in the presence of increasing concentrations of a fragment. An example of such a titration is shown in Fig. 6. Several resonances are perturbed upon addition of the fragment, which allows concentration-dependent changes in the CSP to be determined. Variations on this basic approach include the

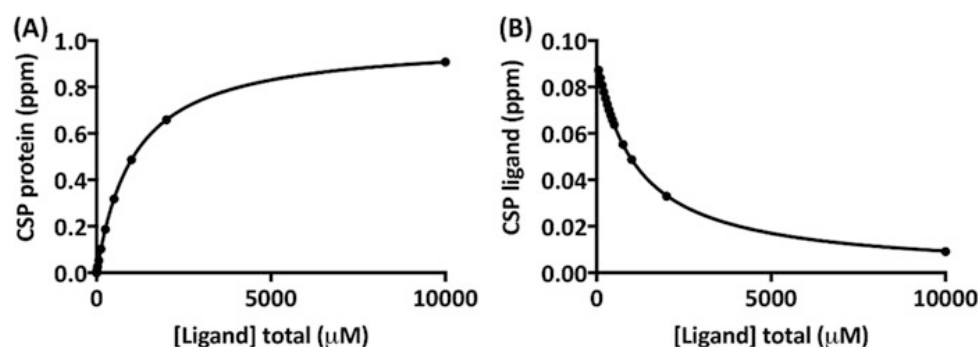


Fig. 5 Calculation of K_D from chemical shift perturbations measured in NMR spectra. The curves were calculated in each case assuming 1:1 stoichiometry, a maximum CSP of 1.0 ppm, a K_D of 1 mM, and a protein concentration of 100 μM. (a) Simulated data for the observed chemical shift perturbation of a **protein** resonance as a function of ligand concentration. (b) Simulated data for the observed chemical shift perturbation of a **ligand** resonance as a function of ligand concentration

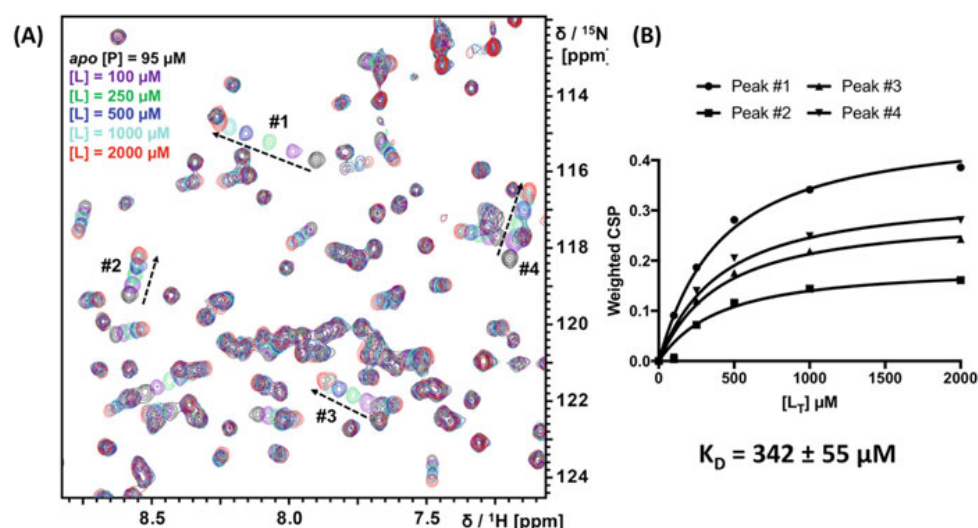


Fig. 6 Chemical shift perturbations observed in a HSQC spectrum upon titration. (a) Spectra acquired with different ligand concentrations are overlaid. The spectrum for each ligand concentration is colored according to the legend and the CSP observed for four residues are highlighted by arrows. (b) The weighted CSP data were fit to a ligand depletion model to estimated K_D for the interaction

use of different experiments (e.g., TROSY for larger proteins) and the measurement of ^{13}C perturbations rather than or as well as those for ^{15}N ; however, the discussion that follows is generally consistent with all of these methods. A *weighted* CSP is calculated to normalize the extent of chemical shift changes in the different frequency dimensions. For the ^1H , ^{15}N HSQC data presented in Fig. 6 the weighted CSP can be calculated using the following eq. [38]:

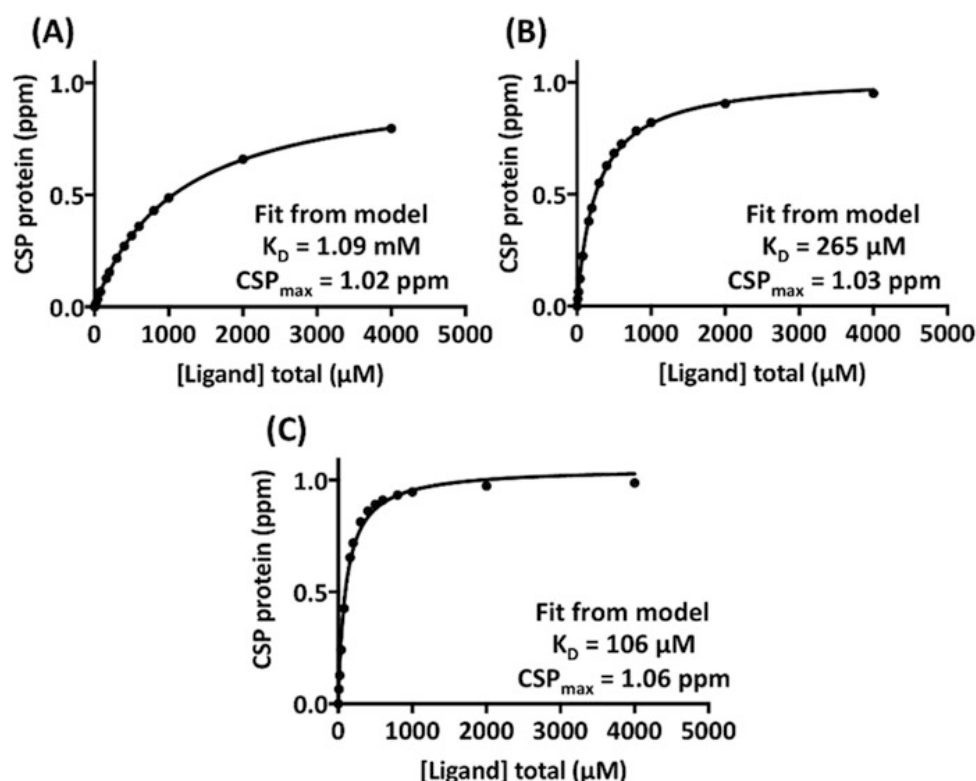


Fig. 7 Fitting CSP data to a simple 1:1 binding model. In each case the data points are calculated to reflect the CSP that would be observed assuming binding with 1:1 stoichiometry, a maximum CSP of 1.0 ppm, and a protein concentration of 100 μM. (a) The data were calculated for (A) $K_D = 1 \text{ mM}$. (b) $K_D = 200 \text{ μM}$. (c) $K_D = 50 \text{ μM}$. The line in each case represents a fit to a simple 1:1 binding model that ignores ligand depletion and assumes that the total ligand concentration is equal to the free ligand concentration. The calculated values from the fit to the model are shown for each case. Although the line of best fit appears reasonable in each case there are significant deviations from the true K_D values, which are more pronounced as the K_D improves

$$CSP = \sqrt{(\Delta\delta_H)^2 + (0.2^*(\Delta\delta_N))^2}$$

where $\Delta\delta_H$ and $\Delta\delta_N$ denote the changes in chemical shift of proton and nitrogen resonances upon addition of the fragment.

Experimentally, it is necessary to have a high enough protein concentration to produce spectra with appropriate sensitivity. The spectra in Fig. 6 were acquired at a protein concentration of $[P] = 95 \text{ μM}$, and required a total acquisition time of 20 min for each HSQC spectrum. For fragments that bind very weakly, it is possible to fit the CSP data to a simple binding model, which makes the assumption that the total ligand concentration (L_T) is equal to the free ligand concentration $[L]$. This simplifies analysis as L_T is known for each point in the titration and the fit does not require knowledge of the protein concentration. Figure 7 shows fits of simulated data to such a model. While the model returns a reasonable fit for a weakly binding fragment with

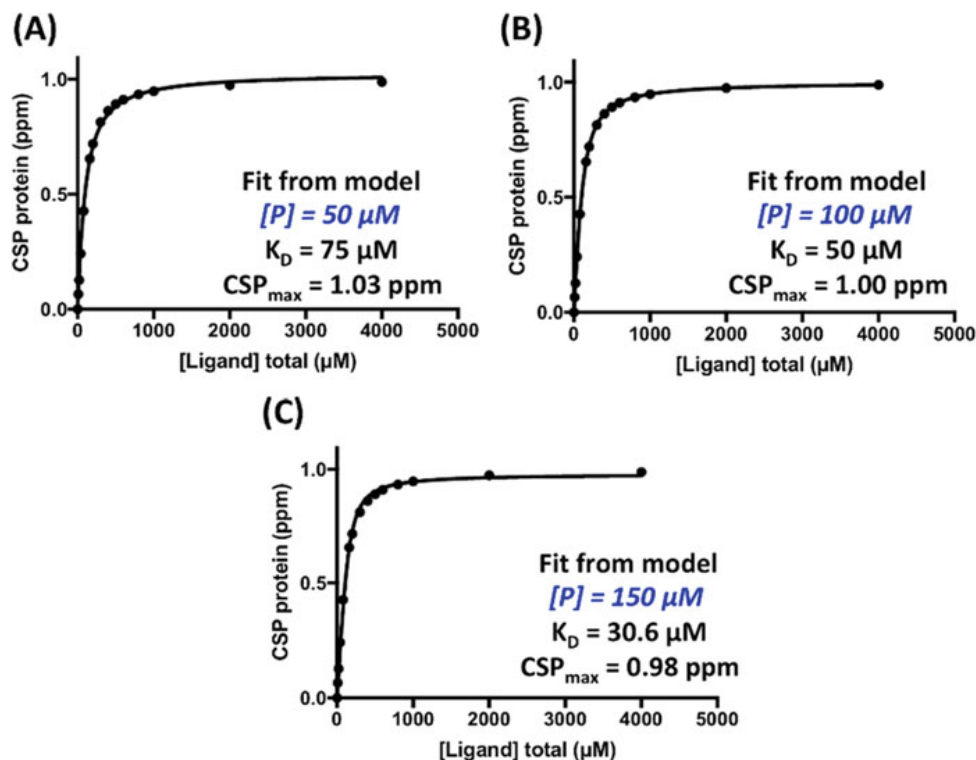


Fig. 8 Fitting CSP data to a 1:1 binding model with ligand depletion. In each case the data points have been calculated to reflect binding with 1:1 stoichiometry, a maximum CSP of 1.0 ppm, a protein concentration of 100 μM, and $K_D = 50 \mu\text{M}$. The line represents the fit to a 1:1 binding model that accounts for ligand depletion, with the protein concentration as shown. (a) The data were fit using a value of $[P] = 50 \mu\text{M}$. (b) The data were fit using the correct value of $[P] = 100 \mu\text{M}$. (c) The data were fit using a value of $[P] = 150 \mu\text{M}$

$K_D = 1 \text{ mM}$, a systematic deviation from the true K_D value is observed as the affinity increases. This arises due to the fact that the assumption that $[L] \approx L_T$ becomes invalid as for fragments that bind with higher affinity a significant amount of the added ligand is bound to the protein. Therefore a more complex model that accounts for this “*ligand depletion*” is required. In practice, a nonlinear regression of the experimental data to Eq. 2 can be used to obtain a more accurate estimate for K_D .

$$[PL] = \frac{(P_T + L_T + K_D) + \sqrt{(P_T + L_T + K_D)^2 - 4P_T L_T}}{2}$$

One complication arising from the use of the *ligand depletion* model is that it requires an accurate measure of protein concentration, which can be difficult to determine. Figure 8 highlights the sensitivity of the ligand depletion model to the measured $[P]$.

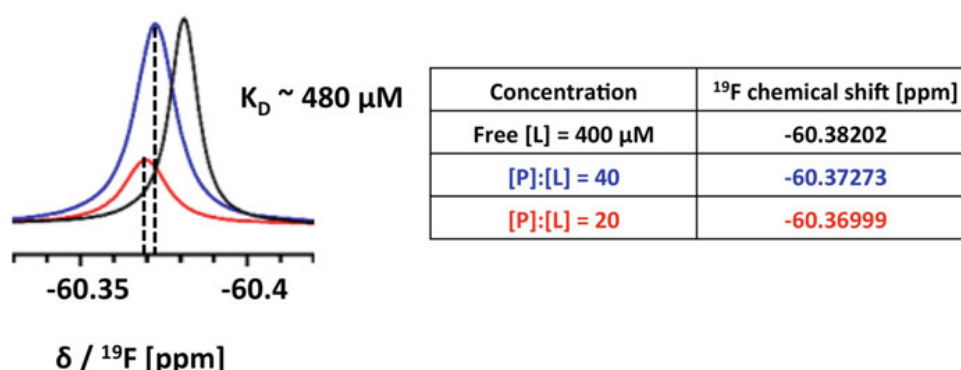


Fig. 9 ¹⁹F spectra for a ligand in the absence of protein and at two different ratios of protein:ligand as shown in the table. The CSP observed in the ligand resonance can be used to estimate K_D for the interaction

Ligand-Detected Approaches

An alternative approach to calculation of K_D has been described based on the measurement of changes in the chemical shift of a ¹⁹F resonance in the ligand upon titration [14]. The ¹⁹F chemical shift of a ligand resonance is measured at two different concentrations under conditions where $[L] \gg [P]$. The frequency shift observed at the two different concentrations can be described as:

$$\gamma = \frac{\Delta\delta F_1}{\Delta\delta F_2}$$

where $\Delta\delta F_1$ and $\Delta\delta F_2$ denote the changes in chemical shift of the fluorine resonance at the two concentrations of ligand. The binding affinity of the interaction can then be determined using the following equation [14, 16]:

$$K_D = \frac{\gamma[L]_1 - \gamma[L]_2}{1 - \gamma}$$

An example of this approach to K_D determination is presented in Fig. 9.

Structure Modeling

Structural details of fragments bound to their targets provide critical information to guide medicinal chemistry that is necessary to elaborate the initial fragment hits into more potent and lead-like compounds. A review of FBDD programs that have been reported in the literature found that the majority utilized structural information, which was most commonly obtained from X-ray crystallography [39]. Crystallography can be particularly challenging with fragments that bind weakly, which dictates that high concentrations of the fragment are required to achieve high levels of

occupancy. In such cases NMR provides a viable alternative but it generally takes longer to obtain structural information compared to X-ray crystallography. This is due in part to the requirement that most NMR methods for structure determination require resonance assignments for protein residues at the ligand-binding site. Additionally NMR structures are less precise than crystal structures; however, high resolution structural data are not necessarily required at the early stages of FBDD [40]. In an attempt to address this problem a number of different approaches to obtain structural details of protein-ligand complexes have been developed, some of which are reviewed below.

Ligand-Observed NMR

Ligand-detected experiments, such as STD-NMR, are sensitive, require relatively little protein, and can identify fragments that bind to a specific protein. However, STD-NMR data does not usually provide information on the specific binding site, except via competition with a high affinity ligand that binds at a known site. Hajduk and coworkers proposed a general method to identify ligand-binding sites in cases where no competitor is available called SOS-NMR (Structural Information using Overhauser Effects and Selective Labeling) [41]. SOS-NMR involves generating a series of protein samples which contain selective methyl/aromatic proton labeling in an otherwise deuterated background and using these to record a series of STD spectra for a ligand. Since observation of a peak in the STD difference spectrum requires intermolecular NOEs to protonated sites in the protein, SOS-NMR makes it possible to generate a set of ambiguous NOE distance constraints between the protein and ligand, which can be used to locate the binding site. This approach was employed to solve the protein-ligand complexes for two systems using NOE-constrained molecular docking: FKBP (MW ~ 14 kDa) in complex with 2-(3'-pyridyl)-benzimidazole and MurA (MW ~ 48 kDa) in complex with uridine diphosphate N-acetylglucosamine. Although this method does not require protein resonance assignments and is not limited by protein size, it requires multiple protein samples with selective labeling and deuteration, which can be expensive and time consuming to produce.

An alternative approach to structure determination from ligand-observed NMR spectra involves measurement of pseudo-contact shifts (PCS) and Paramagnetic Relaxation Enhancements (PRE) from paramagnetic centers in a protein to the ligand, which are commonly either spin labels or lanthanide ions. Both PRE and PCS can be used to generate structural restraints for characterizing weak protein-ligand complexes as they provide information on both the distance and the angle between a nucleus and the center of suitable metal ion. This approach was used by Otting and coworkers [42] to determine the structure of a ligand in complex with a metal-binding protein, and more recently by Siegal and coworkers [43] to determine the structure of ligand in complex with nonmetal-binding protein that had been engineered to contain a lanthanide ion-binding site.

Protein-Observed NMR

Protein-observed NMR methods have been used routinely for determining the ligand-binding modes in weak complexes. In general there are two different sources of structural information that can be used to define the structure of the complex: (i) CSP mapping and (ii) NOE-based methods.

