Understanding the Structure-Activity Relationship (SAR) and PK of PROTAC Linkers

James White

Master of Science by Research in Biochemistry thesis submitted in September 2024

School of Natural Sciences

Word Count: 18580

Page Count: 113

Contents

Figures

Tables

Abstract

Proteolysis targeting chimeras (PROTACs) are a relatively new class of heterobifunctional molecules that have been identified as a potential anti-cancer therapy. They work by bringing a protein of interest (POI) and E3 ligase proteins in proximity and triggering ubiquitination and subsequent targeted degradation of the POI via the proteasome. PROTACs are comprised of a ligand to both the POI and the E3 ligase connected by a linker. The linker is not a directly functional part of the molecule and so many different varieties exist, with varying lengths and physical properties.

However, limited research has been done on whether the linker has an indirect effect on the function and pharmacokinetics of the PROTAC. This project looks to explore the effect of varying the linker on metabolism and activity, with an aliphatic and an ethylene glycol linker tested for their stability in human blood plasma, human liver microsomes and human hepatocytes, with metabolite identification also performed.

This project shows that changing the linker does have a minor effect on stability, with different linkers performing better in different matrices. Stability of the PROTACs was higher in plasma than in HLMs, with anomalous results in hepatocytes. Metabolite identification showed that the likely route of metabolism in plasma was via hydrolysis, and via oxidation in HLMs. However, in HLMs the POI ligand was the primary point of metabolism, meaning that conclusions about linker metabolism couldn't be properly inferred.

Acknowledgements

With thanks to Dr. Dave Beal and Prof. Mark Smales of the University of Kent for their guidance and supervision, as well as Dr. Satish Kurusamy and the rest of the Smales/Beal lab group for their help and support throughout the project.

Also many thanks to Dr. David Palmer, Dr. Angus Nedderman and Mark Savage from Yorkbio in Sandwich for their expertise and the rest of the Yorkbio team for accommodating me when I visited their labs to gather data for my project.

With additional thanks to the technical and administrative staff as well as the general scientific community at the University of Kent.

Abbreviations

Abbreviation	Meaning
BET	Bromodomain and extraterminal
BRD4	Bromodomain-containing protein 4
CRBN	Cereblon
DCM	Dichloromethane
DMF	N,N-dimethyl formamide
DPBS	Dulbecco's phosphate buffered saline
EIC	Extracted ion chromatogram
ER	Endoplasmic reticulum
HLM	Human liver microsome
LCMS	Liquid chromatography-mass spectroscopy
MLM	Mouse liver microsome
PDA	Photodiode array detector
POI	Protein of Interest
PROTAC	Proteolysis Targeting Chimaera
TFA	Trifluoroacetic acid
TIC	Total ion chromatogram
VHL	Von Hippel-Lindau

Introduction

Cancer

Cancer is a devastating condition that is defined as the uncontrolled growth and division of cells, and it affects a substantial number of people worldwide, with 20 million new cases and 9.7 million deaths in 2022 alone[1] and finding effective treatments for it has been a major aim for the biotechnological and pharmacological industries. Survival rates have been increasing in recent years,[1,3] but finding new ways to combat cancer remains an important endeavour, especially as cancer has shown the ability to develop resistances to existing treatments.[2]

Currently, cancer is typically treated in various ways. Notable methods include surgery, immunotherapy, hormone therapy, radiotherapy and chemotherapy.[23] It is very common for multiple methods being employed simultaneously, to maximise efficacy.

As cancer typically forms tumours, surgery can be an effective treatment as it may be possible to physically remove these tumours (or whole tissues when multiple or smaller tumours are present). Advantages to this process are in the typical totality of the surgical removal and in the relative lack of systemic side effects. There are also some drawbacks, as in order to be confident of tumour removal (as tumours can be difficult to distinguish by eye), large amounts of non-cancerous tissue is often removed with the tumour,[24] which in important organs such as the brain or pancreas, can cause more harm than good. Some techniques are being developed to lower this risk, such as utilising intra-operative Raman spectroscopy to distinguish between cancerous and healthy tissue.[24] Furthermore, surgery might not even be possible due to the tumour being non-static (such as in multiple myeloma), physically inaccessible without severe risk, or too distributed or metastasised.[23]

Immunotherapy is a treatment method that utilises the immune system, and as such is a highly specifically targeted methodology. This can involve using antibodies to target tumour-associated surface proteins, such as trastuzumab, which targets HER2, triggering both the downregulation of HER2 levels[25] and also an antibody-mediated immune response;[26] and pembrolizumab, which targets the PD-1 receptor, inhibiting the immune-evasion of tumours.[27]

Immune cells can also be engineered to target cancerous cells, such as in CAR-T therapy. This process involves the removal of T-cells from a patient, before genetically modifying them to express receptors specific to tumours. The engineered cells are then reintroduced to the patient, where they can act to destroy any cancerous cells.[23,28]

The specificity of immune therapy does carry significant advantages, but as antibodies are large proteins (at around 150 kDa) they lack the penetration[29] of a small molecule, and as immune cells are reliant on bodily infrastructure for movement around the body tumours can become invisible to the immune system, whilst still being connected enough to allow smaller molecules such as nutrients to penetrate throughout the tumour.

Hormone therapy works on the endocrine system to affect tumour cell signalling, typically by depriving it of growth promoters, which is achieved either by inhibiting synthesis of hormones, such as the drug letrozole,[30] which inhibits oestrogen synthase aromatase or by antagonising the hormone receptor, as in the case of the drug bicalutamide,[31] which blocks androgen receptors. Upsetting hormone balance can have some drastic effects on health and mood, which can be difficult side effects to deal with.[23]

Radiotherapy involves the use of ionising radiation to directly damage the DNA of cancerous cells, preventing replication.[23] The radiation dose is typically very targeted, in three dimensions, to reduce damage to neighbouring tissue, but similarly to in surgery, areas immediately adjacent to the tumour are often also targeted to ensure that the whole tumour is treated, which can cause damage to otherwise healthy cells.[23,32]

Chemotherapy involves the use of (typically small molecule) chemicals to directly or indirectly damage cancer cells. As cancer cells are human cells, chemicals which are simply cytotoxic will likely cause serious damage to the rest of the body, so chemotherapy targets need to be cancer associated, or cancer specific, such as cells undergoing mitosis (division) or cancer associated proteins.[23]

Targeting dividing cells is an established method and can be achieved in a multitude of ways. One such way is to damage DNA to prevent replication, such as cisplatin[33] or mitomycin C[34] (which both cross link DNA bases, stalling replication forks).

It is also possible to target cell-cycle proteins to inhibit mitosis, such as cyclin dependent kinases. Examples of drugs that work on this pathway include palbociclib[35] (which targets CDK4 and CDK6) and alvocidib (which targets CDK9).[36]

Whilst effective, chemotherapy has an inherent downside of affecting all dividing or growing cells, not just cancerous ones, which causes significant side-effects, such as hair loss and immunosuppression. Considering also that resistance can be developed, it is important to consider alternatives to traditional chemotherapy to enable novel methods to try and limit both side-effects and resistance. One method that shows promise is targeting cancer-associated proteins for degradation via the ubiquitin-proteasome pathway.

Ubiquitin-Proteasome Pathway

Ubiquitin is a small (8.6 kDa) protein that can be attached to other proteins to achieve a variety of functions in the cell, such as signalling and protein activation or deactivation. Both mono- and polyubiquitination can occur and different forms of ubiquitination affect different results. Many different varieties of polyubiquitination exist, such as lysing 48-linked chains (which have proteolytic functions), lysine 63-linked chains (which have functions in signalling) and others that are less well understood.[37]

Regardless of the variety of ubiquitination, it is typically (with some exceptions) initiated by three proteins, E1, E2 and E3. E1 is an ATP-dependent ubiquitin activation enzyme, which activates and binds to a single ubiquitin protein. Ubiquitin is then transferred to an E2 protein, which then interacts with an E3 ligase protein to ubiquitinate a target protein. The ubiquitin is passed directly from the E2 subunit to the target protein, with E3 ligase facilitating the transfer. There are more varieties of E2 than there are E1, and even more E3 varieties than E2. This arrangement allows for effective regulation of the ubiquitination mechanism.[38]

An important role of ubiquitination in the cell is protein degradation via the proteasome, which is a protein complex that causes proteolysis. The 26S proteasome is the kind normally found in eukaryotes and is formed of a core (20S) and a regulatory cap (19S). The core is responsible for the proteolytic activity, and the cap ensures that only proteins targeted for degradation come into contact with the core.[38]

Maintenance of protein levels is a crucial step in homeostasis, as proteins that are undesirable, either due to their function no longer being needed (or potentially now being a detriment) or their damage making them ineffective or harmful. It is also important to recycle the amino acids in these proteins, so an excretion method would not be as suitable as a degradation method. Protein damage can accumulate over time, as well as come from extraneous sources or even be a result of inappropriate protein synthesis.[38]

Certain E3 ligases can recognise these unwanted proteins and coordinate ubiquitination via the ubiquitin-loaded E2 ubiquitin-conjugation enzyme. This then triggers further ubiquitination and allows the polyubiquitinated protein to be recognised by the regulatory cap of the 26S proteasome, catalysing the destruction of the protein. Figure 1 visualises this pathway for a damaged protein.[38]

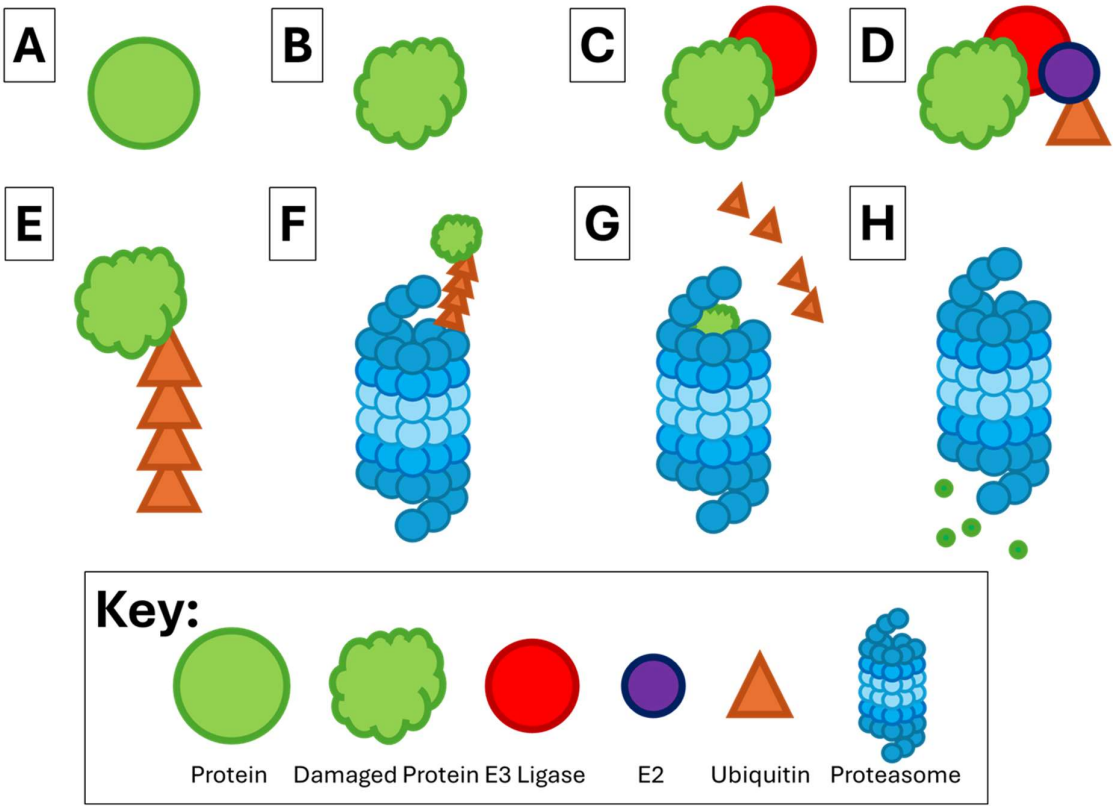


Figure 1: The Ubiquitin Proteasome Pathway (for a damaged protein). A) A protein B) The protein has become too damaged to function. C) E3 ligase recognises this damage and interacts with the damaged protein. D) A ubiquitin-loaded E2 ubiquitin-conjugation enzyme coordinates with E3 ligase, before ubiquitinating the damaged protein. E) After multiple ubiquitinations, the damaged protein now has a poly-ubiquitin chain. F) The poly-ubiquitin chain is recognised by the proteasome. G) The ubiquitin subunits are cleaved, and the damaged protein is passed through the central pore of the proteasome. H) Once passed through the proteasome, the damaged protein is now cleaved into small peptide chains.

PROTACs

Proteolysis targeting chimaeras, (PROTACs) utilise this pathway in order to degrade targeted proteins without needing E3 ligase recognition. This means that cells important to cellular processes and other cells that would not normally be degraded by the UPS system can be artificially down-regulated, modifying cellular behaviour or destroying the cell.

PROTACs are heterobifunctional molecules that are comprised of a ligand to a protein of interest (POI), a ligand to the E3 ligase component of the UPS and a linker part to connect the two other components (figure 2).[4,5,6,7] Different linkers and ligands will have different chemical and physical properties, so careful consideration of these should be taken into account.

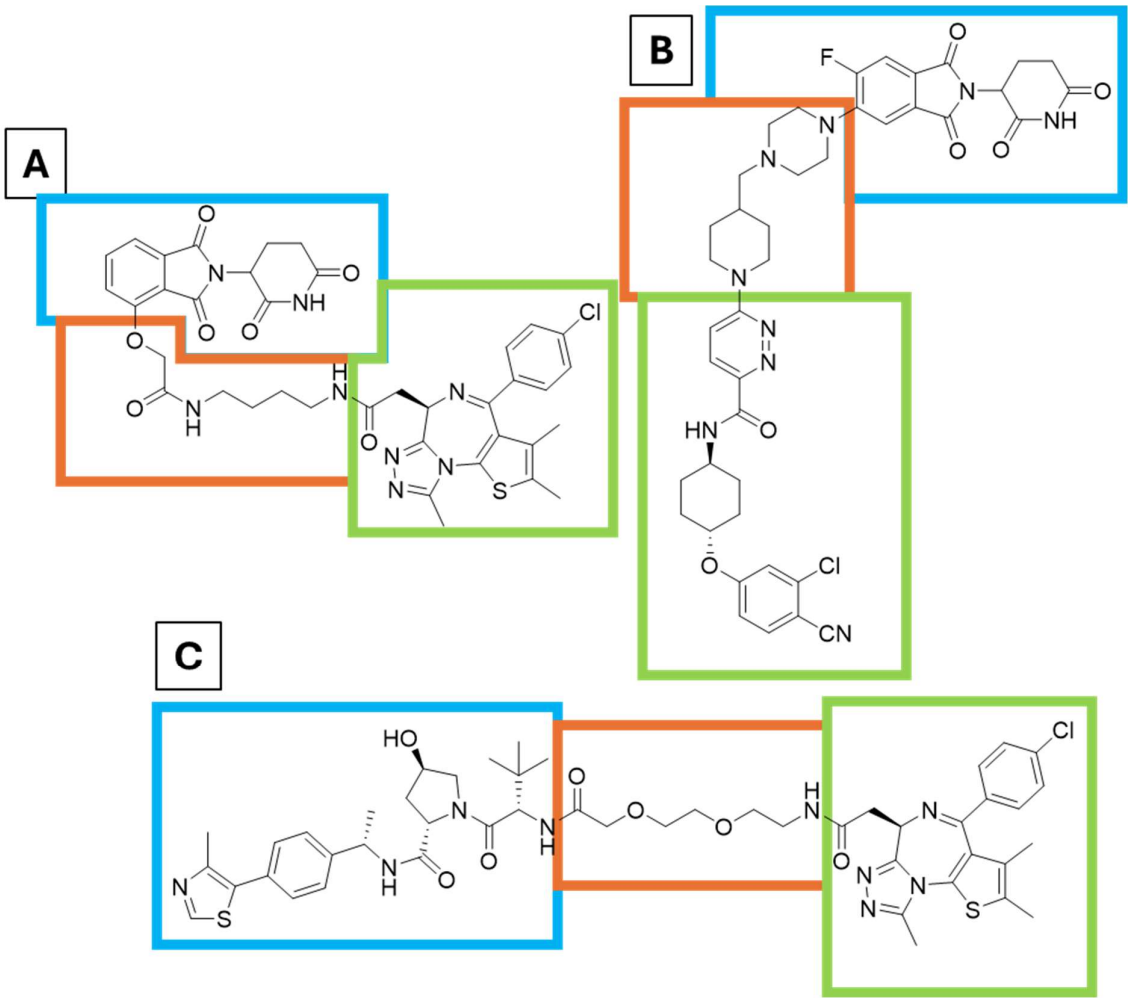


Figure 2: Examples of PROTAC molecules (A: dBET1. B: ARV-110 C: ARV-771) with the three regions highlighted in different colours. Bordered in blue are the E3 ligase ligand regions (A and B target CRBN, C targets VHL); in orange are the linker regions and in green are the POI ligands (A and C target BRD4, B targets androgen receptors)

The fundamental method of action for a PROTAC is to simultaneously interact with both a POI and an E3 ligase, bringing the two proteins in proximity to each other and triggering ubiquitination of the POI, causing the subsequent degradation of the POI.[4,5,6,7] (figure 3).

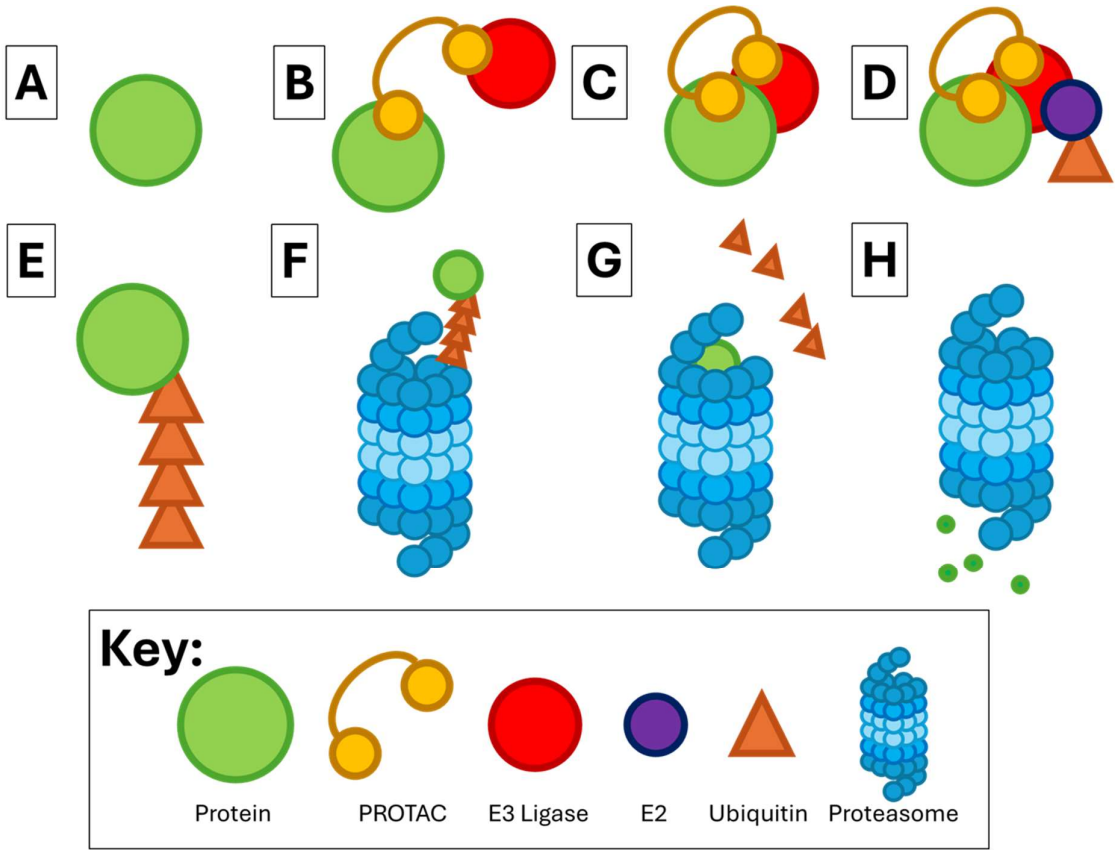


Figure 3: A PROTAC exploiting the Ubiquitin Proteasome Pathway. A) A protein B) A PROTAC binds to both the protein and an E3 ligase molecule C) The E3 ligase molecule and the protein are brought into close proximity by the PROTAC D) A ubiquitin-loaded E2 ubiquitin-conjugation enzyme coordinates with E3 ligase, and ubiquitinates the targeted protein. E) After multiple ubiquitinations, the targeted protein now has a poly-ubiquitin chain. F) The poly-ubiquitin chain is recognised by the proteasome. G) The ubiquitin subunits are cleaved, and the targeted protein is passed through the central pore of the proteasome. H) Once passed through the proteasome, the targeted protein is now cleaved into small peptide chains.

The PROTAC is not degraded and instead recycled, therefore behaving as a catalyst,[5,7] allowing one PROTAC to interact with many POIs, as opposed to conventional inhibitors which can only interact in a 1:1 relationship.[7] Another key advantage of PROTACs is that they do not need to enact any change in activity or function of the POI, it is enough to simply associate with it. This means that proteins without an active site (or with a difficult to access site) can still be targeted as long as you can bind to it, opening up many targets once considered undruggable.

PROTACs have been known to science for over 20 years,[7,8] originating as a useful tool for studying the role of proteins in a cell in situations where genetic knock-outs were either not possible or not desirable. In these instances a PROTAC would be used to degrade the protein of interest after translation had already occurred.[7] It was also hypothesised that it could also be used *in vivo* with the

same effect, potentially unlocking a therapeutic use[7,8] but the early PROTACs contained peptide based linkers which were deemed unsuitable for cell penetration.[8]

BRD4

A protein identified as of interest to cancer[9,10] (among other conditions[11,12]) is bromodomain-containing protein 4 (BRD4). BRD4 is a member of the bromodomain and extraterminal (BET) family of proteins, it is found in the nucleus and is known to have a role in transcriptional regulation[9,14] and histone acetylation[10].

BRD4 has been shown to be deregulated in many kinds of cancerous cells[9] and so is assumed to have a sizeable role in cancer proliferation. It has been shown to be vital in a recurrent translocation in a form of squamous cell carcinoma.[39] BRD4 is also responsible for recruitment of the transcription elongation factor P-TEFb,[40] which has been itself linked with cancer.[36] This makes BRD4 a protein with good potential as a therapeutic target for cancer.

BRD4 has been investigated in the past as a target of inhibition,[9,10,11,12,13,14,15] with drugs such as JQ1 and I-BET showing inhibitory effects, but has more recently been identified as a suitable candidate for PROTAC treatment, with one PROTAC (dBET1) showing strong *in vitro* anti-cancer behaviour.[10]

PROTAC Design

Designing a PROTAC for therapeutic use has typically involved a number of steps. One crucial step is to identify an appropriate POI, which will typically be a protein involved in the condition that is being treated, such as interleukin-1 receptor-associated kinase 4 for auto-immune diseases. In the case of cancer, proteins involved in the cell cycle or gene expression (many of which are already the targets of existing treatments) could be suitable, and PROTACs that target bromodomain-containing protein 9 (BRD9), signal transducer and activator of transcription 3 (STAT3) and androgen receptor (AR) are all currently being investigated in clinical trials.[7]

After a POI has been identified, a ligand for this protein must be chosen. An ideal ligand is specific to only the POI, and it should also have some ability to undergo chemical modification without impacting the affinity to the POI. For proteins that already have an inhibitor available on the market, this inhibitor is likely to be suitable but for other proteins this process may take some time as it requires either designing, synthesising and testing a new molecule capable of binding (no inhibitory effect required) or screening a catalogue of existing molecules for affinity. Both of which can be difficult.

As well as finding a ligand to the POI, a ligand to an E3 ligase must also be selected. There are potentially more than 600 E3 ligases encoded in the human genome,[7,16,17] but very few have been used for PROTACs, with the majority using either cereblon (CRBN) or Von Hippel-Lindau (VHL).[16] Ligands to these ligases are known and established. For CRBN, ligands include thalidomide and its derivatives (such as pomalidomide or lenalidomide) and VHL has molecules in the VH family, such as VH032. If a specific other E3 ligase is required for some reason, a different ligand would need to be found/developed but in most cases CRBN or VHL should be sufficient.

The final step in PROTAC design is to choose a linker to connect the two ligands. This step is in theory the most flexible, with different linkers used in different studies.[10,17] A linker simply needs to join the two other components together without impairing either ligands' affinity, causing insolubility or being too long to bring the two proteins within reasonable proximity. The lack of direct activity would suggest that any linker, as long as it met these parameters, would be acceptable. However, limited research has been done regarding the impact of linker selection, and therefore it cannot be ruled out that linker choice could have a significant impact on stability and function.

Project Aim

This project aims to look at different linkers to attempt to begin to establish a relationship between linker composition and connectivity, and pharmacokinetic properties. This is achieved by standardising the E3 ligase ligand and POI ligand and varying the linker between them. A CRBN ligand (a thalidomide

derivative, either 4-hydroxythalidomide or pomalidomide) is used as the E3 ligase ligand and the POI ligand will be a JQ1 derivative, which binds to BET family proteins, such as BRD4. Synthesis will be analysed primarily by liquid chromatography-mass spectrometry (LCMS) with the chromatography analysed by measuring absorbance over time through a photodiode array detector (PDA), which measures the UV absorbance over a range of wavelengths, over time. The mass spectrometry measured first by a total ion chromatogram (TIC), which measures the total number of mass detections over time, over the whole range of detectable masses, and individual timepoints measured by conventional mass spectrometry.

In order to investigate the stability of the PROTAC linker, this project will treat simulated blood and liver environments (as these are the metabolic environments associated with drug distribution and clearance, respectively – key aspects of the lifetime of a drug in the human body) with PROTAC molecules, and the relative amount of PROTAC molecule remaining over time will be monitored with LC-MS^2. The metabolic pathways will also be looked into to discover how the drug is treated by the body, with metabolites identified by high resolution LC-MS^2.

To simulate blood, pooled blood plasma from human donors will be used as an *in vitro* blood analogue. To replicate the liver, human hepatocytes will be used alongside human liver microsomes, which are small pieces of reconstituted hepatic endoplasmic reticulum (ER), used as the ER is the main metabolic region of hepatocytes.

In order to evaluate the effect on function of the PROTAC linker, this project will attempt to treat cancerous cell lines with developed PROTAC molecules and then look at BRD4 levels over time.

Materials

Reagents, Media and Buffers

Table 1: Reagents, media and buffers used in this project, with the supplier and catalogue number listed. Numbers in brackets are the molecules taken from figure 4 and 6 in the methods section.

Product	Supplier	Catalogue Number
4-hydroxy thalidomide (1)	BLDpharm	BD00756176
tert-butyl bromoacetate (2)	Fluorochem	F043685
N-Boc-1,4,-butanediamine (5)	BLDpharm	BD22164
Triethylamine	Sigma	471283
HATU	Apollo Scientific	PC8718
JQ1 carboxylic acid (8)	BLDpharm	BD630518
COMU	Fluorochem	F044578
N-Boc- 2,2'-(ethylenedioxy)diethylamine (10)	BLDpharm	BD117668
Pomalidomide	BLDpharm	BD235626
Succinic anhydride	Acros Organics	158760500
Succinyl chloride	Fluorochem	F099029
4-bromo-2-(2,6-dioxopiperidin-3-yl)isoindoline-1,3-dione	BLDpharm	BD01196684
4-(tert-Butoxycarbonyl)phenylboronic acid	Apollo Scientific	OR10392
RuPhos Pd G3	Johnson Matthey	BPC-306
Dulbecco's Phosphate Buffered Saline	Sigma	D8537
NADPH	Apollo Scientific	BIB3014
Williams' medium E	Sigma	W1878
RPMI-1640 medium	Gibco	11875093
4X NuPAGE LDS sample buffer	Invitrogen	NP0007
Rabbit anti-BRD4 Recombinant Monoclonal Antibody	Cambridge Biosciences	BL-149-2H5
Pierce ECL Western Blotting Substrate	Thermo Scientific	32209
Penicillin-Streptomycin (10X)	Gibco	11548876
NuPAGE Transfer Buffer (20X)	Invitrogen	NP0006

Matrices

Table 2: Matrices used in this project, with the supplier, catalogue number and lot number listed.

Matrix	Supplier	Product Number	Lot Number
Human Blood Plasma	BioIVT	HUMANPLK2-0110545	HMN1166527
Human Liver Microsomes	XenoTech	H2610 H2620 H2630 H2640	2310132
Human Hepatocytes	BioIVT	X008052	GYO

Cell Lines

A549 lung and MCF7 breast cancerous cell lines were generously donated by Professor Michelle Garrett from the University of Kent

Software

Table 3: Software used in this project, with the supplier and version listed.

Software	Supplier	Version
ChemDraw	PerkinElmer	22.2.0.3300
Chromeleon 7	Dionex	7.1.0.898 7.2 SR5 MUc (9717)
MassLynx	Waters	4.1 SCN 846
DiscoveryQuant	SCIEX	3.0.2.10914
Analyst	SCIEX	1.7.2
FreeStyle	Thermo Scientific	1.8.63.0
Xcalibur	Thermo Scientific	4.5.474.0

Methods

PROTAC Synthesis

PROTAC-1

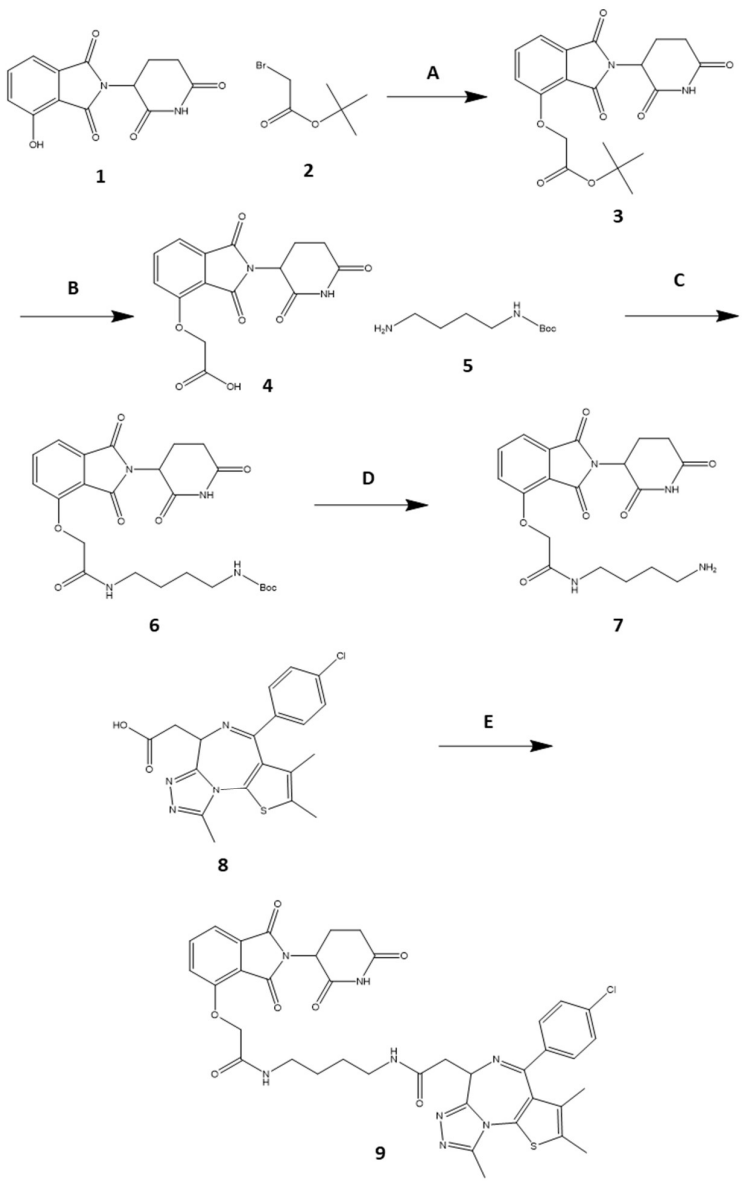


Figure 4: The synthetic route for PROTAC-1 with molecules labelled 1-9 and reactions labelled A-E

The molecules from figure 4 are as follows:

1. 4-hydroxy thalidomide
2. *tert*-butyl bromoacetate
3. *tert*-butyl 2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)oxy)acetate
4. 2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)oxy)acetic acid
5. *N*-Boc-1,4-butanediamine
6. *tert*-butyl (4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)oxy)acetamido)butyl)carbamate
7. *N*-(4-aminobutyl)-2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)oxy)acetamide
8. JQ-1 carboxylic acid
9. 2-(4-(4-chlorophenyl)-2,3,9-trimethyl-6*H*-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)-*N*-(4-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)oxy)acetamido)butyl)acetamide (Known industrially as dBET1, in this document known as **PROTAC-1**)

The reactions from figure 4 are as follows:

A. Molecules 1 (0.2 M) and 2 (0.21 M) along with potassium carbonate (K_2CO_3) (0.3 M) were combined in *N,N*-dimethyl formamide (DMF) at RT overnight. The reaction mixture was washed with water and the organic layer was extracted and dried over magnesium sulfate ($MgSO_4$), filtered and concentrated *in vacuo*. Yield by weight: 38%
B. Molecule 3 was dissolved in dichloromethane (DCM) (2ml) and trifluoro acetic acid (TFA) (1ml) was added and left at RT, overnight. The reaction mixture was then concentrated *in vacuo*. Yield by weight: >99%
C. Molecule 4 (60 mM) was dissolved in DMF (5ml), before adding triethylamine (180 mM) and then molecule 5 (72 mM) followed by HATU (84 mM). The reaction was kept at RT overnight before being concentrated *in vacuo*, re-dissolved in ethyl acetate, filtered and then washed

with water, with the organic layer being extracted and purified by silica chromatography (0-10% methanol in DCM) Yield by weight: 79%

D. Molecule 6 (100% of product from reaction C) was dissolved in DCM (2 ml) and TFA (0.5 ml)was added and left at RT, overnight. The reaction mixture was then concentrated *in vacuo*. Yield by weight: >99%

E. Molecule 7 (50 mM) was dissolved in DMF, before adding triethylamine (100 mM) and then molecule 8 (60 mM) followed by COMU (60 mM). The reaction was kept at RT overnight before being concentrated *in vacuo*, re-dissolved in ethyl acetate, filtered and then washed with water, with the organic layer being extracted and purified by silica chromatography (0-10% methanol in DCM) Yield by weight: 87%. Predicted *m/z* [$M+H^+$] was 785.2267, found 785.2257 (ppm error: 1.27).

