Mechanism of action studies of the utrophin modulator ezutromid for the treatment of Duchenne muscular dystrophy

Isabel Wilkinson

Thesis submitted for the degree of Doctor of Philosophy

Pembroke College, University of Oxford

September 2019

Declaration of Authorship

The work presented in this thesis is my own, except where otherwise stated. No part of this thesis has been submitted for any other degree at this or any other university.

Isabel Wilkinson

September 2019

Abstract

Duchenne muscular dystrophy (DMD) is a muscle wasting disease arising from mutations in the dystrophin gene, affecting about 1 in 5000 boys. Whilst there is currently no available cure for DMD, a number of therapeutic strategies are under development. One such therapy aims to promote expression of utrophin, an autosomal paralogue of dystrophin. Utrophin upregulation rescues the dystrophic phenotype in disease model mice. A novel small molecule utrophin modulator, ezutromid (Summit Therapeutics; formerly SMT C1100), which was identified using a phenotypic screen, progressed to Phase II clinical trials in DMD patients. Interim 24-week data demonstrated reduced muscle fibre damage and increased levels of utrophin, providing the first evidence of ezutromid target engagement and proof of mechanism. However, these effects were not seen after the full 48 weeks of the trial. This work aims to define the mechanism of ezutromid in order to help understand the trial results, and to aid development of new generations of utrophin modulators.

Target identification through affinity based protein profiling (AfBPP) has been carried out in this work alongside phenotypic profiling studies. AfBPP requires bioactive analogues suitable for affinity purification. Biotinylated and photoaffinity-labelled analogues of ezutromid designed to be positive and negative control probes have been synthesised and bioactivity confirmed. These probes were used in AfBPP experiments with disease model mouse myoblasts followed by LC-MS/MS. The results from this experiment and phenotype profiling converged on the aryl hydrocarbon receptor (AhR) as a target of ezutromid. Target validation studies demonstrated that ezutromid binds AhR potently, behaving as an AhR antagonist. Moreover, reported AhR antagonists also upregulate utrophin, showing that this pathway could be exploited in future DMD therapies.

Acknowledgments

I would like to thank Angie for her help and guidance during my PhD and for giving me the freedom and support to take make this project my own. I also want to thank my funding sources, the EPSRC and MRC, without whom this would not have been possible. I am very grateful to Kay and her group members for their advice and help with biological experiments, and especially Adam for introducing me to the world of western blots. Thanks must go to Shabaz for my proteomics training, and to Svenja, Sabrina and Elisabete for guiding me through my mass spec experiments.

I am extremely grateful to Fraydoon for his excellent timing in relocating to Oxford; the AhR protein complex he gave me felt like a precious gift. I would also like to thank Elliot and Frank and his other group members for their training, patience and assistance in the luciferase crystallisation project, even though we never got the result we hoped for. Next, I would like to thank Kevin Henbest and Chris Timmel for giving me time on their fluorescence spectrophotometer and help with my experiments. Finally, I want to thank Kathryn Pugh and Kilian Huber for their help with the CETSA experiments.

DMD project

To Aini, my DMD sister! Thanks for blazing the trail, giving me write-up goals and always being there for a chat. Thank you Jess Reynolds for starting the probe synthesis on C1100 and helping me hit the ground running with my PhD. Thank you Adam for being such a great summer student, and for your highly memorable pavlova and ice cream. I also want to thank Maria, Dan, Nicky and Graham for their helpful advice and contributions over the years.

AJR group

Thanks to everyone in the group for making it a fun and fruitful workplace for the last three years, I'm grateful for all your help along the way. To Carole, thank you for your encouragement and entertaining stories. Thank you also for your leadership in the lab, we're lucky to have you. Thank you Seb for being my career counsellor and keeping me happy with your terrible jokes. If I ever get the human luciferase project off the ground, you'll be the first to know. To Laura, thanks for the fun times we had while you were with us and for passing on the baton, I'll try to make you proud. Thanks Liam for doing the heavy-lifting with proofreading my thesis, I'm mega grateful.

To the DPhil girls, thanks for all the cake and good times (and all the nice things you said about my burnt contributions):

– Fern, we miss you, thanks for instigating the cake and coffee tradition.
– Purnima, the champion in the western blot arena, may the odds be ever in your favour.
– Rachel, you're the boss now. Keep forging the target ID path!
– Ria, the original, thanks for always being on my wavelength.
– and Alice, the best chemist and the most dedicated - never change. And thanks for inspiring me to get into the gym, I haven't looked back.

I'm wishing you girls love and all the best for the rest of your PhDs and I can't wait to read about your achievements.

To my wonderful friends and family, I'm #blessed to have you in my life. Thanks for being my inspiration, sharing my happiness and encouraging me across the finish line. To my parents, thank you for your love and support, and for trying to fix all my problems. I'll never forget your offer to come to lab and do my washing up while I finish my last experiments.

To Darren, my mother has told everyone that you're a saint for putting up with me, so I think that's probably enough thanks already. Seriously though, we've kept each other happy through the good times and the hard. Thanks for being to one to come with me on the next adventure.

Abbreviations

AAV	Adeno-associated virus
AB	Ammonium bicarbonate
aq.	Aqueous
Ac	Acetyl
ABPP	Activity-based protein profiling
AfBPP	Affinity-based protein profiling
AhR	Aryl hydrocarbon receptor
AhRR	Aryl hydrocarbon receptor repressor
ATP	Adenosine triphosphate
ARNT	Aryl hydrocarbon receptor nuclear translocator
bHLH	Basic helix-loop-helix domain
BID	Bis In Die (Latin: Twice a Day)
BMD	Becker muscular dystrophy
brsm	Based on recovered starting material
BSA	Bovine serum albumin
cat.	Catalytic
CETSA	Cellular thermal shift assay
COSY	Correlation spectroscopy
CR	Cysteine-rich domain
CRISPR-Cas9	Clustered regularly interspaced short palindromic repeats - CRISPR associated protein 9
CTD	C-terminal domain
CuAAC	Copper(I)-catalysed azide-alkyne cycloaddition
DAPC	Dystrophin-associated protein complex
DARTS	Drug affinity responsive target stability
DDQ	2,3-Dichloro-5,6-dicyanobenzoquinone
dec.	Decomposed
DIPEA	*N*,*N*-Di*iso*propylethylamine
D-LH_2	D-luciferin
DLSA	5′-O-[N-dehydroluciferyl)-sulfamoyl]adenosine
DMAP	4-Dimethylaminopyridine
DMD	Duchenne muscular dystrophy
DMEM	Dulbecco's modified eagle's medium
DMF	Dimethylformamide
DMP	Dess-Martin Periodinane
DMSO	Dimethyl sulfoxide
DPAG	Department of Physiology, Anatomy and Genetics
DRE	Dioxin response element, interchangeable with XRE, xenobiotic response element
DTT	Dithiothreitol
E	Enzyme

EC_{50}	Half maximal effective concentration
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EMA	European Medicines Agency
EDTA	Ethylenediaminetetraacetic acid
E.S	Enzyme-substrate complex
eq.	Equivalents
ESI	Electrospray ionisation
Et	Ethyl
GCMS	Gas Chromatography Mass Spectrometry
FAC	Final assay concentration
FDA	Food and Drug Administration (US)
FLuc	Firefly luciferase
FT-IR	Fourier transform infrared spectroscopy
h	Hour(s)
HCI	High content imaging
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMBC	Heteronuclear multiple bond correlation spectroscopy
HPLC	High-performance liquid chromatography
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single quantum coherence spectroscopy
IC_{50}	Half maximal inhibitory concentration
iCAT	Isotope-coded affinity tags
ID	Identification
IEDDA	Inverse electron demand Diels-Alder
iPSCs	Induced pluripotent stem cells
ITC	Isothermal Calorimetry
iTRAQ	Isobaric tags for relative and absolute quantitation
J	^{1}H-^{1}H coupling constant
K_M	Half maximal velocity substrate concentration
LFQ	Label-free quantitation
LH_2-AMP	Luciferyl adenylate
LRMS	Low resolution mass spectrometry
m	Meta
M	Molar
Me	Methyl
min	Minute(s)
MoA	Mechanism of action
mp	Melting point
MRI	Magnetic resonance imaging
MS	Mass spectrometry
µw	Microwave
NMR	Nuclear magnetic resonance
nNOS	Neuronal nitric oxide synthase
NTD	N-terminal domain
p	Para
PAL	Photoaffinity label
PAS	Per-ARNT-Sim domain
PBS	Phosphate-buffered saline
PDB	Protein Data Bank
PCR	Polymerase chain reaction

PDD	Phenotypic drug discovery
PE	Petroleum Ether 30-40 ℃ boiling point
PEG	Polyethylene glycol
Ph	Phenyl
PMO	Phosphorodiamidate morpholino oligomer
PVDF	Polyvinylidene difluoride
R	Unspecified substituent
R_f	Retention factor
RIPA	Radioimmunoprecipitation assay
RLU	Relative Luminescence Units
RT-qPCR	Real time quantitative
rt	Room temperature
SAR	Structure-activity relationship
sat.	Saturated
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEA	Similarity ensemble approach
siRNA	Small intefering ribonucleic acid
SILAC	Stable Isotope Labelling by Amino acids in Cell culture
SPAAC	Strain promoted alkyne-azide cycloaddition
SPROX	Stability of proteins from rates of oxidation
TAMRA	Tetramethylrhodamine
TBAF	Tetrabutylammoniumfluoride
TDI	Target Discovery Institute, University of Oxford
TEAB	Triethylammonium bicarbonate
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TMP	Trimethoprim
TMS	Trimethylsilyl
TMT	Tandem mass tags
Tris	Tris(hydroxymethyl)aminomethane
Ts	Tosyl
UtroA	Utrophin A
UV	Ultraviolet
V_{max}	Maximum velocity
WaterLOGSY	Water-ligand observed via gradient spectroscopy

Contents

Chapter 1

Introduction

1.1 Duchenne muscular dystrophy

1.1.1 Disease condition

Duchenne muscular dystrophy (DMD) is a fatal, X-linked muscle wasting disease that affects about 1 in 3500-5000 boys.[1,2] This makes DMD the most common lethal genetic disease diagnosed in childhood, with first symptoms emerging from the ages of 3-6 years.[3] As the disease progresses, patients become wheelchair dependent by their early teens and average life expectancy is reduced to the late 20s to 30s, due to heart and respiratory failure.[4,5]

DMD is caused by loss of function mutations in the *DMD* gene, which encodes the protein dystrophin and which at 2.3 Mb, is the largest known human gene. Thousands of *DMD* gene mutations exist, with an estimated one-third caused by spontaneous mutations[6,7] including deletions, duplications and point mutations. Less commonly, *DMD* mutations occur which maintain the reading frame or are compensated for by exon skipping and result in expression of truncated but semi-functional dystrophin proteins, leading to Becker muscular dystrophy (BMD).[8] BMD symptoms are milder than DMD: the disease onset is typically later and patients can retain ambulation and independence into their 60s-70s.[9]

1.1.2 Dystrophin

Dystrophin protein (427 kDa) is a structural component of skeletal muscle fibres, linking the cytoskeleton to the extracellular matrix. As part of the dystrophin-associated protein complex (DAPC) (Figure 1.1), dystrophin contributes to strength and flexibility in muscle fibres.[11] Dystrophin's N-terminal domain binds actin to absorb shock during muscle contraction, while its cysteine rich domain anchors it to the sarcolemma-bound dystroglycan complex.[12] The dystrophin C-terminal domain interacts with cytosolic proteins such as syntrophins and dys-

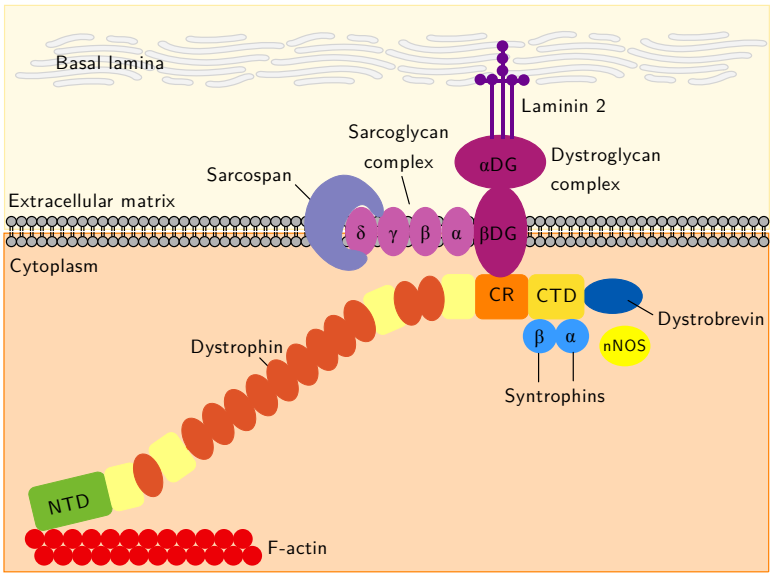


Figure 1.1: Schematic of the dystrophin-associated protein complex (DAPC), adapted from Vuorinen.[10] NTD: N-terminal domain, CR: cysteine-rich domain, CTD: C-terminal domain, nNOS: neuronal nitric oxide synthase

trobrevin, to recruit signalling proteins including nNOS, voltage-gated sodium channels and stress-activated kinases.[13] Thus, the DAPC performs roles in both membrane stabilisation and signal transduction. In DMD patients, loss of dystrophin results in loss of the DAPC, membrane collapse and cell death. Fibre degeneration leads to calcium misregulation, oxidative stress and chronic inflammation which further contribute to the pathophysiology of DMD.[14] Subsequently, cycles of muscle necrosis and regeneration continue until satellite cells are depleted, and the muscle is replaced with connective tissue and fat (fibrosis).[11]

1.1.3 The *mdx* mouse DMD animal model

The *mdx* mouse[15] is the most commonly used DMD animal model, with rat[16] and canine[17] models developing subsequently (reviewed elsewhere[18]). Although the *mdx* mouse *DMD* gene contains a premature stop codon precluding expression of dystrophin, the mice present a milder phenotype due to higher muscle regeneration than seen in DMD patients.[19] Regardless, clear DMD pathogenesis is observed and can be monitored over their lifespan, facilitating research into the mechanisms of the disease.

1.2 Therapeutic strategies for DMD

DMD is a challenging disease to treat as dystrophin is not only expressed in skeletal muscles, but also in cardiac and respiratory muscles, as well as in the brain.[20] Indeed, *DMD* mutations are correlated with an increased risk of cognitive impairment and neurological disorders, including autism spectrum disorder, attention deficit hyperactivity disorder, obsessive-

compulsive disorder and epilepsy.[21–23] Meanwhile, the magnitude of the *DMD* gene poses a delivery challenge for gene therapies, and its high mutation rate results in large genotype variation and therefore clinical heterogeneity.[24] Any DMD therapy must be systemically bioavailable and safe enough for chronic administration. An ideal therapy would be appropriate for the whole patient population, rather than subpopulations with specific genotypes. Sadly these challenges have meant that there is currently no cure for DMD patients.

1.2.1 Current standard of care

Developments in the clinical standard of care have led to improvements in quality of life and longevity.[25,26] Multidisciplinary disease management includes chronic treatment with corticosteriods (typically prednisolone and deflazacort), machine-assisted ventilation and coughing, treatment of complications such as scoliosis, and genetic counselling. Corticosteroids have been shown to improve cardiopulmonary function and prolong ambulation and survival.[27,28] However, there are considerable side effects concomitant with long-term daily use of corticosteriods, including deterioration of bone density, weight gain, height suppression, Cushingoid facies, impairment of glucose tolerance and behaviour changes.[29] Prescription of ACE inhibitors and beta blockers is common to treat cardiomyopathy, while bisphosphonates are used to treat the bone weakness brought on by muscle weakness, reduced mobility and the use of corticosteroids.

However, the current standard of care fails to tackle the causes of DMD, instead mitigating the symptoms of the disease. Extensive research and development has resulted in a variety of therapeutic strategies which are currently undergoing clinical trials.[30] These focus on one of three strategies: (1) restoring dystrophin expression (discussed in sections 1.2.2-1.2.4), (2) upregulating a compensatory related protein utrophin (section 1.2.5) or (3) targeting the consequences of dystrophin deficiency (reviewed elsewhere[30]).

1.2.2 Nonsense mutation stop codon read-through

In 2014, Translarna **1** (Figure 1.2a, formerly Ataluren or PTC124, PTC Therapeutics) became the first licensed treatment for DMD after receiving conditional approval by the EMA for nonsense mutation DMD,[31] found in about 10% of DMD patients. Translarna is a oral small molecule which facilitates ribosomal read-through translation of premature stop codons, and was found to increase dystrophin in 23/38 (61%) of trial patients, with a mean dystrophin increase of 11%.[32] The FDA rejected the PTC Therapeutics New Drug Application

for Translarna based on the failure of two clinical trials (NCT00847379 and NCT01826487) to meet their endpoints.[33,34] The advisory panel concluded that a full placebo-controlled trial was required before approval could be granted.[35]

(a) **(b)**

1 **2**

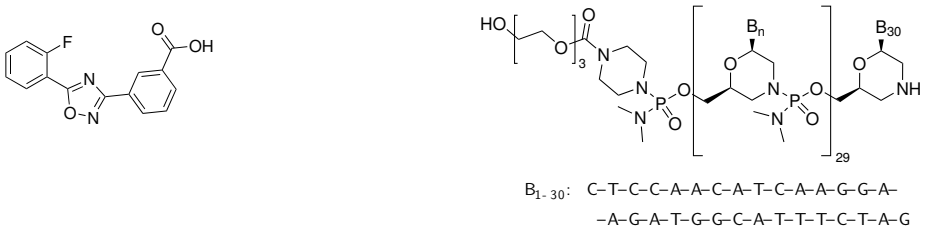


Figure 1.2: (a) Translarna (formerly Ataluren or PTC-124, PTC Therapeutics), an EMA approved oral drug for nonsense mutation DMD; **(b)** Exondys51 (formerly eteplirsen, Sarepta), an FDA approved phosphorodiamidate morpholino oligomer therapy for DMD patients with mutations in exon 51

1.2.3 Exon skipping

The *DMD* coding sequence comprises 79 exons which are spliced together during mRNA processing. Treatment with antisense oligonucleotides complementary to mutation-containing exons could allow these exons to be 'skipped' during splicing, thus bypassing the mutation and allowing patients to express their own Becker-like dystrophins. Drisapersen (PRO051/-GSK2402968, Prosensa/Biomarin and GSK) is an antisense oligonucleotide targeting exon 51 with a 2'-*O*-methyl-phosphorothioate backbone, designed to increase stability and specificity of RNA hybridisation. However, drisapersen failed to secure FDA approval after phase III trial results raised concerns about efficacy and toxicity.[36] In 2016, Exondys51 **2** (Figure 1.2b, formerly eteplirsen, Sarepta), an intravenously delivered phosphorodiamidate morpholino oligomer (PMO) which facilitates reading frame correction of dystrophin pre-mRNA, was granted accelerated FDA approval.[37,38] However, like drisapersen, this treatment is applicable only for the 13% of DMD sufferers who have genetic mutations in exon 51, at an estimated cost of $300,000.[39] Furthermore, eteplirsen's approval was viewed as controversial since it was granted against the recommendations of the FDA's internal and external scientific advisors, and based on a small treatment-dependent increase in dystrophin (0.22 - 0.32% of wildtype levels after one year of treatment[40]) rather than clinical outcomes.[41–43] Exon skipping therapies targeting exons 44, 45 and 53 are currently being evaluated in clinical trials.[44]

1.2.4 Gene therapy

An ideal cure for DMD would be replacement of the faulty dystrophin gene with a fully functional one. Encouragingly, an increase in dystrophin production to approximately 20%

of wildtype levels is sufficient to significantly ameliorate the DMD condition.[45] However, the large 11 kb coding sequence of the *DMD* gene presents a challenge for viral vector delivery. Therefore, inspired by the Becker-like dystrophins, researchers developed truncated forms of dystrophin containing the most essential domains (actin binding, cysteine rich etc.) termed 'microdystrophins' which are small enough to be packaged into adeno-associated viral vectors (AAV).[46] Successful trials in mouse[47,48] and canine DMD models[49,50] featuring various configurations of microdystrophins has sparked a race between Pfizer, Sarepta and Solid Biosciences to complete human trials.[51]

Challenges remain for this approach, however, due to the need to improve viral transduction efficiency, maintain dystrophin transgene expression over time and manage the immune responses incurred by the AAV as well as the introduced microdystrophins.[52]

Gene editing using CRISPR-Cas9 is an alternative approach, although it requires improvements to gene editing efficiency and specificity, as well as development of delivery strategies, whether by viral vectors or autologous *ex vivo* gene correction and transplantation. Exon skipping, frameshifting and exon knock-in using CRISPR-Cas9 has successfully restored full-length dystrophin in a DMD-patient-derived induced pluripotent stem cells (iPSCs) with an exon 44 deletion.[53] Furthermore, a 725 kb internal deletion corresponding to exons 45-55 of the *DMD* gene has been carried out, resulting in the production of functional dystrophin *in vitro* and *in vivo*.[54] This gene therapy strategy could restore the *DMD* gene reading frame for approximately 60% of DMD patients.

Dystrophin restoration was demonstrated in *mdx* mice one year after a single intravenous administration of an AAV-CRISPR targeting exon 23, although immune responses and persistent unintended genome modifications were also reported.[55] Administration of an AAV-CRISPR targeting exon 51 in DMD dogs has been recently found to restore dystrophin expression (n = 4) ranging from 3 to 90% of normal, depending on muscle type up to 8 weeks after treatment.[56]

1.2.5 Utrophin modulation

An alternative therapeutic approach is to compensate for the lack of dystrophin with its autosomal paralogue, utrophin.[57] The DNA sequences of the genes share 65% identity and highly similar intron-exon structures, suggesting that the two may have arisen *via* an ancient duplication event.[58]

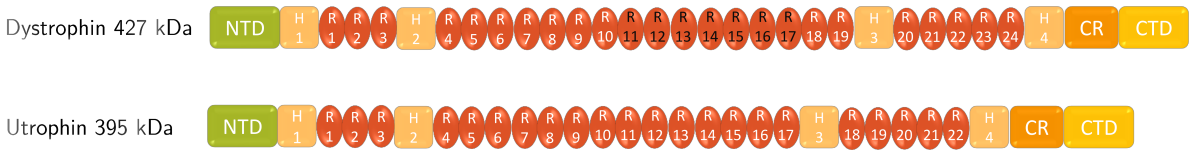


Figure 1.3: Structural domains of dystrophin and utrophin proteins are highly conserved, leading to a similar binding affinity to other DAPC components through the CR and CTD. Both proteins bind to actin via the NTD. The main difference between the two is that dystrophin has an additional actin binding domain in the central rod region (black labelled domains).

As a result, utrophin and dystrophin proteins are highly similar in sequence (80% identity[59]), size and function (Figure 1.3),[60] leading to a similar binding affinity to other DAPC components through the CR and CTD.[61] However, the two proteins have differing spatiotemporal patterns of expression. While dystrophin is expressed in skeletal, smooth and cardiac muscle and brain, utrophin is ubiquitously expressed, including in the lung, liver and kidney.[62,63] In adult skeletal muscle cells, dystrophin is found throughout the sarcolemma. Meanwhile, utrophin is expressed at the sarcolemma during foetal development and in developing or regenerating muscles, after which it is progressively replaced with dystrophin (Figure 1.4).[64,65] In mature skeletal muscle, utrophin is normally localised at neuromuscular and myotendinous junctions, however in *mdx* mice and DMD patients it is expressed along the sarcolemma, at sites normally occupied by dystrophin.[66–68]

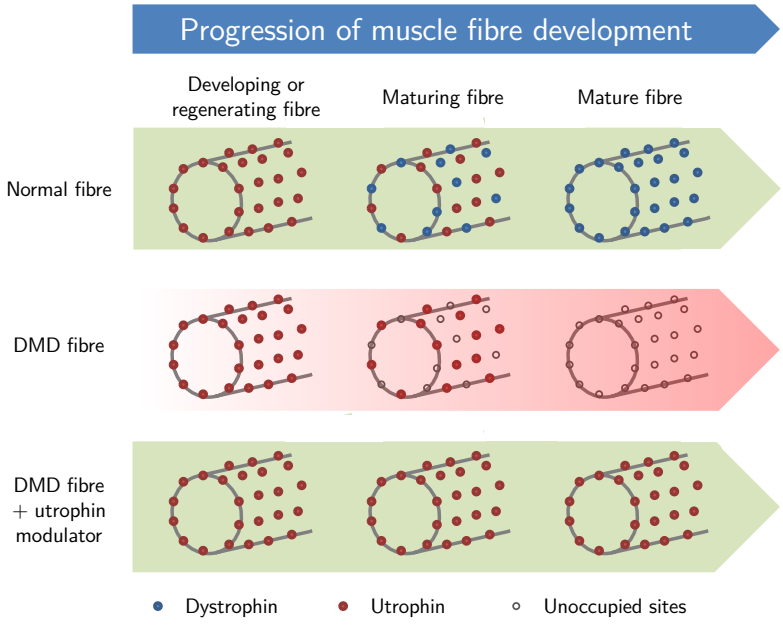


Figure 1.4: Loss of dystrophin from the sarcolemma of maturing DMD muscle fibres could be compensated for by upregulation of the related protein utrophin

The milder phenotype of *mdx* mice compared to DMD patients is correlated with increased production of utrophin, while double knockout mice, deficient in both dystrophin and utrophin, exhibit increased disease severity, comparable to the human DMD condition.[19,69] Similarly, in-

creased utrophin expression was found to correlate with improved clinical outcomes in human DMD patients.[70] Furthermore, overexpression of utrophin in *mdx* mice was shown to rescue the dystrophic phenotype without undesirable side-effects.[71–73] Indeed, increases in utrophin of just 1.5 fold along the sarcolemma were found to be beneficial. Gene therapies comprising AAV-mediated transfer of truncated utrophins have been shown to mitigate the dystrophic phenotype in DMD mice and dogs, while obviating introduction of a immune challenge from exogenous dystrophin.[74–76] Meanwhile, the delivery of Zinc Finger artificial transcription factors designed to target and activate the utrophin promoter has also been shown to improve DMD pathology in *mdx* mice.[77] Attractively, a utrophin modulation approach could offer a potential therapy to all DMD patients regardless of mutation type.[78]

1.3 Small molecule utrophin modulators

Transcriptional upregulation of utrophin using small molecules is an appealing strategy for the clinic, due to their potential for systemic delivery *via* a facile oral administration. This opportunity has sparked research interest into understanding the regulatory mechanisms of utrophin gene expression which could be targeted by a drug discovery campaign.

1.3.1 Regulation of utrophin gene expression

There are two full-length utrophin isoforms, denominated A and B, which differ in their first two exons and are under the control of different, independently regulated promoters.[79] This gives rise to differences in localisation, with utrophin A widely expressed and utrophin B restricted to vascular endothelia.[66] Several transcriptional regulatory elements and associated signalling pathways have now been characterised for the utrophin A promoter, and investigated for their therapeutic potential. These include the N-box *via* extracellular signal-regulated kinases (ERB) and GA-binding protein (GABP),[80] the E-box via MyoD/myogenin[81] and the CG box *via* Sp1 phosphorylation.[82] Additional elements include the calcineurin/nuclear factor of activated T cells (NFAT),[83] peroxisome proliferator-activated receptor γ coactivator 1α (PGC-1α)[83] and peroxisome proliferator-activated receptor β/δ (PPAR-β/δ).[84]

1.3.2 Utrophin modulators in the literature

A large number of therapeutic agents have been developed to target various signalling pathways upstream of the utrophin promoter, each demonstrating a similar capacity to increase utrophin expression 1.5-2 fold at the mRNA and/or protein level. Many of these therapies

are in preclinical trials in *mdx* mice (Table 1.1), while the most advanced is the diabetes drug metformin, which could be repurposed for the treatment of DMD. Heregulin and interleukin-6 are perhaps disadvantaged by the requirement of intraperitoneal injection for administration.

Table 1.1: Summary of the main known utrophin modulators and their mechanism and current status as therapies, adapted from Guiraud *et al.*[57]

Therapy	Type	Mechanism of action	Utrophin A fold change	Status	Ref.
Metformin	Small molecule	AMPK - PGC-1α activation	$\sim 1.5\times$ protein	Phase III clinical trial (NCT01995032)	[85, 86]
AICAR	Small molecule	AMPK - PGC-1α activation	$\sim 1.2\times$ mRNA, $\sim 1.3-2.1\times$ protein	Preclinical (*mdx*)	[87]
GW501516	Small molecule	PPAR-β/δ agonist	$\sim 1.5-1.7\times$ mRNA, $\sim 1.5\times$ protein	Preclinical (*mdx*)	[84]
Heparin	Small molecule	Utrophin mRNA stabilisation through p38 activation	$\sim 1.6\times$ mRNA, $\sim 1.5-3\times$ protein	Preclinical (*mdx*)	[88]
Celecoxib	Small molecule	Utrophin mRNA stabilisation through p38 activation	$\sim 1.5\times$ mRNA, $\sim 1.5-2.5\times$ protein	Preclinical (*mdx*)	[89]
Resveratrol	Small molecule	Proposed Sirt1-PGC-1α pathway	$\sim 1.4\times$ mRNA	Preclinical (*mdx*)	[90]
Heregulin	Peptide	$GABP_{\alpha/\beta}$ activation	$\sim 1.8\times$ mRNA, $\sim 2.7\times$ protein	Preclinical (*mdx*)	[80, 91, 92]
Interleukin-6	Recombinant protein	NRG-1/ErbB signalling	$\sim 1.5\times$ mRNA	Preclinical (*mdx*)	[93]
Nabumetone	Small molecule	Unknown	$\sim 1.8\times$ mRNA, $\sim 1.2\times$ protein	*In vitro*	[94]
Okadaic acid	Small molecule	Sp1 phosphorylation	$\sim 2\times$ mRNA	*In vitro*	[95]

1.3.3 Discovery of novel utrophin modulators by phenotypic screen

Phenotypic drug discovery (PDD) involves screening compounds for the ability to perturb a cell's phenotype in a therapeutically desirable manner. PDD does not assume a target or mechanism of action and has emerged as a popular route to first-in-class drugs.[96] A great advantage of phenotypic over target-based screening is the ability to efficiently discover new compounds

for diseases with complex pathology and regulatory mechanisms that are not fully understood. Furthermore, the efficacy of the compound can be demonstrated in a physiologically- and disease-relevant cellular context at the earliest stages of development.

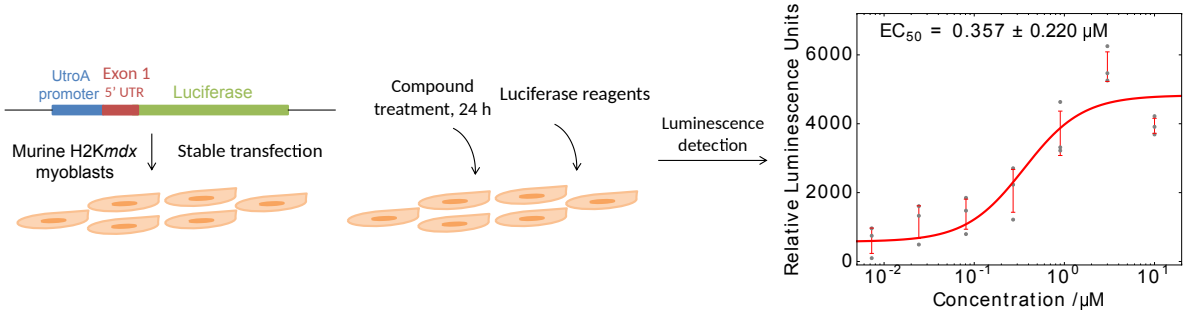


Figure 1.5: Schematic of the H2K *mdx* utrophin A - Firefly luciferase reporter gene assay and representative concentration-response curve.

In the late 1990s, researchers in Prof. Dame Kay Davies' group (Department of Physiology, Anatomy and Genetics, University of Oxford) developed a cell-based Firefly luciferase reporter gene assay as a readout for utrophin gene expression (Figure 1.5).[97,98] 8.4 kb of the mouse utrophin A promoter, 5' to the transcription start site, was transfected into a stable H2K *mdx* myoblast cell line.[99] High-throughput screening of compounds proceeded, measuring the induced bioluminescent readout which depends on the quantity of luciferase, and hence activation of the utrophin promoter. Initial pilot screens identified a stilbene hit series **3** and an isosteric benzoxazole hit **4** (Figure 1.6). Hit-to-lead evaluation gave rise to a family of novel 2-arylbenzoxazoles and ultimately **5** was optimised as the lead compound and named SMT C1100 (later ezutromid).[100] Ezutromid was demonstrated to increase utrophin mRNA 1.4 fold[100] and double utrophin protein expression in a human DMD muscle cell line,[101] warranting assessment of its therapeutic potential in preclinical trials in *mdx* mice.

Meanwhile, further optimisation of ezutromid has led to a second generation of phosphinate utrophin modulators such as **6** with improved physicochemical properties and a more robust metabolism profile.[102] In addition, further screening has engendered new hit series' which have been developed into third generation utrophin modulators.[10]

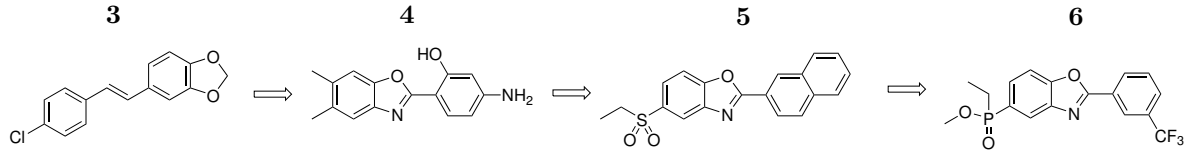


Figure 1.6: Hit-to-lead development of novel utrophin modulators led to selection of SMT C1100/ezutromid **5** as the clinical candidate. Second generation modulator phosphinate **6** was developed subsequently.

1.4 Ezutromid – a first-in-class utrophin modulator

1.4.1 Preclinical trials in *mdx* mice

Daily oral treatment of *mdx* mice with ezutromid led to increased utrophin expression in skeletal, cardiac and respiratory muscles after 28 days, without toxicity.[101] Ezutromid was also found to diminish muscle membrane damage, evidenced by a reduction in force drop after eccentric contractions.[101] Meanwhile, serum creatine kinase, muscle fibrosis and necrosis were also reduced, indicating that ezutromid mitigates the damaging secondary pathology associated with DMD.[101] Furthermore, ezutromid demonstrated synergism with prednisolone, the corticosteroid standardly prescribed for DMD patients.[101]

1.4.2 Clinical trials

Ezutromid's maximum thermodynamic solubility was determined to be a suboptimal 3 µM by Cyprotex (Alderley Park, UK). Optimisation of ezutromid's formulation was required to ameliorate its poor pharmacokinetic profile, much of which derives from its planar, hydrophobic structure. This resulted in the development of two formulations: a microfluidized aqueous oral suspension formulation (F3) and a powder for suspension formulation (F6), which were used in parallel for the later clinical trials.[57]

Phase I trials

In a Phase I trial in healthy volunteers, ezutromid was found to be safe and well-tolerated, and available in plasma concentrations considered sufficient for a therapeutially beneficial 50% increase in utrophin expression.[98] In a subsequent dose-finding Phase Ib trial in DMD patients, ezutromid treatment was again found to be free of significant adverse events, and was found to decrease levels of muscle damage biomarkers such as creatine phosphokinase, alanine aminotransferase and aspartate aminotransferase.[103] However, markedly lower plasma concentrations of ezutromid were observed in the DMD boys compared with the healthy volunteers, along with high patient variability in exposure independent of dose. Additionally, exposure was significantly reduced (56 - 65%) on the final (11th) day of treatment compared to day 1.[103]

Metabolism of ezutromid occurred within hours of dosing, mostly *via* hydroxylation and subsequent hydroxyl conjugations. Ezutromid's primary metabolites were two dihydrodiols formed from dihydroxylation of its naphthyl ring. In total, 17 metabolites were detected in

plasma, and 25 in urine samples.[103]

Phase II trial: PhaseOut DMD

PhaseOut DMD, a multi-centre phase II, open-label proof-of-concept study began in 2016. Ezutromid (parallel F3 and F6 formulations) was administered orally BID in 40 ambulatory paediatric male subjects (aged 5-10 years) with DMD (Table 1.2). The primary endpoint was evaluation of % fat fraction within muscle tissue by quantitative MRI, a validated measure of DMD severity and progression.[104] The secondary endpoint was evidence of an increase in utrophin and decrease in developmental heavy chain myosin, a biomarker for muscle damage[105] in muscle biopsies. In addition, functional endpoints such as the 6-min walk distance and North Star Ambulatory Assessment were monitored throughout the study.

Table 1.2: Study design of the PhaseOut DMD trial, adapted from Guiraud *et al.*[57]

Screening Day −28	Cohort 1: 2.5 g BID F3 formulation, n = 30 Cohort 2: 1.0 g BID F6 formulation, n = 10					Optional extension	
	Day 1	Week 12	Week 24	Week 36	Week 48	Week 72	Week 96+
Biopsy	✓		✓		✓		
MRI	✓	✓	✓	✓	✓	✓	✓
Functional tests	✓	✓	✓	✓	✓	✓	✓

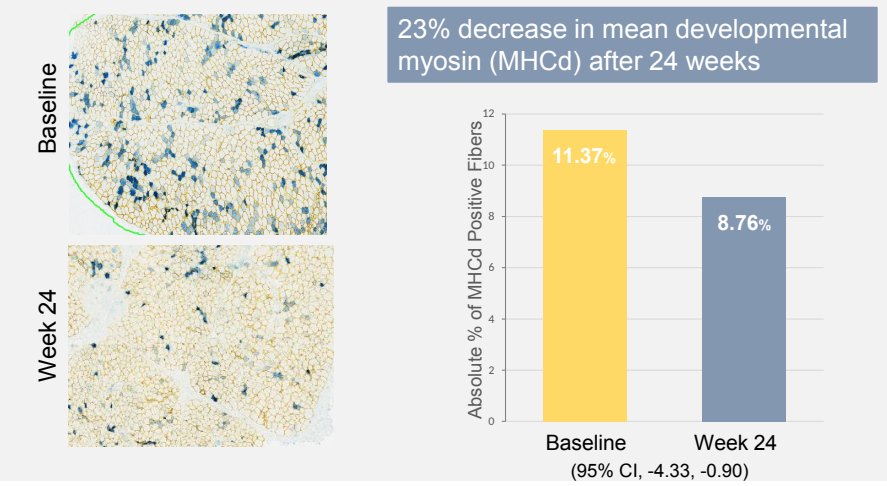


Figure 1.7: Patient muscle biopsies show a significant decrease in the disease biomarker developmental myosin after 24 weeks of ezutromid treatment.

In January 2018, the 24-week interim results were published, finding that all patients achieved plasma levels of ezutromid sufficient to modulate utrophin. Further, ezutromid induced a 7% increase in mean utrophin protein intensity levels in biopsies, alongside a 23% decrease in mean developmental heavy chain myosin (Figure 1.7). Ezutromid also significantly

reduced muscle inflammation as measured by the quantitative MRI.[106] Together, these data provided the first evidence of ezutromid target engagement and proof of mechanism.

However, in July 2018, after the full 48 weeks of trial, the primary and secondary endpoints were not met and development of ezutromid was discontinued. Since the target and molecular mechanism of action of ezutromid were unknown, it was difficult to rationalise this lack of sustained clinical efficacy. Arguably, this is the most significant challenge facing the phenotypic approach to drug discovery.[107]

1.5 Target identification and mechanism of action (MoA) discovery

Target identification is of increasing significance due to the resurging popularity of phenotypic drug discovery, where an unknown target can lead to a difficult blind medicinal chemistry and drug development campaign. Target identification can allow for a rational compound design strategy and provide biomarkers for monitoring efficacy in clinical trials. It can also provide insights into potentially deleterious on- and off-target side effects, which is crucial given that drugs have been found to bind on average between 6 and 12 different proteins.[108] Target identification has the potential to significantly reduce attrition in the drug discovery process. However, its resource- and time-intensive nature represents a major bottleneck.

Modern methods to elucidate direct protein targets and mechanism of action of small molecules comprise a diverse range of fields, and have been extensively reviewed.[109–115] Target ID strategies can be divided into direct methods, which identify the protein interacting partner(s) of a compound, and indirect phenotypic profiling methods, which can infer mechanism by assessing changes in phenotype in response to compound treatment. Direct methods typically lead to a shortlist of targets of the small molecule, without any functional information about the interaction. Determining which are involved in the compound's therapeutic mechanism poses a significant challenge, which is assisted by phenotypic profiling methods. Hence a combination of methods is usually required to deconvolute a compound's mechanism of action. Recent developments in target ID methods will be briefly discussed.

1.5.1 *In silico* target prediction

Computational target inference is based on the principle of chemical similarity (i.e. similar compounds bind similar targets) or on the biophysical calculation of the binding energy between compound and query target. Similarity comparisons can be made between compound

two-dimensional (2D) fingerprint, 3D shape, pharmacophore or phenotype. The Similarity Ensemble Approach (SEA) applies 2D similarity on a systems level, by relating different protein targets based on the whole set of ligands that are known to bind them.[116] In one study, 23 out of 30 SEA-predicted interactions between FDA approved drugs and off-target proteins were validated experimentally.[116] Drug target interactions can also by predicted by network-based[117] and machine learning methods,[118] which mine the vast quantities of genomic, transcriptomic, proteomic and metabolomic along with protein structural data that are now available in public repositories such as PubChem, ChEMBL and the PDB. Across the different strategies for *in silico* target prediction, a variety of accessible tools have been developed including ChemProt, SPiDER, TargetHunter and SwissTargetPrediction.[114] However, while prediction of the protein targets of a compound using cheminformatics can be an excellent place to start, its utility is limited by the need for well-annotated biological targets with known inhibitors.

1.5.2 Chemical genetics

Chemical genetics is the study of compound-induced perturbation of biological systems, in analogy to the perturbations caused by gene mutations. By relating the phenotype of a gene mutation with that of compound treatment, the compound's mechanism of action can be inferred. This section will briefly describe phenotypic profiling and genome-wide screening approaches for MoA discovery.

Gene expression profiling

Improvements in throughput and cost-effectiveness of sequencing and mass-spectrometry methodologies have increased access to global analyses of cell transcriptomes, proteomes and metabolomes. Profiling of differences in gene expression in response to acute or chronic compound treatment can reveal pathways activated or suppressed by compound treatment, which may be related to on- or off-target activity. Drug-induced changes in RNA expression can be readily assessed by microarray (a probe hybridisation technique) or RNA-seq, which can sequence tens of thousands of RNAs. Meanwhile, quantitative shotgun proteomics enables detection of changes in protein abundance as well as post-translational modification status.

Integration of the data that results from these experiments into MoA hypotheses can be challenging, and is perhaps best achieved by comparison to signatures of compounds with known targets. A microarray-based gene expression profiling platform focussed on 1000 'land-

mark' genes (L1000 GEx, Genometry Inc., Cambridge MA) has been used to amass 1.3 million publicly available profiles from 19,811 small molecules, 18,493 shRNAs, 3,462 cDNAs, and 314 biologics tested in 3 to 77 different cell lines into a 'Connectivity Map' (CMap).[119] CMap has been successfully used in academia and industry to determine compound MoA and repurpose drugs.[114,120]

Genetic mutagenesis

Genome- and transcriptome-wide sequencing can also be used to elucidate MoA by identifying resistance-conferring mutations. The identification of resistance mutations can also provide information about compound transport and efflux, helping to inform compound design to prolong the lifespan of the compound once on the market. Resistance mutations can be spontaneously acquired by evolutionary selection from chronic compound treatment, or identified through screening of mutant libraries.

The first chemical genetics MoA studies were carried out in yeast, due to its genetic tractability. Libraries of diploid yeast with one gene copy deleted (i.e., half the gene dosage) were screened for increased drug-induced sensitivity in a process termed haploinsufficiency profiling (HIP).[121] Along with homozygous profiling (HOP) which involves complete gene deletion, HIP has been successfully used to study compound MoAs.[122,123] More recently, combining compound treatment with systematic gene disruption using RNA interference with short hairpin RNAs or genome editing with CRISPR-Cas9 allows for identification of genes that mediate the compound's activity.[124–126]

1.5.3 Target identification

Many techniques to determine a small molecule's direct interacting partners have been developed, all of which in some way exploit the compound's affinity and specificity for its protein target(s). A straightforward approach is to test large panels of prospective targets in biochemical assays for binding or inhibition. This can be achieved conveniently with commercial screens for specific target classes such as kinases or GPCRs. In a recent example,[127] GSK had success by testing a compound with unknown mechanism of action against 800 different proteins ranging from kinases, epigenetic readers and nuclear receptors in target-based assays. However this approach requires large resources, has limited biological relevance and limited target scope. The most commonly used methods for identifying small molecule binding partners are briefly described below and are further reviewed elsewhere.[113,114]

Chemical probe-based target enrichment by affinity purification

In this strategy, chemical probes are designed as 'baits' to capture and enrich their direct binding partners from complex biological systems, including disease model cells, lysates or tissue homogenates. Enriched proteins can then be identified by western blotting or digested and analysed by LC-MS/MS. However, it requires an established SAR of the parent compound to develop active probes bearing tags for pulldown purification or visualisation. These tags can be fluorophores, expression tags or substrates for affinity resins e.g., biotin, -NH_2 etc. but importantly introduction of these linkers and tags must not adversely affect probe activity, localisation, cell permeability or solubility. This can present a significant challenge, especially in developing SAR in the absence of a known target.

A popular method for target ID involves pre-conjugation of the chemical probe to solid supports (typically streptavidin beads or resins such as sepharose or agarose) followed by incubation with cell or tissue homogenates (Figure 1.8a).

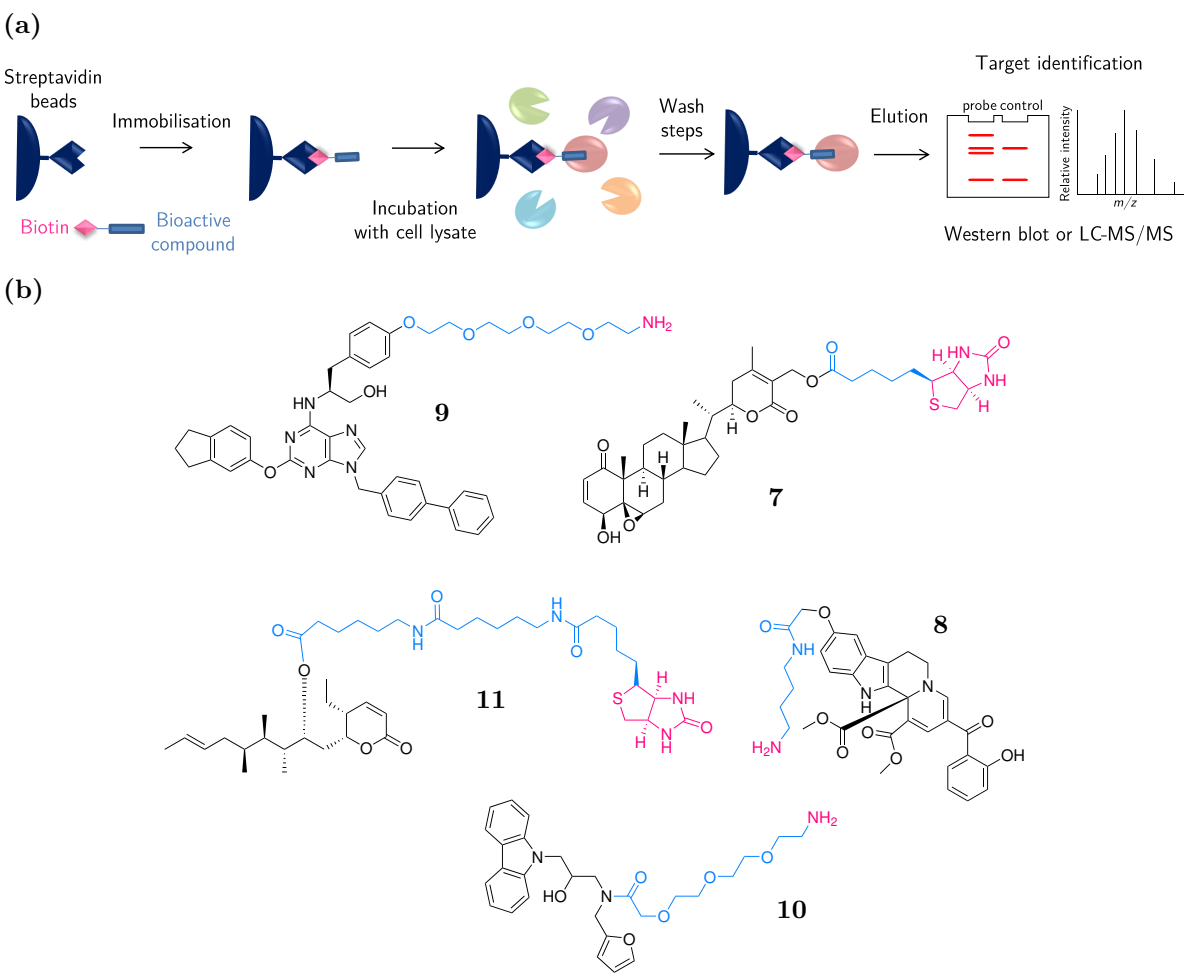


Figure 1.8: **(a)** Schematic overview of chemical probe-based target enrichment with an immobilised affinity matrix. **(b)** Examples of probes that have been used in successful target identification, black: bioactive 'bait' as ligand for the target protein, blue: linker, pink: affinity matrix substrate.

Development of these kind of probes for bioactive compounds such as QS11 **7**,[128] withaferin A[129] **8**, pironetin[130] **9**, centrocoutin[131] **10** and KL001[132] **11** (Figure 1.8b) has led to the successful identification of their molecular targets. However a disadvantage of using immobilised compound matrices is the requirement for cell lysis which may disrupt protein folding, complexation and post-translational modification state.

Activity-based protein profiling (ABPP) exploits the functional reactivity of enzymes (such as proteases and kinases) to design probes containing warheads capable of covalent attachment to the active sites of their protein target. ABPP combined with mass spectrometry has been used to great effect to describe the functionality of uncharacterised proteins[133] and identify compound targets through competition with ABPP reporters.[134] However, ABPP is limited by its inability to target the majority of druggable proteins or allosteric binding sites. As a result, affinity-based protein profiling (AfBPP),[135–137] which relies on the chemical probe's non-covalent affinity for its target, has a much larger target scope.

In situ target enrichment, *i.e.* in live cells in culture or *in vivo*, requires the use of a cell-permeable probe, but addition of enrichment tags including biotin often reduces cell-permeability. In these cases, *in situ* enrichment can be accomplished by a two-step process: first, binding of a cell-permeable probe to its target protein; second, the ligation of a purification tag *via* a bioorthogonal reaction.[138] This requires the incorporation of a bioorthogonal handle into the chemical probe, which importantly must not be buried in the binding site, rather accessible to click reagents in the solvent. *In situ* experiments have been shown to enrich targets which could not be found in *in vitro* experiments.[136]

Bioorthogonal 'click' reactions

A variety of bioorthogonal or 'click' reactions have been developed,[139,140] with the most commonly used in target ID being copper(I)-catalysed alkyne-azide cycloaddition (CuAAC), strain promoted alkyne-azide cycloaddition (SPAAC) and inverse electron demand Diels-Alder (IEDDA) cycloaddition (Figure 1.9). CuAAC and SPAAC use azides **12** as sources of 1,3-dipoles for cycloaddition since they do not naturally exist in cells and are metabolically stable, making them ideal for bioorthogonal transformations. The use of copper(I) to catalyse and improve regioselectivity of the Huisgen 1,3-dipolar cycloaddition between azides and alkynes **15** to give 1,4-substituted 1,2,3-triazoles **16** was reported in 2002 by Sharpless and Meldal[141,142] and since then has become the most widely used bioorthogonal reaction. One drawback of CuAAC which can limit its use in living cells is Cu(I)-mediated generation of reactive oxygen

species from molecular oxygen. Subsequent development of water-soluble ligands to stabilise Cu(I) can restrict this toxicity.[143,144]

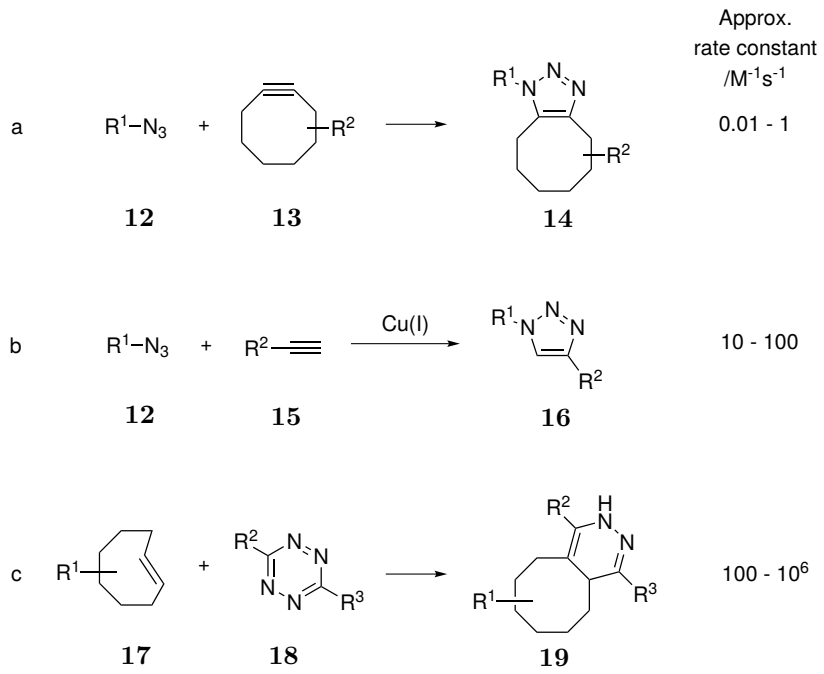


Figure 1.9: Most commonly used bioorthogonal reactions. **(a)** Strain promoted alkyne-azide cycloaddition (SPAAC); **(b)** copper(I)-catalysed alkyne-azide cycloaddition (CuAAC); **(c)** inverse electron demand Diels-Alder (IEDDA) cycloaddition between trans-cyclooctene (TCO) and tetrazine.

Meanwhile, SPAAC was leveraged for bioorthogonal applications by the Bertozzi group in 2004[145] and its significant advantage over CuAAC is its catalyst-free nature, allowing it to be used in live cells. Typically cyclooctynes **13** are used for SPAAC as they are the smallest stable cycloalkyne. Their reactivity is driven by enthalpic release of ring-strain going from a deformed sp-hybridised ring system to a favourable sp^2-hybridised one **14**. However, SPAAC suffers from slow kinetics and side reactions of the cycloalkyne with thiols.

More recently, inverse electron demand [4+2] cycloadditions of 1,2,4,5-tetrazines **18** with various dienophiles have been used for bioorthogonal reactions . IEDDA cycloadditions, particularly using strained dienophiles such as trans-cyclooctenes **17** (TCO), demonstrate extremely fast kinetics (up to 10,000 faster than CuAAC). One study comparing the efficiency of IEDDA (with TCO), SPAAC and CuAAC in target enrichment of PARP1 inhibitor olaparib determined relative efficiencies of 100%, 45% and 9% respectively.[146] However, both tetrazines and TCO are bulky groups and may be difficult to incorporate into a probe without abrogating activity. For this reason, cyclopropenes have been used as IEDDA reactants, for example by the Yao group in the design of a probe for bromodomain inhibitor (+)-JQ1.[147]

Photoaffinity probes

AfBPP can be enhanced by inclusion of a photoaffinity label on the pulldown probe, which facilitates discovery of low affinity binding partners.[148] Photoaffinity groups are stable until activated by irradiation with specific wavelengths of light to reveal highly reactive intermediates, such as radicals and carbenes.[149] Thus, a ligand modified to contain a photoaffinity label can bind to its protein targets *via* its usual reversible intermolecular forces, and after photoactivation, by covalent attachment at the binding site. The concept of photoaffinity labelling was introduced by Westheimer in 1962, with the use of an α-keto diazo group to label chymotrypsin.[150] However, the use of these groups as photoaffinity labels is limited by Wolff-type rearrangements. Since then, several photoaffinity labels (PALs) have been developed and the most commonly used are shown in Figure 1.10.

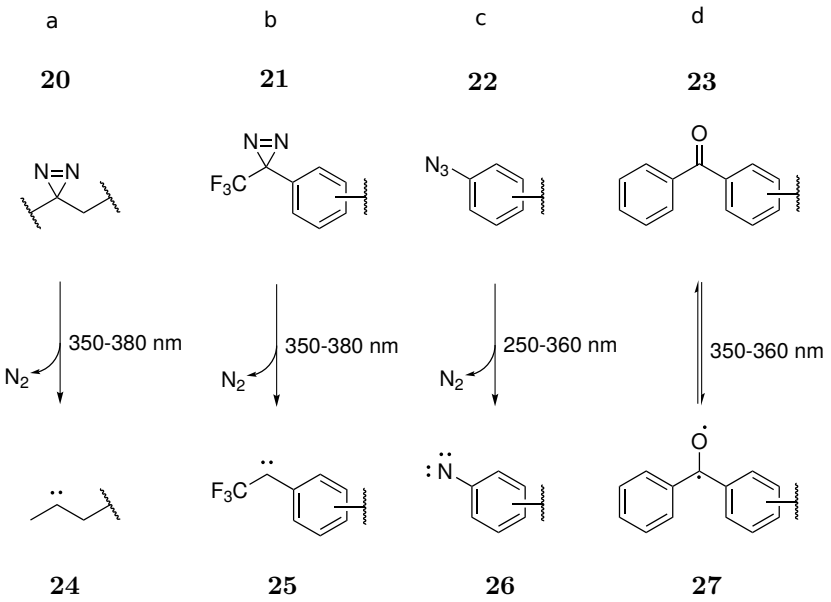


Figure 1.10: Most commonly used photoaffinity labels, their activation and their reactive intermediates. **(a)** Alkyl diazirine **20**; **(b)** trifluoromethylphenyl diazirine **21**; **(c)** aryl azide **22**; **(d)** benzophenone **23**.

In general, the photoaffinity label approach suffers from low crosslinking efficiencies owing to the short half-life of the reactive intermediate and non-specific binding to other proteins, usually driven by the hydrophobic effect.[148] Addition of the PAL to the chemical probe must be at a protein-binding rather than solvent-exposed part of the molecule which can be challenging to achieve without decreasing activity of the probe. However, for dual-tagged probes containing both a PAL and a click handle, after the covalent binding of the ligand to the target the subsequent click ligation can occur under denaturing conditions, improving accessibility of the click handle.

Aryl azides **22** were originally popular as photoaffinity labels, due to their small size and

facile synthesis. However, their activation wavelengths are higher energy than other PALs, ranging from 250-360 nm, meaning a greater risk of UV damage to biological molecules. Furthermore, the nitrene **26** generated with photoirradiation can rearrange to form benzazirines and dehydroazepines/ketenimines leading to undesired side products rather than insertion into the biological target.[148] These rearrangements can be suppressed by the presence of electron-withdrawing groups in the aryl ring, and this solution was used to develop probe **29** containing a tetrafluorophenyl azide and alkyne, and based on the γ-secretase modulator GSM-1 **28**[151] (Figure 1.11).

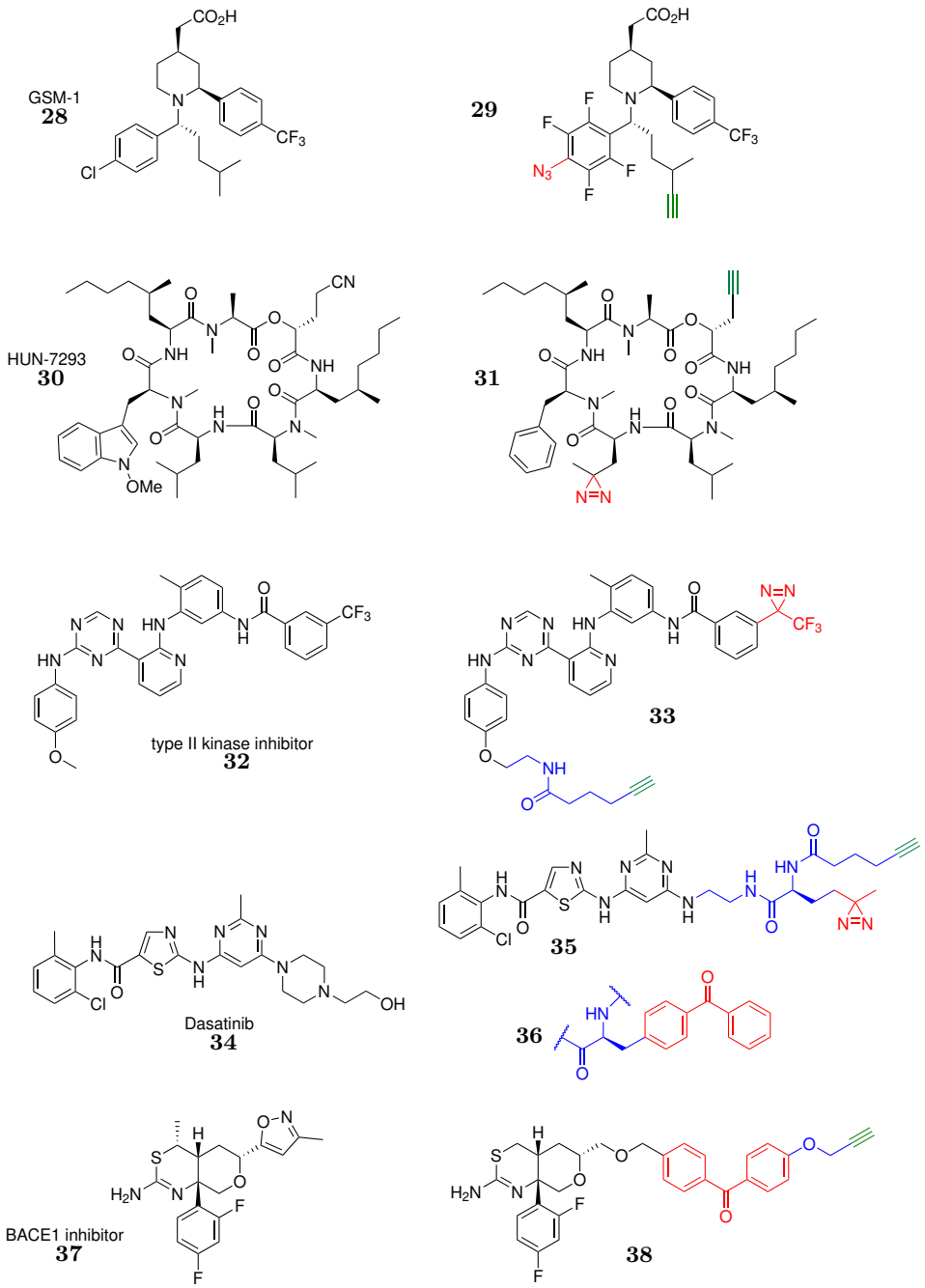


Figure 1.11: Literature examples of photaffinity-labelled probes and the compounds they derive from. Red: photoaffinity label, blue: linker, green: bioorthogonal handle.

Diazirines were first proposed as photoaffinity labels in 1973 by Smith and Knowles with their report on the preparation of 3-aryl-3H-diazirines,[152] and have since grown to be the most popular PAL. However, both aryl and aliphatic diazirines **20** are liable to photoisomerisation to linear diazo groups as well as generation of the desired carbenes **24**.[153] The relative yield of the carbene to diazo isomer upon diazirine photolysis varies depending on the substituents on the diazirine. Brunner *et al.* reported the synthesis of trifluoromethylphenyl diazirines (TPDs) **21** which reduced diazo isomer generation to 35% vs. 65% carbene **25**.[154] This is a higher carbene yield than typically seen for aliphatic diazirines.[155] However, the small size of aliphatic diazirines makes them easy to incorporate into probes and even amino acids, for example, the integration of photo-Leucine into fungal cyclodepsipeptide HUN-7293 **30** to produce photoaffinity probe **31**[156] (Figure 1.11). While TPDs are larger to incorporate into a probe, they can often substitute existing aryl rings, as seen in the design of probe **33** based on type II kinase inhibitor **32**[157] (Figure 1.11).

Benzophenones are another popular PAL **23**, reversibly forming reactive triplet diradicals **27** when photoirradiated. Like diazirines, benzophenones have considerably milder activating wavelengths ($\sim$ 350 nm) than aryl azides, implying less radiation damage when used *in cellulo*.[158] Benzophenones have been used in probes of dasatinib **34**[136] and the BACE-1 inhibitor **37**[159] (Figure 1.11). However, benzophenones tend to require longer irradiation and this together with their hydrophobic bulk often contributes to increased non-specific labelling.[148]

Ligand-directed protein labelling

Ligand-directed (affinity-based) protein labelling is an alternative strategy to PAL. Chemical probes consisting of a protein ligand 'bait' and a reactive group bind to their target before undergoing an S_N2-type reactions with a neighbouring nucleophilic amino acid. Various reactive groups have been applied including tosyl groups,[160] acyl imidazoles,[161] dibromophenyl benzoates[162] and epoxides.[163] These methods have been successfully used to label receptors on the surfaces of live cells as well as intracellular proteins. In another method, DMAP was appended to the small molecule bait and was incubated with cells along with tags (e.g. biotin) primed with a thioester.[164] The DMAP catalyses labelling of the protein target with the tag on a nucleophilic residue in the vicinity (Figure 1.12). The protein can then be enriched and identified by western blot or LC-MS/MS. This method was successfully used by Wang *et al.* to label bradykinin B_2 receptors on the surfaces of live cells,[164] but is limited by the requirement

for a local nucleophilic amino acid sidechain for labelling to occur.

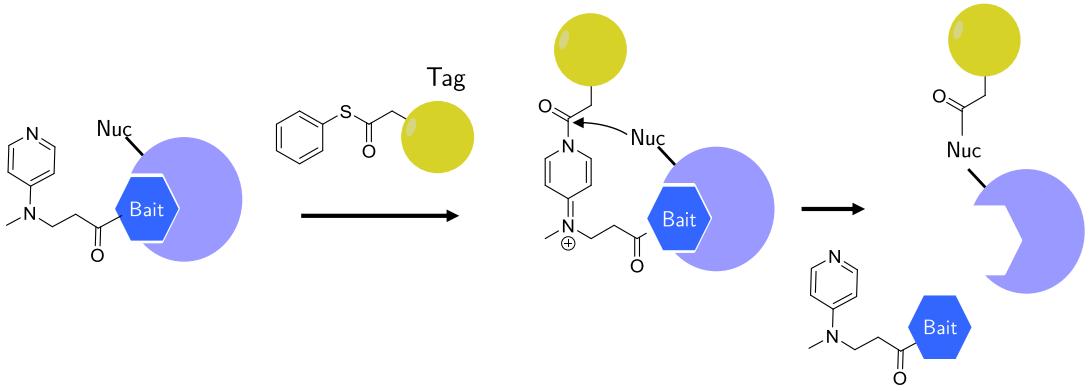


Figure 1.12: Schematic overview of ligand-directed protein labelling using DMAP as a catalyst.

Target display techniques

These strategies involve the artificial enrichment of libraries of proteins followed by screening for those that bind to a particular small molecule. The advantage of this approach is that an interaction can be found even where the protein's physiological abundance is low. Bacteriophage display using full length cDNA libraries has been successfully used for target identification, even of weak affinity targets.[165] However, a limitation of this technique is a dependency on bacterial expression systems, reducing the post-translational modification of proteins.

A related approach, termed the chemical yeast-three hybrid, exploits the binding interaction between target protein and small molecule to bring two parts of transcription machinery in close enough proximity to allow for transcription of an essential gene to occur (see section 1.6.1).[166] In this way, a positive selection pressure is applied and only yeast expressing protein-fusions that bind to the small molecule survive. Limitations include the artificial nature of the gene fusions which may affect conformation, high numbers of false positives and the need to immobilise the small molecule.

Targeted protein degradation

Small-molecule-induced proteasomal degradation of target proteins is a promising new modality for drug discovery.[167] This technique, using proteolysis-targeting chimeras (PROTACs), has also been used to validate the ENL YEATS domain as a novel target for acute leukaemia.[168] It is possible that PROTACs technology could be used in the future for target identification by looking for proteins degraded by PROTACs treatment, although there are challenges in developing an optimised degrader without a known target.

Modification-free target ID methods

The methods described above are complemented by capture-free target ID methods, such as Cellular Thermal Shift Assays (CETSA)[169] (discussed further in section 6.2.1), Stability of Proteins from Rates of Oxidation (SPROX) assays[170] and limitation of proteolysis in Drug Affinity Responsive Target Stability (DARTS) assays[171] (reviewed in detail elsewhere[172]). These methods all exploit the stabilisation imparted by small molecule binding to differentiate between bound and non-bound proteins. Importantly, they can be combined with multiplexed quantitative mass spectrometry to achieve an unbiased, proteome-wide profiling of drug-target interactions in cells or lysates. A significant advantage of these approaches is the ability to use the original compound without modification. However, a challenge is an abundance bias without the enrichment of target proteins. Typically, 5000-8000 proteins can be identified by routine whole cell shotgun proteomics, which currently limits the targets that can be assessed by these methods. However, these experiments can be useful for target engagement and validation experiments by western blot.

1.6 Preliminary target identification studies for ezutromid

As ezutromid progressed through clinical trials, interest in its mechanism of action was growing. This was especially in light of the limitations of ezutromid's pharmacokinetic profile, as a deconvoluted target could help development of improved clinical candidates. Furthermore, it would help differentiate later hit series which would ideally operate through a disparate mechanism. Consequently, three target ID studies were commissioned by Summit Therapeutics, and the results of these are described in the following sections.

1.6.1 Yeast three-hybrid study

In 2014, Hybrigenics Services SAS carried out a yeast three-hybrid experiment to help identify ezutromid's protein targets. 'Prey' proteins were derived from a human adult/foetal skeletal muscle cDNA library and fused to a *HIS3* gene transcriptional activation domain and expressed in yeast (Figure 1.13a). Meanwhile, ezutromid analogues were tethered *via* a PEG chain to trimethoprim (TMP), a potent ligand (K_D = 6 pM) of dihydrofolate reductase (DHFR), which was in turn expressed as a fusion protein with a DNA-binding domain specific for the *HIS3* gene promoter. The yeast were grown in medium lacking histidine to provide a selection pressure for those whose prey proteins are targets of the ezutromid-like bait and therefore capable of forming a trimeric complex and activate *HIS3* gene transcription. Two

ezutromid probes were used in this experiment, **39** and **40** (Figure 1.13b).

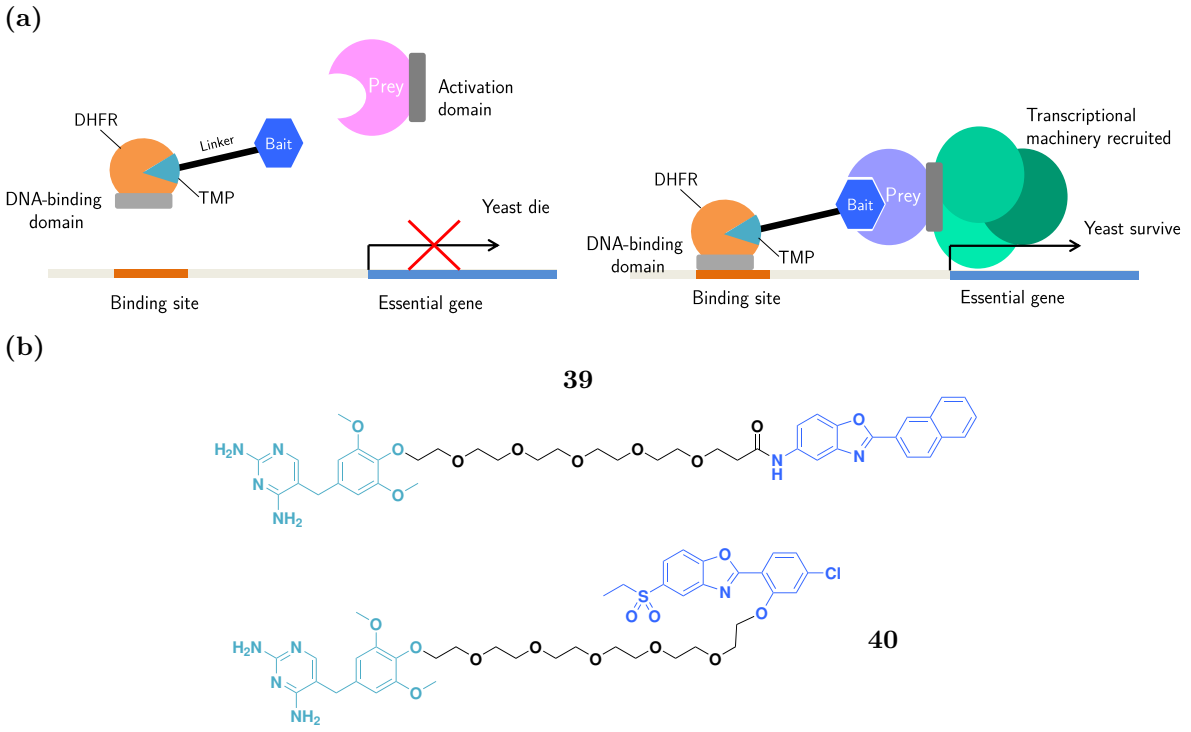


Figure 1.13: **(a)** Schematic overview of the yeast three-hybrid target ID experiment, left: the prey is not a target of the bait, transcription of the essential gene is switched off, right: the prey is a target of the bait, bringing the transcriptional activation domain and recruited transcriptional machinery in proximity of the gene promoter, switching on gene transcription and leading to yeast survival; **(b)** The ezutromid-based probes (linked by a PEG chain to TMP (green)) used in this study. DHFR: dihydrofolate reductase, TMP: trimethoprim

Initially, the yeast were treated with the probes and grown on non-selective medium to confirm their non-toxicity. Then the yeast were grown on the selective medium, and probe **39** was initially found to bind to 18 of the prey proteins. Four of the highest confidence genes (FHL1, CSK, STRADB, MYH1) were re-amplified and freshly transformed into yeast for a repeat experiment as the first validation step. However, in each case, the yeast survival was not recapitulated and resultantly none of these interactions were validated. Probe **40**, which was intended as the negative control, was found to bind BLVRB (biliverdin reductase B) and this interaction was validated. Unfortunately this experiment failed to identify any targets to follow up.

1.6.2 Profiling of ezutromid in a DiscoverX BioMAP® study

Ezutromid was next profiled in a DiscoverX BioMAP® Diversity PLUS study which examines changes in activity of 148 biomarkers across 12 different human primary cell-based co-culture systems upon compound treatment.[173] Ezutromid was found to dose-dependently decrease activity of sIgG, and sIL-17F, indicating downregulation of the immune response (Figure

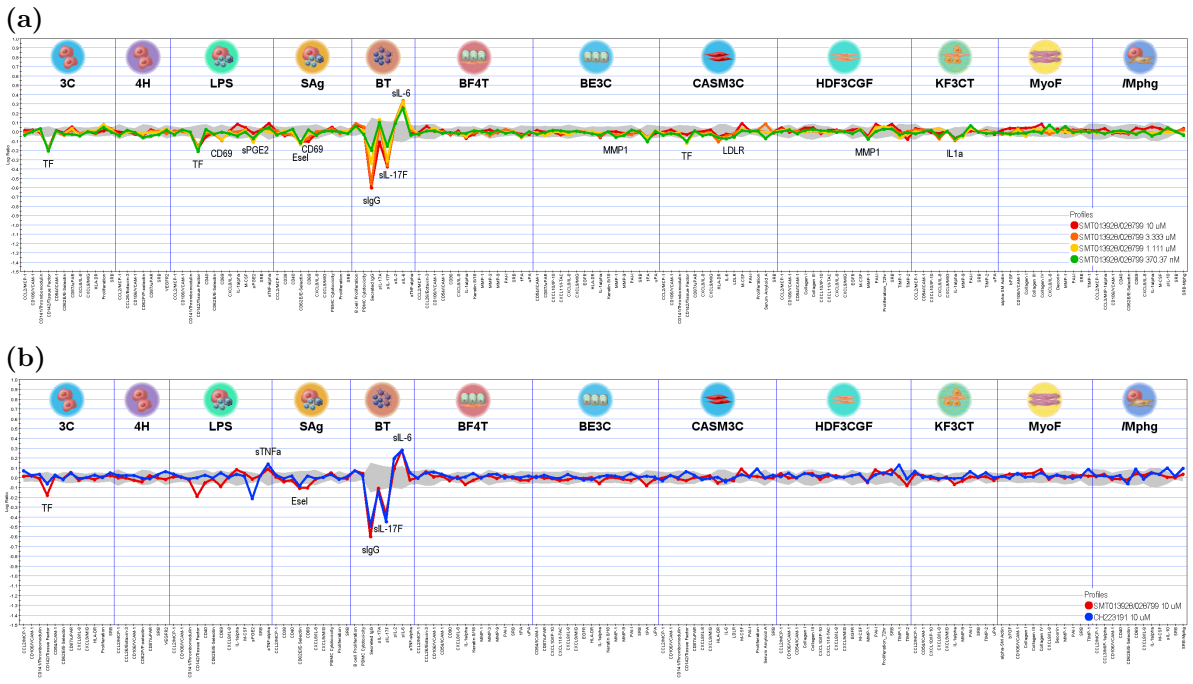


Figure 1.14: **(a)** Profile of ezutromid in the DiscoverX BioMAP Diversity PLUS panel, dosed at 0.37 µM(green), 1.11 µM(yellow), 3.33 µM(orange), 10 µM(red); **(b)** Profile of ezutromid (10 µM, red) overlaid on the profile of AhR antagonist CH223191 (10 µM, blue), Pearson correlation of 0.854. Grey profiles represent baseline variation, biomarkers with activity deviating from the baseline are annotated

1.14a). Meanwhile, ezutromid caused modest decreases in sPGE2, E-selectin and IL1α, which are hallmarks of an anti-inflammatory response. The best match to ezutromid's profile (red) from a database of >4500 reference compounds was that of aryl hydrocarbon receptor (AhR) antagonist CH223191 (blue, Pearson correlation, r = 0.854) (Figure 1.14b). However, there is no known connection between AhR modulation and utrophin expression.

In summary, the 'hits' generated from the preliminary target ID experiments carried out for ezutromid were insufficient to identify a high confidence target or mechanism to follow up in validation studies. However, they could be returned to in the context of results of more target ID experiments.

1.7 Aims of this work

The aim of this work is to uncover ezutromid's molecular mechanism of action to help understand the trial results and thus develop better utrophin modulators for the future. Each target ID technique has its own associated limitations and hence a combination of strategies is prudent to optimise likelihood of success. Moreover, it is likely that ezutromid binds to a large number of protein targets, since ligand promiscuity is associated with linearity, planarity and hydrophobicity.[174,175] Deconvolution of proteins to identify the therapeutically relevant target from non-specific hits would be best facilitated by comparing results from orthogonal experiments.

The starting point for this DPhil research is identification of ezutromid's direct binding partners through affinity purification, to complement phenotypic profiling studies that collaborators were working on: RNA-seq at our partners at Evotec and ChIP-seq and ATAC-seq by the Davies group (DPAG). Chapter 2 discusses the design and synthesis of biotinylated and dual-tagged probes and controls for AfBPP experiments and describes the difficulties encountered in determining probe bioactivity, stemming from artefacts in the primary firefly luciferase reporter gene assay. Chapter 3 explores the causes of the firefly luciferase inhibition which resulted in the assay artefacts. Chapter 4 discusses the secondary assays employed to eventually demonstrate probe bioactivity. Chapter 5 details the affinity capture experiments and results. Chapter 6 discusses hit prioritisation and validation – as a true target of ezutromid as well as a valid target for utrophin modulation.

Chapter 2

Design and Synthesis of Ezutromid-based Probes for Chemoproteomics Target ID Experiments

2.1 Aims

The target(s) and mechanism of action of ezutromid were still unclear after the preliminary experiments described in the section 1.6. In order to shed more light on both target and mechanism, the strategy for this DPhil research was to identify ezutromid's direct binding partners through affinity capture, to complement phenotypic profiling studies. This chapter discusses the design and synthesis of active and inactive ezutromid-based probes for use in *in vitro* and *in situ* affinity purification. The design of negative controls, which are chemically similar to the active probes, is important to help distinguish between specific and non-specific binding proteins in affinity capture. Development of SAR around ezutromid's structure was required to determine which parts of the molecule could be substituted without abolishing activity.

2.2 Design of probes for *in vitro* affinity purification

Kubota and co-workers have suggested that the probability of target identification success can be higher with compound-immobilised matrices for *in vitro* affinity purification than with other chemical proteomics approaches, due to its well-established protocol and less extensive derivatisation required compared to PAL strategies.[115] The use of biotinylated probes for affinity purification for target identification is highly precedented. Biotin-avidin is one of the strongest known non-covalent interactions ($K_D \sim 10^{-15}$ M).[176] Along with other biotin-binding proteins streptavidin and NeutrAvidin™, avidin is a homotetrameric protein, capable of binding up to four biotin molecules. Its high specificity, high on-rate and resilience to changes in pH and temperature have long been exploited in molecular biology. The design of

biotinylated probes was chosen as a starting point for development of ezutromid-based probes.

The yeast three-hybrid experiment carried out by Hybrigenics services (section 1.6.1) used probes comprising ezutromid appended to trimethoprim *via* a polyethylene glycol (PEG) chain. The design elements of these probes was replicated in the design of biotinylated probes **41** and **43** (Figure 2.1), keeping the hydrophilic PEG chain to enhance solubility. Biotinylated probe **42** was envisaged to be a potential alternative active probe, based on the activity of the *meta*-substituted second generation utrophin modulator SMT022357 **6** (Figure 1.6).[102]

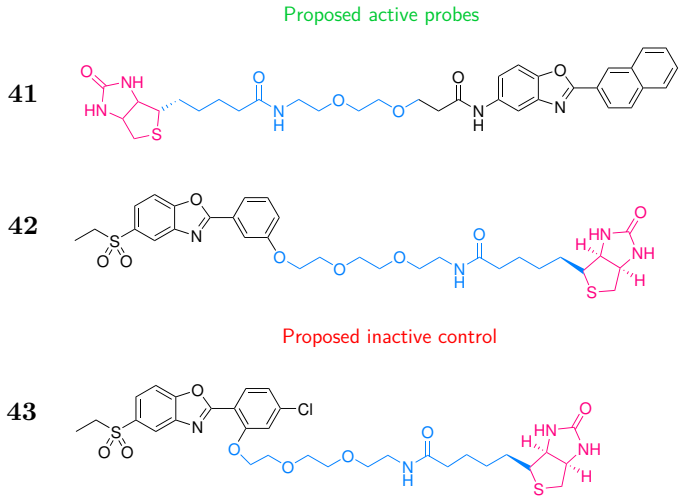


Figure 2.1: Proposed structures of active (**41** and **42**) and inactive (**43**) biotinylated probes. Blue: linker, pink: affinity matrix substrate.

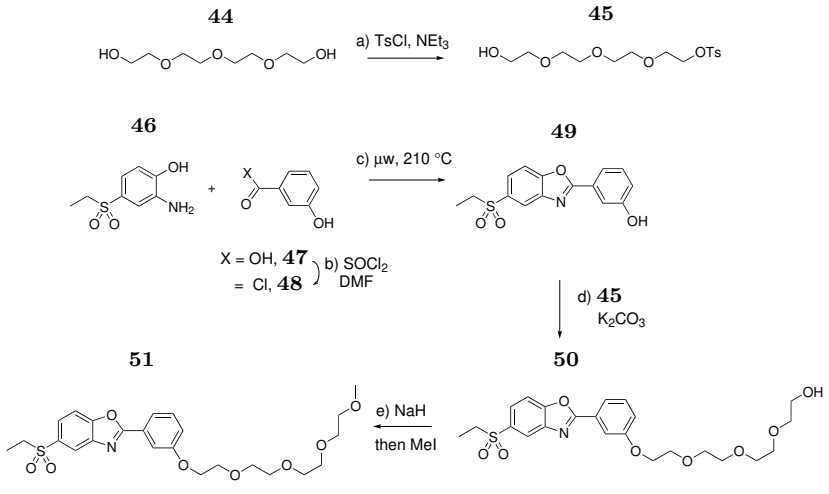
2.3 Synthesis of biotinylated probes

To test whether the *meta*-substitution proposed in probe **42** would retain ezutromid's bioactivity, compound **51** equipped with a *meta*-PEG chain was synthesised according to Scheme 2.1 to serve as a model for **42**.

First, acid chloride **48** was obtained from the corresponding acid **47** using $SOCl_2$ in quantitative yields. Next, benzoxazole ring cyclisation between aminophenol **47** and acid chloride **48** was mediated by microwave activation to afford **49**.[100] Alkylation of the phenol **49** was achieved with mono-tosylated tetraethylene glycol **45** (accessed from tetraethylene glycol **44**). Subsequent methylation yielded probe **51**.

Probe **51** was tested by Sarah Squire (Davies group, DPAG) in the H2K *mdx* cellular utrophin A-firefly luciferase (FLuc) reporter gene assay (Figure 1.5), giving an encouraging EC_{50} of 0.258 ± 0.148 µM (graph presented in Appendix A). Thus synthesis of biotinylated probe **42** was embarked on. Since the attachment of biotin often impedes cell permeability, a truncated analogue **52** containing only the linker was also desired for confirmation of bioactiv-

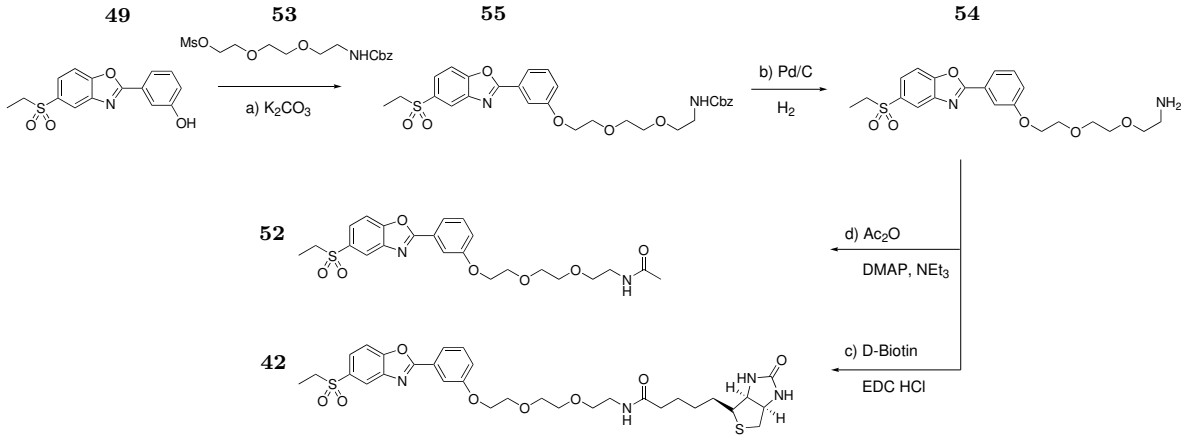
Scheme 2.1: Synthesis of ezutromid analogue **51**, with a PEG chain appended in the *meta* position of the 2-phenyl ring.[a]



[a]Reagents and conditions: (a) *p*-TsCl (0.1 eq.), NEt_3, CH_2Cl_2, 16 h, 73%; (b) $SOCl_2$, DMF, CH_2Cl_2, 0 ℃ → rt, 16 h; (b) 1,4-dioxane, 210 ℃, 15 min under microwave activation, 33% over two steps; (d) **45**, K_2CO_3, DMF, 80 ℃, 80%; (e) NaH, 0 ℃, THF, then MeI, 0 ℃ → rt, 95%.

ity in the cellular reporter gene assay. The planned synthetic route allowed for step-efficient synthesis of both analogues, beginning with alkylation of phenol **49** using linker compound **53** (kindly supplied by Aini Vuorinen[10]), followed by deprotection to yield amine **54** (Scheme 2.2). Amide coupling with D-biotin and with acetic anhydride afforded biotinylated probe **42** and truncated analogue **52** respectively.

Scheme 2.2: Synthesis of biotinylated ezutromid analogue **42**, with a PEG chain appended in the *meta* position of the 2-phenyl ring, and a truncated form without the biotin, **52**, for testing in cellular assays.[a]



[a]Reagents and conditions: (a) K_2CO_3, DMF, 80 ℃, 16 h (**53** kindly provided by Aini Vuorinen[10]); (b) Pd/C, H_2, MeOH, 16 h, 57% over two steps; (c) D-biotin, EDC hydrochloride, HOBt, NEt_3, DMF, 16 h, 71%; (d) Ac_2O, DMAP, NEt_3, $CHCl_3$, 3 h, quant.

The synthesis of biotinylated probe **41** was envisaged using the disconnections shown in Figure 2.2, starting with synthesis of linker **59**, followed by two subsequent amide couplings to attach biotin and ezutromid analogue **63** (Scheme 2.3).

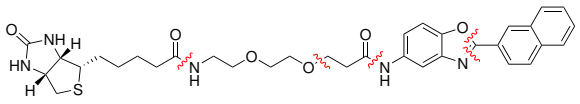
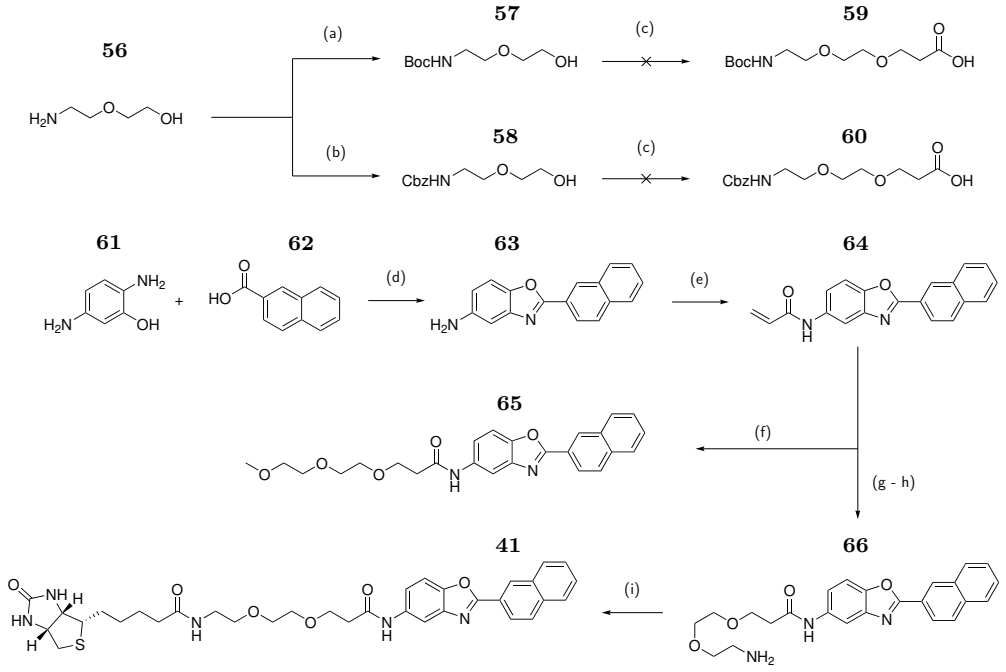


Figure 2.2: Proposed disconnections for the synthesis of biotinylated probe **41**.

However, alkylations of both Boc- and CBz-protected aminoalcohol **56** using 3-bromopropionic acid to form the linkers **59** and **60** were unsuccessful, with no conversion to product. Instead, ezutromid analogue **63**, produced *via* condensation of **61** and **62**, was coupled with acryloyl chloride to form α/β-unsaturated amide **64**. Then, by adapting a literature procedure,[177] attachment of linkers *via* a copper-catalysed Michael addition was achieved, albeit in poor yields (<15%). However, since quantities of biotinylated probe **41** and truncated version **65** sufficient for biological evaluation were obtained, further yield optimisation was not attempted.

Scheme 2.3: Synthesis of biotinylated ezutromid analogue **41** and a truncated form, without the biotin **65**, for testing in cellular assays.[a]

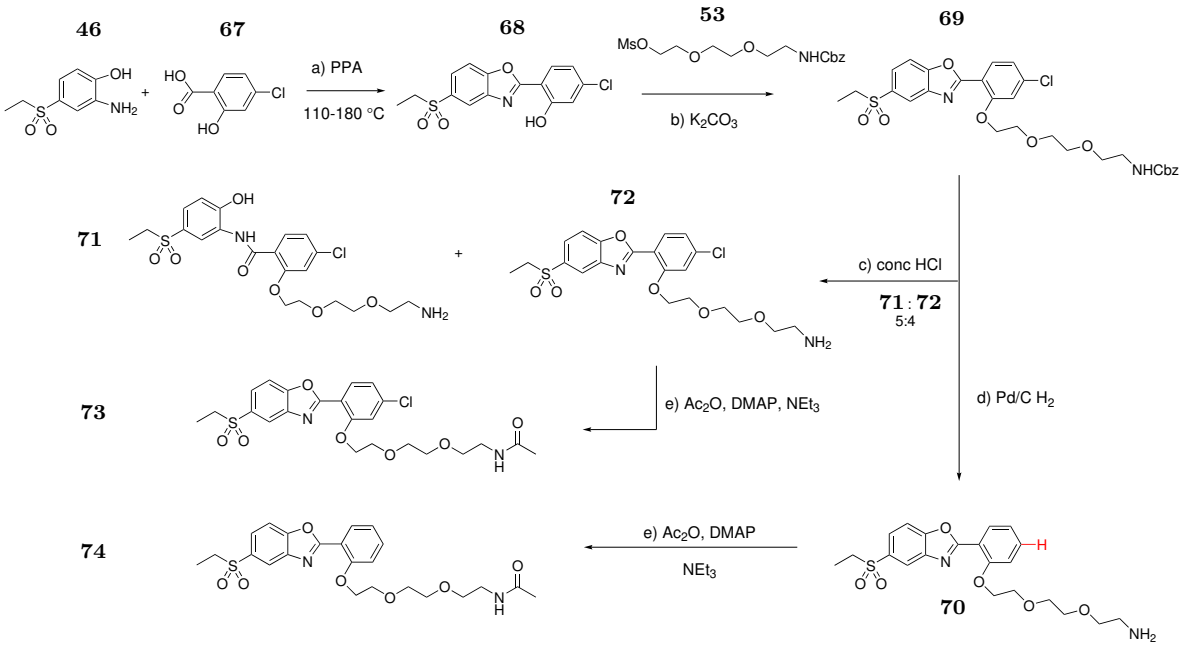


[a]Reagents and conditions: (a) KOH, H_2O, 0 °C then Boc_2O, 1 h, rt, 95%; (b) benzyl chloroformate, NEt_3, THF, 16 h, 73%; (c) NaH (5 eq.), THF, 0 °C, then 3-bromopropionic acid, −5 °C → rt, no reaction; (d) PPA, 110-180 °C, 64 h, 95%; (e) acryloyl chloride, NEt_3, CH_2Cl_2, 0 °C → rt, 16 h, 86%; (f) 2-(2-methoxyethoxy)ethanol, $CuCl_2$, Cs_2CO_3, CH_2Cl_2, reflux, 64 h, 11%; (g) **58**, $CuCl_2$ (0.1 eq.), Cs_2CO_3, CH_2Cl_2, reflux, 64 h; (h) Pd/C, H_2, CH_2Cl_2 : MeOH (1:1), 16 h; (i) D-biotin, EDC hydrochloride, HOBt, NEt_3, DMF, 16 h, 2% over 3 steps.

With biotinylated probes designed to be active in hand, attention was turned to the synthesis of inactive probe **43**. The synthesis was intended to follow an analogous route to that of probe **42**, beginning with benzoxazole cyclisation forming **68**, followed by phenol *O*-alkylation with linker **53** and hydrogenation to liberate the amine from Cbz-protected compound **69**

(Scheme 2.4). However, hydrogenation with palladium on carbon also caused inadvertent dechlorination of the benzoxazole ring, affording **70**. Alternative Cbz-deprotection using concentrated HCl (80 °C, 10 min) resulted in opening of the benzoxazole ring (**71**) as well as the desired product (**72**), in a 5:4 ratio. However, as **71** and **72** proved separable, sufficient quantities of both amines **72** and **70** were isolated for the next step: acetylation to give truncated probes **73** and **74** respectively.

Scheme 2.4: Synthesis of PEGylated ezutromid analogue **73**, and dechlorinated analogue **74** as an unintended side-product.[a]



[a]Reagents and conditions: (a) PPA, 110-180 °C, 5 h, 18%; (b) K_2CO_3, DMF, 80 °C, 16 h, (**53** kindly provided by Aini Vuorinen[10]); (c) conc. HCl, 80 °C, 10 min, 45% of **72** over two steps; (d) Pd/C, H_2, MeOH, 20 h, 65% over two steps; (e) Ac_2O, DMAP, NEt_3, $CHCl_3$, **73**: 76%, **74**: 81%.

Truncated probe analogues **52**, **65**, **73** and **74** were tested in the utrophin-FLuc reporter gene assay (**73** and **74** by collaborators at Evotec). Encouragingly, probe **52** showed activity with an $EC_{50} = 0.0606 \pm 0.0726$ µM (Figure 2.3), although the curve was bell-shaped, impairing the fit (caveats for interpretation discussed in section 2.4). Probe **65** was less potent, with no plateau reached, while probe **73** was confirmed as inactive. Interestingly, the dechlorinated probe **74**, also showed good activity, with an $EC_{50} = 0.113 \pm 0.0211$ µM. As anticipated, the biotinylated probes gave no response in the reporter assay (data presented in Appendix A).

With confirmation of truncated probe **73**'s inactivity, and the possibility of using the dechlorinated analogue as an additional active probe, a higher yielding synthetic strategy towards the full biotinylated probes **75** and **43** was now desired.

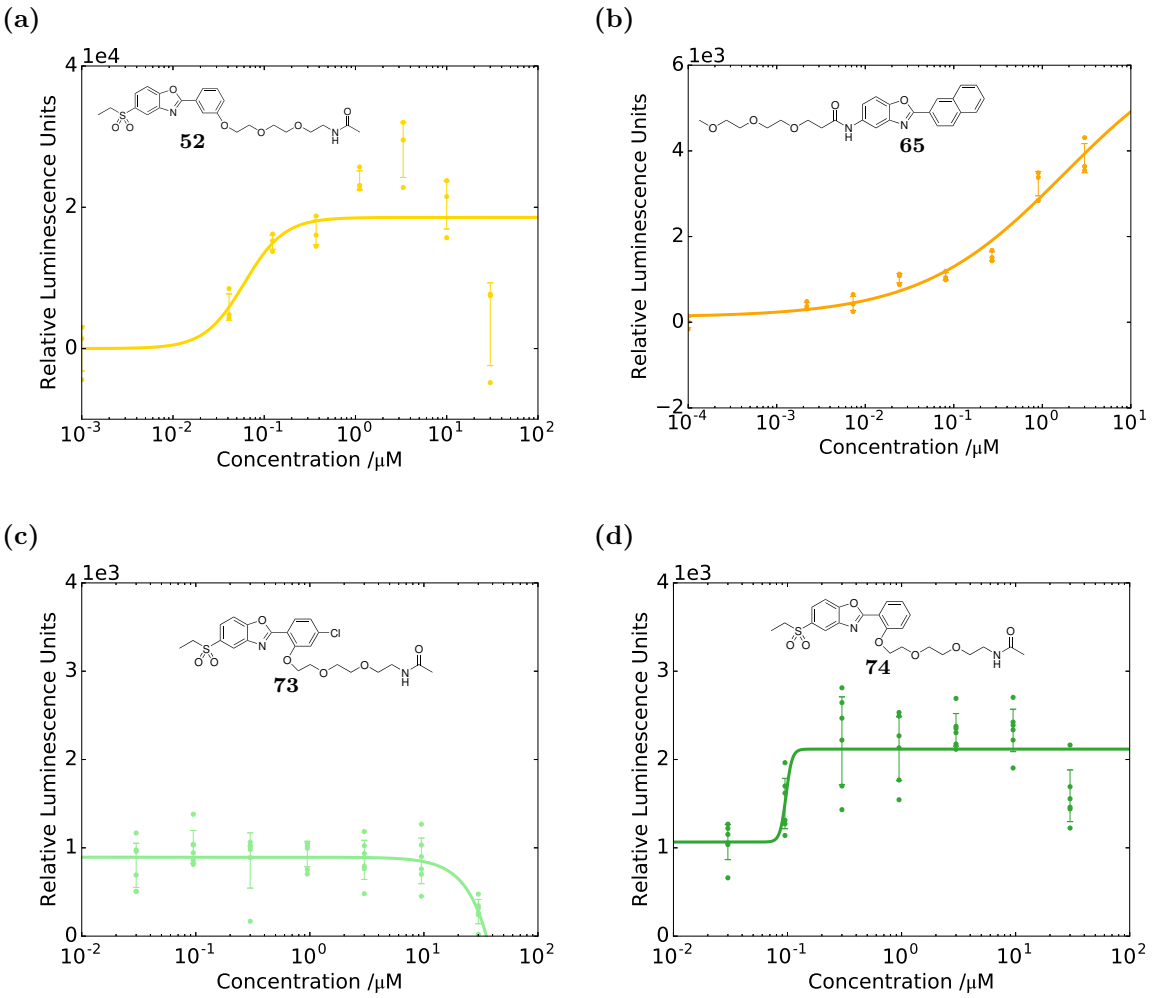
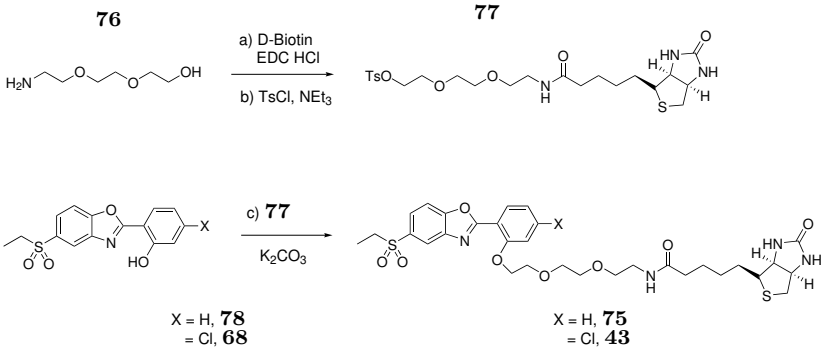


Figure 2.3: Representative concentration-response curves in the H2K *mdx* utrophin-FLuc reporter gene assay for PEGylated ezutromid probes **(a) 52**; **(b) 65**; **(c) 73**; **(d) 74**. Data for **73** and **74** collected by collaborators at Evotec.

Aminoalcohol **76** was first coupled to D-biotin, then activated as the tosylate **77** to enable alkylation of phenols **78** and **68** and yield the final biotinylated probes **75** and **43** respectively (Scheme 2.5).

Scheme 2.5: Synthesis of biotinylated ezutromid analogues **43** and **75**.[a]



[a]Reagents and conditions: (a) D-biotin, EDC hydrochloride, HOBt, NEt_3, DMF, 16 h; (b) *p*-TsCl, NEt_3, CH_2Cl_2, 6 h, 48% over two steps; (c) **77**, K_2CO_3, DMF, 80 °C, 2 h, X = H 85%, X = Cl 12%.

2.4 Caveat for using FLuc reporter gene assays

Firefly luciferase reporter gene assays are susceptible to false positive and negative artefacts caused by small molecule inhibition and stabilisation of the FLuc enzyme (discussed in depth in chapter 3). Ezutromid itself displays a bell-shaped curve, distinctive of FLuc inhibitors which stabilise FLuc and inhibit the bioluminescence reaction at high concentrations (Figure 1.5). Therefore, evaluation of the ezutromid-based probes for FLuc inhibition was required to validate the interpretation of the results from the utrophin-Fluc reporter gene assay. This was assessed using recombinant FLuc *in vitro* and measuring luminescence output in the presence of increasing doses of compound (Figure 2.4). Truncated probes **74** and **73** with *ortho*-substituted PEG chains did not show FLuc inhibition, therefore their results in the reporter gene assay were unlikely to be affected by artefacts. However, the $\leq$ micromolar FLuc inhibition exhibited by probes **52** and **65** meant that their encouraging reporter gene assay results may be false positives.

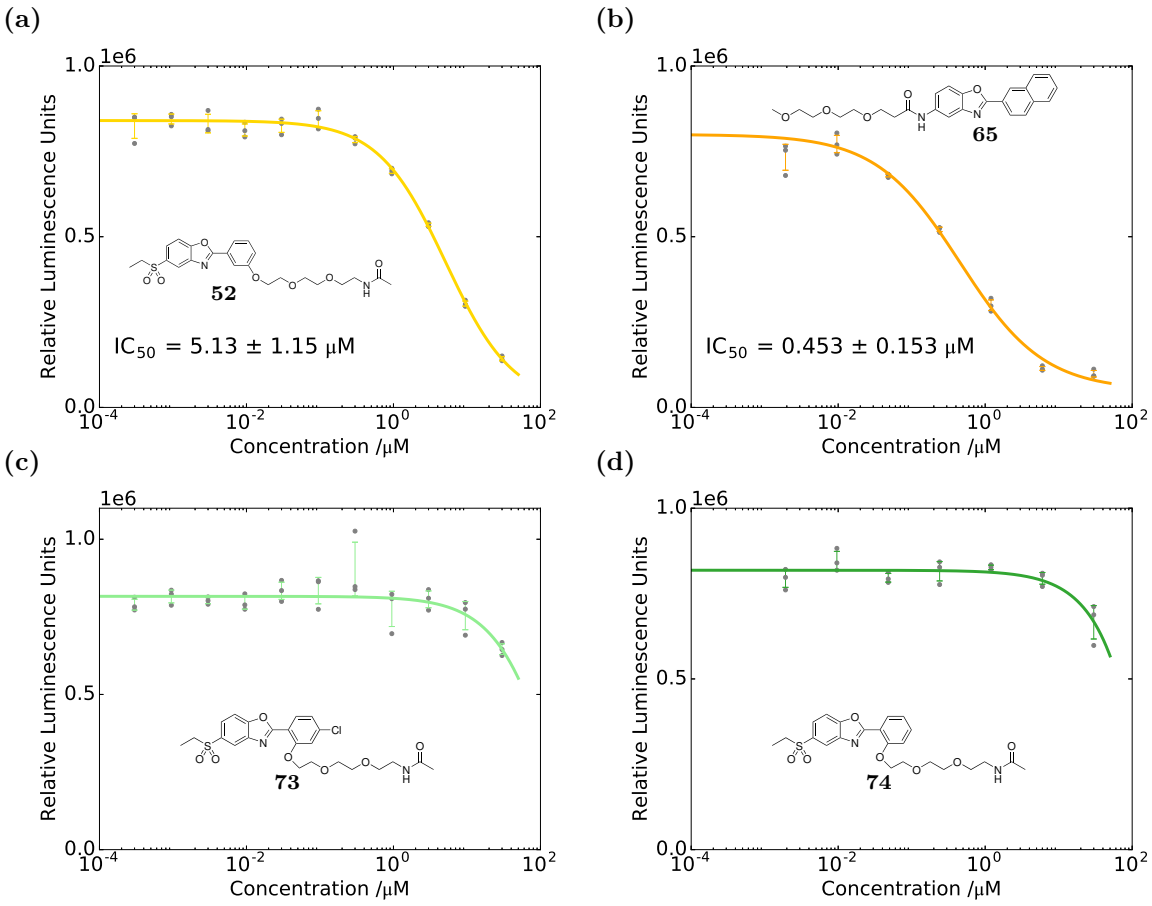


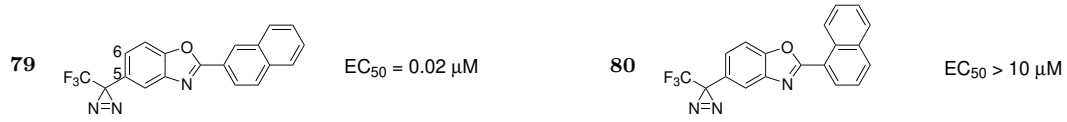
Figure 2.4: **(a)** Inhibition of FLuc activity by PEGylated ezutromid probe **52**; **(b)** submicromolar inhibition of FLuc activity by probe **65**; **(c)**-**(d)** probes **73** and **74** do not inhibit significantly inhibit FLuc below 30 μM.

2.5 Design of cell-permeable photoaffinity probes for *in situ* affinity capture

Initial results from testing non-FLuc inhibiting **74** and **73** in the utrophin reporter gene assay indicated that their biotinylated counterparts **75** and **43** would be suitable for use in AfBPP experiments as active and inactive probes respectively. However, as discussed in section 1.5.3, there are certain limitations to consider with using biotinylated probes: first, cell lysates must be used rather than live cells, decreasing physiological relevance; second, non-covalent binding of probe to target reduces the stringency of wash steps employed and thus can lead to an increase in background of non-specifically binding proteins; third, weakly binding targets may be missed. This is particularly pertinent due to ezutromid's unknown target binding affinity. Therefore, photoaffinity probes for *in situ* affinity capture were desired for use in parallel to aid target identification.