Chemical Shift Perturbation Mapping

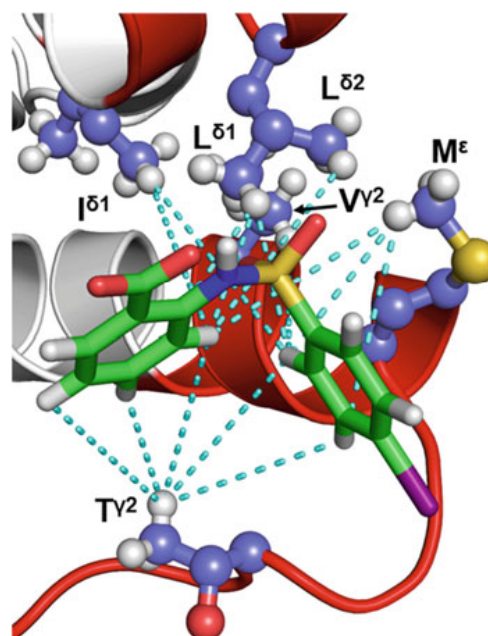
Several different approaches have been described that aim to determine protein-ligand complexes using CSP data for protein resonances, which usually require knowledge of the structure of the protein and sequence-specific resonance assignments. These range from relatively straightforward calculations where clusters of CSP observed on the surface of the protein are used to generate ambiguous distance constraints for docking of the ligand to more complex quantum mechanical calculations which aim to model the size and direction of each CSP at the binding site [44–46].

NOE-Based Protein-Ligand Complexes

More commonly, structures are calculated using distance restraints that are generated from the measurement of intermolecular NOEs between protein and fragment. Traditional NOE-based NMR structure determination requires the complete or almost complete assignment of side-chain resonances for the target protein. This is a significant challenge particularly for proteins with MW > 20 kDa. To accelerate this structure determination process, a number of different approaches have been developed which involve selective labeling of specific sites in the protein. This simplifies the assignment process, but also reduces the number of protonated sites in the protein from which intermolecular NOEs can be observed. For example, Siegal and coworkers reported a method based on a protein sample that contained a subset of protonated methyl groups in an otherwise deuterated protein ([U-²H,¹I^δLV-CH₃]) and unlabeled ligand [47]. These methyl groups could be easily assigned using standard approaches. Subsequently, ILV-methyl-to-ligand NOE and backbone amide-to-ligand NOE data were obtained from 3D ¹³C/¹⁵N-edited NOESY-HSQC spectra. These constraints were then used to generate a structural model of HSP90 (~27 kDa) in complex with an aliphatic ligand.

This method has recently been extended to enable labeling and assignment of all methyl-containing residues (i.e., ILVATM) in a target protein [19, 48], which can be advantageous in cases where there are insufficient ILV residues at the binding site to fully define the structure of a protein-ligand complex. These examples used NOEs from the side chain methyl protons only to generate distance constraints for the structure determination (Fig. 10). Although using NOEs from protein backbone

Fig. 10 The structure of a protein-fragment complex determined using sparse intermolecular NOE data to methyl groups in the protein



amide resonances to the ligand may improve the precision of the structures that are calculated, this can also lead to problems in resolving the intra- and intermolecular NOE cross peaks observed in the experimental data [47].

NMR²

More recently a new approach called “NMR molecular Replacement (NMR²)” has been used to characterize the three-dimensional structure of protein-ligand complexes for weak affinity ligands which requires neither protein resonance assignments nor prior knowledge of protein-binding site/s [49]. The method involves measurement of intermolecular NOEs to generate a set of semi-ambiguous distance constraints. Models of the protein-fragment structure are then generated from unbiased conformational searching using a molecular dynamics simulation in torsional-angle space. By avoiding the need for specific resonance assignments this approach is able to speed up the structure-determination process significantly and has now been applied for a number of different protein-ligand complexes [49, 50].

Concluding Remarks

Since the first description of FBDD, NMR spectroscopy has played a central role in fragment screening, characterization, and structure determination [1]. There are currently two drugs, Vemurafinib [51] and Venetoclax [52], which have been approved for therapeutic use by the FDA. In both cases NMR spectroscopy was

critical both to finding fragments and in guiding their elaboration into more potent molecules. As the technology of FBDD has matured, new NMR-based approaches have been developed to address the significant challenges that are presented in characterizing weak interactions both biophysically and structurally. NMR is very sensitive to the measurement of such weak interactions. Ligand-detected NMR experiments provide not only evidence of fragment binding but also QC data on sample concentration, integrity, and purity, which makes it an extremely powerful approach. New methods for structure-determination are enabling structures to be determined more rapidly and for larger protein systems. It is likely therefore that NMR spectroscopy will remain a key enabling technology for FBDD for the foreseeable future.

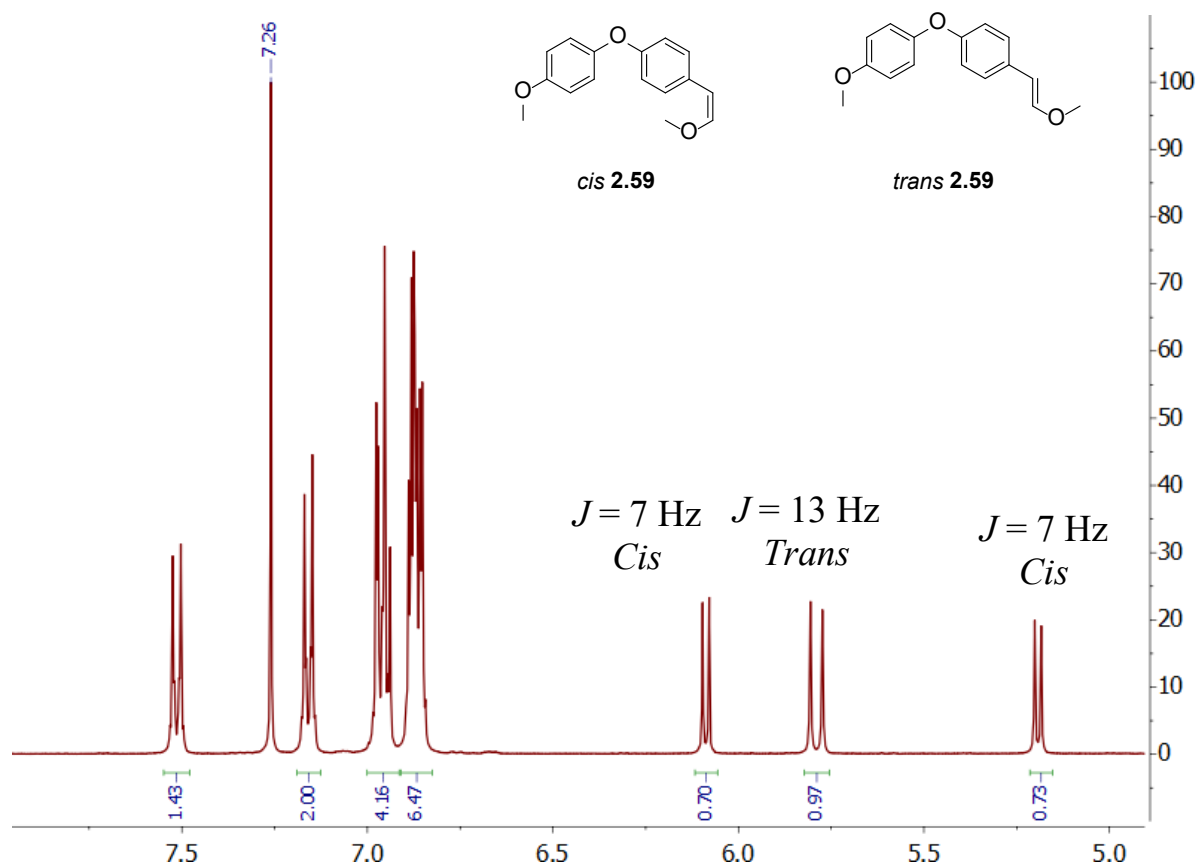
References

1. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. Discovering high-affinity ligands for proteins: SAR by NMR. *Science*. 1996;274:1531–4.
2. Erlanson DA, Fesik SW, Hubbard RE, Jahnke W, Jhoti H. Twenty years on: the impact of fragments on drug discovery. *Nat Rev Drug Discov*. 2016;15:605–19.
3. Blomberg N, Cosgrove DA, Kenny PW, Kolmodin K. Design of compound libraries for fragment screening. *J Comput Aided Mol Des*. 2009;23:513–25.
4. Boyd SM, Turnbull AP, Walse B. Fragment library design considerations. *Wiley Interdiscip Rev-Comput Mol Sci*. 2012;2:868–85.
5. Chen IJ, Hubbard RE. Lessons for fragment library design: analysis of output from multiple screening campaigns. *J Comput Aided Mol Des*. 2009;23:603–20.
6. Doak BC, Morton CJ, Simpson JS, Scanlon MJ. Design and evaluation of the performance of an NMR screening fragment library. *Aust J Chem*. 2013;66:1465–72.
7. Lau WF, Withka JM, Hepworth D, Magee TV, Du YJ, Bakken GA, Miller MD, Hendsch ZS, Thanabal V, Kolodziej SA, Xing L, Hu Q, Narasimhan LS, Love R, Charlton ME, Hughes S, van Hoorn WP, Mills JE. Design of a multi-purpose fragment screening library using molecular complexity and orthogonal diversity metrics. *J Comput Aided Mol Des*. 2011;25:621–36.
8. Keserü GM, Erlanson DA, Ferenczy GG, Hann MM, Murray CW, Pickett SD. Design principles for fragment libraries: maximizing the value of learnings from pharma fragment-based drug discovery (FBDD) programs for use in academia. *J Med Chem*. 2016;59:8189–206.
9. Bharti SK, Roy R. Quantitative ^1H NMR spectroscopy. *TrAC Trends Anal Chem*. 2012;35:5–26.
10. LaPlante SR, Carson R, Gillard J, Aubry N, Coulombe R, Bordeleau S, Bonneau P, Little M, O'Meara J, Beaulieu PL. Compound aggregation in drug discovery: implementing a practical NMR assay for medicinal chemists. *J Med Chem*. 2013;56:5142–50.
11. Dalvit C, Fogliatto G, Stewart A, Veronesi M, Stockman B. WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability. *J Biomol NMR*. 2001;21:349–59.
12. Jordan JB, Poppe L, Xia X, Cheng AC, Sun Y, Michelsen K, Eastwood H, Schnier PD, Nixey T, Zhong W. Fragment based drug discovery: practical implementation based on ^{19}F NMR spectroscopy. *J Med Chem*. 2012;55:678–87.
13. Mayer M, Meyer B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew Chem Int Ed Eng*. 1999;38:1784–8.
14. Jordan JB, Poppe L, Xia X, Cheng AC, Sun Y, Michelsen K, Eastwood H, Schnier PD, Nixey T, Zhong W. Fragment based drug discovery: practical implementation based on $(1)(9)\text{F}$ NMR spectroscopy. *J Med Chem*. 2012;55:678–87.

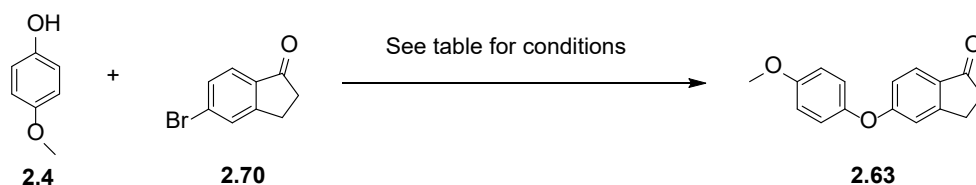
15. Hubbard RE, Davis B, Chen I, Drysdale MJ. The SeeDs approach: integrating fragments into drug discovery. *Curr Top Med Chem*. 2007;7:1568–81.
16. Fielding L. NMR methods for the determination of protein-ligand dissociation constants. *Curr Top Med Chem*. 2003;3:39–53.
17. Guan JY, Keizers PHJ, Liu WM, Lohr F, Skinner SP, Heeneman EA, Schwalbe H, Ubbink M, Siegal G. Small-molecule binding sites on proteins established by paramagnetic NMR spectroscopy. *J Am Chem Soc*. 2013;135:5859–68.
18. Shah DM, Ab E, Diercks T, Hass MAS, van Nuland NAJ, Siegal G. Rapid protein-ligand costructures from sparse NOE data. *J Med Chem*. 2012;55:10,786–90.
19. Mohanty B, Williams ML, Doak BC, Vazirani M, Ilyichova O, Wang G, Bermel W, Simpson JS, Chalmers DK, King GF, Mobli M, Scanlon MJ. Determination of ligand binding modes in weak protein-ligand complexes using sparse NMR data. *J Biomol NMR*. 2016;66:195–208.
20. Davis BJ, Erlanson DA. Learning from our mistakes: the ‘unknown knowns’ in fragment screening. *Bioorg Med Chem Lett*. 2013;23:2844–52.
21. Gossert AD, Jahnke W. NMR in drug discovery: a practical guide to identification and validation of ligands interacting with biological macromolecules. *Prog Nucl Magn Reson Spectrosc*. 2016;97:82–125.
22. Irwin JJ, Duan D, Torosyan H, Doak AK, Ziebart KT, Sterling T, Tumanian G, Shoichet BK. An aggregation advisor for ligand discovery. *J Med Chem*. 2015;58:7076–87.
23. Arrowsmith CH, Audia JE, Austin C, Baell J, Bennett J, Blagg J, Bountra C, Brennan PE, Brown PJ, Bunnage ME, Buser-Doepner C, Campbell RM, Carter AJ, Cohen P, Copeland RA, Cravatt B, Dahlin JL, Dhanak D, Edwards AM, Frederiksen M, Frye SV, Gray N, Grimshaw CE, Hepworth D, Howe T, Huber KV, Jin J, Knapp S, Kotz JD, Kruger RG, Lowe D, Mader MM, Marsden B, Mueller-Farnow A, Muller S, O'Hagan RC, Overington JP, Owen DR, Rosenberg SH, Roth B, Ross R, Schapira M, Schreiber SL, Shoichet B, Sundstrom M, Superti-Furga G, Taunton J, Toledo-Sherman L, Walpole C, Walters MA, Willson TM, Workman P, Young RN, Zuercher WJ. The promise and peril of chemical probes. *Nat Chem Biol*. 2015;11:536–41.
24. Arroyo X, Goldflam M, Feliz M, Belda I, Giralte E. Computer-aided design of fragment mixtures for NMR-based screening. *PLoS One*. 2013;8:e58571.
25. Stark JL, Eghbalnia HR, Lee W, Westler WM, Markley JL. NMRmix: a tool for the optimization of compound mixtures in 1D 1H NMR ligand affinity screens. *J Proteome Res*. 2016;15:1360–8.
26. Lepre CA. Practical aspects of NMR-based fragment screening. *Methods Enzymol*. 2011;493:219–39.
27. Peng C, Frommlet A, Perez M, Cobas C, Blechschmidt A, Dominguez S, Lingel A. Fast and efficient fragment-based lead generation by fully automated processing and analysis of ligand-observed NMR binding data. *J Med Chem*. 2016;59:3303–10.
28. Dalvit C, Pevarello P, Tato M, Veronesi M, Vulpetti A, Sundstrom M. Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. *J Biomol NMR*. 2000;18:65–8.
29. Mashalidis EH, Śledź P, Lang S, Abell C. A three-stage biophysical screening cascade for fragment-based drug discovery. *Nat Protoc*. 2013;8:2309–24.
30. Meiboom S, Gill D. Modified spin-echo method for measuring nuclear relaxation times. *Rev Sci Instrum*. 1958;29:688–91.
31. Carr HY, Purcell EM. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys Rev*. 1954;94:630.
32. Hubbard RE, Murray JB. Experiences in fragment-based lead discovery. *Methods Enzymol*. 2011;493:509–31.
33. Ge X, MacRaild CA, Devine SM, Debono CO, Wang G, Scammells PJ, Scanlon MJ, Anders RF, Foley M, Norton RS. Ligand-induced conformational change of *Plasmodium falciparum* AMA1 detected using 19F NMR. *J Med Chem*. 2014;57:6419–27.