After synthesis, molecule 9 was dissolved in acetonitrile and stored at room temperature.

PROTAC-2

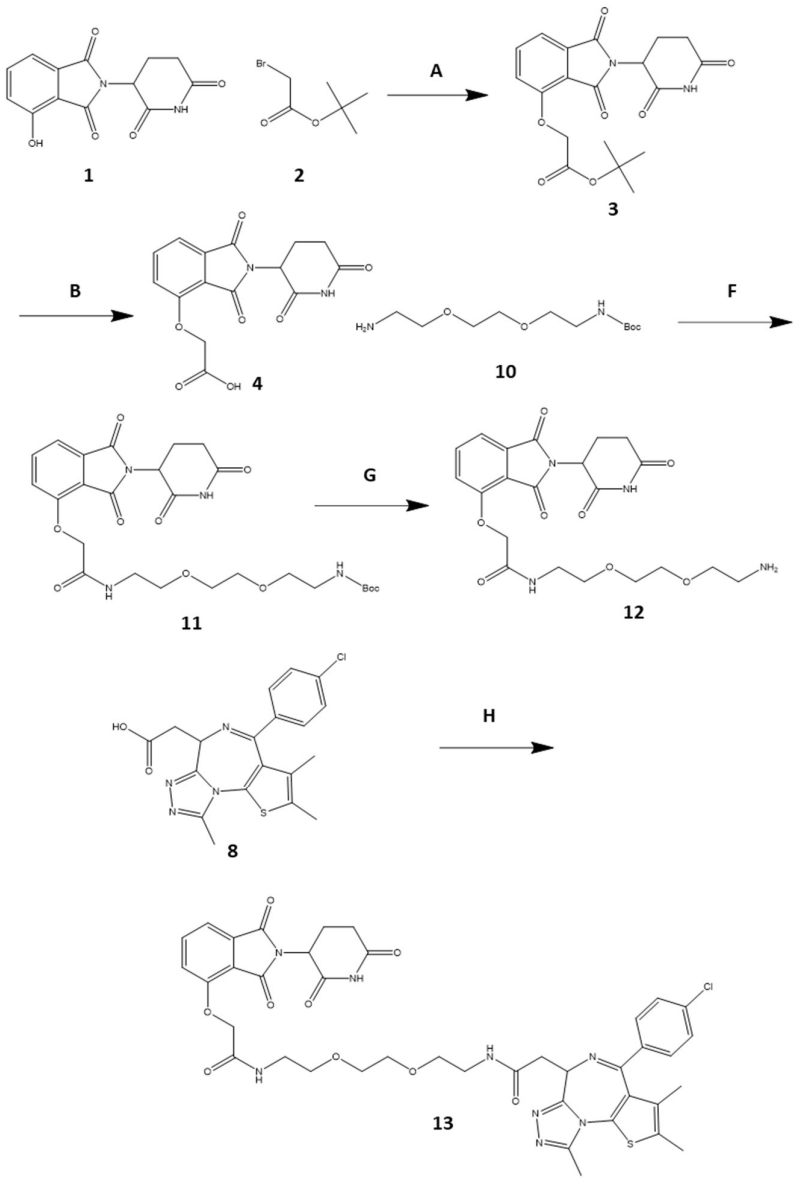


Figure 5: The synthetic route for PROTAC-2 with molecules labelled 1-4, 8 and 10-13 and reactions labelled A, B, F, G and H

The molecules 1-4 and 8 from figure 5 are the same as listed in figure 4, the remaining (10-13) are as follows:

10. *N*-Boc- 2,2'-(ethylenedioxy)diethylamine
11. *tert*-butyl (2-(2-(2-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)oxy)acetamido)ethoxy)ethoxy)ethyl)carbamate
12. *N*-(2-(2-(2-aminoethoxy)ethoxy)ethyl)-2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)oxy)acetamide
13. 2-(4-(4-chlorophenyl)-2,3,9-trimethyl-6*H*-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)-*N*-(2-(2-(2-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)oxy)acetamido)ethoxy)ethoxy)ethyl)acetamide (In this document known as **PROTAC-2**)

The reactions A and B from figure 5 are the same as for PROTAC-1, the remaining reactions are as follows:

F. Molecule 4 (60 mM) was dissolved in DMF (5ml), before adding triethylamine (180 mM) and then molecule 5 (72 mM) followed by HATU (84 mM). The reaction was kept at RT overnight before being concentrated *in vacuo*, re-dissolved in ethyl acetate, filtered and then washed with water, with the organic layer being extracted and purified by silica chromatography (0-10% methanol in DCM) Yield by weight: 81%

F. Molecule 11 was dissolved in DCM and TFA was added to a DCM:TFA ratio of 4:1 and left at RT, overnight. The reaction mixture was then concentrated *in vacuo*. Yield by weight: >99%

G. Molecule 12 (50 mM) was dissolved in DMF, before adding triethylamine (100 mM) and then molecule 8 (60 mM) followed by COMU (60 mM). The reaction was kept at RT overnight before

being concentrated *in vacuo*, re-dissolved in ethyl acetate, filtered and then washed with water, with the organic layer being extracted and purified by silica chromatography (0-10% methanol in DCM) Yield by weight: 92% Predicted *m/z* [$M+H^+$] was 845.2479, found 845.2468 (ppm error: 1.30).

After synthesis, molecule 13 was dissolved in acetonitrile and stored at room temperature.

LC-MS Methods

Synthesis Clarification

Samples were dissolved in methanol (1mg/ml) before being analysed by utilising a Micromass ZQ mass spectrometer (Waters), A 2695 Separations Module (waters) and a 2996 PDA detector (Waters). The mass spectrometer was set to operate in positive ion mode.

Tuning and Clearance Assays

Samples were analysed utilising a API 5000 mass spectrometer (Sciex), 1110/1290 LC pumps (Agilent), 1290 column oven (Agilent), and a HTS PAL autosampler (CTC Analytics). The mass spectrometer was set to operate in positive ion mode. All analytes were monitored using specific MRM transitions. A two-dimensional LC system was utilised, which allowed samples to be injected onto a loading column and back-flushed using gradient elution onto an analytical column for MS analysis.

Metabolite Identification

Samples were analysed by LC-HRMS utilizing an Orbitrap Exploris 120 mass spectrometer, Vanquish pumps, column manager, PDA detector and autosampler (Thermo Fisher Scientific). The mass spectrometer was set to operate in both positive and negative ion modes and analytes were monitored using full scan analysis with data dependent MS^2.

Stability and Metabolism

Human Blood Plasma Clearance

Human Blood Plasma prior to use was thawed at 37 °C and adjusted to a pH of 7.4 (±0.05). Plasma (495 µL/well) were pre-warmed at 37 °C for 5 minutes. 5 µl of test compound solutions (100 µM) were spiked into microsomes to a final concentration of 1 µM and incubated for 1 hour at 37 °C +5% CO_2. Compounds were tested in triplicate. Aliquots (50 µL) of the incubations were taken at 9 time points (0, 3, 6, 10, 30, 45, 60, 90 and 120 minutes) and treated with ice-cold acetonitrile containing propranolol as an internal standard (200 µL). In addition, a Dulbecco's Phosphate Buffered Saline (DPBS) was run in parallel. At the end of the incubation, crashed samples were centrifuged at 2000 xg and 4 °C for 10 minutes, after which the supernatants were analysed by LC-MS. Test compound clearance was measured by determining the peak areas of the test compound compared to the internal standard, and the percentage of parent peak area (taken from the timepoint t=0) was plotted against time. The elimination constant (*k*) was determined by the gradient of ln(percent parent peak area) against time. The half-life was calculated by equation 1

$$(1)\text{Halflife } (min) = -\frac{\ln(2)}{k}$$

Human Liver Microsome Clearance

Human liver microsomes (HLMs) were thawed on ice prior to use. HLMs (495 µL/well; 0.5 mg/mL; mixed gender) containing magnesium chloride ($MgCl_2$) (5 mM) and NADPH (1 mM) in phosphate buffer (pH 7.4) were pre-warmed at 37°C for 5 minutes. 5 µl of test compound solutions (100 µM) were spiked into microsomes to a final concentration of 1 µM and incubated for 1 hour at 37 °C +5% CO_2. Compounds were tested in triplicate. Aliquots (50 µL) of the incubations were taken at 8 time points (0, 3, 5, 10, 15, 30, 45, and 60 minutes) and treated with ice-cold acetonitrile containing propranolol as an internal standard (150 µL). In addition, a no-cofactor (absence of NADPH) control was run in

parallel. At the end of the incubation, crashed samples were centrifuged at 2000 xg and 4 °C for 10 minutes, after which the supernatants were analysed by LC-MS. Test compound clearance was measured by determining the peak areas of the test compound compared to the internal standard, and the percentage of parent peak area (taken from the timepoint t=0) was plotted against time. The elimination constant (*k*) was determined by the gradient of ln(percent parent peak area) against time. The half-life was calculated by equation 1. The intrinsic clearance (Cl_{int}) was calculated using equation 2 where mass protein = 0.5 mg and volume = 1000 µl

$$(2)\ \text{Halflife}\ (min) = -\frac{k}{\text{mass protein/volume}}$$

Human Hepatocyte Clearance

Human hepatocytes were thawed on ice prior to use. Hepatocytes (495 µL/well; $0.5x10^6$ cells/ml; mixed gender in Williams' medium E) were pre-warmed at 37 °C for 5 minutes. 5 µl of test compound solutions (100 µM) were spiked into the hepatocyte mixture to a final concentration of 1 µM and incubated for 1 hour at 37 °C +5% CO_2. Compounds were tested in triplicate. Aliquots (50 µL) of the incubations were taken at 9 time points (0, 3, 5, 10, 15, 30, 60, 90, and 120 minutes) and treated with ice-cold acetonitrile containing propranolol as an internal standard (150 µL). In addition, a heat-inactivated hepatocyte control was run in parallel. At the end of the incubation, crashed samples were centrifuged at 2000 xg and 4 °C for 10 minutes, after which the supernatants were analysed by LC-MS. Test compound clearance was measured by determining the peak areas of the test compound compared to the internal standard, and the percentage of parent peak area (taken from the timepoint t=0) was plotted against time. The elimination constant (*k*) was determined by the gradient of ln(percent parent peak area) against time. The half-life was calculated by equation 1. The intrinsic

clearance (Cl_{int}) was calculated using equation 3 where number of cells = $0.5x10^6$ cells and volume = 1000 µl

$$(3)\ \text{Halflife}\ (min) = -\frac{k}{\text{number of cells/volume}}$$

Human Blood Plasma Metabolite Identification

Human Blood Plasma prior to use was thawed at 37°C and adjusted to a pH of 7.4 (±0.05). Plasma (495 µL/well) was pre-warmed at 37°C for 5 minutes. Test compound solutions (1 mM) were spiked into plasma (5 µL; 10 µM final concentration;) and incubated for 2 hours at 37°C +5% CO_2. Aliquots (100 µL) of the incubations were taken at 2 time points (0 and 120 minutes) and treated with ice-cold acetonitrile (400 µL). In addition, blank plasma and no plasma control (DPBS) were run in parallel. At the end of the incubation, crashed samples were centrifuged at 2000xg and 4°C for 10 minutes, after which the supernatants were analysed by LC-HRMS. Peaks were analysed by mass signature on MS and MS^2 to identify possible routes of primary and secondary metabolism.

Human Liver Microsome Metabolite Identification

Human liver microsomes (HLMs) were thawed on ice prior to use. HLMs (495 µL/well; 0.5 mg/mL; mixed gender) containing MgCl2 (5 mM) and NADPH (1 mM) in phosphate buffer (pH 7.4) were pre-warmed at 37 °C for 5 minutes. 5 µl of test compound solutions (1 mM) were spiked into microsomes to a final concentration of 10 µM and incubated for 30 minutes at 37 °C +5% CO_2. Aliquots (100 µL) of the incubations were taken at 3 time points (0, 5 and 30 minutes) and treated with ice-cold acetonitrile (400 µL). In addition, blank HLMs and no cofactor control were run in parallel. At the end of the incubation, crashed samples were centrifuged at 2000 xg and 4°C for 10 minutes, after which the supernatants were analysed by LC-HRMS. Peaks were analysed by mass signature on MS and MS^2 to identify possible routes of primary and secondary metabolism.

Cell Assays

Tissue Culture

Cancer cell lines A549 (lung) and MCF7 (breast) were grown in RPMI-1640 media (1X) with 10% heat-inactivated fetal bovine serum and penicillin-streptomycin (1000 units/ml) at 37 °C with 5% CO_2 in a humidified incubator. Cells were passaged every 3-4 days.

Compound Treatment and Cell Harvesting

500,000 of A549 cancer cells per 3 ml of media were seeded into each well of a 6 well plate and incubated overnight at 37 °C with 5% CO_2 in a humidified incubator, cells were treated with 167 µl of 20X test compounds in DMSO for final concentrations of 100, 50, 20 and 10 nM with a final DMSO concentration of approximately 5%. A DMSO only sample and an untreated control were run in parallel. The treated cells were incubated for 4 hours at 37 °C with 5% CO_2 in a humidified incubator. After the time had elapsed, the media was removed and the cells were washed with PBS before being lysed with 100 µl of 2X NuPAGE LDS Sample buffer + 5% β-mercaptoethanol. Cell lysates were then boiled for 10 minutes at 95 °C before being cooled at room temperature before freezing at -20 °C until SDS-PAGE could be run.

SDS-PAGE and Western Blot

Samples were thawed at room temperature prior to boiling for 3 minutes at 95°C before being loaded in a 10% Tris-Gly SDS-PAGE gel. Gels were run at 125 V in 1x Tris-Gly running buffer (25mM tris, 200mM glycine, 5mM SDS). 2 gels were run per sample to enable comparison between BRD4 and β-Actin levels. After running the gels, the protein was transferred onto a nitrocellulose membrane at 0.75 A for 1 hour via wet transfer in 1X NuPAGE transfer buffer. The resulting membrane was blocked with 5% skimmed milk powder in TBS for 45 minutes before incubating at 4 °C overnight in the presence of either rabbit

anti-BRD4 (1:1000) or mouse anti-β-Actin (1:5000) primary antibodies. After washing, the membranes were then incubated in anti-rabbit (1:10000) or anti-mouse (1:10000) secondary antibodies, both containing horseradish peroxidase (HRP) tags, which were visualised with Pierce ECL Western Blotting Substrate, which is a mixture consisting of a peroxide and a luminol solution, which were combined prior to use. This mixture reacts with the HRP tags, emitting light which was captured on film.

Results

Synthesis

Initially, several potential targets were identified for investigation, with varying connectivity and linker properties (figure 6).

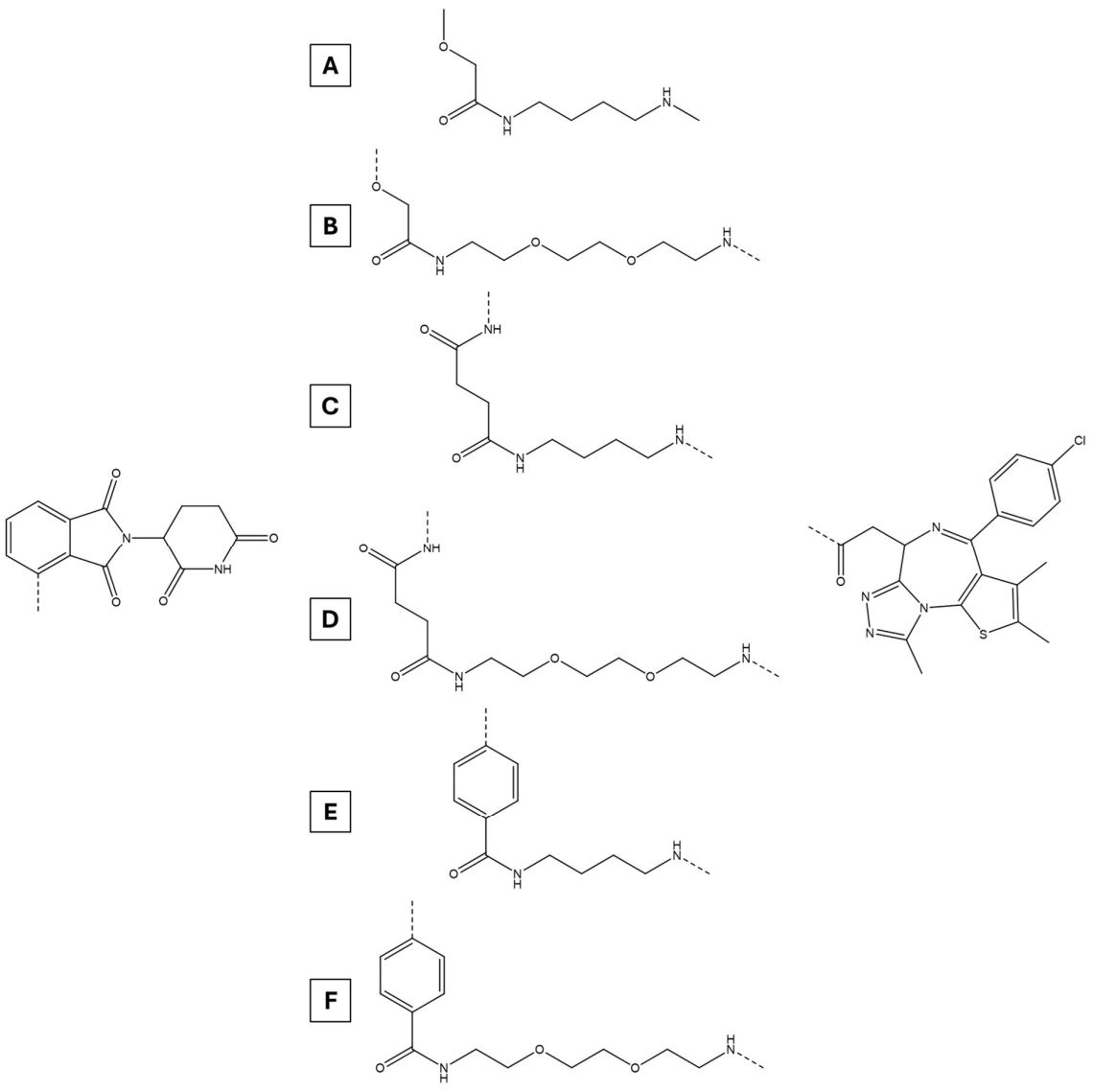


Figure 6: Left: The CRBN ligand. Centre: Six potential linkers, labelled A-F. Right: The BRD4 ligand

The linkers were chosen with the intention of varying the connection to the CRBN ligand as well as length and hydrophobicity. Linkers A and B would be connected by an ether, C and D would be connected by an amide and E and F would be connected by a biphenyl. A, C and E would have an aliphatic linker, and B, D and F would have an ethylene glycol linker.

Ether-Connected Linkers

A route was proposed to synthesise the ether-connected linkers (figure 7). Due to structural similarities between the molecules the first step is shared between the two PROTACs, and although the routes diverge after that point, the reactions involved are similar and require essentially identical conditions.

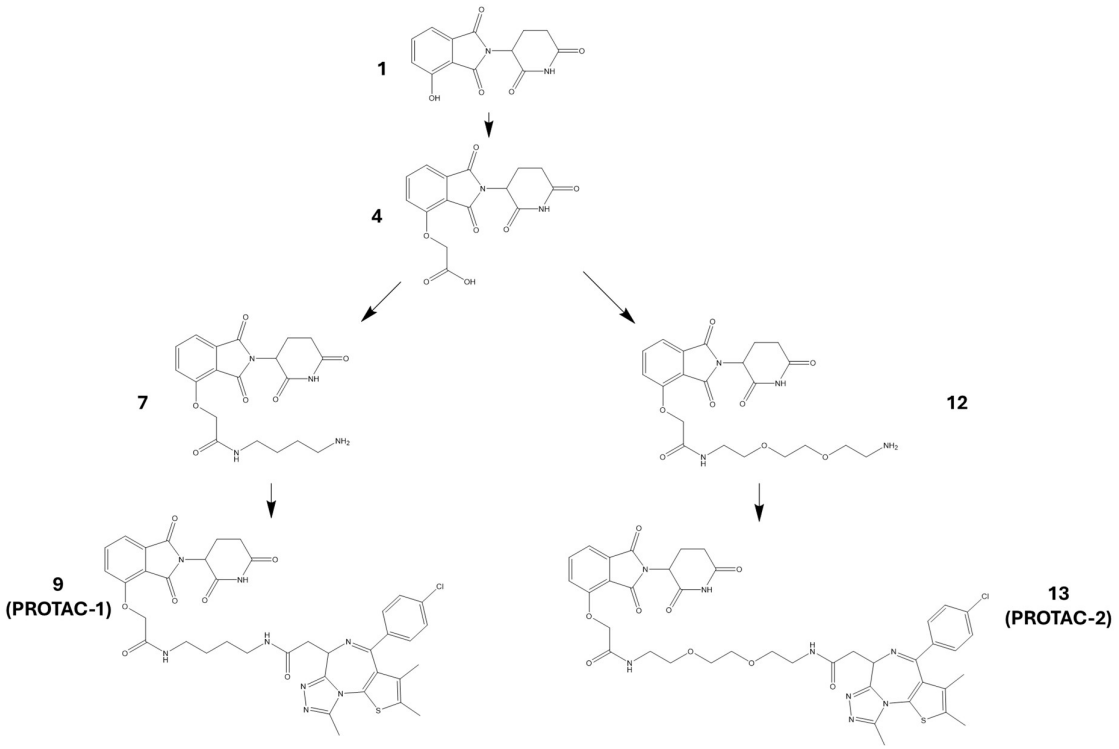


Figure 7: An overview of the steps involved in the synthesis of the ether-connected PROTAC molecules.

The first reaction undertaken was a Williamson ether synthesis. (figure 4, reaction A). This reaction was not the highest yield, and resulted in a substantial quantity of dark green precipitate, as well as a lighter green solution. The precipitate was effectively removed by liquid-liquid extraction, with the

precipitate (and the remaining green colour) removed in the aqueous layer. The organic and aqueous layers were analysed by LC-MS, with the total ion chromatogram (TIC) measuring the sum of all ions being detected by the mass spectrometer, and the photodiode array detector (PDA) measuring the total absorbance over a range of UV wavelengths. From this data, the product was found to be in the organic layer (table 4; figures 8 and 9) with a peak associated with an in-source deprotection on the MS and the precipitate (and the remaining green colour) was removed in the aqueous layer.

Table 4: The predicted masses of molecule 3 from figure 4 and an in-source deprotected molecule C.

Molecule	Predicted Exact Mass of Molecule + H^+ (Da)
Molecule 3	389.1343
Molecule 3 minus *tert*-butyl protecting group	332.0629

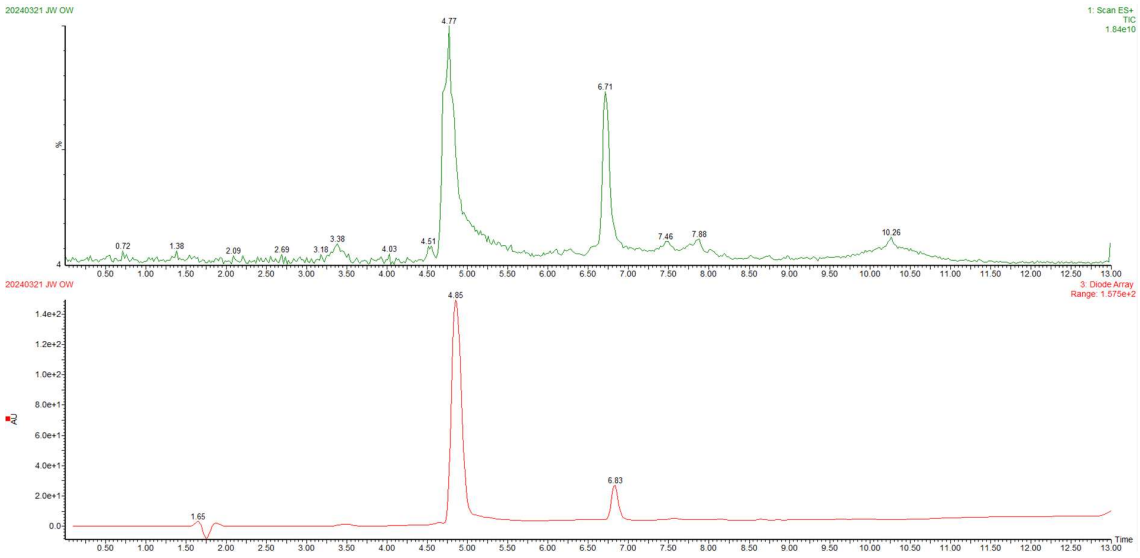


Figure 8: The total ion chromatogram (TIC) (green) and photodiode array detection PDA (red) of a sample taken from the organic layer of unpurified molecule 3.

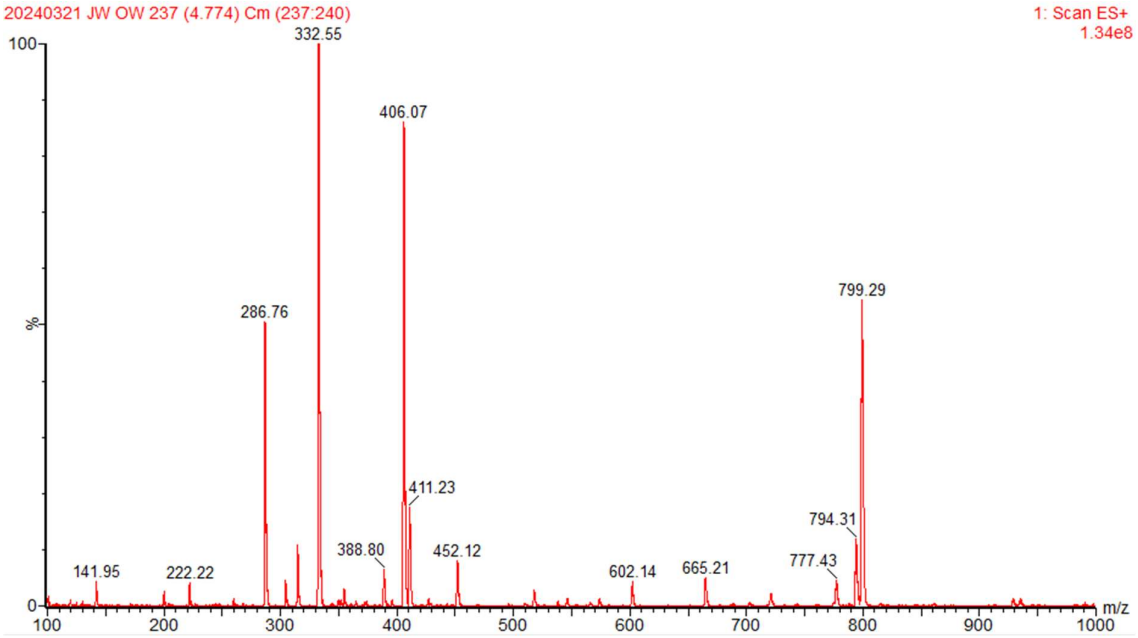


Figure 9: The MS of the peak at 4.77 minutes of a sample taken from the organic layer of unpurified molecule 3

The next reaction was a deprotection step (figure 4, reaction B). *tert*-butyl was used as a protecting group for this reaction as it can be removed using acid hydrolysis, which (unlike base hydrolysis) would prevent the imide from opening.

At this point the methodology split in two, with each following similar reactions but differing in the choice of diamide. The next steps were amide couplings (figure 4, reaction C and figure 5, reaction F) which utilised HATU as a coupling agent. To prevent unwanted interactions, similarly to the Williamson step these reactions used protecting groups. Table 5 along with figures 10, 11, 12 and 13 show the analysis of these reactions by LC-MS, which appear to have been deprotected by hydrolysis during the chromatography (which used H_2O as one of the mobile phases).

Table 5: The predicted masses of molecule 6 from figure 4 and molecule 11 from figure 5 along with the masses of deprotections of both molecules, molecules 7 and 12, after charging with the addition of a proton.

Molecule	**Predicted Exact Mass of Molecule + H^+ (Da)**
Molecule 6	503.2136
Molecule 7	403.1612
Molecule 11	563.2348
Molecule 12	463.1823

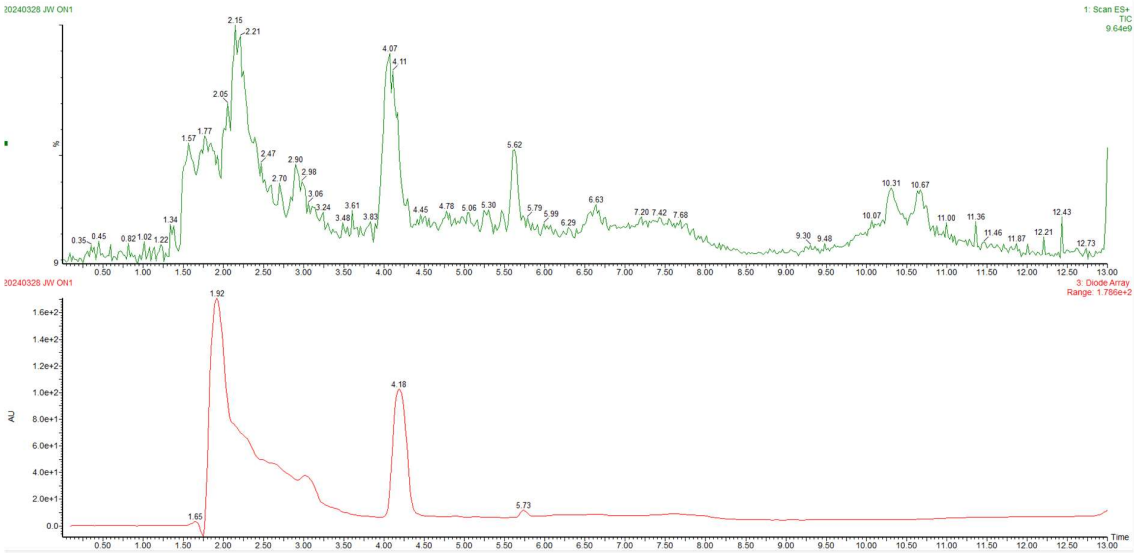


Figure 10: The TIC (green) and PDA (red) of a sample of unpurified molecule 6

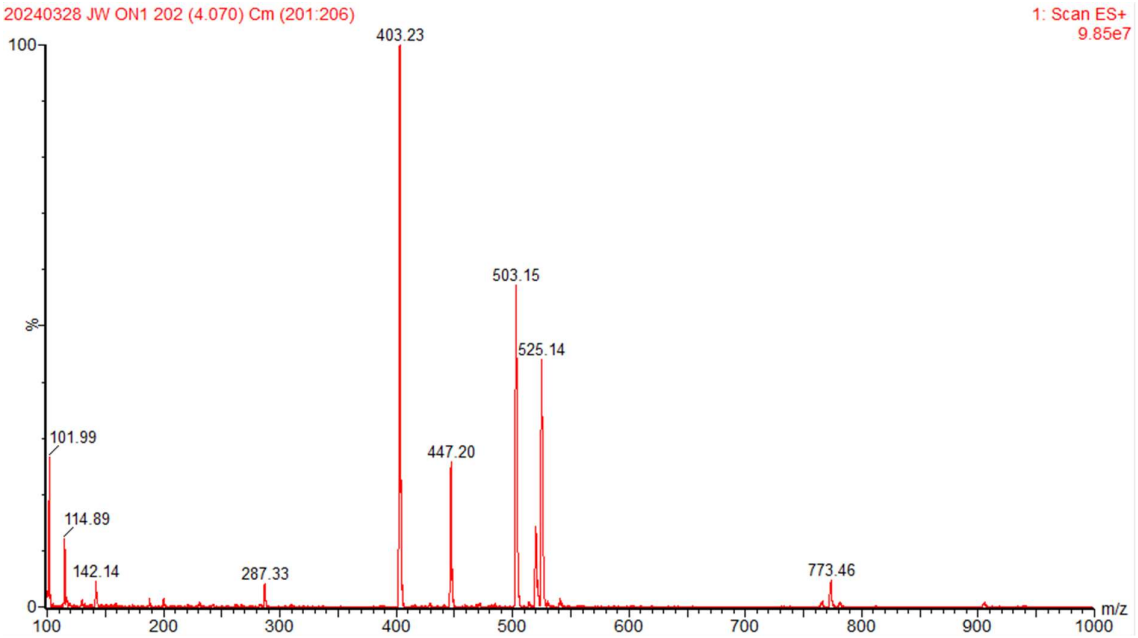


Figure 11: The MS of the peak at 4.07 minutes of a sample of unpurified molecule 6

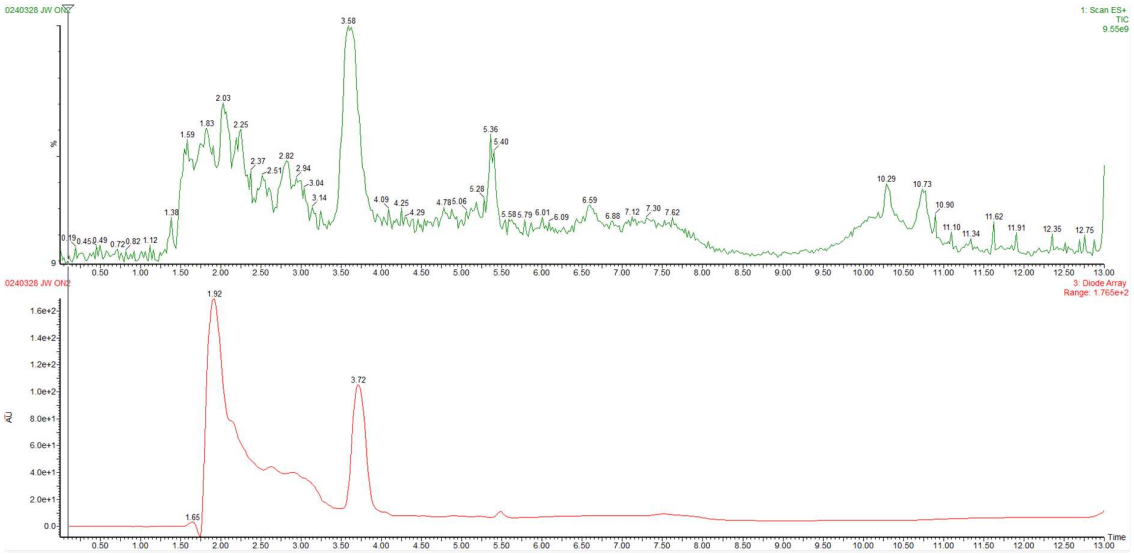


Figure 12: The TIC (green) and PDA (red) of a sample of unpurified molecule 11

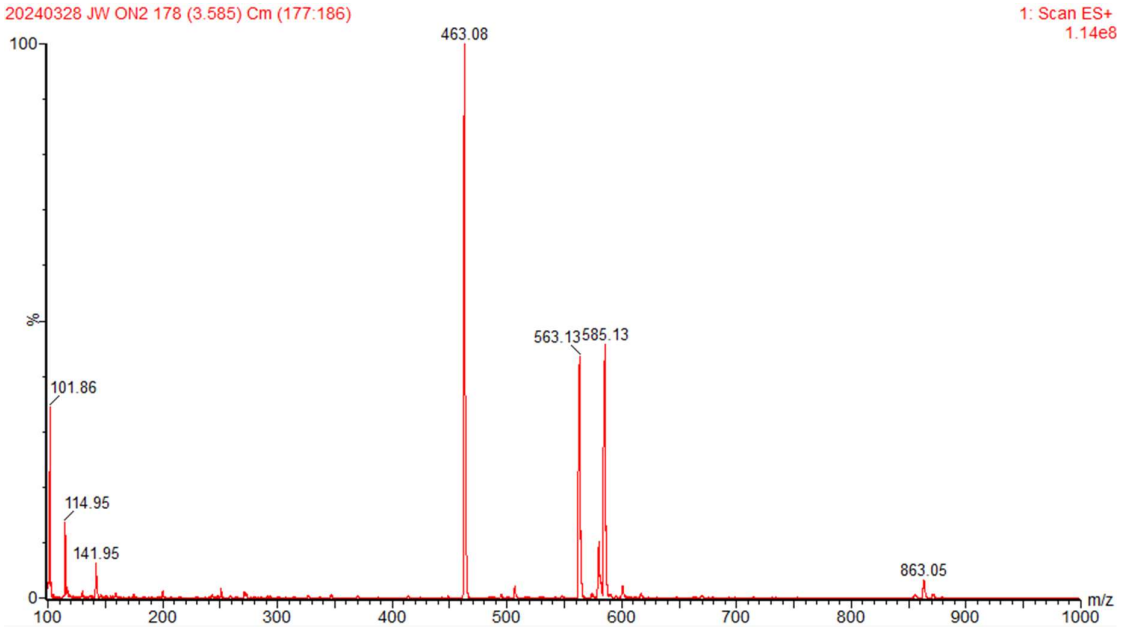


Figure 13: The MS of the peak at 3.58 minutes of a sample of unpurified molecule 11

The Boc groups were used which were removed by acid in a similar manner to the *tert*-butyl groups before, although the acid was more diluted in this instance (figure 4, reaction D and figure 5, reaction G).