2.5.1 Incorporation of a photoaffinity label into ezutromid-based probes

Of popular options for photoaffinity groups (discussed in section 1.5.3), the trifluoromethylphenyl diazirine was chosen for ezutromid due to its small size and excellent PAL profile. Previous work in the group resulted in the development of diazirine analogue **79**, with ezutromid's 5-position ethylsulfonyl moiety substituted for a trifluoromethyl diazirine[178] (Figure 2.5). The rationale for this design was the close match between the sterics and electronics of the diazirine and the sulfonyl group. **79** was found to be active in the utrophin-Fluc reporter gene assay with an EC_{50} comparable to that of ezutromid.[178] An additional analogue, **80**, was designed on the basis of previous SAR studies[100] to be a negative control (Figure 2.5). The chemical properties of **80** are similar to those of PAL-labelled ezutromid analogue **79** yet **80** has been found to be inactive in utrophin upregulation.[178] This negative control could be useful to distinguish non-specifically binding proteins identified by affinity purification, which are especially likely given ezutromid's small size and hydrophobicity. Additional negative controls would aid this target convolution further still.

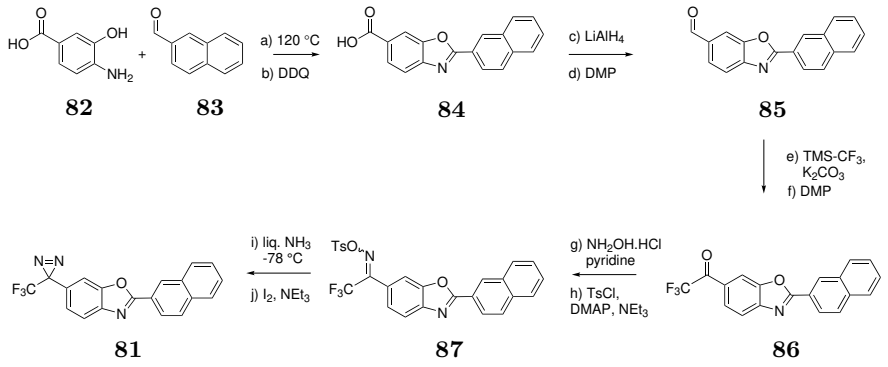


(a) **(b)**

Figure 2.5: Structures of previously synthesised photoaffinity probes for ezutromid. The reporter assay was carried out by Sarah Squire (Davies group, DPAG), data presented in Appendix A.

Since changing regiochemistry from 5-substituted to 6-substituted benzoxazoles was also known to ablate bioactivity,[100] a second negative control **81** was designed, the synthesis for which is described in Scheme 2.6. The disadvantage of the choice of diazirines as PALs is their relatively lengthy synthesis (reviewed extensively elsewhere[179,180]). However, once incorporated they are stable to a variety of conditions (strong acid, base, high temperature, oxidation and reduction), allowing for further synthetic manipulations.

Scheme 2.6: Synthesis of negative control diazirine **81**[a]



[a]Reagents and conditions: (a) *p*-xylene, 140 °C, 5 h; (b) DDQ, 1,4-dioxane, rt, 16 h, 44% over 2 steps; (c) $LiAlH_4$, THF, 0 °C; then reflux, 3 h, 99%; (d) DMP, CH_2Cl_2, rt, 16 h, 75%; (e) TMS-CF_3, K_2CO_3, DMF, rt, 16 h, quant; (f) DMP, MeCN, reflux, 5 h, 79%; (g) $NH_2OH.HCl$, pyridine, EtOH, 3 Å mol. sieves, reflux, 16 h, 91%; (h) *p*-TsCl, DMAP, NEt_3, CH_2Cl_2, 40 °C, 16 h, 85%; (i) liq. NH_3, CH_2Cl_2, −78 °C → rt, 6 h, 76% (94% brsm); (j) I_2, NEt_3, CH_2Cl_2, 0 °C → rt, 15 min, 64% (90% brsm).

Starting materials **82** and **83** were refluxed together to drive imine formation, which was immediately followed by oxidative ring closure mediated by DDQ to give acid **84**. Next, oxidation level manipulation was required, beginning with reduction of acid **84** to the alcohol, followed by DMP oxidation to the aldehyde **85**. Perfluoroalkylation of **85** was achieved smoothly with Ruppert's reagent in quantitative yield,[181] followed by a second DMP oxidation step to yield trifluoromethyl ketone **86**. Compound **86** was converted to an oxime, then activated as the tosyl oxime **87** (geometric isomers (4:7), uncharacterised as *E* or *Z*). Treatment of **87** with liquid ammonia furnished the diaziridine, which was readily oxidised with iodine to deliver diazirine **81** in a yield of 10% over 10 steps. The overall yield was diminished by poor performance of the last two steps, where appreciable quantities of starting material were recovered (19-29%).

With the two photoaffinity-labelled probes designed as negative controls in hand, confirmation of their inactivity was desired. 1-naphthyl probe **80** had already displayed no activity in the primary utrophin-FLuc reporter gene assay (Figure 2.5b). Testing in the FLuc inhibition assay showed that **80** did not inhibit FLuc < 10 µM (data presented in Appendix B), indicating that its reporter gene assay result was a true negative. Evaluation of 6-substituted probe

81 in the utrophin-FLuc reporter gene assay (performed by Sarah Squire, DPAG) showed no activity until the highest dose of 10 µM (Figure 2.6a), but **81** was found to moderately inhibit FLuc (Figure 2.6b), raising concerns about the reporter gene assay result.

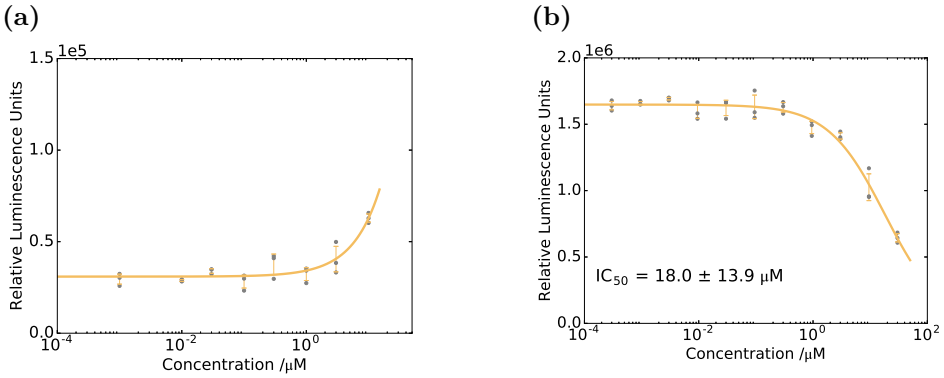


Figure 2.6: **(a)** Concentration-response curve for 6-substituted photoaffinity probe **81** in the utrophin-FLuc reporter gene assay indicates inactivity (data collected by Sarah Squire, DPAG); **(b)** Inhibition of FLuc activity by **81** with an approximate IC_{50} of 18 µM, no plateau reached.

To verify the inactivity of **80** and **81** they were dosed (0.03, 0.3, 3 µM, in triplicate) onto H2K *mdx* myoblasts for 24 h, alongside the vehicle and the endogenous utrophin modulator heregulin as a positive control. Levels of utrophin protein were evaluated by western blot analysis, and fold change relative to DMSO calculated (Figure 2.7). Heregulin was found to upregulate utrophin ~1.6 fold (in keeping with reported results[182]), while probes **80** and **81** appeared inactive, although weak activity (< 1.2 fold) was seen at the highest concentration of probe **80**.

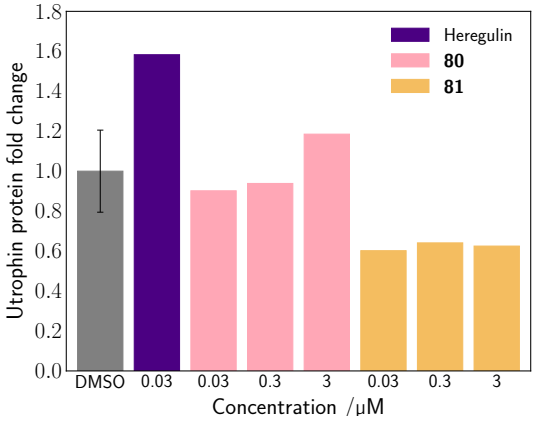


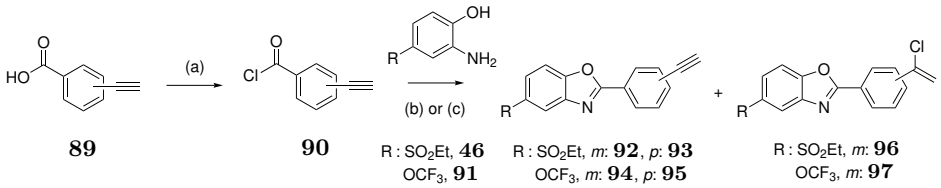
Figure 2.7: Heregulin upregulated utrophin ~1.6 fold in H2K *mdx* myoblasts, while probes **80** and **81** were inactive. Error bars are standard deviation of technical replicates.

Therefore, with sites for photoaffinity label attachment and ablation of activity identified, attention was turned to identifying positions dispensable for bioactivity in order to append a 'click' handle.

2.5.2 Incorporation of a 'click' handle into ezutromid-based probes

Options for click handles were discussed in section 1.5.3, and an alkyne for CuAAC was selected for its small size and well-precedented use in target ID affinity chromatography. *Meta*- and *para*-phenyl acetylene substituents (**92**, **93**, **98** and **99**, Table 2.1) were designed to replace ezutromid's naphthyl ring with minimal impact on the size of the overall compound. A bioisostere for the trifluoromethyl diazirine was desired in the 5 position, to allow approximation of the SAR of the final dual-tagged photoaffinity probes, without the lengthy diazirine syntheses. The trifluoromethoxy group was chosen as this bioisostere, based on its electronic and steric similarity to the trifluoromethyl diazirine, and the commercial availability of starting material **91**. Hence, compounds **94**, **95**, **100** and **101** were synthesised to act as models for the diazirine analogues of ezutromid to help elaborate the probe SAR.

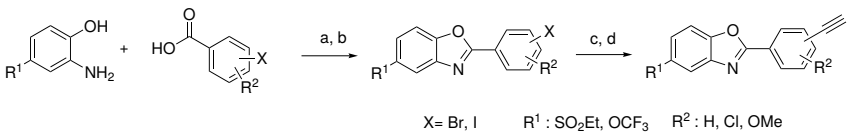
Scheme 2.7: Initial route for the synthesis of alkyne-tagged ezutromid probes.[a]



[a]Reagents and conditions: (a) $SOCl_2$, DMF, CH_2Cl_2, 0 ℃ → rt, 16 h; (b) 1,4-dioxane, 210 ℃, 15 min under microwave activation; (c) *p*-xylene, 110 ℃, 2 h; then $MeSO_3H$ (0.2 eq.) room temperature; then reflux, 16 h.

Initially, the synthesis of these probes was conceived as a two-step process from commercial reagents (Scheme 2.7). Generation of acid chlorides **90** from the corresponding acids **89** ($SOCl_2$, 0 ℃ → rt) proceeded in quantitative yields. However, subjection of these acid chlorides to microwave radiation at 210 ℃ with aminophenols **46** and **91** was met with mixed success. Whilst the *para*-alkynes proceeded to product in acceptable yields (R = $-SO_2Et$, **93**: 40%, R = $-OCF_3$, **95**: 41%) the reactions with the *meta*-alkynes produced significant quantities of side products **96** and **97**. The *meta*-alkynes were found to be prone to Markovnikov side-reactions with addition of HCl (liberated from the acid chloride during cyclisation). This was attributed to the greater availability of the *meta*-alkyne electrons, which cannot be delocalised to the oxygen or nitrogen in the benzoxazole core, compared to those of the *para*-alkyne. The desired alkyne **94** and side-product alkene **97** (8:1 ratio) could be separated on alumina gel. However, for the ethylsulfonyl derivative, the side-product **96** was the major product and the desired alkyne **92** undetectable. It was hypothesised that the microwave activation and extreme high temperature (210 ℃) may be responsible for the observed side-reaction.

Scheme 2.8: General route for the synthesis of alkyne-tagged ezutromid probes.[a]



[a]Reagents and conditions: (a) $SOCl_2$, DMF, CH_2Cl_2, 0 ℃ → rt, 16 h; (b) 1,4-dioxane, 210 ℃, 15 min under microwave activation; (c) TMS-acetylene, $Pd(PPh_3)_2Cl_2$, CuI, NEt_3, THF, 70 ℃, 1 h; (d) TBAF, THF, rt, 30 min.

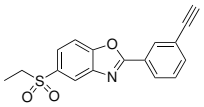
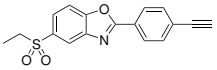
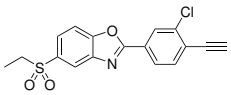
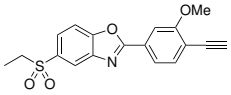
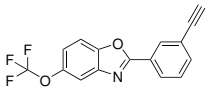
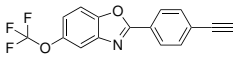
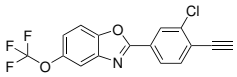
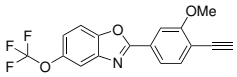
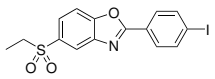
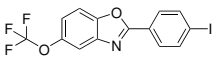
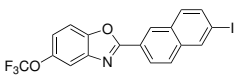
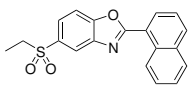
Compound	Structure	Yield /% steps a-b	Yield /% steps c-d	Utrophin A-FLuc reporter assay EC_{50} /µM	FLuc inhibition IC_{50} /µM
92		69	57	0.00686± 0.0335*	0.375 ± 0.0691
93		40†	–	1.22± 1.51	> 30
98		52	62	> 10	> 30
99		78	66	> 10	> 30
94		5†	–	1.86 ± 0.353	4.14 ± 0.828
95		41†	–	> 10	2.58 ± 0.891
100		76	42	> 10	6.73 ± 1.38
101		62	71	> 10	> 10
102		72	–	0.0514 ± 0.0139	4.64 ± 0.867
103		51	–	> 10	> 30
104		35	–	> 10	> 10
105		97	–	> 10	3.86 ± 0.974

Table 2.1: Structure-activity relationships for variation of the 2-aryl substituent in the H2K *mdx* utrophin A-FLuc reporter assay and in FLuc inhibition. The reporter assay was carried out by Sarah Squire (Davies group, DPAG). Graphs for both assays presented in Appendices A and B. Values were calculated by fitting a four-parameter logistic function to the data, error indicates goodness of fit.
† Synthesis according to Scheme 2.7, * bell-shaped result.

Therefore, an alternative approach was trialled: aminophenol **46** and acid chloride **90** were heated (110 ℃), followed by addition of a catalytic quantity of methanesulfonic acid to form the benzoxazole. Unfortunately, this method also gave only the undesired side-product **96**.

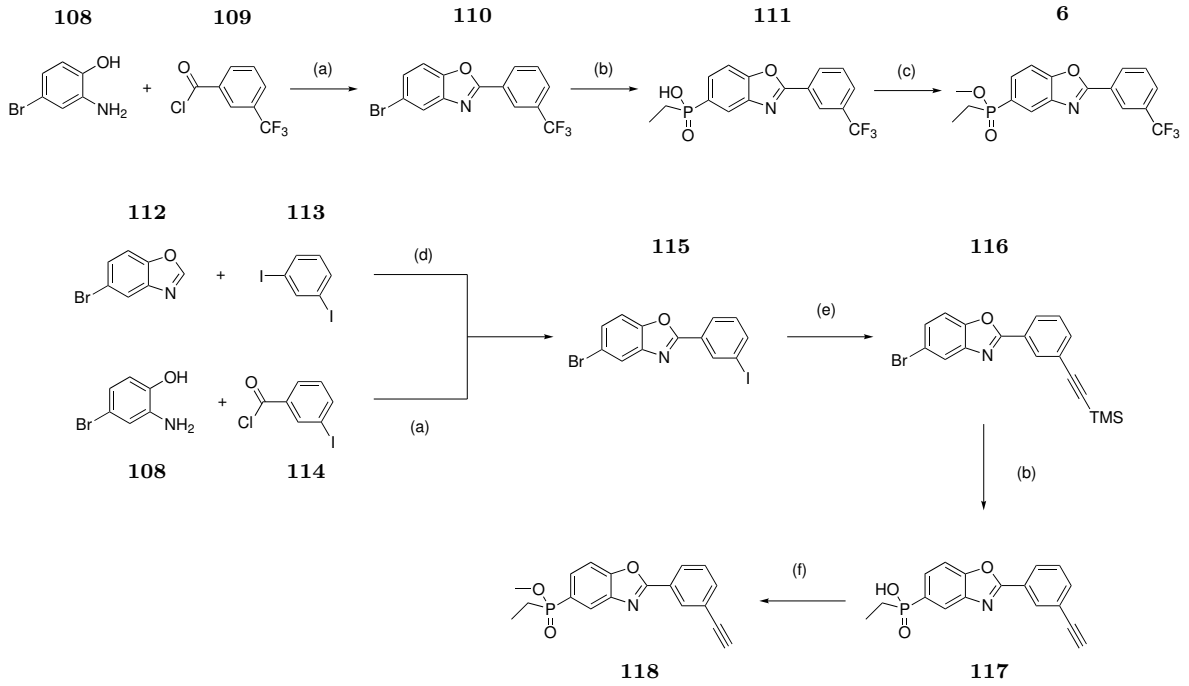
Synthesis of **92** and the remaining alkyne-tagged probes was instead achieved according to general Scheme 2.8, each beginning with formation of an acid chloride from the corresponding halobenzoic acid, followed by microwave-assisted benzoxazole cyclisation with an aminophenol as described previously (yields reported in Table 2.1). Sonagashira couplings with TMS-acetylene and subsequent TMS deprotection with TBAF were then performed to install the alkynes. The activity of the synthesised probes was investigated along with intermediate aryl iodides **102**, **103** and **104** in the utrophin A-FLuc reporter gene assay (performed by Sarah Squire, DPAG). Unfortunately, few alterations to the ezutromid's structure were tolerated (Table 2.1). Ethylsulfonyl-substituted probes were typically more active than their trifluoromethoxy-substituted counterparts, for example for *meta*-alkynes **92** (EC_{50} = 6.7 nM) vs. **94** (1.86 µM) or *para*-alkynes **93** (1.22 µM) vs. **95** (>10 µM). Combining *meta*-substitutions (−OMe and −Cl) with *para*-alkynes to approximate the steric size of the naphthyl ring in the parent compound ezutromid did not yield increased activity. The aryl iodide **102** had a potent EC_{50} of 0.05 µM but the trifluoromethoxy analogues **103** and **104** were far less active.

Assessment of FLuc inhibition revealed that the probes with most promising EC_{50} results were also potent FLuc inhibitors, contravening interpretation of their activity as utrophin modulators. However, interestingly, *para*-alkyne **93** exhibited activity independent of FLuc inhibition, making a *para*-alkyne containing a 5-trifluoromethyl diazirine (**106**, Scheme 2.10) a candidate for an active AfBPP probe. Further, the most active compound, *meta*-alkyne **92**, displayed activity with $EC_{50} > 50$ times lower than its FLuc inhibition IC_{50}, suggesting that a *meta*-alkyne AfBPP probe (**107**) could potentially also be active.

In addition to the negative control probes, a competition control was desired for the chemical proteomics experiments. This competitor compound is co-treated at an excess concentration to occupy the target binding sites and block the binding of the photoaffinity probe. Proteins that are identified in the AfBPP workflow but are absent from the competition control samples are likely to be specific hits. Unfortunately, ezutromid's thermodynamic solubility of 3 µM makes it a poor candidate for the competition control. However, the phosphinate series of second generation utrophin modulators (e.g. SMT022357 **6**, Figure 1.6) had been developed with superior physicochemical properties, and significantly improved aqueous solubility,

allowing for dosage at the excess concentrations appropriate for competition experiments. As a result, the synthesis of **6** and alkyne derivative **118** was undertaken (Scheme 2.9), route adapted from that developed by the group and collaborators at Evotec.

Scheme 2.9: Synthesis of phosphinate second generation utrophin modulators **6** and **118**.[a]



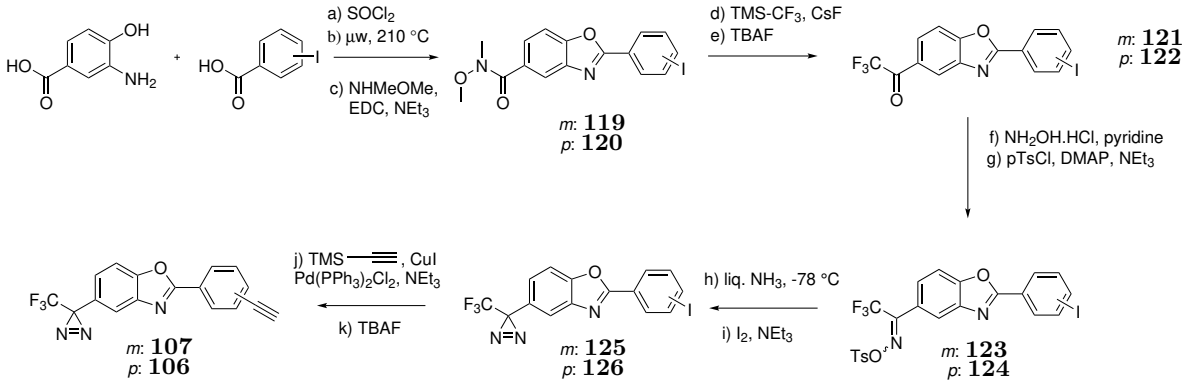
[a]Reagents and conditions: (a) 1,4-dioxane, 210 ℃, 15 min under microwave activation, **110**: 70%, **115**: 72%; (b) ethyl phosphinic acid, Xantphos, palladium(II) acetate, DIPEA in DME and toluene, 105 ℃, 5 h, **111**: 69%; (c) $SOCl_2$, DMF, reflux, 2 h, then MeOH, DMAP, DIPEA, CH_2Cl_2, 0 ℃ → rt, 3 h, 63%; (d) palladium(II) acetate, PPh_3, Cs_2CO_3, copper(I) iodide, DMF, 100 ℃, 6 h, 13%; (e) TMS-acetylene, $Pd(PPh_3)_2Cl_2$, CuI, NEt_3, THF, 70 ℃, 1 h, 90%; (f) TMS-diazomethane, MeOH, toluene, 26% over two steps from **116**.

2-Arylbenzoxazoles **110** and **115** were synthesised from the corresponding aminophenol and acid chloride as described previously. An alternative route to **115**, using palladium-catalysed C-H activation of benzoxazole **112** was also attempted, but proceeded in a poorer yield (13% vs. 72%) and therefore was not pursued further. Phosphinic acids **111** and **117** were installed *via* palladium-catalysed cross-coupling with ethyl phosphinic acid. Formation of the methylated phosphinate was carried out *via* either activation of **111** as the phosphinic chloride, followed by methanol addition-elimination to afford **6** or treating **117** with TMS-diazomethane to give **118**. SMT022357 **6**'s ability to modulate utrophin and diminish dystrophic pathology in *mdx* mice has been previously reported.[102] In the FLuc reporter gene assay, alkyne-tagged phosphinate **118** demonstrated an encouraging EC_{50} of 88 nM (data presented in Appendix A). However, the result was marred by a marked bell-shaped curve, reflecting compound **118**'s potent FLuc IC_{50} of 66 nM (data presented in Appendix B). Given the ambiguity of compound **118**'s activity, SMT022357 **6** was proposed as the competitor for the planned chemical proteomics experiments.

2.6 Synthesis of dual-tagged photoaffinity probes

Given the indications of activity of *para*- and *meta*-alkyne probes **93** and **92**, the synthesis of *para*- and *meta*-alkyne probes **106** and **107** containing 5-trifluoromethylphenyl diazirines was embarked on (Scheme 2.10). The complicating factor of FLuc inhibition by ezutromid analogues required the dual-tagged probes to be tested subsequently in orthogonal activity assays.

Scheme 2.10: Synthesis of dual-tagged ezutromid analogues **106** and **107**[a]



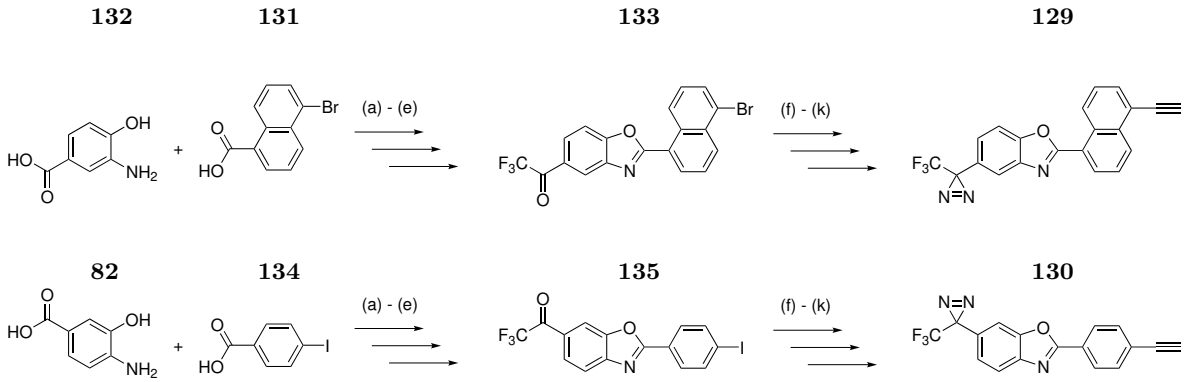
[a]Reagents and conditions: (a) $SOCl_2$, DMF, CH_2Cl_2, 0 ℃ → rt, 16 h; (b) 1,4-dioxane, 210 ℃, 15 min under microwave activation; (c) NHMeOMe.HCl, EDC.HCl, NEt_3, DMF, rt, 16 h, 45-49% over 3 steps; (d) TMS-CF_3, CsF, toluene, rt, 3 days; (e) TBAF, THF, 50 ℃, 3 h; (f) NH_2OH.HCl, pyridine, EtOH, 3 Å mol. sieves, reflux, 16 h; (g) *p*-TsCl, DMAP, NEt_3, CH_2Cl_2, 40 ℃, 16 h, 20-39% over 4 steps; (h) liq. NH_3, CH_2Cl_2, −78 ℃ → rt, 6 h; (i) I_2, NEt_3, CH_2Cl_2, 0 ℃ → rt, 15 min, 95-99% over 2 steps; (j) TMS-acetylene, $Pd(PPh_3)_2Cl_2$, CuI, NEt_3, THF, 70 ℃, 1 h; (k) TBAF, THF, rt, 30 min, 50-70% over 2 steps.

The previous photoaffinity probe synthesis (Scheme 2.6) was significantly lengthened by oxidation and reduction steps, and an alternative approach was desired. It was noticed that formation of Weinreb amides **119** and **120** from the corresponding carboxylic acids could subsequently furnish the trifluoromethyl ketones **127** and **128** in one step using the Ruppert-Prakash reagent,[183] thus reducing overall step count to this intermediate by two. The synthesis of the TPDs from trifluoromethyl ketones proceeded as before, starting from oxime formation and activation (*O*-tosylation) followed by liquid ammonia induced diaziridine cyclisation and lastly oxidation to afford the diazirine. Sonagashira couplings with TMS-acetylene furnished the final dual-tagged probes **106** and **107** in overall yields of 5-13% over 11 steps.

Attention was next turned to the synthesis of negative control dual-tagged probes. 1-naphthyl **80** and 6-substituted **81** photoaffinity probes had been determined as inactive in the utrophin western blot assay (Figure 2.7). Therefore, alkyne-bearing dual-tagged probes **129**, containing a 1-naphthyl ring, and **130**, with the 6-substitution pattern, were synthesised to serve as negative controls in the AfBPP experiments (Scheme 2.11). The placement of the

alkyne in probe **129** was due to the commercial availability of bromo-1-naphthyl starting material **131** and the synthetic route used was performed as previously described in Scheme 2.10. Briefly, the synthesis proceeded from benzoxazole cyclisation, TMS-CF_3-mediated installation of the trifluoromethyl ketone, 4-step diazirine synthesis and finally Sonagashira alkyne coupling. Sonagashira coupling with the aryl bromide substrate to afford **129** proceeded slowly at 70 ℃ and was therefore prematurely halted after 1 h to limit undesired reactions at the diazirine. Probe **129** and **130** were accordingly obtained in overall yields (11 steps) of 8% each.

Scheme 2.11: Synthesis of dual tagged ezutromid analogues **129** and **130**[a]



[a]Reagents and conditions: (a) $SOCl_2$, DMF, CH_2Cl_2, 0 ℃ → rt, 16 h; (b) 1,4-dioxane, 210 ℃, 15 min under microwave activation; (c) NHMeOMe.HCl, EDC.HCl, NEt_3, DMF, rt, 16 h, 58-62% over 3 steps; (d) TMS-CF_3, CsF, toluene, rt, 3 days; (e) TBAF, THF, 50 ℃, 3 h; (f) NH_2OH.HCl, pyridine, EtOH, 3 Å mol. sieves, reflux, 16 h, 29-56% over three steps; (g) *p*-TsCl, DMAP, NEt_3, CH_2Cl_2, 40 ℃, 16 h, 74-77%; (h) liq. NH_3, CH_2Cl_2, −78 ℃ → rt, 6 h, quant.; (i) I_2, NEt_3, CH_2Cl_2, 0 ℃ → rt, 15 min, 94-99%; (j) TMS-acetylene, $Pd(PPh_3)_2Cl_2$, CuI, NEt_3, THF, 70 ℃, 1 h; (k) TBAF, THF, rt, 30 min, 34-59% over 2 steps.

2.7 Initial biological testing of dual-tagged photoaffinity probes

With active and inactive dual-tagged probes prepared, confirmation of bioactivity was required. Testing in the utrophin-FLuc reporter assay showed micromolar activity for probes **106** and **107**, weaker than that of ezutromid and with no plateau reached. No appreciable activity was observed for negative control probes **129** and **130** (Figure 2.8a).

As with the previously synthesised probes, assessment of FLuc inhibition was also required, to validate interpretation of the results from the reporter gene assay. Active probes **106** and **107** had significant FLuc inhibition IC_{50}s of 1.12 ± 0.55 µM and 2.29 ± 0.78 µM while inactive probes **129** and **130** inhibited FLuc to a lesser extent (Figure 2.8b). These results suggest that probes **129** and **130** are suitable inactive controls, but that secondary assays, independent of FLuc reporters, are required to prove bioactivity for compounds **106** and **107**.

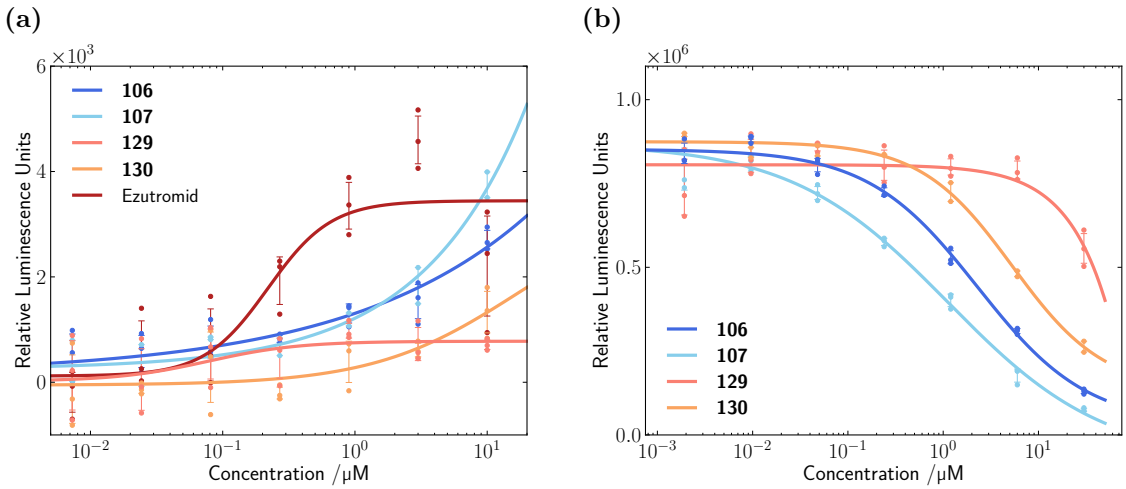


Figure 2.8: **(a)** Probes **106** and **107** show activity in the utrophin-FLuc reporter gene assay, although to a lesser extent than ezutromid. Negative control probes **129** and **130** have negligible activity. N=4, representative graph shown; **(b)** probes **106** and **107** significantly inhibit FLuc below 10 μM. Negative control probes **129** and **130** inhibit FLuc to a lesser extent.

2.8 Summary

In conclusion, probes for both immobilised matrix and *in situ* approaches to target enrichment by affinity capture were synthesised, to maximise chances of successful target identification. Three proposed active and one inactive biotinylated probes were developed, drawing inspiration from existing SAR around ezutromid's structure. After the synthesis of a series of alkyne-bearing ezutromid analogues, two active and two inactive probes equipped with both a photoaffinity label and click handle were synthesised. A phosphinate alkyne analogue was also developed to serve as a competitor control for the affinity capture experiments. All the probes were tested in the primary utrophin reporter gene assay, however, in many cases, particularly those with structures most similar to ezutromid's, valid interpretation of their bioactivity was prevented by assay artefacts caused by FLuc inhibition. For these compounds, additional testing for bioactivity in secondary assays was required, the results of which are discussed in chapter 4. To improve the probe development process, it was of interest to characterise ezutromid's inhibition of firefly luciferase. The experiments carried out to characterise the mechanism of ezutromid's FLuc inhibition are described in chapter 3.

Chapter 3

Characterisation of Ezutromid as a Firefly Luciferase Inhibitor

3.1 Aims

The discovery of FLuc-related artefacts in the primary utrophin reporter gene assay results for the probes developed in the previous chapter led to the desire to investigate and overcome their cause. Understanding the mechanism by which ezutromid inhibits firefly luciferase has the potential to improve the probe development process, by allowing rational design of probes free from FLuc inhibition. Further, this knowledge could inform the development of another series of utrophin modulators that was undergoing preclinical development and was also affected by FLuc inhibition. Finally, FLuc could serve as a model system for techniques to determine ligand binding sites, which could be later applied to validate targets identified through the AfBPP experiments.

3.2 Introduction to Firefly Luciferase (FLuc)

3.2.1 FLuc as a tool for molecular biology and drug discovery

Firefly luciferase (FLuc) is widely used in academia and industry due to its excellent sensitivity and dynamic range, and its ease of use.[184] Firefly luciferase (*Photinus pyralis*) is a bioluminescent enzyme, enabling sensitive detection because a light excitation source is not required, reducing the background signal.[185] Dynamic changes in luciferase abundance can be detected due to its short half-life of 3-4 hrs in mammalian cells.[186] FLuc is used in a broad range of often high-throughput applications including reporter gene assays, *in vivo* imaging, cell viability assays, cAMP sensing for cell signalling assays and more.[185]

However, FLuc assays are sensitive to positive and negative artefacts arising from compound interferences.[184] Studies conducted by Auld *et al.* have shown that ~5% of compounds

in large libraries interfere with FLuc at 11 µM,[187] while ∼6% of the GSK published protein kinase inhibitor set are FLuc inhibitors.[188] Furthermore, up to 98% of compounds identified as high throughput screen (HTS) hits from FLuc reporter assays have been found to be FLuc inhibitors.[189,190] Many commonly used tool compounds such as p53-inhibitor pifithrin-α,[191] resveratrol[192] and aryl hydrocarbon receptor inhibitor β-naphthoflavone[193] have also been found to be FLuc inhibitors. Consequently, hit and tool compounds should be routinely assessed for FLuc inhibition to validate interpretation of results from FLuc assays.

FLuc interferences are typically due to direct binding of the compound to luciferase rather than through light attenuation or scattering.[187,194] Compound binding has been found to impart a thermal stabilisation on the FLuc enzyme, which increases its half-life independently of transcription or translation.[186,195] This stabilisation effect is amplified by the short half-life of FLuc coupled with long compound incubation times (12-48h) typically seen with reporter gene assays.[196] Increases in luciferase abundance by compound binding can lead to false positives, while false negatives arise from inhibition of the bioluminescence reaction.

Firefly luciferase's catalytic bioluminescence reaction proceeds *via* a relatively well characterised multistep mechanism (Figure 3.1a). FLuc binds substrates D-luciferin (D-LH_2) **136** and ATP to form the reactive intermediate luciferyl-adenylate (LH_2-AMP) **137**.[197] Upon adenylation, FLuc undergoes a conformational change.[198] Then, LH_2-AMP reacts with molecular oxygen to form an electronically excited dioxetanone **138** which collapses to generate oxyluciferin **139** in an excited state and carbon dioxide.[199] Relaxation to the ground state of oxyluciferin releases a photon of light of wavelength 560-610 nm.[199] FLuc can also perform 'dark' reactions such as oxidation of the intermediate LH_2-AMP to form the potent tight-binding inhibitor L-AMP **140** which is responsible for the flashing of firefly lanterns.[194,197,200]

3.2.2 Mechanisms of FLuc inhibition

FLuc inhibitors are typically linear, planar small molecules and some related structure-activity-relationships have been established in the literature.[187,194,201] Interestingly, inhibitors comprise a broad range of chemotypes (Figure 3.1b), revealing FLuc to be a highly promiscuous enzyme. This wide structural diversity complicates prediction of compounds as FLuc inhibitors, hindering recognition of assay artefacts. It also suggests that multiple binding sites might exist on the FLuc enzyme to accommodate such a range of inhibitors.

Reported mechanisms of luciferase inhibition so far include substrate competition, multi-substrate adduct inhibition (MAI) and non-specific unfolding.[202–204] Most common amongst

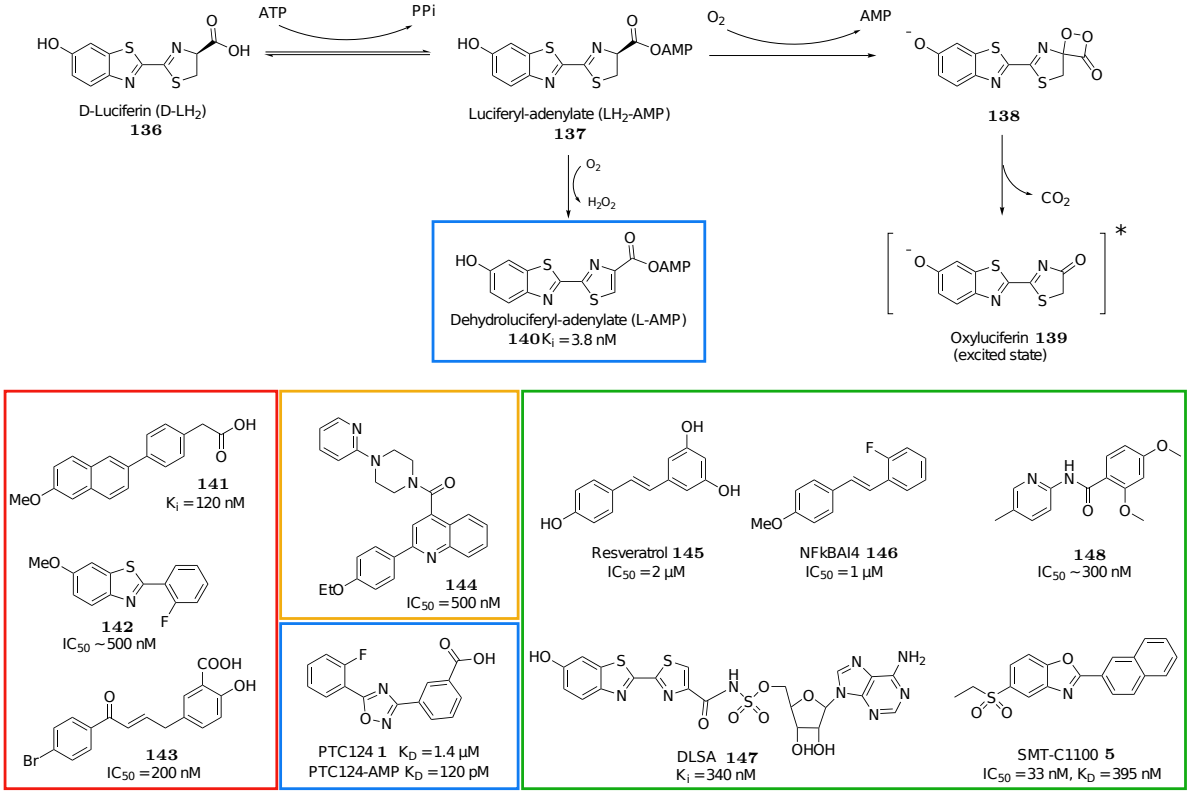


Figure 3.1: (**a**) Mechanism of FLuc bioluminescence reaction (**b**) Examples of FLuc inhibitors. Red box (left): luciferin competitors, orange box: ATP competitor, blue boxes: multisubstrate adduct inhibitors, green box (right): non-competitive inhibitors.

the reported mechanisms are inhibitors that compete with the substrate D-luciferin (FLuc K_M ~10 µM[205,206]). These include the analogues L-luciferin,[207] dehydroluciferin,[208] oxyluciferin,[209] as well as distinct chemotypes such as fatty acids,[210] and 2-phenylnaphthalenes **141**[211] (Figure 3.1b). Three inhibitor-bound FLuc complexes have been reported from soaking of apo crystals, each of which found the inhibitor (benzothiazole derivative **142**,[187] bromoform[212] and chalcone derivative **143**[213]) bound in the luciferin pocket. These findings are in keeping with the hydrophobic nature of the luciferin binding pocket as evidenced by the fluorescence induction of the hydrophobic 2-(p-toluidino)naphthalene-6-sulfonate (TNS) probe.[214] Computational analysis of FLuc inhibitors binding to this pocket has led to predictor tools designed to flag potential HTS false positives.[215,216]

FLuc inhibitors which compete with the ATP substrate have been less frequently observed but include a series of quinoline containing compounds **144** which were also found to compete with luciferin.[194] The lower frequency of identification of inhibitors of this mechanism is likely due to high saturating intracellular concentrations of ATP (~1 mM) which masks the evidence of this inhibition mode.

Multisubstrate adduct inhibitors (MAI) are single molecules comprising two or more substrates of the enzyme.[217,218] This approach often leads to highly potent and selective inhibitors,

with examples for FLuc including the naturally occuring L-AMP inhibitor and PTC124 **1** (Ataluren/Translarna), both of which contain carboxylic acids and form adenylate adducts after reaction with ATP and release of pyrophosphate.[219] As described in section 1.2.2, PTC124 has been granted conditional approval in Europe as a treatment for Duchenne muscular dystrophy in patients with nonsense mutations. It was originally discovered using a cellular FLuc reporter nonsense codon assay[220] but was subsequently found to inhibit FLuc with an IC_{50} of 7 nM, correlating to the activity in the original assay.[203]

Numerous examples of inhibitors of FLuc which are non-competitive with respect to luciferin or ATP or both exist in the literature. These include compounds such as resveratrol **145**[192] and NF-κB inhibitor **146**,[221] which continue to be used in luciferase reporter assays without control for their artefactual behaviour.[222] Other non-competitive inhibitors of different chemotypes have been identified with IC_{50} values as low as 0.26 nM.[194,195,201] However, the binding mode of these inhibitors has not yet been reported. The LH_2-AMP mimic DLSA **147** reported by Branchini *et al.* was found to be non-competitive with respect to the two substrates luciferin and ATP, but competitive with the LH_2-AMP intermediate.[223] This was rationalised by the significant conformational change involving a $\sim$140° rotation of the C terminal domain that occurs after the activating adenylation reaction forming LH_2-AMP, offering different binding environments within the same active site.[198]

Anaesthetics are another class of compounds that have been found to bind and inhibit FLuc with non-competitive kinetics, although via a non-specific mechanism.[224,225] Examples including chloroform, halothane and fluroxene have been shown to irreversibly thermally denature FLuc due to preferential binding to hydrophobic residues exposed in the unfolded protein.[204,226] The same behaviour was seen with detergent molecules such as dodecylamine.[210]

3.3 Investigation of ezutromid's mechanism of FLuc inhibition

Ezutromid was originally discovered using a FLuc reporter gene assay for utrophin upregulation. Although it was later shown to upregulate utrophin at both the mRNA and protein level in DMD mouse models and in human DMD patients' cells,[57,100,101] in dose-response analyses in the original reporter gene assay it displayed a bell-shaped curve, distinctive of FLuc inhibitors which stabilise FLuc and inhibit the bioluminescence reaction at high concentrations (Figure 1.5). This led us to suspect that an interaction between ezutromid and FLuc may be at least contributing to the output, in common with many of the derivatives described

in chapter 2. As anticipated, when subjected to an *in vitro* enzyme assay with recombinant FLuc, ezutromid was found to be a potent FLuc inhibitor, with an IC_{50} of 0.030 µM (Figure 3.2a).

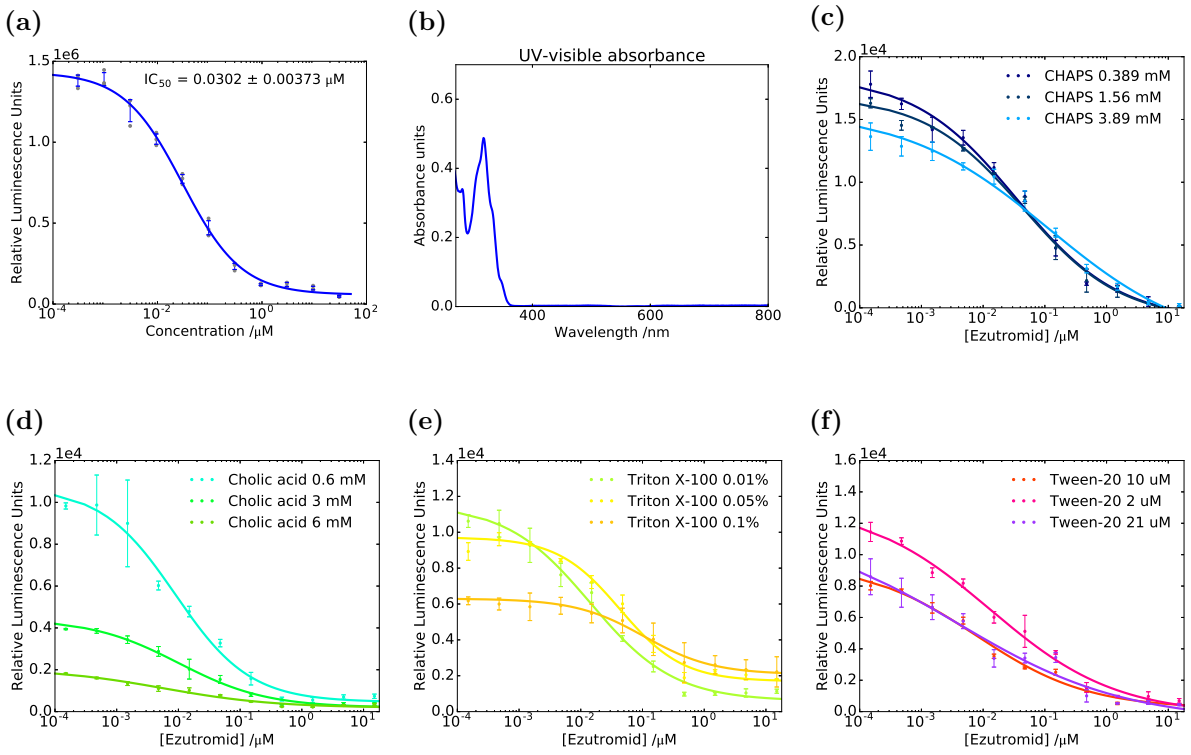


Figure 3.2: Ezutromid inhibits firefly luciferase. **(a)** Inhibition of FLuc luminescence by ezutromid; **(b)** UV-visible absorbance spectrum of 1 µM ezutromid in DMSO; **(c)-(f)** inhibition of FLuc by ezutromid is independent of the presence of a range of detergents CHAPS (zwitterionic), cholic acid (ionic), triton X-100 (non-ionic) and Tween-20 (non-ionic) at concentrations of approximately 10, 25 and 50% of their critical micelle concentrations (the concentration above which micelles form spontaneously).[227]

Since ezutromid does not absorb above 380 nm, it cannot quench the green tone FLuc luminescence (Figure 3.2b), so must be interfering with FLuc *via* a different mechanism. Conservation of ezutromid's inhibition of FLuc in the presence of a range of detergents (Figure 3.2c-3.2f) excluded small molecule aggregation as the mechanism of inhibition.[228]

An approximation of the K_D for the binding interaction between ezutromid and FLuc was now desired, and a fluorescence quenching assay was proposed for this purpose. Ezutromid's planar aromatic structure, in particular its naphthyl ring, suggested that its intrinsic fluorescence could be leveraged for this assay. The absorbance spectrum of ezutromid (40 µM in water with 5% DMSO to ensure compound dissolution) was measured to identify potential excitation wavelengths (Figure 3.3a). Lower compound concentrations (0.3 µM ezutromid, 1% DMSO in water) were used for the more sensitive fluorescence experiments. The peak absorbance wavelength (313 nm) was then used to generate an emission spectrum, revealing a peak emission at 390 nm (Figure 3.3b). The fluorescence excitation spectrum was measured

next, confirming a peak excitation of 310-313 nm (Figure 3.3c).

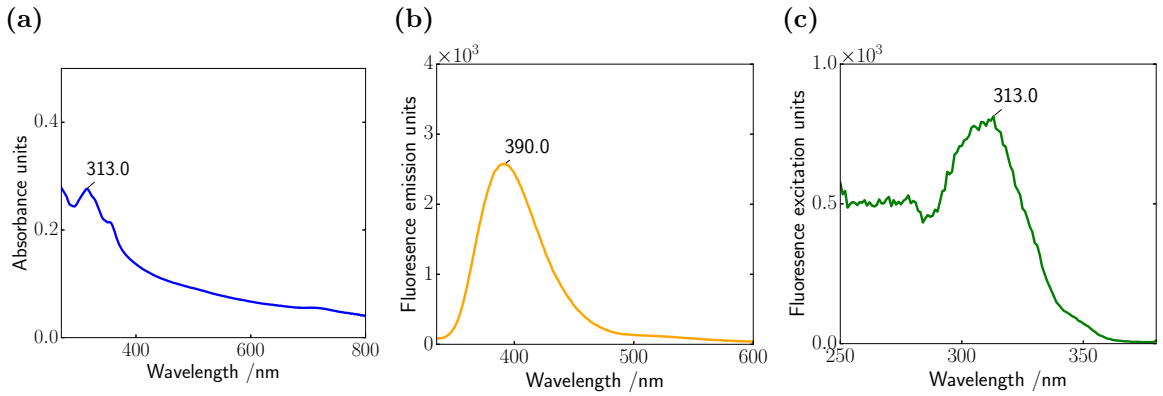


Figure 3.3: **(a)** UV-visible light absorbance spectrum for ezutromid (40 µM in water:DMSO, 19:1); **(b)** fluorescence emission spectrum of ezutromid (0.1 µM, 1% DMSO in Tris buffer, see experimental methods section 8.1), excitation wavelength of 313 nm; **(c)** fluorescence excitation spectrum of ezutromid (0.3 µM, 1% DMSO in water), emission wavelength of 390 nm.

Ezutromid's binding affinity for FLuc was then determined *via* quenching of the its fluorescence upon binding of the protein (Figure 3.4a-3.4b). Fitting of the fluorescence quenching with a four parameter logistic function gave an apparent K_D of 0.40 ± 0.15 µM and Hill coefficient of 0.98 ± 0.24 (Figure 3.4c), indicating a 1:1 stoichiometry of binding. This apparent FLuc K_D correlates well with the EC_{50} from the FLuc reporter gene assay (0.36 ± 0.22 µM, Figure 1.5), indicating that the original result was likely biased by FLuc interference.

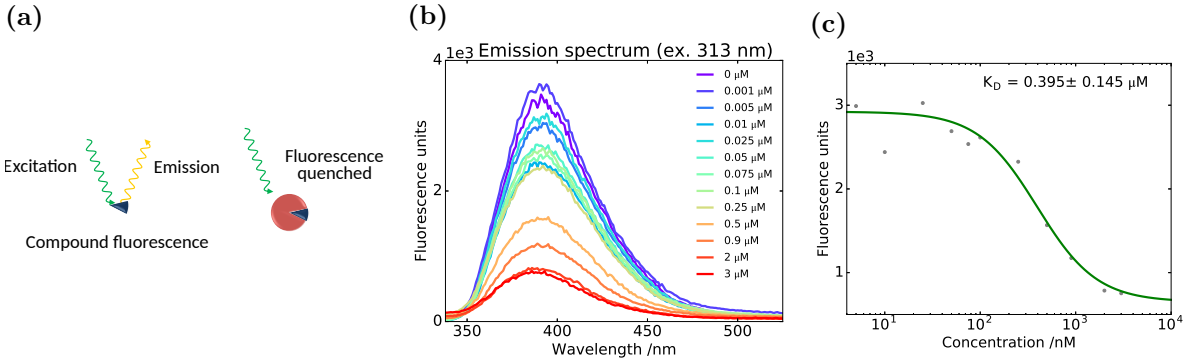


Figure 3.4: **(a)** Schematic of the fluorescence quenching binding assay using ezutromid; **(b)** incubation of ezutromid with FLuc results in quenching of the compound's intrinsic fluorescence (300 nM, 313 nm excitation, 390 nm emission); **(c)** fluorescence of ezutromid (300 nM, 313 nm excitation, 390 nm emission) measured after incubation (15 min) with firefly luciferase (5 − 3000 nM).

3.3.1 Docking studies of ezutromid and FLuc crystal structure

Ezutromid's benzoxazole core and planar, hydrophobic structure are characteristic of luciferin competitors. Docking studies using the benzothiazole inhibitor **142** bound FLuc structure (PDB: 4e5d[187]) in AutoDock Vina (default settings, described in the experimental methods)[229]

supported this hypothesis (Figure 3.5), with ezutromid found to overlay the inhibitor pose as anticipated in the estimated lowest energy binding conformation.

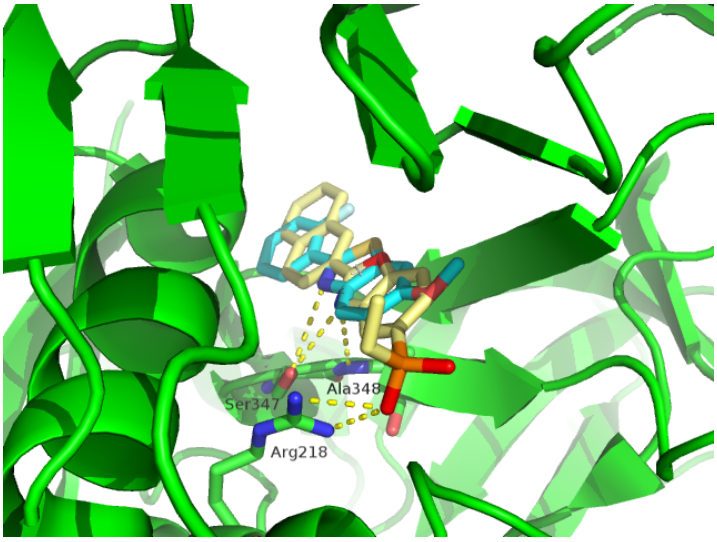


Figure 3.5: Docking studies predict that ezutromid (yellow) overlays the crystallised pose of benzothiazole inhibitor **142** (blue) in the FLuc luciferin pocket (green), PDB deposit: 4e5d.

3.3.2 Kinetic evaluation of ezutromid's FLuc inhibition

However, kinetic experiments varying first luciferin then ATP (with the other substrate held at saturating concentrations) indicated that substrate competition was not the mechanism by which ezutromid operates. Instead, fitting of initial rates to the Michaelis-Menten equation under rapid-equilibrium steady state kinetics led to determination of non-competitive inhibition with respect to luciferin (Figure 3.6a), and uncompetitive with respect to ATP (Figure 3.6b).

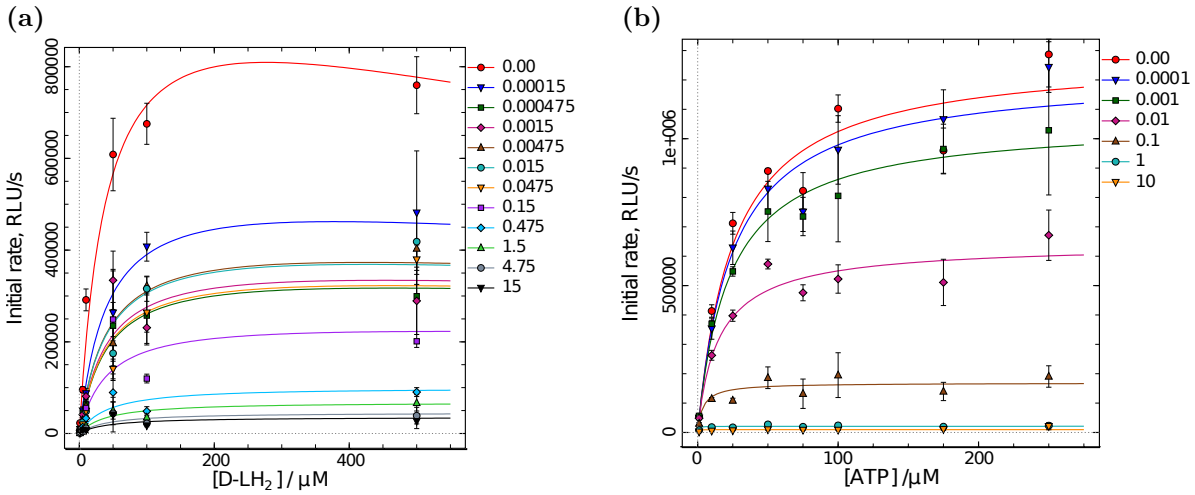


Figure 3.6: The dependence of the initial velocity on the concentration of luciferin **(a)** and ATP **(b)** at different concentrations of ezutromid (µM, coloured lines). Curves fitted using the Michaelis-Menten equation under rapid-equilibrium steady state kinetics to the data. Decrease in V_{max} at 500 µM luciferin (D-LH_2) is caused by substrate inhibition.[230]

These results were supported by Lineweaver-Burk analysis of the results, which showed FLuc's K_M for luciferin unchanged at 18.2 µM (close to the literature value of 16.9 µM[206]) as is expected for non-competitive inhibition (Figure 3.7a) and the K_M for ATP decreasing as seen in uncompetitive inhibition (Figure 3.7b). The K_I was calculated as 0.76 ± 0.43 µM from Dixon plot analysis of the results (Figure 3.7c-3.7d).

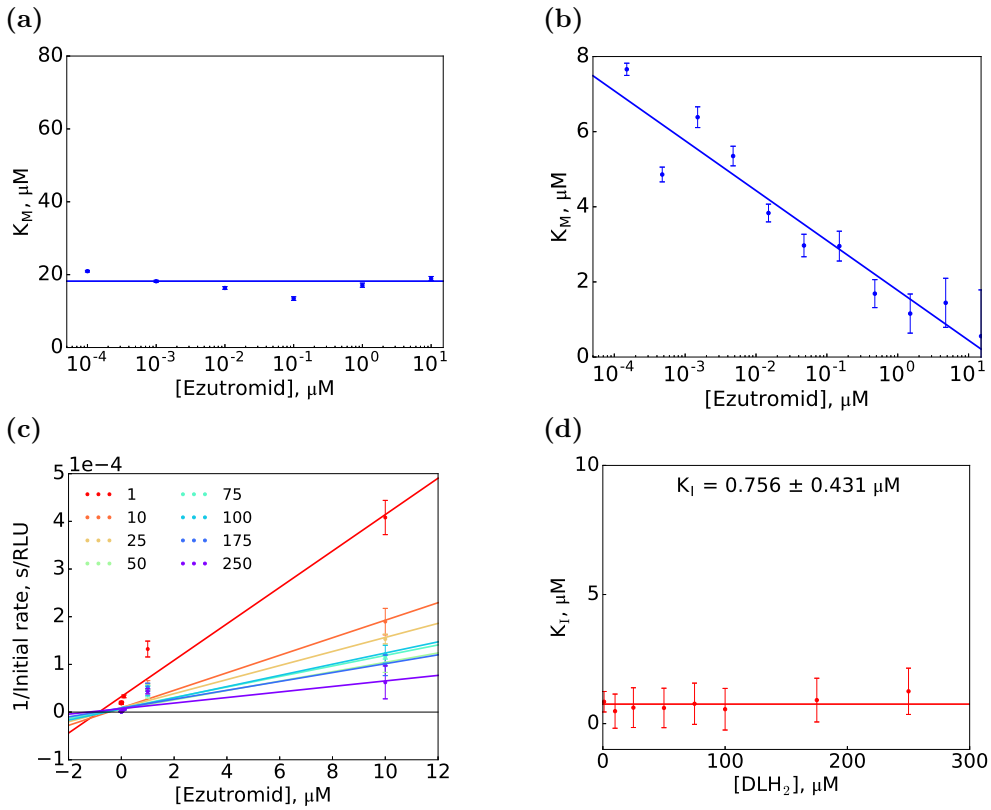
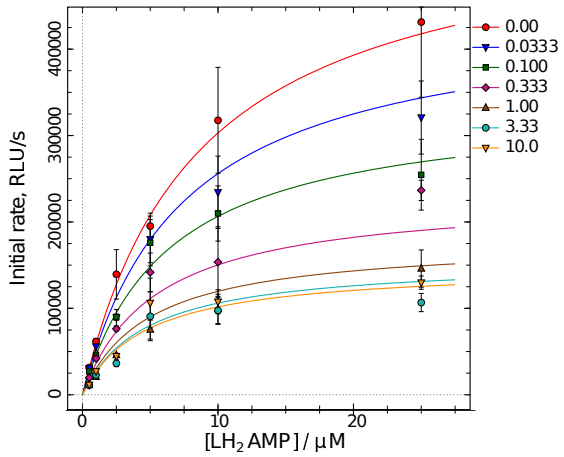


Figure 3.7: **(a)-(b)** FLuc K_M, determined by Lineweaver-Burk analysis of kinetic data, for luciferin and ATP respectively with increasing ezutromid concentration; **(c)** Dixon plot of the reciprocal of the initial velocity against increasing concentration of ezutromid, at different concentrations of luciferin substrate. **(d)** The K_I of ezutromid is estimated as 0.76 ± 0.43 µM, based on the x-axis intersection of lines in the Dixon plot.

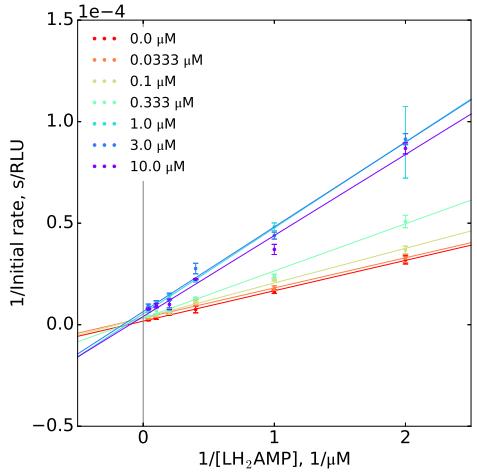
To assess whether ezutromid could be competing with LH_2-AMP for binding to FLuc's second catalytic conformation, further kinetic experiments were carried out, varying LH_2-AMP which was supplied as the sole substrate (initial velocity and Lineweaver-Burk plots in Figures 3.8a and 3.8b respectively).

DLSA was used as a positive control and was confirmed to compete with LH_2-AMP (data presented in Appendix C). Ezutromid was found not to compete with LH_2-AMP, but instead was shown to display partial mixed inhibition, meaning it binds to both the FLuc and the FLuc:LH_2-AMP complex, and also does not fully inhibit light production. Other compounds that have been found to bind *via* mixed type inhibition have also been reported. A class of

(a)



(b)



(c)

Ezutromid	DLSA	Resveratrol	NF-κBAI4	**148**
K_M increased	K_M increased	K_M unchanged	K_M unchanged	K_M decreased
V_{max} decreased	V_{max} unchanged	V_{max} decreased	V_{max} decreased	V_{max} decreased
E.S.I ⇌ E.I + P	E.S.I ⇏ E.I + P	E.S.I ⇏ E.I + P	E.S.I ⇌ E.I + P	E.S.I ⇌ E.I + P
Mixed partial	Competitive	Non-competitive	Partial non-competitive	Mixed partial

Figure 3.8: Mechanistic evaluation of known inhibitors of firefly luciferase with respect to the enzyme intermediate luciferyl adenylate. **(a)** The dependence of the initial velocity on the concentration of luciferyl adenylate (LH_2-AMP) at different concentrations of ezutromid. Curves were fitted using the Michaelis-Menten equation under rapid-equilibrium steady state kinetics to the data using Dynafit. Inhibition model selection was achieved with Akaike and Bayesian Information Criteria; **(b)** The same data presented in double reciprocal Lineweaver-Burk plots; **(c)** Table of the effects on FLuc kinetic parameters K_M and V_{max} induced by DLSA, ezutromid, resveratrol, NF-κBAI4 and **148**, and whether full or partial inhibition is observed. K_M and V_{max} were calculated from slope and the reciprocal of the intercept of Lineweaver-Burk plots (data presented in Appendix C). Mechanism of inhibition with respect to LH_2-AMP (red box) was determined after analysis using both Dynafit and Lineweaver-Burk methods.

aryl triazoles were found to have mixed non-competitive inhibition with respect to substrate analogue aminoluciferin,[211] while some pyrrolo[2,3-*d*]pyrimidines had mixed competitive inhibition with respect to luciferin.[231] To determine whether ezutromid's mode of inhibition is similar to that of other reported non-competitive inhibitors, resveratrol **145**, NF-κBAI4 **146** and compound **148**[232] were included in further experiments (data presented in Appendix C). Resveratrol and structurally related **146** were found to have non-competitive and partial non-competitive inhibition respectively regarding LH_2-AMP, whilst **148** had mixed partial inhibition similar to ezutromid (table, Figure 3.8c). These findings indicate that these non-competitive inhibitors are binding *via* a previously undescribed mechanism. This indicates that FLuc has multiple inhibitor binding sites and conformations beyond that which is currently understood.

3.3.3 Towards determination of ezutromid's FLuc binding mode

Co-crystallisation

X-ray crystallography is a powerful method for determining protein-ligand binding modes. Crystal soaking and co-crystallisation of FLuc with ezutromid alone or in the presence of ATP, luciferin or DLSA was attempted in collaboration with Frank von Delft's group at the SGC to determine the exact location of the ezutromid binding site. However these crystallisation attempts yielded only needle apo structures matching those previously reported in the literature (Figure 3.9a),[212,213] or alternatively crystals of ezutromid (Figures 3.9b-3.9f). Soaking of apo crystals with compound also failed to yield protein-ligand complexes, which suggests that ezutromid may bind to a higher energy conformation of FLuc which we have been unable to crystallise.

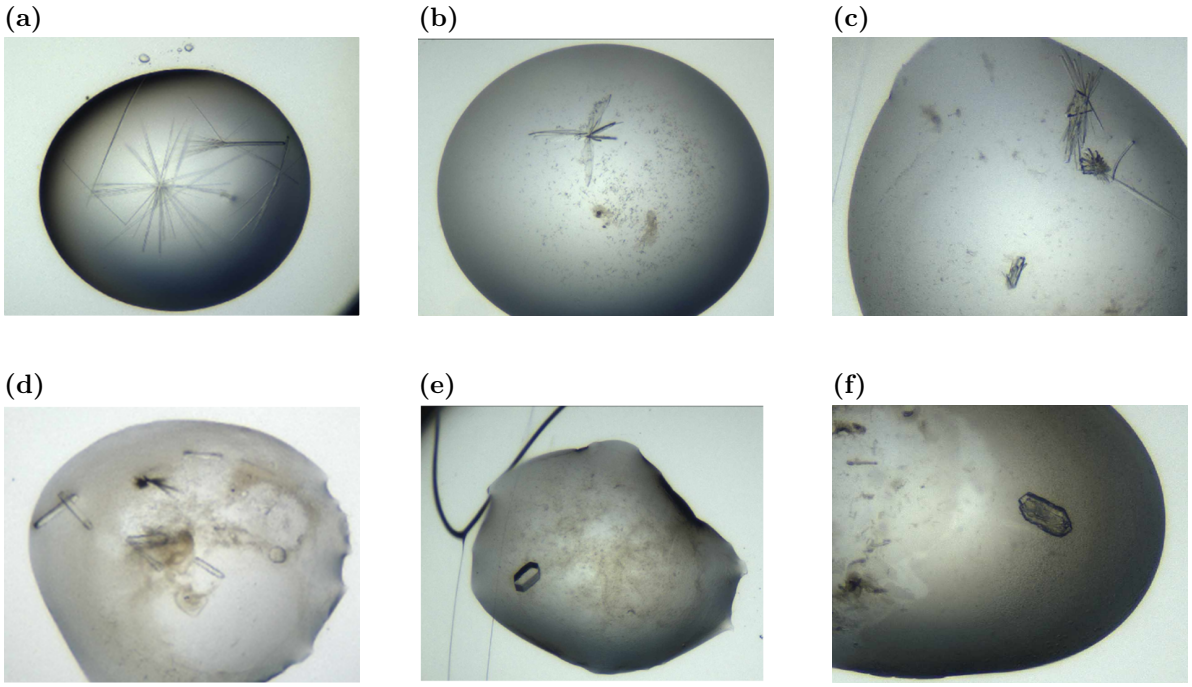
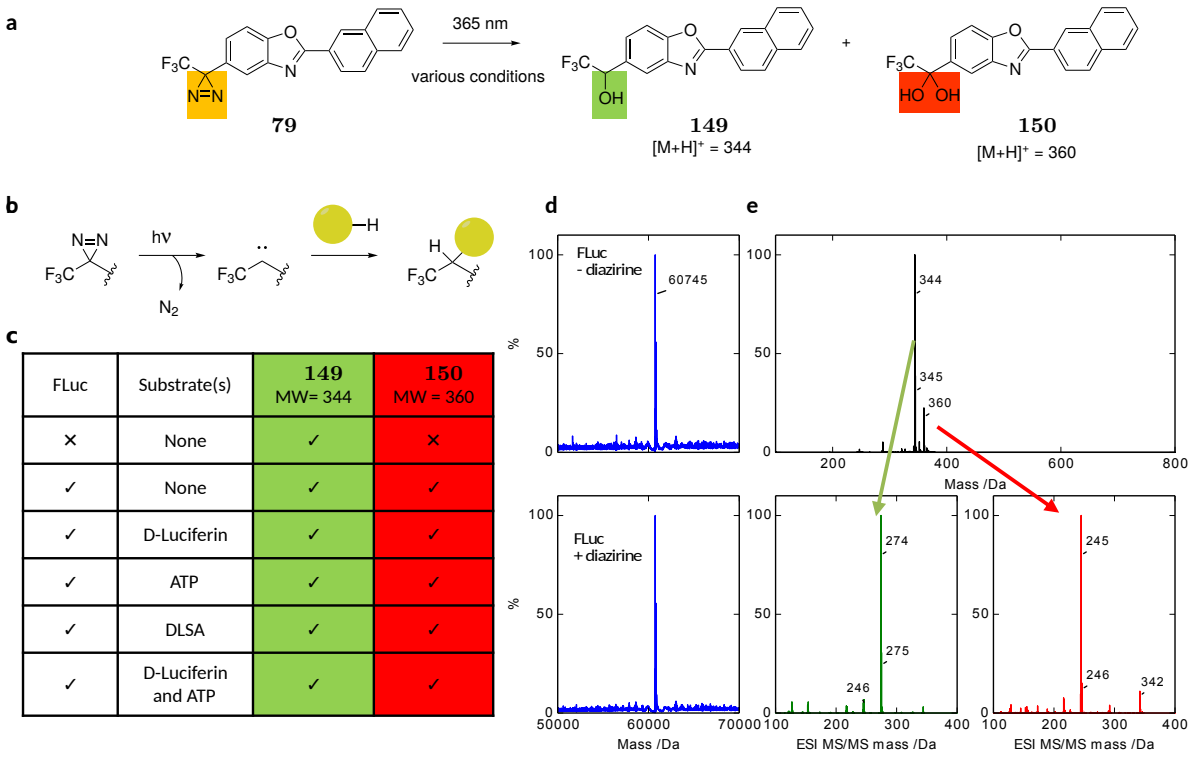


Figure 3.9: Photographs of crystals grown from solutions of FLuc and ezutromid, each were mounted and X-ray data collected at the Diamond Light Source, Harwell, UK (data collected by collaborators in the von Delft group). X-ray data was processed using PHENIX[233] and structures solved by molecular replacement with Phaser[234] using structure 1lci[235] as a model. **(a)** Needle crystals of apo firefly luciferase; **(b)**-**(f)** crystals of small molecules.

Photoaffinity approach

In the absence of X-ray crystal data, a photoaffinity strategy was employed to elucidate the binding mode of ezutromid. This approach involves incorporation of a photoaffinity group into a probe compound then exploiting its affinity for its protein target (Figure 3.10b). Once

bound, UV irradiation to activate the photoaffinity group results in carbene generation, then covalent binding to nearby residues.[149] The site of incorporation can be established by mass spectrometry to reveal the molecule's binding site. Trifluoromethylphenyl diazirines are well-characterised and have been successfully used in the literature for binding site identification studies.[154,236,237] Photoaffinity probe **79**, synthesised by Jess Reynolds, was found to retain FLuc inhibition behaviour, although with a reduced affinity, $IC_{50} = 0.595 \pm 0.112$ µM (graph presented in Appendix B).



FLuc	Substrate(s)	**149** MW= 344	**150** MW = 360
×	None	✓	×
✓	None	✓	✓
✓	D-Luciferin	✓	✓
✓	ATP	✓	✓
✓	DLSA	✓	✓
✓	D-Luciferin and ATP	✓	✓

Figure 3.10: Irradiation studies of ezutromid-related photoaffinity probe with firefly luciferase. **(a)** Scheme of products generated from irradiation of diazirine probe **79** under conditions listed in table **(c)**; **(b)** Schematic of photo-induced covalent crosslinking of trifluorophenyl diazirines into -CH or -XH where X is a heteroatom; **(d)** LC-MS of FLuc with (bottom) and without (top) diazirine probe **79**; **(e)** LC-MS of products formed from diazirine irradiation, bottom left: MS/MS spectrum of peak 344, bottom right: MS/MS spectrum of peak 360.

However, incubation of the probe with FLuc, followed by UV irradiation (365 nm) and LC-MS of the reaction mixture resulted in no observation of probe incorporation into the enzyme (Figure 3.10d, spectrum presented in Appendix D). Instead, two species were generated from the probe, one whose mass corresponds to the addition of H_2O_2 (assigned as hydrate **150**), the other corresponding to the addition of water, assigned as **149** (Figure 3.10a). This finding was replicated with the addition of substrates luciferin, ATP and DLSA. However, interestingly, when the probe was irradiated in the absence of FLuc (Figure 3.10c), only the water adduct was observed, indicating the requirement of the enzyme for the oxidation of the probe (chromatograms for the peak corresponding to the oxidised probe presented in the Appendix D).

ESI MS/MS fingerprint analysis of the putative water **149** and H_2O_2 **150** adducts generated in the irradiation experiments corresponded with the spectra generated from authentic samples of the compounds (Figure 3.10e). These authentic samples were intermediates towards diazirine probe **79** previously synthesised by Jess Reynolds (synthetic route as reported in Scheme 2.6 for regioisomer **81**). The authentic sample of hydrate **150** derives from the corresponding trifluoromethyl ketone, since ketones with strongly electron-withdrawing groups exist predominantly as hydrates in aqueous solutions.[238]

Reaction of molecular oxygen with TPDs has been reported in the literature[239] and FLuc has oxygenase activity (Figure 3.1a), indeed an oxygen/pantetheine tunnel to the active site that has been postulated by Branchini *et al.*.[240,241] Furthermore, the approach of oxygen to the active site has been modelled by molecular dynamics and mechanics, finding that oxygen can move inside FLuc without any energetic barrier.[242] The results of this experiment are not conclusive but suggest that probe **79** might bind to an oxygen binding FLuc conformation.

3.4 Summary

Ezutromid has been identified as an FLuc inhibitor, requiring cautious interpretation of results from its use in FLuc reporter gene assays. Ezutromid is a non-competitive inhibitor with respect to luciferin and uncompetitive with respect to ATP. Alongside three other literature non-competitive inhibitors, ezutromid was found to have previously undescribed non-competitive and mixed inhibition mechanisms with respect to the intermediate LH_2-AMP. Photoirradiation studies to identify the binding site of a ezutromid probe revealed the formation of an FLuc dependent oxygenated species, indicating binding of the probe in the proximity of molecular oxygen. However, the binding site of these non-competitive inhibitors is still uncharacterised. Despite investigations into FLuc's mechanism stretching back 70 years, it seems that much of the dynamic conformational changes involved is still not fully understood. Full characterisation of ezutromid's binding mode in FLuc was not achieved, and ezutromid's non-competitive mode of inhibition means that probe design circumventing FLuc binding is difficult to rationalise. In conclusion, the biological evaluation of ezutromid-related analogues in FLuc reporter gene assays requires experimental counter-screening for FLuc inhibition. Given the significant FLuc inhibition exhibited by 5 of the 8 probes prepared in chapter 2, their evaluation in secondary assay(s) to determine utrophin modulation is essential.

Chapter 4

Biological Evaluation of Probes for Target ID Experiments

4.1 Aims

Chapter 2 discussed the design and synthesis of biotinylated and dual-tagged photoaffinity probes for *in vitro* and *in situ* target identification studies. Cell-permeable truncated versions of the biotinylated probes were also generated for assessment of bioactivity in cellular assays (**65**, **52**, **74** and **73**, Figure 4.1). However, evaluation of probe bioactivity using the primary utrophin-FLuc reporter gene assay was impaired by FLuc-related artefacts. This chapter describes studies to unambiguously ascertain the bioactivity of the prepared active and inactive probes (Figure 4.1), and the competitor control **151**.

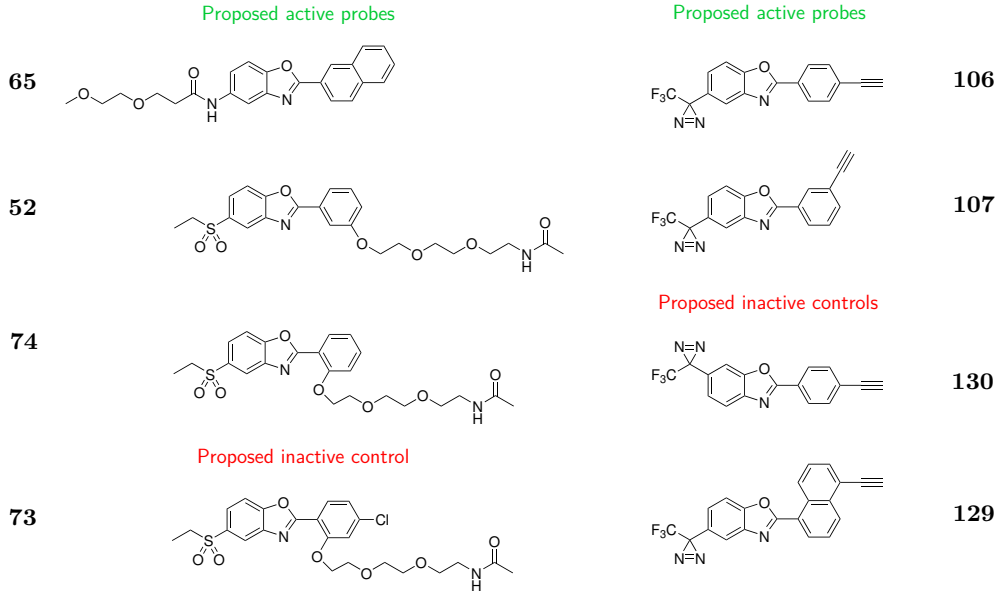


Figure 4.1: Left: structures of the proposed active and inactive PEGylated ezutromid analogues (truncates of the biotinylated probes); right: structures of the proposed active and inactive dual-tagged photoaffinity probes.

4.2 Introduction: assays for utrophin upregulation

There are significant challenges in developing robust and high-throughput assays to quantify utrophin expression in DMD muscle. Firstly, both muscle and the dystrophic phenotype are dynamic systems that are difficult to reproduce in *in vitro* models. Secondly, utrophin is a scaffold protein without enzyme or signaling activity that could be quantified using biochemical or cell-based assays. Thirdly and most importantly, utrophin is a regeneration-associated protein and is increased 2–5-fold over wild type as part of the repair process, following rounds of muscle necrosis in DMD boys and animal models.[243,244] The level of regeneration and utrophin expression in DMD boys and animal models varies with age, individual and muscle type, providing a high and inconsistent background against which small molecule-induced utrophin upregulation needs to be measured.[57] Adding to the complexity, high levels of regeneration are correlated with increased disease severity in DMD boys.[245] Indeed, utrophin related to regeneration is heterogenously expressed in muscle and its increase is associated with a reduction of muscle function.[105] It is the uniform increase of utrophin at the sarcolemma, seen in transgenic *mdx* mice, that reduces the regeneration mechanism and restores muscle function, and is the desired outcome of small molecule utrophin modulator treatment.

The development of assays for quantifying utrophin upregulation has been extensively pursued by our collaborators at DPAG, Evotec and Summit Therapeutics. Assays to examine utrophin promoter activation were developed from the original transfected 8.4 kb utrophin A promoter-FLuc fragment (Figure 1.5) to a FLuc knock-in murine cell line, where one allele of utrophin was replaced with the reporter to allow control of gene expression in the endogenous genomic context. Replacement of FLuc with nanoLuc and mCherry reporters for use in human DMD cell lines to circumvent FLuc-related artefacts was ongoing during this project.

Real time quantitative PCR was used for mRNA quantification, while protein quantification was attempted by western blot, Meso Scale Discovery® (MSD), Homogeneous Time Resolved Fluorescence (HTRF), and High Content Imaging (HCI) assays. However, attempts to increase the throughput of these protein quantification assays *via* miniaturisation have not been successful, with the failure of positive controls to behave as anticipated. Our *in vitro* models of choice were *mdx* and human DMD immortalised myoblasts which offer improved assay flexibility and throughput compared to differentiated myotubes.[14,246]

4.3 Evaluation of the ability of ezutromid-based probes to upregulate utrophin

Additional challenges lie in the biological evaluation of ezutromid-based probes due to the modest activity (1.5-2 fold change) of the parent molecule, which is compounded by the typical reduction in activity following introduction of purification handles or photoaffinity labels. Furthermore, ezutromid's limited aqueous solubility (3 µM) is shared by its photoaffinity-labelled analogues*, restricting possible dosage concentrations and potentially impairing maximal activity calculations required to generate IC_{50} and EC_{50} values.

The primary utrophin-FLuc reporter gene assay protocol involves preparation of a compound dilution plate serially diluted from a maximum compound concentration of 200 µM in 20% DMSO in water, prior to 20× dilution when dosing seeded cells in media. Although ezutromid was discovered in this screen, this protocol results in formation of needle crystals of ezutromid in the high concentration compound plate, which are transferred during cell treatment (Figure 4.2). Crystallisation rather than precipitation gives the appearance of dissolution by eye, but renders meaningful determination of activity impossible.

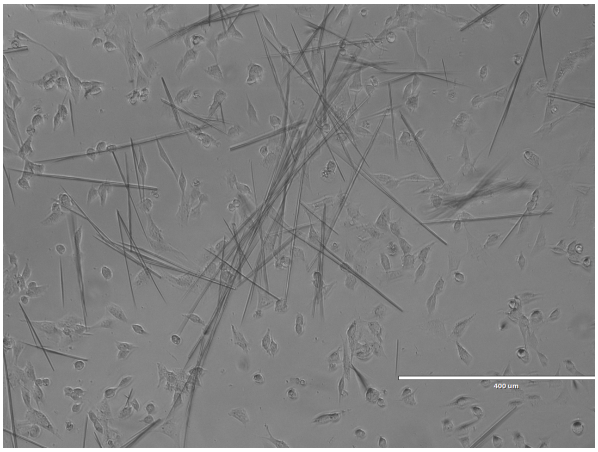


Figure 4.2: Treatment of H2K *mdx* utrnA-luc cells with ezutromid according to the utrophin-FLuc reporter gene assay results in the formation of needle crystals, scale bar: 400 µm.

To accommodate the poor solubility of ezutromid and its dual-tagged photoaffinity analogues during biological testing, compound concentrations never exceeded 3 µM, DMSO concentrations up to 1% were employed and serial dilution was not performed. Experiments to examine the activity of the PEGylated and photoaffinity probes, and phosphinate competitor compounds are described below.

*Probes **106** and **107** are difficult to dissolve in DMSO and were considered to be insoluble in water (semi-thermodynamic solubility < 2 µM) by Cyprotex Discovery Ltd., Alderley Park, Cheshire, SK10 4TG (UK)

4.3.1 Evaluation of probe bioactivity by RT-qPCR

First, the PEGylated and dual-tagged photoaffinity probes were assessed for their ability to upregulate utrophin mRNA in murine *mdx* myoblasts after 24 h treatment (30 nM) using real time quantitative PCR (RT-qPCR). Gratifyingly, PEGylated probes **65** and **74** retained the ability of ezutromid to increase utrophin mRNA levels > 2 fold (Figure 4.3a). *Meta*-substituted probe **52** gave a weaker ∼1.5 fold utrophin increase, while probe **73** was confirmed to be inactive. Further, treatment with photoaffinity probes designed to be active **106** and **107**, and also alkyne-bearing phosphinate probe **118** led to ∼1.5 fold increases in utrophin mRNA expression. However, while 1-naphthyl probe **129** gave a response indistinguishable from the DMSO vehicle, as anticipated for the negative control, 6-substituted **130** unexpectedly displayed activity similar to 5-substituted analogues **106** and **107**, throwing its assignment as an inactive control into doubt. Statistical significance was not calculated as these results were derived from three seeded replicates of cells from the same biological population (same passage number of immortalised H2K *mdx* cell line).

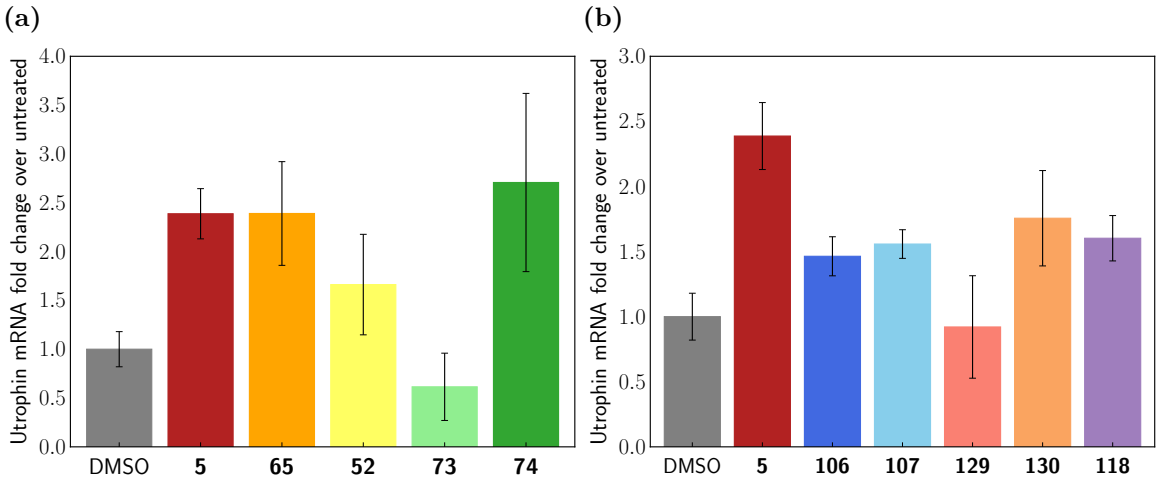


Figure 4.3: **(a)** Probes **65** and **74** retain the ability of ezutromid **5** to increase utrophin mRNA levels after 24 h of compound treatment (30 nM). *Meta*-substituted probe **52** gives a weaker ∼1.5 fold utrophin increase, while probe **73** is inactive; **(b)** probes **106**, **107**, **130** and **118** give ∼1.5 fold increase in utrophin mRNA expression after 24 h of compound treatment (30 nM). Probe **129** gives a response indistinguishable from the DMSO vehicle. Utrophin mRNA expression levels were normalized to S13 and relative to DMSO as the mean ± SD (n = 3, technical replicates).

4.3.2 Evaluation of probe bioactivity by western blot

Next, fold change in utrophin protein levels post probe treatment compared to the vehicle control was evaluated by western blot analysis. Due to the low throughput of this assay, only the dual-tagged photoaffinity probes, which had less conclusive responses in the qPCR assay

were assessed. Compound **130**, which was found to increase utrophin mRNA rather than acting as a inactive control, was not tested. Remaining compounds **106**, **107** and **129** were tested in three concentrations (0.03, 0.3, 3 µM, 24 h incubation) to allow determination of any dose-response. Each experiment (human DMD or mouse (*mdx*) myoblasts) was performed with three replicates, and each experiment carried out three times in total, using different frozen cell stocks. One experiment using the human DMD cell line was carried out by Lee Moir (Davies group, DPAG) and the results combined in the data presented.

Mouse H2K *mdx* experiments

Both probes **106** and **107** were found to upregulate utrophin protein in *mdx* myoblasts after 24 h, with *meta*-substituted **107** demonstrating a response comparable to the parent compound ezutromid (Figure 4.4). At 3 µM, probes **106** and **107** were found to increase utrophin protein ~1.5 fold ($p < 0.05$), similar to other reported utrophin modulators (section 1.3.2) and the endogenous utrophin modulator heregulin. Levels of utrophin protein post treatment with 1-naphthyl probe **129** were indistinguishable from the vehicle control at all concentrations.

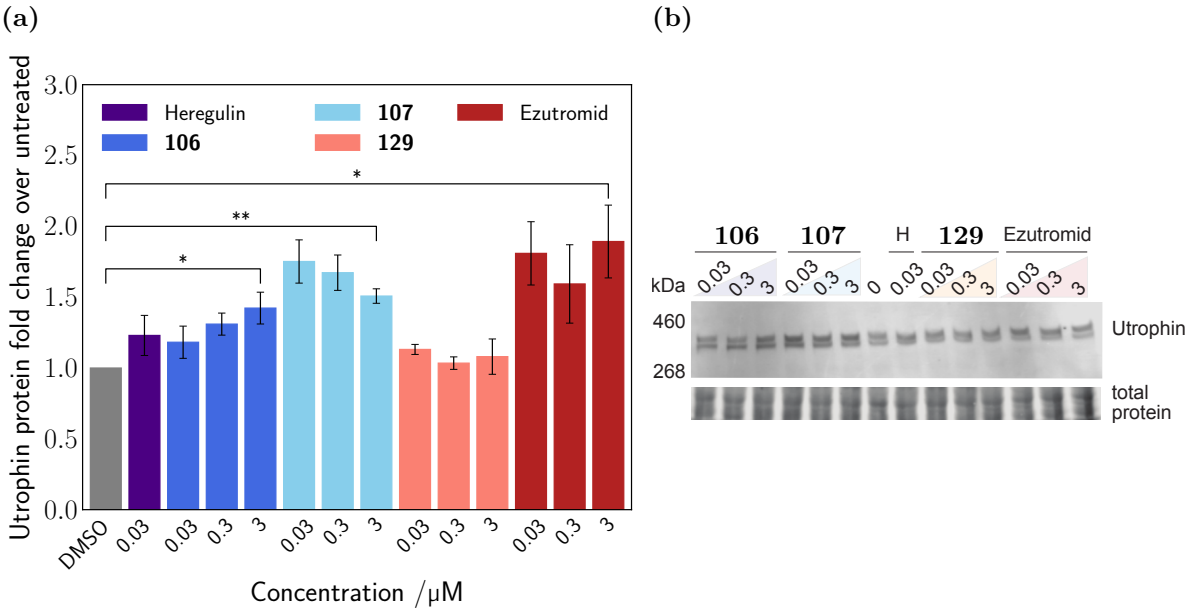


Figure 4.4: **(a)** Probes **106** and **107** increase utrophin protein levels comparable to heregulin and ezutromid after 24 h of compound treatment in H2K *mdx* cells. Inactive probe **129** does not increase utrophin. Bars represent means of utrophin levels normalised to total protein and relative to DMSO, error bars are standard error of the mean, n = 3. * $p < 0.05$, ** $p < 0.005$. **(b)** Representative utrophin western blot from probe treatment of mouse H2K *mdx* myoblasts. H; heregulin. Uncropped blots are presented in Appendix E.

Human DMD myoblast experiments

The activity of probes **106** and **107** was further profiled in human DMD cells. *Meta*-substituted probe **107** increased utrophin protein levels after 24 h to a similar extent as ezutromid (Figure 4.5), with both leading to a ~1.8 fold increase with 3 µM treatment ($p < 0.05$). However, no significant increase in utrophin protein was observed for *para*-substituted **106**. The small compound effect (< 2 fold change) coupled with high variability of the utrophin expression background made determination of significant effects by western blot challenging. A more robust and higher throughput secondary assay would be required for further assessment of probe bioactivity.

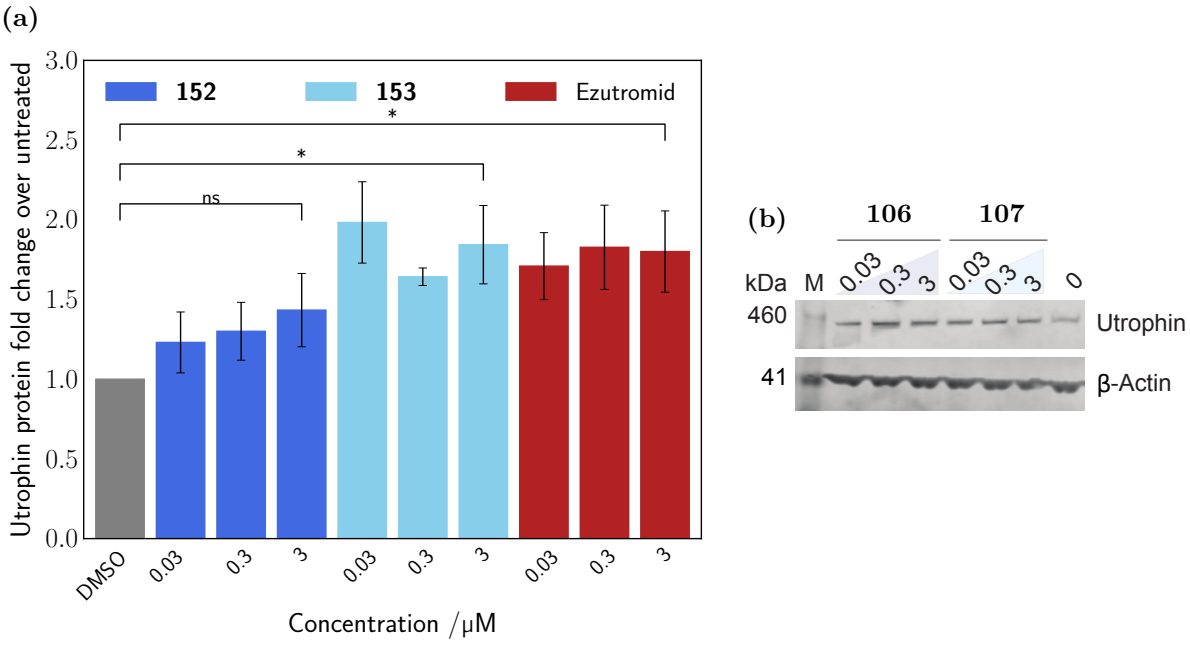


Figure 4.5: **(a)** Probe **107** increases utrophin protein levels comparable to ezutromid after 24 h of compound treatment in human DMD cells. Probe **106** led to no significant increase. Bars represent means of utrophin levels normalised to total protein and relative to DMSO, error bars are standard error of the mean, n = 3. * $p < 0.05$, ns = not significant. One replicate experiment was performed by Lee Moir (Davies group, DPAG); **(b)** Representative utrophin western blot from probe treatment of human DMD myoblasts. M; protein standard (in kDa). Uncropped blots are presented in Appendix E.

4.3.3 Selection of an appropriate competitor control

The alkyne-bearing phosphinate probe **118** had been synthesised to be used as a competitor control for AfBPP experiments, and shown encouraging signs of activity in the preliminary utrophin qPCR assay. However, due to the challenges with low throughput and small assay window in confirming utrophin modulating activity, second generation compound SMT022357 **6** with *in vivo* activity established by collaborators in the Davies group[102] was advanced as a

potential competitor instead. The toxicity of **6** at the high dosage concentrations appropriate for a competitor control was assessed by its effect on cell viability after 2 h incubation (Figure 4.6a). SMT022357 was found to be cytotoxic above 100 µM; indeed it was safety issues that had halted the progression of this compound during maximum tolerated dose determination. However, this hepatoxicity was found to be alleviated after substitution of the *meta*-CF_3 for a *para*-F (Cyprotex Discovery Ltd., Alderley Park, Cheshire, SK10 4TG, UK). Further, the (+)-isomer, SMT022332 **151** (Figure 4.6c) was found to have greater *in vivo* efficacy than its (−)-isomer counterpart (Davies group, DPAG). In this work, SMT022332 was confirmed to be well-tolerated at all concentrations tested (Figure 4.6a), and dose-dependently increased utrophin protein expression in human DMD myoblasts (Figure 4.6b-c).

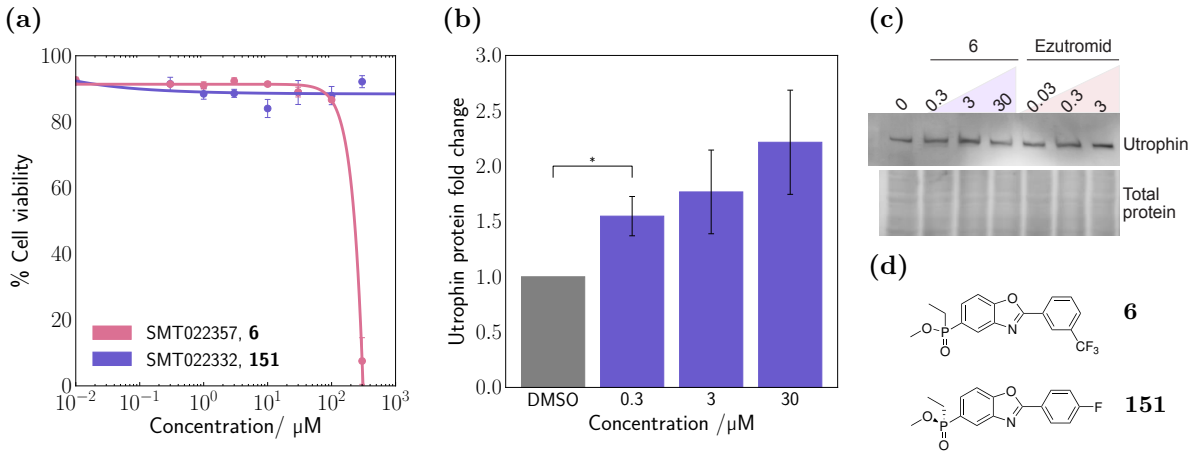


Figure 4.6: **(a)** Second generation compound SMT022357 is cytotoxic above 100 µM, while analogue SMT022332 **151** is well-tolerated at all concentrations tested. SMT022332 dose-dependently increases utrophin protein levels after 24 h of compound treatment in human DMD myoblasts, determined by western blot. **(b)** Bars represent means of utrophin expression normalised to total protein and relative to DMSO, error bars are standard error of the mean, n = 4. One replicate experiment was performed by Lee Moir (Davies group, DPAG). Representative blot shown in **(c)**, uncropped blot presented in appendix E; **(d)** Structures of SMT022357 **6** and SMT022332 **151**.

4.4 Physiological evaluation of bioactivity of ezutromid-based probes

The fusion of myoblasts into mature myotubes is required during muscle development and repair, and is dysregulated in dystrophic muscles.[247] Increases in rate of fusion of treated human DMD myoblasts into mature myotubes (Figure 4.7a) indicates improved myogenesis.[248] The effects of compound treatment were examined in a high content imaging (HCI) assay developed by collaborators at Evotec International GmbH (Hamburg, Germany). Immortalised DMD myoblasts were treated with compound or vehicle (0.1% DMSO) during five days of differentiation, then fixed and stained for utrophin, myosin heavy chain (a myotube marker)

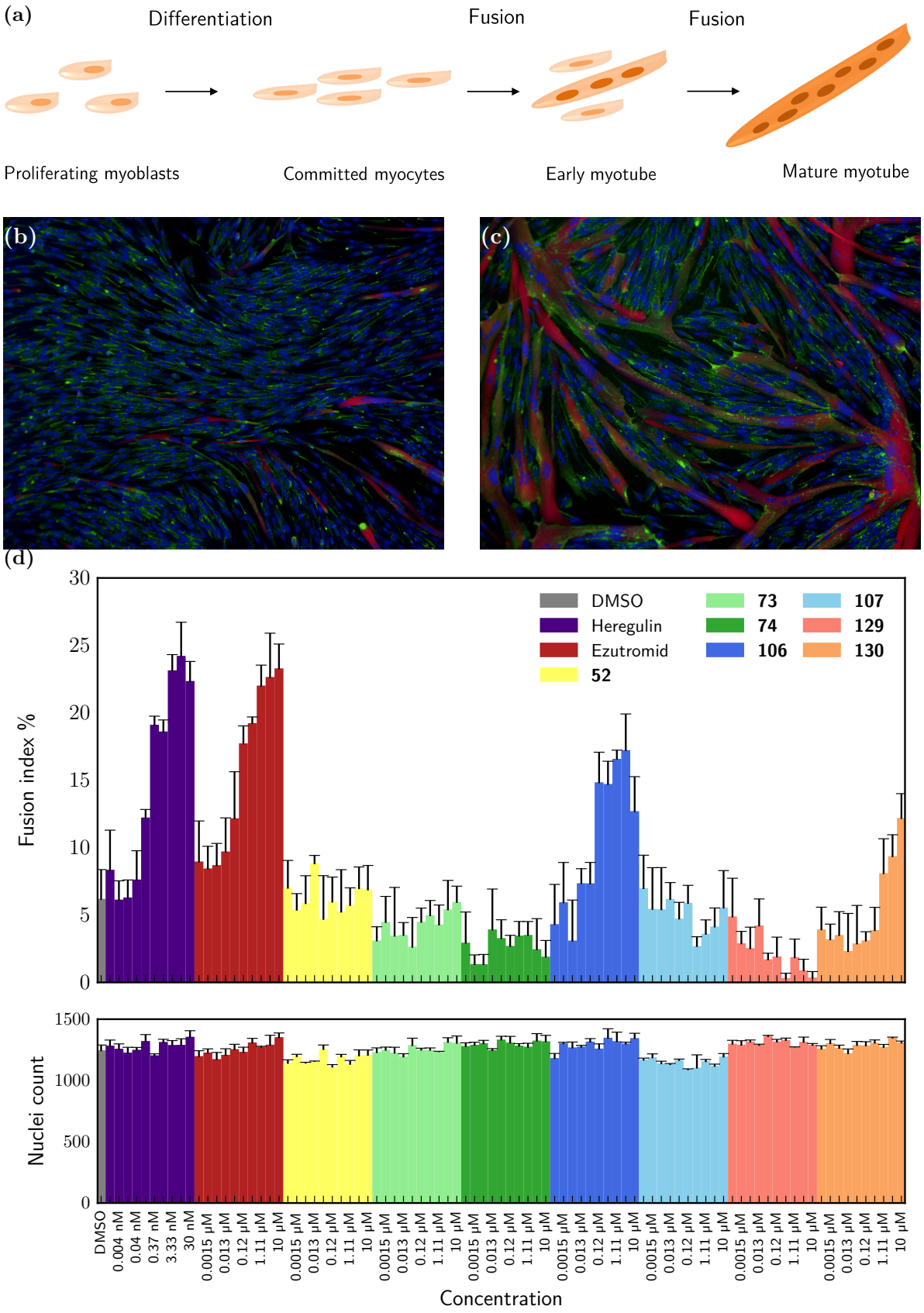


Figure 4.7: Utrophin modulators such as heregulin, ezutromid and probes **106** and **130** increase fusion of DMD myoblasts into mature myotubes. **(a)** Schematic of the process of fusion of proliferating myoblasts into myotubes. Immortalised DMD myoblasts were treated with compound or vehicle (0.1% DMSO) during five days of differentiation, then fixed and stained for utrophin (green), myosin heavy chain (red) and DAPI (blue); **(b)** vehicle control, **(c)** 10 µM ezutromid. **(d)** Dose-dependent increase in fusion of myoblasts to mature myotubes observed after 5 days of heregulin and ezutromid treatment is also observed after treatment with *para*-substituted probes **106** and **130**, but not *meta*-alkyne probe **107**, inactive 1-naphthyl probe **129** or the biotinylated probes. Fusion index is shown (top) as mean ± SD and nuclei count (bottom) mean ± SD, each n = 4.

and DAPI. Images were taken with a Perkin Elmer Operetta high-content imaging system with a 10 × objective; representative images for DMSO and ezutromid (10 µM) are shown in Figure 4.7b-c. Accurate automated determination of utrophin protein fold change which gave expected results for the positive controls was not established in this high content imaging assay format before the end of the collaboration and is therefore not presented here. However, dose-dependent increase in fusion of myoblasts to mature myotubes after 5 days was observed after heregulin and ezutromid treatment (Figure 4.7d). Interestingly, this behaviour was also observed after treatment with *para*-substituted probes **106** and **130**, but not *meta*-alkyne probe **107**, inactive 1-naphthyl probe **129** or the biotinylated probes.

4.5 Summary

The results of this chapter are summarised in Table 4.1. The cell-permeable truncated analogues of the four biotinylated probes and the four dual-tagged photoaffinity probes synthesised in chapter 2 were tested for their ability to upregulate utrophin mRNA levels by RT-qPCR. Encouragingly, three of the truncated biotinylated probes were found to be active, with two retaining the same activity as ezutromid. Further, the probe designed to serve as a negative control was confirmed to be inactive. Although none of these truncated analogues were found to replicate ezutromid's ability to increase fusion of myoblasts into myotubes, all four biotinylated probes were taken forward for target identification studies.

The two dual-tagged photoaffinity probes designed to be active (**106** and **107**) were also found to upregulate utrophin mRNA production, although to a lesser extent than ezutromid ($\sim$1.5 vs. $>$ 2 fold). One of the intended negative control compounds (1-naphthyl probe **129**) was confirmed not to increase utrophin, while the other (6-substituted **130**) was found to be as active as **106** and **107**. The three successful probes were further assessed for their effect on utrophin protein levels by western blot analysis. Pleasingly, probes **106** and **107** were also found to increase utrophin protein $\sim$1.5 fold in both mouse and human DMD cells, while probe **129** was confirmed to be inactive. However, only **106** retained ezutromid's ability to increase the rate of myoblast fusion.

Probe **106**-induced increases in utrophin expression, while small, were consistently observed in orthogonal assay formats. Therefore, based on the strength of this combined activity data, **106** was deemed the most similar to ezutromid's phenotype and selected to be the active *in situ* AfBPP probe in target identification studies. The 6-substitution pattern seen in *para*-alkyne probe **130** had been reported to ablate activity in the utrophin-FLuc reporter gene

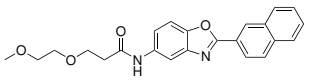
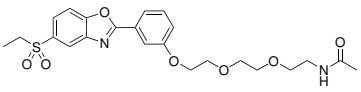
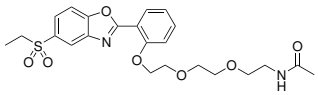
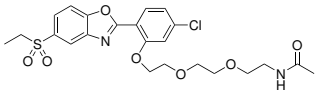
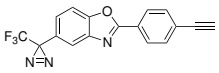
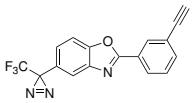
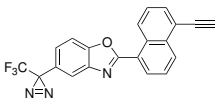
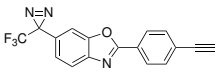
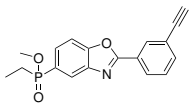
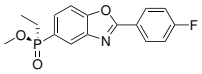
Compound	Structure	Mouse utrophin mRNA fold change	Mouse utrophin protein fold change	Human utrophin protein fold change	Myoblast fusion
65		∼2.2 ×	Not tested	Not tested	Not tested
52		∼1.5 ×	Not tested	Not tested	Inactive
74		∼2.5 ×	Not tested	Not tested	Inactive
73		Inactive	Not tested	Not tested	Inactive
106		∼1.5×	∼1.5 ×	∼1.5 × *	Active
107		∼1.5 ×	∼1.5 ×	∼1.8 ×	Inactive
129		Inactive	Inactive	Not tested	Inactive
130		∼1.5 ×	Not tested	Not tested	Active
118		∼1.5 ×	Not tested	Not tested	Not tested
151		Not tested	Not tested	∼2.0 ×	Not tested

Table 4.1: Summary table of biological activity of ezutromid-based PEGylated, dual-tagged and competitor probes; green: active, red: inactive. Fold changes shown are the highest obtained for each compound, at any concentration. mRNA data are the means of technical replicates (n = 3, no statistical significance calculated). Protein data are the means of biological replicates, n = 3, * denotes statistically non-significant result.

assay.[100] However, in section 2.7, **130** was revealed as an FLuc inhibitor, flagging this result as a potential false negative. With further investigation, it transpired that 6-substituted **130** retained ezutromid's ability to upregulate utrophin mRNA and also to increase fusion index. As a result, only 1-naphthyl probe **129** was taken forward as the negative control. Second generation compound SMT022332 was established as an active, non-toxic ezutromid analogue suitable for use as a competitor control for target identification studies, which are the subject of the next chapter.