34. Leung EWW, Mulcair MD, Yap BK, Nicholson SE, Scanlon MJ, Norton RS. Molecular insights into the interaction between the SPRY domain-containing SOCS box protein SPSB2 and peptides based on the binding motif from iNOS. *Aust J Chem.* 2017;70:191–200.
35. Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. *Drug Discov Today.* 2004;9:430–1.
36. Hopkins AL, Keseru GM, Leeson PD, Rees DC, Reynolds CH. The role of ligand efficiency metrics in drug discovery. *Nat Rev Drug Discov.* 2014;13:105–21.
37. Murray CW, Erlanson DA, Hopkins AL, Keseru GM, Leeson PD, Rees DC, Reynolds CH, Richmond NJ. Validity of ligand efficiency metrics. *ACS Med Chem Lett.* 2014;5:616–8.
38. Ziarek JJ, Peterson FC, Lytle BL, Volkman BF. Binding site identification and structure determination of protein-ligand complexes by NMR a semiautomated approach. *Methods Enzymol.* 2011;493:241–75.
39. Ferenczy GG, Keseru GM. How are fragments optimized? A retrospective analysis of 145 fragment optimizations. *J Med Chem.* 2013;56:2478–86.
40. Jahnke W. Perspectives of biomolecular NMR in drug discovery: the blessing and curse of versatility. *J Biomol NMR.* 2007;39:87–90.
41. Hajduk PJ, Mack JC, Olejniczak ET, Park C, Dandliker PJ, Beutel BA. SOS-NMR: a saturation transfer NMR-based method for determining the structures of protein-ligand complexes. *J Am Chem Soc.* 2004;126:2390–8.
42. John M, Pintacuda G, Park AY, Dixon NE, Otting G. Structure determination of protein-ligand complexes by transferred paramagnetic shifts. *J Am Chem Soc.* 2006;126:2390–8; 128:12910–6.
43. Guan JY, Keizers PH, Liu WM, Lohr F, Skinner SP, Heeneman EA, Schwalbe H, Ubbink M, Siegal G. Small-molecule binding sites on proteins established by paramagnetic NMR spectroscopy. *J Am Chem Soc.* 2013;126:2390–8; 135:5859–68.
44. Aguirre C, ten Brink T, Cala O, Guichou JF, Krimm I. Protein-ligand structure guided by backbone and side-chain proton chemical shift perturbations. *J Biomol NMR.* 2014;60:147–56.
45. Stark J, Powers R. Rapid protein-ligand costructures using chemical shift perturbations. *J Am Chem Soc.* 2008;126:2390–8; 130:535–45.
46. Yu Z, Li P, Merz Jr KM. Using ligand-induced protein chemical shift perturbations to determine protein-ligand structures. *Biochemistry.* 2017;56:2349–62.
47. Shah DM, AB E, Diercks T, Hass MA, van Nuland NA, Siegal G. Rapid protein-ligand costructures from sparse NOE data. *J Med Chem.* 2012;55:10,786–90.
48. Mohanty B, Rimmer K, McMahon RM, Headey SJ, Vazirani M, Shouldice SR, Coincon M, Tay S, Morton CJ, Simpson JS, Martin JL, Scanlon MJ. Fragment library screening identifies hits that bind to the non-catalytic surface of *Pseudomonas aeruginosa* DsbA1. *PLoS One.* 2017;12:e0173436.
49. Orts J, Walti MA, Marsh M, Vera L, Gossert AD, Guntert P, Riek R. NMR-based determination of the 3D structure of the ligand-protein interaction site without protein resonance assignment. *J Am Chem Soc.* 2016;126:2390–8; 138:4393–400.
50. Walti MA, Riek R, Orts J. Fast NMR-based determination of the 3D structure of the binding site of protein-ligand complexes with weak affinity binders. *Angew Chem Int Ed Eng.* 2017;56:5208–11.
51. Bollag G, Tsai J, Zhang J, Zhang C, Ibrahim P, Nolop K, Hirth P. Vemurafenib: the first drug approved for BRAF-mutant cancer. *Nat Rev Drug Discov.* 2012;11:873–86.
52. Souers AJ, Leverson JD, Boghaert ER, Ackler SL, Catron ND, Chen J, Dayton BD, Ding H, Enschede SH, Fairbrother WJ, Huang DC, Hymowitz SG, Jin S, Khaw SL, Kovar PJ, Lam LT, Lee J, Maecker HL, Marsh KC, Mason KD, Mitten MJ, Nimmer PM, Oleksijew A, Park CH, Park CM, Phillips DC, Roberts AW, Sampath D, Seymour JF, Smith ML, Sullivan GM, Tahir SK, Tse C, Wendt MD, Xiao Y, Xue JC, Zhang H, Humerickhouse RA, Rosenberg SH, Elmore SW. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat Med.* 2013;19:202–8.

Appendix 2.1: ^1H NMR spectra of *cis* and *trans* isomers of enol ether **2.59**



Appendix 2.2: Synthesis conditions for bicyclic cyclopentanone analogue **2.63** via copper catalysed Ullmann coupling.



Reagents and conditions: Cs_2CO_3 and CuI . Ligand, solvent, time and temp as per table below.

Reaction conditions for the synthesis of bicyclic cyclopentanone analogue **2.63**.

Solvent	Ligand	Time (d)	Temperature ($^{\circ}\text{C}$)	Conversion (%) [*]
1,4-Dioxane	<i>N,N</i> -dimethylglycine hydrochloride	14	120	0
1,4-Dioxane	Phenanthroline	4	120	20
DMF	<i>N,N</i> -dimethylglycine hydrochloride	1.5	125	20
DMF	<i>N,N</i> -dimethylglycine hydrochloride	8	95	70

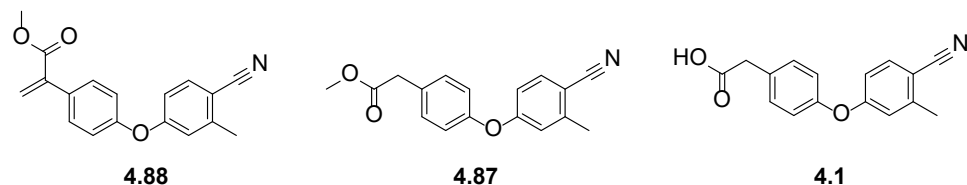
^{*}Conversion estimated by LCMS and HPLC

Appendices

Appendix 4.1: Optimisation of aldol condensation reaction conditions.

Ratios indicated in the time (5h, 24 h and 48 h) are the relative ratio of product **4.88** to starting material **4.87** to carboxylic acid by-product

4.1. Conversions were estimated by crude ^1H NMR and are the relative integrals of each component in the reaction mixture.



Solvent	Temp (°C)	Aldehyde source	Aldehyde (Eq)	5 h	24 h	48 h	Comments
Toluene	50	Formaldehyde (aq.)	3	No product	Minimal	Unclean	By-products observed after 48 h
Toluene	70	Formaldehyde (aq.)	3	No product	33% (1:2:0)	Unclean	By-products observed after 48 h
Toluene	50	Formaldehyde (aq.)	10	No product	No product	No product	No product 4.88 observed
Toluene	70	Formaldehyde (aq.)	10	35% (3:3:1)	60% (12:5:3)	65% (9:3:2)	Conditions used for scale up.
DMF	50	Formaldehyde (aq.)	3	Minimal	Unclean	Unclean	Minimal product 4.88 observed
DMF	70	Formaldehyde (aq.)	3	No product	Unclean	Unclean	No product 4.88 observed
DMF	50	Formaldehyde (aq.)	10	Minimal	Minimal	Minimal	<10% conversion
DMF	70	Formaldehyde (aq.)	10	Minimal	Minimal	Minimal	<10% conversion

Appendices

Toluene	70	Paraformaldehyde	3	60% (3:2:0)	66% (2:1:0)	66% (2:1:0)	Good reaction progress
DMF	70	Paraformaldehyde	3	Minimal	Minimal	Minimal	<10% conversion

Ratios indicated in the time (5h, 24 h and 48 h) are the relative ratio of product **4.88** to starting material **4.87** to carboxylic acid by-product

4.1

Appendix 5.1: Rapid elaboration of fragments into leads by X-ray crystallographic screening of parallel chemical libraries (REFiL_x) supplementary information.

Contents

Table 1. SMILES strings of full amine library and enumerated products	2
Fig 2. Optimisation of microscale reaction conditions	8
Crystallisation and optimisation of crystal soaking conditions for a library of minimally purified reaction products	9
Fig 3. Data collection and processing of the library datasets	10
Fig 4. Refinement and analysis statistics	11
Table 2. Solubility and affinity data for synthesised analogues 2 – 9	12
Table 3. Electron density maps obtained from crude soaks and co-crystals of hits 3 – 6 ...	13
Fig 5. X-ray crystal structure and HSQC titrations of oxadiazole 8	15
Fig 6. Faux background reactions for estimating minimum required conversion.....	16
Fig 7. OMIT maps of pyrazole 6 in the presence and absence of faux background reactions at relative concentrations corresponding to 6, 13 and 28 % reaction conversion.....	17
Experimental methods	20
Protein expression and purification	20
Crystallisation, soaking, data collection and refinement of pure compounds	21
Plate setup	21
Validation of hits and affinity determination by ¹H-¹⁵N HSQC	22
General experimental	22
Experimental and characterisation	24
NMR spectra of all compounds	29
HSQC and affinity data for all analogues.	39
References	47

Appendices

Table 1. SMILES strings of full amine library and enumerated products identified by well position. Conversions, mass observed and new UV vis peak for each well identified. Samples where no UV peak was observed are marked as not observed (N/O).

Well position	Amine SMILES	Product SMILES	Conversion (%)	Mass Observed (Y/N)	New UV Vis peak (Y/N)	PDB code
A01	<chem>Cl.NC(=O)c1nc[nH]c1N</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)NC3=C(N=CN3)C(N)=O)C=C2)=C1)C#N</chem>	60	N	Y	5QKC
A02	<chem>CC(N)CC(O)=O</chem>	<chem>CC(CC(O)=O)NC(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1</chem>	2	Y	Y	5QKH 5QKI 5QKG
A03	<chem>Nc1nc(ccl)C(O)=O</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)NC3=NC(=CO3)C(O)=O)C=C2)=C1)C#N</chem>	N/O	N	N	5QKJ
A04	<chem>CNc1ccc(cn1)C#N</chem>	<chem>CN(C(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1)C1=NC=C(C=C1)C#N</chem>	N/O	N	N	5QKL 5QKK
A05	<chem>OC(=O)C1COCCN1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCOCC3C(O)=O)C=C2)=C1)C#N</chem>	55	Y	Y	5QKM 5QKN
A06	<chem>OCCNC(=O)C1CCNCC1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCC(CC3)C(=O)NCCO)C=C2)=C1)C#N</chem>	60	Y	Y	5QKO
A07	<chem>CN1CCN(CCCCN)CC1</chem>	<chem>CN1CCN(CCCCN(C(=O)CC2=CC=C(OC3=CC(C)=C(C=C3)C#N)C=C2)CC1</chem>	N/O	N	N	5QKQ 5QKR 5QKP
A08	<chem>NCCCCN1CCOCC1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)NCCCCN3CCOCC3)C=C2)=C1)C#N</chem>	34	Y	Y	5QKS

Appendices

A09	<chem>CC(C)(C)OC(=O)N1CCC2CNC2C1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CC4CCNCC34)C=C2)=C1)C#N</chem>	N/O	N	N	5QKT
A10	<chem>CC(CO)CCN</chem>	<chem>CC(CO)CCNC(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1</chem>	N/O	Y	N	5QKD
A11	<chem>CNCCC(=O)NC</chem>	<chem>CNC(=O)CCN(C)C(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1</chem>	28	Y	Y	5QKE
A12	<chem>OS(=O)(=O)C1=NCCN1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCN=C3S(O)(=O)=O)C=C2)=C1)C#N</chem>	N/O	N	N	5QKF
B01	<chem>N#CC1CNCCO1.Cc1ccc(cc1)S(O)(=O)=O</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCOC(C3)C#N)C=C2)=C1)C#N</chem>	11	Y	Y	5QKU 5QKV
B02	<chem>Cl.OC(=O)C1Cc2nc[nH]c2CN1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CC4=C(CC3C(O)=O)N=CN4)C=C2)=C1)C#N</chem>	N/O	Y	N	5QL0
B03	<chem>OCCN1CCCNCC1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCCN(CCO)CC3)C=C2)=C1)C#N</chem>	13	Y	Y	5QL1
B04	<chem>NCC(O)Cc1cccc1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)NCC(O)CC3=CC=CC=C3)C=C2)=C1)C#N</chem>	54	Y	Y	5QL2
B05	<chem>Cl.NC(=O)N1CCNCC1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCN(CC3)C(N)=O)C=C2)=C1)C#N</chem>	53	Y	Y	5QL3
B06	<chem>CC(C)(C)OC(=O)N1CCCC(CC)N)C1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)NCCC3CCCN3)C=C2)=C1)C#N</chem>	N/O	N	Y	5QL4 5QL5

Appendices

B07	Cl.Cl.OC(=O)CCN1CCNCC1	CC1=C(C=CC(OC2=CC =C(CC(=O)N3CCN(CCC(O)=O)CC3)C =C2)=C1)C#N	1	Y	Y	5QL6
B08	Cl.Cl.Cc1nc(CN2CCNCC2)no1	CC1=NC(CN2CCN(CC2)C(=O)CC2=CC =C(OC3=CC(C)=C(C=C3)C#N)C=C2) =NO1	31	Y	Y	5QL7
B09	Cc1ccc(cc1)-c1nnc(N)[nH]1	CC1=CC=C(C=C1)C1=NN =C(NC(=O)CC2=CC=C(OC3=CC(C) =C(C=C3)C#N)C=C2)N1	34	Y	Y	5QL8 5QL9
B10	CC(=O)N1CCN(CCCN)CC1	CC(=O)N1CCN(CCCNC(=O)CC2=CC =C(OC3=CC(C)=C(C=C3)C#N)C =C2)CC1	14	Y	Y	5QKX 5QKW
B11	CC(C)C(N)(CO)C(O)=O	CC(C)[C@@](CO)(NC(=O)CC1=CC =C(OC2=CC(C)=C(C=C2)C#N)C =C1)C(O)=O	N/O	N	N	
B12	OCCNC(=O)C1CCCNC1	CC1=C(C=CC(OC2=CC =C(CC(=O)N3CCCC(C3)C(=O)NCCO)C =C2)=C1)C#N	7	Y	Y	5QKY 5QKZ
C01	C1CN(CCN1)c1encn1	CC1=C(C=CC(OC2=CC =C(CC(=O)N3CCN(CC3)C3=NC=CN =C3)C=C2)=C1)C#N	90	Y	Y	5QLA
C02	CC(=O)N1CCNCC1	CC(=O)N1CCN(CC1)C(=O)CC1=CC =C(OC2=CC(C)=C(C=C2)C#N)C=C1	86	Y	Y	5QLE
C03	Cl.CNCC(=O)OC	COC(=O)CN(C)C(=O)CC1=CC=C(OC2 =CC(C)=C(C=C2)C#N)C=C1	42	Y	Y	5QLG 5QLF
C04	CN(C)Cc1cccc(N)c1	CN(C)CC1=CC=CC(NC(=O)CC2=CC =C(OC3=CC(C)=C(C=C3)C#N)C=C2) =C1	57	Y	Y	5QLH

Appendices

C05	Cl.CCOC(=O)C1CCNCC1=O	CC1=C(C=CC(OC2=CC =C(CC(=O)N3CCC(C(O)=O)C(=O)C3)C =C2)=C1)C#N	47	Y	Y	5QLI
C06	CC1=NCCN1	CC1=NCCN1C(=O)CC1=CC=C(OC2 =CC(C)=C(C=C2)C#N)C=C1	47	Y	Y	5QLJ
C07	Cl.Cl.Cn1nccc1C1CCCCN1	CN1N=CC=C1C1CCCCN1C(=O)CC1 =CC=C(OC2=CC(C)=C(C=C2)C#N)C =C1	7	Y	Y	5QLK
C08	Br.CC(N)c1ccncc1	CC(NC(=O)CC1=CC=C(OC2=CC(C) =C(C=C2)C#N)C=C1)C1=CC=NC=C1	31	Y	Y	5QLL
C09	Cl.COC(=O)N1CCNCC1	CC1=C(C=CC(OC2=CC =C(CC(=O)N3CCNCC3)C=C2)=C1)C#N	21	Y	Y	5QLM
C10	NC1CCN(Cc2ccccc2)C1	CC1=C(C=CC(OC2=CC =C(CC(=O)NC3CCN(CC4=CC=CC =C4)C3)C=C2)=C1)C#N	22	Y	Y	5QLB
C11	CNCC1CCN(CCO)CC1	CN(CC1CCN(CCO)CC1)C(=O)CC1=CC =C(OC2=CC(C)=C(C=C2)C#N)C=C1	12	Y	Y	5QLC
C12	CC(C)(CO)CCCN	CC1=C(C=CC(OC2=CC =C(CC(=O)NCCCC(C)(C)CO)C=C2) =C1)C#N	3	Y	Y	5QLD
D01	CN1CCCN(CCN)CC1	CN1CCCN(CCN)C(=O)CC2=CC=C(OC3 =CC(C)=C(C=C3)C#N)C=C2)CC1	N/O	N	Y	5QLP 5QLQ 5QLO
D02	CNCCCN(C)C	CN(C)CCCN(C)C(=O)CC1=CC=C(OC2 =CC(C)=C(C=C2)C#N)C=C1	54	Y	Y	5QLV 5QLU
D03	OC(=O)C1CCN1	CC1=C(C=CC(OC2=CC =C(CC(=O)N3CC[C@@H]3C(O)=O)C =C2)=C1)C#N	48	Y	Y	5QLW

Appendices

D04	CN1CCN(CC1)c1nc(N)n[nH]1	CN1CCN(CC1)C1=NC(NC(=O)CC2=CC=CC(OC3=CC(C)=C(C=C3)C#N)C=C2)=NN1	54	Y	Y	5QLX
D05	OS(O)(=O)=O.Nc1ncc[nH]1	CC1=C(C=CC(OC2=CC=C(CC(=O)NC3=NC=CN3)C=C2)=C1)C#N	N/O	N	N	5QLY
D06	CC(C)(C)OC(=O)N1CCC(CCN)CC1	CC1=C(C=CC(OC2=CC=C(CC(=O)NCCC3CCNCC3)C=C2)=C1)C#N	N/O	N	Y	5QLZ
D07	OC1CCCNC1	CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCCC(O)C3)C=C2)=C1)C#N	43	Y	Y	5QM0
D08	CNCC(O)CO	CN(CC(O)CO)C(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1	37	Y	Y	5QM1
D09	Cl.CN[C@@H](CC(O)=O)C(=O)OC	COC(=O)C[C@H](N(C)C(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1)C(O)=O	9	Y	Y	5QM2
D10	NCc1cccc(CO)c1	CC1=C(C=CC(OC2=CC=C(CC(=O)NCC3=CC(CO)=CC=C3)C=C2)=C1)C#N	16	Y	Y	5QLR
D11	Cl.Cl.CN1CCNC(C1)C(O)=O	CN1CCN(C(C1)C(O)=O)C(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1	19	Y	Y	5QLS
D12	CCNC1CCN(C1)C(=O)OC(C)C	CCN(C1CCNC1)C(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1	N/O	N	Y	5QLT
E01	Cl.COC1CNC1	COC1CN(C1)C(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1	59	Y	Y	5QM3
E02	OC(=O)C1CCNC1	CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCC(C3)C(O)=O)C=C2)=C1)C#N	N/O	N	N	5QM7