For the last step on each molecule (figure 4, reaction E and figure 5, reaction F) another amide coupling was required. In this instance, COMU was used over HATU, as COMU has lower risks associated with use (COMU was not available for the previous step).

High resolution LC-MS^2 was used to confirm the synthesis of these molecules. Figure 14 shows the photodiode array (PDA) and an extracted ion chromatogram (EIC) (set to a broad range of 100-1250 Da) of a purified PROTAC-1 sample.

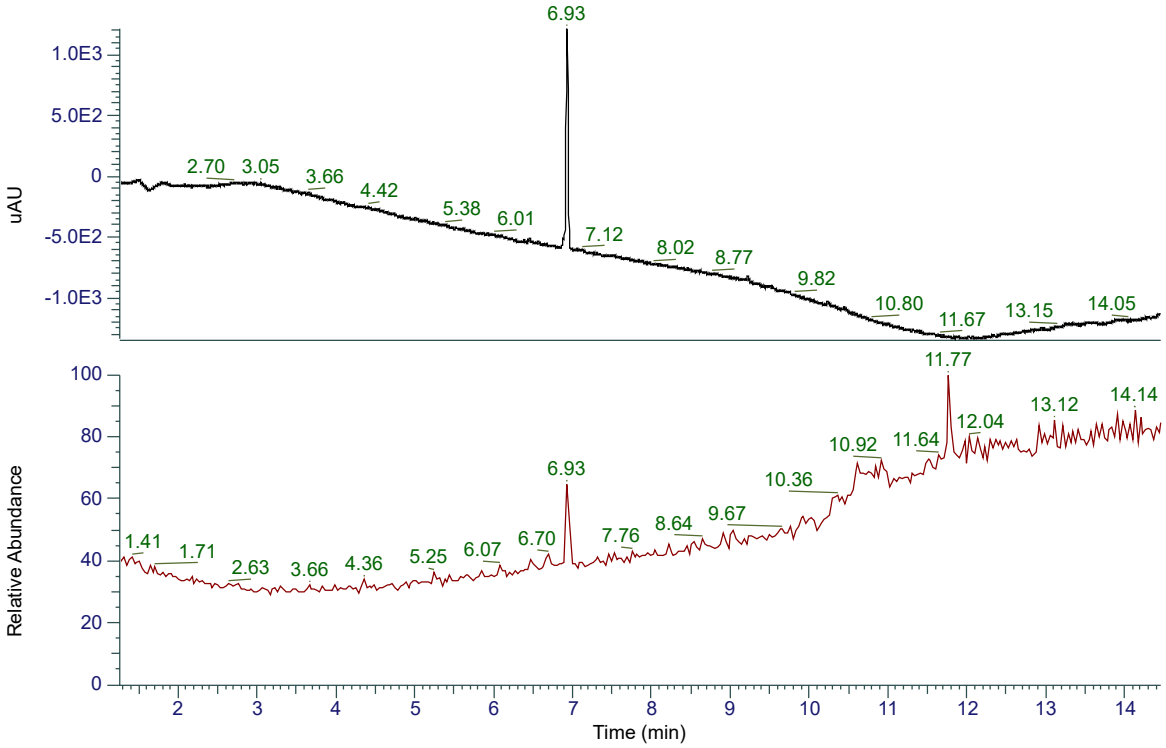


Figure 14: The PDA (300-350nm; black) and EIC (100-1250; red) output of a pure PROTAC-1 sample.

The PDA shows one peak at 6.93, whereas the EIC showed two peaks, one at 6.93 and another at 11.77. The MS of the peak at 6.93 (figure 15) shows an observed mass of 785.2257, which is extremely close to the calculated mass of 785.5567 (1.27 ppm).

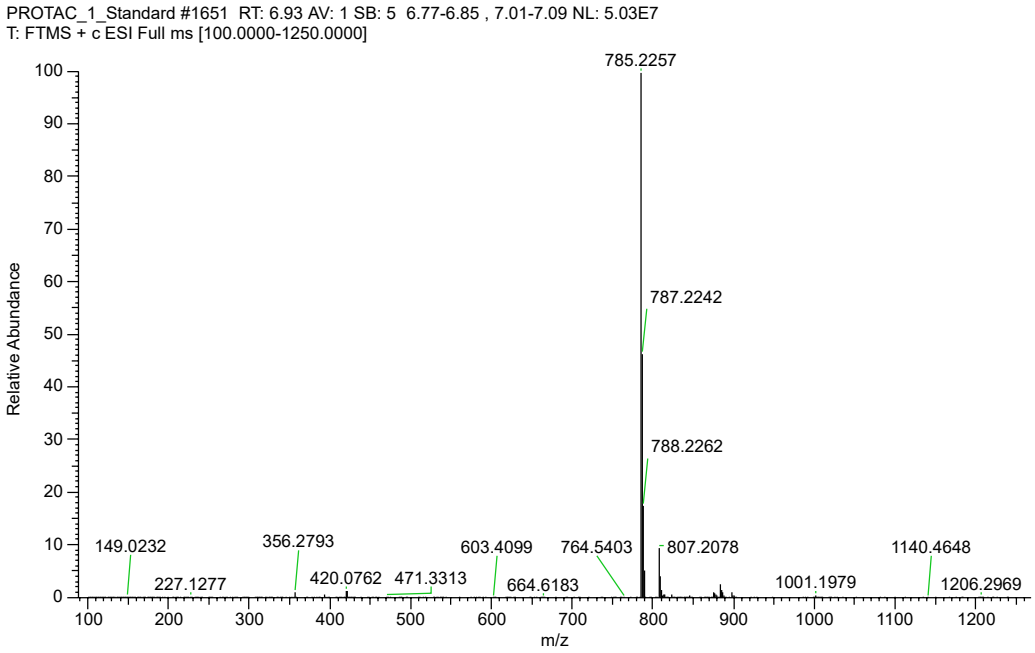


Figure 15: The background-subtracted MS of a pure PROTAC-1 sample at 6.93 minutes.

In order to further support that the 6.93 peak did correspond to PROTAC-1, MS^2 of the 785.2257 peak was undertaken (figure 16).

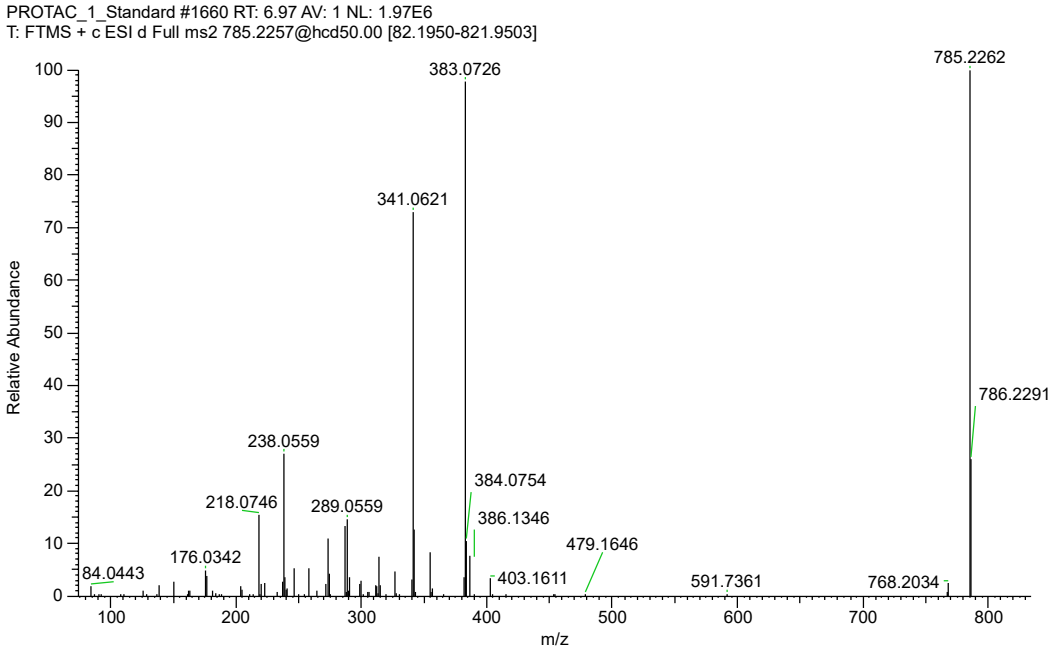


Figure 16: The MS^2 of the 785.2257 peak at 6.93 minutes

The MS^2 showed strong peaks at 383.0726 and 341.0621. These are close in mass to the predicted fragments of 383.0733 and 341.0628 (ppm error of 1.83 and 2.05, respectively) shown in figure 17. Another minor peak was identified at 289.0559, which is quite close to the predicted fragment of 289.0668 but has a ppm error of 37.71 which is well outside the stated accuracy of the instrument (<0.3 ppm) so cannot be used to strongly support the structure of the molecule.

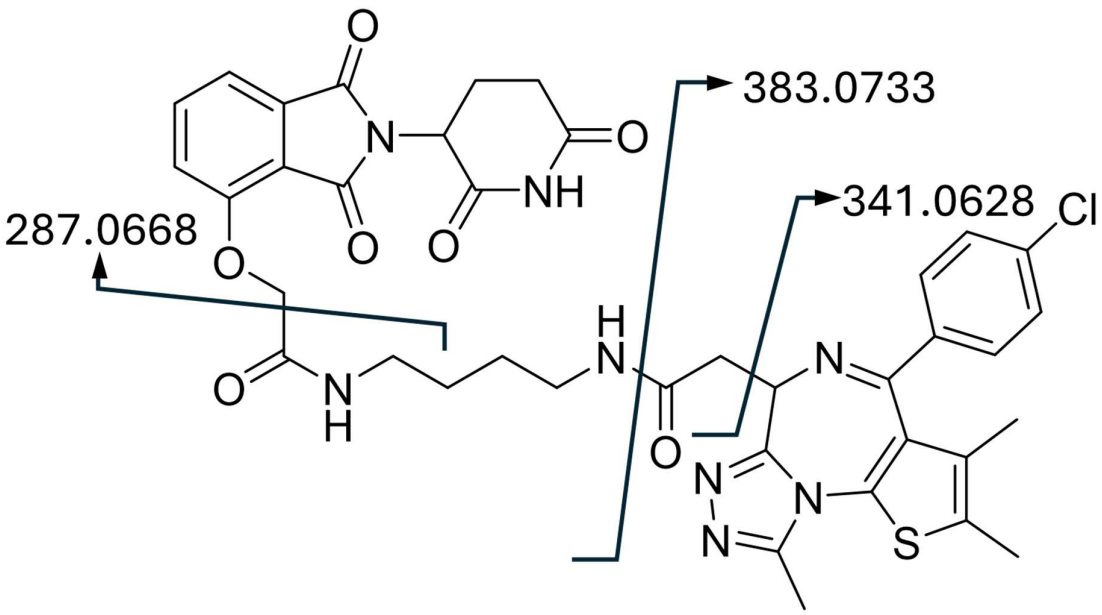


Figure 17: Predicted fragmentation points of PROTAC-1 with the calculated masses of produced fragments.

Any other peaks visible in the MS^2 of PROTAC-1 were not easily identifiable. Possible fragments could be identified (typically requiring two or more fragmentation events to produce) that had calculated masses that were somewhat close to the observed masses, but none near close enough for any certainty.

The same process was repeated for PROTAC-2, with similar results. The PDA and EIC (figure 18) looked largely similar to the ones for PROTAC-1, with the elution peaks now being at 6.90 (PDA) and 6.91 (EIC).

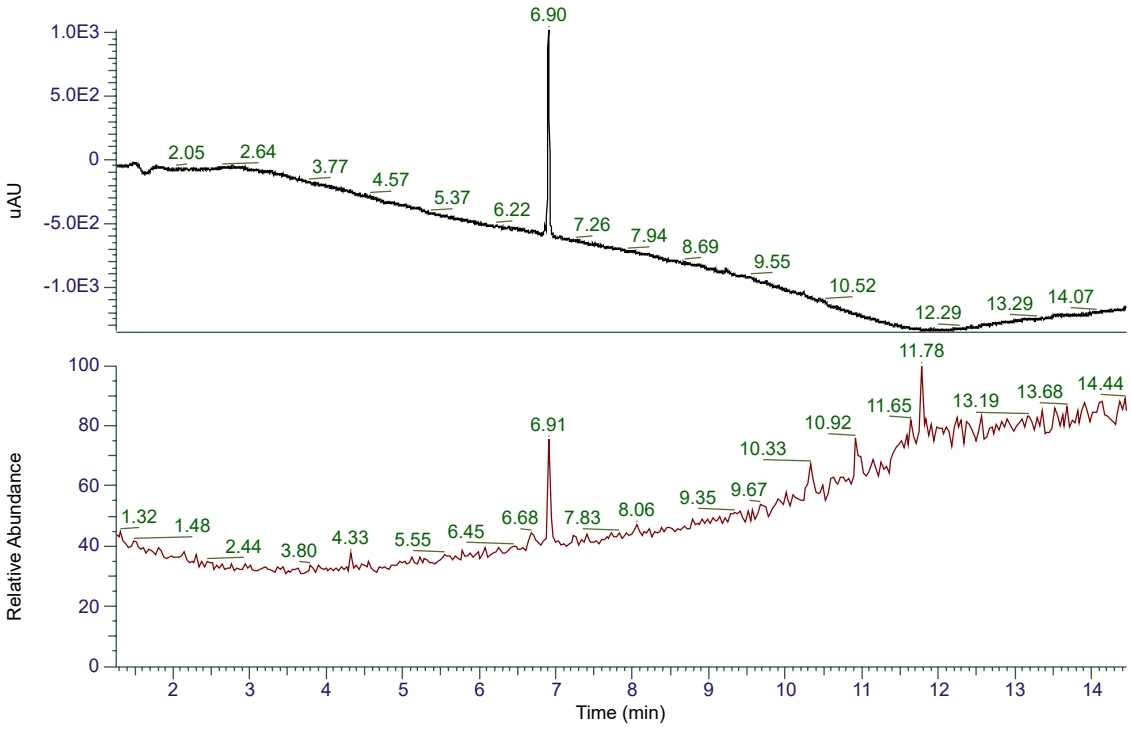


Figure 18: The PDA (300-350nm; black) and EIC (100-1250; red) output of a pure PROTAC-1 sample.

In order to confirm that the peak was associated with PROTAC-2, high resolution MS was taken at that timepoint (figure 19). This showed a mass of 845.2468, which aligns with the calculated mass of 845.2479 (ppm error of 1.30). Figure 20 shows the MS^2 of this mass peak.

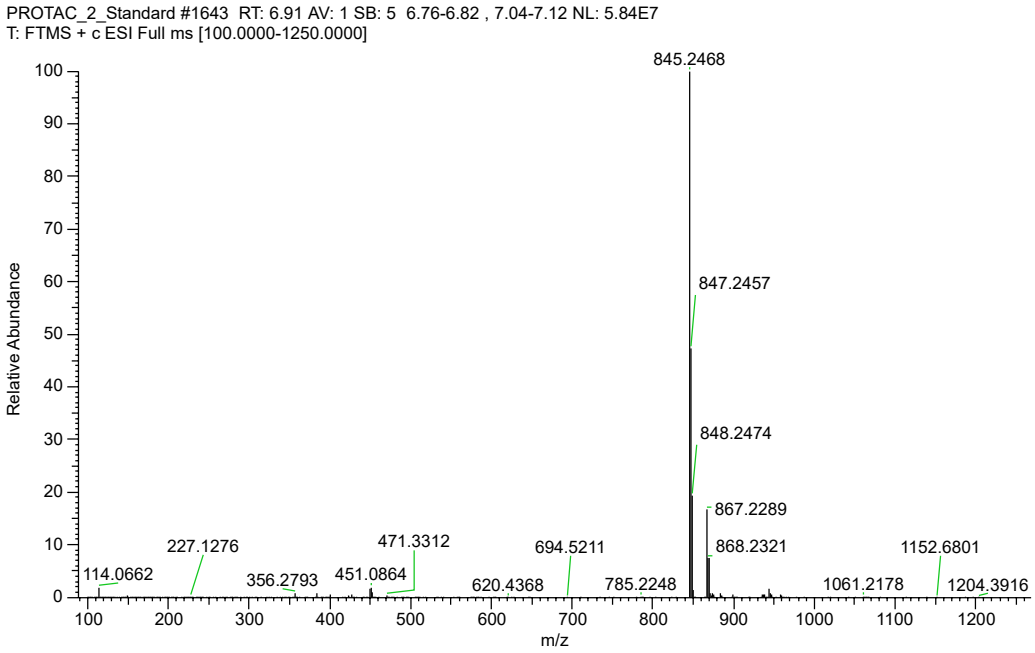


Figure 19: The background-subtracted MS of a pure PROTAC-1 sample at 6.91 minutes.

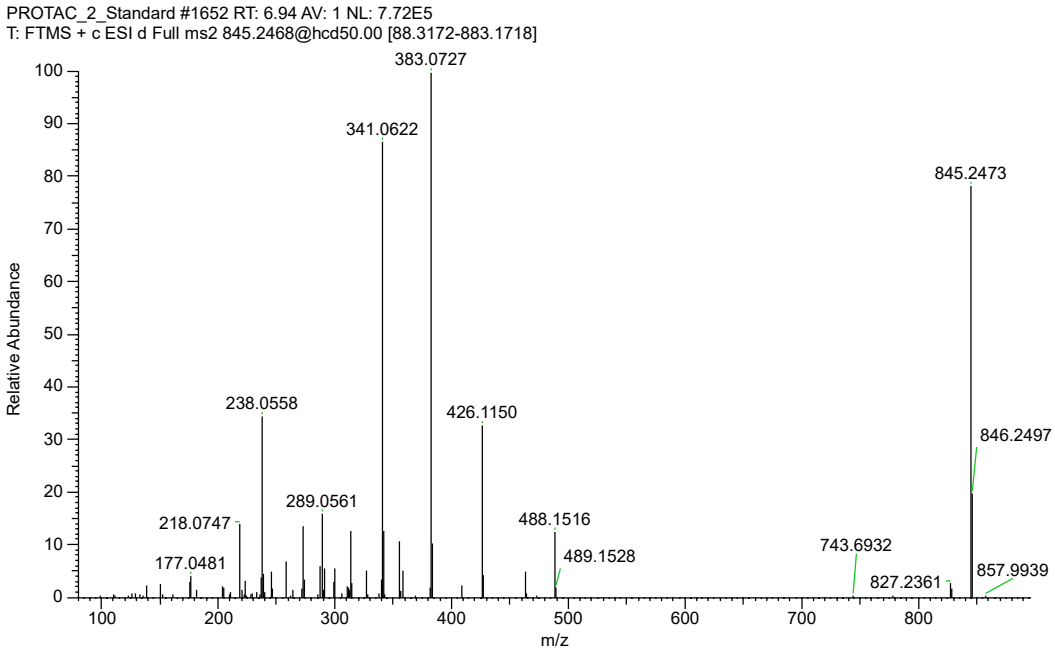


Figure 20: The MS² of the 845.2468 peak at 6.93 minutes

This MS² showed many of the familiar peaks observed in the MS² of PROTAC-1, as well as an additional peak at 426.1150. This is very close to the calculated mass of 426.1155 (ppm error of 1.17) of a predicted fragment that was not seen in PROTAC-1 (figure 21).

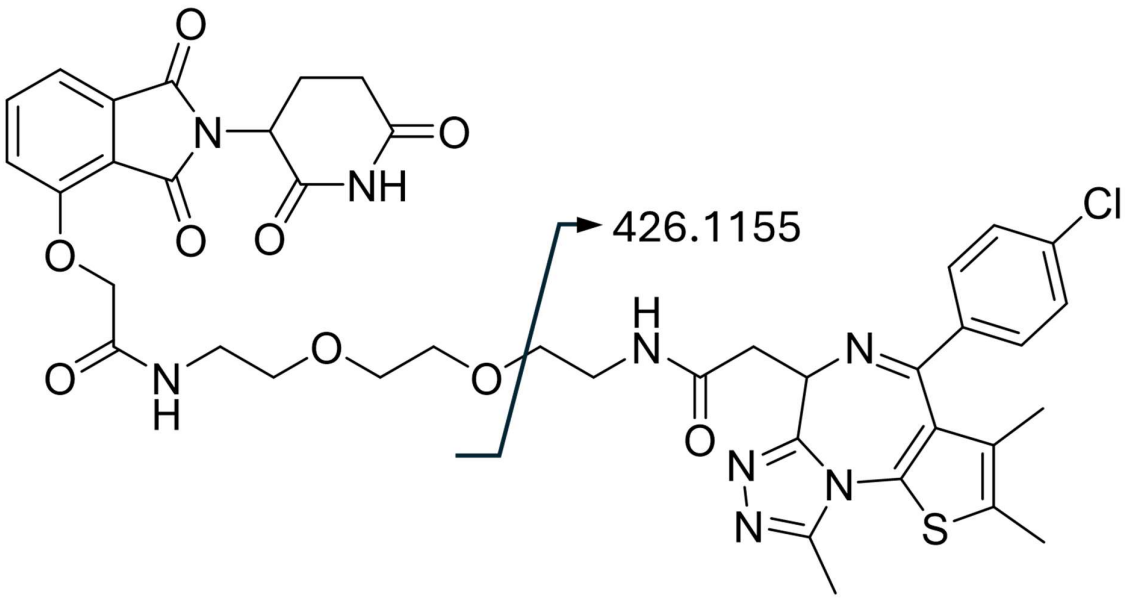


Figure 21: Predicted fragmentation point of PROTAC-2 with the calculated mass of a produced fragment.

Amide-Connected Linkers

In much the same way as for the ether-connected linkers, a pathway was set up for two amide connected linkers (figure 22) that started with a shared initial step.

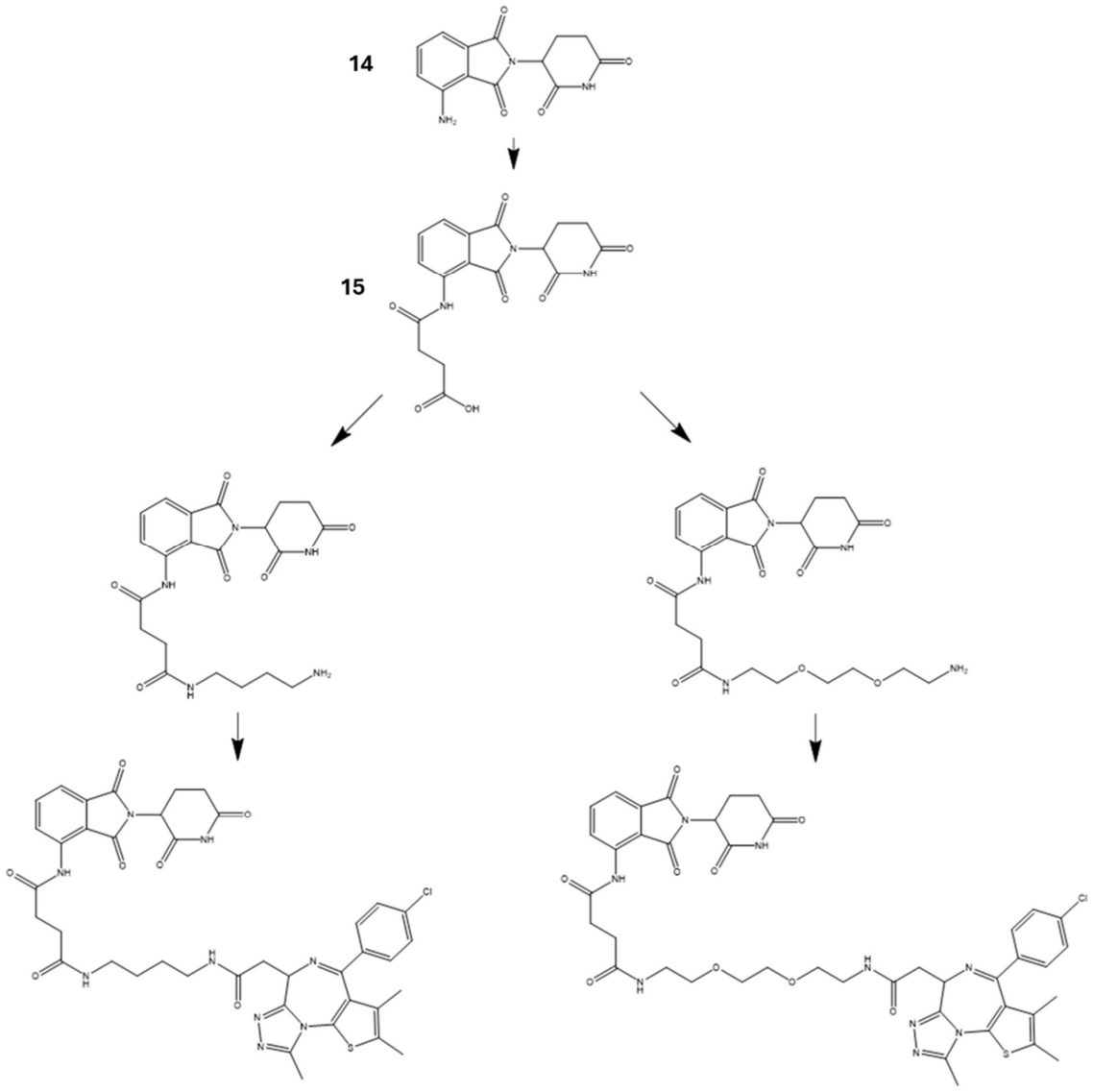


Figure 22: An overview of the potential steps involved in the synthesis of the amide-connected PROTAC molecules. Numbered are pomalidomide (molecule 14) and 4-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)-4-oxobutanoic acid (molecule 15)

The first step, a succinylation of the primary amine on pomalidomide, seemed simple in theory, by reacting with succinic anhydride and potassium acetate (KOAc) acting as a base (figure 23). This reaction was attempted in multiple solvents: acetic acid, acetonitrile and tetrahydrofuran (THF).

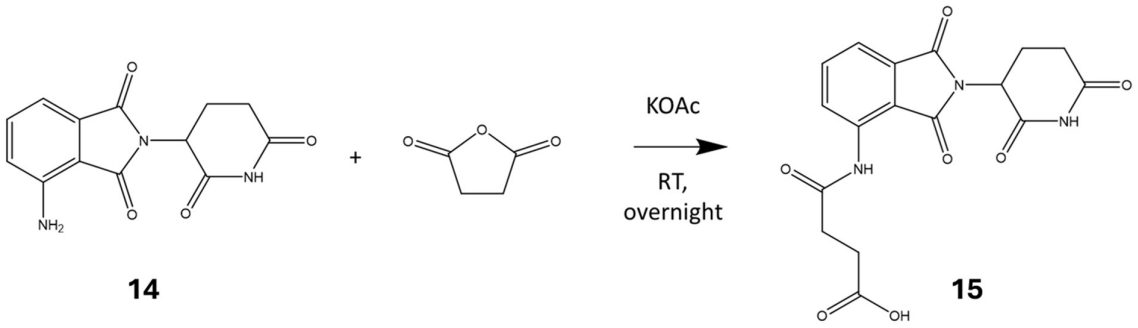


Figure 23: The scheme for a reaction between molecule 14 (0.2 M) and succinic anhydride (0.36 M) in the presence of potassium acetate (KOAc; 0.8 M) to form molecule 15. This reaction was attempted in 1 ml of multiple solvents (acetic acid, acetonitrile and THF)

In THF and acetonitrile, pomalidomide (which is a strong yellow colour) did not go into solution, and did not react with the succinic anhydride. In the acetic acid, the pomalidomide did appear to go into solution, but HPLC data (figure 24) appeared to show that although there did appear to be some product formed, it was minimal. Increasing the temperature to 50°C and 100°C made little or no difference to the yield.

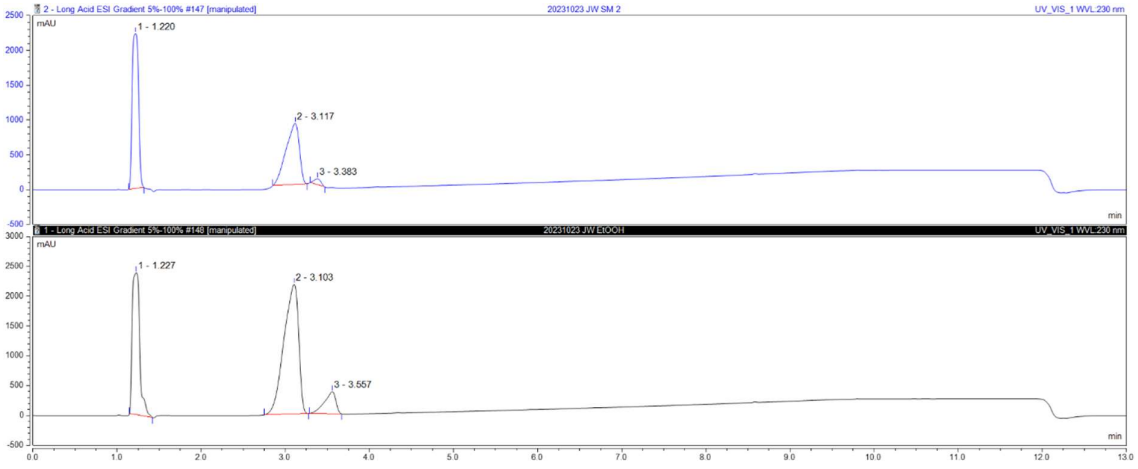


Figure 24: HPLC UV absorbance at 230mm for a pomalidomide standard (blue) and the reaction mix from figure 23 in acetic acid (black).

To increase the chance of a reaction occurring, the succinic anhydride was replaced with succinyl chloride (figure 25), and the solvent was changed to dichloromethane (DCM).

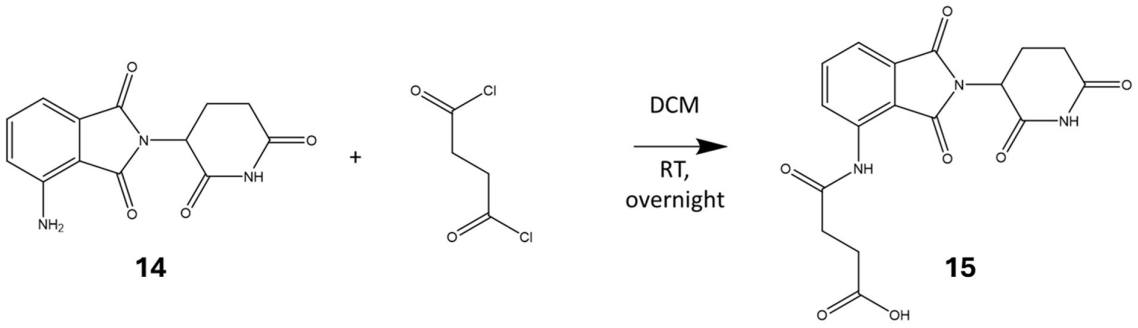


Figure 25: The scheme for a reaction between molecule 14 (0.2 M) and succinyl chloride (0.36 M) in 1 ml of DCM to form molecule 15.

This reaction was initially undertaken in the presence of a weak base (triethylamine), to prevent the -NH_2 on the pomalidomide from becoming an unreactive -NH_3^+, but immediately upon addition of the base the mixture went from yellow to a deep red (almost black). This gave a very cluttered HPLC (figure 26) which may well have involved some product formation but also seemed to involve the formation of a number of by-products which are entirely undesirable.