Chapter 5

Target Identification by Affinity Capture Chemical Proteomics

5.1 Aims

Chapter 4 discussed the selection of biotinylated and dual-tagged probes for AfBPP studies, based on their biological activity. The aim of this chapter is to develop and deliver *in vitro* and *in situ* chemical proteomics strategies to identify probe-interacting protein targets by gel-based and mass spectrometry methods.

5.2 AfBPP experiments using biotinylated probes

In vitro affinity capture experiments using compound-immobilised beads require less optimisation, with fewer steps in the workflow than *in situ* photoaffinity experiments, and were therefore selected as the starting point.

5.2.1 Experimental design

The workflow for affinity purification experiments using a compound-bound matrix was outlined in Figure 1.8. First, probe 'baits' are immobilised on a matrix. Background signals from matrix-binding 'sticky' proteins can be diminished by incubating the protein source with the matrix alone, then exposing the 'pre-cleared' proteins to the probe-functionalised matrix for target binding. Wash steps to remove non-specifically binding proteins must be mild in order not to disrupt target binding. Finally, bound proteins can be eluted and identified by mass spectrometry, either directly on the enriched protein mixture or after SDS-PAGE and gel band excision.

The preparation of complex protein mixtures for mass spectrometry requires disulfide bond reduction and *S*-alkylation to prevent bridge reformation. These steps facilitate protease

digestion to generate peptides, which offer higher accuracy mass detection and identification by fragmentation compared to full length proteins.[249] The peptides can then be desalted on a C18 column, dried *in vacuo* and reconstituted. LC-MS/MS analysis begins with peptide separation using high-performance reverse phase liquid chromatography, from which the peptides are directly sprayed into the mass spectrometer. Upon ionisation, peptides are scanned and high abundance or pre-selected ions are isolated for fragmentation and measurement of the fragment ion mass spectra. Peptide and inferred protein identifications arise from matching the MS and MS/MS mass data to that of known peptides *via* database search. Methods for relative peptide quantitation between MS samples have been extensively reviewed[250,251] and include:

- Label-free quantitation (LFQ)
- Metabolic labelling through SILAC (stable isotope labelling by amino acids in cell culture)
- Isotope labelling of peptides using iCAT (isotope-coded affinity tags), iTRAQ (isobaric tags for relative and absolute quantitation), TMT (tandem mass tags) or dimethylation

Label-free quantitation was used for the MS data generated from the affinity capture experiments described in this chapter. LFQ entails comparison of a peptide's total signal intensity (area under the curve of mass spectra peaks) by searching its m/z and retention time across multiple samples. Assuming consistent sample preparation, stable LC-MS/MS machine and sufficient data redundancy, LFQ offers a robust and straightforward approach to quantitation, applicable for unlimited sample numbers.[252] Developments in software enable normalisation of retention time and peak intensity, suppression of noise and enhanced peptide peak detection, improving quantitation accuracy.[253] However, LFQ is low throughput as samples must be run individually rather than multiplexed, leading to lengthy instrument run times.

5.2.2 Analysis of enriched proteins by SDS-PAGE

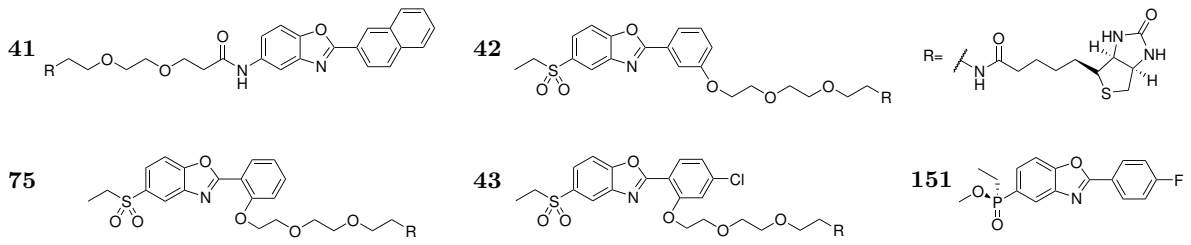


Figure 5.1: Structures of the biotinylated ezutromid-based probes.

Appropriate selection of lysis buffer can be important for efficient affinity capture, which is contingent on the cellular localisation and native-state stability of the target. In the absence

of a known target, preliminary experiments using active and inactive biotinylated probes **75** and **43** were carried out to compare the affinity capture protocol using different reported lysis buffers:

A) 25 mM HEPES pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 0.1% Tween-20, 0.1% Triton X-100[136]

B) 50 mM PIPES pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 5 mM EGTA, 0.1% β-mercaptoethanol, 0.1% NP-40, 0.1% Tween-20, 0.1% Triton X-100[131]

C) 50 mM Tris, 5% glycerol, 100 mM NaCl, 1.5 mM $MgCl_2$, 0.8% NP-40, 1 mM DTT[10]

The biotinylated probes (50 µM) were conjugated to streptavidin beads (1 h, rt), while H2K *mdx* lysate (2 mg/mL in each of buffers A, B and C) was precleared with additional streptavidin beads (1 h, 0 ℃). The probe-functionalised beads were washed then incubated with the precleared lysate (16 h, 4 ℃, 60 rpm). Unbound proteins, contained in the supernatant after pelleting the beads, were then removed and the beads washed three times with lysis buffer. Bound proteins were eluted from the streptavidin beads by heating (95 ℃, 5 min) in 1 × Laemmli buffer, 50 mM DTT and 8 M urea in 50 mM ammonium bicarbonate, then separated by SDS-PAGE (Figure 5.2).

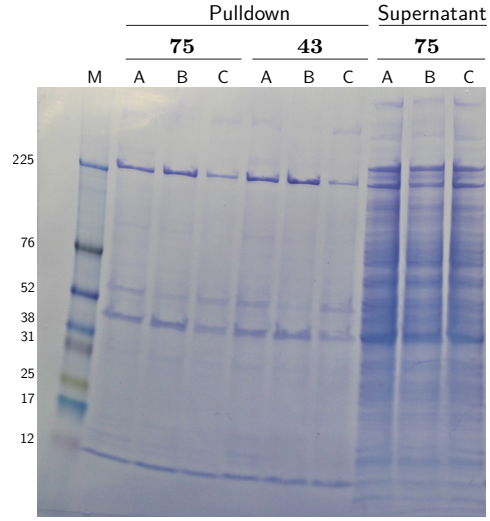


Figure 5.2: Proteins enriched by affinity capture using active and inactive biotinylated probes **75** and **43**; and **75** non-enriched proteins (the supernatant after pelleting of the protein-bound streptavidin beads), Coomassie staining. A, B, C: lysis buffers listed above, M: protein standard in kDa, uncropped gel presented in Appendix E.

Enriched proteins were stained with Coomassie brilliant blue (the ca. 12 kDa molecular weight band corresponds to monomeric streptavidin). Low abundance of binding proteins in the pulldown samples resulted in few observed bands. This could be due to weak binding affinity of proteins for the probes or insufficient sensitivity of the Coomassie stain. The detection limit for silver stains is lower than for Coomassie, and hence could be used to enhance staining,[254] although low abundance proteins would still be difficult to observe. The profiles

for the active and inactive probe appeared the same, with the enriched proteins therefore consistent with non-specific binders. Each of the lysis buffers appeared equally suitable for the protocol, and therefore just one, A, was used in the following experiments.

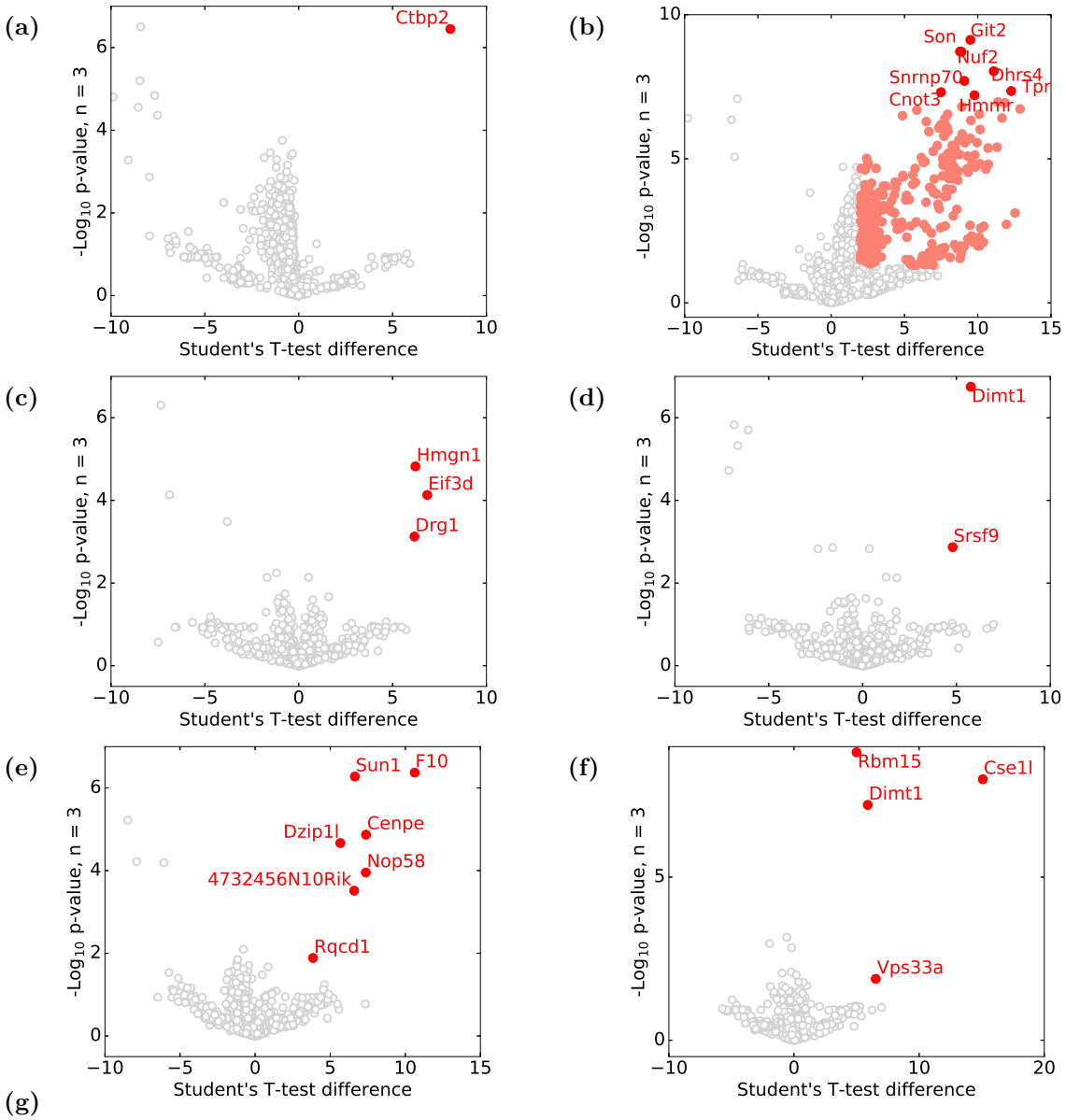
A significant challenge for enrichment of unknown targets is their unknown abundance and localisation. Investigations into target localisation could have been undertaken *via* the development of cell-permeable fluorophore-conjugated probes. Indeed, TAMRA- and BODIPY-based fluorophores exhibit good cellular uptake and intracellular distribution.[255] However, the effects of fluorophore conjugation on the overall probe cell permeability, localisation and activity are unknown and may require significant time and effort to determine and optimise. Therefore, in collaboration with the Mohammed group (Department of Biochemistry, Oxford) it was decided to proceed directly with analysis of affinity-enriched proteins by the more sensitive technique of mass spectrometry.

5.2.3 Analysis of enriched proteins by LC-MS/MS

For the MS experiments, the biotinylated probe 'baits' (active: **41**, **42** and **75**, inactive: **43**, Figure 5.1) were immobilised in triplicate on streptavidin beads, then exposed to cell lysates for target binding, in the presence or absence of competitor **151** (100 µM, Figure 5.1). Sample preparation was performed according to the standard procedures developed in the Mohammed group (described in the experimental methods), followed by analysis using nano liquid chromatography–tandem mass spectrometry (nanoLC–MS/MS). Two sample Student's *t*-tests were performed to distinguish the proteins enriched by the active probes but displaced by the competitor (results for probes **41**, **42** and **75** presented in Figure 5.3a, 5.3c and 5.3e respectively) and not enriched by the inactive control (Figure 5.3b, 5.3d, 5.3f). The most significantly enriched proteins are shown in Table 5.3g.

During the experiment, the beads functionalised with probe **41** appeared to aggregate after overnight incubation with lysate, although the explanation for this behaviour is unclear. MS analysis showed that a higher median peptide intensity was observed in the six probe **41** samples than in the others, presumably due to an increased background of non-specific binding with bead aggregation. This resulted in the majority of proteins appearing upregulated relative to the inactive control (Figure 5.3b). The high abundance of these proteins in the presence of competitor SMT022332, indicate the non-specificity of the interactions (Figure 5.3a). As such the hits from probe **41** were not considered relevant for target identification.

Results analysis for probes **42** and **75** on the other hand gave rise to a short list of



(g)

Gene	Protein	Major protein functions
Ctbp2	C-Terminal Binding Protein 2	Corepressor targeting diverse transcription regulators
Git2	GPCR Kinase Interacting ArfGAP 2	GTPase activator activity
Son	SON DNA Binding Protein	mRNA splicing cofactor
Nuf2	Cell Division Cycle-Associated Protein 1	Chromosome segregation and spindle checkpoint activity
Hmgn1	Non-Histone Chromosomal Protein HMG-14	Nucleosomal DNA binding
Eif3d	Eukaryotic Translation Initiation Factor 3 Subunit D	Translation initiation factor activity
Drg1	Developmentally-regulated GTP-binding protein 1	Translation factor GTPase
Dimt1	Dimethyladenosine transferase 1 homologue	Processing of pre-rRNA
Srsf9	Serine/Arginine-Rich Splicing Factor 9	RNA splicing
F10	Coagulation Factor Xa	Blood clotting factor
Sun1	SUN domain-containing protein 1	Component of the LINC (LInker of Nucleoskeleton and Cytoskeleton) complex
Cenpe	Centromere Protein E	Involved in cell division
Cse1l	Exportin-2	Mediation of nuclear transport
Rbm15	RNA-binding protein 15	Regulation of RNA methylation

Figure 5.3: **(a)**, **(c)** and **(e)**: Volcano plots of proteins enriched by biotinylated probes **41**, **42**, **75** respectively in the presence and absence of competitor **151**; **(b)**, **(d)** and **(f)**: volcano plots of proteins enriched by biotinylated probes **41**, **42**, **75** respectively compared to negative control biotinylated probe **43**. Proteins with $p < 0.05$ and t-test difference > 2 marked in red; **(g)** Top hits from **(a)**-**(f)**.

significantly enriched proteins ($p < 0.05$, marked in red in Figure 5.3 plots). However, no protein hits were identified by an active probe against both its controls. Interestingly, one protein, Dimt1, was observed as a significant hit by both active probes **42** and **75** compared to the inactive **43**, however, its role in processing of ribosomal pre-RNA[256] did not provide an obvious connection to a utrophin modulation mechanism. Similarly, literature searches for the other hits did not yield clear on-target mechanism hypotheses.

Since no one identified protein hit directly warranted a time-consuming target validation campaign, attention was turned to performing the *in situ* photoaffinity capture experiments. Photoaffinity experiments could alleviate problems caused by low affinity binding and may confirm one of these proteins as a hit of interest or provide alternative leads to follow.

5.3 *In situ* AfBPP experiments using photoaffinity probes

5.3.1 Experimental design and preliminary studies

The experimental workflow for the target identification in live cells by AfBPP using photoaffinity probes is outlined in Figure 5.4. Briefly, probe-treated cells are UV irradiated to trigger covalent binding of the probe to its protein target. A click reaction is performed to ligate a fluorophore or enrichment tag, such as biotin, to the probe. After enrichment, proteins are analysed using western blot or mass spectrometry.

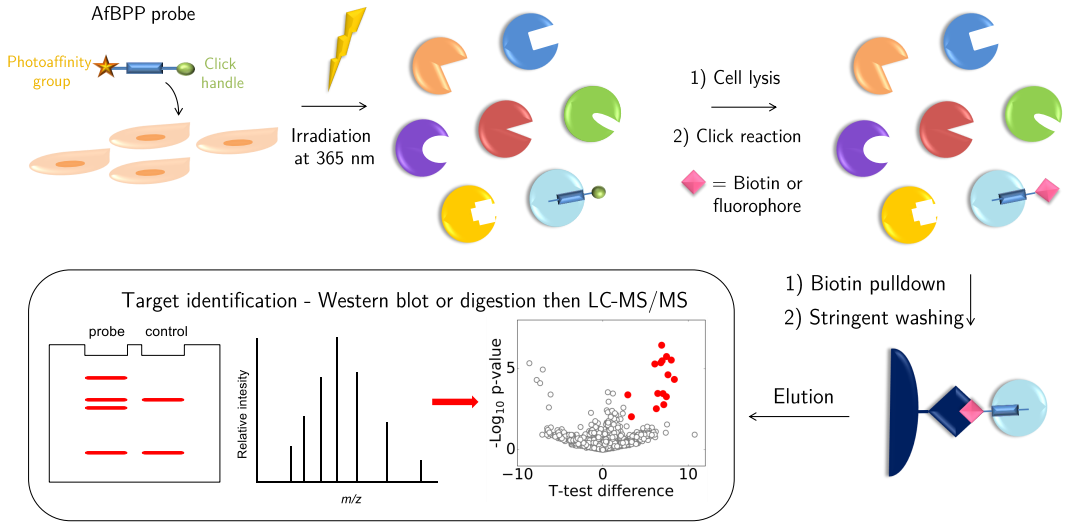


Figure 5.4: Outline of the AfBPP strategy for identification of the targets of small molecule probes in live cells.

With active and inactive dual-tagged ezutromid-based probes in hand (Figure 4.1), preliminary irradiation and click studies were performed to optimise conditions for their use in the AfBPP experiments. Photolysis of trifluoromethylphenyl diazirines can result in insertion into CH, OH, SH and NH bonds as well as undesirable side reactions from insertion into water

or molecular oxygen (Figure 5.5). Of note, insertions into NH bonds can lead to undesirable loss of photoaffinity labels after HF elimination followed by imine hydrolysis (Figure 5.5). As a result of these undesired reactions, productive target capture yields through covalent diazirine crosslinking are typically less than 10%.[257]

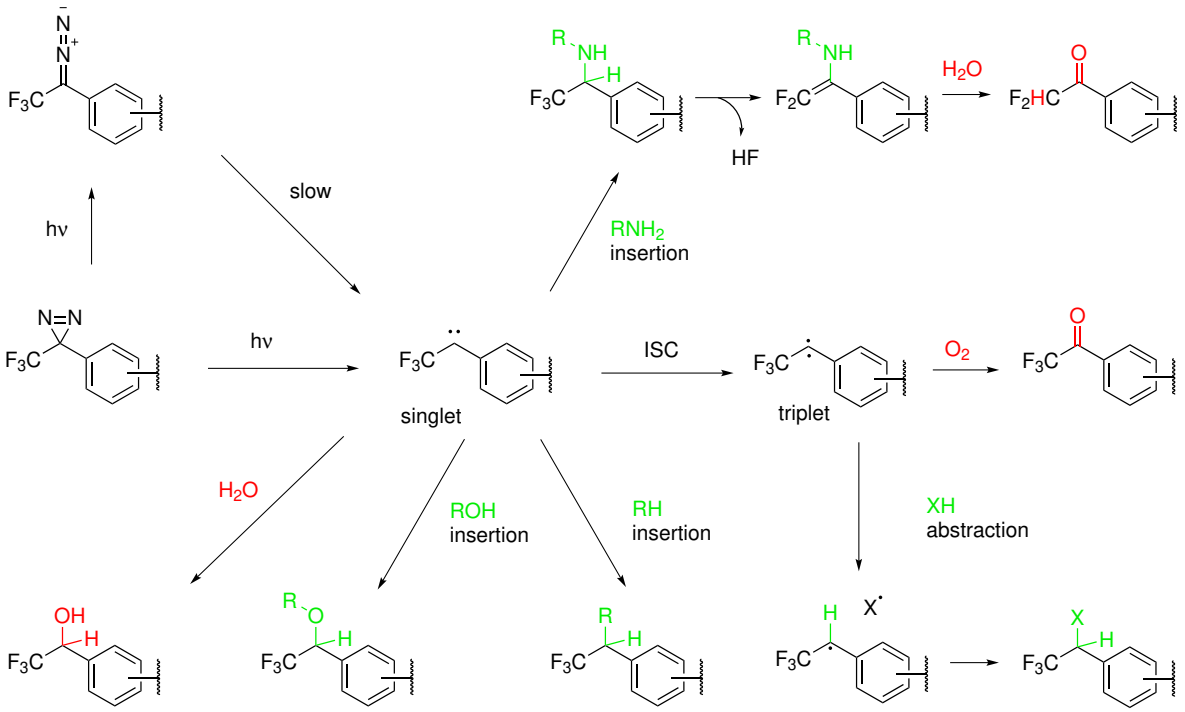
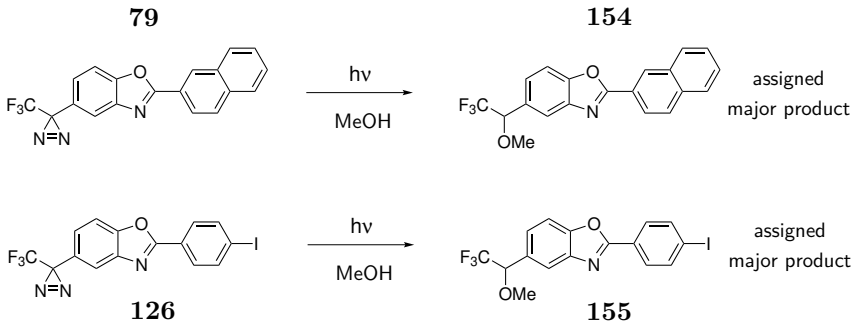


Figure 5.5: Reactions proceeding from photolysis of trifluoromethylphenyl diazirines can result in productive target capture (green) and undesirable side reactions (red). XH: CH, OH, NH, SH; ISC: intersystem crossing. Adapted from Smith and Collins.[148]

The duration of irradiation required to activate PALs depends on the type of PAL and the wavelength and power of the lamp used. Therefore, preliminary irradiation studies were carried out to optimise PAL activation, using diazirine probe **79** (used in chapter 3) and *para*-iodophenyl probe **126**, the precursor to active alkyne probe **106** which was used as a model system. The probes (2 mM in methanol) were irradiated (365 nm, 5 cm distance, 100 W lamp) for 0.5 min or 2.5 min on ice, and the products analysed by LC-MS (Figure 5.6).

The complete loss of the two starting material diazirine probes **79** and **126** was observed after 0.5 min irradiation. The major products had m/z values consistent with reaction with the methanol solvent to form methyl ethers **154** and **155**, although these products were not fully characterised. Additional minor products were not assigned. This experiment showed that the diazirines are rapidly activated, with product distributions similar after 0.5 and 2.5 min irradiation. Resultantly, an irradiation duration of 3 min was selected for use in the scaled-up photaffinity labelling experiments in lysates and live cells.

Attention was then turned to the optimisation of the click reaction required for bioorthog-

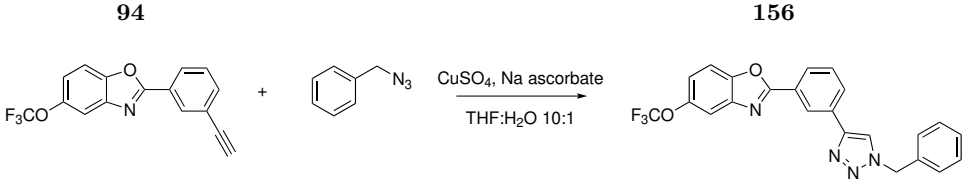


	t = 0		t = 0.5 min		t = 2.5 min		*m/z*	Assignment
	RT /min	Area /%	RT /min	Area /%	RT /min	Area /%		
	5.95	99.57		0		0	353.9	**79**
			6.57	4.65	6.56	2.36	296.2	–
79					6.73	10.79	440.3	–
			7.52	93.46	7.5	74.76	358.1	**154**
			8.21	1.89	8.19	1.11	130.2	–
			6.21	2.67	6.21	4.58	280.2	–
	6.45	100		0		0	N.D.	**126**
			6.74	19.69	6.63		296.2	–
126			6.93	3.1	6.92	5.2	396.3	–
			7.66	66.63	7.64	84.15	433.9	**155**
			8.39	1.04			130.2	–

Figure 5.6: LC-MS results from the preliminary irradiation studies. RT; retention time, N.D.; not detected (compound **126** does not ionise easily). Chromatograms and mass spectra are presented in Appendix F.

onal handle ligation in the AfBPP experimental workflow. Attempts to monitor progress of probe click reactions in lysis buffer by LC-MS were hindered however by poor aqueous solubility of the probe. Therefore, a model reaction using probe **94** and benzyl azide was performed in THF:H_2O (10:1) (Scheme 5.1), according to a literature procedure.[258] Probe **94** was chosen as a model rather than a diazirine-containing probe as the click reaction was envisaged to occur post PAL activation in the AfBPP workflow. The click reaction yielded 89% of product **156**, providing confidence to attempt the reaction on lysates as part of the AfBPP workflow.

Scheme 5.1: Model CuAAC reaction using ezutromid analogue **94** and benzyl azide.



H2K *mdx* cell lysates (4 mg/mL) were treated with diazirine probe **107** (2 µM, 0.5% final assay concentration (FAC) of DMSO) for 30 min at 0 °C. The lysates were then irradiated on ice (365 nm, 5 cm distance, 3 min) to enable target labelling. The proteins were then

denatured (addition of SDS to 1% FAC) and subjected to a CuAAC reaction with biotin azide according to the protocol reported by Heal *et al.*,[259] with a control sample containing no source of copper. Proteins enriched by streptavidin beads (procedure reported in the experimental methods) were analysed by SDS-PAGE (Figure 5.7). Protein labelling was weaker in the copper free sample, indicating a dependence of enrichment on the CuAAC reaction. This experiment indicated that the AfBPP workflow was operating satisfactorily and experiments using active probe **106** were initiated.

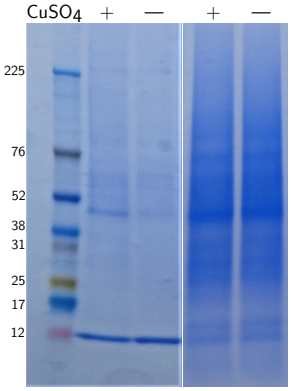


Figure 5.7: Labelling of the H2K *mdx* proteome by diazirine probe **107** is limited in the absence of copper. Left: proteins enriched by affinity capture using **107**, right: non-enriched proteins (the supernatant after pelleting of the protein-bound streptavidin beads); uncropped gel presented in Appendix E.

5.3.2 Analysis of enriched proteins by SDS-PAGE

First, the protein-interaction profile of active dual-tagged probe **106** was assessed with gel-based labelling experiments using live H2K *mdx* cells. This compound was used alongside two controls: (1) negative control probe **129** and (2) SMT022332 **151** dosed at >30 × concentration to competitively block enrichment of **106**-binding proteins.

Active probe **106** (3 µM), alongside DMSO vehicle, negative control dual tagged probe **129** (3 µM) and competitor SMT022332 (100 µM) were dosed for 2 h. After irradiation (365 nm, 0 ℃) and cell lysis, the probe-labelled lysate was subjected to a CuAAC-mediated click reaction with TAMRA-azide followed by SDS-PAGE.

In gel fluorescence showed probe-dependent labelling of proteins (Figure 5.8a), with little to no labelling in the vehicle control. Multiple proteins were labelled in the three conditions, with only a subset of proteins outcompeted by SMT022332 **151**, indicating a shortlist of specifically labelled targets amongst a majority of non-specific hits. The labelling of proteins with probe **106** was also shown to be concentration-dependent (Figure 5.8b), and again few proteins experienced reduced labelling in the presence of the competitor. However, in this subsequent experiment higher background labelling was observed in the DMSO-treated samples.

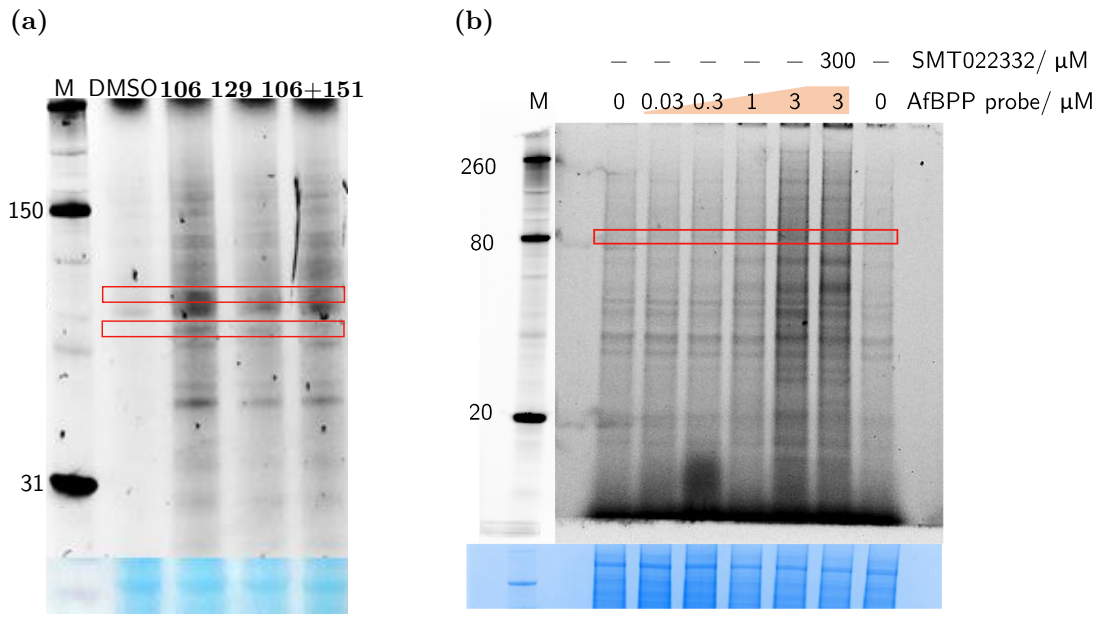


Figure 5.8: **(a)** In gel fluorescence of active probe **106**, inactive probe **129** and active **106** with competitor SMT022332 **151** labelled H2K *mdx* proteome; **(b)** probe **106** shows concentration-dependent labelling of protein targets. H2K *mdx* cells were treated with **106** (2 h) and irradiated (365 nm, 3 min, 0 ℃) to trigger crosslinking. Probe labelled proteins were ligated to TAMRA, separated by SDS-PAGE and visualised by in gel fluorescence. Red boxes, bands that appear stronger in active probe treated samples than in the negative controls; M, protein standard (in kDa); bottom, Coomassie blue stained gel represents total protein.

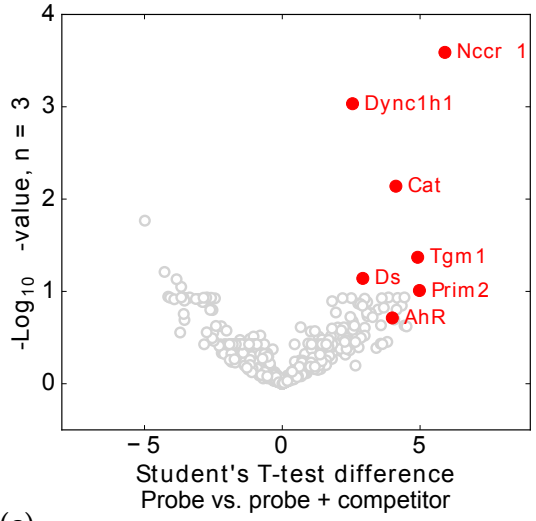
5.3.3 Analysis of enriched proteins by LC-MS/MS

Next, proteome-wide target identification proceeded with click-mediated ligation of probe-labelled proteins to biotin and affinity enrichment with streptavidin beads as above, followed by on-bead tryptic digest. Samples were analysed by LC–MS/MS as described in the experimental methods. Comparison of proteins enriched by active probe **106** but displaced in the presence of competitor SMT022332 gave a list of high occupancy binders (red proteins, Figure 5.9a).

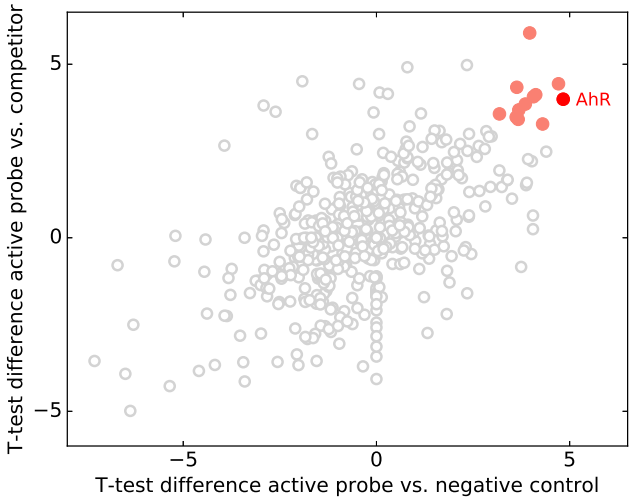
Combining differences in proteins identified in active probe samples but not in the two controls delivered a shortlist of potential targets (red marked proteins in Figure 5.9b and listed in Table 5.9c, full list of proteins identified available in Appendix G). Since none of the shortlisted proteins matched those identified by experiments using the biotinylated probes, the results of the previous target ID experiments were employed to determine the most promising potential targets. The DiscoverX BioMAP phenotypic profiling study, identifying biomarkers which are differentially activated post exposure to ezutromid, indicated a profile match with AhR antagonist CH223191 (previously discussed in section 1.6.2). Interestingly, AhR was found to be one of the most differentially enriched proteins in the *in situ* AfBPP LC-MS/MS experiment, comparing the active probe to the two inactive controls (Figure 5.9b). Although

the *t*-test difference for AhR was large (≥ 4 for the two controls), the p-value was not significant because the protein was found in only two of the three replicates for the active probe. However, it should be noted that AhR is not highly expressed in skeletal muscle cells,[260] which may have led to difficulties for MS detection.

(a)



(b)



(c)

Gene	Protein	Major protein functions
Nccrp1	F-box only protein 50	Positive regulation of cell proliferation
Rpn2	Dolichyl-diphosphooligosaccharide–protein glycosyltransferase subunit 2	Protein glycosylation
AhR	Aryl hydrocarbon receptor	Ligand-binding transcription factor
Cat	Catalase	Peroxidase
Dync1i2	Cytoplasmic dynein 1 intermediate chain 2	Organelle and vesicle transport
Arfip2	Arfaptin-2	Membrane ruffling
Dsg1a	Desmoglein-1-alpha	Cell-cell adhesion
Cyb5b	Cytochrome b5 type B	Electron transport

Figure 5.9: **(a)** Volcano plot of proteins pulled down by probe **106** +/− competitor **151**, top hits in red, n = 3 technical replicates; **(b)** protein enrichment by probe **106** compared to inactive probe **129** and competitor **151**. Proteins strongly enriched by **106** but not the controls in red; **(c)** Top hits from **(b)**.

5.4 Summary

This chapter discussed the use of four ezutromid-based biotinylated probes in affinity capture experiments in H2K *mdx* cell lysates alongside a competitor control. Meanwhile, dual-tagged photoaffinity probes were tested for their suitability as PAL and click reagents, then utilised in AfBPP experiments in live cells. Both experiments led to the identification of a shortlist of potential probe-binding proteins after LC-MS/MS analysis, including the aryl hydrocarbon receptor as a possible lead. Target validation is a time-consuming process and therefore further complementary target ID experiments were desired to aid hit prioritisation. Hit prioritisa-

tion with the support of CETSA-MS and RNA-seq studies and subsequent target validation experiments are the subject of the next chapter.

Chapter 6

Target Validation

6.1 Aims

The previous chapter described the identification of shortlists of potential ezutromid targets from both lysate and *in situ* MS affinity capture experiments. This chapter will discuss the following:

1. Prioritisation of protein hit(s) for target validation
2. Confirmation of binding interaction between the target and ezutromid
3. Validation of the target as a mechanism for utrophin modulation in human DMD myoblasts

6.2 Hit prioritisation

Deconvolution of therapeutically relevant hit(s) from irrelevant and non-specific ones poses a significant challenge. In this work, two additional target ID methods were employed to complement the results of the probe affinity capture experiments: CETSA and RNA-seq. Both are physiologically relevant '*in situ*' global profiling techniques, and were chosen for their orthogonality to affinity purification methods and hence differing limitations. This combination of experiments increases the likelihood of successful target identification and aids prioritisation of the most promising hit for target validation studies.

6.2.1 CETSA

This method relies on the principle of thermodynamic equilibria, whereby proteins which bind compounds are stabilised, and relatively preserved against unfolding and aggregration at high temperatures compared to unbound proteins (Figure 6.1). Aggregated protein masses are insoluble and readily removed by centrifugation.[169]

A CETSA-MS experiment was carried out in collaboration with the Huber group (Target Discovery Institute, University of Oxford), dosing H2K *mdx* myoblasts with ezutromid (3 µM)

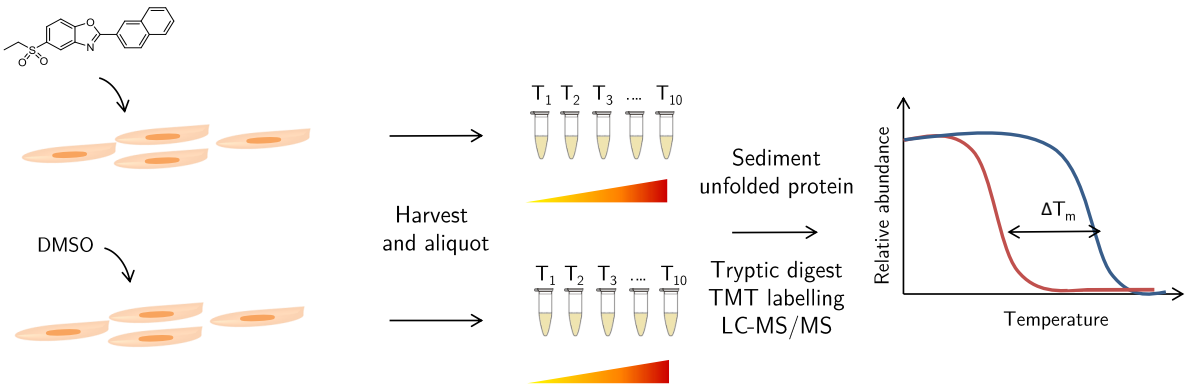


Figure 6.1: Schematic overview of the Cellular Thermal Shift assay (CETSA)

for 30 min at 37 ℃ in duplicate alongside DMSO vehicle control (Figure 6.1). The cells were detached then divided into ten aliquots and heated at one of ten different temperatures, ranging from 37−73 ℃, before lysis and centrifugation to pellet unfolded proteins. The samples were prepared for LC-MS/MS submission with a standard tryptic digest procedure then multiplexed with TMT 10plex labels. Sample submission and results analysis were carried out by Kathryn Pugh and collaborators of the Huber group.

Unfortunately, poor reproducibility between replicates prevented conclusive differentiation between treated and untreated samples in most of the identified proteins. However, one protein, Park7/DJ-1, was identified as a tentative hit and is reported to have a role in muscle growth through positive regulation of the PI3K/AKT pathway in Callipyge sheep,[261] making it a potentially relevant target for DMD. Disappointingly, probing the same CETSA samples with an anti-Park7/DJ-1 antibody did not indicate stabilisation by ezutromid on Park7 relative to DMSO control (Figure 6.2a), nor was it identified in the affinity capture experiments. Therefore Park7/DJ-1 was not validated as a target of ezutromid, and was pursued no further.

A second CETSA experiment was commissioned by Summit Therapeutics plc. to be carried out by Pelago Bioscience AB (Solna, Sweden). This time, '2D CETSA' was carried out with ezutromid on human DMD myoblasts and lysates, where both concentration (0.03, 0.3, 3, 10 µM) and temperature were varied to give EC_{50}s of the identified binding interactions. In total, 6000 proteins were sequenced and approximately 300 weak compound-target interactions were identified. However, very few proteins were significantly stabilised (Figures 6.2c-6.2d), and of these, many were considered to be 'frequent hitters' which are often perturbed by small molecule treatment. For example, the protein identified with highest confidence, ferrochetalase, which was dose-dependently stabilised by ezutromid (Figure 6.2b) is a known frequent hitter. The proteins which were most significantly stabilised in response to ezutromid are listed in Table 6.1. Unfortunately, AhR protein was not identified in the samples.

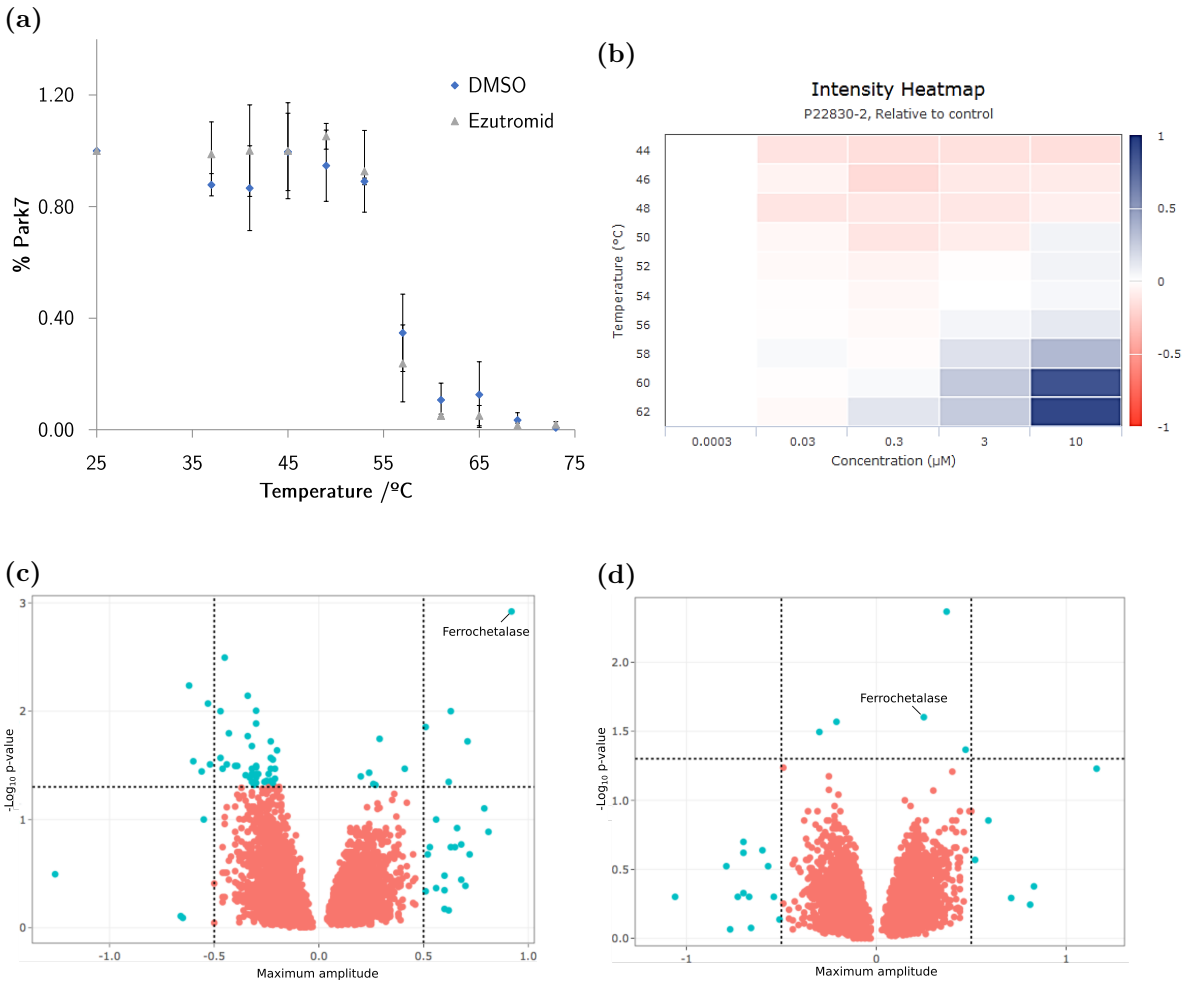


Figure 6.2: **(a)** Plot of western blot derived data of CETSA samples, quantifying for Park7/DJ-1 protein normalised to baseline temperature 37 ℃; **(b)** Heatmap of ferrochetalase protein intensity in the 2D CETSA samples, show dose-dependent stabilisation by ezutromid. Blue: stabilised, red: destabilised, relative to DMSO vehicle control; **(c)** - **(d)** Volcano-plots of CETSA-MS results in intact cells **(c)** and lysate **(d)**. Dashed lines indicate the p-value (<0.05) and amplitude cut-offs (>0.5)

As yet there are few examples of thermal profiling being used for target identification[262–264] and there are a number of limitations of CETSA. These include the assumptions that the ligand binds in a classical reversible sense, as opposed to slow or tight binding, and that the binding has an appreciable effect on protein stabilisation. For instance, multi-protein complexes as targets may not show significant stabilisation if the ligand only binds at a small part of it.[265] Next, the abundance of the target protein in the cell must be considered, due to the detection limit of mass spectrometry. Lastly, whole cell procedures such as CETSA generally identify large numbers of interacting proteins as a result of polypharmacology and false positive effects, of which only a minority are truly involved in the therapeutic mechanism. No obvious connections to utrophin modulation or DMD were observed amongst these proteins, nor any correlation with proteins identified in the affinity purification experiments, impeding MoA hypothesis generation.

Protein	Gene name	Experiment	p-value	Max. Amplitude	Frequent hitter?
Ferrochetalase	FECH	*in situ* and lysate	0.0012 and 0.025	0.92 and 0.25	✓
Pirin	PIR	*in situ* and lysate	0.018 and 0.0043	0.29 and 0.37	×
Nicotinamide N-methyltransferase	NNMT	*in situ*	0.01	0.63	Maybe
Glioblastoma-Amplified Sequence/NipSnap2	GBAS	*in situ*	0.014	0.51	✓
Glyoxylate and Hydroxypyruvate Reductase	GRHPR	*in situ*	0.019	0.71	×
Phosphodiesterase 6D	PDE6D	lysate	0.043	0.47	×
Methylmalonyl-CoA Mutase	MUT	*in situ*	0.037	0.24	×
Methylenetetrahydrofolate Dehydrogenase	MTHFD1	*in situ*	0.04	0.2	✓
Quinoid Dihydropteridine Reductase	QDPR	*in situ*	0.047	0.26	×

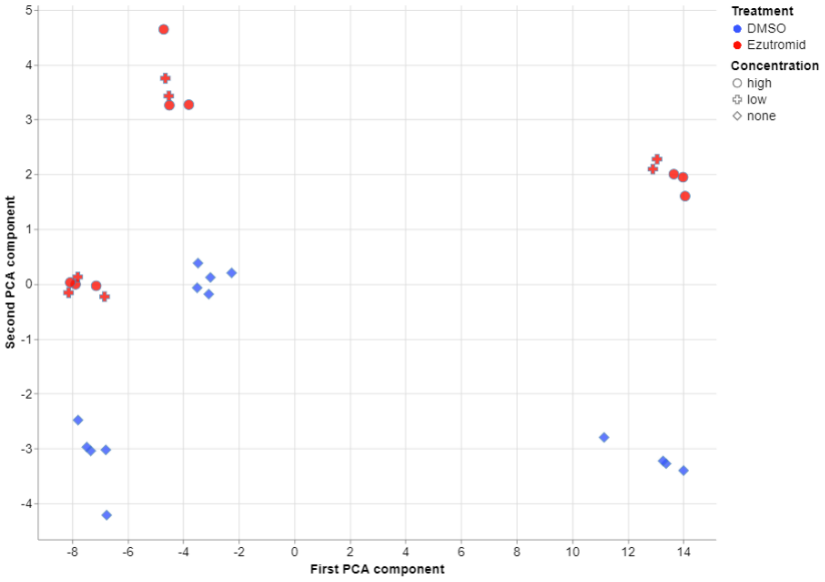
Table 6.1: Manually evaluated proteins that are stabilised by Ezutromid treatment, $p < 0.05$. Max. amplitude is the highest log fold change observed for any temperature, relative to the DMSO vehicle and normalised to the median fold change. The p-value is determined by comparing the amplitudes in a window of four concentration conditions at four consecutive temperatures to the general trend of the entire dataset (assuming a normal distribution around 0).

6.2.2 RNA-seq study

As discussed in section 1.5.2, identifying changes in gene expression that are caused by compound treatment can reveal the mechanism by which the compound operates. Therefore, an RNA-seq study was carried out by Evotec AG to assess changes in RNA expression as a result of ezutromid treatment. Human DMD myoblasts were treated 24 h after seeding with ezutromid (3 or 10 µM) or DMSO vehicle control (0.1%) for 4, 8 or 24 h. After RNA extraction, sample preparation and next generation sequencing, differences in gene expression were calculated for each timepoint and concentration. Overall, treatment of ezutromid resulted in small differences to the transcriptome, reflected in small variations in the principle components of the samples (Figure 6.3a). Of the differentially expressed genes, most significant was the aryl hydrocarbon receptor repressor (AhRR), which was downregulated (log fold change of -0.3 to

−2.0) in almost all ezutromid treatment conditions. Meanwhile, the aryl hydrocarbon receptor (AhR) was found enriched in treated samples compared to the vehicle (Figure 6.3b). These results point to an ability of ezutromid to modulate AhR signalling and expression. Since AhRR is a downstream target of AhR,[266] the downregulation observed indicates antagonism of AhR signalling.

(a)



(b)

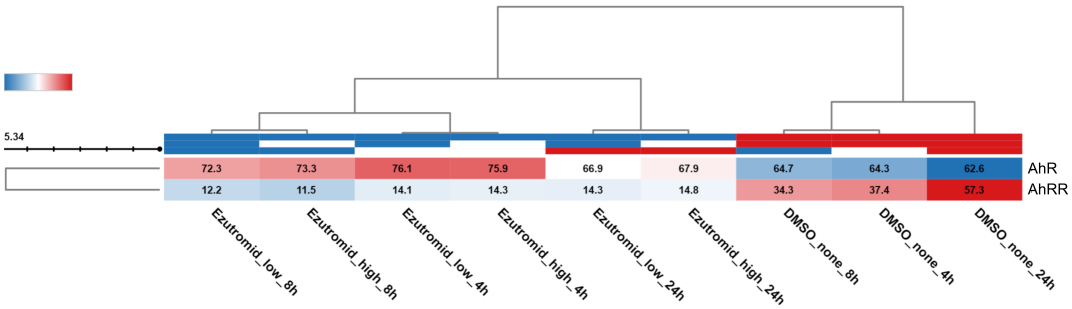


Figure 6.3: **(a)** Principle component analysis of RNA-seq samples, red: ezutromid treatment (crosses: 3 µM, circles: 10 µM), blue: DMSO vehicle; **(b)** heatmap of RNA expression of AhR and AhR repressor, values indicate expression level; blue: downregulated, red: upregulated relative to DMSO vehicle control. Low: 3 µM, high: 10 µM

The DiscoverX phenotypic profiling study described in section 1.6.2 suggested ezutromid behaves similarly to AhR antagonist CH223191. Based on this finding, coupled with the RNA-seq results and the identification of AhR as a possible target in the chemical proteomics experiment, AhR was prioritised for follow-up confirmatory and validation experiments.

6.3 Introduction to the aryl hydrocarbon receptor (AhR)

AhR is a pleiotropic ligand-binding transcription factor which acts as an environmental sensor to control complex transcriptional programmes in a ligand-, cell-type- and context-dependent

manner. AhR has been widely proposed as a potential therapeutic target for immune and inflammatory diseases, cancer chemotherapy, bone remodelling, neurological disorders and more.[267–271]

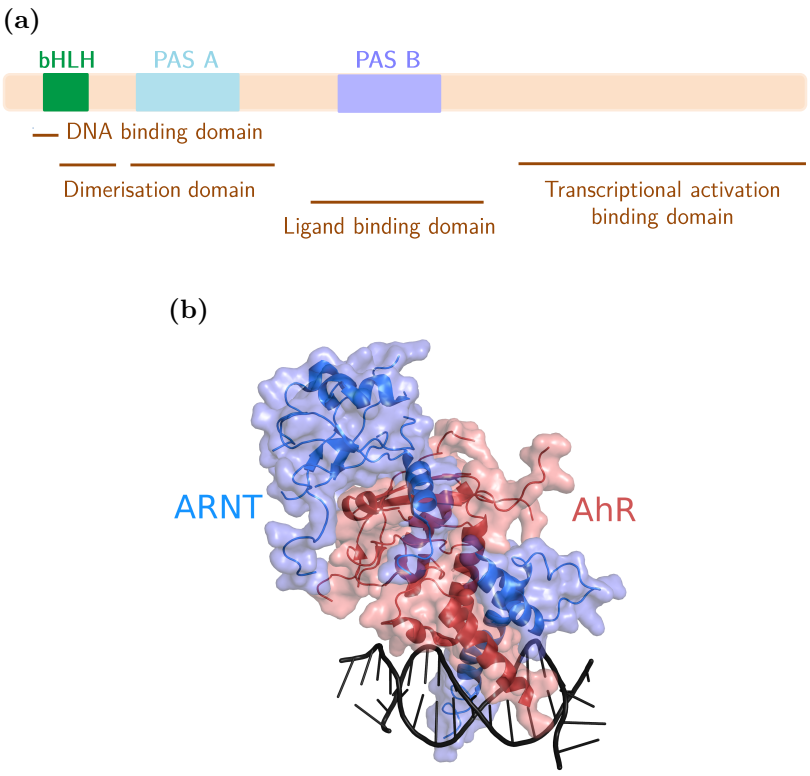


Figure 6.4: **(a)** Architecture and function of aryl hydrocarbon receptor protein domains; bHLH, basic helix-loop-helix domain; PAS, Per-ARNT-Sim domain (A and B repeat regions); **(b)** Crystal structure of truncated AhR (red) and ARNT (blue) proteins in complex with DNA. The two truncates span the start of the bHLH domains and finish before the start of the PAS B domains of their respective proteins, PDB deposit: 5v0l.[272]

AhR is composed of an N-terminal basic helix-loop-helix (bHLH) domain required for DNA binding, two Per-ARNT-Sim domains (PAS A and B) and C-terminal transactivation domain[273] (Figure 6.4a). AhR binds structurally diverse small molecule agonists and antagonists, with its PAS B domain constituting the ligand binding domain.[274,275] AhR is found constitutively inactive, bound to chaperone proteins including Hsp90 and Xap2 in the cytosol. Agonist binding induces conformational changes in AhR, leading to chaperone dissociation and exposure of nuclear localization sequences.[276] Upon nuclear translocation, AhR heterodimerises with the AhR nuclear translocator protein (ARNT, also known as HIF1-β), to form a stable, high-affinity DNA-binding complex[277] (Figure 6.4b).

The classical minimal recognition motif of the AhR:ARNT complex is 5′-GCGTG-3′ and termed either the AhR-, dioxin- or xenobiotic- responsive element (AHRE, DRE or XRE).[278,279] Promoter-bound AhR:ARNT activates transcription of AhR-responsive genes (Figure 6.5a)

involved in a range of processes including development, immunity, cancer and the xenobiotic response. AhR:ARNT also induces expression of the AhR repressor protein (AhRR), a bHLH/PAS family protein which is capable of dimerising with ARNT but lacks ligand-binding and transactivation domains, thus behaving as an AhR-competitive inhibitor.[266] In this way, a negative feedback loop is created to regulate the effects of agonist-binding. Antagonist-binding to AhR results in increased retention of AhR in the cytosol and a decrease in expression of AhR-responsive genes compared to the wild-type baseline (Figure 6.5b). Baseline expression of AhR responsive genes is the result of dynamic basal nucleocytoplasmic shuttling, which was revealed in studies utilising nuclear export inhibitors which reported accumulation of unliganded AhR in the nucleus.[280]

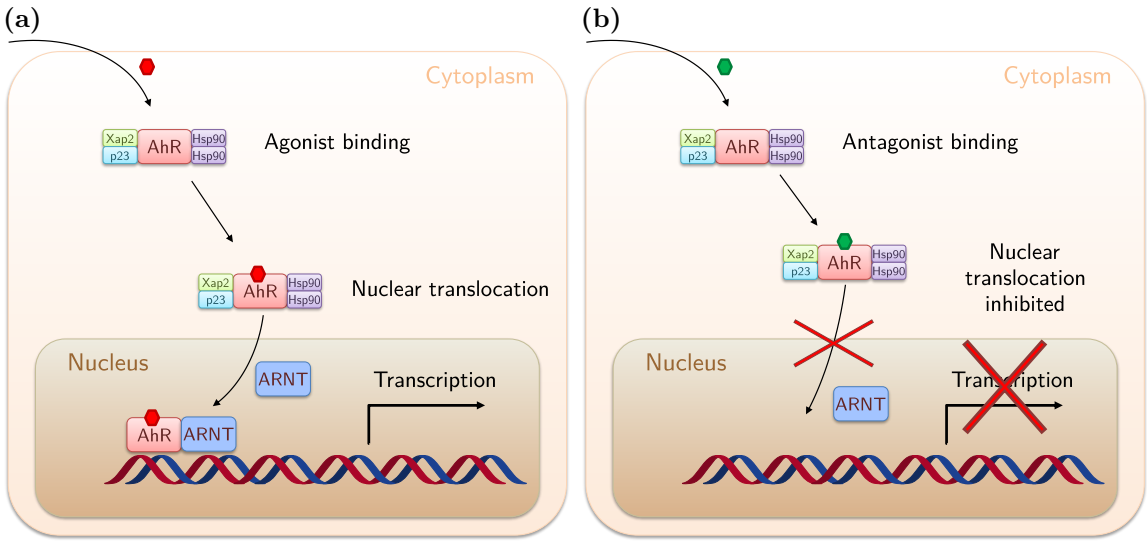


Figure 6.5: Localisation of AhR induced by treatment with an agonist **(a)** and antagonist **(b)**.

AhR binds a range of endogenous, naturally occurring and xenobiotic ligands. Historically, the most studied ligands are AhR agonist environmental contaminants including halogenated or polycyclic aromatic hydrocarbons including dioxins, dibenzofurans, biphenyls such as **158**, pyrenes **159**, anthacenes and fluorenes **160** (Figure 6.6).[281] The prototypical AhR agonist 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) **157** is amongst the most potent AhR ligands, with a K_D of 1 pM – 10 nM (species- and assay-dependent).[282,283]

However, AhR binds structurally diverse ligands, indicating a promiscuous binding site fit for its role as an environmental sensor, and leading to diverse responses, many of which do not induce toxicity.[284] Amongst non-classical AhR agonists are ITE **161**,[285,286] YH439 **162**[287] and leflunomide **163**[288] (Figure 6.6). Many metabolites and derivatives of dietary indoles such as tryptophan and indole-3-carbinol have been found to be AhR ligands,[281] including

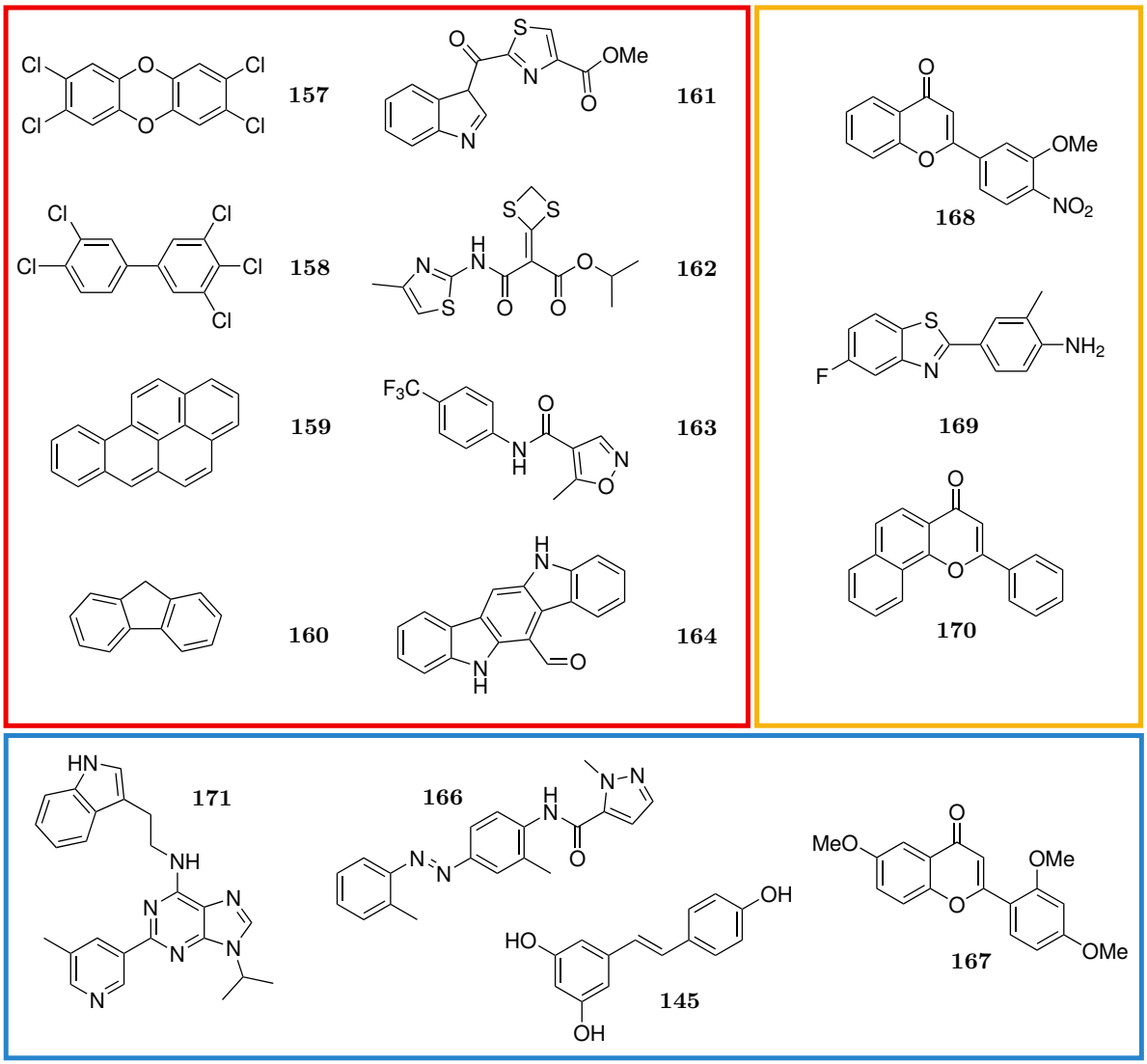


Figure 6.6: Examples of AhR agonists (red box): **157**, TCDD; **158**, PCB126; **159**, benzo[*a*]pyrene; **160**, fluorene; **161**, ITE; **162**, YH439; **163**, leflunomide; **164**, FICZ; partial AhR agonists/antagonists (orange box): **168**, 3′-methoxy-4′-nitroflavone; **169**, 5F203; **170**, α-naphthoflavone; and AhR antagonists (green box): **171**, GNF-351; **166**, CH223191; **145**, resveratrol; **167**, 6,2′,4′-trimethoxyflavone.

the potent endogenous ligand 6-formylindolo[3,2-*b*]carbazole (FICZ) **164**.[289] Reported AhR antagonists include GNF-351 **165**[290] (and structurally related SR1[291]), CH223191 **166**,[292] resveratrol **145**[293] and 6,2′,4′-trimethoxyflavone **167**[294] (Figure 6.6).

However, in the absence of a crystal structure of the ligand binding PAS B domain, there is much unknown about the complex different binding modes and mechanisms of AhR agonists and antagonists. For example, CH223191 **166** and 6,2′,4′-trimethoxyflavone **167** are reportedly capable of inhibiting AhR activation by TCDD **157** and other halogenated aromatic hydrocarbons, but have little effect on other agonists including flavonoids and polycyclic aromatic hydrocarbons, thus exhibiting a ligand-specific antagonism, the mechanism for which is unknown.[284] Concentration-dependent partial AhR agonism/antagonism has also been reported for ligands including α-naphthoflavone **170** and 3′-methoxy-4′-nitroflavone **168**,

both of which exhibit antagonist activity at lower concentrations and full agonism at higher concentrations.[295,296] Furthermore, species-specific partial AhR agonism/antagonism has been reported for benzothiazole 5F 203 **169** (rat vs. human)[297] and synthetic flavones (mouse vs. guinea pig).[298] So far, site directed mutagenesis studies have shown that certain ligand binding domain residues are responsible for ligand selectivity and ligand-dependent Hsp90 displacement and hence determination of agonist/antagonist behaviour.[287,299] In one case, a single mutation (F318L) was found to be sufficient to transform AhR agonist β-naphthoflavone to a partial agonist/antagonist.[299]

In summary, AhR biology is highly complex and only partially understood. However, it is eminently druggable and its potential as a therapeutic agent for cancer, inflammatory disorders and more makes it an exciting target. The remainder of this chapter investigates AhR as a receptor for ezutromid, and as a potential therapeutic target for DMD.

6.4 Validation of the binding between ezutromid and AhR

The MS result presented in section 5.3.3 indicated that ezutromid-based photoaffinity probe **106** binds AhR in mouse myoblasts. Ezutromid's planar hydrophobic structure is characteristic of AhR ligands and reminiscent of benzothiazole ligand 5F 203 **169**. However, experimental evidence of direct binding between ezutromid and AhR was necessary.

6.4.1 Validation of MS pulldown hit by western blot

To validate the MS result from the previous chapter and investigate its relevance for DMD patients, the affinity capture experiment was performed in human DMD myoblasts. Binding of AhR to probe **106** but not negative control probe **129** was verified by western blot, using an AhR-specific antibody (Figure 6.7).

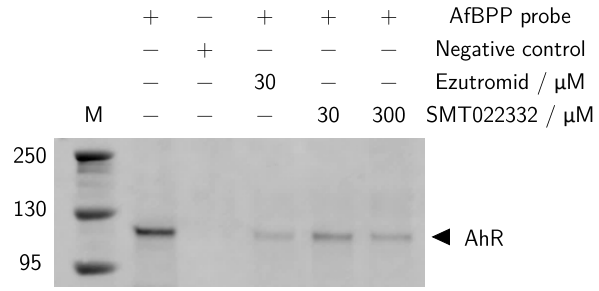


Figure 6.7: Immunoblot of AfBPP samples post biotin enrichment, derived from human DMD myoblast cells, showing selective enrichment of AhR by active probe **106**, but not negative control inactive probe **129** and enrichment competition between ezutromid and SMT022332. M; protein standard (in kDa), uncropped blot is presented in Appendix E.

The binding of competitor SMT022332 to human AhR was confirmed by dose-dependent

disruption of AhR enrichment by probe **106**. Furthermore, AhR enrichment by **106** was markedly reduced in the presence of ezutromid, indicating ezutromid's ability to competitively bind the human aryl hydrocarbon receptor.

6.4.2 Determination of binding affinity between ezutromid and AhR

Competition of affinity capture by AfBPP probe **106** indicated binding between ezutromid and AhR. Next, confirmation and characterisation of the binding interaction was desired. Evaluation of AhR ligands in the literature typically exploits cellular reporter assays, with reporters (usually firefly luciferase) fused to the promoters of AhR-responsive genes like Cyp1a1 or AhRR.[290,292,294,300–305] Furthermore, INDIGO Biosciences Inc. (Pennsylvania, USA) provides commercial kits and services based on a firefly luciferase reporter gene assay. However, ezutromid's FLuc inhibition preclude the utility of FLuc reporter gene assays, and moreover, Cyp1a1 induction has been found to be a non-specific biomarker of direct AhR binding.[306] Characterisation of protein-ligand binding by biophysical methods such as isothermal titration calorimetry (ITC),[307] fluorescence and NMR spectroscopies,[308,309] surface plasmon resonance (SPR),[310] biolayer interferometry (BLI)[311] and microscale thermophoresis (MST)[312] require testing of purified protein. Recombinant full-length AhR protein production *via* baculovirus[313] and TNT® Quick coupled transcription/translation reticulocyte cell-free protein (Promega)[285,303,314] expression systems have been reported, however challenges in production of soluble AhR containing the PAS B domain have meant that a purification method for the full protein has not been published.[272] As a result, AhR binding studies in the literature often use hepatic cytosol extracts as a source of AhR.[290,292,294,306]

Competitive ligand displacement studies have been reported using tritium-labelled or PAL versions of prototypical AhR agonist TCDD **157**. In the former, cytosol extracts are incubated with [^{3}H]TCDD with and without the ligand of interest as a competitor. Subsequently, radioligand signals of the samples are compared after separation of bound and unbound ligand by dextran-charcoal adsorption,[293,297,313] velocity sedimentation on sucrose gradients[285,303,315] or hydroxyapatite adsorption.[299,306,316]

Radiolabelled TCDD-based PAL probe 2-azido-3-[^{125}I]iodo-7,8-dibromodibenzo-*p*-dioxin **172** (Figure 6.8) was synthesised by Poland and co-workers in 1986.[317] Arylazide probe **172** has been used in competitive binding assays with AhR agonists benzo[*a*]pyrene **159** and fluorene as well as antagonists GNF-351 **165** and 6,2′,4′-trimethoxyflavone **167** with radiolabelled proteins analysed by SDS-PAGE and autoradiography.[290,294,317] PAL probe **172** has also been used to

demonstrate increased TCDD-binding affinity of mouse over human AhR.[318]

172

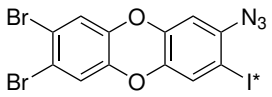


Figure 6.8: Photoaffinity-labelled analogue of AhR agonist TCDD **157**; I*, ^{125}I.

Recently, apparent dissociation constants have been reported for the binding interaction between purified mouse and human truncated AhR:ARNT complexes with resveratrol **145** and structurally related anti-inflammatory agent tapinarof, using a fluorescence quenching assay.[127] This assay exploits the compounds' intrinsic fluorescence in solution which is quenched upon protein binding, and was employed in chapter 3. Purified mouse truncated AhR:ARNT (including the AhR bHLH, PAS A and B domains, kindly provided by Prof. Fraydoon Rastinejad, TDI, Old Road Campus, Oxford) allowed for the binding affinity of ezutromid to be investigated. The assay revealed a potent apparent K_D of 50.1 ± 22.1 nM (Figure 6.9), confirming ezutromid as an AhR ligand. The high affinity of this interaction indicates relevance to ezutromid's treatment phenotype.

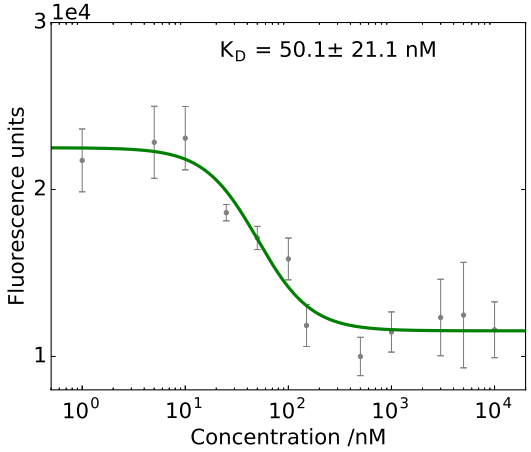


Figure 6.9: Fluorescence of ezutromid (100 nM, 313 nm excitation, 390 nm emission) measured after incubation (16 h) with AhR:ARNT (5 − 10000 nM). Fitting with a four-parameter logistic function provides an approximate K_D of 50.1 ± 22.1 nM.

6.5 Characterisation of ezutromid as an AhR antagonist

Ezutromid's profile match with AhR antagonist CH223191 **166** in the DiscoverX BioMAP study, and the decreased expression of AhR-responsive gene AhRR in the RNA-seq study indicate that ezutromid has AhR antagonist activity. The ability of ezutromid to antagonise AhR activation was further investigated by immunofluorescence of AhR localisation after compound treatment. Human DMD cells were treated with ezutromid (3 µM, 2 h) and imaging

with an AhR-specific antibody revealed no induction of nuclear translocation, unlike that observed with the potent endogenous AhR agonist ITE **161** (1 µM, 2 h), which was used as a positive control (Figure 6.10). Co-treatment of ezutromid with ITE led to retention of AhR in the cytoplasm (Figure 6.10), manifesting the ability of ezutromid to counteract the action of a potent AhR agonist.

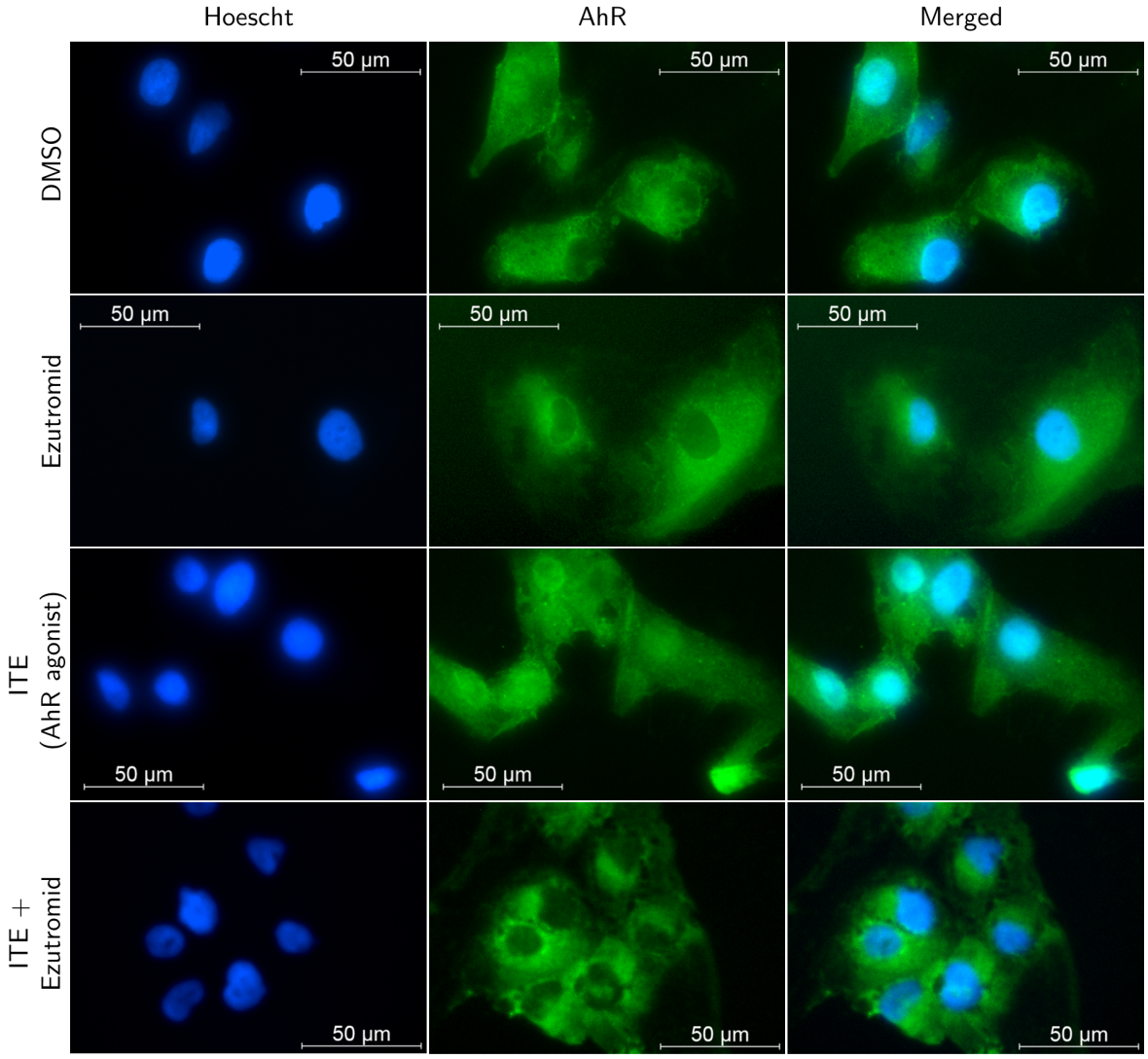


Figure 6.10: Localisation of AhR induced by compound treatment, determined by immunofluorescence. AhR is mostly localised in the cytoplasm of human DMD myoblasts, ezutromid does not increase nuclear translocation, unlike agonist ITE. Co-treatment of ezutromid with ITE results in retention of AhR in the cytoplasm.

6.6 Towards validation of AhR as a target for utrophin modulation

To determine whether AhR antagonism represents a viable mechanism to modulate utrophin in human DMD cells, AhR antagonists GNF-351 **165** and CH223191 **166** were assessed for their effect on utrophin protein levels by western blot (Figure 6.11). Gratifyingly, GNF-351

treatment (24 h) showed a dose-dependent 2.7-fold increase in utrophin protein, which was accompanied by a 3-fold increase in AhR protein (both $p < 0.05$). Meanwhile, dosing with antagonist CH223191 led to a 2-fold increase in utrophin which was not statistically significant, and without a corresponding increase in AhR protein. The result for GNF-351 provides encouraging evidence that AhR antagonism could be exploited as a mechanism for utrophin modulation. The smaller, less significant increase in utrophin observed with CH223191 treatment may be related to the lack of AhR induction, although further studies are required to confirm this.

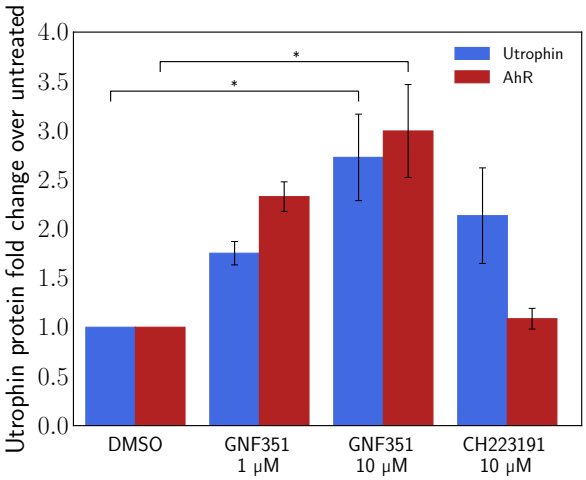


Figure 6.11: AhR antagonist GNF-351 increases AhR and utrophin, while antagonist CH223191 increases utrophin without changing AhR, determined by western blot. Bars represent means of expression normalised to total protein and relative to DMSO, error bars are standard error of the mean, $n = 3$. * denotes $p < 0.05$.

The identification and validation of AhR as a target for ezutromid is an intriguing result given the evidence for utrophin modulation by AhR ligands in the literature. For example, resveratrol **145**, which has AhR antagonist activity,[293] has been shown to increase utrophin expression and decrease inflammation in *mdx* mice.[90] Furthermore, AhR agonist PCB-126 has been shown to decrease utrophin mRNA expression in rat liver progenitor cells.[319]

A mechanism by which AhR antagonism modulates utrophin has not been established. However, prototypical AhR agonist TCDD **157** has been shown to decrease activity and abundance of transcriptional coactivator peroxisome proliferator-activated receptor-γ coactivator-1α (PGC1α),[320] which stimulates utrophin expression at neuromuscular junctions.[83] This dysregulation was removed by co-treatment with an AhR antagonist, leading to stabilisation of active PGC1α. Treatment with AhR antagonists alone could lead to an increase in PGC1α and consequently utrophin.

6.7 Summary

In conclusion, unbiased global CETSA-MS and RNA-seq studies were conducted to generate MoA hypotheses as methods orthogonal to the affinity capture experiments discussed in the previous chapter. While the CETSA experiments did not produce high confidence hits, the most significant changes in mRNA expression observed in the RNA-seq study were consistent with AhR antagonism, supporting the finding of AhR amongst top hits from the *in situ* chemical proteomics experiment and previous phenotypic profiling work performed by DiscoverX. AhR was consequently prioritised for target validation.

The identification of AhR as a target by the photoaffinity affinity capture MS experiment in H2K *mdx* cells was confirmed in human DMD cells. Furthermore, the binding interaction between dual-tagged PAL probe **106** and AhR was displaced by ezutromid treatment. Ezutromid was found to bind AhR:ARNT potently with an apparent K_D of 50.1 $\pm$ 22.1 nM, and to behave as an AhR antagonist, decreasing the expression of AhR-responsive genes and antagonising AhR nuclear translocation. Finally, two additional, structurally distinct AhR antagonists, GNF-351 and CH223191, were found to upregulate utrophin, providing the first evidence that AhR may be a valid target for utrophin modulation.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

The utrophin modulator ezutromid **5** was originally developed using a phenotypic screen, and hence its molecular target(s) were unknown. The aim of this thesis was to uncover ezutromid's mechanism of action to help understand its phase II trial results and thus develop improved utrophin modulators.

7.1.1 Summary of research findings

In this work, biotinylated and dual-tagged photoaffinity ezutromid-based probes for use as active and inactive controls in affinity capture target ID studies were designed and synthesised (Figure 7.1). Truncated forms of the biotinylated probes were also generated for use in cellular assays, thus evading the limited cell permeability of the biotin-containing probes. Recognition of ezutromid's behaviour as an FLuc inhibitor led to detailed kinetic analysis demonstrating its behaviour as a non-competitive inhibitor with a novel but unknown binding site. Many of the ezutromid-based probes were also found to exhibit FLuc inhibition behaviour, and hence their bioactivity could not be accurately determined in the utrophin-FLuc reporter gene assay. As a result, they were tested for their ability to emulate ezutromid's therapeutic effects in *mdx* and human DMD myoblasts in orthogonal experiments including RT-qPCR, western blot and a myoblast fusion assay.

Three of the truncated biotinylated probes (**41**, **42** and **75**) were found to be active, with two (**41** and **75**) retaining similar levels of activity as ezutromid. Further, the probe designed to serve as a negative control (**43**) was confirmed to be inactive. The two dual-tagged photoaffinity probes designed to be active (**106** and **107**) were also found to upregulate utrophin mRNA production, although to a lesser extent than ezutromid. One of the intended negative control compounds (1-naphthyl probe **129**) was confirmed not to increase utrophin,

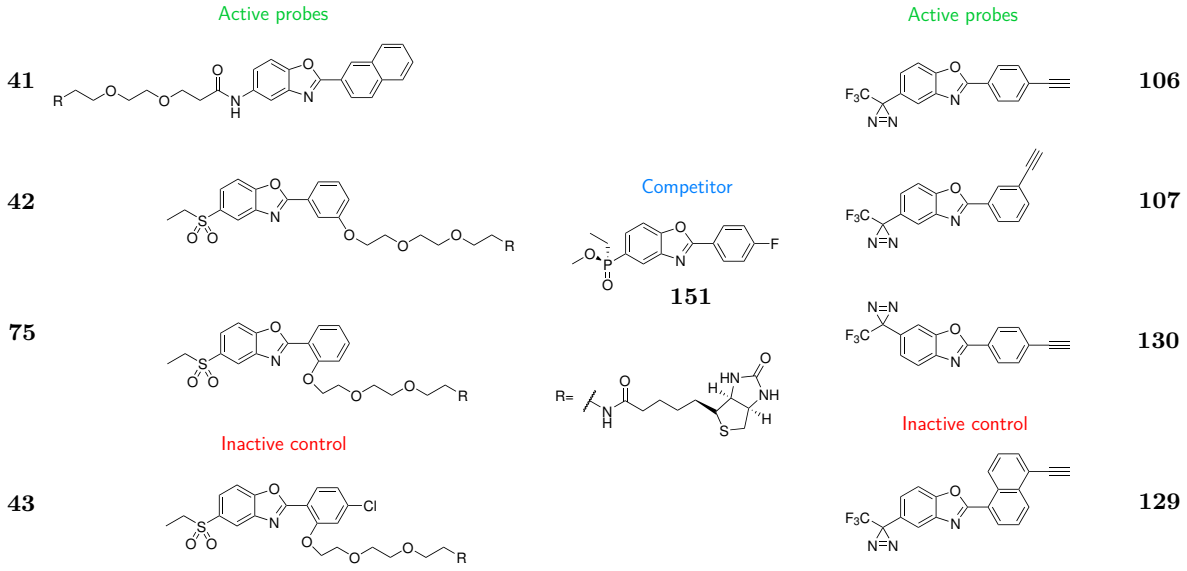


Figure 7.1: Biotinylated and dual-tagged photoaffinity ezutromid-based probes synthesised in this work.

while the other (6-substituted **130**) was found to be as active as **106** and **107**. Probes **106** and **107** were also found to increase utrophin protein in both mouse and human DMD cells, while probe **129** was confirmed to be inactive. However, only **106** retained ezutromid's ability to increase the rate of myoblast fusion. *Para*-alkyne probe **106** had consistently exhibited utrophin modulating activity and was most similar in treatment phenotype to ezutromid and was therefore selected for use as the active AfBPP probe, alongside inactive control **129**. Additionally, second generation phosphinate compound SMT022332 **151** was established as an active, non-toxic ezutromid analogue suitable for use as a competitor control.

Next the protein interactomes of the biotinylated and photoaffinity ezutromid-based probes were profiled in *mdx* myoblasts and analysed by LC-MS/MS. Significantly enriched proteins were identified by comparing the active probe samples with both inactive and competitor controls. The results were then integrated with those obtained from CETSA, RNA-seq and DiscoverX BioMAP studies. Ultimately, the aryl hydrocarbon receptor was identified as a potential target through this combination of chemical proteomics and phenotypic profiling techniques, and prioritised for validation studies.

Ezutromid was found to compete with the active photoaffinity AfBPP probe for binding to AhR in human DMD myoblasts and to bind mouse AhR:ARNT with high affinity (apparent K_D of 50.1 ± 22.1 nM). Ezutromid was also revealed as a novel AhR antagonist, inhibiting nuclear translocation and downregulating the AhR-responsive gene AhRR in human DMD cells. Finally, two additional AhR antagonists were found to upregulate utrophin in human DMD myoblasts, one of which in a dose-dependent manner ($p < 0.05$), suggesting that AhR

may be a valid target for utrophin modulation in DMD patients.

7.1.2 Significance of this work

While numerous AhR agonists have been described, and are additionally found amongst existing pharmaceuticals,[306] antagonists are much fewer in number.[303] Furthermore, while human and mouse AhR are known to have substantial differences in ligand specificity,[321] in this work, ezutromid is shown to bind both species' receptors. AhR antagonists are typically free of the dioxin-like carcinogenicity of some AhR agonists and are increasingly being investigated as possible therapeutic agents in a diverse range of indications including multiple myeloma,[322] treatment-resistant triple-negative breat cancer,[303] glioblastoma,[323] head and neck squamous cell carcinoma,[324] atheroscelerosis,[325] obesity[326] and rheumatoid arthritis.[327] Furthermore, AhR antagonism has a promising application in stem cell expansion and development[291,328,329] and cancer immunotherapy.[330] As a result, the identification of ezutromid as a novel, non-toxic AhR antagonist represents an exciting addition to the literature by expanding the toolbox available to interrogate AhR biology. Additionally, ezutromid's activity as an anti-cancer or anti-inflammatory agent could be investigated.

Phenotypic screening for drug discovery has many advantages and disadvantages, but importantly, when it is combined with target identification, it can lead to the discovery of novel therapeutic mechanisms. Signalling pathways connecting AhR to utrophin modulation now warrant further investigation given that a known AhR antagonist, GNF-351 also upregulates utrophin protein expression, suggesting AhR is a valid target for utrophin functional replacement therapies. This work could pave the way for the first target-based drug discovery program in DMD as well as provide a biomarker for future clinical trials. Furthermore, AhR modulation could be used in combination with other therapeutic approaches, amplifying efficacy. Excitingly, the results of this work justifies research into a new approach to deliver a utrophin modulator to benefit all DMD patients.

7.2 Future work

The discovery of AhR as the target of ezutromid opens up numerous exciting avenues for future work. This section will propose experiments towards achieving the following goals:

- Explore AhR biology using ezutromid as a tool compound
- Determine whether third generation utrophin modulators developed by our team operate by AhR antagonism or by alternative mechanisms
- Define the pathway by which AhR modulates utrophin expression
- Determine whether AhR modulation explains the whole ezutromid treatment phenotype, or if other targets are involved
- Investigate whether AhR 'on-target' activity explains ezutromid's lack of sustained clinical efficacy

7.2.1 Investigation of the mechanism of ezutromid AhR antagonism

Ezutromid has been identified as a novel AhR antagonist, however full characterisation of this activity, including binding site, ligand-selectivity and AhR response, has yet to be accomplished. AhR binding site determination in the literature has been limited to agonist competition (typically using TCDD) and ligand binding domain docking studies. However, the binding site could also be identified using a PAL strategy, and should be particularly feasible using probe **106** which identified AhR in the photoaffinity capture experiment. Incubation of **106** with purified AhR followed by irradiation, digestion and LC-MS/MS analysis could reveal the probe's AhR binding site, assuming covalent addition of the probe to the protein and sufficient protein sequence coverage.

Ezutromid's antagonism of AhR could be further investigated to assess its time- and concentration-dependence using mRNA expression of AhR responsive genes as a readout. This experiment would also be important for eliminating the possibility of ezutromid partial AhR agonism. It would additionally be interesting to determine whether ezutromid antagonises the action of all agonists or is ligand-selective in the manner of CH223191.[292] Competition studies using [^{3}H]TCDD could be performed according to well-precedented procedures in the literature (discussed in section 6.4.2) or by comparing the effect of ezutromid treatment in combination with AhR agonists on AhR:ARNT DNA binding *in vitro* by gel retardation analysis, as performed by Denison and co-workers.[281] Alternatively, competition may be investigated using the fluorescence quenching experiment reported in this thesis, by co-dosing

with excess concentrations of other AhR ligands and assessing the impact on apparent K_D. However, a prerequisite for this experiment is that the chosen competing AhR ligands must not significantly absorb the excitation wavelength.

In this work, retention of AhR in the cytosol upon ezutromid treatment was observed using immunofluorescence. This finding and any attendant ligand-selectivity could be extended by investigating AhR localisation after ezutromid treatment +/− various AhR agonists *via* Hsp90 co-localisation with AhR[301] or western blotting of treated nuclear and cytosolic fractions.

Next, the AhR-mediated response to ezutromid binding should be characterised. Ligand-specific AhR responses are well-documented in the literature, whereby ligands selectively produce some downstream AhR-dependent responses but not others. Ligand-specific differences in AhR co-activator recruitment[331] and AhR targeted DNA sequence have been reported,[332,333] suggesting ligand-induced differences in AhR structure may account for the diversity of ligand-specific AhR responses. Additionally, the generation of an AhR mutant unable to bind to the DRE promoter has identified a ligand-activated DRE-independent role for AhR in suppression of inflammation.[334] This activity operates via suppression of NF-κB regulated-gene expression, especially cytokine-induced acute-phase genes including SAA1, CRP, LBP, C3, C1S, and C1R, and is driven by AhR agonists including TCDD and benzo[*a*]pyrene.[334] Another reported transcription-independent AhR signalling function is mediation of Jak2-Src crosstalk.[335]

GNF-351 **171** has been termed a 'pure' antagonist for its ability to antagonise DRE-dependent AhR activity as well as alleviate suppression of acute-phase gene expression caused by DRE-independent AhR activity, unlike antagonist 6,2′,4′-trimethoxyflavone **167** which only exhibits the former.[290] Meanwhile, most other AhR antagonists, including CH223191 **166**, have only been assessed for their ability to modulate DRE-dependent rather than -independent AhR activity. To determine whether DRE-independent AhR activity is required for utrophin modulation, it would be instructive to investigate whether GNF-351's 'pure' antagonism phenotype is shared by ezutromid. This could be achieved by investigating ezutromid's ability to antagonise AhR agonist driven suppression of NF-κB regulated acute-phase gene expression using RT-qPCR. Additionally, the ability of selective AhR antagonist 6,2′,4′-trimethoxyflavone **167** to modulate utrophin could be assessed by reporter gene assay, RT-qPCR or western blot.

Taken together, future investigations into the mechanism of ezutromid's AhR antagonism could give insights into AhR's complex biology, while also potentially revealing its mechanism of utrophin modulation by analogy with other AhR antagonists. Finally, evidence for AhR antagonism by third generation utrophin modulators developed by our team could be assessed

using DRE-promoter reporter gene assays or RT-qPCR of AhR-responsive genes alongside direct AhR binding studies. Evidence of AhR-independent utrophin modulation would warrant further target identification studies using these compound classes and opens the door to combination therapies.

7.2.2 Investigation of AhR as a therapeutic target for DMD

The work described in this thesis indicates that AhR could be a therapeutic target for utrophin modulation. However, there is an increasing awareness that polypharmacology (i.e. a compound binding multiple targets) can contribute to a therapeutic effect.[336] Indeed, 21% of FDA approved agents in 2015-2017 were multi-targeting drugs.[336] Ezutromid was found to weakly bind ∼300 proteins in the 2D CETSA-MS study (section 6.2.1) and multiple other proteins were identified as significant hits in the affinity capture experiments. Therefore, further studies are required to confirm whether AhR is solely (or principally) responsible for ezutromid's utrophin modulating activity.

Firstly, it would be interesting to determine whether the ability of ezutromid and GNF-351 to upregulate utrophin is ablated in AhR null model systems. Genome editing using CRISPR-Cas9 is a powerful tool in interrogate signalling pathways and evaluate drug targets in living systems. CRISPR-Cas9 deletion of the aryl hydrocarbon receptor is precedented in the literature,[329,337] proving that it is a non-essential gene, but it has never been studied in the context of muscular dystrophy. Alternatively, knockdown of AhR by shRNA as precedented by Boitano and co-workers[291] could be investigated, although achieving loss of AhR expression is reportedly challenging, requiring 48-72 h.[290] Additionally, a PROTACs probe has been reported for AhR antagonist apigenin (4′,5,7-trihydroxyflavone).[338] This strategy for AhR knockdown could also be adapted for ezutromid, potentially exploiting the existing PEG chain linker SAR established in this thesis.

However, one significant consideration for these experiments is whether loss of AhR (either knockout or knockdown) is an appropriate model for AhR antagonism, where AhR is retained in the cytoplasm and possibly performs other roles. Furthermore, in this work, treatment with antagonists GNF-351 and ezutromid was demonstrated to actually lead to increased AhR abundance compared to untreated. Therefore, it may be of more interest to investigate the effect of AhR overexpression (in the absence of agonist treatment) on utrophin modulation. An alternative experiment could investigate the effects on ezutromid-induced utrophin modulation when combining treatment with an excess of AhR agonists or when employing a constitutively

active AhR construct which lacks a ligand binding domain.[291] Observation of a loss in activity in these two cases would indicate that ezutromid's ability to modulate utrophin is AhR-dependent.

Next, the pathway by which AhR modulates utrophin should be investigated. A possible pathway discussed in section 6.6 was AhR-mediated stabilisation of PGC1α, a utrophin transcription factor.[83] This could be investigated by measuring abundance and post-translational modification status of PGC1α by western blot after ezutromid or vehicle control treatment. Furthermore, ezutromid's effect on AhR-responsive gene TIPARP (TCDD-inducible poly(ADP-ribose) polymerase, PARP7) which has also been shown to influence PGC1α abundance could be investigated.[320] Ezutromid treatment could also be combined with inhibitors or activators of signalling pathways influencing utrophin expression, including NFAT, PPAR and MAPK/ERK pathways. Loss of activity in one of these combined treatments could indicate the pathway by which ezutromid is operating. Additionally, assessment of phosphorylation of kinases such as p38 MAPK, AMPK and JNK by western blot could also reveal which utrophin regulation pathways are activated by ezutromid treatment.

In addition to these targeted approaches, the mechanism of ezutromid-induced utrophin transcription could be investigated by the generation of a compound-resistant cell line and the subsequent identification of resistance-conferring mutations by next generation sequencing. Interestingly, the results of the ezutromid phase II trial, although disappointing, offer the opportunity to study loss of compound efficacy in human DMD patients, which could reveal aspects of its mechanism of action (discussed in the next section).

7.2.3 Explanation of PhaseOut DMD trial results

It has not been established whether AhR as ezutromid's target is responsible for its lack of sustained efficacy in the clinical trial or affects its overall exposure. In the short term, antagonism of AhR, which regulates the xenobiotic response, is likely to suppress ezutromid metabolism. However, it is unclear what the long term consequences of an increase in AhR protein caused by antagonism would be.

The study of changes in gene expression in ezutromid-treated liver cells to determine whether induction of xenobiotic metabolism or accommodation to AhR antagonism occurs could help establish whether AhR 'on-target' activity is responsible for ezutromid's limited clinical efficacy. Global transcriptomics or proteomics in tissues of ezutromid-treated *mdx* mice could investigate similar changes in gene expression in an *in vivo* model. These exper-

iments could be additionally performed on the clinical trial biopsies comparing the baseline prior treatment, the trial midpoint when efficacy was observed and the trial endpoint when efficacy was lost. Alternatively, the trial samples could be analysed for specific biomarkers (tissue sectioning and immunofluorescence or homogenisation and western blot), particularly if proteins key to ezutromid-induced utrophin upregulation had been identified in the experiments discussed in the previous section. One protein which could also be investigated further is the AhR repressor (AhRR), the gene found to be most differentially expressed in the RNA-seq study. The downregulation of AhRR observed with compound treatment was an indication of ezutromid's AhR antagonist activity. It would be interesting to establish whether its repression is maintained over chronic treatment.

The rate of metabolism of ezutromid was observed to increase over the course of the phase Ib study in DMD patients (discussed in section 1.4.2) and this progressive loss of compound exposure is likely to contribute to loss of sustained efficacy in the phase II trial. Examination of the phase II clinical trial samples for ezutromid and its metabolites would test this hypothesis. CYP1A1 and CYP1A2 have been determined as the primary P450 isoforms mediating ezutromid metabolism,[339] and were also found among the main enzymes involved in metabolism of GNF-351, leading to its poor absorption and limited *in vivo* efficacy.[340] Induction of these enzymes is a consequence of AhR agonism, but is not expected with treatment of an AhR antagonist. However, CYP1A1/2 induction could explain the increased rate of ezutromid metabolism observed and the subsequent loss of efficacy. Therefore, CYP1A1/2 expression and activity should be monitored in human DMD cells during chronic exposure of ezutromid. Additionally, ezutromid's metabolites should be screened for AhR binding and agonist activity, which could counteract the parent compound's therapeutic effects.

Interestingly, less clearance and markedly higher plasma concentrations of ezutromid were observed in healthy volunteers compared to DMD patients. Investigating whether the AhR antagonism of ezutromid is sustained over chronic treatment in non-DMD models, for example by extending the CYP1A1/2 studies to wildtype mice, could reveal whether ezutromid treatment could be effectively deployed for different indications.

Chapter 8

Experimental Methods

8.1 Biology Experimental Methods

Materials

The following materials were obtained from commercial sources: resveratrol (Santa Cruz Biotechnology #sc-200808), NF-κB inhibitor NF-κBAI4 (Cambridge Bioscience Ltd #10009536), D-luciferin (Promega, #E1602), ATP (Sigma), recombinant firefly luciferase (QuantiLum Promega, E1701), biotin azide (Sigma #762024), TAMRA-biotin-azide (DC Biosciences #CCR-1048), TBTA (Cayman Chemical #18816), TCEP (Sigma #C4706), CH-223191 (Selleckchem #S7711), GNF351 (Sigma #182707), ITE (R&D systems #1803/10) and heregulin (Recombinant Human NRG1-beta 1/HRG1-beta 1 EGF Domain Protein, R&D systems #396-HB-050). Ezutromid, SMT022332 and ethyl phosphinic acid were provided by Summit Therapeutics plc. AhR:ARNT protein complex was kindly provided by Prof. Fraydoon Rastinejad (Target Discovery Institute, Old Road Campus, University of Oxford).

Cell culture

H2K-*mdx*[99] and H2K-*mdx* utrnA-luc cells[97,100] were maintained in DMEM (Life Technologies) supplemented with 20% Fetal Bovine Serum (Life Technologies), 2% CEE (SLI), 2 mM L-Glutamine (Life Technologies), 1% Penicillin Streptomycin (Life Technologies) and 2 µg/500 ml Mouse Interferon-γ (Roche). Cells were maintained at 10% CO_2 at 33 °C.

Immortalised DMD myoblasts isolated from the Fascia lata muscle of a 10 year old male, del 52 DMD (KM571DMD10FL) were acquired through collaboration with Professor Vincent Mouly (Insitut de Myologie, Paris). These were cultured in Skeletal Muscle Cell Growth Medium and Supplement (PromoCell C-23060), 20% Fetal Bovine Serum (Life Technologies) and 1% Penicillin Streptomycin (Life Technologies). Cells were maintained at 5% CO_2 at 37

°C.

H2K *mdx* Utrophin A-FLuc reporter gene assay

White flat bottomed 96 well plates (Corning) were seeded with 5000 H2K *mdx* utrnA-luc cells. After 24 h, cells were dosed with compound in triplicate, in the following concentration series: 0.01, 0.03, 0.1, 0.3, 1.0, 3.0, 10.0 µM from 10 mM solution stocks in DMSO (final DMSO concentration was 0.3%). The cells were incubated for a further 24 h, (10% CO_2, 33 °C). Relative luminescence readout after using the Luciferase Assay System (Promega, E1500) reagents was measured using a FLUOstar Optima plate reader (BMG Labtech). The means from the biological triplicates were fitted with a four parameter logistic function with least squares regression (Levenberg-Marquardt algorithm) to calculate EC_{50} values.

Biochemical FLuc inhibition assay

Recombinant FLuc was assayed at a final concentration of 0.6 nM in a buffer containing 25 mM HEPES, 5 mM $MgCl_2$, 1 mM EDTA, 5 mM DTT and 1 mg/ml BSA. Compounds were diluted in the following concentration series: 0 nM, 0.95 nM, 3 nM, 30 nM, 95 nM, 0.3 µM, 0.95 µM, 3 µM, 9.5 µM, 30 µM, from 10 mM stocks in DMSO (with a final assay concentration of DMSO at 0.3%). PTC124 was used as a positive control.[219] Luciferase substrates ATP and D-luciferin were used in a final assay concentration of 10 µM. Luciferase was pre-incubated with ATP and the query compound at 0 °C for 15 min. D-Luciferin was dispensed and endpoint luminescence output immediately read using a FLUOstar Optima plate reader (BMG Labtech). Luminescence output was fitted with a four parameter logistic function with least squares regression (Levenberg-Marquardt algorithm) to calculate EC_{50} values.

FLuc fluorescence quenching assay

Fluorescence quenching of ezutromid upon binding recombinant FLuc was monitored using a fluorescence spectrophotometer (Hitachi F-4500 FL spectrophotometer, 317 nm excitation/390 nm emission). 0.3 µM ezutromid was incubated with FLuc (various concentrations ranging $0 - 3$ µM) for 15 min in a buffer containing 25 mM HEPES, 5 mM $MgCl_2$, 1 mM EDTA, 5 mM DTT and 1% DMSO. Fluorescence emission was measured in each sample and the K_D calculated by fitting the curve with a four parameter logistic function.

Docking of FLuc luciferin pocket

Docking studies were carried out using AutoDock Vina using default settings.[229] The luciferin pocket of FLuc crystal structure PDB: 4e5d was prepared as follows: bound ligands were extracted using PyMOL, the protonation state relevant for pH 7.4 was applied using a PDB2PQR[341,342] server and the binding box was prepared using AutoDockTools4.[343] Ligand files were prepared from sdf files with 3D coordinates generated using ChemAxon tools[344] then prepared for docking using AutoDockTools4.

FLuc kinetics assay, varying ATP, luciferin or LH_2-AMP as the sole substrate

Recombinant FLuc was assayed at a final concentration of 0.6 nM in a buffer containing 25 mM HEPES, 5 mM $MgCl_2$, 1 mM EDTA, 5 mM DTT and 1 mg/ml BSA. Ezutromid was diluted in a concentration series from 10 mM stocks in DMSO (with a final assay concentration of DMSO at 0.3%). Either D-luciferin or ATP was varied in the following concentration series: 1, 5, 10, 50, 100, 500 µM in triplicate, with the other substrate held at the saturating concentration of 250 µM (added immediately prior to measuring luminescence output). For the LH_2-AMP experiment, a concentration series of 0.5, 1, 2.5, 5, 10, 25 µM LH_2-AMP was used and no other substrate was supplied. Luminescence output was measured every 0.04 s for the first 4 s of the reaction using a FLUOstar Optima plate reader (BMG Labtech) and the initial rate was calculated by linear slope fitting. The initial rates data were plotted and fitted using Dynafit4 for inhibition model discrimination by Akaike and Bayesian Information Criteria.[345]

FLuc crystallisation studies

For the apo crystals, FLuc (13.1 mg/mL as supplied in a buffer containing 25 mM Tris pH 8.5, 0.2 M ammonium sulfate, 16 mM magnesium chloride, 1 mM EDTA, 1 mM DTT, 10% glycerol and 25% ethylene glycol from Promega, E1701) was centrifuged to sediment aggregates and used directly, since gel filtration was not found to improve crystallisation. Crystals were grown in MRC SwissCi 3-well sitting-drop trays, and the well solution was 0.1 M Tris pH 8.5, 0.2 M sodium/potassium tartrate and 30% PEG3350. Drops were set at room temperature with 75 nL protein solution + 75 nL well solution, then trays were incubated at 4 ℃. Crystals appeared within one week.

Co-crystallisation conditions: FLuc was diluted to 1 mg/mL and incubated with ezutromid (40 µM, FAC 5% DMSO) plus of one of DLSA (40 µM), ATP (100 or 200 µM), luciferin

(40 or 100 µM) or nothing at 4 ℃ for 3 h. The protein was then concentrated to 10 mg/mL, reducing excess unbound substrates then seeded into MRC SwissCI 3-well sitting-drop trays as above and incubated at 4 ℃. Crystals appeared within one week. Well conditions of the crystals presented in Figure 3.9 are as follows:

Figure 3.9b: 0.8 M ammonium phosphate dibasic, 0.1 M acetate pH 4.5

Figure 3.9c: 0.1 M tris pH 8.5, 0.2 M magnesium chloride, 30% PEG4000

Figure 3.9d: 0.1 M tris pH 8.5, 0.3 M magnesium chloride, 7.5% PEG3350, 25% PEG400

Figure 3.9e: 0.15 M calcium chloride, 0.1 M acetate pH 4.5, 30% glycerol, 15% 2-propanol

Figure 3.9f: 0.2 M calcium acetate, 0.1 M cacodylate pH 6.5, 18% PEG8000

All crystals were harvested with 20% ethylene glycol as cryoprotection and flash-cooled in liquid nitrogen. All X-ray diffraction data were collected on the beamline I04-1 at the Diamond Light Source (Harwell, UK). Diffraction data were automatically processed by software pipelines at the Diamond Light Source[346] and refined using PHENIX.[233] Structures were solved by molecular replacement with Phaser[234] using structure 1LCI[235] as a model. Manual refinement to build missing residues was performed using Coot.[347]

FLuc diazirine irradiation study

20 µM diazirine probe **79*** was incubated with 10 µM recombinant FLuc in a buffer containing 25 mM HEPES, 5 mM $MgCl_2$, 1 mM EDTA, 5 mM DTT and 5% DMSO. Where applicable, 100 µM ATP, 50 µM luciferin or 20 µM DLSA were included. After 15 mins incubation at 0 °C, the samples were irradiated (365 nm, 100 W UV lamp, UVP™ B-100AP) for 3 mins at 0 °C from a distance of 6 cm. The samples were diluted 1:1 with milli-Q water and submitted for LC-MS analysis on a Waters Acquity 1525 µHPLC system coupled to a LCT Premier XE (Waters) mass spectrometer. A flow rate of 0.3-0.75 mL/min was applied to a mobile phase of A = water + 0.1% formic acid, B = MeCN with a gradient of % A: 0.0 min 97%, 1.0 min 97%, 5.0 min 2%, 6.0 min 2%, 7.0 min 97%, 10.0 min 97%. The electrospray source had a capillary voltage of 3.00 kV and cone voltage of 100 V. Nitrogen was the nebuliser and desolvation gas, flow of 300 L/hr. Total mass spectra were deconvoluted using the MaxEnt algorithm in MassLynx 4.1 (Waters) according to the default settings.

ESI MS/MS analysis was carried out by Elisabete Pires on an Acquity-UPLC system (Waters) connected to a Xevo G2-XS Q-TOF mass spectrometer (Waters) equipped with an electrospray ion source. The analyte separated on an ACE equivalent 3 C18 analytical column

(2.1 mm i.d. × 50 mm, 3 µm, 100 Å) using a linear gradient (length: 10 minutes, 5 % to 95 % solvent B (0.1% formic acid in acetonitrile)) at a flow rate: 0.4 mL/min (solvent A 0.1% formic acid in water). The separated peptides were electrosprayed directly into the mass spectrometer operating in a full scan method using a CID based method. A full scan MS spectrum was operated in electrospray positive mode with a scan range 100-900 m/z and a scan time 1.5 s. Lockspray was used during analysis to maintain mass accuracy. Data was obtained in continuum mode. The electrospray source of the MS was operated with a capillary voltage of 3.00 kV and a cone voltage of 40 V. CID fragmentation was performed at a ramp from 35 to 45 V of normalized collision energy. The MS and MS/MS analysis was processed using Masslynx. The mass window search of precursor ion was set as 5 ppm error. All spectra were manually checked and validated.