Appendices

E03	<chem>CC(N)CN1CCOCC1</chem>	<chem>CC(CN1CCOCC1)NC(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1</chem>	65	Y	Y	5QM8
E04	<chem>COC(=O)C1CCC(C)NC1</chem>	<chem>CC1CCC(CN1C(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1)C(O)=O</chem>	34	Y	Y	5QM9
E05	<chem>CCOC(=O)Cc1ccc(N)c(Cl)c1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)NC3=CC=C(CC(O)=O)C=C3Cl)C=C2)=C1)C#N</chem>	20	Y	Y	5QMA
E06	<chem>Cl.O=S1(=O)CCNCC1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCS(=O)(=O)CC3)C=C2)=C1)C#N</chem>	47	Y	Y	
E07	<chem>NC1(CC1)C(O)=O</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)NC3(CC3)C(O)=O)C=C2)=C1)C#N</chem>	12	Y	Y	5QMB
E08	<chem>Cl.COC(=O)CC1CCNCC1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCC(CC(O)=O)CC3)C=C2)=C1)C#N</chem>	35	Y	Y	5QMC
E09	<chem>NCCCS(O)(=O)=O</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)NCCCS(O)(=O)=O)C=C2)=C1)C#N</chem>	4	Y	Y	5QMD
E10	<chem>OC(=O)CNC(=O)[C@@H]1CCCN1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCC[C@H]3C(=O)NCC(O)=O)C=C2)=C1)C#N</chem>	20	Y	Y	5QM4
E11	<chem>CCOC(=O)N1CCC(N)CC1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)NC3CCN(CC3)C(O)=O)C=C2)=C1)C#N</chem>	8	Y	Y	5QM5
E12	<chem>CC(C)(C)OC(=O)N1CCC(CC1)C1CCNCC1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCC(CC3)C3CCNCC3)C=C2)=C1)C#N</chem>	N/O	N	Y	5QM6

Appendices

F01	CN1CCNCC1=O	CN1CCN(CC1=O)C(=O)CC1=CC=CC(OC2=CC(C)=C(C=C2)C#N)C=C1	74	Y	Y	5QMF
F02	CC(C)CC1(CO)CCCN1	CC(C)CC1(CO)CCCN(C1)C(=O)CC1=CC=CC(OC2=CC(C)=C(C=C2)C#N)C=C1	60	Y	Y	5QMK
F03	NCCCN1CCCC(CO)C1	CC1=C(C=CC(OC2=CC=CC(CC(=O)NCCCN3CCCC(CO)C3)C=C2)=C1)C#N	N/O	N	Y	5QML
F04	CC(C)(C)OC(=O)N1CCN(CCN)CC1	CC1=C(C=CC(OC2=CC=CC(CC(=O)NCCCN3CCNCC3)C=C2)=C1)C#N	N/O	N	Y	5QMM
F05	C(C1CNCCCO1)n1cccn1	CC1=C(C=CC(OC2=CC=CC(CC(=O)N3CCCO(CN4C=CC=N4)C3)C=C2)=C1)C#N	54	Y	Y	5QMN
F06	Cl.Cl.Cc1ncn1CCCN	CC1=NC=CN1CCCN(C(=O)CC1=CC=CC(OC2=CC(C)=C(C=C2)C#N)C=C1	45	Y	Y	5QMO
F07	Cl.CC(C)NCCCc1cccc1	CC(C)N(CCCCC1=CC=CC=C1)C(=O)CC1=CC=CC(OC2=CC(C)=C(C=C2)C#N)C=C1	11	Y	Y	5QMP
F08	Cl.CC(C)(CO)NCc1ccnc1	CC1=C(C=CC(OC2=CC=CC(CC(=O)N(CC3=CC=CN=C3)C(C)(C)CO)C=C2)=C1)C#N	4	Y	Y	5QMQ
F09	Cl.Cl.CCC(N1CCNCC1)C(O)=O	CCC(N1CCN(CC1)C(=O)CC1=CC=CC(OC2=CC(C)=C(C=C2)C#N)C=C1)C(O)=O	2	Y	Y	5QMR
F10	OCCC1NCCNC1=O	CC1=C(C=CC(OC2=CC=CC(CC(=O)N3CCNC(=O)C3CCO)C=C2)=C1)C#N	32	Y	Y	5QMG
F11	Cl.CNCCc1noc(C)n1	CN(CCC1=NOC(C)=N1)C(=O)CC1=CC=CC(OC2=CC(C)=C(C=C2)C#N)C=C1	6	Y	Y	5QMH

Appendices

F12	<chem>Cl.NC(=O)c1cc(N)cc(c1)C(N)=O</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)NC3=CC(=CC(=C3)C(N)=O)C(N)=O)C=C2)=C1)C#N</chem>	N/O	N	N	5QMJ 5QMI
G01	<chem>CC(CN)N1CCOCC1</chem>	<chem>CC(CNC(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1)N1CCOCC1</chem>	N/O	Y	N	5QMS
G02	<chem>COCCN1CC(CN)CC1=O</chem>	<chem>COCCN1CC(CNC(=O)CC2=CC=C(OC3=CC(C)=C(C=C3)C#N)C=C2)CC1=O</chem>	7	Y	Y	5QMW
G03	<chem>NCCCN1CCC(O)CC1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)NCCCN3CCC(O)CC3)C=C2)=C1)C#N</chem>	N/O	N	Y	5QMX
G04	<chem>CC(C)N1CCC(CCN)CC1</chem>	<chem>CC(C)N1CCC(CCNC(=O)CC2=CC=C(OC3=CC(C)=C(C=C3)C#N)C=C2)CC1</chem>	51	Y	Y	5QMY
G05	<chem>OC(=O)C(F)(F)F.CC1CNC(=O)CCN1</chem>	<chem>CC1CNC(=O)CCN1C(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1</chem>	47	Y	Y	5QMZ
G06	<chem>CCNCCn1cccn1</chem>	<chem>CCN(CCN1C=CC=N1)C(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1</chem>	62	Y	Y	5QN2 5QN0 5QN1
G07	<chem>Cl.Cl.C1CC(CCN1)Oc1ccnc1</chem>	<chem>CC1=C(C=CC(OC2=CC=C(CC(=O)N3CCC(CC3)OC3=CC=CN=C3)C=C2)=C1)C#N</chem>	41	Y	Y	5QN3
G08	<chem>CNCCCc1c(C)n[nH]c1C</chem>	<chem>CN(CCCC1=C(C)NN=C1C)C(=O)CC1=CC=C(OC2=CC(C)=C(C=C2)C#N)C=C1</chem>	21	Y	Y	5QN4 5QN5
G09	<chem>CNCC1CCN(CCOC)CC1</chem>	<chem>COCCN1CCC(CN(C)C(=O)CC2=CC=C(OC3=CC(C)=C(C=C3)C#N)C=C2)CC1</chem>	24	Y	Y	5QN6

Appendices

G10	Cl.Cn1nc(C(N)=O)c2CNCCc12	CN1N=C(C(N)=O)C2 =C1CCN(C2)C(=O)CC1=CC=C(OC2 =CC(C)=C(C=C2)C#N)C=C1	12	Y	Y	5QMT
G11	CNCC1CCN(CC1)C(=O)OC(C) (C)C	CN(CC1CCNCC1)C(=O)CC1=CC =C(OC2=CC(C)=C(C=C2)C#N)C=C1	4	Y	Y	5QMU
G12	CCOC(=O)c1cnc(N)[nH]1	CC1=C(C=CC(OC2=CC=C(CC(=O)NC3 =NC=C(N3)C(O)=O)C=C2)=C1)C#N	2	Y	Y	5QMV
H01	Cl.O=C1CCNC1	CC1=C(C=CC(OC2=CC =C(CC(=O)N3CCC(=O)C3)C=C2) =C1)C#N	N/O	N	Y	5QN7
H02	Nc1ncc(en1)C(O)=O	CC1=C(C=CC(OC2=CC=C(CC(=O)NC3 =NC=C(C=N3)C(O)=O)C=C2)=C1)C#N	N/O	N	N	5QN8
H03	OC(=O)CN1CCNCC1=O	CC1=C(C=CC(OC2=CC =C(CC(=O)N3CCN(CC(O)=O)C(=O)C3)C =C2)=C1)C#N	30	Y	Y	5QN9
H04	Cl.O=C1CNCN1	CC1=C(C=CC(OC2=CC =C(CC(=O)N3CNC(=O)C3)C=C2) =C1)C#N	1	Y	Y	5QNA
H05	C1COCCN1	CC1=C(C=CC(OC2=CC =C(CC(=O)N3CCOCC3)C=C2)=C1)C#N	83	N	Y	5QNB
H06	CN1CCNCC1	CN1CCN(CC1)C(=O)CC1=CC=C(OC2 =CC(C)=C(C=C2)C#N)C=C1	20	N	Y	5QNC
H07	NC(=O)C1CCNCC1	CC1=C(C=CC(OC2=CC =C(CC(=O)N3CCC(CC3)C(N)=O)C=C2) =C1)C#N	36	N	Y	5QNE 5QND
H08	CC(N)C(O)=O	CC(NC(=O)CC1=CC=C(OC2=CC(C) =C(C=C2)C#N)C=C1)C(O)=O	N/O	N	Y	

Appendices

H09	OC(=O)C1CCNCC1	CC1=C(C=CC(OC2=CC =C(CC(=O)N3CCC(CC3)C(O)=O)C=C2) =C1)C#N	15	N	Y	5QNF
-----	----------------	---	----	---	---	------

SI Fig 2. Optimisation of microscale reaction conditions

Fourty five microscale amidation reactions were conducted to optimise the reaction solvent, coupling reagent, temperature and amine equivalents (**Table S1**).

Appendices

In a polypropylene 96-well plate, phenylacetic acid **2** (0.2 M in 15% v/v dry Et₃N/solvent, 10 μ L, 1.0 eq., 2 μ mol) and coupling reagent (1 M in dry DMF for HATU/COMU and H₂O for EDC, 4 μ L, 2.0 eq., 4 μ mol) were added to each well. The plate was sealed and left at room temperature for 15 min without agitation. After this time, a 0.5 M solution of amine **10**, **11** or **12** in DMF as per the table below was added to each well and the plate left at temperature for a further 18 h without agitation. The solvent was removed under vacuum and 20 μ L of DMSO added to each well. A 1 μ L aliquot was taken from each well and diluted in MeCN (49 μ L) and analysed by LCMS.

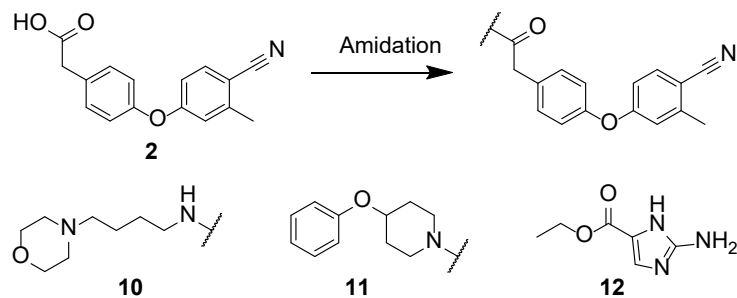


Table S1: LCMS conversion for the chemistry optimisation plate.

Solvent			DMF			MeCN			THF		
Amine			10	11	12	10	11	12	10	11	12
Coupling agent	Temp(°C)	Eq.									
EDC	25	1.0	20	28	6	23	34	9	13	25	0
	25	1.5	23	34	11	24	43	14	19	59	4
	25	3.0	0	46	12	4	53	15	3	61	12
HATU	25	1.5				42	85	20			
COMU	25	1.5				46	0	19			
EDC	25	1.5				28	22	15			
HATU	50	1.5				44	9	47			
COMU	50	1.5				47	0	38			
EDC	50	1.5				18	22	11			

X-ray crystallography optimisation: Crystallisation and optimisation of crystal soaking conditions for a library of minimally purified reaction products

The previously published crystallisation conditions of oxidised *EcDsbA* were used as a guide for obtaining diffraction quality crystals.^[1] Briefly, protein was crystallised using hanging drop vapour diffusion. 1 μ L of 30 mg/mL *EcDsbA* was mixed with an equal volume of crystallisation buffer (11-13% PEG 8000, 5-7.5% glycerol, 1 mM CuCl₂, 100 mM sodium cacodylate pH 6.1) and equilibrated against 0.5 mL of reservoir buffer at 18 °C. Typically, crystals appeared within one to three days and grew to average dimensions of 0.6 mm \times 0.4 mm \times 0.2 mm.

We hypothesised that soaking with unpurified reaction products might present different challenges compared with soaking of purified compounds, and therefore a number of optimisation steps were carried out prior to soaking the library.

Firstly, 200 mM stocks were prepared for three coupling agents (EDC in MilliQ H₂O, COMU in DMF and HATU in DMF) and crystals were soaked with each at a final concentration of 20 mM. Crystals soaked with COMU diffracted at significantly lower resolution compared to those soaked with EDC and HATU (2.4 ± 0.2 Å for COMU vs 1.8 ± 0.1 Å and 1.9 ± 0.1 Å for EDC and HATU respectively).

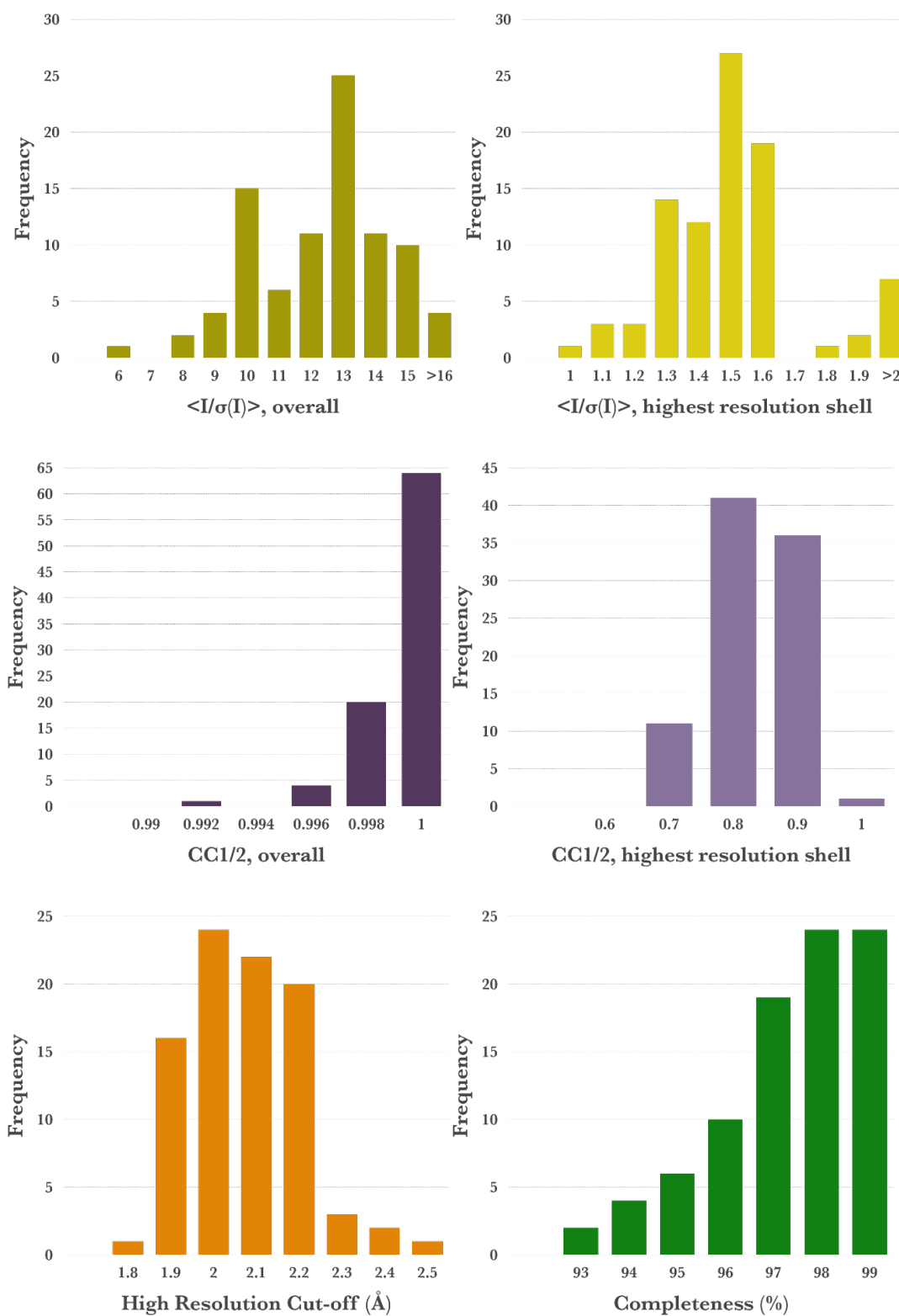
In addition, we sought to test a range of parameters in order to optimise the soaking conditions for the microscale reactions:

- Positive/negative control compounds with background reagents
- Soaking time/concentration
- Solvent sensitivity

Fig 3. Data collection and processing of the library datasets

Diffraction data were collected at the Australian synchrotron MX1 beamline, which is part of ANSTO.^[2] All of the data were processed using the automated data processing pipeline implemented at the beamline, where data were indexed, integrated and scaled with *xdsme* and *AIMLESS*.^[3] The resulting data collection statistics for each dataset were reviewed and if necessary re-processed with *XDS* or *IMOSFLM* and *AIMLESS*. For all the datasets we chose a resolution cut-off based on the following criteria in the highest resolution range met: $CC_{1/2}$ was at least 0.6, $\langle I/\sigma(I) \rangle$ was greater than 1.0 and completeness was greater than 90%.^[4] In cases where data collection statistics did not fulfil our quality requirement or crystals diffracted to lower than 2.5 Å resolution, the crystallographic experiment was repeated (12 crude products required repeat soaks). All datasets were deposited to the PDB and can be downloaded for further evaluation. Library data collection and refinement statistics can be found below.