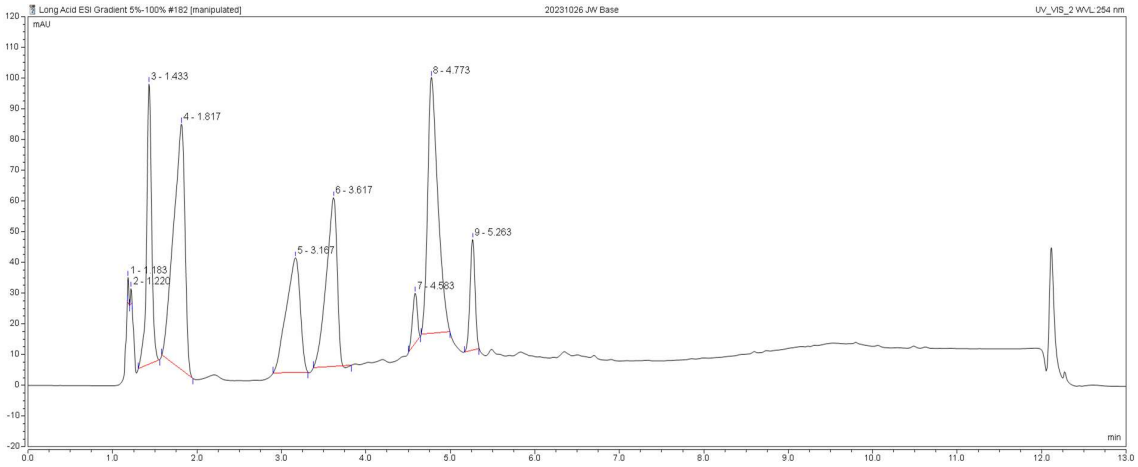


Figure 26: HPLC UV absorbance at 230mm for the reaction from figure 25 in presence of triethylamine.

A potential reason for this could be the unintended formation of highly reactive ketenes facilitated by the base (figure 27). Based on this finding, it was decided to abandon the base.

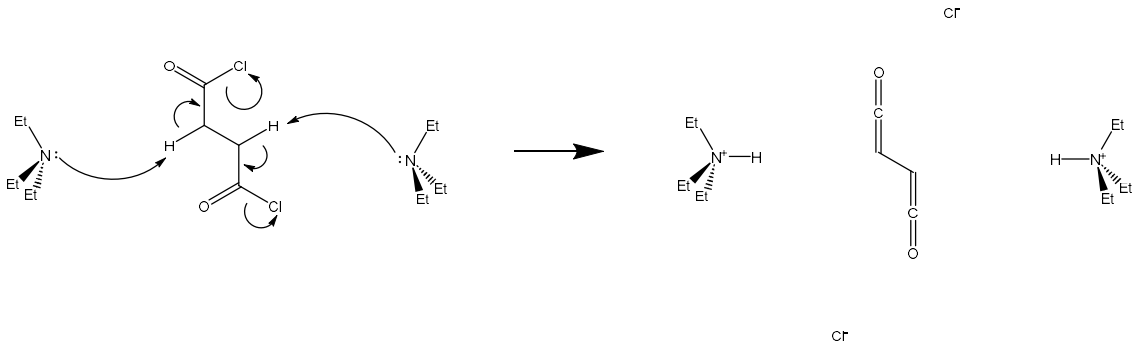
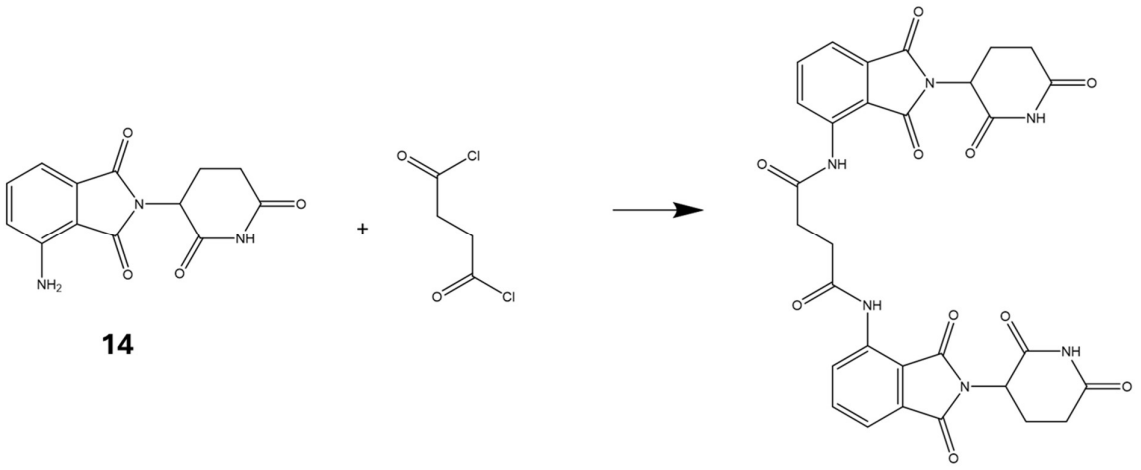


Figure 27: A potential mechanism for ketene formation from succinyl chloride in the presence of triethylamine.

Due to the symmetry of the succinyl chloride, there was some concern about a reaction between one succinyl chloride molecule with two pomalidomide molecules to form a product (referred to as the "dimer", figure 28) that would not be able to be used to make PROTACs



14

Figure 28: The scheme for a reaction between molecule 14 and succinyl chloride forming N^1*,*N^4*-bis(2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)succinamide (the "dimer").*

To mitigate the risk of forming the dimer, the succinyl chloride was kept in a high excess, with addition of the pomalidomide slowly and portion-wise to prevent a high local concentration. Unfortunately, LC-MS data (table 6; figures 29, 30 and 31) showed that although product was being formed, so too was the dimer and there was a not-insignificant amount of unreacted material present. Consequently, it was decided to abandon this connection as there was insufficient time to further optimise the reaction.

Table 6: The predicted masses of pomalidomide and the products of figures 25 and 28 after charging with the addition of a proton.

Molecule	Predicted Exact Mass of Molecule + H^+ (Da)
Molecule 14	274.0822
Molecule 15	374.0983
the “dimer”	629.1627

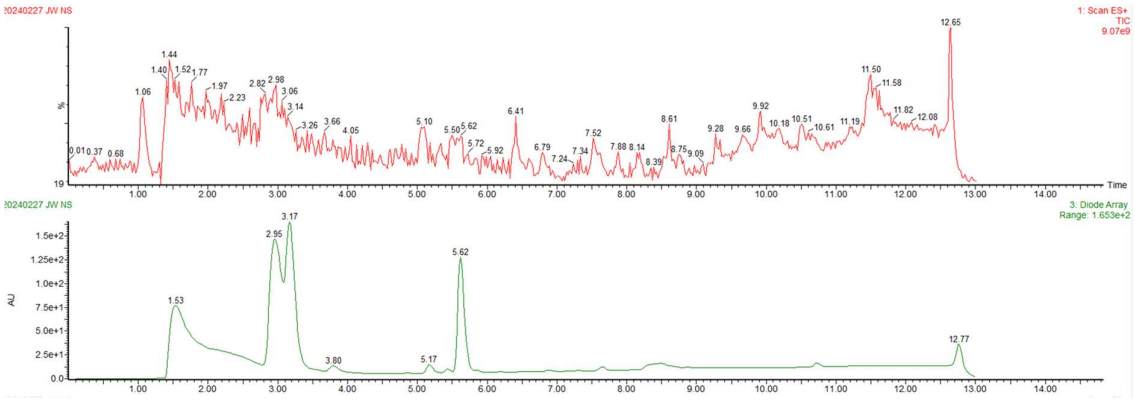


Figure 29: The TIC (red) and PDA (green) of a sample from the reaction in figure 25.

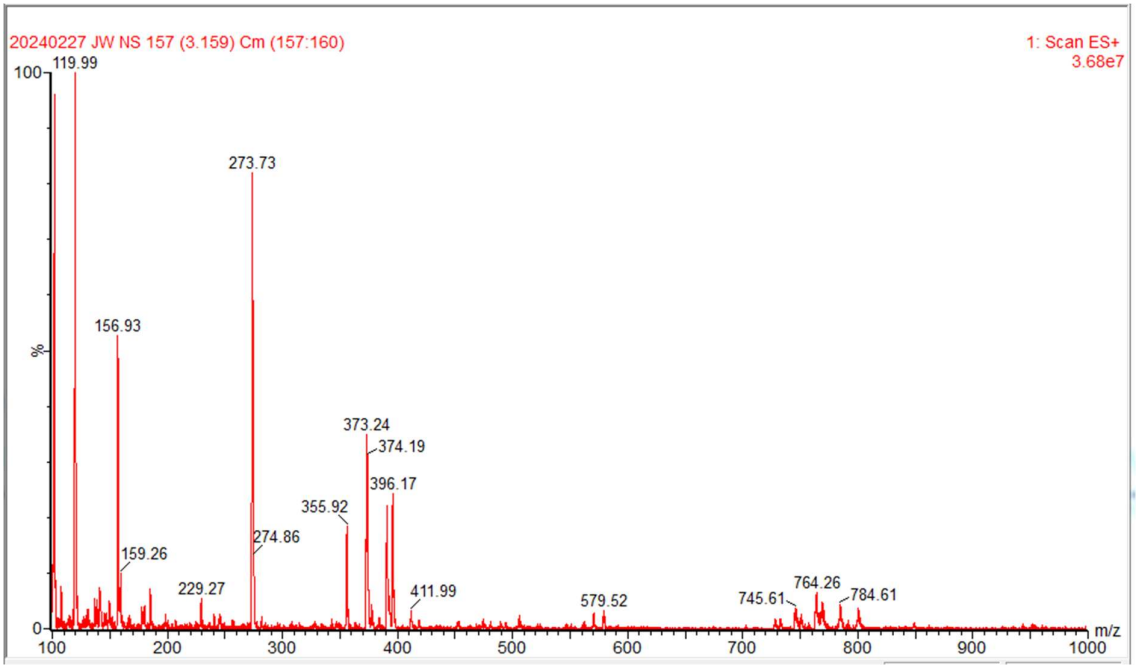


Figure 30: The MS of the peak at 3.17 minutes of a sample from the reaction in figure 25.

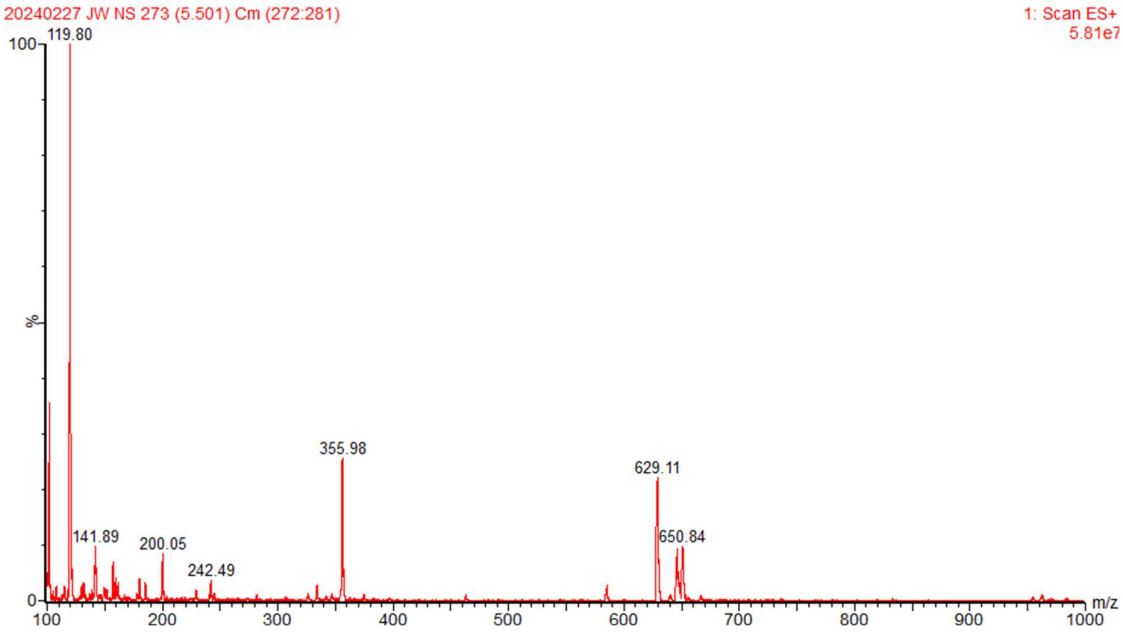


Figure 31: The MS of the peak at 5.62 minutes of a sample from the reaction in figure 25.

Biphenyl-Connected Linkers

Similarly to the plans of the other two connections, a plan was drawn up for synthesis that involved a shared first step before divergence (figure 32)

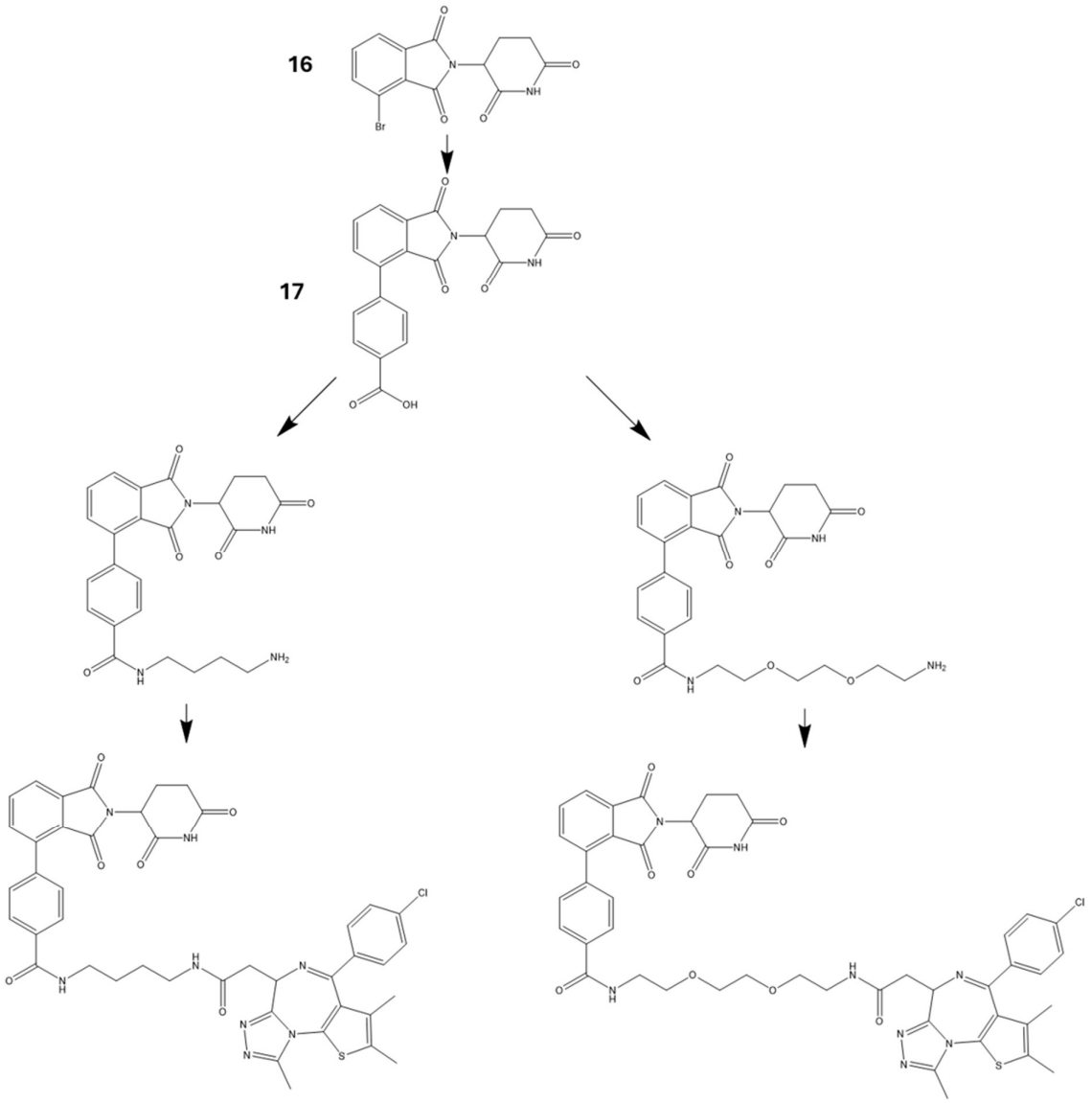


*Figure 32: An overview of the potential steps involved in the synthesis of the biphenyl-connected PROTAC molecules. Numbered are 4-bromo-2-(2,6-dioxopiperidin-3-yl)isoindoline-1,3-dione (molecule 16) and 4-(*tert*-Butoxycarbonyl)phenylboronic acid to form* tert*-butyl 4-(2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)benzoate (molecule 17)*

The first step in this pathway was a suzuki cross-coupling between a 4-bromo thalidomide derivative (molecule 16) and a boronic acid (figure 33). This reaction utilised a *tert*-butyl protecting group in an

attempt to reduce unwanted interactions. The palladium catalyst used was RuPhos Pd G3, which was selected as it was already available in the lab.

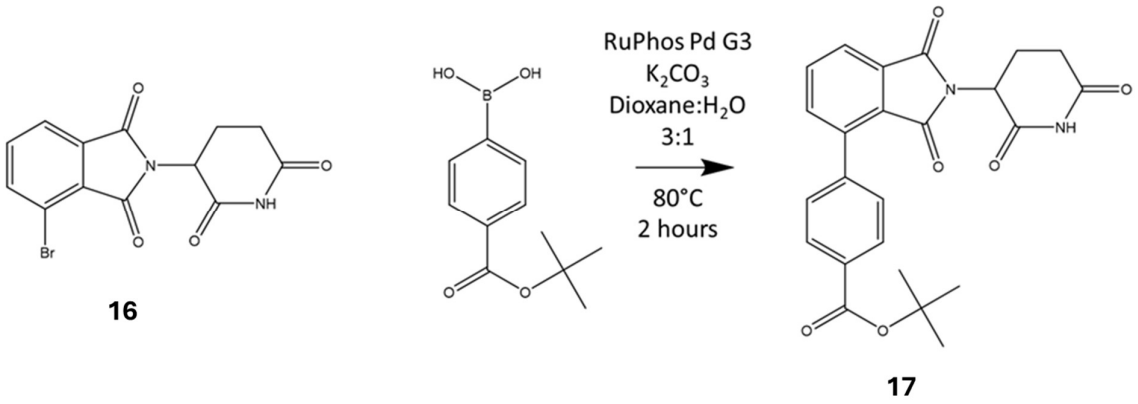


*Figure 33: The scheme for a reaction between molecule 16 (0.2 M) and 4-(*tert*-Butoxycarbonyl)phenylboronic acid (0.25 M) in 1 ml of 3:1 Dioxane:H_2O to form (molecule 17).*

Analysing the reaction mix by LC-MS showed a multitude of peaks (figure 34), implying that there were several by-products involved in the reaction, and based on financial and time constraints it was decided that the effort to optimise this reaction was not worth it, and the biphenyl connection was abandoned.

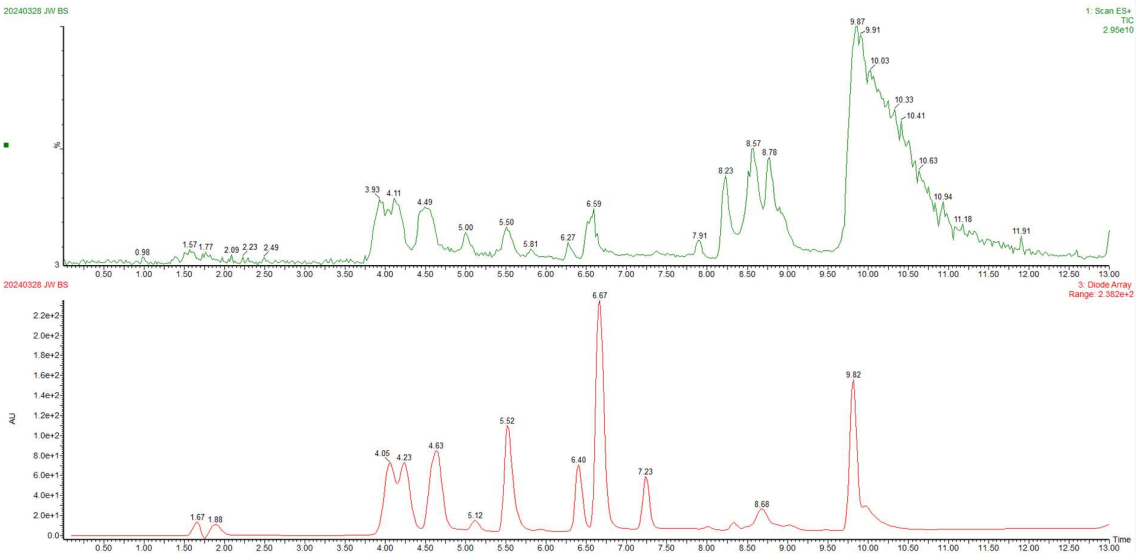


Figure 34: The TIC (green) and PDA (red) of a sample taken from the reaction in figure 33.

Stability and Metabolism

Clearance

Clearance assays were performed in metabolically relevant matrices to understand any possible differences in half-lives of PROTAC-1 and PROTAC-2. The matrices chosen were human blood plasma, human liver microsomes and human hepatocytes. Blood plasma was selected as a matrix because the PROTACs would likely to be distributed in the body via the circulatory system, meaning that they would spend a significant amount of time in blood plasma, so it is crucial to understand the stability of the molecule in this metabolic environment. Liver microsomes and hepatocytes were chosen as the primary method of drug metabolism and clearance is through the liver, so an understanding of the stability of the PROTACs in these matrices allows an understanding of the body's ability to remove the drugs from the system.

Human Blood Plasma

PROTAC-1 and PROTAC-2 showed reasonable stability in blood plasma (figure 35). Table 7 shows a comparison of the two respective half-lives of the molecules, with PROTAC-1 having a considerably longer half-life.

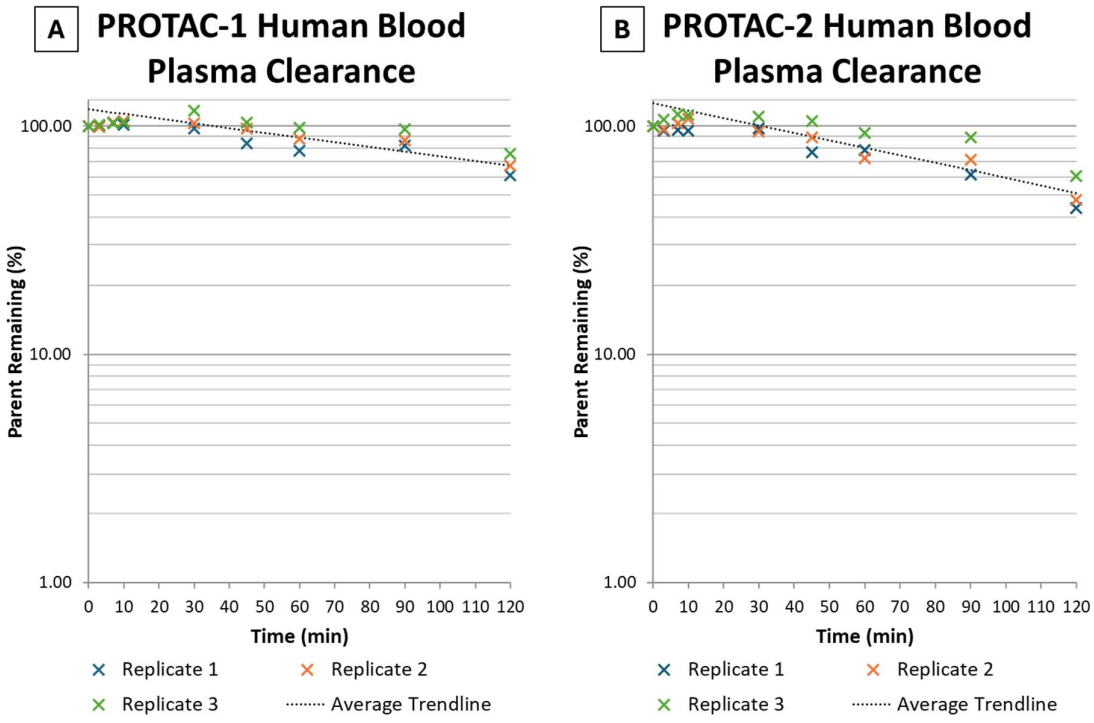


Figure 35: A: A graph showing the proportion of PROTAC-1 remaining over time after spiking into human blood plasma, with an averaged trendline.(n=3) B: A graph showing the proportion of PROTAC-2 remaining over time after spiking into human blood plasma, with an averaged trendline. (n=3)

Table 7: A comparison of the half-life of PROTAC-1 and PROTAC-2 in human blood plasma.

Molecule	Half-Life (min)
PROTAC-1	146.59 ± 3.08
PROTAC-2	92.39 ± 6.46

To confirm that any effect on the molecule was plasma-specific, a Dulbecco's Phosphate Buffered Saline (DPBS) control was also run under the same conditions. (figure 36). DPBS has a similar tonicity (osmotic pressure) and pH to plasma, so is a suitable control as any instability caused by temperature, pH or other non-plasma-specific processes will also be present in the DPBS control, so any differences can be safely assumed to be unique to the blood plasma. The control showed unusual results.

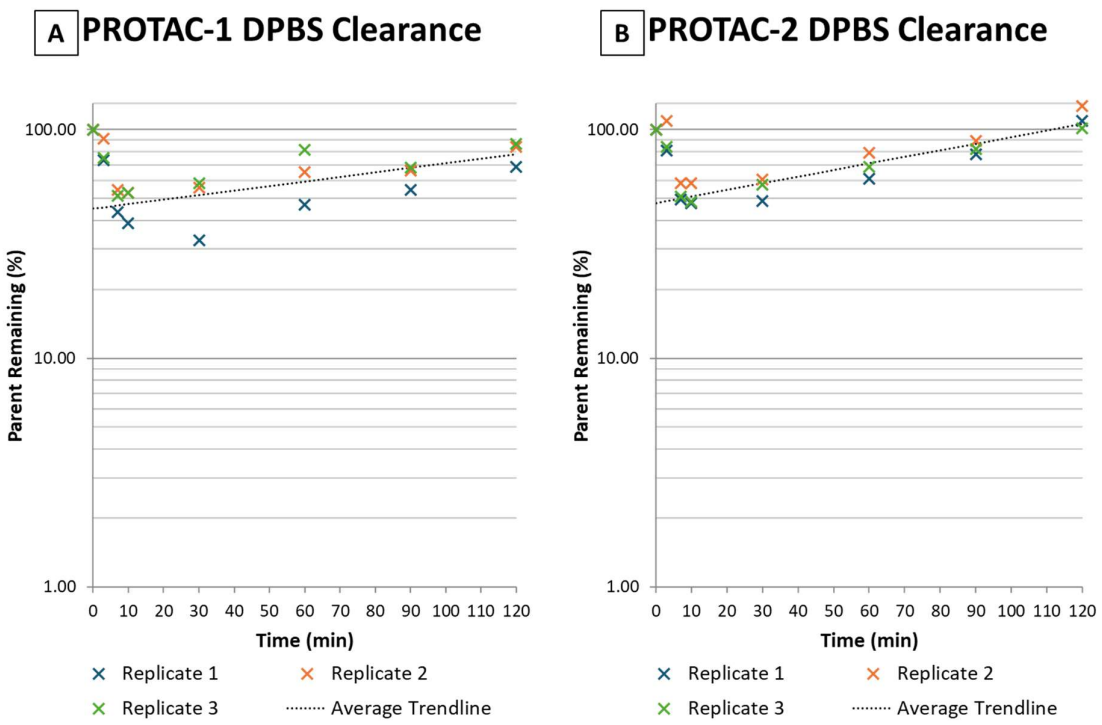


Figure 36: A: A graph showing the proportion of PROTAC-1 remaining over time after spiking into DPBS , with an averaged trendline. (n=3) B: A graph showing the proportion of PROTAC-2 remaining over time after spiking into DPBS, with an averaged trendline. (n=3)

The negative half-life (table 8) isn't easily explained. It could be due to improper mixing or other experimental error, but a similar trend was not observed to the same degree in a control compound, propantheline, which was run on the same plate under the same conditions as the PROTAC samples (figure 37). Propantheline is known to have high plasma stability, so any experimental causes for the upwards trend for the PROTAC would be expected to be observed for propantheline. The stability of propantheline appeared to follow the expected long half-life, and did not show any measurable signs of behaving in a similar manner to the PROTACs. This suggests that some other factor is responsible but it is unclear what it is.

Table 8: A comparison of the half-life of PROTAC-1 and PROTAC-2 in DPBS.

Molecule	Half-Life (min)
PROTAC-1	-153.02 ± 13.05
PROTAC-2	-104.34 ± 4.74

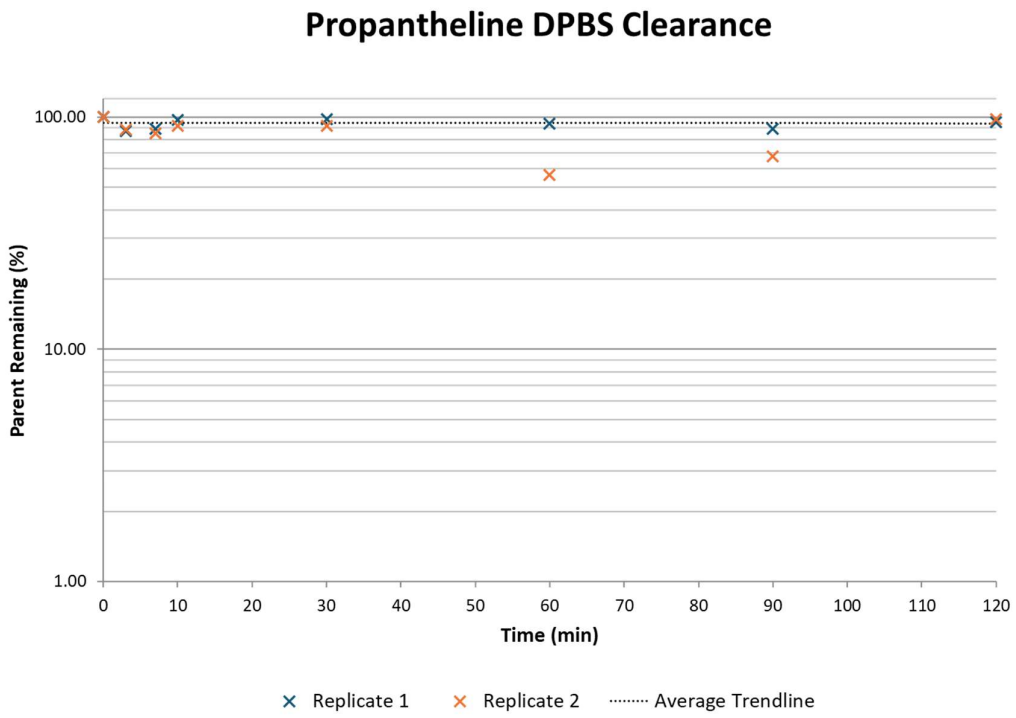


Figure 37: A graph showing the proportion of propantheline remaining over time after spiking into DPBS, with an averaged trendline. (n=2)

It is possible therefore that the half-life obtained in the plasma is unreliable, as the PROTACs may also be behaving in plasma in a way similar to the DPBS alongside the degradation.

Human Liver Microsomes

In human liver microsomes (HLMs) the clearance of both PROTACs was significant (figure 38). Table 9 shows the respective half-lives and intrinsic clearances (Cl_{int}) of each PROTAC. PROTAC-2 has a somewhat longer half-life, and therefore lower Cl_{int}.

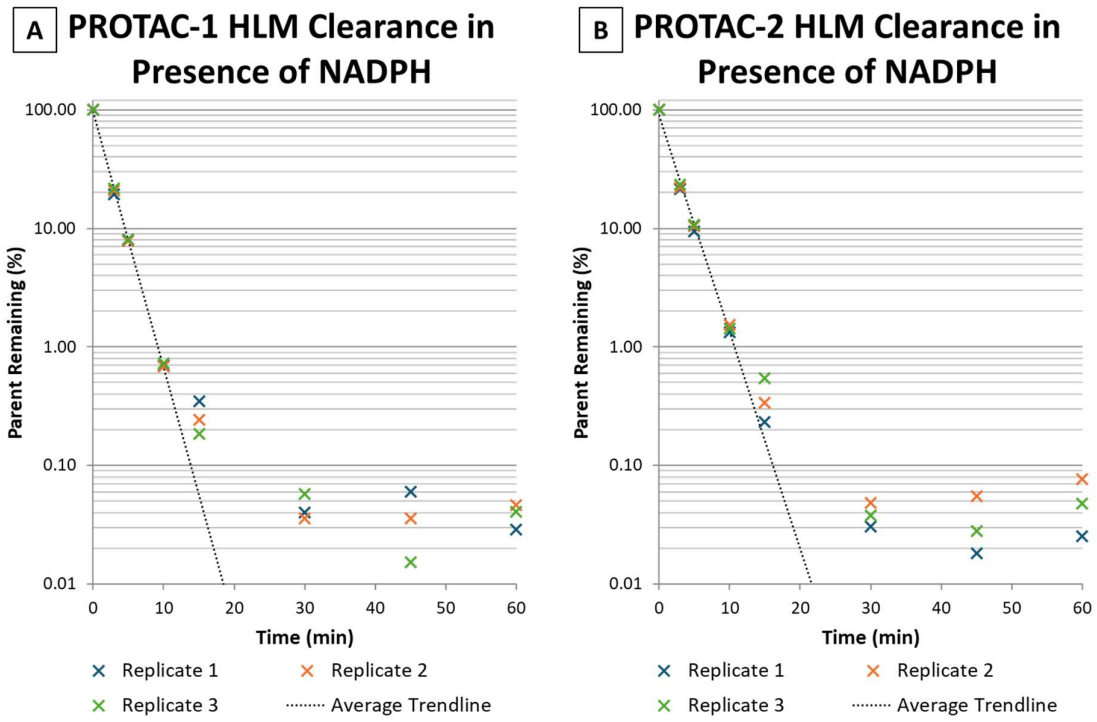


Figure 38: A: A graph showing the proportion of PROTAC-1 remaining over time after spiking into HLMs containing NADPH, with an averaged trendline.(n=3) B: A graph showing the proportion of PROTAC-2 remaining over time after spiking into HLMs containing NADPH, with an averaged trendline. (n=3)

Table 9: A comparison of the half-life and Cl_{int} of PROTAC-1 and PROTAC-2 in HLMs.

Molecule	Half-Life (min)	Cl_{int} (µl min^{-1} mg^{-1})
PROTAC-1	1.39 ± 0.01	994.90 ± 8.19
PROTAC-2	1.65 ± 0.02	839.92 ± 11.39

As a way to show that phase I metabolism is responsible for the degradation of the molecule, an assay was performed simultaneously without NADPH, an important co-factor in the cytochrome P450 (CYP450) and flavin-containing monooxygenase (FMO) metabolic pathways.