*Synthesised by Jess Reynolds.[178]

Utrophin RT-qPCR

6 well plates were seeded with H2K *mdx* cells (1×10^5) per well. After 24 h, cells were dosed with 30 nM query compounds from 10 mM solution stocks in DMSO (0.3% final DMSO concentration) in triplicate. RNA was extracted using TRIzol™ Reagent (Invitrogen) according to the manufacturer's instructions. 500 ng of RNA from each sample was used to generate cDNA with the QuantiTect Reverse Transcription kit (Qiagen). All RT-qPCR reactions were amplified using the StepOne™ Real-Time Polymerase chain reaction system (Applied Biosystems) with Fast SYBR™ Green Master Mix (Thermo Fisher). Utrophin isoform A (mouse: forward primer 5'-ACGAATTCAGTGACATCATTAAGTCC-3' reverse primer 5'-ATCCATTTGGTAAAGGTTTT CTTCTG-3') was normalised to reference S13 (mouse: forward primer 5'-CCCGAGGATCTCTACCATT-3' reverse primer 5'-GCCACTAGACAGAGGCTGT-3') and values obtained according to the $2^{-\Delta\Delta CT}$ method.[348]

Western blot assay

6 well plates were seeded with H2K *mdx* cells (1×10^5) per well. After 24 h, the cells were dosed with query compounds in three concentrations from 10 mM solution stocks in DMSO (0.3% final DMSO concentration) in triplicate. Heregulin at 30 nM was used as a positive control. After 24 h, the cells were harvested (TrypLE Express, Gibco), washed (PBS) and lysed (RIPA buffer and protease inhibitors, Sigma #P8340). Protein content was quantified by a bicinchinonic acid protein assay (Thermo Scientific Pierce). For the western blot, 30

µg lysate was separated by NuPAGE™ 3-8% Tris-Acetate protein gel electrophoresis and transferred to a PVDF membrane (GE Healthcare). Utrophin protein was detected using Mancho-3 antibody (1: 50, kind gift from G.E. Morris, Oswestry, UK) and an AlexaFluor™ 680 anti-mouse antibody (1:10,000, Invitrogen). Blots were imaged with a Licor Odyssey system and relative protein quantitation was performed using Image Studio Lite. REVERT™ total protein stain (Licor) was used as a loading control, along with β-actin detected with anti-β-actin antibody (1: 20,000, Cell Signaling Technology #3700S) and an AlexaFluor™ 680 anti-mouse antibody (1:10,000, Invitrogen). Experiments were carried out at least three times.

Chemoproteomics workflow of immobilised compound affinity capture followed by LC-MS/MS analysis

H2K *mdx* cells were lysed in a buffer containing 25 mM HEPES, 150 mM NaCl, 2 mM $MgCl_2$, 0.1 % Tween-20 and protease inhibitors (0 ℃, 30 min). The lysates were sonicated (4 × 2 s on, 3 s off, 20% amplitude) and centrifuged to pellet cell debris. Protein concentrations were determined by a bicinchinonic acid protein assay (Thermo Scientific Pierce), then adjusted to 2 mg/mL. Streptavidin beads (55 µL per sample, Dynabeads™ MyOne™ T1, Invitrogen) were washed three times (HEPES lysis buffer) and then incubated (1 h, rt, 1000 rpm) with lysate (400 µL per sample) to preclear non-specifically binding proteins. Additional streptavidin beads (15 µL per sample, Dynabeads™ MyOne™ T1, Invitrogen) were washed three times (HEPES lysis buffer) then incubated (1 h, rt, 1000 rpm) with one of the following biotinylated probes: **41**, **42**, **75**, **43** (each 10 µM, 0.3% FAC of DMSO). The beads were then washed three times (HEPES lysis buffer) and added to the precleared lysate with and without competitor compound SMT022332 (100 µM, 0.3% FAC of DMSO) for 16 h at 4 ℃, 1000 rpm. The lysate was then removed and the beads washed three times (HEPES lysis buffer).

On-bead digestion of pulled down proteins was achieved using a Filter Aided Sample Preparation (FASP) protocol.[349] Briefly, Vivacon 500 filters (Sartorius, VN01H02 10 kDa/VNCT01) were washed with 0.1% trifluoroacetic acid in 50% acetonitrile. The beads were loaded on the filter in 8 M urea in 100 mM AB for 30 minutes at rt. On-bead proteins were reduced (10 mM TCEP, 30 minutes, rt), alkylated (50 mM chloroacetamide, 30 min, rt in the dark) and washed (2 × 1 M urea in 50 mM AB). The proteins were subjected to tryptic digestion (0.2 µg enzyme, Promega, 1 M urea in 50 mM AB) overnight at 37°C. Trypsinised peptides collected from the filtrate were dried then desalted using C18 stage tips. Eluted peptides were then

dried and resuspended in 50 µL 5% formic acid and 5% DMSO.

LC-MS/MS analysis was carried out on an Ultimate 3000 ultra-HPLC system (Thermo Fisher) coupled to a QExactive mass spectrometer (Thermo Fisher). The peptides were trapped on a C18 PepMap100 pre-column (300 µm i.d. x 5 mm, 100 Å, Thermo Fisher) using solvent A (0.1% formic acid in water) at a pressure of 500 bar, then separated on an in-house packed analytical column (75 µm i.d. packed with ReproSil-Pur 120 C18-AQ, 1.9 µm, 120 Å, Dr. Maisch GmbH) using a linear gradient (length: 60 minutes, 15% to 35% solvent B (0.1% formic acid in acetonitrile), flow rate: 200 nl/min). Data were acquired in a data-dependent mode (DDA). Full scan MS spectra were acquired in the Orbitrap (scan range 350-1500 m/z, resolution 70000, AGC target 3×10^6, maximum injection time 50 ms). The 10 most intense peaks were selected for HCD fragmentation at 30% of normalised collision energy (resolution 17500, AGC target 5×10^4, maximum injection time 120 ms) with first fixed mass at 180 *m/z*.

Peptide identification and quantification were performed by MaxQuant (version 1.5.0.35i).[350] MS spectra were searched against the *Mus musculus* UniProt Reference proteome (retrieved 12/01/17) alongside a list of common contaminants. The search results were filtered to a 1% false discovery rate (FDR) for proteins, peptides and peptide-spectrum matches (PSM). Protein intensity distributions were log2 transformed using Perseus (version 1.5.5.3) and missing values were imputed with an value lower than the smallest measured intensity, to approximate to background noise. The three replicates for each condition were grouped and Student's *t*-test (s0 = 0, FDR = 0.05) performed between the active probe samples and the two controls (inactive probe and competitor). Results of this analysis were plotted using Python plotting library Matplotlib.

Chemoproteomics workflow of *in situ* labelling followed by LC-MS/MS analysis

H2K *mdx* cells were seeded in 12 × 10 cm diameter dishes and grown to 80% confluency. Cells were treated with 3 µM probe **106** (with and without 100 µM competitor SMT022332), 3 µM probe **129** and the DMSO vehicle, in triplicate for each condition, to a final DMSO concentration of 0.3% in serum-free media, for 2 h. The cells were then washed with 1 × PBS and irradiated (3 min, 5 cm distance, 365 nm 100W lamp, VWR 36595-021) in serum-free media. The cells were washed with 1 × PBS then lysed for 10 min at 4 °C with 300 µL of a buffer containing 1% SDS, 1% triton X-100 and protease inhibitors (Sigma #P8340) in 1

× PBS. The dishes were scraped to collect the lysate, which was sonicated (4 × 2 s on, 3 s off, 20% amplitude) and the total protein concentration determined by a bicinchinonic acid protein assay (Thermo Scientific Pierce). 350 µg of lysate at 0.8 mg/mL concentration per sample was then Clicked to biotin-azide (FAC of 0.1 mM from 10 mM in DMSO stock) with $CuSO_4$ (FAC of 1 mM, from 50 mM in H_2O stock), TCEP (FAC of 1 mM, from 50 mM in H_2O stock) and TBTA (FAC of 0.1 mM, from 10 mM in DMSO stock) for 1 h at rt, 1000 rpm shaking.[259] Excess Click reagents were removed by protein precipitation ($MeOH/CHCl_3$) and washing of the protein pellet (MeOH × 4). The proteins were then solubilized with 2% SDS in PBS before dilution to 0.5 mg/mL protein, 0.5% SDS using 1 × PBS. Streptavidin beads (15 µL per sample, Dynabeads™ MyOne™ T1, Invitrogen) were washed twice (1 × PBS) then added to the labelled lysate samples. After 16 h rocking (65 rpm) at 4 °C, the beads were washed (3 × 1% SDS, 1% triton X-100 in PBS, 3 × 4 M urea in 50 mM ammonium bicarbonate buffer (AB), 3 × 6 M urea in 50 mM AB).

Sample preparation, LC-MS/MS analysis and peptide identification and quantification were performed as above.

Chemoproteomics followed by AhR immunoblotting

Human DMD cells were seeded in 10 cm diameter dishes and grown to 80% confluency. Cells were treated with 3 µM probe **106** (with and without 100 µM competitor SMT022332, 3 µM probe **129** and the DMSO vehicle, in duplicate for each condition, to a final DMSO concentration of 1% in serum-free media, for 2 h. The cells were then irradiated and lysed as above. 500 µg of lysate at 1 mg/mL concentration per sample was then Clicked to TAMRA-biotin-azide (FAC of 0.1 mM from 10 mM in DMSO stock) as above. Excess Click reagents were removed by protein precipitation ($MeOH/CHCl_3$) and washing of the protein pellet (MeOH × 4). The proteins were then solubilized with 2% SDS in PBS before dilution to 1 mg/mL protein, 0.2% SDS using 1 × PBS. Streptavidin beads (30 µL per sample, Dynabeads™ MyOne™ T1, Invitrogen) were washed twice (1 × PBS) then added to the labelled lysate samples. After 16 h rocking (65 rpm) at 4 °C, the beads were washed as above. Proteins were eluted from the beads by heating at 95 °C for 5 min in 1 × Laemmli buffer, 50 mM DTT, 8 M urea in 50 mM AB. The eluted proteins were separated by NuPAGE™ 4-12% Bis-Tris protein gel electrophoresis and transferred to a PVDF membrane (Roche). AhR protein was detected with an anti-AhR antibody (1:5000 in 2% BLOT-QuickBlocker™ (#WB57) solution in PBST, Enzo BML-SA210) followed by an IRDye® 800CW anti-rabbit antibody (1: 5000,

Licor). Blots were imaged with a Licor Odyssey system.

Protein expression and purification of recombinant truncated AhR:ARNT complex

This procedure was carried out by Prof. Fraydoon Rastinejad.

The mAHR (residues 25-433) and mARNT (residues 82-464) proteins were co-expressed in *E. coli*, using an expression and purification strategy similar to the purification of related ARNT heterdimers (HIF-ARNT, NPAS1-ARNT, NPAS3-ARNT) described previously.[351,352] For AHR-ARNT, the ARNT protein construct was expressed using a pMKH vector which did not provide any affinity tag while the AHR protein was expressed simultaneously in the same cells (BL21-CodonPlus DE3-RIL competent cells, from Agilent Technologies, Santa Clara, CA, #230245) using the PSJ2 vector which produced a C-terminal histidine tagged protein. Bacterial cultures were grown at 37℃, induced at 16℃ overnight using 0.2 mM IPTG. Cells were then harvested, lysed using sonication, centrifuged to remove the pellet, and the supernatant applied to a Ni-NTA resin following the manufacturer's recommended procedure. After extensive washing with buffer, the protein complex was eluted using buffer containing 250 mM imidazole, and the heterodimer further purified using a 1-meter Superdex-200) gel filtration column, which eluted at the expected position of the heterodimer. SDS-PAGE gels stained with coommassie blue showed both subunits were co-purified in 1:1 stoichiometric amounts.

AhR:ARNT fluorescence quenching assay

Fluorescence quenching of ezutromid upon binding recombinant AhR:ARNT was monitored using a fluorescence microplate reader (Tecan Spark, 313 nm excitation/390 nm emission). 100 nM compound was incubated with AhR:ARNT (various concentrations ranging $0 - 10$ µM) for 18 h (4 °C) in black 96 well plates, with a buffer containing 20 mM Tris pH 8, 200 mM NaCl and 1% DMSO. Fluorescence emission was measured and the signal background from the protein alone ($0 - 10$ µM) subtracted. The K_D calculated by fitting the curve with a four parameter logistic function.

AhR localisation immunofluorescence

Human DMD cells (3×10^4) were seeded on 18 mm coverslips. After 24 h, the cells were dosed with ezutromid (3 µM), ITE (1 µM) or both (from 10 mM solution stocks in DMSO, 1% final DMSO concentration) in serum-free media alongside a vehicle control. After 2 h, the cells

were fixed (4% paraformaldehyde in PBS, 15 min), permeabilised (0.1% Triton X-100 in PBS) and blocked (10% FBS, 0.1% Tween-20 in PBS, 1 h). The cells were probed with anti-AhR antibody (1:200, Invitrogen, #MA1-514), anti-mouse AlexaFluor™ 488 (1:2000, #A-11059, Invitrogen) and Hoescht 33342 (10 µM, Sigma #14533). The coverslips were mounted onto slides with FluorSave™ reagent. The cells were imaged using an Evos fluorescence microscope with a 40 × objective. Immunofluorescence experiments were carried out three times.

Statistics

Differences between group means were calculated by unpaired, two-tailed t-tests where p values ≤ 0.05 were considered statistically significant.

8.2 Chemistry Experimental Methods

8.2.1 General Chemical Experimental Procedures

Reactions were carried out under inert conditions using an atmosphere of nitrogen, with anhydrous solvents and oven dried glassware, unless aqueous reagents were used or otherwise stated. Microwave reactions were carried out using a Biotage® Initiator Classic microwave synthesiser.

Anhydrous solvents were dried by passing over an activated alumina column, under an inert atmosphere, using a solvent purification system. Water was purified by an Elix® UV-10 system. All other solvents and reagents were used as supplied (analytical or HPLC grade) without further purification. PE refers to the fraction of petroleum ether boiling in the range 30-40 ℃. Aqueous solutions of ammonium chloride (NH_4Cl), sodium carbonate (Na_2CO_3), sodium thiosulfate ($Na_2S_2O_3$) and sodium bicarbonate ($NaHCO_3$) were saturated. Concentration of solvents *in vacuo* was achieved by rotary evaporation using a diaphragm pump. Purification by flash column chromatography was carried out using Kieselgel 60 and a positive solvent pressure. Reverse phase chromatography was carried out on C18 silica using a Biotage® Isolera™ flash system.

8.2.2 Analytical Procedures

LC-MS sample analysis was performed using an Agilent 1260 series machine equipped with a Agilent Poroshell 120 column, EC-C18 2.7 µm, 50 × 4.6 mm. A flow rate of 1.35 mL/min was applied to a mobile phase of A = water + 0.1% formic acid, B = MeCN + 0.1% formic acid at 40 ℃ with a gradient of %B: 0.00 min 20%, 2.45 min 98%, 4.90 min 98%, 5.00 min 20%, 6.00 min 20%. TLC was carried out using Merck Kieselgel 60 F254 plates which were visualised with UV light (254 nm). Melting points were recorded on a EZ-Melt automated melting point apparatus. Fourier transform IR spectroscopy was carried out using a Bruker Tensor 27 FT-IR spectrometer as neat samples or thin films. Wavelengths of peak absorption are given in wavenumbers (cm^{-1}), with broad (br) signals indicated.

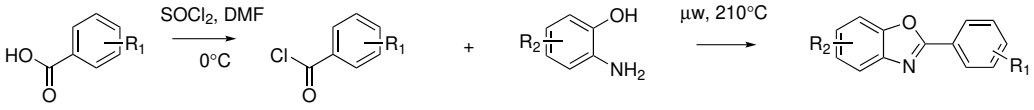
NMR spectra were recorded using Bruker Advance spectrometers (AVII 400 or AVII 500). ^1H NMR and ^{13}C NMR data were recorded at 298 K, locked to the relevant solvent standard. NMR data are presented as: chemical shift δ (in ppm, $\delta_{TMS} = 0$), integration, multiplicity (s = singlet, d = doublet, t = triplet, q =quartet, m = multiplet, br = broad, obs = obscured) and coupling constants (J in Hz). Proton peak assignments were based on 1D data and COSY

analysis and carbon peak assignments were based on 1D data and HSQC and HMBC analysis. Low resolution mass spectroscopy was carried out on an Agilent 6120 mass spectrometer. Accurate mass measurements were run on either a Bruker MicroTOF internally calibrated with polyalanine, or a Micromass GCT instrument fitted with a Scientific Glass Instruments BPX5 column (15 m × 0.25 mm) using amyl acetate as a lock mass. HR ESI and APCI were run on a Waters Acquity Ultraperformance LC system coupled to a Thermo Orbitrap Exactive MS.

8.2.3 Chemical Experimental Procedures and Data

General Procedure A

Benzoxazole ring closure from a benzoic acid and 2-aminophenol derivative

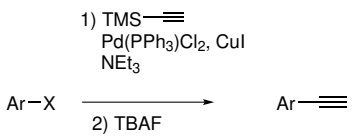


To a suspension of the benzoic acid derivative (1 eq.) in CH_2Cl_2 at 0 ℃ was added thionyl chloride (3 eq.) and dimethylformamide (cat.) The reaction mixture was warmed to rt and stirred for 16 h. When the reaction was complete, the reaction mixture was concentrated *in vacuo*, dissolved in Et_2O and filtered. The filtrate was concentrated *in vacuo* to yield the acid chloride, which was used immediately without further purification.

To a suspension of the 2-aminophenol derivative (1 eq.) in 1,4-dioxane was added a solution of the prepared acid chloride (1 eq.) in 1,4-dioxane at rt. The reaction vessel was heated under microwave activation at 210 ℃ for 15 min, then cooled to rt. The reaction mixture was poured into NaOH solution (1 M aq.), and extracted with EtOAc three times. The combined organic fractions were dried (Na_2SO_4) and concentrated *in vacuo* to yield the crude benzoxazole product.

General procedure B

Sonagashira coupling of aryl halide with TMS-acetylene

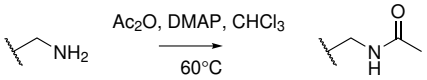


To a solution of aryl halide (1 eq.), $Pd(PPh_3)_2Cl_2$ (0.15 eq.) and copper(I) iodide (0.1 eq.) in THF was added NEt_3 (excess) and trimethylsilylacetylene (1 eq.) dropwise. The

reaction mixture was degassed and stirred at 70 ℃ until the reaction was complete by TLC. The solvent was removed *in vacuo* and the residue purified by flash column chromatography. The isolated TMS-alkyne intermediate was dissolved in THF and TBAF (1.5 eq., 1 M in THF) was added dropwise. The reaction was stirred at rt until completion. The reaction mixture was then diluted with EtOAc, washed with water and the phases separated. The organic layer was dried (Na_2SO_4) and concentrated *in vacuo* to yield the crude alkyne product.

General procedure C

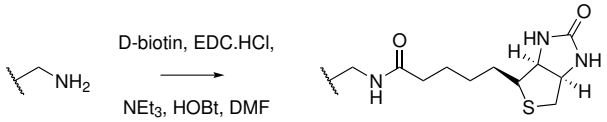
Formation of acetamides from primary amines



To a solution of primary amine (1 eq.), 4-dimethylaminopyridine (0.1 eq.) and acetic anhydride (1.1 eq.) in $CHCl_3$ was added NEt_3 (1.1 eq.). The reaction mixture was heated under reflux for 3 h, then cooled to rt and concentrated *in vacuo* to yield the crude acetamide product.

General procedure D

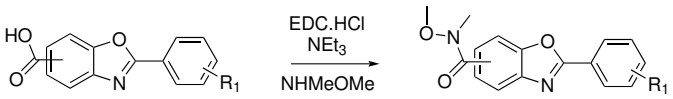
Amide coupling with D-biotin



A solution of primary amine (1 eq.), D-biotin (2 eq.), EDC hydrochloride (2 eq.), HOBt hydrate (0.5 eq.), NEt_3 (2 eq.) in DMF was stirred at rt for 16 h. The reaction mixture was then concentrated *in vacuo* to yield the crude biotinylated product.

General procedure E

Formation of benzoxazole carboxamides

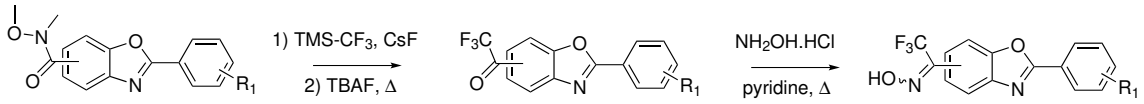


Acid benzoxazole derivatives were obtained by following General Procedure A, with no further purification. To a solution of the crude prepared acid (1 eq.) in DMF was added EDC hydrochloride (2 eq.), *N,O*-dimethylhydroxylamine hydrochloride (2 eq.) and NEt_3 (2 eq.).

The reaction mixture was stirred at rt for 16 h, then concentrated *in vacuo*. The residue was dissolved in EtOAc, then washed with citric acid solution (10% aq.), $NaHCO_3$ solution (10% aq.) and brine. The organic layer was concentrated *in vacuo* to yield the crude carboxamide product.

General procedure F

Formation of trifluoromethylketone oximes from carboxamides

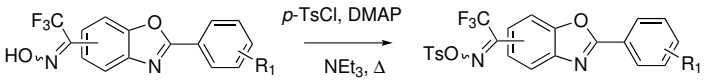


Adapting a known procedure,[183] caesium fluoride (0.2 eq.) and trifluoromethyltrimethylsilane (2 eq.) were added to a solution of Weinreb amide derivative (1 eq.) in toluene. The reaction mixture was stirred at rt with further equivalents of caesium fluoride and trifluoromethyltrimethylsilane added every 24 h until the reaction was complete by TLC. At this point, the reaction was concentrated *in vacuo* and the residue dissolved in THF. TBAF (1.2 eq., 1 M in THF) was added dropwise and the reaction stirred at rt for 3 h at 50 °C. The reaction was cooled to rt, diluted with Et_2O and washed with water and brine. The organic layer was dried (Na_2SO_4) and concentrated *in vacuo* to yield the crude trifluoromethyl ketone product. The crude was filtered through a silica plug, eluting with PE : EtOAc (2:1), and concentrated *in vacuo*.

To a solution of the prepared crude trifluoromethylketone derivative (1 eq.) in pyridine (excess) and ethanol over molecular sieves (3 Å) was added hydroxylammonium chloride (3 eq.). The reaction mixture was heated under reflux for 16 h, then concentrated *in vacuo*. The residue was dissolved in EtOAc, washed with brine/water (1:1), dried (Na_2SO_4) and concentrated *in vacuo* to yield the crude oxime product.

General procedure G

Formation of trifluoromethylketone tosyl oximes

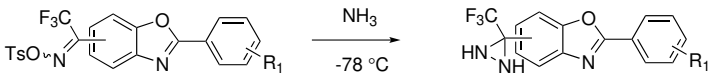


To a solution of oxime (1 eq.), 4-dimethylaminopyridine (0.1 eq.) and NEt_3 (3 eq.) in CH_2Cl_2 was added *p*-toluenesulfonyl chloride (2 eq., recrystallised from PE) portionwise. The reaction mixture was stirred at 40 °C for 16 h, then cooled to rt. The mixture was diluted

with CH_2Cl_2, washed with water then brine, dried (Na_2SO_4) and concentrated *in vacuo* to yield the crude tosyl oxime products as a mixture of E/Z isomers.

General procedure H

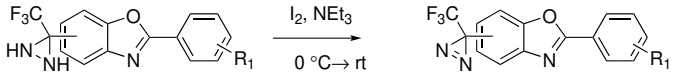
Formation of diaziridines



NH_3 (excess) was condensed into a solution of tosyl oxime (1 eq.) in CH_2Cl_2 at −78 ℃. The reaction mixture was stirred and warmed to rt over 6 h, then concentrated *in vacuo* to yield the crude diaziridine product.

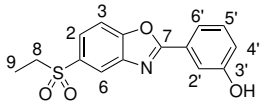
General procedure I

Formation of diazirines



To a solution of diaziridine (1 eq.) and NEt_3 (3 eq.) in CH_2Cl_2 was added iodine (1.2 eq.) portionwise. The reaction mixture was covered in foil to exclude light and stirred for 15 min at rt. The mixture was then concentrated *in vacuo* to yield the crude diazirine product.

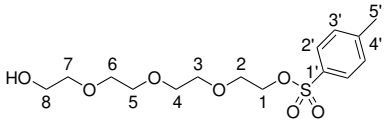
3-(5-(Ethylsulfonyl)benzo[*d*]oxazol-2-yl)phenol 49



Benzoxazole **49** was obtained according to General Procedure A, using 3-hydroxybenzoic acid (350 mg, 2.53 mmol), thionyl chloride (0.55 mL, 7.59 mmol) and DMF (0.10 mL) in CH_2Cl_2 (15 mL). The acid chloride intermediate formed was reacted with 2-amino-4-ethylsulfonylphenol (510 mg, 2.53 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 3:2) to yield benzoxazole **49** (250 mg, 0.824 mmol, 33%) as an off-white solid.

R_f 0.23 (PE : EtOAc, 3:2); **ν_{max}** (cm^{-1}) 3649 (-OH), 2981 (C-H), 1473, 1305 (S=O), 1137 (S=O); **mp** 192-193 °C; **^{1}HNMR** (500 MHz, $CO(CD_3)_2$) δ_H 8.26 (1H, dd, J = 1.6, 0.8 Hz, H_6), 8.00-7.96 (2H, *obs* m, H_2 and H_3), 7.80-7.76 (2H, *obs* m, $H_{2'}$ and $H_{6'}$), 7.47 (1H, ddd, J = 8.2, 7.5, 0.6 Hz, $H_{5'}$), 7.15 (1H, ddd, J = 8.2, 2.5, 1.0 Hz, $H_{4'}$), 3.29 (2H, q, J = 7.4 Hz, H_8), 1.23 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CO(CD_3)_2$) δ_C 165.9 (C_7), 158.8 ($C_{3'}$), 154.5 (C_4), 143.4 (C_5), 137.0 (C_1), 131.4 ($C_{5'}$), 128.4 ($C_{1'}$), 126.4 (C_2), 121.3 (C_6), 120.5 ($C_{4'}$), 119.9 ($C_{6'}$), 115.1 ($C_{2'}$), 112.4 (C_3), 51.1 (C_8), 7.8 (C_9); **HRMS** (ES^+) Calc. for $C_{15}H_{14}NO_4S$ $[M+H]^+$ 304.0638, found 304.0633.

2-(2-(2-(2-Hydroxyethoxy)ethoxy)ethoxy)ethyl 4-methylbenzenesulfonate 45

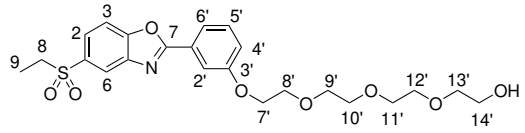


To a solution of tetraethylene glycol (18.1 mL, 105 mmol) in CH_2Cl_2(25 mL) at 0 ℃ was added *p*-toluenesulfonyl chloride (2.00 g, 10.5 mmol) and NEt_3 (2.18 mL, 15.7 mmol). The reaction mixture was warmed to rt, stirred for 18 h then quenched with H_2O (10 mL). The aqueous phase was extracted with CH_2Cl_2 and the combined organic fractions dried (Na_2SO_4) and concentrated *in vacuo*. The crude product was purified by flash column chromatography (PE : EtOAc, 2:3 → 0:1) to yield alcohol **45** (2.69 g, 7.71 mmol, 73%) as a colourless oil.

R_f 0.17 (EtOAc : CH_2Cl_2, 4:1); **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 7.80 (2H, d, J = 8.3 Hz, $H_{2'}$), 7.34 (2H, d, J = 8.1 Hz, $H_{3'}$), 4.16 (2H, dd, J = 5.7, 4.0 Hz, H_1), 3.74-3.62 (14H, m, $H_{2\text{-}8}$), 3.61 (3H, s, $H_{5'}$); **LRMS** (ES^+) Calc. for $C_{15}H_{24}O_7S$ $[M+H]^+$ 349.1, found 349.1.

These data are in accordance with that of Heller *et al.*[353]

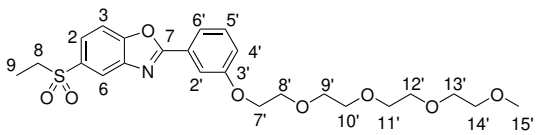
2-(2-(2-(2-(3-(5-(Ethylsulfonyl)benzo[*d*]oxazol-2-yl)phenoxy)ethoxy)ethoxy)-ethoxy)ethan-1-ol 50



To a solution of **49** (100 mg, 0.330 mmol) and K_2CO_3 (91.2 mg, 0.660 mmol) in DMF (10 mL) was added alcohol **45**. The reaction mixture was heated to 80 ℃ and stirred for 4 h. After completion of the reaction, water (5 mL) was added and the mixture extracted with EtOAc (3 × 10 mL). The combined aqueous layers were dried (Na_2SO_4) and concentrated *in vacuo*. The crude product was purified by flash column chromatography (CH_2Cl_2 : EtOAc, 1:3 → 0:1) to yield benzoxazole **50** (115 mg, 0.239 mmol, 80%) as a yellow oil.

R_f 0.13 (EtOAc); **ν_{max}** (cm^{-1}) 3650 (-OH), 2980 (C-H), 2888 (C-H), 1473, 1423, 1303 (S=O), 1138 (S=O), 1071 (C-O); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.32 (1H, d, J = 1.7 Hz, H_6), 7.93 (1H, dd, J = 8.5, 1.7 Hz, H_2), 7.85 (1H, dt, J = 8.0, 1.2 Hz, $H_{6'}$), 7.80 (1H, t, J = 1.2 Hz, $H_{2'}$), 7.74 (1H, d, J = 8.5 Hz, H_3), 7.44 (1H, t, J = 8.0 Hz, $H_{5'}$), 7.16 (1H, dt, J = 8.0, 1.2 Hz, $H_{4'}$), 4.26 (2H, t, J = 4.8 Hz, $H_{7'}$), 3.91 (2H, t, J = 4.8 Hz, $H_{8'}$), 3.77-3.74 (2H, m, $H_{9'}$), 3.73-3.69 (4H, m, $H_{10'}$ and $H_{14'}$), 3.67 (4H, s, $H_{11'}$ and $H_{12'}$), 3.62-3.59 (2H, m, $H_{13'}$), 3.17 (2H, q, J = 7.4 Hz, H_8), 1.30 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 165.4 (C_7), 159.3 ($C_{3'}$), 153.9 (C_4), 142.8 (C_5), 135.3 (C_1), 130.4 ($C_{5'}$), 127.3 ($C_{1'}$), 125.5 (C_2), 121.1 (C_6), 120.8 ($C_{6'}$), 119.8 ($C_{4'}$), 113.3 ($C_{2'}$), 111.5 (C_3), 72.6 ($C_{13'}$), 71.0 ($C_{9'}$), 70.8 ($C_{11'}$ or $C_{12'}$), 70.7 ($C_{11'}$ or $C_{12'}$), 70.4 ($C_{10'}$), 69.8 ($C_{8'}$), 67.9 ($C_{7'}$), 61.9 ($C_{14'}$), 51.1 (C_8), 7.7 (C_9); **HRMS** (ES^+) Calc. for $C_{23}H_{30}NO_8S$ $[M+H]^+$ 480.1687, found 480.1683.

2-(3-((2,5,8,11-Tetraoxatridecan-13-yl)oxy)phenyl)-5-(ethylsulfonyl)benzo[*d*]-oxazole 51



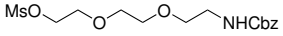
To a solution of **50** (50.0 mg, 0.104 mmol) in THF (3 mL) at 0 ℃ was added sodium hydride (10.0 mg, 60% dispersion in mineral oil, 0.115 mmol). After 30 min, methyl iodide (9 µL, 0.156 mmol) was added dropwise, and the reaction warmed to rt. After reaction completion, the mixture was cooled to 0 ℃ and quenched with water (1 mL) dropwise. The mixture was extracted with EtOAc (2 × 2 mL) and the combined organic fractions washed

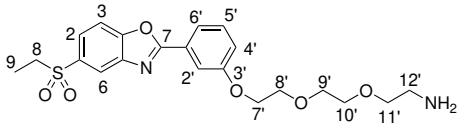
with water (3 mL) and concentrated *in vacuo*. The crude product was purified by flash column chromatography (PE : EtOAc, 9:1) to yield benzoxazole **51** (49.0 mg, 99.3 µmol, 95%) as a yellow oil.

R_f 0.23 (EtOAc : CH_2Cl_2, 9:1); **ν_{max}** (cm^{-1}) 2980 (C-H), 2888 (C-H), 1473, 1382 (S=O), 1251, 1150 (S=O), 1073 (C-O); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.32 (1H, d, J = 1.7 Hz, H_6), 7.93 (1H, dd, J = 8.5, 1.7 Hz, H_2), 7.86 (1H, ddd, J = 8.0, 1.2, 1.2 Hz, $H_{6'}$), 7.80 (1H, dd, J = 2.6, 1.5 Hz, $H_{2'}$), 7.74 (1H, d, J = 8.5 Hz, H_3), 7.45 (1H, dd, J = 8.0, 8.0 Hz, $H_{5'}$), 7.16 (1H, ddd, J = 8.0, 2.6, 1.0 Hz, $H_{4'}$), 4.27-4.24 (2H, m, $H_{7'}$), 3.93-3.90 (2H, m, $H_{8'}$), 3.77-3.73 (2H, m, $H_{9'}$), 3.72-3.62 (8H, m, $H_{10'}$, $H_{11'}$, $H_{12'}$ and $H_{13'}$), 3.55-3.52 (2H, m, $H_{14'}$), 3.36 (3H, s, $H_{15'}$), 3.18 (2H, q, J = 7.4 Hz, H_8), 1.30 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 165.4 (C_7), 159.4 ($C_{3'}$), 153.9 (C_4), 142.8 (C_5), 135.3 (C_1), 130.4 ($C_{5'}$), 127.4 ($C_{1'}$), 125.5 (C_2), 121.1 (C_6), 120.8 ($C_{6'}$), 119.8 ($C_{4'}$), 113.3 ($C_{2'}$), 111.5 (C_3), 72.1 ($C_{14'}$), 71.0 ($C_{9'}$), 70.8 ($H_{10'}$, $H_{11'}$, $H_{12'}$ or $H_{13'}$), 70.8 ($H_{10'}$, $H_{11'}$, $H_{12'}$ or $H_{13'}$), 70.7 ($H_{10'}$, $H_{11'}$, $H_{12'}$ or $H_{13'}$), 70.6 ($H_{10'}$, $H_{11'}$, $H_{12'}$ or $H_{13'}$), 69.8 ($C_{8'}$), 67.9 ($C_{7'}$), 59.2 ($C_{15'}$), 51.2 (C_8), 7.8 (C_9); **HRMS** (ES^+) Calc. for $C_{24}H_{32}NO_8S$ $[M+H]^+$ 494.1843, found 494.1839.

2-(2-(2-(3-(5-(Ethylsulfonyl)benzo[*d*]oxazol-2-yl)phenoxy)ethoxy)ethoxy)ethan-1-amine 54

53





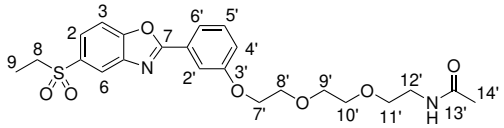
A solution of benzoxazole **49** (75.0 mg, 0.247 mmol), protected linker **53*** (89.3 mg, 0.247 mmol) and K_2CO_3 (68.3 mg, 0.494 mmol) in DMF (10 mL) was stirred at 80 ℃ for 16 h. The reaction mixture was then quenched with water (5 mL) and extracted with EtOAc (2 × 10 mL). The combined organic fractions were dried (Na_2SO_4) and concentrated *in vacuo*. The crude mixture was purified by flash column chromatography (CH_2Cl_2 : MeOH, 99:1 → 49:1). To a solution of the protected amine intermediate in MeOH (10 mL) under an argon atmosphere was added Pd/C (13.1 mg, 10% wt., 0.124 mmol). The reaction mixture was stirred under H_2 at rt for 20 h, then filtered through Celite, eluting with MeOH. The filtrate was concentrated *in vacuo* to yield amine **54** (65.0 mg, 0.151 mmol, 57%) as a yellow oil.

*Linker compound kindly supplied by Aini Vuorinen.[10]

ν_{max} (cm^{-1}) 2886 (C-H), 1553, 1424, 1301 (S=O), 1137 (S=O), 1071 (C-O); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.33 (1H, d, J = 1.7 Hz, H_6), 7.94 (1H, dd, J = 8.5, 1.7 Hz, H_2),

7.87 (1H, ddd, J = 8.0, 1.2, 1.2 Hz, $H_{6'}$), 7.81 (1H, dd, J = 2.6, 1.5 Hz, $H_{2'}$), 7.75 (1H, d, J = 8.5 Hz, H_3), 7.46 (1H, dd, J = 8.0, 8.0 Hz, $H_{5'}$), 7.17 (1H, ddd, J = 8.0, 2.6, 1.0 Hz, $H_{4'}$), 4.28-4.25 (2H, m, $H_{7'}$), 3.94-3.91 (2H, m, $H_{8'}$), 3.78-3.75 (2H, m, $H_{9'}$), 3.69-3.66 (2H, m, $H_{10'}$), 3.53 (2H, t, J = 5.2 Hz, $H_{11'}$), 3.18 (2H, q, J = 7.4 Hz, H_8), 2.88 (2H, t, J = 5.2 Hz, $H_{12'}$), 1.88 (2H, *br* s, -NH_2), 1.31 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 165.4 (C_7), 159.4 ($C_{3'}$), 153.9 (C_4), 142.8 (C_5), 135.4 (C_1), 130.4 ($C_{5'}$), 127.4 ($C_{1'}$), 125.5 (C_2), 121.2 (C_6), 120.8 ($C_{6'}$), 119.8 ($C_{4'}$), 113.3 ($C_{2'}$), 111.5 (C_3), 73.6 ($C_{11'}$), 71.0 ($C_{9'}$), 70.5 ($C_{10'}$), 69.8 ($C_{8'}$), 68.0 ($C_{7'}$), 51.2 (C_8), 41.9 ($C_{12'}$), 7.8 (C_9); **HRMS** (ES^+) Calc. for $C_{21}H_{27}N_2O_6S$ $[M+H]^+$ 435.1584, found 435.1591.

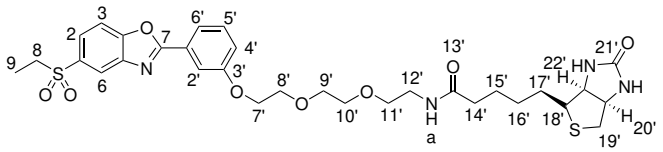
N-(2-(2-(2-(3-(5-(ethylsulfonyl)benzo[*d*]oxazol-2-yl)phenoxy)ethoxy)ethoxy)-ethyl)acetamide 52



Acetamide **52** was obtained according to General Procedure C, using amine **54** (20.0 mg, 46.1 µmol), 4-dimethylaminopyridine (0.563 mg, 4.61 µmol), acetic anhydride (5 µL, 50.7 µmol) and NEt_3 (7 µL, 50.7 µmol) in $CHCl_3$ (1 mL). The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 97:3) to yield acetamide **52** (21.9 mg, 46.0 µmol, quant.) as a white foam.

R_f 0.23 (CH_2Cl_2 : MeOH, 97:3); **ν_{max}** (cm^{-1}) 2923 (C-H), 1654 ((C=O)NH), 1553, 1454, 1301 (S=O), 1153 (S=O), 1048 (C-O); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.32 (1H, d, J = 1.7 Hz, H_6), 7.93 (1H, dd, J = 8.5, 1.7 Hz, H_2), 7.86 (1H, ddd, J = 8.0, 1.2, 1.2 Hz, $H_{6'}$), 7.80 (1H, dd, J = 2.7, 1.5 Hz, $H_{2'}$), 7.74 (1H, d, J = 8.5 Hz, H_3), 7.46 (1H, dd, J = 8.0, 8.0 Hz, $H_{5'}$), 7.15 (1H, ddd, J = 8.0, 2.6, 1.0 Hz, $H_{4'}$), 6.06 (1H, *br* s, -N*H*CO), 4.28-4.25 (2H, m, $H_{7'}$), 3.93-3.90 (2H, m, $H_{8'}$), 3.76-3.73 (2H, m, $H_{9'}$), 3.69-3.66 (2H, m, $H_{10'}$), 3.59-3.56 (2H, m, $H_{11'}$), 3.45 (2H, m, $H_{12'}$), 3.18 (2H, q, J = 7.4 Hz, H_8), 1.95 (3H, s, $H_{14'}$), 1.30 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 170.3 ($C_{13'}$), 165.3 (C_7), 159.3 ($C_{3'}$), 153.9 (C_4), 142.7 (C_5), 135.4 (C_1), 130.4 ($C_{5'}$), 127.4 ($C_{1'}$), 121.1 (C_6), 120.9 ($C_{6'}$), 119.8 ($C_{4'}$), 113.2 ($C_{2'}$), 111.6 (C_3), 70.9 ($C_{9'}$), 70.4 ($C_{10'}$), 70.0 ($C_{11'}$), 69.8 ($C_{8'}$), 67.9 ($C_{7'}$), 51.2 (C_8), 39.4 ($C_{12'}$), 23.4 ($C_{14'}$), 7.8 (C_9); **HRMS** (ES^+) Calc. for $C_{23}H_{29}N_2O_7S$ $[M+H]^+$ 477.1690, found 477.1689.

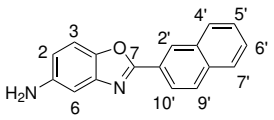
***N*-(2-(2-(2-(3-(5-(ethylsulfonyl)benzo[d]oxazol-2-yl)phenoxy)ethoxy)ethoxy)-ethyl)-5-((3a*S*,4*R*,6a*R*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)-pentanamide 42**



Biotinylated compound **42** was obtained according to General Procedure D, using amine **54** (30.0 mg, 69.0 µmol), D-biotin (33.7 mg, 0.138 mmol), EDC hydrochloride (26.5 mg, 0.138 mmol) and HOBt hydrate (4.66 mg, 34.5 µmol) in DMF (3 mL). The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 95:5 → 93:7) to yield product **42** (32.5 mg, 49.2 µmol, 71%) as a white solid.

R_f 0.19 (CH_2Cl_2 : MeOH, 93:7); **ν_{max}** (cm^{-1}) 2930 (C-H), 1697 ((C=O)NH), 1653 ((C=O)NH), 1553, 1455, 1301 (S=O), 1135 (S=O), 1049 (C-O); **mp** 102-104 ℃; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.31 (1H, d, J = 1.8 Hz, H_6), 7.92 (1H, dd, J = 8.5, 1.8 Hz, H_2), 7.85 (1H, ddd, J = 8.0, 1.2, 1.2 Hz, $H_{6'}$), 7.78 (1H, dd, J = 2.6, 1.5 Hz, $H_{2'}$), 7.74 (1H, d, J = 8.5 Hz, H_3), 7.45 (1H, dd, J = 8.0, 8.0 Hz, $H_{5'}$), 7.14 (1H, ddd, J = 8.0, 2.6, 1.0 Hz, $H_{4'}$), 6.75 (1H, t, J = 5.6 Hz, -NH_a), 6.62 (1H, s, -NH), 5.68 (1H, s, -NH), 4.45 (1H, dd, J = 7.8, 4.9 Hz, $H_{20'}$), 4.24 (3H, m, $H_{7'}$ and $H_{22'}$), 3.92-3.88 (2H, m, $H_{8'}$), 3.74-3.71 (2H, m, $H_{9'}$), 3.67-3.64 (2H, m, $H_{10'}$), 3.58 (2H, t, J = 5.1 Hz, $H_{11'}$), 3.48-3.36 (2H, m, $H_{12'}$), 3.17 (2H, q, J = 7.4 Hz, H_8), 3.05 (1H, td, J = 7.3, 4.4 Hz, $H_{18'}$), 2.83 (1H, dd, J = 12.8, 4.9 Hz, $H_{19'a}$), 2.69 (1H, d, J = 12.8 Hz, $H_{19'b}$), 2.17 (2H, m, $H_{14'}$), 1.73-1.55 (4H, m, $H_{15'}$ and $H_{17'}$), 1.39-1.33 (2H, m, $H_{16'}$), 1.29 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 173.4 ($C_{13'}$), 165.3 (C_7), 164.2 ($C_{21'}$), 159.2 ($C_{3'}$), 153.8 (C_4), 142.7 (C_5), 135.4 (C_1), 130.5 ($C_{5'}$), 127.4 ($C_{1'}$), 125.5 (C_2), 121.1 (C_6), 120.8 ($C_{6'}$), 119.7 ($C_{4'}$), 113.2 ($C_{2'}$), 111.6 (C_3), 70.8 ($C_{9'}$), 70.2 ($C_{10'}$), 70.1 ($C_{11'}$), 69.7 ($C_{8'}$), 67.9 ($C_{7'}$), 61.8 ($C_{22'}$), 60.3 ($C_{20'}$), 55.7 ($C_{18'}$), 51.1 (C_8), 40.6 ($C_{19'}$), 39.3 ($C_{12'}$), 36.1 ($C_{14'}$), 28.3 ($C_{15'}$ or $C_{17'}$), 28.2 ($C_{16'}$), 25.7 ($C_{15'}$ or $C_{17'}$), 7.7 (C_9); **HRMS** (ES^+) Calc. for $C_{31}H_{41}N_4O_7S_2$ $[M+H]^+$ 661.2360, found 661.2357.

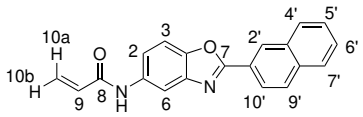
2-(Naphthalen-2-yl)benzo[*d*]oxazol-5-amine 63



2-Naphthoic acid (873 mg, 5.07 mmol) was added to 2,4-diaminophenol dihydrochloride (1.00 g, 5.07 mmol) in polyphosphoric acid (20 mL) at 110 ℃. The reaction mixture was heated for 64 h at 180 ℃ then cooled to rt and basified to pH 10 (2 M NaOH aq.). The precipitate formed was collected and washed with NaOH (20 mL, 1 M aq.) to yield benzoxazole **63** (1.25 g, 4.81 mmol, 95%) as a yellow solid.

R_f 0.45 (EtOAc : PE, 7:3); **ν_{max}** (cm^{-1}) 3324 (N-H), 1546, 1452, 1382, 1184; **mp** 169-170 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.73 (1H, s, $H_{2'}$), 8.28 (1H, dd, J = 8.5, 1.7 Hz, $H_{10'}$), 7.99-7.94 (2H, *obs* m, $H_{9'}$ and $H_{4'}$), 7.88 (1H, dd, J = 7.8, 1.9 Hz, $H_{7'}$), 7.59-7.53 (2H, m, $H_{5'}$ and $H_{6'}$), 7.39 (1H, d, J = 8.6 Hz, H_3), 7.08 (1H, d, J = 2.2 Hz, H_6), 6.73 (1H, dd, J = 8.6, 2.2 Hz, H_2), 3.74 (2H, s, $-NH_2$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.6 (C_7), 144.9 (C_4), 144.0 (C_1), 143.3 (C_5), 134.7 ($C_{8'}$), 133.0 ($C_{3'}$), 128.9 ($C_{9'}$ or $C_{4'}$), 128.7 ($C_{9'}$ or $C_{4'}$), 127.9 ($C_{2'}$), 127.9 ($C_{7'}$), 127.7 ($C_{5'}$ or $C_{6'}$), 126.9 ($C_{1'}$), 124.68 ($C_{5'}$ or $C_{6'}$), 124.0 ($C_{10'}$), 113.8 (C_2), 110.7 (C_3), 105.0 (C_6); **HRMS** (ES^+) Calc. for $C_{17}H_{13}N_2O$ $[M+H]^+$ 261.1022, found 261.1023.

N-(2-(naphthalen-2-yl)benzo[*d*]oxazol-5-yl)acrylamide 64

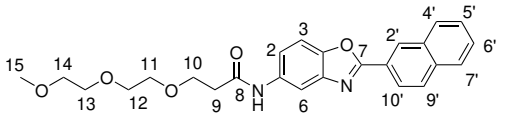


To a solution of aniline **63** (268 mg, 1.03 mmol) and NEt_3 (0.29 mL, 2.06 mmol) in CH_2Cl_2 (10 mL) at 0 ℃ was added acryloyl chloride (84 µL) dropwise. The reaction mixture was warmed to rt, stirred for 16 h, then concentrated *in vacuo*. The residue was dissolved in HCl (10 mL, 1 M aq.) and extracted with CH_2Cl_2 (3 × 10 mL). The combined organic fractions were washed with Na_2CO_3 (20 mL, aq.), dried (Na_2SO_4) and concentrated *in vacuo*. The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 199:1 → 49:1) to yield acrylamide **64** (279 mg, 0.888 mmol, 86%) as an off-white solid.

R_f 0.19 (CH_2Cl_2 : MeOH, 97:3); **ν_{max}** (cm^{-1}) 3105 (alkene C-H), 1663 ((C=O)NH), 1576, 1479, 1234; **mp** 205-206 °C; **^{1}HNMR** (500 MHz, $(CD_3)_2SO$) δ_H 10.37 (1H, s, -NH), 8.83 (1H, d, J = 1.6 Hz, $H_{2'}$), 8.29 (1H, d, J = 2.0 Hz, H_6), 8.26 (1H, dd, J = 8.6, 1.6 Hz, $H_{10'}$), 8.18 (1H, dd, J = 7.0, 1.7 Hz, $H_{4'}$), 8.13 (1H, d, J = 8.6 Hz, $H_{9'}$), 8.04 (1H, dd, J = 7.4, 1.9 Hz, $H_{7'}$), 7.79 (1H, d, J = 8.7 Hz, H_3), 7.69-7.63 (2H, m, $H_{5'}$ and $H_{6'}$), 7.62 (1H, dd, J = 8.7, 2.0 Hz, H_2), 6.48 (1H, dd, J = 17.0, 10.1 Hz, H_9), 6.31 (1H, dd, J = 17.0, 1.9 Hz, H_{10a}), 5.80 (1H, dd, J = 10.1, 1.9 Hz, H_{10b}); **^{13}CNMR** (125 MHz, $(CD_3)_2SO$) δ_C 163.2 (C_8), 163.1 (C_7), 146.6 (C_4), 141.8 (C_5), 136.4 (C_1), 134.3 ($C_{8'}$), 132.6 ($C_{3'}$), 131.8 (C_9), 129.0 ($C_{4'}$), 129.0

($C_{9'}$), 128.2 ($C_{6'}$), 127.9 ($C_{7'}$), 127.8 ($C_{2'}$), 127.2 (C_{10}), 127.0 ($C_{5'}$), 123.7 ($C_{1'}$), 123.6 ($C_{10'}$), 117.8 (C_2), 110.8 (C_3), 110.2 (C_6); **HRMS** (ES^+) Calc. for $C_{20}H_{14}N_2O_2$ $[M+Na]^+$ 337.0948, found 337.0943.

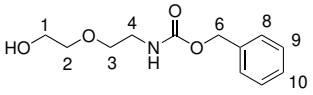
3-(2-(2-Methoxyethoxy)ethoxy)-*N*-(2-(naphthalen-2-yl)benzo[*d*]oxazol-5-yl)-propanamide 65



Adapting a known procedure,[177] copper(II) chloride (10.7 mg, 79.5 µmol) and caesium carbonate (259 mg, 0.795 mmol) were added to a solution of acrylamide **64** (250 mg, 0.795 mmol) and 2-(2-methoxyethoxy)ethanol (0.19 mL, 1.59 mmol) in CH_2Cl_2 (10 mL). The reaction mixture was heated under reflux for 64 h, then concentrated *in vacuo*. The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 199:1 → 49:1) and reverse phase column chromatography (0.1% formic acid in H_2O : MeCN, 13:7 → 11:9) to yield product **65** (39.0 mg, 90.0 µmol, 11%) as a white solid.

R_f 0.31 (CH_2Cl_2 : MeOH, 97:3); **ν_{max}** (cm^{-1}) 2980 (C-H), 2888 (C-H), 1668 ((C=O)NH), 1547, 1481, 1103, 1050 (C-O); **mp** 120-121 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.86 (1H, s, -NH), 8.72 (1H, s, $H_{2'}$), 8.27 (1H, dd, J = 8.5, 1.7 Hz, $H_{10'}$), 8.06 (1H, d, J = 2.0 Hz, H_6), 7.98-7.90 (2H, m, $H_{4'}$ and $H_{9'}$), 7.86 (1H, dd, J = 7.0, 2.2 Hz, $H_{7'}$), 7.58-7.52 (3H, m, H_2, $H_{5'}$ and $H_{6'}$), 7.50 (1H, d, J = 8.8 Hz, H_3), 3.85 (2H, dd, J = 6.1, 5.0 Hz, H_{10}), 3.76-3.71 (4H, m, H_{11} and H_{12}), 3.69-3.67 (2H, m, H_{13}), 3.55-3.53 (2H, m, H_{14}), 3.33 (3H, s, H_{15}), 2.69 (2H, dd, J = 6.1, 5.0 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 170.2 (C_8), 163.9 (C_7), 147.6 (C_4), 142.6 (C_5), 135.6 (C_1), 134.8 ($C_{8'}$), 133.0 ($C_{3'}$), 129.0 ($C_{9'}$), 128.8 ($C_{4'}$), 128.2 ($C_{2'}$), 128.0 ($C_{7'}$), 127.9 ($C_{6'}$), 127.0 ($C_{5'}$), 124.5 ($C_{1'}$), 124.0 ($C_{10'}$), 118.5 (C_2), 111.7 (C_6), 110.3 (C_3), 72.0 (C_{14}), 70.7 (C_{13}), 70.5 (C_{12}), 70.5 (C_{11}), 67.3 (C_{10}), 59.1 (C_{15}), 38.0 (C_9); **HRMS** (ES^+) Calc. for $C_{25}H_{27}N_2O_5$ $[M+H]^+$ 435.1915, found 435.1910.

Benzyl(2-(2-hydroxyethoxy)ethyl)carbamate 58



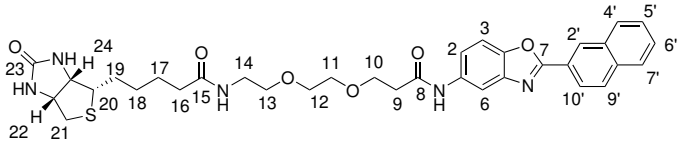
To a solution of 2-(2-aminoethoxy)ethanol (1.00 mL, 9.97 mmol) and NEt_3 (1.39 mL, 9.97 mmol) in THF (25 mL) was added benzyl chloroformate (1.57 mL, 10.96 mmol) dropwise. The

reaction mixture was stirred for 16 h at rt, then diluted with EtOAc (25 mL) and washed with water (2 × 20 mL) and brine (20 mL). The organic layer was dried (Na_2SO_4) and concentrated *in vacuo* to yield protected aminoalcohol **58** (1.75 g, 7.31 mmol, 73%) as a colourless oil.

R_f 0.34 (CH_2Cl_2 : MeOH, 19:1); **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 7.40-7.30 (5H, m, $H_{8\text{-}10}$), 5.22 (1H, *br* s, -NH), 5.10 (2H, s, H_6), 3.75-3.70 (2H, m, H_1), 3.58-3.53 (4H, m, H_2 and H_3), 3.43-3.48 (2H, m, H_4), 2.20 (1H, *br* s, -OH); **LRMS** (ES^+) Calc. for $C_{12}H_{17}NO_4$ $[M+H]^+$ 240.0, found 240.0.

These data are in accordance with that of Sissi *et al.*[354]

***N*-(2-(2-(3-((2-(naphthalen-2-yl)benzo[*d*]oxazol-5-yl)amino)-3-oxopropoxy)-ethoxy)ethyl)-5-((3a*S*,4*S*,6a*R*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)-pentanamide 41**



Adapting a known procedure,[177] copper(II) chloride (16.3 mg, 0.121 mmol) and caesium carbonate (394 mg, 1.21 mmol) were added to a solution of acrylamide **64** (380 mg, 1.21 mmol) and alcohol **58** (577 mg, 2.42 mmol) in CH_2Cl_2 (10 mL). The reaction mixture was heated under reflux for 64 h, then concentrated *in vacuo*.

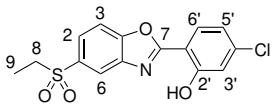
The crude product was purified by flash column chromatography (toluene : MeOH, 19:1). To a solution of the protected amine intermediate in CH_2Cl_2 (5 mL) and MeOH (5 mL) under an argon atmosphere was added Pd/C (64.4 mg, 10% wt., 0.605 mmol). The reaction mixture was stirred under H_2 at rt for 20 h, then filtered through Celite, eluting with MeOH. The filtrate was concentrated *in vacuo* to yield the amine which was used directly in the next step.

Biotinylated compound **41** was obtained according to General Procedure D, using the prepared amine, D-biotin (296 mg, 1.21 mmol), EDC hydrochloride (464 mg, 2.42 mmol), NEt_3 (0.337 mL, 2.42 mmol) and HOBt hydrate (16.3 mg, 0.121 mmol) in DMF (10 mL). The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 49:1 → 23:2) and reverse phase column chromatography (0.1% formic acid in H_2O : MeCN, 3:1 → 3:2) to yield product **41** (15.6 mg, 24.2 µmol, 2%) as a white foam.

R_f 0.20 (CH_2Cl_2 : MeOH, 37:3); **ν_{max}** (cm^{-1}) 3280 (N-H), 2924 (C-H), 1697 ((C=O)NH), 1684 ((C=O)NH), 1653 ((C=O)NH), 1558, 1481, 1262 (C-N), 1118 (C-O); **^{1}HNMR** (500 MHz, CD_3OD) δ_H 8.78 (1H, d, J = 1.7 Hz, $H_{2'}$), 8.27 (1H, dd, J = 8.7, 1.7 Hz, $H_{10'}$), 8.18

(1H, d, J = 2.0 Hz, H_6), 8.06 (1H, d, J = 6.8, 2.0 Hz, $H_{4'}$), 8.05 (1H, d, J = 8.7 Hz, $H_{9'}$), 7.96 (1H, dd, J = 6.8, 2.1 Hz, $H_{7'}$), 7.66 (1H, d, J = 8.8 Hz, H_3), 7.65-7.59 (2H, m, $H_{5'}$ and $H_{6'}$), 7.54 (1H, dd, J = 8.8, 2.0 Hz, H_2), 4.40 (1H, ddd, J = 7.9, 5.0, 0.9 Hz, H_{22}), 4.20 (1H, dd, J = 7.9, 4.4 Hz, H_{24}), 3.88 (2H, t, J = 6.0 Hz, H_{10}), 3.68 (2H, dt, J = 5.0, 1.8 Hz, H_{11}), 3.66-3.63 (2H, m, H_{12}), 3.55 (2H, t, J = 5.4 Hz, H_{13}), 3.33 (2H, *obs* m, H_{14}), 3.09 (1H, ddd, J = 8.8, 5.9, 4.4 Hz, H_{20}), 2.83 (1H, dd, J = 12.7, 5.0 Hz, H_{21a}), 2.69 (2H, t, J = 6.0 Hz, H_9), 2.64 (1H, d, J = 12.7 Hz, H_{21b}), 2.17-2.13 (2H, m, H_{16}), 1.70-1.49 (4H, m, H_{17} and H_{19}), 1.38-1.33 (2H, m, H_{18}); **^{13}CNMR** (125 MHz, CD_3OD) δ_C 176.1 (C_{15}), 172.4 (C_8), 166.1 (C_{23}), 165.6 (C_7), 148.8 (C_4), 143.2 (C_5), 137.4 (C_1), 136.4 ($C_{8'}$), 134.4 ($C_{3'}$), 130.1 ($C_{4'}$), 130.1 ($C_{9'}$), 129.3 ($C_{2'}$), 129.3 ($C_{5'}$ or $C_{6'}$), 129.0 ($C_{7'}$), 128.3 ($C_{5'}$ or $C_{6'}$), 125.2 ($C_{1'}$), 124.7 ($C_{10'}$), 119.7 (C_2), 112.3 (C_6), 111.7 (C_3), 71.5 (C_{11}), 71.3 (C_{12}), 70.6 (C_{13}), 68.2 (C_{10}), 63.3 (C_{24}), 61.6 (C_{22}), 57.0 (C_{20}), 41.0 (C_{21}), 40.4 (C_{14}), 38.5 (C_9), 36.7 (C_{16}), 29.7 (C_{18}), 29.5 (C_{17} or C_{19}), 26.8 (C_{17} or C_{19}); **HRMS** (ES^+) Calc. for $C_{34}H_{40}N_5O_6S$ $[M+H]^+$ 646.2694, found 646.2692.

5-Chloro-2-(5-(ethylsulfonyl)benzo[*d*]oxazol-2-yl)phenol 68



Benzoxazole **68** was synthesised by two methods:

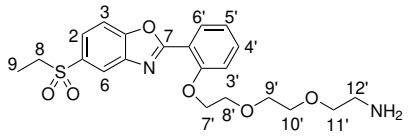
Method A, according to General Procedure A, using 4-chlorosalicylic acid (431 mg, 2.50 mmol), thionyl chloride (0.54 mL, 7.50 mmol) and DMF (0.10 mL) in CH_2Cl_2 (15 mL). The acid chloride intermediate formed was reacted with 2-amino-4-ethylsulfonylphenol (503 mg, 2.50 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 3:2) to yield benzoxazole **68** (127 mg, 0.375 mmol, 15%) as an off-white solid.

Method B, 4-chlorosalicylic acid (1.00 g, 5.79 mmol) was added to 2-amino-4-ethylsulfonylphenol (1.17 g, 5.79 mmol) in polyphosphoric acid (20 mL) at 110 ℃. The reaction mixture was heated for 5 h at 180 ℃ then cooled to rt and diluted with water (20 mL). The mixture was extracted with EtOAc (3 × 30 mL) and the combined organic fractions concentrated *in vacuo*. The crude product was purified by flash column chromatography (PE : EtOAc, 3:2) to yield benzoxazole **68** (355 mg, 1.05 mmol, 18%) as an off-white solid.

R_f 0.36 (PE : EtOAc, 3:2); **ν_{max}** (cm^{-1}) 2942 (C-H), 1630, 1455, 1309 (S=O), 1133 (S=O), 740 (C-Cl); **mp** 170-171 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 11.19 (1H, *br* s, -OH), 8.30 (1H, d, J = 1.9 Hz, H_6), 7.98 (1H, dd, J = 8.4, 1.9 Hz, H_2), 7.95 (1H, d, J = 8.5 Hz, $H_{6'}$), 7.78 (1H,

d, J = 8.4 Hz, H_3), 7.16 (1H, d, J = 2.0 Hz, $H_{3'}$), 7.03 (1H, dd, J = 8.5, 2.0 Hz, $H_{5'}$), 3.19 (2H, q, J = 7.4 Hz, H_8), 1.32 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.6 (C_7), 159.7 ($C_{2'}$), 152.1 (C_4), 140.6 (C_5), 136.1 (C_1), 128.4 ($C_{6'}$), 126.0 (C_2), 120.8 ($C_{5'}$), 120.3 (C_6), 118.2 ($C_{1'}$ and $C_{3'}$), 111.7 (C_3), 108.4 ($C_{4'}$), 51.1 (C_8), 7.7 (C_9); **HRMS** (ES^+) Calc. for $C_{15}H_{13}ClNO_4S$ $[M+H]^+$ 338.0248, found 338.0248.

2-(2-(2-(2-(5-(Ethylsulfonyl)benzo[*d*]oxazol-2-yl)phenoxy)ethoxy)ethoxy)ethan-1-amine 70

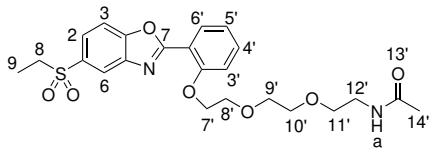


A solution of benzoxazole **68** (100 mg, 0.296 mmol), protected linker **53*** (107 mg, 0.296 mmol) and K_2CO_3 (81.8 mg, 0.592 mmol) in DMF (10 mL) was stirred at 80 ℃ for 16 h. The reaction mixture was then quenched with water (5 mL) and extracted with EtOAc (2 × 10 mL). The combined organic fractions were dried (Na_2SO_4) and concentrated *in vacuo*. The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 99:1). To a solution of the protected amine intermediate in MeOH (10 mL) under an argon atmosphere was added Pd/C (15.7 mg, 10% wt., 0.148 mmol). The reaction mixture was stirred under H_2 at rt for 20 h, then filtered through Celite, eluting with MeOH. The filtrate was concentrated *in vacuo* to yield amine **70** (83.7 mg, 0.193 mmol, 65%) as a yellow oil.

*Linker compound kindly supplied by Aini Vuorinen.[10]

ν_{max} (cm^{-1}) 3398 (N-H), 2980 (C-H), 1613, 1455, 1304 (S=O), 1240 (C-N), 1134 (S=O), 1054 (C-O); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.42 (1H, dd, J = 1.9, 0.6 Hz, H_6), 8.12 (1H, dd, J = 8.0, 1.7 Hz, $H_{6'}$), 7.94 (1H, dd, J = 8.5, 1.9 Hz, H_2), 7.74 (1H, dd, J = 8.5, 0.6 Hz, H_3), 7.54 (1H, ddd, J = 8.0, 7.4, 1.8 Hz, $H_{5'}$), 7.14 (2H, *obs* m, $H_{3'}$ and $H_{4'}$), 4.39-4.36 (2H, m, $H_{7'}$), 3.98-3.95 (2H, m, $H_{8'}$), 3.80-3.78 (2H, m, H), 3.77-3.74 (2H, m, $H_{11'}$), 3.71-3.69 (2H, m, H), 3.25 (2H, q, J = 7.4 Hz, H_8), 3.09-3.06 (2H, m, $H_{12'}$), 1.30 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.9 (C_7), 158.1 ($C_{2'}$), 153.4 (C_4), 142.5 (C_5), 135.3 (C_1), 134.0 ($C_{5'}$), 131.7 ($C_{6'}$), 125.5 (C_2), 121.6 ($C_{4'}$), 121.1 (C_6), 115.5 ($C_{1'}$), 114.2 ($C_{3'}$), 111.4 (C_3), 71.1 ($C_{9'}$), 70.4 ($C_{10'}$), 69.8 ($C_{8'}$), 69.3 ($C_{7'}$), 68.5 ($C_{11'}$), 51.1 (C_8), 40.2 ($C_{12'}$), 7.7 (C_9); **HRMS** (ES^+) Calc. for $C_{21}H_{27}N_2O_6S$ $[M+H]^+$ 435.1595, found 435.1585.

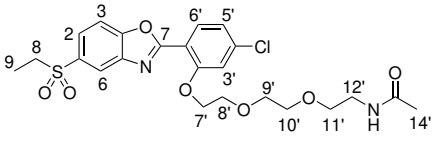
***N*-(2-(2-(2-(2-(5-(ethylsulfonyl)benzo[*d*]oxazol-2-yl)phenoxy)ethoxy)ethoxy)-ethyl)acetamide 74**



Acetamide **74** was obtained according to General Procedure C, using amine **70** (21.0 mg, 48.3 µmol), 4-dimethylaminopyridine (0.590 mg, 4.83 µmol), acetic anhydride (5 µL, 53.2 µmol) and NEt_3 (7 µL, 53.2 µmol) in $CHCl_3$ (1 mL). The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 49:1 → 97:3) to yield acetamide **74** (18.7 mg, 39.2 µmol, 81%) as a clear oil.

R_f 0.33 (CH_2Cl_2 : MeOH, 19:1); **ν_{max}** (cm^{-1}) 2980 (C-H), 1654 ((C=O)NH), 1549, 1455, 1305 (S=O), 1239 (C-N), 1134 (S=O), 1054 (C-O); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.33 (1H, d, J = 1.7 Hz, H_6), 8.14 (1H, dd, J = 7.8, 1.7 Hz, $H_{6'}$), 7.93 (1H, dd, J = 8.5, 1.8 Hz, H_2), 7.73 (1H, d, J = 8.5 Hz, H_3), 7.54 (1H, ddd, J = 8.4, 7.4, 1.7 Hz, $H_{4'}$), 7.16-7.11 (2H, *obs* m, $H_{3'}$ and $H_{5'}$), 6.09 (1H, *br* s, -NH), 4.34 (2H, t, J = 4.8 Hz, $H_{7'}$), 4.00-3.97 (2H, m, $H_{8'}$), 3.82-3.79 (2H, m, $H_{9'}$), 3.63-3.60 (2H, m, $H_{10'}$), 3.52 (2H, t, J = 5.3 Hz, $H_{11'}$), 3.38 (2H, q, J = 5.3 Hz, $H_{12'}$), 3.18 (2H, q, J = 7.4 Hz, H_8), 1.90 (3H, s, $H_{14'}$), 1.31 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 170.2 ($C_{13'}$), 164.2 (C_7), 158.2 ($C_{2'}$), 153.6 (C_4), 142.8 (C_5), 135.1 (C_1), 133.8 ($C_{4'}$), 131.7 ($C_{6'}$), 125.3 (C_2), 121.5 ($C_{5'}$), 121.1 (C_6), 115.9 ($C_{1'}$), 114.0 ($C_{3'}$), 111.3 (C_3), 71.1 ($C_{9'}$), 70.4 ($C_{10'}$), 70.0 ($C_{11'}$), 69.8 ($C_{8'}$), 69.4 ($C_{7'}$), 51.2 (C_8), 39.4 ($C_{12'}$), 23.3 ($C_{14'}$), 7.8 (C_9); **HRMS** (ES^+) Calc. for $C_{23}H_{29}N_2O_7S$ $[M+H]^+$ 477.1690, found 477.1687.

***N*-(2-(2-(2-(5-chloro-2-(5-(ethylsulfonyl)benzo[*d*]oxazol-2-yl)phenoxy)ethoxy)-ethoxy)ethyl)acetamide 73**



A solution of benzoxazole **68** (80 mg, 0.237 mmol), protected linker **53*** (86 mg, 0.237 mmol) and K_2CO_3 (65.5 mg, 0.474 mmol) in DMF (5 mL) was stirred at 80 ℃ for 16 h. The reaction mixture was then quenched with water (5 mL) and extracted with EtOAc (2 × 10 mL). The combined organic fractions were dried (Na_2SO_4) and concentrated *in vacuo*. The

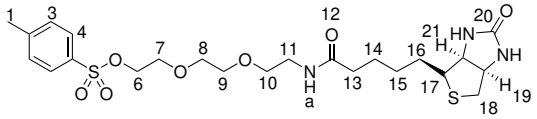
crude mixture was purified by flash column chromatography (CH_2Cl_2 : MeOH, 99:1). The protected amine intermediate was then added to conc. HCl (5 mL) and stirred at 80 ℃ for 10 min. The reaction mixture was cooled, diluted with water (5 mL) and extracted with EtOAc (2 × 10 mL). The aqueous phase was then basified to pH 10 (2 M NaOH aq.) and extracted with EtOAc (3 × 10 mL). The organic fraction was concentrated *in vacuo* and used directly in the next step.

Acetamide **73** was obtained according to General Procedure C, using the prepared amine, 4-dimethylaminopyridine (2.90 mg, 23.7 µmol), acetic anhydride (25 µL, 0.261 mmol) and NEt_3 (36 µL, 0.261 mmol) in $CHCl_3$ (1 mL). The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 49:1 → 24:1) to yield acetamide **73** (41.2 mg, 80.6 µmol, 34% over two steps) as a yellow oil.

*Linker compound kindly supplied by Aini Vuorinen.[10]

R_f 0.25 (CH_2Cl_2 : MeOH, 24:1); **ν_{max}** (cm^{-1}) 2885 (C-H), 1654 ((C=O)NH), 1595, 1306 (S=O), 1253 (C-N), 1137 (S=O); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.31 (1H, d, J = 1.8 Hz, H_6), 8.07 (1H, d, J = 8.2 Hz, $H_{6'}$), 7.91 (1H, dd, J = 8.5, 1.8 Hz, H_2), 7.71 (1H, d, J = 8.5 Hz, H_3), 7.12 (1H, *obs* d, $H_{3'}$), 7.11 (1H, dd, J = 8.2, 1.8 Hz, $H_{5'}$), 6.13 (1H, *br* s, -NH), 4.32 (2H, t, J = 4.7 Hz, $H_{7'}$), 3.97 (2H, t, J = 4.7 Hz, $H_{8'}$), 3.81-3.78 (2H, m, $H_{9'}$), 3.63-3.60 (2H, m, $H_{10'}$), 3.51 (2H, t, J = 5.3 Hz, $H_{11'}$), 3.37 (2H, q, J = 5.3 Hz, $H_{12'}$), 3.17 (2H, q, J = 7.4 Hz, H_8), 1.90 (3H, s, $H_{14'}$), 1.28 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 170.2 ($C_{13'}$), 163.3 (C_7), 158.6 ($C_{2'}$), 153.5 (C_4), 142.6 (C_5), 139.6 ($C_{4'}$), 135.2 (C_1), 132.4 ($C_{6'}$), 125.4 (C_2), 121.8 ($C_{5'}$), 121.1 (C_6), 114.6 ($C_{3'}$), 114.3 ($C_{1'}$), 111.4 (C_3), 71.1 ($C_{9'}$), 70.4 ($C_{10'}$), 70.0 ($C_{11'}$), 69.6 ($C_{7'}$), 69.6 ($C_{8'}$), 51.1 (C_8), 39.3 ($C_{12'}$), 23.3 ($C_{14'}$), 7.7 (C_9); **HRMS** (ES^+) Calc. for $C_{23}H_{28}ClN_2O_7S$ $[M+H]^+$ 511.1300, found 511.1301.

2-(2-(2-(5-((3a*S*,4*R*,6a*R*)-2-Oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)-pentanamido)ethoxy)ethoxy)ethyl 4-methylbenzenesulfonate 77

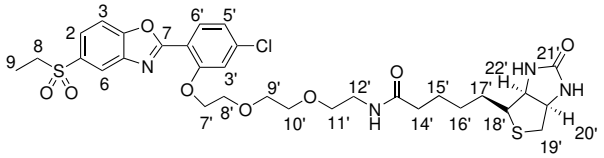


To a solution of 2-(2-(2-aminoethoxy)ethoxy)ethanol (250 mg, 1.68 mmol), D-biotin (372 mg, 1.52 mmol), EDC hydrochloride (437 mg, 2.28 mmol) and HOBt hydrate (103 mg, 0.760 mmol) in DMF (20 mL) was added NEt_3 (0.32 mL, 2.28 mmol) dropwise. The reaction mixture was stirred for 16 h at rt, then concentrated *in vacuo*. The crude mixture was dissolved in

CH_2Cl_2 (20 mL), cooled to 0 ℃ and *p*-toluenesulfonyl chloride (320 mg, 1.68 mmol) and NEt_3 (0.23 mL, 1.68 mmol) were added. The reaction mixture was warmed to rt, stirred for 6 h then washed with $NaHCO_3$ solution (20 mL), dried (Na_2SO_4) and concentrated *in vacuo*. The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 49:1 → 19:1) to yield product **77** (386 mg, 0.728 mmol, 48%) as a yellow oil.

R_f 0.20 (PE : EtOAc, 4:1); **ν_{max}** (cm^{-1}) 1694 ((C=O)NH), 1648 ((C=O)NH), 1541, 1457, 1351 (S=O), 1174 (S=O), 1096 (C-O); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 7.78 (2H, d, J = 8.2 Hz, H_4), 7.34 (2H, d, J = 8.2 Hz, H_3), 6.69 (1H, t, J = 5.7 Hz, -NH_a), 6.53 (1H, *br* s, -NH), 5.68 (1H, *br* s, -NH), 4.48 (1H, dd, J = 7.8, 4.9 Hz, H_{19}), 4.29 (1H, dd, J = 7.8, 4.5 Hz, H_{21}), 4.16-4.13 (2H, m, H_6), 3.69-3.67 (2H, m, H_7), 3.60-3.57 (2H, m, H_8), 3.57-3.55 (2H, m, H_9), 3.53 (2H, t, J = 5.1 Hz, H_{10}), 3.40 (2H, m, H_{11}), 3.12 (1H, td, J = 7.4, 4.5 Hz, H_{17}), 2.87 (1H, dd, J = 12.8, 4.9 Hz, H_{18a}), 2.70 (1H, d, J = 12.8 Hz, H_{18b}), 2.44 (3H, s, H_1), 2.21 (2H, t, J = 7.5 Hz, H_{13}), 1.77-1.59 (4H, m, H_{14} and H_{16}), 1.46-1.37 (2H, m, J = 6.7 Hz, H_{15}); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 173.5 (C_{12}), 164.2 (C_{20}), 145.1 (C_2), 133.0 (C_5), 130.0 (C_3), 128.0 (C_4), 70.7 (C_8), 70.1 (C_9 and C_{10}), 69.4 (C_6), 68.8 (C_7), 61.9 (C_{21}), 60.3 (C_{19}), 55.8 (C_{17}), 40.7 (C_{18}), 39.2 (C_{11}), 36.1 (C_{13}), 28.3 (C_{15}), 28.2 (C_{14} or C_{16}), 25.7 (C_{14} or C_{16}), 21.8 (C_1); **HRMS** (ES^+) Calc. for $C_{23}H_{36}N_3O_7S_2$ $[M+H]^+$ 530.1989, found 530.1987.

***N*-(2-(2-(2-(5-chloro-2-(5-(ethylsulfonyl)benzo[*d*]oxazol-2-yl)phenoxy)ethoxy)-ethoxy)ethyl)-5-((3a*S*,4*R*,6a*R*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)-pentanamide 43**

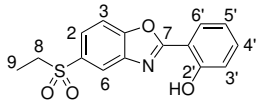


To a solution of phenol **68** (100 mg, 0.189 mmol) and K_2CO_3 (52.2 mg, 0.378 mmol) in DMF (5 mL) was added tosylate **77** (63.8 mg, 0.189 mmol). The reaction mixture was stirred at 100 ℃ for 2 h, then concentrated *in vacuo*. The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 24:1 → 19:1) and reverse phase column chromatography (0.1% formic acid in H_2O : MeCN, 3:1 → 3:2) to yield product **43** (15.5 mg, 22.3 µmol, 12%) as a white foam.

R_f 0.18 (CH_2Cl_2 : MeOH, 19:1); **ν_{max}** (cm^{-1}) 2985 (C-H), 1698 ((C=O)NH), 1656 ((C=O)NH), 1457, 1382, 1305 (S=O), 1258 (C-N), 1135 (S=O); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.32 (1H, d, J = 1.8 Hz, H_6), 8.09 (1H, d, J = 8.3 Hz, $H_{6'}$), 7.93 (1H, dd, J = 8.5, 1.8 Hz,

H_2), 7.73 (1H, d, J = 8.5 Hz, H_3), 7.16-7.09 (2H, *obs* m, $H_{3'}$ and $H_{5'}$), 6.64 (1H, t, J = 5.5 Hz, -NH_a), 6.44 (1H, *br* s, -NH), 5.52 (1H, *br* s, -NH), 4.45 (1H, dd, J = 7.8, 4.9 Hz, $H_{20'}$), 4.32 (2H, t, J = 5.0 Hz, $H_{7'}$), 4.28-4.22 (1H, m, $H_{22'}$), 3.98 (2H, dd, J = 5.0, 3.6 Hz, $H_{8'}$), 3.80 (2H, dd, J = 5.8, 3.6 Hz, $H_{9'}$), 3.65-3.59 (2H, m, $H_{10'}$), 3.53 (2H, t, J = 5.1 Hz, $H_{11'}$), 3.37 (2H, p, J = 5.1 Hz, $H_{12'}$), 3.18 (2H, q, J = 7.4 Hz, H_8), 3.08 (1H, td, J = 7.4, 4.7 Hz, $H_{18'}$), 2.85 (1H, dd, J = 12.8, 4.7 Hz, $H_{19'a}$), 2.70 (1H, d, J = 12.8 Hz, $H_{19'b}$), 2.18-2.11 (2H, m, $H_{14'}$), 1.72-1.62 (2H, m, $H_{17'}$), 1.64-1.54 (2H, m, $H_{15'}$), 1.36 (2H, tt, J = 7.5, 7.5 Hz, $H_{16'}$), 1.29 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 173.4 ($C_{13'}$), 164.0 ($C_{21'}$), 163.4 (C_7), 158.7 ($C_{2'}$), 153.5 (C_4), 142.6 (C_5), 139.7 (C_1), 135.2 ($C_{4'}$), 132.4 ($C_{6'}$), 125.5 (C_2), 121.8 ($C_{3'}$ or $C_{5'}$), 121.1 (C_6), 114.6 ($C_{3'}$ or $C_{5'}$), 114.3 ($C_{1'}$), 111.5 (C_3), 71.0 ($C_{9'}$), 70.7 ($C_{10'}$), 70.1 ($C_{11'}$), 69.7 ($C_{8'}$), 69.5 ($C_{7'}$), 61.8 ($C_{22'}$), 60.3 ($C_{20'}$), 55.7 ($C_{18'}$), 51.1 (C_8), 40.6 ($C_{19'}$), 39.2 ($C_{12'}$), 36.0 ($C_{14'}$), 28.3 ($C_{16'}$), 28.2 ($C_{17'}$), 25.7 ($C_{15'}$), 7.7 (C_9); **HRMS** (ES^+) Calc. for $C_{31}H_{39}N_4O_8S_2Cl$ $[M+H]^+$ 717.1790, found 717.1789.

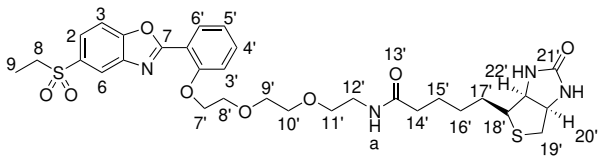
2-(5-(Ethylsulfonyl)benzo[*d*]oxazol-2-yl)phenol 78



Benzoxazole **78** was obtained according to General Procedure A, using salicylic acid (276 mg, 2.00 mmol), thionyl chloride (0.44 mL, 6.00 mmol) and DMF (0.10 mL) in CH_2Cl_2 (15 mL). The acid chloride intermediate formed was reacted with 2-amino-4-ethylsulfonylphenol (403 mg, 2.00 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 1:0 → 99:1) to yield benzoxazole **78** (256 mg, 0.844 mmol, 42%) as an off-white solid.

R_f 0.35 (CH_2Cl_2 : MeOH, 99:1); **ν_{max}** (cm^{-1}) 1631, 1485, 1298 (S=O), 1280, 1140 (S=O); **mp** 167-168 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 10.98 (1H, s, -OH), 8.25 (1H, d, J = 1.9 Hz, H_6), 7.97 (1H, dd, J = 8.2, 1.7 Hz, $H_{6'}$), 7.93 (1H, dd, J = 8.5, 1.9 Hz, H_2), 7.74 (d, J = 8.5 Hz, H_3), 7.45 (1H, ddd, J = 8.4, 7.2, 1.7 Hz, $H_{4'}$), 7.08 (1H, dd, J = 8.4, 1.1 Hz, $H_{3'}$), 7.00 (1H, ddd, J = 8.2, 7.2, 1.1 Hz, $H_{5'}$), 3.18 (2H, q, J = 7.4 Hz, H_8), 1.30 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 165.1 (C_7), 159.1 ($C_{2'}$), 152.1 (C_4), 140.7 (C_5), 135.8 (C_1), 134.7 ($C_{4'}$), 127.4 ($C_{6'}$), 125.7 (C_2), 120.1 (C_6), 120.0 ($C_{5'}$), 117.7 ($C_{3'}$), 111.6 (C_3), 109.6 ($C_{1'}$), 51.0 (C_8), 7.7 (C_9); **HRMS** (ES^+) Calc. for $C_{15}H_{14}NO_4S$ $[M+H]^+$ 304.0638, found 304.0637.

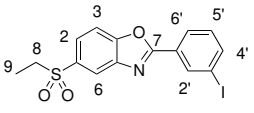
***N*-(2-(2-(2-(2-(5-(ethylsulfonyl)benzo[*d*]oxazol-2-yl)phenoxy)ethoxy)ethoxy)-ethyl)-5-((3a*S*,4*R*,6a*R*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)-pentanamide 75**



To a solution of phenol **78** (136 mg, 0.450 mmol) and K_2CO_3 (124 mg, 0.900 mmol) in DMF (10 mL) was added tosylate **77** (238 mg, 0.450 mmol). The reaction mixture was stirred at 100 ℃ for 2 h, then concentrated *in vacuo*. The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 24:1 → 19:1) to yield product **75** (254 mg, 0.384 mmol, 85%) as a white foam.

R_f 0.28 (CH_2Cl_2 : MeOH, 19:1); **ν_{max}** (cm^{-1}) 2985 (C-H), 1696 ((C=O)NH), 1654 ((C=O)NH), 1382, 1306 (S=O), 1252 (C-N), 1137 (S=O); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.32 (1H, d, J = 1.8 Hz, H_6), 8.12 (1H, dd, J = 7.7, 1.8 Hz, $H_{6'}$), 7.91 (1H, dd, J = 8.5, 1.8 Hz, H_2), 7.73 (1H, d, J = 8.5 Hz, H_3), 7.56-7.48 (1H, ddd, J = 8.0, 7.2, 1.8 Hz, $H_{4'}$), 7.15-7.08 (2H, *obs* m, $H_{3'}$ and $H_{5'}$), 6.76 (1H, t, J = 5.6 Hz, -NH_a), 6.58 (1H, *br* s, -NH), 5.77 (1H, *br* s, -NH), 4.42 (1H, dd, J = 7.8, 4.9 Hz, $H_{20'}$), 4.31 (2H, t, J = 4.7 Hz, $H_{7'}$), 4.22 (1H, ddd, J = 7.8, 4.6, 1.4 Hz, $H_{22'}$), 3.96 (2H, t, J = 4.7 Hz, $H_{8'}$), 3.80-3.76 (2H, m, $H_{9'}$), 3.62-3.57 (2H, m, $H_{10'}$), 3.52 (2H, t, J = 5.1 Hz, $H_{11'}$), 3.36 (2H, dt, J = 5.1, 5.1 Hz, $H_{12'}$), 3.18 (2H, q, J = 7.4 Hz, H_8), 3.05 (1H, td, J = 7.4, 4.6 Hz, $H_{18'}$), 2.82 (1H, dd, J = 12.8, 4.9 Hz, $H_{19'a}$), 2.68 (1H, d, J = 12.8 Hz, $H_{19'b}$), 2.12 (2H, t, J = 7.5 Hz, $H_{14'}$), 1.69-1.52 (4H, m, $H_{15'}$ and $H_{17'}$), 1.34 (2H, m, $H_{16'}$), 1.28 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 173.4 ($C_{13'}$), 164.2 (C_7), 164.2 ($C_{21'}$), 158.1 ($C_{2'}$), 153.6 (C_4), 142.7 (C_5), 135.0 (C_1), 133.9 ($C_{4'}$), 131.6 ($C_{6'}$), 125.2 (C_2), 121.4 ($C_{5'}$), 121.0 (C_6), 115.7 ($C_{1'}$), 114.0 ($C_{3'}$), 111.4 (C_3), 70.9 ($C_{9'}$), 70.2 ($C_{10'}$), 70.0 ($C_{11'}$), 69.6 ($C_{8'}$), 69.2 ($C_{7'}$), 61.8 ($C_{22'}$), 60.3 ($C_{20'}$), 55.7 ($C_{18'}$), 51.1 (C_8), 40.6 ($C_{19'}$), 39.2 ($C_{12'}$), 36.0 ($C_{14'}$), 28.3 ($C_{16'}$), 28.1 ($C_{15'}$ or $C_{17'}$), 25.6 ($C_{15'}$ or $C_{17'}$), 7.7 (C_9); **HRMS** (ES^+) Calc. for $C_{31}H_{41}N_4O_8S_2$ $[M+H]^+$ 661.2360, found 661.2359.

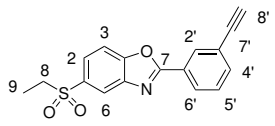
5-(Ethylsulfonyl)-2-(3-iodophenyl)benzo[*d*]oxazole 173



Benzoxazole **174** was obtained according to General Procedure A, using 3-iodobenzoic acid (500 mg, 2.02 mmol), thionyl chloride (0.44 mL, 6.06 mmol) and DMF (0.10 mL) in CH_2Cl_2 (25 mL). The acid chloride intermediate formed was reacted with 2-amino-4-ethylsulfonylphenol (406 mg, 2.02 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 1:1) to yield benzoxazole **173** (575 mg, 1.39 mmol, 69%) as a white solid.

R_f 0.63 (PE : EtOAc, 1:1); **ν_{max}** (cm^{-1}) 2980 (C-H), 1616, 1546, 1346 (S=O), 1260, 1137 (S=O), 1045, 777; **mp** 171-172 ℃; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.63 (1H, dd, J = 1.7, 1.7 Hz, $H_{2'}$), 8.34 (1H, d, J = 1.8 Hz, H_6), 8.24 (1H, ddd, J = 7.9, 1.7, 1.7 Hz, $H_{6'}$), 7.96 (1H, dd, J = 8.4, 1.8 Hz, H_2), 7.92 (1H, ddd, J = 7.9, 1.7, 1.7 Hz, $H_{4'}$), 7.76 (1H, d, J = 8.4 Hz, H_3), 7.30 (1H, dd, J = 7.9, 7.9 Hz, $H_{5'}$), 3.18 (2H, q, J = 7.4 Hz, H_8), 1.31 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.6 (C_7), 153.7 (C_4), 142.5 (C_5), 141.3 ($C_{4'}$), 136.7 ($C_{2'}$), 135.5 (C_1), 130.7 ($C_{5'}$), 128.0 ($C_{1'}$), 127.1 ($C_{6'}$), 125.7 (C_2), 121.3 (C_6), 111.5 (C_3), 94.5 ($C_{3'}$), 51.0 (C_8), 7.6 (C_9); **HRMS** (ES^+) Calc. for $C_{15}H_{13}IN_1O_3S$ $[M+H]^+$ 413.9655, found 413.9636.

5-(Ethylsulfonyl)-2-(3-ethynylphenyl)benzo[*d*]oxazole 92

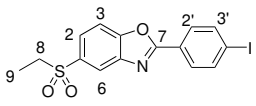


Alkyne **92** was obtained according to General Procedure B, using benzoxazole **173** (550 mg, 1.33 mmol), $Pd(PPh_3)_2Cl_2$ (140 mg, 0.200 mmol), copper(I) iodide (25.3 mg, 0.133 mmol), NEt_3 (3 mL) and trimethylsilylacetylene (0.19 mL, 1.33 mmol) in THF (20 mL). The TMS-protected alkyne intermediate formed was purified by flash column chromatography (PE : EtOAc, 3:1 → 2:1), then deprotected with TBAF (3.99 mL, 1 M in THF) in THF (35 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 7:3) to yield alkyne **92** (237 mg, 0.762 mmol, 57%) as a white solid.

R_f 0.37 (PE : EtOAc, 7:3); **ν_{max}** (cm^{-1}) 3266 (alkyne C-H), 2981 (C-H), 1346 (S=O), 1137 (S-O); **mp** 184-185 ℃; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.39 (1H, dd, J = 1.4, 1.4 Hz, $H_{2'}$), 8.33 (1H, d, J = 1.7 Hz, H_6), 8.25 (1H, ddd, J = 7.7, 1.4, 1.4 Hz, $H_{6'}$), 7.95 (1H, dd, J = 8.5, 1.7 Hz, H_2), 7.75 (1H, d, J = 8.5 Hz, H_3), 7.69 (1H, ddd, J = 7.7, 1.4, 1.4 Hz, $H_{4'}$), 7.52 (1H, dd, J = 7.7, 7.7 Hz, $H_{5'}$), 3.21 (1H, s, $H_{8'}$), 7.50 (1H, q, J = 7.4 Hz, H_8), 1.33 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.4 (C_7), 153.7 (C_4), 142.6 (C_5), 135.7

($C_{4'}$), 135.4 (C_1), 131.5 ($C_{2'}$), 129.2 ($C_{5'}$), 128.1 ($C_{6'}$), 126.5 ($C_{1'}$), 125.6 (C_2), 123.4 ($C_{3'}$), 121.2 (C_6), 111.5 (C_3), 82.2 ($C_{7'}$), 78.8 ($C_{8'}$), 51.0 (C_8), 7.6 (C_9); **HRMS** (CI^+) Calc. for $C_{17}H_{13}NO_3S$ [M^+] 311.0611, found 311.0608.

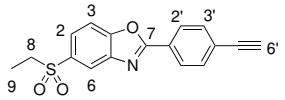
5-(Ethylsulfonyl)-2-(4-iodophenyl)benzo[*d*]oxazole 102



Benzoxazole **102** was obtained according to General Procedure A, using 4-iodobenzoic acid (300 mg, 1.21 mmol), thionyl chloride (0.26 mL, 3.63 mmol) and DMF (0.10 mL) in CH_2Cl_2 (15 mL). The acid chloride intermediate formed was reacted with 2-amino-4-ethylsulfonylphenol (244 mg, 1.21 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 4:1) to yield benzoxazole **102** (361 mg, 0.873 mmol, 72%) as a white solid.

R_f 0.16 (PE: EtOAc, 4:1); **ν_{max}** (cm^{-1}) 2970 (C-H), 2924 (C-H), 1616, 1479, 1309 (S=O), 1137 (S=O), 739; **mp** 170-171 ℃; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.34 (IH, d, J = 1.8 Hz, H_6), 8.00 (2H, m, $H_{2'}$), 7.95 (1H, dd, J = 1.8, 8.5 Hz, H_2), 7.93 (2H, m, $H_{3'}$), 7.75 (1H, d, J = 8.5 Hz, H_3), 3.18 (2H, q, J = 7.4 Hz, H_8), 1.30 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (100 MHz, $(CD_3)_2SO$) δ_C 164.1 (C_7), 153.2 (C_1), 141.8 (C_4), 138.4 ($C_{2'}$), 135.5 (C_5), 129.3 ($C_{3'}$), 125.6 (C_2), 125.1 ($C_{4'}$), 120.1 (C_6), 112.1 (C_3), 100.8 ($C_{1'}$), 49.5 (C_8), 7.25 (C_9); **HRMS** (ES^+) Calc. for $C_{15}H_{13}INO_3S$ $[M+H]^+$ 413.9655, found 413.9654.

5-(Ethylsulfonyl)-2-(4-ethynylphenyl)benzo[*d*]oxazole 93

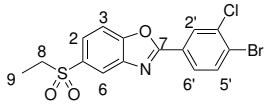


Benzoxazole **93** was obtained according to General Procedure A, using 4-ethynylbenzoic acid (250 mg, 1.71 mmol), thionyl chloride (0.37 mL, 5.13 mmol) and DMF (0.10 mL) in CH_2Cl_2 (15 mL). The acid chloride intermediate formed was reacted with 2-amino-4-ethylsulfonylphenol (344 mg, 1.71 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (Pentane : EtOAc, 4:1 → 7:3) to yield benzoxazole **93** (215 mg, 0.691 mmol, 40%) as a white solid.

R_f 0.12 (Pentane : EtOAc, 4:1); **ν_{max}** (cm^{-1}) 3260 (alkyne C-H), 1618, 1345 (S=O), 1136 (S=O), 1045, 744; **mp** 158-159 ℃; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.28 (1H, d, J = 1.6 Hz,

H_6), 8.18 (2H, J = 8.5 Hz, $H_{2'}$), 7.89 (1H, dd, J = 8.6, 1.6 Hz, H_2), 7.69 (1H, d, J = 8.6 Hz, H_3), 7.61 (2H, J = 8.5 Hz, $H_{3'}$), 3.22 (1H, s, $H_{6'}$), 3.12 (2H, q, J = 7.5 Hz, H_8), 1.24 (3H, t, J = 7.5 Hz, H_9); **^{13}CNMR** (100 MHz, $CDCl_3$) δ_C 164.6 (C_7), 153.8 (C_1), 142.7 (C_4), 135.4 (C_5), 132.8 ($C_{3'}$), 127.8 ($C_{2'}$), 126.3 ($C_{4'}$), 126.2 ($C_{1'}$), 125.6 (C_2), 121.2 (C_6), 111.5 (C_3), 82.7 ($C_{5'}$), 80.6 ($C_{6'}$), 51.0 (C_7), 7.6 (C_8); **HRMS** (ES^+) Calc. for $C_{17}H_{13}NO_3S$ [M+H^+] 312.0689, found 312.0689.

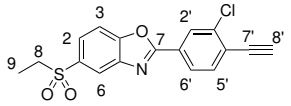
2-(4-Bromo-3-chlorophenyl)-5-(ethylsulfonyl)benzo[*d*]oxazole 175



Benzoxazole **175** was obtained according to General Procedure A, using 4-bromo-3-chlorobenzoic acid (300 mg, 1.27 mmol), thionyl chloride (0.28 mL, 3.82 mmol) and DMF (0.10 mL) in CH_2Cl_2 (15 mL). The acid chloride intermediate formed was reacted with 2-amino-4-ethylsulfonylphenol (256 mg, 1.27 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 4:1 → 3:1) to yield benzoxazole **175** (267 mg, 0.667 mmol, 52%) as a white solid.

R_f 0.47 (PE : EtOAc, 3:1); **ν_{max}** (cm^{-1}) 2956 (C-H), 1308 (S=O), 1138 (S=O), 739 (C-Cl), 668 (C-Br); **mp** 173-174 ℃; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.37 (1H, d, J = 2.0 Hz, $H_{2'}$), 8.35 (1H, d, J = 1.8 Hz, H_6), 8.02 (1H, dd, J = 8.4, 2.0 Hz, $H_{6'}$), 7.98 (1H, dd, J = 8.5, 1.8 Hz, H_2), 7.83 (1H, d, J = 8.4 Hz, H_3), 7.77 (1H, d, J = 8.5 Hz, $H_{5'}$), 3.18 (2H, q, J = 7.5 Hz, H_8), 1.31 (3H, t, J = 7.5 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.3 (C_7), 153.7 (C_4), 142.5 (C_5), 135.8 ($C_{3'}$), 135.7 (C_1), 134.7 (C_3), 129.4 ($C_{2'}$), 127.4 ($C_{1'}$), 126.9 ($C_{6'}$), 126.7 ($C_{4'}$), 125.9 (C_2), 121.4 (C_6), 111.6 ($C_{5'}$), 51.0 (C_8), 7.6 (C_9); **HRMS** (ES^+) Calc. for $C_{15}H_{12}{}^{79}Br^{35}ClNO_3S$ [M+H]$^+$ 399.9404, found 399.9404.

2-(3-Chloro-4-ethynylphenyl)-5-(ethylsulfonyl)benzo[*d*]oxazole 98

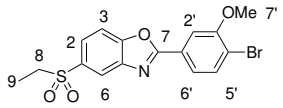


Alkyne **98** was obtained according to General Procedure B, using benzoxazole **175** (230 mg, 0.574 mmol), $Pd(PPh_3)_2Cl_2$ (60.4 mg, 86.1 µmol), copper(I) iodide (10.9 mg, 57.4 µmol), NEt_3 (3 mL) and trimethylsilylacetylene (80 µL, 0.574 mmol) in THF (20 mL). The TMS-protected alkyne intermediate formed was purified by flash column chromatography (PE :

EtOAc, 9:1 → 3:1), then deprotected with TBAF (1.56 mL, 1 M in THF) in THF (15 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 7:3) to yield alkyne **98** (123 mg, 0.356 mmol, 62%) as a white solid.

R_f 0.59 (PE : EtOAc, 7:3); **ν_{max}** (cm^{-1}) 3250 (alkyne C-H), 2956 (C-H), 2925 (C-H), 1310 (S=O), 1138 (S=O), 1051, 748; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.35 (1H, d, J = 1.7 Hz, H_6), 8.34 (1H, d, J = 1.6 Hz, $H_{2'}$), 8.13 (1H, dd, J = 8.0, 1.6 Hz, $H_{6'}$), 7.97 (1H, dd, J = 8.4, 1.7 Hz, H_2), 7.78 (1H, d, J = 8.4 Hz, H_3), 7.71 (1H, d, J = 8.0 Hz, $H_{5'}$), 3.58 (1H, s, $H_{8'}$), 3.19 (2H, q, J = 7.4 Hz, H_8), 1.31 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.5 (C_7), 153.9 (C_1), 142.6 (C_4), 137.4 ($C_{3'}$), 135.8 (C_5), 134.7 ($C_{5'}$), 128.6 ($C_{2'}$), 127.6 ($C_{4'}$), 126.2 ($C_{1'}$), 126.0 (C_2), 125.8 ($C_{6'}$), 121.6 (C_6), 111.7 (C_3), 85.8 ($C_{8'}$), 79.7 ($C_{7'}$), 51.2 (C_8), 7.7 (C_9); **HRMS** (ES^+) Calc. for $C_{17}H_{13}{}^{35}ClNO_3S$ $[M+H]^+$ 346.0299, found 346.0299.

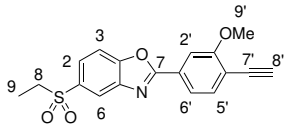
2-(4-Bromo-3-methoxyphenyl)-5-(ethylsulfonyl)benzo[*d*]oxazole 176



Benzoxazole **176** was obtained according to General Procedure A, using 4-bromo-3-methoxybenzoic acid (300 mg, 1.30 mmol), thionyl chloride (0.28 mL, 3.90 mmol) and DMF (0.10 mL) in CH_2Cl_2 (15 mL). The acid chloride intermediate formed was reacted with 2-amino-4-ethylsulfonylphenol (262 mg, 1.30 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 4:1) to yield benzoxazole **176** (401 mg, 1.01 mmol, 78%) as a white solid.

R_f 0.16 (PE : EtOAc, 4:1); **ν_{max}** (cm^{-1}) 1485, 1305 (S=O), 1240, 1126 (S=O), 1071, 690; **mp** 165-166 °C; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.34 (1H, d, J = 1.7 Hz, H_6), 7.95 (1H, dd, J = 8.5, 1.8 Hz, H_2), 7.81 (1H, s, $H_{2'}$), 7.78 (1H, dd, J = 8.5, 1.7 Hz, H_3), 7.74 (1H, s, $H_{6'}$), 7.73 (1H, s, $H_{5'}$), 4.05 (3H, s, $H_{7'}$), 3.18 (2H, q, J = 7.4 Hz, H_8), 1.31 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.6 (C_7), 156.5 ($C_{3'}$), 153.7 (C_4), 142.6 (C_5), 135.5 (C_1), 134.1 ($C_{5'}$), 126.5 ($C_{1'}$), 125.6 (C_2), 121.2 (C_6), 121.1 ($C_{6'}$), 117.1 ($C_{4'}$), 111.5 (C_3), 110.6 ($C_{2'}$), 56.6 ($C_{7'}$), 51.0 (C_8), 7.63 (C_9); **HRMS** (ES^+) Calc. for $C_{16}H_{15}{}^{79}BrNO_4S$ $[M+H]^+$ 395.9900, found 395.9901.

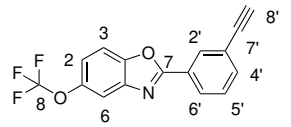
5-(Ethylsulfonyl)-2-(4-ethynyl-3-methoxyphenyl)benzo[*d*]oxazole 99



Alkyne **99** was obtained according to General Procedure B, using benzoxazole **176** (350 mg, 0.883 mmol), $Pd(PPh_3)_2Cl_2$ (93.0 mg, 0.132 mmol), copper(I) iodide (16.8 mg, 88.3 µmol), NEt_3 (3 mL) and trimethylsilylacetylene (0.12 mL, 0.882 mmol) in THF (20 mL). The TMS-protected alkyne intermediate formed was purified by flash column chromatography (PE : EtOAc, 4:1), then deprotected with TBAF (2.64 mL, 1 M in THF) in THF (15 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 4:1) to yield alkyne **99** (199 mg, 0.583 mmol, 66%) as a pale brown solid.

R_f 0.32 (PE : EtOAc, 4:1); **ν_{max}** (cm^{-1}) 2961 (C-H), 2875 (C-H), 1465, 1305 (S=O), 1137 (S=O), 1055; **mp** 161-163 °C; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.35 (1H, d, J = 1.8 Hz, H_6), 7.96 (1H, dd, J = 8.6, 1.8 Hz, H_2), 7.83 (1H, dd, J = 8.0, 1.6 Hz, $H_{6'}$), 7.79 (1H, d, J = 1.6 Hz, $H_{2'}$), 7.76 (1H, d, J = 8.6 Hz, H_3), 7.63 (1H, d, J = 8.0 Hz, $H_{5'}$), 4.06 (3H, s, $H_{7'}$), 3.50 (1H, s, J = 1.8 Hz, $H_{9'}$), 3.19 (2H, q, J = 7.3 Hz, H_8), 1.31 (3H, t, J = 7.3 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.8 (C_7), 161.0 ($C_{3'}$), 153.9 (C_1), 142.7 (C_4), 135.6 (C_5), 134.9 ($C_{5'}$), 127.6 ($C_{4'}$), 125.8 (C_2),121.3 (C_6), 120.2 ($C_{6'}$), 115.8 ($C_{1'}$), 111.6 (C_3), 109.7 ($C_{2'}$), 84.4 ($C_{9'}$), 79.4 ($C_{8'}$), 56.4 ($C_{7'}$), 51.2 (C_8), 7.8 (C_9); **HRMS** (ES^+) Calc. for $C_{18}H_{16}NO_4S$ $[M+H]^+$ 342.0795, found 342.0795.

2-(3-Ethynylphenyl)-5-(trifluoromethoxy)benzo[*d*]oxazole 94

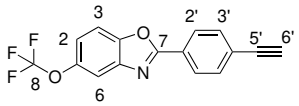


Benzoxazole **94** was obtained according to General Procedure A, using 3-ethynylbenzoic acid (362 mg, 2.48 mmol), thionyl chloride (0.54 mL, 7.44 mmol) and DMF (0.10 mL) in CH_2Cl_2 (20 mL). The acid chloride intermediate formed was reacted with 2-amino-4-trifluoromethoxy-phenol (479 mg, 2.48 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (Pentane : PhMe, 9:1 → 4:1) to yield benzoxazole **94** (37.7 mg, 0.124 mmol, 5%) as a white solid.

R_f 0.32 (Pentane : PhMe, 4:1); **ν_{max}** (cm^{-1}) 3231 (alkyne C-H), 1465, 1298, 1167; **mp**

85-86 ℃; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.37 (1H, t, J = 1.5 Hz, $H_{2'}$), 8.22 (1H, ddd, J = 7.9, 1.5, 1.5 Hz, $H_{6'}$), 7.67 (1H, ddd, J = 7.9, 1.5, 1.5 Hz, $H_{4'}$), 7.65 (1H, d, J = 1.2 Hz, H_6), 7.58 (1H, d, J = 8.8 Hz, H_3), 7.50 (1H, dd, J = 7.9, 7.9 Hz, $H_{5'}$), 7.25 (1H, ddd, J = 8.8, 2.4, 0.7 Hz, H_2), 3.17 (1H, s, $H_{8'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.9 (C_7), 149.1 (C_4), 146.2 (q, $^3J_{C\text{-}F}$ = 1.7, C_1), 142.8 (C_5), 135.3 ($C_{4'}$), 131.3 ($C_{2'}$), 129.1 ($C_{5'}$), 127.9 ($C_{6'}$), 127.0 ($C_{4'}$), 123.3 ($C_{1'}$), 120.6 (q, $^1J_{C\text{-}F}$ = 257.1, C_8), 119.0 (C_2), 113.3 (C_6), 111.1 (C_3), 82.4 ($C_{7'}$), 78.5 ($C_{8'}$); **HRMS** (ES^+) Calc. for $C_{16}H_9F_3NO_2$ [M+H^+] 304.0580, found 304.0581.

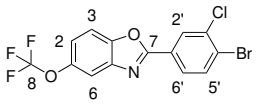
2-(4-Ethynylphenyl)-5-(trifluoromethoxy)benzo[*d*]oxazole 95



Benzoxazole **95** was obtained according to General Procedure A, using 4-ethynylbenzoic acid (438 mg, 3.00 mmol), thionyl chloride (0.65 mL, 9.00 mmol) and DMF (0.10 mL) in CH_2Cl_2 (20 mL). The acid chloride intermediate formed was reacted with 2-amino-4-trifluoromethoxyphenol (579 mg, 3.00 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (Pentane : EtOAc, 20:1) to yield benzoxazole **95** (373 mg, 1.23 mmol, 41%) as a white solid.