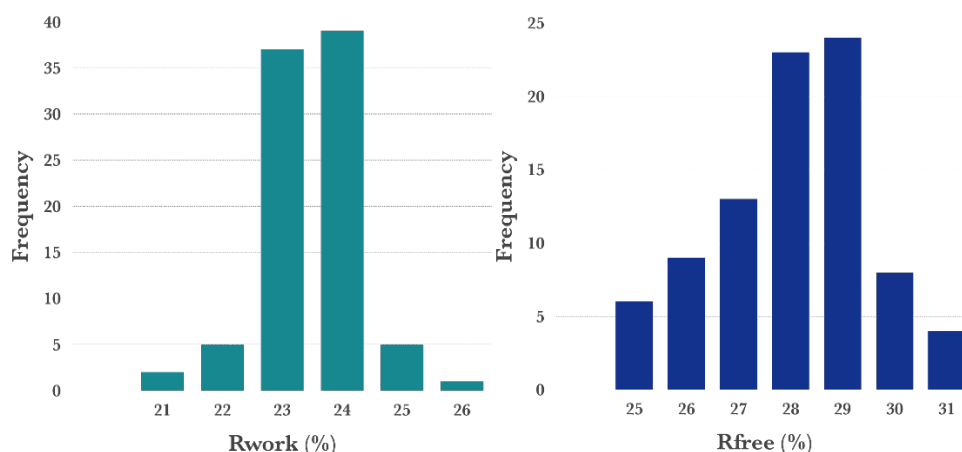
Appendices



Data collection statistics. Distribution for (A) overall $\langle I/\sigma(I) \rangle$ and (B) the highest resolution shell $\langle I/\sigma(I) \rangle$; (C) overall $CC_{1/2}$ and (D) the highest resolution shell $CC_{1/2}$; (E) Distribution of the resolution and (F) completeness are shown for 90 datasets.

Fig 4. Refinement and analysis statistics

All datasets were phased using molecular replacement (MR) using PDB structure 1FVK as a search model. MR and automated refinement steps were performed using the Auto-Rickshaw molecular replacement automated pipeline.^[5] Statistics of the results after the automated refinement can be found in **Figure 2**. The resulting models were inspected manually and, alongside DMSO-soaked apo-EcDsbA datasets, were used as input to PanDDA.^[6] For each ‘event’, the PanDDA maps were inspected. In several cases phenylacetic acid **2** was binding in the active site, and therefore many of the ‘events’ did not correspond to the elaborated ligand binding. Nonetheless, four unpurified reaction products were identified as hits since there was additional electron density present that allowed us to confidently model a corresponding amide extension from starting material. These 4 datasets were manually refined and deposited to the protein databank (PDB).



Statistic for the Results from the Automated Refinement. (A)-(B) histograms are shown for the Rwork and Rfree values for all 90 library structures.

Table 2. Solubility and affinity data for synthesised analogues 2 – 9.

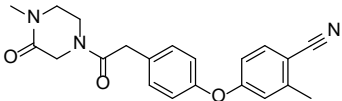
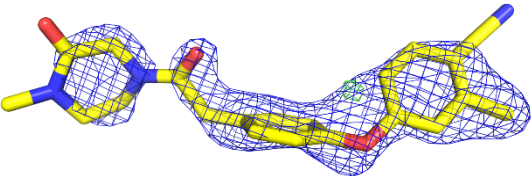
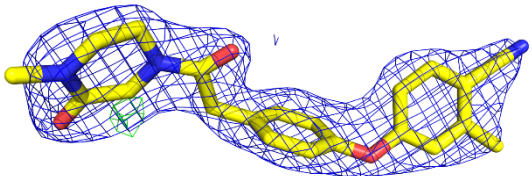
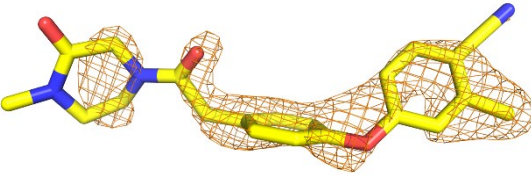
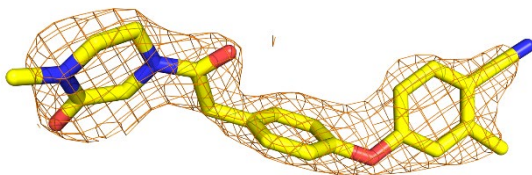
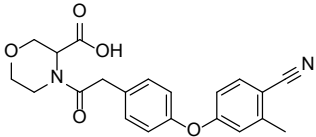
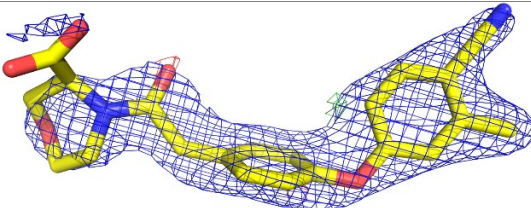
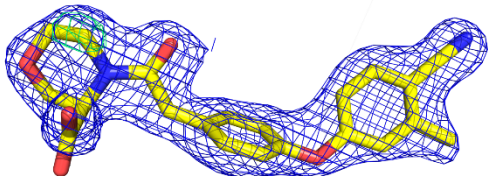
Compound number	Structure	Solubility (μM)	Affinity (K_D , μM)	PDB code crudes	PDB code co-crystal
2		1000	490 ± 20	-	6N9C
3		1000	480 ± 20	6PIQ	6PDH
4		500	290 ± 10	6PG1	6PC9
5		250	365 ± 30	6PGJ	6PD7
6		1000	$62 \pm 9^*$	6PG2	6PBI
7		1000	430 ± 20	-	-
8		500	230 ± 20	-	6PLI
9		1000	1100 ± 80	-	-

Presented as the mean \pm SEM from 3 independent experiments

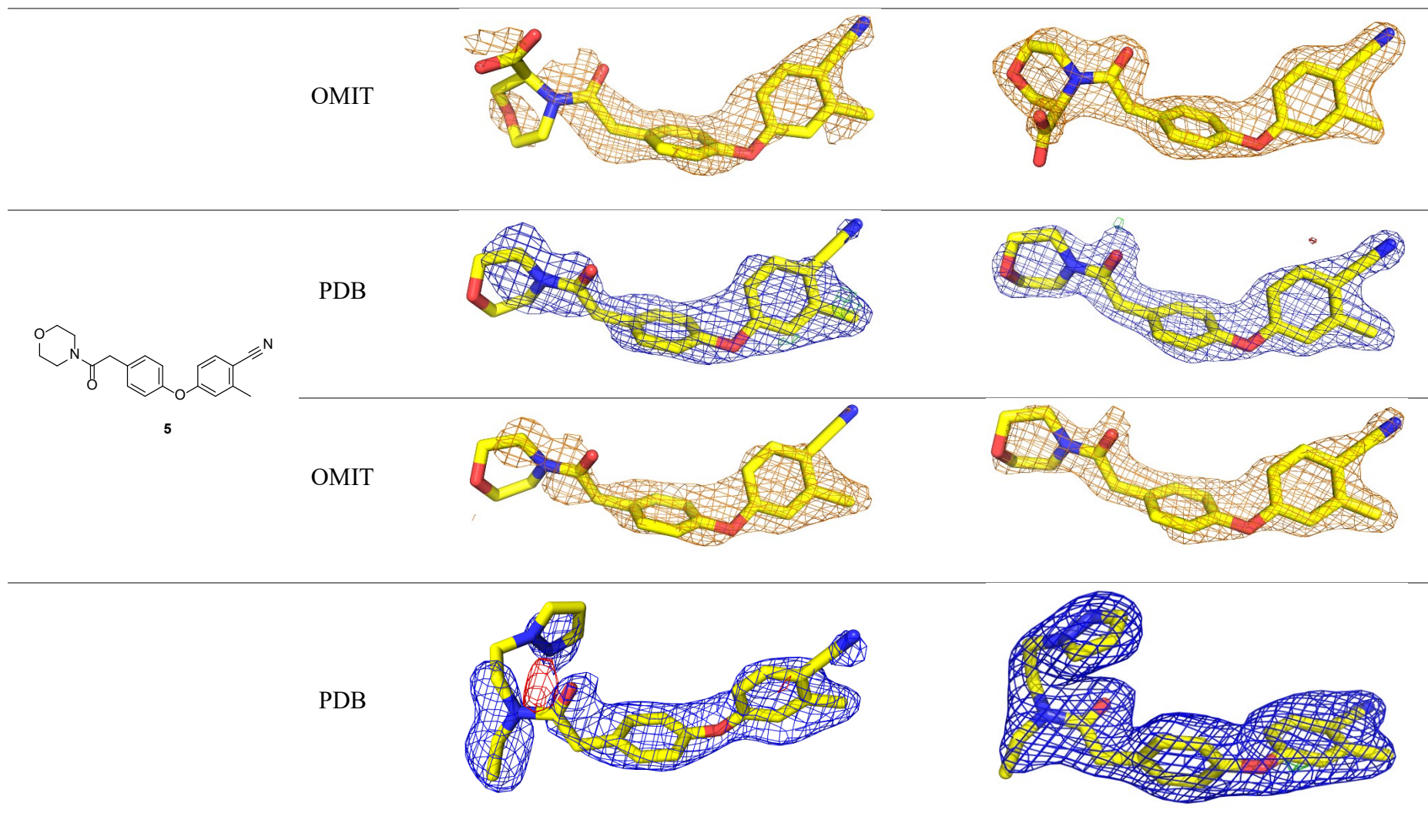
*

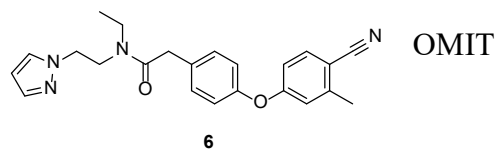
Appendices

Table 3. Electron density maps obtained from crude soaks and co-crystals of hits 3 – 6. Contouring level of 2mFo-DFc maps set at 1 σ (blue) and mFo-DFc maps are shown at $\pm 3 \sigma$ (positive green and negative red). Omit maps were generated by simulated annealing and are shown as 2mFo-DFc at 1 σ (orange). Omit electron densities have ligands shown in yellow sticks for reference only.

Structure		Crude	Co-crystal
 3	PDB		
	OMIT		
 4	PDB		

Appendices





OMIT

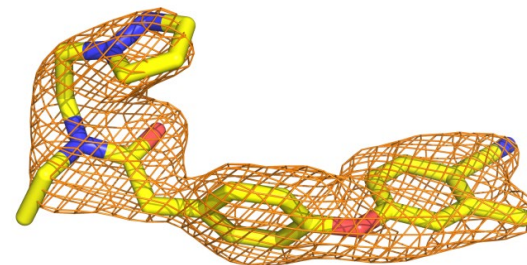
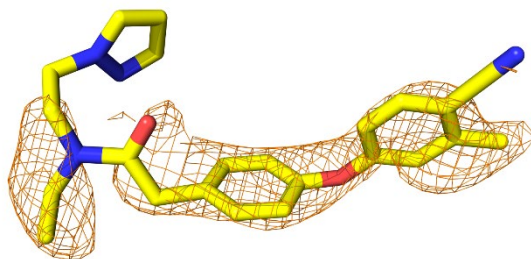
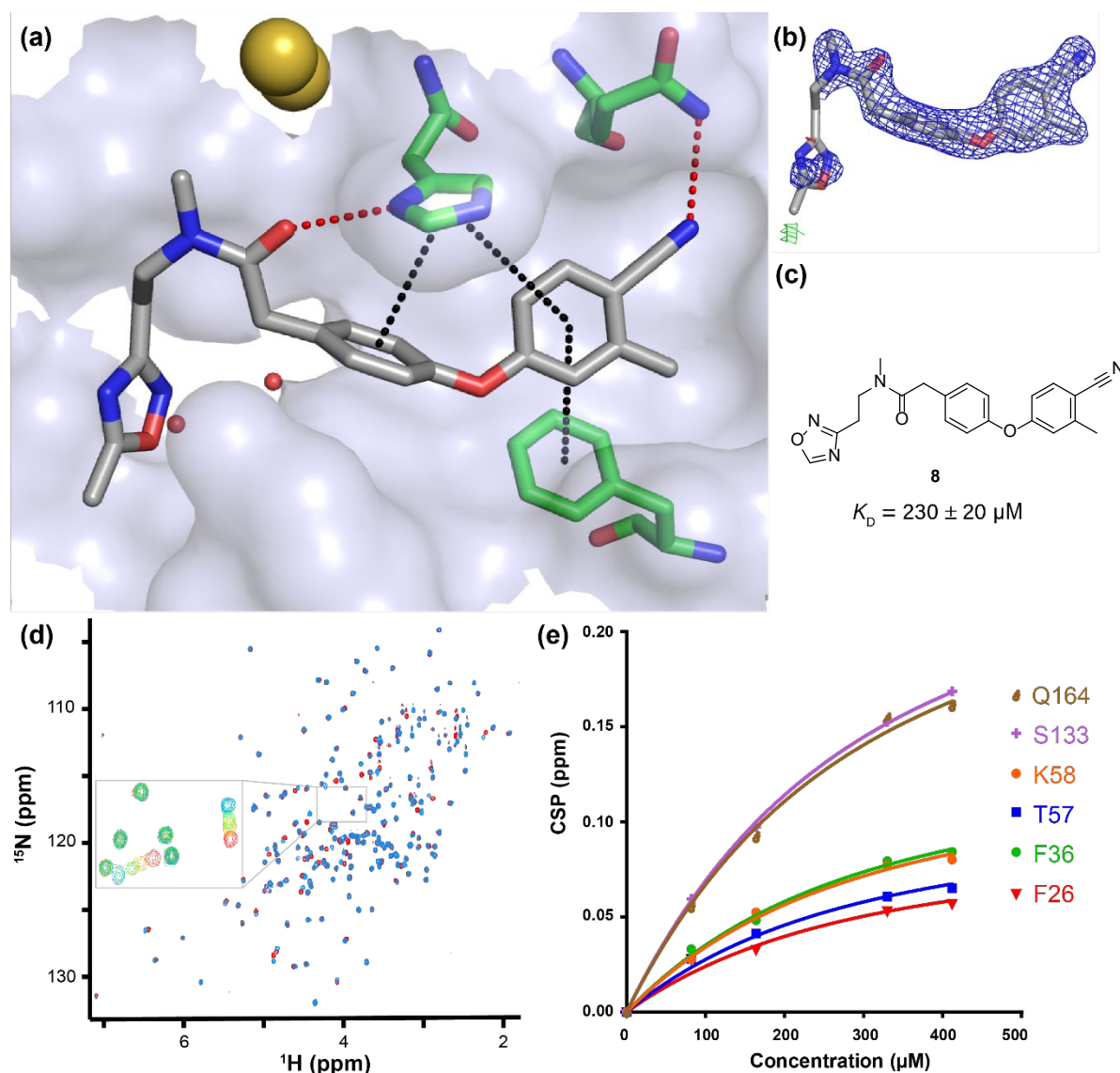
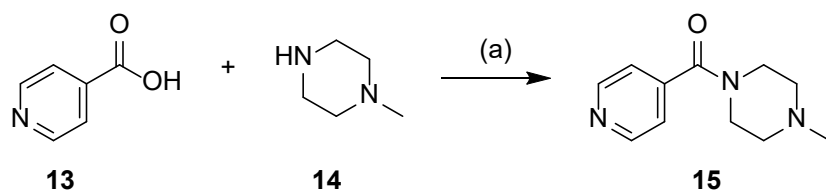


Fig 5. X-ray crystal structure and HSQC titrations of oxadiazole **8**.



Structure and biophysical characterisation of oxadiazole **8** binding to *EcDsbA*. (a) X-ray co-crystal structure of oxadiazole **8** bound to oxidised *EcDsbA*. Red dashed lines represent hydrogen-bond and polar interactions, black dashed lines show π -stacking interactions. Active site cysteines (C30 and C33) are shown as yellow spheres and oxadiazole **8** is shown as grey sticks. (b) Electron density map of oxadiazole **8**, 2mFo-DFc map contoured at 1 σ (blue) and mFo-DFc maps are shown at $\pm 3 \sigma$ (positive green and negative red). (c) Structure of oxadiazole **8** and affinity calculated by NMR with a single site binding model with ligand depletion. Displayed as $K_D \pm$ error of fit. (d) ^1H - ^{15}N HSQC spectra of oxadiazole **8** against oxidised $u\text{-}^{15}\text{N}$ *EcDsbA* at concentrations of 0, 100, 200, 400 and 500 μM shown in red, orange, light green, dark green and blue respectively. (e) ^1H - ^{15}N HSQC titrations of oxadiazole **8** plotting concentration (μM) vs CSP (ppm).

Fig 6. Faux background reactions for estimating minimum required conversion.