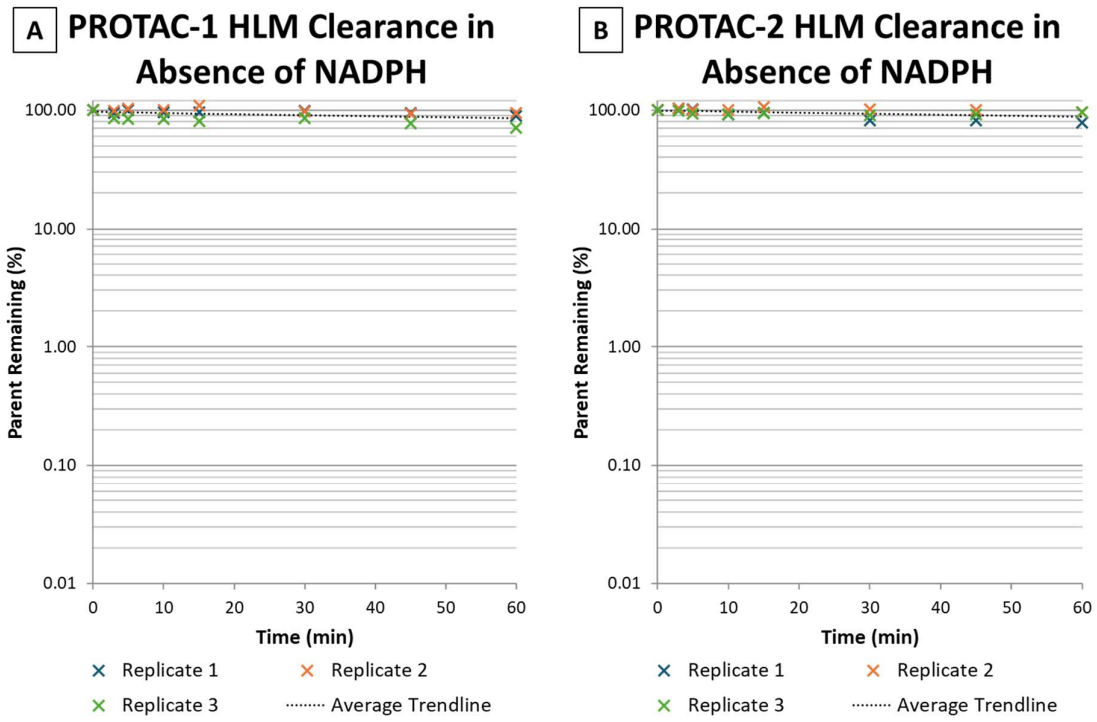


Figure 39: A: A graph showing the proportion of PROTAC-1 remaining over time after spiking into HLMs without any NADPH, with an averaged trendline. (n=3) B: A graph showing the proportion of PROTAC-2 remaining over time after spiking into HLMs without any NADPH, with an averaged trendline. (n=3)

Figure 39 shows the degradation of the PROTACs in the no co-factor (NCF) assay, showing minimal change over time. Due to the slow rate of degradation, the half-life was greater than can be confidently calculated, and therefore the Cl_{int} was also not calculable. From this data, it can safely be concluded that NADPH-mediated enzymes are responsible for the degradation seen in HLMs.

Human Hepatocytes

In human hepatocytes both PROTACs showed strong clearance (figure 40), the half-lives were similar but PROTAC-1 showed a marginally shorter half-life (table 10) and therefore higher intrinsic clearance (Cl_{int})

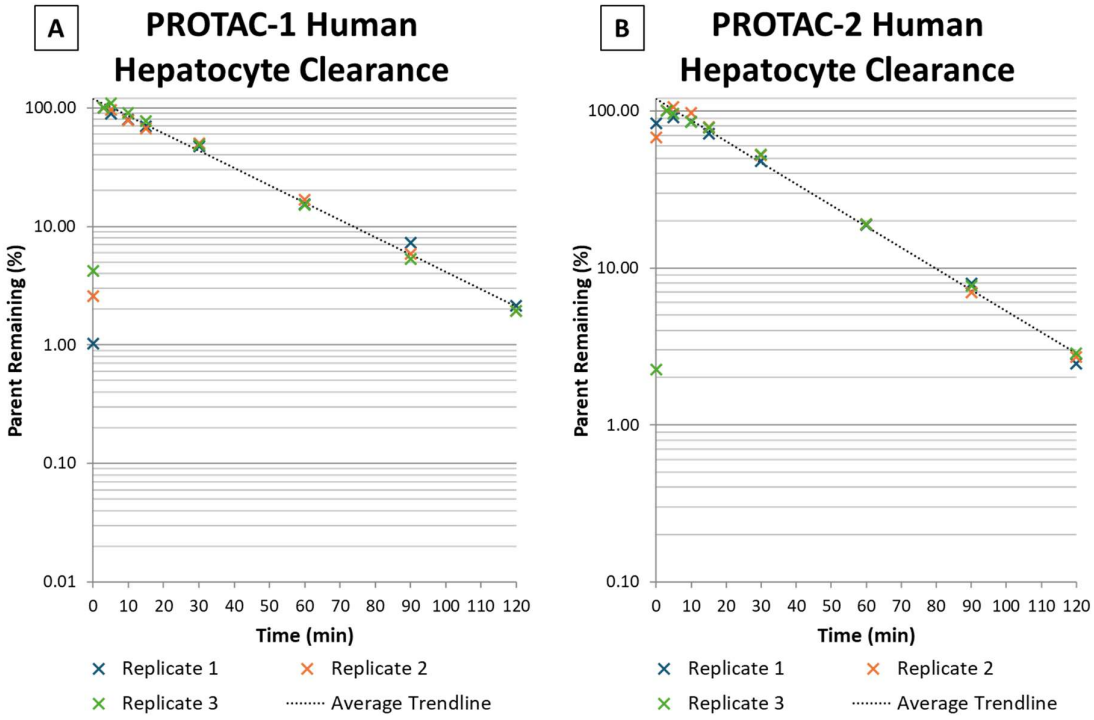


Figure 40: A: A graph showing the proportion of PROTAC-1 remaining over time after spiking into human hepatocytes, with an averaged trendline. (n=3) B: A graph showing the proportion of PROTAC-2 remaining over time after spiking into human hepatocytes, with an averaged trendline. (n=3)

Table 10: A comparison of the half-life and Cl_{int} of PROTAC-1 and PROTAC-2 in human hepatocytes.

Molecule	Half-Life (min)	Cl_{int} (µl min^{-1} 10^6 $cells^{-1}$)
PROTAC-1	20.66 ± 0.83	67.21 ± 2.72
PROTAC-2	22.29 ± 0.36	62.21 ± 1.01

As a control, a parallel assay using hepatocytes that had been heat-inactivated. In this assay both PROTACs showed significant degradation, far higher than expected (figure 41). The half-life of both molecules was shorter in heat-inactivated hepatocytes than in active hepatocytes (table 11), with a similar difference between the PROTACs.

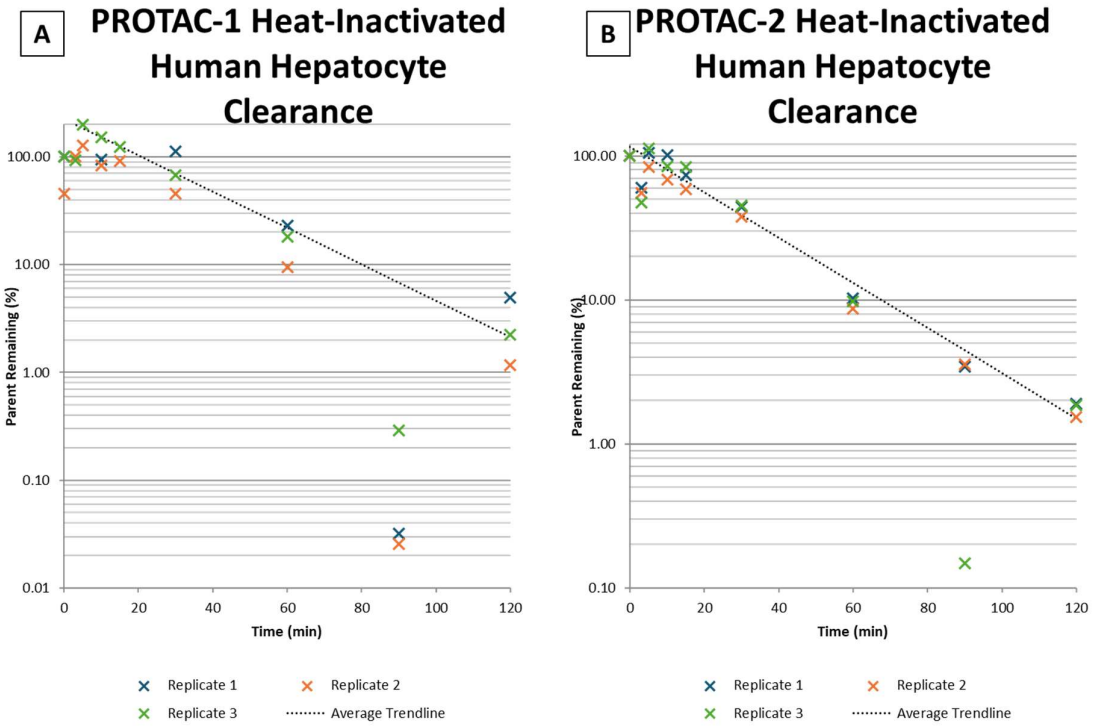


Figure 41: A: A graph showing the proportion of PROTAC-1 remaining over time after spiking into heat-inactivated human hepatocytes, with an averaged trendline. (n=3) B: A graph showing the proportion of PROTAC-2 remaining over time after spiking into heat-inactivated human hepatocytes, with an averaged trendline. (n=3)

Table 11: A comparison of the half-life and Cl_{int} of PROTAC-1 and PROTAC-2 in heat-inactivated human hepatocytes

Molecule	Half-Life (min)	Cl_{int} (µl min^{-1} 10^6 $cells^{-1}$)
PROTAC-1	17.93 ± 0.64	77.42 ± 2.73
PROTAC-2	19.19 ± 0.21	72.24 ± 0.80

From this data it cannot be safely concluded that hepatocytic metabolism is responsible for the degradation of the PROTACs, as in the matrix that has been heat treated, enzymatic activity should be essentially eliminated.

Metabolite Identification

Human Blood Plasma

To understand what mechanisms are influencing the degradation of the PROTACs in blood plasma, metabolite identification was undertaken. All photodiode array detectors (PDAs) were set to filter wavelengths outside of 300-350 to enable easier isolation of the test compounds.

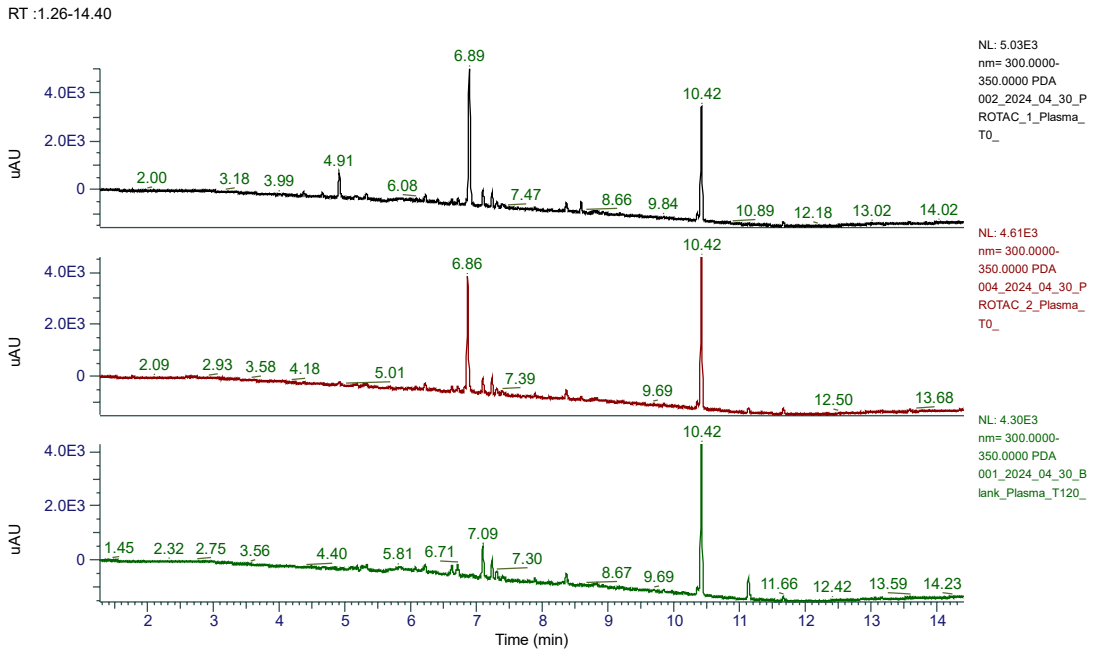


Figure 42: A comparison between the PDA (300-350nm) outputs of PROTAC-1 (black) and PROTAC-2 (red), both in human blood plasma alongside blank plasma (green). PROTAC samples are taken at t=0 min, blank plasma at t=120 min.

Figure 42 shows the t=0 min of both PROTACs compared to a blank plasma sample. This allowed the peak at 10.42 to be identified as intrinsic to plasma and therefore can be ruled out of further investigation as it is not of interest to the test compounds. The peaks that appeared to be the most interesting were the peaks at 4.91 minutes and 6.89 minutes for PROTAC-1 and 6.86 minutes for PROTAC-2, although the peak at 4.91 is unlikely to be of interest as it was not observed in the high-

resolution LC-MS^2 of pure compound and as this is the t=0 min and the PROTAC has been shown not to have a short half-life in plasma it is unlikely that it could be a result of degradation.

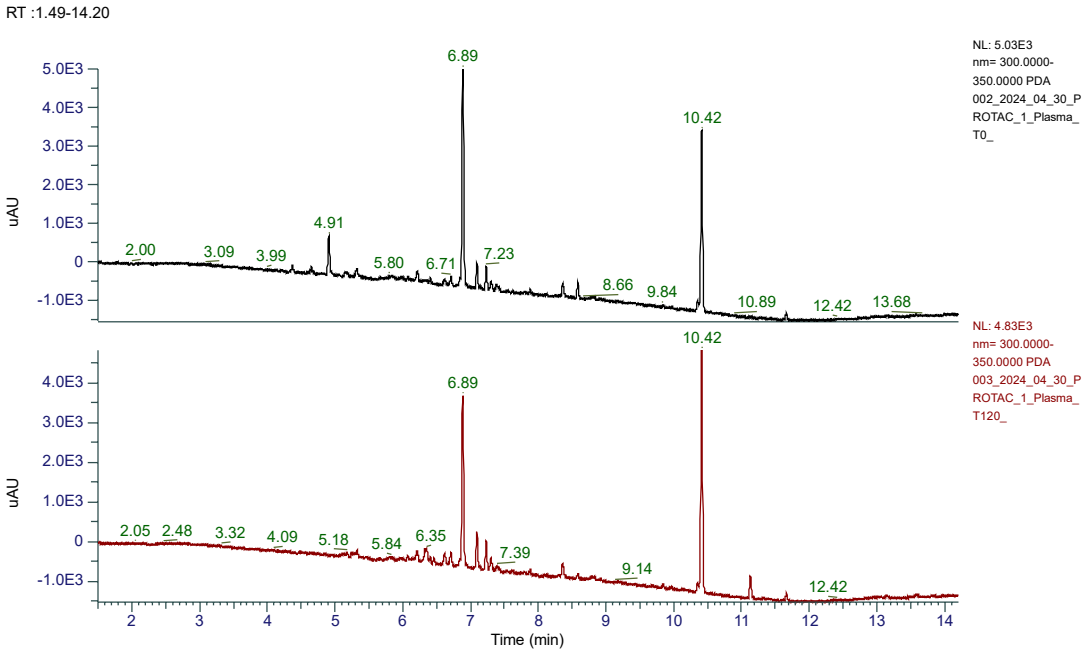


Figure 43: A comparison between the PDA (300-350nm) outputs of PROTAC-1 at t=0 min (black) and t=120 min (red) in human blood plasma.

Figure 43 shows the change in PDA between t=0 min and t=120 min. The peak at 4.91 min is completely absent at the second timepoint, which could be due to complete degradation (which would suggest that it is not PROTAC related, as even if it were a product that had a shorter half-life it would be being

replaced by degradation of the base peak.) or it could be due to it being a contaminant unique to that timepoint. Either way it can be fairly confidently disregarded for investigation.

To confirm that the 6.89 minute peak is PROTAC-1, an extracted ion chromatogram (EIC) around the 785.2257 mass found from the high resolution MS earlier (figure 44). This puts beyond a reasonable doubt that the 6.89 (6.91 in the mass chromatogram) peak is PROTAC-1.

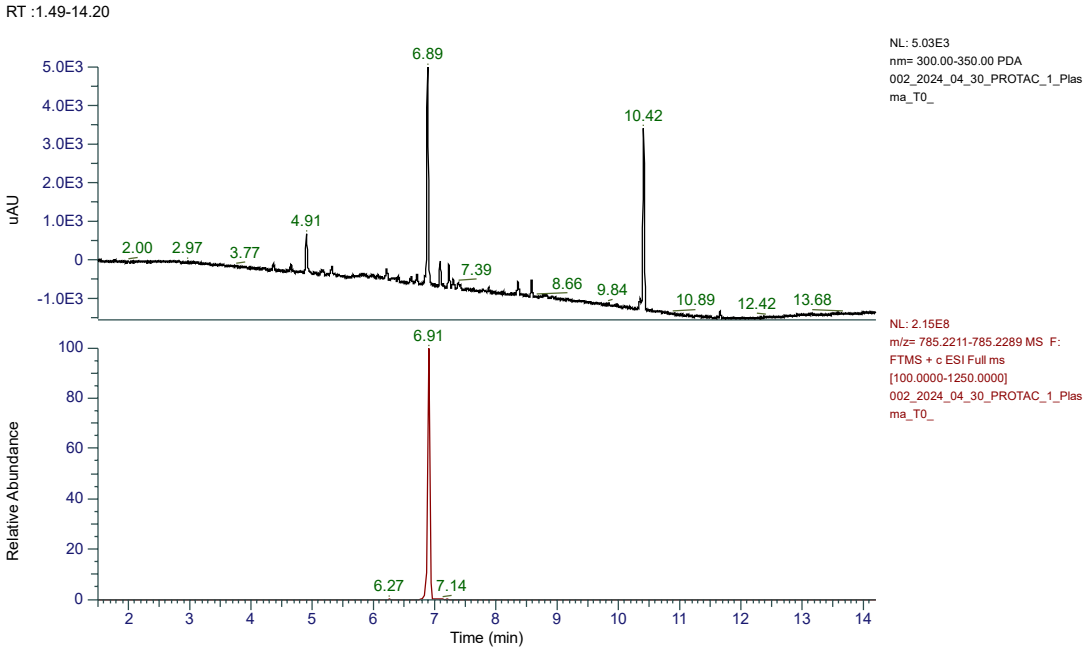


Figure 44: The PDA (300-350nm; black) and EIC (785.2211-785.2289; red) output of PROTAC-1 at t=0 min in human blood plasma.

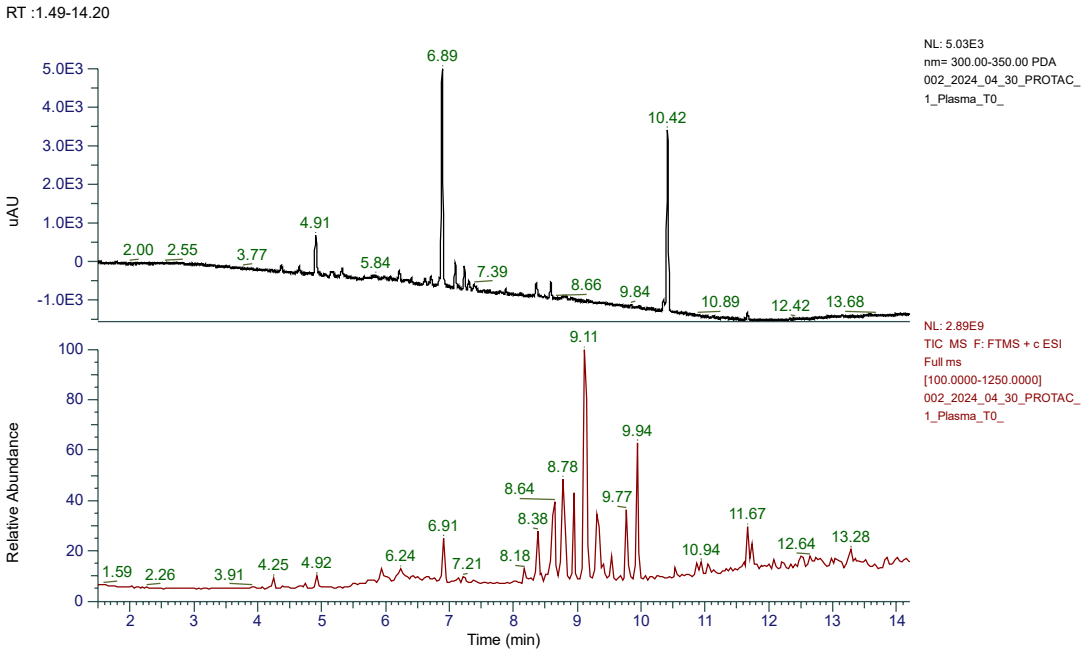


Figure 45: The PDA (300-350nm; black) and TIC (red) output of PROTAC-1 at t=0 min in human blood plasma.

Figure 45 shows the PDA and total ion chromatogram (TIC) of PROTAC-1 at t=0 min. The TIC was entirely too cluttered to be able to pick out potential degradation products as they appeared by identifying new peaks on the TIC, especially when factoring in the relatively small size of the 6.91 peak and the

apparent stability of the PROTACs in plasma . Therefore, EICs of probable products would need to be used.

In plasma, hydrolysis is most likely to be the key route of degradation as hydrolases make up a significant portion of plasma enzymes. Figure 46 shows possible products of hydrolysis, and table 12 lists their respective predicted masses when charged with a proton.

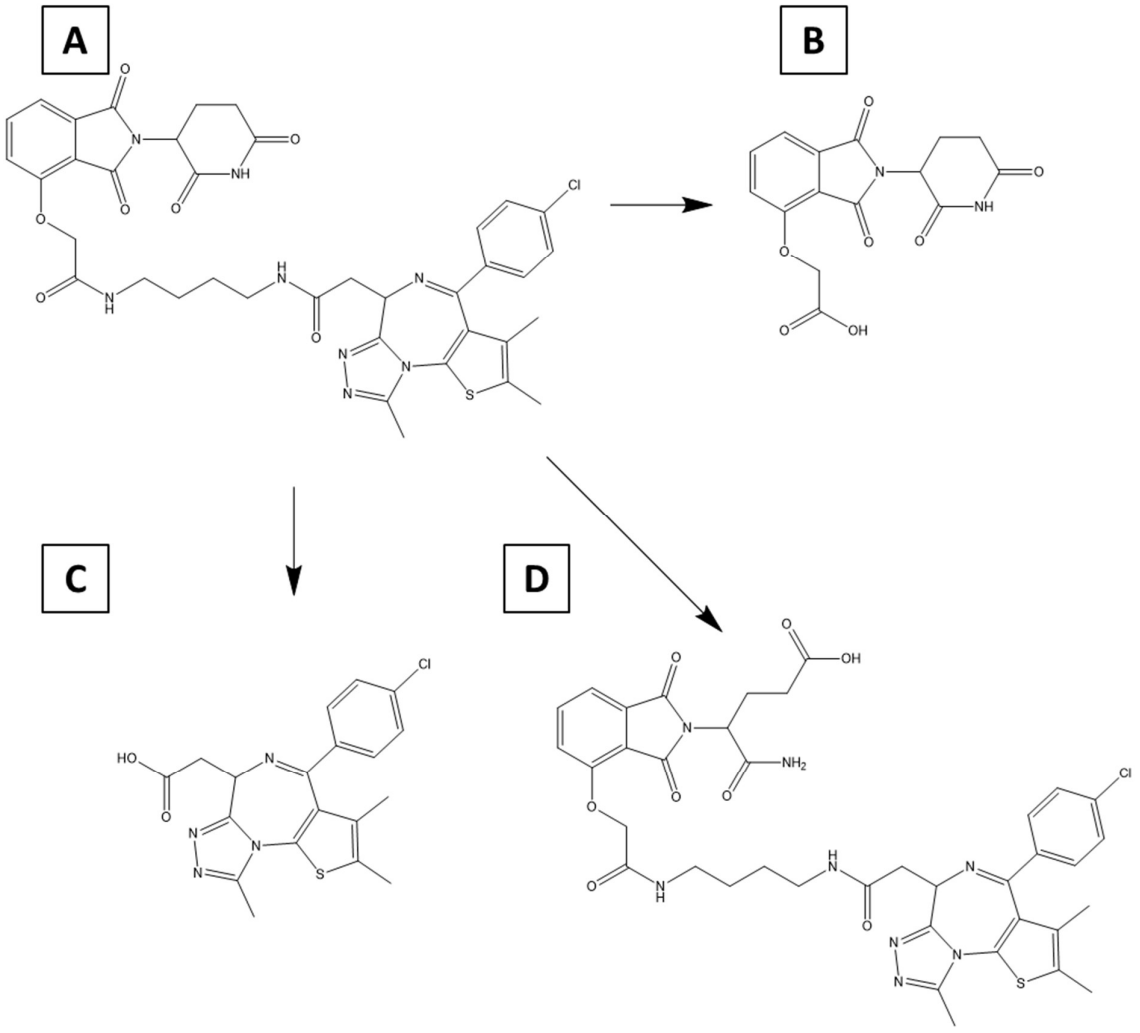


Figure 46: A: The skeletal structure of PROTAC-1 B, C and D: Possible hydrolysis products of PROTAC-1 (in situations where hydrolysis would create two products, only one is shown)

Table 12: The predicted masses of the hydrolysis products shown in figure 46 when charged with the addition of a proton.

Hydrolysis Product	Predicted Exact Mass of Molecule + H^+ (Da)
B	333.0717
C	401.0834
D	803.2373

It is worth noting that B and C are not mutually exclusive; it is possible for both to be formed from the original molecule if hydrolysis were to occur at two separate points. B and C also involve a second molecule each being formed in their hydrolytic event. B and D can undergo further hydrolysis. For this study, however, only the three molecules listed were investigated based on the assumption that any other hydrolysis products would involve at least one of these being formed at some point.

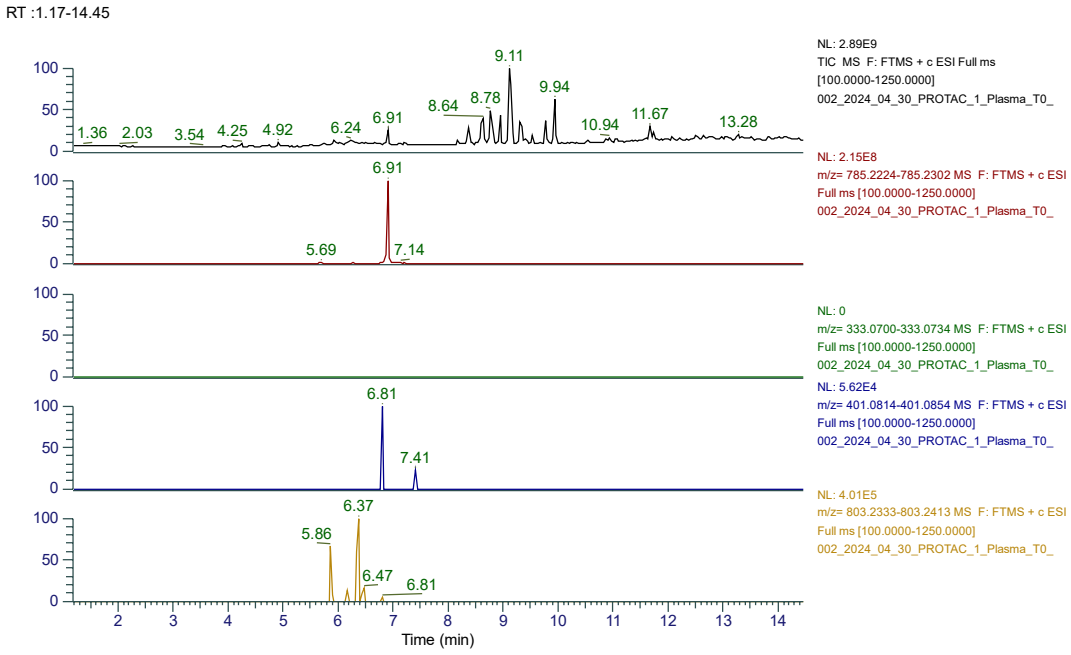


Figure 47: From the t=0 min sample of PROTAC-1 in human blood plasma. The TIC (black), as well as the EICs of the parent ion (785.2211-785.2289; red) and possible hydrolysis products (333.0700-333.0734; green, 401.0814-401.0854; blue, 803.2333-802.2413; yellow).

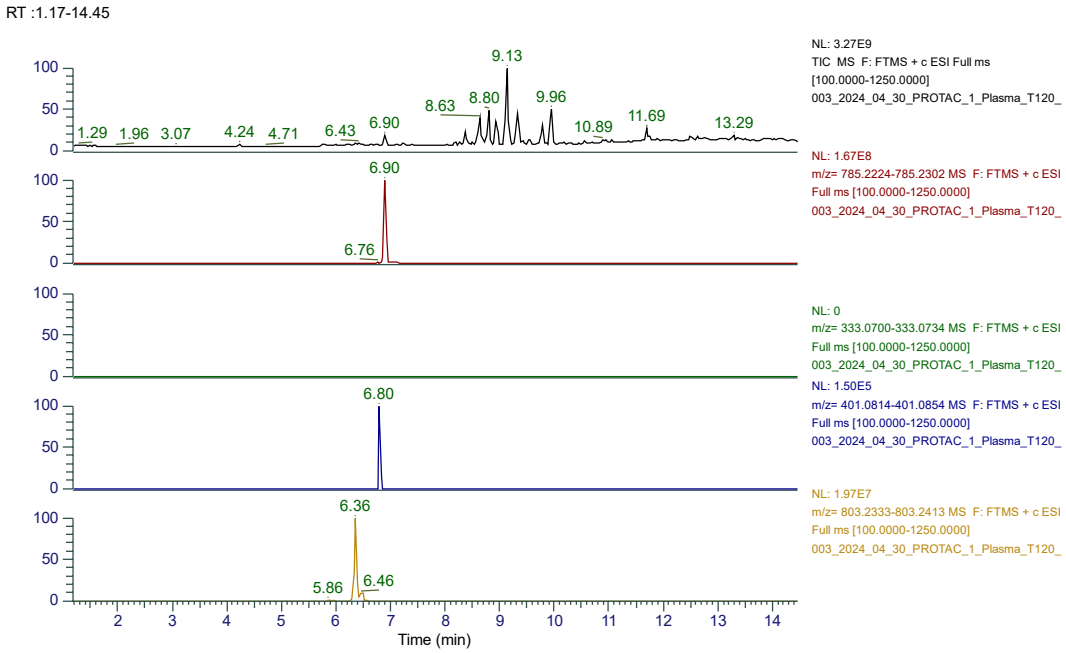


Figure 48: From the t=120 min sample of PROTAC-1 in human blood plasma. The TIC (black), as well as the EICs of the parent ion (785.2211-785.2289; red) and possible hydrolysis products (333.0700-333.0734; green, 401.0814-401.0854; blue, 803.2333-802.2413; yellow).

Figure 47 shows the possible product EICs at t=0 min, and figure 48 shows the same at t=120 min. The EIC data is useful in that it shows that it is highly unlikely that formation of product B is occurring to any real degree. It is possible that it is being formed at levels too low for the MS to pick up the signal, or that B has a very short half-life and is being cleared very quickly upon formation. C and D appear to be present in both samples, and using the amount of signal that the MS picked up as an approximate quantifier of amount of product, C increases a small amount between the two time points (5.65×10^4 at t=0 min and 1.50×10^5 at t=120 min) but this change isn't large enough to rule out experimental

error such as pipetting inconsistency, and in absence of an internal standard it cannot be compared against anything.

Product D sees a much larger increase in magnitude (4.01×10^5 and 1.97×10^7 at t=0 min and t=120 min, respectively), this change is significantly less likely to be caused by experimental error, and could therefore be the primary product from PROTAC-1 degradation.

The PROTAC peak does show a decrease, as expected, but the decrease in PROTAC-1 does not necessarily line up with the increase in product D. This could be down to experimental error or it could be a result of the lack of accuracy in the measurement choice. To make sure that peak for 8.03 is PROTAC-1 related, MS^2 was used (figure 49).

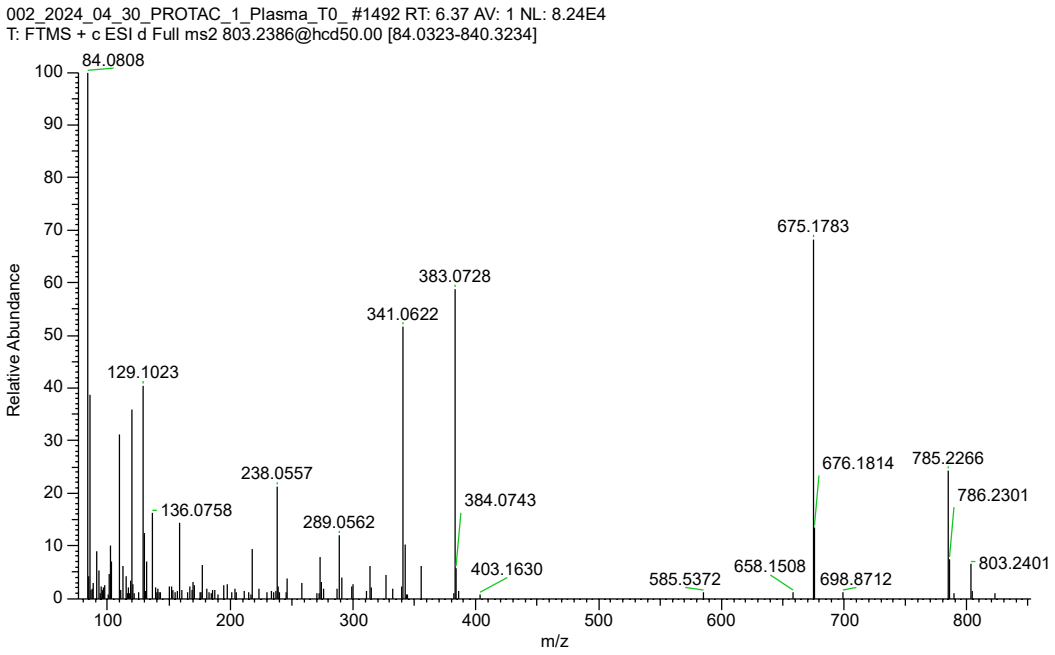


Figure 49: MS^2 of the 803.2386 ion at a retention time of 6.37 minutes from the PROTAC-1 in human blood plasma at t=0 min.