R_f 0.19 (Pentane); **ν_{max}** (cm^{-1}) 3313 (alkyne C-H), 1468, 1244, 1150, 1055; **mp** 143-144 ℃; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.21 (2H, d, J = 8.4 Hz, $H_{2'}$), 7.66 (2H, d, J = 8.4 Hz, $H_{3'}$), 7.65 (1H, m, H_2), 7.58 (1H, d, J = 8.8 Hz, H_3), 7.25 (1H, ddd, J = 8.8, 2.4, 0.7 Hz, H_4), 3.27 (1H, s, $H_{6'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.1 (C_7), 149.1 (C_4), 146.2 (q, $^3J_{C\text{-}F}$ = 1.9, C_1), 142.8 (C_5), 132.7 ($C_{2'}$), 127.6 ($C_{3'}$), 126.6 ($C_{4'}$), 125.8 ($C_{1'}$), 120.5 (q, $^1J_{C\text{-}F}$ = 256.9, C_8), 119.0 (C_2), 113.2 (C_6), 111.1 (C_3), 82.8 ($C_{5'}$), 80.2 ($C_{6'}$); **HRMS** (ES^+) Calc. for $C_{16}H_9F_3NO_2$ [M+H^+] 304.0580, found 304.0582.

2-(4-Bromo-3-chlorophenyl)-5-(trifluoromethoxy)benzo[*d*]oxazole 177

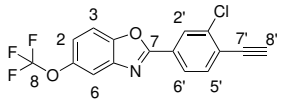


Benzoxazole **177** was obtained according to General Procedure A, using 4-bromo-3-chlorobenzoic acid (300 mg, 1.27 mmol), thionyl chloride (0.28 mL, 3.82 mmol) and DMF (0.10 mL) in CH_2Cl_2 (15 mL). The acid chloride intermediate formed was reacted with 2-amino-4-trifluoromethoxyphenol (245 mg, 1.27 mmol) in 1,4-dioxane (5 mL). The crude product was

purified by flash column chromatography (PE : EtOAc, 19:1) to yield benzoxazole **177** (379 mg, 0.966 mmol, 76%) as an orange solid.

R_f 0.38 (PE : EtOAc, 19:1); **ν_{max}** (cm^{-1}) 1543 (C-H), 1251 (C-F), 1164, 727, 620; **mp** 102-104 °C; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.34 (1H, d, J = 2.1 Hz, $H_{2'}$), 7.99 (1H, dd, J = 8.4, 2.1 Hz, $H_{6'}$), 7.81 (1H, d, J = 8.4 Hz, $H_{5'}$), 7.66 (1H, d, J = 2.0 Hz, H_6), 7.59 (1H, d, J = 8.9 Hz, H_3), 7.28 (1H, dd, J = 8.9, 2.0 Hz, H_2); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 162.9 (C_7), 149.1 (C_4), 146.4 (C_1), 142.7 (C_5), 135.7 ($C_{3'}$), 134.6 ($C_{5'}$), 129.2 ($C_{2'}$), 127.2 ($C_{4'}$), 126.9 ($C_{1'}$), 126.7 ($C_{6'}$), 120.7 (q, $^1J_{C\text{-}F}$ = 257.3, C_8), 119.5 (C_2), 113.4 (C_6), 111.3 (C_3); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −58.3 (s); **HRMS** (ES^+) Calc. for $C_{14}H_7{}^{79}Br^{35}ClF_3NO_2$ $[M+H]^+$ 391.9294, found 391.9295.

2-(3-Chloro-4-ethynylphenyl)-5-(trifluoromethoxy)benzo[*d*]oxazole 100

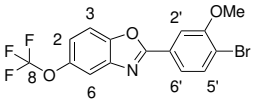


Alkyne **100** was obtained according to General Procedure B, using benzoxazole **177** (300 mg, 0.764 mmol), $Pd(PPh_3)_2Cl_2$ (73.0 mg, 0.104 mmol), copper(I) iodide (13.2 mg, 69.1 µmol), NEt_3 (3 mL) and trimethylsilylacetylene (0.96 mL, 0.691 mmol) in THF (20 mL). The TMS-protected alkyne intermediate formed was purified by flash column chromatography (PE : EtOAc, 1:0 → 49:1), then deprotected with TBAF (1.04 mL, 1 M in THF) in THF (15 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 49:1) to yield alkyne **100** (98.1 mg, 0.291 mmol, 42%) as a pale orange solid.

R_f 0.79 (PE : EtOAc, 49:1); **ν_{max}** (cm^{-1}) 3291 (alkyne C-H), 1260, 1165, 697; **mp** 97-98 °C; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.31 (1H, d, J = 1.7 Hz, $H_{2'}$), 8.10 (1H, dd, J = 8.2, 1.7 Hz, $H_{6'}$), 7.69 (1H, d, J = 8.2 Hz, $H_{5'}$), 7.66 (1H, d, J = 1.9 Hz, H_6), 7.60 (1H, d, J = 8.8 Hz, H_3), 7.28 (1H, dd, J = 8.8, 1.9 Hz, H_2), 3.56 (1H, s, $H_{8'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.0 (C_7), 149.3 (C_4), 146.5 (q, $^3J_{C\text{-}F}$ = 2.2 Hz, C_1), 142.8 (C_5), 137.3 ($C_{3'}$), 134.6 ($C_{5'}$), 128.4 ($C_{2'}$), 128.0 ($C_{1'}$), 125.6 ($C_{8'}$), 125.5 ($C_{6'}$), 120.7 (q, $^1J_{C\text{-}F}$ = 257.2 Hz, C_8), 119.6 (C_2), 113.6 (C_6), 111.4 (C_3), 85.5 ($C_{8'}$), 79.8 ($C_{7'}$); **^{19}FNMR** (400 MHz, $CDCl_3$) δ_F −58.3 (s); **HRMS** (ES^+) Calc. for $C_{16}H_8{}^{35}ClF_3NO_2$ $[M+H]^+$ 338.0190, found 338.0189.

2-(4-Bromo-3-methoxyphenyl)-5-(trifluoromethoxy)benzo[*d*]oxazole 178

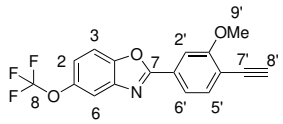
Benzoxazole **178** was obtained according to General Procedure A, using 4-bromo-3-methoxybenzoic acid (300 mg, 1.30 mmol), thionyl chloride (0.28 mL, 3.89 mmol) and DMF (0.10



mL) in CH_2Cl_2 (15 mL). The acid chloride intermediate formed was reacted with 2-amino-4-trifluoromethoxyphenol (251 mg, 1.30 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 49:1) to yield benzoxazole **178** (313 mg, 0.806 mmol, 62%) as an off-white solid.

R_f 0.61 (PE : EtOAc, 49:1); **ν_{max}** (cm^{-1}) 2970 (C-H), 1472 (C-H), 1259, 1167, 668; **mp** 140-141 °C; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 7.73 (1H, s, $H_{2'}$), 7.69 (1H, s, $H_{5'}$), 7.68 (1H, s, $H_{6'}$), 7.64 (1H, d, J = 2.0 Hz, H_6), 7.56 (1H, d, J = 8.9 Hz, H_3), 7.24 (1H, dd, J = 8.9, 2.0 Hz, H_2), 4.03 (3H, s, $H_{7'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.2 (C_7), 156.5 (C_5), 149.1 (C_4), 146.3 (C_1), 142.8 ($C_{3'}$), 134.1 ($C_{5'}$), 127.3 ($C_{4'}$), 121.1 ($C_{6'}$), 120.7 (q, $^1J_{C\text{-}F}$ = 257.2, C_8), 119.1 (C_2), 116.6 ($C_{1'}$), 113.3 (C_6), 111.2 (C_3), 110.5 ($C_{2'}$), 56.6 ($C_{7'}$); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −58.3 (s); **HRMS** (ES^+) Calc. for $C_{15}H_{10}{}^{79}BrF_3NO_3$ $[M+H]^+$ 387.9791, found 387.9792.

2-(4-Ethynyl-3-methoxyphenyl)-5-(trifluoromethoxy)benzo[*d*]oxazole 101

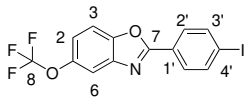


Alkyne **101** was obtained according to General Procedure B, using benzoxazole **178** (280 mg, 0.721 mmol), $Pd(PPh_3)_2Cl_2$ (69.1 mg, 98.4 µmol), copper(I) iodide (12.4 mg, 65.6 µmol), NEt_3 (3 mL) and trimethylsilylacetylene (92 µL, 0.656 mmol) in THF (20 mL). The TMS-protected alkyne intermediate formed was purified by flash column chromatography (PE : EtOAc, 1:0 → 49:1), then deprotected with TBAF (1.97 mL, 1 M in THF) in THF (15 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 99:1 → 49:1) to yield alkyne **101** (156 mg, 0.468 mmol, 71%) as an off-white solid.

R_f 0.93 (PE : EtOAc, 49:1); **ν_{max}** (cm^{-1}) 3271 (alkyne C-H), 2970 (C-H), 1253, 1217, 1153; **mp** 116-117 °C; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 7.81 (1H, dd, J = 7.9, 1.5 Hz, $H_{6'}$), 7.77 (1H, d, J = 1.6 Hz, $H_{2'}$), 7.66 (1H, d, J = 1.8 Hz, H_6), 7.62 (1H, d, J = 8.0 Hz, $H_{5'}$), 7.59 (1H, d, J = 8.8 Hz, H_3), 7.27 (1H, dd, J = 8.8, 1.8 Hz, H_2), 4.05 (3H, s, $H_{7'}$), 3.42 (1H, s, $H_{9'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.1 (C_7), 160.8 ($C_{3'}$), 149.1 (C_4), 146.2 (C_1), 142.7 (C_5), 134.7 ($C_{5'}$), 128.0 ($C_{1'}$), 120.6 (q, $^1J_{C\text{-}F}$ = 257.2 Hz, C_8), 119.8 ($C_{6'}$), 119.1 (C_2), 115.2 ($C_{4'}$), 113.2 (C_6), 111.1 (C_3), 109.4 ($C_{2'}$), 84.0 ($C_{9'}$), 79.4 ($C_{8'}$), 56.2 ($C_{7'}$); **^{19}FNMR** (400

MHz, $CDCl_3$) δ_F −58.3 (s); **HRMS** (ES^+) Calc. for $C_{17}H_{11}F_3NO_3$ $[M+H]^+$ 334.0685, found 334.0683.

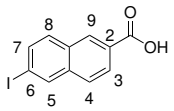
2-(4-Iodophenyl)-5-(trifluoromethoxy)benzo[*d*]oxazole 103



Benzoxazole **103** was obtained according to General Procedure A, using 4-iodobenzoic acid (250 mg, 1.01 mmol), thionyl chloride (0.22 mL, 3.03 mmol) and DMF (0.10 mL) in CH_2Cl_2 (15 mL). The acid chloride intermediate formed was reacted with 2-amino-4-trifluoromethoxyphenol (195 mg, 1.01 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 19:1) to yield benzoxazole **103** (209 mg, 0.515 mmol, 51%) as a pale pink solid.

R_f 0.75 (PE : EtOAc, 9:1); **ν_{max}** (cm^{-1}) 2963 (C-H), 1263 (C-F), 1029 (C-O); **mp** 144-145 ℃; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 7.96 (2H, m, $H_{3'}$), 7.91 (2H, m, $H_{2'}$), 7.64 (1H, d, J = 2.0 Hz, H_6), 7.58 (1H, d, J = 8.8 Hz, H_3), 7.25 (1H, dd, J = 8.8, 2.0 Hz, H_2); **^{13}CNMR** (125 MHz, $(CD_3)_2SO$) δ_C 163.9 (C_7), 148.9 (C_4), 145.4 (C_1), 142.3 (C_5), 138.4 ($C_{2'}$), 129.2 ($C_{3'}$), 125.4 ($C_{4'}$), 120.2 (q, $^1J_{C\text{-}F}$ = 256.2, C_8), 119.3 (C_2), 113.2 (C_6), 112.2 (C_3), 100.5 ($C_{1'}$); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −58.3 (s); **HRMS** (ES^+) Calc. for $C_{14}H_7F_3INO_2$ $[M+H]^+$ 405.9552, found 405.9542.

6-Iodo-2-naphthoic acid 179

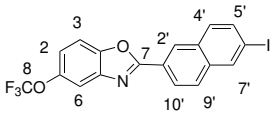


To a solution of 6-amino-2-naphthoic acid (500 mg, 2.67 mmol) in H_2SO_4 (10 mL, 10% aq.) at 0 ℃ was added dropwise a solution of sodium nitrite (221 mg, 3.21 mmol) in water (1 mL). The reaction was stirred for 30 min at rt. A solution of potassium iodide (886 mg, 5.34 mmol) in water (3 mL) cooled to 0 ℃ was then added dropwise and the reaction stirred for 3 h at 0 ℃. The reaction mixture was extracted with EtOAc (3 × 15 mL) and the combined organic layers washed with HCl (2 × 25 mL, 5% aq.) and $Na_2S_2O_3$ solution (25 mL), then dried (Na_2SO_4) and concentrated *in vacuo* to yield carboxylic acid **179** (675 mg, 2.66 mmol, 85%) as a brown solid.

mp: 301-303 °C; **^{1}HNMR** (400 MHz, $(CD_3)_2SO$) δ_H 8.60 (1H, d, J = 1.5 Hz, H_5), 8.50 (1H, d, J = 1.8 Hz, H_9), 8.00 (1H, dd, J = 8.6, 1.5 Hz, H_7), 7.96 (1H, d, J = 8.6 Hz, H_8), 7.92 (1H, d, J = 8.9 Hz, H_4), 7.87 (1H, dd, J = 8.9, 1.8 Hz, H_3); **LRMS** (ESI) Calc. for $C_{11}H_7IO_2$ $[M+H]^+$ 296.9, found 296.9.

These data are in accordance with those presented by Irvine *et al.*[355]

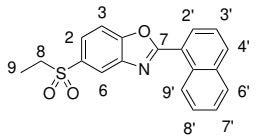
2-(6-Iodonaphthalen-2-yl)-5-(trifluoromethoxy)benzo[*d*]oxazole 104



Benzoxazole **104** was obtained according to General Procedure A, using carboxylic acid **179** (300 mg, 1.01 mmol), thionyl chloride (0.22 mL, 3.02 mmol) and DMF (0.10 mL) in CH_2Cl_2 (15 mL). The acid chloride intermediate formed was reacted with 2-amino-4-trifluoromethoxyphenol (193 mg, 1.01 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 49:1) to yield benzoxazole **104** (158 mg, 0.347 mmol, 35%) as a white solid.

R_f 0.37 (PE : EtOAc, 99:1); **ν_{max}** (cm^{-1}) 1248 (C-F), 1182 (C-O); **mp** 181-182 °C; **^{1}HNMR** (500 MHz, $(CD_3)_2SO$) δ_H 8.85 (1H, d, J = 1.6 Hz, H_6), 8.55 (1H, s, $H_{2'}$), 8.28 (1H, dd, J = 8.5, 1.6 Hz, H_2), 8.10 (1H, d, J = 8.7 Hz, H_3), 8.00 (1H, d, J = 8.7 Hz, $H_{9'}$), 7.96 (1H, d, J = 8.9 Hz, $H_{4'}$), 7.93 (1H, d, J = 1.7 Hz, $H_{7'}$), 7.92 (1H, dd, J = 8.9, 1.8 Hz, $H_{10'}$), 7.49 (1H, dd, J = 8.9, 2.0 Hz, $H_{5'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.7 (C_7), 149.5 (C_4), 145.8 (C_1), 142.9 (C_5), 136.8 ($C_{2'}$), 136.4 ($C_{3'}$), 136.1 ($C_{10'}$), 131.7 ($C_{1'}$), 131.3 ($C_{9'}$), 128.7 (C_6), 128.6 (C_3), 125.0 (C_2), 124.2 ($C_{8'}$), 120.7 (q, $^1J_{C\text{-}F}$ = 256.1, CF_3), 119.7 ($C_{5'}$), 113.7 ($C_{7'}$), 112.6 ($C_{4'}$), 96.1 ($C_{6'}$); **^{19}FNMR** (400 MHz, $CDCl_3$) δ_F −58.3 (s); **HRMS** (ES^+) Calc. for $C_{18}H_{10}F_3INO_2$ $[M+H]^+$ 455.9703, found 455.9699.

5-(Ethylsulfonyl)-2-(naphthalen-1-yl)benzo[*d*]oxazole 105

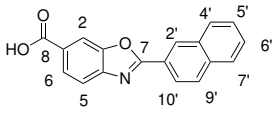


To a suspension of 2-amino-4-ethylsulfonylphenol (300 mg, 1.49 mmol) in 1,4-dioxane (5 mL) was added 1-naphthoyl chloride (0.23 mL, 1.49 mmol). The reaction vessel was heated under microwave activation at 210 °C for 15 min, then cooled to rt. The reaction mixture

was poured into NaOH solution (10 mL, 1 M aq.), and extracted with EtOAc three times. The combined organic fractions were dried (Na_2SO_4) and concentrated *in vacuo*. The crude product was purified by flash column chromatography (PE : EtOAc, 17:3 → 4:1) to yield benzoxazole **105** (496 mg, 1.47 mmol, 97%) as a white solid.

R_f 0.18 (PE : EtOAc, 4:1); **ν_{max}** (cm^{-1}) 1542, 1296 (S=O), 1278, 1243, 1132 (S=O), 1111, 769; **mp** 144-145 ℃; **^{1}HNMR** (500 MHz, $(CD_3)_2SO$) δ_H 9.40 (1H, d, J = 8.7 Hz, $H_{9'}$), 8.50 (1H, dd, J = 8.0, 1.1 Hz, $H_{7'}$), 8.43 (1H, d, J = 1.8 Hz, H_6), 8.28 (1H, d, J = 8.0 Hz, $H_{6'}$), 8.13 (1H, d, J = 8.5 Hz, H_3), 8.12 (1H, d, J = 7.7 Hz, $H_{4'}$), 8.00 (1H, dd, J = 8.5, 1.8 Hz, H_2), 7.77 (2H, m, $H_{2'}$ and $H_{8'}$), 7.70 (1H, dd, J = 7.7, 7.7 Hz, $H_{3'}$), 3.42 (2H, q, J = 7.4 Hz, H_8), 1.15 (3H, t, J = 7.4 Hz, H_9); **^{13}CNMR** (125 MHz, $(CD_3)_2SO$) δ_C 164.2 (C_7), 152.5 (C_1), 142.0 (C_4), 135.3 (C_5), 133.6 ($C_{5'}$), 133.4 ($C_{6'}$), 130.1 ($C_{7'}$), 129.8 ($C_{10'}$), 129.0 ($C_{4'}$), 128.4 ($C_{8'}$), 126.9 ($C_{3'}$), 125.6 (C_2), 125.5 ($C_{2'}$), 125.4 ($C_{9'}$), 121.7 ($C_{1'}$), 120.2 (C_6), 112.0 (C_3), 49.5 (C_8), 7.3 (C_9); **HRMS** (ES^+) Calc. for $C_{19}H_{16}NO_3S$ $[M+H]^+$ 338.0845, found 338.0844.

2-(Naphthalen-2-yl)benzo[*d*]oxazole-6-carboxylic acid 84

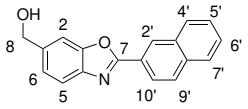


To a solution of 2-naphthaldehyde (1.02 g, 6.53 mmol) in *p*-xylenes (40 mL) was added 4-amino-3-hydroxybenzoic acid (1.00 g, 6.53 mmol). The reaction mixture was heated to 120 ℃, stirred for 4 h, then cooled to rt and concentrated *in vacuo*. The residue was dissolved in 1,4-dioxane (40 mL), DDQ (1.70 g, 7.49 mmol) was added and the reaction mixture stirred for 16 h at rt. The reaction mixture was then filtered and washed with CH_2Cl_2, and the filtrate collected and concentrated *in vacuo*. The crude product was purified by flash column chromatography (CH_2Cl_2: MeOH: AcOH, 98:2:0.1 → 95:5:0.1) to yield acid **84** (954 mg, 3.30 mmol, 44%) as a white solid.

R_f 0.13 (CH_2Cl_2: MeOH: AcOH, 99:1:0.1); **ν_{max}** (cm^{-1}) 2962 (COOH), 1698 (COOH), 1423, 1297, 821, 773; **mp** 258-259 ℃; **^{1}HNMR** (400 MHz, $(CD_3)_2SO$) δ_H 13.2 (1H, *br* s, COOH), 8.91 (1H, d, J = 1.2 Hz, $H_{2'}$), 8.31 (1H, d, J = 1.4 Hz, H_2), 8.30 (1H, dd, J = 8.6, 1.4 Hz, $H_{10'}$), 8.22 (1H, dd, J = 7.5, 1.2 Hz, $H_{4'}$), 8.17 (1H, d, J = 8.6 Hz, $H_{9'}$), 8.06 (1H, dd, J = 7.5, 1.5 Hz, $H_{7'}$), 8.05 (1H, dd, J = 8.4, 1.5 Hz, H_6), 7.94 (1H, d, J = 8.4 Hz, H_5), 7.68 (2H, m, $H_{5'}$, $H_{6'}$); **^{13}CNMR** (125 MHz, $(CD_3)_2SO$) δ_C 167.3 (C_8), 165.3 (C_7), 150.5 (C_1), 145.8 (C_3), 135.0 (C_4), 132.9 ($C_{3'}$), 129.6 ($C_{4'}$), 128.9 ($C_{9'}$), 128.8 ($C_{6'}$), 128.5 ($C_{2'}$), 128.3 ($C_{5'}$), 127.8 ($C_{8'}$), 127.7 (C_6), 126.8 ($C_{7'}$), 124.1 ($C_{10'}$), 123.7 ($C_{1'}$), 120.0 (C_5), 112.4 (C_2);

LRMS (ESI) Calc. for $C_{18}H_{11}NO_3$ $[M-H]^-$ 288.1, found 288.1.

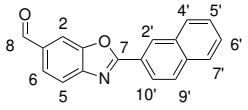
(2-(Naphthalen-2-yl)benzo[*d*]oxazol-6-yl)methanol 180



To a solution of acid **84** in THF (50 mL) at 0 ℃ was added $LiAlH_4$ (3.10 mL, 1 M in THF, 3.10 mmol) dropwise. When no more H_2 gas was evolved, the reaction mixture was heated under reflux for 3 h. When complete by TLC, the reaction was cooled to 0 ℃, quenched with ice water (30 mL) and EtOAc (30 mL). The phases were separated and the aqueous layer extracted with EtOAc (4 × 50 mL). The combined organic layers were washed with brine (50 mL), dried (Na_2SO_4) and concentrated *in vacuo* to yield alcohol **180** (558 mg, 2.03 mmol, 99%) as a brown solid.

R_f 0.24 (CH_2Cl_2:MeOH, 49:1); **ν_{max}** (cm^{-1}) 3302 (OH); **mp** 177-178 ℃; **^{1}HNMR** (400 MHz, $(CD_3)_2SO$) δ_H 8.84 (1H, d, J = 1.4 Hz, $H_{2'}$), 8.28 (1H, dd, J = 8.6, 1.4 Hz, $H_{10'}$), 8.20 (1H, dd, J = 6.8, 2.5 Hz, $H_{4'}$), 8.15 (1H, d, J = 8.6 Hz, $H_{9'}$), 8.05 (1H, dd, J = 6.8, 2.5 Hz, $H_{7'}$), 7.79 (1H, d, J = 8.2 Hz, H_5), 7.75 (1H, d, J = 1.5 Hz, H_2), 7.67 (2H, m, $H_{5'}$, $H_{6'}$), 7.39 (1H, dd, J = 8.2, 1.5 Hz, H_6), 5.41 (1H, t, J = 4.7 Hz, OH), 4.67 (2H, d, J = 4.7 Hz, H_8); **^{13}CNMR** (125 MHz, $(CD_3)_2SO$) δ_C 162.8 (C_7), 151.0 (C_3), 141.4 (C_4), 140.8 (C_1), 134.7 ($C_{3'}$), 133.0 ($C_{8'}$), 129.5 ($C_{9'}$), 129.4 ($C_{4'}$), 128.7 ($C_{6'}$), 128.3 ($C_{7'}$), 128.0 ($C_{2'}$), 127.8 ($C_{5'}$), 127.6 ($C_{1'}$), 124.2 (C_6), 124.0 ($C_{10'}$), 119.7 (C_5), 109.0 (C_2), 63.3 (C_8); **LRMS** (ESI) Calc. for $C_{18}H_{13}NO_2$ $[M-H]^-$ 274.1, found 274.1.

2-(Naphthalen-2-yl)benzo[*d*]oxazole-6-carbaldehyde 85

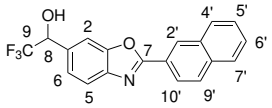


To a solution of alcohol **180** (2.82 g, 10.3 mmol) in wet CH_2Cl_2 (60 mL) was added Dess Martin periodinane (6.50 g, 15.4 mmol) portionwise. The reaction mixture was stirred at rt for 16 h, then diluted with CH_2Cl_2 (30 mL) and quenched with $NaHCO_3$ solution (30 mL) and $Na_2S_2O_3$ solution (30 mL). The organic phase was washed with water (40 mL) and brine (40 mL), dried (Na_2SO_4) and concentrated *in vacuo* to yield aldehyde **85** (2.10 g, 7.67 mmol, 75%) as a white solid.

R_f 0.22 (CH_2Cl_2); **ν_{max}** (cm^{-1}) 1683 (C=O), 1611, 1538, 1428, 815, 779; **mp** 171-172 ℃;

^{1}HNMR (400 MHz, $CDCl_3$) δ_H 10.05 (1H, s, H_8), 8.77 (1H, d, J = 1.0 Hz, $H_{2'}$), 8.28 (1H, dd, J = 8.6, 1.7 Hz, $H_{10'}$), 8.09 (1H, dd, J = 1.0, 1.0 Hz, H_2), 7.96 (1H, dd, J = 7.0, 2.0 Hz, $H_{4'}$), 7.95 (1H, d, J = 8.6 Hz, $H_{9'}$), 7.84-7.88 (3H, m, H_5, H_6, $H_{7'}$), 7.55 (2H, m, $H_{5'}$, $H_{6'}$); **^{13}CNMR** (100 MHz, CD_3Cl) δ_C 191.1 (C_8), 166.5 (C_7), 151.0 (C_3), 147.6 (C_4), 135.2 (C_1), 133.7 ($C_{8'}$), 132.9 ($C_{3'}$), 129.2 ($C_{2'}$), 129.1 ($C_{4'}$), 129.0 ($C_{9'}$), 128.4 ($C_{6'}$), 128.0 (C_6), 127.5 (C_5), 127.2 ($C_{5'}$), 123.9 ($C_{10'}$), 123.6 ($C_{1'}$), 120.3 ($C_{7'}$), 111.3 (C_2); **HRMS** (ES^+) Calc. for $C_{18}H_{12}NO_2$ [M+H^+] 274.0868, found 274.0868.

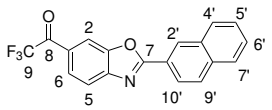
2,2,2-Trifluoro-1-(2-(naphthalen-2-yl)benzo[*d*]oxazol-6-yl)ethan-1-ol 181



To a solution of aldehyde **85** (2.00 g, 7.65 mmol) and trifluoromethyltrimethylsilane (1.47 mL, 9.95 mmol) in DMF (50 mL) was added K_2CO_3 (15 mg, cat.). The reaction mixture was stirred at rt for 16 h, then quenched with HCl (5 mL, 1 M aq.) and diluted with CH_2Cl_2 (100 mL). The mixture was washed with brine/water (1:1, 4 × 100 mL), dried (Na_2SO_4) and concentrated *in vacuo* to yield alcohol **181** (2.62 g, 7.65 mmol, quant.) as an orange solid.

R_f 0.16 (Pentane : EtOAc, 9:1); **ν_{max}** (cm^{-1}) 3208 (OH), 2959 (C-H), 1257 (C-F), 1127 (C-O); **mp** 168-169 °C; **^{1}HNMR** (400 MHz, $(CD_3)_2SO$) δ_H 8.86 (1H, d, J = 1.2 Hz, $H_{2'}$), 8.28 (1H, dd, J = 8.6, 1.7 Hz, $H_{10'}$), 8.20 (1H, dd, J = 7.0, 2.4 Hz, $H_{4'}$), 8.15 (1H, d, J = 8.6 Hz, $H_{9'}$), 8.05 (1H, dd, J = 7.0, 2.2 Hz, $H_{7'}$), 7.94 (1H, s, H_2), 7.88 (1H, d, J = 8.2 Hz, H_6), 7.67 (2H, m, $H_{5'}$, $H_{6'}$), 7.58 (1H, d, J = 8.2 Hz, H_5), 7.07 (1H, d, J = 5.6 Hz, -OH), 5.40 (1H, m, H_8); **^{13}CNMR** (125 MHz, CD_3Cl) δ_C 163.1 (C_7), 149.9 (C_3), 141.9 (C_4), 134.2 ($C_{3'}$), 133.6 (C_1), 132.4 ($C_{8'}$), 128.9 ($C_{9'}$), 128.8 ($C_{4'}$), 128.1 ($C_{6'}$), 127.8 ($C_{2'}$), 127.7 ($C_{7'}$), 127.1 ($C_{5'}$), 124.8 (q, $^1J_{C\text{-}F}$ = 283.0, C_9), 124.4 ($C_{1'}$), 124.3 (C_5), 123.4 ($C_{10'}$), 119.2 (C_6), 110.0 (C_2), 70.1 (q, $^2J_{C\text{-}F}$ = 30.5, C_8); **LRMS** (ESI) Calc. for $C_{19}H_{12}F_3NO_2$ $[M+H]^+$ 344.1, found 344.1.

2,2,2-Trifluoro-1-(2-(naphthalen-2-yl)benzo[*d*]oxazol-6-yl)ethan-1-one 86

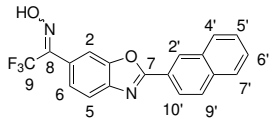


To a solution of alcohol **181** (2.60 g, 7.58 mmol) in MeCN (80 mL) was added Dess Martin periodinane (4.82 g, 11.36 mmol). The mixture was heated under reflux for 4 h, then

the reaction quenched with $NaHCO_3$ solution (40 mL) and $Na_2S_2O_3$ solution (40 mL), and diluted with CH_2Cl_2 (50 mL). The organic phase was washed with $NaHCO_3$ solution (50 mL) and brine (50 mL), dried (Na_2SO_4) and concentrated *in vacuo* to yield ketone **86** (2.02 g, 5.96 mmol, 79%) as a white solid.

ν_{max} (cm^{-1}) 1704 (C=O), 1596, 1210 (C-F), 1142, 766; **mp** 161-162 ℃; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.87 (1H, d, J = 1.4 Hz, $H_{2'}$), 8.37 (1H, s, H_2), 8.34 (1H, dd, J = 8.6, 1.4 Hz, $H_{10'}$), 8.17 (1H, d, J = 1.4 Hz, $H_{4'}$), 8.03 (1H, d, J = 8.4 Hz, $H_{9'}$), 8.03 (1H, m, H_6), 7.93 (1H, d, J = 8.4 Hz, H_5), 7.93 (1H, m, $H_{7'}$), 7.63 (2H, m, $H_{5'}$, $H_{6'}$); **^{13}CNMR** (125 MHz, $(CD_3)_2SO$) δ_C 179.7 (q, $^2J_{C\text{-}F}$ = 35.1, C_8), 167.5 (C_7), 150.8 (C_3), 148.5 (C_4), 135.5 ($C_{3'}$), 133.0 ($C_{8'}$), 129.6 (C_6), 129.4 ($C_{9'}$), 129.3 ($C_{2'}$), 128.7 ($C_{6'}$), 128.2 (C_5), 127.4 ($C_{4'}$), 127.2 ($C_{5'}$), 124.1 ($C_{1'}$), 123.4 ($C_{10'}$), 120.6 ($C_{7'}$), 117.0 (q, $^1J_{C\text{-}F}$ = 291.2, C_9), 112.9 (q, $^3J_{C\text{-}F}$ = 1.8, C_1), 112.2 (C_2); **LRMS** (ESI) Calc. for $C_{19}H_{10}F_3NO_2$ $[M+H]^+$ 342.1, found 342.1.

2,2,2-trifluoro-1-(2-(naphthalen-2-yl)benzo[d]oxazol-6-yl)ethan-1-one oxime 182



To a solution of ketone **86** (2.00g, 5.90 mmol) in pyridine (15 mL) and ethanol (120 mL) over molecular sieves (3 Å) was added hydroxylammonium chloride (1.23 g, 17.7 mmol). The reaction mixture was heated under reflux for 16 h, then concentrated *in vacuo*. The residue was dissolved in EtOAc (60 mL), washed with brine/water (1:1, 45 mL), dried (Na_2SO_4) and concentrated *in vacuo* to yield oxime **182** (1.91 g, 5.36 mmol, 91%) as a pale green solid in a mixture of isomers (5:6).

ν_{max} (cm^{-1}) 3060 (OH), 2924 (C-H), 1608 (C=N), 1255 (C-F), 951; **mp** 248-249 ℃;

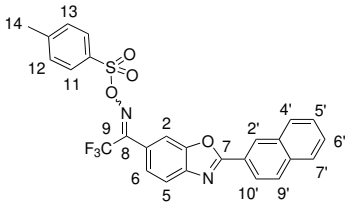
Major isomer: **R_f** 0.18 (Pentane : EtOAc, 9:1); **^{1}HNMR** (400 MHz, $(CD_3)_2SO$) δ_H 13.25 (1H, s, -OH), 8.87 (1H, br s, $H_{2'}$), 8.29 (1H, dd, J = 8.6, 1.2 Hz, $H_{10'}$), 8.21 (1H, d, J = 7.7 Hz, $H_{4'}$), 8.17 (1H, d, J = 8.7 Hz, $H_{9'}$), 8.06 (1H, d, J = 7.4, 1.4 Hz, $H_{7'}$), 8.02-7.97 (1H, m, H_2), 7.93 (1H, d, J = 8.4 Hz, H_5), 7.69 (2H, m, $H_{5'}$, $H_{6'}$), 7.56 (1H, d, J = 8.3 Hz, H_6); **^{13}CNMR** (125 MHz, $(CD_3)_2SO$) δ_C 164.5 (C_7), 150.6 (C_3), 144.9 (q, $^2J_{C\text{-}F}$ = 28.8, C_8), 143.4 (C_4), 134.9 ($C_{3'}$), 133.0 ($C_{8'}$), 129.6 ($C_{4'}$), 129.5 ($C_{9'}$), 128.9 ($C_{6'}$), 128.7 ($C_{2'}$), 128.7 ($C_{1'}$), 128.4 ($C_{7'}$), 127.8 ($C_{5'}$), 125.8 (C_6), 124.1 ($C_{10'}$), 123.9 (C_1), 120.5 (C_5), 118.9 (q, $^1J_{C\text{-}F}$ = 283.0, C_9), 112.2 (C_2)

Minor isomer: **R_f** 0.28 (Pentane : EtOAc, 9:1); **^{1}HNMR** (400 MHz, $(CD_3)_2SO$) δ_H 12.94

(1H, s, -OH), 8.87 (1H, br s, $H_{2'}$), 8.29 (1H, dd, J = 8.6, 1.2 Hz, $H_{10'}$), 8.21 (1H, d, J = 7.7 Hz, $H_{4'}$), 8.17 (1H, d, J = 8.7 Hz, $H_{9'}$), 8.06 (1H, d, J = 7.4, 1.4 Hz, $H_{7'}$), 8.02-7.97 (2H, m, H_2, H_5), 7.69 (2H, m, $H_{5'}$, $H_{6'}$), 7.51 (1H, dd, J = 8.3, 1.0 Hz, H_6); **^{13}CNMR** (125 MHz, $(CD_3)_2SO$) δ_C 164.4 (C_7), 150.5 (C_3), 144.7 (q, $^2J_{C\text{-}F}$ = 31.3, C_8), 143.4 (C_4), 134.9 ($C_{3'}$), 133.0 ($C_{8'}$), 129.6 ($C_{4'}$), 129.5 ($C_{9'}$), 128.9 ($C_{6'}$), 128.7 ($C_{2'}$), 128.7 ($C_{1'}$), 128.4 ($C_{7'}$), 127.8 ($C_{5'}$), 125.8 (C_6), 124.1 ($C_{10'}$), 123.8 (C_1), 120.3 (C_5), 118.9 (q, $^1J_{C\text{-}F}$ = 283.0, C_9), 111.6 (C_2);

HRMS (ESI) Calc. for $C_{19}H_{10}F_3N_2O_2$ $[M-H]^-$ 355.0700, found 355.0702.

2,2,2-Trifluoro-1-(2-(naphthalen-2-yl)benzo[*d*]oxazol-6-yl)ethan-1-one *O*-tosyl oxime 87



Tosyl oxime **87** was obtained according to General Procedure G, using oxime **182** (1.10 g, 3.09 mmol), *p*-toluenesulfonyl chloride (1.18 g, 6.17 mmol), 4-dimethylaminopyridine (37.9 mg, 0.31 mmol) and NEt_3 (1.28 mL, 9.27 mmol) in CH_2Cl_2 (25 mL). The crude product was purified by flash column chromatography (Pentane : CH_2Cl_2, 4:1 → 0:1) to yield tosyl oxime **87** (1.34 g, 2.63 mmol, 85%) as a pale yellow solid in a mixture of isomers (4:7).

ν_{max} (cm^{-1}) 2926 (C-H), 1612 (C=N), 1346 (S=O), 1246 (C-F), 1136 (S=O), 952; **mp** 171-172 ℃;

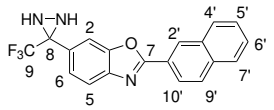
Major isomer: **R_f** 0.21 (Pentane : EtOAc, 9:1); **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.81 (1H, d, J = 1.0 Hz, $H_{2'}$), 8.31 (1H, dd, J = 8.7, 1.7 Hz, $H_{10'}$), 8.01 (1H, dd, J = 6.9, 1.6 Hz, $H_{4'}$), 8.01 (1H, d, J = 8.7 Hz, $H_{9'}$), 7.94 (2H, d, J = 8.3 Hz, H_{11}), 7.92 (1H, dd, J = 6.8, 1.7 Hz, $H_{7'}$), 7.82 (1H, dd, J = 8.3, 0.5 Hz, H_6), 7.77 (1H, d, J = 0.5 Hz, H_2), 7.61 (2H, m, $H_{5'}$ and $H_{6'}$), 7.48 (1H, d, J = 8.3 Hz, H_5), 7.40 (2H, d, J = 8.3 Hz, H_{12}), 2.48 (3H, s, H_{14}); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 165.6 (C_7), 153.6 (q, $^2J_{C\text{-}F}$ = 36.3, C_8), 150.5 (C_3), 145.3 (C_4), 135.1 ($C_{3'}$), 132.9 ($C_{8'}$), 131.4 (C_1), 130.0 (C_{12}), 129.2 (C_{11}), 129.1 ($C_{4'}$), 129.1 ($C_{9'}$), 129.0 ($C_{2'}$), 128.9 (C_{10}), 128.3 ($C_{6'}$), 128.0 ($C_{7'}$), 127.2 ($C_{5'}$), 125.8 (C_5), 124.3 ($C_{1'}$), 123.9 ($C_{10'}$), 123.5 (C_{13}), 120.2 (C_6), 117.4 (q, $^1J_{C\text{-}F}$ = 283.9, C_9), 111.6 (C_2), 21.8 (C_{14});

minor isomer: **R_f** 0.29 (Pentane : EtOAc, 9:1); **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.83 (1H, d, J = 1.2 Hz, $H_{2'}$), 8.32 (1H, dd, J = 8.7, 1.8 Hz, $H_{10'}$), 8.01 (1H, dd, J = 6.9, 1.6 Hz, $H_{4'}$),

(1H, d, J = 8.7 Hz, $H_{9'}$), 7.93 (1H, dd, J = 6.8, 1.7 Hz, $H_{7'}$), 7.92 (2H, d, J = 8.3 Hz, H_{11}), 7.89 (1H, dd, J = 8.3, 0.5 Hz, H_6), 7.76 (1H, d, J = 0.5 Hz, H_2), 7.61 (2H, m, $H_{5'}$ and $H_{6'}$), 7.44 (1H, d, J = 8.3 Hz, H_5), 7.42 (2H, d, J = 8.3 Hz, H_{12}), 2.50 (3H, s, H_{14});

^{13}CNMR (125 MHz, $CDCl_3$) δ_C 165.6 (C_7), 153.6 (q, $^2J_{C\text{-}F}$ = 36.3, C_8), 150.5 (C_3), 145.3 (C_4), 135.1 ($C_{3'}$), 132.9 ($C_{8'}$), 131.4 (C_1), 130.0 (C_{12}), 129.2 (C_{11}), 129.1 ($C_{4'}$), 129.1 ($C_{9'}$), 129.0 ($C_{2'}$), 128.9 (C_{10}), 128.3 ($C_{6'}$), 128.0 ($C_{7'}$), 127.2 ($C_{5'}$), 125.8 (C_5), 124.3 ($C_{1'}$), 123.9 ($C_{10'}$), 123.5 (C_{13}), 120.2 (C_6), 117.4 (q, $^1J_{C\text{-}F}$ = 283.9, C_9), 111.6 (C_2), 21.8 (C_{14}); **LRMS** (ESI) Calc. for $C_{26}H_{17}F_3N_2O_4S$ $[M+H]^+$ 511.1, found 511.1.

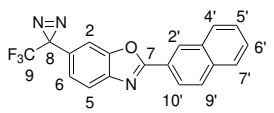
2-(Naphthalen-2-yl)-6-(3-(trifluoromethyl)diaziridin-3-yl)benzo[*d*]oxazole 183



Diaziridine **183** was obtained according to General Procedure H, using tosyl oxime **87** (2.24 g, 4.39 mmol) and NH_3 (15 mL) in CH_2Cl_2 (25 mL). The crude product was purified by flash column chromatography (Pentane : CH_2Cl_2, 1:1) to yield diaziridine **183** (1.19 g, 3.35 mmol, 76%) as white solid.

R_f 0.34 (Pentane : EtOAc, 4:1); **ν_{max}** (cm^{-1}) 3341 (N-H), 1482, 1252 (C-F), 1049; **mp** 145-147 °C; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.80 (1H, d, J = 1.4 Hz, $H_{2'}$), 8.32 (1H, dd, J = 8.6, 1.4 Hz, $H_{10'}$), 8.01 (1H, dd, J = 7.4, 2.5 Hz, $H_{4'}$), 8.00 (1H, d, J = 8.6 Hz, $H_{9'}$), 7.93 (1H, d, J = 1.0 Hz, H_2), 7.92 (1H, dd, J = 7.1, 2.4 Hz, $H_{7'}$), 7.85 (1H, dd, J = 8.3, 0.4 Hz, H_5), 7.67 (1H, dd, J = 8.3, 1.0 Hz, H_6), 7.60 (2H, m, $H_{5'}$ and $H_{6'}$), 2.91 (1H, d, J = 8.7 Hz, NH), 2.34 (1H, d, J = 8.7 Hz, NH); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.9 (C_7), 149.7 (C_3), 142.8 (C_4), 134.1 ($C_{3'}$), 132.0 ($C_{8'}$), 128.9 ($C_{2'}$), 128.4 ($C_{6'}$), 128.0 ($C_{5'}$), 127.9 (C_1), 127.7 ($C_{7'}$), 127.2 ($C_{9'}$), 127.0 ($C_{4'}$), 124.2 (C_6), 123.7 ($C_{10'}$), 122.8 ($C_{1'}$), 122.5 (q, $^1J_{C\text{-}F}$ = 278.5, C_9), 119.2 (C_5), 110.0 (C_2), 57.2 (q, $^2J_{C\text{-}F}$ = 36.3, C_8); **LRMS** (ESI) Calc. for $C_{19}H_{12}F_3N_3O$ $[M+H]^+$ 356.1, found 356.1.

2-(Naphthalen-2-yl)-6-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]oxazole 81

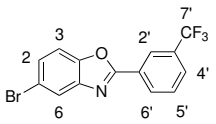


Diazirine **81** was obtained according to General Procedure I, using diaziridine **183** (220 mg, 0.623 mmol), iodine (190 mg, 0.748 mmol) and NEt_3 (0.26 mL, 1.87 mmol) in CH_2Cl_2 (10

mL). The crude product was purified by flash column chromatography (Pentane : CH_2Cl_2, 2:1) to yield diazirine **81** (141 mg, 0.399 mmol, 64%) as a yellow solid.

R_f 0.22 (Pentane : CH_2Cl_2, 4:1); **ν_{max}** (cm^{-1}) 1611, 1253 (C-F), 1187, 1152, 816, 753; **mp** 114-115 ℃; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.76 (1H, s, $H_{2'}$), 8.28 (1H, dd, J = 8.6, 1.7 Hz, $H_{10'}$), 7.98 (1H, dd, J = 6.7, 2.3 Hz, $H_{4'}$), 7.97 (1H, d, J = 8.6 Hz, $H_{9'}$), 7.90 (1H, dd, J = 6.9, 2.3 Hz, $H_{7'}$), 7.79 (1H, d, J = 8.3 Hz, H_5), 7.59 (2H, m, $H_{5'}$ and $H_{6'}$), 7.55 (1H, d, J = 1.0 Hz, H_2), 7.18 (1H, dd, J = 8.3, 1.0 Hz, H_6); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.8 (C_7), 149.7 (C_3), 142.5 (C_4), 133.9 ($C_{3'}$), 131.8 ($C_{8'}$), 128.0 ($C_{4'}$), 127.9 ($C_{9'}$), 127.6 ($C_{2'}$), 127.1 ($C_{6'}$), 126.9 ($C_{7'}$), 126.0 ($C_{5'}$), 124.9 (C_1), 122.8 ($C_{10'}$), 122.6 ($C_{1'}$), 122.1 (C_6), 121.1 (q, $^1J_{C\text{-}F}$ = 274.7, C_9), 119.3 (C_5), 108.5 (C_2), 27.7 (q, $^2J_{C\text{-}F}$ = 40.6, C_8); **HRMS** (ES^+) Calc. for $C_{19}H_{11}F_3N_3O$ $[M+H]^+$ 354.0854, found 354.0854.

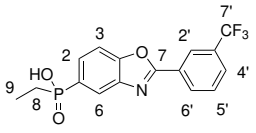
5-Bromo-2-(3-(trifluoromethyl)phenyl)benzo[*d*]oxazole 110



Benzoxazole **110** was obtained according to General Procedure A, using 3-trifluoromethoxybenzoyl chloride (0.30 mL 2.00 mmol) and 2-amino-4-bromophenol (376 mg, 2.00 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 9:1) to yield benzoxazole **110** (478 mg, 1.40 mmol, 70%) as a pale pink solid.

R_f 0.65 (PE : EtOAc, 9:1); **ν_{max}** (cm^{-1}) 1553 (aromatic C=C), 1449 (aromatic C=C), 1257 (C-F), 665 (C-Br); **mp** 102-104 ℃; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.20 (1H, ddd, J = 8.0, 1.1, 1.1 Hz, $H_{6'}$), 8.12 (1H, dd, J = 1.1, 1.1 Hz, $H_{2'}$), 7.95 (1H, dd, J = 1.8, 0.6 Hz, H_6), 7.60 (1H, dd, J = 8.0, 8.0 Hz, $H_{5'}$), 7.53 (1H, dd, J = 8.7, 1.8 Hz, H_2), 7.50 (1H, dd, J = 8.7, 0.6 Hz, H_3), 7.44 (1H, ddd, J = 8.0, 1.1, 1.1 Hz, $H_{4'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 162.6 (C_7), 147.9 (C_4), 149.6 (q, $^2J_{C\text{-}F}$ = 1.9 Hz, $C_{3'}$), 143.5 (C_5), 130.6 ($C_{5'}$), 128.7 (C_2), 128.6 ($C_{1'}$), 126.0 ($C_{6'}$), 124.3 ($C_{4'}$), 123.3 (C_6), 120.4 (q, $^1J_{C\text{-}F}$ = 258.2 Hz, $C_{7'}$), 120.2 ($C_{2'}$), 117.6 (C_1), 112.0 (C_3); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −57.9 (s); **LRMS** (ES^+) Calc. for $C_{14}H_7{}^{79}BrF_3NO$ $[M+H]^+$ 341.9, found 341.9.

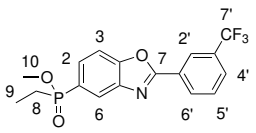
Ethyl(2-(3-(trifluoromethyl)phenyl)benzo[*d*]oxazol-5-yl)phosphinic acid 111



To a solution of **110** (250 mg, 0.731 mmol), ethyl phosphinic acid (137 mg, 1.46 mmol), Xantphos (21.1 mg, 36.6 µmol) and palladium(II) acetate (8.21 mg, 36.6 µmol) in DME (2 mL) and toluene (6 mL) was added DIPEA (0.25 mL, 1.46 mmol) dropwise. The reaction mixture was degassed, and stirred at 105 ℃ until complete. The solvent was removed *in vacuo* and the residue partitioned between EtOAc (20 mL) and NaOH (20 mL, 1 M aq.). The aqueous layer was extracted with EtOAc (3 × 10 mL), then acidified with HCl (1 M aq.) and extracted with EtOAc (3 × 20 mL). The organic layer was washed with brine (20 mL) and dried (Na_2SO_4) then concentrated *in vacuo*. The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 9:1) to yield phosphinic acid **111** (180 mg, 0.507 mmol, 69%) as a pale brown solid.

R_f 0.09 (CH_2Cl_2 : MeOH, 10:1); **ν_{max}** (cm^{-1}) 2980 (C-H), 1238 (C-F), 1126 (P=O), 956 (P-OC); **^{1}HNMR** (500 MHz, $(CD_3)_2SO$) δ_H 8.53 (1H, d, J = 8.0 Hz, $H_{6'}$), 8.46 (1H, s, $H_{2'}$), 8.15 (1H, dd, J = 11.5, 1.5 Hz, H_6), 8.06 (1H, d, J = 8.0 Hz, $H_{4'}$), 7.98 (1H, dd, J = 8.4, 2.1 Hz, H_3), 7.91 (1H, dd, J = 7.8, 7.8 Hz, $H_{5'}$), 7.83 (1H, ddd, J = 10.7, 8.4, 1.5 Hz, H_2), 3.61 (1H, *br* s, P-OH), 1.84 (2H, dq, J = 15.1, 7.7 Hz, H_8), 0.96 (3H, dt, J = 18.5, 7.7 Hz, H_9); **^{13}CNMR** (125 MHz, $(CD_3)_2SO$) δ_C 162.3 (C_7), 152.6 (C_4), 141.7 (C_5), 131.8 ($C_{6'}$), 131.7 (d, $^1J_{C\text{-}P}$ = 123.3 Hz, C_1), 130.6 (q, $^2J_{C\text{-}F}$ = 32.4 Hz, $C_{3'}$), 131.3 ($C_{5'}$), 127.6 ($C_{1'}$), 129.3 (d, $^2J_{C\text{-}P}$ = 11.2 Hz, C_2), 129.2 (q, $^3J_{C\text{-}F}$ = 3.7 Hz, $C_{4'}$), 124.2 (q, $^3J_{C\text{-}F}$ = 3.8 Hz, $C_{2'}$), 124.1 (q, $^1J_{C\text{-}F}$ = 272.5 Hz, $C_{7'}$), 123.5 (d, $^2J_{C\text{-}P}$ = 11.0 Hz, C_6), 111.9 (d, $^3J_{C\text{-}P}$ = 13.4 Hz, C_3), 23.6 (d, $^1J_{C\text{-}P}$ = 99.9 Hz, C_8), 6.7 (d, $^2J_{C\text{-}P}$ = 4.7 Hz, C_9); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −62.9 (s); **HRMS** (ES^+) Calc. for $C_{16}H_{14}F_3NO_3P$ $[M+H]^+$ 356.0658, found 356.0658.

Methyl ethyl(2-(3-(trifluoromethyl)phenyl)benzo[*d*]oxazol-5-yl)phosphinate 6

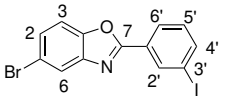


To a solution of phosphinic acid **111** (160 mg, 0.450 mmol) in thionyl chloride (2 mL) was added DMF (0.10 mL). The reaction mixture was heated under reflux for 2 h, con-

centrated *in vacuo* then re-dissolved in CH_2Cl_2 (2 mL) and cooled to 0 ℃. A solution of 4-dimethylaminopyridine (193 mg, 1.58 mmol), DIPEA (0.15 mL, 0.900 mmol), MeOH (2 mL) in CH_2Cl_2 (4 mL) was added dropwise and the reaction warmed to rt. After 3 h, the reaction mixture was concentrated *in vacuo*, dissolved in EtOAc (10 mL) and washed with HCl (5 mL, 1 M aq.), $NaHCO_3$ solution (5 mL) and brine (5 mL). The organic layer was dried (Na_2SO_4) and concentrated *in vacuo* to yield phosphinate **184** as an off-white solid (106 mg, 0.287 mmol, 63%).

R_f 0.73 (CH_2Cl_2 : MeOH, 10:1); **ν_{max}** (cm^{-1}) 2980 (C-H), 1461, 1382, 1251, 1152, 954; **mp** 90-91 ℃; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.51 (1H, dd, J = 1.8, 1.8 Hz, $H_{2'}$), 8.43 (1H, ddd, J = 7.9, 1.8, 1.8 Hz, $H_{6'}$), 8.17 (1H, dd, J = 11.6, 1.4 Hz, H_6), 7.84 (1H, ddd, J = 10.8, 8.3, 1.4 Hz, H_2), 7.80 (1H, ddd, J = 7.9, 1.8, 1.8 Hz, $H_{4'}$), 7.72 (1H, dd, J = 8.3, 2.5 Hz, H_3), 7.67 (1H, dd, J = 7.9, 7.9 Hz, $H_{5'}$), 3.66 (3H, d, J = 11.0 Hz, H_{10}), 2.04-1.86 (2H, m, H_8), 1.11 (3H, dt, J = 19.1, 7.7 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 162.8 (C_7), 153.3 (C_4), 142.3 (d, $^3J_{C\text{-}P}$ = 17.1 Hz, C_5), 131.8 (q, $^2J_{C\text{-}F}$ = 33.1 Hz, $C_{3'}$), 130.9 ($C_{6'}$), 129.8 ($C_{5'}$), 129.5 (d, $^2J_{C\text{-}P}$ = 11.1 Hz, C_2), 128.6 (q, $^3J_{C\text{-}F}$ = 3.6 Hz, $C_{4'}$), 127.5 ($C_{1'}$), 126.9 (d, $^1J_{C\text{-}P}$ = 122.7 Hz, C_1), 124.8 (q, $^3J_{C\text{-}F}$ = 3.8 Hz, $C_{2'}$), 124.2 (d, $^2J_{C\text{-}P}$ = 11.1 Hz, C_6), 123.7 (q, $^1J_{C\text{-}F}$ = 272.5 Hz, $C_{7'}$), 111.5 (d, $^3J_{C\text{-}P}$ = 14.3 Hz, C_3), 51.3 (d, $^2J_{C\text{-}P}$ = 6.7 Hz, C_{10}), 22.8 (d, $^1J_{C\text{-}P}$ = 103.1 Hz, C_8), 5.98 (d, $^2J_{C\text{-}P}$ = 5.0 Hz, C_9); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −62.9 (s); **HRMS** (ES^+) Calc. for $C_{17}H_{16}F_3NO_3P$ $[M+H]^+$ 370.0814, found 370.0815.

5-Bromo-2-(3-iodophenyl)benzo[*d*]oxazole 115



Benzoxazole **115** was synthesised by two methods:

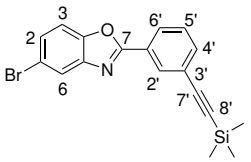
Method A, 1,3-diiodobenzene (835 mg, 2.53 mmol) was added to a solution of 5-bromobenzoxazole (250 mg, 1.26 mmol), palladium(II) acetate (7.07 mg, 31.5 µmol), PPh_3 (33.0 mg, 0.126 mmol), Cs_2CO_3 (823 mg, 2.53 mmol) and copper(I) iodide (6.00 mg, 31.5 µmol) in DMF (20 mL). The reaction mixture was degassed, stirred at 100 ℃ for 6 h then diluted with water (20 mL). The aqueous phase was extracted with CH_2Cl_2 (3 × 20 mL), and the combined organic layers concentrated *in vacuo*. The crude product was purified by flash column chromatography (PE : EtOAc, 49:1 → 19:1) to yield benzoxazole **115** (82.1 mg, 0.205 mmol, 13%) as an off-white solid.

Method B, according to General Procedure A, using 3-iodobenzoic acid (379 mg, 2.02

mmol), thionyl chloride (0.43 mL, 6.06 mmol) and DMF (0.10 mL) in CH_2Cl_2 (15 mL). The acid chloride intermediate formed was reacted with 2-amino-4-bromophenol (379 mg, 2.02 mmol) in 1,4-dioxane (5 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 19:1 → 25:2) to yield benzoxazole **115** (578 mg, 1.45 mmol, 72%) as an off-white solid.

R_f 0.7 (PE : EtOAc, 19:1); **ν_{max}** (cm^{-1}) 1542, 1445, 1161, 680 (C-Br); **mp** 158-159 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.62 (1H, dd, J = 1.7, 1.7 Hz, $H_{2'}$), 8.22 (1H, ddd, J = 7.8, 1.7, 1.1 Hz, $H_{6'}$), 7.93 (1H, dd, J = 1.8, 0.6 Hz, H_6), 7.91 (1H, ddd, J = 7.9, 1.8, 1.1 Hz, $H_{4'}$), 7.52 (1H, dd, J = 8.6, 1.8 Hz, H_2), 7.49 (1H, dd, J = 8.6, 0.6 Hz, H_3), 7.30 (1H, dd, J = 7.9, 7.9 Hz, $H_{5'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 162.6 (C_7), 149.9 (C_4), 143.7 (C_5), 140.9 ($C_{4'}$), 136.6 ($C_{2'}$), 130.7 ($C_{5'}$), 128.7 (C_2 and $C_{1'}$), 127.0 ($C_{6'}$), 123.3 (C_6), 117.7 (C_1), 112.1 (C_3), 94.5 ($C_{3'}$); **HRMS** (APCI) Calc. for $C_{13}H_8{}^{79}BrNOI$ $[M+H]^+$ 399.8829, found 399.8821.

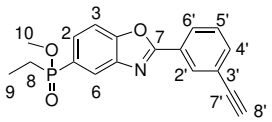
5-Bromo-2-(3-((trimethylsilyl)ethynyl)phenyl)benzo[*d*]oxazole 116



To a solution of aryl iodide **115** (560 mg, 1.40 mmol), $Pd(PPh_3)_2Cl_2$ (147 mg, 0.210 mmol) and copper(I) iodide (26.7 mg, 0.140 mmol) in THF was added NEt_3 (2.00 mL) and trimethylsilylacetylene (0.20 mL, 1.40 mmol) dropwise. The reaction mixture was degassed, and stirred at 80 ℃ for 30 min. The solvent was removed *in vacuo* and the residue purified by flash column chromatography (PE : EtOAc, 1:0 → 99:1) to yield alkyne **116** (467 mg, 1.26 mmol, 90%) as a white solid.

R_f 0.77 (PE : EtOAc, 199:1); **ν_{max}** (cm^{-1}) 2980 (C-H), 2927 (C-H), 2161 (C≡C), 1544, 1470, 1250, 680; **mp** 138-140 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.35 (1H, dd, J = 1.4, 1.4 Hz, $H_{2'}$), 8.17 (1H, ddd, J = 7.7, 1.4, 1.4 Hz, $H_{6'}$), 7.90 (1H, dd, J = 1.7, 0.6 Hz, H_6), 7.63 (1H, ddd, J = 7.7, 1.4, 1.4 Hz, $H_{4'}$), 7.49-7.45 (3H, *obs* m, $H_{5'}$, H_2 and H_3), 0.28 (9H, s, -$SiMe_3$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.5 (C_7), 149.9 (C_4), 143.8 (C_5), 135.1 ($C_{4'}$), 131.4 ($C_{2'}$), 129.1 ($C_{5'}$), 128.5 (C_2), 127.6 ($C_{6'}$), 127.0 ($C_{1'}$), 124.4 ($C_{3'}$), 123.3 (C_6), 117.6 (C_1), 112.0 (C_3), 103.8 ($C_{7'}$), 96.0 ($C_{8'}$), 0.03 (-$SiMe_3$); **HRMS** (APCI) Calc. for $C_{18}H_{17}{}^{79}BrNOSi$ $[M+H]^+$ 370.0257, found 370.0251.

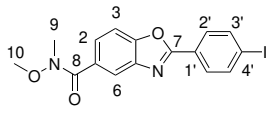
Methyl ethyl(2-(3-ethynylphenyl)benzo[*d*]oxazol-5-yl)phosphinate 118



To a solution of **116** (200 mg, 0.540 mmol), ethyl phosphinic acid (102 mg, 1.08 mmol), Xantphos (15.6 mg, 27.0 µmol) and palladium(II) acetate (6.06 mg, 27.0 µmol) in DME (2 mL) and toluene (10 mL) was added DIPEA (0.19 mL, 1.08 mmol) dropwise. The reaction mixture was degassed, and stirred at 105 ℃ until complete. The solvent was removed *in vacuo* and the residue partitioned between CH_2Cl_2 (20 mL) and NaOH (20 mL, 1 M aq.). The aqueous layer was extracted with CH_2Cl_2 (3 × 10 mL), then acidified with HCl (1 M aq.) and extracted with CH_2Cl_2 (3 × 20 mL). The organic layer was concentrated *in vacuo* and the residue purified by reverse phase flash column chromatography (MeCN : 0.1% formic acid in H_2O, 1:19 → 1:1). To a solution of the phosphinic acid intermediate in methanol (4 mL) and toluene (6 mL) was added TMS-diazomethane (0.68 mL, 2 M in hexane, 1.35 mmol) dropwise. The reaction mixture was stirred for 30 min before dropwise addition of acetic acid (5 mL) and concentration *in vacuo*. The crude product was purified by flash column chromatography (EtOAc) and recrystallised from Et_2O to yield alkyne **118** (46 mg, 0.141 mmol, 26%) as cream needle crystals.

R_f 0.068 (EtOAc); **ν_{max}** (cm^{-1}) 2980 (C-H), 1472, 1382, 1249, 1151 (P=O), 954 (P-OCH_3); **mp** 113-114 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.39 (1H, dd, J = 1.6, 1.6 Hz, $H_{2'}$), 8.24 (1H, ddd, J = 7.8, 1.4, 1.4 Hz, $H_{6'}$), 8.16 (1H, ddd, J = 11.6, 1.4, 0.7 Hz, H_6), 7.84 (1H, ddd, J = 10.9, 8.2, 1.4 Hz, H_2), 7.71 (1H, ddd, J = 8.2, 2.4, 0.7 Hz, H_3), 7.67 (1H, ddd, J = 7.8, 1.4, 1.4 Hz, $H_{4'}$), 7.51 (1H, dd, J = 7.8, 7.8 Hz, $H_{5'}$), 3.67 (3H, d, J = 11.0 Hz, H_{10}), 3.18 (1H, s, $H_{8'}$), 2.07-1.90 (2H, m, H_8), 1.13 (3H, dt, J = 19.1, 7.7 Hz, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.4 (C_7), 153.4 (d, $^4J_{C\text{-}P}$ = 2.8 Hz, C_4), 142.5 (d, $^3J_{C\text{-}P}$ = 17.1 Hz, C_5), 135.5 ($C_{4'}$), 131.5 ($C_{2'}$), 129.4 (d, $^2J_{C\text{-}P}$ = 11.2 Hz, C_2), 129.3 ($C_{5'}$), 128.1 ($C_{6'}$), 127.0 ($C_{1'}$), 126.7 (d, $^1J_{C\text{-}P}$ = 122.3 Hz, C_1), 124.3 (d, $^2J_{C\text{-}P}$ = 10.9 Hz, C_6), 123.3 ($C_{3'}$), 111.4 (d, $^3J_{C\text{-}P}$ = 14.1 Hz, C_3), 82.5 ($C_{7'}$), 78.7 ($C_{8'}$), 51.3 (d, $^2J_{C\text{-}P}$ = 6.7 Hz, C_{10}), 22.9 (d, $^1J_{C\text{-}P}$ = 103.1 Hz, C_8), 6.1 (d, $^2J_{C\text{-}P}$ = 4.8 Hz, C_9); **HRMS** (ES^+) Calc. for $C_{18}H_{17}NO_3P$ $[M+H]^+$ 326.0941, found 326.0940.

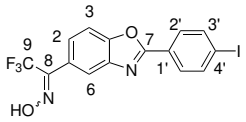
2-(4-Iodophenyl)-*N*-methoxy-*N*-methylbenzo[*d*]oxazole-5-carboxamide 120



Carboxamide **120** was obtained according to General Procedure E, using 4-iodobenzoic acid (500 mg, 2.01 mmol), thionyl chloride (0.44 mL, 6.03 mmol) and DMF (0.10 mL) in CH_2Cl_2 (20 mL). The acid chloride intermediate formed was added to a solution of 3-amino-4-hydroxybenzoic acid (277 mg, 1.81 mmol, 0.9 eq.) in 1,4-dioxane (5 mL). Amide coupling was performed on the crude product using EDC hydrochloride (769 mg, 4.02 mmol), *N*,*O*-dimethylhydroxylamine hydrochloride (392 mg, 4.02 mmol) and NEt_3 (0.56 mL, 4.01 mmol) in DMF (25 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 3:1) to yield carboxamide **120** (281 mg, 0.688 mmol, 38%) as a white solid.

R_f 0.11 (PE : EtOAc, 9:1); **ν_{max}** (cm^{-1}) 2980 (C-H), 1636 ((C=O)NH), 1396, 1261, 1060; **mp** 143-144 ℃; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.13 (1H, d, J = 1.6 Hz, H_6), 7.95 (2H, m, $H_{2'}$), 7.87 (2H, m, $H_{3'}$), 7.75 (1H, dd, J = 1.6, 8.6 Hz, H_2), 7.58 (1H, d, J = 8.6 Hz, H_3), 3.55 (3H, s, H_{10}), 3.39 (3H, s, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 169.2 (C_8), 163.4 (C_7), 152.1 (C_4), 141.7 (C_5), 138.4 ($C_{3'}$), 130.9 (C_1), 129.2 ($C_{2'}$), 126.5 (C_2), 126.3 ($C_{1'}$), 120.7 (C_6), 110.4 (C_3), 99.1 ($C_{4'}$), 61.2 (C_{10}), 33.9 (C_9); **HRMS** (ES^+) Calc. for $C_{16}H_{14}IN_2O_3$ $[M+H]^+$ 409.0043, found 409.0041.

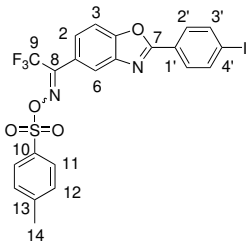
2,2,2-Trifluoro-1-(2-(4-iodophenyl)benzo[*d*]oxazol-5-yl)ethan-1-one oxime 185



Oxime **185** was obtained according to General Procedure F, using carboxamide **120** (1.55 g, 3.81 mmol), trifluoromethyltrimethylsilane (1.13 mL, 7.62 mmol) and caesium fluoride (116 mg, 0.762 mmol) in toluene (50 mL). The reaction was stirred for three days, with further trifluoromethyltrimethylsilane (1.13 mL, 7.62 mmol) and caesium fluoride (116 mg, 0.762 mmol) added every 24 h. TBAF (4.57 mL, 1 M in THF) was used for desilylation of the intermediate formed. The crude trifluoromethylketone product was dissolved in pyridine (5 mL) and ethanol (25 mL) and refluxed after addition of hydroxylammonium chloride (794 mg, 11.4 mmol). The crude product was purified by flash column chromatography (PE : EtOAc, 9:1 → 4:1) to yield oxime **185** (527 mg, 1.22 mmol, 32%) as a pale green solid.

R_f 0.69 (PE : EtOAc, 4:1); **ν_{max}** (cm^{-1}) 1476 (aromatic C=C), 1259 (C-F), 1151 (C-F), 970 (N-O); **mp** 197-199 °C; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 7.97 (2H, m, $H_{3'}$), 7.95 (1H, d, J = 1.2 Hz, H_6), 7.91 (2H, m, $H_{2'}$), 7.68 (1H, d, J = 8.4 Hz, H_3), 7.51 (1H, dd, J = 8.4, 1.2 Hz, H_2); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 165.1 (C_7), 152.9 (C_4), 145.1 (q, $^2J_{C\text{-}F}$ = 32.6 Hz, C_8), 143.3 (C_5), 139.9 ($C_{3'}$), 130.4 ($C_{2'}$), 128.0 (C_1), 127.5 (C_2), 125.8 ($C_{1'}$), 121.9 (C_6), 120.9 (q, $^1J_{C\text{-}F}$ = 271.9 Hz, C_9), 112.3 (C_3), 100.3 ($C_{4'}$); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −67.5 (s); **HRMS** (ES^+) Calc. for $C_{15}H_7F_3IN_2O_2$ $[M-H]^-$ 430.9510, found 430.9513.

2,2,2-Trifluoro-1-(2-(4-iodophenyl)benzo[*d*]oxazol-5-yl)ethan-1-one *O*-tosyl oxime 124



Tosyl oxime **124** was obtained according to General Procedure G, using oxime **185** (1.80 g, 4.17 mmol), *p*-toluenesulfonyl chloride (1.59 g, 8.33 mmol), 4-dimethylaminopyridine (50.9 mg, 0.417 mmol) and NEt_3 (1.73 mL, 12.5 mmol) in CH_2Cl_2 (50 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 19:1 → 4:1) to yield tosyl oxime **124** (1.35 g, 2.30 mmol, 55%) as a pale yellow solid in a mixture of isomers (2:1).

ν_{max} (cm^{-1}) 2981 (C-H), 1622 (C=N-O), 1394 (S=O), 1245 (C-F), 1194 (S=O), 1159; **mp** 171-173 °C;

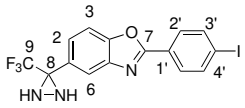
Major isomer: **R_f** 0.45 (PE : EtOAc, 17:3); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 7.97 (2H, m, $H_{3'}$), 7.92 (2H, m, $H_{2'}$), 7.90 (2H, d, J = 8.1 Hz, H_{11}), 7.81 (1H, d, J = 1.6 Hz, H_6), 7.68 (1H, d, J = 8.5 Hz, H_3), 7.42 (1H, dd, J = 8.5, 1.6 Hz, H_2), 7.40 (2H, d, J = 8.1 Hz, H_{12}), 2.49 (3H, s, H_{14}). **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.9 (C_7), 153.6 (q, $^2J_{C\text{-}F}$ = 33.5 Hz, C_8), 152.3 (C_4), 146.3 (C_{13}), 142.3 (C_5), 138.4 ($C_{3'}$), 131.1 (C_{10}), 129.9 (C_{12}), 129.3 (C_{11}), 129.2 ($C_{2'}$), 125.9 (C_2), 125.8 (C_{10}), 121.1 (C_1), 120.9 (C_6), 119.6 (q, $^1J_{C\text{-}F}$ = 281.3 Hz, C_9), 111.4 (C_3), 99.4 ($C_{4'}$), 21.8 (C_{14}); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −66.7 (s).

Minor isomer: **R_f** 0.39 (PE : EtOAc, 17:3); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 7.96 (2H, m, $H_{3'}$), 7.91 (2H, m, $H_{2'}$), 7.89 (2H, d, J = 8.1 Hz, H_{11}), 7.85 (1H, d, J = 1.6 Hz, H_6), 7.63 (1H, d, J = 8.5 Hz, H_3), 7.48 (1H, dd, J = 8.5, 1.6 Hz, H_2), 7.39 (2H, d, J = 8.1 Hz, H_{12}), 2.48 (3H, s, H_{14}). **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.9 (C_7), 153.6 (q, $^2J_{C\text{-}F}$ = 33.5 Hz,

C_8), 152.7 (C_4), 146.1 (C_{13}), 142.3 (C_5), 138.4 ($C_{3'}$), 131.4 (C_{10}), 123.0 (C_{12}), 129.3 (C_{11}), 129.1 ($C_{2'}$), 125.9 (C_2), 125.8 (C_{10}), 121.2 (C_1), 120.9 (C_6), 119.6 (q, $^1J_{C\text{-}F}$ = 281.3 Hz, C_9), 111.1 (C_3), 99.5 ($C_{4'}$), 21.8 (C_{14}); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −66.7 (s);

HRMS (ES^+) Calc. for $C_{22}H_{15}F_3IN_2O_4S$ $[M+H]^+$ 586.9744, found 586.9741.

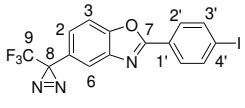
2-(4-Iodophenyl)-5-(3-(trifluoromethyl)diaziridin-3-yl)benzo[*d*]oxazole 186



Diaziridine **186** was obtained according to General Procedure H, using tosyl oxime **124** (1.82 g, 3.11 mmol) and NH_3 (5 mL) in CH_2Cl_2 (25 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 4:1) to yield diaziridine **186** (1.05 g, 2.43 mmol, 78%) as a pale yellow solid.

R_f 0.39 (PE : EtOAc, 4:1); **ν_{max}** (cm^{-1}) 1593 (N-H), 1479, 1397, 1216 (C-F), 1150; **mp** 152-154 ℃; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.05 (1H, d, J = 1.5 Hz, H_6), 7.97 (2H, m, $H_{3'}$), 7.90 (2H, m, $H_{2'}$), 7.66 (1H, dd, J = 8.5, 1.5 Hz, H_2), 7.62 (1H, d, J = 8.5 Hz, H_3), 2.89 (1H, d, J = 8.9 Hz, -NH), 2.32 (1H, d, J = 8.9 Hz, -NH); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.0 (C_7), 152.0 (C_4), 142.7 (C_5), 138.7 ($C_{2'}$), 129.5 ($C_{3'}$), 129.0 (C_1), 126.5 ($C_{1'}$), 125.8 (C_2), 123.5 (q, $^1J_{C\text{-}F}$ = 277.6 Hz, C_9), 120.9 (C_6), 111.4 (C_3), 99.5 ($C_{4'}$), 58.2 (q, $^2J_{C\text{-}F}$ = 36.2 Hz, C_8); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −75.5 (s); **HRMS** (ES^+) Calc. for $C_{15}H_{10}F_3IN_3O$ $[M+H]^+$ 431.9815, found 431.9815.

2-(4-Iodophenyl)-5-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]oxazole 126

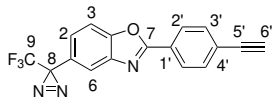


Diazirine **126** was obtained according to General Procedure I, using diaziridine **186** (580 mg, 1.35 mmol), iodine (409 mg, 1.61 mmol) and NEt_3 (0.56 mL, 4.04 mmol) in CH_2Cl_2 (25 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 19:1 → 9:1) to yield diazirine **126** (564 mg, 1.31 mmol, 98%) as a white solid.

R_f 0.84 (PE : EtOAc, 9:1); **ν_{max}** (cm^{-1}) 1478, 1397, 1252 (C-F), 1150, 810; **mp** 117-119 ℃; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 7.95 (2H, m, $H_{3'}$), 7.89 (2H, m, $H_{2'}$), 7.71 (1H, d, J = 1.8 Hz, H_6), 7.59 (1H, d, J = 8.5 Hz, H_3), 7.20 (1H, dd, J = 8.5, 1.8 Hz, H_2); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.0 (C_7), 151.5 (C_4), 142.7 (C_5), 138.5 ($C_{2'}$), 129.3 ($C_{3'}$), 126.1 ($C_{1'}$),

125.9 (C_1), 124.0 (C_2), 122.2 (q, $^1J_{C\text{-}F}$ = 274.7 Hz, C_9), 119.3 (C_6), 111.4 (C_3), 99.4 ($C_{4'}$). 28.7 (q, $^2J_{C\text{-}F}$ = 40.7 Hz, C_8); **^{19}FNMR** (376 MHz, CD_3OD) δ_F −67.4 (s); **HRMS** (ES^+) Calc. for $C_{15}H_8F_3IN_3O$ $[M+H]^+$ 429.9659, found 429.9655.

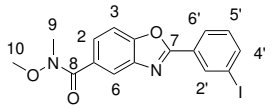
2-(4-Ethynylphenyl)-5-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]oxazole 106



Diazirine **106** was obtained according to General Procedure B, using diazirine **126** (100 mg, 0.233 mmol), $Pd(PPh_3)_2Cl_2$ (24.5 mg, 35.0 µmol), copper(I) iodide (4.44 mg, 23.3 µmol), NEt_3 (0.70 mL) and trimethylsilylacetylene (39 µL, 0.280 mmol) in THF (10 mL). The reaction was complete after 1 h. The TMS-protected alkyne intermediate formed was purified by flash column chromatography (PE : EtOAc, 1:0 → 49:1), then deprotected with TBAF (0.70 mL, 1 M in THF) in THF (5 mL). The reaction was complete after 30 min. The crude product was purified by flash column chromatography (PE : EtOAc, 1:0 → 49:1) to yield alkyne **106** (38.2 mg, 0.117 mmol, 50%) as a white solid.

R_f 0.69 (PE : EtOAc, 49:1); **ν_{max}** (cm^{-1}) 3311 (sharp, alkyne C-H), 1262 (C-F), 1149, 810, 740; **mp** 125-127 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.20 (2H, m, $H_{2'}$), 7.72 (1H, d, J = 1.8 Hz, H_6), 7.64 (2H, m, $H_{3'}$), 7.60 (1H, d, J = 8.6 Hz, H_3), 7.20 (1H, dd, J = 8.6, 1.8 Hz, H_2), 3.26 (1H, s, $H_{6'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.9 (C_7), 151.5 (C_4), 142.8 (C_5), 132.9 ($C_{3'}$), 127.8 ($C_{2'}$), 126.7 ($C_{4'}$), 126.0 ($C_{1'}$), 125.9 (C_1), 124.0 (C_2), 122.3 (q, $^1J_{C\text{-}F}$ = 274.7 Hz, C_9), 119.4 (C_6), 111.4 (C_3), 82.9 ($C_{5'}$), 80.4 ($C_{6'}$), 28.8 (q, $^2J_{C\text{-}F}$ = 40.6 Hz, C_8); **^{19}FNMR** (376 MHz, CD_3OD) δ_F −67.4 (s); **HRMS** (ES^+) Calc. for $C_{17}H_9F_3N_3O$ $[M+H]^+$ 328.0692, found 328.0691.

2-(3-Iodophenyl)-*N*-methoxy-*N*-methylbenzo[*d*]oxazole-5-carboxamide 119

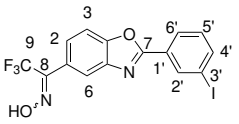


Carboxamide **119** was obtained according to General Procedure E, using 3-iodobenzoic acid (486 mg, 1.96 mmol), thionyl chloride (0.43 mL, 5.88 mmol) and DMF (0.10 mL) in CH_2Cl_2 (20 mL). The acid chloride intermediate formed was reacted with 3-amino-4-hydroxybenzoic acid (300 mg, 1.96 mmol) in 1,4-dioxane (5 mL). Amide coupling was performed on the crude product using EDC hydrochloride (752 mg, 3.92 mmol), *N*,*O*-dimethylhydroxyl-

amine hydrochloride (382 mg, 3.92 mmol) and NEt_3 (0.60 mL, 4.31 mmol) in DMF (25 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 7:3) to yield carboxamide **119** (392 mg, 0.960 mmol, 49%) as a white solid.

R_f 0.21 (PE : EtOAc, 7:3); **ν_{max}** (cm^{-1}) 1635 ((C=O)NH), 1450, 1107 (C-N), 1085; **mp** 113-115 ℃; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.62 (1H, dd, J = 1.3, 1.3 Hz, $H_{2'}$), 8.22 (1H, ddd, J = 7.8, 1.3, 1.3 Hz, $H_{6'}$), 8.15 (1H, d, J = 1.7 Hz, H_6), 7.88 (1H, 1H, ddd, J = 7.8, 1.3, 1.3 Hz, $H_{4'}$), 7.77 (1H, dd, J = 8.5, 1.7 Hz, H_2), 7.61 (1H, d, J = 8.5 Hz, H_3), 7.27 (1H, dd, J = 7.8, 7.8 Hz, $H_{5'}$), 3.57 (3H, s, H_{10}), 3.41 (3H, s, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 169.2 (C_8), 162.5 (C_7), 152.1 (C_4), 141.6 ($C_{4'}$), 140.8 (C_5), 136.6 ($C_{2'}$), 131.0 (C_1), 130.7 ($C_{5'}$), 128.8 ($C_{1'}$), 126.9 ($C_{6'}$), 126.6 (C_2), 120.8 (C_6), 110.5 (C_3), 94.5 ($C_{3'}$), 61.3 (C_{10}), 33.9 (C_9); **HRMS** (APCI$^+$) Calc. for $C_{16}H_{14}IN_2O_3$ $[M+H]^+$ 409.0043, found 409.0035.

2,2,2-Trifluoro-1-(2-(3-iodophenyl)benzo[*d*]oxazol-5-yl)ethan-1-one oxime 187

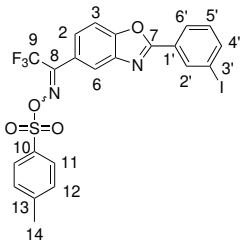


Oxime **187** was obtained according to General Procedure F, using carboxamide **119** (1.45 g, 3.55 mmol), trifluoromethyltrimethylsilane (1.05 mL, 7.10 mmol) and caesium fluoride (108 mg, 0.710 mmol) in toluene (40 mL). The reaction was stirred for three days, with further trifluoromethyltrimethylsilane (1.05 mL, 7.10 mmol) and caesium fluoride (108 mg, 0.710 mmol) added every 24 h. TBAF (4.26 mL, 1 M in THF) was used for desilylation of the intermediate formed. The crude trifluoromethylketone product was dissolved in pyridine (5 mL) and ethanol (25 mL) and refluxed with hydroxylammonium chloride (740 mg, 10.7 mmol). The crude product was purified by flash column chromatography (PE : EtOAc, 9:1) to yield oxime **187** (358 mg, 0.828 mmol, 23%) as a white solid.

R_f 0.28 (PE : EtOAc, 9:1); **ν_{max}** (cm^{-1}) 1547, 1471, 1250 (C-F), 1151 (C-F), 973 (N-O); **mp** 182-184 ℃; **^{1}HNMR** (500 MHz, CD_3OD) δ_H 8.60 (1H, dd, J = 1.6, 1.6 Hz, $H_{2'}$), 8.25 (1H, ddd, J = 7.8, 1.6, 1.0 Hz, $H_{6'}$), 7.97 (1H, ddd, J = 7.8, 1.6, 1.0 Hz, $H_{4'}$), 7.86 (1H, s, H_6), 7.80 (1H, d, J = 8.5 Hz, H_3), 7.53 (1H, dd, J = 8.5, 1.3 Hz, H_2), 7.36 (1H, dd, J = 7.8, 7.8 Hz, $H_{5'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.8 (C_7), 152.7 (C_4), 146.5 (q, $^2J_{C\text{-}F}$ = 32.6 Hz, C_8), 142.9 (C_5), 142.2 ($C_{4'}$), 137.5 ($C_{2'}$), 132.0 ($C_{5'}$), 129.7 ($C_{1'}$), 127.9 ($C_{6'}$), 127.8 (C_2), 125.5 (C_1), 122.5 (q, $^1J_{C\text{-}F}$ = 272.5 Hz, C_9), 121.8 (C_6), 112.1 (C_3), 95.1 ($C_{3'}$); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −66.5 (s); **HRMS** (ES$^+$) Calc. for $C_{15}H_7F_3IN_2O_2$ $[M-H]^-$ 430.9510,

found 430.9513.

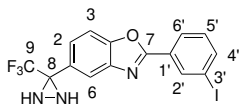
2,2,2-Trifluoro-1-(2-(3-iodophenyl)benzo[*d*]oxazol-5-yl)ethan-1-one *O*-tosyl oxime 123



Tosyl oxime **123** was obtained according to General Procedure G, using oxime **187** (358 mg, 0.828 mmol), *p*-toluenesulfonyl chloride (316 mg, 1.66 mmol), 4-dimethylaminopyridine (10.1 mg, 82.8 µmol) and NEt_3 (0.35 mL, 2.48 mmol) in CH_2Cl_2 (20 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 37:3) to yield tosyl oxime **123** (402 mg, 0.685 mmol, 83%) as a pale yellow solid in a mixture of isomers (7:1).

R_f 0.24 (PE : EtOAc, 9:1); **ν_{max}** (cm^{-1}) 2932 (C-H), 1622 (C=N-O), 1391 (S=O), 1250 (C-F), 1181 (S=O), 1149; **mp** 229-231 ℃; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.64 (1H, dd, J = 1.6, 1.6 Hz, $H_{2'}$), 8.25 (1H, ddd, J = 7.9, 1.6, 1.0 Hz, $H_{6'}$), 7.94 (1H, ddd, J = 7.9, 1.6, 1.0 Hz, $H_{4'}$), 7.93 (2H, d, J = 8.4 Hz, H_{11}), 7.82 (1H, d, J = 1.7 Hz, H_6), 7.72 (1H, d, J = 8.5 Hz, H_3), 7.46 (1H, dd, J = 8.5, 1.7 Hz, H_2), 7.43 (2H, d, J = 8.4 Hz, H_{12}), 7.32 (1H, dd, J = 7.9, 7.9 Hz, $H_{5'}$), 2.52 (3H, s, H_{14}); **^{13}CNMR** (125 MHz, $CDCl_3$)δ_C 162.9 (C_7), 153.5 (q, $^2J_{C\text{-}F}$ = 33.7 Hz, C_8), 152.3 (C_4), 146.3 (C_{13}), 142.2 (C_5), 141.0 ($C_{4'}$), 136.6 ($C_{2'}$), 131.1 (C_{10}), 130.7 ($C_{5'}$), 129.9 (C_{12}), 129.3 (C_{11}), 128.2 ($C_{1'}$), 127.0 ($C_{6'}$), 126.0 (C_2), 121.2 (C_1), 121.0 (C_6), 119.6 (q, $^1J_{C\text{-}F}$ = 277.6 Hz, C_9), 111.5 (C_3), 94.4 ($C_{3'}$), 21.8 (C_{14}); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −66.7 (s); **HRMS** (CI^+) Calc. for $C_{22}H_{15}F_3IN_2O_4$ $[M+H]^+$ 586.9744, found 586.9759.

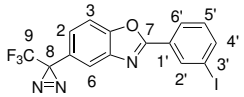
2-(3-Iodophenyl)-5-(3-(trifluoromethyl)diaziridin-3-yl)benzo[*d*]oxazole 188



Diaziridine **188** was obtained according to General Procedure H, using tosyl oxime **123** (375 mg, 0.640 mmol) and NH_3 (5 mL) in CH_2Cl_2 (5 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 4:1) to yield diaziridine **188** (270 mg, 0.626 mmol, 98%) as a white solid.

R_f 0.16 (PE : EtOAc, 9:1); **ν_{max}** (cm^{-1}) 3243 (N-H), 1570 (N-H), 1216 (C-F), 1150; **mp** 128-130 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.62 (1H, dd, J = 1.7, 1.7 Hz, $H_{2'}$), 8.22 (1H, ddd, J = 7.9, 1.7, 1.1 Hz, $H_{6'}$), 8.06 (1H, d, J = 1.6 Hz, H_6), 7.90 (1H, ddd, J = 7.9, 1.7, 1.1 Hz, $H_{4'}$), 7.66 (1H, dd, J = 8.5, 1.6 Hz, H_2), 7.63 (1H, d, J = 8.5 Hz, H_3), 7.28 (1H, dd, J = 7.9, 7.9 Hz, $H_{5'}$), 2.89 (1H, d, J = 8.7 Hz, -NH), 2.32 (1H, d, J = 8.7 Hz, -NH); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 162.6 (C_7), 151.6 (C_4), 142.2 (C_5), 140.8 ($C_{4'}$), 136.5 ($C_{2'}$), 130.6 ($C_{5'}$), 128.7 (C_1), 128.5 ($C_{1'}$), 126.9 ($C_{6'}$), 125.5 (C_2), 123.5 (q, $^1J_{C\text{-}F}$ = 278.2 Hz, C_9), 120.6 (C_6), 111.1 (C_3), 94.4 ($C_{3'}$), 58.2 (q, $^2J_{C\text{-}F}$ = 36.2 Hz, C_8); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −75.5 (s); **HRMS** (ES^+) Calc. for $C_{15}H_{10}F_3IN_3O$ $[M+H]^+$ 431.9815, found 431.9815.

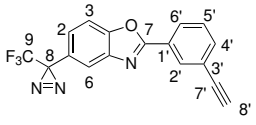
2-(3-Iodophenyl)-5-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]oxazole 125



Diazirine **125** was obtained according to General Procedure I, using diaziridine **188** (250 mg, 0.580 mmol), iodine (176 mg, 0.696 mmol) and NEt_3 (0.24 mL, 1.74 mmol) in CH_2Cl_2 (15 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 19:1 → 9:1) to yield diazirine **125** (242 mg, 0.564 mmol, 97%) as a white solid.

R_f 0.81 (PE : EtOAc, 9:1); **ν_{max}** (cm^{-1}) 1547, 1254 (C-F), 1194, 1154 (C-F), 723; **mp** 101-103 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.60 (1H, t, J = 1.7 Hz, $H_{2'}$), 8.21 (1H, ddd, J = 7.9, 1.7, 1.0 Hz, $H_{6'}$), 7.90 (1H, ddd, J = 7.9, 1.7, 1.0 Hz, $H_{4'}$), 7.72 (1H, d, J = 1.8 Hz, H_6), 7.61 (1H, d, J = 8.6 Hz, H_3), 7.28 (1H, t, J = 7.9 Hz, $H_{5'}$), 7.20 (1H, dd, J = 8.6, 1.8 Hz, H_2); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 162.9 (C_7), 151.5 (C_4), 142.6 (C_5), 141.0 ($C_{4'}$), 136.7 ($C_{2'}$), 130.8 ($C_{5'}$), 128.5 ($C_{1'}$), 127.0 ($C_{6'}$), 126.0 (C_1), 124.1 (C_2), 122.3 (q, $^1J_{C\text{-}F}$ = 274.7 Hz, C_9), 119.5 (C_6), 111.5 (C_3), 94.6 ($C_{3'}$), 28.7 (q, $^2J_{C\text{-}F}$ = 40.6 Hz, C_8); **^{19}FNMR** (376 MHz, CD_3OD) δ_F −67.4 (s); **HRMS** (ES^+) Calc. for $C_{15}H_8F_3IN_3O$ $[M+H]^+$ 429.9659, found 429.9655.

2-(3-Ethynylphenyl)-5-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]oxazole 107



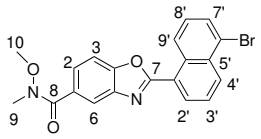
Diazirine **107** was obtained according to General Procedure B, using diazirine **125** (100 mg, 0.233 mmol), $Pd(PPh_3)_2Cl_2$ (24.5 mg, 35.0 µmol), copper(I) iodide (4.44 mg, 23.3 µmol),

NEt_3 (0.70 mL) and trimethylsilylacetylene (40 µL, 0.280 mmol) in THF (10 mL). The reaction was complete after 1 h. The TMS-protected alkyne intermediate formed was purified by flash column chromatography (PE : EtOAc, 1:0 → 99:1), then deprotected with TBAF (0.70 mL, 1 M in THF) in THF (5 mL). The reaction was complete after 30 min. The crude product was purified by flash column chromatography (PE : EtOAc, 1:0 → 49:1) to yield alkyne **107** (53.2 mg, 0.163 mmol, 70%) as a white solid.

R_f 0.62 (PE : EtOAc, 49:1); **ν_{max}** (cm^{-1}) 3310 (sharp, alkyne C-H), 1556, 1259 (C-F), 1197, 1154, 803; **mp** 105-107 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.40 (1H, dd, J = 1.7, 1.7 Hz, $H_{2'}$), 8.25 (1H, ddd, J = 7.9, 1.4, 1.4 Hz, $H_{6'}$), 7.75 (1H, d, J = 1.3 Hz, H_6), 7.70 (1H, ddd, J = 7.9, 1.4, 1.4 Hz, $H_{4'}$), 7.64 (1H, d, J = 8.6 Hz, H_3), 7.54 (1H, dd, J = 7.9, 7.9 Hz, $H_{5'}$), 7.24 (1H, dd, J = 8.6, 1.3 Hz, H_2), 3.20 (1H, s, $H_{8'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.6 (C_7), 151.4 (C_4), 142.6 (C_5), 135.4 ($C_{4'}$), 131.4 ($C_{2'}$), 129.1 ($C_{5'}$), 127.9 ($C_{6'}$), 126.9 ($C_{1'}$), 125.8 (C_1), 123.9 (C_2), 123.3 ($C_{3'}$), 122.2 (q, $^1J_{C\text{-}F}$ = 274.7 Hz, C_9), 119.3 (C_6), 111.3 (C_3), 82.4 ($C_{7'}$), 78.6 ($C_{8'}$), 28.7 (q, $^2J_{C\text{-}F}$ = 40.7 Hz, C_8); **^{19}FNMR** (376 MHz, CD_3OD) δ_F −67.4 (s); **HRMS** (ES^+) Calc. for $C_{17}H_9F_3N_3O$ $[M+H]^+$ 328.0692, found 328.0691.

2-(5-Bromonaphthalen-1-yl)-*N*-methoxy-*N*-methylbenzo[*d*]oxazole-5-carboxamide

189

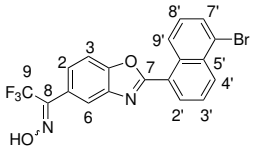


Carboxamide **189** was obtained by a procedure based on General Procedure D except with the addition of 1-hydroxybenzotriazole in the amide coupling. 5-Bromonaphthoic acid (502 mg, 2.00 mmol), thionyl chloride (0.44 mL, 6.00 mmol) and DMF (0.10 mL) were used in CH_2Cl_2 (20 mL). The acid chloride intermediate formed was reacted with 3-amino-4-hydroxybenzoic acid (277 mg, 1.80 mmol, 0.9 eq.) in 1,4-dioxane (5 mL). Amide coupling was performed on the crude product using EDC hydrochloride (575 mg, 3.00 mmol, 1.5 eq.), *N,O*-dimethylhydroxylamine hydrochloride (293 mg, 3.00 mmol, 1.5 eq.), 1-hydroxybenzotriazole (135 mg, 1.00 mmol, 0.5 eq.) and NEt_3 (0.42 mL, 3.00 mmol, 1.5 eq.) in DMF (25 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 1:1) to yield carboxamide **190** (459 mg, 1.12 mmol, 62%) as a white solid.

R_f 0.54 (PE : EtOAc, 1:1); **ν_{max}** (cm^{-1}) 2930 (C-H), 1679 ((C=O)NH), 1460, 1251; **mp**

137-138 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 9.48 (1H, ddd, J = 8.7, 1.0, 1.0 Hz, $H_{9'}$), 8.48 (1H, ddd, J = 8.6, 1.1, 1.1 Hz, $H_{4'}$), 8.43 (1H, dd, J = 7.3, 1.1 Hz, $H_{2'}$), 8.24 (1H, d, J = 1.6 Hz, H_6), 7.87 (1H, dd, J = 7.4, 1.1 Hz, $H_{7'}$), 7.80 (1H, dd, J = 8.4, 1.6 Hz, H_2), 7.67 (1H, dd, J = 8.5, 7.3 Hz, $H_{3'}$), 7.64 (1H, d, J = 8.4 Hz, H_3), 7.50 (1H, dd, J = 8.7, 7.4 Hz, $H_{8'}$), 3.59 (3H, s, H_{10}), 3.42 (3H, s, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 169.2 (C_8), 163.2 (C_7), 151.4 (C_4), 141.8 (C_5), 132.4 ($C_{5'}$), 132.0 ($C_{10'}$), 131.6 ($C_{4'}$), 130.9 ($C_{7'}$), 130.7 (C_1), 130.2 ($C_{2'}$), 128.2 ($C_{8'}$), 126.5 (C_2), 126.3 ($C_{3'}$), 126.2 ($C_{9'}$), 123.6 ($C_{1'}$ or $C_{6'}$), 123.5 ($C_{1'}$ or $C_{6'}$), 120.8 (C_6), 110.2 (C_3), 61.1 (C_{10}), 33.9 (C_9); **HRMS** (ES^+) Calc. for $C_{20}H_{16}{}^{79}BrN_2O_3$ $[M+H]^+$ 411.0339, found 411.0339.

1-(2-(5-Bromonaphthalen-1-yl)benzo[*d*]oxazol-5-yl)-2,2,2-trifluoroethan-1-one oxime 191

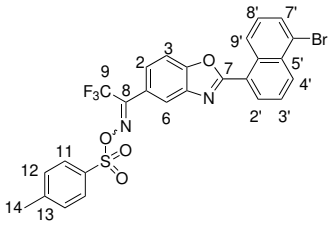


Oxime **191** was obtained according to General Procedure F, using carboxamide **192** (1.18 g, 2.87 mmol), trifluoromethyltrimethylsilane (0.85 mL, 5.74 mmol) and caesium fluoride (87.2 mg, 0.574 mmol) in toluene (40 mL). The reaction was stirred for three days, with further trifluoromethyltrimethylsilane (0.85 mL, 5.74 mmol) and caesium fluoride (87.2 mg, 0.574 mmol) added every 24 h. TBAF (3.44 mL, 1 M in THF) was used for desilylation of the intermediate formed. The crude trifluoromethylketone product was dissolved in pyridine (5 mL) and ethanol (25 mL) and refluxed with hydroxylammonium chloride (598 mg, 8.61 mmol). The crude product was purified by flash column chromatography (PE : EtOAc, 9:1) to yield oxime **191** (699 mg, 1.61 mmol, 56%) as a yellow solid.

R_f 0.31 (PE : EtOAc, 9:1); **ν_{max}** (cm^{-1}) 3223 (-OH), 2981 (C-H), 1696 (C=N-OH), 1253 (C-F), 1154; **mp** 182-183 °C; **^{1}HNMR** (500 MHz, $(CD_3)_2CO$) δ_H 11.96 (1H, *br* s, -OH), 9.63 (1H, ddd, J = 8.7, 1.1, 1.1 Hz, $H_{9'}$), 8.61 (1H, dd, J = 7.4, 1.1 Hz, $H_{2'}$), 8.57 (1H, ddd, J = 8.6, 1.1, 1.1 Hz, $H_{4'}$), 8.06 (1H, s, H_6), 8.05 (1H, *obs* m, $H_{7'}$), 7.81 (1H, d, J = 8.4 Hz, H_3), 7.76 (1H, dd, J = 8.6, 7.4 Hz, $H_{3'}$), 7.67 (1H, dd, J = 8.7, 7.6 Hz, $H_{8'}$), 7.65 (1H, dd, J = 8.4, 1.6 Hz, H_2); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.8 (C_7), 151.9 (C_8), 151.7 (C_4), 143.1 (C_5), 133.1 (C_1), 132.8 ($C_{5'}$), 132.2 ($C_{10'}$), 132.0 ($C_{4'}$), 131.5 ($C_{7'}$), 131.4 ($C_{2'}$), 129.3 ($C_{8'}$), 127.8 ($C_{3'}$), 127.5 (C_2), 127.4 ($C_{9'}$), 124.6 ($C_{1'}$), 123.7 ($C_{6'}$), 122.2 (q, $^1J_{C\text{-}F}$ = 273.1 Hz, C_9), 121.9 (C_6), 111.8 (C_3); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −67.6 (s); **HRMS** (ES^+) Calc. for

$C_{19}H_{11}F_3{}^{79}BrN_2O_2$ $[M+H]^+$ 434.9951, found 434.9951.

1-(2-(5-Bromonaphthalen-1-yl)benzo[*d*]oxazol-5-yl)-2,2,2-trifluoroethan-1-one *O*-tosyl oxime 193



Tosyl oxime **193** was obtained according to General Procedure G, using oxime **194** (818 mg, 1.88 mmol), *p*-toluenesulfonyl chloride (717 mg, 3.76 mmol), 4-dimethylaminopyridine (23.0 mg, 0.188 mmol) and NEt_3 (0.78 mL, 5.64 mmol) in CH_2Cl_2 (30 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 19:1) to yield tosyl oxime **193** (820 mg, 1.39 mmol, 74%) as a pale yellow solid in a mixture of isomers (50:50).

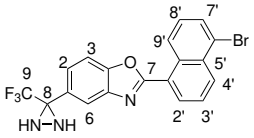
$\mathbf{\nu_{max}}$ (cm^{-1}) 2980 (C-H), 1334 (S=O), 1255 (C-F), 1196 (S=O), 1154; **mp** 148-150 ℃;

Minor isomer: $\mathbf{R_f}$ (0.35, PE : EtOAc, 19:1); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 9.48 (1H, ddd, J = 8.8, 1.0, 1.0 Hz, $H_{9'}$), 8.54 (1H, ddd, J = 8.6, 1.1, 1.1 Hz, $H_{4'}$), 8.46 (1H, dd, J = 7.3, 1.2 Hz, $H_{2'}$), 7.94-7.89 (4H, m, H_6, $H_{7'}$ and H_{11}), 7.73 (1H, d, J = 8.4 Hz, H_3), 7.71 (1H, dd, J = 8.6, 7.3 Hz, $H_{3'}$), 7.54 (1H, dd, J = 8.8, 7.4 Hz, $H_{8'}$), 7.47 (1H, dd, J = 8.5, 1.7 Hz, H_2), 7.41 (2H, m, H_{12}), 2.49 (3H, s, H_{14}); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.8 (C_7), 153.7 (q, $^2J_{C\text{-}F}$ = 34.1 Hz, C_8), 151.8 (C_4), 146.2 (C_{14}), 142.6 (C_5), 132.6 ($C_{5'}$), 132.2 ($C_{10'}$), 132.1 ($C_{4'}$), 131.3 ($C_{1'}$), 131.1 ($C_{7'}$), 130.6 ($C_{2'}$), 130.1 (C_{12}), 129.5 (C_{10}), 129.3 (C_{11}), 128.5 ($C_{8'}$), 126.4 ($C_{3'}$), 126.2 (C_2), 126.2 ($C_{9'}$), 123.7 ($C_{6'}$), 123.4 (C_1), 121.3 (C_6), 119.8 (q, $^1J_{C\text{-}F}$ = 277.5 Hz, C_9), 111.2 (C_3), 22.0 (C_{14}); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −66.6 (s);

Major isomer: $\mathbf{R_f}$ (0.28, PE : EtOAc, 19:1); **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 9.49 (1H, ddd, J = 8.8, 1.0, 1.0 Hz, $H_{9'}$), 8.54 (1H, ddd, J = 8.6, 1.1, 1.1 Hz, $H_{4'}$), 8.46 (1H, dd, J = 7.3, 1.2 Hz, $H_{2'}$), 7.96-7.91 (4H, m, H_6, $H_{7'}$ and H_{11}), 7.71 (1H, dd, J = 8.6, 7.3 Hz, $H_{3'}$), 7.68 (1H, d, J = 8.4 Hz, H_3), 7.55 (1H, dd, J = 8.8, 7.4 Hz, $H_{8'}$), 7.52 (1H, dd, J = 8.5, 1.7 Hz, H_2), 7.40 (2H, m, H_{12}), 2.47 (3H, s, H_{14}); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.9 (C_7), 153.9 (q, $^2J_{C\text{-}F}$ = 32.9 Hz, C_8), 152.2 (C_4), 146.4 (C_{13}), 142.7 (C_5), 132.6 ($C_{5'}$), 132.2 ($C_{10'}$), 132.1 ($C_{4'}$), 131.3 ($C_{1'}$), 131.1 ($C_{7'}$), 130.6 ($C_{2'}$), 130.1 (C_{12}), 129.5 (C_{10}), 129.3 (C_{11}), 128.5 ($C_{8'}$), 126.4 ($C_{3'}$), 126.3 (C_2), 126.2 ($C_{9'}$), 123.7 ($C_{6'}$), 123.6 (C_1), 121.6 (C_6), 119.8 (q, $^1J_{C\text{-}F}$ = 277.5 Hz, C_9), 111.5 (C_3), 22.0 (C_{14}); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −66.6 (s);

HRMS (ES^+) Calc. for $C_{26}H_{17}F_3{}^{79}BrN_2O_4S$ $[M+H]^+$ 589.0039, found 589.0035.

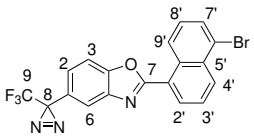
2-(5-Bromonaphthalen-1-yl)-5-(3-(trifluoromethyl)diaziridin-3-yl)benzo[*d*]oxazole 195



Diaziridine **195** was obtained according to General Procedure H, using tosyl oxime **193** (630 mg, 1.07 mmol) and NH_3 (5 mL) in CH_2Cl_2 (10 mL). The crude product was used without further purification, yielding diaziridine **195** (469 mg, 1.07 mmol, quant.) as a white solid.

R_f (0.67, PE : EtOAc, 4:1); **ν_{max}** (cm^{-1}) 2981 (C-H), 1457, 1157, 1036; **mp** 149-150 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 9.51 (1H, ddd, J = 8.5, 1.1, 1.1 Hz, $H_{9'}$), 8.55 (1H, ddd, J = 8.6, 1.1, 1.1 Hz, $H_{4'}$), 8.49 (1H, dd, J = 7.4, 1.1 Hz, $H_{2'}$), 8.17 (1H, s, H_6), 7.92 (1H, dd, J = 7.4, 1.1 Hz, $H_{7'}$), 7.73 (1H, dd, J = 8.6, 7.4 Hz, $H_{3'}$), 7.70 (2H, s, H_2 and H_3), 7.55 (1H, dd, J = 8.5, 7.4 Hz, $H_{8'}$), 2.91 (1H, d, J = 8.9 Hz, -NH), 2.34 (1H, d, J = 8.9 Hz, -NH); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.5 (C_7), 151.0 (C_4), 142.5 (C_5), 132.5 ($C_{5'}$), 132.0 ($C_{10'}$), 131.9 ($C_{4'}$), 130.9 ($C_{7'}$), 130.3 ($C_{2'}$), 129.7 (C_1), 128.3 ($C_{8'}$), 127.2 ($C_{1'}$), 126.3 ($C_{3'}$ or $C_{9'}$), 126.2 ($C_{3'}$ or $C_{9'}$), 125.6 (C_2), 124.5 (q, $^1J_{C\text{-}F}$ = 276.6 Hz, C_9), 123.6 ($C_{6'}$), 120.8 (C_6), 110.9 (C_3), 58.3 (q, $^2J_{C\text{-}F}$ = 36.1 Hz, C_8);**^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −75.5 (s); **HRMS** (ES^+) Calc. for $C_{19}H_{12}{}^{79}BrF_3N_3O$ $[M+H]^+$ 434.0105, found 434.0110.

2-(5-Bromonaphthalen-1-yl)-5-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]-oxazole 196

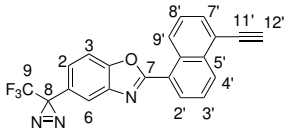


Diazirine **196** was obtained according to General Procedure I, using diaziridine **195** (439 mg, 1.01 mmol), iodine (308 mg, 1.21 mmol) and NEt_3 (0.42 mL, 3.03 mmol) in CH_2Cl_2 (20 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 49:1 → 19:1) to yield diazirine **196** (407 mg, 0.945 mmol, 94%) as a white solid.

R_f (0.53, PE : EtOAc, 49:1); **ν_{max}** (cm^{-1}) 2981 (C-H), 1258 (C-F), 1154, 773; **mp** 104-106 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 9.50 (1H, ddd, J = 8.7, 1.1, 1.1 Hz, $H_{9'}$), 8.54 (1H, ddd,

J = 8.6, 1.1, 1.1 Hz, $H_{4'}$), 8.47 (1H, dd, J = 7.3, 1.1 Hz, $H_{2'}$), 7.92 (1H, dd, J = 7.4, 1.0 Hz, $H_{7'}$), 7.83 (1H, dd, J = 1.8, 0.6 Hz, H_6), 7.72 (1H, dd, J = 8.6, 7.3 Hz, $H_{3'}$), 7.66 (1H, d, J = 8.6, 0.6 Hz, H_3), 7.54 (1H, dd, J = 8.7, 7.4 Hz, $H_{8'}$), 7.24 (1H, dd, J = 8.6, 1.8 Hz, H_2); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.6 (C_7), 150.7 (C_4), 142.8 (C_5), 132.4 ($C_{5'}$), 132.1 ($C_{10'}$), 132.0 ($C_{4'}$), 131.0 ($C_{7'}$), 130.4 ($C_{2'}$), 128.3 ($C_{8'}$), 126.3 ($C_{3'}$), 126.1 ($C_{9'}$), 125.6 (C_1), 124.0 (C_2), 123.6 ($C_{1'}$ or $C_{6'}$), 123.4 ($C_{1'}$ or $C_{6'}$), 122.3 (q, $^1J_{C\text{-}F}$ = 274.7 Hz, C_9), 119.5 (C_6), 111.2 (C_3), 28.8 (q, $^2J_{C\text{-}F}$ = 40.6 Hz, C_8); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −65.4 (s); **HRMS** (ES^+) Calc. for $C_{19}H_{10}{}^{79}BrF_3N_3O$ $[M+H]^+$ 431.9954, found 431.9950.