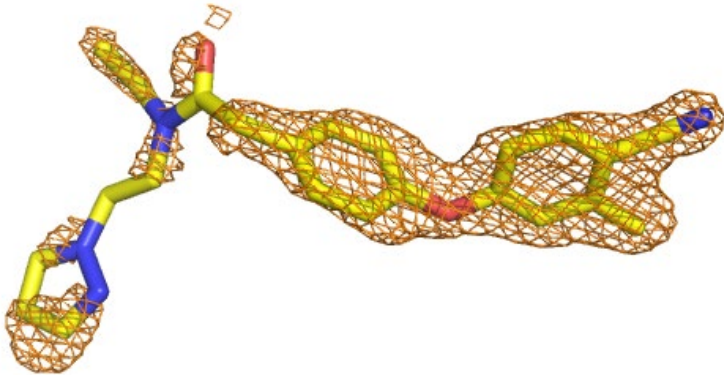
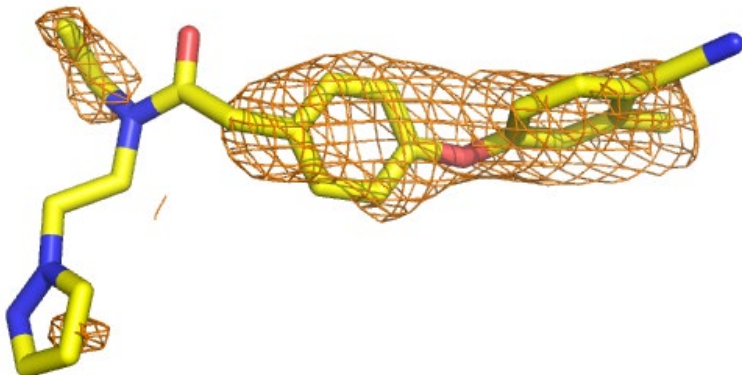
Isonicotinic acid (**13**) in 15% v/v dry Et₃N/MeCN (0.2 M, 125 μ L 1 eq., 25 μ mol) and a solution of HATU in dry DMF (1 M, 50 μ L, 2.0 eq., 50 μ mol) were combined and the reaction left at room temperature for 15 min without agitation. After this time, a solution of *N*-methylpiperazine (**14**) in DMF (0.5 M, 75 μ L, 1.5 eq., 38 μ mol) was added and the reaction left at room temperature for a further 18 h without agitation. 10 μ L aliquots of this reaction were dispensed into a 96 well plate as per table below and the solvents were removed under vacuum. Starting material phenylacetic acid **2** and pyrazole product **6** were then added to each well as per table below, mixed and soaked into *Ec*DsbA crystals (see methods).

Well	Starting material 2 (μ L)	Product 6 (μ L)	Background reaction (Y/N)	Conversion (%)
A1	9.4	0.6	Y	6
A2	8.7	1.3	Y	13
A3	7.2	2.8	Y	28
A4	9.4	0.6	N	6
A5	8.7	1.3	N	13
A6	7.2	2.8	N	28

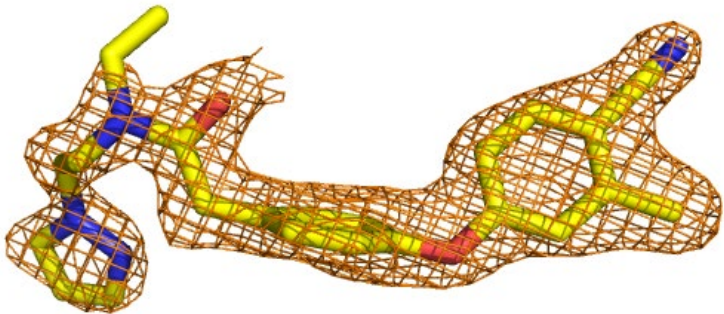
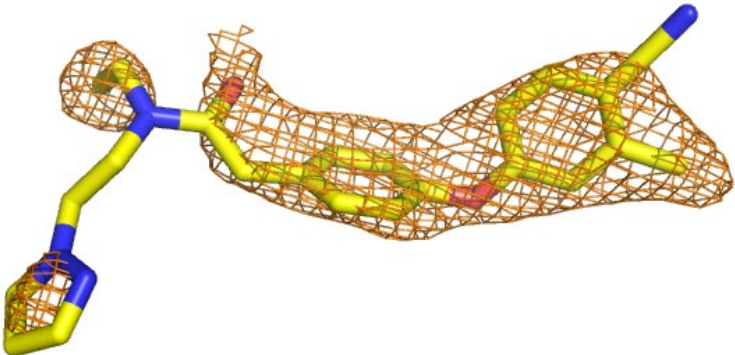
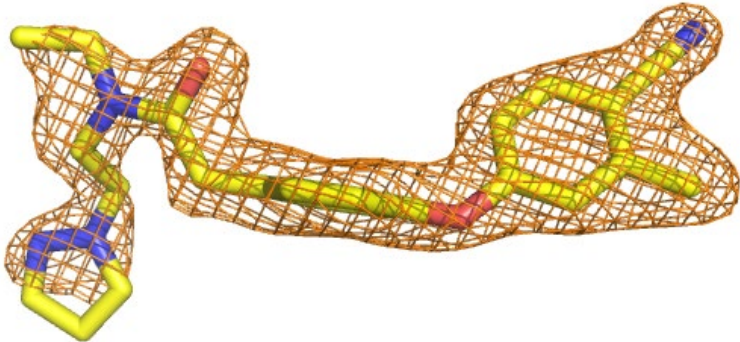
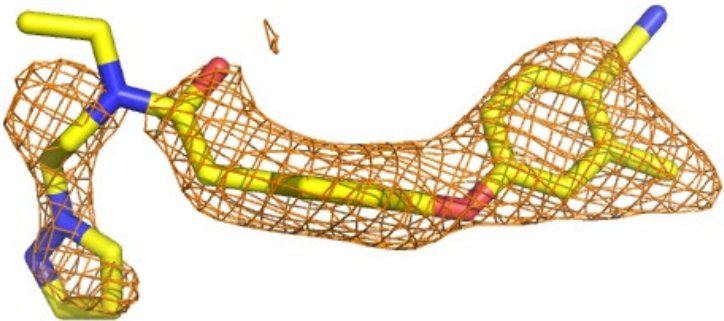
Compounds were made up as 100 mM stocks in DMSO.

Appendices

Fig 7. OMIT maps of pyrazole **6** in the presence and absence of faux background reactions at relative concentrations corresponding to 6, 13 and 28 % reaction conversion. Omit maps were generated by simulated annealing and are shown as 2mFo-DFc at 1 σ (orange). Ligand shown on OMIT densities yellow sticks for reference only.

Conversion (%)	OMIT density no background	OMIT density with background reaction
6		

Appendices

13		
28		

Experimental methods

Protein expression and purification

Unlabelled and ^{15}N -labelled *EcDsbA* were expressed as described previously using autoinduction.^[7] All media components were sterilised prior to use by autoclave or 0.22 μm filtration as appropriate. Pre-expression cultures were prepared by inoculating 10 mL Luria-Bertani broth supplemented with 50 $\mu\text{g}/\text{ml}$ kanamycin from frozen glycerol stocks of *E. coli* BL21(DE3) carrying the plasmid B0013-(5644bb) coding for *EcDsbA*, as described previously,^[8] and incubating for 16 h at 37 °C with agitation at 220 rpm. ZYM-5052 or N-5052 media supplemented with 50 $\mu\text{g}/\text{ml}$ kanamycin in baffled conical flasks were inoculated with 1% v/v of pre-culture and incubated for 24-30 h at 37 °C with agitation at 170 rpm, after which the cells were harvested by centrifugation at 3200 \times g at 4 °C for 20 min. The supernatant was discarded and the pellets were stored at -20 °C for subsequent protein extraction and purification.

Each frozen bacterial pellet was thawed, resuspended in an equal volume of osmotic shock buffer (20 mM Tris (pH 9.0), 10 mM EDTA, 50% w/v sucrose) and maintained at 4 °C with gentle stirring for 60 min, after which the cell suspension was rapidly diluted to 10 times its volume with deionised H_2O at 4 °C. The suspension was maintained at 4 °C and stirred for a further 90 min before the cell pellet and lysate were separated by centrifugation at 50000 \times g at 4 °C for 30 min. The lysate was carefully decanted from the pellet and stored at -20 °C. The cell pellet was resuspended in 10 times its volume of lysis buffer (20 mM Tris (pH 8.0), 25 mM NaCl, 4 mg/mL colistin sulfate) and maintained at room temperature with gentle stirring for 18-24 h. This suspension was then centrifuged at 50000 \times g at 4 °C for 30 min and the lysate was decanted and added to the thawed osmotic shock lysate. The pellet was discarded and the target protein was purified from the combined lysate. $(\text{NH}_4)_2\text{SO}_4$ was added to the combined lysate with gentle stirring to a concentration of 0.8 M and the solution was syringe filtered (0.22 μm , Millipore) for loading onto a HiLoad 1610 Phenyl Sepharose HP hydrophobic interaction chromatography column (GE Healthcare) using 20 mM Tris (pH 8.0), 50 mM NaCl, 1 M $(\text{NH}_4)_2\text{SO}_4$. The bound proteins were

eluted on a gradient from 1 – 0 M $(\text{NH}_4)_2\text{SO}_4$. Fractions were analysed by SDS-PAGE and those containing target protein were pooled and concentrated to 10 mL using an Amicon centrifugal diafiltration unit (10000 MWCO, Millipore) and buffer exchanged to 50 mM HEPES (pH 6.8), 50 mM NaCl using a HiPrep 2610 desalting column (GE Healthcare). Following buffer exchange, the fractions containing target protein were pooled and loaded onto a MonoQ HR 10/10 anion-exchange column (GE Healthcare). In these buffer conditions, pure target protein was collected in the flowthrough fraction and impurities were eluted using a gradient from 50 mM – 1 M NaCl. Protein purity was confirmed by SDS-PAGE.

Following purification by anion-exchange chromatography, each *EcDsbA* solution was concentrated to approximately 8 – 9 mL by diafiltration (Amicon 3000 MWCO, Millipore) and treated with 0.1 volumes of freshly prepared 15 mM copper-phenanthroline solution. After 1 h reaction time at 4 °C, the copper-phenanthroline was removed by buffer exchange (50 mM HEPES pH 6.8, 50 mM NaCl) using a HiPrep 2610 desalting column.

Protein samples were adjusted to their required final concentrations by concentration using Amicon centrifugal diafiltration units (3000 MWCO, Millipore) or dilution using storage/NMR buffer (50 mM HEPES (pH 6.8), 50 mM NaCl). For NMR, D_2O was added to a final concentration of 10% (v/v). Final protein concentrations were quantified using a NanoVue UV spectrophotometer (GE Healthcare) based on a calculated extinction coefficient for oxidized *EcDsbA* of $\epsilon = 28545 \text{ M}^{-1}\text{cm}^{-1}$.

Crystallisation, soaking, data collection and refinement of pure compounds

All the resynthesised pure compounds were co-crystallised with *EcDsbA*. For co-crystallisation experiments, compounds were solubilised in 100% DMSO to a final concentration of 100 mM. Prior to the crystallisation experiment, 95 μL of 30 mg/mL *EcDsbA* was mixed with 5 μL of a compound (where 5 mM is the final concentration) and incubated at 4 °C overnight. A previously identified crystallisation condition (100-300 mM KBr, 25-35% PEG 2000 MME) was used to obtain co-crystals. Briefly, 200 nL:200 nL of protein-compound mixture and crystallisation buffer drops were set up utilising sitting-drop vapor diffusion method using an automated robotic facility

in the CSIRO Collaborative Crystallisation Centre (www.csiro.au/C3), Melbourne, Australia. Plates were incubated at 18 °C and crystals appeared in 1-5 days. Prior to data collection crystals were briefly swiped through reservoir buffer and flash-cooled in liquid nitrogen.

Datasets were collected at the Australian Synchrotron, part of ANSTO, on MX1 and MX2 beamlines.^[2, 9] MX1 beamline was equipped with an ADSC Quantum 210r detector and MX2 with an EIGER × 16M pixel detector (Dectris Ltd). All datasets were processed with *XDS* or *IMOSFLM* and *AIMLESS*. Phasing was performed by molecular replacement with *Phaser* using previously solved structure of *EcDsbA* as a search model, PDB code 1FVK. The final structure was obtained after several rounds of manual refinement using *Coot* and refinement in *phenix.refine*.^[10] Data collection and refinement statistics are summarised at the end of this document. Structures, factors and coordinates have been deposited in the Protein Data Bank under the accession codes 6PBI, 6PC9, 6PD7 and 6PDH

Plate setup

In a 96-well plate, acid **2** in 15% v/v dry Et₃N/MeCN (0.2 M, 10 µL 1 eq., 2 µmol) and HATU in dry DMF (1 M, 4 µL, 2.0 eq., 4 µmol) were added to each well. The plate was sealed and left at room temperature for 15 min without agitation. After this time, amine in DMF (0.5 M, 6 µL, 1.5 eq., 3 µmol) was added to each well and the plate left at room temperature for a further 18 h without agitation. Some wells required Boc deprotection or ester hydrolysis using 1:1 DCM:TFA (20 µL) or THF (10 µL) and NaOH (aq., 0.6 M, 10 µL, 3.0 eq., 6 µmol). The plate was covered and left at room temperature for a further 3 h without agitation. The solvent was then removed under vacuum and 20 µL of DMSO added to each well. A 1 µL aliquot was taken from each well and diluted in MeCN (49 µL) to 2 mM and analysed on LCMS 1. The remaining DMSO samples was stored at -78 °C until crystallographic soaking.

Assessing solubility of synthesised analogues by q-NMR

Purified compounds had their solubility and aggregation evaluated via a 1D ^1H -qNMR titration in aqueous buffer. Compounds were accurately weighed and dissolved in d_6 -DMSO at 100 mM. D_2O buffer was prepared with 50 mM sodium phosphate, 25 mM NaCl at pH 7.4 with 100 μM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) internal standard. Serial dilution of the d_6 -DMSO stock was used to prepare a 4-point two-fold dilution series with final concentrations of 125, 250, 500 and 1000 μM with 2 % d_6 -DMSO in phosphate buffer and a total volume of 600 μL . NMR spectra were recorded on a Bruker 400 MHz spectrometer with $\text{TD} = 32 \text{ K}$, $\text{D1} = 5 \text{ s}$, 64 scans and water suppression at 4.7 ppm. NMR spectra were processed in MNOVA and referenced to the DSS peak at 0.030 ppm. Compounds failed solubility at a given concentration if they did not show a doubling in intensity concordant with doubling concentration or showed any chemical shift changes $>0.004 \text{ ppm}$ (1.6 Hz) from the lowest recorded concentration. Ligand concentrations were calculated relative to the DSS integral in the 1D ^1H -qNMR spectra and used to calculate the stock concentration which was used in K_D determination experiments.

Validation of hits and affinity determination by ^1H ^{15}N -HSQC

Hits from the REFIL_X screen were further validated by measuring CSP in ^1H - ^{15}N HSQC NMR. A reference ^1H - ^{15}N HSQC spectrum of uniformly labelled ^{15}N oxidised *EcDsbA* (100 μM protein, 2% d_6 -DMSO, 50 mM sodium phosphate, 25 mM NaCl, pH 6.8 in 10 % D_2O , 90% H_2O) was acquired. A series of ^1H - ^{15}N HSQC spectra were acquired under the same conditions in the presence of a sample fragment at and below its maximum solubility. CSP induced by fragment binding were calculated using the following equation

$$\text{CSP} = \sqrt{\Delta H^2 + (0.2 \times \Delta N)^2}$$

where ΔH and ΔN are the measured change in chemical shift of the backbone amide protons in the absence (*apo*) and presence of the fragment.

Backbone amide assignments were taken from the biological magnetic resonance data bank with a BMRB code of 19838. All assigned peaks showing $>0.040 \text{ ppm}$ shift within the ^1H - ^{15}N HSQC spectra

at their highest concentration were used for affinity determination. CSP below 0.023 ppm shift, ambiguous merging peaks, poorly fitting residues and residues with less than 4 data points were excluded. Affinities were determined using the following equation where [P] is protein concentration and [L] is ligand concentration.

$$Y = \frac{d_{max}}{2 \times [P]} \times ([P] + [L] + K_D) - \sqrt{([P] + [L] + K_D)^2 - 4 \times [P] \times [L]}$$

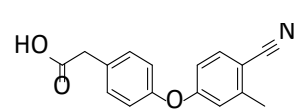
For general experimental see experimental section of the thesis.

Experimental and characterisation

General procedure 1: Acid chloride formation

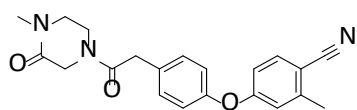
Oxalyl chloride (8 - 16 eq.) was dissolved in CDCl₃ (0.3 M) under N₂. DMF (cat. 2 - 3 drops) was added at 0 °C and the reaction mixture stirred for 30 min. After this time, acid (1.00 eq.) was added and the reaction stirred for a further 4 - 6 h at 0 °C. The reaction mixture was evaporated to dryness and the residue dissolved in dry DMF or CDCl₃ (0.3 M). Amine (1.2 - 2 eq.) and DIPEA or Et₃N (3 - 5 eq.) were added and the reaction mixture was stirred at room temperature for a further 18 - 24 h. The reaction mixture was diluted in ether (50 mL) and washed with hydrochloride (aq., 1 M, 2 × 50 mL), NaHCO₃ (aq., sat., 2 × 50 mL), brine (3 × 30 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetic acid (2)

 Methyl 2-(4-(4-cyano-3-methylphenoxy)phenyl)acetate (**16**) (4.56 g, 1.00 eq., 16.2 mmol) dissolved in MeOH (50 mL) and 2 M NaOH (aq., 30 mL, 3.70 eq., 60.0 mmol). The reaction was stirred for 3 h at room temperature. After this time, the MeOH was removed under vacuum and the aqueous layer acidified with hydrochloride (aq., 2 M, 50 mL) and diluted in EtOAc (150 mL). The layers were separated and the aqueous layer further extracted with EtOAc (2 × 100 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered and evaporated to dryness to give the product as a white solid (3.89 g, 90 %).