Bizarrely, the MS^2 did not pick up the peak at t=120 min, but it did at the lower signal t=0 min. The fragmentation pattern aligns with that of PROTAC-1, (341.0622 and 383.0728 for example) suggesting that D is a product of PROTAC-1. However, 785.2266 corresponds to PROTAC-1 without the hydrolysis of the imide. As the retention time for product D is 6.37 minutes, and the LC-MS shows that PROTAC-

1 is not present at that time, it is unlikely that this could be due to contamination. This means that either condensation is happening inside the MS, or the molecule in question is not product D, but some other unanticipated molecule with a similar mass.

For PROTAC-2, the same procedure was followed. Firstly, a comparison between the PDAs of t=0 min and t=120 min (figure 50) showed nothing unexpected, the 6.86 peak closely lined up with the PROTAC-2 high resolution LC-MS peak, and to confirm it an EIC around the observed 845.2468 mass was performed (figure 51) which lined up with the 6.86 peak, supporting the conclusion that it is the PROTAC-2 peak.

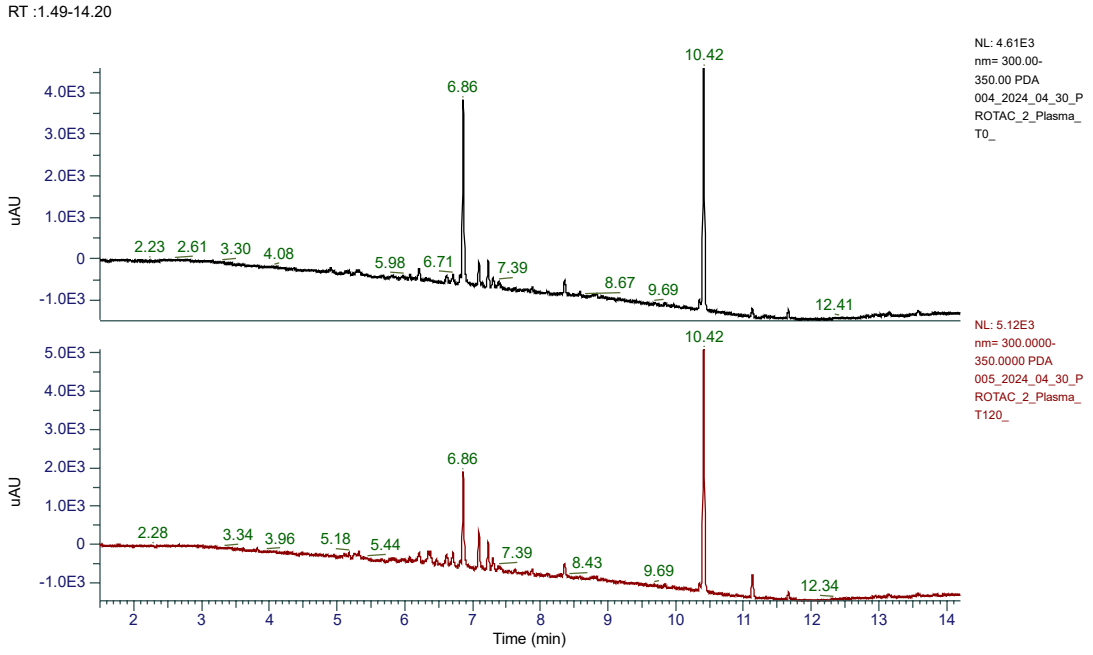


Figure 50: A comparison between the PDA (300-350nm) output of the LCMS of PROTAC-2 at t=0 min (black) and t=120 min (red) in human blood plasma.

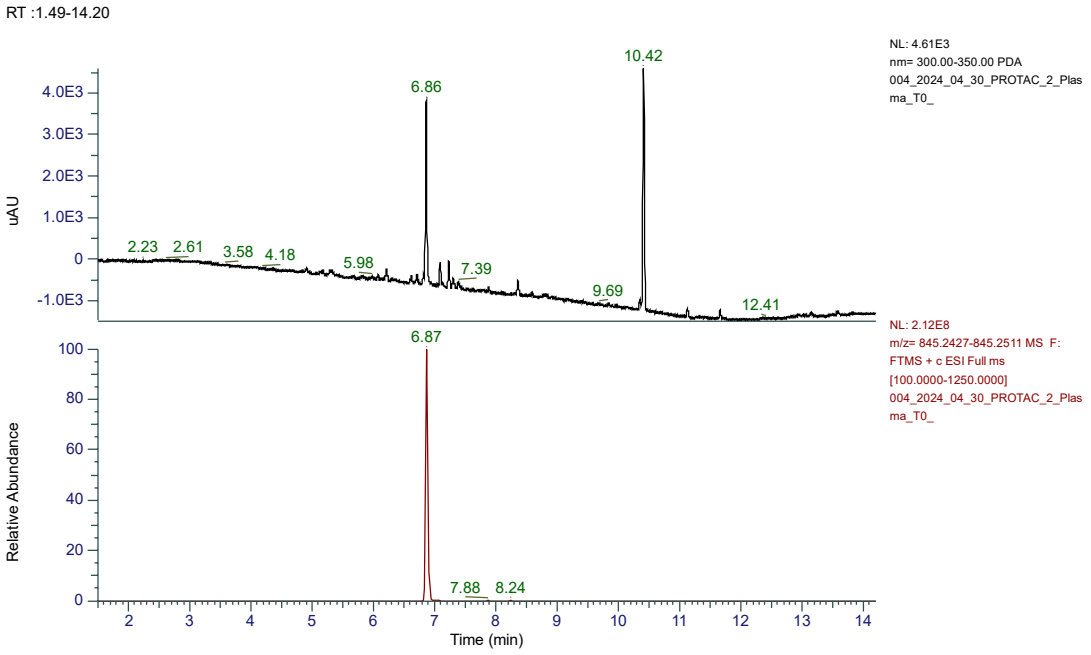


Figure 51: The PDA (300-350nm; black) and EIC (845.2427-845.2511; red) output of PROTAC-1 at t=0 min in human blood plasma.

Similarly to PROTAC-1, in order to observe potential products, possible hydrolysis products were produced (figure 52) and their protonated masses predicted (table 13) to allow EICs to be made.

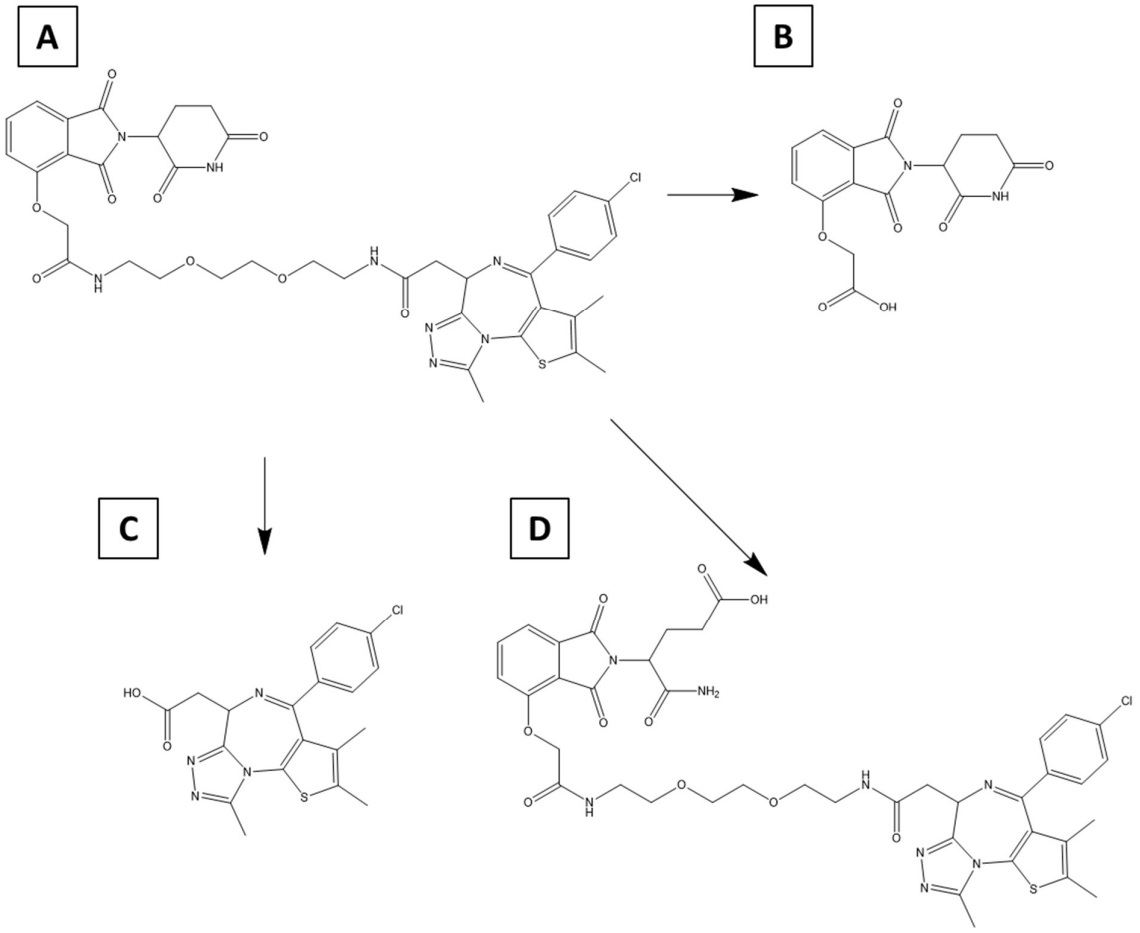


Figure 52: A: The skeletal structure of PROTAC-2 B, C and D: Possible hydrolysis products of PROTAC-2 (in situations where hydrolysis would create two products, only one is shown)

Table 13: The predicted masses of the hydrolysis products shown in figure 52 when charged by the addition of a proton.

Hydrolysis Product	Predicted Exact Mass of Molecule + H^+ (Da)
B	333.0717
C	401.0834
D	863.2584

Due to the similarity of the molecules, the products are similar in PROTAC-2 to how they were in PROTAC-1. Figures 53 and 54 show the TICs for t=0 min and t=120 min respectively.

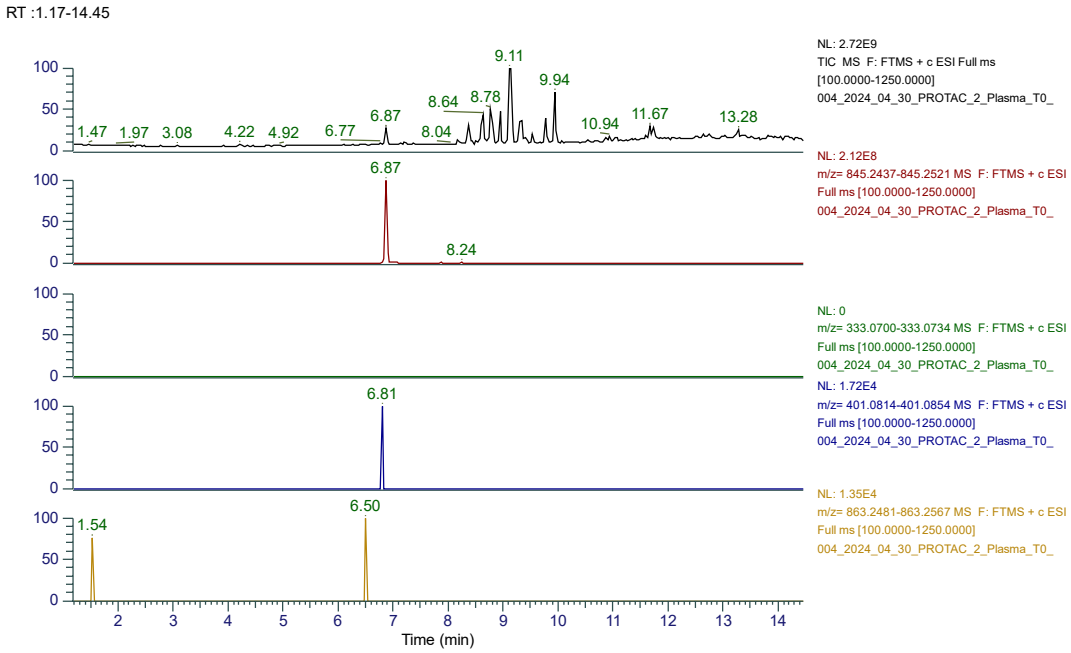


Figure 53: From the t=0 min sample of PROTAC-2 in human blood plasma. The TIC (black), as well as the EICs of the parent ion (845.2437-845.2521; red) and possible hydrolysis products (333.0700-333.0734; green, 401.0814-401.0854; blue, 863.2481-863.2567; yellow).

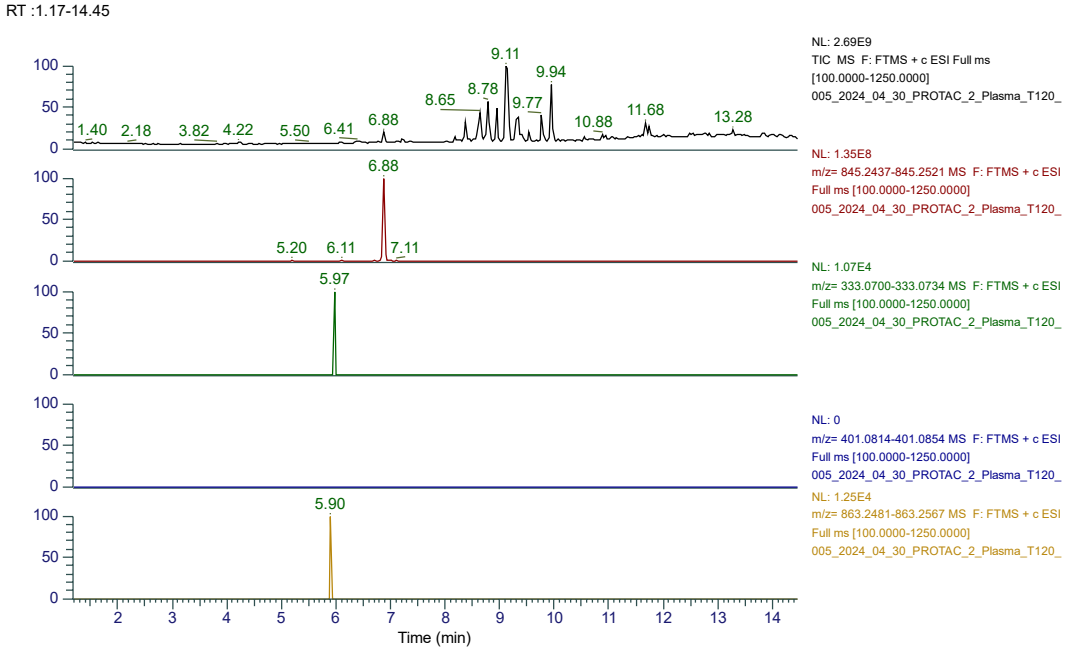


Figure 54: From the t=120 min sample of PROTAC-2 in human blood plasma. The TIC (black), as well as the EICs of the parent ion (845.2437-845.2521; red) and possible hydrolysis products (333.0700-333.0734; green, 401.0814-401.0854; blue, 863.2481-863.2567; yellow).

The TICs in PROTAC-2 give a very different result to those for PROTAC-1. Whereas in PROTAC-1, product B is absent from both samples, it is present (albeit in a small amount) in the t=120 min sample in PROTAC-2. Product C differs in PROTAC-2 as well, being absent in the t=120 min sample. Product D is present in both samples, similar to PROTAC-1, but an increase in magnitude of signal over time is not observed to the same degree in PROTAC-2.

For PROTAC-2, there is not enough in the MS data to make any meaningful conclusions about the route of degradation in human blood plasma.

Human Liver Microsomes

From the clearance assay, HLMs are known to degrade PROTACs at a significant rate. This potentially enables some further insight into the metabolic pathway, as it can allow secondary and potentially tertiary metabolites to be observed. For this reason, three timepoints were investigated for each PROTAC: t=0 min, t=5 min and t=30 min.

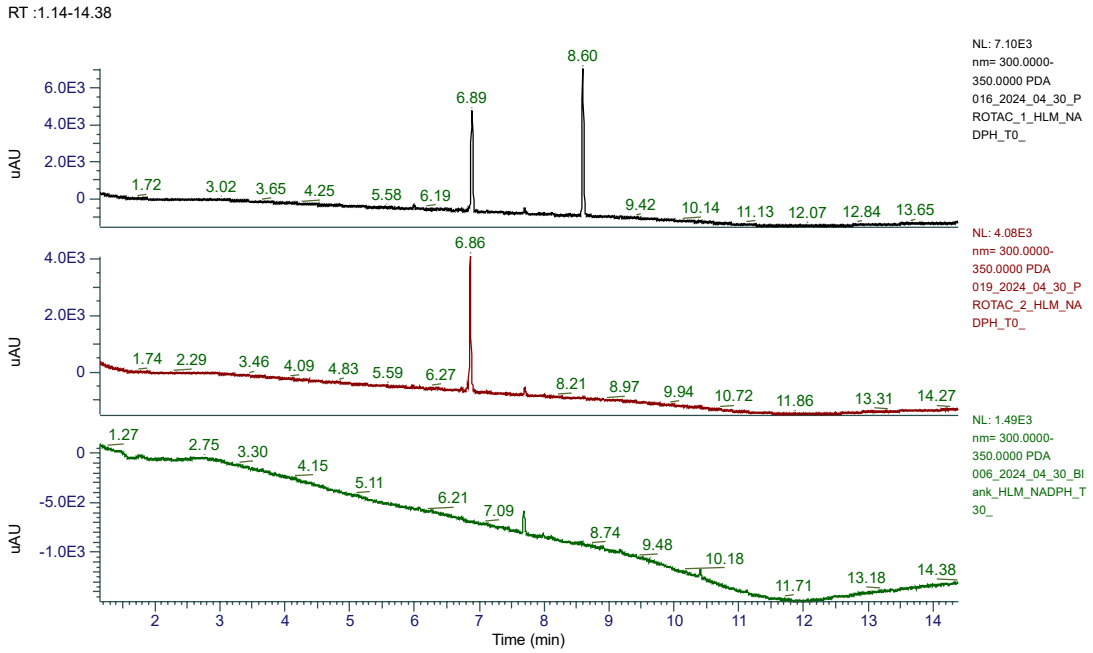


Figure 55: A comparison between the PDA (300-350nm) outputs of PROTAC-1 (black) and PROTAC-2 (red), both HLMs alongside a blank HLM sample. PROTAC samples are taken at t=0 min, blank plasma at t=30 min.

Figure 55 shows the t=0 min of both PROTACs alongside a blank HLM sample, to rule out any peaks that come from the matrix. From this data, it can be concluded that the 6.89 and 8.60 peaks in the PROTAC-1 sample and the 6.86 peak in the PROTAC-2 sample are of interest.

For PROTAC-1, a comparison of the three timepoints (figure 56) shows the both the 6.89 and 8.60 peaks being degraded over time, (although the peak display size is similar, as the axis auto-scales, the 8.60 peak magnitude drops from 7.10 x 10^3 to 4.50 x 10^3 uAU) as well as newer peaks showing up (5.99 at t=5 min and numerous peaks between 5.50 and 6.50 for t=30 min). This is promising as it suggests that there would be no issue in locating metabolites in the MS.

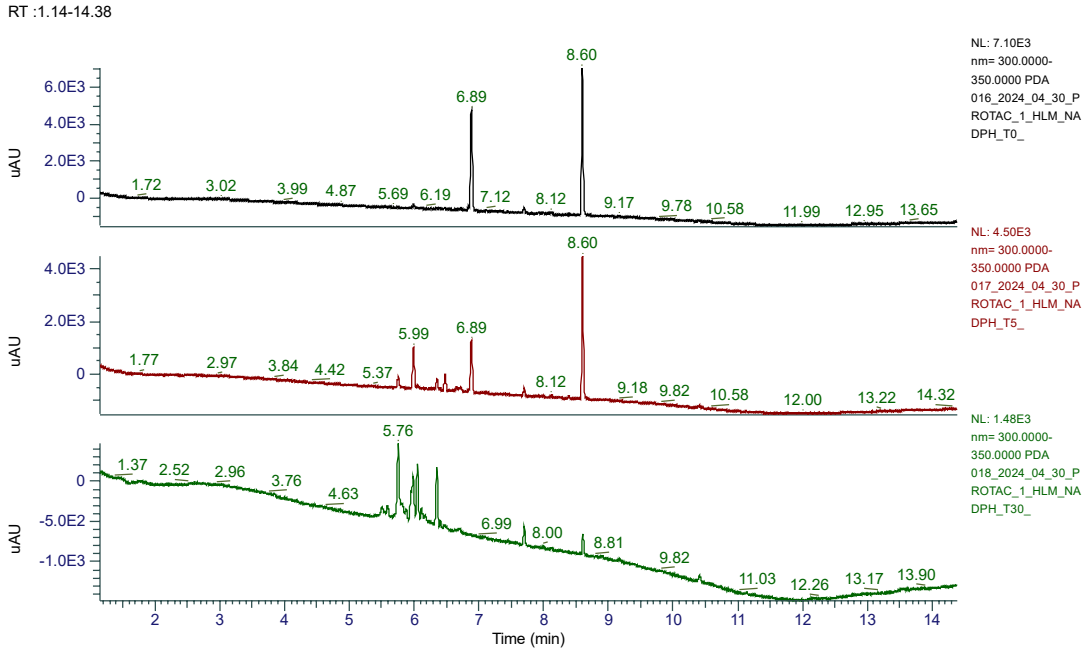


Figure 56: A comparison between the PDA (300-350nm) outputs of PROTAC-1 at t=0 min (black), t=5 min (red) and t=30 min (green) in HLMs.

Contrary to the blood plasma, the MS TIC for HLMs was comparatively clean (figure 57), with both peaks observed on the PDA being clearly visible, as well as peaks at 9.14 and 11.72, but those are unlikely to be the PROTAC-1 peak or related as it is known that the PROTACs show up at the absorbances covered by the PDA.

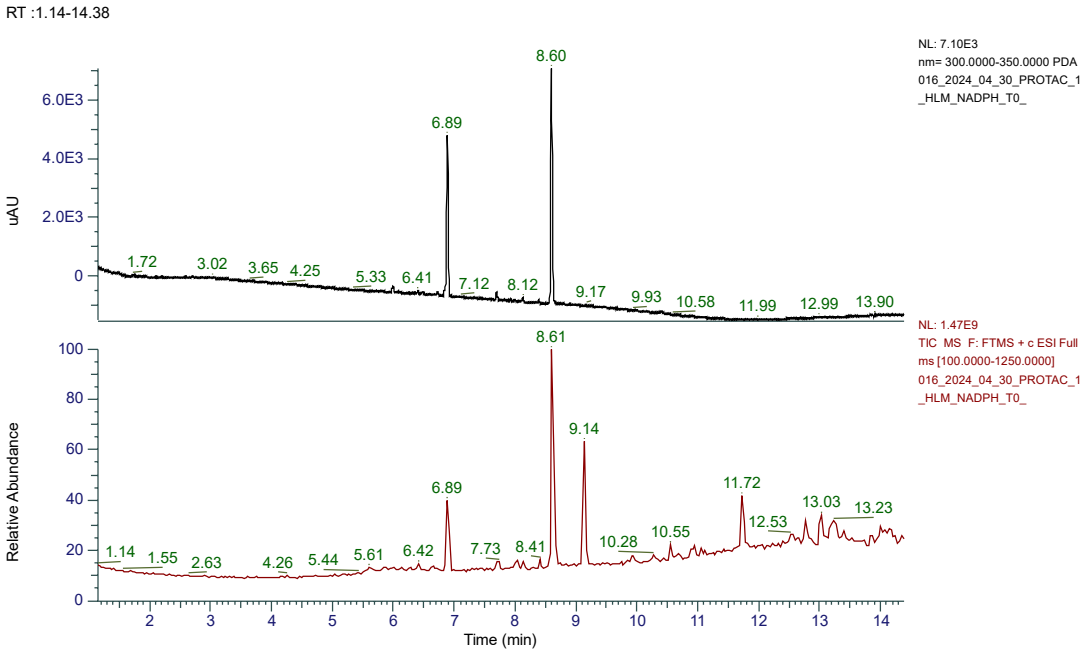


Figure 57: The PDA (300-350nm; black) and TIC (red) output of PROTAC-1 at t=0 min in HLMs

Due to having a retention time similar to previous LCMS runs, the 6.89 peak was likely to be the PROTAC-1 peak, which was confirmed by the MS at 6.89, which showed a peak at 785.2257, alongside a peak explained by chlorine isotope mass. The MS^2 of this peak (figure 58) showed fragmentation that corresponds to PROTAC-1, putting it beyond doubt that the 6.89 peak is PROTAC-1.

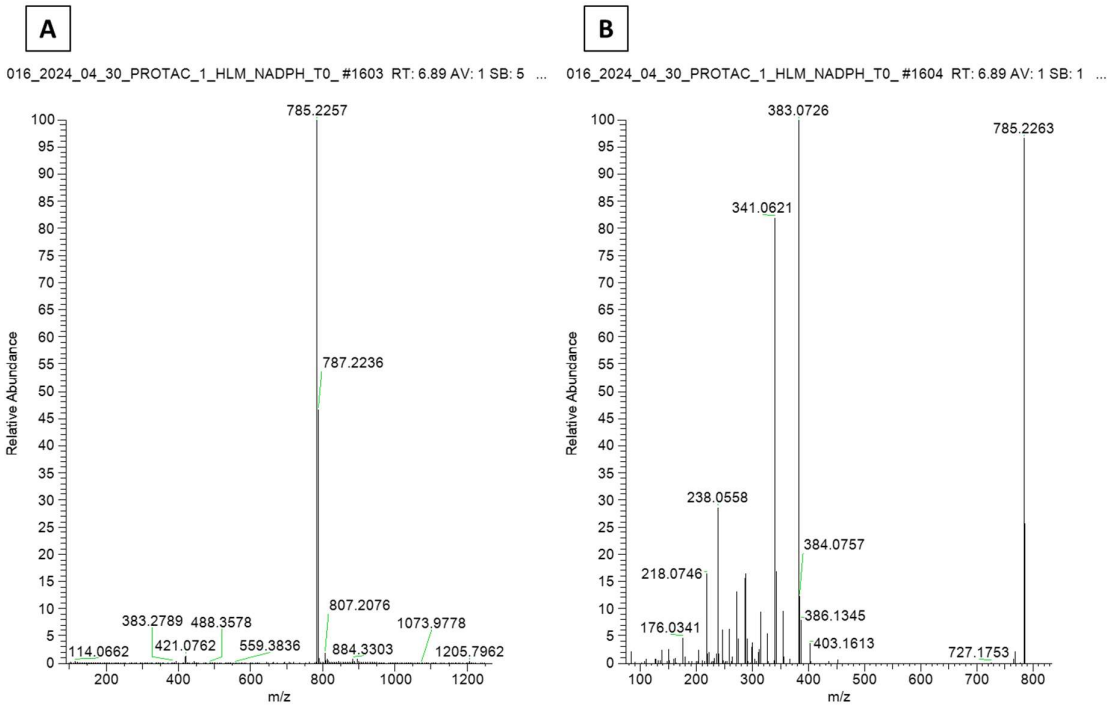


Figure 58: A: The background-subtracted MS of t=0 min PROTAC-1 in HLMs at 6.89 minutes. B: The MS^2 at 785.2260 and 6.89 minutes of PROTAC-1 in HLMs at t=0 min.

As 6.89 appears to be the peak for PROTAC-1, it does raise the question of what the peak at 8.60 could be. Looking at the MS and MS^2 (figure 59) gave no useful insights as to what it is, as the mass in question doesn't align with anything involved, and it doesn't fragment well under these conditions. It is most likely a contaminant of some sort.

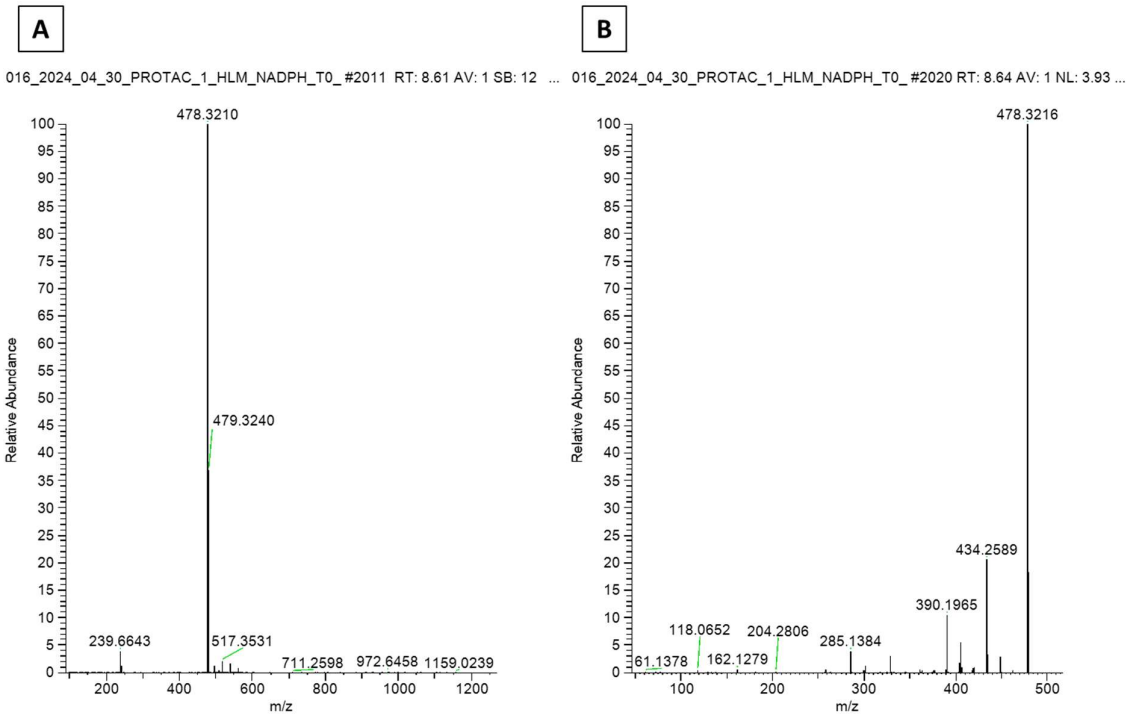


Figure 59: A: The background-subtracted MS of t=0 min PROTAC-1 in HLMs at 8.61 minutes. B: The MS² at 478.3210 and 8.64 minutes of PROTAC-1 in HLMs at t=0 min.

At t=5 min, from the PDA and MS TIC (figure 60) show a peak at 5.99 (PDA) and 6.01 (TIC) that looks worth investigating. Other peaks either don't show up on the PDA or the TIC suggesting that they are not of interest.

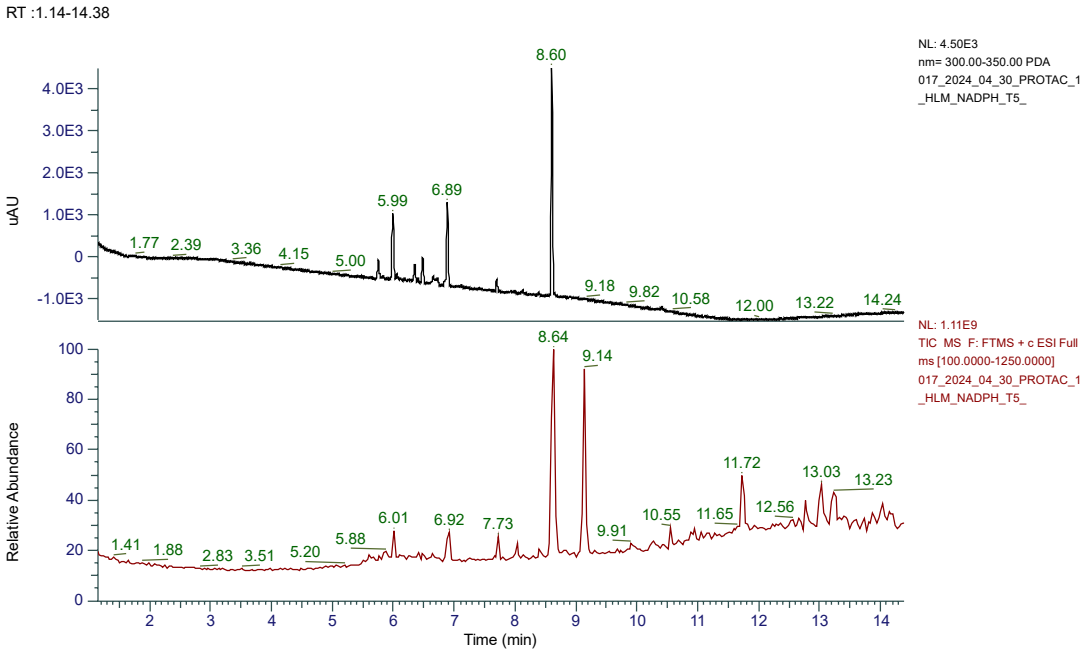


Figure 60: The PDA (300-350nm; black) and TIC (red) output of PROTAC-1 at t=5 min in HLMs

Figure 61 shows the MS at 6.01 compared to the MS at 6.92 (parent molecule peak) in the t=5 min sample. The 6.01 peak shows as approximately 16 daltons higher than the parent peak, which suggests either the addition of oxygen (via a carbonyl) or the substitution of a proton with a hydroxyl group.