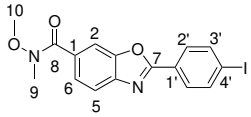
2-(5-Ethynylnaphthalen-1-yl)-5-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]-oxazole 129



Diazirine **129** was obtained according to General Procedure B, using diazirine **196** (200 mg, 0.465 mmol), $Pd(PPh_3)_2Cl_2$ (48.9 mg, 69.7 µmol), copper(I) iodide (8.86 mg, 46.5 µmol), NEt_3 (1.50 mL) and trimethylsilylacetylene (78.2 µL, 0.558 mmol) in THF (10 mL). The reaction was worked up after 6 h, although starting material was still present. The TMS-protected alkyne intermediate formed was purified by flash column chromatography (PE : EtOAc, 1:0 → 99:1), then deprotected with TBAF (0.93 mL, 1 M in THF) in THF (5 mL). The reaction was complete after 30 min. The crude product was purified by flash column chromatography (pentane : toluene, 19:1) to yield alkyne **129** (58.9 mg, 0.156 mmol, 34%) as an off-white solid.

R_f (0.16, pentane : toluene, 19:1); **ν_{max}** (cm^{-1}) 2981 (C-H), 1256 (C-F), 1154, 773; **mp** 87-89 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 9.52 (1H, ddd, J = 8.6, 1.0, 1.0 Hz, $H_{9'}$), 8.64 (1H, ddd, J = 8.4, 1.1, 1.1 Hz, $H_{4'}$), 8.46 (1H, dd, J = 7.3, 1.3 Hz, $H_{2'}$), 7.85 (1H, dd, J = 7.1, 1.2 Hz, $H_{7'}$), 7.82 (1H, d, J = 1.7 Hz, H_6), 7.70 (1H, dd, J = 8.4, 7.3 Hz, $H_{3'}$), 7.65 (1H, dd, J = 8.7, 7.1 Hz, $H_{8'}$), 7.64 (1H, d, J = 8.7 Hz, H_3), 7.24 (1H, dd, J = 8.4, 1.9 Hz, H_2), 3.53 (1H, s, $H_{12'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.9 (C_7), 150.7 (C_4), 142.8 (C_5), 134.0 ($C_{5'}$), 131.9 ($C_{7'}$), 130.9 ($C_{4'}$), 130.5 ($C_{2'}$), 130.2 ($C_{10'}$), 127.4 ($C_{9'}$), 127.3 ($C_{8'}$), 125.9 ($C_{3'}$), 125.6 (C_1), 123.9 (C_2), 123.2 ($C_{1'}$), 122.3 (q, $^1J_{C\text{-}F}$ = 274.7 Hz, C_9), 120.6 ($C_{6'}$), 119.4 (C_6), 111.1 (C_3), 82.7 ($C_{12'}$), 81.5 ($C_{11'}$), 28.8 (q, $^2J_{C\text{-}F}$ = 40.6 Hz, C_8); **^{19}FNMR** (376 MHz, CD_3OD) δ_F −65.4 (s); **HRMS** (CI^+) Calc. for $C_{21}H_{10}F_3N_3O$ $[M^+]$ 377.0770 found 377.0778.

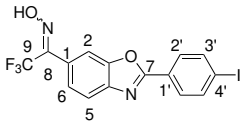
2-(4-Iodophenyl)-*N*-methoxy-*N*-methylbenzo[*d*]oxazole-6-carboxamide 190



Carboxamide **190** was obtained according to General Procedure E, using 4-iodobenzoic acid (500 mg, 2.01 mmol), thionyl chloride (0.44 mL, 6.03 mmol) and DMF (0.10 mL) in CH_2Cl_2 (20 mL). The acid chloride intermediate formed was reacted with 4-amino-3-hydroxybenzoic acid (277 mg, 1.81 mmol, 0.9 eq.) in 1,4-dioxane (5 mL). Amide coupling was performed on the crude product using EDC hydrochloride (769 mg, 4.02 mmol), *N*,*O*-dimethylhydroxylamine hydrochloride (392 mg, 4.02 mmol) and NEt_3 (0.56 mL, 4.01 mmol) in DMF (25 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 4:1 → 3:1) to yield carboxamide **190** (260 mg, 0.638 mmol, 58%) as a white solid.

R_f 0.30 (PE : EtOAc, 4:1); **ν_{max}** (cm^{-1}) 2889 (C-H), 1636 ((C=O)NH), 1380, 1219, 772; **mp** 168-169 ℃; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 7.98 (2H, m, $H_{2'}$), 7.97 (1H, s, H_2), 7.90 (2H, m, $H_{3'}$), 7.77 (1H, d, J = 8.3 Hz, H_5), 7.76 (1H, dd, J = 8.3, 1.3 Hz, H_6), 3.57 (3H, s, H_{10}), 3.41 (3H, s, H_9); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 169.1 (C_8), 164.2 (C_7), 150.2 (C_3), 144.0 (C_4), 138.4 ($C_{3'}$), 131.2 (C_1), 129.3 ($C_{2'}$), 126.4 ($C_{1'}$), 125.7 (C_6), 119.6 (C_5), 111.4 (C_2), 99.2 ($C_{4'}$), 61.3 (C_{10}), 33.9 (C_9); **HRMS** (ES^+) Calc. for $C_{16}H_{14}IN_2O_3$ $[M+H]^+$ 409.0044, found 409.0043.

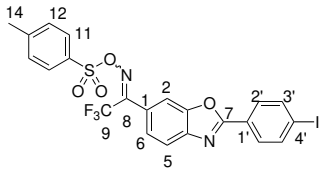
2,2,2-Trifluoro-1-(2-(4-iodophenyl)benzo[*d*]oxazol-6-yl)ethan-1-one oxime 197



Oxime **197** was obtained according to General Procedure F, using carboxamide **198** (487 mg, 1.19 mmol), trifluoromethyltrimethylsilane (0.35 mL, 2.39 mmol) and caesium fluoride (36.3 mg, 0.239 mmol) in toluene (25 mL). The reaction was stirred for two days, with further trifluoromethyltrimethylsilane (0.35 mL, 2.39 mmol) and caesium fluoride (36.3 mg, 0.239 mmol) added every 24 h. TBAF (1.43 mL, 1 M in THF) was used for desilylation of the intermediate formed. The crude trifluoromethylketone product was dissolved in pyridine (3 mL) and ethanol (20 mL) and refluxed with hydroxylammonium chloride (248 mg, 3.57 mmol). The crude product was purified by flash column chromatography (PE : EtOAc, 9:1) to yield oxime **199** (149 mg, 0.345 mmol, 29%) as a yellow solid.

R_f 0.26 (PE : EtOAc, 9:1); **ν_{max}** (cm^{-1}) 2981 (C-H), 1610 (C=N-OH), 1479, 1245 (C-F), 1148; **mp** 185-186 ℃; **^{1}HNMR** (500 MHz, CD_3OD) δ_H 7.97 (2H, m, $H_{2'}$), 7.96 (2H, m, $H_{3'}$), 7.80 (1H, s, H_2), 7.76 (d, J = 8.4 Hz, H_5), 7.53 (1H, d, J = 8.4 Hz, H_6); **^{13}CNMR** (125 MHz, CD_3OD)δ_C 165.3 (C_7), 151.7 (C_3), 146.3 (q, $^2J_{C\text{-}F}$ = 29.5 Hz, C_8), 144.1 (C_4), 139.7 ($C_{3'}$), 130.2 ($C_{2'}$), 130.0 (C_1), 127.2 ($C_{1'}$), 126.7 (C_6), 120.6 (C_5), 120.2 (q, $^1J_{C\text{-}F}$ = 282.1 Hz, C_9), 112.1 (C_2), 100.1 ($C_{4'}$); **^{19}FNMR** (376 MHz, CD_3OD) δ_F −63.7 (s); **HRMS** (ES^+) Calc. for $C_{15}H_9F_3IN_2O_2$ $[M+H]^+$ 432.9655, found 432.9655.

2,2,2-Trifluoro-1-(2-(4-iodophenyl)benzo[*d*]oxazol-6-yl)ethan-1-one *O*-tosyl oxime 200



Tosyl oxime **200** was obtained according to General Procedure G, using oxime **197** (482 mg, 1.12 mmol), *p*-toluenesulfonyl chloride (425 mg, 2.23 mmol), 4-dimethylaminopyridine (13.6 mg, 0.112 mmol) and NEt_3 (0.47 mL, 3.35 mmol) in CH_2Cl_2 (35 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 19:1) to yield tosyl oxime **200** (505 mg, 0.862 mmol, 77%) as a yellow solid in a mixture of isomers (5:6).

ν_{max} (cm^{-1}) 2980 (C-H), 1610 (C=N-O), 1478, 1393 (S=O), 1245 (C-F), 1195 (S=O), 1149; **mp** 162-164 ℃;

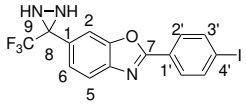
Minor isomer: **R_f** 0.15 (PE : EtOAc, 19:1); **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 7.93 (2H, m, $H_{2'}$), 7.90 (2H, m, H_{11}), 7.86 (2H, m, $H_{3'}$), 7.80 (1H, dd, J = 8.3, 0.6 Hz, H_5), 7.68 (1H, d, J = 1.6 Hz, H_2), 7.39 (3H, m, H_6 and H_{12}), 2.46 (3H, s, H_{14}); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.9 (C_7), 153.4 (q, $^2J_{C\text{-}F}$ = 33.8 Hz, C_8), 150.5 (C_3), 146.5 (C_{13}), 145.1 (C_4), 138.6 ($C_{3'}$), 131.5 (C_{10}), 130.1 ($C_{1'}$), 129.5 ($C_{2'}$), 129.4 (C_{12}), 129.3 (C_{11}), 127.4 (q, $^3J_{C\text{-}F}$ = 1.56 Hz, C_1), 125.9 (C_6), 120.7 (C_5), 119.7 (q, $^1J_{C\text{-}F}$ = 277.9 Hz, C_9), 111.8 (C_2), 99.8 ($C_{4'}$), 22.0 (C_{14}); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −61.2 (s);;

Major isomer: **R_f** 0.10 (PE : EtOAc, 19:1); **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 7.95 (2H, m, $H_{2'}$), 7.92 (2H, m, H_{11}), 7.89 (2H, m, $H_{3'}$), 7.77 (1H, dd, J = 8.4, 0.6 Hz, H_5), 7.70 (dd, J = 2.1, 0.6 Hz, H_2), 7.45 (dd, J = 8.4, 2.1 Hz, H_6), 7.38 (2H, m, H_{12}), 2.47 (3H, s, H_{14}); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.8 (C_7), 153.6 (q, $^2J_{C\text{-}F}$ = 32.4 Hz, C_8), 150.4 (C_3), 146.3 (C_{13}), 144.8 (C_4), 138.6 ($C_{3'}$), 131.2 (C_{10}), 130.1 ($C_{1'}$), 129.5 ($C_{2'}$), 129.4 (C_{12}), 129.3

(C_{11}), 127.4 (q, $^{3}J_{C\text{-}F}$ = 1.56 Hz, C_1), 126.0 (C_6), 120.5 (C_5), 119.7 (q, $^{1}J_{C\text{-}F}$ = 277.9 Hz, C_9), 111.8 (C_2), 99.8 ($C_{4'}$), 22.0 (C_{14}); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −61.2 (s);

HRMS (ES^+) Calc. for $C_{22}H_{15}F_3IN_2O_4S$ $[M+H]^+$ 586.9744, found 586.9739.

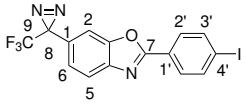
2-(4-Iodophenyl)-6-(3-(trifluoromethyl)diaziridin-3-yl)benzo[*d*]oxazole 201



Diaziridine **201** was obtained according to General Procedure H, using tosyl oxime **200** (480 mg, 0.819 mmol) and NH_3 (5 mL) in CH_2Cl_2 (10 mL). The crude product was used without further purification, yielding diaziridine **201** (353 mg, 0.819 mmol, quant.) as a white solid.

R_f (0.67, PE : EtOAc, 4:1); **ν_{max}** (cm^{-1}) 2952 (C-H), 1478, 1152, 1035, 816; **mp** 163-164 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.00 (2H, m, $H_{2'}$), 7.93 (2H, m, $H_{3'}$), 7.91 (1H, d, J = 1.2 Hz, H_2), 7.83 (1H, d, J = 8.4 Hz, H_5), 7.67 (1H, dd, J = 8.4, 1.2 Hz, H_6), 2.92 (1H, d, J = 9.2 Hz, -NH), 2.34 (1H, d, J = 9.2 Hz, -NH); **^{13}CNMR** (125 MHz, $CDCl_3$)δ_C 164.1 (C_7), 150.6 (C_3), 143.7 (C_4), 138.5 ($C_{3'}$), 129.6 ($C_{2'}$), 127.4 (q, $^{3}J_{C\text{-}F}$ = 1.79 Hz, C_1), 126.2 ($C_{1'}$), 125.0 (C_6), 123.6 (q, $^{1}J_{C\text{-}F}$ = 278.2 Hz, C_9), 120.5 (C_5), 111.2 (C_2), 99.3 ($C_{4'}$), 58.3 (q, $^{2}J_{C\text{-}F}$ = 36.2 Hz, C_8); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −75.3 (s); **HRMS** (ES^+) Calc. for $C_{15}H_{10}F_3IN_3O$ $[M+H]^+$ 431.9815, found 431.9820.

2-(4-Iodophenyl)-6-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]oxazole 202

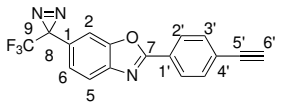


Diazirine **202** was obtained according to General Procedure I, using diaziridine **201** (333 mg, 0.772 mmol), iodine (235 mg, 0.927 mmol) and NEt_3 (0.32 mL, 2.32 mmol) in CH_2Cl_2 (20 mL). The crude product was purified by flash column chromatography (PE : EtOAc, 49:1 → 19:1) to yield diazirine **202** (328 mg, 0.764 mmol, 99%) as a white solid.

R_f 0.86 (PE : EtOAc, 19:1); **ν_{max}** (cm^{-1}) 2981 (C-H), 1251 (C-F), 1144, 955; **mp** 120-121 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 7.98 (2H, m, $H_{2'}$), 7.93 (2H, m, $H_{3'}$), 7.80 (1H, d, J = 8.3 Hz, H_5), 7.55 (1H, d, J = 1.7 Hz, H_2), 7.20 (1H, dd, J = 8.3, 1.7 Hz, H_6); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 164.0 (C_7), 150.7 (C_3), 143.3 (C_4), 138.4 ($C_{3'}$), 129.1 ($C_{2'}$), 126.3 (C_1),

126.0 ($C_{1'}$), 123.3 (C_6), 122.2 (q, $^1J_{C\text{-}F}$ = 278.2 Hz, C_9), 120.5 (C_5), 109.7 (C_2), 99.3 ($C_{4'}$), 28.9 (q, $^2J_{C\text{-}F}$ = 40.7 Hz, C_8); **^{19}FNMR** (376 MHz, $CDCl_3$) δ_F −65.3 (s); **HRMS** (ES^+) Calc. for $C_{15}H_8F_3IN_3O$ $[M+H]^+$ 429.9659, found 429.9652.

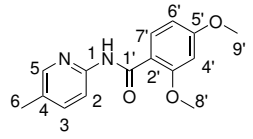
2-(4-Ethynylphenyl)-6-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]oxazole 130



Diazirine **130** was obtained according to General Procedure B, using diazirine **202** (218 mg, 0.508 mmol), $Pd(PPh_3)_2Cl_2$ (53.4 mg, 76.2 µmol), copper(I) iodide (9.67 mg, 50.8 µmol), NEt_3 (1.50 mL) and trimethylsilylacetylene (71 µL, 0.508 mmol) in THF (10 mL). The reaction was complete after 1 h. The TMS-protected alkyne intermediate formed was purified by flash column chromatography (PE : EtOAc, 49:1), then deprotected with TBAF (1.02 mL, 1 M in THF) in THF (5 mL). The reaction was complete after 30 min. The crude product was purified by flash column chromatography (PE : EtOAc, 99:1) to yield alkyne **130** (98.4 mg, 0.301 mmol, 59%) as a white solid.

R_f 0.53 (PE : EtOAc, 49:1); **ν_{max}** (cm^{-1}) 2980 (C-H), 1219, 1145 (C-F), 773; **mp** 115-116 °C; **^{1}HNMR** (500 MHz, $CDCl_3$) δ_H 8.23 (2H, m, $H_{2'}$), 7.80 (1H, dd, J = 8.4, 0.6 Hz, H_5), 7.68 (2H, m, $H_{3'}$), 7.55 (1H, dd, J = 1.7, 0.6 Hz, H_2), 7.21 (1H, dd, J = 8.4, 1.7 Hz, H_6), 3.29 (1H, s, $H_{6'}$); **^{13}CNMR** (125 MHz, $CDCl_3$) δ_C 163.9 (C_7), 150.7 (C_3), 143.4 (C_4), 132.7 ($C_{3'}$), 127.7 ($C_{2'}$), 126.5 ($C_{4'}$), 126.3 (C_1), 126.0 ($C_{1'}$), 123.2 (C_6), 122.2 (q, $^1J_{C\text{-}F}$ = 274.7 Hz, C_9), 120.6 (C_5), 109.7 (C_2), 82.8 ($C_{5'}$), 80.3 ($C_{6'}$), 28.9 (q, $^2J_{C\text{-}F}$ = 40.6 Hz, C_8); **^{19}FNMR** (400 MHz, $CDCl_3$) δ_F −65.3 (s); **HRMS** (ES^+) Calc. for $C_{17}H_9F_3N_3O$ $[M+H]^+$ 328.0692, found 328.0693.

2,4-Dimethoxy-*N*-(5-methylpyridin-2-yl)benzamide 148



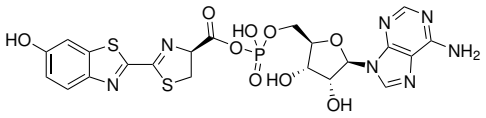
To a solution of 2,4-dimethoxybenzoic acid (500 mg, 2.74 mmol) in CH_2Cl_2 (25 mL) was added EDC hydrochloride (789 mg, 4.12 mmol), 2-amino-5-methylpyridine (296 mg, 2.74 mmol), HOBt (185 mg, 1.37 mmol) and NEt_3 (0.57 mL, 4.12 mmol). The reaction mixture was stirred at rt for 16 h, then washed with citric acid solution (10% aq.), $NaHCO_3$ solution

(10% aq.) and brine. The organic layer was concentrated *in vacuo*. The crude product was purified by flash column chromatography (CH_2Cl_2 : $(CH_3)_2CO$, 97:3 → 24:1) to yield amide **148** (386 mg, 1.42 mmol, 52%) as a white crystalline solid.

R_f 0.25 (CH_2Cl_2 : $(CH_3)_2CO$, 24:1); **mp** 104-105 °C; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 10.17 (1H, s, -NH), 8.31 (1H, d, J = 8.4 Hz, H_2), 8.24 (1H, d, J = 8.8 Hz, $H_{7'}$), 8.13 (1H, d, J = 2.4 Hz, H_5), 7.52 (1H, dd, J = 8.4, 2.4 Hz, H_3), 6.64 (1H, dd, J = 8.8, 2.3 Hz, $H_{6'}$), 6.53 (1H, d, J = 2.3 Hz, $H_{4'}$), 4.05 (3H, s, -OMe), 3.87 (3H, s, -OMe), 2.30 (3H, s, H_6); **LRMS** (ES^+) Calc. for $C_{15}H_{16}N_2O_3$ $[M+H]^+$ 273.1, found 273.1.

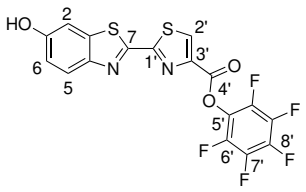
These data are in accordance with that of Heitman *et al.*[232]

Luciferyl adenylate (LH_2-AMP) 137



This compound was prepared as reported previously with some modifications.[356,357] Adenosine 5'-monophosphate sodium salt (18.8 mg, 48.1 µmol) was dissolved in 0.50 mL of degassed pyridine : 0.5 M HCl solution (4:1) at 40 °C and cooled to 0 °C. D-luciferin potassium salt (9.0 mg, 28.3 µmol) was added to the reaction mixture, followed by dicyclohexylcarbodiimide (250 mg, 1.21 mmol) in 0.40 mL pyridine. The reaction mixture was stirred at 0 °C for 30 min, followed by the addition of CH_2Cl_2 (1 mL) and 20 mM NaCl sodium acetate buffer (1 mL, pH 4.8). After IST phase separation, the aqueous layer was purified by reverse phase column chromatography (1% TFA in water : MeCN, 19:1 → 4:1) and lyophilised to yield LH_2-AMP as pale yellow crystalline solid (1.90 mg, 3.12 µmol, 11%). This compound has a half-life of 2.5 h at rt[357] as it readily oxidises, epimerises and hydrolyses and therefore was not fully characterised. LC-MS: RT (min) 1.6, MS (ES^+) Calc. for $C_{21}H_{20}N_7O_9PS_2$ $[M+H^+]$ 610.0 found 610.0.

Perfluorophenyl 2-(6-hydroxybenzo[*d*]thiazol-2-yl)thiazole-4-carboxylate 203



To a solution of dehydroluciferin* (40 mg, 0.144 mmol) and EDC hydrochloride (34.4 mg,

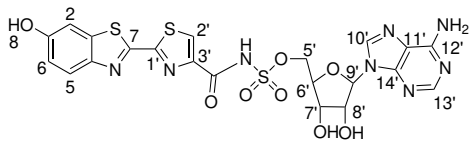
0.180 mmol) in pyridine (2 mL) was added pentafluorophenol (33.1 mg, 0.180 mmol). The reaction mixture was stirred for 16 h at rt, then concentrated *in vacuo*. The crude product was purified by flash column chromatography (PE : EtOAc, 7:3) to yield activated ester **203** (31.4 mg, 70.7 µmol, 49%) as a pale yellow solid.

*Synthesised by Adam Sills (summer student under my supervision) according to the procedure reported by Ciuffreda *et al.*.[358]

R_f 0.34 (EtOAc : PE, 7:3); **mp** >350 °C; **^{1}HNMR** (400 MHz, CD_3OD) δ_H 8.95 (1H, s, $H_{2'}$), 7.92 (1H, d, J = 8.9 Hz, H_3), 7.39 (1H, d, J = 2.4 Hz, H_6), 7.09 (1H, dd, J = 8.9, 2.4 Hz, H_1); **LRMS** (ES^+) Calc. for $C_{17}H_5F_5N_2O_3S_2$ $[M+H]^+$ 444.9, found 444.9.

These data are in accordance with that of Nakatsu *et al.*[359]

5-*O*-[(*N*-dehydroluciferyl)-sulfamoyl]adenosine (DLSA) 147



To a solution of 2′,3′-*O*-isopropylidene-5′-*O*-sulfamoyladenosine* (21.7 mg, 56.3 µmol) and DBU (12.8 mg, 84.5 µmol) in DMF (1 mL) stirred for 10 min at rt was added activated ester **203** (25 mg, 56.3 µmol) in DMF (0.50 mL). The reaction mixture was stirred for 16 h at rt, then pyridine (0.50 mL) was added and the mixture stirred for a further 4 h, followed by concentration *in vacuo*. The crude product was purified by flash column chromatography (CH_2Cl_2 : MeOH, 9:1 → 4:1) to yield isopropylidene-DLSA. This intermediate was dissolved in TFA (1 mL) and water (0.50 mL) and stirred for 1 h at rt before concentration *in vacuo*. The residue was azeotroped with ethanol and toluene, then washed with ice cold ethanol (4 × 0.50 mL) to yield **147** (3.46 mg, 5.71 µmol, 10%) as a pale yellow solid.

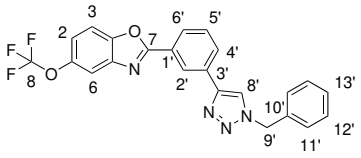
*Synthesised by Adam Sills (summer student under my supervision) according to the procedure reported by Heacock *et al.*.[360]

mp 229-230 °C (dec.); **^{1}HNMR** (500 MHz, $(CD_3)_2SO$) δ_H 10.16 (1H, *br* s, -OH), 9.36 (1H, *br* s, -NH), 9.25 (2H, *br* s, $-NH_2$), 8.68 (1H, s, $H_{2'}$), 8.55 (1H, s, $H_{10'}$), 8.45 (1H, s, $H_{13'}$), 7.94 (1H, d, J = 8.9 Hz, H_5), 7.48 (1H, d, J = 2.4 Hz, H_2), 7.05 (1H, dd, J = 8.9, 2.4 Hz, H_6), 6.40 (1H, s, $H_{9'}$), 5.74 (1H, d, J = 5.3 Hz, -OH), 5.47 (1H, d, J = 7.2 Hz, -OH), 4.92 (1H, m, $H_{5'a}$), 4.71-4.69 (1H, m, $H_{6'}$), 4.63-4.59 (1H, m, $H_{5'b}$), 4.29 (1H, m, $H_{7'}$), 3.97 (1H, dd, J = 5.3, 5.3 Hz, $H_{8'}$); **^{13}CNMR** (125 MHz, $(CD_3)_2SO$) δ_C 160.9 ($C_{14'}$), 157.0 (C_7), 156.9 (C_1), 156.6 ($C_{4'}$), 156.3 ($C_{1'}$), 150.8 ($C_{11'}$), 149.2 ($C_{2'}$), 146.4 (C_4), 139.4 ($C_{3'}$), 139.3 ($C_{13'}$), 136.8

(C_3), 127.1 ($C_{10'}$), 124.2 (C_5), 119.5 ($C_{12'}$), 117.1 (C_6), 107.0 (C_2), 93.2 ($C_{9'}$), 83.2 ($C_{6'}$), 75.4 ($C_{8'}$), 70.0 ($C_{7'}$), 58.2 ($C_{5'}$); **LRMS** (ES^+) Calc. for $[M+H]^+$ 607.0, found 607.0.

These data are in accordance with that presented by Branchini *et al.*[223]

2-(3-(1-Benzyl-1*H*-1,2,3-triazol-4-yl)phenyl)-5-(trifluoromethoxy)benzo[*d*]-oxazole 156



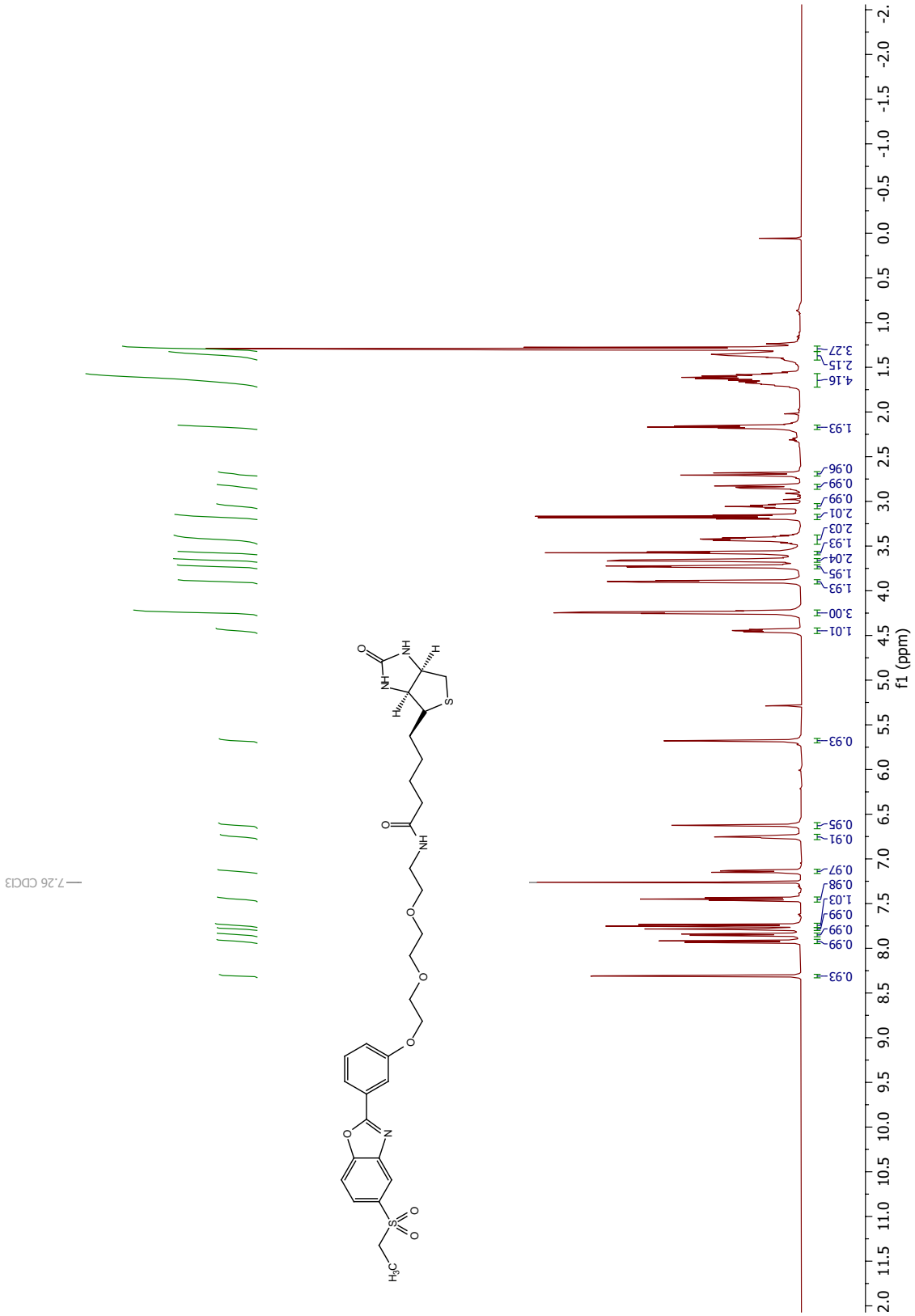
To a solution of alkyne **94** (29.4 mg, 97.0 µmol), anhydrous copper(II) sulphate (3.03 mg, 19.4 µmol) and sodium ascorbate (3.76 mg, 19.4 µmol) in THF:H_2O (2 mL, 10:1) was added benzyl azide (12.0 µL, 97.0 µmol). The reaction mixture was stirred for 2 h, then concentrated *in vacuo*. The crude product was purified by flash column chromatography (PE : EtOAc, 9:1 → 3:1) to yield triazole **156** (36.5 mg, 83.7 µmol, 86%) as a white solid.

R_f 0.12 (PE : EtOAc, 9:1); **ν_{max}** (cm^{-1}) 3135 (alkene C-H), 2980 (C-H), 1340, 1253; **mp** 151-153 °C; **^{1}HNMR** (400 MHz, $CDCl_3$) δ_H 8.57 (1H, dd, J = 1.7, 1.7 Hz, $H_{2'}$), 8.18 (1H, ddd, J = 7.8, 1.2, 1.2 Hz, $H_{6'}$), 8.10 (1H, ddd, J = 7.8, 1.4, 1.4 Hz, $H_{4'}$), 7.81 (1H, s, $H_{8'}$), 7.63 (1H, d, J = 2.5 Hz, H_6), 7.58 (1H, dd, J = 8.1, 8.1 Hz, $H_{5'}$), 7.57 (1H, d, J = 8.5 Hz, H_3), 7.44-7.38 (3H, m, $H_{11',13'}$), 7.36-7.32 (2H, m, $H_{12'}$), 7.24 (1H, dd, J = 8.5, 2.5 Hz, H_2), 5.61 (2H, s, $H_{9'}$); **^{13}CNMR** (100 MHz, $CDCl_3$) δ_C 164.7 (C_7), 149.2 ($C_{7'}$), 147.3 (C_4), 146.3 (C_1), 142.9 (C_5), 134.6 ($C_{10'}$), 131.7 ($C_{3'}$), 129.8 ($C_{5'}$), 129.4 ($C_{11'}$), 129.3 ($C_{13'}$), 129.1 ($C_{1'}$), 128.3 ($C_{12'}$), 127.4 ($C_{6'}$), 127.2 ($C_{4'}$), 124.9 ($C_{2'}$), 120.4 (q, $^1J_{C\text{-}F}$ = 257.2, C_8), 120.1 ($C_{8'}$), 119.0 (C_2), 113.2 (C_6), 111.2 (C_3), 54.5 ($C_{9'}$); **HRMS** (ES^+) Calc. for $C_{23}H_{16}F_3N_4O_2$ $[M+H]^+$ 437.1219, found 437.1215.

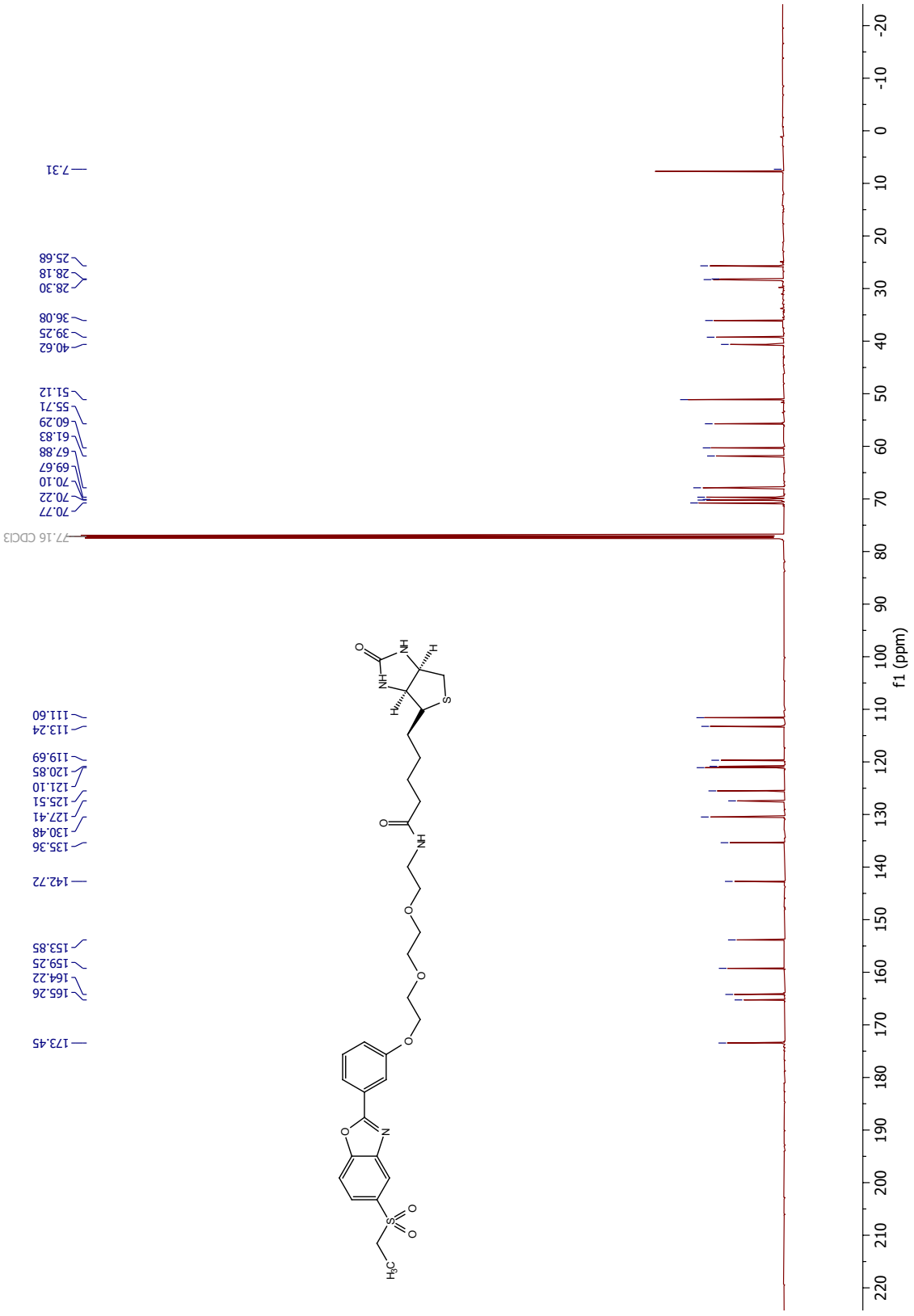
8.2.4 Spectra of key compounds

Biotinylated probe 42

^{1}H spectrum

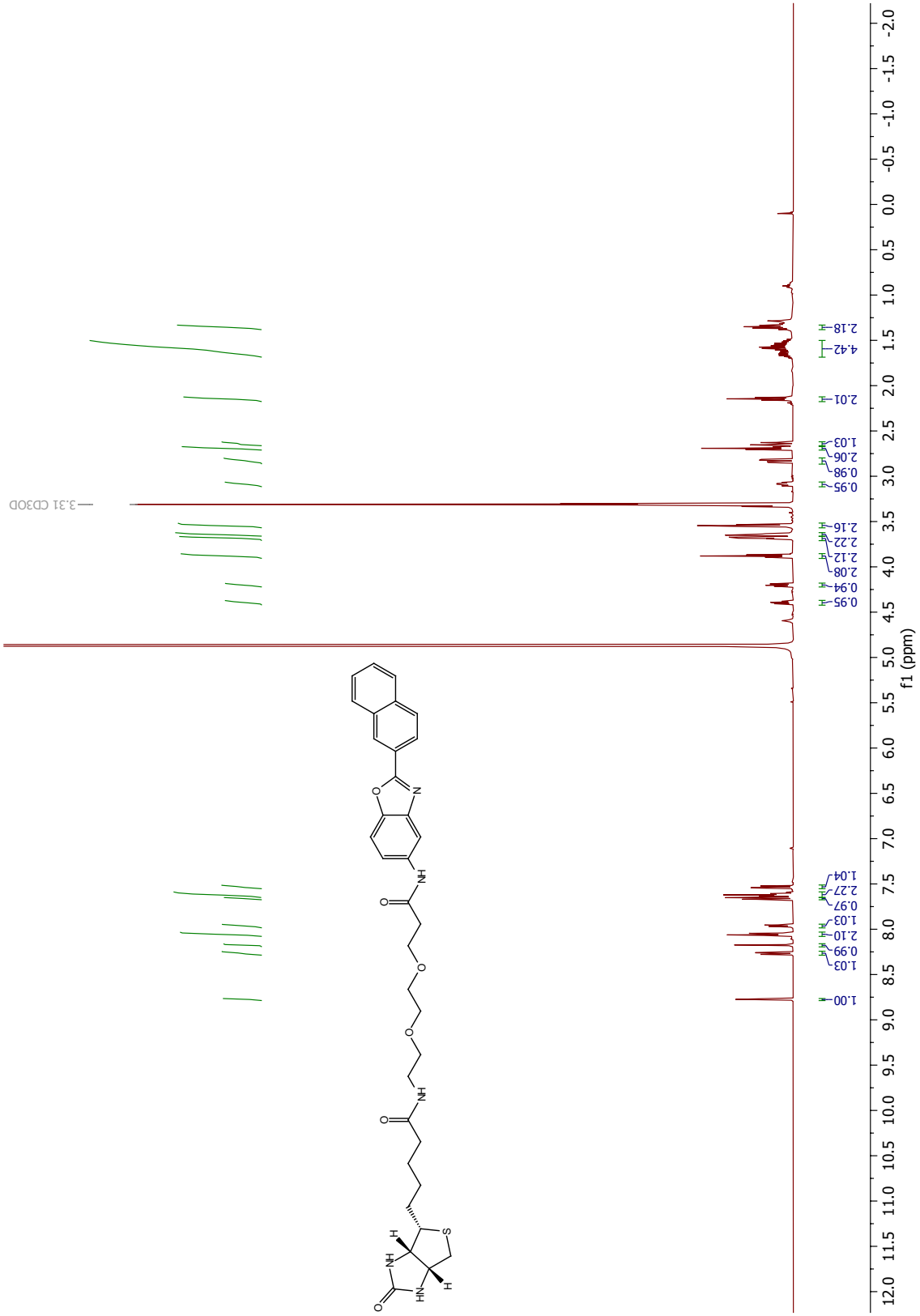


^{13}C spectrum

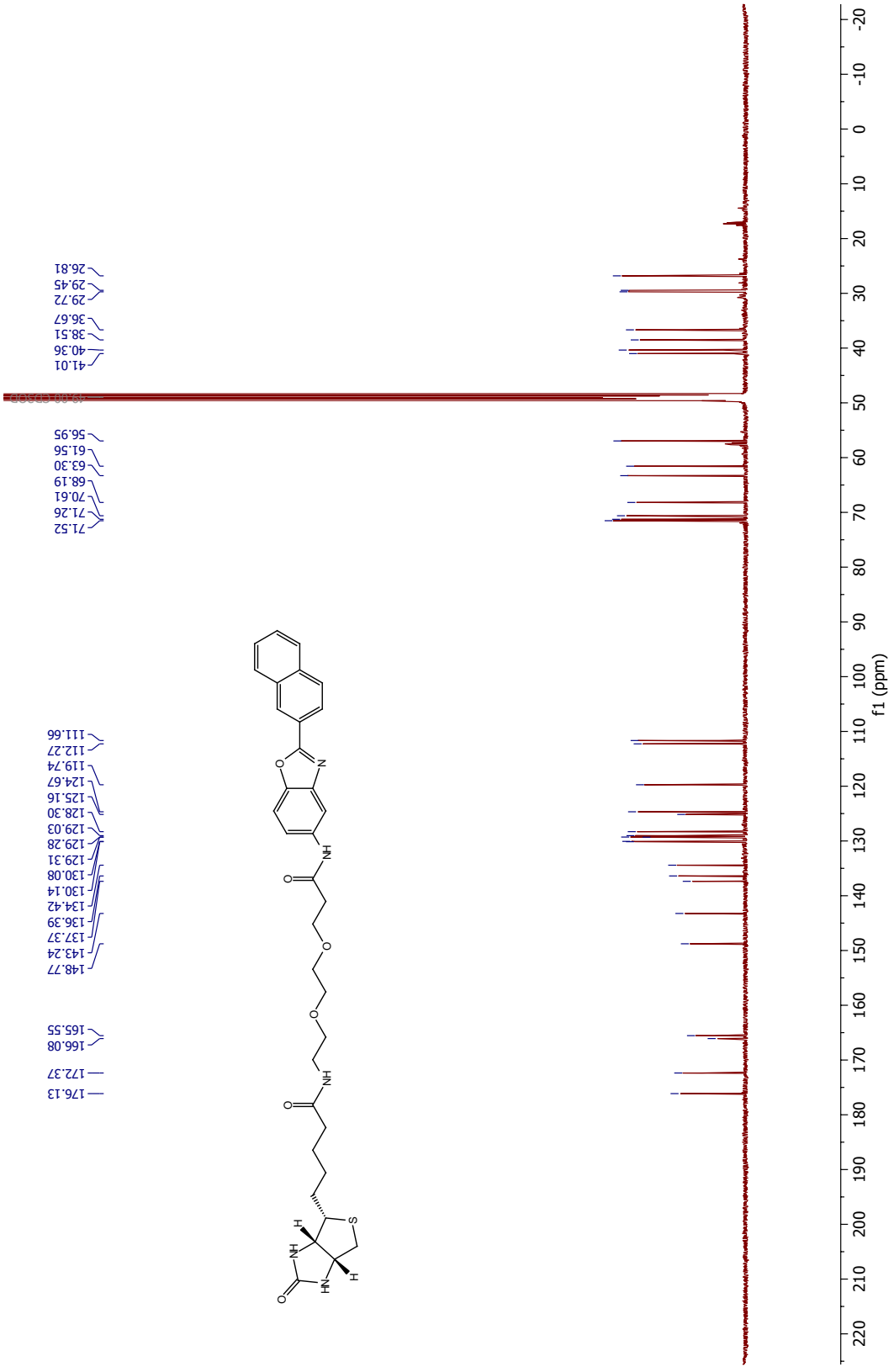


Biotinylated probe 41

^{1}H spectrum

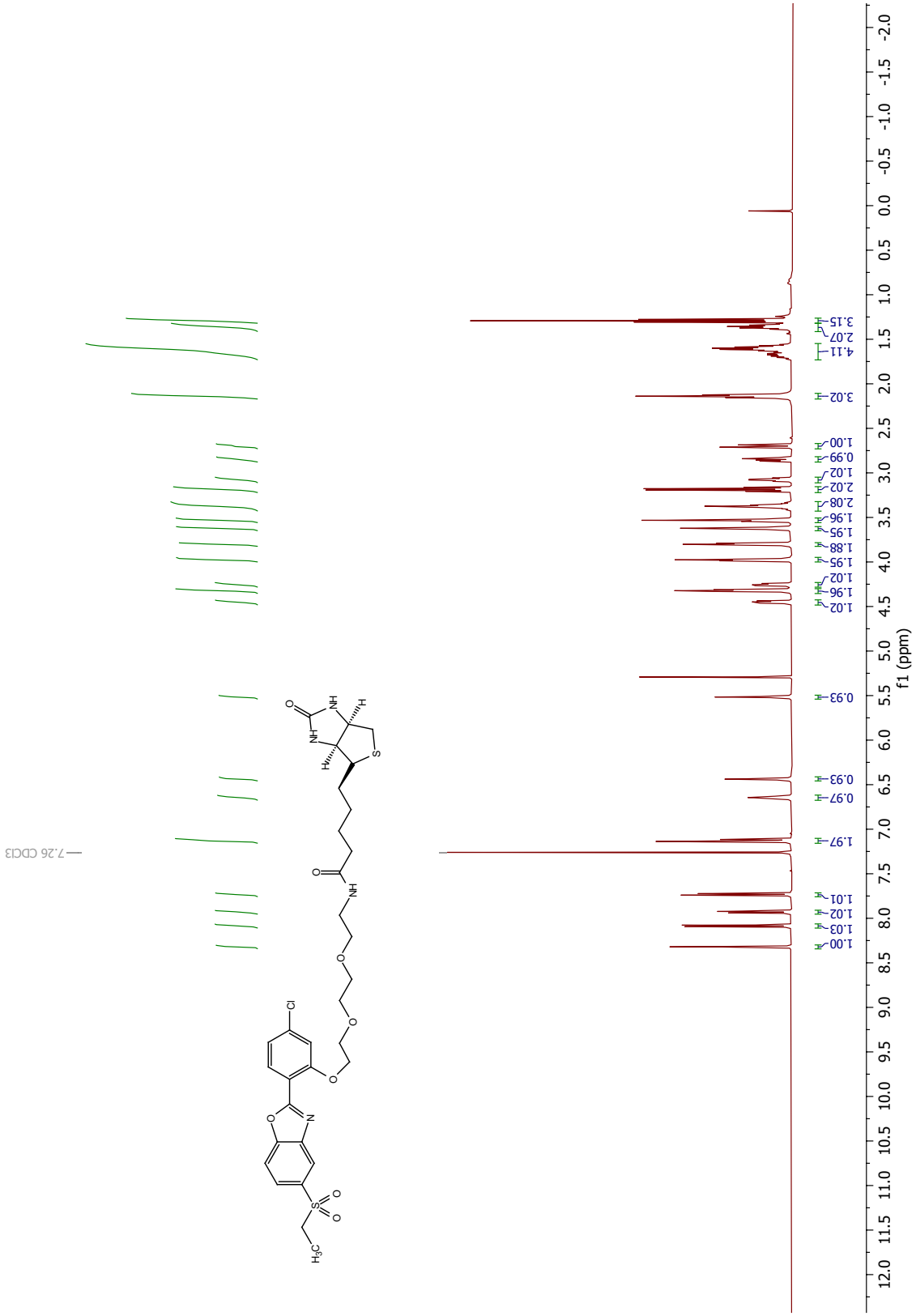


^{13}C spectrum

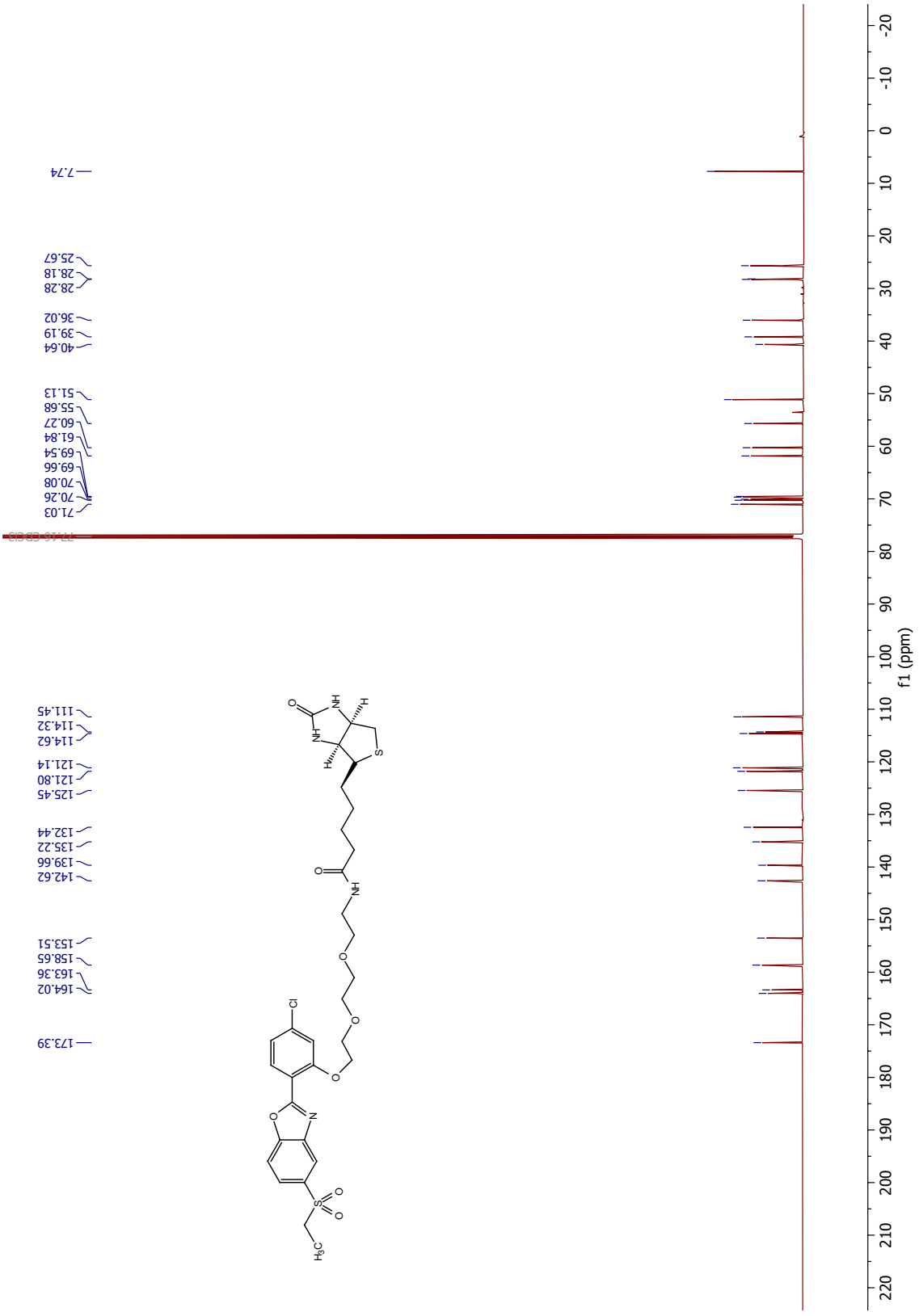


Biotinylated probe 43

^{1}H spectrum

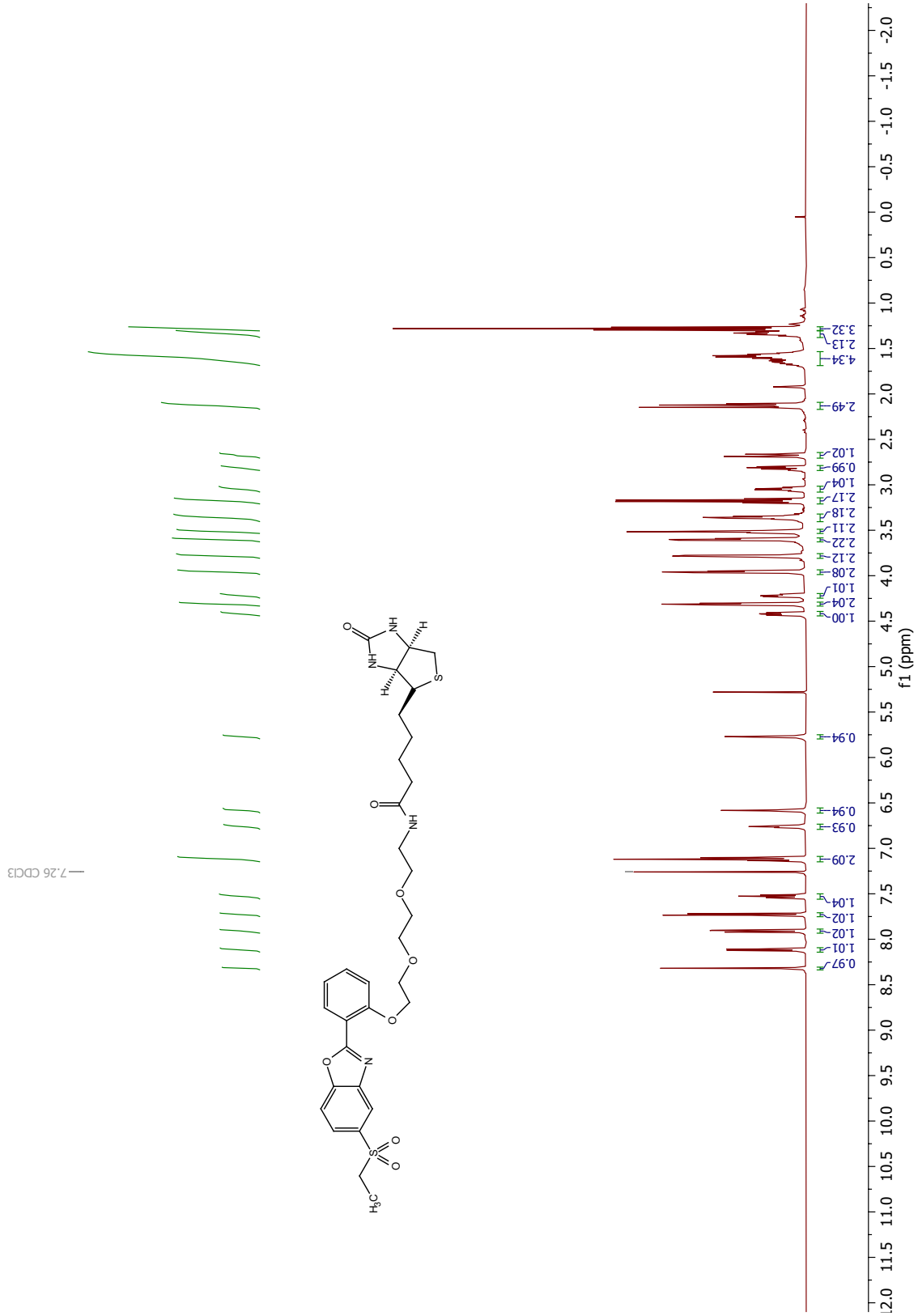


^{13}C spectrum

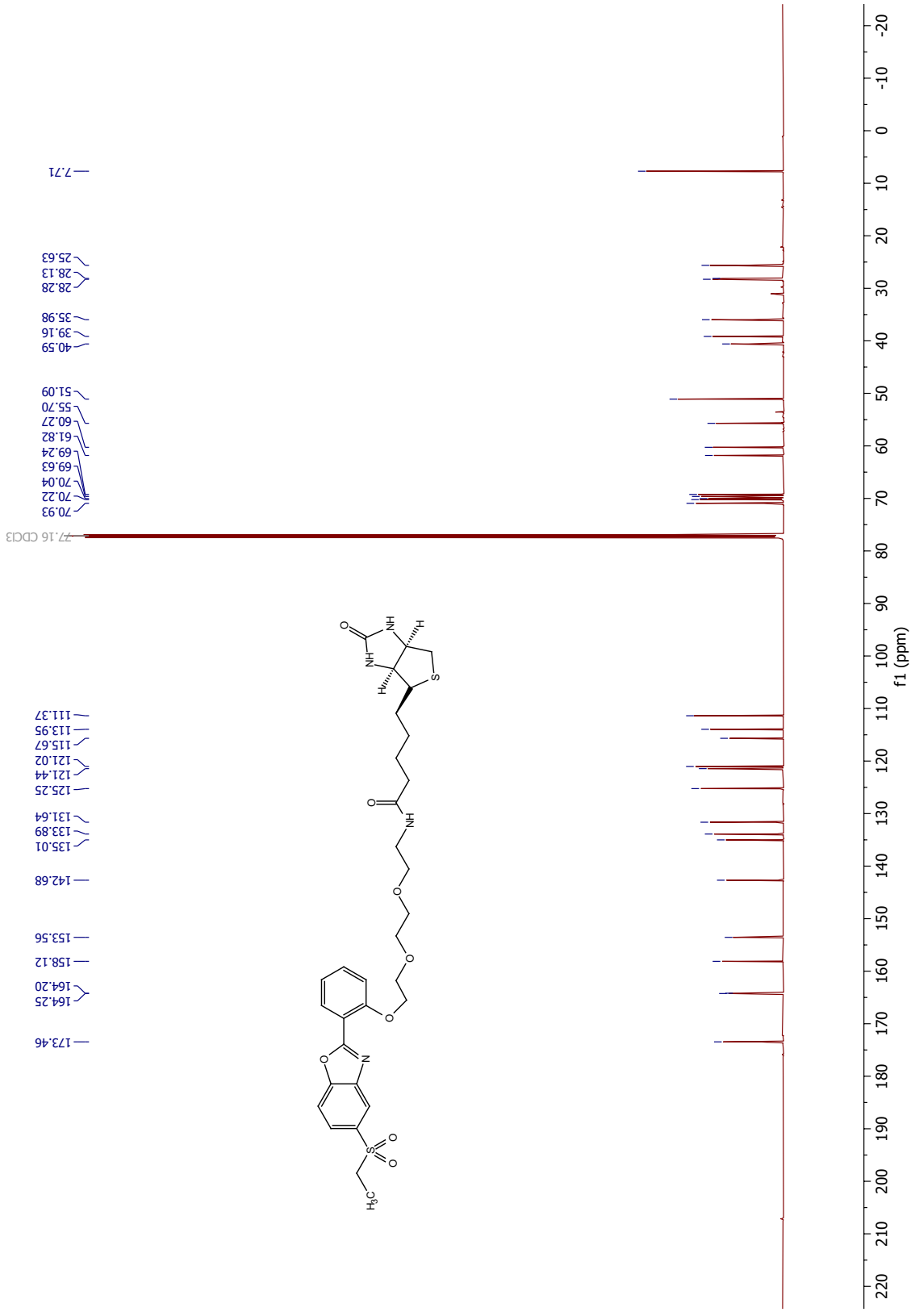


Biotinylated probe 75

^{1}H spectrum

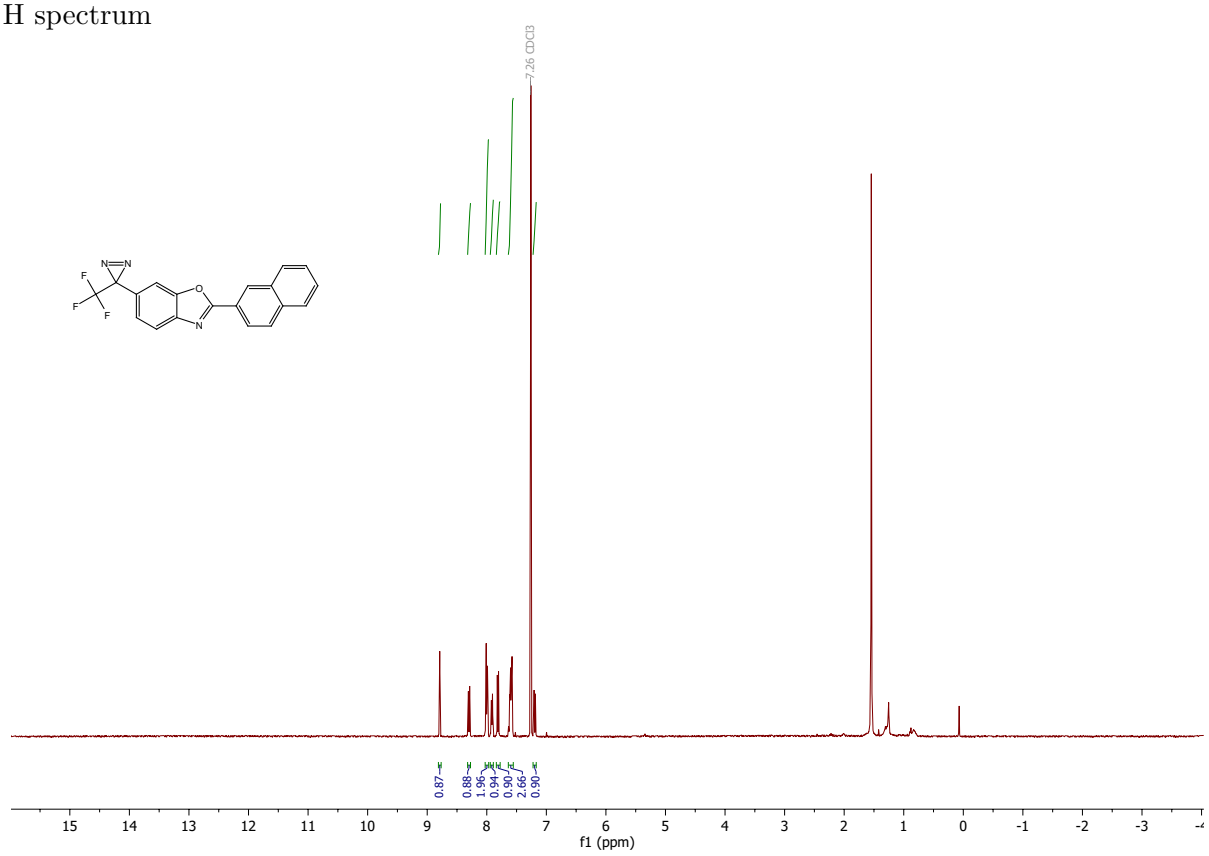


^{13}C spectrum

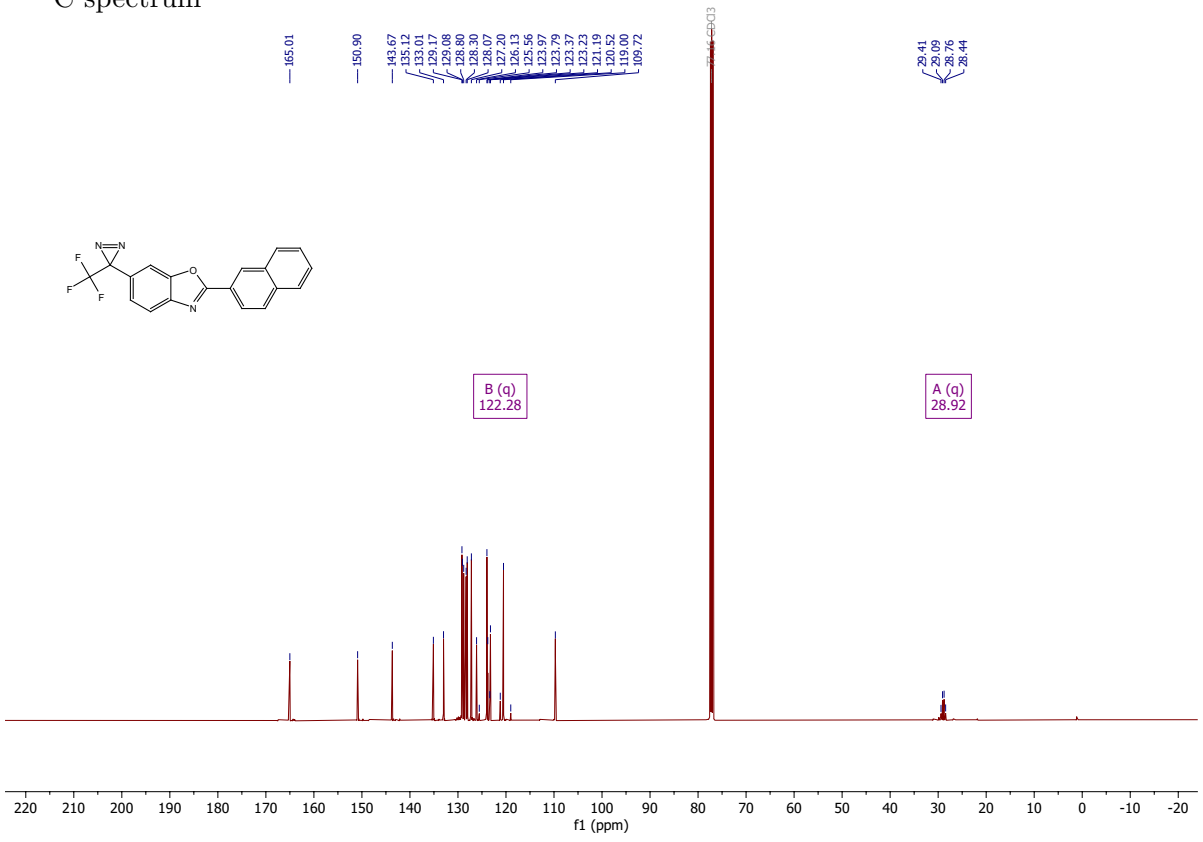


2-(Naphthalen-2-yl)-6-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]oxazole 81

^{1}H spectrum

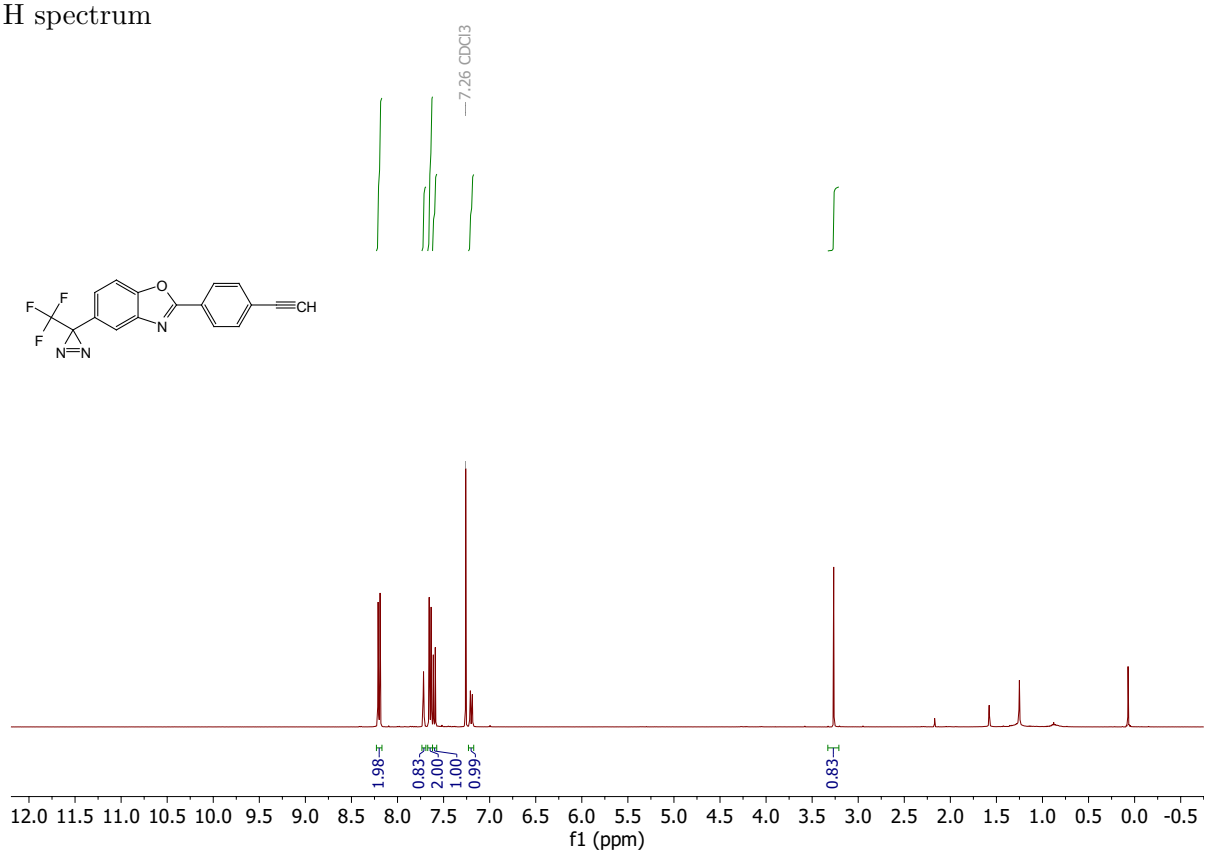


^{13}C spectrum

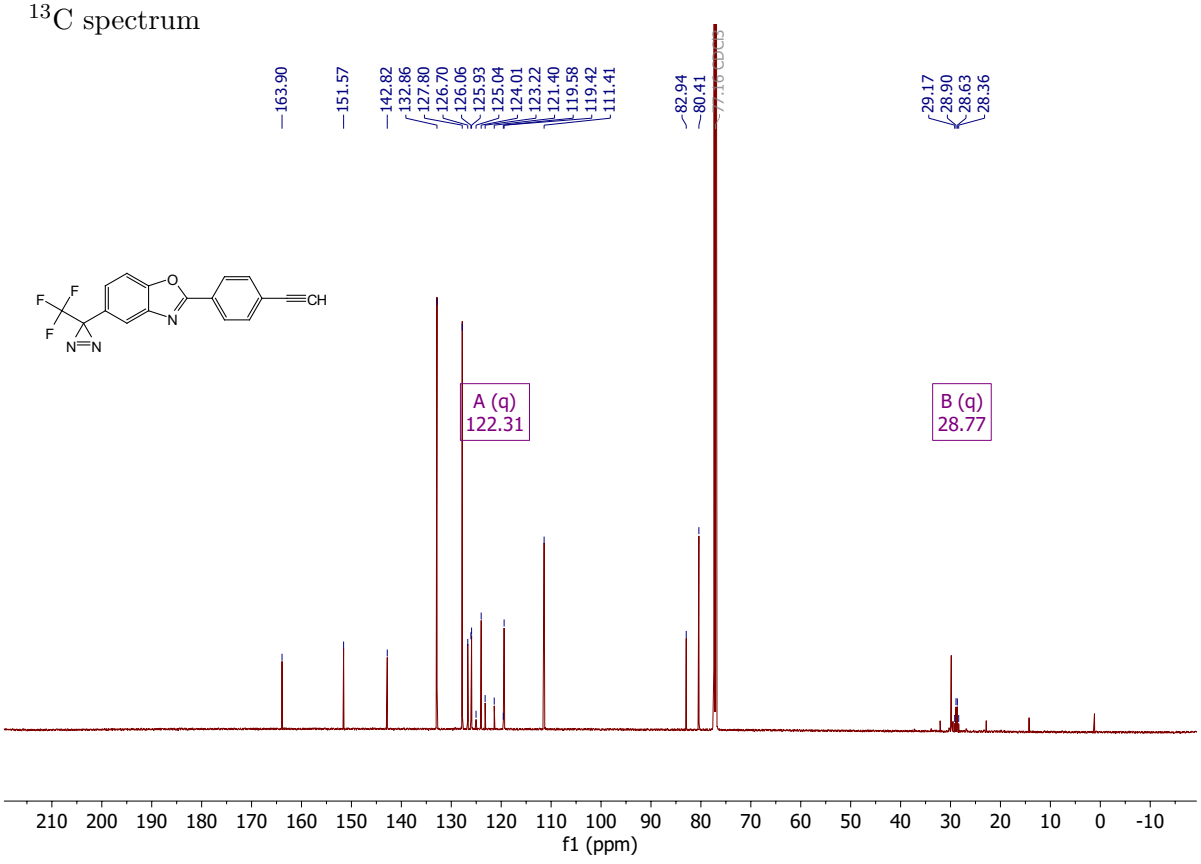


2-(4-Ethynylphenyl)-5-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]oxazole 106

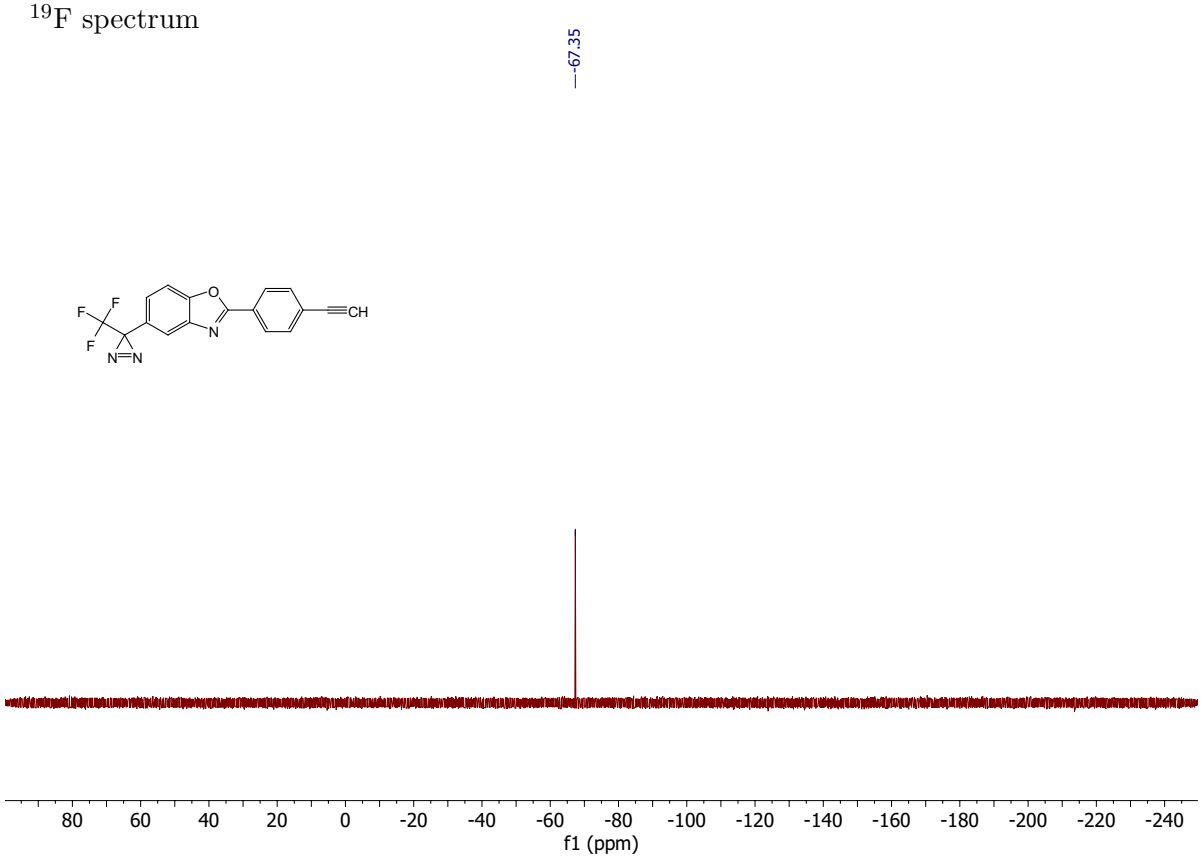
^{1}H spectrum



^{13}C spectrum

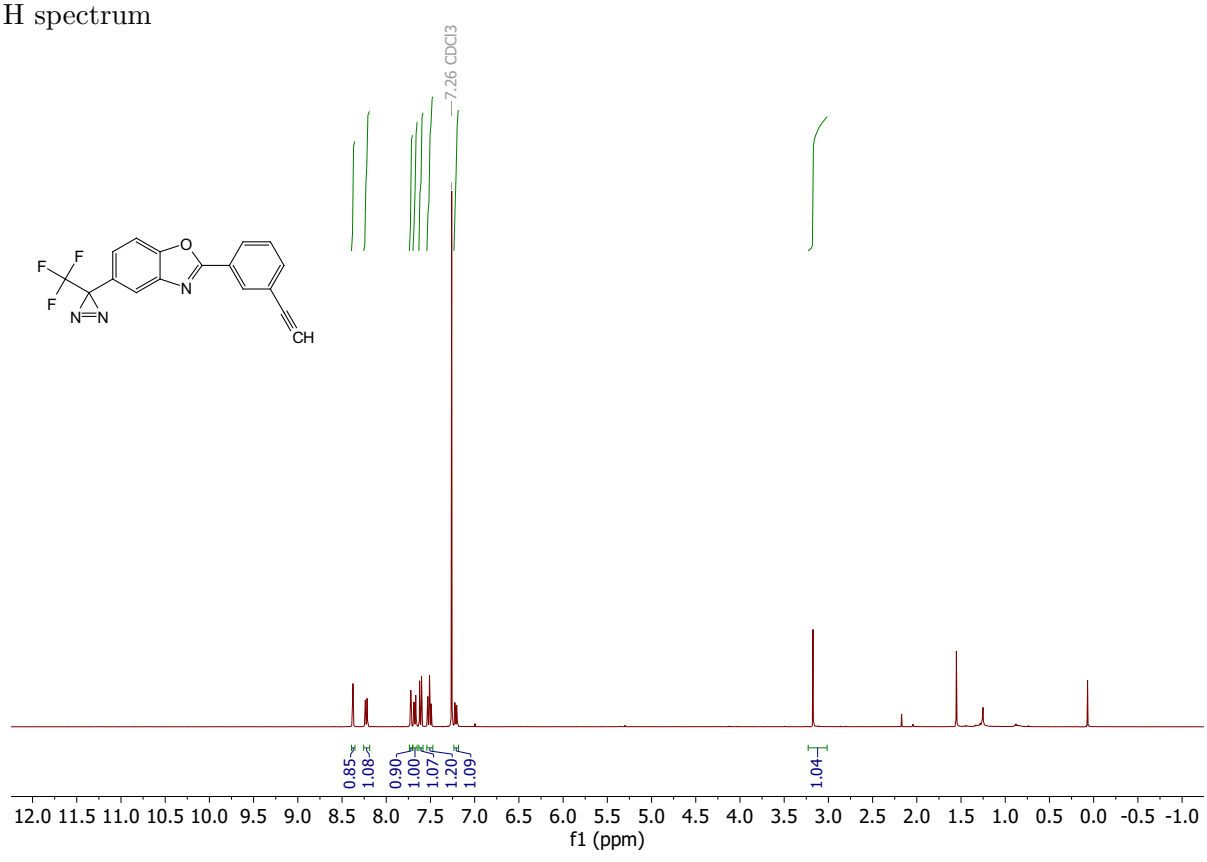


^{19}F spectrum

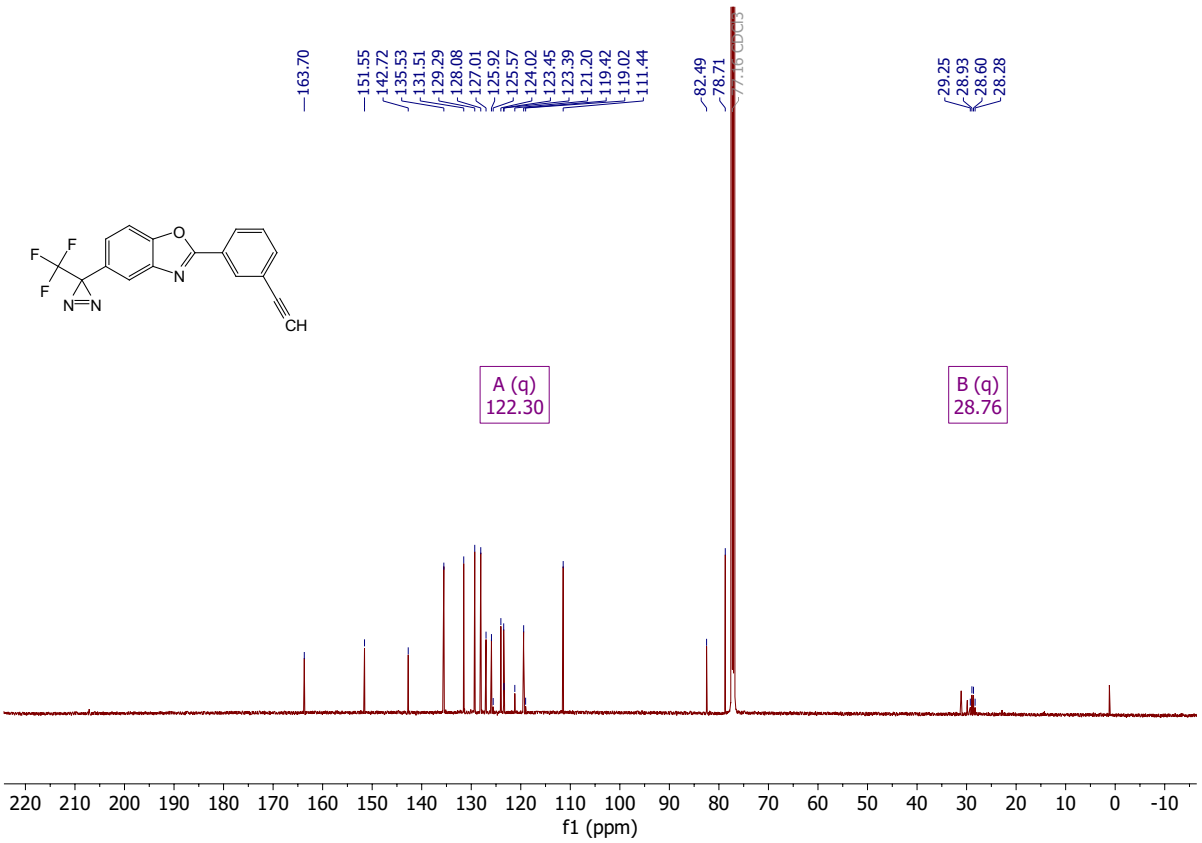


2-(3-Ethynylphenyl)-5-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]oxazole 107

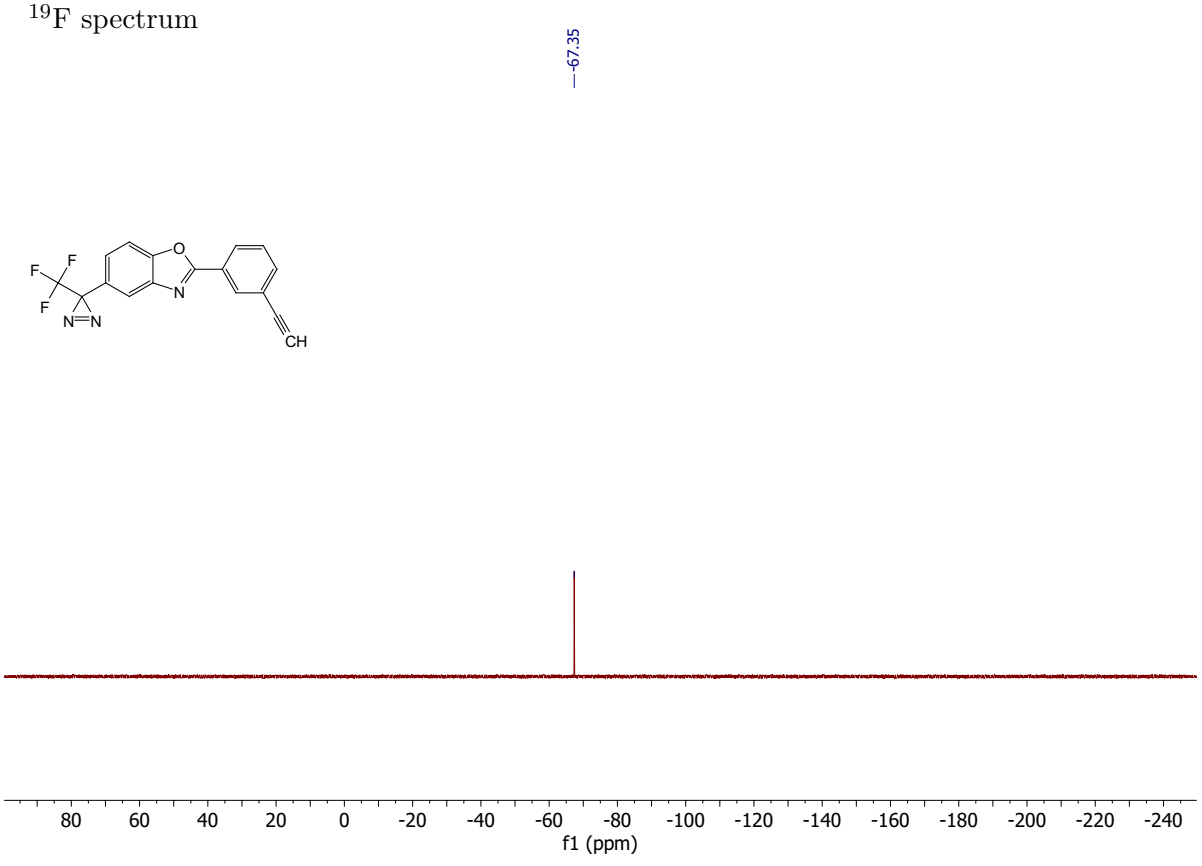
^{1}H spectrum



^{13}C spectrum

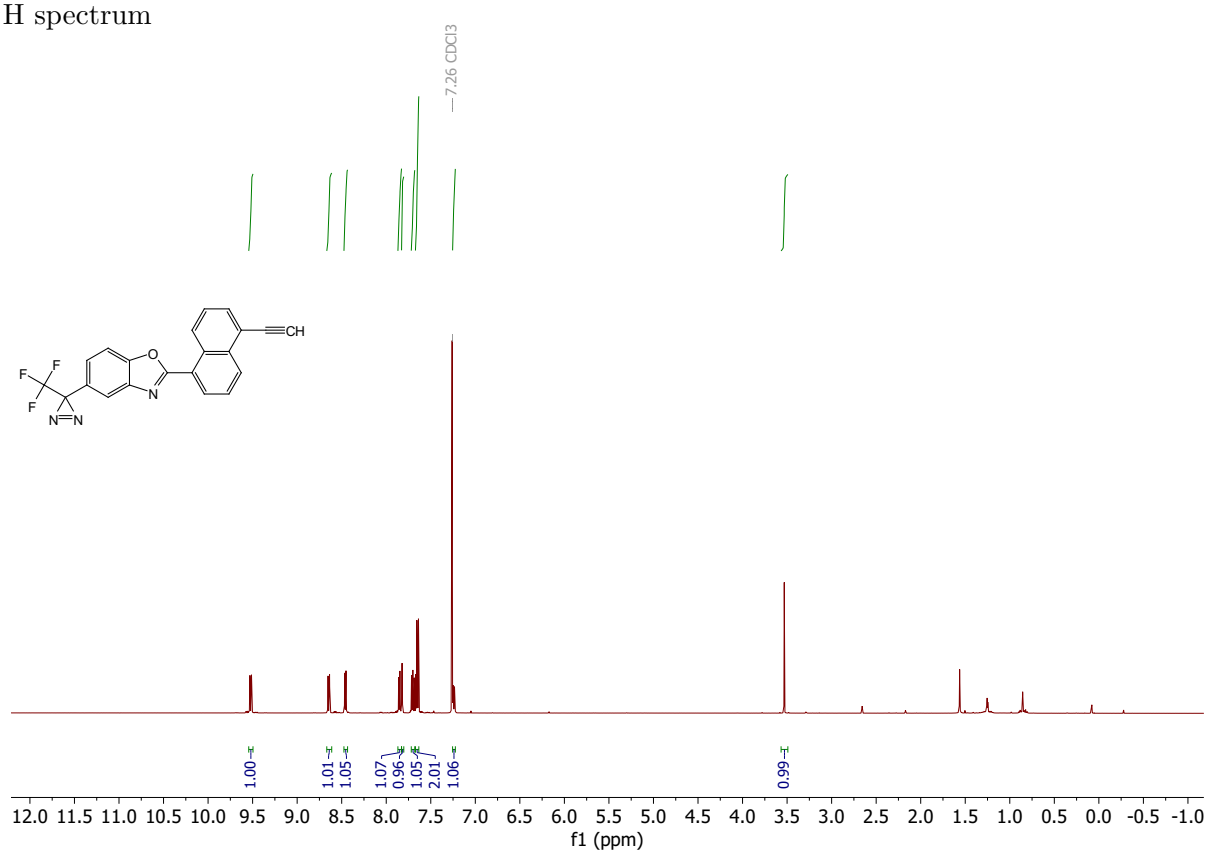


^{19}F spectrum

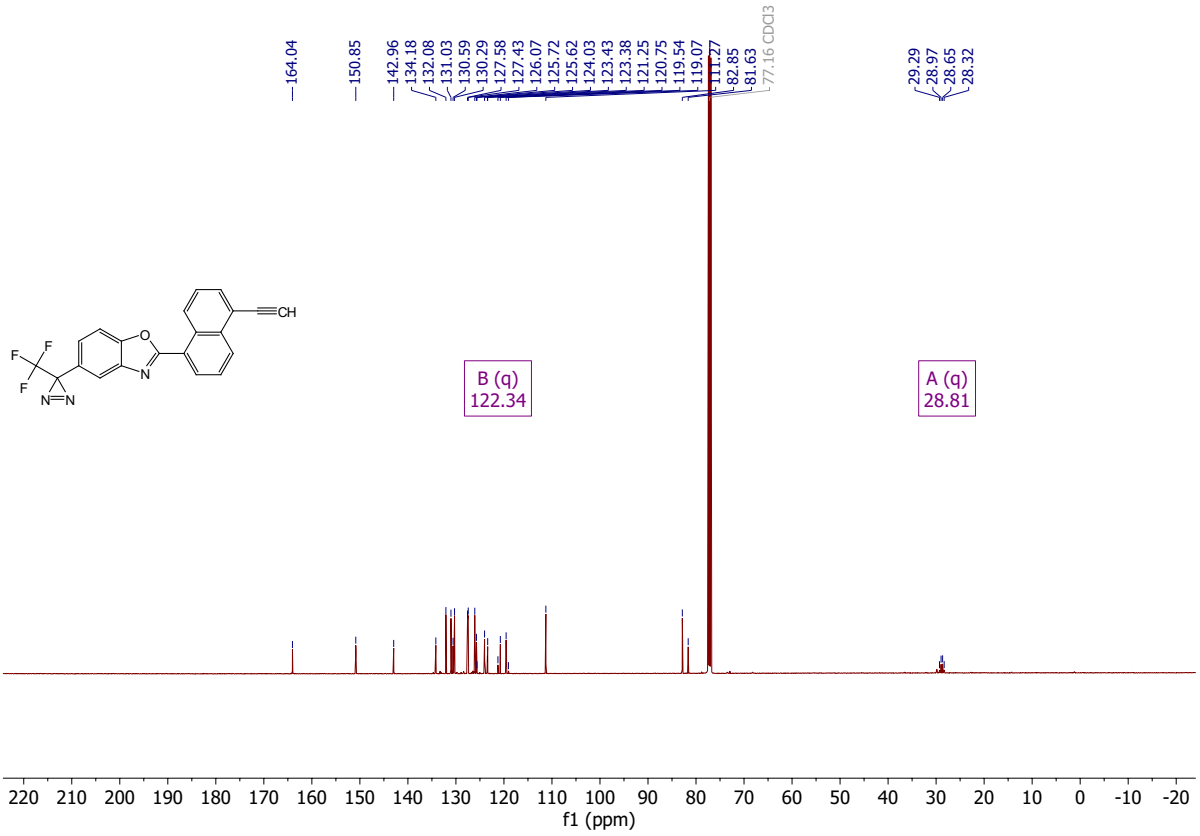


2-(5-Ethynylnaphthalen-1-yl)-5-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzo[*d*]-oxazole 129

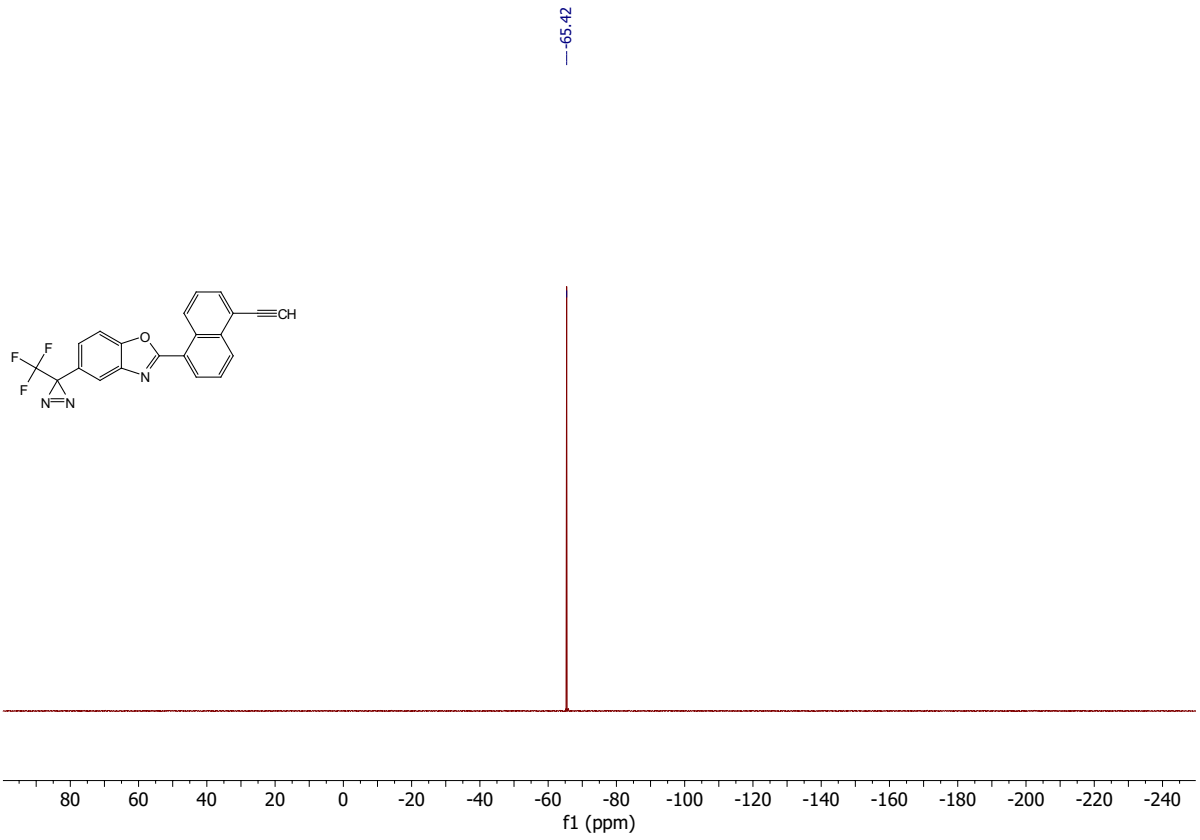
^{1}H spectrum



^{13}C spectrum

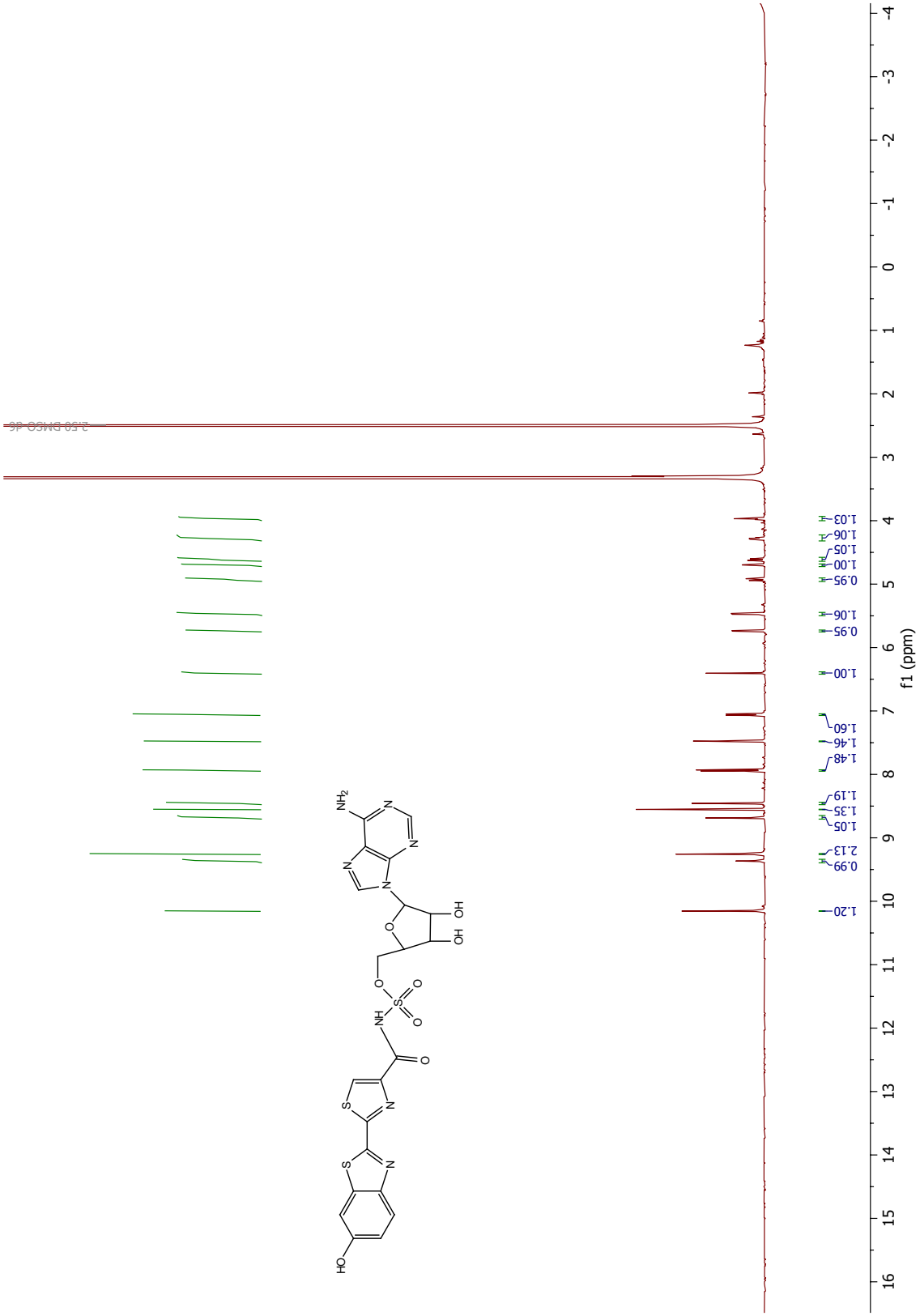


^{19}F spectrum

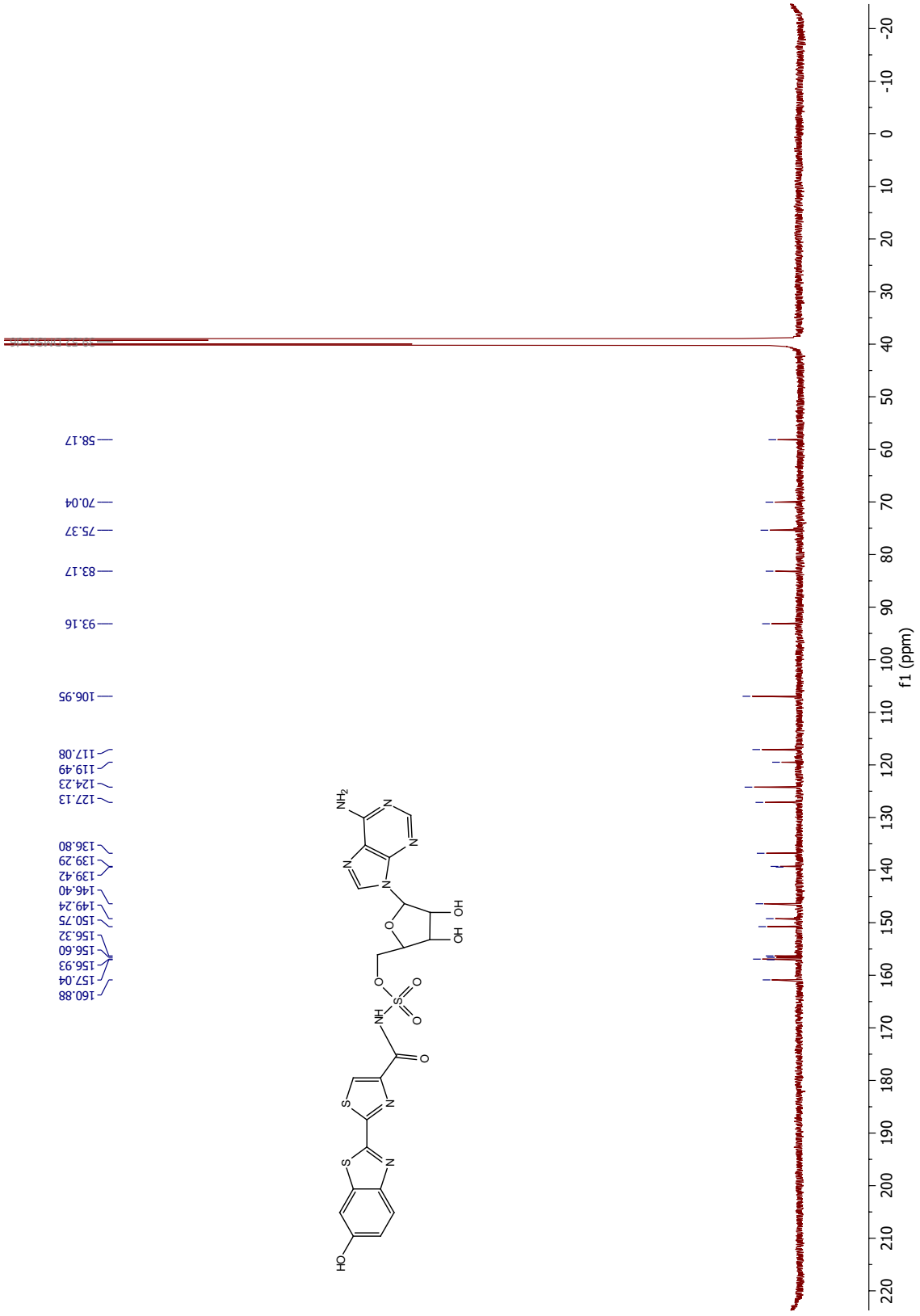


5-*O*-[*N*-Dehydroluciferyl)-sulfamoyl]adenosine (DLSA) 147

^{1}H spectrum



^{13}C spectrum



Bibliography

(1) Emery, A. E. Population frequencies of inherited neuromuscular diseases–a world survey. *Neuromuscul. Disord.* **1991**, *1*, 19–29.

(2) Mendell, J. R.; Shilling, C.; Leslie, N. D.; Flanigan, K. M., et al. Evidence-based path to newborn screening for Duchenne muscular dystrophy. *Ann. Neurol.* **2012**, *71*, 304–313.

(3) Van Ruiten, H. J. A.; Straub, V.; Bushby, K.; Guglieri, M. Improving recognition of Duchenne muscular dystrophy: a retrospective case note review. *Arch. Dis. Child.* **2014**, *99*, 1074–7.

(4) Kohler, M.; Clarenbach, C. F.; Bahler, C.; Brack, T., et al. Disability and survival in Duchenne muscular dystrophy. *J. Neurol. Neurosurg. Psychiatry* **2009**, *80*, 320–325.

(5) McDonald, C. M.; Henricson, E. K.; Abresch, R. T.; Han, J. J., et al. The cooperative international neuromuscular research group duchenne natural history study-a longitudinal investigation in the era of glucocorticoid therapy: Design of protocol and the methods used. *Muscle Nerve* **2013**, *48*, 32–54.

(6) Bladen, C. L.; Salgado, D.; Monges, S.; Foncuberta, M. E., et al. The TREAT-NMD DMD Global Database: Analysis of More than 7,000 Duchenne Muscular Dystrophy Mutations. *Hum. Mutat.* **2015**, *36*, 395–402.

(7) Davie, A. M.; Emery, A. E. Estimation of proportion of new mutants among cases of Duchenne muscular dystrophy. *J. Med. Genet.* **1978**, *15*, 339.

(8) Koenig, M.; Beggs, A. H.; Moyer, M.; Scherpf, S., et al. The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. *Am. J. Hum. Genet.* **1989**, *45*, 498–506.

(9) Yazaki, M.; Yoshida, K.; Nakamura, A.; Koyama, J., et al. Clinical Characteristics of Aged Becker Muscular Dystrophy Patients with Onset after 30 Years. *Eur. Neurol.* **1999**, *42*, 145–149.

(10) Vuorinen, A. Towards the Elucidation of the Mechanism of Action of Small Molecule Upregulators of Utrophin using Chemical Proteomics. PhD thesis, 2017, University of Oxford, UK.

(11) Cohn, R. D.; Campbell, K. P. Molecular basis of muscular dystrophies. *Muscle Nerve* **2000**, *23*, 1456–71.

(12) Huang, X.; Poy, F.; Zhang, R.; Joachimiak, A., et al. Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. *Nat. Struct. Biol.* **2000**, *7*, 634–638.

(13) Albrecht, D. E.; Froehner, S. C. Syntrophins and Dystrobrevins: Defining the Dystrophin Scaffold at Synapses. *Neurosignals* **2002**, *11*, 123–129.

(14) Blat, Y.; Blat, S. Drug Discovery of Therapies for Duchenne Muscular Dystrophy. *J. Biomol. Screen.* **2015**, *20*, 1189–1203.

(15) Bulfield, G.; Siller, W. G.; Wight, P. A.; Moore, K. J. X chromosome-linked muscular dystrophy (mdx) in the mouse. *Proc. Natl. Acad. Sci.* **1984**, *81*, 1189–1192.

(16) Larcher, T.; Lafoux, A.; Tesson, L.; Remy, S., et al. Characterization of Dystrophin Deficient Rats: A New Model for Duchenne Muscular Dystrophy. *PLoS One* **2014**, *9*, ed. by Asakura, A., e110371.

(17) Cooper, B. J.; Winand, N. J.; Stedman, H.; Valentine, B. A., et al. The homologue of the Duchenne locus is defective in X-linked muscular dystrophy of dogs. *Nature* **1988**, *334*, 154–156.

(18) McGreevy, J. W.; Hakim, C. H.; McIntosh, M. A.; Duan, D. Animal models of Duchenne muscular dystrophy: from basic mechanisms to gene therapy. *Dis. Model. Mech.* **2015**, *8*, 195–213.

(19) Deconinck, A. E.; Rafael, J. A.; Skinner, J. A.; Brown, S. C., et al. Utrophin-Dystrophin-Deficient Mice as a Model for Duchenne Muscular Dystrophy. *Cell* **1997**, *90*, 717–727.

(20) Chamberlain, J.; Pearlman, J.; Muzny, D.; Gibbs, R., et al. Expression of the murine Duchenne muscular dystrophy gene in muscle and brain. *Science (80-.).* **1988**, *239*, 1416–1418.

(21) Bushby, K. M.; Appleton, R.; Anderson, L. V.; Welch, J. L., et al. Deletion status and intellectual impairment in Duchenne muscular dystrophy. *Dev. Med. Child Neurol.* **1995**, *37*, 260–9.

(22) Milic Rasic, V.; Vojinovic, D.; Pesovic, J.; Mijalkovic, G., et al. Intellectual ability in the Duchenne muscular dystrophy and dystrophin gene mutation location. *Balkan J. Med. Genet.* **2014**, *17*, 25–35.

(23) Hendriksen, R. G.; Vles, J. S.; Aalbers, M. W.; Chin, R. F., et al. Brain-related co-morbidities in boys and men with Duchenne Muscular Dystrophy: A descriptive study. *Eur. J. Paediatr. Neurol.* **2018**, *22*, 488–497.

(24) Desguerre, I.; Christov, C.; Mayer, M.; Zeller, R., et al. Clinical heterogeneity of duchenne muscular dystrophy (DMD): definition of sub-phenotypes and predictive criteria by long-term follow-up. *PLoS One* **2009**, *4*, e4347.

(25) Bushby, K.; Finkel, R.; Birnkrant, D. J.; Case, L. E., et al. Diagnosis and management of Duchenne muscular dystrophy, part 2: implementation of multidisciplinary care. *Lancet Neurol.* **2010**, *9*, 177–189.

(26) Bushby, K.; Finkel, R.; Birnkrant, D. J.; Case, L. E., et al. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. *Lancet Neurol.* **2010**, *9*, 77–93.

(27) Moxley, R. T.; Pandya, S.; Ciafaloni, E.; Fox, D. J., et al. Change in Natural History of Duchenne Muscular Dystrophy With Long-term Corticosteroid Treatment: Implications for Management. *J. Child Neurol.* **2010**, *25*, 1116–1129.

(28) Matthews, E.; Brassington, R.; Kuntzer, T.; Jichi, F., et al. Corticosteroids for the treatment of Duchenne muscular dystrophy. *Cochrane Database Syst. Rev.* **2016**.

(29) Angelini, C. The role of corticosteroids in muscular dystrophy: A critical appraisal. *Muscle Nerve* **2007**, *36*, 424–435.

(30) Guiraud, S.; Davies, K. E. Pharmacological advances for treatment in Duchenne muscular dystrophy. *Curr. Opin. Pharmacol.* **2017**, *34*, 36–48.

(31) Translarna | European Medicines Agency. https://www.ema.europa.eu/en/medicines/human/EPAR/translarna. Accessed 14/01/19.

(32) Finkel, R. S.; Flanigan, K. M.; Wong, B.; Bönnemann, C., et al. Phase 2a Study of Ataluren-Mediated Dystrophin Production in Patients with Nonsense Mutation Duchenne Muscular Dystrophy. *PLoS One* **2013**, *8*, ed. by Sawada, H., e81302.