¹H NMR (401 MHz, *d*₆-DMSO) δ 12.39 (s, 1H), 7.72 (d, *J* = 8.6 Hz, 1H), 7.39 – 7.30 (m, 2H), 7.09 – 7.03 (m, 2H), 7.01 (d, *J* = 2.4 Hz, 1H), 6.86 (dd, *J* = 8.6, 2.4 Hz, 1H), 3.60 (s, 2H), 2.42 (s, 3H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 172.69, 160.98, 153.28, 144.37, 134.73, 131.86, 131.38, 119.98, 118.78, 117.97, 115.36, 105.75, 39.93, 20.03. **LCMS** (*m/z*): 265.9 [M+H]⁺, *t*_R = 3.31 min, standard method. **HRMS** (*m/z*): C₁₆H₁₃NO₃ requires 268.0968 [M+H]⁺; found 268.0965. **HPLC**: *t*_R = 5.74 min, >99 %, standard method.

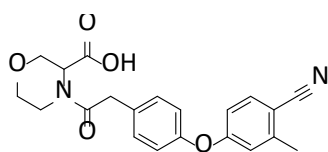
2-Methyl-4-(4-(2-(4-methyl-3-oxopiperazin-1-yl)-2-oxoethyl)phenoxy)benzonitrile (3)



Was prepared according to **General procedure 1** with oxalyl chloride (2.00 mL, 7.80 eq., 23.3 mmol), dry CDCl₃ (12 mL), **2** (800 mg, 1.00 eq., 3.00 mmol), 1-methylpiperazin-2-one (513 mg, 1.5 eq. 4.49 mmol) and DIPEA (2.00 mL, 3.75 eq., 11.2 mmol). The reaction mixture was diluted in ether (150 mL) and washed with hydrochloride (aq., 1 M, 2 × 150 mL), NaHCO₃ (aq., sat., 2 × 150 mL) and brine (3 × 50 mL), dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 10:90:0 pet spirits:EtOAc:MeOH to 0:90:10 pet spirits:EtOAc:MeOH give the product as an orange oil (57 mg, 5.2 %).

¹H NMR (401 MHz, CDCl₃) δ 7.51 (d, *J* = 8.5 Hz, 1H), 7.29 – 7.22 (m, 2H), 7.02 – 6.96 (m, 2H), 6.85 (d, *J* = 2.1 Hz, 1H), 6.80 (dd, *J* = 8.5, 1.9 Hz, 1H), 4.25 (s, 0.6H), 4.14 (s, 1.4H), 3.86 (app t, *J* = 5.4 Hz, 1.4H), 3.77 – 3.67 (m, 2.6H), 3.35 (app t, *J* = 5.4 Hz, 1.4H), 3.30 – 3.23 (m, 0.6H), 2.98 (m, 3.00 – 2.96, 3H), 2.47 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 169.38, 168.95, 165.48, 164.36, 161.31, 161.22, 154.15, 154.13, 144.55, 134.41, 130.96, 130.76, 130.69, 130.64, 120.68, 119.01, 118.98, 118.24, 118.19, 115.50, 115.48, 106.70, 106.59, 49.37, 48.31, 47.89, 46.41, 42.95, 39.91, 39.86, 38.90, 34.61, 34.40, 20.68. **¹H NMR and ¹³C NMR spectra were consistent with rotamers in a 2:1 ratio.** **LCMS 2** (*m/z*): 364.0 [M+H]⁺, *t*_R = 4.24 min, standard method. **HRMS** (*m/z*): C₂₁H₂₁N₃O₃ requires 364.1656 [M+H]⁺; found 364.1666. **HPLC**: *t*_R = 5.72 min, >99 %, standard method.

4-(2-(4-(4-Cyano-3-methylphenoxy)phenyl)acetyl)morpholine-3-carboxylic acid (4)

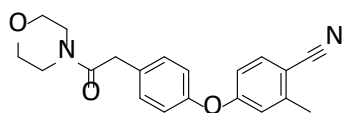


Prepared and worked up according to **General procedure 1** with oxalyl chloride (1.00 mL, 15.6 eq., 11.7 mmol), dry CDCl_3 (2.5 mL), **2** (200 mg, 1.00 eq., 0.75 mmol), dry DMF (3 mL), morpholine-3-carboxylic acid

(118 mg, 1.2 eq. 0.90 mmol) and DIPEA (666 μL , 5 eq., 3.75 mmol). The residue was purified by flash column chromatography using a gradient of 75:25:0 pet spirits:EtOAc:AcOH to 74.5:25:0.5 pet spirits:EtOAc:AcOH to give the product as a colourless oil (10 mg, 3.51 %).

^1H NMR (401 MHz, CD_3OD) δ 7.50 (d, J = 8.6 Hz, 1H), 7.28 – 7.22 (m, 1.4H), 7.22 – 7.17 (m, 0.6H), 6.97 – 6.89 (m, 2H), 6.85 (d, J = 2.0 Hz, 1H), 6.77 (dd, J = 8.6, 2.4 Hz, 1H), 4.86 (s, 0.7H), 4.57 (s, 0.3H), 4.32 (s, 0.4H), 4.29 (s, 0.6H), 4.10 (d, J = 12.7 Hz, 0.4H), 3.81 (dd, J = 11.6, 3.4 Hz, 0.4H), 3.76 (s, 1.5H), 3.75 – 3.64 (m, 1.5H), 3.62 – 3.50 (m, 1.4H), 3.48 – 3.30 (m, 1.8H), 2.36 (s, 3H). **^{13}C NMR** (101 MHz, CD_3OD) δ 173.93, 163.19, 163.16, 155.26, 155.19, 145.77, 135.57, 133.04, 132.97, 132.30, 132.14, 121.54, 119.86, 119.84, 118.96, 118.94, 116.46, 116.44, 107.23, 107.19, 68.80, 68.67, 67.62, 67.36 (two peaks confirmed by HSQC), 67.29, 44.84, 40.83, 40.10, 39.83, 20.51. **^1H NMR and ^{13}C NMR spectra were consistent with rotamers in a 2:1 ratio. LCMS** 2 (m/z): 381.0 $[\text{M}+\text{H}]^+$, t_{R} = 3.82 min, standard method. **HRMS** (m/z): $\text{C}_{21}\text{H}_{20}\text{N}_2\text{O}_5$ requires 381.1445 $[\text{M}+\text{H}]^+$; found 381.1458. **HPLC**: t_{R} = 5.69 min, >99 %, standard method.

2-Methyl-4-(4-(2-morpholino-2-oxoethyl)phenoxy)benzonitrile (5)



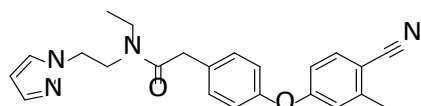
To a solution of **2** (100 mg, 1.00 eq., 0.37 mmol), and HATU (213 mg, 1.5 eq., 0.56 mmol) dissolved in dry 1:1 MeCN:DMF (4 mL) under

nitrogen was added Et_3N (0.27 mL, 5.3 eq., 1.98 mmol). The reaction mixture was stirred for 15 min at room temperature. After this time, morpholine (65 μL , 2.0 eq., 0.75 mmol) was added and the reaction stirred for a further 24 h at room temperature. The reaction was then diluted in ether (40 mL) and H_2O (50 mL). The layers were separated and the organic layer washed with hydrochloride (aq., 0.1 M, 50 mL) and brine (2×50 mL), dried over anhydrous Na_2SO_4 , filtered and evaporated to

dryness. The residue was purified by flash column chromatography using a gradient of 25:75 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give the product as a white solid (38 mg, 30 %).

¹H NMR (401 MHz, *d*₆-DMSO) δ 7.76 (d, *J* = 8.6 Hz, 1H), 7.33 – 7.27 (m, 2H), 7.09 – 7.04 (m, 2H), 7.02 (d, *J* = 2.4 Hz, 1H), 6.87 (dd, *J* = 8.6, 2.6 Hz, 1H), 3.74 (s, 2H), 3.57 – 3.50 (m, 6H), 3.48 – 3.44 (m, 2H), 2.44 (s, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 169.50, 161.34, 153.95, 144.54, 134.41, 131.56, 130.63, 120.62, 118.95, 118.22, 115.44, 106.58, 66.88, 66.56, 46.54, 42.28, 39.76, 20.69. **LCMS 1** (*m/z*): 337.15 [M+H]⁺, *t*_R = 2.11 min. **HRMS** (*m/z*): C₂₀H₂₀N₂O₃ requires 337.1547 [M+H]⁺; found 337.1555. **HPLC**: *t*_R = 5.93 min, >99 %, standard method.

***N*-(2-(1*H*-pyrazol-1-yl)ethyl)-2-(4-(4-cyano-3-methylphenoxy)phenyl)-*N*-ethylacetamide (6)**



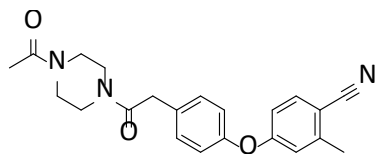
Prepared and worked up according to **General procedure 1** with oxalyl chloride (1.00 mL, 15.6 eq., 11.7 mmol), dry CDCl₃ (2.5 mL), **2** (200 mg, 1.00 eq., 0.75 mmol), dry DMF (3 mL), *N*-ethyl-2-pyrazol-1-yl-ethanamine (125 mg, 1.2 eq. 0.90 mmol) and DIPEA (400 μL, 3.0 eq., 2.24 mmol). The residue was purified by flash column chromatography using 75:25 pet spirits:EtOAc to give the product as an orange oil (74 mg, 25 %).

¹H NMR (401 MHz, *d*₆-DMSO) δ 7.77 (d, *J* = 6.7 Hz, 0.5H), 7.75 (d, *J* = 6.7 Hz, 0.5H), 7.73 (d, *J* = 1.7 Hz, 0.5H), 7.60 (d, *J* = 1.7 Hz, 0.5H), 7.54 (d, *J* = 1.2 Hz, 0.5H), 7.45 (d, *J* = 1.2 Hz, 0.5H), 7.35 – 7.29 (m, 1H), 7.15 – 7.10 (m, 1H), 7.10 – 7.06 (m, 1H), 7.06 – 7.00 (m, 2H), 6.90 – 6.84 (m, 1H), 6.30 (t, *J* = 2.0 Hz, 0.5H), 6.20 (t, *J* = 2.0 Hz, 0.5H), 4.32 (t, *J* = 5.8 Hz, 1H), 4.25 (t, *J* = 6.2 Hz, 1H), 3.74 – 3.68 (m, 2H), 3.61 (t, *J* = 6.1 Hz, 1H), 3.24 (q, *J* = 7.2 Hz, 1H), 3.20 (s, 1H), 3.06 (q, *J* = 7.2 Hz, 1H), 2.45 (s, 1.5H), 2.44 (s, 1.5H), 1.01 (t, *J* = 7.1 Hz, 1.5H), 0.91 (t, *J* = 7.1 Hz, 1.5H). **¹³C NMR** (101 MHz, CD₃OD) δ 173.54, 173.32, 163.12, 163.12, 155.26, 155.16, 145.75, 145.73, 141.15, 140.64, 135.57, 135.54, 133.44, 133.41, 132.47, 132.10, 132.06, 131.96, 121.61, 121.47, 119.82, 119.81, 118.94, 118.93, 116.44, 116.43, 107.26, 107.20, 107.16, 106.73, 50.97, 50.35, 48.83, 47.58, 44.89, 41.64, 40.33, 39.46, 20.54, 20.51, 13.88, 12.66. **¹H NMR and ¹³C NMR spectra were consistent with rotamers in a 1:1 and 3:2 ratio respectively.** **LCMS 2** (*m/z*): 389.0 [M+H]⁺, *t*_R =

3.79 min, standard method. **HRMS** (m/z): $C_{23}H_{24}N_4O_2$ requires 411.1791 $[M+Na]^+$; found 411.1793.

HPLC: t_R = 6.35 min, >99 %, standard method.

4-(4-(2-(4-Acetylpiperazin-1-yl)-2-oxoethyl)phenoxy)-2-methylbenzonitrile (7)

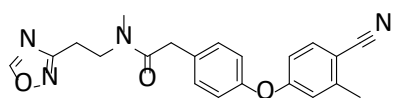


Was prepared according to **General procedure 1** with oxalyl chloride (1.0 mL, 7.8 eq., 11.6 mmol), dry $CDCl_3$ (5 mL), **2** (400 mg, 1.0 eq., 1.50 mmol), dry $CDCl_3$ (5 mL), *1*-piperazin-1-ylethanone (384 mg,

2.0 eq. 3.00 mmol) and Et_3N (626 μ L, 3 eq., 4.49 mmol). The reaction mixture was diluted in CH_2Cl_2 (50 mL) and washed with H_2O (2×50 mL), hydrochloride (aq., 1 M, 2×50 mL), brine (30 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using 90:10 EtOAc:MeOH to give the product as a white solid (359 mg, 64 %).

1H NMR (401 MHz, d_6 -DMSO) δ 7.76 (d, J = 8.6 Hz, 1H), 7.34 – 7.29 (m, 2H), 7.11 – 7.05 (m, 2H), 7.02 (s, 1H), 6.88 (dd, J = 8.6, 2.5 Hz, 1H), 3.78 (d, J = 4.1 Hz, 2H), 3.57 – 3.47 (m, 3H), 3.47 – 3.37 (m, 5H), 2.44 (s, 3H), 2.02 (s, 3H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 169.74, 169.45, 169.42, 169.09, 161.29, 154.14, 144.61, 134.46, 131.32, 130.65, 130.58, 120.67, 119.08, 118.24, 115.53, 106.71, 46.17, 46.09, 45.99, 45.83, 41.92, 41.70, 41.34, 41.21, 40.20, 40.02, 21.47, 21.45, 20.74. **LCMS 2** (m/z): 378.0 $[M+H]^+$, t_R = 3.32 min, standard method. **HRMS** (m/z): $C_{22}H_{23}N_3O_3$ requires 378.1812 $[M+H]^+$; found 378.1820. **HPLC**: t_R = 5.80 min, >99 %, standard method.

2-(4-(4-Cyano-3-methylphenoxy)phenyl)-*N*-methyl-*N*-(2-(5-methyl-1,2,4-oxadiazol-3-yl)ethyl)acetamide (8)



Was prepared and worked up according to **General procedure 1** with oxalyl chloride (1.0 mL, 7.8 eq., 11.6 mmol), dry $CDCl_3$ (2.5

mL), **2** (200 mg, 1.0 eq., 0.75 mmol), *N*-methyl-2-(5-methyl-1,2,4-oxadiazol-3-yl)ethanamine hydrochloride (160 mg, 1.2 eq. 0.90 mmol) and Et_3N (521 μ L, 5 eq., 3.74 mmol). The residue was purified by flash column chromatography using a gradient of 25:75 pet spirits:EtOAc to 0:100 pet spirits:EtOAc to give the product as a colourless oil (50 mg, 17 %).

¹H NMR (401 MHz, *d*₆-DMSO) δ 7.75 (app dd, *J* = 8.6, 2.1 Hz, 1H), 7.29 – 7.22 (m, 2H), 7.08 – 7.03 (m, 2H), 7.02 (app t, *J* = 2.7 Hz, 1H), 6.88 (dd, *J* = 8.6, 2.5 Hz, 1H), 3.75 (t, *J* = 7.2 Hz, 0.8H), 3.70 (app d, *J* = 3.7 Hz, 2H), 3.62 (t, *J* = 7.2 Hz, 1.2H), 3.01 (s, 1.8H), 2.93 (t, *J* = 7.1 Hz, 0.8H), 2.86 (t, *J* = 7.2 Hz, 1.2H), 2.83 (s, 1.2H), 2.57 (s, 1.2H), 2.54 (s, 1.8H), 2.44 (app d, *J* = 2.1 Hz, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 176.92, 176.48, 170.93, 170.81, 168.54, 167.60, 161.47, 161.46, 153.85, 153.80, 144.52, 134.39, 131.88, 131.66, 130.82, 130.81, 120.60, 120.53, 118.88, 118.87, 118.27, 115.41, 115.41 (**confirmed via HMBC**), 106.47, 47.46, 46.12, 40.21, 39.68, 36.58, 33.34, 25.33, 24.43, 20.68, 12.39, 12.37. **¹H NMR and ¹³C NMR spectra were consistent with rotamers in a 3:2 ratio.** **LCMS 2** (*m/z*): 391.0 [M+H]⁺, *t*_R = 3.41 min, standard method. **HRMS** (*m/z*): C₂₂H₂₂N₄O₃ requires 391.1765 [M+H]⁺; found 391.1774. **HPLC**: *t*_R = 6.12 min, >99 %, standard method.