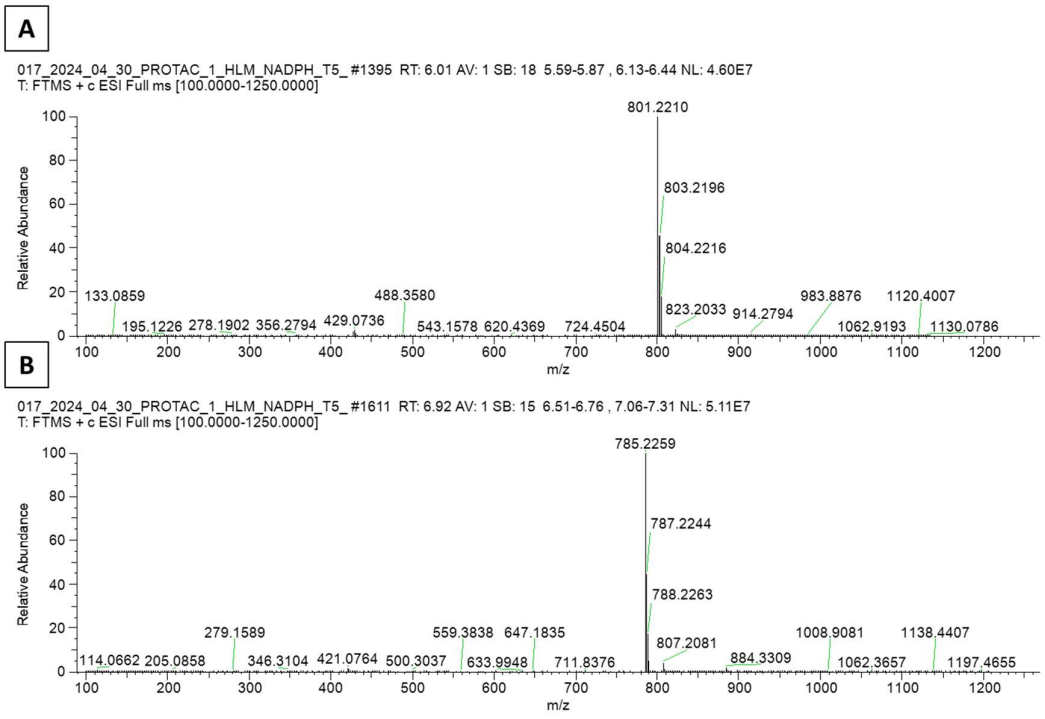


Figure 61: A: The background-subtracted MS of t=5 min PROTAC-1 in HLMs at 6.01 minutes. B: The background-subtracted MS of t=5 min PROTAC-1 in HLMs at 6.92 minutes.

Using the MS^2 data (figure 62), it can be both confirmed as a PROTAC-1 metabolite and further information can be deduced about the location of the modification, as the 357 and 399 fragments observed in figure 62 A are also 16 up from the familiar 341 and 383 fragments seen in figure 62 B and previous PROTAC samples, suggesting that the modification is on the JQ1 section of the molecule.

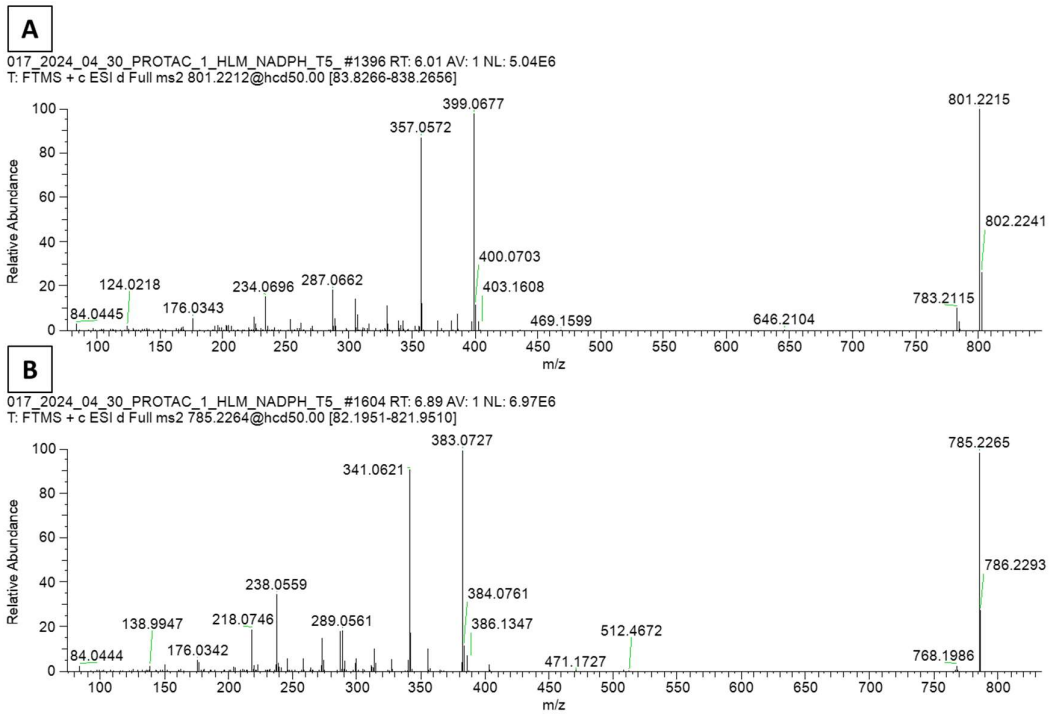


Figure 62: A: The MS² at 801.2212 and 6.01 minutes of PROTAC-1 in HLMs at t=5 min. B: The MS² at 785.2264 and 6.89 minutes of PROTAC-1 in HLMs at t=5 min.

At t=30 min, the parent peak is no longer visible on either the PDA or the TIC (figure 63), and the region around 5.50-6.50 that has numerous peaks present on the PDA does not appear to have as many peaks on the TIC, and the peaks in question appear lower in magnitude relative to the background than on the PDA. Figure 64 shows a clearer view of specific peaks in that region, showing two main peaks on the TIC, at 5.76 and 6.00. The 5.76 peak appears to correlate with the PDA peak at the same retention time and the 6.00 peak which is likely to correspond to the 5.99 PDA peak.

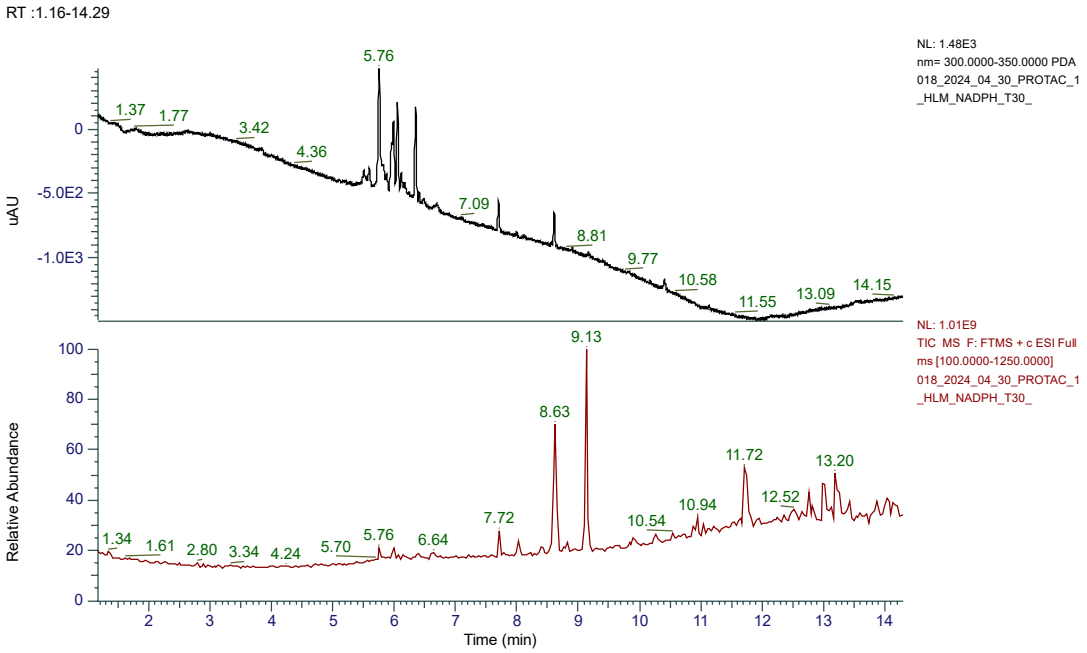


Figure 63: The PDA (300-350nm; black) and TIC (red) output of PROTAC-1 at t=30 min in HLMs

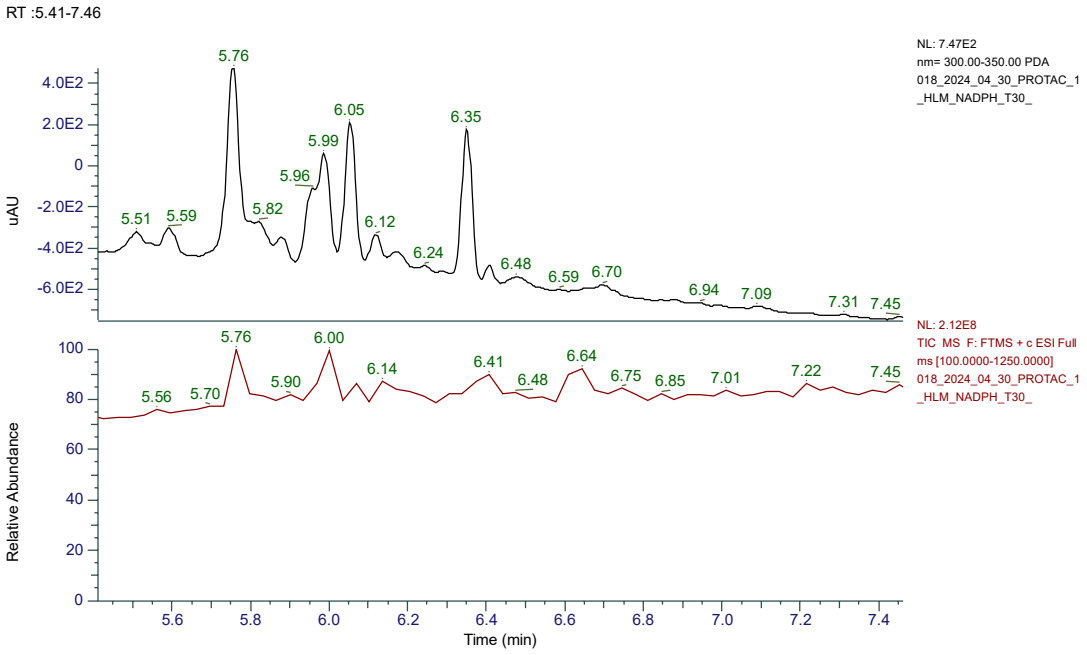


Figure 64: The PDA (300-350nm; black) and TIC (red) output of PROTAC-1 at t=30 min in HLMs at retention times between 5.41 and 7.46 minutes

Figure 65 shows the MS at the 5.76 and 6.00 timepoints, which shows that the 6.00 peak is most likely the +16 molecules observed in t=5 min sample, and the 5.76 peak appears to be a second +16 from that. This suggests that a similar modification is occurring at another point on the molecule.

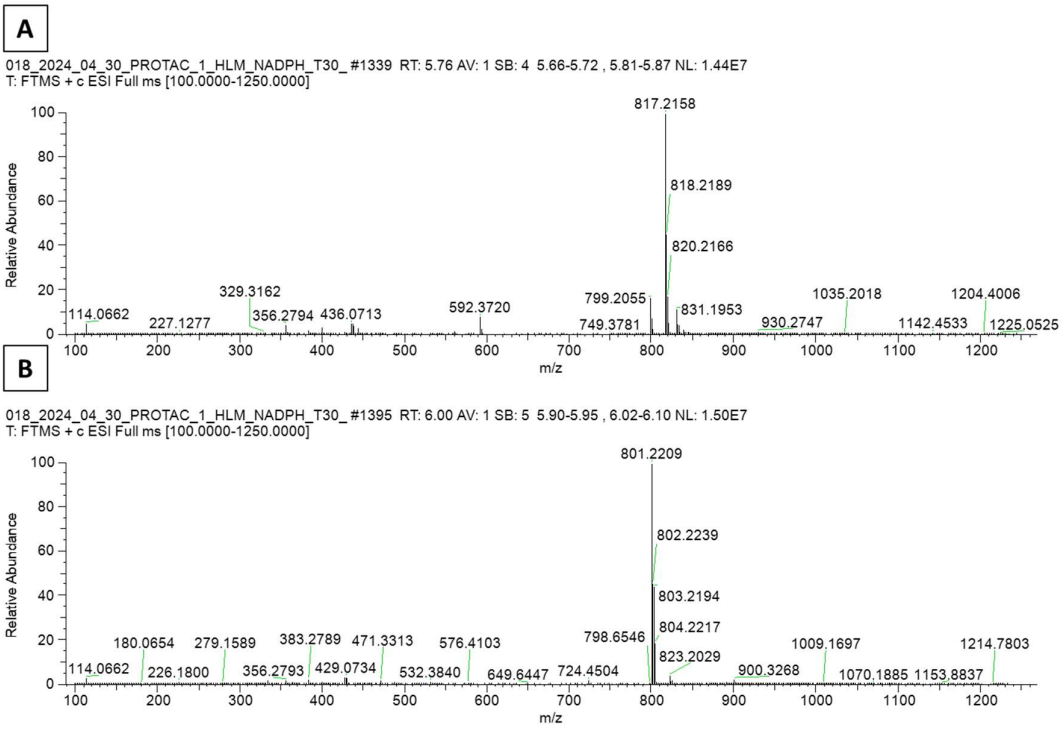


Figure 65: A: The background-subtracted MS of t=30 min PROTAC-1 in HLMs at 5.76 minutes. B: The background-subtracted MS of t=30 min PROTAC-1 in HLMs at 6.00 minutes.

The location of the second modification appears to be also on the JQ1 region of the PROTAC, with figure 66 showing the fragmentation pattern, where fragments consistent with the JQ1 associated fragments +32 daltons appear (373 and 415).

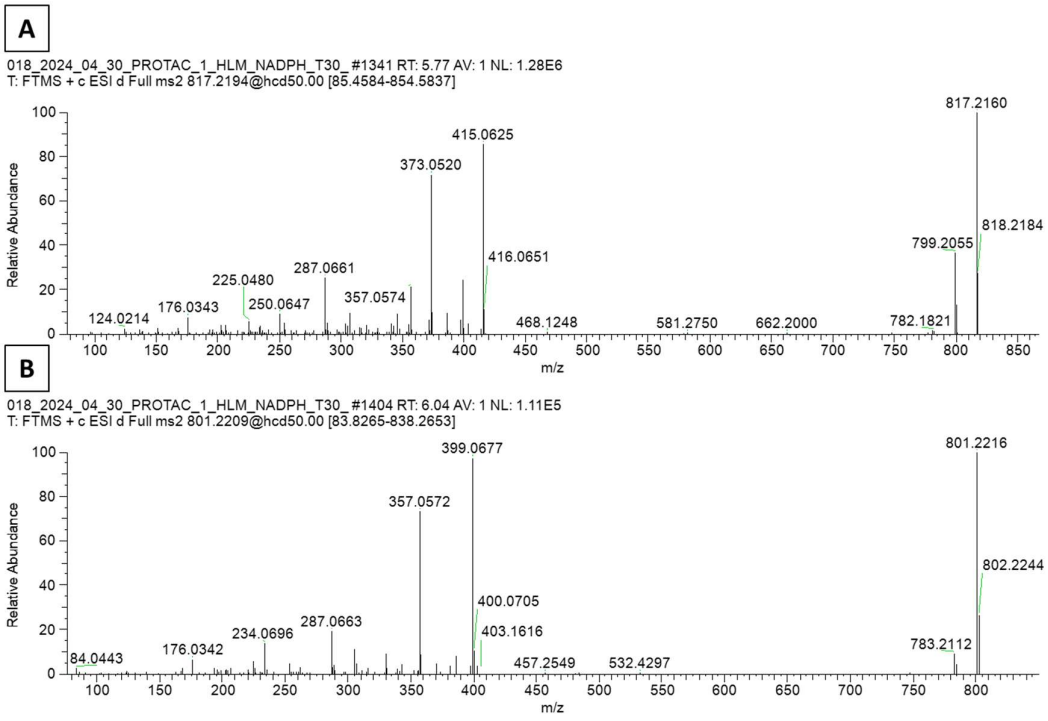


Figure 66: A: The MS² at 817.2194 and 5.77 minutes of PROTAC-1 in HLMs at t=30 min. B: The MS² at 801.2209 and 6.00 minutes of PROTAC-1 in HLMs at t=30 min.

The 6.05 peak on the PDA (figure 64) did align with a small peak on the TIC, unlabelled but at 6.07. This peak has a mass of PROTAC-1 +30, (figure 67 A) and fragments that suggest (similarly to other metabolites) that the JQ1 region is the modified section.

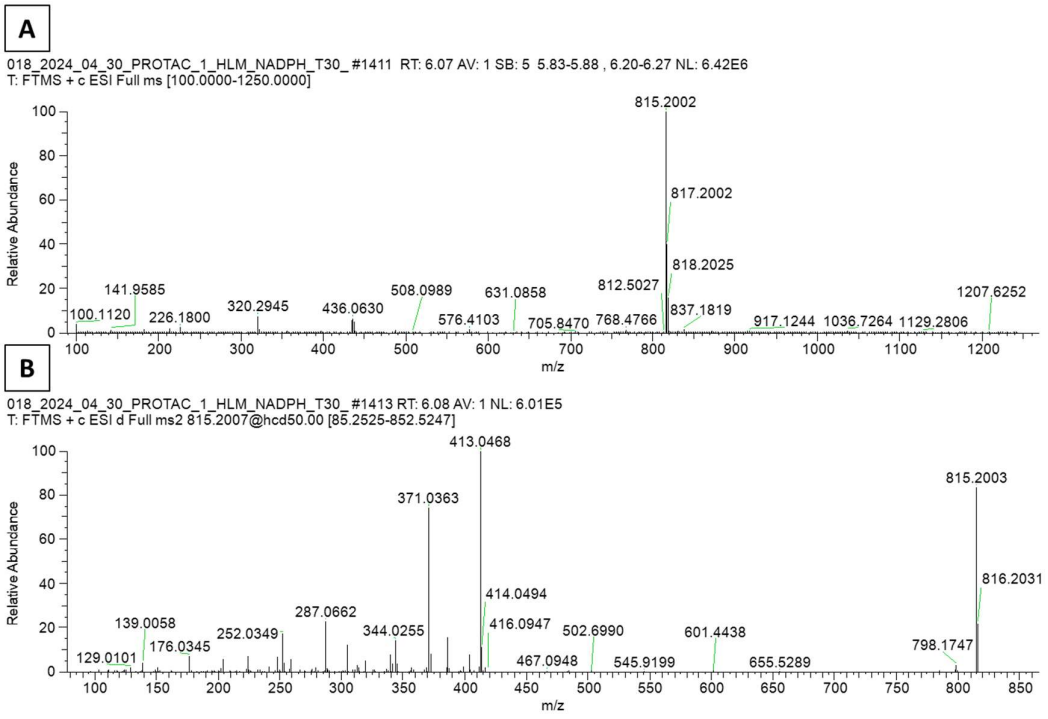


Figure 67: A: The background-subtracted MS of t=30 min PROTAC-1 in HLMs at 6.07 minutes. B: The MS^2 at 815.2007 and 6.08 minutes of PROTAC-1 in HLMs at t=30 min.

A +30 mass is consistent with a +32 modification that has lost a hydrogen molecule, (mass of approximately 2) which could be the result of an oxidation of a hydroxyl to a carbonyl. This suggests that at least one of the previous +16 modifications was the substitution of a proton with a hydroxyl group.

The same process was undertaken for PROTAC-2, figure 68 shows the PDA for the three timepoints tested, as noted before the peak seen in the PROTAC-1 sample at 8.60 is not present in this sample, and there are differences with the pattern of metabolites appearing.

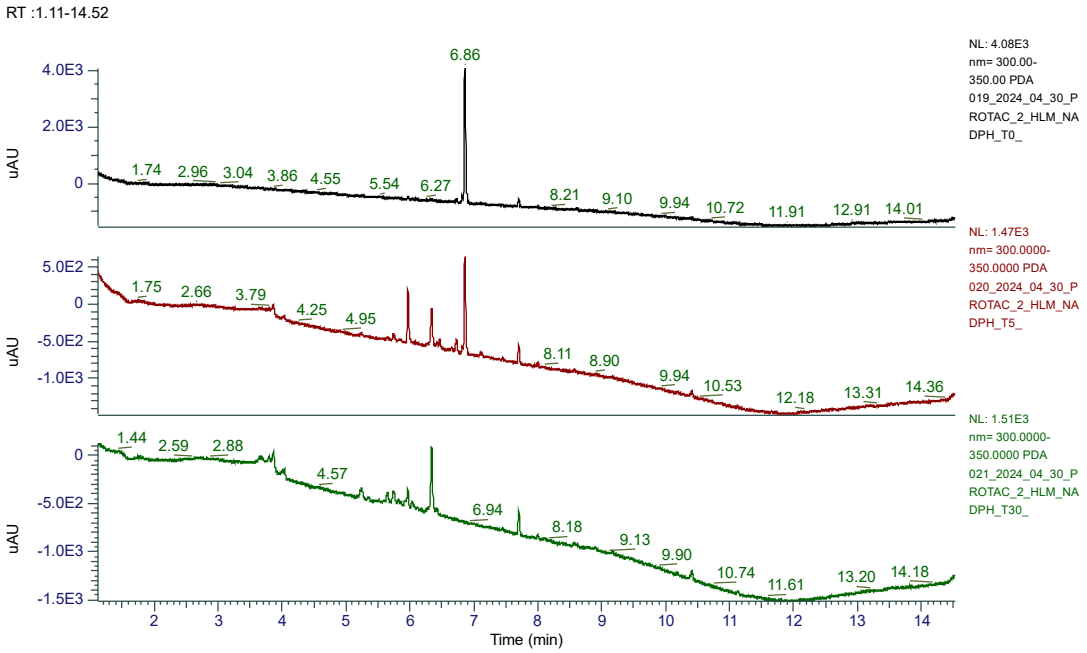


Figure 68: A comparison between the PDA (300-350nm) outputs of PROTAC-2 at t=0 min (black), t=5 min (red) and t=30 min (green) in HLMs

At the first timepoint (t=0 min), the TIC (figure 69) showed a peak that appeared to match with the 8.63 peak observed in the PROTAC-1 sample, which may indicate that the mass observed at 8.63 earlier isn't responsible for the PDA signal seen at 8.60.

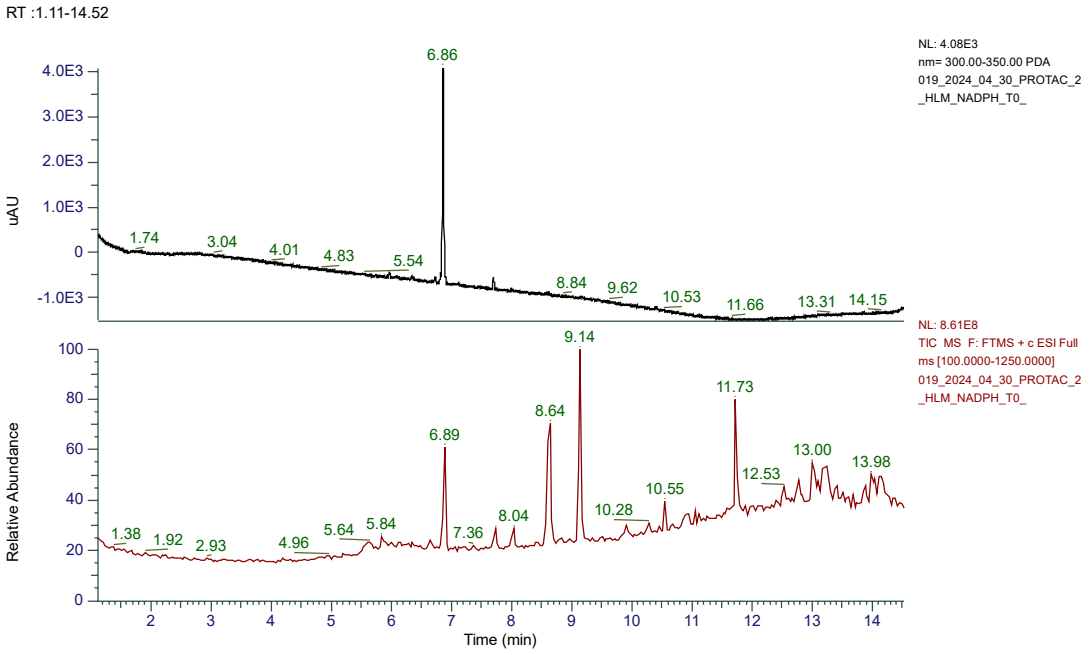


Figure 69: The PDA (300-350nm; black) and TIC (red) output of PROTAC-2 at t=0 min in HLMs

The peak at 6.89 would appear to be the parent PROTAC-2 peak, which is supported by the MS data for that retention time (figure 70), which closely matches the mass seen in the LC-MS earlier.

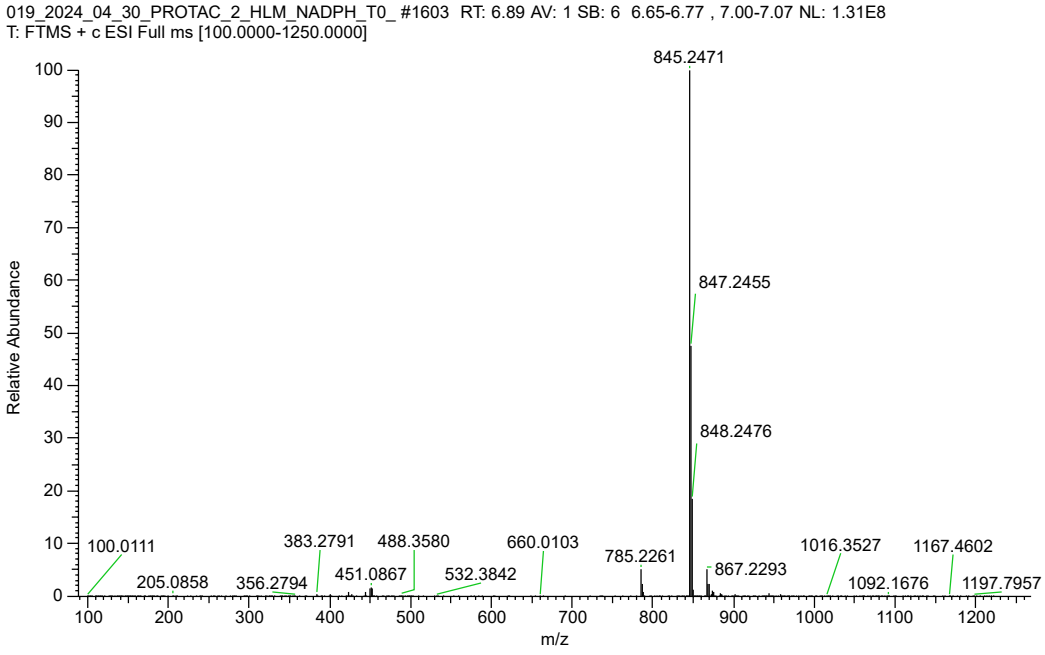


Figure 70: The background-subtracted MS of t=0 min PROTAC-2 in HLMs at 6.89 minutes

In the t=5 min sample, the peak assignment on the PDA (figure 71) wouldn't recognise the three major peaks properly, which were (left to right) 5.97, 6.34 and 6.86.

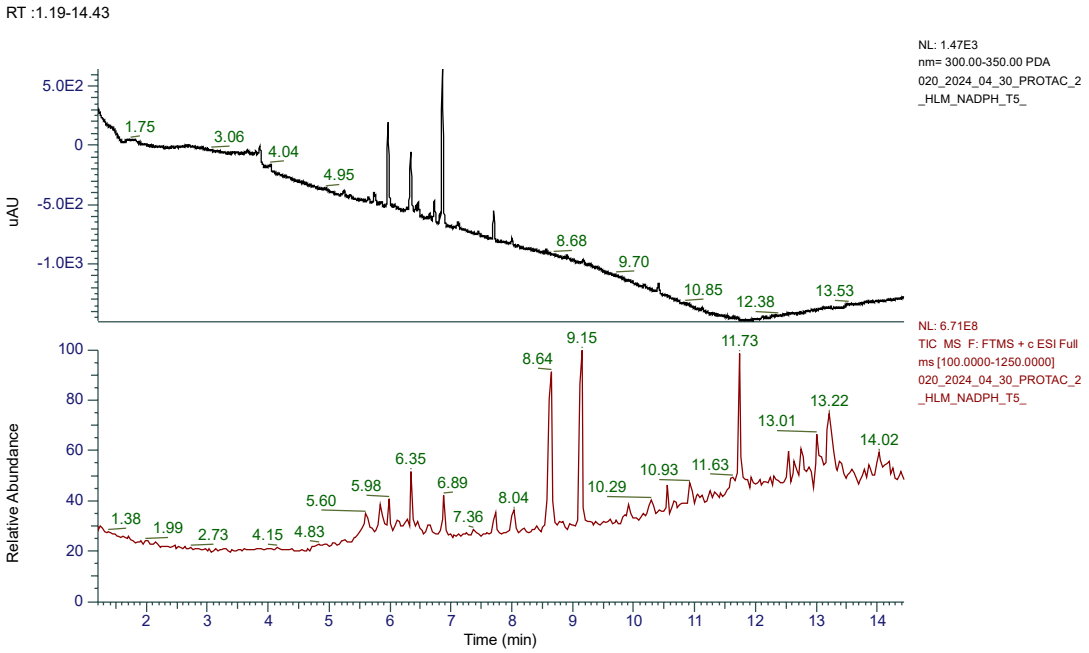


Figure 71: The PDA (300-350nm; black) and TIC (red) output of PROTAC-2 at t=5 min in HLMs.

The TIC peaks at 5.98 and 6.35 would seem the most likely to be metabolites. Comparing the 5.98 and 6.89 peaks (figure 72) shows a similar +16 pattern to what was observed in PROTAC-1, a net gain of one oxygen molecule.

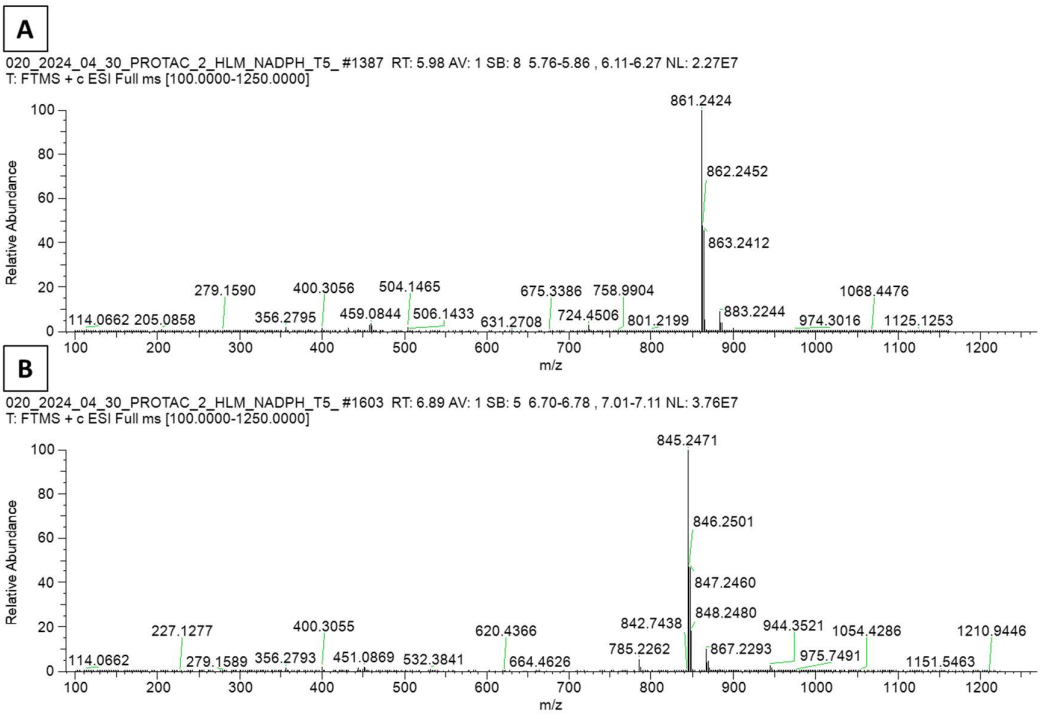


Figure 72: A: The background-subtracted MS of t=5 min PROTAC-2 in HLMs at 5.98 minutes. B: The background-subtracted MS of t=5 min PROTAC-2 in HLMs at 6.89 minutes.

MS^2 (figure 73) of the 5.98 peak shows that, again like PROTAC-1, that the JQ1 section of the molecule is the bit undergoing modification, showing the same 357 and 399 fragments observed earlier in the PROTAC-1 metabolite.

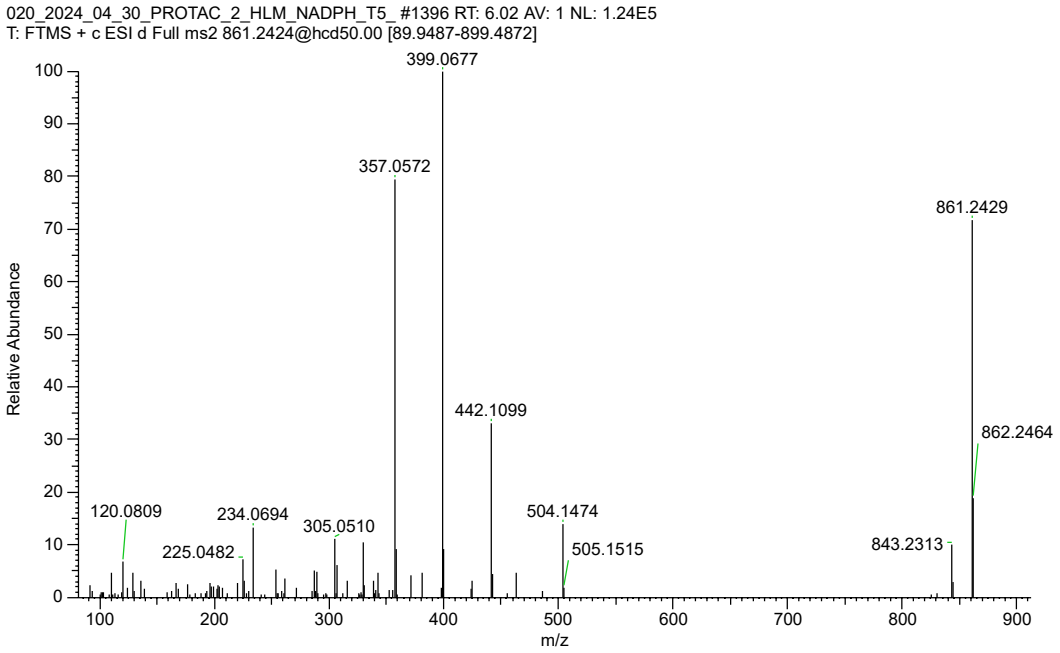


Figure 73: The MS² at 861.2424 and 6.02 minutes of PROTAC-2 in HLMs at t=5 min.