(33) Bushby, K.; Finkel, R.; Wong, B.; Barohn, R., et al. Ataluren treatment of patients with nonsense mutation dystrophinopathy. *Muscle Nerve* **2014**, *50*, 477–487.

(34) McDonald, C. M.; Campbell, C.; Torricelli, R. E.; Finkel, R. S., et al. Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* **2017**, *390*, 1489–1498.

(35) Landfeldt, E.; Sejersen, T.; Tulinius, M. A mini-review and implementation model for using ataluren to treat nonsense mutation Duchenne muscular dystrophy. *Acta Paediatr.* **2019**, *108*, 224–230.

(36) Lu, Q.-L.; Cirak, S.; Partridge, T. What Can We Learn From Clinical Trials of Exon Skipping for DMD? *Mol. Ther. Nucleic Acids* **2014**, *3*, e152.

(37) Mendell, J. R.; Rodino-Klapac, L. R.; Sahenk, Z.; Roush, K., et al. Eteplirsen for the treatment of Duchenne muscular dystrophy. *Ann. Neurol.* **2013**, *74*, 637–647.

(38) Charleston, J. S.; Schnell, F. J.; Dworzak, J.; Donoghue, C., et al. Eteplirsen treatment for Duchenne muscular dystrophy: Exon skipping and dystrophin production. *Neurology* **2018**, *90*, e2146–e2154.

(39) McNally, E. M.; Wyatt, E. J. Mutation-Based Therapy for Duchenne Muscular Dystrophy: Antisense Treatment Arrives in the Clinic. *Circulation* **2017**, *136*, 979–981.

(40) Lim, K. R. Q.; Maruyama, R.; Yokota, T. Eteplirsen in the treatment of Duchenne muscular dystrophy. *Drug Des. Devel. Ther.* **2017**, *11*, 533–545.

(41) Dowling, J. J. Eteplirsen therapy for Duchenne muscular dystrophy: skipping to the front of the line. *Nat. Rev. Neurol.* **2016**, *12*, 675–676.

(42) Unger, E. F.; Califf, R. M. Regarding "Eteplirsen for the treatment of Duchenne muscular dystrophy". *Ann. Neurol.* **2017**, *81*, 162–164.

(43) Dalakas, M. C. Gene therapy for Duchenne muscular dystrophy: balancing good science, marginal efficacy, high emotions and excessive cost. *Ther. Adv. Neurol. Disord.* **2017**, *10*, 293–296.

(44) Servais, L.; Montus, M.; Guiner, C. L.; Ben Yaou, R., et al. Non-Ambulant Duchenne Patients Theoretically Treatable by Exon 53 Skipping have Severe Phenotype. *J. Neuromuscul. Dis.* **2015**, *2*, 269–279.

(45) Chamberlain, J. S. Dystrophin levels required for genetic correction of Duchenne muscular dystrophy. *Basic Appl Myol* **1997**, *7*, 255–257.

(46) Athanasopoulos, T.; Graham, I.; Foster, H.; Dickson, G. Recombinant adeno-associated viral (rAAV) vectors as therapeutic tools for Duchenne muscular dystrophy (DMD). *Gene Ther.* **2004**, *11*, S109–S121.

(47) Wang, B.; Li, J.; Xiao, X. Adeno-associated virus vector carrying human minidystrophin genes effectively ameliorates muscular dystrophy in mdx mouse model. *Proc. Natl. Acad. Sci.* **2000**, *97*, 13714–13719.

(48) Watchko, J.; O'Day, T.; Wang, B.; Zhou, L., et al. Adeno-Associated Virus Vector-Mediated Minidystrophin Gene Therapy Improves Dystrophic Muscle Contractile Function in mdx Mice. *Hum. Gene Ther.* **2002**, *13*, 1451–1460.

(49) Shin, J.-H.; Pan, X.; Hakim, C. H.; Yang, H. T., et al. Microdystrophin Ameliorates Muscular Dystrophy in the Canine Model of Duchenne Muscular Dystrophy. *Mol. Ther.* **2013**, *21*, 750–757.

(50) Le Guiner, C.; Servais, L.; Montus, M.; Larcher, T., et al. Long-term microdystrophin gene therapy is effective in a canine model of Duchenne muscular dystrophy. *Nat. Commun.* **2017**, *8*, 16105.

(51) Davies, K. E.; Guiraud, S. Micro-dystrophin Genes Bring Hope of an Effective Therapy for Duchenne Muscular Dystrophy. *Mol. Ther.* **2019**, *27*, 486–488.

(52) Shimizu-Motohashi, Y.; Komaki, H.; Motohashi, N.; Takeda, S., et al. Restoring Dystrophin Expression in Duchenne Muscular Dystrophy: Current Status of Therapeutic Approaches. *J. Pers. Med.* **2019**, *9*, 1.

(53) Li, H. L.; Fujimoto, N.; Sasakawa, N.; Shirai, S., et al. Precise Correction of the Dystrophin Gene in Duchenne Muscular Dystrophy Patient Induced Pluripotent Stem Cells by TALEN and CRISPR-Cas9. *Stem Cell Reports* **2015**, *4*, 143–154.

(54) Young, C. S.; Hicks, M. R.; Ermolova, N. V.; Nakano, H., et al. A Single CRISPR-Cas9 Deletion Strategy that Targets the Majority of DMD Patients Restores Dystrophin Function in hiPSC-Derived Muscle Cells. *Cell Stem Cell* **2016**, *18*, 533–540.

(55) Nelson, C. E.; Wu, Y.; Gemberling, M. P.; Oliver, M. L., et al. Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. *Nat. Med.* **2019**, *25*, 427–432.

(56) Amoasii, L.; Hildyard, J. C. W.; Li, H.; Sanchez-Ortiz, E., et al. Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy. *Science* **2018**, *362*, 86–91.

(57) Guiraud, S.; Roblin, D.; Davies, K. E. The potential of utrophin modulators for the treatment of Duchenne muscular dystrophy. *Expert Opin. Orphan Drugs* **2018**, *6*, 179–192.

(58) Pearce, M.; Blake, D. J.; Tinsley, J. M.; Byth, B. C., et al. The utrophin and dystrophin genes share similarities in genomic structure. *Hum. Mol. Genet.* **1993**, *2*, 1765–1772.

(59) Tinsley, J. M.; Blake, D. J.; Roche, A.; Fairbrother, U., et al. Primary structure of dystrophin-related protein. *Nature* **1992**, *360*, 591–593.

(60) Blake, D. J.; Tinsley, J. M.; Davies, K. E. Utrophin: a structural and functional comparison to dystrophin. *Brain Pathol.* **1996**, *6*, 37–47.

(61) Ishikawa-Sakurai, M.; Yoshida, M.; Imamura, M.; Davies, K. E., et al. ZZ domain is essentially required for the physiological binding of dystrophin and utrophin to beta dystroglycan. *Hum. Mol. Genet.* **2004**, *13*, 693–702.

(62) Matsumura, K.; Shasby, D. M.; Campbell, K. P. Purification of dystrophin-related protein (utrophin) from lung and its identification in pulmonary artery endothelial cells. *FEBS Lett.* **1993**, *326*, 289–93.

(63) Lumeng, C. N.; Phelps, S. F.; Rafael, J. A.; Cox, G. A., et al. Characterization of Dystrophin and Utrophin Diversity in the Mouse. *Hum. Mol. Genet.* **1999**, *8*, 593–599.

(64) Khurana, T. S.; Watkins, S. C.; Chafey, P.; Chelly, J., et al. Immunolocalization and developmental expression of dystrophin related protein in skeletal muscle. *Neuromuscul. Disord.* **1991**, *1*, 185–94.

(65) Tomé, F. M.; Matsumura, K.; Chevallay, M.; Campbell, K. P., et al. Expression of dystrophin-associated glycoproteins during human fetal muscle development: a preliminary immunocytochemical study. *Neuromuscul. Disord.* **1994**, *4*, 343–8.

(66) Weir, A. P.; Burton, E. A.; Harrod, G.; Davies, K. E. A- and B-utrophin Have Different Expression Patterns and Are Differentially Up-regulated in mdx Muscle. *J. Biol. Chem.* **2002**, *277*, 45285–45290.

(67) Pons, F.; Nicholson, L. V.; Robert, A.; Voit, T., et al. Dystrophin and dystrophin-related protein (utrophin) distribution in normal and dystrophin-deficient skeletal muscles. *Neuromuscul. Disord.* **1993**, *3*, 507–14.

(68) Mizuno, Y.; Nonaka, I.; Hirai, S.; Ozawa, E. Reciprocal expression of dystrophin and utrophin in muscles of Duchenne muscular dystrophy patients, female DMD-carriers and control subjects. *J. Neurol. Sci.* **1993**, *119*, 43–52.

(69) Grady, R. M.; Teng, H.; Nichol, M. C.; Cunningham, J. C., et al. Skeletal and cardiac myopathies in mice lacking utrophin and dystrophin: a model for Duchenne muscular dystrophy. *Cell* **1997**, *90*, 729–38.

(70) Kleopa, K. A.; Drousiotou, A.; Mavrikiou, E.; Ormiston, A., et al. Naturally occurring utrophin correlates with disease severity in Duchenne muscular dystrophy. *Hum. Mol. Genet.* **2006**, *15*, 1623–1628.

(71) Tinsley, J.; Deconinck, N.; Fisher, R.; Kahn, D., et al. Expression of full-length utrophin prevents muscular dystrophy in mdx mice. *Nat. Med.* **1998**, *4*, 1441–1444.

(72) Fisher, R.; Tinsley, J. M.; Phelps, S. R.; Squire, S. E., et al. Non-toxic ubiquitous over-expression of utrophin in the mdx mouse. *Neuromuscul. Disord.* **2001**, *11*, 713–21.

(73) Squire, S.; Raymackers, J. M.; Vandebrouck, C.; Potter, A., et al. Prevention of pathology in mdx mice by expression of utrophin: analysis using an inducible transgenic expression system. *Hum. Mol. Genet.* **2002**, *11*, 3333–3344.

(74) Gilbert, R.; Nalbantoglu, J.; Petrof, B. J.; Ebihara, S., et al. Adenovirus-Mediated Utrophin Gene Transfer Mitigates the Dystrophic Phenotype of mdx Mouse Muscles. *Hum. Gene Ther.* **1999**, *10*, 1299–1310.

(75) Cerletti, M.; Negri, T.; Cozzi, F.; Colpo, R., et al. Dystrophic phenotype of canine X-linked muscular dystrophy is mitigated by adenovirus-mediated utrophin gene transfer. *Gene Ther.* **2003**, *10*, 750–757.

(76) Odom, G. L.; Gregorevic, P.; Allen, J. M.; Finn, E., et al. Microutrophin delivery through rAAV6 increases lifespan and improves muscle function in dystrophic dystrophin/utrophin-deficient mice. *Mol. Ther.* **2008**, *16*, 1539–45.

(77) Pisani, C.; Strimpakos, G.; Gabanella, F.; Di Certo, M. G., et al. Utrophin up-regulation by artificial transcription factors induces muscle rescue and impacts the neuromuscular junction in mdx mice. *Biochim. Biophys. Acta - Mol. Basis Dis.* **2018**, *1864*, 1172–1182.

(78) Tinsley, J. M.; Davies, K. E. Utrophin: a potential replacement for dystrophin? *Neuromuscul. Disord. 3*, 537–9.

(79) Burton, E. A.; Tinsley, J. M.; Holzfeind, P. J.; Rodrigues, N. R., et al. A second promoter provides an alternative target for therapeutic up-regulation of utrophin in Duchenne muscular dystrophy. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 14025–30.

(80) Khurana, T. S.; Rosmarin, A. G.; Shang, J.; Krag, T. O. B., et al. Activation of Utrophin Promoter by Heregulin via the ets-related Transcription Factor Complex GA-binding Protein α/β. *Mol. Biol. Cell* **1999**, *10*, ed. by Yamamoto, K. R., 2075–2086.

(81) Perkins, K. J.; Burton, E. A.; Davies, K. E. The role of basal and myogenic factors in the transcriptional activation of utrophin promoter A: implications for therapeutic up-regulation in Duchenne muscular dystrophy. *Nucleic Acids Res.* **2001**, *29*, 4843–50.

(82) Gyrd-Hansen, M.; Krag, T. O.; Rosmarin, A. G.; Khurana, T. S. Sp1 and the ets-related transcription factor complex GABPα/β functionally cooperate to activate the utrophin promoter. *J. Neurol. Sci.* **2002**, *197*, 27–35.

(83) Angus, L. M.; Chakkalakal, J. V.; Méjat, A.; Eibl, J. K., et al. Calcineurin-NFAT signaling, together with GABP and peroxisome PGC-1α, drives utrophin gene expression at the neuromuscular junction. *Am. J. Physiol. Physiol.* **2005**, *289*, C908–C917.

(84) Miura, P.; Chakkalakal, J. V.; Boudreault, L.; Belanger, G., et al. Pharmacological activation of PPARβ/δ stimulates utrophin A expression in skeletal muscle fibers and restores sarcolemmal integrity in mature mdx mice. *Hum. Mol. Genet.* **2009**, *18*, 4640–4649.

(85) Ljubicic, V.; Jasmin, B. J. Metformin increases peroxisome proliferator-activated receptor γ Co-activator-1α and utrophin a expression in dystrophic skeletal muscle. *Muscle Nerve* **2015**, *52*, 139–142.

(86) Hafner, P.; Bonati, U.; Rubino, D.; Gocheva, V., et al. Treatment with l-citrulline and metformin in Duchenne muscular dystrophy: study protocol for a single-centre, randomised, placebo-controlled trial. *Trials* **2016**, *17*, 389.

(87) Ljubicic, V.; Miura, P.; Burt, M.; Boudreault, L., et al. Chronic AMPK activation evokes the slow, oxidative myogenic program and triggers beneficial adaptations in mdx mouse skeletal muscle. *Hum. Mol. Genet.* **2011**, *20*, 3478–3493.

(88) Péladeau, C.; Ahmed, A.; Amirouche, A.; Crawford Parks, T. E., et al. Combinatorial therapeutic activation with heparin and AICAR stimulates additive effects on utrophin A expression in dystrophic muscles. *Hum. Mol. Genet.* **2016**, *25*, 24–43.

(89) Péladeau, C.; Adam, N. J.; Jasmin, B. J. Celecoxib treatment improves muscle function in mdx mice and increases utrophin A expression. *FASEB J.* **2018**, *32*, 5090–5103.

(90) Gordon, B. S.; Delgado Díaz, D. C.; Kostek, M. C. Resveratrol decreases inflammation and increases utrophin gene expression in the mdx mouse model of Duchenne muscular dystrophy. *Clin. Nutr.* **2013**, *32*, 104–111.

(91) Gramolini, A. O.; Angus, L. M.; Schaeffer, L.; Burton, E. A., et al. Induction of utrophin gene expression by heregulin in skeletal muscle cells: Role of the N-box motif and GA binding protein. *PNAS* **1999**, *96*, 3223–3227.

(92) Krag, T. O. B.; Bogdanovich, S.; Jensen, C. J.; Fischer, M. D., et al. Heregulin ameliorates the dystrophic phenotype in mdx mice. *Proc. Natl. Acad. Sci.* **2004**, *101*, 13856–13860.

(93) Fujimori, K.; Itoh, Y.; Yamamoto, K.; Miyagoe-Suzuki, Y., et al. Interleukin 6 Induces Overexpression of the Sarcolemmal Utrophin in Neonatal mdx Skeletal Muscle. *Hum. Gene Ther.* **2002**, *13*, 509–518.

(94) Moorwood, C.; Lozynska, O.; Suri, N.; Napper, A. D., et al. Drug discovery for Duchenne muscular dystrophy via utrophin promoter activation screening. *PLoS One* **2011**, *6*, e26169.

(95) Rodova, M.; Brownback, K.; Werle, M. J. Okadaic acid augments utrophin in myogenic cells. *Neurosci. Lett.* **2004**, *363*, 163–167.

(96) Swinney, D. C.; Anthony, J. How were new medicines discovered? *Nat. Rev. Drug Discov.* **2011**, *10*, 507–519.

(97) Wynne, G. M.; De Moor, O.; Lecci, C.; Van Well, R., et al. Benzoxazole, benzimidazole and benzothiazole derivatives and their use in the treatment of Duchenne Muscular Dystrophy. 2009, WO/2009/021750.

(98) Tinsley, J.; Robinson, N.; Davies, K. E. Safety, tolerability, and pharmacokinetics of SMT C1100, a 2-arylbenzoxazole utrophin modulator, following single- and multiple-dose administration to healthy male adult volunteers. *J. Clin. Pharmacol.* **2015**, *55*, 698–707.

(99) Morgan, J.; Beauchamp, J.; Pagel, C.; Peckham, M., et al. Myogenic Cell Lines Derived from Transgenic Mice Carrying a Thermolabile T Antigen: A Model System for the Derivation of Tissue-Specific and Mutation-Specific Cell Lines. *Dev. Biol.* **1994**, *162*, 486–498.

(100) Chancellor, D. R.; Davies, K. E.; De Moor, O.; Dorgan, C. R., et al. Discovery of 2-Arylbenzoxazoles as Upregulators of Utrophin Production for the Treatment of Duchenne Muscular Dystrophy. *J. Med. Chem.* **2011**, *54*, 3241–3250.

(101) Tinsley, J. M.; Fairclough, R. J.; Storer, R.; Wilkes, F. J., et al. Daily Treatment with SMTC1100, a Novel Small Molecule Utrophin Upregulator, Dramatically Reduces the Dystrophic Symptoms in the mdx Mouse. *PLoS One* **2011**, *6*, ed. by Dent, P., e19189.

(102) Guiraud, S.; Squire, S. E.; Edwards, B.; Chen, H., et al. Second-generation compound for the modulation of utrophin in the therapy of DMD. *Hum. Mol. Genet.* **2015**, *24*, 4212–4224.

(103) Ricotti, V.; Spinty, S.; Roper, H.; Hughes, I., et al. Safety, Tolerability, and Pharmacokinetics of SMT C1100, a 2-Arylbenzoxazole Utrophin Modulator, following Single- and Multiple-Dose Administration to Pediatric Patients with Duchenne Muscular Dystrophy. *PLoS One* **2016**, *11*, ed. by Choonara, I., e0152840.

(104) Forbes, S. C.; Walter, G. A.; Rooney, W. D.; Wang, D.-J., et al. Skeletal Muscles of Ambulant Children with Duchenne Muscular Dystrophy: Validation of Multicenter Study of Evaluation with MR Imaging and MR Spectroscopy. *Radiology* **2013**, *269*, 198–207.

(105) Guiraud, S.; Edwards, B.; Squire, S. E.; Moir, L., et al. Embryonic myosin is a regeneration marker to monitor utrophin-based therapies for DMD. *Hum. Mol. Genet.* **2018**, *28*, 307–319.

(106) Summit Presents New 24-Week Analyses from PhaseOut DMD at the 2018 AAN Conference. https://otp.investis.com/clients/uk/summit_corporation_plc/rns/regulatory-story.aspx?cid=1575&newsid=1001571. Accessed 05/07/2019.

(107) Moffat, J. G.; Vincent, F.; Lee, J. A.; Eder, J., et al. Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nat. Rev. Drug Discov.* **2017**, *16*, 531–543.

(108) Trosset, J.-Y.; Cavé, C. In *In Silico Drug–Target Profiling*; Humana Press, New York, NY: 2019, pp 89–103.

(109) Schenone, M.; Dančík, V.; Wagner, B. K.; Clemons, P. A. Target identification and mechanism of action in chemical biology and drug discovery. *Nat. Chem. Biol.* **2013**, *9*, 232–240.

(110) Wagner, B. K.; Schreiber, S. L. The Power of Sophisticated Phenotypic Screening and Modern Mechanism-of-Action Methods. *Cell Chem. Biol.* **2016**, *23*, 3–9.

(111) Schirle, M.; Jenkins, J. L. Identifying compound efficacy targets in phenotypic drug discovery. *Drug Discov. Today* **2016**, *21*, 82–89.

(112) Wright, M. H.; Sieber, S. A. Chemical proteomics approaches for identifying the cellular targets of natural products. *Nat. Prod. Rep.* **2016**, *33*, 681–708.

(113) Tulloch, L. B.; Menzies, S. K.; Coron, R. P.; Roberts, M. D., et al. Direct and indirect approaches to identify drug modes of action. *IUBMB Life* **2018**, *70*, 9–22.

(114) Comess, K. M.; McLoughlin, S. M.; Oyer, J. A.; Richardson, P. L., et al. Emerging Approaches for the Identification of Protein Targets of Small Molecules - A Practitioners' Perspective. *J. Med. Chem.* **2018**, *61*, 8504–8535.

(115) Kubota, K.; Funabashi, M.; Ogura, Y. Target deconvolution from phenotype-based drug discovery by using chemical proteomics approaches. *Biochim. Biophys. Acta - Proteins Proteomics* **2019**, *1867*, 22–27.

(116) Keiser, M. J.; Setola, V.; Irwin, J. J.; Laggner, C., et al. Predicting new molecular targets for known drugs. *Nature* **2009**, *462*, 175–181.

(117) Wu, Z.; Li, W.; Liu, G.; Tang, Y. Network-Based Methods for Prediction of Drug-Target Interactions. *Front. Pharmacol.* **2018**, *9*, 1134.

(118) Chen, R.; Liu, X.; Jin, S.; Lin, J., et al. Machine Learning for Drug-Target Interaction Prediction. *Molecules* **2018**, *23*, 2208.

(119) Subramanian, A.; Narayan, R.; Corsello, S. M.; Peck, D. D., et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **2017**, *171*, 1437–1452.e17.

(120) Musa, A.; Ghoraie, L. S.; Zhang, S.-D.; Glazko, G., et al. A review of connectivity map and computational approaches in pharmacogenomics. *Brief. Bioinform.* **2018**, *19*, 506–523.

(121) Giaever, G.; Shoemaker, D. D.; Jones, T. W.; Liang, H., et al. Genomic profiling of drug sensitivities via induced haploinsufficiency. *Nat. Genet.* **1999**, *21*, 278–283.

(122) Ho, C. H.; Piotrowski, J.; Dixon, S. J.; Baryshnikova, A., et al. Combining functional genomics and chemical biology to identify targets of bioactive compounds. *Curr. Opin. Chem. Biol.* **2011**, *15*, 66–78.

(123) Parsons, A. B.; Lopez, A.; Givoni, I. E.; Williams, D. E., et al. Exploring the Mode-of-Action of Bioactive Compounds by Chemical-Genetic Profiling in Yeast. *Cell* **2006**, *126*, 611–625.

(124) Deans, R. M.; Morgens, D. W.; Ökesli, A.; Pillay, S., et al. Parallel shRNA and CRISPR-Cas9 screens enable antiviral drug target identification. *Nat. Chem. Biol.* **2016**, *12*, 361–366.

(125) Jost, M.; Chen, Y.; Gilbert, L. A.; Horlbeck, M. A., et al. Combined CRISPRi/a-Based Chemical Genetic Screens Reveal that Rigosertib Is a Microtubule-Destabilizing Agent. *Mol. Cell* **2017**, *68*, 210–223.e6.

(126) Jost, M.; Weissman, J. S. CRISPR Approaches to Small Molecule Target Identification. *ACS Chem. Biol.* **2018**, *13*, 366–375.

(127) Smith, S. H.; Jayawickreme, C.; Rickard, D. J.; Nicodeme, E., et al. Tapinarof Is a Natural AhR Agonist that Resolves Skin Inflammation in Mice and Humans. *J. Invest. Dermatol.* **2017**, *137*, 2110–2119.

(128) Zhang, Q.; Major, M. B.; Takanashi, S.; Camp, N. D., et al. Small-molecule synergist of the Wnt/beta-catenin signaling pathway. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 7444–8.

(129) Falsey, R. R.; Marron, M. T.; Gunaherath, G. M. K. B.; Shirahatti, N., et al. Actin microfilament aggregation induced by withaferin A is mediated by annexin II. *Nat. Chem. Biol.* **2006**, *2*, 33–38.

(130) Usui, T.; Watanabe, H.; Nakayama, H.; Tada, Y., et al. The Anticancer Natural Product Pironetin Selectively Targets Lys352 of α-Tubulin. *Chem. Biol.* **2004**, *11*, 799–806.

(131) Dückert, H.; Pries, V.; Khedkar, V.; Menninger, S., et al. Natural product–inspired cascade synthesis yields modulators of centrosome integrity. *Nat. Chem. Biol.* **2012**, *8*, 179–184.

(132) Hirota, T.; Lee, J. W.; St John, P. C.; Sawa, M., et al. Identification of small molecule activators of cryptochrome. *Science* **2012**, *337*, 1094–7.

(133) Barglow, K. T.; Cravatt, B. F. Activity-based protein profiling for the functional annotation of enzymes. *Nat. Methods* **2007**, *4*, 822–827.

(134) Ahn, K.; Johnson, D. S.; Mileni, M.; Beidler, D., et al. Discovery and Characterization of a Highly Selective FAAH Inhibitor that Reduces Inflammatory Pain. *Chem. Biol.* **2009**, *16*, 411–420.

(135) Medvedev, A.; Kopylov, A.; Buneeva, O.; Zgoda, V., et al. Affinity-based proteomic profiling: Problems and achievements. *Proteomics* **2012**, *12*, 621–637.

(136) Shi, H.; Zhang, C.-J.; Chen, G. Y. J.; Yao, S. Q. Cell-Based Proteome Profiling of Potential Dasatinib Targets by Use of Affinity-Based Probes. *J. Am. Chem. Soc.* **2012**, *134*, 3001–3014.

(137) Won, S. J.; Eschweiler, J. D.; Majmudar, J. D.; Chong, F. S., et al. Affinity-Based Selectivity Profiling of an In-Class Selective Competitive Inhibitor of Acyl Protein Thioesterase 2. *ACS Med. Chem. Lett.* **2017**, *8*, 215–220.

(138) Martell, J.; Weerapana, E. Applications of Copper-Catalyzed Click Chemistry in Activity-Based Protein Profiling. *Molecules* **2014**, *19*, 1378–1393.

(139) Lang, K.; Chin, J. W. Bioorthogonal Reactions for Labeling Proteins. *ACS Chem. Biol.* **2014**, *9*, 16–20.

(140) Row, R. D.; Prescher, J. A. Constructing New Bioorthogonal Reagents and Reactions. *Acc. Chem. Res.* **2018**, *51*, 1073–1081.

(141) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem. Int. Ed. Engl.* **2002**, *41*, 2596–9.

(142) Tornøe, C. W.; Christensen, C.; Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(i)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **2002**, *67*, 3057–64.

(143) Presolski, S. I.; Hong, V.; Cho, S.-H.; Finn, M. Tailored Ligand Acceleration of the Cu-Catalyzed Azide−Alkyne Cycloaddition Reaction: Practical and Mechanistic Implications. *J. Am. Chem. Soc.* **2010**, *132*, 14570–14576.

(144) Soriano Del Amo, D.; Wang, W.; Jiang, H.; Besanceney, C., et al. Biocompatible copper(I) catalysts for in vivo imaging of glycans. *J. Am. Chem. Soc.* **2010**, *132*, 16893–9.

(145) Nicholas J. Agard; Prescher, J. A.; Bertozzi, C. R. A Strain-Promoted [3 + 2] Azide−Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. **2004**.

(146) Rutkowska, A.; Thomson, D. W.; Vappiani, J.; Werner, T., et al. A Modular Probe Strategy for Drug Localization, Target Identification and Target Occupancy Measurement on Single Cell Level. *ACS Chem. Biol.* **2016**, *11*, 2541–2550.

(147) Li, Z.; Wang, D.; Li, L.; Pan, S., et al. "Minimalist" Cyclopropene-Containing Photo-Cross-Linkers Suitable for Live-Cell Imaging and Affinity-Based Protein Labeling. **2014**.

(148) Smith, E.; Collins, I. Photoaffinity labeling in target- and binding-site identification. *Future Med. Chem.* **2015**, *7*, 159–183.

(149) Ruoho, A. E.; Kiefer, H.; Roeder, P. E.; Singer, S. J. The mechanism of photoaffinity labeling. *Proc. Natl. Acad. Sci. U. S. A.* **1973**, *70*, 2567–71.

(150) Singh, A.; Thornton, E. R.; Westheimer, F. H. The photolysis of diazoacetylchymotrypsin. *J. Biol. Chem.* **1962**, *237*, 3006–8.

(151) Crump, C. J.; Fish, B. A.; Castro, S. V.; Chau, D.-M., et al. Piperidine Acetic Acid Based γ-Secretase Modulators Directly Bind to Presenilin-1. *ACS Chem. Neurosci.* **2011**, *2*, 705–710.

(152) Smith, R. A. G.; Knowles, J. R. Aryldiazirines. Potential reagents for photolabeling of biological receptor sites. *J. Am. Chem. Soc.* **1973**, *95*, 5072–5073.

(153) Korneev, S. M. Valence Isomerization between Diazo Compounds and Diazirines. *European J. Org. Chem.* **2011**, *2011*, 6153–6175.

(154) Brunner, J.; Senn, H.; Richards, F. M. 3-Trifluoromethyl-3-phenyldiazirine. A new carbene generating group for photolabeling reagents. *J. Biol. Chem.* **1980**, *255*, 3313–3318.

(155) Das, J. Aliphatic Diazirines as Photoaffinity Probes for Proteins: Recent Developments. *Chem. Rev.* **2011**, *111*, 4405–4417.

(156) MacKinnon, A. L.; Garrison, J. L.; Hegde, R. S.; Taunton, J. Photo-Leucine Incorporation Reveals the Target of a Cyclodepsipeptide Inhibitor of Cotranslational Translocation. **2007**.

(157) Ranjitkar, P.; Perera, B. G. K.; Swaney, D. L.; Hari, S. B., et al. Affinity-Based Probes Based on Type II Kinase Inhibitors. *J. Am. Chem. Soc.* **2012**, *134*, 19017–19025.

(158) Lapinsky, D. J. Tandem photoaffinity labeling–bioorthogonal conjugation in medicinal chemistry. *Bioorg. Med. Chem.* **2012**, *20*, 6237–6247.

(159) Gerritz, S. W.; Zhai, W.; Shi, S.; Zhu, S., et al. Acyl guanidine inhibitors of β-secretase (BACE-1): Optimization of a micromolar hit to a nanomolar lead via iterative solid- and solution-phase library synthesis. *J. Med. Chem.* **2012**, *55*, 9208–9223.

(160) Tsukiji, S.; Miyagawa, M.; Takaoka, Y.; Tamura, T., et al. Ligand-directed tosyl chemistry for protein labeling in vivo. *Nat. Chem. Biol.* **2009**, *5*, 341–343.

(161) Fujishima, S.-h.; Yasui, R.; Miki, T.; Ojida, A., et al. Ligand-Directed Acyl Imidazole Chemistry for Labeling of Membrane-Bound Proteins on Live Cells. *J. Am. Chem. Soc.* **2012**, *134*, 3961–3964.

(162) Takaoka, Y.; Nishikawa, Y.; Hashimoto, Y.; Sasaki, K., et al. Ligand-directed dibromophenyl benzoate chemistry for rapid and selective acylation of intracellular natural proteins. *Chem. Sci.* **2015**, *6*, 3217–3224.

(163) Takaoka, Y.; Nukadzuka, Y.; Ueda, M. Reactive group-embedded affinity labeling reagent for efficient intracellular protein labeling. *Bioorg. Med. Chem.* **2017**, *25*, 2888–2894.

(164) Wang, H.; Koshi, Y.; Minato, D.; Nonaka, H., et al. Chemical Cell-Surface Receptor Engineering Using Affinity-Guided, Multivalent Organocatalysts. *J. Am. Chem. Soc.* **2011**, *133*, 12220–12228.

(165) Takakusagi, Y.; Takakusagi, K.; Sugawara, F.; Sakaguchi, K. Use of phage display technology for the determination of the targets for small-molecule therapeutics. *Expert Opin. Drug Discov.* **2010**, *5*, 361–389.

(166) Licitra, E. J.; Liu, J. O. A three-hybrid system for detecting small ligand-protein receptor interactions. *Proc. Natl. Acad. Sci.* **1996**, *93*, 12817–12821.

(167) Cromm, P. M.; Crews, C. M. Targeted Protein Degradation: from Chemical Biology to Drug Discovery. *Cell Chem. Biol.* **2017**, *24*, 1181–1190.

(168) Erb, M. A.; Scott, T. G.; Li, B. E.; Xie, H., et al. Transcription control by the ENL YEATS domain in acute leukaemia. *Nature* **2017**, *543*, 270–274.

(169) Huber, K. V. M.; Olek, K. M.; Muller, A. C.; Tan, C. S. H., et al. Proteome-wide drug and metabolite interaction mapping by thermal-stability profiling. *Nat. Meth.* **2015**, *12*, 1055–1057.

(170) Strickland, E. C.; Geer, M. A.; Tran, D. T.; Adhikari, J., et al. Thermodynamic analysis of protein-ligand binding interactions in complex biological mixtures using the stability of proteins from rates of oxidation. *Nat. Protoc.* **2013**, *8*, 148–161.

(171) Lomenick, B.; Hao, R.; Jonai, N.; Chin, R. M., et al. Target identification using drug affinity responsive target stability (DARTS). *PNAS* **2009**, *106*, 21984–21989.

(172) Park, H.; Ha, J.; Park, S. B. Label-free target identification in drug discovery via phenotypic screening. *Curr. Opin. Chem. Biol.* **2019**, *50*, 66–72.

(173) Bergamini, G.; Bell, K.; Shimamura, S.; Werner, T., et al. A selective inhibitor reveals PI3Kγ dependence of TH17 cell differentiation. *Nat. Chem. Biol.* **2012**, *8*, 576–582.

(174) Lovering, F. Escape from Flatland 2: complexity and promiscuity. *Medchemcomm* **2013**, *4*, 515.

(175) Park, H.; Koo, J. Y.; Srikanth, Y. V. V.; Oh, S., et al. Nonspecific protein labeling of photoaffinity linkers correlates with their molecular shapes in living cells. *Chem. Commun.* **2016**, *52*, 5828–5831.

(176) Green, N. M. Avidin. 3. The nature of the biotin-binding site. *Biochem. J.* **1963**, *89*, 599–609.

(177) Wang, F.; Yang, H.; Fu, H.; Pei, Z. Efficient copper-catalyzed Michael addition of acrylic derivatives with primary alcohols in the presence of base. *Chem. Commun.* **2013**, *49*, 517–519.

(178) Reynolds, J. K. Synthesis of photoaffinity probes to elucidate the method of action of a utrophin modulator drug candidate in the treatment of Duchenne muscular dystrophy. MChem thesis, 2016, University of Oxford, UK.

(179) Hill, J. R.; Robertson, A. A. B. Fishing for Drug Targets: A Focus on Diazirine Photoaffinity Probe Synthesis. *J. Med. Chem.* **2018**, *61*, 6945–6963.

(180) Ge, S.-S.; Chen, B.; Wu, Y.-Y.; Long, Q.-S., et al. Current advances of carbene-mediated photoaffinity labeling in medicinal chemistry. *RSC Adv.* **2018**, *8*, 29428–29454.

(181) Surya Prakash, G. K.; Yudin, A. K. Perfluoroalkylation with Organosilicon Reagents. **1997**, *97*, 757–786.

(182) Basu, U.; Gyrd-Hansen, M.; Baby, S. M.; Lozynska, O., et al. Heregulin-induced epigenetic regulation of the utrophin-A promoter. *FEBS Lett.* **2007**, *581*, 4153–4158.

(183) Rudzinski, D. M.; Kelly, C. B.; Leadbeater, N. E. A Weinreb amide approach to the synthesis of trifluoromethylketones. *Chem. Commun.* **2012**, *48*, 9610–9612.

(184) Thorne, N.; Inglese, J.; Auld, D. S. Illuminating Insights into Firefly Luciferase and Other Bioluminescent Reporters Used in Chemical Biology. *Chem. Biol.* **2010**, *17*, 646–657.

(185) Fan, F.; Wood, K. V. Bioluminescent Assays for High-Throughput Screening. *Assay Drug Dev. Technol.* **2007**, *5*, 127–136.

(186) Thompson, J. F.; Hayes, L. S.; Lloyd, D. B. Modulation of firefly luciferase stability and impact on studies of gene regulation. *Gene* **1991**, *103*, 171–177.

(187) Thorne, N.; Shen, M.; Lea, W. A.; Simeonov, A., et al. Firefly Luciferase in Chemical Biology: A Compendium of Inhibitors, Mechanistic Evaluation of Chemotypes, and Suggested Use As a Reporter. *Chem. Biol.* **2012**, *19*, 1060–1072.

(188) Dranchak, P.; MacArthur, R.; Guha, R.; Zuercher, W. J., et al. Profile of the GSK Published Protein Kinase Inhibitor Set Across ATP-Dependent and -Independent Luciferases: Implications for Reporter-Gene Assays. *PLoS One* **2013**, *8*, ed. by Jung, M., e57888.

(189) Lyssiotis, C. A.; Foreman, R. K.; Staerk, J.; Garcia, M., et al. Reprogramming of murine fibroblasts to induced pluripotent stem cells with chemical complementation of Klf4. *Proc. Natl. Acad. Sci.* **2009**, *106*, 8912–8917.

(190) Younis, I.; Berg, M.; Kaida, D.; Dittmar, K., et al. Rapid-response splicing reporter screens identify differential regulators of constitutive and alternative splicing. *Mol. Cell. Biol.* **2010**, *30*, 1718–28.

(191) Rocha, S.; Campbell, K. J.; Roche, K. C.; Perkins, N. D. The p53-inhibitor Pifithrin-α inhibits Firefly Luciferase activity in vivo and in vitro. *BMC Mol. Biol.* **2003**, *4*, 9.

(192) Bakhtiarova, A.; Taslimi, P.; Elliman, S. J.; Kosinski, P. A., et al. Resveratrol inhibits firefly luciferase. *Biochem. Biophys. Res. Commun.* **2006**, *351*, 481–484.

(193) Wang, T. T. β-Naphthoflavone, an Inducer of Xenobiotic Metabolizing Enzymes, Inhibits Firefly Luciferase Activity. *Anal. Biochem.* **2002**, *304*, 122–126.

(194) Auld, D. S.; Southall, N. T.; Jadhav, A.; Johnson, R. L., et al. Characterization of chemical libraries for luciferase inhibitory activity. *J. Med. Chem.* **2008**, *51*, 2372–2386.

(195) Auld, D. S.; Thorne, N.; Nguyen, D.-T.; Inglese, J. A Specific Mechanism for Nonspecific Activation in Reporter-Gene Assays. *ACS Chem. Biol.* **2008**, *3*, 463–470.

(196) Auld, D. S.; Inglese, J. Interferences with Luciferase Reporter Enzymes Assay Guidance Manual. *Assay Guid. Man. [Internet]* **2016**, *1*, 1–14.

(197) McElroy, W.; Green, A. Function of adenosine triphosphate in the activation of luciferin. *Arch. Biochem. Biophys.* **1956**, *64*, 257–271.

(198) Sundlov, J. A.; Fontaine, D. M.; Southworth, T. L.; Branchini, B. R., et al. Crystal structure of firefly luciferase in a second catalytic conformation supports a domain alternation mechanism. *Biochemistry* **2012**, *51*, 6493–5.

(199) McElroy, W. D.; Seliger, H. H.; White, E. H. Mechanism of bioluminescence, chemiluminescence and enzyme function in the oxidation of firefly luciferin. *Photochem. Photobiol.* **1969**, *10*, 153–170.

(200) Fontes, R.; Dukhovich, A.; Sillero, A.; Sillero, M. A. Synthesis of Dehydroluciferin by Firefly Luciferase: Effect of Dehydroluciferin, Coenzyme A and Nucleoside Triphosphates on the Luminescent Reaction. *Biochem. Biophys. Res. Commun.* **1997**, *237*, 445–450.

(201) Poutiainen, P. K.; Palvimo, J. J.; Hinkkanen, A. E.; Valkonen, A., et al. Discovery of 5-Benzyl-3-phenyl-4,5-dihydroisoxazoles and 5-Benzyl-3-phenyl-1,4,2-dioxazoles as Potent Firefly Luciferase Inhibitors. *J. Med. Chem.* **2013**, *56*, 1064–1073.

(202) Leitão, J. M.; Esteves da Silva, J. C. Firefly luciferase inhibition. *J. Photochem. Photobiol. B Biol.* **2010**, *101*, 1–8.

(203) Auld, D. S.; Thorne, N.; Maguire, W. F.; Inglese, J. Mechanism of PTC124 activity in cell-based luciferase assays of nonsense codon suppression. *PNAS* **2009**, *106*, 3585–3590.

(204) Ueda, I.; Kamaya, H. Kinetic and thermodynamic aspects of the mechanism of general anesthesia in a model system of firefly luminescence in vitro. *Anesthesiology* **1973**, *38*, 425–36.

(205) Ignowski, J. M.; Schaffer, D. V. Kinetic analysis and modeling of firefly luciferase as a quantitative reporter gene in live mammalian cells. *Biotechnol. Bioeng.* **2004**, *86*, 827–834.

(206) Feeney, K. A.; Putker, M.; Brancaccio, M.; O'Neill, J. S., et al. In-depth Characterization of Firefly Luciferase as a Reporter of Circadian Gene Expression in Mammalian Cells. *J. Biol. Rhythm.* **2016**, *31*, 540–550.

(207) Nakamura, M.; Maki, S.; Amano, Y.; Ohkita, Y., et al. Firefly luciferase exhibits bimodal action depending on the luciferin chirality. *Biochem. Biophys. Res. Commun.* **2005**, *331*, 471–5.

(208) Denburg, J. L.; Lee, R. T.; McElroy, W. Substrate-binding properties of firefly luciferase: I. Luciferin-binding site. *Arch. Biochem. Biophys.* **1969**, *134*, 381–394.

(209) Ribeiro, C.; Esteves da Silva, J. C. G. Kinetics of inhibition of firefly luciferase by oxyluciferin and dehydroluciferyl-adenylate. *Photochem. Photobiol. Sci.* **2008**, *7*, 1085.

(210) Takehara, K.; Kamaya, H.; Ueda, I. Inhibition of firefly luciferase by alkane analogues. *Biochim. Biophys. Acta - Gen. Subj.* **2005**, *1721*, 124–129.

(211) Bai, H.; Chen, W.; Wu, W.; Ma, Z., et al. Discovery of a series of 2-phenylnaphthalenes as firefly luciferase inhibitors. *RSC Adv.* **2015**, *5*, 63450–63457.

(212) Franks, N.; Jenkins, A.; Conti, E.; Lieb, W., et al. Structural Basis for the Inhibition of Firefly Luciferase by a General Anesthetic. *Biophys. J.* **1998**, *75*, 2205–2211.

(213) Zhang, H.; Su, J.; Lin, Y.; Bai, H., et al. Inhibiting Firefly Bioluminescence by Chalcones. *Anal. Chem.* **2017**, *89*, 6099–6105.

(214) DeLuca, M. Hydrophobic nature of the active site of firefly luciferase. *Biochemistry* **1969**, *8*, 160–166.

(215) Chen, F.; Sun, H.; Liu, H.; Li, D., et al. Prediction of luciferase inhibitors by the high-performance MIEC-GBDT approach based on interaction energetic patterns. *Phys. Chem. Chem. Phys.* **2017**, *19*, 10163–10176.

(216) Ghosh, D.; Koch, U.; Hadian, K.; Sattler, M., et al. Luciferase Advisor: High-Accuracy Model To Flag False Positive Hits in Luciferase HTS Assays. *J. Chem. Inf. Model.* **2018**, *58*, 933–942.

(217) Broom, A. D. Rational design of enzyme inhibitors: multisubstrate analog inhibitors. *J. Med. Chem.* **1989**, *32*, 2–7.

(218) Le Calvez, P. B.; Scott, C. J.; Migaud, M. Multisubstrate adduct inhibitors: Drug design and biological tools. *J. Enzyme Inhib. Med. Chem.* **2009**, *24*, 1291–1318.

(219) Auld, D. S.; Lovell, S.; Thorne, N.; Lea, W. A., et al. Molecular basis for the high-affinity binding and stabilization of firefly luciferase by PTC124. *Proc. Natl. Acad. Sci.* **2010**, *107*, 4878–4883.

(220) Welch, E. M.; Barton, E. R.; Zhuo, J.; Tomizawa, Y., et al. PTC124 targets genetic disorders caused by nonsense mutations. *Nature* **2007**, *447*, 87–91.

(221) Braeuning, A.; Vetter, S. The nuclear factor κB inhibitor (E)-2-fluoro-4′-methoxystilbene inhibits firefly luciferase. *Biosci. Rep.* **2012**, *32*, 531–537.

(222) Braeuning, A. Firefly luciferase inhibition: a widely neglected problem. *Arch. Toxicol.* **2015**, *89*, 141–142.

(223) Branchini, B. R.; Murtiashaw, M. H.; Carmody, J. N.; Mygatt, E. E., et al. Synthesis of an N-acyl sulfamate analog of luciferyl-AMP: A stable and potent inhibitor of firefly luciferase. *Bioorganic Med. Chem. Lett.* **2005**, *15*, 3860–3864.

(224) Ueda, I.; Kamaya, H. Effects of Diethyl Ether and Halothane on Firefly Luciferin Bioluminescence. *Anesthesiology* **1965**, *26*, 603–606.

(225) Keyaerts, M.; Remory, I.; Caveliers, V.; Breckpot, K., et al. Inhibition of Firefly Luciferase by General Anesthetics: Effect on In Vitro and In Vivo Bioluminescence Imaging. *PLoS One* **2012**, *7*, ed. by Hamblin, M., e30061.

(226) Ueda, I.; Suzuki, A.; Kamaya, H. Do anesthetics act by competitive binding to specific receptors? Phase transition of firefly luciferase. *Toxicol. Lett.* **1998**, *100-101*, 405–411.

(227) Ryan, A. J.; Gray, N. M.; Lowe, P. N.; Chung, C.-w. Effect of detergent on "promiscuous" inhibitors. *J Med Chem* **2003**, *46*, 3448–3451.

(228) Feng, B. Y.; Shoichet, B. K. A detergent-based assay for the detection of promiscuous inhibitors. *Nat. Protoc.* **2006**, *1*, 550–553.

(229) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–61.

(230) Lembert, N.; Idahl, L. A. Regulatory effects of ATP and luciferin on firefly luciferase activity. *Biochem. J.* **1995**, *305*, 929–33.

(231) Liu, Y.; Fang, J.; Cai, H.; Xiao, F., et al. Identification and synthesis of substituted pyrrolo[2,3-d]pyrimidines as novel firefly luciferase inhibitors. *Bioorg. Med. Chem.* **2012**, *20*, 5473–5482.

(232) Heitman, L. H.; van Veldhoven, J. P. D.; Zweemer, A. M.; Ye, K., et al. False Positives in a Reporter Gene Assay: Identification and Synthesis of Substituted N-Pyridin-2-ylbenzamides as Competitive Inhibitors of Firefly Luciferase. *J. Med. Chem.* **2008**, *51*, 4724–4729.

(233) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B., et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 213–221.

(234) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D., et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.

(235) Conti, E.; Franks, N. P.; Brick, P. Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure* **1996**, *4*, 287–98.

(236) Seifert, T.; Malo, M.; Lengqvist, J.; Sihlbom, C., et al. Identification of the Binding Site of Chroman-4-one-Based Sirtuin 2-Selective Inhibitors using Photoaffinity Labeling in Combination with Tandem Mass Spectrometry. *J. Med. Chem.* **2016**, *59*, 10794–10799.

(237) Muranaka, H.; Momose, T.; Handa, C.; Ozawa, T. Photoaffinity Labeling of the Human A_2A Adenosine Receptor and Cross-link Position Analysis by Mass Spectrometry. *ACS Med. Chem. Lett.* **2017**, *8*, 660–665.

(238) Guthrie, J. P. Carbonyl Addition Reactions: Factors Affecting the Hydrate–Hemiacetal and Hemiacetal–Acetal Equilibrium Constants. *Can. J. Chem.* **1975**, *53*, 898–906.

(239) Nassal, M. 4'-(1-Azi-2,2,2-trifluoroethyl)phenylalanine, a photolabile carbene-generating analog of phenylalanine. *J. Am. Chem. Soc.* **1984**, *106*, 7540–7545.

(240) Branchini, B. R.; Southworth, T. L.; Murtiashaw, M. H.; Wilkinson, S. R., et al. Mutagenesis Evidence that the Partial Reactions of Firefly Bioluminescence Are Catalyzed by Different Conformations of the Luciferase C-Terminal Domain. *Biochemistry* **2005**, *44*, 1385–1393.

(241) Branchini, B. R.; Behney, C. E.; Southworth, T. L.; Fontaine, D. M., et al. Experimental Support for a Single Electron-Transfer Oxidation Mechanism in Firefly Bioluminescence. *J. Am. Chem. Soc.* **2015**, *137*, 7592–7595.

(242) Berraud-Pache, R.; Lindh, R.; Navizet, I. QM/MM Study of the Formation of the Dioxetanone Ring in Fireflies through a Superoxide Ion. *J. Phys. Chem. B* **2018**, *122*, 5173–5182.

(243) Helliwell, T. R.; Man, N. T.; Morris, G. E.; Davies, K. E. The dystrophin-related protein, utrophin, is expressed on the sarcolemma of regenerating human skeletal muscle fibres in dystrophies and inflammatory myopathies. *Neuromuscul. Disord.* **1992**, *2*, 177–84.

(244) Arechavala-Gomeza, V.; Kinali, M.; Feng, L.; Brown, S. C., et al. Immunohistological intensity measurements as a tool to assess sarcolemma-associated protein expression. *Neuropathol. Appl. Neurobiol.* **2009**, *36*, 265–274.

(245) Janghra, N.; Morgan, J. E.; Sewry, C. A.; Wilson, F. X., et al. Correlation of Utrophin Levels with the Dystrophin Protein Complex and Muscle Fibre Regeneration in Duchenne and Becker Muscular Dystrophy Muscle Biopsies. *PLoS One* **2016**, *11*, ed. by Cohn, R., e0150818.

(246) Rando, T. A.; Blau, H. M. Primary mouse myoblast purification, characterization, and transplantation for cell-mediated gene therapy. *J. Cell Biol.* **1994**, *125*, 1275–1287.

(247) Choi, I. Y.; Lim, H.; Estrellas, K.; Mula, J., et al. Concordant but Varied Phenotypes among Duchenne Muscular Dystrophy Patient-Specific Myoblasts Derived using a Human iPSC-Based Model. *Cell Rep.* **2016**, *15*, 2301–2312.

(248) McRae, N.; Forgan, L.; McNeill, B.; Addinsall, A., et al. Glucocorticoids Improve Myogenic Differentiation In Vitro by Suppressing the Synthesis of Versican, a Transitional Matrix Protein Overexpressed in Dystrophic Skeletal Muscles. *Int. J. Mol. Sci.* **2017**, *18*.

(249) Cox, J.; Mann, M. Quantitative, High-Resolution Proteomics for Data-Driven Systems Biology. *Annu. Rev. Biochem.* **2011**, *80*, 273–299.

(250) Smith, J. C.; Macklin, A. M. In *Adv. LC-MS Appl. proteomics*, Pennington, S. R., Ed.; Future Science group: London, 2016; Chapter 5.

(251) McClure, R. A.; Williams, J. D. Impact of Mass Spectrometry-Based Technologies and Strategies on Chemoproteomics as a Tool for Drug Discovery. *ACS Med. Chem. Lett.* **2018**, *9*, 785–791.

(252) Wong, J. W. H.; Cagney, G. In *Methods Mol. Biol.* Humana Press: 2010; Vol. 604, pp 273–283.

(253) Neilson, K. A.; Ali, N. A.; Muralidharan, S.; Mirzaei, M., et al. Less label, more free: Approaches in label-free quantitative mass spectrometry. *Proteomics* **2011**, *11*, 535–553.

(254) Winkler, C.; Denker, K.; Wortelkamp, S.; Sickmann, A. Silver- and Coomassie-staining protocols: Detection limits and compatibility with ESI MS. *Electrophoresis* **2007**, *28*, 2095–2099.

(255) Cunningham, C. W.; Mukhopadhyay, A.; Lushington, G. H.; Blagg, B. S. J., et al. Uptake, distribution and diffusivity of reactive fluorophores in cells: implications toward target identification. *Mol. Pharm.* **2010**, *7*, 1301–10.

(256) DIMT1 Gene ID: 27292. Gene [Internet], 2004, Bethesda (MD): National Library of Medicine (US), NCBI. Available from: https://www.ncbi.nlm.nih.gov/gene/27292. Accessed: 27/08/19.

(257) Fleming, S. A. Chemical reagents in photoaffinity labeling. *Tetrahedron* **1995**, *51*, 12479–12520.

(258) Wurz, R. P.; Dellamaggiore, K.; Dou, H.; Javier, N., et al. A "Click Chemistry Platform" for the Rapid Synthesis of Bispecific Molecules for Inducing Protein Degradation. *J. Med. Chem.* **2018**, *61*, 453–461.

(259) Heal, W. P.; Wright, M. H.; Thinon, E.; Tate, E. W. Multifunctional protein labeling via enzymatic N-terminal tagging and elaboration by click chemistry. *Nat. Protoc.* **2012**, *7*, 105–117.

(260) Petryszak, R.; Keays, M.; Tang, Y. A.; Fonseca, N. A., et al. Expression Atlas update—an integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Res.* **2016**, *44*, D746–D752.

(261) Yu, H.; Waddell, J. N.; Kuang, S.; Bidwell, C. A. Park7 Expression Influences Myotube Size and Myosin Expression in Muscle. *PLoS One* **2014**, *9*, 1–10.

(262) Axelsson, H.; Almqvist, H.; Otrocka, M.; Vallin, M., et al. In Situ Target Engagement Studies in Adherent Cells. *ACS Chem. Biol.* **2018**, *13*, 942–950.

(263) Dziekan, J. M.; Yu, H.; Chen, D.; Dai, L., et al. Identifying purine nucleoside phosphorylase as the target of quinine using cellular thermal shift assay. *Sci. Transl. Med.* **2019**, *11*, eaau3174.

(264) Shaw, J.; Dale, I.; Hemsley, P.; Leach, L., et al. Positioning High-Throughput CETSA in Early Drug Discovery through Screening against B-Raf and PARP1. *SLAS Discov. Adv. Life Sci. R&D* **2019**, *24*, 121–132.

(265) Schürmann, M.; Janning, P.; Ziegler, S.; Waldmann, H. Small-Molecule Target Engagement in Cells. *Cell Chem. Biol.* **2016**, *23*, 435–441.

(266) Vogel, C. F. A.; Haarmann-Stemmann, T. The aryl hydrocarbon receptor repressor - More than a simple feedback inhibitor of AhR signaling: Clues for its role in inflammation and cancer. *Curr. Opin. Toxicol.* **2017**, *2*, 109–119.

(267) Neavin, D.; Liu, D.; Ray, B.; Weinshilboum, R. The Role of the Aryl Hydrocarbon Receptor (AHR) in Immune and Inflammatory Diseases. *Int. J. Mol. Sci.* **2018**, *19*, 3851.

(268) Rothhammer, V.; Quintana, F. J. The aryl hydrocarbon receptor: an environmental sensor integrating immune responses in health and disease. *Nat. Rev. Immunol.* **2019**, *19*, 184–197.

(269) Kolluri, S. K.; Jin, U.-H.; Safe, S. Role of the aryl hydrocarbon receptor in carcinogenesis and potential as an anti-cancer drug target. *Arch. Toxicol.* **2017**, *91*, 2497–2513.

(270) Yu, H.; Jiang, L.; Wan, B.; Zhang, W., et al. The role of aryl hydrocarbon receptor in bone remodeling. *Prog. Biophys. Mol. Biol.* **2018**, *134*, 44–49.

(271) Dietrich, C. Antioxidant Functions of the Aryl Hydrocarbon Receptor. *Stem Cells Int.* **2016**, *2016*, 1–10.

(272) Schulte, K. W.; Green, E.; Wilz, A.; Platten, M., et al. Structural Basis for Aryl Hydrocarbon Receptor-Mediated Gene Activation. *Structure* **2017**, *25*, 1025–1033.e3.

(273) Fukunaga, B. N.; Probst, M. R.; Reisz-Porszasz, S.; Hankinson, O. Identification of Functional Domains of the Aryl Hydrocarbon Receptor. *J. Biol. Chem.* **1995**, *270*, 29270–29278.

(274) Coumailleau, P.; Poellinger, L.; Gustafsson, J. A.; Whitelaw, M. L. Definition of a minimal domain of the dioxin receptor that is associated with Hsp90 and maintains wild type ligand binding affinity and specificity. *J. Biol. Chem.* **1995**, *270*, 25291–300.

(275) Pandini, A.; Denison, M. S.; Song, Y.; Soshilov, A. A., et al. Structural and functional characterization of the aryl hydrocarbon receptor ligand binding domain by homology modeling and mutational analysis. *Biochemistry* **2007**, *46*, 696–708.

(276) Ikuta, T.; Eguchi, H.; Tachibana, T.; Yoneda, Y., et al. Nuclear localization and export signals of the human aryl hydrocarbon receptor. *J. Biol. Chem.* **1998**, *273*, 2895–904.

(277) Reyes, H.; Reisz-Porszasz, S.; Hankinson, O. Identification of the Ah receptor nuclear translocator protein (Arnt) as a component of the DNA binding form of the Ah receptor. *Science* **1992**, *256*, 1193–5.

(278) Yao, E. F.; Denison, M. S. DNA sequence determinants for binding of transformed Ah receptor to a dioxin-responsive enhancer. *Biochemistry* **1992**, *31*, 5060–7.

(279) Swanson, H. I.; Chan, W. K.; Bradfield, C. A. DNA Binding Specificities and Pairing Rules of the Ah Receptor, ARNT, and SIM Proteins. *J. Biol. Chem.* **1995**, *270*, 26292–26302.

(280) Richter, C. A.; Tillitt, D. E.; Hannink, M. Regulation of Subcellular Localization of the Aryl Hydrocarbon Receptor (AhR). *Arch. Biochem. Biophys.* **2001**, *389*, 207–217.

(281) Denison, M. S.; Nagy, S. R. Activation of the Aryl Hydrocarbon Receptor by Structurally Diverse Exogenous and Endogenous Chemicals. *Annu. Rev. Pharmacol. Toxicol.* **2003**, *43*, 309–334.

(282) Bradfield, C. A.; Kende, A. S.; Poland, A. Kinetic and equilibrium studies of Ah receptor-ligand binding: use of [125I]2-iodo-7,8-dibromodibenzo-p-dioxin. *Mol. Pharmacol.* **1988**, *34*, 229–37.

(283) Denison, M. S.; Wilkinson, C. F.; Okey, A. B. Ah receptor for 2,3,7,8-tetrachlorodibenzo-p-dioxin: Comparative studies in mammalian and nonmammalian species. *Chemosphere* **1986**, *15*, 1665–1672.

(284) DeGroot, D.; He, G.; Fraccalvieri, D.; Bonati, L., et al. In *AH Recept. Biol. Toxicol.* Pohjanvirta, R., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2011; Chapter 4, pp 63–79.

(285) Song, J.; Clagett-Dame, M.; Peterson, R. E.; Hahn, M. E., et al. A ligand for the aryl hydrocarbon receptor isolated from lung. *Proc. Natl. Acad. Sci.* **2002**, *99*, 14694–14699.

(286) Henry, E.; Bemis, J.; Henry, O.; Kende, A., et al. A potential endogenous ligand for the aryl hydrocarbon receptor has potent agonist activity in vitro and in vivo. *Arch. Biochem. Biophys.* **2006**, *450*, 67–77.

(287) Whelan, F.; Hao, N.; Furness, S. G. B.; Whitelaw, M. L., et al. Amino Acid Substitutions in the Aryl Hydrocarbon Receptor Ligand Binding Domain Reveal YH439 As an Atypical AhR Activator. *Mol. Pharmacol.* **2010**, *77*, 1037–1046.

(288) O'Donnell, E. F.; Saili, K. S.; Koch, D. C.; Kopparapu, P. R., et al. The Anti-Inflammatory Drug Leflunomide Is an Agonist of the Aryl Hydrocarbon Receptor. *PLoS One* **2010**, *5*, ed. by Rodrigues-Lima, F., e13128.

(289) Rannug, U.; Rannug, A.; Sjöberg, U.; Li, H., et al. Structure elucidation of two tryptophan-derived, high affinity Ah receptor ligands. *Chem. Biol.* **1995**, *2*, 841–5.

(290) Smith, K. J.; Murray, I. A.; Tanos, R.; Tellew, J., et al. Identification of a High-Affinity Ligand That Exhibits Complete Aryl Hydrocarbon Receptor Antagonism. *J. Pharmacol. Exp. Ther.* **2011**, *338*, 318–327.

(291) Boitano, A. E.; Wang, J.; Romeo, R.; Bouchez, L. C., et al. Aryl hydrocarbon receptor antagonists promote the expansion of human hematopoietic stem cells. *Science* **2010**, *329*, 1345–8.

(292) Zhao, B.; DeGroot, D. E.; Hayashi, A.; He, G., et al. CH223191 Is a Ligand-Selective Antagonist of the Ah (Dioxin) Receptor. *Toxicol. Sci.* **2010**, *117*, 393–403.

(293) Casper, R. F.; Quesne, M.; Rogers, I. M.; Shirota, T., et al. Resveratrol has antagonist activity on the aryl hydrocarbon receptor: implications for prevention of dioxin toxicity. *Mol. Pharmacol.* **1999**, *56*, 784–90.

(294) Murray, I. A.; Flaveny, C. A.; DiNatale, B. C.; Chairo, C. R., et al. Antagonism of Aryl Hydrocarbon Receptor Signaling by 6,2′,4′-Trimethoxyflavone. *J. Pharmacol. Exp. Ther.* **2010**, *332*, 135–144.

(295) Santostefano, M.; Merchant, M.; Arellano, L.; Morrison, V., et al. alpha-Naphthoflavone-induced CYP1A1 gene expression and cytosolic aryl hydrocarbon receptor transformation. *Mol. Pharmacol.* **1993**, *43*, 200–6.

(296) Zhou, J.; Gasiewicz, T. A. 3'-methoxy-4'-nitroflavone, a reported aryl hydrocarbon receptor antagonist, enhances Cyp1a1 transcription by a dioxin responsive element-dependent mechanism. *Arch. Biochem. Biophys.* **2003**, *416*, 68–80.

(297) Bazzi, R.; Bradshaw, T. D.; Rowlands, J. C.; Stevens, M. F., et al. 2-(4-Amino-3-methylphenyl)-5-fluorobenzothiazole is a ligand and shows species-specific partial agonism of the aryl hydrocarbon receptor. *Toxicol. Appl. Pharmacol.* **2009**, *237*, 102–110.

(298) Zhou, J.-G.; Henry, E. C.; Palermo, C. M.; Dertinger, S. D., et al. Species-specific transcriptional activity of synthetic flavonoids in guinea pig and mouse cells as a result of differential activation of the aryl hydrocarbon receptor to interact with dioxin-responsive elements. *Mol. Pharmacol.* **2003**, *63*, 915–24.

(299) Soshilov, A. A.; Denison, M. S. Ligand Promiscuity of Aryl Hydrocarbon Receptor Agonists and Antagonists Revealed by Site-Directed Mutagenesis. *Mol. Cell. Biol.* **2014**, *34*, 1707–1719.

(300) Behnisch, P. A.; Hosoe, K.; Sakai, S.-i. Bioanalytical screening methods for dioxins and dioxin-like compounds — a review of bioassay/biomarker technology. *Environ. Int.* **2001**, *27*, 413–439.

(301) Soshilov, A.; Denison, M. S. Ligand Displaces Heat Shock Protein 90 from Overlapping Binding Sites within the Aryl Hydrocarbon Receptor Ligand-binding Domain. *J. Biol. Chem.* **2011**, *286*, 35275–35282.

(302) Hao, N.; Whitelaw, M. L.; Shearwin, K. E.; Dodd, I. B., et al. Identification of residues in the N-terminal PAS domains important for dimerization of Arnt and AhR. *Nucleic Acids Res.* **2011**, *39*, 3695–3709.

(303) Parks, A. J.; Pollastri, M. P.; Hahn, M. E.; Stanford, E. A., et al. In Silico Identification of an Aryl Hydrocarbon Receptor Antagonist with Biological Activity In Vitro and In Vivo. *Mol. Pharmacol.* **2014**, *86*, 593–608.

(304) Perkins, A.; Phillips, J.; Kerkvliet, N.; Tanguay, R., et al. A Structural Switch between Agonist and Antagonist Bound Conformations for a Ligand-Optimized Model of the Human Aryl Hydrocarbon Receptor Ligand Binding Domain. *Biology (Basel).* **2014**, *3*, 645–669.

(305) Dolciami, D.; Gargaro, M.; Cerra, B.; Scalisi, G., et al. Binding Mode and Structure-Activity Relationships of ITE as an Aryl Hydrocarbon Receptor (AhR) Agonist. *ChemMedChem* **2018**, *13*, 270–279.

(306) Hu, W.; Sorrentino, C.; Denison, M. S.; Kolaja, K., et al. Induction of Cyp1a1 Is a Nonspecific Biomarker of Aryl Hydrocarbon Receptor Activation: Results of Large Scale Screening of Pharmaceuticals and Toxicants in Vivo and in Vitro. *Mol. Pharmacol.* **2007**, *71*, 1475–1486.

(307) Leavitt, S.; Freire, E. Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr. Opin. Struct. Biol.* **2001**, *11*, 560–6.

(308) Mocz, G.; Ross, J. A. In *Protein-Ligand Interact. Methods Mol. Biol.* Williams, M., Daviter, T., Eds.; Humana Press, Totowa, NJ: 2013; Vol. 1008, pp 169–210.

(309) Maity, S.; Gundampati, R. K.; Suresh Kumar, T. K. NMR Methods to Characterize Protein-Ligand Interactions. *Nat. Prod. Commun.* **2019**, *14*, 1934578X1984929.

(310) McDonnell, J. M. Surface plasmon resonance: towards an understanding of the mechanisms of biological molecular recognition. *Curr. Opin. Chem. Biol.* **2001**, *5*, 572–577.

(311) Wilson, J. L.; Scott, I. M.; McMurry, J. L. Optical biosensing: Kinetics of protein A-IGG binding using biolayer interferometry. *Biochem. Mol. Biol. Educ.* **2010**, *38*, 400–407.

(312) Wienken, C. J.; Baaske, P.; Rothbauer, U.; Braun, D., et al. Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.* **2010**, *1*, 100.

(313) Fan, M. Q.; Bell, A. R.; Bell, D. R.; Clode, S., et al. Recombinant expression of aryl hydrocarbon receptor for quantitative ligand-binding analysis. *Anal. Biochem.* **2009**, *384*, 279–287.

(314) Pandini, A.; Soshilov, A. A.; Song, Y.; Zhao, J., et al. Detection of the TCDD Binding-Fingerprint within the Ah Receptor Ligand Binding Domain by Structurally Driven Mutagenesis and Functional Analysis. *Biochemistry* **2009**, *48*, 5972–5983.

(315) Lee, J.-S.; Iwabuchi, K.; Nomaru, K.; Nagahama, N., et al. Molecular and Functional Characterization of a Novel Aryl Hydrocarbon Receptor Isoform, AHR1β, in the Chicken (Gallus gallus). *Toxicol. Sci.* **2013**, *136*, 450–466.

(316) Gasiewicz, T.; Neal, R. The examination and quantitation of tissue cytosolic receptors for 2,3,7,8-tetrachlorodibenzo-p-dioxin using hydroxylapatite. *Anal. Biochem.* **1982**, *124*, 1–11.

(317) Poland, A.; Glover, E.; Ebetino, F. H.; Kende, A. S. Photoaffinity labeling of the Ah receptor. *J. Biol. Chem.* **1986**, *261*, 6352–65.

(318) Ramadoss, P.; Perdew, G. H. Use of 2-Azido-3-[125I]iodo-7,8-dibromodibenzo-p-dioxin as a Probe to Determine the Relative Ligand Affinity of Human versus Mouse Aryl Hydrocarbon Receptor in Cultured Cells. *Mol. Pharmacol.* **2004**, *66*, 129–136.

(319) Faust, D.; Vondráček, J.; Krčmář, P.; Šmerdová, L., et al. AhR-mediated changes in global gene expression in rat liver progenitor cells. *Arch. Toxicol.* **2013**, *87*, 681–698.

(320) Diani-Moore, S.; Ram, P.; Li, X.; Mondal, P., et al. Identification of the aryl hydrocarbon receptor target gene TiPARP as a mediator of suppression of hepatic gluconeogenesis by 2,3,7,8-tetrachlorodibenzo-p-dioxin and of nicotinamide as a corrective agent for this effect. *J. Biol. Chem.* **2010**, *285*, 38801–10.

(321) Murray, I. A.; Patterson, A. D.; Perdew, G. H. Aryl hydrocarbon receptor ligands in cancer: friend and foe. *Nat. Rev. Cancer* **2014**, *14*, 801–814.

(322) Gonzalez, T.; Catton, E. C.; Rosko, A. E.; Efebera, Y., et al. Aryl Hydrocarbon Receptor (AHR) Antagonism As a Transformative, Dual-Mechanism Novel Therapy for Multiple Myeloma. *Blood* **2018**, *132*, 1933–1933.

(323) Opitz, C. A.; Litzenburger, U. M.; Sahm, F.; Ott, M., et al. An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor. *Nature* **2011**, *478*, 197–203.

(324) DiNatale, B. C.; Smith, K.; John, K.; Krishnegowda, G., et al. Ah Receptor Antagonism Represses Head and Neck Tumor Cell Aggressive Phenotype. *Mol. Cancer Res.* **2012**, *10*, 1369–1379.

(325) Pernomian, L.; da Silva, C. H. Current basis for discovery and development of aryl hydrocarbon receptor antagonists for experimental and therapeutic use in atherosclerosis. *Eur. J. Pharmacol.* **2015**, *764*, 118–123.

(326) Moyer, B. J.; Rojas, I. Y.; Kerley-Hamilton, J. S.; Nemani, K. V., et al. Obesity and fatty liver are prevented by inhibition of the aryl hydrocarbon receptor in both female and male mice. *Nutr. Res.* **2017**, *44*, 38–50.

(327) Nguyen, N. T.; Nakahama, T.; Nguyen, C. H.; Tran, T. T., et al. Aryl hydrocarbon receptor antagonism and its role in rheumatoid arthritis. *J. Exp. Pharmacol.* **2015**, *7*, 29–35.

(328) Smith, B. W.; Rozelle, S. S.; Leung, A.; Ubellacker, J., et al. The aryl hydrocarbon receptor directs hematopoietic progenitor cell expansion and differentiation. *Blood* **2013**, *122*, 376–385.

(329) Angelos, M. G.; Ruh, P. N.; Webber, B. R.; Blum, R. H., et al. Aryl hydrocarbon receptor inhibition promotes hematolymphoid development from human pluripotent stem cells. *Blood* **2017**, *129*, 3428–3439.

(330) Schmees, N.; Gutcher, I.; Roehn, U.; Irlbacher, H., et al. In *Cancer Chem.* American Association for Cancer Research: 2019; Vol. 79, pp 4454–4454.

(331) Zhang, S.; Rowlands, C.; Safe, S. Ligand-dependent interactions of the Ah receptor with coactivators in a mammalian two-hybrid assay. *Toxicol. Appl. Pharmacol.* **2008**, *227*, 196–206.

(332) Matikainen, T.; Perez, G. I.; Jurisicova, A.; Pru, J. K., et al. Aromatic hydrocarbon receptor-driven Bax gene expression is required for premature ovarian failure caused by biohazardous environmental chemicals. *Nat. Genet.* **2001**, *28*, 355–360.

(333) Gouedard, C.; Barouki, R.; Morel, Y. Dietary Polyphenols Increase Paraoxonase 1 Gene Expression by an Aryl Hydrocarbon Receptor-Dependent Mechanism. *Mol. Cell. Biol.* **2004**, *24*, 5209–5222.

(334) Patel, R. D.; Murray, I. A.; Flaveny, C. A.; Kusnadi, A., et al. Ah receptor represses acute-phase response gene expression without binding to its cognate response element. *Lab. Investig.* **2009**, *89*, 695–707.

(335) Ye, M.; Zhang, Y.; Gao, H.; Xu, Y., et al. Activation of the Aryl Hydrocarbon Receptor Leads to Resistance to EGFR TKIs in Non–Small Cell Lung Cancer by Activating Src-mediated Bypass Signaling. *Clin. Cancer Res.* **2018**, *24*, 1227–1239.

(336) Ramsay, R. R.; Popovic-Nikolic, M. R.; Nikolic, K.; Uliassi, E., et al. A perspective on multi-target drug discovery and design for complex diseases. *Clin. Transl. Med.* **2018**, *7*.

(337) Rothhammer, V.; Borucki, D. M.; Kenison, J. E.; Hewson, P., et al. Detection of aryl hydrocarbon receptor agonists in human samples. *Sci. Rep.* **2018**, *8*, 4970.

(338) Puppala, D.; Lee, H.; Kim, K. B.; Swanson, H. I. Development of an Aryl Hydrocarbon Receptor Antagonist Using the Proteolysis-Targeting Chimeric Molecules Approach: A Potential Tool for Chemoprevention. *Mol. Pharmacol.* **2008**, *52*, 202–211.

(339) Chatzopoulou, M.; Claridge, T. D.; Davies, K. E.; Davies, S. G., et al. Isolation, structural identification, synthesis, and pharmacological profiling of 1,2-trans-dihydro-1,2-diol metabolites of the utrophin modulator ezutromid. *Submitted* **2019**.

(340) Fang, Z.-Z.; Krausz, K. W.; Nagaoka, K.; Tanaka, N., et al. In vivo effects of the pure aryl hydrocarbon receptor antagonist GNF-351 after oral administration are limited to the gastrointestinal tract. *Br. J. Pharmacol.* **2014**, *171*, 1735–1746.

(341) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **2004**, *32*.

(342) Dolinsky, T. J.; Czodrowski, P.; Li, H.; Nielsen, J. E., et al. PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **2007**, *35*.

(343) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F., et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.

(344) Marvin16.16.6.0 ChemAxon (http://www.chemaxon.com).

(345) Kuzmič, P. Program DYNAFIT for the Analysis of Enzyme Kinetic Data: Application to HIV Proteinase. *Anal. Biochem.* **1996**, *237*, 260–273.

(346) Winter, G.; McAuley, K. E. Automated data collection for macromolecular crystallography. *Methods* **2011**, *55*, 81–93.

(347) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 486–501.

(348) Livak, K. J.; Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2-\Delta\Delta$CT Method. *Methods* **2001**, *25*, 402–408.

(349) Wiśniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **2009**, *6*, 359–362.

(350) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26*, 1367–1372.

(351) Wu, D.; Potluri, N.; Lu, J.; Kim, Y., et al. Structural integration in hypoxia-inducible factors. *Nature* **2015**, *524*, 303–308.

(352) Wu, D.; Su, X.; Potluri, N.; Kim, Y., et al. NPAS1-ARNT and NPAS3-ARNT crystal structures implicate the bHLH-PAS family as multi-ligand binding transcription factors. *Elife* **2016**, *5*.

(353) Heller, K.; Ochtrop, P.; Albers, M. F.; Zauner, F. B., et al. Covalent Protein Labeling by Enzymatic Phosphocholination. *Angew. Chemie Int. Ed.* **2015**, *54*, 10327–10330.

(354) Sissi, C.; Dovigo, L.; Greco, M. L.; Ciancetta, A., et al. Conjugates between minor groove binders and Zn(II)-tach complexes: Synthesis, characterization, and interaction with plasmid DNA. *Tetrahedron* **2017**, *73*, 3014–3024.

(355) Irvine, M. W.; Costa, B. M.; Dlaboga, D.; Culley, G. R., et al. Piperazine-2,3-dicarboxylic Acid Derivatives as Dual Antagonists of NMDA and GluK1-Containing Kainate Receptors. *J. Med. Chem.* **2012**, *55*, 327–341.

(356) Rhodes, W. C.; McElroy, W. D. The Synthesis and Function of Luciferyl-adenylate and Oxyluciferyl-adenylate. *J. Biol. Chem.* **1958**, *233*, 1528.

(357) Imai, K.; Goto, T. Improved synthesis of firefly D-luciferyl-D-adenylate-A key intermediate of firefly bioluminescence. *Agric. Biol. Chem.* **1988**, *52*, 2803–2809.

(358) Ciuffreda, P.; Casati, S.; Meroni, G.; Santaniello, E. A new synthesis of dehydroluciferin [2-(6′-hydroxy-2′-benzothiazolyl)-thiazole-4-carboxylic acid] from 1,4-benzoquinone. *Tetrahedron* **2013**, *69*, 5893–5897.

(359) Nakatsu, T.; Ichiyama, S.; Hiratake, J.; Saldanha, A., et al. Structural basis for the spectral difference in luciferase bioluminescence. *Nature* **2006**, *440*, 372–376.

(360) Heacock, D.; Forsyth, C. J.; Shiba, K.; Musier-Forsyth, K. Synthesis and Aminoacyl-tRNA Synthetase Inhibitory Activity of Prolyl Adenylate Analogs. *Bioorg. Chem.* **1996**, *24*, 273–289.

Appendix A

Utrophin-luciferase reporter gene assay results

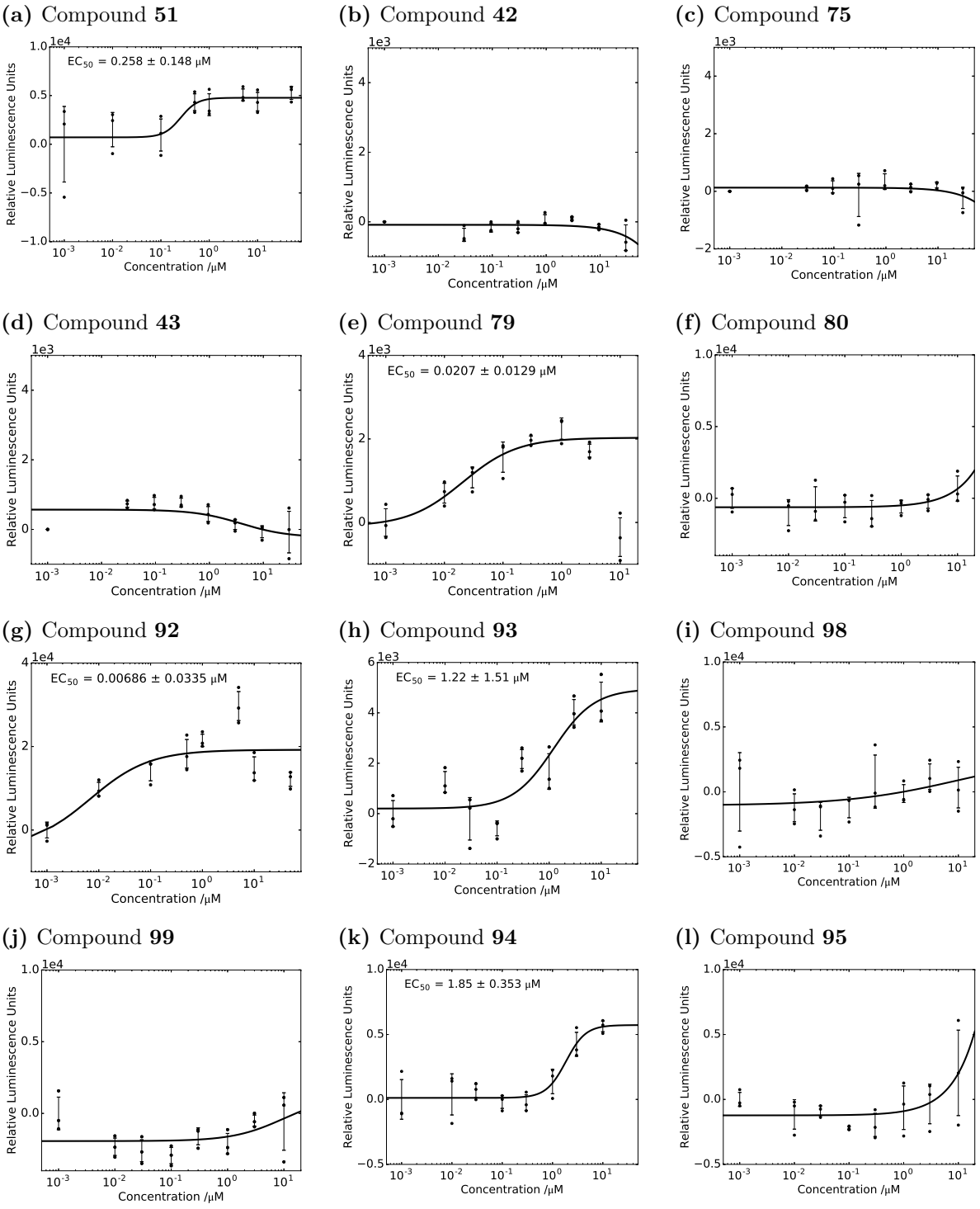


Figure A.1: Results of compound testing in the H2K *mdx* utrophin A-FLuc reporter assay. Data collected by Sarah Squire (Davies group, DPAG) and collaborators at Evotec. EC_{50} were calculated by fitting a four-parameter logistic function to the data, error indicates goodness of fit.

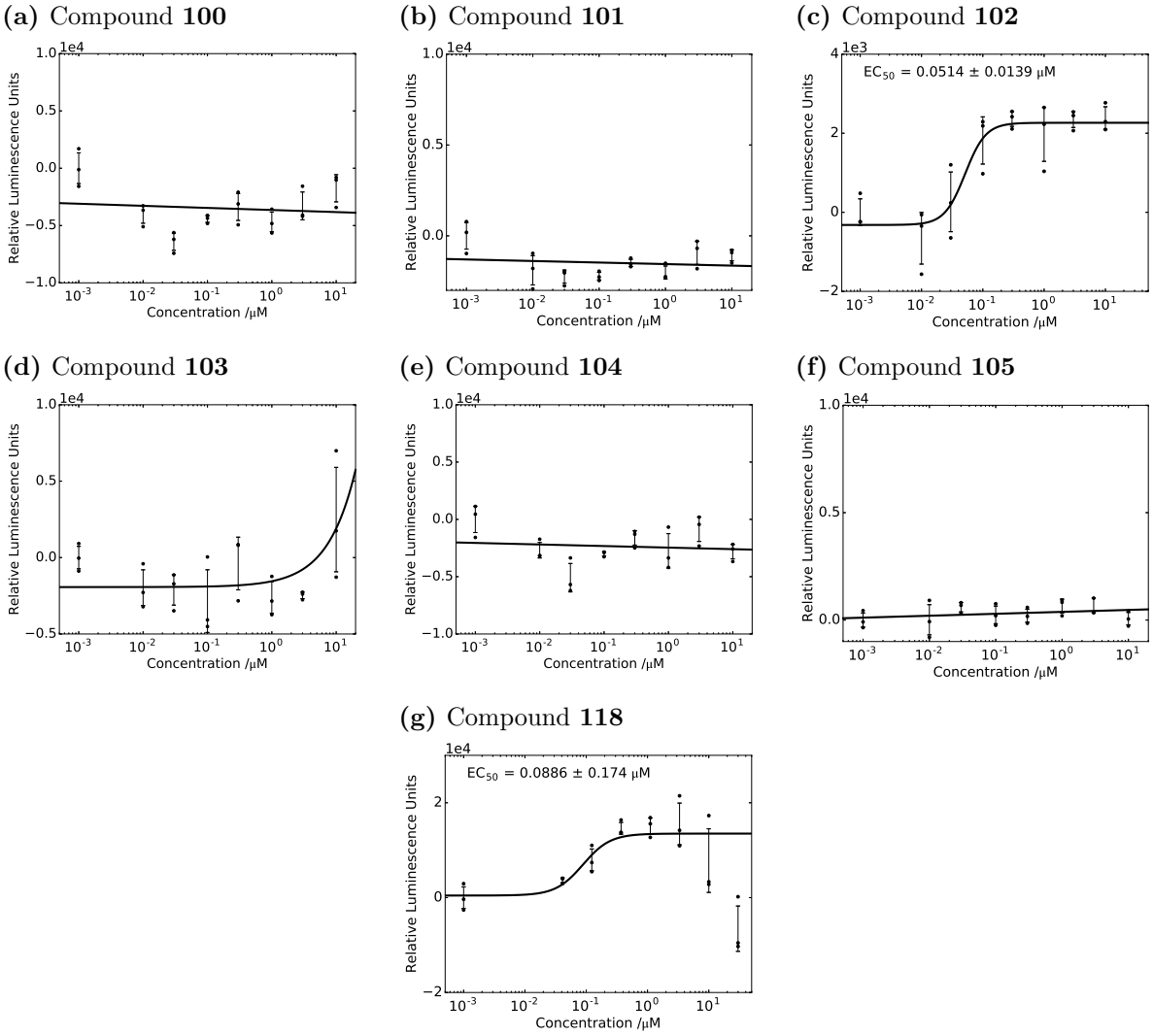


Figure A.2: Results of compound testing in the H2K *mdx* utrophin A-FLuc reporter assay continued. Data collected by Sarah Squire (Davies group, DPAG) and collaborators at Evotec. EC_{50} were calculated by fitting a four-parameter logistic function to the data, error indicates goodness of fit.

Appendix B

FLuc inhibition assay results

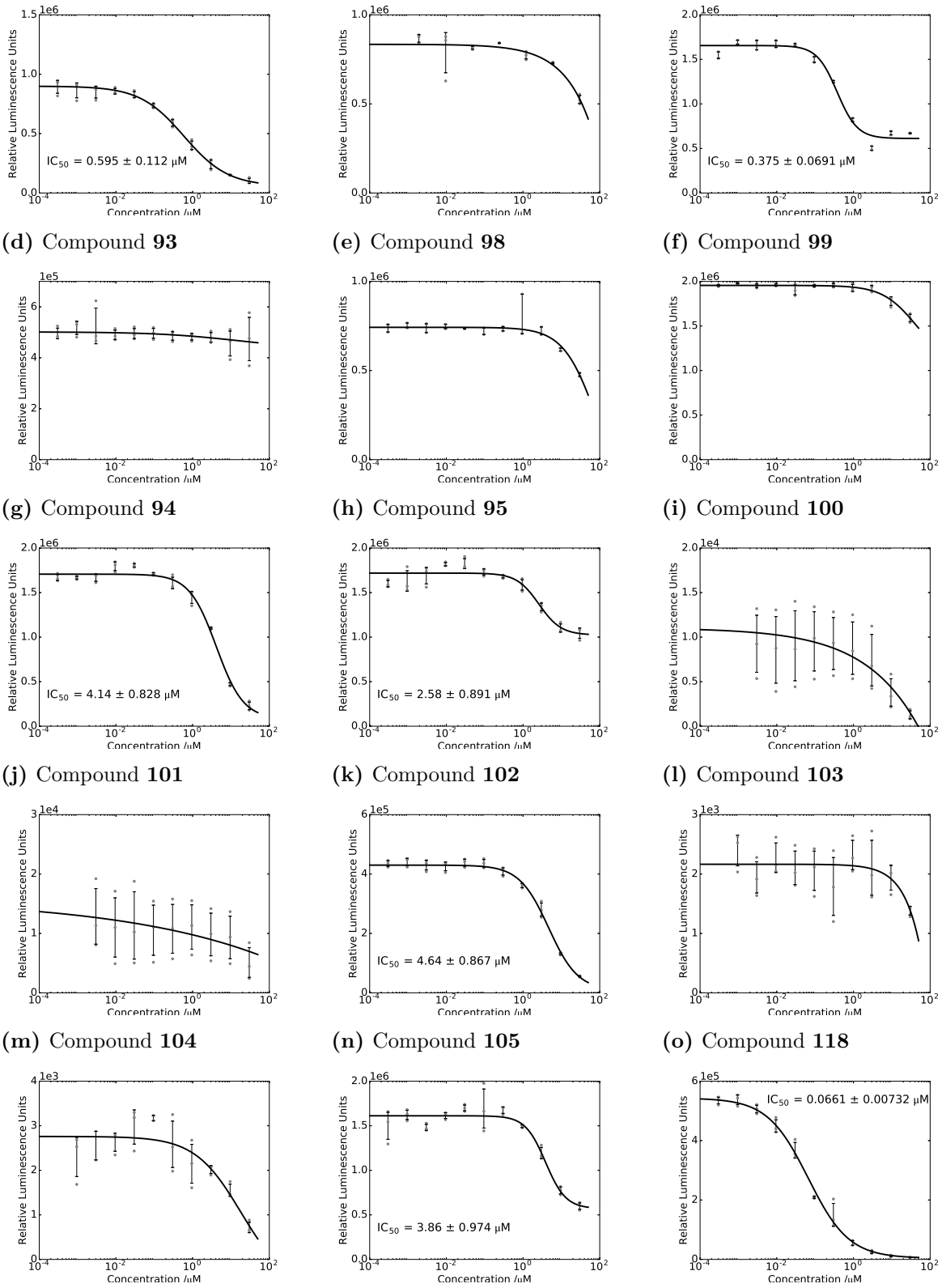


Figure B.1: Results of compound testing in the biochemical FLuc inhibition assay. IC_{50} were calculated by fitting a four-parameter logistic function to the data, error indicates goodness of fit.

Appendix C

Supplementary kinetic data for chapter 3

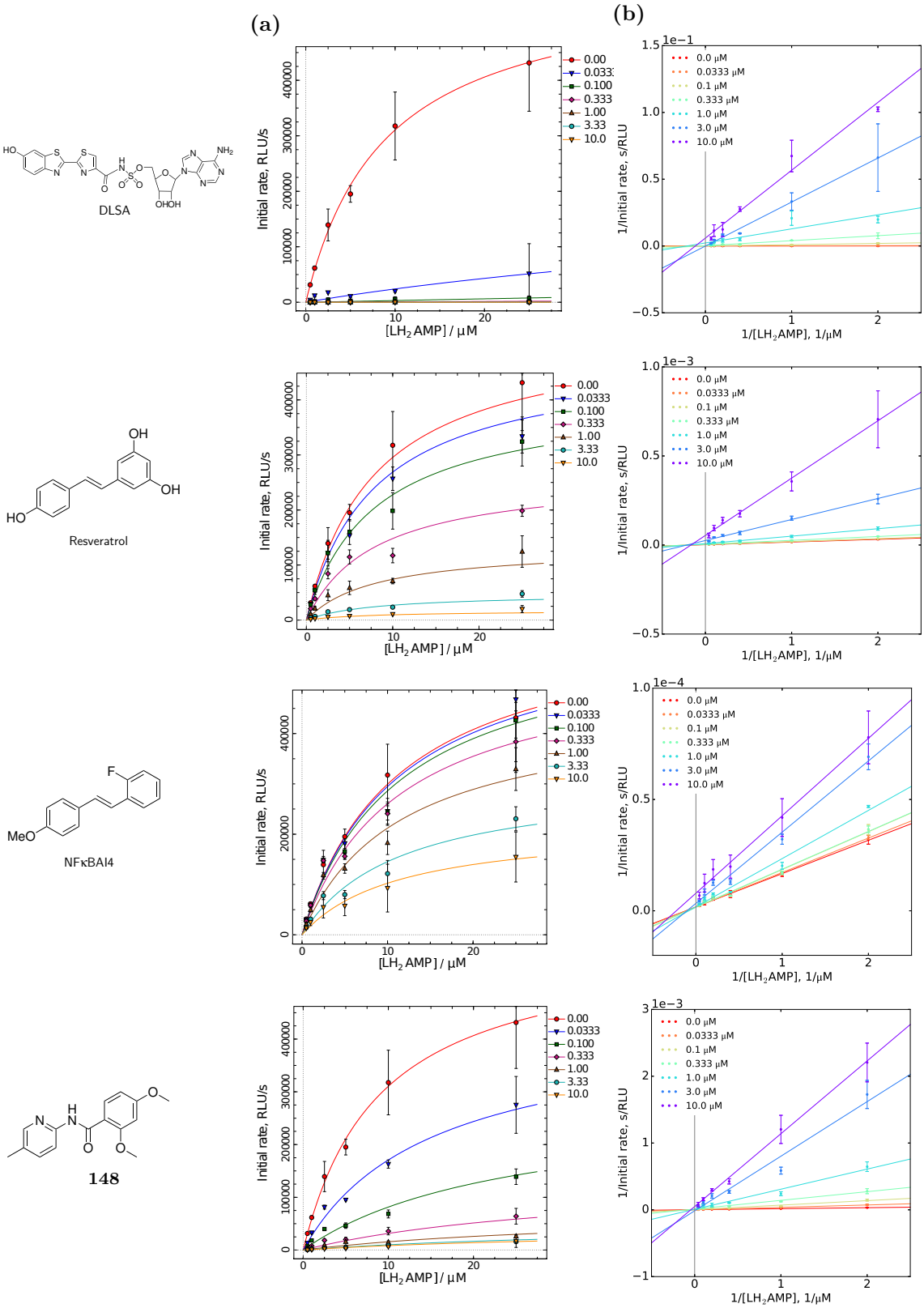


Figure C.1: The dependence of the initial velocity on the concentration of luciferyl adenylate (LH_2-AMP) at different concentrations of DLSA, resveratrol, NF-κBAI4 and **148**. **(a)** Curves were fitted using the Michaelis-Menten equation under rapid-equilibrium steady state kinetics to the data using Dynafit. Inhibition model selection was achieved with Akaike and Bayesian Information Criteria; **(b)** The same data presented in double reciprocal Lineweaver-Burk plots.

Appendix D

MS spectra of luciferase irradiation study

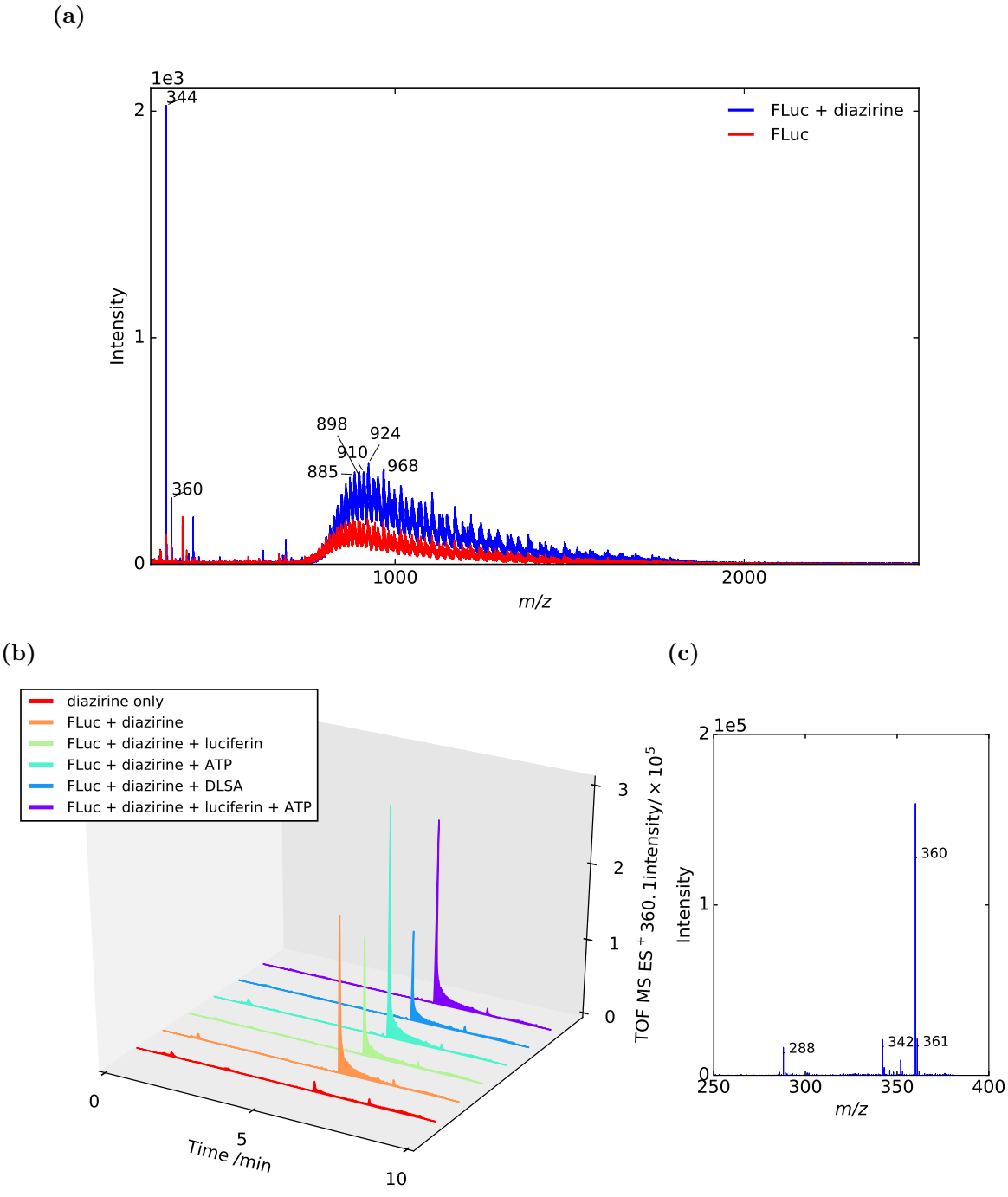


Figure D.1: **(a)** Irradiation of diazirine **79** in the presence of firefly luciferase results in no incorporation into the protein (blue spectrum), which is unchanged from the non-treated control (red spectrum), instead generating water (m/z 344) and H_2O_2 (m/z 360) adducts; **(b)** chromatograms of diazirine irradiated samples in the presence and absence of FLuc and its substrates, extracted for ES^+ ions with m/z of 360.1 (example spectrum **(c)**), which show the dependence of the 360.1 ion occurrence on the presence of FLuc.

Appendix E

Uncropped gels and blots

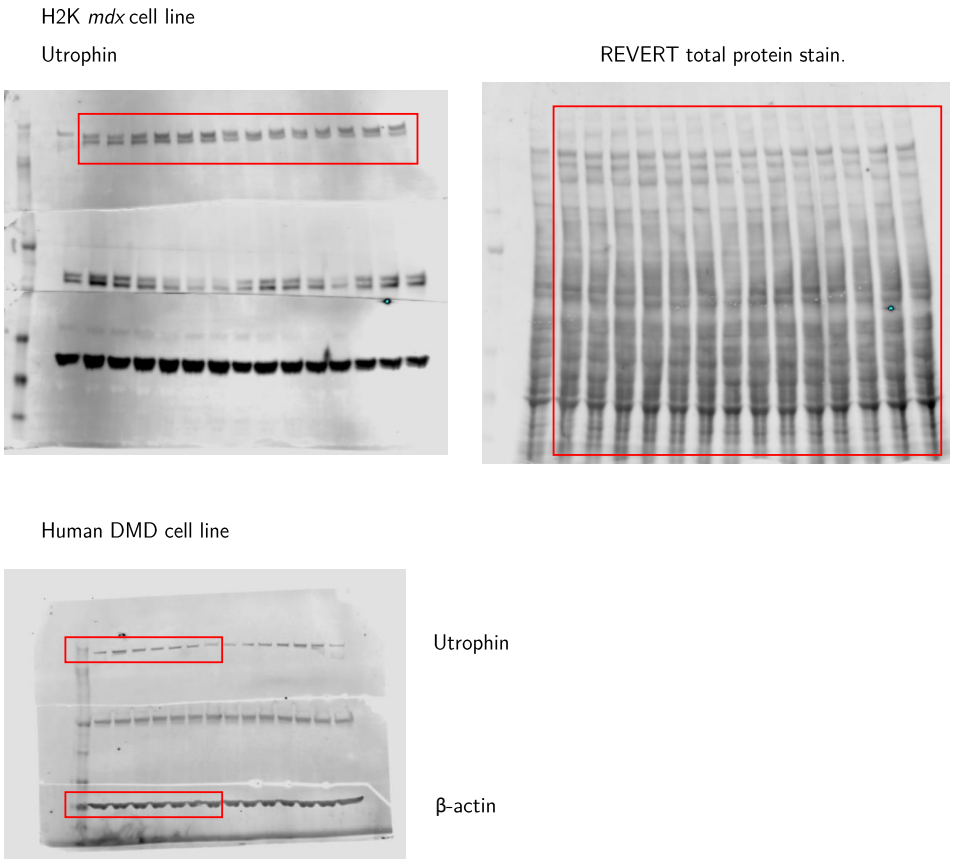


Figure E.1: Uncropped image of immunoblots presented in Figure 4.4 and Figure 4.5.

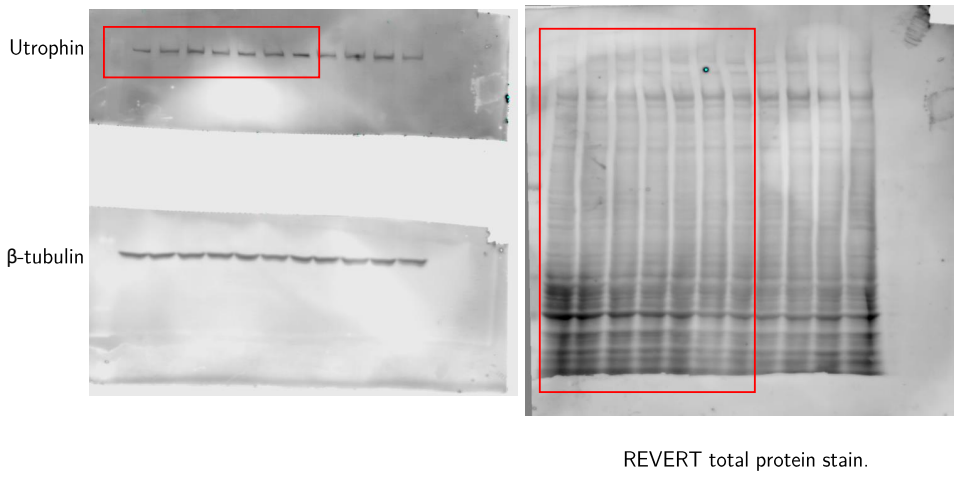


Figure E.2: Uncropped image of immunoblot presented in Figure 4.6.

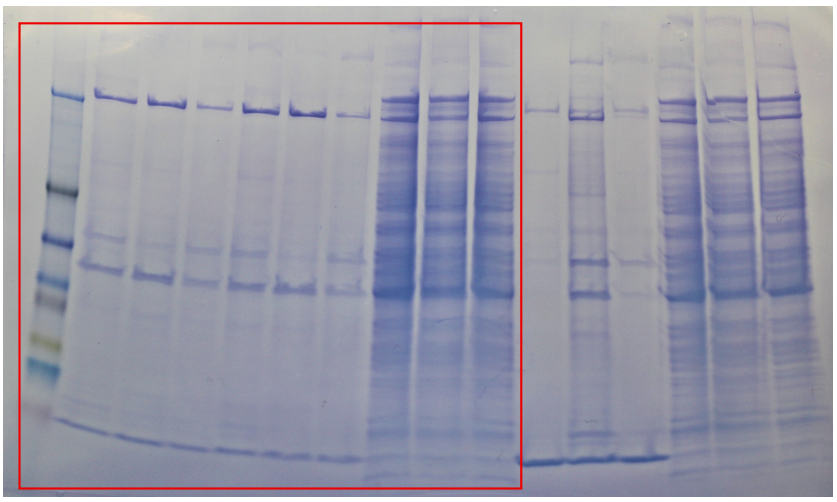


Figure E.3: Uncropped image of gel presented in Figure 5.2.

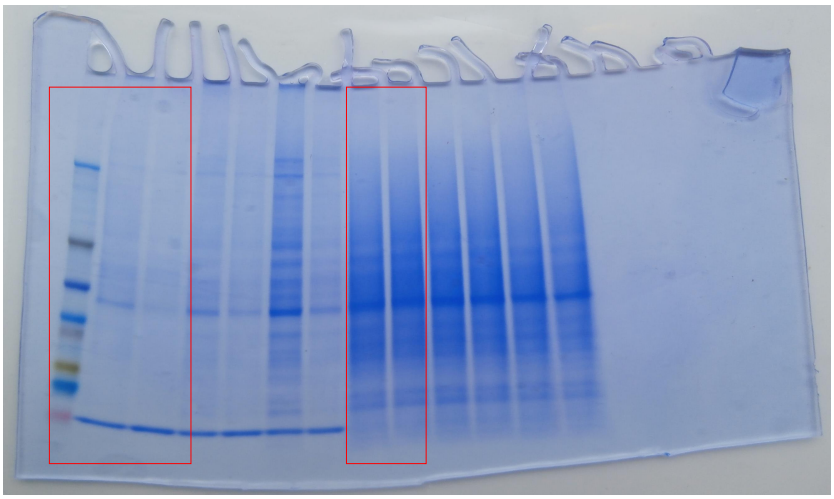


Figure E.4: Uncropped image of gel presented in Figure 5.7.

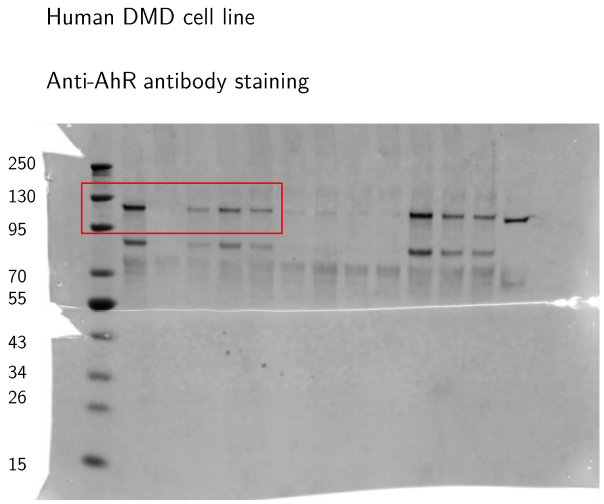


Figure E.5: Uncropped image of immunoblot presented in Figure 6.7.

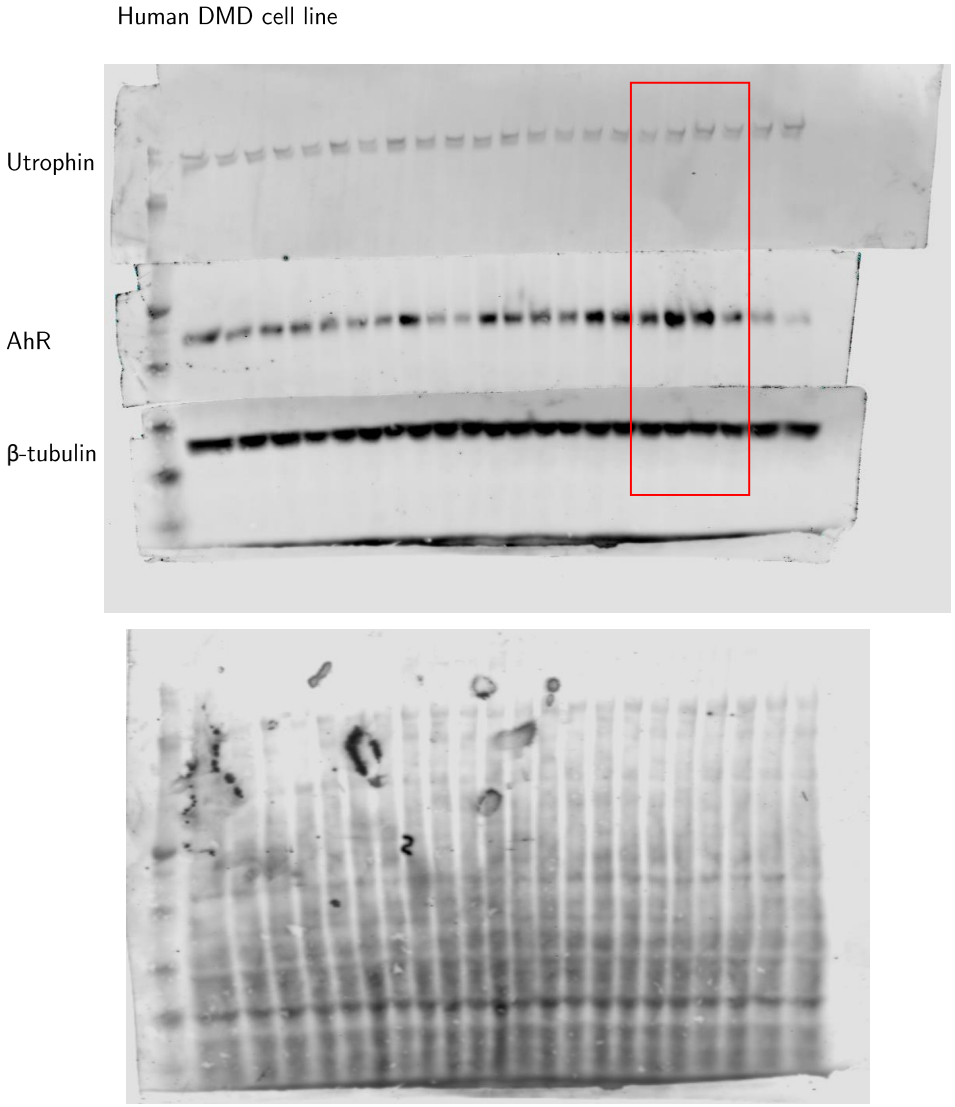


Figure E.6: Uncropped image of immunoblot presented in Figure 6.11.