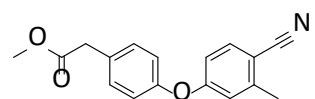
2-(4-(4-Cyano-3-methylphenoxy)phenyl)-N-(2-(4-methyl-1,4-diazepan-1-yl)ethyl)acetamide (9)

Oxalyl chloride (1.00 mL, 7.8 eq., 11.7 mmol) was dissolved in dry CDCl₃ (5 mL). DMF (cat. 2 - 3 drops) was added at 0 °C and the reaction mixture stirred for 30 min under N₂. After this time, **2** (200 mg, 1.00 eq., 0.94 mmol) was added and the reaction stirred for a further 4 h at 0 °C. 2-(4-Methyl-1,4-diazepan-1-yl)ethanamine (282 mg, 1.2 eq. 1.8 mmol) and Et₃N (1000 μL, 4.8 eq., 7.2 mmol) were added and the reaction stirred at room temperature for a further 24 h. After this time, the reaction mixture was diluted in C hydrochloride₃ (40 mL) and washed with NaHCO₃ (aq., sat., 2 × 20 mL) and dried over anhydrous Na₂SO₄, filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 EtOAc:MeOH to 0:100 EtOAc:MeOH. The silica was resuspended in 2:98 Et₃N:MeOH, filtered and evaporated to dryness to give the product as a colourless oil (65 mg, 11 %).

¹H NMR (401 MHz, *d*₆-DMSO) δ 7.92 (t, *J* = 5.3 Hz, 1H), 7.75 (d, *J* = 8.6 Hz, 1H), 7.38 – 7.31 (m, 2H), 7.09 – 7.04 (m, 2H), 7.02 (d, *J* = 2.2 Hz, 1H), 6.88 (dd, *J* = 8.6, 2.4 Hz, 1H), 3.44 (s, 2H), 3.13 (dd, *J* = 12.5, 6.5 Hz, 2H), 2.68 – 2.59 (m, 4H), 2.55 – 2.46 (m, 6H), 2.44 (s, 3H), 2.23 (s, 3H), 1.67 (dt, *J* = 11.8, 6.0 Hz, 2H). **¹³C NMR** (101 MHz, *d*₆-DMSO) δ 169.82, 161.02, 152.97, 144.30, 134.70,

133.31, 130.87, 119.95, 118.64, 117.91, 115.28, 105.59, 57.41, 56.59, 56.17, 53.96, 53.67, 46.46, 41.64, 37.03, 26.86, 19.98. **LCMS** (m/z): 407.1 $[M+H]^+$, t_R = 3.25 min, standard method. **HRMS** (m/z): $C_{24}H_{30}N_4O_2$ requires 407.2442 $[M+H]^+$; found 407.2454. **HPLC**: t_R = 4.726 min, >99 %, standard method.

Methyl 2-(4-(4-cyano-3-methylphenoxy)phenyl)acetate (16)



4-Bromo-2-methylbenzonitrile (10.0 g, 1.00 eq., 51 mmol), methyl 2-(4-hydroxyphenyl)acetate (17.0 g, 2.0 eq., 102 mmol), Cs_2CO_3 (33.2 g, 2.0 eq., 102 mmol), *N,N*-dimethylglycine hydrochloride (2.85 g, 0.20 eq., 20.4 mmol) and copper(I) iodide (1.94 g, 0.10 eq., 10.2 mmol) were sealed in a 250 mL round bottom flask and vacuum purged with nitrogen (3 times). Degassed dry DMF (60 mL) was added and the reaction mixture heated at 100 °C for 18 h. The reaction was then cooled to room temperature, diluted with H_2O (200 mL) and Et_2O (300 mL). The layers were separated and the aqueous layer further extracted with Et_2O (2×100 mL). The combined organic layers were washed with H_2O (2×200 mL), brine (2×125 mL) and dried over anhydrous Na_2SO_4 , filtered and evaporated to dryness. The residue was purified by flash column chromatography using a gradient of 100:0 toluene:acetone to 90:10 toluene:acetone to give the product as a colourless oil (7.74 g, 54 %).

1H NMR (401 MHz, d_6 -DMSO) δ 7.73 (d, J = 8.6 Hz, 1H), 7.39 – 7.31 (m, 2H), 7.11 – 7.03 (m, 2H), 7.02 (d, J = 2.2 Hz, 1H), 6.87 (dd, J = 8.6, 2.3 Hz, 1H), 3.71 (s, 2H), 3.63 (s, 3H), 2.43 (s, 3H). **^{13}C NMR** (101 MHz, d_6 -DMSO) δ 171.51, 160.81, 153.41, 144.31, 134.68, 131.29, 131.06, 119.96, 118.79, 117.87, 115.37, 105.74, 51.69, 39.26, 19.94. **LCMS** (m/z): 281.9 $[M+H]^+$, t_R = 3.59 min, standard method. **HRMS** (m/z): $C_{17}H_{15}NO_3$ requires 282.1125 $[M+H]^+$; found 282.1124. **HPLC**: t_R = 6.74 min, >99 %, standard method.

Appendices

Table 4: Table 1 statistics for all crude and purified analogues.

Compound #	2 pure	3 crude	3 pure	4 crude	4 pure	5 crude	5 pure	6 crude	6 pure	8 pure
Wavelength		0.9537	0.9537	0.9537	0.9537					
Resolution range	32.25 - 1.99 (2.061 - 1.99)	30.96 - 2.005 (2.077 - 2.005)	38.36 - 2.301 (2.383 - 2.301)	32.41 - 1.896 (1.964 - 1.896)	43.19 - 1.921 (1.99 - 1.921)	34.72 - 1.907 (1.975 - 1.907)	34.79 - 1.899 (1.967 - 1.899)	32.21 - 2.007 (2.079 - 2.007)	36.89 - 1.96 (2.03 - 1.96)	32.14 - 1.931 (2 - 1.931)
Space group	C 1 2 1	C 1 2 1	P 2 ₁ 2 ₁ 21	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1
Unit cell	116.77 64.5 74.21 90 125.89 90	117.996 64.41 74.433 90 125.854 90	60.534 76.711 86.205 90 90 90	117.608 64.827 74.131 90 125.93 90	119.339 48.822 91.538 90 117.029 90	117.674 63.912 74.455 90 125.693 90	117.734 64.624 74.404 90 125.783 90	117.979 64.426 74.434 90 125.87 90	119.214 50.295 91.98 90 118.113 90	117.745 64.286 74.444 90 125.907 90
Total reflections		59247 (5680)	36581 (3495)	68576 (6772)	71591 (6708)	68616 (6778)	70038 (6888)	58770 (5721)	68987 (6832)	66904 (6018)

Appendices

Unique reflections	30744 (3070)	30129 (2888)	18314 (1767)	35343 (3447)	35960 (3428)	34979 (3358)	35586 (3485)	30046 (2864)	34708 (3421)	33580 (3055)
Multiplicity		2.0 (2.0)	2.0 (2.0)	1.9 (1.9)	2.0 (2.0)	2.0 (2.0)	2.0 (2.0)	2.0 (2.0)	2.0 (2.0)	2.0 (2.0)
Completeness (%)	99.78 (99.84)	97.31 (96.04)	99.41 (96.61)	97.37 (96.66)	99.46 (96.10)	97.15 (96.58)	99.19 (97.78)	96.81 (95.71)	99.43 (99.65)	98.77 (90.05)
Mean I/sigma(I)		9.85 (1.40)	18.88 (3.47)	10.68 (1.33)	13.44 (2.66)	13.70 (1.62)	17.37 (2.08)	12.31 (1.52)	11.10 (3.20)	11.20 (1.94)
Wilson B-factor	30.24	36.94	44.36	33.06	28.36	32.16	31.48	37.18	32.97	37.69
R-merge		0.02813 (0.3865)	0.01842 (0.1985)	0.02372 (0.3865)	0.02854 (0.2388)	0.02319 (0.3547)	0.02309 (0.3694)	0.02719 (0.4031)	0.02493 (0.1312)	0.02508 (0.2593)
R-meas		0.03979 (0.5466)	0.02606 (0.2807)	0.03355 (0.5467)	0.04036 (0.3377)	0.03279 (0.5016)	0.03265 (0.5224)	0.03846 (0.57)	0.03526 (0.1855)	0.03547 (0.3668)
R-pim		0.02813 (0.3865)	0.01842 (0.1985)	0.02372 (0.3865)	0.02854 (0.2388)	0.02319 (0.3547)	0.02309 (0.3694)	0.02719 (0.4031)	0.02493 (0.1312)	0.02508 (0.2593)
CC1/2		0.999 (0.8)	0.999 (0.915)	0.999 (0.785)	0.999 (0.915)	0.999 (0.805)	1 (0.785)	0.999 (0.744)	0.999 (0.98)	0.999 (0.925)
CC*		1 (0.943)	1 (0.978)	1 (0.938)	1 (0.977)	1 (0.944)	1 (0.938)	1 (0.924)	1 (0.995)	1 (0.98)

Appendices

Reflections used in refinement	30733 (3068)	29657 (2886)	18308 (1766)	34977 (3447)	35924 (3427)	34112 (3358)	35578 (3485)	29407 (2857)	34601 (3411)	33540 (3050)
Reflections used for R-free	1526 (162)	959 (74)	888 (101)	944 (77)	1836 (191)	1012 (83)	1737 (149)	964 (74)	1715 (151)	1582 (140)
R-work	0.1782 (0.2521)	0.2008 (0.3078)	0.1768 (0.2351)	0.1932 (0.2865)	0.1874 (0.3105)	0.1844 (0.2753)	0.1905 (0.2746)	0.2129 (0.3422)	0.1986 (0.2975)	0.1979 (0.3072)
R-free	0.2167 (0.2674)	0.2737 (0.3182)	0.2427 (0.3007)	0.2273 (0.3507)	0.2314 (0.3909)	0.2232 (0.3310)	0.2282 (0.2982)	0.2863 (0.3801)	0.2450 (0.3366)	0.2274 (0.3787)
CC(work)		0.963 (0.830)	0.968 (0.876)	0.964 (0.836)	0.968 (0.895)	0.968 (0.861)	0.961 (0.853)	0.958 (0.772)	0.965 (0.939)	0.967 (0.885)
CC(free)		0.908 (0.778)	0.943 (0.881)	0.941 (0.689)	0.955 (0.874)	0.948 (0.885)	0.957 (0.836)	0.895 (0.739)	0.936 (0.873)	0.963 (0.795)
Number of non- hydrogen atoms	3260	3184	3021	3225	3282	3258	3244	3146	3051	3115
macromolecules	2924	2931	2886	2895	2951	2937	2912	2910	2894	2902

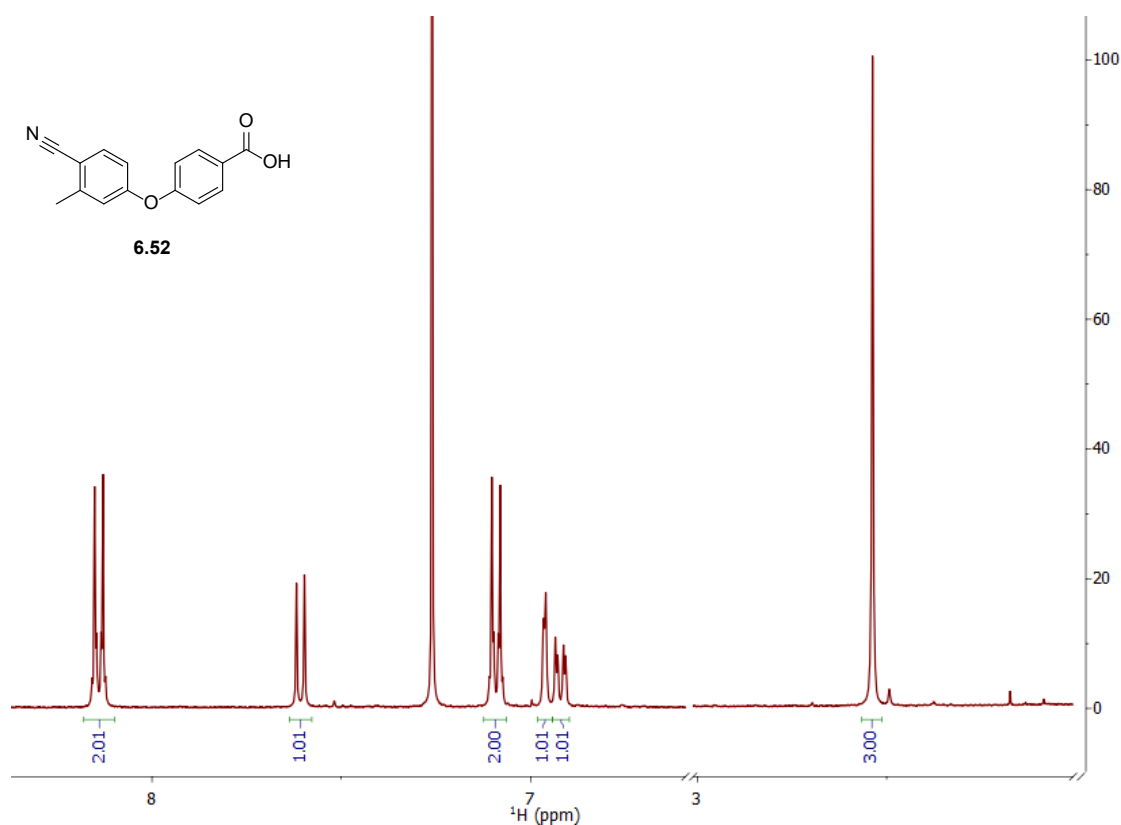
Appendices

ligands	46	28	37	58	28	26	26	30	39	30
solvent	290	225	98	272	303	295	306	206	118	183
Protein residues	376	376	375	376	376	376	376	376	374	376
RMS(bonds)	0.007	0.009	0.014	0.008	0.009	0.007	0.007	0.007	0.006	0.009
RMS(angles)	0.76	0.84	0.94	0.81	0.76	0.76	0.79	0.87	0.76	0.74
Ramachandran favoured (%)	97.31	97.04	97.84	97.85	98.66	98.12	96.77	96.77	98.11	97.58
Ramachandran allowed (%)	2.69	2.96	2.16	2.15	1.34	1.88	2.69	2.96	1.89	2.42
Ramachandran outliers (%)	0	0	0	0	0	0	0.54	0.27	0	0
Rotamer outliers (%)	0.97	0.66	1.02	0	0.65	1.65	0	1.01	0.34	0
Clashscore	3.97	3.12	4.76	3.16	1.9	3.47	3.34	4.37	3.16	2.81
Average B- factor	42.03	49.23	53.97	46.29	42.15	41.79	42.11	50	55.32	53.62

Appendices

macromolecules	41.49	49.15	54.06	45.68	41.88	41.18	41.72	50	55.54	53.51
ligands	59.14	71.41	55.61	69.36	45.28	67.11	51.4	65.47	54.6	77.58
solvent	44.74	47.48	50.74	47.83	44.59	45.66	45.02	47.73	50.17	51.52

Appendix 6.1: ^1H NMR spectra of benzoic acid **6.52** in CDCl_3 .



References

- [1] J. L. Martin, G. Waksman, J. C. A. Bardwell, J. Beckwith, J. Kuriyan, *J. Mol. Biol.* **1993**, 230, 1097-1100.
- [2] N. P. Cowieson, D. Aragao, M. Clift, D. J. Ericsson, C. Gee, S. J. Harrop, N. Mudie, S. Panjekar, J. R. Price, A. Riboldi-Tunncliffe, R. Williamson, T. Caradoc-Davies, *J. Synchrotron Radiat.* **2015**, 22, 187-190.
- [3] P. R. Evans, G. N. Murshudov, *Acta Crystallogr. D Biol. Crystallogr.* **2013**, 69, 1204-1214.
- [4] A. P. Karplus, K. Diederichs, *Science* **2012**, 336, 1030.
- [5] S. Panjekar, V. Parthasarathy, V. S. Lamzin, M. S. Weiss, P. A. Tucker, *Acta Crystallogr. D Biol. Crystallogr.* **2005**, 61, 449-457.
- [6] N. M. Pearce, T. Krojer, A. R. Bradley, P. Collins, R. P. Nowak, R. Talon, B. D. Marsden, S. Kelm, J. Shi, C. M. Deane, F. von Delft, *Nat. Commun.* **2017**, 8, 15123.

- [7] aL. A. Adams, P. Sharma, B. Mohanty, O. V. Ilyichova, M. D. Mulcair, M. L. Williams, E. C. Gleeson, M. Totsika, B. C. Doak, S. Caria, K. Rimmer, J. Horne, S. R. Shouldice, M. Vazirani, S. J. Headey, B. R. Plumb, J. L. Martin, B. Heras, J. S. Simpson, M. J. Scanlon, *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 2179-2184; bF. W. Studier, *Protein Expr. Purif.* **2005**, *41*, 207-234.
- [8] J. J. Paxman, N. A. Borg, J. Horne, P. E. Thompson, Y. Chin, P. Sharma, J. S. Simpson, J. Wielens, S. Piek, C. M. Kahler, H. Sakellaris, M. Pearce, S. P. Bottomley, J. Rossjohn, M. J. Scanlon, *J. Biol. Chem.* **2009**, *284*, 17835-17845.
- [9] D. Aragão, J. Aishima, H. Cherukuvada, R. Clarken, M. Clift, N. P. Cowieson, D. J. Ericsson, C. L. Gee, S. Macedo, N. Mudie, S. Panjekar, J. R. Price, A. Riboldi-Tunnicliffe, R. Rostan, R. Williamson, T. T. Caradoc-Davies, *J. Synchrotron Radiat.* **2018**, *25*, 885-891.
- [10] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66*, 213-221.
- [11] H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, *62*, 7512-7515.