The peak at 6.38 is not as readily identifiable, the MS (figure 74) shows a mass of 444.1252, which doesn't neatly align to any theoretical product or derivative from PROTAC-2. It is possible that it is a contaminant or a multiply-charged molecule but without further data it can't be determined what the peak is. Unfortunately, for whatever reason the peak wasn't selected for MS².

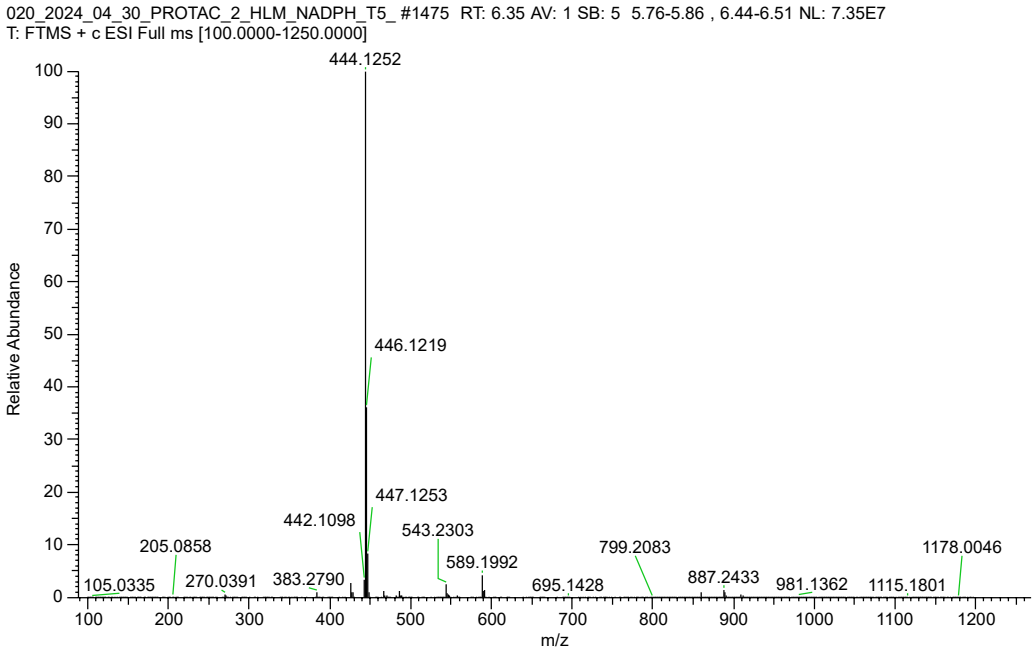


Figure 74: The background-subtracted MS of t=5 min PROTAC-2 in HLMs at 6.35 minutes.

Figure 75 shows that at t=30 min there is a region between 5 and 6 minutes that seems to be comprised of several peaks on the PDA, although none are clearly visible on the TIC, the peak at 6.34, although unlabelled on the PDA, is present in both PDA and TIC.

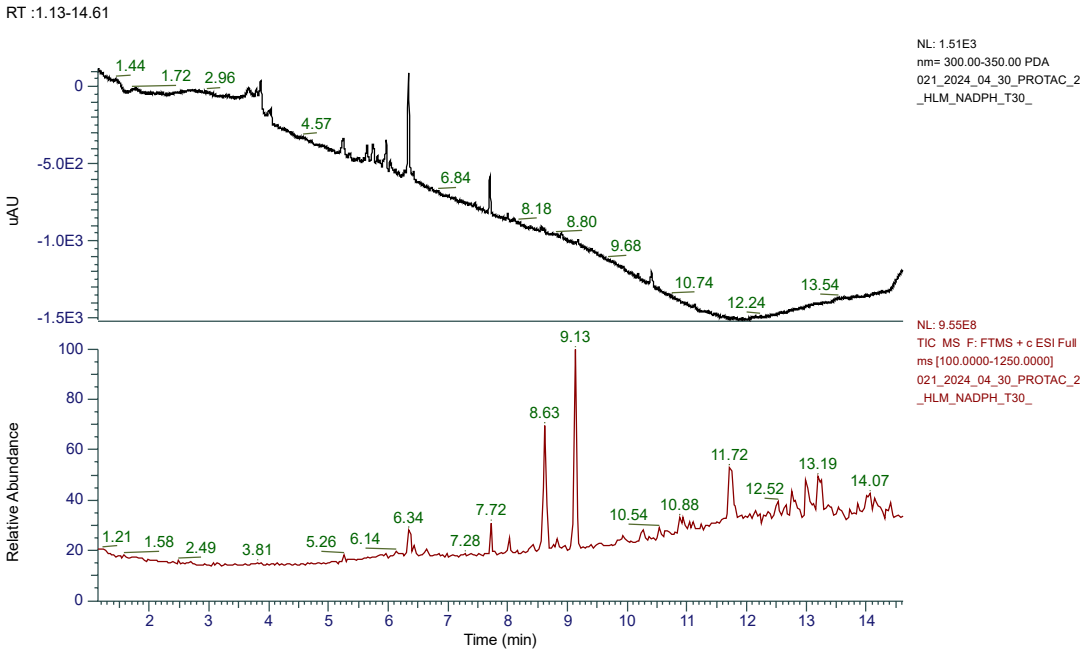


Figure 75: The PDA (300-350nm; black) and TIC (red) output of PROTAC-2 at t=30 min in HLMs.

The 6.34 peak had a mass of 444.1252 (figure 76), which supports the idea that it is the same peak as in the t=5 min sample.

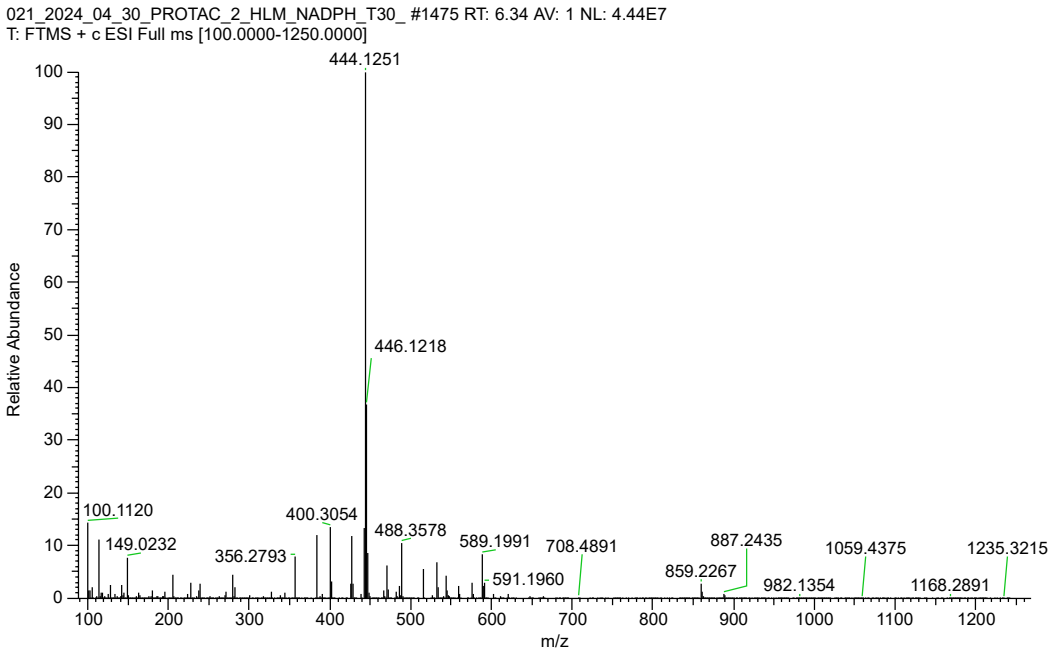


Figure 76: The background-subtracted MS of t=30 min PROTAC-2 in HLMs at 6.34 minutes.

Unlike in the t=5 min sample, however, this time the peak was selected for MS^2 (figure 77), which showed fragments at 341, 383 and 426, similar to PROTAC-2. This suggests that it is some kind of PROTAC derivative, but it is unclear what it could actually be.

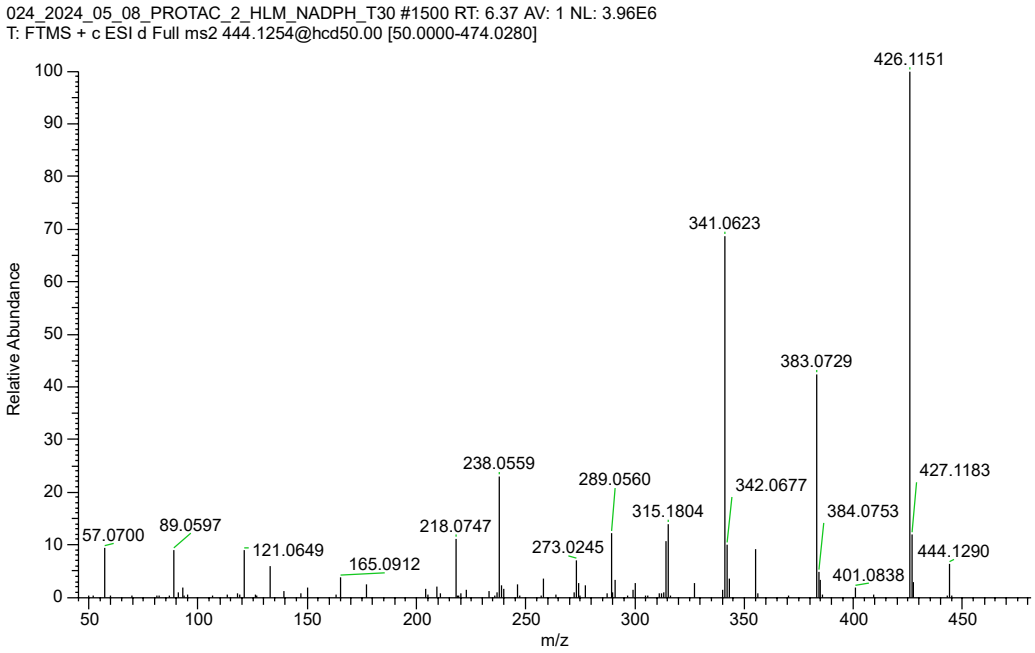


Figure 77: The MS^2 at 444.1254 and 6.37 minutes of PROTAC-2 in HLMs at t=30 min.

Human Hepatocytes

Due to the availability of resources and the time frame being worked in, a separate metID assay for hepatocytes could not be performed. However, due to the unexpected results of the clearance assay, an attempt was made to try and use the samples used in the clearance assay in a metID experiment but unfortunately the concentration was too low to identify any metabolites.

Cell Assays

In order to compare their function, assays were designed to measure the effect of the PROTACs on tissue cultures. The BET family proteins that JQ1 binds to are ubiquitously expressed across various tissue types, so it wasn't necessary to design or select cultures with expression in mind. The cancerous cell lines A549 (lung) and MCF7 (breast) were chosen based on availability and that they had been used in a previous study using dBET1 (PROTAC-1).[19]

Unfortunately, it took some time to procure the cells, which put some time pressure on performing the assay, and after some passaging it was observed that MCF7 did not grow quickly enough to have sufficient cells available for the assay, and consequently it was abandoned and the A549 line became the sole focus.

The assay design was not complicated, BRD4 levels would be monitored via western blot in varying concentrations of PROTAC dose. A blank DMSO and untreated controls would also be included. Each sample would also have the amount of β-actin measured as a baseline protein to control for differences in cell number. This assay would allow a comparison of cellular BRD4 levels with varying levels of PROTAC, allowing the effectiveness of the PROTAC to be measured at causing the degradation of the BRD4. This would allow an insight into whether the linker choice would have an impact on PROTAC function.

The first assay to be undertaken was done with PROTAC-2. The blot of this assay (figure 78) was inconclusive. BRD4 has a mass of around 150kDa, but the bands on this gel (figure 78, A) ran above 250kDa. In absence of a positive control, it cannot be safely concluded that this protein is the protein of interest. The BRD4 band widths seem to correlate closely with the widths of β-actin, with no obvious differences.

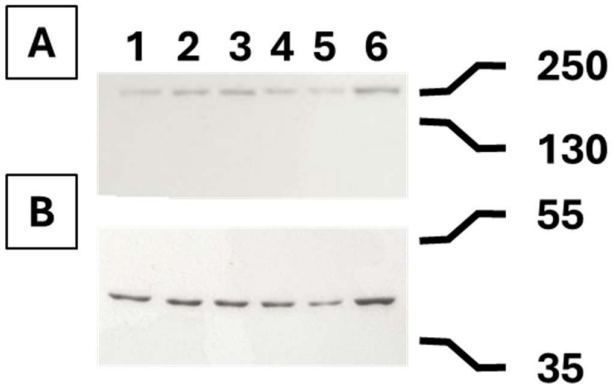


Figure 78: Western blots (A: BRD4 and B: β-actin) after treating A549 (lung) cell tissue with PROTAC-2 for 4 hours. Lanes 1-4 are 100 nM, 50 nM, 20 nM, 10 nM of PROTAC in the treatment (in that order)|, lane 5 is a DMSO only treatment and lane 6 is an untreated (no DMSO) sample. The numbers to the right of the bands indicate approximate protein masses in kDa.

Based on this western blot, no conclusions can be made on the function of PROTAC-2, and ideally the assay would be repeated with a positive control (using a BRD4 standard) and varying the concentrations.

Due to time constraints, it became a choice between repeating the PROTAC-2 and then not having a PROTAC-1 sample to compare it to, or attempting the same methodology for PROTAC-1 and not having a decent PROTAC-2 sample to compare it to. Based on the fact that either decision would not result in useful data, the cell assays were abandoned with the hope that they can be repeated in the future.

Discussion

Synthesis

Ether-Connected Linkers

This project set out to design a synthetic route to build multiple different PROTACs, and the route developed can theoretically be adapted to create a cereblon-binding PROTAC for any protein that has an established ligand/inhibitor. By stopping short of the last amide coupling step (reactions E and H from figures 4 and 5 respectively) a stock of a molecule comprised of cereblon-ligand and linker can be built up which enables any molecule with a carboxylic acid to substitute for the JQ1 derivative. This is useful as it reduces the amount of time required in any future projects that require a PROTAC, by reducing the design requirements. In order to make a new ether-linked PROTAC for a new protein of interest, the majority of the synthesis of the new PROTAC is shared with the ones developed in this project, so no time and resources need to be spent on re-development.

Amide-Connected and Biphenyl-Connected Linkers

Optimisation of these linkers is worth revisiting, as different connection types would allow a different dimension to linker comparison.

The primary issue to overcome with the amide connection is the lack of reactivity of pomalidomide. Even with a highly reactive acyl halide, there was still substantial unreacted material left over. This could be down to $-NH_3^+$ formation, which a base might prevent but due to the unwelcome ketene formation it would have to be weaker than triethylamine and some experimentation with various bases might be required to find a suitable one.

With the biphenyl connection, attempting various catalysts, conditions and ratios and selecting those that produce the highest yields of the desired product could result in a route to synthesis, although it is worth noting that it might not be possible to get it to a point where it is economically viable.

Stability and Metabolism

Linkers

The main aim of the project was to evaluate the metabolic differences between the two linkers, but due to the JQ1 molecule's drastically shorter half-life than the linker connections appeared to be, no meaningful difference between the two can be ascertained through this investigation. It appears that neither has a rapid degradation so both would appear suitable for a drug but without further testing this cannot be concluded with any certainty.

However, haemolysis has been observed for flexible linkers such as these,[22,41] although it was not seen in the plasma assays undertaken in this project. This could, however, have been a result of an insufficient assay time, and a longer assay may reveal a difference between the linkers that cannot be observed in the limited timeframe used.

In mouse liver microsomes (MLMs), half-lives have been observed for a similar linker to PROTAC-1 at 2.4 minutes, which is approximately 1.7 times longer than the 1.39 minutes observed in this project.[41] The CRBN ligand differed (figure 79), however, between the two molecules which may explain the difference to some extent, but it could also be that the different matrix (as this project used HLMs) explains the difference.

Interestingly, changing the linker (and the linker alone) does appear to have an effect on PROTAC stability in MLMs, which seems counterintuitive as all of the metabolism observed in this project occurred on the BRD4 ligand, not on the linker.

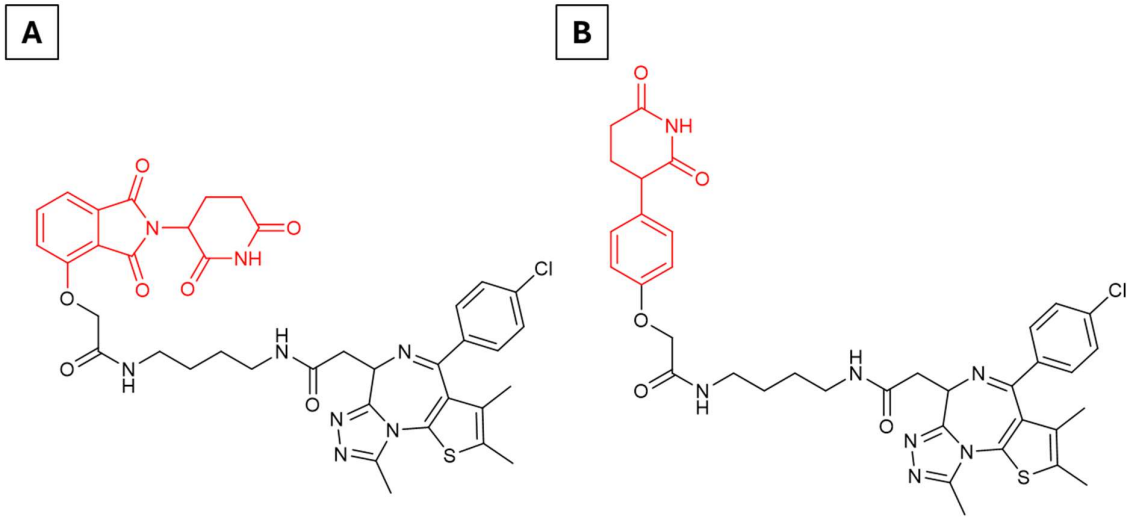


Figure 79: A: PROTAC-1 B: The PROTAC investigated in Actis et al.[41] *Coloured in red are the differing CRBN ligands between the two molecules.*

Thalidomide

The imide region of thalidomide has been known to spontaneously hydrolyse in aqueous environments for a long period of time,[18] and this looked to be observed to a small degree in this study. Further metabolism work on the PROTACs – a longer incubation in blood plasma – could show this to a greater degree, and due to the difficulty in finding metabolites in this project for the plasma assay it might be beneficial to undertake this anyway.

JQ1

JQ1 is known to have a very low half-life *in vivo*,[14,20] which the findings of this study support as in microsomes it seems that the main site of metabolism of the PROTACs was on the JQ1 region of the molecules.

Metabolic work on JQ1[20] aligns with this project's findings, with monohydroxylation being the primary route of metabolism in mouse and human liver microsomes.

The monohydroxylation and subsequent metabolic modifications observed in this project raise the hydrophilicity of the molecule, which is a known method of clearance by the liver, with hydrophilic molecules more soluble in blood plasma and also capable of being excreted by the kidneys.

The second phase in liver drug metabolism is the conjugation of very hydrophilic, often charged groups to the parent molecule, such as glutathione or sulfate. This is followed by a third phase, which involves potential further modification and excretion, utilising the modifications as ligands to membrane transport proteins to facilitate removal from the cell and in the case of the kidneys, the body (via urine).

As the second phase occurs not in the ER, but in the cytoplasm, the second phase of drug metabolism was not able to be observed in this project, but if the hepatocyte sample were to be incubated for a longer time with higher concentrations it could be possible to observe this second phase and gain a better understanding of the metabolic process.

Although, as a result of JQ1 being the sole region of metabolism observed in HLMs in our assay, it made comparison between the two linkers less feasible as neither was directly involved in the observed metabolic pathway.

In order to get a better idea of the roles of the linkers in metabolism, substitution of JQ1 with a different BET ligand (many alternative ligands exist)[13] with a significantly longer half-life should enable an observation of the metabolism of the two linkers as opposed to just the metabolism of one of the ligand regions.

If metabolism were the only factor of interest, and the function of the PROTAC were no considered, then the substitute does not necessarily have to be a ligand, any non-cytotoxic molecule with a long half-life can be substituted here to achieve the same effect.

Hepatocyte Irregularity

The irregularity of the PROTACs being degraded by heat-inactivated hepatocytes at such a significant rate warrants revisiting. The buffer pH was no different from that used in other assays, in fact no other conditions besides the matrix differed, so it likely is due to the effect of the hepatocytes.

Heat inactivation should, in theory, denature the vast majority (if not all) of the proteins present. This should prevent any metabolism from being able to be performed, and the remaining degradation effects being solely due to small-molecule and intra-molecular interactions. The degradation even in heat-inactivated hepatocytes invalidates any conclusions from this stability assay as nothing can be attributed solely to the normal cellular function of the hepatocytes.

This would not seem to account for the results of the assay, and understanding the reasons could enable better molecular design. For this reason, it would be ideal if metabolite identification could be performed on both active and heat-inactivated hepatocytes.

Relevance of Molecules

dBET-1

The linker choice at the start of this project was driven by the prevalence of dBET1 (PROTAC-1 in this project), which was an established PROTAC that had been shown to degrade BRD4 in a wide variety of human cell lines.[19]

However, despite the *in vitro* impact of dBET1 being very high, and it being initially a promising candidate[10] for anti-cancer drug development, recent developments suggest that flexible linkers like the ones used in PROTAC-1 and PROTAC-2 suffer from poor pharmacokinetic properties *in vivo* compared to stiffer linkers formed of nitrogen heterocycles.[22] Although one of the purported benefits of the stiffer linker is a lower haemolysis rate, this project did not observe haemolysis to be much of

an issue, with microsome stability seeming to have much more of an impact. If revisiting this project, it might be beneficial to instead focus on a clinically relevant linker, with use of ARV-110 (which uses nitrogen heterocycles) as a basis to start instead of dBET1, as ARV-110 is currently undergoing clinical trials, and is a leading candidate for eventually reaching the market.[7]

Lipinski's Rule of 5

A commonly used method for screening orally bioavailable compounds is via Lipinski's rule of 5. This is a set of rules first published in 1997 that set parameters for drugs that have poor bioavailability. The rules state that to be orally active, a drug should have a molecular weight lower than 500 Da, 5 or fewer H-bond donors (calculated as the number of -OH or -NH groups), 10 or fewer H-bond acceptors (calculated as the number of O and N atoms) and the logarithm of the octanol/water partition coefficient (log P) should be lower than 5. These rules do not apply to substrates of transporters (such as vitamins).[42]

The PROTACs investigated in this project violate the molecular weight and hydrogen bond acceptor rules, which would suggest that they would not be orally bioavailable. A significant hurdle with making a PROTAC conform to the rule of 5 is in the minimum necessary size of the molecule. If you were to remove the linker entirely, the resulting molecule would still have a molecular weight of 657.1 Da. Similarly, either or both of the number of hydrogen bond donors or acceptors is likely to be over their respective limit.

In order to make a PROTAC more bioavailable, then a new modality needs to be explored. A different set of criteria for oral bioavailability is using the number of rotatable bonds, wherein 10 or fewer rotatable bonds corresponds to good oral bioavailability, alongside a low polar surface area (≤ 140 $Å^2$).[43] These different criteria set no upper limit on molecular weight (although the larger a molecule is, the probability of it violating the rules also increases).

The PROTACs investigated in this project do not conform to this criteria either, but replacing the linkers with a stiffer one (such as the nitrogen heterocycle linker found on ARV-110) can move them close to

if not within the criteria. Although in order to be fully orally active changing the E3 ligase or BRD4 ligands may be necessary.

Design of the PROTAC as a whole (with the E3 ligand, linker and POI ligand all considered together) has been recommended,[41] although if the goal is simply to minimise the total rotatable bonds and polar surface area, all three should be able to be optimised separately, and then only connectivity needs to be considered.

Conclusion

Due to issues with synthesis, linker connectivity couldn't be compared in this project beyond the comparison in synthesis difficulty. It cannot be ignored that the ether connection had a significantly simpler synthesis than the other two connections explored, as manufacture is a significant part of any drug. If it is complicated to make, the expense will likely be higher and keeping cost of manufacture low is important as it keeps the cost of any treatment low.

Between the ether-connected linkers, a slightly higher yield was observed in the ethylene glycol-linked PROTAC than in the aliphatic-linked PROTAC. However, these yields are still reasonably close together and consequently could likely differ only by chance.

When it comes to metabolism, both PROTACs showed somewhat similar patterns of behaviour, with PROTAC-1 being more stable in plasma and PROTAC-2 being more stable in human liver microsomes. Both PROTACs metabolised via oxygen addition (probably via hydroxylation) in HLMs, although this only occurred in a region of the PROTAC that was shared between the two. The patterns of hydrolysis in the plasma did appear to differ, but as it was on the limit of detection it cannot be concluded that the difference was meaningful.

No difference in function could be explored as the necessary assays were incomplete.

In summation, from the data gathered in this project, there is limited evidence of a significant effect on PROTAC synthesis or metabolism between an aliphatic linker and an ethylene glycol linker, suggesting that slight variation in the linker does not have a major impact on a PROTAC.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: A Cancer Journal for Clinicians. 2024 Apr 4;74(3):229–63.
2. Iwasa Y, Nowak MA, Michor F. Evolution of Resistance During Clonal Expansion. Genetics. 2006 Apr 1;172(4):2557–66.
3. Bray F, Piñeros M. Cancer patterns, trends and projections in Latin America and the Caribbean: a global context. Salud Publica Mex. 2016 Mar 9;58(2):104117.
4. Cyrus K, Wehenkel M, Choi E, Lee H, Swanson H, Kim K. Jostling for Position: Optimizing Linker Location in the Design of Estrogen ReceptorTargeting PROTACs. ChemMedChem. 2010;5(7):979–85.
5. Bondeson DP, Mares A, Ian, Ko E, Campos S, Miah AH, et al. Catalytic *in vivo* protein knockdown by smallmolecule PROTACs. Nature Chemical Biology. 2015;11(8):611–7.
6. Lai AC, Toure M, Hellerschmied D, Salami J, Jaime-Figueroa S, Ko E, et al. Modular PROTAC Design for the Degradation of Oncogenic BCR-ABL. Angewandte Chemie (International ed in English). 2016 Jan 11;55(2):807–10.
7. Békés M, Langley DR, Crews CM. PROTAC targeted protein degraders: the past is prologue. Nature Reviews Drug Discovery. 2022;21(3):181–200.
8. Sakamoto KM, Kim KB, Kumagai A, Mercurio F, Crews CM, Deshaies RJ. Protacs: Chimeric molecules that target proteins to the Skp1–Cullin–F box complex for ubiquitination and degradation. Proceedings of the National Academy of Sciences. 2001;98(15):8554–9.
9. Wu S, Lee Ay, Lai H, Zhang H, Chiang C. Phospho Switch Triggers Brd4 Chromatin Binding and Activator Recruitment for GeneSpecific Targeting. Molecular Cell. 2013;49(5):843–57.

10. Zhang K, Gao L, Wang J, Chu X, Zhang Z, Zhang Y, et al. A Novel BRD Family PROTAC Inhibitor dBET1 Exerts Great AntiCancer Effects by Targeting cMYC in Acute Myeloid Leukemia Cells. Pathology and Oncology Research. 2022;28.
11. Hua T, Wang H, Fan X, An N, Li J, Song H, et al. BRD4 inhibition attenuates inflammatory pain by ameliorating NLRP3 inflammasome-induced pyroptosis. Frontiers in Immunology. 2022;13.
12. Hao K, Jiang W, Zhou M, Li H, Chen Y, Jiang F, et al. Targeting BRD4 prevents acute gouty arthritis by regulating pyroptosis. International Journal of Biological Sciences. 2020 Jan 1;16(16):3163–73.
13. Kenneth, Xing E, Larue RC, Li P. BET Bromodomain Inhibitors: Novel Design Strategies and Therapeutic Applications. Molecules. 2023;28(7).
14. Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, et al. Selective inhibition of BET bromodomains. Nature. 2010;468(7327):1067–73.
15. Trabucco SE, Gerstein RM, Evens AM, Bradner JE, Shultz LD, Greiner DL, et al. Inhibition of Bromodomain Proteins for the Treatment of Human Diffuse Large Bcell Lymphoma. Clin Cancer Res. 2015 Jan 5;21(1):113–22.
16. Clague MJ, Heride C, Urbé S. The demographics of the ubiquitin system. Trends in Cell Biology. 2015;25(7):417–26.
17. Ottis P, Toure M, Cromm, Philipp M, Ko E, Gustafson JL, Crews CM. Assessing Different E3 Ligases for Small Molecule Induced Protein Ubiquitination and Degradation. ACS Chem Biol. 2017;12(10):2570–8.
18. Lu J, Qian Y, Altieri M, Dong H, Wang J, Raina K, et al. Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. Chemistry & Biology. 2015;22(6):755–63.
19. Luo X, Archibeque I, Dellamaggiore K, Smither K, Homann O, Lipford JR, et al. Profiling of diverse tumor types establishes the broad utility of VHLbased ProTaCs and triages candidate ubiquitin ligases. iScience. 2022;25(3).

20. SCHUMACHER H, L SR, T WR. THE METABOLISM OF THALIDOMIDE: THE SPONTANEOUS HYDROLYSIS OF THALIDOMIDE IN SOLUTION. British Journal of Pharmacology and Chemotherapy. 1965;25(2):324–37.
21. Li F, MacKenzie KR, Jain P, Santini C, Young DW, Matzuk MM. Metabolism of JQ1, an inhibitor of bromodomain and extra terminal bromodomain proteins, in human and mouse liver microsomes. Biology of Reproduction. 2020 Apr 14;103(2):427–36.
22. Li Y, Qu J, Jiang L, Peng X, Wu K, Chen M, et al. Application and challenges of nitrogen heterocycles in PROTAC linker. European Journal of Medicinal Chemistry. 2024;273:116520.
23. Cancer Research UK. Treatment for Cancer | Cancer in General | Cancer Research UK [Internet]. www.cancerresearchuk.org. 2021. Available from: https://www.cancerresearchuk.org/about-cancer/treatment
24. Zhang L, Zhou Y, Wu B, Zhang S, Zhu K, Liu CH, et al. A Handheld Visible Resonance Raman Analyzer Used in Intraoperative Detection of Human Glioma. Cancers. 2023 Mar 14;15(6):1752–2.
25. Baselga J, Albanell J. Mechanism of action of anti-HER2 monoclonal antibodies. Annals of Oncology. 2001 Jan 1;12(suppl 1):S35–41.
26. Cooley S, Burns LJ, Repka T, Miller JS. Natural killer cell cytotoxicity of breast cancer targets is enhanced by two distinct mechanisms of antibody-dependent cellular cytotoxicity against LFA-3 and HER2/neu. Experimental Hematology. 1999 Oct;27(10):1533–41.
27. Robert C, Soria JC, Eggermont AMM. Drug of the year: programmed death-1 receptor/programmed death-1 ligand-1 receptor monoclonal antibodies. European journal of cancer (Oxford, England : 1990). 2013 Sep;49(14):2968–71.
28. Singh H, Manuri PR, Olivares S, Dara N, Dawson MJ, Huls H, et al. Redirecting Specificity of T-Cell Populations For CD19 Using the Sleeping Beauty System. Cancer Research. 2008 Apr 15;68(8):2961–71.

29. Thirion S, Motmans K, Heyligen H, Janssens J, Raus J, Vandevyver C. Mono- and bispecific single-chain antibody fragments for cancer therapy. PubMed. 1996 Dec 1;5(6):507–11.
30. GeislerJ, Helle H, Ekse D, Duong NK, Evans DB, Nordbø Y, et al. Letrozole is Superior to Anastrozole in Suppressing Breast Cancer Tissue and Plasma Estrogen Levels. Clinical Cancer Research. 2008 Sep 30;14(19):6330–5.
31. Eri LM, Tveter KJ. A Prospective, Placebo-Controlled Study of the Antiandrogen Casodex as Treatment for Patients with Benign Prostatic Hyperplasia. The Journal of Urology. 1993 Jul 1;150(1):90–4.
32. Jacky J, Kalet I, Chen J, Coggins J, Cousins S, Drzymala R, et al. Portable software tools for 3D radiation therapy planning. International journal of radiation oncology, biology, physics. 1994 Autumn;30(4):921–8.
33. Osieka R, Bruntsch U, Gallmeier WM, Seeber S, Schmidt CG. cis-Diamino-dichloro-platinum (II) in the treatment of otherwise treatment-resistant malignant testicular teratoma (author's transl). Deutsche medizinische Wochenschrift (1946). 1976 Jun;101(6):191–5, 199.
34. Tomasz M. Mitomycin C: small, fast and deadly (but very selective). Chemistry & Biology. 1995 Sep 1;2(9):575–9.
35. Fry DW, Harvey PJ, Keller PR, Elliott WL, Meade M, Trachet E, et al. Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts. Molecular Cancer Therapeutics. 2004 Nov 1;3(11):1427–38.
36. Phelps MA, Lin TS, Johnson AJ, Hurh E, Rozewski DM, Farley KL, et al. Clinical response and pharmacokinetics from a phase 1 study of an active dosing schedule of flavopiridol in relapsed chronic lymphocytic leukemia. Blood. 2009 Mar 19;113(12):2637–45.
37. Hofmann RM, Pickart CM. In Vitro Assembly and Recognition of Lys-63 Polyubiquitin Chains. Journal of Biological Chemistry. 2001 Jul;276(30):27936–43.
38. Nelson DL, Cox MM. Lehninger Principles of Biochemistry. 7th ed. Macmillan Higher Education: Basingstoke; 2017.

39. French CA, Miyoshi I, Aster JC, Kubonishi I, Kroll TG, Dal Cin P, et al. BRD4 Bromodomain Gene Rearrangement in Aggressive Carcinoma with Translocation t(15;19). The American Journal of Pathology. 2001 Dec;159(6):1987–92.
40. Yang Z, Yik JHN, Chen R, He N, Jang MK, Ozato K, et al. Recruitment of P-TEFb for Stimulation of Transcriptional Elongation by the Bromodomain Protein Brd4. Molecular Cell. 2005 Aug;19(4):535–45.
41. Actis M, Cresser-Brown J, Caine EA, Steger M, Aggarwal A, Anand Mayasundari, et al. Evaluation of Cereblon-Directing Warheads for the Development of Orally Bioavailable PROTACs. Journal of Medicinal Chemistry. 2025 Jan 15;
42. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews. 2001 Mar;46(1-3):3–26.
43. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. Journal of Medicinal Chemistry. 2002 Jun;45(12):2615–23.