Appendix F

LC-MS chromatograms and mass spectra

Data accompanying Figure 5.6

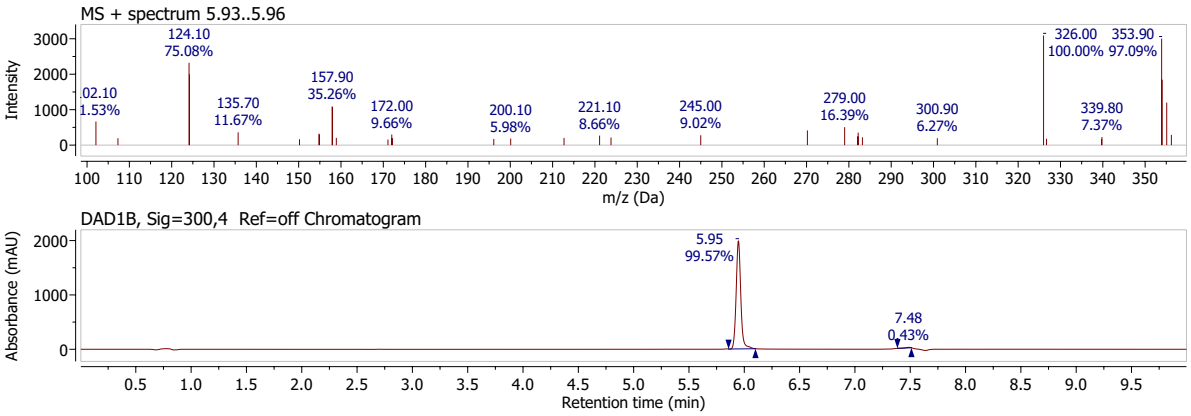


Figure F.1: Compound **79** before irradiation (t = 0). Top: positive ion mass spectrum for the RT = 5.93-5.96 min peak, bottom: full UV absorbance (300 nm) chromatogram.

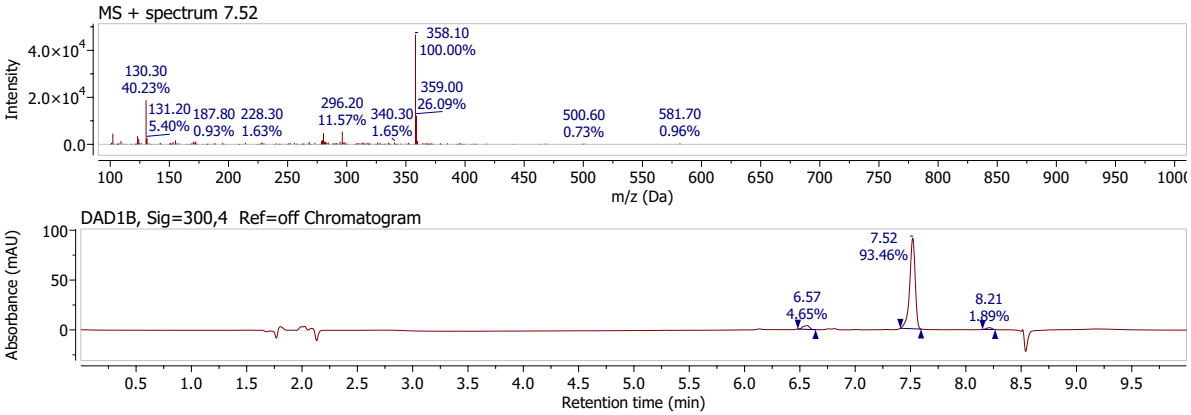


Figure F.2: Compound **79** after 0.5 min irradiation (365 nm, 0 °C, 100 W). Top: positive ion mass spectrum for the RT = 7.52 min peak, bottom: full UV absorbance (300 nm) chromatogram.

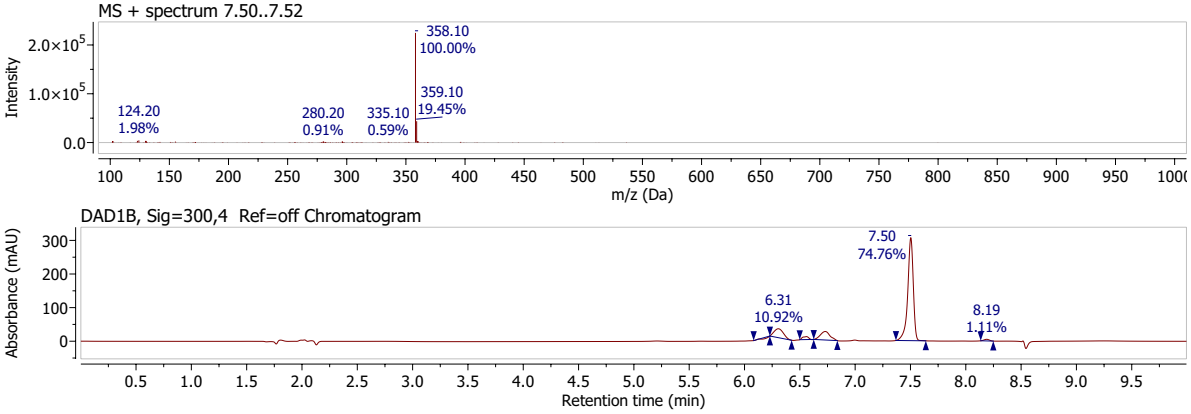


Figure F.3: Compound **79** after 2.5 min irradiation (365 nm, 0 °C, 100 W). Top: positive ion mass spectrum for the RT = 7.50-7.52 min peak, bottom: full UV absorbance (300 nm) chromatogram.

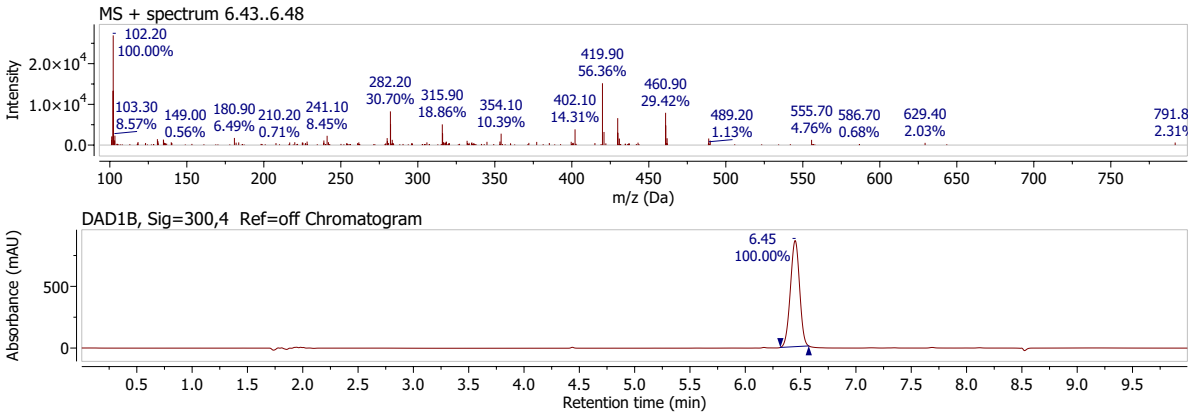


Figure F.4: Compound **126** before irradiation (t = 0). Top: positive ion mass spectrum for the RT = 6.43-6.48 min peak (compound **126** does not ionise well), bottom: full UV absorbance (300 nm) chromatogram.

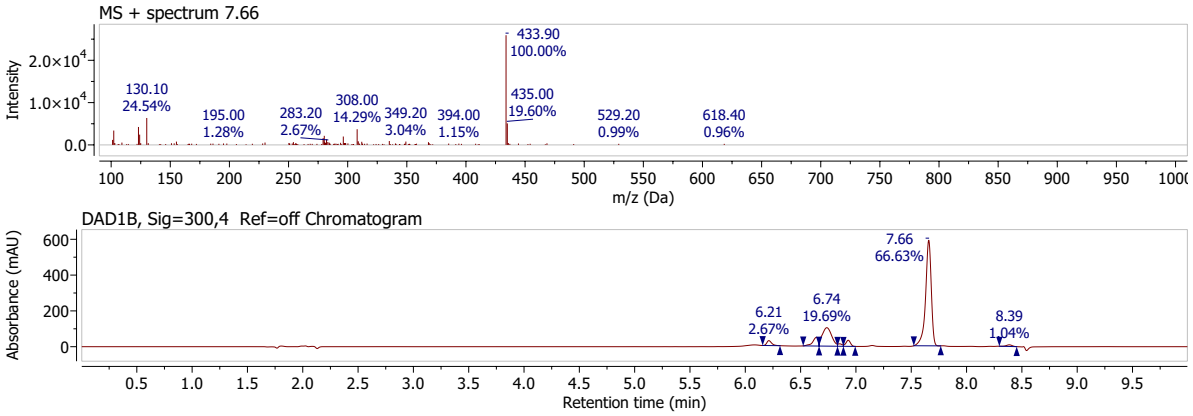


Figure F.5: Compound **126** after 0.5 min irradiation (365 nm, 0 °C, 100 W). Top: positive ion mass spectrum for the RT = 7.66 min peak, bottom: full UV absorbance (300 nm) chromatogram.

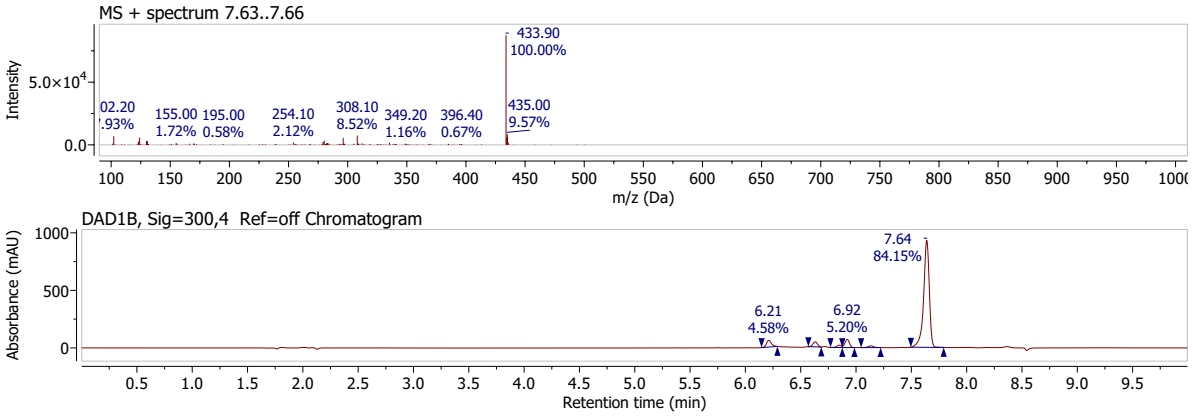


Figure F.6: Compound **126** after 2.5 min irradiation (365 nm, 0 °C, 100 W). Top: positive ion mass spectrum for the RT = 7.63-7.66 min peak, bottom: full UV absorbance (300 nm) chromatogram.

Appendix G

Proteomics data

Table G.1: Full proteomics data for the AfBPP experiment, including log10 transformed intensity values for each replicate and calculated Student's t-test -log10(p-values) and differences (Perseus). Active probe (A) is **106**, inactive probe (IA) is **129** and competitor (C) is **151**. Pept.; peptides, cov.; coverage.

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Acaca	Acetyl-CoA carboxylase 1; Biotin carboxylase	17.5	30.4	30.3	29.6	30.3	30.1	26.6	17.5	26.0	62	28.8	0	0.388	-3.935	0.197	2.655
Acot7	Cytosolic acyl coenzyme A thioester hydrolase	17.5	22.6	17.5	17.5	17.5	17.5	17.5	17.5	22.5	2	3.1	0.0040568	0.427	1.711	0.006	0.043
Acta1; Actc1; Acta2	Actin, alpha skeletal muscle;Actin, alpha cardiac muscle 1	23.4	24.7	24.4	25.7	25.6	27.0	20.2	23.7	24.1	22	19.9	0	1.503	-1.968	0.479	1.442
Actb; Actg1	Actin, cytoplasmic 1; Actin, cytoplasmic 1, N-terminally processed	29.1	31.0	31.6	31.5	31.9	31.3	30.5	31.1	32.2	28	36.8	0	0.566	-0.978	0.308	-0.673
Actn1	Alpha-actinin-1	20.5	23.0	22.8	23.4	17.5	23.2	17.5	17.5	24.9	12	11	0	0.126	0.722	0.349	2.167
Actn4	Alpha-actinin-4	21.5	24.5	25.0	25.6	23.6	24.7	17.5	24.1	26.0	13	12.6	0	0.316	-0.948	0.158	1.172
Aga	N(4)-(beta-N-acetyl-glucosaminyl)-L-asparaginase; Glycosylasparaginase alpha chain	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	1	3.6	0.004065	0.000	0.000	0.000	0.000
Ahcy	Adenosylhomocysteinase	21.5	17.5	23.4	23.6	17.5	17.5	17.5	22.4	23.5	2	5.6	0	0.179	1.260	0.049	-0.363
Ahnak		27.9	27.9	28.1	28.2	27.0	28.0	25.0	28.1	28.2	66	38.1	0	0.193	0.181	0.323	0.820
Ahnak2		17.5	17.5	17.5	17.5	17.5	22.8	17.5	22.6	21.1	4	40.9	0	0.427	-1.783	0.892	-2.917
Ahr	Aryl hydrocarbon receptor	17.5	24.7	24.8	17.5	17.5	17.5	20.0	17.5	17.5	3	5	0	0.935	4.833	0.713	3.992
Akr1b1	Aldose reductase	22.3	22.0	21.9	23.9	23.6	23.5	17.5	24.2	24.7	1	4.4	0	3.086	-1.605	0.009	-0.062
Alb	Serum albumin	17.5	24.5	17.5	17.5	28.2	17.5	17.5	17.5	17.5	6	10.2	0	0.104	-1.236	0.427	2.332
Aldh2	Aldehyde dehydrogenase, mitochondrial	17.5	17.5	17.5	21.9	17.5	17.5	17.5	17.5	22.6	1	2.7	0	0.427	-1.477	0.427	-1.688
Aldh9a1	4-trimethylamino-butyraldehyde dehydrogenase	17.5	17.5	17.5	17.5	20.6	17.5	17.5	17.5	17.5	1	3	0.003992	0.427	-1.034	0.000	0.000

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Aldoa; Aldoart1	Fructose-bisphosphate aldolase A; Fructose-bisphosphate aldolase	23.2	25.8	26.2	26.3	24.0	22.4	21.8	25.8	27.2	12	38.7	0	0.219	0.824	0.024	0.131
Anxa1	Annexin A1	24.6	25.9	25.0	25.5	24.6	23.6	23.5	24.9	25.7	7	30.9	0	0.396	0.626	0.240	0.456
Anxa2	Annexin A2	25.4	26.7	28.2	26.9	26.1	26.3	24.8	27.0	27.0	16	42.5	0	0.144	0.324	0.179	0.516
Anxa5	Annexin A5	17.5	17.5	17.5	20.4	17.5	17.5	17.5	21.6	22.6	2	6.6	0.0022075	0.427	-0.961	0.918	-3.053
Arf3; Arf1; Arf2	ADP-ribosylation factor 3; ADP-ribosylation factor 1	17.5	17.5	17.5	23.5	17.5	23.3	17.5	17.5	24.2	2	7.7	0	0.934	-3.931	0.427	-2.245
Arfip1		19.8	28.1	28.8	22.0	21.9	21.9	24.5	17.5	24.5	11	49.1	0	0.563	3.667	0.387	3.412
Arfip2	Arfaptin-2	17.5	23.3	23.2	17.5	17.5	17.5	17.5	17.5	17.5	2	7.9	0	0.935	3.853	0.935	3.853
Arg1	Arginase-1	21.9	23.9	27.4	17.5	23.8	22.3	20.3	23.7	23.3	2	6.8	0	0.568	3.181	0.433	1.940
Arglu1	Arginine and glutamate-rich protein 1	21.5	17.5	17.5	22.1	17.5	21.8	17.5	17.5	17.5	2	5.5	0.00409	0.338	-1.631	0.427	1.330
Arhgdia	Rho GDP-dissociation inhibitor 1	21.0	23.4	23.5	23.9	21.6	21.6	22.6	22.1	24.0	3	23.5	0	0.080	0.254	0.097	-0.271
Arpc1b	Actin-related protein 2/3 complex subunit 1B	17.5	21.9	17.5	20.8	21.5	21.1	19.7	21.5	17.5	2	8.3	0.0039293	0.639	-2.113	0.112	-0.578
Atic	Bifunctional purine biosynthesis protein PURH; Phosphoribosyl-aminoimidazolecarboxamide formyltransferase	17.5	21.6	21.3	23.1	17.5	20.8	17.5	17.5	22.4	2	5.2	0.0042105	0.052	-0.316	0.185	1.015
Atp1a1	Sodium/potassium-transporting ATPase subunit alpha-1	17.5	23.1	24.0	22.6	17.5	22.6	17.5	21.4	23.5	3	4.9	0	0.083	0.621	0.098	0.732
Atp2a2	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	17.5	24.8	25.1	23.6	17.5	23.9	19.9	17.5	23.1	8	12	0	0.088	0.801	0.315	2.289
Atp5a1	ATP synthase subunit alpha, mitochondrial; ATP synthase subunit alpha	24.4	22.9	23.7	26.0	25.1	26.2	17.5	24.4	27.1	9	20.1	0	1.706	-2.094	0.085	0.689
Atp5b	ATP synthase subunit beta, mitochondrial	22.7	23.1	23.0	23.7	17.5	17.5	17.5	21.9	24.4	4	12.7	0	0.751	3.382	0.337	1.656
Atp5o	ATP synthase subunit O, mitochondrial	17.5	17.5	22.6	17.5	17.5	23.4	17.5	17.5	22.1	1	5.2	0.0039604	0.033	-0.256	0.022	0.155
Atpif1	ATPase inhibitor, mitochondrial	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	22.3	2	17.9	0	0.000	0.000	0.427	-1.614

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Atxn2l	Ataxin-2-like protein	20.7	20.6	17.5	17.5	17.5	17.5	17.5	17.5	22.2	1	1.6	0.0043764	0.935	2.078	0.100	0.521
Basp1	Brain acid soluble protein 1	22.2	23.5	17.5	24.3	17.5	17.5	21.5	24.6	25.4	6	61.9	0	0.171	1.321	0.557	-2.744
Bclaf1	Bcl-2-associated transcription factor 1	21.6	22.1	22.6	22.8	17.5	21.9	21.0	23.0	22.8	6	8.5	0	0.348	1.388	0.090	-0.178
Bst2	Bone marrow stromal antigen 2	21.5	24.1	17.5	23.9	23.7	22.9	26.6	17.5	22.4	2	17.4	0	0.557	-2.453	0.125	-1.106
C030048-H21Rik		21.2	23.2	25.5	22.4	22.0	21.0	17.5	22.6	20.6	1	5	0.0038095	0.494	1.505	0.733	3.091
Cad		21.3	20.4	17.5	22.9	17.5	24.9	17.5	17.5	24.8	6	4.7	0	0.338	-2.038	0.024	-0.196
Cald1		24.6	24.3	24.2	25.5	24.3	24.4	22.1	25.6	26.8	11	2.4	0	0.390	-0.383	0.114	-0.448
Calr	Calreticulin	20.3	22.8	22.5	24.5	17.5	22.1	17.5	23.1	21.7	3	11.8	0	0.086	0.531	0.231	1.094
Camk2a	Calcium/calmodulin-dependent protein kinase type II subunit alpha	25.4	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	1	4.2	0.0055866	0.427	2.647	0.427	2.647
Cand1	Cullin-associated NEDD8-dissociated protein 1	17.5	21.1	21.5	22.8	22.2	23.9	20.6	17.5	23.1	4	4.4	0	1.003	-2.922	0.066	-0.390
Cap1	Adenylyl cyclase-associated protein 1	17.5	17.5	17.5	21.0	17.5	17.5	17.5	17.5	23.7	1	3.8	0.0056818	0.427	-1.182	0.427	-2.064
Capg	Macrophage-capping protein	17.5	20.4	21.4	21.2	17.5	20.9	17.5	17.5	21.2	1	3.7	0.0039841	0.013	-0.064	0.245	1.050
Caprin1	Caprin-1	20.8	22.7	22.8	22.8	17.5	17.5	17.5	22.5	23.6	2	3.3	0.0023364	0.679	2.817	0.173	0.904
Capza2	F-actin-capping protein subunit alpha-2	21.8	23.4	22.8	22.4	17.5	20.9	17.5	17.5	24.0	3	14.7	0	0.724	2.396	0.607	2.993
Cat	Catalase	20.9	20.7	23.3	17.5	17.5	17.5	17.5	17.5	17.5	2	4.4	0	2.140	4.122	2.140	4.122
Cav1; Cav3	Caveolin; Caveolin-1/3	20.6	23.0	24.6	21.2	17.5	17.5	21.6	17.5	22.5	2	36.2	0	1.121	4.023	0.498	2.204
Cbx3	Chromobox protein homolog 3	17.5	22.1	23.0	21.9	22.2	21.4	17.5	22.7	23.1	1	7.7	0	0.214	-0.948	0.030	-0.219
Cct2	T-complex protein 1 subunit beta	20.2	22.4	23.6	24.8	21.1	17.5	17.5	21.1	24.2	4	13.1	0	0.150	0.931	0.201	1.134
Cct3	T-complex protein 1 subunit gamma	26.3	30.6	31.2	31.1	29.6	28.0	28.3	30.8	30.9	9	22.4	0	0.040	-0.209	0.126	-0.600
Cct4	T-complex protein 1 subunit delta	20.4	22.7	17.5	25.0	25.4	26.2	23.0	25.2	25.1	3	7.8	0	1.598	-5.350	1.211	-4.274
Cct5	T-complex protein 1 subunit epsilon	17.5	22.6	17.5	23.2	17.5	22.8	17.5	17.5	24.3	4	9.2	0	0.325	-1.988	0.068	-0.550

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Cct6a	T-complex protein 1 subunit zeta	20.1	22.4	17.5	23.7	23.5	24.0	20.9	17.5	24.3	5	13.6	0	1.244	-3.746	0.137	-0.896
Cct7	T-complex protein 1 subunit eta	20.9	21.4	20.8	24.1	22.5	22.9	17.5	20.2	23.4	4	10.7	0	1.813	-2.111	0.141	0.657
Cct8	T-complex protein 1 subunit theta	20.9	22.7	23.5	24.1	25.2	25.1	21.8	23.7	25.1	8	17.9	0	1.327	-2.457	0.403	-1.179
Cd44	CD44 antigen	17.5	21.0	17.5	21.8	17.5	20.5	18.7	20.4	21.4	1	1.7	0.0042918	0.307	-1.296	0.474	-1.534
Cdc42	Cell division control protein 42 homolog	17.5	17.5	21.0	23.0	17.5	17.5	17.5	17.5	17.5	1	11	0.0042463	0.109	-0.661	0.427	1.178
Cdsn	Corneodesmosin	20.1	21.8	24.1	22.1	20.5	17.5	17.5	21.0	20.2	1	1.6	0.0038986	0.493	2.006	0.723	2.470
Cdv3	Protein CDV3	23.0	17.5	17.5	17.5	17.5	17.5	17.5	23.1	24.0	3	30.2	0	0.427	1.819	0.333	-2.198
Cfdp1	Craniofacial development protein 1	20.0	17.5	22.4	21.3	17.5	17.5	17.5	21.8	21.9	2	12.5	0.0022272	0.246	1.184	0.078	-0.448
Cfl1	Cofilin-1	24.9	25.4	27.0	26.0	25.2	25.6	23.4	27.2	27.4	8	60.8	0	0.076	0.149	0.056	-0.241
Chchd1; Pip5kl1	Coiled-coil-helix-coiled-coil-helix domain-containing protein 1; Phosphatidylinositol 4-phosphate 5-kinase-like protein 1	23.0	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	1	5.9	1	0.427	1.823	0.427	1.823
Ckap4	Cytoskeleton-associated protein 4	22.9	26.1	26.9	25.1	24.3	24.0	24.6	22.3	25.7	17	40.2	0	0.252	0.800	0.283	1.108
Clic1	Chloride intracellular channel protein 1	21.3	23.1	17.5	17.5	21.3	17.5	17.5	23.3	25.0	3	23.7	0	0.383	1.891	0.173	-1.287
Clic4	Chloride intracellular channel protein 4	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	23.2	1	4.7	0	0.000	0.000	0.427	-1.914
Cltc	Clathrin heavy chain 1	21.6	25.2	25.1	26.1	22.9	25.9	17.5	23.2	26.3	12	11.3	0	0.259	-1.010	0.218	1.592
Cnn3	Calponin-3	17.5	17.5	17.5	17.5	17.5	22.0	17.5	17.5	17.5	1	5.8	0.0042017	0.427	-1.506	0.000	0.000
Col1a1	Collagen alpha-1(I) chain	22.3	22.6	17.5	22.3	22.4	22.3	17.5	24.0	24.2	10	14.6	0	0.389	-1.518	0.149	-1.095
Copa	Coatomer subunit alpha;Xenin	17.5	17.5	22.2	23.1	23.8	26.0	17.5	17.5	24.2	5	5.1	0	1.357	-5.229	0.088	-0.679
Copg1	Coatomer subunit gamma-1	17.5	17.5	17.5	17.5	17.5	23.0	21.6	17.5	22.0	2	3.3	0.0042553	0.427	-1.840	0.932	-2.857
Cyb5b	Cytochrome b5 type B	21.4	17.5	24.7	17.5	17.5	17.5	17.5	17.5	17.5	1	23.3	0	0.824	3.684	0.824	3.684
Cyb5r3	NADH-cytochrome b5 reductase 3; NADH-cytochrome b5 reductase 3 membrane-bound form	19.9	21.4	22.7	23.4	17.5	21.5	17.5	17.5	22.4	2	9.6	0	0.105	0.557	0.541	2.218

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Dad1	Dolichyl-diphosphooligosaccharide–protein glycosyltransferase subunit DAD1	17.5	20.7	22.8	17.5	17.5	21.7	17.5	17.5	22.3	1	10.6	0	0.278	1.441	0.222	1.259
Ddx21	Nucleolar RNA helicase 2	18.7	24.3	23.0	24.6	23.4	25.3	17.5	22.8	24.1	11	17.7	0	0.618	-2.448	0.071	0.536
Ddx39b	Spliceosome RNA helicase Ddx39b	19.6	22.6	17.5	22.8	21.6	17.5	21.3	21.5	24.1	3	3.3	0	0.127	-0.756	0.626	-2.409
Ddx3x; D1Pas1; Ddx3y	ATP-dependent RNA helicase DDX3X; Putative ATP-dependent RNA helicase Pl10	17.5	20.8	21.4	24.9	24.7	25.7	17.5	17.5	24.5	7	15	0	1.860	-5.202	0.007	0.056
Ddx46	Probable ATP-dependent RNA helicase DDX46	25.1	24.8	24.2	24.9	24.8	24.7	25.1	25.4	24.5	2	2.1	0.0022989	0.137	-0.106	0.373	-0.346
Ddx5	Probable ATP-dependent RNA helicase DDX5	22.4	17.5	17.5	24.4	23.6	28.2	20.0	21.0	23.9	6	9.3	0	1.372	-6.290	0.553	-2.506
Ddx6	Probable ATP-dependent RNA helicase DDX6	17.5	17.5	17.5	21.7	17.5	21.8	17.5	17.5	22.6	1	3.9	0.0038388	0.935	-2.830	0.427	-1.687
Degs1	Sphingolipid delta(4)-desaturase DES1	21.8	22.2	22.1	22.6	17.5	17.5	17.5	17.5	23.4	2	10.5	0	0.753	2.804	0.584	2.559
Dek	Protein DEK	21.3	21.6	17.5	24.3	17.5	21.5	17.5	20.5	22.8	6	13.7	0	0.153	-0.974	0.022	-0.133
Dhx15	Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	17.5	17.5	17.5	17.5	17.5	17.5	17.5	21.5	20.9	2	3.3	0.0043197	0.000	0.000	0.925	-2.443
Dhx9	ATP-dependent RNA helicase A	17.5	17.5	22.8	17.5	17.5	22.4	17.5	17.5	21.3	1	1.2	0	0.019	0.136	0.080	0.494
Dnase1	Deoxyribonuclease-1	20.6	17.5	17.5	21.9	17.5	20.3	17.5	21.4	21.6	1	3.2	0.0038911	0.340	-1.355	0.406	-1.611
Dpysl2; Dpysl3; Crmp1	Dihydropyrimidinase-related protein 2; Dihydropyrimidinase-related protein 3	17.5	20.7	22.6	17.5	21.2	20.9	17.5	20.7	23.4	2	7.5	0.0056075	0.077	0.416	0.041	-0.273
Dsg1b; Dsg1a	Desmoglein-1-beta; Desmoglein-1-alpha	26.1	25.4	27.7	23.1	23.8	21.4	17.5	24.3	24.4	4	7.3	0	1.652	3.632	0.846	4.336
Dsp	Desmoplakin	29.7	30.6	32.5	28.7	28.9	26.2	26.4	29.4	28.2	83	23.1	0	1.158	2.981	1.141	2.921
Dstn	Destrin	17.5	21.5	22.7	22.4	17.5	17.5	17.5	17.5	23.1	2	18.2	0	0.242	1.396	0.186	1.182
Dync1h1	Cytoplasmic dynein 1 heavy chain 1	19.9	20.6	19.6	22.5	17.5	21.7	17.5	17.5	17.5	4	1.6	0	0.118	-0.510	3.033	2.549

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Dync1i2	Cytoplasmic dynein 1 intermediate chain 2	17.5	20.6	26.6	17.5	17.5	17.5	17.5	17.5	17.5	2	4.9	0	0.696	4.063	0.696	4.063
Ebp	3-beta-hydroxysteroid-Delta(8), Delta(7)-isomerase	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	20.9	1	5.7	0.004008	0.000	0.000	0.427	-1.121
Eef1a1	Elongation factor 1-alpha 1	26.9	28.3	28.8	30.5	28.9	30.7	27.8	28.8	29.0	19	35.5	0	1.168	-2.027	0.327	-0.541
Eef1b; Eef1b2	Elongation factor 1-beta	17.5	18.1	22.5	21.6	17.5	17.5	17.5	17.5	23.0	1	6.7	0.0041152	0.084	0.494	0.004	0.028
Eef1d	Elongation factor 1-delta	17.5	21.4	20.7	17.5	17.5	17.5	17.5	23.1	23.7	2	12.8	0	0.921	2.349	0.278	-1.599
Eef1g	Elongation factor 1-gamma	23.0	24.5	25.1	26.0	24.9	24.6	22.9	24.5	25.9	8	18.5	0	0.575	-0.982	0.073	-0.220
Eef2	Elongation factor 2	24.3	26.5	26.7	27.6	27.0	25.5	25.3	25.9	27.6	17	21.7	0	0.381	-0.878	0.167	-0.447
Eif2s1	Eukaryotic translation initiation factor 2 subunit 1	20.7	21.7	17.5	22.1	22.7	23.8	17.5	22.5	23.4	2	13	0	1.003	-2.904	0.203	-1.169
Eif2s2	Eukaryotic translation initiation factor 2 subunit 2	17.5	17.5	20.3	17.5	17.5	22.5	17.5	17.5	19.0	2	7.6	0.0043573	0.146	-0.752	0.147	0.418
Eif2s3y; Eif2s3x	Eukaryotic translation initiation factor 2 subunit 3, Y-linked; Eukaryotic translation initiation factor 2 subunit 3, X-linked	17.5	17.5	22.1	22.7	17.5	20.9	17.5	17.5	21.5	3	11.7	0	0.243	-1.326	0.031	0.185
Eif3a	Eukaryotic translation initiation factor 3 subunit A	21.4	22.6	22.9	22.9	21.0	22.8	17.5	20.4	24.3	5	4.2	0	0.029	0.067	0.310	1.542
Eif3b	Eukaryotic translation initiation factor 3 subunit B	17.5	17.5	17.5	23.2	24.3	24.1	21.4	21.1	25.0	3	6.4	0	4.385	-6.370	1.766	-4.988
Eif3c	Eukaryotic translation initiation factor 3 subunit C	23.5	24.2	23.7	24.5	20.4	22.9	17.5	22.5	23.6	7	9	0	0.425	1.193	0.628	2.616
Eif3d	Eukaryotic translation initiation factor 3 subunit D	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	1	4.6	0	0.000	0.000	0.000	0.000
Eif3l	Eukaryotic translation initiation factor 3 subunit L	17.5	17.5	17.5	21.1	17.5	22.0	20.4	17.5	17.5	1	2.8	0.004386	0.918	-2.691	0.427	-0.966
Eif4a1; Eif4a2	Eukaryotic initiation factor 4A-I; Eukaryotic initiation factor 4A-II	24.0	25.9	25.5	25.5	24.6	24.8	17.5	23.7	26.9	8	29.3	0	0.098	0.174	0.364	2.447
Eif4b	Eukaryotic translation initiation factor 4B	21.5	22.2	23.6	22.1	17.5	17.5	17.5	22.5	22.3	1	2.6	0.0041929	0.948	3.396	0.403	1.673
Eif4g1	Eukaryotic translation initiation factor 4 gamma 1	22.7	23.4	22.4	25.5	22.5	25.3	17.5	24.2	25.4	8	8.1	0	0.699	-1.566	0.070	0.496

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Eif4h	Eukaryotic translation initiation factor 4H	17.5	20.3	22.5	22.7	17.5	22.0	17.5	23.6	24.4	2	13.3	0	0.104	-0.626	0.265	-1.733
Eif5a; Eif5a2	Eukaryotic translation initiation factor 5A-1; Eukaryotic translation initiation factor 5A-2	17.5	20.7	23.3	22.6	22.5	25.5	17.5	23.1	24.6	1	7.8	0	0.723	-3.062	0.171	-1.235
Eif5b	Eukaryotic translation initiation factor 5B	20.7	17.5	17.5	17.5	20.8	17.5	17.5	17.5	24.1	2	3.1	0	0.011	-0.049	0.177	-1.143
Eif6	Eukaryotic translation initiation factor 6	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	22.2	1	7.3	0	0.000	0.000	0.427	-1.553
Elavl1	ELAV-like protein 1	19.4	19.9	20.9	20.3	17.5	22.1	17.5	20.7	21.2	1	3.4	0.0038241	0.021	0.090	0.086	0.300
Eno1	Alpha-enolase;Enolase	25.9	27.3	26.8	27.8	26.9	27.0	25.6	27.7	28.5	20	55.8	0	0.476	-0.555	0.240	-0.585
Eppk1	Epiplakin	17.5	17.5	17.5	20.5	17.5	17.5	17.5	17.5	17.5	1	2.6	0	0.427	-1.009	0.000	0.000
Eprs	Bifunctional glutamate/proline–tRNA ligase; Glutamate–tRNA ligase	21.1	17.5	17.5	17.5	17.5	24.1	17.5	22.6	23.6	2	2.6	0	0.147	-0.995	0.494	-2.542
Erh	Enhancer of rudimentary homolog	17.5	21.3	22.5	21.9	17.5	19.9	17.5	17.5	21.0	1	10.6	0.003861	0.124	0.666	0.392	1.765
Esd	S-formylglutathione hydrolase	21.3	22.3	23.2	23.1	21.4	22.3	17.5	20.6	24.0	1	7.8	0	0.014	-0.030	0.321	1.536
Fasn	Fatty acid synthase; [Acyl-carrier-protein] S-acetyltransferase	19.9	24.2	23.6	24.4	17.5	22.8	20.4	22.4	25.1	9	4.8	0	0.151	1.006	0.016	-0.092
Flna	Filamin-A	26.6	28.2	28.2	29.4	28.4	29.3	25.7	27.7	29.6	62	36.2	0	1.036	-1.334	0.008	0.030
Flnb	Filamin-B	22.5	24.9	24.6	26.9	24.9	27.8	18.5	24.4	27.1	21	13.6	0	1.041	-2.502	0.096	0.706
Flnc	Filamin-C	23.3	25.3	24.5	26.1	25.0	26.4	22.6	22.7	26.8	26	13.2	0	0.964	-1.486	0.079	0.335
Fscn1	Fascin	22.2	21.9	22.5	24.4	23.0	22.5	21.8	22.4	24.5	4	12.6	0	0.781	-1.044	0.322	-0.652
Fxyd5	FXYD domain-containing ion transport regulator 5	17.5	24.5	26.1	17.5	23.9	21.9	17.5	23.7	22.8	1	5.1	0.003876	0.189	1.603	0.160	1.394
Gapdh; Gm7293	Glyceraldehyde-3-phosphate dehydrogenase	25.8	26.8	28.0	27.9	28.2	28.1	25.9	26.5	28.4	15	44.7	0	0.865	-1.191	0.029	-0.086
Gbp2	Interferon-induced guanylate-binding protein 2	17.5	17.5	17.5	17.5	23.8	23.8	17.5	23.0	23.0	2	3.7	0	0.935	-4.181	0.935	-3.664
Gdi2	Rab GDP dissociation inhibitor beta	17.5	21.2	17.5	22.5	17.5	17.5	20.1	17.5	22.1	2	5.8	0	0.076	-0.445	0.258	-1.172

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Gfpt1	Glutamine–fructose-6-phosphate aminotransferase [isomerizing] 1	17.5	17.5	17.5	21.8	17.5	17.5	17.5	17.5	20.6	1	2.3	0	0.427	-1.447	0.427	-1.029
Glud1	Glutamate dehydrogenase 1, mitochondrial	19.9	17.5	17.5	22.3	17.5	19.6	17.5	17.5	21.3	1	3.8	0	0.397	-1.501	0.117	-0.484
Gm5414		23.1	27.2	30.3	25.4	26.9	25.0	23.3	26.8	24.2	9	0	0	0.203	1.140	0.378	2.106
Gnai2; Gnao1; Gnas	Guanine nucleotide-binding protein G(i) subunit alpha-2; Guanine nucleotide-binding protein G(o) subunit alpha	19.0	22.3	17.5	22.8	17.5	21.3	17.5	17.5	21.1	2	7.3	0.0022173	0.168	-0.945	0.190	0.915
Gnb1	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	17.5	19.3	17.5	22.9	17.5	17.5	17.5	17.5	22.8	1	5.3	0.0022371	0.249	-1.187	0.247	-1.162
Gnb2; Gnb4	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2; Guanine nucleotide-binding protein subunit beta-4	17.5	22.7	17.5	22.5	17.5	21.0	17.5	17.5	17.5	1	4.1	0	0.186	-1.110	0.427	1.728
Gnb2l1	Guanine nucleotide-binding protein subunit beta-2-like 1	22.4	23.6	25.3	24.3	24.3	25.3	21.7	23.9	25.5	5	15.8	0	0.425	-0.886	0.013	0.052
Golgb1		17.5	17.5	28.2	24.4	24.2	22.1	17.5	24.6	23.2	1	0.2	0	0.275	-2.504	0.057	-0.695
Got2	Aspartate aminotransferase, mitochondrial	17.5	22.1	21.0	22.9	17.5	20.8	17.5	17.5	23.1	3	9.5	0	0.031	-0.190	0.133	0.840
Gpi	Glucose-6-phosphate isomerase	19.7	20.2	17.5	21.2	17.5	17.5	17.5	17.5	21.1	2	7.3	0.0022472	0.090	0.377	0.107	0.431
Gsdma2; Gsdma	Gasdermin-A2; Gasdermin-A	18.8	17.5	17.5	23.9	17.5	17.5	17.5	17.5	17.5	1	3.2	0	0.316	-1.670	0.427	0.446
Gstp1; Gstp2	Glutathione S-transferase P 1; Glutathione S-transferase P 2	19.4	23.9	20.8	24.4	23.2	23.7	20.9	17.5	24.4	3	22.4	0	0.813	-2.386	0.062	0.428
Gtf2f1	General transcription factor IIF subunit 1	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	1	3.1	0.0022422	0.000	0.000	0.000	0.000
Gvin1	Interferon-induced very large GTPase 1	17.5	25.3	26.6	17.5	21.5	20.9	17.5	17.5	23.7	3	2.1	0	0.439	3.183	0.433	3.569
H2-D1	H-2 class I histocompatibility antigen, D-B alpha chain	17.5	22.7	19.1	23.3	17.5	17.5	17.5	17.5	23.4	3	11.6	0	0.051	0.364	0.044	0.324

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Hadha	Trifunctional enzyme subunit alpha, mitochondrial; Long-chain enoyl-CoA hydratase	17.5	20.7	23.0	17.5	17.5	17.5	17.5	17.5	17.5	2	3.7	0	0.845	2.911	0.845	2.911
Hal	Histidine ammonia-lyase	17.5	17.5	23.2	21.4	17.5	17.5	17.5	17.5	17.5	3	5.6	0	0.094	0.605	0.427	1.903
Hbb-bh1; Hbb-bh0	Hemoglobin subunit beta-H1; Hemoglobin subunit beta-H0	19.7	21.8	17.5	22.1	17.5	17.5	23.6	17.5	17.5	1	6.8	0.0022124	0.115	0.629	0.021	0.153
Hdlbp	Vigilin	17.5	23.2	22.3	23.1	21.7	24.4	17.5	17.5	23.5	6	7.2	0	0.465	-2.071	0.218	1.497
Hint1	Histidine triad nucleotide-binding protein 1	20.2	21.5	22.8	22.1	17.5	17.5	20.9	22.9	22.1	1	11.1	0.0041408	0.646	2.454	0.186	-0.469
Hist1h1a	Histone H1.1	24.0	17.5	21.8	17.5	23.4	22.6	17.5	23.2	23.2	5	17.4	0	0.007	-0.059	0.024	-0.195
Hist1h1b	Histone H1.5	25.5	25.1	26.4	25.7	23.7	24.3	23.2	25.1	25.6	9	24.2	0	0.734	1.099	0.565	1.031
Hist1h1c	Histone H1.2	17.5	17.5	17.5	17.5	22.6	22.7	17.5	22.7	23.1	9	9.4	0	0.935	-3.447	0.933	-3.586
Hist1h1d	Histone H1.3	26.2	24.0	24.5	25.4	25.1	25.1	24.6	25.2	24.6	12	9.5	0	0.165	-0.289	0.042	0.085
Hist1h1e	Histone H1.4	28.1	25.0	28.1	27.4	27.0	25.9	25.7	27.2	27.3	13	19.6	0	0.105	0.336	0.105	0.340
Hist1h2ab; Hist1h2af; Hist1h2ak	Histone H2A type 1;Histone H2A type 1-F	25.0	26.0	25.4	26.3	26.7	25.9	24.0	26.0	27.6	5	13.8	0	1.082	-0.857	0.149	-0.433
Hist1h2bf; Hist1h2bc; Hist1h2bh	Histone H2B type 1-F/J/L;Histone H2B type 1-C/E/G	26.5	27.1	27.4	28.5	27.7	28.2	26.0	26.7	28.9	7	43.7	0	1.569	-1.170	0.098	-0.248
Hist1h3a; H3f3a; Hist1h3b	Histone H3.1; Histone H3.3	25.5	27.7	28.8	27.9	26.6	27.4	26.2	26.7	27.2	4	19.9	0	0.021	0.064	0.247	0.642
Hist1h4a	Histone H4	25.2	27.5	28.4	27.5	26.5	27.0	26.1	27.8	28.6	8	56.3	0	0.009	0.026	0.138	-0.450
Hmga1	High mobility group protein HMG-I/HMG-Y	17.5	17.5	26.1	24.7	17.5	20.0	22.7	25.0	24.6	1	14.6	0	0.033	-0.348	0.555	-3.706
Hmga1	High mobility group protein HMG-I/HMG-Y	17.5	23.9	24.8	21.9	17.5	23.8	17.5	25.2	24.5	2	23.4	0	0.122	0.995	0.031	-0.314
Hmga2	High mobility group protein HMGI-C	23.4	24.2	17.5	23.5	17.5	23.8	17.5	17.5	17.5	1	21.3	0	0.009	0.081	0.929	4.185
Hmgb1	High mobility group protein B1	24.1	21.4	25.0	25.8	20.9	25.7	17.5	24.4	25.9	9	39.1	0	0.116	-0.619	0.114	0.885
Hmgn1	Non-histone chromosomal protein HMG-14	17.5	17.5	17.5	17.5	17.5	17.5	17.5	22.3	23.0	2	36.5	0.0056285	0.000	0.000	0.929	-3.423
Hmox2	Heme oxygenase 2	17.5	17.5	17.5	22.3	17.5	24.1	17.5	17.5	22.4	2	14.3	0	0.900	-3.779	0.427	-1.646

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Hn1l	Hematological and neurological expressed 1-like protein	17.5	17.5	17.5	17.5	17.5	21.6	17.5	17.5	21.4	2	18.9	0.0038314	0.427	-1.364	0.427	-1.305
Hnrnpa1	Heterogeneous nuclear ribonucleoprotein A1	22.5	22.7	24.0	23.2	23.5	25.6	17.5	24.8	25.8	5	24.1	0	0.489	-1.001	0.050	0.386
Hnrnpa2b1	Heterogeneous nuclear ribonucleoproteins A2/B1	23.1	23.7	23.8	25.1	23.0	23.2	22.1	23.6	23.0	2	9.9	0	0.104	-0.211	0.612	0.667
Hnrnpa3; Gm6793; Gm9242	Heterogeneous nuclear ribonucleoprotein A3	22.5	24.1	23.3	24.4	22.6	26.4	21.1	23.9	25.2	7	29.6	0	0.413	-1.181	0.030	-0.114
Hnrnpab	Heterogeneous nuclear ribonucleoprotein A/B	24.0	17.5	17.5	17.5	17.5	23.6	17.5	23.8	22.9	1	4.6	0	0.015	0.138	0.229	-1.711
Hnrnpc	Heterogeneous nuclear ribonucleoproteins C1/C2	23.0	20.8	17.5	22.1	22.2	22.9	22.1	17.5	23.3	2	7.3	0	0.547	-1.990	0.081	-0.544
Hnrnpd	Heterogeneous nuclear ribonucleoprotein D0	21.6	19.6	20.7	20.0	22.2	23.3	19.1	22.2	23.9	2	6.8	0	0.469	-1.235	0.299	-1.118
Hnrnph1; Hnrnph2	Heterogeneous nuclear ribonucleoprotein H; Heterogeneous nuclear ribonucleoprotein H, N-terminally processed	21.4	22.6	22.2	23.4	21.0	23.9	17.5	23.1	23.9	2	3.8	0	0.278	-0.659	0.103	0.585
Hnrnpk	Heterogeneous nuclear ribonucleoprotein K	21.6	24.9	23.8	24.8	24.5	25.4	23.3	23.8	25.1	5	18.8	0	0.654	-1.447	0.213	-0.607
Hnrnpl	Heterogeneous nuclear ribonucleoprotein L	17.5	17.5	17.5	21.3	17.5	21.3	17.5	17.5	23.1	2	6.8	0.005597	0.935	-2.513	0.427	-1.852
Hnrnpm	Heterogeneous nuclear ribonucleoprotein M	17.5	21.8	17.5	17.5	17.5	17.5	17.5	17.5	17.5	1	2.2	0	0.427	1.445	0.427	1.445
Hnrnpu	Heterogeneous nuclear ribonucleoprotein U	24.6	26.2	25.6	26.8	24.9	26.6	23.4	26.4	27.5	12	21.5	0	0.352	-0.639	0.071	-0.268
Hp1bp3	Heterochromatin protein 1-binding protein 3	17.5	17.5	17.5	17.5	17.5	22.9	17.5	17.5	17.5	1	2.7	0.0040816	0.427	-1.810	0.000	0.000
Hsp90aa1	Heat shock protein HSP 90-alpha	21.8	23.3	25.1	26.1	24.2	23.7	21.7	25.6	27.1	21	15.8	0	0.431	-1.219	0.303	-1.383
Hsp90ab1	Heat shock protein HSP 90-beta	26.0	28.4	28.1	28.4	27.6	27.3	25.3	27.8	28.9	29	23.5	0	0.102	-0.234	0.049	0.187
Hsp90b1	Endoplasmin	22.6	24.1	22.7	25.3	22.8	23.5	21.9	23.1	25.8	10	16	0	0.330	-0.715	0.135	-0.458

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Hspa1a; Hspa1b	Heat shock 70 kDa protein 1A; Heat shock 70 kDa protein 1B	19.9	22.4	21.3	22.4	17.5	17.5	17.5	17.5	21.9	6	2.5	0	0.514	2.084	0.611	2.235
Hspa4	Heat shock 70 kDa protein 4	17.5	21.5	20.2	23.2	17.5	21.5	17.5	20.9	23.4	5	8.4	0	0.186	-0.997	0.163	-0.900
Hspa5	78 kDa glucose-regulated protein	24.8	26.7	26.4	27.0	25.0	25.2	22.2	25.5	27.2	23	39.5	0	0.091	0.219	0.246	0.992
Hspa8	Heat shock cognate 71 kDa protein	27.1	28.6	29.2	29.1	28.5	28.1	26.5	28.2	29.4	25	34.8	0	0.146	-0.266	0.085	0.248
Hspa9	Stress-70 protein, mitochondrial	17.5	23.9	21.2	25.0	23.5	25.7	20.2	23.7	25.3	8	18	0	0.926	-3.871	0.400	-2.255
Hspd1	60 kDa heat shock protein, mitochondrial	22.1	23.5	23.7	24.6	23.9	23.9	20.8	22.3	25.4	7	19.4	0	0.898	-1.031	0.072	0.294
Hspe1	10 kDa heat shock protein, mitochondrial	17.5	17.5	17.5	17.5	17.5	17.5	17.5	21.6	21.2	1	13.7	0.0055659	0.000	0.000	0.931	-2.583
Ifi205a	Interferon-activable protein 205-A	17.5	17.5	22.8	21.7	17.5	22.0	17.5	22.8	23.5	2	6.9	0	0.192	-1.136	0.320	-2.015
Ifitm3	Interferon-induced transmembrane protein 3	20.2	23.8	26.0	24.6	23.4	21.1	17.5	17.5	22.6	4	48.2	0	0.051	0.296	0.803	4.147
Igtp		22.3	25.2	25.1	25.1	23.7	23.4	20.2	22.8	25.4	7	22.7	0	0.043	0.136	0.316	1.370
Iigp1	Interferon-inducible GTPase 1	21.5	24.4	24.6	25.4	24.3	25.8	21.4	23.5	25.5	7	24	0	0.709	-1.686	0.001	0.003
Impdh2	Inosine-5-monophosphate dehydrogenase 2	17.5	19.8	17.5	22.4	17.5	22.1	17.5	17.5	23.1	3	7.2	0	0.613	-2.403	0.212	-1.102
Ipo5	Importin-5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	1	1.5	0	0.000	0.000	0.000	0.000
Ipo9	Importin-9	17.5	17.5	22.4	17.5	17.5	22.0	17.5	17.5	17.5	1	2	0	0.020	0.133	0.427	1.618
Iqgap1	Ras GTPase-activating-like protein IQGAP1	20.6	19.1	17.5	17.5	17.5	23.9	17.5	17.5	20.3	2	2.1	0	0.088	-0.572	0.184	0.620
Irgm1	Immunity-related GTPase family M protein 1	17.5	23.3	23.4	21.4	22.3	23.6	17.5	21.7	19.9	3	9	0	0.188	-1.013	0.302	1.710
Itgb1	Integrin beta-1	21.4	21.8	22.1	22.5	17.5	23.0	17.5	17.5	22.6	2	4.6	0	0.161	0.757	0.676	2.557
Jup	Junction plakoglobin	29.0	30.1	32.2	28.4	28.9	26.9	26.5	29.4	27.8	29	42.4	0	0.993	2.356	0.928	2.504
Kars	Lysine–tRNA ligase	20.2	21.6	22.0	21.7	22.3	21.9	21.0	21.5	21.7	1	2.5	0	0.531	-0.703	0.069	-0.116
Khsrp	Far upstream element-binding protein 2	22.3	23.8	24.1	22.6	23.1	24.4	20.8	22.9	24.5	6	11	0	0.026	0.060	0.231	0.714
Kpna2	Importin subunit alpha-1	20.5	20.6	21.2	20.3	22.1	22.7	17.5	17.5	22.0	3	8.3	0	0.553	-0.931	0.516	1.763
Kpnb1	Importin subunit beta-1	21.1	22.9	21.4	24.2	22.8	23.7	19.8	23.0	24.1	4	6.1	0	1.216	-1.770	0.127	-0.480
Krt27	Keratin, type I cytoskeletal 27	17.5	20.7	17.5	22.1	17.5	17.5	17.5	17.5	17.5	2	2	1	0.091	-0.475	0.427	1.063

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Krt76	Keratin, type II cytoskeletal 2 oral	31.0	31.6	32.6	30.7	31.4	29.6	23.2	31.2	30.1	17	1.9	0	0.757	1.152	0.633	3.581
Lap3	Cytosol aminopeptidase	17.5	17.5	17.5	21.7	17.5	21.6	17.5	17.5	21.9	1	2.3	0.0042644	0.935	-2.774	0.427	-1.466
Lasp1	LIM and SH3 domain protein 1	25.4	22.8	22.9	22.9	17.5	17.5	17.5	21.6	24.6	6	27	0	1.032	4.394	0.484	2.481
Ldha	L-lactate dehydrogenase A chain; L-lactate dehydrogenase	24.0	25.4	25.3	26.5	25.7	24.9	25.9	24.7	26.5	9	23.5	0	0.544	-0.784	0.522	-0.796
Lgals1	Galectin-1	20.1	22.7	22.6	22.1	17.5	23.4	17.5	17.5	23.4	1	13.3	0	0.152	0.802	0.470	2.322
Lgals3	Galectin-3	17.5	21.9	23.4	21.8	17.5	17.5	17.5	21.7	21.4	1	4.2	0.0040733	0.367	2.004	0.120	0.734
Lmna	Prelamin-A/C;Lamin-A/C	26.0	26.0	26.6	26.0	25.7	24.5	17.5	26.4	26.4	19	34.7	0	0.684	0.775	0.394	2.770
Lrrc59	Leucine-rich repeat-containing protein 59	17.5	21.5	25.8	23.5	20.1	20.0	17.5	19.4	20.9	3	12.4	0.0023202	0.049	0.378	0.377	2.334
Lrrfip1	Leucine-rich repeat flightless-interacting protein 1	17.5	17.5	17.5	21.2	17.5	17.5	17.5	21.1	21.8	1	4.3	0.0041667	0.427	-1.232	0.924	-2.631
Luc7l2	Putative RNA-binding protein Luc7-like 2	21.1	21.9	22.6	23.4	21.1	21.8	17.5	21.9	24.7	5	15.6	0	0.091	-0.203	0.088	0.525
Luc7l3	Luc7-like protein 3	20.7	17.5	23.5	23.6	17.5	20.5	17.5	22.1	23.2	5	13	0	0.005	0.040	0.053	-0.380
Lyz1;Lyz2	Lysozyme C-1;Lysozyme C-2	22.1	23.9	25.1	24.0	17.5	17.5	20.4	24.3	24.6	2	14.2	0	0.797	4.051	0.148	0.641
Map4	Microtubule-associated protein 4; Microtubule-associated protein	17.5	17.5	17.5	21.5	17.5	17.5	17.5	20.9	21.7	2	2.3	0.0041322	0.427	-1.329	0.918	-2.547
Marcks	Myristoylated alanine-rich C-kinase substrate	17.5	17.5	17.5	17.5	17.5	17.5	17.5	23.3	23.9	3	28.5	0	0.000	0.000	0.932	-4.070
Marcksl1	MARCKS-related protein	21.8	20.6	17.5	22.1	17.5	21.8	20.2	22.8	23.0	2	27.5	0	0.089	-0.492	0.575	-2.025
Mars	Methionine–tRNA ligase, cytoplasmic	17.5	17.5	17.5	21.7	17.5	21.8	17.5	17.5	20.7	3	5.1	0	0.935	-2.827	0.427	-1.052
Mat2a	S-adenosylmethionine synthase isoform type-2	17.5	17.5	21.2	17.5	17.5	17.5	17.5	21.4	22.5	2	7.6	0	0.427	1.231	0.368	-1.711
Matr3	Matrin-3	17.5	17.5	17.5	21.2	21.2	20.7	17.5	20.9	22.5	2	3.9	0.0056604	4.418	-3.529	0.885	-2.822
Mccc1	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial	18.8	28.4	29.0	27.7	28.3	28.1	23.3	18.8	23.2	27	45.9	0	0.326	-2.610	0.429	3.632
Mcm3	DNA replication licensing factor MCM3	17.5	17.5	20.8	17.5	17.5	21.3	17.5	17.5	21.8	2	3.9	0	0.033	-0.161	0.064	-0.329

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Mcm4	DNA replication licensing factor MCM4	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	22.5	1	3.1	0.0038835	0.000	0.000	0.427	-1.680
Mdh2	Malate dehydrogenase, mitochondrial	22.0	23.2	24.8	25.6	21.6	25.3	23.7	22.7	26.1	5	20.4	0	0.217	-0.835	0.257	-0.827
Mif	Macrophage migration inhibitory factor	21.2	18.8	23.3	23.6	17.5	21.9	20.2	24.1	23.2	2	15.7	0	0.011	0.075	0.335	-1.421
Msn	Moesin	23.8	24.4	24.4	25.8	23.4	25.7	18.2	21.7	27.4	19	31.7	0	0.383	-0.738	0.270	1.803
Mtap	S-methyl-5-thioadenosine phosphorylase	17.5	18.5	20.2	17.5	20.7	17.5	17.5	17.5	22.2	2	12.4	0	0.048	0.185	0.063	-0.317
Mthfd1	C-1-tetrahydrofolate synthase, cytoplasmic; Methylenetetrahydrofolate dehydrogenase	17.5	20.6	17.5	21.4	21.1	21.9	17.5	17.5	20.0	3	3.6	0	1.292	-2.921	0.055	0.211
Mtpn	Myotrophin	21.7	22.2	22.8	22.9	17.5	17.5	17.5	21.0	23.9	2	30.5	0	0.737	2.950	0.321	1.466
Myadm	Myeloid-associated differentiation marker	21.0	23.0	23.3	23.0	22.9	22.8	20.6	21.8	23.0	1	5.3	0	0.240	-0.446	0.251	0.638
Mybbp1a	Myb-binding protein 1A	21.7	23.4	24.2	24.2	24.2	26.0	22.4	23.7	24.0	12	10.6	0	0.816	-1.700	0.115	-0.283
Myh14	Myosin-14	17.5	24.8	25.8	24.6	17.5	24.1	21.5	23.0	24.4	4	0.6	0.004329	0.063	0.629	0.039	-0.313
Myh7; Myh6; Myh4	Myosin-7/6/4/1/8/3	17.5	17.5	17.5	17.5	17.5	17.5	26.0	17.5	17.5	5	2.9	0	0.000	0.000	0.427	-2.836
Myh9	Myosin-9	28.3	29.1	29.5	30.2	29.1	30.0	27.2	28.8	30.3	87	45.9	0	0.818	-0.828	0.072	0.197
Myl12b; Myl12a	Myosin regulatory light chain 12B	21.3	17.5	23.7	22.4	22.2	23.9	17.5	22.7	25.7	3	23.8	0	0.464	-2.011	0.141	-1.139
Myl6	Myosin light polypeptide 6	17.5	20.4	19.3	24.1	17.5	17.5	17.5	22.4	24.8	3	24.5	0	0.094	-0.620	0.464	-2.487
Myo1c	Unconventional myosin-Ic	17.5	17.5	19.3	23.2	17.5	21.3	17.5	23.9	23.6	2	3	0	0.653	-2.559	0.756	-3.577
Naca	Nascent polypeptide-associated complex subunit alpha, muscle-specific form; Nascent polypeptide-associated complex subunit alpha	23.3	23.9	24.7	23.7	22.0	23.7	22.1	23.4	25.4	3	1.9	0	0.510	0.825	0.110	0.310
Nap1l1	Nucleosome assembly protein 1-like 1	17.5	21.8	23.2	22.9	17.5	22.6	17.5	23.6	24.7	1	4.3	0	0.023	-0.166	0.146	-1.106
Nccrp1	F-box only protein 50	23.3	22.6	24.3	17.5	17.5	23.3	17.5	17.5	17.5	1	4.5	0	0.929	3.966	3.589	5.901
Ncl	Nucleolin	27.5	27.3	28.3	28.5	27.0	27.0	26.0	28.5	28.3	29	37.5	0	0.127	0.203	0.037	0.096
Nes	Nestin	17.5	17.5	17.5	22.1	17.5	17.5	17.5	17.5	22.5	2	2	0	0.427	-1.541	0.427	-1.652

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Nkap	NF-kappa-B-activating protein	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	1	3.9	0.0040323	0.000	0.000	0.000	0.000
Nme1	Nucleoside diphosphate kinase A	17.5	17.5	17.5	22.3	17.5	17.5	17.5	21.2	22.1	3	6.6	1	0.427	-1.585	0.919	-2.768
Nme2; Gm20390	Nucleoside diphosphate kinase B; Nucleoside diphosphate kinase	22.0	24.2	25.3	25.3	22.3	22.1	20.9	24.6	23.3	3	9.2	0	0.153	0.585	0.235	0.873
Nmt1	Glycylpeptide N-tetradecanoyltransferase 1	17.5	17.5	17.5	17.5	17.5	21.2	17.5	17.5	21.7	1	2.6	0.0092081	0.427	-1.241	0.427	-1.410
Nolc1		25.2	24.8	25.9	25.6	23.7	24.3	22.3	26.1	26.8	16	17.1	0	0.555	0.795	0.061	0.249
Nop58	Nucleolar protein 58	17.5	21.2	17.5	17.5	17.5	22.3	17.5	17.5	17.5	2	5.4	0	0.067	-0.388	0.427	1.225
Npm1; Gm5611	Nucleophosmin	26.2	26.7	25.5	26.7	26.9	26.8	24.6	26.1	27.3	12	35.3	0	0.852	-0.636	0.066	0.161
Npm3	Nucleoplasmin-3	22.9	23.7	17.5	17.5	17.5	17.5	17.5	17.5	24.4	1	8.6	0.0039761	0.928	3.893	0.202	1.586
Ntf3	Neurotrophin-3	17.5	22.1	17.5	17.5	22.2	21.6	21.8	17.5	21.5	1	4.7	0.0038023	0.257	-1.381	0.231	-1.220
Nucks1	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	17.5	21.9	17.5	21.8	21.6	20.2	17.5	22.2	22.7	2	16.2	0.0022321	0.646	-2.233	0.341	-1.836
P4hb	Protein disulfide-isomerase	21.4	23.9	17.5	23.8	23.3	21.7	17.5	21.6	24.1	2	4.9	0	0.445	-2.039	0.020	-0.162
Pa2g4	Proliferation-associated protein 2G4	21.1	23.8	23.6	23.5	20.6	22.8	18.1	20.1	24.8	6	18.3	0	0.165	0.545	0.350	1.844
Pabpc1; Pabpc6	Polyadenylate-binding protein 1	21.5	24.1	23.6	24.4	22.8	22.3	20.5	21.6	25.2	8	14.8	0	0.045	-0.134	0.141	0.610
Paics	Multifunctional protein ADE2; Phosphoribosylaminoimidazole-succinocarboxamide synthase	17.5	22.3	17.5	17.5	22.3	22.9	17.5	17.5	17.5	3	10.4	0	0.316	-1.805	0.427	1.596
Pc;Pcx	Pyruvate carboxylase, mitochondrial; Pyruvate carboxylase	19.4	28.4	28.4	27.8	28.4	28.8	23.1	17.5	24.1	29	29.5	0	0.415	-2.920	0.453	3.810
Pcbp1	Poly(rC)-binding protein 1	20.5	22.6	22.1	22.8	20.9	24.0	17.5	18.9	24.2	4	7.6	0	0.303	-0.811	0.292	1.539
Pcbp2; Pcbp3	Poly(rC)-binding protein 2; Poly(rC)-binding protein 3	17.5	17.5	17.5	20.6	21.9	23.3	17.5	17.5	20.4	4	6.9	0.0023256	2.341	-4.453	0.427	-0.979
Pcca	Propionyl-CoA carboxylase alpha chain, mitochondrial	20.3	28.4	28.1	27.1	27.9	27.6	23.8	17.5	22.0	26	40.5	0	0.289	-1.919	0.621	4.509

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Pcna	Proliferating cell nuclear antigen	17.5	20.8	17.5	17.5	17.5	17.5	17.5	17.5	23.5	2	12.3	0	0.427	1.108	0.146	-0.899
Pdap1	28 kDa heat- and acid-stable phosphoprotein	22.3	23.5	23.6	21.1	17.5	21.1	17.5	23.2	24.3	3	22.1	0	1.201	3.242	0.272	1.463
Pdcd5	Programmed cell death protein 5	21.9	21.0	22.4	23.5	17.5	17.5	17.5	23.5	25.4	3	27.8	0	0.480	2.261	0.052	-0.367
Pdia3	Protein disulfide-isomerase A3	24.3	23.8	23.6	26.0	23.7	24.9	19.9	23.7	26.1	11	27.3	0	0.614	-0.968	0.133	0.668
Pdia6	Protein disulfide-isomerase A6	19.8	21.3	17.5	21.9	17.5	21.7	17.5	17.5	21.7	1	3.9	0	0.175	-0.836	0.129	0.627
Pebp1	Phosphatidylethanolamine-binding protein 1; Hippocampal cholinergic neurostimulating peptide	17.5	17.5	17.5	22.2	17.5	17.5	17.5	17.5	23.2	1	13.9	0.0022523	0.427	-1.551	0.427	-1.899
Pgam1	Phosphoglycerate mutase 1	22.2	23.7	24.8	23.9	17.5	17.5	17.5	22.0	24.4	4	28	0	0.803	3.923	0.454	2.265
Pgd	6-phosphogluconate dehydrogenase, decarboxylating	17.5	23.2	17.5	23.0	17.5	22.7	17.5	17.5	23.6	2	6.4	0	0.249	-1.640	0.015	-0.126
Pgk1	Phosphoglycerate kinase 1	23.9	24.8	25.6	27.6	25.8	25.7	22.4	26.4	27.4	7	28.1	0	0.912	-1.569	0.148	-0.634
Phb	Prohibitin	17.5	17.5	21.9	22.9	17.5	22.6	23.4	21.1	23.4	3	13.6	0	0.371	-2.018	1.050	-3.669
Phgdh	D-3-phosphoglycerate dehydrogenase	22.5	24.0	23.1	24.0	23.1	23.9	23.9	17.5	24.0	5	12.2	0	0.392	-0.494	0.247	1.374
Pkm	Pyruvate kinase PKM	27.3	28.3	29.0	29.5	29.4	29.3	27.6	28.9	30.2	24	55.6	0	1.171	-1.209	0.329	-0.714
Pkp1	Plakophilin-1	24.7	25.9	28.2	24.5	25.0	22.9	23.7	25.3	23.9	14	24.9	0	0.799	2.106	0.772	1.944
Plec	Plectin	23.3	24.4	24.6	26.1	17.5	26.2	17.5	20.8	26.0	29	7.9	0	0.104	0.837	0.452	2.651
Plin2	Perilipin-2	17.5	20.7	23.3	21.4	17.5	22.0	17.5	17.5	20.9	2	7.3	0	0.029	0.190	0.385	1.861
Plp2	Proteolipid protein 2	22.2	17.5	17.5	22.2	22.1	21.8	17.5	17.5	22.1	1	7.9	0.0023148	0.872	-2.962	0.008	0.055
Pnn	Pinin	17.5	17.5	17.5	23.7	17.5	17.5	17.5	22.4	22.6	1	1.5	0	0.427	-2.060	0.935	-3.361
Pof1b	Protein POF1B	17.5	17.5	21.8	17.5	17.5	17.5	17.5	17.5	17.5	2	4.4	0.0043478	0.427	1.423	0.427	1.423
Ppa1	Inorganic pyrophosphatase	17.5	21.1	17.5	21.9	17.5	21.2	17.5	17.5	21.9	1	7.6	0	0.337	-1.494	0.043	-0.239
Ppia; Gm5160	Peptidyl-prolyl cis-trans isomerase A; Peptidyl-prolyl cis-trans isomerase A, N-terminally processed	17.5	20.6	22.9	23.6	22.8	22.0	20.6	17.5	25.9	3	17.7	0	0.705	-2.503	0.132	-1.047
Ppm1g	Protein phosphatase 1G	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	1	2.2	0	0.000	0.000	0.000	0.000
Ppp1ca	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit	22.2	22.6	17.5	22.9	17.5	19.3	17.5	20.8	23.8	2	7.6	0	0.140	0.863	0.010	0.076

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Ppp1cb	Serine/threonine-protein phosphatase PP1-beta catalytic subunit	17.5	17.5	17.5	17.5	22.5	22.3	17.5	17.5	22.3	1	4	0.0042735	0.934	-3.274	0.427	-1.594
Ppp2r1a; Ppp2r1b	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform;Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A beta isoform	20.7	17.5	24.3	23.4	20.9	24.1	21.0	17.5	24.1	2	5.4	0	0.376	-1.951	0.004	-0.032
Pqbp1	Polyglutamine-binding protein 1	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	1	9.5	0.0022831	0.000	0.000	0.000	0.000
Prdx1	Peroxiredoxin-1	24.2	23.7	27.0	23.4	24.7	19.4	21.2	25.8	25.9	4	20.1	0	0.562	2.394	0.128	0.645
Prdx2	Peroxiredoxin-2	17.5	23.5	25.1	23.1	23.1	21.0	20.7	23.2	22.8	3	24.2	0	0.051	-0.360	0.035	-0.250
Prdx4	Peroxiredoxin-4	18.1	23.1	22.3	23.0	21.4	17.5	17.5	21.4	23.9	4	16.1	0	0.093	0.584	0.039	0.277
Prdx6	Peroxiredoxin-6	17.5	22.1	17.5	22.6	17.5	22.1	17.5	17.5	22.9	1	9.4	0	0.314	-1.703	0.040	-0.280
Prep	Prolyl endopeptidase	17.5	17.5	17.5	22.7	17.5	17.5	17.5	17.5	22.9	2	4.5	0	0.427	-1.727	0.427	-1.798
Prim2	DNA primase large subunit	25.7	26.5	28.0	24.9	24.7	23.5	22.8	24.9	17.5	2	4.8	0.0039139	1.351	2.342	1.010	4.976
Prpf38a	Pre-mRNA-splicing factor 38A	20.3	17.5	17.5	17.5	17.5	20.7	17.5	17.5	22.2	1	3.8	0	0.031	-0.131	0.128	-0.638
Prpf40a	Pre-mRNA-processing factor 40 homolog A	17.5	17.5	17.5	20.6	17.5	17.5	17.5	17.5	17.5	1	1.7	0	0.427	-1.025	0.000	0.000
Psma1	Proteasome subunit alpha type-1	20.5	21.0	17.5	20.9	17.5	17.5	17.5	17.5	17.5	1	5.7	0	0.273	1.069	0.927	2.186
Psma3	Proteasome subunit alpha type-3	21.5	21.8	17.5	21.4	17.5	21.8	17.5	17.5	21.7	1	7.5	0	0.002	0.010	0.281	1.373
Psma6	Proteasome subunit alpha type-6	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	23.8	1	7.3	0	0.000	0.000	0.427	-2.090
Psma7; Psma8	Proteasome subunit alpha type-7; Proteasome subunit alpha type-7-like	17.5	22.5	23.7	23.4	22.1	22.5	17.5	22.4	22.1	2	14.9	0	0.298	-1.420	0.081	0.568
Psmb5	Proteasome subunit beta type-5	17.5	20.8	22.7	20.6	17.5	17.5	17.5	17.5	17.5	1	5.3	0	0.419	1.806	0.863	2.830
Psmc2	26S protease regulatory subunit 7	17.5	17.5	17.5	17.5	17.5	22.4	21.0	17.5	22.7	1	3	0	0.427	-1.640	0.885	-2.891
Psmc6	26S protease regulatory subunit 10B	19.8	21.3	19.0	21.5	17.5	17.5	17.5	19.3	21.5	2	5.9	0	0.324	1.186	0.175	0.615

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Psmd11	26S proteasome non-ATPase regulatory subunit 11	17.5	21.6	17.5	17.5	17.5	17.5	17.5	17.5	23.4	1	2.8	0.0073937	0.427	1.362	0.090	-0.602
Psmd12	26S proteasome non-ATPase regulatory subunit 12	19.3	20.6	17.5	21.4	17.5	17.5	17.5	17.5	22.1	1	4.6	0.002331	0.075	0.334	0.017	0.092
Psmd2; Gm5422	26S proteasome non-ATPase regulatory subunit 2	17.5	21.1	17.5	22.3	17.5	21.9	19.9	17.5	24.3	4	6.1	0	0.398	-1.843	0.325	-1.849
Psmd3	26S proteasome non-ATPase regulatory subunit 3	17.5	21.9	17.5	22.3	17.5	22.1	17.5	17.5	23.3	1	2.8	0.0039526	0.313	-1.642	0.066	-0.464
Psmd6	26S proteasome non-ATPase regulatory subunit 6	17.5	17.5	17.5	21.9	17.5	17.5	17.5	17.5	22.7	1	4.4	0.0043384	0.427	-1.453	0.427	-1.742
Psmd7	26S proteasome non-ATPase regulatory subunit 7	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	2	15	0	0.000	0.000	0.000	0.000
Psme2	Proteasome activator complex subunit 2	17.5	21.0	24.7	22.1	17.5	17.5	17.5	21.7	17.5	2	5	0	0.327	2.046	0.363	2.174
Ptbp1	Polypyrimidine tract-binding protein 1	17.5	21.1	23.1	24.0	17.5	21.1	17.5	21.3	24.5	4	11	0	0.043	-0.311	0.071	-0.529
Ptges3	Prostaglandin E synthase 3	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	21.8	1	6.2	0.0041841	0.000	0.000	0.427	-1.448
Ptgs2	Prostaglandin G/H synthase 2	17.5	17.5	22.4	17.5	17.5	17.5	17.5	17.5	21.1	3	7.3	0	0.427	1.638	0.076	0.438
Ptma	Prothymosin alpha; Prothymosin alpha, N-terminally processed	23.2	21.3	17.5	20.1	17.5	23.5	17.5	17.5	23.3	2	15.3	0	0.047	0.332	0.186	1.247
Ptrh2	Peptidyl-tRNA hydrolase 2, mitochondrial	17.5	17.5	17.5	23.1	17.5	17.5	17.5	17.5	22.8	1	7.7	0.0023095	0.427	-1.866	0.427	-1.763
Qars		17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	22.0	1	2.5	0.0038685	0.000	0.000	0.427	-1.510
Rab10	Ras-related protein Rab-10	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	2	6	0.0092251	0.000	0.000	0.000	0.000
Rab1A; Rab1b; Rab1	Ras-related protein Rab-1A; Ras-related protein Rab-1B	17.5	21.7	23.8	22.5	17.5	17.5	17.5	17.5	21.5	3	13.2	0	0.298	1.829	0.404	2.168
Rab2a	Ras-related protein Rab-2A	17.5	21.4	23.5	23.0	17.5	17.5	17.5	17.5	20.9	2	14.2	0	0.227	1.470	0.440	2.152
Rab7a	Ras-related protein Rab-7a	17.5	24.2	17.5	24.8	23.4	24.3	17.5	17.5	24.4	1	6.8	0	0.907	-4.424	0.005	-0.050

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Ran; 1700009-N14Rik	GTP-binding nuclear protein Ran	22.1	23.2	22.6	22.6	17.5	23.1	21.6	17.5	24.7	5	20.8	0	0.357	1.566	0.263	1.384
Rars	Arginine–tRNA ligase, cytoplasmic	17.5	21.1	21.1	22.2	17.5	23.3	17.5	17.5	23.7	5	10.6	0	0.204	-1.133	0.044	0.308
Rb1cc1	RB1-inducible coiled-coil protein 1	23.7	22.9	17.5	22.6	17.5	17.5	17.5	17.5	21.0	2	1.3	0.0055556	0.346	2.166	0.522	2.705
Rbbp4	Histone-binding protein RBBP4	17.5	17.5	17.5	21.7	21.9	17.5	17.5	17.5	21.8	1	3.3	0.0042827	0.934	-2.847	0.427	-1.445
Rbm3	Putative RNA-binding protein 3	20.4	20.7	17.5	17.5	17.5	17.5	17.5	17.5	23.1	3	26.1	0	0.932	2.051	0.030	0.188
Rbm39	RNA-binding protein 39	21.9	22.1	24.3	23.3	22.2	23.8	20.9	24.0	23.9	3	7.9	0	0.130	-0.319	0.046	-0.170
Reep5	Receptor expression-enhancing protein 5	17.5	22.7	17.5	22.3	17.5	17.5	17.5	17.5	17.5	1	5.4	0.0042194	0.022	0.159	0.427	1.748
Rnf17	RING finger protein 17	17.5	21.8	17.5	22.5	17.5	17.5	17.5	17.5	17.5	1	0.5	1	0.035	-0.226	0.427	1.441
Rnf213	E3 ubiquitin-protein ligase RNF213	17.5	20.4	17.5	20.4	17.5	23.2	17.5	17.5	21.1	2	0.6	0	0.427	-1.920	0.047	-0.212
Rnps1	RNA-binding protein with serine-rich domain 1	23.9	17.5	24.3	24.9	17.5	21.5	19.2	23.4	25.4	5	22.6	0	0.070	0.618	0.095	-0.760
Rpl10; Rpl10l	60S ribosomal protein L10; 60S ribosomal protein L10-like	18.4	17.5	22.1	24.6	21.8	24.6	17.5	23.0	24.0	4	16.4	0	1.226	-4.385	0.371	-2.184
Rpl10a	60S ribosomal protein L10a; Ribosomal protein	20.9	17.5	24.4	24.6	20.8	23.4	20.8	17.5	25.1	4	25.3	0	0.371	-2.011	0.026	-0.226
Rpl11; Gm10036; Gm5093	60S ribosomal protein L11	21.7	24.4	24.4	24.1	22.7	25.6	17.5	23.1	24.1	2	14.6	0	0.197	-0.634	0.354	1.915
Rpl12	60S ribosomal protein L12	22.9	17.5	23.7	24.5	17.5	24.2	17.5	17.5	25.2	2	14.5	0	0.083	-0.698	0.150	1.292
Rpl13	60S ribosomal protein L13	21.6	22.9	23.8	24.0	17.5	24.6	17.5	22.1	25.3	7	34.6	0	0.110	0.717	0.177	1.100
Rpl13a	60S ribosomal protein L13a	22.8	23.3	25.6	25.2	25.2	25.9	24.5	25.2	25.0	4	20.7	0	0.789	-1.536	0.511	-1.043
Rpl14	60S ribosomal protein L14	21.9	22.6	24.4	24.6	17.5	24.6	22.1	24.4	24.5	3	17.1	0	0.104	0.718	0.273	-0.719
Rpl15; Gm10020	60S ribosomal protein L15; Ribosomal protein L15	23.8	23.4	26.0	25.3	17.5	24.5	22.8	24.3	27.6	9	35.8	0	0.304	1.945	0.117	-0.530
Rpl17	60S ribosomal protein L17	19.1	20.8	22.5	24.4	17.5	26.8	21.7	22.8	25.6	8	43.5	0	0.288	-2.102	0.800	-2.613
Rpl18	60S ribosomal protein L18	25.5	26.2	27.0	28.3	27.3	27.6	25.5	28.2	28.5	7	31.4	0	1.348	-1.510	0.485	-1.174
Rpl18a; Gm17541	60S ribosomal protein L18a	21.1	22.1	22.9	23.1	23.8	24.6	22.7	17.5	22.6	4	21.6	0	1.233	-1.798	0.242	1.105

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Rpl19	60S ribosomal protein L19	22.6	24.0	24.5	24.4	21.5	24.2	19.9	23.1	25.6	4	18.4	0	0.114	0.344	0.188	0.852
Rpl21	60S ribosomal protein L21	18.7	20.7	23.4	23.2	17.5	17.5	17.5	20.7	24.0	2	16.2	0	0.263	1.548	0.031	0.211
Rpl22	60S ribosomal protein L22	22.0	17.5	17.5	23.4	23.6	23.8	21.8	22.9	23.7	2	19.5	0	1.427	-4.594	1.132	-3.837
Rpl23	60S ribosomal protein L23	17.5	18.2	21.8	21.9	17.5	24.9	20.0	17.5	22.6	2	25	0	0.376	-2.255	0.167	-0.878
Rpl23a	60S ribosomal protein L23a	20.3	24.6	25.4	24.4	17.5	17.5	17.5	20.8	21.5	2	15.4	0.0039683	0.579	3.616	0.803	3.476
Rpl24	60S ribosomal protein L24	19.8	20.6	22.7	22.0	21.2	23.4	17.5	17.5	21.7	2	14	0	0.471	-1.172	0.578	2.140
Rpl27a	60S ribosomal protein L27a	21.0	22.6	22.5	23.5	21.8	24.0	17.5	22.7	23.2	3	14.9	0	0.536	-1.026	0.178	0.893
Rpl28	60S ribosomal protein L28	17.5	17.5	17.5	25.9	24.1	24.3	17.5	20.2	25.5	5	29.2	0	3.693	-7.296	0.690	-3.550
Rpl29; Gm17669; Gm10709	60S ribosomal protein L29	23.2	21.8	22.7	24.3	17.5	23.1	17.5	22.0	23.2	4	13.8	0	0.167	0.943	0.397	1.671
Rpl3	60S ribosomal protein L3	23.4	17.5	23.4	25.2	23.2	24.9	20.9	22.6	24.8	4	16.1	0	0.658	-2.997	0.236	-1.362
Rpl30	60S ribosomal protein L30	17.5	22.5	17.5	25.0	17.5	17.5	21.5	17.5	24.8	3	34.8	0	0.100	-0.834	0.324	-2.115
Rpl34; Gm2178	60S ribosomal protein L34	19.3	21.9	22.1	22.9	21.8	24.5	17.5	17.5	23.9	4	21.4	0	0.765	-1.969	0.245	1.436
Rpl38	60S ribosomal protein L38	17.5	19.8	21.5	21.6	17.5	17.5	17.5	17.5	23.3	2	35.7	0	0.150	0.723	0.025	0.169
Rpl4	60S ribosomal protein L4	17.5	25.0	25.9	25.0	24.8	26.3	22.4	25.2	26.5	8	29.1	0	0.395	-2.538	0.259	-1.909
Rpl5	60S ribosomal protein L5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	20.8	1	4.4	0.0041494	0.000	0.000	0.427	-1.112
Rpl6; Gm5428	60S ribosomal protein L6	24.2	23.8	24.7	25.4	23.4	23.4	22.5	24.3	25.9	9	26	0	0.085	0.178	0.006	-0.020
Rpl7	60S ribosomal protein L7	26.4	25.9	27.2	26.9	21.0	24.4	23.1	26.5	26.0	12	39.6	0	0.619	2.403	0.511	1.320
Rpl7a	60S ribosomal protein L7a	23.0	24.3	25.4	24.7	24.1	25.5	23.5	25.0	25.5	6	18.8	0	0.264	-0.531	0.181	-0.437
Rpl8	60S ribosomal protein L8	21.7	22.9	23.9	23.4	23.1	24.9	17.5	23.6	24.3	4	14	0	0.498	-0.967	0.164	0.984
Rpl9	60S ribosomal protein L9	17.5	17.5	21.1	24.8	24.3	27.1	17.5	17.5	23.5	3	22.4	0	1.984	-6.699	0.124	-0.790
Rplp0	60S acidic ribosomal protein P0	17.5	24.2	23.5	25.2	23.2	23.3	17.5	25.5	24.7	3	20.2	0	0.412	-2.161	0.089	-0.834
Rplp2	60S acidic ribosomal protein P2	21.0	24.2	25.4	21.8	17.5	20.2	22.2	24.5	26.5	7	64.3	0	0.969	3.745	0.175	-0.840
Rpn1	Dolichyl-diphosphooligosaccharide–protein glycosyltransferase subunit 1	17.5	21.2	17.5	21.7	17.5	20.5	17.5	17.5	20.6	2	3.9	0.0043956	0.262	-1.154	0.045	0.213
Rpn2	Dolichyl-diphosphooligosaccharide–protein glycosyltransferase subunit 2	17.5	25.7	26.6	20.6	17.5	17.5	17.5	17.5	21.4	2	6.3	0	0.701	4.711	0.630	4.437
Rps11	40S ribosomal protein S11	20.9	17.5	22.2	24.0	17.5	23.8	21.4	24.6	25.2	6	38.6	0	0.240	-1.548	0.900	-3.549
Rps13	40S ribosomal protein S13	22.4	23.4	24.1	24.8	22.9	23.2	21.6	22.9	24.8	8	40.4	0	0.143	-0.300	0.068	0.203

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Rps14	40S ribosomal protein S14	22.7	23.4	24.4	23.4	23.6	24.6	17.5	23.2	24.8	5	18.5	0	0.230	-0.372	0.302	1.684
Rps16	40S ribosomal protein S16	17.5	20.7	20.9	21.5	17.5	20.9	17.5	17.5	22.3	2	15.1	0.0022883	0.049	-0.238	0.111	0.602
Rps18; Gm10260	40S ribosomal protein S18	20.3	20.6	21.8	21.1	21.2	21.1	17.5	21.3	22.4	1	7.2	0.0041754	0.211	-0.238	0.122	0.519
Rps19	40S ribosomal protein S19	17.5	17.5	21.6	17.5	17.5	17.5	17.5	17.5	21.4	2	13.8	0.0022936	0.427	1.351	0.012	0.066
Rps2; Gm6576; Gm5786	40S ribosomal protein S2	20.7	22.7	24.1	23.4	22.4	21.8	22.3	22.7	24.4	4	15.4	0	0.018	-0.058	0.214	-0.645
Rps20	40S ribosomal protein S20	20.3	21.6	17.5	22.9	17.5	17.5	17.5	17.5	24.0	2	22.7	0	0.081	0.499	0.018	0.134
Rps21	40S ribosomal protein S21	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	22.6	2	34.9	0	0.000	0.000	0.427	-1.686
Rps23	40S ribosomal protein S23	22.8	23.2	23.1	24.8	21.7	23.4	19.4	23.2	24.2	4	21.7	0	0.100	-0.254	0.199	0.767
Rps25	40S ribosomal protein S25	17.5	23.2	23.7	24.0	22.7	23.2	17.5	21.6	24.0	3	22.4	0	0.386	-1.850	0.051	0.408
Rps27; Rps27l	40S ribosomal protein S27;40S ribosomal protein S27-like	17.5	22.5	23.6	23.8	17.5	22.7	17.5	17.5	23.9	2	31	0	0.013	-0.107	0.221	1.606
Rps27a; Ubc; Ubb	Ubiquitin-40S ribosomal protein S27a;Ubiquitin	27.0	27.8	27.6	28.3	27.9	27.8	26.1	27.1	28.5	7	39.1	0	0.855	-0.554	0.090	0.190
Rps3	40S ribosomal protein S3	23.1	24.5	25.3	25.7	25.3	25.6	24.1	24.6	26.7	8	46.9	0	0.834	-1.219	0.319	-0.804
Rps3a	40S ribosomal protein S3a	23.3	24.3	24.4	26.4	24.2	26.8	21.7	25.4	26.8	6	26.9	0	0.944	-1.768	0.153	-0.643
Rps4x	40S ribosomal protein S4, X isoform	17.5	17.5	22.6	22.9	17.5	23.8	17.5	17.5	23.7	4	13.7	0	0.357	-2.225	0.048	-0.374
Rps5	40S ribosomal protein S5;40S ribosomal protein S5, N-terminally processed	17.5	17.5	23.3	23.4	17.5	23.3	17.5	17.5	24.9	4	27.5	0	0.287	-1.958	0.059	-0.532
Rps6	40S ribosomal protein S6	24.1	23.1	25.6	25.5	22.2	25.3	21.8	22.4	25.2	4	22.1	0	0.029	-0.112	0.370	1.118
Rps8	40S ribosomal protein S8	24.6	24.3	25.7	25.5	25.0	26.0	22.9	25.1	26.1	7	33.2	0	0.542	-0.645	0.049	0.150
Rps9	40S ribosomal protein S9	22.7	24.2	24.9	25.3	23.5	24.2	17.5	24.4	25.1	6	24.7	0	0.198	-0.432	0.253	1.597
Rpsa	40S ribosomal protein SA	24.8	25.5	26.0	25.2	23.8	24.9	22.6	25.9	26.7	8	42	0	0.660	0.812	0.105	0.377
Rrbp1	Ribosome-binding protein 1	24.8	24.3	26.2	26.6	28.1	26.8	24.8	26.1	25.9	23	32.3	0	1.306	-2.065	0.294	-0.509
Rrm2	Ribonucleoside-diphosphate reductase subunit M2	17.5	17.5	17.5	21.8	17.5	17.5	17.5	20.0	22.6	2	7.7	0	0.427	-1.443	0.798	-2.532
Rtn4	Reticulon-4	17.5	24.2	24.1	23.0	21.8	22.4	17.5	17.5	17.5	3	5	0	0.070	-0.450	0.935	4.433
S100a11	Protein S100-A11	24.0	26.0	25.5	25.3	23.2	22.8	17.5	24.4	25.1	1	10.2	0.0040161	0.661	1.429	0.497	2.851
S100a14	Protein S100-A14	22.4	23.7	26.6	22.6	22.5	21.2	21.4	23.1	22.7	3	40.4	0	0.738	2.128	0.598	1.805
S100a4	Protein S100-A4	17.5	17.5	22.0	17.5	17.5	17.5	17.5	17.5	23.0	2	15.8	0.0041068	0.427	1.502	0.050	-0.344
S100a6	Protein S100-A6	24.8	25.4	25.8	26.1	25.5	25.0	22.5	25.1	26.3	3	25.8	0	0.159	-0.178	0.251	0.730
Sdc1	Syndecan-1	17.5	17.5	17.5	17.5	21.0	17.5	17.5	21.4	21.7	1	6.4	0.0056497	0.427	-1.178	0.933	-2.714

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Sec22b	Vesicle-trafficking protein SEC22b	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	22.4	1	6.5	0.0043103	0.000	0.000	0.427	-1.623
Sec61a1	Protein transport protein Sec61 subunit alpha isoform 1	17.5	22.2	24.3	22.7	17.5	21.8	17.5	17.5	21.7	3	6.7	0	0.098	0.706	0.421	2.429
Sec61b	Protein transport protein Sec61 subunit beta	17.5	17.5	17.5	23.8	24.0	24.1	22.7	17.5	24.1	1	15.6	0	6.564	-6.498	0.915	-3.920
Serbp1	Plasminogen activator inhibitor 1 RNA-binding protein	23.4	24.3	24.1	24.3	17.5	22.0	17.5	24.2	26.0	7	24.6	0	0.589	2.665	0.199	1.349
Serpinh1	Serpin H1	23.1	23.1	23.6	23.9	23.2	26.2	17.5	24.6	25.3	3	9.4	0	0.558	-1.157	0.120	0.820
Set	Protein SET	17.5	17.5	22.6	17.5	17.5	17.5	17.5	17.5	17.5	1	4.5	0	0.427	1.684	0.427	1.684
Sf1	Splicing factor 1	20.0	17.5	17.5	23.0	17.5	17.5	17.5	17.5	22.4	1	2	0.0043668	0.189	-0.995	0.163	-0.794
Sf3b3	Splicing factor 3B subunit 3	17.5	20.8	21.1	22.6	17.5	22.7	17.5	17.5	24.4	3	3.7	0	0.211	-1.127	0.001	-0.004
Sf3b4	Splicing factor 3B subunit 4	17.5	17.5	17.5	20.3	17.5	20.4	17.5	17.5	20.3	1	3.3	0.0043011	0.935	-1.894	0.427	-0.924
Sfpq	Splicing factor, proline- and glutamine-rich	17.5	17.5	22.9	22.9	17.5	23.8	17.5	17.5	23.0	1	2.4	0.0022676	0.320	-2.078	0.005	-0.037
Shmt2	Serine hydroxymethyltransferase	17.5	17.5	17.5	21.8	22.2	17.5	17.5	21.5	21.8	1	2.2	0.0022624	0.932	-2.987	0.933	-2.761
Slc25a24	Calcium-binding mitochondrial carrier protein SCaMC-1	17.5	20.9	21.7	21.1	19.9	17.5	17.5	17.5	20.0	1	2.7	0.0091912	0.118	0.541	0.475	1.687
Slc25a3	Phosphate carrier protein, mitochondrial	23.2	25.8	26.5	26.5	26.1	24.3	24.4	23.8	25.4	6	15.7	0	0.142	-0.463	0.231	0.652
Slc25a5; Slc25a4	ADP/ATP translocase 2; ADP/ATP translocase 2, N-terminally processed	21.4	25.0	27.3	25.4	26.2	26.2	24.5	23.8	26.5	10	30.9	0	0.314	-1.330	0.061	-0.336
Slc39a7	Zinc transporter SLC39A7	17.5	17.5	17.5	21.5	17.5	17.5	17.5	17.5	17.5	2	7.1	0.0040984	0.427	-1.337	0.000	0.000
Slc3a2	4F2 cell-surface antigen heavy chain	21.2	23.0	17.5	17.5	17.5	21.8	17.5	17.5	17.5	1	3.2	0.0039062	0.306	1.634	0.885	3.081
Smarca5	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5	17.5	17.5	17.5	17.5	17.5	23.1	17.5	17.5	17.5	1	1.4	0	0.427	-1.870	0.000	0.000
Snd1	Staphylococcal nuclease domain-containing protein 1	19.9	21.4	24.0	23.2	22.0	21.9	17.5	22.0	23.5	6	8.9	0	0.183	-0.611	0.127	0.752

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Snrnp27	U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein	20.0	20.8	23.2	22.2	17.5	20.9	17.5	20.6	22.7	3	17.4	0	0.255	1.088	0.230	1.046
Snrnp70	U1 small nuclear ribonucleoprotein 70 kDa	17.5	17.5	17.5	21.7	17.5	17.5	17.5	17.5	17.5	1	2.2	0.0056391	0.427	-1.407	0.000	0.000
Srek1	Splicing regulatory glutamine/lysine-rich protein 1	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	22.2	1	2.8	0	0.000	0.000	0.427	-1.564
Srrm1	Serine/arginine repetitive matrix protein 1	17.5	20.9	23.9	17.5	17.5	22.1	17.5	22.6	23.0	5	8.2	0	0.294	1.742	0.032	-0.246
Srrm2	Serine/arginine repetitive matrix protein 2	24.8	22.9	23.6	25.2	24.5	24.0	23.3	24.6	25.0	14	7.7	0	0.562	-0.818	0.293	-0.537
Srsf1	Serine/arginine-rich splicing factor 1	25.9	25.5	25.2	25.6	20.8	23.6	19.9	25.4	26.2	10	42.3	0	0.731	2.212	0.355	1.716
Srsf10	Serine/arginine-rich splicing factor 10	17.5	23.7	17.5	24.3	17.5	21.5	17.5	17.5	23.6	3	17.9	0	0.213	-1.566	0.001	0.008
Srsf11		17.5	19.7	20.3	19.6	17.5	17.5	17.5	17.5	22.0	4	13.9	0	0.366	0.956	0.028	0.144
Srsf2	Serine/arginine-rich splicing factor 2	25.3	24.2	25.4	25.3	22.8	24.7	23.9	25.7	26.1	4	21.7	0	0.324	0.665	0.124	-0.272
Srsf3; Gm12355	Serine/arginine-rich splicing factor 3	26.7	26.1	26.4	27.4	24.0	27.5	23.6	27.9	28.0	8	44.5	0	0.036	0.123	0.019	-0.085
Srsf4	Serine/arginine-rich splicing factor 4	17.5	17.5	17.5	17.5	17.5	17.5	21.7	17.5	22.7	2	2.7	0	0.000	0.000	0.918	-3.136
Srsf5	Serine/arginine-rich splicing factor 5	21.6	22.6	17.5	24.1	17.5	24.2	17.5	23.2	26.5	3	7.4	0	0.196	-1.386	0.237	-1.839
Srsf6		24.2	22.3	23.8	24.6	23.3	21.9	22.8	24.3	24.6	5	11.8	0	0.048	0.135	0.220	-0.447
Srsf7	Serine/arginine-rich splicing factor 7	24.6	23.9	23.0	26.6	19.7	24.1	22.1	23.3	26.6	6	24.7	0	0.060	0.354	0.039	-0.166
Ssr3	Translocon-associated protein subunit gamma	17.5	17.5	17.5	21.4	17.5	17.5	17.5	17.5	21.1	1	7.6	0.0037951	0.427	-1.312	0.427	-1.206
Ssrp1	FACT complex subunit SSRP1	17.5	17.5	17.5	22.1	17.5	17.5	17.5	17.5	19.5	2	2.8	0	0.427	-1.541	0.427	-0.654
Stat1	Signal transducer and activator of transcription 1	17.5	22.1	22.1	23.7	17.5	21.4	19.4	20.2	22.0	4	6.1	0	0.043	-0.298	0.009	0.049
Stip1	Stress-induced-phosphoprotein 1	21.3	22.5	22.7	22.5	21.9	23.2	20.0	23.1	23.5	2	4.2	0	0.266	-0.377	0.018	-0.064
Stmn1	Stathmin	22.3	17.5	22.2	21.9	17.5	23.8	17.5	17.5	25.0	2	18.1	0	0.053	-0.377	0.081	0.678

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Stmn2	Stathmin-2	17.5	20.4	17.5	21.2	17.5	20.7	17.5	17.5	17.5	1	5.6	0.0042373	0.365	-1.332	0.427	0.980
Stt3a	Dolichyl-diphosphooligosaccharide–protein glycosyltransferase subunit STT3A	17.5	20.8	21.7	20.4	17.5	17.5	17.5	17.5	21.6	1	1.6	0.0038536	0.413	1.547	0.242	1.143
Sumo2; Sumo3	Small ubiquitin-related modifier 2; Small ubiquitin-related modifier 3	22.1	22.4	23.5	22.5	17.5	17.5	17.5	24.0	23.1	1	12.6	0.0022573	0.965	3.508	0.211	1.127
Supt16h; Supt16	FACT complex subunit SPT16	17.5	21.5	17.5	22.2	17.5	22.7	17.5	17.5	22.3	1	1.4	0.0022727	0.388	-1.954	0.045	-0.275
Syncrip	Heterogeneous nuclear ribonucleoprotein Q	19.7	17.5	17.5	22.9	17.5	22.8	17.5	17.5	22.7	1	2.1	0.0041237	0.668	-2.846	0.207	-1.010
Tagln2	Transgelin-2	24.1	25.6	24.5	25.1	23.7	23.5	22.7	24.6	25.8	9	44.7	0	0.379	0.603	0.131	0.358
Tap1	Antigen peptide transporter 1	17.5	23.6	24.3	23.0	22.3	25.0	17.5	17.5	21.4	3	6.4	0	0.293	-1.663	0.521	2.989
Tap2	Antigen peptide transporter 2	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	21.0	1	2.6	0.0055762	0.000	0.000	0.427	-1.171
Tardbp	TAR DNA-binding protein 43	21.5	22.6	22.9	22.7	17.5	22.4	17.5	21.5	22.0	3	10.1	0	0.362	1.500	0.609	2.013
Tcof1	Treacle protein	26.1	25.2	26.4	26.8	24.0	24.3	23.7	26.2	26.6	21	20.5	0	0.351	0.812	0.132	0.346
Tcp1	T-complex protein 1 subunit alpha	21.5	22.4	23.7	24.3	17.5	23.2	17.5	17.5	22.9	4	10.1	0	0.144	0.857	0.765	3.210
Tgm1	Protein-glutamine gamma-glutamyltransferase K	23.7	23.0	25.2	23.2	23.8	22.5	17.5	17.5	22.1	4	6.1	0	0.475	0.800	1.370	4.913
Tgm3	Protein-glutamine gamma-glutamyltransferase E	21.6	23.0	17.5	20.3	19.5	18.2	17.5	22.8	21.7	1	1.9	0	0.318	1.369	0.006	0.041
Tgtp; Tgtp2		17.5	17.5	17.5	23.0	17.5	17.5	17.5	17.5	21.9	1	2.4	0.0040404	0.427	-1.850	0.427	-1.456
Thrap3	Thyroid hormone receptor-associated protein 3	17.5	21.3	22.5	20.4	22.5	19.8	17.5	23.0	24.8	4	6.7	0	0.104	-0.495	0.197	-1.367
Tkt	Transketolase	21.3	22.2	21.3	23.4	22.7	22.7	17.5	22.1	23.8	3	6.4	0	1.549	-1.324	0.091	0.488
Tln1	Talin-1	22.0	25.0	24.1	25.2	26.2	25.7	19.9	22.7	25.8	16	9.8	0	0.993	-1.998	0.178	0.900
Tmem184c		17.5	24.0	21.5	20.9	26.6	25.8	26.6	25.0	23.9	1	3.5	1	0.583	-3.408	0.943	-4.140

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Tmpo	Lamina-associated polypeptide 2, isoforms beta/delta/epsilon/gamma; Lamina-associated polypeptide 2, isoforms alpha/zeta	22.1	22.9	24.6	17.5	17.5	23.0	17.5	22.4	25.2	6	18.6	0	0.915	3.872	0.256	1.516
Top1	DNA topoisomerase 1	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	2	3.3	0	0.000	0.000	0.000	0.000
Top2a	DNA topoisomerase 2-alpha	21.1	20.8	17.5	21.3	17.5	20.4	17.5	20.8	21.1	4	3.9	0	0.019	0.093	0.002	-0.011
Tpi1	Triosephosphate isomerase	20.6	23.0	23.1	23.1	17.5	17.5	17.5	17.5	22.5	4	15.4	0	0.622	2.842	0.766	3.073
Tpm3; Tpm3-rs7	Tropomyosin alpha-3 chain	24.6	24.5	24.5	25.0	22.3	25.1	23.5	25.8	26.7	13	34.7	0	0.169	0.416	0.324	-0.754
Tpm4	Tropomyosin alpha-4 chain	20.9	17.5	17.5	22.9	17.5	22.2	17.5	17.5	24.2	5	19	0	0.469	-2.218	0.161	-1.077
Tpt1	Translationally-controlled tumor protein	17.5	21.2	17.5	23.4	17.5	24.0	17.5	17.5	24.0	1	8.1	0	0.527	-2.901	0.137	-0.925
Tra2a	Transformer-2 protein homolog alpha	21.1	21.3	22.6	21.6	17.5	20.7	17.5	17.5	21.7	2	5	0.0023419	0.558	1.690	0.869	2.754
Tra2b	Transformer-2 protein homolog beta	22.0	17.5	17.5	23.1	17.5	22.0	17.5	17.5	22.6	4	12.5	0	0.333	-1.843	0.027	-0.184
Trim28	Transcription intermediary factor 1-beta	17.5	17.5	17.5	17.5	20.0	25.2	17.5	17.5	17.5	2	4.1	0	0.684	-3.412	0.000	0.000
Try4;Try5		26.5	27.1	26.9	27.4	24.6	24.5	24.5	25.5	27.6	1	8.1	0.0039216	0.599	1.311	0.436	0.943
Tuba1a; Tuba1c; Tuba3a	Tubulin alpha-1A chain;Tubulin alpha-1C chain	26.1	27.8	28.4	28.0	28.8	29.6	27.5	27.9	28.8	14	2.9	0	0.731	-1.343	0.317	-0.618
Tuba1b; Tuba4a	Tubulin alpha-1B chain;Tubulin alpha-4A chain	20.2	18.1	20.6	23.3	23.4	17.5	17.5	17.5	22.4	14	2.9	0.0042283	0.350	-1.766	0.099	0.499
Tubb2b; Tubb2a	Tubulin beta-2B chain;Tubulin beta-2A chain	17.5	17.5	17.5	24.0	17.5	22.5	21.0	17.5	17.5	9	3.4	0	0.912	-3.829	0.427	-1.159
Tubb4b; Tubb4a	Tubulin beta-4B chain;Tubulin beta-4A chain	22.5	24.1	24.8	24.8	22.5	25.6	21.7	22.3	23.8	13	7	0	0.175	-0.523	0.570	1.183
Tubb5	Tubulin beta-5 chain	26.3	27.6	28.7	29.2	28.3	29.3	27.4	27.8	29.0	14	13.1	0	0.861	-1.390	0.269	-0.570
Tubb6	Tubulin beta-6 chain	17.5	17.5	21.8	21.5	17.5	25.2	17.5	17.5	21.4	10	11.9	0	0.390	-2.456	0.023	0.136
Txn	Thioredoxin	22.9	25.3	26.1	25.2	24.6	17.5	24.6	25.0	25.6	4	31.4	0	0.362	2.309	0.098	-0.270
Txnl1	Thioredoxin-like protein 1	17.5	21.4	17.5	21.5	21.4	21.5	17.5	17.5	22.0	2	12.8	0.004158	0.948	-2.654	0.034	-0.199

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
U2surp	U2 snRNP-associated SURP motif-containing protein	21.2	21.6	17.5	22.4	17.5	17.5	17.5	17.5	22.8	3	6	0	0.182	0.996	0.143	0.849
Uba1	Ubiquitin-like modifier-activating enzyme 1	21.6	24.7	25.1	26.1	24.1	23.5	21.2	25.9	25.2	7	10.2	0	0.219	-0.767	0.059	-0.313
Ugdh	UDP-glucose 6-dehydrogenase	21.3	22.0	17.5	22.4	17.5	17.5	17.5	21.4	21.0	2	6.7	0	0.200	1.119	0.051	0.279
Upp2	Uridine phosphorylase 2	23.6	27.7	29.3	26.6	17.5	23.6	21.2	25.0	24.5	1	4.1	0	0.607	4.300	0.722	3.278
Vars	Valine–tRNA ligase	17.5	19.5	17.5	20.5	17.5	17.5	21.9	17.5	21.8	2	2.1	0	0.092	-0.309	0.623	-2.217
Vcl	Vinculin	24.0	26.0	26.4	25.8	24.8	25.3	22.6	23.8	24.6	11	12.4	0	0.088	0.195	0.894	1.809
Vcp	Transitional endoplasmic reticulum ATPase	23.0	24.0	25.1	26.5	22.7	24.1	22.9	23.1	26.4	10	17.2	0	0.117	-0.409	0.027	-0.106
Vdac1	Voltage-dependent anion-selective channel protein 1	17.5	23.0	25.0	23.9	23.0	22.7	22.3	21.0	22.7	5	20.5	0	0.239	-1.383	0.024	-0.169
Vdac2	Voltage-dependent anion-selective channel protein 2	17.5	17.5	24.5	22.7	23.3	22.9	21.9	17.5	23.8	2	10.8	0	0.599	-3.124	0.155	-1.240
Vim	Vimentin	28.2	28.8	29.2	29.1	28.8	29.6	26.4	29.0	30.4	43	79.4	0	0.489	-0.418	0.049	0.171
Wdr96		17.5	17.5	23.1	17.5	17.5	17.5	17.5	17.5	17.5	1	1	0	0.427	1.870	0.427	1.870
Xpo1	Exportin-1	17.5	18.3	22.8	23.9	17.5	20.9	17.5	17.5	23.5	2	2.8	0.0023041	0.195	-1.267	0.002	0.013
Ybx1	Nuclease-sensitive element-binding protein 1	22.1	21.5	24.4	22.4	17.5	22.5	21.0	24.1	25.7	7	51.2	0	0.414	1.828	0.228	-0.952
Ywhab	14-3-3 protein beta/alpha; 14-3-3 protein beta/alpha, N-terminally processed	17.5	17.5	17.5	17.5	17.5	21.9	17.5	17.5	23.3	2	5.7	0	0.427	-1.451	0.427	-1.922
Ywhae	14-3-3 protein epsilon	22.3	24.0	24.3	24.0	23.2	22.1	21.7	23.3	24.9	4	18.8	0	0.185	0.410	0.066	0.210
Ywhag	14-3-3 protein gamma; 14-3-3 protein gamma, N-terminally processed	22.1	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	1	5.7	0.0039448	0.427	1.525	0.427	1.525
Ywhah	14-3-3 protein eta	21.9	17.5	17.5	17.5	21.0	21.1	17.5	23.0	23.2	3	17.1	0	0.191	-0.936	0.411	-2.291
Ywhaq	14-3-3 protein theta	21.4	21.6	22.5	22.8	17.5	17.5	17.5	17.5	23.6	3	13.9	0	0.658	2.596	0.497	2.331
Ywhaz	14-3-3 protein zeta/delta	24.8	24.9	26.2	25.5	24.3	24.6	22.3	24.4	26.1	10	38.4	0	0.377	0.506	0.375	1.059
Zc3h18	Zinc finger CCCH domain-containing protein 18	17.5	17.5	17.5	17.5	17.5	19.4	17.5	17.5	20.4	1	1.6	0.0038168	0.427	-0.644	0.427	-0.979

Gene names	Protein names	Active probe (A)			Inactive probe			A + competitor			Pept.	Unique sequence cov. (%)	Q-value	-Log p-value A vs IA	Difference A vs IA	-Log p-value A vs A + C	Difference A vs A + C
Znf143; Znf76; Zfp523	Zinc finger protein 143; Zinc finger protein 76	17.5	23.8	23.3	17.5	17.5	17.5	23.4	17.5	23.1	2	3.4	0.0056711	0.933	4.050	0.030	0.248
Zyx	Zyxin	24.1	24.7	23.1	25.1	17.5	24.2	21.0	24.9	25.4	4	10.7	0	0.280	1.695	0.045	0.195
	UPF0444 transmembrane protein C12orf23 homolog	17.5	17.5	24.5	24.4	17.5	21.8	17.5	17.5	23.2	1	22.6	0.0022222	0.171	-1.390	0.047	0.414
		17.5	22.6	23.4	17.5	17.5	21.4	21.8	22.4	17.5	2	0.6	0.0040486	0.457	2.390	0.089	0.605